Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Targeting RNA G-Quadruplex in SARS-CoV-2: A Promising Therapeutic Target for COVID-19?

*Chuanqi Zhao[+], Geng Qin[+], Jingsheng Niu, Zhao Wang, Chunyu Wang, Jinsong Ren, and Xiaogang Qu**

anie_202011419_sm_miscellaneous_information.pdf

Supporting information

Materials

RNAs were synthesized by Shanghai Sangon Biological Engineering Technology & Services (Shanghai, China). Concentrations of the oligomers were determined by measuring the absorbance at 260 nm after melting. All DNA samples were heated at 95 °C for 5 min, and then slowly cooled to room temperature unless otherwise indicated. Chemicals were purchased from Sigma–Aldrich and used without further purification. All water used to prepare buffer solutions was obtained by using a Milli-Q water system.

Methods

Circular dichroism (CD) measurements. CD spectra and CD melting experiments were carried out on a JASCO J-810 spectropolarimeter equipped with a temperature-controlled water bath. The optical chamber of CD spectrometer was deoxygenated with dry purified nitrogen (99.99%) for 45 min before use, and kept the nitrogen atmosphere during experiments. Three scans were accumulated and automatically averaged. In CD melting experiments, signal were collected at a heating rate of 1 °C min^{-1}. RNA samples were obtained by incubation at 4℃ for 24 h. **Thermodynamic parameters obtained by melting analysis:** In order to calculate thermodynamic parameters, including the enthalpy change (ΔH^o), the entropy change (ΔS^o), and the free energy change (ΔG^o) for intramolecular RNA G-quadruplex formation, the melting curves were fit to the theoretical equation. ΔH^o was the slope of the ln Ka versus $1/T$ plot according to the equation $\ln K\mathrm{a} = -(\Delta H^o), /\mathrm{R}T) + \Delta S^o /\mathrm{R}$, where ΔS^o was the entropy change that was calculated according to the y axis intercept. The free energychange (ΔG^o_{25}) was calculated from the standard Gibbs's

equation, $\Delta G^{o}_{25} = \Delta H^{o} - T\Delta S^{o}$.

Fluorescence assays. The assays were carried out on a JASCO FP-6500 spectrofluorometer at 25 °C. For NMM treated samples, emission spectra were measured by using an excitation wavelength of 399 nm. The concentration of NMM was fixed at 0.6 μM in strand. For ThT treated samples, emission spectra were measured by using an excitation wavelength of 442 nm. The concentration of ThT was fixed at 0.6 μM in strand. The concentration of DNA was fixed at 0.3 μM in strand. RNA samples were obtained by incubation at 4℃ for 24 h.

Stopped-flow experiments: Fluorescence stopped-flow experiments were carried out by using an SX20 Stopped-Flow Spectrometer. In these assays, NMM treated samples (NMM/F-RG-1 or NMM/F-RG-1-Mut in 10 mM Tris-HCl buffer, 0 mM KCl, pH = 7.2) were mixed with K^{+}-containing buffer (10 mM Tris-HCl buffer, 200 mM KCl, pH = 7.2). Fluorescence changes were monitored at a wavelength of 661 nm with an excitation wavelength of 395 nm.

Nondenaturing polyacrylamide gel electrophoresis experiments: Native gel electrophoresis was carried out on acrylamide gel (15 %) and run at 25 °C, 1 × TB buffer containing 10 mM KCl and was silver stained. To prepare the final loading sample, the PDP were added to the annealed DNA samples and incubated at 4 °C for 12 h.

NMR spectroscopy: ^{1}H NMR spectra were carried out on a Bruker Avance 600 MHz NMR Spectrometer. The RNA samples were dissolved in 10 mM phosphate-buffered saline (PBS) (pH 7.0) containing 100 mM KCl and 10% D_2O at a final concentration of 0.3 mM in strand.

Cells culture: The human cervical carcinoma cell HeLa and the human embryonic kidney cell line HEK293T were a generous gift from Dr. Qian Jin (Beijing Institute of Lifeomics, Beijing, China). All these cell lines were maintained in high-glucose Dulbecco's modified Eagle's medium (DMEM; HyClone, USA) supplemented with 10% fetal bovine serum (FBS; HyClone, USA), 100 U/mL penicillin, and 100 μg/mL streptomycin. The cells were incubated at 37 °C in a humidified incubator containing 5% CO_2.

Vectors construction and transfection: A reporter vector with the pLV-EGFP-N backbone (Inovogen Tech. Co., Beijing, China) was generated to encode the 15 base pairs (bp) RG-1 (position at 28,903-28,917 bp of SARS-CoV genome), and was named as pLV-RG-1-WT-EGFP. By using the Stratagene QuikChange Site-Directed Mutagenesis kit (Agilent Technologies, Palo Alto, CA, USA), we point mutated the RG-1-WT, which was named as pLV-RG-1-Mut-EGFP.

The cDNAs encoding the N protein were PCR-amplified and subcloned into pCAG-Flag vector, and were named as pCAG-N-WT-Flag. And the pCAG-N-WT-Flag is a generous gift from Prof. Peihui Wang (Cheeloo College of Medicine, Shandong University, Jinan). By using the Stratagene QuikChange Site-Directed Mutagenesis kit, we point mutated the pCAG-N-WT-Flag in the RG-1 site, which was named as pCAG-N-Mut-Flag.

For viral packaging and infection were performed as previously described. Briefly, the vectors along with viral packaging plasmids (PMD and SPA) were transfected into human embryonic kidney 293T cells using Lipofectamine 2000 (Invitrogen, USA), following the manufacturer's instructions. Virus supernatant was harvested after 48 hr, filtered through a 0.45 μM filter, and incubated on target cells for 6 hr at a 1:10 dilution with 8 μg/mL polybrene. Infected cells were selected with 200 μg/mL puromycin for 2 weeks before evaluation. For transfection, the cells were transfected with the indicated vectors using Lipofectamine 2000 (Invitrogen, CA, USA) following the manufacturer's protocol. The transfected cells were harvested after 48 to 72 h.

Immunofluorescence assays: Cells were washed twice in 1 × PBS, fixed with 4% paraformaldehyde for 10 minutes, permeabilized with 0.5% Triton X-100-PBS for 15 minutes, and then blocked for 30 minutes with 3% BSA-0.2% Triton X-100-PBS. Furher, cells were incubated with primary antibody for 16 hours at 4 °C. Fluorescent secondary antibodies were used in a 1:1,000 dilution for 45 minutes at 37 °C. DAPI was added for 2 minutes in the dark before taken images with confocal laser-scanning microscope.

Western blotting assays: For protein analysis of whole-cell lysates, cells were lysed

in RIPA (CWBIO, Beijing, China) buffer. Total proteins (10 - 20 μg) were re-suspended in Laemmli buffer (63 mM Tris-HCl, 10% glycerol, 2% SDS, 0.0025% bromophenolblue, pH 6.8) and electrophoresed on SDS-polyacrylamide gels. Then, proteins were transferred to polyvinylidene difluoride membrane. After an incubation with antibodies specific for Flag (bs-0965R; Bioss, Beijing, China) or GAPDH (#CW0098M; CWBIO, Beijing, China), the blots were incubated with anti-mouse (#CW0102M; CWBIO, Beijing, China) or anti-rabbit (#CW0103M; CWBIO, Beijing, China) secondary antibodies conjugated to horseradish peroxidase (HRP). The immunoreactive bands were detected using SuperSignal™ West Pico chemiluminescent substrate kit (Thermo Fisher Scientific, Waltham, MA, USA) and Western blotting detection system (BioRad, Hercules, CA, USA). The GAPDH was used as a loading control for Western blotting assays.

In vitro **translation (IVT) assay**s: To generate the N transcripts for the biophysical studies, the full length of native N sequence (N-RG-1-WT) or RG-1 G4 mutant N sequence (N-RG-1-Mut) were PCR-amplified by specific primers with T7 promoter (Forward primer: AGTAATACGACTCACTATAGGGATGTCTGATAATGGACCC-CA; Reverse primer: GGCCTGAGTTGAGTCAGCAC). For IVT assays, The PCR products were transcribed and translated using TNT® Quick Coupled Transcription/Translation System (Promega). Lysates were subjected to SDS-PAGE and autoradiography. IVT assays were also performed in the presence or absence of PDP.

Table S1. RNA sequences used in our studies.

Oligomer	Sequence (from 5′ to 3′)	Structure
RG-1	5’-GGCUGGCAAUGGCGG-3’	G-quadruplex
F-RG-1	5’-cy5-AGGCUGGCAAUGGCGG-3’	G-quadruplex
RG-1-Mut	5’-GACUGACAAUGACGA-3’	Single Strand
F-RG-1-Mut	5’-cy5-AGACUGACAAUGACGA-3’	Single Strand
AS-RG-1	5’-CCGCCAUUGCCAGCC-3’	Single Strand
L-RG-1	5’-AGAAUGGCUGGCAAUGGCGGUGAUG-3’	G-quadruplex
L-RG-1-Mut	5’-AGAAUGACUGACAAUGACGAUGAUG-3’	Single Strand
RG-2	5’-GGUAUGUGGAAAGGUUAUGG-3’	G-quadruplex
RG-3	5’-GGCUUAUAGGUUUAAUGGUAUUGG-3’	Single Strand
RG-4	5’-GGCCAUGGUACAUUUGGCUAGG-3’	Single Strand
NRAS	5’-UGUGGGAGGGGCGGGUCUGGG-3’	G-quadruplex
ssRNA	5’-UUCUCCGAACGUGUCACGUTT-3’	Single Strand

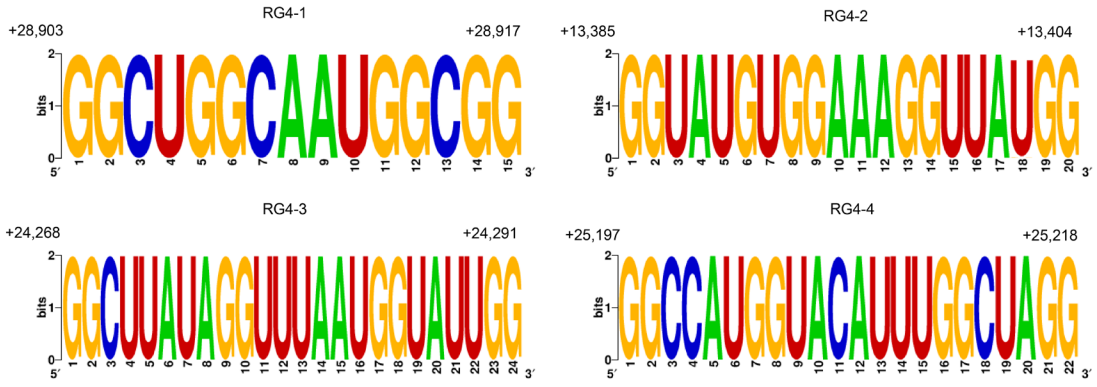


Figure S1. A total of 3365 complete genomic sequences of SARS-CoV-2 genome were retrieved from the University of California Santa Cruz (UCSC) genome browser (https://genome.ucsc.edu/) and the conservation of the four PQSs were aligned using WebLogo software.

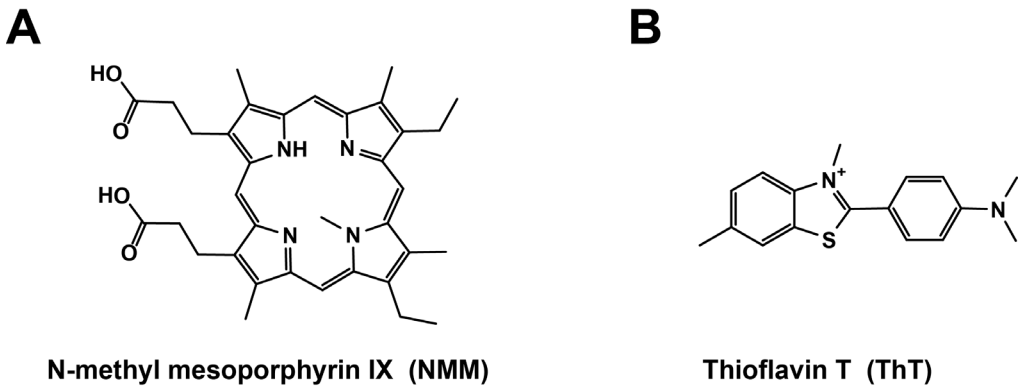


Figure S2. (A) Chemical structures of G4 specific small molecule N-methyl mesoporphyrin IX (NMM) and (B) Thioflavin T (ThT).

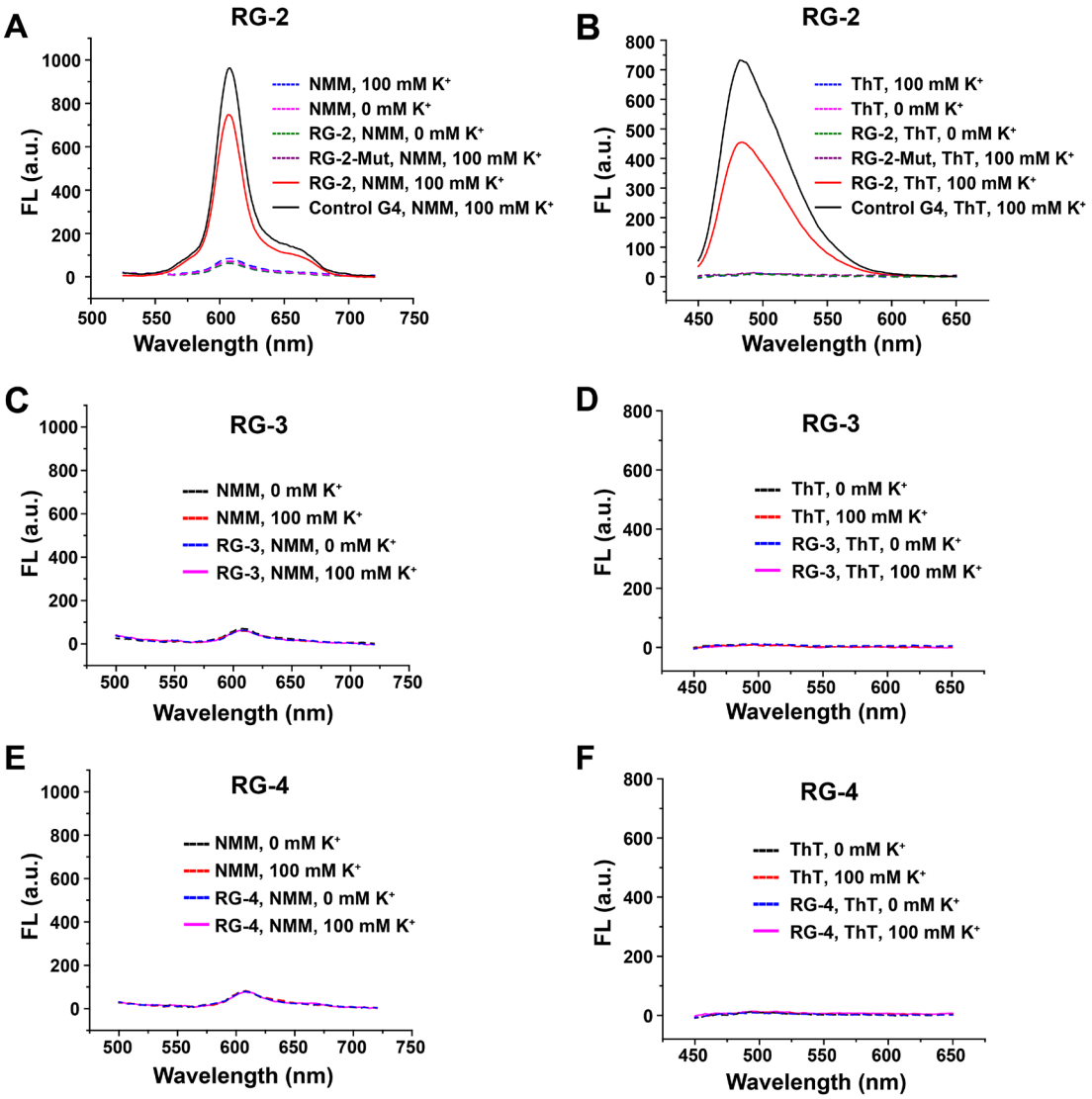


Figure S3. (A) Fluorescence turn-on assays of NMM in the absence or presence of RG-2 under different conditions. NRAS G4 was used as control. (B) Fluorescence turn-on assays of ThT in the absence or presence of RG-2 under different conditions. NRAS G4 was used as control. (C) Fluorescence assays of NMM in the absence or presence of RG-3. (D) Fluorescence assays of ThT in the absence or presence of RG-3. (E) Fluorescence assays of NMM in the absence or presence of RG-4. (F) Fluorescence assays of ThT in the absence or presence of RG-4. RNA was 0.3 μM in strand. NMM was 0.6 μM. ThT was 0.6 μM. Excitation wavelength is 399 nm for NMM. Excitation wavelength is 442 nm for ThT. Assays were carried out in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer.

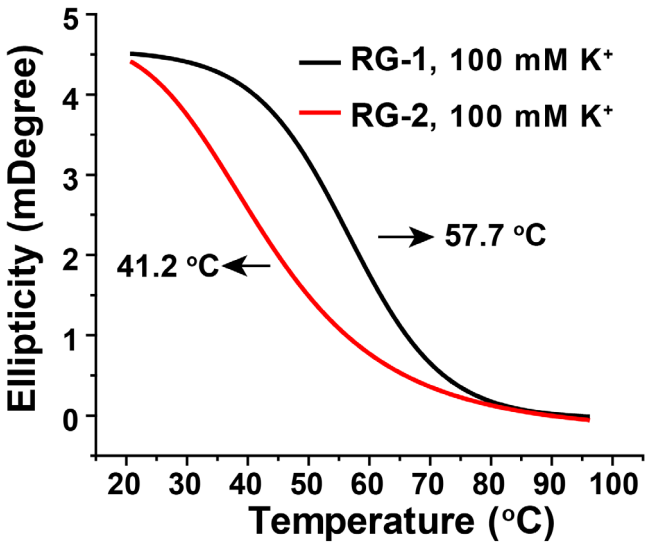


Figure S4. CD melting curves of RG-1 (1.3 μM) and RG-2 (1.3 μM) G-quadruplexes. For RG-1, 267 nm versus temperature were collected. For RG-2, 263 nm versus temperature were collected. Assays were carried out in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer.

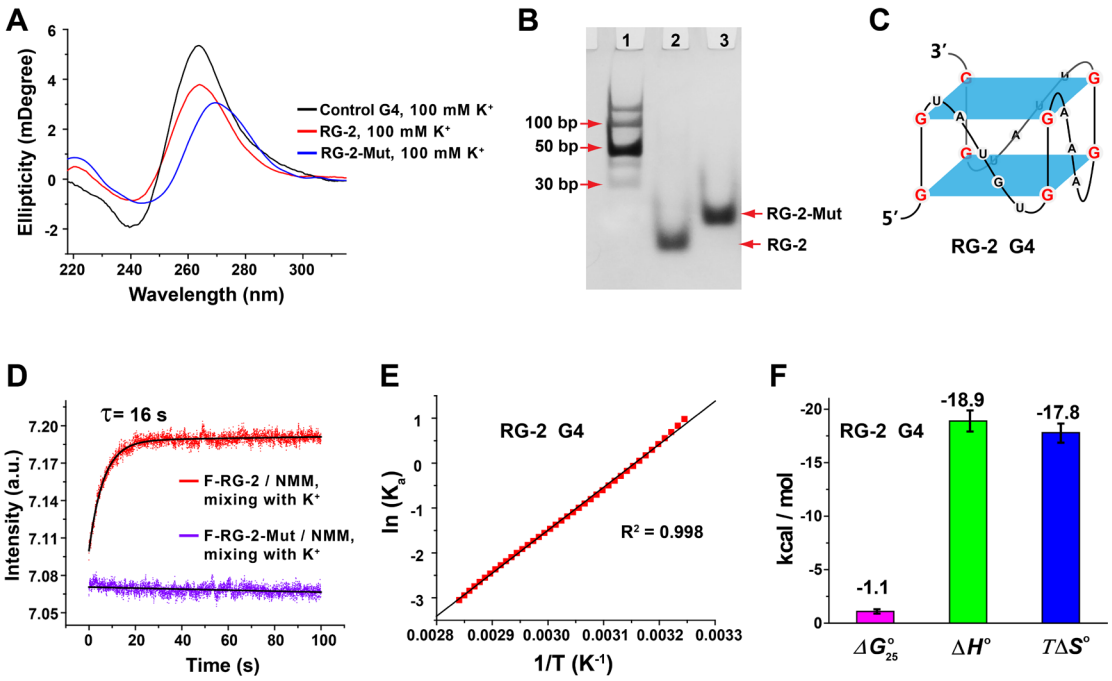


Figure S5. RG-2 G4 formation of and its kinetic and thermodynamic analysis. (A) CD spectra of RG-2 and RG-2-Mut. NRAS G4 was used as control. (B) Native gel electrophoretic analysis of the RG-2 G4 formation. Lane 1, DNA ladder; Lane 2, RG-2; Lane 3, RG-2-Mut. RNA is 2 μM. Samples were prepared in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer. (C) Schematic representation of the proposed RG-2 G4 structure. (D) Typical stopped-flow traces of F-RG-2/NMM or F-RG-2-Mut/NMM samples mixing with 200 mM K^+ buffer. Excitation wavelength was 395 nm, data recorded at 661 nm. (E-F) Determination of $\Delta G°_{25}$, $\Delta H°$ and $\Delta S°$ for RG-2 G4 formation. The thermodynamic parameters were estimated from the melting measurements. Experiments were carried out in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer.

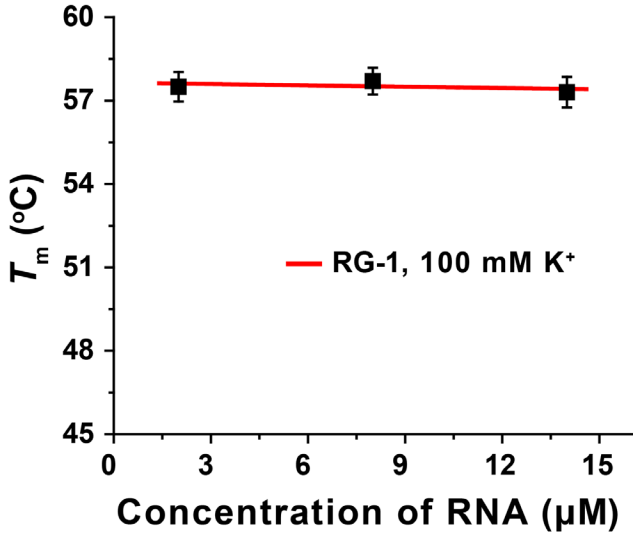


Figure S6. Plots of melting temperature (T_m) versus concentration of RG-1. Melting Assays were carried out in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer.

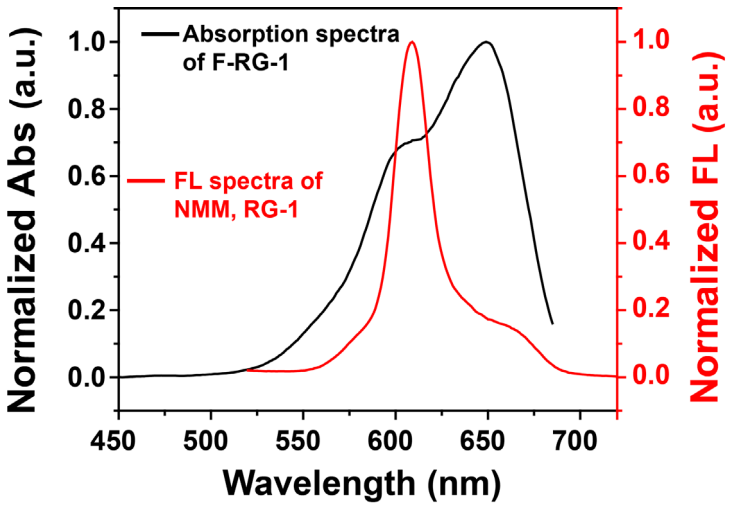


Figure S7. FL spectra of NMM in the presence of RG-1 G4 excited at 980 nm and absorption spectra of Cy5-labeled F-RG-1. Assays were carried out in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer.

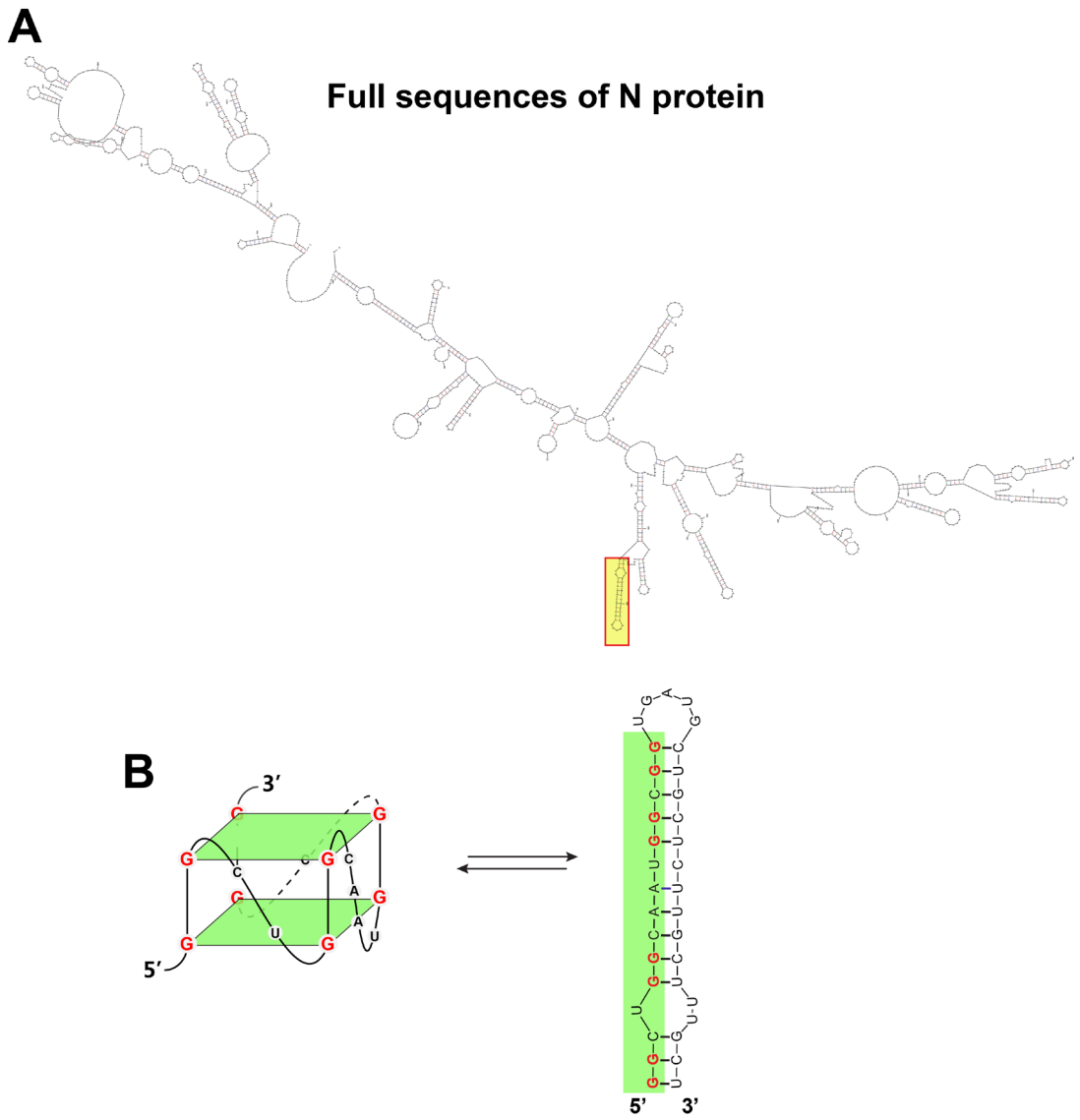


Figure S8. (A) Prediction of the RNA secondary structure of the N gene (SARS-CoV-2) using free-energy minimization. As shown in the yellow area, RG-1 is located in a hairpin region. (B) Schematic illustration of the potential competition between RG-1 G4 and the hairpin structure.

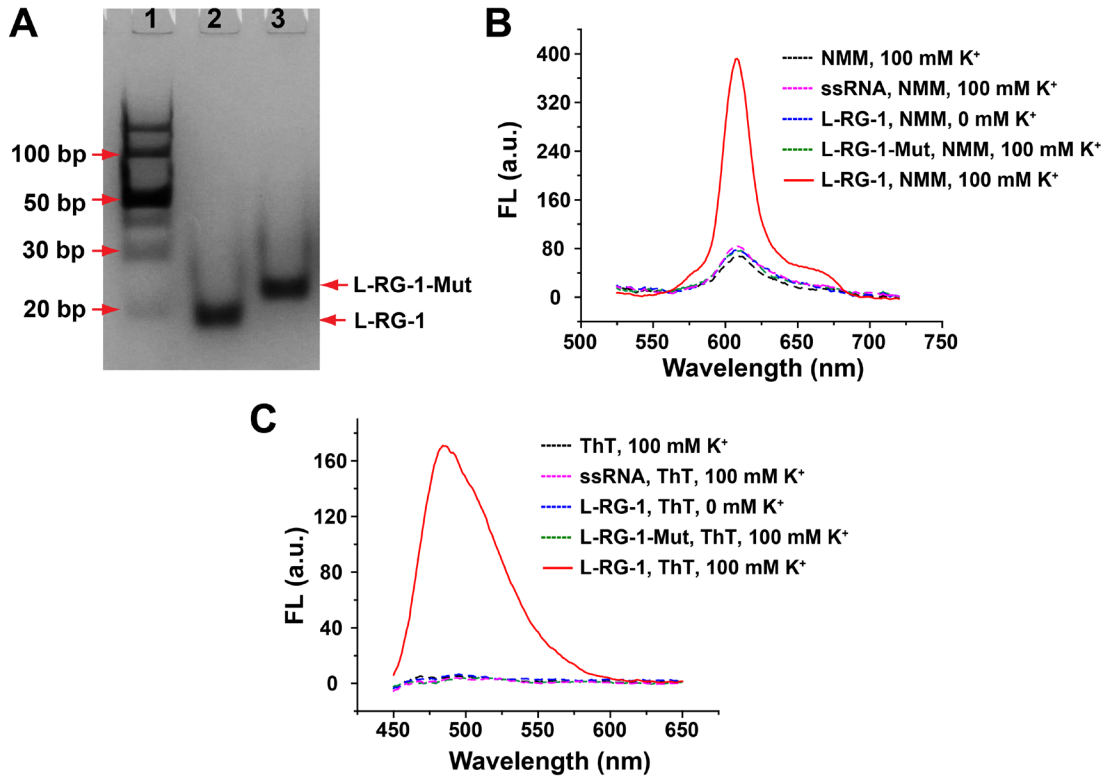


Figure S9. RNA G4 formation of L-RG-1. (A) Native gel electrophoretic analysis of the G4 formation of L-RG-1. Lane 1, DNA ladder; Lane 2, L-RG-1; Lane 3, L-RG-1-Mut. RNA is 2 μM. (B) Fluorescence turn-on assays of NMM (0.6 μM) in the absence or presence of L-RG-1 (0.3 μM) under different conditions. Excitation wavelength was 399 nm. (C) Fluorescence turn-on assays of ThT (0.6 μM) in the absence or presence of L-RG-1 (0.3 μM) under different conditions. Excitation wavelength was 442 nm. Samples were prepared in 10 mM Tris-HCl, 100 mM KCl, pH=7.2 buffer.

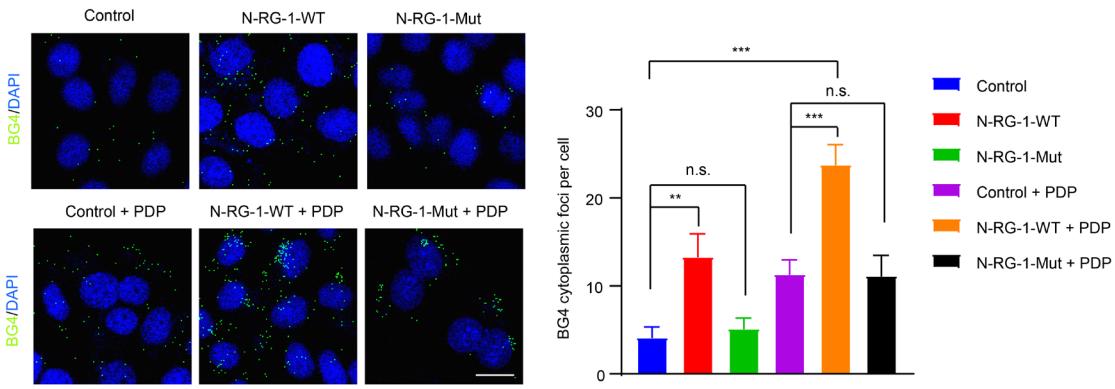


Figure S10. Immunofluorescence microscopy shows detection of endogenous RNA G4 by the G4-specific antibody BG4 in the cytoplasm of HeLa cells transfecting with empty vector, full length of N-RG-1-WT and full length of N-RG-1-Mut without or with PDP treatment (Left). BG4 foci appear green because of FITC fluorescence. Maintenance of cytoplasmic, but not nuclear, BG4 foci after DNase treatment. Nuclei are coloured blue by counterstaining with the general DNA dye DAPI and, outside the nuclei, the cytoplasm is dark. Quantification of BG4 cytoplasmic foci number (Right). For each condition 100 cells were counted and the standard error of the mean was calculated from three replicates. $^{**}P < 0.01$. $^{***}P < 0.001$. n.s., no significance. Scale bar, 20 μm.

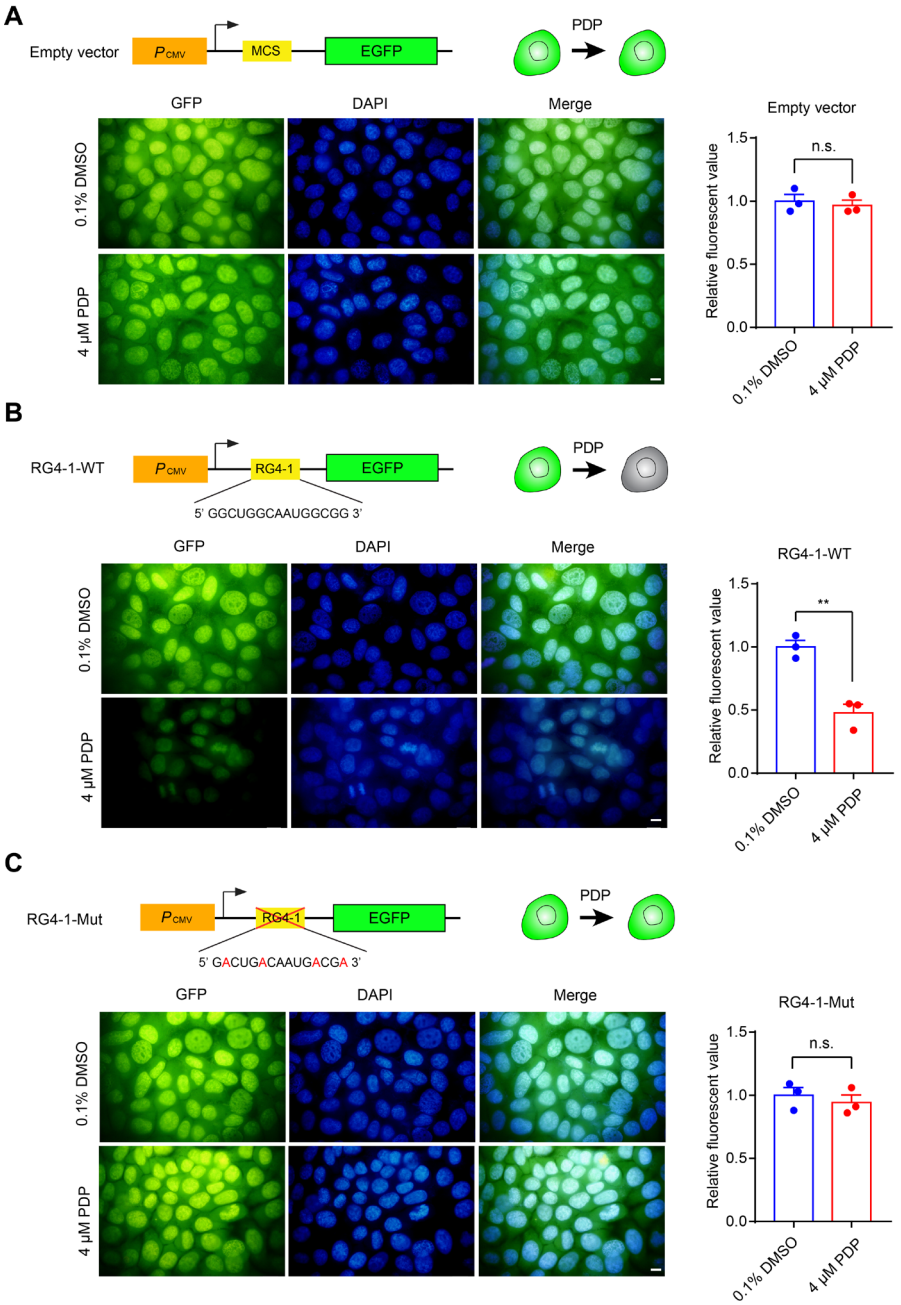


Figure S11. PDP treatment suppresses the expression of EGFP by targeting RG-1 G4. (A-C) HeLa cells steadily transfected with EGFP (A), RG-1-WT-EGFP (B) and RG-1-Mut-EGFP (C) were treated with PDP or DMSO control. Confocal fluorescence microphotographs of the same cells with GFP (left), DAPI (middle) or Merge (right) were demonstrated (scale bar = 20 μm). The relative fluorescent value was measured by flow cytometry. Data are shown as mean ± SEM of three independent experiments, two-tailed Student's t test. SEM, standard error of mean. n.s., not significant. $^{**}P < 0.01$.

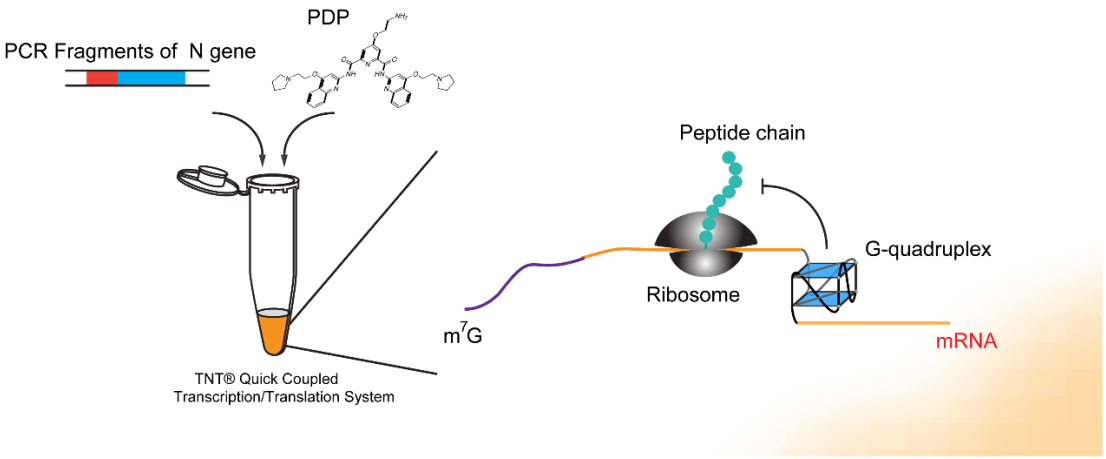
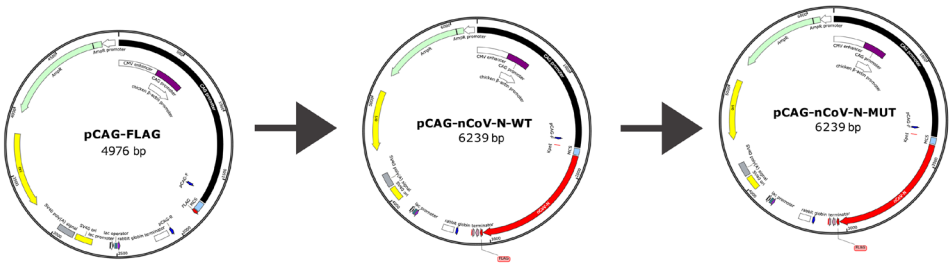


Figure S12. Schematic representation of *in vitro* translation (IVT) assays.



Coding sequence of SARS-CoV-2 N protein

```
1     ATGTCTGATA ATGGACCCCA AAATCAGCGA AATGCACCCC GCATTACGTT TGGTGGACCC
61    TCAGATTCAA CTGGCAGTAA CCAGAATGGA GAACGCAGTG GGGCGCGATC AAAACAACGT
121   CGGCCCCAAG GTTTACCCAA TAATACTGCG TCTTGGTTCA CCGCTCTCAC TCAACATGGC
181   AAGGAAGACC TTAAATTCCC TCGAGGACAA GGCGTTCCAA TTAACACCAA TAGCAGTCCA
241   GATGACCAAA TTGGCTACTA CCGAAGAGCT ACCAGACGAA TTCGTGGTGG TGACGGTAAA
301   ATGAAAGATC TCAGTCCAAG ATGGTATTTC TACTACCTAG GAACTGGGCC AGAAGCTGGA
361   CTTCCCTATG GTGCTAACAA AGACGGCATC ATATGGGTTG CAACTGAGGG AGCCTTGAAT
421   ACACCAAAAG ATCACATTGG CACCCGCAAT CCTGCTAACA ATGCTGCAAT CGTGCTACAA
481   CTTCCTCAAG GAACAACATT GCCAAAAGGC TTCTACGCAG AAGGGAGCAG AGGCGGCAGT
541   CAAGCCTCTT CTCGTTCCTC ATCACGTAGT CGCAACAGTT CAAGAAATTC AACTCCAGGC
601   AGCAGTAGGG GAACTTCTCC TGCTAGAATG GCTGGCAATG GCGGTGATGC TGCTCTTGCT
661   TTGCTGCTGC TTGACAGATT GAACCAGCTT GAGAGCAAAA TGTCTGGTAA AGGCCAACAA
721   CAACAAGGCC AAACTGTCAC TAAGAAATCT GCTGCTGAGG CTTCTAAGAA GCCTCGGCAA
781   AAACGTACTG CCACTAAAGC ATACAATGTA ACACAAGCTT TCGGCAGACG TGGTCCAGAA
841   CAAACCCAAG GAAATTTTGG GGACCAGGAA CTAATCAGAC AAGGAACTGA TTACAAACAT
901   TGGCCGCAAA TTGCACAATT TGCCCCCAGC GCTTCAGCGT TCTTCGGAAT GTCGCGCATT
961   GGCATGGAAG TCACACCTTC GGGAACGTGG TTGACCTACA CAGGTGCCAT CAAATTGGAT
1021  GACAAAGATC CAAATTTCAA AGATCAAGTC ATTTTGCTGA ATAAGCATAT TGACGCATAC
1081  AAAACATTCC CACCAACAGA GCCTAAAAAG GACAAAAAGA AGAAGGCTGA TGAAACTCAA
1141  GCCTTACCGC AGAGACAGAA GAAACAGCAA ACTGTGACTC TTCTTCCTGC TGCAGATTTG
1201  GATGATTTCT CCAAACAATT GCAACAATCC ATGAGCAGTG CTGACTCAAC TCAGGCC
```

Coding sequence of SARS-CoV-2 N protein with RG4 mutation

```
1     ATGTCTGATA ATGGACCCCA AAATCAGCGA AATGCACCCC GCATTACGTT TGGTGGACCC
61    TCAGATTCAA CTGGCAGTAA CCAGAATGGA GAACGCAGTG GGGCGCGATC AAAACAACGT
121   CGGCCCCAAG GTTTACCCAA TAATACTGCG TCTTGGTTCA CCGCTCTCAC TCAACATGGC
181   AAGGAAGACC TTAAATTCCC TCGAGGACAA GGCGTTCCAA TTAACACCAA TAGCAGTCCA
241   GATGACCAAA TTGGCTACTA CCGAAGAGCT ACCAGACGAA TTCGTGGTGG TGACGGTAAA
301   ATGAAAGATC TCAGTCCAAG ATGGTATTTC TACTACCTAG GAACTGGGCC AGAAGCTGGA
361   CTTCCCTATG GTGCTAACAA AGACGGCATC ATATGGGTTG CAACTGAGGG AGCCTTGAAT
421   ACACCAAAAG ATCACATTGG CACCCGCAAT CCTGCTAACA ATGCTGCAAT CGTGCTACAA
481   CTTCCTCAAG GAACAACATT GCCAAAAGGC TTCTACGCAG AAGGGAGCAG AGGCGGCAGT
541   CAAGCCTCTT CTCGTTCCTC ATCACGTAGT CGCAACAGTT CAAGAAATTC AACTCCAGGC
601   AGCAGTAGGG GAACTTCTCC TGCTAGAATG ACTGACAATG ACGATGATGC TGCTCTTGCT
661   TTGCTGCTGC TTGACAGATT GAACCAGCTT GAGAGCAAAA TGTCTGGTAA AGGCCAACAA
721   CAACAAGGCC AAACTGTCAC TAAGAAATCT GCTGCTGAGG CTTCTAAGAA GCCTCGGCAA
781   AAACGTACTG CCACTAAAGC ATACAATGTA ACACAAGCTT TCGGCAGACG TGGTCCAGAA
841   CAAACCCAAG GAAATTTTGG GGACCAGGAA CTAATCAGAC AAGGAACTGA TTACAAACAT
901   TGGCCGCAAA TTGCACAATT TGCCCCCAGC GCTTCAGCGT TCTTCGGAAT GTCGCGCATT
961   GGCATGGAAG TCACACCTTC GGGAACGTGG TTGACCTACA CAGGTGCCAT CAAATTGGAT
1021  GACAAAGATC CAAATTTCAA AGATCAAGTC ATTTTGCTGA ATAAGCATAT TGACGCATAC
1081  AAAACATTCC CACCAACAGA GCCTAAAAAG GACAAAAAGA AGAAGGCTGA TGAAACTCAA
1141  GCCTTACCGC AGAGACAGAA GAAACAGCAA ACTGTGACTC TTCTTCCTGC TGCAGATTTG
1201  GATGATTTCT CCAAACAATT GCAACAATCC ATGAGCAGTG CTGACTCAAC TCAGGCC
```

Figure S13. Map of the plasmid pCAG-Flag and a sequenced portion of this plasmid for verification. The sequenced fragment of the plasmid pCAG-Flag contained the N gene (SARS-CoV-2) and the RG-1-WT and RG-1-Mut sequences were highlighted.