Supplementary Information for

Bacterial translation machinery for deliberate mistranslation of the genetic code

Oscar Vargas-Rodriguez[a,1], Ahmed H. Badran[b], Kyle S. Hoffman[c], Manyun Chen[d], Ana Crnković[a,†], Yousong Ding[d], Jonathan R. Krieger[c], Eric Westhof[e], Dieter Söll[a,f,1], Sergey Melnikov[a,#]

[a]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA
[b]Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA
[c]Bioinformatics Solutions Inc., Waterloo, ON N2L 6J2, Canada
[d]Department of Medicinal Chemistry, University of Florida, Gainesville, FL 32610, USA
[e]Université de Strasbourg, Institut de biologie moléculaire et cellulaire du CNRS, 67084 Strasbourg, France
[f]Department of Chemistry, Yale University, New Haven, CT 06511, USA

[1] To whom correspondence may be addressed.
Email: dieter.soll@yale.edu or oscar.vargas@yale.edu

Present address:
[†] Laboratory for Molecular Biology and Nanobiotechnology, National Institute of Chemistry, Ljubljana, Slovenia
[#] Biosciences Institute, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

This PDF file includes
Supplementary text
Table S1
Figures S1-S5

LC-MS/MS analysis

Samples were then prepared for LC-MS/MS analysis at Bioinformatics Solutions Inc. (Waterloo, Ontario, Canada). Briefly, samples were reduced with 10mM DTT (Sigma-Aldrich, Missouri, USA) alkylated with 20mM Iodoacetamide (Sigma-Aldrich, Missouri, USA), acetone precipitated and digested overnight with either MS grade trypsin (Promega, Wisconsin, USA), or chymotrypsin (Promega, Wisconsin, USA). Digested samples were lyophilized. Lyophilized samples were resuspended in 0.2% TFA and desalted using a homemade C18 spin column. C18 desalted samples were resuspended in 12 µl buffer A (0.1% FA). Six µl of each sample was injected into the *tims*TOF Pro (Bruker Daltronics, Bremen, Germany) by nanoflow liquid chromatography using a Bruker NanoElute chromatography system (Bruker Daltronics, Bremen, Germany). Liquid chromatograpny was preformed using a constant flow of 300 µl/min and a 15cm reversed-phased column with a 75 µm inner diameter filled with Reprosil C18 (PepSEP, Odense Denmark). Mobile phase A was 0.1% Formic Acid and Mobile phase B was 99.9% Acetonitrile, 0.1% Formic Acid. Peptide separation was carried out over 21 minutes as follows; linearly 5% A to 30% B over 18 minutes with an increase to 95% B over 30 seconds, and held constant for 2.5 minutes to clean the column. Column equilibration was done prior to sample loading automatically. The column was heated to 60°C. The *tims*TOF Pro was outfitted with a Captivespray source (Bruker Daltronics, Bremen, Germany), operated in PASEF mode. Trapped ion mobility separation was achieved by using an accumiation time of 100ms in the first TIMS region and ramps of the TIMS region from 0.85 V·s/cm^2 to 1.30 V·s/cm^2, with each ramp lasting 100ms. MS and MS/MS scans were limited to 100 m/z to 1700 m/z, and a polygon filter was applied to the m/z and ion mobility dimensions to select for multiple charged ions most likely to be peptide precursors. Due to the use of chymotrypsin, this polygon filter did not exclude singly charged precursors, as these are quite common with a chymotrypsin digestion. Collision energy was applied as a function of ion mobility with a linear regression using the following paremeter settings: 0.85 V·s/cm^2 → 27eV, 1.30 V·s/cm^2 → 45eV. TIMS voltage was calibrated using ions from the Agilent Tune Mix (m/z 622, 922, 1222). Active Exclusion of MS/MS scans was enabled at a setting of 0.40min. Quadrupole isolation was set to 2m/z for m/z less than 700, and 3.0 m/z for ions with an m/z greater than 800. A linear regression calculation was done automatically for ions in between m/z 700 and 800. All mass spectrometry experiments were completed at the mass spectrometry lab of Bioinformatics Solutions Inc (Waterloo, Ontario, Canada). MS Raw Files were processed using PEAKS XPro (v10.6, Bioinformatics Solutions Inc., Ontario, Canada). The data was searched against a custom database containing the sfGFP sequence with the T65A mutation, in conjunction with the *E. coli* K12 Uniprot reviewed database. Precursor ion mass error tolerance was set to 20 ppm and fragment ion mass error tolerance was set to 0.02 Da. Semi-specific cleavage with trypsin or chymotrypsin was selected with a maximum of 3 missed cleavages. A fixed modification of carbamidomethylation (+57.02 Da) on cysteine residues were specified. Variable modifications of deamidation (+0.98 Da) on asparagine and glutamine, as well as oxidation (15.99 Da) on methionine, were specified. The false discovery rate threshold was set to 1% for the database search and only mutations supported by relative fragment ion intensities equal to or greater than 1% were considered. In other words, a pair of major fragment ions (b- or y-ions) must be found before and after the amino acid with a minimum relative ion intensity of 1% in each spectrum to ensure confidence in the identity of the mutation. MS Raw files and Search files were deposited into the PRIDE Repository (identifier PXD026636).

Table S1. List of organisms that encode ProRSx and tRNA^{ProA}.

Organism	Genome ID	ProRS genes
Streptomyces turgidiscabies Car8	698760.3	3
Streptomyces turgidiscabies	85558.6	3
Streptomyces reticuliscabiei strain NRRL B-24446	146821.3	3
Streptomyces scabiei strain NRRL B-2795	1930.35	3
Streptomyces sp. PsTaAH-130	1305828.3	2
Streptomyces sp. CNH099	1137269.3	3
Streptomyces sp. CNT371	1136433.3	3
Streptomyces ipomoeae strain 88-35	103232.16	3
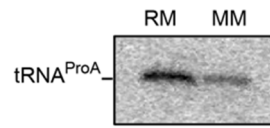
Figure S1. Northern blotting revealed that tRNA^{ProA} is present in extracts from laboratory cultures of *S. turgidiscabies* that were grown in either a rich medium (RM) or a minimal medium (MM).
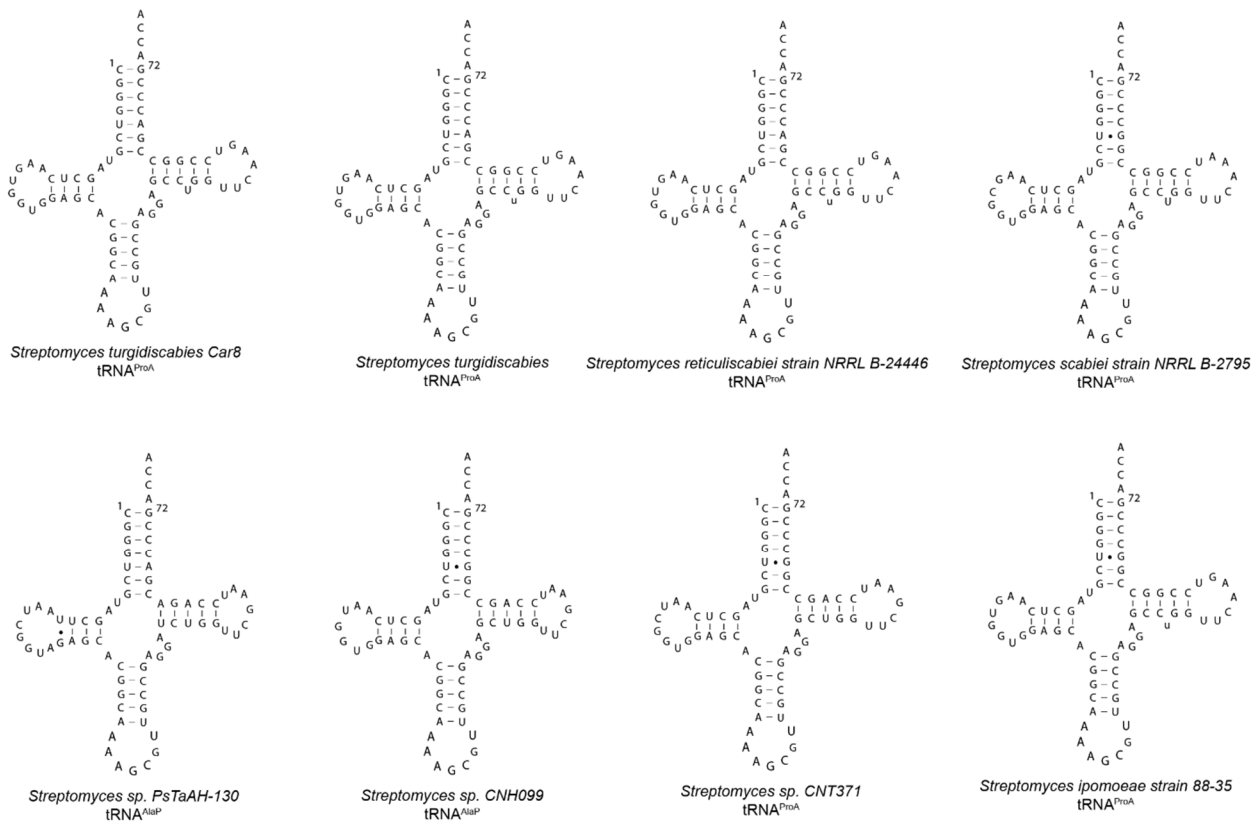
Figure S2. Predicted secondary structures of identified tRNAProA genes. The structures were generated using tRNAscan-SE.
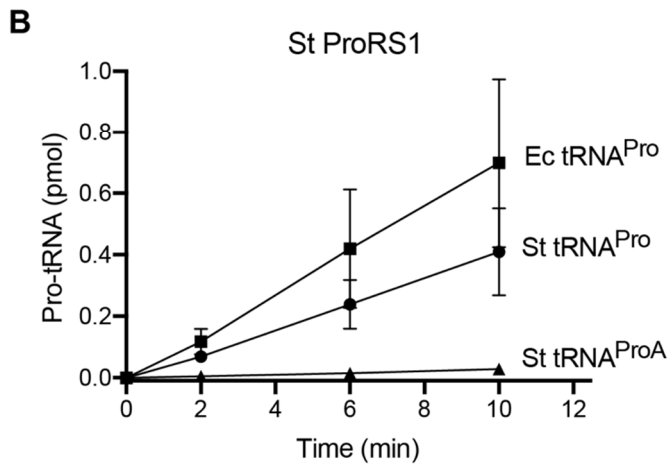
A

30 °C   42 °C

1. Vector
2. Ec ProRS
3. St ProRS1
4. St ProRSx

B

St ProRS1

Ec tRNA^{Pro}

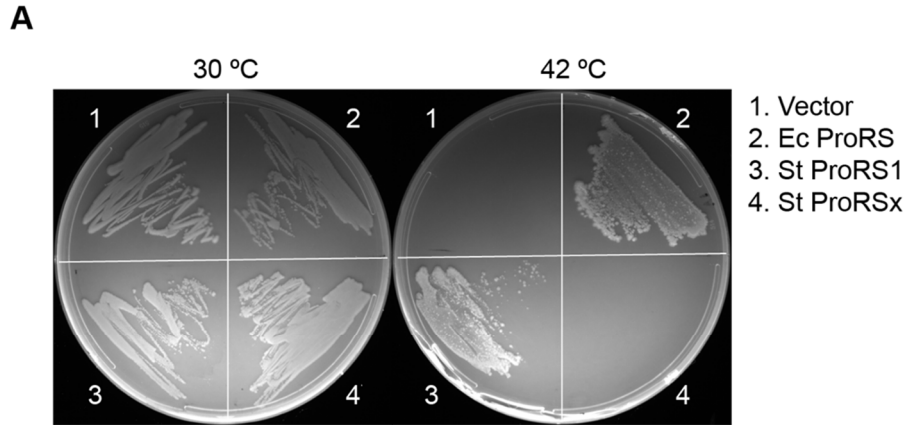St tRNA^{Pro}

St tRNA^{ProA}

Figure S3. (*A*) *E. coli* strain QU27 coding for a temperature-sensitive ProRS mutant was used to assess the ability of *S. turgidiscabies* ProRS1 and ProRSx to aminoacylate endogenous *E. coli* tRNA^{Pro}. In this assay, the *E. coli* cells expressed either plasmid-encoded *E. coli* ProRS, or bacterial-like *S. turgidiscabies* ProRS, or *S. turgidiscabies* ProRSx, or were transformed with an empty vector used as a negative control. Cell growth at 42 °C was expected only when a plasmid-encoded ProRS was able to aminoacylate endogenous *E. coli* tRNA^{Pro}. (*B*) Aminoacylation of *S. turgidiscabies* tRNA^{Pro}, tRNA^{ProA}, and *E. coli* tRNA^{Pro} by *S. turgidiscabies* ProRS1. Each time point represents the average of three independent trials and the error bars indicate the standard deviation.
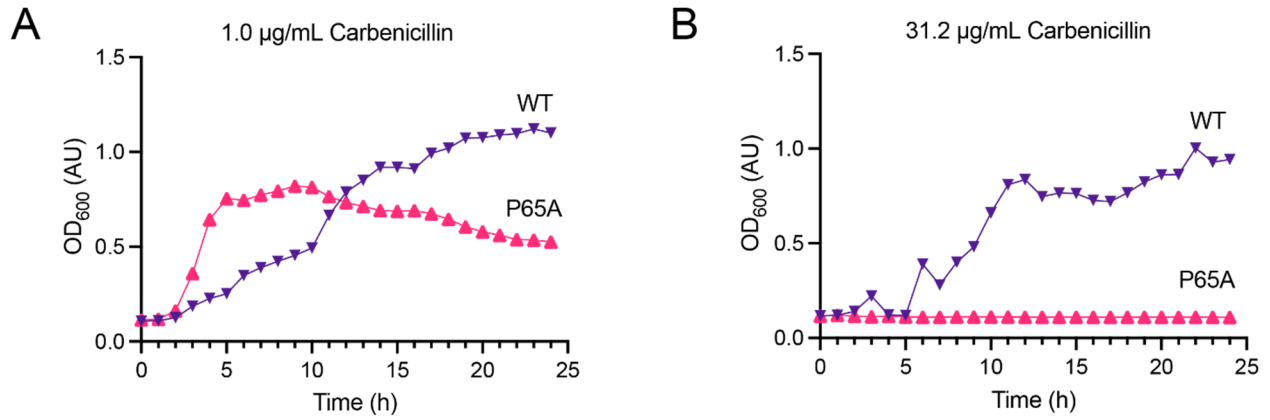
Figure S4. Growth curve assays with *E. coli* S2060 cells harboring the β-lactamase reporter plasmid alone in the presence of (*A*) 1.0 µg/mL or (*B*) 31.2 µg/mL carbenicillin. WT and P65A denotes wild-type and P65A β-lactamase variants, respectively. Cells were grown in LB media with 30 µg/mL kanamycin and the indicated carbenicillin concentration.
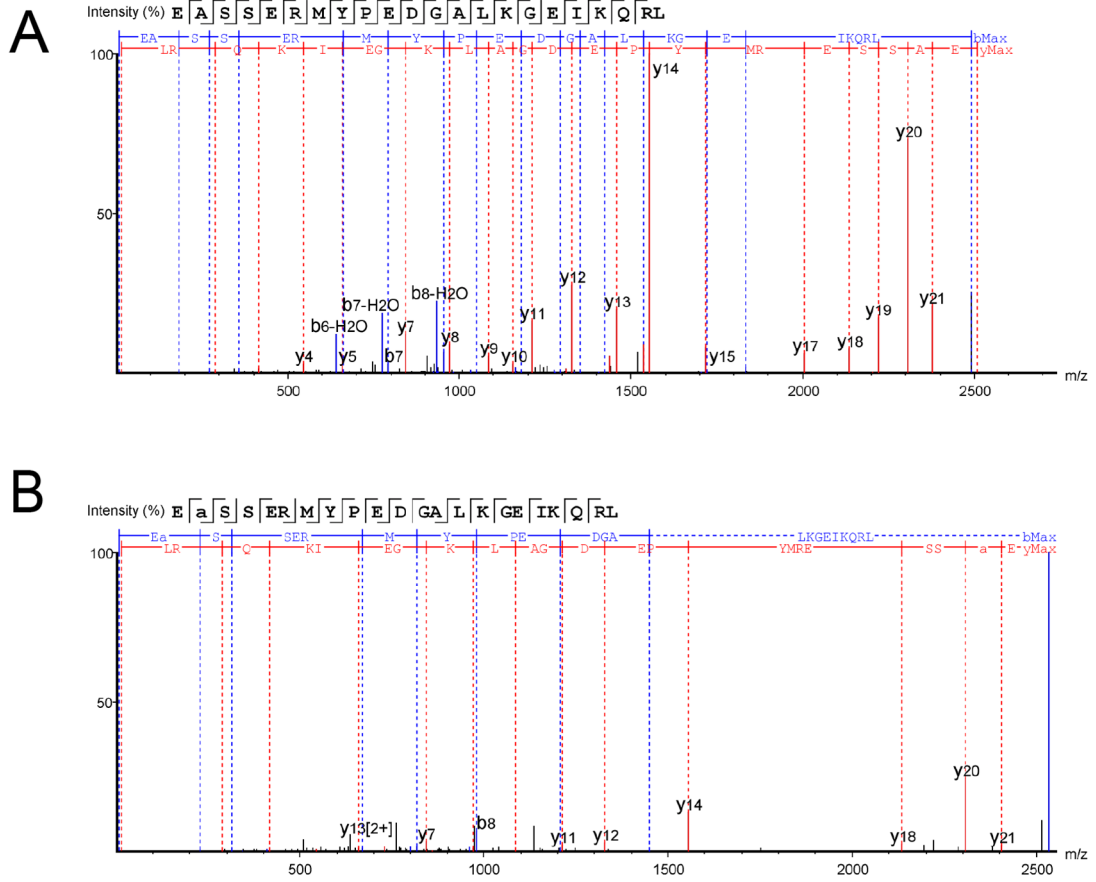
Figure S5. Tandem mass spectrometry identified peptide EASSERMYPEDGALKGEIKQRL of the GFP-mCherry reporter that is also found in a mistranslated form with proline replacing alanine E(a)SSERMYPEDGALKGEIKQRL; proline is indicated as (*a*).