Supporting Information for

Topological effect of intramolecular split G-quadruplex on Thioflavin T binding and fluorescence light-up

Mengmeng Lv,^{a,b} Jiangtao Ren,^{b*} Erkang Wang^{a,b*}

^a College of Chemistry, Jilin University, Changchun, Jilin 130012, China.

^b State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry,

Chinese Academy of Sciences, Changchun, Jilin 130022, China.

* Corresponding author E-mail: jiangtaoren@ciac.ac.cn; ekwang@ciac.ac.cn

Table of contents:

- Oligonucleotides used in this study. (Table S1)
- The gene fragments of this work (**Table S2**).
- Eleven Intra-SG structures (S1-S11) derived from T30695 G-quadruplex. (Figure S1)
- Fluorescence responses of ThT with different types of DNA strands (S8, C/S8 or C/C').

(Figure S2)

UV-vis spectra of ThT incubated with different Intra-SG strands or C/Intra-SG hybrids.

```
(Figure S3)
```

- Fluorescence spectra of different ThT/Intra-SG systems (*S8Ha'*, *S8Hb'*, *S8Hc'* and *S8Hd'*).
 (Figure S4)
- CD spectra and normalized UV-melting profiles of *S8Ha*' without or with *Ha*. (Figure S5)
- The fluorescence intensity of ThT/Intra-SG in the presence of target or different target variants generated through grafting different numbers of A or T bases at the 5' or 3' end.

(Figure S6)

■ The fluorescence intensity of ThT/*S*8*Ha*′ with *R11* or *R11* variants generated through grafting different numbers (*k*) of T bases at the 3′ terminus (*RTk*). (**Figure S7**)

- The fluorescence intensity of ThT/*S*8*Ha*' in the presence of *Ha* or different *Ha* variants generated through tailoring different numbers of bases from the 5' end (*Ly*). (Figure S8)
- Fluorescence spectra of different ThT/Intra-SG systems (*S8Ra'*, *S8Rb'*, *S8Rc'*, *S8Rd'*, *S8Re'* and *S8Rf*). (Figure S9)
- Fluorescence data on the topological effect study of ThT/S8Ra' system. (Figure S10)

Experimental section

Asymmetric PCR amplification (A-PCR). The gene sequences of Hepatitis B (GenBank accession number AB116094), Ebola (GenBank accession number KY786026.1) and Nova (GenBank accession number KB270442.1) viruses can be retrieved from the GenBank database of the National Center for Biotechnology Information. A-PCR primers including a forward primer (*FP*) and a reverse primer (*RP*) were designed for HBV fragment (*HF*) using the Snap gene sequence analysis and design software.

As for the developed LABP (Figure 6D-E), an optimum molar ratio of 10:1 was adopted. Each A-PCR sample (20 μ L) was prepared by mixing 10 μ L of Taq PCR Master Mix consisting of Taq DNA polymerase, dNTPs mix and MgCl₂, 3.6 μ L of *FP* (20 μ M), 1.8 μ L of *RP* (4 μ M), and 5 μ L of target template (*HF*, *EF* or *NF*) of different concentrations. The template was replaced by buffer as a negative control. All the reaction mixtures were subjected to repeated cycles of 10 s at 98°C, 15 s at 60°C, and 15 s at 72°C, on an Applied Biosystems Veriti Thermal Cycler (Life Technologies Holdings Pte Ltd., Singapore). The A-PCR conditions used included an initial denaturation step at 98°C for 30 s, followed by 35 cycles of denaturation at 98°C for 10 s, annealing at 60°C for 15 s, and an extension at 72°C for 15 s. In the end, there was a final extension step at 72°C for 7 min. The A-PCR amplicons were subjected to subsequent ligation and RCA reaction.

Ligation and RCA reaction. 20 μ l of A-PCR product was mixed with 2 μ l of T4 DNA ligase reaction buffer (400 mM Tris-HCl, pH 7.8, 100 mM MgCl₂, 100 mM DTT, 5 mM ATP and 50% PEG4000), 2.5 μ l of padlock (*Pa*, 3.2 μ M) and 0.75 μ l of ddH₂O.

After annealing for 5 min at 75°C, 2 μ l of T4 DNA ligase (4 U/ μ l) was added into the mixture, and the ligation reaction proceeded for 2 h at 37°C. Subsequently, the polymerization reaction was initiated by adding 1.75 μ L of dNTPs (10 mM), 2 μ L of phi29 DNA polymerase (5 U/ μ L) and 4 μ L of phi29 buffer (330 mM Tris-HCl, 660 mM KCl, 100 mM MgCl₂, 1% (v/v) Tween20 and 10 mM DTT, pH 7.5) into the ligation solution. After 2 h at 37°C, it was terminated through a thermal treatment for 5 min at 95°C.

Signal transduction. 44 μ L of Tris/Mg buffer (25 mM Tris-HCl buffer, 10 mM Mg²⁺, pH 8.0) was added to each RCA sample. After sonication for 10 min at room temperature, *S8Ra'* (10 μ L, 9 μ M) in Tris/K buffer and *Rv/BL* conversion probe (5 μ L, 24 μ M) in Tris/Mg buffer were added and incubated for 1 h at room temperature. Finally, the solution was mixed with ThT (6 μ L, 250 μ M) and 204 μ L of Tris/Mg buffer with 30 mM K⁺, and then fluorescence spectra were collected after 20 min of incubation at room temperature.

Gel electrophoresis. As for agarose gel electrophoresis in Figure 6A, an agarose gel (3% (w/v)) was prepared in 1×TAE buffer (40 mM Tris-acetic acid, 2 mM EDTA, pH 7.5).10 μ L of each A-PCR sample was mixed with 4 μ L of loading buffer (1×TAE with 36% glycerol) and loaded into the gel. As for denaturing polyacrylamide gel electrophoresis in Figure 6B, a polyacrylamide gel (12% (w/v)) was prepared within 1×TBE buffer with urea (89 mM Tris, 89 mM boric acid, 2 mM EDTA and 7 M urea, pH 8.3). 5 μ L of each ligation sample was mixed with 2 μ L of loading buffer (1×TBE with 7 M urea and 36% glycerol, pH 8.3) and loaded into the gel. As for native polyacrylamide gel electrophoresis in Figure 6C, a polyacrylamide gel (12% (w/v)) was prepared within 1×TBE buffer. 5 μ L of each sample was mixed with 2 μ L of loading buffer (1×TBE with 36% glycerol, pH 8.3) and loaded into the gel. All the gels were run at 60 V for 80-100 min, stained with Gel-Red (0.2‰) for 40 min and photographed using a fluorescence imaging system, iBrightTM FL1000 (Thermo Fisher Scientific, USA).

 Table S1. Oligonucleotides used in this study. (Note: Spacer sequences are indicated by italic letter. The mutated bases were indicated by red letters.)

Oligo	Sequence (5' to 3')
T30695	GGGTGGGTGGGTGGGT
SI	G TTCTTTTCTTTTCTTTCTT GGTGGGTGGGTGGGT
S2	GG TTCTTTTCTTTCTTTCTT GTGGGTGGGTGGGT
<i>S3</i>	GGG TTTCTTTCTTTCTTTCTT GGGTGGGTGGGT
<i>S4</i>	GGGTG TTCTTTCTTTCTTTCTT GGTGGGTGGGT
<i>S5</i>	GGGTGG TTCTTTTCTTTCTTTCTT GTGGGTGGGT
<i>S6</i>	GGGTGGGT TTCTTTTCTTTCTTTCTT GGGTGGGT
<i>S7</i>	GGGTGGGTG TTCTTTCTTTCTTTCTT GGTGGGT
<i>S8</i>	GGGTGGGTGG TTCTTTTCTTTCTTTCTT GTGGGT
<i>S9</i>	GGGTGGGTGGGT TTCTTTTCTTTTCTTTCTT GGGT
<i>S10</i>	GGGTGGGTGGGTG TTCTTTTCTTTCTTTCTT GGT
<i>S11</i>	GGGTGGGTGGGTGG TTCTTTTCTTTCTTTCTTTCTT GT
<i>C</i> ′	ТТСТТТТСТТТТСТТ
С	AAGAAAAGAAAAGAAAAGAA
S8Ha'	GGGTGGGTGG TCCTAGGAATCCTGATGTGA GTGGGT
На	TCACATCAGGATTCCTAGGA
S8Hb'	GGGTGGGTGG TAAGAAGATGAGGCATAGCA GTGGGT
Hb	TGCTATGCCTCATCTTCTTA
S8Hc'	GGGTGGGTGGAGAAGTCCACCACGAGTCTAGTGGGT
Нс	TCCTCACAATACCGCAGAGT
S8Hd'	GGGTGGGTGGTCCTGGTTATCGCTGGATGT GTGGGT
Hd	ACATCCAGCGATAACCAGGA
5HT1	tTCACATCAGGATTCCTAGGA
5HT2	ttTCACATCAGGATTCCTAGGA
5HT4	ttttTCACATCAGGATTCCTAGGA
5HT6	ttttttTCACATCAGGATTCCTAGGA
5HT8	tttttttTCACATCAGGATTCCTAGGA

5HT10	tttttttttTCACATCAGGATTCCTAGGA
3HT1	TCACATCAGGATTCCTAGGAt
<i>3HT2</i>	TCACATCAGGATTCCTAGGAtt
<i>3HT4</i>	TCACATCAGGATTCCTAGGAtttt
<i>3HT6</i>	TCACATCAGGATTCCTAGGAtttttt
<i>3HT8</i>	TCACATCAGGATTCCTAGGAttttttt
3HT10	TCACATCAGGATTCCTAGGAttttttttt
5HA1	aTCACATCAGGATTCCTAGGA
5HA2	aaTCACATCAGGATTCCTAGGA
5HA4	aaaaTCACATCAGGATTCCTAGGA
5HA6	aaaaaaTCACATCAGGATTCCTAGGA
5HA8	aaaaaaaTCACATCAGGATTCCTAGGA
5HA10	aaaaaaaaaTCACATCAGGATTCCTAGGA
3HA1	TCACATCAGGATTCCTAGGAa
3HA2	TCACATCAGGATTCCTAGGAaa
3HA4	TCACATCAGGATTCCTAGGAaaaa
<i>3HA6</i>	TCACATCAGGATTCCTAGGAaaaaaa
<i>3HA8</i>	TCACATCAGGATTCCTAGGAaaaaaaaa
3HA10	TCACATCAGGATTCCTAGGAaaaaaaaaaaaaaaaaaaaa
5 <i>ST1</i>	tAAGAAAAGAAAAGAAAAGAA
5 <i>ST2</i>	ttAAGAAAAGAAAAGAAAAGAA
5 <i>ST</i> 4	ttttAAGAAAAGAAAAGAAAAGAA
5 <i>ST</i> 6	ttttttAAGAAAAGAAAAGAAAAGAA
5 <i>ST</i> 8	ttttttttAAGAAAAGAAAAGAAAAGAA
5ST10	tttttttttAAGAAAAGAAAAGAAAAGAA
3ST1	AAGAAAAGAAAAGAAAAGAAt
3 <i>ST</i> 2	AAGAAAAGAAAAGAAAAGAAtt
3ST4	AAGAAAAGAAAAGAAAAGAAtttt
3ST6	AAGAAAAGAAAAGAAAAGAAtttttt

3ST8	AAGAAAAGAAAAGAAAAGAAtttttttt
3ST10	AAGAAAAGAAAAGAAAAGAAtttttttttt
MI	ACACATCAGGATTCCTAGGA
M2	TGACATCAGGATTCCTAGGA
M3	TCTCATCAGGATTCCTAGGA
M4	TCAGATCAGGATTCCTAGGA
M5	TCACTTCAGGATTCCTAGGA
M6	TCACAACAGGATTCCTAGGA
M11	TCACATCAGGTTTCCTAGGA
M16	TCACATCAGGATTCCAAGGA
M20	TCACATCAGGATTCCTAGGT
<i>R6</i>	TCACAT
<i>R7</i>	TCACATC
R9	TCACATCA
R10	TCACATCAG
R11	TCACATCAGG
R16	TCACATCAGGATTCC
R19	TCACATCAGGATTCCTAGG
L6	CTAGGA
L7	CCTAGGA
L9	TTCCTAGGA
L10	ATTCCTAGGA
L11	GATTCCTAGGA
L16	ATCAGGATTCCTAGGA
L19	CACATCAGGATTCCTAGGA
RT5	TCACATCAGGttttt
RT10	TCACATCAGGttttttttt
RT20	TCACATCAGGuuuuuuuu
RT30	TCACATCAGGuuuuuuuuuuuuuuu

S8Ra'	GGGTGGGTGGTTAATAATATCCAGATAGTTGTGGGT
Ra	AACTATCTGGATATTATTAA
S8Rb'	GGGTGGGTGGTCCTGAGACACCTATCTCATGTGGGT
Rb	ATGAGATAGGTGTCTCAGGA
S8Rc'	GGGTGGGTGGTAATAATATACCTATCTCATGTGGGT
Rc	ATGAGATAGGTATATTATTA
S8Rd'	GGGTGGGTGGTCCTATGACTCTCATTGTGAGTGGGT
Rd	TCACAATGAGAGTCATAGGA
S8Re'	GGGTGGGTGG <i>TCATTACAGTCCACATAGTT</i> GTGGGT
Re	AACTATGTGGACTGTAATGA
S8Rf	GGGTGGGTGGTCCTATGACTTAGTTTGTGAGTGGGT
Rf	TCACAAACTAAGTCATAGGA
5RT1	tAACTATCTGGATATTATTAA
5RT2	ttAACTATCTGGATATTATTAA
5RT4	ttttAACTATCTGGATATTATTAA
5RT6	ttttttAACTATCTGGATATTATTAA
5RT8	tttttttAACTATCTGGATATTATTAA
5RT10	tttttttttAACTATCTGGATATTATTAA
3RT1	AACTATCTGGATATTATTAAt
3RT2	AACTATCTGGATATTATTAAtt
3RT4	AACTATCTGGATATTATTAAtttt
3RT6	AACTATCTGGATATTATTAAtttttt
3RT8	AACTATCTGGATATTATTAAttttttt
3RT10	AACTATCTGGATATTATTAAttuttutt
5RA1	aAACTATCTGGATATTATTAA
5RA2	aaAACTATCTGGATATTATTAA
5RA4	aaaaAACTATCTGGATATTATTAA
5RA6	aaaaaaAACTATCTGGATATTATTAA
5RA8	aaaaaaaAACTATCTGGATATTATTAA

5RA10	aaaaaaaaAACTATCTGGATATTATTAA
3RA1	AACTATCTGGATATTATTAAa
3RA2	AACTATCTGGATATTATTAAaa
3RA4	AACTATCTGGATATTATTAAaaaa
3RA6	AACTATCTGGATATTATTAAaaaaaa
3RA8	AACTATCTGGATATTATTAAaaaaaaaaa
3RA10	AACTATCTGGATATTATTAAaaaaaaaaaaaaaaaaaaaa
MRI	TACTATCTGGATATTATTAA
MR2	ATCTATCTGGATATTATTAA
MR3	AAGTATCTGGATATTATTAA
MR4	AACAATCTGGATATTATTAA
MR5	AACTTTCTGGATATTATTAA
MR6	AACTAACTGGATATTATTAA
MR7	AACTATGTGGATATTATTAA
MR9	AACTATCTCGATATTATTAA
MR10	AACTATCTGCATATTATTAA
MR11	AACTATCTGGTTATTATTAA
MR16	AACTATCTGGATATTTTAA
MR20	AACTATCTGGATATTATTAT
D6	AACTAT
D7	AACTATC
D9	AACTATCTG
D10	AACTATCTGG
DII	AACTATCTGGA
D16	AACTATCTGGATATTA
D19	AACTATCTGGATATTATTA
RD5	AACTATCTGGAttttt
RD10	AACTATCTGGAttttttttt
RD20	AACTATCTGGAttttttttttttttttt

RD30	AACTATCTGGAttttttttttttttttttttttttttttt
Rv	AACTATCTGGAATGAT
FP	ACCGAACATGGAGAACATCACA
RP	GAGGACAAACGGGCAACATAC
SC	GGACCCTGCACCGAACATGGAGAACATCACATCAGGATTCCTAGG
	ACCCCTGCTCGTGTTACAGGCGGGGTTTTTC
Pa	Phosphate-
	CTGATGTGATGTTCTTTCCTTTCTCATGTTTAGATAGTTCGAGACAC
	TGACATTTGGTCCTAGGAATC
BL	ATCATTGTAGATAGTTCGAGACACTGACA

Table S2. The sequences of three double-stranded DNA targets (*HF*, *EF* and *NF*)derived from genes of Hepatitis B, Ebola and Nova viruses, respectively.

Name	Information (5' to 3')
HF	AATCTCCTCGAGGACTGGGGACCCTGCACCGAACATGGAGAA
	CATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCG
	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTC
	TAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGGATCACC
	CGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCAC
	TCACCAACCTCCTGTCCTCCAATTTGTCCTGGTTATCGCTGGAT
	GTGTCTGCGGCGTTTTATCATATTCCTCTTCATCCTGCTGCTAT
	GCCTCATCTTCTTATTGGTTCTTCTGGATTATCAGGGTATGTTG
	CCCGTTTGTCCTCT
EF	CTTCAGCTCACACCCCTTGAGAGAGCCGGTCAATGCAACGGA
	GGACCCGTCCAGTGGCTACTATTCTACCACAATTAGATATCAG
	GCTACCGGTTTTGGAACCAATGAGACGGAGTACTTGTTCGAG
	GTTGACAATTTGACCTACGTCCAACTTGAATCAAGATTCACGC
	CACAGTTTTTGCTCCAGCTGAATGAGACAATATATGCAAGTGG
	GAAAAGGAGCAACACCACGGGAAAACTAATTTGGAAGGTCA
	ACCCCGAAATTGATACAACAATCGGGGAGTGGGCCTTCTGGG

	AAACTAAAAAAACCTCACTAGAAAAATTCGCAGT
NF	TACTCCTGATTTTGAGTTCACTTATTTAGTGCCACCTTCTGTTG
	AATCTAAAACCAAGCCTTTTTCCTTACCTATTTTAACCCTTTCT
	GAGCTCACAAATTCGAGGTTCCCAGTCCCCATCGATTCGCTTT
	TCACCGCCCAGAATAATGTGTTGCAGGTGCAGTGTCAAAATG
	GCAGGTGTACACTTGATGGTGAGTTACAAGGCACAACCCAGT
	TGCTCCCATCTGGCATCTGTGCATTCAGAGGACGGGTGACAGC
	ACAAATTAACCAACGTGACAGGTGGCACATGCAACTGCAAAA
	CCTCAATGGTACAACATATGACCCAACTGATGATGTGCCAGC
	CCCGCTGG



Figure S1. Eleven Intra-SG structures (S1-S11) derived from T30695 G-quadruplex.



Figure S2. Fluorescence responses of ThT with different types of DNA strands (*S8*, *C/S8* or *C/C'*). The concentration of each DNA strand and ThT was 300 nM and 5 μ M, respectively, in Tris/K buffer.



Figure S3. (A-H) UV-vis spectra of ThT incubated with different Intra-SG strands (blue curves) or *C*/Intra-SG hybrids (red curves). The black curves represent ThT only. The final concentrations of ThT and DNA were 2.5 μ M and 5 μ M, respectively, in

Tris/K buffer.



Figure S4. Fluorescence spectra of different ThT/Intra-SG systems. Four new Intra-SG structures (*S8Ha'*, *S8Hb'*, *S8Hc'* and *S8Hd'*) derived from *S8* were designed for recognizing four short DNA segments (*Ha*, *Hb*, *Hc* and *Hd*, respectively) from HBV gene. The calculated R values were indicated in the graphs. The concentration of each DNA strand and ThT in Tris/K buffer were 300 nM and 5 μ M, respectively.



Figure S5. (A) CD spectra of *S8Ha'* without (black curve) or with (red curve) *Ha*. The blue curve represents the duplex *Ha/Ha'*. The final concentration of each indicated strand in Tris/K buffer was 7 μ M. (B) Normalized UV-melting profiles of *S8Ha'* at 295 nm (red curve) and duplex *Ha/Ha'* at 260 nm (black curve). The corresponding T_m values were shown in the inserted table. The concentration of each DNA strand in Tris/K buffer was 2 μ M.



Figure S6. (A-B) The fluorescence intensity of ThT/*S8Ha*' (Blank) in the presence of *Ha* or different *Ha* variants generated through grafting different numbers (m, n) of A bases at the 5' (*5HAm*, A) or 3' end (*3HAn*, B). The value. (C-H) The fluorescence intensity of ThT/*S2* (Blank, C-D), ThT/*S5* (Blank, E-F), and ThT/*S8* (Blank, G-H), in

the presence of *C* or different *C* variants generated through grafting different numbers (m, n) of T bases at the 5' (*5CTm*, C, E and G) or 3' end (*3CTn*, D, F and H). The concentrations of ThT and each DNA strand were 5 μ M and 300 nM, respectively, in Tris/K buffer.



Figure S7. The fluorescence intensity of ThT/*S8Ha'* with *R11* or *R11* variants generated through grafting different numbers (k) of T bases at the 3' terminus (*RTk*). The concentration of each DNA strand and ThT was 300 nM and 5 μ M, respectively, in Tris/K buffer.



Figure S8. The fluorescence intensity of ThT/*S8Ha'* (Blank) in the presence of *Ha* or different *Ha* variants generated through tailoring different numbers of bases from the 5' end (*Ly*). The letter "*y*" in *Ly* represents base numbers of target DNA after tailoring. A schematic diagram is inserted in each figure for illustrating the target DNA tailoring. The concentration of each DNA strand and ThT was 300 nM and 5 μ M, respectively,



Figure S9. Fluorescence spectra of different ThT/Intra-SG systems. These new six Intra-SG structures (*S8Ra'*, *S8Rb'*, *S8Rc'*, *S8Rd'*, *S8Re'* and *S8Rf'*) derived from *S8* were designed for recognizing six random DNA target strands, *Ra*, *Rb*, *Rc*, *Rd*, *Re* and *Rf*, respectively. The calculated R values were indicated in the graphs. The concentration of each DNA strand and ThT was 300 nM and 5 μ M, respectively, in Tris/K buffer.



Figure S10. (A-D) The fluorescence intensity of ThT/*S8Ra'* (Blank) in the presence of *Ra* or different *Ra* variants generated through grafting different numbers (m, n) of A bases (A-B) or T bases (C-D) at the 5' (*5RAm*, A and C) or 3' end (*3RAn*, B and D), or single-base mutation (*MRx*, E), or tailoring different numbers of bases from the 3' end

(*Dy*, F). The letter "*x*" in *MRx* represents the mutation site from the 5' to 3' end. The letter "*y*" in *Dy* represents base numbers of target DNA after tailoring. (G) The fluorescence intensity of ThT/*S8Ra*' with *D11* or *D11* variants generated through grafting different numbers (*k*) of T bases at the 3' terminus (*RDk*). (H) Fluorescence intensity of ThT/*S8Ra*' without or with different targets (*Ra*, *D11* or *Rv*). The concentrations of each DNA strand and ThT were 300 nM and 5 μ M, respectively.