Supplementary information

Enhanced codon-anticodon interaction at in-frame UAG stop codon improves efficiency of non-natural amino acid mutagenesis

Purnima Mala[+] and Ishu Saraogi[+,*]

[+]Department of Biological Sciences, Indian Institute of Science Education and Research Bhopal, Bhopal 462066, India

[*]Department of Chemistry, Indian Institute of Science Education and Research Bhopal, Bhopal 462066, India

Corresponding author email: ishu@iiserb.ac.in

Scheme S1. The DNA oligonucleotides used for the synthesis of the RNA fragments F1 and F3.

S1a Double stranded F1 DNA template – T7 promoter + *Hammerhead Ribozyme* + F1 DNA

5'TAATACGACTCACTATAGGG*CCGG*GACCGGCTGATGAGTCCGTGAGGACGAAACGGTACCCGGTACCGTCCCGGCGGTAGTTCAGCAGGGCAGAAC3'
3'ATTATGCTGAGTGATATCCC*GGCC*CTGGCCGACTACTCAGGCACTCCTGCTTTGCCATGGGCCATGGCAGGGCCGCCATCAAGTCGTCCCGTCT$_m$T$_m$G5'

The double stranded F1 DNA template was prepared by overlapping PCR to perform *in vitro* transcription. The underlined sequence in the ribozyme permits the ribozyme to specifically anneal to underlined sequences in the target RNA to cleave F1 RNA specifically after transcription. The cis-cleaving ribozyme generates the correct 5'-end transcript, thus removing 5'-end heterogeneity. The $_m$T and $_m$G bases represent 2'-O-methylated thymine and guanine at the last two positions in the DNA template to generate the correct 3'-end transcript, thus removing 3'-end heterogeneity.

S1b Double stranded F3 DNA template – T7 promoter + *Hammerhead Ribozyme* + F3 DNA

5'TAATACGACTCACTATAGGG*GATT*GACCGGCTGATGAGTCCGTGAGGACGAAACGGTACCCGGTACCGTCAATCCGCATGGCAGGGGTTCAAATCCCCTCCGCCGGACCA3'
3'ATTATGCTGAGTGATATCCC*CTAA*CTGGCCGACTACTCAGGCACTCCTGCTTTGCCATGGGCCATGGCAGTTAGGCGTACCGTCCCCAAGTTTAGGGGAGGCGGCCTG$_m$G$_m$T5'

The double stranded F3 DNA template was prepared by overlapping PCR to perform *in vitro* transcription. The underlined sequence in the ribozyme permits the ribozyme to specifically anneal to underlined sequences in the target RNA to cleave F3 RNA specifically after transcription. The cis-cleaving ribozyme generates the correct 5'-end transcript, thus removing 5'-end heterogeneity. The $_m$G and $_m$T bases represent 2'-O-methylated guanine and thymine at the last two positions in the DNA template to generate the correct 3'-end transcript, thus removing 3'-end heterogeneity.

S1c DNA splint sequence is complementary to RNA fragments

3'- AAGTCGTCCCGTCTTGCCGCCTGAGATTTAGGCGTACCGTC -5' (41nt DNA splint)
5'CCGGCGGUAGUUCAGCAGGGCAGAACGGCGGDCUCUDAAUCCGCAUGGCAGGGGUUCAAAUCCCCUCCGCCGGACCA3'
(full-length tRNA after ligation)

The 41nt DNA splint was designed to be complementary to 16nt of F1 RNA, complete 11nt of F2 RNA, and 14nt of F3 RNA for specific annealing of RNA fragments allowing correct end-joining of $tRNA_{CUD}$ using T4 DNA ligase.

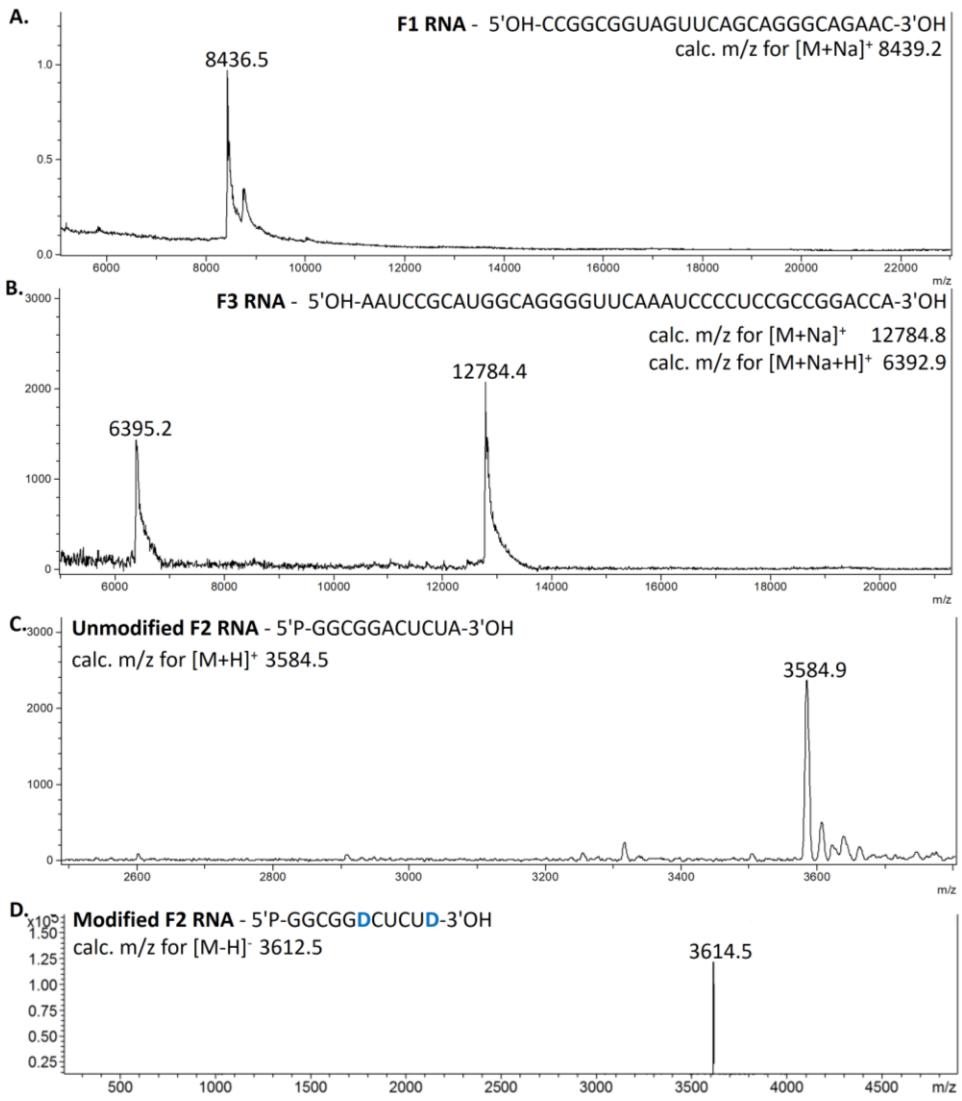

Fig. S1. Mass spectrometric characterization of purified RNA fragments. MALDI-TOF-MS of A. F1 RNA and B. F3 RNA fragment. C. F2 RNA fragment. D. Deconvoluted ESI-MS spectrum of two 2,6-diaminopurine (D) containing F2 RNA fragment.

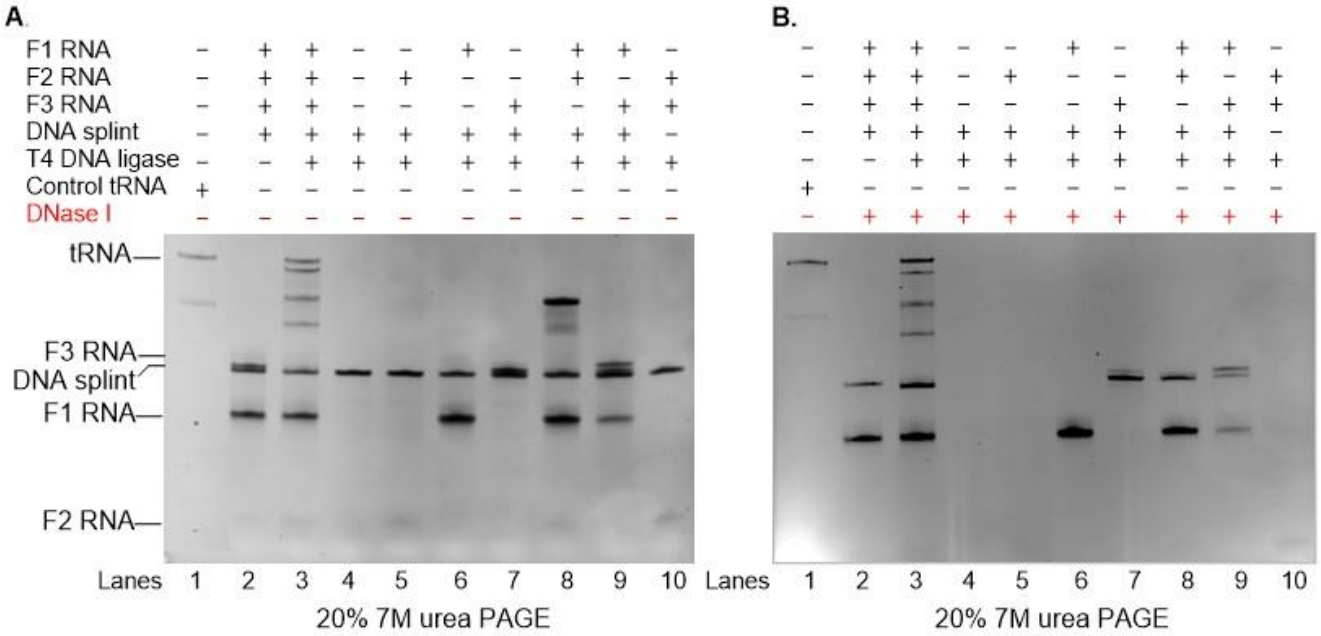

Fig. S2. Analysis of *in vitro* ligation reaction using F1 RNA (5 μM), F2 RNA (10 μM), F3 RNA (5 μM), and DNA splint (2.5 μM) on 20% 7 M urea PAGE. A. The gel was run without DNase I treatment. B. The gel was run after DNase I treatment.

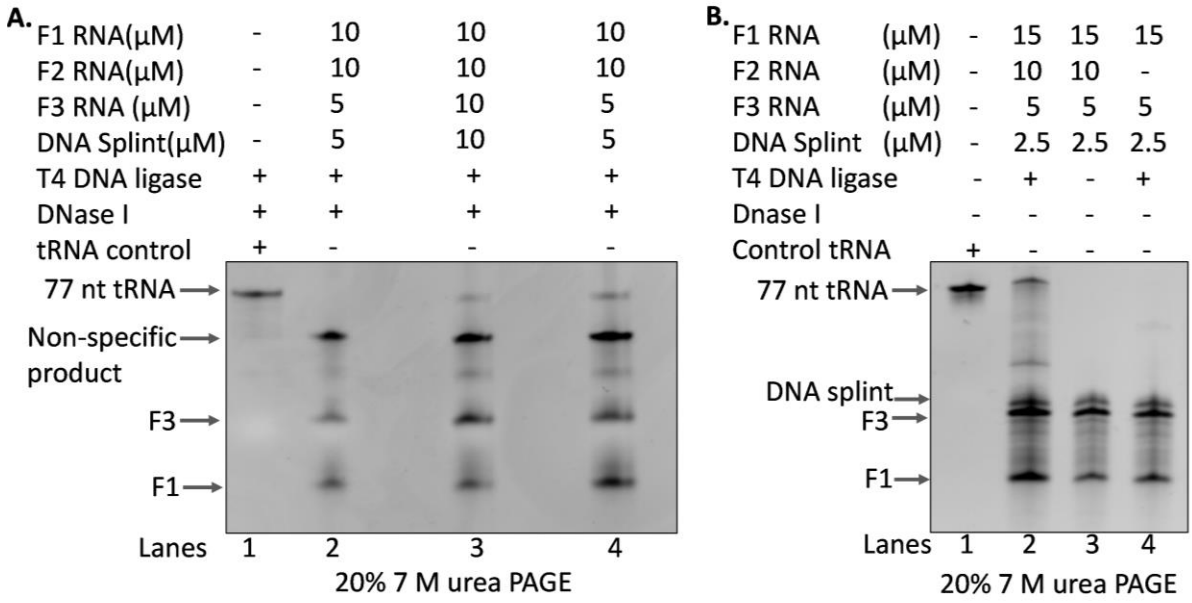

Fig. S3. Optimization of *in vitro* ligation reaction for generation of $tRNA_{CUD}$ with different concentrations of RNA fragments and DNA splint. Lane 1: 77nt tRNA was used as a positive control.

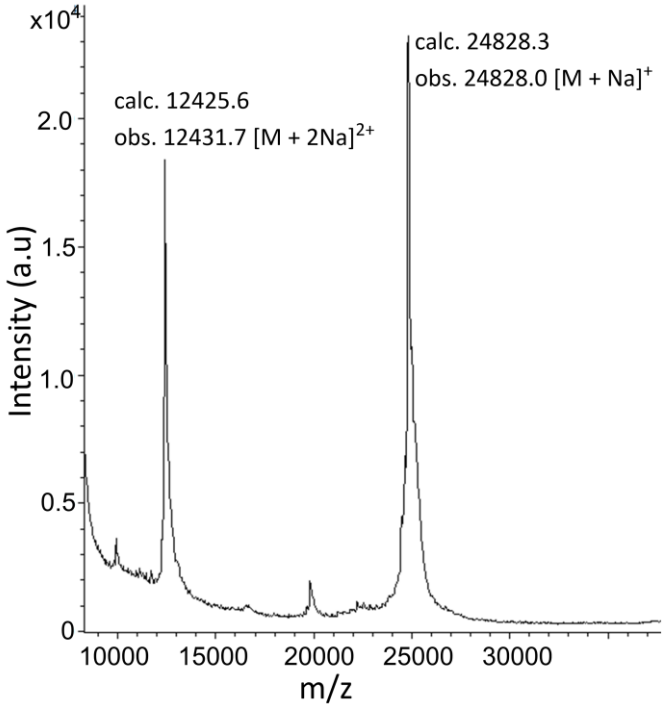

Fig. S4. MALDI-TOF-MS spectra of purified full-length $tRNA_{CUA}$ generated through *in vitro* transcription.

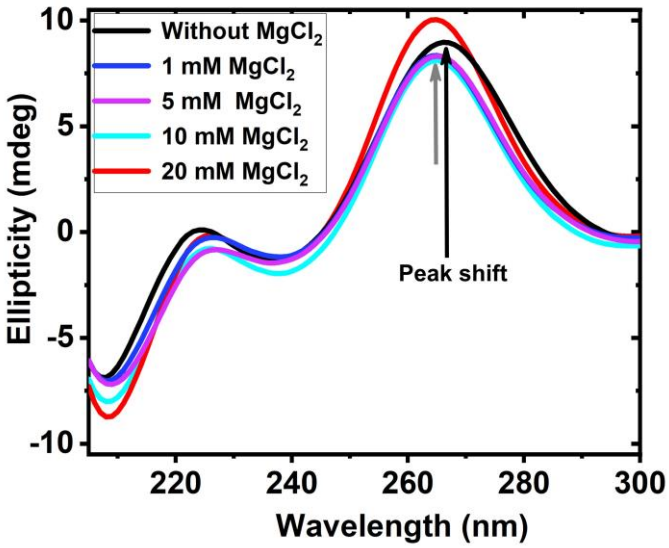

Fig. S5. CD spectra at 25°C for 3 μM refolded native $tRNA^{Tyr}_{CUA}$ at different $MgCl_2$ concentrations.

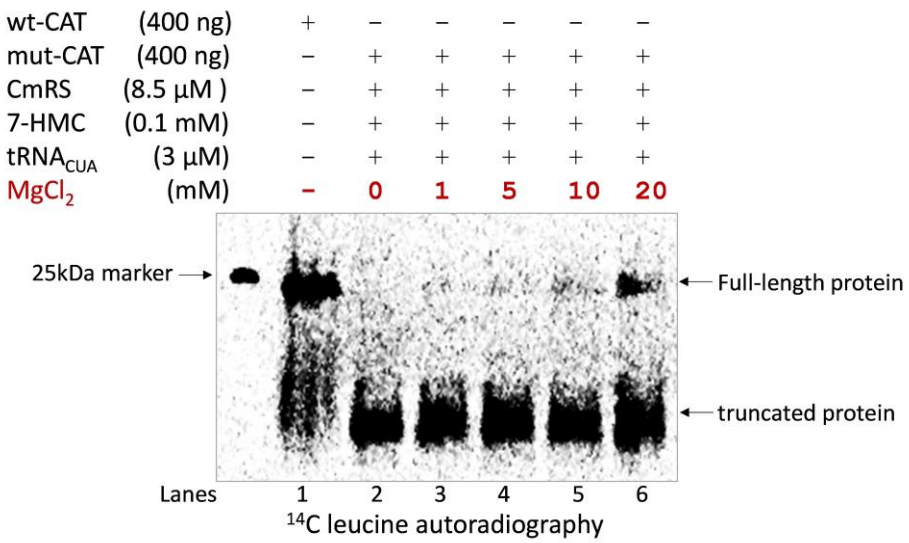

Fig. S6. ^{14}C leucine autoradiography of an *in vitro* translation reaction using tRNA$^{Tyr}_{CUA}$ that was refolded with different $MgCl_2$ concentrations. Full length product is only expected when the natively folded tRNA$_{CUA}$ is present along with the non-natural amino acid (7-HMC) and cognate aminoacyl tRNA synthetase (CmRS).

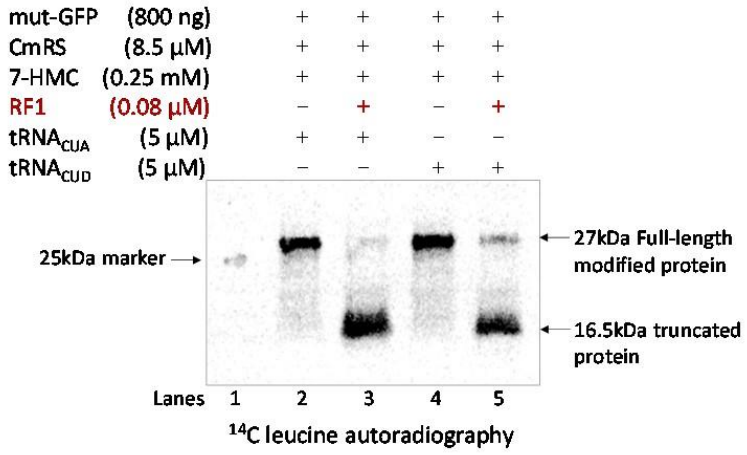

Fig. S7. ^{14}C leucine autoradiography of an *in vitro* translation reaction in presence and absence of RF 1 using NEB PURExpress kit with tRNA analogs. Lane 2; tRNA$^{Tyr}_{CUA}$ in the absence of RF1. Lane 3; tRNA$^{Tyr}_{CUA}$ in the presence of RF1. Lane 4; tRNA$^{Tyr}_{CUD}$ in the absence of RF1. Lane 5; tRNA$^{Tyr}_{CUD}$ in the presence of RF1.