Published in final edited form as:

Chembiochem. 2013 October 11; 14(15): . doi:10.1002/cbic.201300444.

Transfer RNA misidentification scrambles sense codon recoding

Radha Krishnakumar^{[a],[+]}, Laure Prat^{[b],[+]}, Hans Aerni^{[c],[+]}, Jiqiang Ling^{[b],[+]}, Chuck Merryman^[a], John I. Glass^[a], Jesse Rinehart^[c], and Dieter Söll^[b]

John I. Glass: jglass@jcv.i.org; Jesse Rinehart: jesse.rinehart@yale.edu; Dieter Söll: dieter.soll@yale.edu

^[a]Synthetic Biology and Bioenergy, J. Craig Venter Institute, Rockville, MD 20850 (USA)

^[b]Department of Molecular Biophysics and Biochemistry, and Chemistry, Yale University New Haven, CT 06520 (USA)

^[c]Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06520 (USA), Systems Biology Institute, Yale University, West Haven, CT 06516 (USA)

Abstract

Sense codon recoding is the basis for genetic code expansion with more than two different noncanonical amino acids. It requires an unused or rarely used codon, and an orthogonal tRNA synthetase:tRNA pair with the complementary anticodon. *Mycoplasma capricolum* contains only 6 CGG arginine codons without a dedicated tRNA^{Arg}. We wanted to reassign this codon to pyrrolysine by providing *M. capricolum* with pyrrolysyl-tRNA synthetase, a synthetic tRNA with a CCG anticodon (tRNA^{Pyl}_{CCG}), and the genes for pyrrolysine biosynthesis. Here we show that tRNA^{Pyl}_{CCG} is efficiently recognized by the endogenous arginyl-tRNA synthetase, presumably at the anticodon. Mass spectrometry reveals that in the presence of tRNA^{Pyl}_{CCG}, CGG codons are translated as arginine. This result is not unexpected as most tRNA synthetases use the anticodon as a recognition element. The data suggest that tRNA misidentification by endogenous aminoacyl-tRNA synthetases needs to be overcome for sense codon recoding.

Keywords

sense codon recoding; anticodon; tRNA identity; aminoacyl-tRNA synthetase recognition

Introduction

The genetic code was thought to be universal and frozen^[1] when discovered in the 1960s.^[2] This view was later shown to be incomplete with the discovery of multiple recoding events.^[3] This inspired chemists and synthetic biologists to engineer the genetic code and expand protein building blocks with noncanonical amino acids (ncAAs).^[4] Such recoding systems provided powerful tools to install chemical handles into proteins to study post-translational modifications.^[5] To date, recoding efforts have mostly focused on the three stop codons (UAG, UAA, and UGA), limiting the number of different ncAAs that can be simultaneously incorporated into proteins. Reassigning sense codons to ncAAs thus has emerged as an intriguing topic in chemical and synthetic biology. Given the degeneracy of the genetic code, our current understanding of how many codons a single tRNA can decode,^[6] the small number (only 22) of mammalian mitochondrial tRNA genes,^[7] and some special tRNA decoding properties of ribosomes from different organisms,^[8] the notion

Correspondence to: John I. Glass, jglass@jcv.i.org; Jesse Rinehart, jesse.rinehart@yale.edu; Dieter Söll, dieter.soll@yale.edu.

^[+]These authors contributed equally to this work.

emerged that between 30–40 sense codons are required to encode an organism's genetic information. Thus, a large number of sense codons (>20) should be available to be recoded.

Apart from opening up the dispensable codons, a major hurdle for sense codon recoding is the development of orthogonal tRNA/aminoacyl-tRNA synthetase (AARS) pairs for ncAAs. An AARS specifically recognizes its cognate AA and attaches it to the corresponding tRNA(s); the resulting aminoacyl-tRNA (AA-tRNA) is then delivered by an elongation factor to the ribosome, where the tRNA anticodon pairs with the correct mRNA codon for accurate protein synthesis.^[9] Codon reassignment requires evolution or engineering of orthogonal AARS/tRNA pairs that do not cross-react with endogenous AARSs and tRNAs. While many studies focused on developing orthogonal AARS/tRNA pairs,^[4] their orthogonality is often insufficiently characterized.^[10] Here we show that misidentification of the synthetic tRNA by endogenous AARSs poses a significant challenge for sense coding recoding, and suggest future directions to circumvent this barrier.

Results

To completely recode a sense codon with an ncAA, which requires tRNA anticodon reassignment, we first searched for organisms that lack the endogenous tRNA for a specific codon. *Mycoplasma capricolum* (NCBI Reference Sequence: NC_007633.1) contains only 6 CGG arginine (Arg) codons in its genome and apparently lacks a dedicated tRNA for decoding CGG. This codon is not efficiently translated nor recognized by a release factor, and as such it can be considered an open sense codon.^[11] Only a small fraction of *M. capricolum* tRNA^{Arg}_{ACG} does not undergo an adenosine (A)-to-inosine (I) editing event in the anticodon.^[12] Whereas edited tRNA^{Arg}_{ICG} recognizes CGA, CGU, CGC but not CGG, unmodified tRNA^{Arg}_{ACG} can weakly decode CGG. *Mycoplasma* species are ideal organisms for synthetic biology studies due to their small genome size.^[13] Previously we developed techniques to completely synthesize, assemble and transplant *Mycoplasma* genomes, and have successfully created new *Mycoplasma mycoides* cells that are controlled by a chemically synthesized genome,^[13–14] suggesting an avenue to create synthetic organisms with various open sense codons for recoding. In this work, we decided to recode the *M. capricolum* CGG codon. Archaeal pyrrolysyl-tRNA synthetase (PylRS) is well-suited for sense codon reassignment because: (a) PylRS does not recognize the tRNA^{Pyl} anticodon,^[15] which then can be changed to match the sense codon of choice; (b) PylRS has proven to be orthogonal in *Escherichia coli* and *Saccharomyces cerevisiae*,^[16] and may not recognize *Mycoplasma* tRNAs; (c) PylRS has been engineered to recognize a wide range of ncAAs.^[17]

Our *in vitro* aminoacylation experiments showed that *Methanosarcina mazei* PylRS attached the pyrrolysine (Pyl) analog N^ε-cyclopentylloxycarbonyl-L-lysine (Cyc) onto tRNA^{Pyl} variants with CCG, GAG or CGU anticodons (Figure S1), consistent with previous studies that the tRNA anticodon is not critical for recognition by PylRS.^[15a] For *in vivo* tests, the PylRS/tRNA^{Pyl}_{CCG} pair was expressed from a plasmid in *M. capricolum* cells that were grown in SP4 medium with and without Cyc. The level of *in vivo* tRNA^{Pyl} aminoacylation was determined by acid urea gel electrophoresis followed by Northern blot hybridization with a tRNA^{Pyl}-specific probe.^[18] The result revealed that tRNA^{Pyl}_{CCG} was well-expressed in *M. capricolum* (Figure 1A). Unexpectedly, the tRNA^{Pyl}_{CCG} extracted from the cells showed aminoacylation in the absence of Cyc, and the AA-tRNA^{Pyl}_{CCG} exhibited a different migration pattern than Cyc-tRNA^{Pyl} prepared in *E. coli* (Figure 1A). This suggests that tRNA^{Pyl}_{CCG} is charged with an endogenous amino acid, presumably by a *Mycoplasma* AARS. Further addition of Cyc in the media did not alter the migration pattern of the aminoacylated tRNA^{Pyl}_{CCG}, indicating that either Cyc was not taken up into *M. capricolum* cells, or PylRS was not able to effectively compete with the endogenous AARS for tRNA^{Pyl}_{CCG}.

Mycoplasma (including *M. capricolum* and *M. mycoides*) require some amino acids in the form of short peptides, which are hydrolyzed to free amino acids in the cell.^[19] To test whether a dipeptide form of a Pyl analog would be taken up, we purchased the dipeptide NH₂-Lys(Boc)-Lys(Boc)-OH (Biosynthesis, Inc). Boc-lysine (BocK) was reported to be recognized by wild-type PylRS.^[10b] *M. capricolum* cells carrying PylRS/tRNA^{Pyl}_{CCG} were grown in SP4 medium supplemented with 5 mM of BocK and the dipeptide NH₂-Lys(Boc)-Lys(Boc)-OH. As with Cys, addition of BocK dipeptide did not appear to increase the level of AA-tRNA^{Pyl}_{CCG} (Figure 1B).

To circumvent the problem of ncAA uptake, we decided to express from a plasmid the *Methanosarcina acetivorans* Pyl biosynthesis genes (*pylBCD*) in *M. capricolum*. Previous studies showed that *E. coli* harboring this pathway grown in the presence of ornithine (Orn) produces pyrroline-carboxy-lysine (Pcl), a demethylated form of Pyl.^[20] Pcl was also shown to be readily incorporated by the PylRS/tRNA^{Pyl} pair into proteins expressed in *E. coli* and mammalian cells.^[21] In the presence of *pylBCD* and Orn, tRNA^{Pyl}_{CCG} still appeared to be predominantly acylated with the endogenous amino acid (Figure 1C).

We next tested whether cross-charging of tRNA^{Pyl}_{CCG} by endogenous AARSs also occurs in *E. coli*. As shown in Figure 1D, tRNA^{Pyl}_{CCG} was aminoacylated in *E. coli* in the absence of Cys. In contrast, tRNA^{Pyl} variants with CUA and GAG anticodons were not significantly aminoacylated without Cys, suggesting that the CCG anticodon confers recognition of tRNA^{Pyl} by an endogenous AARS. Addition of Cys led to a band shift of the AA-tRNA^{Pyl}_{CCG}, indicating that PylRS outcompetes the endogenous AARS for tRNA^{Pyl}_{CCG} in *E. coli*.

To identify the amino acid incorporated into CGG codons, we expressed b-galactosidase, which contains 6 CGG codons, in *M. capricolum* cells with PylRS, tRNA^{Pyl}_{CCG}, *pylBCD*, with and without Orn, as well as in otherwise wild type *M. capricolum* cells. To maximize resolution and sequence coverage of the recombinant protein, b-galactosidase was purified from *M. capricolum* cell extracts with specific antibodies. Immuno-purified proteins were separated on SDS-PAGE for subsequent LC-MS/MS experiments. Protein bands at the molecular weight for b-galactosidase were visible in all preparations, which indicated that the full length protein was expressed and all CGG codons were efficiently translated. The representative bands were isolated, digested with trypsin, and subjected to LC-MS/MS. Highly significant peptides were identified that spanned ~50% of the full length b-galactosidase protein and peptides covering three of the 6 CGG codons were unambiguously assigned to the amino acid arginine (Figure 2, Figure S2–S5). Arg was found in all conditions tested and no significant peptides containing Pyl or Pcl were observed in the recombinant b-galactosidase proteins. In an earlier experiment we extracted from wild-type *M. capricolum* cells (containing PylRS and tRNA^{Pyl}_{CCG} but lacking *pylBCD*) a small amount of b-galactosidase and performed LC-MS/MS. Due to the very low protein concentration only one decoded CGG was observed; again Arg was inserted at this position (Figure S6). Although it is possible the endogenous Arg-tRNA^{Arg}_{ACG} might weakly decode CGG codons, we think tRNA^{Pyl}_{CCG} may be recognized by the *Mycoplasma* arginyl-tRNA synthetase (ArgRS) and that Arg-tRNA^{Pyl} is responsible for reading CGG as Arg (Table 1).

ArgRSs from *E. coli* and *Saccharomyces cerevisiae* use the tRNA anticodon, particularly C35, as the major identity element.^[22] Changing the anticodon of tRNA^{Val} to ACG converts it to a substrate for *E. coli* ArgRS.^[22b] It is thus likely that tRNA^{Pyl}_{CCG} would be recognized by *M. capricolum* and *E. coli* ArgRSs. Indeed, we showed in the *in vitro* aminoacylation experiment that tRNA^{Pyl}_{CCG} was aminoacylated with Arg by recombinant *M. capricolum* and *E. coli* ArgRSs, although at a lower efficiency than tRNA^{Arg} (Figure 3 and Table 1). Aminoacylation of tRNA^{Pyl}_{CCG} by the *M. capricolum* S100 cell extract was

also dependent on Arg (Table 1). These results clearly demonstrate that ArgRS cross-reacts with the synthetic tRNA^{Pyl}_{CCG}, which impedes recoding of CGG codons in *M. capricolum*. Anticodon nucleotides are frequently recognized by AARSs as important identity elements.^[23]

Discussion

Our work highlights that anticodon misidentification by host AARSs introduces undesired amino acids into the codons to be reassigned. Several approaches might be used—either separately or in combination—to overcome this hurdle. First, the anticodons of certain tRNAs (e.g., tRNA^{Ser}, tRNA^{Leu} and tRNA^{Ala}) are known not to be recognized by their cognate AARSs.^[23] Synthetic tRNAs bearing these anticodons are therefore less likely to be acylated with endogenous amino acids. For recoding studies that require editing or rewriting of the synthetic genome, one would also need to consider the codon usage. In *M. capricolum*, the UCC serine codon and CUC leucine codon are used only 105 and 47 times, respectively. In the essential genome of *M. mycoides* JCVI syn 1.0,^[13] the UCC and CUC codons occur only 55 and 38 times respectively; and this number is reducing as we continue to minimize the genome. We showed that tRNA^{Pyl}_{GAG}, which reads the CUC codon, was not aminoacylated by *E. coli* AARSs (Figure 1D). This tRNA^{Pyl}_{GAG} is also likely to be an orthogonal tRNA in *Mycoplasma* and could be used in future recoding work. Second, antideterminants in tRNA play important roles in preventing misacylation by non-cognate AARSs.^[24] It is known that a single methylation event (m¹G37) prevents yeast tRNA^{Asp} from being recognized by ArgRS.^[25] Introducing antideterminants into synthetic tRNAs is thus a feasible strategy to establish orthogonality. This, however, would require further characterization of the largely unknown tRNA antideterminants combining computational and biochemical approaches. Third, efficient recognition of the synthetic tRNA by its ‘cognate’ AARS would likely diminish its binding by endogenous AARSs (Figure 4). We showed in *E. coli* that expression of PylRS in the presence of Cyc eliminated charging of tRNA^{Pyl}_{CCG} with Arg. In *M. capricolum*, the PylRS system is less efficient in competing with ArgRS leading to formation of Arg-tRNA^{Pyl}, possibly due to lower ncAA, Pyl, or PylRS levels. Thus, future work is needed to develop PylRS variants that will efficiently acylate *Mycoplasma* cell-permeable ncAAs, under conditions that lead to high levels of these AAs in the cell.^[26] Engineering AARS variants with high catalytic efficiencies for ncAAs is particularly important, because many currently available orthogonal pairs are inefficient compared with natural AARSs.^[27]

Conclusion

This work illustrates that sense codon recoding will be much more challenging than imagined; the orthogonal tRNA/AARS pairs that provided acceptable stop codon recoding will need sophisticated further engineering to satisfy the still incompletely known rules of tRNA identity,^[23] tRNA modification and coding response,^[28] dissimilar ribosomal recognition in different organisms,^[8] and quality control mechanisms,^[29] in order to provide the orthogonality required for successful sense codon recoding.

Experimental Section

Strains and plasmids

M. capricolum subsp. *capricolum* California Kid strain ATCC 27343 (referred to as *M. capricolum* in this study) was grown in SP4 medium^[30] at 37 °C, supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and tetracycline (4 mg/ml) when required. Cyc and BocK were purchased from Sigma, and Bachem, respectively. Dipeptide NH₂-Lys(Boc)-Lys(Boc)-OH was synthesized by Biosynthesis Inc. For production of

pyrroline-carboxy-lysine (Pcl) in *M. capricolum*, the *M. acetivorans* *pylBCD* genes were expressed from derivatives of plasmid pMYCO1^[31] grown in SP4 supplemented with D-Orn and L-Lys (10 mM each).

Purification of b-galactosidase and mass spectrometry

M. capricolum containing the b-galactosidase transgene, *pylS*, *pylBCD*, and *tRNA^{Pyl}* genes described above were grown in 500 ml SP4 medium (with and without Cyc or Orn and Lys) until pH of 5–6 was reached. The cells were collected by centrifugation, washed three times with phosphate-buffered saline (PBS 1×) (13.7 mM NaCl, 0.27 mM KCl, 0.15 mM KH₂PO₄, 0.8 mM Na₂HPO₄, 12 H₂O [pH 7.4]) and lysed using gentle sonication (Sonic Dismembrator, model 500; 10% power, 10 sec each for 6 times) in 150 mM NaCl, protease inhibitor, 50 mM Tris-HCl pH7.5, 5% glycerol and 1 mM 2-mercaptoethanol. The b-galactosidase protein was purified with anti-b-galactosidase antibodies (Abcam #ab616 and Covance #MMS-168P) complex to a 1:1 mixture of protein agarose A and G (Pierce). The immunoprecipitated proteins were resolved on SDS-PAGE and Coomassie stained gel bands were excised and digested with trypsin. Mass spectrometry was carried out as described previously.^[32] Data were analyzed with MASCOT using an *M. capricolum* genome database appended with a b-galactosidase ORF with CGG codons as Arg (Figure S2) or Lys to enable Pyl and Pcl searches. Search results (Table S1) were reported for significant peptides below a global 1.0% FDR cutoff.

Expression and purification of PylRS and ArgRS

The *M. mazei* *pylS* gene was PCR amplified using primers 5' gcctgacgcatatgga-taaaaaacctaaac and 5' cgttagatcctaca-ggtggtagaaatc and cloned into pET16b at the *NdeI* and *BamHI* sites. The full length protein was expressed as a fusion with an N-terminal His tag in *E. coli* BL21 (DE3) Codon Plus RIL (Stratagene) and purified using published protocols^[33]. Briefly, 400 ml cells were grown until A₆₀₀ = 0.2 was reached, induced with IPTG to a final concentration of 0.5 mM and then continued to grow at 30 °C overnight. Cells were spun down and resuspended in lysis buffer (50 mM potassium phosphate buffer pH7.4, 500 mM NaCl, 2.5 mM imidazole, 5 mM 2-mercaptoethanol, and 10% glycerol). Cells were lysed by sonication and centrifuged to remove cell debris. The supernatant was added to 4 ml Ni-NTA agarose and binding was carried out for 1 hour at 4 °C. The resin was washed twice with buffer containing 25 mM imidazole and proteins were eluted with buffer containing 300 mM imidazole. Analysis by SDS-PAGE showed that the protein was present in sufficient quantity and purity. Fractions containing the protein were pooled and dialyzed against 20 mM potassium phosphate buffer pH7.4, 300 mM NaCl, 5 mM MgCl₂, 10 mM 2-mercaptoethanol and 10% glycerol.

The codon optimized gene encoding *M. capricolum* ArgRS for expression in *E. coli* was purchased from DNA 2.0. The full length protein was expressed as a fusion with an N-terminal His-tag in *E. coli* BL21 (DE3) (Stratagene). Protein purification was standardized to the methods described briefly here. 500 ml LB was inoculated at 0.01 A₆₀₀ with an overnight cultures of *E. coli* BL21 DE3 carrying the expression plasmid. When an A₆₀₀ of 0.5 was reached, IPTG was added to a final concentration of 0.1 mM and growth was continued for 4 hours. Cells were collected by centrifugation and resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM Imidazole, 7 mM 2-mercaptoethanol, pH8.0). Cells were lysed by sonication and cellular debris was removed by centrifugation. The supernatant was used for batch binding (5 ml Ni-NTA resin for 20 ml lysate, 1 hour, 4 °C). The column resin was washed twice with the same buffer as before, but with 25 mM imidazole, and proteins were eluted with buffer containing 250 mM imidazole. Following SDS-PAGE analysis, fractions were pooled and dialyzed against buffer containing 50 mM

sodium phosphate buffer pH8.0, 10 mM MgCl₂, 7 mM 2-mercaptoethanol, 30% glycerol. The purified enzymes was tested for activity, quantified and stored at –20 °C.

Isolation of total tRNA and S100 fractions from *M. capricolum*

For total tRNA extraction from *M. capricolum*, cells were grown until early mid-log phase in SP4, and resuspended in buffer A (25 mM Tris-HCl pH7.5, 8 mM MgCl₂, 100 mM NH₄Cl, 30 mM KCl, 10 mM DTT). Nucleic acids were extracted with acid phenol and precipitated with 20% potassium acetate and ethanol. The pellet obtained was resuspended in 100 mM Tris. HCl, pH 7.5 and loaded onto a 5 ml DEAE-Sepharose column pre-equilibrated with 100 mM Tris-HCl pH7.5 and washed with the same buffer until absorbance at A₂₆₀ dropped below 0.1. The tRNA was eluted with 5 column volumes (25 ml) of 100 mM Tris-HCl pH7.5 containing 1 M NaCl. The tRNA was precipitated with ethanol and resuspended in buffer A. For deacylation of aminoacyl-tRNAs, the tRNA was incubated at 37°C for 30 min in 0.5 M Tris-HCl pH9.5. S-100 fractions from *M. capricolum* were prepared according to previously described methods.^[34] *M. capricolum* cells from 250 ml cells grown to mid-logarithmic phase were collected by centrifugation, resuspended buffer A containing 10 mM Tris-HCl (pH7.8), 10mM MgCl₂, 7 mM 2-mecaptoethanol, 10% glycerol, and lysed using sonication. The cell lysate was centrifuged in three stages- 20,000 g for 10 min, 37,000 g for 20 min and 126,000 g for 3 hours. The supernatant from the third step was loaded onto a DEAE-cellulose column (1.5 cm x 30 cm) equilibrated with buffer A and proteins were eluted with a linear gradient of NaCl from 0 - 0.4 M in the same buffer to a total volume of 120 ml. The fractions with highest concentration of protein were pooled and precipitated with 80% NH₄SO₄. The precipitate was dissolved in buffer A, dialyzed against buffer A + 50% glycerol and stored in aliquots at –80°C.

In vitro transcription and tRNA aminoacylation

Unless mentioned otherwise, the transcription of tRNAs and aminoacylation of tRNA^{Pyl} and tRNA^{Arg} variants with PylRS and ArgRS was performed as described in published protocols^[10b, 35] (37°C, in 100 mM Na-HEPES (pH 7.2), 30 mM KCl (or 60 mM KCl for PylRS), 10 mM MgCl₂ (or 25 mM MgCl₂ for PylRS), 10 mM AARS, 5 mM [α -³²P]-labeled tRNA, 10 mM amino acid, 2 mM ATP, and 4 mM DTT. The reaction was stopped with the addition of 1.5 volumes of 0.66 mg/ml nuclease P1 (Sigma) in 100 mM sodium citrate, pH 5.49. The reaction products were separated on PEI- cellulose plates in 0.1M ammonium acetate and 5% acetic acid for quantification using a phosphorimager.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Ilka Heinemann and Patrick O'Donoghue for inspired discussions, and Rembert Pieper and Moojin Suh (JCVI Proteomics Core) for their assistance. We gratefully acknowledge support from the National Institutes of Health (GM22854, to DS; NIDDK-K01DK089006, to JR) and by DARPA contract N66001-12-C-4211 (to JR and DS), and by Synthetic Genomics Inc. (to the Synthetic Biology group at JCVI).

References

1. Crick FHC. J Mol Biol. 1968; 38:367. [PubMed: 4887876]
2. a) Nirenberg M, Leder P, Bernfield M, Brimacombe R, Trupin J, Rottman F, O'Neal C. Proc Natl Acad Sci USA. 1965; 53:1161. [PubMed: 5330357] b) Söll D, Ohtsuka E, Jones DS, Lohrmann R, Hayatsu H, Nishimura S, Khorana HG. Proc Natl Acad Sci USA. 1965; 54:1378. [PubMed: 5325653]

3. Ambrogelly A, Palioura S, Söll D. *Nat Chem Biol.* 2007; 3:29. [PubMed: 17173027]
4. Liu CC, Schultz PG. *Annu Rev Biochem.* 2010; 79:413. [PubMed: 20307192]
5. a) Wan W, Huang Y, Wang Z, Russell WK, Pai PJ, Russell DH, Liu WR. *Angew Chem Int Ed Engl.* 2010; 49:3211. [PubMed: 20340150] b) Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. *J Am Chem Soc.* 2012; 134:10317. [PubMed: 22694658] c) Lee S, Oh S, Yang A, Kim J, Söll D, Lee D, Park HS. *Angew Chem Int Ed Engl.* 2013; 52:5771. [PubMed: 23533151]
6. a) Crick FHC. *J Mol Biol.* 1966; 19:548. [PubMed: 5969078] b) Söll D, Jones DS, Ohtsuka E, Faulkner RD, Lohrmann R, Hayatsu H, Khorana HG. *J Mol Biol.* 1966; 19:556. [PubMed: 5338858] c) Agris PF, Vendeix FA, Graham WD. *J Mol Biol.* 2007; 366:1. [PubMed: 17187822]
7. Suzuki T, Nagao A, Suzuki T. *Annu Rev Genet.* 2011; 45:299. [PubMed: 21910628]
8. Taniguchi T, Miyauchi K, Nakane D, Miyata M, Muto A, Nishimura S, Suzuki T. *Nucleic Acids Res.* 2013; 41:2621. [PubMed: 23295668]
9. Ibba M, Söll D. *Annu Rev Biochem.* 2000; 69:617. [PubMed: 10966471]
10. a) Heinemann IU, Rovner AJ, Aerni HR, Rogulina S, Cheng L, Olds W, Fischer JT, Söll D, Isaacs FJ, Rinehart J. *FEBS Lett.* 2012; 586:3716. [PubMed: 22982858] b) O'Donoghue P, Prat L, Heinemann IU, Ling J, Odoi K, Liu WR, Söll D. *FEBS Lett.* 2012; 586:3931. [PubMed: 23036644]
11. Oba T, Andachi Y, Muto A, Osawa S. *Proc Natl Acad Sci USA.* 1991; 88:921. [PubMed: 1992485]
12. Yokobori SI, Kitamura A, Grosjean H, Bessho Y. *Nucleic Acids Res.* 2013; 41:1093. [PubMed: 2356356]
13. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA, Young L, Qi ZQ, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA 3rd, Smith HO, Venter JC. *Science.* 2010; 329:52. [PubMed: 20488990]
14. a) Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA 3rd, Smith HO. *Science.* 2008; 319:1215. [PubMed: 18218864] b) Lartigue C, Vashee S, Algire MA, Chuang RY, Benders GA, Ma L, Noskov VN, Denisova EA, Gibson DG, Assad-Garcia N, Alperovich N, Thomas DW, Merryman C, Hutchison CA 3rd, Smith HO, Venter JC, Glass JI. *Science.* 2009; 325:1693. [PubMed: 19696314]
15. a) Ambrogelly A, Gundllapalli S, Herring S, Polycarpo C, Frauer C, Söll D. *Proc Natl Acad Sci USA.* 2007; 104:3141. [PubMed: 17360621] b) Polycarpo C, Ambrogelly A, Berube A, Winbush SM, McCloskey JA, Crain PF, Wood JL, Söll D. *Proc Natl Acad Sci USA.* 2004; 101:12450. [PubMed: 15314242]
16. a) Nozawa K, O'Donoghue P, Gundllapalli S, Arais Y, Ishitani R, Umehara T, Söll D, Nureki O. *Nature.* 2009; 457:1163. [PubMed: 19118381] b) Hancock SM, Uprety R, Deiters A, Chin JW. *J Am Chem Soc.* 2010; 132:14819. [PubMed: 20925334]
17. Wang YS, Fang X, Wallace AL, Wu B, Liu WR. *J Am Chem Soc.* 2012; 134:2950. [PubMed: 22289053]
18. Varshney U, Lee CP, RajBhandary UL. *J Biol Chem.* 1991; 266:24712. [PubMed: 1761566]
19. Yus E, Maier T, Michalodimitrakakis K, van Noort V, Yamada T, Chen WH, Wodke JA, Guell M, Martinez S, Bourgeois R, Kuhner S, Raineri E, Letunic I, Kalinina OV, Rode M, Herrmann R, Gutierrez-Gallego R, Russell RB, Gavin AC, Bork P, Serrano L. *Science.* 2009; 326:1263. [PubMed: 19965476]
20. Gaston MA, Zhang L, Green-Church KB, Krzycki JA. *Nature.* 2011; 471:647. [PubMed: 21455182]
21. Ou W, Uno T, Chiu HP, Grunewald J, Cellitti SE, Crossgrove T, Hao X, Fan Q, Quinn LL, Patterson P, Okach L, Jones DH, Lesley SA, Brock A, Geierstanger BH. *Proc Natl Acad Sci USA.* 2011; 108:10437. [PubMed: 21670250]
22. a) McClain WH, Foss K, Jenkins RA, Schneider J. *Proc Natl Acad Sci USA.* 1990; 87:9260. [PubMed: 2251270] b) Tamura K, Himeno H, Asahara H, Hasegawa T, Shimizu M. *Nucleic Acids Res.* 1992; 20:2335. [PubMed: 1375736] Sissler, cM; Giege, R.; Florentz, C. *EMBO J.* 1996; 15:5069. [PubMed: 8890180]

23. Giegé R, Sissler M, Florentz C. *Nucleic Acids Res.* 1998; 26:5017. [PubMed: 9801296]
24. Freyhult E, Moulton V, Ardell DH. *Nucleic Acids Res.* 2006; 34:905. [PubMed: 16473848]
25. Pütz J, Florentz C, Benseler F, Giege R. *Nat Struct Biol.* 1994; 1:580. [PubMed: 7634096]
26. Giese C, Lepthien S, Metzner L, Brandsch M, Budisa N, Lilie H. *ChemMedChem.* 2008; 3:1449. [PubMed: 18752223]
27. a) Umehara T, Kim J, Lee S, Guo LT, Söll D, Park HS. *FEBS Lett.* 2012; 586:729. [PubMed: 22289181] b) Nehring S, Budisa N, Wiltschi B. *PLoS One.* 2012; 7:e31992. [PubMed: 22493661]
28. Gustilo EM, Vendeix FA, Agris PF. *Curr Opin Microbiol.* 2008; 11:134. [PubMed: 18378185]
29. Sprinzl M. *Angew Chem Int Ed Engl.* 2009; 48:3738. [PubMed: 19308938]
30. Tully JG, Whitcomb RF, Clark HF, Williamson DL. *Science.* 1977; 195:892. [PubMed: 841314]
31. Lartigue C, Blanchard A, Renaudin J, Thiaucourt F, Sirand-Pugnet P. *Nucleic Acids Res.* 2003; 31:6610. [PubMed: 14602921]
32. Prat L, Heinemann IU, Aerni HR, Rinehart J, O'Donoghue P, Söll D. *Proc Natl Acad Sci USA.* 2012; 109:21070. [PubMed: 23185002]
33. Yanagisawa T, Ishii R, Fukunaga R, Nureki O, Yokoyama S. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2006; 62:1031.
34. Andachi Y, Yamao F, Muto A, Osawa S. *J Mol Biol.* 1989; 209:37. [PubMed: 2478713]
35. Wolfson AD, Pleiss JA, Uhlenbeck OC. *RNA.* 1998; 4:1019. [PubMed: 9701292]
36. Sampson JR, Uhlenbeck OC. *Proc Natl Acad Sci USA.* 1988; 85:1033. [PubMed: 3277187]

Figure 1.

Aminoacylation of tRNA^{Pyl} variants *in vivo* analyzed by acid PAGE and Northern blot. (A) tRNA^{Pyl}_{CCG} in *M. capricolum* with and without Cyc. Non-specific charging of the tRNA^{Pyl}_{CCG} was observed; there appeared to be no significant difference in the aminoacylation upon addition of Cyc to the growth medium. The alkaline condition (OH⁻) was used to remove the amino acid from the tRNA. (B) tRNA^{Pyl}_{CCG} in *M. capricolum* with and without BocK dipeptide. Addition of dipeptide did not increase aminoacylation of tRNA^{Pyl}_{CCG}. (C) tRNA^{Pyl}_{CCG} in *M. capricolum* with and without *pylBCD*, which forms the Pyl analog Pcl in the presence of Orn.^[21] (D) tRNA^{Pyl} variants in *E. coli* with and without Cyc. Only tRNA^{Pyl}_{CCG} showed non-specific charging in the absence of Cyc.

Figure 2.

MS/MS spectrum of a peptide encoding arginine at position 201 in β -galactosidase. This graph presents a higher energy collisional dissociation fragment ion spectrum and matched y and b ions showing a high scoring peptide identification:

IDGSGQMAITVDVEVASDTPHPAR, Arg underline encoded at GCG in the *M. capricolum* expressed β -galactosidase.

Figure 3. Aminoacylation of tRNA^{Arg} and tRNA^{Pyl} by ArgRSs and *M. capricolum* S100 cell extracts. The secondary structures of tRNA^{Arg} and tRNA^{Pyl} are shown on the left. The arrows indicate positions that are important for recognition by respective AARSs. Aminoacylation was performed in the presence of 1 μ M ArgRS, 5 μ M [³²P]tRNA, and 10 mM Arg. The final product was digested with nuclease P1, separated with thin-layer chromatography, and visualized with phosphorimaging. Following digestion, Arg-tRNA was converted to Arg-AMP, and uncharged tRNA was converted to AMP. Both *E. coli* and *M. capricolum* ArgRSs acylated Arg onto tRNA^{Pyl}_{CCG}. *Ec*, *Escherichia coli*; *Mc*, *Mycoplasma capricolum*; *Mb*, *Methanosarcina barkeri*.

Figure 4. Competition for tRNA^{Pyl}_{CCG} between PylRS and ArgRS during sense codon recoding. tRNA^{Pyl}_{CCG} is recognized by both PylRS and ArgRS. At a low cellular concentration of the PylRS substrate (*e.g.*, Cys), tRNA^{Pyl}_{CCG} is mostly acylated by the endogenous ArgRS. Increasing the concentrations of PylRS and its amino acid substrate would drive tRNA^{Pyl}_{CCG} acylation by PylRS.

Table 1

Percentage of *in vitro* aminoacylation after a 20 min reaction as shown in Figure 3.

AARS	Amino acid	<i>Mc</i> tRNA ^{Arg} _{ACG}	<i>Ec</i> tRNA ^{Arg} _{CCU}	<i>Mb</i> tRNA ^{Pyl} _{CUA}	<i>Mb</i> tRNA ^{Pyl} _{CCG}
<i>Mc</i> ArgRS	Arg	25 ± 4	49 ± 1	1	3 ± 0
<i>Ec</i> ArgRS	Arg	22 ± 1	44 ± 1	1	3 ± 1
<i>Mc</i> S100	20 AA	23 ± 5	ND	1	5 ± 2
<i>Mc</i> S100	19 AA (– Arg)	8 ± 3	ND	1	1
PylRS	BocK	ND	ND	36 ± 2	6 ± 1