# The Role of Loop Stacking in the Dynamics of DNA Hairpin Formation

Majid Mosayebi,*,[†] Flavio Romano,[†] Thomas E. Ouldridge,[‡] Ard A. Louis,[‡] and Jonathan P. K. Doye*,[†]

[†]Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

[‡]Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford OX1 3NP, United Kingdom

ABSTRACT: We study the dynamics of DNA hairpin formation using oxDNA, a nucleotide-level coarse-grained model of DNA. In particular, we explore the effects of the loop stacking interactions and non-native base pairing on the hairpin closing times. We find a nonmonotonic variation of the hairpin closing time with temperature, in agreement with the experimental work of Wallace et al. (*Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5584−5589). The hairpin closing process involves the formation of an initial nucleus of one or two bonds in the stem followed by a rapid zippering of the stem. At high temperatures, typically above the hairpin melting temperature, an effective negative activation enthalpy is observed because the nucleus has a lower enthalpy than the open state. By contrast, at low temperatures, the activation enthalpy becomes positive mainly due to the increasing energetic cost of bending a loop that becomes increasingly highly stacked as the temperature decreases. We show that stacking must be very strong to induce this experimentally observed behavior, and that the existence of just a few weak stacking points along the loop can substantially suppress it. Non-native base pairs are observed to have only a small effect, slightly accelerating hairpin formation.
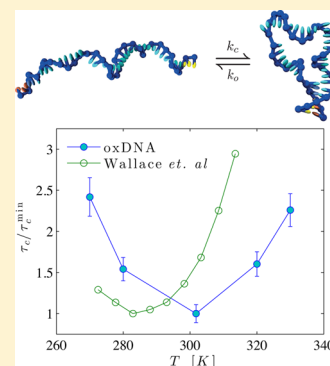
INTRODUCTION

Nucleic acid hairpins have diverse biological functions. For example, DNA hairpins play important roles in gene expression, DNA recombination, and transposition.[1−4] In RNA, hairpins are a common secondary structure motif and can serve as nucleating sites for higher order RNA structures.[5] In addition, hairpins are commonly used in DNA-based nanotechnology, for instance, as fuels in motors,[6,7] as biosensors,[8] or for controlling self-assembly pathways.[9] Together with duplex hybridization, hairpin formation is one of the most basic dynamical processes involving nucleic acids, and therefore, a fundamental understanding of this process is of wide-ranging importance.

There have been many experimental and theoretical studies on the effects of different parameters such as the loop length, the sequence, and the ionic strength of the solvent on the dynamics and thermodynamics of nucleic acid hairpins.[10−20] DNA hairpin formation is considered to be predominantly a two-state process with a rate-limiting step which involves loop closure and a subsequent rapid zippering of the stem.[21] This hypothesis has been challenged by several recent studies that suggest a more complex pathway, involving intermediate states containing misfolded or partially folded configurations.[22−24] Additionally, there have been diverse, and sometimes contradictory, experimental observations and interpretations of hairpin formation kinetics. Fluorescence energy transfer and fluorescence correlation spectroscopy measurements[25] have reported an Arrhenius temperature dependency with a positive activation enthalpy for the hairpin closing rates. Other studies[10,26,27] have reported a nonmonotonic variation of the

closing time with temperature and an effective activation enthalpy that switches sign from positive to negative close to the hairpin melting temperature. The reason for the positive activation enthalpy at low temperature has been posited to be the slow configurational diffusion of the single-stranded DNA (ssDNA) due to the intrachain interactions.[10,21,26] For example, in ref 21, it was argued that the intrachain interactions, including non-native Watson−Crick (WC) bonds and also misstacked bases in the loop, increase the ruggedness of the free-energy surface at the early stages of hairpin formation, which in turn decreases the configurational diffusion coefficient and makes it significantly temperature dependent. An alternative explanation is that the positive activation enthalpy is a result of an increase in the ssDNA stiffness due to increasingly strong stacking in the loop as the temperature is decreased.[25] Evidence for this position comes also from experiments on the possible effects of the loop sequence on hairpin thermodynamics and dynamics.[11] These experiments found that the melting temperature of a hairpin with a more strongly stacking poly(dA) loop is lower than that for the equivalent hairpin with a poly(dT) loop, and that the hairpin closing times are longer for the hairpin with the poly(dA) loop.

To investigate in detail the effects of base stacking and non-native WC base pairing (i.e., misbonding) on the stability and dynamics of hairpins, we have performed extensive simulations of the kinetics and thermodynamics of the hairpin studied by

Wallace et al.,[26] using oxDNA,[28−30] a coarse-grained model at the nucleotide level that has been shown to be capable of describing single- and double-stranded DNA[31] and basic processes such as hybridization[32] and toe-hold mediated strand displacement.[33] Importantly, oxDNA incorporates single-stranded stacking interactions and allows for the formation of non-native base pairs (WC base pairs that are not intended to be present in the stem). Coarse-grained modeling of hairpin formation has been studied previously in refs 14, 16, 17, and 31 for DNA and in refs 19, 20, and 34 for RNA, and has considered, for example, mechanical unzipping of hairpins[19] and the scaling of the zipping up time of a stem with stem length,[17] but none of these studies has probed possible nonmonotonicity in closing rates due to loop properties. In this paper, we first investigate whether oxDNA reproduces the nonmonotonic dependence of hairpin closing times on temperature,[26] and explore the role of stacking and non-native base pairs in the observed behavior. We then consider perturbed stacking strengths to establish the sensitivity of hairpin closing times to this physical parameter.

■ METHODS

Coarse-Grained Model. In oxDNA, each nucleotide is modeled as a rigid body with three interaction sites (see Figure 1). The interaction potential consists of terms representing the
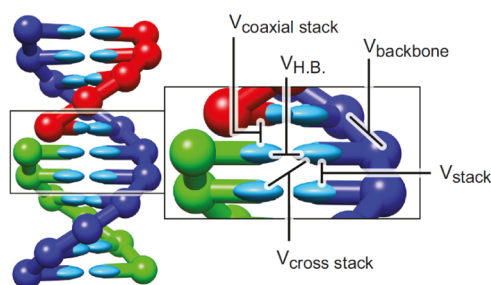
Figure 1. Illustration of a DNA double helix in the oxDNA model and the different interaction terms that stabilize the structure. The bases are represented by cyan ellipsoids and the backbone sites by spheres.

backbone connectivity, excluded volume, hydrogen bonding between WC complementary base pairs, stacking between adjacent bases along the chain, coaxial stacking between nonadjacent bases, and cross stacking. Aside from backbone connectivity and excluded volume, all interactions are anisotropic, depending on the relative orientation of the nucleotides. The interactions are designed to favor the formation of a right-handed double helix at low temperature. Orientational modulations of the stacking potential encourage the bases to form coplanar stacks, and hydrogen bonding can occur between complementary WC base pairs when they are antialigned, leading to the formation of double helical structures for which the helical twist arises from the different length scales of the backbone separation and the optimal stacking separation. It should be noted that oxDNA neglects noncanonical base pairing interactions, e.g., non-AT/GC interactions between the WC edges of bases, and interactions involving other edges such as Hoogsteen pairing and bonds with the sugar edge.[35]

The oxDNA model has been previously described in detail[28−30] and is implemented in a simulation package which is available for download.[36] The model has been parametrized for a salt concentration of 0.5 M, where the Debye screening length is short and it is reasonable to incorporate the electrostatic interactions in a soft excluded volume. This is the regime where most DNA nanotechnological experiments are carried out. Here we use the sequence-dependent parametrization of the model,[30] where the strengths of hydrogen bonding and stacking interactions depend on the identities of the interacting bases.

OxDNA is particularly suited to the present study, as it has been designed to reproduce the thermodynamic and mechanical properties of both single- and double-stranded DNA.[29,30] As oxDNA incorporates single-stranded stacking interactions, and allows for the formation of non-native base pairs, simulating the model allows us to explore the consequences of these physical aspects of DNA for hairpin formation kinetics. In addition, oxDNA has been shown to accurately reproduce the most important aspects of the kinetics of DNA duplex hybridization[32] and toehold-mediated DNA strand displacement.[33] OxDNA has also been successfully applied to study a variety of DNA biophysical properties such as duplex overstretching[37] and plectoneme formation,[38] as well as DNA nanotechnology systems, such as nanotweezers[39] and walkers.[40]

Simulation Methods. We perform dynamics simulations, where the mass, energy, and length units are chosen to be $m_0 = 315.75$ Da, $l_0 = 8.518$ Å, and $\epsilon_0 = 4.142 \times 10^{-20}$ J, respectively, implying a time unit of $\tau_0 = (m_0 l_0^2 / \epsilon_0)^{-1/2} = 3.03 \times 10^{-12}$ s. Each nucleotide has a mass of m_0 and a moment of inertia of $I = 0.138\ m_0 l_0^2$. We use a weakly coupled Andersen-like thermostat as implemented by Russo et al.:[41] the system is evolved by solving Newton's equation of motion using a velocity Verlet algorithm for ~100 steps with an integration time step of $0.005\tau_0$, and then, the velocities and angular velocities of each particle are updated from a Maxwell−Boltzmann distribution at the given temperature T with probabilities $p_v = 0.02$ and $p_\omega = 0.0067$, respectively. On time scales much longer than 5000 steps, the thermostat produces diffusive motion.

We study the kinetics of hairpin formation by performing brute-force closing simulations, as well as more complex simulations using the forward flux sampling (FFS)[42] method, which is a rare event method for accelerating kinetic measurements. It should be noted that, as is common for most coarse-grained models, it is not straightforward to map absolute time scales in our model onto experiment.[31] We expect, however, the relative time scales of similar processes in oxDNA to be directly comparable with experiments, as has been shown recently for duplex hybridization[18,32] and strand displacement.[33]

We obtain free-energy profiles using the virtual move Monte Carlo (VMMC) algorithm of Whitelam and Geissler.[43] We have found that VMMC greatly improves the equilibration speed in oxDNA.[28] In addition, in order to efficiently sample the free-energy landscapes of these hairpin systems, we use the umbrella sampling technique.[44] Umbrella sampling allows the system to overcome free-energy barriers by artificially biasing the system to sample states with higher free energy more frequently.

Base stacking and hydrogen bonding between complementary base pairs are the two crucial interactions that cause DNA to behave differently from a normal polymer chain. Here, to investigate the role of stacking interactions on the dynamics of hairpin formation, we perform additional simulations in which all stacking interactions in the loop are scaled by a factor λ_{st} relative to their (sequence-dependent) values from ref 30; $\lambda_{st} =$

14327

dx.doi.org/10.1021/jp510061f | J. Phys. Chem. B 2014, 118, 14326−14335

1 therefore corresponds to the canonical values quoted in ref 30. To clarify the role of misbonding, we also simulate hairpins in which the complementary hydrogen-bonding interactions are switched off either (i) completely for all non-native base pairs or (ii) just for non-native base pairs where at least one base belongs to the loop. In the latter case, the hairpin can still form misbonded base pairs in the stem.

■ RESULTS

We initially consider two hairpins, each with a 30-base loop and a 5-base-pair stem. The sequences that we study are

- (S_1) 3′-CCCAA(A)$_{30}$TTGGG-5′
- (S_2) 3′-CGCTA(A)$_{30}$TAGCG-5′

in which the sequence S_1 is chosen to be the same as the hairpin studied by Wallace et al.[26] To study the effect of misbonding on the kinetics, we also consider the second hairpin S_2, that is very similar to S_1 but with a slightly altered sequence in order to reduce the amount of misbonding between the two stems. While the hairpin S_1 is able to make eight different stem–stem misbonded base pairs, there are only two misbonding possibilities available to S_2.

Hairpin Closing Time. Figure 2 illustrates typical configurations for the open, closed, and misbonded states as well as the temporal evolution of the number of base pairs (a base pair being defined by a hydrogen-bonding energy of less than -0.596 kcal mol^{-1}, where the typical hydrogen-bonding energy for a base pair is ∼6 times this threshold), and the end-to-end distance R_{ee} for the hairpin S_1. The open and closed states can be clearly distinguished. Furthermore, the transition from one state to the other is very fast compared to the time scale that the hairpin spends in each of the states.

For hairpin formation, we define the average closing time τ_c as the average time it takes for an open hairpin to form all its native base pairs in the stem for the first time. The open hairpin belongs to an ensemble of equilibrated configurations with no base pairs. The distribution of closing times for the S_1 hairpin at $T = 280$ K is shown in Figure 3a. The distribution follows an exponential form, with a characteristic time scale that matches the average closing time τ_c of the hairpin at $T = 280$ K. The single exponential distribution that we observe for τ_c is typical of two-state reactions, where the dynamics are governed by transitions between two well-defined states (i.e., the open and closed states) and for which there exists a well-defined closing rate constant $k_c = 1/\tau_c$.

The kinetics of simple two-state reactions are expected to follow the Arrhenius law, in which the reaction rate constant k_c is exponentially related to a temperature-independent activation enthalpy H_a through the relation $k_c \propto \exp(-H_a/k_B T)$, where k_B is the Boltzmann constant. Figure 3b shows the closing time for the S_1 hairpin as a function of temperature (filled circles). This hairpin forms most efficiently at $T \approx 300$ K. τ_c has a minimum at this temperature and increases on both raising or lowering the temperature. The existence of this minimum clearly shows that the formation of this hairpin is a non-Arrhenius process with an apparent activation enthalpy $H_a = d(\ln \tau_c)/d(1/T)$ that changes sign at $T \approx 300$ K and becomes larger in magnitude as the temperature deviates more from $T \approx 300$ K. This observation is in qualitative agreement with the experimental measurements of ref 26 on the same hairpin (see Figure 3b). Quantitative agreement with the experiment of ref 26 is difficult to achieve because oxDNA is parametrized at a higher salt concentration, and the rates are sensitive to this. Nevertheless,

(a)

(b)

(c)

(d)
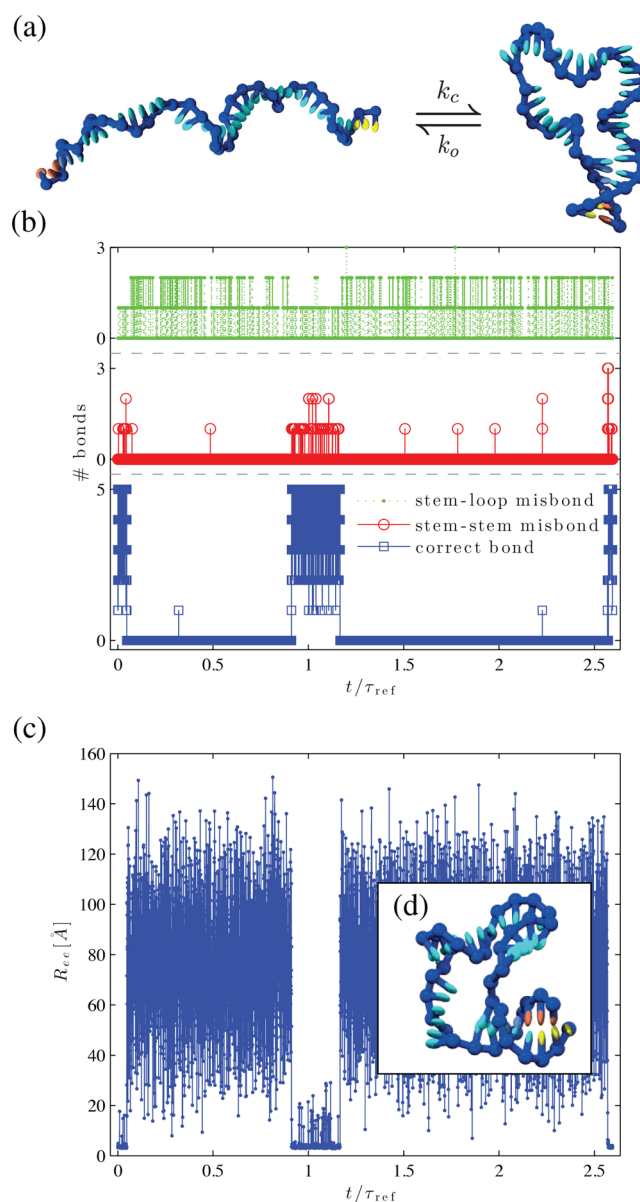
Figure 2. (a) Typical snapshots of the S_1 hairpin at $T = 300$ K in open and closed states. (b) Number of base pairs and (c) the end-to-end distance as a function of time at $T = 310$ K. Two different closing pathways occur. In the first closing event at $t/\tau_{ref} \approx 0.91$, the hairpin folds by first forming correct base pairs in the stem, while in the second at $t/\tau_{ref} \approx 2.57$ an initial stable closed loop is formed by misbonded base pairs which then rearrange to form a complete hairpin. (d) S_1 hairpin with two stem–stem misbonds and two stem–loop misbonds. In parts b and c, the time is normalized with respect to τ_{ref}, the average closing time of the S_1 hairpin at $T = 300$ K.

the changes we measure in our relative rates are of a similar order to those measured in the experiments.[10,26,27,45] In addition, the overall shift of the oxDNA curve to higher temperatures compared to the experimental curve may in part reflect the further stabilization of the S_1 hairpin at higher salt concentrations. For example, the relative melting temperatures are 305 K at 0.5 M in oxDNA and 291.6 K at 0.1 M in the experiment. We note that duplex hybridization in the oxDNA model also exhibits non-Arrhenius behavior, where the apparent activation enthalpy, although always negative, increases in magnitude with increasing temperature.[32]
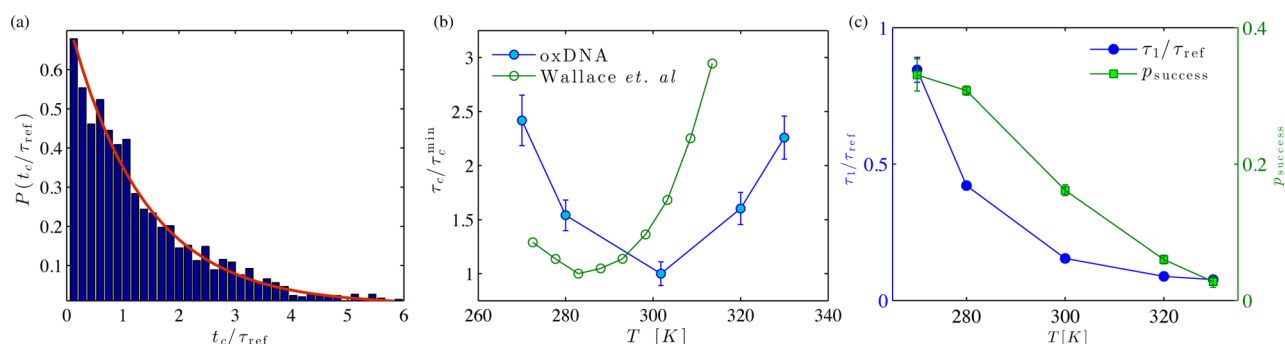
Figure 3. (a) Distribution of closing times t_c/τ_{ref} for the S_1 hairpin at $T = 280$ K. Closing times are normalized with respect to τ_{ref}, the average closing time of the S_1 hairpin at $T = 300$ K. The solid line is an exponential function $\tau_c^{-1} \exp(-t_c/\tau_c)$. (b) Normalized closing time τ_c as a function of temperature for the S_1 hairpin simulated with oxDNA at a salt concentration of 0.5 M (filled circles), and the experimental results of ref 26 for the same hairpin but at a salt concentration of 0.1 M (open circles). (c) First contact formation time τ_1 and success probability $P_{success}$ for the S_1 hairpin, where $P_{success}$ is defined as the probability that a loop with one base pair leads to a fully formed hairpin before opening.
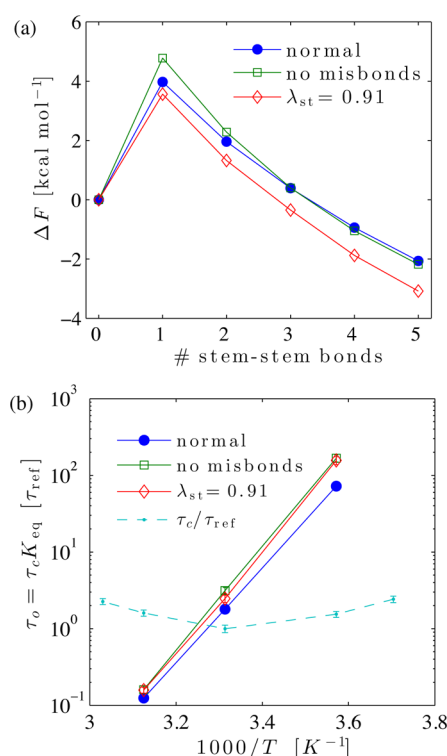
Figure 4. (a) Free-energy profile as a function of the number of stem–stem base pairs for the S_1 hairpin at $T = 280$ K. ΔF is measured relative to the open state with no stem–stem bonds. (b) Arrhenius plots of the opening time (solid lines) and closing time (dashed line) for the S_1 hairpin. τ_{ref} is the average closing time of the S_1 hairpin at $T = 300$ K. In both parts a and b, results are compared to when misbonding is switched off and when the stacking in the loop is reduced.

To better understand the non-Arrhenius behavior associated with hairpin formation, we write the closing time as $\tau_c \approx \tau_1/P_{success}$, where τ_1 is the average time it takes to form the first stem–stem base pair with a hydrogen-bonding energy of less than -0.596 kcal mol^{-1} and $P_{success}$ is the probability of successful formation of the hairpin (i.e., with all native base pairs in the stem) starting from a state with one stem–stem base pair before returning back to the open state. The values of τ_1 and $P_{success}$ are obtained using the FFS technique.[42] We note that the basic assumption in the application of FFS to measure rates of hairpin formation is that the zippering of the stem is

much faster than the loop closure. This assumption is valid for the hairpin studied here, since τ_c computed with brute-force simulations that do include the zippering time agrees with the τ_c obtained with the FFS method. Figure 3c shows that the success probability $P_{success}$ is significantly reduced when the temperature is raised. At high temperatures, a single base pair is not sufficiently stable to ensure formation of the complete hairpin and most of the times the loop opens up before zippering of the rest of the stem occurs. The rate limiting step for the formation of the hairpin at those temperatures therefore involves the search for a state which has on average more than one base pair. Although hairpin formation is predominantly a two-state process, it is characterized by a complex set of transition states that have on average a larger number of base pairs at higher temperatures. As these transition states are enthalpically more stable (due to the base pairing) than the open state, a negative activation enthalpy is observed at high temperatures (where τ_1 is relatively temperature independent) whose magnitude increases with increasing temperature. Negative activation enthalpies were previously observed for hybridization of oxDNA duplexes for similar reasons.[32]

On lowering the temperature, $P_{success}$ increases until it reaches a plateau at very low temperatures. In this regime, the temperature dependence is dominated by the change in the average time it takes to form the first stem–stem base pair (i.e., τ_1). The rapid increase of τ_1 with decreasing temperature causes the closing time to pass through a minimum and then increase again. Thus, the apparent activation enthalpy becomes positive at low temperatures. The positive activation enthalpy can be caused by variation in the enthalpy of the open state as well as the enthalpy of transition state ensemble. We will explore the microscopic origins of the low-temperature rise of τ_c later.

Thermodynamics of Hairpin Formation. To gain further insight into the mechanism of hairpin formation, we compute the free energies of the hairpin using the umbrella sampling technique. In Figure 4a, we plot the free energy as a function of the number of stem–stem bonds at $T = 280$ K, showing that the S_1 hairpin is more stable than the open state by ~2 kcal mol^{-1} at this temperature.

For the hairpins that are studied here, the opening rates vary by several orders of magnitude within the temperature range of our study. At low temperatures, the hairpin opening process is slower than it could be conveniently measured with a brute-force method. Therefore, we calculate the opening time with the help of the free-energy profiles that we obtain with umbrella
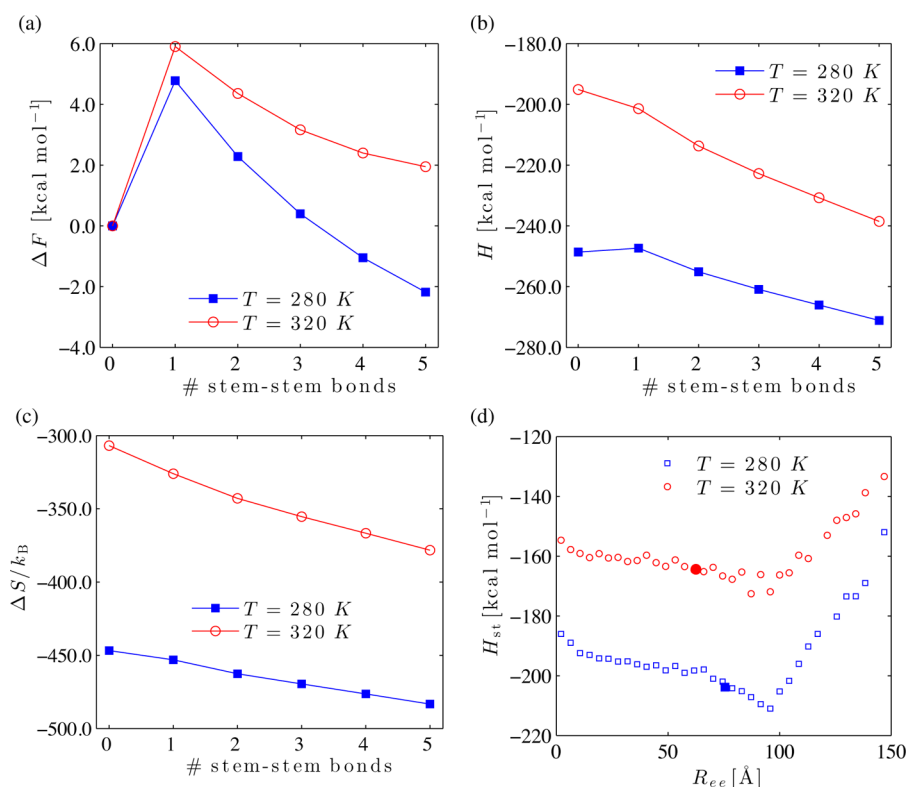
Figure 5. (a) Free energy, (b) enthalpy, and (c) entropy of the S_1 hairpin as a function of the number of bonds in the stem at $T = 280$ K (solid squares) and at $T = 320$ K (open circles) when misbonding is not permitted. (d) Stacking enthalpy H_{st} as a function of end-to-end distance R_{ee} for a 30-base poly(dA) ssDNA at two temperatures. Filled symbols mark the respective mean values of $\langle H_{st} \rangle$ and $\langle R_{ee} \rangle$.

sampling. In particular, we calculate the average opening time τ_o from the relation $K_{eq} = \tau_o/\tau_c$, where the equilibrium constant K_{eq} is the ratio of the partition functions of the closed (with at least one stem−stem base pair) and open states. The ratio can readily be obtained from free-energy profiles, such as those illustrated in Figure 4a. The Arrhenius plot in Figure 4b illustrates that, in contrast to hairpin formation, the opening of the S_1 hairpin exhibits Arrhenius behavior in the temperature range of our study.

We showed that the apparent activation enthalpy of the formation of the S_1 hairpin inferred from the kinetic data is positive at low temperatures, implying that the enthalpy of the open state is lower than the enthalpy of the transition state. This is in contrast to the apparent negative activation enthalpy at high temperatures. In Figure 5b, we decompose the free-energy profiles at high and low temperatures into their enthalpic and entropic components when misbonding is not allowed. Misbonding is forbidden to simplify the analysis. We observe that the enthalpy difference between the state with one base pair and the open state is positive at $T = 280$ K, $\Delta H = H(1) - H(0) \approx 1.28$ kcal mol^{-1}. While the sign of ΔH is in agreement with our kinetic data, the apparent activation enthalpy $H_a = (\mathrm{d} \ln \tau_c)/(\mathrm{d}(1/T)) \approx 3.33$ kcal mol^{-1} is somewhat larger than ΔH. This apparent discrepancy probably arises because the number of stem−stem base pairs is not a perfect reaction coordinate; i.e., the transition state for hairpin formation at low temperatures does not coincide with an equilibrium population of states having one base pair by our energy criterion.

We can also see from Figure 5b that, at $T = 320$ K, the enthalpy difference $\Delta H \approx -6.28$ kcal mol^{-1} is less negative than the apparent activation enthalpy $H_a \approx -7.19$ kcal mol^{-1},

implying that the transition state has on average more than one base pair in this regime. Consistent with this argument, the success probability after forming one base pair is very small at $T = 320$ K (Figure 3c).

Hairpin loop closure of course involves a large loss of conformational entropy, but there is also an enthalpic cost to bringing the two ends of the chain together, because states with small R_{ee} are on average less stacked than those typical of the open state. From Figure 5d, we can see that this enthalpy cost is greater at low temperatures, when the bases are more strongly stacked and the loop is less flexible. Consistent with our results, a positive activation enthalpy, which increases with decreasing temperature, has been reported experimentally for the end-to-end collisions of poly(dA) strands.[46] Breaking of stacking interactions to bring the two ends of the strand together is therefore the origin of the positive apparent activation enthalpy at low temperatures. Note that in this way the ssDNA in the loop region behaves differently from a freely jointed chain, a commonly used model for ssDNA, for which the cost of bringing the chain ends together would be purely entropic.

Role of Misbonding. Non-native base pairs have the potential both to enhance and to hinder the formation of a target hairpin. They can provide alternative pathways where misbonded base pairs form first, followed by an internal rearrangement to form the correct structure. Whether such pathways enhance the rate of hairpin formation depends on the ease with which misbonded structures can be resolved. For short duplexes, it has been previously shown that misbonding enhances the rate of hybridization because the time scale for internal displacement reactions that allow the correct duplex to be achieved after misbonding are much faster than the diffusion time associated with base-pair forming encounters at typical
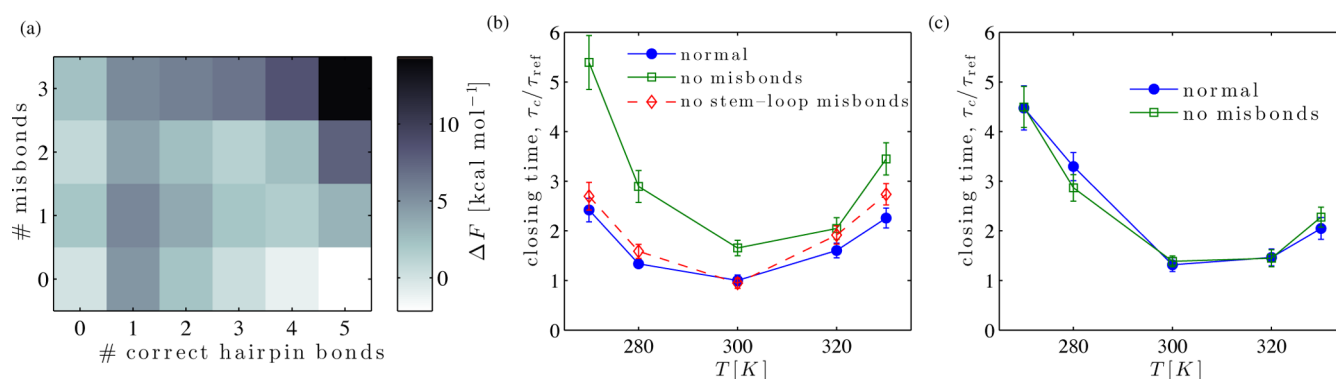
Figure 6. (a) Two-dimensional free-energy landscape showing the free energy of the S_1 hairpin at $T = 280$ K as a function of the number of misbonds and the number of correct hairpin bonds. ΔF is measured relative to the open state with no base pairs. The average closing time of (b) the S_1 hairpin and (c) the S_2 hairpin as a function of the temperature. In parts b and c, the results are compared to those when misbonding is not permitted. All times are normalized with respect to τ_{ref}, the average closing time of the S_1 hairpin at $T = 300$ K.

(low) strand concentrations.[32] However, whether a similar time scale separation holds for hairpin formation is less obvious because of the unimolecular nature of the process, and is likely to depend on the structure of the hairpin (e.g., loop length), sequence, and parameters such as temperature that affect the ease with which free-energy barriers can be overcome.

A two-dimensional free-energy landscape as a function of the number of correct bonds in the stem and the total number of misbonds is plotted for the S_1 hairpin at $T = 280$ K in Figure 6a. The profile shows two minima corresponding to the fully open and closed states. In addition, we observe that various misbonded configurations are possible. In particular, there is a local minimum involving two non-native WC base pairs (and no native base pairs). Moreover, the apparent free-energy barrier separating such a configuration from the fully formed hairpin is actually smaller than the barrier of the direct path from the fully open state to the fully formed hairpin. It is thus plausible that folding of the hairpin can occur through an alternative pathway initially involving non-native base pairs.

In kinetic measurements of the hairpin S_1, which has a short stem and a relatively long loop, we have considered the effects of completely switching off WC misbonding and only switching off WC misbonding between bases in the stem and in the loop. Figure 6b shows that the closing time becomes longer, particularly at low temperatures, when misbonding is not allowed for the S_1 hairpin. However, stem−loop misbonds turn out to be less relevant in determining the closing time for this hairpin. It is interesting to note that, for the hairpin S_1, misbonding in the stem significantly reduces τ_1 (from $\sim 1.25\tau_{\mathrm{ref}}$ to $\sim 0.42\tau_{\mathrm{ref}}$ at $T = 280$ K), but it only reduces P_{success} slightly (from ~ 0.37 to ~ 0.31 at $T = 280$ K) as the transient entrapment of the hairpin in misbonded states only makes the zippering stage of the hairpin formation slightly less likely. Note that the effect of misbonding that we observe here works against the nonmonotonic behavior of τ_c, rather than causing it. For hairpin S_2, on the other hand, there are fewer misbonding possibilities and therefore misbonding does not play an important role in the hairpin closing process, as is shown in Figure 6c.

Figure 4a also compares the free-energy profiles for hairpin formation to the case when misbonding is not allowed. We observe that misbonding stabilizes the states with high free energy and therefore reduces the effective free-energy barriers of both closing and opening reactions. This explains, once more, why misbonding is assisting the hairpin formation/

opening for the sequence S_1. Note that misbonds can only slow down hairpin formation at low temperature if the following conditions hold: (a) the misbonded configurations must be stable relative to the open state, and (b) the misbonded configuration should be unable to easily transition to the folded state (e.g., via internal displacement[32]). In oxDNA, for the hairpins studied here, these conditions do not hold. It is also difficult to see how WC interactions could cause such behavior for physical DNA.

Role of Stacking. To further investigate the nature of the growth of τ_c at low temperatures, we systematically vary the stacking strength between the bases in the loop from no stacking ($\lambda_{\mathrm{st}} = 0$) to normal stacking ($\lambda_{\mathrm{st}} = 1$) and measure the closing time. The results are shown in Figure 7a. Note that
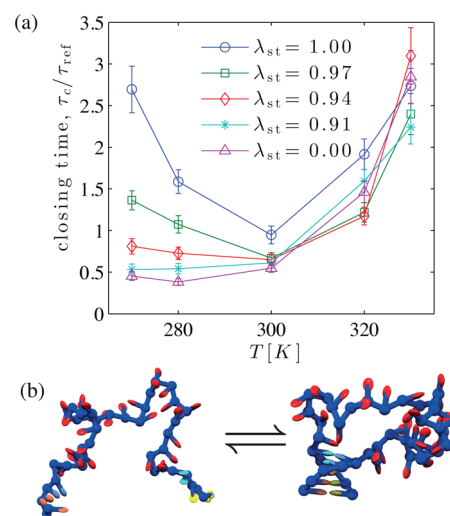
Figure 7. (a) Dependence of the closing time on λ_{st}, the strength of base stacking in the loop. (b) Typical snapshots of the S_1 hairpin at $T = 300$ K in closed and open states when there is no base stacking in the loop, $\lambda_{\mathrm{st}} = 0$.

stem−loop misbonding is also forbidden, but this imposes only a negligible perturbation to the dynamics of the system, as was shown earlier. We observe that, for large λ_{st} and at low T, where the bases in the loop stack strongly, the closing time grows. However, for $\lambda_{\mathrm{st}} \approx 0$, no such rise is observed.

The free-energy profiles in Figure 4a also indicate that decreasing the stacking strength in the loop stabilizes the closed

states and thus shifts the melting point to a slightly higher temperature; such a shift was also observed in experiment when a loop sequence was replaced by one believed to stack less strongly[11] and also studied previously for oxDNA.[30] The opening of the more stable hairpin is slower compared to the normal hairpin (see Figure 4b). In addition, the reduction of the effective height of the free-energy barriers to hairpin formation that we see in Figure 4b further suggests that the rapid growth of the closing time observed at low temperatures is actually due to the strong stacking in the loop.

To further quantify the effect of the stacking interactions, we measure the average stack length $\langle l \rangle$, the probability of stacking P_{st}, and the average end-to-end distance $\langle R_{ee} \rangle$ for a 30-nucleotide poly(dA) ssDNA as the strength of AA stacking is varied. The stack length $\langle l \rangle$ is defined as the average over stacks of contiguously stacked bases (a stack of length l consists of $l + 1$ bases). Two neighboring bases are considered to be stacked if the stacking interaction between them is less than or equal to $\epsilon_c = -(0.596 \lambda_{st})$ (note that the typical stacking energy between stacked bases is ~10 times this threshold).

The results as a function of temperature and stacking strength are summarized in Figure 8. For small P_{st}, the average
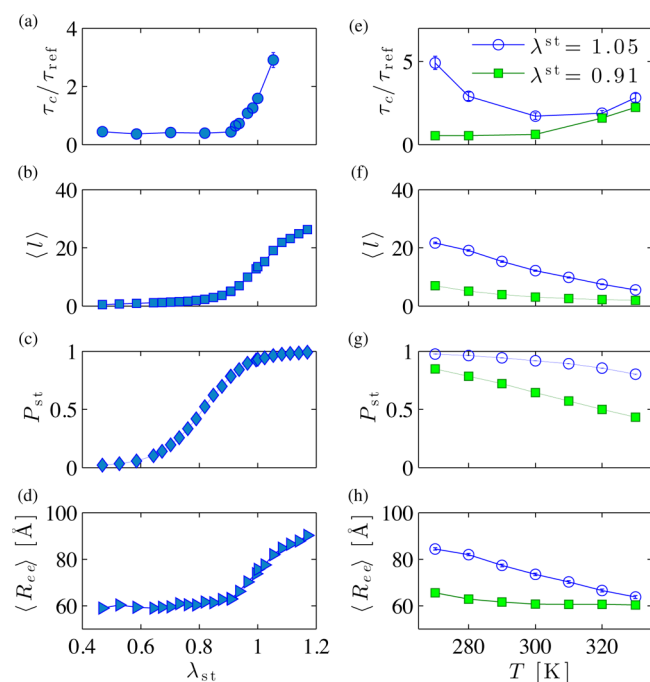
Figure 8. Closing time τ_c, average stacked length $\langle l \rangle$, probability of stacking P_{st}, and end-to-end distance R_{ee} as a function of (a−d) the strength of stacking interaction when $T = 280$ K and as a function of (e−h) the temperature when $\lambda_{st} = 1.05$ (open circles) and when $\lambda_{st} = 0.91$ (filled squares).

stack length is close to zero, and since there is little base stacking, the DNA chain is very flexible. All quantities that are plotted in Figure 8 remain almost constant when $P_{st} < 0.7$, and above this threshold, they start to increase significantly. The stacking length for an infinite chain has previously been shown to behave as $\langle l \rangle = P_{st}/(1 - P_{st})$ in an uncooperative model.[30] By lowering the temperature and/or increasing the stacking strength, P_{st} and therefore $\langle l \rangle$ increase. Note that, in the strong stacking regime, due to the finite length of our ssDNA chain, $\langle l \rangle$ deviates from the above formula, which predicts $\langle l \rangle$ to

diverge as $P_{st} \rightarrow 1$. In this regime, the strong stacking of the bases leads to a stiff ssDNA chain. Comparing the plots of τ_c and $\langle l \rangle$ (Figure 8a, b, e, and f) shows that the increase of τ_c for the hairpin S_1 at low temperatures coincides with the stiffening of ssDNA, whereas when the stacking is weaker $\langle l \rangle$ is always much shorter than the loop length and τ_c does not rise at low temperatures. It is interesting to note that the AA-stacking is the strongest stacking interaction within oxDNA, and the consequence of slightly reducing its strength is to suppress the low-temperature rise of the closing time. Therefore, if stacking is the root cause of the observed behavior in experiments,[26] it must be very strong (corresponding to high stacking probabilities ~$P_{st} > 0.9$).

Sequence Heterogeneity. All the hairpins that we have studied so far have a homogeneous loop with identical bases. While the stacking among different bases does break and form over time (temporal heterogeneity), the average local stiffness of the ssDNA chain is homogeneous along the loop. Intuitively, one can imagine that introducing a few weak stacking points along the loop should effectively reduce the loop stiffness and therefore facilitate hairpin formation at low temperatures. To test this hypothesis, we replaced three of the adenine bases in the loop of the S_1 hairpin with "dummy" bases (D) for which we have chosen the strength of the stacking interactions to be $\lambda_{st}^{DA} = \lambda_{st}^{AD} = \lambda_{st}^{DD} = 0.91$, which would be similar in strength to our parametrization of AT stacking in ref 30. Moreover, we vary the positions of the D bases in the loop to highlight any position-specific effects. The sequences that we consider are

- (I) 3′-CCCAA DAAAAAAAAAAAAADAAAAAAAAAA-AAAADA TTGGG-5′
- (II) 3′-CCCAA AAAAAADAAAAAAADAAAAAAAD-AAAAAAA TTGGG-5′
- (III) 3′-CCCAA AAAAAAAAAADAAADAAADAAAA-AAAAAAA TTGGG-5′

Figure 9a displays the closing time for the three modified hairpins compared to the normal S_1 hairpin and the hairpin with $\lambda_{st} = 0$. Our results show that the presence of just a few weak stacking points along the loop significantly affects the hairpin closing dynamics, making τ_c shorter. In addition, the positions of the weaker stacking points are important. We find that hairpin formation is easier when the weak stacking points are moved toward the middle of the loop, consistent with the results of Goddard et al.[11] who found that replacement of an adenine base with a cytosine base in the middle of the loop of a similar hairpin was most effective at reducing the low-temperature closing time. For sequence III, where the weaker points are closer to the middle of the loop, the closing process is enhanced the most. For this hairpin, the three modified bases effectively reduce τ_c to a value comparable to the closing time of the hairpin with no base stacking in the loop, except for the lowest temperature that we consider. These results are consistent with the intuitive expectation that the stem-forming ends of the strands are more likely to come close together if the strand can bend back on itself more easily at the middle of the hairpin loop.

■ DISCUSSION AND CONCLUSIONS

We have studied thermodynamic and kinetic properties of DNA hairpin formation with oxDNA, a coarse-grained model of DNA at the nucleotide level. The hairpins we consider, some of which have been studied experimentally, have a relatively long (30 bases) loop and a rather short (5 bases) stem. Our
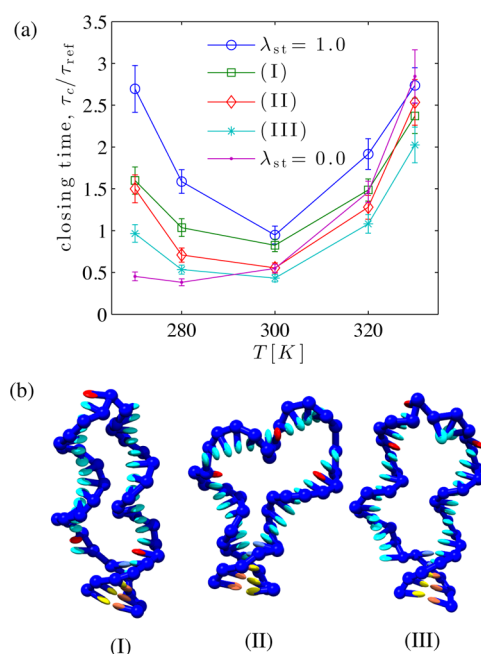
Figure 9. (a) Introduction of three weak stacking points along the loop significantly reduces τ_c at low temperatures. (b) Snapshots of the three types of the modified loop hairpins. Red colored bases mark the positions of the weak stacking points.

results indicate that the formation of such hairpins at high salt concentrations is well-approximated by a two-state reaction with a rate-determining step that involves the formation of a loop with a nucleus of one or two stem–stem base pairs. The formation of the hairpin is then achieved by a rapid zippering of the stem. In addition, the closing time shows a non-Arrhenius temperature dependence with a minimum close to the melting temperature, consistent with some previous experiments.[10,26,27,45]

At high temperatures, the temperature dependence of the closing time shows a negative apparent activation enthalpy (i.e., increases with temperature), implying that the set of transition states have on average a lower enthalpy (due to base pairing) than the open state. In this regime, the free-energy barrier to hairpin formation is mainly due to a significant loss of conformational entropy associated with the formation of the initial stem–stem contact. Moreover, similar to what has been recently observed for duplex hybridization[32] with oxDNA, we found that at high temperatures a loop that is stabilized by just one base pair is unlikely to succeed in forming a complete hairpin. Instead, a larger number of stem–stem base pairs is required to make zippering most likely at high temperatures. Hence, the apparent activation enthalpy for the hairpin formation becomes more negative as the temperature increases, in agreement with several experiments[10,26,45] but in contrast with ref 25.

The behavior changes at temperatures below the melting point. At these temperatures, the closing time shows a positive apparent activation enthalpy (i.e., increases with decreasing temperature), consistent with several studies.[10,26,45] As the temperature is decreased, the bases in the loop stack more strongly together. To form an initial stem–stem contact at low temperatures requires the bending rigidity of the ssDNA loop to be overcome and involves a loss of stacking interactions, thus increasing the enthalpy of the transition states. At some point,

the enthalpic cost of forming a loop becomes larger than the negative enthalpy gain due to base pairing at the transition state, and as a result, the apparent activation enthalpy becomes positive. The positive activation enthalpy has previously been attributed to the roughness of the free-energy surface, due to the misbonded (WC and non-WC misbonds including misstacked bases) configurations in the early stages of the formation of the loop.[21,45] For oxDNA at least, we have shown here that this phenomenon is mainly due to strong base stacking in the loop and that the WC misbonding when present acts to suppress the nonmonotonic behavior observed for the hairpin closing time.

In addition, we have shown that by decreasing the stacking strength of the bases in the loop, or by introducing a few more weakly stacking bases in the loop, the hairpin closing time decreases significantly at low temperatures, in agreement with experimental observations.[11] The significance of introducing sites of enhanced flexibility for the hairpin dynamics provides further evidence for the key role of loop rigidity in determining the hairpin closing time at low temperatures. For example, although such a hairpin loop is more flexible than our standard case, the overall roughness of the free-energy landscape is likely to be almost the same for the two hairpins.

It should be noted that in our study we used an implicit solvent model. Therefore, the temperature dependency of the rate constants due to the change of the solvent viscosity is not captured. As the viscosity of the solvent decreases with increasing temperature, its effect would be to move the minimum in the closing time to slightly higher temperature. Furthermore, hydrodynamic effects are neglected in our simulations. However, as all the strands that are considered here are the same length (40 bases), they are likely to experience similar hydrodynamic effects. Such effects are likely to be more relevant when studying the kinetics of hairpins with different sizes.

Overall, our results suggest that strong stacking within the loop of a hairpin can produce hairpin folding kinetics consistent with experimental studies that have reported nonmonotonic dependencies on temperature. We note that, if neighboring bases stack in a largely uncooperative fashion, the overall stacking strength must be very large (corresponding to high stacking probabilities $\sim P_{st} > 0.9$) to produce a strong signal. Such strong stacking is found in oxDNA for polyA at low temperatures. While we do not find similar effects arising from non-native WC base pairs (since those misbonded configurations are not more stable than the open state, and they can easily transition to the folded state), there could possibly be other contributions to the observed behavior. In particular, there may be non-WC interactions (e.g., alternative base pairings on the WC edge or interactions on the Hoogsteen and sugar edges) that are neglected in oxDNA but are sufficiently stable at low temperature to inhibit folding. Alternatively, single-stranded stem sections might form structures at low temperatures that are incompatible with nucleating the correct stem and that must be disrupted prior to hairpin formation.[32] Finally, changing interactions with the solvent or increased viscosity of water at lower temperatures may play a role. We note, however, that these alternative explanations are unlikely to be strongly sensitive to substitutions of a few bases within the loop, as is observed in oxDNA when strong stacking is responsible. The fact that such changes are known to influence folding dynamics at low temperature, therefore, supports the stacking-based explanation rather than the suggestion of Ansari

and co-workers[21,45] that an increased roughness of the free-energy surface arising from intrachain misbonding controls the hairpin formation dynamics.

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: majid.mosayebi@chem.ox.ac.uk.
*E-mail: jonathan.doye@chem.ox.ac.uk.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Bikard, D.; Loot, C.; Baharoglu, Z.; Mazel, D. Folded DNA in Action: Hairpin Formation and Biological Functions in Prokaryotes. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 570−588.

(2) Sassone-Corsi, P.; Zazopoulos, E.; Lalli, E.; Stocco, D. M. DNA Binding and Transcriptional Repression by DAX-1 Blocks Steroido-genesis. *Nature* **1997**, *390*, 311−315.

(3) Roth, D. B.; Menetski, J. P.; Nakajima, P. B.; Bosma, M. J.; Gellert, M. V(D)J Recombination: Broken DNA Molecules with Covalently Sealed (Hairpin) Coding Ends in Scid Mouse Thymocytes. *Cell* **1992**, *70*, 983−991.

(4) Kennedy, A. K.; Guhathakurta, A.; Kleckner, N.; Haniford, D. B. Tn10 Transposition via a DNA Hairpin Intermediate. *Cell* **1998**, *95*, 125−134.

(5) Varani, G. Exceptionally Stable Nucleic Acid Hairpins. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379−404.

(6) Green, S. J.; Lubrich, D.; Turberfield, A. J. DNA Hairpins: Fuel for Autonomous DNA Devices. *Biophys. J.* **2006**, *91*, 2966−2975.

(7) Tomov, T. E.; Tsukanov, R.; Liber, M.; Masoud, R.; Plavner, N.; Nir, E. Rational Design of DNA Motors: Fuel Optimization through Single-Molecule Fluorescence. *J. Am. Chem. Soc.* **2013**, *135*, 11935−11941.

(8) Tyagi, S.; Kramer, F. R. Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nat. Biotechnol.* **1996**, *14*, 303−308.

(9) Yin, P.; Choi, H. M.; Calvert, C. R.; Pierce, N. A. Programming Biomolecular Self-Assembly Pathways. *Nature* **2008**, *451*, 318−322.

(10) Shen, Y.; Kuznetsov, S. V.; Ansari, A. Loop Dependence of the Dynamics of DNA Hairpins. *J. Phys. Chem. B* **2001**, *105*, 12202−12211.

(11) Goddard, N. L.; Bonnet, G.; Krichevsky, O.; Libchaber, A. Sequence Dependent Rigidity of Single Stranded DNA. *Phys. Rev. Lett.* **2000**, *85*, 2400−2403.

(12) Woodside, M. T.; Anthony, P. C.; Behnke-Parks, W. M.; Larizadeh, K.; Herschlag, D.; Block, S. M. Direct Measurement of the Full, Sequence-Dependent Folding Landscape of a Nucleic Acid. *Science* **2006**, *314*, 1001−1004.

(13) Tsukanov, R.; Tomov, T. E.; Berger, Y.; Liber, M.; Nir, E. Conformational Dynamics of DNA Hairpins at Millisecond Resolution Obtained from Analysis of Single-Molecule FRET Histograms. *J. Phys. Chem. B* **2013**, *117*, 16105−16109.

(14) Kenward, M.; Dorfman, K. D. Brownian Dynamics Simulations of Single-Stranded DNA Hairpins. *J. Chem. Phys.* **2009**, *130*, 095101.

(15) Bowman, G. R.; Huang, X.; Yao, Y.; Sun, J.; Carlsson, G.; Guibas, L. J.; Pande, V. S. Structural Insight into RNA Hairpin Folding Intermediates. *J. Am. Chem. Soc.* **2008**, *130*, 9676−9678.

(16) Sales-Pardo, M.; Guimerà, R.; Moreira, A.; Widom, J.; Amaral, L. Mesoscopic Modeling for Nucleic Acid Chain Dynamics. *Phys. Rev. E* **2005**, *71*, 051902.

(17) Frederickx, R.; In't Veld, T.; Carlon, E. Anomalous Dynamics of DNA Hairpin Folding. *Phys. Rev. Lett.* **2014**, *112*, 198102.

(18) Schreck, J. S.; Ouldridge, T. E.; Romano, F.; Šulc, P.; Shaw, L.; Louis, A. A.; Doye, J. P. K. DNA Hairpins Primarily Promote Duplex Melting rather than Inhibiting Hybridization. *arXiv:1408.4401*, 2014.

(19) Hyeon, C.; Thirumalai, D. Mechanical Unfolding of RNA Hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6789−6794.

(20) Liu, L.; Chen, S.-J. Coarse-grained Prediction of RNA Loop Structures. *PLoS One* **2012**, *7*, e48460.

(21) Kuznetsov, S. V.; Ansari, A. A Kinetic Zipper Model with Intrachain Interactions Applied to Nucleic Acid Hairpin Folding Kinetics. *Biophys. J.* **2012**, *102*, 101−111.

(22) Orden, A. V.; Jung, J. Fluorescence Correlation Spectroscopy for Probing the Kinetics and Mechanisms of DNA Hairpin Formation. *Biopolymers* **2008**, *89*, 1−16.

(23) Jung, J.; Van Orden, A. A Three-State Mechanism for DNA Hairpin Folding Characterized by Multiparameter Fluorescence Fluctuation Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128*, 1240−1249.

(24) Ma, H.; Wan, C.; Wu, A.; Zewail, A. H. DNA Folding and Melting Observed in Real Time Redefine the Energy Landscape. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 712−716.

(25) Bonnet, G.; Krichevsky, O.; Libchaber, A. Kinetics of Conformational Fluctuations in DNA Hairpin-Loops. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8602−8606.

(26) Wallace, M. I.; Ying, L.; Balasubramanian, S.; Klenerman, D. Non-Arrhenius Kinetics for the Loop Closure of a DNA Hairpin. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5584−5589.

(27) Wallace, M. I.; Ying, L.; Balasubramanian, S.; Klenerman, D. FRET Fluctuation Spectroscopy: Exploring the Conformational Dynamics of a DNA Hairpin Loop. *J. Phys. Chem. B* **2000**, *104*, 11551−11555.

(28) Ouldridge, T. E. Coarse-grained Modelling of DNA and DNA Self-Assembly. Ph.D. Thesis, University of Oxford, 2011.

(29) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. Structural, Mechanical, and Thermodynamic Properties of a Coarse-Grained DNA Model. *J. Chem. Phys.* **2011**, *134*, 085101.

(30) Šulc, P.; Romano, F.; Ouldridge, T. E.; Rovigatti, L.; Doye, J. P. K.; Louis, A. A. Sequence-Dependent Thermodynamics of a Coarse-Grained DNA Model. *J. Chem. Phys.* **2012**, *137*, 135101.

(31) Doye, J. P. K.; Ouldridge, T. E.; Louis, A. A.; Romano, F.; Šulc, P.; Matek, C.; Snodin, B. E. K.; Rovigatti, L.; Schreck, J. S.; Harrison, R. M.; et al. Coarse-Graining DNA for Simulations of DNA Nanotechnology. *Phys. Chem. Chem. Phys.* **2013**, *15*, 20395−20414.

(32) Ouldridge, T. E.; Šulc, P.; Romano, F.; Doye, J. P. K.; Louis, A. A. DNA Hybridization Kinetics: Zippering, Internal Displacement and Sequence Dependence. *Nucleic Acids Res.* **2013**, *41*, 8886−8895.

(33) Srinivas, N.; Ouldridge, T. E.; Šulc, P.; Schaeffer, J. M.; Yurke, B.; Louis, A. A.; Doye, J. P. K.; Winfree, E. On the Biophysics and Kinetics of Toehold-Mediated DNA Strand Displacement. *Nucleic Acids Res.* **2013**, *41*, 10641−10658.

(34) Šulc, P.; Romano, F.; Ouldridge, T. E.; Doye, J. P. K.; Louis, A. A. A Nucleotide-Level Coarse-Grained Model of RNA. *J. Chem. Phys.* **2014**, *140*, 235102.

(35) Leontis, N. B.; Stombaugh, J.; Westhof, E. The Non-Watson-Crick Base Pairs and their Associated Isostericity Matrices. *Nucleic Acids Res.* **2002**, *30*, 3497−3531.

(36) http://dna.physics.ox.ac.uk/.

(37) Romano, F.; Chakraborty, D.; Doye, J. P. K.; Ouldridge, T. E.; Louis, A. A. Coarse-Grained Simulations of DNA Overstretching. *J. Chem. Phys.* **2013**, *138*, 085101.

(38) Matek, C.; Ouldridge, T. E.; Doye, J. P. K.; Louis, A. A. Plectoneme Tip Bubbles: Coupled Denaturation and Writhing in Supercoiled DNA. *arXiv:1404.2869*, 2014.

(39) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. DNA Nanotweezers Studied with a Coarse-Grained Model of DNA. *Phys. Rev. Lett.* **2010**, *104*, 178101.

(40) Ouldridge, T. E.; Hoare, R. L.; Louis, A. A.; Doye, J. P. K.; Bath, J.; Turberfield, A. J. Optimizing DNA Nanotechnology through

Coarse-Grained Modeling: A Two-Footed DNA Walker. *ACS Nano* **2013**, *7*, 2479−2490.

(41) Russo, J.; Tartaglia, P.; Sciortino, F. Reversible Gels of Patchy Particles: Role of the Valence. *J. Chem. Phys.* **2009**, *131*, 014504.

(42) Allen, R. J.; Warren, P. B.; ten Wolde, P. R. Sampling Rare Switching Events in Biochemical Networks. *Phys. Rev. Lett.* **2005**, *94*, 018104.

(43) Whitelam, S.; Geissler, P. L. Avoiding Unphysical Kinetic Traps in Monte Carlo Simulations of Strongly Attractive Particles. *J. Chem. Phys.* **2007**, *127*, 154101.

(44) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(45) Ansari, A.; Kuznetsov, S. V.; Shen, Y. Configurational Diffusion down a Folding Funnel Describes the Dynamics of DNA Hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7771−7776.

(46) Uzawa, T.; Isoshima, T.; Ito, Y.; Ishimori, K.; Makarov, D. E.; Plaxco, K. W. Sequence and Temperature Dependence of the End-to-End Collision Dynamics of Single-Stranded DNA. *Biophys. J.* **2013**, *104*, 2485−2492.