tRNA Methylation Resolves Codon Usage Bias

at the Limit of Cell Viability

Isao Masuda[1], Yuka Yamaki[1], Rajesh Detroja[2], Somnath Tagore[2,3], Henry Moore[4,5],

Sunita Maharjan[1], Yuko Nakano[1], Thomas Christian[1], Ryuma Matsubara[1,6], Todd Lowe[4,5],

Milana Frenkel-Morgenstern[2], and Ya-Ming Hou[1,*]

[1]Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia,

PA 19107, USA

[2]The Azrieli Faculty of Medicine, Bar-Ilan University, Safed 1311502, Israel

[3]Department of Systems Biology, Columbia University Medical Center, Herbert Irving Cancer

Research Center, New York, NY 10032, USA

[4]Department of Biomolecular Engineering, Baskin School of Engineering, University of California

Santa Cruz, Santa Cruz, CA 95064, USA

[5]UCSC Genomics Institute, University of California, Santa Cruz, Santa Cruz, CA 95064, USA

[6]Institute of Cell Biology, University of Bern, 3012 Bern, Switzerland

[*]Corresponding author: ya-ming.hou@jefferson.edu; (T) 215-503-4480; (F) 215-503-4954

Lead contact: Ya-Ming Hou (ya-ming.hou@jefferson.edu)

Key words: TrmD, m[1]G37-tRNA, proS, cmo[5]U34, Pro(UGG) isoacceptor

Summary:

The codon usage of each genome is closely correlated with the abundance of tRNA isoacceptors. How codon-usage bias is resolved by tRNA post-transcriptional modifications is largely unknown. Here we demonstrate that the N^1-methylation of guanosine at position 37 (m^{1}G37) on the 3'-side of the anticodon, while not directly responsible for reading of codons, is a neutralizer that resolves differential decoding of proline codons. A genome-wide suppressor screen of a non-viable *E. coli* strain, m^{1}G37-lacking, identifies *proS* suppressor mutations, indicating a coupling of methylation with tRNA prolyl-aminoacylation that sets the limit of cell viability. Using these suppressors, where prolyl-aminoacylation is decoupled from tRNA methylation, we show that m^{1}G37 neutralizes differential translation of proline codons by the major isoacceptor. Lacking of m^{1}G37 inactivates this neutralization and exposes the need for a minor isoacceptor for cell viability. This work has medical implications for bacterial genomes that exclusively use the major isoacceptor for survival.

Introduction

Codon usage is a unique feature of each gene and each genome and it impacts the fitness of each cell. In the degeneracy of the genetic code, proteins can be coded in multiple ways using different sets of synonymous codons, which are not translated equally in speed or quality. Each codon choice makes a specific demand for the supply of the tRNA isoacceptors that have the matching anticodons. This supply-to-demand balance is broadly observed throughout the domains of life (e.g., *Escherichia coli* (Ikemura, 1981), *Saccharomyces cerevisiae* (dos Reis et al., 2004), *Caenorhabditis elegans* (Duret, 2000), *Drosophila spp.* (Moriyama and Powell, 1997), and eukaryotes in general (Sabi and Tuller, 2014)). It ensures that frequent codons are translated by abundant tRNAs to speed up local translation, whereas rare codons are met with rare tRNAs to slow down local translation (Hanson and Coller, 2018). The impact of this balance is manifested in multiple aspects of a cellular life, including expression level of protein, accuracy of translation, maturity of protein folding, and stability of mRNA (Rak et al., 2018). However, the quality of a codon-anticodon pair is determined not only by the abundance of each isoacceptor, but also by the epigenetic modifications to the anticodon of the isoacceptor that are synthesized post-transcriptionally (de Crecy-Lagard and Jaroch, 2021; Maraia and Arimbasseri, 2017; Pan, 2018). The evolution of epigenetic modifications is an important mechanism to explore the wobble rule of the genetic code, where the wobble nucleotide of the anticodon (at position 34 of tRNA) can form non-canonical base pairing interactions with the 3rd nucleotide of a codon, thus enabling one isoacceptor to translate multiple codons (Agris et al., 2018). Indeed, many genomes, across all three domains of life, encode just one isoacceptor for all synonymous codons. How a single isoacceptor translates all synonymous codons with the speed and quality that is necessary to support life is poorly understood.

The wobble nucleotide at position 34 of tRNA hosts one of the highest chemical diversity and density of post-transcriptional modifications. At this position, several modifications have the

ability to expand the anticodon-codon pairing from a canonical W-C structure to wobble structures, thus impacting translation across multiple families of tRNAs. For example, the position-34 modification of adenosine-to-inosine (A34-to-I34), expanding base pairing from A34 with U to I34 with C, U, and A (Torres et al., 2014), is conserved in Eukarya in eight families of tRNAs (Ala, Arg, Ile, Leu, Pro. Ser, Thr, and Val). This expansion is important for neuron health in human (Ramos et al., 2020), and is most critical in organisms that use NNC as the preferred codons, but lack the corresponding tRNA(GNN) isoacceptors, implying that the I34-modified tRNA(ANN) isoacceptors are solely responsible for translation of NNC codons. Indeed, I34-knockout (KO) is lethal to cells, while I34-deficiency (KD) changes the profile of tRNA distribution and re-programs expression of NNC-enriched mRNAs (Lyu et al., 2020). Similarly, several other position-34 modifications in tRNA are also known to expand the capacity of decoding (Agris *et al.*, 2018). Deficiency of these modifications also produces changes of gene expression (Begley et al., 2007; Chionh et al., 2016; Nedialkova and Leidel, 2015; Zinshteyn and Gilbert, 2013). However, not all wobble base pairings are created equal, raising the question of whether some wobble base pairings may be deficient in thermodynamic stability and thus would be discriminated against during the high demand of global protein synthesis to support life. This question is most relevant to the deficiency that threatens cell viability, which impacts on the genome structure in evolution.

We hypothesize that the potential deficiency of a wobble base pairing can be compensated by additional post-transcriptional modifications in tRNA. Of interest is the position-37 modification on the 3'-side of the anticodon, which is similarly populated in the density and diversity of chemical structures as the position-34 modification. Importantly, while the position-37 modification is not within the anticodon triplet and is not directly involved in the anticodon-codon pairing, it is often conserved with the position-34 modification or other anticodon-loop modifications in a so-called "modification circuit" (Han and Phizicky, 2018). For example, the

t^6A37 (N^6-threonyl-carbamoyl adenosine) modification in *E. coli* Leu(CAA) and Leu(UAA) is in a circuit with the 2'-*O*-methylation of C34/U34 (Cm/Um); the yW37 (wybutosine) modification in yeast and human $tRNA^{Phe}$ is in a circuit with Cm32 and Gm34; the t^6A37 modification in *S. cerevisiae* $tRNA^{Thr}$ is in a circuit with m^3C32; and the i^6A37 (N^6-isopentenyl adenosine) modification in *S. pombe* and *S. cerevisiae* is in a circuit with m^3C32 of $tRNA^{Ser}$. These circuits are thought to enhance enzyme modification specificity, such that the first modification in the circuit is a recognition element that facilitates the second modification (Han and Phizicky, 2018). While this notion rationalizes the ordered enzymatic synthesis of the modification circuit, which is supported by some evidence (Han and Phizicky, 2018; Masuda et al., 2019), the biological question of why the circuit is formed is unanswered. Particularly, whether the position-37 modification compensates the deficiency of the position-34 modification to improve the quality of anticodon-codon pairing at the level necessary for cell viability is unknown.

Here we address this question focusing on the m^1G37 methylation that is invariably associated with tRNAs of the Pro, Leu, and Arg families, and is specifically associated with isoacceptors of Pro(CGG, GGG, and UGG), Leu(CAG, GAG, and UAG), and Arg(CCG) (Bjork and Hagervall, 2014). Deficiency of m^1G37 stalls ribosomes at most of the associated codons (Masuda et al., 2021), clearly indicating a role of the methylation in regulating the quality of anticodon-codon pairing. However, the biology of m^1G37 is complex. It is required for accuracy of the translational reading frame (Gamper et al., 2021a; Gamper et al., 2021b; Gamper et al., 2015; Hoffer et al., 2020); loss of m^1G37 results in +1 frameshifting and premature termination of protein synthesis, leading to cell death (Bjork et al., 1989; Masuda *et al.*, 2019). It is also required for efficient and specific aminoacylation of tRNA (Masuda *et al.*, 2021; Perret *et al.*, 1990); loss of m^1G37 results in uncharged tRNA that stalls the ribosome at m^1G37-dependent codons (Masuda *et al.*, 2021), potentially also leading to cell death. In this complex biology of m^1G37, we explored its essentiality for cell viability to elucidate a potential role in anticodon-

5

codon pairing by removing the methylation to "expose" any deficiency that would be otherwise hidden. We performed a genome-wide suppressor screen to isolate suppressors that support cell growth when m^1G37 is absent. This is a systematic and unbiased approach to identify the gene network that operates at the limit of cell viability. Unexpectedly, despite the association of m^1G37 with multiple tRNAs and multiple functions, all of the suppressors in our screen are mapped to a single gene: *proS* for prolyl-tRNA synthetase (ProRS) responsible for prolyl-aminoacylation of tRNA. This exclusiveness indicates that it is the coupling of methylation with prolyl-aminoacylation that distinguishes life from death. In the *proS* suppressors, however, this coupling is disrupted, enabling us to examine anticodon-codon pairings by prolyl isoacceptors that are now "charged" but "not methylated". We show that a minor prolyl isoacceptor, which is normally dispensable when m^1G37 is present, has become indispensable when m^1G37 is absent, revealing a deficiency of the anticodon-codon pairing upon removal of the methylation that needs to be addressed by the minor isoacceptor. This deficiency results from the inability of the major prolyl isoacceptor to use wobble pairings to decode a subset of codons. Thus, m^1G37 is the determinant that resolves codon-reading bias of the major prolyl isoacceptor. We find that many bacterial pathogens of serious threat to human health do not encode the minor prolyl isoacceptor, suggesting that cell fitness and viability of these pathogens would be severely compromised in m^1G37 deficiency.

Results

A genome-wide suppressor screen identifies *proS* mutations in m^1G37 deficiency

We used *E. coli* as a genetic model, because all of the isoacceptors that are associated with m^1G37 are well defined and experimentally tractable (Bjork and Hagervall, 2014). As with most bacterial species, *E. coli* encodes a complete set of all essential isoacceptors, as well as an additional set of non-essential tRNAs, which together assemble into a more-than-sufficient

ribosome-tRNA apparatus that readily translates all of the sense codons of the genetic code. In bacteria, m^1G37 is synthesized by the conserved TrmD methyl transferase that is also essential for cell viability (Bystrom and Bjork, 1982) (Figure 1A). This essentiality suggests that, to generate an m^1G37-deficient condition, a simple *trmD*-KO cannot be made. We therefore created a modified *E. coli trmD*-KO strain, in which the chromosomal *trmD* was replaced by the antibiotic Kan marker, while cell survival was maintained by expression of the gene under the control of an arabinose (Ara)-inducible promoter from a plasmid that had a temperature-sensitive (*ts*) origin of replication (Figure 1B). This strain was grown overnight in the presence of Ara at the permissive temperature 30 °C, and was then diluted into an Ara-free medium at the non-permissive temperature 43 °C for multiple cycles of growth and dilution. This cycling between growth and dilution at 43 °C was designed to irradicate the *trmD*-carrying plasmid and to eliminate residual levels of the TrmD enzyme and its m^1G37-tRNA products. After three cycles of growth and dilution (~20 h) at 43 °C, when cell counts were no longer visible indicating cell death, an aliquot of the culture was streaked on an Ara-free plate and incubated at 43 °C to screen for suppressors. Importantly, to minimize a potential bias of the screen, due to differences in the genetic composition of the *E. coli* host, we performed the screen in parallel using three well-characterized strains (see Methods), each constructed and grown in the same modified *trmD*-KO condition. We isolated 5 colonies from MG1655, 4 colonies from BW25113, and 3 colonies from XAC. Each isolate was streaked-purified, confirmed for the loss of the maintenance plasmid (Figure S2A), and subjected to whole-genome sequencing analysis.

To our surprise, all but one clones mapped their single-nucleotide substitutions to the gene *proS*, totaling in 9 unique mutations (Figure 1C). The exception was a single clone from BW25113 that did not yield a clear whole-genome sequencing result. The exclusive mapping of all suppressor mutations to *proS*, occurring across the three strains, indicates a connection between m^1G37 methylation and prolyl-aminoacylation of tRNA that is required for cell viability.

7

In this connection, m^{1}G37 promotes efficient prolyl-aminoacylation to generate Pro-charged tRNA^{Pro} isoacceptors that support ribosome translation of Pro codons. This connection is consistent with our published work showing that loss of m^{1}G37 reduces prolyl-aminoacylation, causing ribosome stalling at Pro codons and arresting cell growth (Masuda *et al.*, 2021). Thus, while m^{1}G37 is associated with multiple tRNAs in multiple aspects of translation, it is most intimately associated with prolyl-aminoacylation of tRNA at the nexus that determines cell death and survival.

Mapping of each unique suppressor mutation to a known crystal structure of the *proS* enzyme in complex with a prolyl analog identified three clusters (Yaremchuk et al., 2000) (Figure 1D, E). Cluster I mutations (D98N, G107C, S123R) are localized between motif 1 and motif 2 of the catalytic domain, which is formed at the monomer-monomer interface of the obligated dimer structure of the enzyme (Yaremchuk et al., 2001) (Figure S2B). Cluster II mutations (T199S, T199I, and Q211L) are between the catalytic motif 2 and an insertion domain (also known as the editing domain) (Beuning and Musier-Forsyth, 2000; Vargas-Rodriguez and Musier-Forsyth, 2013), the latter of which provides a proof-reading activity to remove mis-acylated tRNA^{Pro}. Cluster III mutations (A231E, N232S, and A236E) are near the start of the insertion domain. Some mutations in clusters II and III are also adjacent to the Pro-binding pocket of the enzyme (Figure S2B). Thus, all mutations are mapped to functionally important regions of *proS*. Additionally, all mutations are mapped to highly conserved positions of *proS* enzymes across a wide range of bacterial species (Figure 1D), indicating that they arose from the necessity to restore cell viability, even against a strong selective pressure of sequence conservation.

Each unique *proS* suppressor mutation was reconstructed by genome editing in the native MG1655 strain (Figure 2A), expressing *trmD*. In each edited strain, the chromosomal *trmD* was then deleted and confirmed for the loss of the *trmD* enzyme (Figure S2C). Note that the symbols

8

(+) and (−) indicate the presence and absence of *trmD,* respectively. Importantly, all reconstructed suppressors supported cell viability (Figure 2B), validating that each is a bona fide suppressor of the loss of viability upon *trmD*-KO. As shown for the *proS*-D98N and *proS*-A231E suppressors (Figure 2C, Figure S2D), however, each had reduced fitness over time in a growth competition assay with the WT strain. This reduced fitness was correlated with an increased frequency of ribosomal +1 frameshifting in a reporter assay based on expression of the nLuc gene that carried a frameshift-prone insertion (Figure 2D). Thus, consistent with the notion that m^{1}G37 ensures accuracy of the translational reading frame (Gamper *et al.*, 2021a; Gamper *et al.*, 2021b; Gamper *et al.*, 2015), these m^{1}G37-lacking suppressors are prone to +1 frameshifting, providing an explanation for their compromised cell fitness relative to the *trmD*-expressing WT strain. The slow fitness loss of these suppressors over the course of 24 h (Figure 2C) is likely due to the generally low efficiency of +1 frameshifting in the absence of m^{1}G37 (Gamper *et al.*, 2015).

Decoupling of m^{1}G37 methylation from tRNA aminoacylation in *proS* suppressors

As we showed recently (Masuda *et al.*, 2021), the WT *proS* enzyme requires the presence of m^{1}G37 for efficient prolyl-aminoacylation of tRNA^{Pro}, indicating a mechanistic coupling between tRNA methylation and prolyl-aminoacylation. The isolation of *proS* suppressors that are viable but lack m^{1}G37 suggests that this coupling is disrupted, such that each suppressor restores prolyl-aminoacylation without the need for m^{1}G37. We tested this possibility by probing the intracellular prolyl-aminoacylation status of each *proS* suppressor, lacking *trmD* (Figure 3A). Total tRNA was isolated from each suppressor and the prolyl-aminoacylation status at the time of cell harvest was monitored for the Pro(UGG) isoacceptor using a tRNA-specific oligonucleotide. This experiment was performed in an acid condition to preserve the charged state, starting with isolation of total tRNA to separation of the charged from the uncharged

states on a denaturing gel (Masuda *et al.*, 2021). The prolyl-aminoacylation status was calculated by the band intensity of the charged state over the sum of both the charged and uncharged state. The result showed that the prolyl-aminoacylation status was generally high across all of the *proS* suppressors, ranging from 63% for the A231E (–) suppressor to 38% and 42% for the T199S (–) and S123R (–) suppressors (Figure 3A). The high charging status of the A231E (–) suppressor at 63% is comparable to the status of 73% of the WT (+) strain that expresses *trmD*.

Notably, both a negative and a positive control of total tRNA were isolated from a separately constructed *trmD*-KD strain (Gamper *et al.*, 2015), where the chromosomal *trmD* was removed, while a plasmid-expressed human counterpart gene *trm5* under the control of an Ara-inducible promoter was used to maintain cell viability. In this *trmD*-KD strain, addition of Ara to the medium induced the m^{1}G37+ condition, producing a charging status of Pro at 58%, while depletion of Ara from the medium induced the m^{1}G37– condition, producing a charging status at 6% (Figure 3A). The choice of human *trm5* as the maintenance gene was because the protein product is rapidly degraded upon turning off the Ara-controlled promoter (Christian et al., 2013), thus immediately generating the m^{1}G37– condition. In contrast, a *trmD*-KD strain maintained by *E. coli trmD* would take much longer time to generate an m^{1}G37– condition (Masuda *et al.*, 2021), due to the high stability of the protein product in cells. We found that, relative to these controls, even the low charging status of the T199S (–) and S123R (–) suppressors (at 38-42%) is considerably higher than the charging status of the negative control and is within 2-fold of the positive control (Figure 3A). Thus, all of the isolated *proS* suppressors maintain a high intracellular prolyl-aminoacylation status, despite the absence of m^{1}G37.

We next probed the enzymatic prolyl-aminoacylation activity of each *proS* suppressor mutant, using *in vitro* kinetic assays to generate intrinsic parameters that define the catalytic efficiency of each enzyme. These kinetic assays served as an independent approach from cell-

based assays to verify the decoupling of aminoacylation from m^1G37 methylation. They also offered the kinetic sensitivity to distinguish among all prolyl isoacceptors for the level of decoupling. The *E. coli* genome expresses a complete set of prolyl isoacceptors, consisting of Pro(UGG), Pro(GGG), and Pro(CGG). To focus on the effect of m^1G37, each isoacceptor was prepared in the G37-state and in the m^1G37-state, differing in the presence and absence of the methylation. The G37-state was transcribed from each gene with natural nucleotides, lacking m^1G37 and any other modification, whereas the m^1G37-state was transcribed and was post-transcriptionally modified by TrmD with the single methylation (Gamper *et al.*, 2015).

Representative members of *proS* suppressor enzymes, each purified in the recombinant form to near homogeneity, were assayed for prolyl-aminoacylation with all three isoacceptors. The assay was performed in multiple-turnover condition as a function of tRNA concentration, yielding kinetic parameters K_m (the concentration of the tRNA that produces half-maximum of the enzyme activity), k_{cat} (the catalytic turnover), and k_{cat}/K_m (the catalytic efficiency per tRNA-binding event). For prolyl-aminoacylation of the Pro(UGG) isoacceptor, while the k_{cat}/K_m of the WT enzyme decreased by 20.7-fold upon loss of m^1G37, it decreased by no more than 3-fold for three *proS* suppressor enzymes (Figure 3B, Figure S3A). Specifically, the decrease was by 2.9-fold for the *proS*-D98N enzyme (a member of cluster I), by 3.0-fold for the *proS*-T199I enzyme (a member of cluster II), and by 2.5-fold for the *proS*-N232S enzyme (a member of cluster III). The large decrease of k_{cat}/K_m of the WT enzyme upon loss of m^1G37 is consistent with the decrease we observed previously (Masuda *et al.*, 2021), confirming the dependence on the methylation for aminoacylation, whereas the much smaller decrease of k_{cat}/K_m of the three suppressor enzymes indicates a loss of the dependence. Thus, across all three clusters of *proS* suppressor mutants, prolyl-aminoacylation is no longer tightly coupled to m^1G37 methylation.

The loss of coupling was also observed for prolyl-aminoacylation of the other two isoacceptors, using the *proS*-D98N mutant enzyme as an example. For prolyl-aminoacylation of

11

the Pro(GGG) isoacceptor, loss of m^1G37 decreased k_{cat}/K_m by 4.4-fold for the mutant enzyme, but by 20.8-fold for the WT enzyme (Figure 3C, Figure S3B). Similarly, for prolyl-aminoacylation of the Pro(CGG) isoacceptor, loss of m^1G37 decreased k_{cat}/K_m by 3.4-fold for the mutant enzyme, but by 18.8-fold for the WT enzyme (Figure 3D, Figure S3C). Thus, consistent across all three isoacceptors of *E. coli*, while prolyl-aminoacylation by the WT *proS* enzyme is strongly coupled to the presence of m^1G37, this coupling is weakened in suppressor mutants, resulting in synthesis of Pro-charged tRNA[Pro] in the absence of m^1G37, which underlies the basis for sustaining cell viability in *trmD*-KO.

The Pro(GGG) isoacceptor is required for cell viability of *proS* suppressors

The prolyl-aminoacylation activity of *proS* suppressors without the need for m^1G37 offered a way to generate Pro-charged tRNAs that were aminoacylated but not methylated. These tRNAs would then be competent to enter the ribosomal machinery, allowing us to investigate the quality of anticodon-codon pairing in the absence of the methylation. Without this prolyl-aminoacylation, none of the prolyl isoacceptors would be permitted to perform anticodon-codon pairing on the ribosome. The lack of m^1G37 from each isoacceptor then provided a framework to identify any bias against anticodon-codon pairing. In the native state of *E. coli* with m^1G37, each of the three prolyl isoacceptors has a dedicated responsibility for decoding. Pro(CGG) (encoded by *proK*) is dedicated for reading the CCG codon; Pro(GGG) (encoded by *proL*) is dedicated for reading the CCC and CCU (CC[C/U]) codons; and Pro(UGG) (encoded by *proM*) is capable of reading all four CCN codons, due to the presence of the cmo^5U34 (5-caroxy-methoxyl uridine 34) modification that expands the decoding capacity to all four nucleotides (Nasvall et al., 2004; 2007) (Figure 4A, Figure S1). Of the three, the UGG isoacceptor is the major species due to its high abundance in the prolyl family (Wei et al., 2019). It is also the only one that is essential for viability in the family and cannot be deleted (Masuda *et al.*, 2019; Nasvall *et al.*, 2004). We

therefore used the genetic tools available to *E. coli* and deleted either the GGG or the CGG isoacceptor from a *proS* suppressor strain. We then determined, upon removal of m^{1}G37, if any bias against an anticodon-codon pairing could be exposed.

We chose the previously constructed *trmD*-KD strain for this purpose (Gamper *et al.*, 2021a; Gamper *et al.*, 2015), considering that deletion of the GGG or the CGG isoacceptor in the absence of m^{1}G37 might be lethal if a bias exists against an anticodon-codon pairing that would block genome-wide translation. In this *trmD*-KD strain, we reconstructed a *proS* mutation to the host genome, where *trmD* was removed from the chromosome and cell viability was maintained by expression of a plasmid-borne human *trm5* under the control of an Ara-inducible promoter (Figure 4B). The benefit of using Ara to control the biosynthesis of m^{1}G37 is the ability to identify biases that could otherwise threaten cell viability in the m^{1}G37– condition.

The results showed that the *trmD*-KD strain, when supported by the WT *proS*, was viable in the m^{1}G37+ condition, but was non-viable in the m^{1}G37– condition, consistent with the notion that m^{1}G37 is required for efficient prolyl-aminoacylation to confer cell growth (Figure 4C, first row). In contrast, the *trmD*-KD strain, when supported by the *proS*-D98N suppressor, was viable in the m^{1}G37– condition, but this viability was observed only when the minor isoacceptor GGG was present (second row). Thus, while the *proS*-D98N suppressor was able to support cell viability without the need for m^{1}G37, it failed to do so when the GGG isoacceptor was also lacking. This dependence on the GGG isoacceptor for viability was also observed for the *proS*-T199I and *proS*-N232S suppressors (Figure S4), indicating a shared mechanism. Upon over-expression of a plasmid-borne GGG isoacceptor, however, cell viability was restored in the m^{1}G37– condition (Figure 4D, second row). This restoration was not observed in the *trmD*-KD strain supported by the WT *proS* (first row), or in the *proS*-D98N suppressor supported by over-expression of a variant of the GGG isoacceptor that harbored the UGG anticodon, or in the suppressor strain supported by over-expression of a CAU isoacceptor specific for reading of the

Met AUG codons (Figure 4D, right column). Thus, the *proS*-D98N suppressor requires Pro(GGG) for cell viability, implying that the lack of the isoacceptor has rendered decoding of the associated CC[C/U] codons insufficient. Normally, decoding of CC[C/U] is a shared responsibility between the minor GGG isoacceptor and the major UGG isoacceptor that is modified with cmo^5U34 (Figure 4A). The dependence on the minor isoacceptor for viability of the *proS*-D98N suppressor implies that decoding of CC[C/U] codons is no longer supported by the major isoacceptor in the absence of m^1G37.

Codon usage bias for expression of cold-shock response genes

We tested the hypothesis that, in the absence of m^1G37, translation of CC[C/U] codons would be dependent on the availability of the minor isoacceptor Pro(GGG). A genome-wide codon usage analysis of *E. coli* protein-coding genes showed an average usage of CC[C/U] at 0.31 when normalized by the sum of all Pro codons that occur in the same gene (Figure 5A). Interestingly, 20.4% of *E. coli* protein-coding genes display an average usage of CC[C/U] higher than 0.5, indicating an enrichment of these codons. In GO analysis, these CC[C/U]-enriched genes are most populated in the group involved in organization of cellular components in response to stress with a frequency of 4.2-fold above the background frequency (Figure 5B). Notably, many genes in this group encode cold-shock protein (Csp) products in response to temperature downshift (Figure S4A), possibly serving as chaperones to prevent mRNA from aggregation and to facilitate initiation of protein synthesis (Keto-Timonen et al., 2016). We tested the possibility that the high enrichment of CC[C/U] codons in *csp*-related genes would render the *E. coli* response to cold-shock dependent on Pro(GGG) when m^1G37 is absent. Indeed, the WT strain, when supported by the native *proS* while expressing *trmD* and synthesizing m^1G37, remained viable upon temperature shift from 37 °C to 22 °C, whereas several *proS* suppressors, lacking *trmD* and unable to synthesize m^1G37, did not (Figure 5C).

The loss of viability was notable for the *proS*-D98N mutant even with a milder temperature downshift to 30 °C (Figure S4B). The high sensitivity of these suppressors to the cold shock stress is consistent with the notion that, in the absence of m^{1}G37, translation of CC[C/U]-enriched *csp*-related genes is compromised, likely due to insufficient supply of the Pro(GGG) isoacceptor.

In *E. coli*, CC[C/U] codons are rare (average usage at 0.31, Figure 5A) relative to the more abundant CC[A/G] codons for Pro, raising the possibility that the intrinsic level of the corresponding ProGGG) would be similarly low relative to the others. This possibility was confirmed by over-expression of the isoacceptor from a multi-copy plasmid, showing restoration of viability at 22 °C (Figure 5D, Figure S5B). In contrast, no restoration of viability was observed by over-expression of the other two isoacceptors. The specific requirement of Pro(GGG), but not others, to restore cell viability at 22 °C in m^{1}G37-lacking suppressors supports the notion that it is the GGG isoacceptor that is responsible for translation of the highly enriched CC[C/U] codons that are necessary to produce the cold-shock response.

Absence of the Pro(GGG) isoacceptor from bacterial species

The requirement of the Pro(GGG) isoacceptor for cell viability in m^{1}G37-lacking suppressors is important for fundamental understanding of the genome structure across the diversity of bacterial species. While m^{1}G37 is strictly conserved throughout the bacterial domain (Bystrom and Bjork, 1982), not all bacterial species encode the minor Pro(GGG) isoacceptor. In such a simplified genome structure, decoding of CC[C/U] codons could only be achieved by the cmo^{5}U34-modified major isoacceptor Pro(UGG), while decoding of the CCG codon could be achieved by both the major isoacceptor and the other minor isoacceptor Pro(CGG). We thus surveyed the genomic tRNA database (GtRNAdb) to understand expression of all prolyl isoacceptors. We compared the presence and absence of each isoacceptor with a combined

15

16S and 23S ribosomal RNA (rRNA) phylogeny of selected genera in GtRNAdb (Figure 6A). Nearly all examined bacterial genomes encode the major isoacceptor Pro(UGG), supporting the essentiality of the isoacceptor for cell viability, while a few exceptions lacking the isoacceptor are likely due to incomplete analysis of the genomes. Although some bacterial species also encode a minor Pro(AGG) isoacceptor, this is extremely rare in occurrence (12 occurrences across the species surveyed) and has no pattern of evolutionary conservation, suggesting that they are likely tRNA pseudogenes and are therefore not considered further. The resulting phylogenetic tree revealed a bifurcation of bacterial species into phyla with either the presence or absence of the minor Pro(GGG) isoacceptor. Notably, almost all the clades that are missing Pro(GGG) are also missing Pro(CGG), indicating a co-existence relationship between these two minor isoacceptors (Figure S5). The absence of both isoacceptors from a genome thus signifies the sole dependence on the major isoacceptor for translation of all Pro codons. This is most notable in the entire class of Mycoplasma, which are intracellular parasites with minimal sets of tRNA isoacceptors (Razin *et al.*, 1998). The only exception to the co-existence rule of both minor isoacceptors is in the family Lactobacillaceae, which lack Pro(GGG) but contain Pro(CGG) (Figure 6A), an observation of interest for future studies.

Many bacterial species lacking the Pro(GGG) isoacceptor are infectious human pathogens. In Gram-positive genera, 55 of the 171 species (32%) lack Pro(GGG), and 23 of the 55 are human pathogens, including species of the phyla *Enterococcus* and *Staphylococcus* (Figure 6A, B). In Gram-negative genera, 57 of the 424 species (13%) lack Pro(GGG), and 25 of the 57 are human pathogens, including species of the phylum *Acinetobacter* (Figure 6A, C). Notably, the pathogens mentioned here constitute the ESKAPE family, posing an additional threat to human health, due to their acquisition of resistance to almost all currently available antibiotics (De Oliveira et al., 2020). The prevalence of lacking Pro(GGG) from these pathogens suggests a medical relevance that can be explored for antibiotic targeting.

16

Discussion

Codon usage is well known to be closely correlated with the abundance and composition of the cellular tRNA pool (Rak *et al.*, 2018). Much less known is how codon usage may be controlled or modulated by various tRNA post-transcriptional modifications. Even less known is how codon usage is determined by position-37 post-transcriptional modifications, which occur on the 3'-side of the anticodon and are not directly involved in the anticodon-codon pairing. Yet, elimination of a position-37 modification from a cell has resulted in ribosome stalling and programmatic changes of gene expression (Lamichhane et al., 2016; Masuda *et al.*, 2021; Thiaville *et al.*, 2016), strongly implicating that the position-37 modification modulates the quality of anticodon-codon pairing. This modulation however has never been directly tested. Here we show that the position-37 modification with m^1G37 is an important neutralizer that addresses the deficiency in, and bias against, specific anticodon-codon pairings established by the position-34 wobble modification in the major Pro(UGG) isoacceptor. Typically, U34 of the major isoacceptor is modified into cmo^5U34, which expands the decoding to all four nucleotides (Nasvall *et al.*, 2004; 2007). We show that, while this expansion supports cell viability when m^1G37 is present, it no longer can do so when m^1G37 is absent (Figure 4D). Thus, m^1G37 is the determinant that enables the cmo^5U34-modified UGG anticodon to expand decoding, whereas lacking m^1G37, the modified UGG anticodon is unable to read CC[C/U] codons, requiring the minor isoacceptor Pro(GGG) to assume the responsibility (Figure 7A). Notably, the minor isoacceptor Pro(GGG) is designed to only read CC[C/U] codons, although the reading is non-essential in m^1G37 abundance, due to coverage of the codons by the cmo^5U34-modified major isoacceptor. However, the reading of the minor isoacceptor becomes essential in m^1G37 deficiency, emphasizing the notion that m^1G37 is the neutralizer that resolves the deficiency of the cmo^5U34-modified major isoacceptor to read CC[C/U] codons. Importantly, m^1G37 is in a

conserved modification circuit with the chemical moiety of cmo^{5}U34 in the major isoacceptor Pro(UGG) throughout evolution. This includes the methyl addition to the cmo^{5}U34 moiety in *E. coli* Pro(UGG), generating mcmo^{5}U34 (5-methoxy-carbonyl-methoxy-uridine 34) in bacteria, and the ncm^{5}U34 (5-carbamoyl-methyl-uridine 34) modification in eukaryotes, sharing the carbamoyl chemical moiety at the 5-position of the uridine (Figure S1). This broad conservation indicates that the functional compensation of m^{1}G37 for the deficiency of cmo^{5}U34-mediated decoding of Pro codons as we report here has sustained a strong selective pressure over the evolution time.

The ability of m^{1}G37 to resolve the deficiency of the cmo^{5}U34-modified major isoacceptor now provides new insight into the long-standing question of how cmo^{5}U34 expands the pairing capacity to all four nucleotides. In crystal structures, the base pairing of cmo^{5}U34 with C and U is mediated by just one hydrogen (H)-bond each, whereas those with A and G are mediated by three H-bonds each (Figure S1) (Weixlbaumer et al., 2007). While these structures were obtained with a synthetic anticodon stem-loop (ASL), lacking the position-37 modification, they nonetheless raise the question of how the single H-bond can provide the same stability as three H-bonds. This question has remained unsolved for decades, due to the lack of a tool to probe the question at a biological scale. Here we use a genetic tool to remove m^{1}G37 and, upon removal, we find that the single H-bond in the cmo^{5}U34-C/U pairing is indeed insufficient and unable to satisfy the high demand of genome-wide translation required for cell viability. In the major isoacceptor Pro(UGG), stabilization of the cmo^{5}U34-C/U pairing by m^{1}G37 is likely rooted in the ability of the methylation to remodel the ASL structure, which adopts a disordered form when lacking m^{1}G37 (Maehigashi et al., 2014). This implication is presumably far reaching, given that the cmo^{5}U34-mediated anticodon-codon pairing is wide-spread among 5 other tRNA families, in addition to Pro(UGG). In *E. coli*, cmo^{5}U34 co-exists with m^{1}G37 in Leu(UAG), with ms^{2}i^{6}A37 in Ser(UGA), with m^{6}A37 in Val(UAC), and with ct^{6}A37 (cyclic N^6-threonyl-carbamoyl-adenosine 37) in Thr(UGU). Although cmo^{5}U34 in Ala(UGC) does not co-exist with a position-37

modification, the deficiency of its pairing with C/U may be resolved by other post-transcriptional modifications in the tRNA. Further structural analysis of a cmo^{5}U34-C/U base pairing in the presence of a position-37 modification, or accompanied by the full complement of all post-transcriptional modifications in a tRNA, will shed more light on how other tRNA modifications stabilize the cmo^{5}U34-C/U base pairing.

While we emphasize the importance of m^{1}G37 in codon-reading of prolyl isoacceptors, this does not mean that m^{1}G37 is not important for other tRNAs that are associated with the methylation (e.g., isoacceptors of Leu and Arg). The distinction is that the importance of m^{1}G37 in prolyl isoacceptors is established at the limit of cell viability, whereas the importance of m^{1}G37 in others is not as critical. Indeed, all of the mutations in our genome-wide screen of suppressors of m^{1}G37-KO are mapped to the single gene *proS*, indicating a unique dependence of m^{1}G37 for prolyl-aminoacylation that is required for cell viability. In contrast, while we showed previously a similar dependence on m^{1}G37 for aminoacylation of Arg(CCG) and Leu(UAG) (Masuda et al., 2021), no suppressor mutations are mapped to *argS* or *leuS*. These results suggest that, in the decision of life and death, cells choose to introduce mutations into *proS* at highly conserved and functionally important regions that would compromise activity of the enzyme in exchange for viability. A detailed analysis of selected *proS* suppressor mutants verifies that each has a kinetic deficiency relative to the WT enzyme (Figure S3), ranging in a loss of k_{cat}/K_m from 4- to 170-fold. This kinetic deficiency is apparently small enough of a price for the trade-off to have the ability to continue synthesis of Pro-charged tRNA^{Pro} to support protein synthesis and cell viability. Nonetheless, the trade-off does incur a fitness cost. One reason is the increased propensity of ribosomal +1 frameshifting (Figure 2D), consistent with the notion that loss of m^{1}G37 induces the ribosome to shift to the +1-frame, resulting in changes of the reading frame and cell death (Gamper *et al.*, 2021a; Gamper *et al.*, 2015).

The dependence of *proS* suppressors on the minor isoacceptor Pro(GGG) for cell viability has important medical implication (Figure 7B). The methyl transferase TrmD has long been ranked as a high-priority antibacterial target (White and Kell, 2004), due to its essentiality for bacterial viability, conservation across the bacterial domain, fundamental distinction from its human counterpart Trm5 in structure and mechanism (Ahn et al., 2003; Christian et al., 2004; Christian and Hou, 2007; Christian et al., 2010; Christian et al., 2016; Goto-Ito et al., 2008; Goto-Ito et al., 2009; Sakaguchi et al., 2014), and possession of a small-molecule binding pocket for the methyl donor that can be targeted by new compounds. Notably, the methyl-donor binding pocket in TrmD is unique, adopting a rare protein topological knot-fold that is absent from the great majority of methyl transferases (Ahn *et al.*, 2003; Christian *et al.*, 2016; Elkins et al., 2003), indicating the potential of bacteria-specific drug targeting in novel chemical space and diversity. The importance of targeting TrmD is further emphasized by our finding that m^{1}G37-tRNA is required for translation of membrane-associated genes in the bacterial cell-envelop structure (Masuda *et al.*, 2019), due to an enrichment of CC[C/U] codons in these genes, suggesting that targeting can disrupt the membrane-controlled mechanisms to facilitate drug entry and intracellular retention for faster bactericidal action. Indeed, developing TrmD inhibitors is currently an active pursuit (Hill et al., 2013; Thomas et al., 2020; Whitehouse et al., 2019; Zhong et al., 2019a; Zhong et al., 2019b). However, in any antibacterial strategy that targets a single gene, the risk of resistance is a major concern. The discovery of *proS* suppressors in *trmD*-KO raises the concern of a potential resistance mechanism. This concern is now ameliorated by our subsequent discovery that these suppressors require the minor isoacceptor Pro(GGG) for survival. The prevalence of lacking the minor isoacceptor in many clinically relevant bacterial pathogens (Figure 6A) suggests that these pathogens would be attractive candidates for targeting *trmD* in a new antimicrobial strategy that would not encounter a resistance mechanism due to suppressor mutations of *proS*.

20

A recent study reported the isolation of suppressors of m^1G37 deficiency from several *E. coli* strains (Clifton et al., 2021), each harboring a chromosomal mutation in *trmD*. Although the majority of the isolated suppressor mutations are also mapped to *proS*, with some overlapping with the suppressors we report here, the approach and focus of the recent study is fundamentally different from ours (Table S2). Specifically, the recent study was set out to isolate suppressors of *trmD*-KD and concluded that mutations of *proS* constitute the dominant mechanism that repairs the deficiency of m^1G37. In contrast, we use *trmD*-KO to eliminate m^1G37, so as to isolate suppressors that would expose differential codon-reading by various tRNA isoacceptors that would have been otherwise hidden by the presence of m^1G37. From the exposed differential codon-reading upon deletion of m^1G37, we gain insight into how the methylation resolves the bias. We show that, while m^1G37 is normally associated with decoding of Pro, Leu, and Arg codons in the genetic code, it is most critically associated with resolving codon bias against decoding of pyrimidine-ending codons of Pro, serving as an essential neutralizer that operates at the limit of cell survival. This discovery has important medical implication for antibacterial targeting of bacterial pathogens that are enriched with pyrimidine-ending codons of Pro and are also lacking the minor isoacceptor of Pro(GGG).

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE

- CONTACT FOR REAGENT AND RESOURCE SHARING

- METHOD DETAILS

  ○ Genome-wide screen of *trmD*-KO suppressors

  ○ Whole-genome sequencing and identification of variants

  ○ Structural mapping and amino-acid sequence alignment

  ○ CRISPR reconstruction of each *proS* suppressor

  ○ Cellular prolyl-aminoacylation status by acid urea gel analysis

  ○ Frameshift reporter assay

  ○ Competition fitness assay

  ○ Kinetic assay of prolyl-aminoacylation of tRNA^{Pro} *in vitro*

  ○ Evaluation of tRNA^{Pro} isoacceptors for cell viability

  ○ Codon usage analysis of CC[C/U] codons in *E. coli*

  ○ Temperature-sensitivity assay

  ○ Phylogenetic analysis of prolyl isoacceptors across bacterial species

- QUANTIFICATION AND STATISTICAL ANALYSES

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and can be found with this article online.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

I.M. constructed the *trmD*-KO strain and performed the genome-wide suppressor screen, Y.Y. and T.C. performed kinetic assays of prolyl-aminoacylation, R.D., S.T., R.M., and M.F.M. performed whole-genome sequencing analysis and variant calling, H.M. and T.L. performed phylogenetic analysis, S.M. generated data of acid-urea gel analysis, Y.N. performed codon usage analysis. I.M. prepared figures, legends, and methods, and Y.M.H. wrote the manuscript. All authors have given approval to the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing financial interests.

Figure legends

Figure 1. Whole-genome sequencing analysis revealed *proS* suppressor mutations of the *E. coli trmD*-KO strains.

(A) Synthesis of m^{1}G37 (red circle) by the bacterial TrmD enzyme on the 3'-side of the anticodon located at positions 34-36 (blue circles) of tRNA.

(B) Three *E. coli trmD*-KO strains (MG1655, BW25113, and XAC) for isolation of suppressors of m^{1}G37 deficiency. In each strain, the chromosomal gene was replaced with a Kan marker by P1 transduction and cell viability was supported by a plasmid-borne *trmD* under the control of the Ara-inducible pBAD promoter and a *temperature-sensitive* (*ts*) origin of replication. Culturing each strain in an Ara-free LB medium at 43 °C cured the plasmid and depleted the *trmD* enzyme and m^{1}G37-tRNA. After three cycles of growth and dilution at 43 °C, cells were streaked on an Ara-free LB plate and incubated at 43 °C until appearance of suppressor colonies, each of which was streak-purified, verified for the absence of the *ts* plasmid, and subjected to whole-genome sequencing.

(C) Exclusive mapping of suppressor mutations to *proS*, totaling in 9 unique mutations. Each unique mutation is listed for the *E. coli* strain(s) from which it was isolated, the nucleotide substitution in the chromosome, and the corresponding change of the amino acid.

(D) Three clusters of suppressor mutations superimposed on the crystal structure of the *Enterococcus faecalis proS* enzyme in complex with the prolyl adenylate analog 5'-O-(N-(L-prolyl)-sulfamoyl) adenosine (PDB: 2J3L). In the obligate dimer structure of the enzyme, mutations are mapped to residues shown in red spheres in monomer #1, while the prolyl analog is shown in white.

(E) A linear diagram of the structural domains of the *E. coli proS* enzyme (top) is shown with the 9 unique suppressor mutations by yellow triangles. A multi-sequence alignment of the enzyme

24

among bacterial species (bottom) displays residues harboring mutations in yellow. Amino acids boxed in blue are conserved, while those in blue letters share structural similarity.

Figure 2. Each *proS* suppressor mutant supports cell viability without *trmD.*

(A) Reconstruction of each *proS* suppressor mutant in MG1655. Each mutation was introduced to the chromosome by CRISPR-Cas9, followed by replacement of *trmD* with a Kan marker using P1 transduction, and by removal of the Kan marker using FLP recombination.

(B) Viability assay of each reconstructed *proS* suppressor mutant upon a serial dilution on an LB plate incubated at 37 °C. WT is represented by the native MG1655 strain, while each suppressor is indicated by the mutation in the *proS* gene. The symbols (+) and (–) indicate the presence and absence of the *trmD* gene on the chromosome.

(C) A cell fitness assay of a 1:1 mixture of the MG1655 WT strain (labeled with an mCherry-expressing plasmid) and the *proS*-D98N suppressor strain (labeled with a YFP-expressing plasmid) monitored over time at 37 °C in LB. Relative fraction of each strain at a given time point is based on the number of colonies of each as visualized by the fluorescence color of mCherry or YFP.

(D) A +1-frameshifting assay measuring expression of a plasmid-borne nLuc reporter in the MG1655 WT strain or a reconstructed *proS* suppressor strain. Each strain expresses the reporter harboring an in-frame CCC triplet or a +1-frame CCC-C quadruplet next to the initiation codon AUG. The frequency of +1 frameshifting is the ratio of the reporter expression harboring the CCC-C quadruplet relative to the CCC triplet, each normalized by an internal expression of mCherry. Data are represented as mean ± SD ($n = 4$). **, $p < 0.05$; ***, $p < 0.01$.

Figure 3. *proS* suppressor mutants are less dependent on m^{1}G37 for prolyl-aminoacylation of tRNA.

(A) Acid urea gel analysis of the prolyl-aminoacylation status of each reconstructed *proS* suppressor strain. Total RNA extracted in an acidic condition from each strain in early-log phase was run on a denaturing 12% PAGE/7M urea gel at pH 5.0 and probed for the Pro(UGG) tRNA by Northern blots. A pair of control tRNA samples isolated from $m^1G37(+)$ and m^1G37 (–) conditions yielded a prolyl-aminoacylation status of 58% and 6%, respectively, whereas all reconstructed *proS* suppressors showed a status significantly higher than the negative control, ranging in 38-63%, indicating the ability to perform prolyl-aminoacylation in the absence of m^1G37.

(B) Kinetics of prolyl-aminoacylation on the isoacceptor Pro(UGG), comparing the loss of k_{cat}/K_m of the WT enzyme upon removal of m^1G37 (from the m^1G37-state to the G37-state) relative to three of the reconstructed suppressor mutants *proS*-D98N, T199I, and N232S.

(C) Kinetics of prolyl-aminoacylation on the isoacceptor Pro(GGG), comparing the loss of k_{cat}/K_m of the WT enzyme relative to the suppressor mutant *proS*-D98N.

(D) Kinetics of prolyl-aminoacylation on the isoacceptor Pro(CGG), comparing the loss of k_{cat}/K_m of the WT enzyme relative to the suppressor mutant *proS*-D98N.

All data are represented as mean ± SD. (*n* = 3). See Figure S3 for detailed kinetic parameters.

Figure 4. *proS* suppressor mutants require the Pro(GGG) isoacceptor for survival.

(A) The three prolyl isoacceptors in *E. coli*, consisting of Pro(CGG) encoded by *proK* and responsible for reading the CCG codon; Pro(GGG) encoded by *proL* and responsible for reading CC[C/U] codons, and Pro(UGG) encoded by *proM* and capable of reading CCN codons due to the cmo^5U34 modification.

(B) Construction of a *trmD*-KD strain with a *proS* suppressor mutation. A *proS* mutation was reconstructed on the chromosome of MG1655, followed by removal of *trmD*, while cell viability

was maintained by a plasmid-borne human *trm5* under the control of the Ara-inducible promoter pBAD.

(C) Depletion analysis of Pro(GGG) or Pro(CGG) in the *trmD*-KO strain. While the strain maintained by the WT *proS* is viable in the m^{1}G37+ condition, it is not viable in the m^{1}G37– condition. In contrast, the strain maintained by the *proS*-D98N suppressor is viable in the m^{1}G37– condition, but only when the Pro(GGG) isoacceptor is present.

(D) Isoacceptor complementation analysis in the *trmD*-KO strain. The strain maintained by the *proS*-D98N mutation is complemented by over-expression (O/E) of the Pro(GGG) tRNA (Left), but not by over-expression of the Pro(GGG) tRNA where the anticodon is mutated to UGG, or by over-expression of the Met(CAU) tRNA. Each tRNA is expressed from the pKK223-3 plasmid.

Figure 5. Expression of *E. coli csp* genes is sensitive to the cellular level of Pro(GGG).

(A) A violin plot of codon usage frequency of CC[C/U] in *E. coli* protein-coding genes. About 20.4% of total protein-coding genes have a higher codon usage frequency than the average, which are used for Gene Ontology (GO) analysis.

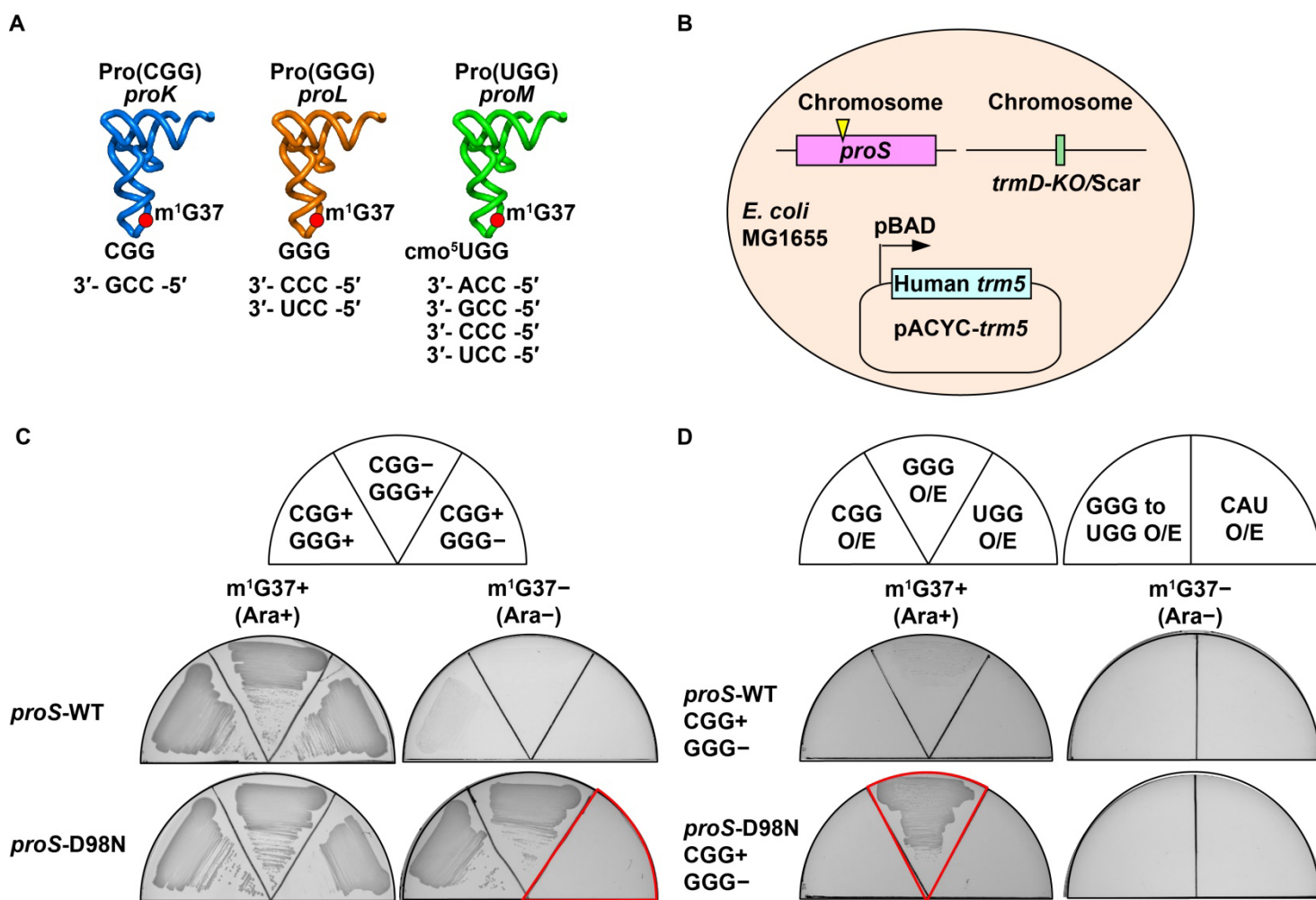**(B)** Identification of four groups of *E. coli* protein-coding genes enriched with CC[C/U] codons. Of the 16 genes categorized in the top group, 12 (in red) are enriched with CC{C/U} codons, while 8 of the 12 are of the *csp* family.

(C) Cold sensitivity assay. The MG1655 WT (+) strain and the reconstructed *proS*-D98N suppressor strain (–) (Figure 2A) were serially diluted and tested for viability at 37 °C (one overnight) or at 22 °C (two overnights).

(D) Rescue of cold sensitivity by expression of tRNA isoacceptors for the reconstructed *proS*-D98N mutant. Each isoacceptor was expressed from the pKK223-3 plasmid in the WT (+) strain or the *proS*-D98N strain.

Figure 6. Absence of Pro(GGG) from some bacterial species.

(A) A maximum-likelihood phylogenetic tree showing the presence and absence of prolyl tRNA genes in genomes of representative bacterial genera that are available at GtRNAdb. This is generated with the concatenated 16S and 23S alignments with higher classifications as taxonomic groups shown in various colors. Gram-positive (Gram (+)) genera are shown in bold face, and those of human pathogens are indicated by a red circle. Inner rings represent the presence and absence of prolyl isoacceptor genes: the blue ring for Pro(CGG), the green ring for Pro(GGG), and the yellow ring for Pro(UGG). The gradient of color darkness indicates the frequency of the species lacking the corresponding isoacceptor gene within the genus. Darker color means more species lacking the isoacceptor gene, with the complete black indicating the complete absence.

(B) A Venn diagram depicting the distribution of medically relevant (pathogenic) genera with and without Pro(GGG) within Gram-positive genera. Among 171 genera, 51 genera contain pathogenic species (orange), 55 genera lack Pro(GGG (blue), and 23 are in common, indicating pathogenic species that lack Pro(GGG) (blue).

(C) A Venn diagram depicting the distribution of medically relevant (pathogenic) genera with and without Pro(GGG) within Gram-negative genera. Among 424 genera, 93 genera contain pathogenic species (orange), 57 genera lack Pro(GGG) (blue), and 25 are in common, indicating pathogenic species that lack Pro(GGG).
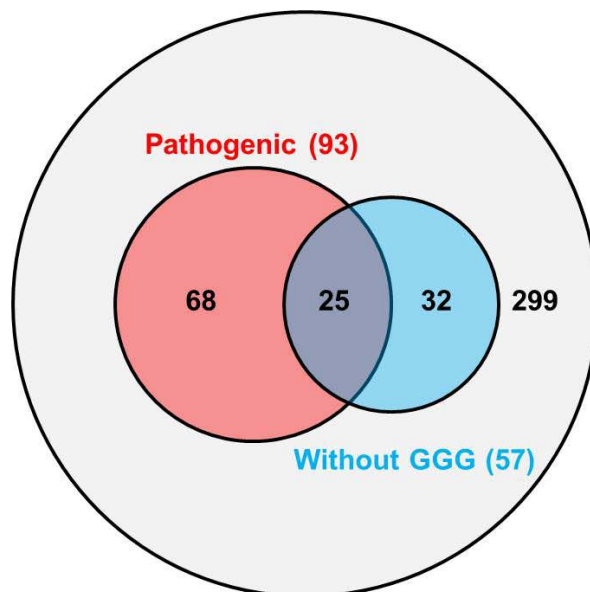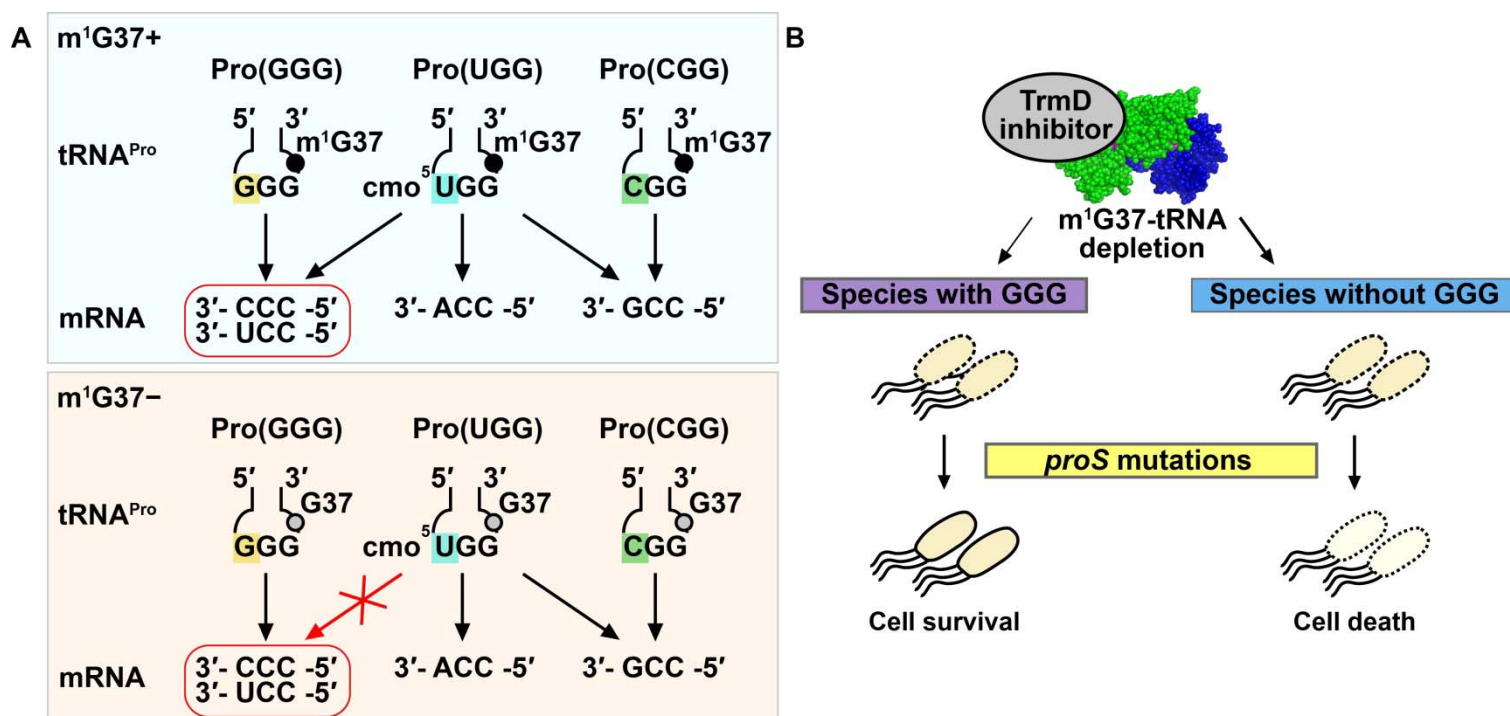
Figure 7. tRNA methylation with m^{1}G37 resolves codon bias of the cmo^{5}U34-modified major isoacceptor Pro(UGG) with medical relevance.

(A) m^{1}G37 resolves bias of the major prolyl isoacceptor. (Top) In the presence of m^{1}G37 (a black circle), the major isoacceptor Pro(UGG), modified with cmo^{5}U34, is capable of reading all four Pro codons, including using wobble base pairings to read CC[C/U]. (Bottom) In the absence of m^{1}G37– (an open circle), the cmo^{5}U34-modified major isoacceptor Pro(UGG) is no longer capable of reading CC[C/U], thus requiring the minor isoacceptor Pro(GGG) to assume the role and rendering the latter essential for cell viability.

(B) Risk of resistance of targeting TrmD. As a high-priority antibacterial target, TrmD is being actively pursued for discovery of new antibiotics. Upon targeting TrmD by an inhibitor, while cell viability is diminished, a potential resistance mechanism could occur by suppressor mutations in *proS* as shown here. However, this resistance mechanism would be thwarted if the bacterial species lacks the Pro(GGG) isoacceptor, which becomes essential for viability in the absence of TrmD. Attractive candidates for this anti-TrmD strategy would include the Gram-negative *Acinetobacter baumannii*, and the Gram-positive *Enterococcus faecalis* and *Staphylococcus aureus*.

29

Figure 1. Whole-genome sequencing analysis revealed *proS* suppressor mutations of the *E. coli* *trmD*-KO strains.

Figure 2. Each *proS* suppressor mutant supports cell viability without *trmD*.
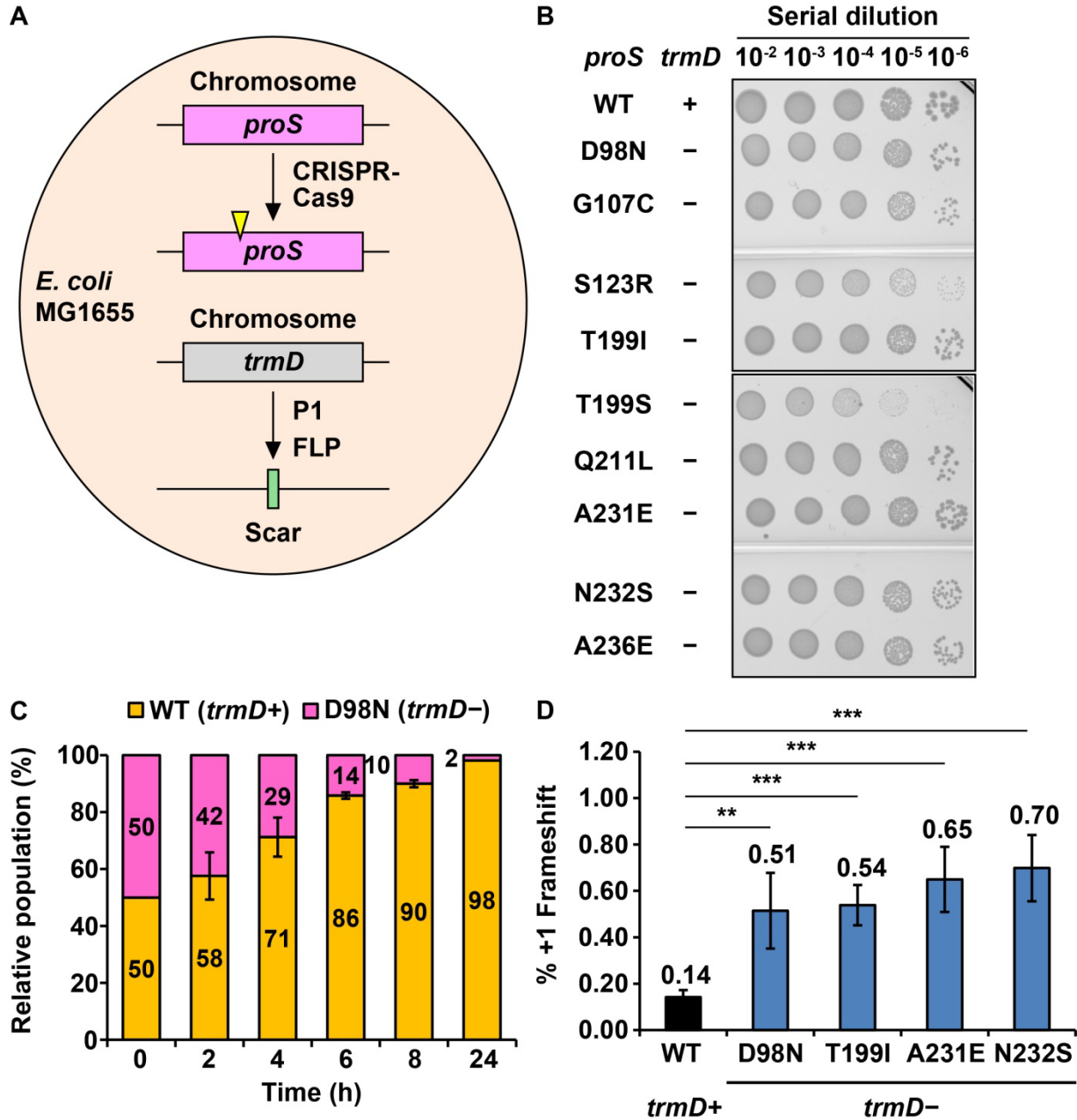
Figure 3. *proS* suppressor mutants are less dependent on m^{1}G37 for prolyl-aminoacylation of tRNA.
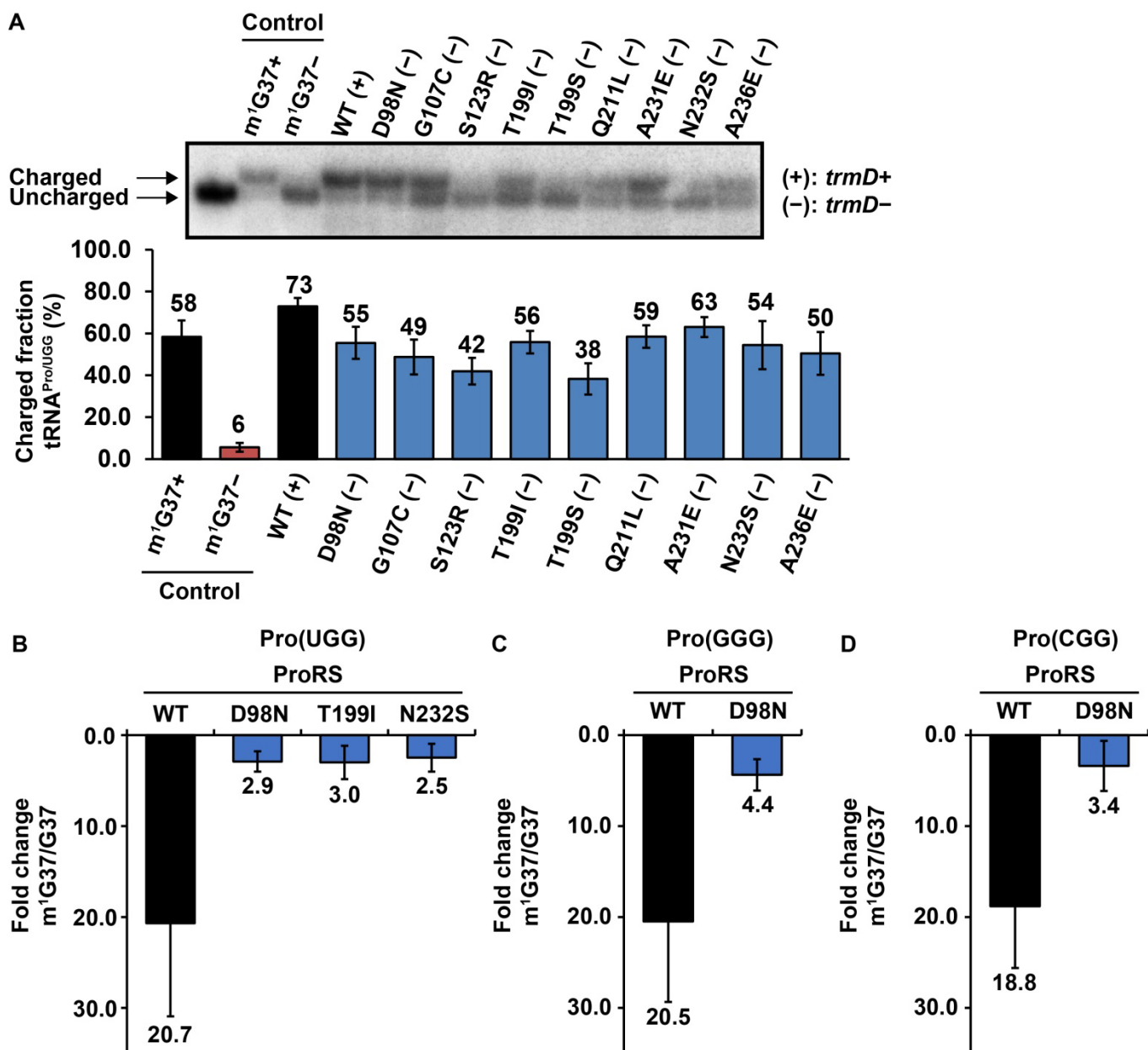
Figure 4. *proS* suppressor mutants require the Pro(GGG) isoacceptor for survival.

Figure 5. Expression of *E. coli csp* genes is sensitive to the cellular level of Pro(GGG).

A — Violin plot of Frequency of CC[C/U] relative to all Pro codon per gene vs. *E. coli* gene number. Enriched fraction (≥0.5): 20.4% of total genes. Average: 0.31, Median: 0.29.

B — GO analysis (Biological process), Fold-enrichment:
- Negative regulation of cellular component organization: 4.2, $P = 4.0 \times 10^{-4}$
- Transposition, DNA-mediated: 2.8, $P = 6.8 \times 10^{-5}$
- Regulation of transcription, DNA-templated: 1.6, $P = 3.3 \times 10^{-5}$
- Unclassified: 1.4, $P = 2.1 \times 10^{-8}$

Gene list: *cbtA cspI, cspA dgcZ, cspB fliT, cspC quuD, cspE slmA, cspF sulA, cspG ybjN, cspH yfbV*

C — Spot assays at 37°C (1 overnight) and 22°C (2 overnights), dilutions 10^{-3} 10^{-4} 10^{-5} 10^{-6} 10^{-7}:

proS	trmD
WT	+
D98N	−
T199I	−
A231E	−
N232S	−

D — Spot assays at 37°C (1 overnight) and 22°C (2 overnights), dilutions 10^{-3} 10^{-4} 10^{-5} 10^{-6}:

proS	trmD	Plasmid
WT	+	+Empty
	+	+CGG
	+	+GGG
	+	+UGG
D98N	−	+Empty
	−	+CGG
	−	+GGG
	−	+UGG

Figure 6. Absence of Pro(GGG) from some bacterial species.

Figure 7. tRNA methylation with m^{1}G37 resolves codon bias of the cmo^{5}U34-modified major isoacceptor Pro(UGG) with medical relevance.

References:

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 Biol *15*, 537-553. 10.1080/15476286.2017.1356562.

Ahn, H.J., Kim, H.W., Yoon, H.J., Lee, B.I., Suh, S.W., and Yang, J.K. (2003). Crystal structure of tRNA(m1G37)methyltransferase: insights into tRNA recognition. EMBO J *22*, 2593-2603. 10.1093/emboj/cdg269.

Begley, U., Dyavaiah, M., Patil, A., Rooney, J.P., DiRenzo, D., Young, C.M., Conklin, D.S., Zitomer, R.S., and Begley, T.J. (2007). Trm9-catalyzed tRNA modifications link translation to the DNA damage response. Mol Cell *28*, 860-870. 10.1016/j.molcel.2007.09.021.

Beuning, P.J., and Musier-Forsyth, K. (2000). Hydrolytic editing by a class II aminoacyl-tRNA synthetase. Proc Natl Acad Sci U S A *97*, 8916-8920. 10.1073/pnas.97.16.8916.

Bjork, G.R., and Hagervall, T.G. (2014). Transfer RNA Modification: Presence, Synthesis, and Function. EcoSal Plus *6*. 10.1128/ecosalplus.ESP-0007-2013.

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.

Bystrom, A.S., and Bjork, G.R. (1982). The structural gene (trmD) for the tRNA(m1G)methyltransferase is part of a four polypeptide operon in Escherichia coli K-12. Mol Gen Genet *188*, 447-454.

Chan, P.P., and Lowe, T.M. (2016). GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. Nucleic Acids Res *44*, D184-189. 10.1093/nar/gkv1309.

Chionh, Y.H., McBee, M., Babu, I.R., Hia, F., Lin, W., Zhao, W., Cao, J., Dziergowska, A., Malkiewicz, A., Begley, T.J., et al. (2016). tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence. Nat Commun *7*, 13302. 10.1038/ncomms13302.

Christian, T., Evilia, C., Williams, S., and Hou, Y.M. (2004). Distinct origins of tRNA(m1G37) methyltransferase. J Mol Biol *339*, 707-719. 10.1016/j.jmb.2004.04.025.

Christian, T., Gamper, H., and Hou, Y.M. (2013). Conservation of structure and mechanism by Trm5 enzymes. RNA *19*, 1192-1199. 10.1261/rna.039503.113.

Christian, T., and Hou, Y.M. (2007). Distinct determinants of tRNA recognition by the TrmD and Trm5 methyl transferases. J Mol Biol *373*, 623-632. 10.1016/j.jmb.2007.08.010.

Christian, T., Lahoud, G., Liu, C., and Hou, Y.M. (2010). Control of catalytic cycle by a pair of analogous tRNA modification enzymes. J Mol Biol *400*, 204-217. 10.1016/j.jmb.2010.05.003.

Christian, T., Sakaguchi, R., Perlinska, A.P., Lahoud, G., Ito, T., Taylor, E.A., Yokoyama, S., Sulkowska, J.I., and Hou, Y.M. (2016). Methyl transfer by substrate signaling from a knotted protein fold. Nat Struct Mol Biol *23*, 941-948. 10.1038/nsmb.3282.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80-92. 10.4161/fly.19695.

Clifton, B.E., Fariz, M.A., Uechi, G.I., and Laurino, P. (2021). Evolutionary repair reveals an unexpected role of the tRNA modification m1G37 in aminoacylation. Nucleic Acids Res *49*, 12467-12485. 10.1093/nar/gkab1067.

Crepin, T., Yaremchuk, A., Tukalo, M., and Cusack, S. (2006). Structures of two bacterial prolyl-tRNA synthetases with and without a cis-editing domain. Structure *14*, 1511-1525. 10.1016/j.str.2006.08.007.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A *97*, 6640-6645. 10.1073/pnas.120163297.

de Crecy-Lagard, V., and Jaroch, M. (2021). Functions of Bacterial tRNA Modifications: From Ubiquity to Diversity. Trends Microbiol *29*, 41-53. 10.1016/j.tim.2020.06.010.

De Oliveira, D.M.P., Forde, B.M., Kidd, T.J., Harris, P.N.A., Schembri, M.A., Beatson, S.A., Paterson, D.L., and Walker, M.J. (2020). Antimicrobial Resistance in ESKAPE Pathogens. Clin Microbiol Rev *33*. 10.1128/CMR.00181-19.

Dincbas, V., Heurgue-Hamard, V., Buckingham, R.H., Karimi, R., and Ehrenberg, M. (1999). Shutdown in protein synthesis due to the expression of mini-genes in bacteria. J Mol Biol *291*, 745-759. 10.1006/jmbi.1999.3028.

dos Reis, M., Savva, R., and Wernisch, L. (2004). Solving the riddle of codon usage preferences: a test for translational selection. Nucleic Acids Res *32*, 5036-5044. 10.1093/nar/gkh834.

Duret, L. (2000). tRNA gene number and codon usage in the C. elegans genome are co-adapted for optimal translation of highly expressed genes. Trends Genet *16*, 287-289. 10.1016/s0168-9525(00)02041-2.

Elkins, P.A., Watts, J.M., Zalacain, M., van Thiel, A., Vitazka, P.R., Redlak, M., Andraos-Selim, C., Rastinejad, F., and Holmes, W.M. (2003). Insights into catalysis by a knotted TrmD tRNA methyltransferase. J Mol Biol *333*, 931-949.

Gamper, H., Li, H., Masuda, I., Miklos Robkis, D., Christian, T., Conn, A.B., Blaha, G., Petersson, E.J., Gonzalez, R.L., Jr., and Hou, Y.M. (2021a). Insights into genome recoding from the mechanism of a classic +1-frameshifting tRNA. Nat Commun *12*, 328. 10.1038/s41467-020-20373-z.

Gamper, H., Mao, Y., Masuda, I., McGuigan, H., Blaha, G., Wang, Y., Xu, S., and Hou, Y.M. (2021b). Twice exploration of tRNA +1 frameshifting in an elongation cycle of protein synthesis. Nucleic Acids Res *49*, 10046-10060. 10.1093/nar/gkab734.

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. Nat Commun *6*, 7226. 10.1038/ncomms8226.

Goto-Ito, S., Ito, T., Ishii, R., Muto, Y., Bessho, Y., and Yokoyama, S. (2008). Crystal structure of archaeal tRNA(m(1)G37)methyltransferase aTrm5. Proteins *72*, 1274-1289. 10.1002/prot.22019.

Goto-Ito, S., Ito, T., Kuratani, M., Bessho, Y., and Yokoyama, S. (2009). Tertiary structure checkpoint at anticodon loop modification in tRNA functional maturation. Nat Struct Mol Biol *16*, 1109-1115. 10.1038/nsmb.1653.

Grenier, F., Matteau, D., Baby, V., and Rodrigue, S. (2014). Complete Genome Sequence of Escherichia coli BW25113. Genome Announc *2*. 10.1128/genomeA.01038-14.

Han, L., and Phizicky, E.M. (2018). A rationale for tRNA modification circuits in the anticodon loop. RNA *24*, 1277-1284. 10.1261/rna.067736.118.

Hanson, G., and Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. Nat Rev Mol Cell Biol *19*, 20-30. 10.1038/nrm.2017.91.

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B.L., Mori, H., and Horiuchi, T. (2006). Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110. Mol Syst Biol *2*, 2006 0007. 10.1038/msb4100049.

Hill, P.J., Abibi, A., Albert, R., Andrews, B., Gagnon, M.M., Gao, N., Grebe, T., Hajec, L.I., Huang, J., Livchak, S., et al. (2013). Selective inhibitors of bacterial t-RNA-(N(1)G37) methyltransferase (TrmD) that demonstrate novel ordering of the lid domain. J Med Chem *56*, 7278-7288. 10.1021/jm400718n.

Hoffer, E.D., Hong, S., Sunita, S., Maehigashi, T., Gonzalez, R.L.J., Whitford, P.C., and Dunham, C.M. (2020). Structural insights into mRNA reading frame regulation by tRNA modification and slippery codon-anticodon pairing. Elife *9*. 10.7554/eLife.51898.

Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Mol Biol Evol *33*, 1635-1638. 10.1093/molbev/msw046. Ikemura, T. (1981). Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the E. coli translational system. J Mol Biol *151*, 389-409. 10.1016/0022-2836(81)90003-6.

Jakowec, M., Prentki, P., Chandler, M., and Galas, D.J. (1988). Mutational analysis of the open reading frames in the transposable element IS1. Genetics *120*, 47-55. 10.1093/genetics/120.1.47.

Kelsic, E.D., Zhao, J., Vetsigian, K., and Kishony, R. (2015). Counteraction of antibiotic production and degradation stabilizes microbial communities. Nature *521*, 516-519. 10.1038/nature14485.

Keto-Timonen, R., Hietala, N., Palonen, E., Hakakorpi, A., Lindstrom, M., and Korkeala, H. (2016). Cold Shock Proteins: A Minireview with Special Emphasis on Csp-family of Enteropathogenic Yersinia. Front Microbiol *7*, 1151. 10.3389/fmicb.2016.01151.

Lamichhane, T.N., Arimbasseri, A.G., Rijal, K., Iben, J.R., Wei, F.Y., Tomizawa, K., and Maraia, R.J. (2016). Lack of tRNA-i6A modification causes mitochondrial-like metabolic deficiency in S. pombe by limiting activity of cytosolic tRNATyr, not mito-tRNA. RNA *22*, 583-596. 10.1261/rna.054064.115.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948. 10.1093/bioinformatics/btm404.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics *27*, 2987-2993. 10.1093/bioinformatics/btr509.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760. 10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.

Li, J.N., and Bjork, G.R. (1999). Structural alterations of the tRNA(m1G37)methyltransferase from Salmonella typhimurium affect tRNA substrate specificity. RNA *5*, 395-408.

Lyu, X., Yang, Q., Li, L., Dang, Y., Zhou, Z., Chen, S., and Liu, Y. (2020). Adaptation of codon usage to tRNA I34 modification controls translation kinetics and proteome landscape. PLoS Genet *16*, e1008836. 10.1371/journal.pgen.1008836.

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. 10.1073/pnas.1409436111.

Maraia, R.J., and Arimbasseri, A.G. (2017). Factors That Shape Eukaryotic tRNAomes: Processing, Modification and Anticodon-Codon Use. Biomolecules *7*. 10.3390/biom7010026.

Masuda, I., Hwang, J.Y., Christian, T., Maharjan, S., Mohammad, F., Gamper, H., Buskirk, A.R., and Hou, Y.M. (2021). Loss of N(1)-methylation of G37 in tRNA induces ribosome stalling and reprograms gene expression. Elife *10*. 10.7554/eLife.70619.

Masuda, I., Matsubara, R., Christian, T., Rojas, E.R., Yadavalli, S.S., Zhang, L., Goulian, M., Foster, L.J., Huang, K.C., and Hou, Y.M. (2019). tRNA Methylation Is a Global Determinant of Bacterial Multi-drug Resistance. Cell Syst *8*, 302-314 e308. 10.1016/j.cels.2019.03.008.

Masuda, I., Takase, R., Matsubara, R., Paulines, M.J., Gamper, H., Limbach, P.A., and Hou, Y.M. (2018). Selective terminal methylation of a tRNA wobble base. Nucleic Acids Res *46*, e37. 10.1093/nar/gky013.

Moriyama, E.N., and Powell, J.R. (1997). Codon usage bias and tRNA abundance in Drosophila. J Mol Evol *45*, 514-523. 10.1007/pl00006256.

Nasvall, S.J., Chen, P., and Bjork, G.R. (2004). The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons in vivo. RNA *10*, 1662-1673. 10.1261/rna.7106404.

Nasvall, S.J., Chen, P., and Bjork, G.R. (2007). The wobble hypothesis revisited: uridine-5-oxyacetic acid is critical for reading of G-ending codons. RNA *13*, 2151-2164. 10.1261/rna.731007.

Nedialkova, D.D., and Leidel, S.A. (2015). Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. Cell *161*, 1606-1618. 10.1016/j.cell.2015.05.022.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res *44*, D733-745. 10.1093/nar/gkv1189.

Pan, T. (2018). Modifications and functional genomics of human transfer RNA. Cell Res *28*, 395-404. 10.1038/s41422-018-0013-y.

Perret, V., Garcia, A., Grosjean, H., Ebel, J.P., Florentz, C., and Giege, R. (1990). Relaxation of a transfer RNA specificity by removal of modified nucleotides. Nature *344*, 787-789. 10.1038/344787a0.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol *26*, 1641-1650. 10.1093/molbev/msp077.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One *5*, e9490. 10.1371/journal.pone.0009490.

Pruesse, E., Peplies, J., and Glockner, F.O. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics *28*, 1823-1829. 10.1093/bioinformatics/bts252.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glockner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res *41*, D590-596. 10.1093/nar/gks1219.

Rak, R., Dahan, O., and Pilpel, Y. (2018). Repertoires of tRNAs: The Couplers of Genomics and Proteomics. Annu Rev Cell Dev Biol *34*, 239-264. 10.1146/annurev-cellbio-100617-062754.

Ramos, J., Proven, M., Halvardson, J., Hagelskamp, F., Kuchinskaya, E., Phelan, B., Bell, R., Kellner, S.M., Feuk, L., Thuresson, A.C., and Fu, D. (2020). Identification and rescue of a tRNA wobble inosine deficiency causing intellectual disability disorder. RNA *26*, 1654-1666. 10.1261/rna.076380.120.

Razin, S., Yogev, D., and Naot, Y. (1998). Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev *62*, 1094-1156. 10.1128/MMBR.62.4.1094-1156.1998.

Reisch, C.R., and Prather, K.L. (2015). The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli. Sci Rep *5*, 15096. 10.1038/srep15096.

Reisch, C.R., and Prather, K.L.J. (2017). Scarless Cas9 Assisted Recombineering (no-SCAR) in Escherichia coli, an Easy-to-Use System for Genome Editing. Curr Protoc Mol Biol *117*, 31 38 31-31 38 20. 10.1002/cpmb.29.

Sabi, R., and Tuller, T. (2014). Modelling the efficiency of codon-tRNA interactions based on codon usage bias. DNA Res *21*, 511-526. 10.1093/dnares/dsu017.

Sakaguchi, R., Lahoud, G., Christian, T., Gamper, H., and Hou, Y.M. (2014). A divalent metal ion-dependent N(1)-methyl transfer to G37-tRNA. Chem Biol *21*, 1351-1360. 10.1016/j.chembiol.2014.07.023.

Sayers, E.W., Cavanaugh, M., Clark, K., Pruitt, K.D., Schoch, C.L., Sherry, S.T., and Karsch-Mizrachi, I. (2022). GenBank. Nucleic Acids Res *50*, D161-D164. 10.1093/nar/gkab1135.

Thiaville, P.C., Legendre, R., Rojas-Benitez, D., Baudin-Baillieu, A., Hatin, I., Chalancon, G., Glavic, A., Namy, O., and de Crecy-Lagard, V. (2016). Global translational impacts of the loss of the tRNA modification t(6)A in yeast. Microb Cell *3*, 29-45. 10.15698/mic2016.01.473.

Thomas, S.E., Whitehouse, A.J., Brown, K., Burbaud, S., Belardinelli, J.M., Sangen, J., Lahiri, R., Libardo, M.D.J., Gupta, P., Malhotra, S., et al. (2020). Fragment-based discovery of a new class of inhibitors targeting mycobacterial tRNA modification. Nucleic Acids Res *48*, 8099-8112. 10.1093/nar/gkaa539.

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

Vargas-Rodriguez, O., and Musier-Forsyth, K. (2013). Exclusive use of trans-editing domains prevents proline mistranslation. J Biol Chem *288*, 14391-14399. 10.1074/jbc.M113.467795.

Wei, Y., Silke, J.R., and Xia, X. (2019). An improved estimation of tRNA expression to better elucidate the coevolution between tRNA abundance and codon usage in bacteria. Sci Rep *9*, 3184. 10.1038/s41598-019-39369-x.

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. 10.1038/nsmb1242.

42

White, T.A., and Kell, D.B. (2004). Comparative genomic assessment of novel broad-spectrum targets for antibacterial drugs. Comp Funct Genomics *5*, 304-327. 10.1002/cfg.411.

Whitehouse, A.J., Thomas, S.E., Brown, K.P., Fanourakis, A., Chan, D.S., Libardo, M.D.J., Mendes, V., Boshoff, H.I.M., Floto, R.A., Abell, C., et al. (2019). Development of Inhibitors against Mycobacterium abscessus tRNA (m(1)G37) Methyltransferase (TrmD) Using Fragment-Based Approaches. J Med Chem *62*, 7210-7232. 10.1021/acs.jmedchem.9b00809.

Yaremchuk, A., Cusack, S., and Tukalo, M. (2000). Crystal structure of a eukaryote/archaeon-like protyl-tRNA synthetase and its complex with tRNAPro(CGG). EMBO J *19*, 4745-4758. 10.1093/emboj/19.17.4745.

Yaremchuk, A., Tukalo, M., Grotli, M., and Cusack, S. (2001). A succession of substrate induced conformational changes ensures the amino acid specificity of Thermus thermophilus prolyl-tRNA synthetase: comparison with histidyl-tRNA synthetase. J Mol Biol *309*, 989-1002. 10.1006/jmbi.2001.4712.

Zhong, W., Koay, A., Ngo, A., Li, Y., Nah, Q., Wong, Y.H., Chionh, Y.H., Ng, H.Q., Koh-Stenta, X., Poulsen, A., et al. (2019a). Targeting the Bacterial Epitranscriptome for Antibiotic Development: Discovery of Novel tRNA-(N(1)G37) Methyltransferase (TrmD) Inhibitors. ACS Infect Dis *5*, 326-335. 10.1021/acsinfecdis.8b00275.

Zhong, W., Pasunooti, K.K., Balamkundu, S., Wong, Y.H., Nah, Q., Gadi, V., Gnanakalai, S., Chionh, Y.H., McBee, M.E., Gopal, P., et al. (2019b). Thienopyrimidinone Derivatives That Inhibit Bacterial tRNA (Guanine37-N(1))-Methyltransferase (TrmD) by Restructuring the Active Site with a Tyrosine-Flipping Mechanism. J Med Chem *62*, 7788-7805. 10.1021/acs.jmedchem.9b00582.

Zinshteyn, B., and Gilbert, W.V. (2013). Loss of a conserved tRNA anticodon modification perturbs cellular signaling. PLoS Genet *9*, e1003675. 10.1371/journal.pgen.1003675.

Supplemental Information

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CysRS antibodies	(Masuda *et al.*, 2021)	N/A
Rabbit polyclonal anti-TrmD antibodies	(Li and Bjork, 1999)	N/A
Goat polyclonal anti-rabbit IgG antibodies peroxidase conjugate	Sigma-Aldrich	Cat. #A0545
Bacterial Strains		
Escherichia coli strain K-12 substrain MG1655	ATCC	Cat. #700926
E. coli strain BW25113	The Coli Genetic Stock Center (CGSC)	CGSC#: 7636
E. coli strain XAC	(Jakowec *et al.*, 1988)	N/A
E. coli strain BL21(DE3)	Millipore	Cat. #69450
E. coli strain ArcticExpress (DE3)	Agilent Technologies	Cat. #230191
E. coli strain K-12 substrain MG1655, *proS* D98N (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* G107C (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* S123R (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* T199I (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* T199S (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* Q211L (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* A231E (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* N232S (*trmD+* or *trmD–*)	This paper	N/A
E. coli strain K-12 substrain MG1655, *proS* A236E (*trmD+* or *trmD–*)	This paper	N/A
Bacteriophage P1vir	The Coli Genetic Stock Center (CGSC)	CGSC#: 12133
Chemicals, Peptides, and Recombinant Proteins		
L-Arabinose	Gold Biotechnology	Cat. #A-300-250
BamHI	New England BioLabs	Cat. #R0136
EcoRI	New England BioLabs	Cat. #R0101
NdeI	New England BioLabs	Cat. #R0111
PstI	New England BioLabs	Cat. #R0140
XhoI	New England BioLabs	Cat. #R0146
PfuUltraII fusion HS DNA polymerase	Agilent Technologies	Cat. #600670
Phusion™ High-Fidelity DNA Polymerase	ThermoFisher Scientific	Cat. #F530
NEBuilder® HiFi DNA Assembly Cloning Kit	New England BioLabs	Cat. #E5520S
T4 DNA ligase	New England Biolabs	Cat. #M0202
T4 polynucleotide kinase	New England Biolabs	Cat. #M0201

Ampicillin	Fisher Scientific	Cat. #BP1760
Chloramphenicol	Gold Biotechnology	Cat. #G-105
Kanamycin	Gemini Bio-products	Cat. #400-114P
Anhydrous tetracycline hydrochloride	Cayman Chemicals	Cat. #10009542
$[\gamma\text{-}^{32}\text{P}]$-ATP	PerkinElmer	Cat. #NEG002A
$[^{3}\text{H}]$-Proline	PerkinElmer	Cat. #NET483005MC
Critical Commercial Assays		
SuperSignal West Pico Chemiluminescent Substrate	ThermoFisher Scientific	Cat. #34080
ONE-Glo™ Luciferase Assay System	Promega	Cat. #E6120
Experimental Models: Organisms/Strains		
E. coli strain K-12 substrain MG1655	ATCC	Cat. #700926
Oligonucleotides		
Oligo DNAs for bacterial strain construction, plasmid construction, CRISPR mutagenesis, and Northern blotting	Table S1	N/A
Recombinant DNA		
pKD46	CGSC	CGSC #7634
pKD46-*EctrmD*	This paper	N/A
pCP20	CGSC	CGSC #7629
pACYC-*araC*-P_C-P_{BAD}-human *trm5*	(Gamper *et al.*, 2015)	N/A
pCas9cr4	(Reisch and Prather, 2015)	N/A
pKDsgRNA	(Reisch and Prather, 2015)	N/A
pZS2R-mCherry	(Kelsic *et al.*, 2015)	N/A
pZS2R-YFP	(Kelsic *et al.*, 2015)	N/A
pKK223-3 *E. coli* Pro(CGG) (*proK*)	This paper	N/A
pKK223-3 *E. coli* Pro(GGG) (*proL*)	(Gamper *et al.*, 2015)	N/A
pKK223-3 *E. coli* Pro(UGG) (*proM*)	(Masuda *et al.*, 2021)	N/A
pKK223-3 *E. coli* Pro(GGG) (*proL*) with anticodon UGG	(Masuda *et al.*, 2018)	N/A
pKK223-3 *E. col* Met(CAU) (*metV*)	(Gamper *et al.*, 2015)	N/A
Software and Algorithms		
FASTQC v 0.11.5	Andrews S. (2010)	http://www.bioinformatics.babraham.ac.uk/projects/fastqc
Burrows-Wheeler Alignment (BWA) tool v 0.7.15 - r1140	(Li and Durbin, 2009)	N/A
SAMtools v 1.3.1	(Li *et al.*, 2009)	N/A
bcftools	(Li, 2011)	https://samtools.github.io/bcftools/bcftools.html
SnpEff	(Cingolani *et al.*, 2012)	N/A
Image Lab v. 6.0	BIO-RAD	http://www.bio-rad.com/en-us/product/image-lab-software
PyMOL v. 2.4.1	Schrödinger	https://pymol.org/
Clustal X v. 2.1	(Larkin *et al.*, 2007)	N/A
Kaleidagraph v. 4.5	Synergy software	https://www.synergy.com/
ImageJ v. 1.51	National Institutes of Health	https://imagej.nih.gov

Python v. 3.7.7	Python Software Foundation	https://www.python.org/
ETE3 v. 3.1.2	(Huerta-Cepas *et al.*, 2016)	http://etetoolkit.org/
RefSeq release 211	(O'Leary *et al.*, 2016)	https://www.ncbi.nlm.nih.gov/refseq/
GenBank release 248	(Sayers *et al.*, 2022)	https://www.ncbi.nlm.nih.gov/genbank/
FastTree v. 2.1.10	(Price *et al.*, 2009; 2010)	http://www.microbesonline.org/fasttree/
GtRNAdb release 19	(Chan and Lowe, 2016)	http://gtrnadb.ucsc.edu/
matplotlib v. 3.5.0	10.1109/MCSE.2007.55	https://matplotlib.org/3.1.1/index.html
SINA v. 1.2.11	(Pruesse *et al.*, 2012)	https://www.arb-silva.de/aligner/
SILVA release 89	(Quast *et al.*, 2013)	https://www.arb-silva.de/
Other		
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel	Cat. #740609
Hybond^TM-N+ Nitrocellulose Membrane	GE healthcare	Cat. #RPN203B
Nextera XT DNA Library Preparation Kit	Illumina	Cat. #FC-131-1024
Trans-Blot® TurboTM Transfer System	BIO-RAD	Cat. #1704150
MicroPulser Electroporator	BIO-RAD	Cat. #1652100
UV Crosslinker	Fisher Scientific	Cat. #FB-UVXL-1000
Tri-Carb 4910 TR Scintillation Counter	PerkinElmer	N/A
Immobilon-P PVDF Membrane	Millipore	IPVH00010
ChemiDoc MP Imaging System	BIO-RAD	Cat. #12003154
Typhoon IP Imaging System	GE Healthcare	N/A
Infinite M200 PRO Plate Reader	Tecan	N/A

Method Details

Genome-wide screen of *trmD-KO* suppressors

Because *trmD* is essential for viability, a simple KO cannot be made. A maintenance plasmid carrying *E. coli trmD* was made with a *ts* origin of replication. The gene together with a partial sequence of P_{BAD} was isolated from pACYC-*araC*-P_C-P_{BAD}-*EctrmD* (Gamper et al., 2015; Masuda *et al.*, 2019) by excision with BamHI and XhoI (NEB) and gel purified by NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). The *ts* origin of replication of pKD46 (Datsenko and Wanner, 2000) was amplified with PfuUltraII fusion HS DNA polymerase (Agilent Technologies) between the BamHI and an XhoI sites, excluding the λRed recombinase coding sequence. The PCR product was digested by BamHI and XhoI, then gel-purified by NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). The *trmD* gene and the amplified region of pKD46 were ligated by T4 DNA ligase (NEB) and transformed into competent cells of one of the three *E. coli* strains of this work. The ligated plasmid (pKD46-*EctrmD*) was verified by Sanger sequencing.

The three *E. coli* strains in this work were MG1655 (Hayashi *et al.*, 2006), BW25113 (Grenier *et al.*, 2014), and XAC (Dincbas *et al.*, 1999). After transformation with pKD46-*EctrmD,* each strain was made into *trmD*-KO by P1 transduction to replace *trmD* with a Kan marker using the P1 lysate of the *trmD*-KO strain (Gamper *et al.*, 2015; Masuda *et al.*, 2019). Each strain was grown overnight at 30 °C in LB with the presence of Ara 0.2% (w/v), ampicillin (Amp) (100 µg/mL), and Kan (50 µg/mL), allowing stable replication and expression of pKD46-*EctrmD* from the pBAD promoter. Cells were diluted at 1:100 to fresh LB + Kan and grown at 43 °C until OD_{600} = 1.0 for the first cycle, during which expression of *trmD* was turned off, replication of the *ts*-plasmid was inhibited, and cellular TrmD and m^1G37-tRNA was reduced with each cell division. Cell cultures at OD_{600} = 1.0 were diluted (1:20) and passed through two more rounds of growth and dilution to pre-warmed LB + Kan (at 1:20 each) until cell counts were not visible (~4 h). An aliquot (10 µL) of each culture was spread on an LB + Kan plate and incubated at 43 °C

48

until distinct colonies were formed. In total, 5 colonies of MG1655, 4 colonies of BW25113, and 3 colonies of XAC were isolated, streak-purified at 43 °C, verified for the absence of pKD46-*EctrmD*, and subjected to whole-genome sequencing.

Whole-genome sequencing and identification of variants

For whole-genome sequencing analysis, we chose two suppressor colonies from each strain. After streak-purification, each was grown in 3 mL LB + Kan at 37 °C overnight and cells were harvested and resuspended in 500 μL of 50 mM NaCl and 25 mM EDTA. Cell lysis was performed with the addition of lysozyme (1 mg/mL) at 37 °C for 30 min, and further by the addition of 1.5% sarcosyl for 10 min at 37 °C. Cell lysates were extracted by phenol twice, phenol-chloroform-isoamyl alcohol once, and ethanol precipitated. Chromosomal DNA was re-suspended in 200 μL TE buffer (pH = 8.0) and was used to generate a genomic library using Nextera XT DNA Library Prep Kit (Illumina). A DNA library from each WT was also prepared as a reference. Each library was paired-end sequenced by MiSeq at 2 x 150 bp, yielding 80x coverage of reads relative to the entire genome.

Sequencing reads were pre-processed using FASTQC (Version 0.11.5) (Andrews, 2010), and mapped to the *E. coli* BW25113 genome sequence (Grenier *et al.*, 2014) using BWA (Version 0.7.15 - r1140) (Li and Durbin, 2009) and SAMtools (Version 1.3.1) (Li *et al.*, 2009). Single nucleotide polymorphisms (SNPs) in each sample were retrieved as Variant Call Format (VCF) files. A comparative analysis of SNPs between all samples with the BW25113 genome was then performed using bcftools (Li, 2011). Further annotation of SNPs for each comparison was performed using SnpEff (Cingolani *et al.*, 2012). Finally, SNPs with missense and frameshift variants were considered to generate final Excel files. The SNPs only present in the suppressor genome but not in the parental WT genome were identified as variants.

Of the 6 suppressor colonies subjected to whole-genome sequencing, 5 contained a single nucleotide substitution that was mapped to *proS*, while one was unclear in sequence analysis. We therefore focused on the *proS* gene, PCR-amplified it from all suppressor clones, and performed Sanger sequence analysis of the amplified gene. We confirmed that all suppressor clones contained a single nucleotide substitution in *proS*, noting that G107C and N232S were identified independently from both MG1655 and BW25113.

Structural mapping and amino-acid sequence alignment

Each suppressor mutation was mapped to the crystal structure of *E. faecalis* ProRS (PDB 2J3L) (Crepin *et al.*, 2006) using PyMOL (Schrödinger). The amino acid sequence of the enzyme from selected bacterial species was retrieved from KEGG database (https://www.genome.jp/kegg/) and aligned by Clustal X 2.1 (Larkin *et al.*, 2007).

CRISPR reconstruction of each *proS* suppressor

Each *proS* suppressor mutation was reconstructed in the WT MG1655 using the scar-less CRISPR-Cas9 technology (Reisch and Prather, 2017). The specific sgRNA for each mutation was expressed from the pKDsgRNA plasmid, which was constructed by round-the-horn PCR using Phusion™ High-Fidelity DNA Polymerase (ThermoFischer Scientific) and self-ligation using T4 DNA ligase (NEB) (primers listed in Table S1). The sgRNA plasmid was introduced to the WT MG1655 strain that already contained the plasmid pCas9cr4 encoding a Cas9 protein. CRISPR-Cas9 system was induced on an LB plate containing anhydrous tetracycline (aTc) (100 ng/mL). Each *proS* mutation was assessed by PCR using allele-specific primers, and verified by Sanger sequencing. Subsequently, each strain was made into *trmD*-KO by P1 transduction without any maintenance plasmid, followed by removal of the Kan marker by FLP recombination with pCP20.

All reconstructed suppressors were confirmed for viability without the support of *trmD*. Early-log phase cultures were serially diluted, spotted on an LB plate, and grown overnight at 37 °C. To confirm the absence of TrmD from each suppressor, cell lysate proteins (15 µg) from a culture of 3 mL LB and harvested at OD = 0.4 were run on a 12% SDS-PAGE gel for Western blot analysis using an anti-TrmD primary antibody (Li and Bjork, 1999), an anti-CysRS primary antibody for a loading control, and an anti-rabbit IgG secondary antibody (Sigma-Aldrich).

Cellular prolyl-aminoacylation status by acid urea gel analysis

Total RNA was extracted in an acidic condition to maintain the aminoacylation status of tRNAs (Masuda *et al.*, 2021). Each reconstructed *proS* suppressor lacking *trmD* was grown in LB at 37 °C to OD = 0.4, upon which cells were mixed with 10% (w/v) trichloroacetic acid (TCA) at an equal volume and incubated on ice for 10 min. Cells were harvested and resuspended in the extraction buffer (0.3 M NaOAc, pH 4.5 and 10 mM EDTA), and extracted with equal volume of ice-cold phenol-chloroform-isoamyl alcohol (25:24:1), pH 4.5. The mixtures were vortexed for 3 cycles of 1 min vortex and 1 min rest on ice, and spun at 12,000 rpm for 10 min at 4 °C. The RNA-containing aqueous phase was stored, while a second extraction of the organic phase with an equal volume of fresh extraction buffer was performed. The combined aqueous phase from the two extractions was mixed with an equal volume of cold isopropanol and incubated at –20 °C for precipitation. Precipitated RNA was collected at 14,000 rpm for 20 min at 4 °C, and rinsed with wash buffer (70% ethanol and 30% 10 mM NaOAc, pH 4.5). Pellets were dried and dissolved in the buffer containing 10 mM NaOAc, pH 4.5, and 1 mM EDTA.

From each isolated total RNA sample, 10% was estimated to represent total tRNA (Masuda *et al.*, 2021). Approximately 700 ng of total tRNA of each sample was mixed with a 2X loading buffer (0.1 M NaOAc, pH 5.0, 9 M Urea, 0.05% bromophenol blue (BPB), and 0.05% xylene cyanol (XC)) and separated on a 6.5%PAGE/7M urea (14 x 17 cm) in 0.1 M NaOAc, pH 5.0 at

51

250 V for 3 h 45 min at 4 °C. Subsequently, the gel region between XC and BPB was excised, washed in 1X Tris-borate pH 8.0 and EDTA (TBE) buffer, transferred to a wet Hybond™-N+ nitrocellulose membrane (GE healthcare) in 1X TBE using Trans-Blot® TurboTM Transfer System (BIO-RAD) at constant 25 V for 20 min. The membranes were briefly air-dried, crosslinked with the bound RNA in a UV cross-linker (FB-UVXL-1000, Fisher Scientific), and then probed with a ^{32}P-labeled DNA oligonucleotide targeting positions 18 to 36 of *E. coli* Pro(UGG) in a Northern blot. The membrane was imaged by a phosphorimager (Typhoon FLA 9500, GE) and bands corresponding to the charged and uncharged tRNA were quantified by ImageJ (NIH). The prolyl-aminoacylation status was calculated as the band corresponding to the charged tRNA in the sum of the bands for both charged and uncharged tRNAs.

Frameshift reporter assay

The nano-luciferase (nLuc) reporter gene was subcloned into plasmid pKK223-3 using NEBuilder® HiFi DNA Assembly Cloning Kit (NEB). One reporter had an insertion of the CCC triplet for Pro, while a separate reporter had an insertion of the quadruplet CCC-C codon motif (primers in Table S1). Each reporter was introduced to a reconstructed *proS* mutant expressing a normalization plasmid pZS2R-mCherry. An overnight culture in LB + Amp + Kan at 37 °C was inoculated to a fresh medium at 1:200 for WT and 1:100 for mutants. Each fresh culture was grown at 37 °C for 6 h and measured for the luciferase activity using ONE-Glo™ Luciferase Assay System (Promega) on an Infinite M200 PRO plate reader (Tecan). The luminescence at 617 nm, normalized by the mCherry fluorescence at excitation 586 nm, was used to determine the nLuc activity. The frequency of +1-frameshifting was the ratio of the CCC-C mediated nLuc activity over the CCC-mediated activity.

Competition fitness assay

A fitness assay was developed by mixing a reconstructed *proS* suppressor expressing the plasmid pZS2R-mCherry, and the MG1655 WT strain expressing pZS2R-YFP (Kelsic *et al.*, 2015). Each strain was separately grown in LB + Kan overnight at 37 °C, and diluted 10-fold in LB. According to the OD_{600} of each, the same number of cells of each culture was inoculated to a single culture of 3 mL LB + Kan at ~10^6 colony-forming unit (CFU)/mL The 1:1 mixture was grown at 37 °C shaking, and an aliquot of the mixture was removed over time, diluted, and spread on an LB + Kan plate. After incubation at 37 °C overnight, the number of yellow colonies for WT and pink colonies for the mutant was counted. The plates were incubated at 4 °C as needed to enhance development of colors. The CFU of each culture at T = 0 was adjusted to 50:50, which was used to normalize the fractional distribution of the two strains over time.

Kinetic assay of prolyl-aminoacylation of tRNA^{Pro} *in vitro*

Each *E. coli* tRNA^{Pro} in the G37-state was synthesized by transcription, while each in the m^{1}G37-state was synthesized by post-transcriptional methylation of the transcript with TrmD (Christian *et al.*, 2004). Recombinant WT enzyme of *E. coli proS* was affinity-purified from *E. coli* BL21(DE3) expressing pET22b-*Ec* ProRS (Masuda *et al.*, 2021). Suppressor enzymes of *proS*-D98N, *proS*-T199I, and *proS*-N232S were purified similarly from strains harboring the designated mutations but expressed in ArcticExpress (DE3) (Agilent). Each tRNA substrate was heat-denatured at 85 °C for 3 min and re-annealed at 37 °C for 15 min. Aminoacylation in steady-state conditions was performed at 37 °C in a 24 μL reaction of 0.2 to 20 μM tRNA, 0.5 to 40 nM ProRS, and 20 μM [^{3}H]-Proline (Perkin Elmer, 7.5 Ci/mmol) in the aminoacylation buffer of 20 mM KCl, 10 mM $MgCl_2$, 4 mM DTT, 0.15 mg/mL bovine serum albumin, 2 mM ATP (pH 8.0), and 50 mM Tris-HCl, pH 7.5. Reaction aliquots of 4 μL were removed at different time intervals and precipitated with 5% (w/v) TCA on filter pads for 10 min twice. The filter pads were washed with 95% ethanol twice, with ether once, air-dried, and measured for radioactivity in Tri-

Carb 4910 TR scintillation counter (Perkin Elmer). Counts were converted to pmoles based on the specific activity of the [^{3}H]-Proline after correction for signal quenching by filter pads. Data corresponding to the initial rate of aminoacylation as a function of tRNA concentration were fit to the Michaelis-Menten equation to derive the K_m (tRNA), k_{cat} (catalytic turnover of the enzyme), and k_{cat}/K_m (tRNA) (the catalytic efficiency of aminoacylation) using Kaleidagraph v. 4.5 (Synergy software).

Evaluation of tRNA^{Pro} isoacceptors for cell viability

In *E. coli*, the Pro(GGG) and Pro(CGG) isoacceptors are non-essential and can be deleted. To test their role in cell viability, several reconstructed *proS* suppressor strains were transformed to harbor the maintenance plasmid pACYC-*araC*-P$_C$-P$_{BAD}$-human *trm5* (Masuda *et al.*, 2019) to afford analysis of both the m^{1}G37+ and m^{1}G37– conditions. Deletion of one of these isoacceptors was made by λRed recombination (primers in Table S1) on the WT MG1655 strain to replace the respective gene with a Kan marker (*proL* for Pro(GGG) and *proK* for Pro(CGG)). P1 lysates of the *proL*-KO strain (Masuda *et al.*, 2019) and of the *proK*-KO strain were used to transduce the reconstructed *proS* suppressor strains harboring the *trm5*-maintenance plasmid. After removal of the Kan marker by FLP recombination with pCP20, each deletion strain was maintained in the presence of Ara until analysis of cell viability in the m^{1}G37– condition.

To test cell viability, each strain was grown overnight at 37 °C in LB + chloramphenicol (Cm) (34 μg/mL) + Ara (0.2% (w/v)). Each strain was inoculated 1:100 to an Ara-free medium of LB + Cm and grown at 37 °C for 5 h for depletion of the Trm5 protein and m^{1}G37-tRNAs. Cells were then streaked out in the m^{1}G37+ condition (LB + Cm + Ara) or the m^{1}G37– condition (LB + Cm) and grown overnight at 37 °C.

Growth complementation analysis was achieved by introducing each isoacceptor with an over-expression pKK223-3 plasmid. The *proK* gene was synthesized by Sequenase extension and cloned into pKK223-3 (primers in Table S1). The resultant *proK*-pKK223-3 plasmid, and *proL*-pKK223-3 and *proM*-pKK223-3 plasmids (Gamper *et al.*, 2015; Masuda *et al.*, 2021), was separately introduced to reconstructed *proS* suppressors harboring the *trm5* maintenance plasmid. A variant of the *proM*-pKK223-3 plasmid, carrying the GGG anticodon replacing the UGG anticodon was made by site-direct mutagenesis (Masuda *et al.*, 2018). A control plasmid, over-expressing Met(CAU) (initiator tRNA^{Met} encoded by *metV*) was chosen. Expression of each tRNA in MG1655 was constitutive, due to lack of the *lacI* repressor. Growth complementation was tested in both m^{1}G37+ and m^{1}G37– conditions at 37 °C.

Codon usage analysis of CC[C/U] codons in *E. coli*

The coding sequences (CDS) of *E. coli* str. K-12 substr. MG1655 were obtained from NCBI (GenBank assembly accession: GCA_000005845.2). The frequencies of each Pro codon were calculated relative to the total number of all Pro codons in each CDS using Python (version 3.7.7). GO analysis was used to predict pathways with enriched codon usage of CC[C/U]. From the 3,876 genes analyzed, those with the frequency of CC[C/U] usage equal to or higher than 0.5 (792 genes, the top 20.4%) were selected and categorized into 4 groups according to their biological processes (http://geneontology.org/).

Temperature-sensitivity assay

The *E. coli* MG1655 WT strain and reconstructed *proS* suppressor strains were grown in LB overnight, then spotted in serial dilution on an LB plate and grown at various temperatures (37, 30, 22 (room temp), and 18 °C) for one to three overnights. Rescue of cold sensitivity was determined by over-expression of individual tRNA^{Pro} species from the pKK223-3 plasmid.

55

Phylogenetic analysis of prolyl isoacceptors across bacterial species

Bacterial tRNA genes were retrieved from the Genomic tRNA Database (GtRNAdb) (Chan and Lowe, 2016). The number of each tRNA gene was counted for each genome. The Biopython Entrez package designed for searching genome information in NCBI was used in a custom Python script to obtain the taxonomy classifications (phylum and genus) of each genome based on the taxonomy IDs. Assemblies lacking Pro(UGG) isoacceptors for 4 box tRNAs were removed from analysis, which was likely caused by an incomplete assembly (3 assemblies). Genera with three or more representative species were selected for further analysis. Genus-level frequencies of each prolyl isoacceptor were computed by normalizing the number of species with each isoacceptor by the total number of species in the genus within this curated set. Genera were categorized into Gram-positive and Gram-negative using information obtained from PubMed literature search. The same search was used to highlight genera containing medically significant species.

To present the distribution of prolyl isoacceptors across the phylogenetic tree of bacteria, sequences of the 16S and 23S ribosomal RNA (rRNA) of each genome were retrieved from NCBI Refseq (O'Leary *et al.*, 2016) or GenBank (Sayers *et al.*, 2022). If rRNA annotations of a particular genome were not available, it was excluded from further analysis (3958 retrieved, 86 skipped). Sequences of 16S and 23S rRNAs were aligned using the SINA aligner (Pruesse *et al.*, 2012) with the non-redundant SSU (small subunit of the ribosome) and LSU (large subunit of the ribosome) reference datasets (Ref NR 99) from SILVA (Quast *et al.*, 2013) respectively. The 16S and 23S alignments of each genome were concatenated to a single FASTA entry. A single representative sequence was chosen arbitrarily from each genus selected for frequency analysis of a prolyl tRNA gene. A maximum-likelihood phylogenetic tree was generated with the

prepared concatenated 16S and 23S alignments using FastTree (Price *et al.*, 2009; 2010). Tree

image was created using matplotlib with the ETE toolkit (Huerta-Cepas *et al.*, 2016).

SUPPLEMENTAL FIGURES

Figure S1. (Related to Figure 1) The cmo^{5}U34 modification expands the decoding capacity using various wobble base pairing structures.

(A) Chemical structures of cmo^{5}U, mcmo^{5}U, and ncm^{5}U.

(B) A model of expanded base-pairing by cmo^{5}U with A, G, C, and U. The hydrogen bonds are shown by dotted lines.
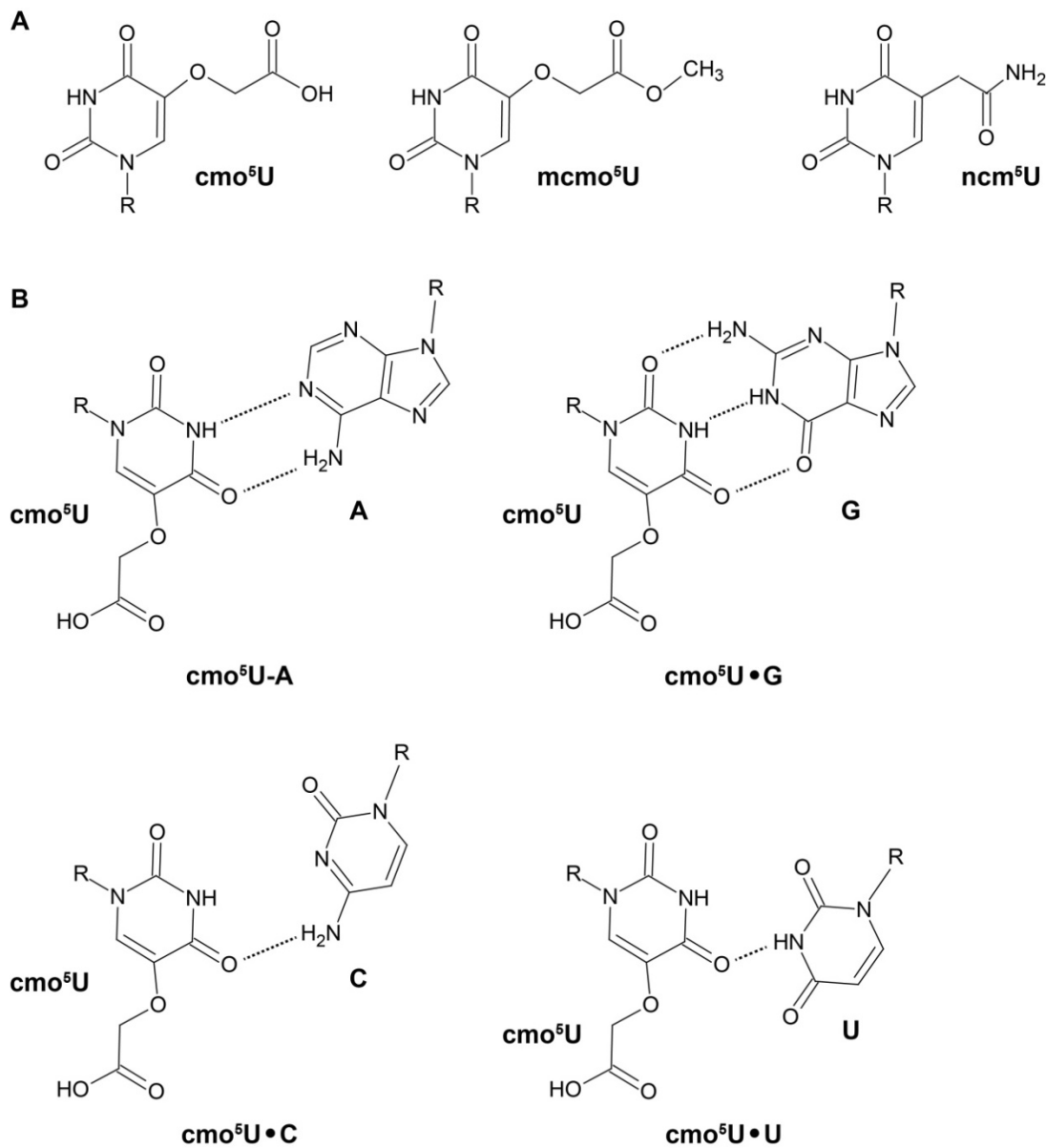
Figure S2. (Related to Figures 1 and 2) The isolated *proS* suppressor mutants are viable without TrmD but grow slower than the WT expressing TrmD.
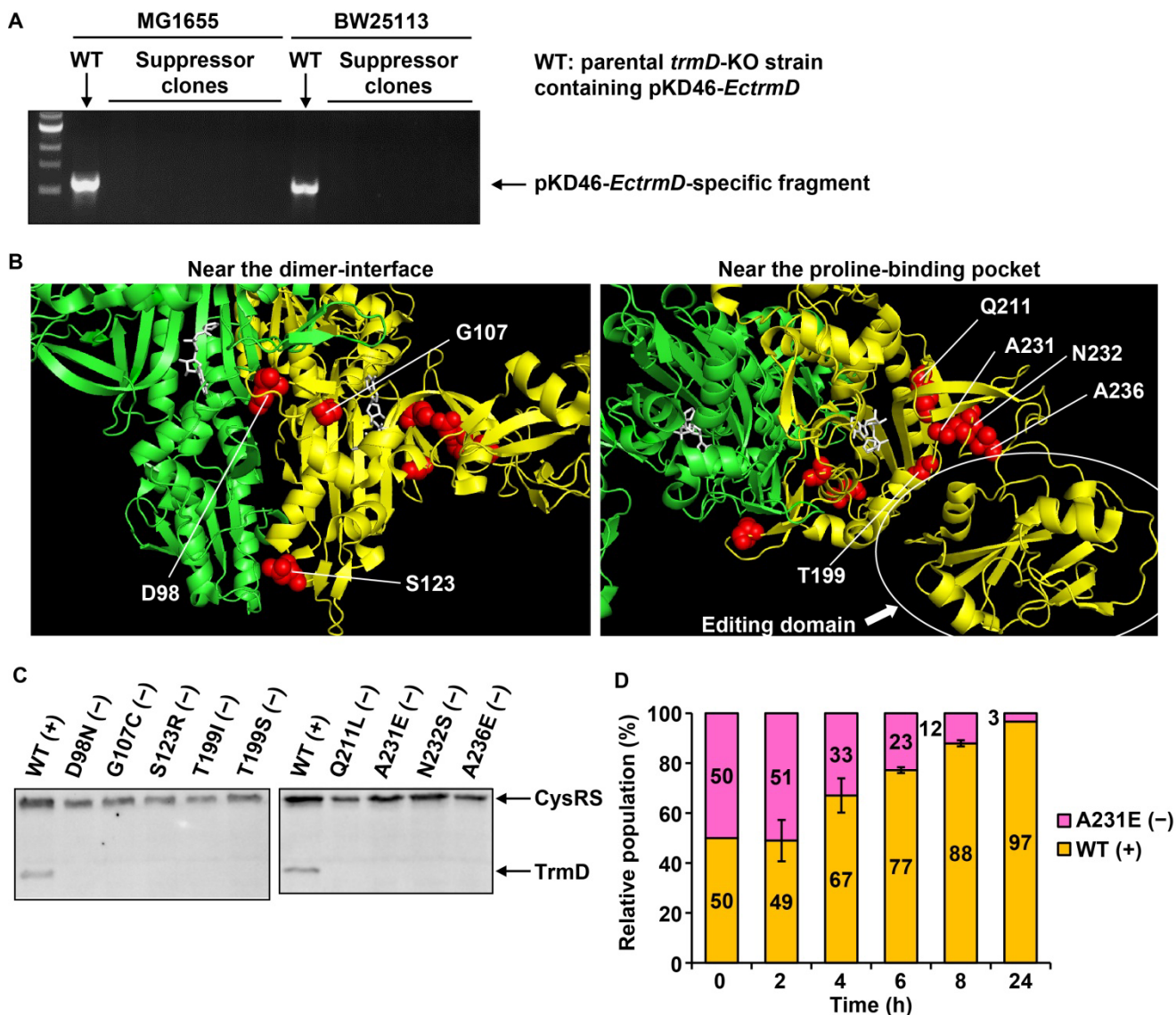
(A) Absence of the pKD46-*EctrmD* maintenance plasmid from suppressors of the *trmD*-KO allele. Suppressors were analyzed for the 5 colonies isolated from MG1655 and the 4 colonies isolated from BW25113. Absence of the plasmid was analyzed by PCR using primers specific to pKD46-*EctrmD*. The WT sample was isolated from the parental strain containing pKD46-*EctrmD*.

(B) Mapping of *proS* suppressor mutations to the structure of *E. faecalis proS* enzyme as in Figure 1D. The images are blown up for Cluster I mutations near the monomer-monomer interphase (left) and for Clusters II and III mutations near the proline-binding pocket adjacent to the proof-reading domain (right).

(C) Absence of TrmD in the reconstructed *proS* suppressors in Western blots. Whole-cell lysates of early-log phase cultures were run on a 10% SDS-PAGE gel. Proteins transferred to a PVDF membrane were probed by anti-TrmD and anti-CysRS antibodies.

(D) The *proS*-A231E (–) suppressor is out-competed by the WT (+) strain in a cell fitness assay as in Figure 2C. The bar graph shows the relative abundance of the two strains in a 1:1 mixture over the time course of 24 h.

Figure S3. (Related to Figure 3) Kinetic parameters of prolyl-aminoacylation.

A Pro(UGG)

ProRS	State of G37	K_m (µM)	k_{cat} (s^{-1})	k_{cat} / K_m (µM^{-1} s^{-1})	Fold change in k_{cat} / K_m Relative to G37-state
WT	m^{1}G37	0.96 ± 0.15	4.92 ± 0.78	5.14 ± 0.09	20.7
	G37	11.8 ± 2.24	2.91 ± 1.71	0.25 ± 0.12	1.0
D98N	m^{1}G37	0.29 ± 0.08	0.04 ± 0.003	0.13 ± 0.03	2.9
	G37	1.59 ± 0.41	0.07 ± 0.005	0.05 ± 0.1	1.0
T199I	m^{1}G37	0.61 ± 0.19	0.02 ± 0.007	0.03 ± 0.003	3.0
	G37	4.50 ± 1.90	0.05 ± 0.03	0.01 ± 0.006	1.0
N232S	m^{1}G37	0.25 ± 0.04	0.03 ± 0.01	0.13 ± 0.05	2.5
	G37	1.71 ± 0.88	0.07 ± 0.005	0.05 ± 0.03	1.0

B Pro(GGG)

ProRS	State of G37	K_m (µM)	k_{cat} (s^{-1})	k_{cat} / K_m (µM^{-1} s^{-1})	Fold change in k_{cat} / K_m Relative to G37-state
WT	m^{1}G37	0.70 ± 0.25	1.97 ± 0.62	3.12 ± 1.14	20.5
	G37	9.10 ± 2.97	1.28 ± 0.19	0.15 ± 0.03	1.0
D98N	m^{1}G37	0.26 ± 0.08	0.04 ± 0.007	0.14 ± 0.04	4.4
	G37	2.15 ± 0.47	0.07 ± 0.007	0.03 ± 0.008	1.0

C Pro(CGG)

ProRS	State of G37	K_m (µM)	k_{cat} (s^{-1})	k_{cat} / K_m (µM^{-1} s^{-1})	Fold change in k_{cat} / K_m Relative to G37-state
WT	m^{1}G37	8.33 ± 3.27	3.59 ± 1.15	0.45 ± 0.14	18.8
	G37	14.8 ± 1.27	0.35 ± 0.07	0.02 ± 0.005	1.0
D98N	m^{1}G37	0.53 ± 0.29	0.02 ± 0.003	0.07 ± 0.05	3.4
	G37	4.71 ± 0.47	0.10 ± 0.03	0.02 ± 0.006	1.0

(A) Kinetic parameters of prolyl-aminoacylation of Pro(UGG) shown in Figure 3B. For each enzyme, K_m (µM) for the tRNA substrate, k_{cat} (s^{-1}), k_{cat}/K_m (µM^{-1}s^{-1}), and the fold-change of k_{cat}/K_m of the G37-state relative to the m^{1}G37-state are shown.

(B) Kinetic parameters of prolyl-aminoacylation of Pro(GGG) shown in Figure 3C.

(C) Kinetic parameters of prolyl-aminoacylation of Pro(CGG) shown in Figure 3D.

Figure S4. (Related to Figure 4) Requirement of Pro(GGG) for cell viability of *proS*-T199I and *proS*-N232S mutants in m^1G37- condition.
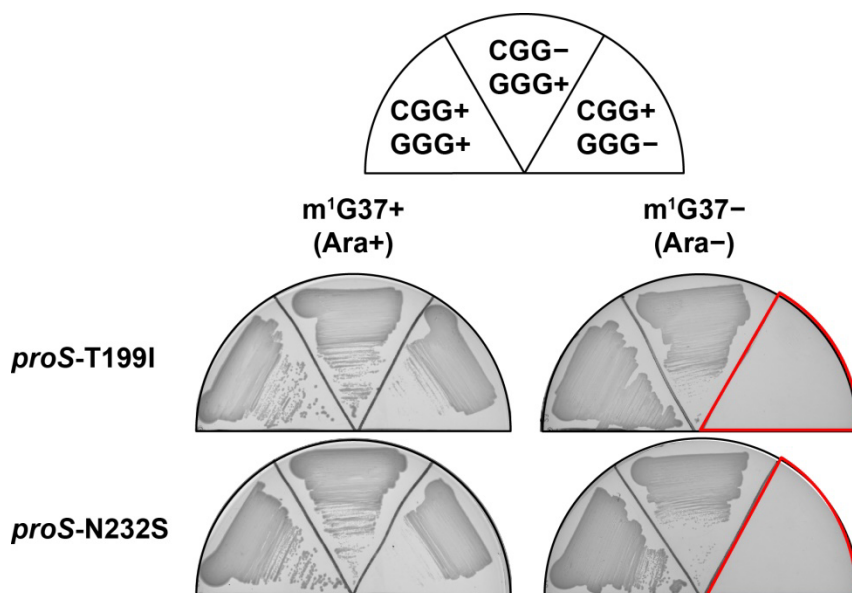
Pro(CGG) or Pro(GGG) was removed from *proS*-T199I and *proS*-N232S mutants lacking *trmD* and cells were grown in m^1G37+ or m^1G37- condition as in Figure 4C. No growth was observed when both m^1G37 and Pro(GGG) were absent.

Figure S5. (Related to Figure 5) High frequency of CC[C/U] codons in *E. coli csp* genes.
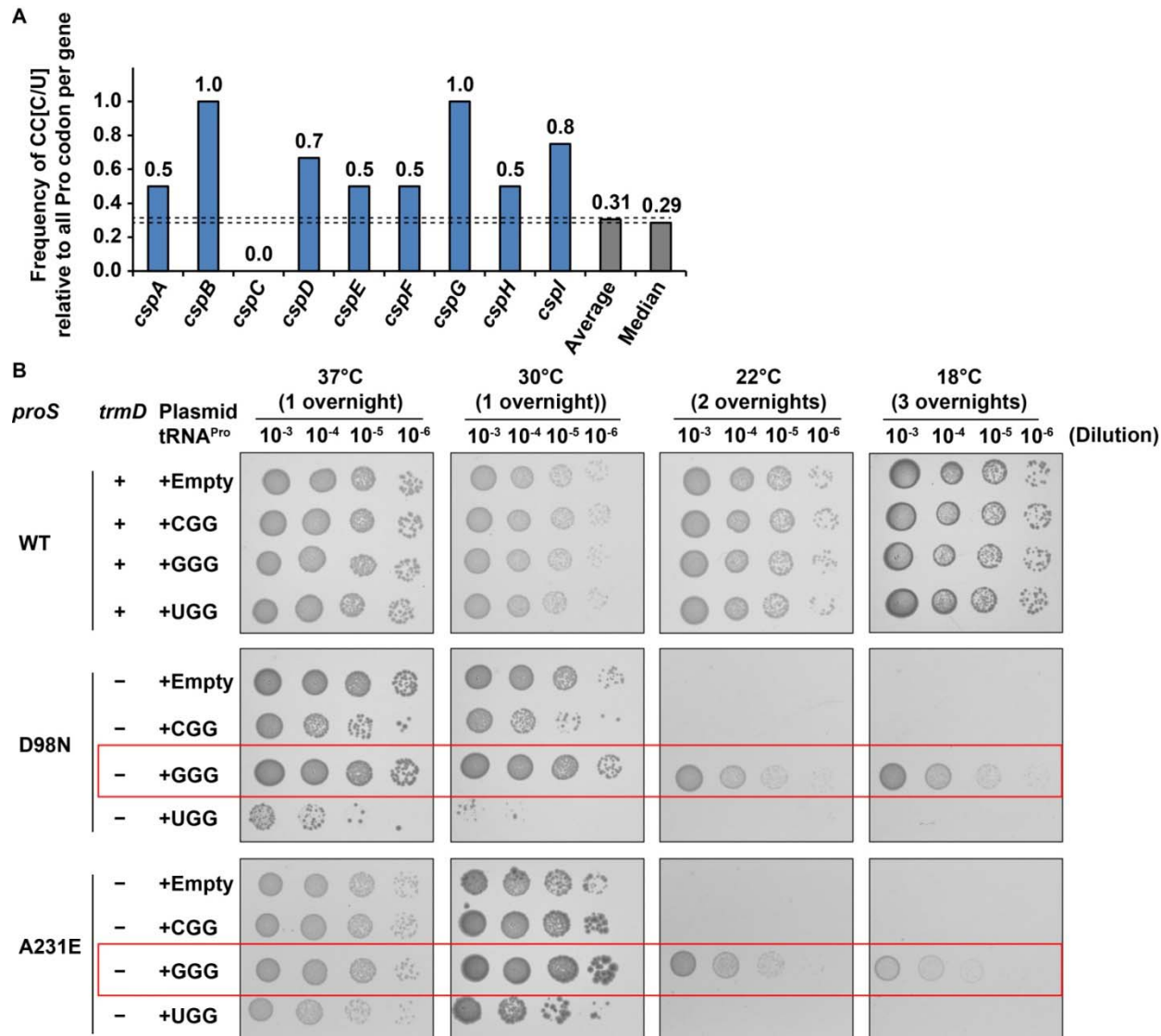
(A) The frequency of CC[C/U] codons is shown for selected *csp* genes, showing higher frequencies than the average and median (0.31 and 0.29, respectively) for most of these genes.

(B) Rescue of cold sensitivity by expression of Pro(GGG) in *proS*-D98N and *proS*-A231E suppressors as in Figure 5D. The *proS*-A231E mutant was incubated at 30 °C for 2 overnights due to a slower growth phenotype.

Figure S6. (Related to Figure 6) Co-occurrence of Pro(GGG) and Pro(CGG) across bacterial

species.

Heat-map table showing frequencies of co-occurrence of Pro(GGG) and Pro(CGG) across all

bacterial species that are available in GtRNAdb. Counts are species within GtRNAdb beyond

the curated set used in Figure 6A. Few bacterial species have Pro(CGG) but lack Pro(GGG),

whereas many have Pro(GGG) but lack Pro(CGG).

SUPPLEMENTAL TABLES

Table S1. Primers used in this study.

Category	Sequence (5'-3')	Purpose
Gene knockout		
E. coli proK-KO	AGGGCGTATCTGCGCAGTAAGATGCGCCCCGCATTGTGTAGGCTGGAGCTGCTTC	Forward to amplify *Kan* from pKD4 with homologous extension
	AAAAAGCCTGCTCGTTGAGCAGGCTTTTCGAATTATGGGAATTAGCCATGGTCC	Reverse to amplify *Kan* from pKD4 with homologous extension
CRISPR mutation		
E. coli proS	GTGCTCAGTATCTCTATCACTGA	Reverse to amplify pKDsgRNA
	CTGCTGCGTTTTGTTGACCGGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for D98N
	TGAGTTGGGCCGAGTACGAAGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for G107C
	AGCTAAGCTCGTTACGAATCGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for creating a new PAM site to make S123R
	TGTAAGAGCTAAGCTCGTTAGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for S123R
	CAAGCCGACACCGGTTCTATGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for T199I and T199S
	CTCTCACGAATTCCAGGTGCGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for Q211L
	CTATGCAGCGAACATTGAACGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for A231E and N232S
	CTATGCAGCGAACATTGAACGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA for creating a new PAM site for making A236E
	AGCTTCTGCCAGTTCAATATGTTTTAGAGCTAGAAATAGCAAG	Forward to amplify pKDsgRNA or A236E
	GGGAACAGTACGGTCCGGAACTGCTGCGTTTTGTGAATCGTGGCGAGCGTCCGTTCGTACTCGGCCCAACTCATGAAGAA	Mutation donor for making D98N
	CTGCGTTTTGTTGACCGTGGCGAGCGTCCGTTCGTACTGTGTCCAACTCATGAAGAAGTTATCACTGACCTGATTCGTAA	Mutation donor for making G107C
	GTACTCGGCCCAACTCATGAAGAAGTTATCACTGACCTTATCCGTAACGAGCTTAGCTCTTACAAACAGCTGCCGCTGAA	Mutation donor for making a new PAM site for making S123R
	CTCATGAAGAAGTTATCACTGACCTTATCCGTAACGAGTTGCGCTCTTACAAACAGCTGCCGCTGAACTTCTATCAGATC	Mutation donor for making S123R
	CCGCATGGGGCTGGATTTCCGCGCCGTACAAGCCGATATTGGTTCTATCGGCGGCAGCGCCTCTCACGAATTCCAGGTGC	Mutation donor for making T199I
	CCGCATGGGGCTGGATTTCCGCGCCGTACAAGCCGATTCTGGTTCTATCGGCGGCAGCGCCTCTCACGAATTCCAGGTGC	Mutation donor for making T199S
	CAAGCCGACACCGGTTCTATCGGCGGCAGCGCCTCTCACGAATTTCTTGTGCTGGCGCAGAGCGGTGAAGACGATGTGGT	Mutation donor for making Q211L

	GACGATGTGGTCTTCTCCGACACCTCTGACTATGC AGAAAATATTGAACTGGCAGAAGCTATCGCGCCGA AAGAACCGCG	Mutation donor for making A231E
	GACGATGTGGTCTTCTCCGACACCTCTGACTATGC AGCCAGTATTGAACTGGCAGAAGCTATCGCGCCGA AAGAACCGCG	Mutation donor for making N232S
	GACGATGTGGTCTTCTCCGACACCTCTGACTATGC AGCCAATATTGAACTGGCAGAAGCTATCGCGCCGA AAGAACCGCG	Mutation donor for making a new PAM site for making A236E
	TCTCCGACACCTCTGACTATGCAGCCAATATTGAAT TAGAAGAAGCTATCGCGCCGAAAGAACCGCGCGCT GCTGCTACC	Mutation donor for making A236E
Probe for Northern blotting		
E. coli Pro(UGG) tRNA	CCAAACCAGTTGCGCTACCA	Annealing to the tRNA region from the D-loop to the anticodon stem-loop
Plasmid construction		
pKD46-*EctrmD*	GCCATTCTCGAGAATGGCGATGACGCATCCTCAC	Forward to amplify pKD46 with an XhoI site
	AAGCGTCAGGTAGGATCCGCTAATCTTATGG	Reverse to amplify pKD46 with a BamHI site
pET22b-*EcproS* mutant	AAGGAGATATACATATGCGTACTAGCCAATACCTGC TCTC	Forward to amplify *proS* with an NdeI site
	GGTGGTGGTGCTCGAGGCCTTTAATCTGTTTCACC AGATAT	Reverse to amplify *proS* with an XhoI site
pKK223-3-*proK*	TCACCGGAATTCCGGTGATTGGCGCAGCCTGGTAG CGCACTTCGTTCGGGACGAAG	Forward to construct *proK* with EcoRI site
	CGACCACTGCAGTGGTCGGTGATAGAGGATTCGAA CCTCCGACCCCTTCGTCCCGAACGAAG	Reverse to construct *proK* with PstI site
pKK223-3-N-luc	AAGCTTGGCTGTTTTGGCGGATG	Forward to amplify the pKK223-3 backbone
	GAATTCTGTTTCCTGTGTGAAATTGTTATC	Reverse to amplify the pKK223-3 backbone
	CACACAGGAAACAGAATTCATGCCCGGTGTCTTCA CACTCGAAGATTTCG	Forward to amplify nLuc with a CCC at the second codon position
	CACACAGGAAACAGAATTCATGCCCCGGTGTCTTC ACACTCGAAGATTTCG	Forward to amplify nLuc with a CCC-C codon at the second position
	CGCCAAAACAGCCAAGCTTTTACGCCAGAATGCGT TCGCACAGC	Reverse to amplify nLuc
Confirmation of plasmid curing		
pKD46-*EctrmD*	CATATTGCATCAGACATTGCCG	Forward to amplify a partial fragment of pKD46-*EctrmD*
	CTATGTGCCATCTCGATACTCG	Reverse to amplify a partial fragment of pKD46-*EctrmD*

Table S2. Comparison between this study by Masuda *et al*. and the study by Clifton *et al* (Clifton *et al*., 2021).

	Masuda *et al.* (2022)	Clifton *et al.* (2021)									
Chromosomal *trmD*	KO	Y86*	R154A	S165L	G117N	R154A	S165L	Y86*	Y115A	G141A	D169A
Mutation description	Complete deletion	Y86 changed to a stop codon	Activity largely abolished	Colony size reduced by 90%	Colony size reduced by 90%	Activity largely abolished	Colony size reduced by 90%	Y86 changes to a stop codon	Catalytically compromised	Catalytically compromised	Activity largely abolished
Level of growth defect	Severe	Severe			Intermediate	Severe		Severe	Little	Little	Severe
Growth condition	LB, 43 °C	LB, 37 °C									
Days of cell culture	2 to 3	9			15	22		30	–	–	–
proS mutations	D98N G107C S123R T199I T199S Q211L A231E N232S A236E	A231E	E19K A36V Y89N G107C G200S A231E E514V	D98N G200S G200A A231E L235Q M441I		E19K G200S A231E E514V	G200S E514K	V195L G200A A231E			
trmD coding region mutation (Yes or No)	N.A.	Yes	Yes	Yes	Yes	Yes	Yes	Yes			Yes (A169E)
rimM-trmD intergenic region mutation (Yes or No)	N.A.			Yes	Yes			Yes			
tRNA^Arg gene mutations						argQ argW argY argZ	argV	argQ argV argY argZ			
Other gene mutations including intergenic region				acrD fecA cfa nirB ulaG		spoT yfbM-yfbN intergenic yifK-argX intergenic		insH1 intergenic lrhA-alaA intergenic rpoC spoT yifK-argX intergenic ytfK			

Shown in red are *proS* suppressor mutations in common between the two studies.