

Supporting Information

for manuscript entitled

Programming folding cooperativity of dimeric i-motif with DNA frameworks for sensing small pH variations

by

Shiliang He,^{†a} Mengmeng Liu,^{†b} Fangfei Yin,^{ac} Jiangbo Liu,^a Zhilei Ge,^d Fan Li,^a Min Li,^{1a} Jiye Shi,^c
Lihua Wang,^c Xiuhai Mao,^a Xiaolei Zuo^a and Qian Li^{*d}

a. Institute of Molecular Medicine, Shanghai Key Laboratory for Nucleic Acid Chemistry and Nanomedicine, Renji Hospital, School of Medicine and School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, Shanghai 200127, China

b. Shanghai Key Laboratory of Green Chemistry and Chemical Processes, Department of Chemistry, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200241, China

c. Division of Physical Biology, CAS Key Laboratory of Interfacial Physics and Technology, Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China

d. School of Chemistry and Chemical Engineering, Frontiers Science Center for Transformative Molecules and National Center for Translational Medicine, Shanghai Jiao Tong University, Shanghai 200240, China

Table of Contents

Experimental methods	S1
Table S1. DNA sequences used to assemble TDF.....	S5
Table S2. Sequences of dimeric C ₅ T ₃ C ₄ i-motifs with duplexes of different lengths.....	S7
Table S3. Sequences of dimeric C ₅ T ₃ C ₄ i-motifs with duplexes of different GC contents.....	S8
Table S4. Sequences of duplexed i-motifs with polyT spacers of different lengths.....	S9
Positive cooperativity: intermediate states and preorganization	S10
Figure S1. Preorganization can reduce the number of intermediate states in the folding of bimolecular i-motif and thus enhance its folding cooperativity.....	S11
Figure S2. Preorganization effect of TDF in the folding of dimeric i-motif	S12
Figure S3. Proposed folding pathway of dimeric i-motif pre-organized by different elements...	S13
Figure S4. Electrophoresis characterization of the assembly of TDF containing i-motif-forming strands.....	S14
Figure S5. TEM images of TDF-(C ₅ T ₃ C ₄) ₂ prepared in PBS buffer of pH 5.5 and 7.5.....	S15
Figure S6. Thermal melting profiles and van't Hoff plots of 10 bp duplex-assisted C ₅ T ₃ C ₄ i-motif and TDF-(C ₅ T ₃ C ₄) ₂	S16
Figure S7. Migration behavior of i-motif-tethered TDFs in non-denaturing polyacrylamide gel (8%) at pH 8.0 and 5.5, respectively.....	S17
Figure S8. Dimeric i-motif formed by C ₅ T ₃ C ₄ was chosen as the parent structure to be modified in in this work.....	S18
Table S5. pH denaturation data of dimeric i-motifs formed by C-rich oligonucleotides in the form of C _m X ₃ C _n (X = A or T; m, n = 4, 5 or 6)	S19
Figure S9. The folding cooperativity of unmodified bimolecular i-motifs.....	S21
Table S6. pH denaturation data of hairpin-assisted i-motifs (hairpins of different sizes were tested)	S22
Figure S10. Effects of the integration of small DNA hairpins on the apparent folding cooperativity of the dimeric i-motif formed by C ₅ T ₃ C ₄	S24
Table S7. pH denaturation data of hairpin-assisted i-motifs (hairpins with stems of different GC contents were tested)	S25
Table S8. pH denaturation data of hairpin-assisted i-motifs (hairpins with loops of different base compositions were tested)	S26
Figure S11. The observed Hill coefficient (n _H) of hairpin-assisted i-motifs showed dependence on hairpin stem's GC content and loop's base composition.....	S27
Table S9. pH denaturation data of dimeric i-motifs assisted by short duplexes of varying length.....	S28
Figure S12. CD spectra comparison for i-motifs before and after short duplex integration.....	S30
Table S10. pH denaturation data of dimeric i-motifs assisted by 8-bp duplex of different GC contents.....	S31
Table S11. pH denaturation data of dimeric i-motifs assisted by 10-bp duplex with ploy(dT) spacer of different lengths.....	S33
Figure S13. The mean distance between endpoints of ployT spacer and FWHM derived from Monte Carlo simulation, as a function of the number of T in the spacer.....	S34
Table S12. pH denaturation data of TDF-assisted C ₅ T ₃ C ₄ i-motif with polyT spacer of different	

lengths.....	S35
Table S13. pH denaturation data of TDF-assisted $C_mX_3C_n$ ($m, n = 2, 3, 4, 5$) i-motifs.....	S36
References	S37

Experimental Methods:

Chemicals and instruments.

All the chemical reagents were purchased from commercial suppliers (i.e., Sigma-Aldrich (St. Louis, MO), Fisher Scientific (Pittsburg, PA) and others). All oligonucleotides were synthesized and purified by Sangon Inc. Nuclease-free water (Sigma-Aldrich) was used to prepare DNA solutions and buffers. A FE28 pH-meter and calibration buffers (4.01, 7.00 and 9.21) were purchased from Mettler Toledo. Gel filtration chromatography was carried out with a ÄKTA chromatography system (GE Healthcare) equipped with Yarra™ 3 μ M SEC-4000 LC column 300 \times 7.8 mm (Phenomenex). The UV-Vis spectra were recorded on a Cary100 spectrophotometer (Agilent), using quartz cuvettes of 1 cm path length. The Fluorescence excitation and emission spectra were recorded on a FluoroMax-4 fluorescence spectrometer (Horiba), using quartz cuvettes of 1 cm path length. CD spectra were recorded on a J-1500 CD spectrometer (JASCO) with a 1 cm quartz cell.

Preparation of TDF.

The TDFs were self-assembled by mixing the 5 component strands in an equimolar ratio in TM buffer (10 mM Tris-HCl, pH 8.0, 50 mM MgCl₂), followed by heating to 95 °C for 5 min and then quickly cooled to 4 °C. Subsequently, the synthesized TDFs were applied to HPLC purification. Isocratic elution was performed using a buffer (25 mM Tris-HCl, 450 mM NaCl, pH 7.1) at room temperature and with detection at 260 nm. The fractions were collected and concentrated with ultrafiltration (100 kDa). Further, in order to be used in following tests, the buffer of the concentrated sample was replaced by corresponding buffer (400 μ L \times 3) with the help of ultrafiltration.

Nondenaturing gel electrophoresis.

TDF samples (1 μ M) were firstly prepared in TM buffer by fast annealing as described above. To confirm the formation of TDF assembly, non-denaturing polyacrylamide gel electrophoresis (PAGE) was conducted in 1xTBE buffer at 85 V for 2h at 4°C. On the other hand, in order to confirm the monodispersity of TDF at both basic and acidic conditions, the assembled TDF were purified with HPLC, then subjected to ultrafiltration and TDF samples of different concentrations (0.1 μ M, 1.0 μ M and 10 μ M) were prepared in citric acid-Na₂HPO₄ buffer (20 mM, pH 8.0/5.5, containing 100 mM NaCl), which were incubated at 37 °C for 30 min and then cooled to 15 °C with a gradient of 1.0 °C/min to allow the folding of i-motif. The samples were then mixed with loading buffer (Takara) and subjected to 8% PAGE in Tris-AcOH buffer (40 mM, pH 8.0/5.5, containing 1.5 mM Mg(OAc)₂). Gel electrophoresis was carried out at 140 V for 210 min in the same buffer at 0 °C. The gel was stained in a solution of 0.01% YeaRed (Yeasen) for 30 min and photographed.

Transmission electron microscopy (TEM) imaging.

Samples of frameworked i-motif (10 nM) were prepared in PBS buffer (pH 7.5 and 5.5), and heated to 37 °C for 30 min, and then cooled to 15 °C with a gradient of 1.0 °C/min to allow the folding of i-motif. On copper grids (coated with carbon), the solution (5 μ L) of TDF was placed and excess liquid was removed after 10 min by a filter paper. Then, uranyl formate aqueous solution (5 μ L, 1%) was added on the grids and removed after 1min with a filter paper, followed by washing with Milli-Q water 3 times. After drying, the grids were used for TEM measurement, which was carried out on an HT7700 electron microscope (Hitachi) at an acceleration voltage of 100 kV.

Atomic force microscopy (AFM) imaging.

Samples of frameworked i-motif (10 nM) in PBS buffer (pH 7.5) were first deposited on mica surfaces for 5 mins. AFM images were then taken in tapping mode in fluid mode on a Cypher ES (Asylum Research) using BL-AC40TS tips (Olympus). Typical scanning parameters were: scan rate = 1-2 Hz, lines = 512, amplitude set point = 150-300 mV, drive amplitude = 180-300 mV, integral gain = 18.

UV-melting curves and thermodynamic analysis.

Prior to melting curve acquisition, duplex-forming C-rich oligonucleotides (2.0 μM) in PBS buffer (pH 6.2 or 7.4) was annealed from 85 $^{\circ}\text{C}$ to 15 $^{\circ}\text{C}$ with a gradient of 1.0 $^{\circ}\text{C}/\text{min}$, and HPLC-purified TDF bearing i-motif-forming oligonucleotides (0.2 μM) was incubated at 37 $^{\circ}\text{C}$ for 30 min and then annealed from 37 $^{\circ}\text{C}$ to 15 $^{\circ}\text{C}$ with a gradient of 1.0 $^{\circ}\text{C}/\text{min}$. UV-melting curves were recorded at 260 nm as the solution was heated from 15 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$. To limit the evaporation of the sample during the experiment, the top cover of quartz cuvette was wrapped with parafilm. Thermodynamic parameters were determined according to methods described in the literatures,^{1,2} assuming a two-state equilibrium.

Sequence design requirements for hairpin-assisted i-motif structures.

In order to ensure that the investigated hairpins to be reasonably stable, C-G was taken as the hairpin closing base pair because it provides the most thermal stability to DNA hairpins among the four variations of base pairing (i.e., C-G, G-C, T-A and A-T).^{3,4} Similarly, oligodeoxythymidylate strands with a length of 3-5 nucleotides were taken as the loop because among hairpins with loops of homogeneous sequence, the ones with loops consisting of homo-thymidine oligonucleotides were found most stable.⁵ Moreover, in a given hairpin with ploy(dT) loop, once the number of thymidine was increased beyond 5, the melting temperature of the hairpin correlates negatively with its loop length,^{6,7,8} thus only hairpins with triloop, tetraloop and pentaloop were evaluated in this study.

pH Denaturation studies by UV spectrometry.

Solutions of C-rich DNA (100 μL) were prepared in citric acid- Na_2HPO_4 buffer (20 mM, pH 4.0-8.0) with 100 mM NaCl. The concentration of C-rich DNA in the form of $\text{C}_m\text{X}_3\text{C}_n$ ($\text{X} = \text{A}, \text{T}; m, n = 4, 5$ or 6) was 5 μM , and so as that of those containing hairpin fragments. For the C-rich DNA that can form duplex, a 2.5 μM of either of the two complementary single-stranded DNA was used to ensure the end concentration to be 5 μM for the $\text{C}_m\text{X}_3\text{C}_n$ oligonucleotide fragment in the solution. The as-prepared solutions were then placed in a thermocycler and heated to 85 $^{\circ}\text{C}$ for 5 minutes, followed by cooling to 15 $^{\circ}\text{C}$ with a gradient of 1.0 $^{\circ}\text{C}/\text{min}$. The prepared samples were then applied to UV-Vis absorption measurements using a 10 mm optical path cuvette at 25 $^{\circ}\text{C}$. The structural switch of i-motif was monitored at 295 nm. Hypochromicity at 295 nm is a characteristic of i-motif unfolding,⁹ which has been widely adopted to monitor i-motif's denaturing process.^{10,11,12,13,14}

pH Denaturation studies by circular dichroism (CD).

Prior to circular dichroism (CD) measurement, a solution of the C-rich DNA (end concentration of the two-C-tracts fragment was 5.0 μM , 0.4 mL) was prepared in citric acid- Na_2HPO_4 buffer (20 mM, pH 5.3-7.9) with 100 mM NaCl, and was heated to 85 $^{\circ}\text{C}$, then cooled to room temperature

with a gradient of 1.0 °C/min. The annealed samples were then applied to CD spectra acquisition. Each measurement was taken from 230 to 340 nm at a scanning speed of 30 nm/min and a bandwidth of 10 nm. The baseline was corrected by scanning the buffer solution.

pH Denaturation studies by fluorescence spectrometry.

Prior to fluorescence spectra measurement, as for duplex-forming C-rich oligonucleotides, two complementary single-stranded DNAs were mixed at a molar ratio of 1:1 in PBS buffers of different pH values to prepare a solution (80 µL, with a final concentration of 100 nM for each ssDNA) which was first heated to 85 °C for 5 minutes and then cooled to 15 °C with a gradient of 1.0 °C/min. As for TDF bearing i-motif-forming oligonucleotides, a PBS solution (80 µL) of 100 nM for the tetrahedron was first incubated at 37 °C for 30 min and then cooled to 15 °C with a gradient of 1.0 °C/min before fluorescence measurements. Then the emission spectra were recorded from 530 to 750 nm at room temperature with an excitation wavelength of 510 nm.

Intracellular pH calibration and fluorescence imaging.

MCF-7 cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin solution at 37 °C in a humidified incubator containing 5% CO₂. After removing the medium, cells were incubated with fresh medium containing TDF-assisted i-motif (100 nM) for 4 h. Then cells were washed three times by washing buffer (WB; 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, 5 mM MgCl₂, and 250 mM glucose, pH 7.4), and treated with high K⁺ HEPES-buffered solution (30 mM NaCl, 120 mM KCl, 1 mM CaCl₂, 0.5 mM MgSO₄, 1 mM NaH₂PO₄, 5 mM glucose, 20 mM HEPES, and 20 mM NaOAc) of various pH values in the presence of 10 µM nigericin. After 20 min, cells were imaged on a laser-scanning confocal microscope. Cy3-Cy5 FRET signal was measured by excitation at 543 nm and collection of Cy3 (549-606 nm) and FRET (650-730 nm) emission.

Data processing.

The midpoint of pH denaturation curve (pH_m), transition width (ΔpH, the pH difference between the two pH values at which the bound species is 10% and 90%, respectively), dynamic range and Hill coefficient (n_H) were derived from three independent pH-denaturation profiles. Dynamic range was defined as the ratio of the target concentrations (H⁺, in this study) at which occupancy is 90% and 10%, respectively (denoted here as [H⁺]₉₀ and [H⁺]₁₀).¹⁵ Dynamic range = [H⁺]₉₀ / [H⁺]₁₀ = 81^(1/n_H). Considering pH = - log [H⁺] and ΔpH = pH₁₀ - pH₉₀ = log {[H⁺]₉₀ / [H⁺]₁₀}, dynamic range = 10^{ΔpH}. By plotting the UV absorption at 295 nm (or in the case of fluorescent assays, the fluorescent D/A ratio at 564 nm/665 nm) against pH data points, and fitting with Boltzmann sigmoidal fits, the value for the curve's inflection point "x0" was taken as pH_m. The width of pH intervals over which signal changes from 10 percent of maximum value to 90 percent of maximum value was taken as ΔpH. By plotting the normalized data of the UV absorption at 295 nm (or D/A ratio at 564 nm/665 nm) against H⁺ concentration data points, and fitting with a two-state Hill equation, the value for the parameter "n" was taken as n_H.

Relative distance prediction for the endpoints of ployT spacers.

In order to predict the distance between the most outward bases of the ployT spacer, the position of each nucleobase in duplex strand is calculated via Monte Carlo simulation. Generally, in a single

strand, each nucleobase is presented as a hard sphere monomer, which is linked to each other with fixed bond length (5.6Å for single strand and 3.3Å base for double strand). After one initial monomer is deployed, the following ones are added sequentially to link to the previous monomer, with a random polar angle. During this adding process, the position of new added base must not overlap with previous monomers. If overlap happens, the addition is dismissed and a ‘fail counter’ is increased. To avoid that the simulated DNA strand being trapped in a cavity, we defined a ‘fail counter’ threshold to stop adding process and re-start from the initial point. After we produced a serial of coordinates for these nucleobases, we record the position of each monomer and determine the length following the equation:

$$L = \max [\text{sqrt}((x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2)]$$

$$i = 2, 3, \dots, N, j = 1, 2, \dots, N-1;$$

For the frameworked i-motif structure, we generated coordinates for the polyT strands of length of 1, 2, 3, 4, 5, 6 base with (0,0,0) as initial monomer for strand A and (0.81nm,0,0) as initial monomer for strand B, respectively. For the duplexed i-motif structure, the position of initial monomer in strand B was changed to (1.61nm,0,0). Then, the spatial distance of the endpoints of polyT spacers was calculated following the equation:

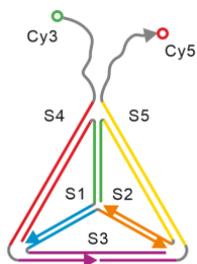
$$D = \text{sqrt}((x_{An} - x_{Bn})^2 + (y_{An} - y_{Bn})^2 + (z_{An} - z_{Bn})^2)$$

Detailed discussions for Figure 3:

Dimeric i-motif folds into compact structure with a spatial distance of 13.8 Å between the 5' end and 3' end (Figure 3a). For dsDNA, the minimum spatial distance between its endpoints is 16.1 Å. Instead, the spatial distance between the endpoints in TDF is calculated to be ~8.1 Å, which almost reaches the spatial regulation resolution limit of ssDNA (~7 Å). Hence, TDF has the chance to position the endpoints of i-motif to the optimum distance (i.e., 13.8 Å). Besides, TDF can regulate the endpoints distance in a broader range by starting from 8.1 Å, compared with dsDNA (starting from 16.1 Å).

Prior to n_H value measurement, we carried out Monte Carlo stimulation to predict the location of terminal oligonucleotide of the polyT spacer (in both Strand A and Strand B), and calculate the D_x and F_x (sample number = 1000, Figure 3b). As shown in Figure 3c, in the case of integrating one thymidine oligonucleotide ($N = 1T$), F_x is significantly smaller than D_x . As the number of thymidine oligonucleotide in the polyT spacer increasing, both D_x and F_x increased accordingly, and the gap between them gradually narrowed (Figure S13). These results indicated that altering the spacer length is a reliable approach for fine-tuning the relative distance between the two endpoints of i-motif. Next, we measured the folding cooperativity of duplexed i-motif and frameworked i-motif bearing different polyT spacers. As shown in Figure 3d, as the length of polyT increasing, the n_H value of duplexed i-motif gradually decreased, reaching a plateau of 4.2. While for frameworked i-motif, its n_H value increased with the spacer extending from 1T to 3T, and decreased afterwards, giving a maximum n_H value when the spacer is of 3 thymidine oligonucleotides ($F_x = 1.56$ nm).

Table S1. DNA sequences used to assemble TDF. TDF-(C_mT₃C_n)₂ (m, n = 2, 3, 4, 5) in this study was assembled from three component strands (S1, S2 and S3) and two fluorescently modified strands (S4 and S5).



S1		GCCTGGAGATACATGCACATTACGGCTTTCCTATTAGAAGGTCTCAGGTGC GCGTTTCGGTAAGTAGACGGGACCAGTTCGCC
S2		CGCGCACCTGAGACCTTCTAATAGGGTTTGCGACAGTCGTTCAACTAGAAT GCCCTTTGGGCTGTTCCGGGTGTGGCTCGTCGG
S3		GGCCGAGGACTCCGTCTCCGCTGCGGTTTGCGAACTGGTCCCCTCTACTT ACCGTTTCCGACGAGCCACACCCGGAACAGCCC
TDF-(C ₅ T ₃ C ₄) ₂ with 1T spacer	S4	Cy3- CCCCCTTCCCCTGCCGTAATGTGCATGTATCTCCAGGC TTTCCGCAGCGGA GAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTGCGTCCCCCTTT CCCC-Cy5
TDF-(C ₅ T ₃ C ₄) ₂ with 2T spacer	S4	Cy3- CCCCCTTCCCCTT GCCGTAATGTGCATGTATCTCCAGGC TTTCCGCAGCGG AGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTGCGTCCCCCTTT CCCC-Cy5
TDF-(C ₅ T ₃ C ₄) ₂ with 3T spacer	S4	Cy3- CCCCCTTCCCCTTTGCCGTAATGTGCATGTATCTCCAGGC TTTCCGCAGCG GAGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTGCGTCCCCCTTT TCCCC-Cy5
TDF-(C ₅ T ₃ C ₄) ₂ with 4T spacer	S4	Cy3- CCCCCTTCCCCTTTT GCCGTAATGTGCATGTATCTCCAGGC TTTCCGCAGC GGAGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTGCGTTTTCCCCT TTCCCC-Cy5
TDF-(C ₅ T ₃ C ₄) ₂ with 5T spacer	S4	Cy3- CCCCCTTCCCCTTTTT GCCGTAATGTGCATGTATCTCCAGGC TTTCCGCAG CGGAGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTGCGTTTTCCCCT TTTTCCCC-Cy5
TDF-(C ₅ T ₃ C ₄) ₂ with 6T spacer	S4	Cy3- CCCCCTTCCCCTTTTTT GCCGTAATGTGCATGTATCTCCAGGC TTTCCGCA GCGGAGAC

	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGCCTTTTTTCCCC CTTTCCCC-Cy5
TDF-(C ₅ T ₃ C ₃) ₂	S4	Cy3- CCCCCTTTCCCCCTTTGCCGTAATGTGCATGTATCTCCAGGCTTTCCGCAGC GGAGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGCCTTTCCCCCTT TCCCCC-Cy5
TDF-(C ₄ T ₃ C ₄) ₂	S4	Cy3- CCCCTTTCCCCCTTTGCCGTAATGTGCATGTATCTCCAGGCTTTCCGCAGCGG AGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGCCTTTCCCCTTT CCCC-Cy5
TDF-(C ₃ T ₃ C ₃) ₂	S4	Cy3- CCCTTTCCCTTTGCCGTAATGTGCATGTATCTCCAGGCTTTCCGCAGCGGAG AC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGCCTTTCCCTTTC CC-Cy5
TDF-(C ₃ T ₃ C ₂) ₂	S4	Cy3- CCCTTTCCTTTGCCGTAATGTGCATGTATCTCCAGGCTTTCCGCAGCGGAGA C
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGCCTTTCCCTTTC C-Cy5
TDF-(C ₂ T ₃ C ₂) ₂	S4	Cy3- CCTTTCCTTTGCCGTAATGTGCATGTATCTCCAGGCTTTCCGCAGCGGAGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGCCTTTCCTTTC -Cy5
TDF-C ₅ T ₃ C ₄ (control TDF in PAGE assay)	S4	CCCCCTTTCCCCCTTTGCCGTAATGTGCATGTATCTCCAGGCTTTCCGCAGCG GAGAC
	S5	GGAGTCCTCGGCCTTTGGGCATTCTAGTTGAACGACTGTCGC

Table S2. Sequences of dimeric C₅T₃C₄ i-motifs with duplexes of different lengths. Fragments composing complementary sequences are underlined.

Duplex length	Sequence (5'→3')	ΔG^0 (kcal/mol)
3-bp	<u>CGCT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGCG</u>	-2.7
4-bp	<u>CGGCT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGCCG</u>	-4.6
5-bp	<u>CGCAGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCTGCG</u>	-5.5
6-bp	<u>CGCAACTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGTTGCG</u>	-6.8
7-bp	<u>GCAGTCGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGACTGC</u>	-8.4
8-bp	<u>AGCACTCGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGAGTGCT</u>	-9.6
9-bp	<u>GCATCAACGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGTTGATGC</u>	-10.5
10-bp	<u>GCATCAGACGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGTCTGATGC</u>	-12.1
11-bp	<u>GCTGTTATGCGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGCATAACAGC</u>	-13.4
12-bp	<u>GCTAGTCAGACGTC</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGTCTGACTAGC</u>	-14.5

Table S3. Sequences of dimeric C₅T₃C₄ i-motifs with duplexes of different GC contents. Fragments composing complementary sequences are underlined.

GC% of 8bp-duplex	Sequence (5'→3')	ΔG ⁰ (kcal/mol)
0	<u>ATAAAATAT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTATTTTAT</u>	-4.1
13	<u>ATACTATAT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTATAGTAT</u>	-4.9
25	<u>AATTCATCT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGATGAATT</u>	-6.1
38	<u>ACTTCATCT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGATGAAGT</u>	-7.1
50	<u>GAACAACG</u> T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGTTGTTT</u>	-8.3
63	<u>AGCACTCG</u> T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGAGTGCT</u>	-9.6
75	<u>GCACGACG</u> T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGTCGTGC</u>	-10.8
88	<u>GCACGGCG</u> T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGCCGTGC</u>	-12.2
100	<u>GCGCCGCCT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGGCGGCGC</u>	-13.3

Table S4. Sequences of duplexed i-motifs with polyT spacers of different lengths. Fragments composing complementary sequences are underlined.

Number of T in PolyT spacer	Sequence (5'→3')
1	<u>GCATCAGACGTC</u> ₅ T ₃ C ₄ -Cy5 / Cy3-C ₅ T ₃ C ₄ <u>TCGTCTGATGC</u>
2	<u>GCATCAGACGTTTC</u> ₅ T ₃ C ₄ -Cy5 / Cy3-C ₅ T ₃ C ₄ <u>TTCGTCTGATGC</u>
3	<u>GCATCAGACGTTTTC</u> ₅ T ₃ C ₄ -Cy5 / Cy3-C ₅ T ₃ C ₄ <u>TTTCGTCTGATGC</u>
4	<u>GCATCAGACGTTTTTC</u> ₅ T ₃ C ₄ -Cy5 / Cy3-C ₅ T ₃ C ₄ <u>TTTTTCGTCTGATGC</u>
5	<u>GCATCAGACGTTTTTTC</u> ₅ T ₃ C ₄ -Cy5 / Cy3-C ₅ T ₃ C ₄ <u>TTTTTTCGTCTGATGC</u>
6	<u>GCATCAGACGTTTTTTTC</u> ₅ T ₃ C ₄ -Cy5 / Cy3-C ₅ T ₃ C ₄ <u>TTTTTTTCGTCTGATGC</u>

Positive cooperativity: intermediate states and preorganization

Positive cooperativity presents a fundamental property of multivalent biological systems in implementing molecular recognition and macromolecular self-assembly.¹⁶ Cooperativity is intrinsically associated with the amount of intermediate states in the binding/folding process: receptor-ligands binding/folding with weak positive cooperativity experiences a significant amount of partially bound/folded states, while strong positive cooperativity shift the receptor population from predominantly free to predominantly bound/folded over only a small change in ligand concentration (Figure S1).^{16,17,18}

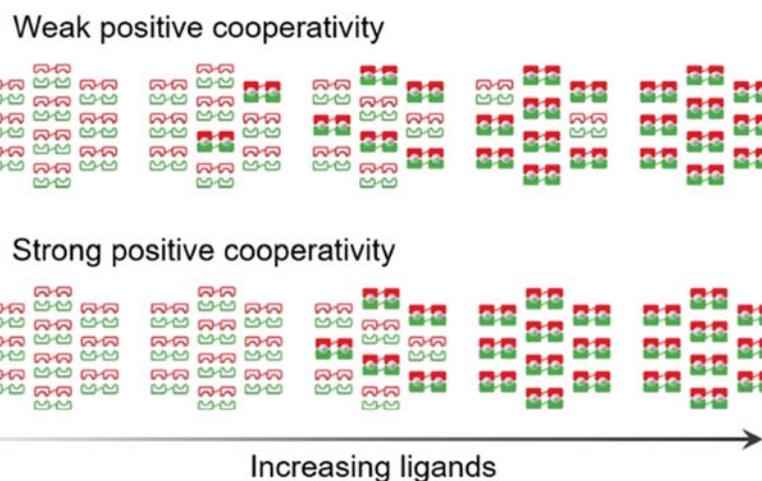
As for bimolecular i-motif, Leroy¹⁹ speculated that it undergoes a complicated three-step folding pathway: (i) two C-rich oligonucleotides ‘nucleate’ by forming several hemiprotonated CH⁺-C base pairs, (ii) a third C-tract intercalates into the parallel duplex during the zippering of the remaining bonds, (iii) a fourth C-tract hybridizes with the third one in parallel orientation which leads to the final thermodynamically most stable structure (Figure S1c). In fact, a partially folded three-strands species was recently identified to coexist with i-motif tetraplex in the pH range of 5.5-7.0.²⁰ The complexity of this multi-step folding pathway increases the occurrence possibility of intermediate species, and any mispairing that fails to contribute to the correct folding should diminish the folding cooperativity. Indeed, weak positive cooperativity was observed for commonly used bimolecular i-motifs in this study.

The folding cooperativity will be enhanced when the folding is pre-organized, in which some structural components assemble early in the folding process, resembling the nucleus in crystallization, which can greatly reduce the number of nonproductive conformations that the assembly can occupy. Watson-Crick duplex folds with an association rate of 10^6 - 10^7 M⁻¹ s⁻¹, much faster than bimolecular i-motif (1 - 100 M⁻¹ s⁻¹).^{19,21} In this study, three ‘preorganization elements’ based on Watson-Crick base-pairing, namely, small DNA hairpins, short duplexes and TDF, were integrated in the bimolecular i-motif system in order to enhance and regulate its folding cooperativity (Figure S2 and Figure S3).

a. Bimolecular receptor-ligand cooperative binding/folding:



b. Free-bound/unfolded-folded state switching:



c. Speculated folding pathway of bimolecular i-motif:

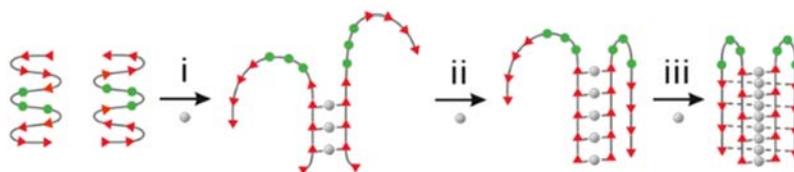


Figure S1. Preorganization can reduce the number of intermediate states in the folding of bimolecular i-motif and thus enhance its folding cooperativity. (a) Representative divalent receptor-ligand models with positive cooperativity ($K_{d2} < K_{d1}$) for bimolecular receptors. (b) Weak positive cooperativity was observed in this study for dimeric DNA i-motif proton-binding, in which the receptors switch from mainly free to mainly bound over a broad change in ligand concentration; Strong positive cooperativity can be achieved by introducing preorganization in receptor-ligand binding, which is characterized by steeper sigmoidal binding curve and narrower transition width. (c) Speculated folding pathway of bimolecular i-motif: (i) two C-rich oligonucleotides 'nucleate' by forming several hemiprotonated $\text{CH}^+\text{-C}$ base pairs, (ii) a third C-tract collides with the nucleus, along with the 'zippering' of the nucleus, (iii) a fourth C-tract hybridizes with the third one in parallel orientation, which leads to the final thermodynamically most stable i-motif. Cytosines are drawn as red triangles, thymines as green circles, and proton ions as gray circles.

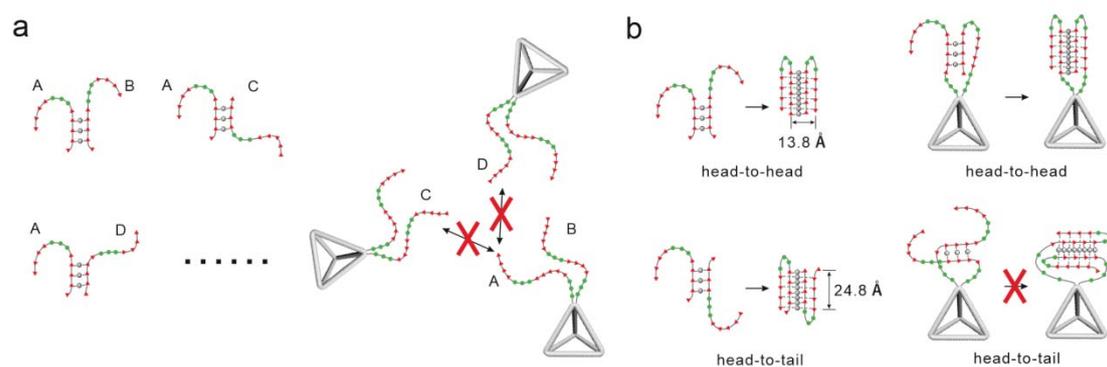


Figure S2. Preorganization effect of TDF in the folding of dimeric i-motif. (a) Left: In the folding pathway of dimeric i-motif, it is possible for strand A to base-pair with multiple strands (strand B, C, D, etc.), which leads to a significant number of intermediates; Right: After anchoring at the vertex, strand A preferentially base-pairs with strand B (the other strand at the same vertex), and no longer with other strands in their surroundings (strand C, D, etc.), which is indicated by PAGE analysis (Figure 1d). (b) Left: intermediates that can lead to head-to-head and head-to-tail conformations coexist in the folding pathway of dimeric i-motif; Right: As predicted from Monte Carlo simulations, the mean distance between endpoints of ployT spacer on TDF (1.04-1.85 nm, from 1T to 5T, Figure 3c) is smaller than the height of folded i-motif (eg., 2.48 nm for i-motif with 8 CH⁺-C base pairs), and thus prevent the complete folding of i-motif structure with head-to-tail conformation on TDF. In this way, TDF is able to diminish the intermediate species leading to head-to-tail conformation, and promote the formation of a single conformation (i.e., head-to-head). Cytosines are drawn as red triangles, thymines as green circles, and proton ions as gray circles.

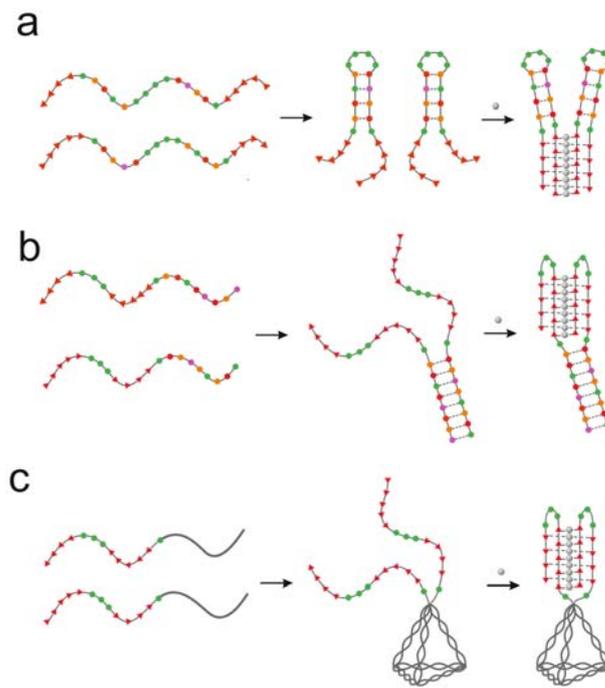


Figure S3. Proposed folding pathway of dimeric i-motif pre-organized by (a) small hairpin, (b) short duplex, and (c) TDF. Structures based on Watson-Crick base-pairing formed firstly, which brings the two C-rich strands into close proximity, and thereby leads to the final folded i-motif. The color coding of residues is as follows: red for cytidine, green for thymidine, orange for guanine, and purple for adenine.

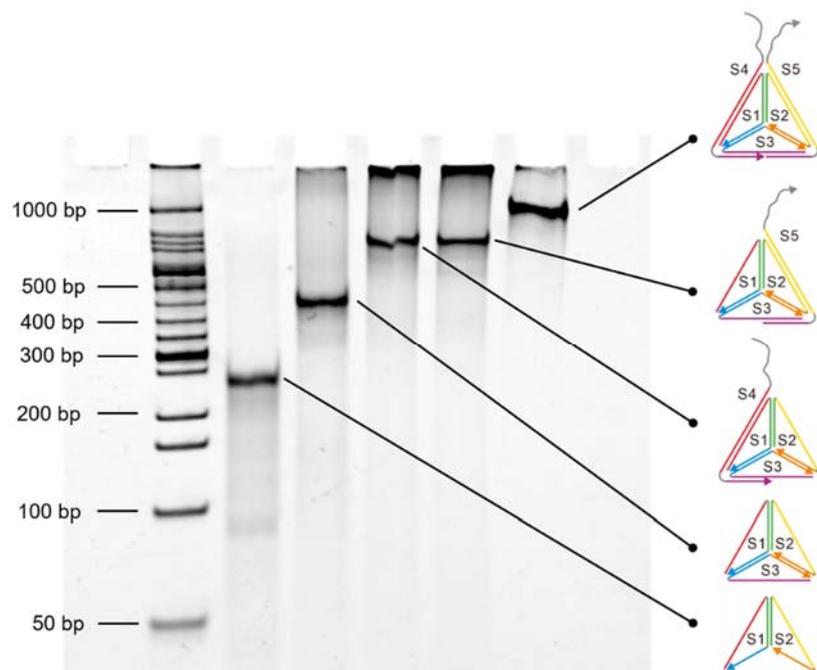


Figure S4. Electrophoresis characterization of the assembly of TDF containing i-motif-forming strands by 8% native PAGE.

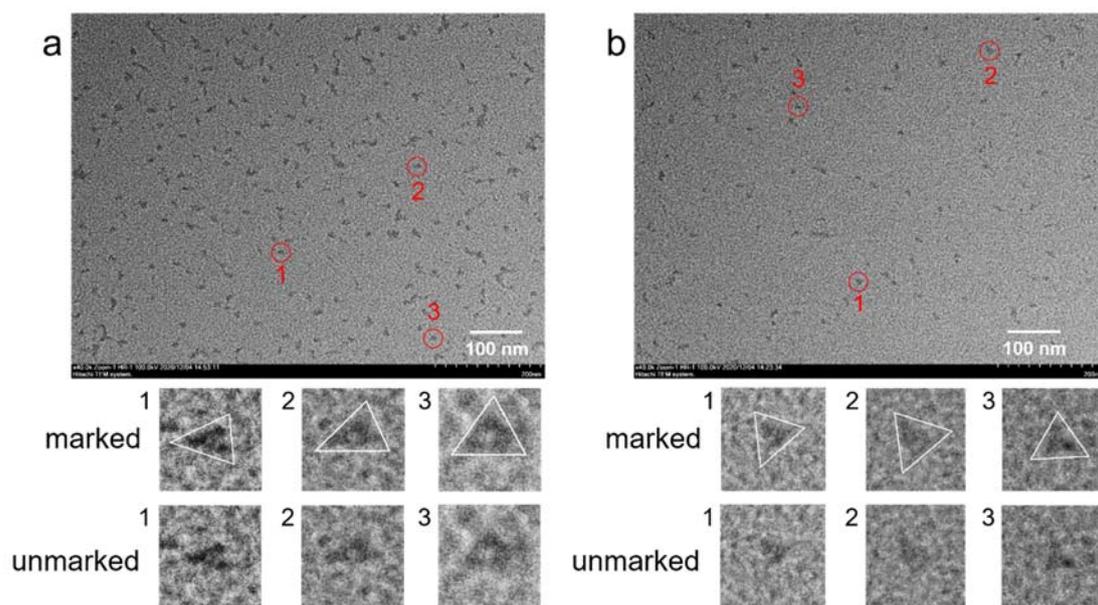


Figure S5. TEM images of TDF-(C₅T₃C₄)₂ prepared in PBS buffer of pH 5.5 (a) and 7.5 (b).

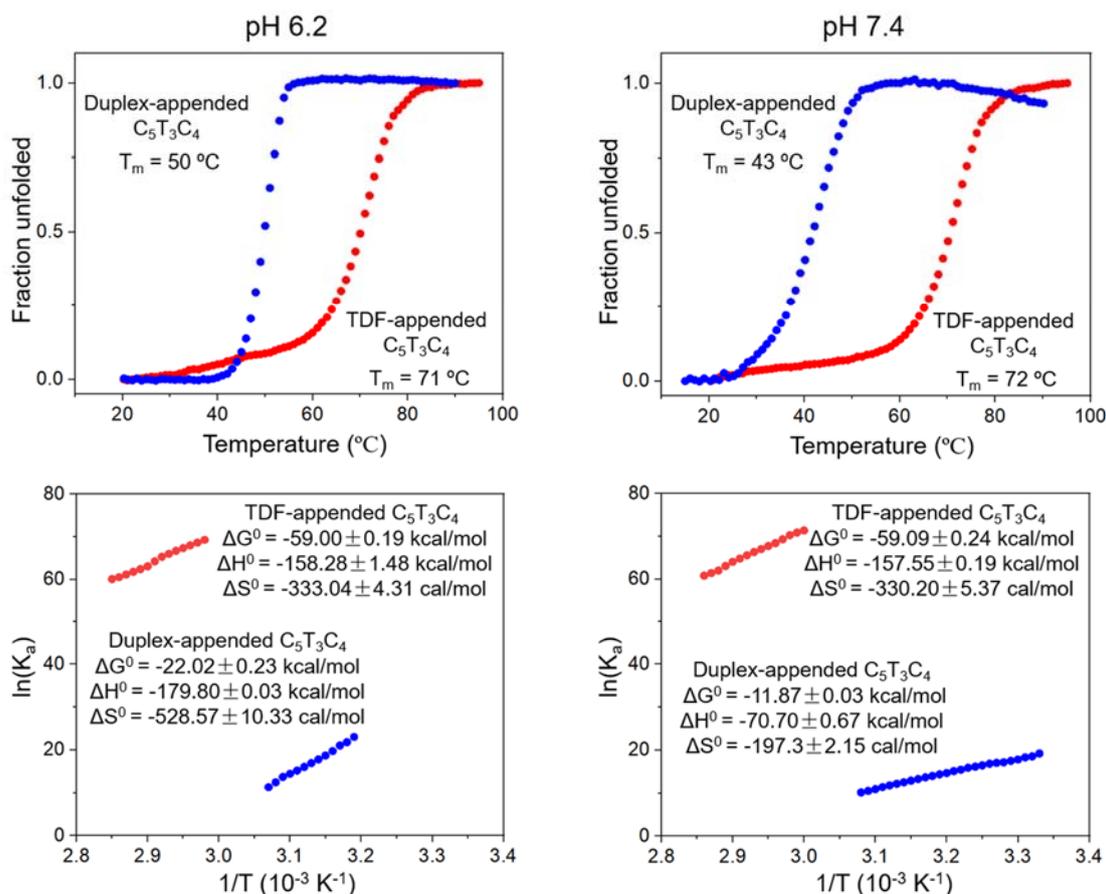


Figure S6. Thermal melting profiles and van't Hoff plots of 10 bp duplex-attached $C_5T_3C_4$ (2.0 μM) and TDF- $(C_5T_3C_4)_2$ (0.2 μM) in PBS buffer at pH 6.2 and 7.4, respectively. The 10 bp duplex provided the highest n_H value among the screened duplexes and its sequence is given in Table S2. Note that in order to acquire UV absorption of appropriate intensity, different concentrations were adopted for the two i-motifs because of large difference in their absorbance coefficients. Please note $C_5T_3C_4$ assisted by either duplex or TDF will not fold into i-motif structure at pH 7.4 under the current condition, and the melting at pH 7.4 is due to solely the dissociation of the appended duplex/TDF.

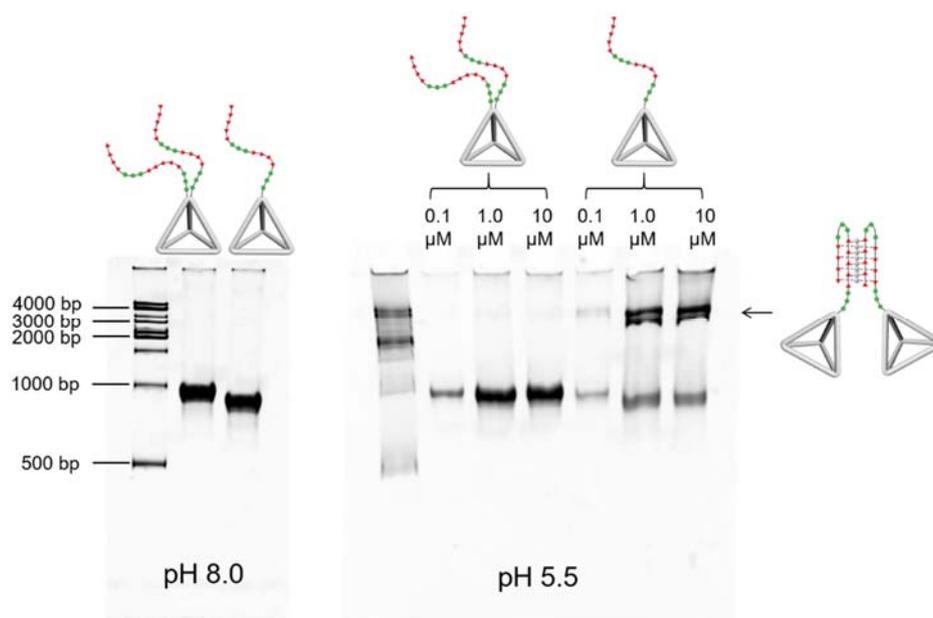


Figure S7. Migration behavior of i-motif-tethered TDFs in nondenaturing polyacrylamide gel (8%) at pH 8.0 and 5.5, respectively. 500 bp DNA ladder (Takara) was used as marker. At pH 8.0, both of the two types of TDFs showed single bands, indicating their monomolecular status; while at pH 5.5, bands with slower mobility dominated in the lanes of the control TDF with only one C-rich strand (having two C-tracts) extended from the vertex, abbreviated as TDF-C₅T₃C₄. Based on its mobility pattern with respect to the DNA ladder, the structure adopted by TDF-C₅T₃C₄ in the lower band was assumed to be dimeric TDF assemblies. On the contract, the TDF with two C-rich strands (each having two C-tracts) extended from the vertex, namely TDF-(C₅T₃C₄)₂, exhibited neglectable dimerization in the concentration range from 0.1 μM to 10 μM. Cytosines are drawn as red triangles, and thymines as green circles.

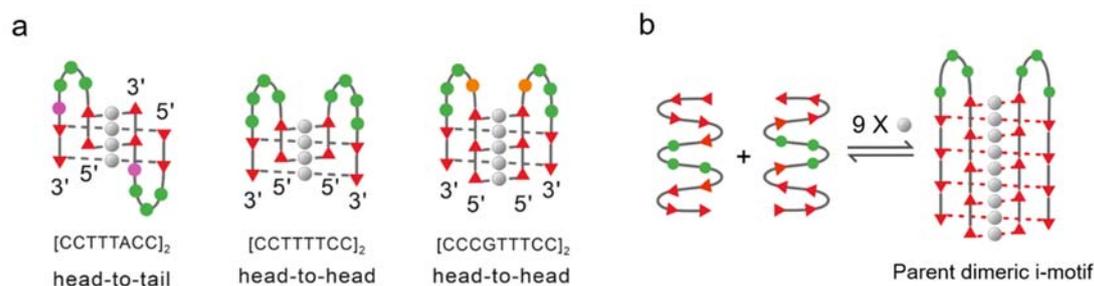
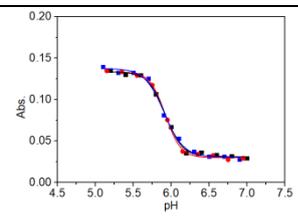
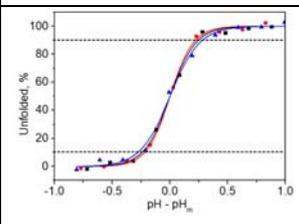
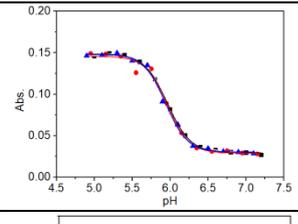
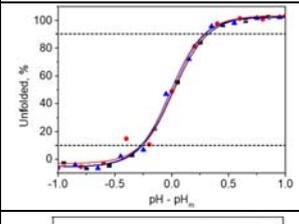
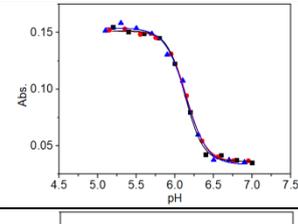
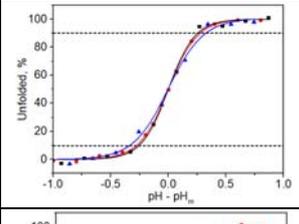
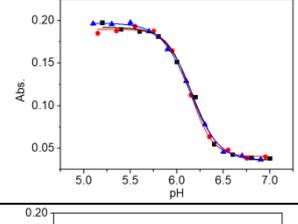
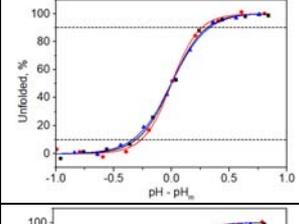
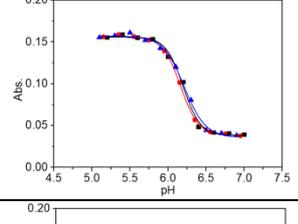
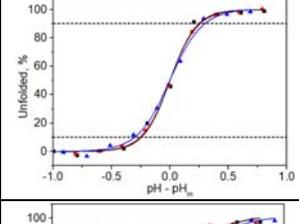
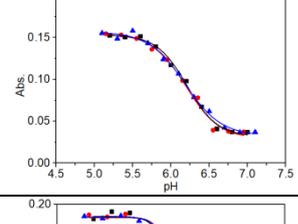
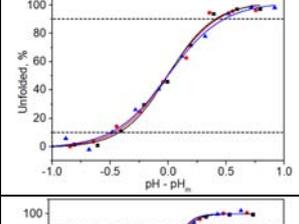
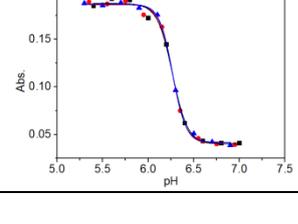
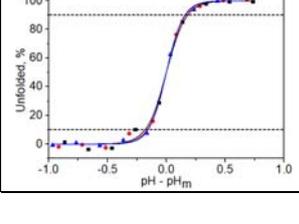
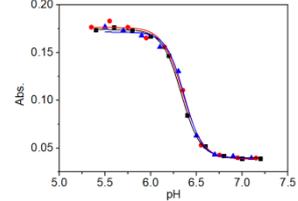
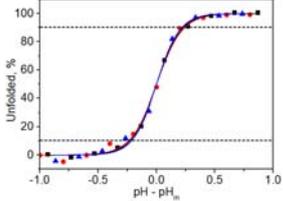
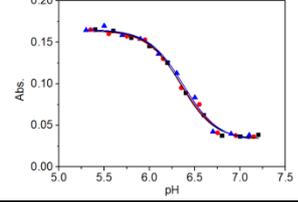
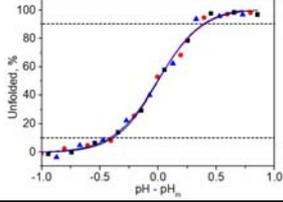
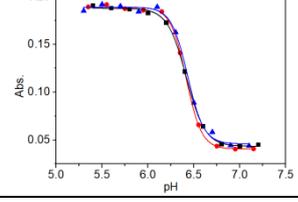
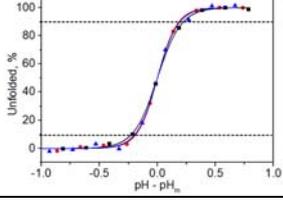
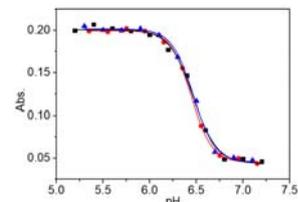
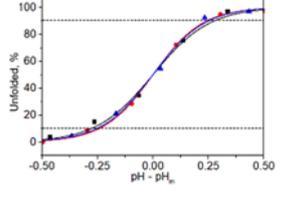
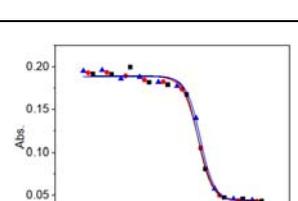
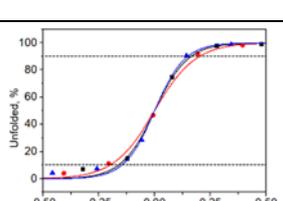


Figure S8. Dimeric i-motif formed by C₅T₃C₄ was chosen as the parent structure to be modified in in this work. (a) Dimeric i-motif can exist in two possible topologies: dimeric i-motifs bearing C-tracts of the same length (eg. CCTTTACC, CCTTTTCC and C₄TGTC₄) can fold into either head-to-head or head-to-tail topology,^{22,23} depending on its loop composition, but if one of the C-tracts contains one more C-residue than that of the other (eg. CCCGTTTCC), only the head-to-head arrangement is thermodynamically favored.²⁴ Therefore, in order to eliminate the complication caused by the potential coexistence of different topologies, bimolecular i-motif formed by a oligonucleotide bearing C-tracts of unequal lengths, C₅T₃C₄ (b), was thus chosen as the parent structure to be modified in different ways in this work. The color coding of residues is as follows: red for cytosine, green for thymine, orange for guanine, purple for adenine, and gray for proton ion.

Table S5. pH denaturation data of dimeric i-motifs formed by C-rich oligonucleotides in the form of $C_mX_3C_n$ ($X = A$ or T ; $m, n = 4, 5$ or 6). The data are arranged with increasing pH_m . The n_H data are depicted in Figure S9.

DNA	pH denaturation profile	pH_m	Transition width	ΔpH	n_H
C_4AATC_4		5.92 ± 0.01		0.47 ± 0.03	3.9 ± 0.5
C_4AAAC_4		5.95 ± 0.02		0.65 ± 0.02	3.0 ± 0.1
C_4TTTC_4		6.15 ± 0.02		0.53 ± 0.02	3.5 ± 0.3
C_5AAAC_5		6.15 ± 0.01		0.64 ± 0.02	3.2 ± 0.4
C_4TTTC_5		6.19 ± 0.03		0.51 ± 0.01	3.6 ± 0.4
C_4TTAC_4		6.20 ± 0.02		0.86 ± 0.03	2.1 ± 0.2
C_5AATC_5		6.27 ± 0.01		0.33 ± 0.01	5.7 ± 0.4

C ₅ TTAC ₅		6.35±0.02		0.46±0.03	4.2±0.2
C ₅ TTTC ₄		6.35±0.02		0.81±0.01	2.4±0.1
C ₅ TTTC ₅		6.42±0.01		0.38±0.01	4.8±0.5
C ₅ TTTC ₆		6.47±0.01		0.53±0.03	3.5±0.2
C ₆ TTTC ₅		6.53±0.01		0.34±0.02	5.4±0.4

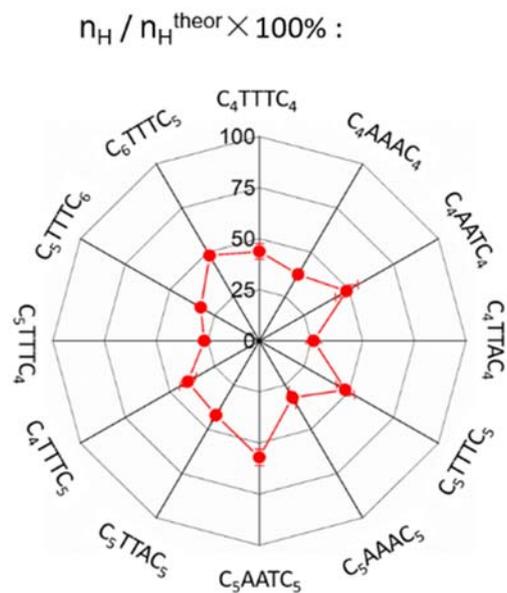
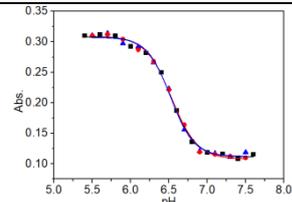
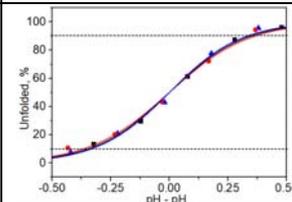
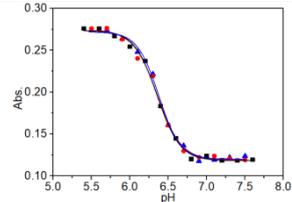
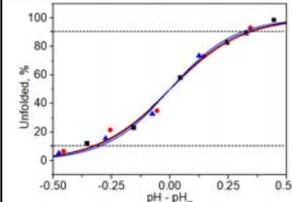
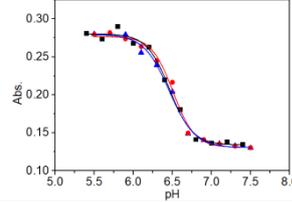
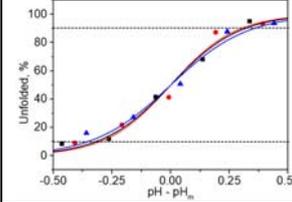
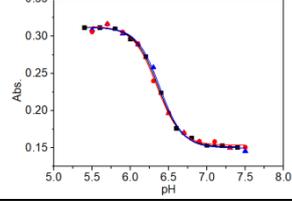
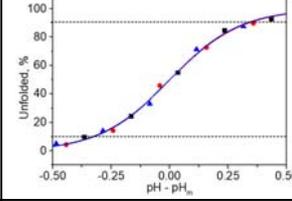
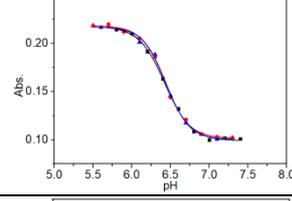
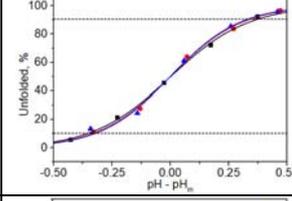
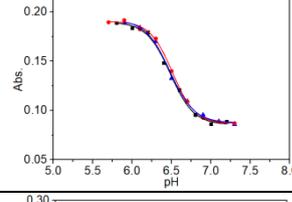
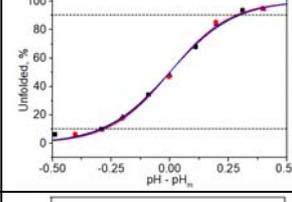
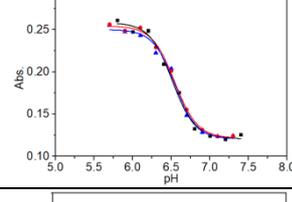
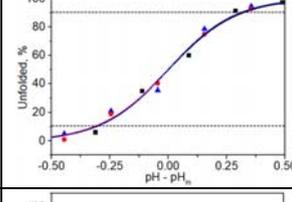
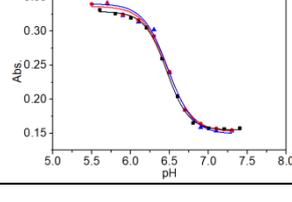
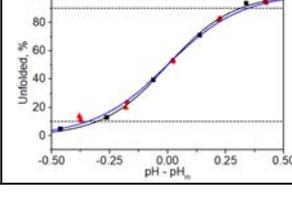


Figure S9. The folding cooperativity of unmodified bimolecular i-motifs. The observed Hill coefficients (n_H) of commonly used bimolecular i-motif are within 50% of the theoretical maximum. The observed n_H value for each oligonucleotide was listed in Table S5. The maximum theoretically possible n_H value (n_H^{theor}) equals the number of binding sites.

Table S6. pH denaturation data of hairpin-assisted i-motifs (hairpins of different sizes were tested). Fragments composing hairpin stems are underlined.

DNA	pH denaturation profile	pH _m	Transition width	ΔpH	n _H
CCCCCTG <u>ACTTTGT</u> CTCCCC		6.52±0.01		0.70±0.0 4	2.7± 0.1
CCCCCTC <u>GACTTTG</u> TCGTCCC C		6.36±0.01		0.65±0.0 4	2.9± 0.2
CCCCCTG <u>CGACTTT</u> <u>GTCGCTC</u> CCC		6.48±0.03		0.65±0.0 6	2.9± 0.3
CCCCCTC <u>GCGACTT</u> <u>TGTCGCG</u> TCCCC		6.36±0.03		0.66±0.0 1	3.0± 0.1
CCCCCTG <u>ACTTTTG</u> <u>TCTCCCC</u>		6.43±0.01		0.68±0.0 4	3.0± 0.1
CCCCCTC <u>GACTTTT</u> <u>GTCGTCC</u> CC		6.50±0.01		0.51±0.0 3	3.7± 0.1
CCCCCTG <u>CGACTTT</u> <u>TGTCGCT</u> CCCC		6.54±0.02		0.63±0.0 1	3.4± 0.1
CCCCCTC <u>GCGACTT</u> <u>TTGTCGC</u> <u>GTCCCC</u>		6.47±0.01		0.67±0.0 5	3.3± 0.1

CCCCCTG <u>A</u> CTTTT GTCTCCC C		6.63 ± 0.01		0.61 ± 0.0 2	3.0 ± 0.1
CCCCCTC <u>G</u> ACTTTT TGTCGTC CCC		6.49 ± 0.02		0.55 ± 0.0 1	3.6 ± 0.1
CCCCCTG <u>C</u> GACTTT TTGTCGC TCCCC		6.44 ± 0.02		0.57 ± 0.0 4	3.3 ± 0.2
CCCCCTC <u>G</u> CGACTT TTTGTCG <u>C</u> GTCCCC		6.38 ± 0.02		0.58 ± 0.0 4	3.3 ± 0.2

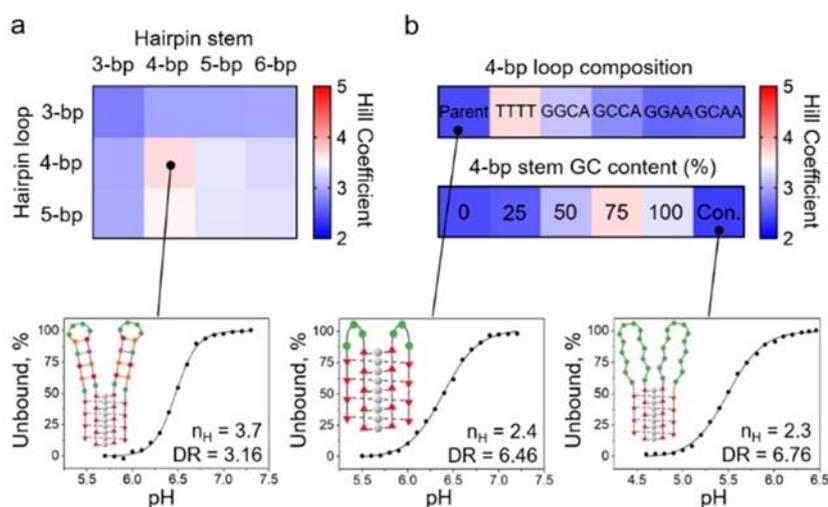


Figure S10. Effects of the integration of small DNA hairpins on the apparent folding cooperativity (in terms of Hill coefficient, n_H) of the dimeric *i*-motif formed by $C_5T_3C_4$. (a) Small hairpins with stems and loops of different lengths were incorporated in the loop of $C_5T_3C_4$, and the n_H values for the resulting DNA structures were shown in the heatmap (data from Table S6). As for the influence of the hairpin's loop length and stem length on dimeric *i*-motif's folding cooperativity, tetraloop and pentaloop hairpins generally increased the n_H value to a larger extent than triloop hairpins, and hairpins whose stem consisting of 4 base pairs exhibited larger n_H value than those with either shorter (3-bp) or longer (5-bp and 6-bp) stem, probably because though the presence of longer stem gives higher thermodynamic stability, it increases the likelihood of self-dimer formation. (b) The influence of hairpin's loop and stem base compositions on the n_H value (data from Table S7 and S8): hairpin stems of high G-C content generally contributed to large n_H values, while loops recently screened to be thermodynamically most stable in the whole tetraloop hairpin library (i.e., GGCA, GCCA, GGAA and GCAA)³ gave somewhat smaller n_H values than that of those with ploy(dT) loops. Typical pH denaturation curves are included for the parent *i*-motif ($n_H = 2.4$), a 4-bp stem/4-bp loop-hairpin-modified dimeric *i*-motif ($n_H = 3.7$), and the control *i*-motif with 14-bp ploy T loops ($n_H = 2.3$). Please note that the maximum theoretically possible n_H value for this group of *i*-motif structures is 9. The color coding of residues in the structure scheme is as follows: red for cytosine, green for thymine, orange for guanine, purple for adenine, and gray for proton ion. DR denotes dynamic range.

Table S7. pH denaturation data of hairpin-assisted i-motifs (hairpins with stems of different GC contents were tested). Fragments composing hairpin stems are underlined.

DNA	pH denaturation profile	pH _m	Transition width	ΔpH	n _H
CCCCCTG <u>CGCTTTT</u> <u>GCGCTCC</u> CC (100%)		6.43±0.02		0.57±0.0 4	3.3± 0.2
CCCCCTC <u>GACTTTT</u> <u>GTCGTCC</u> CC (75%)		6.50±0.01		0.58±0.0 2	3.5± 0.1
CCCCCTC <u>AACTTTT</u> <u>GTGTCC</u> CC (50%)		6.34±0.02		0.62±0.0 3	3.1± 0.2
CCCCCTC <u>AAITTTT</u> <u>ATTGTCC</u> CC (25%)		6.35±0.02		0.76±0.0 4	2.5± 0.1
CCCCCTT <u>AAITTTT</u> <u>ATTATCC</u> CC (0%)		6.18±0.02		0.81±0.0 4	2.4± 0.2
CCCCCTT TTTTTTT TTTTTCC CC (control)		5.52±0.01		0.83±0.0 3	2.3± 0.1

Table S8. pH denaturation data of hairpin-assisted i-motifs (hairpins with loops of different base compositions were tested). Adopted were loops recently screened to be thermodynamically most stable in the whole tetraloop hairpin library (i.e., GGCA, GCCA, GGAA and GCAA).³ Fragments composing hairpin stems are underlined.

DNA	pH denaturation profile	pH _m	Transition widths	ΔpH	n _H
CCCCCTC <u>GACGGC</u> AGTCGTC CCC		6.18±0.02		0.61±0.0 5	3.1± 0.3
CCCCCTC <u>GACGCC</u> AGTCGTC CCC		6.24±0.03		0.68±0.0 6	2.8± 0.2
CCCCCTC <u>GACGGA</u> AGTCGTC CCC		6.08±0.04		0.74±0.1	2.6± 0.4
CCCCCTC <u>GACGCA</u> AGTCGTC CCC		6.25±0.04		0.74±0.0 1	2.6± 0.1

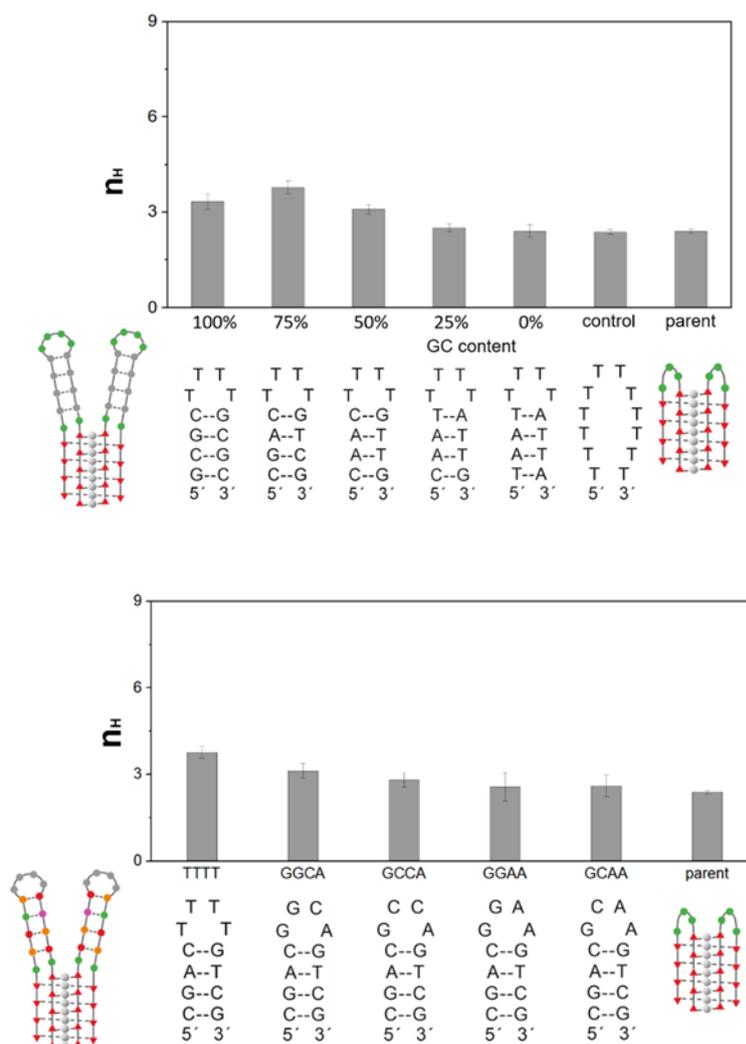


Figure S11. The observed Hill coefficient (n_H) of hairpin-assisted i-motifs showed dependence on hairpin stem's GC content (top) and loop's base composition (bottom). The full sequences of hairpins integrated in the two loops of dimeric $C_5T_3C_4$ i-motif are illustrated. It should be noted that the maximum theoretically possible n_H value for this group of i-motifs is 9.

Table S9. pH denaturation data of dimeric i-motifs assisted by short duplexes of varying length.

Fragments composing complementary sequences are underlined.

DNA	ΔG^0 of duplex	pH denaturation profile	pH_m	Transition width	ΔpH	Dynamic range	nH
<u>CGCTC</u> ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGCC</u>	-2.7 (3-bp)		6.42 ± 0.01		0.60 ± 0.06	3.98 ± 1.15	3.1 ± 0.2
<u>CGGCTC</u> ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGCC</u> <u>G</u>	-4.6 (4-bp)		6.53 ± 0.02		0.54 ± 0.05	3.47 ± 1.12	3.6 ± 0.4
<u>CGCAGTC</u> ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCIG</u> <u>CG</u>	-5.5 (5-bp)		6.64 ± 0.02		0.46 ± 0.07	2.88 ± 1.17	4.1 ± 0.5
<u>CGCAACTC</u> ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TGTT</u> <u>GCG</u>	-6.8 (6-bp)		6.71 ± 0.01		0.41 ± 0.04	2.57 ± 1.10	4.5 ± 0.4
<u>GCAGTCGT</u> C ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGA</u> <u>CTGC</u>	-8.4 (7-bp)		6.84 ± 0.02		0.36 ± 0.04	2.29 ± 1.10	5.4 ± 0.4
<u>AGCACTCG</u> TC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGA</u> <u>GTGCT</u>	-9.6 (8-bp)		6.94 ± 0.01		0.32 ± 0.02	2.11 ± 1.05	5.9 ± 0.7
<u>GCATCAAC</u> <u>GTC</u> ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGT</u> <u>TGATGC</u>	-10.5 (9-bp)		7.04 ± 0.02		0.30 ± 0.02	1.99 ± 1.04	6.3 ± 0.6
<u>GCATCAGA</u> <u>CGTC</u> ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGT</u> <u>CTGATGC</u>	-12.1 (10-bp)		7.09 ± 0.02		0.27 ± 0.03	1.86 ± 1.07	7.1 ± 0.6

<u>GCTGTTAT</u> <u>GCGTC₅T₃C₄</u> / <u>C₅T₃C₄TCGC</u> <u>ATAACAGC</u>	-13.4 (11-bp)		7.06± 0.01		0.28 ±0.03	1.91 ±1.07	6.9 ±0.6
<u>GCTAGTCA</u> <u>GACGTC₅T₃</u> C ₄ / <u>C₅T₃C₄TCGT</u> <u>CTGACTAG</u> <u>C</u>	-14.5 (12-bp)		7.05± 0.01		0.29 ±0.04	1.95 ±1.09	6.8 ±0.5
<u>TTTTTTTTT</u> <u>ITC₅T₃C₄</u> / <u>C₅T₃C₄TTTT</u> <u>TTTTTTT</u>	Control		6.21 ±0.02		0.70 ±0.09	5.01 ±1.23	2.7 ±0.4

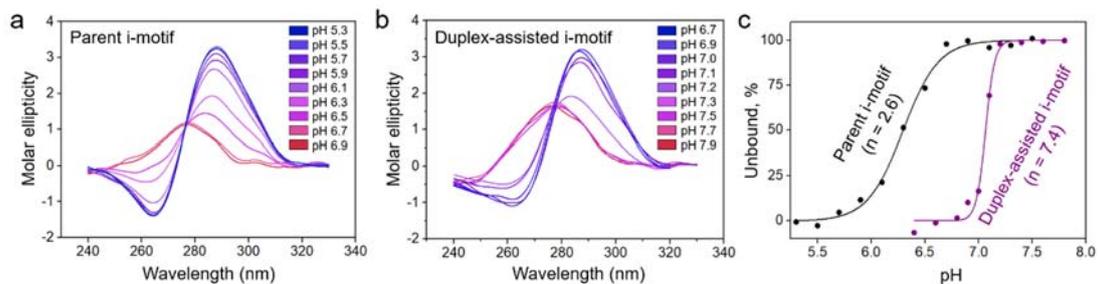
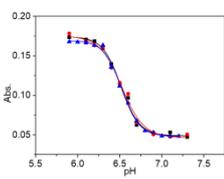
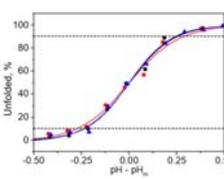
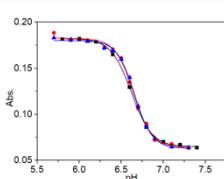
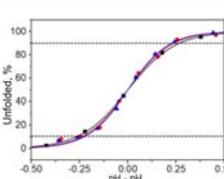
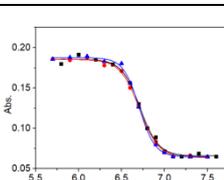
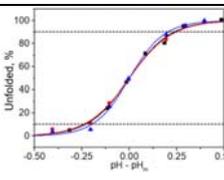
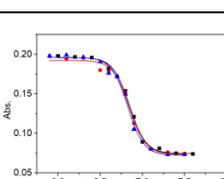
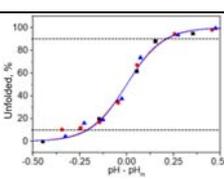
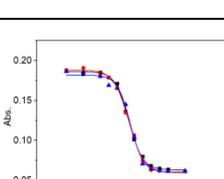
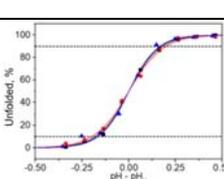
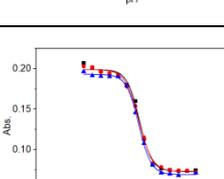
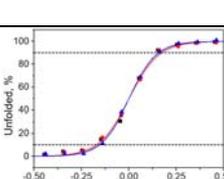
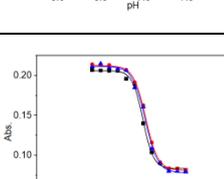
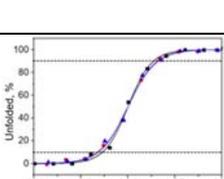
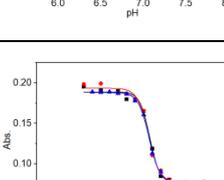
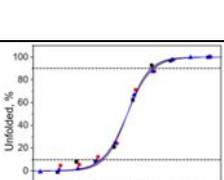


Figure S12. CD spectra comparison for i-motifs before and after integration of short duplex. CD spectra of (a) parent i-motif formed by C₅T₃C₄ (5 μM) and (b) its 10-bp duplex assisted variation, GCATCAGACGTC₅T₃C₄/C₅T₃C₄TCGTCTGATGC, (2.5 μM for each oligonucleotide) were recorded at varying pH. In acidic buffer, a positive band at 287 nm and a negative band at 264 nm were observed, indicating the formation of i-motif structure.^{25,26} When the pH was increased, the positive band was shifted to 277 nm. (c) The molar ellipticity at 287 nm was normalized and plotted against pH value. The n_H for the C₅T₃C₄ i-motif before and after the duplex integration was calculated to be 2.6 and 7.4, respectively, which was in accordance with the UV study.

Table S10. pH denaturation data of dimeric i-motifs assisted by 8-bp duplex of different GC contents.

DNA	ΔG^0 of duplex	pH denaturation profile	pH_m	Transition width	ΔpH	Dynamic range	nH
<u>ATAAAAT</u> <u>ATC₅T₃C₄</u> / <u>C₅T₃C₄TTA</u> <u>TTTTAT</u>	-4.1 (0%GC)		6.52 ± 0.01		0.51 ± 0.05	3.24 ± 1.12	3.6 ± 0.3
<u>ATACTAT</u> <u>ATC₅T₃C₄</u> / <u>C₅T₃C₄TTA</u> <u>TAGTAT</u>	-4.9 (13%GC)		6.64 ± 0.02		0.49 ± 0.04	3.08 ± 1.09	3.9 ± 0.3
<u>AATTCAT</u> <u>CTC₅T₃C₄</u> / <u>C₅T₃C₄TG</u> <u>ATGAATT</u>	-6.1 (25%GC)		6.71 ± 0.01		0.45 ± 0.05	2.84 ± 1.12	4.2 ± 0.5
<u>ACTTCAT</u> <u>CTC₅T₃C₄</u> / <u>C₅T₃C₄TG</u> <u>ATGAAGT</u>	-7.1 (38%GC)		6.84 ± 0.01		0.42 ± 0.01	2.65 ± 1.01	4.5 ± 0.2
<u>GAACAA</u> <u>CGTC₅T₃C</u> 4 / <u>C₅T₃C₄TC</u> <u>GTIGTTC</u>	-8.3 (50%GC)		6.87 ± 0.02		0.36 ± 0.03	2.30 ± 1.07	5.3 ± 0.5
<u>AGCACTC</u> <u>GTC₅T₃C₄</u> / <u>C₅T₃C₄TC</u> <u>GAGTGCT</u>	-9.6 (63%GC)		6.94 ± 0.01		0.32 ± 0.02	2.11 ± 1.05	5.9 ± 0.7
<u>GCACGA</u> <u>CGTC₅T₃C</u> 4 / <u>C₅T₃C₄TC</u> <u>GTCGTGC</u>	-10.8 (75%GC)		7.02 ± 0.02		0.30 ± 0.03	2.01 ± 1.07	6.3 ± 0.7
<u>GCACGG</u> <u>CGTC₅T₃C</u> 4 / <u>C₅T₃C₄TC</u> <u>GCCGTGC</u>	-12.2 (88%GC)		7.06 ± 0.01		0.27 ± 0.02	1.87 ± 1.04	7.0 ± 0.5

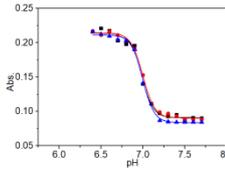
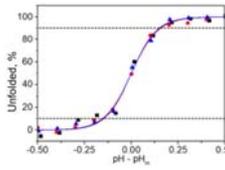
<u>GCGCCGC</u> <u>CTC₅T₃C₄</u> / <u>C₅T₃C₄TG</u> <u>GCGGCG</u> <u>C</u>	-13.3 (100%G C)		6.99 ±0.02		0.30 ±0.01	2.01 ±1.02	6.4 ±0.4
---	-----------------------	---	---------------	--	---------------	---------------	-------------

Table S11. pH denaturation data of dimeric i-motifs assisted by 10-bp duplex with ployT spacer of different lengths. Fragments composing complementary sequences are underlined.

DNA	pH denaturation profile ^[a]	pH _m	Transition width	ΔpH	n _H
<u>GCATCAGACG</u> TC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TCGTCT</u> <u>GATGC</u>		7.01±0.01		0.27±0.03	7.1±0.6
<u>GCATCAGACG</u> TTC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTCGTC</u> <u>TGATGC</u>		7.07±0.02		0.31±0.03	6.1±0.5
<u>GCATCAGACG</u> TTTC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTTCGT</u> <u>CTGATGC</u>		7.10±0.01		0.42±0.05	4.6±0.5
<u>GCATCAGACG</u> TTTTTC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTTTCGT</u> <u>CTGATGC</u>		7.09±0.02		0.48±0.07	4.3±0.4
<u>GCATCAGACG</u> TTTTTC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTTTTCG</u> <u>TCTGATGC</u>		7.09±0.01		0.46±0.04	4.2±0.3
<u>GCATCAGACG</u> TTTTTTC ₅ T ₃ C ₄ / C ₅ T ₃ C ₄ <u>TTTTTTC</u> <u>GTCTGATGC</u>		7.07±0.02		0.45±0.06	4.2±0.3

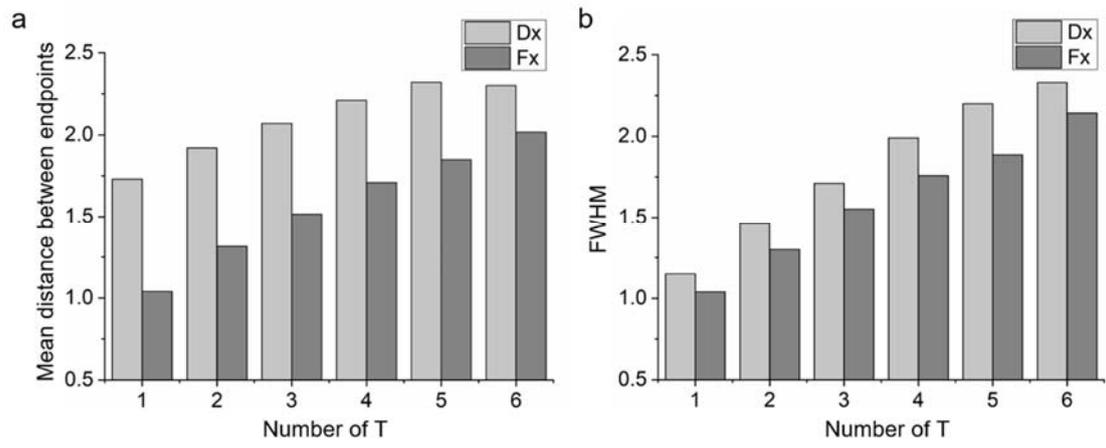


Figure S13. The mean distance between endpoints of ployT spacer (a) and FWHM (full width at half maxima, b) derived from Monte Carlo simulation, as a function of the number of T in the spacer.

Table S12. pH denaturation data of TDF-assisted C₅T₃C₄ i-motif with polyT spacer of different lengths.

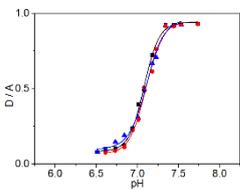
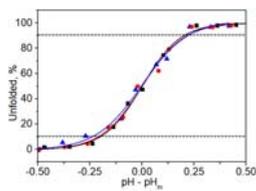
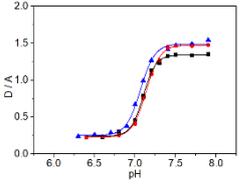
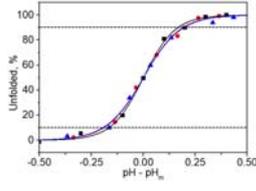
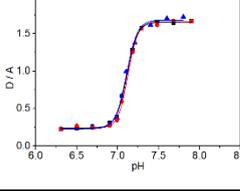
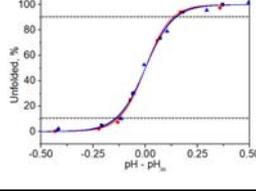
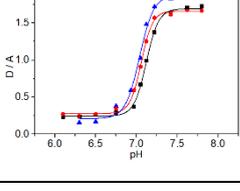
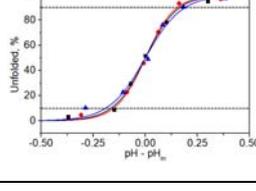
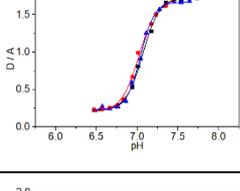
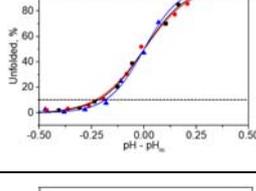
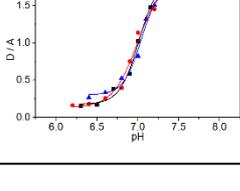
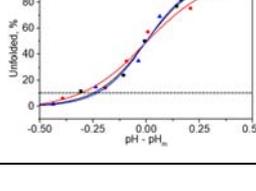
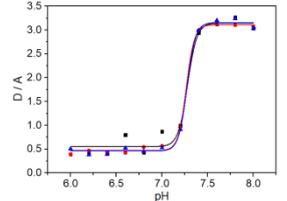
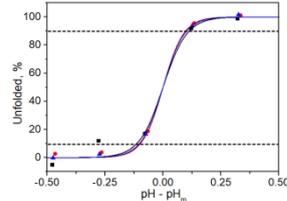
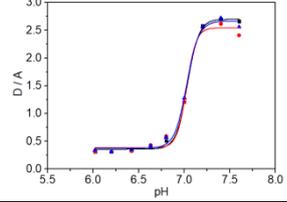
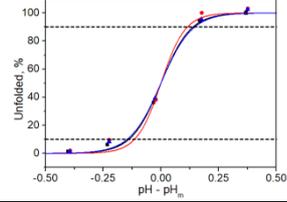
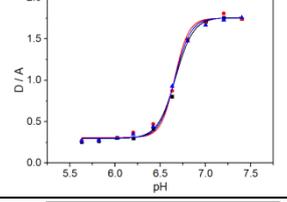
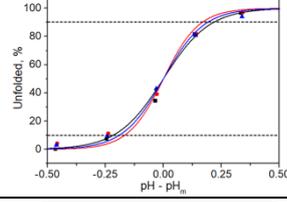
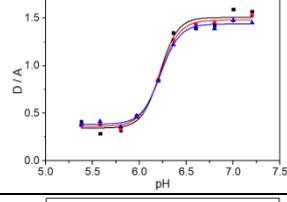
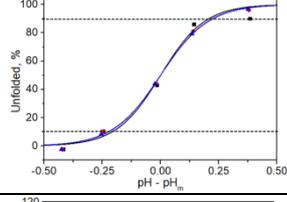
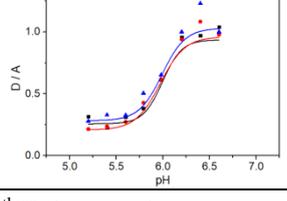
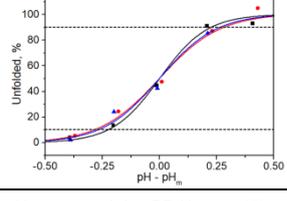
DNA	pH denaturation profiles	pH _m	Transition widths	ΔpH	n _H
TDF- (C ₅ T ₃ C ₄) ₂ with 1T spacer		7.09±0.02		0.43±0.05	4.5±0.5
TDF- (C ₅ T ₃ C ₄) ₂ with 2T spacer		7.10±0.03		0.39±0.05	5.1±0.3
TDF- (C ₅ T ₃ C ₄) ₂ with 3T spacer		7.11±0.01		0.29±0.02	6.8±0.4
TDF- (C ₅ T ₃ C ₄) ₂ with 4T spacer		7.07±0.04		0.31±0.03	6.1±0.5
TDF- (C ₅ T ₃ C ₄) ₂ with 5T spacer		7.04±0.03		0.41±0.06	4.5±0.6
TDF- (C ₅ T ₃ C ₄) ₂ with 6T spacer		7.01±0.03		0.55±0.06	3.6±0.5

Table S13. pH denaturation data of TDF-assisted $C_mX_3C_n$ ($m, n = 2, 3, 4, 5$) i-motifs.

DNA	pH denaturation profiles	pH_m	Transition widths	ΔpH	n_H	$n_H^{\text{theor [a]}}$
TDF- ($C_5T_3C_5$) ₂		7.27 ± 0.01		0.21 ± 0.02	9.2 ± 0.8	10
TDF- ($C_4T_3C_4$) ₂		7.03 ± 0.01		0.27 ± 0.04	7.1 ± 0.9	8
TDF- ($C_3T_3C_3$) ₂		6.66 ± 0.01		0.38 ± 0.05	5.0 ± 0.7	6
TDF- ($C_3T_3C_2$) ₂		6.22 ± 0.01		0.44 ± 0.02	4.4 ± 0.2	5
TDF- ($C_2T_3C_2$) ₂		5.98 ± 0.02		0.52 ± 0.07	3.7 ± 0.5	4

Note: [a] n_H^{theor} denotes the maximum theoretically possible Hill coefficient.

References:

- (1) J. L. Mergny, L. Lacroix, Analysis of thermal melting curves. *Oligonucleotides* **2003**, *13*, 515–537.
- (2) L. A. Marky, K. J. Breslauer, Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **1987**, *26*, 1601–1620.
- (3) J. Wang, P. Dong, W. Wu, X. Pan, X. Liang. High-throughput thermal stability assessment of DNA hairpins based on high resolution melting. *J. Biomol. Struct. Dyn.* **2018**, *36*, 1–13.
- (4) E. M. Moody, P. C. Bevilacqua. Thermodynamic coupling of the loop and stem in unusually stable DNA hairpins closed by CG base pairs. *J. Am. Chem. Soc.* **2003**, *125*, 2032–2033.
- (5) P. M. Vallone, T. M. Paner, J. Hilario, M. J. Lane, B. D. Faldasz, A. S. Benight. Melting studies of short DNA hairpins: Influence of loop sequence and adjoining base pair identity on hairpin thermodynamic stability. *Biopolymers* **1999**, *50*, 425–442.
- (6) Y. Shen, S. V. Kuznetsov, A. Ansari. Loop dependence of the dynamics of DNA hairpins. *J. Phys. Chem. B* **2001**, *105*, 12202–12211.
- (7) S. V. Kuznetsov, C. C. Ren, S. A. Woodson, A. Ansari. Loop dependence of the stability and dynamics of nucleic acid hairpins. *Nucleic Acids Res.* **2008**, *36*, 1098–1112.
- (8) S. V. Kuznetsov, Y. Shen, A. S. Benight, A. Ansari. A semiflexible polymer model applied to loop formation in DNA hairpins. *Biophys. J.* **2001**, *81*, 2864–2875.
- (9) J. L. Mergny, L. Lacroix, C. Hélène, X. Han, J. L. Leroy, Intramolecular folding of pyrimidine oligodeoxynucleotides into an I-DNA motif. *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898.
- (10) T. Li, M. Famulok, I-motif-programmed functionalization of DNA nanocircles. *J. Am. Chem. Soc.* **2013**, *135*, 1593–1599.
- (11) A. M. Fleming, Y. Ding, R. A. Rogers, J. Zhu, J. Zhu, A. D. Burton, C. B. Carlisle, C. J. Burrows, 4n-1 Is a “sweet spot” in DNA i-motif folding of 2'-deoxycytidine homopolymers. *J. Am. Chem. Soc.* **2017**, *139*, 4682–4689.
- (12) P. Wright, M. A. S. Abdelhamid, M. O. Ehiabor, C. Grigg, K. Irving, N. M. Smith, Epigenetic modification of cytosines fine tunes the stability of i-motif DNA. **2020**, *48*, 55–62.
- (13) S. P. Gurung, C. Schwarz, J. P. Hall, C. J. Cardin, J. A. Brazier, The importance of loop length on the stability of i-motif structures. *Chem. Commun.* **2015**, *51*, 5630–5632.
- (14) J. L. Mergny, L. Lacroix, Kinetics and thermodynamics of i-DNA formation: phosphodiester versus modified oligodeoxynucleotides. *Nucleic Acids Res.* **1998**, *26*, 4797–4803.
- (15) A. J. Simon, A. Vallée-Bélisle, F. Ricci, H. M. Watkins, K. W. Plaxco, Using the population-shift mechanism to rationally introduce “Hill-type” cooperativity into a normally non-cooperative receptor, *Angew. Chem. Int. Ed.* **2014**, *53*, 9471-9475.
- (16) C. A. Hunter, H. L. Anderson. What is cooperativity? *Angew. Chem. Int. Ed.* **2009**, *48*, 7488–7499.
- (17) M. A. Mullen, S. M. Assmann, P. C. Bevilacqua. Toward a digital gene response: RNA G-quadruplexes with fewer quartets fold with higher cooperativity. *J. Am. Chem. Soc.* **2012**, *134*, 812–815.
- (18) C. K. Kwok, M. E. Sherlock, P. C. Bevilacqua. Decrease in RNA folding cooperativity by deliberate population of intermediates in RNA G-quadruplexes, *Angew. Chem. Int. Ed.* **2013**, *52*, 683-686.
- (19) M. Canalia, J. L. Leroy. Structure, internal motions and association-dissociation kinetics of the i-motif dimer of d(5mCCTCACTCC). *Nucleic Acids Res.* **2005**, *33*, 5471–5481.

- (20) G. Devi, L. He, B. Xu, T. Li, F. Shao, In-stem thiazole orange reveals the same triplex intermediate for pH and thermal unfolding of i-motifs. *Chem. Commun.* **2016**, 52, 7261–7264.
- (21) T. E. Ouldrige, P. Šulc, F. Romano, J. P. K. Doye, A. A. Louis. DNA hybridization kinetics: Zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* **2013**, 41, 8886–8895.
- (22) S. Nonin, A. T. Phan, J. L. Leroy. Solution structure and base pair opening kinetics of the i-motif dimer of d(5mCCTTTACC): A noncanonical structure with possible roles in chromosome stability. *Structure* **1997**, 5, 1231–1247.
- (23) P. Catasti, X. Chen, L. L. Deaven, R. K. Moyzis, E. M. Bradbury, G. Gupta. Cytosine-rich strands of the insulin minisatellite adopt hairpins with intercalated cytosine⁺·cytosine pairs. *J. Mol. Biol.* **1997**, 272, 369–382.
- (24) J. Gallego, S. H. Chou, B. R. Reid. Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad: Solution structure of the d(TCCCGTTTCCA) dimer. *J. Mol. Biol.* **1997**, 273, 840–856.
- (25) E. P. Wright, J. L. Huppert, Z. A. E. Waller, Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* **2017**, 45, 2951–2959.
- (26) J. Kypr, I. Kejnovská, D. Renčiuk, M. Vorlíčková, Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **2009**, 37, 1713–1725.