Electronic Supplementary Information (ESI)

Small molecule binding to G-hairpin and G-triplex: A new insight in anticancer drug design targeting G-rich regions

Arivazhagan Rajendran,[‡a] Masayuki Endo,*[bc] Kumi Hidaka,[a] Marie-Paule Teulade-Fichou,[d] Jean-Louis Mergny[e] and Hiroshi Sugiyama*[abc]

[a] *Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa-oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan.*
[b] *Institute for Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Yoshida-ushinomiyacho, Sakyo-ku, Kyoto 606-8501, Japan.*
[c] *CREST, JST, Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan.*
[d] *Institut Curie, UMR 176 CNRS, Campus Universitaire Paris-Sud, 91405 Orsay, France.*
[e] *Univ. Bordeaux, INSERM, U869, ARNA Laboratory, 2 rue Robert Escarpit, Pessac, F-33607, France.*
[‡] *Current address: Institute of Advanced Energy, Kyoto University, Gokasho, Uji-shi, Kyoto-fu, 611-0011, Japan.*

EXPERIMENTAL SECTION

Chemicals and reagents. Tris-HCl, EDTA, and $MgCl_2$ were purchased from Nacalai Tesque, Inc. (Kyoto, Japan). STV was purchased from Sigma-Aldrich (Japan). PDC-biotin ligand was prepared based on the synthetic procedures that we have reported recently (Faverie *et al., Biochimie*, **2011**, *93*, 1357.). Single-stranded M13mp18 DNA was obtained from New England Biolabs, Inc. (Ipswich, MA). The staple strands (most of them are 32-mer) for the fabrication of the DNA origami frame, and the oligomers for the formation of intermediate structures were received from Sigma Genosys (Hokkaido, Japan). The gel-filtration column and sephacryl S-300 were purchased from Bio-Rad Laboratories, Inc. (Hercules, CA) and GE Healthcare UK Ltd. (Buckinghamshire, UK), respectively. Water was deionized (≥18.0 MΩ cm specific resistance at 25 °C) by a Milli-Q system (Millipore Corp., Bedford, MA).

Preparation of the origami frame and incorporation of the duplexes. Origami frame was prepared by annealing the solution of M13mp18 DNA (final concentration of 0.01 μM), staple DNAs (0.04 μM), Tris-HCl (20 mM, pH 7.6), EDTA (1 mM) and $MgCl_2$ (5 or 10 mM) from 85 to 15 °C at a rate of -1.0 °C/min. The duplex DNAs (final concentration of 0.1 μM each) were also prepared using the same condition with that of the origami frame. 10-fold excess of each duplex was then mixed with the origami frame. Self-assembly of these duplexes inside the origami frame was carried out by re-annealing the solution from 50 to 15 °C at a rate of -1.0 °C/min. The duplexes incorporated origami was purified using sephacryl S-300 gel-filtration column before HS-AFM imaging. For the experiments in the presence of PDC-biotin, the ligand (1 μM of final concentration) was added to the origami solution before immobilization on the mica surface. For the design of origami frame and the sequence of staple strands, see our previous publication (Endo *et al.*, *J. Am. Chem. Soc.,* **2010**, *132*, 1592.).

AFM imaging. AFM images were recorded using a fast-scanning AFM system (Nano Live Vision, RIBM Co. Ltd., Tsukuba, Japan) with a silicon nitride cantilever (resonant frequency 1.0-2.0 MHz, spring constant 0.1-0.3 N/m, EBD tip radius <15 nm, Olympus BL-AC10EGS-A2). 2 μL of sample was adsorbed onto a freshly cleaved mica plate (ϕ 1.5 mm, RIBM Co. Ltd., Tsukuba, Japan) for 5 min at room temperature and then the surface was gently washed 3-5 times using 20 mM Tris-HCl buffer solution with the same concentration of salt in which origami was prepared. For the experiments in the presence of streptavidin (STV), STV (0.2 μM) was added on the mica surface and incubated for 5 min. Images were recorded after gently washing the surface. Scanning was performed using the tapping mode. All images reported here were recorded with an image acquisition speed of 0.2 frame/s. The yield calculations of the parallel and X-shapes were carried out by counting the shapes in the AFM images. Note, the term "parallel-shape" doesn't mean the strand polarity such as parallel G-quadruplex, it rather represents the parallel orientation of the incorporated duplexes in the AFM topographic image.

1) Tetramolecular G-hairpin: Six contiguous Gs

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTA**GGGGGG**TCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′

Bottom duplex:
5′- GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG-3′
3′-CTAAGAGA ACAAACGA CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT **GGGGGG** AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC **TACAACAA ATAACAGC-5′**

2) Tetramolecular G-triplex: Six contiguous Gs

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTA**GGGGGG**TCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT**GGGGGG**AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′

Bottom duplex:
5′- GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG-3′
3′-CTAAGAGA ACAAACGA CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT **GGGGGG** AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC **TACAACAA ATAACAGC-5′**

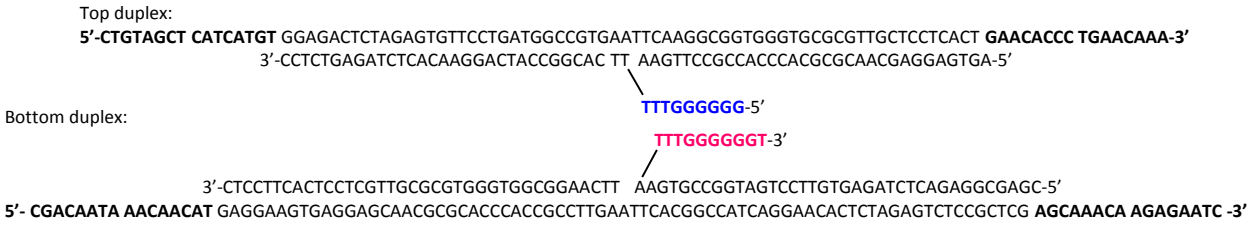
3) Tetramolecular G-quadruplex: Six contiguous Gs

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTA**GGGGGG**TCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT**GGGGGG**AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′

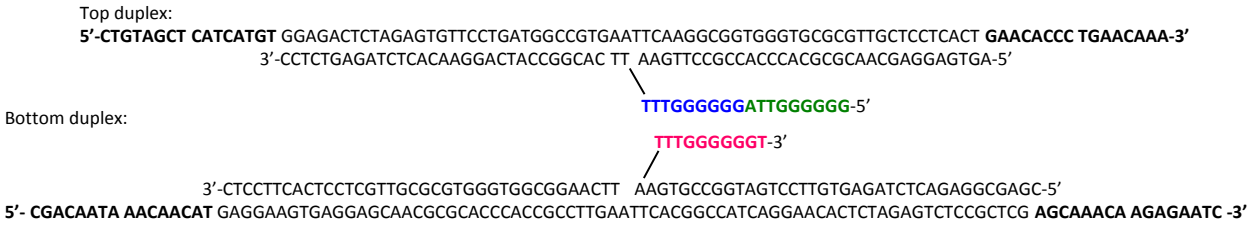
Bottom duplex:
5′-CGACAATA AACAACAT GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA**GGGGGG**TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG **AGCAAACA AGAGAATC-3′**
3′-CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT**GGGGGG**AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC-5′

Fig. S1. The DNA sequences used for the preparation of tetramolecular antiparallel G-hairpin, G-triplex and G-quadruplexes. The colored regions at the middle are the G-repeat sequences that form the notable structure. The bold letter regions at both the termini indicate the single-stranded regions that are needed to attach these duplexes inside the origami frame. This figure is taken from the supporting information of our recent report [Rajendran *et al.*, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4107.] as we have used the same sequences in this study.

1) (3+1)-type G-hairpin: Six contiguous Gs



2) (3+1)-type G-triplex: Six contiguous Gs



3) (3+1)-type G-quadruplex: Six contiguous Gs

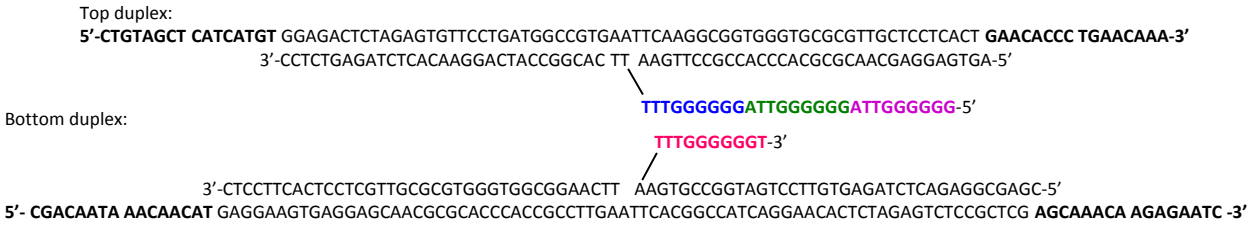
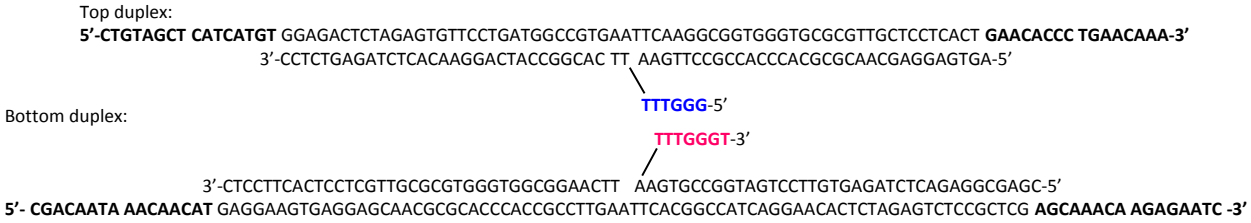
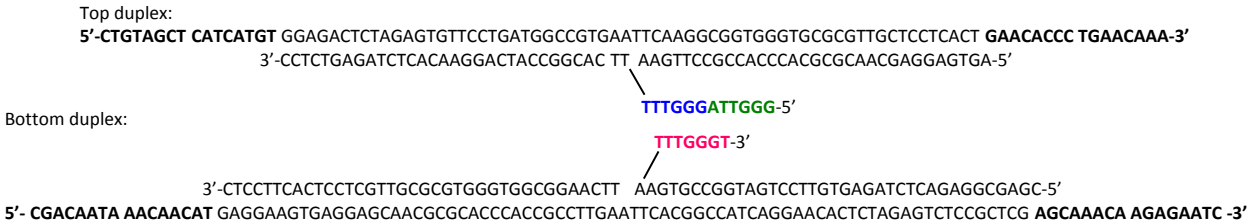


Fig. S2. The DNA sequences used for the preparation of (3+1)-type G-hairpin, G-triplex and G-quadruplexes. Six contiguous Gs were used in this case. The colored regions at the middle are the G-repeat sequences that form the notable structure. The bold letter regions at both the termini indicate the single-stranded regions that are needed to attach these duplexes inside the origami frame. This figure is taken from the supporting information of our recent report [Rajendran *et al.*, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4107.] as we have used the same sequences in this study.

1) (3+1)-type G-hairpin: Three contiguous Gs



2) (3+1)-type G-triplex: Three contiguous Gs



3) (3+1)-type G-quadruplex: Three contiguous Gs

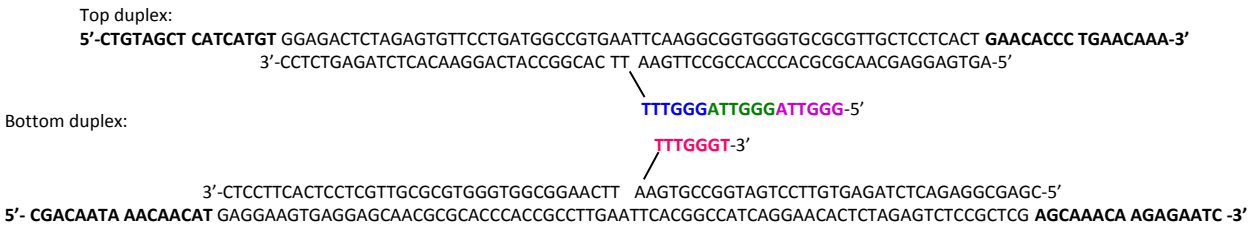


Fig. S3. The DNA sequences used for the preparation of (3+1)-type G-hairpin, G-triplex and G-quadruplexes. Three contiguous Gs were used here. The colored regions at the middle are the G-repeat sequences that form the notable structure. The bold letter regions at both the termini indicate the single-stranded regions that are needed to attach these duplexes inside the origami frame. This figure is taken from the supporting information of our recent report [Rajendran *et al.*, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4107.] as we have used the same sequences in this study.

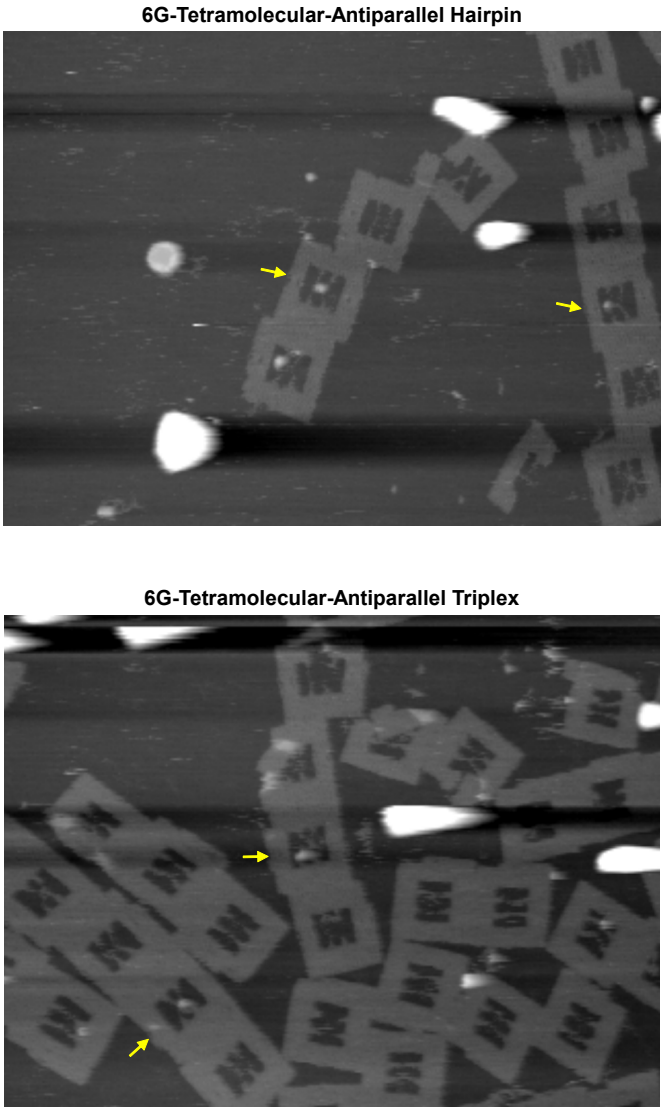


Fig. S4. Representative zoom-out images of the PDC-biotin ligand-induced formation of tetramolecular antiparallel G-hairpin (top) and G-triplex (bottom) structures inside a DNA origami frame. Six contiguous Gs were used in this case. Arrows indicate the X-shaped DNA strands with bound STV inside the origami. Image size: 800 × 600 nm. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.

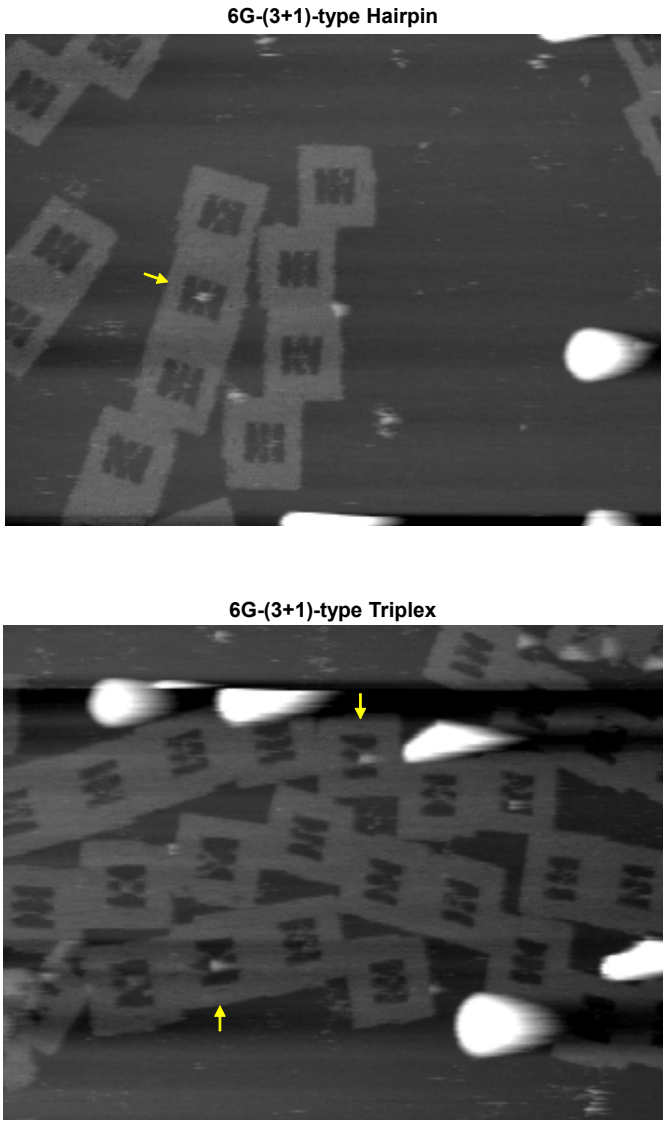


Fig. S5. Representative zoom-out images of the PDC-biotin ligand-induced formation of (3+1)-type G-hairpin (top) and G-triplex (bottom) structures formed inside a DNA origami frame. Six contiguous Gs were used in this case. Arrows indicate the X-shaped DNA strands with bound STV inside the origami. Image size: 800 × 600 nm. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.

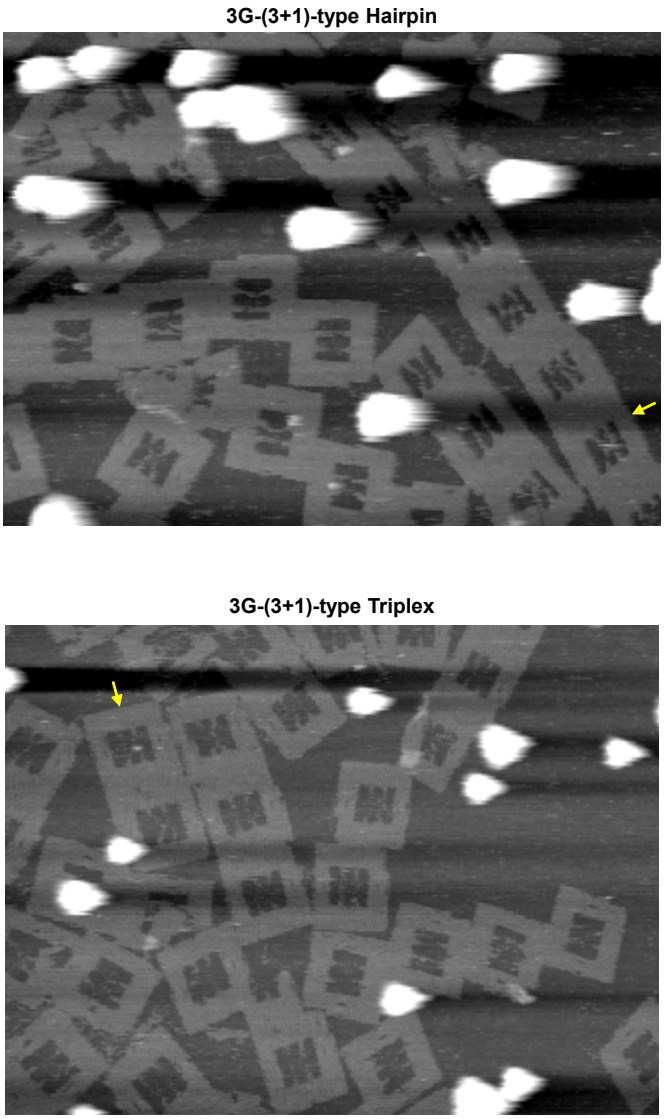
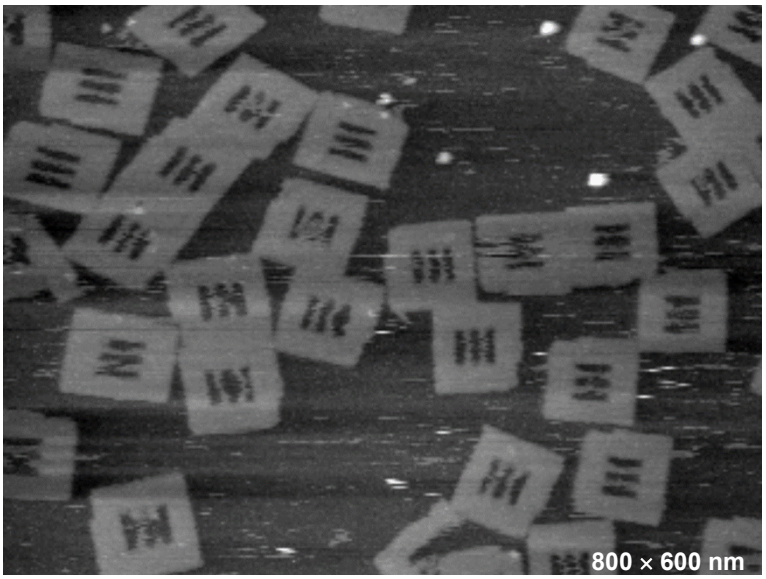


Fig. S6. Representative zoom-out images of the PDC-biotin ligand-induced formation of (3+1)-type G-hairpin (top) and G-triplex (bottom) structures formed inside a DNA origami frame. Three contiguous Gs were used in this case. Arrows indicate the X-shaped DNA strands with bound STV inside the origami. Image size: 800 × 600 nm. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.

Mutated sequence - Tetramolecular hairpin: Six contiguous Gs/Ts

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTAGGGGGGTCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′

Bottom duplex:
5′- GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG-3′
3′-CTAAGAGA ACAAACGA CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT TTTTTT AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC **TACAACAA ATAACAGC-5′**



Mutated sequences - (3+1)-type triplex: Contiguous Ts

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTGAATTCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAC TT AAGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′

Bottom duplex:
TTTTTTTTTTTTTTTT-5′
TTTTTTTTT-3′
3′-CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAACTT AAGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC-5′
5′- CGACAATA AACAACAT GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTGAATTCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG **AGCAAACA AGAGAATC -3′**

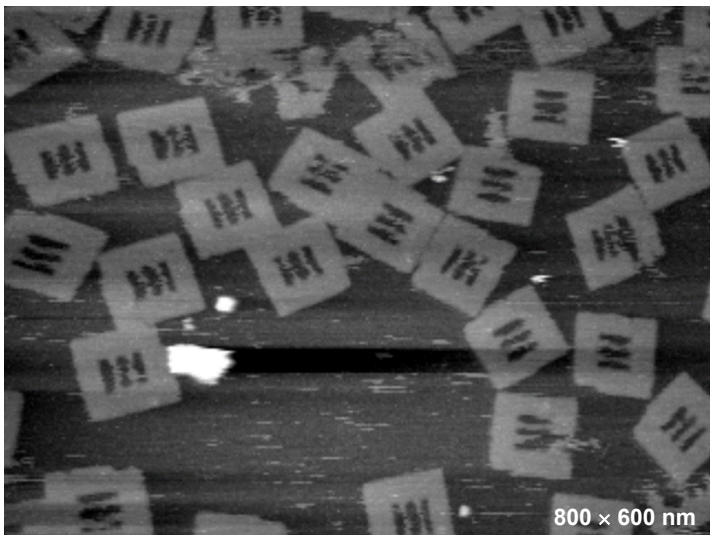


Fig. S7. Control experiments performed by using the sequences that contained the G-T and T-T mismatches. Top: Tetramolecular hairpin; Bottom: (3+1)-type triplex. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM. [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM. Image size is given at the bottom of each image.