Open Peer Review on Qeios

RESEARCH ARTICLE

Functional Split-tRNA: A New Perspective on the Codon Decoding Mechanism

Sergey Mureev[1,2,3], Yue Wu[2,4], Zhenling Cui[1,2,3], Kirill Alexandrov[1,2,3]

1 ARC Centre of Excellence in Synthetic Biology, Sydney, Australia
2 Centre for Agriculture and the Bioeconomy, Queensland University of Technology, Australia
3 School of Biology and Environmental Science, Queensland University of Technology, Australia
4 Cold Spring Harbor Laboratory, Cold Spring Harbor, United States

Funding: This work was supported in part by the Australian Research Council Discovery Projects DP 210104020 as well as ARC Centre of Excellence in Synthetic Biology CE200100029 to KA. KA gratefully acknowledges financial support of CSIRO-QUT Synthetic Biology Alliance.

Potential competing interests: No potential competing interests to declare.

Abstract

The translation machinery must rapidly and accurately process all codon triplets despite large differences in the stability of codon:anticodon duplexes. The constrained structure and intramolecular cooperativity of tRNA complicate understanding how its structural elements influence the thermodynamics and kinetics of the selection process. Specifically, it remains unclear whether codon:anticodon complex stability controls kinetics of tRNA selection beyond the codon recognition step. To address this we engineered fully functional split-tRNAs with a dangling anticodon triplet instead of an anticodon loop. Using this tool, we demonstrated that codon-anticodon complex stability is primarily influenced by the dipole moments of adjacent nucleobases and does not control the rate of GTP hydrolysis by EF-Tu. We conclude that the codon-anticodon minihelix functions as a passive steric gate of decoding site closure.

Corresponding authors: Kirill Alexandrov, kirill.alexandrov@qut.edu.au; Sergey Mureev, ziraffa81@gmail.com

A trade-off between speed and accuracy is inherent to most biological processes [1][2]. This is particularly true for translation, which must operate sufficiently fast to self-replicate and support cellular homeostasis [3]. As a consequence, the ribosome must have evolved to maximise discrimination between cognate and near-cognate codon-anticodon interactions within the critical time window of tRNA selection. This process consists of an initial assessment of codon-anticodon interaction, known as initial selection, followed by a proofreading step. Initial selection involves the codon-independent binding of an aminoacylated tRNA (aa-tRNA), in complex with a GTP-loaded elongation factor (EF-Tu), to the small ribosomal subunit, followed by codon-anticodon recognition. The stability of the codon-anticodon complex is cooperative function of complementarity of nucleotides' hydrogen-bonding faces and stacking interactions. Mismatches compromise this cooperativity and can result in up to 1,000-fold difference in affinity between cognate and near-cognate pairs [4][5][6][7][8][9][10][11][12][13][14].

In the next step, the closure of the ribosomal decoding site around the codon-anticodon minihelix triggers EF-Tu's

GTPase activity. In the closed complex, the decoding site scans the first and second codon positions of the minihelix by interacting with the backbone of its minor groove [8]. Notably, this interaction creates a new molecular interface that imposes energetic costs distinct from those associated with codon-anticodon interaction. The combined energetic effects of both interactions results in a 10,000-fold selectivity for cognate over near-cognate aa-tRNA [12][14]. However, it is widely recognized that the forward rate of GTPase activation must be rapid to sustain high translation speed, thereby biasing the speed-accuracy trade-off towards speed [9][10][14][15]. This requirement places the codon recognition step under kinetic control, limiting its ability to fully exploit the affinity difference between cognate and near-cognate codon-anticodon pairs. There is, however, one caveat: although mismatches at the first or second codon positions can be effectively discriminated during the GTPase activation step, sterically neutral mismatches at the third codon position may evade surveillance by the decoding site. Consequently, for most readers of split codon families, these mismatches must either be discriminated based on reduced codon-anticodon complex stability or result in high error rates during the initial selection stage [16][17][18][19][20] (table S1).

Despite decades of structural, biochemical, and biophysical analysis, there is no consensus on the extent to which differences in thermodynamic stability between cognate and near-cognate codon-anticodon complexes contribute to their discrimination, and whether initial selection effectively eliminates most near-cognate interactions [21].

One strategy to reconcile the fast forward rate with effective differentiation among suppressors of split codon families could be to couple codon-anticodon complex stability to the rate of GTPase activation or GTP hydrolysis by EF-Tu [5]. Such coupling could occur through the tRNA itself or via some component of the ribosome. The former mechanism is intuitively appealing given the obligatory distortion of the tRNA core during codon recognition [22][23][24]. Indeed, mutations within the core of $tRNA^{Trp}$, such as G24A, induce miscoding by accelerating the GTPase reaction[25][26][27]. Consequently, it has been proposed that strain within tRNA plays a crucial role in positioning its acceptor end for productive GTP hydrolysis [28][29][30]. An alternative mechanism involving a ribosomal component may rely on an ‘induced fit’, where part of the binding energy between the codon-anticodon minihelix and the decoding site is used to rearrange the latter [9][10]. In this scenario, stabilization of the codon-anticodon complex accelerates the rate of decoding site closure by reducing the entropic component of the corresponding activation barrier. Determining whether the rate of tRNA selection beyond the codon recognition step is influenced by codon-anticodon complex stability is essential for identifying the relevant mechanism.

The ribosome must process all 61 triplets, which display a 10,000-fold difference in complementary duplex stabilities [31] (table S1). Specifically, codons with G/C in the first position, which form stable to moderately stable codon-anticodon duplexes, are mostly found in unsplit codon families. In contrast, codons with A/U in the first position, which form duplexes of moderate to weak stability, are predominantly associated with split codon families [32] (table S1). The evolutionary rational for this partitioning likely arises from the need to facilitate the rejection of 3^{rd}-position mismatches in split codon families. A 10,000-fold stability range of cognate duplexes exceeds the 1,000-fold difference in affinity between cognate and near-cognate tRNA substrates. If uncompensated, this disparity would significantly undermine stability-based discrimination between cognate and near-cognate substrates. To this end, the Extended Anticodon Hypothesis postulates that the anticodon loop adjusts the strength of codon-anticodon interactions by varying nucleotides in the flanks [33].

Understanding the molecular basis of such adjustment would clarify how the anticodon loop narrows the stability range of codon-anticodon duplexes and whether the codon-anticodon complexes formed by readers of split and unsplit codon families remain segregated into lower and higher stability ranks. These results will shed light on the balance between thermodynamic and kinetic control during the codon recognition step.

However, prediction of codon-anticodon complex stabilities is challenging because multiple factors contribute to the properties of the constrained anticodon loop. Although the impact of a weak stacking context at N38 or the complementarity between N32 and N38 in reducing stability can be rationalized [13][34][35][36][37], the contributions of other determinants remain unclear. For example, it is unclear why an optimal loop context is likely established when A occupies positions 37 and 38, whereas G, despite its favourable stacking properties, is rarely found at position 38 and tends to destabilize the complex when it occurs at position 37 [38].

To address the questions raised above, we engineered a functional split-tRNA with a discontinuous anticodon arm, allowing modulation of codon-anticodon complex stability without altering the geometry of the codon-anticodon minihelix. Our findings indicate that the anticodon loop is not essential for translation and challenge the notion that the strain within tRNA is crucial for triggering GTP hydrolysis. This new decoding architecture, combined with a novel codon suppression assay, allowed us to evaluate the decoding efficiency of all N37N38 split-tRNA variants. We found that codon-anticodon complex stability is primarily influenced by the dipole moment of adjacent nucleobases. Our data allowed us to analyse the impact of compromised stability in the N36-N39 segment on the kinetic parameters of tRNA selection stages. From this analysis, we concluded that the stability of codon-anticodon complex does not influence the rate of GTP conversion and that the codon-anticodon minihelix primarily serves as a passive steric gate for decoding site closure.

Results

Rational for anticodon loop splitting

RNA maximizes base stacking by ordering intraloop nucleotides on the 3’ side of the stem and using coaxial stacking to prevent the exposure of bases and base pair planes to the solvent [39][40]. This results in the formation of a quasi-continues helix-like structure, propagating from the codon-anticodon minihelix to the anticodon stem across the 3’ anticodon flank (nucleotides 37, 38) [41][42] (Fig. 1, A and D). Nucleotides 32 and 33 in the 5’ anticodon flank mediate loop closure and do not engage in stacking interactions or make contact with the ribosome’s decoding site [43] (Fig. 1D). We hypothesized that the stacking column involving the 3’ anticodon flank makes the largest contribution to codon-anticodon complex stability. This renders the 5’ anticodon flank functionally dispensable and potentially enables its deletion, resulting in a discontinuous sequence (Fig. 1B). The loss of rigidity in discontinuous anticodon arm is expected to impose a significant entropic penalty on the stability of the codon-anticodon complex. This, however, may not be detrimental to the functionality of the split system, as the stability of cognate codon-anticodon duplexes naturally varies by nearly 10,000-fold (table S1). Hence, the entropic penalty could potentially be compensated by combining a stable GC-rich codon-anticodon

duplex with an optimal 3'-flanking sequence. Splitting the anticodon loop should not interfere with the ability of the elongation factor (EF-Tu) to bind and deliver such tRNA to the ribosome, given the remote position of the tRNA:EF-Tu binding interface.

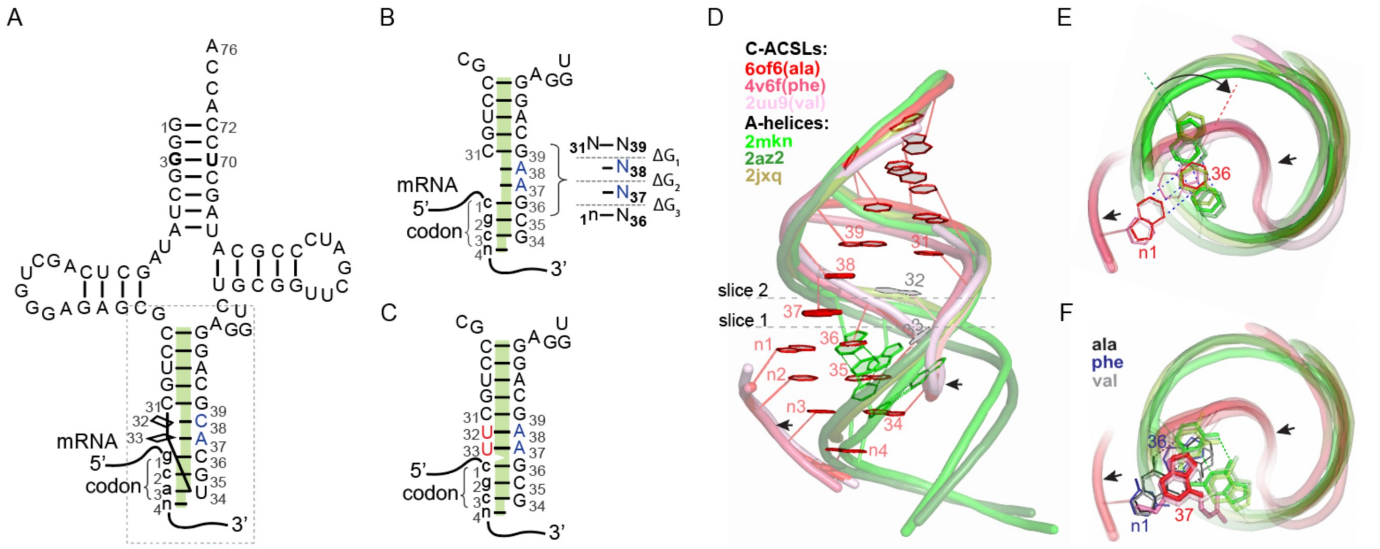

Figure 1. Structural basis of codon-anticodon interactions and its implications for constructing functional split-tRNA. (**A**) Schematic of wild-type $tRNA^{Ala}UGC$ aligned with its codon, highlighting a quasi-continuous stacking column formed by the codon-anticodon minihelix, N37N38 flank, and the anticodon stem (green). The dashed box outlines the codon-anticodon stem-loop complex shown in panels (B) and (C). (**B**) Discontinuous anticodon arm of the chimeric split-$tRNA^{Ala}GCG$, with the deleted N32N33 flank, in complex with the arginine codon CGC. Stacking interfaces within the N36-N39 segment and the corresponding free energy changes (ΔG_1-ΔG_3) are shown in the right inset. (**C**) As in (B), but with an extended anticodon stem (red) including the A37A38 flank. (**D**) A side view of high-resolution structures showing three codon-anticodon stem-loop complexes (in shades of red) superimposed with canonical RNA A-helices (in shades of green). The alignment pattern and sequence information are shown in table S2. Molecular backbones are depicted as worms, with nucleobases shown as wires, corresponding to PDB entries 6of6 (for Ala C-ACSL) and 2mkn. The inset displays the color coding for the PDB entries. An arrow indicates the protrusion of the codon-anticodon base pairs beyond A-helical boundaries. (**E**) Top-down view corresponding to optical slice 1 from (D), illustrating the positions of n1:N36 and the overlapping base pair in A-helices, as well as the angle of overtwist (arc arrow). (**F**) Similar to (E), but for optical slice 2. The n1:N36 base pairs are color-coded according to their assignment to specific C-ACSL (left inset). The figure shows that in the intact, overwound anticodon loop, N37 serves as a stacking platform for both n1 and N36 nucleotides of the first C-AC base pair, whereas the A-helix counterpart of N37 overlaps only with the proximal part of N36.

With this in mind, we designed a split-tRNA based on the $tRNA^{Ala}$, as its anticodon arm is not involved in aminoacylation process. We introduced the arginine anticodon GCG, flanked by the common A37A38 sequence [38][41][42], at the dangling end (Fig. 1B).

Establishing a fluorescence rescue assay for split-tRNA functionality analysis

To assess the functionality of split-tRNA we developed a novel fluorescence rescue assay that monitors decoding of a specific codon by a specific tRNA. The assay is based on the translation of a mutant, non-fluorescent GFP in which alanine 226 is replaced by arginine. The effect of the mutation can be reversed by supplementing an *in vitro* translation

reaction with a chimeric tRNA featuring an arginine anticodon but charged with a function-restoring amino acid such as alanine (Fig. 2A). The six-fold degeneracy of arginine codons enables construction of a codon-biased GFP ORF, wherein codons from two distinct arginine codon families are used to encode the amino acid at position 226 and the remaining arginines in the GFP ORF (Fig. 2A). The functionality of the assay was confirmed using a coupled *E. coli* S30 *in vitro* transcription/translation system, programmed with DNA templates containing the A226R mutant GFP-coding ORF [44].

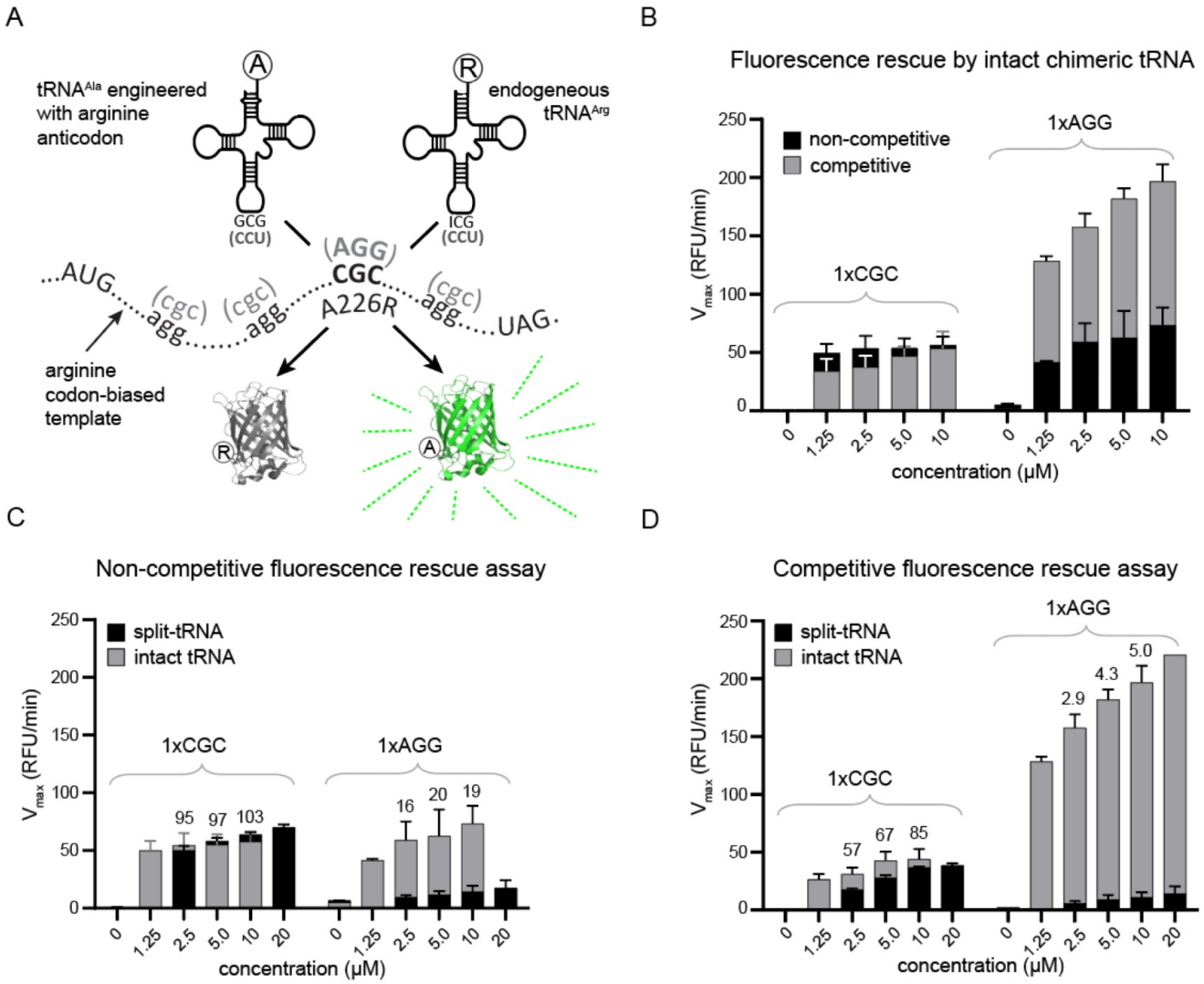

Figure 2. Assessment of split-tRNA functionality using the fluorescence rescue assay. (**A**) A schematic of the assay. The arginine codon at position 226 of a codon-biased template encoding the non-fluorescent GFPA226R mutant that can be decoded by either chimeric $tRNA^{Ala}$ that incorporates alanine and rescues fluorescence or endogenous $tRNA^{Arg}$ that incorporates arginine and results in non-fluorescent GFP. The 'cloverleaf' structures of chimeric $tRNA^{Ala}$ and native $tRNA^{Arg}ICG$ are shown above the template. Alternative codon/anticodon combination is indicated in parentheses. (**B**) A bar chart showing maximal fluorescence accumulation rates (V_{max}) in *in vitro* translation reactions programmed with GFP A226R-coding templates and supplemented with chimeric $tRNA^{Ala}$ featuring GCG or CCU anticodons. The reactions were performed under competitive conditions (grey bars) using S30 *E. coli* extract or noncompetitive conditions (black bars) using extracts depleted of endogenous $tRNA^{Arg}$ (fig. S1). Error bars represent standard deviations from 5 pseudo-replicates for CGC and 3 technical replicates for AGG. (**C**) Translation efficiency of split-tRNAs with GCGaa or CCUaa dangling ends (Fig. 1B) compared to intact chimeric tRNAs under noncompetitive translation conditions. The values above bars indicate the percentage of V_{max} for split-tRNAs relative to intact $tRNA^{Ala}$. The error bars represent standard deviations from 3 independent translation experiments. (**D**) As in (C), but under competitive conditions.

In vitro translation reactions supplemented with purified chimeric $tRNA^{Ala}$ carrying GCG or CCU anticodons displayed dose-dependent increase in fluorescence yield (Fig. 2B). Under these conditions, hereafter referred to as 'competitive', the chimeric tRNA competes with the endogenous arginine isoacceptor for the codon of interest. To estimate the decoding efficiency of chimeric tRNA, we compared fluorescence yields under competitive and noncompetitive reaction conditions. The noncompetitive condition was established by supplementing a translation reaction, based on S30 lysate depleted of total tRNA, with a tRNA complement specifically depleted of competing endogenous tRNA species [45] (fig. S1). As shown in Fig. 2B, fluorescence yields of the CGC- and AGG-templates exhibited opposing trends under under competitive and noncompetitive reaction conditions. This can be attributed to the ~10-fold excess of competing endogenous $tRNA^{Arg}ICG$ over $tRNA^{Arg}CCU$ [46].

To assess the reading efficiency of codons other than arginine, we developed a generalizable assay that monitors the accumulation of a fluorescent product in response to the complementation of two consecutive codons of interest by chimeric or wt tRNAs (fig. S2, A and B). The assay can only be performed in the lysate depleted for endogenous suppressor(s) of the codons analysed. Hereafter it is referred to as 'two-codon' assay.

Construction and functional characterization of tRNA with a split anticodon loop

The split-tRNA based on the $tRNA^{Ala}UCG$ isoacceptor was constructed by annealing synthetic RNA fragments spanning nucleotides 1-31 and 34-76 (Fig. 1B and fig. S2C). We evaluated the codon suppression efficiency of split- and intact tRNAs featuring GCG and CCU anticodons in a fluorescence rescue assay under noncompetitive reaction conditions. Surprisingly, fluorescence yields of reactions containing split-$tRNA^{Ala}GCG$ were comparable to those of reactions primed by intact chimeric $tRNA^{Ala}GCG$ (Fig. 2C).

We confirmed that the observed translation was indeed mediated by the split-tRNA, rather than its derivative with an enzymatically repaired anticodon loop, by assessing the structural integrity of the split-tRNA during the course of the translation reaction (fig. S3, A to D). To ascertain the generalizability of the developed approach, we evaluated the functionality of the split-tRNA in a eukaryotic translation system using *Leishmania tarentolae* cell extract (LTE) [47][48]. We found that the split-tRNA exhibited 30% of the decoding efficiency of the intact chimeric tRNA under competitive conditions in the LTE system (fig. S4).

We next tested the tolerance of split-tRNAs with GCG or UGC anticodons to wobble pairing and mismatches at the 3rd codon position. Decoding efficiency decreased by one to two orders of magnitude compared to the corresponding intact tRNAs (fig. S5, A and B). Additionally, the presence of a deoxynucleotide at position 37 reduced decoding efficiency, most likely due to increased ribose flexibility (fig. S5D). These findings suggest that the stacking contribution of the correct N3:N34 base pair has a greater impact on codon-anticodon complex stability in split-tRNA than in intact tRNA. This can be attributed to the higher entropic cost associated with accommodating the free dangling end within the A-site.

We also tested whether extending the double-stranded region of the anticodon stem in the split-tRNA across the A37A38

flank would enhance codon-anticodon complex stability by introducing a U33:A37 base pair, rather than relying on A37 alone for stacking with the N1-N36 (Fig. 1C). Compared to the nonextended variant, the resulting 1-33/34-76 split-tRNA exhibited either no change or a decrease in decoding efficiency, depending on the codon-anticodon interaction strength and the presence of 5'-ribose modifications at N34 (fig. S3B and fig. S5). This can be explained by the overtwisted loop configuration (Fig. 1E), which reduces the overlap between the N1:N36 codon-anticodon base pair and N37 when the latter conforms to A-helical geometry (Fig. 1F). This finding aligns with the observation that intraloop complementarity in intact tRNA destabilizes the codon-anticodon complex. Accordingly, the insertion of an additional nucleotide – specifically adenine - into the 3' anticodon flank is well tolerated, as it relaxes the overtwist (fig. S6).

In subsequent experiments, we found that the decoding efficiency of the split-tRNA with a CCU anticodon for the alternative arginine codon AGG was significantly lower, relative to intact $tRNA^{Ala}$CCU, then the relative efficiency of split-$tRNA^{Ala}$GCG (Fig. 2C and fig. S2F). This discrepancy can be attributed to the diminished stacking strength of uridine (U36) in the 'cardinal' position of split-$tRNA^{Ala}$CCU compared to guanine (G36) in split-$tRNA^{Ala}$GCG [33]. Notably, split-$tRNA^{Ala}$UCU, with A3:U34 instead of C3:G34 at the third codon position, exhibited negligible decoding activity for its cognate codon (fig. S5C). A similar trend of reduced decoding activity, relative to the intact tRNA, was observed for both chimeric split-$tRNA^{Ala}$GCU and split-$tRNA^{Ser}$GCU, which also feature U36 in the cardinal position (Fig. S2, D to G).

Based on these observations, we concluded that in split-tRNA, the high entropy loss associated with the overwinding of the free dangling end in the A-site makes the enthalpic contribution of stacking interactions a critical component of codon-anticodon complex stability. In contrast, intact tRNAs exhibit a significantly lower entropic penalty during codon-anticodon complex formation, resulting in a less pronounced dependence on stacking interactions.

Comprehensive analysis of N37N38 combinations in split-tRNA

Removal of the intraloop constraint creates a novel system where the stability of codon-anticodon complex is governed by stacking interactions within the N36-N39 segment (Fig. 1B). This enables analysis of the relationship between the sequence of this segment and the decoding efficiency of split-tRNA. To this end, we constructed 16 variants of split-$tRNA^{Ala}$GCG corresponding to all N37N38 combinations (Fig. 1B), and tested them in a fluorescence rescue assay under competitive and noncompetitive conditions. Control reactions supplemented with intact chimeric $tRNA^{Ala}$GCG were used to normalize the fluorescence yields across these conditions.

The observed reduction in fluorescence yield among the tested split-tRNAs ranged from negligible to pronounced upon transitioning from noncompetitive to competitive reaction conditions (Fig. 3A). This variation likely reflects differences in dissociation rates among the N37N38 split-tRNA variants, which become more apparent in the presence of competing native $tRNA^{Arg}$ICG. To further analyze these results, we plotted the ratios of relative fluorescence accumulation velocities in competitive versus noncompetitive environments against the estimated stabilities of the N36-N39 segment (Fig. 3A inset). The stabilities were calculated as the exponential of the sum of free energy changes for the formation of N1-N36/N37, N38/N31-N39, and N37/N38 stacks (Fig. 1B inset, table S3).

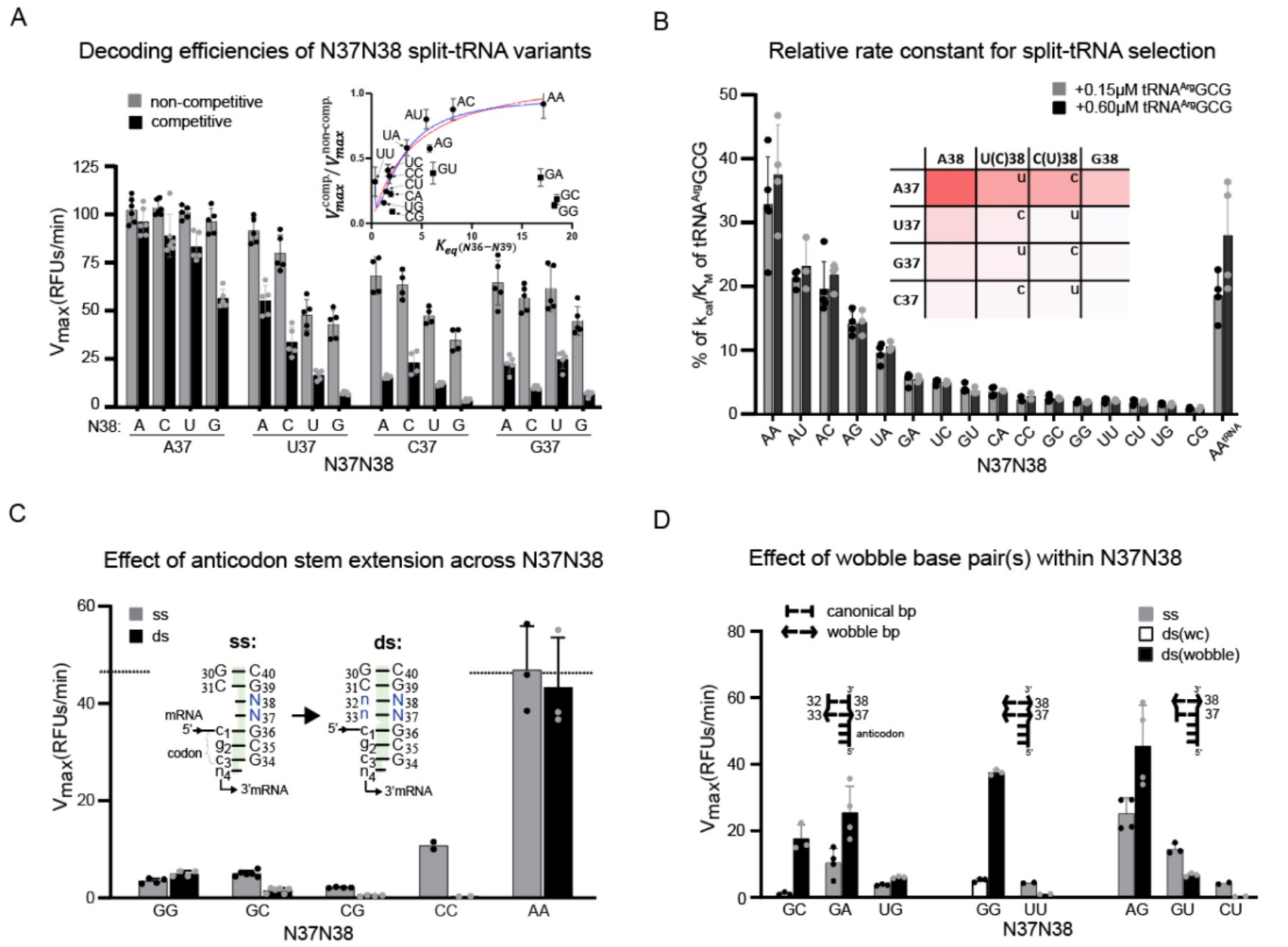

Figure 3. Impact of N37N38 identity on split-tRNA ^{Ala}GCG decoding efficiency. (**A**) A bar chart plot showing the percentage of maximal fluorescence accumulation rate (V_{max}) for various N37N38 split-tRNA variants relative to intact chimeric tRNA ^{Ala}GCG, under competitive and noncompetitive translation conditions. Data are shown as mean ± standard error from five independent experiments. The inset plots decoding efficiency against estimated stability within N36-N39. Data points (black circles) are fitted to quadratic (blue) or hyperbolic (red) equations (Eq. S11), with outliers (black squares) excluded. The fit is shown as a solid line. (**B**) A plot of normalized k_{cat}/K_m for various N37N38 split-tRNA variants relative to k_{cat}/K_m for intact tRNA^{Arg}GCG. Bars represent average values from experiments using S30 lysate depleted of endogenous tRNA ^{Arg}ICG and reconstituted with synthetic tRNA^{Arg}GCG competitor at 0.15 μM and 0.6 μM. Error bars show standard deviations from three and four independent experiments, respectively. The inset presents a heatmap summarizing nucleobase preferences at positions 37 and 38 for optimal decoding efficiency. (**C**) Comparison of decoding efficiencies of split-tRNA variants with single or double-stranded N37N38 segments, as shown in the inset. The dashed line indicates the V_{max} value for intact tRNA. Data are based on two to six independent measurements. (**D**) Effect of wobble base pairs within the N37N38 segment on split-tRNA decoding efficiency. The efficiencies of split-tRNAs with wobble base pairs are compared to those with single-stranded or canonical double-stranded N37N38 configurations. Schematic representations of the terminal N34-N38 section of the split anticodon arm are shown above the respective bars for clarity. Data are based on two to four independent measurements.

Stacking energies for the first two segments were estimated from Dangling End stabilities, while those for the N37/N38 interface were derived from previously reported experimental measurements or molecular dynamics simulations of dinucleotide monophosphates (tables S4 to S6). The plot revealed only a partial correlation between the estimated stabilities and decoding efficiency (Fig. 3A inset). Notably, despite the high stacking stabilization energy associated with G-containing interfaces, split-tRNA variants with G37 exhibited reduced decoding efficiency.

The ratio of fluorescence accumulation velocities (V_{max}) under competitive versus noncompetitive reaction conditions reflects the likelihood of productive selection of split-tRNA^{Ala}GCG on the CGC codon in the presence of competing tRNA^{Arg} (supplementary text 1). This ratio can be expressed using the selection rates for split-tRNA ($Rate^{Ala}$) and competing tRNA^{Arg} ($Rate^{Arg}$), as shown in Equation 1:

$$\frac{V_{\max}^{\text{comp.}}}{V_{\max}^{\text{non-comp.}}} = \frac{\text{Rate}^{\text{Ala}}}{\text{Rate}^{\text{Ala}} + \text{Rate}^{\text{Arg}}}$$

Substituting the rates in Equation 1 with the apparent second-order rate constants $\left(\frac{k_{cat}}{K_m}\right)$ and concentrations of ternary complexes (TC), as detailed in supplementary text 1 (eq. S1 to eq. S3), yields Equation 2:

$$\frac{V_{\max}^{\text{comp.}}}{V_{\max}^{\text{non-comp.}}} = \frac{\left(\frac{k_{cat}}{K_m}\right)^{Ala} \cdot \left[\text{TC}^{Ala}\right]}{\left(\frac{k_{cat}}{K_m}\right)^{Ala} \cdot \left[\text{TC}^{Ala}\right] + \left(\frac{k_{cat}}{K_m}\right)^{Arg} \cdot \left[\text{TC}^{Arg}\right]}$$

Using Equation 2 and the concentrations of ternary complexes, we can derive the second-order rate constants for the selection of N37N38 split-tRNA variants, normalized to the intact tRNA competitor $\frac{\left(k_{cat}/K_m\right)^{Ala}}{\left(k_{cat}/K_m\right)^{Arg}}$ (supplementary text 1, eq. S4). We used the concentration of aa-tRNAs as a proxy to estimate the concentration of ternary complexes, which was found to be approximately 4 µM during the log phase of product accumulation (fig. S7 and fig. S8). Instead of measuring the concentration of competing native tRNA^{Arg}ICG (TC^{Arg}) in the S30 lysate, we introduced T7-transcribed tRNA^{Arg}GCG into a tRNA-depleted lysate to reproduce competitive reaction conditions. Since the introduced tRNA^{Arg}GCG effectively competed with split-tRNA at submicromolar concentrations, it was presumed to be fully aminoacylated. Performing both competitive and noncompetitive translation reactions in the same depleted lysate eliminated the need for fluorescence normalization, allowing for direct calculation of the normalized k_{cat}/K_m values for all N37N38 variants (supplementary text 3, eq. S4). The similar k_{cat}/K_m values obtained with two different concentrations of competing tRNA^{Arg}GCG validate the robustness of the assay and its suitability for quantitative analysis of decoding efficiency (Fig. 3B). The corresponding plot shows that the presence of an adenine moiety at either position 37 or 38 resulted in the highest decoding efficiency within each set, whereas the A37A38 combination exhibited a synergistic increase in decoding efficiency (Fig. 3B inset).

Compared to variants featuring GCG anticodon, the variants with CCU anticodon showed similar preferences for N37N38 combinations to achieve optimal decoding efficiency for the AGG codon, although their normalized k_{cat}/K_m values were over an order of magnitude lower (fig. S9A). In addition to the weak stacking context of U36 in the cardinal anticodon position, this difference may also be attributed to the unusually high decoding efficiency of wild-type tRNA^{Arg}CCU, which could be an adaptation to mitigate high rates of miscoding at AGG codons (fig. S5C inset). Similar to the AGG codon,

much lower decoding efficiency for AGC was observed, particularly for G37-containing split-tRNAs based on $tRNA^{Ser}GCU$ (fig. S9B).

To determine the cause of the low decoding efficiencies observed for G37/G38-containing split-tRNA variants, we first ruled out extrinsic factors, such as N37N38-specific degradation of the anticodon termini or variations in aminoacylation levels (supplementary text 2, fig. S7 and fig. S8).

Intrinsic factors that may contribute to the low decoding efficiency of G37/G38-containing split-tRNA variants include the distinct energetic effects of specific N37 and N38 combinations, formation of inter- or intramolecular secondary structures by protruding N34-N38 strands, and N37N38-dependent coordination of Mg^{2+} [49]. Experimental testing ruled out secondary structure formation as a major factor contributing to the reduced decoding efficiency of G/C-rich N37N38 split-tRNA variants (supplementary text 2, fig. S10). Additionally, N37N38 split-tRNA variants generally exhibited similar relative decoding efficiencies for the CGC codon across different Mg^{2+} concentrations. This observation supports the notion that Mg^{2+} primarily facilitates the initial, codon-independent stage of ternary complex association with the ribosome [12][14] (supplementary text 2, fig. S11).

Next, we examined how thermodynamic factors may influence the observed decoding efficiency patterns. We initially investigated whether the activation barrier for transitioning of the N34-N38 region from its free state to the codon-bound state affects the forward rate of codon-anticodon complex formation. To address this, we standardized the unbound states by extending the anticodon stem across the N37N38 segment (Fig. 1C and Fig. 3C, inset). The resulting variants exhibited either unchanged or significantly reduced decoding efficiency (Fig. 3C), suggesting that variation in the forward rate are unlikely to account for the observed differences in decoding efficiency.

Remarkably, G37 involvement in wobble pairing with U33 improved split-tRNA performance by over 10-fold compared to nonextended or canonically extended variants (Fig. 3D). The wobble geometry likely stabilizes G37/38 in a configuration that introduces new stacking edges, enhancing continuous stacking within the codon-anticodon stem with lower entropy loss (supplementary text 3, fig. S12). The inverse orientation of wobble pairs in U37C38 or U37U38 variants did not enhance decoding efficiency, likely due to the nonisosteric nature of wobble pairing (Fig. 3D).

The orientations of adjacent nucleotides are governed by the balance between the repulsion of their individual dipoles and the solvation energy of their net dipole [50]. Polarizability, a critical factor for effective stacking, follows the order G > A > C > U [51][52], whereas the dipole moment of individual nucleobases decreases in the order G ~ C > U > A[53][54][55]. Consequently, positioning G37 between the C1:G36 base pair and G38 or C38 is expected to result in increased repulsion between the large dipole moments of the three residues in the N36-N38 stack. This effect may be more pronounced in the overtwisted context of a dangling end. Adding an extra A to the N37N38 flanking sequence, which relaxing the overtwist, improved the decoding efficiency of split-tRNA containing G37 compared to those containing U37 or C37 (fig. S6).

Unlike other nucleobases, A lacks strong electronegative ring substituents, resulting in the smallest dipole moment. Simultaneously, its polarizability is comparable to that of G [51][52][56]. This unique combination allows A-containing

N37N38 variants to adopt a broader range of orientations, providing optimal stacking with neighboring nucleobases regardless of their identity. The ordering of nucleobases at positions 37 and 38 favors smaller dipole moments, with decoding efficiency following the order A >> U(C) > G for the given N38 or N37, respectively (Fig. 3B inset).

Stability of codon-anticodon complexes and its effect on tRNA selection kinetics

As outlined in the introduction, the thermodynamic differences between cognate and near-cognate codon-anticodon complexes can ensure high translational accuracy at the codon recognition stage, provided the system reaches equilibrium. This is particularly relevant for the accurate decoding of split-codon families, where mismatches at the 3rd codon position may evade geometric proofreading, leading to miscoding.

However, according to current views, prioritizing translation speed over accuracy shifts initial tRNA selection to kinetic control, preventing the system from reaching equilibrium during codon recognition. This implies a mechanism that couples accurate codon-anticodon recognition with accelerated GTP hydrolysis by EF-Tu. Such coupling may involve either the tRNA or a ribosomal component of the decoding site. The latter scenario gave rise to the induced fit hypothesis, which implies that the codon-anticodon complex controls the energy barrier for decoding site closure (supplementary text 4).

The split-tRNA framework offers an experimental platform for manipulating stability of the codon-anticodon complex through various N37N38 combinations, without altering the geometric component of selection. This allows for the analysis of how changes in codon-anticodon complex stability affect different stages of tRNA selection (supplementary text 4).

The tRNA selection process is described by a four-stage kinetic model[12]. It begins with the codon-independent association of the ternary complex with the ribosome to form complex C_0. It is followed by codon recognition, resulting in the formation of codon-anticodon complex C_1. The subsequent stage involves the closure of the decoding site around the correct codon-anticodon minihelix, forming the pre-catalytic complex C_2 (Scheme 1, Fig. 4A).

$$R + T_3 \overset{k_1}{\underset{k_{-1}}{\rightleftarrows}} C_0 \overset{k_2}{\underset{k_{-2}}{\rightleftarrows}} C_1 \overset{k_3}{\underset{k_{-3}}{\rightleftarrows}} C_2 \xrightarrow{k_4} GDP + P_i$$

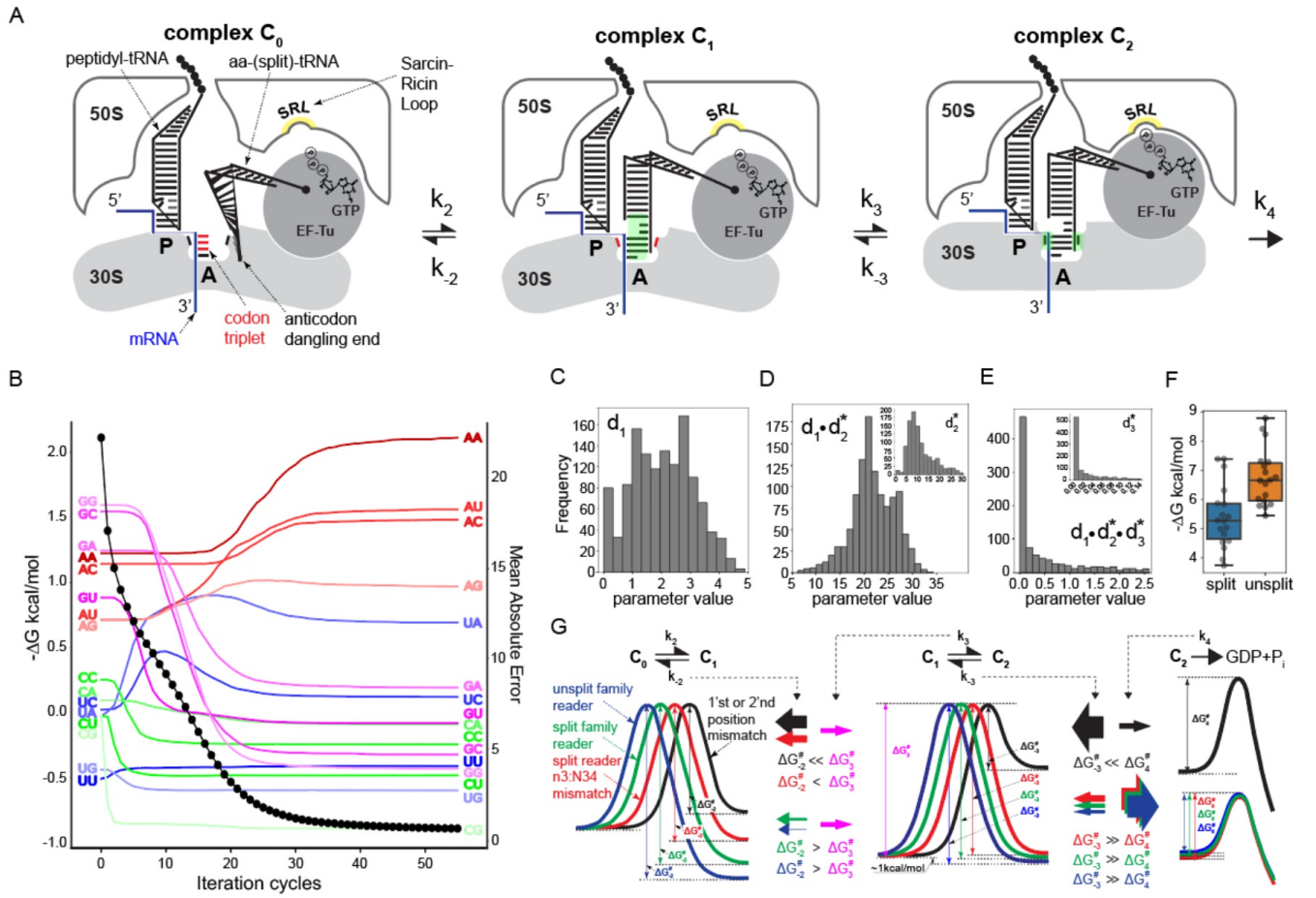

Figure 4. Influence of codon-anticodon complex stability on split-tRNA selection kinetics. (**A**) Schematic representation of tRNA selection on the ribosome, illustrating changes in interaction interfaces between the ribosome and the ternary complex at different selection steps (see Scheme 1). In complex C_0, the mRNA ('cofactor') presents the codon triplet (red) as a recognition surface for the anticodon ('ligand's binding site'). The stability of the interaction interface (green) is assessed in complex C_1. In this complex, the codon-anticodon minihelix acts as an allosteric effector, providing a new recognition surface to the monitoring bases (red) of the decoding site. The free energy of this new interaction interface in complex C_2 (green) influences the assembly of the transition state for GTP hydrolysis. (**B**) Evolution of stability terms from fitting Eq. 5 to the regression model used for analysis (fig. S13). The traces represent averages from 1000 simulations. The X axis shows the number of fitting iterations, the left Y axis denotes stabilities of the N36-N39 segment in kcal/mol, and the right Y axis shows the mean absolute error of the fit at each iteration (filled black circles). (**C**) Frequency distribution of predicted d_1 values from regression analysis. Histogram based on 1000 simulations, with frequency shown on the Y axis. The optimal bin number was determined using the Freedman-Diaconis rule (see Git repository). (**D**) Frequency distribution of the linear regression coefficient d_1d_2, with an inset histogram for d_2. (**E**) Frequency distribution of the quadratic regression coefficient $d_1d_2d_3$, with an inset histogram for d_3. (**F**) Distribution of stabilities for codon-anticodon complexes formed by decoders of split and unsplit codon families (data from table S1). Each box represents the interquartile range (IQR) with the median stability indicated by the horizontal line inside the box. Outliers for $tRNA^{Trp}$CCA and $tRNA^{Leu}$UAG are excluded for clarity. (**G**) Free energy profiles for the stages of tRNA selection, as indicated above each profile.

The above scheme can be described by Equation 3, which relates the apparent second-order rate constant $\left(\frac{k_{cat}}{K_m}\right)$ for the GTPase reaction to the rate parameters of different selection stages [12]:

$$\frac{k_{cat}}{K_m} = \frac{k_1}{1 + d_1\left(1 + d_2\left(1 + d_3\right)\right)}$$

In this equation, k_1 is the codon-independent rate constant for the association of the ternary complex with the ribosome, which corresponds to the maximum achievable rate of tRNA selection. The discard parameters d_1, d_2, and d_3 are defined by the ratios $\frac{k_{-1}}{k_2}$, $\frac{k_{-2}}{k_3}$, and $\frac{k_{-3}}{k_4}$, respectively (Scheme 1). These parameters indicate the probabilities of dissociation for the complexes C_0, C_1, and the reversal of the C_2 complex.

According to Eq. 3, while d_1 is independent of codon-anticodon interaction, the stability of the codon-anticodon complex may affect k_{cat}/K_m through d_2, d_3, or both discard parameters (supplementary text 5). Thus, modeling these discard parameters by fitting experimental data to Eq. 3 may reveal which tRNA selection stages are sensitive to variations in codon-anticodon complex stability.

To incorporate the stabilities of different N36-N39 segments as independent variables in Eq. 3, we introduced universal surrogate discard parameters d_2^* and d_3^*, which correspond to a hypothetical split-tRNA* that is devoid of any stacking interactions within the N36-N39 segment. By adjusting these surrogate parameters with the stabilities of different N36-N39 segments, we can express the discard parameters d_2 and d_3 for each N37N38 split-tRNA variant (supplementary text 5). Substituting these into Eq. 3 yields a system of 16 equations in the form of Equation 4:

$$\frac{k_{cat}}{K_m} = \frac{k_1}{1 + d_1\left(1 + d_2^* \cdot x\left(1 + d_3^* \cdot x\right)\right)}$$

In this equation, x represents the stability factor from the stacking interactions within N36-N39 for the given N37N38 split-tRNA variant, expressed as $e^{\frac{\Delta G_{N36-N39}}{RT}}$. This approach follows the nearest-neighbour model for nucleic acid stability, where overall stability is determined by the sum of individual stacking contributions.

As described, the $\Delta G_{N36-N39}$ was initially calculated as the sum of free energy changes for each stacking unit (Fig. 1B), assuming unit independence and optimal nucleotide overlaps (table S3 to S6). However, in the constrained context of an accommodated dangling end, $\Delta G_{N36-N39}$ is affected by the specific repulsion and attraction forces of each N37N38 combination. Therefore, we approximated the actual $\Delta G_{N36-N39}$ values in our linear regression model by iteratively perturbing the initially estimated $\Delta G_{N36-N39}$ with randomly selected weights (fig. S13). Using experimentally determined parameters, such as decoding efficiencies and estimated concentrations of competing ternary complexes, we transformed Eq. 4 into a quadratic polynomial form (Equation 5) suitable for regression analysis (supplementary text 5):

$$\frac{\left[c_s\right] \cdot (1 - f)}{\left[c_{wt}\right] \cdot f} = d_1 + d_1 d_2^* \cdot x + d_1 d_2^* d_3^* \cdot x^2$$

In this equation, $[c_s]$ and $[c_{wt}]$ represent the concentrations of aminoacylated N37N38 split-tRNA variants and competing synthetic $tRNA^{Arg}GCG$, respectively. The parameter f denotes the ratio of maximal fluorescence accumulation rates $\left(\frac{V_{max}^{comp.}}{V_{max}^{non\text{-}comp.}}\right)$ in reactions with or without competing $tRNA^{Arg}GCG$.

The system of 16 equations for each N37N38 variant was subjected to 56 fitting iterations, each involving 100 weight perturbations (fig. S13). Figure 4B illustrates the evolution of free energy terms for each variant across 56 iterations, averaged over 1000 repeats. The improvement in the metric primarily stems from reducing the stabilities of G37- or C37-containing split-tRNA variants relative to initial estimates. For most variants, excluding A37N38, the stability gain from stacking interactions within N36-N39 was moderate and often insufficient to offset the entropy loss associated with dangling end accommodation in the A-site (Fig. 4B). Simulations showed that the A37A38 variant contributes more than 2 kcal/mol to codon-anticodon stability, whereas G37-containing variants either contribute none or destabilize it by 0.5 kcal/mol.

Since the intact anticodon loop is likely to adopt a configuration similar to that of the dangling end when aligned with the conformationally restrained codon triplet, the energy profile of the native codon-anticodon complex must be influenced by a similar interplay of dipole moments and polarizabilities within the N36-N39 segment. As a result, the 10,000-fold variation in the stabilities of codon-anticodon duplexes can be reduced by more than 100-fold within the context of the complex by varying the identities of N37 or N38, the presence of N37 modification, and intraloop complementarity (table S1). This would allow effective discrimination against near-cognate ternary complexes based on codon-anticodon stabilities.

As the regression metric improved, the model converged to coefficients corresponding to d_1, $d_1 d_2^*$ and $d_1 d_2^* d_3^*$ from which the individual discard parameters can be calculated (Eq. 5). Figures 4C and 4D display histograms for d_1 and $d_1 d_2^*$ respectively. The distribution of d_1 appears bimodal, centered around 1.5 and 2.5, while the d_2^* distribution shows a major peak around 8 with a right-skewed tail extending to larger values (Fig. 4D inset). While d_1 remains consistent with its expected value [57], unaffected by the codon-anticodon complex stability, d_2^*, estimated to be a fraction of 1 for intact tRNA [14], exceeds the anticipated value by approximately 100-fold. This discrepancy aligns with the predicted loss of several kcal/mol in stabilization energy due to the absence of stacking interactions within the N36-N39 segment in the hypothetical split-tRNA*.

The distribution of the d_3^* parameter shows pronounced skewness towards zero (Fig. 4E). This may suggest that the quadratic stability term has a minimal impact on the model's fit. Thermodynamically, it indicates that the codon-anticodon binding energy is fully realized in the dissociation constant (k_2/k_{-2}) during the codon recognition step, with no additional contribution to selectivity before GTP conversion. This challenges the induced fit hypothesis, which postulates that some binding energy is used to facilitate the closure of the decoding site (supplementary text 4). Alternatively, the small d_3^* value might reflect the dominance of k_4 over k_{-3}, regardless of the selection mechanism (Eq. 3, supplementary text 4).

Within the constraints of the mathematical model presented, we can conclude that the stability of the codon-anticodon

complex does not influence the rate of GTP conversion, provided the Watson-Crick geometry of the codon-anticodon minihelix is maintained. This suggests that the canonical minihelix likely acts as a passive steric gate, not obstructing the closure of the small ribosomal subunit, rather than stimulating it or stabilizing the closed state. Accordingly, the small ribosomal subunit undergoes spontaneous, thermally-driven fluctuations between open and closed states. These states allow for independent evaluation of the energetic penalties incurred at the codon-anticodon interaction interface and at the interface between the decoding site and the minor groove backbone spanning the 1st and 2nd codon positions (Fig. 4A). Therefore, while mismatches in the first two codon positions can be effectively rejected due to steric incompatibility, discrimination against 3rd position mismatches depends on the timing of 30S subunit fluctuations. This is consistent with split codon readers occupying a lower stability rank (Fig. 4F), allowing dissociation of their 3rd position mismatches before 30S subunit closure (Fig. 4G).

Initiation of translation with split tRNA-mRNA fusion

Our observation that split-tRNA supports translation as efficiently as its wild-type counterpart highlights its potential for engineering the translational machinery. A split-tRNA architecture offers the advantage of functioning in both *trans* and *cis* configurations with mRNA, which could be leveraged to manipulate translation initiation, elongation, and termination (Fig. 5A). To explore this, we tested whether split-tRNA could endow mRNA with a tertiary structure capable of mediating translation initiation. We hypothesized that a continuous A-helix formed by split-tRNA could mimic the codon-anticodon complex and serve as its surrogate, thereby driving translation initiation.

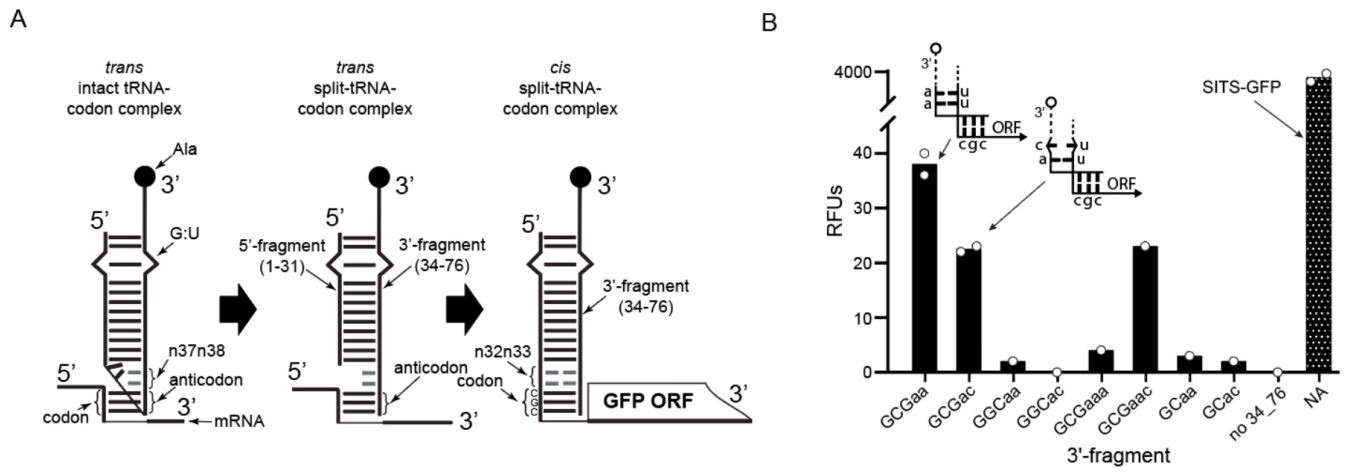

Figure 5. Split-tRNA-mRNA fusion as a mimic of the translation initiation complex. (**A**) Schematic showing the transition from *trans*-architecture codon-anticodon complexes (formed by intact and split-tRNAs) to *cis*-architecture in 5'-split-tRNA-mRNA fusion. The 1-33 5'-fragment is fused to a GFP ORF beginning with CGC codon and annealed with the 3'-fragment [34][35][36][37][38][39][40][41][42][43][44][45][46][47][48][49][50][51][52][53][54][55][56][57][58][59][60][61][62][63][64][65][66][67][68][69][70][71][72][73][74][75][76] of split-tRNA, mimicking the initiator tRNA in the ribosomal P-site. (**B**) Endpoint fluorescence counts for split-tRNA/mRNA fusion constructs translated in an *E. coli in vitro* translation system. The 1-33-CGC-eGFP-mRNA was annealed with various 3'-fragments, differing by anticodon triplets and nucleotide residues at position 38 (X-axis). Insets show stylized initiation complexes for the canonical A-U pair and C-U mismatch at positions N32-N38 (the 3'-fragments carrying Ala are marked as empty circles, with dashed lines indicating the remainder of the split-tRNA). SITS-GFP is a positive control containing SITS translation enhancer [48]. Bars represent average values from two replicates.

To test this hypothesis, we fused a 1-33 split-tRNA fragment with U32U33 to the 5' end of an eGFP-coding ORF prefaced with a CGC codon and evaluated the resulting mRNA in an *in vitro* translation reaction (Fig. 5, A and B). Although this mRNA by itself was translationally inactive, annealing with the complementary 34-76 fragment ('GCGaa') or the fragment yielding a U32•C38 mismatch ('GCGac') resulted in low but detectable GFP expression (Fig. 5B). Interestingly, the fragment with the C-extended flank ('GCGaac'), but not the A-extended flank, also produced functional hybrid mRNA (Fig. 5B). Additional controls, including 34-76 RNA fragments mismatched at two positions of the codon-anticodon duplex and fragments with the anticodon truncated at position 34 (Fig. 5B), confirmed that observed expression was attributed to split-tRNA-mediated translation initiation. These observations strongly suggest that using an exogenous RNA fragment allowed us to reconstruct a structure sufficiently similar to the initiation complex. The low efficiency of this process is expected, as the native codon-anticodon complex is overwound relative to a classical A-helix. This may be addressed by adjusting the identity and number of codon- and/or anticodon-flanking bases, potentially incorporating wobble base pairs. In addition to the demonstrated translation initiation using 5'-mRNA-split-tRNA fusion, 3'-fusions may enable the co-translational formation of mRNA-protein fusions without chemical linkers. Furthermore, removing steric constraints within the anticodon loop should enhance the efficiency of codon reassignment for quadruplet codons and unnatural bases. Recent findings that split-tRNA structures can be assembled and maintained *in vivo* are promising, as they provide avenues for the deployment of split-tRNAs in living organisms [58].

Discussion

Translation machinery must operate sufficiently fast to support self-replication and maintain intracellular protein homeostasis. It must process the entire repertoire of codon triplets, which vary by orders of magnitude in affinity for complementary anticodons depending on their G/C or A/U content. Therefore, achieving a high selection rate requires a universal mechanism that minimizes misincorporation without relying solely on the strength of codon-anticodon interactions. Sequence-to-shape complementarity provides an additional selection mechanism for the rapid geometric matching of codon/anticodon pairs. However, this process disregards the geometry at the third codon position, likely to avoid delays from rejecting isoaccepting species that rely on wobble base pairing. Consequently, for the majority of split codon family readers, mismatches at the third codon position would lead to miscoding and must be discriminated against based on the stability of codon-anticodon interactions. Therefore, the stability range among correct codon-anticodon pairs must be narrower than the stability difference between cognate and near-cognate complexes to effectively distinguish between them. The anticodon loop evolved to modulate this range by increasing stability at the lower end while decreasing it at the higher end, although the specific mechanisms it employs are not fully understood. The intramolecular cooperativity tRNA complicates defining the specific functions of its structural elements and how they influence the thermodynamics and kinetics of the selection process.

To create an experimental system where codon-anticodon complex stability is uncoupled from intra-tRNA strain, we designed a split-tRNA with a dangling end functionally replacing the anticodon loop. Additionally, we developed a robust high-resolution fluorescence rescue *in vitro* translation assay for codon-specific evaluation of tRNA functionality.

This assay demonstrated that split-tRNAs are able to support translation with efficiency comparable to intact synthetic tRNA, challenging the notion that tRNA strain is crucial for activation of GTPase hydrolysis by EF-Tu [25][28][30]. This finding aligns with previous results from cryo-EM-informed molecular dynamics simulations and smFRET studies [59][60][61], which demonstrated that both cognate and near-cognate tRNA molecules spontaneously sample distorted conformations.

Using the fluorescence rescue assay, we observed a pronounced dependence of split-tRNA decoding efficiency on the identity of the 'cardinal' nucleotide at position 36, the G/C content of the codon-anticodon duplex, and the presence of a wobble at the 3'rd codon position. These observations reflect an increased relative contribution of the enthalpy of stacking interactions to the stabilization energy of the codon-anticodon complex, due to the significant entropic losses associated with accommodating the dangling end in the A-site. The increased stacking contribution suggests a strong dependence of codon-anticodon complex stability on the identities of nucleobases at positions 37 and 38. Thus, the split-tRNA system provides a unique opportunity to investigate how variations in codon-anticodon stability affect decoding efficiency without altering codon-anticodon duplex geometry. We analyzed split-tRNAs with each of the 16 possible N37N38 combinations and found that their decoding activity declined in the following order: A37N38 > U37A38 > G37A38 > C37A38. Despite the greater stacking propensity of G and C, split-tRNA variants with these nucleotides at position 37 were least competitive (Fig. 3B). Further experiments revealed a preference for smaller dipole moments over greater polarizability in nucleobases at positions 37 and 38 with respect to decoding efficiency (Fig. 3C). We concluded that the large dipole moment of G causes repulsion with adjacent residues, limiting its conformational space and increasing the free energy of the codon-anticodon complex. In contrast, the positive synergy between dipole moment and polarizability enables A to achieve optimal stacking overlap with adjacent nucleotides.

Next, we used decoding efficiency data for all N37N38 split-tRNA variants to evaluate how codon-anticodon complex stability affects tRNA selection stages. Specifically, we modeled the impact of compromised stability in the N36-N39 segment on the kinetic parameters for codon recognition and GTPase activation [12] (fig. S13). Our results indicate that compromised stability of codon-anticodon complex does not affect the rate of GTPase activation (Fig. 4, C to E). Thus, we conclude that the canonical codon-anticodon minihelix likely acts as a steric gate by allowing the closure of the decoding site without actively stimulating it. This would explain why sterically neutral mismatches formed by $tRNA^{Lys}UUU$ lead to high background frequency of miscoding [17][19]. The ribosome's ability to hydrolyze GTP in the EF-Tu-GTP complex without tRNA [62][63] supports the notion of spontaneous decoding site fluctuations. Cryo-EM studies revealed a physical coupling between decoding site closure and EF-Tu-GTP docking at the GTPase activation site [61][64][65]. Further support for our model comes from experiments in which the replacement of atomic groups interacting with the codon-anticodon minihelix did not significantly affect the efficiency of tRNA selection process [66][67].

We used model-derived free energy changes within the N36-N39 segments to estimate the stabilities of codon-anticodon complexes formed by wt tRNAs. We assumed that the intact anticodon loop likely adopts a configuration similar to that of the dangling end when aligned with the conformationally restrained codon triplet, suggesting a similar interplay of repulsion and stacking forces. The analysis indicates that the anticodon loop narrows the stability range of cognate codon-

anticodon duplexes from ~10,000-fold to ~100-fold through variations in the identities of N37 and N38, which is further modulated by the presence of N37 modification and intraloop complementarity (table S1). This range is smaller than 1000-fold difference in strength between cognate and near-cognate interactions [4][5][6][7][8][9][10], enabling stability-based discrimination of the latter. Accordingly, most of split codon family readers fall within the lower rank of cognate codon-anticodon stabilities (Fig. 4F). Consequently, third-position mismatches formed by these readers fall into the lowest stability rank. Therefore, if the activation barrier for spontaneous decoding site closure is set below some average barrier for the dissociation of cognate split codon readers, their 3rd-position mismatches would be selected under thermodynamic control (Fig. 4G).

The fidelity of the translational machinery appears to rely primarily on the conservative mechanisms governing RNA-RNA interactions and the intrinsic dynamics of macromolecular RNA structures, potentially tracing back to a pre-protein world. Our findings challenge the notion that tRNA plays an active role in selection process. Instead, we propose that the information storage capacity of tRNA sequences is fully expended to maintaining the universal core structure and specific aminoacylation determinants [68]. Its primary functions are to balance the strength of codon-anticodon interactions, allowing the rejection of 3rd-position mismatches under thermodynamic control, and to convert sequence information into a molecular shape that enables rapid rejection of 1st- and 2nd-position mismatches based on geometric fit.

Statements and Declarations

Funding

This work was supported in part by the Australian Research Council Discovery Projects DP 210104020 as well as ARC Centre of Excellence in Synthetic Biology CE200100029 to KA. KA gratefully acknowledges financial support of CSIRO-QUT Synthetic Biology Alliance.

Author Contributions

- Conceptualization: SM
- Methodology: SM, ZC
- Investigation: SM, YW
- Visualization: SM
- Funding acquisition: KA
- Project administration: KA, SM
- Supervision: SM, KA
- Writing – original draft: SM, KA
- Writing – review & editing: SM, KA

Conflicts of Interest

Authors declare that they have no competing interests.

Data and materials availability

The code used for analyzing trends in kinetic parameters of split-tRNA selection is available on GitHub at https://github.com/mureich81/tRNASelectionModel.git. For detailed information, refer to the README file in the GitHub repository or to the step-by-step description of the algorithm in fig. S13.

The structures of canonical and wobble-containing A-helices are available under the following PDB entries: 2mkn[69], 2az2[70], 2jxq[71], 6db9[72], 433D[73], 472d[74], 5e7k[75]. The structures of codon-anticodon stem-loop complexes correspond to the following PDB entries: 6of6 (ala)[76], 4v6f (phe)[77], 2uu9 (val)[78].

Acknowledgments

The authors would like to acknowledge Dr. Patricia Walden and Micaela Fiorito for their excellent organizational assistance. The authors also thank Dr. Elena Eremeeva for her help with RNA fractionation, Roxane Mutschler for providing reverse transcriptase, and Dr. Zhong Guo for discussions on some aspects of enzyme kinetics.

References

1. ^*J. D. Mallory, A. B. Kolomeisky, O. A. Igoshin, Trade-Offs between Error, Speed, Noise, and Energy Dissipation in Biological Processes with Proofreading. J. Phys. Chem. B 123, 4718–4725 (2019). doi:10.1021/acs.jpcb.9b03757*
2. ^*K. Banerjee, A. B. Kolomeisky, O. A. Igoshin, Elucidating interplay of speed and accuracy in biological error correction. Proc. Natl. Acad. Sci. U. S. A. 114, 5183–5188 (2017). doi:10.1073/pnas.1614838114*
3. ^*I. Tubulekas, D. Hughes, Suppression of rpsL phenotypes by tuf mutations reveals a unique relationship between translation elongation and growth rate. Mol. Microbiol. 7, 275–284 (1993). doi:10.1111/J.1365-2958.1993.TB01118.X*
4. a, b*R. C. Thompson, A. M. Karim, The accuracy of protein biosynthesis is limited by its speed: High fidelity selection by ribosomes of aminoacyl-tRNA ternary complexes containing GTP[yS]. Proc. Natl. Acad. Sci. U. S. A. 79, 4922–4926 (1982). doi:10.1073/pnas.79.16.4922*
5. a, b, c*R. C. Thompson, D. B. Dix, Accuracy of protein biosynthesis. A kinetic study of the reaction of poly (U)-programmed ribosomes with a leucyl-tRNA2-elongation factor Tu-GTP complex. J. Biol. Chem. 257, 6677–6682 (1982). doi:10.1016/s0021-9258(18)34482-x*
6. a, b*U. Kothe, M. V. Rodnina, Codon Reading by tRNAAla with Modified Uridine in the Wobble Position. Mol. Cell 25, 167–174 (2007). doi:10.1016/j.molcel.2006.11.014*
7. a, b*F. A. P. Vendeix, A. Dziergowska, E. M. Gustilo, W. D. Graham, B. Sproat, A. Malkiewicz, P. F. Agris, Anticodon domain modifications contribute order to tRNA for ribosome-mediated codon binding. Biochemistry 47, 6117–6129 (2008). doi:10.1021/bi702356j*

8. [a, b, c]*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 interactions trigger the hydrolysis of GTP by EF-Tu. Cell 111, 721–732 (2002). doi:10.1016/s0092-8674(02)01086-3*
9. [a, b, c, d]*K. B. Gromadski, M. V. Rodnina, Kinetic Determinants of High-Fidelity tRNA Discrimination on the Ribosome. Mol. Cell 13, 191–200 (2004). doi:10.1016/s1097-2765(04)00005-x*
10. [a, b, c, d]*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). doi:10.1016/j.molcel.2005.12.018*
11. [^]*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). doi:10.1073/pnas.1116480109*
12. [a, b, c, d, e, f]*J. Zhang, M. Y. Pavlov, M. Ehrenberg, Accuracy of genetic code translation and its orthogonal corruption by aminoglycosides and Mg2+ ions. Nucleic Acids Res. 46, 1362–1374 (2018). doi:10.1093/nar/gkx1256*
13. [a, b]*S. Ledoux, M. Olejniczak, O. C. Uhlenbeck, A sequence element that tunes Escherichia coli tRNAGGCAla to ensure accurate decoding. Nat. Struct. Mol. Biol. 16, 359–364 (2009). doi:10.1038/nsmb.1581*
14. [a, b, c, d, e]*P. Satpati, J. Sund, J. Åqvist, Structure-based energetics of mRNA decoding on the ribosome. Biochemistry 53, 1714–22 (2014). doi:10.1021/bi5000355*
15. [^]*M. Yarus, Proofreading, NTPases and translation: successful increase in specificity. Trends Biochem. Sci. 17, 171–174 (1992). doi:10.1016/0968-0004(92)90257-A*
16. [^]*V. I. Lim, J. F. Curran, Analysis of codon:anticodon interactions within the ribosome provides new insights into codon reading and the genetic code structure. RNA 7, 942–957 (2001). doi:10.1017/S135583820100214X*
17. [a, b]*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). doi:10.1261/rna.039792.113*
18. [^]*K. Y. Sanbonmatsu, S. Joseph, Understanding discrimination by the ribosome: Stability testing and groove measurement of codon-anticodon pairs. J. Mol. Biol. 328, 33–47 (2003). doi:10.1261/rna.039792.113*
19. [a, b]*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). doi:10.1261/rna.294907*
20. [^]*J. Zhang, K. W. Ieong, M. Johansson, M. Ehrenberg, Accuracy of initial codon selection by aminoacyl-tRNAs on the mRNA-programmed bacterial ribosome. Proc. Natl. Acad. Sci. U. S. A. 112, 9602–9607 (2015). doi:10.1073/pnas.1506823112*
21. [^]*M. Y. Pavlov, A. Liljas, M. Ehrenberg, A recent intermezzo at the Ribosome Club. Philos Trans R Soc Lond B Biol Sci 372, 1–10 (2017). doi: 10.1098/rstb.2016.0185.*
22. [^]*J. M. Ogle, V. Ramakrishnan, Structural insights into translational fidelity. Annu Rev Biochem 74, 129–177 (2005). doi: 10.1146/annurev.biochem.74.061903.155440.*
23. [^]*T. Martin Schmeing, R. M. Voorhees, A. C. Kelley, Y. G. Gao, F. V. Murphy IV, J. R. Weir, V. Ramakrishnan, The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science 326, 688–694 (2009). doi:*

10.1126/science.1179700.

24. ^*M. Valle, A. Zavialov, W. Li, S. M. Stagg, J. Sengupta, R. C. Nielsen, P. Nissen, S. C. Harvey, M. Ehrenberg, J. Frank, Erratum: Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy Nature Structural Biology 10, 899–906 (2003). doi:10.1038/nsb1203-1074.*
25. a, b*L. Cochella, R. Green, An active role for tRNA in decoding beyond codon:anticodon pairing. Science 308, 1178–1180 (2005). doi: 10.1126/science.1111408.*
26. ^*R. F. Ortiz-Meoz, R. Green, Functional elucidation of a key contact between tRNA and the large ribosomal subunit rRNA during decoding. RNA 16, 2002–2013 (2010). doi: 10.1261/rna.2232710.*
27. ^*D. W. Schultz, M. Yarus, tRNA structure and ribosomal function. I. tRNA nucleotide 27-43 mutations enhance first position wobble. J. Mol. Biol. 235, 1381–1394 (1994). doi: 10.1006/jmbi.1994.1095.*
28. a, b*K. Y. Sanbonmatsu, Alignment/misalignment hypothesis for tRNA selection by the ribosome. Biochimie 88, 1075–1089 (2006). doi: 10.1016/j.biochi.2006.07.002*
29. ^*M. Yarus, D. Smith, tRNA: Structure, Biosynthesis, and Function. tRNA on the Ribosome: a Waggle Theory, 443–469 (1994). doi:10.1016/j.biochi.2006.07.002*
30. a, b*J. Frank, J. Sengupta, H. Gao, W. Li, M. Valle, A. Zavialov, M. Ehrenberg, The role of tRNA as a molecular spring in decoding, accommodation, and peptidyl transfer. FEBS Letters 579, 959–962 (2005) doi: 10.1016/j.febslet.2004.10.105.*
31. ^*H. Grosjean, E. Westhof, An integrated, structure- and energy-based view of the genetic code. Nucleic Acids Res. 44, 8020–8040 (2016). doi: 10.1093/nar/gkw608.*
32. ^*U. Lagerkvist, "Two out of three": An alternative method for codon reading. Proc. Natl. Acad. Sci. U. S. A. 75, 1759–1762 (1978). doi: 10.1073/pnas.75.4.1759.*
33. a, b*M. Yarus, S. W. Cline, P. Wier, L. Breeden, R. C. Thompson, Actions of the anticodon arm in translation on the phenotypes of RNA mutants. J. Mol. Biol. 192, 235–255 (1986). doi: 10.1016/0022-2836(86)90362-1.*
34. a, b*W. H. Mcclain, J. Schneider, S. Bhattacharya, K. Gabriel, The importance of tRNA backbone-mediated interactions with synthetase for aminoacylation. Proc. Natl. Acad. Sci. U. S. A. 95, 460–465 (1998). doi: 10.1073/pnas.95.2.460.*
35. a, b*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). doi: 10.1038/nsmb978*
36. a, b*C. Claesson, F. Lustig, T. Borén, C. Simonsson, M. Barciszewska, U. Lagerkvist, Glycine codon discrimination and the nucleotide in position 32 of the anticodon loop. J. Mol. Biol. 247, 191–196 (1995). doi: 10.1006/jmbi.1994.0132.*
37. a, b*F. Lustig, T. Borén, C. Claesson, C. Simonsson, M. Barciszewska, U. Lagerkvist, The nucleotide in position 32 of the tRNA anticodon loop determines ability of anticodon UCC to discriminate among glycine codons. Proc. Natl. Acad. Sci. U. S. A. 90, 3343–7 (1993). doi: 10.1073/pnas.90.8.3343.*
38. a, b, c*M. Sprinzl, K. S. Vassilenko, Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. 33, r53–188 (2005). doi: 10.1093/nar/15.suppl.r53.*
39. a, b*G. Varani, Exceptionally stable nucleic acid hairpins. Annu Rev Biophys Biomol Struct 24, 379–404 (1995). doi: 10.1146/annurev.bb.24.060195.002115.*

40. [a, b]*J. F. Curran, M. Yarus, Reading frame selection and transfer RNA anticodon loop stacking. Science 238, 1545–1550 (1987). doi: 10.1126/science.3685992.*
41. [a, b, c]*W. Saenger, Principles of Nucleic Acid Structure (Springer New York, New York, NY, 1984).*
42. [a, b, c]*P. N. Borer, B. Dengler, I. Tinoco, O. C. Uhlenbeck, Stability of ribonucleic acid double-stranded helices. J. Mol. Biol. 86, 843–853 (1974). doi: 10.1016/0022-2836(74)90357-x.*
43. [a, b]*C. M. Dunham, M. Selmer, S. S. Phelps, A. C. Kelley, T. Suzuki, S. Joseph, V. Ramakrishnan, Structures of tRNAs with an expanded anticodon loop in the decoding center of the 30S ribosomal subunit. RNA 13, 817–823 (2007). doi: 10.1261/rna.367307.*
44. [a, b]*D. Schwarz, F. Junge, F. Durst, N. Frölich, B. Schneider, S. Reckel, S. Sobhanifar, V. Dötsch, F. Bernhard, Preparative scale expression of membrane proteins in Escherichia coli-based continuous exchange cell-free systems. Nat. Protoc. 2, 2945–2957 (2007). doi: 10.1038/nprot.2007.426.*
45. [a, b]*Z. Cui, V. Stein, Z. Tnimov, S. Mureev, K. Alexandrov, Semisynthetic tRNA Complement Mediates in Vitro Protein Synthesis. J. Am. Chem. Soc. 137, 4404–4413 (2015). doi: 10.1021/ja5131963.*
46. [a, b]*H. Dong, L. Nilsson, C. G. Kurland, Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. J. Mol. Biol. 260, 649–663 (1996). doi: 10.1006/jmbi.1996.0428.*
47. [a, b]*O. Kovtun, S. Mureev, W. R. Jung, M. H. Kubala, W. Johnston, K. Alexandrov, Leishmania cell-free protein expression system. Methods 55, 58–64 (2011). doi: 10.1016/j.ymeth.2011.06.006.*
48. [a, b, c]*S. Mureev, O. Kovtun, U. T. T. Nguyen, K. Alexandrov, Species-independent translational leaders facilitate cell-free expression. Nat. Biotechnol. 27, 747–752 (2009). doi: 10.1038/nbt.1556.*
49. [a, b]*A. L. Konevega, N. G. Soboleva, V. I. Makhno, Y. P. Semenkov, W. Wintermeyer, M. V. Rodnina, V. I. Katunin, 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 (2004). doi: 10.1261/rna.5142404.*
50. [a, b]*S. Jafilan, L. Klein, C. Hyun, J. Florián, Intramolecular base stacking of dinucleoside monophosphate anions in aqueous solution. J. Phys. Chem. B 116, 3613–3618 (2012). doi: 10.1021/jp209986y.*
51. [a, b, c]*L. C. Sowers, B. R. Shaw, W. D. Sedwick, Base stacking and molecular polarizability: Effect of a methyl group in the 5-position of pyrimidines. Biochem. Biophys. Res. Commun. 148, 790–794 (1987). doi: 10.1016/0006-291x(87)90945-4.*
52. [a, b, c]*K. M. Guckian, B. A. Schweitzer, R. X. F. Ren, C. J. Sheils, D. C. Tahmassebi, E. T. Kool, Factors contributing to aromatic stacking in water: Evaluation in the context of DNA. J. Am. Chem. Soc. 122, 2213–2222 (2000). doi: 10.1021/ja9934854.*
53. [a, b]*D. Svozil, P. Jungwirth, Z. Havlas, Electron binding to nucleic acid bases. Experimental and theoretical studies. A review. Annu Rev Phys Chem 65, 275–292 (2004). doi: 10.1146/annurev-physchem-040513-103626.*
54. [a, b]*J. Šponer, J. Leszczynski, P. Hobza, Structures and energies of hydrogen-bonded DNA base pairs. A nonempirical study with inclusion of electron correlation. J. Phys. Chem. 100, 1965–1974 (1996). doi: 10.1021/jp952760f*
55. [a, b]*R. Luchowski, S. Krawczyk, Electroabsorption (Stark effect) spectroscopy of monomeric purine and pyrimidine*

bases. Chem. Phys. 314, 309–316 (2005). doi: 10.1016/j.chemphys.2005.03.012

56. [a, b]*H. Rosemeyer, F. Seela, Modified purine nucleosides as dangling ends of DNA duplexes: The effect of the nucleobase polarizability on stacking interactions. J. Chem. Soc. Perkin Trans. 2 2 (2002). doi: 10.1039/b200077f*
57. [a, b]*J. Zhang, K. W. Ieong, H. Mellenius, M. Ehrenberg, Proofreading neutralizes potential error hotspots in genetic code translation by transfer RNAs. RNA 22, 896–904 (2016). doi: 10.1261/rna.055632.115.*
58. [a, b]*D. L. Dunkelmann, C. Piedrafita, A. Dickson, K. C. Liu, T. S. Elliott, M. Fiedler, D. Bellini, A. Zhou, D. Cervettini, J. W. Chin, Adding α,α-disubstituted and β-linked monomers to the genetic code of an organism. Nature 625, 603–610 (2024). doi: 10.1038/s41586-023-06897-6.*
59. [a, b]*J. Mittelstaet, A. L. Konevega, M. V. Rodnina, Distortion of tRNA upon near-cognate codon recognition on the ribosome. J. Biol. Chem. 286, 8158–8164 (2011). doi: 10.1074/jbc.M110.210021.*
60. [a, b]*P. Geggier, R. Dave, M. B. Feldman, D. S. Terry, R. B. Altman, J. B. Munro, S. C. Blanchard, Conformational sampling of aminoacyl-tRNA during selection on the bacterial ribosome. J. Mol. Biol. 399, 576–95 (2010). doi: 10.1016/j.jmb.2010.04.038.*
61. [a, b, c]*N. Fischer, P. Neumann, L. V. Bock, C. Maracci, Z. Wang, A. Paleskava, A. L. Konevega, G. F. Schröder, H. Grubmüller, R. Ficner, M. V. Rodnina, H. Stark, The pathway to GTPase activation of elongation factor SelB on the ribosome. Nature 540, 80–85 (2016). doi: 10.1038/nature20560.*
62. [a, b]*R. C. Thompson, D. B. Dix, A. M. Karim, The reaction of ribosomes with elongation factor Tu·GTP complexes. Aminoacyl-tRNA-independent reactions in the elongation cycle determine the accuracy of protein synthesis. J. Biol. Chem. 261, 4868–74 (1986). doi: 10.1016/s0021-9258(19)89185-8*
63. [a, b]*G. Sander, R. Ivell, J. B. Crechet, A. Parmeggiani, Interaction of Elongation Factor Tu with the Ribosome. A Study Using the Antibiotic Kirromycin. Biochemistry 19, 865–70 (1980). doi: 10.1021/bi00546a006.*
64. [a, b]*A. B. Loveland, G. Demo, N. Grigorieff, A. A. Korostelev, Ensemble cryo-EM elucidates the mechanism of translation fidelity. Nature 546, 113–117 (2017). doi: 10.1038/nature22397.*
65. [a, b]*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). doi: 10.1038/s41586-020-2447-x.*
66. [a, b]*P. Schrode, P. Huter, N. Clementi, M. Erlacher, Atomic mutagenesis at the ribosomal decoding site. RNA Biol. 14, 104–112 (2017). doi: 10.1080/15476286.2016.1256535.*
67. [a, b]*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. Mol. Cell 10, 799–807 (2002). doi: 10.1016/s1097-2765(02)00686-x.*
68. [a, b]*A. Saint-Léger, C. Bello, P. D. Dans, A. G. Torres, E. M. Novoa, N. Camacho, M. Orozco, F. A. Kondrashov, L. R. De Pouplana, Saturation of recognition elements blocks evolution of new tRNA identities. Sci. Adv. 2, 1–9 (2016). doi: 10.1126/sciadv.1501860.*
69. [a, b]*R. G. Burge, M. A. Martinez-Yamout, H. J. Dyson, P. E. Wright, Structural characterization of interactions between the double-stranded RNA-binding zinc finger protein JAZ and nucleic acids. Biochemistry 53, 1495–1510 (2014). doi: 10.1021/bi401675h.*

70. [a, b]*J. A. Chao, H. L. June, B. R. Chapados, E. W. Debler, A. Schneemann, J. R. Williamson, Dual modes of RNA-silencing suppression by Flock House virus protein B2. Nat. Struct. Mol. Biol. 12, 952–957 (2005). doi: 10.1038/nsmb1005.*

71. [a, b]*L. Popenda, R. W. Adamiak, Z. Gdaniec, Bulged adenosine influence on the RNA duplex conformation in solution. Biochemistry 47, 5059–5067 (2008). doi: 10.1021/bi7024904.*

72. [a, b]*S. A. Shelke, Y. Shao, A. Laski, D. Koirala, B. P. Weissman, J. R. Fuller, X. Tan, T. P. Constantin, A. S. Waggoner, M. P. Bruchez, B. A. Armitage, J. A. Piccirilli, Structural basis for activation of fluorogenic dyes by an RNA aptamer lacking a G-quadruplex motif. Nat. Commun. 9, 4542 (2018). doi: 10.1038/s41467-018-06942-3.*

73. [a, b]*J. Trikha, D. J. Filman, J. M. Hogle, Crystal structure of a 14 bp RNA duplex with non-symmetrical tandem G·U wobble base pairs. Nucleic Acids Res. 27, 1728–1739 (1999). doi: 10.1093/nar/27.7.1728.*

74. [a, b]*J. Deng, M. Sundaralingam, Synthesis and crystal structure of an octamer RNA r (guguuuac)/r (guaggcac) with G·G/U·U tandem wobble base pairs: Comparison with other tandem G·U pairs. Nucleic Acids Res. 28, 4376–4381 (2000). doi: 10.1093/nar/28.21.4376.*

75. [a, b]*A. Rozov, N. Demeshkina, I. Khusainov, E. Westhof, M. Yusupov, G. Yusupova, Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. Nat. Commun. 7, 10457 (2016). doi: 10.1038/ncomms10457.*

76. [a, b]*H. A. Nguyen, S. Sunita, C. M. Dunham, Disruption of evolutionarily correlated tRNA elements impairs accurate decoding. Proc. Natl. Acad. Sci. U. S. A. 117, 16333–16338 (2020). doi: 10.1073/pnas.2004170117.*

77. [^]*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). doi: 10.1038/nature10913.*

78. [^]*A. Weixlbaumer, F. V. Murphy IV, A. Dziergowska, A. Malkiewicz, F. A. P. Vendeix, P. F. Agris, V. Ramakrishnan, Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. Nat. Struct. Mol. Biol. 14, 498–502 (2007). doi: 10.1038/nsmb1242.*