Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

Summer 10-2-2014

Translational Fidelity of a Eukaryotic Glutaminyl-tRNA Synthetase with an N-terminal Domain Appendage

Aaron Bethea Rogers
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Chemistry Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Rogers, Aaron Bethea, "Translational Fidelity of a Eukaryotic Glutaminyl-tRNA Synthetase with an N-terminal Domain Appendage" (2014). *Dissertations and Theses*. Paper 2006.
<https://doi.org/10.15760/etd.2005>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Translational Fidelity of a Eukaryotic Glutaminyl-tRNA Synthetase
with an N-terminal Domain Appendage

by

Aaron Bethea Rogers

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Chemistry

Thesis Committee:
John Perona, Chair
Niles Lehman
David Peyton

Portland State University
2014

© 2014 Aaron Bethea Rogers

ABSTRACT

Several *Saccharomyces cerevisiae* mutant tRNA^{Q2} species (glutamine isoacceptor, CUG anticodon) were synthesized and assayed for aminoacylation activity with *Saccharomyces cerevisiae* glutaminyl-tRNA synthetase. The derived steady state parameters were compared to similar datasets from the literature. The mutants behaved analogously to similar mutant species based on tRNA from *Escherichia coli*, but with slightly relaxed specificity as revealed by comparison of $k_{\text{cat}}/K_{\text{M}}$ values relative to wild type *in vitro* transcribed tRNA. Additionally the eukaryotic N-terminal domain appendage, as found in *Sce* glutaminyl-tRNA synthetase, is considered in light of the discovery of non-canonical aminoacyl-tRNA synthetase functions, including its role in the assembly of the multiple aminoacyl-tRNA synthetase complex.

Dedication

For my **very** patient wife, Adriana

Table of Contents

Abstract.....	i
Dedication.....	ii
List of Tables.....	iv
List of Figures.....	v
List of Schemes.....	vi
 Chapter 1	
Introduction.....	1
 Chapter 2	
Experimental.....	19
 Chapter 3	
Results & Discussion.....	28
 Chapter 4	
Conclusion.....	41
 References.....	42
 Appendices	
A. Supplemental Figures.....	48
B. DNA Oligomers used in tRNA Synthesis	62
C. Michaelis-Menten Plots.....	64
D. Supplemental File Descriptions.....	74

List of Tables

Table 1. tRNA ^{Q2} consensus nucleotides & mutant <i>Sce</i> tRNA ^{Q2} species	28
Table 2. Steady state parameters for glutamine, compared to literature values	29
Table 3. Steady state parameters for tRNA, compared to literature values	29
Table 4. Steady state parameters for Gln, compared to <i>Eco</i> QRS with <i>Eco</i> tRNA ^{Q2} mutants	31
Table 5. Steady state parameters for tRNA, compared to <i>Eco</i> QRS with <i>Eco</i> tRNA ^{Q2} mutants	31
Table 6. Further information, including UniProtKB accession numbers, on predicted association partners, as predicted by STRING and BioGRID	38

List of Figures

Figure 1. <i>a)</i> tRNA 2° structure cloverleaf <i>b)</i> tRNA 3° fold	3
Figure 2. Class I (left) and class II (right) common tRNA identity elements	6
Figure 3. <i>Eco</i> QRS, colored according to description in text, in complex with tRNA ^{Q2} (PDB id 1QTQ)	8
Figure 4. Structural rearrangements of <i>Eco</i> QRS upon tRNA binding	10
Figure 5. The ribonucleoprotein <i>Eco</i> tRNA ^{Q2} bound to <i>Eco</i> QRS (PDB id 1QTQ)	11
Figure 6. Domain expansion in <i>Sce</i> QRS, as illustrated in the partial crystal structures of the C-terminal catalytic core (residues 215 to 809; PDB 4H3S) & the helical N-terminal appendage (residues 1 to 187; PDB 3TL4)	13
Figure 7. <i>Eco</i> (left) and <i>Sce</i> (right) tRNA ^{Q2} sequences, compared to <i>Sce</i> mt-tRNA ^{Q1}	15
Figure 8. Comparison of <i>Sce</i> (left) and <i>Hsa</i> (right) QRS and Arc1p or AIMP1 associations, respectively, as predicted by STRING	36
Figure 9. Comparison of <i>Sce</i> (left) and <i>Hsa</i> (right) QRS associations as predicted by BioGRID	37

List of Schemes

Scheme 1. Generic mechanism of tRNA aminoacylation via class II aaRSs (top) and class I aaRSs (bottom)	5
Scheme 2. tRNA ^{Q2} charging and nucleolytic liberation of modified ³² P labeled nucleoside	24

1. INTRODUCTION

i. molecular information in terrestrial life

The canonical role of nuclear/genetic polynucleotides made from deoxyribonucleic acid (DNA) contain the information necessary for the production of proteins. This is initially achieved via transcription of discrete segments of DNA, known as genes, into the complementary polynucleotide comprised of ribonucleic acids (RNA), commonly referred to as messenger RNA (mRNA). These mRNAs are subsequently translated into protein/polypeptide gene products via the ribosomal assembly, a ribonucleoprotein (RNP) complex.

Specific interactions between transfer RNA (tRNA) and mRNA facilitate the processes of genetic information being translated into peptide/protein product. The nature of these interactions give rise to a phenomenon commonly referred to the genetic code. The letters of this "alphabet" are not simply the single constitutive canonical nucleobases (adenine, uridine, cytidine, and guanine), but rather groupings of three consecutive nucleobase residues. These give rise to codons, which are "interpreted" into meaning (protein) by the translational apparatus of the cell. Most of the 64 codons correspond to one of the twenty canonical amino acid monomers, but three are reserved as signals indicating that transcription should terminate and mRNA be released. Just downstream of transcription factor promoter sequence, the "START" codon of a contiguous DNA sequence coding for gene product (peptide/protein) regions also serves as the codon for methionine. As such, all synthesis of proteins and peptides begin with methionine, but in many cases the methionine is subsequently removed by enzymatically mediated post-translational modification.

The genetic code is nearly universal, implying that all life on this planet ultimately originated from a single form of early life, be it either a single discrete universal common ancestor or communal prebiotic entities coexisting, competing, and freely exchanging amongst the group. Codons are not equally distributed amongst the twenty amino acids, and as such some amino acids correspond to a single codon (Met, Trp), while others correspond to six codons (Leu, Ser, and, Arg). These give rise to the one-box, two-box, etc. codon sets. While the genetic code rarely varies in amino acid meaning, codons are not utilized at the same frequency across the various phyla within the three main evolutionary branches of terrestrial life.

Assembly of polypeptide is achieved with amino acids delivered to the ribosomal assembly via transfer RNA (tRNA). Amino acids arrive covalently attached to the 2' or 3' hydroxyl oxygen of the tRNA 3' terminus, where they are subsequently hydrolyzed off and covalently added by the ribosome assembly to the C-terminus of the nascent polypeptide. Hybridization between the codon of an mRNA and the anti-codon (reverse complement) of tRNA facilitates correct amino acid utilization by the ribosome. Ostensibly, in a genome there is at least one tRNA species per the 64 codons, but many species possess a wealth of paralogous tRNA genes, all with the same anticodon. For example, *Homo sapiens* (*Hsa*) has 26 genes that appear to code for tRNA^{Q2} (Gln CUG) alone (Supplemental data 4). Essentially nothing is understood regarding the role of these extra sets *in vivo*.

ii. tRNA commonalities (consensus of shape)

While tRNAs must have enough structural and sequence diversity to be differentiated amongst one another, they are quite similar in secondary and tertiary structure. All tRNAs must be able conform to the ribosome binding pockets for protein translation to proceed.¹ The classic depiction of tRNA is via the so-called clover leaf structure (Figure 1a) that illustrates the four major arms arising from internal Watson-Crick (WC) base pairing. However, tRNAs are replete with non-WC base pairings far apart in sequence that, in conjunction with stacking in the core, give rise to the three-dimensional fold of tRNA that overall resembles an L-shape (Figure 1).

Figure 1.¹ a) tRNA 2° structure cloverleaf, with standard numbering b) tRNA 3° fold

tRNA can be thought of as having five major secondary structural features: the acceptor stem, the D-stem loop, the anticodon stem loop, the variable loop, and the T-stem loop.¹ The length of the variable loop defines the class of tRNA: shorter variable arms define class I tRNAs, while longer variable loops are characteristic of class II tRNAs. While many tRNAs adhere to this general feature set, deviant species with stems

and loops varying in length (or missing altogether) emerge from the massive datasets of the modern genomic and bioinformatic sciences. Many of these non-standard instances are found in tRNAs from organellar genomes.¹

Following the standard numbering system for tRNAs (Figure 1), the bases comprising the anticodon are found at N34, N35, and N36 relative to the 5' terminus. These correspond to the reverse complement of the codon found on mRNA, with the wobble position being located at N34 (corresponding to complement base of the appropriate 3' mRNA codon residue). Adhering to the same numbering system, the single stranded 3' terminus invariably consists of C74, C75, and A76 in mature tRNAs.² It is at the ubiquitous A76 that amino acid is covalently attached via the sugar hydroxyl moiety.

iii. aaRSs: chemistry, quality control, and evolutionarily directed multi-purposing

tRNA functions not only as a molecular shuttle, but also participates in one of the principal quality control checkpoints of translation: the covalent attachment of cognate amino acid to the correct tRNA by an **aminoacyl-tRNA synthetase (aaRS)**. This family of proteins, organized into two main classes (I & II), demonstrate remarkable diversity in structure and organization across species and classes, but all contain highly conserved catalytic cores (within their class) that effect same fundamental chemistry, between amino acid and tRNA. All aaRSs achieve "charging" of tRNA by means of a two step reaction pathway (Scheme 1). In the first step, an aaRS activates amino acid by hydrolysis of rATP, with concomitant release of inorganic pyrophosphate (PP_i). As rAMP is a good leaving group, the 3' terminal A₇₆ covalently attaches aminoacyl moiety

via attack of the 2' or 3' OH (Class I or Class II aaRS, respectively) group of the backbone carboxyl carbon of the adenylated amino acid intermediate, with loss of rAMP.²

Scheme 1. Generic mechanism of tRNA aminoacylation via class II aaRSs (top) and class I aaRSs (bottom)

Cells produce protein with an estimated error rate of 10^3 - 10^4 ,³, and much of this fidelity comes from the ability of aaRSs to selectively charge cognate tRNA with cognate amino acid. This is achieved by a vast species and aaRS dependent toolbox including direct readout of consensus tRNA residue moieties via the major and minor grooves, kinetic mechanism, induced fit, negative determinants, obligate quaternary structure beyond the simple 1:1 aaRS:tRNA complex, and indirect readout of tRNA shape via direct and solvent mediated protein-RNA contacts (Figure 2). Additionally there are

several examples of aaRSs that possess editing domains that are capable of further quality check via hydrolysis of activated non-cognate adenylated amino acid (pre-transfer editing) or via hydrolysis of the non-cognate aminoacyl moiety from tRNA just after erroneous attachment (post-transfer editing).⁴

Figure 2.⁵ Class I (left) and class II (right) common tRNA identity elements

In species- and organelle-dependent fashion, proper charging of tRNA is sometimes achieved via an indirect pathway. In these instances, non-cognate amino acid is "incorrectly" incorporated into tRNA corresponding to another amino acid. Another protein subsequently effects chemistry upon the mischarged aminoacyl-tRNA at the precursor amino acid side chain, yielding the correctly charged aminoacyl-tRNA. One example is glutamyl-tRNA^{Qx} (Gln-tRNA^Q, representing both Gln anticodons: UUG and CUG) synthesis in Archaea, organelles, and most bacteria outside Proteobacteria; these organisms' genomes lack a glutamyl(**Q**)-tRNA synthetase (QRS). However, they still incorporate glutamine into their proteins and utilize the glutamine codon set. In many of these cases the solution comes by way of a **non-discriminating** (ND) glutamyl(**E**)-tRNA synthetase (ERS) that lacks the ability to distinguish between tRNA^Q and tRNA^E species, indiscriminately charging both with glutamate.² Mischarged glutamyl-tRNA^Q (Glu-

tRNA^Q) is transamidated to Gln-tRNA^Q by a separate enzymatically catalyzed reaction. For example, the mitochondria in yeast utilize this pathway by importation of the cytosolic nuclear-encoded ND-ERS, which charges mitochondrial tRNA^Q with glutamate. The tRNA-dependent amidotransferase assembly GatFAB then effects the transamidation of Glu-tRNA^Q to Gln-tRNA^Q.⁶ Erroneous incorporation of incorrect amino acid from mischarged tRNA into nascent polypeptide is limited by discrimination from elongation factor channeling.⁷

Beyond similar tertiary folds and (generally) highly conserved active site residues, the aaRSs are amongst the most rapidly evolved group of enzymes across phylogenetic distances. The aaRSs are replete with domain appendages, **insertions/deletions** (indel), and functional expansions.⁸ Furthermore, many human aaRSs have been implicated in a number of cytoplasmic, nuclear, and extracellular roles including cytokine signaling, apoptosis, tumorigenesis, inflammation, translational suppression, transcriptional regulation, and more.⁹

The evolutionary expansion of aaRS function is perhaps most dramatically exhibited by the emergence of the **m**ultiple **a**minocyl-tRNA synthetase complex (MARS) found in some Eukarya, including humans.¹⁰ The core complex consists of nine aaRSs and at least¹¹ three accessory/scaffold proteins, hitherto referenced as **a**minoacyl-tRNA complex-**i**nteracting **m**ultifunctional **p**rotein (AIMP). It appears that MARS exists in loose association with the ribosome, in part through extensive interactions with elongation factors.¹² However, there are possibly other weakly interacting partners.¹³ The AIMPs have been implicated in a broad array of regulatory and signaling roles similar in scope to the non-canonical functions adopted in eukaryotic aaRSs.¹⁴

iv. Eco QRS

The enzyme responsible for glutamyl-tRNA^Q synthesis in *Escherichia coli* (*Eco*) has been extensively studied¹⁵ using numerous techniques including (but not limited to) crystallographic^{16–19} and kinetic^{20,21} methods. The 553 amino acid, 63 kDa monomeric protein consists of four principal domains with specific roles (Figure 3).

Figure 3. *Eco* QRS, colored according to description in text, in complex with tRNA^{Q2} (PDB id 1QTQ)

The active site lies closer to the N-terminus, comprising the two halves of the **R**ossmann fold (RF, residues 8-100; 210-270; orange) separated by the **a**cceptor stem **b**inding

domain (ASB, residues 101-209; lime green). The **h**elical subdomain (HS, residues 271-339; purple). The **a**nticodon **b**inding domain (ACB, residues 340-547) is made up by two subdomains, the proximal β -barrel (residues 340-346; 461-547; mauve) separated by the distal β -barrel (residues 347-460; pale green). Of interest, the anticodon binding domain shares structural homology with Ribosomal protein L25.²² The seven N-terminal residues and six C-terminal residues have not demonstrated ordered crystal structure.¹⁵

Eco QRS has been visualized via X-ray crystallography as both tRNA co-crystal (with various ligation partners) and apoenzyme.¹⁶ Rearrangements of *Eco* QRS upon tRNA binding can be demonstrated via structural alignments (Figure 4). These shifts of structure are required for the active site residues of *Eco* QRS to adopt a catalytically proficient arrangement.¹⁵ As such, in the absence of bound tRNA^Q *Eco* QRS, like only a few other Class I aaRSs, is unable to catalyze the first step in aminoacylation,²³ the activation of amino acid with rATP. As A76 of tRNA^{Qx} participates in proper orientation of rATP substrate in the active site cleft,² the 1:1 QRS:tRNA^{Qx} complex can be considered an obligate ribonucleoprotein.

The identity elements required for proper *Eco* QRS recognition of tRNA^Q have been scrutinized. The identity and role of specific residues in both biopolymers are well established²⁴ (see Table 1, Results & Discussion). There may yet be unknown determinants, as many studies have used unmodified *in vitro* transcribed tRNA, and the shape of the 3° tRNA fold have been shown to affect aminoacylation even when the full complement of identity nucleotides are still present.²⁵ In short, the most important

Figure 4. Structural rearrangements of *Eco* QRS upon tRNA binding
top, detail of structural alignments for several *Eco* QRS structures, colored by RMSD;
bottom, residue number vs. RMSD (Å) plot of the same data in the top

residues in *Eco* tRNA^Q required for correct *Eco* QRS recognition are found at the discriminator base (N73), the acceptor stem (N1:N72, N2:N71, and N3:N70), in the D arm stem (N10:N25, not detailed in Figure 5), and throughout the anticodon loop (N34, N35, N36, etc.). Figure 5 details the spatial orientation of the anticodon and the acceptor stem relative to protein. Other bases are known to contribute to proper recognition, just not at the level of those mentioned above.

Figure 5. The ribonucleoprotein *Eco* tRNA^{Q2} bound to QRS (PDB id 1QTQ)
top-left, N- and C-termini (red and green, respectively), as viewed looking down the T- and acceptor-stems
top-right, acceptor stem; first three base pairs are colored red, discriminator blue, and the CCA tail cyan
bottom-right, additional detail of acceptor end, illustrating discriminator base stacking with C₇₅ and A₇₆
bottom-left, anticodon; C₃₄, U₃₅, and G₃₆ (cyan, orange, and red, respectively)

v. Sce QRS

Like all Eukarya, the genome of *Saccharomyces cerevisiae* (*Sce*) contains a coding sequence for QRS, located within its *GLN4* gene.²⁶ While the C-terminal 595 residues of the gene product contain the common features of a class I aaRS, there is also a 224 residue N-terminal extension that is absent in Bacteria, including *Eco* (Figure S1). This domain appendage is present in most Eukarya, especially through Opisthokonta, but the exact role of this appendage is poorly understood.

Recently, two separate partial structures of *Sce* QRS were solved. The most recent²⁷ is of the C-terminal, catalytically active (residues 215-809) portion of the molecule and contains the same overall domain architecture as found in *Eco* QRS, expected given significant fold and ~38% sequence identity of the catalytic cores (Figure 6). Beyond the addition of the N-terminal domain, two insertions are found in residues 364-368 (within the ASB) & 751-770 (within the ACB), the former of which appears to be a putative zinc binding motif specific to Ascomycota.

Previously, the N-terminal appendage was crystallized (residues 1-187).²⁸ The first half of the Rossmann fold is preceded (progressing towards the N-terminus) by a flexible linker connecting to a four-coil helix-turn-helix subdomain bundle, followed by another flexible linker leading to a separate seven-coil helix-turn-helix subdomain bundle (Figure 6). The C-terminal four helix subdomain demonstrated significant fold similarity when structurally compared to the tail domains of the GatB subunits of the *Staphylococcus aureus* and *Thermotoga maritima* amidotransferases despite low sequence identity. Additionally, the electrostatic potential surface calculated from the QRS NTD structure implies a co-localized positively charged partial surface suitable for

nucleic acid binding. This arises in part from two clusters of lysine residues near the N-terminal end of the NTD (residues 19, 20, 28, and 29) and near the C-terminal end of the NTD (residues 188, 189, 190, 196, 197, and 198).

Figure 6. Domain expansion in *Sce* QRS, as illustrated in the partial crystal structures of the C-terminal catalytic core (residues 215 to 809; PDB 4H3S) & the helical N-terminal appendage (residues 1 to 187; PDB 3TL4)

top-center, crystal structure of the *Sce* QRS C-terminal catalytic core, oriented & domain color-coded analogously to *Eco* QRS (see Figure 3); tRNA would bind on the left side of the enzyme
top-right, detail of the two *Sce* QRS insertions (relative to *Eco* QRS), slight rotation to the left
top-left, detail of the resolved N-terminal (red) & C-terminal (green) residues, slight rotation to the right
bottom-left, crystal structure of the *Sce* QRS N-terminal appendage, with detail of the resolved N-terminal (red) & C-terminal (green) residues
bottom-right, 180° rotation from left, illustration of the two helical bundles (ochre, green) connected by a flexible linker (red, dark space)

In the same study²⁸, the role of the NTD in *Sce* QRS was examined via steady state kinetic analysis (see lower sections of Tables 2 and 3) and *in vivo* viability screenings. Mutant yeast cells expressing only an N-terminal domain (NTD) deletion mutant of *Sce* QRS (Δ NTD-QRS, residues 211-809) exhibited slower growth at lower temperatures and increased susceptibility to L-methionine sulfoximine, a glutamine synthase inhibitor. However, the mutant cell lines remained viable, and at higher culture temperatures (30°C), growth rates were comparable to cell lines expressing full length QRS in the same genetic construct, implying that the domain is dispensable. The kinetics demonstrate that the NTD has little effect on catalytic rate (k_{cat}), but K_M values for tRNA and glutamine are significantly increased, with the tRNA K_M value being the more greatly perturbed. Additionally, mutations were introduced into the NTD. Alignments of eukaryotic QRS NTD appendages demonstrate little sequence identity. However, some well conserved residues are present; these are located in a disordered flexible linker region between the N-terminal seven-helix subdomain and the C-terminal four-helix domain. The mutant G112P (located in the flexible linker) was also examined with steady state analysis, yielding a 10-fold increase in K_M for tRNA and a doubling of k_{cat} (see Tables 2 and 3, Results & Discussion). This mutant, in addition to three more mutant constructs (G112A/V113A/G114A, W160F, and W160A) all exhibited decreased tRNA binding as evidenced by electrophoretic mobility shifts in the presence of varied amounts of protein. None of these mutants exhibited loss of tRNA binding capacity at the level of the Δ NTD *Sce* QRS variant.

vi. *Eco* tRNA^{Q2} vs. *Sce* tRNA^{Q2} vs. *Sce* mitochondrial tRNA^{Q1}

tRNA identity elements are not always conserved across species. This could be due in part to evolutionary pressure on tRNA sequence/structure brought on by divergence of their corresponding aaRSs, perhaps via the emergence of new domains mediated by aaRSs adopting novel non-canonical function within (and without) the cell. Comparison of the *Eco* & *Sce* tRNA^{Q2} cloverleaf representations (Figure 7) illustrate changes in a number of tRNA identity elements important for recognition by *Eco* QRS. Trans-kingdom identity shifts at the discriminator base and in the acceptor stem (*Eco* to *Sce*: G73U, U1G/A72C, and G3U/C70A) should be sufficient to preclude appreciable activity of *Eco* QRS upon *Sce* tRNA^{Q2} (see Tables 4 and 5, Results & Discussion).^{24,29,30} However, the G10:C25 identity is conserved between the nuclear *Sce* and *Eco* tRNAs, as are the seven (unmodified) bases of the anticodon loop. *Sce* mitochondrial tRNA^{Q1} (mt-tRNA) has a number of commonalities and deviations relative to , perhaps most notably at the discriminator base (A73).

Figure 7.³¹ *Eco* (left) and *Sce* (middle) tRNA^{Q2} sequences, compared to *Sce* mt-tRNA^{Q1}

Accordingly, *Eco* QRS is incapable of charging bulk *Sce* tRNA, nor is *Sce* QRS (as present in cell extract) able to charge bulk *Eco* tRNA.³² A hybrid construct of *Eco* QRS with the NTD of *Sce* QRS appended to its N-terminus is able to rescue *GLN4* deficient mutants constitutively expressing the hybrid product via a secondary gene. The authors of this study indicate that the hybrid *Eco* QRS exhibited some lack of specificity in *Sce* bulk tRNA (cross charging was reportedly observed for *Sce* tRNA^K & tRNA^E), but appeared to retain specificity for tRNA^Q as found in bulk *Eco* tRNA. Furthermore, the authors indicated that the *Eco* QRS NTD hybrid construct continued to exhibit "preference" for *Eco* tRNA, as evinced by faster aminoacylation rates compared to bulk *Sce* tRNA.

A subsequent study from the same group³³ examined which residue subset of the NTD was responsible for the trans-kingdom rescue. This was achieved by way of similar *in vivo* experiments where several truncations of the *Sce* QRS NTD was appended to *Eco* QRS. The results demonstrated that *Eco* QRS appended with *Sce* NTD residues 1-228 (full length), 1-179, and 43-228 were able to rescue *Sce* QRS knockout alleles. However, similar appendage of *Sce* NTD residues 1-42, 180-228, 1-106, and 43-179 were not sufficient to rescue these knockout strains; appendage of residues 107-228 yielded cultures with extreme growth retardation. In the same study, an array of double and double-double alanine mutations were introduced into the lysine rich regions of the NTD and demonstrated that none of these residues alone are able to attenuate aminoacylation beyond detectable levels. However, some of the double-double mutants demonstrated no activity, but only when one set of double mutants was in the N-terminal NTD helical bundle and the other set was in the C-terminal NTD helical bundle. It has also been

reported³⁴ that the *Sce* QRS NTD alone (residues 1-228) binds yeast tRNAs, a viral pseudoknot, and a single-stranded poly-A 35-mer with roughly similar affinity, implying that the NTD non-specifically binds double- and single-stranded nucleic acids, seemingly regardless of 3° fold.

The *Sce* protein Arc1p (i.e. GU4 nucleic-binding protein 1, G4p1) shares approximately 27% identity with the *Hsa* MARS scaffold protein AIMP1 and, like the eukaryotic QRS NTD, contains putative nucleic acid binding motifs. The protein also serves as scaffold in the association of *Sce* MRS and ERS into a three membered complex reminiscent of the MARS complex observed in "higher" Eukarya. It also appears to enhance binding of cognate tRNA isoacceptors for both aaRS members of this complex.³⁵ The association appears to occur via folds (in all members) similar to glutathione S-transferase (GST), resulting in GST dimerization interfaces, a common protein-protein binding mode.³⁶ Arc1p has also been implicated in preferential binding of G4 quadruplexes.³⁷ Additionally, fusion of *Sce* Arc1p to *Eco* QRS, like the chimeric NTD_{*Sce*}-QRS, is able to rescue QRS knockout yeast strains upon supplementation *in vivo* and also demonstrates a marked increase in the rate of aminoacylation of *Sce* tRNA relative to WT *Eco* QRS *in vitro*.³⁴ Furthermore, fusion of the (unrelated) eukaryotic nonspecific nucleic acid binding N-terminal appendage of *Sce* or *Hsa* valyl-tRNA synthetase (VRS) to *Eco* QRS also rescued QRS knockout yeast strains.³⁸ Interestingly, the native *Sce* QRS exhibited varied binding affinities to a mixed tRNA population, whereas the two N-terminal appended *Eco* QRS constructs appeared to bind across the population more or less equivalently. Finally, another group demonstrated³¹ that *Eco* QRS N-terminally fused with Arc1p is able to rescue the indirect *Sce* mitochondrial

glutaminyl-tRNA^Q synthesis pathway *in vivo* (via mitochondrial amidotransferase knockout alleles) and exhibited significant aminoacylation towards mt-tRNA^{Q1}, whereas the WT *Eco* and *Sce* did not.

vii. unknown molecular recognition, direct or otherwise?

The functional role of domain appendages for many aaRSs have been studied, and in many cases within *Hsa* those appended domains serve regulatory roles separate from canonical aminoacylation activity.⁸ Prior work^{39,40} with other domain-appended aaRSs establishes precedent for more closely examining the role of the N-terminal appendage as found within *Sce* QRS, perhaps in its primitive form. Additionally, molecular dynamics simulations of modeled full length *Sce* QRS (3TL4 and 4H3S), in part based on solution small angle X-ray scattering, suggest dynamics previously not observed in QRS systems.²⁷ Specifically, the NTD was predicted to potentially make contacts with the catalytic core. Additionally, the simulation predicts distortion of tRNA upon NTD binding, compelling U73 of tRNA^{Q2} to flip inwards toward enzyme, potentially making meaningful contacts with Leu463.

Herein is described a study investigating what, if any, influence trans-kingdom isoacceptor tRNA identity element shifts have regarding the N-terminal extension of *Sce* QRS. This is achieved via steady state analysis of *Sce* tRNA^{Q2} species mutated at the discriminator base, the acceptor stem, and the anticodon. Additionally, the nature of the N-terminal extension is briefly considered, with discussion focused on *Hsa*.

2. EXPERIMENTAL

i. protein

Full length *Sce* QRS (residues 1-809), provided by labmate Andrew Hadd, were expressed and purified as previously described.²⁸ Protein was stored at -20 °C in 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 5 mM β -mercaptoethanol (β -ME), and 50% (v/v) glycerol.

ii. DNA template synthesis via Klenow fragment

tRNA was enzymatically synthesized via chemically synthesized DNA oligomers, principally via adoption of the methods previously described.⁴¹ In summary, a DNA oligomer corresponding to the tRNA gene, with bacteriophage T7 RNA polymerase promoter and TATA binding box appended upstream of tRNA sequence, is synthesized via Klenow fragment utilizing two DNA oligomer templates, one corresponding to roughly half of the 5' forward sequence and the other corresponding to the first 5' half of the sequence reverse complement, such that there is an approximately ten base pair overlap to ensure cross-primer hybridization. Additionally, the two terminal 5' residues of the reverse complement DNA template are 2' *O*-methylated to ensure product tRNA 3' homogeneity. Transfer RNA is then produced via T7 transcription of the ~100 base pair DNA product. Mutation of *Sce* tRNA^{Q2} sequence was achieved by altering the design of the DNA templates at the appropriate position. Specific DNA oligomer sequences are detailed in Appendix B.

Specifically, Klenow reaction was performed at equimolar DNA template concentrations (4 μ M), 10 units of Klenow fragment (Thermo Scientific), 10 mM each

dNTPs, and the manufacturer supplied enzyme buffer stock at the appropriate dilution. Normally, reactions were executed at 2 nmol scale (500 μ L reaction volume). Reaction was effected via PCR block with the following program: hold at 10 $^{\circ}$ C for 45 seconds, ramp up to 37 $^{\circ}$ C at 0.2 $^{\circ}$ C/sec, hold for 45 seconds, ramp down to 10 $^{\circ}$ C at 0.2 $^{\circ}$ C/sec, and repeat 15 cycles. Upon program completion, the block is held indefinitely at 4 $^{\circ}$ C. Lid temperature was constant at 100 $^{\circ}$ C. DNA was then precipitated by diluting 1:1 (v/v) with ethanol and adding sufficient sodium acetate and acetic acid to bring their concentrations to $\sim$ 300 mM and $\sim$ 200 mM, respectively. This solution was then placed in the freezer at -20 $^{\circ}$ C at least overnight. Before use, Klenow product is spun at 14000 RPM for 30 minutes to pellet DNA. Supernatant is removed, the pellet is washed with 80% (v/v) cold ethanol, and re-pelleted at 15000 CFG for 10 minutes. Supernatant is removed again, the pellet is dried under vacuum centrifugation for $\sim$ 10 minutes, and resuspended in 18 M Ω water filtered through .22 μ m PES syringe filter (VWR). Product is verified on 2% (w/v) agarose gel against 100 base pair DNA ladder (Figure S2). As no evidence of DNA oligomer template primers is observed in these gel analyses, reaction is assumed to have progressed to completion. As such, each normal scale reaction yields approximately 1 nmol of DNA duplex template.

iii. tRNA synthesis and purification

T7 transcription reaction was carried out with half ($\sim$ 0.5 nmol) of the Klenow-produced DNA template product in solution containing 40 mM Tris-HCl (pH 8.0), 25 mM MgCl₂, 2 μ M spermidine, 0.01% (v/v) Triton X-100, 40 mM dithiothreitol (DTT), 4 mM each Mg:rNTPs, 10 μ L containing 0.01 mg of yeast PP_iase (Roche), and 200 μ L T7

RNA polymerase (5.99 mg/mL, Perona lab stock; strain 87) per 4 mL reaction. After reaction overnight at 37 °C, the product tRNA is recovered from the upper phenolic/aqueous layer upon addition of an equal volume of biphasic extraction mixture: 25:24:1 phenol:chloroform:isoamyl-alcohol, pH 4.1 (Sigma). Care was taken to ensure crashed salts and protein at the biphasic interface weren't aspirated along with upper layer. A second extraction was then carried out on the first aqueous layer in the exact same fashion. The second extract was concentrated to ~250-500 µL via 10 kDa molecular weight cutoff (MWCO) ultracentrifugation device (EMD Millipore), mixed 1:1 (v/v) with 2x RNA loading dye {15% (v/v) formamide, 5 mM EDTA (pH 8.0), 0.025% w/v sodium dodecyl sulfate, bromophenol blue, and xylene cyanol}, and loaded onto a 250 mL 10% (v/v) 19:1 acylamide:bisacrylamide 7 M urea-PAGE preparatory gel pre-run for 30 minutes at 30 Watts in Tris-HCl/boric acid/EDTA (TBE) running buffer.

Separation is achieved by running at 50 Watts for approximately 4 hours. tRNA is identified by UV shadow (Figure S3) and excised from bulk gel using a sterilized scalpel. The tRNA containing gel piece is then placed in a falcon tube, crushed, and submerged with extraction solution of 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (TE). After shaking overnight at room temperature, extraction solution is recovered and filtered through 0.45 µm PES syringe filter (VWR) followed by a second filtration through 0.22 µm PES syringe filter (VWR). The filtrate is concentrated with 10 kDa MWCO ultracentrifugation device (EMD Millipore), followed by several buffer exchanges with TE in 3 kDa MWCO microfuge ultracentrifugation device (EMD Millipore).

The tRNA is subsequently precipitated, pelleted, dried, and reconstituted into TE solution using the same procedure as described for the DNA template Klenow product in

the preceding section. Concentration determination was achieved by measuring UV absorption at 260 nm, utilizing the molar absorptivity coefficient ($\epsilon_{260 \text{ nm}}$) calculated by OligoCalc (<http://www.basic.northwestern.edu/biotools/oligocalc.html>),⁴² as based upon mutant sequence input. The ratio A_{260}/A_{280} was always >2. Fidelity of product was ascertained by analytical urea-PAGE separation, visualized with toluidine staining, and compared to another co-run tRNA, typically *Eco* tRNA^{Q2} (Figure S4). Yields of purified tRNA were typically 0.5 to 1.0 mg (~20 to 40 nmol) per 0.5 nmol DNA template.

iv. ³²P labeling of tRNA

Incorporating ³²P radionuclide into the 5' internucleotide linkage between the 3' terminal A76 and C75 (N-1) is achieved by the *Eco* CCA-adding nucleotidyltransferase in the presence of [α -³²P] rATP, as previously described.²¹ The labeling reaction is carried out in 25 mM HEPES (pH 7.4), 6.25 mM MgCl₂, 0.3125 mM DTT, 2 μ M tRNA, 2.5 mM PP_i, 5 μ L [α -³²P] rATP (PerkinElmer) at 3000 Ci/mmol (or, 10 mCi/mL, per assay date), and 6 μ L CCA enzyme (5.99 mg/mL, Perona lab stock, strain 26) per 80 μ L. Reactions are typically at a total volume of 80 μ L and executed at 37 °C.

Before reaction, tRNA is refolded by heating tRNA stock and water to 80 °C for three minutes, transferred to 65 °C for five minutes, the magnesium containing CCA buffer (also at 65 °C) is added, then the metal block is removed from the heat plate and allowed to slow cool on the bench. PP_i and [α -³²P] rATP are added once folded tRNA has cooled to 37 °C; reaction is then initiated by addition of CCA enzyme. The reaction is allowed to proceed for one minute, and 2 μ L containing 0.002 mg of yeast PP_iase (Roche) is added and allowed to incubate for another four minutes at 37 °C. Reaction is

terminated by addition of 25:24:1 phenol:chloroform:isoamyl-alcohol, pH 4.1 (Sigma). After recovery of the phenol-aqueous layer (top) followed by second biphasic extraction, labeled tRNA is purified and concentrated as with T7 transcripts. Purification was achieved via 55 mL purification gel (~11% (v/v) 29:1 acylamide:bisacrylamide), containing no urea. Visualization of labeled tRNA is achieved by brief exposure of gel to phosphor storage screen, printing out 1:1 scale to use as guide (Figure S5). Labeled tRNA is then extracted overnight at 4 °C with TE buffer (in the same fashion described above for T7 transcript purification), filtered through a 0.45 µm PES centrifugal filter, buffer exchanged, ethanol precipitated, vacuum dried, reconstituted in TE buffer, stored in aliquots at -20 °C, and used within two weeks. As the radionuclide is located at the internucleotide linkage between the unpaired 3' terminal A76 and C75 residues, damage due to ionizing radiation should be readily visualized by TLC due to potential for liberation of mono- and di-nucleotides. Fidelity was estimated with every set of assays by spotting undigested ³²P labeled tRNA (directly from freezer stock and after folding) onto PEI and running in parallel to digested assay time points.

v. steady state aminoacylation assay

Steady state assay was adopted from a previously described method,²¹ and were performed per Scheme 2.

Scheme 2. tRNA^{Q2} charging and nucleolytic liberation of modified ³²P labeled nucleoside

All reactions were executed at 37 °C in 50 mM Tris-HCl (pH 7.48 at 37 °C), 10 mM MgCl₂, 10 mM β-ME, and 5 mM Mg:rATP. Reaction volume was 20 μL in all cases. Glutamine and rATP freezer stocks were brought to pH ~7 by addition of 1 M Tris-HCl and 50% (w/v) NaOH, respectively. Glutamine concentrations were varied from 0.1 to 60 mM for K_M^Q measurements. tRNA concentrations were varied from 0.02 to 40 μM for K_M^{tRNA} determinations. Trace ³²P labeled tRNA was volumetrically constant in all time courses at 1 μL freezer stock per 20 μL reaction. Substrate saturation

was ensured for glutamine and tRNA (Figures S6, S7), but was assumed for rATP. Enzyme concentrations remained constant within each set of substrate range time courses, but were suitably varied to accommodate tRNA mutants. As such, QRS concentrations were varied from 1 nM to 100 nM, but the ratio of tRNA:QRS was maintained above 20:1, ensuring that measurements were within the multiple turnover regime. All tRNAs were refolded prior to assay as described for the CCA labeling reaction and left at room temperature just prior to use.

After reactions were initiated in individual time courses by addition of 10 μ L of 2x tRNA solution to 10 μ L of 2x {QRS, Glutamine, rATP}. Both solutions contained equal concentrations of the remaining buffer components. Aliquots (2 μ L) were taken every 10 seconds and diluted five fold into 8 μ L of solution containing 400 mM sodium acetate (pH 5.2) with 0.25 units nuclease P₁ (Sigma-Aldrich) to quench the reaction and digest tRNA. Subsequently, one μ L of this solution was spotted onto scored pre-run **polyethyleneimine** (PEI) cellulose plates and developed with solution containing 400 mM ammonium acetate with 5% (v/v) acetic acid to separate pA₇₆-Gln and pA₇₆-OH, corresponding to reacted and unreacted tRNA respectively. Plates were dried, wrapped in plastic wrap, and exposed overnight to phosphor screen. The exposed phosphor screen was then scanned using Typhoon scanner (GE Healthcare Life Sciences) and spot intensities were quantitated using ImageQuant (GE Healthcare Life Sciences). The ratios of pA₇₆-Gln to unreacted pA₇₆-OH can then be used to calculate moles of product formed. Individual time course initial rates are determined via linear regression analysis, as included with Excel (Microsoft); only aminoacylation values corresponding to less than 20% total reacted tRNA substrate were used. Non-linear fitting of initial rate time

courses to the Michaelis-Menten equation were achieved via KaleidaGraph (Synergy Software), yielding the Michaelis-Menten constant (K_M) and maximal velocity (V_{max}), which in turn yield k_{cat} , given enzyme concentration. Initial estimates for the fits were estimated by visual inspection of plotted data; maximum allowable error was defined at 0.1%, per KaleidaGraph dialogue entry. MM plots can be found in Appendix C. Raw data for reported values and FASTA files can be found in Appendix D.

vi. alignments (sequence- & structure-based), visualizations, data mining, and association networks

Multiple sequence alignments were performed in the Visual Molecular Dynamics (VMD) environment via the MultiSeq⁴³ plug-in. Protein sequence based alignments were performed with CLUSTAL-W.⁴⁴ Structure based alignment was achieved using the STAMP⁴⁵ algorithm. All structures used were obtained from either the RCSB PDB (<http://www.rcsb.org>) or from the Protein Data Bank in Europe (<http://www.ebi.ac.uk/pdbe>). All sequences used came from the UniProtKB (<http://www.uniprot.org/help/uniprotkb>). Homology models were constructed on UniProtKB reviewed entries using the SWISS-MODEL web server (<http://swissmodel.expasy.org>). All tRNA sequences were obtained from the UCSC tRNADB⁴⁶ (<http://gtrnadb.ucsc.edu>), and tRNA alignments were the product of the deposited COVE alignments of the predicted gene products located on that site.

Protein-protein associations were predicted by STRING v9.1⁴⁷ (<http://string-db.org>) by the multiple names search mode at the default (0.400) confidence level. For *Hsa* results, the genes QARS and AIMP1 were the input; results for coexpression, textmining, and neighborhood were excluded for the sake of clarity. For *Sce* results, the

genes GLN4 and ARC1 were used as input; results for coexpression and textmining were excluded. Additional protein-protein interaction predictions were provided by BioGRID⁴⁸ v3.2.114 (<http://thebiogrid.org>) using the input QARS and GLN4 for *Hsa* and *Sce*, respectively. Intrinsic disorder predictions were provided via the DISOPRED3/DISOPRED2 algorithm provided by the PSIPRED Protein Sequence Analysis Workbench⁴⁹ (<http://bioinf.cs.ucl.ac.uk/psipred>) webserver using the canonical FASTA sequences for *Hsa* QRS (full length: 1-775; NTD: 1-225) and *Sce* QRS (NTD: 1-228) were used as input.

3. RESULTS & DISCUSSION

i. does the Sce QRS NTD influence tRNA recognition?

To ascertain the extent of influence the QRS NTD has in relation to tRNA^Q identity elements, eight mutant *Sce* tRNA^{Q2} species were constructed, of which five were subjected to steady state analysis. Wild type glutamine-based measurements were performed for internal comparison (Figures S8 and S9). Mutations were introduced in the acceptor stem, the anticodon loop, the D-stem, and at the discriminator base (N73). Table 1 lists the homologous tRNA^{Qx} bases for *Eco* and *Hsa* as compared to *Sce*; mutations introduced into *Sce* tRNA^{Q2} are included in the same table.

Table 1. tRNA^{Q2} consensus nucleotides & mutant *Sce* tRNA^{Q2} species used in this study

	<i>Eco</i> tRNA ^{Qx}	<i>Sce</i> tRNA ^{Qx}	<i>Hsa</i> tRNA ^{Qx}	Mutant <i>Sce</i> tRNA ^{Q2}	
	N ₇₃	G ₇₃	U ₇₃	(U/G/A/C) ₇₃	U73A, U73G, U73C
ACCEPTOR STEM	N ₁ :N ₇₂	U ₁ :A ₇₂	G ₁ :C ₇₂	(G/U/C) ₁ :(C/A/G) ₇₂	G1U/C72A
	N ₂ :N ₇₁	G ₂ :C ₇₁	G ₂ :C ₇₁	(G/U/C) ₂ :(C/A/G) ₇₁	G2A/C71U (untested)
	N ₃ :N ₇₀	G ₃ :C ₇₀	U ₃ :A ₇₀	(G/U/C/A) ₃ :(C/A/G/A) ₇₀	U3G/A70C (untested)
	...	...	...	...	...
D-STEM	N ₁₀ :N ₂₅	G ₁₀ :C ₂₅	G ₁₀ :C ₂₅	(G/U/A/C) ₁₀ :(C\U\G) ₂₅	G10A/C25U (untested)
	...	...	...	...	...
ANTICODON	N ₃₅	U ₃₅	U ₃₅	U ₃₅	none
	N ₃₆	G ₃₆	G ₃₆	G ₃₆	G36C

Between *Eco* and *Sce*, major identity shifts are evident at the discriminator base (N73) and within the acceptor arm (N1:N72 and N3:N70). The remaining major identity elements are in consensus. Of note is the *Eco* U1:A72 pair (compare to G1:C72 in *Sce*), as the majority of *Eco* tRNA utilize a G1:C72 pair at this position. Additionally, structural evidence¹⁹ and kinetic analysis^{24,30} (partially depicted in Table 4) implicate this identity as one of the major determinants for *Eco* QRS selecting against tRNA^E. Specifically, Leu136, located within CP1, appears to be involved in disrupting the base

pairing between U1 and A72, such that the 3'-terminal GCCA_{OH} (N73-N76) is free to form a hairpin back upon itself, as observed in the active site (Figure 5).

Eco tRNA^{Q2} G73 is homologous to *Sce* tRNA^{Q2} U73. Substitution of tRNA^{Q2} G73 to U73 in *Eco* has been shown²⁴ to have a pronounced detriment upon tRNA "specificity" (k_{cat}/K_M) comparable to mutations within the anticodon (see Table 5). Another study, utilizing glutamine-varied steady state analysis,²⁴ also demonstrated the ability of this mutant to have significant effect on specificity (see Table 4). The other two possible mutations of *Eco* G73 to either A or C also had significant effect.

Steady state parameters for five *Sce* tRNA^{Q2} mutants (U73A, U73G, U73C, G1U/C72A, and G36C) were determined as described and are presented in Tables 2 and 3.

Table 2. Steady state parameters for glutamine, compared to literature values

System	K_M^{Gln} (M)	$k_{\text{cat}}^{\text{Gln}}$ (s ⁻¹)	k_{cat}/K_M (M ⁻¹ ·s ⁻¹)	% relative to WT	% relative to WT
WT	0.0016 ± 0.0007	2.57 ± 0.25	1.57 × 10 ³	100.0%	190.8%
U73A	0.0230 ± 0.0006	23.03 ± 0.87	1.00 × 10 ³	63.7%	121.5%
U73G	0.0343 ± 0.0255	5.27 ± 2.56	1.54 × 10 ²	9.8%	18.6%
U73C	0.0023 ± 0.0003	0.42 ± 0.01	1.79 × 10 ²	11.4%	21.8%
G1U:C72A	0.0220 ± 0.0082	0.65 ± 0.14	2.97 × 10 ¹	1.9%	3.6%
G36C	0.0266 ± 0.0074	0.26 ± 0.05	9.89 × 10 ⁰	0.6%	1.2%
WT QRS ²⁸	0.0017 ± 0.0002	1.4 ± 0.2	8.24 × 10 ²	N/A	100.0%
ΔNTD-QRS ²⁸	0.0093 ± 0.0003	1.7 ± 0.3	1.80 × 10 ²	N/A	21.9%
G112P QRS ²⁸	ND	ND	ND	N/A	ND
WT QRS (native tRNA) ²⁸	ND	ND	ND	N/A	ND

Table 3. Steady state parameters for tRNA, compared to literature values

System	K_M^{tRNA} (M)	$k_{\text{cat}}^{\text{tRNA}}$ (s ⁻¹)	k_{cat}/K_M (M ⁻¹ ·s ⁻¹)	% relative to WT	% relative to WT
WT	ND	ND	ND	ND	ND
U73A	(1.70 ± 0.21) × 10 ⁻⁶	16.52 ± 5.24	9.71 × 10 ⁶	ND	131.8%
U73G	(1.23 ± 0.52) × 10 ⁻⁵	2.32 ± 0.80	1.89 × 10 ⁵	ND	2.6%
U73C	(5.02 ± 0.17) × 10 ⁻⁶	0.50 ± 0.07	9.96 × 10 ⁴	ND	1.4%
G1U:C72A	(2.56 ± 0.49) × 10 ⁻⁶	1.48 ± 0.54	5.78 × 10 ⁵	ND	7.8%
G36C	ND	ND	ND	ND	ND
WT QRS ²⁸	(1.90 ± 0.40) × 10 ⁻⁷	1.4 ± 0.2	7.37 × 10 ⁶	N/A	100.0%
ΔNTD-QRS ²⁸	(5.85 ± 0.52) × 10 ⁻⁶	1.7 ± 0.3	2.91 × 10 ⁵	N/A	3.9%
G112P QRS ²⁸	(1.55 ± 0.51) × 10 ⁻⁶	2.8 ± 0.6	1.81 × 10 ⁶	N/A	24.5%
WT QRS (native tRNA) ²⁸	(1.40 ± 0.01) × 10 ⁻⁷	1.7 ± 0.1	1.21 × 10 ⁷	N/A	164.8%

For the purpose of discussion, WT *Sce* tRNA^{Q2} experimental and literature²⁸ k_{cat}/K_M discrepancies were assumed to be similar for tRNA- and glutamine-derived value sets (note that ratio $(k_{\text{cat}}/K_M)_{\text{exp}}/(k_{\text{cat}}/K_M)_{\text{lit}}$ is greater than one, Table 2, right column). Of interest, the k_{cat} values for U73A and U73G were greater than WT; the remaining mutants (U73C, G1U/C72A, and G36C) exhibited rate reductions of about 10-fold. All observed mutant K_M values were elevated relative to WT. Steady state analyses with tRNA were not performed on WT or G36C, due to time constraints. As such, only comparison to tRNA-derived values of WT tRNA^{Q2} from the literature²⁸ are provided for comparison with this dataset. All measurements were performed in triplicate, with the exceptions of U73G and U73C, which were measured in duplicate.

Comparison of this data with similar mutational studies^{24,30} in *Eco* tRNA^{Q2} is found in Tables 4 and 5 for glutamine- and tRNA-derived parameters, respectively. Inspection indicates that both systems tolerate N73A mutants better N73C mutants. Additionally, both systems exhibit an elevation in k_{cat} with N73A mutants. Increases in K_M compensate for this, resulting in lowered specificity, more pronouncedly in specificity comparisons derived from tRNA measurements. However, it should be noted that the tRNA derived values for *Eco*³⁰ may have been determined from assay conditions in which glutamine was not at saturating concentrations, as ¹⁴C labeled glutamine substrate was employed at 0.1 mM, well below the glutamine K_M values subsequently reported.²⁴ N73C mutants were detrimental in both systems, but relatively more tolerated by *Sce* QRS, at least in glutamine-based measurements. Mutating *Sce* tRNA^{Q2} at the discriminator base to the consensus *Eco* tRNA^{Q2} N73 (U73G) negatively impacted

specificity via increases in K_M^Q and K_M^{tRNA} ; observed k_{cat} values were comparable or faster (as with glutamine values).

Table 4. Steady state parameters for Gln, compared to *Eco*²⁴ QRS with *Eco* tRNA^{Q2} mutants

System	K_M^{Gln} (M)	$k_{\text{cat}}^{\text{Gln}}$ (s ⁻¹)	k_{cat}/K_M (M ⁻¹ ·s ⁻¹)	% relative to WT (experiment)	% relative to WT (2012 NAR) ²⁸	
WT	0.0016 ± 0.0007	2.57 ± 0.25	1.57 × 10 ³	100.0%	190.8%	
U73A	0.2300 ± 0.0006	23.03 ± 0.87	1.00 × 10 ³	63.7%	121.5%	
U73G	0.0343 ± 0.0255	5.27 ± 2.56	1.54 × 10 ²	9.8%	18.6%	
U73C	0.0023 ± 0.0003	0.42 ± 0.01	1.79 × 10 ²	11.4%	21.8%	
G1U:C72A	0.0220 ± 0.0082	0.65 ± 0.14	2.97 × 10 ¹	1.9%	3.6%	
G36C	0.0266 ± 0.0074	0.26 ± 0.05	9.89 × 10 ⁰	0.6%	1.2%	
Escherichia coli	In vivo ²⁴	2.14 × 10 ⁻⁴	2.80	13084.1	76.1%	N/A
	In vitro (G1) ²⁴	2.89 × 10 ⁻⁴	4.97	17185.7	100.0%	N/A
	G73A ²⁴	1.61 × 10 ⁻³	12.70	7888.2	45.9%	N/A
	G73U ²⁴	1.06 × 10 ⁻³	2.60	2459.8	14.3%	N/A
	G73C ²⁴	1.56 × 10 ⁻³	1.82	1164.5	6.8%	N/A
	U35A ²⁴	5.62 × 10 ⁻³	0.15	27.0	0.2%	N/A
	U35C ²⁴	4.88 × 10 ⁻³	0.23	47.8	0.3%	N/A
	G36A ²⁴	1.82 × 10 ⁻³	0.77	421.7	2.5%	N/A
	G36U ²⁴	1.12 × 10 ⁻³	6.22	5565.5	32.4%	N/A

Table 5. Steady state parameters for tRNA, compared to *Eco*³⁰ QRS with *Eco* tRNA^{Q2} mutants

System		K_M^{tRNA} (M)	$k_{\text{cat}}^{\text{tRNA}}$ (s ⁻¹)	k_{cat}/K_M (M ⁻¹ ·s ⁻¹)	% relative to WT (experiment)	% relative to WT (2012 NAR) ²⁸
	WT	ND	ND	ND	ND	ND
	U73A	$(1.70 \pm 0.21) \times 10^{-6}$	16.52 ± 5.24	9.71×10^6	ND	131.8%
	U73G	$(1.23 \pm 0.52) \times 10^{-5}$	2.32 ± 0.80	1.89×10^5	ND	2.6%
	U73C	$(5.02 \pm 0.17) \times 10^{-6}$	0.50 ± 0.07	9.96×10^4	ND	1.4%
	G1U:C72A	$(2.56 \pm 0.49) \times 10^{-6}$	1.48 ± 0.54	5.78×10^5	ND	7.8%
	G36C	ND	ND	ND	ND	ND
	Escherichia coli	WT (U1 in vivo) ³⁰	5.20×10^{-7}	2.6	5.00×10^6	358.7%
WT (U1 in vitro) ³⁰		1.50×10^{-7}	0.2	1.33×10^6	95.7%	N/A
G73A ³⁰		2.00×10^{-7}	0.14	7.00×10^5	50.2%	N/A
G73C ³⁰		7.00×10^{-6}	0.03	4.29×10^3	0.3%	N/A
G73U ³⁰		8.00×10^{-6}	0.0068	8.50×10^2	0.06%	N/A
"WT" (U1G, in vitro) ³⁰		6.60×10^{-7}	0.92	1.39×10^6	100.0%	N/A
C72 ³⁰		1.30×10^{-6}	0.17	1.31×10^5	9.4%	N/A
G36A ³⁰		6.00×10^{-6}	0.005	8.33×10^2	0.1%	N/A
G36U ³⁰		5.00×10^{-7}	0.01	2.00×10^4	1.4%	N/A
U1; G36A ³⁰		6.60×10^{-6}	0.0036	5.45×10^2	0.04%	N/A

Sce tRNA^{Q2}-G36C proved to be the poorest substrate out of the mutants tested.

Unaware of any previous G36C cognate tRNA measurements in *Eco*, comparisons had to be made with other mutants at the same position. Not surprisingly, the *Eco* system

apparently strongly discriminates against mutations at this position, especially against G36A; G36U is also a very poor substrate for *Eco* QRS. Subsequent work²¹ with *Eco* tRNA^{Q2}-G36U reaffirms this trend with constants determined under substrate saturation. Finally, the *Sce* tRNA^{Q2}-G1U/C72A double mutant exhibited pronounced decreases in k_{cat}/K_M in both tRNA and glutamine measurements, as shown by elevations in both k_{cat} and K_M^Q .

It has been suggested⁵⁰ that the presence of a pyrimidine following either G1:C72 or C1:G72 in tRNA has a destabilizing effect on the base pair, lowering the energy required to disrupt WC geometry (or perhaps promote subsequent stacking interactions?). Most *Eco* tRNAs have a G1:C72 base pair in tRNA^{Q2}, as does *Sce* tRNA^{Q2}. Unlike *Eco* tRNA^{Q2}, *Sce* tRNA^{Q2} contains U73, a pyrimidine, perhaps to assist the protein in destabilizing the N1:N72 base pair. However, U73C is a poor substrate. This could yet be due to unknown protein:RNA, intra-RNA:RNA (protein mediated?), and/or protein:protein (tRNA mediated?) interactions within the system.

There has been suggestion²⁸ that the shift to U73 in Eukaryotic tRNA^{Qx} is linked to adoption of the N-terminal QRS appendage. The data presented here support that U73 is indeed preferred by *Sce* QRS, but as to whether this is a consequence of evolutionary mutation across the older, more conserved "core" protein, or due to interactions of the NTD with the tRNA core and/or the core protein, remains elusive. It is not known for certain whether analogous decoupling even occurs in the *Sce* QRS system. Revisiting these mutants with a mutant *Sce* QRS lacking its N-terminal domain appendage (Δ NTD-QRS) would potentially provide insight into intra-complex signaling if different trends in k_{cat}/K_M were observed. Additionally, similar kinetic comparisons of WT *Sce* QRS to

Δ NTD-QRS with additional tRNA^{Q2} mutants, such as RNA microhelices or deletion mutants in the 5' stand of the acceptor stem (i.e. Δ N1 to assess importance of N72 identity independent of decoupling). Additional point mutations across the "elbow" of the 3° fold of the tRNA, especially those that would introduce little perturbation to the stacking core but potentially in proximity to the NTD, could possibly uncover additional identity elements.

If anything, it would appear that the *Sce* system is *more* relaxed with regards to discriminator identity (Tables 4 and 5). Averaging the relative $(k_{\text{cat}}/K_{\text{M}})_{\text{mutant}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT}}$ glutamine-derived values for the three discriminator mutants yields a crude composite $k_{\text{cat}}/K_{\text{M}}$ relative to WT of 28.3% vs. 22.3% for *Sce* and *Eco*, respectively. As tRNA-derived steady state values were not determined in this study, comparison with WT literature $k_{\text{cat}}/K_{\text{M}}$ must be made. The WT-tRNA^{Q2} glutamine-derived constants presented herein exhibit a $(k_{\text{cat}}/K_{\text{M}})_{\text{exp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_lit}}$ ratio of 190.8%. Similarly, the glutamine-derived U73A $(k_{\text{cat}}/K_{\text{M}})_{\text{exp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_lit}}$ ratio is 121.5%, but the $(k_{\text{cat}}/K_{\text{M}})_{\text{exp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_exp}}$ ratio is 63.7%. Assuming $(k_{\text{cat}}/K_{\text{M}})_{\text{exp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_lit}}$ and $(k_{\text{cat}}/K_{\text{M}})_{\text{exp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_exp}}$ (the latter of which wasn't determined) discrepancy is similar for tRNA-derived constants, the values in Table 5 (right-most column) can be scaled accordingly. The corrected average of $(k_{\text{cat}}/K_{\text{M}})_{\text{mutant}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_lit}}$ for the *Sce* tRNA^{Q2} discriminator mutants thus becomes 23.7% compared to an *Eco* average of 16.9%. Without discrepancy scaling, the *Sce* tRNA^{Q2} mutants yield an average $(k_{\text{cat}}/K_{\text{M}})_{\text{exp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{WT_lit}}$ of 45.3%.

In light of the fact that *Hsa* tRNA^{Qx} sequences are quite promiscuous (Table 1, Figures S11 and S12) with consensus identities, this seems reasonable, as *Hsa* QRS has

also adopted the eukaryotic N-terminal appendage. Across the three principal species discussed in this document (*Eco*, *Sce*, and *Hsa*), tRNA containing C73 is only known to exist in *Eco* tRNA^H, *Sce* tRNA^P, and *Hsa* tRNA^P species, in addition to few others in *Hsa* (Figures S11 and S12): 2x tRNA^Q (chr3.tRNA9, chr6.tRNA131), 1x tRNA^L (chr10.tRNA3), 1x tRNA^F (chr6.tRNA112), and 1x tRNA^W (chr9.tRNA3). As to whether all, or any, of these gene products are utilized for translation is entirely unknown.

Hsa has an abundance of paralogous tRNA genes (see Introduction, section *i*), as predicted by the tRNAscan-SE⁵¹ analyses deposited in the UCSC Genomic tRNA Database⁴⁶ (<http://gtrnadb.ucsc.edu>). According to the most recently deposited analysis (hg19 - NCBI Build 37.1 Feb 2009), there are 506 "standard" tRNAs (putatively accepting one of the twenty canonical amino acids), three selenocysteine tRNAs, three possible suppressor tRNAs, three "tRNAs" with undetermined/unknown isotype(s), and 110 predicted pseudogenes. Excluding predicted **pseudogenes** (PG), this totals to 515 genes that appear to code for putatively expressed tRNA (or tRNA-like) products. By comparison, *Eco* (K12) encodes for 87 tRNA products (+1 PG) and *Sce* (S288c Apr 2011) encodes 289 (+6 PG) tRNA products.

Additional sets are not necessarily indicative of "higher" evolution or complexity of organism; ploidity may play a role as well. The complement of genomic tRNA products varies wildly: *Leishmania major* (Version 5.0) encodes 82 (+1 PG), *Drosophila melanogaster* (Release 6 Apr 2006) encodes 300 (+4 PG), *Felis catus* (catChrV17e Dec 2008) encodes 3095 (+0 PG), and *Danio rerio* (Zv9 Jul 2010) encodes 12844 (+0 PG). Of note among the species mentioned here, *Hsa* has a disproportionately high number of

predicted pseudogenes. Brief examination of a few other deposited analyses from other simian species (chimpanzee, gorilla, rhesus monkey, baboon, and orangutan) do not indicate the presence of any pseudogenes, but all encode for a similar number of tRNAs (~380-470) relative to *Hsa*.

ii. evolutionary repurposing?

The Search Tool for the Retrieval of Interacting Genes (STRING)⁴⁷ meta-analysis webserver parses the information from a number of principal databases^{47,52-55} in addition to the information stored in numerous other databases, including known associations (MINT,⁵⁶ KEGG,⁵⁷ etc.), text mining on PubMed, and predictive methods. In doing so, a network of protein-protein association predictions is generated, with color-coded links between partner nodes representing the data type used for prediction. The default search includes results related to genomic context, high throughput, conserved expression, and textmining. All associations are assigned a confidence score; the mathematical details of which can be found within the appropriate reference. A conservative results-limited query (excluding coexpression and textmining associations) on the STRING webserver for the *Sce* genes GLN4 (GLN4p = QRS) and Arc1 (Arc1p is the yeast orthologue of AIMP1, a MARS scaffold protein) compared to a similar search for the *Hsa* genes QARS (additionally excluding "neighborhood" results as well, for simplicity) and AIMP1 suggest the nature of interconnectivity between the members of MARS and how they potentially relate to diverse cellular functions (Figure 7, Appendix D). A search against *Hsa* QARS alone with all associations enabled illustrates the complexity of results with textmining and coexpression associations (Figure S13). Not

surprisingly, coexpression data produces a lot of noise, given that QRS is part of the translation apparatus and expression is certainly coregulated along with other translational and transcriptional proteins as a function of cell cycle, DNA damage, etc.

Figure 8. Comparison of *Sce* (left) and *Hsa* (right) QRS and Arc1p or AIMP1 associations, respectively, as predicted by STRING⁴⁷

genetic neighborhood (green), probable gene fusion events (red), cooccurrence (blue), conserved coexpression (black), high-throughput experiments (magenta), databases (cyan), textmining (green), and homology (mauve)

A different data mining webserver, The **B**iological **G**eneral **R**epository for **I**nteraction **D**atasets (BioGRID)⁴⁸ is limited to experimental data (affinity capture, co-fractionation, two-hybrid, etc.) and includes low-throughput studies. A comparison of *Sce* GLN4 vs. *Hsa* QARS results (without altering any settings) also provides a stark depiction of the level of complexity arising in "higher" Eukarya (Figure 8, Appendix D). Certainly both of these servers are biased by the amount of data produced by high-throughput experiments, and likely contain many false associations, due to both empirical

(i.e. database entries) and other incorrect hits. However, there are many intriguing connections amongst the various nodes, hinting at countless unknown cellular subtleties.

Figure 9. Comparison of *Sce* (left) and *Hsa* (right) QRS associations as predicted by BioGRID⁴⁸

Partial details regarding the nature of some of these interacting partners can be found in Table 6. One of the entries, MAP3K5 (i.e. ASK1), had previously been implicated in glutamine-dependent interactions with *Hsa* QRS.⁵⁸ In short, *Hsa* QRS appears to inhibit the kinase activity of ASK1 given sufficient cytoplasmic glutamine concentrations. The same study also implicated that residues 237-775 (the catalytic core and ABD) of *Hsa* QRS were necessary for association with MAP3K5, potentially precluding the N-terminal domain extension in this interaction. Also amongst the predicted STRING associations is TRAF2, another protein involved in cell cycle and apoptosis. Additionally there are predicted associations with transcription factors

Table 6. Further information on predicted association/interaction partners with *Hsa* QRS, as predicted by STRING (top) and BioGRID (bottom)

	Identifier	UniProt Accession	Full Name	Cellular Role, Predicted Function, and/or Comments	Cellular Localization
STRING association predictions	TRAF2	Q12933	TNF receptor-associated factor 2	Regulation of cell survival and apoptosis	Cytoplasm
	MAP3K5	Q99683	Mitogen-activated protein kinase kinase kinase (<i>aka</i> ASK1)	Ser/Thr kinase; apoptosis, signal transduction, immune response	Cytoplasm; endoplasmic reticulum
	GATA5	Q9BWX5	Transcription factor GATA-5	Binds CEF-1; transcriptional program related to smooth muscle	Nucleus (nucleoplasm)
	RSBN1	Q5VWQ0	Round spermatid basic protein 1	Unknown; possible protein binding	Nucleus (predicted)
	D6S51	P48634	Protein PRRC2A (proline-rich coiled-coil 2A)	Unknown; possible regulation of pre-mRNA splicing	Cytoplasm; nucleus
	BAT2	P48634	PRRC2A (may be an isoform of D6S51)	Unknown; possible regulation of pre-mRNA splicing	Cytoplasm; nucleus
	ERCC3	P19447	TFIIH basal transcription factor complex helicase XPB subunit	ATP-dependent 3'-5' DNA helicase; core component of TFIIH; nucleotide excision repair or RNA transcription, dependent on protein binding partner(s)	Nucleus
	ASL	P04424	Argininosuccinate lyase	Urea cycle	Cytoplasm; extra-cellular vesicular exosome
	ENSG-00000231825	P48634	PRRC2A (may be an isoform of D6S51)	Unknown; possible regulation of pre-mRNA splicing	Cytoplasm; nucleus
	HSP90AA1	P07900	Heat shock protein HSP 90-alpha	ATPase; molecular chaperone; cell cycle control, signal transduction	Cytoplasm; cell membrane
BioGRID association predictions	...	...	...	...	...
	UBC	P0CG48	Polyubiquitin-C	Exists as free-form or covalent attachment to protein; DNA damage response; JNK cascade; signaling; activation of protein kinases	Cytoplasm; nucleus
	RNF2	Q99496	E3 ubiquitin-protein ligase RING2 (Huntingtin-interacting protein 2-interacting protein 3)	Mediates monoubiquitination of histone H2A (histone code and gene regulation); essential component of polycomb group multiprotein PRC1-like complex; dosage compensation	Nucleus; chromosome
	RNF11	Q9Y3C5	RING finger protein 11	Component of a ubiquitin-editing complex; inflammation	Cytoplasm; endosome; nucleus
	HDAC5	Q9UQL6	Histone deacetylase 5	Lysine deacetylation of core histones; transcriptional regulation; cell cycle; development & differentiation (from low-throughput dataset)	Nucleus; cytoplasm
	gag	Q71B39	Gag protein (Human immunodeficiency virus 1)	Viral infection and proliferation	N/A
	EBNA-LP	Q8AZK7	Epstein-Barr nuclear antigen leader protein (Human herpesvirus 4)	Assists establishment of B-cell lymphocyte immortalization	N/A

(GATA5 and ERCC3), a urea cycle protein (ASL), and a heat shock protein (HSP90AA1). The entries D6S51, BAT2, and ENSG00000231825 may be redundant, isoforms, or erroneous database entries, as some (but not all) of the various accession numbers have been deleted on some of the databases. The RSBN1 association is amongst the highest scoring, higher than even some of the MARS members, but is a protein of unknown function. Examination of the BioGRID results exhibit similar diversity of nuclear and cytoplasmic interaction partners, including two viral proteins.

It has been proposed that the MARS complex begins in Arthropoda and is limited to Metazoa (i.e. "drosophila to humans") and is possibly linked with multicellularity and the development of more complex anatomies, including development,¹⁴ but examples of simpler complexes (in terms of number observed interaction partners) can be found throughout Opisthokonta. The simplest example of a possible evolutionary MARS precursor analogue can be found in the *Sce* minimal multi-aaRS synthetase complex, which consists of ERS, MRS, and Arc1p, a homolog of the MARS scaffold protein AIMP1. It has also been speculated that adoption of the eukaryotic N-terminal appendage is somehow co-evolutionarily related to the emergence of MARS and/or the recruitment of QRS into the complex.⁵⁹

Putative protein-protein binding motifs in either the *Sce* or *Hsa* QRS N-terminal domains are not known, unlike in several of the other MARS-associated synthetases⁸. Additionally, *Sce* QRS is not known to be a member of the *Sce* ERS:MRS:Arc1p mini-MARS assembly, but it possesses the extension nonetheless. However, a prediction of intrinsic protein disorder (and protein-protein binding that that sometimes imparts) provided by PSIPRED predicts that residues ~175-225 could be involved in protein

binding (Figures S14 and S15). Subjecting the NTD of *Sce* QRS to the same analysis provides similar results (Figure S16). A pull-down study⁶⁰ of an array of Myc epitope-tagged *Hsa* QRS progressive deletion mutants transiently expressed in HEK 293 cell line demonstrate that residues 237-457 (corresponding to N-terminal portion of the catalytic core of the protein) were required for MARS association, albeit at the loss of endogenous (full-length) QRS, arginyl-tRNA synthetase (RRS), and AIMP1 from the complex. It could be that the disordered residues interact with AIMP1 in a fashion such that it adopts the proper conformation required for binding with RRS and another of its interaction partners, AIMP2. However, if the association is limited to the catalytic core, it would not be the first time the catalytically active domain of an *Hsa* protein has apparently adopted expanded function.⁶¹ Of interest, the non-canonical functions assigned to AIMP1 require its dissociation from the MARS complex.⁶²

4. CONCLUSION

There is still much to be learned about the aaRSs and their interacting partners, the AIMP. The MARS interaction network has been studied,⁶³ but at low resolution. Additionally, the MARS complex has been studied with cryo-EM,^{64,65} but not at the high resolution afforded by the newest generation of instrumentation. Undoubtedly a chemical cross linking study coupled with tryptic digest and LC-MS/MS, similar to that performed on transcription initiation complexes,⁶⁶ would provide a wealth of insight. Similar analyses on "lower" Eukarya, such as *Caenorhabditis elegans* which has a MARS assembly of intermediate complexity,⁶⁷ could aid in drug development by targeting the translation machinery in novel ways to defend against Trypanosomes and other problematic Eukaryotic parasites, known to have unique aaRS features and MARS like aspects.⁶⁸⁻⁷⁰

REFERENCES

1. Giegé, R. *et al.* Structure of transfer RNAs: similarity and variability. *Wiley Interdiscip. Rev. RNA* **3**, 37–61
2. Perona, J. J. & Hadd, A. Structural diversity and protein engineering of the aminoacyl-tRNA synthetases. *Biochemistry* **51**, 8705–29 (2012).
3. Ling, J. & Söll, D. Severe oxidative stress induces protein mistranslation through impairment of an aminoacyl-tRNA synthetase editing site. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4028–33 (2010).
4. Geslain, R. & Ribas de Pouplana, L. Regulation of RNA function by aminoacylation and editing? *Trends Genet.* **20**, 604–10 (2004).
5. Giege, R., Sissler, M. & Florentz, C. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **26**, 5017–5035 (1998).
6. Frechin, M. *et al.* Yeast mitochondrial Gln-tRNA(Gln) is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS. *Genes Dev.* **23**, 1119–30 (2009).
7. Guo, M., Yang, X.-L. & Schimmel, P. New functions of aminoacyl-tRNA synthetases beyond translation. *Nat. Rev. Mol. Cell Biol.* **11**, 668–74 (2010).
8. Bullwinkle, T. J. & Ibba, M. Emergence and evolution. *Top. Curr. Chem.* **344**, 43–87 (2014).
9. Brown, M. V, Reader, J. S. & Tzima, E. Mammalian aminoacyl-tRNA synthetases: cell signaling functions of the protein translation machinery. *Vascul. Pharmacol.* **52**, 21–6 (2010).
10. Mirande, M., Le Corre, D. & Waller, J. P. A complex from cultured Chinese hamster ovary cells containing nine aminoacyl-tRNA synthetases. Thermolabile leucyl-tRNA synthetase from the tsH1 mutant cell line is an integral component of this complex. *Eur. J. Biochem.* **147**, 281–9 (1985).
11. Quevillon, S., Robinson, J. C., Berthonneau, E., Siatecka, M. & Mirande, M. Macromolecular assemblage of aminoacyl-tRNA synthetases: identification of protein-protein interactions and characterization of a core protein. *J. Mol. Biol.* **285**, 183–95 (1999).

12. David, A. *et al.* RNA binding targets aminoacyl-tRNA synthetases to translating ribosomes. *J. Biol. Chem.* **286**, 20688–700 (2011).
13. Kim, J. H., Han, J. M. & Kim, S. Protein-Protein Interactions and Multi-component Complexes of Aminoacyl-tRNA Synthetases. *Top. Curr. Chem.* **344**, 119–44 (2014).
14. Park, S. G., Choi, E.-C. & Kim, S. Aminoacyl-tRNA synthetase-interacting multifunctional proteins (AIMPs): a triad for cellular homeostasis. *IUBMB Life* **62**, 296–302 (2010).
15. Ibba, M. *The Aminoacyl-tRNA synthetases*. 73–88 (Landes Bioscience, 2005).
16. Bullock, T. L., Uter, N., Nissan, T. A. & Perona, J. J. Amino acid discrimination by a class I aminoacyl-tRNA synthetase specified by negative determinants. *J. Mol. Biol.* **328**, 395–408 (2003).
17. Rodriguez-Hernandez, A. *et al.* Structural and mechanistic basis for enhanced translational efficiency by 2-thiouridine at the tRNA anticodon wobble position. *J. Mol. Biol.* **425**, 3888–906 (2013).
18. Uter, N. T. & Perona, J. J. Active-site assembly in glutamyl-tRNA synthetase by tRNA-mediated induced fit. *Biochemistry* **45**, 6858–65 (2006).
19. Rould, M. A., Perona, J. J., Söll, D. & Steitz, T. A. Structure of E. coli glutamyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution. *Science* **246**, 1135–42 (1989).
20. Rodríguez-Hernández, A. & Perona, J. J. Heat maps for intramolecular communication in an RNP enzyme encoding glutamine. *Structure* **19**, 386–96 (2011).
21. Uter, N. T. & Perona, J. J. Long-range intramolecular signaling in a tRNA synthetase complex revealed by pre-steady-state kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 14396–401 (2004).
22. Stoldt, M., Wöhnert, J., Görlach, M. & Brown, L. R. The NMR structure of Escherichia coli ribosomal protein L25 shows homology to general stress proteins and glutamyl-tRNA synthetases. *EMBO J.* **17**, 6377–84 (1998).
23. Fersht, A. Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding, by A. Fersht. 1999. New York: Freeman. *Protein Sci.* (2008). doi:10.1110/ps.9.1.207

24. Ibba, M., Hong, K. W., Sherman, J. M., Sever, S. & Söll, D. Interactions between tRNA identity nucleotides and their recognition sites in glutaminyl-tRNA synthetase determine the cognate amino acid affinity of the enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6953–8 (1996).
25. Sherlin, L. D. *et al.* Influence of transfer RNA tertiary structure on aminoacylation efficiency by glutaminyl and cysteinyl-tRNA synthetases. *J. Mol. Biol.* **299**, 431–46 (2000).
26. Ludmerer, S. W. & Schimmel, P. Gene for yeast glutamine tRNA synthetase encodes a large amino-terminal extension and provides a strong confirmation of the signature sequence for a group of the aminoacyl-tRNA synthetases. *J. Biol. Chem.* **262**, 10801–6 (1987).
27. Grant, T. D. *et al.* The structure of yeast glutaminyl-tRNA synthetase and modeling of its interaction with tRNA. *J. Mol. Biol.* **425**, 2480–93 (2013).
28. Grant, T. D. *et al.* Structural conservation of an ancient tRNA sensor in eukaryotic glutaminyl-tRNA synthetase. *Nucleic Acids Res.* **40**, 3723–31 (2012).
29. Hayase, Y. *et al.* Recognition of bases in Escherichia coli tRNA(Gln) by glutaminyl-tRNA synthetase: a complete identity set. *EMBO J.* **11**, 4159–65 (1992).
30. Jahn, M., Rogers, M. J. & Söll, D. Anticodon and acceptor stem nucleotides in tRNA(Gln) are major recognition elements for E. coli glutaminyl-tRNA synthetase. *Nature* **352**, 258–60 (1991).
31. Liao, C.-C., Lin, C.-H., Chen, S.-J. & Wang, C.-C. Trans-kingdom rescue of Gln-tRNA^{Gln} synthesis in yeast cytoplasm and mitochondria. *Nucleic Acids Res.* **40**, 9171–81 (2012).
32. Whelihan, E. F. & Schimmel, P. Rescuing an essential enzyme-RNA complex with a non-essential appended domain. *EMBO J.* **16**, 2968–74 (1997).
33. Wang, C. C., Morales, A. J. & Schimmel, P. Functional redundancy in the nonspecific RNA binding domain of a class I tRNA synthetase. *J. Biol. Chem.* **275**, 17180–6 (2000).
34. Wang, C. C. & Schimmel, P. Species barrier to RNA recognition overcome with nonspecific RNA binding domains. *J. Biol. Chem.* **274**, 16508–12 (1999).
35. Deinert, K., Fasiolo, F., Hurt, E. C. & Simos, G. Arc1p organizes the yeast aminoacyl-tRNA synthetase complex and stabilizes its interaction with the cognate tRNAs. *J. Biol. Chem.* **276**, 6000–8 (2001).

36. Simader, H. *et al.* Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes. *Nucleic Acids Res.* **34**, 3968–79 (2006).
37. Frantz, J. D. & Gilbert, W. A novel yeast gene product, G4p1, with a specific affinity for quadruplex nucleic acids. *J. Biol. Chem.* **270**, 20692–7 (1995).
38. Chang, C.-P. *et al.* Promoting the formation of an active synthetase/tRNA complex by a nonspecific tRNA-binding domain. *J. Biol. Chem.* **283**, 30699–706 (2008).
39. Liu, C. *et al.* Kinetic quality control of anticodon recognition by a eukaryotic aminoacyl-tRNA synthetase. *J. Mol. Biol.* **367**, 1063–78 (2007).
40. Deniziak, M. *et al.* Deinococcus glutaminyl-tRNA synthetase is a chimer between proteins from an ancient and the modern pathways of aminoacyl-tRNA formation. *Nucleic Acids Res.* **35**, 1421–31 (2007).
41. Sherlin, L. D. *et al.* Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA* **7**, 1671–8 (2001).
42. Kibbe, W. A. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* **35**, W43–6 (2007).
43. Roberts, E., Eargle, J., Wright, D. & Luthey-Schulten, Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* **7**, 382 (2006).
44. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–80 (1994).
45. Russell, R. B. & Barton, G. J. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins* **14**, 309–23 (1992).
46. Chan, P. P. & Lowe, T. M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* **37**, D93–7 (2009).
47. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–15 (2013).
48. Chatr-Aryamontri, A. *et al.* The BioGRID interaction database: 2013 update. *Nucleic Acids Res.* **41**, D816–23 (2013).

49. Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349–57 (2013).
50. Lee, C. P., Mandal, N., Dyson, M. R. & RajBhandary, U. L. The discriminator base influences tRNA structure at the end of the acceptor stem and possibly its interaction with proteins. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7149–52 (1993).
51. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–64 (1997).
52. Stark, C. *et al.* The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.* **39**, D698–704 (2011).
53. Kerrien, S. *et al.* The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* **40**, D841–6 (2012).
54. Salwinski, L. *et al.* The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.* **32**, D449–51 (2004).
55. Goll, J. *et al.* MPIDB: the microbial protein interaction database. *Bioinformatics* **24**, 1743–4 (2008).
56. Licata, L. *et al.* MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res.* **40**, D857–61 (2012).
57. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–14 (2012).
58. Ko, Y. G. *et al.* Glutamine-dependent antiapoptotic interaction of human glutaminyl-tRNA synthetase with apoptosis signal-regulating kinase 1. *J. Biol. Chem.* **276**, 6030–6 (2001).
59. Guo, M. & Yang, X.-L. Architecture and metamorphosis. *Top. Curr. Chem.* **344**, 89–118 (2014).
60. Kim, T. *et al.* Catalytic peptide of human glutaminyl-tRNA synthetase is essential for its assembly to the aminoacyl-tRNA synthetase complex. *J. Biol. Chem.* **275**, 21768–72 (2000).
61. Zhou, Q. *et al.* Orthogonal use of a human tRNA synthetase active site to achieve multifunctionality. *Nat. Struct. Mol. Biol.* **17**, 57–61 (2010).

62. Park, S. G., Ewalt, K. L. & Kim, S. Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers. *Trends Biochem. Sci.* **30**, 569–74 (2005).
63. Kaminska, M. *et al.* Dissection of the structural organization of the aminoacyl-tRNA synthetase complex. *J. Biol. Chem.* **284**, 6053–60 (2009).
64. Norcum, M. T. & Boisset, N. Three-dimensional architecture of the eukaryotic multisynthetase complex determined from negatively stained and cryoelectron micrographs. *FEBS Lett.* **512**, 298–302 (2002).
65. Wolfe, C. L., Warrington, J. A., Davis, S., Green, S. & Norcum, M. T. Isolation and characterization of human nuclear and cytosolic multisynthetase complexes and the intracellular distribution of p43/EMAPII. *Protein Sci.* **12**, 2282–90 (2003).
66. Chen, Z. A. *et al.* Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29**, 717–26 (2010).
67. Havrylenko, S., Legouis, R., Negrutskii, B. & Mirande, M. *Caenorhabditis elegans* evolves a new architecture for the multi-aminoacyl-tRNA synthetase complex. *J. Biol. Chem.* **286**, 28476–87 (2011).
68. Gowri, V. S., Ghosh, I., Sharma, A. & Madhubala, R. Unusual domain architecture of aminoacyl tRNA synthetases and their paralogs from *Leishmania major*. *BMC Genomics* **13**, 621 (2012).
69. Rinehart, J., Horn, E. K., Wei, D., Soll, D. & Schneider, A. Non-canonical eukaryotic glutamyl- and glutamyl-tRNA synthetases form mitochondrial aminoacyl-tRNA in *Trypanosoma brucei*. *J. Biol. Chem.* **279**, 1161–6 (2004).
70. Cestari, I. *et al.* A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes. *Mol. Cell. Biol.* **33**, 4872–88 (2013).

APPENDIX A - Supplemental Figures

List of Supplemental Figures

Figure S1. Alignment of eukaryotic (top) and bacterial QRS sequences (bottom), colored by BLOSUM30 scores	49
Figure S2. Agarose gel of template DNA produced with Klenow fragment	50
Figure S3. Purification gel, showing UV shadow	50
Figure S4. Analytical gel of purified tRNA	51
Figure S5. Purification gel after ³² P labeling of tRNA via <i>Eco</i> CCA nucleotidyl-transferase	51
Figure S6. Matrix strategy to ensure saturation	52
Figure S7. Example of experiments to ascertain substrate saturation in <i>Sce</i> tRNA ^{Q2} -U73A	52
Figure S8. Raw data as obtained for <i>Sce</i> tRNA Q2 wild type, with volume quantitation boxes illustrating Gln-AMP:AMP ratio determination & background subtraction	53
Figure S9. Michaelis-Menten fits of <i>Sce</i> tRNA ^{Q2} initial rate data	54
Figure S10. <i>Sce</i> tRNA ^{Qx} sequences, aligned with COVE	55
Figure S11. <i>Hsa</i> tRNA ^{Q1} sequences aligned with COVE	56
Figure S12. <i>Hsa</i> tRNA ^{Q2} sequences aligned with COVE	57
Figure S13. <i>Hsa</i> QARS STRING association predictions	58
Figure S14. <i>Hsa</i> QRS intrinsic disorder profile, as predicted by DISOPRED, residues 1-775	59
Figure S15. <i>Hsa</i> QRS intrinsic disorder profile, as predicted by DISOPRED, residues 1-225	60
Figure S16. <i>Sce</i> QRS intrinsic disorder profile, as predicted by DISOPRED, residues 1-228	61

Figure S1. Alignment of Eukaryotic (top) and Bacterial QRS sequences (bottom), colored by BLOSUM30 scores

Figure S2. Agarose gel of template DNA produced with Klenow fragment

Figure S3. Purification gel, showing UV shadow

Figure S4. Analytical gel of purified tRNA

Figure S5. Purification gel after ^{32}P labeling of tRNA via *Eco* CCA nucleotidyl-transferase

Figure S6. Matrix strategy to ensure saturation

Figure S7. Example of experiments to ascertain substrate saturation

in *Sce* tRNA^{Q2}-U73A

Figure S8. Raw data as obtained for *Sce* tRNA^{Q2} wild type, with volume quantitation boxes illustrating Gln-AMP:AMP ratio determination & background subtraction

20130924 - Sce Q2-WT ss-Gln @ 1 uM tRNA,
TRIPPLICATE MEASURE

y = m1*x/(m2+x)		
	Value	Error
m1	2.8406e-9	3.274e-10
m2	0.0023955	0.00070276
Chisq	6.1469e-20	NA
R ²	0.98059	NA

y = m1*x/(m2+x)		
	Value	Error
m1	2.5278e-9	2.8137e-10
m2	0.0013898	0.00045931
Chisq	9.5353e-20	NA
R ²	0.96722	NA

y = m1*x/(m2+x)		
	Value	Error
m1	2.3485e-9	1.5058e-10
m2	0.0011247	0.00022668
Chisq	3.4311e-20	NA
R ²	0.98574	NA

20130924 - Sce Q2-WT ss-Gln @ 1 uM tRNA,
TRIPPLICATE AVERAGED

y = m1*x/(m2+x)		
	Value	Error
m1	2.5511e-9	1.7955e-10
m2	0.0015645	0.0003148
Chisq	3.3331e-20	NA
R ²	0.98801	NA

Figure S9. Michaelis-Menten fits of *Sce* tRNA^{Q2} initial rate data

Gln tRNAs

[illegible]

Figure S10. *See* tRNA^{Qx} sequences, aligned with COVE

Figure S11. *Hsa* tRNA^{Q1} sequences aligned with COVE

Gln tRNAs

Figure S12. *Hsa* tRNA^{Q2} sequences aligned with COVE

Figure S13. *Hsa* QARS STRING association predictions

Figure S14. *Hsa* QRS intrinsic disorder profile, as predicted by DISOPRED, residues 1-775

Figure S15. *Hsa* QRS intrinsic disorder profile, as predicted by DISOPRED, residues 1-225

Figure S16. *Sce* QRS intrinsic disorder profile, as predicted by DISOPRED, residues 1-228

APPENDIX B - DNA oligomers used in synthesis of *Sce* tRNA^{Q2} mutants

WT sequence (without CCA)

>*Sce*_chr13.trna8-GlnCTG (808316-808245) 72 bp Sc: 64.54

GGTCCTATAGTGTAGTGGTTATCACTTTCGGTTCTGATCCGAACAACCCAGTTCGAATCCG
GGTGGGACCT 72 nt

U73A

Promoter & 3' CCA sequence appended; incorporate U73A mutation:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTCGGTTC
TGATCCGAACAACCCAGTTCGAATCCGGTGGGACCACCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGTGGTCCCACCCGGATTCTGAAGTGGGGTTGTTCCGATCAGAACCGAAAGTGATAA

U73G

Promoter & 3' CCA sequence appended; incorporate U73G mutation:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTCGGTTC
TGATCCGAACAACCCAGTTCGAATCCGGTGGGACCGCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGCGGTCCCACCCGGATTCTGAAGTGGGGTTGTTCCGATCAGAACCGAAAGTGATAA

U73C

Promoter & 3' CCA sequence appended; incorporate U73C mutation:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTCGGTTC
TGATCCGAACAACCCAGTTCGAATCCGGTGGGACCCCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGCGGTCCCACCCGGATTCTGAAGTGGGGTTGTTCCGATCAGAACCGAAAGTGATAA

G1U/C72A

5' Promoter & 3' CCA sequences appended; incorporate G1U/C72A mutations:

AATTCCTGCAGTAATACGACTCACTATAGTGCCTATAGTGTAGTGGTTATCACTTTCGGTTC
TGATCCGAACAACCCAGTTCGAATCCGGTGGGACATCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGTGCCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGATGTCCCACCCGGATTCTGAAGTGGGGTTGTTCCGATCAGAACCGAAAGTGATAA

G36C

Promoter & 3' CCA sequence appended; incorporate G36C mutation:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTTCGGTTTCATCCGAACAACCCAGTTCGAATCCGGGTGGGACCTCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGAGGTCCCACCCGGATTCTGAAGTGGGGTTGTTCCGGATCAGAACCGAAAGTGATAA

G2A/C71U

5' Promoter & 3' CCA sequences appended; incorporate G2A/C71U mutations:

AATTCCTGCAGTAATACGACTCACTATAGATCCTATAGTGTAGTGGTTATCACTTTTCGGTTTCATCCGAACAACCCAGTTCGAATCCGGGTGGGATCTCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGATCCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGAGATCCCACCCGGATTCTGAAGTGGGGTTGTTCCGGATCAGAACCGAAAGTGATAA

U3G/A70C

5' Promoter & 3' CCA sequences appended; incorporate U3G/A70C mutations:

AATTCCTGCAGTAATACGACTCACTATAGGCGCTATAGTGTAGTGGTTATCACTTTTCGGTTCTGATCCGAACAACCCAGTTCGAATCCGGGTGGGCCCTCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGGCGCTATAGTGTAGTGGTTATCACTTTC

Final Reverse:

mUmGGAGGCGCCCACCCGGATTCTGAAGTGGGGTTGTTCCGGATCAGAACCGAAAGTGATAA

G10A/C25U

5' Promoter & 3' CCA sequences appended; incorporate G10A/C25U mutations:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAATGTAGTGGTTATCAITTTTCGGTTTCATCCGAACAACCCAGTTCGAATCCGGGTGGGACCTCCA 103 nt

Final Forward:

AATTCCTGCAGTAATACGACTCACTATAGGTCCTATAATGTAGTGGTTATCAITTTTC

Final Reverse:

mUmGGAGGTCCCACCCGGATTCTGAAGTGGGGTTGTTCCGGATCAGAACCGAAAATGATAA

APPENDIX C - Michaelis-Menten Fits of mutant *Sce* tRNA^{Q2} initial rates

i) U73A - triplicate K_M^{Gln} measurement

Q2-U73A		[Gln] (M)								Km-Q (M)		Vmax		Error		R ²		Kcat (1/sec)		Kcat/Km (1/M*sec)	
		5.00E-03	1.00E-02	1.50E-02	2.00E-02	2.50E-02	3.50E-02	5.00E-02	6.00E-02		Error	Vmax	Error	R ²							
2014.03.06		4.628E-08	6.366E-08	8.114E-08	1.115E-07	1.219E-07		0.000000159	1.61E-07	0.023361	0.0033926	2.29E-07	1.45E-08	0.98343	22.90	9.80E+02					
2014.03.11		2.77E-09	5.4044E-09	8.6043E-09	1.2494E-08			1.5489E-08	1.5063E-08	0.023384	0.0084143	2.22E-08	3.44E-09	0.935	22.22	9.50E+02					
2014.03.13		2.4969E-08	3.6423E-08	5.182E-08		6.7075E-08	7.391E-08	8.645E-08	9.7571E-08	0.02227	0.0023074	1.20E-07	5.09E-09	0.99253	23.95	1.08E+03					
MEAN										0.023005	0.0047014	1.24E-07	7.69E-09	0.97032	23.03	1.00E+03					
STD DEV										0.00063663	0.0032601	1.03E-07	5.99E-09	0.0309246	0.87	6.54E+01					
%RSD										2.8%	69.3%	83.7%	77.9%	3.2%	3.8%	6.5%

ii) U73A - K_M^{tRNA} measurements

[tRNA] (M)	Q2-U73A					
	2014.03.24	2014.03.26.1	2014.03.26.2	2014.03.26.3	MEAN	STD DEV
1.00E-07	2.9482E-10	5.8914E-10	5.237E-10	8.85E-10		
1.00E-06	1.9829E-09	4.7586E-09	6.2057E-09	2.2483E-09		
5.00E-06	4.3043E-09	1.3364E-08		1.5229E-08		
7.50E-06	4.5214E-09	8.9912E-09	9.1929E-09	1.2207E-08		
1.50E-05	4.1829E-09	1.3527E-08	1.5175E-08	2.2004E-08		
2.50E-05	5.2221E-09					
K_M^{tRNA} (M)	1.48E-06	1.74E-06	1.88E-06	9.34E-06	3.61E-06	3.82E-06
Error	4.60E-07	1.42E-06	1.75E-06	7.40E-06	2.76E-06	
V_{max} (M/sec)	5.24E-09	1.45E-08	1.48E-08	3.46E-08	1.73E-08	1.24E-08
Error	3.18E-10	1.42E-08	1.75E-08	1.37E-08	7.96E-07	
R^2	0.970	0.879	0.890	0.930	0.918	0.041
k_{cat} (1/sec)	10.48	19.30	19.78	46.13	23.92	15.41
k_{cat}/K_M (1/M*sec)	7.1E+06	1.1E+07	1.0E+07	4.9E+06	8.4E+06	2.9E+06
						34.6%

iv) U73G - K_M^{tRNA} measurements

Q2-U73G		[U73G]											
		2013.12.13.1	2013.12.13.2	2013.12.13.3	2013.12.13.4	2014.01.23.1	2014.01.23.2	2014.01.23.3	2014.01.23.4	MEAN	STD DEV	%RSD	
[tRNA] (M)	1.00E-06	1.056E-09	3.8152E-09	3.9E-09	7.00E-09	6.5229E-10	2.2857E-10	9.4857E-10	3.1486E-10				
	1.00E-05	2.3522E-09	8.44E-09		2.5053E-08	3.1486E-09	1.8649E-08	2.5263E-08	1.9986E-08				
	2.00E-05	2.1905E-09	1.3495E-08		1.6762E-08	2.3878E-08	1.8663E-08	2.6606E-08	3.49E-08				
	4.00E-05	2.2629E-09	9.3752E-09	2.1749E-08	1.6745E-08	1.14E-09	1.0474E-08	2.2003E-08	3.8383E-08				
K_M^{tRNA} (M)		1.18E-06	2.09E-06	8.58E-06	JUNK	3.92E-07	2.82E-06	3.77E-06	1.59E-05	4.97E-06	5.41E-06	108.9%	
Error		1.22E-10	2.30E-06	6.37E-06	JUNK	3.71E-06	5.87E-08	4.54E-06	8.70E-06	4.50E-06			
V_{max} (M/sec)		2.41E-09	1.17E-06	2.55E-08	JUNK	1.18E-09	1.76E-08	2.89E-08	5.61E-08	2.05E-08	1.90E-08	92.6%	
Error		1.22E-10	2.26E-09	4.74E-09	JUNK	1.25E-09	6.90E-09	7.21E-09	1.25E-08	4.99E-09			
R^2		0.951	0.726	0.981	JUNK	0.013	0.610	0.840	0.969	0.727	0.409	56.3%	
k_{cat} (1/sec)		0.24	1.17	2.55		0.12	1.76	2.89	5.61	2.05	1.89	92.4%	
k_{cat}/K_M (1/M*sec)		2.0E+05	5.6E+05	3.0E+05		3.0E+05	6.2E+05	7.7E+05	3.5E+05	4.4E+05	2.1E+05	46.8%

v) U73C - K_M^{Gln} measurements

Q2-U73C		[Gln] (M)		Km-Q (M)		Error	Vmax	Error	R ²	Kcat (1/sec)	Kcat/Km (1/M*sec)
		1.00E-03	1.00E-02	2.00E-02	5.00E-02	JUNK	JUNK	JUNK	JUNK		
2013.12.13.a		2.419E-10	3.030E-10	1.659E-10	1.8687E-10	0.025986	6.03E-09	0.019794	0.89365	0.60	2.32E+01
2013.12.13b		8.8147E-10	1.8314E-09	2.2983E-09	4.082E-09	0.00593	3.42E-09	8.73E-11	0.99824	0.34	5.77E+01
2013.12.13c		4.2533E-10	2.181E-09	2.6581E-09	3.0295E-09	0.063886	1.02E-08	0.028726	0.98639	1.02	1.59E+01
2013.12.13d		4.8343E-10	1.2624E-09	2.4743E-09	4.4552E-09	0.0006526	5.15E-10	3.41E-11	0.96891	0.05	7.90E+02
2014.01.23a		1.0566E-09	4.88E-10	4.7257E-10	5.7629E-10	0.0025162	4.25E-09	0.0011344	0.95219	0.42	1.69E+02
2014.01.23b		9.9543E-10	4.25E-09	3.0543E-09	3.8971E-09	0.0021452	4.11E-09	7.00E-10	0.80062	0.41	1.92E+02
2014.01.23c		1.0126E-09	3.5614E-09	4.5143E-09	5.0743E-09	0.0052019	5.60E-09	1.74E-10	0.99671	0.56	1.08E+02
MEAN		0.01510437	0.007598781	4.87E-09	9.11E-10	0.042387143	0.49	1.93E+02	0.49	1.93E+02	
STD DEV		0.02214397	0.01147197	3.22E-09	1.10E-09	0.339823776	0.29	2.72E+02	0.29	2.72E+02	
%RSD		146.6%	146.7%	66.2%	121.1%	36.1%	60.4%	140.3%	60.4%	140.3%

vi) $U73C - K_M^{tRNA}$ measurements

Q2-U73C		[5f3d]											MEAN	STD DEV	%RSD
		2013.12.13.1	2013.12.13.2	2013.12.13.3	2013.12.13.4	2014.01.23.1	2014.01.23.2	2014.01.23.3	2014.01.23.4	2014.01.23.5					
[RNA] (M)	1.00E-06	2.419E-10	3.0305E-10	1.659E-10	1.8667E-10	4.8743E-10	4.88E-10	4.7257E-10	5.7629E-10						
	1.00E-05	8.8147E-10	1.8314E-09	2.2993E-09	4.082E-09	1.0566E-09	3.4611E-09	4.1271E-09	3.6971E-09						
	2.00E-05	4.2533E-10	2.181E-09	2.6581E-09	3.0295E-09	9.9543E-10	4.25E-09	3.0543E-09	3.92E-09						
	4.00E-05	4.8343E-10	1.2624E-09	2.4743E-09	4.4552E-09	1.0126E-09	3.5614E-09	4.5143E-09	5.0743E-09						
	K _m -[RNA] (M)	9.61E-07	2.40E-06	4.78E-06	5.03E-06	1.13E-06	3.90E-06	4.23E-06	6.95E-06	3.67E-06	2.05E-06		55.9%		
Error	1.99E-06	3.74E-06	3.69E-06	6.33E-06	3.37E-07	3.19E-06	4.88E-06	2.90E-06	3.38E-06						
	V _{max} (M/sec)	6.12E-10	1.94E-09	3.07E-09	4.81E-09	1.08E-09	4.49E-09	4.73E-09	5.78E-09	3.31E-09	1.92E-09		58.1%		
	Error	1.94E-10	5.42E-10	5.42E-10	1.42E-09	5.17E-11	7.79E-10	1.18E-08	6.61E-10	6.72E-10					
	R ²	0.322	0.670	0.929	0.835	0.954	0.909	0.838	0.976	0.804	0.217		27.0%		
	k _{cat} (1/sec)	0.06	0.19	0.31	0.48	0.11	0.45	0.47	0.58	0.33	0.19		58.1%		
k _{cat} /K _m (1/M·sec)	6.2E+04	8.1E+04	6.4E+04	9.6E+04	9.6E+04	1.2E+05	1.1E+05	8.3E+04	8.9E+04	2.0E+04		22.3%

vii) $G1U/C72A - K_M^{Gln}$ measurements

		$[Gln] (M)$										$Kcat (1/sec)$	$Kcat/Km (1/M^*sec)$
		5.00E-03	7.50E-03	1.00E-02	2.00E-02	3.00E-02	5.00E-02	6.00E-02	7.50E-02	1.00E-01	1.25E-01		
Q2-G1U/C72A	C	3.7660E-09	4.4308E-09	6.2629E-09	6.0280E-09	6.0280E-09	6.0280E-09	6.0280E-09	6.0280E-09	6.0280E-09	6.0280E-09	0.77	2.56E+01
	C'	3.127E-09	3.7957E-09	4.61E-09	5.8529E-09	5.8414E-09	1.0191E-08	1.0191E-08	1.0191E-08	1.0191E-08	1.0191E-08	0.61	3.13E+01
	C''	3.068E-09	3.7209E-09	5.0051E-09	6.0014E-09	5.42E-09	1.0103E-08	1.0103E-08	1.0103E-08	1.0103E-08	1.0103E-08	0.57	3.38E+01
	MEAN	3.0975E-09	3.7411E-09	4.6819E-09	6.03907E-09	5.76333E-09	1.0431E-08	1.0431E-08	1.0431E-08	1.0431E-08	1.0431E-08	0.63	3.10E+01
	STD DEV	4.1719E-11	2.3318E-11	2.9392E-10	2.07579E-10	3.1172E-10	4.9444E-10	4.9444E-10	4.9444E-10	4.9444E-10	4.9444E-10	0.00	#DIV/0!
	%RSD	1.3%	0.6%	6.3%	3.4%	5.4%	4.7%	4.7%	4.7%	4.7%	4.7%	0.00	#DIV/0!
	CUT AVG											0.69	3.31E+01
												30 mM Gln timecourses discarded

viii) G1U/C72A - K_M^{tRNA} measurements

Q2-G1U/C72A		Trial			MEAN	STD DEV	%RSD
		1	2	3			
[tRNA] (M)	1.00E-07	3.0943E-10	3.2371E-10	3.7857E-10			
	1.00E-06	6.0857E-09	5.7829E-09	5.74E-09			
	2.00E-06	6.06E-09	1.3606E-08	1.3686E-08			
	5.00E-06	1.3391E-08	2.2157E-08	2.13E-08			
	7.50E-06	1.694E-08	3.7046E-08	3.6794E-08			
	1.50E-05	1.2043E-08	2.3026E-08	2.3469E-08			
Km-tRNA (M)		1.99E-06	2.79E-06	2.89E-06	2.56E-06	4.90E-07	19.2%
Error		1.40E-06	2.55E-06	2.58E-06	2.18E-06		
Vmax (M/sec)		1.70E-08	3.57E-08	3.59E-08	2.95E-08	1.09E-08	36.9%
Error		3.47E-09	1.08E-08	1.08E-08	8.34E-09		
R^2		0.842	0.786	0.794	0.808	0.030	3.7%
kcat (1/sec)		0.85	1.79	1.80	1.48	0.54	36.9%
kcat/Km (1/M*sec)		4.3E+05	6.4E+05	6.2E+05	5.6E+05	1.2E+05	21.2%

viii) G1U/C72A - K_M^{tRNA} measurements (cont.)

Q2-G1U/C72A		Trial			MEAN	STD DEV	%RSD
		1	2	3			
[tRNA] (M)	1.00E-07	3.0943E-10	3.2371E-10	3.7857E-10			
	1.00E-06	6.0857E-09	5.7829E-09	5.74E-09			
	2.00E-06	6.06E-09	1.3606E-08	1.3686E-08			
	5.00E-06	1.3391E-08	2.2157E-08	2.13E-08			
	7.50E-06	1.694E-08	3.7046E-08	3.6794E-08			
	1.50E-05	1.2043E-08	2.3026E-08	2.3469E-08			
Km-tRNA (M)		1.99E-06	2.79E-06	2.89E-06	2.56E-06	4.90E-07	19.2%
Error		1.40E-06	2.55E-06	2.58E-06	2.18E-06		
Vmax (M/sec)		1.70E-08	3.57E-08	3.59E-08	2.95E-08	1.09E-08	36.9%
Error		3.47E-09	1.08E-08	1.08E-08	8.34E-09		
R^2		0.842	0.786	0.794	0.808	0.030	3.7%
kcat (1/sec)		0.85	1.79	1.80	1.48	0.54	36.9%
kcat/Km (1/M*sec)		4.3E+05	6.4E+05	6.2E+05	5.6E+05	1.2E+05	21.2%

APPENDIX D - Supplemental Data (SD) file descriptions

SD1 (143 KB) - FASTA sequences of Eukaryotic and Bacterial QRS sequences used for Figure S1; can be opened with text editor

SD2 (15 KB) - FASTA sequences of *Eco* tRNA; can be opened with text editor

SD3 (51 KB) - FASTA sequences of *Sce* tRNA; can be opened with text editor

SD4 (99 KB) - FASTA sequences of *Hsa* tRNA; can be opened with text editor

SD5 (289,388 KB) - Zip file of raw kinetic data (.gel); requires ImageQuant to open

SD6 (68 KB) - Spreadsheet (.xml) legend to eukaryotic QRS sequences in SD1; requires a spreadsheet application to open