Supplementary Information

Site-specific functionalisation of RNA by an unnatural base pair transcription system via click chemistry

Takumi Ishizuka,[a] Michiko Kimoto,[a,b] Akira Sato[a] and Ichiro Hirao*[a,b]

[a] RIKEN Systems and Structural Biology Center (SSBC) and [b] TagCyx Biotechnologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. Fax: +81-45-503-9645; Tel: +81-45-503-9644; E-mail: ihirao@riken.jp

Materials and Methods

Materials and Methods

1. Chemical syntheses

General.

^{1}H NMR, ^{13}C NMR and ^{31}P NMR spectra were recorded on a BRUKER (300-AVM) magnetic resonance spectrometer. DMSO-d_6 and D_2O were used as the solvents. Coupling constants (J) values are given in Hz and are correct to within 0.5 Hz. Signal patterns are indicated as br, broad; s, singlet; d, doublet; t, triplet; m, multiplet. All reagents were purchased from Aldrich, TCI (Tokyo Chemical Industry Co., Ltd.) or Wako (Wako Pure Chemical Industries, Ltd.). Thin layer chromatography was performed using TLC Silica gel 60 F_{254} (Merck). Compounds were visualized by staining with a potassium permanganate solution. Nucleoside purification was performed on a Gilson HPLC system with a preparative C18 column (μ-BONDASPHERE, Waters, 19 mm × 150 mm). The triphosphate derivatives were purified by chromatography on a DEAE-Sephadex A-25 column (300 × 15 mm) and a C18 column (CAPCELL PAK MG III, SHISEIDO, 4.6 mm × 250 mm). High resolution mass spectra (HRMS) and electrospray ionization mass spectra (ESI-MS) were recorded on a JEOL JM 700 mass spectrometer and a Waters micromass ZMD 4000 mass detector equipped with a Waters 2690 LC system, respectively.

1-(β-D-Ribofuranosyl)-4-iodopyrrole-2-carbaldehyde (4)

1-(β-D-Ribofuranosyl)-4-iodopyrrole-2-carbaldehyde (**4**) was synthesized by coupling 4-iodopyrrole-2-carbaldehyde[1] (**3**) with 2,3,5-tri-*O*-benzyl-D-ribofuranosyl chloride, which was prepared by the chlorination[2] of 2,3,5-tri-*O*-benzyl-D-ribofuranose, using CCl_4 and tris(dimethylamino)phosphane, followed by a treatment with BBr_3 as follows.[3] To a THF solution (9.2 ml) containing 2,3,5-tri-*O*-benzyl-D-ribofuranose (2.0 g, 4.76 mmol) and CCl_4 (693 μl, 7.13 mmol), hexamethylphosphorous triamide (1076 μl, 6.18 mmol) was added at –78 °C. The solution was stirred for 2 h at –78 °C and for 40 min at room temperature (solution A). To a solution of **3** (1.37 g, 6.18 mmol) in CH_3CN (50 mL) was added NaH (60% oil dispersion, 247 mg, 6.18 mmol). The reaction mixture was stirred at room temperature for 40 min. A solution of 2,3,5-tri-*O*-benzyl-D-ribofuranosyl chloride in THF (solution A) was then added. The reaction mixture was stirred at room temperature for 14 h. The product was separated by ethyl acetate and water. The organic layer was washed with saturated NH_4Cl (×2), dried with Na_2SO_4, and evaporated *in vacuo*. The product was purified by silica gel column chromatography (CH_2Cl_2 : MeOH = 19 : 1) to give 1-(2,3,5-tri-*O*-benzyl-D-ribofuranosyl)-4-iodopyrrole-2-carbaldehyde. To a solution of 1-(2,3,5-tri-*O*-benzyl-D-ribofuranosyl)-4-iodopyrrole-2-carbaldehyde in CH_2Cl_2 (28 ml) was added BBr_3 (1 M in CH_2Cl_2 solution, 15 ml, 15 mmol) at –78 °C. The reaction mixture was stirred for 3 h, and then 50% MeOH in CH_2Cl_2 (60 ml) was added. After stirring the solution at –78 °C for 10 min, 28% NH_3 aq. (10 ml) was added, and the reaction mixture was stirred until

it reached room temperature. The solution was added to CH_2Cl_2 (30 ml) and H_2O (30 ml). The water layer was isolated and washed with CH_2Cl_2 (×2), and the residue was evaporated *in vacuo*. The product was purified by reversed phase C18 HPLC to give compound **4** (237 mg, 14%, 2 steps total yield). ^{1}H NMR (DMSO-d_6, 300 MHz) δ 9.50 (d, J = 0.6 Hz, 1H), 7.89 (s, 1H), 7.21 (d, J = 1.8 Hz, 1H), 6.34 (d, J = 3.9 Hz, 1H), 5.33 (d, J = 5.7 Hz, 1H), 5.09-5.06 (m, 2H), 4.05-3.98 (m, 2H), 3.88-3.84 (m, 1H), 3.66 (ddd, J = 3.6, 5.4, 12.0 Hz, 1H), 3.54 (ddd, J = 3.3, 4.8, 12.0 Hz, 1H). ESI-MS for $C_{10}H_{11}INO_5$ $[M–H]^-$: calcd, 352.11; found, 352.04.

1-(β-D-Ribofuranosyl)-4-ethynylpyrrole-2-carbaldehyde (5)

1-(β-D-Ribofuranosyl)-4-iodopyrrole-2-carbaldehyde **4** (209 mg, 0.59 mmol) was co-evaporated with pyridine and toluene. To a solution of **4**, tetrakis (triphenylphosphine)palladium (34.3 mg, 0.030 mmol), CuI (18.4 mg, 0.095 mmol), and triethylamine (123 µl, 0.89 mmol) in DMF (3 ml), trimethylsilylacetylene (167 µl, 0.89 mmol) was added. The reaction was stirred at room temperature for 12 h. The product was evaporated *in vacuo*, and purified by silica gel column chromatography (CH_2Cl_2 : MeOH = 90 : 10) to give 1-(β-D-ribofuranosyl)-4-trimethylsilylethynylpyrrole-2-carbaldehyde. To a solution of 1-(β-D-ribofuranosyl)-4-trimethylsilylethynylpyrrole-2-carbaldehyde in THF (6 ml), a 1 M solution of TBAF (700 µl, 0.70 mmol) in THF was added. After 1 h at room temperature, the reaction mixture was evaporated *in vacuo*. The product was purified by silica gel column chromatography (CH_2Cl_2 : MeOH = 90 : 10) and RP-HPLC to give compound **5** (68.0 mg, 46%, 2 steps total yield). ^{1}H NMR (DMSO-d_6, 300 MHz) δ 9.53 (d, J = 0.6 Hz, 1H), 8.03 (s, 1H), 7.20 (d, J = 1.8 Hz, 1H), 6.33 (d, J = 3.3 Hz, 1H), 5.35 (d, J = 5.4 Hz, 1H), 5.09 (t, J = 5.1 Hz, 1H), 5.07 (d, J = 4.8 Hz, 1H), 4.03-4.00 (m, 2H), 3.96 (s, 1H), 3.88-3.86 (m, 1H), 3.68 (ddd, J = 3.3, 5.4, 12.0 Hz, 1H), 3.56 (ddd, J = 3.3, 4.8, 12.0 Hz, 1H). ^{13}C NMR (DMSO-d_6, 75 MHz) δ 180.30, 131.81, 131.79, 127.07, 105.06, 90.22, 85.02, 80.56, 78.01, 76.35, 69.71, 60.96. HRMS (FAB, 3-NBA matrix) for $C_{12}H_{14}NO_5$ $[M+H]^+$: calcd, 252.0872; found, 252.0867.

1-(β-D-Ribofuranosyl)-4-(octa-1,7-diynyl)pyrrole-2-carbaldehyde (6)

1-(β-D-Ribofuranosyl)-4-iodopyrrole-2-carbaldehyde **4** (158 mg, 0.45 mmol) was co-evaporated with pyridine and toluene. To a solution of **4**, tetrakis (triphenyl phosphine)palladium (25.2 mg, 0.022 mmol), CuI (14.0 mg, 0.072 mmol), and triethylamine (92.6 µl, 0.67 mmol) in DMF (3 ml), 1,7-octadiyne (290 µl, 2.24 mmol) was added. The reaction was stirred at room temperature for 12 h. The product was evaporated *in vacuo*, and purified by silica gel column chromatography (CH_2Cl_2 : MeOH = 90 : 10) and RP-HPLC to give compound **6** (36.4 mg, 25%). ^{1}H NMR (DMSO-d_6, 300 MHz) δ 9.51 (d, J = 0.3 Hz, 1H), 7.90 (s, 1H), 7.10 (d, J = 1.5 Hz, 1H), 6.33 (d, J = 3.9 Hz, 1H), 5.32 (d, J = 5.1 Hz, 1H), 5.06-5.05 (m, 2H), 4.05-3.97 (m, 2H), 3.88-3.83 (m, 1H), 3.66 (ddd, J = 3.3, 8.7, 12.0 Hz, 1H), 3.54 (ddd, J = 3.3, 8.7, 12.0 Hz, 1H), 2.76 (t, J = 2.6 Hz, 1H), 2.40-2.36 (m, 2H), 2.23-2.16 (m, 2H), 1.62-1.52 (m, 4H).

^{13}C NMR (DMSO-d_6, 75 MHz) δ 180.13, 131.71, 130.97, 126.83, 106.60, 90.03, 89.63, 85.03, 84.75, 76.27, 74.62, 71.82, 69.85, 61.08, 27.81, 27.65, 18.62, 17.71. HRMS (FAB, 3-NBA matrix) for $C_{18}H_{12}NO_5$ $[M+H]^+$: calcd, 332.1498; found, 332.1477.

Synthesis of nucleoside 5′-triphosphates (1) and (2)

To a solution containing a proton sponge (33 mg, 0.15 mmol) and either 1-(β-D-ribofuranosyl)-4-ethynylpyrrole-2-carbaldehyde **5** or 1-(β-D-ribofuranosyl)-4-(octa-1,7-diynyl)pyrrole-2-carbaldehyde **6** (0.1 mmol) in trimethyl phosphate (500 µl), $POCl_3$ (12 µl, 0.13 mmol) was added at 0 °C. The reaction mixture was stirred at 0 °C for 2 h. Tri-*n*-butylamine (120 µl, 0.5 mmol) was added to the reaction mixture, followed by 0.5 M bis(tributylammonium) pyrophosphate in a DMF solution (1.0 ml, 0.5 mmol). After 5 min, the reaction was quenched by the addition of 0.5 M triethylammonium bicarbonate (TEAB, 500 µl). The resulting crude product was purified by DEAE Sephadex A-25 column chromatography (eluted by a linear gradient of 50 mM to 1 M TEAB), and then by C18-HPLC (eluted by a gradient of CH_3CN, 5% to 50% for **1** or 20% to 50% for **2** in 100 mM triethylammonium acetate) to give either compound **1** (40.1 µmol, 40%) or compound **2** (33.6 µmol, 34%).

4-Ethynylpyrrole-2-carbaldehyde ribonucleoside 5′-triphosphate (1)

^{1}H NMR (D_2O, 300 MHz) δ 9.46 (d, J = 0.6 Hz, 1H), 7.99 (s, 1H), 7.34 (d, J = 1.8 Hz, 1H), 6.55 (d, J = 3.9 Hz, 1H), 4.46-4.42 (m, 2H), 4.32-4.27 (m, 3H), 3.41 (s, 1H). ^{31}P NMR (D_2O, 121 MHz) δ −9.58 (d, J = 20.0 Hz, 1P), −10.70 (d, J = 20.1 Hz, 1P), −22.46 (t, J = 20.0 Hz, 1P). ESI-MS for $C_{12}H_{15}NO_{14}P_3$ $[M–H]^–$: calcd, 490.17; found, 490.07. UV-vis spectrum (in 10 mM sodium phosphate buffer, pH 7.0), λ_{max} = 257 nm ($\varepsilon = 9.2 \times 10^3$), 303 nm ($\varepsilon = 8.9 \times 10^3$).

4-(Octa-1,7-diynyl)pyrrole-2-carbaldehyde ribonucleoside 5′-triphosphate (2)

^{1}H NMR (D_2O, 300 MHz) δ 9.46 (s, 1H), 7.84 (s, 1H), 7.25 (d, J = 1.5 Hz, 1H), 6.55 (d, J = 3.9 Hz, 1H), 4.46-4.43 (m, 2H), 4.31-4.21 (m, 3H), 2.47-2.43 (m, 2H), 2.37 (t, J = 2.4 Hz, 1H), 2.29-2.28 (m, 2H), 1.70-1.68 (m, 2H). ^{31}P NMR (D_2O, 121 MHz) δ −9.94 (d, J = 20.6 Hz, 1P), −10.63 (d, J = 19.4 Hz, 1P), −22.53 (t, J = 19.4 Hz, 1P). ESI-MS for $C_{18}H_{23}NO_{14}P_3$ $[M–H]^–$: calcd, 570.30; found, 570.30. UV-vis spectrum (in 10 mM sodium phosphate buffer, pH 7.0), λ_{max} = 258 nm ($\varepsilon = 1.05 \times 10^4$), 312 nm ($\varepsilon = 7.7 \times 10^3$).

Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) and 3-Azido-7-hydroxycoumarin

Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) as the ligand and 3-azido-7-hydroxycoumarin, used in the click reactions, were prepared according to previous reports.[4,5]

Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA)

^{1}H NMR (DMSO-d_6, 300 MHz) δ 8.02 (s, 3H), 4.66 (t, J = 4.8 Hz, 3H), 4.40 (t, J = 7.2 Hz, 6H), 3.62 (s, 6H), 3.38 (m, 6H), 1.95 (m, 6H). ^{13}C NMR (75 MHz, DMSO-d_6) δ 143.4, 123.9, 57.4, 47.0, 46.5, 32.9. HRMS (FAB, 3-NBA matrix) for $C_{18}H_{31}N_{10}O_3$ $[M+H]^+$: calcd, 435.2580; found, 435.2605.

3-Azido-7-hydroxycoumarin

^{1}H NMR (DMSO-d_6, 300 MHz) δ 7.60 (s, 1 H), 7.48 (d, J = 8.4 Hz, 1 H), 6.81 (d, J = 8.7, Hz, 1 H), 6.76 (s, 1H). ^{13}C NMR (75 MHz, DMSO-d_6) δ 160.3, 157.5, 152.8, 129.1, 127.9, 121.2, 113.8, 111.3, 102.0. HRMS (FAB, 3-NBA matrix) for $C_9H_6N_3O_3$ $[M+H]^+$: calcd, 204.0409; found, 204.0418.

References

1. Mitsui, T., Kimoto, M., Sato, A., Yokoyama, S. & Hirao, I. An unnatural hydrophobic base, 4-propynylpyrrole-2-carbaldehyde, as an efficient pairing partner of 9-methylimidazo[(4,5)-*b*]pyridine. *Bioorg. Med. Chem. Lett.* **13**, 4515–4518 (2003).
2. Rosemeyer, H. & Seela, F. Stereoselective synthesis of pyrrolo[2,3-*d*]pyrimidine α- and β-D-ribonucleosides from anomerically pure D-ribofuranosyl chlorides: solid-liquid phase-transfer glycosylation and ^{15}N-NMR spectra. *Helv. Chim. Acta* **71**, 1573–1585 (1988).
3. Hirao, I., Kimoto, M., Mitsui, T., Fujiwara, T., Kawai, R., Sato, A., Harada, Y. & Yokoyama, S. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **3**, 729–735 (2006).
4. Hong, V., Presolski, S. I., Ma, C. & Finn, M. G. Analysis and optimization of copper-catalyzed azide–alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed.* **48**, 9879–9883 (2009).
5. Sivakumar, K., Xie, F., Cash, B. M., Long, S., Barnhill, H. N. & Wang, Q. A fluorogenic 1,3-dipolar cycloaddition reaction of 3-azidocoumarins and acetylenes. *Org. Lett.* **6**, 4603–4606 (2004).

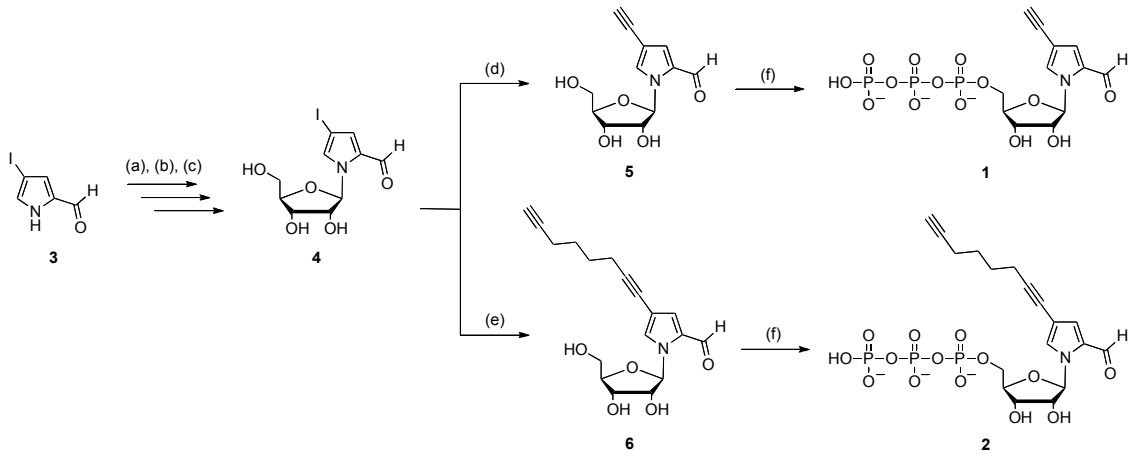


Figure S1. Synthesis of Eth-**Pa**TP and Eth-C4-**Pa**TP. Conditions and reagents: (a) NaH in CH_3CN, then 2,3,5-tri-*O*-benzyl-D-ribofuranosyl chloride; (b) BBr_3 in CH_2Cl_2; (c) 28% NH_3 aq. in CH_2Cl_2 : MeOH = 1:1; (d) trimethylsilyl acetylene, $Pd(PPh_3)_4$, CuI, triethylamine in DMF then TBAF in THF; (e) 1,7-octadiyne, $Pd(PPh_3)_4$, CuI, triethylamine in DMF; (f) $POCl_3$, proton sponge in $PO(OCH_3)_3$ then tri-*n*-butylamine, bis(tributylammonium)pyrophosphate in DMF.

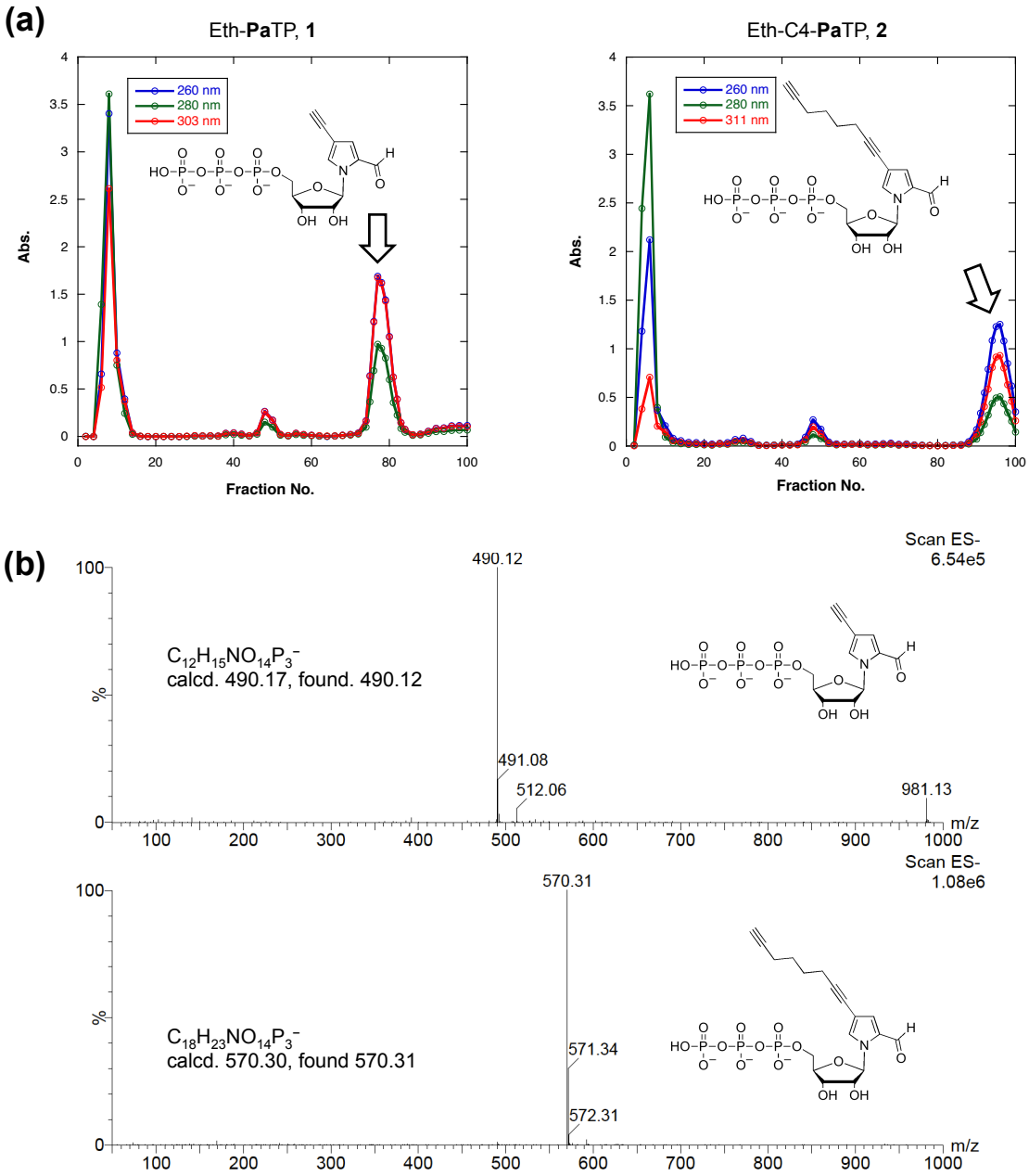
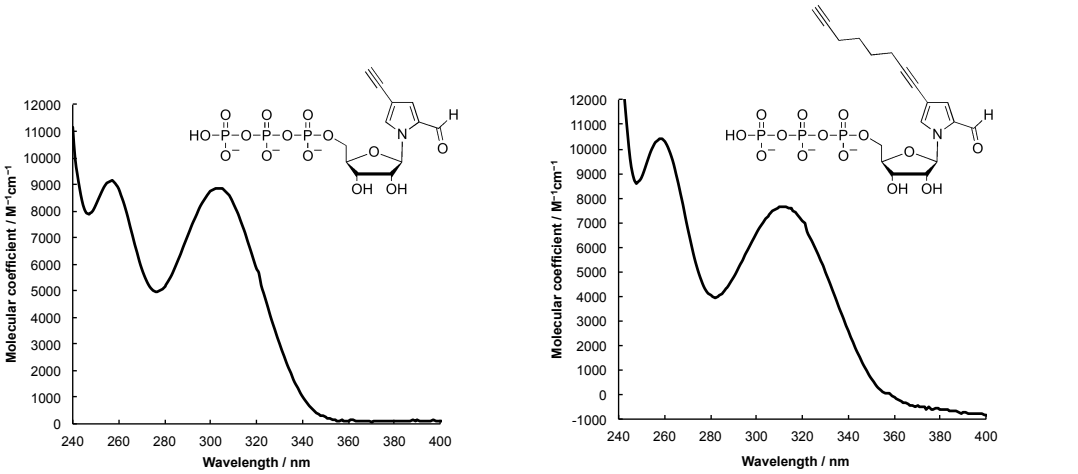


Figure S2. DEAE Sephadex ion-exchange column elution patterns and ESI-MS spectra of 5′-triphosphates. (a) DEAE Sephadex ion-exchange column elution patterns of compounds **1** and **2**. (b) ESI-mass spectra of compounds **1** (upper) and **2** (lower). (c) Molar absorption coefficients (ε) of the 5′-triphosphates of compounds **1** (left) and **2** (right).

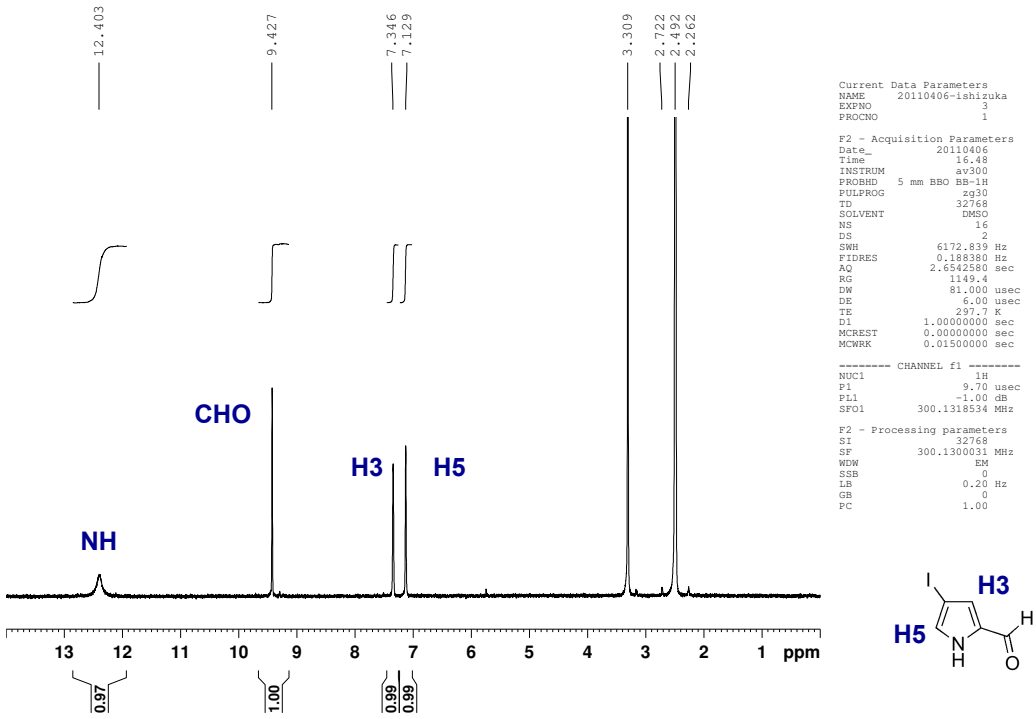
(c)

	Wavelength (nm)		ε
Eth-**Pa**TP, **1**	260		8,900
	257	max	9,200
	303	max	8,900
Eth-C4-**Pa**TP, **2**	260		10,400
	258	max	10,500
	312	max	7,700

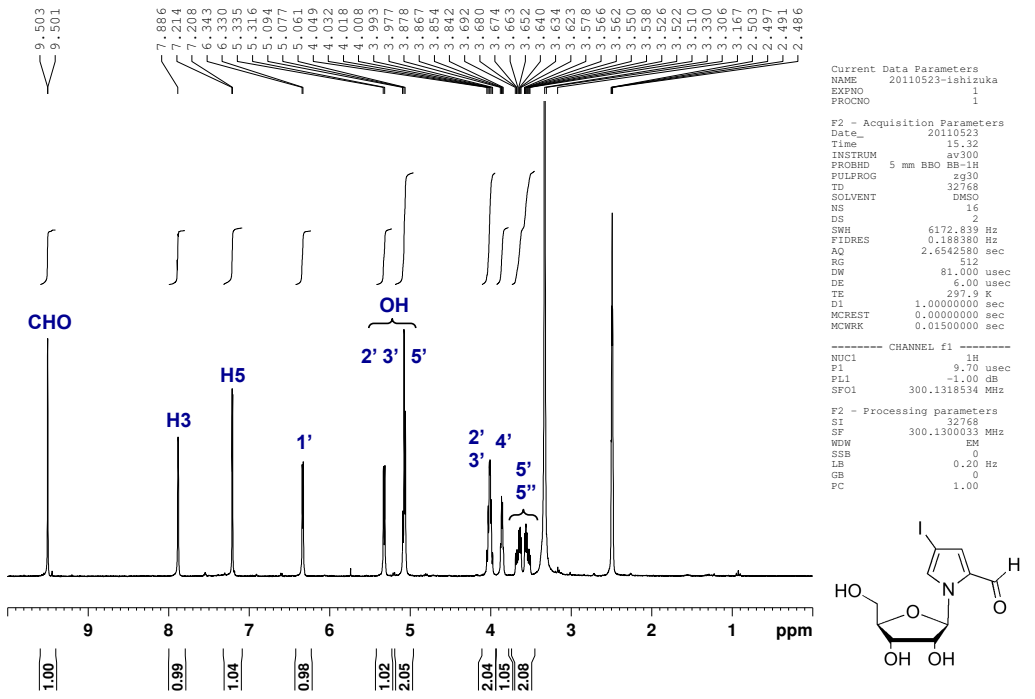


in 10 mM sodium phosphate buffer (pH7.0)

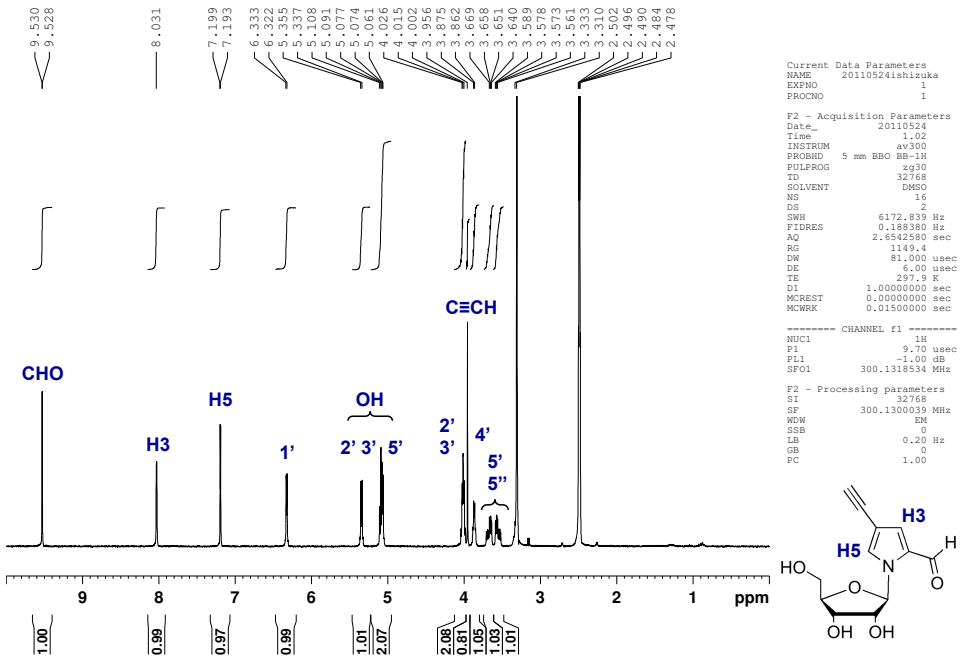
Figure S2. Continued.



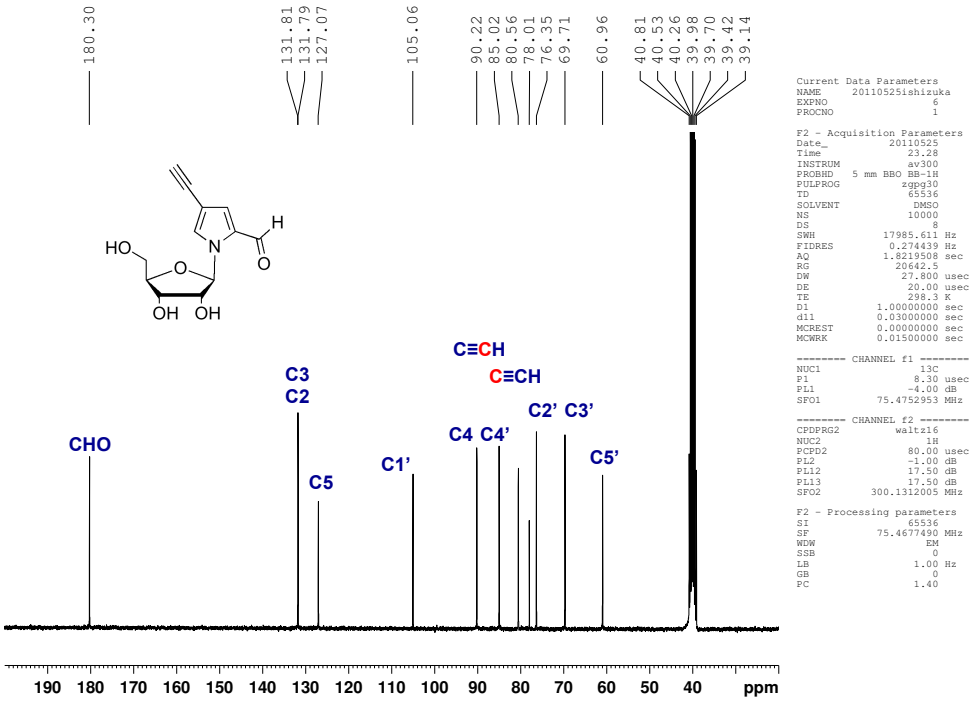
F1. ^{1}H NMR spectrum of compound **3**



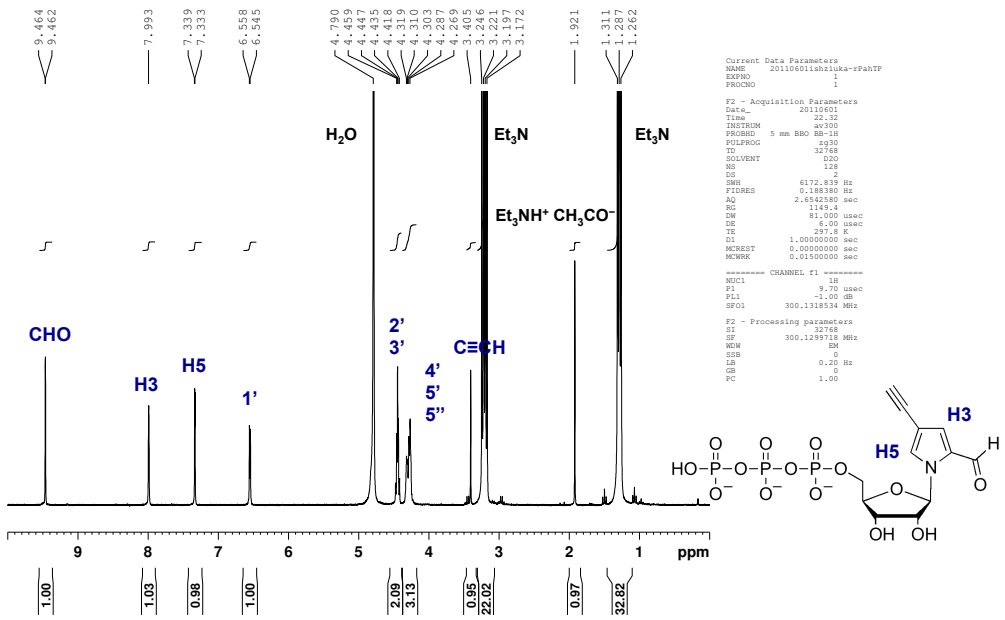
F2. ^{1}H NMR spectrum of compound **4**



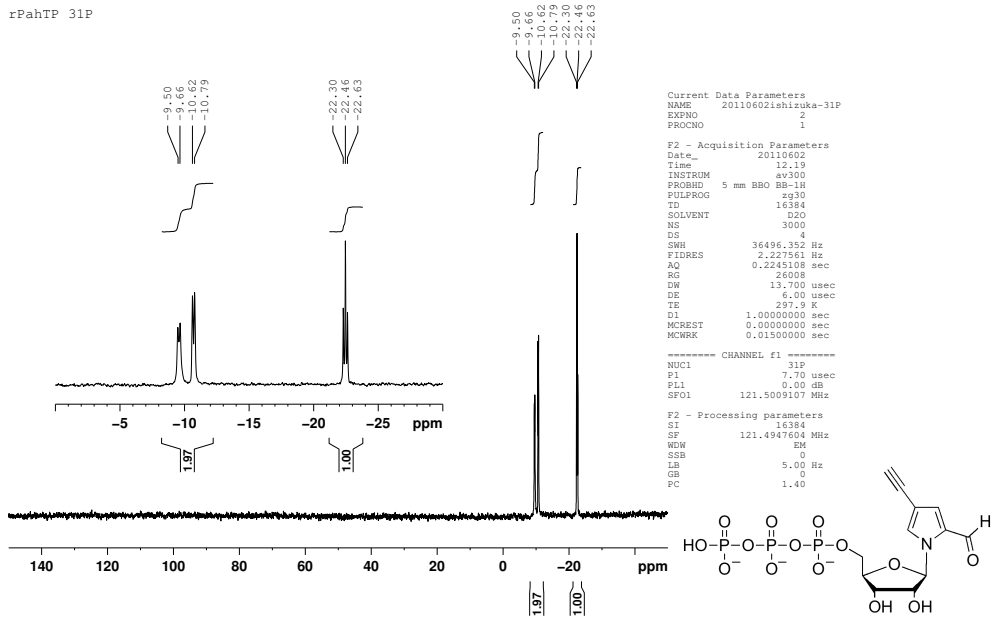
F3. ^{1}H NMR spectrum of compound **5**



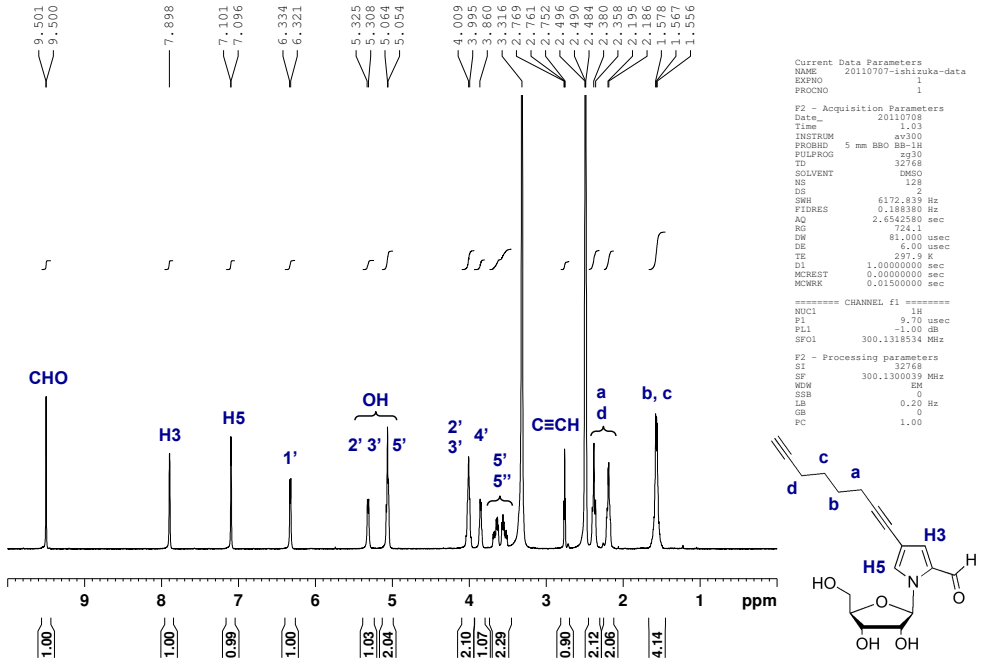
F4. ^{13}C NMR spectrum of compound **5**



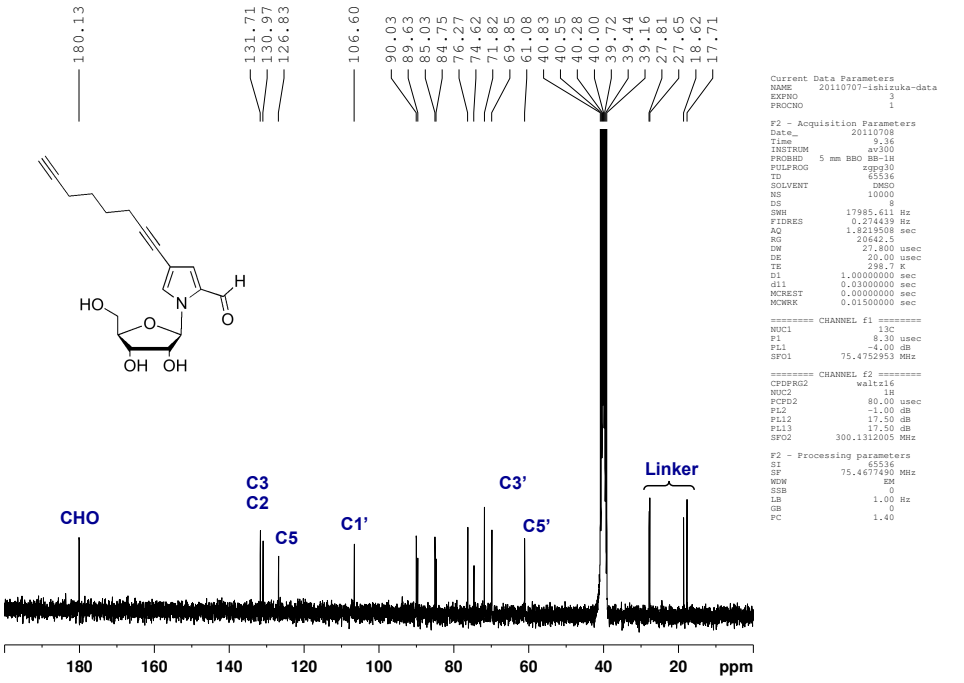
F5. ^{1}H NMR spectrum of compound **1**



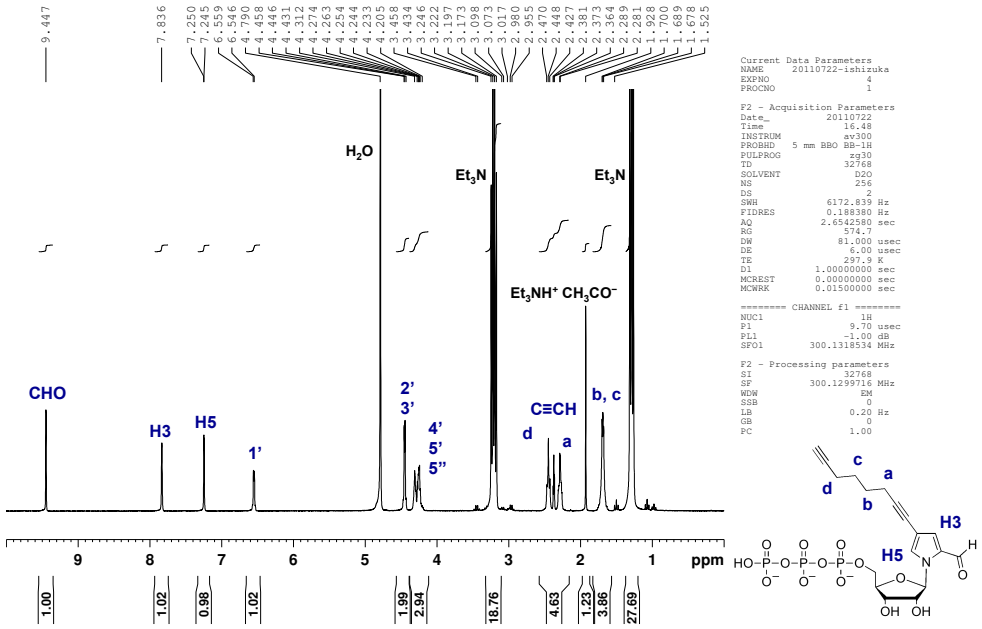
F6. ^{31}P NMR spectrum of compound **1**



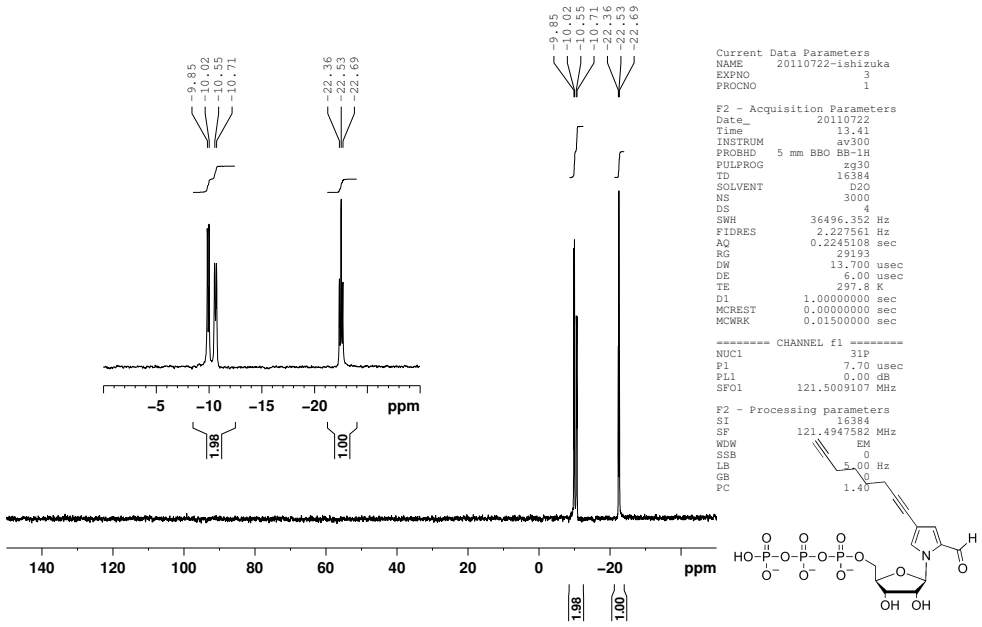
F7. ^{1}H NMR spectrum of compound **6**



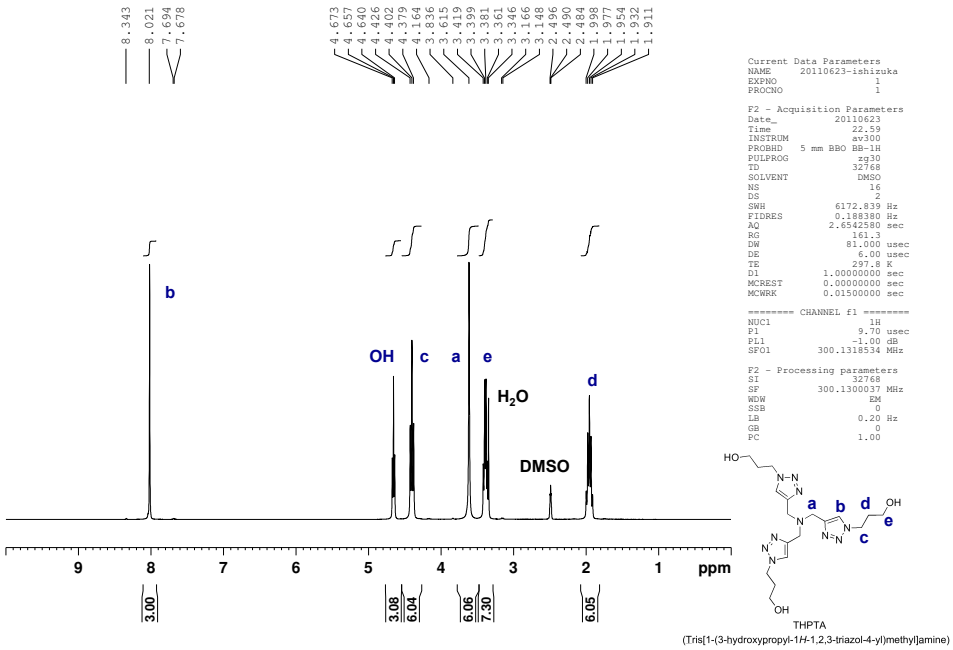
F8. ^{13}C NMR spectrum of compound **6**



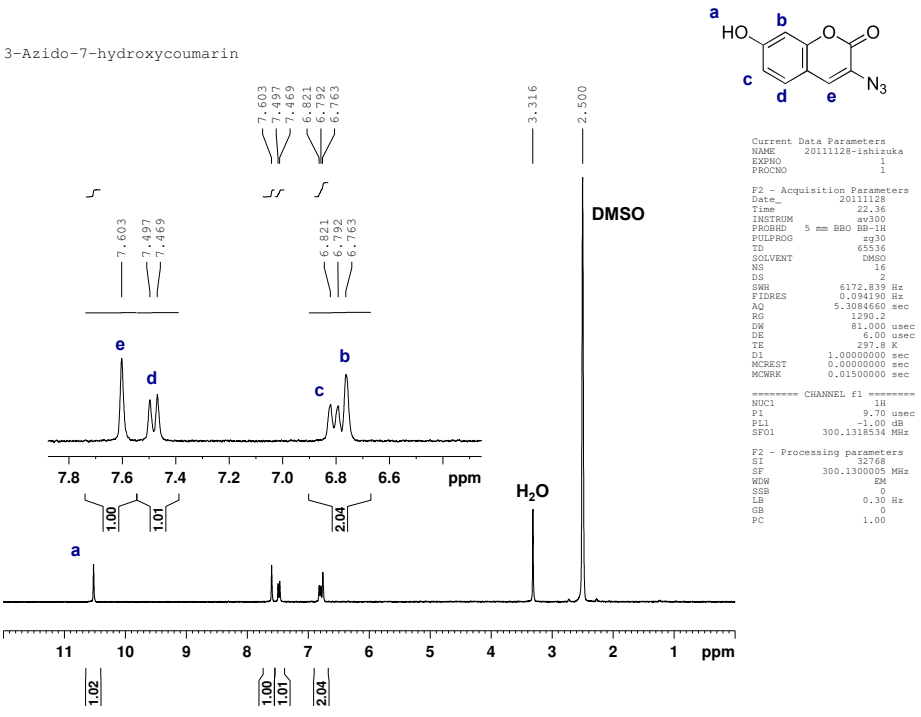
F9. ^{1}H NMR spectrum of compound **2**



F10. ^{31}P NMR spectrum of compound **2**



F11. ^{1}H NMR spectrum of THPTA



F12. ^{1}H NMR spectrum of 3-azido-7-hydroxycoumarin

2. Biological experiments

General.

Reagents and solvents were purchased from standard suppliers. The DNA templates were chemically synthesized with an automated DNA synthesizer (model 392, PerkinElmer Applied Biosystems) using the phosphoramidites of **Ds** (Glen Research) and the natural bases. The oligonucleotides were purified by gel electrophoresis. The substrates of Eth-**Pa** and Eth-C4-**Pa** were chemically synthesized from its ribonucleoside, as described above.

DNA templates for T7 transcription

The chemically synthesized DNA template (10 μM of a 35-mer template strand and a 21-mer non-template strand for 17-mer RNA synthesis) was annealed in a buffer containing 10 mM Tris-HCl (pH 7.6) and 10 mM NaCl, by heating at 95 °C and slow cooling to 4 °C. The sequences of the DNA fragments were as follows: 5′-ATAATACGACTCACTATAGGG (the non-template DNA, 21-mer), 5′-TATTATGCTGAGTGATATCCCTTAGGGCTCNTCAC (the template DNA, 35-mer, N = **Ds** or A). To prepare DNA templates for tRNA synthesis, the DNA fragments were chemically synthesized with the 2′-*O*-methylribonucleoside amidites (Glen Research) and the natural bases. The sequences of the DNA fragments were as follows:
5′-GATAATACGACTCACTATAGCCAGGGTGGCAGAGGGGCTATGCGGCGGA CTCTAGATCCGCTTTACCCCGG (5′-non-template DNA, 71-mer), 5′-TmGmGAGCCAGGGCCCGGATTCGAACCGGGGTAAAGCGGATCTANAGTC CGCCGCATAGCCC (3′-template DNA, 60-mer, N = **Ds** or G, Tm = 2′-*O*-methylthymidine, Gm = 2′-*O*-methylguanosine). The double-stranded DNA template (94 bp) was prepared by annealing the 5′-non-template and 5′-template DNA fragments, followed by primer extension with the Klenow fragment (Takara). The extended product was purified with Micropure-EZ and Microcon 30 filters (Millipore).

T7 transcription

Transcription was performed in a reaction buffer (20 μl), containing 40 mM Tris-HCl (pH 8.0), 24 mM $MgCl_2$, 2 mM spermidine, 5 mM DTT and 0.01% Triton X-100, in the presence of 1 mM natural NTPs, 0 or 1 mM Eth-**Pa**TP and Eth-C4-**Pa**TP, 2 μM DNA template and 50 U T7 RNA polymerase (Takara), unless otherwise indicated. Transcription of the 17-mer RNA with the optimized conditions for Eth-C4-**Pa**TP was performed using 2 mM Eth-C4-**Pa**TP and 4 μM DNA template, and transcription for tRNA was performed in the presence of 2 mM natural NTPs and 0 or 2 mM Eth-C4-**Pa**TP. To prepare internally ^{32}P-labeled transcripts, 2 to 3 μCi [α-^{32}P]ATP (PerkinElmer) was added to the transcription reaction. To prepare transcripts labeled with ^{32}P at the 5′-end, the transcription was performed in the presence of 3 μCi

[γ-^{32}P]GTP (PerkinElmer), or the gel-purified transcripts were labeled with [γ-^{32}P]ATP (PerkinElmer) and T4 polynucleotide kinase (Takara) after treatment with Antarctic phosphatase (New England Biolabs). The use of [α-^{32}P]ATP facilitated the analyses of the transcription yields and the nucleotide-compositions. After an incubation at 37 °C for 3 h (for 17-mer) or 6 h (for tRNA), the reaction was quenched by the addition of a dye solution (20 μl) containing 10 M urea and 0.05% BPB. The mixture was heated at 75 °C for 3 min, and the products were fractionated on a 20% polyacrylamide–7 M urea gel (for 17-mer) or a 10% polyacrylamide–7 M urea gel (for tRNA) and analyzed with a bio-imaging analyzer, FLA-7000 (FUJIFILM).

Nucleotide-composition analysis in T7 transcription (17-mer RNA)

The full-length 17-mer products, internally labeled with [α-^{32}P]ATP, were purified on a 15% polyacrylamide–7 M urea gel, eluted from the gel with water, and precipitated with ethanol and 0.05 A_{260} units of *E. coli* tRNA (Sigma-Aldrich). The transcripts were digested by 0.075 U/μl RNase T2 (Sigma-Aldrich) at 37 °C for 2 h, in 15 mM sodium acetate buffer (pH 4.5). The digestion products were analyzed by 2D-TLC, using a Merck HPTLC plate (10 × 10 cm, Merck) with the following developing solvents: isobutyric acid/NH_4OH/H_2O (66:1:33 v/v/v) for the first dimension, and isopropyl alcohol/HCl/H_2O (70:15:15 v/v/v for Eth-**Pa** transcripts and 75:15:10 v/v/v for Eth-C4-**Pa** transcripts) for the second dimension. The products on the gels and the TLC plates were analyzed with the FLA-7000 bioimager.

Click reaction of the transcripts

To 20 μL (total volume) of the THPTA ligand (Tris[1-(3-hydroxypropyl-1*H*-1,2,3-triazol-4-yl) methyl]amine; 500 μM), sodium ascorbate (1 mM) and $CuSO_4$ (100 μM) were added sequentially. To the above solution, the transcribed RNAs (5 μM for 17-mer RNA, 3 μM for 75-mer tRNA) and 30 molar equivalents of azide reagents (Alexa Fluor 488 azide, Alexa Fluor 594 azide or Biotin azide; Invitrogen) were added, unless otherwise indicated. After an incubation at 37 °C for 3 h, the products were fractionated on a gel and analyzed with the FLA-7000 bioimager. The products stained with SYBR Gold (Lonza) were analyzed with a bio-imaging analyzer, LAS-4000 (FUJIFILM).

Analysis of the bioinylated 75-mer transcripts

The biotinylated transcript was detected by a gel mobility-shift assay, using streptavidin (Promega). The mixture (10 μl), containing 2 pmol ^{32}P-labeled transcripts and 100 pmol streptavidin, was incubated for 1 h at 20 °C in 10 mM Tris-HCl buffer (pH 7.6), containing 50 mM NaCl and 10 mM EDTA, and the biotinylated tRNA–streptavidin complexes were analyzed by gel electrophoresis. To determine the clicked Biotin-**Pa** insertion position, each tRNA (75-mer) labeled with ^{32}P at the 5′-end was partially digested by RNase T1 (0.17 U/μl, Amersham USB) for 15 min at 55 °C, in 13.3 mM sodium citrate buffer (pH 5.0), containing 4.7 M urea, 0.7 mM EDTA and 0.17 mg/ml *E.*

coli tRNA, and by alkali for 15 min at 90 °C, in 32 mM sodium carbonate buffer (pH 9.1) containing 0.6 mM EDTA. The alkali-digested RNA solution (9 µl) was mixed with 11 µl of 20 mM Tris-HCl buffer (pH 7.6) containing 150 mM NaCl, and a 10 µl portion was incubated with 0.4 mg of Streptavidin magnetic beads (New England Biolabs) for 5 min at room temperature (25 °C). Then, 5 µl aliquots of the supernatant and the other digested samples were analyzed on a 10% polyacrylamide–7 M urea gel, and the band patterns were detected with the FLA-7000 bioimager.

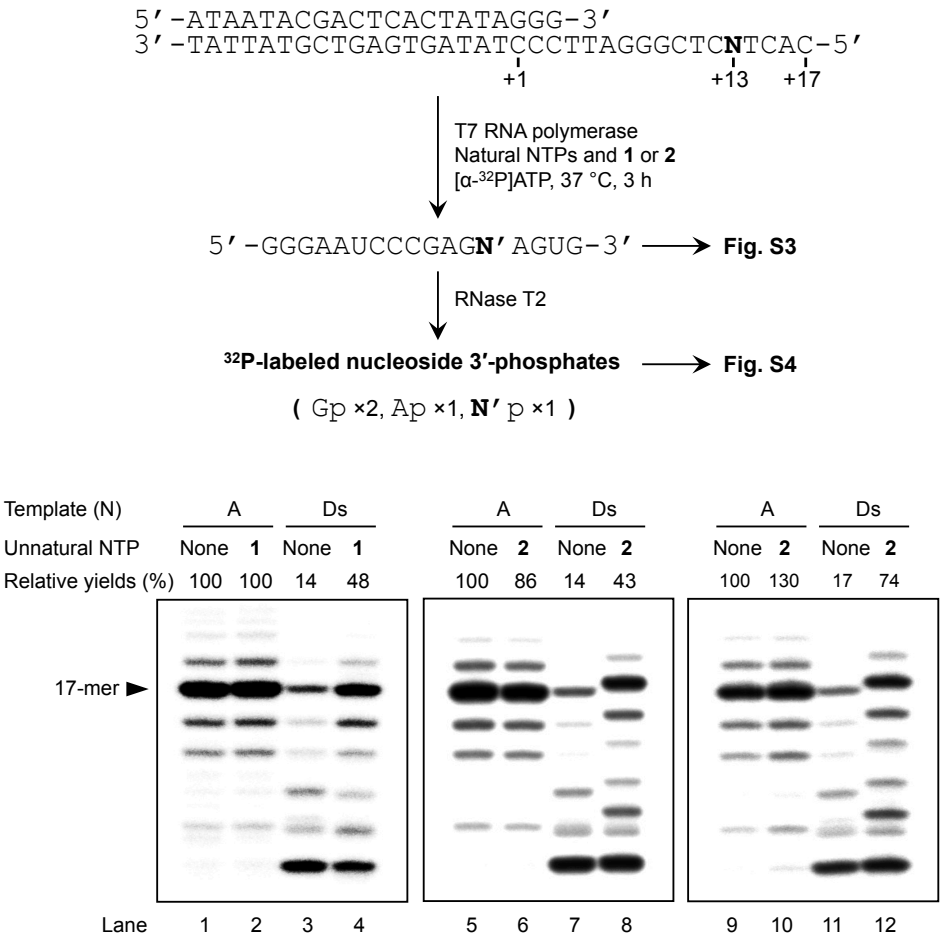


Figure S3. Gel electrophoresis of transcripts using the templates (N = **Ds** or A) with the natural NTPs in the presence or absence of **1** or **2**. Transcripts were internally labeled with [α-^{32}P]ATP. The relative yields of each transcript were determined by comparison to the yields of native transcripts from templates (N = A) consisting of the natural bases. Conditions (for lanes 1-8): [NTPs] = 1 mM, 2 μCi [α-^{32}P] ATP, [**1** or **2**] = 0 or 1 mM, [Template] = 2 μM, [Non-template] = 2 μM, [$MgCl_2$] = 24 mM, [spermidine] = 2 mM, [DTT] = 5 mM, 0.01% Triton X-100, and 2.5 U/μl T7 RNAP in 40 mM Tris-HCl buffer (pH 8.0), 37 °C, 3 h. The transcripts were analyzed on a 20% denaturing polyacrylamide-7 M urea gel. Conditions (for lanes 9-12): [**2**] = 0 or 2 mM, [Template] = 4 μM, and [Non-template] = 4 μM. Except for these conditions, the others are the same as described above.

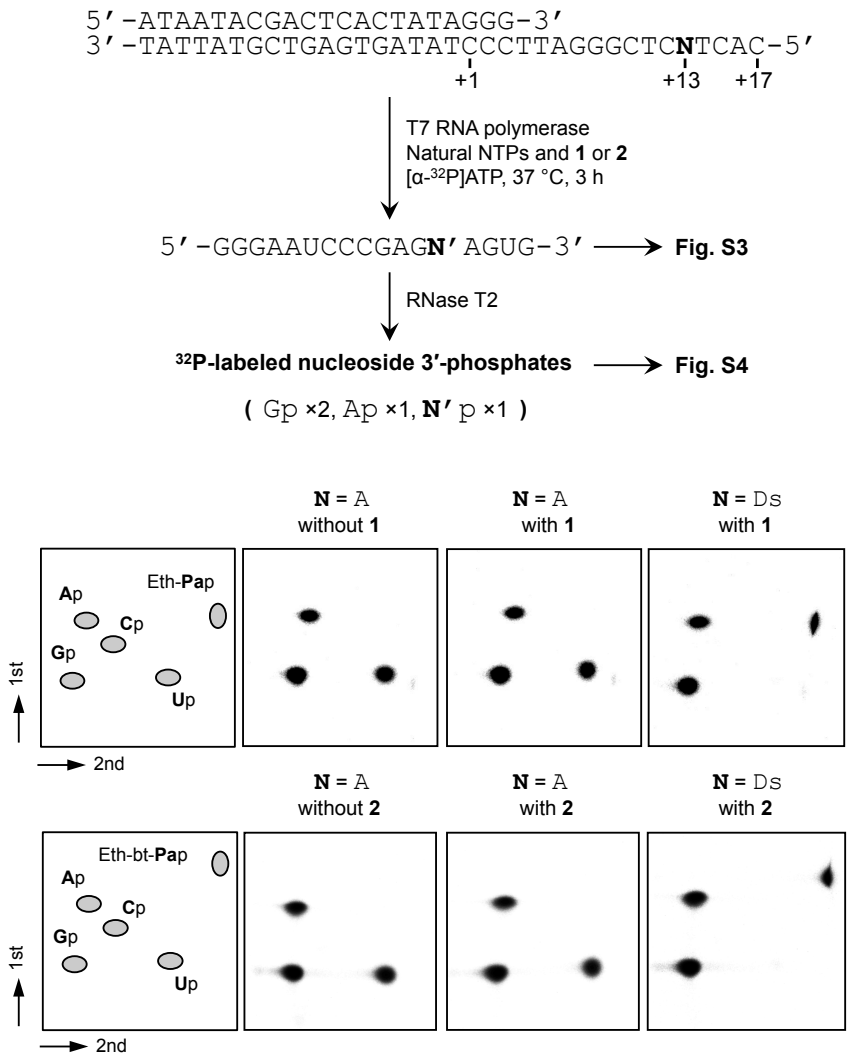


Figure S4. 2D-TLC analysis of the labeled ribonucleoside 3′-phosphates obtained from the nuclease digestion of the transcripts (17-mer). The digestion products were analyzed by 2D-TLC, using an HPTLC plate with the following developing solvents: isobutyric acid/NH_4OH/H_2O (66:1:33 v/v/v) for the first dimension, and isopropyl alcohol/HCl/H_2O (70:15:15 v/v/v for Eth-**Pa** transcripts or 75:15:10 v/v/v for Eth-C4-**Pa** transcripts) for the second dimension. All experiments were performed in triplicate.

Table S1 Nucleotide-composition analysis of T7 transcripts.

Entry	Template (N)[a]	Unnatural NTP	(mM)	Composition of nucleotides incorporated as 5' neighbor of A					
				Ap	Gp	Cp	Up	Eth-**Pa**p	Eth-C4-**Pa**p
1	N = A	–	0	1.00[b] [1][c] (0.01)[d]	2.00 [2] (0.01)	n.d.[e] [0] (<0.01)	1.00 [1] (0.01)	n.d. [0] (–)	n.d. [0] (–)
2		**1**	1	0.98 [1] (<0.01)	1.99 [2] (0.01)	n.d. [0] (<0.01)	1.02 [1] (<0.01)	n.d. [0] (<0.01)	–
3	N = Ds		1	0.97 [1] (0.04)	2.02 [2] (0.03)	n.d. [0] (<0.01)	0.01 [0] (<0.01)	0.99 [1] (0.01)	–
4	N = A	**2**	1	1.01 [1] (<0.01)	1.99 [2] (0.01)	n.d. [0] (<0.01)	1.00 [1] (0.01)	–	n.d. [0] (<0.01)
5	N = Ds		1	1.00 [1] (<0.01)	2.00 [2] (0.01)	n.d. [0] (<0.01)	0.01 [0] (<0.01)	–	0.99 [1] (<0.01)
6	N = A		2	1.01 [1] (<0.01)	1.99 [2] (0.02)	n.d. [0] (<0.01)	1.01 [1] (0.01)	–	n.d. [0] (<0.01)
7	N = Ds		2	0.99 [1] (<0.01)	1.99 [2] (0.01)	0.01 [0] (<0.01)	0.02 [0] (<0.01)	–	0.99 [1] (<0.01)

[a]The nucleotide sequence of the transcription template was 5'-CACT**N**CTCGGGATTCCCTATAGTGAGTCGTATTAT-3'.
[b]The values were determined using the following formula: (radioactivity of each nucleotide)/[total radioactivity of all nucleotides (3'-monophosphates)] × (total number of nucleotides at 5' neighbor of [α-^{32}P]ATP).
[c]The theoretical number of each nucleotide is shown in brackets.
[d]SDs are shown in parentheses.
[e]Not detected.

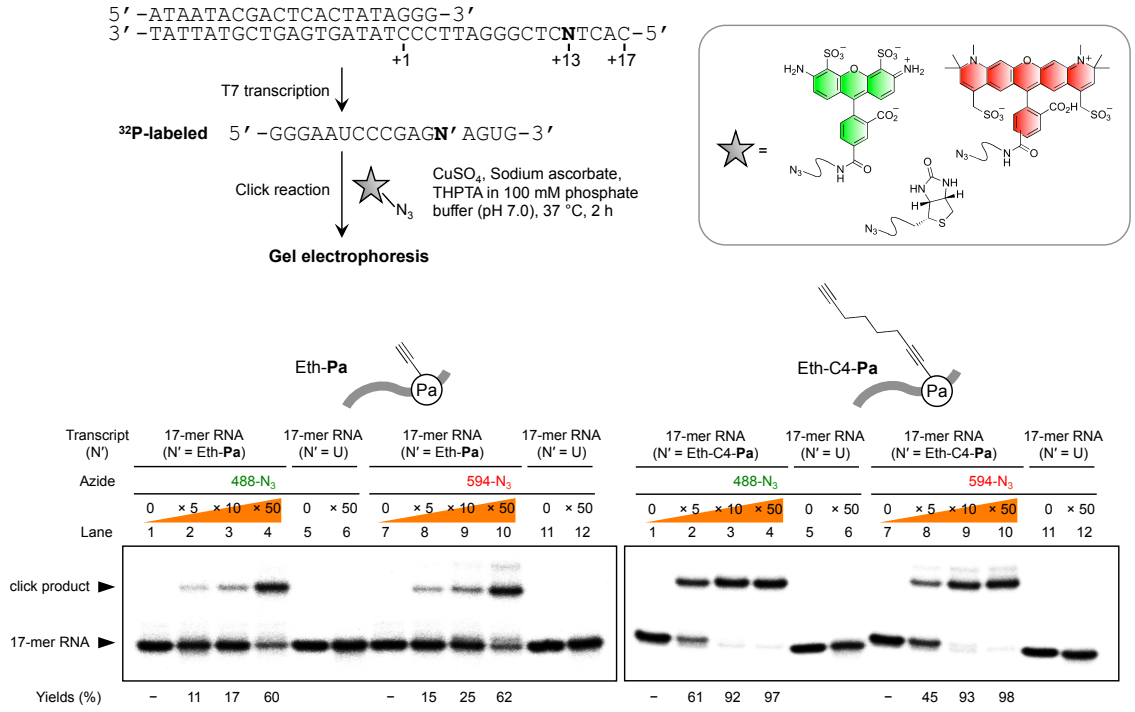


Figure S5. Site-specific labeling of the transcribed 17-mer RNA fragments containing Eth-**Pa** or Eth-C4-**Pa**, using the click reaction. To calculate the efficiency of the modification, a ^{32}P-labeled RNA fragment (17-mer) containing Eth-**Pa** or Eth-C4-**Pa** (5 μM) was labeled with a fluorescent dye azide reagent, Alexa Fluor 488 (488-N_3) or Alexa Fluor 594 (594-N_3), and the biotin azide (Biotin-N_3) reagent (150 μM each) in the presence of $CuSO_4$ (100 μM), sodium ascorbate (1 mM), and THPTA (500 μM), in 100 mM sodium phosphate buffer (pH 7.0) at 37 °C for 3 h. The efficiency (%) of the modification was calculated from the intensities of the bands corresponding to the modified 17-mer and the original 17-mer.

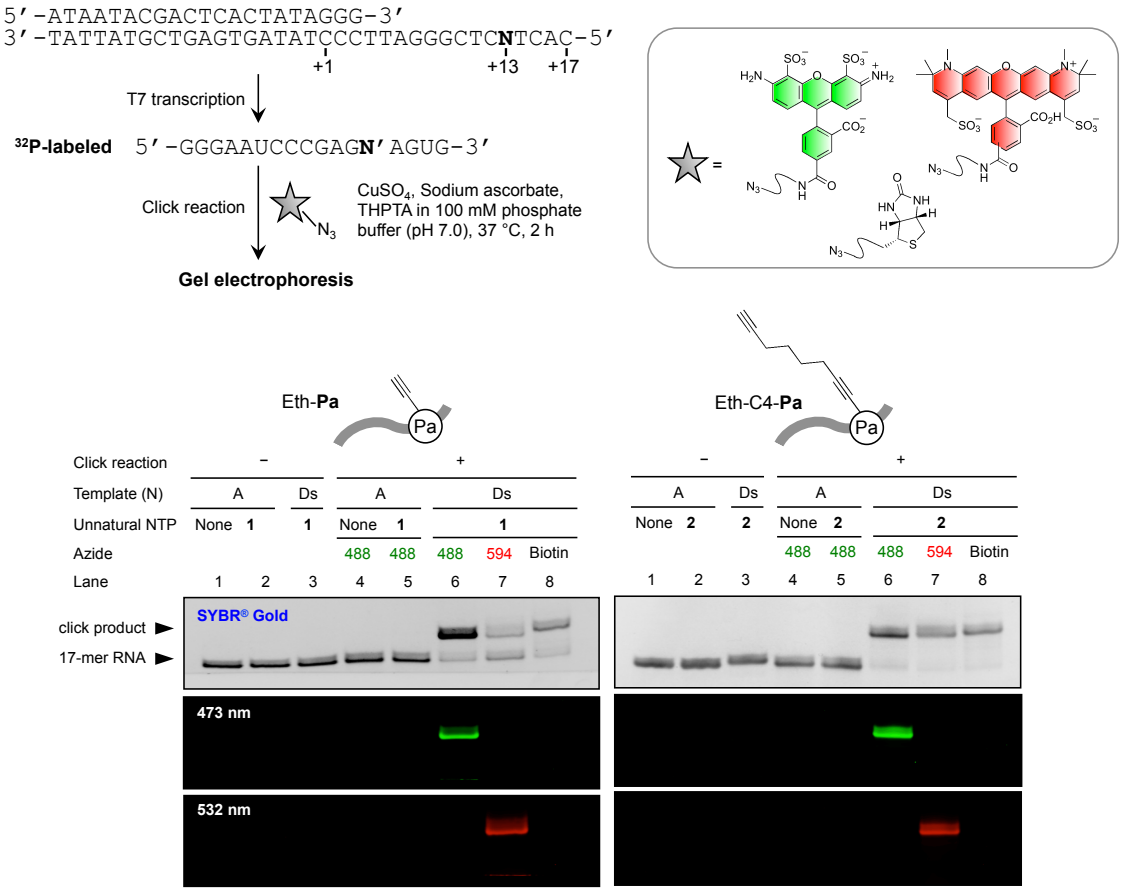


Figure S6. Site-specific labeling of the transcribed 17-mer RNA fragments containing Eth-**Pa** or Eth-C4-**Pa**, using the click reaction. A transcribed RNA fragment (17-mer) containing Eth-**Pa** or Eth-C4-**Pa** (5 μM) was labeled with fluorescent dye azide reagents (150 μM). The click reaction conditions are described in the legend to Figure S5. The clicked products were analyzed by 20% polyacrylamide gel electrophoresis. The clicked products on the gel were first detected with an FLA-7000 imager (GE Healthcare) in the FAM mode excitation mode (473 nm/emission filter Y520) and the Cy3 mode (excitation 532 nm/emission filter O580). The RNA fragments on the gel were then stained by SYBR Gold (Invitrogen) and detected with an LAS-4000 imager (GE Healthcare) in the SYBR Green mode (excitation 460 nm/emission filter Y515-Di).

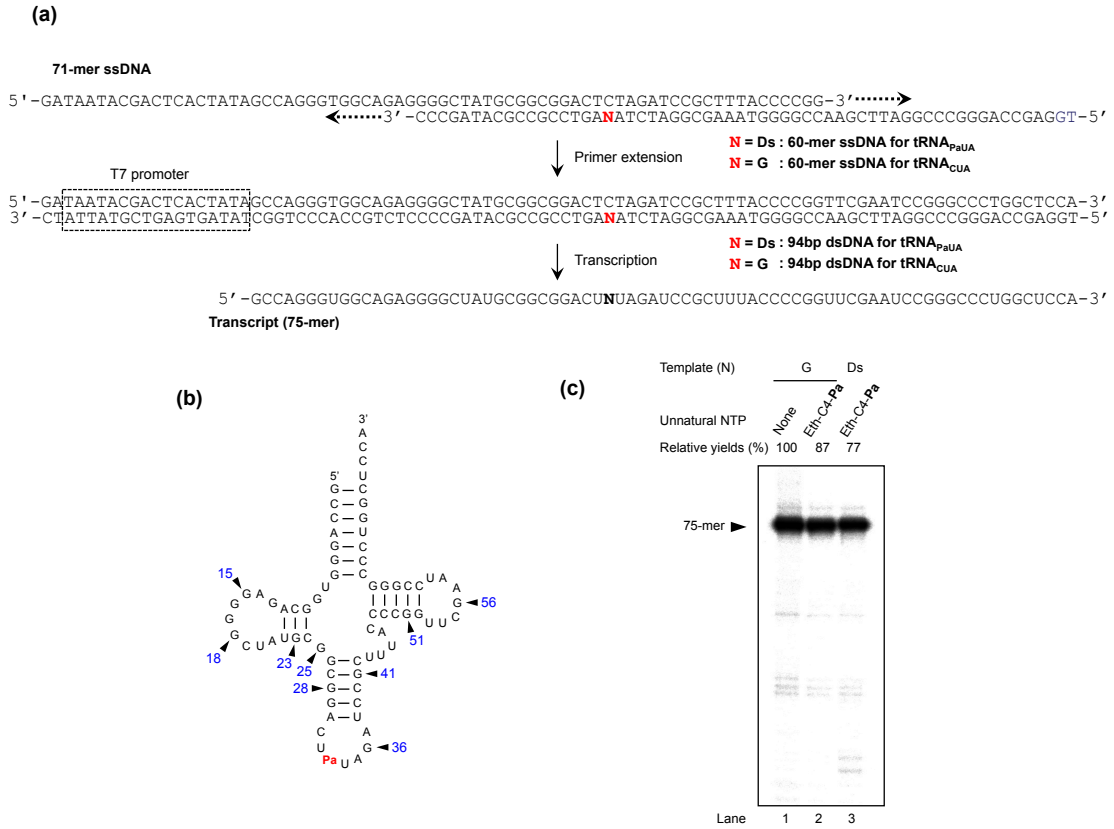


Figure S7. (a) Overview and sequences used in this study of T7 transcription mediated by the **Ds** and modified **Pa** pairing for tRNA (75-mer). (b) Secondary structure of the tRNA used in this study. Some guanines (denoted as G) are numbered according to the sequence analysis in Figure 2e. (c) Gel electrophoresis of transcripts using the templates (N = **Ds** or A) with the natural NTPs (1 mM), in the presence (1 mM) or absence of **2**. Transcripts were internally labeled with [α-^{32}P]ATP. The relative yields of each transcript were determined by comparison to the yields of native transcripts from templates (N = A) consisting of the natural bases. Experiments were performed in triplicate.

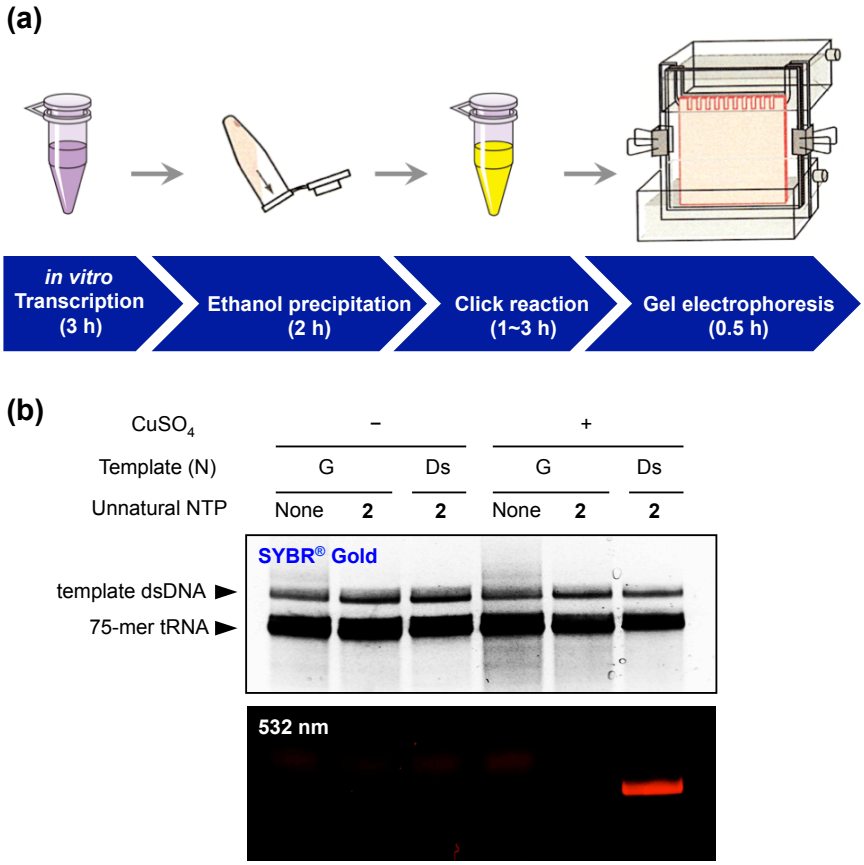


Figure S8. One-pot preparation of tRNA site-specifically modified with fluorescent dye. (a) The experimental scheme. The conditions for the click reaction are the same as shown in Figure 2b. (b) The transcribed tRNA (75-mer) containing Eth-C4-**Pa** was labeled with 594-N_3.