Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Data

Site Specifically Probing the Unfolding Process of Human Telomere i-motif DNA Using Vibrationally Enhanced Alkynyl Stretch

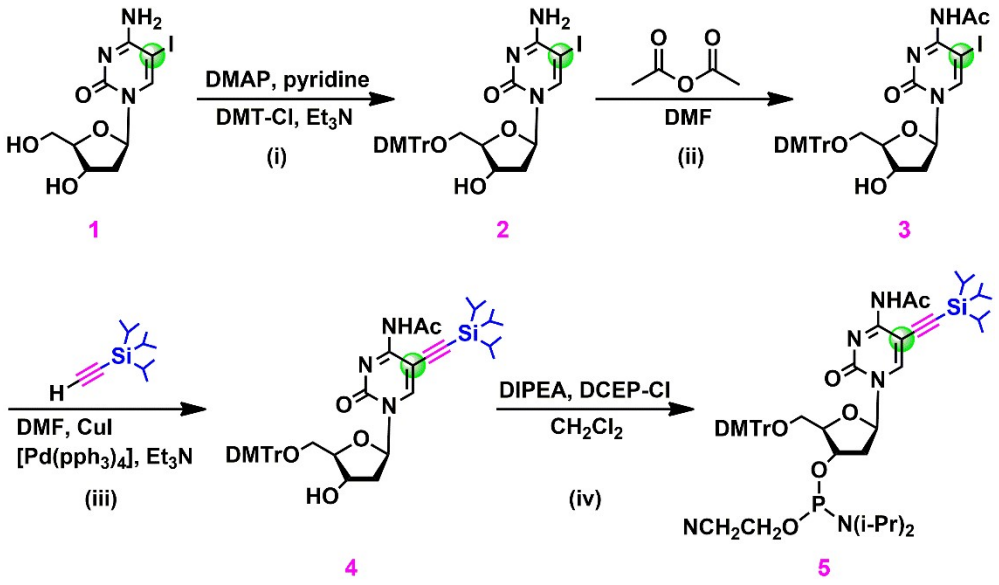
Tiantian Dong[†‡], *Pengyun Yu*[†‡], *Juan Zhao*[†‡], *and Jianping Wang*[†‡*]

[†] *Beijing National Laboratory for Molecular Sciences, Molecular Reaction Dynamics Laboratory, CAS Research/Education Center for Excellence in Molecular Sciences, Institute of Chemistry, Chinese Academy of Sciences, Beijing, 100190, P. R. China*

[‡] *University of Chinese Academy of Sciences, Beijing 100049, P. R. China*

***** *Corresponding author, Tel: (+86)-010-62656806; Fax: (+86)-010-62563167; E-mail: jwang@iccas.ac.cn*

A. Synthesis of ^{TIPS}EdC phosphoramidite



Scheme S1. Chemical reaction reagents and conditions. (i) DMAP, pyridine, DMT-Cl, Et_3N, DMF, r.t.76%; (ii) $(CH_3CO)_2O$, DMF, r.t.71%; (iii) TIPSE, CuI, $[Pd^0[P(Ph_3)]_4]$, Et_3N, DMF, r.t.60%; (iv) DIPEA, DCEP-Cl, CH_2Cl_2, r.t. 47%; DMAP = 4-Dimethylaminopyridine, and DMT-Cl = 4,4-Dimethoxytriphenylmethyl chloride.

Synthesis of compound 2. A solution of compound **1** (5.00 g, 142 mmol) in anhydrous pyridine (80 ml) was bubbled with a stream of nitrogen for 30 min. To this solution at 0 °C under N_2, Et_3N (3.94ml, 284 mmol), DMAP (0.425 g, 35.4 mmol) and 4,4'-dimethoxytrityl chloride (9.6 g, 284 mmol) were added and stirred for first 15 min at 0 °C and then at room temperature for 4 hours. The reaction mixture was quenched with MeOH (30 ml) and the solvent was evaporated. The residue was extracted with EtOAc, washed with aqueous $NaHCO_3$ solution for several times. After filtered and vacuum dried, the compound **2** is isolated as white solid (7.02g, 76%). ^{1}H-NMR (DMSO-d_6, δ=2.50, 300 MHz): δ=2.05-2.12 (m, 2H, 2′), 3.14-3.20 (m, 2H, 5′), 3.74 (s, 6H, OMe), 3.89-3.91 (m, 1H, 4′), 4.18-4.20 (m, 1H, 3′) ,5.27 (d, *J*=4.29Hz, 1H, 3′-OH), 6.10 (t, *J*=6.66Hz, 1H, H-1′), 6.64 (bs, 1H, -NH), 6.91 (d, 4H, aromatic), 7.20-7.41 (m, 9H, aromatic), 7.87 (bs, 1H, NH), 7.97 (s, 1H, H-6). MS (ESI-MALDI): calculated for $C_{30}H_{30}IN_4O$ $[M + Na^+]$ 678.12, found 678.10737.

Synthesis of compound 3. A solution of compound **2** (4.00 g, 1.37 mmol) in *N,N*-dimethylformamide (30 ml) was added acetic anhydride (1.45 ml, 15.40 mmol), and the reaction mixture was stirred at room temperature for 24 h. The mixture was added saturated NaCl solution (120 ml) and precipitated yellow crude solid. The filtered solid was dissolved in ethyl acetate (30 ml) and extracted with saturated NaCl solution (10 ml); the combined organic layers were dried over Na_2SO_4 and evaporated. The residue was purified by FC (silica gel, CH_2Cl_2/acetone, 80/20). After evaporation of the solvent

from the main zone, compound **2** was isolated as a colorless foam (3.02g, 71%). ^{1}H-NMR (DMSO-d_6, δ=2.50, 300 MHz): δ=2.13-2.20 (m, 1H, H_α-2′), 2.24 (s, 3H, CH_3), 2.34-2.38 (m, 1H, H_β-2′), 3.18-3.22 (m, 2H, H-5′), 3.74 (s, 6H, 2×OCH_3), 3.96-4.01 (m, 1H, H-4′), 4.16-4.23 (m, 1H, H-3′), 5.32 (d, *J*=3.27Hz, 1H, 3′-OH), 6.05 (t, *J*=4.86Hz, 1H, H-1′), 6.88-6.91 (m, 4H, Ar-H), 7.21-7.40 (m, 9H, Ar-H), 8.30 (s, 1H, H-C6), 9.44 (s, 1H, NH). MS (ESI): calculated for $C_{32}H_{32}IN_3O_7$ [M + Na^+] 720.13, found 720.9.

Synthesis of compound 4. A solution of compound **3** (1.50g, 2.69 mmol), [$Pd(pph_3)_4$] (0.374 g, 0.319 mmol), CuI (83 mg, 0.42 mmol) in anhydrous DMF (30 ml) was bubbled with a stream of nitrogen for 30 min. After being stirred for 5 min, Et_3N (0.80 ml, 5.75 mmol) and Triisopropylsilylacetylene (1.205 ml, 8.44 mmol) was added to the solution. The resulting mixture was stirred at 25 °C for 24 h under nitrogen. The mixture was diluted with saturated NaCl solution (120 ml) and extracted with ethyl acetate; the combined organic layers were dried over Na_2SO_4 and evaporated. The residue was purified by FC (silica gel, CH_2Cl_2/acetone, 140/10). After evaporation of the solvent from the main zone, compound **4** was isolated as a slightly yellow amorphous solid (0.97g, 60%). ^{1}H-NMR (DMSO-d_6, δ=2.50, 300 MHz): δ=0.91-0.97 (m, 21H, 6×CH_3, 3×CH), 2.11-2.18 (m, 1H, H_α-2′), 2.31-2.39 (m, 4H, CH3, H_β-2′), 3.07-3.12 (m, 2H, H-5′), 3.72 (s, 6H, 2×OCH_3), 4.01-4.04 (m, 1H, H-4′), 4.10-4.18 (m, 1H, H-3′), 5.34 (d, *J*=4.29Hz, 1H, 3′-OH), 6.03 (t, *J*=6.39Hz, 1H, H-1′), 6.85-6.88 (m, 4H, Ar-H), 7.18-7.40 (m, 9H, Ar-H), 8.22 (s, 1H, H-6), 9.13 (s, 1H, NH). MS (ESI-MALDI): calculated for $C_{43}H_{53}N_3O_7Si$ [M + H^+] 752.37, found 752.37236.

Synthesis of compound 5. A stirred solution of 4 (0.20 g, 0.26 mmol) in anhydrous CH_2Cl_2 (10 ml) was treated with (*i*-Pr)$_2$EtN (100 μL, 0.59 mmol) and 2-cyanoethyl-*N,N*-dissopropylphosphoramido chloridite (100 μL, 0.45 mmol). The resulting mixture was stirred for 1 h under nitrogen and ice bath. Then the solution was diluted with CH_2Cl_2 (10 ml) and extracted with 5% aqueous $NaHCO_3$ solution (10 ml). The organic layer was dried over Na_2SO_4 and concentrated. The residue was purified by FC (silica gel, refrigerate for several hours, PE/$CHCl_3$, 10/10). After evaporation of the solvent from the main zone, compound **5** was isolated as a colorless foam (0.12g, 47%). ^{31}P-NMR ($CDCl_3$, 400 MHz): δ=148.17, 148.91. MS (ESI-MALDI): calculated for $C_{52}H_{70}N_5O_8PSi$ [M + Na^+] 974.47, found 974.48201.

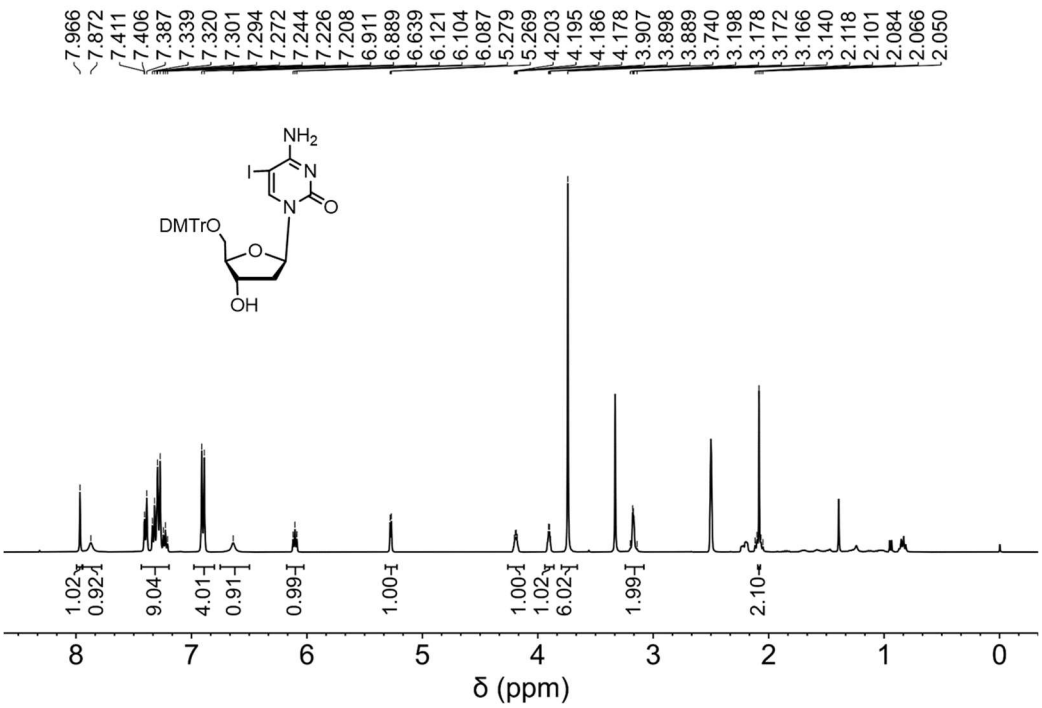


Figure S1. [1]H NMR spectrum of compound **2**.

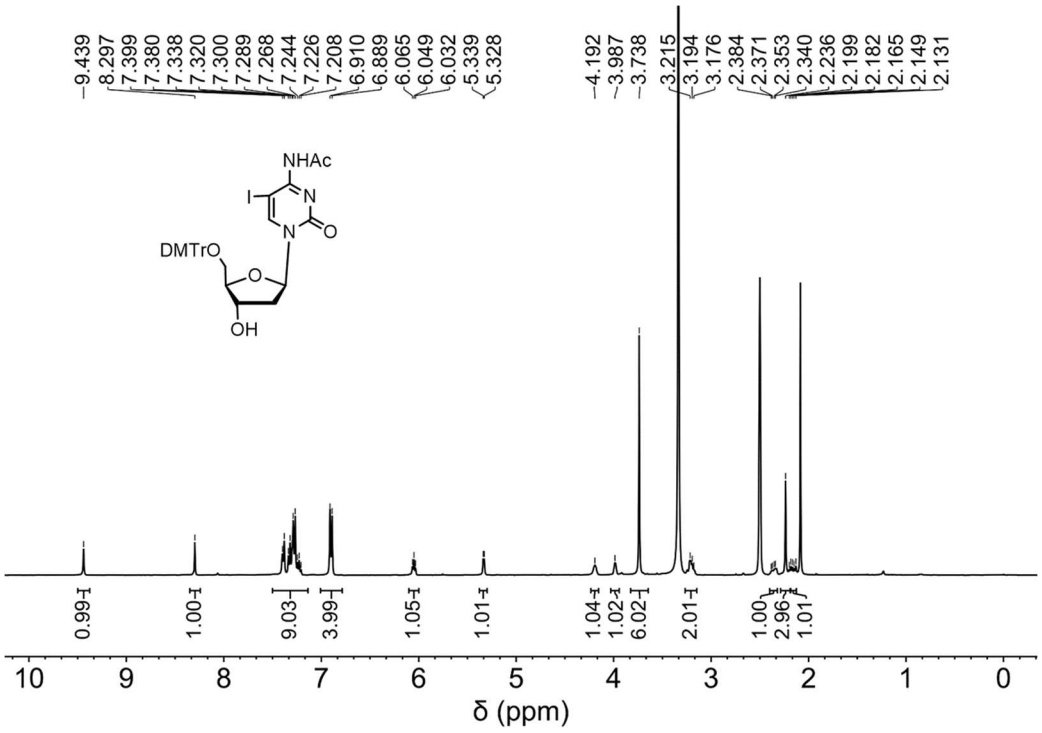


Figure S2. [1]H NMR spectrum of compound **3**.

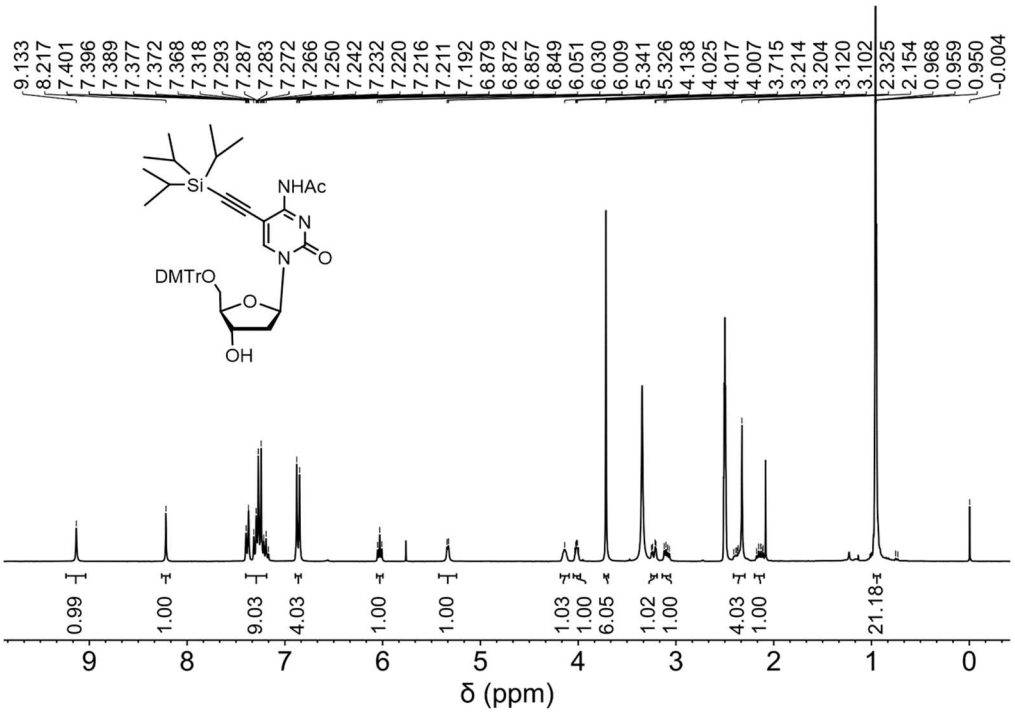


Figure S3. [1]H NMR spectrum of compound **4**.

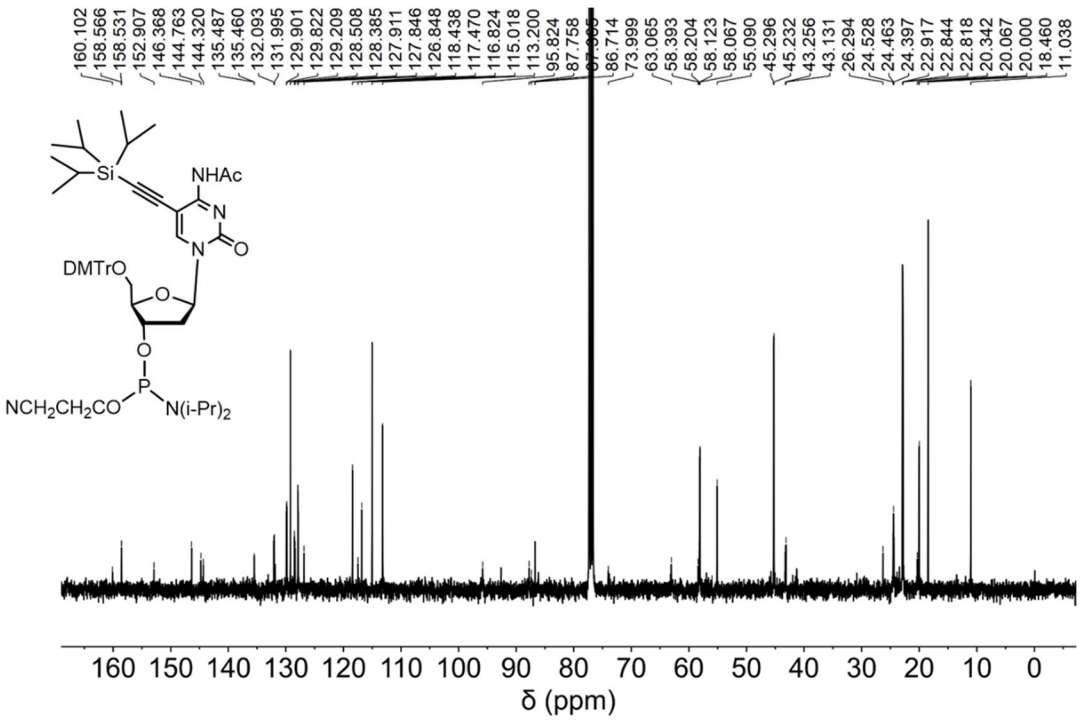


Figure S4. [13]H NMR spectrum of compound **5**.

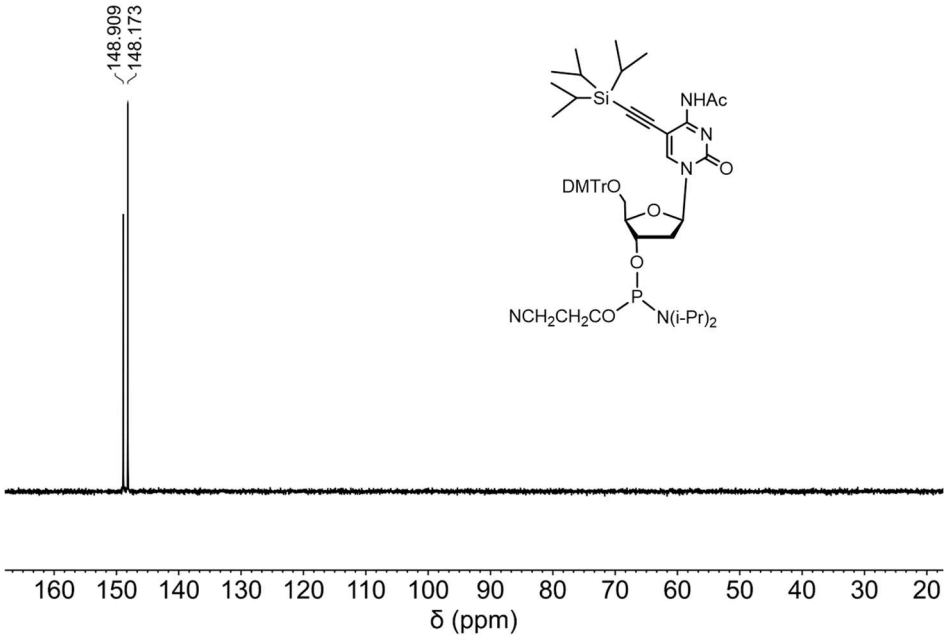


Figure S5. ^{13}H NMR spectrum of compound **5**.

B. MS characterization of the C≡C-labeled DNA strand

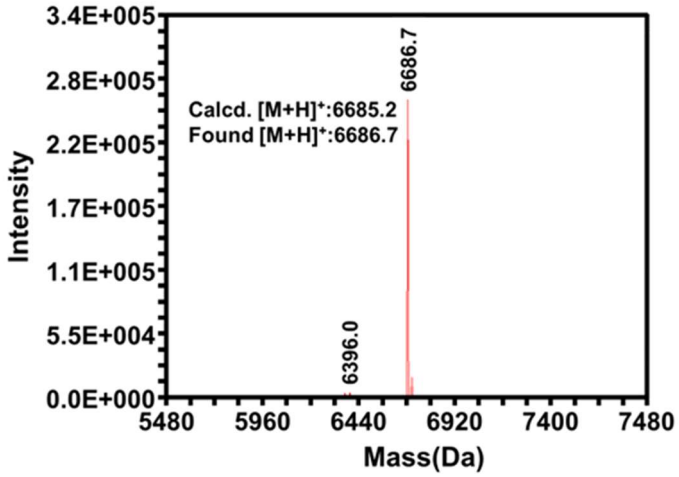


Figure S6. MS data for the characterization of the C≡C-labeled DNA strand.

C. CD spectra of the unlabeled and C≡C-labeled DNA strands

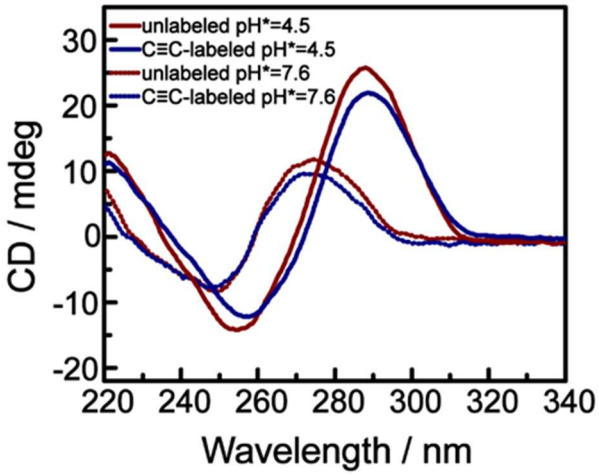


Figure S7. Far-UV CD spectra of the unlabeled DNA strand and the C≡C-labeled DNA strand in pH* 4.5 and 7.6 DPBS buffer solution.

Circular dichroism (CD) spectra were collected on a Jasco J-1700 CD spectrometer with wavelength ranging from 220 to 350 nm by using a quartz cell with a path length of 1 mm. The unlabeled and C≡C-labeled strands were dissolved in pH*=4.5, 7.6 DPBS buffer solution with the concentration of 50μm to be prepared for CD measurement. Each CD spectrum was accumulated over three runs at a scanning speed of 100 nm/min with a band width of 1.0 nm. As illustrated in Figure S1, at pH* 4.5, the CD spectra showed a strong positive peak near 288 nm and a negative peak near 258 nm with a crossover at around 272 nm, indicating that the i-motif structure was formed for both strands.[1,2] In contrast, at pH* 7.6 the characteristic i-motif peak at 288 nm and 258 nm disappeared, but appeared at 273 nm and 248 nm instead, indicating the partially folded or/and unfolded structures were formed for both strands under such condition.

D. FTIR spectra of nucleosides with band assignment

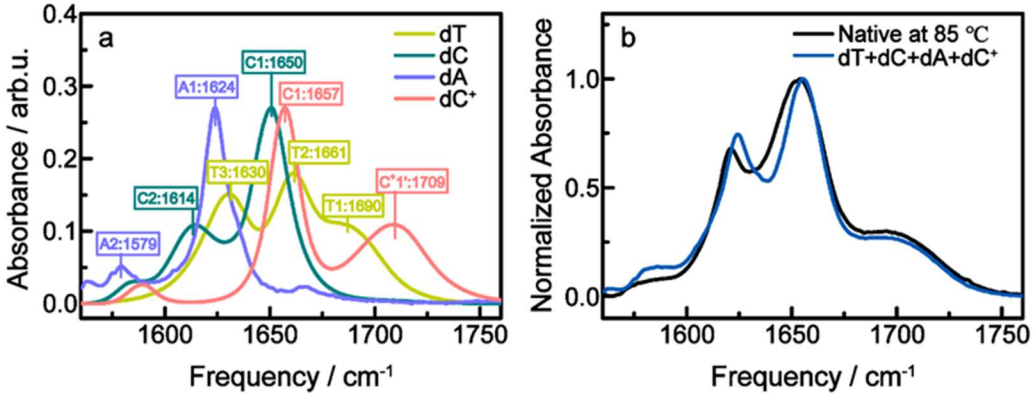


Figure S8. (a) FTIR spectra of deoxythymidine (dT), deoxycytidine (dC), deoxyadenosine (dA) and protonated deoxycytidine (dC^+) in pH* 4.5 DPBS buffer

solution in the 6-μm region. (b) Absorbance normalized FTIR spectra of the unlabeled strand in the pH* 4.5 DPBS buffer solution at 85 ℃ *vs* a superimposed spectrum of monomer (shown in (a)) by the ratio of 4:6:6:6. In (a), each spectrum is absorbance normalized then is scaled by the ratio of 4:6:6:6.

Table S1 Peak assignment of the i-motif DNA in the 6-μm region. "ν" represents stretching vibrational mode and "δ" represents bending vibrational mode.

Frequency / cm^{-1}	Peak	Peak assignment
1575	A2	ν (N^1=C^6), ν (N^3-C^4), δ (C^2-H), δ (C^8-H)
1602	C2	ν (N^3=C^4-C^5=C^6), δ (C^5-H), δ (C^6-H)
1626	A1	ν (C^4=C^5, C^5-C^6 out-of-phase), δ (C^2-H)
1638	T3	ν (C^5=C^6), δ (C^6-H)
1655	C1′	unfold state ν (C^2=O), ν (N^1=C^6), δ (C^6-H)
1665	C1	fold state ν (C^2=O), ν (N^1=C^6), δ (C^6-H)
1667	T2	ν (C^4=O), δ (N^3-D)
1689	T1	ν (C^2=O), δ (N^3-D)
1697	C^+1	fold state and protonic cytosine ν (N^1=C^6), ν (C^2=O)
1710-1720	C^+1′	unfold state and protonic cytosine ν (N^1=C^6), ν (C^2=O)

E. Temperature-dependent FTIR difference spectra

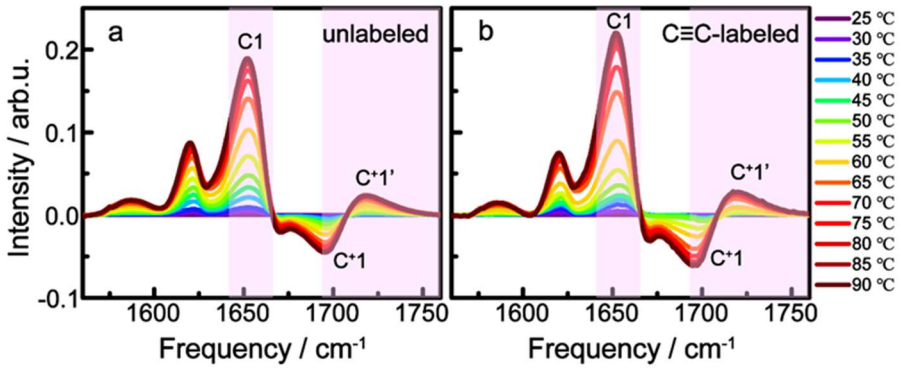


Figure S9. FTIR difference spectra of the unlabeled h-telo i-motif DNA strand (a) and C≡C-labeled strand (b) in pH* 4.5 DPBS buffer solution in the 6-μm region. Each spectrum is subtracted from the 25 ℃ spectrum. Shading indicates the IR peaks coming from C1, C^+1 and C^+1'.

F. Plot of the C≡C band frequency position *vs* intensity

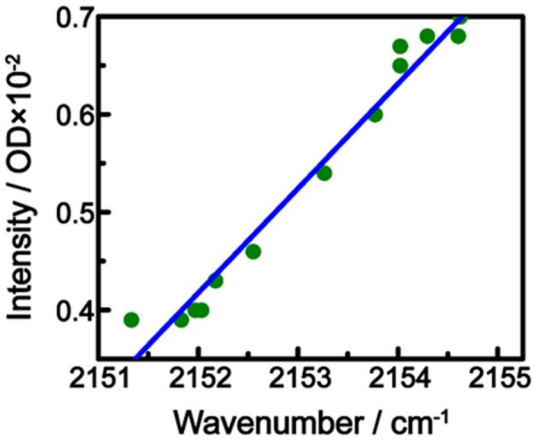


Figure S10. A plot of the C≡C band intensity *vs* its frequency during the unfolding process for the C≡C-labeled strand.

G. Temperature-dependent FTIR spectra of ^{TMS}EdC and ^{TMS}EdC^{+}

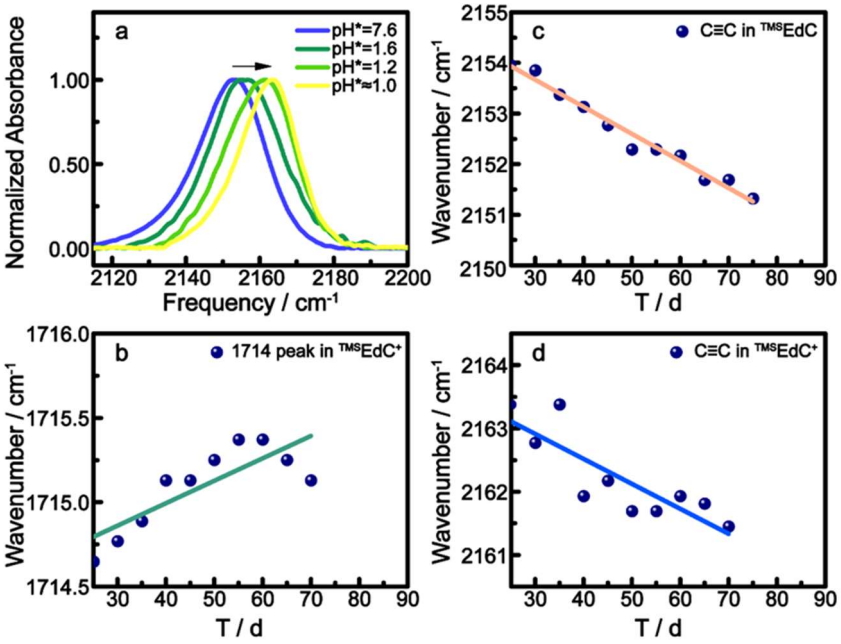


Figure S11. (a) FTIR spectra of ^{TMS}EdC measured under different pH conditions (pH* 7.6, 1.6, 1.2 and ~1.0 DPBS buffer solution) in the C≡C stretching region. (b) Temperature dependence of 1714 cm^{-1} peak of molecule ^{TMS}EdC dissolved in pH* ~1.0 DPBS buffer solution (the protonated form, ^{TMS}EdC^{+}). (c) Temperature dependence of the C≡C stretching frequency of ^{TMS}EdC dissolved in pH* 7.6 DPBS buffer solution (the non-protonated form, ^{TMS}EdC). (d) Temperature dependence of the C≡C stretching frequency of ^{TMS}EdC dissolved in pH* ~1.0 DPBS buffer solution (the protonated form ^{TMS}EdC^{+}).

As illustrated in Figure S11a, from the non-protonated form of ^{TMS}EdC (dissolved in pH* 7.6 DPBS) to the protonated form of ^{TMS}EdC (^{TMS}EdC$^+$, dissolved in pH* ~1.0 DPBS), the C≡C stretching band position blue shifts significantly, from 2153 to 2163 cm^{-1}. With temperature increasing, the C≡C band position of ^{TMS}EdC red shifts linearly (see Figure S11c) and that of the C≡C band of ^{TMS}EdC$^+$ fluctuates slightly (see Figure S11d). An absorption peak at 1714 cm^{-1} for ^{TMS}EdC$^+$ also fluctuates slightly (see Figure S11b).

H. Fitting results of van't Hoff function

Table S2. Fitting results of thermodynamic parameters shown in Figure 4.

sample	T_m (°C) (averaged)	ΔG_m (kcal/mol)	ΔH_m^{vH} (kcal/mol)	ΔS_m (kcal/mol•K)
Unlabeled_C1 peak [a]	62.8	-0.63	29.80	0.090
C≡C-labeled_C1 peak [a]	63.1	-0.37	28.75	0.087
C≡C-labeled_C≡C probe [a]	61.4	-0.12	18.30	0.055
$(CCCTAA)_4$ [b]	66.2	-0.46	36.54	0.109
i-motif different sequence [c]	10.2~54.9	-0.6~5.6	5*n [d]	—

[a] All parameters from Figure 4 measured in a pH* 4.5 DPBS buffer solution.
[b] Thermodynamic parameters derived from fitting the van't Hoff function.[3]
[c] Thermodynamic parameters from the i-motif DNA with different lengths and loop sequences.[4]
[d] Approximately 5.0 kcal/mol melting enthalpy per C-C$^+$ bond in the i-motif DNA.[4] Here n represents the number of the C-C$^+$ bonds in the i-motif DNA.

I. Location of the six adenines in the i-motif DNA

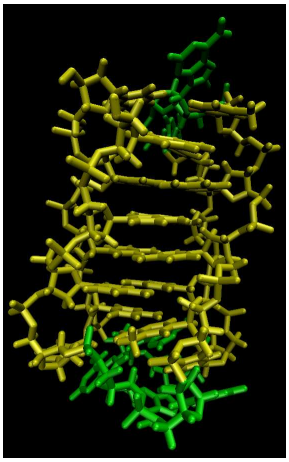


Figure S12. An illustration of the locations of six adenines (highlighted in green) in the i-motif DNA structure.

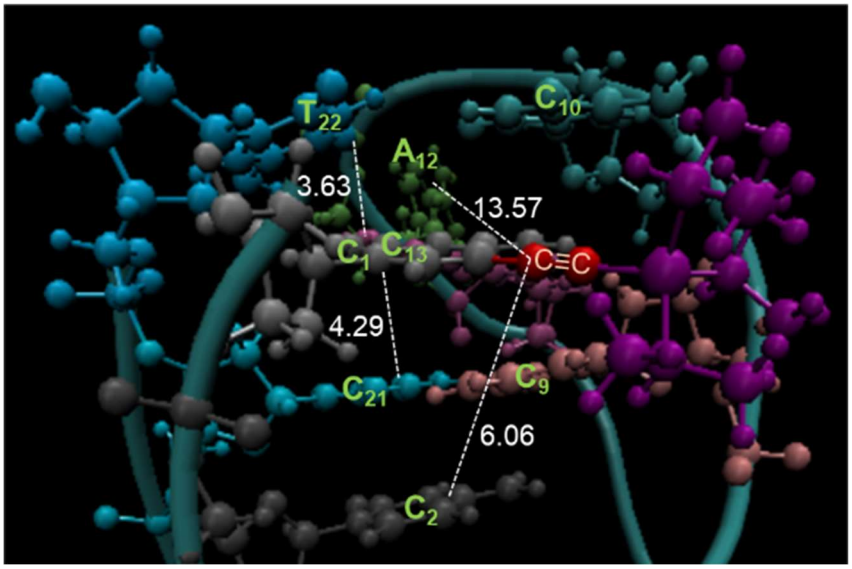


Figure S13. Chemical environment of the C≡C-labeled C_1-C_{13} base pair using the PDB structure of hTelo i-motif DNA.[5] The base-base distance is marked with white dashed line (in Å). Relevant bases are marked in green.

J. Comparison of the unlabeled and C≡C-labeled strands by pump-probe IR spectroscopy

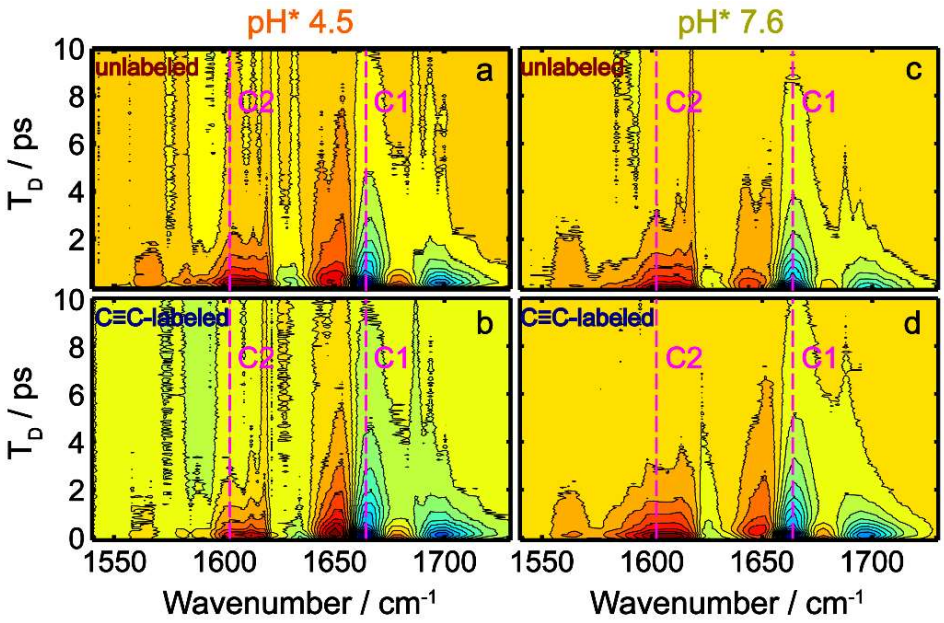


Figure S14. Magic-angle polarization pump-probe spectra of the unlabeled (top row) (a) and (c) and the C≡C-labeled i-motif (bottom row) (b) and (d) at pH* 4.5 (left row) and pH* 7.6 (right row).Dashed lines indicate the values of ω_τ for C1 ($0 \rightarrow 1$ transition) peaked at 1665 cm^{-1} and for C2 ($1 \rightarrow 2$ transition) peaked at 1603 cm^{-1}.

K. Nonlinear IR results of the unlabeled and C≡C-labeled strands at pH* 7.6

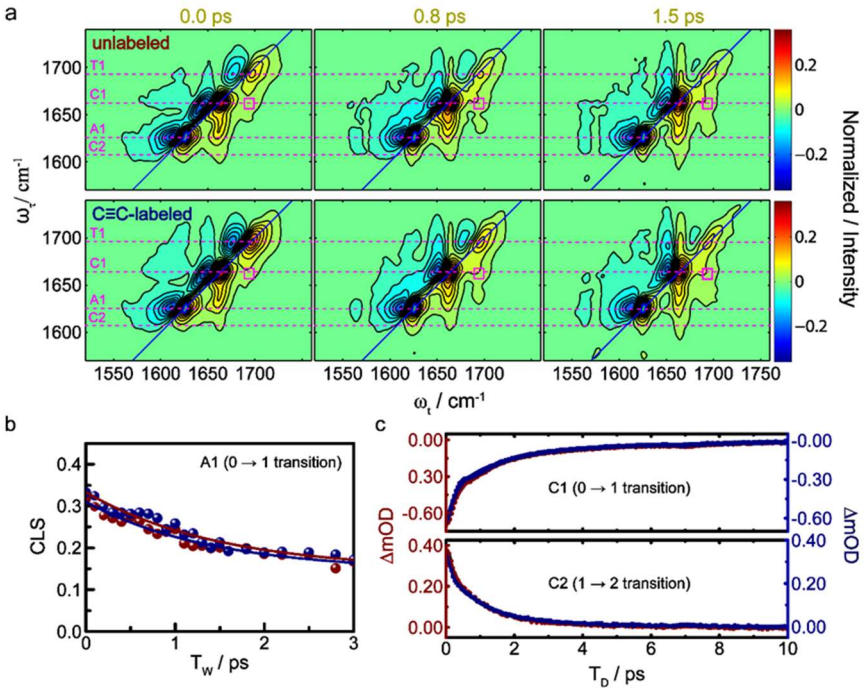


Figure S15. (a) A comparison of the unlabeled and C≡C-labeled strands monitored by time-dependent 2D IR spectroscopy in the 6-μm wavelength region. (b) CLS of the A1 band (0 → 1 transition). (c) Selected delay-time dependent vibration population dynamics. Wine color indicates the unlabeled strand and navy color indicates the C≡C-labeled strand. In (b), the solid curves represent the least-square fitting with single exponential function. In (c), the solid curves represent least-square fitting with di-exponential function.

All spectral characteristics from whether from the steady-state IR spectra or time-resolved IR spectra of the unlabeled and those of the C≡C-labeled strands at pH* 7.6 are nearly the same. It should be mentioned that both strands are more likely to be partially folded structures or mixed structures dissolved in the pH* 7.6 DPBS at the concentration of 10 mM. This does not agree with the CD spectroscopic results, because the concentration is much lower in the CD measurement.

L. Fitting results of the population relaxation dynamics

Table S3 Fitting parameters of the population relaxation dynamics of the unlabeled strand and the C≡C-labeled strand in pH* 4.5 and pH* 7.6 DPBS buffer solution.

Samples	C1		C2	
	0 → 1 peak		1 → 2 peak	
	T_1^a / ps	T_1^b / ps	T_2^a / ps	T_2^b / ps
Unlabeled in pH* 4.5	0.26 ± 0.02	1.95 ± 0.05	0.27 ± 0.03	1.37 ± 0.05
C≡C-labeled in pH* 4.5	0.25 ± 0.01	1.85 ± 0.02	0.12 ± 0.01	1.27 ± 0.01
Unlabeled in pH* 7.6	0.37 ± 0.02	2.05 ± 0.05	0.37 ± 0.03	1.41 ± 0.07
C≡C-labeled in pH* 7.6	0.28 ± 0.01	1.97 ± 0.01	0.18 ± 0.01	1.28 ± 0.01

References

1. C. Chen, M. Li, Y. Xing, Y. Li, C. C. Joedecke, J. Jin, Z. Yang and D. Liu, *Langmuir*, 2012, **28**, 17743-17748.
2. J. Choi, S. Kim, T. Tachikawa, M. Fujitsuka and T. Majima, *J. Am. Chem. Soc.*, 2011, **133**, 16146-16153.
3. M. McKim, A. Buxton, C. Johnson, A. Metz and R. D. Sheardy, *J. Phys. Chem. B*, 2016, **120**, 7652-7661.
4. S. M. Reilly, R. K. Morgan, T. A. Brooks and R. M. Wadkins, *Biochemistry*, 2015, **54**, 1364-1370.
5. A. T. Phan, M. Guéron and J.-L. Leroy, *J. Mol. Biol.*, 2000, **299**, 123-144.