Electronic Supplementary Material (ESI) for Chemical Science.

Advancing the Enzymatic Toolkit for 2'-Fluoro Arabino Nucleic Acid (FANA) manipulation: Phosphorylation, Ligation, Replication, and Templating RNA Transcription

Yingyu Liu[1, 2, #], Jun Wang[1, 2, #], Yashu Wu[2, 3], Yajun Wang[*2, 3]

[1]College of Pharmaceutical Sciences, Soochow University, Suzhou 215123, China

[2]Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences; Hangzhou, Zhejiang 310000, China

[3]The Cancer Hospital of the University of Chinese Academy of Science (Zhejiang Cancer Hospital), Hangzhou, Zhejiang 310022, China

[*]Corresponding email: wangyj@ibmc.ac.cn

[#] These authors contributed equally

1. Materials

FANA NTPs (fATP, fCTP, fGTP, fUTP) were obtained from Metkinen Chemistry (Kuusisto, Finland). FANA phosphoramidites were purchased from Glen Research (Sterling, Virginia). FANA oligonucleotides were synthesized on an ABI3400 DNA synthesizer using chemical synthesis reagents purchased from Glen Research (Sterling, Virginia). dNTPs, NTPs, DNA and RNA oligonucleotides were purchased from GenScript, purified by denaturing polyacrylamide gel electrophoresis, and quantified by UV absorbance. DH5α competent cells were obtained from Sangon. YM-3 and YM-10 microcentrifugal concentrators were purchased from EMD Millipore. Other molecular biology reagents included T4 PNK (NEB), T4 DNA Ligase (NEB), Gibson assembly kit (NEB), T7 RNA polymerase (NEB), Pyrophosphatase, inorganic (yeast) (NEB), DNase I RNase free (BBI) and RNase Inhibitor (BBI).

2. Biochemical reactions and data analysis

Tgo (exo-) polymerase purification

Polymerases were expressed and purified as reported previously[1]. Briefly, DH5α *E. coli* cells carrying Tgo (exo-) expression plasmid were inoculated in 1 L of LB-ampicillin (100 μg/mL) liquid medium and grown at 37°C with shaking at 225 rpm. At OD600 = 0.6, the expression culture was cooled to 15°C and induced with IPTG at a final concentration of 0.5 mM and incubated overnight at 15°C with shaking at 225 rpm. Cells were harvested by centrifugation for 20 min at 3,315 x g at 4°C and lysed in 40 mL buffer (10 mM Tris pH 8.0, 500 mM KCl, 10% glycerol) by sonication on ice. The cell lysate was centrifuged for 30 min at 23,708 x g at 4°C, and the clarified supernatant was heat for 30 min at 80°C then immediately cooling for 30 min on ice. The lysate was clarified again by centrifugation for 20 min at 23,708 x g at 4°C. Nucleic acids were precipitated by adding 10% (v/v) polyethyleneimine (PEI) to a final concentration of 0.5% and incubating for 30 min on ice, then centrifuging for 30 min at 23,708 x g at 4°C. The supernatant was recovered, and the polymerase was precipitated by adding 60% (w/v) ammonium sulfate, incubating for 30 min on ice, and then centrifuging for 30 min at 23,708 x g at 4°C. Protein pellets were suspended in 4°C buffer (10 mM Tris pH 8.0, 50 mM KCl, 10% glycerol). Particulates were removed by centrifuging for 10 min at 23,708 x g at 4°C. Tgo (exo-) was then purified by 5 mL heparin HP affinity chromatography with step elutions of 100, 250, 500, and 1000 mM KCl. Fractions corresponding to protein of the correct size were verified by SDS PAGE, combined, quantified by UV absorbance at 280 nm, and stored at 4°C.

T4 PNK catalyzed 5’ phosphorylation reaction of FANA

FANA **ON 1** was treated in 1x T4 PNK reaction buffer containing 0.4 U/μL of T4 PNK supplemented with 1 mM ATP for 18 h or 1 h at 37°C. The outputs were purified through phenol-chloroform extraction and ethanol precipitation with sodium acetate (0.3 M, pH 5.2). The pellet was suspended in DEPC-treated water and ultraviolet quantified by NanoDrop. Samples were subjected to matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry for identity confirmation.

T4 DNA ligase catalyzed FANA ligation

The ligation reaction was performed in 1x T4 DNA ligase reaction buffer containing 5 μM of Donor, 10 μM of Acceptor, 20 μM of splint, and 20 U/μL of T4 DNA ligase for 18 h at 24°C. Multiple time points were collected by quenching 1 μL of the reaction using 15 μL (15 equiv., v/v) of formamide stop buffer (99% deionized formamide, 25 mM EDTA) and cooling on ice. Samples were denatured for 10 min at 95°C and analyzed by 15% denaturing PAGE. Gels were visualized using an Amersham Typhoon laser-scanner platform.

Kinetic cleavage reactions for FANAzyme 12-7 obtained by ligation

Single-turnover kinetic cleavage reactions were conducted under single-turnover conditions in 50 mM Tris-HCl buffer (pH 8.5) containing 25 mM $MgCl_2$ and 200 mM NaCl, 0.5 μM substrate, and 2.5 μM FANAzyme 12-7 at 24°C. Purified enzymes and substrates were annealed in 50 mM Tris-HCl buffer (pH 8.5) by heating for 5 min at 95°C and cooling for 5 min on ice. Reactions were initiated by the addition of NaCl and $MgCl_2$ to the reaction. Multiple time points were collected by quenching 1 μL of the reaction using 15 μL (15 equiv., v/v) of formamide stop buffer (99% deionized formamide, 25 mM EDTA) and cooling on ice. Samples were denatured for 10 min at 95°C and analyzed by 15% denaturing PAGE. Gels were visualized using an Amersham Typhoon laser-scanner platform.

Single-nucleotide incorporation kinetics catalyzed by Tgo (exo-)

Single-nucleotide incorporation assays were performed using a mini-scaffold for measuring single-nucleotide insertion kinetics. The mini-scaffold was prepared by annealing the template-strand DNA, non-template-strand DNA, and a Cy5-labeled primer at 2:3:1 molar ratio in ThermoPol buffer. Annealing was performed by heating the scaffold mixture at 80 °C for 5 min and gradually cooling down to room temperature. The final reaction mixture contained 50 nM mini-scaffold, 60 nM Tgo (exo-), and varying concentration of dNTP or fNTP tested. At each time point, 3 μl of the reaction mixture was removed from the reaction into 17 μl of stop buffer (95% formamide, 5% EDTA, pH 8.0). After the reaction, all samples were denatured at 95 ℃ for 15 min and resolved by urea-denaturing PAGE. Values of k_{obs} were calculated by fitting the percentage of incorporation and reaction time (min) to the first-order decay equation using Prism 8, GraphPad. All images were quantitated by using ImageQuantTL software.

DNA-template directed primer extension

DNA template-directed primer extension reactions were performed in 1x ThermoPol buffer supplemented with 100 nM of each DNA template-primer duplex, 0.25 mM of each dNTP or fNTP, 1 mM $MgSO_4$, and 40 nM of Tgo-WT using triphosphate mixtures containing single fNTP substitution. The primer-template complex was annealed in 1× ThermoPol buffer by heating for 5 min at 90 °C and cooling for 10 min at 4 °C. Primer-extension reactions were performed for 1 min at 55 °C before quenching the reactions with 10 equivalents (v/v) of formamide stop buffer (99% deionized formamide, 25 mM EDTA). All quenched

samples were denatured for 15 min at 95 °C before analyzing by denaturing PAGE. Gels were visualized using an Amersham Typhoon laser-scanner platform.

FANA substitutional polymerase chain reaction (PCR) using Tgo (exo-)

All reactions were performed in 1x ThermoPol buffer supplemented with 1 μM of each primer (forward and reverse), 0.25 mM of each dNTP or fNTP, 1 mM $MgSO_4$, and 40 nM of Tgo (exo-). PCR reactions following protocol: step 1: 95°C for 5 min; step 2: 25 cycles of (98°C for 30 s; 55°C for 30 s; 72°C for 60 s); step 3: 72°C for 5 min. Samples were analyzed by 2% agarose gels.

Measurement of FANA replication fidelity

FANA replication fidelity was measured by sequencing the recombined UV-excitable green fluorescent protein expressing plasmid and carrying the inserted DNA-FANA chimeric amplicon. The plasmids were assembled following the Gibson assembly protocol provided by the manufacture and transformed into *E. coli* following standard protocol, and recombinants containing the inserted amplicon were screened on LB-agar plates supplemented with antibiotics. Green colonies were selected under UV transillumination for Sanger sequencing via Genewiz Inc. sequencing services. DNA sequences were aligned with the PCR template and analyzed for point mutations using SnapGene.

T7 RNA polymerase catalyzed in vitro RNA transcription (IVT)

A synthetic DNA template coding the fluorescent spinach RNA aptamer with one of the primers containing the T7 promoter sequence was amplified by single fNTP substitutional PCR for subsequent IVT. RNA was transcribed in a 1x T7 RNAP reaction buffer, supplemented with 1 mM each ATP, UTP, GTP, and CTP, 5 mM dithiothreitol, 2.5 U/mL of Pyrophosphatase, inorganic (yeast), 1 U/μL of RNase inhibitor using 20 ng/μL of the purified DNA amplicon and 5 U/μL of T7 RNA polymerase at 37°C for 16 h. The IVT reaction was terminated by adding 10 U/mL of RNase-free DNase I and incubating at 37°C for 15 min followed by heat inactivation of the DNase I at 75°C for 20 min.

Fluorescent image of DFHBI-activated IVT reactions

The DNase I treated IVT reaction containing the RNA transcript was co-incubated with DFHBI dissolved in DMSO in 40 mM HEPES (pH 7.4) containing a final concentration of 125 mM KCl, 5 mM $MgCl_2$, and 5 μM of DFHBI at 24°C for 30 min[2]. The reactions were then imaged using an Amersham Typhoon laser-scanner platform or quantified for arbitrary units of fluorescence on a TECAN Spark multimode microplate reader with an excitation wavelength of 460 nm and an emission wavelength of 505 nm.

3. Supplementary figures

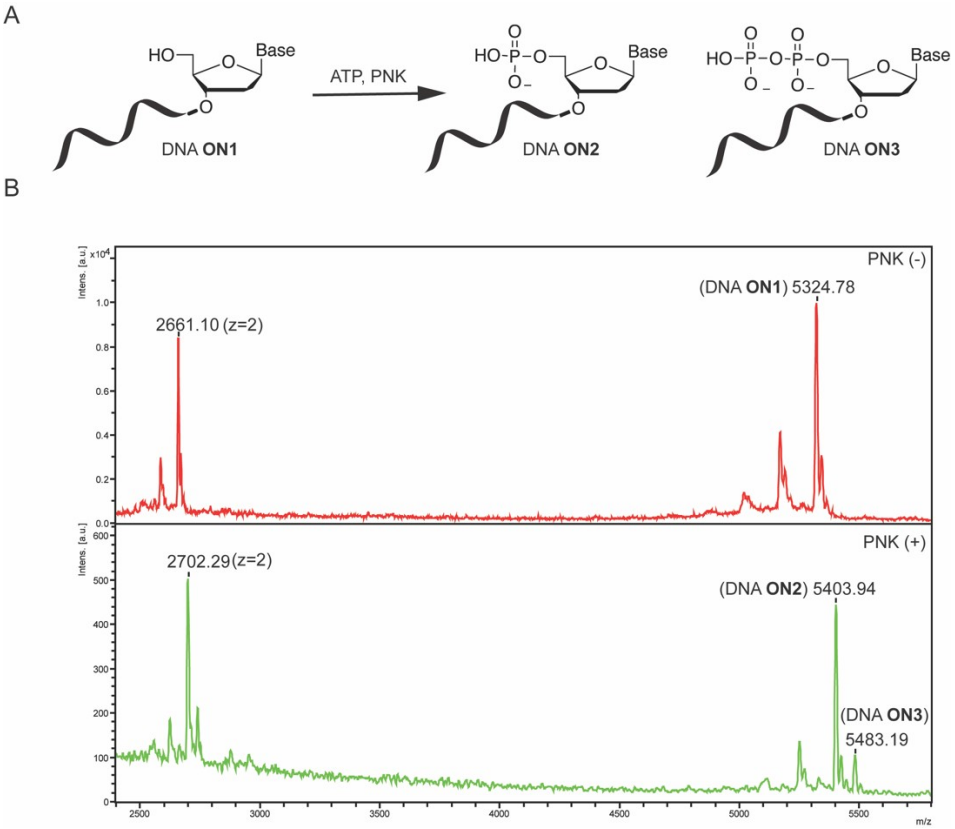

Figure S1. T4 PNK catalyzed 5' phosphorylation of DNA ON**.** (A) Schematic representation of T4 PNK catalyzed 5' phosphorylation reaction of DNA. (B) MALDI-TOF spectra of the DNA ON before and after PNK phosphorylation reaction, which suggested that both 5' mono- (DNA **ON2**) and diphosphorylation (DNA **ON3**) products were generated. The reaction was performed in 1X T4 PNK reaction buffer containing 0.4 U/uL of T4 PNK supplemented with 1 mM ATP for 18 h at 37 °C.

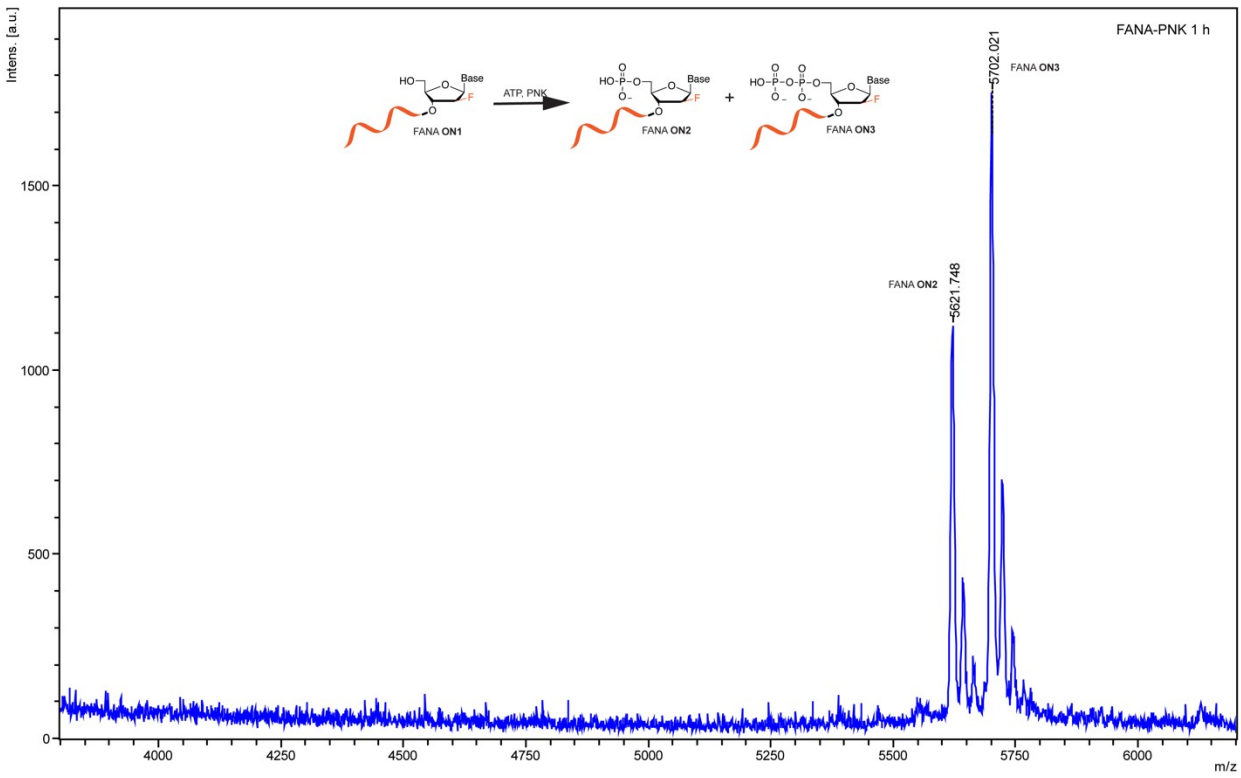

Figure S2. T4 PNK catalyzed 5’ phosphorylation of FANA ON**.** The reaction was performed in 1X T4 PNK reaction buffer containing 0.4 U/uL of T4 PNK supplemented with 1 mM ATP for 1 h at 37 °C. MALDI-TOF spectrum of the FANA ON after PNK phosphorylation reaction revealed the formation of both 5’ mono- (FANA **ON2**) and diphosphorylation (FANA **ON3**) products.

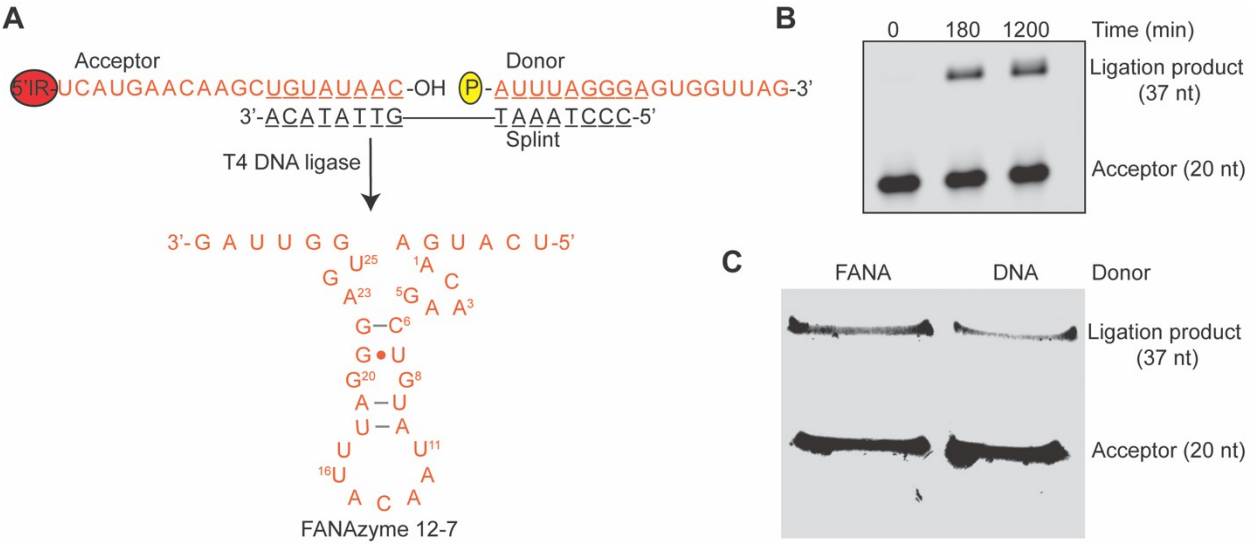

Figure S3. T4 DNA ligase catalyzed FANA ligation**.** (A) The sequence scheme of ONs for T4 DNA ligase catalyzed ligation to generate the full-length FANAzyme 12-7. (B) Representative PAGE showing the ligation reaction at 180 and 1200 min. The reaction was performed in 1X T4 DNA ligase reaction buffer containing 5 μM of Donor, 10 μM of Acceptor, 20 μM of splint, and 20 U/uL of T4 DNA ligase at 24 °C. (C) The purification PAGE for isolating the ligation

product FANAzyme 12-7 used for RNA catalytic cleavage reaction. All ligation reactions were performed in 1X T4 DNA ligase reaction buffer containing 5 μM of Donor, 10 μM of Acceptor, 20 μM of splint, and 20 U/uL of T4 DNA ligase at 24 °C.

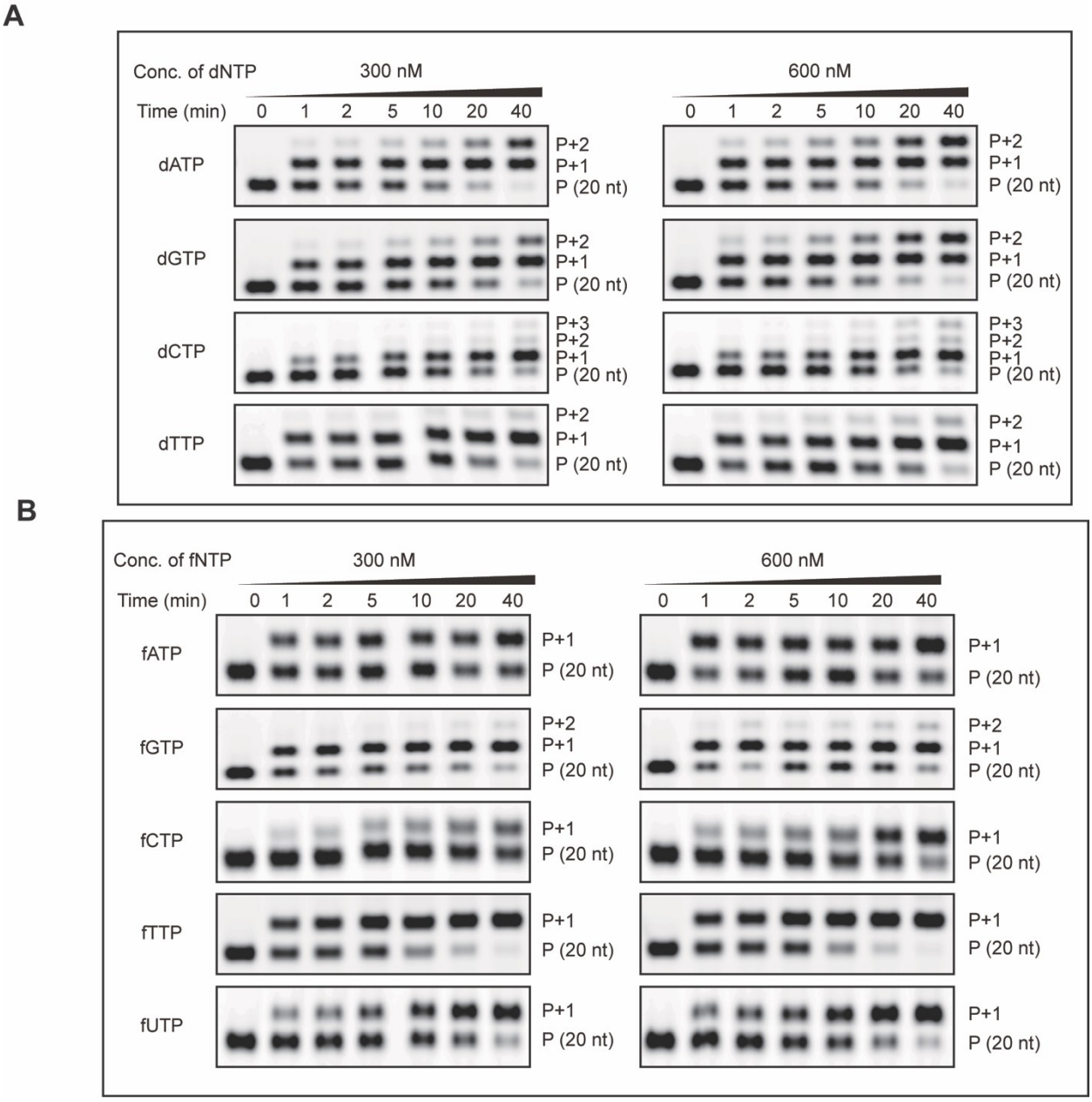

Figure S4. Representative gels showing the single-nucleotide insertion kinetic of dNTPs (**A**) and fNTPs (**B**). All reactions were performed in 1X ThermoPol buffer containing 50 nM of template-primer mini-scaffold, 300 or 600 nM tested dNTP or fNTP and 60 nM of Tgo (exo-) at 55 °C. Values of k_{obs} were calculated by fitting the percentage of incorporation and reaction time (min) to the first-order decay equation using Prism 8, GraphPad.

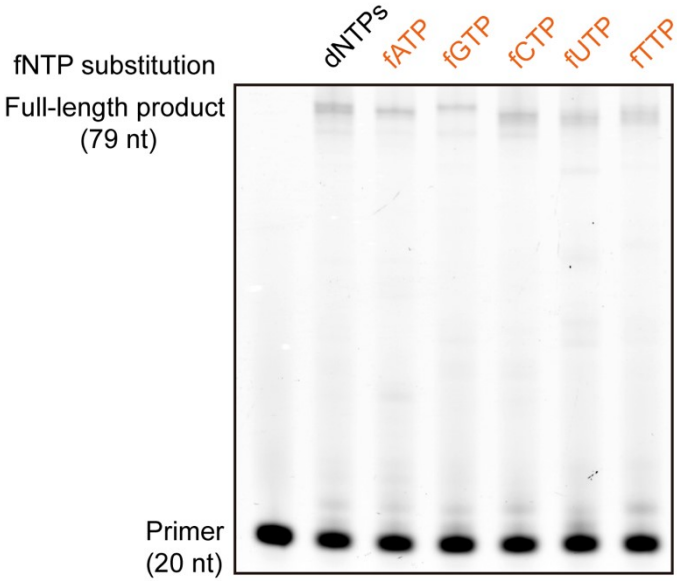

Figure S5. DNA template-directed primer extension reaction using triphosphate mixtures containing single fNTP substitution. All reactions were performed in 1x ThermoPol buffer supplemented with 100 nM of each DNA template-primer duplex, 0.25 mM of each dNTP or fNTP, 1 mM $MgSO_4$, and 40 nM of Tgo-WT.

Figure S6. FANA substitutional PCR of the spinach aptamer template using Tgo (exo-). Representative agarose gel electrophoresis showing the PCR amplicons (denoted by the asterisk) of different dNTP and fNTP combinations. All

reactions were performed in 1x ThermoPol buffer supplemented with 1 μM of each primers (forward and reverse), 0.25 mM of each dNTP or fNTP, 1 mM $MgSO_4$, and 40 nM of Tgo (exo-). PCR protocol: step 1: 95 °C for 5 min; step 2: 25 cycles of (98 °C for 30 s; 55 °C for 30 s; 72 °C for 60 s); step 3: 72 °C for 5 min.

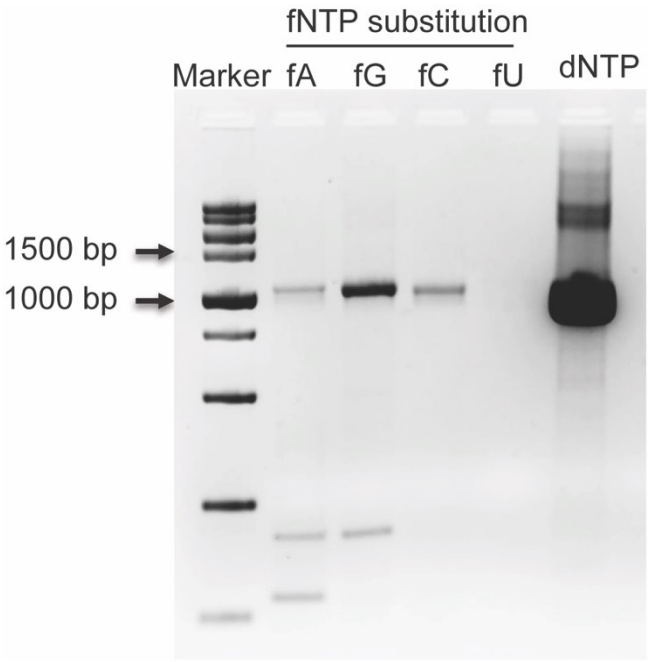

Figure S7. FANA substitutional PCR bp using Tgo (exo-) to amplify amplicons of around 1000. All reactions were performed in 1x ThermoPol buffer supplemented with 1 μM of each primers (forward and reverse), 0.25 mM of each dNTP or fNTP, 1 mM $MgSO_4$, and 40 nM of Tgo-WT. PCR protocol: step 1: 95 °C for 5 min; step 2: 25 cycles of (98 °C for 30 s; 55 °C for 30 s; 72 °C for 60 s); step 3: 72 °C for 5 min.

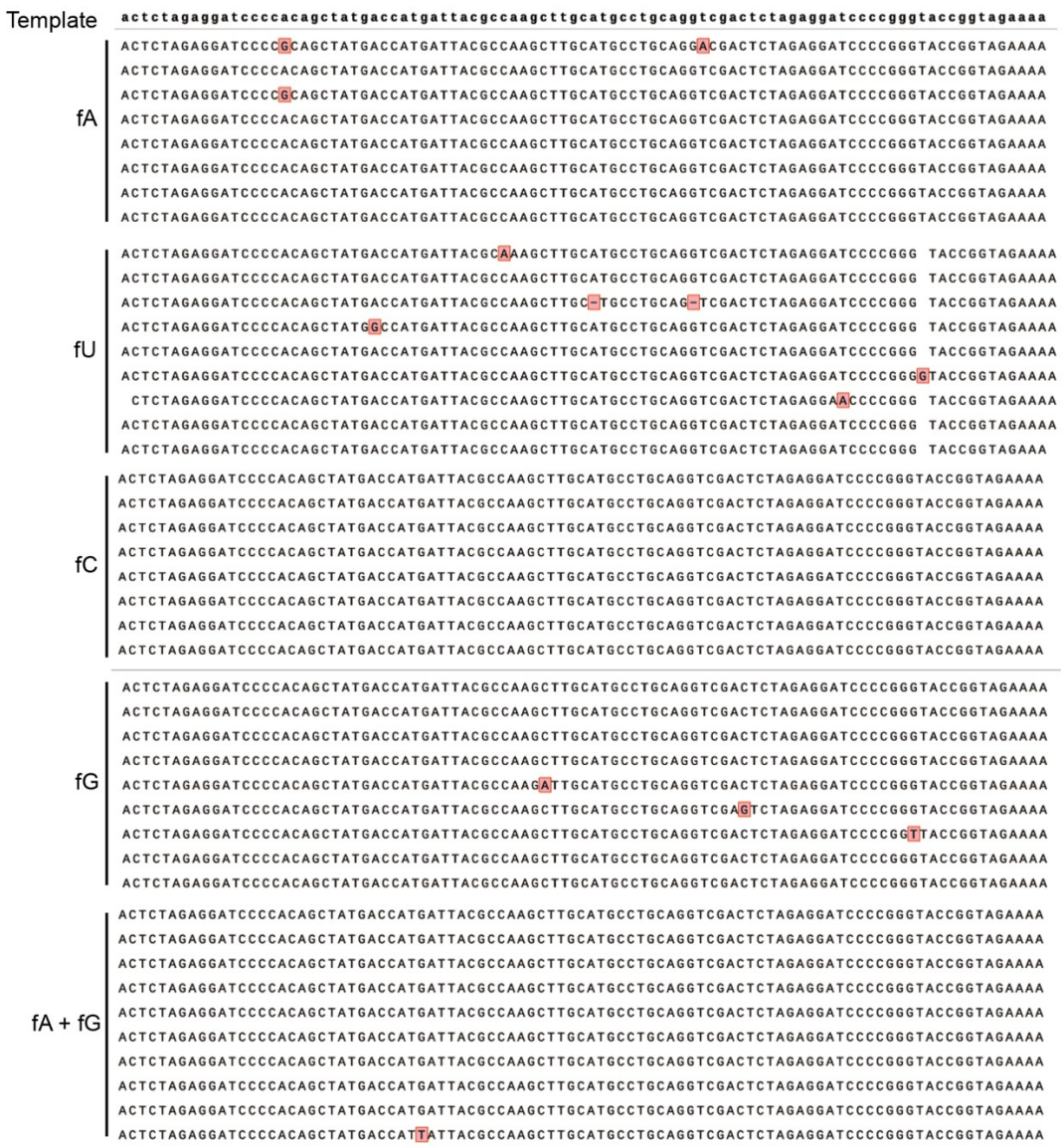

Figure S8. The alignment of sequences obtained by Sanger sequencing of FANA substitutional PCR amplicons.

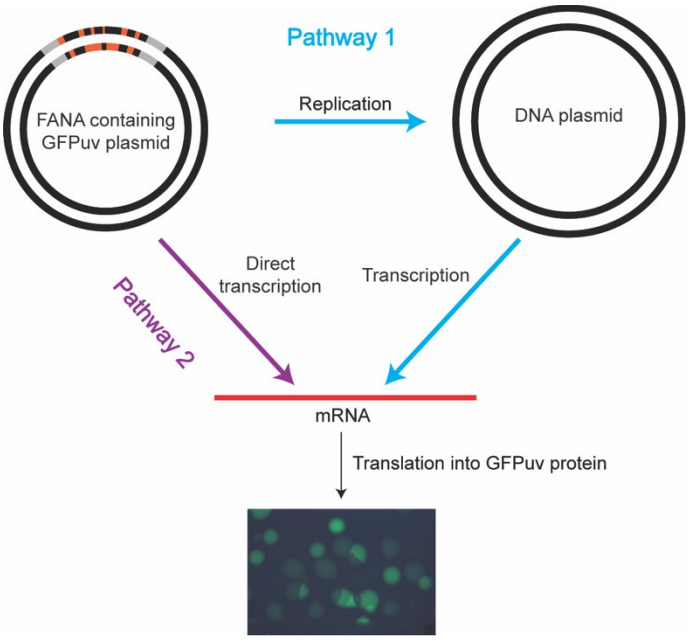

Figure S9. The potential replicational pathways of the FANA residue containing "alien" plasmid in *E. Coli*.

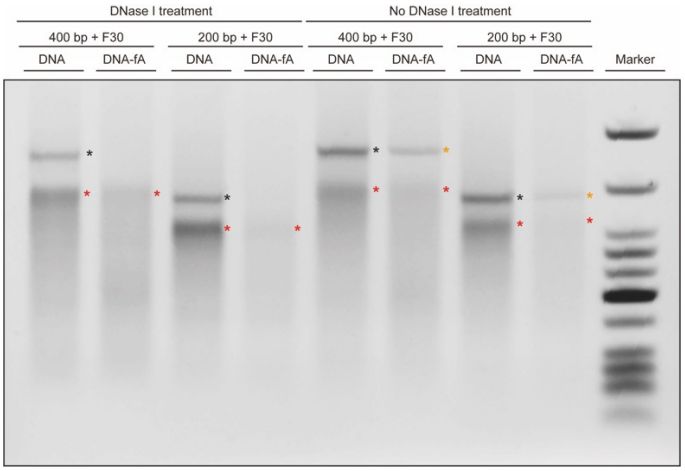

Figure S10. Representative agarose resolving IVT reactions using different DNA-fA chimeric amplicons measuring different length as templates. Asterisks in black marked the DNA template bands, asterisks in orange marked the DNA-fA chimeric template bands, and asterisks in red marked the RNA transcripts. The band intensities of DNA-fA were significantly lower than that of DNA, which might have contributed to the lower yield of RNA transcripts measured by both fluorescence reading and agarose electrophoresis.

4. Table of oligonucleotides

Oligo Name	Sequence (5'-3')
FANA ON1 (Donor)	auuuagggagugguuag
DNA ON1	ATTTAGGGAGTGGTTAG
Acceptor	/5IRD800/ ucaugaacaagcuguauaac
Splint	CCCTAAATGTTATACA
All RNA substrate	/5Cy5/rCrUrArArCrCrGrUrCrArUrGrA
FANAzyme 12-7 ligation product	IR680-ucaugaacaagcuguauaac-auuuagggagugguuag
DNA template for primer extension	GGATCGTCAGTGCATTGAGACCTAACCACTCCCTAAATGTTATACAGCTTGT TCATGACTGTTGGCGTAGGCCGACGC
DNA primer for primer extension (PBS11)	GCGTCGGCCTACGCCAACAG
Template for dATP/fATP incorporation	GGGAATCGATCTGTTGGCGTAGG
Template for dGTP/fGTP incorporation	GGGAATCGACCTGTTGGCGTAGG
Template for dCTP/fCTP incorporation	GGGAATCGAGCTGTTGGCGTAGG
Template for dTTP/fTTP/fUTP incorporation	GGGAATCGTACTGTTGGCGTAGG
Primer for single nucleotide incorporation	IR680-GCGTCGGCCTACGCCAACAG
Non-template strand	ACGATTCCCGA
Synthetic DNA template	TATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATC CCCGGGTACCGGTAGAAAA
Forward primer	ACTCTAGAGGATCCCCACAGCTATGACCATGATTACGCC
Reverse primer	TTTTCTACCGGTACCCGG

T7 promoter forward primer	TAATACGACTCACTATAGG GACGCAACTGAATGAAAT
Spinach reverse primer	GACGCGACTAGTTACGGAGCTCACACTCTACTCAACAAGC
Spinach DNA template	GGGACGCAACTGAATGAAATGGTGAAGGACGGGTCCAGGTGTGGCTGCTTCG GCAGTGCAGCTTGTTGAGTAGAGTGTGA

Letters in lower cases denote FANA nucleotides; RNA residues are denoted by “r”; the underlined letter in the templates for single-nucleotide incorporation denotes the templating nucleotide.

References

1 A. Nikoomanzar, M. R. Dunn and J. C. Chaput, *Curr Protoc Nucleic Acid Chem*, 2017, **69**, 4.75.1-4.75.20.

2 J. S. Paige, K. Y. Wu and S. R. Jaffrey, *Science*, 2011, **333**, 642–646.