Mechano-chemical selections of two competitive unfolding pathways of a single DNA i-motif

2014 Chinese Phys. B 23 068702

(http://iopscience.iop.org/1674-1056/23/6/068702)

Mechano-chemical selections of two competitive unfolding pathways of a single DNA i-motif*

Xu Yue(徐 悦)[a)b)c)e)†], Chen Hu(陈 虎)[b)d)†], Qu Yu-Jie(璩玉杰)[c)], Artem K. Efremov[b)], Li Ming(黎 明)[e)], Ouyang Zhong-Can(欧阳钟灿)[a)], Liu Dong-Sheng(刘冬生)[f)‡], and Yan Jie(严 洁)[b)c)§]

[a)] *State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Beijing* 100190, *China*
[b)] *Mechanobiology Institute, National University of Singapore, Singapore*, 117411
[c)] *Department of Physics, National University of Singapore, Singapore*, 117542
[d)] *Department of Physics, Xiamen University, Xiamen* 361005, *China*
[e)] *School of Physics, University of Chinese Academy of Sciences, Beijing* 100049, *China*
[f)] *Key Laboratory of Organic Optoelectronics & Molecular Engineering of the Ministry of Education, Department of Chemistry, Tsinghua University, Beijing* 100084, *China*

(Received 31 March 2014; revised manuscript received 4 April 2014; published online 20 April 2014)

The DNA i-motif is a quadruplex structure formed in tandem cytosine-rich sequences in slightly acidic conditions. Besides being considered as a building block of DNA nano-devices, it may also play potential roles in regulating chromosome stability and gene transcriptions. The stability of i-motif is crucial for these functions. In this work, we investigated the mechanical stability of a single i-motif formed in the human telomeric sequence 5′-$(CCCTAA)_3CCC$, which revealed a novel pH and loading rate-dependent bimodal unfolding force distribution. Although the cause of the bimodal unfolding force species is not clear, we proposed a phenomenological model involving a direct unfolding favored at lower loading rate or higher pH value, which is subject to competition with another unfolding pathway through a mechanically stable intermediate state whose nature is yet to be determined. Overall, the unique mechano–chemical responses of i-motif-provide a new perspective to its stability, which may be useful to guide designing new i-motif-based DNA mechanical nano-devices.

Keywords: single-molecule techniques, i-motif, folding/structure of biomolecules, mechanical properties/biomolecules

PACS: 87.80.Nj, 87.14.gk, 87.15.Cc, 87.15.La **DOI:** 10.1088/1674-1056/23/6/068702

1. Introduction

DNA often exists as the well-known right-handed antiparallel double helix structure, so-called B-DNA, by Watson–Crick base-pairing interaction. Other than this typical B-form, under certain solution conditions, DNA can also adopt into non-B structures, such as the cruciform, Z-DNA, A-DNA, S-DNA, triplex, tetraplex, etc.[1–4] Of particular interest, the tetraplex structures can be formed by G-rich and C-rich DNA sequences. Repeating G-rich sequences were found to be able to form the G-quadruplex by stacked G-quartets, which are composed of four coplanar guanines linked through Hoogsteen base pairing with a centrally located monovalent ion.[5,6] For C-rich repeating sequences, the tetraplex structures can also form through intercalated $C–C^+$ pairing interaction at slightly acidic condition,[7] often referred to as i-motif. Distinct from the G-quadruplex consisting of stacking planes each containing four guanine bases, i-motif structure is formed by stacking of the $C–C^+$ pairs. It consists of two intercalated parallel duplexes oriented in the opposite directions, each containing two parallel ssDNA strands associated by hemi-protonated $C–C^+$ pairs (Fig. 1(a) Inset).[7–9]

These tetraplex structures have recently drawn extensive attention, due to their biological or pharmaceutical relevance. It has been reported that G-rich and C-rich sequences with the potential to form G-quadruplex and i-motif are frequently found in the human genome, throughout the telomere that caps the end of the linear eukaryotic chromosomes,[10] and also in the regulatory regions of human diseases related specific genes.[1,11,12] The potential regulatory roles of G-quadruplex and i-motif are proposed through inhibition of transcription of the relevant genes.[13–16] A possible mechanism is that these structures form physical blocks to the transcription elongation of RNA polymerase that can exert large force up to $\sim$ 30 pN.[17,18] For i-motif, its tunable mechanical stability by pH also provides unique advantages for designing fast responsive DNA nano-devices[19–24] compared to traditional hybridization principled DNA machines.[25] Therefore, the mechanical stability of these structures may be crucial to their biological functions and applications.

*Project supported by Grants from the National Research Foundation through the Mechanobiology Institute Singapore and the Ministry of Education of Singapore (Grant No. MOE2012-T3-1-001) [to Yan J], the National Basic Research Program of China (Grant No. 2013CB932800), the Major Research Plan of the National Natural Science Foundation of China (Grant Nos. 91027046 and 91027045), and the Fundamental Research Funds for the Central Universities (Grant No. 2013121005).

†These authors contributed equally to this work.
‡Corresponding author. E-mail: liudongsheng@tsinghua.edu.cn
§Corresponding author. E-mail: phyyj@nus.edu.sg

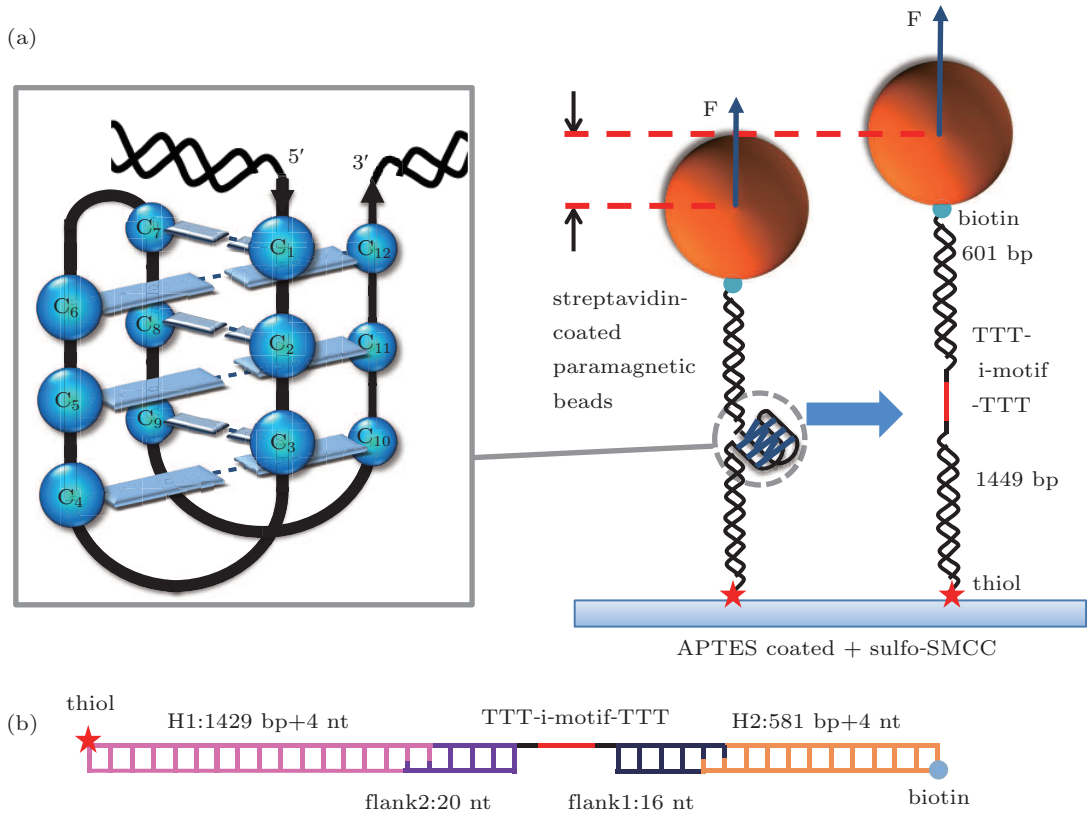

Fig. 1. (color online) (a) Schematic illustration of the experimental principle. The DNA construct is tethered between a paramagnetic bead via biotin/streptavidin interaction and the glass surface through covalent bond via sulfo-SMCC as a cross linker. DNA extension change due to unfolding of i-motif is recorded by the magnetic tweezers setup. The inset shows an enlarged schematic illustration of a folded i-motif structure linked with two DNA handles. It consists of two intercalated parallel C–C^+ paired duplexes oriented with opposite directions. The intercalation topology is 5′E for our telomeric sequence.[9] The arrow indicates the 5′ to 3′ direction. (b) Schematic illustration of the DNA construct, in which the 27-nt ssDNA containing a 21-nt i-motif sequence: 5′-TTT$(CCCTAA)_3$CCCTTT was flanked between two dsDNA handles (1449 bp and 601 bp).

Very recently, Dhakal and his coworkers reported the first single-molecule stretching study on the mechanical stability of a specific i-motif structure formed in the insulin-linked polymorphic region (ILPR) sequence: 5′-y$(TGTCCCCACACCCC)_2$.[26,27] They observed two unfolding species with distinct changes in contour length, which were proposed to be unfolding of the fully folded ILPR i-motif and a partially folded structure.[27] This finding is important, as it reveals complex structural organizations of the ILPR sequence. An important question raised from this finding is whether multiple i-motif conformers can be extended to other C-rich sequences with potential to form i-motif structures. The existence of partially folded telomeric i-motif detected by fluorescence resonance energy transfer (FRET) technique in bulk during the pH-induced conformational changes has also been reported.[28,29] The microscopic mechanism of formation and dissociation on the C–C^+ base pair level has been revealed by recent theoretical and simulation studies.[30] But the unfolding mechanism of intramolecular DNA i-motif is still unclear.

In order to address this question, here we investigate the unfolding of i-motif formed in the 21 nucleotides (nt) telomeric sequence: 5′-$(CCCTAA)_3$CCC. Our experiments were based on measuring the unfolding steps and force as a function of the loading rate using a high force magnetic tweezers setup recently developed in our laboratory,[31,32] which is capable of constant loading rate control in a range from 0.01 pN/s up to 10 pN/s and a force range from 0.1 pN up to more than 100 pN. From the loading rate-dependent studies of unfolding forces, the unfolding kinetics of i-motif structure was obtained in four different pH values 5.0, 5.5, 5.8, and 6.0.

2. Materials and methods

2.1. DNA constructs

The oligonucleotide containing telomeric i-motif sequence with flanking oligonucleotides incorporated to two dsDNA handles of 1429 bp (H1) and 581 bp (H2) resulting in a 27-nt ssDNA in the middle of the DNA construct (Fig. 1(b)).

The thiol-labeled dsDNA handle H1 and the biotin-labeled H2 were amplified using labeled primers by PCR on the template of λ-DNA (48502 bp, New England Biolabs, NEB) from 4479 bp to 5895 bp and from 5331 bp to 5899 bp, respectively. Both dsDNA handles are GC-rich

(GC content> 60%) to favor the B-to-S transition during DNA overstretching.[3,4,33–35] After PCR, two handles were digested with BstXI restriction enzyme (NEB) and purified with PCR purification kits (Qiagen), and then mixed with the middle oligonucleotide containing i-motif sequence (underlined) spanned between two flanking sequences: 5′-phosphate-CTTGTGCACAGACTCGTTTCCCTAACCCTAACCCTAACCCTTTGCAGCCAGGTCAGTAGCGAC, and two oligos that are complementary to the flanking sequences: 5′-phosphate-CTACTGACCTGGCTGC, 5′-CGAGTCTGTGCACAAGGTGC. Following the ligation reaction using T4 DNA ligase (NEB), the final DNA constructs (∼ 2000 bp) were extracted from agarose gel and purified with purification kits (Qiagen).

The GC-rich handles used in our experiments (total size is 2050 bp, ∼ 676 nm) also served as specificity control to ensure a single tether is attached between the bead and the coverslip surface. A single DNA tether was identified by the characteristic B-to-S transition at ∼ 65 pN with the expected elongation by around 70% (∼ 474-nm extension jump at ∼ 65 pN).

2.2. Magnetic tweezers experiments

To apply force to i-motif, the DNA construct was tethered between a coverslip surface and a paramagnetic bead. A single DNA tether was formed by attaching the thiol end of a DNA to an APTES (Thermal scientific) coated coverslip via sulfo-SMCC (Thermal scientific) as a cross linker, and the biotin end to a streptavidin-coated paramagnetic bead with a diameter of 2.8 μm (Dynal beads M270, Invitrogen) (Fig. 1(a)).

Force was applied to the bead by a pair of permanent magnets placed above the sample, and its magnitude was controlled by adjusting the distance between the magnets and cover glass surface. The sampling rate was fixed at 100 Hz in our experiments. The magnetic tweezers setup is able to stretch short DNA or protein tethers with accurate force calibration over a wide force range of more than 100 pN, and is capable of conducting experiments at constant forces or under constant loading rates up to 10 pN/s.[31,32]

To eliminate non-specific interaction between the DNA and the surface, 1 mg/ml BSA in PBS buffer was used to passivate the surface. In neutral pH, these BSA blocked surfaces are negatively charged and do not nonspecifically interact with DNA tethers. However, at low pH value, these surfaces may become locally positively charged and sticky to the DNA tethers.[36] We introduced 1-μM poly-dA_{10} or poly-dT_{15} ssDNA in solution. These short ssDNA are adhesive to the surfaces at low pH and provide an extra blocking. Therefore, the non-specific interactions between DNA tethers and surfaces were eliminated at low pH value.

2.3. Data analysis to identify unfolding steps

Figure 2(a) shows a typical time trace of the DNA extension as the force was increasing at a loading rate of 0.1 pN/s at pH 5.5. A single stepwise unfolding signal marked in the rectangle box indicates the unfolding of i-motif. Detection of such unfolding signals was automated by detecting abrupt increases in the extension change in a 0.02-s sliding time window (Fig. 2(b)). The unfolding step size can be calculated by the subtraction of the linear fittings of the extension in 1-s time window before and after the unfolding (shown in Fig. 2(a) inset).

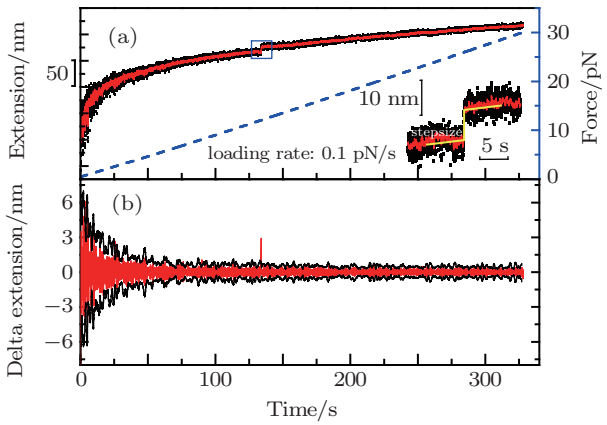

Fig. 2. (color online) (a) I-motif unfolding at a constant loading rate of 0.1 pN/s. Black dots show a typical time trace of the extension recorded with a sampling rate of 100 Hz. The red dots show a smoothed time trace by adjacent average method within 0.1-s time interval. The extension jump marked in a rectangle indicates an unfolding event of i-motif. The inset shows the zoom in of the unfolding event. Yellow lines are the linear fittings of the extension in 1-s window before and after the jump. The unfolding step size was calculated by the subtraction of the two fitting lines at the jump. The dashed blue line shows that force changes linearly with time (right axis). (b) The red curve shows extension change in a 0.02-s sliding time window across the whole time trace of the smoothed data (red) in the top panel. The local fluctuation of the extension change was estimated by calculation of the standard deviation in 1-s sliding time window. The black curve shows six times of the calculated local standard deviation. The single unfolding event is detected by the ratio of the grey (red) data to the black data, which is maximal at the time when unfolding occurred.

3. Results and discussion

3.1. Unfolding force distributions

At each pH and each loading rate, > 100 unfolding events obtained from > 10 independent DNA tethers were used to generate the unfolding force histograms. Figure 3(a) shows the results obtained at pH 5.0. A single unfolding force peak, which could be well fitted by Gaussian distribution, was revealed at each loading rate. The force peak shifted linearly to larger values as loading rate increased (Fig. 4).

Figure 3(b) shows the results obtained at pH 5.5. Surprisingly, a bimodal force distribution profile was revealed. Each unfolding force species could be fitted with Gaussian distribution. Both force peaks shifted to higher values linearly as loading rate increased (Fig. 4). However, the relative probability of the two unfolding force species changed with loading rate. The higher the loading rate, the smaller the probability of the lower unfolding force species, which completely vanished at the loading rate of 10 pN/s.

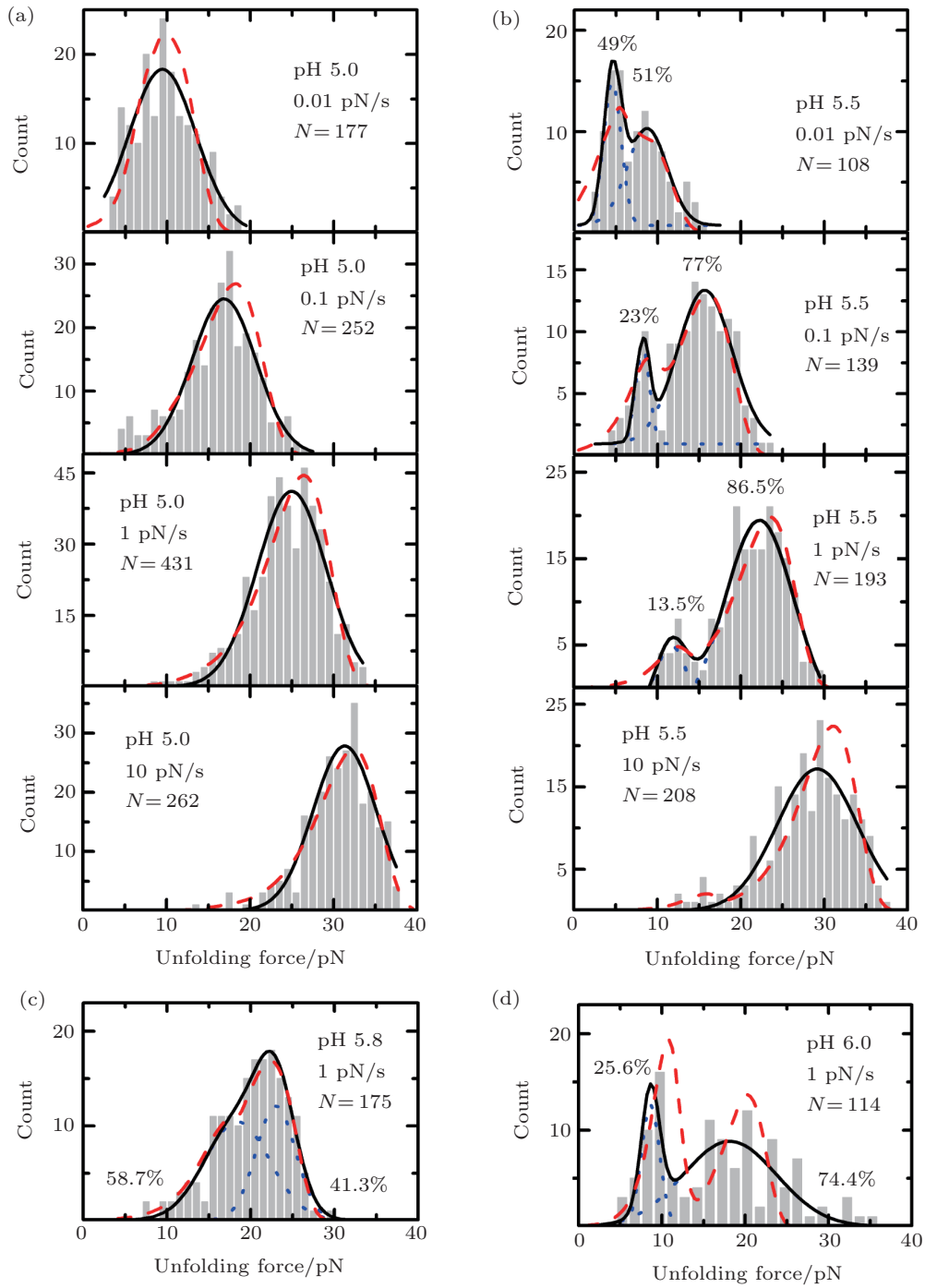

Fig. 3. (color online) Unfolding force distributions at four different pH values, in 100-mM KCl, 10-mM sodium phosphate buffer. (a) Results obtained at loading rates of 0.01 pN/s, 0.1 pN/s, 1 pN/s, and 10 pN/s, respectively, at pH 5.0. (b) Results obtained at pH 5.5 at the same set of loading rates as applied in panel (a). (c) Results obtained at pH 5.8 at loading rate of 1 pN/s. (d) Results obtained at pH 6.0 at loading rate of 1 pN/s. Loading-rate-dependent bimodal unfolding force distributions were observed at pH 5.5, 5.8, 6.0, but not at pH 5.0. At the fix loading rate of 1 pN/s, the unfolding force distributions also revealed pH dependence. The black curves in panel (a) are single Gaussian fittings to the unfolding force distributions; the blue dotted curves in panels (b), (c), and (d) are individual Gaussian fittings to each unfolding force species, while the black curves are the sum of the two. The red dashed curves are the best fittings of our kinetics model with the best fitting kinetics parameters in Table 1.

In order to investigate the pH dependence of i-motif mechanical stability under higher pH conditions, we explore the i-motif stability to a wider pH range where i-motif is supposed to be less stable. At pH 5.8 (Fig. 3(c)) and pH 6.0 (Fig. 3(d)), with the fixed loading rate 1 pN/s, the binomial unfolding force distributions were revealed similarly as at pH 5.5, except more unfolding events occurred at the lower unfolding force species than at pH 5.5.

These results suggest that the two unfolding force species are competitive in nature and can be selected by changes in pH

and loading rate. The lower unfolding force species favored lower loading rate or higher pH value.

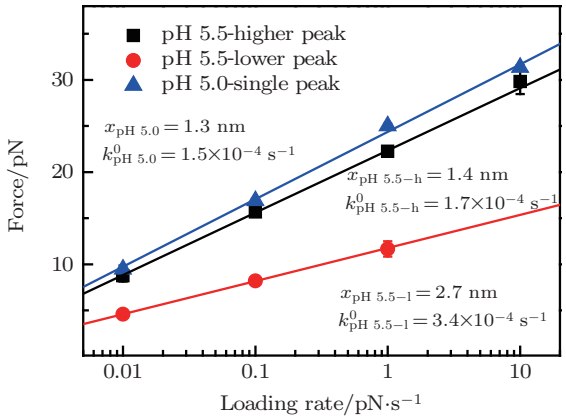

Fig. 4. (color online) The dependence of the unfolding force on the loading rate in the logarithm scale. Symbols indicate data obtained from experiments, and lines are linear fittings for the experimental data, respectively. The black squares are for the higher force peaks at pH 5.5, while the red circles for the lower force peaks at pH 5.5; the blue uptriangles are for the single force peaks at pH 5.0. The kinetics parameters shown inside the figure were determined by linear fittings of the respective unfolding force peaks using Eq. (4).

3.2. Unfolding step size distributions

In order to estimate the number of ssDNA nucleotides released during unfolding of the i-motif, a theoretical model of the ssDNA force-extension curve is needed. However, due to complex stacking and hydrogen bond formation between the bases, the force response of ssDNA cannot be ideally described by the commonly used worm-like-chain (WLC) or the freely-joint-chain (FJC) polymer models. To date, the ssDNA force-extension curves are much better described by a phenomenological force-extension relation proposed by Cocco *et al.*[37]

$$x_{\mathrm{ss}} = h\left[\frac{a_1 \ln(f/f_1)}{1+a_3\,\mathrm{e}^{-f/f_2}} - a_2 - f/f_3\right], \tag{1}$$

whose fitting parameters were determined to be $h = 0.34$ nm, $a_1 = 0.21$, $a_2 = 0.34$, $f_1 = 0.0037$ pN, $f_2 = 2.9$ pN, and $f_3 = 8000$ pN. The salt dependence is described by $a_3 = 2.1\ln([\mathrm{Na}^+]/0.0025)/\ln(0.15/0.0025) - 0.1$, where $[\mathrm{Na}^+]$ is the salt concentration expressed in the unit of Molar.[37] With these parameters, the model is able to fit the ssDNA force-extension curves over a wide force range (1 pN–100 pN) and a wide salt concentration (> 50 mM NaCl).[37] The experimental data (results at each pH values grouped by different colors) and the fitting are shown in Fig. 5(a). Extension released during unfolding was fitted to be 17.8 ± 0.1 nt by *step size* = nx_{ss}, where n is the number of nucleotides released during i-motif unfolding. As a contrast, $n = 19$ nt and 16 nt using the same model are also shown as dashed curves in Fig. 5(a).

We also calculated the number of nucleotides released from all unfolding events at four pH values. The statistic results shown in Fig. 5(b) gave us 17.1 ± 2.8 nt by Gaussian fitting. These two values are reasonably close to the number of nucleotides ($n = 21$) in our i-motif sequence, on account of the size of the folded i-motif.

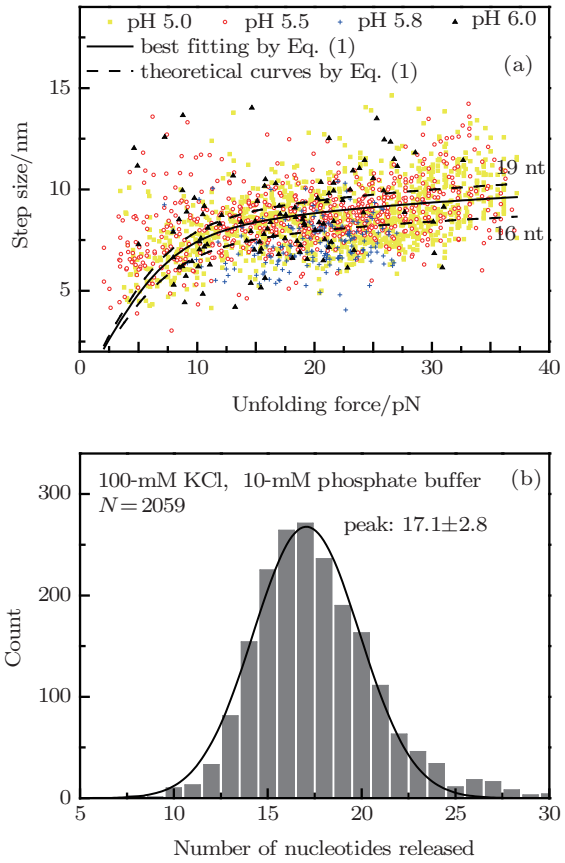

Fig. 5. (color online) (a) Data show the dependence of the extension change versus unfolding force, which can be fitted by the phenomenological model of ssDNA force-extension curve proposed by Cocco *et al.*[37] (black curve). Data obtained at different pH values are grouped using different colors: yellow squares for pH 5.0, red open circles for pH 5.5, blue crosses for pH 5.8, and black uptriangles for pH 6.0. We found 17.8 ± 0.1 nt were released during unfolding by fitting the theoretical curve to the experimental data. Theoretical results of 19 nt and 16 nt using the same model are also shown as dashed curves. (b) The histogram of the number of released nucleotides during i-motif unfolding estimated by the theoretical model Eq. (1). The histogram is based on all unfolding data obtained at all four pH values we tested. Fitting by Gaussian distribution determined that 17.1 ± 2.8 nt of ssDNA were released during i-motif unfolding.

3.3. Kinetics model

The two competitive unfolding force species revealed in Fig. 3 indicate very complex kinetics of i-motif unfolding. The so-called Bell model, a simple two-state transition model with force-dependent transition rate of

$$k(f) = k^0 \exp(xf/k_{\mathrm{B}}T), \tag{2}$$

predicts an unfolding force distribution $P(f)$:[38]

$$P(f) = \frac{k^0}{r}\exp\left\{\frac{xf}{k_{\mathrm{B}}T} + \frac{k_{\mathrm{B}}Tk^0}{xr}\left[1-\exp\left(\frac{xf}{k_{\mathrm{B}}T}\right)\right]\right\}, \tag{3}$$

where r is the loading rate, k_{B} is Boltzmann constant, T is absolute temperature, x is the transition distance to the unfolding barrier along the force direction, and k^0 is the unfolding rate at zero force, which can be conveniently approximated by Gaussian distribution.[38] This model is inadequate to explain

our results, as the distribution only predicts a single force peak located at:

$$f = \frac{k_B T}{x} \ln\left(\frac{xr}{k_B T k^0}\right). \quad (4)$$

The cause of the two unfolding force species is unclear. Some insights may be obtained from a previous study by Dhakal *et al.*[26] In that work, the mechanical stability of an i-motif structure formed by the ILPR sequence: 5′-(TGTCCCCACACCCC)$_2$ was studied at a fixed loading rate of ~ 5.5 pN/s.[26] Two unfolding species were also observed. However, unlike our observations for the telomeric i-motif sequence, two distinct unfolding force species with similar unfolding step sizes, they reported two distinct unfolding step species with similar unfolding forces. Therefore, the causes of the two unfolding species observed in the ILPR i-motif sequence and the telomeric i-motif sequence may be different.

We have considered two possibilities to understand the two unfolding force species observed in our experiments. One is that there may exist two co-existing folded i-motif species whose populations depend on pH values. Assuming different mechanical stability between them, two pH-dependent unfolding force species can be expected in mechanical stretching experiments. However, this picture has difficulty explaining the loading rate dependence of the probability of the two species, as the distribution of pre-existing folded i-motif species should not be affected by the loading rate of the experiments. Thus, we propose another possibility that likely two competitive unfolding pathways from the same folded structure are involved.

In the hypothetical competitive two-pathway unfolding model (Fig. 6), we reason that one pathway is a direct unfolding from the folded structure C to the unfolded ssDNA state U. The other includes two kinetics steps: the first step is transition from the folded structure C to an intermediate state I with an undetectable extension change, and the second step is transition from I to U with a step size similar to the direct unfolding pathway from C to U.

Further, all transition rates were assumed to behave according to Bell formula, Eq. (2). For experiments with a constant loading rate r, one can solve the following set of master equations and find the probabilities $p_i(t)$ for each state i at time t:

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \sum_{j\neq i} p_j k_{ji}(f(t)) - p_i \sum_{j\neq i} k_{ij}(f(t)), \quad (5)$$

where k_{ji} are the transition rates from state j to state i. Then the rupture force probability density function, $\rho(f)$, can be simply calculated as:

$$\rho(f(t)) = A \sum_{j<i} p_j k_{ji}(f(t)), \quad (6)$$

where A is a normalizing constant. Thus, by using Eqs. (2), (5), (6) and given the values of kinetics parameters k^0_{ji}, x_{ji}, it is possible to predict the theoretical force distribution histogram. Inversely, by knowing the unfolding force histogram the values of parameters can be determined. From a set of initial parameters, we applied simplex search method[39] by varying the parameters to minimize the objective function χ^2 in Eq. (7), the difference between the observed frequency O_i and theoretical prediction E_i

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}. \quad (7)$$

In this way, we can search for the best fitting parameters.

In our model, six kinetics parameters are required: $\{k^0_{cu}, x_{cu}, k^0_{ci}, x_{ci}, k^0_{iu}, x_{iu}\}$ (Fig. 4). The magnitudes of initial parameters can be estimated referring to the linear fittings using Eq. (4) as shown in Fig. 6, which are 10^{-4} s^{-1} for k^0, and 1 nm for x. The best fitting parameters corresponding to unfolding force histograms at each pH value can be generated by our search algorithm. All parameters were shown in Table 1, and the unfolding rates at zero force of each transition were plotted with the pH value in Fig. 7. The best fitting curves are shown as red dashed curves in Fig. 3.

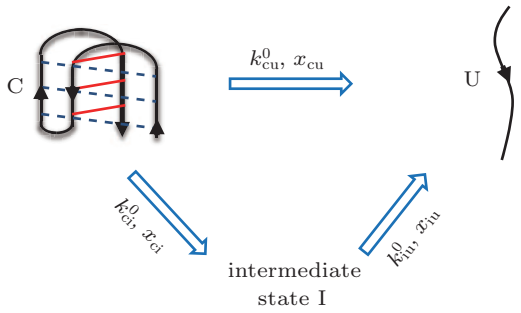

Fig. 6. (color online) Schematics of two kinetically competitive i-motif unfolding pathways: (i) direct unfolding from the folded state C to the single-stranded state U, and (ii) unfolding through a hypothesized intermediate state I, which involves a C→ I transition with a small extension change undetectable in our experiments, followed by a second I→ U transition with a larger step size detected in experiments. The kinetics parameters of each transition pathway can be optimized by our fitting algorithm.

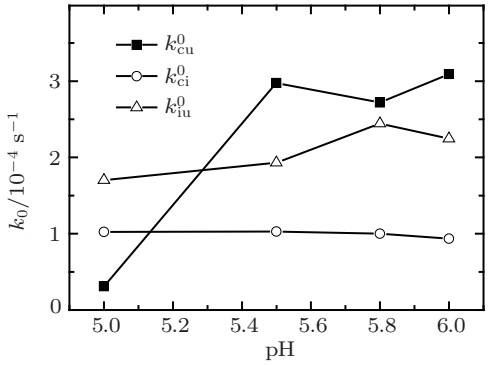

Fig. 7. The best fitted unfolding rates at zero force based on the kinetics model shown in Fig. 6. Solid symbols for pathway C→ U, and open symbols for pathway C→ I→ U. There is a remarkable increase of k^0_{cu} from pH 5.0 to 5.5, implying the pH sensitive dependence of corresponding transition C → U. Using the best fitted kinetics parameters, our predictions agreed with experimental distributions of unfolding forces. All the fittings are shown as red dashed curves in Fig. 3.

Table 1. The best fitting kinetics parameters of our model.

pH	k^0_{cu}/s^{-1}	k^0_{ci}/s^{-1}	k^0_{iu}/s^{-1}	x_{cu}/nm	x_{ci}/nm	x_{iu}/nm
5.0	3.1×10^{-5}	1.0×10^{-4}	1.7×10^{-4}	1.7	3.0	1.1
5.5	3.0×10^{-4}	1.0×10^{-4}	1.9×10^{-4}	1.7	2.8	1.3
5.8	2.7×10^{-4}	1.0×10^{-4}	2.4×10^{-4}	1.5	2.1	1.3
6.0	3.1×10^{-4}	9.3×10^{-5}	2.2×10^{-4}	2.1	2.5	1.3

So far, the correspondence between transition pathways and unfolding force species is unclear. We applied kinetic Monte Carlo simulations to reproduce the observed unfolding force distributions and to look into the competition of two transition pathways. The rate of each transition under force f is determined by Eq. (2). Force is increasing as $f = rt$, linearly with time under a given loading rate r. At each loading rate, we choose a time step of $\Delta t = 0.01$ s, the same as the time resolution of magnetic tweezers experiment, in which the probability of a particular transition i to happen is $\Delta P_i = k_i(f)\Delta t$. The simulation started from the initial conformation C in Fig. 6, which has two transition pathways C $\rightarrow$ U with $\Delta P_{cu} = k_{cu}(f)\Delta t$ and C $\rightarrow$ I with $\Delta P_{ci} = k_{ci}(f)\Delta t$. A random number R uniformly distributed between zero and one was generated and compared with the two transition probabilities. If $R < \Delta P_{cu}$, the C $\rightarrow$ U transition was selected; if $\Delta P_{cu} < R < \Delta P_{cu} + \Delta P_{ci}$, the C $\rightarrow$ I transition was selected; otherwise, the state C remained. From the state I, the transition I $\rightarrow$ U happened if $R < \Delta P_{iu}$. The simulation ended in state U. During the simulations, the unfolding forces were recorded as long as in which transition pathway the unfolding events happened.

We reproduced the experimental observations reasonably by the simulation and it also gave us the details of pathway selections. The observed lower unfolding force species related to the direct unfolding pathway C $\rightarrow$ U and the higher species related to pathway C $\rightarrow$ I $\rightarrow$ U. More specifically, the competitive selections between two pathways relied on the competition between k_{cu} and k_{ci}, while the I $\rightarrow$ U gave us the higher unfolding force species, which indicated the counterintuitive mechanical stability of intermediate state I. Further, we reasoned that transition C $\rightarrow$ I involves a small extension change undetectable in our experiments, and transition I $\rightarrow$ U with a large step size detected in experiments.

As to the conformational nature of the intermediate state I, a partially unfolded i-motif was considered as a candidate, whose conformation is similar to the partially folded ILPR i-motif suggested in Refs. [26] and [27]. However, since we did not observe different unfolding step sizes between the two unfolding species, we had to give up this possibility. We suspect it is a conformer produced by mechanical stretching, which does not exist in bulk experiments where force is absent. In view of the lack of a structural basis to infer the possible structure of such intermediate state, we have refrained from further speculations of its nature.

Among all three zero force transition rates in Fig. 7, one remarkable increase by one magnitude of k^0_{cu} from pH 5.0 to pH 5.5 was noticed, whereas other parameters do not show such pH sensitive behavior. The pH sensitivity of k^0_{cu} implies the pH dependence of C $\rightarrow$ U pathway in a certain pH range. We further assumed that the C $\rightarrow$ U pathway is pH-induced, while the other unfolding pathway is mainly caused by tension. Further, it allowed us to treat the C $\rightarrow$ U pathway as an analogy to the pH-induced i-motif unfolding in bulk experiments.

A recent study demonstrated a cooperative neutralization of two-protons as the only rate-limiting step in the pH-induced i-motif unfolding process,[40] where the two boundary C–C^+ pairs (C_1–C_7 and C_6–C_{12} pairs in Fig. 1(a) inset) are simultaneously disrupted, resulting in a complete unfolding of i-motif. Relating their results to our model, the transition rate k^0_{cu} depends on the proton concentration via $k^0_{cu} \propto [H^+]^{-n}$, where n is the cooperative factor. Here $n = 2$ accounts for cooperative release of two protons. We calculated n in the jump range of k^0_{cu} in Fig. 7, resulting in $n = 1.94$, which is in agreement with the bulk experiment. However, for the pH value above pH 5.5, k^0_{cu} is saturated. More factors may affect the i-motif unfolding, which are beyond our simple two-pathways model.

Our results can still provide some insights into the possible pH-dependent unfolding processes: There are two pathways leading to i-motif unfolding, one is pH-induced cooperative unfolding, the other can be force-induced non-cooperative unfolding, via the intermediate state I. At pH 5.5 or above, two pathways are comparably competitive and lead to the bimodal unfolding force distribution; however, the pathway C $\rightarrow$ I $\rightarrow$ U becomes predominant at pH 5.0 due to decreased rate k^0_{cu}.

4. Conclusion and application

By using magnetic tweezers to apply force to a single i-motif, the mechanical stability of a single i-motif formed in the human telomeric sequence 5′-$(CCCTAA)_3CCC$ was investigated. A novel pH and loading rate-dependent bimodal unfolding force distribution was found, which indicates there are two competing unfolding pathways: one direct cooperative unfolding pathway and one pathway with an intermediate state. Overall, our results revealed the mechanical stability complexity of the i-motif structure formed on the telomeric sequence. The findings also suggest a possible intermediate

mechanically stable i-motif structure, whose nature requires future studies.

The switchable folding and unfolding of i-motif by simply changing the solution pH over a narrow range has made it an attractive building block for designing nanometer size devices for various biosensing and molecular assembly applications, such as serving as a reporter of the intracellular pH[41] and promoting a pH-dependent self-assembly of gold nano particles.[42] In addition to these pH-sensitive applications, i-motif has a great potential to serve as a mechanical nano-device for various purposes. Figure 8 shows an example, which is an ongoing project in our lab, to use i-motif to actively produce a sharply bent dsDNA in order to address several open questions regarding the micromechanics of sharply bent dsDNA.[43–45] The design includes hybridizing a short oligonucleotide to the complementary region of a longer circular ssDNA at pH 7.5. The non-hybridized region contains an i-motif sequence, which is designed to fold upon switching to a low pH and hence bend the dsDNA region.

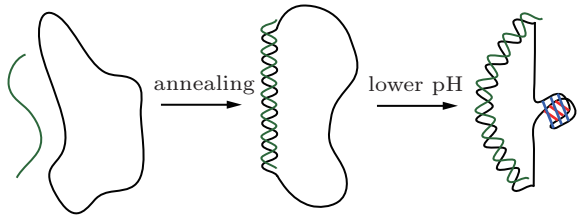

Fig. 8. (color online) Illustration of a dsDNA bent by an i-motif. An oligonucleotide annealed with a circular ssDNA to form a hybridizing dsDNA, which is bent by an i-motif in the ssDNA region when changing to lower pH values which favor the i-motif formation.

Acknowledgments

We thank Zhang Xing-Hua (Singapore-MIT Alliance for Research and Technology), Chen Chen, Xing Yong-Zheng, Yang Yang (Tsinghua University) for DNA constructs design, Chen Chun (Tsinghua University) for the discussion on bulk experiments of i-motif, Lim Ci-Ji, You Hui-Juan and Toh Kee-Chua (National University of Singapore) for proofreading the manuscript.

References

[1] Bacolla A and Wells R D 2004 *Biol. Chem.* **279** 47411
[2] Wells R D 2007 *Trends Biochem. Sci.* **32** 271
[3] Fu H X, Chen H, Zhang X H, Qu Y Y, Marko J F and Yan J 2011 *Nucleic Acids Res.* **39** 3473
[4] Zhang X H, Chen H, Fu H X, Doyle P S and Yan J 2012 *Proc. Natl. Acad. Sci.* **109** 8103
[5] Keniry M A 2001 *Biopolymers* **56** 123
[6] Burge S, Parkinson G N, Hazel P, Todd A K and Neidle S 2006 *Nucleic Acids Res.* **34** 5402
[7] Gehring K, Leroy J L and Gueron M 1993 *Nature* **363** 561
[8] Leroy J L, Gueron M, Mergny J L and Helene C 1994 *Nucleic Acids Res.* **22** 1600
[9] Phan A T, Gueron M and Leroy J L 2000 *J. Mol. Biol.* **299** 123
[10] Blackburn E H 1991 *Nature* **350** 569
[11] Xu Y and Sugiyama H 2006 *Nucleic Acids Res.* **34** 949
[12] Guo K X, Gokhale V, Hurley L H and Sun D 2008 *Nucleic Acids Res.* **36** 4598
[13] Siddiqui-Jain A, Grand C L, Bearss D J and Hurley L H 2002 *Proc. Natl. Acad. Sci.* **99** 11593
[14] Huppert J L and Balasubramanian S 2007 *Nucleic Acids Res.* **35** 406
[15] Kendrick S, Akiyama Y, Hecht S M and Hurley L H 2009 *J. Am. Chem. Soc.* **131** 17667
[16] Dai J X, Hatzakis E, Hurley L H and Yang D Z 2010 *PLoS One* **5** e11647
[17] Wang M D, Schnitzer M J, Yin H, Landick R, Gelles J and Block S M 1998 *Science* **282** 902
[18] Wang H Y, Elston T, Mogilner A and Oster G 1998 *Biophys. J.* **74** 1186
[19] Liu D S and Balasubramanian S 2003 *Angew. Chem. Int. Ed.* **42** 5734
[20] Shu W M, Liu D S, Watari M, Riener C K, Strunz T, Welland M E, Balasubramanian S and McKendry R A 2005 *J. Am. Chem. Soc.* **127** 17054
[21] Liu D S, Cheng E J and Yang Z Q 2011 *NPG Asia Mater.* **3** 109
[22] Liu D S, Bruckbauer A, Abell C, Balasubramanian S, Kang D J, Klenerman D and Zhou D J 2006 *J. Am. Chem. Soc.* **128** 2067
[23] Liu H J, Zhou Y C, Yang Y, Wang W X, Qu L, Chen C, Liu D S, Zhang D Q and Zhu D B 2008 *J. Phys. Chem. B* **112** 6893
[24] Sun Y W, Liu H J, Xu L J, Wang L Y, Fan Q H and Liu D S 2010 *Langmuir* **26** 12496
[25] Seeman N C 2003 *Nature* **421** 427
[26] Dhakal S, Schonhoft J D, Koirala D, Yu Z B, Basu S and Mao H B 2010 *J. Am. Chem. Soc.* **132** 8991
[27] Dhakal S, Lafontaine J L, Yu Z B, Koirala D and Mao H B 2012 *PLoS One* **7** e39271
[28] Choi J K and Majima T 2013 *Photochem. Photobiol.* **89** 513
[29] Choi J K, Kim S Y, Tachikawa T, Fujitsuka M and Majima T 2011 *J. Am. Chem. Soc.* **133** 16146
[30] Zhang X H, Li M, Wang Y T and Ouyang Z C 2014 *Chin. Phys. B* **23** 020702
[31] Chen H, Fu H X, Zhu X Y, Cong P W, Nakamura F and Yan J 2011 *Biophys. J.* **100** 517
[32] Chen H, Zhu X Y, Cong P W, Sheetz M P, Nakamura F and Yan J 2011 *Biophys. J.* **101** 1231
[33] Smith S B, Cui Y J and Bustamante C 1996 *Science* **271** 795
[34] Cluzel P, Lebrun A, Heller C, Lavery R, Viovy J L, Chatenay D and Caron F 1996 *Science* **271** 792
[35] Fu H X, Chen H, Marko J F and Yan J 2010 *Nucleic Acids Res.* **38** 5594
[36] Cooper C L, Goulding A, Kayitmazer A B, Ulrich S, Stoll S, Turksen S, Yusa S, Kumar A and Dubin P L 2006 *Biomacromolecules* **7** 1025
[37] Cocco S, Yan J, Leger J F, Chatenay D and Marko J F 2004 *Phys. Rev. E* **70** 011910
[38] Anderson K L, Radford S E and Brockwell D J 2008 *The Dynamical Response of Proteins under Force*, in *Handbook of Molecular Force Spectroscopy* (New York: Springer) pp. 205–249
[39] Lagarias J C, Reeds J A, Wright M H and Wright P E 1998 *SIAM J. Optimiz* **9** 112
[40] Chen C, Li M, Xing Y Z, Li Y M, Joedecke C C, Jin J, Yang Z Q and Liu D S 2012 *Langmuir* **28** 17743
[41] Modi S, Swetha M G, Goswami D, Gupta G D, Mayor S and Krishnan Y 2009 *Nat. Nanotechnol.* **4** 325
[42] Sharma J, Chhabra R, Yan H and Liu Y 2007 *Chem. Commun.* **5** 477
[43] Cloutier T E and Widom J 2004 *Mol. Cell* **14** 355
[44] Yan J and Marko J F 2004 *Phys. Rev. Lett.* **93** 108108
[45] Du Q, Smith C, Shiffeldrim N, Vologodskaia M and Vologodskii A 2005 *Proc. Natl. Acad. Sci.* **102** 5397