Supporting Information

Unusual topological RNA G-quadruplex formed of RNA duplex: Implications for dimerization of COVID-19 RNA

Shiyu Wang,[a,e] Yi Song,[a,e] Zhiyong He,[a] Hisao Saneyoshi,[a] Rie Iwakiri,[a] Pengyu Xu,[b] Chuanqi Zhao,[c] Xiaogang Qu[c,d] and Yan Xu[*a]

[a]Division of Chemistry, Department of Medical Sciences, Faculty of Medicine, University of Miyazaki, 5200 Kihara, Kiyotake, Miyazaki 889-1692, Japan.
[b]Shonan Laboratory, Corporate R&D Headquarters, Otsuka Chemical Co., Ltd. 2Chome26-1, Muraokahigashi, Fujisawa, Kanagawa 251-8555, Japan.
[c]Laboratory of Chemical Biology and State Key Laboratory of Rare Earth Resource Utilization, Changchun Institute of Applied Chemistry, Chinese Academy of Science, Changchun, Jilin 130022, P. R. China.
[d]School of Applied Chemistry and Engineering, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China.
[e]The two authors contributed equally: Yi Song, Shiyu Wang

* Corresponding authors: Yan Xu

Yan Xu: xuyan@med.miyazaki-u.ac.jp

Experimental section

RNA sample preparation

The native, Cy3-, Cy5-, and trifluoromethyl- labeled RNAs were synthesized by using an automatic solid-phase phosphoramidite chemistry and DNA/RNA synthesizer. RNAs were cleavage from the column and deprotected by using AMA (Ammonium Hydroxide/40% aqueous methylamine 1:1 v/v) at room temperature for 20 minutes and at 65°C or 10 minutes. TBDMS protections were removed by treatment with triethylamine trihydrofluoride followed by filtration through an ion exchange cartridge. The oligomers were further purified by HPLC in a linear gradient of 50 mM ammonium formate in 1:1 acetonitrile/H_2O and 50 mM ammonium formate in H_2O. The oligomers were desalted through a NAP 10 column (GE Healthcare) and identified by MALDI-TOF-MS on an Autoflex III smart beam mass spectrometer (negative mode). RNA sequences used in the study:

17 mer 5'-UGGCUGGCAAUGGCGGU-3' (ORN-1),

Cy3-5'-UGGCUGGCAAUGGCGGU-3' (Cy3-RNA),

Cy3-5'-UGUCUGUCAAUGGCAGC-3' (Cy3-RNA1),

Cy5-5'-UGGCUGGCAAUGGCGGU-3' (Cy5-RNA),

5'-UGGCUGGCAAUGGCG(^{F}rG)U-3' (ORN-2).

5'-ACUUCUCCUGCUAGAAUGGCUGGCAAUGGCGGUUGAGCUGCUCUUGCUUU-3' (ORN-3).

5'-UGGCUGGCAAUGGCGAU-3' (Mutant-RNA1).

5'-UAACUAACAAUAGCAGU-3' (Mutant-RNA2).

5'-UAGGGUUAGGGU-3' r$(UAGGGU)_2$

5'-UAGGGUUAGGGUUAGGGUUAGGGU-3' r$(UAGGGU)_4$

RNase T1 footprint assay

The Rnase T1 was purchased from ambion®, and the RNA structure analysis proceed was followed the product manual. 5′ end FAM labeled RNAs were heated at 90°C for 5 min and subsequently cooled to 4°C in 100 mM KCl or 100 mM NaCl in buffer to assist the structure formation. The structured RNA was digested with RNase T1 (0.025 units) for 5 min at room temperature. ALK ladder was generated by alkaline hydrolysis of RNA. Denaturing gel electrophoresis experiment was performed on 10 M urea/20% polyacrylamide gel in 1 × TBE buffer.

CD measurement

CD experiments were performed by using a JASCO model J-820 CD spectrophotometer (JASCO Corporation, Tokyo, Japan). The RNA samples were prepared at a 10 μM concentration in the presence of 100 mM KCl or LiCl and 20 mM K-PO_4 buffer (pH 7.0). The melting curve was obtained by monitoring 270 nm between 10 °C - 90 °C, 0.3 mL samples at 10 μM concentration, 100 mM KCl or LiCl and 20 mM K-PO_4 buffer (pH 7.0).

Fluorescent measurement

Fluorescence spectroscopy was carried out on a JASCO FP-8200 spectrofluorometer. The spectra were recorded using a 1-cm path-length cell. The excitation wavelength was selected at 540 nm (Cy3). The sample solutions were as follows: 10 mM Tris-HCl buffer (pH 7.0), 100 mM KCl, and 1.0 μM each RNA strand. Fluorescence imaging was performed using an imaging system (Lumazone CMS, Shoshin EM, Japan), the excitation/emission filter of 555−605 nm (Cy3)/ 660−780 nm (Cy5) were used. The sample solutions were as follows: 1.0 μM RNA in 100 mM KCl and 10 mM Tris-HCl buffer pH 7.0.

Gel shift assay

In general, 20 μM RNAs in desired condition treated with heat denaturation, renaturation and was stored at 25 ℃ at least for 12 h. Then 3 μL of 30% glycerol was added before loading onto gel. Native gel electrophoresis (PAGE) was run on 20% polyacrylamide gel (acrylamide/bisacrylamide, 29:1) at 4 ℃, in 1X TBE buffer. Subsequently, RNAs were stained with DMSO and imaged using excitation/emission filter of 488 nm/515 nm. For the FRET experiments, the Cy3-labeled RNA (20 μM) with addition of Cy5-labeled RNA were performed in a buffer, 10 mM Tris−HCl, pH 7.0, 200 mM KCl. Each sample was treated with heat denaturation, renaturation and was stored at 4°C at least for 12 h. Then 3 μl of 30% glycerol was added before loading onto gel. Native gel electrophoresis (PAGE) was run on 20% polyacrylamide gel at 4°C, in 1X TBE buffer (Tris/borate/EDTA) with the addition of 20 mM KCl. The green bond was obtained by using excitation/emission filter of 550 nm/575 nm (Cy3 channel). The red bond was obtained by using excitation/emission filter of 635 nm/670 nm (Cy5 channel). The yellow bond was the result of overlapping Cy3 channel and Cy5 channel manually.

^{1}H-NMR spectroscopy

NMR data were recorded on a BRUKER AVANCE 400 and 600 MHz spectrometer. For spectra

recorded in 90%H_2O/10% D_2O water signal was suppressed using the 3–9–19 WATERGATE pulse sequence or excitation sculpting with gradient pulse. The data were processed with TopSpin 3.0 (Bruker BioSpin Gmbh) software and analyzed with MestReNova software. For 1D NMR measurement, 1.0 mM native RNAs were dissolved in 150 μl of designed solution containing 10% D_2O, 100 mM KCl or LiCl and 20 mM K-PO_4 or Na-PO_4 buffer (pH 7.0). The spectrum were collected by 2000 scans at 10 °C. The assignment of exchangeable resonances was based on homonuclear ^{1}H-^{1}H NOESY. The 2D NMR spectrum in 90% H_2O/10% D_2O was collected from 360 scans with 150 ms mixing time at 10 °C. On average, 2048 complex points and 512 FIDs were collected within the spectral width of 14097 Hz. The sample solutions were as follows: 3 mM native RNA was dissolved in 150 μl of designed solution containing 10% D_2O, 100 mM KCl and 20 mM Na-PO_4 buffer (pH 7.0). Samples were prepared by heating the oligonucleotides at 85 °C for 3 minutes and gradually cooling to room temperature.

^{19}F-NMR spectroscopy

For ^{19}F NMR measurement, RNA samples of 1 mM concentration were dissolved in 150 μl of designed solution containing 10% D_2O, 200 mM KCl, 20 mM K-PO_4 buffer (pH 7.0). Samples were prepared by heating the ^{19}F-labeled oligonucleotides at 85 ℃ for 3 minutes and gradually cooling to room temperature. The ^{19}F NMR spectrum was measured on a Bruker AVANCE 400 MHz spectrometer at a frequency of 376.05 MHz and referenced to the internal standard CF_3COOH (–75.66 ppm). The experimental parameters are recorded as follows: spectral width 89.3 kHz, ^{19}F excitation pulse 15.0 μs, relaxation delay 1.5 s, acquisition time 0.73 s, scan numbers 256, and line width 3. At temperature-dependent experiment, the sample is kept for 10 ~ 20 minutes at each temperature.

In-cell ^{19}F-NMR spectroscopy

The experiments accessible of ^{19}F-labeled RNA entry into HeLa cells by SLO treatment is detected by comparing the fluorescence intensity in SLO-treated cells and SLO-untreated cells as shown in our reported article.[1] For in-cell ^{19}F NMR measurement, the transfected cells were suspended in 200 μl of DMEM containing 10% D_2O and transferred to a Shigemi tube (Shigemi 5 mm Symmetrical NMR microtube). The experiment was performed at 10°C with a scan numbers value in 10^4. After the intracellular NMR measurement, 100 μL of DMEM was added to the cell suspension, and the supernatant was collected by centrifugation at 400 × g for 3 minutes. The ^{19}F NMR spectrum of the supernatant was measured with the same number of scans as the in-cell ^{19}F NMR measurement. The remaining cells are mixed with 200 μl of DMEM containing 10% D_2O and disrupted by ultrasound, the cell lysate was obtained and

transferred to a Shigemi tube for performing ^{19}F NMR spectrum with the same number of scans as the in-cell ^{19}F NMR measurement.

ESI-MS

ESI-MS experiments were performed on Q Exactive Quadrupole-Orbitrap LC-MS System (Termo Scientifc) equipped with a XBridge Oligonucleotide Separation Technology C18 column (2.5 μm, 4.6 mm×50 mm, Waters). Data was acquired using Xcalibur sofware. Sample was prepared at 0.1 mM concentration in 150 mM ammonium acetate at 4 °C. A spray voltage 3.5kV was used to generate the negative ion mode with a heated capillary temperature of 250 °C. An ion accumulation time was 100 ms.

Fluorescence imaging in HeLa cells

HeLa cells were grown at 37°C in a humidified atmosphere of 5% CO_2. Cells were maintained by regular passage in DMEM (Nissui). The medium was supplemented with 10% heat-inactivated fetal bovine serum (FBS), 4 mM L-glutamine, 100 U/ml penicillin and 100 μg/mL streptomycin (Gibco by Life Technologies). For chemical transfection, HeLa cells were grown on Ø35 mm coverslips for 24 h in complete DMEM. Transfection was carried out using RNAs with desirable concentrations, DMEM medium and Lipofectamine 3000 (Invitrogen™). Transfection was carried out over 16 h at 37°C. Cells were fixed with 4% formaldehyde. Confocal images were acquired with a laser of TCS SP8 for 20× HC PL APO CS2 20×/0.75 Dry and for 63× was HC PL APO CS2 63×/1.40 OIL. The green fluorescence for Cy3 was using excitation wavelength of 550 nm. The red fluorescence for Cy5 was an excitation wavelength of 650 nm. The excitation/emission filter of 530-550 nm/650-700 nm was used for FRET detection. The sample solutions were as follows: 10 or 25 or 100 μM Cy3-RNA with 100 or 250 or 1000 μM Cy5-RNA in 100 mM KCl, 10 mM Tris-HCl buffer pH 7.0 before use.

Molecular Modeling

We manually generated the model of RNA structure based on the reported structure (PDB code 3NZ7, 3ES0, 3EQW and 3EM2) using the BIOVIA Discovery Studio 4.5. The molecular dynamics simulation was performed by the standard dynamics cascade in BIOVIA Discovery Studio 4.5 with some modifications. The structure was heated from 50 K to 283 K over 4 ps and equilibration at 283 K with 100 ps simulation time. The save results interval in the production step was 2 ps during 100 ps simulation time at 283 K. 10 best conformations generated by simulation were further energy minimized, and the conformation with lowest energy was selected.

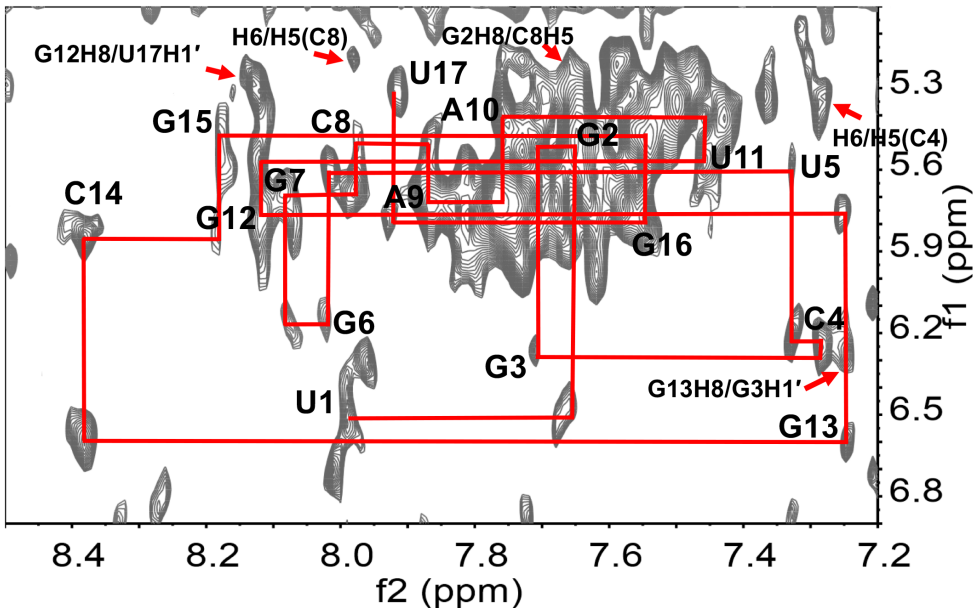

Figure S1. The H8/H6-H1′ proton region of NOESY spectrum (mixing time, 250 ms) of ORN-1 in K^+ solution. The sequential pathway is shown (red line). Intraresidue H6/H8−H1′ NOE cross-peaks are labeled with residue numbers. Cross peaks G12H8/U17H1′, H6/H5(C8), G2H8/C8H5, H6/H5(C4) and G13H8/G3H1′ are indicated by red arrows.

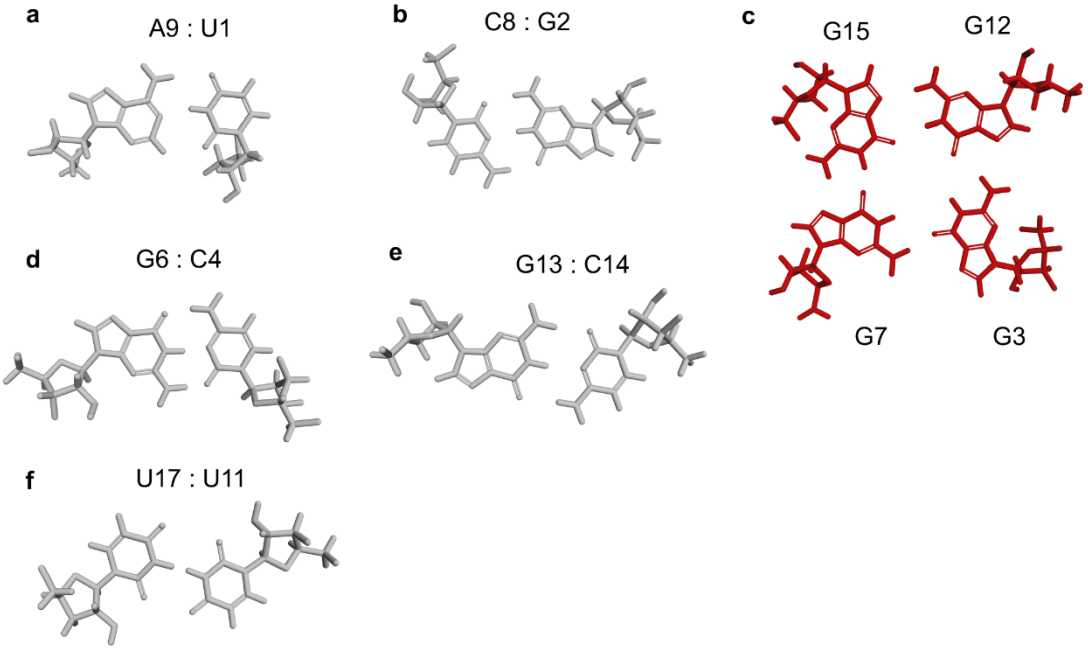

Figure S2. Watson–Crick and Hoogsteen base pairs. (a) A9:U1, (b) G2:C8, (c) G-tetrad G3:G7:G15:G12, (d) G6:C4, (e) G13:C14, and (f) U17:U11.

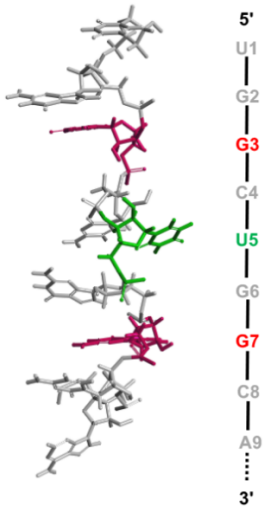

Figure S3. A stereoview of U5 (green) that is bulged out of the strand.

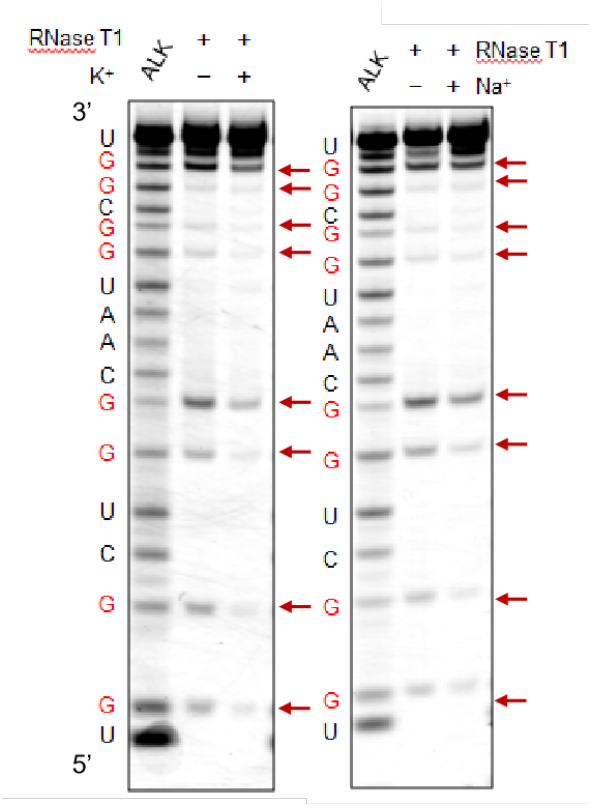

Figure S4. RNase T1 footprinting assay of ORN-1 in presence of different conditions. ORN-1 labeled with FAM fluorescence in the presence or absence of K^+ or Na^+-containing binding buffer was treated with RNase T1. ALK indicated base hydrolysis ladders of ORN-1 with alkali. rG residues that involved in the structure are indicated by arrows.

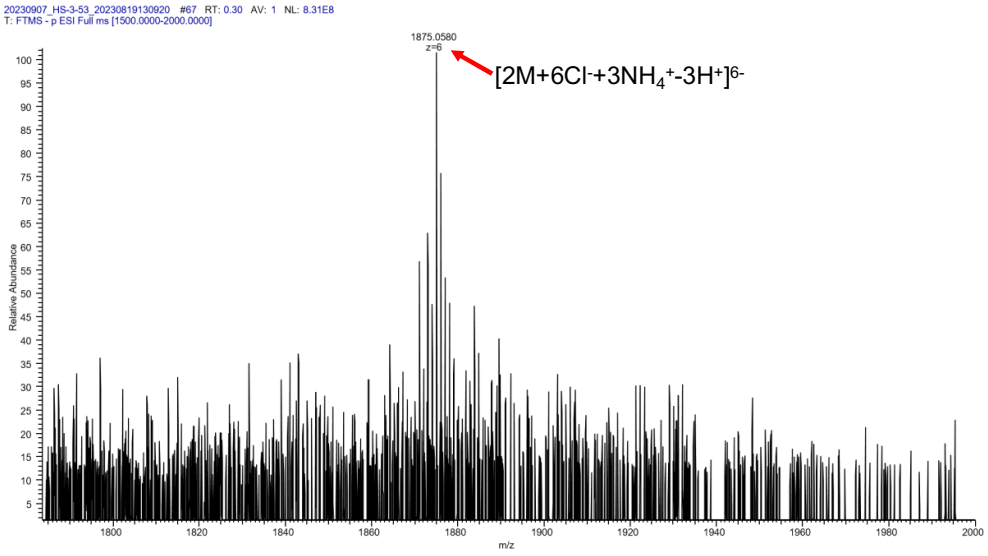

Figure S5. ESI-MS spectrum of ORN-1 molecular ions (8^- and 9^-) are directly observed for dimer G-quadruplex by electrospray ionization MS. 0.1 mM RNA in 150 mM ammonium acetate kept at 4 °C before used. M: molecular weight.

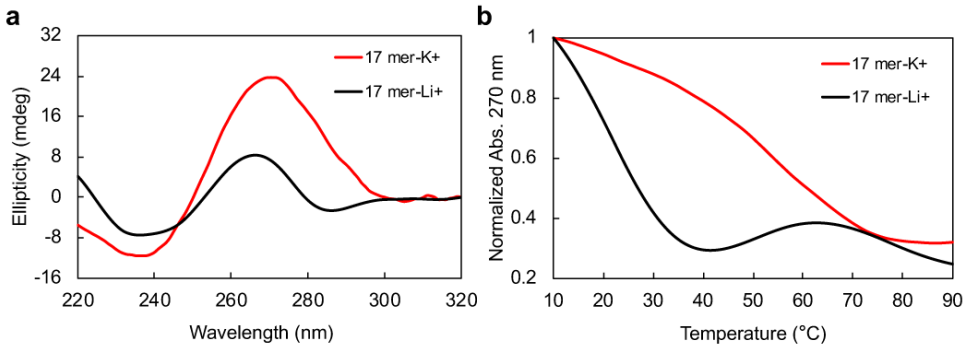

Figure S6. (a) CD spectra of ORN-1 (10 μM) in 100 mM KCl or LiCl. (b) CD thermal melting curves of ORN-1 (10 μM) monitored at 270 nm in the presence of 100 mM KCl or LiCl.

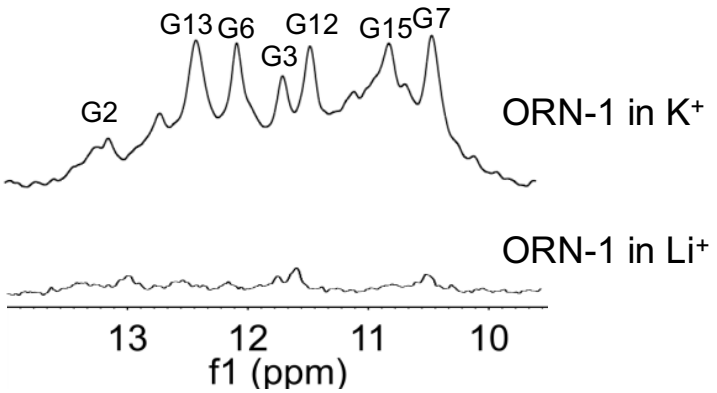

Figure S7. Imino-proton NMR spectrum of ORN-1 in 100 mM LiCl and KCl with 10 mM Tris-HCl, pH 7.0, 10 ℃.

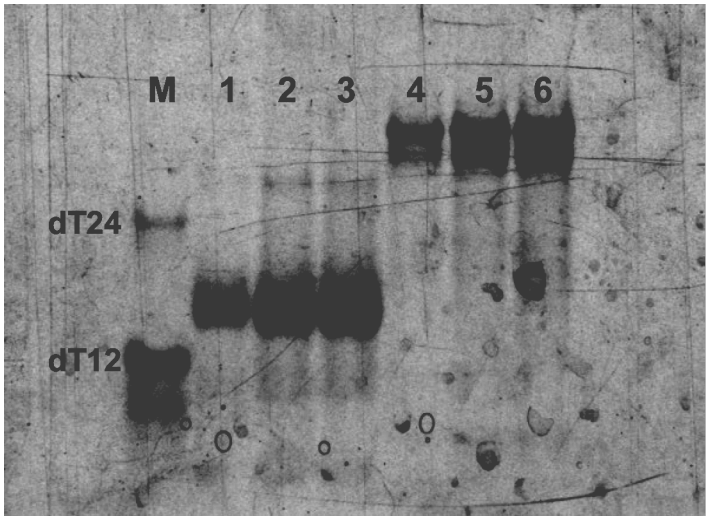

Figure S8. Non-denaturing gel electrophoresis of ORN-1 in 100 mM KCl or LiCl at 4 °C. Lane M, makers dT12 and dT24; lane 1, 2 and 3: 5, 10 and 15 μM ORN-1 in 100 mM LiCl; lane 4, 5 and 6: 5, 10 and 15 μM ORN-1 in 100 mM KCl.

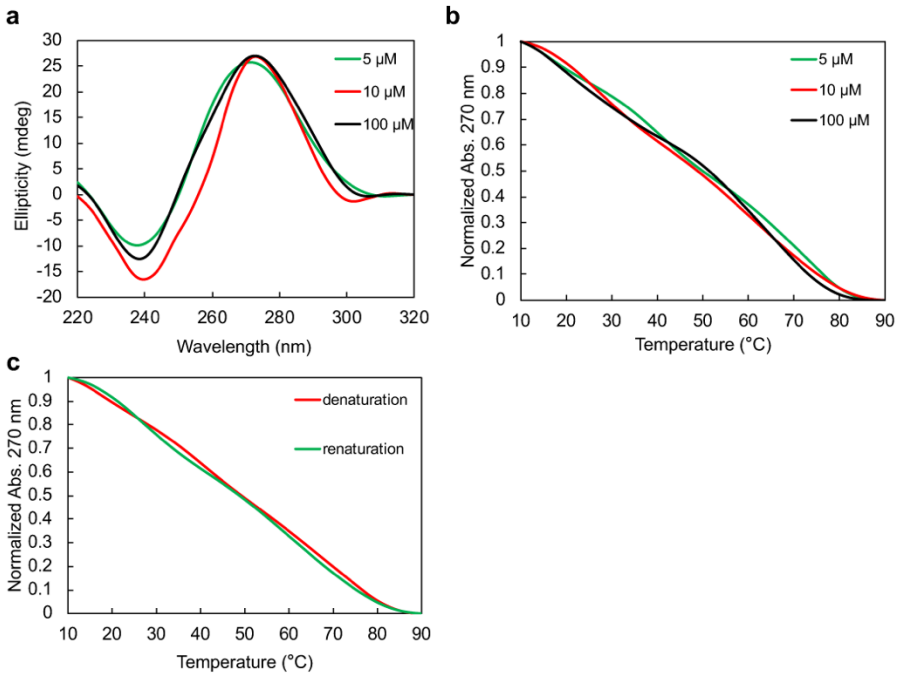

Figure S9. (a) CD spectra of 5, 10 and 100 μM 17 mer-RNA in the presence of 100 mM KCl at 10 ℃. 10 mM Tris-HCl (pH 7.0). (b) CD melting curves of 5, 10 and 100 μM 17 mer-RNA in the presence of 100 mM KCl with 10 mM Tris-HCl (pH 7.0). (c) CD Curves of 10 μM 17 mer-RNA at 270 nm absorbance intensity with heating from 10 °C to 90 °C (red line) and cooling from 90 °C to 10 °C (green line). Condition: 100 mM KCl and 10 mM Tris-HCl (pH 7.0).

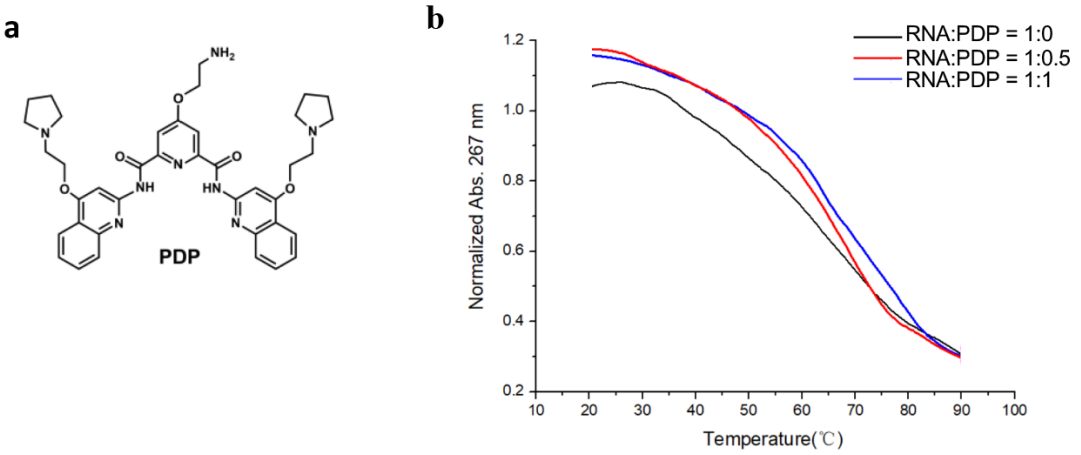

Figure S10. (a) Chemical structures of PDP. (b) CD thermal melting curves of ORN-1 (15 μM) monitored at 270 nm in the absence or presence of PDP (15 and 30 μM).

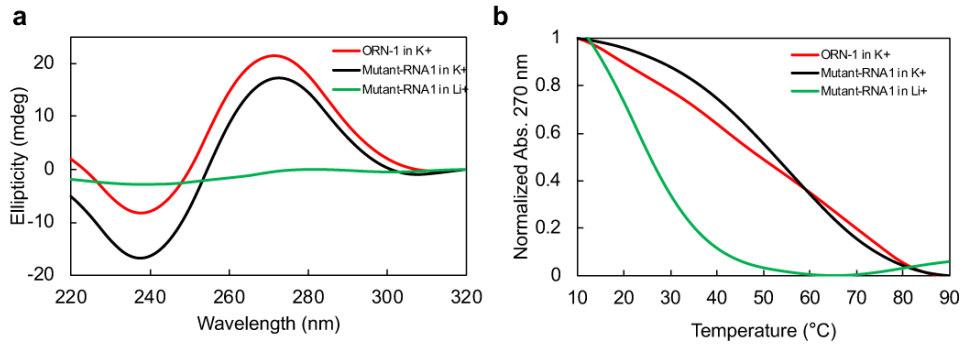

Figure S11. (a) CD spectra of ORN-1 (10 μM) in 100 mM KCl as well as mutant-RNA1 (10 μM) in 100 mM KCl or LiCl. (b) CD thermal melting curves of ORN-1 (10 μM) in 100 mM KCl as well as mutant-RNA1 (10 μM) in 100 mM KCl or LiCl.

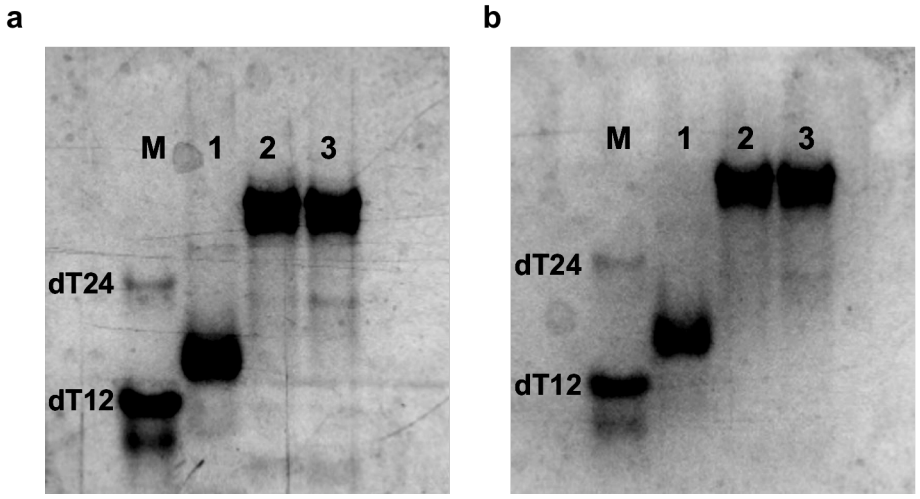

Figure S12. (a) Non-denaturing gel electrophoresis of mutant-RNA1 and ORN-1 at 4 °C. Lane M: makers dT12 and dT24; lane 1: mutant-RNA1 in 100 mM LiCl; lane 2: mutant-RNA1 in 100 mM KCl; lane 3: ORN-1 in 100 mM KCl. (b) Non-denaturing gel electrophoresis of mutant-RNA1, mutant-RNA2 and ORN-1 at 4 °C. Lane M: makers dT12 and dT24; lane 1: mutant-RNA2 in 100 mM KCl; lane 2: mutant-RNA1 in 100 mM KCl; lane 3: ORN-1 in 100 mM KCl.

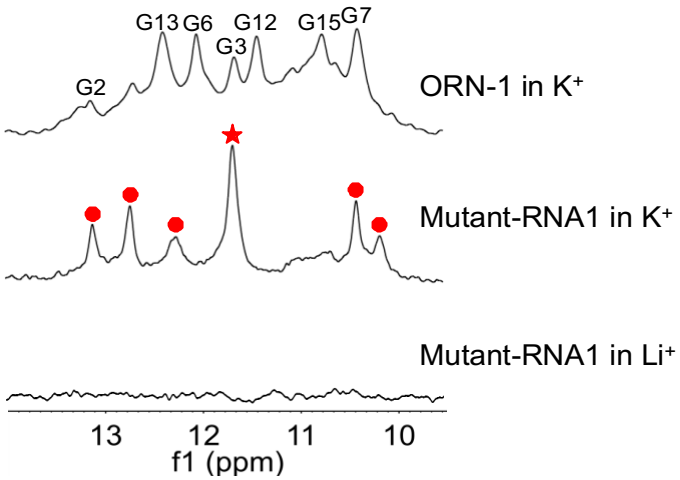

Figure S13. Comparison of imino-proton NMR spectrum of ORN-1 in 100 mM KCl, mutant-RNA1 in 100 mM LiCl and KCl, 10 mM Tris-HCl, pH 7.0, 10 °C. Red balls represent imino proton resonances were found. Asterisk indicates an overlapping of two peaks.

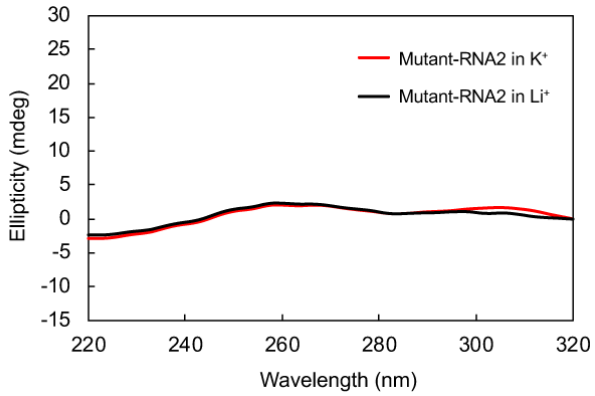

Figure S14. (a) CD spectra of mutant-RNA2 (10 μM) in 100 mM KCl or LiCl.

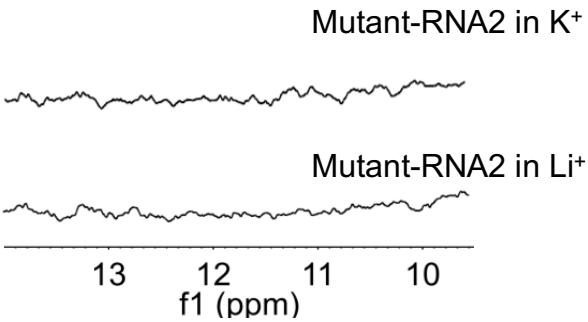

Figure S15. Imino-proton NMR spectra of mutant-RNA2 in 100 mM LiCl and KCl with 10 mM Tris-HCl, pH 7.0, 10 °C.

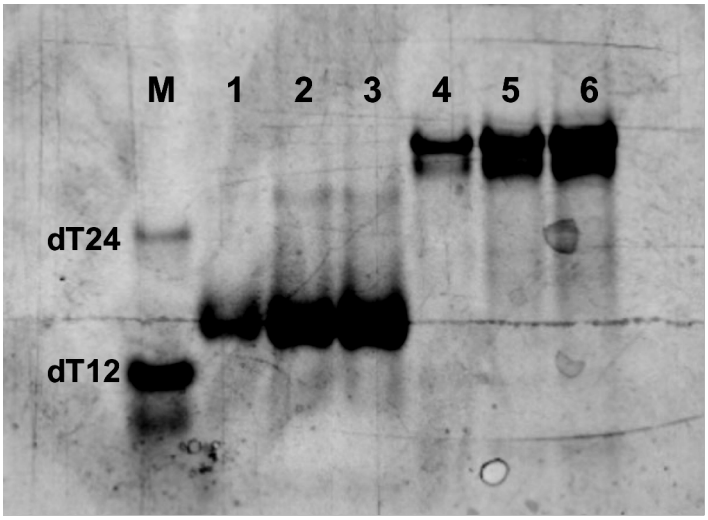

Figure S16. Non-denaturing gel electrophoresis of mutant-RNA2 and ORN-1 in 100 mM KCl at 4 °C. Lane M, makers dT12 and dT24; lane 1, 2 and 3: mutant-RNA2 (5, 10 and 15 μM); lane 4, 5 and 6: ORN-1(5, 10 and 15 μM).

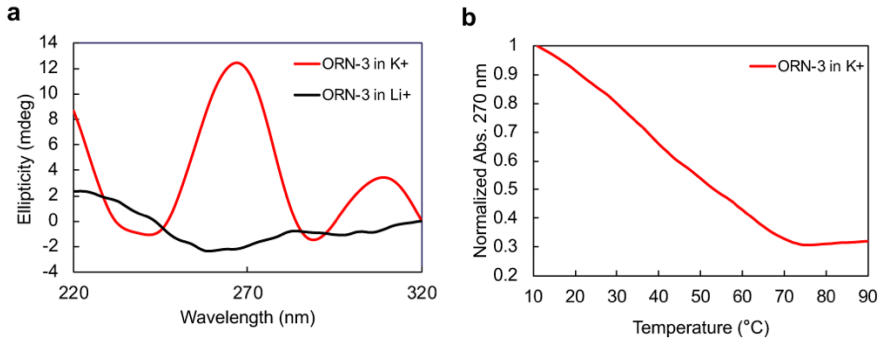

Figure S17. (a) CD spectrum of 50 mer ORN-3 (10 μM) in 100 mM KCl or LiCl. (b) CD thermal melting curves of ORN-3 (10 μM) monitored at 270 nm in 100 mM KCl.

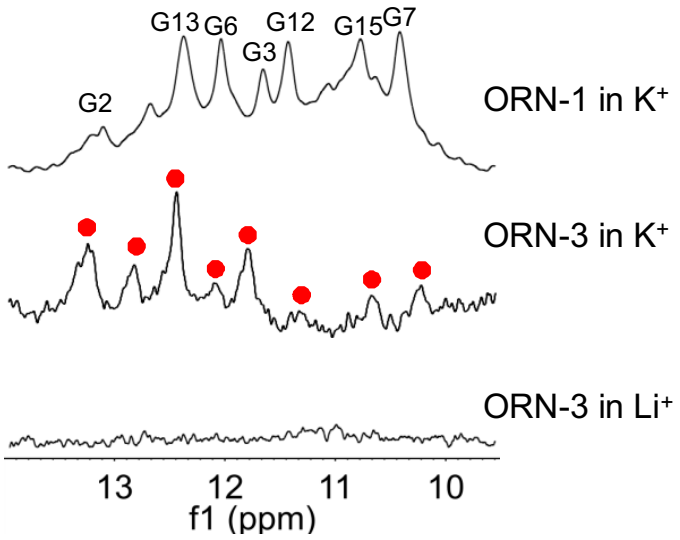

Figure S18. Comparison of imino-proton NMR spectrum of ORN-1 in 100 mM KCl, ORN-3 in 100 mM LiCl and KCl, 10 mM Tris-HCl, pH 7.0, 10 °C. Red balls represent imino proton resonances were found.

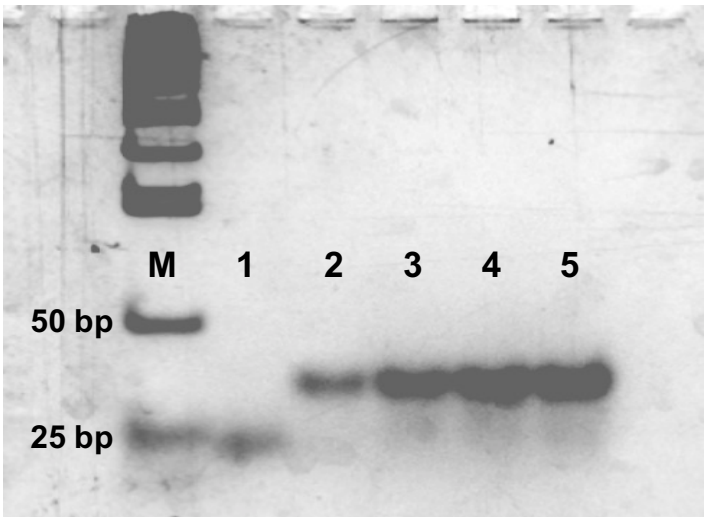

Figure S19. Non-denaturing gel electrophoresis of ORN-3 in 100 mM KCl or LiCl at 4 °C. M: DNA ladders show 25 and 50 bp; lane 1: 10 μM ORN-3 in LiCl; lane 2, 3, 4 and 5 show ORN-3 with 10, 15, 20 and 25 μM in KCl.

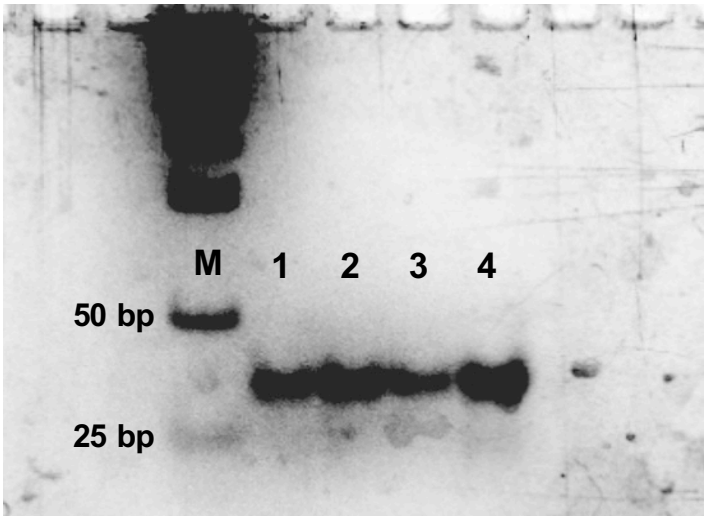

Figure S20. Non-denaturing gel electrophoresis of ORN-3 in metal ions at 4 ℃. M: DNA ladders show 25 and 50 bp. lane 1 and 2 represent 10 μM and 20 μM of ORN-3 in the 100 mM mixture of LiCl and KCl (50/50, mol/mol). lane 3 and 4 represent 10 μM and 20 μM of ORN-3 in the 100 mM KCl.

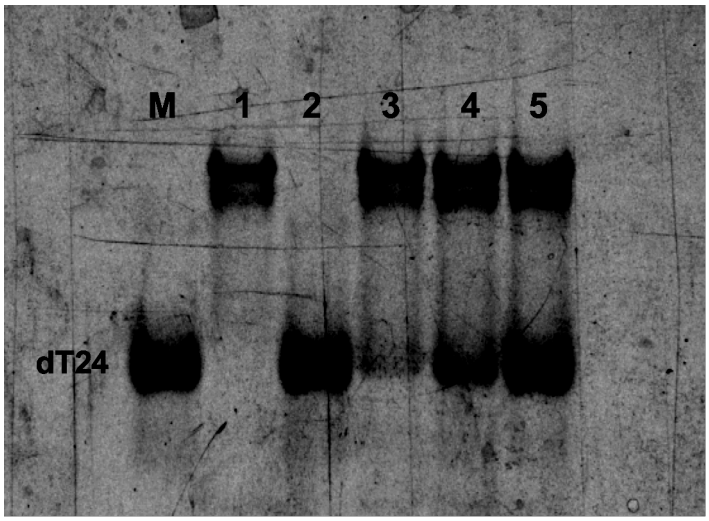

Figure S21. Non-denaturing gel electrophoresis of ORN-1, r(UAGGGU)$_2$ and the mixture of ORN-1 and r(UAGGGU)$_2$ in 100 mM KCl at 4 °C. Lane M: maker dT24; lane 1: 15 μM ORN-1; lane 2: 15 μM r(UAGGGU)$_2$; lane 3, 4 and 5 represent 15 μM ORN-1 with r(UAGGGU)$_2$ at different ratio, 1:0.2, 1:0.5 and 1:1 respectively.

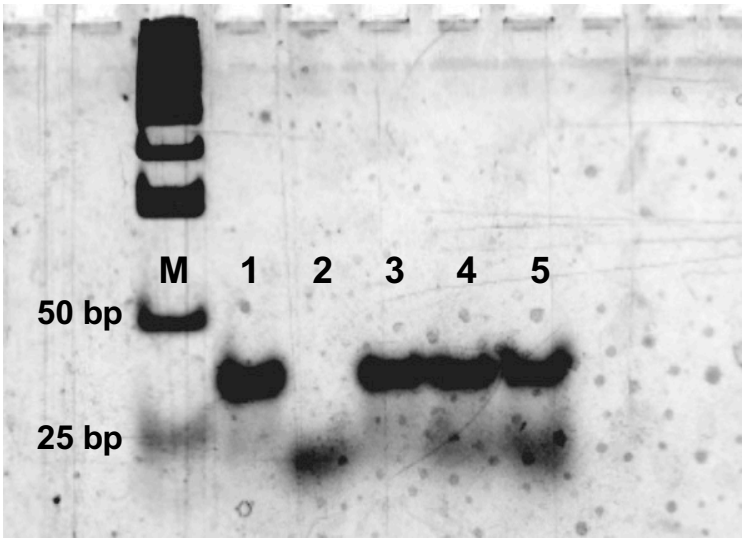

Figure S22. Non-denaturing gel electrophoresis of ORN-3, r(UAGGGU)$_2$ and the mixture of ORN-3 and r(UAGGGU)$_2$ in 100 mM KCl at 4 °C. Lane M: DNA ladders show 25 and 50 bp. lane 1: 15 μM ORN-3; lane 2: 15 μM r(UAGGGU)$_2$; lane 3, 4 and 5 represent 15 μM ORN-3 with r(UAGGGU)$_2$ at different ratio, 1:0.2, 1:0.5 and 1:1 respectively.

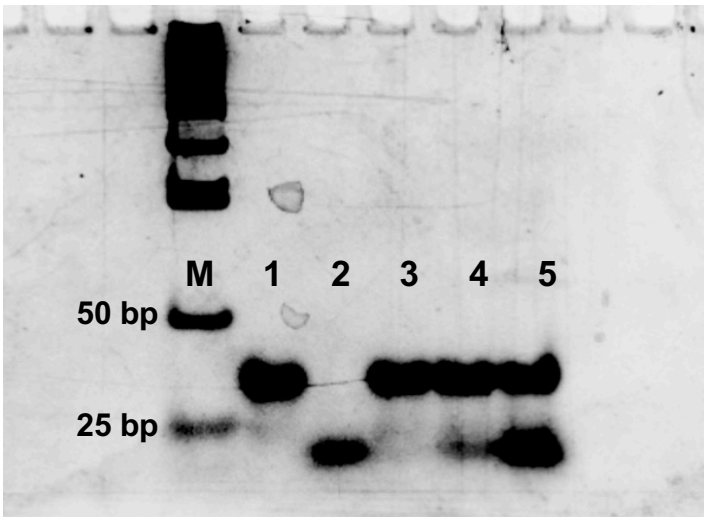

Figure S23. Non-denaturing gel electrophoresis of ORN-3, r(UAGGGU)$_4$ and the mixture of ORN-3 and r(UAGGGU)$_4$ in 100 mM KCl at 4 °C. Lane M: DNA ladders show 25 and 50 bp. lane 1: 15 μM ORN-3; lane 2: 15 μM r(UAGGGU)$_4$; lane 3, 4 and 5 represent 15 μM ORN-3 with r(UAGGGU)$_4$ at different ratio, 1:0.2, 1:0.5 and 1:1 respectively.

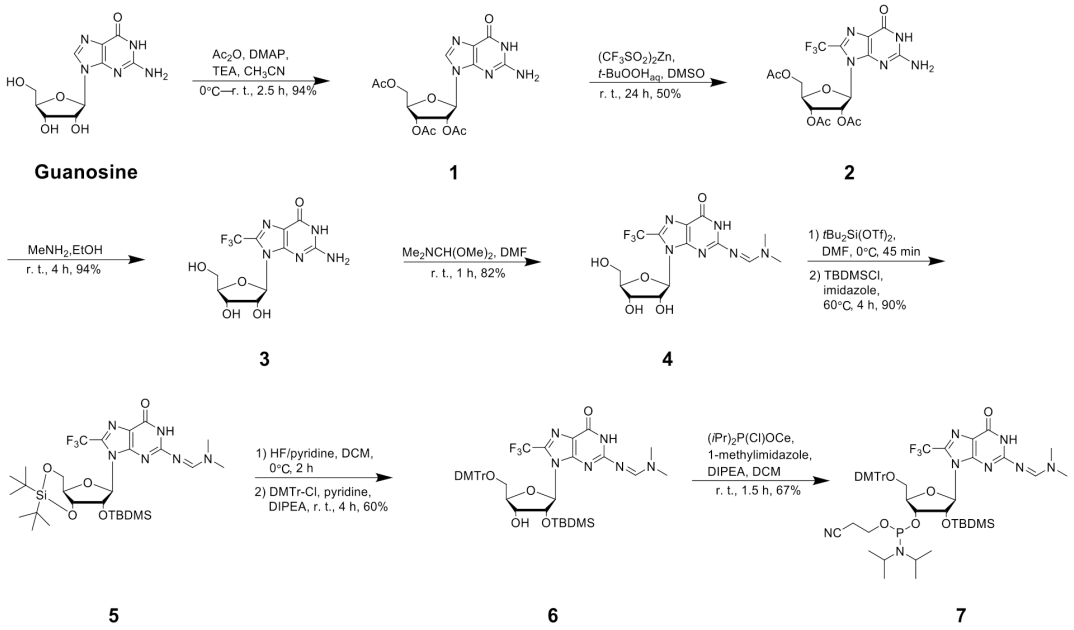

Figure S24. Approach for synthesis of 8-trifluoromethyl-guanosine (^{F}rG) nucleoside and incorporated it into RNA sequences by phosphoramidite chemistry. ^{F}rG is site-specific and incorporated at the G16 position of 5'-UGGCUGGCAAUGGCG(^{F}rG)U-3' (ORN-2), in which G16 does not form base pair with others and has no effect on the structure.

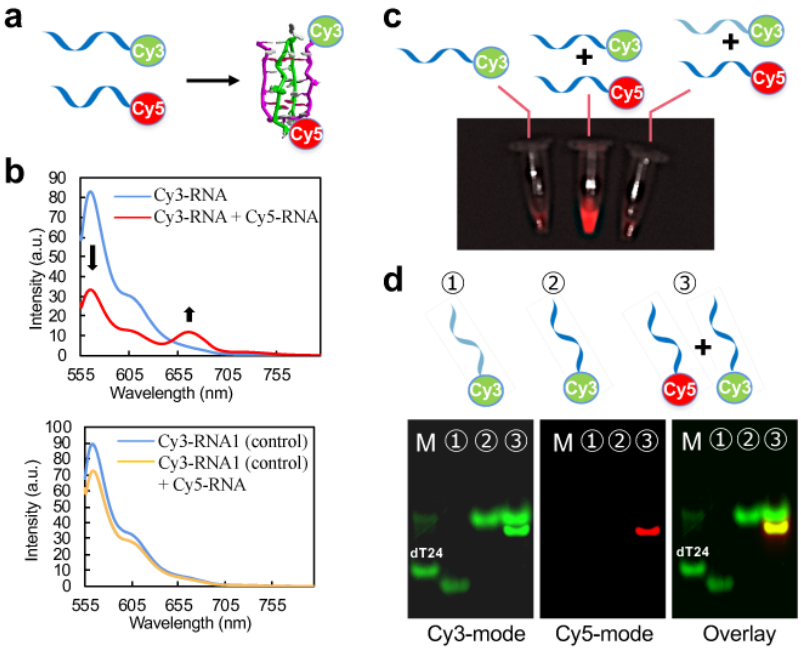

Figure S25. FRET was employed to monitor dimer formation. (a) Use of Cy3 and Cy5 fluorophores labeled RNA to probe dimerization RNA structure. G-quadruplex formation brings Cy3 and Cy5 molecules close to each other to display FRET. (b) Emission spectra of Cy3-, Cy5-labeled 17-mer RNA and a control Cy3-RNA1 without SARS-COV-2 RNA sequence. Addition of Cy5-RNA to Cy3-RNA (upper), showing resulting Cy3-Cy5 FRET upon dimeric G-quadruplex formation (excitation wavelength = 550 nm) but not in using control Cy3-RNA1 (lower). The emission intensities at 570 nm and 670 nm correspond to emission from Cy3 and Cy5 respectively. (In general, Cy3 excitation wavelength = 550 nm, emission wavelength = 570 nm; Cy5 excitation wavelength = 650 nm, emission wavelength = 670 nm). (c) Fluorescence image of Cy3-, Cy5-labeled 17-mer RNA and control Cy3-RNA1 (as an orange strand) after observation at an imaging system. (d) Visualization of the dimerization RNA G-quadruplex structure with Cy3 and Cy5 labeled RNA. Lane M: maker dT24; Lane 1, control Cy3-RNA1 (as an orange strand); Lane 2, Cy3-RNA; Lane 3, Cy3-RNA /Cy5-RNA. Cy3-mode: excitation wavelength = 550 nm, emission wavelength = 570 nm, Cy5-mode: excitation wavelength = 650 nm, emission wavelength = 670 nm. Maker is not chromogenic at Cy5 channel (670 nm) even after staining.

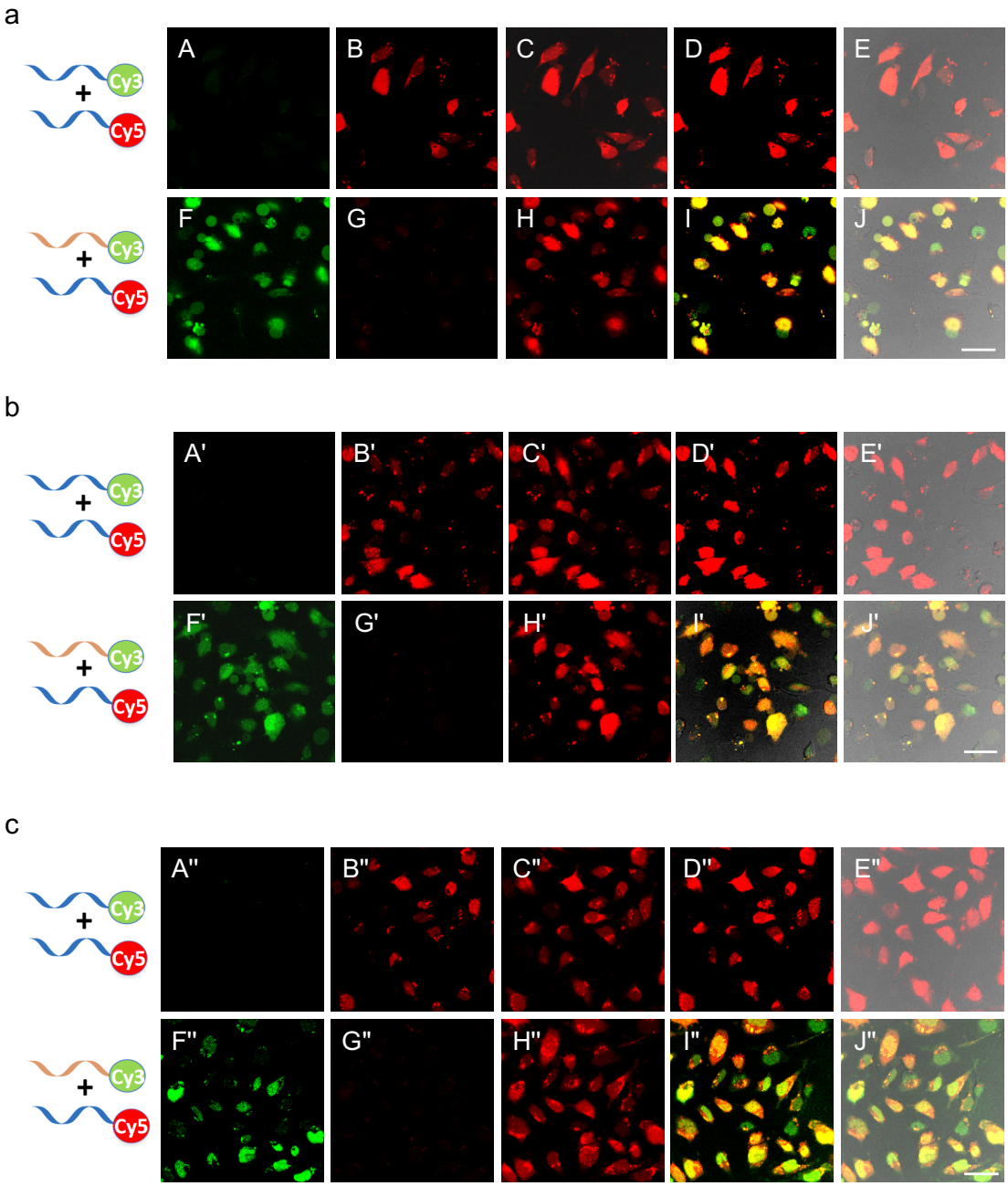

Figure S26. FRET was employed to monitor dimer formation in cells. HeLa cells were incubated with Cy3-, Cy5-labeled 17-mer RNA/and control Cy3-RNA1 (as an orange strand) imaged using confocal microscopy. Images A/A′/A″ and F/F′/F″ represents the Cy3 channel, B/B′/B″, C/C′/C″, G/G′/G″ and H/H′/H″ represent the Cy5 channel. Images A/A′/A″, B/B′/B″, F/F′/F″ and G/G′/G″ are excited with a 550 nm laser, C/C′/C″ and H/H′/H″ are excited with a 650 nm laser, D/D′/D″ and I/I′/I″ represent overlay of all the channels respectively for control group or experiment group, E/E′/E″ and J/J′/J″ represent overlay results in bright-field. Scale bar: 50 μm. Condition: (a) 1 μM Cy3-labeled RNA and 10 μM Cy5-labeled RNA. (b) 2.5 μM Cy3-labeled RNA and 25 μM Cy5-labeled RNA. (a) 10 μM Cy3-labeled RNA and 100 μM Cy5-labeled RNA.

Synthetic Methods

General

^{1}H-NMR, ^{19}F-NMR and ^{31}P-NMR spectra were recorded on a BRUKER (AV-400M) magnetic resonance spectrometer. DMSO-d_6 was 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 s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet. All reagents were purchased from Sigma-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). High-resolution mass spectra (HRMS) were recorded by electrospray ionization (ESI) on a Thermo Scientific Q Exactive instrument.

2',3',5'-tri-*O*-acetylguanosine (1)

Guanosine (1500 mg, 5.2 mmol), trimethylamine (5.85 mL, 41.4 mmol) and 4-dimethylaminopyridine (69 mg, 0.56 mmol) were dissolved in 20.25 mL anhydrous acetonitrile, acetic anhydride (1.65 mL, 16.5 mmol) was added dropwise and the mixture reacted for 1.5 h at 0 ℃ and another 1 hour at room temperature. The reaction was quenched by additional methanol (1.725 mL, 42.68 mmol). The volume was reduced to 1/3 using a rotary evaporator and diethyl ether was added dropwise to induce precipitation of a fine white powder. The product was collected by filtration, washed with diethyl ether, and then stirred for 2 h in acetone (30 mL) at 50 ℃. The filtrate produced 1800 mg (94%) of a fine white powder. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.73 (s, 1H), 7.94 (s, 1H), 6.55 (s, 2H), 5.99 (d, J = 8.0 Hz, 1H), 5.80 (t, J = 8.0 Hz, 1H), 5.50 (dd, J = 4.0 Hz, 1H), 4.41-4.25 (m, 3H), 2.12-2.05 (m, 9H); HRMS (ESI) for $C_{22}H_{35}N_6O_8$ $[M+TEA+H]^+$: Calcd. 511.2507; Found. 511.2494.

2',3',5'-tri-*O*-acetyl-8-trifluoromethylguanosine (2)

2',3',5'-tri-*O*-acetyl-guanosine (1800 mg, 3.8 mmol) and Zinc Trifluoromethanesulfinate (4500 mg, 12.2 mmol) were dissolved in dimethylsulfoxide (23 ml) and vigorous stirring. When the clear solution was formed (15-20 min), tert-butyl hydroperoxide (70% solution in water, 2.56 ml, 19.2 mmol) in 10 aliquots (256 µl each) in 20 min. intervals. The reaction mixture gradually turn yellow during addition of tert-Butyl hydroperoxide and continue reacted for 24 h at room temperature. The mixture was poured into 300 mL water and extracted with dichloromethane (3 × 80 mL). The combine organic layers were washed with water (3 × approx. 50 ml), brine (approx. 50 ml) and dried over sodium sulfate. The drying agent was filtered off, washed with dichloromethane and the filtrate was concentrated in vacuo. The oily residue was purified by MPLC (medium pressure liquid chromatography) with the mixture of methanol in dichloromethane (5%, v/v) to give product 2 (1600 mg, 50%). ^{1}H NMR (400 MHz, DMSO-d6) δ 12.20 (s, 1H), 6.29 (s, 1H), 5.95 (t, J = 6.4 Hz, 2H), 4.55-4.52 (m, 1H), 4.46-4.38 (m, 2H), 2.15-2.04 (m, 9H); ^{19}F NMR (400 MHz, DMSO-d_6) δ -61.1843 (s, 3F); HRMS (ESI) for C17H15O8N5F3Na $[M+Na]^+$: Calcd. 500.1065; Found. 500.0981.

8-trifluoromethylguanosine (3)

2',3',5'-tri-*O*-acetyl-8-trifluoromethylguanosine (1600 mg, 3.34 mmol) was placed in 200 mL round bottom flask. Methylamine (33% in ethanol, 17.76 mL, 170.25 mmol) was added for

getting reaction mixture, and the mixture reacted for 4 h at room temperature. The reaction solution was concentrated *in vacuo* and the residue was purified by MPLC with the mixture of methanol in dichloromethane (10%, v/v). The product 3 was giving as yellowish solid (995mg, 94%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.05 (s, 1H), 6.72 (s, 2H), 5.64 (d, *J* = 6.0 Hz, 1H), 5.50 (d, *J* = 4.4 Hz, 1H), 5.14-4.92 (m, 3H), 4.17 (dd, *J* = 4.8 Hz, 1H), 3.91 (dd, *J* = 3.2 Hz, 1H), 3.71-3.52 (m, 2H); ^{19}F NMR (400 MHz, DMSO-d_6) δ -59.8351 (s, 3F); HRMS (ESI) for $C_{11}H_{11}O_5N_5F_3$ $[M-H]^-$: Calcd. 350.0688; Found. 350.0695.

***N2*-dimethylformamidyl-8-trifluoromethylguanosine (4)**

8-trifluoromethylguanosine (995 mg, 2.84 mmol) and N,N-dimethylformamide dimethyl acetal (2.71 mL, 20.1 mmol) was dissolved in anhydrous dimethylformamide (20 mL). The mixture was reacted for 1 hour at room temperature and the solvent was evaporated *in vacuo* for getting residue. The residue was purified by MPLC with mixture of methanol in dichloromethane (20%, v/v), the product **4** was giving as white foam (960 mg, 82%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.76 (s, 1H), 8.54 (s, 1H), 5.68 (d, *J* = 6.0 Hz, 1H), 5.42 (d, *J* = 6.4 Hz, 1H), 5.24 (d, *J* = 4.8 Hz, 1H), 5.05 (*J* = 6.0 Hz, 1H), 4.89 (*J* = 4.6 Hz, 1H), 4.27 (dd, *J* = 5.2 Hz, 1H), 3.94 (dd, *J* = 3.6 Hz, 1H), 3.72-3.56 (m, 2H), 3.18 (s, 3H), 3.08 (s, 3H); HRMS (ESI) for $C_{14}H_{16}O_5N_6F_3$ $[M-H]^-$: Calcd. 405.1231; Found. 405.1225.

***N2*-dimethylformamidyl-8-trifluoromethyl-2'-*O*-(tert-butyldimethylsilyl)-3',5'-*O*-(di-tert-butylsilylene)guanosine (5)**

N^2-dimethylformamidyl-8-trifluoromethylguanosine (200 mg, 0.48 mmol) was dissolved in anhydrous dimethylformamide (2 mL), and di-tert-butylsilyl bis (trifuoromethanesulfonate) (230 mg, 0.53 mmol) was added dropwise. The mixture reacted for 15 minutes at 0 ℃. Imidazole (163.5 mg, 2.4 mmol) also was added and the mixture continue reacted for 15 minutes at 0 ℃ and another 15 minutes at room temperature. tert-Butyldimethylsilyl chloride (345.6 mg, 2.3 mmol) was added, the reaction mixture reacted for 4 h at 60 ℃. The reaction solvent was evaporated *in vacuo* and the residue was purified by MPLC with mixture of chloroform in ethyl acetate (25%, v/v). The product was giving as white foam (260 mg, 90%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.19 (s, 1H), 8.38 (s, 1H), 8.01 (s, 1H), 5.82 (d, *J* = 1.2 Hz, 1H), 5.09 (dd, *J* = 2.0 Hz, 1H), 4.52 (dd, *J* = 6.0 Hz, 1H), 4.42 (dd, *J* = 4.8 Hz, 1H), 4.06 (dt, *J* = 5.2 Hz, 1H), 3.93 (t, *J* = 6.8 Hz, 1H), 3.18-3.14 (m, 7H), 2.95-2.87 (m, 6H), 1.68 (s, 6H), 1.06-0.86 (m, 45H), 0.08-0.04 (m, 12H); HRMS (ESI) for $C_{28}H_{48}O_5N_6F_3Si_2$ $[M+H]^+$: Calcd. 661.3099; Found. 661.3160.

***N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (6)**

Compound 5 (330 mg, 0.5 mmol) was dissolved in 2 mL dichloromethane, and 54.6 μL hydrofluoric acid-pyridine solution (70% hydrofluoric acid, 30% pyridine) in 0.34 mL pyridine was added for getting the mixture, stirred at 0 ℃ for 2 h. The reaction solution was extracted by dichloromethane and combined organic layer evaporated *in vacuo* for giving crude product 260 mg. No further purification, the crude product (260 mg) and 4,4'-dimethoxytrityl chloride (241.6 mg, 0.715 mmol) were dissolved in 2 mL anhydrous pyridine and the mixture stirred for 4 h at room temperature. The solvent was concentrated *in vacuo* and the residue was purified by MPLC with mixture of dichloromethane in ethyl acetate (25%, v/v). The product was giving

as white foam (200 mg, 60%).^{1}H NMR (400 MHz, DMSO-d_6) δ 9.05 (s, 1H), 8.61 (dt, *J* = 12 Hz, 1H), 8.20 (s, 1H), 7.41-7.14 (m, 8H), 6.77-6.73 (m, 4H), 5.81 (d, *J* = 4.0 Hz, 1H), 5.14 (dd, *J* = 4.0 Hz, 1H), 4.54 (dd, *J* = 6.4 Hz, 1H), 4.02 (dd, *J* = 12.4 Hz, 1H), 3.76 (s, 6H), 3.44-3.40 (m, 2H), 2.99 (s, 3H), 2.63 (s, 3H), 2.04 (s, 1H), 0.86 (s, 10H), 0.118 (s, 3H), 0.008 (s, 3H); HRMS (ESI) for $C_{41}H_{49}O_7N_6F_3SiNa$ $[M+Na]^+$: Calcd.845.3384; Found. 845.3332.

3'-*O*-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-*N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (7)

Compound 6 (1000 mg, 1.21 mmol) co-evaporated with 5 mL anhydrous acetonitrile by three times and dissolved in 8 mL anhydrous dichloromethane. Diisopropylethylamine (0.78 mL, 4.5 mmol) and 1-methylimidazole (0.09 mL, 1.125 mmol) were added. After 5 minutes, 2-cyanoethyl-N,N-diisopropylamidochlorophosphoramidite (0.75 mL, 3.63 mmol) also was added, the mixture was reacted for 1.5 h at room temperature. The reaction mixture was extracted by dichloromethane and organic layer concentrated *in vacuo*, the residue was purified by MPLC with mixture of ethyl acetate in dichloromethane (25%, v/v). The product was giving as white foam (803 mg, 67%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.76 (s, 1H), 8.03 (s, 1H), 7.92 (s, 1H), 7.48-7.17 (m, 12H), 6.78-6.72 (m, 5H), 5.91-5.88 (m, 1H), 5.30 (m, 1H), 4.58-4.31 (m, 2H), 4.11 (s, 1H), 3.76-3.47 (m, 16H), 2.98-2.56 (m, 5H), 2.31 (s, 1H), 2.10-2.02 (m, 2H), 1.29-0.73 (m, 30H), 0.08-0.00 (m, 11H); ^{31}P NMR (161 MHz, DMSO-d_6) δ 150.86, 148.17; HRMS (ESI) for $C_{50}H_{66}O_8N_8F_3Na$ $[M+Na]^+$: Calcd. 1045.4363; Found. 1045.4348. HRMS (ESI) for $C_{56}H_{82}O_8N_9F_3$ $[M+TEA+H]^+$: Calcd. 1124.5787; Found. 1124.5729.

Organic compounds and oligonucleotide characterization
HRMS and NMR spectrum of necessarily synthetic compounds in this study

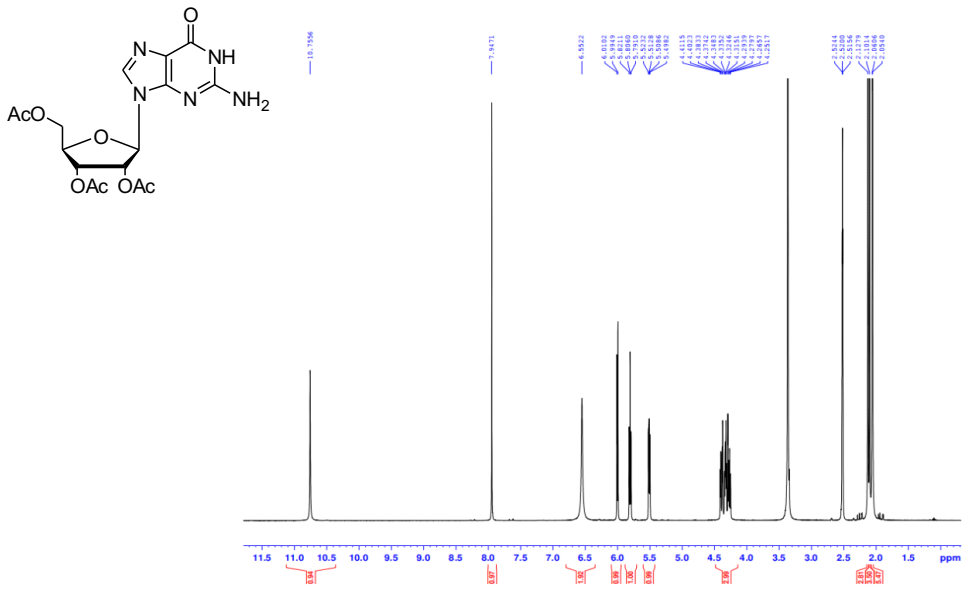

Figure S27. ^{1}H NMR spectrum of 2',3',5'-tri-*O*-acetylguanosine (**1**).

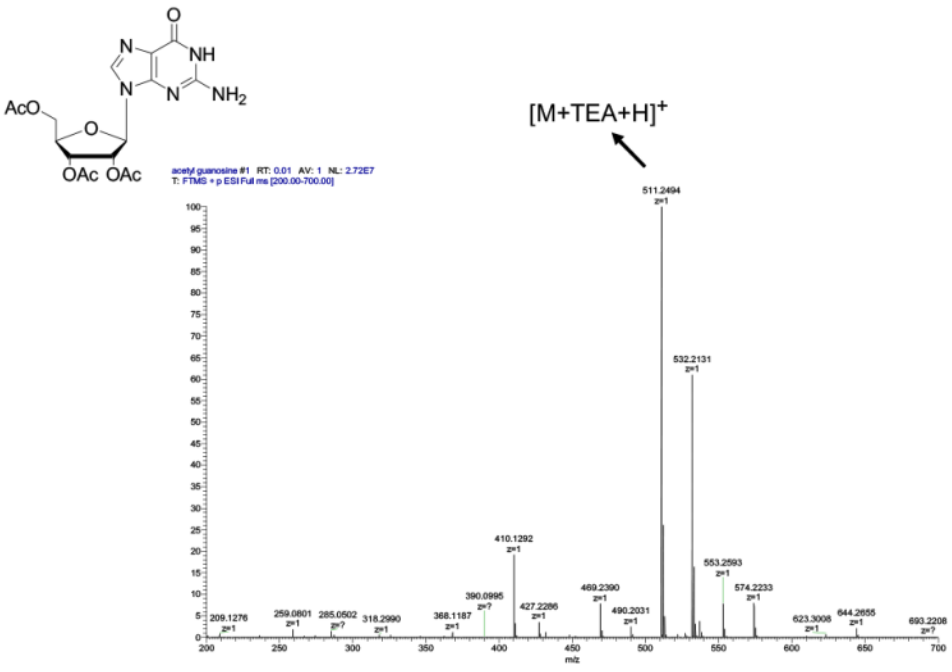

Figure S28. HRMS spectrum of 2',3',5'-tri-*O*-acetylguanosine (**1**).

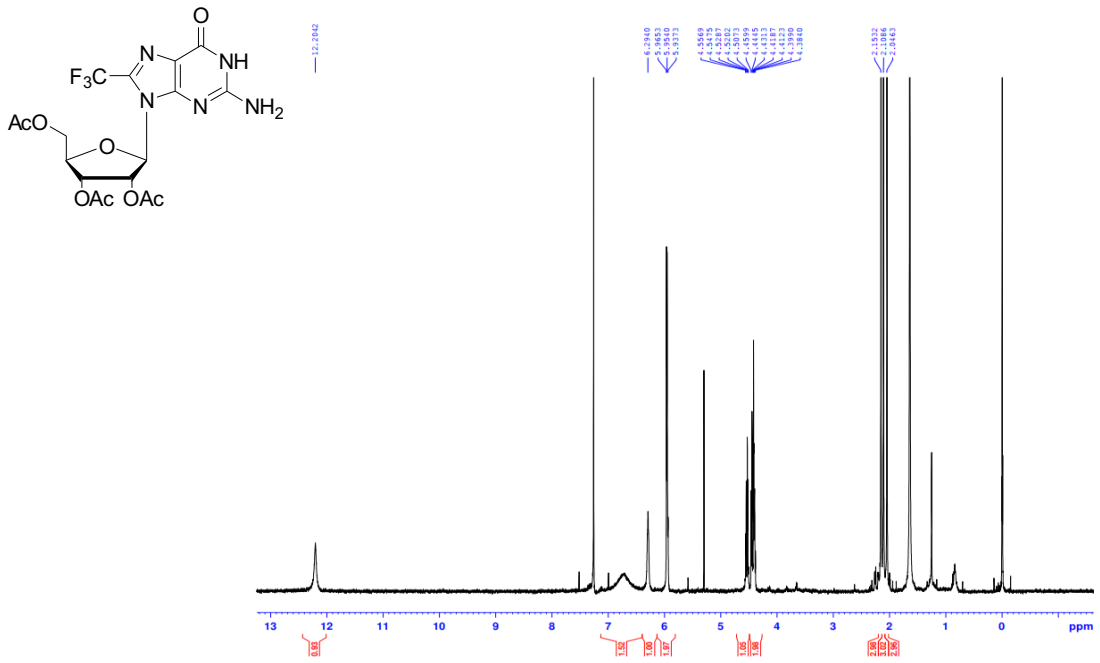

Figure S29. ^{1}H NMR spectrum of 2',3',5'-tri-*O*-acetyl-8-trifluoromethylguanosine (**2**).

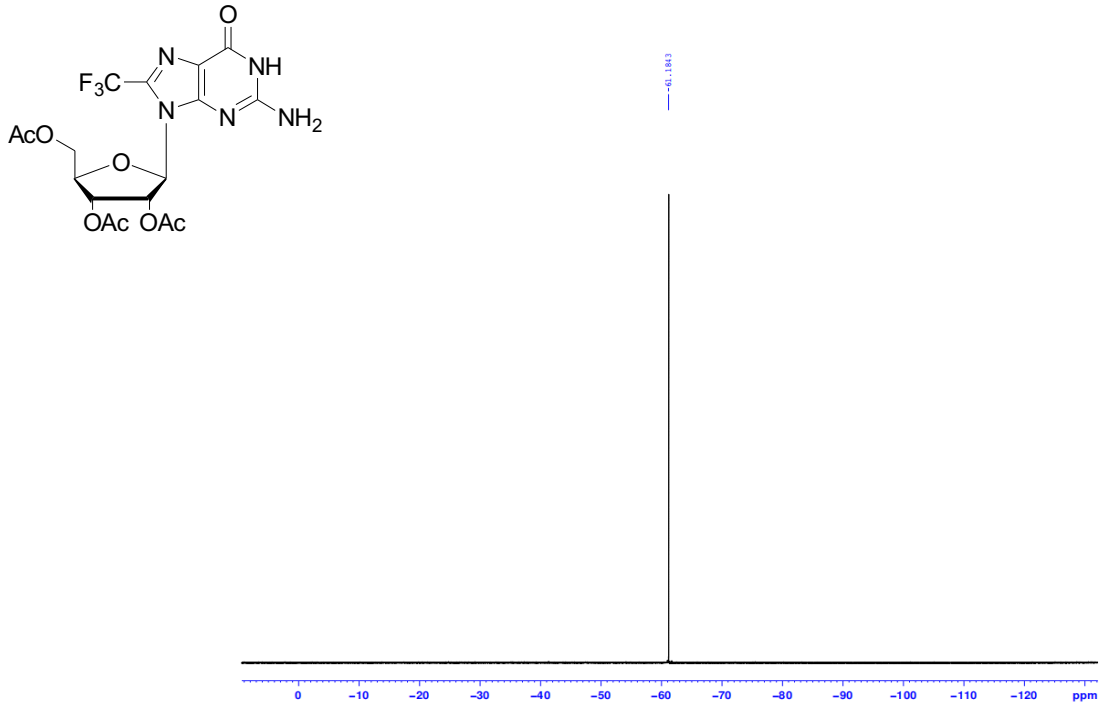

Figure S30. ^{19}F NMR spectrum of 2',3',5'-tri-*O*-acetyl-8-trifluoromethylguanosine (**2**).

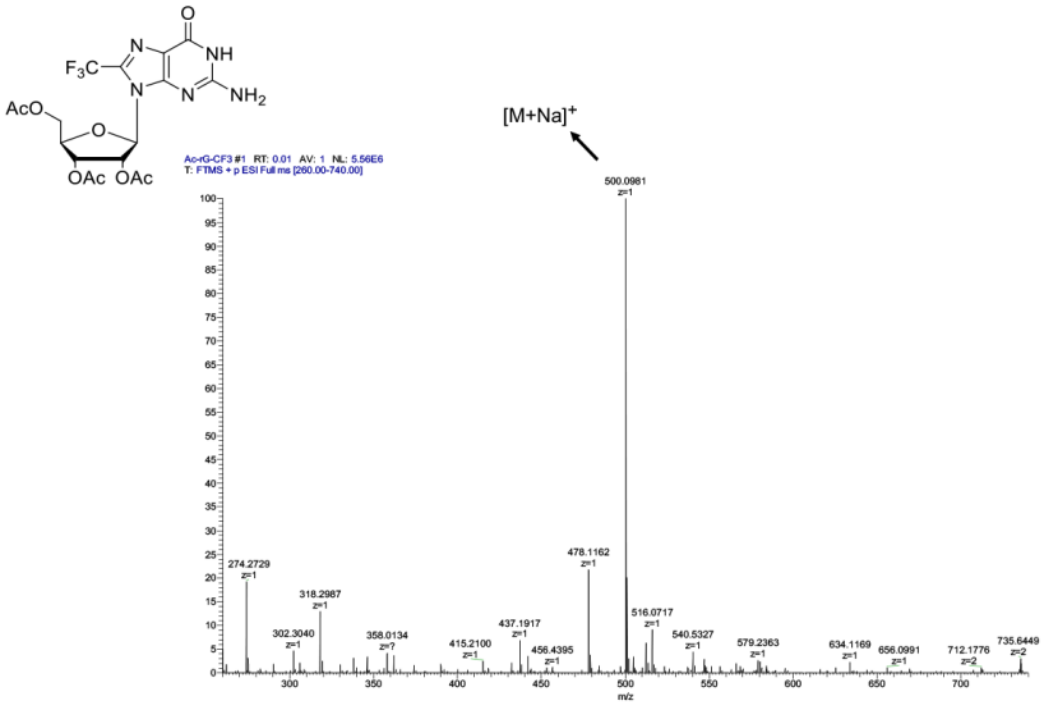

Figure S31. HRMS spectrum of 2',3',5'-tri-*O*-acetyl-8-trifluoromethylguanosine (**2**).

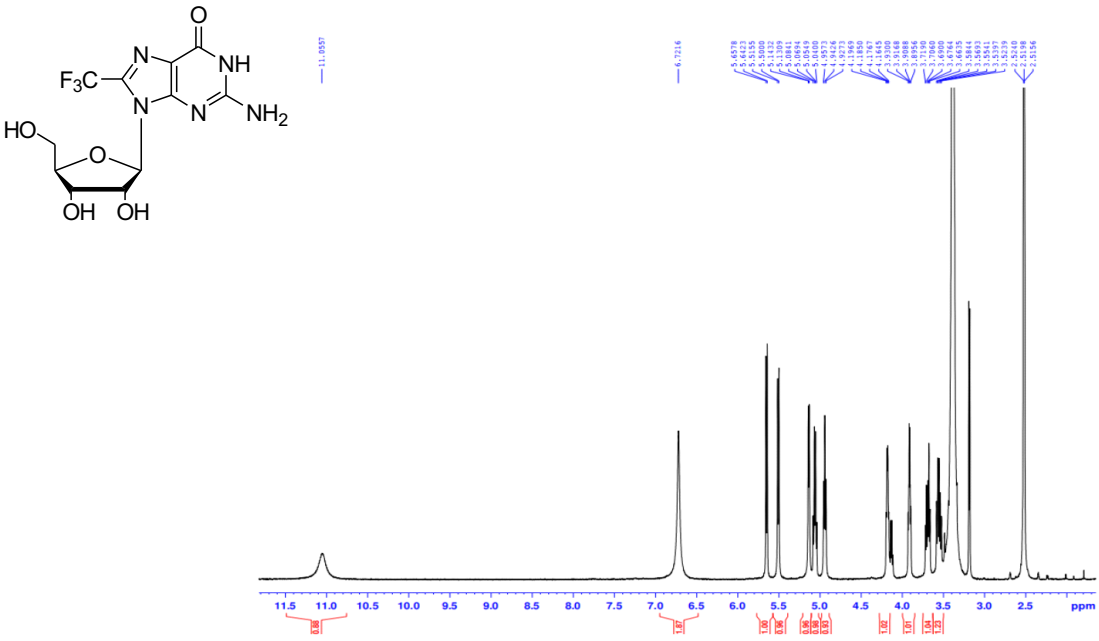

Figure S32. ^{1}H NMR spectrum of 8-trifluoromethylguanosine (**3**).

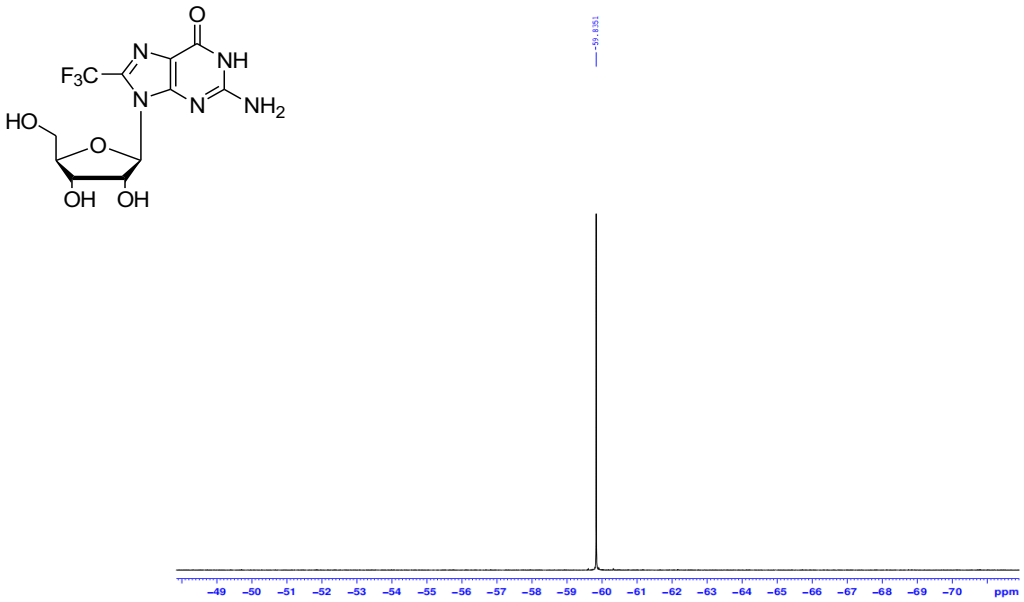

Figure S33. ^{19}F NMR spectrum of 8-trifluoromethylguanosine (**3**).

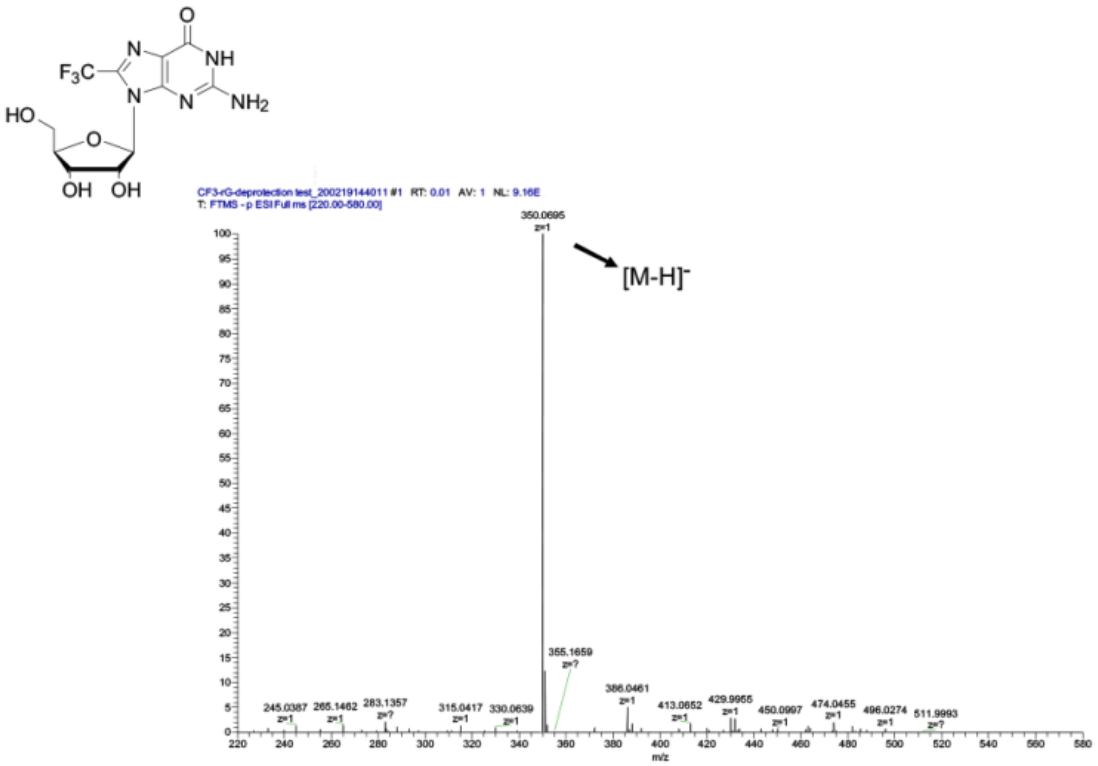

Figure S34. HRMS spectrum of 8-trifluoromethylguanosine (**3**).

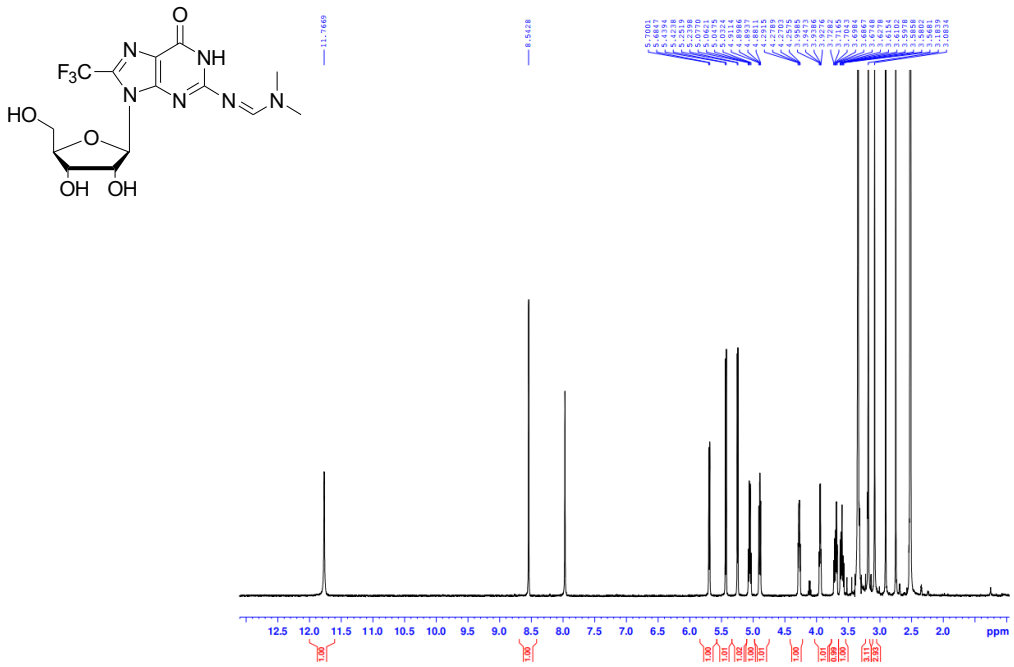

Figure S35. ^{1}H NMR spectrum of *N2*-dimethylformamidyl-8-trifluoromethylguanosine (**4**).

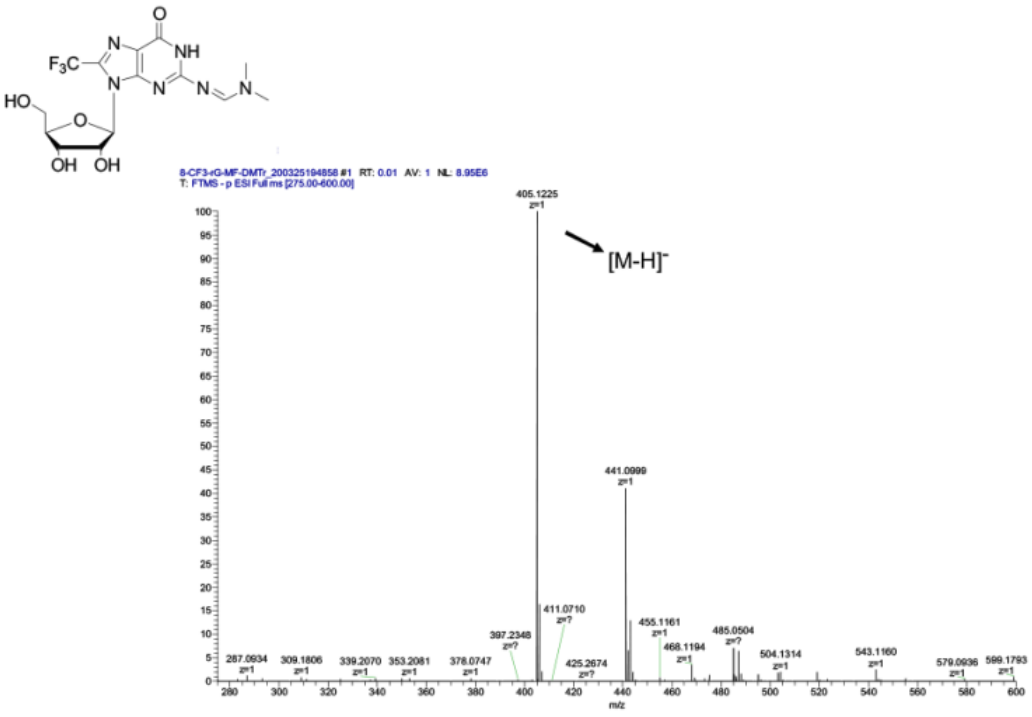

Figure S36. HRMS spectrum of *N2*-dimethylformamidyl-8-trifluoromethylguanosine (**4**).

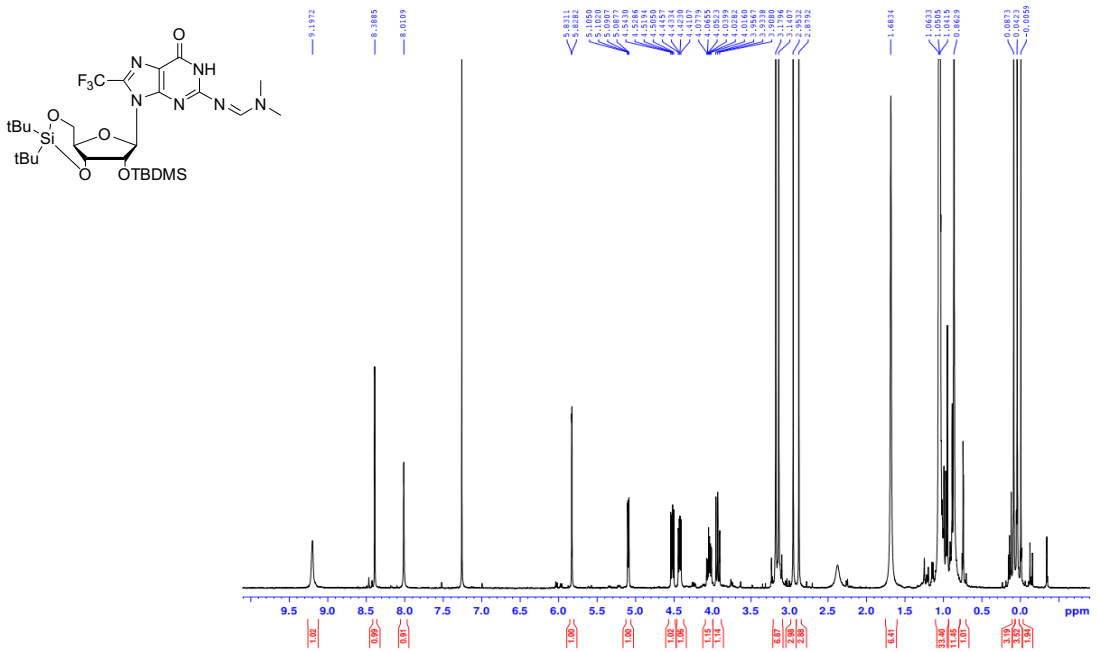

Figure S37. ^{1}H NMR spectrum of *N2*-dimethylformamidyl-8-trifluoromethyl-2'-*O*-(tert-butyldimethylsilyl)-3',5'-*O*-(di-tert-butylsilylene)guanosine (**5**).

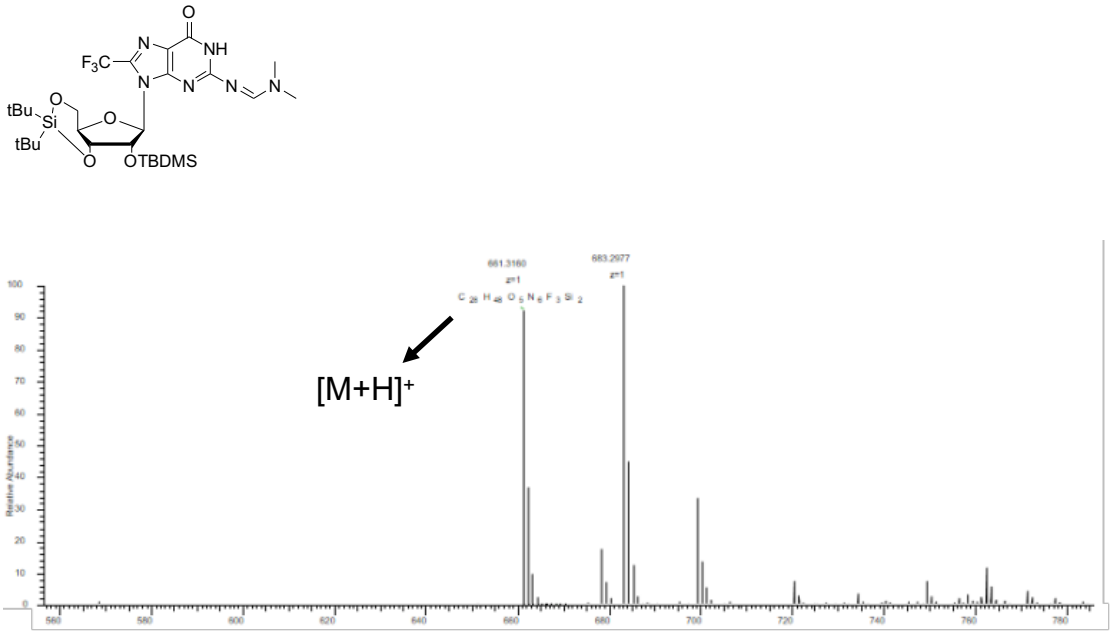

Figure S38. HRMS spectrum of *N2*-dimethylformamidyl-8-trifluoromethyl-2'-*O*-(tert-butyldimethylsilyl)-3',5'-*O*-(di-tert-butylsilylene)guanosine (**5**).

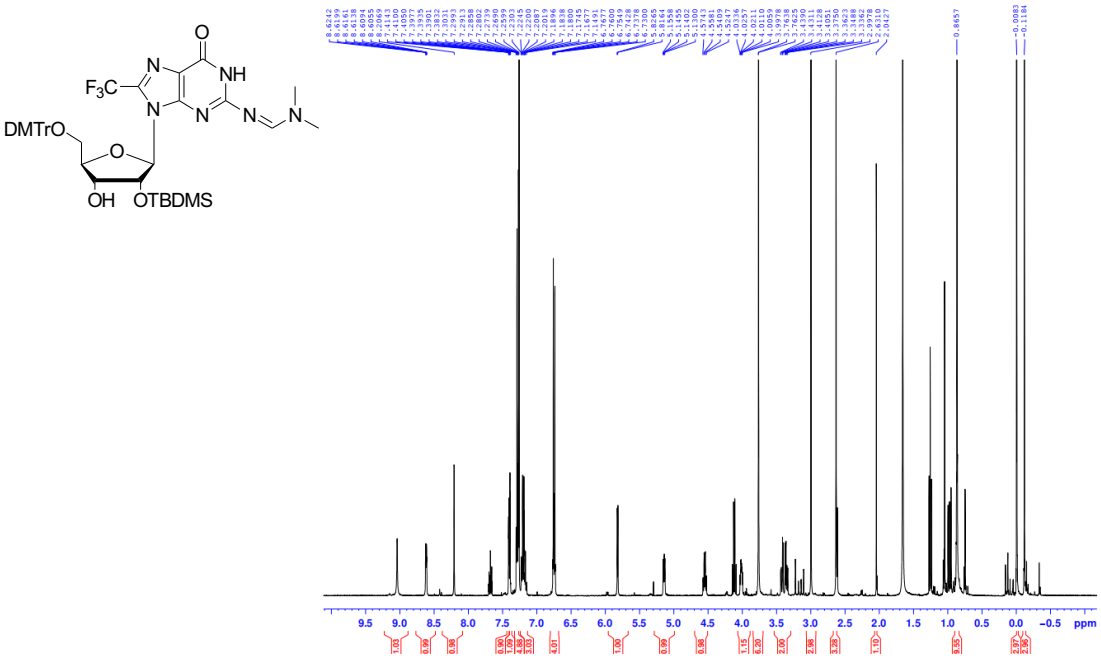

Figure S39. [1]H NMR spectrum of *N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (**6**).

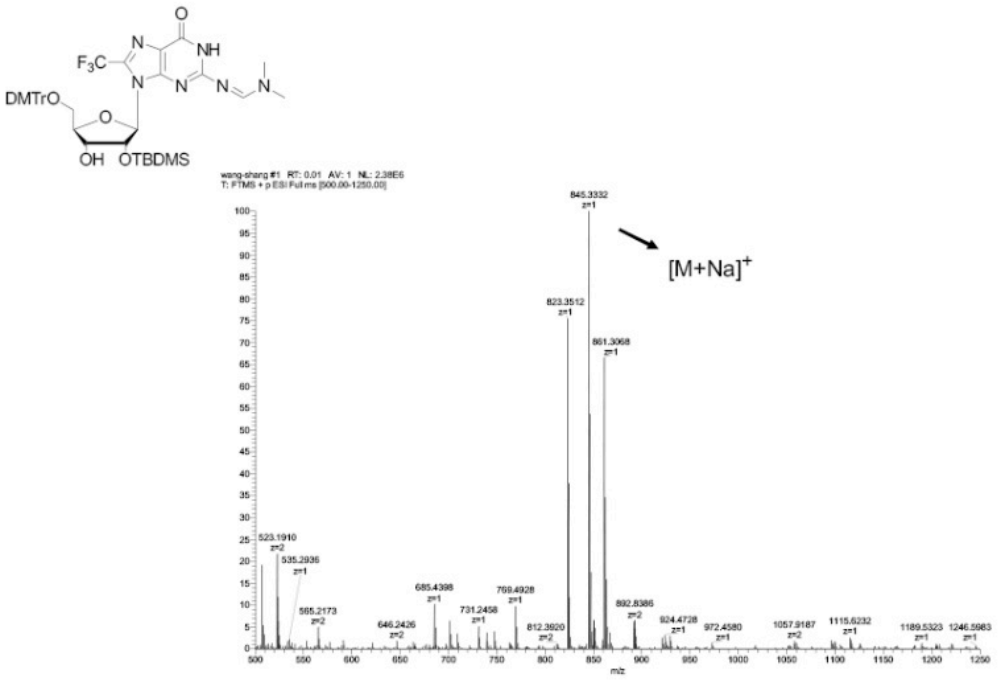

Figure S40. HRMS spectrum of *N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (**6**).

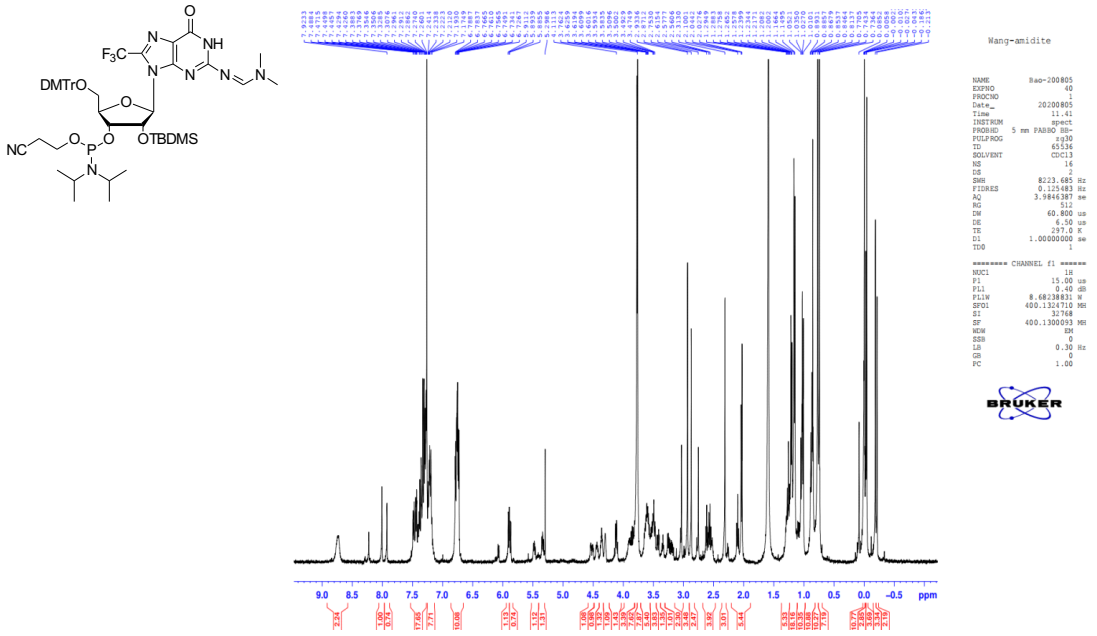

Figure S41. ^{1}H NMR spectrum of 3'-*O*-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-*N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (**7**).

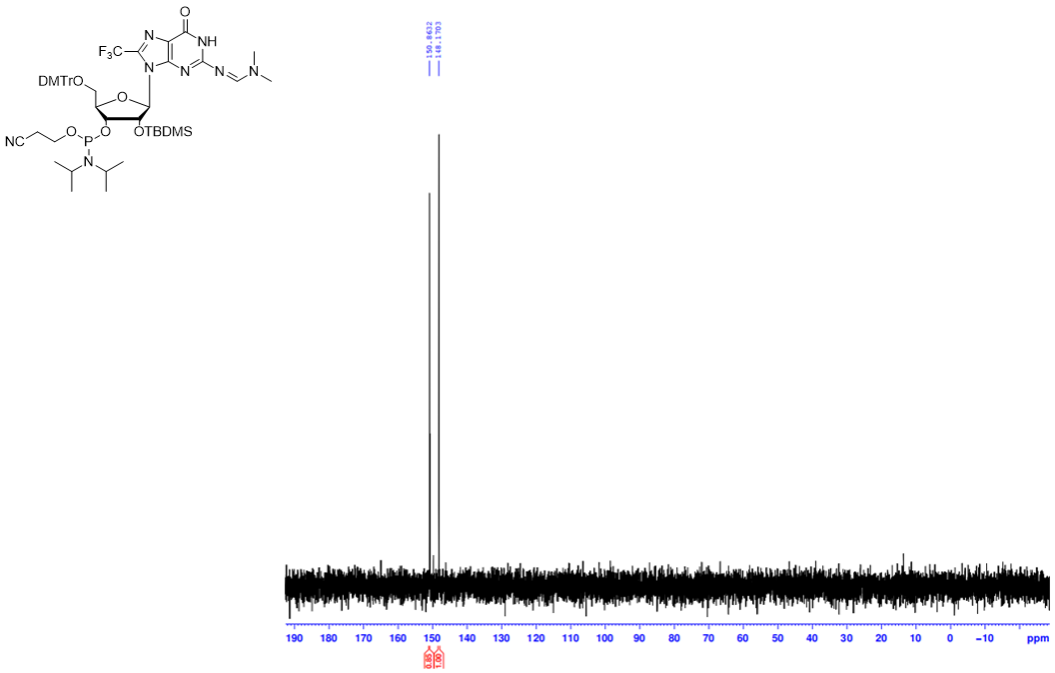

Figure S42. ^{31}P NMR spectrum of 3'-*O*-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-*N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (**7**).

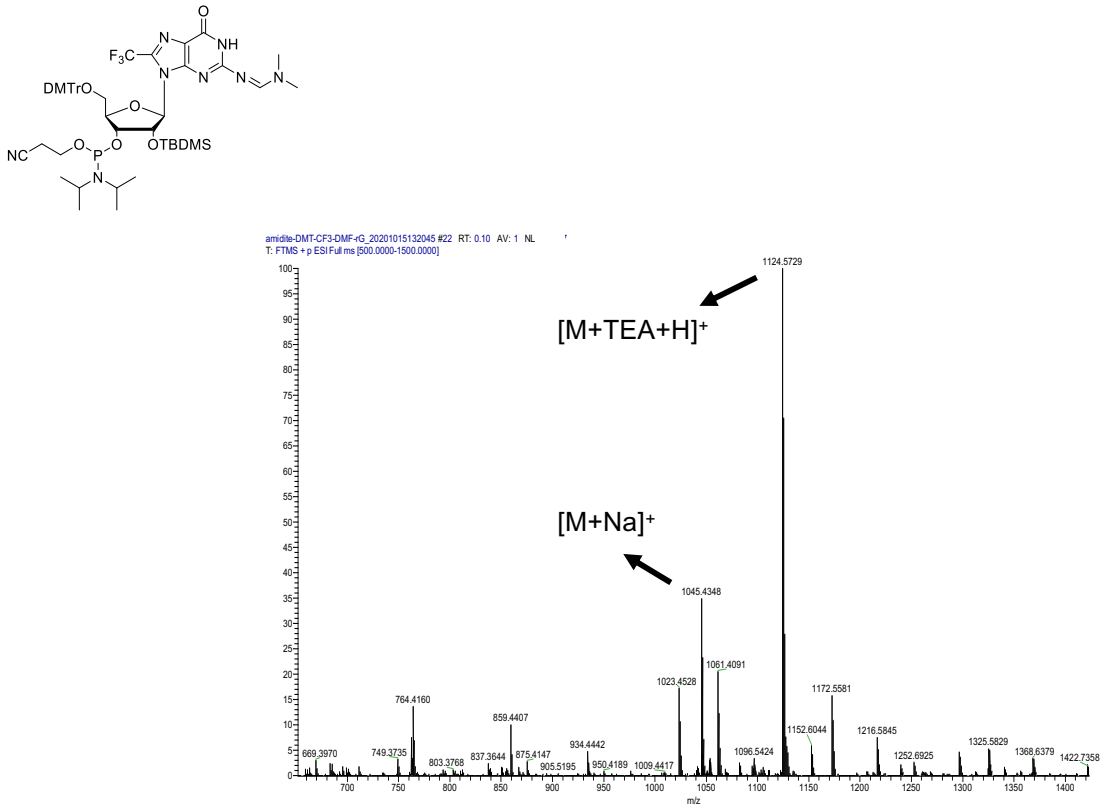

Figure S43. HRMS spectrum of 3'-*O*-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-*N2*-dimethylformamidyl-8-trifluoromethyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-tert-butyldimethylsilylguanosine (**7**).

MALDI-MS and ESI-MS spectrum of RNAs used in this study

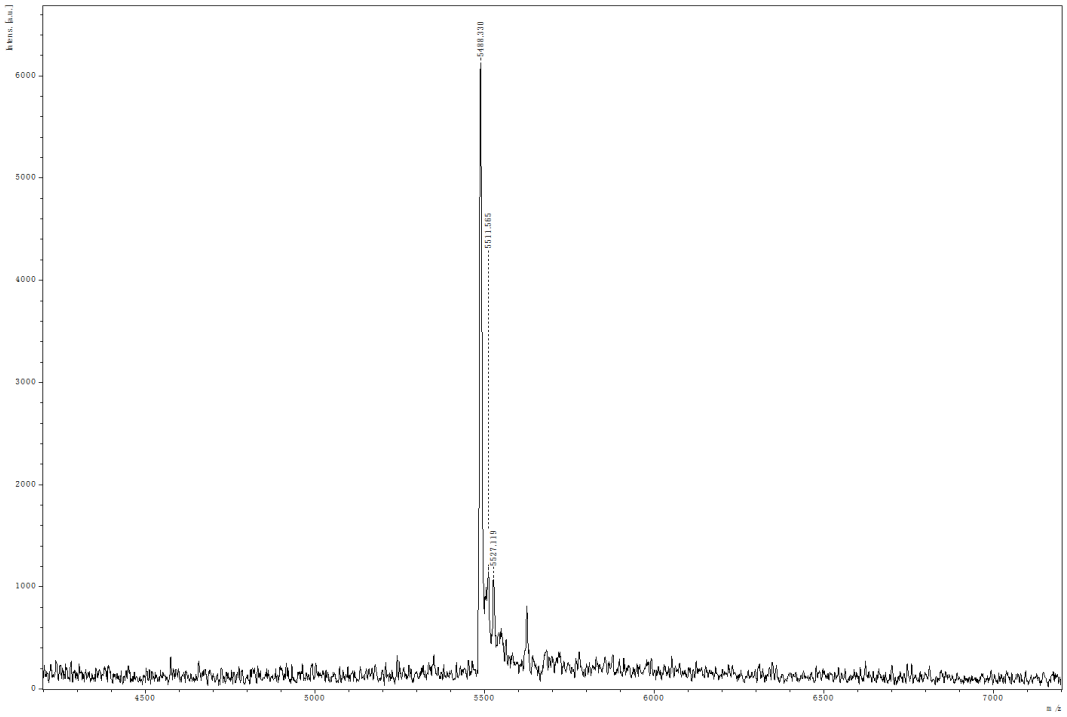

Figure S44. MALDI TOF MS of 5'-UGGCUGGCAAUGGCGGU-3'.

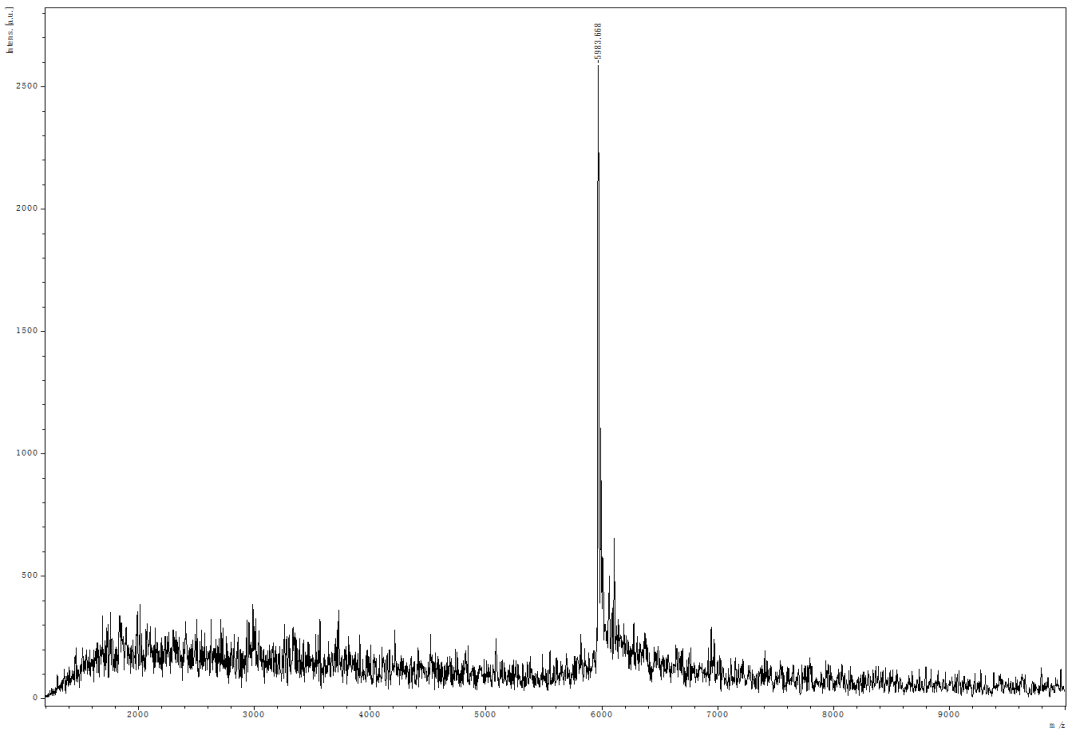

Figure S45. MALDI TOF MS of Cy3-5'-UGGCUGGCAAUGGCGGU-3'.

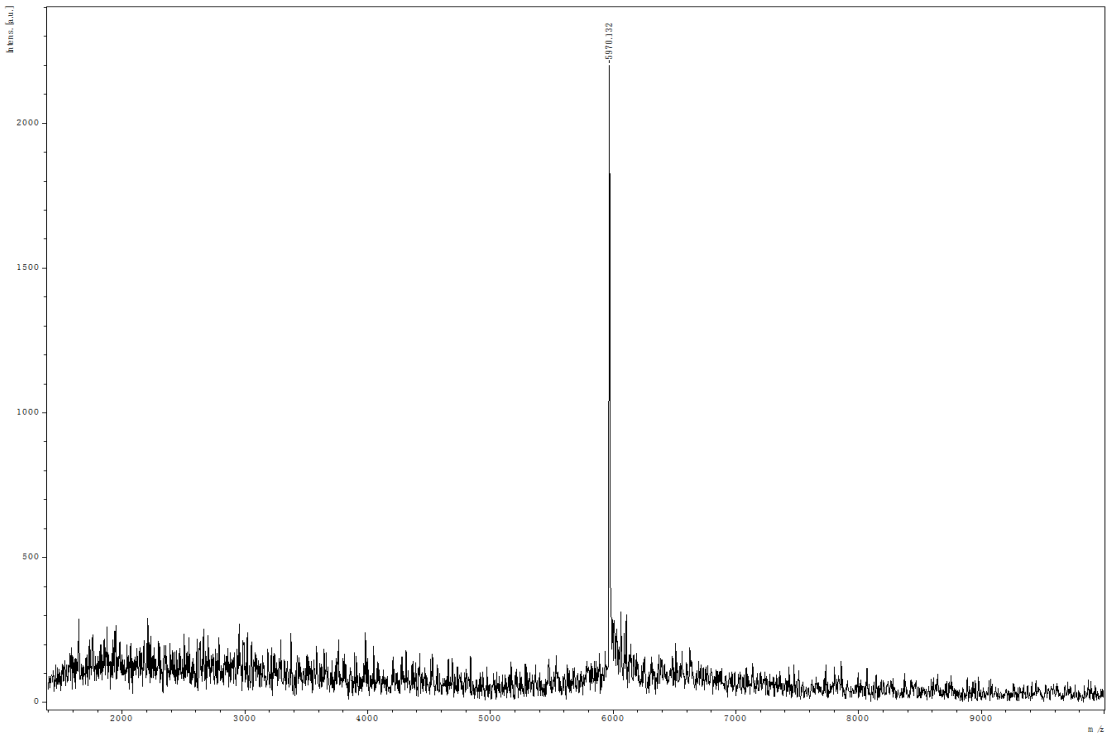

Figure S46. MALDI TOF MS of Cy3-5'-UGUCUGUCAAUGGCAGC-3'.

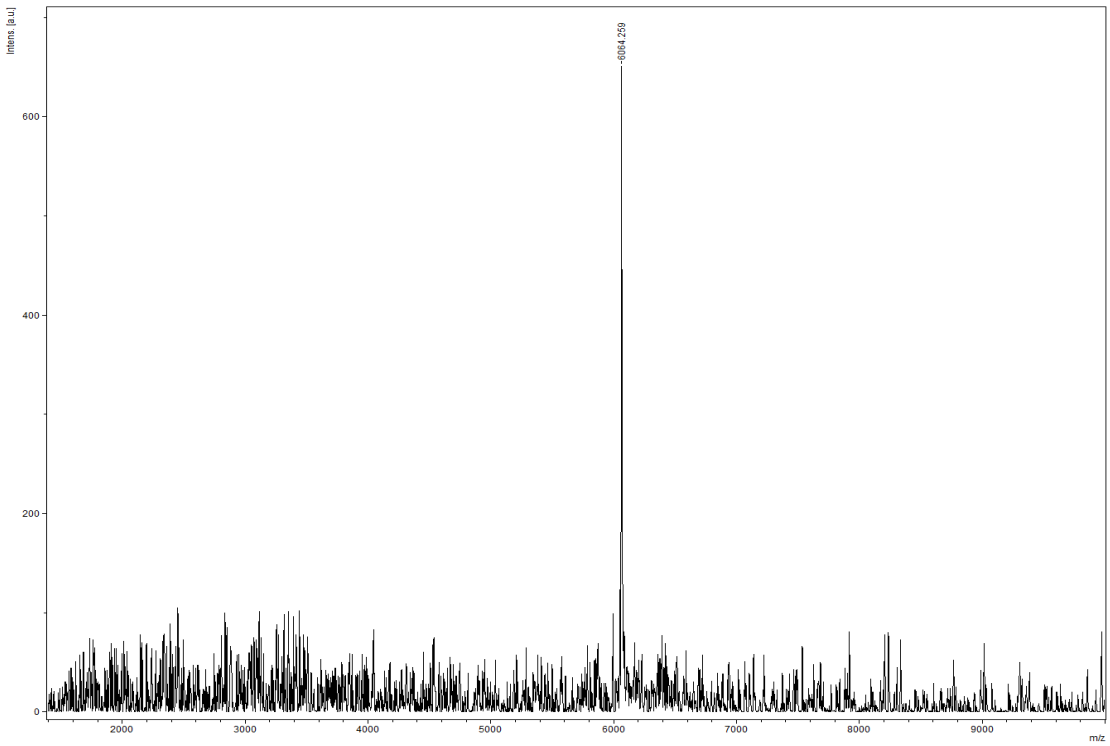

Figure S47. MALDI TOF MS of Cy5-5'-UGGCUGGCAAUGGCGGU-3'.

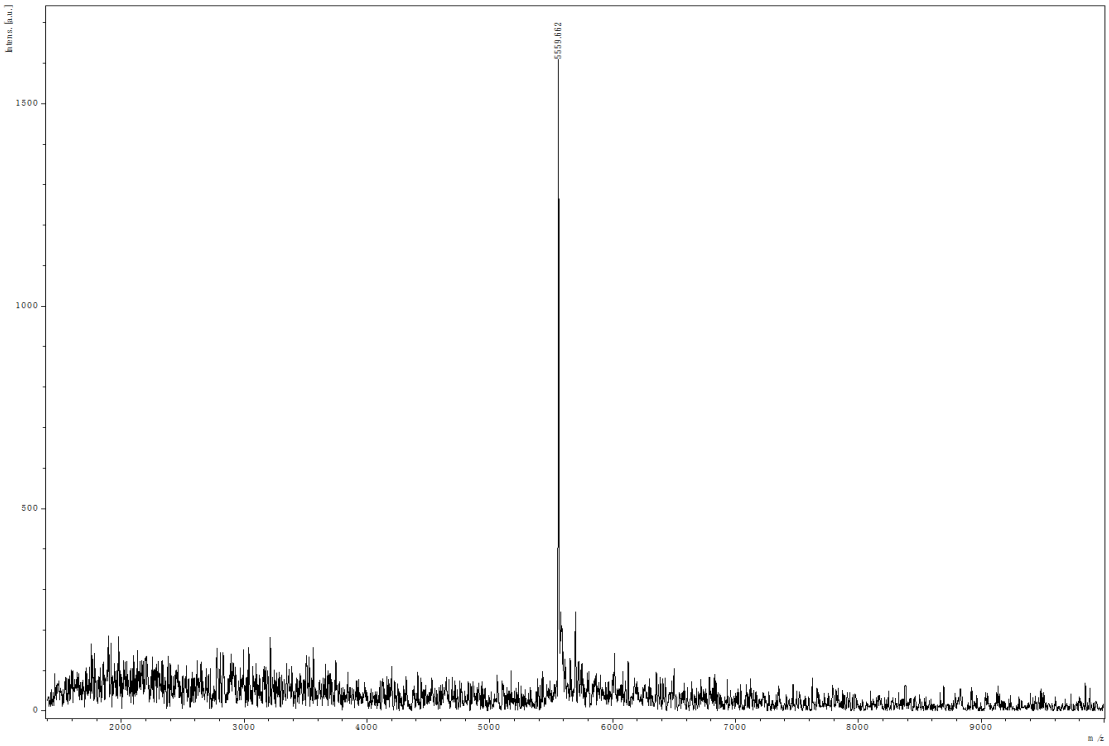

Figure S48. MALDI TOF MS of 5'-UGGCUGGCAAUGGCG^{F}GU-3'.

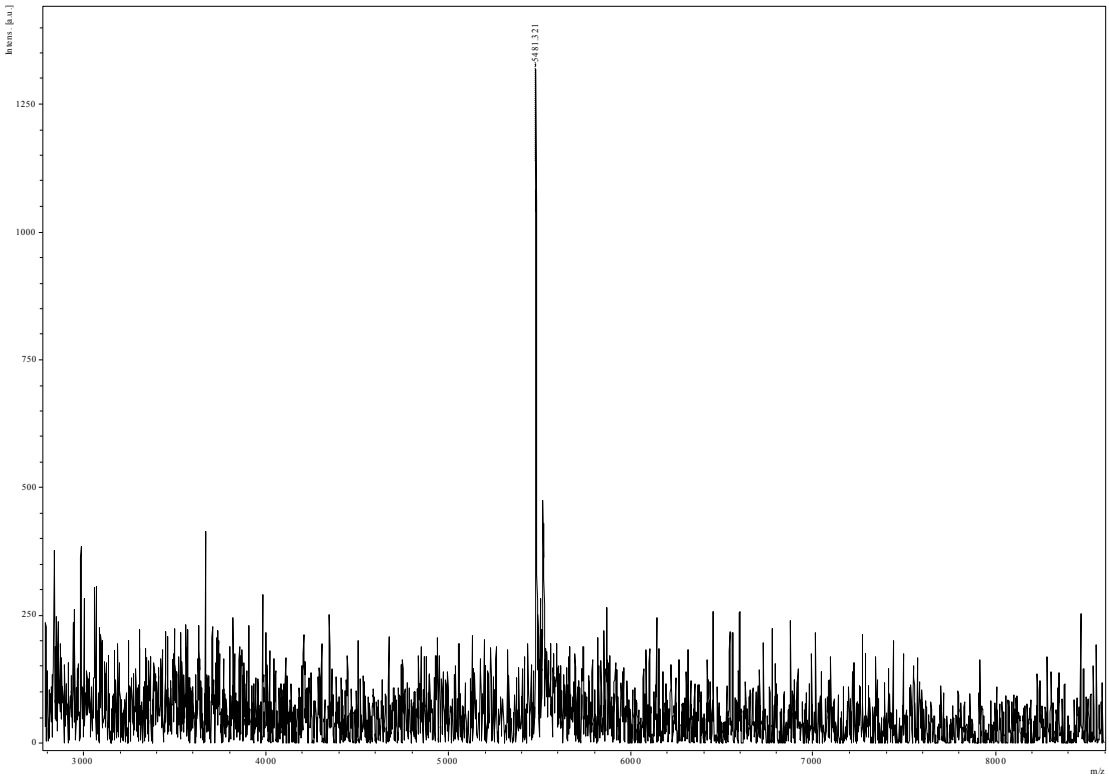

Figure S49. MALDI TOF MS of 5'-UGGCUGGCAAUGGCGAU-3'.

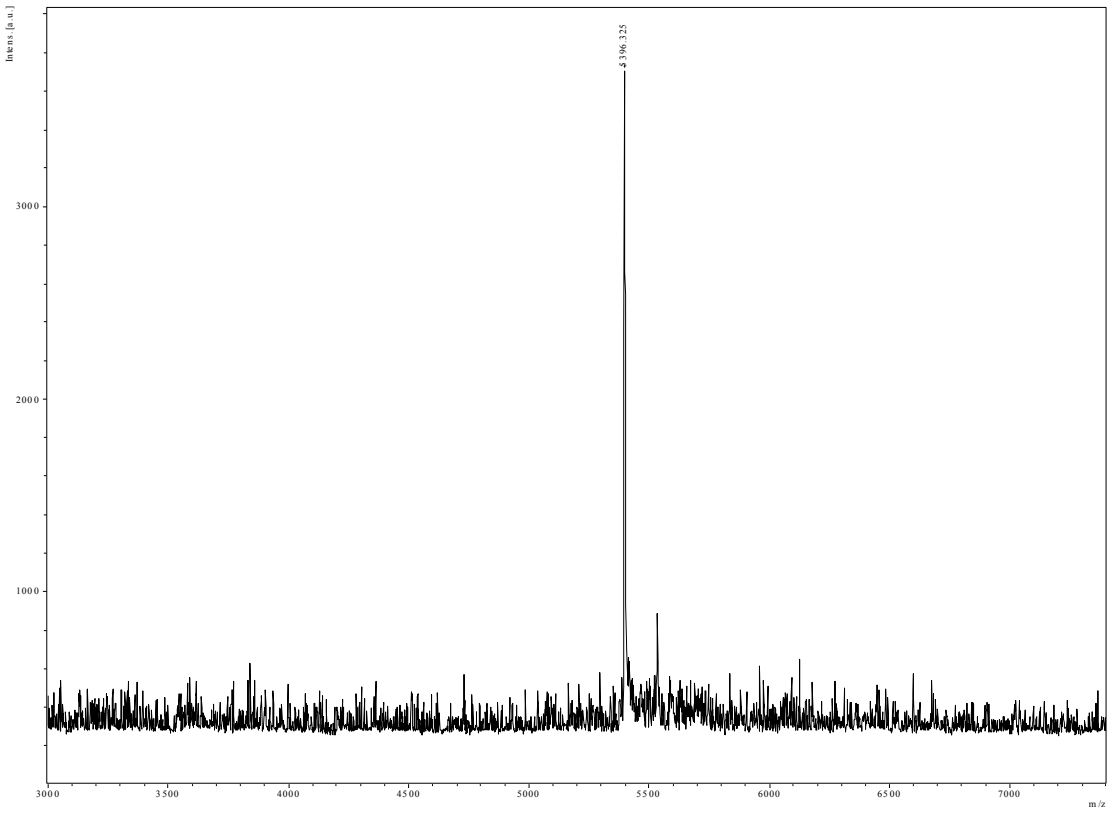

Figure S50. MALDI TOF MS of 5'-UAACUAACAAUAGCAGU-3'.

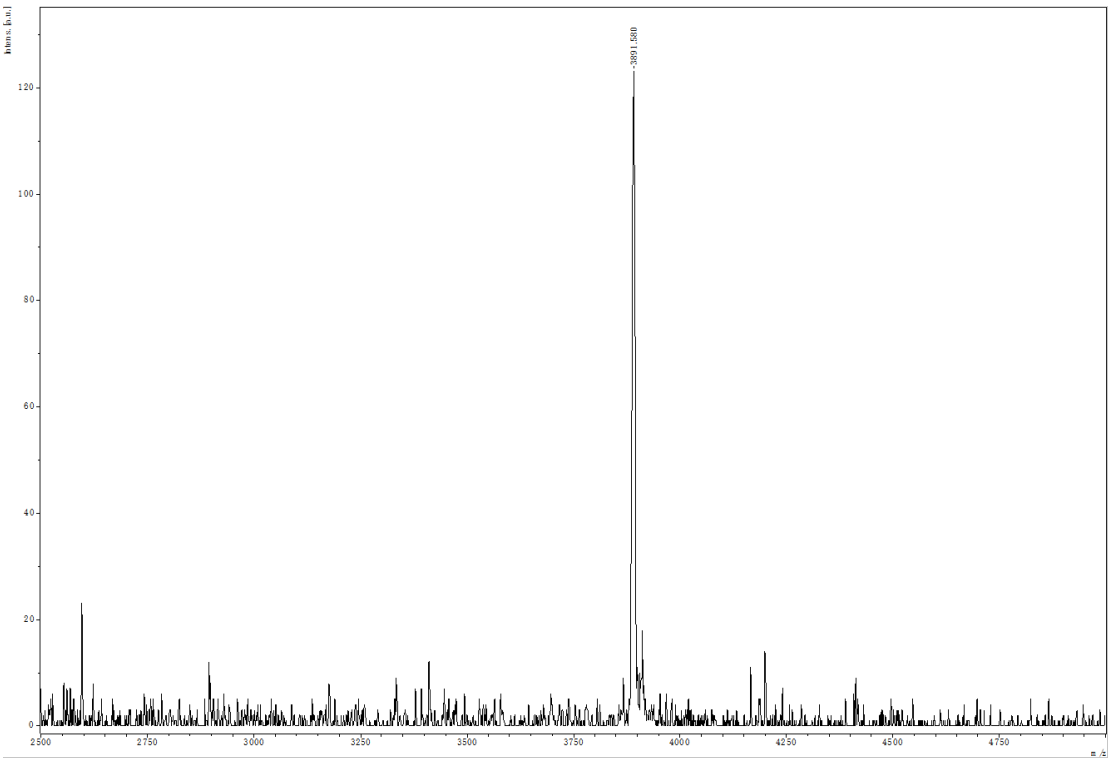

Figure S51. MALDI TOF MS of 5'-UAGGGUUAGGGU-3'.

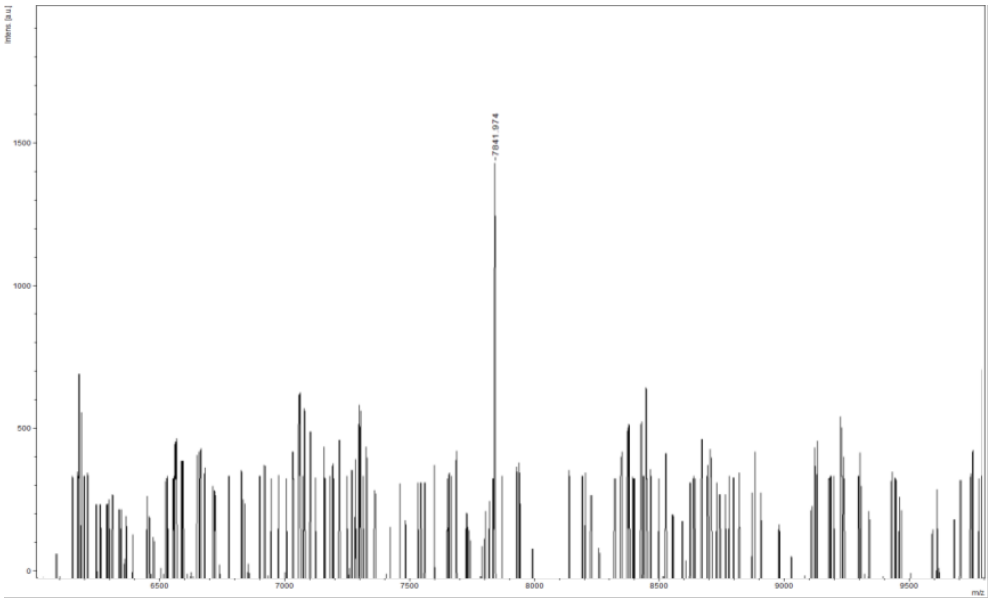

Figure S52. MALDI TOF MS of 5'-UAGGGUUAGGGUUAGGGUUAGGGU-3'.

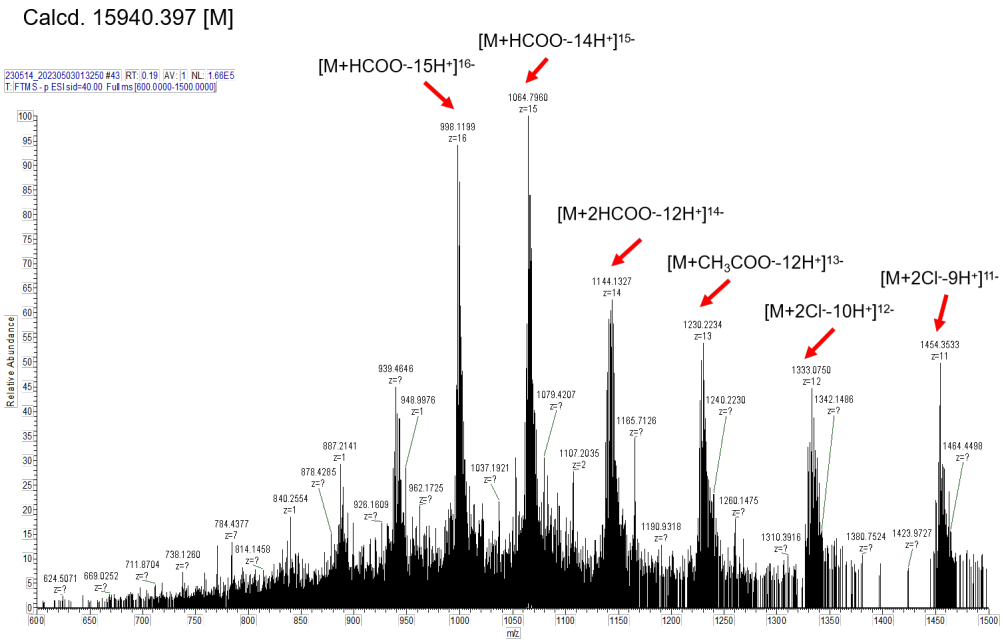

Figure S53. ESI-MS of 5'-ACUUCUCCUGCUAGAAUGGCUGGCAAUGGCGGUUGAGCUGCUCUUGCUUU-3'.

References

1. (a) H. L. Bao, T. Masuzawa, T. Oyoshi and Y. Xu, *Nucleic Acids Res.*, 2020, **48**, 7041-7051. (b) H. L. Bao, H. S. Liu and Y. Xu, *Nucleic Acids Res.*, 2019, **47**, 4940-4947. (c) H. L. Bao, T. Ishizuka, T. Sakamoto, K. Fujimoto, T. Uechi, N. Kenmochi and Y. Xu, *Nucleic Acids Res.*, 2017, **45**, 5501-5511. (d) H. L. Bao and Y. Xu, Nat. Protoc., 2018, **13**, 652-665.