

Role of Hoogsteen interaction in the stability of different phases of triplex DNA

Tanmoy Pal ^{*}, Keerti Chauhan, and Sanjay Kumar
Banaras Hindu University, Varanasi 221005, India

 (Received 24 October 2021; accepted 5 April 2022; published 21 April 2022)

A simple coarse-grained model of DNA which includes both Watson-Crick and Hoogsteen base pairing has been used to study the melting and unzipping of triplex DNA. Using Langevin dynamics simulations, we reproduce the qualitative features of one-step and two-step thermal melting of triplex as seen in experiments. The thermal melting phase diagram shows the existence of a stable interchain three-strand complex (bubble-bound state). Our studies based on the mechanical unzipping of a triplex revealed that it is mechanically more stable compared to an isolated duplex-DNA.

DOI: [10.1103/PhysRevE.105.044407](https://doi.org/10.1103/PhysRevE.105.044407)

I. INTRODUCTION

The structure of DNA is critical to its biological function. Beside the canonical B-DNA structure, DNA is capable of adopting several noncanonical forms, e.g., triplex, G-quadruplex etc. A DNA triple helix can form when a single strand of appropriate complementarity binds with a Watson-Crick (WC) paired duplex DNA via Hoogsteen (HO) base pairing mechanism [1]. Triple helix have been found both *in vivo* and *in vitro* conditions [2]. Three strand complexes can also form sans HO base pairing. A situation of this kind appears during homologous recombination [1]. There is also a proposal of a different type of bound state for three stranded DNA, named Efimov-DNA, which exists due to the thermal fluctuations near the duplex melting point [3–6]. The triplexes can be inter- or intramolecular. RNAs have also been shown to be involved in stable triplex structures [2]. Triplexes form with high probability when there are presence of mirror repeats [7]. The mirror repeats are widespread in Eukaryote genomes; as a result, triplex structures have been deemed to play controlling roles in gene expression, DNA replication and genome instability [2,8,9] etc. The sequence specific binding of the third strand has been utilized to recognize DNA sequences without disturbing the duplex structure and to design potential anti-gene strategy drugs. Synthetically designed triplex forming oligonucleotides (TFOs) can bind with specific DNA sequences to inhibit DNA protein interaction. This makes TFOs useful tools to engineer gene expression *in vivo* and a candidate for universal gene targeting drugs [10].

Due to their importance in biology [7], drug designing, and nanotechnology [10] it is important to understand the thermodynamic stability of a triplex. When heated, a stable triplex melts to three single strands at a critical temperature. Depending on the experimental conditions, the melting can be either a one-step (triplex directly melts to single strands)

process or a two-step (duplex and a single strand at low temperature and three strands at a higher temperature) process [11–15]. When the solution condition is such that no HO bond can form but WC bonds are stable, then also an inter-three-strand stable bound state can form. A similar dynamic state where one single strand oscillates between the other two has been reported earlier [16]. However, the direct monitoring of different kinds of base pairing can be difficult experimentally. Our aim is to better understand the stability of triplex by separately monitoring WC and HO base pairing.

During duplex DNA replication *in vivo*, the DNA is opened up by the class of enzymes called helicases which apply mechanical force at a constant pH/temperature. It has been shown in model systems that a DNA opens up abruptly when the unzipping force is increased beyond a critical value [17–19]. Like the duplex-DNA helicases, molecules have been detected which unwind triplex-DNA structures [20,21]. Helicases exert forces \sim pN. Through single molecule force spectroscopy (SMFS) techniques, it is now possible to generate such a small force. Manipulation of a DNA by these techniques provides important information about elastic properties of single and double-stranded DNA [22]. They also reveal structural properties in the form of many phases of dsDNA, e.g., S (overstretched), P (Pauling), SC (super-coiled) in addition to A, B, and Z DNA [22]. Although there are many studies which explore mechanical unwinding of duplex DNA, there are too few for the triplex unwinding case [23]. The aim of this communication is to first understand the role of different interactions in the melting of a triplex and then explore its mechanical stability.

Here, we introduce a simple coarse-grained model of triplex DNA which incorporates both the WC and HO types of interactions in its description. We use Langevin equation of motion to study the equilibrium properties of the triplex. First we study the zero force thermal melting and compare our results with available experiments. We show that the model is rich enough to reproduce the qualitative features of one-step and two-step melting as seen in Refs. [11–14]. We also report the phase diagram of the thermal melting of a triplex. Then, we study the triplex unzipping under the applied mechanical

^{*}tanmoy.pal@utoronto.ca; Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

force. Through force-temperature phase diagram, we provide evidence of the enhanced mechanical stability of triplex in comparison to duplex under unzipping force.

II. MODELS AND METHODS

To model duplex and triplex DNA in three-dimension, we use a coarse-grained approach by retaining only the essential features of DNA, e.g., native base-pairing, and ignore other microscopic details of nucleotides, Coulomb interaction, etc. Nucleotides are taken as beads, and the covalent bonds which form the backbone are modeled by harmonic springs. For our models described below, ϵ and a are the energy and length scales, respectively.

A. Isolated duplex

To model duplex-DNA in our study, we use almost the same model of Refs. [24,25], with only minor modifications. The total number of beads, $N = 64$, is always a multiple of 2, and the first [strand-1, Poly(A)] and the second [strand-2, Poly(T)] strand each consists of $n = N/2$ beads. The Hamiltonian of the system is given by

$$H = \sum_{\substack{i=1 \\ i \neq n}}^{N-1} \frac{K_{\text{bond}}}{2} (r_{i,i+1} - r_0)^2 + \sum_{i=1}^{N-1} \sum_{j \geq i+1}^N 4\epsilon \left[B_{i,j} \left(\frac{a}{r_{i,j}} \right)^{12} - A_{i,j} \left(\frac{a}{r_{i,j}} \right)^6 \right], \quad (1)$$

where $r_{i,j}$ is the distance between the i th bead and the j th bead, K_{bond} is the spring constant and r_0 is the equilibrium distance between two adjacent beads along a strand. The first term in Eq. (1) provides chain connectivity and the second term is the familiar Lennard-Jones (LJ) potential. The first term in LJ provides excluded volume interactions among the beads. We set $B_{i,j} = 1$ for all the bead-pairs except for the pairs at the same covalent bond for which $B_{i,j} = 0$. The attractive second term in LJ corresponds to the WC hydrogen-bonding interaction. We set $A_{i,j} = 1$ when i and j correspond to a native base-pair and $A_{i,j} = 0$ otherwise.

To study unzipping of the isolated duplex, we apply a constant force between the end monomers of the two strands. Numerically, we implement this by adding an energy $-g_* X_{32,33}$ to the Hamiltonian Eq. (1), where g_* is the magnitude of the force and $X_{32,33}$ is the absolute distance between the beads 32 and 33 along the x axis. For both melting and unzipping studies, the starting monomers are anchored at the r_0 distance apart.

B. Triplex

We adapt the isolated duplex model to introduce a model for a triplex. The total number of beads, $N (= 96)$, is always a multiple of 3, and the first [strand-1, Poly(T)], second [strand-2, Poly(A)], and third strand [strand-3, Poly(T)] each consists

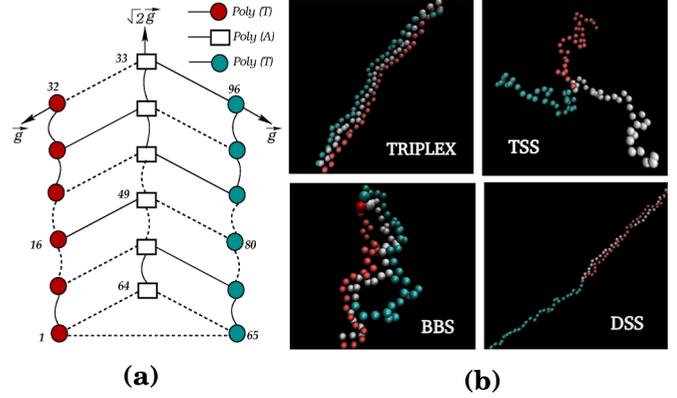


FIG. 1. (a) The bead-spring description of triplex DNA used in this study. The starting monomers of each of the strands are anchored at the vertices of an equilateral triangle. Solid and dashed black lines represent Watson-Crick and Hoogsteen bonds, respectively. Unzipping force is applied at the end monomers. The force is applied in the X - Y plane in such a manner that the two pairs of strands (strand-1, strand-2) and (strand-2, strand-3) always feel the same magnitude of unzipping force. (b) Representative snapshots obtained from our simulations showing TRIPLEX (at $T \approx 312\text{K}$, $g = 0$, $\epsilon_H = 1$), three single strands (TSS, at $T \approx 390\text{K}$, $g = 0$, $\epsilon_H = 0.6$), bubble-bound state (BBS, $T \approx 335\text{K}$, $g = 0$, $\epsilon_H = 0.6$), and duplex + single-strand (DSS, $T \approx 312\text{K}$, $g \approx 56\text{pN}$, $\epsilon_H = 0.6$). The color convention is of panel (a).

of $n = N/3$ beads. The Hamiltonian of the system is given by

$$H = \sum_{\substack{i=1 \\ i \neq n, i \neq 2n}}^{N-1} \frac{K_{\text{bond}}}{2} (r_{i,i+1} - r_0)^2 + \sum_{i=1}^{N-1} \sum_{j \geq i+1}^N 4\epsilon \left[B_{i,j} \left(\frac{a}{r_{i,j}} \right)^{12} - A_{i,j} \left(\frac{a}{r_{i,j}} \right)^6 \right], \quad (2)$$

where $r_{i,j}$ is the distance between the i th bead and the j th bead. The first term in Eq. (2) provides chain connectivity (bonded interaction) and the second term is the familiar Lennard-Jones (LJ) potential (nonbonded interaction). The first term in LJ provides excluded volume interactions among beads. We set $B_{i,j} = 1$ for all the bead-pairs except for the pairs at the same covalent bond for which $B_{i,j} = 0$. $A_{i,j}$ correspond to both WC ($\epsilon_{WC} = 1$) and HO (ϵ_H) type interaction strengths between the i th and the j th beads. As the strand-1 and strand-3 are Poly(T) and the strand-2 is a Poly(A) (Fig. 1), strand-2 can form native base-pairs with both strand-1 and strand-3. Once a bead of strand-2 has formed WC-bond with its native pair on, say, strand-1, it cannot form WC bond with its other native pair on strand-3; however, it can form HO-bond with its strand-3 native pair, if allowed. To incorporate this numerically, we chose a common Hydrogen bonding distance $r_c = 1.5a$. If the distance between any two native base-pairs is $< r_c$, we say that they are Hydrogen-bonded. This criteria for base-pairing also applies to the isolated duplex model. For a system with $N = 96$, 16th (T), 49th (A), and 80th (T) beads, for example, are native base pairs (Fig. 1). To incorporate both WC and HO interactions during simulation, we implement the strategy shown in Table I. We set

TABLE I. Algorithm: The native base-pairing interactions are tuned at each Langevin step.

Base-pair distance	Interaction strengths
$r_{16,49} < r_c$ & $r_{49,80} > r_c$	$A_{16,49} = 1$ & $A_{49,80} = \varepsilon_H$
$r_{16,49} > r_c$ & $r_{49,80} < r_c$	$A_{16,49} = \varepsilon_H$ & $A_{49,80} = 1$
$r_{16,49} < r_c$ & $r_{49,80} < r_c$ & $r_{16,49} < r_{49,80}$	$A_{16,49} = 1$ & $A_{49,80} = \varepsilon_H$
$r_{16,49} < r_c$ & $r_{49,80} < r_c$ & $r_{16,49} > r_{49,80}$	$A_{16,49} = \varepsilon_H$ & $A_{49,80} = 1$
$r_{16,49} > r_c$ & $r_{49,80} > r_c$	$A_{16,49} = 1$ & $A_{49,80} = 1$

$A_{i,j} = 0$ for all the bead-pairs that are not native. Following this strategy, we reset $A_{i,j}$ at each Langevin time step. To study the force-induced unzipping transition in a triplex, a constant unzipping force of magnitude g_* is the applied at the end monomers between strand-1 (32nd bead) and strand-2 (33rd bead) along the y axis and between strand-2 (33rd bead) and strand-3 (96th bead) along the x axis. The resultant force protocol is shown in Fig. 1. As there is no direct attractive interaction between strand-1 and strand-3 by construction, we chose not to apply any unzipping force between those strands. We implement this numerically by adding the energies $-g_*Y_{32,33}$ and $-g_*X_{33,96}$ to the Hamiltonian Eq. (2) where $Y_{32,33}$ is the absolute distance between beads 32 and 33 along the y axis and where $X_{33,96}$ is the absolute distance between beads 33 and 96 along the x axis. We chose this force protocol to compare triplex unzipping results to the isolated duplex unzipping results. For all the simulations with triplex, the starting monomers of the three strands are anchored at the vertices of an equilateral triangle of side length r_0 . Note that, as a result of this force protocol, the maximum end-to-end distance between strand-1, strand-2 and strand-2, strand-3 pairs could be $Y_{12,\max} = Y_{23,\max} = \sin(3\pi/4)/\sin(\pi/8)L$ and between strand-1, strand-3 could be $Y_{13,\max} = \sqrt{2}L$ where $L = (n-1)r_0$.

For both isolated duplex and triplex, we set the spring constant and equilibrium bond length to $K_{\text{bond}} = 200\epsilon/a^2$ and $r_0 = 1.12a$, respectively, and obtain the time evolution by solving the Langevin equation given by

$$m \frac{d^2 \mathbf{r}}{dt^2} = -\zeta \frac{d\mathbf{r}}{dt} + \mathbf{F}_c + \mathbf{\Gamma},$$

for each bead using the sixth order Gear algorithm with a time step $\Delta t = 0.025\tau$ where $\tau \equiv \sqrt{ma^2/\epsilon}$, m is the mass of each bead. $\zeta = 0.4m/\tau$ is the friction coefficient. $\mathbf{F}_c = -\nabla H$ is the conservative force, and $\mathbf{\Gamma}$ is random force with zero mean. This is related to the friction coefficient by the fluctuation-dissipation relation $\langle \Gamma_i(t) \Gamma_j(t') \rangle = 2\zeta T_* \delta_{i,j} \delta(t-t')$, where T_* is a reduced temperature, $\delta_{i,j}$ is the Kronecker δ function and $\delta(t-t')$ is the Dirac δ function.

Following Ref. [25], we convert our model reduced temperature and force to their real units by choosing $\epsilon = 2.306$ kcal/mol and $a = 1$ Å. The temperature in K is given by $T = 363 + 389 \times (T_* - 0.23)$ and the force in pN is given by $g = 160 \times g_*$. It is to be noted that even though the linear scale used here to convert the reduced temperatures to real temperatures produces experimentally reasonable melting

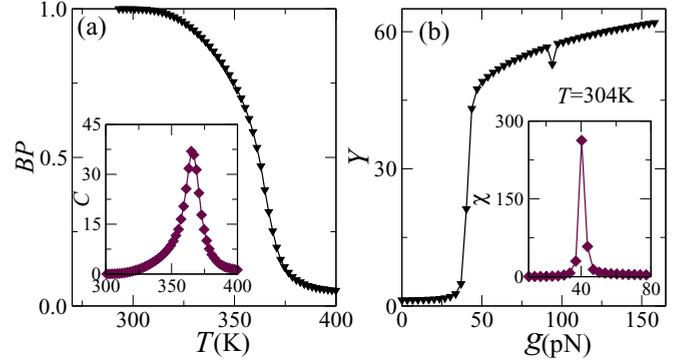


FIG. 2. Zero force melting and forced unzipping of an isolated duplex. (a) Fraction of intact base pairs vs temperature plot with the corresponding specific heat shown in the inset. (b) Separation between the end monomers of the two strands vs unzipping force plot with the corresponding fluctuation in extension shown in the inset. Lines through the data points are guide for the eye.

temperatures, this simple scaling is not valid for all T_* values [25] and could be used in the T_* range considered here to get a qualitative measure of the actual temperature.

The response due to force is given by the distance (Y) between the end monomers of any two strands and the fluctuation in Y is quantified by its variance, χ . The specific heat (C) and χ are defined as

$$C = \frac{\langle E^2 \rangle - \langle E \rangle^2}{T_*^2} \text{ and } \chi = \langle Y^2 \rangle - \langle Y \rangle^2, \quad (3)$$

where E is the total energy and the brackets represent averaging over simulation time.

III. RESULTS

A. Thermal melting and forced unzipping of duplex

To do a comparative study of thermodynamic properties between a triplex and a duplex DNA, first we study thermal melting and force unzipping of an isolated duplex with the interaction Hamiltonian Eq. (1). From hereupon, we represent the fraction of bound native pairs by BP . Corresponding thermal melting and forced unzipping results are shown in Fig. 2. For both duplex and triplex, we identify the temperature at which the specific heat peaks as the critical temperature and the force at which χ peaks as the critical force. With this convention, the isolated duplex melts at $T_c = 363$ K at zero force and unzips at $g \approx 41$ pN at $T = 304$ K. χ versus force plots for other temperatures are shown in Supplemental Material Fig. S1 [26] and the critical force versus temperature phase diagram is shown in Fig. 4 below.

B. Thermal melting of triplex

To study thermal melting, we take a triplex DNA of length $N/3 = 32$. Since the relative HO interaction is variable with respect to the WC strength, we vary ε_H in between 0 to 1, keeping the WC interaction fixed (i.e., $\varepsilon_{WC} = 1$). The average number of intact base-pair fraction (BP), average end-to-end distance (Y) and the corresponding specific heat (C) at different temperatures are shown in Fig. 3 for three different

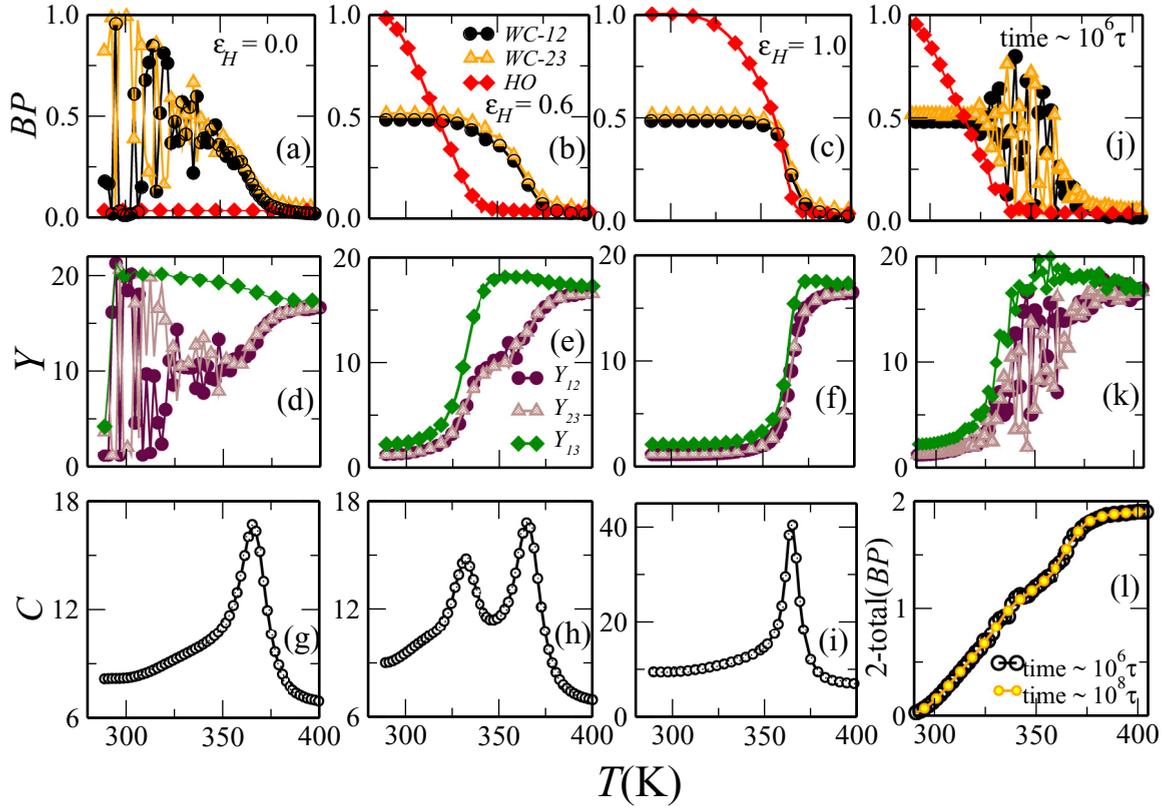


FIG. 3. Zero force melting of a triplex DNA for different ϵ_H . Panels (a–c) show the variation of fraction of bound base-pairs (BP) as a function of temperature. Here, WC-12 represents WC bonds between strand-1 and strand-2 and WC-23 represents WC bonds between strand-2 and strand-3. HO represents HO bonds. End-to-end distances are shown in panels (d–f). Panels (g–i) show the specific heat (C) as a function of temperature for different ϵ_H . Panels (j, k) are same as panels (b, e) but for nonequilibrated case for $\epsilon_H = 0.6$. (l) Comparative plot of UV absorption and $2 - (\text{WC-12} + \text{WC-23} + \text{HO})$ vs temperature. Lines through the data points are guides for the eye.

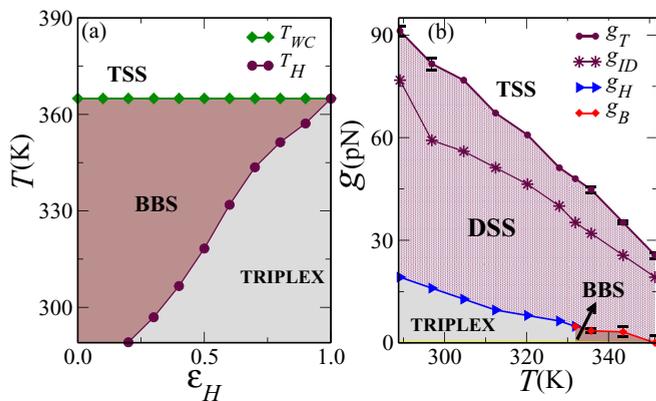


FIG. 4. (a) Zero force melting phase diagram. Here, T_H is the temperature above which no HO-bonds remain intact and above T_{WC} all the WC-bonds break. (b) Force vs temperature phase diagram for $\epsilon_H = 0.6$. The filled blue triangles (g_H) and red diamonds (g_B) separate the triplex bound state phase and the bubble-bound state (BBS) from the single-strand + duplex phase, respectively. The filled maroon circles (g_T) act as the phase boundary between the single-strand + duplex phase and the phase where all three strands are completely separated from each other. The maroon stars (g_D) represent the critical unzipping forces for an isolated duplex.

values of ϵ_H . After an equilibration time of $2 \times 10^8 \Delta t$, configurations were sampled on an interval of Δt for a duration of $4 \times 10^8 \Delta t$. We represent WC bonds between strand-1 and strand-2 as WC-12, etc. and Hogsteen bonds as HO. When the condition is such that $HO \approx 0$ and either $WC-12 \approx 1$ & $WC-23 \approx 0$ or $WC-12 \approx 0$ & $WC-23 \approx 1$, the system stays in the Duplex+Single-Strand (DSS) state. However, when $HO \approx 0$ and both WC-12 and WC-23 are close to 0.5, we say the system is in the bubble-bound state (BBS) [27]. The advantage of our model is that one can compute the fraction of WC base pairs and the HO base pairs separately which is difficult to monitor experimentally.

For $\epsilon_H = 0$, there is no HO bonding contribution and the system is dominated by the WC base-pairing only [Fig. 3(a)]. At high temperatures, $T > 363$ K, all bound base pairs break and the system attains the molten state consisting of three single strands. This one step melting is reflected in the specific heat plot [Fig. 3(g)], which shows a single peak at $T_c \approx 363$ K. At very low temperatures, the WC bonds are quite stable against thermal knocks. As a result, we see an oscillating region in temperature where total WC bonds do not partition equally between WC-12 and WC-23. However, none of WC-12 and WC-23 is close to zero separately. Existence of this oscillating region is more evident from Supplemental Material Fig. S2 [26], where we show data for 10 independent runs.

Data from a longer time simulation run presented in Supplemental Material Fig. S3 [26] also indicate the existence of the oscillatory region. This unequal partitioning at very low temperatures gives rise to the possibility of a DSS to BBS transition. However, absence of any low temperature peak in the corresponding specific heat plot in Fig. 3(g) makes any such transition unlikely. From the symmetry of our model, i.e., the strand-2 is equally likely to form WC base pairs with either one of the other two strands, one expects an equal partitioning of the time-averaged WC bonds in the infinite time limit. Given the finite duration of time averaging used in our simulations, we attribute this apparent discrepancy of un-equal partitioning to the limited simulation time. An increase in temperature admits larger thermal fluctuations and in the intermediate temperatures ($< T_c$) the WC bonds partition almost equally. To clarify the state of the triplex DNA, we plot end-to-end distances in Fig. 3(d). Comparing Figs. 3(a) and 3(d) we see that for the temperature region where the fraction of intact base pairs follow $WC-12 \approx WC-23$, end-to-end distances also follow $Y_{12} \approx Y_{23} \ll Y_{12,\max}$ and $Y_{13} \ll Y_{13,\max}$ where $Y_{i,j,\max}$ is the maximum possible end-to-end distance between strand- i and strand- j . This indicates that the state of the triplex DNA is BBS rather than DSS, as illustrated by the representative simulation snapshots in Fig. 1(b). For convenience, we classify the state of the triplex DNA in the whole oscillatory region for this $\varepsilon_H = 0$ case as BBS, which means BBS is possible up to T_c .

The $\varepsilon_H = 0.6$ case is also quite interesting. Fig. 3(b), together with Fig. 3(e), indicate that at very low temperatures ($T < 328$ K), an H-bonded triplex bound state with no bubbles forms, where the intact base pairs are either WC bonded or Hoogsteen bonded. In the temperature range 328 K $< T < T_c$, one can see the absence of HO bonds and the system stays in the bubble-bound state. The specific heat plot for this case shows two peaks. The first peak corresponds to the transition from the triplex bound state to the bubble-bound state, and the second peak is the signature of the duplex melting [Fig. 3(h)].

Finally, when the HO interaction strength is same as the WC interaction strength ($\varepsilon_H = 1$), HO bonding is possible up to T_c [Fig. 3(c), 3(f) and 3(i)]. Here, the triplex melts directly into the three single strands. The specific heat plot for this case shows a single peak.

Although most of our results are in qualitative agreement with the calorimetry and UV absorbance results of Refs. [13,15], they differ in the following aspect. When HO interaction strength is intermediate, say $\varepsilon_H = 0.6$, for 340 K $< T < T_c$ we get an equilibrium configuration of the system where it stays in the bubble-bound state (Fig. 3), while the results in Refs. [13,15] suggest that the corresponding state is made off duplex plus a single strand (DSS). One of the reason this conclusion was made because of the observation that Triplex to TSS transition temperature was the same as the melting temperature of an isolated duplex, which we also observed in our simulations. To understand this apparent discrepancy, we plot again the fraction of H-bonds and end-to-end distance as a function of temperature in Figs. 3(j) and 3(k), respectively, but this time we have given less time to the system for its equilibration. It is to be noted that DNA melting experiments utilize hyperchromicity of bare nucleic acid bases at wavelength $\lambda = 260$ nm [1]. When the nucleic

acids in the strands are Hydrogen bonded, they absorb less UV light. Accordingly, the absorbance curves obtained from experiments have correspondence with $1 - \langle WC \rangle$ (duplex) and $2 - \langle WC-12 + WC-23 + HO \rangle$ (triplex) obtained from the simulations. There is not much difference between the curves for the equilibration time $\sim 10^6 \Delta t$ and $\sim 10^8 \Delta t$ shown in Fig. 3(l) and they resemble the UV absorption spectra of Refs. [13,15]. It is clear that for the nonequilibrated case, the system is closer to the DSS state compared to the equilibrated system. Thus, it could be tricky to differentiate between BBS and DSS using UV absorption spectroscopy only. Even when differentiation between BBS and DSS is possible, our results suggest that equilibration time could be an issue. Detection of the thermodynamic BBS in calorimetry experiments may require very slow heating rate and large equilibration time. It is also to be noted here that a different sequence of triplex was used in the study of Ref. [13].

By varying ε_H from zero to one and locating the triplex to bubble-bound state transition points and the duplex melting point (see Supplemental Material Fig. S4 [26]), we construct a phase diagram for zero force thermal melting which is shown in Fig. 4(a). Due to large equilibration time at very low temperatures, reliable results are hard to obtain. For this, we avoid temperatures < 289 K. As ε_H increases, the phase boundaries gradually converge. Although the low-temperature critical point depends on ε_H , the high-temperature critical point is practically independent of ε_H . This result is consistent with the ph-dependence of the critical points studied in Ref. [13]. However, for our set up triplex and an isolated duplex melt at the same temperature. This is unlike the results suggested in Refs. [3–6] where the triplex melts at a higher temperature. For $\varepsilon_H = 0.9$, we found the second peak is very hard to detect. To be consistent with the rest of the ε_H cases considered here, we have taken the position of the second peak as T_c .

C. Forced unzipping of triplex

To study the unzipping transition, we consider a triplex of length 32 and keep ε_H constant ($= 0.6$). We apply a force at the end monomers of the single strands in such a way that the same magnitude of the force is experienced between strand-1 & strand-2 and strand-2 & strand-3. The manner in which the unzipping force is applied is shown in Fig. 1. We plot the average distance between the end monomers (Y) of any two pairs of chains as a function of force in Fig. 5 for two different temperatures. After an equilibration time of $1 \times 10^8 \Delta t$, configurations were sampled on an interval of Δt for a duration of $2 \times 10^8 \Delta t$. At zero force, it is evident from Fig. 4(a) that for $T < 331$ K and 331 K $< T < T_c$, triplex bound state and bubble-bound state, respectively, are the stable states. Although no direct H-bonding is allowed between strand-1 and strand-3, they can bind indirectly through HO and WC bonding which is mediated by strand-2. The HO bonding being weaker than the WC bonding in this case, strand-1 gets separated from strand-3 first when the unzipping force is increased. However, as long as the force is not strong enough to break all the WC bonds, the duplex state remains stable and the strand-2 can bind with any of the other two strands. We also calculate the fluctuation in end-monomer

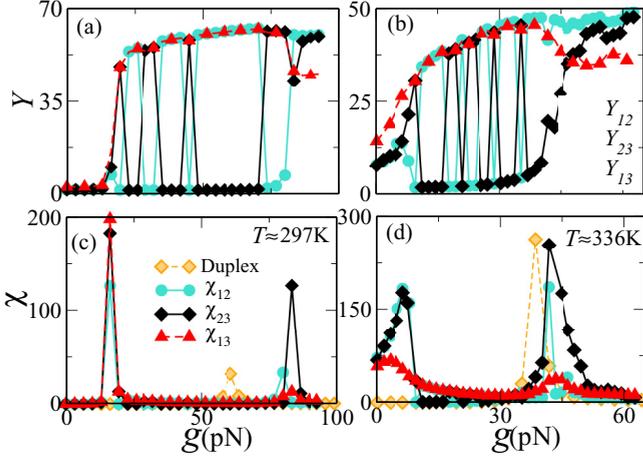


FIG. 5. (a, b) Force-Extension plots for triplex DNA for different temperatures at $\varepsilon_H = 0.6$. Here Y_{12} represents extension between chain-1 and chain-2, etc. Panels (c, d) show the corresponding fluctuation (χ) curves of the extension [Eq. (3)]. In panels (c, d), corresponding data for an isolated duplex is also shown (filled orange diamonds). Lines through the data points are guide for the eye.

distances (χ) and plot them in Fig. 5 as a function of force for two temperatures. At $T < 331$ K, the fluctuation plots in Fig. 5(c) show two peaks. The first peak corresponds to the transition from triplex bound state to duplex + single-strand (DSS) state while the second peak represents the transition to the three completely separated single strands state. Note here that for DSS, $Y_{13} \approx Y_{13,\max}$ and either $Y_{12} \approx 0$ & $Y_{23} \approx Y_{23,\max}$ or $Y_{23} \approx 0$ & $Y_{12} \approx Y_{12,\max}$ which are markedly different from the Y values for BBS found in the zero force thermal melting case in Fig. 3. Representative simulation snapshots of BBS and DSS are shown in Fig. 1(b). Above $T_c > T > 331$ K also, there are two peaks [Fig. 5(d)]. But, here, the first peak corresponds to the transition from the bubble-bound state to the duplex + single-strand state. One important feature of the unzipping plots can be seen if we compare the duplex unzipping force in the presence of a third strand with the unzipping force of an isolated duplex. It turns out that the duplex unzipping force in the presence of a third strand is always greater than the unzipping force of an isolated duplex (peaks in χ for the isolated duplex appear at smaller forces compared to the corresponding peaks for the triplex in Fig. 5). To illustrate this observation further, we plot the force-extension curves for different ε_H values at $T = 304$ K in Fig. 6. For $\varepsilon_H = 1$, the first peak appears at the same force where an isolated duplex unzips, but the position of the second peak, where the triplex unzips to three single strands, is significantly higher than the first peak. Although the critical point at lower force depends on ε_H , the transition at higher force is independent of ε_H . These results indicate that the observed enhanced mechanical stability of the triplex is not due to the HO-interaction, but due to the presence of the third strand. Note that, when the stable state of the triplex is the duplex + single-strand state, strand-2 can form the duplex with either of the other two strands exclusively. This fact is more clearly demonstrated by the Supplemental Material Figs. S5– S8 [26]. We also note

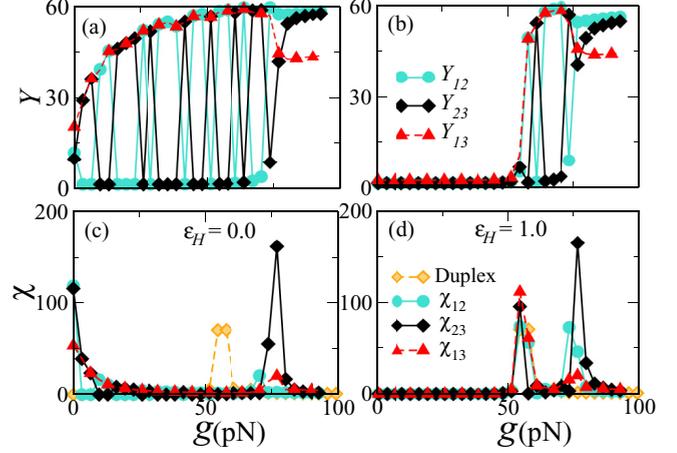


FIG. 6. Same as described in the caption of Fig. 5 but at $T = 304$ K and at $\varepsilon_H = 0$ and 1.

that the appearance of χ peaks at $g = 0$ for $\varepsilon_H = 0$ in Fig. 6(c) points toward a transition between BBS to DSS with $g = 0$ serving as the transition point.

For $\varepsilon_H = 0.6$, we compute the critical unzipping force vs temperature phase diagram [Fig. 4(b)]. By convention, we take the positions of the peaks in the χ_{13} vs g plots as the transition points. The complete set of χ_{13} curves for different temperatures are shown in Supplemental Material Fig. S3 [26]. For comparison, we also superimpose the phase diagram of an isolated duplex in Fig. 4(b). As the zero force melting temperature is approached, the numerical data become increasingly unreliable due to fluctuations. To be safe, we avoid temperatures $T > 351$ K for our mechanical unzipping studies.

IV. CONCLUSION

By introducing a simple coarse-grained model which incorporates both the Watson-Crick and Hoogsteen base-pairing simultaneously, melting and unzipping of triplex DNA have been studied. Here, we have shown that the model is rich enough to capture the qualitative features of the one-step and the two-step experimental triplex-melting results. By directly monitoring the fraction of intact Watson-Crick and Hoogsteen bonds separately, we showed that for a range of parameter values, a stable inter-three-strand bubble-bound state can form. A Hoogsteen interaction strength versus critical temperature phase diagram is also obtained. We also study the mechanical unzipping of a triplex by applying constant unzipping force. The simple model system proposed here clearly demonstrates that the triplex can be mechanically more stable than the isolated duplex. The melting temperature phase diagram exhibits that the bubble-bound state can be stabilized in presence of Hoogsteen interaction. Although there are experimental studies concerning the rupture of a triplex are available [28,29], it is unknown to us whether a triplex unzipping study exists. Thus, at this stage, our simulations warrant further experimental investigations preferably using SMFS devices.

ACKNOWLEDGMENTS

The authors thank an anonymous referee for several critical suggestions that have greatly improved the presentation of the paper. T.P. thanks Garima Mishra for insightful discussions.

K.C. thanks DST, India, for providing the INSPIRE Fellowship. The financial assistance from SERB, India, SPARC, MHRD, India, IoE, MHRD, India are gratefully acknowledged. S.K. thanks ICTP, Trieste, Italy, where a part of the work has been carried out.

-
- [1] R. R. Sinden, *DNA Structure and Function* (Academic Press, San Diego, CA), 1994.
- [2] F. A. Buske, J. S. Mattick, and T. L. Bailey, *RNA Biol.* **8**, 427 (2011).
- [3] J. Maji, S. M. Bhattacharjee, F. Seno, and A. Trovato, *New J. Phys.* **12**, 083057 (2010).
- [4] T. Pal, P. Sadhukhan, and S. M. Bhattacharjee, *Phys. Rev. Lett.* **110**, 028105 (2013).
- [5] T. Pal, P. Sadhukhan, and S. M. Bhattacharjee, *Phys. Rev. E* **91**, 042105 (2015).
- [6] S. M. Bhattacharjee and D. P. Foster, *J. Chem. Phys.* **155**, 064903 (2021).
- [7] M. D. Frank-Kamenetskii and S. M. Mirkin, *Annu. Rev. Biochem.* **64**, 65 (1995).
- [8] M. K. Tiwari, N. Adaku, N. Peart, and F. A. Rogers, *Nucleic Acids Res.* **44**, 7742 (2016).
- [9] I. T. Holder, S. Wagner, P. Xiong, M. Sinn, T. Frickey, A. Mayer, and J. S. Hartig, *Nucleic Acids Res.* **43**, 10126 (2015).
- [10] A. R. Chandrasekaran and D. A. Rusling, *Nucleic Acids Res.* **46**, 1021 (2018).
- [11] D. S. Pilch, C. Levenson, and R. H. Shafer, *Biochemistry* **30**, 6081 (1991).
- [12] F. Svinarchuk, J. Paoletti, and C. Malvy, *J. Biol. Chem.* **270**, 14068 (1995).
- [13] G. E. Plum, Y. W. Park, S. F. Singleton, P. B. Dervan, and K. J. Breslauer, *Proc. Natl. Acad. Sci. USA* **87**, 9436 (1990).
- [14] R. W. Roberts and D. M. Crothers, *Science* **258**, 1463 (1992).
- [15] I. Haq, J. E. Ladbury, B. Z. Chowdhry, and T. C. Jenkins, *J. Am. Chem. Soc.* **118**, 10693 (1996).
- [16] S. K. Jain, M. M. Cox, and R. B. Inman, *J. Biol. Chem.* **270**, 4943 (1995).
- [17] S. M. Bhattacharjee, *J. Phys. A* **33**, L423 (2000).
- [18] D. K. Lubensky and D. R. Nelson, *Phys. Rev. Lett.* **85**, 1572 (2000).
- [19] R. Kapri, S. M. Bhattacharjee, and F. Seno, *Phys. Rev. Lett.* **93**, 248102 (2004).
- [20] R. M. Brosh, Jr., A. Majumder, S. Desai, I. D. Hickson, V. A. Bohr, and M. M. Seidman, *J. Biol. Chem.* **276**, 3024 (2001).
- [21] M. Guo, K. Hundseth, H. Ding, V. Vidyasagar, A. Inoue, C. H. Nguyen, R. Zain, J. S. Lee, and Y. Wu, *J. Biol. Chem.* **290**, 5174 (2015).
- [22] S. Kumar and M. S. Li, *Phys. Rep.* **486**, 1 (2010).
- [23] L. Ling, H. J. Butt, and R. Berger, *J. Am. Chem. Soc.* **126**, 13992 (2004).
- [24] G. Mishra, D. Giri, M. S. Li, and S. Kumar, *J. Chem. Phys.* **135**, 035102 (2011).
- [25] G. Mishra, P. Sadhukhan, S. M. Bhattacharjee, and S. Kumar, *Phys. Rev. E* **87**, 022718 (2013).
- [26] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.105.044407> for supplementary data.
- [27] J. Maji, F. Seno, A. Trovato, and S. M. Bhattacharjee, *J. Stat. Mech.* (2017) 073203.
- [28] C-C. Chan, P-Y. Lin, Y-F. Chen, C-S. Chang and L-S. Kan, *Appl. Phys. Lett.* **91**, 203901 (2007).
- [29] L. Ling, H-J. Butt and R. Berger, *Appl. Phys. Lett.* **89**, 113902 (2006).