Computational Analysis of Looping of a Large Family of Highly Bent DNA by LacI

Todd D. Lillian,* Sachin Goyal,† Jason D. Kahn,‡ Edgar Meyhöfer,* and N. C. Perkins*
*Mechanical Engineering, University of Michigan, Ann Arbor, Michigan; †Theoretical & Applied Mechanics, Cornell University, Ithaca, New York; and ‡Chemistry and Biochemistry, University of Maryland, College Park, Maryland

ABSTRACT Sequence-dependent intrinsic curvature of DNA influences looping by regulatory proteins such as LacI and NtrC. Curvature can enhance stability and control shape, as observed in LacI loops formed with three designed sequences with operators bracketing an A-tract bend. We explore geometric, topological, and energetic effects of curvature with an analysis of a family of highly bent sequences, using the elastic rod model from previous work. A unifying straight-helical-straight representation uses two phasing parameters to describe sequences composed of two straight segments that flank a common helically supercoiled segment. We exercise the rod model over this two-dimensional space of phasing parameters to evaluate looping behaviors. This design space is found to comprise two subspaces that prefer parallel versus anti-parallel binding topologies. The energetic cost of looping varies from 4 to 12 *kT*. Molecules can be designed to yield distinct binding topologies as well as hyperstable or hypostable loops and potentially loops that can switch conformations. Loop switching could be a mechanism for control of gene expression. Model predictions for linking numbers and sizes of LacI-DNA loops can be tested using multiple experimental approaches, which coupled with theory could address whether proteins or DNA provide the observed flexibility of protein-DNA loops.

INTRODUCTION

Proteins often induce DNA bending and torsion to influence or control cellular processes. In protein-mediated DNA looping, proteins bind simultaneously to separated sites on a DNA molecule to form a loop from the intervening DNA. The loop can both respond to changes in bending and torsion or induce them to affect downstream functions, such as regulating transcription or facilitating protein-DNA assembly during replication, recombination, and condensation (1,2). For example, the activator protein NtrC facilitates transcription by interacting with the RNA polymerase holoenzyme via looping, and the loop can respond to DNA bending. Repressor proteins such as the Lac and Gal repressors (LacI and GalR) of the bacterium *Escherichia coli* repress transcription upon introducing a DNA loop, and GalR may act by establishing a promoter conformation inconsistent with initiation. For these and other systems, the intrinsic curvature of DNA may facilitate looping (3,4) by essentially prebending the helical axis of otherwise unstressed DNA. Similarly, proteins such as CAP and IHF induce large DNA bends (~90° for CAP, 180° for IHF) to facilitate looping (see, for example (5–7)). Herein we focus on looping of intrinsically curved DNA by LacI, but the method and conclusions should be relevant to other DNA-protein complexes forming similarly short DNA loops.

Substantial understanding of protein looping interactions with DNA has been obtained from biochemical studies (8,9) and the x-ray crystal structures of LacI and LacI bound to oligonucleotide operators are available (6,10). Despite these significant advances, many questions remain concerning fundamental aspects of looping mechanisms, in particular how looping is influenced by the sequence of the interoperator DNA. Experimental and theoretical methods have been employed to probe the thermodynamics and topology of DNA loops as functions of DNA sequence. Experiments are challenged by the absence of complete high-resolution data for these large, flexible, and dynamic structures, while theories are challenged in characterizing the many physical interactions that govern looping. This article implements a theoretical model to explore looping by the Lac repressor (LacI) protein for a large family of bent interoperator sequences related to an existing model system. The model makes a direct connection between basepair level DNA structure/flexibility parameters and the results to be expected from looping experiments, and the analysis is readily extendable to additional interactions such as protein flexibility.

The theoretical framework employs an elastic rod model for DNA that explicitly incorporates the intrinsic curvature of the helical axis and its sequence-dependent anisotropic flexibility, as determined by the basepair sequence (11,12). The rod approximation is a coarse-grained model that necessarily averages the elasticity properties of DNA over at least a helical turn, so it is most useful for describing the long-length scale mechanics of the molecule (13,14). The computational efficiency of rod models has been recognized in previous studies of DNA looping (11,15–17). Finer-grained discrete basepair models in which the basepairs are approximated as rigid bodies connected to their nearest neighbors by translational and rotational springs have also been applied to looping (18–21). Direct simulation of DNA looping at the

Submitted July 16, 2008, and accepted for publication September 5, 2008.

Address reprint requests to N. C. Perkins, Tel.: 734-936-0403; E-mail: ncp@umich.edu.

Editor: Ruth Nussinov.

© 2008 by the Biophysical Society
0006-3495/08/12/5832/11 $2.00

doi: 10.1529/biophysj.108.142471

atomistic scale by molecular dynamics remains computationally challenging due to the systems' size and the disparate time- and length-scales needed to resolve DNA and protein dynamics (13). Nevertheless, MD simulations of the related problem of minicircle conformation have been achieved (22). A multiscale method which couples a rod formulation for the DNA with an MD formulation of LacI has been presented (23,24). Potentially limiting the rod model is its exclusion of entropic effects, which are otherwise included in statistical approaches (see, for example, (21,25,26)). However, an extension of the Shimada-Yamakawa theory for DNA cyclization (27) leads to a computational method for evaluating the entropic contribution to the free energy of looping (28).

Recent models of the DNA-LacI complex use the crystal structure of LacI bound to DNA operators (6) to determine positional and orientational boundary conditions for the structure of the interoperator DNA. However, the protein structure in solution may differ from the crystal structure due to protein flexibility or alternative stable conformations (3,10,29,30). Therefore, a challenge remains in defining boundary conditions with imprecise knowledge of the protein structure for the looped complex. Recent studies have explored the role of protein flexibility. For example, Zhang et al. (21,31) employ a simplified protein model with varying amounts of stiffness while Swigon et al. (16) considers the limiting case of vanishing stiffness for prescribed modes of protein deformation. Recently, we extended our elastic rod model to represent the protein as an elastic body (an extended rod) having nonuniform elasticity that can account for both stiff and flexible protein domains (32). Advantages of this approach include the ability to approximate distributed stiffness and anisotropic bending. The multiscale model of Villa et al. (23,24) explicitly accounts for protein flexibility over short timescales.

Experimentally, one of the best-characterized features of LacI looping is the oscillatory dependence of looping free energy (as measured by repression level) on the length of the interoperator DNA (33–35). As seen for cyclization, the period of the oscillatory component is about a helical turn, arising from the change in torsional alignment or phasing of the operator sites upon the addition or subtraction of base-pairs. However, as cautioned in the literature (12,21), the distinct closure requirements for looping versus cyclization yield significant quantitative differences in the thermodynamics, particularly for the subpersistence length of the interoperator DNA of the DNA-LacI complex. DNA supercoiling, protein flexibility, DNA intrinsic curvature, sequence-dependent stiffness, changes in loop topology, and the presence of IPTG can also affect operator-phasing patterns (16,21,31,35–37).

The Kahn lab (3,38,39) has probed the combined effects of operator phasing and intrinsic curvature on protein-mediated looping using designed sequences that contain A-tract bends in the interoperator DNA (3). Three such bent sequences, referred to as 11C12, 9C14, and 7C16, form hyperstable loops proposed to have markedly different conformations, although they are not necessarily optimal representatives of any particular conformation (3). In a recent study (12), we demonstrate that upon explicitly incorporating sequence-dependent intrinsic curvature in the DNA constitutive law, the computational rod model correctly predicts a broad range of the experimental results for the Kahn constructs. In particular, the theory successfully predicts the operator orientations (loop topologies) known from FRET measurements, the linking number distribution known from cyclization assays of the LacI-DNA complex, the relative loop stabilities known from competition assays, and the relative loop size inferred from gel mobility assays.

The bent sequences above are but three examples from a large family of molecules having the embedded A-tract. Molecules within this family are distinguished by the phasing of the A-tract relative to the terminal operators which bind to LacI. The energetics and topology of the looped DNA-LacI complexes depend critically upon this phasing, and there is no reason to believe that any of the three sequences synthesized to date have yielded optimal characteristics such as stability or conformational uniformity. In this study, we materially extend the work of Goyal et al. (12) by establishing a two-parameter representation for a complete description of the operator phasing of an entire family of related molecules. We exercise the computational rod model over the two-dimensional design space of sequences to evaluate the landscape of possible looping characteristics including loop energy/stability. Doing so illustrates that combining the operator phasing length experiment with the unique variety of experiments possible on hyperstable loops may allow us to design loops with intriguing conformational switching properties as well as to test the model's assumptions about protein and DNA flexibility.

METHODS

We employ a computational rod model to explore the effects of intrinsic curvature on looping for an entire family of bent sequences that incorporates as special cases the three examples synthesized in the Kahn lab (3,38,39). A derivation of this computational rod model is provided in Goyal et al. (40) and its application to looping of the DNA-LacI complex is summarized in Goyal et al. (12). Herein, we describe how this formulation is modified to analyze a family of sequences.

As reviewed above, the continuum rod is fundamentally a course-grain model of DNA (13), which describes the three-dimensional bending and twisting of the helical axis of the molecule on length scales of approximately a helical turn (3 nm) and longer. As in the literature (11,15), we use averaged stiffness properties of DNA as determined by commonly accepted values of the bending and torsional persistence lengths, 50 nm and 75 nm, respectively (13,41). We explicitly incorporate the sequence-dependent intrinsic curvature of the helical axis in the rod formulation (12) with the rod elastic energy functional

$$E(t) = \int_0^L \frac{1}{2}[\{\vec{\kappa}(s,t) - \vec{\kappa}_o(s)\}^T \mathbf{B}\{\vec{\kappa}(s,t) - \vec{\kappa}_o(s)\}]ds. \quad (1)$$

Here, $E(t)$ is the elastic deformation energy (in units of kT), t is the independent time variable, s is the independent spatial (contour length) variable (nm), $\mathbf{B}$ is a diagonal tensor which defines the bending and torsional

stiffness of DNA, assumed sequence-independent (nm – kT), and $\vec{\kappa}(s,t)$ is the curvature/twist vector of the helical axis (nm^{-1}). The vector $\vec{\kappa}_o(s)$ defines the curvature/twist (nm^{-1}) of the stress-free conformation, which depends upon s and thus accounts for the sequence-dependent intrinsic curvature of the molecule, as detailed further below.

By allowing off-diagonal elements and including explicit dependence on s in **B**, the rod formulation has been extended in Goyal et al. (40) to account for sequence-dependent stiffness properties, anisotropy, and tension-torsion coupling (42,43). In Goyal et al. (40), we suggest a further extension to approximate electrostatics and thermal kinetics by incorporating distributed forces and moments along the length of the rod. Since accounting for each additional effect requires specifying several frequently uncertain parameters, currently we ignore these additional influences (12). Despite these approximations, the resulting theory incorporating intrinsic curvature (12) predicts the major experimental results for the three highly bent sequences synthesized to date (3,38,39).

A comprehensive study of looping for an entire family of bent DNA sequences requires significant computational effort because numerous combinations of parameters are required, including those defining the following features: 1), the geometry of the intrinsic bend; 2), the location (phasing) of the intrinsic bend relative to the two operator sites; and 3), the manner in which LacI binds to the operators. We discuss next a number of simplifications that yield an efficient analysis.

Fig. 1 *A* introduces a simplified straight-helical-straight (SHS) representation for describing intrinsic curvature with the rod model (12). Bent sequences related to the three synthesized sequences introduced in Mehta and Kahn (3) are modeled as a helically supercoiled bent A-tract domain (*dark shaded*) flanked by two straight linker domains (*light shaded*). Fig. 1 *A* also illustrates the SHS rod representation superimposed upon an atomistic representation of the stress-free and zero temperature conformation of one such sequence, 11C12 from Mehta and Kahn (3). The close agreement between the two representations follows from the fact that repeating A-tracts largely bend the helical axis of DNA into a helical supercoil (see, for example, (44,45)). The pitch and the radius of the superhelix depend on the details of the dependence of curvature on sequence. In Data S1 in the Supplementary Material, we show that the SHS representation approximates the helical axis of all three bent sequences of Mehta and Kahn (3) to within an RMS error of less than the radius of DNA (<10 Å).

The SHS representation is symmetric in that the right and left halves are identical to within a 180° rotation about the out-of-plane axis at the midpoint. This symmetry is not a fundamental limitation of the rod model, which is capable of accounting for arbitrary intrinsic curvature (12), but here we use the symmetry to reduce the computational cost of evaluating an entire family of bent sequences. The mathematical definition of the SHS representation, as provided in Data S1, ultimately defines the intrinsic curvature/twist $\vec{\kappa}_o(s)$ employed in the above elastic energy functional for the computational rod model. The SHS representation of the initial DNA conformation provides the initial conditions for the differential equations describing the dynamics of the rod. The subsequent computation of the conformation of the looped DNA-LacI complex follows the methods described in Goyal et al. (12).

Individual molecules within the family of bent sequences are distinguished by the phasing of their operators relative to the A-tract bend, which determines the orientation of the bound protein relative to the bend center. (For straight DNA, only a single phasing parameter is required, describing simply the torsional alignment of the terminal operator sites.) The orientation of the helix is defined by a vector triad, the three principal directions of the space curve corresponding to the tangential, normal and bi-normal vectors. Fig. 2 defines the two phasing parameters Θ_1 and Θ_2 by illustrating the torsional alignment of two triads. The shaded triad represents the orientation of the helix at the end of the A-tract domain, and the solid triad the orientation of the first basepair of the operator within the linker domain. The angle Θ_1 between the triads defines the torsional alignment or phasing of the A-tract to this operator. An analogous definition holds for Θ_2 at the opposite end.

For intrinsically straight DNA, the torsional phasing of the operators has a much larger effect on the elastic energy of looping than the modest effect of small changes (<3 nm) in contour length (12). Accordingly, we simplify our calculations by exploring up to complete 360° changes of the phasing parameters (Θ_1 and Θ_2) while keeping the overall contour length of the SHS representation constant. Experimentally, the phasing parameters are changed by adding or subtracting basepairs in the linker domains (3), where adding a

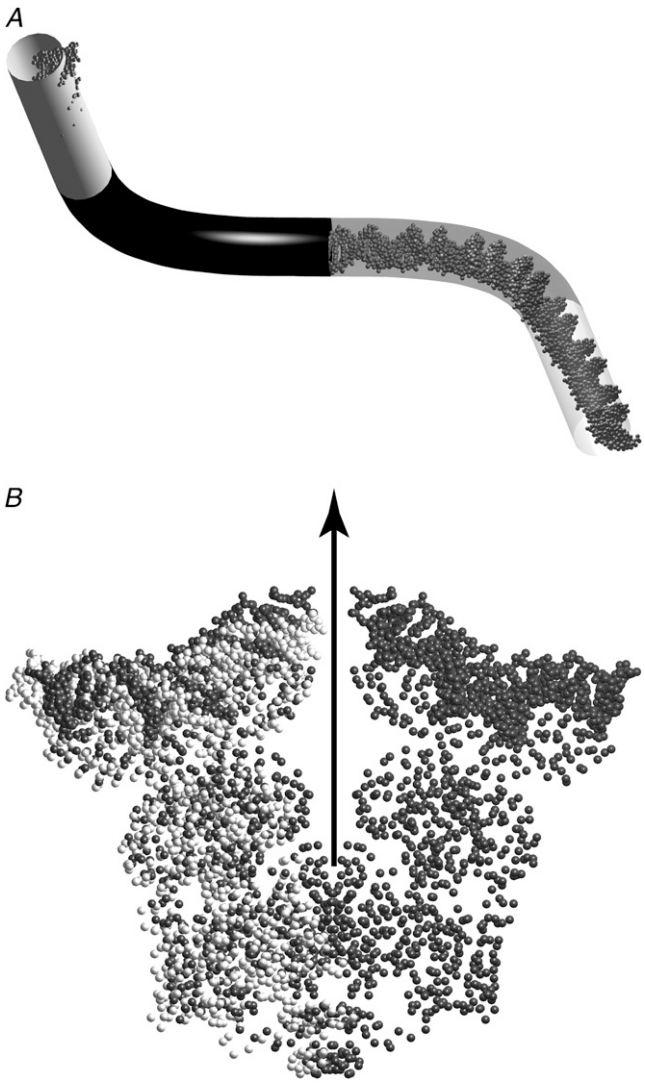

FIGURE 1 (*A*) The straight-helical-straight (SHS) approximation for highly bent DNA sequences. The light- and dark-shaded tube illustrates the SHS approximation to the stress-free, zero temperature conformation of the bent sequence 11C12 introduced in Mehta and Kahn (3). The light shaded portions of the tube are the straight ''linker'' domains and the dark-shaded portion is the superhelically curved A-tract domain. The transparent half of the tube reveals an atomistic representation of the 11C12 sequence provided by the web-based tool ''model.it'' (46). The SHS representation is symmetric: the left and right halves are related by a 180° rotation about the out-of-plane axis at the midpoint. Data S1 provides the mathematical definition of the SHS representation. The SHS representation approximates the helical axes of the bent sequences (3) to within an RMS error of less than the radius of DNA (<10 Å). (*B*) The x-ray co-crystal structure of the LacI-DNA complex (6). In this figure, the right half (*dark shaded*) of the crystal structure is rotated about the solid axis by 180° and then superimposed on the left half (*light shaded*). The differences between the light- and dark-shaded C-α atoms on the left side illustrate the very slight asymmetry in the crystal structure.

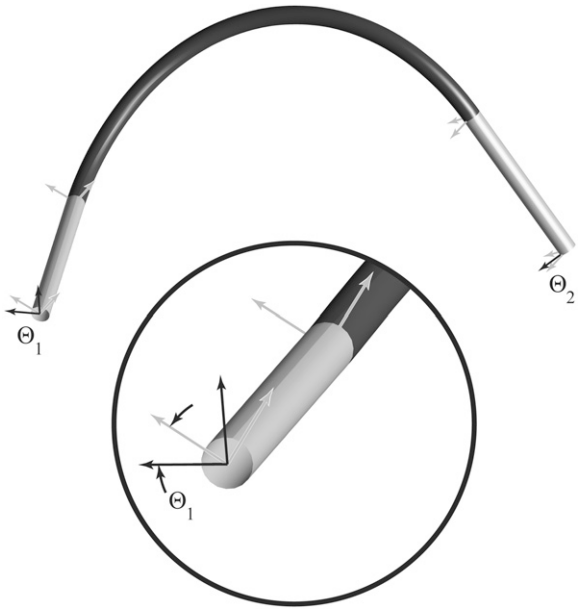

FIGURE 2 SHS representation showing the definition of the phasing parameters Θ_1 and Θ_2 in terms of triads aligned with the principal directions (tangential, normal, and bi-normal unit vectors) of the DNA. The shaded triad is aligned with the end of the helical (A-tract) segment, and the solid one is at the first basepair of the operator at the end of the straight (linker) segment. The enlarged view shows that Θ_1 is the angle formed between corresponding vectors for the shaded and solid triads. The Θ_1 and Θ_2 phasing parameters vary between 0 and 1 helical turns upon adding or subtracting basepairs to the linker domain, as in the three molecules in Mehta and Kahn (3). Θ_1 and Θ_2 define the torsional alignment of the operators relative to each other and to the central superhelical domain. The first and second numbers in the designations of the three bent sequences (11C12, 9C14, and 7C16) define their linker lengths and correspond to the phasing parameters Θ_1 and Θ_2, respectively.

single basepair changes the associated phasing parameter by ~34°. The first and second numbers in the designations of the three bent sequences (11C12, 9C14, and 7C16) define their linker lengths and correspond to the phasing parameters Θ_1 and Θ_2, respectively. We report Θ_1 and Θ_2 in units of helical turns where one helical turn represents a 360° change in the torsional alignment of the associated operator relative to the A-tract domain. Thus, the family of bent sequences with a common A-tract domain reduces to a two-parameter family of curves distinguished by Θ_1 and Θ_2. Allowing these phasing parameters to vary through their entire range (0–1) yields a large family of sequences, the majority of which have not been synthesized, but which includes the three synthesized sequences of Mehta and Kahn (3).

We estimate the positions of the designed sequences in the Θ_1-Θ_2 design space by mapping the rod representations of the stress-free conformations (12) onto the SHS representation introduced in this article. To this end, we first superimpose the helical axes of the SHS representation with that obtained in Goyal et al. (12) based on a consensus tri-nucleotide model for DNA (46). Next, we project the triads defining the operator phasing from the rod representations onto the SHS representation as in Fig. 2. This projection determines the values of the phasing parameters (Θ_1 and Θ_2) for the SHS representation that best approximate the corresponding rod representation. Note that since the rod representation for the helical axis (12) originates from an assumed DNA model, the consensus tri-nucleotide model (46), the approximate values of Θ_1-Θ_2 are estimates that will vary whether an alternative DNA model is used or the DNA helical repeat changes significantly.

The protein crystal structure dictates the position and orientation of the operators and thus the boundary conditions for the rod model. For non-symmetric DNA and protein, there exist eight possible binding topologies, where topology refers here to the orientation of each operator relative to the protein and the direction of the connecting DNA (12,47). Fig. 3 illustrates four topologies and reviews our notation for their classification (12). In addition, for each binding topology, one must further consider at least two twist isomers (one over- and one undertwisted). Hence, in the absence of symmetry, each sequence in the two-parameter family of sequences requires simultaneous consideration of eight possible binding topologies with both over- and undertwisted topoisomers. However, by assuming symmetry of the protein or DNA about their respective dyadic axes, the eight possible binding topologies reduce to four unique topologies. (For example, a symmetric protein results in four pairs of indistinguishable binding topologies: P1F and P1R; P2F and P2R; A1F and A1R; and A2F and A2R.) The four binding topologies further reduce to three upon assuming both symmetric protein and DNA. (Specifically, we have three sets of indistinguishable binding topologies: P1F and P1R; P2F and P2R; and A1F, A1R, A2F, and A2R.) Hence, analysis of a single molecule within this family requires only six computations (three binding topologies × two topoisomers) relative to the 16 computations (eight binding topologies × two topoisomers) required otherwise.

Because the Lac repressor is a homotetramer, we expect its structure to be symmetric. Fig. 1 *B* illustrates the near-symmetry of the protein crystal structure: a 180° rotation of the right half (*dark shaded*), about the dyadic axis (*solid arrow*), superimposes nearly perfectly on the left half (*light shaded*). Quantitatively, the boundary conditions for the rod model resulting from this assumption of protein symmetry deviate by <10 Å (in position) and 10° (in orientation) from those computed exactly from the crystal structure (6). Looping energetics were found to be relatively insensitive to a similar simplifying assumption used in Zhang et al. (31), in which the three-dimensional protein structure considered herein was approximated by a two-dimensional structure.

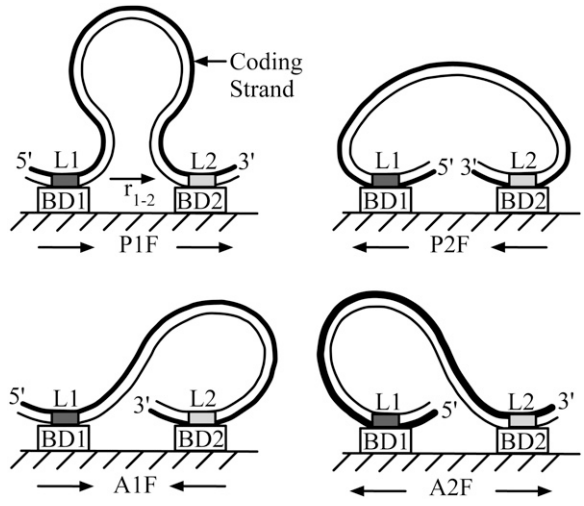

FIGURE 3 Classification of binding topologies for the LacI-DNA complex (12,47). There are eight possible ways for the two protein binding domains (BD1 and BD2) to bind to the two DNA operator locations (L1 and L2), four of which are illustrated here. These eight possibilities are distinguished using a three-character code (12) which slightly extends the two-character code of Geanacopoulos et al. (47). The first character (P or A) indicates if the operators are parallel (P) or anti-parallel (A) with respect to each other along the 5′ to 3′ direction. The second character (1 or 2) indicates the orientation (5′ to 3′) of the operator at L1; if it points toward the inside of the protein, it is assigned the number 1; if it points toward the outside, it is assigned the number 2. The third character (F or R) indicates if L1 binds to BD1 (*F* for forward) or to BD2 (*R* for reverse). All four F-topologies are illustrated here. With the assumption of a symmetric LacI, the four R-topologies are indistinguishable from the F-topologies.

RESULTS

The looping of three highly bent DNA sequences by the Lac repressor (LacI) protein has been characterized using a wide array of experimental techniques (3,38,39). In a recent theoretical study (12), we demonstrate that the computational rod model correctly predicts the major experimental findings for these three specific sequences. We now significantly extend these predictions to the entire two-parameter family of related bent sequences, as described by the SHS approximation defined above and in Data S1. While many of these sequences have not yet been synthesized, the theoretical results below reveal intriguing possibilities for future experiments.

Computational (e.g., (12,16,21,31)) and experimental (e.g., (33–35)) studies of phasing effects for straight DNA often report the free energy (or repression level) as a function of a single independent phasing parameter (often the contour length in basepairs). Such results must now be extended to demonstrate the simultaneous dependence on two independent phasing parameters for the highly bent sequences. The two phasing parameters Θ_1 and Θ_2 (see Fig. 2), given in units of helical turns, distinguish bent sequences within this family.

Fig. 4 illustrates the computed elastic energy cost of looping for all bent sequences over the possible sequence design space: $0 \leq \Theta_1 \leq 1$ and $0 \leq \Theta_2 \leq 1$. The result is an energy contour map where the contours map the loci of designs having equivalent elastic energy (kT) cost for loop formation. For each specified value of Θ_1 and Θ_2, we report the elastic energy of the minimum energy looped conformation, from among all three possible binding topologies and considering both over- and undertwisted topoisomers. (In Fig. 4, the computed elastic energy is symmetric about the diagonal $\Theta_1 = \Theta_2$. This symmetry is a direct result of the prescribed symmetries of the protein and bent DNA.) Also illustrated in Fig. 4 are three specific computed loops: the minimum energy loop that forms with the A (anti-parallel) binding topology sketched in Fig. 3 (4.0 kT), the minimum energy loop that forms with the P1 (parallel 1) binding topology (4.2 kT), and the overall maximum energy loop, which also forms with the A-binding topology (12.3 kT). These energetic differences are experimentally testable as in Mehta and Kahn (3).

The looped complexes formed by the original three bent sequences exhibited significant differences in gel migration behavior (3). Although we cannot predict quantitative electrophoretic mobilities for these complexes, we expect the radius of gyration (R_g) to correlate to the migration speed, as suggested previously for the three synthesized sequences (12). In Fig. 5, we report the radius of gyration of the minimum-energy DNA-LacI complex for each value of Θ_1 and Θ_2. (The R_g is estimated as the root mean-square distance of each atom in the DNA-LacI complex from the average position of all the atoms. We lump the atoms for each basepair into the associated spatial grid point in our discretized rod model, and we lump the atoms of each amino acid into a

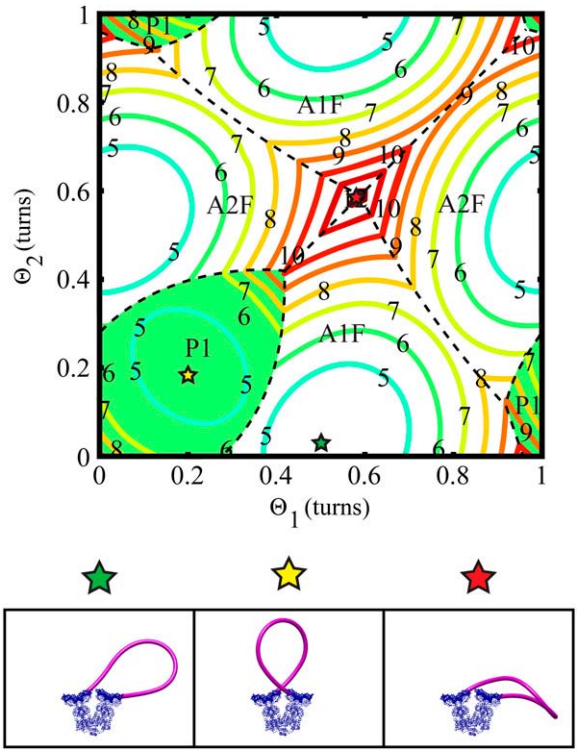

FIGURE 4 The energetic cost of looping over the entire Θ_1 and Θ_2 sequence design space of highly bent sequences with a common bent A-tract domain. The elastic energy (kT) of looping for the minimum energy conformations is reported, considering all binding topologies for both over- and undertwisted topoisomers. Distinct (minimum-energy) binding topologies are delineated with dashed dark lines. The green regions correspond to bent sequences that preferentially bind with parallel (P1) binding topologies. All other sequences, the open region, preferentially bind with anti-parallel (A) binding topologies. Three computed loops are shown: the minimum energy loop with the A-binding topology (4.0 kT), the minimum energy loop with the P1-binding topology (4.2 kT), and the overall maximum energy loop, which adopts the A-binding topology (12.3 kT). (The assumed DNA and protein symmetries make the A1- and A2-binding topologies indistinguishable.)

single point located at the associated C-α atom in the crystal structure because only the C-α atoms are resolved in the crystal structure (6).) In Fig. 5, we again demarcate the regions of the energetically preferred binding topologies as identified in Fig. 4. These borders between energetically preferred binding topologies often result in discontinuities in R_g. In addition to discontinuities arising from a change in preferred binding topology, discontinuities may also arise from a change in preferred topoisomer (i.e., over- or undertwisted). Fig. 5 illustrates two sample looped complexes, one on each side of a discontinuity. ($\Delta Tw = -0.13$ and $\Delta Tw = 0.07$ for the loops designated by the *open* and *solid stars*, respectively.)

Mehta and Kahn (3) carried out ligation reactions that yielded minicircles formed by cyclizing the free ends of ~350 bp DNA molecules after formation of embedded ~150 bp DNA-LacI loops. They measured ΔLk, the difference between the Lk of DNA cyclized while bound to LacI and the

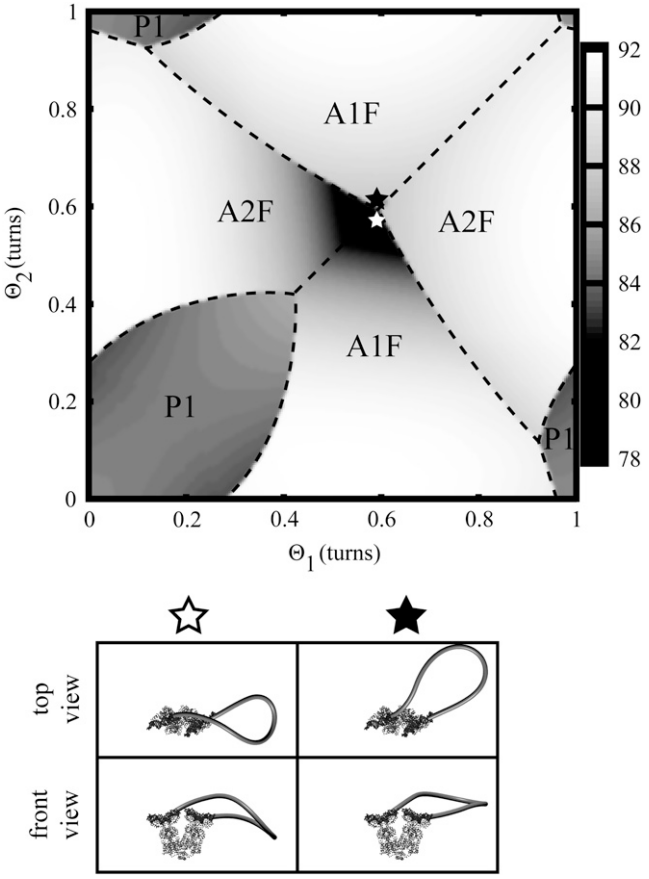

FIGURE 5 Computed radius of gyration (Å) of the minimum energy looped complexes reported in Fig. 4. Preferred binding topologies are demarcated with dashed lines. Some of these borders between binding topologies result in discontinuities in the R_g. Discontinuities may also arise from the change in preferred topoisomer (i.e., over- versus undertwisted) as in the two example loops illustrated below ($\Delta Tw = -0.13$ and $\Delta Tw = 0.07$ for the loops designated by the *open* and *solid stars*, respectively).

Lk of DNA cyclized in the absence of LacI. They interpreted the results in terms of changes in the LacI conformation (open and closed forms), but our earlier work showed that the results can also be explained by considering over- or undertwisting in the loop as well as various binding topologies of the loop. Extending this analysis and the experiments to the entire family of bent sequences should resolve this issue of protein bending flexibility versus DNA twisting and binding topology changes. Fig. 6 *A* illustrates DNA cyclized in the absence of LacI, which establishes the baseline *Lk* for computing ΔLk. Fig. 6 *B* illustrates the minicircle formed by cyclizing the free ends of the DNA after first forming the DNA-LacI complex for the minimum energy loop that forms with the A (anti-parallel) binding topology (and also illustrated in Fig. 4).

The cyclization experiment was modeled numerically using the computational rod model (Fig. 6) by dividing a minicircle into an SHS domain and two straight domains. The SHS domain, with its prescribed intrinsic curvature (*light shaded segments* in Fig. 6), ultimately forms the primary DNA loop, and the straight domain (*dark shaded segments* in Fig. 6) represents the DNA tails outside the interoperator region (with a length of 211 bp). For a given set of phasing parameters Θ_1 and Θ_2, we first simulate the formation of a minicircle in the absence of LacI, as in Fig. 6 *A*. Because the two ends of the SHS representation are connected by a straight segment of DNA and there is no bound protein, only one phasing parameter is required, the sum of Θ_1 and Θ_2. Next, we select the minimum energy looped DNA-LacI complex as in Fig. 4. The operator DNA is considered fixed by the protein crystal structure, and based on the kinetic stability of these complexes we assume that the geometry and topology of this loop remain unchanged during ligation. Starting from this state, we simulate the formation of a secondary loop formed by the tails, which now ligate using the only available binding topology. After choosing the energetic minimum from the under- and overtwisted topoisomers, we compute the *Lk* of the final minicircle with LacI (refer to the computed minicircle example in Fig. 6 *B*). Because the length and phasing of the straight DNA tails that form the secondary loop are held constant, only one minimum energy loop of each binding topology must be computed. As a result, all A1 primary loops (formed with arbitrary phasing of a bent sequence) are closed with identical A2 secondary loops, with a constant calculated elastic energy of 15.5 *kT*. All P1 primary loops are closed with identical P2 secondary loops, with energy 19.6 *kT*. We report in Fig. 7 the computed ΔLk between the loops formed with and without LacI and over the entire design space. The preferred binding topology of the primary loop is denoted in the figure (*thin dashed lines*). In addition, we identify in Fig. 7 the discrete changes in *Lk* of the energetically preferred minicircle formed in the absence of LacI (*thick dashed lines*), which affects the baseline for calculating ΔLk.

DISCUSSION

In a previous article (12), we established that the computational rod model predicts the major experimental findings for the looping of three bent sequences synthesized to date (3,38,39). In this article, we introduce the SHS representation for an entire family of these bent sequences and systematically explore possible looping behaviors. Doing so reveals new insights on how intrinsic curvature influences the looping of DNA, which in turn motivates future experimental studies.

The differences in the looping behaviors of the three highly bent sequences (7C16, 9C14, and 11C12) (3) originate from the differences in phasing of the A-tract domain. In particular, the three sequences shift the A-tract domain by 2 bp relative to the ends of the linker domains, which leads to a substantial (~70°) torsional phase difference of the A-tract from one sequence to another. The energy contour plot of Fig. 8 illustrates the approximate positions of the three synthesized sequences within the Θ_1-Θ_2 design space, estimated as de-

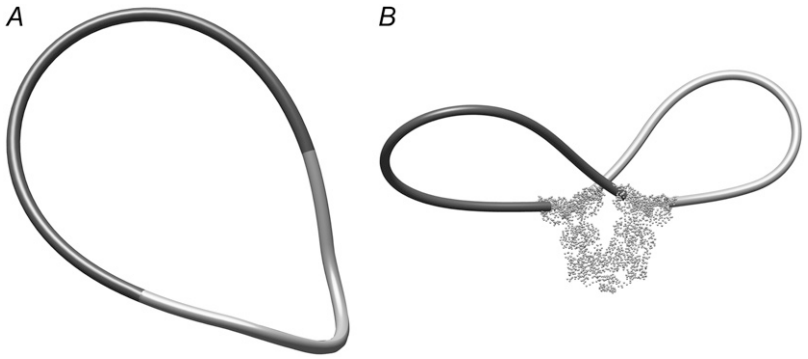

FIGURE 6 Computed minicircles mimicking the experimental procedure of Mehta and Kahn (3). Light-shaded segments represent the highly bent sequence (SHS representation) and the dark-shaded segments represent the intrinsically straight sequences forming the tails. Since the operator DNA is considered fixed by the protein crystal, there is a visible gap between the light- and dark-shaded rods. (*A*) Computed minicircle cyclized in the absence of LacI. For this sample minicircle, $\Delta Tw^{A} = 0.38$ and $Wr^{A} = 0.16$. (*B*) Minicircle cyclized subsequent to the formation of the looped DNA-LacI complex. ΔLk is defined as the Lk of the minicircle cyclized in the absence of LacI minus the Lk of the minicircle cyclized subsequent to the formation of the LacI-DNA looped. This sample minicircle is formed from the minimum energy loop illustrated in Fig. 4 with $\Delta Tw^{B} = 0.14$ and $Wr^{B} = -0.56$. The resultant change in Lk is $\Delta Lk = \Delta Tw^{B} + Wr^{B} - (\Delta Tw^{A} + Wr^{A}) = -1$.

scribed in Methods. Note that the three synthesized sequences span only a modest fraction of the overall design space, which has been largely unexplored to date. The assignment of Θ_1 and Θ_2 depends on an assumed helical repeat and DNA curvature model, but the relative properties of DNA constructs differing by the indicated number of turns would be unaffected by errors in the absolute positions of the existing molecules.

The larger view of the design space illustrated in Fig. 8 is obtained by tiling the energy contour plot of Fig. 4, assuming that the loop elastic energy remains periodic in the two phasing parameters Θ_1 and Θ_2. (We again emphasize that any variation in the loop elastic energy due to the small changes in the sequence contour length are inconsequential relative to those induced by changes in phasing (12)). To highlight this periodicity, Fig. 9 presents one cross section of the energy landscape of Fig. 8. The periodic variation in elastic energy for this family of bent sequences is strikingly similar to the periodic variation in repression level (or elastic energy and free energy) for nominally straight DNA due to torsional phasing as discussed, for example, in the literature (33–35,48). Fig. 8 illustrates that there exist distinct regions of the design space where the preferred binding topology is P1 (*outlined* with *thin dashed lines*) versus A. Thus, by appropriate experimental design, one may bias the preferred binding topology. Also, the elastic energy cost of looping varies significantly—by more than a factor of two—over the design space. Again, this suggests that by appropriate experimental design, one may significantly influence the loop stability. We explore these possibilities further below.

Although it is evident from Fig. 8 that the A-region (*outlined* with *thin dashed lines*) is larger than the P1-region, both occupy significant fractions of the design space (i.e., both regions encompass many possible bent sequences). By contrast, only the A-binding topology is preferred for intrinsically straight DNA of similar contour length, assuming no protein flexibility (16,31). Thus, for the P1-region of the design space, properly phased intrinsic curvature is necessary and sufficient to decrease the elastic energy of P1 enough to overcome the bias toward the A-binding topology. As with intrinsically straight DNA on these length scales (~140 bp), the P2-binding topology is never preferred (31).

The experimental and computational data for the three synthesized sequences already provide evidence for two distinct looped states: qualitatively different topoisomer distributions and FRET behavior have been observed for the 9C14 and 11C12 molecules. The existence of the two states was proposed to arise primarily from flexibility of LacI at the base of the ''V'' (3,26,39). The results presented here and in Goyal et al. (12) suggest an alternative explanation, that distinct

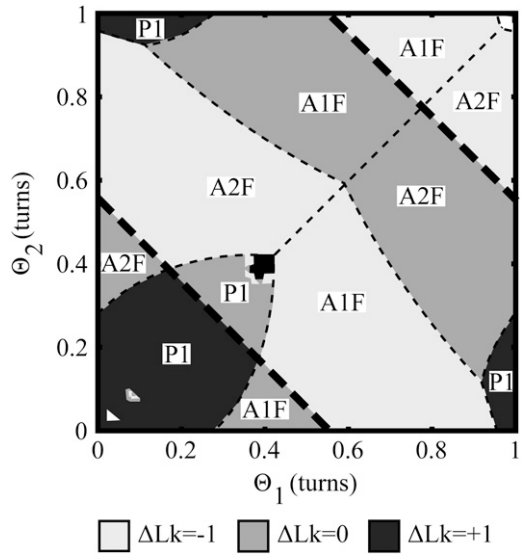

FIGURE 7 Map of relative linking number (ΔLk) for minicircles formed from the DNA-LacI complex, with ΔLk defined as in Fig. 6. ΔLk is computed over the entire design space. Preferred binding topologies are demarcated with dashed lines as in Fig. 5. The thick dashed lines denote discrete changes in ΔLk resulting from discrete changes in Lk of the DNA minicircle formed in the absence of LacI. Isolated discontinuities in ΔLk within the P1-region along the diagonal ($\Theta_1 = \Theta_2$) arise from changes in writhe (± 2) due to changes in the handedness (sign) of the crossings of the primary and secondary loops (see, for example, (44)).

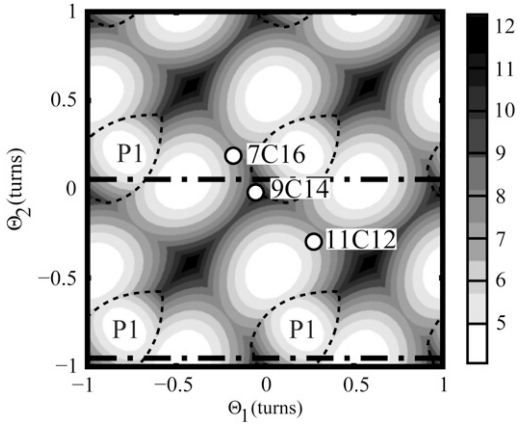

FIGURE 8 Elastic energy contour map (Fig. 4) extended using tiling. The thin dashed lines outline the regions corresponding to the P1-binding topology; the remaining region corresponds to the A-binding topology. The indicated locations of the synthesized sequences are a result of mapping the representation of intrinsically curved DNA used in Goyal et al. (12) onto the SHS representation. The thick dashed lines indicates the plane for the cross section of Fig. 9. (These two lines are separated by one turn and therefore result in identical cross sections.)

looped states could arise from distinct preferred binding topologies, as also proposed in Saiz and Vilar (37) for straight DNA. The role of LacI flexibility has been examined experimentally and computationally but a clear consensus has not yet emerged. For example, an MD model of the LacI protein suggests a stiff protein ''V'' (24) with flexible head domains. By contrast, electron microscopy and x-ray scattering demonstrate a flexible V-region in solution (29,30). DNA looping models that include approximate treatments of protein flexibility align reasonably well with experimental data (16,26,31) and thereby support a flexible V.

Whether protein flexibility or binding topology is the root cause of the observed multiple looped states is a persistent question that could be resolved by further experiment. Some of the most compelling experimental data derive from FRET assays on DNA molecules where donor and acceptor fluorophores are attached at opposite ends of the interoperator DNA sequence. Single molecule measurements on the looped LacI-9C14 complex give a high FRET efficiency, corresponding to a distance of ~35 Å, indicating that the 9C14 sequence forms primarily P1-loops with no observable protein deformation (38). This experiment is well suited to detect the P1-topology, but is unable to distinguish other binding topologies or protein flexibility because of their expected low FRET efficiency. Repositioning one fluorophore to the DNA tail just outside the operator domain could allow FRET detection of the A-binding topology, as also suggested in Swigon et al. (16). A high FRET efficiency would confirm the A-binding topology, whereas a low FRET efficiency would confirm protein flexibility. Another experiment to probe protein flexibility would start with a loop suspected of having large protein deformation and exhibiting low FRET efficiency. The loop could then be cleaved using a restriction enzyme or alternatively nicked at selected sites, significantly reducing its stiffness. Negligible protein deformation would be implicated if the FRET efficiency did not change. Conversely, a significantly increased FRET efficiency would confirm protein deformation in the initial loop. However, if the system equilibrates rapidly after cleaving the DNA, the sandwich complex could also form at random and thereby confound the experiment.

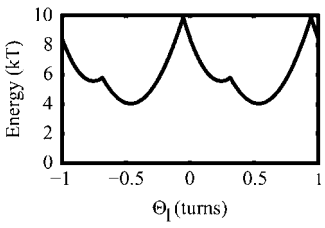

FIGURE 9 Cross section of the energy landscape along the thick dashed lines of Fig. 8.

A second major result apparent in Fig. 4 is that the energetic cost of looping varies by more than a factor of two (from ~4 kT to 12 kT) across the family of bent sequences. Consequently, one would expect large differences in loop formation and breakdown rates and loop stability for representative bent sequences which span the design space. (If we assume that the entropic contribution to the free energy of looping is nearly the same for all sequences in the family of bent sequences, we calculate the relative equilibrium constants between the minimum and maximum elastic energy loops to be: $k_{\mathrm{loop,min}}/k_{\mathrm{loop,max}} = \exp[12 - 4] \sim 3000$.) The energy contour map in Fig. 4 could be used as a guide to design new sequences to minimize (or maximize) the elastic energy cost of looping, starting from the three previously synthesized sequences. Sequences could be synthesized specifically to probe the energetic extremes leading to comparatively hyper- and hypostable loops. Interestingly, the elastic energy cost of looping for the local minima within the P1- and A-regions are nearly equivalent, suggesting that hyperstable P1- and A-loops could each be produced. They would prove useful in future experiments due to their extreme stability.

Another possibility arises from considering the borders between the P1- and A-regions. The bent sequences near these borders may form loops with distinct binding topologies but with near-equivalent energetic costs, so that one would expect to observe a near-equal distribution of loops having two distinct binding topologies. In single molecule experiments (e.g., SM-FRET), one might further hope to observe interconversions among the different looped states and the unlooped state. Though Edelman et al. (39) suggest that such interconversions occur on a long timescale for hyperstable loops because of their energetic stability, the interconversions could be accelerated with low concentrations of the inducer IPTG. Any such transitions were not observable in the SM-FRET study of Morgan et al. (38) because the looped complexes freely diffused through the microscope's field of view.

The radius of gyration (R_g) for the computed loops reported in Fig. 5 is a measure of loop size and thus one factor influencing the speed of a looped complex through a gel. However, the gel matrix or ionic conditions in the gel may alter the preferred loop topology. Among the three sequences synthesized to date, the experimental mobilities for loops that form the P1-binding topology are larger (suggesting a more compact loop) than the mobilities of loops believed to have the A-binding topology or an open form LacI (12). This result holds true for much of the bent sequence design space, as seen by the dashed lines of Fig. 5 demarcating the preferred binding topology. (An exception exists in a very small region in the center of Fig. 5 where compact loops with the A-binding topology are predicted.) We further observe in Fig. 5 that discontinuities in R_g are a result of changes in preferred binding topologies and topoisomers. The discontinuities in loop size suggest the design of gel assays to detect distinct binding topologies and/or topoisomers as a function of the two phasing parameters Θ_1 and Θ_2.

The computed change in linking number ΔLk is also discontinuous across regions of preferred binding topologies as well as preferred topoisomers. In addition, we observe that $\Delta Lk = +1$ occurs only within the P1-region. This fact suggests yet another experiment to detect loops that preferentially form with the P1-binding topology as the discrete changes in ΔLk are relatively easy to detect.

Calculations of ΔLk depend on the topologies of three important components: the minicircle formed in the absence of LacI, the primary loop, and the secondary loop. The influence of the minicircle formed in the absence of LacI becomes pronounced at the discontinuities marked by the thick dashed lines in Fig. 7. These discontinuities result from a change in the energetically preferred topoisomer as the total length of the minicircle changes. As previously discussed, this length depends upon the sum of Θ_1 and Θ_2. Thus, these discontinuities follow the diagonal lines ($\Theta_1 + \Theta_2 =$ constant) in Fig. 7. The primary loop has the most pronounced effect on ΔLk. The primary loop is formed by the most energetically favored binding topology and topoisomer available for a pair of phasing parameters Θ_1 and Θ_2. Changes in preferred binding topology or topoisomer with changes in these phasing parameters result in the majority of the discontinuities observable in Fig. 7. Finally, the topology of the secondary loop has relatively little impact on ΔLk because it remains constant, regardless of phasing, among all complexes in which the primary loops have the same binding topology. The role of the secondary loop is, however, quite pronounced at very isolated points within the P1-region along the diagonal ($\Theta_1 = \Theta_2$) of Fig. 7. These points arise from the changes in writhe (± 2) due to changes in the handedness (sign) of the crossings of the primary and secondary loops (see, for example, (44)). In these regions the effect of electrostatic repulsion, which we do not consider, may be important.

In addition to measuring the ΔLk of the minicircles, Mehta and Kahn (3) also measured the corresponding cyclization rates. These rates are strongly influenced by the energetic cost of forming the secondary loop from the DNA tails outside the operator region. Experiments on the three synthesized sequences show that loops having $\Delta Lk = +1$ form more slowly than all others. Fig. 7 supports this observation. In particular, note that $\Delta Lk = +1$ arises only for loops with the P1-binding topology. During cyclization, the tails must therefore form a secondary loop with the (remaining) P2-binding topology. However, as noted in the discussion above of the energetic cost of looping, the P2-binding topology leads to the least energetically favorable loops, a finding also in agreement with Zhang et al. (31). In fact, we calculate the energetic cost of forming a loop from the DNA tails with the A- and the P2-binding topologies to be 12.5 kT and 19.6 kT, respectively.

The landscapes presented here suggest that it would be useful to synthesize larger sets of molecules in which Θ_1 and Θ_2 are varied systematically. Significant differences in gel migration speed between sequences differing by a single basepair could indicate a sudden change in preferred binding topology or preferred topoisomer. Gel assays offer a simple means to distinguish loops that generate different topoisomers. FRET or tethered particle microscopy techniques can address loop topology and the conformation of the protein.

These computational results also support suggestions based on experiment that LacI-DNA loops are dynamic entities, and that intrinsic DNA curvature in the loops can control their shapes. Cellular DNA bending proteins could certainly have the same effect. Dynamic transitions arising from compositional, chemical, or thermal signals in the cell might play an important role in the control of gene expression.

CONCLUSIONS

This study materially extends the theoretical results of Goyal et al. (12) by analyzing the properties of a family of LacI-DNA loops, using a computational rod model for highly bent interoperator DNA. The extension follows from considering bent sequences composed of two straight linker domains that flank a common helically supercoiled (A-tract) domain. A unifying straight-helical-straight (SHS) representation for the unstressed state leads to a family of molecules that are distinguished by two phasing parameters, Θ_1 and Θ_2, specifying the torsional phase of the A-tract relative to each of the operators. By exploiting the near-symmetry of LacI and the interoperator DNA, we significantly reduce the computational effort in analyzing looping for the entire space of molecules (i.e., over all possible torsional phases $\Theta_1 \in [0, 1]$ and $\Theta_2 \in [0, 1]$ turns). The resulting computations of loop energy and topology reveal new observations of looping mechanics and further suggest compelling experimental studies.

First, the entire two-parameter design space is composed of two subspaces of bent sequences that prefer to bind to LacI with parallel (P1) versus anti-parallel (A) binding topologies. Thus, the distinct looped states observed for three previously synthesized bent sequences (3,39) could arise from differ-

ences in preferred binding topology instead of solely from protein flexibility. Second, sequences located near the parallel versus anti-parallel borders in the design space form loops with distinct binding topologies but with near-equivalent energetic cost. Thus, such loops might readily interconvert between P1- and A-binding topologies (and unlooped states) due to thermal energy. Third, the energetic cost of looping varies by more than a factor of two (from ~4 kT to 12 kT) over the design space, confirming the significant influence of DNA intrinsic curvature on looping. As a result, it should be possible to design bent sequences with controlled looping kinetics and stability.

The ability to synthesize any sequence in the family of bent sequences discussed above creates several possibilities for future experimental studies. For instance, the energy contour map (Fig. 8) may guide the design of new sequences that minimize (or maximize) the elastic energy cost of looping leading to hyperstable (or hypostable) looped complexes. Another possibility is to probe whether multiple looped states (3,39) develop from changes in binding topology or protein flexibility, one may extend the previous FRET experiments (38,39) by repositioning one fluorophore just outside the operator domain to detect the A-binding topology. Gel assays may provide a ready means to observe the distinct (discontinuous) changes in the size of the looped complexes (R_g) and/or the distinct changes in linking number (ΔLk) within known regions of the design space. Finally, possible biological roles of putative loop switching could be addressed in vivo with repression experiments similar to those originally used to identify the presence of LacI-anchored DNA loops.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

The authors recognize the following individuals for engaging discussions on these and related topics: Ioan Andricioaei (Biochemistry, University of Michigan); David Wilson, Chris Meiners, and Alexei Tkachenko (Physics, University of Michigan); Troy Lionberger (Cell and Molecular Biology, University of Michigan); and Seth Blumberg (University of Michigan).

This work is funded by The National Science Foundation under grant Nos. CMS-0510266 and CMMI-0825488.

REFERENCES

1. Semsey, S., K. Virnik, and S. Adhya. 2005. A gamut of loops: meandering DNA. *Trends Biochem. Sci.* 30:334–341.
2. Echols, H. 1990. Nucleoprotein structures initiating DNA replication, transcription, and site-specific recombination. *J. Biol. Chem.* 265: 14697–14700.
3. 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.
4. Cheema, A. K., N. R. Choudhury, and H. K. Das. 1999. A- and T-tract-mediated intrinsic curvature in native DNA between the binding site of the upstream activator NtrC and the nifLA promoter of *Klebsiella pneumoniae* facilitates transcription. *J. Bacteriol.* 181:5296–5302.
5. Balaeff, A., L. Mahadevan, and K. Schulten. 2004. Structural basis for cooperative DNA binding by CAP and Lac repressor. *Structure.* 12: 123–132.
6. Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, and P. Z. Lu. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science.* 271:1247–1254.
7. Swigon, D., and W. K. Olson. 2008. Mesoscale modeling of multi-protein-DNA assemblies: the role of the catabolic activator protein in Lac-repressor-mediated looping. *Intl. J. Non-Linear Mechanics.* In press.
8. Schleif, R. 1992. DNA looping. *Annu. Rev. Biochem.* 61:199–223.
9. Matthews, K. S. 1992. DNA looping. *Microbiol. Mol. Biol. Rev.* 56: 123–136.
10. 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.
11. 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.
12. Goyal, S., T. Lillian, S. Blumberg, J.-C. Meiners, E. Meyhöfer, and N. C. Perkins. 2007. Intrinsic curvature of DNA influences LacR-mediated looping. *Biophys. J.* 93:4342–4359.
13. Schlick, T. 1995. Modeling superhelical DNA: recent analytical and dynamic approaches. *Curr. Opin. Struct. Biol.* 5:245–262.
14. Olson, W. K. 1996. Simulating DNA at low resolution. *Curr. Opin. Struct. Biol.* 6:242–256.
15. Balaeff, A., L. Mahadevan, and K. Schulten. 2006. Modeling DNA loops using the theory of elasticity. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 73:031919.
16. 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.
17. Purohit, P. K., and P. C. Nelson. 2006. Effect of supercoiling on formation of protein-mediated DNA loops. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 74:061907.
18. Olson, W. K., D. Swigon, and B. D. Coleman. 2004. Implications of the dependence of the elastic properties of DNA on nucleotide sequence. *Philos. T. Roy. Soc. A.* 362:1403–1422.
19. Coleman, B. D., W. K. Olson, and D. Swigon. 2003. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* 118:7127–7140.
20. 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.
21. Zhang, Y., A. E. McEwen, D. M. Crothers, and S. D. Levene. 2006. Statistical-mechanical theory of DNA looping. *Biophys. J.* 90:1903–1912.
22. Lankas, F., R. Lavery, and J. H. Maddocks. 2006. Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure.* 14:1527–1534.
23. Villa, E., A. Balaeff, L. Mahadevan, and K. Schulten. 2004. Multiscale method for simulating protein-DNA complexes. *Multiscale Model Simul.* 2:527–553.
24. 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.
25. Podtelezhnikov, A. A., C. D. Mao, N. C. Seeman, and A. Vologodskii. 2000. Multimerization-cyclization of DNA fragments as a method of conformational analysis. *Biophys. J.* 79:2692–2704.
26. Kahn, J. D., R. Cheong, L. M. Edelman, R. A. Mehta, and M. A. Morgan. 2006. Flexibility and control of protein-DNA loops. *Biophys. Rev. Lett.* 1:327–341.
27. Shimada, J., and H. Yamakawa. 1984. Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules.* 17: 689–698.

28. Wilson, D. P., T. Lillian, S. Goyal, A. V. Tkachenko, N. C. Perkins, and J.-C. Meiners. 2007. Understanding the role of thermal fluctuations in DNA looping. *In* Proceedings of SPIE, Vol. 6602. S. M. Bezrukov, editor. SPIE, #660208.

29. 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.

30. Taraban, M., H. Zhan, A. E. Whitten, D. B. Langley, K. S. Matthews, L. Swint-Kruse, and J. Trewhella. 2008. Ligand-induced conformational changes and conformational dynamics in the solution structure of the lactose repressor protein. *J. Mol. Biol.* 376:466–481.

31. Zhang, Y., A. E. McEwen, D. M. Crothers, and S. D. Levene. 2006. Analysis of in-vivo LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS ONE.* 1:e136.

32. Lillian, T. D., S. Goyal, E. Meyhofer, J. D. Kahn, and N. C. Perkins. 2008. Computational elastic rod theory captures DNA and protein flexibility in the Lac-repressor complex. (Abstr.). Proceedings 52nd Annual Meeting, Biophysical Society, Long Beach, CA.

33. Bellomy, 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.

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

35. Becker, N. A., J. D. Kahn, and L. J. Maher. 2005. Bacterial repression loops require enhanced DNA flexibility. *J. Mol. Biol.* 349:716–730.

36. Saiz, L., and J. M. Vilar. 2007. Multilevel deconstruction of the in vivo behavior of looped DNA-protein complexes. *PLoS ONE.* 2:e355.

37. Saiz, L., and J. M. G. Vilar. 2006. In vivo evidence of alternative loop geometries in DNA-protein complexes. *arXiv:q-bio.BM/0602012.*

38. 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.

39. Edelman, L. M., R. Cheong, and J. D. Kahn. 2003. Fluorescence resonance energy transfer over ~130 basepairs in hyperstable Lac repressor-DNA loops. *Biophys. J.* 84:1131–1145.

40. 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.

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

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

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

44. Calladine, C. R., H. R. Drew, B. F. Luisi, and A. A. Travers. 2004. Understanding DNA: The Molecule and How It Works, 3rd Ed. Elsevier/Academic Press, San Diego, CA.

45. Drak, J., and D. M. Crothers. 1991. Helical repeat and chirality effects on DNA gel electrophoretic mobility. *Proc. Natl. Acad. Sci. USA.* 88: 3074–3078.

46. Vlahovicek, K., L. Kajan, and S. Pongor. 2003. DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res.* 31:3686–3687.

47. Geanacopoulos, M., G. Vasmatzis, V. B. Zhurkin, and S. Adhya. 2001. Gal repressosome contains an antiparallel DNA loop. *Nat. Struct. Biol.* 8:432–436.

48. Krämer, H., M. Niemöller, M. Amouyal, B. Revet, B. von Wilcken-Bergmann, and B. Müller-Hill. 1987. Lac repressor forms loops with linear DNA carrying two suitably spaced Lac operators. *EMBO J.* 6: 1481–1491.