Supplementary files to

Hyperaccurate ribosomes for improved genetic code reprogramming

Bipasana Shakya[1,2], and Olivia Joyner and Matthew C.T Hartman[1,2]

[1] Department of Chemistry, Virginia Commonwealth University, Richmond, Virginia, 23220, U.S.A.
[2] Massey Cancer Center, Virginia Commonwealth University, Richmond, Virginia, 23220, U.S.A.

This supplement contains:

Supporting Figures S1 to S8

Supporting Tables S1 to S10

Contents

Supporting figures and tables

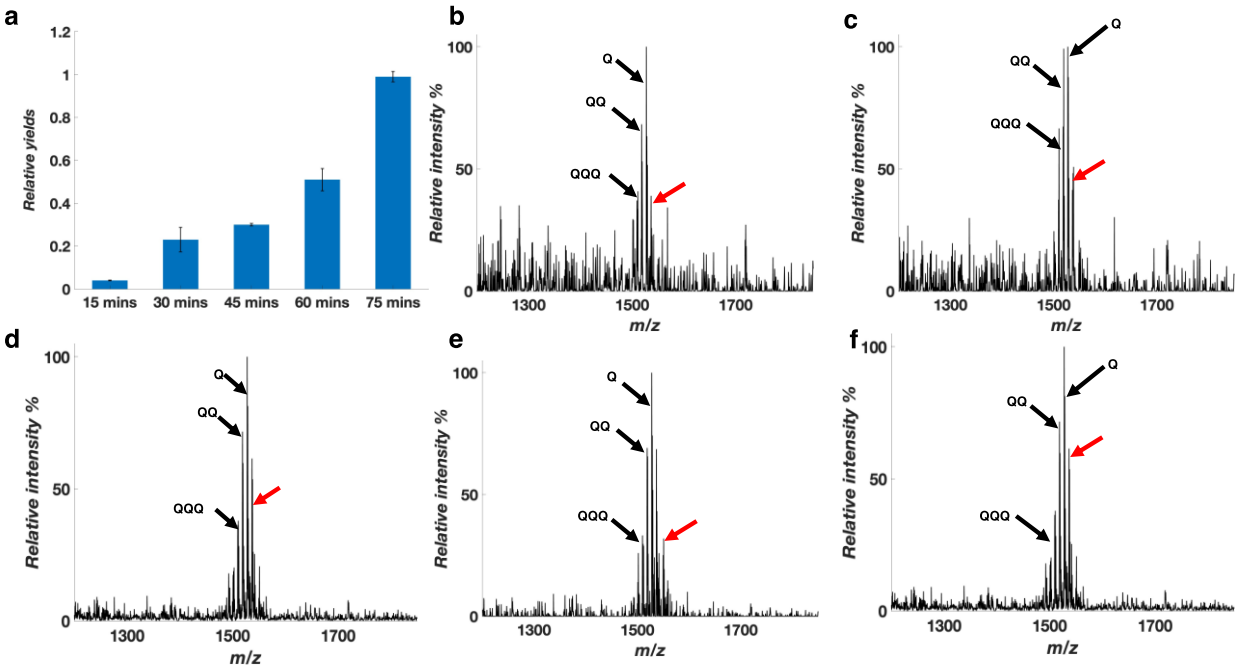


Figure S1: Early time points in translation do not lead to cleaner products. Translation of mRNA 2 in the presence of required AARS/AAs and GlnRS/Gln was followed over time (a) relative yield of translation of mRNA 2 obtained by quantifying the 35-S Met containing peptides bound to Ni-NTA at specified times. MALDI-MS spectra acquired at (b) 15 min, (c) 30 mins, (d) 45 mins, and (e) 60 mins. The correct peak is shown with a red arrow. Glutamine substitutions are indicated with a "Q."

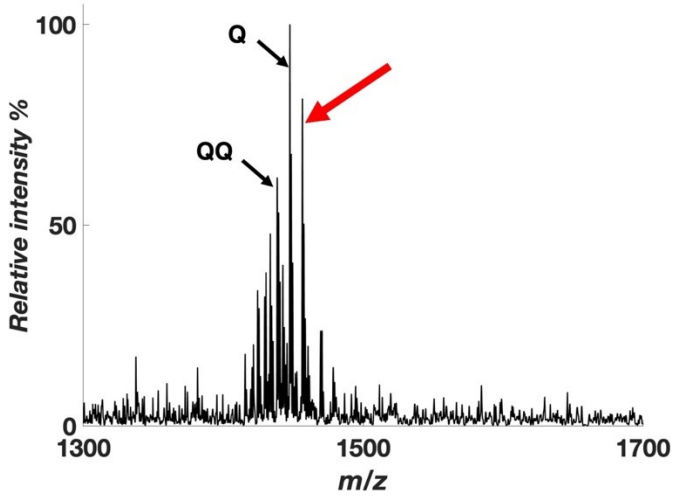


Figure S2: Polymix buffer does not improve fidelity of WT ribosomes. Polymix buffer was prepared as standard protocols[1]. In-vitro translation of peptide MHHHHHHMVEP from mRNA 1 using WT ribosomes, in the presence of Gln/GlnRS and Polymix buffer. The MALDI-TOF spectrum for the peptide is shown above. The red arrow represents target peptide, while black arrows with 'n' number of Q shows n number of Glutamine for Histidine misincorporations.

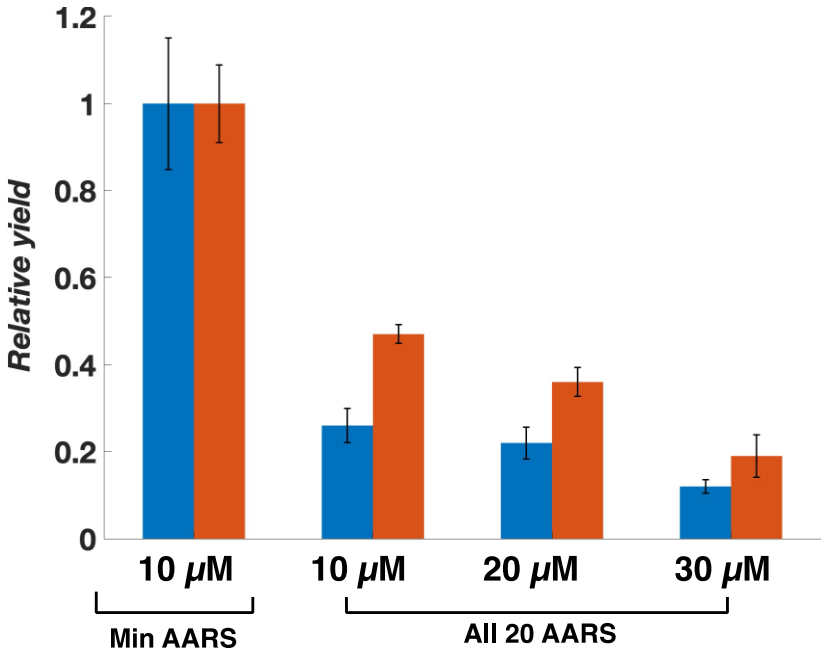
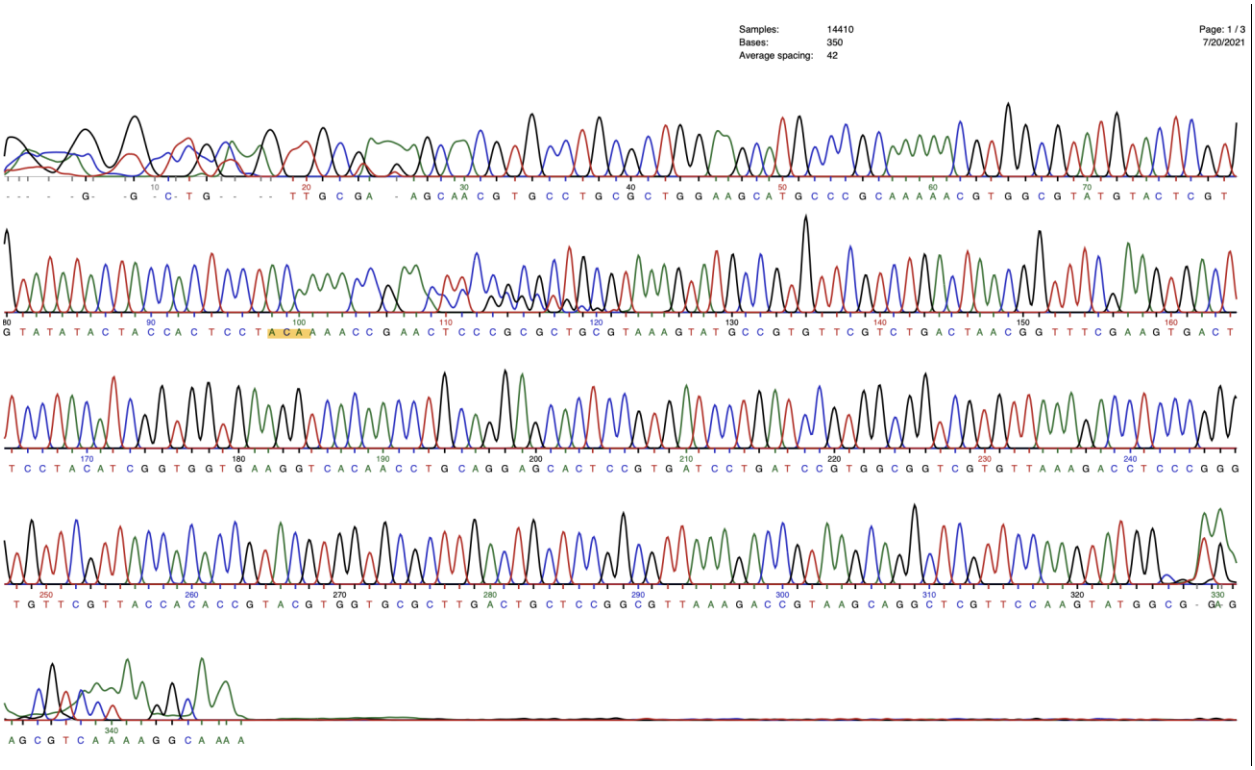


Figure S3: Addition of more EF-Tu does not rescue translation yields in the presence of extraneous AA-tRNAs. In vitro translations of mRNA templates 3 (blue) and 4 (orange) at specified concentrations of Ef-Tu and with either minimal AARS or all 20 AARS. Each bar is representative of at least 4 individual translation reactions quantified separately, error bars represent the standard deviation.



Predicted Sequence of rpsL K42T

atg GCA ACA GTT AAC CAG CTG GTA CGC AAA CCA CGT GCT CGC AAA GTT GCG AAA AGC AAC GTG CCT GCG CTG GAA GCA TGC CCG CAA AAA CGT GGC GTA TGT ACT CGT GTA TAT ACT ACC ACT CCT **ACA** AAA CCG AAC TCC GCG CTG CGT AAA GTA TGC CGT GTT CGT CTG ACT AAC GGT TTC GAA GTG ACT TCC TAC ATC GGT GGT GAA GGT CAC AAC CTG CAG GAG CAC TCC GTG ATC CTG ATC CGT GGC GGT CGT GTT AAA GAC CTC CCG GGT GTT CGT TAC CAC ACC GTA CGT GGT GCG CTT GAC TGC TCC GGC GTT AAA GAC CGT AAG CAG GCT CGT TCC AAG TAT GGC GTG AAG CGT CCT AAG GCT **taa**

Figure S4: Sequencing chromatogram for MAGE of RpsL gene. The highlighted bases correspond to the K to T mutation (AAA to ACA) and match the predicted sequence.

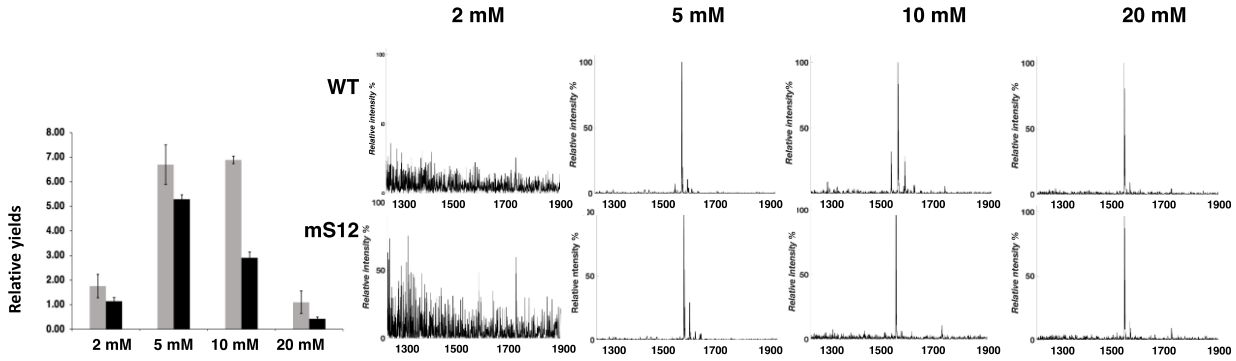


Figure S5: Effects of Mg^{++} on the efficiency of translation. (a) Yields for in vitro translations with template mRNA 2 at different Mg^{++} concentrations by WT (grey) and mS12 ribosomes (black). Values represent averages and error bars are ± standard deviations, with n=3. (b) MALDI-TOF spectra for each condition, Na + adducts are also observed in addition to the target peak of [M+H] 1537.68.

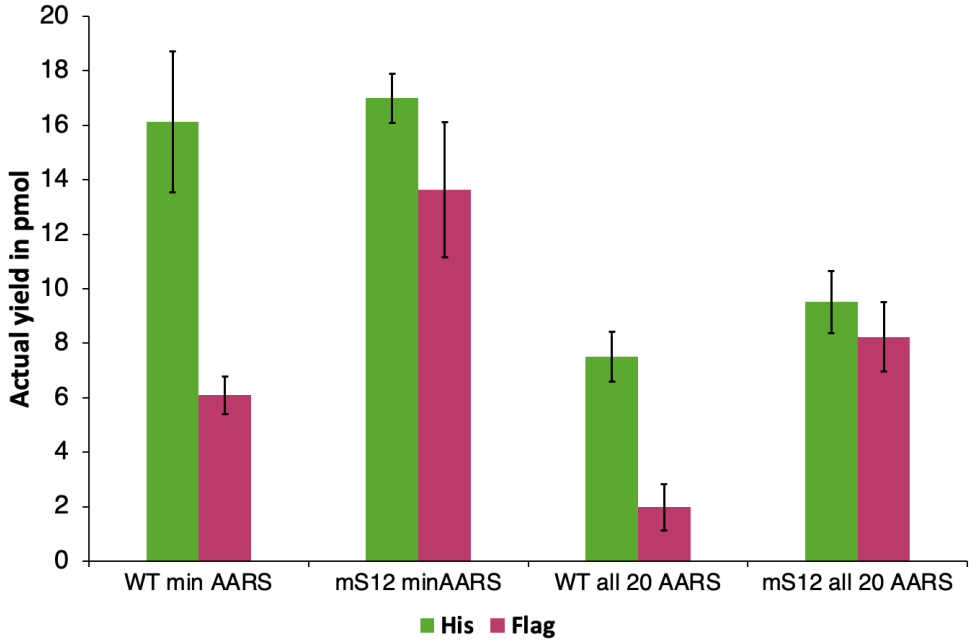


Figure S6: mS12 ribosomes show increased full length peptide yields. Actual yields obtained for in vitro translation reactions containing WT and mS12 with both minimal and all 20 AARS conditions as purified by either the N-terminal His_6 tag (green) or C-terminal FLAG tag (magenta). Values represent averages and error bars are ± standard deviations, with n=4 (* : $p < 0.005$, NS= not significant).

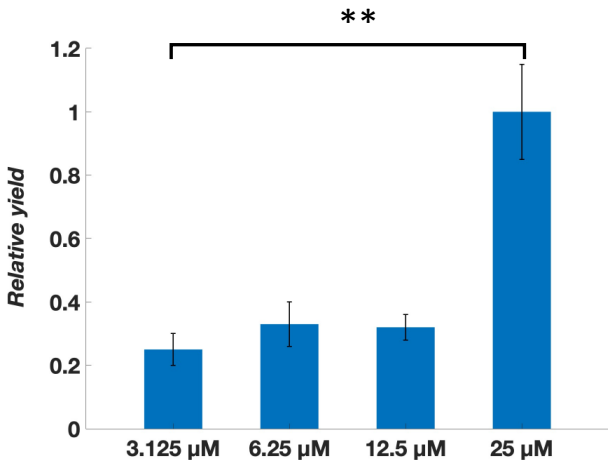


Figure S 7: Dose response for titration of α-methyl cysteine tRNA. Yields of translation for mRNA 5 at increasing concentrations of added α-methyl cysteine-$tRNA^{Val}_{GUA}$. Values represent averages and error bars are ± standard deviations, with n=4, (** : p< 0.00001, NS= not significant).

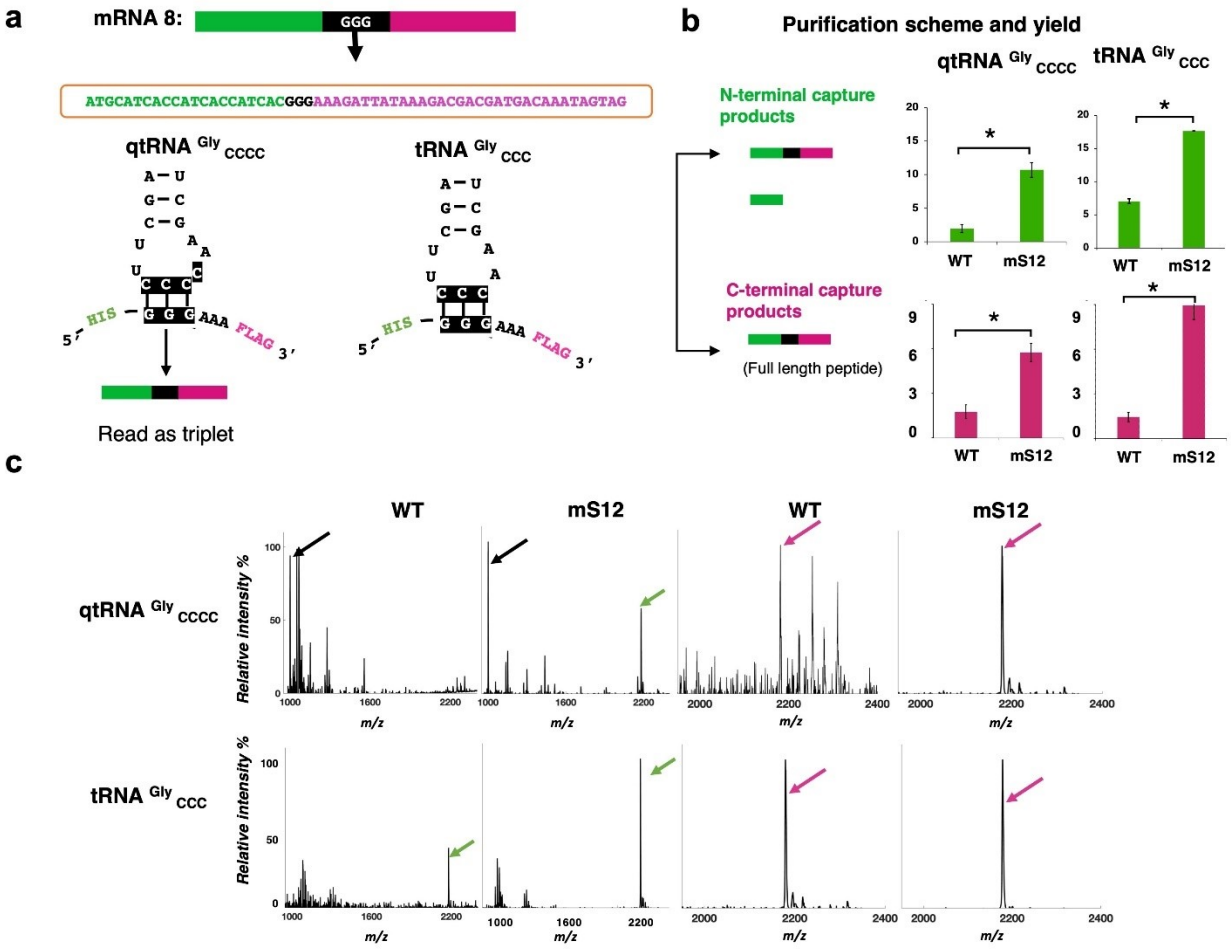


Figure S8: qtRNA$^{Gly}_{cccc}$ can efficiently read triplet codons. (a) The design of mRNA 8 is similar to mRNA 7 in the main text except it has GGG codon instead of GGGG. mRNA 8 is translated in presence Gly-qtRNA$^{Gly}_{cccc}$ or Gly-tRNA$^{Gly}_{ccc}$ (b) Purification with N-terminal his capture and FLAG capture and the expected products, alongside the yields obtained for each purification as produced by WT and mS12 ribosomes (green and magenta respectively) and either qt or standard tRNAs. Translation experiments were performed with all of the components necessary to translate the full-length peptide MHHHHHHGKDYKDDDDK. Values represent averages and error bars are ± standard deviations, with n=4 or 6. (*: $p < 0.005$). (c) MALDI spectra for WT and mS12 for respective captures. Note: Green and magenta arrows represent full length peptides from His and FLAG tag respectively. Black arrows represent truncated peptides containing MHHHHHH (truncation before the GGG codon).

Supporting Tables:

Table S1: Primer sequences used for sequencing the K42T rpsL mutation

Primer name	Sequence from 5'-3'
S12IDforward	5′ ATGGCAACAGTTAACCAGCTG
S12IDReverse	5′ GCCTTAGGACGCTTCACG

MAGE oligo for rpsL mutations

Using the Modest MAGE oligo design tool (http://modest.biosustain.dtu.dk)[2], the oligo sequence determined to use was:

Oligo name	Sequence from 5'-3'
rpsL MAGE oligo	TTCGAAACCGTTAGTCAGACGAACACGGCATACTTTACGCAGCGCGGAGTTCGGtt ttgtAGGAGTGGTAGTATATACACGAGTACATAC

Table S2: mRNAs used table:

S. N	mRNA name	mRNA sequence	Peptide sequence	[M+H] calc.
1	mRNA 1	GGGAGACCACAACGGUUUCCCUCUAGAAAUAAUUUU GUUUAACUUUAAGAAGGAGAUAUACC**AUG**CAUCACC AUCACCACCAUAUGGUAGAACCGUAAUAAGGAUCCG GCUGCUAACAAAGCCCGAAAGGAAGCUGAGUUGGCU GCUGCCACCGCUGAGCAAUAACUAGC	MHHHHHHMVEP	1456.61
2	mRNA 2	GGGAGACCACAACGGUUUCCCUCUAGAAAUAAUUUU GUUUAACUUUAAGAAGGAGAUAUACCAUGGAUUAUA AAAUGCAUCACCAUCACCAUCACUAGUAGGAUUAUA AAGACGACGAUGACAAAUAGUAGGGAUCCGGCUGCU AACAAAGCCCGAAAGGAAGCUGAGUUGGCUGCUGCC ACCGCUGAGCAAUAACUAGC	MDYKMHHHHHH	1537.63
3	mRNA 3	GGGAGACCACAACGGUUUCCCUCUAGAAAUAAUUUU GUUUAACUUUAAGAAGGAGAUAUACAUAUGACCAUU AACCGUGAUUAUAAAGACGACGAUGACAAAUAGUAG GGAUCCGGCUGCUAACAAAGCCCGAAAGGAAGCUGA GUUGGCUGCUGCCACCGCUGAGCAAUAACUAGC	MTINRDYKDDDDK	1656.71
4	mRNA 4	GGGAGACCACAACGGUUUCCCUCUAGAAAUAAUUUU GUUUAACUUUAAGAAGGAGAUAUACAUAUGCUGGAA CCGCAGGAUUAUAAAGACGACGAUGACAAAUAGUAG GGAUCCGGCUGCUAACAAAGCCCGAAAGGAAGCUGA GUUGGCUGCUGCCACCGCUGAGCAAUAACUAGC	MLEPQDYKDDDDK	1639.68
5	mRNA 5	GGGAGACCACAACGGUUUCCCUCUAGAAAUAAUUUU GUUUAACUUUAAGAAGGAGAUAUACAUAUGCAUUUU AGCUGGGAUUAUAAAGACGACGAUGACAAAUAGUAG GGAUCCGGCUGCUAACAAAGCCCGAAAGGAAGCUGA GUUGGCUGCUGCCACCGCUGAGCAAUAACUAGC	MHFSWDYKDDDDK	1729.68
6	mRNA 6	GGGAGACCACAACGGUUUCCCUCUAGAAAUAAUUUU GUUUAACUUUAAGAAGGAGAUAUACAUAUGCAUCAC CAUCACCACCAUAUGCAUUUUAGCUGGGAUUAUAAA GACGACGAUGACAAAUAGUAGGGAUCCGGCUGCUAA CAAAGCCCGAAAGGAAGCUGAGUUGGCUGCUGCCAC CGCUGAGCAAUAACUAGC	MHHHHHHMHFSW DYKDDDDK	2683.07
7	mRNA 7	GGCGTAATACGACTCACTATAGGGTTAACTTTACCG AAGGAGGAAAGAATGCATCACCATCACCATCACGGG GAAAGATTATAAAGACGACGATGACAAATAGTAG	MHHHHHH**G**KDYKD DDDK	2179.91
8	mRNA 8	GGCGTAATACGACTCACTATAGGGTTAACTTTACCG AAGGAGGAAAGAATGCATCACCATCACCATCACGGG AAAGATTATAAAGACGACGATGACAAATAGTAG	MHHHHHHGKDYKD DDDK	2179.91

Table S3: Supplementary Table for Fig 1a

	Target $[M+H]^+$	Misincorporation $[M+H]^+$			
	No Q	**H→ Q**	**HH→ QQ**	**HHH→ QQQ**	**HHHH→ QQQQ**
Expected M+H	**1456.67**	1447.61	1438.61	1429.61	1420.61
Observed M+H	**1456.67**	1447.93	1438.67	1429.82	1420.92

Table S4:Supplementary Table for Fig 1c

	Target $[M+H]^+$	Adducts formed $[M+H]^+$		
	Target	**+ bME**	**H→E mis-incorporation**	**H→E + Na**
Expected M+H	**1474.66**	1550.67	1486.59	1508.89
Observed M+H	**1474.91**	1550.77	1486.87	1508.86

Table S5: Supplementary Table for Fig 2

Ribosomes	**Expected $[M+H]^+$**	**Observed $[M+H]^+$**				
	0-500 μM	**0 μM**	**50 μM**	**100 μM**	**200 μM**	**500 μM**
WT	**1656.717**	1656.72	1656.96	1656.88	1656.97	-
mS12	**1656.717**	1656.96	1656.98	1656.73	1656.81	1656.89

Table S6:Supplementary Table for Fig 3

Ribosomes	**Expected $[M+H]^+$**	**Observed $[M+H]^+$**			
		Min aaRS His	**Min aaRS Flag**	**All 20 aaRS His**	**All 20 aaRS Flag**
WT	**2683.074**	2683.08	2683.50	-	2683.18
mS12	**2683.074**	2683.71	2683.18	2683.60	2683.50

Table S7:Supplementary Table for Fig 4

Ribosomes	Expected $[M+H]^+$	Peptide	Observed $[M+H]^+$			
			3.125 μM	**6.25 μM**	**12.50 μM**	**25 μM**
WT	1474.66	Target peptide	-	-	1474.91	1474.22
	1550.66	+ bME	-	-	1550.77	1550.22
	1486.58	*E misincorporation	1486.43	1486.79	1486.87	-
	1508.58	**E + Na adduct	-	-	1508.88	-
mS12	1474.661	Target peptide	-	1475.06	1474.42	1474.41
	1550.66	+ bME	-	1550.97	1550.43	1550.38
	1486.58	*E misincorporation	-	-	-	-
	1508.58	**E + Na adduct	-	-	-	-

Table S8: Supplementary Table for Fig 5

Ribosomes	Calculated Masses $[M+H]^+$			Observed $[M+H]^+$	
			Expected full length [M+H]+	**His**	**Flag**
WT	MHHHHHHGK-FLAG	Full length product	2179.911	**-**	**2180.29**
	MHHHHHHGERL	Triplet codon suppression	1455.65	1455.57	-
	MHHHHHH	Truncation before Gly codon	1000.40	-	-
mS12	MHHHHHHGK-FLAG	Full length product	2179.911	2179.65	**2180.00**
	MHHHHHHGERL	Triplet codon suppression	1455.65	1456.03	-
	MHHHHHH	Truncation before Gly codon	1000.40	1000.22	-

Oligos used for in vitro transcription of tRNAs:

i) $tRNA^{Val}_{GUA}$: This design was adapted from Iqbal (2018)[3]

ii) tRNA Gly CCCC:

Table S9: qtRNAGlyCCCC construction

RNA	**5'GCGGGCGUAGUUCAAUGGUAGAACGAGAGCUUCCCCAAGCUCUAUACGAGGGUUCGAUUCCCUUCGCCCGCUCCA 3'**			
	Extension primers		**1st PCR primers**	
	Forward	**Reverse**	**Forward**	**Reverse**
	GlytRNACCCfwd001	**GlytRNACCCCmid002**	**GlytRNACCCfwd001**	**GlytRNACCCrev003**
Oligo names:	**Sequences 5' to 3'**			
GlytRNACCCfwd001	GGCGTAATACGACTCACTATAGCGGGCGTAGTTCAATGGTAGAACG			
GlytRNACCCCmid002	CCCTCGTATAGAGCTTGGGGAAGCTCTCGTTCTACCATTGAAC			
GlytRNACCCrev003	T**mG**GAGCGGGCGAAGGGAATCGAACCCTCGTATAG			

iii) tRNA Gly CCC:

Table S10: tRNA GlyCCC construction

RNA	**5'GCGGGCGUAGUUCAAUGGUAGAACGAGAGCUUCCCAAGCUCUAUACGAGGGUUCGAUUCCCUUCGCCCGCUCCA 3'**			
	Extension primers		**1st PCR primers**	
	Forward	**Reverse**	**Forward**	**Reverse**
	GlytRNACCCfwd001	**GlytRNACCCmid002**	**GlytRNACCCfwd001**	**GlytRNACCCrev003**
Oligo names:	**Sequences 5' to 3'**			
GlytRNACCCfwd001	GGCGTAATACGACTCACTATAGCGGGCGTAGTTCAATGGTAGAACG			
GlytRNACCCmid002	CCCTCGTATAGAGCTTGGGAAGCTCTCGTTCTACCATTGAAC			
GlytRNACCCrev003	T**mG**GAGCGGGCGAAGGGAATCGAACCCTCGTATAG			

PCR Primers used for qtRNA and tRNA Gly DNA template construction. GlytRNACCCrev003 contains a 2' O-methylated guanine base ("mG") to reduce non-templated nucleotide addition at the 3' terminus of the DNA template to be used for tRNA transcription[4].

References:

1. Jelenc, P. C. & Kurland, C. G. Nucleoside triphosphate regeneration decreases the frequency of translation errors. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 3174–3178 (1979).
2. Bonde, M. T. *et al.* MODEST: a web-based design tool for oligonucleotide-mediated genome engineering and recombineering. *Nucleic Acids Res.* **42**, W408–W415 (2014).
3. Iqbal, E. S., Dods, K. K. & Hartman, M. C. T. Ribosomal incorporation of backbone modified amino acids via an editing-deficient aminoacyl-tRNA synthetase. *Org. Biomol. Chem.* **16**, 1073–1078 (2018).
4. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268–1272 (1999).