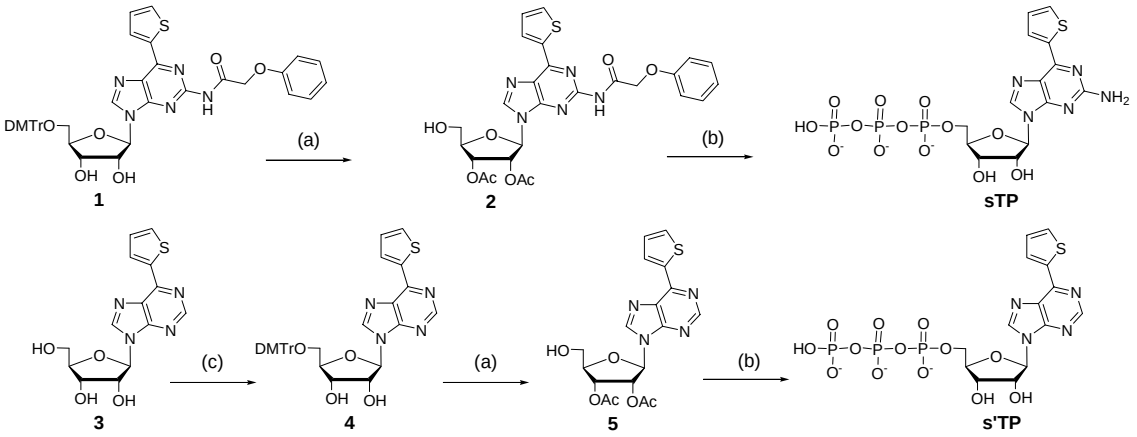
Supporting materials

Fluorescent probing for RNA molecules by an unnatural base pair system

Michiko Kimoto[1], Tsuneo Mitsui[1], Yoko Harada[1], Akira Sato[1], Shigeyuki Yokoyama[1,2,3,*], Ichiro Hirao[1*]

[1]Protein Research Group, RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, [2]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, and [3]RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan

I. Synthesis of the substrate of s (sTP)



Reagents and abbreviations: (a) acetic anhydride, pyridine, then dichloroacetic acid, dichloromethane, 93% for **2** and 87% for **5**; (b) 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one, dioxane, pyridine, tri-*n*-butylamine, bis(tributylammonium)pyrophosphate, DMF, then I_2/pyridine, water, NH_4OH, 44% for **s**TP and 48% for **s'**TP. DMTr: 4,4′-dimethoxytrityl, Ac: acetyl; (c) 4,4′-dimethoxytrityl chloride, pyridine, 90%. DMTr: 4,4′-dimethoxytrityl, Ac: acetyl.

General methods and synthesis of compounds. Reagents and solvents were purchased from standard suppliers and used without further purification. Reactions were monitored by thin-layer chromatography (TLC), using 0.25 mm silica gel 60 plates impregnated with 254 nm fluorescent indicator (Merck). ^{1}H NMR, ^{13}C NMR, and ^{31}P NMR spectra were recorded on JEOL EX270 and BRUKER (300-AVM) magnetic resonance spectrometers. Nucleoside purification was performed on a Gilson HPLC system with a preparative C18 column (Waters Microbond Sphere, 150 × 19 mm). The triphosphate derivatives were purified with a DEAE-Sephadex A-25 column (300 × 15 mm) and a C18 column (Synchropak RPP, 250 × 4.6 mm, Eichrom Technologies). 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.

2-*N*-Phenoxyacetyl-6-(2-thienyl)-9-[2,3-di-*O*-(acetyl)-1-β-D-ribofuranosyl]purine (2). Compound **1**[1,2] (790 mg, 0.98 mmol) was co-evaporated with dry pyridine three times. To the residue in dry pyridine (10 ml), acetic anhydride (370 μl, 4.0 mmol) was added. The mixture was stirred at room temperature for 12 h. The solution was poured into 5% $NaHCO_3$, and then the product was extracted with ethyl acetate. After drying over Na_2SO_4, the solvent was evaporated *in vacuo*. To the residue in dry methylene chloride (100 ml) was added dichloroacetic acid (1.0 ml) at 0°C. The solution was stirred for 15 min at 0°C. The mixture was poured into 5% aqueous sodium hydrogen carbonate, and the product was extracted with methylene chloride. After drying over Na_2SO_4, the solvent was evaporated *in vacuo*. The residue was subjected to silica gel column chromatography using methylene chloride : ethylacetate (2:1, v/v) as an eluent to afford 519 mg of **2** in 93% yield.

^{1}H NMR (300 MHz, DMSO-*d*6) δ 10.82 (s, 1H), 8.79 (s, 1H), 8.63 (d, 1H, J=3.7 Hz), 7.96 (d, 1H, J=4.0 Hz), 7.37-7.28 (m, 3H), 6.98 (m, 3H), 6.27 (d, 1H, J=6.5 Hz), 5.91 (t, 1H, J=6.1 Hz), 5.57 (dd, 1H, J=3.0 and 5.6 Hz), 5.33 (t, 1H, J=5.4 Hz), 5.12 (s, 2H), 4.27-4.24 (m, 1H), 3.79-3.66 (m, 2H), 2.13 (s, 3H), 1.99 (s, 3H). ^{13}C NMR (75 MHz, DMSO-*d*6) δ 169.49, 169.22, 167.79, 157.97, 152.79, 152.37, 149.40, 143.77, 139.05, 132.88, 132.27, 129.41, 129.10, 125.26, 120.85, 114.48, 84.68, 83.95, 72.85, 71.24, 67.35, 60.95, 20.45, 20.14. ESI-MS for $C_{26}H_{25}N_5O_8S$: calcd. 566.13 $[M-H]^-$, found. 565.74 $[M-H]^-$.

2-Amino-6-(2-thienyl)-9-(1-β-D-ribofuranosyl)purine 5'-triphosphate (sTP). Compound **2** (57 mg, 0.1 mmol) was dissolved in pyridine and evaporated to dryness *in vacuo*. The residue was dissolved in pyridine (100 μl) and dioxane (300 μl). A 1 M solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one in dioxane (110 μl, 0.11 mmol) was then added. After 10 min, tri-*n*-butylamine (100 μl) and 0.5 M bis(tributylammonium)pyrophosphate in DMF (300 μl) were added to the mixture. The reaction mixture was stirred at room temperature for 10 min. A solution of 1% iodine in pyridine/water (2.0 ml, 98/2, v/v) was then added. After 15 min, a 5% aqueous solution (150 μl) of $NaHSO_3$, followed by 5.0 ml of water, was added to the reaction mixture. The solution was stirred at room temperature for 30 min, and then concentrated ammonia (20 ml) was added. Ammonolysis was carried out at 55°C for 12 h. The solution was concentrated *in vacuo*, and the 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 0% to 30% CH_3CN in 100 mM triethylammonium acetate) to give the nucleoside 5′-triphosphate (**s**TP) in 44% yield. ^{1}H NMR (270 MHz, D_2O) δ 8.33 (s, 1H), 8.16 (d, 1H, J=4.0 Hz), 7.66 (d, 1H, J=4.9 Hz), 7.19 (dd, 1H, J=4.0 and 4.9 Hz), 5.94 (d, 1H, J=5.9 Hz), 4.62-4.44 (m, 4H), 4.32 (m, 1H), 4.19 (m, 2H), 3.05 (q, 24H, J=7.3 and 14.8 Hz), 1.13 (t, 36H, J=7.3 Hz). ^{31}P NMR (109 MHz, D_2O) δ -9.24 (d, 1H, J=20.1 Hz), -10.68 (d, 1H, 19.5 Hz), -22.38 (t, 1H, J=19.5 and 20.1 Hz). ESI-MS for $C_{14}H_{18}N_5O_{13}P_3S$: calcd. 587.96 $[M-H]^-$, found. 587.70 $[M-H]^-$.

6-(2-Thienyl)-9-[2,3-di-*O*-(acetyl)-1-β-D-ribofuranosyl]purine (5). Compound **3** (167 mg, 0.5 mmol), which was synthesized according to Hocek's method[3], was co-evaporated with dry pyridine three times. To the residue in dry pyridine (5.0 ml) was added 178 mg (0.53 mmol) of 4,4′-dimethoxytrityl chloride. The resulting mixture was stirred for 2 h at room temperature. The reaction mixture was separated by ethyl acetate and water. The organic phase was washed with 5% $NaHCO_3$ and saturated NaCl, dried with Na_2SO_4, and evaporated *in vacuo*. The product was purified by silica gel column chromatography to give **4** (286 mg, 90%). After the evaporation of compound **4** (286 mg, 0.45 mmol) with dry pyridine three times, the residue was dissolved in pyridine (4.5 ml). To the mixture was added acetic anhydride (170 μl, 1.8 mmol). The reaction mixture was stirred at room temperature for 13 h. The solution was poured into 5% $NaHCO_3$, and then the product was extracted with ethyl acetate. After drying over Na_2SO_4, the solvent was evaporated *in vacuo*. The residue was co-evaporated with dry toluene three times. To the residue in dry methylene chloride (44.5 ml) was added dichloroacetic acid (450 μl) at 0°C. The solution was stirred for 30 min at 0°C. The mixture was poured into 5% $NaHCO_3$, and the product was extracted with methylene chloride. After drying over Na_2SO_4, the solvent was evaporated *in vacuo*. The residue was subjected to silica gel column chromatography to afford 164 mg of **5** in 87% yield. Compound **5**: ^{1}H NMR (270 MHz, DMSO-*d*6) δ 8.90 (s, 1H), 8.87 (s, 1H), 8.62 (d, 1H, J=3.6 Hz), 7.94 (d, 1H, J=4.9 Hz), 7.34 (t, 1H, J=4.0 and 4.6 Hz), 6.35 (d, 1H, J=6.6 Hz), 5.98 (t, 1H, J=5.9 Hz), 5.56 (m, 1H), 5.42 (m, 1H), 4.28 (m, 1H), 3.80-3.63 (m, 2H), 2.14 (s, 3H), 1.98 (s, 3H). ESI-MS for $C_{18}H_{18}N_4O_6S$: calcd. 419.10 $[M+H]^+$, found. 419.02 $[M+H]^+$.

6-(2-Thienyl)-9-(1-β-D-ribofuranosyl)purine 5'-triphosphate (s'TP). Compound **5** (42 mg, 0.1 mmol) was dissolved in pyridine and evaporated to dryness *in vacuo*. The residue was dissolved in pyridine (100 μl) and dioxane (300 μl). A 1 M solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one in dioxane (110 μl, 0.11 mmol) was added. After 10 min, tri-*n*-butylamine (100 μl) and 0.5 M bis(tributylammonium)pyrophosphate in DMF (300 μl, 0.15 mmol) were added to the mixture. The reaction mixture was stirred at room temperature for 10 min. A solution of 1% iodine in pyridine/water (2.0ml, 98/2, v/v) was then added. After 15 min, a 5% aqueous solution (150 μl) of $NaHSO_3$, followed by 5.0 ml of water, was added to the reaction mixture. The solution was stirred at room temperature for 30 min, and then concentrated ammonia (20 ml) was added. Ammonolysis was carried out at room temperature for 4 h. The solution was concentrated *in vacuo*, and the 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 0% to 30% CH_3CN in 100 mM triethylammonium acetate) to give the nucleoside 5'-triphosphate (**s'TP**) in 48% yield. ^{1}H NMR (270 MHz, D_2O) δ 8.72 (s, 1H), 8.62 (s, 1H), 8.25 (d, 1H, J=3.0 Hz), 7.70 (d, 1H, J=4.9 Hz), 7.21 (dd, 1H, J=3.0 and 4.9 Hz), 6.14 (d, 1H, J=5.9 Hz), 4.74 (m, 1H), 4.51 (m, 1H), 4.31 (m, 1H), 4.17 (m, 2H), 3.05 (q, 24H, J=7.3 Hz), 1.13 (t, 36H, J=7.3 Hz). ^{31}P NMR (109 MHz, D_2O) δ -9.38 (d, 1H, J=17.1 Hz), -10.56 (d, 1H, J=19.5 Hz), -22.30 (t, 1H, J=17.1 and 19.5 Hz). ESI-MS for $C_{14}H_{17}N_4O_{13}P_3S$: calcd. 572.96 $[M-H]^-$, found. 572.70 $[M-H]^-$.

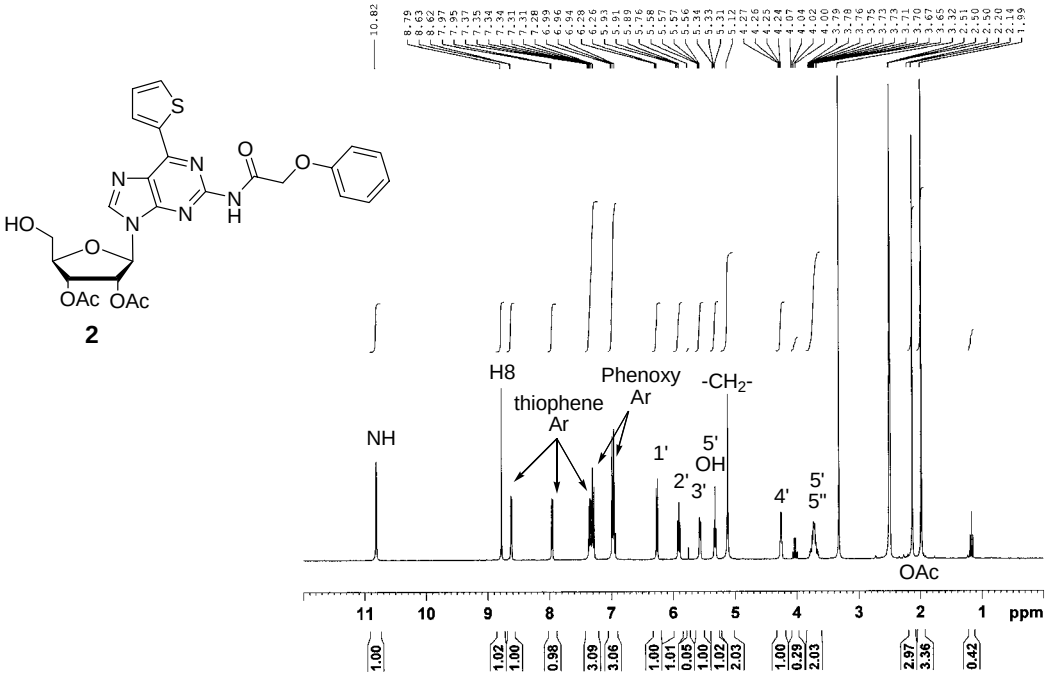
References.

(1) Fujiwara, T., Kimoto, M., Sugiyama, H., Hirao, I. and Yokoyama, S. (2001) Synthesis of 6-(2-thienyl)purine nucleoside derivatives that form unnatural base pairs with pyridin-2-one nucleosides. *Bioorg. Med. Chem. Lett.* **11**, 2221-2223.

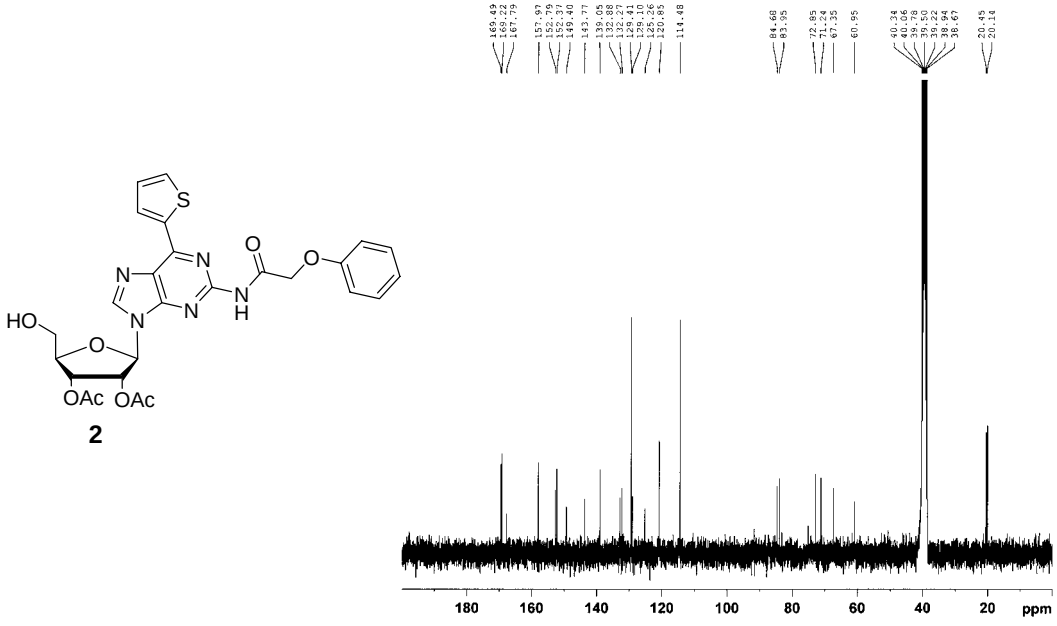
(3) Hirao, I., Ohtsuki, T., Fujiwara, T., Mitsui, T., Yokogawa, T., Okuni, H., Nakayama, K., Takio, K., Yabuki, T., Kigawa, T. *et al.* (2002) An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **20**, 177-182.

(2) Hocek, M., Holý, A., Votruba, I. and Dvořáková, H. (2001) Cytostatic 6-arylpurine nucleosides III. Synthesis and structure-activity relationship study in cytostatic activity of 6-aryl-, 6-hetaryl- and 6-benzylpurine ribonucleosides. *Collect. Czexh. Chem. Commun.* **66**, 483-499.

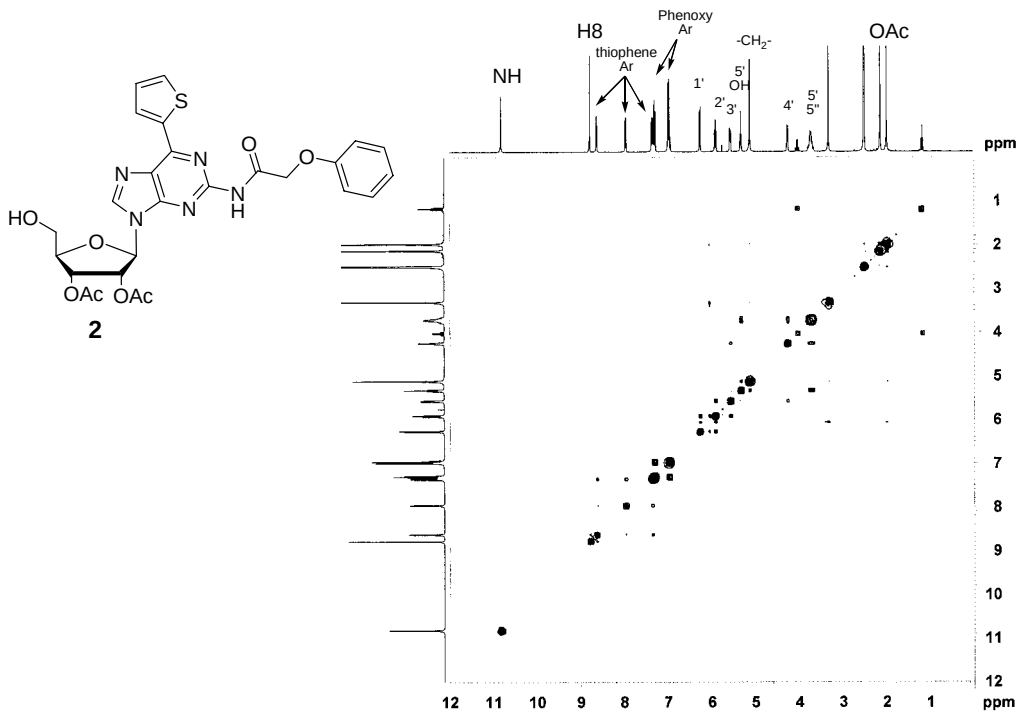
II. NMR spectra of the nucleoside derivatives of s and s'.



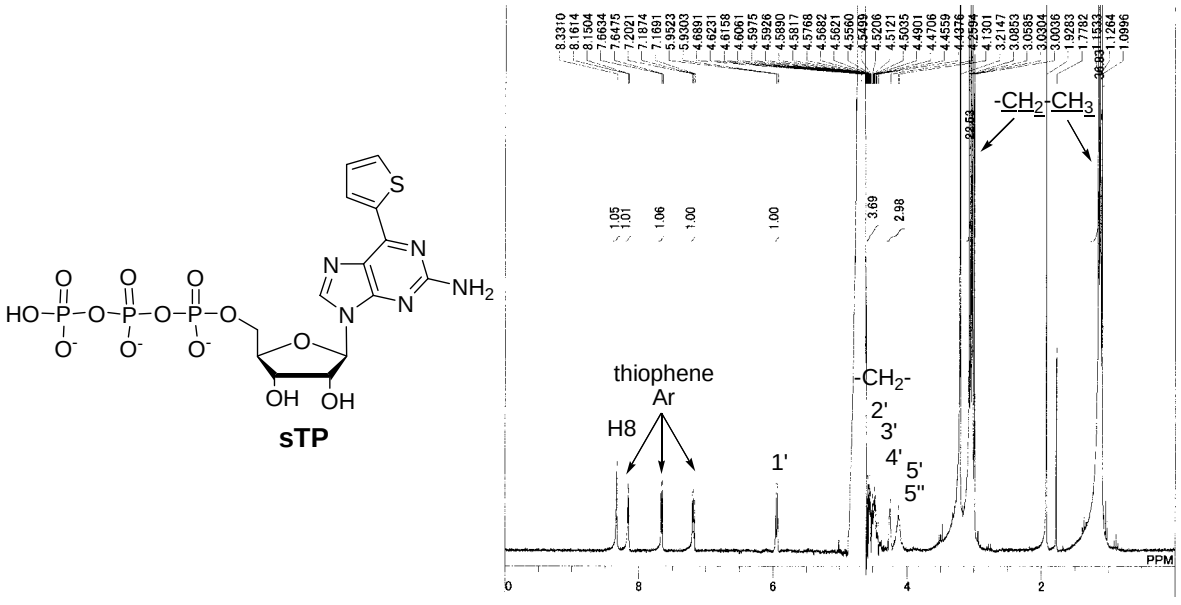
^{1}H NMR (300 MHz, DMSO-*d*6) spectrum of 2-*N*-phenoxyacetyl-6-(2-thienyl)-9-[2, 3-di-*O*-(acetyl)-1-β-D-ribofuranosyl]purine (**2**).



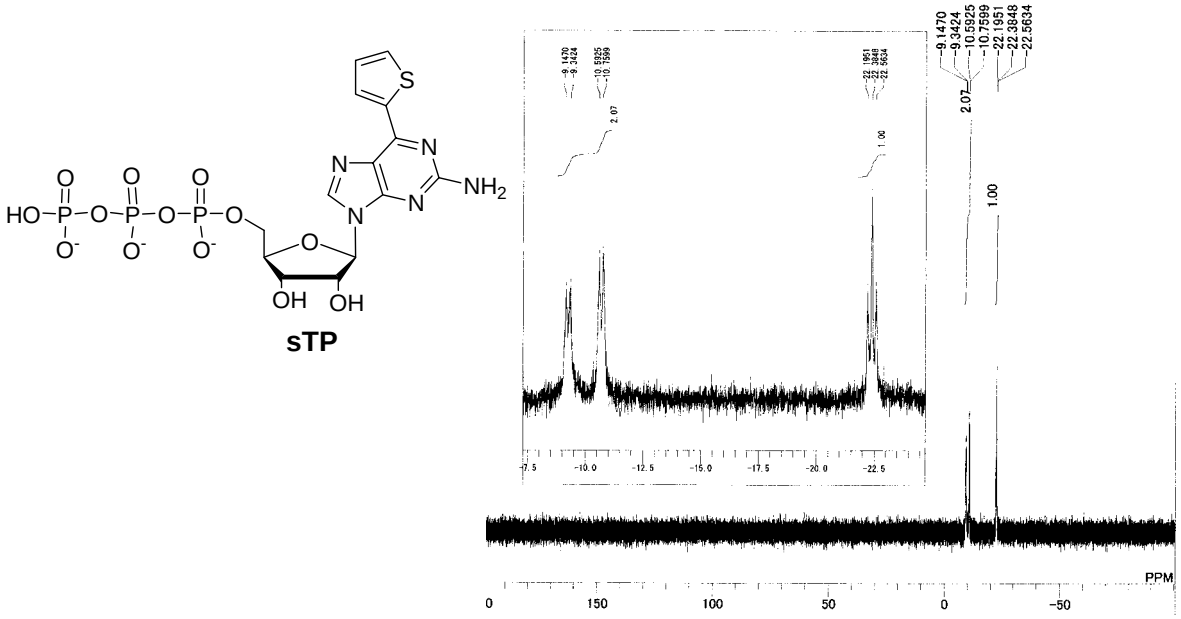
^{13}C-NMR (75 MHz, DMSO-*d*6) spectrum of 2-*N*-phenoxyacetyl-6-(2-thienyl)-9-[2, 3 di-*O*-(acetyl)-1-β-D-ribofuranosyl]purine (**2**).



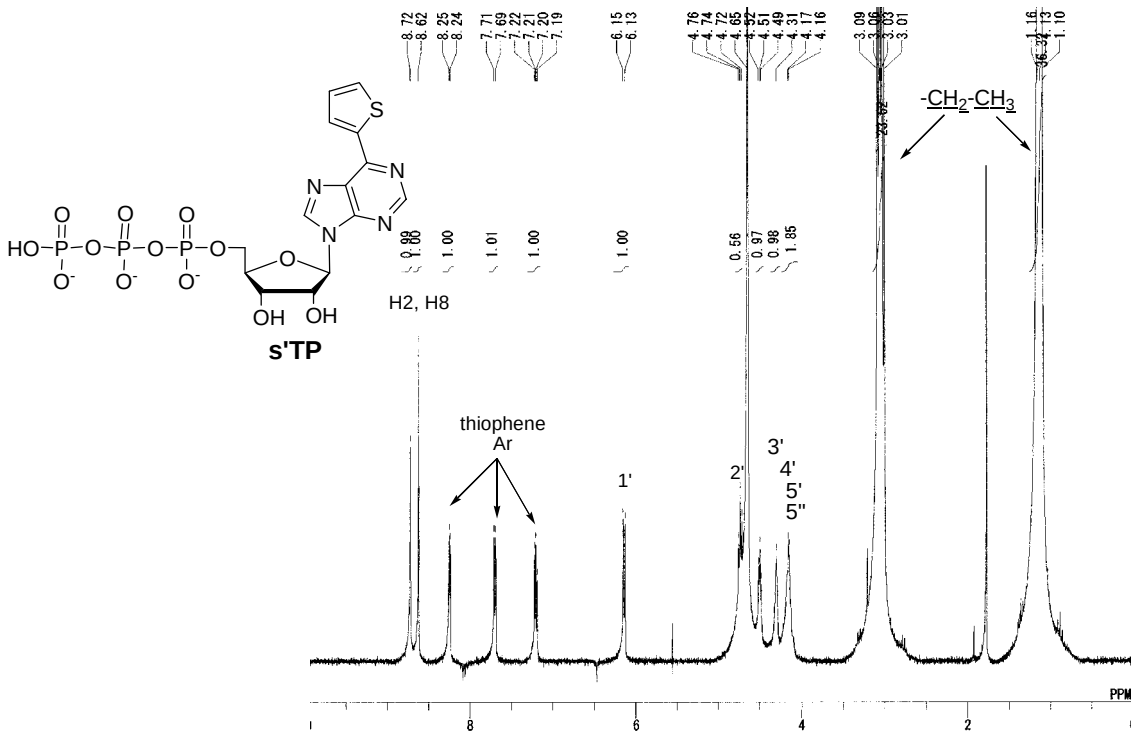
^{1}H-^{1}H COSY spectrum of 2-*N*-phenoxyacetyl-6-(2-thienyl)-9-[2, 3-di-*O*-(acetyl)-1-β-D-ribofuranosyl]purine (**2**).



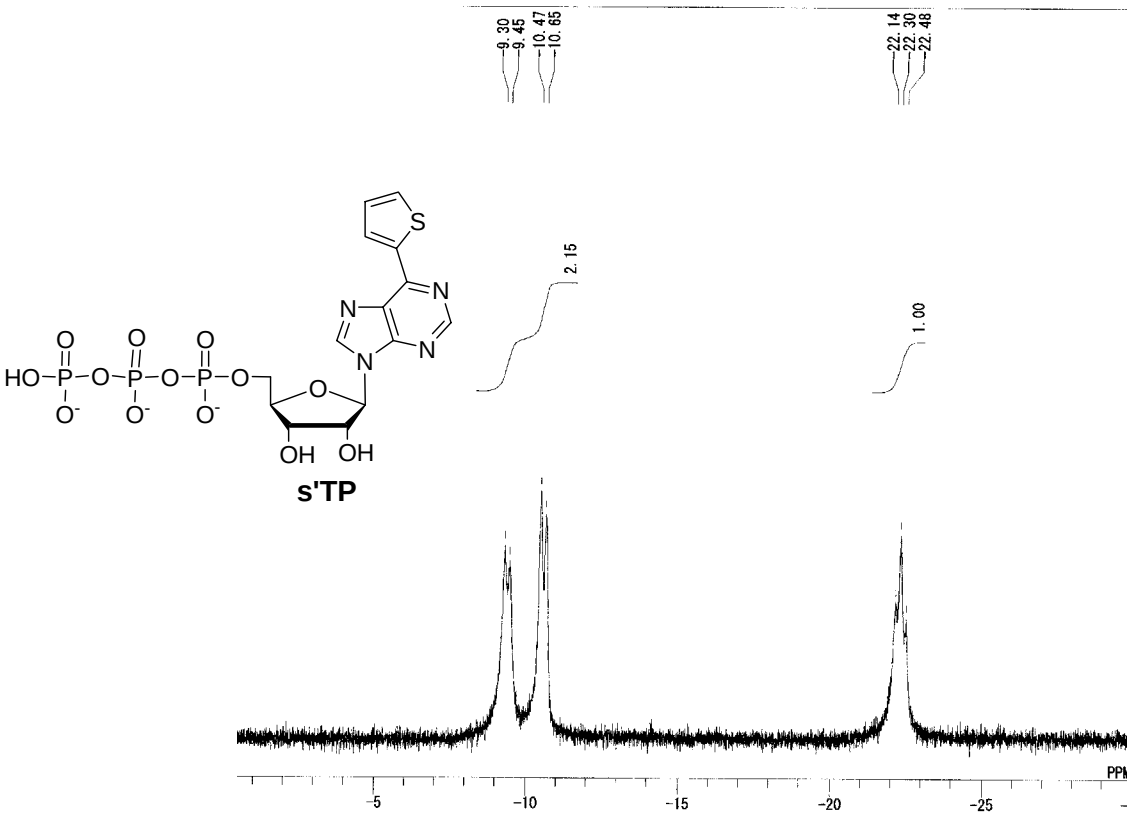
^{1}H NMR (270 MHz, D_2O) spectrum of 6-(2-thienyl)-9-(1-β-D-ribofuranosyl)purine 5′-triphosphate (**s**TP).



^{31}P NMR (109 MHz, D_2O) spectrum of 6-(2-thienyl)-9-(1-β-D-ribofuranosyl)purine 5′-triphosphate (**s**TP).



^{1}H NMR (270 MHz, D_2O) spectrum of 6-(2-thienyl)-9-(1-β-D-ribofuranosyl)purine 5'-triphosphate (**s'TP**).



^{31}P NMR (109 MHz, D_2O) spectrum of 6-(2-thienyl)-9-(1-β-D-ribofuranosyl)purine 5'-triphosphate (**s'TP**).

III. Sequences of DNA fragments for transcription

III-1. Sequences of DNA templates for s-containing hairpin RNA transcripts.

Non-template strand (23-mer) for all hairpin transcripts
5'-ATAATACGACTCACTATAGGGAG

Template strand (37-mer) for hairpin s10
5'-TCAGCGCTT**Pa**CGCACTCCCTATAGTGAGTCGTATTAT

Template strand (37-mer) for hairpin s11
5'-TCAGCGCT**Pa**TCGCACTCCCTATAGTGAGTCGTATTAT

III-2. Sequences of DNA templates for s-containing tRNA transcripts.

Non-template strand (94-mer) for all tRNA transcripts
5'-ATAATACGACTCACTATAGGGGATTTAGCTCAGTTGGGAGAGCGCCAGACTGAAGA
TCTGGAGGTCCTGTGTTCGATCCACAGAATTCCCACCA

Template strand (94-mer) for the original tRNA transcript (Tm=2'-OMe-T, Gm=2'-OMe-G)
5'-TmGmGTGGGAATTCTGTGGATCGAACACAGGACCTCCAGATCTTCAGTCTGGCGCTC
TCCCAACTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

Template strand (94-mer) for tRNA 16s (Tm=2'-OMe-T, Gm=2'-OMe-G)
5'-TmGmGTGGGAATTCTGTGGATCGAACACAGGACCTCCAGATCTTCAGTCTGGCGCTC
TCCCA**Pa**CTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

Template strand (94-mer) for tRNA 17s (Tm=2'-OMe-T, Gm=2'-OMe-G)
5'-TmGmGTGGGAATTCTGTGGATCGAACACAGGACCTCCAGATCTTCAGTCTGGCGCTC
TCCC**Pa**ACTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

Template strand (94-mer) for tRNA 36s (Tm=2'-OMe-T, Gm=2'-OMe-G)
5'-TmGmGTGGGAATTCTGTGGATCGAACACAGGACCTCCAGATC**Pa**TCAGTCTGGCGCT
CTCCCAACTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

Template strand (94-mer) for tRNA 47s (Tm=2'-OMe-T, Gm=2'-OMe-G)
5'-TmGmGTGGGAATTCTGTGGATCGAACACAGG**Pa**CCTCCAGATCTTCAGTCTGGCGCT
CTCCCAACTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

Template strand (94-mer) for tRNA 57s (Tm=2'-OMe-T, Gm=2'-OMe-G)

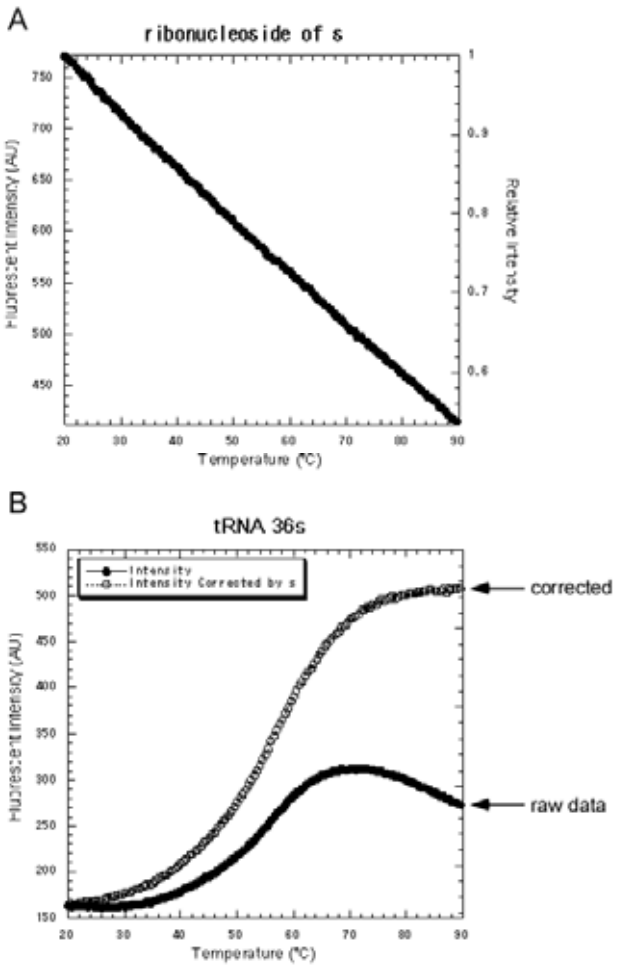
5'-TmGmGTGGGAATTCTGTGGAT**Pa**GAACACAGGACCTCCAGATCTTCAGTCTGGCGCT
CTCCCAACTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

Template strand (94-mer) for tRNA 59s (Tm=2'-OMe-T, Gm=2'-OMe-G)
5'-TmGmGTGGGAATTCTGTGG**Pa**TCGAACACAGGACCTCCAGATCTTCAGTCTGGCGCT
CTCCCAACTGAGCTAAATCCCCTATAGTGAGTCGTATTAT

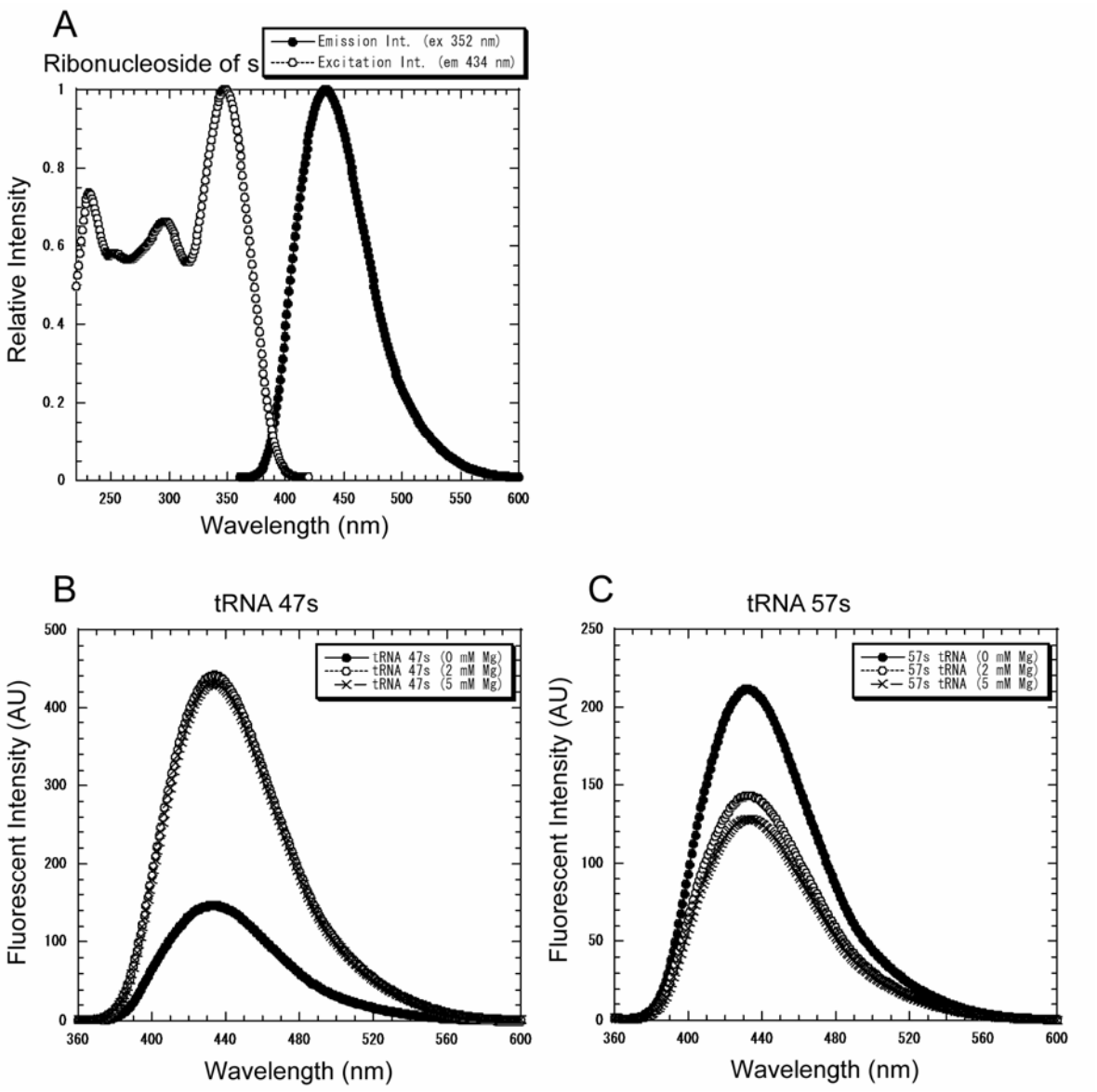
IV. Correction of the fluorescent profiles of RNA molecules containing s

Temperature dependency of the fluorescent intensity of the **s** ribonucleoside (A) and correction of the fluorescent intensity of tRNA 36**s** by that of the **s** ribonucleoside (B).

Fluorescent intensities of tRNA transcripts were corrected by using the equation, *Yct* = *Yt* / (*Rt* / *R20*), where *Yct* is the corrected intensity of each tRNA at t°C, *Yt* is the measured fluorescent intensity of tRNA at t°C, *Rt* is the observed fluorescent intensity of the ribonucleoside of **s** at t°C, and *R20* is the observed fluorescent intensity of the **s** ribonucleoside at 20°C.



V. Excitation and emission scans of ribonucleoside s and s in the tRNA molecules



Excitation and emission scans of the **s** ribonucleoside (1 μM) (A) and emission scans of tRNA 47s (1 μM) (B) and tRNA 57s (1 μM) (C) with excitation at 352 nm in 50 mM sodium cacodylate (pH 7.2), 50 mM KCl, and 0.1 mM EDTA at 20°C.