Intrinsic Curvature of DNA Influences LacR-Mediated Looping

Sachin Goyal,* Todd Lillian,* Seth Blumberg,[†‡] Jens-Christian Meiners,[†‡] Edgar Meyhöfer,* and N. C. Perkins*
*Department of Mechanical Engineering, [†]Department of Physics, and [‡]Biophysics Research Division, University of Michigan, Ann Arbor, Michigan

ABSTRACT Protein-mediated DNA looping is a common mechanism for regulating gene expression. Loops occur when a protein binds to two operators on the same DNA molecule. The probability of looping is controlled, in part, by the basepair sequence of inter-operator DNA, which influences its structural properties. One structural property is the intrinsic or stress-free curvature. In this article, we explore the influence of sequence-dependent intrinsic curvature by exercising a computational rod model for the inter-operator DNA as applied to looping of the LacR-DNA complex. Starting with known sequences for the inter-operator DNA, we first compute the intrinsic curvature of the helical axis as input to the rod model. The crystal structure of the LacR (with bound operators) then defines the requisite boundary conditions needed for the dynamic rod model that predicts the energetics and topology of the intervening DNA loop. A major contribution of this model is its ability to predict a broad range of published experimental data for highly bent (designed) sequences. The model successfully predicts the loop topologies known from fluorescence resonance energy transfer measurements, the linking number distribution known from cyclization assays with the LacR-DNA complex, the relative loop stability known from competition assays, and the relative loop size known from gel mobility assays. In addition, the computations reveal that highly curved sequences tend to lower the energetic cost of loop formation, widen the energy distribution among stable and meta-stable looped states, and substantially alter loop topology. The inclusion of sequence-dependent intrinsic curvature also leads to nonuniform twist and necessitates consideration of eight distinct binding topologies from the known crystal structure of the LacR-DNA complex.

INTRODUCTION

DNA is often viewed as a static structure, whose primary role is to store the genetic code of the cell. In addition to this static picture, the structural flexibility and sequence-dependent mechanical properties of DNA enable the dynamic formation of complex protein-DNA assemblies responsible for gene regulation, DNA replication, and DNA repair. It is therefore important to consider the interplay between sequence, mechanical properties, and dynamics of DNA to fully understand its biological functions.

One way in which the structure and mechanical properties of DNA can influence biomolecular activity is by forming protein-mediated DNA loops; see, for example, Schleif et al. (1). In such instances, a protein or protein complex binds simultaneously to (at least) two noncontiguous operator sites on a DNA molecule, thereby forcing the intervening DNA into a loop. Depending on the specific proteins and sequences involved, a DNA loop can affect transcription by either repressing or promoting the binding and activity of RNA polymerase (1,2).

In this article, we employ a computational rod model of the inter-operator DNA as a means to explore sequence-dependent effects on looping. In particular, our objective is to understand how the looping energy and topology are influenced by the sequence-dependent intrinsic curvature (or stress-free curvature) of the substrate DNA. We also recognize the importance of sequence-dependent stiffness in this context as discussed in the literature (3–11). However, our objective is to explore the role of sequence-dependent intrinsic curvature which, while frequently addressed in experimental studies (12–14), has received relatively little attention from the modeling community. For instance, in the framework of rod models for DNA, we are only aware of the work of Manning et al. (6) that addresses sequence-dependent intrinsic curvature in the related problem of cyclization. However, looping mechanics is very different from cyclization mechanics as recently emphasized by Zhang et al. (15).

An overview of our goal, as well as our computational method, is illustrated in Fig. 1. We adopt the lactose repressor protein DNA complex (LacR) found in the bacterium *Escherichia coli* as our example. As illustrated in Fig. 1, we begin by specifying the sequence of the substrate DNA from which we compute its zero-temperature, stress-free conformation (via consensus tri-nucleotide model (16,17)) and, subsequently, the intrinsic curvature of the helical axis as input to the rod model. We then employ the known crystal structure of the LacR protein bound to the operators (4.80 Å resolution as reported by Lewis et al. (18)) to compute the position and orientation of the rod (boundary conditions) at the operator sites. The dynamic computational rod model (19) is then used to predict the topology and energetics of the resulting inter-operator loop.

To explore how the energy and topology of DNA loops are sensitive to the sequence-dependent intrinsic curvature, we consider both wild-type and curved variants of the inter-operator DNA for the LacR-DNA complex. The convenience of this example is that Kahn and co-workers have already studied LacR looping with a set of designed constructs whose

Submitted May 7, 2007, and accepted for publication July 12, 2007.
Address reprint requests to N. C. Perkins, E-mail: ncp@umich.edu.
Editor: Tamar Schlick.

0006-3495/07/12/4342/18 $2.00
doi: 10.1529/biophysj.107.112268

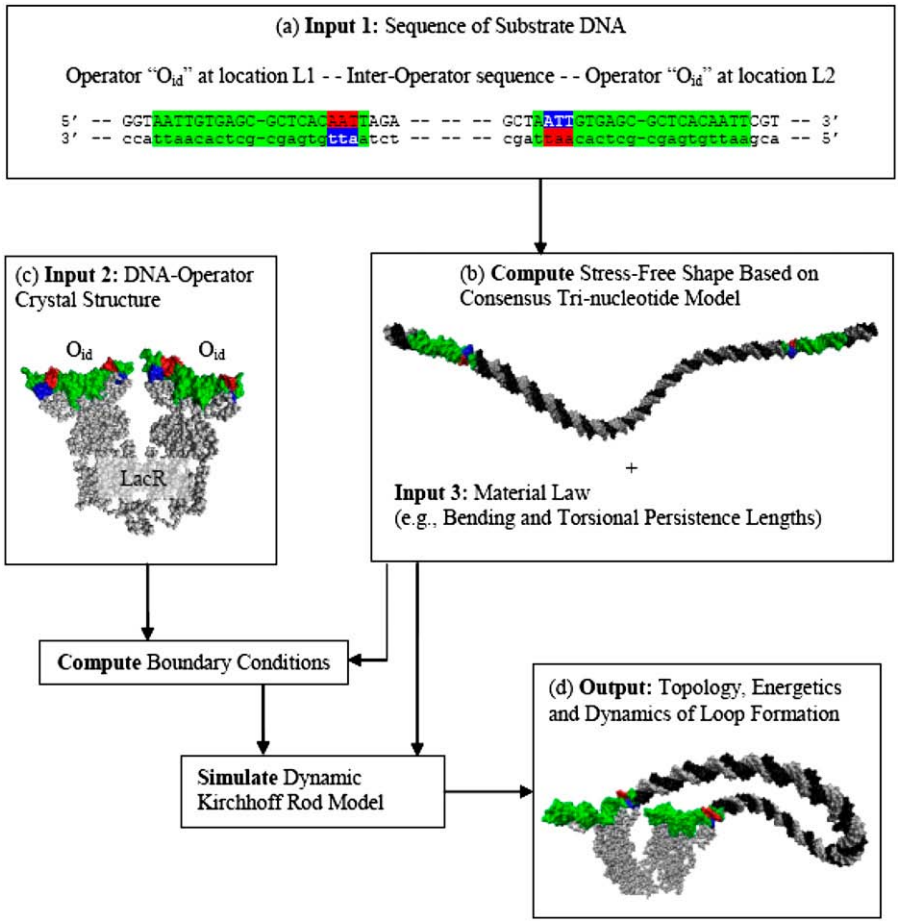


FIGURE 1 Modeling the effects of sequence-dependent, intrinsic curvature in looping of LacR-DNA. (*a*) Begin with specifying operator and inter-operator sequences (*green* denotes operators, capital case denotes the primary coding strand). (*b*) Construct zero-temperature, stress-free conformation using the consensus tri-nucleotide model (16,17) and compute intrinsic shape for rod model (twist and curvature of helical axis and inclination of the basepair planes with respect to the helical axis). (*c*) Employ known crystal structure of the LacR protein bound to the operators (18) and intrinsic shape to compute boundary conditions for rod model of looped DNA. (*d*) Input boundary conditions, intrinsic shape, and DNA material law to rod model (19) to compute inter-operator loop.

highly curved inter-operator sequences contain A-tracts with known and distinct helical phases with respect to the operators (12–14). Their studies, using gel-electrophoresis and fluorescence resonance energy transfer (FRET) experiments, provide experimental evidence that A-tract bends increase LacR loop stability and alter loop topology. By employing their inter-operator sequences as inputs to our computational model, we can probe these and other findings. Although our present focus is on LacR-mediated looping, the methods described herein can be generally applied to other examples (1,20) of looping behavior such as arising in GalR (21,22), Ara (23), SfI (24,25), and ntrC (26) in addition to other (nonlooping) behavior such as plectoneme formation in supercoiled DNA.

Our computational approach builds upon a long history of coarse-grain models for DNA dynamics that include Brownian dynamics simulations, Monte Carlo methods, and other statistical models (4,15,27–29) that are also reviewed in Olson (30) and Schick (31). A recent approach called the ''naturally discrete model'' has also been applied to LacR looping (15,32). The Kirchhoff rod approach leads to detailed descriptions of loop topology and internal (elastic) energy with modest computational effort. Our inclusion of sequence-dependent intrinsic curvature also builds upon the prior work of Schulten and co-workers (33–38) who employ a homogeneous elastic rod model to analyze the mechanics of LacR looping. The sequence-dependent intrinsic curvature included herein, renders the rod model nonhomogenous, and leads to substantial differences (both qualitative and quantitative) in the predictions of loop topology and internal energy.

The computational model used in this study requires three major inputs; namely, 1), the sequence of substrate DNA; 2), the crystal structure of LacR-operator complex; and 3), the material law for DNA; refer to Fig. 1. By material law, we refer to the elastic properties (that includes stiffness and intrinsic curvature) of DNA, which themselves can be sequence-dependent (3–11). We presently ignore sequence-dependent stiffness and focus instead on the effects of sequence-dependent intrinsic curvature. To this end, we employ averaged stiffness constants using published values of bending and torsional persistence lengths (39–41). The computational model, however, provides the framework for incorporating both sequence-dependent linear elastic material laws as well as nonlinear (and inelastic) laws (9,10,42,43), should they someday become well-characterized. We also treat the LacR as rigid and thereby ignore effects of protein flexibility. Molecular dynamics (MD) simulations have suggested that flexibility of the LacR derives primarily from the head regions (38) while

flexibility in the V-region has also been suggested in Friedman et al. (44) and Ruben and Roos (45).

We emphasize again that intrinsic curvature is only one manifestation of sequence-dependent behavior and properly accounting for other physical behaviors in a model of the LacR-DNA complex will also influence the computed loop topology and energy. For example, including protein flexibility and twist-extension coupling in DNA would lower the loop strain energy relative to that computed herein. We discuss the current limitations of our model and several extensions in detail after presenting our results.

With the above assumptions duly noted however, the computational model successfully predicts major experimental findings for the LacR-DNA complexes formed with the highly curved sequences studied by the Kahn lab (12–14) as detailed herein. First, the model predicts the binding topologies of the energetically favorable loops consistent with FRET measurements (13,14). Second, the model predicts the relative linking number distributions observed in cyclization assays formed with the LacR-DNA complex (12). In particular, the model predicts that mini-circles with relative linking number +1 are energetically favorable only in parallel binding topologies and that the associated closure energy reduces the cyclization rate. Third, the model largely predicts the relative loop stabilities as observed in competition assays (12). Fourth, the model also correctly predicts the relative speeds of looped sequences in gel mobility assays (12) upon computing and comparing the radius of gyration of the looped LacR-DNA complex. The model confirms that complexes formed with anti-parallel binding topologies move slower through the gel than those of comparable length with parallel binding topologies due to the greater compaction of the latter.

In addition, the computational model reveals the following major influences of sequence-dependent intrinsic curvature on looping in the LacR-DNA complex. First, the highly curved sequences of Mehta and Kahn (12) tend to lower the energetic cost of the (lowest energy) stable loops, widen the energy distribution among stable and meta-stable loops, and substantially alter loop topology. Second, the inclusion of sequence-dependent intrinsic curvature leads to nonuniform twist (or twist deficit) as recognized in Tobias and Olson (46) and also necessitates consideration of eight distinct binding topologies consistent with the known crystal structure of the LacR complex.

METHODS

In the Kirchhoff rod model, ds-DNA is approximated as a flexible rod having elastic properties as determined from single molecule experiments (47–50), MD simulations (5), and other biophysical techniques. We begin by reviewing the salient features of the computational rod model (19) for use in this study. The interested reader is referred to Goyal et al. (19) for a comprehensive development of this model, its relationship to other rod models, and benchmark results that confirm its accuracy. We then detail how we incorporate sequence-dependent intrinsic curvature in our formulation starting from knowledge of the inter-operator sequence.

Nonhomogeneous rod model for DNA

Fig. 2 illustrates a segment of ds-DNA with its helical axis defining the centerline of an equivalent rod. The shape of ds-DNA is parameterized by the three-dimensional centerline curve $R(s,t)$ and the cross-section fixed frame $\{a_i,(s,t)\}$, where s denotes the contour-length coordinate measured from one operator site and t denotes time. This equivalent rod model can be used to study the energetics and topology of DNA looping by formulating its mechanical properties (described below) based on experimental data and/or MD simulations.

The shape of the rod is also determined by the curvature and twist vector $\kappa(s,t)$ (defined as the spatial rate of rotation of $\{a_i,(s,t)\}$ (19)). Under stress-free conditions, the helical axis is not straight but conforms to a curved/twisted space curve. This intrinsic curvature of ds-DNA is captured by $\kappa_0(s)$ and it depends on the basepair sequence. The change in curvature/twist, $\kappa(s,t) - \kappa_0(s)$, produced by any subsequent deformation of the helical axis (e.g., by protein binding), generates an internal moment $q(s,t)$ and internal force $f(s,t)$. This response is governed by the long-length scale material law, which can be estimated from experiments or MD simulations. The interatomic interactions conspire to yield the long-length scale material law which is often assumed to be linearly elastic (see, for example, (6,7,17,30,31,33–38,51–53)). An exception is the nonlinear law proposed in the literature (9,10,43) for highly kinked strands, which has also been questioned in subsequent studies (42). Here, we shall adopt a linear elastic law

$$q(s,t) = B(\kappa(s,t) - \kappa_0(s)), \tag{1}$$

where the stiffness tensor B includes both bending and torsion stiffness. Commonly used values of the bending and torsional stiffness can be found from experimental measurements of the persistence lengths for bending/torsion (39–41). The above law renders the rod model nonhomogenous, that is, sequence-dependent by capturing the effects of intrinsic curvature/twist $\kappa_0(s)$. The associated elastic strain energy density follows from

$$S_e(s,t) = \frac{1}{2}(\kappa(s,t) - \kappa_0(s))^T B(\kappa(s,t) - \kappa_0(s)), \tag{2}$$

where the superscript T denotes transpose. This result can be readily used to understand how the elastic energy is distributed along the looped inter-operator DNA and its decomposition into components due to bending and twisting.

The deformation of the rod is governed by a set of differential equations (below) that are integrated using specified boundary conditions (19). For example, the boundary conditions for the inter-operator DNA loop define the relative position and orientation of the LacR operators known from the crystal structure (18) as detailed later. We describe the kinematics of this deformation by the linear velocity $v(s,t)$ and the angular velocity $\omega(s,t)$ of the rod cross section. The following four vector equations of rod theory (19) can be numerically integrated to solve for the four vector unknowns (f,q,v,ω) when combined with Eq. 1:

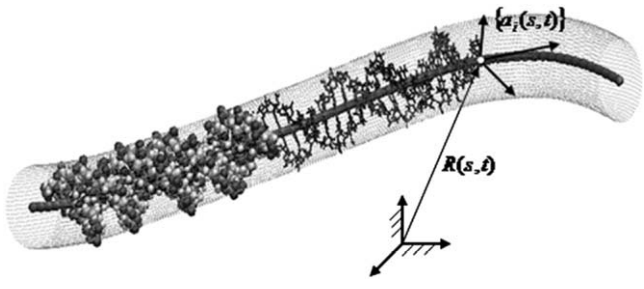


FIGURE 2 Rod model of (ds) DNA on long-length scales. Helical axis of duplex defines the rod centerline, which forms a three-dimensional space curve located by $R(s,t)$.

$$\frac{\partial f}{\partial s} + \kappa \times f = m\left(\frac{\partial v}{\partial t} + \omega \times v\right) - F, \tag{3}$$

$$\frac{\partial q}{\partial s} + \kappa \times q = I\frac{\partial \omega}{\partial t} + \omega \times I\omega + f \times \hat{t} - Q, \tag{4}$$

$$\frac{\partial v}{\partial s} + \kappa \times v = \omega \times \hat{t}, \tag{5}$$

$$\frac{\partial \omega}{\partial s} + \kappa \times \omega = \frac{\partial \kappa}{\partial t}. \tag{6}$$

Equations 3 and 4 represent the balance laws for linear and angular momentum of an element of ds-DNA, respectively. Equations 5 and 6 are kinematical constraints that describe the (assumed) inextensible helical axis and the required compatibility between curvature and angular velocity, respectively. The compatibility between the curvature and angular velocity arises from the fact that the orientation of the rod cross section is smooth in both space s and time t. A detailed description and derivation of these balance laws and the associated constraint equations is provided in Goyal et al. (19,54). In this dynamic formulation, $m(s)$ denotes the ds-DNA mass per unit contour length, $I(s)$ denotes the tensor of principal mass moments of inertia per unit contour length, $F(s,t)$ denotes any external forces (body force per unit contour length), $Q(s,t)$ denotes any distributed external moments (body moment per unit contour length), and $\hat{t}(s,t)$ denotes the helical axis unit tangent vector.

Note that the above formulation is dynamical in that we track the rod deformation in time from an assumed initial state (initial condition). Doing so allows the solution to relax to equilibrium ($v = \omega = 0$) under the influence of hydrodynamic dissipation and, in the process, confirms the mechanical stability of the computed equilibrium (19). Employing a dynamic formulation is advantageous because the solution dynamically relaxes to a (looped) equilibrium that is mechanically stable against small perturbations. The nonlinear theory also admits multiple (mechanically stable and looped) solutions which arise through sampling the multiple boundary conditions associated with all possible binding topologies of the operators with the LacR as detailed below. The resulting multiple equilibria are distinguished by differing elastic energies. By contrast, note that predictions of looped or supercoiled states directly from equilibrium rod theory (e.g., (6,33,34,37, 52,53) and citations in (55)) require a subsequent analysis of loop mechanical stability. Furthermore, equilibrium theories cannot capture possible dynamic transitions between equilibrium states as highlighted in Goyal et al. (19). Finally, note that while we employ a specific (linearly elastic) material law for ds-DNA, the formulation above is general in that Eq. 1 may be replaced with any other proposed material laws including those that capture sequence-dependent stiffness (4,5,8,11) and nonlinear material behavior (43).

In this study, the dynamical formulation is used as a numerical means to converge to the final equilibrium (looped) states, and it is not used to study or represent the dynamic pathway for looping in the presence of thermal kinetics. We accomplish this by integrating from the initial stress-free shape of the inter-operator DNA and then slowly transforming the operators from their stress-free conformation to their (final) position and orientation when bound to the LacR. The final loop topology and elastic energy are then computed. In other words, the boundary conditions for the rod are slowly varying and prescribed functions of time that begin with those of the stress-free state and end with those of the (final) looped state. We detail in Appendix A how we define the boundary conditions for the (final) looped state for the LacR-DNA complex. The inter-operator DNA modeled here as a rod includes three basepairs from each operator site (see Appendix A), as also assumed in the literature (33–38).

In aligning the boundary basepairs with the known crystal structure (18), one can consider eight possible binding topologies that distinguish how the operators bind to the binding domains. Note from Fig. 1 that the operators are identical and palindromic. Because the operators at locations L1 and L2 are palindromes, we first consider four distinct ways to attach them to the two binding domains BD1 and BD2 as illustrated in Fig. 3, where we also arbitrarily assumed that L1 always binds to BD1 (and L2 with BD2). (These four binding topologies were suggested to us by Prof. W. K. Olson, Department of Chemistry and Chemical Biology, Rutgers University.) By then allowing L1 to bind to BD2 and L2 to BD1, we arrive at a total of eight possible binding topologies. Since the crystal structure of the LacR protein given by PDB ID: 1LBG (18,56) appears to be asymmetric (by our calculations of data in the literature (18,56)) and so is the inter-operator DNA, the eight topologies are unique. For special cases (including palindromic inter-operator sequences and/or symmetry in the orientation of the boundary domains), one may arrive at fewer than eight (unique) binding topologies.

To distinguish the binding topologies, two conventions have been proposed in the literature (21,36). Here, we elect to extend the original notation of Geanacopoulos et al. (21) to a three-character binary notation

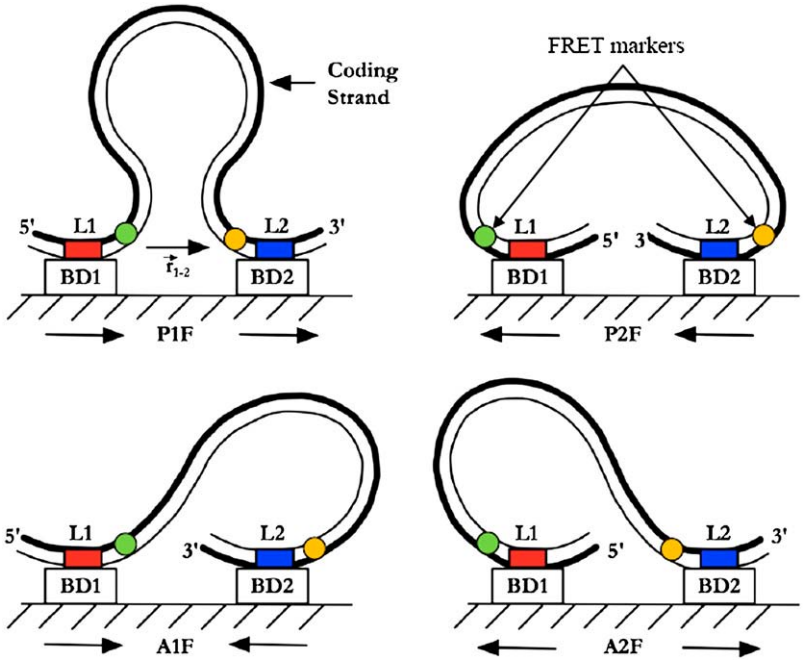


FIGURE 3 Four of eight possible binding topologies. The operator locations L1 and L2 on the substrate DNA may bind to the protein-binding domains BD1 and BD2. The operators at L1 and L2 are identical and palindromic. A three-character binary notation is used to distinguish all eight possible binding topologies and all forward (F) binding topologies are illustrated here. The strategic positioning of FRET markers along a DNA strand in recent experiments (13,14), as illustrated, provides a method to experimentally probe binding topology.

to distinguish all eight binding topologies. According to Geanacopoulos et al., (21), the first character describes the relative orientation of the 5′-3′ direction of the coding strand at the two boundary domains. If the dot product of these two directions is positive, the two directions are closer to being parallel than to being anti-parallel and the first character is assigned the letter *P*. If the dot product is negative, the two directions are closer to being anti-parallel than to being parallel and the first character is *A*. Next define the position vector $\vec{r}_{1-2}$ extending from L1 (the operator location at the 5′ end of the coding strand) to L2 (the operator location at the 3′ end of the coding strand) as illustrated in Fig. 3. The second character is chosen to be 1 if the 5′-3′ direction of the coding strand at L1 points toward the interior of the V-shaped protein, otherwise it is chosen to be 2. In other words, the second character is 1 if the dot product of $\vec{r}_{1-2}$ and the 5′-3′ direction at L1 is positive, or 2 if negative. Villa et al. (36) used *O* and *I* in their two character binary notation resulting in II = A1, OO = A2, IO = P1, and OI = P2. For the third character, we define BD1 as the protein headgroup bound to the strands labeled *H* and *G* in the LacR crystal structure given by PDB ID: 1LBG (18,56). The boundary domain BD2 is then the other headgroup. The third character distinguishes whether BD1 binds to L1 and BD2 binds to L2, as denoted by *F* for forward, or the opposite case denoted by *R* for reverse. All possible forward binding topologies are illustrated and notated in Fig. 3. If the inter-operator DNA is modeled as a homogeneous rod, the forward and reverse topologies are indistinguishable.

As discussed in greater detail in the following, Kahn and co-workers (13,14) used FRET experiments to resolve some of the binding topologies above. They bound two FRET markers on the inter-operator DNA close to the operators as shown in Fig. 3 and in Appendix A. The FRET markers included Cy3 (donor) and Cy5 (receptor). By measuring the FRET efficiencies, they estimated the distance between the two markers in the looped complex, which depends on both binding topology and possible deformation of the protein. We neglect protein deformations in this study and use these FRET measurements to infer binding topologies. In support of this assumption, we demonstrate that the differences in FRET efficiency are consistent with the computed differences due to binding topologies. Although the experimental method cannot resolve all eight binding topologies, it readily detects the P1 binding topology (which corresponds to the smallest distance between the FRET markers of ~3.5 nm). The experimental data is also consistent with our predictions of the anti-parallel topologies (which correspond to a distance of ~8.0 nm and a FRET efficiency of ~10%). The FRET efficiency becomes very poor beyond 10.0 nm.

The binding topologies discussed above determine the relative position and orientation of the two operators to within a single 2π rotation of one operator about any axis. In other words, the two operators achieve the same relative orientation after one is rotated about any axis by any whole number of turns. The additional turns produce an infinity of boundary conditions (57), corresponding to different topoisomers. Highly overwound and underwound topoisomers are expected to have high energetic cost in the LacR-DNA complex. Thus, while highly overwound and underwound topoisomers may be mechanically stable to small perturbations as determined by our computational method, they are unlikely to form in the thermal ensemble. Clearly, the topoisomer with the minimum energetic cost will be favored. Moreover, should the energetic cost of that minimum energy topoisomer remain sufficiently high, the probability of loop formation with that binding topology will diminish accordingly. Thus, as in prior predictions of looping for the LacR (33–38), we exclude all cases of linking numbers sufficiently large to generate self-contact of the inter-operator DNA. That said, computations with self-contact and even the formation of plectonemes are possible using this computational rod model upon the addition of a suitable contact law as demonstrated in Goyal et al. (58).

Including sequence-dependent, intrinsic curvature

We now turn our attention to defining the intrinsic curvature/twist of the inter-operator DNA from knowledge of its sequence following the three steps below:

1. A web tool (16,17) is used to construct the stress-free all-atom representation (PDB file) of the entire sequence of each DNA given in Appendix A at zero temperature based on the consensus tri-nucleotide model (16). This web tool outputs a protein data-bank file giving the coordinates of each atom.
2. A smooth (at least C^3 continuous) curve $R_0(s)$ is interpolated through the chain of atoms to approximate the helical axis averaging over basepair origins (59) as detailed in Appendix B.
3. We define the cross-section fixed unit vectors $a_1(s,t)$, $a_2(s,t)$, and $a_3(s,t)$, such that they align with the normal $\hat{n}(s)$, binormal $\hat{b}(s)$, and tangent $\hat{t}(s)$ unit vectors, respectively of $R_0(s)$ (60). The intrinsic curvature and twist of the helical axis are determined by the principal curvature $\kappa_P(s)$ and geometric torsion $\tau(s)$ of $R_0(s)$ (60). To understand the term ''principal curvature'', imagine the best-fit circle that most closely approximates the space curve of the helical axis at a given position *s*. The radius of this circle is referred to as the radius of curvature at *s* and its reciprocal is referred to as the principal curvature at *s*. Alternatively, the principal curvature measures the change in orientation of the tangent $\hat{t}(s)$ local to *s*. Moreover, the plane of the best-fit circle proposed above (referred to as the osculating plane) changes orientation with position *s* along the space curve. The geometric torsion measures the change in orientation of this plane with position *s*. The components of the vector $\kappa_0(s)$ with respect to the triad $\{a_i,(s,t)\}$ are $\{0\ \kappa_P(s)\ \tau(s)\}$ and they are employed in Eq. 1 to capture the effects of sequence-dependent intrinsic curvature/twist on looping.

The steps above can also be reversed and doing so allows one to reconstruct an approximate, all-atom representation of the deformed inter-operator DNA, from the computed helical axis of the rod model. To this end, we assume that the basepair atoms can only undergo a rigid body motion and therefore their positions remain fixed with respect to the triad $\{a_i,(s,t)\}$ attached to the helical axis. Thus, the locations of the basepair atoms can be computed by tracking the position and orientation of the triad $\{a_i,(s,t)\}$, which are known directly from the output of the computational rod model. We emphasize that this procedure leads only to an estimate of the final conformation and further refinements might also be possible via subsequent relaxation through MD simulation.

RESULTS

The methods above are used to explore the topology and energetics of LacR-DNA loops with different inter-operator sequences. We include results from three numerical studies that in combination reveal the overall effects of sequence-dependent intrinsic curvature and corroborate three major conclusions from experimental studies (12,61); specifically:

The sequence-dependent intrinsic curvature can reduce the energetic cost of looping (12).

The sequence-dependent intrinsic curvature influences loop topology and the distribution of topoisomers (12).

Looping energy depends strongly on operator orientation as opposed to operator separation (61).

To start, we first reexamine the computed loops for the wild-type sequence (see Appendix A) and contrast the results with those in Balaeff et al. (37), where sequence-dependent intrinsic curvature was not incorporated in the computed results. We then explore looping in four other sequences (see Appendix A) with designed A-tract bends (12). Finally, we evaluate how the loop elastic energy depends on operator

separation (length of the inter-operator sequence) as well as operator orientation.

Looping in wild-type sequence

Fig. 4 *A* depicts candidate stress-free, zero temperature conformations of the 77-bp inter-operator DNA (defined in Appendix A) for the wild-type sequence as predicted by four different models from the web tool (17). These include electrophoresis di-nucleotide (62), straight B-DNA, consensus tri-nucleotide (16), and NMR di-nucleotide (63). In general, it is expected that different models will predict different intrinsic curvature which ultimately influences looping behaviors. For the wild-type sequence, however, the intrinsic curvatures predicted by the di-nucleotide and the tri-nucleotide models are all modest and similar, yet these differences can influence the looping energy and topology as discussed further in this article. To demonstrate the sensitivity of looping on the sequence-dependent curvature, we contrast results based on the consensus tri-nucleotide model with those of the straight B-DNA model. The straight B-DNA with 3.46 Å height and 34.6° twist per basepair step correspond to the homogeneous rod model used in the literature (33–38). The consensus tri-nucleotide model (16) accounts for sequence-dependent shape modeled in the nonhomogeneous rod. The helical axis of the B-DNA is straight, which renders the principal curvature and geometric torsion identically zero, i.e., $\kappa_P(s) \equiv 0$ and $\tau(s) \equiv 0$. By contrast, $\kappa_P(s) \neq 0$ and $\tau(s) \neq 0$ for the consensus tri-nucleotide model (16), which yields a distinct three-dimensional curve for the helical axis. Fig. 4 *B* illustrates the resulting principal curvature and geometric torsion computed from the consensus tri-nucleotide model (as functions of nondimensional contour length s), which are then input to the computational rod model for studying looping.

For the two models we first examine the loops with P1 binding topologies (refer to Fig. 3 for definition of P1 binding topology). Fig. 5, *A* and *B*, illustrates the loops for both models—homogenous B-DNA (*red*) and nonhomogenous consensus tri-nucleotide model (*blue*, P1F and *green*, P1R). Note that for homogeneous B-DNA, the binding topologies P1F and P1R yield identical loops and hence we designate them simply by P1. Computations reveal two loops (without self-contact) of the inter-operator DNA for each binding topology, one undertwisted (Fig. 5 *A*) and the other overtwisted (Fig. 5 *B*). The principal curvature $|\kappa \times a_3|$ and overtwist density $(\kappa_3 - \tau)$ for each loop are reported in Fig. 5, *C* and *D*, together with the intrinsic (principal) curvature of the stress-free, zero temperature state (*black*) for reference. For the case of vanishing intrinsic curvature (homogeneous B-DNA), the above formulation should replicate the results of Balaeff et al. (37). Indeed, the computations shown in Fig. 5, *C* and *D*, for homogeneous B-DNA (*red*) faithfully reproduce the principal curvature and overtwist density reported in Balaeff et al. (37) to within 0.2 deg/bp. The model in Balaeff et al. (37) also takes into account bending anisotropy and electrostatics of the DNA loop, while this model describes an isotropic rod with nonhomogeneous intrinsic curvature but without electrostatics. Thus, the associated homogeneous rods for these two formulations, while very similar, are not strictly equivalent. Nevertheless, their predictions under these circumstances for the example of Fig. 5 are in very close agreement as indicated.

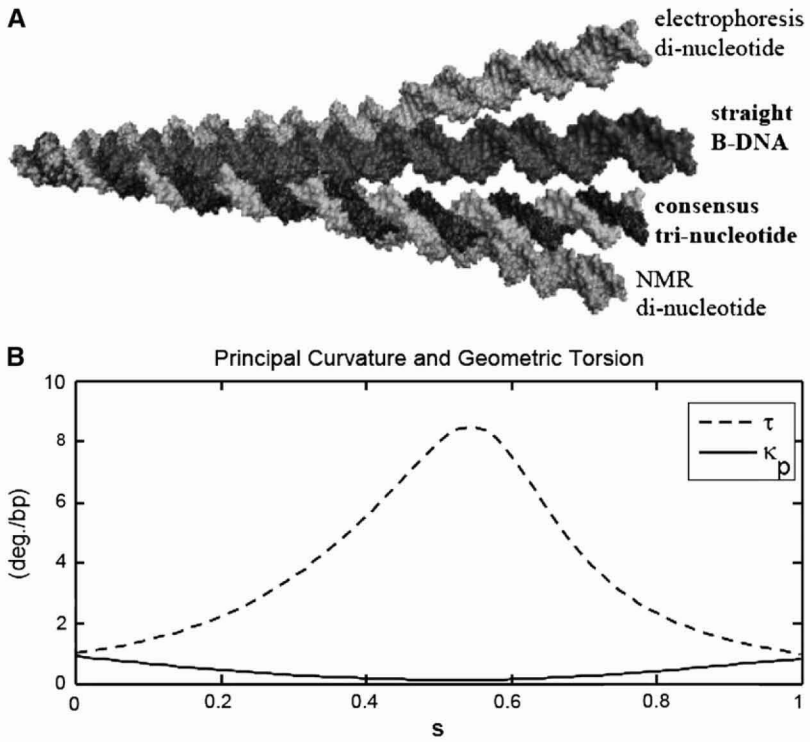
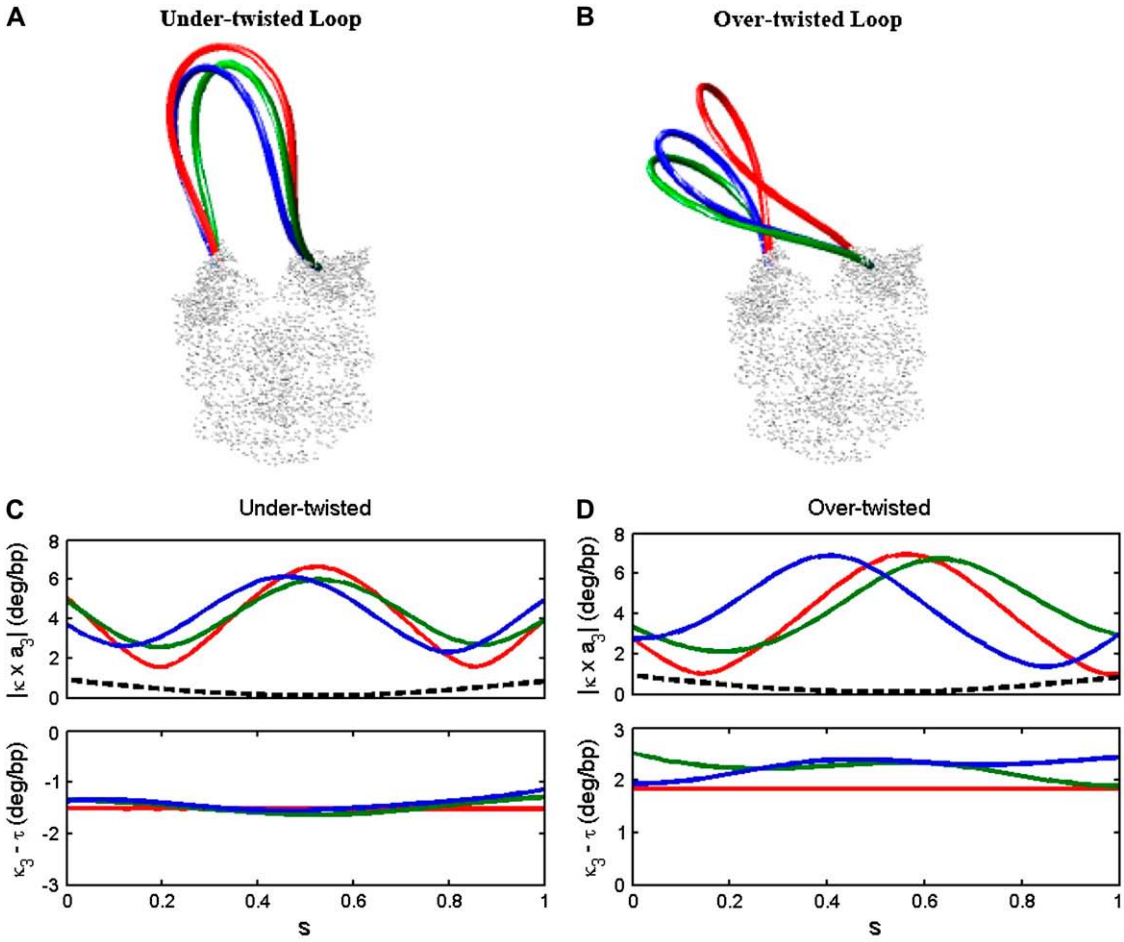


FIGURE 4 (*A*) Comparison of four different models of stress-free, zero-temperature, wild-type, inter-operator DNA: electrophoresis di-nucleotide (62), straight B-DNA, consensus tri-nucleotide model (16), and NMR di-nucleotide (63). The left boundary basepair for the four models are aligned. (*B*) Principal curvature and geometric torsion of the helical axis for the consensus tri-nucleotide model (16) as a function of (nondimensional) contour length s.



E

Binding Topology	Under-twisted Loop					Over-twisted Loop				
	E(kT)	ΔTw	Wr	ΔLk	Inter-ference	E(kT)	ΔTw	Wr	ΔLk	Inter-ference
Homogeneous B-DNA										
P1	34.0	-0.32	-0.01	-0.33	No	38.0	0.39	0.25	0.64	No
A1	26.5	-0.27	-0.10	-0.37	No	34.0	0.36	0.23	0.59	No
A2	30.5	-0.42	-0.08	-0.50	No	30.0	0.41	0.09	0.50	No
P2	76.0	-0.66	-0.20	-0.86	No	35.5	0.06	0.07	0.13	No
Non-homogeneous consensus tri-nucleotide model (16)										
P1F	35.0	-0.32	-0.03	-0.35	No	50.5	0.55	0.11	0.66	No
A1F	26.0	-0.16	-0.11	-0.27	No	41.5	0.40	0.27	0.67	No
A2F	27.0	-0.29	-0.34	-0.63	No	39.0	0.47	0.13	0.60	No
P2F	81.5	-0.30	-0.39	-0.69	Yes	52.5	0.07	0.47	0.54	No
P1R	34.5	-0.29	-0.02	-0.31	No	50.0	0.52	0.14	0.66	No
A2R	25.5	-0.13	-0.08	-0.21	No	43.5	0.34	0.40	0.74	No
A1R	28.5	-0.30	-0.11	-0.41	No	37.5	0.48	0.08	0.56	No
P2R	76.0	-0.33	-0.14	-0.47	Yes	50.5	0.12	0.17	0.29	No

FIGURE 5 (*A* and *B*) Computed LacR loops for wild-type, inter-operator DNA for LacR. Loops accounting for intrinsic shape (binding topology P1R is shown in *blue* and binding topology P1F is shown in *green*) differ from those that ignore intrinsic shape (homogeneous B-DNA, binding topology P1 shown in *red*). Two solutions for the loop exist for each binding topology (ignoring self-contact)—one is undertwisted (*A*) while the other is overtwisted (*B*). (*C* and *D*) Principal curvature and overtwist density of all loops above shown as functions of (nondimensional) contour length coordinate *s*. The principal curvature for the (stress-free) consensus model (*black*) is reproduced for comparison. (*E*) Table summarizes the total overtwist (above the natural helical twist) ΔTw, writhe Wr, linking number ΔLk, and loop elastic energy E for all the binding topologies. The writhe Wr is computed using Method 1a described by Klenin and Langowski (74). We form a closed loop for calculating writhe by adding a straight segment $\vec{r}_{1-2}$ that connects the two ends of the DNA bound to the protein in Fig. 3. The stress-free B-DNA is characterized by a uniform twist of 34.6°/bp, zero principal curvature, and rise of 3.46 Å/bp. The bending and torsional persistence lengths are assumed to be 50 nm and 75 nm (39–41), respectively, yielding a bending to torsional stiffness ratio of 2/3. The term *Interference* is used whenever a visual check reveals DNA-protein steric interference.

A summary of the total overtwist ΔTw, writhe Wr (computed as discussed in the caption to Fig. 5), link ΔLk, and loop elastic energy E for all binding topologies is provided in Fig. 5 *E*. The lowest elastic energy (again without steric interference) is highlighted in blue font, the second lowest is in green font, and the highest in red font.

Looping in four sequences with designed A-tract bends

We now utilize the same methods with the consensus tri-nucleotide model to explore the role of intrinsic curvature in the four highly curved sequences with phased A-tract bends introduced in Mehta and Kahn (12). The four sequences, denoted by control, 11C12, 7C16, and 9C14, are defined in Appendix A and their predicted stress-free, zero-temperature conformations are illustrated in Fig. 6. The graphics showing all atom details were created from PDB files using Visual Molecular Dynamics software (64). Note that the control sequence is nearly straight while the other three sequences have similar A-tract bends but with vary large differences in the helical phase (of ~70°).

It is important to understand that the experimental results from the Kahn sequences (12) in principle include effects from both sequence-dependent elasticity as well as sequence-dependent intrinsic curvature. Note, however, that the four sequences (Appendix A) are very similar and thus it is likely that only very minor differences exist in the sequence-dependent elasticity of the molecules. By contrast, there exist very major differences in the phasing of the A-tract relative to the operators and hence major differences exist in the sequence-dependent intrinsic curvature of the molecules.

For each sequence, we used the computational rod model to compute the inter-operator loops formed by LacR binding

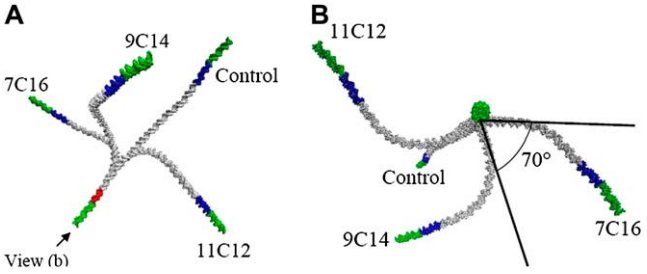


FIGURE 6 Two views of the stress-free, zero-temperature conformations of four designed inter-operator DNA sequences (12) as computed using the consensus tri-nucleotide model (16). The first basepair of each sequence is assigned the same position and orientation. The operator regions are shown in green and the red and blue segments are the same in all the four constructs, but the silver segments are different in each of them. In the control sequence, the silver segment is nearly straight, while in the others it has A-tract bends between two straight linkers of different lengths (refer to Appendix A). The control sequence is nearly straight as best observed in view *A*. For the three variants, the inter-operator sequences contain a series of A-tract bends between two nearly straight linker regions of differing lengths. The different length linker regions lead to bends that are phased by ~70° about the helical axis of the control as best observed in view *B*.

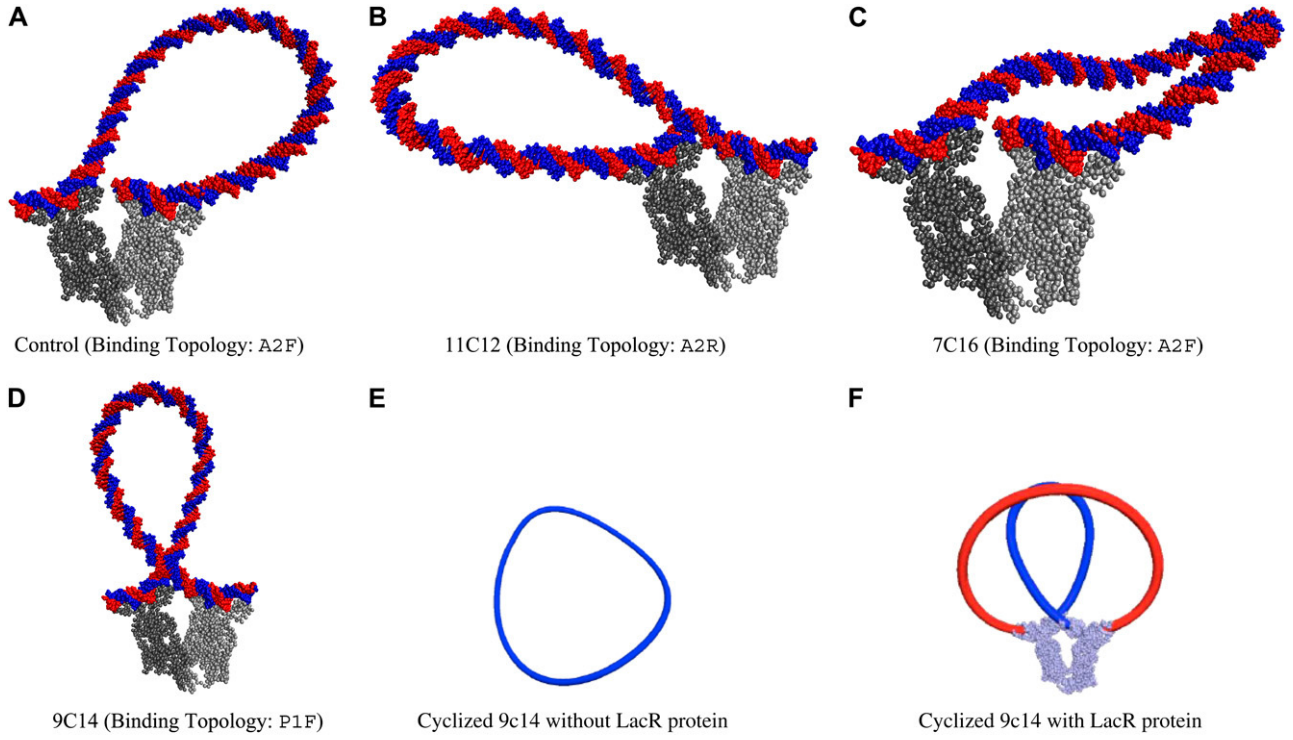
and for all possible (eight) binding topologies (refer to Fig. 3). As in the wild-type case, multiple stable loops are possible for each binding topology. Fig. 7, *A–D*, illustrates the loop that achieves the minimum elastic energy for each sequence. In the table in Fig. 7 *G* we report the number of bp for the inter-operator sequence (as defined in Appendix A), binding topology, loop elastic energy, total overtwist (ΔTw), writhe, and radius of gyration of the LacR-DNA complex for the minimum energy loops (illustrated) as well as those having the second lowest elastic energy (not illustrated). The loops with the second lowest elastic energies might correspond to the most probable meta-stable states of the Boltzmann distribution. If their free energies are close to those of the stable states (lowest energy states), the meta-stable states may coexist with the stable states in a thermal environment with some likelihood of inter-conversion. In fact, if one were to account for all other components of the free energy (see Model Limitations and Extensions below), one may arrive at a different conclusion regarding which equilibrium yields the global free energy minimum relative to the first approximation here based on the elastic energy of the inter-operator DNA. Note also that while some loops have very comparable elastic energies, their binding topologies and geometrical properties (e.g., whether overtwisted or undertwisted) are often altogether different.

Next, we return to the cyclization experiments conducted by Mehta and Kahn (12), which provide an opportunity to compare experimental linking number (Lk) measurements using gel assays with the predictions from the model. We begin by reviewing the datum used to measure ΔLk in Mehta and Kahn (12), which is the Lk of the DNA cyclized into a mini-circle in the absence of LacR as illustrated for the 9C14 sequence in Fig. 7 *E*. Next, consider the formation of the looped LacR-DNA complex with the bent DNA as illustrated in Fig. 7 *D*. In Mehta and Kahn (12), DNA tails outside the inter-operator loop (refer to Appendix A) are ligated to form a mini-circle as depicted for the 9C14 sequence in Fig. 7 *F*. Finally, the ΔLk of this mini-circle was measured experimentally with respect to the prior datum (12).

Having established the experimental procedure for measuring ΔLk, we now perform the analogous procedure using the computational rod model. In particular, we first compute the mini-circles to provide the datum for subsequent computations of ΔLk. We then simulate cyclization after the numerical computation of the looped LacR-DNA complex. To this end, the tails are deformed into a loop of minimum elastic energy and occupy the only remaining binding topology available with the LacR. The energies of each closure segment and ΔLk for each cyclized LacR-DNA complex are reported in the table in Fig. 7 *G* for all four sequences.

Influence of inter-operator length and phase

The sequences of Mehta and Kahn (12) considered above differ both in the location/phase of the A-tract bend as well as



G

	Control (144 bp)		11C12 (141 bp)		7C16 (131 bp)		9C14 (142 bp)	
	min	2nd min	min	2nd min	min	2nd min	min	2nd min
Inter-operator Binding Topology	A2F	P1R	A2R	A1R	A2F	A1F	P1F	P1R
E (kT)	12.0	12.5	7.5	10.5	8.5	11.5	11.0	11.5
ΔTw	0.03	-0.07	-0.17	0.17	0.14	-0.28	0.12	0.11
Wr	-0.01	0.15	-0.21	0.18	-0.04	-0.16	0.27	0.37
Rad. of Gyr. (nm)	8.44	7.68	8.05	8.19	7.55	7.69	7.66	7.61
Closure Segment Binding Topology	A2R	P2F	A2F	A1F	A2R	A1R	P2R	P2F
E (kT)	12.5	16.0	10.0	14.0	12.5	11.0	16.5	16.0
ΔLk	0	1	-1	0	0	-1	1	1

FIGURE 7 (*A–D*) Lowest energy solutions for four designed sequences and the associated binding topology. (*E*) Computed geometry of the 9C14 DNA mini-circle cyclized in the absence of LacR. The *Lk* of this mini-circle is used as a datum for the computation of Δ*Lk* of the mini-circle cyclized in the presence of LacR as measured by experiment. (*F*) Computed geometry of the mini-circle cyclized in the presence of LacR for the 9C14 sequence. Here the blue loop represents the inter-operator DNA, while the red loop represents the loop formed by ligation of DNA tails outside the inter-operator DNA. (*G*) Table summarizing energetic and topological data for the loops. The upper portion reports the binding topology, loop elastic energy, twist relative to the stress-free configuration, writhe for inter-operator loops (computed as discussed in the caption to Fig. 5), and radius of gyration of LacR-DNA complex for the minimum and second minimum elastic energy conformations. The largest of all the minimum energies is highlighted in red and the lowest in blue. The middle portion reports the binding topology and elastic energy of the loop formed by ligation of the DNA tails. The bottom portion (the *last row*) reports the theoretically computed Δ*Lk*, which is the change in the linking number of the mini-circle cyclized in the presence of LacR from the linking number of the corresponding mini-circle cyclized in the absence of LacR. This theoretically computed Δ*Lk* is analogous to the experimentally observed Δ*Lk* (12).

the number of basepairs of the inter-operator DNA, which range from 131 bp (7C16) to 144 bp (Control). The elastic energy of the resulting loops is influenced by both factors. Therefore, the gel shift experiments on the looping of the four sequences in Mehta and Kahn (12) were influenced by both effects. In the following results, we isolate these influences on elastic loop energy.

Adding (or subtracting) a single basepair is expected to change the loop energy by changing 1), the length of inter-operator DNA (e.g., by 3.46 Å/basepair); and 2), the relative orientation of the operators by one unit of basepair twist (e.g., by 34.6°/basepair). The first effect is negligible for the four designed sequences considered above, since the relatively small differences in contour length (<8%) generate negligibly small changes in the stiffness of the inter-operator DNA. By contrast, changes in the relative orientation of the operators may yield as much as a 50% change in elastic energy as shown in the results below.

Fig. 8 illustrates the computed loop elastic energy for the Control sequence (modeled as straight B-DNA) with the P1 binding topology. The line curves (*solid* and *dashed*) represent the energy computed by simply rotating one operator about its tangent vector $\hat{t}$ in increments of the nominal basepair twist (34.6°) while holding the number of basepairs constant (= 142). The two curves distinguish two computed loops; one undertwisted (*solid*) and one overtwisted (*dashed*). The markers (*circles* and *stars*) represent the energy computed by adding basepairs and thereby simultaneously increasing the length of the inter-operator DNA as well as changing the relative orientation of the operators. The two sets of markers distinguish two computed loops; one undertwisted (*circles*) and one overtwisted (*stars*).

DISCUSSION

We open this discussion by first establishing the predictive ability of the computational model to describe the effects of intrinsic curvature in the LacR-DNA complex. To this end, we immediately draw from a wealth of experimental data from the Kahn lab on highly curved DNA sequences (12–14) and discuss both quantitative and qualitative agreements between theory and experiment. We then discuss overall effects of the sequence-dependent intrinsic curvature and analyze three major conclusions based on the model simulations as well as the experimental studies of the LacR-DNA complex (12,61). The three major conclusions address: 1) how sequence-dependent intrinsic curvature reduces the energetic cost of looping; 2) how sequence-dependent intrinsic curvature influences loop topology and the distribution of topoisomers; and 3) how looping energy is influenced by operator orientation and separation. Finally, we duly note both limitations and extensions of the computational model.

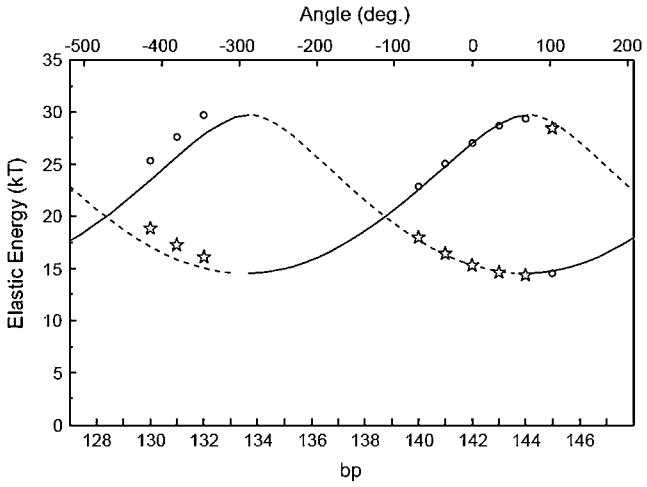


FIGURE 8 The influence of operator orientation and inter-operator length on loop elastic energy for straight B-DNA (Control) with P1F/P1R binding topology. The line curves (*solid* and *dashed*) illustrate the periodic variation in elastic energy obtained by rotating one operator about the helical axis in increments of the basepair twist (in this model 34.6°/basepair) while keeping the inter-operator length constant (142 bp); refer to scale on top for relative angular orientation of operators. The markers (*circles* and *stars*) illustrate the same variation obtained by adding basepairs and thereby both rotating one operator as well as increasing the inter-operator length (in this model 3.46 Å/basepair); refer to scale on bottom for bp number. (Overtwisted solutions are denoted by *dashed-line curves* and *stars*, whereas undertwisted solutions are denoted by *solid-line curves* and *circles*.)

Comparisons with experiments on highly bent sequences

The numerous experimental observations made by Kahn and co-workers on highly bent sequences provide abundant data for comparison with the model developed herein. These experimental results, which are derived from gel electrophoresis (12), bulk FRET studies (13), and recent single-molecule FRET studies (14), are discussed vis-à-vis predictions from the computational model beginning with the single-molecule FRET experiments.

Single-molecule FRET experiments

A promising comparison follows from the recent experiments from Kahn and co-workers (14), who infer loop and binding topology using single-molecule FRET techniques for one of the highly bent sequences (9C14) evaluated herein. They observe that ''the LacI-9C14 loop exists exclusively in a single closed form exhibiting essentially 100% ET.'' Here the ''single closed form'' is equivalent to the P1F and P1R binding topologies. The tabulated loop strain energies in Fig. 7 confirm that the lowest two energy conformations for the sequence 9C14 correspond to the P1F and P1R binding topologies in solid agreement with the SM-FRET measurements of Morgan et al. (14).

Bulk FRET experiments

The bulk FRET experiments of Edelman et al. (13) focus on the 11C12 sequence in the looped complex. The measurements reveal a FRET efficiency of ~10% that corresponds to a fluorophore separation of ~8 nm. Our model also predicts this separation to be 8 nm as arising from two anti-parallel loop topologies for the closed V-conformation of the protein. In particular, note from Fig. 7 that the two lowest energy loops of the 11C12 sequence are both anti-parallel (A2R and A1R) binding topologies and both yield the expected 8 nm fluorophore separation observed experimentally.

Edleman et al. (13) also report that the average bulk FRET efficiency of an unlooped sandwich complex is nearly the same as that of the looped 11C12 complex. The sandwich complex occurs when two fragments of DNA bind to the LacR protein, one on each head region, and thus no intervening loop is formed. With no loop, there are no reaction forces on the protein, and the protein is therefore stress-free. Under these circumstances, we expect 1), that the stress-free shape of the

protein will coincide with the crystal structure (closed V-form); and 2), that the FRET efficiency for the sandwich complex will be close to that of the looped complex with anti-parallel binding topology. In the sandwich complex, if we assume all binding topologies are equally likely then the average fluorophore-separation would be about the same as that of the anti-parallel topologies. The experimental fact that the sandwich and 11C12 looped complexes yield nearly equivalent FRET distances (~8 nm), and that this distance agrees with our computed anti-parallel binding topology provides further evidence that 1), the 11C12 loop binds with an anti-parallel binding topology; and 2), the protein does not deform significantly when bound to the 11C12 sequence.

Cyclization assays formed with LacR-DNA complex

Mehta and Kahn (12) performed cyclization assays for their designed sequences. In these assays, they compare the linking number distribution for DNA cyclized after the formation of the looped LacR-DNA complex to that of DNA cyclized in the absence of LacR. Thus, they measure a relative linking number ΔLk of the DNA-protein mini-circle with respect to the mini-circle of the DNA alone. The distributions of the relative linking numbers observed in the cyclization assays are reproduced in Table 1. We tabulate in Fig. 7 *G* our computed ΔLk of the stable and the first meta-stable loops for each sequence. The computed results immediately reveal that only loops formed with the P1 binding topologies express the topoisomer with $\Delta Lk = +1$. This prediction agrees with the experimental results (12) that show that the sequence 9C14, which preferentially forms closed loops (i.e., loops with the P1 binding topology as discussed above (14)), is the only sequence that significantly expresses $\Delta Lk = +1$ (50%).

Measurements of cyclization kinetics (12) provide yet another opportunity for model comparison. In particular, the experiments reveal that mini-circles with $\Delta Lk = +1$ take the longest to form. The computational model provides a clear explanation for this observation. As discussed above, the topoisomer with $\Delta Lk = +1$ arises when the inter-operator loop forms with the P1 binding topology. The P1 binding topology for the inter-operator DNA then dictates that the DNA tails must cyclize with the remaining P2 binding topology. The computational rod model predicts that such P2 closing segments consistently incur the greatest elastic energy cost and hence yield a reduced rate of cyclization (refer to Fig. 7 *G* and the reported elastic energy for closure of the DNA tail segment). This also supports the arguments provided in Mehta and Kahn (12) to describe the reduced rate of cyclization for the $\Delta Lk = +1$ mini-circles.

TABLE 1 Experimentally measured topoisomer distribution by relative link ΔLk as reported in Mehta and Kahn (12)

ΔLk	Control	11C12	7C16	9C14
−1	0%	10%	5%	10%
0	100%	90%	85%	40%
+1	0%	0%	10%	50%

Competition assays

Mehta and Kahn (12) measured the stability of LacR-DNA loops formed by each sequence via competition reactions in which labeled DNA in the looped complex was reacted with a 50-fold excess of unlabeled DNA. In this competition reaction, the unlabeled DNA tends to replace the labeled DNA in the looped complex until equal proportions of labeled and unlabeled DNA is reached in both looped and unlooped states. The more stable the loop, the longer the time to achieve this equilibrium. In particular, the relative stability of the loops formed by each sequence was assessed by measuring the percentage of residual labeled DNA in the looped complex after 24 h. Mehta and Kahn (12) report that the (straight) control sequence forms the least stable loops with only 3.8% labeled DNA remaining bound to LacR, while 11C12 forms the most stable loop with 63.3% labeled DNA remaining bound. These experimental observations align with the computations of loop elastic energies for these two sequences. As reported in Fig. 7 *G*, the 11C12 sequence has the lowest elastic energy cost of looping (confirming the most stable looped complex) while the control sequence has the highest elastic energy cost of looping (confirming the least stable looped complex). The remaining two sequences (7C16 and 9C14) form minimum energy loops having intermediate elastic energies relative to the extremes (control and 11C12), as shown in Fig. 7 *G*. However, in this instance, the model does not correctly predict their relative stabilities.

Gel mobility assays versus loop size

The gel mobility assays of Mehta and Kahn (12) reveal the relative speeds and hence the relative sizes of the looped LacR-DNA complexes formed by the designed (bent) sequences. In particular, the looped 7C16 sequence moves fastest through the gel with the 9C14 sequence coming in a close second. The 11C12 and the straight (control) sequences exhibit similar and slower mobilities and therefore likely form loops of similar and relatively larger size. Consider next Fig. 3, which suggests that the P1 binding topology is likely to yield the smallest loop compared to all other topologies followed by the two anti-parallel topologies (A1 or A2). The computed results of Fig. 7, *A–D*, can thus be used to infer the relative sizes (hence relative speeds in the gel) by noting the preferred binding topologies. In particular, the straight (control) and 11C12 sequences preferentially form loops with the anti-parallel binding topologies whereas the 9C14 sequence preferentially forms loops with the P1 binding topology. Thus, the 9C14 sequence is expected to have a greater speed through the gel relative to either the straight (control) or the 11C12 sequence and the two latter are expected to have nearly the same speed. An exception, however, is the 7C16

sequence which, despite its computed preference for an antiparallel binding topology, still moves through the gel with the greatest relative speed. Note however, that this same sequence also has a shorter contour length (shorter by ~10 bp from the other ~142 bp sequences), which may unduly bias it to have a smaller overall loop size relative to the other sequences.

This suspicion is confirmed by a more quantitative analysis of loop size from the model. The computed topology of the intervening DNA loop and the known structure of the LacR can now be used to compute the radius of gyration of the entire looped LacR-DNA complex. We report the computed radius of gyration for the minimum energy (and second minimum energy) looped complexes for each sequence in Fig. 7 *G*. The predicted sizes of the minimum energy loops strongly correlate to the measured mobilities (12). In particular, the minimum energy looped complexes formed with the control and 11C12 sequences have the greatest radii of gyration and experimentally have the slowest mobilities. By contrast, the fastest mobilities were observed for the looped complexes formed with the 9C14 and 7C16 sequences which have the smallest computed radii of gyration. We again note that the 7C16 sequence is also shorter by 10 bp and this difference strongly influences its computed radius of gyration, despite its preferred A2 binding topology. In interpreting these results, note that computed radius of gyration for the protein alone is ~4.6 nm, so the contribution by the intervening DNA loop to the overall radius of gyration of the looped LacR-DNA complex is quite significant.

Overall effects of sequence-dependent intrinsic curvature

Having established the predictive capability of the computational model, we now discuss the overall effects of sequence-dependent intrinsic curvature by drawing on both computational and experimental findings.

First, sequence-dependent intrinsic curvature necessitates the consideration of eight distinct binding topologies, four of which are illustrated in Fig. 3. As a result of intrinsic curvature and protein asymmetry, reversing the order of the binding domains yields loops with distinct topologies; for example, compare the loops for P1F (*green*) versus P1R (*blue*) in Fig. 5 *A*. Second, sequence-dependent intrinsic curvature may greatly alter the topology of the loop relative to that predicted for homogeneous B-DNA. For the wild-type sequence, which has modest intrinsic curvature, one might not expect significant changes in writhe between loops that include or ignore this intrinsic curvature. This, however, is not always the case, as seen, for example, in Fig. 5 *E*, where the writhe of the A2F undertwisted loop ($Wr = -0.34$) (tri-nucleotide) has significantly greater magnitude than that of the A2 undertwisted loop ($Wr = -0.08$) (homogeneous B-DNA) that ignores intrinsic curvature. For the designed sequences, we do expect to see large changes in writhe due to their significant intrinsic curvature. For example, the writhe of the P1R loops for the sequence 9C14 ($Wr = 0.15$) in Fig. 7 *G* is less than one-half that of the control sequence ($Wr = 0.37$). The associated impact that these topological changes have on the energetics of looping can be substantial, as discussed in detail below. Third, sequence-dependent intrinsic curvature qualitatively alters the distribution of twist along the inter-operator DNA. While the over- or undertwist remains uniform for the (homogeneous) model for the straight B-DNA, it becomes nonuniform for (nonhomogeneous) models (46) that include intrinsic curvature (refer to Fig. 5, *C* and *D*). This general observation may open further questions about possible sequence-dependent localization of over- and undertwist and its impact in biological processes, such as facilitating or impeding promoter melting.

Sequence-dependent intrinsic curvature reduces energetic cost of looping

Our computations show that the addition of A-tract bends into the three designed sequences substantially reduces the loop elastic energy in comparison to that of control sequence (refer to Fig. 7 *G*). For example, an energy reduction of nearly 40% occurs between the minimum energy loop of the curved sequence 11C12 compared to that of the unbent control. Thus, we support the conclusion based on gel shift assays in Mehta and Kahn (12), ''Free energy cost can be decreased by incorporating designed DNA bends into looped complexes.'' Intuitively, one would expect that the sequence-dependent intrinsic curvature may conform (to some degree) to the final loop shape and particularly given the freedom afforded by eight binding topologies and the number of topoisomers occurring for each. Thus, as stated in Mehta and Kahn (12), ''the DNA whose initial structure most closely matches the optimum structure preferred by the LacI protein will form the most stable looped complex''. We can demonstrate this clearly by comparing the sequence 11C12 and the control sequence which exhibit the largest energy difference (12.0 kT vs. 7.5 kT) as shown in Fig. 9. Illustrated are the stress-free and (lowest energy) looped conformations for the sequence 11C12, Fig. 9 *A*, and the control sequence, Fig. 9 *B*, with the color scale indicating the strain energy density of the looped conformations. For the sequence 11C12, notice that modest twisting near the middle of the strand allows it to quickly conform to the looped configuration and with minimal strain energy (which is also dominated by twisting). By contrast, the nearly unbent control requires substantial bending (largely planar) to arrive at the looped conformation and with significantly greater (bending) strain energy density in the middle portion.

The conclusion that sequence-dependent intrinsic curvature reduces loop energy is also supported by the computations for the wild-type sequence, though to a lesser degree. Observe from Fig. 5 *E* the 3% reduction in elastic energy of the minimum energy loop (highlighted in *blue*) that accounts for sequence-dependent intrinsic curvature from that of the

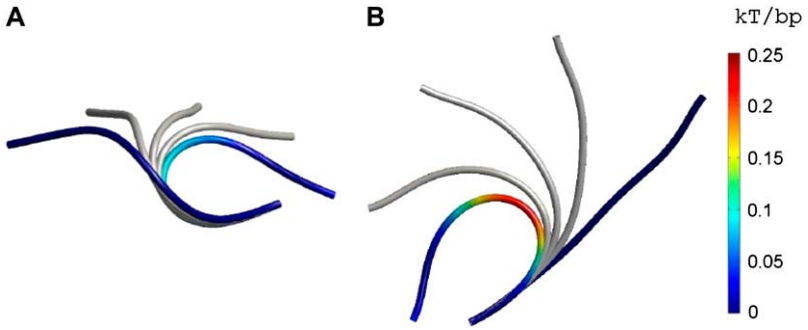


FIGURE 9 The transition from stress-free shape to looped conformation. The stress-free shapes are given in blue. The final loop geometries are shaded as a function of strain energy density (kT/bp). (*A*) 11C12; (*B*) control.

minimum energy loop that ignores the intrinsic curvature (homogeneous B-DNA). The rather modest energy reduction in this example is expected, given the very modest curvature of the stress-free shape compared to the straight B-DNA (refer to Fig. 4 *A*).

Sequence-dependent intrinsic curvature influences loop topology and distribution of topoisomers

The minimum energy loops computed for the wild-type sequence (Fig. 5) and the designed sequences (Fig. 7) reveal a wide range of binding and loop topologies. For example, note that the minimum energy loops for three sequences in Fig. 7 (including the control sequence and the bent sequences 7C16, 9C14) are all overtwisted, while the minimum energy loops for one designed sequence 11C12 and the wild-type sequence are both undertwisted. These observations support findings from the gel shift assay experiments of Mehta and Kahn (12), who state, ''Designed DNA bends can also control the shape of a DNA loop formed by Lac repressor.'' Second, there are large variations in the preferred (minimum energy) binding topologies among the four designed and the wild-type sequences. This observation suggests that energetically favorable binding topologies are in part determined by inter-operator sequence.

The sequence and associated intrinsic curvature may also strongly influence the distribution of topoisomers as suggested by the computed elastic energies reported in Figs. 5 and 7. For example, certain sequences exhibit only modest differences in elastic energy between the minimum and second lowest energy loops. Only a 2% energy difference separates these states for the wild-type sequence when accounting for the sequence-dependent intrinsic curvature (refer to Fig. 5 *E*). Likewise, the analogous energy differences for the control and 9C14 sequences in Fig. 7 are 5%. Thus, per the Boltzmann distribution, one may anticipate nearly equal concentrations of these topoisomers in experiments. By contrast, the 30% energy difference for the sequence 7C16 and 40% for 11C12 suggest substantially different topoisomer concentrations. Likewise, a large (40%) energy difference separates the undertwisted (lower energy) from the overtwisted (higher energy) topoisomers of the wild-type sequence with P1 binding topology when the sequence-dependent intrinsic curvature is included in the model (refer to Fig. 5 *E*). However, when this sequence-dependent intrinsic curvature is ignored, the energy difference is substantially reduced (to 13%), suggesting a significantly different distribution of these topoisomers. For the other binding topologies, sequence-dependent intrinsic curvature also substantially widens the energetic gap between the two topoisomers as observed in Fig. 5 *E*.

In addition, for the wild-type sequence the major trends we predict for homogeneous (straight) B-DNA are a minimum (elastic) energy conformation with 26.5 kT, a maximum with 76.0 kT, and a resulting energy range of 49.5 kT for the four possible binding topologies. By contrast, when intrinsic curvature is accounted for, the energetic minimum is lowered (albeit modestly) to 25.5 kT, the maximum is raised to 81.5 kT, and the resulting range is now 56 kT for the eight binding topologies now possible. Thus, even the modest intrinsic curvature in the wild-type widens the energy distribution among the possible loops, rendering the energetic minimum highly preferable over the others.

It is also intuitive that the intrinsic curvature may selectively raise the energetic cost of looping and that this possibility is determined as much by the binding topology as it is by the intrinsic curvature. For instance, if in one binding topology the intervening DNA is bent largely in the same direction of the intrinsic curvature, then the intrinsic curvature is likely to reduce the energetic cost of loop formation. However, a different binding topology may require the intervening DNA to bend largely in the opposite direction of the intrinsic curvature leading to an increase in energetic cost.

Looping energy: roles of operator orientation and separation

There is a considerable and expected overall reduction in the elastic energies of all of the lowest energy loops for the designed sequences relative to those of the wild-type. The four designed inter-operator sequences range from 144 bp (control) to 131 bp (7C16), and these longer inter-operator sequences relative to the wild-type (77 bp) lead to a far more flexible inter-operator DNA. For instance, elementary beam theory (65) predicts that the bending elastic energy E developed when bending an initially straight elastic beam of length L into a complete circle of radius $L/2\pi$ scales as

$E \sim 1/L$. Thus, ignoring all other complications (e.g., coupled bending/torsion leading to three-dimensional deformation, nonuniform curvature, intrinsic curvature, etc.), a sequence of 77 bp requiring 34 kT to form a circular loop would then only require 17 kT to form (a larger diameter) loop if it were 154-bp long instead. This approximate 50% reduction in elastic energy is not unlike the large elastic energy reductions observed for the far more refined computations reported in Figs. 5 and 7.

As previously noted, adding (or subtracting) a single basepair may alter the elastic energy through changing the length of inter-operator DNA by one unit of basepair rise, and/or the changing the relative orientation of the operators by one unit of basepair twist. For the relatively long sequences of Mehta and Kahn (12), the small differences in inter-operator lengths lead to negligible changes in elastic energy compared to the associated changes in the relative orientation of the two operators (61).

To understand this conclusion, refer again to Fig. 8, which shows the elastic energy for the control sequence (modeled as homogeneous B-DNA with the P1 binding topology). The energy computed by simply rotating one operator in increments of the basepair twist (*solid* and *dashed curves*) closely approximate the energy (*circles* and *stars*) computed when also allowing the inter-operator DNA to increase by increments of the basepair rise. This close agreement between these two calculations provides strong support for the claim that changes in operator orientation brought about by adding/subtracting basepairs have a far greater influence on loop energy than the associated changes in length of the inter-operator DNA. This conclusion also supports the experimental finding in Bellamy et al. (61) as discussed in Mehta and Kahn (12), ''The in vivo probability of loop formation depends strongly on the torsional phasing of the operators but relatively weakly on their separation.''

Note also the obvious periodic variation in elastic energy illustrated in Fig. 8. This computed result using rod theory supports the experimental observations that looping probability is a periodic function of the inter-operator distance (22,23,66). The period of 10.5 bp corresponds to a complete helical turn of DNA and, the results of Fig. 8 demonstrate that specific helical orientations of the operators may significantly reduce the energetic cost of loop formation by up to 50%.

Model limitations and extensions

The energy computations herein are solely restricted to the elastic (or strain) energy of the loop. To assess thermal stability, one needs to determine the free energy difference between looped and unlooped states. Major contributions to the free energy include: 1), loop elastic energy; 2), protein deformation energy; 3), entropy; 4), DNA-protein surface binding energy; and 5), electrostatic potential between the negatively charged phosphates in the DNA backbone. Of these, we believe that entropy and the surface binding energy would remain relatively constant for variations within a class of sequences (i.e., for topoisomers of the wild-type or for topoisomers of the designed sequences considered herein). We also emphasize the importance of elastic energy relative to entropy in looping of short DNA segments by noting Balaeff et al. (33), who conclude that looping of short DNA (less than three persistence lengths) is dominated by the elastic energy of DNA.

By contrast, contributions from the loop and protein deformation energies might vary significantly and, as a result of associated conformation changes, so might the electrostatic repulsion. For instance, the binding co-cooperativity of the two operator sites depends on their electrostatic repulsion (67), which decays exponentially with operator separation per the Debye-Hückel approximation. Some of these additional influences could, in fact, be approximated in the context of a computational rod model for DNA (19).

For example, the formulation herein tacitly assumes a rigid protein as determined from the crystal structure. However, the effects of protein flexibility on the loop could be captured by replacing the fixed (Dirichlet) boundary conditions with elastic (mixed Neumann-Dirichlet) boundary conditions that model the equivalent flexibility at the DNA-protein interface. Molecular dynamics (MD) simulations have suggested that flexibility of the LacR derives primarily from the head regions (38) while the possibility of flexibility in the V-region has also been suggested in prior studies (44,45). Similarly, the entire LacR might also be approximated by a small number of rigid bodies with concentrated flexibility (stiffness) at the V-region and at the protein heads. Coupling this low-dimensional protein model with the elastic rod model of DNA would allow one to capture the elastic deformation of the entire protein/DNA complex in an approximate manner. This might provide initial conditions for MD simulations of the complex or possibly obviate the need for full MD simulations altogether (38). Protein flexibility will likely lower the free energy cost of looping, may facilitate loop formation in shorter DNA segments, and may also affect the preferred binding topology and topoisomer. For example, the model simulations of the literature (32,38) confirm a lowering of the free energy for LacR-mediated looping. While Villa et al. (38) attribute this influence to protein flexibility in the head regions, Swigon et al. (32) attributed it to flexibility at the base of the V-region. Swigon et al. (32) further demonstrate that loops formed with short DNA segments (50–180 bp) tend to open the LacR V-region, while loops formed by negatively supercoiled plasmids tend to close the V-region. Along similar lines, Gemmen et al. (68) attribute the ease of formation of short loops with restriction enzymes to the protein flexibility and span, where they imagine DNA to wrap around the proteins. Similarly, protein flexibility may also affect the preferred binding topologies as explained by Adhya and co-workers (20–22,69) in the context of two GalR dimers interacting with each other in multiple ways to form different stacked V-shaped tetramers that accommodate parallel and anti-parallel topologies.

It is recognized that any long-length scale material law for DNA will surely influence the loop topology and elastic energy computed from rod theory and that further advances in determining accurate material laws are likely to follow from single-molecule experiments and MD simulations. For instance, recent MD simulations (5,8,11) have begun to reveal the sequence-dependent stiffness parameters for linear elastic behavior, while other studies (9,10,43) have begun to explore nonlinear (and inelastic) behavior, though this has also been questioned (42). It is also recognized that DNA must exhibit a strong coupling between twist and extension (70) due to its chiral (helical) construction (71) and this requires a modification of the material law used herein as proposed in Goyal et al. (19). Overall, the sequence-dependent bending and torsional stiffnesses affect computed properties of the Lac-repressor loop (3–10) and these can also be accommodated herein by accounting for spatial variations in the stiffness tensor $B(s)$ employed in Eq. 1; see, for example, Goyal et al. (58,72). Likewise, the sequence-dependent stress-free shape (or intrinsic curvature) surely affects the mechanics of looping for the LacR-DNA complex and the results herein suggest its dominant role for the sequences with designed A-tract bends. In this regard, we also recognize that the intrinsic curvature predicted using the consensus tri-nucleotide model (16) is necessarily approximate and that further efforts, such as those employing the tetra-nucleotide model (5,11), may ultimately yield more refined estimates of intrinsic curvature. The computational algorithm developed herein, however, remains independent of the means by which one first computes or measures intrinsic curvature, and thus it provides a systematic procedure for accommodating future advancements.

Despite the limitations duly listed above, the model successfully predicts major experimental findings for the LacR-DNA complexes formed with the highly curved sequences studied by the Kahn lab (12–14) and provides experimentally testable and biologically relevant insights into the effects of sequence-dependent intrinsic curvature.

CONCLUSIONS

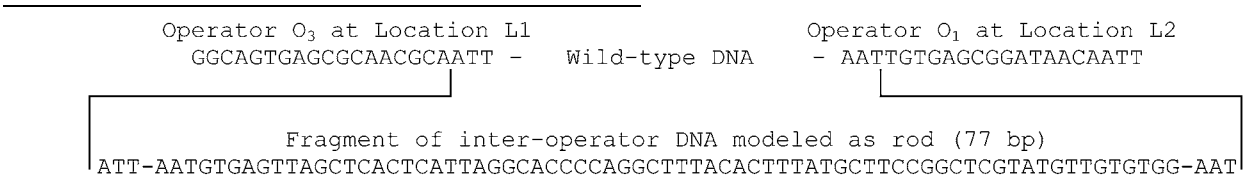
This article employs a computational rod model for the long-length scale structure of DNA as a means to explore the mechanics of protein-mediated DNA looping. Our specific objective is to understand how looping energy and topology are influenced by the sequence-dependent intrinsic curvature of the substrate DNA. We adopt the lactose repressor (LacR) protein-DNA complex as our example and consider both the wild-type sequence possessing relatively little intrinsic curvature and the highly curved sequences with designed A-tract bends introduced by Mehta and Kahn (12). Our method uses the known sequence of the inter-operator DNA to construct the intrinsic curvature of the helical axis as input to the computational rod model. Simulations allow us to predict the elastic (strain) energy required to transform the stress-free conformation into a looped conformation that complies with the known LacR-operator crystal structure.

The model successfully predicts major experimental findings for the LacR-DNA complexes formed with the highly curved sequences studied by the Kahn lab (12–14). First, the model predicts the binding topologies of the energetically favorable loops consistent with FRET measurements (13,14). Second, the model predicts the relative linking number distributions observed in cyclization assays formed with the LacR-DNA complex (12). In particular, the model predicts that mini-circles with relative linking number $+1$ are energetically favorable only in parallel binding topologies and that the associated closure energy reduces the cyclization rate. Third, the model largely predicts the relative loop stabilities as observed in competition assays (12). Fourth, the model also correctly predicts the relative speeds of looped sequences in gel mobility assays (12) upon computing and comparing the radius of gyration of the looped LacR-DNA complex. The model confirms that complexes formed with anti-parallel binding topologies move slower through the gel than those of comparable length with parallel binding topologies due to the greater compaction of the latter.

In addition, numerical studies of loop energetics and topology reveal the following major influences of sequence-dependent intrinsic curvature on the LacR-DNA complex. First, the highly curved sequences of Mehta and Kahn (12) tend to lower the energetic cost of looping, widen the energy distribution among stable and meta-stable loops, and substantially alter loop topology. Qualitatively, the inclusion of sequence-dependent intrinsic curvature also leads to nonuniform twist (or twist deficit) (46) and necessitates consideration of eight distinct binding topologies from the known crystal structure of the LacR complex. The generality and several extensions of the computational rod model are also discussed for other looping and nonlooping behaviors of DNA.

APPENDIX A: SEQUENCES AND BOUNDARY CONDITIONS

Wild-type LacR sequence (73):



Designed sequences (12) (personal communication, 2005, with Prof. J. D. Kahn, Department of Chemistry and Biochemistry, University of Maryland):

Operator O_{id}
AATTGTGAGCGCTCACAATT

Cy3 (donor fluorophore) (acceptor fluorophore) Cy5
L1 L2
Tail 1-O_{id}-AGATCTCAATTC- Linker – A tracts – Linker -TAGAATCGAAGCTAGCT-O_{id}-Tail 2

11C12
CTGTACGGATC – CGC(A)$_6$CGGGC(A)$_6$CGGC(A)$_5${CGGGC(A)$_6$CGGC(A)$_6$}$_2$CGGGC(A)$_6$C – CGCTACGCGTCC

9C14
GTACGGATC – CGG(T)$_6${GCCCG(T)$_6$GCCG(T)$_6$}$_3$GCCCG(T)$_6$G – CGCTGAACGCGTCC

7C16
ACGGATC – CGG(T)$_6$GCCCG(T)$_6$GCCG(T)$_6${GCCG(T)$_6$GCCCG(T)$_6$}$_2$G – CGCTGACAACGCGTCC

Control
CTGTACGGATCCACTGAATCCGGTGAGAATGGCAAAAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAG
CCATTACGCTCGTCATCAAAATCACTACGCGTCC

Tail 1: GCGCGC-AATTAACCCTCACTAAAGGGAACAAAAGCTGGGTACCGATATCTGCAGGTCAGTCTAGGT
Tail 2: CGTTGTGGTAAAGCTTTGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCGCCACCGCGGTGG
AGCTCCAATTCGCCCTATAGTGAGTCGTATTAC-GCGCGC
GCGCGC: BssH II sequence (sticky ends for cyclization)

In arriving at the final looped state, the boundary basepairs are made to align with their corresponding configurations known from the LacR crystal structure given by PDB ID: 1LBG (18,56). This alignment is achieved in the calculations by slowly translating and rotating the boundary basepairs (rigid body motion) from the initial stress-free configuration to the final protein-bound configuration. As a consequence, the inter-operator DNA deforms into a loop. Note that the boundary conditions account for the basepair inclination with respect to the ds-DNA helical axis. We also verified that these

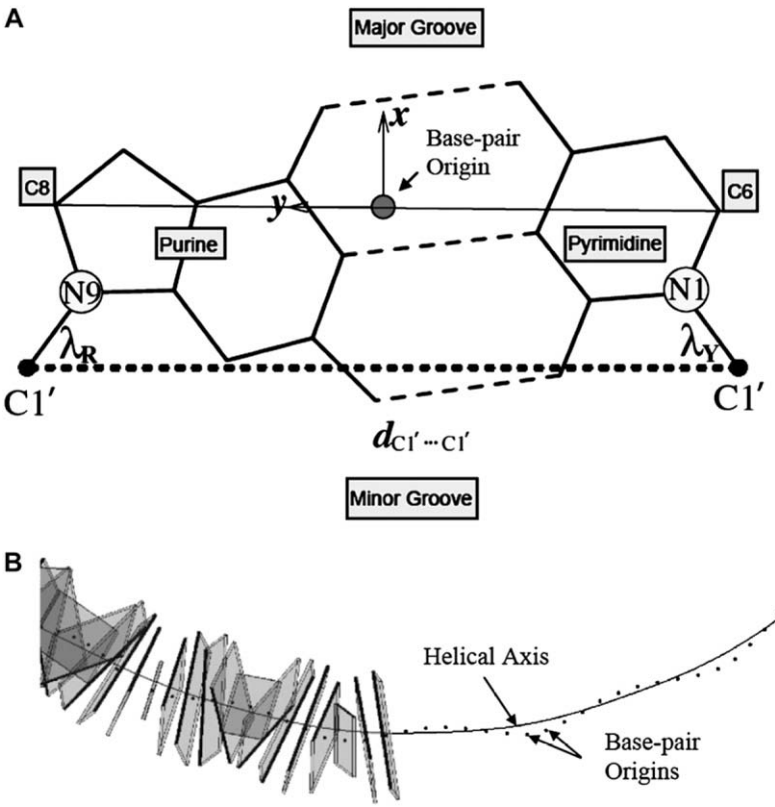


FIGURE 10 (*A*) The origin and standard basepair fixed reference frame described in Olson et. al. (59). (This figure is a modification of Fig. 1 from (59).) (*B*) Basepairs represented as semitransparent blocks with the minor-groove face shaded black. Black dots represent the basepair origins, and the solid curve represents the helical axis as computed by the moving average over a helical turn.

boundary conditions are insensitive to the choice of boundary basepairs within their immediate neighbors. To this end, we used commercial software NX Imageware (UGS, Plano, TX) to estimate the rigid body motion needed for the best alignment of the tri-nucleotide set of atoms around the chosen boundary basepair (three basepairs highlighted in *blue* and *red* in Fig. 1) and found it to be the same (within numerical tolerance) as that needed for the alignment of the chosen boundary basepair.

APPENDIX B: SEQUENCE-DEPENDENT INTRINSIC CURVATURE

The following steps were used to compute the approximate helical axis $R_0(s)$ from the stress-free all-atom representation (PDB file) of the inter-operator DNA at zero temperature (consensus tri-nucleotide model (16)).

1. Following Olson et al. (59), we first compute the origin of each basepair as the midpoint of the C6 atom of the pyrimidine and the C8 atom of the purine (see Fig. 10 *A*). A curve interpolated through the basepair origins forms an approximate helix of radius $r \approx 2.0$ Å and helical pitch ≈ 10.3 basepair (see Fig. 10 *B*). The helical axis of this curve is not straight in general due to the intrinsic (stress-free) curvature of the molecule (see Fig. 10 *B*).
2. An approximation to the helical axis $\bar{R}_0(s)$ follows from averaging the positions of the origins of the basepairs. We begin at one operator and then average the positions of the origins of the first 10 basepairs for the inter-operator DNA (see Fig. 10 *B*). We then increment by one basepair and repeat this (moving average) computation and continue to the other operator thereby developing a pointwise approximation to the helical axis $\bar{R}_0(s)$.
3. A continuous (differentiable at least three times) curve $R_0(s)$ is sought to approximate $\bar{R}_0(s)$ to compute the intrinsic curvature and torsion. We use the MatLab curve-fitting toolbox (The MathWorks, Natick, MA) to construct a C^∞ continuous curve $R_0(s)$.

The smoothing algorithm above essentially filters out nonsmooth, small-length scale effects arising from randomness of the sequence/basepair orientations. We emphasize that the computed results are insensitive to the specific approximations described in Steps 1–3 above. In particular, we have employed alternative curve fitting algorithms for Steps 1 and 3 and alternative moving averaging algorithms for Step 2. The resulting loop elastic energies typically differ by $<2\%$.

The authors acknowledge Professor Ioan Andricioaei (Biochemistry, University of Michigan) for fruitful discussions, Professor Jason D. Kahn (Department of Chemistry and Biochemistry, University of Maryland) for fruitful discussions and for providing the sequences used in his experiments, and Professor Wilma K. Olson (Department of Chemistry and Chemical Biology, Rutgers University) for suggesting alternative binding topologies. We also acknowledge comments offered by the reviewers.

The co-authors S.G., T.L., N.C.P., and E.M. gratefully acknowledge support for this research from the US National Science Foundation under grants No. CMS-0439574 and CMS-0510266. The co-authors J.-C.M. and S.B. gratefully acknowledge support from the National Institutes of Health under grant No. GM65934.

REFERENCES

1. Schleif, R. 1992. DNA looping. *Annu. Rev. Biochem.* 61:199–223.
2. Ptashne, M., and A. Gann. 2002. Genes and Signals. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
3. Coleman, B. D., W. K. Olson, and D. Swigon. 2003. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* 118:7127–7140.
4. Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA.* 95:11163–11168.
5. Beveridge, D. L., G. Barreiro, K. S. Byun, D. A. Case, T. E. Cheatham, S. B. Dixit, E. Giudice, F. Lankas, R. Lavery, J. H. Maddocks, R. Osman, E. Seibert, H. Sklenar, G. Stoll, K. M. Thayer, P. Varnai, and M. A. Young. 2004. Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(C(p)G) steps. *Biophys. J.* 87:3799–3813.
6. Manning, R. S., J. H. Maddocks, and J. D. Kahn. 1996. A continuum rod model of sequence-dependent DNA structure. *J. Chem. Phys.* 105: 5626–5646.
7. Olson, W. K., D. Swigon, and B. D. Coleman. 2004. Implications of the dependence of the elastic properties of DNA on nucleotide sequence. *Philos. Trans. Roy. Soc. London A Math. Phys. Eng. Sci.* 362:1403–1422.
8. Lankas, F. 2004. DNA sequence-dependent deformability—insights from computer simulations. *Biopolymers.* 73:327–339.
9. Cloutier, T. E., and J. Widom. 2005. DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc. Natl. Acad. Sci. USA.* 102:3645–3650.
10. Cloutier, T. E., and J. Widom. 2004. Spontaneous sharp bending of double-stranded DNA. *Mol. Cell.* 14:355–362.
11. Dixit, S. B., D. L. Beveridge, D. A. Case, T. E. Cheatham, E. Giudice, F. Lankas, R. Lavery, J. H. Maddocks, R. Osman, H. Sklenar, K. M. Thayer, and P. Varnai. 2005. Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II. Sequence context effects on the dynamical structures of the 10 unique dinucleotide steps. *Biophys. J.* 89:3721–3740.
12. Mehta, R. A., and J. D. Kahn. 1999. Designed hyperstable Lac repressor. DNA loop topologies suggest alternative loop geometries. *J. Mol. Biol.* 294:67–77.
13. Edelman, L. M., R. Cheong, and J. D. Kahn. 2003. Fluorescence resonance energy transfer over similar to 130 basepairs in hyperstable Lac repressor-DNA loops. *Biophys. J.* 84:1131–1145.
14. Morgan, M. A., K. Okamoto, J. D. Kahn, and D. S. English. 2005. Single-molecule spectroscopic determination of Lac repressor-DNA loop conformation. *Biophys. J.* 89:2588–2596.
15. Zhang, Y., A. E. McEwen, D. M. Crothers, and S. D. Levene. 2006. Statistical-mechanical theory of DNA looping. *Biophys. J.* 90:1903–1912.
16. Gabrielian, A., and S. Pongor. 1996. Correlation of intrinsic DNA curvature with DNA property periodicity. *FEBS Lett.* 393:65–68.
17. Munteanu, M. G., K. Vlahovicek, S. Parthasarathy, I. Simon, and S. Pongor. 1998. Rod models of DNA: sequence-dependent anisotropic elastic modeling of local bending phenomena. *Trends Biochem. Sci.* 23:341–347.
18. Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, and P. Lu. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science.* 271:1247–1254.
19. Goyal, S., N. C. Perkins, and C. L. Lee. 2005. Nonlinear dynamics and loop formation in Kirchhoff rods with implications to the mechanics of DNA and cables. *J. Comput. Phys.* 209:371–389.
20. Semsey, S., K. Virnik, and S. Adhya. 2005. A gamut of loops: meandering DNA. *Trends Biochem. Sci.* 30:334–341.
21. Geanacopoulos, M., G. Vasmatzis, V. B. Zhurkin, and S. Adhya. 2001. Gal repressosome contains an antiparallel DNA loop. *Nat. Struct. Biol.* 8:432–436.
22. Lewis, D. E. A., and S. Adhya. 2002. In vitro repression of the Gal promoters by GalR and HU depends on the proper helical phasing of the two operators. *J. Biol. Chem.* 277:2498–2504.
23. Dunn, T. M., S. Hahn, S. Ogden, and R. F. Schleif. 1984. An operator at −280 basepairs that is required for repression of Arabad operon promoter—addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc. Natl. Acad. Sci. USA.* 81:5017–5020.
24. Watson, M. A., D. M. Gowers, and S. E. Halford. 2000. Alternative geometries of DNA looping: an analysis using the SfiI endonuclease. *J. Mol. Biol.* 298:461–475.
25. Embleton, M. L., A. V. Vologodskii, and S. E. Halford. 2004. Dynamics of DNA loop capture by the Sfil restriction endonuclease on supercoiled and relaxed DNA. *J. Mol. Biol.* 339:53–66.
26. Schulz, A., J. Langowski, and K. Rippe. 2000. The effect of the DNA conformation on the rate of NtrC activated transcription of *Escherichia coli* RNA polymerase $\sigma(54)$ holoenzyme. *J. Mol. Biol.* 300:709–725.

27. Zhang, Y. L., and D. M. Crothers. 2003. Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* 84:136–153.

28. Merlitz, H., K. Rippe, K. V. Klenin, and J. Langowski. 1998. Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation. *Biophys. J.* 74:773–779.

29. Rippe, K., P. H. von Hippel, and J. Langowski. 1995. Action at a distance—DNA-looping and initiation of transcription. *Trends Biochem. Sci.* 20:500–506.

30. Olson, W. K. 1996. Simulating DNA at low resolution. *Curr. Opin. Struct. Biol.* 6:242–256.

31. Schlick, T. 1995. Modeling superhelical DNA—recent analytical and dynamic approaches. *Curr. Opin. Struct. Biol.* 5:245–262.

32. Swigon, D., B. D. Coleman, and W. K. Olson. 2006. Modeling the Lac repressor-operator assembly: the influence of DNA looping on Lac repressor conformation. *Proc. Natl. Acad. Sci. USA.* 103:9879–9884.

33. Balaeff, A., C. R. Koudella, L. Mahadevan, and K. Schulten. 2004. Modeling DNA loops using continuum and statistical mechanics. *Philos. Trans. Roy. Soc. London A Math. Phys. Eng. Sci.* 362:1355–1371.

34. Balaeff, A., L. Mahadevan, and K. Schulten. 1999. Elastic rod model of a DNA loop in the Lac operon. *Phys. Rev. Lett.* 83:4900–4903.

35. Balaeff, A., L. Mahadevan, and K. Schulten. 2004. Structural basis for cooperative DNA binding by CAP and Lac repressor. *Structure.* 12:123–132.

36. Villa, E., A. Balaeff, L. Mahadevan, and K. Schulten. 2004. Multiscale method for simulating protein-DNA complexes. *Multiscale Model. Sim.* 2:527–553.

37. Balaeff, A., L. Mahadevan, and K. Schulten. 2006. Modeling DNA loops using the theory of elasticity. *Phys. Rev. E.* 73.

38. Villa, E., A. Balaeff, and K. Schulten. 2005. Structural dynamics of the Lac repressor-DNA complex revealed by a multiscale simulation. *Proc. Natl. Acad. Sci. USA.* 102:6783–6788.

39. Baumann, C. G., S. B. Smith, V. A. Bloomfield, and C. Bustamante. 1997. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. USA.* 94:6185–6190.

40. Hagerman, P. J. 1988. Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* 17:265–286.

41. Strick, T. R., J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. 1996. The elasticity of a single supercoiled DNA molecule. *Science.* 271:1835–1837.

42. Du, Q., C. Smith, N. Shiffeldrim, M. Vologodskaia, and A. Vologodskii. 2005. Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci. USA.* 102:5397–5402.

43. Wiggins, P. A., R. Phillips, and P. C. Nelson. 2005. Exact theory of kinkable elastic polymers. *Phys. Rev. E.* 71.

44. Friedman, A. M., T. O. Fischmann, and T. A. Steitz. 1995. Crystal-structure of Lac repressor core tetramer and its implications for DNA looping. *Science.* 268:1721–1727.

45. Ruben, G. C., and T. B. Roos. 1997. Conformation of Lac repressor tetramer in solution, bound and unbound to operator DNA. *Microsc. Res. Tech.* 36:400–416.

46. Tobias, I., and W. K. Olson. 1993. The effect of intrinsic curvature on supercoiling—predictions of elasticity theory. *Biopolymers.* 33:639–646.

47. Strick, T., J. F. Allemand, V. Croquette, and D. Bensimon. 2000. Twisting and stretching single DNA molecules. *Prog. Biophys. Mol. Biol.* 74:115–140.

48. Bustamante, C., J. F. Marko, E. D. Siggia, and S. Smith. 1994. Entropic elasticity of λ-phage DNA. *Science.* 265:1599–1600.

49. Bryant, Z., M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli, and C. Bustamante. 2003. Structural transitions and elasticity from torque measurements on DNA. *Nature.* 424:338–341.

50. Bustamante, C., Z. Bryant, and S. B. Smith. 2003. Ten years of tension: single-molecule DNA mechanics. *Nature.* 421:423–427.

51. Olson, W. K., and V. B. Zhurkin. 2000. Modeling DNA deformations. *Curr. Opin. Struct. Biol.* 10:286–297.

52. Tobias, I., D. Swigon, and B. D. Coleman. 2000. Elastic stability of DNA configurations. I. General theory. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics.* 61:747–758.

53. Coleman, B. D., D. Swigon, and I. Tobias. 2000. Elastic stability of DNA configurations. II. Supercoiled plasmids with self-contact. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics.* 61:759–770.

54. Goyal, S. 2006. A dynamic rod model to simulate mechanics of cables and DNA. PhD Dissertation in Mechanical Engineering. University of Michigan, Ann Arbor.

55. Goyal, S., N. C. Perkins, and C. L. Lee. 2003. Torsional buckling and writhing dynamics of elastic cables and DNA. *In* 2003 ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Sept. 2–6, 2003. American Society of Mechanical Engineers, Chicago, IL, United States. 183–191.

56. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.

57. Alexander, J. C., and S. S. Antman. 1982. The ambiguous twist of Love. *Quart. Appl. Math.* 40:83–92.

58. Goyal, S., N. C. Perkins, and C. L. Lee. 2007. Nonlinear dynamic intertwining of rods with self-contact. *Int. J. Nonlinear Mechanics.* In press.

59. Olson, W. K., M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X. J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C. S. Tung, E. Westhof, C. Wolberger, and H. M. Berman. 2001. A standard reference frame for the description of nucleic acid basepair geometry. *J. Mol. Biol.* 313:229–237.

60. Stoker, J. J. 1969. Differential Geometry. Wiley-Interscience, New York.

61. Bellamy, G. R., M. C. Mossing, and M. T. Record. 1988. Physical-properties of DNA in-vivo as probed by the length dependence of the Lac operator looping process. *Biochemistry.* 27:3900–3906.

62. Bolshoy, A., P. McNamara, R. E. Harrington, and E. N. Trifonov. 1991. Curved DNA without A-A—experimental estimation of all 16 DNA wedge angles. *Proc. Natl. Acad. Sci. USA.* 88:2312–2316.

63. Ulyanov, N. B., and T. L. James. 1995. Statistical analysis of DNA duplex structural features. *Methods Enzymol.* 261:90–120.

64. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.

65. Timoshenko, S. P., and J. M. Gere. 1961. Theory of Elastic Stability. McGraw-Hill, New York.

66. Muller, J., S. Oehler, and B. Muller-Hill. 1996. Repression of Lac-I promoter as a function of distance, phase and quality of an auxiliary Lac operator. *J. Mol. Biol.* 257:21–29.

67. Levandoski, M. M., O. V. Tsodikov, D. E. Frank, S. E. Melcher, R. M. Saecker, and M. T. Record. 1996. Cooperative and anticooperative effects in binding of the first and second plasmid O-sym operators to a LacI tetramer: evidence for contributions of non-operator DNA binding by wrapping and looping. *J. Mol. Biol.* 260:697–717.

68. Gemmen, G. J., R. Millin, and D. E. Smith. 2006. Dynamics of single DNA looping and cleavage by Sau3AI and effect of tension applied to the DNA. *Biophys. J.* 91:4154–4165.

69. Lia, G., D. Bensimon, V. Croquette, J. F. Allemand, D. Dunlap, D. E. A. Lewis, S. C. Adhya, and L. Finzi. 2003. Supercoiling and denaturation in Gal repressor/heat unstable nucleoid protein (HU)-mediated DNA looping. *Proc. Natl. Acad. Sci. USA.* 100:11373–11377.

70. Marko, J. F. 1997. Stretching must twist DNA. *Europhys. Lett.* 38:183–188.

71. Healey, T. J. 2002. Material symmetry and chirality in nonlinearly elastic rods. *Math. Mech. Solids.* 7:405–420.

72. Goyal, S., N. C. Perkins, and J. C. Meiners. 2007. Resolving the sequence-dependent stiffness of DNA using cyclization experiments and a computational rod model. *ASME J. Computational and Nonlinear Dynamics.* In press.

73. Dickson, R. C., J. Abelson, W. M. Barnes, and W. S. Reznikoff. 1975. Genetic-regulation—Lac control region. *Science.* 187:27–35.

74. Klenin, K., and J. Langowski. 2000. Computation of writhe in modeling of supercoiled DNA. *Biopolymers.* 54:307–317.