Effect of loop sequence on unzipping of short DNA hairpins

Anurag Upadhyaya and Sanjay Kumar

Department of Physics, Banaras Hindu University, Varanasi, 221 005, India

The dependence of stability on the sequence of a DNA hairpin has been investigated through atomistic simulations. For this, a sequence of 16 bases of a hairpin, which consists of a loop of four bases and a stem of six base pairs, has been considered. We have taken eight different sequences, where the first five base pairs were kept fixed in all sequences, whereas the loop sequence and the identity of the duplex base pair closing the loop have been varied. For these hairpin structures, force-induced melting (unzipping) studies were carried out to investigate the effect of the variables on the stability of hairpin. The temperature at which half of the base pairs are open is termed the melting temperature. We defined the unzipping force F_h (half of the base pairs are open) and showed that it may not provide the effect of closing the base pair or loop sequence on the stability of the DNA hairpin. In order to have a better understanding of the stability of a DNA hairpin, the closing base pair or hairpin loop must be open. This requires complete opening of the stem. We defined a force F_c at which all base pairs of the stem are open, and we showed that the F_c gives better understanding of DNA hairpin stability.

I. INTRODUCTION

Recent advances in the understanding of the melting behavior of nucleic acids (RNA, single-stranded DNA) with computer simulations have led to improved insights for processes like transcription and replication. These findings were further supplemented by single-molecule force-spectroscopic (SMFS) techniques, which measured the molecular forces at play during these processes [1–13]. It is now realized that conformations that can be adopted by nucleic acids are much more diverse, complex, and curious than the regular double helix originally envisioned. Earlier it was thought that the conformational space available to nucleic acids will be crucial in cellular processes. Motivated by this, attempts have been made to explain the structural properties based on the nucleotide sequences.

It is well known that a single nucleotide change in a particular sequence of a double-stranded DNA (dsDNA) cannot be detected by electrophoresis. It was found that the physical properties of the double strands are almost the same for both cases [14]. However, after denaturation, a single-stranded DNA (ssDNA) may acquire a unique conformational state based on its sequence. The difference in shape between two ssDNA strands with different sequences may cause them to move differently on an electrophoresis gel, although the number of nucleotides is the same [15]. Among such polymorphic conformations, hairpin loop structures in RNA and ssDNA play a special role in biological functions such as the regulation of gene expression, DNA recombination, and facilitation of mutagenic events [16–27]. The stability and conformational fluctuations of the hairpin structure have recently been explored by designing simple ssDNA oligonucleotides that have n ($\sim$47) complementary bases at each end called a stem and one type of m ($\sim$430) nucleotides in the middle of the chain referred to as a loop [28]. The stem shows the same response to a change in solution conditions as a dsDNA oligomer. However, the loop region shows a wide range of folding patterns that depend on the number and sequence in the loop. Portella and Orozco studied a short DNA hairpin (dGCGAAGC) and observed multiple routes of folding [29]. Their results indicated that hairpin folding is a complex process even for a simple system.

Goddard *et al.* [30] investigated sequence dependent rigidity of ssDNA that can form a hairpin. Keeping the stem sequence fixed, they considered loops made up of either poly T or poly A. They observed that the ssDNA distortion is purely entropic for poly T but requires an additional enthalpy for poly A. Vallone *et al.* [31] have studied the thermodynamic effect of a loop sequence and the base pair closing the loop on the stability of short DNA hairpins, which consist of a six base pairs stem and a loop of four bases. In the stem, five base pairs were kept fixed, and a base pair closing the loop has been changed for different loop sequences. They observed that 18 out of 28 melting profiles can be well represented by the two-state model [32]; however, the rest showed deviations from the two-state model. However, their measurements enabled a thermodynamic characterization of the loop sequence dependence of these hairpins but was lacking in several aspects. These measurements are indirect and provide overall information of the system. For example, melting is usually defined when half of the base pairs are open or corresponding to differential melting curves [33]. It is almost difficult to pinpoint whether a particular base pair is open or not. Hence, getting precise information about the role of a closing base pair from such a study is quite challenging.

In this respect, SMFS experiments have provided unprecedented information about the cellular processes, which directly measured the forces at the molecular level [3,9,34–36]. In particular, optical tweezers and scanning force microscopy revealed the unusual elastic properties of nucleic acids, e.g., semimicroscopic changes in the monomer (nucleotide) are found to influence the elastic property of

the nucleic acids. Attempts were made to obtain the force-extension (*F*-*x*) curve of ssDNA and RNA consisting of only one type of nucleotide [37–39]. The elastic properties of nucleic acid made up of adenine [poly(dA)] is found to be significantly different from that of thymine [poly(dT)] or uracil [poly(rU)] [37]. The *F*-*x* curve for poly(rU) or poly (dT) showed the entropic response, whereas poly(dA) exhibits a plateau in the *F*-*x* curve due to intrastrand stacking interaction. Ke *et al.* [38] studied the elastic properties of ssDNA and found the existence of multistep plateaus in the case of poly(dA) [38]. Woodside *et al.* [40] considered 20 different hairpin sequences and studied the folding-unfolding transition of a hairpin. They obtained the essential features required for the construction of an energy landscape. In most cases of the DNA hairpin, the applied force is kept close to a critical value. As a result, the hairpin fluctuates between the closed and open state. Therefore, efforts were mainly focused to understand the kinetics of the hairpin in the presence of the applied force.

Most of the theoretical works focused on simple models, which are either analytically solvable or an accurate solution is possible through the extensive numerical simulations. Theoretical analysis of the elasticity of a polymer chain with hairpins as secondary structures by Montanari and Mezard [41] reproduces the experimental *F*-*x* curve measured on the ssDNA chains, whose nucleotide bases are arranged in a relatively random order. The force-induced transition in the hairpin appears to be of second order and characterized by a gradual decrease in the number of base pairs (*m*) as the external force increases. Zhou *et al.* [42] studied the secondary structure formation of the ssDNA (or RNA) both analytically as well numerically. They showed that the force-induced transition is continuous from the hairpin-I (small base stacking interaction) to the coil, while of a first order for the hairpin-II (large base stacking interaction). Hugel *et al.* [43] investigated three different chains, namely, ssDNA, poly vinylamine, and peptide at very high force (2 nN). At such a high force, conformational entropy does not have a significant role, therefore, a zero temperature *ab initio* calculation has been applied to compare the experimental results.

Mishra *et al.* [44] developed a simple coarse-grained model of polymers and performed Langevin dynamics simulations to obtain the force-temperature diagram of DNA and the DNA hairpin. They successively demonstrated that at low temperature, the hairpin-coil transition is force driven, while at high temperature, the transition is entropy driven. They measured the entropic force as a function of the loop length. However, they considered the homosequence of loop and stem and, therefore, could capture only the limited insight in the mechanism involved in the stability of the DNA hairpin. The more accurate picture of the stability of the DNA hairpin may be probed through the atomistic simulations.

Before studying the stability and structural properties of the DNA hairpin through atomistic simulations, it is important to note that the hairpin-coil transition falls in the large-scale motion of DNA. The typical time and length scales involved are 10^{-7} to 10^1 s and more than 10 Å, respectively. The thermal stability of the DNA hairpin in such a timescale is difficult to achieve with our computer resources. Therefore, we resort to force-induced melting of the DNA hairpin to overcome the requirement of large computational time. The
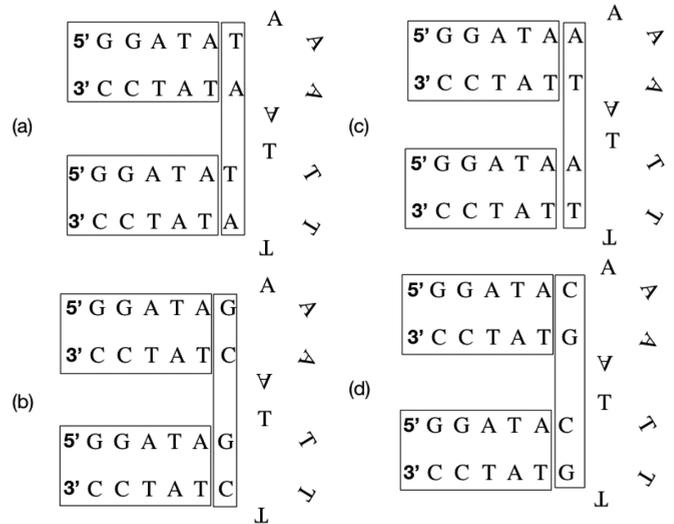
FIG. 1. Schematic representations of the eight different sequences of DNA hairpins of 16 bases. The stem consists of six base pairs, out of which the first five base pairs remain (shown in the horizontal boxes) and are the same in all sequences. The loop consists of four bases of either A or T. The closing base pair remains the same (shown in the vertical boxes), whereas loop sequences are different. We chose different combinations of closing base pair and loop sequence to explore the stability of a DNA hairpin.

critical force, where the hairpin unfolds, is the direct measure of the stability. However, before the hairpin unfolds, it is important to know how the structure changes and whether the closing base pair is open or not. Such microscopic information cannot be detected through experiments.

The aim of the present work is to provide an atomistic description of the stability of the DNA hairpin arising due to the sequence. For this, we choose some of the sequences (Fig. 1), whose melting profiles were studied by Vallone *et al.* [31]. We perform molecular dynamics simulations (AMBER [45]) to obtain the force-extension curve, and we explore the stability of these short DNA hairpins. The outline of the paper is as follows. In Sec. II we introduce the model used for the hairpin. We provide details of the simulations and the response of structural changes against the applied force in Sec. III. We performed 100 simulations using different seeds and analyzed results in terms of force-extension curves obtained for different sequences in Sec. IV. The paper ends with a short summary and conclusion in Sec. V.

II. MODELING OF THE DNA HAIRPIN

We used the AMBER10 software package [45] with all-atom (ff99SB) force field [46] to study the dynamics of the DNA hairpin. The model for the DNA hairpin in the present study has the following effective Hamiltonian:

$$E_{\text{total}} = \sum_{\text{bonds}} K_r (r - r_{eq})^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2$$

$$+ \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

$$+ \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]. \quad (1)$$

The first term (sum over bonds) of Eq. (1) represents the energy between covalently bonded atoms, where r and r_{eq} are the interparticle separation and the equilibrium interparticle separation, respectively. This harmonic (ideal spring) term is a good approximation near the equilibrium bond length. The second term (sum over angles) represents the energy due to the geometry of electron orbitals involved in covalent bonding, where θ_{eq} corresponds to the equilibrium potential for the bending angle (θ) potential. K_r and K_θ are the spring constants associated with stretching and bending, respectively. The third term (sum over torsions) corresponds to the energy for twisting a bond due to bond order (e.g., double bonds) and neighboring bonds or lone pairs of electrons. One bond may have more than one of these terms, such that the total torsional energy is expressed as a Fourier series. V_n is a force constant, where n and γ are multiplicity and phase angle, respectively. The fourth term mimics the nonbonded energy between all atom pairs, which can be decomposed into van der Waals (first term of the summation) and electrostatic (second term of summation) energies. Here R_{ij} is the separation between ith and jth particles. $A = 4\epsilon\sigma^{12}$ and $B = 4\epsilon\sigma^6$, where σ and ϵ are the distance at which the intermolecular potential between the ith and jth particles is zero (van der Waals radius) and well depth, respectively. q_i and q_j are the electric charge on ith and jth particles, respectively. ϵ is absolute permittivity.

In the following, we describe the computational process leading to the DNA hairpin formation. We have taken eight different sequences of 16-base DNA oligomer strands in their folded hairpin configurations shown in Fig. 1, which contain a loop consisting of four homosequence nucleotides made up of either A or T and a stem of six base pairs. The first five base pairs remain the same in all sequences; however, the sixth one, which joins the loop to the stem, is different combinations of AT [Figs. 1(a) and 1(c)] and CG [Figs. 1(b) and 1(d)] for different loops of A and T. We highlight the distinguishing sequence features in the stem with the rectangular boxes in Fig. 1. DNA hairpins were built using the Nucleic Acid Builder (NAB) module in the Assisted Model Building with Energy Restraints (AMBER) suite of programs [45] and 3DNA server [47]. The hairpin structure has been merged with water. As pointed out earlier, simulations are significantly more computationally expensive, therefore, it is essential to reduce the computational complexity as much as possible. To achieve this, we used the triangulated water, where the angle between the hydrogen atoms is kept fixed. One such model is the TIP3P water model [48]. Using the LEaP module in AMBER, 15 Na+ (counterions) are added to the system to neutralize the negative charges on the DNA hairpin. The force F is applied perpendicular to the helical (x) direction at 5′-3′ ends at a temperature 300 K to study the force-induced melting of the DNA hairpin (Fig. 2). We have chosen the box dimension $100 \times 50 \times 50$ Å³ with three-dimensional periodic boundary conditions in such a way that the unzipped structure always remains inside the box. The unzipping happens in such a controlled manner that the end-to-end distance extends only along the force direction, i.e., in the x direction. Since the transverse fluctuations (in y and z directions) are much less than 50 Å, therefore, the DNA hairpin always remains inside the box in all simulations. The box contains 7211 water molecules, which are free to move inside the box along with ions.
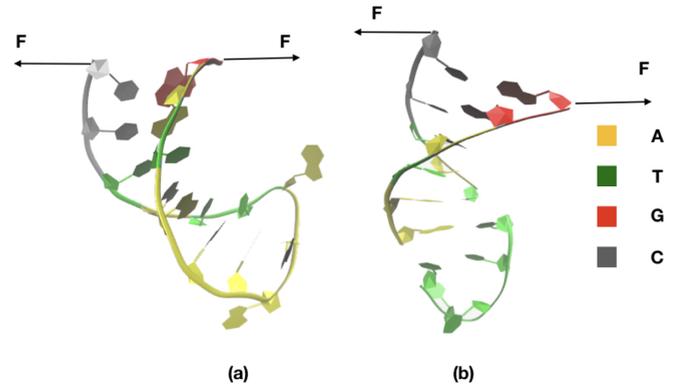
FIG. 2. The DNA hairpin in the folded state. Force has been applied perpendicular to the helix direction (x direction) at the opposite ends at a temperature of 300 K. Bases G, C, A, and T are shown in red, metal, yellow, and green, respectively. Forces were applied on two backbone atoms of the first base pair: the O5′ atom of the G1 residue at the 5′ end of one strand and on the O3′ atom of the C16 residue at the 3′ end in the simulations. Water molecules and counterions have not been shown here for clarity.

We used Particle Mesh Ewald (PME) method to calculate the electrostatic interactions [49,50] using the Cubic B-spline interpolation of fourth order, and a tolerance value of 10^{-5} Å is set for the direct space sum cutoff. A real space cutoff of 10 Å is used for both the nonbonded interactions and the electrostatics interactions. Initial energy minimization of the solvated systems was performed in two steps: in the first stage, we keep the DNA hairpin fixed and just minimize the positions of water molecules and ions. Then in the second stage, we minimize the entire system. After minimization of the energy, the system is gradually heated from 0 to 300 K. We used the Langevin thermostat to control the temperature with a collision frequency of 1.0 ps^{-1} [51–53]. For the zero force, the root mean square deviations (RMSDs) as a function of time are shown in Fig. 3 for different closing base pairs and loop sequences. One can notice that for all cases the system is well equilibrated within 3 ns. The deviation in RMSD due to loop and closing base pairs on the stability of the DNA
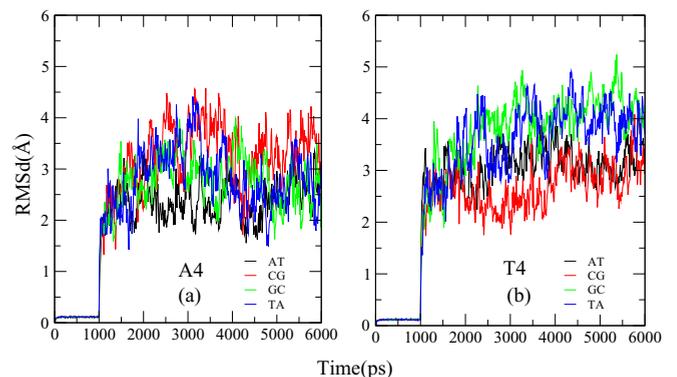
FIG. 3. The variation of RMSD as a function of time (a) for the A4 loop and (b) for the T4 loop of different closing base pairs. It is evident from these plots that the system is well equilibrated in a specific time.
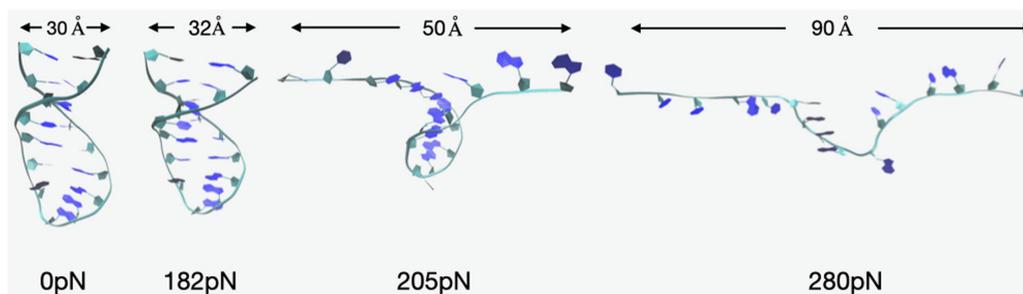
FIG. 4. Snapshots of some of the intermediate structures of the DNA hairpin at different forces 0 pN, 182 pN, 205 pN, and 280 pN and corresponding extension at $T = 300$ K for the sequence GGATAA$(A)_4$TTATCC [Fig. 1(c)] that forms during unzipping. These snapshots have been generated from the simulations using VMD. The force (not shown) has been applied along the x (perpendicular to helical) direction at $5'$-$3'$ ends. These snapshots provide atomistic descriptions of the unzipping processes at different stages. For example, we note that at extension $\sim$50 Å, half of the base pairs are open, whereas at $\sim$90 Å, all base pairs of the stem are open. We have not shown water molecules and counterions for clarity.

hairpin is also apparent from the plots. We have calculated the persistence length at zero force for all DNA hairpin sequences and found that the values lie in the range of 1.3–1.5 nm. These values are consistent with previous estimates [40].

III. UNZIPPING OF HAIRPINS: COMPUTATIONAL DETAIL

Temporal and spatial resolutions of molecular motion as the position of each atom including solvent molecules have been monitored within a picosecond-scale range. We used a constant force ensemble to study the unzipping of the DNA hairpin. For this, a force routine has been included in AMBER10 [54,55]. This allowed us to mimic the single-molecule experiments *in silico* and enabled us to visualize the biomolecular mechanics in action. Simulations have been carried out in isothermal-isobaric ensembles using a time step of 1 fs. A constant pressure is maintained by the isotropic position scaling with a reference pressure of 1 atm and a relaxation time of 2 ps. This procedure was repeated for the same starting structure, but with different (100) randomly assigned initial velocities to each atom.

The unzipping force depends linearly on the loading rate. For example, in the AFM experiments typical pulling rates range from 20 to 4000 pN/s. However, because of limitations in computer resources, the slowest stretching velocity used in the present simulations is 0.0001 pN fs^{-1}. However, to study force-induced melting of the DNA hairpin, the system has been first equilibrated for 5 ns at zero force then pulling simulation has been performed for 3 ns, where the external force started at 0 pN and increased linearly with time steps resulting in a total simulation time 8 ns for the AT/TA closing base pairs. However, for the GC/CG closing base pairs, the simulation is carried out for additional 2 ns. As a result, forcing protocol is several orders of magnitude faster than that used in AFM. Hence the magnitude of force required for the unzipping of the DNA hairpin would be larger compared to those observed experimentally [7,12,13,54].

In order to have further insight into the process, the deformations in the DNA hairpin of a particular sequence [Fig. 1(c)] at different forces have been monitored. In Fig. 4 we show some of the snapshots of conformations at different forces. The visual molecular dynamics (VMD) program was used for the visualization of trajectories and preparation of figures [56]. It can be seen from these plots that initially the DNA hairpin is in the zipped state. As force increases, the conformation of the hairpin remains in the zipped state up to $\sim$182 pN. Around 205 pN, there is a distortion in hydrogen bonds along the force direction. As a result, conformation of the folded state changed. Above the force $\sim$200 pN, hydrogen bonds start breaking, and around $\sim$280 pN, the hairpin unzips completely. It is interesting to note that at the extension $\sim$50 Å half of the base pairs get open, whereas around $\sim$90 Å all the base pairs of the stem are open. This is in accordance with previous results, where one observed about $\sim$7 Å extension per nucleotide (see Supplemental Material Ref. [57]). We note in some of the simulations that the extension exceeds the box size, but due to the periodic boundary conditions, the hairpin remains inside the box. In analyzing the data, it was ensured that the distance between the first nucleotide and the reentered end of the hairpin remains above the cutoff distance of van der Waals and Coulomb interactions (Fig. S1 of the Supplemental Material [57]).

IV. RESULTS AND DISCUSSION

In any experiment on systems of small size the outcome depends crucially on whether the control parameter is the force or the extension. For example, optical tweezers and AFM essentially control the position of the end monomer where a force is applied. On the other hand, magnetic tweezers setups provide a constant force at the end monomer. In a constant force ensemble (CFE) the control parameter is the average extension $\langle x \rangle$ [3]. Many biological reactions involve large conformational fluctuations that provide well-defined mechanical reaction coordinates. Therefore, choosing a suitable reaction coordinate is the first step in a free-energy profile. The applied force "tilts" the free-energy surface along the reaction coordinate by an amount linearly dependent on the end-to-end distance. Therefore, the end-to-end distance of a biopolymer has been used quite frequently to measure the progress of the reaction in all single-molecule experiments. Keeping this in view, we used CFE to obtain the force-extension (F-$\langle x \rangle$) curves for eight different sequences, which are compared in Fig. 5. The averaging has been performed over 100 independent simulations with different seeds.
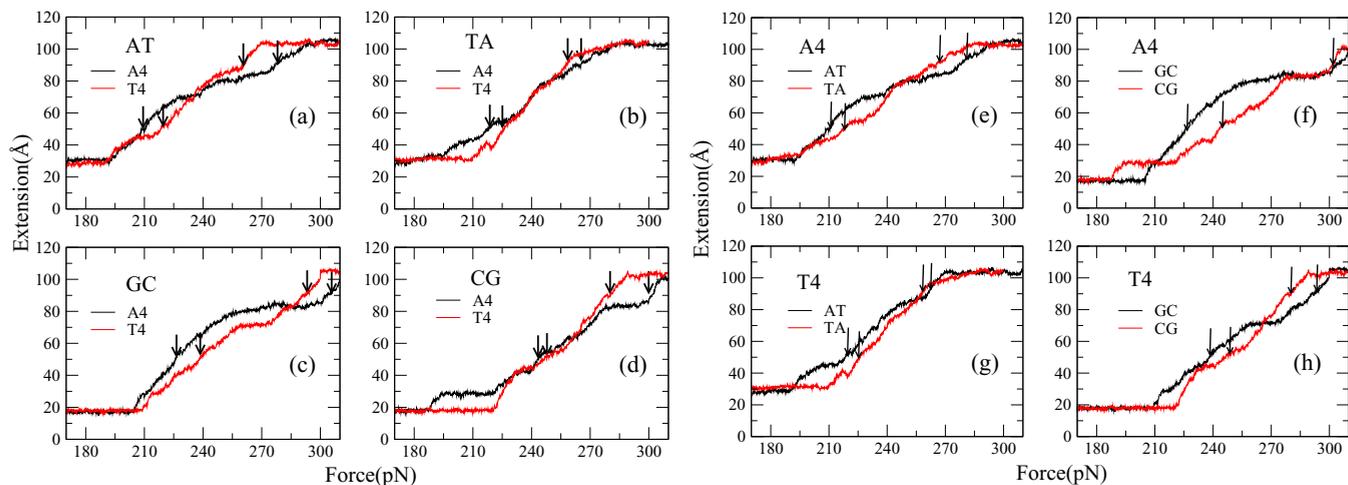
FIG. 5. (a)–(d) The force-extension curves for the DNA hairpin with the same closing base pair but of different loops. (e)–(h) The force-extension curve for the DNA hairpin with the same loop but different closing base pair. The extension has been averaged over 100 simulations with different seeds. Arrows are shown to indicate the extension at which half of the base pairs and all base pairs of the stem are open.

Some of the representative plots of the force-extension curves are shown in Figs. S2 and S3 of the Supplemental Material [57]. The total area under the curve F-$\langle x \rangle$ curve should be attributed to entropic, enthalpic, structural, and elastic contributions [38,39]. Thus keeping everything fixed, one can extract the difference in energy associated with the change in either loop sequence or closing base pair sequence. Keeping this in view, we have shown the F-$\langle x \rangle$ curves of hairpins having the same closing base pair (AT/TA, GC/CG) but of different loops in Figs. 5(a)–5(d). In Figs. 5(e)–5(h), we have

shown F-$\langle x \rangle$ curves of the same loop (A4/T4) to study the effect of different closing base pair (AT/TA, GC/CG). These curves show many interesting features. For example, one can notice that the loop sequence has a significant impact on the stability of hairpins. This is evident from Figs. 5(a)–5(d), where the force-extension curves of A4 are different from T4. This may be attributed to the stacking energy associated with adenine, which is absent in thymine [37–39]. In fact, the force-extension curve of thymine is entropic in nature and is well represented by the FJC model of polymer [3], whereas
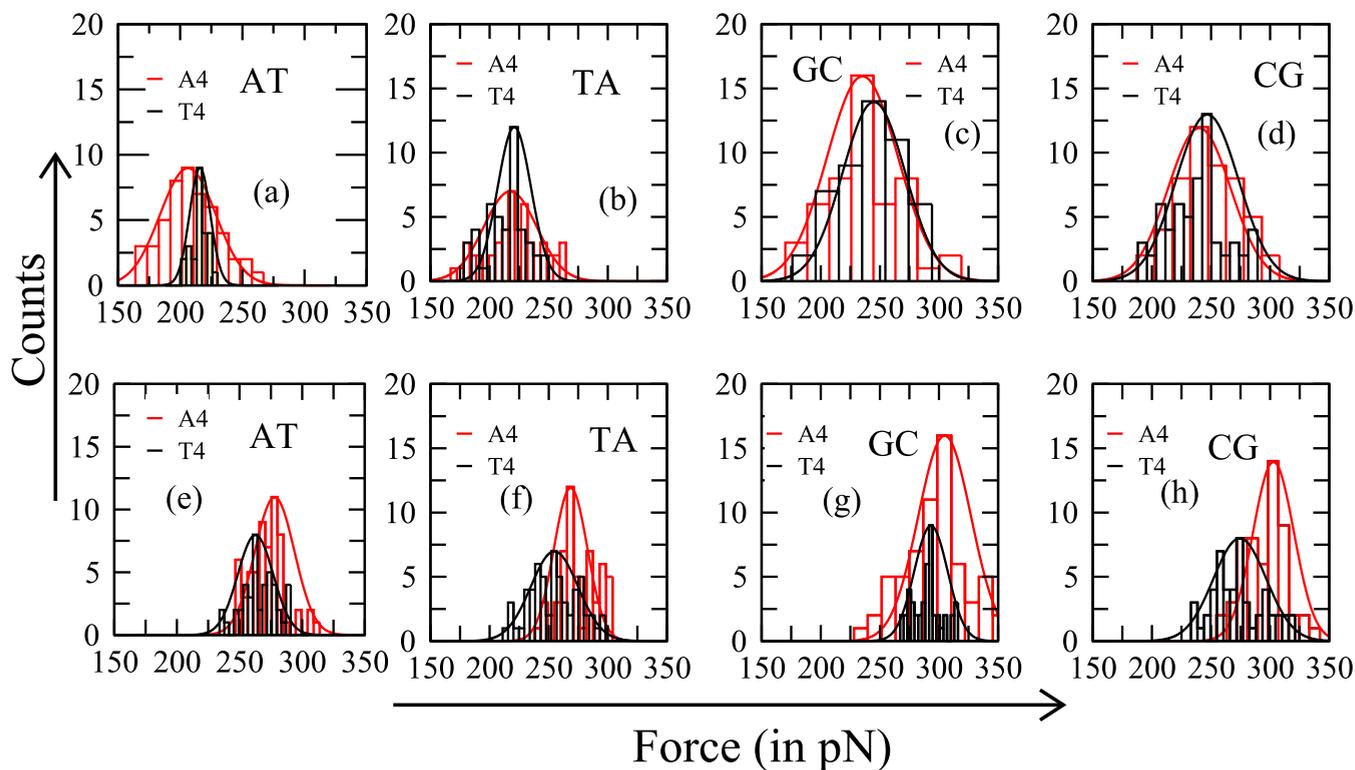
FIG. 6. Probability distribution of the unzipping force of the DNA hairpins with same closing base pair but different loops. The upper panels (a)–(d) correspond to F_h, whereas the lower panels (e)–(h) correspond to F_c.
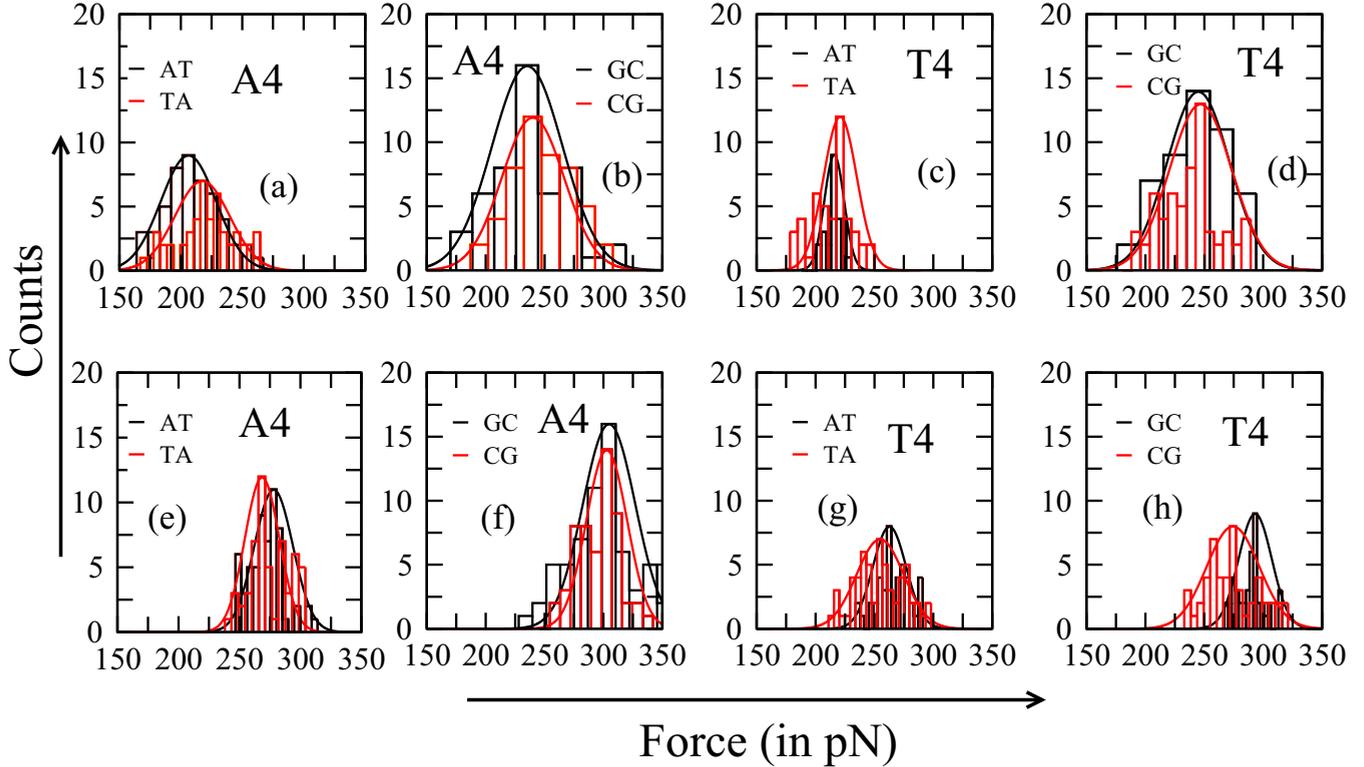
FIG. 7. Probability distribution of the unzipping force of the DNA hairpins for the same loop with different closing base pairs. The upper panels (a)–(d) correspond to F_h, whereas the lower panels (e–h) correspond to F_c.

the force-extension curve of adenine is enthalpic in nature and differs from the FJC.

Another interesting observation is that below 240 pN [Fig. 5(a)], the extension of A4 is higher than T4, whereas above this force, the extension of T4 is larger than A4. This implies that the inference drawn about the stability of the DNA hairpin based on the melting may differ from the force-induced melting. To have the semimicroscopic view of it, we use two definitions of the unzipping force: (1) force (F_h) at which half of base pairs are open similar to the DNA melting and (2) force (F_c) at which all the base pairs including closing base pairs are open. We note that this happens at ≈ 50 Å ± 3 Å and at ≈ 90 Å ± 3 Å for F_h and F_c, respectively. Furthermore, the force-extension curve usually shows two jumps. Here the first jump corresponds to the breaking of the hydrogen bond between the first base pair of the stem, whereas the second jump is associated with the gradual opening of base pairs. Most of the curves start from the zipped state, i.e., ≈ 20 Å for

GC/CG and ≈ 28 Å for AT/TA, and remain in the zipped state up to the certain applied force. This indicates that the nature of closing base pairs also affects the stability of the DNA hairpin even at zero force. In the case of the AT closing base pair, the unzipping force at which half of base pairs gets open for the A4 loop is ~ 208 pN, which is lower than the T4 loop (~ 215 pN). However, the unzipping force F_c for the A4 loop (~ 280 pN) is much higher than for the T4 loop (~ 260 pN). For the GC closing base pair, we observed that the unzipping force F_h (~ 242 pN) for the T4 loop is higher than the A4 loop (~ 235 pN). The unzipping force F_c (~ 290 pN) for T4 loop is lower than the A4 loop (~ 305 pN). In the case of the CG closing base pair, the unzipping forces F_h are ~ 240 pN and ~ 250 pN for the A4 and T4 loop, respectively. On the contrary, the unzipping forces F_c are found to be ~ 305 pN for the A4 loop, whereas ~ 270 pN for the T4 loop.

It has been reported in the literature that for certain DNA and RNA hairpin loops, a CG closing base pair provides

TABLE I. Unzipping force F_h and F_c of eight studied DNA hairpins for different cases. The error is found to be less than 3.5 pN for all cases [60,61].

	Closing base pair							
	AT		TA		GC		CG	
Loop	F_h (pN)	F_c (pN)	F_h (pN)	F_c (pN)	F_h (pN)	F_c (pN)	F_h (pN)	F_c (pN)
A4	205	280	215	270	235	305	240	300
T4	215	260	220	230	245	290	250	270
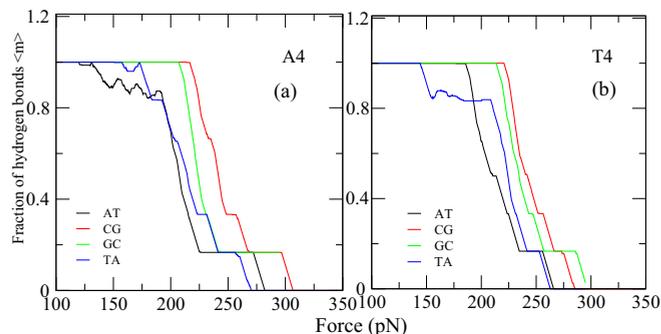
FIG. 8. Fraction of hydrogen bonds $\langle m \rangle$ as a function of the applied force for various loop closing base pairs: (a) loop consisting of adenine, (b) loop consisting of thymine.

enhanced stability in comparison to GC closing base pair [58,59]. For example, changing the closing base pair from CG to GC hairpins destabilized the structure by ~ 2 kcal mol^{-1} and caused a notable reduction in the melting temperature. However, such studies have been performed for a folded structure only, not for the unfolded structure. In Figs. 5(f) and 5(h), we have shown the comparative force-extension curves for GC and CG base pairs for the same loop. One can notice that there is a change in the unzipping force F_h for the A4 and T4 loop, which shows that CG is more stable than GC, similar to the one observed in the case of melting. The change in free energy is found to be nearly equal to ~ 2 kcal mol^{-1}. It is interesting to note that for the complete unzipping F_c for GC is larger than CG for the T4 loop [Fig. 5(h)]. It is pertinent to mention here that the stability data based on F_h [Figs. 5(e)–5(h)] are consistent with the findings of Vallone et al., whereas the F_c gives new insights into the stability of the DNA hairpin.

To further substantiate our findings, we show the histograms of unzipping force (F_c and F_h) of DNA hairpins with the same closing base pair but of different loops in Fig. 6, whereas in Fig. 7 we show the histograms of the unzipping force of DNA hairpins with the same loop but different closing base pairs. In Figs. 6 and 7, the upper row corresponds to F_h and the bottom row corresponds to F_c. For a given loading rate, the most probable unzipping force for the hairpin is obtained by the Gaussian fit of distribution of the unzipping force for every studied DNA conformation. For a given loading rate, the most probable unzipping force for the same closing base pair but with different loops and the same loop but with the different closing base pair are shown in Table I. Errors are estimated using the covariance matrix [60,61]. We observed that the error decreases as sample size increases, and for the present data set (100 simulations) the error is less than 3.5 pN for all cases.

Following the protocol of unzipping discussed earlier, we further substantiate these findings by extending our study to

obtain the fraction of hydrogen bonds as a function of the applied force for different closing base pairs in Fig. 8. In Figs. 8(a) and 8(b) we consider loops of A4 and T4, respectively for different closing base pairs. Here we also investigate two cases: (1) when half of the base pairs are open corresponding to DNA melting and (2) when all the base pairs are open including the closing ones. The values obtained from this method are also consistent with the one given in Table I.

V. CONCLUSION

We performed atomistic molecular dynamics simulation of eight designed DNA hairpins, which consist of a stem of six base pairs and a homosequence loop (A4 and T4) of four bases. The simulation provided an atomistic description of the stability of hairpin on the loop sequence and effects of closing the base pair that closes the loop. The histograms of unzipping force revealed the difference between F_h and F_c. The order of stability of eight hairpins can be concluded from these histograms, which have been tabulated in Table I. If the unzipping force (F_h) is defined at a force half of the base pairs are open similar to the definition of DNA melting, present simulations indicate that A4 is less stable than T4. The largest difference in F_h is 10 pN between the A4 loop and T4 loops, which is higher than the error bar, and single-molecule experiments will be able to observe experimentally. Regardless of the loop sequence, in terms of closing base pair, the order of stability is CG > GC > TA > AT, which is in agreement with Vallone et al. [31]. However, if F_c is taken as the direct measurement of the stability of the hairpin, the order of stability of eight DNA hairpins in terms of the loop sequence is A4 > T4. The difference in the stability arising due to the definition of F_c and F_h may be attributed to the competition between entropy associated with T4 loop and enthalpy (stacking energy) associated with stacked bases in the A4 loop. When only half of the base pairs are open, the free energy of the hairpin having the T4 loop is higher compared to the A4 loop. Whenever a loop opens (i.e., complete unzipping due to F_c) because of stacking energy associated with A4 its free energy is more than the T4 loop. Moreover in this case, the order of stability is found to be different than that of F_h, i.e., GC > CG > AT > TA. Our results open up the question of the stability of the DNA hairpin when subjected to an unzipping force. We hope that single-molecule experiments will be able to explore experimentally and resolve the stability issue.

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell* (Garland, New York, 1994).

[2] J. N. Israelachvili, *Intermolecular and Surface Forces*, 2nd ed. (Academic Press, London, 1992).

[3] S. Kumar and M. S. Li, Phys. Rep. **486**, 1 (2010).

[4] B. Essevaz-Roulet, U. Bockelmann, and F. Heslot, Proc. Natl. Acad. Sci. USA **94**, 11935 (1997).

[5] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot, Phys. Rev. Lett. **79**, 4489 (1997).

[6] G. U. Lee, L. A. Chrisey, and R. J. Colton, Science **266**, 771 (1994).

[7] T. Strunz, K. Oroszlan, R. Schafer, and H. J. G. Uentherodt, Proc. Natl. Acad. Sci. USA **96**, 11277 (1999).

[8] I. Schumakovitch, W. Grange, T. Strunz, P. Bertoncini, H. J. Gntherodt, and M. Hegner, Biophys. J. **82**, 517 (2002).

[9] C. Danilowicz, Y. Kafri, R. S. Conroy, V. W. Coljee, J. Weeks, and M. Prentiss, Phys. Rev. Lett. **93**, 078101 (2004).

[10] K. Hatch, C. Danilowicz, V. Coljee, and M. Prentiss, Phys. Rev. E **78**, 011920 (2008).

[11] C. Danilowicz, C. Limouse, K. Hatch, A. Conover, V. W. Coljee, N. Klecknerb, and M. Prentiss, Proc. Natl. Acad. Sci. USA **106**, 13196 (2009).

[12] F. K. Kueher, J. Morfill, R. A. Neher, K. Blank, and H. E. Gaub, Biophys. J. **92**, 2491 (2007).

[13] A. M. Naserrian-Nik, M. Tahani, and M. Karttunen, RSC Adv. **26**, 10516 (2013).

[14] A. P. Drabovich and S. N. Krylov, Anal. Chem. **78**, 2035 (2006).

[15] M. Orita, H. Iwahana, H. Kanazawa, K. Hayashi, and T. Sekiya, Proc. Natl. Acad. Sci. USA **86**, 2766 (1989).

[16] X. Dai, M. B. Greizerstein, K. Nadas-Chinni, and L. B. Roth-Man-Denes, Proc. Natl. Acad. Sci. USA **94**, 2174 (1997).

[17] A. M. Gacy and C. T. McMurray, Biochemistry **33**, 11951 (1994).

[18] C. T. McMurray, W. D. Wilson, and J. O. Douglass, Proc. Natl. Acad. Sci. USA **88**, 666 (1991).

[19] U. R. Muller and W. M. Fitch, Nature (London) **298**, 582 (1982).

[20] J. M. Zengel and L. J. Lindahl, J. Bacteriol. **178**, 2383 (1996).

[21] R. D. Wells, T. C. Goodman, W. Hillen, G. T. Horn, R. D. Klein, J. E. Larson, U. R. Müller, S. K. Neuendorf, N. Panayotatos, and S. M. Stirdivant, Prog. Nucleic Acid Res. Mol. Biol. **24**, 167 (1980).

[22] J. J. Tyson, K. C. Chen, M. Lederman, and R. C. J. Bates, Theor. Biol. **144**, 155 (1990).

[23] M. Rosenberg and D. Court, Annu. Rev. Genet **13**, 319 (1979).

[24] G. Hobom, R. Grosschedl, M. Lusky, G. Scherer, E. Schwarz, and H. Kossel, Cold Spring Harb. Symp. Quant. **43**, 165 (1979).

[25] A. Gonzalez, A. Talavera, J. M. Almendral, and E. Vinuela, Nucleic Acids Res. **14**, 6835 (1986).

[26] K. C. Chen, J. J. Tyson, M. Lederman, E. R. Stout, and R. C. Bates, J. Mol. Biol. **208**, 283 (1989).

[27] C. L. Chan, D. Wang, and R. J. Landick, Mol. Biol. **268**, 54 (1997).

[28] S. Svoboda and A. Di Cara, Cell. Mol. Life Sci. **63**, 901 (2006).

[29] G. Portella and M. Orozco, Angew. Chemie **122**, 7673 (2010).

[30] N. L. Goddard, G. Bonnet, O. Krichevsky, and A. Libchaber, Phys. Rev. Lett. **85**, 2400 (2000).

[31] P. M. Vallone, T. M. Paner, J. Hilario, M. J. Lane, B. D. Faldasz, and A. S. Benight, Biopolymers **50**, 425 (1999).

[32] L. A. Marky and K. J. Breslauer, Biopolymers **26**, 1601 (1987).

[33] M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985).

[34] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, Science **276**, 1109 (1997).

[35] S. Kumar and G. Mishra, Phys. Rev. Lett. **110**, 258102 (2013).

[36] D. Marenduzzo, S. M. Bhattacharjee, A. Maritan, E. Orlandini, and F. Seno, Phys. Rev. Lett. **88**, 028102 (2001).

[37] Y. Seol, G.M. Skinner, K. Visscher, A. Buhot, and A. Halperin, Phys. Rev. Lett. **98**, 158103 (2007).

[38] C. Ke, M. Humeniuk, H. S-Gracz, and P. E. Marszalek, Phys. Rev. Lett. **99**, 018302 (2007).

[39] G. Mishra, D. Giri, and S. Kumar, Phys. Rev. E **79**, 031930 (2009).

[40] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, and S. M. Block, Proc. Natl. Acad. Sci. USA **103**, 6191 (2006).

[41] A. Montanari and M. Mezard, Phys. Rev. Lett. **86**, 2178 (2001).

[42] H. Zhou, Y. Zhang, and Z.-C. Ou-Yang, Phys. Rev. Lett. **86**, 356 (2001).

[43] T. Hugel, M. Grosholz, H. Clausen-Schaumann, A. Pfau, G. Hermann, and M. Seitz, Macromolecules **34**, 1039 (2001).

[44] G. Mishra, D. Giri, M. S. Li, and S. Kumar, J. Chem. Phys. **135**, 035102 (2001).

[45] D. A. Case *et al.*, AMBER 10, University of California, San Francisco (2008).

[46] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman, J. Comput. Chem. **24**, 1999 (2003).

[47] X.-J, Lu and W. K. Olson, Nucleic Acids Res. **31**, 5108 (2003).

[48] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, J. Chem. Phys. **79**, 926 (1983).

[49] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).

[50] U. Essmann, L. Perera, and M. L. Berkowitz, J. Chem. Phys. **103**, 8577 (1995).

[51] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science, Oxford, 1987).

[52] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, London, 2002).

[53] J. Schluttig, M. Bachmann, and W. Janke, J. Comput. Chem. **29**, 2603 (2008).

[54] M. Santosh and P. K. Maiti, J. Phys: Cond. Matt. **21**, 034113 (2009).

[55] S. Izrailev, *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Lecture Notes in Computational Science and Engineering, Vol. 4, edited by P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel (Springer, Berlin, 1999).

[56] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).

[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.103.062411 for more details about the data used in figures.

[58] E. M. Moody and P. C. Bevilacqua, J. Am. Chem. Soc. **125**, 2032 (2003).

[59] S. Kannan and M. Zacharias, Nucl. Acid. Res. **39**, 8271 (2011).

[60] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed. (McGraw-Hill, New York, 1991).

[61] C. Ray, J. R. Brown, and B. B. Akhremitchev, Langmuir **23**, 6076 (2007).