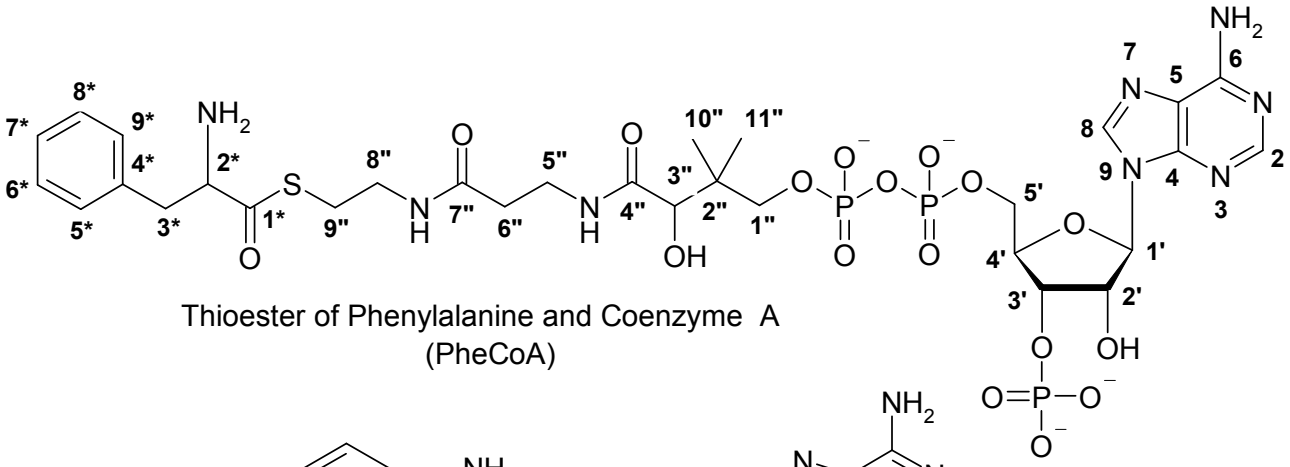
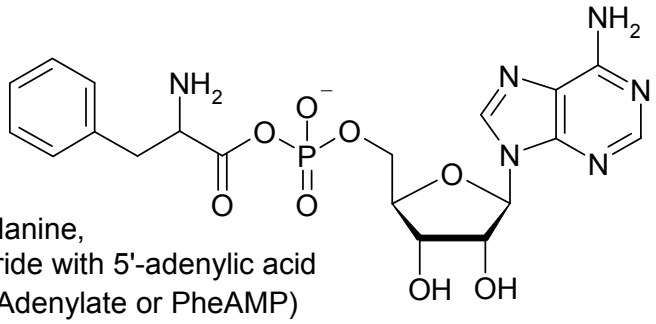
Supporting information

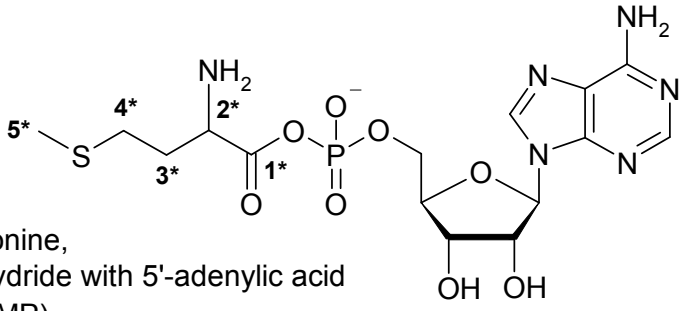
Materials and Methods.



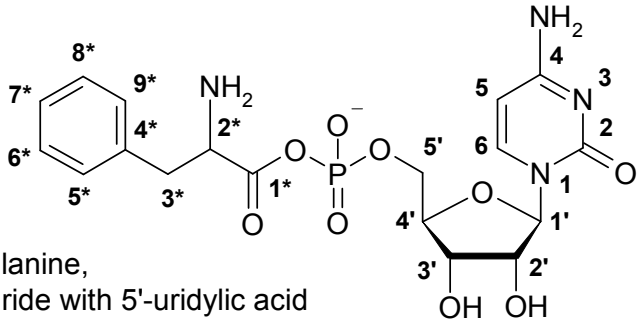
Thioester of Phenylalanine and Coenzyme A
(PheCoA)



Phenylalanine,
1-anhydride with 5'-adenylic acid
(Phenyl Adenylate or PheAMP)



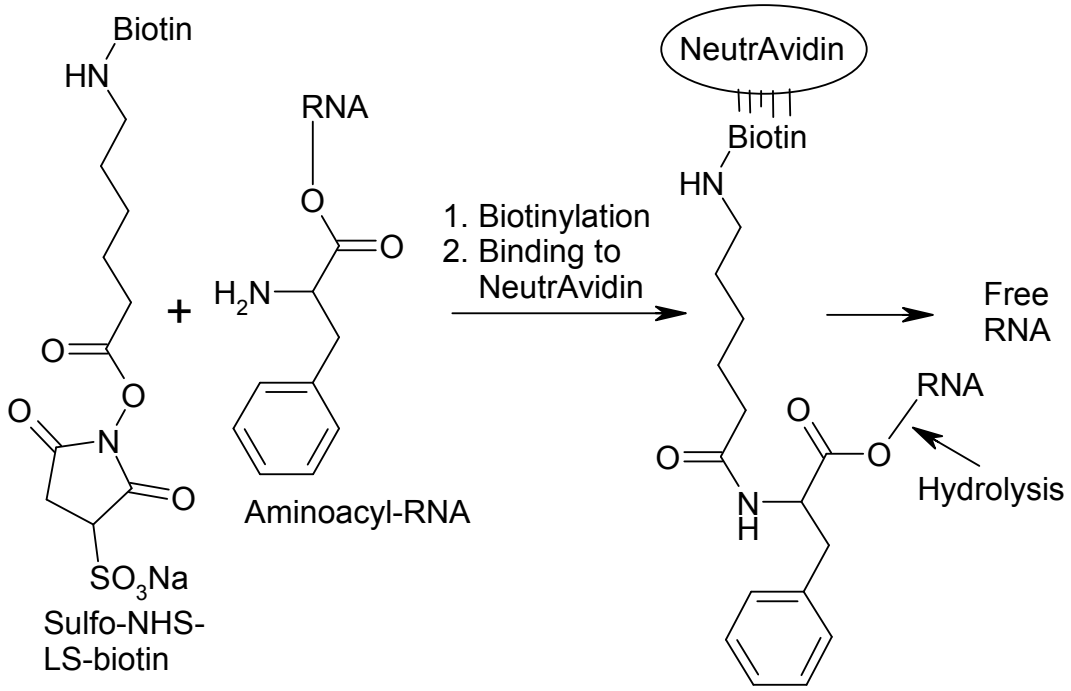
Methionine,
1-anhydride with 5'-adenylic acid
(MetAMP)



Phenylalanine,
1-anhydride with 5'-uridylic acid
(PheUMP)

Supplementary figure 1. Structures of activated derivatives of phenylalanine and methionine used in selection and assay. Nomenclature system used for the adenylic and

uridylic acids is standard. That used for the pantetheine portion of CoA is adapted from [30].



Supplementary figure 2. Selection of aminoacyl-RNA by Biotin-NeutrAvidin binding of aminoacylated, biotinylated RNA. NeutrAvidin beads retain biotinylated RNA, whereas unreacted RNA can be washed away.

The modified selection procedure.

Biotinylation reached a maximum 90% of the initial amino acid ester, reflecting partial deacylation of aminoacylated tRNAs prior to biotinylation and/or incomplete biotinylation [18].

Ligation required 15-20 times excess of the RNA-19 compared to the selected ribozymes. Reverse transcription utilized 8-9 times excess of 3′-primer (Rz-G45) over RNA-19 because their sequences can hybridize to each other. This resulted in *ca* 120-240 times excess of Rz-G45 over cDNA to be amplified by PCR-1. To achieve reasonable PCR-1 volumes, we used an unusually high (~5 μM) concentration of 3′-primer (Rz-G45) in the PCR reaction. The 5′-primer (T7-36) concentration was adjusted to ~3 μM to match.

The final RNA/HDV crossover construct (173 bp) was made by overlap extension PCR [14, 31, 32], which usually works well for *two* overlapping sequences. Two complementary strands of elongated DS DNA hybridize to each other on lowering the temperature and produce double stranded PCR product. In our random pool, once an elongated DS DNA is melted, it is unlikely for the two strands to find each other and the PCR results in a mixture of single strand DNAs. To produce a pool of DS DNA in PCR-2 both 5′-primer T7-36 and 3′-primer Rz-D32 are required.

NMR data was collected on Varian INOVA 400MHz spectrometer (deuterium resonance from the solvent was used as a standard).

L-Phenylalanine 1-anhydride with 5'-adenylic acid (L-PheAMP). Crude PheAMP was prepared by the Berg procedure [33], however the reaction mixture was filtered before the precipitation with acetone to remove N,N'-dicyclohexyl urea. The yield of dry crude product is 850 mg from 2.0 mmol of Phe. Crude product (45-50 mg) was dissolved in 0.45 mL of KH_2PO_4 (50 mM, pH 4.5), filtered and transferred to a Sep-Pak C18 Cartridge pre-saturated with water (Waters, Part N WAT051910). The cartridge was washed with 5 mL of water over ~1 min then with acetonitrile. The first 400 μL were discarded and the following 800 μL were collected, divided equally in 8 tubes, immediately frozen in dry ice, lyophilized overnight, and stored at -20 ^oC.

Immediately before use, the content of one tube was dissolved in 70 μL of cold water and filtered to give a 40-60 mM solution of PheAMP. The total concentration of adenosine in the solution was determined by absorbance at 260 nm (ε = 15400 L/mol*cm). ^1H NMR, collected in approximately 5 min after dissolving the sample in D_2O, showed a mixture of PheAMP/AMP/Phe in 1/0.1-0.3/0.2-0.4 molar ratio, and no other resonances. This mixture was used in all experiments. Selected RNA pools show no reactivity with free amino acids.

D-Phenylalanine 1-anhydride with 5'-adenylic acid (D-PheAMP). Obtained as for L-PheAMP above, in the same crude yield. ^1H NMR, collected in approximately 5 min after dissolving the sample in D_2O, showed a mixture of PheAMP/AMP/Phe in 1/0.13/0.13 molar ratio, and no other resonances. ^1H NMR (D_2O): δ 8.21 (s, 1H, H8), 7.98 (s, 1H, H2), 7.09-7.00 (m, 3H, Ph), 6.83-6.80 (m, 2H, Ph), 5.93 (d, *J* = 5.6 Hz, 1H, H1'), 4.75-4.55 (broad H_2O signal masks H2'), 4.30 (t, *J* = 4.7, 1H, H3'), 4.20-4.14 (m, 1H, H4') overlaps with 4.14 (t, *J* = 6.7 Hz, 1H, H2*), 4.04 (app. dd, *J* = 2.9, 5.2 Hz, 2H, H5'), 2.84-2.79 (m, 2H, H3*). P^{31} NMR (D_2O): δ -7.12 (PheAMP), 1.26 (AMP) in ratio 1/0.09.

Methionine, 1-anhydride with 5'-adenylic acid (PheAMP). Obtained as for L-PheAMP above, in the same crude yield. ^{1}HNMR (D_2O): δ 8.40 (s, 1H, H8), 8.24 (s, 1H, H2), 6.10 (d, *J* = 5.7 Hz, 1H, H1'), 4.75-4.55 (broad H_2O signal masks H2'), 4.48 (t, *J* = 4.6, 1H, H3'), 4.36-4.32 (m, 1H, H4'), 4.24 (t, *J* = 6.8 Hz, 1H, H2*) overlapd with 4.24-4.20 (m, 2H, H5'), 2.46 (t, *J* = 7,3 Hz, 2H, H4*), 2.02-2.19 (m, 2H, H3*), 1.88 (s, 3H, H5*). P^{31} NMR (D_2O): δ -7.07 (MetAMP), 1.26 (AMP) in ratio 1/0.25.

Phenylalanine, 1-anhydride with 5'-uridylic acid (PheUMP). Obtained as PheAMP above, in the same crude yield. ^1H NMR, collected in approximately 5 min after dissolving the sample in D_2O, showed a mixture of PheUMP/Phe in ~1/1 molar ratio, and no other resonances. ^{1}HNMR (D_2O): δ 7.59 (d, *J* = 8.1 Hz, 1H, H6), 7.27-7.07 (m, 10H, Ph PheAMP and Phe), 5.75 (d, *J* = 4.6 Hz, 1H, H1'), 5.63 (d, *J* = 8.1 Hz, 1H, H5), 4.34 (app. t, *J* = 6.6 Hz, 1H, H2* PheAMP), 4.13 (t, *J* = 4.6, 1H, H2'), 4.09-4.01 (m, 3H, H3', H4', H5'), 3.97-3.90 (m, 1H, H5'), 3.81 (dd, *J* = 5.2, 8.0 Hz, 1H, H2* Phe), 3.15 (dd, *J* = 6.1, 14.7 Hz, 1H, H3* PheAMP), 3.07 (dd, *J* = 7.2, 14.7 Hz, 1H, H3* PheAMP) overlaps

with 3.10 (dd, *J* = 5.2, 14.4 Hz, 1H, H3* Phe), 2.93 (dd, *J* = 8.1, 14.4 Hz, 1H, H3* Phe). P^{31} NMR (D_2O): δ -7.15.

Hydrolysis rate constant for PheAMP under selection conditions was determined by P^{31} NMR. Spectra were taken at 15 ^oC in an NMR tube (1 mL reaction: 100 µl 10x selection buffer, 100 µl D_2O, 800 µL H_2O) in the presence of an internal standard $(MeO)_3PO$, 8.66 mM (1 µL added to the 1 mL reaction, d = 1.213, FW = 140.08). The ratios of the integrated signals of $(MeO)_3PO$ (-4.014 ppm) and PheAMP (-7.119 ppm) gave the concentration of PheAMP (C^t_{PheAMP}). The beginning and end times of the (ca. 90 sec) scan were averaged to give the time point (t) used for calculations, and k_{HYDR} = 0.0129 min^{-1} was found by least squares fitting (R^2=0.9902) to:

$$\ln(C^t_{PheAMP}) = k_{HYDR}\, t + \ln(C^0_{PheAMP}).$$

Phe-CoA. The reaction was done according to [34]. Purified PheAMP (28 mg) and CoA Li salt (10 mg, 13 µmol) in water (200 µL) were treated with 390 µL of imidazole*HCl (400 mM, pH 7.0). After 40 min the reaction mixture was purified as described above for PheAMP (First 200 µL of acetonitrile were discarded, the following 1 mL was collected). Depending on the run, the product contained up to 0.12 equivalents of each: non-modified CoA and/or AMP and/or free Phe. Yield 7 µmol, 54%.^{1}HNMR (D_2O): δ 8.41 (s, 1H, H8), 8.10 (s, 1H, H2), 7.24-7.08 (m, 3H, Ph), 7.07-7.05 (m, 2H, Ph), 5.98 (d, *J* = 6.0 Hz, 1H, H1'), 4.70-4.48 (H_2O signal masks H2' and H3'), 4.50-4.38 (m, 1H, H4') overlaps with 4.38 (t, *J* = 7.5 Hz, 1H, H2*), 4.06 (br. s, 2H, H5'), 3.86 (s, 1H, H3"), 3.66 (dd, J = 4.6, 9.8 Hz, 1H, H1"), 3.42 (dd, J = 4.7, 9.8 Hz, 1H, H1"), 3.34-3.06 (m, 5H, H5", H8", H9"), 2.78-3.00 (m, 3H, H9" and H3*), 2.26 (t, J = 6.4, 2H, H6"), 0.75 (s, 3H, H10/11), 0.64 (s, 3H, H10/11).

DNA Oligonucleotides (Synthesized by Integrated DNA Technologies, Inc.): **50NC**, 5'-GAG ATG CCA TGC CGA CCC N_{50} AGC CTG TCA ACC GTC TTC C; **T7-36**, 5'-TAA TAC GAC TCA CTA TAG GAA GAC GGT TGA CAG GCT; **Rz-G45**, 5'- TAG CCC AGG TCG GAC CGC GAG GAG GTG GAG ATG CCA TGC CGA CCC; **Rz-A45** 5'- GGG TCG GCA TGG CAT CTC CAC CTC CTC GCG GTC CGA CCT GGG CTA; **Rz-B45**, 5'- CTT CTC CCT TAG CCT ACC GAA GTA GCC CAG GTC GGA CCG CGA GGA; **RzC21**, 5'- GGG TCG GCA TGG CAT CTC CAC; **Rz-D32**, 5'- AAA CGA CGG CCA GTG CCA AGC TTC TCC CTT AG; **Rz67**, 5'- CTT CTC CCT TAG CCT ACC GAA GTA GCC CAG GTC GGA CCG CGA GGA GGT GGA GAT GCC ATG CCG ACC C; **T7-19**, 5'- TAA TAC GAC TCA CTA TAG G; **E-18**, 5'- CTT CTC CCT TAG CCT ACC. Oligonucleotides for the assays of truncated ribozymes: 5'-TAA TAC GAC TCA CTA TA (the sequence of the truncated ribozyme) GGG TCG GCA TGG CAT CTC.

RNA Oligonucleotides (Synthesized by Dharmacon RNA Technologies, Inc.): **RNA-19**, 5'-P- GGG UCG GCA UGG CAU CUC ddC; **C3-cg**, 5'-OH- CGA CUU CGG UCG UGU GGC GAA AGC CU.

Throughout the text "P30 columns" refer to Micro Bio-Spin P30 columns in Tris buffer, Bio-Rad, Cat.# 732-6224. Filtration and NeutrAvidin bead manipulations were done using 0.2 mL Spin-X$^{(R)}$ centrifuge tube filters, Corning Inc., COSTAR, Cat.# 8161. Unless stated otherwise PAGE was done in 89 mM Tris-borate and 2 mM EDTA, pH 8.3 (10x Concentrate, Sigma, T4415).

All PCR reactions are (in mM): 20 Tris-HCl, 10 KCl, 10 $(NH_4)_2SO_4$, 2 $MgSO_4$, 0.1% Triton X-100, pH 8.8 at rt and 0.2 mM dNTPs, 0.025 U/μL of *Taq* DNA Polymerase (5 U/μL, New England BioLabs).

All transcriptions used the AmpliScribe^{TM} T7 Transcription Kit (Epicentre Biotechnologies): 1X buffer, 10 mM DTT, 7.5 mM NTPs, 7μL of T7-*Flash* Enzyme Solution/100 μL reaction. Reactions were incubated at 37 ^{o}C (2 hours for the selection cycles or 3.5 hours for RNA assays), 7 μL of RNase-Free DNase 1 (1MBU/μL) added per 100 mL reaction, incubated 40 min at 37 ^{o}C, purified through a P30 column (1 column per 50-60 μL reaction), then purified via 10% PAGE (8 M Urea).

3′-dephosphorylation reactions were done using a slight modification of [18]: up to 300 pmol of RNA was incubated in 100 mM morpholinoethanesulfonate-NaOH pH 5.5, 10 mM $MgCl_2$, 10 mM DTT, 300 mM NaCl, 14 U T4 polynucleotide kinase (New England BioLabs) and 20 U RNasin (Promega, 40U/μL) in a final volume of 20 μL for 6 h at 37 ^{o}C. The reaction mixture was diluted with sodium acetate (final 0.5 M NaOAc, pH 4.5), ethanol-precipitated and gel-purified.

HDV-Ribozyme (HDV-67) was constructed by overlap PCR-01 using DNA oligos Rz-A45 (0.5 μM) and Rz-B45 (0.5 μM) which have a 23 nt overlap: 15 cycles of 1 min/94 ^{o}C, 1 min/60 ^{o}C, 30 s/72 ^{o}C.

Elongated HDV-Ribozyme (HDV-67) was constructed/amplified by PCR-02 of HVD-67 (0.05 μM) with primers Rz-C21 (1 μM) and Rz-D32 (1 μM): 2 min/94 ^{o}C; 40 cycles of 1 min/94 ^{o}C, 1 min/56 ^{o}C, 30 s/72 ^{o}C, followed by native 8% PAGE purification.

The initial RNA-pool was constructed by two sequential PCRs followed by transcription and 3′-dephosphorylation.

PCR-1 gives DNA-131 (4 mL): 50NC (0.125 μM), 5′-T7 (1 μM), Rz-G45 (1 μM): 1.5 min/94 ^{o}C; 5 cycles of 30 s/94 ^{o}C, 2 min/68 ^{o}C, followed by native 8% PAGE purification. Yield 1150 pmol.

PCR-2 gives DNA-173 (4 mL): DNA-131 (0.1 μM), HDV-87 (0.33 μM), 5'-T7 (1 μM), Rz-D32 (1 μM): 1.5 min/94 ^{o}C; 4 cycles of 1 min/94 ^{o}C, 5 min/68 ^{o}C, followed by native 5% PAGE purification. Yield 500 pmol. PCR-2 was done twice and the products were mixed.

Transcription (2h, 100 μL) of 500 pmol of DNA-173 gave 6.5 nmol of gel-purified RNA, which was 3′-dephosphorylated, gel-purified and used in the 1st cycle (4.5 nmol).

Selection Procedure. Selection buffer (in mM): 100 Pipes, 100 NaCl, 100 KCl, 5 $CaCl_2$, 5 $MgCl_2$, 0.1 $MnCl_2$, final pH 6.4. RNA in water was incubated at 75 °C for 4 min, then Pipes (1 M, pH 6.65) and salts (10X stock solutions) were added. The mixture was left to cool at rt for 10 min, incubated at 15°C for 10 min, treated with the appropriate amounts of PheAMP and PheCoA (stock, 90 mM) solutions and incubated at 15 °C for the indicated time. The reaction was stopped by removal of aminoacylating reagents using two subsequent P30 columns, pre-saturated with water. Purified product was diluted with sodium acetate (final 0.3 M NaOAc, pH 4.5) and ethanol-precipitated.

Supplementary Table 1. Selection.

Cycle N	Time, min	Amount of RNA, nmol	Concentration			RNA recovered		RNA R-transcribed	
			RNA, μM	PheAMP, mM	PheCoA, mM	pmol	%	pmol	%
1	60	4.5	22.5	~8.0	9.0	32	0.71	22	0.49
2	60	4.6	23	~8.0	8.0	39	0.85	30	0.65
3	60	5.1	25.5	~8.5	8.6	130	2.5	105	2.1
4	60	1.27	25.4	~8.3	8.6	178	14	143	11.3
5	45	0.64	12.8	~2.8	2.9	130	20	105	16

Biotinylation. All manipulations were done in a cold room at 4 °C. The RNA pellet was dissolved in water (14 μL) and immediately treated with 6 μL of freshly prepared 36 mM aqueous solution of Sulfo-NHS-LS-biotin (Pierce), followed by Hepes (5 μL, 1 M, pH 8). The reaction was incubated for 1 h, brought to rt and immediately stopped by removal of biotinylating reagent using two subsequent P30 columns, pre-equilibrated with water. The product was diluted with sodium acetate (final 0.3 M NaOAc, pH 4.5) and ethanol-precipitated.

Binding to NeutrAvidin. UltraLink(R) Immobilized NeutrAvidin™ Protein Plus (0.6 mL of 50% slurry, Pierce) was applied to a Spin-X(R) centrifuge tube filter (Corning Inc.), filtered (1000 rcf, 50 sec) and washed on the filter using 5 x 0.5 mL of Binding buffer [50 mM Hepes, pH 7.5, 500 mM KCl, 5 mM EDTA]. The RNA pellet was dissolved in Binding buffer (200 μL) and applied to the filter, which was then shaken at rt for 40 min.

Removal of unreacted RNA. Each washing was done by applying the corresponding buffer solution (0.5 mL) on the filter, vortexing the filter for 20 sec, then removing the liquid using a table-top centrifuge (1000 rcf, 50-60 sec). Washing order: 8 x Binding buffer; 8 x Urea (3M Urea, 5 mM Hepes, 50 mM KCl, 0.5 mM EDTA, pH 7.7); 8 x H_2O; 6 x Binding buffer, 8 x H_2O, 4 x Hydrolysis buffer (100 mM Tris*HCl, pH 8.37, 0.1 mM EDTA, 5% Glycerol).

Recovering self-aminoacylated RNA[19]. The NeutrAvidin slurry was transferred into a 1.5 mL eppendorf tube using the Hydrolysis buffer. Excess buffer was removed to make the apparent volume 0.4 mL. The tube was treated with RNAsin (12 μL, Promega, 40U/μL) and incubated on a shaker at 37°C for 6 hours. The slurry was filtered, washed on filter with 4 x 50 μL of water, diluted with sodium acetate (final 0.5 M NaOAc, pH 4.5) and ethanol-precipitated.

Ligation. Reaction buffer (in mM): 50 Tris-HCl, 10 $MgCl_2$, 10 DTT, 1 ATP, 0.1 mg/mL BSA, pH 7.8 at 25 °C, 20% PEG-8000, 2U/μL T4 RNA Ligase. Selected RNA, RNA-oligo and BSA in water were incubated at 65 °C for 3 min, then cooled in ice for 3 min. Appropriate amounts of 10X T4 RNA Ligase 1 reaction buffer (New England Bio-Labs), 40% PEG and T4 RNA Ligase were added, and the reaction was incubated at 37 °C for 28-30 min. Longer reaction lowers the yield. The enzyme was removed by phenol-chloroform extraction, the product was diluted with sodium acetate (final 0.3 M NaOAc, pH 4.5) and ethanol-precipitated.

Supplementary Table 2. Ligation and Reverse transcription

Cycle #	Ligation			Reverse transcription		
	Reaction volume, μL	Concentration, μM		Reaction volume, μL	Concentration, μM	
		RNA	RNA-17		RNA	Rz-G45
1	80	0.325	9.23	80	0.275	79.8
2	80	0.45	9.23	80	0.375	79.8
3	110	1.12	16.4	200	0.525	78.9
4	150	1.13	16.4	270	0.530	79.3
5	120	0.96	16.4	210	0.500	81.8

.

Reverse Transcription. Reaction conditions (in mM): 50 Tris-HCl (pH 8.3 at rt, 75 KCl, 3 $MgCl_2$, 5 DTT, 1 dNTPs, 8 U/μL SuperScript III Reverse Transcriptase (Invitrogen, 200 U/μL). RNA was combined with DNA 3'-primer in 8.7 times molar excess in water, incubated at 85 °C for 2 min, at 42 °C for 2 min, then treated with an appropriate amount of the mixture of 5X reaction buffer, DTT, dNTPs and SS III. The reaction mixture was further incubated at 42 °C for 3 min, at 55 °C for 30 min, cooled to 37°C, treated with RNase H (Invitrogen, 2 U/μL, final concentration 0.04 U/μL), incubated at 37 °C for 20 min, at 75 °C for 15 min, cooled to rt and used for PCR-1 without further manipulations.

PCR, Transcription and 3'-Dephosphorylation of the selected sequences were done as described above.

Supplementary Table 3. PCR-1 (after the cycle).

Cycle #	ssDNA, pmol	Concentration, μM		# of cycles	Volume, mL	DNA-131, pmol
		5'-T7	Rz-G45			
1	~22	3.0	4.0	8	1.6	850
2	~30	3.0	4.0	8	1.6	870
3	~105	3.0	4.9	6	3.2	1970

4	~143	3.1	4.8	5	4.5	2060
5	~105	3.1	4.8	4	3.6	890

PCR-2 after cycles 1-5 was done as PCR-2 for the pool, 4 ml, 4 cycles, 5′-T7 (1-1.2 μM), Rz-D32 (1-1.2 μM).

Supplementary Table 4. PCR-2 and Transcription.

Cycle #	PCR-2 (after the cycle)			Transcription (after the cycle)		
	DNA-131, pmol	HDV-Rz, pmol	DNA-173, pmol	Volume, μL	DNA-173, pmol	RNA, pmol
1	400	1200	500	102	400	8100
2	420	1320	580	102	510	7800
3	640	1300	600	51	380	3600
4	640	1300	570	55	300	3300
5	450	900	600	60	200	2900

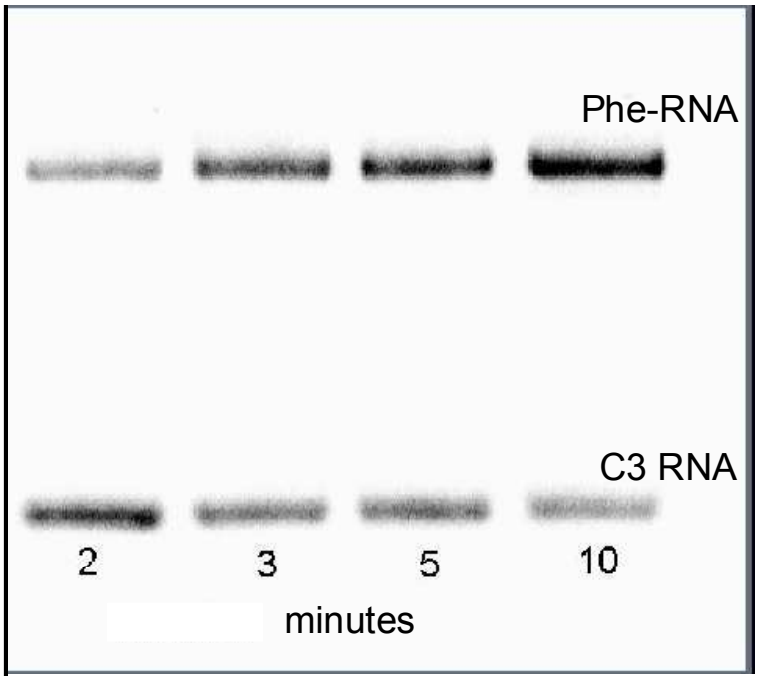
Preparation of shortened ribozymes for the assays.
PCR-3. Appropriate DNA-oligo (1 μM), Rz67 (1 μL) (18 bases overlap) in 100 μL: 9 cycles of 1 min/94 ^{o}C, 1 min/56 ^{o}C, 30 s/72 ^{o}C. PCR-4. Primer E-18 (1.75 μM), Primer T7-19 (1.75 μM), reaction mixture after PCR-3 (12 μL) in 800 μL: 2 min/94 ^{o}C, then 20 cycles of 1 min/94 ^{o}C, 1 min/55 ^{o}C, 2 min/72 ^{o}C. The reaction mixture was diluted with sodium acetate (final 0.3 M NaOAc, pH 4.5), ethanol-precipitated, dissolved in 100 μL of water, purified using two P30 columns pre-equilibrated in water, precipitated (as above). One half of the product was used for the following transcription. Transcription (50 μl) followed by dephosphorylation (200 μl) yielded, on average, 2.5 nmol of RNA.

5′-Dephosphorylation of C3 was done for 15 min at 37 ^{o}C: 50 mM Tris-HCl (pH 9.0 at 37 ^{o}C), 10 mM $MgCl_2$, 10 μM RNA, 0.15 U/μL of Shrimp Alkaline Phosphatase (Promega, 1 U/μL). The dephosphorylated product runs *below* the original RNA on 7 M Urea PAGE.

Preparation of C3-cg. The oligo RNA-cg was deprotected according to the manufacturer's procedure and gel-purified (10% PAGE, 7 M Urea). RNA-cg (100 pmol) was kinased with 7.5 pmol of [γ-^{32}P] ATP (6000 Ci/mmol) in 20 μL reaction (50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 100 mM KCl, 1 mM 2-mercaptoethanol, 0.5 U/μl of T4 PNK, Invitrogen). The reaction mixture was purified by two subsequent P30 columns, pre-equilibrated in water. Hot and cold RNA-cg were mixed to achieve reasonable specific radioactivity (1 μL of hot RNA-cg gave ~1.4*10^6 CPM for 60 μl assay).

Assays of self-aminoacylating ribozymes. From the reaction mixture of RNA (~11 mM) and PheAMP (~3 mM) in the selection buffer at 15 ^{o}C, timed aliquots (7 μL) were taken and immediately frozen in dry ice. Cold Hepes (35 μl, 1 M, pH 8.46) was mixed with freshly prepared cold solution of Sulfo-NHS-LS-biotin (70 μL, 36 mM) and used immediately. The reaction aliquots were transferred in ice, immediately treated with 11.5 μL of cold Hepes/Biotin mixture, stirred, incubated in ice for 25 min, stirred, incubated in ice for 35 min. Each aliquot was diluted with 15 μL of water and purified by P30 column.

Each purified aliquot (2.5 μL) was treated with streptavidin solution (2 μl of 1mg/1mL, New England BioLabs), incubated in ice for 10 min, diluted with 5 μL of DMF-based dyes (Bromophenol Blue & Xylene Cyanol) and assayed on 8% PAGE (6 M Urea, 0.1 M Sodium acetate, pH 4.5). Unreacted RNA migrated with Bromophenol Blue, streptavidin-bound reacted RNA ran above Xylene Cyanol. The gel was run for 1 hour, dried and quantified using a Bio-Rad Molecular Imager FX.



Supplementary Figure 3 – Gel separation of the products of self-aminoacylation of C3 RNA by 0.07 mM (L)Phe-AMP at 15° under standard assay conditions. With no incubation or no Phe-AMP, there is no observable Phe-RNA (upper) band.

Jump Kinetics

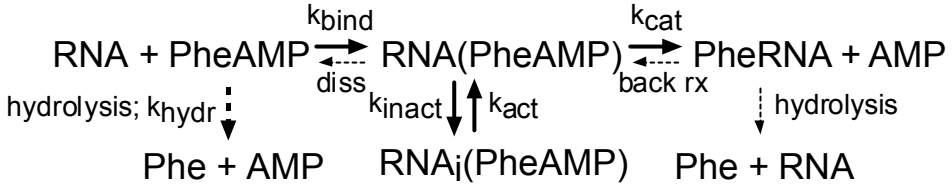
The challenge presented by these kinetics of RNA aminoacylation can be seen in the observation that PheRNA does not smoothly increase with time. Instead there is a fast reaction (a "jump") at early times, followed by a slower reaction, which sometimes carries the more rapidly acylated RNAs to near-complete aminoacylation. This is particularly evident at low concentrations of adenylate, where acylated RNA has jumped to a slower rate before the first point can be taken (≤ 20 seconds; Figure 2A). For low substrate, the early jump is smaller in magnitude. As adenylate substrate concentration is increased, the jump increases in size, ultimately saturating (not shown here). We have searched for a plausible mechanism which has these non-Michaelian properties, notably fast and slow phases whose magnitude increases with substrate concentration. RNA species responsible for the slow phase must be stable for the period of assay, to explain the persistence of the slow reaction.

This section's purpose is to show that such properties can be plausibly accounted for, without proposing extraordinary new RNA properties.

Such an accounting is particularly needed because protein enzyme jumps ("burst kinetics"; [35] are well-known, but this RNA example must be a completely different phenomenon. Protein enzyme 'bursts' are due to slow turnover of a covalent intermediate. In these RNA self-aminoacylation reactions no covalent intermediate is likely and there is no turnover at all. Therefore a novel reaction scheme is implied, and we will avoid the word 'burst', calling these 'jump' kinetics.

The simplest scheme that we have found to exhibit jump kinetics and to fit observed data, without requiring added assumptions is shown as Supplementary Scheme I. Here adenylate substrate binds rapidly and effectively irreversibly (more below) to the RNA active site (k_{bind}), which then can partition into an inactive form (k_{inact}) or alternatively, rapidly go on to the product aminoacyl-RNA (k_{cat}). The inactive form is not lost, but slowly reactivates (k_{act}), and when it does can rapidly transacylate to PheRNA. The slow reactivation is presumably a conformational change.

Discussion of the reaction mechanism is usually based on numerical integration of the system of equations for the kinetic hypothesis shown in Supplementary Scheme I:



Supplementary scheme I - the jump scheme, with additional tested reactions (see text)

$$\frac{dRNA}{dt} = -k_{bind} \times RNA \times PheAMP$$

$$\frac{dAMP}{dt} = k_{hydr} \times PheAMP + k_{cat} \times RNA(PheAMP)$$

$$\frac{dPheAMP}{dt} = -k_{hydr} \times PheAMP - k_{bind} \times PheAMP \times RNA$$

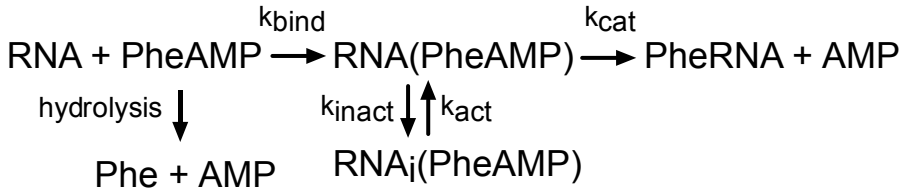
$$\frac{dRNA(PheAMP)}{dt} = k_{bind} \times RNA \times PheAMP - k_{inact} \times RNA(PheAMP) + k_{act} \times RNA_i(PheAMP)$$
$$-k_{cat} \times RNA(PheAMP)$$

$$\frac{dRNA_i(PheAMP)}{dt} = k_{inact} \times RNA(PheAMP) - k_{act} \times RNA_i(PheAMP)$$

$$\frac{dPheRNA}{dt} = k_{cat} \times RNA(PheAMP)$$

Major reactions in Supplementary scheme I – as a system of differential equations

Hydrolysis of the unstable substrate PheAMP was slowed by the use of pH 6.4 and 15° - however the adenylate has a half-life of 54 minutes, and this must sometimes be accounted for in longer kinetic experiments. The dashed arrow represents PheAMP hydrolysis, a reaction which was included in numerical fitting. The reaction scheme representing the above reactions (Supplementary Scheme I) also shows additional reactions (as thinner dotted arrows) that were tested and discarded. These reactions were either reproducibly assigned very low rates when actual kinetic data was fit, or could not be reproducibly assigned a rate constant at all. This implies that our data show no support for these reactions, and so the dotted arrows have not been included when quantitative fitting was used to get rate constants. Reactions for which our kinetic data provide no evidence are the destruction of aminoacyl-RNA product by back reaction with AMP (released by PheRNA synthesis and PheAMP hydrolysis; marked "back rx" in Supplementary Scheme I) and the spontaneous hydrolysis of the RNA aminoacyl ester (marked "hydrolysis"). The dissociation of the initial RNA complex with PheAMP (marked "diss"; makes the first step a binding equilibrium) is in a special class. Adding it does not give a reproducible fitted rate nor does it improve the overall fit to data, but it seems to help to make other rates more reproducibly fit across independent experiments. However, adding another adjustable parameter might improve the fit without itself being a valid mechanistic addition. Thus, our evidence seems equivocal about dissociation, and it has been omitted for simplicity. The result is shown below, with only the reactions that were usually used.

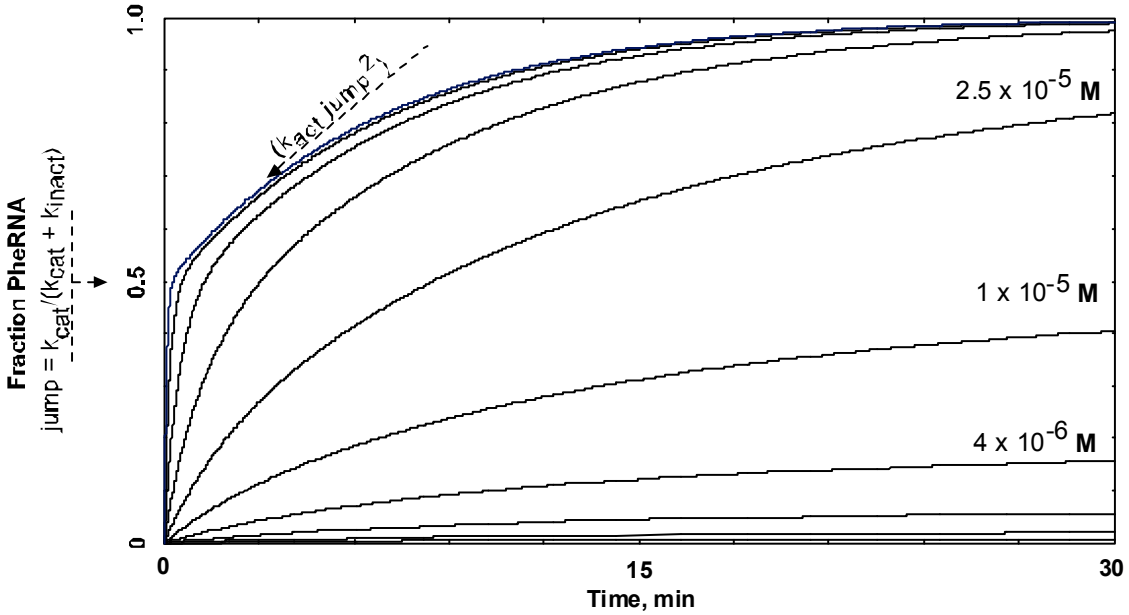


Supplementary scheme II – The simplified reaction scheme usually fit to jump kinetics.

When alternative reactions were considered (for example, hydrolysis of PheRNA) the corresponding equations and/or terms were added to the system shown above. Such systems were analyzed in TwenteSim Pro v2.3 (University of Twente / Controllab Products), which automates testing of kinetic hypotheses against numerically integrated data, and optimizes fitting of kinetic courses to suites of rate constants.

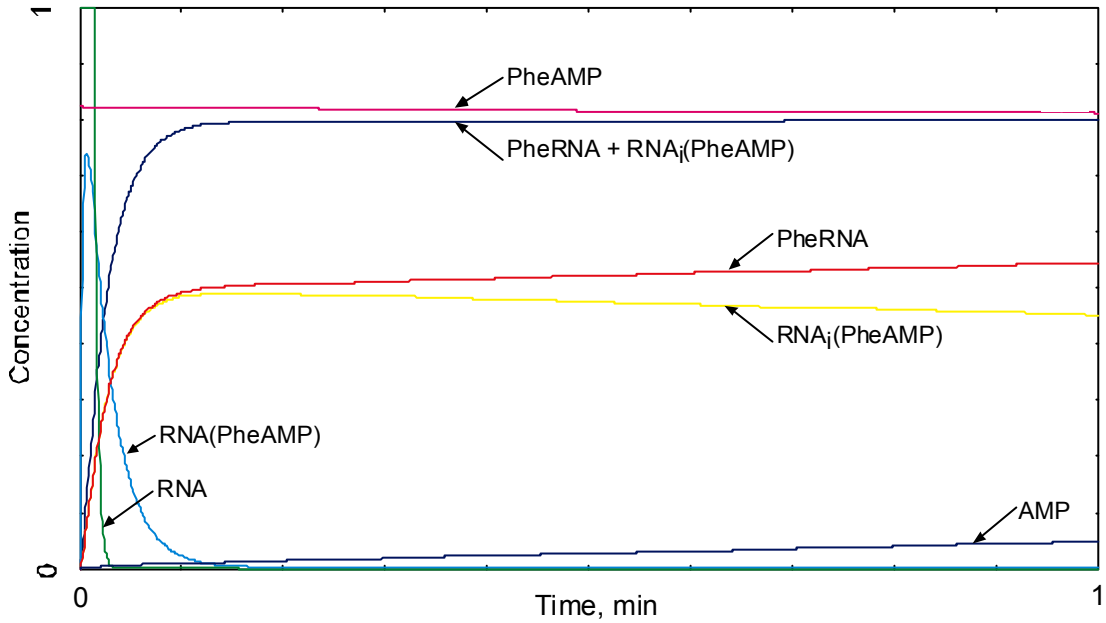
Illustration of jump kinetics

The figure below summarizes a reaction at varied PheAMP concentration using typical rates calculated by numerical integration of the system of equations for this mechanism, from low (10^{-7} M) to saturating concentrations (10^{-3} M) of PheAMP. At lower adenylate, the break to the slow reaction is less sharp and levels off to produce small steps of Phe-RNA. The partition of the common RNA(adenylate) intermediate produces a jump equal to $k_{cat}/k_{cat} + k_{inact}$. Alternatively, $k_{inact}/k_{cat} + k_{inact}$ of the starting RNA is partitioned to the inactive form RNA_i(PheAMP). After such a partition at an early time, the steady slow phase when product is derived by reactivation of RNA_i(Phe-AMP) has a rate of (k_{act} / k_{act} + k_{cat} + k_{inact}) k_{cat} (RNAo – PheRNA/RNAo), or approximately k_{act} $jump^2$, as shown in the panel (derived below). As product is progressively formed, available RNA (RNAo-PheRNA), gets smaller, so the reaction slows progressively after the jump, as seen in the Figure.



Supplementary figure 4 – Simulated C3 RNA reactions with 10 concentrations of PheAMP, increasing in 2.5-fold increments. k_{bind}= 95,600 M^{-1} min^{-1}, k_{inact} = 20 min^{-1}, k_{act} = 0.2 min^{-1}, k_{cat} = 20 min^{-1}, k_{hydr} = 0.0129 min^{-1}.

The Supplementary Figure below shows numerical integration of the above equations for the first minute of a reaction at 3.3 mM PheAMP and 12 μM C3 RNA, in order to show explicitly how jumps are produced by the above Scheme.



Supplementary Figure 5 – The first minute of the jump system in Scheme I: k_{bind} = 95,600 M^{-1} min^{-1}, k_{inact} = 20 min^{-1}, k_{act} = 0.2 min^{-1}, k_{cat} = 20 min^{-1}, k_{hydr} = 0.0129 min^{-1}. Full scale values: AMP, 0.001 M; PheAMP, 0.004 mM; PheRNA, 1.2 x 10^{-5} M (100% acylation); RNA, 1.2 x 10^{-7} M (1%); RNA(PheAMP), 1.2 x 10^{-5} M, RNA_i(PheAMP), 1.2 x 10^{-5} M, PheRNA + RNAi(PheAMP), 1.5 x 10^{-5} M.

Firstly, the implied system is fast; free RNA disappears by 2 sec (99.97 % consumed) and the system has settled into the late, slow state by 6 sec. Secondly, release of AMP from PheAMP instability and aminoacylation is small but visible, even in this first minute. Following the direct path to acylated RNA, at 0.45 sec active RNA(PheAMP) has reached its peak (74 % of RNAo), and this is quickly partitioned to product PheRNA and RNA_i(PheAMP). By 9 sec, the active RNA(PheAMP) complex has declined to very low concentration, the early jump of PheRNA has been created, and the inactive form, RNA_i(PheAMP), is at a maximum (49% of RNAo). From this point, product will come from reactivation of RNA_i(PheAMP). This is demonstrated in the complementarity of the RNA_i(PheAMP) curve and the PheRNA product curves, but particularly by the summed curve for RNA_i(PheAMP) + PheRNA. The sum is constant from this early point to the end of the reaction, manifesting interconversion of these two species.

Fitted rate constants:

k_{bind} – the rate at which RNA and PheAMP encounter each other and bind must be high enough to account for the quick early jumps in acylated RNA at low adenylate concentrations, as displayed in kinetics Figure panel A. Neither these nor other data, however, include points early enough to actually define the high early rate, as this era is over by the time of the first point (see Supplementary Figure 1). Therefore the rate is not determined but must be at least of the order 10^5 M^{-1} min^{-1}; 95,600 M^{-1} min^{-1} is used for the calculated curves above. This is fast, but still plausible without special assumptions because it is 1000 to 10000-fold slower than the diffusion-controlled reaction limit. Nevertheless, the implication is that the active site in rapidly reacting RNA is poised near the transfer conformation (Figure 4 – 5).

k_{hydr} – the hydrolysis of PheAMP is fixed at 0.0129 min^{-1} ($t_{½}$ = 54 min), determined following phosphorus NMR of PheAMP incubated at pH 6.4 and 15° under aminoacylation conditions.

k_{cat} – curve fitting for C3 RNA gives values of 11 – 22 min^{-1}. This value is fit reproducibly, but is defined less well in these experiments than the ratio of k_{cat} to k_{inact} because the jump = $k_{cat}/(k_{cat} + k_{inact})$, a crucial and obvious characteristic which can be read off kinetic courses at saturating substrate almost without analysis (Supplementary Figure 3 and 4). It is not clear whether differences between fitted k_{cat} in different experiments are due to error, or to a potentially real dependence of RNA properties on the history of a given RNA preparation.

k_{inact} – following the paragraph just above, if C3 RNA acylations seem to break in slope (rate) at 50% acylation and k_{cat} has been taken to be 20 min^{-1} (as in the simulation, Supplementary Figure 3-4) then it is implied that the first order rate constant for transition to the inactive form of the RNA must be 20 min^{-1} also. These values also (see Supplementary Figure 3) must come from an extensive curve at saturating (PheAMP)

k_{act} – curve fitting gives 0.04 to 0.19 min^{-1} for the slow reactivation of C3 RNA_i(PheAMP) that controls the late, slow phase of aminoacyl-RNA production (see Supplementary Figure 1). This corresponds plausibly to a slow conformational transition, and is derived from fitting experiments with long distinct slow phases. C3 RNA exhibits curves with a long late slow phase (Figure 2, Supp. Figure 3), which is sensitive to k_{act}. However, k_{act}, like the jump, can be gotten directly from inspection of overall kinetics.

In the late, slow phase of reaction (Supplementary Figure 4) at high (PheAMP), free RNA has disappeared, and product PheRNA is drawn from the major remaining RNA form, RNA_i(PheAMP). Conservation of RNAo (the initial RNA concentration) yields:

$$RNAo = RNA(PheAMP) + RNA_i(PheAMP) + PheRNA; RNA \rightarrow 0$$

Derivation of PheRNA from the inactive RNA pool implies:

$$-\frac{dPheRNA}{dt} = -k_{cat} \times RNA(PheAMP) = \frac{dRNA_i(PheAMP)}{dt} = k_{inact} \times RNA(PheAMP)$$
$$-k_{act} \times RNA_i(PheAMP)$$

Combining these two equations:

$$v_L = k_{cat} \times RNA(PheAMP) = (k_{act} / k_{act} + k_{inact} + k_{cat}) \times k_{cat}(RNAo - PheRNA)$$

where the subscript L has been added to the velocity to symbolize limitation to the late, slow reaction only. This expanded expression for the rate of the slow phase contains three parts, each readily intelligible. The first parenthesis is the small late steady state fraction of the RNA in active form, RNA(PheAMP), the second term is the first order rate of product production from RNA(PheAMP), (k_{cat}), and the third part, (RNAo-PheRNA) is the RNA left to react, because free RNA has disappeared and most remaining RNA exists as inactive RNA_i(Phe-AMP), whose reactivation will support the slow phase of reaction.

Taking account of activation as a slow step: $k_{act} << (k_{inact} + k_{cat})$ (ultimately required by the need for a distinct fast and slow phase of reaction - see fitted values just above) and normalizing to RNAo as is appropriate for plots of fraction reacted vs time, we rearrange the above to give:

late fractional velocity =

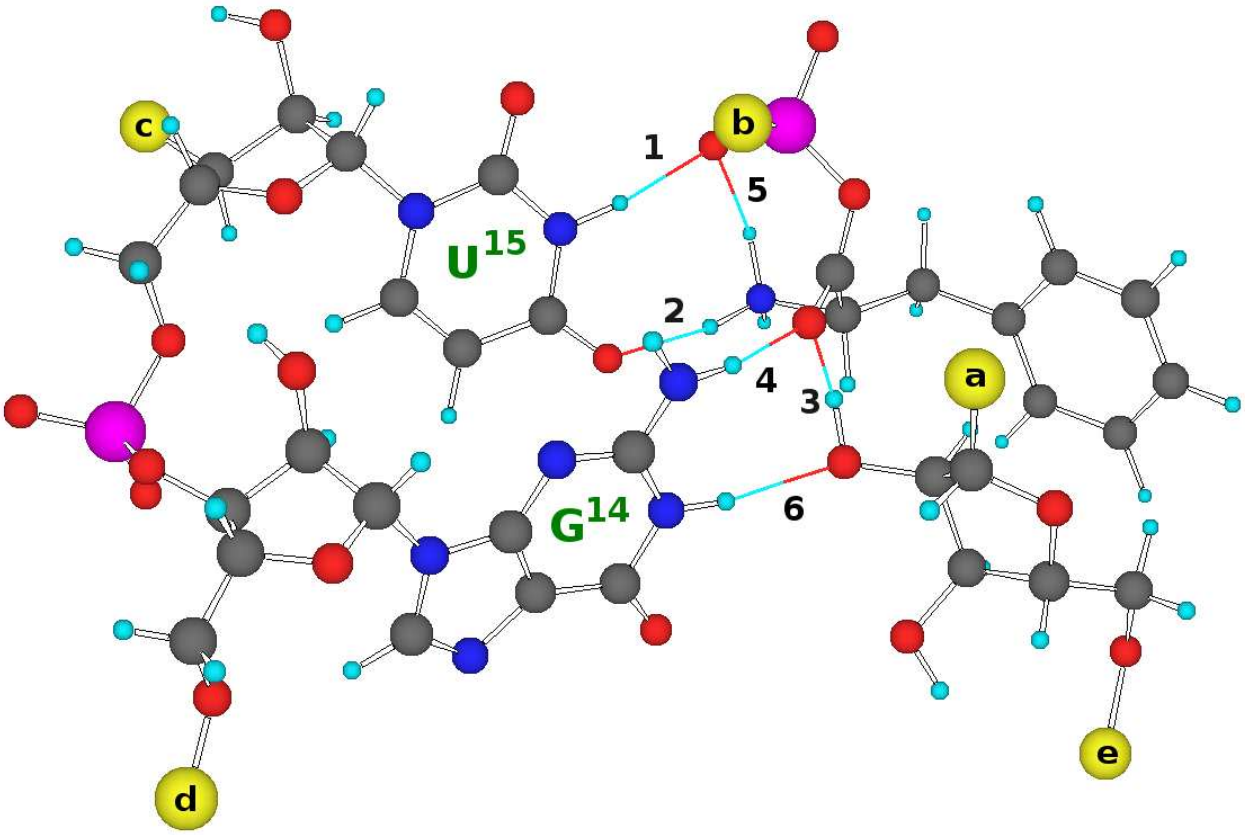
$$v_L / RNAo = k_{act} \times (\frac{k_{cat}}{k_{cat} + k_{inact}}) \times (\frac{RNAo - PheRNA}{RNAo})$$

Again, this expression is composed of three intelligible parts. The first parenthetical expression is the fractional size of the jump and, given that we are going to examine the rate at the jump, so is the second parenthetical expression. Thus given a kinetic course plotted as fraction PheRNA vs time, thought to be at high, saturating (PheAMP), one can take the slope after the break to the slow rate, divide it by the size of the jump squared, and deduce k_{act} directly from experimental data:

late fractional velocity =

$$v_L / RNAo = k_{act} \times jump^2$$

This procedure gives 0.07 to 0.18 min^{-1} for various C3 RNAs, in agreement with more complex curve fitting described above. As shown by Figure2, RNAs with mutational changes in the active site nucleotides tend to have small jumps, and then long linear slow reactions with substantial rates. This means, from the analysis just above, that mutated active sites tend to give large inactive fractions, but fast reactivation of early-forming inactive RNA_i(PheAMP). In fact, k_{act} for mutant C3 RNAs for which complete kinetics can be defined, k_{act} ranges from 0.3 to 13 min^{-1}, both faster and more variable than C3 reference RNAs. This strongly supports the natural assumption that transition from inactive to active RNA acylation should involve transconformational change of the mutated nucleotides of the active site loop.



Supplementary figure 6. The reaction model - approach of PheAMP to the terminal RNA ribose in the minimized C3 RNA reaction center: (L)-5'-Phenylalanyl adenylate, (L)-PheAMP, bound to the C3 ribozyme. Color code: gray – C, cyan – H, red – O, dark blue – N, magenta – P. Labeled yellow balls show the directions toward: a – Uridine ring at the 3'-terminus, U^{26}; b – adenosine of (L)-PheAMP; c – G^{16}; d – U^{13}, e – C^{25}. Essential model H-bonds are numbered as in the text.