INTRINSIC CURVATURE OF DNA INFLUENCES LacR-MEDIATED LOOPING

RUNNING TITLE: INTRINSIC CURVATURE INFLUENCES LOOPING

AUTHORS:

Sachin Goyal\textsuperscript{1} (sgoyal@umich.edu)
Todd Lillian\textsuperscript{1} (tlillian@umich.edu)
Seth Blumberg\textsuperscript{2,3} (blumberg@caltech.edu)
Jens-Christian Meiners\textsuperscript{2,3} (meiners@umich.edu)
Edgar Meyhöfer\textsuperscript{1} (meyhofer@umich.edu)
N. C. Perkins\textsuperscript{1} (ncp@umich.edu) – Corresponding Author

AFFILIATIONS:

\textsuperscript{1}Department of Mechanical Engineering
\textsuperscript{2}Department of Physics
\textsuperscript{3}Biophysics Research Division

University of Michigan, Ann Arbor MI 48109
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 base-pair sequence of inter-operator DNA which influences its structural properties. One structural property is the ‘intrinsic’ or ‘stress-free’ curvature. In this paper, we explore this 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 our dynamic rod model that predicts the energetics and topology of the intervening DNA loop. Our simulations 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 non-uniform twist and necessitates consideration of eight distinct binding topologies from the known crystal structure of the LacR-DNA complex.

KEY WORDS
DNA Mechanics, DNA-Protein Interactions, Gene Regulation, Kirchhoff Rod, Lac Repressor

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 (1). In such instances, a protein or protein complex binds simultaneously to (at least) two non-contiguous 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 paper, 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
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. 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 *E. 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 trinucleotide model (15,16)) 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 (4.80Å resolution as reported by Lewis et al. (17)) to compute the position and orientation of the rod (boundary conditions) at the operator sites. The dynamic computational rod model (18) 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 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-electrophoretic and 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 (1,19) of looping behavior such as arising in GalR (20,21), Ara (22), SfI (23,24) and ntrC (25) in addition to other (non-looping) 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 (4,26-28) that are also reviewed in (29,30). 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 (31-36) 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 non-homogenous 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 (37-39). The computational model, however, provides the framework for incorporating both sequence-dependent linear elastic material laws as well as nonlinear (and inelastic) laws (9,10,40,41) should they someday become well-characterized. We also treat the LacR as rigid and thereby ignore effects of protein flexibility as suggested in (36,42,43). With the above assumptions duly noted, 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 (14) 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 non-uniform twist (or twist deficit) as recognized in (44) and also necessitates consideration of eight distinct binding topologies from the known crystal structure of the LacR complex.

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, twist-extension coupling in DNA, and sequence-dependent stiffness parameters would all 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.

METHODS

In the Kirchhoff rod model, ds-DNA is approximated as a flexible rod having elastic properties as determined from single molecule experiments (45-48), MD simulations (5) and other biophysical techniques. We begin by reviewing the salient features of the computational rod model (18) for use in this study. The interested reader is referred to (18) 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.

Non-homogeneous Rod Model for DNA

Figure 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 $\{a_i(s,t)\}$ (18)]. 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 base-pair 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 inter-atomic interactions conspire to yield the long-length scale material law which is often assumed to be linearly elastic (see, for example, \((6,7,16,29-36,49-51)\)). An exception is the nonlinear law proposed in \((9,10,41)\) for highly kinked strands, which has also been questioned in subsequent studies \((40)\). Here, we shall adopt a linear elastic law

\[
q(s,t) = B(\kappa(s,t) - \kappa_0(s)),
\]

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 \((37-39)\). The above law renders the rod model non-homogenous, that is, sequence-dependent by capturing the effects of intrinsic curvature/twist \( \kappa_0(s) \). The associated elastic strain energy density follows from

\[
S_e(s,t) = \frac{1}{2}(\kappa(s,t) - \kappa_0(s))^\top B(\kappa(s,t) - \kappa_0(s))
\]

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 \((18)\). 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 \((17)\) 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 \((18)\) can be numerically integrated to solve for the four vector unknowns \((f, q, v, \omega)\) when combined with Eq. 1:

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

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

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

\[
\frac{\partial \omega}{\partial s} + \kappa \times \omega = \frac{\partial \kappa}{\partial t}. \tag{6}
\]
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. 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 stability of the computed equilibrium \((18)\). Employing a dynamic formulation is advantageous because the solution dynamically relaxes to a stable (looped) equilibrium. The predictions of looped or supercoiled states directly from equilibrium rod theory (e.g. \((6,31,32,35,50,51)\) and citations in \((52)\)) require a subsequent analysis of loop stability. Furthermore, equilibrium theories cannot capture possible dynamic transitions between equilibrium states as highlighted in \((18)\). 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 \((41)\).

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 DNA-LacR complex. The inter-operator DNA modeled here as a rod includes three base-pairs from each operator site (see Appendix A) as also assumed in \((31-36)\).

In aligning the boundary base-pairs with the known crystal structure \((17)\), we 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 (17,53) appears to be asymmetric (by our calculations of data in (17,53)) 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 four binding topologies, two conventions have been proposed in the literature (20,34). Here, we elect to extend the original notation of Geanacopoulos et al. (20) to a three-digit binary notation to distinguish all eight binding topologies. According to (20), the first digit 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 digit 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 digit 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 3’ end of the coding strand) as illustrated in Fig. 3. The second digit is chosen to be 1 if the 5'-3' direction of the coding strand at L1 points towards the interior of the V-shaped protein, otherwise it is chosen to be 2. In other words, the second digit 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. (34) used ‘O’ and ‘I’ in their two digit binary notation resulting in ‘II’ = ‘A1’, ‘OO’ = ‘A2’, ‘IO’ = ‘P1’ and ‘OI’ = ‘P2’. For the third digit, we define BD1 as the protein head group bound to the strands labeled H and G in the LacR crystal structure given by PDB ID: 1LBG (17,53). The boundary domain BD2 is then the other head group. The third digit 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.

The above arguments determine the relative position and orientation of the two operators to within a single \( 2\pi \) 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 (54), corresponding to different topoisomers. Highly over-wound and under-wound topoisomers are expected to have high energetic cost in the LacR-DNA complex. Thus, as in prior predictions of looping for the LacR (31-36), 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 (55).

**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 (15,16) 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 (15). This web tool outputs a protein data-bank file giving the co-ordinates 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 base-pair origins (56) 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 $R_0(s)$ (57). The intrinsic curvature and twist of the helical axis are determined by the “principal curvature” $\kappa_p(s)$ and “geometric torsion” $\tau(s)$ of $R_0(s)$ (57). 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 re-construct 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 base-pair 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 base-pair 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 (14,58); specifically,

- the sequence-dependent intrinsic curvature can reduce the energetic cost of looping (14),
- the sequence-dependent intrinsic curvature influences loop topology and the distribution of topoisomers (14),
- looping energy depends strongly on operator orientation compared to operator separation (58).
To this end, we first re-examine the computed loops for the wild-type sequence (see Appendix A) and contrast our results with those in (35) 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 (14). 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**

Figure 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. The straight B-DNA (red) with 3.46Å height and 34.6° twist per base-pair step correspond to the homogeneous rod model used in (31-36). The consensus tri-nucleotide model (15) (blue/green) accounts for sequence-dependent shape modeled in our non-homogeneous 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 (15) which yields a distinct three-dimensional curve for the helical axis. Figure 4(b) illustrates the resulting principal curvature and geometric torsion computed from the consensus tri-nucleotide model (as functions of non-dimensional 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). Figures 5(a,b) illustrate the loops for both models – homogenous B-DNA (red) and non-homogenous 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 under-twisted (Fig. 5(a)) and the other over-twisted (Fig. 5(b)). The principal curvature \( |k \times a_3| \) and over-twist density \( \kappa_3 - \tau \) for each loop are reported in Figs. 5(c,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 (35). Indeed, the computations shown in Figs. 5(c,d) for homogeneous B-DNA (red) faithfully reproduce the principal curvature and over-twist density reported in (35) to within 0.2 deg/ bp. A summary of the total over-twist \( \Delta Tw \), writhe \( Wr \), link \( Lk \), 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 and the second lowest is 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 shape in the four highly curved sequences with phased A-tract bends introduced in (14). 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 control sequence is nearly straight while the other three have similar A-tract bends with helical phase differences of approximately 70°.

For each sequence, we used the computational rod model to compute the inter-operator loops formed by LacR binding and for all possible (eight) binding topologies; refer to Fig. 3. As in the wild-type case, multiple (mechanically) stable loops are possible for each binding topology. Figure 7 illustrates the loop that achieves the minimum elastic energy for each sequence. In the table below we report the number of bp for the inter-operator sequence (as defined in Appendix A), loop elastic energy, total over-twist, writhe and link for these 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 co-exist with the stable states in a thermal environment with a high likelihood of inter-conversion. In fact, if one were to account for the other components of the free energy, a state having the 2nd lowest elastic energy may well yield the global free energy minimum. Note also that while some loops have very comparable elastic energies, their binding topologies and geometrical properties (e.g. whether over-twisted or under-twisted) can be altogether different.

**Influence of Inter-operator Length and Phase**

The sequences of (14) considered above differ both in the location/phase of the A-tract bend as well as the number of base-pairs of the inter-operator DNA which range from 131 bp (7C16) to 144 bp (Control). The elastic energy of the resulting loops is certainly influenced by both factors. Therefore, the gel shift experiments on the looping of the four sequences in (14) were influenced by both effects, not just the helical phasing of A-tracts with respect to the operators. In the following results, we isolate these influences on elastic loop energy.

Adding (or subtracting) a single base-pair is expected to change the loop energy by changing (a) the length of inter-operator DNA (e.g., by 3.46 Å/base-pair), and (b) the relative orientation of the operators by one unit of base-pair twist (e.g., by 34.6°/base-pair). The first effect is negligible for the four designed sequences considered above since the relatively small differences in contour length (less than 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.

Figure 8 illustrates the computed loop elastic energy for the Control sequence (modeled as straight B-DNA) with the P1 binding topology. The solid curves represent the energy computed by simply rotating one operator about its tangent vector \( \hat{t} \) in increments of the nominal base-pair twist (34.6°) while holding the number of base-pairs constant (142). The two curves distinguish two computed loops; one under-twisted (blue) and one over-twisted (red). The circles represent the energy computed by adding base-pairs 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 circles distinguish two computed loops; one under-twisted (blue) and one over-twisted (red).

**DISCUSSION**

We open this discussion by describing the overall effects of sequence-dependent intrinsic curvature on the mechanics of looping. We then discuss how our computational results support three major conclusions drawn from experimental studies of the LacR-DNA complex (14,58) including: 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 note several limitations and extensions of our computational model.

**Overall Effects of Sequence-dependent Intrinsic Curvature**

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 under-twisted loop (W_r = -0.34) (tri-nucleotide) has significantly greater magnitude than that of the A2 under-twisted loop (W_r = -0.08) (homogeneous B-DNA) that ignores intrinsic curvature. For the designed sequences, we do expect to see large changes in writhe due their significant intrinsic curvature. For example, the writhe of the P1R loops for the sequence 9C14 (W_r=0.15) in Fig. 7(e) is less than one-half that of the control sequence (W_r = 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 under-twist remains uniform for the (homogeneous) model for the straight B-DNA, it becomes non-uniform for (non-homogeneous) models (44) that include intrinsic curvature, refer to Figs. 5(c,d). This general observation may open further questions about possible sequence-dependent localization of over- and under-twist 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 Fig. 7(e). 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 (14), “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 (14), “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 (11.93 kT versus 7.38 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 (that 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 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 over-twisted, while the minimum energy loops for one designed sequence 11C12 and the wild-type sequence are both under-twisted. These observations support findings from the gel shift assay experiments of (14) which 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 2nd 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 7C12 and 40% for 11C12 suggest substantially different topoisomer concentrations. Likewise a large (40%) energy difference separates the under-twisted (lower energy) from the over-twisted (higher energy) topoisomers of the wild-type sequence with P1 binding topology when the sequence-dependent intrinsic curvature is captured by 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. Even for the other binding topologies, consideration of sequence-dependent intrinsic curvature, as observed in Fig. 5(e), seems to substantially widen the gap between the energies of the two topoisomers in general.

**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 (59) 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, non-uniform 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 base-pair may alter the elastic energy through changing the length of inter-operator DNA by one unit of base-pair rise, and/or (b) the changing the relative orientation of the operators by one unit of base-pair twist. For the relatively long sequences of (14), 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 (58).

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 base-pair twist (solid curves), closely approximate the energy (circles) computed when also allowing the inter-operator DNA to increase by increments of the base-pair rise. This close agreement between these two calculations provides strong support for the claim that changes in operator orientation brought about by adding/subtracting base-pairs 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 (58) as discussed in (14),
“The \textit{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 (21,22,60). 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%.

\textbf{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). By contrast, contributions from the loop and protein deformation energies may vary significantly and, as a result of associated conformation changes, so might the electrostatic repulsion. For instance, the binding co-operativity of the two operator sites depends on their electrostatic repulsion (61) 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 (18).

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 (36) while the possibility of flexibility in the V-region has also been suggested in prior studies (42,43). Similarly, the entire Lac-R 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 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 (36).

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 (9,10,41) have begun to explore nonlinear (and inelastic) behavior, though this has also been questioned (40). It is also recognized that DNA must exhibit a strong coupling between twist and extension (62) due to its chiral (helical) construction (63) and this requires a modification of the material law used herein as proposed in (18). 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, (55). 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.

CONCLUSIONS

This paper 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 (14). 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. 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 (14) 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 non-uniform twist (or twist deficit) (44) 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 non-looping behaviors of DNA.
APPENDIX A: SEQUENCES AND BOUNDARY CONDITIONS

(a) Wild-type LacR sequence (64)

Operator O\textsubscript{1} at Location L\textsubscript{1}

\[
\begin{align*}
\text{GGCAGTGAGCGCAACGCAATT} & \quad \text{Wild-type DNA} & \quad \text{ATTTGTGACCCGATAACAAAT} \\
\end{align*}
\]

Fragment of inter-operator DNA modeled as rod (77 bp)

\[
\begin{align*}
\text{ATTAATGTGAGTTAGCTCACTCAT} & \quad \text{Wild-type DNA} & \quad \text{AATTGTGAGCGGATAACAATT} \\
\end{align*}
\]

(b) Designed sequences (14) (personal communication with Prof. J. D. Kahn, Department of Chemistry and Biochemistry, University of Maryland)

\[
\begin{align*}
\text{Operator O}\textsubscript{id} \\
\text{L} & \quad \text{Linker} & \quad \text{A tracts} & \quad \text{Linker} & \quad \text{TAGAATGAACCTAGCT} \\
\text{L} & \quad \text{Linker} & \quad \text{A tracts} & \quad \text{Linker} & \quad \text{ATTTGTGACCCGATAACAAAT} \\
\end{align*}
\]

In arriving at the final looped state, the boundary base-pairs are made to align with their corresponding configurations known from the LacR crystal structure given by PDB ID: 1LBG (17,53). This alignment is achieved in our calculations by slowly translating and rotating the boundary base-pairs (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 our boundary conditions account for the base-pair inclination with respect to the ds-DNA helical axis. We also verified that these boundary conditions are insensitive to the choice of boundary base-pairs 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 base-pair (3 base-pairs 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 base-pair.

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 (15)).

1. Following (56), we first compute the origin of each base-pair as the mid-point of the C6 atom of the pyrimidine and the C8 atom of the purine (see Fig. 10(a)). A curve interpolated through the base-pair origins forms an approximate helix of radius $r \approx 2.0 \, \text{Å}$ and helical pitch $\approx 10.3 \, \text{base-pair}$ (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 $\tilde{R}_0(s)$ follows from averaging the positions of the origins of the base-pairs. We begin at one operator and then average the positions of the origins of the first ten base-pairs for the inter-operator DNA (see Fig 10(b)). We then increment by one base-pair and repeat this (moving average) computation and continue to the other operator thereby developing a point-wise approximation to the helical axis $\tilde{R}_0(s)$.

3. A continuous (differentiable at least three times) curve $R_0(s)$ is sought to approximate $\tilde{R}_0(s)$ in order to compute the intrinsic curvature and torsion. We use the MATLAB curve-fitting toolbox (The MathWorks, Natick, MA) to construct a $C^\infty$ continuous curve $R_0(s)$.

We emphasize that our 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 less than 2%.

ACKNOWLEDGEMENTS

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 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. Goyal, T. Lillian, N. C. Perkins and E. Meyhöfer gratefully acknowledge support for this research from the US National Science Foundation under grants CMS-0439574 and CMS-0510266. The co-authors J. C. Meiners and S. Blumberg gratefully acknowledge support from the National Institutes of Health under grant GM65934.
REFERENCES


FIGURE CAPTIONS

**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 Consensus Tri-nucleotide model (15,16) and compute intrinsic shape for rod model (twist and curvature of helical axis and inclination of the base-pair planes with respect to the helical axis). (c) Employ known crystal structure of the LacR protein bound to the operators (17) 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 (18) to compute inter-operator loop.

**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 $R(s,t)$.

**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-digit binary notation is used to distinguish all eight possible binding topologies and all “forward” (F) binding topologies are illustrated here.

**Figure 4:** (a) Comparison of two different models of stress-free, zero-temperature, wild-type, inter-operator DNA: Red – straight B-DNA and Blue/ Green – consensus tri-
nucleotide model (15). The left boundary base-pair for the two models are aligned. (b) Principal curvature and geometric torsion of the helical axis for the consensus tri-nucleotide model (15) as a function of (non-dimensional) contour length $s$.

Figure 5: (a) & (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 under-twisted (a) while the other is over-twisted (b). (c) & (d) Principal curvature and over-twist density of all loops above shown as functions of (non-dimensional) contour length coordinate $s$. The principal curvature for the (stress-free) consensus model (black) is reproduced for comparison. (e) Table summarizes the total over-twist (above the natural helical twist) $\Delta Tw$, writhe $W_r$, linking number $L_k$, and loop elastic energy $E$ for all the binding topologies. The writhe $W_r$ is computed using “Method 1a” described by Klenin and Langowski (65). We form a closed loop for calculating writhe by adding a straight segment $\tilde{r}_{i-2}$ 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 $34.6^\circ$/bp, zero principal curvature, and rise of 3.46 Å/bp. The bending and torsional persistence lengths are assumed to be 50nm and 75nm (37-39) 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.

25
Figure 6: Two views of the stress-free, zero-temperature conformations of four designed inter-operator DNA sequences (14) as computed using the consensus tri-nucleotide model (15). The first base-pair of each sequence is assigned the same position and orientation. The operator regions are shown in green, the red and blue segments are 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 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 approximately 70° about the helical axis of the control as best observed in view (b).

Figure 7: (a)-(d) Lowest energy solutions for four designed sequences and the associated binding topology. (e) Table summarizing binding topology, loop elastic energy, over-twist, writhe and link for loops with the minimum and second smallest elastic energies. The largest of all the minimum energies is denoted in red font and the lowest in blue font.

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. Solid curves illustrate the periodic variation in elastic energy obtained by rotating one operator about the helical axis in increments of the base-pair twist (in this model 34.6°/base-pair) while keeping the inter-operator length constant (142 bp); refer to scale on top for relative
angular orientation of operators. The circles illustrate the same variation obtained by adding base-pairs and thereby both rotating one operator as well as increasing the inter-operator length (in this model 3.46 Å/base-pair); refer to scale on bottom for bp number. Over-twisted solutions denoted by red, under-twisted solutions by blue.

**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.

**Figure 10:** (a) The origin and standard base-pair fixed reference frame described in Olson et. al. (56). (This figure is a modification of Fig. 1 from (56)). (b) Base-pairs represented as transparent red blocks with the minor-groove face shaded black. Black dots represent the base-pair origins, and the blue line represents the helical axis as computed by the moving average over a helical turn.
(a) **Input 1:** Sequence of Substrate DNA

Operator “O₁” at location L₁ - - Inter-Operator sequence - - Operator “O₂” at location L₂

5’ ... GGTAAATTGAGC-GCTCAACATAAGA ... ... ... ... GCTAATTGAGC-GCTCAAGTTCGT ... 3’

3’ ... ccattaacactcg-cgagtgttaatct ... ... ... ... cgattaacactcg-cgagtgttaagca ... 5’

(b) **Compute** Stress-Free Shape Based on Consensus Tri-nucleotide Model

Input 3: Material Law
(e.g., Bending and Torsional Persistence Lengths)

(c) **Input 2:** DNA-Operator Crystal Structure

(d) **Output:** Topology, Energetics and Dynamics of Loop Formation

**Figure 1**
Figure 2
Figure 4

Principal Curvature and Geometric Torsion

(a) consensus tri-nucleotide model

(b) straight B-DNA

Figure 4
### Figure 5

#### Table

<table>
<thead>
<tr>
<th>Binding Topology</th>
<th>Under-twisted Loop</th>
<th>Over-twisted Loop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E(kT)</td>
<td>$\Delta Tw$</td>
</tr>
<tr>
<td>Homogeneous B-DNA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>34.06</td>
<td>-0.32</td>
</tr>
<tr>
<td>A1</td>
<td>26.25</td>
<td>-0.27</td>
</tr>
<tr>
<td>A2</td>
<td>30.26</td>
<td>-0.42</td>
</tr>
<tr>
<td>P2</td>
<td>75.77</td>
<td>-0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-homogeneous consensus tri-nucleotide model (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1F</td>
<td>35.23</td>
<td>-0.32</td>
</tr>
<tr>
<td>A1F</td>
<td>25.94</td>
<td>-0.16</td>
</tr>
<tr>
<td>A2F</td>
<td>27.14</td>
<td>-0.29</td>
</tr>
<tr>
<td>P2F</td>
<td>81.42</td>
<td>-0.30</td>
</tr>
<tr>
<td>P1R</td>
<td>34.48</td>
<td>-0.29</td>
</tr>
<tr>
<td>A2R</td>
<td>25.41</td>
<td>-0.13</td>
</tr>
<tr>
<td>A1R</td>
<td>28.27</td>
<td>-0.30</td>
</tr>
<tr>
<td>P2R</td>
<td>75.9</td>
<td>-0.33</td>
</tr>
</tbody>
</table>
Figure 6
(a) Control (Binding Topology: A2F)

(b) 11C12 (Binding Topology: A2R)

(c) 7C16 (Binding Topology: A2F)

(d) 9C14 (Binding Topology: P1F)

<table>
<thead>
<tr>
<th>Binding Topology</th>
<th>Control (144 bp)</th>
<th>11C12 (141 bp)</th>
<th>7C16 (131 bp)</th>
<th>9C14 (142 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min</td>
<td>2nd</td>
<td>min</td>
<td>2nd</td>
</tr>
<tr>
<td>E (kT)</td>
<td>11.93</td>
<td>12.55</td>
<td>7.38</td>
<td>10.42</td>
</tr>
<tr>
<td>ΔTw</td>
<td>0.03</td>
<td>-0.07</td>
<td>-0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Wr</td>
<td>-0.01</td>
<td>0.15</td>
<td>-0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Lk</td>
<td>0.01</td>
<td>0.08</td>
<td>-0.37</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Figure 7
Figure 8
Figure 9
Figure 10

Base-pair Origin

Helical Axis

Minor Groove

Major Groove