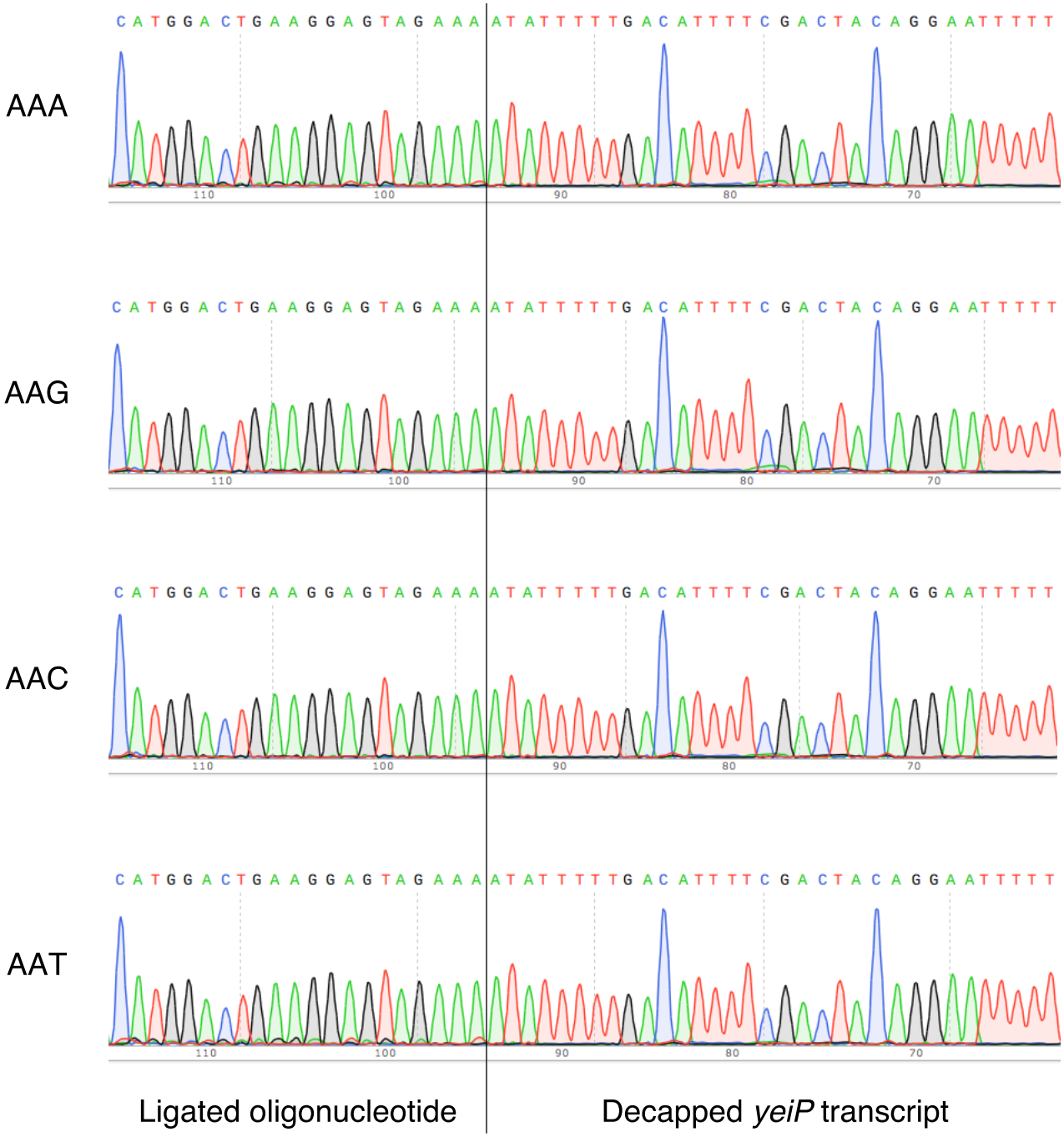


www.pnas.org/cgi/doi/10.1073/pnas.1914229117

Figure S1. Uniformity of the capped 5′ terminus of *yeiP* mRNA in Δ*apaH* cells, irrespective of the preceding promoter nucleotide.

Total RNA was extracted from unstressed Δ*apaH* cells that lacked the chromosomal *yeiP* gene and instead contained a plasmid-borne *yeiP* gene in which the sequence of the coding strand at positions –3, –2, and –1 of the promoter had been changed to AAN, where N = A, G, C, or T. The RNA was treated with alkaline phosphatase to convert uncapped 5′ ends to unligatable hydroxyls, and the capped 5′ ends were subsequently converted to monophosphates by treatment with RppH and joined to a chimeric DNA-RNA oligonucleotide by T4 RNA ligase. The ligation products were reverse transcribed, and the resulting cDNA was amplified by PCR with primers surrounding the transcription initiation site and sequenced. A portion of each cDNA sequence chromatogram is shown. Vertical line, boundary between the sequence of the decapped transcript and the oligonucleotide to which it was ligated.

Figure S2

A

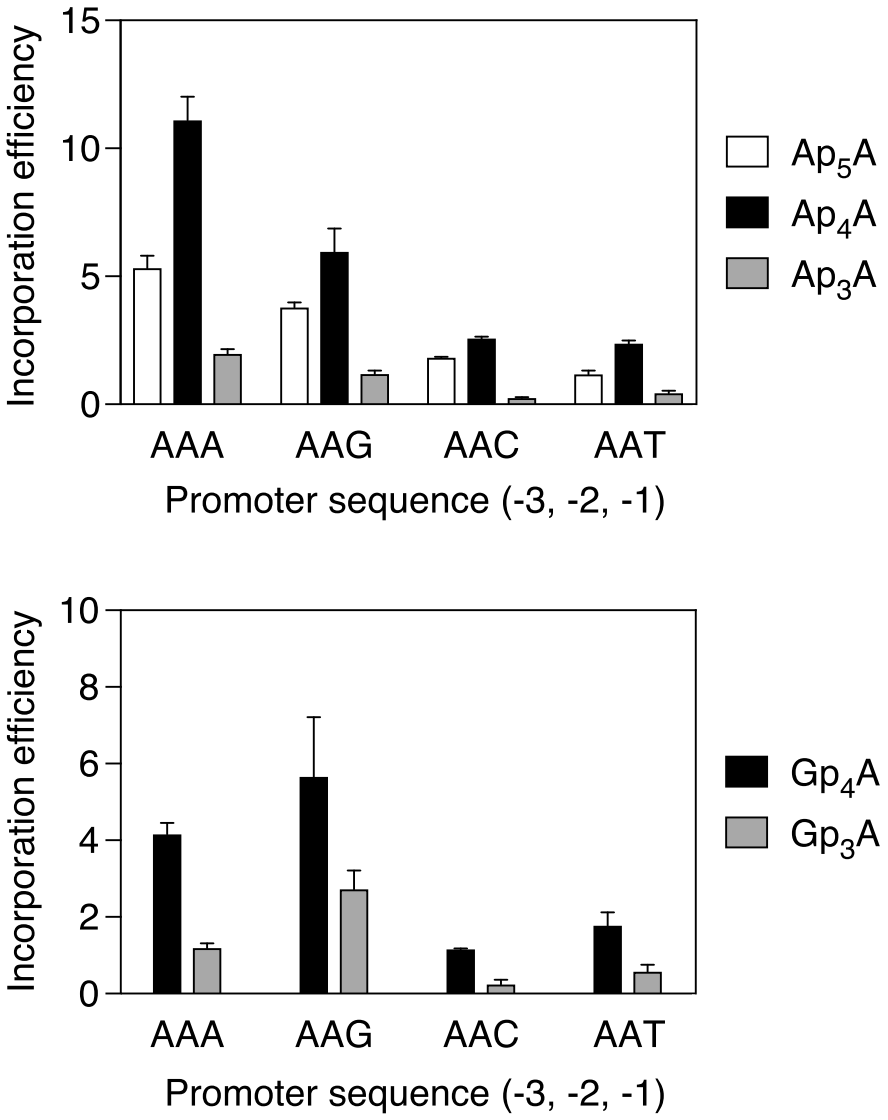
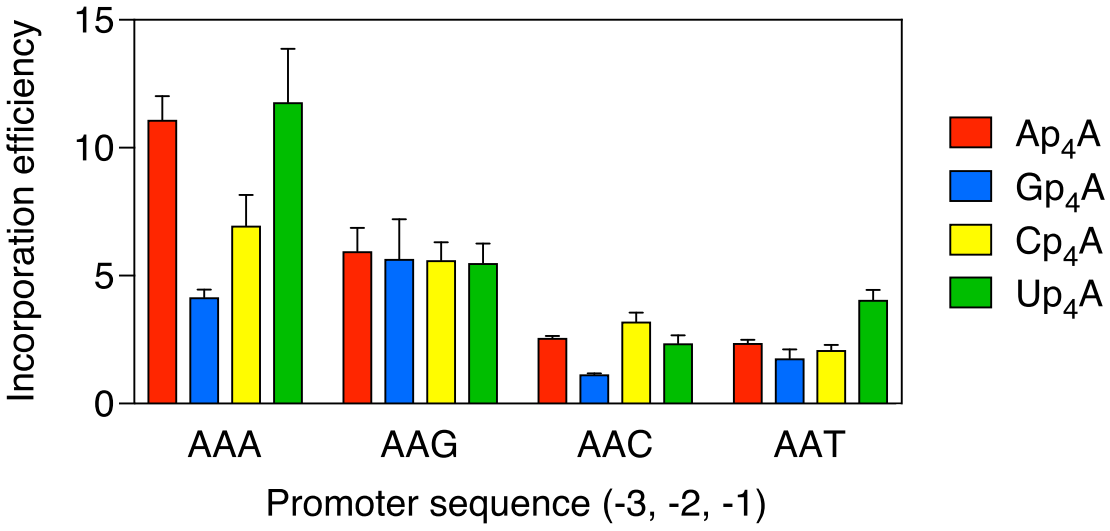
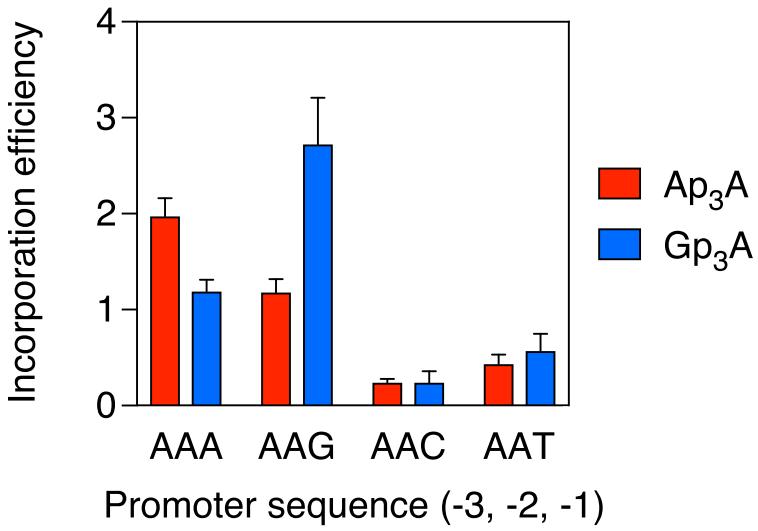

Incorporation efficiency
15
10
5
0
AAA
AAG
AAC
AAT
Promoter sequence (-3, -2, -1)
Ap5A
Ap4A
Ap3A
Incorporation efficiency
10
8
6
4
2
0
AAA
AAG
AAC
AAT
Promoter sequence (-3, -2, -1)
Gp4A
Gp3A

Figure S2

B



C



D

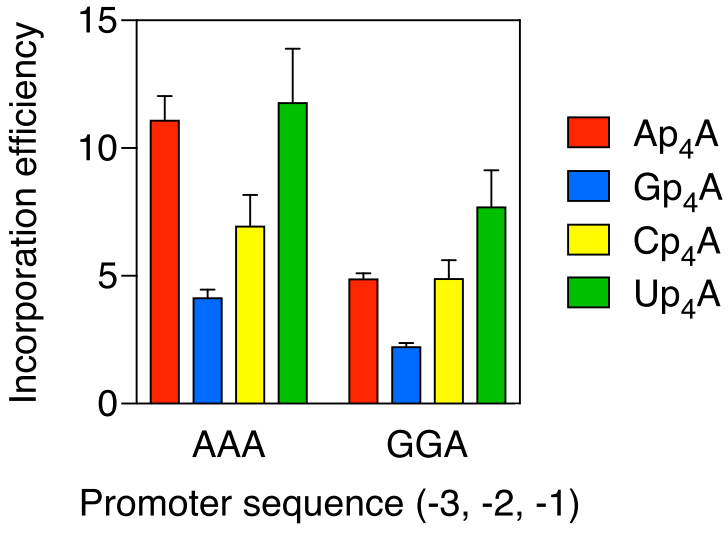


Figure S2. Efficiency of dinucleoside polyphosphate incorporation by purified *E. coli* RNA polymerase at an Np_nA:ATP ratio of 0.25:1.

(A) Effect of the number of bridging phosphates on the efficiency of dinucleoside polyphosphate incorporation.

(B) Comparative efficiency of Ap_4A, Gp_4A, Cp_4A, and Up_4A incorporation.

(C) Comparative efficiency of Ap_3A and Gp_3A incorporation.

(D) Effect of the promoter sequence at positions –3 and –2 on the efficiency of Ap_4A, Gp_4A, Cp_4A, and Up_4A incorporation.

RNA was synthesized *in vitro* by *E. coli* RNA polymerase holoenzyme in the presence of a 0.25:1 mixture of each dinucleoside polyphosphate and ATP. The radiolabeled products were then cleaved site-specifically with a 10-23 deoxyribozyme (1) and analyzed by boronate gel electrophoresis and autoradiography (2). The DNA sequence at positions –3, –2, and –1 of the coding strand of the promoter was AAN or GGA, followed by A at position +1 (the first transcribed nucleotide). In each case, the product ratio (capped transcripts/uncapped transcripts) was determined by boronate gel electrophoresis and normalized to the substrate ratio (Np_4A/ATP = 0.25). Each value corresponds to the mean of three technical replicates. Error bars represent one standard deviation.

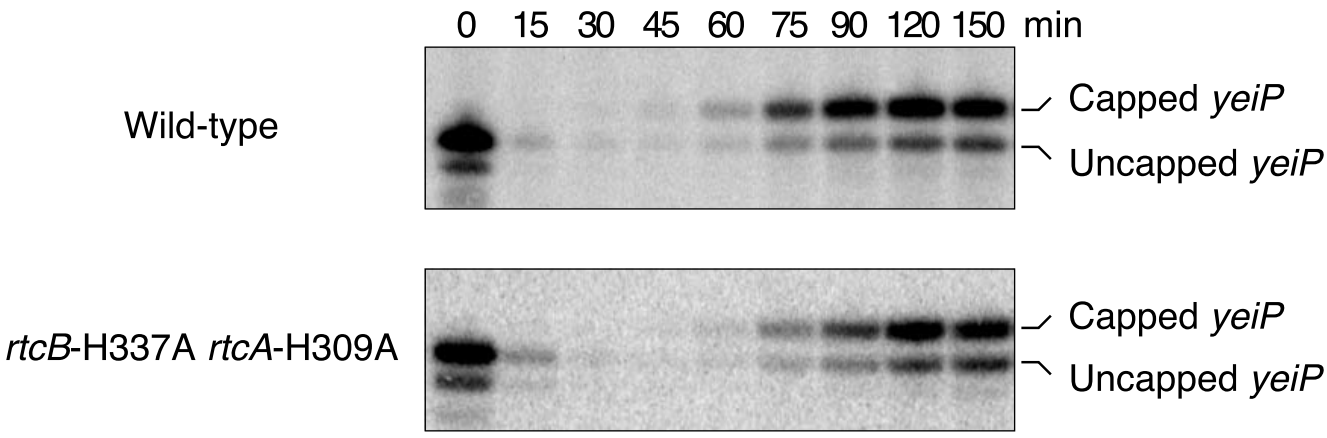


Figure S3. Insensitivity of Np_4 capping to RtcB and RtcA inactivation in cadmium-stressed *E. coli*.

Wild-type *E. coli* BW25113 and an isogenic derivative bearing chromosomal *rtcB*-H337A and *rtcA*-H309A mutations introduced by allelic exchange were grown to mid-log phase at 37°C in MOPS-glucose medium. Cadmium chloride (0.2 mM) was added to induce disulfide stress, and cellular RNA was extracted at time intervals thereafter. *yeiP* mRNA was visualized by boronate gel electrophoresis and northern blotting.

Table S1. Promoter sequences of the capped transcripts examined in *E. coli*.

Promoter	Sequence [a]
yeiP (AA) [b]	TACTTTGTCAANATA
yeiP-A1G (AA) [b]	TACTTTGTCAANGTA
yeiP-A1G (GG) [b]	TACTTTGTCGGNGTA
slyB P1 [c]	TATGATTTGCTNATA
slyB P2 [c]	TATAATTGGTTNATC

[a] The sequence of the coding strand is shown, starting with the –10 promoter region and ending with the third transcribed nucleotide. The nucleotides at positions –3, –2, and –1 are highlighted in red, and the transcribed region is highlighted in blue. The identity of the nucleotide at position –1 was varied (N = A, G, C, or T).
[b] The sequence of these three promoters at positions –3 and –2 (AA or GG) differs from that of wild-type *yeiP* (GC).
[c] The sequence of these two promoters at positions –3 and –2 (CT or TT) is the same as that of wild-type *slyB* P1 and *slyB* P2, respectively.

Table S2. Oligonucleotides used in these studies.

Name	Sequence [a]	Purpose
DZyeiP69	GTAATTCAGTAGGCTAGCTACAACGAC ATACCTTTT	10-23 deoxyribozyme for cleaving *yeiP* and *yeiP*-A1G mRNA 69 nucleotides from the 5′ end
DZyeiP573	GCGTTCTTCGAGGCTAGCTACAACGAA TGGATACGA	10-23 deoxyribozyme for cleaving *in vitro* transcription products of *E. coli* RNA polymerase 106 nucleotides from the 5′ end
yeiP probe	TTCGTTCGCTCTTGGCATCG	Northern blot probe for *yeiP* and *yeiP*-A1G mRNA
DZslyB128	TACGCGTTTAAGGCTAGCTACAACGAC ATTGAAAAA	10-23 deoxyribozyme for cleaving *slyB* P1 and *slyB* P2 mRNA 103 or 128 nucleotides, respectively, from the the 5′ end
slyB probe	GGCATATGTGAATCCATGGTTAC	Northern blot probe for *slyB* P1 and *slyB* P2 mRNA
RACE 1	CGACTGGAGCACGAGGACACTGACAT GGACTGAAGGAGTAGrArArA [b]	DNA-RNA chimera used for 5′ RACE ligation
RACEplus	CGACTGGAGCACGAGGACACTGA	PCR primer for first round of nested PCR for 5′ RACE
RACEnest	GGACACTGACATGGACTGAAGGAGTA	PCR primer for second round of nested PCR for 5′ RACE
yeiP 444 rev	TTTCCAGATCAACGGTCTGCG	Primer for reverse transcription and first round of nested PCR for 5′ RACE of *yeiP* mRNA
yeiP 292 rev	TCCATAAAGACATATTCGTTGCC	PCR primer for second round of nested PCR for 5′ RACE of *yeiP* mRNA
yeiP 128 rev	AGTGGGCGACTGAATATC	Sequencing primer for 5′ RACE of *yeiP* mRNA

[a] All oligonucleotide sequences are written from 5′ to 3′.
[b] rA, adenylate ribonucleotide.

Table S3. Percentage of *yeiP* transcripts that are Np_4-capped in unstressed Δ*apaH* cells.

yeiP (% capped)			
AAA	AAG	AAC	AAT
43.6 ± 1.1	38.2 ± 3.3	14.4 ± 2.0	21.0 ± 2.7

AAA, AAG, AAC, and AAT are the sequence of the coding strand at positions –3, –2, and –1 of the *yeiP* promoter. Each value corresponds to the mean and standard deviation of four or five biological replicates. These data are graphed in Figure 2A.

Table S4. Percentage of *yeiP*-A1G transcripts that are Np_4-capped in unstressed Δ*apaH* cells.

yeiP-A1G (% capped)							
AAA	AAG	AAC	AAT	GGA	GGG	GGC	GGT
9.5 ± 0.8	7.9 ± 0.3	5.4 ± 0.6	6.1 ± 0.7	6.8 ± 1.3	5.5 ± 0.2	4.3 ± 0.9	4.7 ± 0.6

AAA, AAG, AAC, AAT, GGA, GGG, GGC, and GGT are the sequence of the coding strand at positions –3, –2, and –1 of the *yeiP* promoter. Each value corresponds to the mean and standard deviation of five biological replicates. These data are graphed in Figure 2B.

Table S5. Percentage of *slyB* P1 and P2 transcripts that are Np_4-capped in unstressed Δ*apaH* cells.

slyB P1 (% capped)				*slyB* P2 (% capped)			
CTA	CTG	CTC	CTT	TTA	TTG	TTC	TTT
40.1 ± 3.5	38.8 ± 0.4	22.5 ± 1.0	24.4 ± 0.4	39.8 ± 2.1	14.9 ± 0.8	17.9 ± 1.3	15.7 ± 1.4

CTA, CTG, CTC, CTT, TTA, TTG, TTC, and TTT are the sequence of the coding strand at positions –3, –2, and –1 of the *slyB* P1 and P2 promoters. Each value corresponds to the mean and standard deviation of five biological replicates. These data are graphed in Figure 2C.

Table S6. Absolute efficiency of Np_nA incorporation (Np_nA/ATP = 0.5 or 2.0).

	Ap_5A	Ap_4A	Gp_4A	Cp_4A	Up_4A	Ap_3A	Gp_3A
AAA	4.01 ± 0.50	7.85 ± 0.85	3.92 ± 0.64	5.23 ± 0.17	8.73 ± 1.36	1.46 ± 0.05	0.93 ± 0.06
AAG	2.95 ± 0.02	5.22 ± 0.16	4.59 ± 0.18	5.26 ± 0.47	4.98 ± 0.05	0.89 ± 0.07	1.80 ± 0.07
AAC	1.50 ± 0.11	2.25 ± 0.10	1.13 ± 0.07	2.67 ± 0.08	2.00 ± 0.02	0.14 ± 0.01	0.17 ± 0.01
AAT	1.08 ± 0.06	2.14 ± 0.08	1.44 ± 0.01	1.93 ± 0.03	3.34 ± 0.09	0.25 ± 0.02	0.30 ± 0.01
GGA	n.d.	4.79 ± 0.30	2.30 ± 0.06	4.49 ± 0.12	7.35 ± 0.60	n.d.	n.d.

AAA, AAG, AAC, AAT, and GGA are the sequence of the coding strand at positions –3, –2, and –1 of the DNA template used for *in vitro* transcription with *E. coli* RNA polymerase σ^{70} holoenzyme. The product ratio (capped transcripts/uncapped transcripts) was normalized to the substrate ratio (Np_nA/ATP = 0.50 for Ap_5A, Ap_4A, Gp_4A, Cp_4A, and Up_4A and 2.00 for Ap_3A and Gp_3A). Each value corresponds to the mean and standard deviation of three technical replicates. n.d., not determined. These data are graphed in Figures 3B, 4A, 5, and 6.

Table S7. Relative efficiency of Np_4A incorporation (Np_4A/ATP = 0.5).

	AAA	AAG	AAC	AAT
Ap_4A	1.22 ± 0.12	1.04 ± 0.04	1.12 ± 0.04	0.97 ± 0.03
Gp_4A	0.61 ± 0.09	0.91 ± 0.04	0.56 ± 0.03	0.65 ± 0.01
Cp_4A	0.81 ± 0.06	1.05 ± 0.07	1.33 ± 0.03	0.87 ± 0.02
Up_4A	1.36 ± 0.15	0.99 ± 0.03	0.99 ± 0.02	1.51 ± 0.03

AAA, AAG, AAC, and AAT are the sequence of the coding strand at positions –3, –2, and –1 of the DNA template used for *in vitro* transcription with *E. coli* RNA polymerase σ^{70} holoenzyme. The efficiencies of Ap_4A, Gp_4A, Cp_4A, and Up_4A incorporation were each normalized to the mean efficiency of all four with the same promoter. The relative efficiencies calculated in this manner can range from 0 to 4, where a relative efficiency of 1 is average for that promoter. Each value was calculated by using data from three technical replicates for each of four Np_4As and a particular promoter. Errors correspond to a confidence level of 68.3%, equivalent to one standard deviation. These data are graphed in Figure 4B.

Table S8. Absolute efficiency of Np_nA incorporation (Np_nA/ATP = 0.25).

	Ap_5A	Ap_4A	Gp_4A	Cp_4A	Up_4A	Ap_3A	Gp_3A
AAA	5.32 ± 0.49	11.08 ± 0.92	4.15 ± 0.30	6.95 ± 1.20	11.78 ± 2.09	1.97 ± 0.19	1.19 ± 0.12
AAG	3.78 ± 0.20	5.94 ± 0.91	5.65 ± 1.56	5.61 ± 0.69	5.49 ± 0.77	1.18 ± 0.14	2.72 ± 0.49
AAC	1.82 ± 0.03	2.56 ± 0.07	1.15 ± 0.03	3.20 ± 0.35	2.35 ± 0.32	0.24 ± 0.04	0.24 ± 0.12
AAT	1.17 ± 0.15	2.35 ± 0.13	1.77 ± 0.35	2.09 ± 0.20	4.06 ± 0.38	0.43 ± 0.10	0.57 ± 0.18
GGA	n.d.	4.88 ± 0.20	2.24 ± 0.13	4.91 ± 0.70	7.70 ± 1.42	n.d.	n.d.

AAA, AAG, AAC, AAT, and GGA are the sequence of the coding strand at positions –3, –2, and –1 of the DNA template used for *in vitro* transcription with *E. coli* RNA polymerase σ^{70} holoenzyme. The product ratio (capped transcripts/uncapped transcripts) was normalized to the substrate ratio (Np_nA/ATP = 0.25 for all Np_nAs). Each value corresponds to the mean and standard deviation of three technical replicates. n.d., not determined. These data are graphed in Figure S2.

References

1. S. W. Santoro, G. F. Joyce, A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci. USA* **94**, 4262-4266 (1997).
2. D. J. Luciano, J. G. Belasco, Analysis of RNA 5' ends: Phosphate enumeration and cap characterization. *Methods* **155**, 3-9 (2019).