Coarse-grained simulations of DNA overstretching

Flavio Romano, Debayan Chakraborty, Jonathan P. K. Doye, Thomas E. Ouldridge, and Ard A. Louis

Citation: J. Chem. Phys. **138**, 085101 (2013); doi: 10.1063/1.4792252
View online: http://dx.doi.org/10.1063/1.4792252
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v138/i8
Published by the AIP Publishing LLC.

Coarse-grained simulations of DNA overstretching

Flavio Romano,[1] Debayan Chakraborty,[1,a)] Jonathan P. K. Doye,[1,b)]
Thomas E. Ouldridge,[2] and Ard A. Louis[2]
[1]*Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom*
[2]*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, United Kingdom*

(Received 26 September 2012; accepted 31 January 2013; published online 22 February 2013)

We use a recently developed coarse-grained model to simulate the overstretching of duplex DNA. Overstretching at 23 °C occurs at 74 pN in the model, about 6–7 pN higher than the experimental value at equivalent salt conditions. Furthermore, the model reproduces the temperature dependence of the overstretching force well. The mechanism of overstretching is always force-induced melting by unpeeling from the free ends. That we never see S-DNA (overstretched duplex DNA), even though there is clear experimental evidence for this mode of overstretching under certain conditions, suggests that S-DNA is not simply an unstacked but hydrogen-bonded duplex, but instead probably has a more exotic structure. © *2013 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4792252]

I. INTRODUCTION

DNA *in vivo* is not just a passive molecular bearer of information, but is an active molecule that is able to respond structurally to cellular signals, be it through changes in solution conditions, protein binding, or the action of molecular machines. Some of these changes are mediated by very specific responses to the biochemical details of the protein binding, say, but in many instances this control is achieved through the mechanical response of DNA, e.g., enzymes that apply structural control by adjusting the supercoiling of DNA. For these reasons, there has been much interest in the fundamental mechanical properties of DNA,[1–3] particularly as there are now the means to study these properties in unprecedented detail using single-molecule techniques.

One particular focus has been the response of DNA to tension. At low forces the B-DNA duplex responds elastically and is well described by the worm-like chain model.[4] However, at higher forces, typically in the 60–70 pN range at room temperature (although the precise value depends on solution conditions[5–7]), DNA undergoes a dramatic overstretching in which it extends by about 70% over a few pN. Perhaps somewhat surprisingly, even though the first detailed results on DNA overstretching were reported in 1996,[8,9] the nature of the overstretched state is only now beginning to be fully resolved and has been a controversial topic for much of this period.

The two main proposals are that yielding corresponds to a transition to a new overstretched form of double-stranded DNA (normally termed S-DNA)[8,9] or to force-induced melting.[10–12] Features interpreted as pointing to an S-DNA mechanism include (i) the lack of complete strand separation after overstretching, as might be expected for force-induced melting, (ii) that the force-extension curves after overstretching often do not initially follow that expected for single-stranded DNA (ssDNA), (iii) that further transitions can sometimes be seen at a higher force (that are taken to correspond to force-induced melting and strand dissociation), and (iv) the reversibility of the transition under some conditions. By contrast, the dependence of the overstretching force F_{over} on temperature[7,13] and solution conditions (e.g., pH[5] and salt concentration[6,7]) fits well with the force-induced melting picture; i.e., conditions which destabilize the duplex (e.g., higher temperature, lower salt) lead to a lower overstretching force. Furthermore, the hysteresis sometimes seen has been interpreted as due to the slow nature of reassociation after the melting of a long DNA molecule.

New impetus has come to this debate from the recent fluorescence experiments of van Mameren *et al*. that clearly showed that at room temperature and at 5–150 mM salt overstretching occurred by unpeeling from the free ends.[14] Although this experiment was interpreted by some as proof of the force-induced melting hypothesis (and hence that the S-DNA hypothesis was unnecessary),[15,16] this conclusion was disputed by others.[17–19] Furthermore, subsequent experiments[7,20–28] have shown this interpretation to be too simplistic and provide a basis for a more balanced position. In particular, these experiments have shown that the two modes of overstretching can occur depending on conditions, with S-DNA being favoured at low temperature, high salt, high G-C content, and short time scales. One of the complications is that at room temperature the two mechanisms can compete. So for lower salt concentrations unpeeling is dominant, as in the experiments of van Mameren *et al*.[14] and the impressive high-resolution follow-up study of Gross *et al*. which resolved a saw-tooth structure in the overstretching force-extension curves that correlated with the ease with which different parts of the sequence could unpeel.[27] However, at higher salt, an S-DNA mechanism becomes more feasible. For example, in the elegant experiment of Paik *et al*. at 150 mM salt

a)Present address: Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.
b)Author to whom correspondence should be addressed. Electronic mail: jonathan.doye@chem.ox.ac.uk.

the transition occurred with some hysteresis when free ends were present (presumably due to unpeeling), but occurred reversibly at the same overstretching force for a DNA construct that had no free ends and was still torsionally unconstrained (presumably by an S-DNA mechanism).

There has also been a lot of work using theory and simulation that has aimed to provide further insights into the nature of the overstretching transition. The theoretical approaches usually use simple polymer models to describe each of the different possible states with parameters fitted to reproduce experimental behaviour.[17,29–43] In particular, the work of Cocco and Marko[32] and Whitelam and co-workers[17,35,36] has provided important insights into the competition between different processes, such as S-DNA formation versus unpeeling. Furthermore, such approaches are particularly well-suited to the long length and time scales that are typical of experiments on DNA stretching. However, as these models lack a detailed molecular representation of the underlying phenomena, the questions these approaches can address is inevitably limited.

At the other extreme, one has simulations of DNA stretching using fully-atomistic models.[44–54] These have the potential to provide answers to detailed structural questions that are inaccessible to experiment. However, first, the computational cost of such simulations means that the pulling rates are often extremely fast, making it likely that much of what is observed is not at equilibrium. Second, it is not clear how well the atomistic potentials will perform under the more extreme conditions associated with overstretching. Third, in some of the simulations, constraints or boundary conditions are used that restrict the range of possible behaviour, e.g., torsional constraints or the absence of free ends. For these reasons, it is perhaps not surprising that these simulations only observe overstretching at forces well above the experimental values (all above 100 pN, some substantially more) and that the transitions are much broader. Furthermore, although they have the potential to identify the structure of S-DNA, no consensus has been reached.

In the middle are simulations using coarse-grained models that have a simplified representation of DNA and so allow longer time-scale and length-scale processes to be more easily accessible. These can be divided into two types. First, simple models that attempt to obtain qualitative insights into the stretching behaviour that stem from capturing some of the basic physical ingredients of DNA.[38,55,56] Second, there are coarse-grained DNA models that are fitted in order to provide a quantitatively accurate description of DNA's behaviour.[57–59] For some of these models, the stretching behaviour has been simulated.[60–63] However, the observed overstretching forces are also typically much higher than experiment (e.g., 320 pN in Ref. 61, 500-1100 pN in Ref. 62, and 400-550 pN in Ref. 63).

Here, we simulate DNA overstretching for our recently developed coarse-grained DNA potential, which provides a quantitative description of many of the structural, thermodynamic, and mechanical properties of DNA.[64,65] One aim of the work is to provide an independent test of the robustness of the model by applying it to phenomena to which it has not been fitted. A second aim is to provide fundamental insights into the nature of the overstretching transition, both by visualizing some of the mechanistic aspects of the transitions and by deducing from the successes and failures of the model which ingredients are needed to capture different aspects of the transition. We proceed as follows. In Section II, we describe our DNA potential, the simulation techniques used, and some of the theory of the thermodynamics of pulling. In Section III, we focus in detail on the room temperature overstretching behaviour, before more briefly examining the temperature dependence of overstretching, including looking for the re-entrant behaviour that has been predicted to occur near melting.[10,13,33,34,38,39,55]

II. METHODS

A. DNA model

We use the coarse-grained model of DNA recently developed in our group.[64,65] In particular, the model has been designed to provide a good description of the structural, thermodynamic, and mechanical properties of both double- and single-stranded DNA, all features that are very important if we are to be able to describe DNA overstretching. In addition, the wide range of applications for which the model has been successfully used give us further confidence in its robustness. These applications so far include DNA nanotweezers,[64] kissing hairpins,[66] DNA walkers,[67–69] the formation of cruciform DNA,[70] and the nematic transition of dense solutions of short duplexes;[71] furthermore, investigations into hybridization, strand-exchange reactions, plectonemes, DNA origami, and phase transitions in solutions of DNA junctions are ongoing.

The model describes each nucleotide as a rigid object (Fig. 1(a)) that has interactions corresponding to backbone connectivity, excluded volume, stacking, hydrogen-bonding between complementary bases, and cross-stacking. Each nucleotide has three collinear interaction sites (corresponding to the centres of the backbone, stacking, and hydrogen-bonding interactions) and a vector perpendicular to this axis to capture the planarity of the bases. We would like to emphasise that the attractive interactions in the model are not isotropic, but depend on the relative orientation of the nucleotides. It is these angular modulations that ensure that the DNA helix is

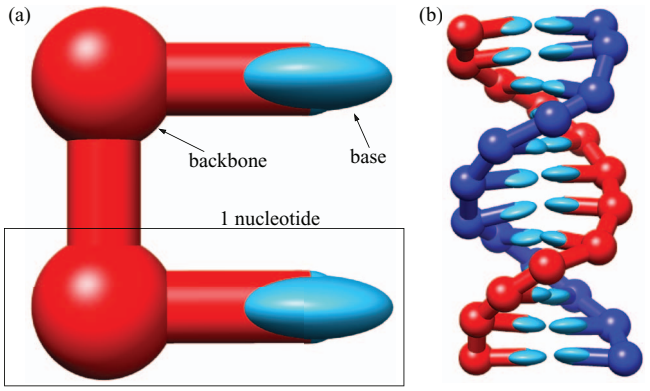

FIG. 1. (a) Two nucleotides represented by our model, showing the rigid nucleotide unit and the backbone and base regions. (b) A short DNA duplex as represented by our model.

right-handed and pairs in an anti-parallel manner. The full form of the potential has been given in Refs. 65 and 67. A simulation code incorporating the potential is available to download.[72]

We should note a number of simplifying features of the model that are relevant to the current study. First, we do not attempt to model the electrostatics explicitly, but instead have fitted the potential parameters for a particular salt concentration, namely 500 mM. At this concentration the Debye screening length is short-ranged and the effects of the electrostatics are included in the excluded-volume interactions. Fortunately, this salt concentration is also one of those commonly used in experiments on overstretching. However, as a consequence of this simplification we cannot of course examine how overstretching depends on salt concentration. At room temperature and 500 mM salt, experiments indicate that both the force-induced melting and S-DNA modes of overstretching compete.[7,22,23]

A second of the model's simplifications is that it ignores sequence dependence in the interactions except for the Watson-Crick nature of the base pairing (i.e., the hydrogen bonding term in the potential only occurs for complementary bases). Therefore, each base pair has the same average interaction strength irrespective of its identity and its neighbours. This "average-base" approximation can be an advantage when one is interested in the generic behaviour of DNA, as for the most part we are here, but of course it excludes us from, for example, examining how the overstretching depends on G-C content. A sequence-dependent version of the model has very recently been developed.[73]

Third, the double helix in our model (Fig. 1(b)) is symmetrical with the helical grooves being of the same size. This simplification, however, is likely to be relatively unimportant for the current study.

B. Pulling schemes

The three ways of pulling DNA that we will consider are shown in Fig. 2. Our main focus will be on scheme I, in which a force is imposed on both ends of the same strand and the free ends allow the duplex to overstretch by unpeeling. Scheme II is similar in that a force is exerted on only one of the two strands at each end of the duplex, thus again allowing unpeeling. However, as the forces are applied to different strands, dissociation will be completely irreversible when pulled at constant force. In our simulations, we pull on both 3′ ends of each strand, but given the symmetric nature of the helix in our model and the simplified representation of the nucleotides, we do not expect any differences for pulling both 5′ ends. In experiments differences between 5′-5′ and 3′-3′ pulling are only found at forces well above that for overstretching.[74]

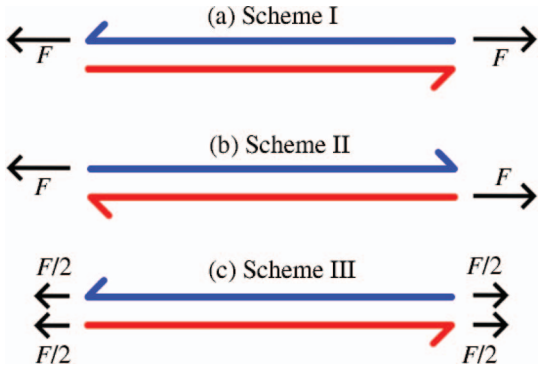

FIG. 2. Schematic representation of the three pulling schemes used. The arrows on the DNA are in the 3′-5′ direction.

Finally, in scheme III, as both ends of both strands are pulled, force-induced melting by unpeeling is suppressed. Therefore, if the DNA is to melt, it would have to be by bubble formation. Note the system in scheme III is not torsionally constrained, i.e., the ends are free to rotate.

C. Thermodynamics of pulling

Before we consider our pulling simulations, it is important to understand the effect of imposing a force on the thermodynamics of our systems.[10,32] The force provides an additional source of work being done on the system. Assuming the force F acts along the z direction, this gives an additional term Fdz to the change in internal energy and hence to the change in Helmholtz free energy dA, and A is therefore a function of z. However, it will usually be more convenient to consider a system under constant force and so we introduce the free energy $A' = A - Fz$ to describe such a constant-force ensemble. It follows that the change in free energy on applying a force is

$$A'(F) - A'(F = 0) = -\int_0^F z(F)dF. \qquad (1)$$

Therefore, the force-extension curves of dsDNA and ssDNA determine the relative stabilization of these two forms by force. Furthermore, assuming overstretching occurs by force-induced melting, one just needs a correct description of the zero-force thermodynamics and the force-extension curves to correctly predict overstretching.

Force-induced unpeeling of DNA occurs at the force at which the free energy per base pair of the ssDNA and dsDNA states are equal, leading to coexistence of the two forms. In other words, at this overstretching force $F_{\text{over}}(T)$, the average free energy change for a duplex to unpeel by one more base pair is zero. We wish to emphasise that overstretching, in the force-induced melting picture, is not determined by an equilibrium between dsDNA and fully-dissociated ssDNA, but rather by when the duplex becomes unstable with respect to unpeeling. In this sense overstretching is distinct from melting, as the translational entropy gain from complete dissociation of the two strands plays no role. However, in the bulk limit of infinitely long chains, the contribution of translational entropy is negligible and melting and unpeeling transitions are equivalent. In this limit, the melting temperature $T_m(F)$ can be defined as the temperature at which the average free energy change for the loss of one base pair from a duplex is zero at a given force F, and the curves $F_{\text{over}}(T)$ and $T_m(F)$ then give equivalent representations of the overstretching transition in the force-temperature plane. However, for finite chains $T_m(F)$

will depend on concentration and this equivalence will be broken.

We also note that care should be taken when applying this approach to the different schemes in Fig. 2. For scheme I only one of the strands is force-bearing when single stranded. By contrast, in scheme III the force on each single strand is $F/2$. Finally, in scheme II, equilibrium is not well-defined, as for any non-zero force the most stable state is two dissociated single strands because they can then increase their separation without limit.

D. Simulation methods

We use a mixture of Monte Carlo and Brownian dynamics to simulate DNA stretching for our model. We use Monte Carlo to obtain the equilibrium behaviour of the system at a particular force and Brownian dynamics to study the stretching dynamics of DNA when the force is increased or decreased (linearly) as a function of time.

To aid the equilibration in our Monte Carlo simulations, we use the virtual move Monte Carlo approach of Whitelam and co-workers.[75] This algorithm is a type of "cluster-move" Monte Carlo that allows clusters of nucleotides to be moved at each step, where the clusters that are constructed reflect both the configuration and the proposed move. The collective motion of nucleotides that this algorithm introduces allows more efficient sampling for our model compared to Monte Carlo with single-particle moves.

For the dynamics simulations we use the simple Brownian thermostat introduced in Ref. 76. The system is evolved according to Newtonian dynamics, but every n_{Newt} time steps a fraction p of the particles have all three components of their velocities drawn from the Maxwell distribution. On time scales longer than pn_{Newt}, the dynamics is Brownian.

Modifying these simulation algorithms to incorporate the forces exerted on the terminal bases consistent with the different schemes in Fig. 2 is straightforward. To perform Monte Carlo at constant force it is necessary to include an additional term in the energy $-\sum_i F_i z_i$, where F_i is the external force imposed on nucleotide i. In Monte Carlo, this gives rise to an extra term in the Boltzmann factor in the standard Metropolis acceptance criterion, namely $\exp(\beta \sum_i F_i \Delta z_i)$. For the Brownian dynamics, implementing the constant force is just a matter of adding the external forces to the forces acting on the relevant nucleotides.

III. RESULTS

In all cases, we consider a 100 base-pair duplex with a random sequence. We first consider pulling by scheme I in detail. Force-extension curves at "room temperature" (23 °C) are presented in Fig. 3 for Monte Carlo simulations at different forces and for dynamics simulations at different pulling rates. This data shows a number of clear features. First, the MC data (which we expect to be closer to "equilibrium") shows a clear and narrow overstretching transition at a force of about 77 pN. At the slowest pulling rates, the dynamics results also show a clear overstretching plateau at a very similar force. However, as the pulling rate increases, there is an increasing tendency to overshoot this transition, and for overstretching to start at a higher force and for the transition to be spread over a wider range of force, simply because the time scale associated with the overstretching transition is no longer fast compared to the pulling rate. These effects of pulling rate are also relevant to the all-atom simulations as they provide an indication of the consequences of being increasingly far from equilibrium. Indeed, our results at faster pulling rates somewhat resemble the force-extension curves seen in all-atom simulations.[45–54] We should also note that although our slowest pulling rates are significantly slower than those used in all-atom simulations, they are still much faster than in typical experiments.

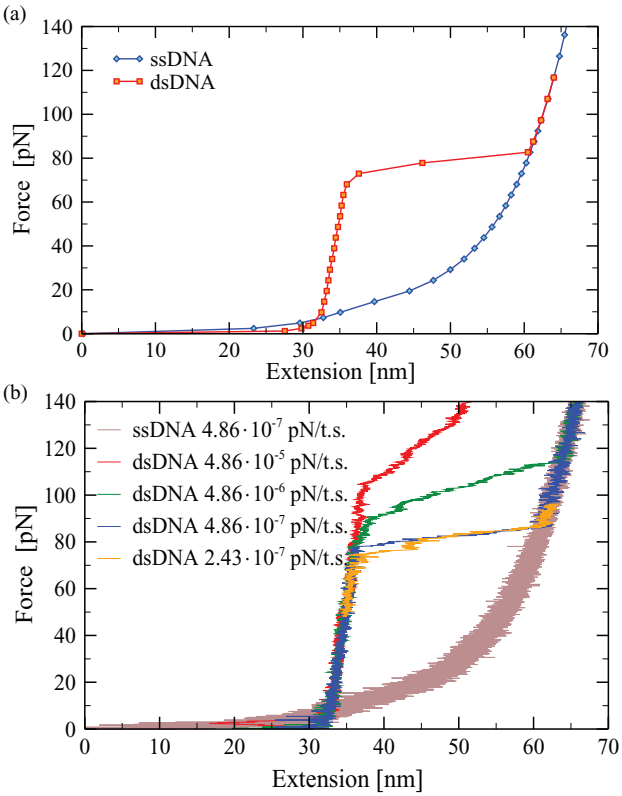

FIG. 3. Force-extension curves for DNA at 23 °C for pulling scheme I. In (a) results from a series of Monte Carlo simulations at constant force are presented. For comparison the curve for single-stranded DNA has been added. In (b) results from dynamics simulations are presented at a number of different pulling rates. The pulling rates are given in pN/time step. These units can be converted into pN ns^{-1} by multiplying by 1.17×10^5 time steps/ns, but as with any coarse-grained model absolute values of time should be treated with caution. Using this conversion, our slowest and fastest rates correspond to 0.0284 pN ns^{-1} and 5.69 pN ns^{-1}, respectively.

Second, beyond the overstretching transition, the force-extension curves follow that for ssDNA, thus indicating that DNA overstretching in our model is a result of force-induced melting. This conclusion is confirmed in Fig. 4 which depicts typical configurations as the system passes through the overstretching transition. From these snapshots, it can be clearly seen that the increase in extension is a result of DNA unpeeling from the free ends. In particular, there is no sign of any S-DNA-like state. Experiments suggest that the transition from B- to S-DNA is able to occur more rapidly than

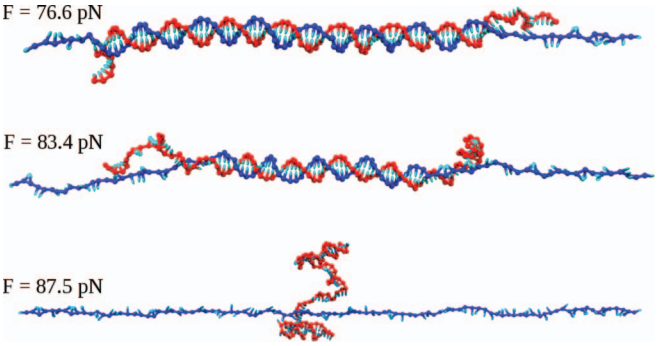

FIG. 4. Snapshots of typical DNA configurations as the system passes through the overstretching transition at 23 °C for scheme I. The snapshots are from the dynamics run at a pulling rate of 2.43×10^{-7} pN/time step.

force-induced melting.[17,22,23] Therefore, if S-DNA exists for our model, one might expect that it is more likely to be seen in the faster pulling simulations where the overstretching is observed at higher force. However, in all the simulations depicted in Fig. 3(b) we never found any evidence of an S-DNA-like state. Since the mechanism of overstretching for our model is always unpeeling, we should note that we expect the dynamical behaviour of our model to depend on chain length with slower pulling rates needed for longer molecules if overshooting of the transition is not to occur.

There are a number of more minor features that are apparent from the configurations in Fig. 4. First, immediately beyond the overstretching transition, the duplex does not fully dissociate, because for short (e.g., 4 base pairs) intermolecular helices, the force-bearing strand can still approximately align itself along the direction of the force, and so the extension gain associated with the loss of the base pairs is lower for these short duplexes. Therefore, a somewhat higher force is required to dissociate these last few bases. This effect can be seen in Fig. 5(a) where overstretching leads to the loss of most, but not quite all, of the energy associated with intermolecular hydrogen bonding—only at slightly higher forces does this term go to zero.

Second, from a careful inspection of Fig. 4, one can see that the orientation of the bases in the overstretched force-bearing strand is not random, but there seem to be short runs

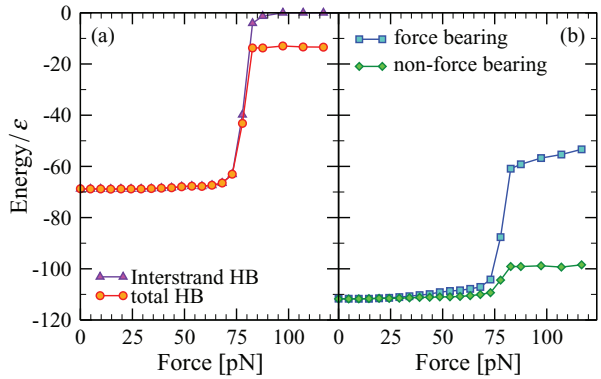

FIG. 5. Contributions to the potential energy from (a) hydrogen bonding and (b) stacking as a function of force. In (b) the contributions from the two individual strands are included. $k_B T/\epsilon = 0.1$ at $T = 300$ K.

of stacked bases. Although completely stacked ssDNA in our model is helical, it is possible for short sections of 3–4 stacked bases to orient the backbone approximately along the direction of force.[73] For this reason, the force-bearing strand is able to retain a significant fraction of its stacking interactions after overstretching. Figure 5(b) indicates that just after overstretching is complete, the stacking energy for the force-bearing strand is 57% of that for B-DNA and 61% of that for ssDNA at zero force (assuming the unpeeled non-force-bearing strand is typical). Interestingly, ssDNA bound to oligomers of DnaA proteins adopts a similar stretched geometry with short runs of three stacked bases.[77]

Third, it is clear from Fig. 4 and the appearance of intramolecular base pairs in Fig. 5(a) that the unpeeled strand exhibits secondary structure, in particular hairpins. Even though the sequence is random, hairpins can be formed with stems having a sufficient number of Watson-Crick base pairs (as well as some mismatches) to be stable.

Figure 3(a) only provides a rough guide to the position of the overstretching transition, because it is hard to fully equilibrate the simulations in the vicinity of the transition. One approach to locate F_{over}, as noted in Sec. II C, is to find the force at which the average free energy for unpeeling a base is zero. However, this approach is complicated by secondary structure formation. First, it makes thermodynamic sampling of the transition more difficult, because secondary structure in the unpeeled state leads to free energy barriers between different states. For example, it might be that for intermolecular base pairs to form a hairpin must first open or as the double helix unpeels the identity of the most stable hairpin in the unpeeled strand changes. Second, it obscures some of the basic physics of the overstretching transition, as hairpins will stabilize the unpeeled state, but the extent to which this occurs (and hence the overstretching force) will be sequence dependent (even for our "average base" model where all the interactions are assumed to be of the same strength).

Therefore, in our simulations to locate the overstretching force, we turn off intramolecular base pairing[78] and, but less importantly, non-native intermolecular base pairing.[79] As a consequence, the free energy change for unpeeling every base pair is now zero (not just on average) at the overstretching transition of our model and so we only need to sample a local section of the free energy landscape, rather than that for the whole transition. However, this approximation will also lead to an overestimation of the overstretching force, because of the stabilization of the unpeeled state by secondary structure. Later, we will estimate the magnitude of the resulting overestimation.

In Fig. 6(a) we show the free energy landscape for a 10-base section of the 100-base-pair duplex for a number of forces close to F_{over}. In real DNA the free energy landscape for unpeeling will have a saw-tooth structure due to the sequence-dependence of the thermodynamics of base-pairing (e.g., G-C base pairs are stronger than A-T base pairs) and the overstretching force will be where the average free-energy change for forming a base pair is zero. However, because of the "average-base" nature of our model, the free energy is roughly linear in the number of base pairs in the duplex. When the slope is positive, the fully associated duplex is most stable,

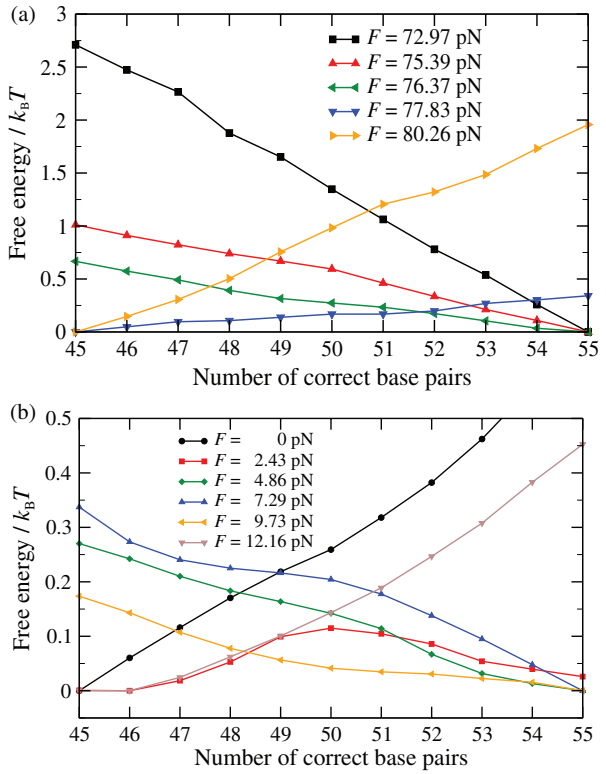

FIG. 6. Free energy profiles as a function of the number of base pairs for different forces at (a) room temperature ($T = 23\,^{\circ}$C) and (b) $T = 94\,^{\circ}$C. The latter is just above the zero-force bulk melting temperature and exhibits reentrance. Note the relative flatness of the free energy profiles at this temperature.

whereas when the slope is negative the fully unpeeled state is most stable. Thus, by estimating the force for which the slope is zero, we obtain a value of $F_{\rm over} \approx 77$ pN for our model at 23 °C when intramolecular base-pairing is turned off.

On the longer time scales in experiment, when unpeeling occurs at equilibrium, the competing states are dsDNA and ssDNA in which the free strand can form secondary structure. Therefore, it would be particularly useful for the comparison with experiment if we could estimate the error in the overstretching force that results from our constraint preventing secondary structure in the unpeeled state. The stabilization of ssDNA by secondary structure can be computed from the difference in the ssDNA force-extension curves when intramolecular base-pairing is allowed or forbidden using Eq. (1) and hence we can obtain an estimate of the correction to the overstretching force.

Although computing the force-extension curve of ssDNA with secondary structure is less difficult than simulating overstretching in the presence of secondary structure, it is nonetheless challenging, because of the potentially large free-energy barriers between different states. We therefore use parallel tempering, in which exchanges are attempted between simulations at different temperatures,[80] to help equilibration. Furthermore, as the stabilization will be sequence-dependent, one would also want to average the correction over different sequences when comparing to experiments for long DNAs. We therefore performed our simulations for five different sequences. However, for the sequence that formed the strongest secondary structure, we were unable to achieve equilibrium.

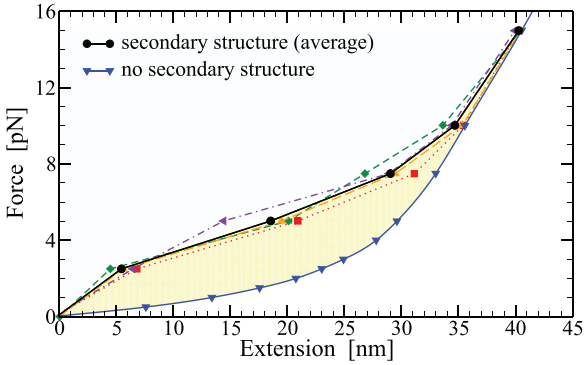

FIG. 7. Force-extension curves for ssDNA in the presence or absence of secondary structure. As the effects of secondary structure depend on sequence, the results for four different random sequences (as well as their average) are depicted. The area of the shaded region corresponds to the average free energy of stabilization of ssDNA by secondary structure.

The force-extension curves of the four other sequences are depicted in Fig. 7 (along with their average) and compared to that for ssDNA with no secondary structure. In particular, the secondary structure makes the ssDNA less extensible at low force. However, as the force increases the secondary structure becomes destabilized such that beyond 10 pN there is little difference between the force-extension curves when secondary structure is or is not allowed. This secondary-structure induced low-force feature is also apparent in experimental pulling curves for ssDNA at 500 mM salt, but disappears at lower salt.[82] For comparison, fully complementary hairpins are destabilized at about 16–17pN for the same salt conditions as here.[82]

The area between the curve for ssDNA with intramolecular base-pairing forbidden and the average curve with secondary structure present gives the average stabilization of a 100-base strand due to secondary structure. At room temperature, this stabilization is 0.23 kT per base on average. Given this value and the free-energy profiles in Fig. 6(a), we can then estimate the force at which the average free-energy change for unpeeling a base pair is zero when secondary structure in the unpeeled strand is allowed. This gives a correction of about 3 pN at room temperature, although this may be a slight underestimate, as we have not included the sequence that had the strongest secondary structure in our estimate.

Hence, our best estimate of the overstretching force for our model at 23 °C is $F_{\rm over} \approx 74$ pN. This value is satisfyingly close to the oft-quoted 65 pN for room temperature overstretching and much closer than has been previously achieved for any coarse-grained model or atomistic simulation. More precisely, at 500 mM salt, the concentration at which our coarse-grained model has been fitted, Wenner *et al.* obtain a value of 67 pN,[6] and in the more recent study of Zhang *et al.* the onset and mid-point of the transitions were found to occur at 65 pN and 68 pN, respectively.[7]

The same approach can be used to study the temperature dependence of $F_{\rm over}$, including the secondary structure correction factor for room temperature and above the secondary structure correction factor—note that above 45 °C the correction factor goes to zero, because by this temperature all of

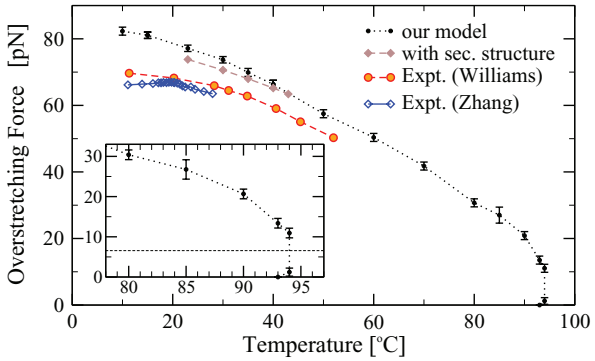

FIG. 8. Dependence of the overstretching force on temperature. The main results for our model are from simulations where intramolecular base-pairing was turned off, but we also include results where a correction for secondary structure formation in the unpeeled chain has been applied for temperatures between 23 °C and 43 °C—above the latter temperature there is no need for a correction as the secondary structure is thermally unstable. Also included are the experimental results from Refs. 7 and 13 at a salt concentration of 500 mM. Note that the results of Zhang *et al*. are for the onset (not the midpoint) of the transition. The inset provides an expansion of the high temperature region where non-monotonic behaviour is observed. The horizontal line in the inset is the force at which ssDNA and dsDNA have the same extension.

the secondary structure in the single strand has melted. As the temperature is increased, the duplex is destabilized relative to ssDNA, so F_{over} is expected to decrease. This behaviour is exactly what we see for our model in Fig. 8. Also included in this figure is the experimental data of Refs. 7 and 13. Our results show a very similar temperature dependence over the temperature range studied in experiments, albeit with our results being displaced to a slightly higher force by a similar amount to that already seen at room temperature. In particular, Zhang *et al*. find a slope of -0.44 pN K^{-1} at and above room temperature,[7] which is similar to the -0.46 pN K^{-1} that we obtain for our model. Interestingly, Zhang *et al*. find a change in the sign of the slope below 20 °C which they interpret as a crossover to an S-DNA mechanism. Our results show no such change in slope, consistent with the fact that overstretching always occurs by unpeeling in our model.

One particularly interesting feature at high temperature and low force is the non-monotonic behaviour of $F_{\text{over}}(T)$. There is a narrow range of temperature just above the bulk (i.e., infinitely long strands) zero-force melting temperature where there are two transitions as a function of force. This behaviour is illustrated in the inset to Fig. 8 and also in Fig. 6(b) where there are two changes of sign in the slope of the free-energy profile as a function of force for a temperature in this window. At low and high force it is favourable for the duplex to unpeel, but there is an intermediate force range (∼2.5–10 pN at 94 °C) where duplex DNA is most stable. This type of behaviour has been predicted for a number of theoretical models[10,13,33,34,38,39,55] and occurs because which of dsDNA or ssDNA is more extensible depends on the magnitude of force. At low force dsDNA is more extensible because it has a larger persistence length and so has a smaller entropy cost for aligning with the force. However, when the extension approaches the contour length of dsDNA the force required to further extend the duplex grows rapidly and ssDNA becomes more extensible. The crossover between these two regimes occurs at approximately 6.5 pN (i.e., when the two force-extension curves in Fig. 3(a) cross). For forces below 6.5 pN, dsDNA is stabilized with respect to ssDNA by increasing force, but above this value of the force, the opposite is true. Hence, the turning point in $F_{\text{over}}(T)$ is expected to occur at this force.

One of the features in experiments when overstretching occurs by unpeeling is that the transition is not fully reversible and that on reducing the force from above F_{over} the system does not trace out exactly the same force-extension curve. One can envisage two possible sources for this lack of reversibility. First, if the unpeeled strand has formed some secondary structure (e.g., hairpins), the intramolecular base pairs involved would first have to be broken. Second, even for an unstructured unpeeled strand the reassociation dynamics have the potential to be slow, particularly if the strands are long, as is typically the case in experiments.

In Fig. 9(a) we show force-extension curves when the force is first increased and then decreased after reaching different stages in the overstretching process. If the force is decreased after overstretching when the non-force-bearing strand has dissociated, the resulting force-extension curve just follows that for ssDNA (case A), because reassociation of the separated chains does not occur. Also depicted in Fig. 9(a) are two cases when the force decrease starts roughly halfway through the overstretching transition. In both cases the extension, perhaps somewhat surprisingly, initially continues to increase. As mentioned earlier, the dynamic pulling simulations somewhat overshoot the overstretching transition. So, even though the force is decreasing, initially it is still above the equilibrium overstretching force and so the molecule continues to unpeel. However, although the extension nearly reaches that for ssDNA, the chains do not dissociate, because, as noted earlier, the last few base pairs are harder to melt. As the

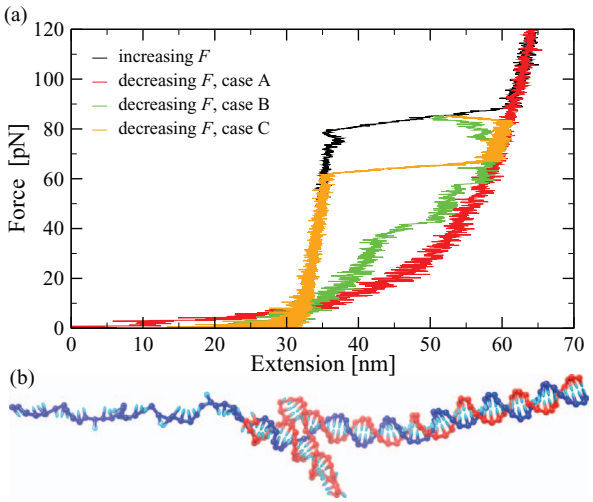

FIG. 9. (a) Force-extension curves illustrating the irreversibility in our model. In all the simulations the magnitude of the pulling rate is 2.43×10^{-6} pN/time step. The increasing force run starts from a fully bonded duplex and provides the starting configurations for the decreasing force simulations. In case A, the force is decreased after reaching 146 pN, whereas in cases B and C the force is decreased after reaching 85.3 pN. In case C intramolecular base pairing is not allowed, preventing secondary structure formation in the unpeeled chain. (b) The final zero-force configuration for case B.

force decreases further, the two cases show quite different behaviour. In case B reassociation does occur but only partially and in "bursts." Figure 9(b) depicts the final configuration that results, clearly showing that full reassociation has been blocked by secondary structure formation in the non-force-bearing strand. Similarly, the bursts occur when sections involved in secondary structure open up, allowing further reassociation.

By contrast, in case C we prevent secondary structure in the unpeeled strand by turning off intramolecular base pairing. Now when reassociation occurs, it does so relatively quickly and goes to completion, albeit at a lower force than for overstretching. Thus, for our model, secondary structure formation is a much more substantial contributor to irreversibility than the underlying dynamics of reassociation. This finding is in agreement with Gross *et al*. who found that prolonged stalls in the reannealing occurred at points at which particularly stable secondary structure was possible in the unpeeled strand.[27] However, we should note that our system is much smaller than those typically studied in overstretching experiments (for example, λ-DNA, which is often used in experiments, has 48 502 base pairs).

Having considered pulling scheme I in considerable detail, we also examine the other two pulling schemes in Fig. 2. First, for scheme II we expect a similar behaviour to scheme I—there are still two free ends that allow unpeeling—except beyond the overstretching transition, where after dissociation of the two strands the system will no longer be able to bear a force. In Fig. 10 we see that the overstretching transition occurs at approximately the same force as for scheme I and again occurs by unpeeling. Note that there are no points above the midpoint of overstretching due to this instability.

Scheme III is more interesting, because now overstretching can no longer occur by unpeeling. Two remaining possible mechanisms for overstretching are force-induced melting by bubble formation and an S-DNA-like transition to an overstretched but still associated duplex form. The former is expected to require a larger force than unpeeling because both strands are still force-bearing after bubble formation and so have less entropy than an unpeeled strand. Interestingly, in room temperature experiments at 150 mM NaCl where unpeeling cannot occur and the system is not torsionally constrained, reversible overstretching still occurs at about 65 pN.[24]

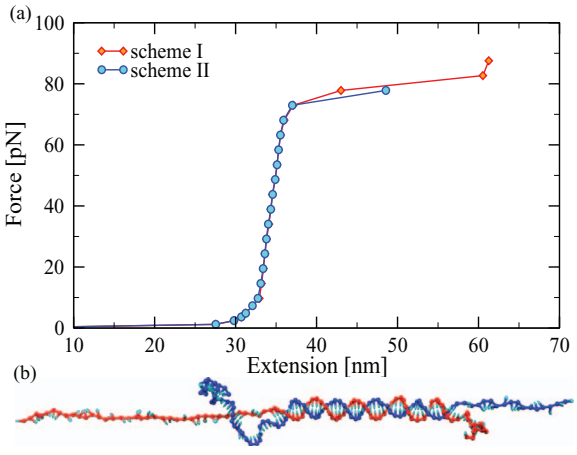

FIG. 10. (a) Force-extension curve for pulling scheme II obtained from Monte Carlo simulations compared to that for scheme I. There are no points above the midpoint of overstretching because the system can no longer bear a force after dissociation. (b) Snapshot near the middle of the transition at $F = 78$ pN.

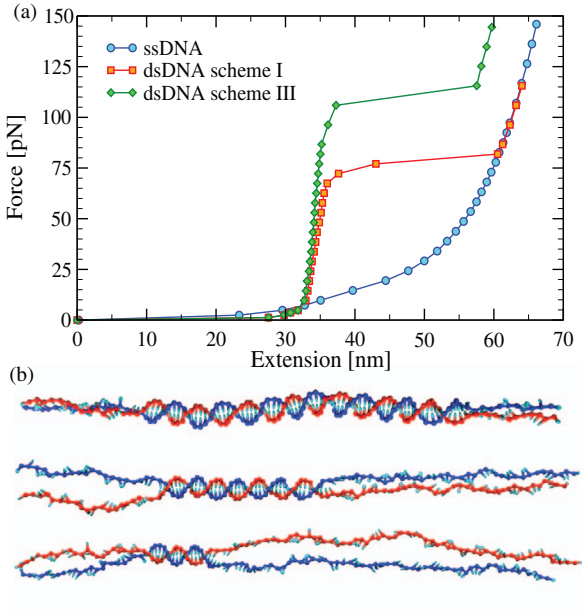

FIG. 11. (a) Force-extension curves and (b) snapshots of the DNA configuration at $F = 116$ pN as it undergoes overstretching for pulling scheme III obtained from Monte Carlo simulations. For comparison, curves for pulling ssDNA and dsDNA in pulling scheme I have been added in (a).

Figure 11(a) shows the force-extension curve for our model for pulling scheme III. The clear difference from scheme I and II is that overstretching now occurs at significantly higher force, namely at about 110–115 pN. Although overstretching can no longer occur by unpeeling, Fig. 11(b) shows that it is still by force-induced melting, but in this case by bubble formation. Notably, even though the duplex reaches higher forces than in scheme I and II, there is still no sign of a transition to an S-DNA-like overstretched form. Also, that the force-extension curve above overstretching does not follow that for ssDNA is simply because each individual strand only bears a force $F/2$ (Fig. 2(c)) and therefore the curve can be made to overlap with that for ssDNA by shifting it down by a factor of 2.

IV. CONCLUSIONS

In this paper we study the overstretching transition in a recently developed coarse-grained model of DNA, which has previously been shown to give an excellent description of the thermodynamic, structural, and mechanical properties of DNA.[64,65] At room temperature and 500 mM salt our model undergoes overstretching by unpeeling at 74 pN, just somewhat higher than the experimentally observed overstretching force and much closer than the predictions for any other coarse-grained potential model[60–63] or any atomistic

simulations.[45–54] We are also able to reproduce the temperature dependence of the overstretching force well. This ability to accurately reproduce phenomena to which it was not directly fitted provides further validation of our model.

It is also important that we understand why our coarse-grained model does so well in describing the overstretching transition. As emphasised in Sec. II C to predict force-induced melting correctly, one just needs to know the zero-force thermodynamics of melting and the force-extension curves of ds-DNA and ssDNA. As our model describes all of the above reasonably well, its success in describing overstretching by force-induced melting is perhaps not surprising.

The 6–7 pN overestimation of the room temperature overstretching force in our model is probably due to a couple of factors. First, our description of the ssDNA force-extension curve deviates slightly from fits to experiment,[32] in particular, underestimating the extension at larger forces. The net effect is to somewhat underestimate the stabilization of ssDNA by force near to overstretching. Second, the zero-force thermodynamics in our model has been designed to fit the thermodynamic predictions of the SantaLucia nearest-neighbour model[81] and so any errors in this model will be carried over to our results. For example, Huguet *et al*. have shown that using the SantaLucia model and fits to the force-extension curves of ssDNA and dsDNA gives an approximately 1 pN overestimation of DNA unzipping at room temperature.[82] As the difference in extension between the duplex and ssDNA is less for the overstretching than for the unzipping geometry, one would expect the error due to the use of the SantaLucia model to be larger for overstretching, with the above result implying an overestimation of the order of 4 pN. However, when Gross *et al*. applied a similar approach to model their experimental overstretching force-extension curves, the model actually underestimated the experimental results.[27]

Although in our simulations to locate the overstretching force, we had to turn off secondary structure in the unpeeled strand in order to equilibrate the system, we were able to estimate a correction to F_{over} from calculations of the amount by which ssDNA is stabilized by secondary structure. The correction factor for an average random sequence was 3 pN at the solution conditions for which our model has been fitted. However, for a sequence specifically designed to form strong hairpins, the effect could be much larger, and lead to a significant lowering of the overstretching force. Secondary structure in the unpeeled strand also had a significant effect on the dynamics with the associated free-energy barriers preventing the stretching simulations reaching equilibrium and leading to significant hysteresis on decreasing the force.

In all simulations where free ends are present, we see a force-induced melting mechanism for overstretching occurring by unpeeling, one of the confirmed experimental mechanisms.[14] However, no sign is ever seen of an S-DNA-like transition to an overstretched duplex form. This shortcoming leads to an overstretching force when there are no free ends (i.e., scheme III) that is much higher than experiment, because the mechanism is force-induced melting by bubble formation. The reason for our model's inability to exhibit an S-DNA-like state is harder to identify, as not much is known about the detailed structure of S-DNA. One suggestion is that S-DNA might adopt a "ladder-like" unstacked but base-paired configuration.[8,83] However, if this were the case one might expect our model to be able to reproduce this, as it describes the thermodynamics of stacking and base-pairing well. Furthermore, the reason why such a form is not stable in our model is simply that stacking provides a larger free energetic contribution to the stability of B-DNA than base pairing and that the overstretched unpeeled state retains a significant fraction of its stacking (Fig. 5). This stacking occurs both in the non-force-bearing unpeeled strand that is totally free to stack and in the force-bearing strand where short runs of stacked bases can align their backbones along the force axis. Even when there are no free ends, for which force-induced melting in our model occurs at higher force by bubble formation with less stacking retained because both strands are force-bearing, an unstacked but base-paired duplex is still not observed. Our results strongly suggest that S-DNA does not correspond to such an unstacked but base-paired configuration, but instead probably to a more exotic structure. A plausible candidate structure for S-DNA might be "zip-DNA,"[46,52] where the chains are no longer base-paired but interdigitate allowing stacking between bases on different strands. Our model does not include the possibility of this more unusual stacking interaction and so, of course, would not be able to reproduce any such behaviour.

ACKNOWLEDGMENTS

The authors are grateful for financial support from the EPSRC and University College, and for helpful discussions with Felix Ritort.

[1]C. Bustamante, Z. Bryant, and S. B. Smith, Nature (London) **421**, 423 (2003).
[2]J. P. Peters and L. J. Maher III, Q. Rev. Biophys. **43**, 23 (2010).
[3]R. Vafabakhsh, K. S. Lee, and T. Ha, Adv. Chem. Phys. **150**, 169 (2012).
[4]C. Bustamante, J. F. Marko, E. D. Siggia, and S. B. Smith, Science **265**, 1599 (1994).
[5]M. C. Williams, J. R. Wenner, I. Rouzina, and V. A. Bloomfield, Biophys. J. **80**, 874 (2001).
[6]J. R. Wenner, M. C. Williams, I. Rouzina, and V. A. Bloomfield, Biophys. J. **82**, 3160 (2002).
[7]X. Zhang, H. Chen, H. Fu, P. S. Doyle, and J. Yan, Proc. Natl. Acad. Sci. U.S.A. **109**, 8103 (2012).
[8]S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[9]P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron, Science **271**, 792 (1996).
[10]I. Rouzina and V. A. Bloomfield, Biophys. J. **80**, 882 (2001).
[11]I. Rouzina and V. A. Bloomfield, Biophys. J. **80**, 894 (2001).
[12]M. C. Williams, I. Rouzina, and V. A. Bloomfield, Acc. Chem. Res. **35**, 159 (2002).
[13]M. C. Williams, J. R. Wenner, I. Rouzina, and V. A. Bloomfield, Biophys. J. **80**, 1932 (2001).
[14]J. van Mameren, P. Gross, G. Farge, P. Hooijman, M. Modesti, M. Falkenberg, G. J. L. Wuite, and E. J. G. Peterman, Proc. Natl. Acad. Sci. U.S.A. **106**, 18231 (2009).
[15]M. C. Williams, I. Rouzina, and M. J. McCauley, Proc. Natl. Acad. Sci. U.S.A. **106**, 18047 (2009).
[16]K. R. Chaurasiya, T. Paramanathan, M. J. McCauley, and M. C. Williams, Phys. Life Rev. **7**, 299 (2010).
[17]S. Whitelam, S. Pronk, and P. Geissler, Phys. Rev. E **82**, 021907 (2010).
[18]S. Whitelam, Phys. Life Rev. **7**, 348 (2010).
[19]O. Krichevsky, Phys. Life Rev. **7**, 350 (2010).
[20]C. P. Calderon, W.-H. Chen, K.-J. Lin, N. C. Harris, and C.-H. Kiang, J. Phys.: Condens. Matter **21**, 034114 (2009).

[21]C. U. Murade, V. Subramaniam, C. Otto, and M. L. Bennink, Nucleic Acids Res. **38**, 3423 (2010).
[22]H. Fu, H. Chen, J. F. Marko, and J. Yan, Nucleic Acids Res. **38**, 5594 (2010).
[23]H. Fu, H. Chen, X. Zhang, Y. Qu, J. F. Marko, and J. Yan, Nucleic Acids Res. **39**, 3473 (2011).
[24]D. H. Paik and T. T. Perkins, J. Am. Chem. Soc. **133**, 3219 (2011).
[25]P. Bianco, L. Bongini, L. Melli, M. Dolfi, and V. Lombardi, Biophys. J. **101**, 866 (2011).
[26]M. Maaloum, A.-F. Beker, and P. Muller, Phys. Rev. E **83**, 031903 (2011).
[27]P. Gross, N. Laurens, L. B. Oddershede, U. Bockelmann, E. J. G. Peterman, and G. J. L. Wuite, Nat. Phys. **7**, 731 (2011).
[28]N. Bosaeus, A. H. El-Sagheer, T. Brown, S. B. Smith, B. Akerman, C. Bustamante, and B. Norden, Proc. Natl. Acad. Sci. U.S.A. **109**, 15179 (2012).
[29]P. Cizeau and J.-L. Viovy, Biopolymers **42**, 383 (1997).
[30]A. Ahsan, J. Rudnick, and R. Bruinsma, Biophys. J. **74**, 132 (1998).
[31]C. Storm and P. C. Nelson, Phys. Rev. E **67**, 051906 (2003).
[32]S. Cocco, J. Yan, J.-F. Léger, D. Chatenay, and J. F. Marko, Phys. Rev. E **70**, 011910 (2004).
[33]A. Hanke, M. G. Ochoa, and R. Metzler, Phys. Rev. Lett. **100**, 018106 (2008).
[34]S. J. Rahi, M. P. Hertzbert, and M. Kardar, Phys. Rev. E **78**, 051910 (2008).
[35]S. Whitelam, S. Pronk, and P. Geissler, J. Chem. Phys. **129**, 205101 (2008).
[36]S. Whitelam, S. Pronk, and P. Geissler, Biophys. J. **94**, 2452 (2008).
[37]B. Chakrabarti and D. R. Nelson, J. Phys. Chem. B **113**, 3831 (2009).
[38]D. Marenduzzo, A. Maritan, E. Orlandini, F. Seno, and A. Trovato, J. Stat. Mech.: Theory Exp. (**2009**) L04001.
[39]T. R. Einert, D. B. Staple, H.-J. Kreuzer, and R. R. Netz, Biophys. J. **99**, 578 (2010).
[40]F. A. Massuci, I. P. Castillo, and C. P. Vicente, e-print arXiv:1006.2327.
[41]C. Nisoli and A. R. Bishop, Phys. Rev. Lett. **107**, 068102 (2011).
[42]A. Fiasconaro and F. Falo, Phys. Rev. E **86**, 032902 (2012).
[43]M. Manghi, N. Destainville, and J. Palmeri, Eur. Phys. J. E **35**, 110 (2012).
[44]A. Lebrun and R. Lavery, Nucleic Acids Res. **24**, 2260 (1996).
[45]S. A. Harris, Z. A. Sands, and C. A. Laughton, Biophys. J. **88**, 1684 (2005).
[46]R. Lohikoski, J. Timonen, and A. Laaksonen, Chem. Phys. Lett. **407**, 23 (2005).
[47]J. Wereszczynski and I. Andricioaei, Proc. Natl. Acad. Sci. U.S.A. **103**, 16200 (2006).
[48]B. Luan and A. Aksimentiev, Phys. Rev. Lett. **101**, 118101 (2008).
[49]H. Li and T. Gisler, Eur. Phys. J. E **30**, 325 (2009).
[50]M. Santosh and P. K. Maiti, J. Phys.: Condens. Matter **21**, 034113 (2009).
[51]J. Řezáč, P. Hobza, and S. A. Harris, Biophys. J. **98**, 101 (2010).
[52]A. Balaeff, S. L. Craig, and D. N. Beratan, J. Phys. Chem. A **115**, 9377 (2011).
[53]M. Santosh and P. K. Maiti, Biophys. J. **101**, 1393 (2011).
[54]M. Wolter, M. Elstner, and T. Kubař, J. Phys. Chem. A **115**, 11238 (2011).
[55]D. Marenduzzo, E. Orlandini, F. Seno, and A. Trovato, Phys. Rev. E **81**, 051926 (2010).
[56]A. R. Singh, D. Giri, and S. Kumar, J. Chem. Phys. **132**, 235105 (2010).
[57]J. J. de Pablo, Annu. Rev. Phys. Chem. **62**, 555 (2011).
[58]F. Lankas, in *Innovations in Biomolecular Modelling and Simulation*, edited by T. Schlick (Royal Society of Chemistry, 2012), Vol. 2, pp. 3–32.
[59]J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Šulc, C. Matek, and B. Snodin, "Coarse-graining DNA for simulations of DNA nanotechnology," Phys. Chem. Chem. Phys. (unpublished).
[60]S. Niewieczerzał and M. Cieplak, J. Phys.: Condens. Matter **21**, 474221 (2009).
[61]A. V. Savin, M. A. Mazo, I. P. Kikot, and A. V. Onufriev, Proc. Natl. Acad. Sci. USA **110** (2013).
[62]A. Kocsis and D. Swigon, Int. J. Non-Linear Mech. **47**, 639 (2012).
[63]C. W. Hsu, M. Fyta, G. Lakatos, S. Melchionna, and E. Kaxiras, J. Chem. Phys. **137**, 105102 (2012).
[64]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, Phys. Rev. Lett. **104**, 178101 (2010).
[65]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, J. Chem. Phys. **134**, 085101 (2011).
[66]F. Romano, A. Hudson, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis, J. Chem. Phys. **136**, 215102 (2012).
[67]T. E. Ouldridge, "Coarse-grained modelling of DNA and DNA self-assembly," Ph.D. dissertation (University of Oxford, 2011); see http://tinyurl.com/7ycbx7c.
[68]T. E. Ouldridge, R. Hoare, A. A. Louis, J. P. K. Doye, J. Bath, and A. J. Turberfield, "Optimizing DNA nanotechnology through coarse-grained modelling: A two-footed DNA walker," ACS Nano (to be published).
[69]P. Šulc, T. E. Ouldridge, F. Romano, J. P. K. Doye, and A. A. Louis, "Simulating a burnt-bridges motor with a coarse-grained DNA model," Nat. Comput. (submitted); e-print arXiv:1212.4536.
[70]C. Matek, T. E. Ouldridge, A. Levy, J. P. K. Doye, and A. A. Louis, J. Phys. Chem. B **116**, 11616 (2012).
[71]C. de Michele, L. Rovigatti, T. Bellini, and F. Sciortino, Soft Matter **8**, 8388 (2012).
[72]The oxDNA program is available at http://dna.physics.ox.ac.uk.
[73]P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, and A. A. Louis, J. Chem. Phys. **137**, 135101 (2012).
[74]C. Danilowicz, C. Limouse, K. Hatch, A. Conover, V. W. Coljee, N. Kleckner, and M. Prentiss, Proc. Natl. Acad. Sci. U.S.A. **106**, 13196 (2009).
[75]S. Whitelam and P. Geissler, J. Chem. Phys. **127**, 154101 (2007).
[76]J. Russo, P. Tartaglia, and F. Sciortino, J. Chem. Phys. **131**, 014504 (2009).
[77]K. E. Duderstadt, K. Chuang, and J. M. Berger, Nature (London) **478**, 209 (2011).
[78]This could also be achieved by choosing a sequence with no complementary bases in the same strand.
[79]In some simulations with repetitive sequences (not reported here) we have seen a tendency for the unpeeled strand to rebind to the force-bearing strand in a non-native site for short stretches (short so that the helix of the force-bearing strand can still approximately align along the direction of force).
[80]D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys. **7**, 3910 (2005).
[81]J. SantaLucia, Jr. and D. Hicks, Annu. Rev. Biophys. Biomol. Struct. **33**, 415 (2004).
[82]J. M. Huguet, C. V. Bizarro, N. Forns, S. B. Smith, C. Bustamante, and F. Ritort, Proc. Natl. Acad. Sci. U.S.A. **107**, 15431 (2010).
[83]T. Thundat, D. P. Allison, and R. J. Warmack, Nucleic Acids Res. **22**, 4224 (1994).