

# Identifying Physical Causes of Apparent Enhanced Cyclization of Short DNA Molecules with a Coarse-Grained Model

Ryan M. Harrison,<sup>†</sup> Flavio Romano,<sup>‡</sup> Thomas E. Ouldridge,<sup>§</sup> Ard A. Louis,<sup>||</sup> and Jonathan P. K. Doye<sup>\*,†</sup>

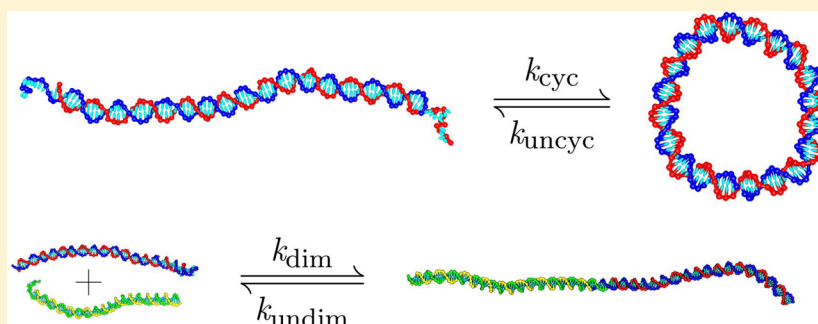
<sup>†</sup>Physical & Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

<sup>‡</sup>Dipartimento di Scienze Molecolari e Nanosistemi, Università Ca' Foscari Venezia, I-30123 Venezia, Italy

<sup>§</sup>Imperial College Centre for Synthetic Biology and Department of Bioengineering, Imperial College London, 180 Queen's Road, London SW7 2AZ, United Kingdom

<sup>||</sup>Rudolf Peierls Centre for Theoretical Physics, Department of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, United Kingdom

## Supporting Information



**ABSTRACT:** DNA cyclization is a powerful technique to gain insight into the nature of DNA bending. While the wormlike chain model provides a good description of small to moderate bending fluctuations, it is expected to break down for large bending. Recent cyclization experiments on strongly bent shorter molecules indeed suggest enhanced flexibility over and above that expected from the wormlike chain. Here, we use a coarse-grained model of DNA to investigate the subtle thermodynamics of DNA cyclization for molecules ranging from 30 to 210 base pairs. As the molecules get shorter, we find increasing deviations between our computed equilibrium  $j$ -factor and the classic wormlike chain predictions of Shimada and Yamakawa for a torsionally aligned looped molecule. These deviations are due to sharp kinking, first at nicks, and only subsequently in the body of the duplex. At the shortest lengths, substantial fraying at the ends of duplex domains is the dominant method of relaxation. We also estimate the dynamic  $j$ -factor measured in recent FRET experiments. We find that the dynamic  $j$ -factor is systematically larger than its equilibrium counterpart—with the deviation larger for shorter molecules—because not all the stress present in the fully cyclized state is present in the transition state. These observations are important for the interpretation of recent cyclization experiments, suggesting that measured anomalously high  $j$ -factors may not necessarily indicate non-WLC behavior in the body of duplexes.

## I. INTRODUCTION

The mechanics of DNA plays an important role in its biological capacities: looping in the Lac operon regulates gene expression;<sup>1</sup> DNA supercoiling is part of the circadian cycle in cyanobacteria;<sup>2</sup> and the dynamically variable wrapping of DNA around histone proteins.<sup>3</sup> Consequently, there has been much activity to accurately characterize these properties, not only in the elastic regime of small to moderate fluctuations but also for more strongly stressed systems. For example, DNA is found to overstretch beyond a salt-dependent critical force.<sup>4</sup> Similarly, in response to twist DNA forms plectonemes beyond a critical buckling superhelical density.<sup>5</sup> Here, we are interested in the response of DNA to strong bending.

The wormlike chain (WLC) model provides a good description of small-to-moderate bending fluctuations in DNA.<sup>6–9</sup> However, although there is a consensus that for sufficiently strong bending the stress will be localized within small regions, often termed “kinks”, much about this crossover to non-WLC behavior remains controversial. For example, a recent review by Vologodskii et al.<sup>10</sup> highlighted a number of open questions, including what is the free energy cost of kink formation, how does the free-energy of a kink depend on bend angle, what is the critical curvature that causes the double helix

Received: February 3, 2019

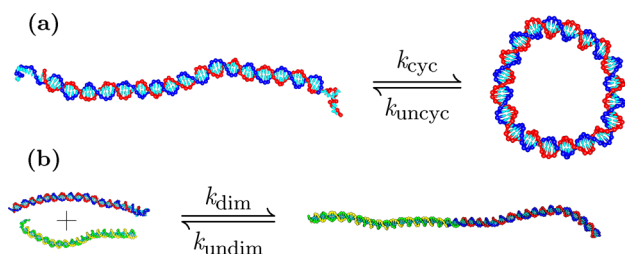
Published: July 8, 2019

to kink? These questions are particularly subtle because kinking is postulated to arise in a range of distinct contexts, involving a combination of stretching,<sup>11,12</sup> bending,<sup>13–27</sup> and torsional stresses,<sup>11–13,16,20,22,24,25,27</sup> as well as for specific sequences in relaxed conditions.<sup>28</sup> In this work, we use oxDNA,<sup>29–31</sup> a coarse-grained model of DNA, to probe whether recent data on the cyclization of short molecules is indicative of non-WLC behavior, and if so, whether kinking plays a role.

**I.A. DNA Cyclization.** DNA cyclization is a convenient model system used to probe DNA bending. Cyclization experiments were first reported in 1966, albeit on 48 500 base pair (bp)  $\lambda$ -DNA.<sup>32,33</sup> In 1981, Shore et al. developed a method to probe the bending of shorter 126–4361 bp fragments,<sup>34</sup> later noting periodicity in the cyclization efficiency of 237–254 bp fragments.<sup>6</sup>

More recently, there has been a particular interest in probing the cyclization of sub-persistence-length DNA, to explore whether this regime is accurately described by the WLC model. For example, in 2004 Cloutier and Widom (C&W)<sup>35</sup> challenged the conventional wisdom of WLC flexibility established by Shore et al.,<sup>6,34</sup> claiming much greater than predicted cyclization efficiency in 93–95 bp DNA fragments. This controversial finding spurred debate on the characteristic length at which DNA cyclization efficiency deviates from the predictions of the WLC model. Despite much experimental<sup>8,18,20,23,35–37</sup> and theoretical effort,<sup>10,24,27,38–55</sup> a consensus has not yet been established.

A typical cyclization experiment, as depicted in Figure 1a, uses a cyclization substrate with complementary sticky ends,  $N_s$



**Figure 1.** Schematic representations of (a) cyclization where  $k_{cyc}$  and  $k_{uncyc}$  are the forward and reverse rate constants, respectively, and (b) dimerization, where the rate constants are  $k_{dim}$  and  $k_{undim}$ . Note that, for the dimerization system, there is only one complementary sticky end per monomer, with the other end being blunt to allow for only one reaction product, a linear dimer. Figures are oxDNA representations for monomers of length  $N_{bp} = 101$ , including complementary sticky ends of length  $N_s = 10$ . (Dimer length is  $2N_d + N_s$ ).

bases in length, on both ends of a  $N_d$  base-pair duplex. Cyclization leads to the formation of a  $N_{bp}$ -base-pair duplex, where  $N_{bp} = N_s + N_d$ . The resultant structure is not a closed minicircle—two backbone “nicks” are present at either end of the sticky ends. Either the forward rate or equilibrium constant of the cyclization reaction is reported. Experiments differ in how exactly they probe cyclization: methods based on ligation,<sup>6,8,34,35</sup> FRET,<sup>18,20,37,53</sup> and multimerization<sup>56</sup> have been reported.

In ligase-based experiments, cyclized molecules are fixed in the cyclized state by ligation of the two backbone nicks. The open and ligated cyclized molecules can then be resolved by gel electrophoresis, and the concentration of different products

can be measured. FRET-based experiments can be performed in equilibrium, with the molecules allowed to cyclize and uncyclize indefinitely. Fluorophores are attached to both ends of the molecule as FRET reporters: a high FRET signal will be reported when the duplex ends are in close proximity (cyclized) and low FRET when apart (open).

Non-WLC behavior has been suggested by the ligase-based experiments of C&W<sup>35</sup> and the FRET-based experiments of Vafabakhsh and Ha (V&H),<sup>18</sup> but there is not yet a consensus on whether these experimental results clearly indicate non-WLC behavior within the body of a typical DNA duplex. Some authors have used these data to support models that incorporate non-WLC flexibility within the body of a duplex, both with<sup>39–41,54,55</sup> and without<sup>45–47,49,51</sup> sharp local disruptions of the helix structure. Others have instead criticized the original work of both C&W<sup>8,10</sup> and V&H,<sup>10,52,53,57</sup> pointing to the possibility of enhanced flexibility at duplex ends that is not reflective of the bulk WLC behavior,<sup>10,52,57</sup> the possible existence of defects in synthetically prepared strands,<sup>10</sup> and inappropriate interpretation of the experimentally measured quantities.<sup>8,10,53</sup>

To understand these issues, we first consider the  $j$ -factor, introduced by Jacobson and Stockmayer,<sup>58</sup> which is a measure of the effective local concentration of duplex ends at zero end-to-end separation. The  $j$ -factor enables the ring closure probability to be calculated, and importantly, may be related to a ratio of equilibrium constants

$$j\text{-factor} = j_{eq} \equiv \frac{K_{eq}^{cyc}}{K_{eq}^{dim}} \quad (1)$$

where  $K_{eq}^{cyc}$  and  $K_{eq}^{dim}$  are the equilibrium constants for cyclization and dimerization, respectively. These equilibrium constants are simply related to the rate constants in Figure 1, where  $K_{eq}^{cyc} = k_{cyc}/k_{uncyc}$  and  $K_{eq}^{dim} = k_{dim}/k_{undim}$ .

Typical experiments seek to probe cyclization and dimerization processes and, hence, obtain an estimate of the  $j$ -factor that can be compared to a polymer model. This is often done by mixing many cyclization substrates at finite concentration and observing a competition between linear and circular products.<sup>8,35</sup> It is also possible to study a separate dimerization substrate with the same sequence as the cyclization substrate, but with only one  $N_s$ -base complementary sticky end per molecule<sup>18</sup> (Figure 1b). Following hybridization, the total length of the system is then  $2N_d + N_s$  base pairs, with blunt as opposed to sticky ends. The consequences of this choice for  $K_{eq}^{dim}$  and, hence,  $j_{eq}$  are discussed in section S1.C.

Some of the confusion surrounding claims of cyclization efficiency greater than that predicted by the WLC model revolves around the use and interpretation of  $j$ -factors. While the  $j$ -factor relation using the ratio of cyclization to dimerization equilibrium constants is well established, experimental studies usually report the ratio of forward rate constants. In the case of ligase-based assays (reviewed in ref 59)

$$j_{dyn}^{ligase} = \frac{k_{cyc}^{ligase}}{k_{dim}^{ligase}} \quad (2)$$

where  $k_{cyc}^{ligase}$  and  $k_{dim}^{ligase}$  are the forward rate constants for the formation of the ligated circle and dimer, respectively.

In the experimental limit where the ligation rate is very slow compared to the rate constants for uncyclization ( $k_{\text{uncyc}}$ ) and undimerization ( $k_{\text{undim}}$ ), the concentrations of unligated circles and dimers will reach an equilibrium with reactants. If this condition is met,  $j_{\text{dyn}}^{\text{ligase}}$  should be equivalent to  $j_{\text{eq}}$ . In practice, this limit is valid for low ligase concentrations ( $[\text{Ligase}] \ll [\text{DNA}]$ ) and short sticky ends.<sup>8,42</sup> The importance of this condition is illustrated by Du et al.,<sup>8</sup> who suggested that the apparent non-WLC behavior in the C&W experiments was rather due to an insufficiently low ligase concentration (reviewed in ref 10.).

The FRET experiments of V&H<sup>18</sup> have the advantage of directly monitoring the transition between cyclized and open states. While it is possible to report thermodynamics from FRET experiments, V&H report dynamics:

$$j_{\text{dyn}}^{\text{FRET}} = \frac{k_{\text{cyc}}}{k_{\text{dim}}} \quad (3)$$

where  $k_{\text{cyc}}$  and  $k_{\text{dim}}$  are the forward rate constants for the formation of the unligated circle and dimer, respectively. In fact, ref 18 is ambiguous as to whether the data reported is  $j_{\text{dyn}}^{\text{FRET}}$  as defined in eq 3 or  $(k_{\text{cyc}} + k_{\text{uncyc}})/k_{\text{dim}} \geq j_{\text{dyn}}^{\text{FRET}}$  with  $k_{\text{uncyc}}$  being the rate constant for opening up of the cyclized state.<sup>18,53</sup> The latter quantity is not obviously a good measure of cyclization efficiency, and we shall assume that V&H indeed report  $j_{\text{dyn}}^{\text{FRET}}$ . The difference between the two becomes relevant for the shortest systems in which  $k_{\text{uncyc}} > k_{\text{cyc}}$  and we will address the potential consequences of this discrepancy later in the Discussion.

Dynamic  $j$ -factors extracted from FRET-based experiments must also be interpreted with care.<sup>53</sup> By comparing  $j_{\text{dyn}}^{\text{FRET}}$  and a prediction for  $j_{\text{eq}}$  from the WLC,  $j_{\text{eq}}^{\text{WLC}}$ , V&H make a claim of much greater than WLC flexibility at  $N_{\text{bp}} \lesssim 100$  bp. However,  $j_{\text{dyn}}^{\text{FRET}} \approx j_{\text{eq}}$  only in the limit where  $k_{\text{uncyc}} \approx k_{\text{undim}}$ , a condition that V&H, as well as another more recent FRET experiment,<sup>20</sup> have shown not to be met.<sup>53</sup> Given that  $k_{\text{uncyc}} \neq k_{\text{undim}}$ , one should not expect  $j_{\text{dyn}}^{\text{FRET}} \approx j_{\text{eq}}$  at short  $N_{\text{bp}}$ . Thus, the observed deviation of  $j_{\text{dyn}}^{\text{FRET}}$  from  $j_{\text{eq}}^{\text{WLC}}$  is not necessarily an indication of non-WLC behavior.

Even given a well-measured experimental  $j$ -factor, inferring whether that quantity is indicative of non-WLC behavior within the body of a duplex is challenging. By assuming the contribution from base pairing to  $K_{\text{eq}}^{\text{cyc}}$  and  $K_{\text{eq}}^{\text{dim}}$  is the same, the WLC model can be used to estimate  $j_{\text{eq}}$ . If the cyclized state is fully stacked, with coaxial stacking across the two nicks, the analytic expression derived by Shimada and Yamakawa (SY),<sup>60</sup> which includes both the bending energy cost of bringing the two ends together and the twist energy cost of bringing the two helix ends into register, can be used to predict the  $j$ -factor and compare to experiment.

In ligase-based experiments, however, it is unclear whether the ligase will act equally on all nicked duplexes or only the subset that happen to adopt the right configuration at the nick, be that coaxially stacked or kinked, to allow the ligase to bind. Similarly, FRET experiments involve nicked duplexes, which have the potential to break coaxial stacking and alleviate stress,<sup>10,52,57</sup> and successful hybridization can be nucleated by sticky ends that bridge a relatively large gap between duplex ends.<sup>10,57</sup> To account for these subtleties, the Shimada and Yamakawa (SY) prediction<sup>60</sup> can be modified by weakening the constraints on the looped duplex ends when calculating WLC-based  $j$ -factors.<sup>10,18,20,57</sup> However, the resultant addi-

tional free parameters make it difficult to say when a WLC-model for the bulk duplex really fails to predict the observed behavior, particularly, because the microscopic states responsible for the putative non-WLC flexibility cannot be directly observed.

In this work, we use a coarse-grained model, oxDNA,<sup>29–31</sup> to tackle these problems. We present a broad set of simulations that allow full exploration of the V&H experiments and their interpretation in terms of non-WLC behavior. We calculate  $j_{\text{eq}}$  and  $j_{\text{dyn}}^{\text{FRET}}$  for 81 duplex lengths and investigate the typical configurations adopted by the oxDNA model in these systems. We, thereby, probe whether the experimental results of V&H are truly indicative of non-WLC behavior and whether any such behavior can be attributed to certain types of structure, such as kinks. We find that we are able to reproduce the anomalously high values of  $j_{\text{dyn}}^{\text{FRET}}$  observed by V&H without significant non-WLC behavior in the bulk of the duplex for all but the smallest systems considered. Instead, the apparent deviation from WLC behavior is largely due to a combination of kink formation at the nicks in the DNA backbone and the difference between  $j_{\text{dyn}}$  and  $j_{\text{eq}}$ .

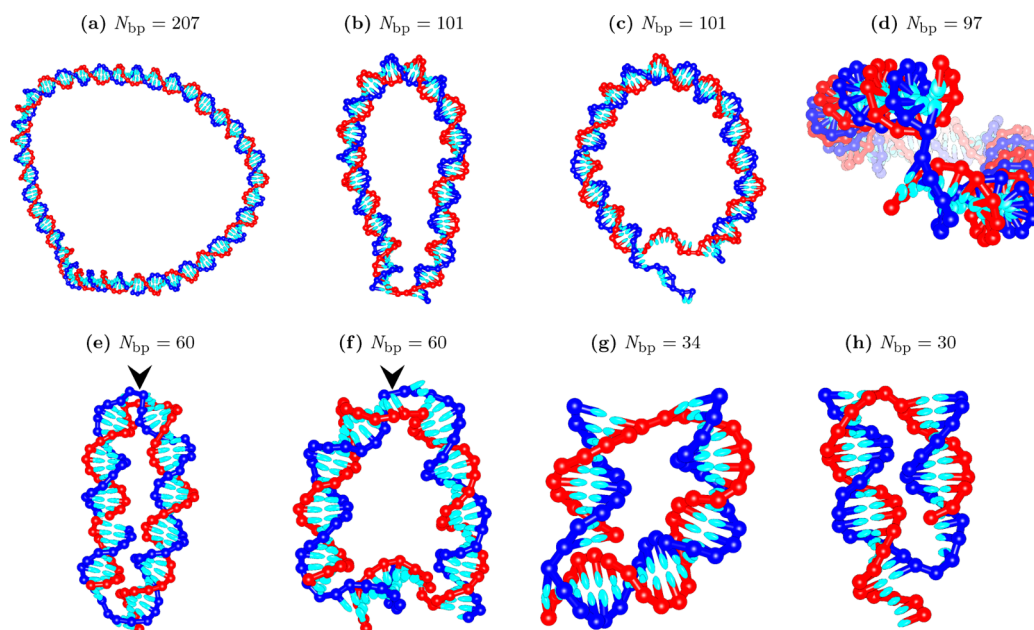
## II. METHODS

**II.A. oxDNA Model.** OxDNA<sup>29–31</sup> is a nucleotide-level coarse-grained model of DNA that has been employed successfully for a wide variety of systems,<sup>61</sup> beginning with the thermodynamic and structural characterization of DNA nanotweezers.<sup>62</sup> It is the most widely used model at this level of coarse-graining and is particularly well-suited for the current task because it reproduces well the thermodynamics of DNA hybridization, the elastic mechanical properties of double-stranded DNA,<sup>29</sup> and the mechanical properties of single-stranded DNA.<sup>31</sup> Briefly, the model consists of rigid nucleotides with three interaction sites per nucleotide, interacting via Watson–Crick base pairing, base stacking, excluded volume, and a potential to represent backbone connectivity. The model is parametrized to reproduce the thermodynamics of duplex melting at high-salt ( $[\text{Na}^+] = 500$  mM), where backbone–backbone electrostatic repulsion is short-ranged due to counterion screening. We use the average-base parametrization,<sup>29,30</sup> in which the strength of the base-pairing and stacking interactions are independent of the identity of the bases, to highlight the basic thermodynamics of DNA cyclization. By contrast, the sequence-dependent parametrization<sup>31</sup> has stacking and base-pairing interactions that depend on the base identity and is used to compare more directly to the V&H experiments. Note that, for simplicity, we do not use the more computationally expensive “oxDNA2” model,<sup>63</sup> as we are not concerned with grooving and only consider results for a single salt concentration.

OxDNA is parametrized to reproduce features associated with the single- to double-stranded transition in DNA, and was not explicitly designed with kinking transitions in mind. Nonetheless, the model captures the onset of kinking in a “molecular vice”,<sup>19,26</sup> and successfully predicted the colocalization of kinked regions and plectoneme loops in stretched and twisted systems,<sup>11,12</sup> as well as describing a range of other stress-induced structural transitions.<sup>64–66</sup>

A number of alternative mesoscopic and statistical models have been proposed explicitly for the purpose of describing apparent non-WLC behavior in DNA cyclization.<sup>20,39,41,45,47,49,51,54</sup> All of these approaches share, with oxDNA, the advantage that model configurations can be





**Figure 2.** OxDNA representations of different cyclized configurations. Kinks in the duplex, which disrupt stacking and induce a 1–3 bp bubble, are indicated with an arrow. All configurations have  $N_s = 10$ . (a) A fully stacked “circle”. (b) A “teardrop” configuration with a kink at one of the nicks. (c) A  $Q_{bp} = 1$  “transition state” configuration. (d) Teardrop configurations with  $N_{bp} \approx (n + 1/2) \times \text{pitch length}$  can reduce the stress associated with chain continuity at the nick by out-of-plane bending. Configurations with a kink in the duplex and either (e) a kink at one of the nicks or (f) kinks at both nicks. For short duplexes, where  $N_d$  is not that much larger than  $N_s$ , the sticky ends can associate either by (g) relatively minor bending of the duplex or (h) fraying a few base pairs.

tracked, and hence, the underlying physical cause of model behavior can be unambiguously determined. There are major advantages, however, in using a more general model, such as oxDNA, that has been designed and validated for other contexts. First, if a model can *predict* experimental behavior without free parameters, it provides stronger support for the underlying biophysical explanation than when a model is explicitly designed with the purpose of reproducing a phenomenon. Second, ambiguities, such as how to set the looping criteria, are resolved since the constraints imposed by cyclization arise naturally from the underlying biophysics of the model. Third, oxDNA is designed around basic and well-known structural, mechanical, and thermodynamic properties of single- and double-stranded DNA. Exploring cyclization with oxDNA is, therefore, a test of whether cyclization (and any concomitant kinking) can be explained within the framework of well-understood DNA biophysics. Finally, oxDNA can be applied across the range of contexts in which kinking has been proposed to test whether these observations are mutually consistent. Without a single model that can be applied to all systems, this comparison is extremely hard.<sup>26</sup>

**II.B. Simulations.** Simulations of the cyclization equilibria in the oxDNA representation of the V&H system, for a range of substrate lengths, were performed with a virtual-move Monte Carlo (VMMC) algorithm<sup>67</sup> at 298 K. As the free-energy barrier between typical open and cyclized states is large, the transition between the two macrostates constitutes a rare-event. Umbrella sampling, a technique that allows for the biased sampling of states with respect to an order parameter,<sup>68</sup> was employed to sample the barrier crossing in reasonable computational time.

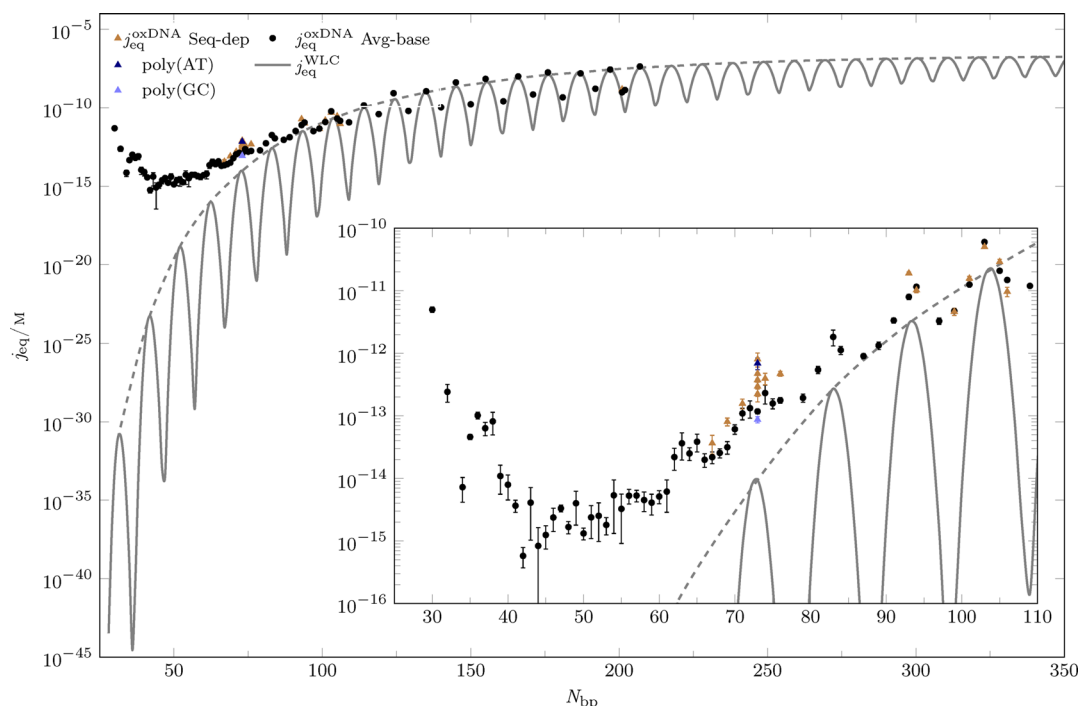
We use a two-dimensional order parameter  $Q = (Q_{ee}, Q_{bp})$  to characterize the transition.  $Q_{ee}$  is a discretized measure of the distance of closest approach between the complementary

sticky ends, with each value corresponding to the range of distances given in section S1.A.  $Q_{bp}$  is the number of base pairs formed between complementary sticky ends, where  $0 \leq Q_{bp} \leq N_s$ .

To further improve computational efficiency, the umbrella sampling was broken into two “windows”, which separately sampled the open and cyclized states of each molecule. For the window associated with the open state, the system was restricted to  $Q_{bp} = 0$ ; for the window associated with the cyclized state, the system was restricted to  $Q_{ee} = Q_{ee}^{\min}$  (the value corresponding to the shortest range of distances between the sticky ends). Simulations were run until convergence to within  $\pm 5\%$  for each window.

The two sampling windows overlap only at  $Q_{ee} = Q_{ee}^{\min}$ ,  $Q_{bp} = 0$  (note that the region of order parameter space with  $Q_{ee} > Q_{ee}^{\min}$  and  $Q_{bp} > 0$  is unfeasible). The results were combined by normalizing each window so that the free energies were equal for this value of the order parameter. As there is only one well-defined overlap between the values of the order parameters for both windows, more complex approaches, such as the weighted histogram analysis method,<sup>69,70</sup> were unnecessary. To further simplify sampling, we forbade the formation of base pairs that are not intended in the design of the system (non-native base pairs). For short strands that are not designed with competing non-native structures, this simplification is a good approximation in the oxDNA model. To compute the equilibrium constants, we deemed all states with  $Q_{bp} \geq 1$  to contribute to the cyclized states. Further details of the cyclization simulations can be found in section S1.A.

Simulations of dimerization were also performed to obtain estimates of  $j$ -factors from oxDNA. The simulation of dimerization equilibrium is roughly analogous to cyclization and is elaborated in section S1.B. A complete list of sequences



**Figure 3.** Measured values of  $j_{\text{eq}}^{\text{oxDNA}}$  for the oxDNA average-base parametrization (black circles) as a function of  $N_{\text{bp}}$  for  $N_s = 10$ . For comparison, the Shimada and Yamakawa (SY) WLC prediction<sup>60</sup>  $j_{\text{eq}}^{\text{WLC}}$  (gray solid line) is plotted using values of torsional stiffness ( $4.75 \times 10^{-28} \text{ J m}^{-1}$ ), persistence length (41.82 nm) and pitch length (10.36 bp/turn) appropriate to oxDNA.<sup>11,29</sup> The dashed line gives the maxima envelope for the SY prediction (gray dashed line). Sequence-variation in  $j_{\text{eq}}^{\text{oxDNA}}$ , computed using oxDNA's sequence-dependent parametrization, is illustrated for sequences considered by V&H (brown triangles), including six at  $N_{\text{bp}} = 73$  bp, as well as poly-AT and poly-GC.

is available in Table S1. Error bars represent the standard error of the mean from 5 independent simulations.

**II.C. Structural Analysis.** For sufficiently strong bending stress, localized structural disruptions to the DNA double-helical structure are expected. We define three such disruptions, namely, fraying, bubble formation, and kinking, which are elaborated in detail in ref 26. Briefly, both fraying and bubble formation involve the breaking of base pairs; the difference is in the location along the duplex. Fraying involves disruption of base pairing at the duplex ends, while bubbles occur in otherwise fully base-paired contiguous stretches away from the duplex ends.

Conceptually, a kink is an area of strong bending localized to a small segment of DNA and can occur both at a nick and within the duplex. When kinking occurs in a duplex region, it is nearly always accompanied by bubble formation. Similarly, kinking at a nick can be accompanied by fraying. We do not attempt to distinguish different types of kinks as has been done when analyzing atomistic molecular dynamics simulations.<sup>16,71</sup> Instead, we simply define a kink as present within a duplex region when there is a change in orientation of consecutive bases of greater than  $90^\circ$  on either strand. For nicked regions, only consecutive bases on the intact (un-nicked) strand are considered. Although this cutoff is somewhat arbitrary, since in the current system kinks arise to localize bending stress, they are usually very sharply bent and the criterion works well.<sup>26</sup> Sometimes, however, it gives rise to false negatives, particularly in the case of kinks at nick sites and when fraying is present. Nonetheless, it is satisfactory as an indicator of behavior for our purposes. More details, along with subtleties related to kinking at a nick, are discussed in section S1.D.

### III. RESULTS

We simulate a large range of system sizes; some illustrative configurations are shown in Figure 2. Quantitatively, we first consider the behavior of  $j_{\text{eq}}$  as a function of length. For the dimerization system we only computed the equilibrium constant  $K_{\text{eq}}^{\text{dim}}$  for a few lengths (monomers  $N_{\text{bp}} = 30, 67, 73, 101$ ). As expected, we found  $K_{\text{eq}}^{\text{dim}}$  to be length-independent to within numerical error (section S2.A); therefore, we use an average value of  $K_{\text{eq}}^{\text{dim}} = 0.92 \pm 0.20 \times 10^{12} \text{ M}^{-1}$  in our  $j$ -factor calculation. In contrast, we found  $K_{\text{eq}}^{\text{cyc}}$  and, thereby,  $j_{\text{eq}}$  (eq 1) to vary substantially with length.

In Figure 3, we show  $j_{\text{eq}}^{\text{oxDNA}}$  values, calculated from our measured values of  $K_{\text{eq}}^{\text{cyc}}$  and  $K_{\text{eq}}^{\text{dim}}$  using eq 1, for 81 different lengths in the range  $N_{\text{bp}} = 30\text{--}207$  for fixed  $N_s = 10$  using the average-base parametrization of oxDNA. These results are compared to  $j_{\text{eq}}^{\text{WLC}}$  predictions based on the Shimada and Yamakawa (SY) expression<sup>60</sup> using previously calculated values for the relevant structural and mechanical properties of oxDNA.<sup>29</sup> The SY expression is appropriate for the formation of a fully stacked “circle” configuration with coaxial stacking at both nicks (Figure 2a). Note that the comparison in Figure 3 is fit-free. The effect of varying  $N_s$  as well as the role of nicks and mismatches are discussed in section S2.C.

The behavior of the SY expression is well understood. In the regime of interest, shortening  $N_{\text{bp}}$  tends to make cyclization less favorable as a result of increased bending stress within the duplex. This effect becomes particularly acute for DNA lengths below the persistence length (41.82 nm/126 bp for the curve plotted in Figure 3). On top of this systematic behavior, a periodic oscillation is associated with the need to over- or under-twist the duplex when the natural twist is not commensurate with that required to form a closed circle.

The magnitude of this oscillation increases at shorter  $N_{bp}$  because a stronger twist per base pair is required.

When considering the behavior of  $j_{eq}^{oxDNA}$  in light of the SY expression, three length-scale dependent regimes become apparent: long ( $N_{bp} \gtrsim 80$ ), intermediate ( $N_{bp} \approx 45\text{--}80$  bp), and short ( $N_{bp} \lesssim 45$ ).

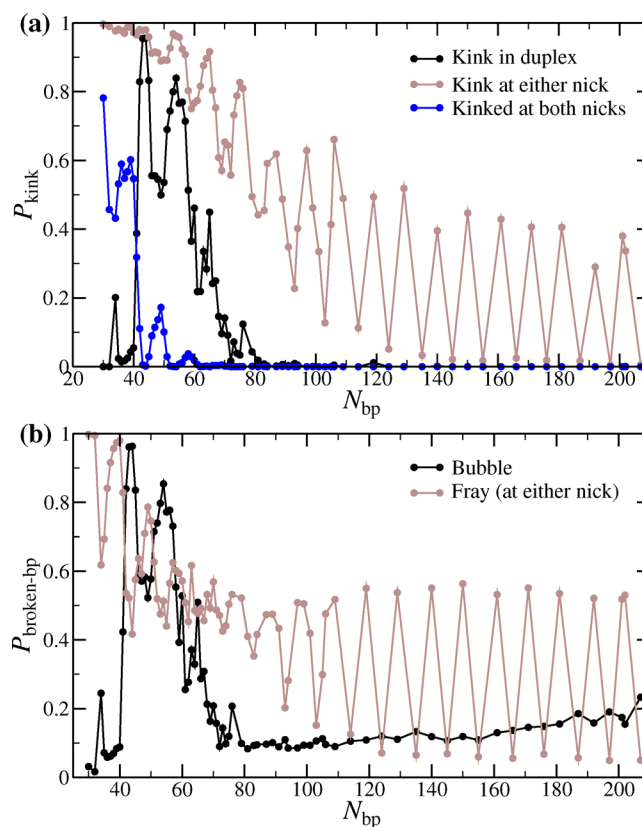
In the long length regime ( $N_{bp} \gtrsim 80$ ), oxDNA reproduces the periodic oscillations predicted by the SY expression, and values of  $j_{eq}$  coincide at the maxima of these oscillations. However, even at the longest lengths we consider, the magnitude of the oscillation is smaller for oxDNA than predicted by the SY expression. At shorter  $N_{bp}$ , the magnitude of the oscillation decreases for oxDNA, at odds with the SY expression. The difference stems from the possibility of adopting alternative “teardrop” configurations (Figure 2b), in which most of the twisting stress, and some of the bending stress can be relieved by kinking at one of the nicks, thus breaking coaxial stacking. In these configurations, there is still the constraint of DNA continuity at the nick and this can now be more easily satisfied, rather than through over- or undertwist, by out-of-plane bending (Figure 2d).

The possibility of adopting this alternative teardrop configuration reduces the free-energy penalty for incommensurate values of  $N_{bp}$  (i.e.,  $(n + 1/2) \times$  pitch length), thus suppressing the oscillations. Further, at shorter  $N_{bp}$ , the bending stress increases more slowly for these kinked teardrop configurations than for the coaxially stacked circles. Thus, the difference in  $j_{eq}$  between the “on-register” (coaxially stacked) and “off-register” (kinked at a nick) molecules decreases at shorter  $N_{bp}$ , rather than increasing as predicted by the SY model. Although not taken into account for most analyses of cyclization, the possibility of the cyclized molecule exhibiting a “teardrop” configuration has previously been suggested by Vologodskii et al.<sup>10</sup>

Further evidence in support of this analysis is given in Figure 4a, which shows the probability of kinking as a function of  $N_{bp}$ . At long lengths ( $N_{bp} \gtrsim 80$ ), kinking does not occur in the duplex regions but can occur at the nick sites. There is clear periodicity in kinking at a nick. For example, at  $N_{bp} = 145 \approx 14 \times$  pitch length, the probability of kinking at either of the nicks is negligible and the system virtually always adopts a coaxially stacked circle configuration; however, at the longer  $N_{bp} = 201 \approx 19.5 \times$  pitch length, the probability is  $\sim 40\%$ . As  $N_{bp}$  is shortened, the probability of kinking at a nick gradually increases for these “off-register” lengths; it is not until  $N_{bp} = 114 \approx 11 \times$  pitch length that the bending stress along the duplex is sufficient to cause  $\sim 10\%$  of molecules to kink at a nick for an “on-register” length.

In the intermediate-length regime ( $N_{bp} \approx 45\text{--}80$  bp), we observe enhanced cyclization efficiency compared to the SY prediction. Although  $j_{eq}^{oxDNA}$  continues to decrease with  $N_{bp}$ , it does so more gradually than the SY expression would predict; consequently,  $j_{eq}^{oxDNA}$  is in excess of the peak envelope of the SY expression. In this regime, bending stress in the circular coaxially stacked configuration is sufficiently large that kinking at one of the nicks occurs for even the on-register systems (Figure 4a).

Oscillations due to on-register effects also seem to contribute to  $j_{eq}^{oxDNA}$  in the upper end of the intermediate-length regime, with shallow maxima occurring at  $N_{bp} = 63$  and  $N_{bp} = 74$ , approximately 6 and 7 times the pitch length. However, these sizes no longer correspond to minima in the probability of kinking at a nick (Figure 4a), so the structural



**Figure 4.** (a) Probability of kinking  $P_{kink}$  as a function of length  $N_{bp}$ , in the duplex, at either nick and at both nicks. (b) Probability of broken base pairs  $P_{broken-bp}$  as a function of length  $N_{bp}$  for fraying (base-pairing disruption at either nick) and bubble formation (base-pairing disruption in the duplex region).

underpinnings of these variations is less clear. For shorter lengths, although there are size-dependent variations in  $j_{eq}^{oxDNA}$  (e.g., maxima at  $N_{bp} = 43$  and  $49$ ), there is no longer a simple relationship to the pitch length, instead reflecting more complex geometric compatibilities that allow cyclized states at these lengths to be particularly stable compared to nearby lengths.

In the intermediate regime, most cyclized molecules are kinked at one of the two nicks. At shorter  $N_{bp}$ , the bending stress in the duplex region of the teardrop configurations increases, with the highest curvature being localized opposite the nick that is kinked. Consequently, it becomes increasingly favorable to localize bending stress into a kink in the duplex (un-nicked) region, with the probability of this duplex kinking (Figure 2e) increasing from near zero at  $N_{bp} = 81$  to near one at  $N_{bp} = 43$  (Figure 4a). A kink in the duplex will generally be located opposite a kink at a nick because this arrangement minimizes the residual bending stress in the unkinked portions of the duplex by equalizing the lengths of the double-helical segments between the two nicks.<sup>25,27</sup> In addition to a loss of stacking, kinking in the duplex typically involves breaking 1–2 base pairs.<sup>26</sup> A typical configuration for such a kinked duplex state is shown in Figure 2e. As with kinking at the nick, the availability of this configuration lowers the free-energy cost of cyclization, raising  $j_{eq}^{oxDNA}$  further above the SY prediction.

At the lower end of the intermediate-length regime, the cost of bending without kinking in the duplex region is so high that molecules with duplex kinking dominate the cyclized state. States with two nicks localize the majority of the bending stress



at the kinking sites, as shown in Figure 2e, with long duplex sections relatively relaxed. Consequently, the free-energy cost of looping is largely independent of  $N_{bp}$  in this regime, causing  $j_{eq}^{oxDNA}$  to level off and reach an approximately constant value. In contrast, the SY  $j_{eq}^{WLC}$  prediction decreases very rapidly. For example, at  $N_{bp} = 43$ ,  $j_{eq}^{oxDNA}$  is  $10^9$  times greater than the SY prediction.

There exists a rich landscape of structures informed by slight differences in local geometries as a function of  $N_{bp}$ . In addition to the canonical two-kink structures, containing one kink in the duplex and one kink at a nick (Figure 2e), we observe several nontrivial arrangements, albeit with relatively low probability. For example, the configuration in Figure 2f contains a kink at both nicks, as well as in the duplex.

In the short-length regime ( $N_{bp} \leq 42$ ), we observe an increase in  $j_{eq}^{oxDNA}$ , in stark contrast to the rapidly decreasing  $j_{eq}^{WLC}$  predicted by the SY expression. Given the many WLC assumptions that are violated at this length scale, a deviation is unsurprising; however, we did not anticipate an increase in  $j_{eq}^{oxDNA}$ .

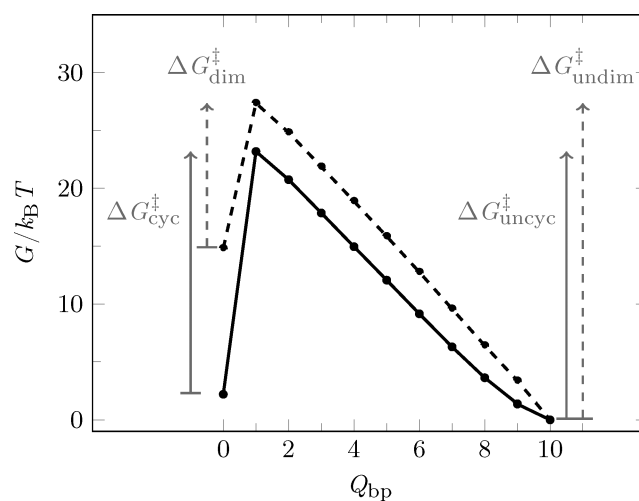
As  $N_d$  is now not that much larger than  $N_s$ , kinking at both nicks allows the single-stranded sticky ends to hybridize without duplex kinking. The system now generally adopts a conformation of two parallel duplexes, with the stress now borne by a mixture of continuous bending (Figure 2g) and fraying of a few base pairs at the ends of the duplexes (Figure 2h). Figure 4a shows a very clear crossover to cyclized states with kinks at both nicks and no kinks present in the duplex, occurring abruptly at  $N_{bp} = 40$ – $42$ . At the same time, fraying at the duplex ends also increases (Figure 4b). In this regime, as  $N_{bp}$  is shortened, the difference in length of the duplexes decreases and the stress in the system tends to drop, leading to higher  $j_{eq}$  (Figure 3). Similar behavior was also observed in an analysis of DNA loops closed by rigid protein linkers.<sup>72</sup> Overlaid on this overall trend are nontrivial geometric effects associated with whether the lengths of the two duplexes relative to the pitch length is convenient for connecting them (the thickness of the double helix is now significant compared to the duplex lengths), which leads to nonmonotonic behavior of  $j_{eq}$  and kinking. Note that the value of  $N_{bp}$  at the crossover between the short- and intermediate-length regimes is expected to be very dependent on  $N_s$ , occurring at smaller  $N_{bp}$  for smaller  $N_s$ .

So far, we have only reported results using the oxDNA average-base parametrization, in which the strength of base pairing and stacking interactions are independent of base identity. As the free-energy cost of disrupting a duplex to form a kink is sequence-dependent, it is important to consider how sequence might perturb the general trends we have elaborated thus far. We, therefore, studied a variety of sequences used in the V&H<sup>18</sup> experiments including six at  $N_{bp} = 73$  using the parametrization of oxDNA that includes sequence-dependent thermodynamics (Figure 3).

As duplex kinking for oxDNA typically involves the breaking of base pairs, we observe that this kinking preferentially occurs at A–T base pairs. Base pairs that are weaker than average introduce preferred locations for kinking; thus, when duplex kinking is relevant,  $j_{eq}^{oxDNA}$  is expected to be larger for the sequence-dependent than for the average-base parametrization. This is indeed the case for  $N_{bp} = 73$ . The sequence-induced variation for the V&H  $N_{bp} = 73$  sequences is a factor of  $\sim 4$  (GC-content 11–52%). This compares to a factor of  $\sim 8$  between the extrema in GC-content, poly(AT) and poly(GC).

That we find sequence heterogeneity generally makes duplex kinking easier is consistent with our previous explicit investigation of the free energy of duplex kinking for oxDNA.<sup>26</sup> Our results imply that the crossover to cyclized configurations with duplex kinking occurs at slightly longer  $N_{bp}$  when sequence-dependence is included than for the average-base parametrization. Sequence also has the potential to impact cyclization in the short-length regime where fraying is important, but this will depend sensitively on the exact sequences at the ends of the duplex region with A–T base pairs generally making fraying and, hence, cyclization more likely.

As well as equilibrium constants, free-energy profiles as a function of the number of bases pairs were computed for each system we considered, as these can give more insight into the pathway for association. Example profiles for cyclization and dimerization at  $N_{bp} = 101$  are illustrated in Figure 5. One



**Figure 5.** Free energy profiles of cyclization (solid) and dimerization (dashed) for a  $N_{bp} = 101$  system ( $N_s = 10$ ,  $N_d = 91$ ). The activation free-energy barriers,  $\Delta G^\ddagger$  for the forward (cyclization, dimerization) and reverse (uncyclization, undimerization) reactions are labeled.  $\Delta G_{cyc}^\ddagger$  reflects the free-energy cost of bending to form the first base pair in a cyclization system, whereas  $\Delta G_{dim}^\ddagger$  reflects the entropic cost of bringing two monomers together within the simulation volume. The dimerization simulations are for a cubic simulation box of dimension 85.18 nm, corresponding to a duplex concentration of  $2.69 \mu\text{M}$ . Note that for clarity, we have depicted  $\Delta G_{uncyc}^\ddagger$  and  $\Delta G_{undim}^\ddagger$  as the free energy difference between the fully base-paired closed state  $Q_{bp} = 10$  and the transition state  $Q_{bp} = 1$ ; however, in practice, frayed states  $Q_{bp} = [2, 9]$  do contribute to the closed state. While the distinction does not significantly impact our results ( $<1/4k_B T$ ), we do include the contribution of frayed states in both  $\Delta G_{uncyc}^\ddagger$  and  $\Delta G_{undim}^\ddagger$ .

interesting feature of the profiles is that the free-energy gain from hybridizing the complementary sticky ends once an initial base pair has formed is less for cyclization than for dimerization:  $\Delta G_{uncyc}^\ddagger < \Delta G_{undim}^\ddagger$ . Physically, this indicates that substantial additional bending stress develops as subsequent base pairs form for the cyclization system, reducing the free-energy gain upon zipping of the sticky ends relative to dimerization. For example, the  $Q_{bp} = 1$  configuration in Figure 2c is clearly less bent than the  $Q_{bp} = 10$  configuration in Figure 2b (Figure S1). The activation energies derivable from these profiles will be particularly useful in the next section when we consider the dynamic  $j$ -factor.

**III.A. Comparison with experiment.** As noted in the Introduction, the results of ligase experiments performed in the low ligase concentration limit should, in principle, be comparable to equilibrium  $j$ -factors (although there are subtleties related to the ensemble of states actually detected by the ligation enzymes). Indeed, for long DNA molecules ( $N_{bp}$  much longer than the persistence length) and low ligase concentrations, there has been consistent agreement between experiment and the WLC model.<sup>6,8,18,20,34,35</sup> However, Cloutier and Widom (C&W)<sup>35</sup> reported results for  $N_{bp}$  shorter than the persistence length ( $N_{bp} = 93, 94, 95, 105, 116$ ), showing an apparent deviation from WLC behavior, with  $j_{dyn}^{ligase}$  (eq 2) enhanced over the SY WLC prediction  $j_{eq}^{WLC60}$  by a factor of  $10^2$ – $10^4$ . The differences between the maxima in  $j_{eq}^{oxDNA}$  and  $j_{eq}^{WLC}$  in this size range are much smaller (Figure 3) and so cannot account for this discrepancy.

In contrast, Du et al.<sup>8</sup> found no deviation from WLC behavior for  $N_{bp} = 105$ – $130$ . Furthermore, they presented evidence suggesting that the C&W experiments used too high a ligase concentration to enable  $j_{dyn}^{ligase}$  to be compared with  $j_{eq}^{WLC}$ . The results of Du et al. are in good agreement with the SY WLC expression, albeit with somewhat different materials parameters (torsional stiffness, persistence length and pitch length) than for oxDNA, owing in part to different buffer conditions (section S2.D).

FRET measurements on DNA cyclization, as pioneered by Vafabakhsh and Ha (V&H),<sup>18</sup> provide a more direct measure of cyclization because  $k_{cyc}$ ,  $k_{uncyc}$ , and  $K_{eq}^{cyc}$  are obtainable. V&H report  $K_{eq}^{cyc}$  for a small number of systems; oxDNA results are in good agreement for those lengths (section S2.C).

V&H's evidence of enhanced cyclization at  $N_{bp} \lesssim 100$ , however, is not based on  $K_{eq}^{cyc}$  but a comparison between their  $j_{dyn}^{FRET}$  (eq 3) and the SY WLC expression  $j_{eq}^{WLC}$ . However, as noted earlier, this is only a fair comparison if  $k_{uncyc} = k_{undim}$ , which, as  $k_{undim}$  is expected to be length independent, also implies that  $k_{uncyc}$  should be independent of  $N_{bp}$ . However, since both V&H<sup>18</sup> and more recent FRET measurements<sup>20</sup> suggest that  $k_{uncyc}$  increases with  $N_{bp}$ , this condition is not met.

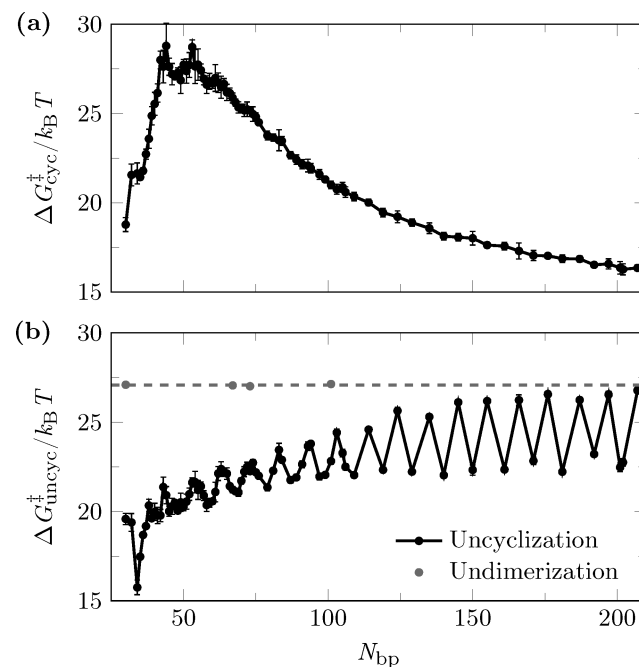
Although we do not directly simulate the dynamics of cyclization, we can estimate the relative rates of processes using free-energy profiles such as those in Figure 5 and unimolecular rate theory for activated processes.<sup>73</sup> In agreement with experimental investigations, previous work on oxDNA has shown that duplex formation has an effective "transition state" involving a very small number of base pairs.<sup>74</sup> We, therefore, make the assumption that uncyclization and undimerization rates are given by

$$k_{uncyc} = A \exp\left(\frac{-\Delta G_{uncyc}^\ddagger}{k_B T}\right) \quad (4)$$

$$k_{undim} = A \exp\left(\frac{-\Delta G_{undim}^\ddagger}{k_B T}\right) \quad (5)$$

where  $A$  is a constant for DNA melting and  $\Delta G_{uncyc}^\ddagger$  and  $\Delta G_{undim}^\ddagger$  are defined in Figure 5. The physical content of this assumption is that an increased favorability of base-pair formation is manifested in slower unbinding rates. This is important because bending stress in cyclized systems reduces  $\Delta G_{uncyc}^\ddagger$  (Figure 5). The rate constants will be equal ( $k_{uncyc} = k_{undim}$ ) in the very long length limit ( $N_{bp}$  much longer than the persistence length).

$\Delta G_{cyc}^\ddagger$  and  $\Delta G_{uncyc}^\ddagger$  are plotted as a function of  $N_{bp}$  in Figure 6. In particular, there is a general decrease in  $\Delta G_{uncyc}^\ddagger$  at shorter



**Figure 6.** Free-energy barriers for (a) cyclization  $\Delta G_{cyc}^\ddagger$ , (b) uncyclization  $\Delta G_{uncyc}^\ddagger$  (black), and undimerization (gray, average highlighted with dashed line), all computed using the oxDNA average-base parametrization.

lengths, suggesting  $k_{uncyc}$  increases with shorter  $N_{bp}$ , in agreement with experimental results.<sup>18,20</sup> Additionally, any torsional stress in the cyclized state is relieved as base pairs are disrupted, leading to the oscillations in  $\Delta G_{uncyc}^\ddagger$  at long lengths, with the minima occurring at the more torsionally stressed off-register lengths.

To compare with V&H's results for  $j_{dyn}^{FRET}$ , we note that

$$j_{dyn} = \frac{k_{cyc}}{k_{dim}} = \frac{K_{eq}^{cyc} k_{uncyc}}{K_{eq}^{dim} k_{undim}} \quad (6)$$

Therefore, using eq 5, our approximation for  $j_{dyn}^{oxDNA}$  is

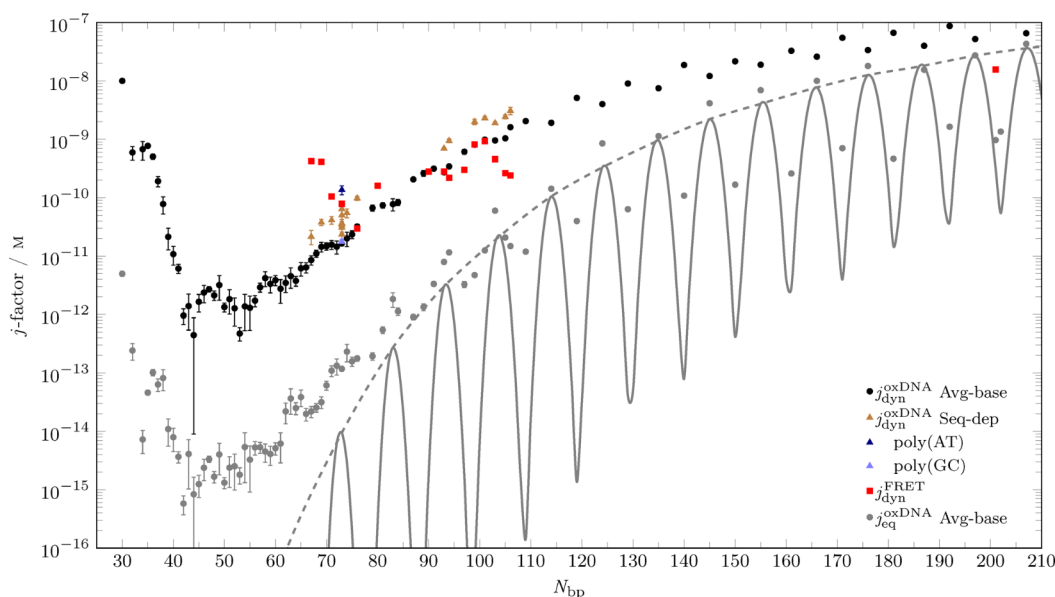
$$j_{dyn}^{oxDNA} = \frac{K_{eq}^{cyc} \exp\left(\frac{-\Delta G_{uncyc}^\ddagger}{k_B T}\right)}{K_{eq}^{dim} \exp\left(\frac{-\Delta G_{undim}^\ddagger}{k_B T}\right)} \quad (7)$$

at each  $N_{bp}$ .

We are now in a position to estimate  $j_{dyn}^{oxDNA}$  from oxDNA's equilibrium constants and activation free-energy barriers. We expect  $K_{eq}^{dim}$  and  $\Delta G_{undim}^\ddagger$  to be length-independent as the excluded volume of the duplex far from the complementary single-stranded sticky ends is likely to have little effect on the dimerization process. Indeed, this appears to be the case to within less than  $0.5k_B T$  in  $\Delta G$  (Figure S2 and Table S2). Our resulting  $j_{dyn}^{oxDNA}$  values are plotted in Figure 7, where they are compared to our  $j_{eq}^{oxDNA}$  values and V&H's  $j_{dyn}^{FRET}$ .

We observe that oxDNA's  $j_{dyn}^{oxDNA}$  values lie substantially above  $j_{eq}^{oxDNA}$ , except in the limit of long  $N_{bp}$  for on-register lengths. Only if all the stress-induced destabilization of the fully cyclized state is exhibited in the forward rate would  $j_{dyn} = j_{eq}$ .





**Figure 7.** OxDNA dynamic  $j$ -factor  $j_{\text{dyn}}^{\text{oxDNA}}$  (black circle) compared to the FRET experiments of V&H  $j_{\text{dyn}}^{\text{FRET}}$  (red squares).<sup>18</sup> Results for the oxDNA sequence-dependent parametrization using V&H 14 variable  $N_{\text{bp}}$  sequences, in addition to 6 sequences at  $N_{\text{bp}} = 73$ , highlight the role of sequence variation. For reference,  $j_{\text{eq}}^{\text{oxDNA}}$  and  $j_{\text{eq}}^{\text{WLC}}$  are also shown (gray).

Although the activation free-energy for cyclization increases as  $N_{\text{bp}}$  decreases (Figure 6a), the full bending and twisting stress present in the cyclized state is not yet present in the transition state where the two complementary sticky ends have formed their first base pair, for example, Figure 2c, and so  $j_{\text{dyn}}^{\text{oxDNA}} > j_{\text{eq}}^{\text{oxDNA}}$ . Instead, stress which is not present in the transition state leads to a decrease in the free-energy barrier for uncyclization (Figure 6b) and accelerates uncyclization relative to undimerization. Alternatively, one can consider that the stress in the cyclized state subjects the duplex formed between the sticky ends to a shear force,<sup>20</sup> which is well-known to lead to a more rapid rupture of a duplex.<sup>66,75</sup>

The observed behavior is similar to that seen by V&H; indeed,  $j_{\text{dyn}}^{\text{oxDNA}}$  and  $j_{\text{dyn}}^{\text{FRET}}$  agree remarkably well in the range  $N_{\text{bp}} = 90$ – $105$ , with reasonable agreement extending to  $N_{\text{bp}} = 70$ . Our results suggest that variation of uncyclization rates with  $N_{\text{bp}}$  may be an important contribution to apparent non-WLC behavior in  $j_{\text{dyn}}^{\text{FRET}}$ .

To begin to understand the difference between V&H's dynamic  $j$ -factor and  $j_{\text{eq}}^{\text{WLC}}$ , a number of authors have tried to account for the role of the single-stranded tails in cyclization by incorporating a “capture radius” into the WLC  $j$ -factor calculation, as an approximation for how close the duplex ends must be for the sticky ends to hybridize.<sup>10,18,20,57,76</sup> Indeed, this approach leads to a significantly enhanced  $j$ -factor when a value that is taken to be roughly appropriate for the 10-base sticky ends of the V&H experiment, namely, 5 nm, is used.<sup>18,57</sup>

Assuming this approach is an attempt to capture the weaker constraints at the transition state and, hence, to estimate the activation free-energy barrier relevant to  $j_{\text{dyn}}$  (note  $j_{\text{dyn}}$  and  $j_{\text{eq}}$  are often not clearly differentiated in discussions of V&H's results), we can use our oxDNA results to test the accuracy of this approximation by measuring the separation of the duplex ends in the transition state ensemble. At long lengths our measured capture radius is approximately constant with a value just under 4 nm but increases at shorter lengths because some of the stress at the transition state is partitioned into stretching

the single-stranded tails, reaching a maximum of about 7 nm at about  $N_{\text{bp}} \approx 40$  (Figure S5). Thus, although the 5 nm value used previously<sup>18,57</sup> is not unreasonable, unsurprisingly this approach does not capture the full complexity of the transition state to cyclization, and nor does it account for the strain present in the sticky ends for them to achieve contact.

Therefore, although explaining a dynamic  $j$ -factor in terms of a capture radius is physically well-motivated, it does not provide a full understanding. Moreover, as its value is not tightly constrained, and certain contributions to stability are neglected, it is difficult to judge whether WLC behavior is violated or not by comparing a calculated  $j$ -factor curve with a given capture radius to an experimental  $j_{\text{dyn}}$ .

We note that  $j_{\text{dyn}}^{\text{oxDNA}}$ , in contrast to  $j_{\text{eq}}^{\text{oxDNA}}$ , varies comparatively smoothly with  $N_{\text{bp}}$ , with only a weak periodicity on the length scale of the pitch length. The shallow maxima at larger lengths occur at  $N_{\text{d}} = (n + 1/2) \times \text{pitch length}$ , because the two sticky ends are then on the same side of a torsionally unstressed duplex. Consistent with this smooth variation,  $\Delta G_{\text{cyc}}^{\ddagger}$  also varies relatively smoothly with  $N_{\text{bp}}$ . The strong periodicity in  $j_{\text{eq}}^{\text{oxDNA}}$  comes from the free-energy gain when zipping up the complementary sticky ends (i.e.,  $\Delta G_{\text{uncyc}}^{\ddagger}$ ), which is greater when  $N_{\text{bp}}$  is an integer multiple of the pitch length, allowing the formation of a relatively relaxed coaxially stacked circle. V&H suggest that their  $j_{\text{dyn}}^{\text{FRET}}$  data at  $N_{\text{bp}} = 93$ – $106$  displays a strong oscillation with a period of about one pitch length (Figure 7). In agreement with Vologodskii et al.,<sup>10</sup> we find no physical mechanism for such a strong oscillation and would suggest that the experimental evidence for this oscillation is not compelling.

In the short-length regime that we earlier identified in our analysis of  $j_{\text{eq}}^{\text{oxDNA}}$ ,  $j_{\text{dyn}}^{\text{oxDNA}}$  also tends to vary more smoothly, probably because the geometric effects responsible for size-specific features in  $j_{\text{eq}}^{\text{oxDNA}}$  are most keenly felt when the sticky ends are fully zipped up. For example, at  $N_{\text{bp}} = 34$ , there is a strong minimum in  $\Delta G_{\text{uncyc}}^{\ddagger}$  and, hence,  $j_{\text{eq}}^{\text{oxDNA}}$ , which is not present in  $j_{\text{dyn}}^{\text{oxDNA}}$ .

At the shortest lengths investigated by V&H  $j_{\text{dyn}}^{\text{FRET}}$  still lies above  $j_{\text{dyn}}^{\text{oxDNA}}$ . Some of this difference might be due to our use of the oxDNA average-base parametrization. Sequence does play a role in DNA flexibility but the impact of sequence variation is nontrivial. In particular, because kinks tend to localize to AT base pairs, kinking in the duplex is easier for a sequence with 50% GC-content than reported for our average-base parametrization.

Using the oxDNA parametrization with sequence-dependent thermodynamics we found  $j_{\text{dyn}}$  to increase compared to the values for the average-base parametrization. Consistent with our results for  $j_{\text{eq}}^{\text{oxDNA}}$ , we find a sequence-induced variation in  $j_{\text{dyn}}^{\text{oxDNA}}$  of a factor of  $\sim 4$  for the six V&H  $N_{\text{bp}} = 73$  sequences (GC-content 11–52%)<sup>18</sup> compared to a factor of  $\sim 8$  between the extrema in GC-content, poly(AT) and poly(GC). This compares to a factor of  $\sim 60$  for experimental looping rates, a discrepancy, which may be explained by V&H's use of poly(A) tracts, a sequence-motif well-known to introduce intrinsic curvature in duplex DNA.<sup>77</sup> As oxDNA's sequence-dependent parametrization is based on the nearest-neighbor thermodynamics of SantaLucia et al.,<sup>78,79</sup> alternative structural motifs, such as poly(A) tracts, are outside the scope of the model. OxDNA also does not reproduce sequence-dependent structural (e.g., the difference in size between purine and pyrimidine) or mechanical (e.g., flexibility) properties.

The origin of the remaining discrepancy between  $j_{\text{dyn}}^{\text{FRET}}$  and  $j_{\text{dyn}}^{\text{oxDNA}}$  at the shortest  $N_{\text{bp}}$  is not yet clear but may indicate enhanced flexibility for V&H versus oxDNA because of WLC (i.e., lower persistence length) or non-WLC behavior (i.e., kinking within the duplex at slightly longer  $N_{\text{bp}}$ ). The oxDNA persistence length of 41.82 nm is within the range of experimental observations at  $[\text{Na}^+] = 500$  mM, but smaller values at  $[\text{Na}^+] = 750$  mM are not implausible.<sup>80,81</sup> It is also possible that oxDNA slightly underestimates the prevalence of kinking within duplex regions; an onset of kinking at slightly lower stress (longer  $N_{\text{bp}}$ ) would make cyclization at shorter lengths more favorable. It is also worth noting both that the presence of fluorophores may cause perturbations in the V&H experiments and that there are also unexplained differences between  $j_{\text{dyn}}^{\text{FRET}}$  and  $j_{\text{dyn}}^{\text{oxDNA}}$  at  $N_{\text{bp}} \approx 105$ .

Finally, we noted in the Introduction that, although V&H infer both  $k_{\text{cyc}}$  and  $k_{\text{uncyc}}$  from their experiments, it is unclear whether they plot  $j_{\text{dyn}}^{\text{FRET}}$  as defined in eq 3 or  $(k_{\text{cyc}} + k_{\text{uncyc}})/k_{\text{dim}} \geq j_{\text{dyn}}^{\text{FRET}}$ .<sup>18,53</sup> For the longer duplex lengths considered by V&H, for which  $k_{\text{cyc}} \gtrsim k_{\text{uncyc}}$ ,  $j_{\text{dyn}}^{\text{FRET}} \approx j_{\text{dyn}}^{\text{FRET}}$  this difference has limited consequences and our comparison to the experimental data is reasonable regardless of this uncertainty. For the shortest lengths, however,  $(k_{\text{cyc}} + k_{\text{uncyc}})/k_{\text{dim}}$  may be significantly larger than the true value of  $j_{\text{dyn}}^{\text{FRET}}$ , perhaps accounting for some of the deviation between  $j_{\text{dyn}}^{\text{oxDNA}}$  and experimental results at the shortest  $N_{\text{bp}}$ .

#### IV. DISCUSSION

Cyclization is a system-dependent manifestation of the general thermodynamics of strong DNA bending.<sup>26</sup> The remarkable range of behavior in cyclized systems is explicable by the interplay between three specific deformation modes of stressed duplexes: continuous bending, kinking and fraying.

OxDNA reveals that each of these modes is present at a characteristic length-scale with respect to cyclization: continuous bending at long lengths ( $N_{\text{bp}} \gtrsim 80$ ), duplex kinking at intermediate lengths ( $N_{\text{bp}} \approx 45\text{--}80$  bp) and fraying at short lengths ( $N_{\text{bp}} \lesssim 45$  bp). In addition, as  $N_{\text{bp}}$  is shortened, there is

an increase in kinking at the two nicks that remain after the hybridization of the sticky ends. At longer lengths, kinking at a nick is only observed for “off-register” molecules that cannot form torsionally relaxed coaxially stacked circles. The ability of said nicks to relax bending, as well as torsional stress, means that kinking at the nicks becomes increasingly prevalent for shorter  $N_{\text{bp}}$ . At the shortest lengths, kinking at both nicks is dominant.

We use oxDNA to probe the reported observation of non-WLC behavior in FRET-based cyclization experiments.<sup>18</sup> In agreement with experiment, we observe that for shorter values of  $N_{\text{bp}}$ , the apparent  $j$ -factor lies substantially above the predictions of the Shimada and Yamakawa (SY)<sup>60</sup> WLC model. We also observe that the periodic oscillations predicted by the SY model are suppressed. This behavior arises from two conceptually distinct phenomena.

First, highly stressed cyclized systems can adopt configurations that relax stress more effectively than through continuous bending, thereby reducing the overall free-energy cost of cyclization relative to a direct estimate based on a simple WLC-based model. At various values of  $N_{\text{bp}}$ , oxDNA identifies kinking at nicks, kinking within the duplex region, and fraying of base pairs as key relaxation modes.

Second, oxDNA suggests that not all of the reduction in  $K_{\text{eq}}^{\text{cyc}}$  relative to  $K_{\text{eq}}^{\text{dim}}$  is due to stress manifest in the cyclization rate; uncyclization rates are also substantially increased relative to undimerization rates. The result is that dynamic  $j$ -factors based on the ratio of cyclization and dimerization rates lie even further above the SY prediction than their equilibrium  $j$ -factor counterpart.

Of the above effects, only kinking within the duplex can reasonably be described as truly non-WLC behavior. WLC and related statistical models do not predict absolute cyclization rates directly. Kinking at nicks and fraying can only occur when the DNA backbone is discontinuous, and any resultant effects are unrelated to whether WLC models accurately describes the body of the DNA duplex. In oxDNA (with the average-base parametrization), kinking in the duplex only has a substantial effect for  $N_{\text{bp}} \lesssim 70$  bp. If sequence dependence is taken into account, kinking occurs for slightly longer lengths.

Various authors have incorporated a “capture radius” into the WLC  $j$ -factor calculation to capture phenomenologically some of the effects listed above.<sup>10,18,20,57</sup> Indeed, this approach leads to an enhanced  $j$ -factor, but the choice of capture radius is somewhat arbitrary and imprecise, making it difficult to assess whether non-WLC behavior is present. Our explicit simulation with a coarse-grained model has removed this major source of ambiguity in interpreting V&H's results, while suggesting that the underlying argument for a capture radius is not unreasonable.

OxDNA is only a model, and good correspondence with experimental results should not be overinterpreted. The model neglects effects, such as sequence-dependent intrinsic curvature,<sup>82,83</sup> which may play a complementary role to the effects identified here in facilitating cyclization. Nonetheless, the mechanisms identified by oxDNA are clearly physically plausible; indeed, enhanced uncyclization rates have previously been noted in the literature<sup>18,20</sup> (and also for transcription-factor mediated looping<sup>84</sup>). Most importantly, it is clear that much, if not all, of the apparent discrepancy between the data of V&H and the predictions of WLC-based models can be explained by effects that are not true violations of the WLC model of duplex DNA flexibility, although we cannot account

for deviations at their very shortest lengths ( $N_{bp} \lesssim 70$ ). In oxDNA, kinking within the duplex region is present, but not completely dominant, in these systems. It is possible that oxDNA slightly overestimates the difficulty of kinking within a duplex—if this is the case, the data for the very smallest values of  $N_{bp}$  studied by V&H may be indicative of duplex flexibility over and above that predicted by the WLC, although we note that to give a substantial effect on  $j_{eq}$ , kinking must not only be present, but must dominate the ensemble.

Our study also helps to reconcile the results of V&H with previous ligase-based assays which saw no evidence of enhanced flexibility at  $N_{bp} \approx 100$ ,<sup>8</sup> and previous studies of minicircles, which detected no evidence of duplex disruption at these length scales.<sup>85</sup> In addition, oxDNA has also been shown to quantitatively reproduce the signature of an apparent kinking transition in a “molecular vice”,<sup>26</sup> as well as supporting the hypothesized kinking explanation in that system. The molecular vice is significantly different from a cyclization experiment. Nonetheless, by comparing the oxDNA results, we can infer that the experimental observation of kinking in the molecular vice is not consistent with the hypothesized kinking in the longer duplexes studied by V&H. Without a model that can simulate both systems, this kind of comparison is extremely hard.<sup>26</sup>

To explore whether the shortest lengths studied by V&H do show evidence of kinking and enhanced flexibility, we propose experiments on shorter sequences and systematic collection of both dynamic and equilibrium data. The latter is extremely important; statistical WLC models make equilibrium predictions, and so the breakdown of a WLC description can only be confirmed with equilibrium data. Indeed, elucidating the subtleties of dynamic and quasi-dynamic (C&W)  $j$ -factors is one of the key issues addressed in this work.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00112.

Additional details about the simulations, equilibrium constant computation, kink detection, and sequences used; additional results on the dimerization and cyclization equilibria, including effects of the lengths of the sticky ends, mismatches and nicks, and estimates of the capture radius; and further discussion of comparisons to experiment (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: jonathan.doye@chem.ox.ac.uk

### ORCID

Jonathan P. K. Doye: 0000-0002-2226-9524

### Notes

The authors declare no competing financial interest. The oxDNA code and documentation is available from <https://dna.physics.ox.ac.uk/>. Initialization files necessary to run the particular simulations reported here, along with representative raw data, are available to download from DOI: 10.5281/zenodo.1753767.

## ■ ACKNOWLEDGMENTS

This work was supported by the Engineering and Physical Sciences Research Council [EP/I001352/1], the National Science Foundation Graduate Research Fellowship Program, the National Institutes of Health National Heart, Lung and Blood Institute, Wolfson College, Oxford, University College, Oxford, and the Royal Society. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work (<http://dx.doi.org/10.5281/zenodo.22558>). The authors also acknowledge the use of the e-Infrastructure South IRIDIS High Performance Computing Facility.

## ■ REFERENCES

- (1) Zhang, Y.; McEwen, A. E.; Crothers, M. D.; Levene, S. D. Analysis of *in-vivo* LacR-Mediated Gene Repression Based on the Mechanics of DNA Looping. *PLoS One* **2006**, *1*, No. e136.
- (2) Vijayan, V.; Zuzow, R.; O’Shea, E. K. Oscillations in Supercoiling Drive Circadian Gene expression in Cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 22564–22568.
- (3) Wolffe, A. P.; Hayes, J. J. Chromatin Disruption and Modification. *Nucleic Acids Res.* **1999**, *27*, 711–720.
- (4) Smith, S. B.; Cui, Y.; Bustamante, C. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science* **1996**, *271*, 795–799.
- (5) Strick, T. R.; Allemand, J.-F.; Bensimon, D.; Bensimon, A.; Croquette, V. The Elasticity of a Single Supercoiled DNA Molecule. *Science* **1996**, *271*, 1835–1837.
- (6) Shore, D.; Baldwin, R. L. Energetics of DNA Twisting: I. Relation Between Twist and Cyclization Probability. *J. Mol. Biol.* **1983**, *170*, 957–981.
- (7) Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S. B. Entropic Elasticity of  $\lambda$ -Phage DNA. *Science* **1994**, *265*, 1599–1600.
- (8) Du, Q.; Smith, C.; Shiffeldrim, N.; Vologodskaya, M.; Vologodskii, A. Cyclization of Short DNA Fragments and Bending Fluctuations of the Double Helix. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 5397–5402.
- (9) Mazur, A. K.; Maaloum, M. DNA Flexibility on Short Length Scales Probed by Atomic Force Microscopy. *Phys. Rev. Lett.* **2014**, *112*, 068104.
- (10) Vologodskii, A.; Frank-Kamenetskii, M. D. Strong Bending of the DNA Double Helix. *Nucleic Acids Res.* **2013**, *41*, 6785–6792.
- (11) Matek, C.; Ouldrige, T. E.; Doye, J. P. K.; Louis, A. A. Plectoneme Tip Bubbles: Coupled Denaturation and Writhing in Supercoiled DNA. *Sci. Rep.* **2015**, *5*, 7655.
- (12) Dittmore, A.; Brahmachari, S.; Takagi, Y.; Marko, J. F.; Neuman, K. C. Supercoiling DNA Locates Mismatches. *Phys. Rev. Lett.* **2017**, *119*, 147801.
- (13) Lankaš, F.; Lavery, R.; Maddocks, J. H. Kinking Occurs During Molecular Dynamics Simulations of Small DNA Minicircles. *Structure* **2006**, *14*, 1527–1534.
- (14) Qu, H.; Tseng, C.-Y.; Wang, Y.; Levine, A. J.; Zocchi, G. The Elastic Energy of Sharply Bent Nicked DNA. *EPL* **2010**, *90*, 18003.
- (15) Qu, H.; Zocchi, G. The Complete Bending Energy Function for Nicked DNA. *EPL* **2011**, *94*, 18003.
- (16) Mitchell, J. S.; Laughton, C. A.; Harris, S. A. Atomistic Simulations Reveal Bubbles, Kinks and Wrinkles in Supercoiled DNA. *Nucleic Acids Res.* **2011**, *39*, 3928–3938.
- (17) Spiriti, J.; Kamberaj, H.; de Graff, A. M. R.; Thorpe, M. F.; van der Vaart, A. DNA Bending Through Large Angles is Aided by Ionic Screening. *J. Chem. Theory Comput.* **2012**, *8*, 2145–2156.
- (18) Vafabakhsh, R.; Ha, T. Extreme Bendability of DNA Less Than 100 Base Pairs Long Revealed by Single-Molecule Cyclization. *Science* **2012**, *337*, 1097–1101.
- (19) Fields, A. P.; Meyer, E. A.; Cohen, A. E. Euler Buckling and Nonlinear Kinking of Double-stranded DNA. *Nucleic Acids Res.* **2013**, *41*, 9881–9890.



- (20) Le, T. T.; Kim, H. D. Probing the Elastic Limit of DNA Bending. *Nucleic Acids Res.* **2014**, *42*, 10786–10794.
- (21) Kim, C.; Lee, O.-C.; Kim, J.-Y.; Sung, W.; Lee, N. K. Dynamic Release of Bending Stress in Short dsDNA by Formation of a Kink and Forks. *Angew. Chem., Int. Ed.* **2015**, *54*, 8943.
- (22) Irobalieva, R. N.; Fogg, J. M.; Catanese, D., Jr; Sutthibutpong, T.; Chen, M.; Barker, A. K.; Ludtke, S. J.; Harris, S. A.; Schmid, M. F.; Chiu, W.; et al. Structural Diversity of Supercoiled DNA. *Nat. Commun.* **2015**, *6*, 8440.
- (23) Lee, O.-C.; Kim, C.; Kim, J.-Y.; Lee, N. K.; Sung, W. Two Conformational States in D-shaped DNA: Effects of Local Denaturation. *Sci. Rep.* **2016**, *6*, 28239.
- (24) Wang, Q.; Pettitt, B. M. Sequence Affects the Cyclization of DNA Minicircles. *J. Phys. Chem. Lett.* **2016**, *7*, 1042–1046.
- (25) Sutthibutpong, T.; Matek, C.; Benham, C.; Slade, G. G.; Noy, A.; Laughton, C.; Doye, J. P. K.; Louis, A. A.; Harris, S. A. Long-range Correlations in the Mechanics of Small DNA Circles under Topological Stress Revealed by Multi-Scale Simulation. *Nucleic Acids Res.* **2016**, *44*, 9121–9130.
- (26) Harrison, R. M.; Romano, F.; Ouldrige, T. E.; Louis, A. A.; Doye, J. P. K. Coarse-Grained Modelling of Strong DNA Bending I: Thermodynamics and Comparison to an Experimental “Molecular Vice”. 2015, arXiv:1506.09005. arXiv.org e-Print archive. <https://arxiv.org/abs/1506.09005>.
- (27) Wang, R. N. Q.; Irobalieva; Chiu, W.; Schmid, M. F.; Fogg, J. M.; Zechiedrich, L.; Pettitt, B. M. Influence of DNA sequence on the structure of minicircles under torsional stress. *Nucleic Acids Res.* **2017**, *45*, 7633–7642.
- (28) Schindler, T.; González, A.; Boopathi, R.; Roda, M. M.; Romero-Santacreu, L.; Wildes, A.; Porcar, L.; Martel, A.; Theodorakopoulos, N.; Cuesta-López, S.; et al. Kinky DNA in Solution: Small-Angle-Scattering Study of a Nucleosome Positioning Sequence. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2018**, *98*, 042417.
- (29) Ouldrige, 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) Ouldrige, T. E. *Coarse-Grained Modelling of DNA and DNA Self-Assembly*; Springer: Berlin, 2012.
- (31) Šulc, P.; Romano, F.; Ouldrige, 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.
- (32) Wang, J. C.; Davidson, N. Thermodynamic and Kinetic Studies on the Interconversion between the Linear and Circular Forms of Phage Lambda DNA. *J. Mol. Biol.* **1966**, *15*, 111–123.
- (33) Wang, J. C.; Davidson, N. On the Probability of Ring Closure of Lambda DNA. *J. Mol. Biol.* **1966**, *19*, 469–482.
- (34) Shore, D.; Langowski, J.; Baldwin, R. L. DNA Flexibility Studied by Covalent Closure of Short Fragments into Circles. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 4833–4837.
- (35) Cloutier, T. E.; Widom, J. Spontaneous Sharp Bending of Double-Stranded DNA. *Mol. Cell* **2004**, *14*, 355–362.
- (36) Zhang, Y.; Crothers, D. M. High-Throughput Approach for Detection of DNA Bending and Flexibility Based on Cyclization. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3161–3166.
- (37) Le, T. T.; Kim, H. D. Measuring Shape-Dependent Looping Probability. *Biophys. J.* **2013**, *104*, 2068–2076.
- (38) Zhang, Y.; Crothers, D. M. Statistical Mechanics of Sequence-Dependent Circular DNA and its Application for DNA Cyclization. *Biophys. J.* **2003**, *84*, 136–153.
- (39) Yan, J.; Marko, J. F. Localized Single-Stranded Bubble Mechanism for Cyclization of Short Double Helix DNA. *Phys. Rev. Lett.* **2004**, *93*, 108108.
- (40) Travers, A. DNA dynamics: Bubble ‘n’ Flip for DNA Cyclisation? *Curr. Biol.* **2005**, *15*, R377–R379.
- (41) Wiggins, P. A.; Phillips, R.; Nelson, P. C. Exact Theory of Kinkable Elastic Polymers. *Phys. Rev. E* **2005**, *71*, 021909.
- (42) Forties, R. A.; Bundschuh, R.; Poirier, M. G. The Flexibility of Locally Melted DNA. *Nucleic Acids Res.* **2009**, *37*, 4580–4586.
- (43) Wilson, D. P.; Tkachenko, A. V.; Meiners, J.-C. A Generalized Theory of DNA Looping and Cyclization. *EPL* **2010**, *89*, 58005.
- (44) Sivak, D. A.; Geissler, P. L. Consequences of Local Inter-Strand Dehybridization for Large-Amplitude Bending Fluctuations of Double-Stranded DNA. *J. Chem. Phys.* **2012**, *136*, 045102.
- (45) Xu, X. L.; Thio, B. J. R.; Cao, J. S. Correlated Local Bending of a DNA Double Helix and its Effect on DNA Flexibility in the Sub-Persistence-Length Regime. *J. Phys. Chem. Lett.* **2014**, *5*, 2868–2873.
- (46) Taranova, M.; Hirsh, A. D.; Perkins, N. C.; Andricioaei, I. Role of Microscopic Flexibility in Tightly Curved DNA. *J. Phys. Chem. B* **2014**, *118*, 11028–11036.
- (47) Chen, B.; Dong, C. Modeling Deoxyribose Nucleic Acid as an Elastic Rod Inlaid with Fibrils. *J. Appl. Mech.* **2014**, *81*, 071005.
- (48) Pollak, Y.; Goldberg, S.; Amit, R. Self-avoiding Wormlike Chain Model for Double-Stranded-DNA Loop Formation. *Phys. Rev. E* **2014**, *90*, 052602.
- (49) Shin, J.; Lee, O.-C.; Sung, W. How a Short Double-Stranded DNA Bends. *J. Chem. Phys.* **2015**, *142*, 155101.
- (50) Naômé, A.; Laaksonen, A.; Vercauteren, D. P. A Coarse-Grained Simulation Study of the Structures, Energetics and Dynamics of Linear and Circular DNA with its Ions. *J. Chem. Theory Comput.* **2015**, *11*, 2813–2826.
- (51) Salari, H.; Eslami-Mossallam, B.; Naderi, M. S.; Ejtehad, M. R. Extreme Bendability of DNA Double Helix due to Bending Asymmetry. *J. Chem. Phys.* **2015**, *143*, 104904.
- (52) Cong, P.; Dai, L.; Chen, H.; van der Maarel, J. R. C.; Doyle, P. S.; Yan, J. Revisiting the Anomalous Bending Elasticity of Sharply Bent DNA. *Biophys. J.* **2015**, *109*, 2338–2351.
- (53) Jeong, J.; Le, T. T.; Kim, H. D. Single-Molecule Fluorescence Studies on DNA Looping. *Methods* **2016**, *105*, 34–43.
- (54) Zoli, M. J-Factors of Short DNA Molecules. *J. Chem. Phys.* **2016**, *144*, 214104.
- (55) Alexandrov, L. B.; Rasmussen, K. O.; Bishop, A. R.; Alexandrov, B. S. Evaluating the Role of Coherent Delocalized Phonon-like Modes in DNA Cyclization. *Sci. Rep.* **2017**, *7*, 9731.
- (56) Podtelezhnikov, A. A.; Mao, C.; Seeman, N. C.; Vologodskii, A. Multimerization-Cyclization of DNA Fragments as a Method of Conformational Analysis. *Biophys. J.* **2000**, *79*, 2692–2704.
- (57) Vologodskii, A.; Du, Q.; Frank-Kamenetskii, M. D. Bending of Short DNA Helices. *Artificial DNA: PNA & XNA* **2013**, *4*, 1–3.
- (58) Jacobson, H.; Stockmayer, W. H. Intramolecular Reaction in Polycondensations. I. The Theory of Linear Systems. *J. Chem. Phys.* **1950**, *18*, 1600–1606.
- (59) Peters, J. P.; Maher, L. J. DNA curvature and flexibility *in vitro* and *in vivo*. *Q. Rev. Biophys.* **2010**, *43*, 23–63.
- (60) Shimada, J.; Yamakawa, H. Ring-Closure Probabilities for Twisted Wormlike Chains. Application to DNA. *Macromolecules* **1984**, *17*, 689–698.
- (61) Doye, J. P. K.; Ouldrige, 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.
- (62) Ouldrige, 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.
- (63) Snodin, B. E. K.; Randisi, F.; Mosayebi, M.; Šulc, P.; Schreck, J. S.; Romano, F.; Ouldrige, T. E.; Tsukanov, R.; Nir, E.; Louis, A. A.; et al. Introducing Improved Structural Properties and Salt Dependence into a Coarse-Grained Model of DNA. *J. Chem. Phys.* **2015**, *142*, 234901.
- (64) Matek, C.; Ouldrige, T. E.; Levy, A.; Doye, J. P. K.; Louis, A. A. DNA Cruciform Arms Nucleate Through a Correlated but Asynchronous Cooperative Mechanism. *J. Phys. Chem. B* **2012**, *116*, 11616–11625.
- (65) Romano, F.; Chakraborty, D.; Doye, J. P. K.; Ouldrige, T. E.; Louis, A. A. Coarse-Grained Simulations of DNA Overstretching. *J. Chem. Phys.* **2013**, *138*, 085101.

(66) Mosayebi, M.; Louis, A. A.; Doye, J. P. K.; Ouldridge, T. E. Force-Induced Rupture of a DNA Duplex: From Fundamentals to Force Sensors. *ACS Nano* **2015**, *9*, 11993–12003.

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

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

(69) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(70) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* **2007**, *3*, 26–41.

(71) Zeida, A.; Machado, M. R.; Dans, P. D.; Pantano, S. Breathing, Bubbling, and Bending: DNA Flexibility from Multimicrosecond Simulations. *Phys. Rev. E* **2012**, *86*, 021903.

(72) Becker, N. B.; Rosa, A.; Everaers, R. The Radial Distribution Function of Worm-Like Chains. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2010**, *32*, 53–69.

(73) Hänggi, P.; Talkner, P.; Borkovec, M. Reaction-Rate Theory: Fifty Years after Kramers. *Rev. Mod. Phys.* **1990**, *62*, 251–341.

(74) Ouldridge, T. E.; Sulc, 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.

(75) Hatch, K.; Danilowicz, C.; Coljee, V. W.; Prentiss, M. Demonstration that the Shear Force Required to Separate Short Double-Stranded DNA does not Increase Significantly with Sequence Length for Sequences Longer than 25 Base Pairs. *Phys. Rev. E* **2008**, *78*, 011920.

(76) Guérin, T. Analytical Expressions for the Closure Probability of a Stiff Wormlike Chain for Finite Capture Radius. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2017**, *96*, 022501.

(77) Rivetti, C.; Walker, C.; Bustamante, C. Polymer Chain Statistics and Conformational Analysis of DNA Molecules with Bends or Sections of Different Flexibility. *J. Mol. Biol.* **1998**, *280*, 41–59.

(78) SantaLucia, J. A Unified View of Polymer, Dumbbell, and Oligonucleotide DNA Nearest-Neighbor Thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 1460–1465.

(79) SantaLucia, J.; Hicks, D. The Thermodynamics of DNA Structural Motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440.

(80) Savelyev, A. Do Monovalent Mobile Ions Affect DNA's Flexibility at High Salt Content? *Phys. Chem. Chem. Phys.* **2012**, *14*, 2250.

(81) Herrero-Galán, E.; Fuentes-Perez, M. E.; Carrasco, C.; Valpuesta, J. M.; Carrascosa, J. L.; Moreno-Herrero, F.; Arias-Gonzalez, J. R. Mechanical Identities of RNA and DNA Double Helices Unveiled at the Single-Molecule Level. *J. Am. Chem. Soc.* **2013**, *135*, 122–131.

(82) Mitchell, J. S.; Glowacki, J.; Grandchamp, A. E.; Manning, R. S.; Maddocks, J. H. Sequence-Dependent Persistence Lengths of DNA. *J. Chem. Theory Comput.* **2017**, *13*, 1539–1555.

(83) Rosanio, G.; Widom, J.; Uhlenbeck, O. C. *In vitro* Selection of DNAs with Increased Propensity to Form Small Circles. *Biopolymers* **2015**, *103*, 303–320.

(84) Chen, Y.-J.; Johnson, S.; Mulligan, P.; Spakowitz, A. J.; Phillips, R. Modulation of DNA Loop Lifetimes by the Free Energy of Loop Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 17396–17401.

(85) Du, Q.; Kotlyar, A.; Vologodskii, A. Kinking the Double Helix by Bending Deformation. *Nucleic Acids Res.* **2008**, *36*, 1120–1128.