Supplementary Information

A Selective HIV-1 RNA G-quadruplex-Targeting L-Aptamer – D-Antisense Conjugate Inhibits HIV-1 Minus Strand Transfer

Hengxin Feng[a], Chun Kit Kwok[a,b,*]

[a]Department of Chemistry and State Key Laboratory of Marine Pollution, City University of Hong Kong, Kowloon Tong, Hong Kong SAR, China

[b]Shenzhen Research Institute of City University of Hong Kong, Shenzhen, China

*Corresponding author. Email: ckkwok42@cityu.edu.hk

Materials and Methods

References

Materials and Methods

Oligonucleotides

D-DNAs and D-RNAs were synthesized by Integrated DNA Technologies, Inc., Suzhou Biosyntech Co., Ltd or Azenta, Inc.. L-RNAs are synthesized by Suzhou Biosyntech Co., Ltd. After nuclease-free water dissolving, all oligonucleotides were stored at -30 °C before experiment.

G4-SELEX-seq

To start G4-SELEX-seq, an ssDNA N_{30} random library with the ratio (A:T:G:C=35:35:15:15) was extended to dsDNA library with Superscript III reverse transcriptase (SSIII RTase) (Thermo Fisher Scientific, U.S.A.). The ssDNA library was mixed with reverse primer and Li^+ RT buffer with dNTPs mix in nuclease-free water, and heated at 75 °C for 5 min. After cooling to 35 °C for 10 min, SSIII RTase was added to the ssDNA extension sample and subsequently incubated at 50 °C for 50 min. The final extension reaction mixture contained 2 μM ssDNA library, 3 μM reverse primer, 20 mM Tris-HCl (pH 7.5), 4 mM $MgCl_2$, 1 mM DTT, 150 mM LiCl, 5 mM dNTPs and 200 U SSIII RTase with a final volume of 20 μL. The dsDNA library was purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research, U.S.A.) following the manufacturer's instructions. The prepared dsDNA library (100–150 ng) was then transcribed into the N_{30} random RNA library at 37 °C for 2.5 hours using HiScribe T7 High Yield RNA Synthesis Kit (NEB, U.S.A.) with a total volume of 40 μL. The DNA templates were removed by adding 4 U TURBO DNase (Thermo Fisher Scientific, U.S.A.). Next, the RNA library was loaded to 10 % denaturing PAGE for electrophoresis after adding the 2x RNA loading dye (NEB, U.S.A.) and heat denaturation at 95 °C for 5 min. The corresponding RNA band was cut out under short wavelength UV and crushed into small pieces, following with soaking in TE-Li 800 buffer (1x TE buffer, 0.8 M LiCl, 320 μL for every 40 μL transcription reaction). After 4 °C overnight incubation at 1300 rpm in a thermo-shaker, the remaining gel pieces were filtered out with a cellulose acetate filter (0.45 μm pore size, Thermo Fisher Scientific, U.S.A., or Corning, U.S.A.) and extracted with RNA clean and concentrator (RNA column) (Zymo Research, U.S.A.). For beads preparation, 0.75 mg of Streptavidin Magnetic Beads (10 mg/mL; MedChemExpress, U.S.A.) were washed twice with solution A (100 mM NaOH and 50 mM NaCl, 600 μL each time) and once with solution B (100 mM NaCl, 600 μL) and subsequently resuspended with 75 μL 2× selection buffer (150 mM KCl, 25 mM Tris-HCl pH 7.5, and 10 mM $MgCl_2$) with the addition of 7.5 μg of yeast tRNA (10 mg/mL, Thermo Fisher Scientific, U.S.A.). The mixture was shaken at 700 rpm for one hour at 25 °C in a thermo-shaker (Allsheng, China). The RNA library pool (final 4.5 μM, first round) was mixed with 2x RNA buffer (300 mM KCl and 50 mM Tris-HCl, pH 7.5) with a total volume of 50 μL, and then heated at 95 °C for 5 min for denaturation, followed by refolding at room temperature for 15 min. Negative selection was performed with the addition of 25 μL tRNA pre-blocked beads to the refolded RNA library and then incubated at 700 rpm for 2 h. Subsequently, the non-specific binding RNAs were removed by discarding the beads.

To start the positive selection, U3-III L-rG4 with 5' biotin (Suzhou Biosyntech, China) was heated at 95 °C and cooled down at room temperature for 15 min in 1x selection buffer (150 mM KCl, 25 mM Tris-HCl pH 7.5, and 5 mM $MgCl_2$) for refolding and mixed with the RNA pool with the final concentration of 0.65 μM (first round). The mixture was incubated at 37 °C with shaking at 300 rpm for 30 min for target binding in the 1x selection buffer. Next, the biotin-labelled U3-III L-rG4 and D-RNA complexes were immobilized with 500 μg beads and washed with 600 μL of 1x selection buffer (150 mM KCl, 25 mM Tris-HCl pH 7.5, and 5 mM $MgCl_2$) five times. For elution, target-binding RNAs were treated with elution buffer twice (25 mM NaOH and 1 mM EDTA, 100 μL each time) and subsequently neutralized with 50 μL of 1M Tris-HCl immediately. The eluted RNAs were purified with the RNA column to reach a final volume of 30 μL by column elution with nuclease-free water. The collected RNAs were then reverse transcribed to ssDNA templates. All 30 μL purified RNAs were mixed with reverse primer and Li^+ RT buffer with dNTPs mix in nuclease-free water, and heated at 95 °C for 5 min. After cooling down at 35 °C for 10 min, SSIII RTase was added to the mixture and subsequently incubated at 50 °C for 15 min. The final reverse transcription reaction mixture contained all selected RNAs, 0.33 μM reverse primer, 20 mM Tris-HCl pH 7.5, 4 mM $MgCl_2$, 1 mM DTT, 150 mM LiCl, 5 mM dNTPs and 400 U SSIII RTase with a volume of 60 μL. The reaction was terminated by adding 3 μL of 2 M NaOH, and heating at 95 °C for 10 min to degrade the RNAs and denature the reverse transcriptase, followed by neutralization with 15 μL 1 M Tris-HCl, pH 7.5 after cooling down to 25 °C. The ssDNA templates were purified with the RNA column to reach a final volume of 30 μL by column elution with nuclease-free water and PCR amplified for next selection round. Small-scale PCR cycle tests were performed to optimize the PCR cycle for large-scale amplification. 8 μL of purified ssDNA was mixed with forward and reverse primers (each with a final concentration of 0.5 μM) and 10 μL of 2x KAPA HiFi HotStart ReadyMix (Roche) to a total volume of 20 μL and aliquoted into five tubes with 4 μL mixture in each. The five tubes were with the amplification cycles of 6, 8, 10, 12 and 14 respectively (PCR program: 3 min at 98 °C, [20 s at 98 °C, 20 s at 60 °C, 20 s at 72 °C]$_n$, 3 min at 72 °C, 4 °C, n = number of cycles). The amplified samples were mixed with 1 μL 6× DNA loading dye (NEB, U.S.A.) respectively and loaded to SYBR Safe (Thermo Fisher Scientific, U.S.A.) stained 2% agarose gel for electrophoresis and scanned with ChemiDoc Touch imaging system (Bio-Rad). The optimal PCR cycle was determined as the lowest PCR cycle that provided a clear band. The large-scale DNA amplification was then performed in a volume of 40 μL under the same conditions. The purification and collection of the selected dsDNAs library was carried out with Zymoclean™ Gel DNA Recovery Kit. Five rounds of selection were performed and the dsDNA from 2nd to 5th rounds were later amplified with NGS adapters for four PCR cycles and pooled together for NGS sequencing (IGE Biotechnology, China) after dehydration with Concentrator Plus (Eppendorf). The NGS result was analyzed with Python and Microsoft Excel.

Electrophoretic mobility shift assay (EMSA)

For each set of EMSA samples, samples of D- or L-Apt.T8 were prepared with two-

fold serial dilution with the highest initial concentration of 8 μM/2 μM in 5 μL, together with one negative control, in binding buffer containing 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$. The FAM labelled L- or D-rG4s were prepared in the same buffer with an initial concentration of 20 nM and a total volume of 50 μL. After heating the aptamers and the rG4s separately at 75 °C for 5 min, all the samples were incubated on ice for 10 min before adding 5 μL rG4 to each aptamer sample to generate a mixture containing 10 nM FAM labelled rG4s and serial diluted aptamers with the highest final concentration of 4 μM/1 μM in the binding buffer. The negative control only contained 10 nM FAM labelled rG4s and the binding buffer. The samples were incubated at 37 °C for one hour and subsequently incubated at 4 °C overnight. Glycerol was added to the samples to achieve a final concentration of 8% before loading samples to an 8% native gel, running in the buffer containing 50 mM potassium acetate, 25 mM Tris-HCl pH 7.5 and 1 mM $MgCl_2$ with the consistent current of 70 mA for 35 to 45 min at 4 °C. After running, the gels were scanned by Amersham TYPHOON Imager (Cytiva) and analyzed by the analysis software, ImageQuantTL.

The fraction bound was calculated from dividing the volume of binding complex by the sum of the volume of binding complex and free targets. K_D was analyzed by GraphPad Prism with the one-site total mode.

In single concentration selectivity test of L-Apt.T8 (Fig. S8), the negative controls contained 10 nM FAM labelled tested oligonucleotides while the testing samples contained 10 nM FAM labelled tested oligonucleotides and 100 nM L-Apt.T8 in the same buffer condition mentioned above.

In selectivity test of L-Apt.T8, the D-rG4s referred to U3-III rG4, *APP* rG4, *hTERC* rG4 and *BCL2* rG4. In selectivity test of L-Apt.T8-10D, the D-rG4-exts referred to U3-III-ext, *APP*-ext, *hTERC*-ext and *BCL2*-ext.

Microscale thermophoresis (MST) binding assay

For each set of MST binding assay, sixteen samples were prepared with two-fold serial dilution with the highest initial concentration of 8 μM in 5 μL in binding buffer containing 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$. The FAM labelled L- or D-rG4s were prepared in the same buffer with the initial concentration of 60 nM and the total volume of 85 μL. After heating the aptamers and the rG4s separately at 75 °C for 5 min, all the samples were incubated on ice for 10 min before adding 5 μL rG4 to each aptamer sample to generate a mixture containing 30 nM FAM labelled rG4s and serial diluted aptamers with the highest final concentration of 4 μM in the binding buffer. The samples were incubated at 37 °C for one hour and subsequently incubated at 4 °C overnight before loading to MST capillary tubes (Nanotemper, Germany) and scanned by Monolith NT.115 (Nanotemper, Germany) under blue light mode. Results were analyzed with the software Nano Temper Analysis (nta).

Circular dichroism (CD) spectroscopy

CD samples contained 5 μM D-Apt.T8, 10 mM LiCac pH 7.0, and 150 mM LiCl or KCl with a total volume of 2 mL. The samples were refolded by heating at 95 °C for 5 min and cooling at room temperature for 15 min. Samples were then loaded to a quartz

cuvette (1 cm path length, Hellma Analytics) and scanned at 25 °C within the wavelength ranging from 220 to 310 nm with J-1500 Circular Dichroism spectrophotometer (Jasco, Japan). The data were collected every 1 nm with a responding time of 0.5 nm/s and all spectra were the average of two times of scanning. The spectra were smoothed every five data points and normalized following previous study[1] with Microsoft Excel.

Thermal Denaturation Monitored by Ultra-Violet Spectroscopy (UV Melting)

Samples for UV melting were with the same preparation as CD spectroscopy. Agilent Cary 100 UV-Vis Spectrophotometer and Cary 3500 UV-Vis Spectrophotometer were used for the experiments. 3 layers of Teflon tape were used to seal the cuvettes after loading the sample solutions. The absorbance at 295 nm were monitored from 20 to 95 °C with 0.5 °C/min temperature increasing. The temperature was hold at 95 °C for 5 min, following by reverse scanning of the samples at 295 nm with a 0.5 °C/min decreasing rate until 20 °C and hold for 5 min. The initial data were corrected according to the background provided by the blank solutions containing 150 mM KCl or LiCl and 10 mM LiCac (pH 7.0) with the same volume of tested samples. The data was then smoothed every ten data points and normalized following previous study[1] with Microsoft Excel.

G4-ligand enhanced fluorescence assay

Sample solutions were prepared with 1 μM D-Apt.T8, 10 mM LiCac pH 7.0), 150 mM KCl/LiCl with the volume of 200 μL. By heating at 95 °C for 5 min and cooling at 25 °C for 15 min, the oligos were refolded. Thioflavin T (ThT) and N-methyl mesoporphyrin IX (NMM) ligands were added to the solutions before transferring to the quartz cuvettes. The excitation wavelength for ThT was 425 nm and and 394 nm for NMM, while the emission spectra were measured within 440-700 nm for ThT and 550-750 nm for NMM at 25 °C with HORIBA FluoroMax-4 Fluorometer. The entrance slit was 5 nm, together with 2 nm exit slit and 2 nm increment. The spectra were smoothed every five data points and plotted with Microsoft Excel.

G4-ligand enhanced fluorescence assay for immobilized U3-III rG4

Sample solutions were prepared with 1 μM D-Apt.T8, 10 mM LiCac pH 7.0), 150 mM KCl/LiCl with the volume of 100 μL. By heating at 95 °C for 5 min and cooling at 25 °C for 15 min, the oligos were refolded. The oligos were then incubated with 10 μL streptavidin magnetic beads (supernatant removed) at 25 °C for 30 min. Thioflavin T (ThT) and N-methyl mesoporphyrin IX (NMM) ligands were added to the samples respectively before transferring to the 96 well microplate (Thermo Fisher Scientific, U.S.A.) and incubated at room temperature for 5 min. The excitation wavelength was 425 nm for ThT and 394 nm for NMM. The endpoint emission was read at the peak of 610 nm for NMM and 490 nm for ThT with a microplate reader (SpectraMax iD5 Multi-Mode Microplate Reader). The signal was corrected to the background of beads with G4 ligands in the respective buffer. Signal of samples in K^+ buffer was normalized to samples in Li^+ buffer accordingly. Three replicates were performed and error bar was

the standard deviation.

L-Apt.T8-10D preparation with click-reaction

For each click-reaction sample, 1 nmol of 5' hexynyl labelled L-Apt.T8 and 4 nmol of 3' azide labelled 10 nt DNA antisense were heated at 95 °C for 5 min and cooling down on ice for 10 min separately. Oligos after cooling were mixed with 2 μL 2M TEAA (pH 7.0) and 2 μL DMSO following by addition of 2 μL freshly prepared 10 mM L-ascorbic acid and 1 μL 10 mM Cu(II)-TBTA (Lumiprobe Corporation, U.S.A.). DMSO was added to reach the total volume of 20 μL. The mixture was incubated at 40 °C, shaking at 1000 rpm with a thermo-shaker for about 5 hours. A small amount, i.e. 1 μL, was aliquoted and diluted for electrophoresis with 10% denaturing urea gel to confirm the completion of reaction. The conjugated product was purified with RNA column and the product was quantified by nanodrop.

EMSA for complementary binding interference test

For each set of complementary binding interference test, ten initial samples of L-Apt.T8-10D/L-Apt.T8 were prepared with the concentration of 20 μM, 10 μM, 2 μM, 1 μM, 200 nM, 100 nM, 20 nM, 10 nM, 2 nM, and 0 nM, in 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$ with the volume of 5 μL. The buffer condition was applied for all following sample preparation. Two control samples were prepared with nuclease-free water only or 142.86 nM L-Apt.T8-10D/L-Apt.T8 with the volume of 7 μL. FAM labelled U3-III-ext was diluted to the concentration of 33 nM with the volume of 39 μL. The complementary $tRNA_3^{Lys}$ region (CompR) was diluted to the the concentration of 750 nM with the volume of 22 μL. The L-Apt.T8-10D/L-Apt.T8 samples, U3-III-ext and CompR were heated for refolding at 95 °C for 5 min and cooling down at room temperature for 10 min respectively. 3 μL of U3-III-ext mixture was aliquoted and added to the two control samples respectively. The CompR mixture was mixed well with the remaining U3-III-ext mixture to a total volume of 55 μL. 5 μL of the mixture of CompR and U3-III-ext was added to the ten samples of L-Apt.T8-10D/L-Apt.T8 respectively. Each sample for the complementary binding interference test, excepting the two control samples, was with the volume of 10 μL, containing 10 nM U3-III-ext, 150 nM CompR, 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$, with the L-Apt.T8-10D/L-Apt.T8 concentrations of 10 μM, 5 μM, 1 μM, 500 nM, 100 nM, 50 nM, 10 nM, 5 nM, 1 nM, and 0 nM respectively. The two control samples were in the same buffer conditions but contained 10 nM U3-III-ext only or 10 nM U3-III-ext and 100 nM L-Apt.T8-10D/L-Apt.T8. The samples were incubated at 37 °C for one hour following by slow cooling down to 4 °C for 30 min. 2.5 μL of 40% glycerol was added to each tube to achieve the total volume of 12.5 μL. Each sample was separated into two halves and loaded to two 8% native gels for electrophoresis separately. The running buffer contained 25 mM potassium acetate, 25 mM Tris-HCl pH 7.5 and 1 mM $MgCl_2$. Each gel was run with the consistent current of 150V for 35 min (all the bands in the gel) or 45 min (better separation between Comp. complex and Apt. complex without free U3-III-ext band) at 4 °C. After running, the gels were scanned by Amersham TYPHOON Imager (Cytiva) and analyzed by the analysis software, ImageQuantTL.

The Comp. complex fraction was calculated from dividing the volume of complementary binding complex by the sum of the volume of complementary binding complex (Comp. complex) and L-Apt.T8-10D binding complex (Apt. complex). IC_{50} was analyzed by GraphPad Prism with the [Inhibitor] vs. response (three parameters) mode.

EMSA for HIV-1 NC interaction inhibition test

For each set of NC interaction inhibition test, ten initial samples of L-Apt.T8-10D were prepared with the concentration of 20 μM, 10 μM, 2 μM, 1 μM, 200 nM, 100 nM, 20 nM, 10 nM, 2 nM, and 0 nM, in 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$ with the volume of 5 μL. The buffer condition was applied for all following sample preparation. Two control samples were prepared with nuclease-free water only or 566 nM L-Apt.T8-10D with the volume of 5.3 μL. FAM labelled U3-III-ext was diluted to the concentration of 31.9 nM with the volume of 65.8 μL. L-Apt.T8-10D and U3-III-ext was heated for refolding at 75 °C for 5 min and cooling down on ice for 10 min respectively. 14.1 μL of U3-III-ext mixture was aliquoted for the two control samples. 3.3 μL of 10 μM HIV-1 NC (Synpeptide, China) was mixed well with the remaining U3-III-ext mixture to a total volume of 55 μL and incubated at 37 °C for 50 min. 5 μL of the mixture of NC and U3-III-ext was added to the ten samples of L-Apt.T8-10D. Each sample for the NC interaction inhibition test, excepting the two control samples, was with the volume of 10 μL, containing 15 nM U3-III-ext, 300 nM NC, 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 1 mM $MgCl_2$, with the L-Apt.T8-10D concentrations of 10 μM, 5 μM, 1 μM, 500 nM, 100 nM, 50 nM, 10 nM, 5 nM, 1 nM, and 0 nM respectively. The two control samples are in the same buffer conditions but contained 15 nM U3-III-ext only or 15 nM U3-III-ext and 300 nM L-Apt.T8-10D. The samples were incubated at 37 °C for one hour following incubation at 4 °C overnight and loading to the 8% native gel for electrophoresis with 0.5x TBE buffer. 2.5 μL of 40% glycerol was added to each tube before 4 °C overnight incubation Each gel was run with the consistent voltage of 150V for 60 min at 4 °C. After running, gels were scanned by Amersham TYPHOON Imager (Cytiva) and analyzed by the analysis software, ImageQuantTL.

The NC interaction ratio was calculated from dividing the volume of NC binding and unwound U3-III-ext bands by the sum of the volume of NC binding, unwound U3-III-ext, Apt. complex and U3-III-ext with G4 bands. IC_{50} was analyzed by GraphPad Prism with the [Inhibitor] vs. response (three parameters) mode.

Similar set up was applied to the BRACO-19 inducing NC unwinding inhibition by replacing the L-Apt.T8-10D with BRACO-19. BRACO-19 was prepared with the two-fold serial dilution with the highest concentration of 20 μM without heating for refolding. The negative controls were with 15 nM U3-III-ext only or 15 nM U3-III-ext and 300 nM BRACO-19 with the same buffer component.

Preparation of FAM labelled $tRNA_3^{Lys}$

To generate the 76 nt FAM $tRNA_3^{Lys}$, the tRNA-F1 with 5' FAM labelling and 3' OH was ligated to the tRNA-F2 with 5' phosphate and 3' Spc3. Ligation reaction samples

contained 2 μM tRNA-F1, 20 μM tRNA-F2, 1x T4 RNA ligase buffer, 10% DMSO, 20% PEG8000, and 0.5U/μL T4 RNA ligase 1 (NEB, U.S.A.). The samples were incubated at 25 °C for ligation for 2 hours. The FAM labelled ligation product was separated from unligated tRNA-F1 under blue light and purified with 10% urea gel and extracted with RNA column.

Preparation of donor and acceptor RNA substrates for minus strand transfer

The dsDNA templates for donor and acceptor RNA were generated by amplification the according sequence from the HIV-1 vector, pNL43 (NovoPro, China) with PCR. D-FP, D-RP, and A-FP, A-RP were the two pairs of primers used for amplification with 2x KAPA HiFi HotStart ReadyMix for donor and acceptor templates respectively. The PCR reactions and following purification were as mentioned previously. Both forward primers, D-FP and A-FP were with T7 RNA promoter. The purified dsDNA templates were then transcribed into donor and acceptor RNA substrates with HiScribe T7 High Yield RNA Synthesis Kit and purified accordingly as mentioned above.

Minus strand transfer inhibition with L-Apt.T8-10D

The primer for minus strand transfer, FAM $tRNA_3^{Lys}$, was annealed to the donor RNA by heating at 95 °C for 5 min and cooling down to 37 °C for 10 min. The acceptor RNA was denatured at 95 °C for 5 min separately for denaturing secondary structure and slow cooled down to room temperature for 10 min. After the cooling down of the two mixtures above, the annealed FAM $tRNA_3^{Lys}$ and donor RNA were added to the acceptor RNA and incubated with 200% NC (100% NC is 7 nt per NC molecule)[2] at 37 °C for 3 min. Subsequently, the HIV-1 reverse transcriptase (Abcam, UK) was added and incubated at 37 °C for 5 min. The mixture was aliquoted and added to L-Apt.T8-10D with the final concentration of 10 μM, 5 μM, 1μM, 500 nM, 100 nM, 50 nM, 10 nM and 0 nM. L-Apt.T8-10D samples with different concentrations were heated at 75 °C for 5 min and on ice incubated for more than 10 min before mixing. The reactions were initiated by adding dNTPs and reverse transcription buffer. Each 10 μL reaction mixtures were with the final components of 200 nM FAM $tRNA_3^{Lys}$, 100 nM donor RNA, 200 nM acceptor RNA, 200% NC, 1U HIV-1 reverse transcriptase, 50 mM Tris-HCl (pH 8.0), 50 mM KCl, 1 mM DTT, 1 mM EDTA, 6 mM $MgCl_2$, and 50 μM dNTPs, together with the corresponding L-Apt.T8-10D concentration. After incubation at 37 °C for 30 min, 0.5 μL proteinase K (Invitrogen, U.S.A.) was added to each 10 μL mixture and incubated at 25 °C for 30 min and terminated by heating at 65 °C for 10 min. Samples were purified with RNA column and eluted with 11 μL H_2O before mixture with 2x RNA loading dye (NEB, U.S.A.) and electrophoresis with 10% denaturing urea gel. for analysis of minus strand transfer ratio. After electrophoresis, the gels were scanned by Amersham TYPHOON Imager (Cytiva) and analyzed by the analysis software, ImageQuantTL.

The minus strand transfer ratio was calculated from dividing the volume of transfer product by the sum of the volume of transfer product and donor extension product. IC_{50} was analyzed by GraphPad Prism with the [Inhibitor] vs. response (three parameters) mode.

Table S1 Sequences of oligonucleotides used in this study.

Name	Sequence (5' to 3')	Nucleotides
Low GC-content DNA N_{30} Library	TTCTAATACGACTCACTATAGGTTACCAGCCTTCACTGC (N_{30}) GCACCACGGTCGGTCACAC	88
Forward Selection Primer	TTCTAATACGACTCACTATAGGTTACCAGCCTTCACTGC	39
Reverse Selection Primer	GTGTGACCGACCGTGGTGC	19
Biotin L-U3-III rG4	Biotin-GGGAGGCGUGGCCUGGGCGGGACUGGGG (All L-RNA bases)	28
FAM L-U3-III rG4	FAM-GGGAGGCGUGGCCUGGGCGGGACUGGGG (All L-RNA bases)	28
D-Apt.4-1c	GCCCUAAAGGUGGUGGUGGGAGGGC	25
D-Apt.4	GGUUACCAGCCUUCACUGCGAAUCUUAAGGUGGAAUGGGGAGGUGAUUUGCACCACGGUCGGUCACAC	68
D-Apt.T4	GAAUCUUAAGGUGGAAUGGGGAGGUGAUUU	30
D-Apt.7	GGUUACCAGCCUUCACUGCUCAUGGGGUACUUAAAUAGGUGGCGGGAGAGCACCACGGUCGGUCACAC	68
D-Apt.T7	UCAUGGGGUACUUAAAUAGGUGGCGGGAGA	30
D-Apt.8	GGUUACCAGCCUUCACUGCUCGAAUGGGGGACAGGGAGUGGGAUUUCGAGCACCACGGUCGGUCACAC	68
D-Apt.T8	UCGAAUGGGGGACAGGGAGUGGGAUUUCGA	30
D-Apt.T8-S1	UCGAUUGGGGGACAGGGAGUGGGAUAUCGA	30
D-Apt.T8-S2	AGCAUUGGGGGACAGGGAGUGGGAUAUGCU	30
D-Apt.T8-S3	GACGUUGGGGGACAGGGAGUGGGAUACGUC	30
D-Apt.T8-G9A	UCGAAUGGAGGACAGGGAGUGGGAUUUCGA	30
D-Apt.T8-G16A	UCGAAUGGGGGACAGAGAGUGGGAUUUCGA	30
D-Apt.T8-G22A	UCGAAUGGGGGACAGGGAGUGAGAUUUCGA	30
L-Apt.T8	UCGAAUGGGGGACAGGGAGUGGGAUUUCGA (All L-RNA bases)	30
FAM U3-III rG4	FAM-GGGAGGCGUGGCCUGGGCGGGACUGGGG	28
FAM- *APP* rG4	FAM-GGGGCGGGUGGGGAGGGG	18
FAM- *hTERC* rG4	FAM-GGGUUGCGGAGGGUGGGCCU	20
FAM- *BCL2* rG4	FAM-GGGGGCCGUGGGGUGGGAGCUGGGG	25
c-Kit dG4	FAM-AGGGAGGGCGCTGGGAGGAGGG	22
RNA hairpin	FAM-CAGUACAGAUCUGUACUG	18
5' hexynyl - L-Apt.T8	Hexynyl-UCGAAUGGGGGACAGGGAGUGGGAUUUCGA (All L-RNA bases)	30
D-DNA antisense - 3' azide	GCTCGCCACT-Azide	10
FAM U3-III-ext	FAM-GGGAGGCGUGGCCUGGGCGGGACUGGGGAGUGGCGAGCCCUCAGAU	46
FAM 18 nt ext	FAM-AGUGGCGAGCCCUCAGAU	18
FAM- *APP*-ext	FAM-	36

	GGGGCGGGUGGGGAGGGGUGCUCUGCUGGUCUUCAA	
FAM- *hTERC*-ext	FAM-GGGUUGCGGAGGGUGGGCCUGGGAGGGGUGGUGGCCAU	38
FAM- *BCL2*-ext	FAM-GGGGGCCGUGGGGUGGGAGCUGGGGCGAGAGGUGCCGUUGGCC	43
CompR ($tRNA_3^{Lys}$ segment)	Sp18-AUCUGAGGGUCCAGGGUUCAAGUCCCUGUUCGGGCGCCA	39
Donor-FP (D-FP)	GGATCCTAATACGACTCACTATAGGGTCTCTCTGGTTAGACCAGATC	47
Donor-RP (D-RP)	GTCCCTGTTCGGGCGCCACTGCTAGAGA	28
Acceptor-FP (A-FP)	GGATCCTAATACGACTCACTATAGGGAGGCGTGGCCTGGGCGGGACTGGGG	51
Acceptor-RP (A-RP)	GGTCTAACCAGAGAGACCCAGTACAGGCAAAAAGCAG	37
tRNA-F1	FAM-GCCCGGAUAGCUCAGUCGGUAGAGCAUCAGACUUUU	36
tRNA-F2	Phos-AAUCUGAGGGUCCAGGGUUCAAGUCCCUGUUCGGGCGCCA-Spc3	40

Table S2 Conditions of the *in vitro* selection in G4-SELEX-seq process.

Cycles	**1**	**2**	**3**	**4**	**5**
$MgCl_2$ Conc º (mM)	5	5	5	5	1
D-RNA poolº (pmol)	1000	500	200	50	25
Target L-U3-III rG4 (pmol)	200	200	200	100	30
Negative selectionº (hr)	2	2	2	1	1
Positive selection (min)	30	30	30	30	30
Washing (min)	Pipette Mix	Pipette Mix	Pipette Mix	10	10
Temperatureº (°C)	37	37	37	37	37

Table S3 NGS results with reads percentage in each round.

	Sequences	R2_reads (%)	R3_reads (%)	R4_reads (%)	R5_reads (%)
1	TGAACTTGGTGGTGGTGGAATTGTTCAGTA	0.0017	0.0271	0.6288	4.6711
2	TCTGGATTGGTGGTGGTGGTATGGTCCAGA	0.0011	0.0160	0.3038	2.3976
3	GGACAGGAAATGGTGGTGGTGGACATGTTC	0.0022	0.0299	0.6374	2.1924
4	ATCTCATTGGTGGTGGTGGAATGTGAGTAT	0.0011	0.0127	0.2545	2.1795
5	TTGGTGCTACAAAAGGTGGTGGTGGATGTA	0.0006	0.0107	0.2213	1.9023
6	ATAGATTACGGTGGTGGTGGAAATGTCTAT	-	0.0107	0.2437	1.5375
7	ATGTGCGTAAGGTGGTGGTGGACTGCACAT	0.0006	0.0160	0.2638	1.3747
8	TTTGAGGTACGGTGGTGGTGGATCATCAAA	0.0011	0.0123	0.2892	1.2947
9	GAATCTTAAGGTGGAATGGGGAGGTGATTT	0.0006	0.0180	0.4030	1.2324
10	GTACGACTCGGTGGTGGTGGAATGACGTGC	0.0017	0.0131	0.2444	1.2060
11	AGTGTTATAAGGTGGTGGTGGGTTAATACT	-	0.0082	0.1746	1.0701
12	GGGTCCCTGGTGGTGGTGGAAATCGGATTC	-	0.0115	0.2209	1.0648
13	ATATGTATAGGTGGTGGTGGAGTGACATAT	0.0006	0.0078	0.1694	1.0277
14	ATTCGAAATGGTGGTGGTGGAACTCGATAT	0.0006	0.0152	0.2049	0.8816
15	AACGTGTCTCTTGGTGGTGGTGGAATAGAG	-	0.0045	0.1515	0.7822
16	ATCAGGATTTGGTGGTGGTGGTTATCCTAT	-	0.0082	0.1250	0.7790
17	ATTATTAATAGGTGGTGGTGGAGAAATAAT	0.0006	0.0086	0.1258	0.6640
18	TGTGGTGCTTTATGGTGGTGGTGGGAAAAT	0.0006	0.0066	0.1194	0.6484
19	AAAGCTATGGTGGTGGTGGAATGTAGCTTC	-	0.0057	0.1168	0.6393
20	AGTTATTCGGTGGTGGTGGATGTGATAATT	0.0006	0.0070	0.1452	0.6114
21	AGTAATTTGGTGGTGGTGGTAAAAATTACT	-	0.0090	0.1310	0.5662
22	AGTAATTATTGGTGGTGGTGGATATGTACT	0.0006	0.0094	0.1541	0.5555
23	GAGCTTAAAGGTGGTGGTGGTTATCAGCTC	-	0.0049	0.1295	0.5335
24	GGTGCACTTTGGTGGTGGTGGATT	0.0006	0.0037	0.0627	0.5270

	GTAGGT				
25	ATTATTCACAGGTGGTGGTGGTAG AATAAT	-	0.0078	0.1209	0.5227
26	GTACGATAGGGTGGTGGTGGAATT ATCGTA	0.0006	0.0062	0.1362	0.5147
27	ATAAAATAGGTGGTGGTGGAAATG TGCAGT	-	0.0053	0.1086	0.5120
28	TGTAAAATGGTGGTGGTGGATATG TTTACA	0.0006	0.0041	0.0884	0.4948
29	ACTATTCAGGTGGTGGTGGTTATG AATAGT	-	0.0062	0.1161	0.4803
30	TTGTCTAACGGTGGTGGTGGATTA GACTAA	0.0006	0.0086	0.2000	0.4792
31	GTTAGAACGGAGGTAAATGGAGGA TTTAAC	0.0006	0.0103	0.1205	0.4685
32	TTAGGATATAATGGTGGTGGTGGA ACCTAA	0.0006	0.0082	0.1347	0.4120
33	ATGCTCATCTGGTGGTGGTGGTAG GAGTAT	0.0011	0.0090	0.1086	0.4115
34	TCATGGGGTACTTAAATAGGTGGC GGGAGA	-	0.0279	0.3758	0.3906
35	TCGATTTACTGGTGGTGGGTGGAT TTTCGA	-	0.0111	0.1355	0.3873
36	AGAACACGGGTGGTGGTGGATGAT TGTTTT	-	0.0053	0.1168	0.3723
37	TTGATTTTGATGGTGGTGGTGGAC AGTCAA	-	0.0082	0.1138	0.3675
38	TCGAATGGGGGACAGGGAGTGGGA TTTCGA	0.0006	0.0139	0.1683	0.3599
39	ATGTTATCAGGTGGTGGTGGAAGT GAACAT	-	0.0033	0.0832	0.3519
40	GGTCACGTTTTGGTGGTGGTGGTT TAACGT	-	0.0021	0.1011	0.3470

Table S4 K_D of binding between L-Apt.T8/L-Apt.T8-10D and rG4s/rG4-ext.

Aptamer/conjugate	**Binding target**	**K_D (nM)**
L-Apt.T8	U3-III rG4	58.1 ± 4.90
	APP rG4	17.9 ± 2.58
	hTERC rG4	11.7 ± 1.86
	BCL2 rG4	8.33 ± 1.54
L-Apt.T8-10D	U3-III-ext	12.5 ± 5.17
	APP-ext	125 ± 20.3
	hTERC-ext	152 ± 70.5
	BCL2-ext	86.6 ± 9.83

* K_D of Fig. 3B (L-Apt.T8) and 3F (L-Apt.T8-10D) binding curves.

Table S5 Sequence of peptide used in this study.

Name	**Sequence (N' to C')**	**Amino acids**
HIV-1 Nucleocapsid (NC)	IQKGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKG CWKCGKEGHQMKDCTERQAN	55

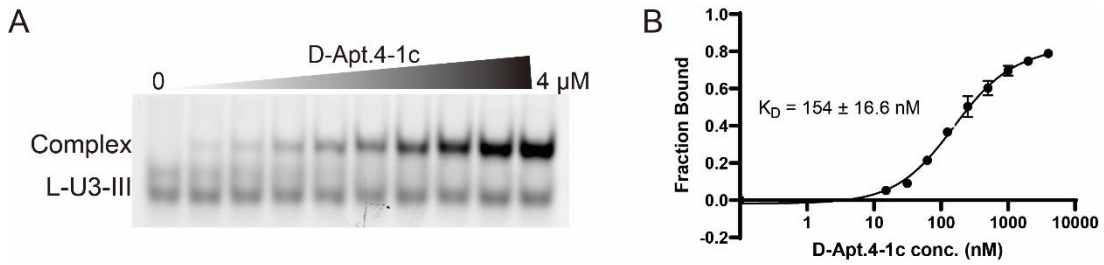

Fig. S1 Binding between D-Apt.4-1c and L-U3-III monitored by EMSA. (A) representative gel. (B) Binding curve generated with three replicates with the K_D of 154 ± 16.6 nM. (Error bars: standard deviation) Binding between D-Apt.4-1c and L-*hTERC* rG4 was with the K_D of 46.0 ± 4.2 nM[1], which was much lower than the K_D of binding between D-Apt.4-1c and L-U3-III, indicating less effective binding.

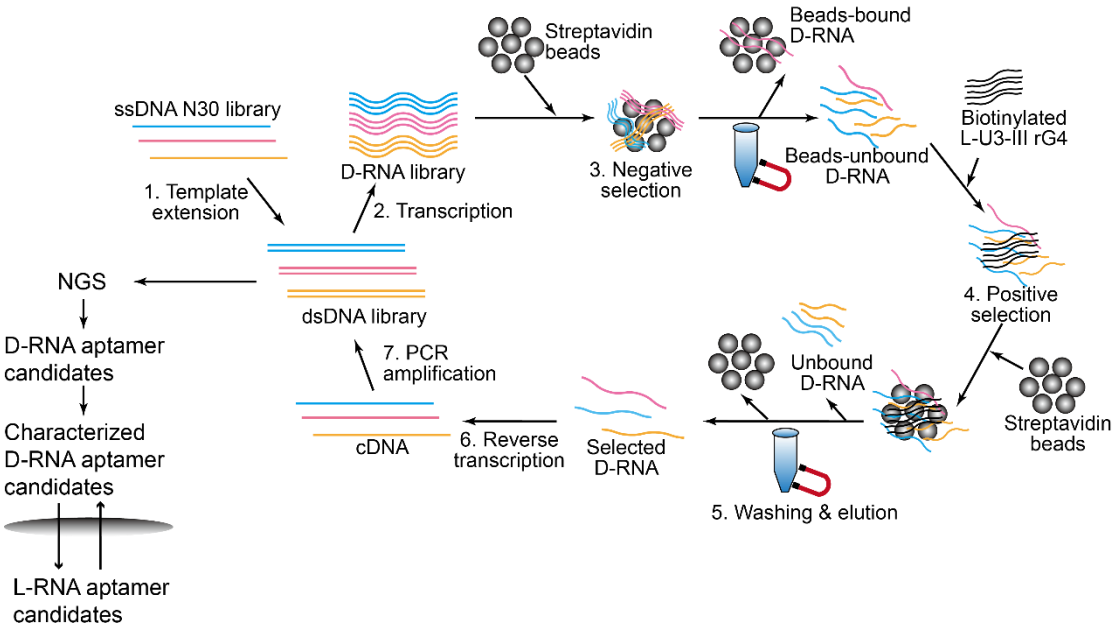

Fig. S2 Schematic of G4-SELEX-seq. For amplification, target G4 is converted to L-form and the RNA pool remains D-form during *in vitro* selection. During the *in vitro* selection, D-RNA candidates that binds to the L-U3-III target has been identified and characterized. After sequencing and initial characterization, the selected candidate is chemically synthesized in L-RNA form and applied for further experiments.

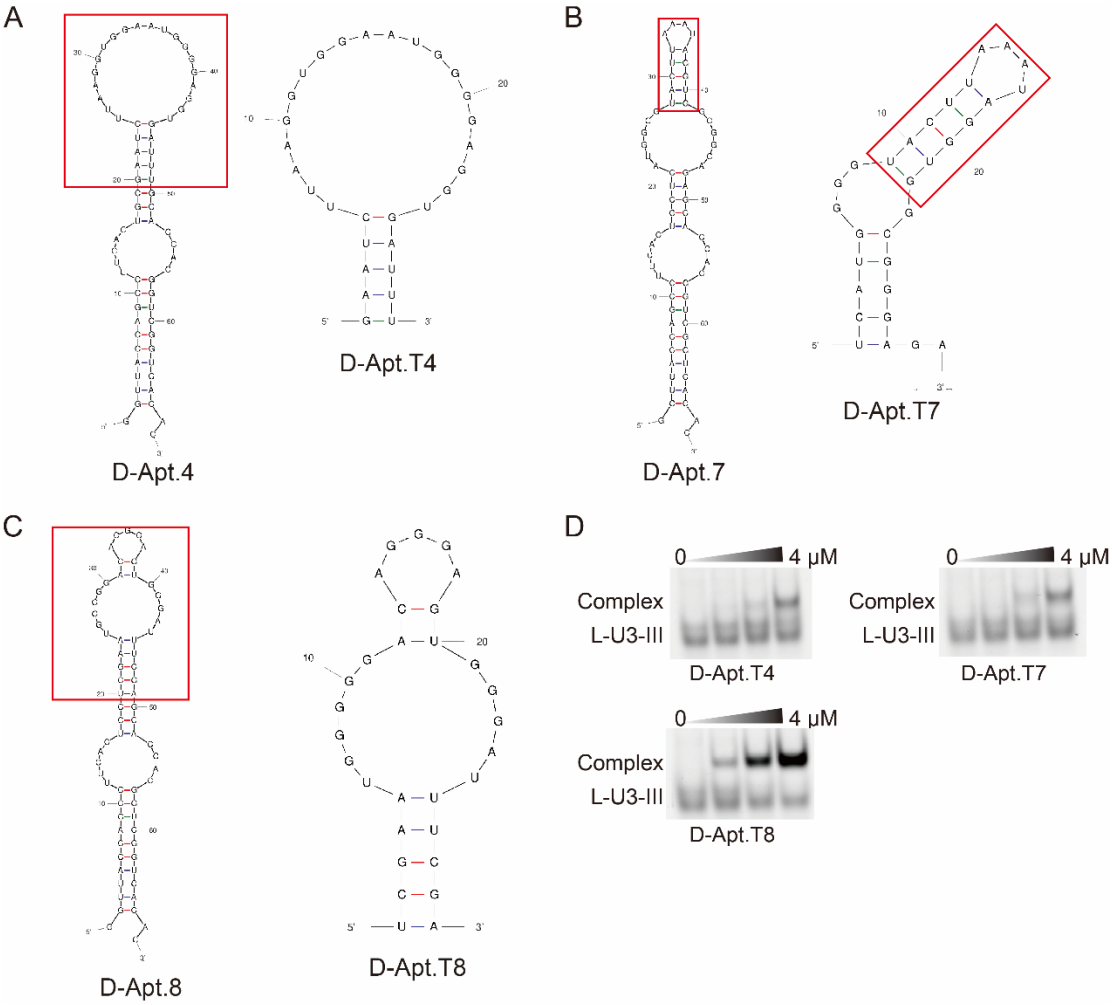

Fig. S3 Mfold predicted secondary structures of chosen candidates and their initial binding test with L-U3-III rG4 target. (A) Predicted secondary structure of D-Apt.4 (full-length) and D-Apt.T4 (adapter-truncated). (B) Predicted secondary structure of D-Apt.7 (full-length) and D-Apt.T7 (adapter-truncated). (C) Predicted secondary structure of D-Apt.8 (full-length) and D-Apt.T8 (adapter-truncated). The consistent structures in the presence or absence of linker predicted are boxed in red. (D) Initial binding test between D-candidates and L-U3-III rG4 monitored by EMSA. All three candidates showed binding to L-U3-III rG4, while D-Apt.T8 showed the strongest binding and therefore picked for further study.

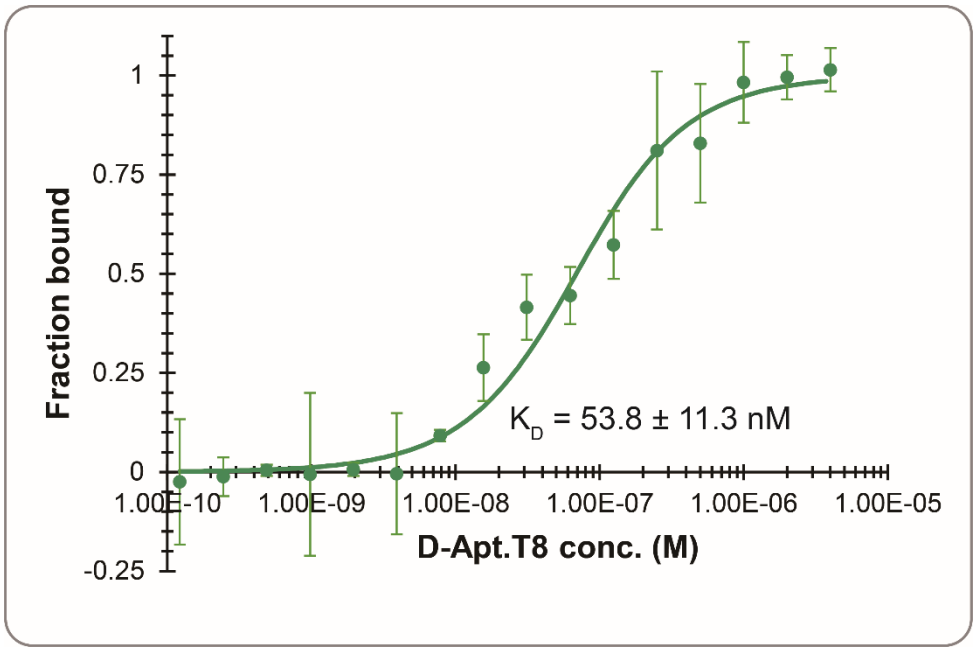

Fig. S4 Binding between D-Apt.T8 and L-U3-III rG4 monitored by MST. The K_D obtained is 53.8 ± 11.3 nM. Error bars stand for standard deviation with three replicates.

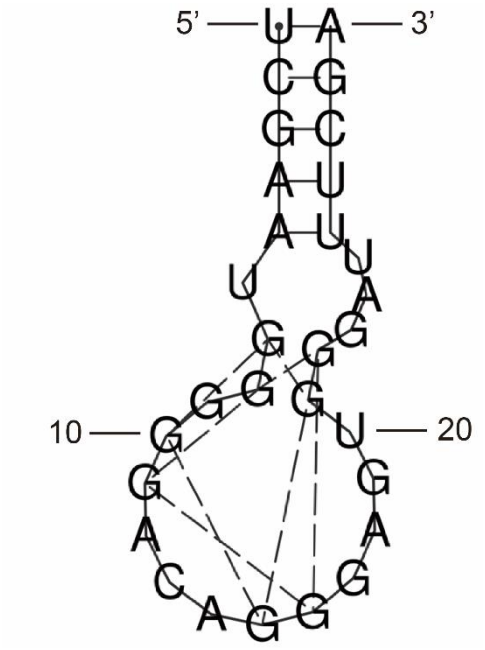

Fig. S5 Predicted G4 formation in Apt.T8. A G4 with two layers of G-quartet was predicted to form in Apt.T8 by Vienna RNAfold webserver[3].

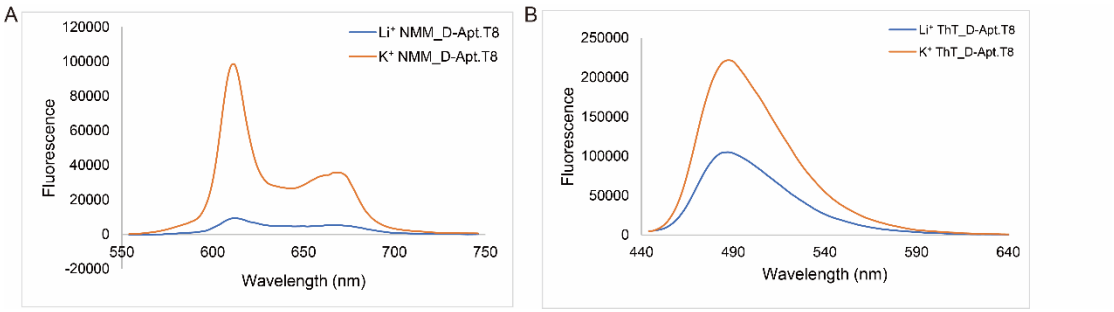

Fig. S6 Ligand-enhanced fluorescence test of D-Apt.T8 with NMM and ThT. NMM and ThT are G4 ligands with enhanced fluorescence during G4 binding[4]. In both NMM (A) and ThT (B) enhanced fluorescence test of D-Apt.T8, the signal obtained under 150 mM K^+ is much stronger than that under 150 mM Li^+, suggesting the formation of G4 in D-Apt.T8.

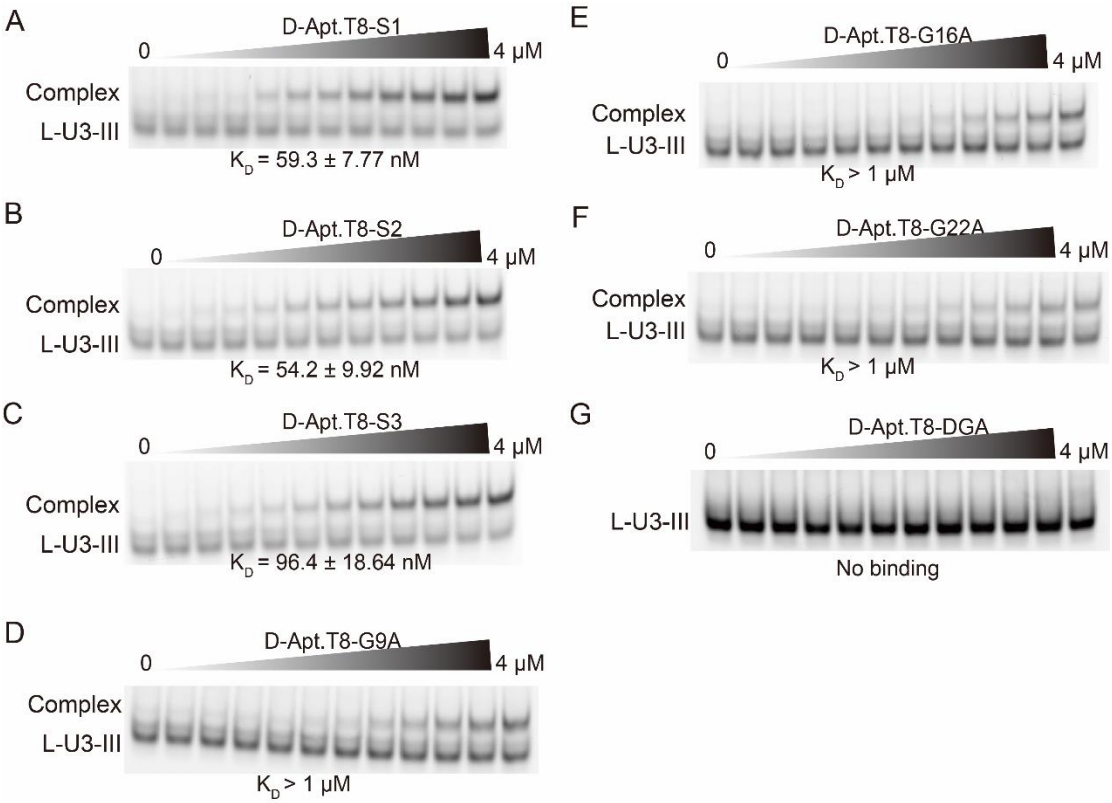

Fig. S7 Representative gels and K_D of binding between mutants of D-Apt.T8-10D and L-U3-III monitored by EMSA. The mutants, S1 (A), S2 (B), S3 (C), G9A (D), G16A (E), G22A (F), DGA (G) were binding to fixed concentration of L-U3-III with increasing concentrations. The binding plots are shown in Fig. 2D.

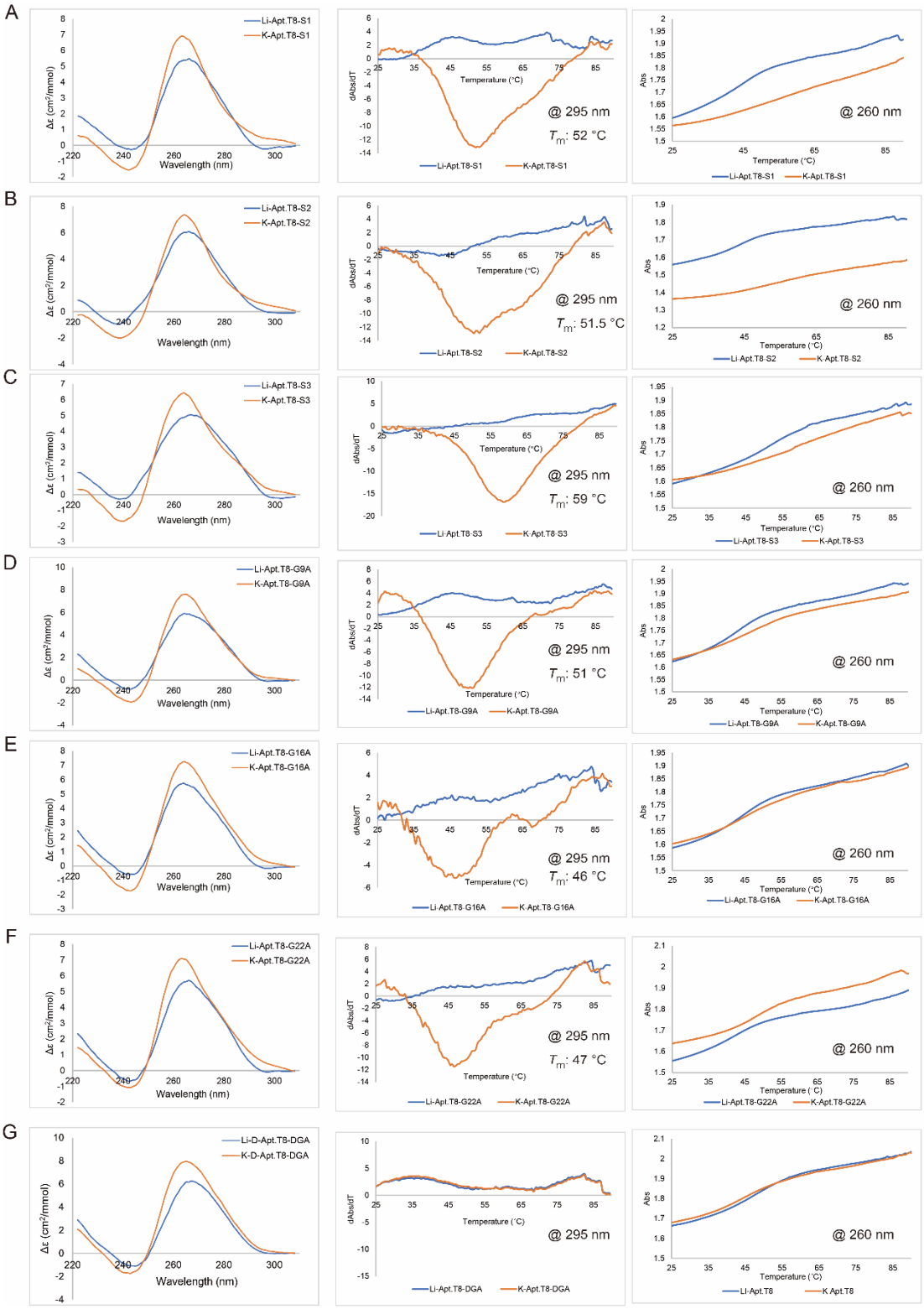

Fig. S8 Spectroscopic analysis of G4 formation in D-Apt.T8 mutants. The CD spectra, UV-melting curves monitored at 295 nm and 260 nm of Apt.T8-S1 (A), Apt.T8-S2 (B), Apt.T8-S3 (C), Apt.T8-G9A (D), Apt.T8-G16A (E), Apt.T8-G22A (F), and Apt.T8-DGA (G) were generated from the same sequence in 150 mM Li^+ (blue) or 150 mM K^+ (orange). In (A) to (F), the stronger peaks of K^+ samples in CD spectra and negative peak at ≥ 45 °C in UV-melting curves in K^+ but not Li^+ monitored at 295 nm, as well as

the similar trends of K^+ and Li^+ samples in UV-melting curves monitored at 260 nm suggested the formation of parallel G4 in stem mutations and single G to A mutations. However, in (G), even with the slightly stronger peak of K^+ sample in CD spectra, the same trend of UV-melting curves monitored at both 295 nm and 260 nm suggested no G4 forming in the double G to A (DGA) mutation.

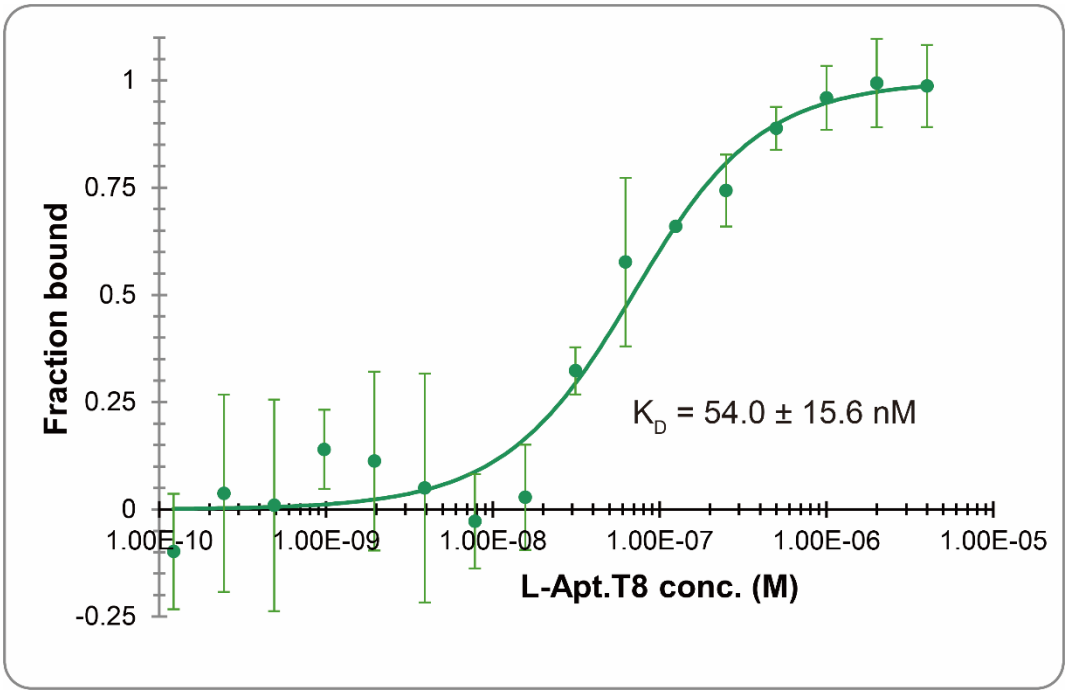

Fig. S9 L-Apt.T8 - U3-III rG4 binding monitored by MST. The K_D obtained is 54.0 ± 15.6 nM. Error bars stand for standard deviation with three replicates.

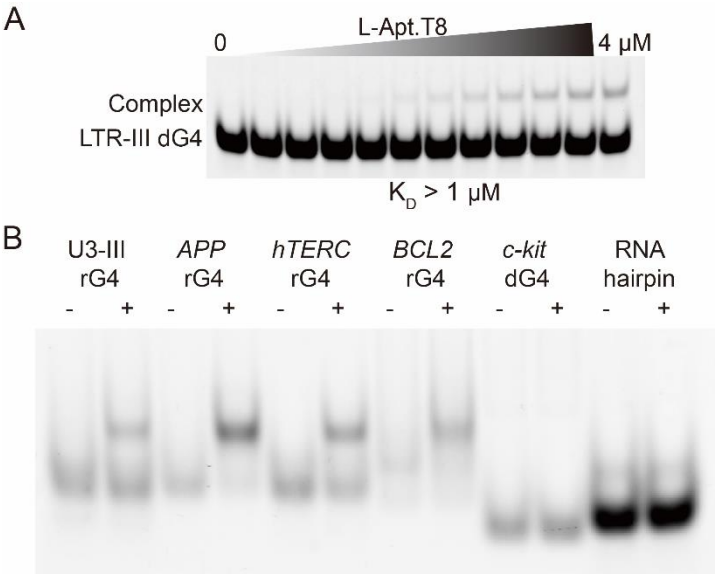

Fig. S10 Binding selectivity test of L-Apt.T8 monitored with EMSA. (A) L-Apt.T8 showed weak binding to LTR-III dG4. (B) Selectivity test of L-Apt.T8 with control targets including canonical rG4s, canonical dG4 and RNA hairpin. L-Apt.T8 binds to all tested rG4s, but no binding is detected to dG4 and RNA hairpin.

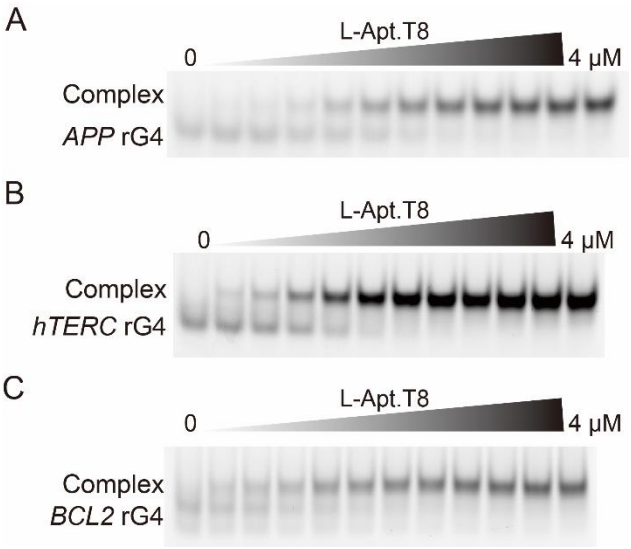

Fig. S11 Representative gels of binding between L-Apt.T8 and control rG4s monitored by EMSA. L-Apt.T8 showed strong binding to *APP* rG4 (A), *hTERC* rG4 (B) and *BCL2* rG4 (C). All the binding tests are performed with the same condition. The binding plots are shown in Fig. 3B and K_D in Table S4.

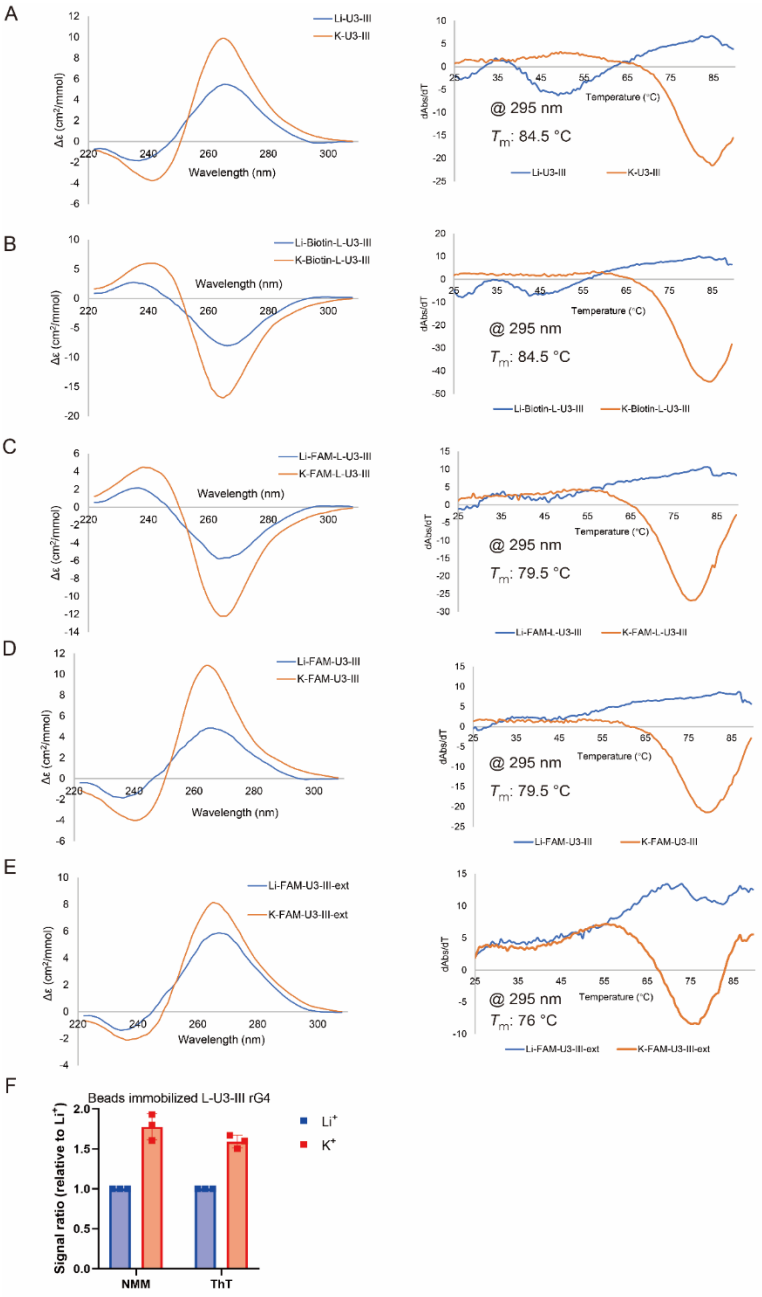

Fig. S12 Spectroscopic analysis of U3-III rG4 formation with labels and extension and during immobilization. The CD spectra and UV-melting curves monitored at 295 nm of U3-III rG4 (A), biotin L-U3-III rG4 (B), FAM L-U3-III rG4 (C), FAM U3-III rG4 (D), FAM U3-III-ext (E), were generated from the same sequence in 150 mM Li^+ (blue) or 150 mM K^+ (orange). In (A) to (E), the stronger peaks of K^+ samples in CD spectra and negative peak at ≥ 76 °C in K^+ UV-melting curves monitored at 295 nm suggested the formation of parallel G4 in all labelled U3-III rG4 and U3-III-ext. As the mirror conformation of natural D-form U3-III rG4, L-U3-III rG4 was with the symmetrical CD spectrum. (F) Ligand enhanced fluorescence test of beads immobilized L-U3-III rG4. The stronger signal of K^+ samples than Li^+ samples incubated with both NMM and ThT suggested the formation of G4 in beads immobilized L-U3-III rG4.

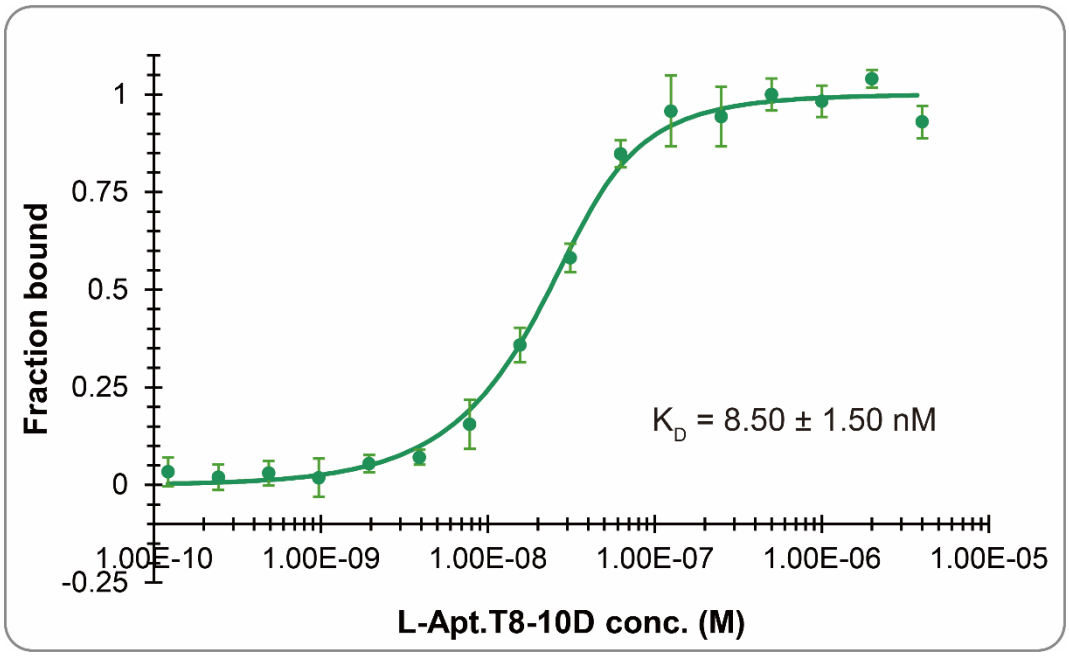

Fig. S13 Binding between L-Apt.T8-10D and U3-III-ext monitored by MST. The K_D obtained is 8.50 ± 1.50 nM. Error bars stand for standard deviation with three replicates.

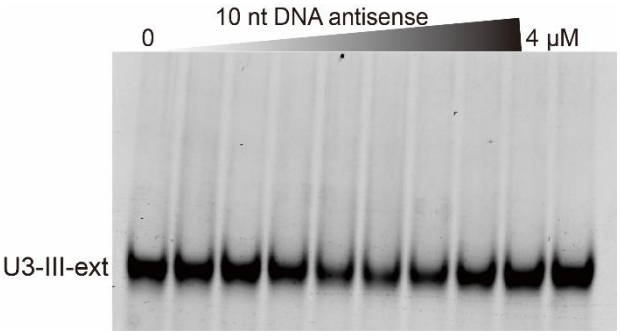

Fig. S14 Binding between 10 nt DNA antisense and U3-III-ext monitored with EMSA. No observable binding is detected from the gel.

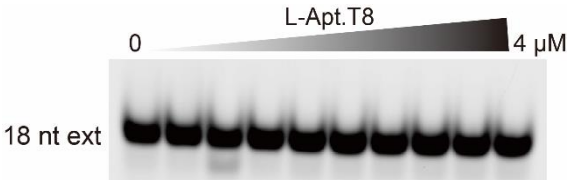

Fig. S15 Binding between L-Apt.T8 and the 18 nt extension region monitored with EMSA. No observable binding is detected from the gel.

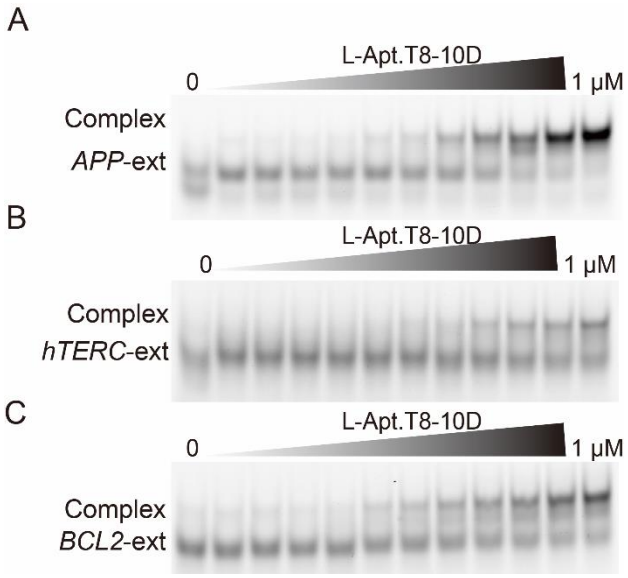

Fig. S16 Binding between L-Apt.T8-10D and control rG4-exts monitored by EMSA. L-Apt.T8-10D showed weak binding to *APP*-ext (A), *hTERC*-ext (B) and *BCL2*-ext (C). All the binding tests are performed with the same condition. The binding plots are shown in Fig. 3F and K_D in Table S4.

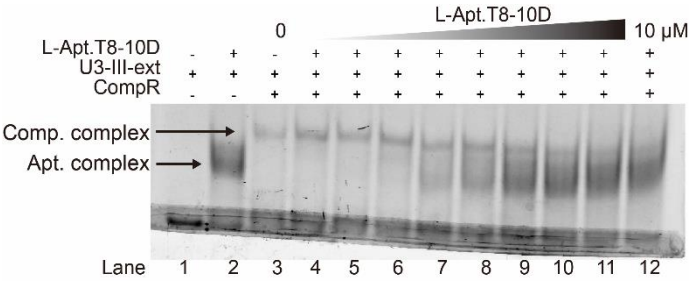

Fig. S17 L-Apt.T8-10D binding interference monitored by EMSA. Half of the samples from the same reactions shown in Fig. 4B were loaded to 10% native gel and run for longer time for better separation of bands and gel image analysis.

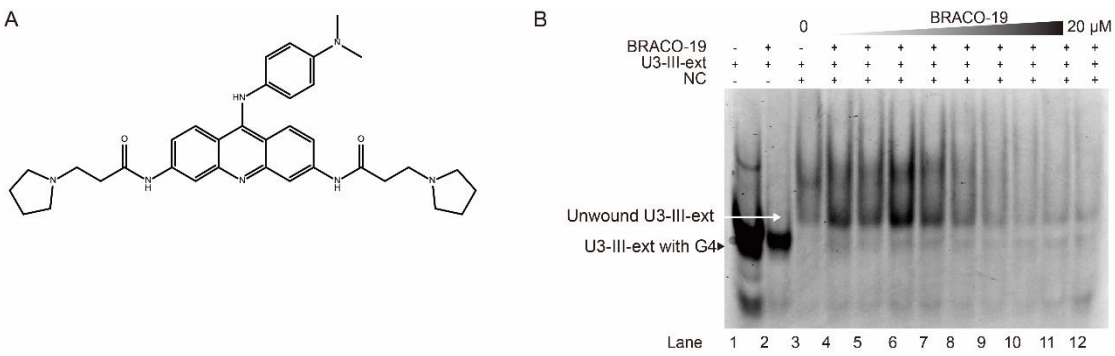

Fig. S18 BRACO-19 inducing NC unwinding inhibition at high concentration. (A) Structure of BRACO-19. (B) Result of BRACO-19 inducing NC unwinding inhibition monitored by EMSA. With the increase of BRACO-19 concentration, the bands of unwound U3-III-ext become fainter, indicating inhibition in NC unwinding of U3-III-ext with the presence of more BRACO-19 in the reaction.

References

1. M. I. Umar and C. K. Kwok, *Nucleic Acids Res.*, 2020, **48**, 10125-10141.
2. D. Piekna-Przybylska, L. DiChiacchio, D. H. Mathews and R. A. Bambara, *Nat. Struct. Mol. Biol.*, 2010, **17**, 83-89.
3. R. Lorenz, S. H. Bernhart, C. H. Zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler and I. L. Hofacker, *Algorithms Mol. Biol.*, 2011, **6**, 26.
4. D. Zhao, X. Dong, N. Jiang, D. Zhang and C. Liu, *Nucleic Acids Res.*, 2014, **42**, 11612-11621.