PNAS

Supplementary Information for:

A distinct RNA recognition mechanism governs Np_4 decapping by RppH

Rose Levenson-Palmer, Daniel J. Luciano, Nikita Vasilyev, Ashok Nuthanakanti, Alexander Serganov, and Joel G. Belasco

Corresponding author: Joel G. Belasco
Email: joel.belasco@med.nyu.edu

This PDF file includes:

Figures S1 to S4
Tables S1 to S3

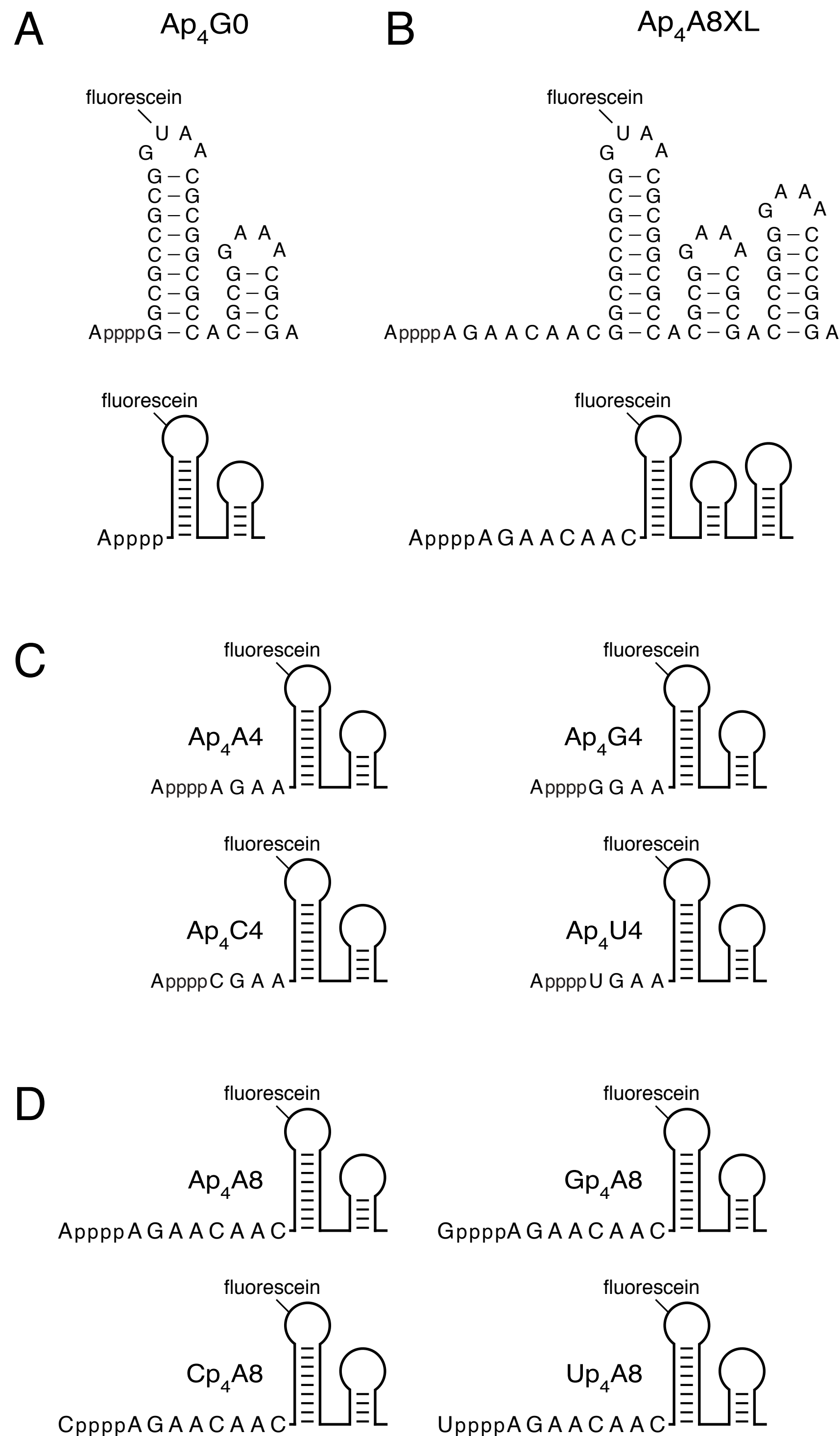

Figure S1. Sequence and expected secondary structure of substrates used to examine the specificity of decapping by RppH *in vitro*.

(A) Ap_4G0.

(B) Ap_4A8XL.

(C) Ap_4N4 RNAs, where N = A, G, C, or U.

(D) Np_4A8 RNAs, where N = A, G, C, or U.

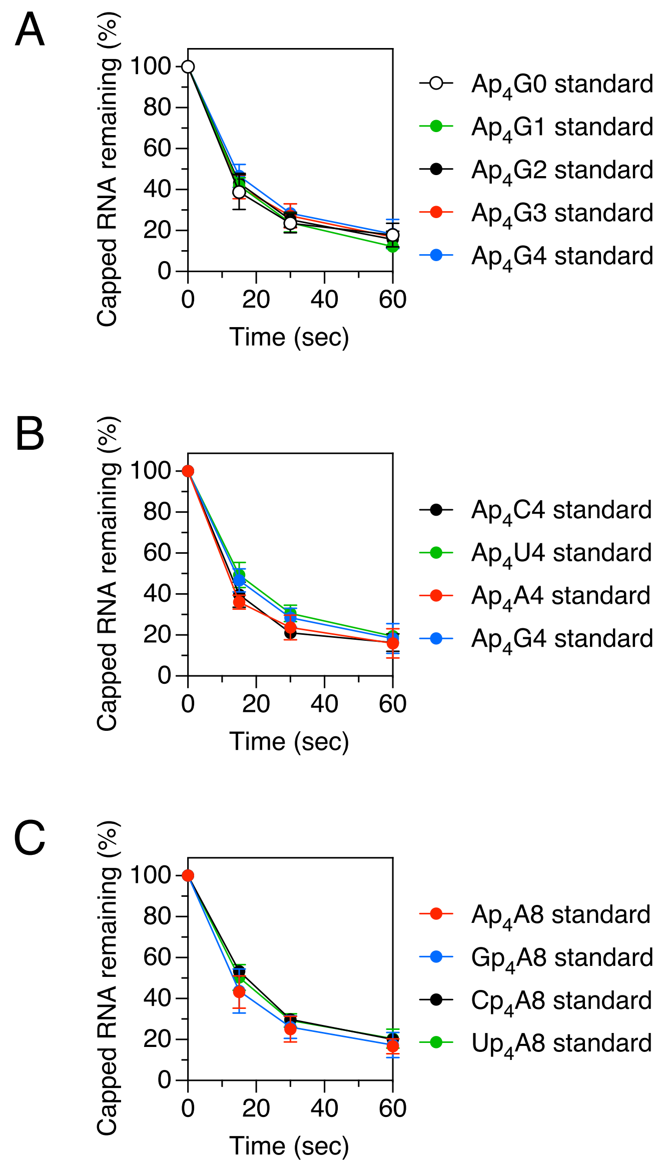

Figure S2. Decapping of the invariant internal standard Ap_4A8XL by RppH.

(A) Decapping of the internal standard in the reactions graphed in Figure 2D.

(B) Decapping of the internal standard in the reactions graphed in Figure 3A.

(C) Decapping of the internal standard in the reactions graphed in Figure 3B.

Ap_4A8XL was included as an internal standard in each reaction in Figures 2 and 3, and its decapping by RppH was monitored as in Figure 2. Each time point is the average of three or more independent measurements. Error bars correspond to standard deviations.

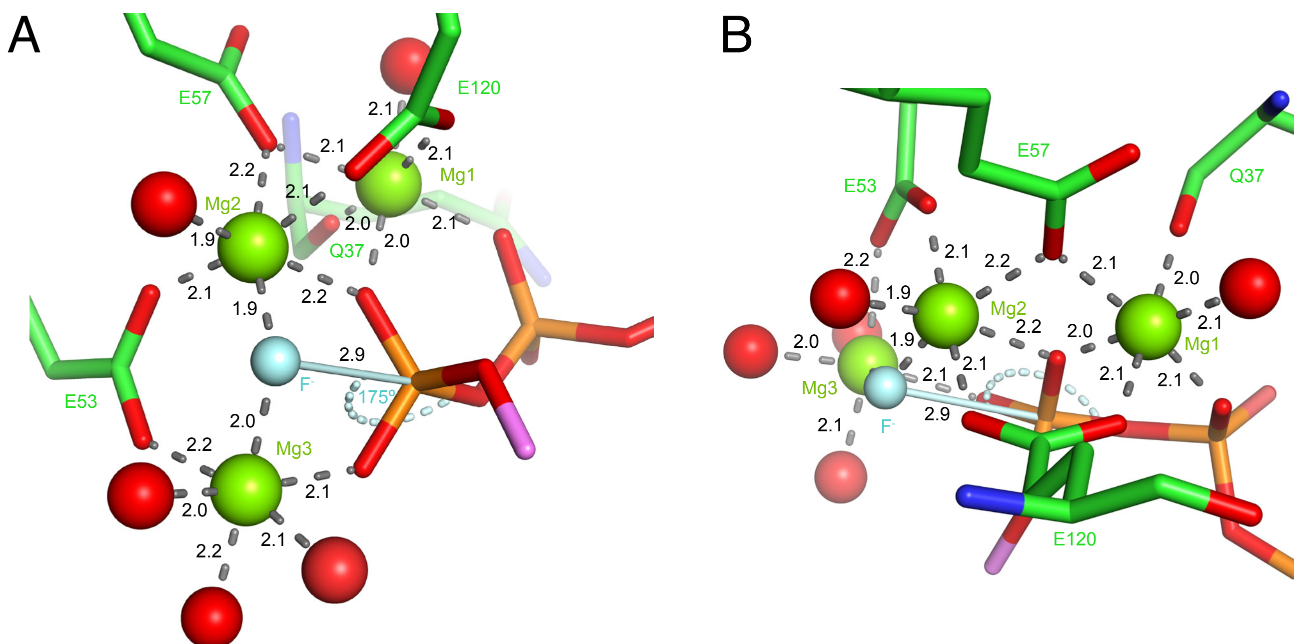

Figure S3. Detailed views of the RppH catalytic center.

Zoomed-in side (A) and top (B) views of the catalytic center of RppH bound to Ap_4A. Green, red, and cyan spheres depict Mg^{2+} ions, water molecules, and a fluoride ion, respectively. (Note that the spheres are not intended to represent the actual sizes of these ions and molecules.) The cyan stick shows the direction of nucleophilic attack by the fluoride ion on the δ phosphate of Ap_4A, and the curved cyan dashed line shows the angle of in-line attack by this anion on the δ phosphorus atom to displace the bridging oxygen atom. The gray dashed lines depict coordination bonds. Coordination distances are indicated in Å. Despite a high concentration of Na^+ ions in the crystallization solution, the large density map peaks in the active site were assigned to three Mg^{2+} ions and not to Na^+ ions on the basis of their octahedral coordination geometry and coordination distances of 1.9-2.2 Å, which are characteristic of Mg^{2+} ions. Na^+ ions typically have longer coordination bonds in the range of 2.3-2.4 Å.

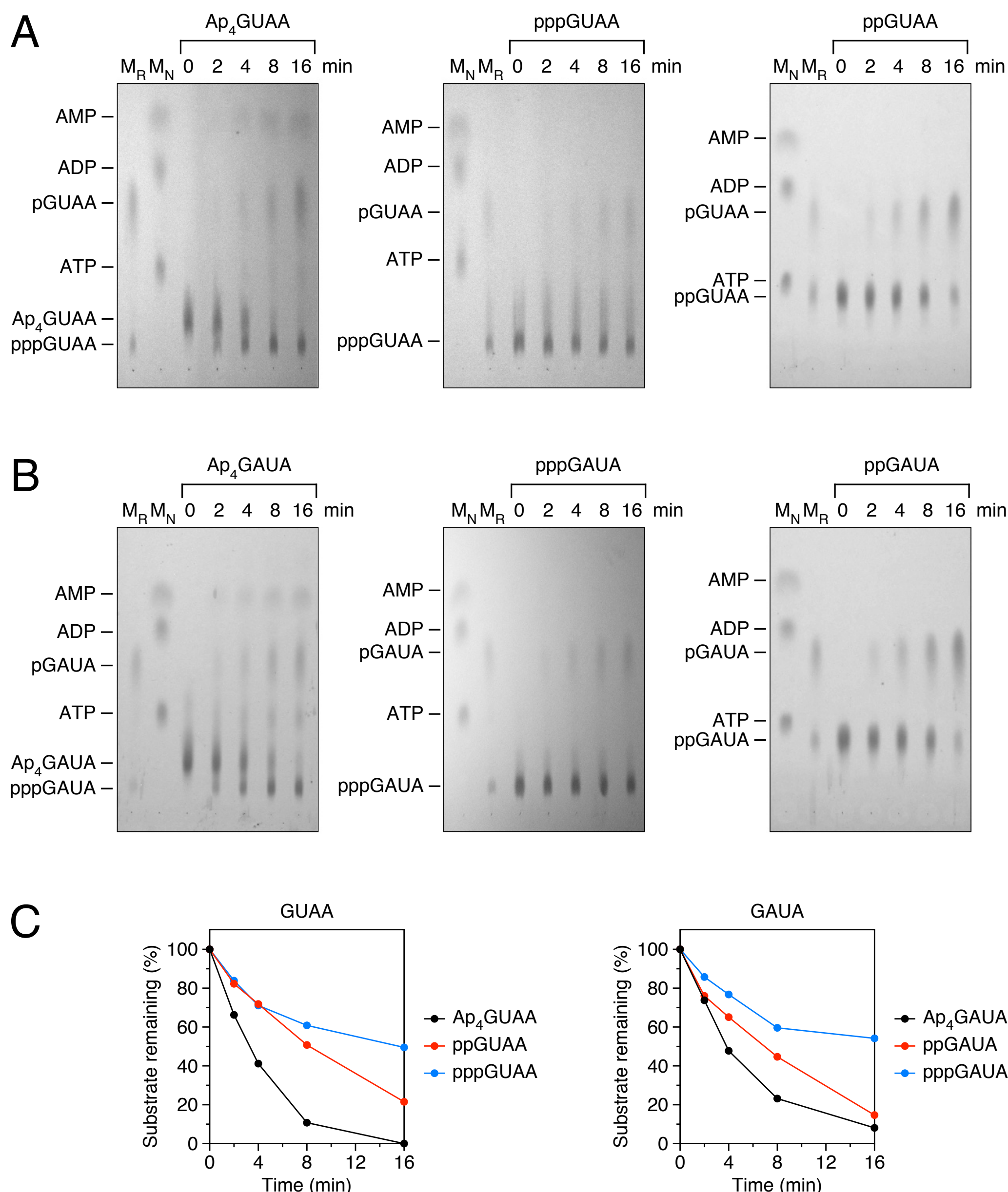

Figure S4. Comparative reactivity of Ap_4-capped, triphosphorylated, and diphosphorylated RNA substrates with RppH *in vitro*.

(A) GUAA. The reaction of Ap_4GUAA, pppGUAA, and ppGUAA with identical amounts of RppH was monitored as a function of time by thin layer chromatography on fluorescent PEI-cellulose plates and UV shadowing. Mixtures of pGUAA and either pppGUAA or ppGUAA (M_R) and of ATP, ADP, and AMP (M_N) were included as markers. Ap_4GUAA was initially converted primarily to pppGUAA, and both pppGUAA and ppGUAA were converted to pGUAA.

(B) GAUA. The reaction of Ap_4GAUA, pppGAUA, and ppGAUA with identical amounts of RppH was monitored as a function of time as in panel A, except that pGAUA, ppGAUA, and pppGAUA were substituted as RNA markers (M_R). Ap_4GAUA was initially converted primarily to pppGAUA, and both pppGAUA and ppGAUA were converted to pGAUA.

(C) Graphical comparisons of reactivity. The extent of reaction at each time point of the experiments in panels A and B was calculated from the molar ratio of substrate to RNA product(s). Differences in reactivity may have been partially obscured by the high RNA concentration required for detection by UV shadowing, which likely exceeded the K_m values of these substrates, thereby diminishing the impact of differences in binding affinity.

Table S1. First-order rate constants measured in *E. coli*.

		First-order rate constants (min^{-1})		
Capped mRNA		Δ*apaH*	Δ*apaH* Δ*rppH*	Difference
yeiP	unstressed	0.178 ± 0.006	0.064 ± 0.005	0.115 ± 0.008
	stressed	0.089 ± 0.004	0.058 ± 0.003	0.031 ± 0.004
efp variants (unstressed)	Wild-type	0.108 ± 0.003	0.074 ± 0.004	0.034 ± 0.005
	C1A	0.227 ± 0.006	0.128 ± 0.007	0.099 ± 0.009
	C1G	0.173 ± 0.011	0.084 ± 0.004	0.089 ± 0.012

Table S2. Data collection and refinement statistics for the RppH-Ap_4A complex.

Dataset	RppH-Ap_4A
Data collection	
Wavelength	1.54
Space group	*C*2
Unit Cell	
a, b, c (Å)	79.4, 36.4, 57.8
α, β, γ (º)	90.0, 102.1, 90.0
Resolution (Å)	19.40-1.60 (1.66-1.60)[a]
R_{merge} [b]	0.043 (0.587)
$CC_{1/2}$	1.000 (0.752)
$I/\sigma(I)$	22.5 (2.6)
Completeness (%)	97.5 (89.3)
Redundancy	5.2 (4.4)
No. of unique reflections	20,993 (1,890)
Refinement	
Resolution (Å)	18.8-1.6
R_{work}/R_{free} (%)	18.28/21.95
No. of atoms	
Protein	1,301
Ligand	37
Water	113
Ions	5
Average B factor ($Å^2$)	
Protein	24.8
Ligand	40.9
Water	32.1
Ions	19.7
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.957
Ramachandran analysis	
Favored (%)	99.3
Outliers (%)	0
Estimated error [c]	0.2

[a] Highest resolution shell (in Å) shown in parentheses.
[b] $Rmerge=\sum_{hkl}\sum_{i=1}^{n}|Ii(hkl)-\bar{I}(hkl)|/\sum_{hkl}\sum_{i=1}^{n}Ii(hkl)$, where Ii (*hkl*) is the *i*th observation of reflection *hkl* and $\bar{I}$(*hkl*) is the weighted average intensity for all *i* observations of reflection *hkl*.
[c] Estimated coordinate error based on maximum likelihood was calculated by Phenix.refine.

Table S3. Oligonucleotides used in these studies.

Name	Sequence [a]	Purpose
DZyeiP69	GTAATTCAGTAGGCTAGCTACAACGACATACCTTTT	10-23 deoxyribozyme for cleaving *yeiP* mRNA 69 nucleotides from the 5′ end
DZefp87	TCTAACATGAGGCTAGCTACAACGATTTAAGACCA	10-23 deoxyribozyme for cleaving *efp*, *efp*-C1A, and *efp*-C1G mRNA 87 nucleotides from the 5′ end
yeiP probe	TTCGTTCGCTCTTGGCATCG	Northern blot probe for *yeiP* mRNA
efp probe	ACGTTGCCATAAGGCCCTCT	Northern blot probe for *efp*, *efp*-C1A, and *efp*-C1G mRNA
A8-5	AATTCCTGCAGTAATACGACTCACTATTAGAACAACG GCG	Template for Ap_4A8, Gp_4A8, Cp_4A8, Up_4A8, and Ap_4A8XL
A8-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCGTTGTTCT [b]	Template for Ap_4A8, Gp_4A8, Cp_4A8, and Up_4A8
A8XL-3	mUmCCCGGGTTTCCCCGGGTCGCGTTTCCGCGTGGCG CCGCGTTACCGCGGCGCCGTTGTTC [b]	Template for Ap_4A8XL
A4-5	AATTCCTGCAGTAATACGACTCACTATTAGAAGGCGC	Template for Ap_4A4
A4-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCTTCTAATA [b]	Template for Ap_4A4
G4-5	AATTCCTGCAGTAATACGACTCACTATAGGAAGGCGC	Template for Ap_4G4
G4-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCTTCCTATA [b]	Template for Ap_4G4
C4-5	AATTCCTGCAGTAATACGACTCACTATACGAAGGCGC	Template for Ap_4C4
C4-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCTTCGTATA [b]	Template for Ap_4C4
U4-5	AATTCCTGCAGTAATACGACTCACTATATGAAGGCGC	Template for Ap_4U4
U4-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCTTCATATA [b]	Template for Ap_4U4
G3-5	AATTCCTGCAGTAATACGACTCACTATAGGAGGCGC	Template for Ap_4G3
G3-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCTCCTATAG [b]	Template for Ap_4G3
G2-5	AATTCCTGCAGTAATACGACTCACTATAGGGGCGC	Template for Ap_4G2
G2-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCCCTATAGT [b]	Template for Ap_4G2
G1-5	AATTCCTGCAGTAATACGACTCACTATAGGGCGC	Template for Ap_4G1

G1-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCCTATAGTG [b]	Template for Ap_4G1
G0-5	AATTCCTGCAGTAATACGACTCACTATAGGCGC	Template for Ap_4G0
G0-3	mUmCGCGTTTCCGCGTGGCGCCGCGTTACCGCGGCG CCTATAGTGAG [b]	Template for Ap_4G0
GUAA	GGTTACAATAGTGAGTCGTATTACTG	Template for Ap_4GUAA, pppGUAA, ppGUAA
GAUA	GGTATCAATAGTGAGTCGTATTACTG	Template for Ap_4GAUA, pppGAUA, ppGAUA
P_{T7}	CAGTAATACGACTCACTATT	Template for Ap_4GUAA, Ap_4GAUA, etc.

[a] All oligonucleotide sequences are written 5′ to 3′.
[b] mU or mC, 2′-O-methyl nucleotide.