Supporting information

Structure-dependent binding of hnRNP A1 to telomere RNA

Xiao Liu[1], Takumi Ishizuka[1], Hong-Liang Bao[1], Kei Wada[2], Yuuma Takada[1], Keisuke Iida[3], Kazuo Nagasawa[3], Danzhou Yang[4], and Yan Xu[1]*

[1]Division of Chemistry, Department of Medical Sciences, Faculty of Medicine, University of Miyazaki, 5200 Kihara, Kiyotake, Miyazaki 889-1692, Japan

[2]Organization for Promotion of Tenure Track, University of Miyazaki, Japan

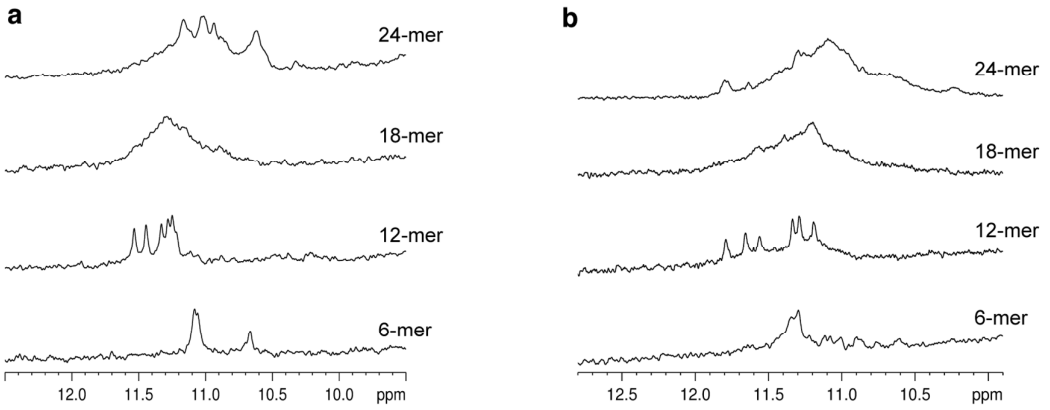
[3] Department of Biotechnology and Life Science Tokyo University of Agriculture and Technology 2-24-16, Naka-cho Koganei city, Tokyo, 184-8588 Japan

[4]Department of Medicinal Chemistry and Molecular Pharmacology College of Pharmacy Purdue University 201 S. University St. West Lafayette, IN 47907, United States

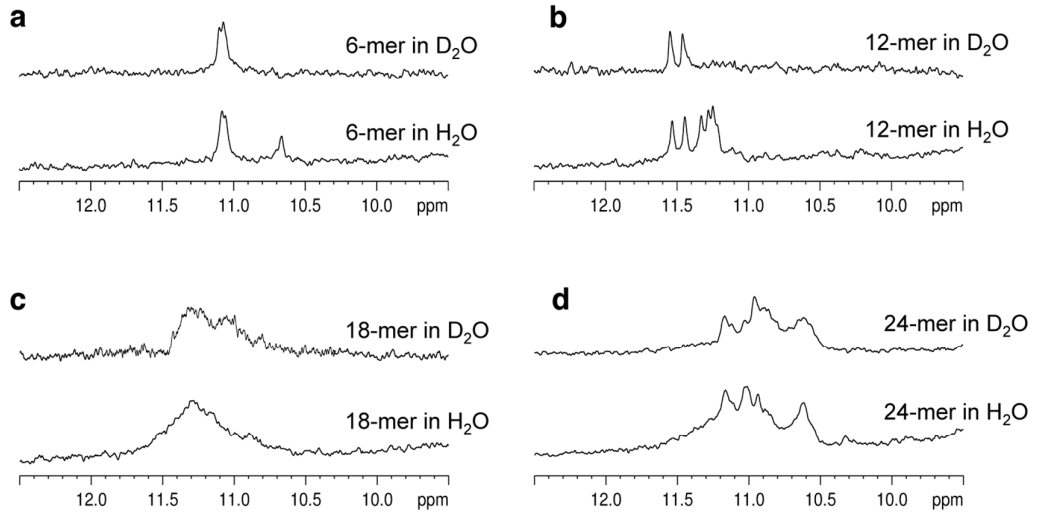
Email: xuyan@med.miyazaki-u.ac.jp

*Corresponding author.

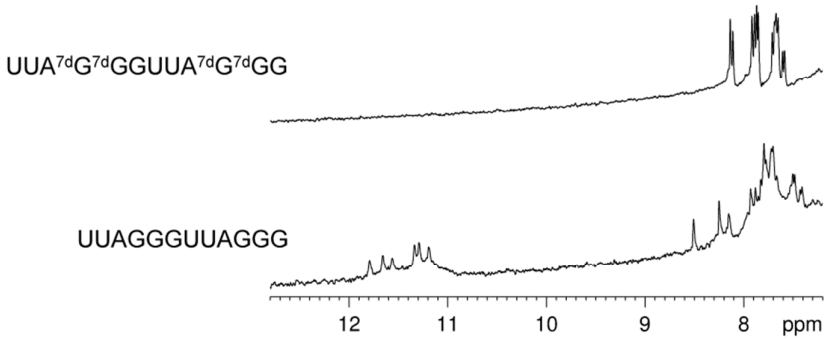
This Supplementary Information section includes:



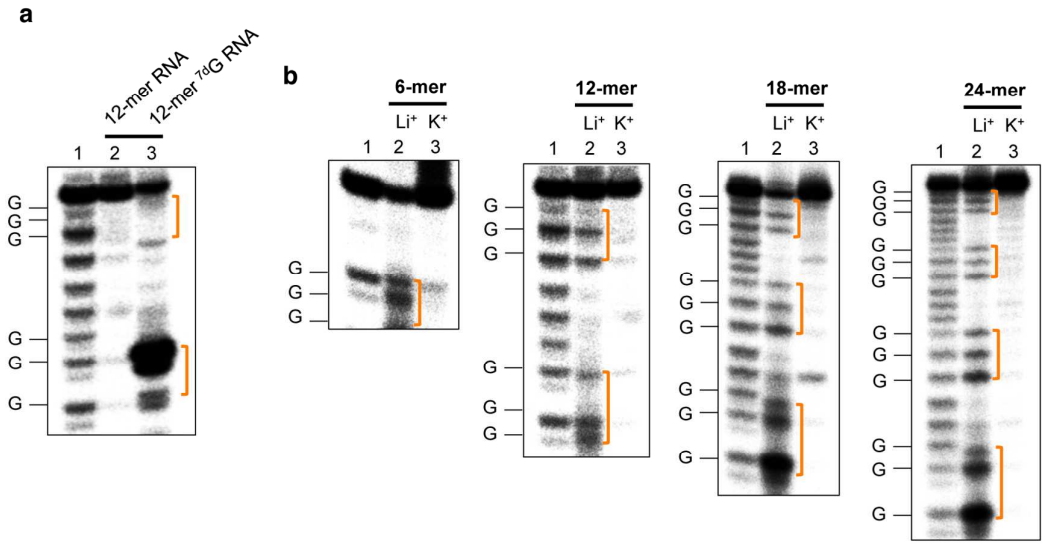
Supplementary Figure S1. Imino proton NMR spectra of 6-mer UUAGGG, 12-mer UUAGGGUUAGGG, 18-mer UUAGGGUUAGGGUUAGGG, and 24-mer UUAGGGUUAGGGUUAGGGUUAGGG RNA in the presence of 150 mM KCl and K-phosphate buffer (a), in the presence of 150 mM NaCl and Na-phosphate buffer (b).



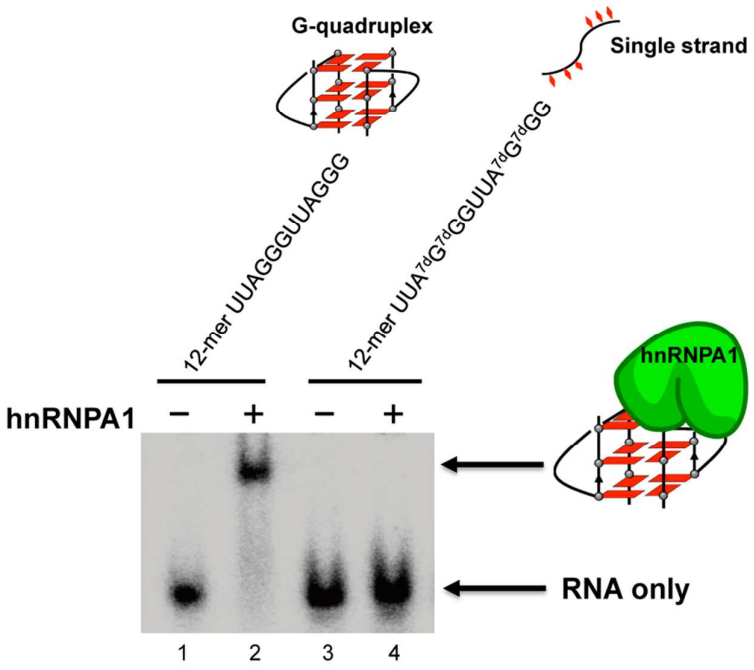
Supplementary Figure S2. Imino proton NMR spectra recorded in H_2O and after hydrogen-deuterium exchange in D_2O for 6-mer UUAGGG (a), 12-mer UUAGGGUUAGGG (b), 18-mer UUAGGGUUAGGGUUAGGG (c), and 24-mer UUAGGGUUAGGGUUAGGGUUAGGG (d) RNA in the presence of 150 mM KCl and K-phosphate buffer.



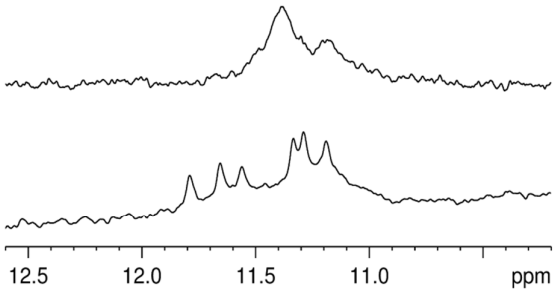
Supplementary Figure S3. Proton spectrum (7.0 – 12 ppm) of 12-mer UUAGGGUUAGGG RNA (bottom) and 7-deazaguanosine-substituted RNA 12-mer UUA^{7d}G^{7d}GGUUA^{7d}G^{7d}GG (top) in the presence of 150 mM NaCl and Na-phosphate buffer. ^{7d}G indicated 7-deazaguanosine.



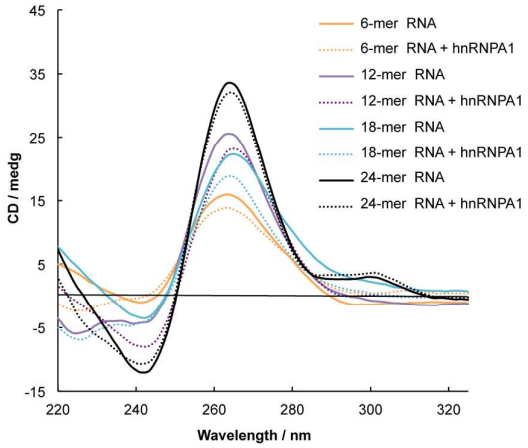
Supplementary Figure S4. RNase T1 footprinting assay for analysis of G-quadruplex formation. (a) ^{32}P-labeled 12-mer UUAGGGUUAGGG RNA (lane 2) and 7-deazaguanosine-substituted 12-mer UUA^{7d}G^{7d}GGUUA^{7d}G^{7d}GG RNA (lane 3) were digested with RNase T1 in K$^+$-containing binding buffer. Base hydrolysis ladders (lane 1) (b) ^{32}P-labeled 6-mer UUAGGG, 12-mer UUAGGGUUAGGG, 18-mer UUAGGGUUAGGGUUAGGG, and 24-mer UUAGGGUUAGGGUUAGGGUUAGGG RNA were digested with RNase T1 in either K$^+$- or Li$^+$-containing binding buffer (lane 2 and 3). Base hydrolysis ladders (lane 1). G residues that involved in G-quadruplex are indicated by brackets. The position of every G is indicated by bars.



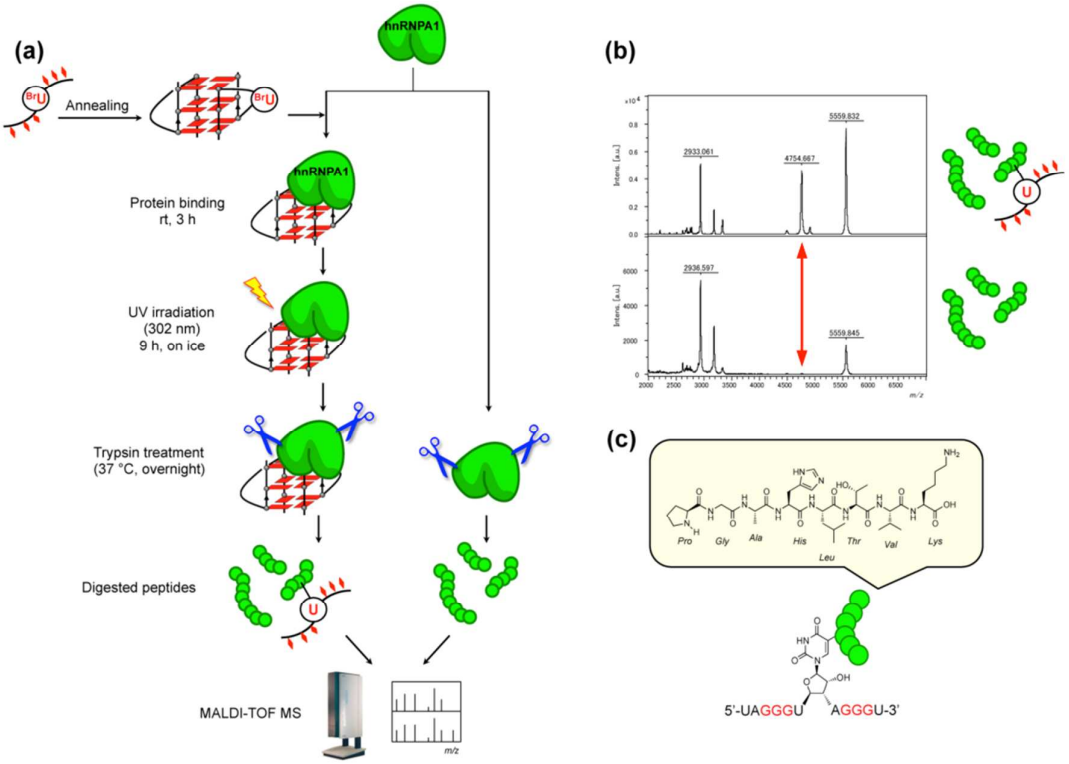
Supplementary Figure S5. Telomere RNA G-quadruplex binding selectivity of hnRNPA1. Experiment was performed with ^{32}P-labeled 12-mer UUAGGGUUAGGG RNA (lane 1, 2) and 7-deazaguanosine-substituted 12-mer UUA^{7d}G^{7d}GGUUA^{7d}G^{7d}GG RNA (lane 3, 4) in the presence or absence of hnRNPA1. Using 6% PAGE in 1×TBE buffer with 100 mM NaCl in 4 °C for 2 h (80V). ^{7d}G indicated 7-deazaguanosine.



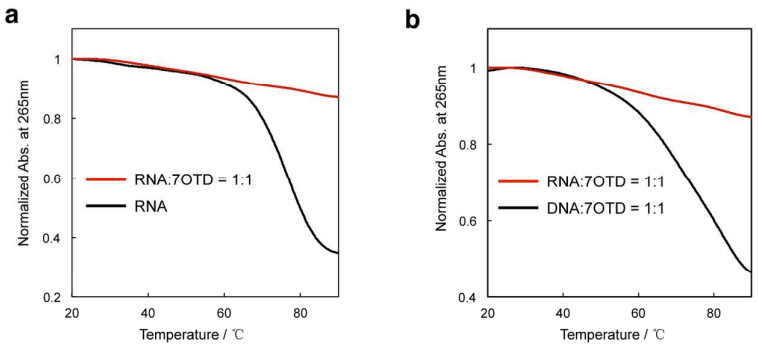
Supplementary Figure S6. Imino proton spectra of 12-mer UUAGGGUUAGGG RNA free (Bottom) and in the presence of hnRNPA1 (Top) (at the 1:1 RNA/protein ratio).



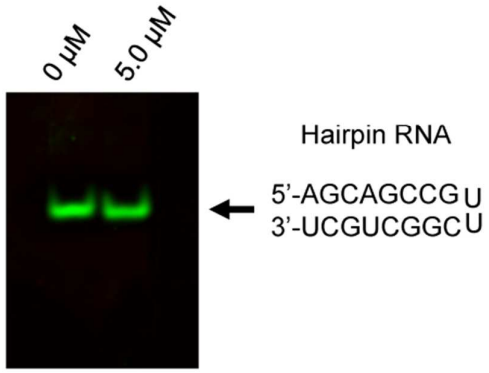
Supplementary Figure S7. CD spectra of 6-mer UUAGGG RNA, 12-mer UAGGG UUAGGGU RNA, 18-mer UUAGGGUUAGGG RNA, and 24-mer UUAGGGUUAGGGUUAGGGUUAGGG RNA with (dot-line) and without (solid line) hnRNPA1 in 100 mM KCl, 10 mM Tris-HCl buffer (PH 7.0). The 5 μM hnRNPA1 was added to 5 μM telomere RNAs and incubated 1h at room temperature before spectra were obtained.



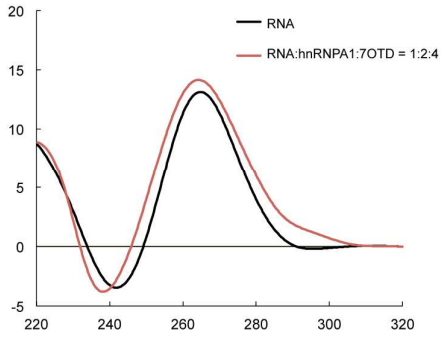
Supplementary Figure S8. (a) Schematic overview of MALDI-TOF mass spectral characterization of hnRNAPA1 and telomere RNA G-quadruplex containing a 5-bromouracil at loop. The crosslinked RNA-protein complex (binding with RNA G-quadruplex) and free hnRNPA1 (without RNA G-quadruplex) were enzymatically digested with trypsin and then spectral characterization by MALDI-TOF mass. (b) MALDI-TOF MS spectra of the crosslinked RNA-protein complex (upper) or free hnRNPA1 (bottom) after enzymatically digested. Red arrow indicates a peak difference between the mass spectra. (c) Digested product derived from telomere RNA G-quadruplex-hnRNPA1 photocrosslinking treated with trypsin.



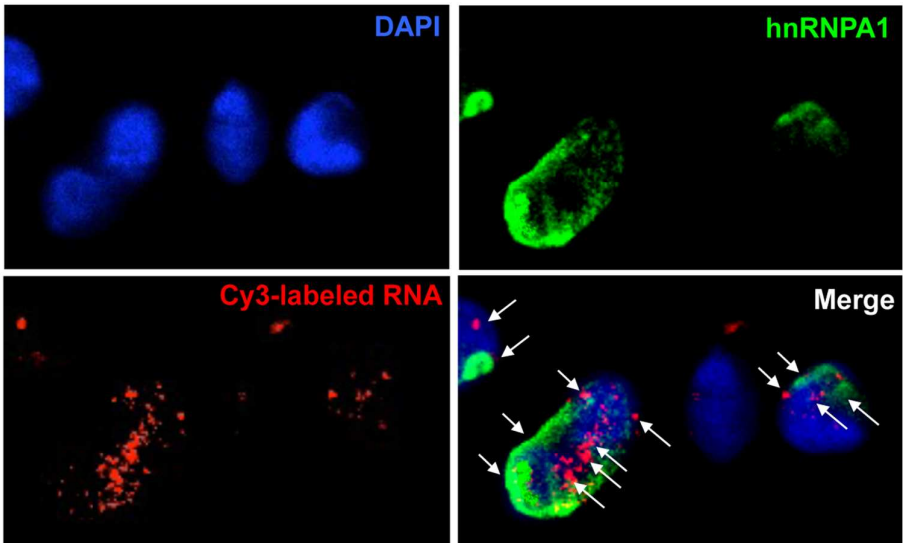
Supplementary Figure S9. (a) CD melting curves of 12-mer UAGGGUUAGGGU RNA (5 μM) with (red line) and without 7OTD (5 μM) (back line) in 10 mM Tris-HCl buffer pH 7.0, 50 mM KCl. (b) CD melting curves of 12-mer UAGGGUUAGGGU RNA (5 μM) (red line) and 12-mer TAGGGTTAGGGT DNA (5 μM) (back line) in 10 mM Tris-HCl buffer pH 7.0, 50 mM KCl.



Supplementary Figure S10. Visualization of hairpin RNA AGCAGCCGUUCCGGCUGCU with Cy5-7OTD. RNA bands were stained with GelStar^{TM} (green). Concentrations of Cy5-7OTD show in upper of gel. Using 20% PAGE in 1×TBE buffer with 20 mM NaCl in 4 °C for 8 h (80V). RNA concentration, 5 μM.



Supplementary Figure S11. CD spectra of 12-mer UAGGGUUAGGGU RNA (back line) and the complex of RNA (5 μM), hnRNPA1 (10 μM) and 7OTD (20 μM) (red line) in 10 mM Tris-HCl buffer (pH 7.0), 100 mM NaCl, 0.1 mg/mL BSA, 5 mM dithiothreitol, and 10% (v/v) glycerol. Concentration ratio shows on red line.

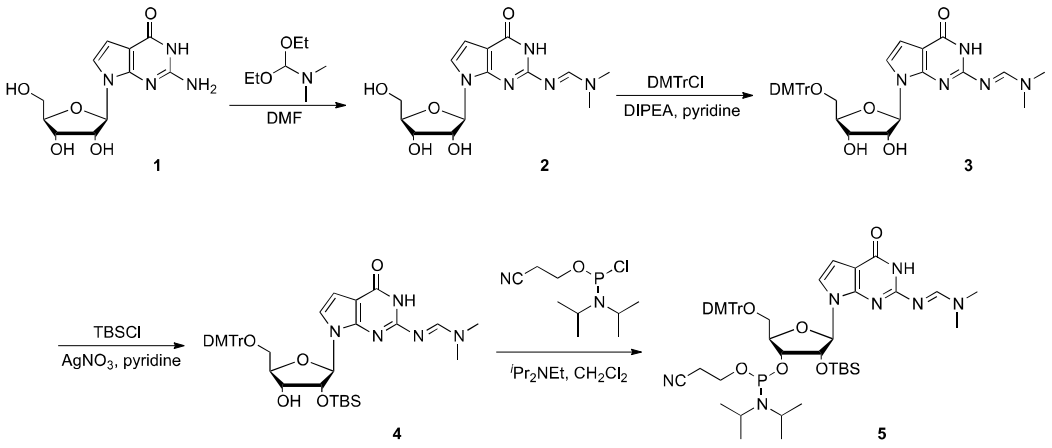


Supplementary Figure S12. Fluorescence microscopy imaging of Hela cells. Cy3-labeled 7-deazaguanosine-containing probe RNA 5'-UUA^{7d}G^{7d}GGUUA^{7d}G^{7d}G)G-Cy3-3' (red) was transfected into living cells. hnRNPA1 was immunostained using antibodies anti human hnRNPA1 (green). Blue fluorescence foci derived from DNA stained by DAPI. There is no colocalization of fluorescent signals (merged panel). The allows indicated the single-strand probe RNA (red) and hnRNPA1 (green).

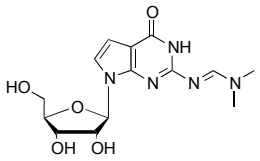
Synthesis of 7-deazaguanosine-containing RNA by phosphoramidite chemistry

General

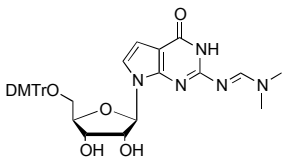
^{1}H-NMR and ^{31}P-NMR spectra were recorded on a BRUKER (AV-400M) magnetic resonance spectrometer. DMSO-d_6 and $CDCl_3$ were used as the solvents. ^{1}H spectra chemical shifts (δ) are reported in parts per million (ppm) referenced to residual protonated solvent peak (DMSO-d_6, δ = 2.50, $CDCl_3$, δ = 7.26). Coupling constants (J) values are given in hertz (Hz). Signal patterns are indicated as br (broad), s (singlet), d (doublet), t (triplet), sept (septet), m (multiplet). All reagents were purchased from Aldrich, TCI (Tokyo Chemical Industry) or Wako (Wako Pure Chemical Industries). Thin layer chromatography (TLC) was performed using TLC Silica gel 60 F_{254} (Merck). Compounds were visualized using a UV lamp (254 nm) or staining with a potassium permanganate solution. Silica gel (Wakogel® C-300, 200-325 mesh) was used for column chromatography. Purification of products was also performed on a middle pressure liquid chromatography (MPLC) systems (EPCLC-AI-580S, Yamazen Corporation) equipped with silica gel column (Hi-Flash Column, Yamazen Corporation). High-resolution mass spectra (HRMS) were recorded by electrospray ionization (ESI) on an Exactive Orbitrap mass spectrometer instrument (Thermo Scientific). 7-Deazaguanosine **1** was purchased from ChemGenes (#RP-2313).



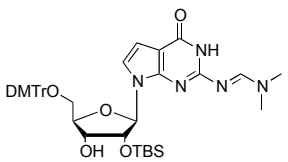
Scheme S1. Synthetic scheme of 7-deazaguanosine (^{7d}G) phosphoramadite **5**



***N*^{2}-Dimethylformamidyl-7-deazaguanosine (2)** 7-Deazaguanosine **1** (500 mg, 1.77 mmol) was co-evaporated with anhydrous DMF (10 mL) three times, followed by suspension in anhydrous DMF (5 mL), and *N*,*N*-dimethylformamide diethyl acetal (1.8 mL, 10.6 mmol) was added. The mixture was stirred for 1 h at 50 °C under argon, then the solvent was evaporated. Pale-yellow crystals of **2** (508 mg, 85 %) were obtained from MeOH. ^{1}H-NMR (400 MHz, DMSO-d_6) δ 11.02 (s, 1H), 8.52 (s, 1H), 7.07 (d, *J* = 3.6 Hz, 1H), 6.34 (d, *J* = 3.6 Hz, 1H), 5.96 (d, *J* = 6.1 Hz, 1H), 5.25 (br s, 1H), 5.10 (br s, 1H), 4.96 (t, *J* = 5.3 Hz, 1H), 4.30 (t, *J* = 5.3 Hz, 1H), 4.06 (m, 1H), 3.84 (q, *J* = 3.6 Hz, 1H), 3.55 (m, 2H), 3.15 (s, 3H), 3.02 (s, 3H).

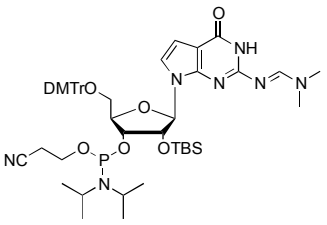


5'-*O*-Dimethoxytrityl-*N*^{2}-dimethylformamidyl-7-deazaguanosine (3) To a compound **2** (500 mg, 1.48 mmol) dried three times by co-evaporation of anhydrous pyridine (20 ml) and dissolved in anhydrous pyridine (15 ml) was added *N*,*N*-diisopropylethylamine (1.0 ml, 5.93 mmol) and 4,4'-dimethoxytritylchloride (753 mg, 2.22 mmol). The mixture was stirred for 2 h at 50 °C under argon, then the added 5% $NaHCO_3$ aqueous solution. The mixture was extracted three times with CH_2Cl_2. The organic layer was dried over Na_2SO_4, and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (CH_2Cl_2:MeOH=9:1) to give the compound **3** (743 mg, 78%) as a solid. ^{1}H-NMR (400 MHz, DMSO-d_6) δ 11.05 (s, 1H), 8.54 (s, 1H), 7.39-7.19 (m, 9H), 6.93 (d, *J* = 3.6 Hz, 1H), 6.86-6.84 (m, 4H), 6.34 (d, *J* = 3.6 Hz, 1H), 6.02 (d, *J* = 5.0 Hz, 1H), 5.39 (d, *J* = 5.5 Hz, 1H), 5.14 (d, *J* = 5.5 Hz, 1H), 4.29 (q, *J* = 5.3 Hz, 1H), 4.12 (q, *J* = 5.1 Hz, 1H), 3.96 (q, *J* = 3.3 Hz, 1H), 3.73 (s, 6H), 3.20-3.16 (m, 2H), 3.13 (s, 3H), 3.02 (s, 3H). HRMS (ESI) for $C_{35}H_{36}O_7N_5$ $[M-H]^-$: Calcd. 638.2609; Found. 638.2619.

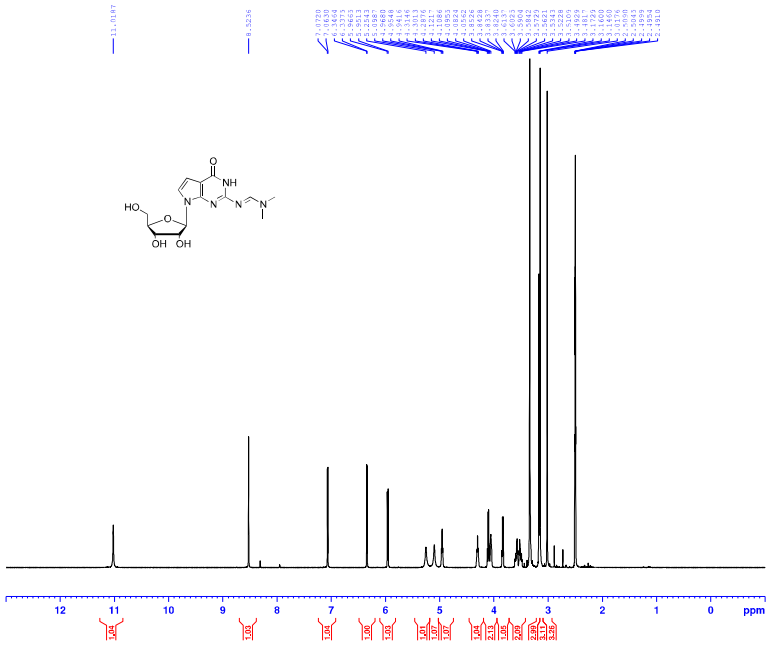


2'-*O*-*tert*-Butyldimethylsilyl-5'-*O*-dimethoxytrityl-*N*^{2}-dimethylformamidyl-7-deazaguanosine (4) To a solution of **3** (740 mg, 1.16 mmol) in anhydrous pyridine (5 ml), $AgNO_3$ (197 mg, 1.16 mmol) was added under stirring at room temperature. Stirring was continued for 5 min, and a solution of *tert*-butyldimethylchloro

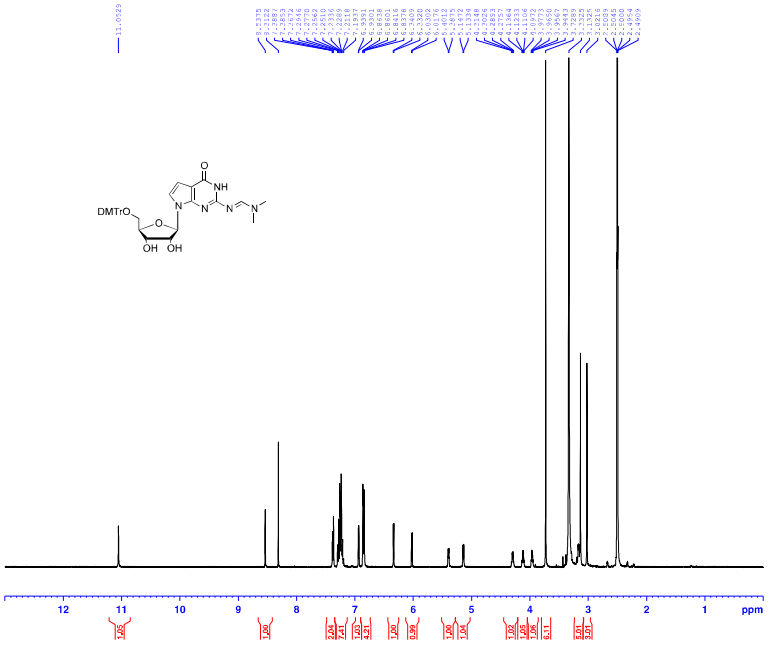
silane (175 mg, 1.16 mmol) in anhydrous THF (10 ml) was introduced under exclusion of light and moisture. A second portion of $AgNO_3$, (99 mg, 0.58 mmol) and *tert*-butyldimethylchlorosilane (88 mg, 0.58 mmol) was added after 10 h. Stirring was continued for another 14 h, AgCl was filtered off, the filtrate treated with 5% $NaHCO_3$ aqueous solution (20 ml), the aqueous layer extracted twice with CH_2Cl_2, and the combined organic layer dried over Na_2SO_4, and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (AcOEt only) to give the compound **4** (258 mg, 30%) as a solid. ^{1}H-NMR (400 MHz, DMSO-d_6), 11.07 (s, 1H), 8.51 (s, 1H), 7.42-7.40 (m, 2H), 7.31-7.23 (m, 7H), 6.98 (d, J = 3.6 Hz, 1H), 6.88-6.86 (m, 4H), 6.33 (d, J = 3.6 Hz, 1H), 6.08 (d, J = 4.8 Hz, 1H), 5.03 (d, J = 5.6 Hz, 1H), 4.34 (t, J = 4.8 Hz, 1H), 4.09 (q, J = 5.2 Hz, 1H), 4.02 (m, 1H), 3.73 (s, 6H), 3.23 (m, 2H), 3.11 (s, 3H), 3.02 (s, 3H), 0.78 (s, 9H), −0.03 (s, 3H), −0.12 (s, 3H). HRMS (ESI) for $C_{41}H_{50}O_7N_5Si$ $[M-H]^-$: Calcd. 752.3474; Found. 752.3488.



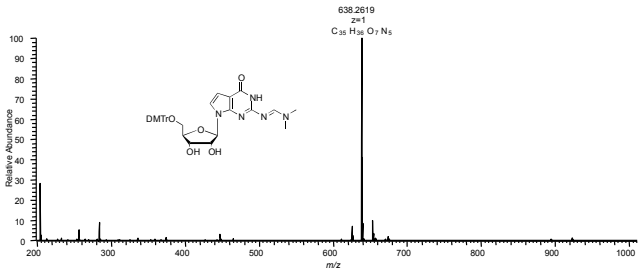
3'-*O*-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-2'-*O*-*tert*-butyldimethylsilyl-5'-*O*-dimethoxytrityl-N^2-dimethylformamidyl-7-deazaguanosine (5) The compound **4** (250 mg, 0.33 mmol) was treated with dry *N*,*N*-diisopropylethylamine (231 µl, 1.33 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (222 µl, 1.00 mmol) in anhydrous acetonitrile (10 ml) and anhydrous CH_2Cl_2 (4 ml). The mixture was stirred for 2 h at room temperature under argon. After addition of CH_2Cl_2 (50 ml), the reaction was stopped by adding a 5% $NaHCO_3$ aqueous solution (50 ml). The aqueous layer was extracted three times with CH_2Cl_2. The combined organic layers were dried over Na_2SO_4 and the solvent was evaporated *in vacuo*. The residue was purified by silica gel column chromatography (AcOEt only) to give the compound **5** (241 mg, 76%) as a solid. ^{1}H-NMR (400 MHz, $CDCl_3$), 8.64 (s, 1H), 8.48 (br s, 1H), 7.48 (m, 2H), 7.38-7.21 (m, 7H), 7.04 (d, J = 3.6 Hz, 1H), 6.84-6.80 (m, 5H), 6.58 (m, 1H), 6.26 (m, 1H), 4.61 (m, 1H), 4.34-3.86 (m, 5H), 3.78 (s, 6H), 3.66-3.44 (m, 4H), 3.28-3.18 (m, 1H), 3.15-3.08 (s, 6H), 2.78-2.59 (m, 1H), 2.35-2.24 (m, 1H), 1.29-1.16 (s, 9H), 1.00 (d, J = 6.7 Hz, 3H), 0.83-0.78 (s, 12H). ^{31}P-NMR (161 MHz, $CDCl_3$) δ 150.72, 149.43. HRMS (ESI) for $C_{50}H_{69}O_8N_7PSi$ $[M+H]^+$: Calcd. 954.4709; Found. 954.4694.



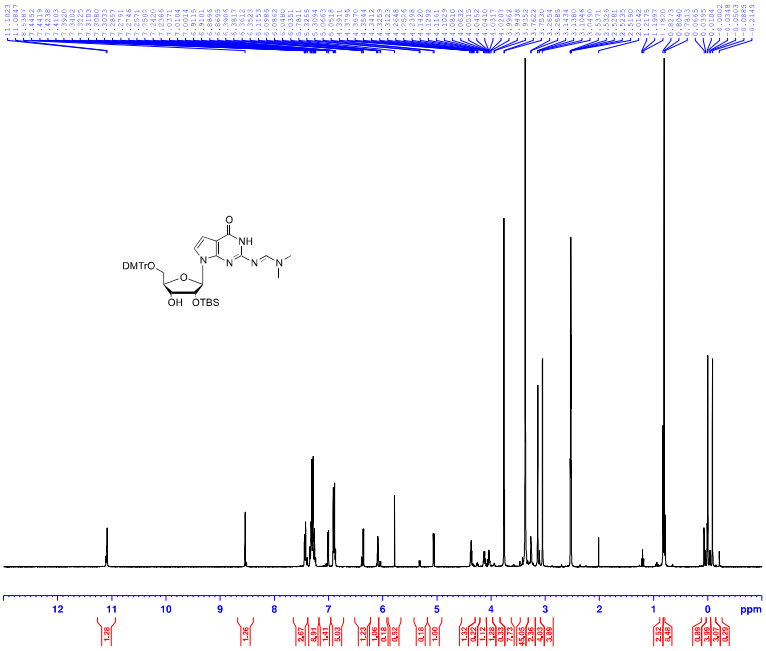
Supplementary Figure S13. ^{1}H NMR spectrum of compound **2**



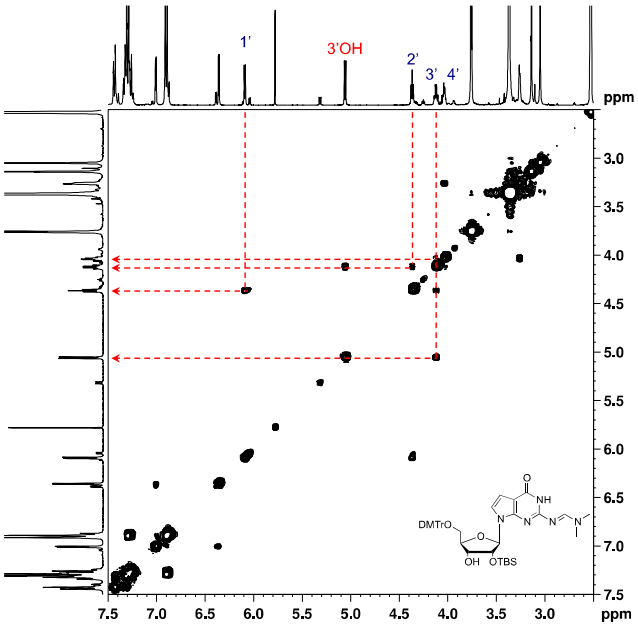
Supplementary Figure S14. ^{1}H NMR spectrum of compound **3**



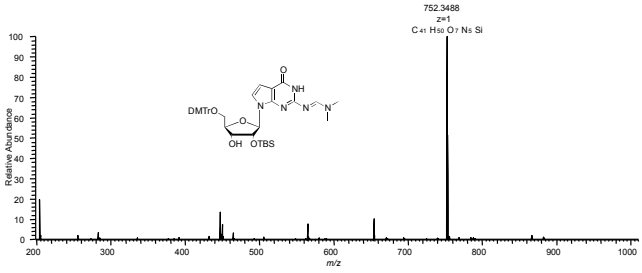
Supplementary Figure S15. HRMS spectrum of compound **3**



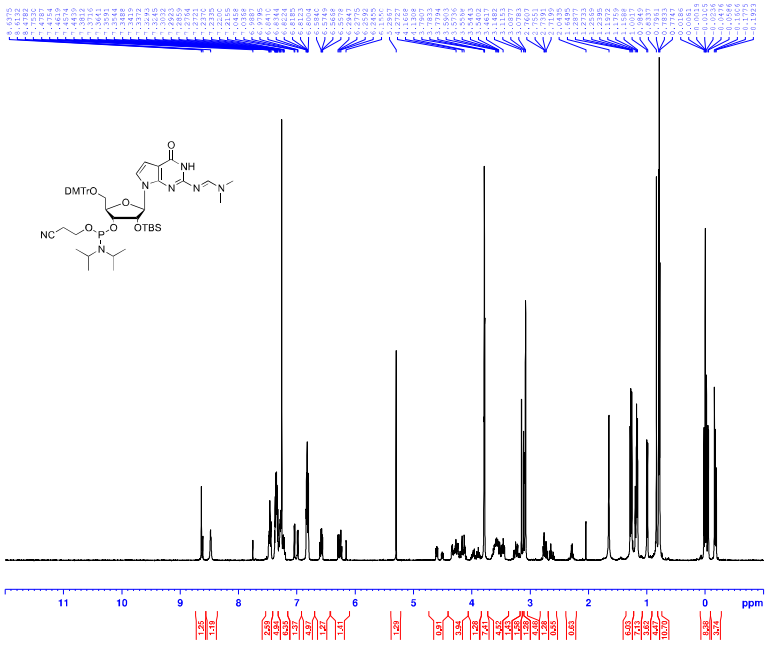
Supplementary Figure S16. ^{1}H NMR spectrum of compound **4**



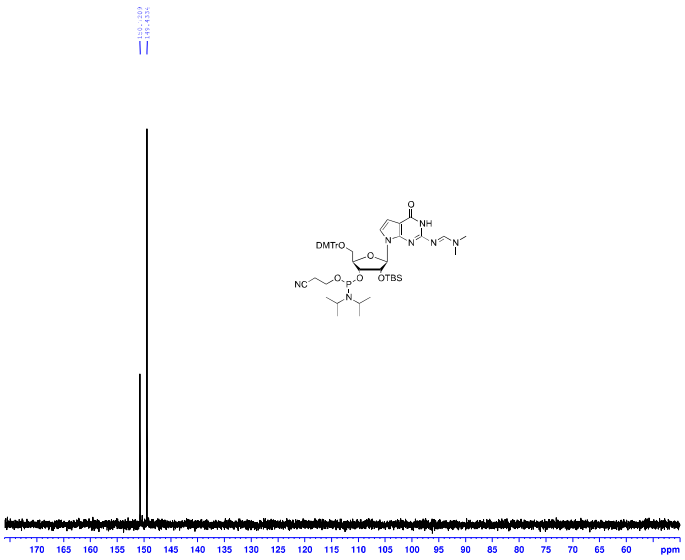
Supplementary Figure S17. COSY spectrum of compound **4**



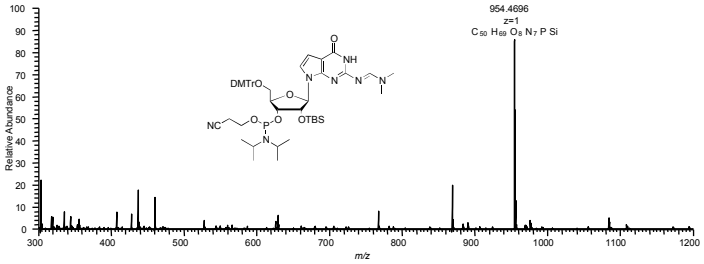
Supplementary Figure S18. HRMS spectrum of compound **4**



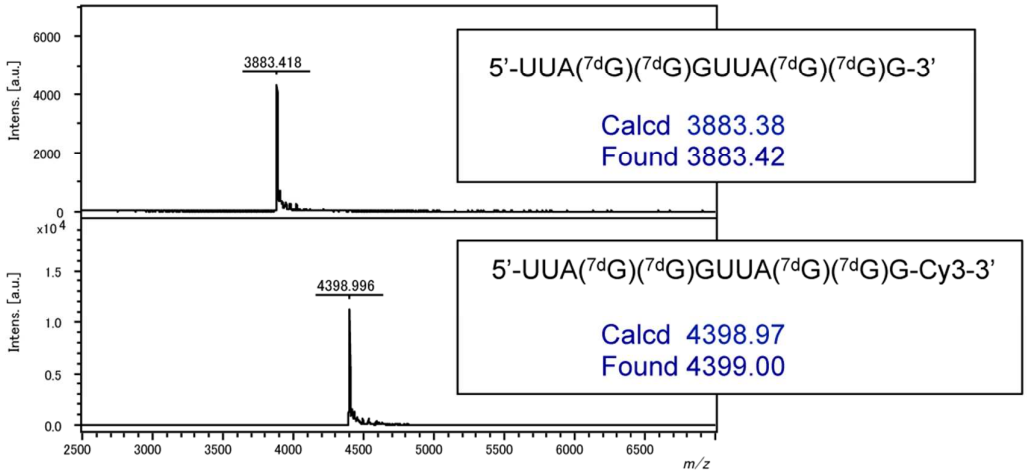
Supplementary Figure S19. ^{1}H NMR spectrum of compound **5**



Supplementary Figure S20. ^{31}P NMR spectrum of compound **5**



Supplementary Figure S21. HRMS spectrum of compound **5**



Supplementary Figure S22. MALDI-TOF MS data of 7-deazaguanosine-substituted RNA 5'-UUA(^{7d}G)(^{7d}G)GUUA(^{7d}G)(^{7d}G)G-3' and 5'-5'-UUA(^{7d}G)(^{7d}G)GUUA(^{7d}G)(^{7d}G)G-Cy3-3'.