

# Spontaneous Sharp Bending of Double-Stranded DNA

Timothy E. Cloutier and Jonathan Widom\*

Department of Biochemistry, Molecular Biology  
and Cell Biology  
Northwestern University  
Evanston, Illinois 60208

## Summary

Sharply bent DNA is essential for gene regulation in prokaryotes and is a major feature of eukaryotic nucleosomes and viruses. The explanation normally given for these phenomena is that specific proteins sharply bend DNA by application of large forces, while the DNA follows despite its intrinsic inflexibility. Here we show that DNAs that are 94 bp in length—comparable to sharply looped DNAs in vivo—spontaneously bend into circles. Proteins can enhance the stability of such loops, but the loops occur spontaneously even in naked DNA. Random DNA sequences cyclize  $10^2$ – $10^4$  times more easily than predicted from current theories of DNA bending, while DNA sequences that position nucleosomes cyclize up to  $10^5$  times more easily. These unexpected results establish DNA as an active participant in the formation of looped regulatory complexes in vivo, and they point to a need for new theories of DNA bending.

## Introduction

Experimental (Crothers et al., 1992; Hagerman, 1988; Shore et al., 1981; Taylor and Hagerman, 1990; Vologodskaya and Vologodskii, 2002) and theoretical (Crothers et al., 1992; Hagerman and Ramadevi, 1990; Merlitz et al., 1998; Podtelezhnikov and Vologodskii, 2000; Shimada and Yamakawa, 1984; Zhang and Crothers, 2003) studies led to a unified picture of the behavior of DNA. DNA lengths that are greater than a bending persistence length ( $P$ ,  $\sim 150$  bp or 50 nm) are, on average, spontaneously gently bent and require relatively little force to bend significantly. In contrast, DNA lengths that are shorter than  $P$  are nearly straight and require great force to bend significantly. Thus, a circle of length 500 bp easily forms spontaneously, while a circle of length 100 bp does not.

Other experiments, however, show that sharply bent DNA plays important roles in biology (Schleif, 1992). DNA is sharply bent in prokaryotic regulatory complexes, such as in a 92 bp loop between  $O_3$  and  $O_1$  of the *lac* operon (Oehler et al., 1994), a 93 bp loop between  $\text{BoxE}$  and  $\text{BoxB}$  in the *nagE-B* operon (Plumbridge and Kolb, 1998), or a 113 bp loop between  $O_E$  and  $O_i$  in the *gal* operon (Geanacopoulos et al., 2001) of *Escherichia coli*. In these examples, the sharp DNA looping contributes importantly to gene regulation. Looping allows proteins bound at distant DNA sites to act synergistically, thereby increasing steady-state occupancies (Rippe et

al., 1995) and decreasing the statistical noise in occupancy (Vilar and Leibler, 2003), beyond those that would otherwise obtain. Most of the genomic DNA of eukaryotes is sharply bent by structural proteins in nucleosomes (80 bp superhelical loops) (Richmond and Davey, 2003), which regulate the accessibility and proximity of other DNA functional sites. Double-stranded DNA is also sharply bent inside viruses, where its stiffness contributes importantly to the energetics and dynamics of viral packaging and DNA ejection (Kindt et al., 2001; Purohit et al., 2003).

The explanation normally given for these phenomena is that such sharp DNA bending is achieved in vivo by specific proteins that overwhelm the DNA's inherent inflexibility with large force. In this view, proteins dictate how and when the loops will form, while the DNA follows passively.

We show here that this view is incorrect; rather, we find that DNA is inherently very flexible, and sharp DNA bends occur spontaneously. Our results do not conflict with earlier experiments because this regime of sharp DNA bending was not previously investigated by experiment. Our results do, however, conflict with current theories of DNA bending, which are based on linear elasticity of a continuous material or on harmonic bending of base steps.

Our attention was focused on the problem of sharp DNA bending by the discovery of unexpected  $\sim 10$  bp periodic sequence motifs in nucleosome positioning sequences—DNAs that preferentially form and stabilize nucleosomes (Lowary and Widom, 1998). This discovery highlighted deficiencies in our understanding of the sequence-dependent bendability of DNA itself (Widom, 2001). We therefore sought to analyze the bendability of random DNA sequences and nucleosome positioning sequences in a regime of sharp DNA bending comparable to that of nucleosomes and of DNA loops in gene regulatory complexes.

DNA cyclization assays are particularly well suited to this purpose because they are specifically sensitive to strongly bent conformations, in contrast to most other methods used to analyze DNA mechanics, which are sensitive to the most probable (i.e., weakly bent) conformations (Crothers et al., 1992; Hagerman, 1988; Shore et al., 1981). In the cyclization assay, 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. 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. 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 ( $j$ ). For a given DNA length,  $j$  is a direct measure of flexibility, mathematically related to  $P$  (Crothers et al., 1992; Hagerman and Ramadevi, 1990; Merlitz et al., 1998; Podtelezhnikov and Vologodskii, 2000; Shimada and Yamakawa, 1984).

\*Correspondence: j-widom@northwestern.edu

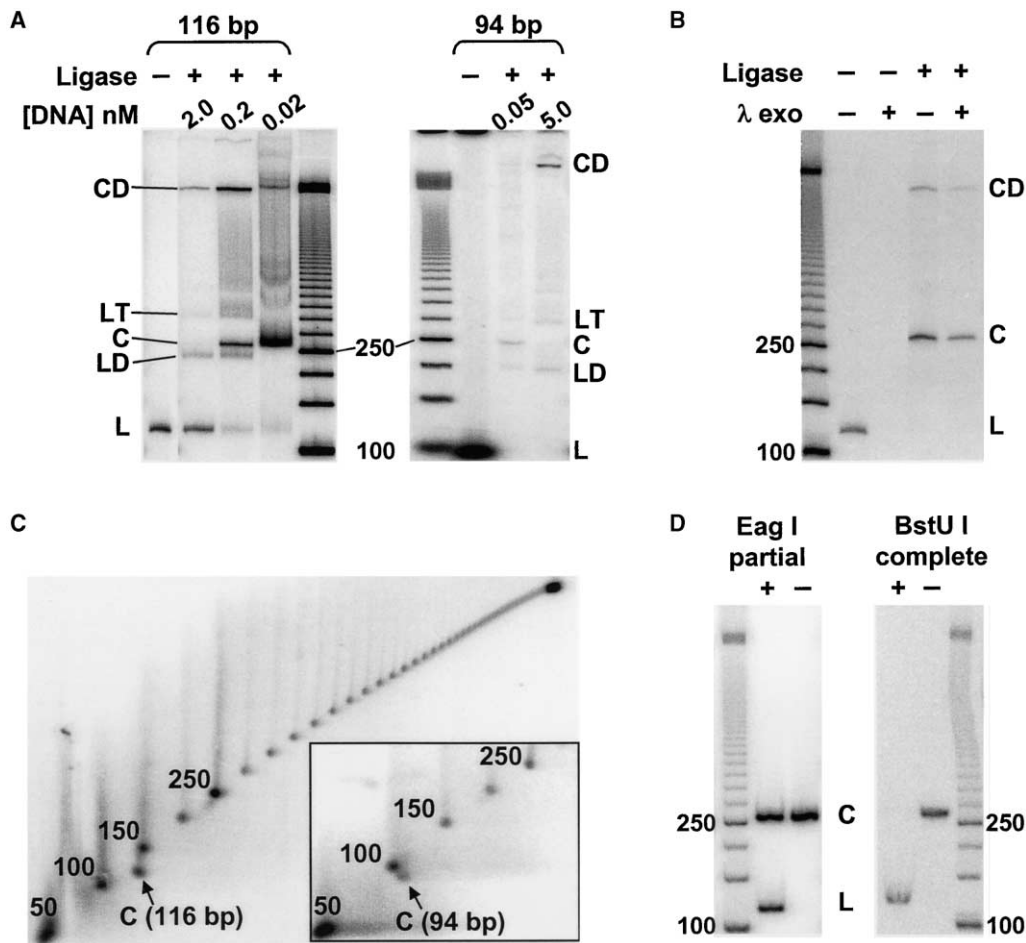


Figure 1. Ninety-Four Base Pair DNAs Cyclize Spontaneously

(A) Cyclization reactions on 116 bp random sequence DNA E6-116 and a 94 bp subfragment, E6-94, ligated at 30°C and analyzed by native PAGE. Linear monomer (L), dimer (LD), and trimer (LT), and circular monomer (C), and dimer (CD) DNA species are indicated. [DNA], in nM, are indicated; [ligase] = 0 or 200 units ml<sup>-1</sup>. Fifty base pair ladders are included as markers. The yield of band C increases at low [DNA], and its mobility lies off the ladder for the linear oligomeric species, as expected for a monomeric DNA circle.

(B)  $\lambda$  exonuclease digestion of 116 bp DNA E8-116, before or after ligation. [DNA] = 20 pM. Unligated (linear) species are rapidly degraded by the exonuclease, while circular species resist degradation.

(C) Two-dimensional agarose gel assay for circular products. Purified band C from a ligation reaction of DNA E6-116 is mixed with 50 bp ladder and analyzed by two-dimensional agarose gel electrophoresis. Inset: Equivalent experiment with DNA E6-94. The 50 bp ladder indicates the behavior of linear DNAs. Circular species run off this diagonal.

(D) Restriction digestion assays for monomer versus oligomeric circular products. Purified band C from a ligation of DNA E6-116 is shown before and after partial digestion with EagI or complete digestion with BstUI, which cleaves nearby the EagI site. Partial digestion of oligomeric circles by EagI yields oligomeric linear species, yet none are detected from purified band C. Oligomeric circles will include head-head dimers. Complete digestion of oligomeric 116 bp circles by BstUI therefore yields a short fragment lost from the gel and an ~200 bp fragment; however, no such product is detected from purified band C.

Importantly, ligase does not contribute to the DNA cyclization process itself; rather, DNA conformational changes that bring the two cohesive ends into transient base-paired contact occur spontaneously in a ligase-independent process (Shore et al., 1981).

## Results and Discussion

We began by investigating the cyclizability of 116 bp DNAs (11 helical repeats) having cohesive ends generated by restriction endonuclease digestion but soon learned that DNAs as short as 94 bp (nine helical repeats) cyclized efficiently enough to allow quantitative analysis (Figure 1A). Ligation of a random sequence 116 bp DNA

(E6-116) at 2 nM DNA reveals a ladder of fragments having mobilities expected for linear monomer, dimer, and trimer (L, LD, and LT, respectively), together with a faint band (C) whose mobility is off the linear-oligomer ligation ladder. The same DNA ligated at concentrations of 0.2 nM and 20 pM reveals analogous products but in very different relative amounts, such that C becomes the predominant product. Analogous results are obtained with a 94 bp subfragment of this DNA (E6-94).

These and other results establish that C is a monomeric circle. C runs off the ladder of linear ligation products because of its topology, and its production is favored at very low concentrations by mass action because it is unimolecular. Consistent with its identification as a

monomer circle, C is resistant to digestion by bacteriophage  $\lambda$  exonuclease (Figure 1B), and it runs off the diagonal in a topology-sensitive two-dimensional agarose gel assay (Figure 1C). Two different assays based on partial or complete restriction digestion of purified circular products (Figure 1D) provide further proof that C is a monomeric circle, not an oligomer. Parallel experiments establish band CD to be a circular dimer (data not shown).

Since the DNAs used in the cyclization reactions are produced by *E. coli*, not by chemical synthesis, and essentially all of the molecules in the sample are converted to monomer circle at low [DNA] (Figure 1A), we conclude that the ability to spontaneously form sharp loops is an inherent property of natural DNA—not an artifact from a small fraction of degraded species.

Our results establish that random sequence DNAs as short as 94 bp spontaneously bend sharply into full circles. This conclusion is striking because such short DNAs were previously expected to be incapable of producing circles in detectable amounts. When [DNA] is  $\ll j$ , monomer circles will predominate—as observed for bands C—while, when [DNA] is  $> j$ , dimeric and higher order oligomers will predominate, as observed for bands LD and LT (Shore et al., 1981). However, because of the presumed inflexibility of DNAs shorter than P, the  $j$  factor expected for a 94 bp DNA is extremely low, only 200 or 10 fM for the Shimada-Yamakawa (Shimada and Yamakawa, 1984) or Zhang-Crothers (Zhang and Crothers, 2003) theories, respectively, whereas the ligations in Figure 1A were carried out in the pM–nM range. The detection of such short circular DNA products in these ligation reactions means that sharp looping of these DNAs occurs orders of magnitude more easily than predicted from our current understanding.

We used quantitative kinetic assays to measure the  $j$  factors for several short DNAs. Nucleosome positioning sequences can demonstrate enhanced flexibility (Roychoudhury et al., 2000). We therefore included two such sequences in our analysis, along with three random sequence DNAs, to test whether cyclizability in this regime of sharp DNA bending is significantly dependent on specific DNA sequences. We analyzed 94 bp DNAs derived from the natural sea urchin 5S rDNA nucleosome positioning sequence (Simpson and Stafford, 1983) and from a selected high-affinity nonnatural nucleosome positioning sequence (Lowary and Widom, 1998).

We first tested whether these reactions occur in the rapid preequilibrium regime, as required for the kinetic analysis (Shore et al., 1981). Ligation reactions were set up using a low [DNA] such that monomer circle products predominate, and the rate constant for monomer circle formation was measured as a function of [ligase] over the range 5–600 units  $\text{ml}^{-1}$ . The observed rate constant is first order in [ligase] over this full range (Figure 2A), confirming that our subsequent assays are safely in the rapid preequilibrium regime.

To measure  $j$  factors, the ratio of monomer circle to dimeric species is determined as a function of time in a ligation reaction (Taylor and Hagerman, 1990). Typical data for a random sequence 94 bp DNA and a nucleosome positioning sequence are illustrated in Figures 2B and 2C). Significant fractions of the DNA are converted over time to monomeric circle, despite the large DNA

concentrations used (0.04 and 0.4 nM for the random sequence and nucleosome positioning sequence, respectively). Quantitative values of  $j$  are obtained from these data by linear extrapolation of the ratio of monomeric to dimeric products back to zero time (Figure 2D). The resulting  $j$  factors for these and several other DNAs are plotted as a function of DNA length in Figures 3 and 4.

The five different 94 bp DNAs exhibit an  $\sim 25$ -fold range of  $j$  factors. These differences in  $j$  factor for the differing 94 bp DNAs are not simply attributable to a known dependence of  $j$  on DNA helical twist, which can depend on DNA sequence, because 93 bp and 95 bp variants of these sequences had lower  $j$  factors than do the 94 bp fragments (Figure 3). Instead, these data imply that DNA sequence is an important determinant of the ability of DNA molecules to bend sharply.

All five 94 bp DNAs have  $j$  factors that are orders of magnitude greater than predicted based on the current theories of DNA cyclization (Figure 4A). The three 94 bp random sequence DNAs have  $j$  factors that are  $10^2$ - to  $10^3$ -fold greater than predicted by the classical Shimada-Yamakawa theory (Shimada and Yamakawa, 1984), and  $10^3$ - to  $10^4$ -fold greater than predicted by the newer Zhang-Crothers theory (Zhang and Crothers, 2003); the 94 bp nucleosome positioning sequences have  $j$  factors that are up to  $10^5$ -fold greater than predicted.

The observed disparity between observed and predicted  $j$  factors is not attributable to a problem with the measurements themselves. The cyclization technique is a standard experiment, widely used in studies of DNA bending in a less sharply bending regime. Moreover, our results obtained for generic sequence 200, 322, and 326 bp fragments agree closely with earlier experimental studies of cyclization on diverse generic DNA sequences in that regime of gentler bending (Crothers et al., 1992; Hagerman, 1988; Shore et al., 1981; Taylor and Hagerman, 1990; Vologodskaja and Vologodskii, 2002), and they agree relatively well with the theoretical curves (Figure 4A), using the canonical value of  $P \sim 50$  nm. That is, the longer generic fragments studied here behave in full accord with expectation, assuming they have the canonical P.

In order to fit the 94 bp generic sequence data to the Shimada-Yamakawa theory (Figure 4A), a persistence length of 40 nm is required (results not shown). Such a small value of P is lower than is justified by any experiment. Moreover, whereas a single chemically random sequence might (unexpectedly) have such an unprecedented property, it is not plausible that all three of them do. We conclude that specific sequence effects do not account for the discrepancy between these theories and experiment for these generic sequence 94 bp DNAs. Instead, we conclude that the observed easy bendability of generic sequence 94 bp DNAs points to deficiencies in both theories for this sharp bending regime.

The theoretical predictions of Figure 4A assume that ligation of DNA ends can only proceed if the two ends approach with zero tangent angles. This condition may be overly restrictive. The Shimada-Yamakawa theory allows prediction also for the opposite limit, in which ligation could hypothetically proceed with no constraints at all on angles of approach of the two DNA ends. The predictions of this alternate theory are shown

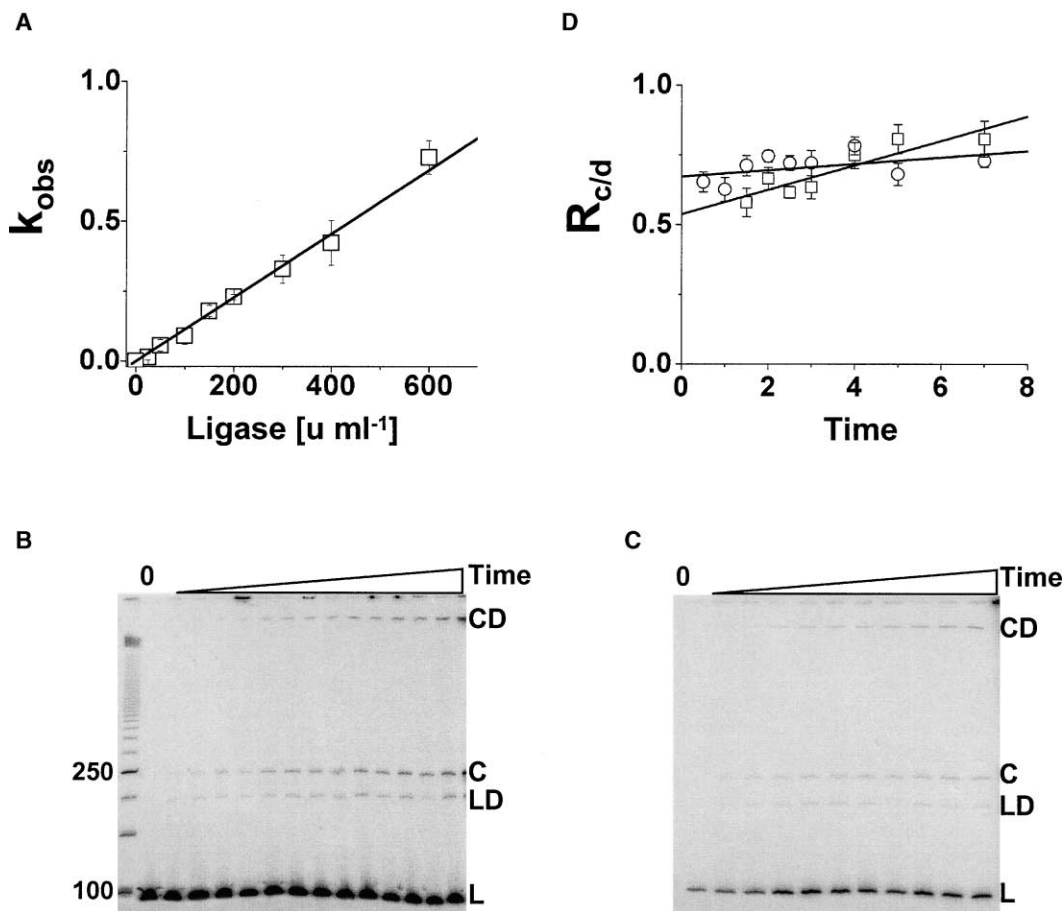


Figure 2. Cyclization Kinetics of 94 bp DNAs

(A) Test for rapid preequilibrium. The observed rate constant for accumulation of monomer circles from a ligation reaction with E6-94 DNA at 50 pM, 30°C, is plotted as a function of [ligase]. The rate of cyclization is strictly first order in [ligase] over this wide range of [ligase]. (B and C) Phosphorimages of ligation reactions on (B) DNA E6-94 ([DNA] = 40 pM; [ligase] = 250 units ml<sup>-1</sup>) and (C) DNA 601TA-94 ([DNA] = 400 pM; [ligase] = 150 units ml<sup>-1</sup>). (D) Quantitative analysis of phosphorimages. The ratio of counts in monomer circle (C) to dimeric species (LD + CD) is shown. □, E6-94; ○, 601TA-94. Lines through the data show extrapolations to time = 0.

in Figure 4B. Using  $P = 50$  nm, the predicted  $j$  factor comes closer to, but still falls short of, the data for 94 bp DNAs, by factors of 10- to 100-fold, while at the same time overestimating the  $j$  factors for the longer fragments. Using a larger value of  $P = 55$  nm (larger than suggested by any experiment for any generic sequence DNA in these solution conditions), this theory roughly fits the  $j$  factors of longer DNAs but now falls even further short of the experimental results for the 94 bp fragments. That is, this theory too has the wrong functional form: it does not describe the data. Moreover, it falls short of the 94 bp data despite being overgenerous in its assumption that ligation could proceed with no angular constraints at all. We conclude that the facile cyclizabilities of the diverse random sequence 94 bp DNAs studied here cannot be explained by existing theories of DNA cyclization.

Two lines of evidence imply that the facile cyclizabilities of 94 bp DNAs measured in this study represent inherent bending flexibility, but not permanent bendedness. First, the regime studied here involves far sharper bending than that of the most-bent of all sequences,

which are phased A tracts (Beutel and Gold, 1992); these require  $\sim 230$  bp or  $\sim 400$  bp (for A6 or A4 tracts, respectively) for a complete 360° bend (Barbic et al., 2003). Thus, permanent bendedness, on its own, cannot account for the ability of any of these 94 bp DNAs to cyclize easily. Second, a recent study reveals that the cyclizabilities of random sequence  $\sim 200$  bp DNAs and inherently straight DNAs are essentially identical, implying that the random sequence DNAs have negligible permanent bendedness (Vologodskaja and Vologodskii, 2002). Our sequences E6-94, E8-94, and E13-94 are (distinct) chemically random sequences; moreover, the variance of bendedness in random sequence 94 bp DNAs will be lower even than in the random sequence 200 bp DNAs. Thus, it is unlikely that any of our 94 bp random DNAs have significant permanent bendedness, and it is even more unlikely that all three of them do at once. We conclude for these reasons that permanent bendedness does not plausibly account for the disparity between theory and experiment.

The enhanced cyclizability of the nucleosome positioning sequences relative to random DNA sequences

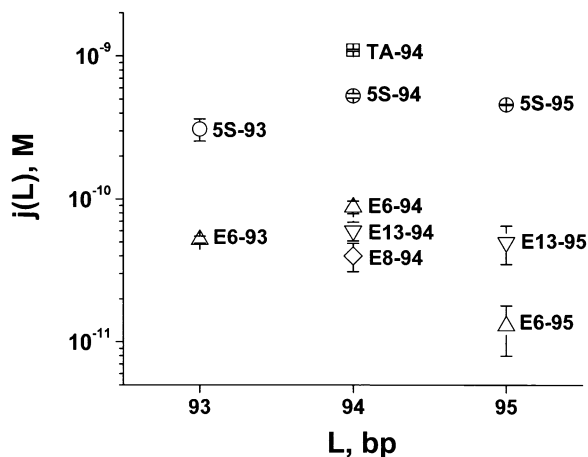


Figure 3. Measured  $j$  Factors Depend Strongly on DNA Sequence Datapoints, means, and standard deviations from measurements on five different 94 bp DNAs plus 93 and 95 bp variants of some of these DNAs. Note the log scale of  $j$  factor. The  $j$  factors of 94 bp fragments exceed those of the 93 and 95 bp variants, implying that differences in intrinsic helical twist between differing 94 bp DNA sequences do not account for the observed differences in  $j$  factor.

suggests that their cyclizability might contribute importantly to their high affinity for histones and their nucleosome positioning power. Ninety-four base pairs suffices to cover the DNA binding surface of the histone H3<sub>2</sub>H4<sub>2</sub> tetramer, which protects 70–80 bp DNA (Richmond and Davey, 2003). We used a competitive reconstitution experiment (Lowary and Widom, 1997) to measure the relative affinities of these same five DNAs for the histone tetramer (Alilat et al., 1999) (Figure 5A). The five DNAs

investigated reveal an  $\sim 25$ -fold range of affinities. (The corresponding free energies are shown in Figure 5B). The free energies of cyclization of the same 94 bp DNAs are calculated from the measured  $j$  factors (Figure 3). The difference free energies for histone-DNA interactions in tetrasomes for the five 94 bp DNAs are plotted against their free energies of cyclization in Figure 5B.

The free energy of histone-DNA interactions in tetrasomes is strongly correlated with the free energy of cyclization (linear slope = 0.97, correlation coefficient  $r = 0.98$ , probability of achieving this correlation coefficient by random chance  $< 0.005$ ). The linear slope is  $\sim 1$  because the  $\sim 25$ -fold range in relative cyclizability of the DNAs is matched by an  $\sim 25$ -fold range in their affinities for wrapping by histones in tetrasomes. Thus, the results in Figure 5 imply that sequence-dependent differences in the cost of DNA bending dominate the interaction of differing DNA sequences with the histone tetramer, while DNA twisting or specific protein-DNA interactions contribute little. We conclude that DNA sequences that are especially flexible for sharp looping will have high affinity for histones in nucleosome reconstitution and thus will act as nucleosome positioning sequences.

Our findings have several important ramifications. First, they show that DNA is highly flexible for sharp bending. Circles as short as 94 bp form spontaneously. This finding differs from the phenomenon of protein-induced sharp DNA kinking because the DNA looping process studied here occurs spontaneously, with no contribution from other proteins.

This finding of spontaneous sharp DNA looping provides a simpler explanation for a paradox in gene regulation: DNA looping is believed to contribute to gene regu-

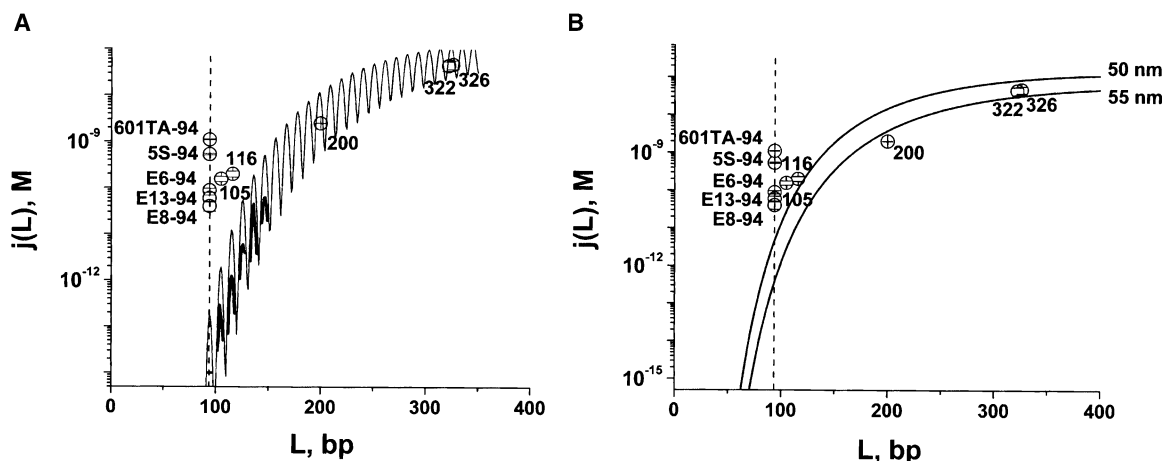


Figure 4. Measured  $j$  Factors Greatly Exceed Theoretical Predictions

(A) Datapoints, means, and standard deviations from measurements on five different 94 bp DNAs plus 105, 116, 200, 322, and 326 bp DNAs. The dashed vertical line indicates 94 bp. Note that the log scale for  $j$  spans over seven orders of magnitude. The thin line represents the prediction of the Shimada-Yamakawa theory (Shimada and Yamakawa, 1984) for dependence of  $j$  factor as a function of DNA length (in bp), assuming a standard DNA persistence length of 51 nm; computer program courtesy of Dr. Y. Zhang. The bold line represents the prediction of the Zhang-Crothers theory (Zhang and Crothers, 2003); calculations courtesy of Dr. Y. Zhang.  $P = 51$  nm corresponds to the latest value of persistence length for generic sequence DNAs (Zhang and Crothers, 2003) for the high cation-concentration regime (Hagerman, 1988), as appropriate for the solution conditions used in these cyclization reactions.

(B) Theoretical prediction for cyclization with no constraints on angles of end-end joining (Shimada and Yamakawa, 1984). Predictions are shown for the canonical persistence length (50 nm) and for a larger  $P$  (55 nm, not justified by any experiment) that better fits the experimental results for the longer DNAs but falls further short of the data at 94 bp.

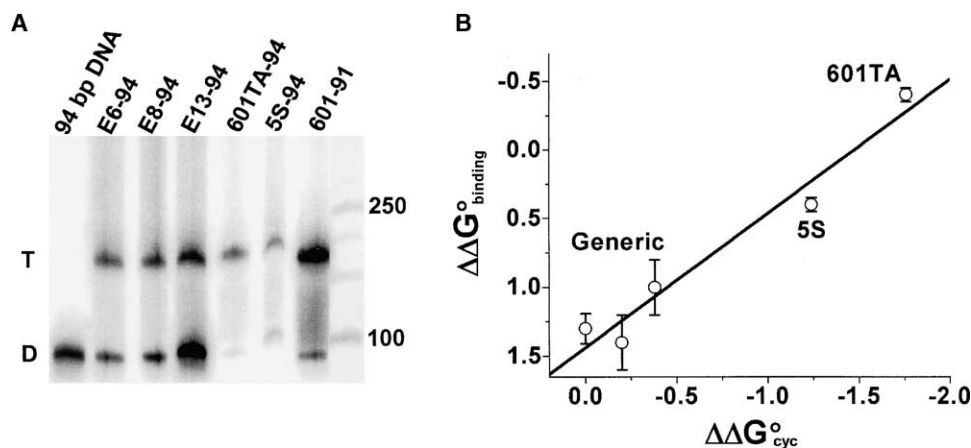


Figure 5. The Free Energy of DNA Wrapping on the Histone H<sub>3</sub>H<sub>4</sub><sub>2</sub> Tetramer Is Strongly Correlated with the Free Energy of Cyclization

(A) Native gel analysis of competitive tetrasome reconstitution assays. Free 94 bp E6-94 DNA and a 50 bp ladder are included as size markers. Reconstitution experiments with histone H<sub>3</sub>H<sub>4</sub><sub>2</sub> tetramer and radiolabeled tracer DNAs E6-94, E8-94, E13-94, 5S-94, 601TA-94, and 601TA-91, respectively. 601TA-91 tracer is included as a reference affinity standard to relate these measurements to others in our laboratory. Mobilities of free DNA (D) and tetrasomes (T) are indicated.

(B) Difference free energies for binding of 94 bp DNAs to H<sub>3</sub>H<sub>4</sub><sub>2</sub> tetramer are plotted against the difference free energies of cyclization for these same DNAs (see Experimental Procedures). The line illustrates the least-squares fit to the data.

lation by increasing the local concentrations of gene regulatory proteins, yet the presumed inflexibility of DNA for sharp bending prohibited such a role (Rippe et al., 1995). The previous explanation of this paradox was that sharp DNA looping is achieved *in vivo* by specific proteins that overwhelm the DNA's inherent inflexibility (Rippe et al., 1995). Our results show that the needed sharp DNA loops occur spontaneously. Further, they imply that spontaneous sharp DNA looping will greatly increase the effective concentration of a regulatory protein bound at a site  $\sim$ 94 bp away along the DNA to  $>1 \mu\text{M}$  (see Experimental Procedures). Since this effective concentration greatly exceeds typical bulk concentrations of regulatory proteins (Rippe et al., 1995; Menlitz et al., 1998), spontaneous sharp DNA looping can greatly increase both occupancies and rates of regulatory protein binding.

Second, the spontaneous looping of 94 bp DNAs and their unexpectedly large *j* factors mean that cyclized or looped DNA conformations can be stabilized with relatively little cost in energy. The equilibrium constant and corresponding free energy cost for looping a given DNA to a 10 nm end-to-end capture radius can be estimated from the measured *j* factors (Merlitz et al., 1998). The *j* factor of 1.1 nM measured for the 94 bp DNA 601TA-94 equates to a cost of  $\sim 8 \text{ kcal mol}^{-1}$  or less. Such a modest cost could be accommodated by a protein that made just a few good hydrogen bonds or ionic interactions with the DNA. On this basis we predict that there exist protein-DNA complexes having sharply bent DNA stabilized by just a few protein-DNA contacts.

Third, the large disparity between theoretical prediction and our experimental observations implies either that random sequence DNAs have significant permanent bendedness, or else that DNA bending in this tight bending regime is not well described by existing theories. The available evidence argues against the permanent bendedness interpretation; thus, our findings point to a

need for a new theory for sharp DNA bending. The greatly enhanced *j* factors, compared to expectation, suggest that, as DNA is asked to bend increasingly sharply, it finds pathways for bending that greatly reduce the energetic cost relative to that assumed by current theory. Distinctive local structural distortions of the DNA observed in the high resolution crystallographic structure of the nucleosome (Richmond and Davey, 2003) may already be providing glimpses of such behavior. More broadly, the discrepancy between theory and experiment could represent a breakdown of the linear approximation to mechanical elasticity theory. Alternatively, it could arise from inhomogeneities in DNA sequence, or it could indicate that an entirely distinct mode of DNA bending, such as localized buckling, dominates sharp looping. Structural studies of the sharply looped DNA will be valuable.

Fourth, we predict that the greatly enhanced observed *j* factors will be accompanied by comparably greatly enhanced rates of cyclization (P. Wiggins, personal communication; see Experimental Procedures). Brownian dynamics simulations on DNA cyclization, which were tuned to fit the linear elastic rod model for DNA bending, suggest that the spontaneous looping rate of DNAs shorter than *P* can approach the cell lifetime (Podtelezhnikov and Vologodskii, 2000). Our results imply that such loops will form much faster. A rapid rate of looping is needed for looping to contribute significantly to a reduction in stochastic noise in regulatory systems (Vilar and Leibler, 2003).

Fifth, the 25-fold or greater difference in cyclizability between differing DNA sequences suggests that the sequences of looped DNAs are likely to be tuned in evolution, in concert with the tuning of protein-DNA and protein-protein binding affinities, such that the stability of a complex as an integrated whole falls within a desired range. On this basis we predict the existence of a new class of gene regulatory mutations, in which changes

to the sequence of the looped DNA that fall outside any protein binding site should influence activation/repression importantly through their effects on the intrinsic bendability of the looped DNA.

Sixth, differences in the bendability of differing DNA sequences account quantitatively for the phenomenon of nucleosome positioning—that is, for the preference of nucleosomes to form on one DNA sequence versus another.

Taken together, these findings demonstrate an active and sequence-specific role for DNA in the establishment and maintenance of regulatory and structural complexes involving sharp DNA loops.

### Experimental Procedures

#### 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 *EagI* sites. Sequences E6-94, E8-94, and E13-94, and their 93 or 95 bp variants, are internal subfragments derived from their 116 bp parents by PCR. 5S-94 and its 93 and 95 bp variants derive from the sea urchin 5S rRNA gene nucleosome positioning sequence (Simpson and Stafford, 1983). 601TA-91 and 601TA-94 derive from the selected nonnatural nucleosome positioning sequence 601 (Lowary and Widom, 1998). These DNAs were isolated from purified plasmid by restriction enzyme digestion with *EagI*. The 322 bp fragment was obtained by *EaeI* digestion of plasmid pRS304 (this leaves *EagI* cohesive ends); the 326 bp fragment was obtained by *Apo I* digestion of plasmid pYES2 (which leaves *EcoRI* cohesive ends, useful for comparison with earlier cyclization studies [Shore et al., 1981]). These DNA fragments were purified by agarose gel electrophoresis. The 200 bp fragment was provided in a plasmid by Prof. A. Vologodskii and was prepared and analyzed as described (Vologodskia and Vologodskii, 2002). Concentrations of purified DNA stocks were determined from their absorbance at 260 nm. DNAs were 5'-end labeled with  $\gamma$ -<sup>32</sup>P-ATP and T4 polynucleotide kinase, followed by a chase with 1 mM unlabeled ATP. DNA sequences are available upon request.

#### Ligation Reactions, Exonuclease Digestions, Two-Dimensional Gels

Ligation reactions were carried out using T4 DNA ligase in standard buffer (New England Biolabs). Reactions with DNAs having *EagI* ends were performed at 30°C, while those with *ApoI* ends were performed at 20°C. Products were analyzed on polyacrylamide gels and quantified by phosphorimager.  $\lambda$  exonuclease digestions were carried out as recommended (NEB) using 5 units ml<sup>-1</sup> enzyme and [DNA] = 100 pM. Two-dimensional agarose gel electrophoresis was carried out using 2.5% agarose in standard 0.5 × TAE buffer. Radiolabeled samples were mixed with radiolabeled 50 bp ladder. The first dimension was prepared with no ethidium bromide. Gels of 20 cm × 20 cm were run at 250 volts for 2 hr; they were then soaked in 10  $\mu$ g ml<sup>-1</sup> ethidium bromide for 1 hr, rotated 90° clockwise, and subjected to a second dimension of electrophoresis at 275 volts for 1–2 hr. Gels were visualized by phosphorimager. For each sample a control gel with only the 50 bp ladder was run to confirm which spot(s) derived from the test sample itself.

#### Quantitative Kinetic Assays

Quantitative cyclization assays were analyzed as described (Taylor and Hagerman, 1990). Counts in monomer circle product (band C) and in linear and circular dimeric species (bands LD + CD) were obtained by phosphorimager analysis of native polyacrylamide gels. (In this tight bending regime, the *j* factors for linear dimers are greater than for linear monomer, hence much of the linear dimer cyclizes.) The measured ratio of these quantities ( $R_{c/d}$ ) was linearly extrapolated to time  $t = 0$ , yielding  $R_0$ , and *j* factors were calculated as  $j = 4 M_0 R_0$ , where  $M_0$  is [DNA] at  $t = 0$ . For analysis of the [ligase]-dependence of the observed rate, ligation reactions contained 50

pM DNA E6-94. The apparent first-order cyclization rate constant was determined from the counts incorporated into monomer circle over 0–20 min.

#### Histone Binding Affinity Measurements

Competitive reconstitution experiments were performed as described (Lowary and Widom, 1997, 1998) except using purified H3<sub>2</sub>H4<sub>2</sub> tetramer instead of histone octamer. Samples were recovered and analyzed on native 5% polyacrylamide gels containing 0.3 × TBE, and the gels were quantified by phosphorimager.  $K_{eq}$ 's were calculated as ratios of background subtracted counts in tetrasome complexes to counts in free DNA.  $\Delta G^{\circ}_{binding} = -RT \ln(K_{eq})$ , where *R* is the gas constant and *T* the absolute temperature (277 K).  $\Delta\Delta G^{\circ}_{binding}$  values are differences in  $\Delta G^{\circ}$ 's between a given sample and the 601TA-91 reference molecule, measured in an identical competitive reaction.  $\Delta\Delta G^{\circ}_{cyclization}$  values are calculated from the measured *j* factors for cyclization, using the approximation (Kahn and Crothers, 1992; Merlitz et al., 1998; Podtelezhnikov and Vologodskii, 2000) that  $j_{sample}/j_{reference} = K_{c,sample}/K_{c,reference}$ , where  $K_c$  represents the cyclization equilibrium constants for a given sample and an arbitrary reference species. Thus,  $\Delta\Delta G^{\circ}_{cyclization} = -RT \ln(K_{c,sample}/K_{c,reference})$ .  $\Delta\Delta G^{\circ}_{cyclization}$  values are reported relative to an arbitrary zero.

#### Numerical Evaluations of the Shimada-Yamakawa and Zhang-Crothers Theories

The numerical evaluations of the Zhang-Crothers theory were carried out for us by Dr. Y. Zhang exactly as described (Zhang and Crothers, 2003). The algorithm fails to converge for some short DNA lengths, hence the gaps in these numerical predictions. The predictions for the Shimada-Yamakawa theory in Figure 4A (Shimada and Yamakawa, 1984) were evaluated using a computer program provided to us by Dr. Zhang, as described (Zhang and Crothers, 2003). The predictions for the variant Shimada-Yamakawa theory allowing for cyclization without constraints on end-end angle (Figure 4B) were evaluated from the analytical expression (Shimada and Yamakawa, 1984) as described (Rippe et al., 1995).

#### Effective Concentration from a 94 bp Loop

We fit our measured *j* factors for the 94 bp random sequence DNAs to the Shimada-Yamakawa theory for ligase-mediated DNA cyclization (Figure 4A), with the bending persistence length as an adjustable parameter, resulting in an apparent persistence length of ~40 nm for these DNAs. Protein-protein interactions mediated by looped DNA take place with fewer constraints on the angles of approach between the two DNA ends because of flexibility in the proteins themselves. (This stands in contrast to ligase-mediated DNA cyclization, in which the angle of approach between the two DNA ends is likely to be strongly constrained.) A 40 nm persistence length equates to a *j* factor of ~1 nM using the angle-independent theory for loop formation (Rippe et al., 1995; Shimada and Yamakawa, 1984). Protein-protein interactions mediated by DNA looping occur with the two DNA target sites separated by some distance, e.g., 10 nm, in contrast to ligase-mediated DNA cyclization which requires direct end-to-end contact. Brownian dynamics simulations show that looping of a 120 bp DNA to a separation of 10 nm occurs with ~200-fold greater probability than looping to contact (0 separation) (Merlitz et al., 1998); for 94 bp DNAs the effects will be much greater still. Thus, an estimated angle-independent *j* factor of ~1 nM for a random sequence 94 bp DNA translates to a >1  $\mu$ M effective concentration from looping.

#### Rate of Cyclization

The elastic strain energy that is released by a small increase in end-to-end separation, sufficient to break the base pairs at the cohesive ends, is small in comparison to the energy of thermal fluctuations,  $k_B T$  ( $k_B$  is Boltzmann's constant and *T* the absolute temperature), even using the existing theory for this elastic cost, which our results show is plainly an overestimate. Therefore, elastic strain should contribute little to the rate of end-to-end dissociation, while it directly affects the probability and hence rate constant of end-to-end approach.

## Acknowledgments

We thank A. Matouschek, A. Mondragon, R. Phillips, O. Uhlenbeck, and P. Wiggins for discussion and critical comments on the manuscript, and A. Vologodskii for the 200 bp DNA fragment. We gratefully acknowledge Y. Zhang for the calculations of the Zhang-Crothers theory in Figure 4A and for the computer program used for numerical evaluation of the Shimada-Yamakawa theory in Figure 4A. We acknowledge the use of instruments in the Keck Biophysics Facility at Northwestern University. This work was supported by NIH grants R01-GM054692 and R01-GM058617 to J.W.

Received: December 30, 2003

Revised: March 24, 2004

Accepted: March 28, 2004

Published: May 6, 2004

## References

- Allat, M., Sivolob, A., Revet, B., and Prunell, A. (1999). Nucleosome dynamics. Protein and DNA contributions in the chiral transition of the tetrasome, the histone (H3-H4)<sub>2</sub> tetramer-DNA particle. *J. Mol. Biol.* **291**, 815–841.
- Barbic, A., Zimmer, D.P., and Crothers, D.M. (2003). Structural origins of adenine-tract bending. *Proc. Natl. Acad. Sci. USA* **100**, 2369–2373.
- Beutel, B.A., and Gold, L. (1992). In vitro evolution of intrinsically bent DNA. *J. Mol. Biol.* **228**, 803–812.
- Crothers, D.M., Drak, J., Kahn, J.D., and Levene, S.D. (1992). DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* **212**, 3–29.
- Geanacopoulos, M., Vasmatzis, G., Zhurkin, V.B., and Adhya, S. (2001). Gal repressosome contains an antiparallel DNA loop. *Nat. Struct. Biol.* **8**, 432–436.
- Hagerman, P.J. (1988). Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 265–286.
- Hagerman, P.J., and Ramadevi, V.A. (1990). Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. I. Computational analysis. *J. Mol. Biol.* **212**, 351–362.
- Kahn, J.D., and Crothers, D.M. (1992). Protein-induced bending and DNA cyclization. *Proc. Natl. Acad. Sci. USA* **89**, 6343–6347.
- Kindt, J., Tzllil, S., Ben-Shaul, A., and Gelbart, W.M. (2001). DNA packaging and ejection forces in bacteriophage. *Proc. Natl. Acad. Sci. USA* **98**, 13671–13674.
- Lowary, P.T., and Widom, J. (1997). Nucleosome packaging and nucleosome positioning of genomic DNA. *Proc. Natl. Acad. Sci. USA* **94**, 1183–1188.
- Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42.
- Merlitz, H., Rippe, K., Klenin, K.V., and Langowski, J. (1998). Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation. *Biophys. J.* **74**, 773–779.
- Oehler, S., Amouyal, M., Kolkhof, P., Wilcken-Bergmann, B., and Muller-Hill, B. (1994). Quality and position of the three lac operators of *E. coli* define efficiency of repression. *EMBO J.* **13**, 3348–3355.
- Plumbridge, J., and Kolb, A. (1998). DNA bending and expression of the divergent nagE-B operons. *Nucleic Acids Res.* **26**, 1254–1260.
- Podtelezchnikov, A.A., and Vologodskii, A.V. (2000). Dynamics of small loops in DNA molecules. *Macromolecules* **33**, 2767–2771.
- Purohit, P.K., Kondey, J., and Phillips, R. (2003). Mechanics of DNA packaging in viruses. *Proc. Natl. Acad. Sci. USA* **100**, 3173–3178.
- Richmond, T.J., and Davey, C.A. (2003). The structure of DNA in the nucleosome core. *Nature* **423**, 145–150.
- Rippe, K., von Hippel, P.H., and Langowski, J. (1995). Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.* **20**, 500–506.
- Roychoudhury, M., Sitali, A., Lapham, J., and Crothers, D.M. (2000). Global structure and mechanical properties of a 10-bp nucleosome positioning motif. *Proc. Natl. Acad. Sci. USA* **97**, 13608–13613.
- Schleif, R. (1992). DNA looping. *Annu. Rev. Biochem.* **61**, 199–223.
- Shimada, J., and Yamakawa, H. (1984). Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules* **17**, 689–698.
- Shore, D., Langowski, J., and Baldwin, R.L. (1981). DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. USA* **78**, 4833–4837.
- Simpson, R.T., and Stafford, D.W. (1983). Structural features of a phased nucleosome core particle. *Proc. Natl. Acad. Sci. USA* **80**, 51–55.
- Taylor, W.H., and Hagerman, P.J. (1990). Application of the method of T4 phage DNA ligase-catalyzed ring-closure to the study of DNA structure. II. Na-Cl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* **212**, 363–376.
- Vilar, J.M.G., and Leibler, S. (2003). DNA looping and physical constraints on transcription regulation. *J. Mol. Biol.* **331**, 981–989.
- Vologodskaya, M., and Vologodskii, A. (2002). Contribution of the intrinsic curvature to the measured DNA persistence length. *J. Mol. Biol.* **317**, 205–213.
- Widom, J. (2001). Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* **34**, 269–324.
- Zhang, Y., and Crothers, D.M. (2003). Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* **84**, 136–153.