Supplementary Information

MATERIAL AND METHODS

Preparation of T7 RNA polymerase variants, DNA fragments, and unnatural triphosphates

Wild-type T7 RNA polymerase (WT-T7RNAP) and the T7 RNA polymerase (T7RNAP) variants used in this study were prepared as described previously.[1] DNA fragments were chemically synthesized with an Oligonucleotide Synthesizer nS-8 (Gene Design), using phosphoramidites of the natural and Ds bases (Glen Research), and purified by denaturing gel electrophoresis. Unnatural nucleobase substrates, PaTP, Pa'TP, Bio-PaTP, TMR-PaTP and Cy3-PaTP, were prepared as described previously.[2, 3]

DNA templates for T7 transcription

For the 37-mer RNA synthesis, partially double-stranded DNA fragments (10 μM each of a 52-mer non-template strand and a 58-mer template strand, Fig. S1A) were annealed in buffer containing 10 mM Tris–HCl (pH 7.6) and 10 mM NaCl, by heating at 95°C and slow cooling to 25°C. For the 106-mer RNA synthesis, partially double-stranded DNA templates were similarly prepared, using a 52-mer non-template strand and a 127-mer template strand (Fig. S3). Each 127-mer template strand was prepared by the ligation of a phosphorylated 58-mer DNA and a non-phosphorylated Temp3-69 (69-mer, 5'-GACGCGACCAGTTACGGAGCTCACACTCTACTCAACAGTGCCGAAGCACTGGACCCGTCCTTCACCATT) in the presence of SplintDNA-22 (22-mer, 5'-GACCGAAATGGTGAAGGACGGG), with T4 DNA ligase (New England Biolabs), at 14°C overnight, and then the ligated 127-mer single-stranded DNA was purified by denaturing gel electrophoresis (Fig. S3).

T7 transcription

Transcription for the 37-mer RNA was performed in reaction buffer (10 μl), containing 40 mM Tris–HCl (pH 8.0), 24 mM $MgCl_2$, 5 mM DTT, 2 mM spermidine, 0.01% Triton X-100, and 0.05 mg/ml T7RNAP variant, in the presence of 1 mM natural NTPs, 0 or 1 mM modified PaTP, 0.05 μCi/μl γ-^{32}P-GTP, and 1 μM DNA template. After an incubation at 37°C for 3 h, the reaction was quenched

by adding an equivalent volume of dye solution (95% formamide, 25 mM EDTA, with BPB). The reaction mixtures were heated at 75°C for 3 min, and the products were analyzed on a 20% denaturing polyacrylamide gel containing 7 M urea. The band patterns on the gel were detected by autoradiography. Transcription for the 106-mer RNA was performed in the presence of 1 or 2 mM natural NTPs (or 2'-fluoro-modified CTP and UTP, instead of the natural CTP and UTP), 0 or 1 mM modified PaTP, and 0.5 μM DNA template, and the products were analyzed on a 10% denaturing polyacrylamide gel containing 7 M urea. The band patterns on the gel were detected by UV shadowing on a silica gel TLC plate coated with the fluorescent indicator F254. The fluorescent images were analyzed with a STORM bio-imager or a FluoroChem Q bio-imager (Protein Simple). The full-length transcripts (106-mer) were also purified by denaturing PAGE, and the amounts of recovered RNA were measured with a NanoDrop spectrophotometer (ThermoFisher). The recovered RNA was subjected to further analysis (gel mobility shift assays for biotin-Pa containing transcripts and fluorescence detection for TMR-Pa and Cy3-Pa containing transcripts) to assess the incorporation efficiencies of the modified-Pa substrates. In fluorescence detection, 2 pmol of each purified RNA was subjected to denaturing PAGE, and the fluorescent images of the gel band patterns were detected with the fluorescent bio-imager. Gel band intensities were quantified with the ImageJ software.

Gel-Mobility Shift Assays

We detected the biotinylated RNA transcripts (106-mer, purified) by gel-mobility shift assays, using streptavidin. We incubated the mixture (6 μl) of 0.8 pmol transcripts and excess amounts of streptavidin (2 μg) for 1 h at room temperature, in buffer containing 10 mM Tris–HCl (pH 7.5) and 50 mM NaCl. After mixing with an equal volume of a denaturing reagent (10 M urea in 1× TBE), the biotinylated RNA-streptavidin complexes were separated from the free RNAs on a denaturing 10% polyacrylamide gel, and the RNAs on the gel were stained with SYBR Gold and detected with the FluoroChem Q bio-imager. Gel band intensities were quantified with the ImageJ software.

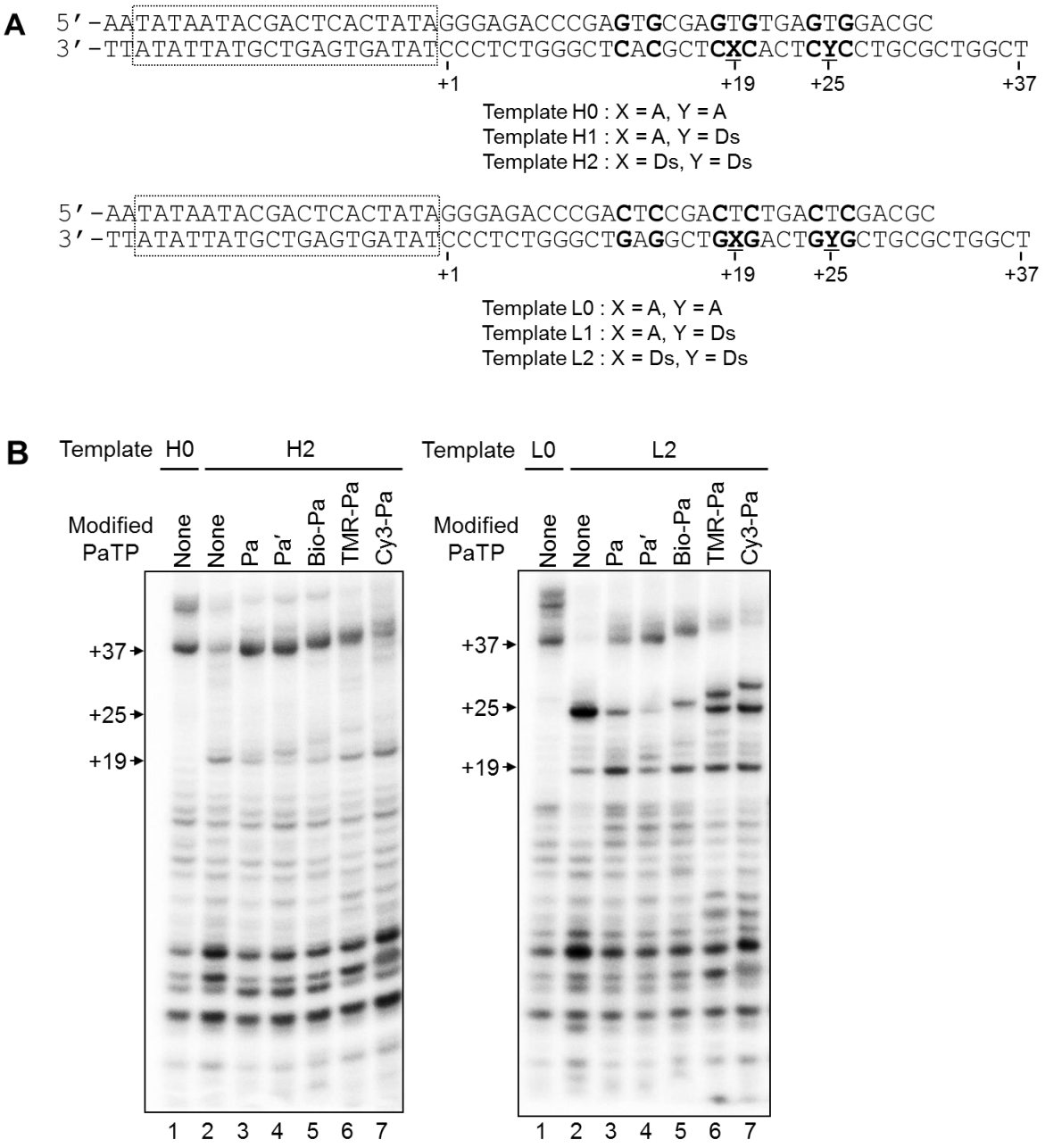
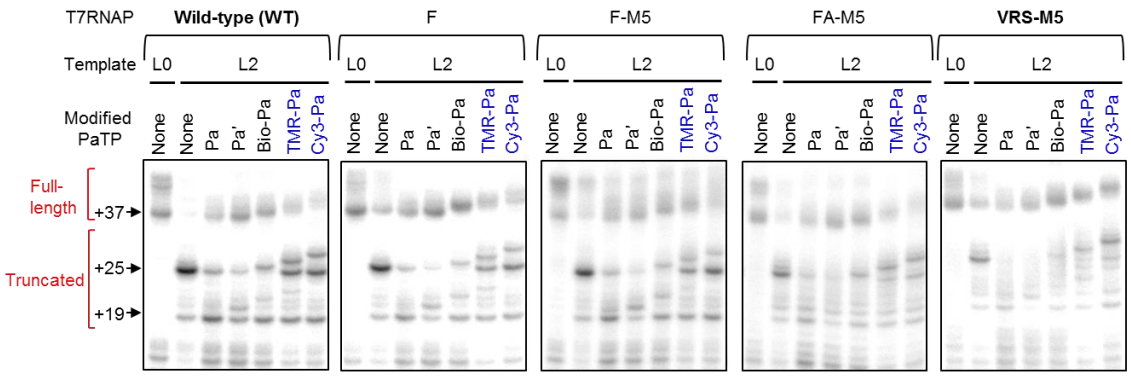


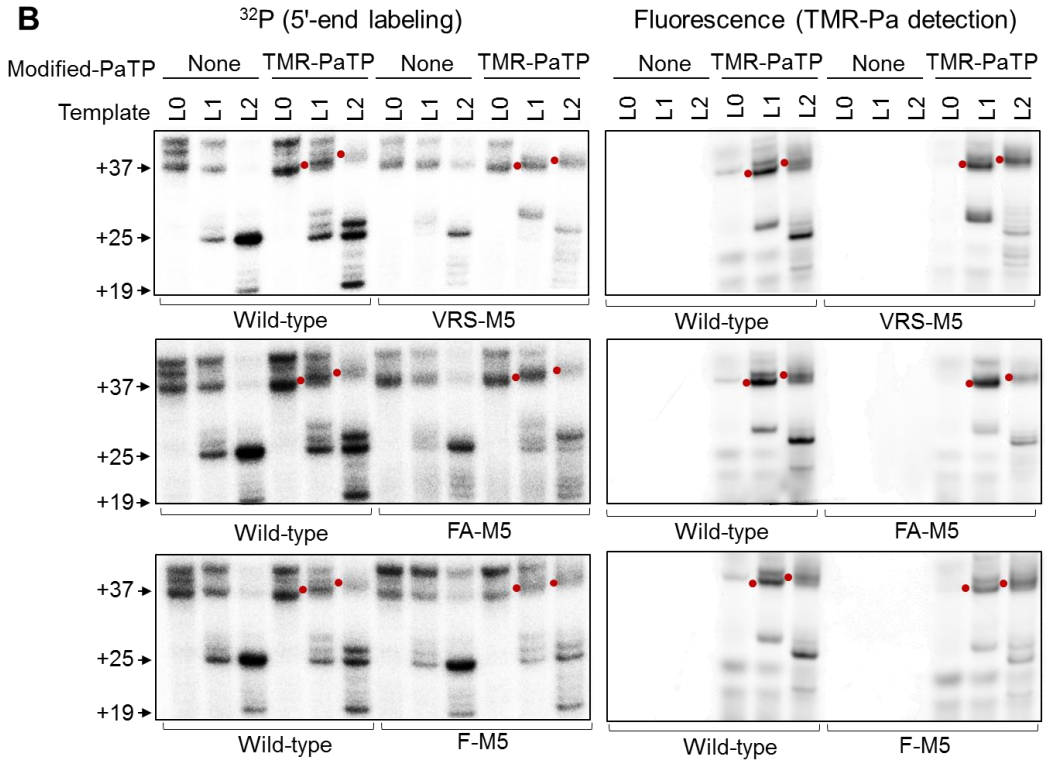
Fig. S1. (A) DNA template sequences (non-template 52-mer DNA and template 58-mer DNA) used in the T7 transcription for 37-mer RNA transcripts. (B) Gel electrophoresis of transcripts from 58-mer DNA templates containing two Ds bases (H2 and L2) or no Ds base (H0 or L0) with WT-T7RNAP, in the presence of natural NTPs (1 mM each) and each modified-PaTP, as indicated (1 mM). The transcripts were labeled at the 5'-end through reactions with γ-^{32}P-GTP**.**

A



Unnatural	T7RNAP	Full-length / Truncated (19-25 nts)	Relative yield (%)
TMR-PaTP	**WT**	**0.38 ± 0.06**	**48 ± 5**
	F	0.46	53
	F-M5	0.55	69
	FA-M5	0.34	50
	FA-M6	0.27	31
	VRS-M5	**0.60**	**63**
Cy3-PaTP	**WT**	**0.36 ± 0.06**	**46 ± 4**
	F	0.48	54
	F-M5	0.47	63
	FA-M5	0.33	47
	FA-M6	0.32	42
	VRS-M5	**0.53**	**67**

B



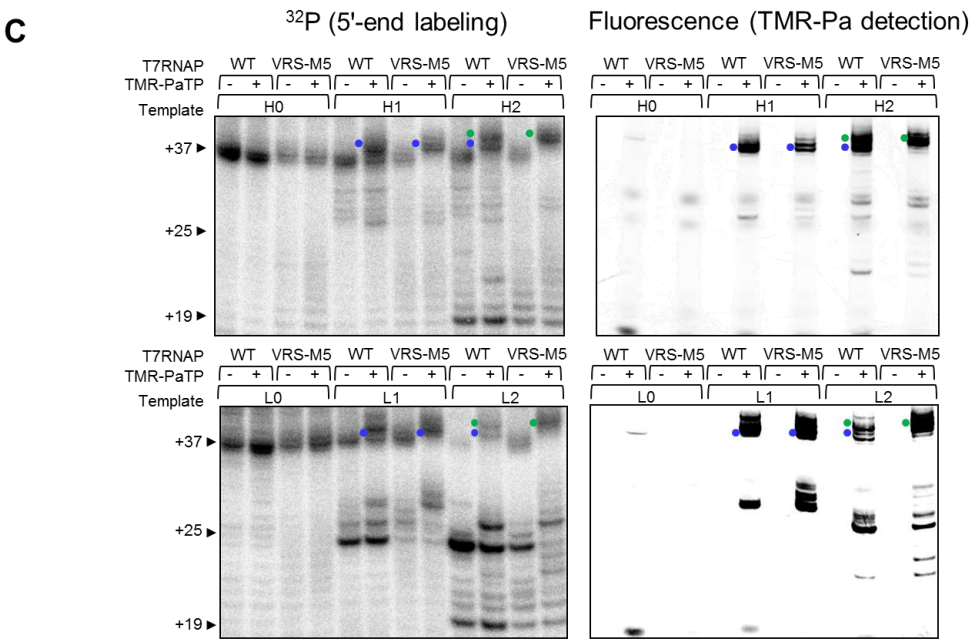


Fig. S2. Gel electrophoresis of transcripts from 58-mer DNA templates with WT-T7RNAP or each T7 RNAP variant, in the presence of natural NTPs (1 mM each) and each modified-PaTP, as indicated (1 mM). The transcripts were labeled at the 5'-end through reactions with γ-^{32}P-GTP. (A) Transcription with WT, F, F-M5, FA-M5, and VRS-M5, in the presence of each modified-PaTP. Yields of each full-length transcript, as well as each truncated product, were calculated and summarized. Relative yields (%) of each full-length transcript were calculated by comparing the yield of the full-length native transcripts using template L0, which consists of only the natural bases. (B) Transcription with WT, VRS-M5, FA-M5, and F-M5, in the presence of TMR-PaTP. (C) Transcription with WT and VRS-M5, in the presence of TMR-PaTP. Left panels of (B) and (C): detected by ^{32}P; Right panels of (B) and (C): detected by the TMR fluorescence. The product bands to compare at the same mobility are indicated by red circles (B), and blue and green circles (C). In (C), to more clearly examine the mobility differences depending on the presence of none, one or two TMR-Pa nucleotides, the gels were run for a longer time to allow higher resolutions, as compared to (B).

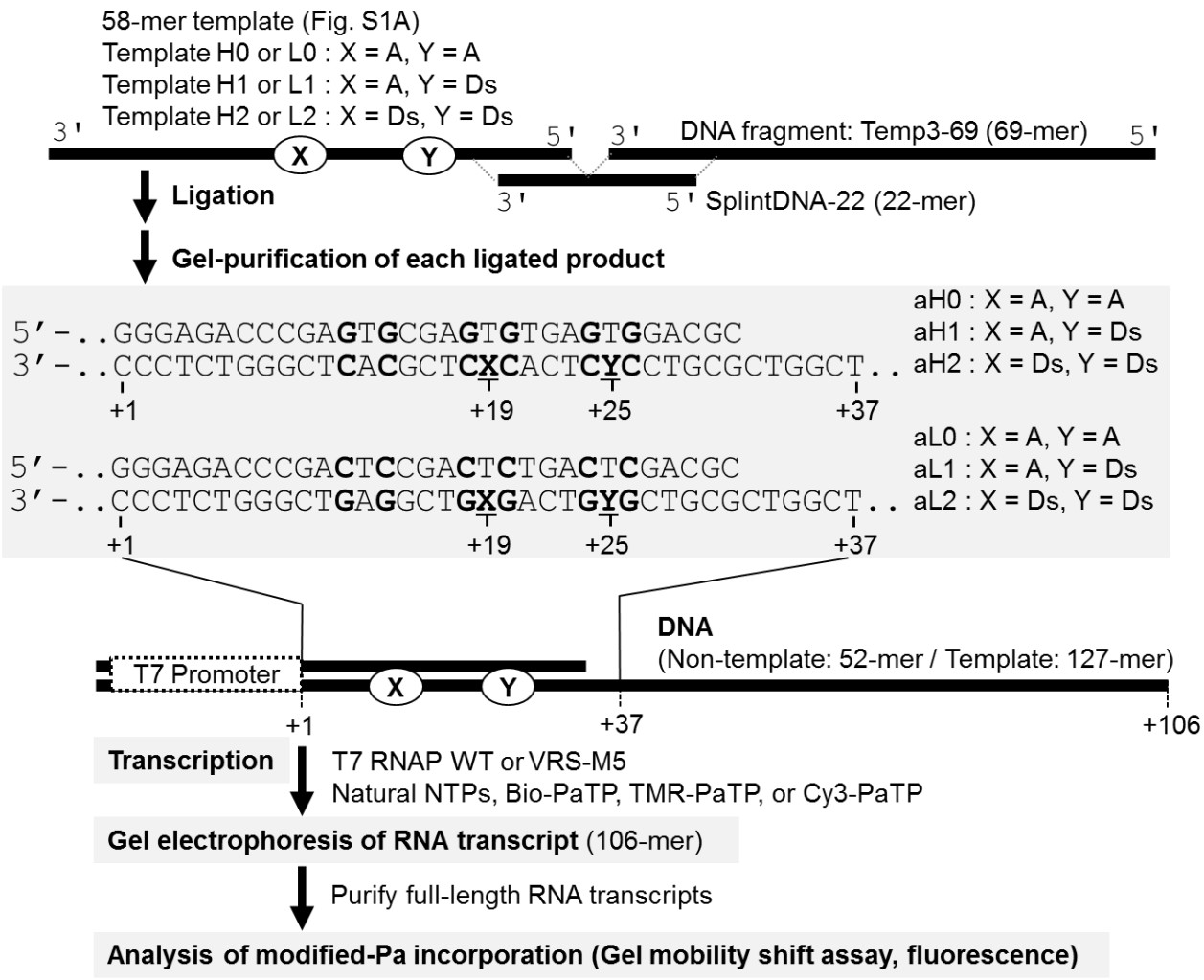
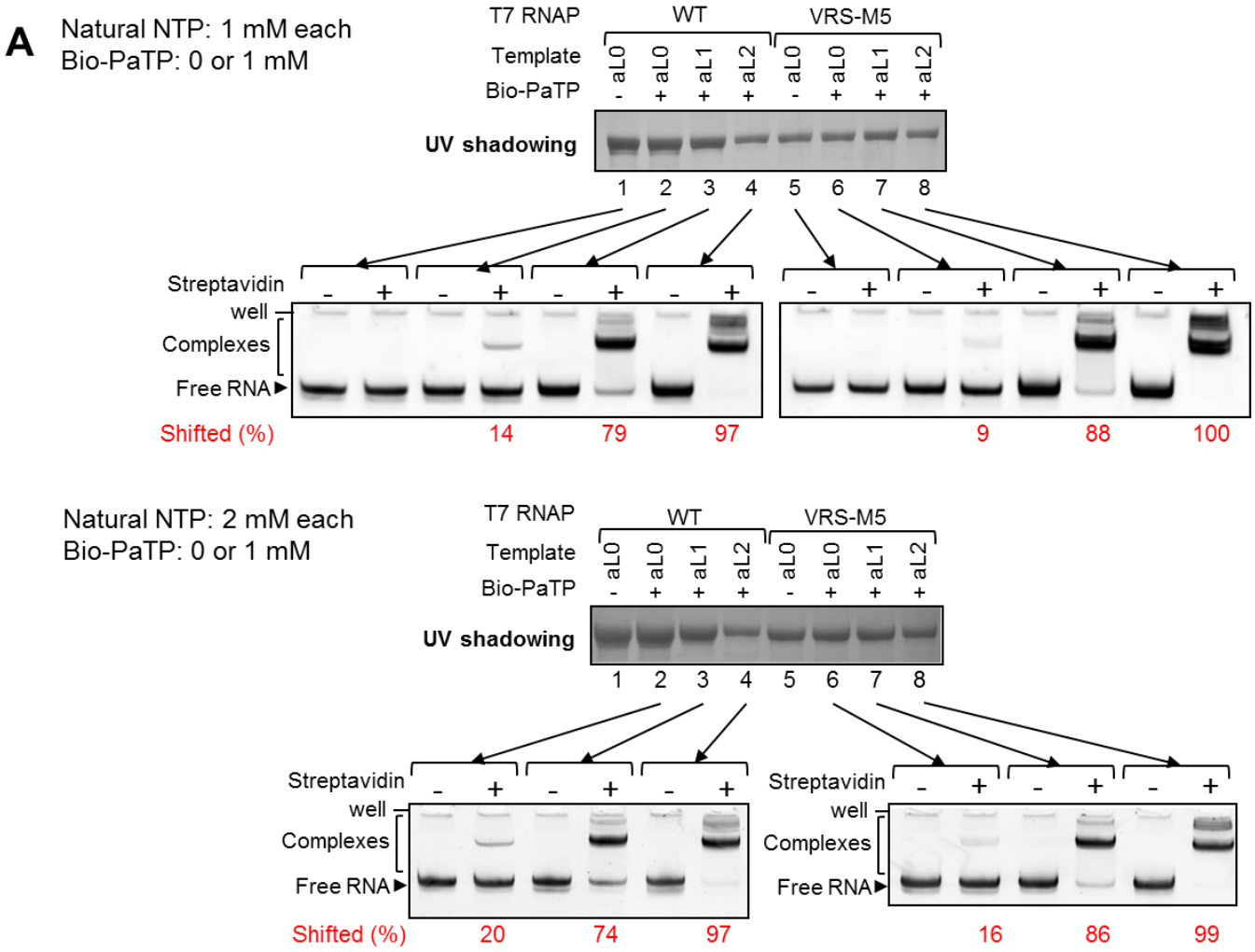
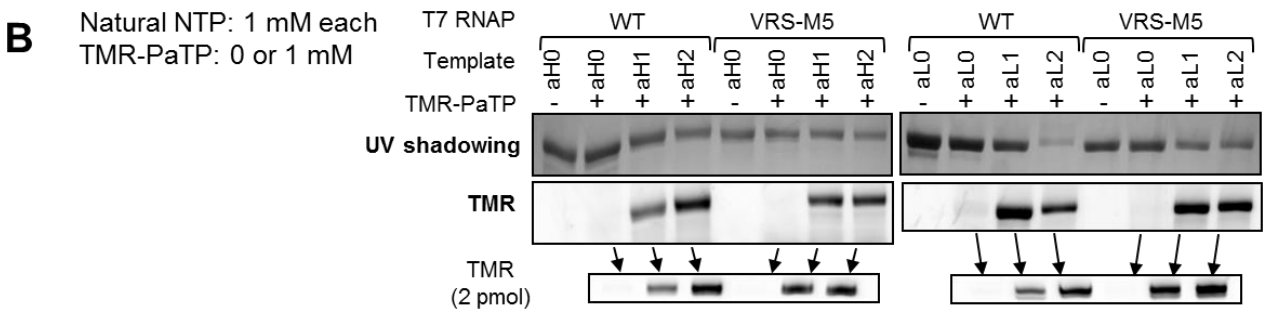
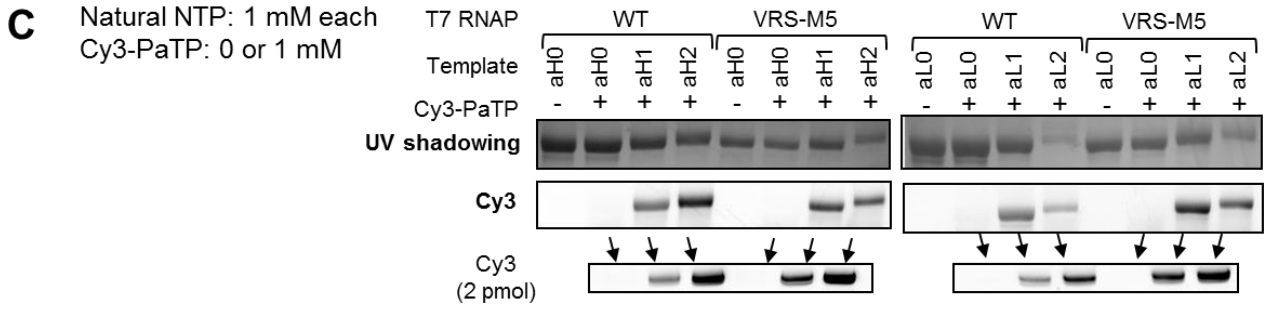


Fig. S3. Scheme of T7 transcription for 106-mer RNA and further characterization.


A
Natural NTP: 1 mM each
Bio-PaTP: 0 or 1 mM
T7 RNAP
WT
VRS-M5
Template
- aL0
+ aL0
+ aL1
+ aL2
- aL0
+ aL0
+ aL1
+ aL2
Bio-PaTP
UV shadowing
1 2 3 4 5 6 7 8
Streptavidin well
Complexes
Free RNA
Shifted (%)
14
79
97
9
88
100
Natural NTP: 2 mM each
Bio-PaTP: 0 or 1 mM
20
74
97
16
86
99


B
Natural NTP: 1 mM each
TMR-PaTP: 0 or 1 mM
T7 RNAP
WT
VRS-M5
Template
- aH0
+ aH0
+ aH1
+ aH2
- aH0
+ aH0
+ aH1
+ aH2
- aL0
+ aL0
+ aL1
+ aL2
- aL0
+ aL0
+ aL1
+ aL2
TMR-PaTP
UV shadowing
TMR
TMR (2 pmol)


C
Natural NTP: 1 mM each
Cy3-PaTP: 0 or 1 mM
T7 RNAP
WT
VRS-M5
Template
- aH0
+ aH0
+ aH1
+ aH2
- aH0
+ aH0
+ aH1
+ aH2
- aL0
+ aL0
+ aL1
+ aL2
- aL0
+ aL0
+ aL1
+ aL2
Cy3-PaTP
UV shadowing
Cy3
Cy3 (2 pmol)

D

Template DNA	Natural		Unnatural (1 mM)	T7RNAP	Relative fluorescence intensity
aL0	1 mM	OH	TMR-PaTP	WT	0.05 ± 0.04
aL1					**1.00**
aL2					1.92 ± 0.29
aL0				VRS-M5	0.08 ± 0.10
aL1					**2.13 ± 0.28**
aL2					3.04 ± 0.80
aL0	1 mM	OH	Cy3-PaTP	WT	0.05
aL1					**1.00**
aL2					1.98
aL0				VRS-M5	0.01
aL1					**2.26**
aL2					3.20
aH0	1 mM	OH	TMR-PaTP	WT	0.03
aH1					**1.00**
aH2					2.63
aH1	2 mM				0.70
aH0	1 mM	OH		VRS-M5	0.06
aH1					**1.69 ± 0.13**
aH2					3.55
aH1	2 mM				1.15
aH0	2 mM	F		VRS-M5	0.13
aH1					1.81

Fig. S4. Gel analysis of transcripts generated with WT-T7RNAP or VRS-M5 in the presence of the natural NTPs (1 or 2 mM) and modified-PaTP (1 mM) as substrates, and further characterization of the purified transcripts, to assess the efficiency and selectivity of the modified-Pa incorporation opposite Ds in DNA templates. (A) Analysis of Bio-Pa incorporation into transcripts and gel-mobility shift assays to detect biotinylated transcripts through complex formation with streptavidin. The upper panels (UV shadowing) show the band patterns of the transcribed RNAs, detected by UV shadowing. The lower panels show the results of the gel-mobility shift assays. Gel-shifted band patterns of the full-length RNAs purified by gel electrophoresis, in the presence and absence of streptavidin, are presented. Biotinylated RNA-streptavidin complexes were separated from free RNAs on a 10% polyacrylamide–7 M urea gel. The gels were stained with SYBR Gold, and the images were obtained with a FluoroChem Q bio-imager. The amounts (Shifted, %) of the shifted bands corresponding to the complex between biotinylated transcripts and streptavidin were calculated from the band intensities. The data confirmed the highly-specific incorporation of Bio-PaTP opposite Ds. (B) Analysis of transcripts with TMR-PaTP. The transcribed RNAs were detected by UV shadowing or with a FluoroChem Q bio-imager (TMR). The lower panels show the fluorescence detection of each purified product (2 pmol per well) analyzed on a gel. (C) Analysis of transcripts with Cy3-PaTP. The transcribed RNAs were detected by UV shadowing (RNA) or with a FluoroChem Q bio-imager (Cy3). The lower panels show the fluorescence detection of each purified products (2 pmol per well) analyzed on the gel. (D) Comparison of relative fluorescent band intensities of each purified product on the gel images (Fig. S4B and S4C, and Fig. 4C).

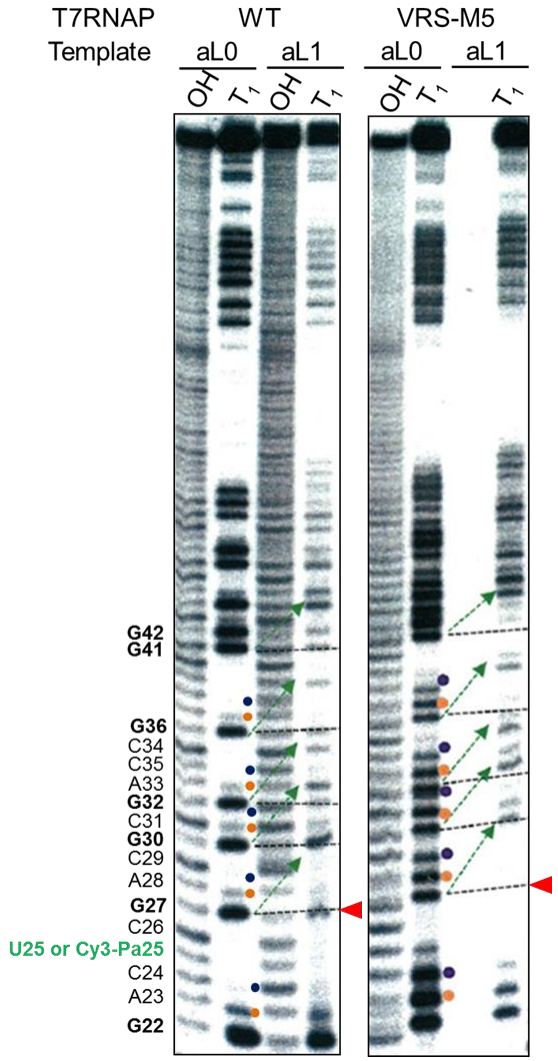


Fig. S5. Sequence analysis of the 106-mer transcripts in the presence of 1 mM each Cy3-PaTP and natural NTPs, using templates aL0 and aL1 with WT-T7RNAP or VRS-M5. The Cy3-Pa incorporation using template aL1 was located at position 25. The full-length transcripts were labeled with γ-^{32}P-ATP at the 5′-end with T4 polynucleotide kinase. The labeled transcripts were purified on a gel and partially digested with alkali (OH) or G-specific RNase T_1 (T_1). The digestion patterns of each transcript were analyzed on a 10% denaturing polyacrylamide gel. The band positions of the T_1-digested products containing Cy3-Pa were largely shifted, as compared to those of the transcripts from template aL0 (indicated by green dotted arrows for shifted bands/with Cy3-Pa and black dotted lines for original bands/without Cy3-Pa). Note that a band corresponding to the 27-mer fragment without Cy3-Pa (red triangle of WT) was generated by WT-T7RNAP, indicating that the transcripts were a mixture with (shifted) and without (unshifted) Cy3-Pa. In contrast, the corresponding band (red triangle) was not detected when VRS-M5 was used, indicating the high efficiency and specificity of Cy3-Pa incorporation opposite Ds when using VRS-M5.

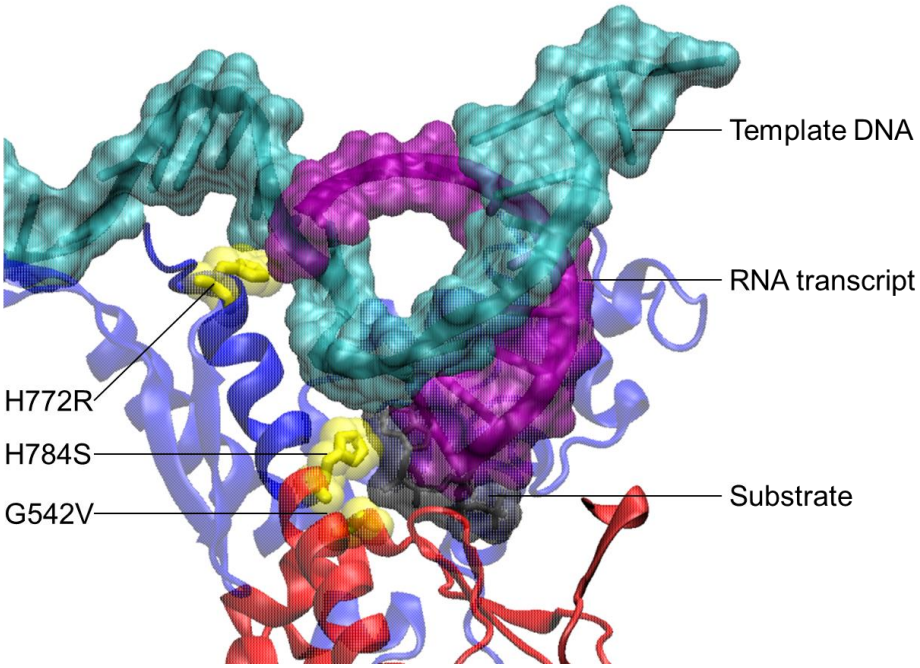


Fig. S6 The three mutations in the VRS T7RNAP variant, highlighted on the tertiary structure of T7RNAP (PDB ID 1S76).

REFERENCES

1. A. J. Meyer, D. J. Garry, B. Hall, M. M. Byrom, H. G. McDonald, X. Yang, Y. W. Yin and A. D. Ellington, *Nucleic Acids Res*, 2015, **43**, 7480-7488.
2. I. Hirao, M. Kimoto, T. Mitsui, T. Fujiwara, R. Kawai, A. Sato, Y. Harada and S. Yokoyama, *Nat Methods*, 2006, **3**, 729-735.
3. N. Morohashi, M. Kimoto, A. Sato, R. Kawai and I. Hirao, *Molecules*, 2012, **17**, 2855-2876.