Supplementary Information

Molecular Evolution of Protein-RNA Mimicry as a Mechanism for Translational Control

Assaf Katz[a], Lindsey Solden[a], S.Betty Zou[b], Will Navarre[b] and Michael Ibba[a,c]

[a]Department of Microbiology, Ohio State University, Columbus, OH 43210, USA

[b]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada M5S 1A8

[c]Ohio State Biochemistry Program, Center for RNA Biology, Ohio State University, Columbus, OH 43210, USA

*Corresponding author: Michael Ibba
Department of Microbiology
The Ohio State University
318 West 12th Avenue
Columbus, OH 43210-1292
USA

Tel: +1 (614) 292-2120
E-mail: ibba.1@osu.edu

Table S1. Primers used for *poxA* mutagenesis.

Name	Sequence [a]	Position mutated to alanine
mut_His52Ala_fd	GCGACGGTAACCGATATTgcTTTGGTCCCGTTTGAGACAC	His52
mut_His52Ala_rv	GTGTCTCAAACGGGACCAAAgcAATATCGGTTACCGTCGC	His52
mut_Glu102Ala_fd	CGCAGCTTCCGTAATGcAGAGATGGGGCGTTATC	Glu102
mut_Glu102Ala_rv	GATAACGCCCCATCTCTgCATTACGGAAGCTGCG	Glu102
mut_Glu103Ala_fd	CAGCTTCCGTAATGAAGcGATGGGGCGTTATCAC	Glu103
mut_Glu103Ala_rv	GTGATAACGCCCCATCgCTTCATTACGGAAGCTG	Glu103
mut_Arg106Ala_fd	CCGTAATGAAGAGATGGGGgcTTATCACAACCCTGAGTTCAC	Arg106
mut_Arg106Ala_rv	GTGAACTCAGGGTTGTGATAAgcCCCCATCTCTTCATTACGG	Arg106
mut_His108Ala_fd	GAAGAGATGGGGCGTTATgcCAACCCTGAGTTCACTATGC	Hist108
mut_His108Ala_rv	GCATAGTGAACTCAGGGTTGgcATAACGCCCCATCTCTTC	Hist108
mut_Asp177Ala_fd	GTCGCAGCGAAACTGGcTTTGAGCAATGTTGCTG	Asp177
mut_Asp177Ala_rv	CAGCAACATTGCTCAAAgCCAGTTTCGCTGCGAC	Asp177
mut_Asn180Ala_fd	CAGCGAAACTGGATTTGAGCgcTGTTGCTGATACCGAAGAAG	Asn180
mut_Asn180Ala_rv	CTTCTTCGGTATCAGCAACAgcGCTCAAATCCAGTTTCGCTG	Aspn180
mut_Glu185Ala_fd	CAATGTTGCTGATACCGcAGAAGACCGCGACACG	Glu185
mut_Glu185Ala_rv	CGTGTCGCGGTCTTCTgCGGTATCAGCAACATTG	Glu185
mut_Gln193Ala_fd	GACCGCGACACGCTGCTAgcATTGCTGTTTACCTTTGGC	Gln193
mut_Gln193Ala_rv	GCCAAAGGTAAACAGCAATgcTAGCAGCGTGTCGCGGTC	Gln193
mut_Ser218Ala_fd	GTACCACTTTCCAGCCgcCCAGGCATCACTGGCG	Ser218
mut_Ser218Ala_rv	CGCCAGTGATGCCTGGgcGGCTGGAAAGTGGTAC	Ser218
mut_Arg235Ala_fd	CGAAGATCATCGGGTCGCTGAAgcCTTTGAGGTTTATTATAAAGG	Arg235
mut_Arg235Ala_rv	CCTTTATAATAAACCTCAAAGgcTTCAGCGACCCGATGATCTTCG	Arg235
mut_Glu244Ala_fd	GGTTTATTATAAAGGTATTGcGCTGGCGAATGGTTTCCATG	Glu244
mut_Glu244Ala_rv	CATGGAAACCATTCGCCAGCgCAATACCTTTATAATAAACC	Glu244
mut_Arg303Ala_fd	GTGGCATTAGGTGTTGATgcTCTGGTGATGTTGGCGCTG	Arg303
mut_Arg303Ala_rv	CAGCGCCAACATCACCAGAgcATCAACACCTAATGCCAC	Arg303

[a] Mutation site is marked in lower case.

Table S2. Primers used for EF-P mutagenesis.

Name	Sequence [a]	Position mutated to alanine
mut_Phe29Ala_rv	CCTTTACCCGGTTTTACGGCTTCACTCGCTTCAACCGG	Phenylalanine 29
mut_Phe29Ala_fd	CCGGTTGAAGCGAGTGAAGCCGTAAAACCGGGTAAAGG	Phenylalanine 29
mut_Lys31Ala_rv	CCTGGCCTTTACCCGGTGCTACGAATTCACTCGCTTC	Lysine 31
mut_Lys31Ala_fd	GAAGCGAGTGAATTCGTAGCACCGGGTAAAGGCCAGG	Lysine 31
mut_Gly33Ala_rv	GCAATGCCTGGCCTTTAGCCGGTTTTACGAATTCAC	Glycine 33
mut_Gly33Ala_fd	GTGAATTCGTAAAACCGGCTAAAGGCCAGGCATTTGC	Glycine 33

[a] Mutation site is marked in lower case.

Table S3. Species with *poxA* used for alignments.

	TaxID	Species name	Protein GI
γ proteobacteria			
	386585	*Escherichia coli*	15834382
	62977	*Acinetobacter sp.*	50085316
	243277	*Vibrio cholerae*	15642655
	190485	*Xanthomonas campestris*	21231710
	272843	*Pasteurella multocida*	15601965
	177416	*Francisella tularensis*	56707386
	297246	*Legionella pneumophila*	54296336
δ proteobacteria			
	246197	*Myxococcus xanthus*	108762276
	443143	*Geobacter sp.*	322419666
	448385	*Sorangium cellulosum*	162457520
α proteobacteria			
	634452	*Acetobacter Pasteurianus*	258542853
	342108	*Magnetospirillum_magneticum*	83310384
Spirochaetales			
	243276	*Treponema pallidum*	15639515
	189518	*Leptospira interrogans*	24217355

Table S4. Species without *poxA* used for alignments.

	TaxID	Species name	Protein GI
γ proteobacteria			
	380394	*Acidithiobacillus ferrooxidans*	198283109
δ proteobacteria			
	882	*Desulfovibrio vulgaris*	46580076
	644282	*Desulfarculus_baarsii_DSM_2075*	302343635
ε proteobacteria			
	85962	*Helicobacter pilory*	15644806
	195099	*Campylobacter jejuni*	57238249
	387092	*Nitratiruptor sp.*	152990419
α proteobacteria			
	315456	*Rickettsia felis*	67459441
	262698	*Brucella abortus*	62290582
Firmicutes			
	169963	*Listeria monocytogenes*	16803395
	170187	*Streptococcus pneumoniae*	15900353
	413999	*Clostridium botulinum*	148379858
Actinobacteria			
	83332	*Mycobacterium tuberculosis*	15609671
	525909	*Acidimicrobium ferrooxidans*	256372478
Bacteroidetes			
	402612	*Flavobacterium psychrophilum*	150025056
	295405	*Bacteroides fragilis*	53711777
Cyanobacteria			
	103690	*Nostoc sp.*	17232550
	269084	*Synechococcus elongatus*	56751554

Fig S1. Contacts between EF-P and PoxA. A) EF-P contacts both PoxA subunits of the dimer. Figure shows the complex of 2 PoxA (in green and magenta) and 2 EF-P (in blue and red) as observed in pdb 3a5z (chains A, B, C and D). Arrows indicate area where EF-P contacts the other PoxA of the complex. **B) and C)** Amino acids involved in contacts between EF-P and PoxA. Figure shows amino acids on EF-P (B) and PoxA (C) that make either polar or non-polar contacts. Coloring of figure B and C is as in Figure 1. Additional non-polar contacts are highlighted in blue while non-polar contacts to the other subunit of the complex are marked in cyan.

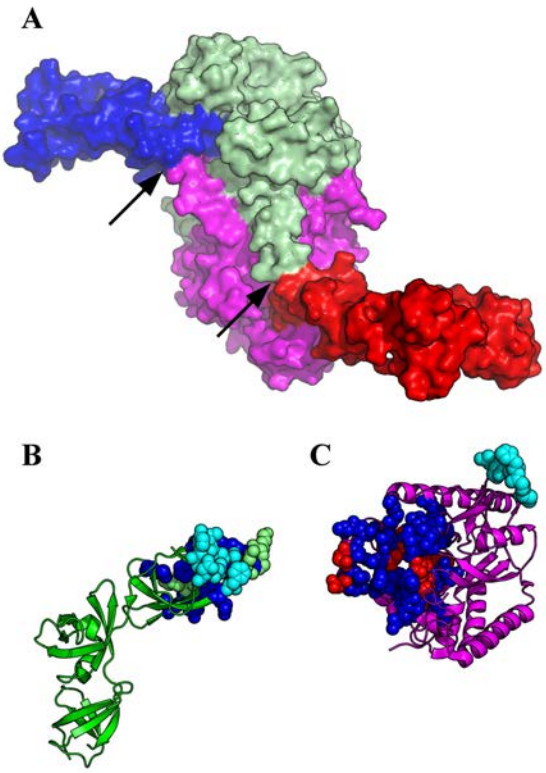

Fig S2. Superposition of EFP/PoxA complex with tRNA^{Asp}/AspRS complex. Structures of EFP/PoxA (pdb 3a5z chains C and D) and tRNA/AspRS (pdb 1asy chains A and R) were superposed guided by PoxA and AspRS structures. PoxA is shown in magenta and EF-P in green while AspRS is in blue and tRNA^{Asp} in orange.

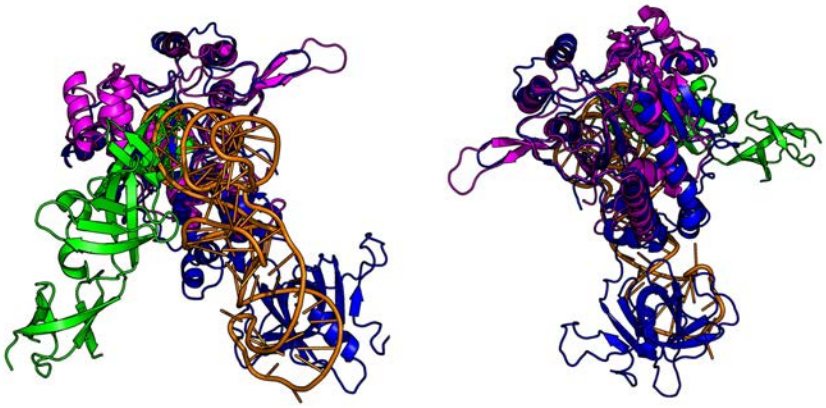

Fig S3. Analysis of EF-P aminoacylation. After EF-P aminoacylation, the sample was separated by isoelectric focusing (inset). Aminoacylated EF-P (K-EF-P) migrated lower in the gel due to the addtional positive charge, which allowed quantification and analysis.

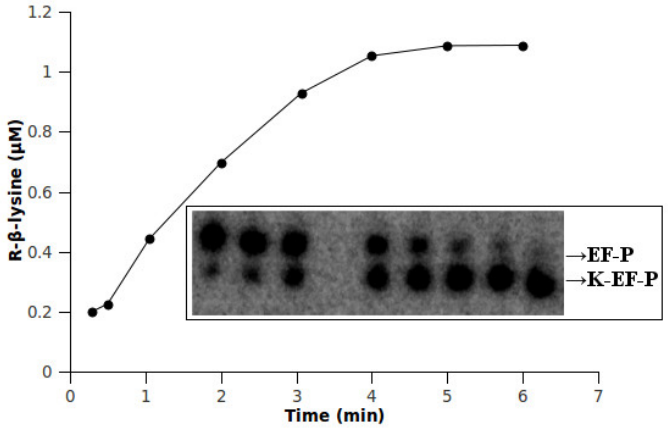

Fig. S4. Alignment of EF-P sequences from organisms with or without a *poxA* gene. Alignment of 14 sequences of EF-P from organisms that have a *poxA* gene (top) and 17 organisms that do not have *poxA* (bottom). Lines named PoxA and ribosome indicate the contact positions to PoxA, or ribosome plus tRNA, respectively. Sequences are named by their NCBI taxID. Correspondence between taxID, species name and the GI accession code of the protein sequence used are given in the tables S3 and S4.

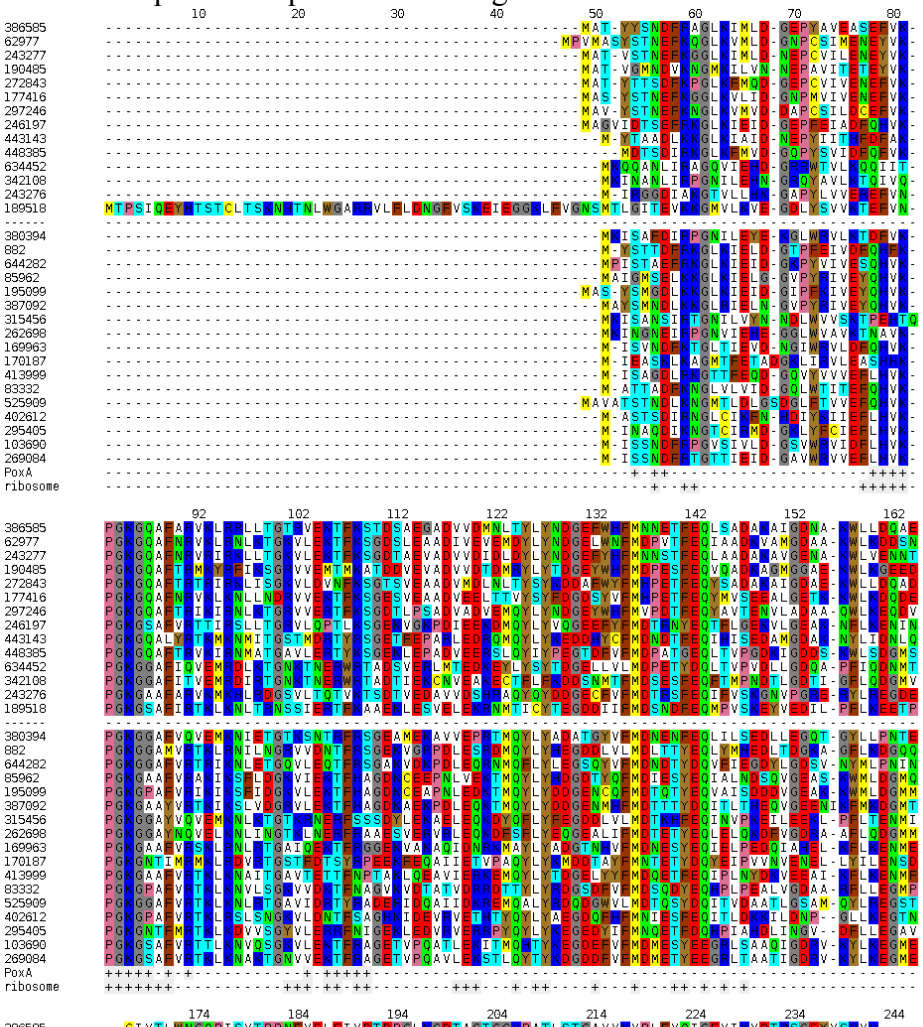

Fig S5. EF-P contacts with PoxA and the ribosome. WebLogo representation of the alignment of 112 EF-P sequences. Residues that interact with PoxA in pdb 3a5z are marked with red triangles while those that interact with the ribosome or $tRNA^{fMet}$ in pdb files 3huw or 3hux are marked with light blue triangles. The acceptor loop is marked with a black line, with a black triangle highlighting the aminoacylation position.