SUPPORTING INFORMATION

For

Nonalternating purine pyrimidine sequences can form stable left-handed DNA duplex by strong topological constraint

Lin Li[1], Yaping Zhang[1,§], Wanzhi Ma[1], Hui Chen[1], Mengqin Liu[1], Ran An[1,2], Bingxiao Cheng[1], Xingguo Liang[1,2,*]

1 College of Food Science and Engineering, Ocean University of China, Qingdao 266003, China

2 Laboratory for Marine Drugs and Bioproducts, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266235, China

§ Present Address: State Key Laboratory of Chemical Oncogenomics, The School of Chemical Biology and Biotechnology, Peking University, Shenzhen Graduate School, Shenzhen 518055, China

*To whom correspondence should be addressed. Tel: +86 532 82031086; Fax: +86 532 82031086; Email: liangxg@ouc.edu.cn.

Table S1. The sequences of 74 bp, 92 bp, and 109 bp

Name	Sequences（5′→3′）	Length (nt)
Is74α_0	AGGATCAGAGTGAGAGTTAGAGAGAGACGTTACATCAGGTACGTGTACATATACAT ACGCATGCGCATGCACAG	74
Is74β_0	ATGTAACGTCTCTCTCTAACTCTCACTCTGATCCTCTGTGCATGCGCATGCGTATGTA TATGTACACGTACCTG	74
Sp74α_0	CTGATCCT CTGTGCAT	16
Sp74β_0	CGTTACATCAGGTACG	16
Is74α_1	AGCTAGAGCAGAGCAGGAATTCAGAGTGAGAGAGCGAGAGGATCAGAGTGAGAG TTAGAGAGAGACGTTACATC	74
Is74β_1	ATGTAACGTCTCTCTCTAACTCTCACTCTGATCCTCTCGCTCTCTCACTCTGAATTCC TGCTCTGCTCTAGCTG	74
Sp74α_1	CTCTAGCTGATGTAAC	16
Sp74β_1	TTACATCAGCTA	12
Is74$\alpha_{1\text{-}20}$	AGCTAGAGCCACAGGACTTAAGTATAAATGAGAGCGAGAGGATCAGAGTGAGAGT TAGAGAGAGACGTTACATC	74
Is74$\alpha_{1\text{-}35}$	AGGATATAGCACAGGACTTAAGTATAAATAGACAGATGAGGATCAGAGTGAGAGTT AGAGAGAGACGTTACATC	74
Sp74$\alpha_{1\text{-}20}$	CTCTAGCTGATGTAAC	16
Sp74$\alpha_{1\text{-}35}$	TATATCCT GATGTAAC	16
Is74α_2	CGAGAGCAAGAGAGAGGAGAAGATAGAAGAAGAGGAGCGAGGAGGAGAGAAGG AGTGAGAAAGAAGGGAGTGAA	74
Is74β_2	TTCACTCCCTTCTTTCTCACTCCTTCTCTCCTCCTCGCTCCTCTTCTTCTATCTTCTCC TCTCTCTTGCTCTCG	74
Sp74α_2	CTTGCTCTCGTTCACTCCCT	20
Sp74β_2	GGAGTGAACGAGAGCA	16
Is74α_3	AGTTGATATTTCGCTAAGTGATCTATGAAAGTAAGGTGGCCTCAGACGTTTGAGTCA AGGTGTGAATTCACGGC	74
Is74β_3	CGTGAATTCACACCTTGACTCAAACGTCTGAGGCCACCTTACTTTCATAGATCACTT AGCGAAATATCAACTGC	74
Sp74α_3	TTCACGGCAGTT	12
Sp74β_3	TCAACTGCCGTG	12
Is92α_A	ATATGCCTAGTCTTCCATGACACTGAGACA ACTGGAAACCACTCATGCGT	50
Is92α_B	AGATGCGAACTAGCACAGTGAACTCATCCACTAGAACTTGAC	42
Is92β_A	GTTCGCATCTACGCATGAGTGGTTTCCAGTTGTCTCAGTGTCA	43
Is92β_B	TGGAAGACTAGGCATATGTCAAGTTCTAGTGGATGAGTTCACTGTGCTA	49
Sp92α_A	GCATATGTCAAG	12
Sp92α_B	GCATCTACGCAT	12
Sp92β_A	GCGAACTAGCAC	12
Sp92β_B	CTTCCATGACAC	12
Is109α	GAAGGACCAAGTCTGTCATGCACTGAAATCAGTCTCATTGCTTTATAAACAACCAG CTAAGACACTGCCATACCCTGTAGAACCGAATTTGTGCAGACTCCGGTGGAAT	109
Is109β	ATTCCACCGGAGTCTGCACAAATTCGGTTCTACAGGGTATGGCAGTGTCTTAGCTG GTTGTTTATAAAGCAATGAGACTGATTTCAGTGCATGACAGACTTGGTCCTTC	109
Sp109α	GTCCTTCATTCCAC	14
Sp109β	GTGGAATGAAGGAC	14

Table S2. T_ms (**ºC**) for LR-chimera and lc in various concentrations of $MgCl_2$ and KCl*

DNA \ $MgCl_2$ (mM)	10	2.5	1.0	140 mM KCl
cc-74$_0$	**67.4**	**66.3**	**65.3**	**54.7**
lc-74$_0$	**76.3**	**75.8**	**75.5**	**75.9**
cc-74$_1$	**65.8**	**65.5**	**65.0**	**60.0**
lc-74$_1$	**77.6**	**77.1**	**76.4**	**76.2**
cc-74$_{1\text{-}20}$	**61.2**	**60.7**	**59.4**	**50.0**
lc-74$_{1\text{-}20}$	**72.9**	**72.3**	**72.0**	**71.8**
cc-74$_{1\text{-}35}$	**57.4**	**55.4**	**53.4**	**45.0**
lc-74$_{1\text{-}35}$	**70.3**	**69.3**	**68.5**	**68.3**
cc-74$_2$	**72.1**	**72.2**	**72.7**	**63.5**
lc-74$_2$	**76.9**	**76.4**	**76.3**	**75.2**

*All solutions containing 0.25 μM of DNA and 10 mM HEPES (pH7.5).

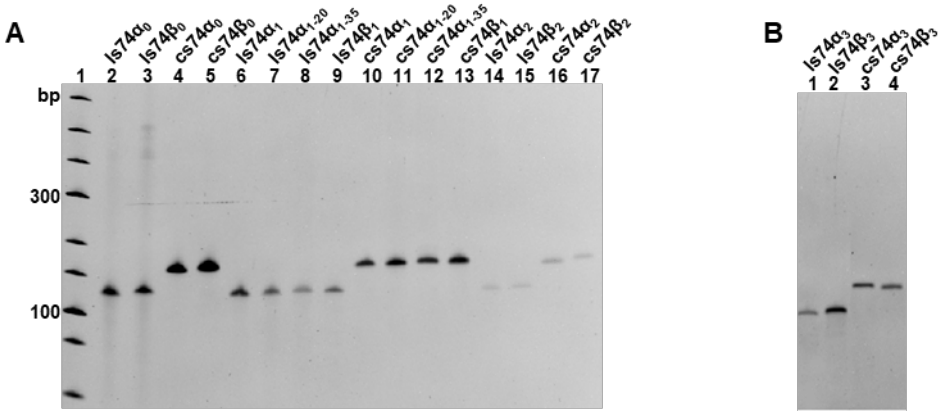

Figure S1. Analysis of purified circular ssDNA by PAGE. A) Purified circular ssDNA $cs74_0$, $cs74_1$, $cs74_{1\text{-}20}$, $cs74_{1\text{-}35}$, and cc-74_2. Lanes 2, 3: linear ssDNA $ls74\alpha_0$, $ls74\beta_0$; lanes 4, 5: purified circular ssDNA $cs74\alpha_0$, $cs74\beta_0$; lanes 6-9: linear ssDNA $ls74\alpha_1$, $ls74\alpha_{1\text{-}20}$, $ls74\alpha_{1\text{-}35}$, $ls74\beta_1$; lanes 10-13: purified circular ssDNA $cs74\alpha_1$, $cs74\alpha_{1\text{-}20}$, $cs74\alpha_{1\text{-}35}$, $cs74\beta_1$; lanes 14, 15: linear ssDNA $ls74\alpha_2$, $ls74\beta_2$; lanes 16, 17: purified circular ssDNA $cs74\alpha_2$, $cs74\beta_2$. (8% dPAGE containing 8.0 M urea, lanes 1-13: ssDNA is 2.0 pmol, lanes 14-17: ssDNA is 4.0 pmol). B) Purified circular ssDNA $cs74_3$. Lanes 1, 2: linear ssDNA $ls74\alpha_3$, $ls74\beta_3$; lanes 3, 4: purified circular ssDNA $cs74\alpha_3$, $cs74\beta_3$;

Fluorescent dyes are difficult to stain ssDNA composed of only purine (A/G) and pyrimidine (C/T), which leads to the single stranded DNA α_2 and β_2 bands is not obvious (lanes 14-17). This has been reported previously (Han,X., Wang,E., Cui,Y., Lin,Y., Chen,H., An,R., Liang,X., Komiyama,M. (2019) The staining efficiency of cyanine dyes for single-stranded DNA is enormously dependent on nucleotide composition. *Electrophoresis*, **40**, 1708-1714).

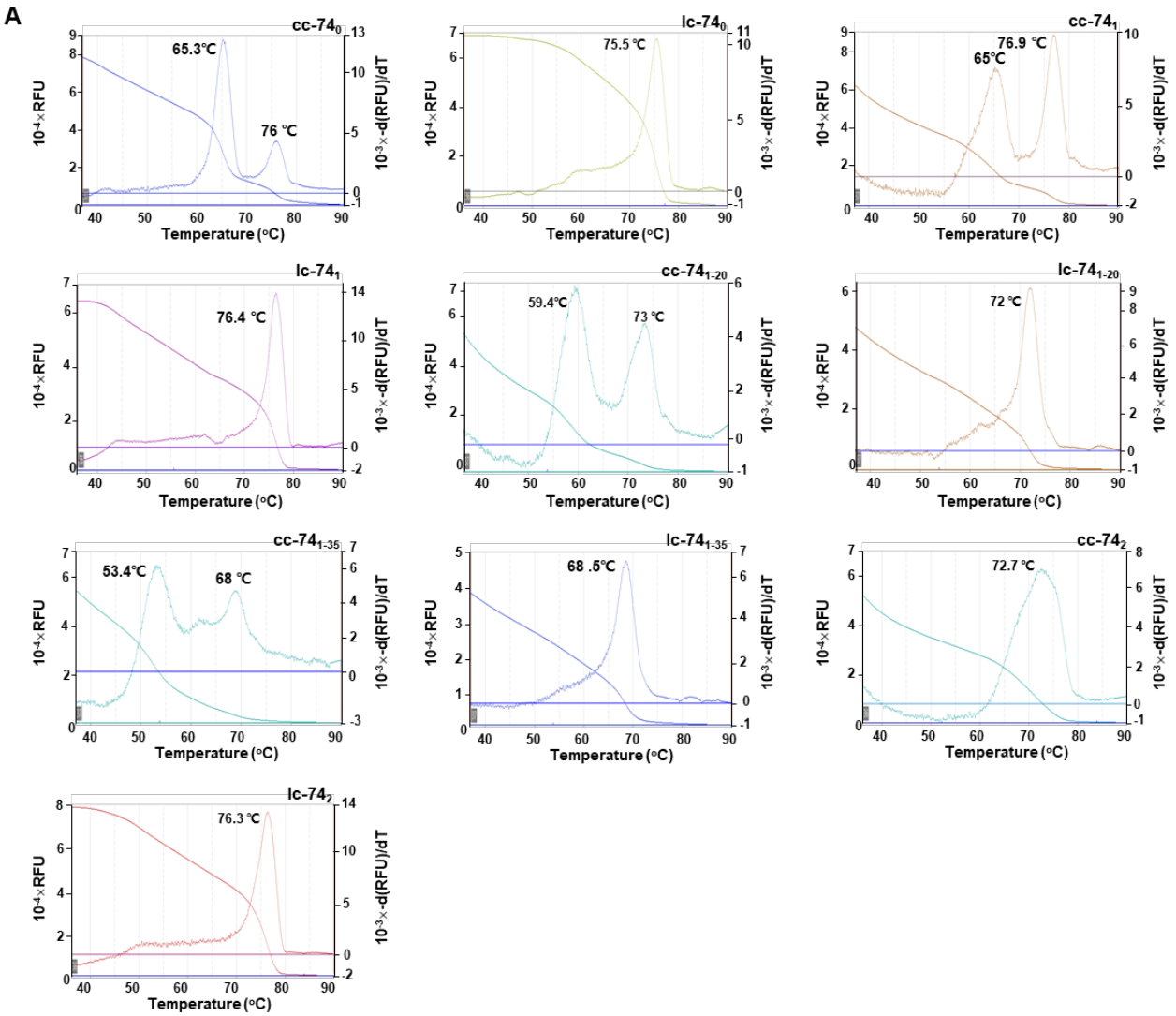

Figure S2A. T_m curves and calculus curves measured by HRM in 1.0 mM $MgCl_2$. [dsDNA] = 0.25 μM, [HEPES] = 10 mM, pH7.5.

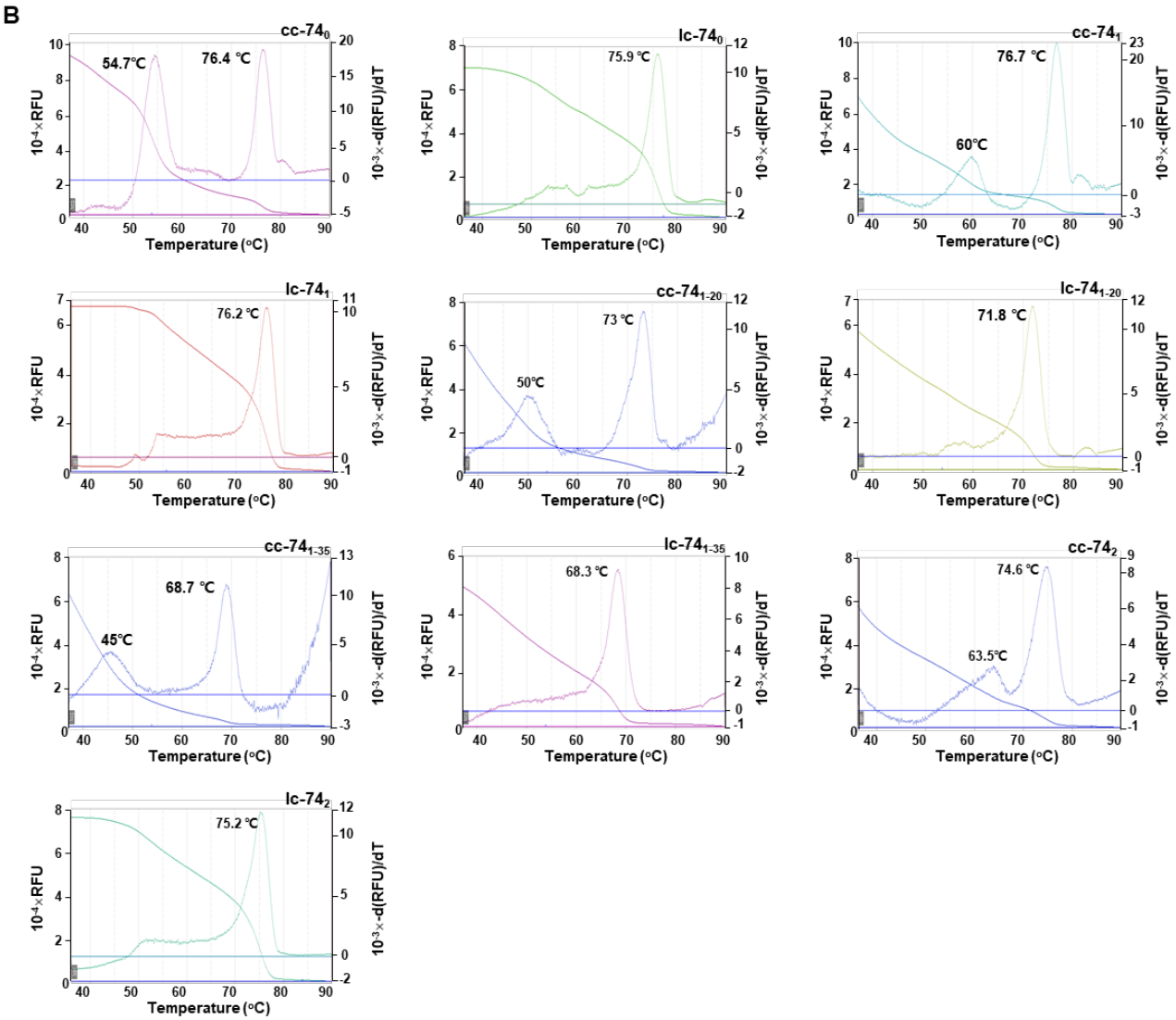

Figure S2B. T_m curves and calculus curves measured by HRM in 140 mM KCl. [dsDNA] = 0.25 μM, [HEPES] = 10 mM, pH7.5.

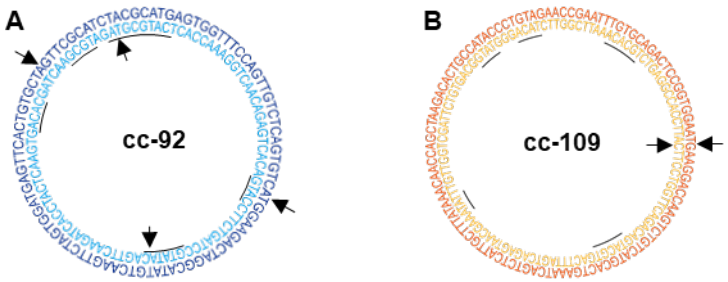

Figure S3. Schematic diagram and sequences for other non-APP sequences of 92 and 109 nt long. A) Non-APP sequence cc-92 (containing a 6 bp APP sequence and an 8 bp one). B) Non-APP sequence cc-109 (containing only two 6 bp APP sequences). The underlines show continuous APP sequences. Arrows indicate the terminals before circularization.

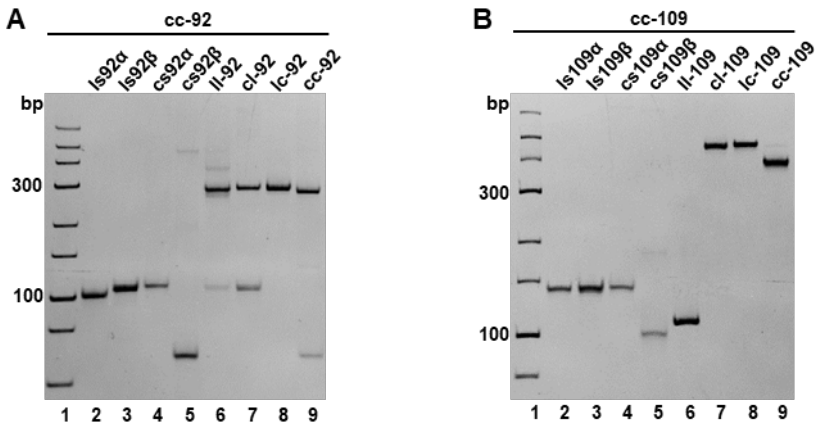

Figure S4. PAGE (8%) analysis of hybridization for cc-92 (A) and cc-109 (B), respectively. Name of samples are shown on top of the gel. The abbreviations are shown as follows. ls: linear ssDNA strand; cs: circular ssDNA strand; lc: hybridization of liner α strand and circular β strand; cl: hybridization of circular α strand and linear β strand; cc: hybridization of two circular strands. The gels were kept at 20^{o}C during electrophoresis.

After two complementary strands of purified circular ssDNAs were hybridized at ionic strength close to physiological conditions (10 mM HEPES, 10 mM $MgCl_2$, pH7.5), the formation of LR-chimera was analyzed by 8% PAGE (Figure S4). As expected, new bands assigned as dsDNA of cc-92 (lane 9, Figure S4A), and cc-109 (lane 9, Figure S4B) were observed, which have similar mobility as the dsDNA formed from the circular ssDNA and its complementary linear one (lanes 7, 8, Figure S4A, S4B). It should be noted that during PAGE analysis, only 89 mM Tris-H_3BO_4 and no $MgCl_2$ were present, indicating that the formed cc-92 and cc-109 were even stable under these low ionic strength conditions.

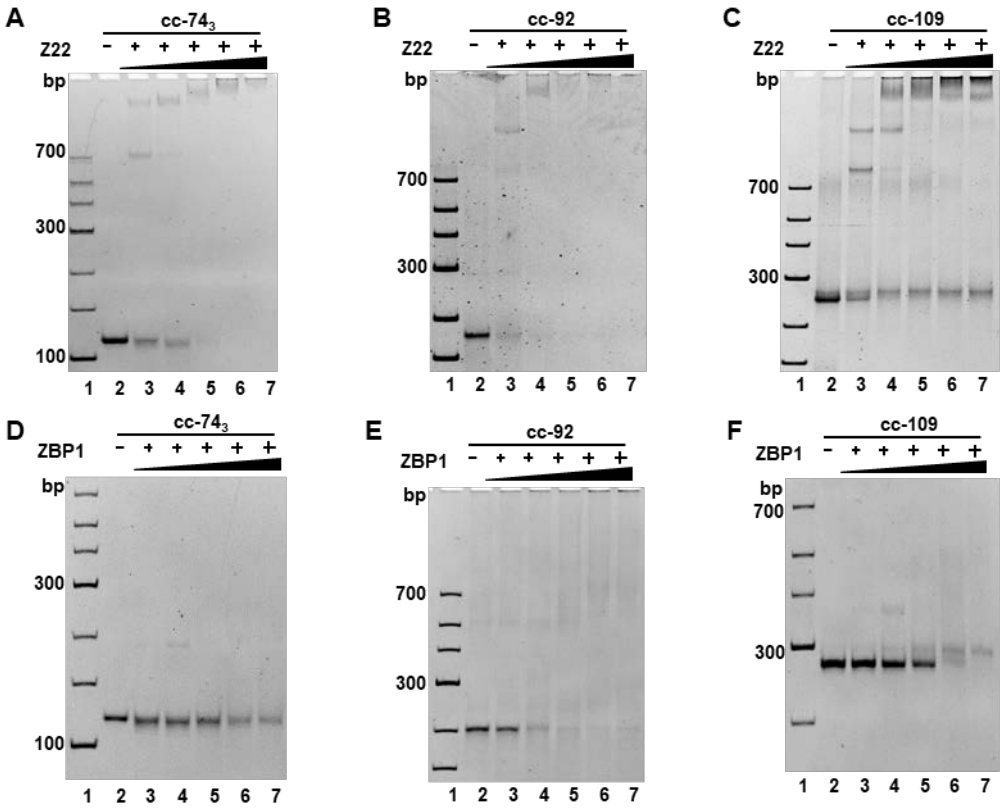

Figure S5. Gel shift assay for binding of Z22 (Z-DNA-specific antibody) and ZBP1 to LR-chimera of various sequences. A) Z22 to cc-74$_3$; B) Z22 to cc-92; C) Z22 to cc-109; D) ZBP1 to cc-74$_3$; E) ZBP1 to cc-92; F) ZBP1 to cc-109. 0.25 μM of LR-chimeras was used. For A-C) 0.25, 0.5, 0.75, 1.0, and 1.25 μM of Z22 were used; For D-F) 0.25, 0.625, 1.25, 1.875, and 2.5 μM of ZBP1 were used. The gel of cc-74$_3$ (A, C) was kept at 10^{o}C during running the gel, the gel of cc-92 (B, E) and cc-109 (C, F) was kept at 20^{o}C during running the gel.

The results of Z22 (0.25, 0.5, 0.75, 1.0, 1.25 μM) binding to cc-74$_3$, cc-92, and cc-109 (0.25 μM) in 1×HEPES buffer (10 mM HEPES, 10 mM $MgCl_2$, pH7.5) are shown in Figure S5A-C. The band for cc-74$_3$ (Figure S5A), cc-92 (Figure S5B), and cc-109 (Figure S5C) disappeared gradually with the increase of Z22 concentration, and new bands were observed, indicating that Z22 binds strongly. Indicating that Z-conformation formed in both cc-74$_3$, cc-92, and cc-109. When another Z-DNA binding protein ZBP1 was used, similar results were obtained but with weaker binding activities (Figure S5D-F). Also suggesting that Z-DNA-like conformation formed in cc-74$_3$, cc-92, and cc-109.

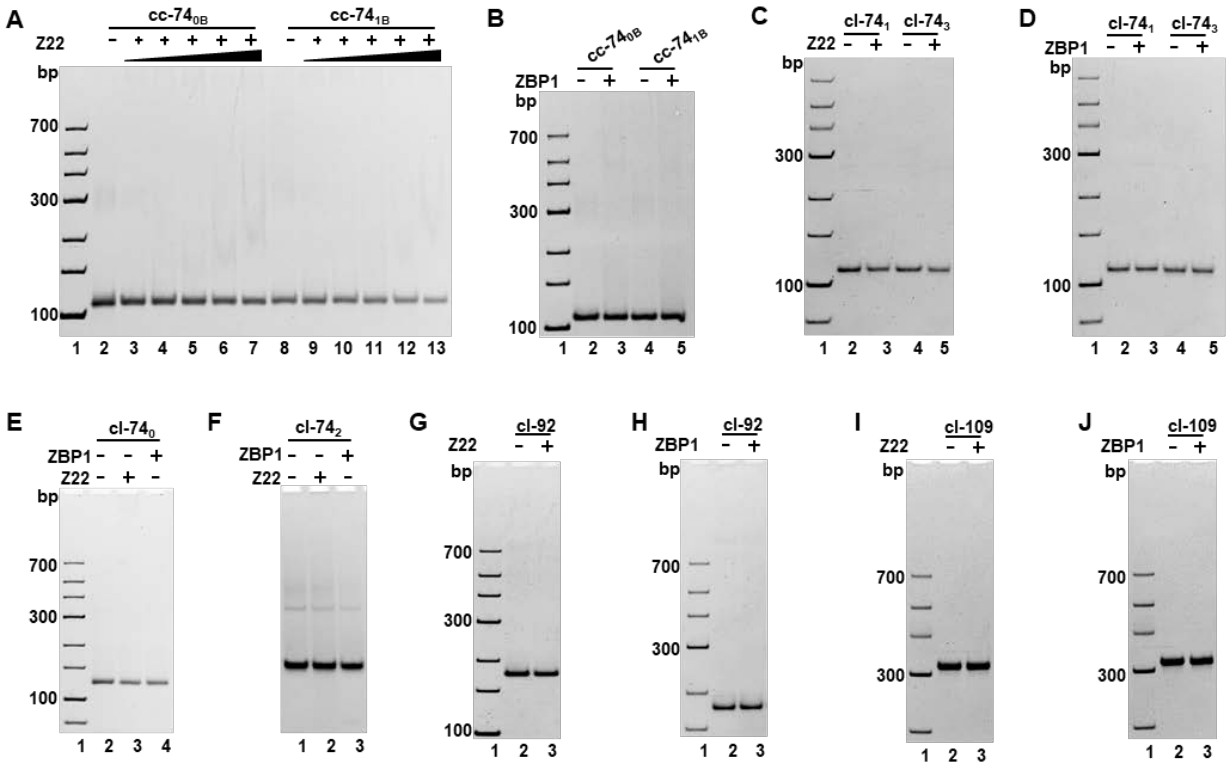

Figure S6. Gel shift assay for binding of Z22 (Z-DNA-specific antibody) and ZBP1 to B-form cc and cl (hybridization of circular α strand and linear β strand) of various sequences. A) Z22 to cc-74$_{0B}$ and cc-74$_{1B}$; B) ZBP1 to cc-74$_{0B}$ and cc-74$_{1B}$; C) Z22 to cl-74$_1$ and cl-74$_3$; D) ZBP1 to cl-74$_1$ and cl-74$_3$; E) Z22 and ZBP1 to cl-74$_0$; F) Z22 and ZBP1 to cl-74$_2$; G) Z22 to cl-92; H) ZBP1 to cl-92; I) Z22 to cl-109; J) ZBP1 to cl-109. In A), the concentrations of cc-74$_{0B}$ and cc-74$_{1B}$ were 0.25 μM, the Z22 were 0.25, 0.5, 0.75, 1.0 and 2.5 μM, respectively. In B-J), the concentrations of DNA, Z22, and ZBP1 were 0.25 μM, 1.25 μM, and 2.5 μM, respectively. The gel for cc-74$_{0B}$, cc-74$_{1B}$, cl-92 and cl-109 (A, B, G-J) were kept at 20^{o}C during running the gel. The gel for cl-74$_0$, cl-74$_1$, cl-74$_2$, and cl-74$_3$ (C-F) were kept at 10^{o}C during running the gel.

For cc-74$_{0B}$ and cc-74$_{1B}$ (0.25 μM), almost no binding was observed even for 1.25 μM of Z22 and 2.5 μM of ZBP1 (Figure S6A, S6B). For cl-74, cl-92, and cl-109 (0.25 μM) without topological constraint, no binding was observed even for 1.25 μM of Z22 (Figure S6C, S6E, S6F, S6G, S6I). Similarly, no binding was observed even for 2.5 μM of ZBP1 to cl-74, cl-92, and cl-109 (0.25 μM) (Figure S6D-F, S6H, S6J).

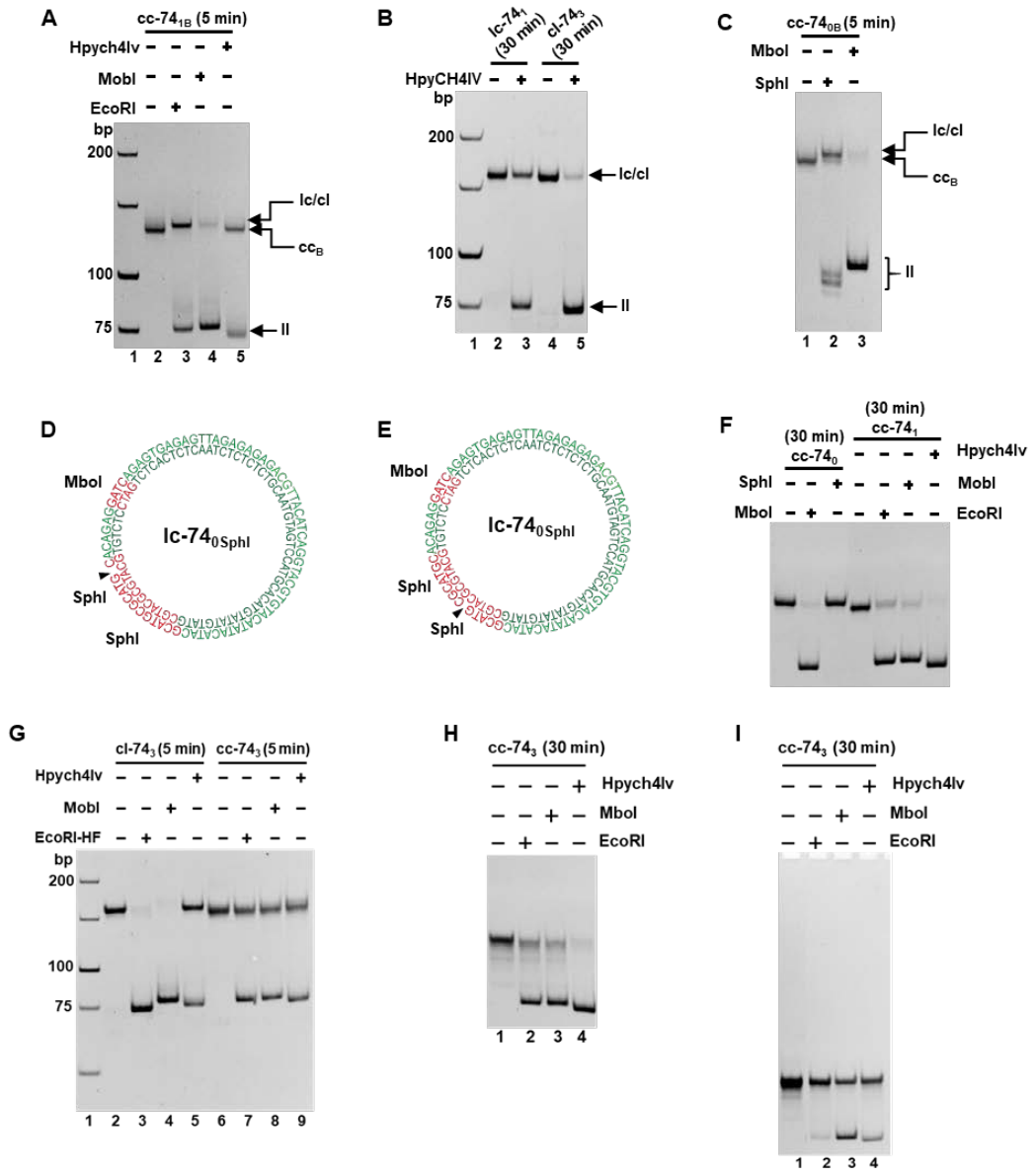

Figure S7. Analysis of the positions of left-handed DNA in the LR-chimeras by restriction enzyme cleavage (8% PAGE). A) Cleavage (5 min) of cc-74$_{1B}$. Lane 2: cc-74$_{1B}$; lanes 3-5: cc-74$_{1B}$ cleaved by EcoRI, MobI, and Hpych4Iv, respectively. B) Cleavage (30 min) by Hpych4Iv using lc-74$_1$ and lc-74$_3$ with a nick. Lane 2: lc-74$_1$; lane 3: Hpych4Iv digestion of lc-74$_1$; lane 4: cl-74$_3$; lane 5: Hpych4Iv digestion of cl-74$_3$. C) Cleavage (5 min) of cc-74$_{0B}$. Lane 1: cc-74$_{0B}$; lanes 2 and 3: cc-74$_{0B}$ cleaved by SphI and MobI, respectively. D and E) Schematic illustration of the product lc-74$_{0SphI}$ from cc-74$_{0B}$ cleaved by SphI. F) Cleavage (30 min) of cc-74$_0$ and cc-74$_1$. Lane 1: cc-74$_0$; lanes 2 and 3: cc-74$_0$ cleaved by MboI and SphI, respectively; lane 4: cc-74$_1$; lanes 5-7: cc-74$_1$ cleaved by EcoRI, MboI, and Hpych4Iv, respectively. G) Cleavage (5 min) of cl-74$_3$ and cc-74$_3$. Lane 2: lc-74$_3$; lanes 3-5: lc-74$_3$ cleaved by EcoRI, MobI, and Hpych4Iv, respectively; lane 6: cc-74$_3$; lanes 7-9: cc-74$_3$ cleaved by EcoRI, MobI, and Hpych4Iv, respectively. H) Cleavage (30 min) of cc-74$_3$. Lane 1: cc-74$_3$; lanes 2-4: cc-74$_3$ cleaved EcoRI, MboI, and Hpych4Iv, respectively. I) Cleavage (30 min) of cc-74$_3$ after binding to Z22. Lane 1: control of cc-74$_3$ without RE; lanes 2-4: cleavage by EcoRI, MboI, and Hpych4Iv, respectively.

Cleavage results of lc-74_3 and cc-74_3 (almost no APP sequences) by EcoRI, MboI, and Hpych4Iv (5 min) are shown in Figure S7G. For EcoRI and MboI, almost all lc-74_3 disappeared (lanes 3 and 4), and about 50% of cc-74_3 still existed (lanes 7 and 8). For Hpych4Iv, about 60% of lc-74_3 and cc-74_3 were remaining (lanes 5 and 9). When the cleavage time is extended to 30 min (Figure S7H), EcoRI and MboI cleaved about 80% of cc-74_3 into linear dsDNA (lanes 2 and 3), but almost no cc-74_3 was left for Hpych4Iv (lane 4). Above the results suggest that the recognition sites for EcoRI and MboI may be prefer to form lh-DNA. To clarify this, effect of Z22 binding on cleavage of cc-74_3 by REs was further investigated (Figure S7I). The results showed that, even after 30 min, EcoRI, MboI, and Hpych4Iv cleaved about 5%, 40% and 20% of cc-74_3, respectively. Obviously, Z22 binding affects greatly the activity of cc-74_3 cleavage by all these three REs. Accordingly, the recognition site of EcoRI may prefer to form lh-DNA.