Supporting Information

O'Donoghue et al. 10.1073/pnas.1117294108

SI Methods

Bacterial Strains. *Escherichia coli* BL21/DE3 with pSJS1240 [encoding rare tRNA genes that increases the translation efficiency of *Methanothermobacter thermautotrophicus* glutamyl-tRNA synthetase (GluRS) in *E. coli* ; ref. 1], which is maintained with 50 μg/mL spectinomycin, was the genetic background for protein expression. The pTYB1-GluRS plasmid was maintained with 100 μg/mL ampicillin for protein production.

The tRNA genes ($tRNA^{Glu}_{UUC}$, $tRNA^{Gln1}_{UUG}$, $tRNA^{Gln2}_{CUG}$ from *M. thermautotrophicus*) were previously cloned into constructs for in vitro transcription of the tRNA in a pUC18 background (2). To enhance in vitro transcription efficiency of the two $tRNA^{Gln}$ species, their constructs include a self-cleavable hammer head ribozyme directly upstream of the tRNA, which was implemented as detailed previously (3).

All mutant tRNAs and GluRS variants were produced with the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies).

Protein Purification. *E. coli* strains harboring the pTYB1-GluRS variants were inoculated (10 mL preculture) into 1 L batches of standard LB medium supplemented at 1× final concentrations with 20× NPS [0.5 M $(NH_4)_2SO_4$, 1 M KH_2PO_4, 1 M Na_2HPO_4] and 50 × 5,052 (25% glycerol, 2.5% glucose, and 10% α-lactose monohydrate) solutions for auto-induction of protein synthesis (4) and grown 24 h with shaking at 37 °C.

Harvested cells were resuspended in chitin-column buffer (20 mM Hepes-KOH, pH 7.2, 500 nM NaCl, 1 mM EDTA, 0.1% Triton X-100, 20 μM PMSF), and subject to sonication (60% amplitude, 0.5 s pulse, 0.5 s pause, 5 × 1 min) on ice. Cleared lysate was produced by centrifugation of the cell lysate at 13,000 × *g* for 1 h at 4 °C. Column chromatography was performed under gravity flow at room temperature. Columns were filled with 50 mL of chitin beads suspended in 20% ethanol (New England Biolabs) and equilibrated with 10 column volumes of chitin-column buffer before the cleared lysate was applied. The column was washed with 10 additional volumes of column buffer, flushed with 3 volumes of cleavage buffer (20 mM Tris·HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA, 50 mM DTT), and incubated overnight at 30 °C to stimulate cleavage of the inteine tag. The column was flushed with elution buffer (20 mM Hepes-KOH, pH 7.2, 20 mM KCl, 0.2 mM EDTA, 5 mM ß-mercaptoethanol). Fractions that were greater than 95% pure as estimated by SDS-PAGE were pooled together. These pooled fractions were concentrated with an Amicon Ultra 30 K molecular weight cutoff centrifugal filters (Millipore), and the protein solution dialyzed overnight into storage buffer (25 mM Hepes-KOH, pH 7.2, 500 mM NaCl, 1 mM DTT, 50% glycerol). No 17-kDa contaminant (observed in ref. 5) was detected in the pooled fractions.

tRNA Transcription and Purification. In vitro transcription of tRNA was performed following standard procedures (6). Template plasmids containing tRNA genes were purified with the plasmid maxi kit (Qiagen), and 100 μg of plasmid was digested with BstNI (New England Biolabs). The BstNI digested template DNA was purified by phenol chloroform extraction, followed by ethanol precipitation and resolved in double distilled water. A His-tagged T7 RNA polymerase was purified over column of Ni/nitrilotriacetate agarose beads according to manufacturer's instructions (Qiagen). Transcription reactions [40 mM Tris·HCl, pH 8.0; 4 mM each of UTP, CTP, GTP, and ATP at pH 7.0; 22 mM $MgCl_2$; 2 mM spermidine; 10 mM DTT; 6 μg pyrophosphatase (Roche); 60 μg/mL BstNI digested DNA template, approximately 0.2 mg/mL T7 RNA polymerase] were performed in 1 mL reaction volumes for 6 h at 37 °C. The tRNAs were purified on 12% denaturing polyacrylamide gel containing 8 M urea and 1 × TBE buffer (90 mM Tris, 90 mM boric acid, 2 mM EDTA). UV shadowing illuminates the pure tRNA band, which is excised and extracted three times with 1 M sodium acetate pH 5.3 at 4 °C. The tRNA extractions are then ethanol precipitated, resolved in RNase-free distilled water, pooled, and finally desalted using a Biospin 30 column (BioRad).

tRNA Folding and ^{32}P Labeling. tRNAs were refolded by heating to 100 °C for 5 min and slow cooling to room temperature. At 65 °C, $MgCl_2$ was added to a final concentration of 10 mM to aid folding. Folded tRNA was then stored at −20 °C. Freshly folded tRNA was directly used for ^{32}P labeling of the terminal adenosine of the tRNA with the CCA adding enzyme and [α-^{32}P]-labeled ATP (PerkinElmer), which was performed as described (7).

Aminoacylation Assay. Formation of Glu-tRNA was monitored with the Wolfson assay (8). The aminoacylation reaction contains the following components: 1× buffer [50 mM Hepes-KOH (pH 7.2), 25 mM KCl, 10 mM $MgCl_2$, 5 mM DTT], 10 mM ATP (pH 7.0), 27 μg/mL pyrophosphatase (Roche Applied Science), 50 mM L-Glu. The saturating concentrations of ATP and Glu (10 and 50 mM, respectively) for the aminoacylation reaction are exactly as those used and determined previously (9).

All plateau tRNA aminoacylation levels were determined in a 40 min time course at 37 °C according to the reactions conditions described above with 1 μM enzyme, 10 μM tRNA plus 400 nM ^{32}P-labeled tRNA. All aminoacylation reactions were carried out in a 15 μL volume in 1.5 mL eppendorf tubes, and the reaction mixture was preincubated at 37 °C for 15 s before adding enzyme from a 10× stock to initiate the time course. Time points were taken each minute by removing 2 μL aliquots from the reaction and immediately quenching the reaction into an ice-cold 3 μL quench solution [0.66 μg/μL nuclease P_1 (Sigma) in 100 mM sodium citrate, pH 5.49]. For each reaction, 2 μL of blank reaction mixture (containing no enzyme) was added to the quench solution as the $t = 0$ s time point.

The nuclease P_1 mixture, now containing unaminoacylated tRNA and glutamyl-tRNA, was then incubated a room temperature for 35 min and 2 μL aliquots were spotted on polyethyleneimine-cellulose thin layer chromatography plates (EMD Chemicals) and developed in running buffer containing 5% acetic acid and 100 mM ammonium acetate, as in ref. 10. During development radioactive spots for AMP and Glu-AMP (representing free tRNA and Glu-tRNA, respectively) were separated and then visualized and quantified by phosphorimaging. The ratio of aminoacylated tRNA to total tRNA was determined to monitor reaction progress.

Determination of Enzyme Kinetics. All reactions included 20 nM ^{32}P-labeled tRNA, whereas the unlabeled tRNA and enzyme concentrations were adjusted for each enzyme tRNA pair such that in a 5 min time course (with 1 min time points) the reaction progressed linearly and less than 10% of the total tRNA became aminoacylated. For the WT-GluRS 2 nM of enzyme was used and the $tRNA^{Gln2}$ concentration was varied over seven different tRNA concentrations from 0.02 to 10.02 μM. For the all remaining enzyme and tRNA pairs, kinetics were determined with the

following optimal conditions: WT-GluRS 10 nM, $tRNA^{Glu}$ 0.02–10.02 μM; WT-GluRS 100 nM, $tRNA^{Gln1}$ 0.02–10.02 μM; WT-GluRS 50 nM, $tRNA^{Gln1}$ $G_1:C_{72}$ 0.17–20.04 μM; WT-GluRS 5 nM, $tRNA^{Gln1}$ $U_1:A_{72}$ 0.04–20.04 μM; anticodon loop (AL) 3-GluRS 2 nM, $tRNA^{Gln2}$ 0.18–10.02 μM; AL3-GluRS 10 nM, $tRNA^{Glu}$ 0.15–10.02 μM; acceptor stem loop (ASL) 1-GluRS 20 nM, $tRNA^{Gln2}$ 0.18–10.02 μM; ASL1-GluRS 500 nM, $tRNA^{Glu}$ 0.14–40.04 μM; WT-GluRS 150 nM, $tRNA^{Gln1}$ $G_1:C_{72}$ 0.17–53.04 μM; WT-GluRS 20 nM, $tRNA^{Gln1}$ $U_1:A_{72}$ 0.04–20.04 μM; ASL2-GluRS 570 nM, $tRNA^{Gln2}$ 0.27–37.33 μM; ASL2-GluRS 1,140 nM, $tRNA^{Glu}$ 0.15–10.02 μM; WT-GluRS 1 μM, $tRNA^{Gln1}$ $G_1:C_{72}$ 0.14–42.38 μM; WT-GluRS 570 nM, $tRNA^{Gln1}$ $U_1:A_{72}$ 0.27–40.02 μM.

Kinetic constants were derived from plotting initial velocity of a series of aminoacylation reactions that contained varied amounts of unlabeled tRNA substrate. The velocity versus [tRNA] plots were fitted to the standard Michaelis–Menten curve using Kaleida Graph 3.1 (Synergy Software).

Phylogeny and Bioinformatics. All annotated $tRNA^{Gln}$ and $tRNA^{Glu}$ sequences were used in the phylogenetic calculations, but a number of potential tRNA pseudogenes, which did not group among the typical $tRNA^{Gln}$ or $tRNA^{Glu}$ sequences, were removed from the analysis [i.e., with transfer RNA database (tRNAdb) ID codes: tdbD00008758, tdbD00008618, tdbD00002257, tdbD00002382, tdbD00002406, tdbD00004806, tdbD00004808, tdbD00000524, tdbD00004637, tdbD00004755, tdbD00008570, tdbD00004908, tdbD00004902, tdbD00000642, tdbD00008587, tdbD00008601, tdbD00008598, tdbD00008600, tdbD00008599, tdbD00008795, tdbD00008796, tdbD00008797]. A third *M. thermautotrophicus* $tRNA^{Gln}$ isoacceptor is annotated in the tRNA database (11) (tRNAdb identifier tdbD00008570), but inspection of the genomic context and sequence of the tRNA body indicate this to be a $tRNA^{Pro}$ that could result from a different intron splicing than the annotated version. In total, 735 $tRNA^{Glu}$ and $tRNA^{Gln}$ sequences were included. Representative sets (determined as previously and implemented in Multiseq 2.0; refs. 12–14) of $tRNA^{Cys}$, $tRNA^{Lys}$, $tRNA^{Met}$ sequences, each containing approximately 30 tRNA sequences, were used to generate independent root points in the $tRNA^{Glu}$ and $tRNA^{Gln}$ phylogeny. Unfortunately, the procedure produced inconsistent root points that were not statistically supported. For this reason, only unrooted tRNA phylogenies are presented in Fig 5*B* and Dataset S1.

Phylogenetic calculations were performed with PHYML version 3.0.1 (15), with the following specifications: A GTR substitution model, all rate parameters optimized by maximum likelihood analysis, four rate categories, and initial trees were generated by maximum likelihood subtree pruning and rebranching search algorithm and optimized with a nearest neighbor interchange tree search. The bootstrap values were generated for the resulting phylogenetic trees by the Shimodaira–Hasegawa method implemented in PHYML (with option -b -4).

Sequence similarity (shown in Fig. S1) was calculated based on a Blosum60 matrix as implemented in Multiseq 2.0 (13). Sequences were downloaded from the integrated microbial genomes system (16) and structures from the protein databank (17).

Multiseq 2.0 (13) and Visual Molecular Dynamics (VMD) 1.8.6 (18) were used for GluRS and glutaminyl-tRNA synthetase (GlnRS) sequence and structure alignments and graphic rendering (Figs. 3 and 4, and Fig. S1).

SI Results.

Behavior of GluRS Variants Toward Acceptor Stem tRNA Mutants. Given that the first base pair could be a critical element in $tRNA^{Gln}$ versus $tRNA^{Glu}$ discrimination, we constructed $tRNA^{Gln2}$ mutants in which the wild-type $A_1:U_{72}$ base pair is mutated to $G_1:C_{72}$ (as in *M. thermautotrophicus* $tRNA^{Glu}$) and to $U_1:A_{72}$ as in *E. coli* $tRNA^{Gln}$.

WT and ASL-GluRSs all showed reduced activity toward the $tRNA^{Gln2}$ $U_1:A_{72}$ mutant (Table S1). For the $tRNA^{Gln2}$ $G_1:C_{72}$ mutant, we hypothesized that the GC pair would not open as easily as an UA pair, and might lead to a greater reduction in aminoacylation efficiency by the ASL1 and ASL2 variants compared to WT-GluRS. The $G_1:C_{72}$ base pair reduces catalytic efficiency for the ASL-GluRS variants, but also for the WT-GluRS (Table S1 and Fig. S3). Whereas the WT-GluRS shows a fivefold reduced k_{cat} for the $tRNA^{Gln2}$ $G_1:C_{72}$ substrate and the K_M is increased fourfold, the ASL-GluRS variants experience a greater change in K_M (up to 33.6 μM for ALS2) compared to no change in k_{cat} for $tRNA^{Gln2}$ $G_1:C_{72}$ versus $tRNA^{Gln2}$.

tRNA Phylogeny. A remarkable feature of the tRNA tree (Fig. 5*B*, Dataset S1) is that the $tRNA^{Gln}$ and $tRNA^{Glu}$ species from archaea are highly similar to one another, and are, in fact, more similar than they are to their closely related eukaryotic counterparts. Given that the $tRNA^{Glu}$ and $tRNA^{Gln}$ in archaea are under consistent evolutionary pressure to serve as substrates to the same nondiscriminating (ND)-GluRS enzyme, it is possible that evolutionary convergence is in part responsible for the high similarity of these archaeal tRNA species. Whatever the mechanism, the archaeal $tRNA^{Glu}$ and $tRNA^{Gln}$ sequences occupy shorter branches on the tree and appear to have evolved at a slower rate than the bacterial and eukaryotic tRNAs. The observation is further supported by the fact that the archaea (lacking GlnRS and encoding an ND-GluRS and the Glu-$tRNA^{Glu}$ amidotransferase) still retain the pathway for Gln-$tRNA^{Gln}$ formation that was utilized by the last common ancestor (19, 20).

The higher evolutionary tempo in the bacterial $tRNA^{Glu}$ and $tRNA^{Gln}$ and the eukaryotic $tRNA^{Glu}$ is evident from the long branch lengths leading to these taxa compared to those seen among archaeal tRNAs (Fig. 5*B*). The increased evolutionary rate is a potential imprint of adaptation in eukaryotes and bacteria to the introduction of GlnRS that would naturally be absent from archaea.

Dataset S1: Phylogeny of $tRNA^{Glu}$ and $tRNA^{Gln}$ Sequences

Taxa are labeled according to the unique accession code from the tRNAdb (11), and all bootstrap values (on a scale from 0.0 to 1.0) are included as node data. The tree, which is in Newick (.tre) file format, can be viewed with a number of free software programs, including TreeView X (21) or Multiseq 2.0 (13) in VMD 1.8.6 (18).

1. Kim R, et al. (1998) Overexpression of archaeal proteins in *Escherichia coli*. *Biotechnol Lett* 20:207–210.
2. Feng L, Sheppard K, Tumbula-Hansen D, Söll D (2005) Gln-$tRNA^{Gln}$ formation from Glu-$tRNA^{Gln}$ requires cooperation of an asparaginase and a Glu-$tRNA^{Gln}$ kinase. *J Biol Chem* 280:8150–8155.
3. Fechter P, Rudinger J, Giegé R, Théobald-Dietrich A (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: Production and activity. *FEBS Lett* 436:99–103.
4. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41:207–234.
5. Bhaskaran H, Perona JJ (2011) Two-step aminoacylation of trna without channeling in archaea. *J Mol Biol* 411:854–869.
6. Milligan JF, Uhlenbeck OC (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol* 180:51–62.
7. Oshikane H, et al. (2006) Structural basis of RNA-dependent recruitment of glutamine to the genetic code. *Science* 312:1950–1954.
8. Wolfson AD, Pleiss JA, Uhlenbeck OC (1998) A new assay for tRNA aminoacylation kinetics. *RNA* 4:1019–1023.
9. Rodríguez-Hernández A, Bhaskaran H, Hadd A, Perona JJ (2010) Synthesis of Glu-$tRNA^{Gln}$ by engineered and natural aminoacyl-tRNA synthetases. *Biochemistry* 49:6727–6736.
10. Bullock TL, Uter N, Nissan TA, Perona JJ (2003) Amino acid discrimination by a class I aminoacyl-tRNA synthetase specified by negative determinants. *J Mol Biol* 328:395–408.

11. Jühling F, et al. (2009) tRNAdb 2009: Compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res* 37:D159–162.
12. O'Donoghue P, Luthey-Schulten Z (2005) Evolutionary profiles derived from the QR factorization of multiple structural alignments gives an economy of information. *J Mol Biol* 346:875–894.
13. Roberts E, Eargle J, Wright D, Luthey-Schulten Z (2006) MultiSeq: Unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* 7:382.
14. Sethi A, O'Donoghue P, Luthey-Schulten Z (2005) Evolutionary profiles from the QR factorization of multiple sequence alignments. *Proc Natl Acad Sci USA* 102:4045–4050.
15. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321.
16. Markowitz VM, et al. (2008) The integrated microbial genomes (IMG) system in 2007: Data content and analysis tool extensions. *Nucleic Acids Res* 36:D528–533.
17. Berman HM, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242.
18. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:27–38.
19. Nureki O, et al. (2010) Structure of an archaeal non-discriminating glutamyl-tRNA synthetase: A missing link in the evolution of Gln-tRNA^{Gln} formation. *Nucleic Acids Res* 38:7286–7297.
20. Sheppard K, Söll D (2008) On the evolution of the tRNA-dependent amidotransferases, GatCAB and GatDE. *J Mol Biol* 377:831–844.
21. Page RD (2002) Visualizing phylogenetic trees using TreeView. *Curr Protoc Bioinformatics* 6.2.1–6.2.15.

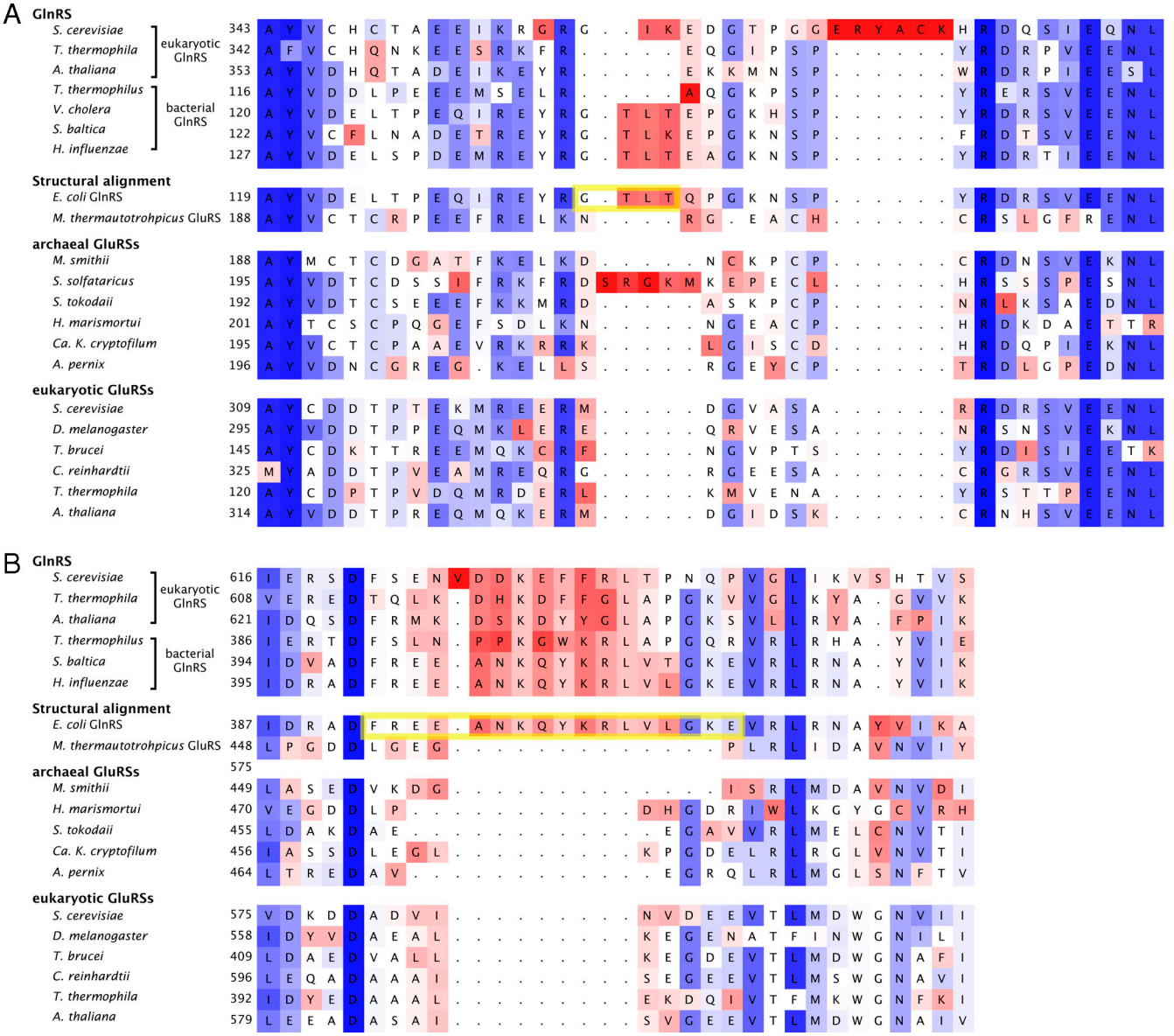

Fig. S1. Alignment of representative GluRS and GlnRS sequences showing only the region nearby the (*A*) acceptor stem loop (ASL) and (*B*) the anticodon loop (AL). The sequence is color coded according to amino acid similarity (descending from blue to red).

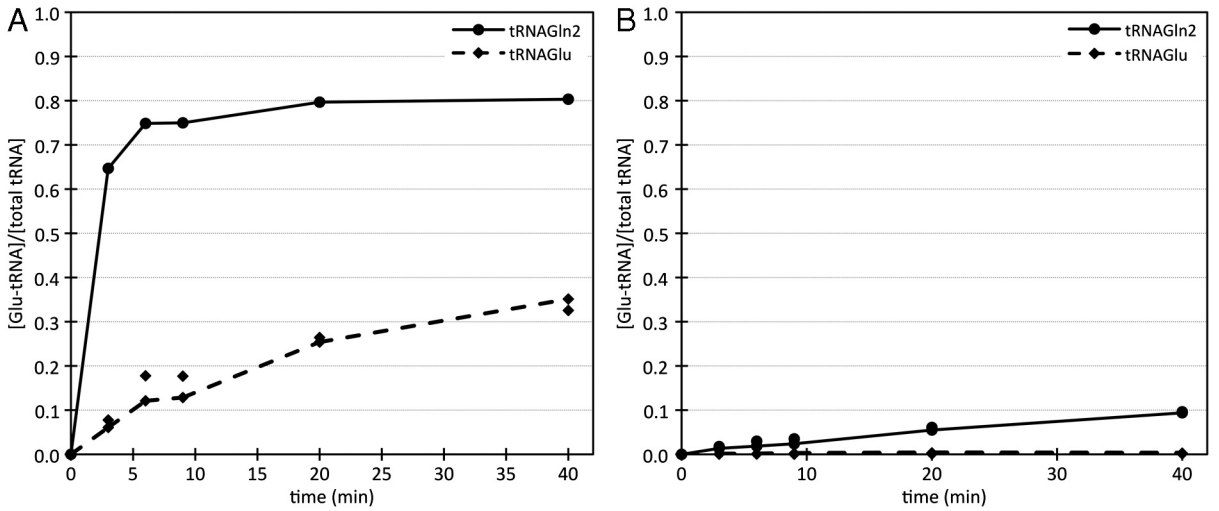

Fig. S2. Plateau tRNA charging curve for AL3-ASL1-GluRS (*A*) and AL3-ASL2-GluRS (*B*) enzymes. Aminoacylation time courses are shown for $tRNA^{Gln2}$ (●) and $tRNA^{Glu}$ (◆). In the reactions, 1 μM enzyme and 10.4 μM tRNA were used.

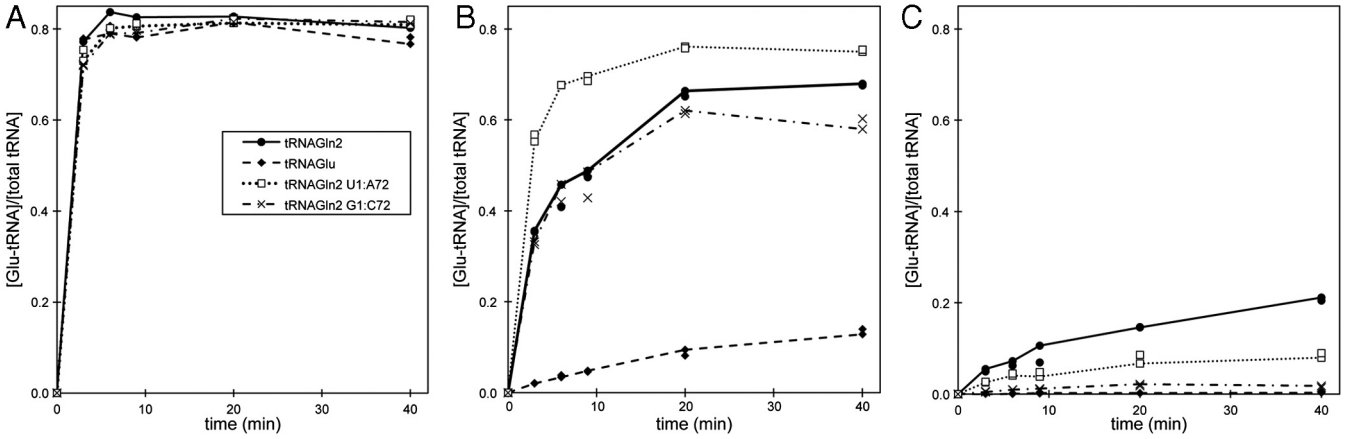

Fig. S3. Plateau tRNA charging curve for GluRS variants and tRNA mutants. Plateau charging levels were measured for the enzymes WT-GluRS (*A*), ASL1-GluRS (*B*), and ASL2-GluRS (*C*) with $tRNA^{Gln2}$ (●), $tRNA^{Glu}$ (◆), and mutant tRNAs: $tRNA^{Gln2}$ $U_1:A_{72}$ (□), and $tRNA^{Gln2}$ $G_1:C_{72}$ (×). In the reactions, 1 μM enzyme and 10.4 μM tRNA were used.

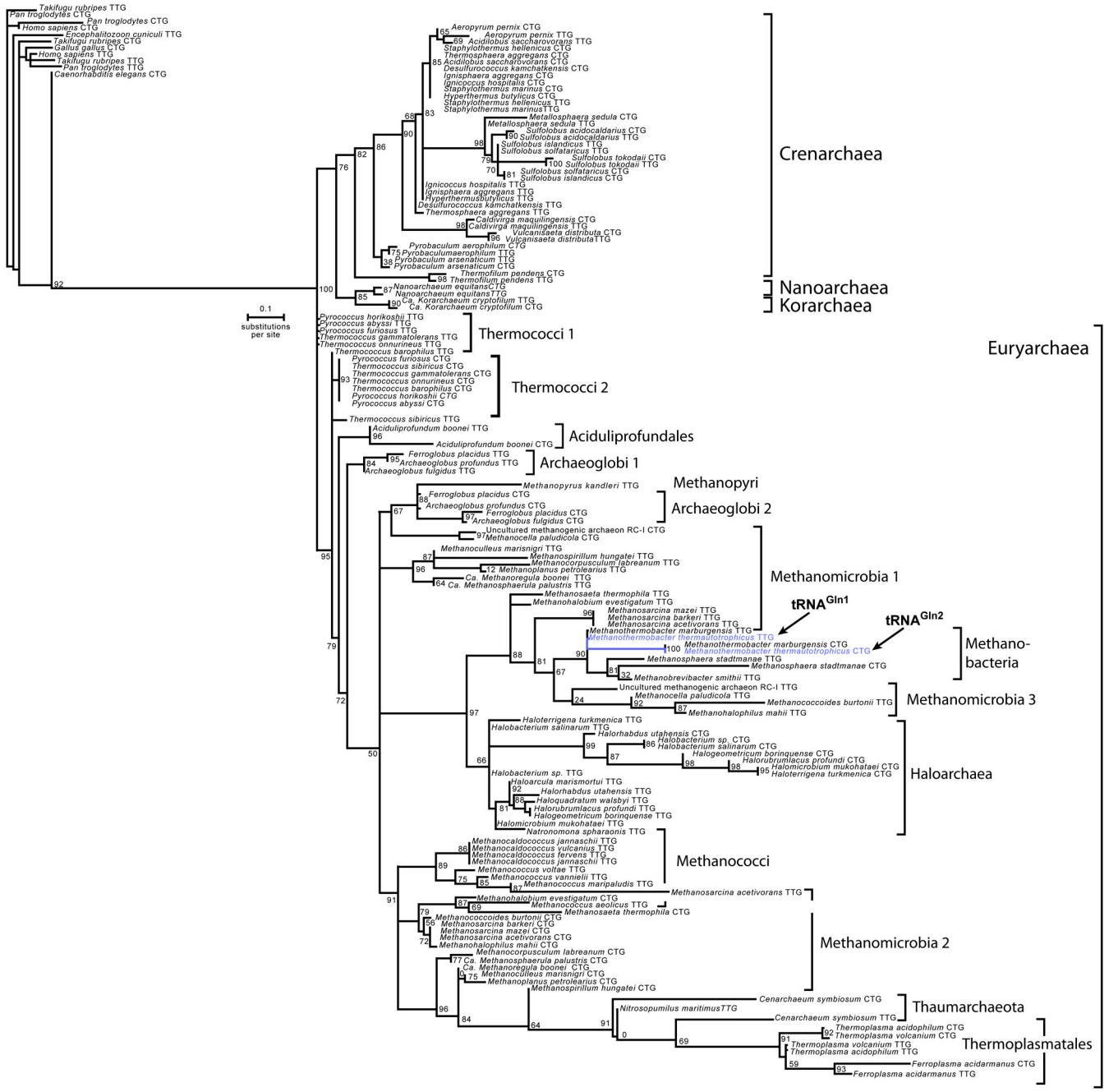

Fig. S4. A maximum likelihood phylogeny of archaeal tRNA^{Gln} sequences. A selection of eukaryotic tRNA^{Gln} sequences were used as an outgroup. The evolutionary relationship between the tRNA^{Gln} isoacceptors of *M. thermautotrophicus* is highlighted in blue. The anticodon (as it appears encoded in the tRNA gene) of the tRNA^{Gln} isoacceptor is indicated after each species name. Bootstrap values are given at each node, and branches without statistical support (bootstrap value of zero) have been collapsed. The scale bar shows the number of inferred nucleotide substitutions per site.

Table S1. Aminoacylation kinetics of Mt-GluRS variants for mutant tRNA substrates

	k_{cat}, s^{-1}	K_M, μM	k_{cat}/K_M, s^{-1} μM^{-1}	Loss of efficiency,* *x* fold	Loss of efficiency,† *x* fold
WT-GluRS					
$tRNA^{Gln2}$‡	0.41 ± 0.04	1.33 ± 0.40	0.31 ± 0.10	1.0	1.0
$tRNA^{Gln2}$ $U_1:A_{72}$	0.36 ± 0.04	2.61 ± 0.90	0.14 ± 0.05	2.2	2.2
$tRNA^{Gln2}$ $G_1:C_{72}$	0.09 ± 0.01	5.25 ± 1.67	0.02 ± 0.01	18.2	18.2
ASL1-GluRS					
$tRNA^{Gln2}$‡	0.12± < 0.01	2.36 ± 0.28	0.05 ± 0.01	1.0	6.1
$tRNA^{Gln2}$ $U_1:A_{72}$	0.04 ± 0.01	13.3 ± 8.2	0.003 ± 0.002	17.6	106.8
$tRNA^{Gln2}$ $G_1:C_{72}$	0.10 ± 0.03	22.0 ± 3.2	0.005 ± 0.001	11.1	67.4
ASL2-GluRS					
$tRNA^{Gln2}$‡	0.004 ± 0.0002	6.49 ± 0.90	$(5.8 \pm 0.8) \times 10^{-4}$	1.0	534.5
$tRNA^{Gln2}$ $U_1:A_{72}$	0.007 ± 0.002	29.8 ± 12.4	$(2.2 \pm 1.0) \times 10^{-4}$	2.6	1.41×10^3
$tRNA^{Gln2}$ $G_1:C_{72}$	0.003 ± 0.001	33.6 ± 18.3	$(9.6 \pm 6.0) \times 10^{-5}$	6.0	3.23×10^3

*Relative loss of catalytic efficiency is the ratio of k_{cat}/K_M for each GluRS variant for $tRNA^{Gln2}$ over the k_{cat}/K_M of the same GluRS variant for the indicated $tRNA^{Gln2}$ mutant.

†Relative loss of catalytic efficiency is the ratio of k_{cat}/K_M of WT-GluRS for $tRNA^{Gln2}$ over the k_{cat}/K_M of the indicated GluRS and $tRNA^{Gln2}$ variants.

‡Kinetics values for wild-type $tRNA^{Gln2}$ (with a $A_1:U_{72}$ base pair) from Tables 1 and 2 are reproduced for comparison.

Other Supporting Information Files

Dataset S1 (TXT)