# DNA twisting flexibility and the formation of sharply looped protein–DNA complexes

**T. E. Cloutier and J. Widom***

Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, IL 60208-3500

Edited by Roger D. Kornberg, Stanford University School of Medicine, Stanford, CA, and approved January 19, 2005 (received for review December 6, 2004)

**Gene-regulatory complexes often require that pairs of DNA-bound proteins interact by looping-out short (often ≈100-bp) stretches of DNA. The loops can vary in detailed length and sequence and, thus, in total helical twist, which radically alters their geometry. How this variability is accommodated structurally is not known. Here we show that the inherent twistability of 89- to 105-bp DNA circles exceeds theoretical expectation by up to 400-fold. These results can be explained only by greatly enhanced DNA flexibility, not by permanent bends. They invalidate the use of classic theories of flexibility for understanding sharp DNA looping but support predictions of two recent theories. Our findings imply an active role for DNA flexibility in loop formation and suggest that variability in the detailed helical twist of regulatory loops is accommodated naturally by the inherent twistability of the DNA.**

activation | gene regulation | repression

Double-stranded DNA *in vivo* is often sharply distorted away from its classic B-form conformations. In many prokaryotic gene-regulatory complexes, short stretches of DNA, ≈100 bp in length, are bent sharply into circular or antiparallel (U-shaped or teardrop-shaped) loops (1, 2). Looping allows synergy between proteins bound at distant DNA sites (3) and decreases the statistical noise in their occupancy (4). DNA looping is also important in eukaryotic regulatory systems (5–7). A striking recent example is the discovery of ligand-dependent DNA looping by the RXR receptor, which may play an important regulatory role at as many as 172 different locations, genome wide, in the mouse (8). In addition, most (≈75–80%) of the length of eukaryotic genomic DNA is sharply bent into nucleosomes (80-bp superhelical loops) (9), which regulate the accessibility and proximity of other DNA-functional sites (10, 11).

Prokaryotic and eukaryotic regulatory complexes involving short DNA loops (12–18) place strong constraints on the helical twist of the looped DNA (2). For two DNA-bound proteins to interact when they are separated along the DNA, the protein-binding sites need to occur on mutually compatible faces of the DNA double helix. This requirement is satisfied by a set of lengths for the intervening DNA that differ from one another by integral multiples of the DNA helical repeat, ≈10.5 bp. When the exact length of the intervening DNA in such complexes is suboptimal, the DNA may be under- or overtwisted to allow the protein–protein interaction. The DNA helical twist is altered in other biological systems as well, most notably in the nucleosome, in which the wrapped DNA is under- or overtwisted for most of its length (9).

DNA bending and twisting deformations that are required for protein–DNA complex formation come at a cost in free energy, which contributes importantly to the net stabilities and functions of the resulting complexes. For these reasons, the inherent bendability and twistability of DNA itself have been a focus of experimental and theoretical investigation. Classic studies revealed double-stranded DNA to behave as a semiflexible polymer, characterized by a bending-persistence length $P \approx 50$ nm (≈150 bp) (19–26). DNAs that are longer than $P$ are expected to be gently bent spontaneously and to require relatively little force to bend significantly, whereas DNAs that are shorter than $P$ are expected to be nearly straight and to require great force to bend significantly. Similarly, studies of DNA twistability have shown the DNA to behave as an elastically twistable rod, with a torsional modulus of $\approx 2.4–4.5 \times 10^{-19}$ erg·cm (1 erg = 0.1 mJ) (22, 27–29), implying that large amounts of force are also required to significantly twist a DNA of length shorter than the corresponding torsional persistence length, ≈170–320 bp. These findings led to a picture in which sharp DNA bending, and any substantial under- or overtwisting, is achieved *in vivo* by specific proteins that overwhelm the inherent inflexibility of DNA with large force.

Although these classic studies of DNA mechanics were motivated in part by a need to understand the behavior of DNA in sharply looped regulatory complexes and in nucleosomes, the actual behavior of DNA in this regime of sharp bending had never been investigated experimentally. Rather, experiments were carried out in a regime of gentler bending, and theories based on linear elasticity of continuous materials [the Shimada–Yamakawa (SY) theory (19)] or harmonic deformations of base steps [the Zhang–Crothers (ZC) theory (20)] were used to extrapolate the expected behavior into more sharply bent regimes.

We recently used the ligase-mediated cyclization method (21, 23, 26) to directly quantify the ability of 94- and 116-bp DNAs to spontaneously bend and be ligated into covalently closed circles. (We refer to the resulting products as "circles" to indicate their topology and distinguish them from linear oligomers; however, they may be kinked or irregularly curved, not uniformly round; see *Discussion*.) Whereas the current understanding of DNA bendability predicted that such small circles would essentially never form, we found that they formed easily, with a probability that exceeded the predictions of the SY and ZC theories by as much as 100,000-fold.

In the present study we extend our analysis of sharp DNA looping to measure the twistability of sharply looped DNA, which is as important as DNA bendability in determining the stability and function of looped protein–DNA complexes. Our results challenge the classic understanding of DNA mechanics but support two recent theories that invoke a localized DNA melting or kinking in the sharply looped DNA; they also reveal roles for DNA as an active participant in the formation and function of looped regulatory complexes *in vivo*.

## Methods

**DNAs.** Sequences E6-116, E8-116, and E13-116 are randomly chosen clones from a chemical synthesis of 116-bp DNAs having random sequences except for short, defined sequence ends containing *Eag*I sites. Variants in the range of 105–89 bp were created from them by deletion of internal sequence by using PCR. 5S and TA sequences similarly derive from nucleosome-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: SY, Shimada–Yamakawa; ZC, Zhang–Crothers; KWLC, kinkable wormlike chain.

*To whom correspondence should be addressed. E-mail: j-widom@northwestern.edu.

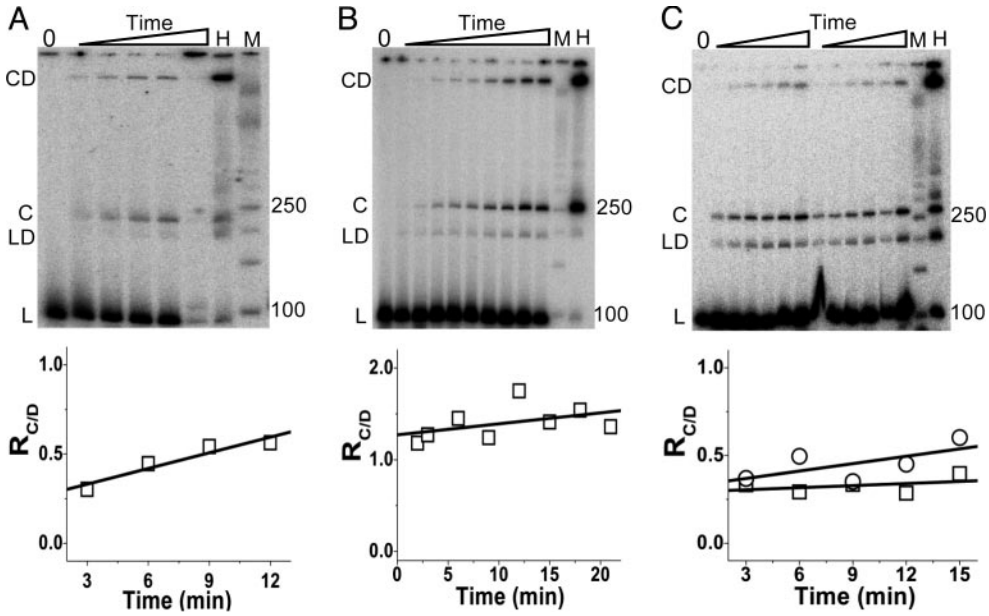


**Fig. 1.** Short DNAs are both highly bendable and highly twistable. Data from cyclization reactions on several DNAs having worst-case lengths for cyclization are shown. Comparable data for best-case lengths are shown in ref. 30. (*Upper*) Phosphorimages. (*Lower*) Quantitative analyses. The graphs plot the ratio of counts in monomer circle to linear and circular dimeric species, as a function of time, for the data from the corresponding phosphorimage indicated above. L, mobility of linear monomer; LD, linear dimer; C, circular monomer; CD, circular dimer; M, molecular weight markers; H, a control reaction carried out at elevated [DNA]; 0, a control with no ligase added. (*A*) DNA sequence E8-89 (the 89-bp-long variant of sequence E8). [DNA] = 2 pM; [ligase] = 300 units·ml$^{-1}$; $T$ = 30°C. (*B*) DNA sequence E13-99; [DNA] = 5 pM; [ligase]= 250 units·ml$^{-1}$; $T$ = 30°C. (*C*) Left-hand tracks on phosphorimage and data points (□) in the graph: DNA sequence TA-100. [DNA]= 25 pM; [ligase]= 250 units·ml$^{-1}$; $T$ = 30°C. Right-hand tracks on phosphorimage and data points (○): DNA sequence E8-100. [DNA]= 5 pM; [ligase]= 250 units·ml$^{-1}$; $T$ = 30°C.

positioning sequences (30). Sequences are listed in Table 1, which is published as supporting information on the PNAS web site. DNAs were isolated from purified plasmid by digestion with *Eag*I, followed by agarose gel electrophoresis and extraction with Elu-tubes (Fermentas, Hanover, MD). Concentrations of purified DNA stocks were determined from their absorbance at 260 nm. DNAs were 5′-end-labeled with [$\gamma$-$^{32}$P]ATP and T4 polynucleotide kinase, followed by a chase with 1 mM unlabeled ATP.

**Cyclization Assays.** Methods and conditions used for quantitative ligase-mediated cyclization assays and methods used to distinguish circular products from linear, and monomeric products from oligomeric, are described in ref. 30.

**Bending and Torsional Elastic Constants.** We estimated the apparent bending and twisting stiffnesses from our cyclization data by least-squares fitting to the SY theory (19). We do not attempt to fit to the ZC theory (20), because it does not converge for short DNA lengths. The DNA helical twist was kept fixed at the canonical value [34.45°/bp (20)], as justified by our finding that, for all sequences examined, 94 bp represented a locally optimal length for cyclization. Bending and twisting stiffnesses were refined iteratively, minimizing the sum of the squared deviations of $\log(J_{\text{measured}})$ from $\log(J_{\text{predicted}})$. The SY predictions were calculated as described (30).

## Results

**Ligation-Mediated Cyclization Assay for DNA Flexibility.** Cyclization assays are particularly well suited to an analysis of sharp looping, because they are specifically sensitive to strongly bent conformations. DNA restriction fragments having self-complementary ends are in rapid equilibrium with base-paired but noncovalently linked circular and linear oligomeric forms. This equilibrium is sampled and trapped by addition of T4 DNA ligase and ATP. Ligase does not contribute to the DNA cyclization process itself; rather, DNA conformational changes that bring the two ends into transient base-paired contact occur spontaneously (23).

The ratio of the rate of formation of covalently closed monomer circles to the rate of formation of covalently joined dimeric species yields the ratio of cyclization to dimerization equilibrium constants (21, 23, 26). This ratio equals the molar concentration of one end of a linear DNA molecule in the vicinity of its own other end, also known as the $J$ factor, a direct measure of flexibility, mathematically related to $P$.

When the length of the DNA is varied in 1-bp increments, $J$ varies with a period equal to the DNA helical repeat (21, 26, 29). This oscillation in $J$ reflects a requirement for the two DNA-cohesive ends to come into proximity with proper rotational alignment, which is needed for the juxtaposed ends to base-pair and be ligatable. DNAs that are longer or shorter than an optimal length need to under- or overtwist to allow base pairing, which occurs with increased energetic cost and corresponding decreased equilibrium probability. The amplitude of the oscillations in $J$ as the length is changed depends on the torsional modulus $C$, allowing $C$ to be calculated.

**Twistability of Sharply Looped DNA.** We used the cyclization assay to quantify the cyclizabilities of 89- to 105-bp-long DNAs. We investigated two nucleosome-positioning DNA sequences and three random DNA sequences. All sequences carried cohesive ends generated by restriction-enzyme digestion.

Cyclization reactions (Fig. 1 *A–C Upper*) show monomer circles to form readily even when the total twist is worst-case (one half of a helical turn away from a nearby optimum), implying that sharply bent DNA is highly flexible for twisting. Quantitative values of $J$ are obtained from these data by linear extrapolation of the ratio of counts in monomeric and dimeric products back to zero time (Fig. 1 *A–C Lower*). The results from many such experiments show a clear oscillation of $J$ with DNA length for all five sequences examined (Fig. 2*A*).

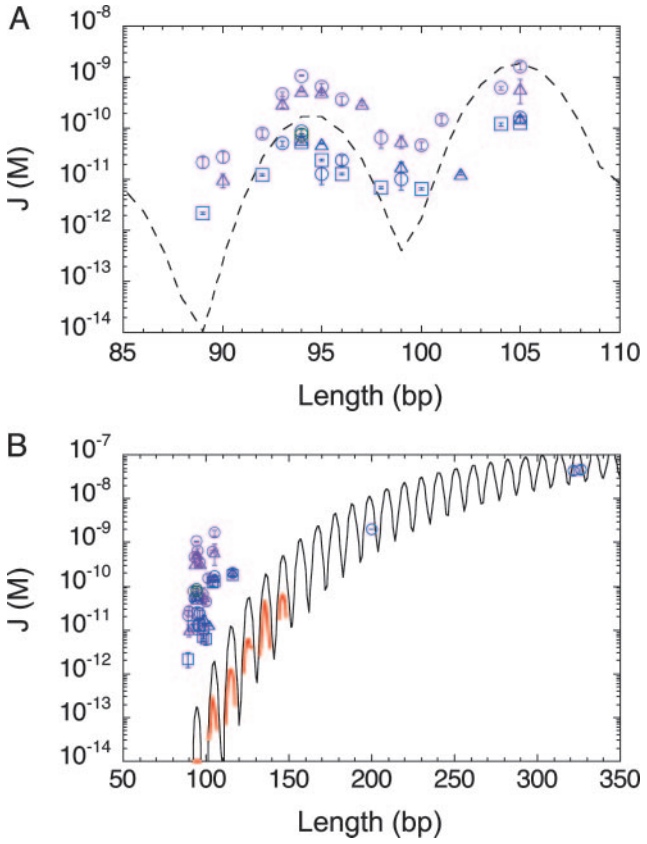


**Fig. 2.** Twistability of sharply looped DNAs. (*A*) *J* factors for five DNA sequences as a function of length. Data points, means and standard deviations for each sample; purple, nucleosome-positioning sequences TA (○) and 5S (△); blue, random sequences E6 (○), E8 (□), and E13 (△); green, sequence CA94 (○). Data for 94-bp fragments of TA, 5S, E6, E8, and E13 and for a subset of the 93- and 95-bp variants are from our earlier study (30). Dashed line, predictions of the SY theory (with $P = 51$ nm and $C = 2.4 \times 10^{-19}$ erg·cm) multiplied by an arbitrary factor of 1,000 to translate the predictions vertically onto the scale of the data. The amplitude of oscillatory changes in *J* predicted by the SY theory greatly exceeds the observed amplitude. (*B*) Measured *J* factors compared to the predictions of the SY (black curve) and ZC (red curve) theories for DNA cyclization. The predictions use the same canonical parameters for DNA bending and twisting flexibility as in *A*. Note the log scale for the ordinate, which covers 7 orders of magnitude. The ZC prediction is discontinuous, because the numerical evaluation of this theory does not converge for short DNAs having lengths that deviate too much from a local optimum. Data points come from the present work and our earlier study (30). Both theories fit the experimental data for longer lengths (19, 20), but both have the wrong functional form to simultaneously describe the cyclization of long and short fragments. For 89- to 94-bp DNAs, the deviation between experiment and theory can exceed several million-fold.

Several aspects of these results are striking. The measured *J* factors greatly exceed the values predicted by the classical elastic or harmonic theories (Fig. 2*B*) [by as much as several million-fold for sequence TA-89 (Fig. 2*A*)], confirming and extending our earlier findings (30).

Not only are all DNAs examined far more easily cyclizable than expected theoretically, but they also are far more easily twistable. The dashed line in Fig. 2*A* shows the prediction of the SY theory for canonical values of the DNA bendability and twistability, multiplied by an arbitrary factor (1,000) to translate the curve upward for comparison with our data. Whereas the SY theory predicts that *J* will decrease by ≈10,000-fold between 94 and 89 bp, the measured *J* factors for sequences TA and E8 decrease by only ≈50-fold and ≈25-fold, respectively; that is, these 89-bp DNAs are 200- and 400-fold more easily twistable than predicted. Moreover, these calculations assume the lower canonical value for the torsional modulus of DNA, $\approx 2.4 \times 10^{-19}$ erg·cm (22, 29); using the value $\approx 4.5 \times 10^{-19}$ erg·cm, which comes from more direct measurements (27, 28), would increase even further the disparity between the predicted decrease in *J* and that which is actually observed. Similar effects are observed for all five sequences when the lengths are increased from 94 bp to the next larger minimum at 99–100 bp. Sharply bent DNAs are far more easily twistable than is currently understood.

*J* values for the nucleosome-positioning sequences (5S, TA) systematically exceed those for the random sequences (E6, E8, E13), which is consistent with our earlier conclusion that the nucleosome-positioning sequences are unusually flexible for sharp bending (30). However, the amplitude of the oscillation in *J* for the nucleosome-positioning sequences is at least as great as that for the random sequences, implying that the nucleosome-positioning sequences are not more flexible for twisting than are the random sequences.

Replacing periodic TA steps in sequence TA94 with CA/TG steps reduced the *J* factor by ≈14-fold, down close to the values of the chemically random sequences. It is evident that the TA steps in sequence TA94 contribute importantly to the flexibility and nucleosome-positioning power of this high-affinity nucleosome-positioning sequence (31).

We asked what values of bending and twisting stiffnesses were required to fit these data to the SY theory. Including all data in the range of 89–105 bp (all sequences simultaneously) yielded best-fit parameters of bending-persistence length $P = 113$ bp and torsional modulus $C = 7.4 \times 10^{-20}$ erg·cm (Fig. 3*A*). Although these parameters allow *J* factors for the short DNAs to be reasonably well fit, the resulting curves deviate substantially from the data for longer DNAs (Fig. 3*B*; see also ref. 30). Because the *J* factors of nucleosome-positioning and random sequences differed systematically from each other, we also fit them separately, obtaining values of $P = 105$ bp and $C = 1.03 \times 10^{-19}$ erg·cm for the nucleosome-positioning sequences and $P = 119$ bp and $C = 7.4 \times 10^{-20}$ erg·cm for the random sequences; however, these fits, too, deviate systematically from the experimental measurements of *J* factors for longer fragments (results not shown).

## Discussion

**Failure of Classic Theories of DNA Flexibility.** Permanent bendedness in short DNAs can greatly enhance their cyclization probabilities (20); hence, our earlier results could be attributed to either a breakdown of the SY and ZC theories in the sharply bending regime or significant permanent bendedness in the DNAs. However, our current results on DNA twistability can only be explained by a breakdown of the SY and ZC theories. Permanent bendedness cannot explain the greatly enhanced twistability of DNA in this regime, which exceeds theoretical prediction by up to 200- to 400-fold (Fig. 2*A*). Moreover, both theories fail to simultaneously describe the behavior of DNA in the gently bending regime studied earlier and in the sharply looping regime (Fig. 3*B*). These failures of the two theories must represent a breakdown, in the sharply bending regime, of the assumptions of linear elasticity or harmonic bending on which the theories are founded.

**Recent Theories of DNA Flexibility.** In response to our earlier findings of exceptional cyclizabilities of 94-bp DNAs, two new models for DNA bendability were developed, having in common an assumption of conformational changes in the DNA that occur with significant energetic cost but that create sites of enhanced local flexibility. One model (ref. 32 and P. Ranjith, P. B. Sunil Kumar, and G. I. Menon, personal communication), henceforth referred to as the melted-bubble model, supposes that short melted bubbles, present at equilibrium in double-stranded DNA

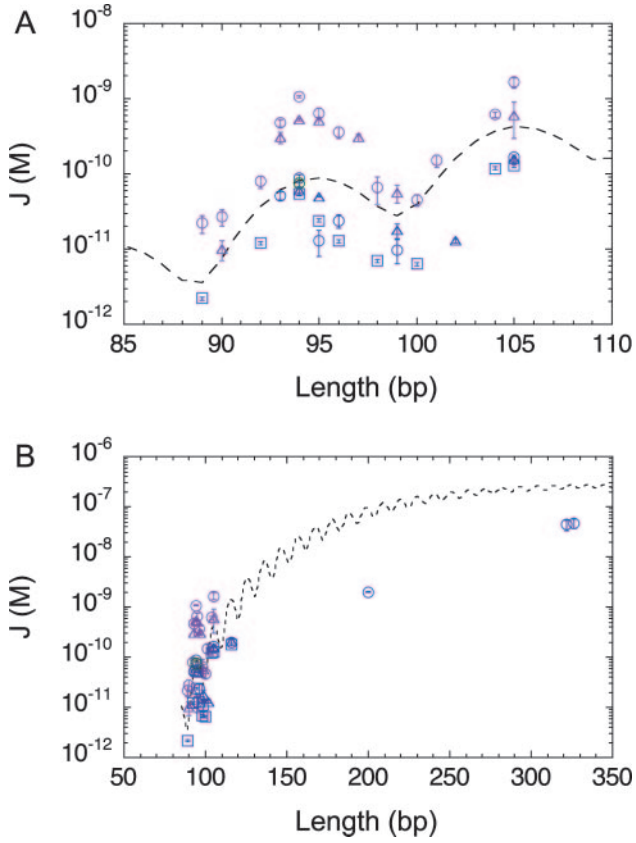


**Fig. 3.** Apparent bendability and twistability of sharply looped DNAs. (*A*) Nonlinear least-squares fitting of the SY theory to the experimental data for 89- to 105-bp DNAs, with bending-persistence length $P$ and torsional modulus $C$ as variables. Optimum values obtained are $P = 113$ bp and $C = 7.4 \times 10^{-20}$ erg·cm. See *Results* for results of fits to the nucleosome-positioning sequences and the random sequences independently. (*B*) Predictions of the SY theory using these optimal $P$ and $C$ values, extended to longer lengths, compared to our data for short and longer DNAs. Using the best-fit $P$ and $C$ values, obtained for the random-sequence data only, negligibly changes the predictions (results not shown).

in ambient conditions, allow for enhanced DNA flexibility at the melted sites; the other model (33), henceforth referred to as the kinkable wormlike chain (KWLC) model, invokes an unspecified local conformation that has enhanced flexibility. [Two additional theories are related to these theories but have not been applied to our cyclization data (34, 35).] The KWLC theory predicts that, in the sharply looping regime studied here, cyclization is dominated by doubly kinked (or two-bubble) conformations. Both models explain the cyclizabilities of the 94- and 116-bp DNAs measured in our earlier study using reasonable values for the energetic cost of the proposed conformational transitions: $\approx 11\ k_BT$ for the melted-bubble model and $\approx 15\ k_BT$ per kink for the KWLC model ($k_B$ is Boltzmann's constant and $T$ is the temperature).

Both the melted-bubble and KWLC theories seem consistent with our data (J. Marko, P. Nelson, and M. Schurr, personal communications; see also *Supporting Text*, which is published as supporting information on the PNAS web site). Our measured torsional moduli represent (in the context of these theories) a weighted reciprocal average of the moduli for the predominant B-form DNA plus a small amount of short melted/kinked DNA and are quantitatively consistent with the presence of a few base pairs of melted/flexibly kinked DNA, as assumed in the melted-bubble and KWLC models. Conversely, given the experimental moduli for short melted regions, our measured moduli imply the presence of (or the equivalent of) ≈2–4 melted bp in the sharply looped DNAs, and the corresponding free-energy cost of such melting is in good agreement with the kink (or bubble) free energies that were estimated from our original data on cyclization of 94-bp DNAs, i.e., based on torsion-independent loop-closure probability considerations only.

Beyond complete local melting, the phenomenon of base "flip-out" also allows our data to be understood with both recent theories. Individual bases in DNA transiently rotate out of the double helix on the millisecond time scale (36), creating localized highly flexible defects (37). The corresponding equilibrium constants (38) imply a typical energetic cost for base flip-out of $\approx 12$–$14\ k_BT$, again in remarkable agreement with the kink energies estimated from fitting the two theories to our original data.

Although our data are broadly consistent with the melted-bubble and KWLC theories, neither theory as currently developed accounts for the nonzero stiffnesses of the melted/flexible regions observed in this study, nor can they quantitatively account for the dependence of $J$ on the detailed length of DNA through a helical turn. Additional theoretical development would be useful.

**Molecular Mechanics of Sharp DNA Looping.** Our results establish that, as DNA is asked to bend ever-more sharply, it finds ways of doing so at a reduced energetic cost relative to that expected for a continuous elastic or discrete harmonic material. Such behavior is not surprising in principle. Crick and Klug (39) noted that this regime of sharp looping is distinctive in that the radius of curvature of the looped DNA is comparable to the molecular radius of the DNA itself. It follows that bending strains are likely to be large, and thus it would not be surprising if the atomicity of DNA allowed sharp bending and twisting to occur at reduced cost through localized structural deformations. Real materials under bending force typically fail by localized buckling, relieving most of the bending strain at the buckled sites. This fact underlies and motivates the recent theories for DNA bending.

Localized melted/kinked regions may be expected in sharply bent DNAs because, for example, the energetic cost of bending a 94-bp DNA into a circle, assuming linear elastic behavior, is $\approx 31\ k_BT$ (32), far greater than the energetic cost of forming a small melted bubble or flipping out one base, either of which would greatly relieve the bending strain because both melted bubbles and DNA with a flipped-out base are highly flexible (37, 40).

Beyond just DNA melting or base flip-out, however, there are many other possibilities for localized structural defects that could potentially account for our observations. The only unifying requirement among these possibilities is that they confer greatly enhanced bending and torsional flexibility, not merely permanent bendedness. TA steps exhibit an unusual NMR line broadening of the adenine N2 proton, indicating large amplitude motions that explore long-lived altered conformational states (41). Zn-dependent kinks have been visualized in DNA by atomic force microscopy (42). Another atomic force microscopy study reveals a systematic disparity between the fourth moment of the tangent angle distribution and the square of the second moment, exceeding that attributable to image processing (43), suggesting a softening of DNA on short length scales that is consistent with localized flexible regions. Specific flexibly kinked DNA structures have been observed in all-atom molecular dynamics simulations of a 94-bp DNA minicircle (F. Lankas, R. Lavery, and J. H. Maddocks, personal communication). Localized structural deformations have been detected biochemically in sharp DNA loops stabilized by Lac repressor (44), lambda repressor (45), and nucleosomes (46), although their mechanical consequences are not known. Many specific kinked structures in protein–DNA complexes are understood at atomic

resolution, although, again, their mechanical consequences are not known. It remains to be determined which, if any, of these modes of localized structural deformation contribute importantly to the ability of naked DNA to sharply bend and twist spontaneously.

Note, too, that melted/kinked regions may be mobile. The KWLC model implies that melted/kinked regions will be present, on average, twice per sharply looped DNA for DNAs in the size range that we investigated. Even if the melted/kinked regions occur with enhanced probability at specific locations in the cyclized sequences (e.g., regions of lower thermodynamic stability against melting or base flip-out), they will be free to diffuse around the chain, subject only to their varying intrinsic local probabilities and to any restrictions on their mutual locations in the cyclized conformations.

Although localized flexible defects are consistent with our observations and are expected theoretically, models for enhanced DNA flexibility that are based on continuous deformation of the DNA are not excluded by our results. The only clear requirement is that the energetic cost of deforming the DNA must grow much more slowly than quadratic in the curvature. Indeed, the structure of the nucleosome itself can be viewed in this light. Although the nucleosome structure does reveal specific points with substantial localized structural distortion, much of the wrapped length is reasonably well described by a continuously but slightly deformed structure, in which the base roll and tilt angles oscillate continuously between modest positive and negative values (9).

**Implications for Tests of DNA Looping *in Vivo*.** A hallmark of looped regulatory complexes is that their probability (occupancy) oscillates as the length of the intervening DNA is changed in 1-bp intervals, with the period of oscillation equal to the DNA helical repeat (2). Our data on cyclization reveal only ≈25-fold changes in the *J* factor for cyclization of random-sequence 90- to 100-bp DNAs, and earlier studies reveal only ≈25-fold changes for cyclization of ≈200-bp DNAs (22), which means that effects on occupancy from changing the detailed length of sharply looped DNA will not ordinarily exceed ≈25-fold.

Qualitative bioassays for looping *in vivo* could miss such finite-sized effects and thereby mistakenly lead to the conclusion that looping did not occur. For example, transcriptional silencing of a *URA3* gene in *Saccharomyces cerevisiae* reduces *URA3* transcription by 30-fold, yet cells with the silenced gene can still grow in media lacking uracil (47). Quantitative rather than qualitative assays are required for definitive tests of looping *in vivo*.

**Biological Roles of Sharp DNA Looping.** Our results imply an active role for the formation and function of looped regulatory complexes *in vivo*. The easy cyclizability of DNAs having suboptimal twists cannot be attributed to permanent bendedness but must represent exceptional flexibility inherent even to random-sequence DNAs, which implies that essentially any DNA will be capable of forming sharply looped complexes; yet, differing sequence-dependent DNA bendabilities and twistabilities can contribute to important differences in the stabilities of these complexes.

The exceptional flexibility of DNA for sharp looping explains how sharp loops can form in times that are far shorter than the cell-doubling time, in contrast to previous theoretical expectation (48), and it implies that DNA looping can contribute importantly to both increasing the occupancy of proteins in looped regulatory complexes (30) and to decreasing the inevitable statistical noise in their occupancy (4).

Our findings provide a simple explanation for the ability of two gal-repressor dimers to form a tetrameric repressosome even in the absence of heat-unstable nucleoid (HU) protein (16): the free energy of protein–protein interaction suffices to overcome the real cost in free energy of looping the DNA, which is far less than predicted from the classic theories of DNA flexibility. More generally, our findings suggest that HU and analogous proteins, rather than being responsible for deciding which loops will form in the first place, may instead function primarily to lock-in and stabilize looped states that form spontaneously by site-specific binding proteins interacting on inherently flexible DNA.

Our results provide an explanation for a puzzling aspect of existing *in vivo* looping data, namely, an unexpectedly high relative probability of complex formation even for worst-case lengths of the intervening DNA. Measurements of looping probabilities *in vivo* reveal surprisingly small dependences for changing the length of intervening DNA from a local optimum to the adjacent worst case: only an ≈5- to 20-fold range for AraC (15) and only an ≈20-fold (14) or ≈3- to 20-fold (13) range for Lac repressor, whereas far larger (e.g., 10,000-fold) effects were predicted based on the classical theories of DNA flexibility. These smaller-than-expected dependences of looping probability on twist are not simply attributable to the supercoiled state of DNA *in vivo* (49), because equally small changes (only ≈2- to 5-fold) were also observed for NTRC activation by using relaxed DNAs *in vitro* (12); nor are the smaller-than-expected dependences simply an artifact of large background activities masking the detection of a far larger dependence: the relative probabilities measured at the local minima are adequately larger than the

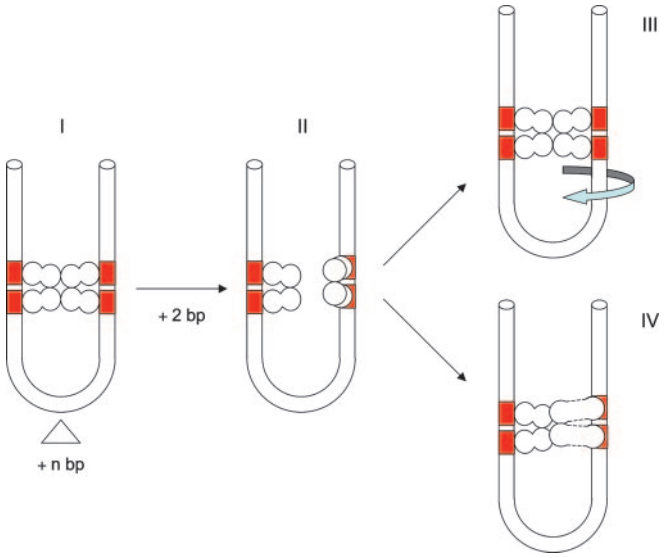


**Fig. 4.** Role for DNA twisting flexibility in the stabilization of sharply looped protein–DNA complexes. A DNA molecule is shown having two palindromic target sites (each represented as a pair of red rectangles). Each site is bound by a protein dimer, with each protein represented as having two domains, a DNA-binding domain and a dimerization domain, which are shown here in intimate contact, as is often the case (e.g., ref. 51). In state I the length and sequence of the looped DNA are such that its total twist allows two protein dimers to interact favorably. Changing the length of the looped DNA even by as little as one or two base pairs or, equivalently, changing its sequence to result in the same change in twist (52) causes a large-scale rotation of one protein dimer about the DNA helix axis, disrupting the protein–protein contacts (state II). If the protein were sufficiently flexible, then a large-scale protein conformational change could restore the protein–protein contacts (state IV). However, such hypothetical flexibility has not been documented, and the available evidence argues that it sometimes may not exist (50). Alternatively, under- or overtwisting the looped DNA (state III) restores the protein–protein interactions, with no requirement for flexibility in the protein. Our findings imply that even worst-case changes to the twist (one half of a helical turn) come at a surprisingly low cost in free energy and corresponding probability.

uncertainty in the measured background (e.g., basal promoter) activities.

Such observations must mean that sharply looped protein–DNA complexes *in vivo* have a remarkable ability to deform structurally to counteract the unavoidable geometric consequences of changing the total helical twist of the looped DNA and that they do so with remarkably little cost in free energy (probability). In principle, the structural effects of small changes in twist of the looped DNA could be accommodated with flexibility in either the proteins or the DNA (Fig. 4). However, because of the nature of the structural change, which involves a rigid body rotation of one of the interacting proteins around the DNA helix axis, a large structural alteration of the protein would be required to restore the original protein–protein contacts (Fig. 4, state IV). Such large-scale conformational changes are possible in principle but have not been documented. Indeed, the available evidence argues in favor of rigid, not flexible, regulatory protein monomers (50). Our findings establish that sharply looped DNA itself inherently has the flexibility needed to restore the protein–protein interactions at a low cost in free energy simply by under- or overtwisting the DNA to restore the total twist back to its original value (Fig. 4, state III). Additional flexibility in the protein, if any, could lower this cost even further, but the enhanced real twisting flexibility of sharply looped DNA revealed in our study explains most, if not all, of the observed insensitivity of looping probability to the detailed total twist of looped DNA *in vivo*.

Until now, sharp looping of DNA in gene-regulatory complexes has been described and analyzed chiefly in prokaryotic systems, with only a few examples documented in eukaryotes. However, the recent discovery of ligand-dependent looping by the mouse RXR receptor, together with the finding of 172 sites genome-wide at which RXR DNA-binding sites occur in pairs spaced 30 to 500 bp apart (8), suggests that sharply looped gene-regulatory complexes may prove to be far more common and important in eukaryotic systems than had been imagined previously.

We thank J. Marko, P. Nelson, R. Phillips, M. Schurr, and P. Wiggins for valuable discussions and comments on the manuscript; A. Vologodskii for the 200-bp DNA fragment; and Y. Zhang for the calculations of the Zhang–Crothers theory and a computer program used for numerical evaluation of the Shimada–Yamakawa theory. We acknowledge the use of instruments in the Keck Biophysics Facility at Northwestern University. This work was supported by National Institutes of Health Grants R01-GM054692 and R01-GM058617 (to J.W.).

1. Gralla, J. D. (1991) *Cell* **66,** 415–418.
2. Schleif, R. (1992) *Annu. Rev. Biochem.* **61,** 199–223.
3. Rippe, K., von Hippel, P. H. & Langowski, J. (1995) *Trends Biochem. Sci.* **20,** 500–506.
4. Vilar, J. M. G. & Leibler, S. (2003) *J. Mol. Biol.* **331,** 981–989.
5. Blackwood, E. M. & Kadonaga, J. T. (1998) *Science* **281,** 60–63.
6. Zeller, R. W., Griffith, J. D., Moore, J. G., Kirchhamer, C. V., Britten, R. J. & Davidson, E. H. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 2989–2993.
7. Su, W., Jackson, S., Tjian, R. & Echols, H. (1991) *Genes Dev.* **5,** 820–826.
8. Yasmin, R., Yeung, K. T., Chung, R. H., Gaczynska, M. E., Osmulski, P. A. & Noy, N. (2004) *J. Mol. Biol.* **343,** 327–338.
9. Richmond, T. J. & Davey, C. A. (2003) *Nature* **423,** 145–150.
10. Lu, Q., Wallrath, L. L. & Elgin, S. C. (1995) *EMBO J.* **14,** 4738–4746.
11. Schild, C., Claret, F.-X., Wahli, W. & Wolffe, A. P. (1993) *EMBO J.* **12,** 423–433.
12. Minchin, S. D., Austin, S. & Dixon, R. A. (1989) *EMBO J.* **8,** 3491–3499.
13. Law, S. M., Bellomy, G. R., Schlax, P. J. & Record, M. T. (1993) *J. Mol. Biol.* **230,** 161–173.
14. Muller, J., Oehler, S. & Muller-Hill, B. (1996) *J. Mol. Biol.* **257,** 21–29.
15. Lee, D. H. & Schleif, R. F. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 476–480.
16. Semsey, S., Tolstorukov, M. Y., Virnik, K., Zhurkin, V. B. & Adhya, S. (2004) *Genes Dev.* **18,** 1898–1907.
17. Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. & Chambon, P. (1986) *Nature* **319,** 121–126.
18. Cohen, R. S. & Meselson, M. (1988) *Nature* **332,** 856–858.
19. Shimada, J. & Yamakawa, H. (1984) *Macromolecules* **17,** 689–698.
20. Zhang, Y. & Crothers, D. M. (2003) *Biophys. J.* **84,** 136–153.
21. Crothers, D. M., Drak, J., Kahn, J. D. & Levene, S. D. (1992) *Methods Enzymol.* **212,** 3–29.
22. Vologodskaia, M. & Vologodskii, A. (2002) *J. Mol. Biol.* **317,** 205–213.
23. Shore, D., Langowski, J. & Baldwin, R. L. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 4833–4837.
24. Hagerman, P. J. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17,** 265–286.
25. Hagerman, P. J. & Ramadevi, V. A. (1990) *J. Mol. Biol.* **212,** 351–362.
26. Taylor, W. H. & Hagerman, P. J. (1990) *J. Mol. Biol.* **212,** 363–376.
27. Moroz, J. D. & Nelson, P. (1998) *Macromolecules* **31,** 6333–6347.
28. Bryant, Z., Stone, M. D., Gore, J., Smith, S. B., Cozzarelli, N. R. & Bustamante, C. (2003) *Nature* **424,** 338–341.
29. Shore, D. & Baldwin, R. L. (1983) *J. Mol. Biol.* **170,** 957–981.
30. Cloutier, T. E. & Widom, J. (2004) *Mol. Cell* **14,** 355–362.
31. Widom, J. (2001) *Q. Rev. Biophys.* **34,** 269–324.
32. Yan, J. & Marko, J. F. (2004) *Phys. Rev. Lett.* **93,** 108108.
33. Wiggins, P. A., Phillips, R. & Nelson, P. C. (2005) *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, in press.
34. Chakrabarti, B. & Levine, A. J. (2004) e-Print Archive, http://xxx.lanl.gov, cond-mat/0411358 v1.
35. Popov, Y. O. & Tkachenko, A. V. (2004) e-Print Archive, http://xxx.lanl.gov, cond-mat/0410591 v1.
36. Guéron, M. & Leroy, J.-L. (1995) *Methods Enzymol.* **261,** 383–413.
37. Ramstein, J. & Lavery, R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7231–7235.
38. Guéron, M. & Leroy, J.-L. (1992) *Nucleic Acids Mol. Biol.* **6,** 1–22.
39. Crick, F. H. C. & Klug, A. (1975) *Nature* **255,** 530–533.
40. Kahn, J. D., Yun, E. & Crothers, D. M. (1994) *Nature* **368,** 163–166.
41. McAteer, K., Ellis, P. D. & Kennedy, M. A. (1995) *Nucleic Acids Res.* **23,** 3962–3966.
42. Han, W., Lindsay, S. M., Dlakic, M. & Harrington, R. E. (1997) *Nature* **386,** 563.
43. Rivetti, C., Guthold, M. & Bustamante, C. (1996) *J. Mol. Biol.* **264,** 919–932.
44. Borowiec, J. A., Zhang, L., Sasse-Dwight, S. & Gralla, J. D. (1987) *J. Mol. Biol.* **196,** 101–111.
45. Hochschild, A. & Ptashne, M. (1986) *Cell* **44,** 681–687.
46. Fitzgerald, D. J. & Anderson, J. N. (1999) *J. Mol. Biol.* **293,** 477–491.
47. Chen, L. & Widom, J. (2005) *Cell* **120,** 37–48.
48. Podtelezhnikov, A. A. & Vologodskii, A. V. (2000) *Macromolecules* **33,** 2767–2771.
49. Schulz, A., Langowski, J. & Rippe, K. (2000) *J. Mol. Biol.* **300,** 709–725.
50. Harmer, T., Wu, M. & Schleif, R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 427–431.
51. Bell, C. E. & Lewis, M. (2000) *Nat. Struct. Biol.* **7,** 209–214.
52. Olson, W. K., Gorin, A. A., Lu, X., Hock, L. M. & Zhurkin, V. B. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 11163–11168.