Cyclization of short DNA fragments and bending fluctuations of the double helix

Quan Du, Chaim Smith, Nahum Shiffeldrim, Maria Vologodskaia, and Alexander Vologodskii*

Department of Chemistry, New York University, New York, NY 10003

Edited by Nicholas R. Cozzarelli, University of California, Berkeley, CA, and approved March 8, 2005 (received for review February 4, 2005)

Cloutier and Widom [Cloutier, T. E. & Widom, J. (2004) *Mol. Cell* 14, 355–362] recently reported that the cyclization efficiency of short DNA fragments, about 100 bp in length, exceeds theoretical expectations by three orders of magnitude. In an effort to resolve this discrepancy, we tried modifying the theory. We investigated how the distribution of the angles between adjacent base pairs of the double helix affects the cyclization efficiency. We found that only the incorporation of sharp kinks in the angle distribution provides the desired increase of the cyclization efficiency. We did not find a model, however, that fits all cyclization data for DNA fragments of different lengths. Therefore, we carefully reinvestigated the cyclization of 100-bp DNA fragments experimentally and found their cyclization efficiency to be in remarkable agreement with the traditional model of DNA bending. We also found an explanation for the discrepancy between our results and those of Cloutier and Widom.

DNA bending | DNA cyclization | DNA flexibility | DNA kinks

The flexibility of the DNA double helix is extremely important for its functioning and has been studied for nearly 50 years (reviewed in ref. 1). It is now generally accepted that the major mechanisms of DNA bending are the small fluctuations between the planes of adjacent base pairs (2). Correspondingly, DNA conformational properties are well described by the worm-like chain (WLC) model (3). Calculations based on this model accurately reproduce experimental data on hydrodynamic properties of DNA molecules (4–6), equilibrium distributions of topological states (7–12), and light and neutron scattering data on supercoiled DNA (13–16). One of the most impressive tests of the WLC model was the single-molecule measurement of DNA extension under the action of a force applied to the ends of the double helix (17). Initially, Bustamante and coworkers (17) tried to fit the experimental results by applying the theory for a freely jointed chain, the only theory being considered at that time, but found a large discrepancy between theoretical and experimental results for large extensions of the molecule. It was soon understood that the WLC model gives a different result in the case of large extensions. The force-extension dependence for the WLC was found to be in excellent agreement with the experimental data (18, 19).

Theoretical analysis and computations based on the WLC also accurately predict the cyclization efficiency of small DNA fragments, 200–350 bp in length (20–24). For even shorter fragments, this model predicts a very low efficiency of cyclization, and until recently, quantitative measurements of this efficiency had not been attempted. Therefore, it was a complete surprise when Cloutier and Widom (hereafter referred as CW) reported that DNA fragments of ≈100 bp in length are cyclized several orders of magnitude more efficiently than the current theory predicts (25). It is known, on the other hand, that microscopically different models of polymer chains can give similar or even identical results for many properties. So, even excellent agreement between a model prediction and a set of experimental data does not necessary mean that the model provides an adequate microscopic description of polymer flexibility. The data of CW (25) suggest that an important aspect of DNA flexibility has yet to be understood. Here, we investigate this issue in detail. We analyzed, by computer simulations, whether the model of DNA flexibility can be modified to give a much higher cyclization efficiency for such short fragments. We concluded that the very high cyclization efficiency of short fragments can be explained by the transient appearance of sharp kinks in the double helix. Such kinks have been found in cocrystals of DNA–protein complexes. It turns out, however, that the model incorporating kinks in DNA predicts a higher cyclization efficiency for longer DNA fragments than was reported earlier. To clarify this issue, we reinvestigated the cyclization efficiency of DNA fragments 105–130 bp in length experimentally. We found, in sharp disagreement with CW (25), that the cyclization efficiency for these fragments *does not* deviate from the theoretical prediction based on WLC. We also found an explanation for the discrepancy between our data and the results reported by CW (25). Finally, our theoretical analysis and experimental data allowed us to estimate an upper limit for the frequency/amplitude of DNA kinks.

Computational and Experimental Procedures

DNA Models and *j*-Factor Calculation. We performed computations for two DNA models, with smooth bending potentials and for a model incorporating kinks in the double helix. For the first set of computations, a DNA molecule of N base pairs in length was modeled as a discrete WLC composed of N rigid segments. The bending elastic energy of the chain, E_b, is computed as

$$E_b = k_B T \sum_{i=1}^{N-1} U(\theta_i), \qquad [1]$$

where the summation extends over all of the joints between the elementary segments, θ_i is the angular displacement of segment $i + 1$ relative to segment i, k_BT is the Boltzmann temperature factor, and $U(\theta) = g_2\theta^2 + g_3\theta^3 + g_4\theta^4$ is a bending potential. Three different sets of g_1, g_2, g_3 were used in the computations: (*i*) the quadratic potential: $g_2 = 69.66$, $g_3 = g_4 = 0$; (*ii*) x-scaled Protein Data Bank (PDB) potential: $g_2 = 123.4$, $g_3 = -225.0$, $g_4 = 113.7$; (*iii*) y-scaled PDB potential: $g_2 = 203.1$, $g_3 = -552.7$, $g_4 = 416.8$. Potential $U(\theta)$ specifies the equilibrium distribution of θ, $P(\theta)$, the average value of cos θ, $\langle\cos\theta\rangle$, and, consequently, the value of DNA persistence length, a (26):

$$a = l/2\cdot(1 + \langle\cos\theta\rangle)/(1 - \langle\cos\theta\rangle), \qquad [2]$$

where l is the segment length. All these potentials give the same value of a, 48nm.

This model does not account for the torsional orientation of DNA ends responsible for j-factor oscillation with the helical periodicity of DNA. For intrinsically straight DNA fragments,

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

Abbreviations: CW, Cloutier and Widom; PDB, Protein Data Bank; WLC, worm-like chain.

*To whom correspondence should be addressed. E-mail: alex.vologodskii@nyu.edu.

considered here, this factor can be easily added to the computed values of j-factor (20).

It should be noted that conformational properties of the model chain with a quadratic bending potential can be approximated very accurately by a chain with longer segments, corresponding to a few base pairs, with a redefined value of g_2 (27). This is not the case, however, for models with nonquadratic terms in the potential.

To calculate j-factors for these model chains, we used an algorithm, based on a sequence of conditional probabilities (23, 28). In this algorithm, restricted sets of chain conformations are generated by the Metropolis procedure. The algorithm allows accurate and efficient calculation of j-factors regardless of their magnitudes.

The second set of computations was performed for a DNA model with a quadratic bending potential and rare sharp kinks in the double helix. All kink directions were assumed to have an equal probability of occurrence. There are three parameters in this model: the kink angle, ϕ, the probability of the kink appearance at a particular stack of base pairs, p, and the bending rigidity of the quadratic potential, g_2. For a particular pair of ϕ and p, the value of g_2 was chosen so that the persistence length of the model chain, specified by Eq. **2**, was equal to 48 nm. For this model,

$$\langle \cos\theta \rangle = (1 - p)\langle \cos\theta_{\mathrm{WLC}} \rangle + p\cos\phi, \qquad [3]$$

where θ_{WLC} is the angle between adjacent base pairs in the absence of the kink. Eqs. **2** and **3** allow the determination of $\langle \cos\theta_{\mathrm{WLC}} \rangle$, which defines g_2.

Since our version of the Metropolis procedure is not suitable for generation of the equilibrium conformational ensemble for this model, a straightforward way was used to calculate j-factors in this case. A large set of independent chain conformations was simulated, and the number of conformations in this set that satisfied the chain closing conditions was counted. The probability of closing, thus determined, was converted to a j-factor value (see ref. 28 for details of the conversion procedure). Although this approach is very time-consuming for small values of j-factor, the calculations are possible after scaling the size of the chain segments. We found that the scaling does not change the value of j-factor if the resulting chain contains >30 straight segments. During the scaling, the value of ϕ was held constant, p was multiplied on the scaling coefficient, and g_2 was recalculated correspondingly. As many as 2×10^{10} chain conformations were generated during a single simulation run.

DNA Preparation. Four sets of DNA fragments, 105–107 bp, 114–118 bp, 123–130 bp, and 199 bp, were prepared by PCR amplification from a fragment of λ DNA starting from nucleotide 29,853. This sequence was chosen because it does not contain any known intrinsically curved elements of the double helix, AAAA, ATTT, AATT, GGGCCC. The PCR primers were designed to produce DNA subfragments of proper lengths between two *Hin*dIII restriction sites. All PCR products were cleaned by using a QIAquick PCR purification kit (Qiagen) and subsequently digested by *Hin*dIII restriction enzyme. The resulting fragments with the desired lengths were cloned into the *Hin*dIII site of pUC19 plasmid vector. These constructs were then transformed into DH5α competent cells, and the lengths and sequences of all inserts were verified by a DNA sequencing facility. The plasmid DNA was isolated from bacterial cultures by using the QIAprep Spin Miniprep kit (Qiagen). Concentrations of purified DNA plasmids were determined from their absorbance at 260 nm. DNA sequences are available upon request.

Radioactive Labeling of DNA Fragments. The plasmids carrying the short DNA inserts were digested by *Hin*dIII restriction endonuclease (NEB). Each digested DNA was 5′-end-labeled with ^{32}P by T4 polynucleotide kinase (NEB) in a 12-μl total volume, containing 7 μl of [γ-^{32}P]ATP [10 mCi/ml, 6,000 Ci/mmol (1 Ci = 37 GBq); PerkinElmer]. The DNA concentrations were 5–20 nM. The labeling was carried out at 37°C for 40 min, followed by heat inactivation at 65°C for 20 min.

Ligation Time Course. The plasmid vectors remained in solution, along with the excised short DNA inserts, during the subsequent labeling and the ligation time course. Leaving the vector DNA in the reaction mixture does not affect the ratio of circular and dimeric forms of the insert at the early stage of ligation while greatly facilitating sample preparation. As an additional control, the ligation experiments for the 106-bp fragment were also performed with a sample prepared by PCR amplification. These experiments gave the same value of j-factor as the sample obtained by plasmid restriction.

Ligation experiments were performed in 100- to 1,000-μl volumes, using T4 DNA ligase (400 units/μl; NEB) and its standard buffer at 21°C. The final concentrations of DNA substrates in ligation buffer were 0.025 nM for the 100-bp series, 0.1 nM for 110- and 120-bp series, and 1 nM for the 199-bp fragment. Each reaction was initiated by the addition of ligase diluted from stock with its standard buffer just before the ligation experiments. Aliquots of the ligation mixtures were withdrawn at specific time intervals and quenched with EDTA at a 50 mM final concentration. Unincorporated [γ-^{32}P]ATP from the ligation samples was removed with Spin-50 Sephadex minicolumns in TE buffer (Biomax). Alternatively, desalting of ligation samples was carried out by using Microcon YM-30 centrifugal filters (Millipore) when concentration of the DNA samples was necessary. Before applying the samples to Microcon filters, the ligase was heat inactivated at 65°C for 10 min. All purification procedures followed the protocols of the manufacturers.

Gel Electrophoresis. Ligation products from 105- to 130-bp fragments were separated by the use of 4–12% gradient Novex polyacrylamide gels (Invitrogen) at 130 V in TBE buffer (89 mM Tris borate/2 mM EDTA, pH 8.3) (see the supporting information, which is published on the PNAS web site). Ligation products from 199-bp experiments were separated on 2.2% MetaPhor agarose gel (Cambrex) at 5 V/cm in TBE buffer for 8 h. The gels were equilibrated in an ethanol/glycerol solution and dried. The bands were subsequently quantified by using a phosphorimager.

Results

Theoretical Analysis of the DNA Flexibility. The discrete WLC is a chain of N rigid segments of equal length l. The bending energy of the chain, E_{b}, is specified as

$$E_{\mathrm{b}} = g \sum_{n=1}^{N-1} \theta_i^2, \qquad [4]$$

where θ_i is the angular displacement of segment $i + 1$ relative to segment i, and g is the bending rigidity constant. Each segment of the chain corresponds to a certain number of base pairs, k. The value of g is chosen to correspond to the DNA bending rigidity, akT, where a is the persistence length of the double helix, which is close to 50 nm (1). In the limit $l \to 0$, the discrete WLC is converted to the WLC model. The distribution of the angles θ_i in a model chain, $P(\theta)$, is specified by the Boltzmann distribution:

$$P(\theta) \propto \sin\theta \, \exp(-g\theta^2/k_{\mathrm{B}}T), \qquad [5]$$

where sin θ accounts for the number of conformations with angle θ (see ref. 29 for an example) and k_BT is the Boltzmann temperature factor.

Clearly, the discrete model is more convenient for computations. As the value of l decreases, each conformational property approaches its limiting value. Which value of l provides the limiting property depends on the conformational property of interest. It is noteworthy that, although the model does not account for local anisotropy of DNA bending rigidity, the anisotropy does not affect properties of DNA fragments if they are long enough to include a few turns of the double helix (30).

This model of the double helix, sometimes with additional features, is widely used in computer simulations of DNA properties. In particular, it is used to calculate the j-factor, the effective concentration of one DNA end in the vicinity of the other, that specifies the fragment cyclization efficiency (20, 21, 23, 30–32). It is important to note, however, that the great majority of DNA statistical features are insensitive to details of the model such as a specific form of the bending energy function (see Eq. **1**). In particular, a depends only on the average value of cos θ rather than on $P(\theta)$. The WLC was initially introduced as a limit of the model with fixed values of θ_i and free rotation of each segment around the direction of the previous segment (3), and it is hard to find a statistical property that is affected by this difference in $P(\theta)$. Furthermore, many properties of large DNA molecules, longer than a few kilobases, can be equally well described by a much simpler model, the freely jointed chain (see ref. 29, for example). There are properties, however, that are sensitive to model details. We show below that this is the case for the j-factor values of very short DNA fragments. A similar conclusion has recently been reached by two other groups (33, 34).

The right side of Eq. **4**, which specifies $P(\theta)$, should be considered as the first meaningful term of the Taylor expansion of the bending energy. Thus, the equation is accurate as long as the values of θ_i are sufficiently small. However, the next terms of the energy expansion could be important for the bend angle values playing a role in the cyclization of short DNA fragments. It has also been suggested that rare sharp kinks of the double helix may contribute to the j-factor value of short DNA fragments (25). Here, we investigate two such options for $P(\theta)$. First, we analyze the possible effect of a nonquadratic potential for the bending energy. Second, we investigate how the transient appearance of sharp kinks in the double helix could affect the j-factor values. Both options address the effect of larger angles in $P(\theta)$. During this analysis we assumed, regardless of the microscopic mechanism of bending, that DNA persistence length always equals 48 nm (1, 22, 24).

In the choice of a nonquadratic potential for the bending energy many options are available. As a guide for a reasonable choice, we extracted the distribution $P(\theta)$ from the database of DNA–protein complexes (35). This PDB distribution contains large angles between adjacent base pairs, differing in this respect from the angle distribution in DNA crystals. Although there is no reason to assume that the PDB distribution corresponds to the Boltzmann distribution at room temperature (Eq. **5**), we suggest that it reflects features of the potential we want to approximate. The distribution is shown in Fig. 1*A* together with the distribution for the WLC corresponding to Eqs. **4** and **5**. There is a pronounced additional peak in the PDB distribution with a maximum at 48°, and smaller isolated peaks corresponding to still larger bend angles, which we consider as kinks of the double helix. Even if the kinks are ignored, the distribution corresponds to a value of a that is two times smaller than the experimental value. This means that the double helix is more strongly perturbed, on average, in these complexes than DNA in solution. Therefore, we adjusted the PDB distribution of θ to have the known value of a. First, we approximated the PDB distribution by an equilibrium distribution $P_{apr}(\theta)$ with the bending energy $E_{apr}(\theta)$, specified by a fourth order polynomial (kinks were omitted in this approximation). Two adjusted distributions $P(\theta)$ were obtained by scaling $E_{apr}(\theta)$ along the x or along the y axis to get the desired value of a. The resulting energy functions are shown in Fig. 1*B*. These functions were used to calculate the values of j-factors as a function of DNA length. It can be seen from the results, shown in Fig. 2*A*, that both potentials increase the j-factor values for DNA fragments shorter than 200 bp. In particular, for the fragments of ≈100 bp in length the increase is close to a factor of 10. We tested several other modifications of the PDB distribution and obtained very similar results. We conclude from these data that the small probability of larger bend angles, up to 40–50°, results in a substantial increase of j-factor values for short model chains, although the increase is insufficient to explain the experimental data, reported by CW (25). It is important that the j-factor values for longer chains are not affected by these modifications of $P(\theta)$.

To increase the effect of larger bend angles on the j-factor value of 100-bp fragments, one must increase the probability of their appearance or the magnitude of these angles. We investi-

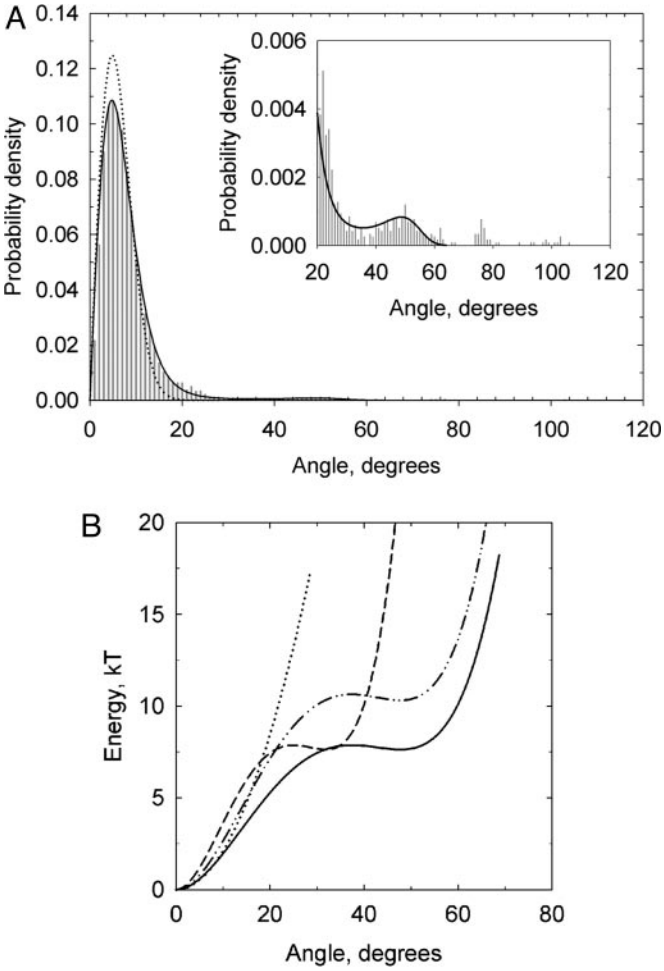

Fig. 1. Distribution of angles between adjacent base pairs, $P(\theta)$. (*A*) The distribution extracted from the 650 structures of DNA–protein complexes that give 11,732 different values for the angles (35), and its smoothed approximation, P_{PDB} (solid line). $P(\theta)$ for the energy function specified by Eqs. **4** and **5** is shown for comparison (dotted line). (*Inset*) Part of the same distribution for larger angles that have low probabilities of appearance. (*B*) Dependence of the bending energy on the bend angle. The function that corresponds to P_{PDB} (solid line), obtained by applying Eq. **5** to P_{PDB}, is shown together with the quadratic potential (dotted line), x-scaled P_{PDB} (dashed line), and y-scaled P_{PDB} (dashed–dotted line).

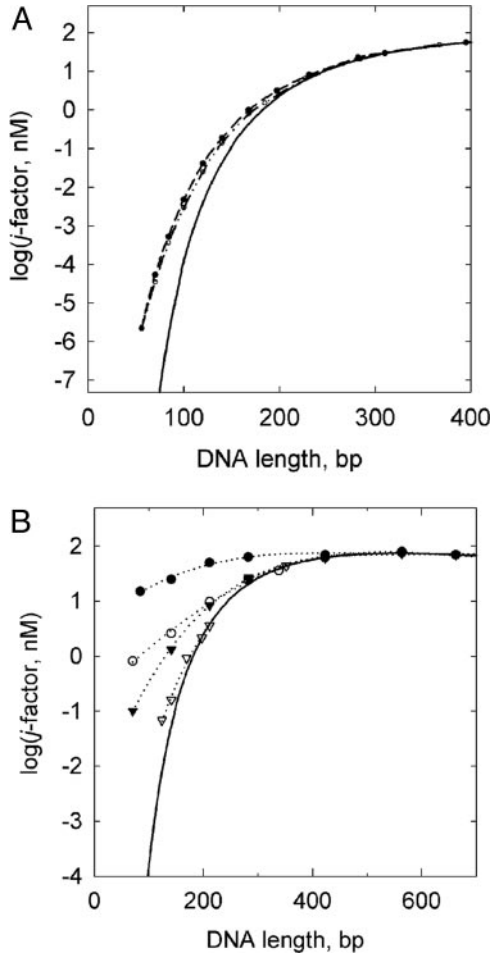

Fig. 2. Computed *j*-factors for different models of DNA bending. (*A*) The values of *j*-factors calculated for *x*-scaled PDB potential (dashed line) and *y*-scaled PDB potential (dashed–dotted line) are shown together with the theoretical data for the WLC with quadratic bending potential (solid line). The later dependence, taken from ref. 21, was also reproduced by our computation. (*B*) The values of *j*-factors for the WLC DNA model that incorporates randomly directed kinks of 70° (triangles) and 100° (circles). The data correspond to two different probabilities of the kinks, p_k, 0.002 (filled symbols) and 0.0005 (open symbols). DNA persistence length, which depends on both the quadratic potential with a certain value of the bending rigidity and on p_k and θ_k of the kinks, was equal to 48 nm. To eliminate oscillation of the *j*-factor with DNA helical periodicity, all calculations were performed for zero DNA torsional rigidity.

gated the second option, adding to the WLC, with potential specified by Eq. **4**, the possibility of forming sharp kinks of the double helix. We assumed that kinks with angle θ_k could appear at any stack of the base pairs with probability p_k and that their directions are uniformly distributed. For each particular value of p_k and θ_k, the bending rigidity of the discrete WLC was chosen so that the total persistence length equals 48 nm (see *Computational and Experimental Procedures* for details). We performed computations for different values of p_k and θ_k, and found that both parameters strongly affect the *j*-factors for short model chains. Computed values of *j*-factors for four pairs of p_k and θ_k values are shown in Fig. 2*B*. It can be seen from the figure that the appearance of kinks of 100° with $p_k = 0.002$ increases the value of *j*-factor for 100-kb DNA by five orders of magnitude. Kinks of 100° and greater were found in the PDB (Fig. 1*A*). The effect of kinks reduces quickly, however, as the values of p_k and θ_k decrease. We conclude that sharp kinks of the double helix can explain a few orders of magnitude increase in *j*-factors for 100-bp

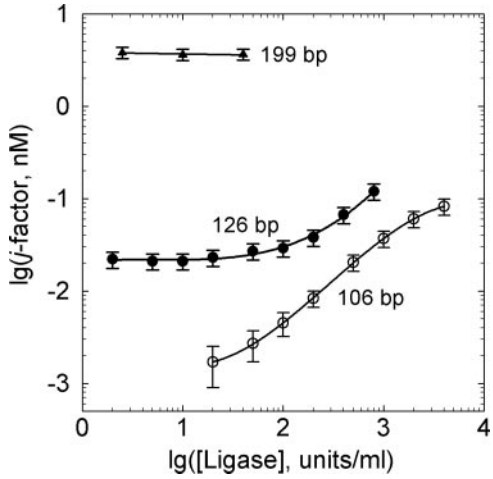

Fig. 3. Measured values of *j*-factors as a function of [ligase]. The data for the fragments of 106 bp (open circles), 126 bp (filled circles), and 199 bp (filled triangles) are shown. [DNA] was equal to 25 pM (106 bp), 100 pM (126 bp), and 1 nM (199 bp).

DNA fragments, and the probability of appearance of such sharp kinks should be ≈0.001. Clearly, there is nothing unreasonable in this model of DNA bending.

Our analysis, however, highlighted a problem. The experimentally measured *j*-factors for DNA fragments of larger lengths, 200–250 bp (24, 25, 36), follow the theoretical predictions for the WLC with the quadratic bending potential specified by Eq. **1** perfectly. Kinks with parameters that produce an increase of the *j*-factor value for 100-bp DNA fragments by three orders of magnitude, however, would also increase the *j*-factor for 200-bp fragments, as is clearly shown in Fig. 2*B*. This ≈3-fold increase is incompatible with experimental data obtained for the 200- to 250-bp fragments (20, 24). In an attempt to resolve this discrepancy, we decided to repeat the cyclization experiments of CW (25) on DNA fragments ≈100 bp in length.

Experimental Measurements of *j*-Factor for Short DNA Fragments. Ligation of DNA fragments with cohesive ends can produce many different products. The major products produced during the initial stage of the ligation reaction, however, are circular monomers (CM) and linear and circular dimers (LD and CD) of the fragments. Following previous studies (22, 24), we determined the *j*-factor of a DNA fragment as the ratio of the amounts of CM, denoted by $C(t)$, and LD and CD, denoted by $D(t)$, formed during the early stages of fragment ligation:

$$j = 2M_0 \lim_{t \to 0} C(t)/D(t), \qquad [6]$$

where M_0 is the initial concentration of the fragments. To measure $C(t)/D(t)$, DNA samples were end-labeled by ^{32}P and separated by gel electrophoresis after ligation (see ref. 24 and supporting information for details).

In performing our experiments with 105–130 bp fragments we unexpectedly found that the measured values of *j*-factors depend on [ligase] (Fig. 3). The *j*-factor values increased with an increase in [ligase] and, for the shortest set of fragments, represented in the figure by the 106 bp fragment, the change approaches 2 orders of magnitude over the range of [ligase] suitable for the measurements. To understand the meaning of this finding we have to return to the original analysis of the *j*-factor determination through fragment ligation of Baldwin and coworkers (20, 36). Their analysis defined the reaction conditions under which the *j*-factor is specified by the measured rates of fragment

cyclization and dimerization, k_1 and k_2 [the value of k_2 here corresponds to the dimerization rate constant for fragments with two identical cohesive ends and is four times larger than the constant used by (20)]. The ligation reaction was considered as a three-step process (20, 36):

$$L \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} S$$

$$S + E \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} SE \overset{k_{34}}{\longrightarrow} E + P,$$

where L is a linear DNA fragment, S is a substrate for DNA ligase (E), SE is a complex of ligase with jointed DNA cohesive ends, and P is a reaction product. Here, S corresponds to the circular form of the fragment or the fragment dimer with the conformation of the cohesive ends suitable for ligation. The main conclusion of the analysis is that k_1 and k_2 specify the j-factor only if

$$k_{21} \gg k_{23}[E], \qquad [7]$$

so that the rate of substrate decay is much higher than the rate of ligase binding with one of two nicks at the joined ends. Under this condition the rate of product accumulation is proportional to [E]. In the opposite case, where $k_{21} \ll k_{23}[E]$, the rate of the product accumulation does not depend on [E], since the first joining of the sticky ends results in their ligation. It is important to note that the dissociation rate k_{21} can be different for circles and dimers formed by a fragment.

The very small j-factor values of these short DNA fragments motivate using higher [ligase] to reduce the ligation time course. Thus, although condition **7** was satisfied in earlier experiments with longer fragments (20, 22, 24, 32), it may fail for fragments of about 100 bp in length at higher [ligase]. CW (25) tested that k_1, the rate of fragment cyclization, is proportional to [ligase] over a range of 20–600 units/ml. Such linear dependence means that condition **7** is satisfied. We measured the same dependence and also found that it is linear over the broad range of [ligase] (Fig. 4*A*). We found, however, that the rate of dimerization, k_2, does not depend on [ligase] if [ligase] $>$ 100 unit/ml (Fig. 4*B*). Thus, condition **7** is not satisfied for k_2 if [ligase] $>$ 100 nM for cohesive ends used in our study, with the sequence of AGCT. These results indicate that the rates of dissociation of the joined cohesive ends, k_{21}, are different for dimers and circles of these short fragments. CW (25) did not measure k_2 for dimers, implicitly assuming that if **7** is satisfied for circles, it is also satisfied for dimers.

The above analysis shows that, for the *Hin*dIII sticky ends used in our study, AGCT, one has to use a [ligase] a few times $<$100 units/ml to determine the j-factor from a ligation experiment. This restriction creates a problem for the smallest circles used in this study, since the enzyme produces few circles before losing its activity. However, since k_{21} is many times faster for circles than for dimers, k_1 and k_2 can be measured in separate experiments, using higher [ligase] for k_1 determination. Then, the j-factor can be calculated as $2k_1/k_2$ (20). This increases the experimental error, however, since it requires precise determination of the ligase concentration and activity, which is not needed for j-factor calculation from Eq. **6**. Running the reaction in a large volume and then concentrating the reaction mixture before gel electrophoresis is another method we used to overcome this problem.

The j-factor values obtained for DNA fragments 105–130 bp in length are shown in Fig. 5. The data were fitted by the theoretical values, calculated for the WLC with a quadratic potential. Here, torsional orientation of the fragment ends is taken into account, as opposed to the data presented in Fig. 2,

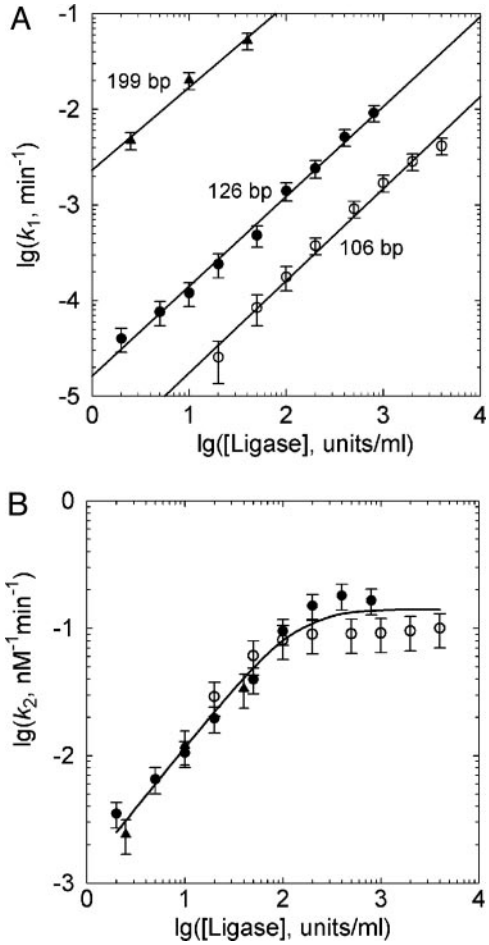

Fig. 4. The rate constants of cyclization, k1, and dimerization, k2, as a function of [ligase]. The constants are specified by equations $dC/dt = k_1C$ and $dD/dt = k_2D^2$. The data for the fragments of 106 bp (open circles), 126 bp (filled circles), and 199 bp (filled triangles) were obtained in the same experiments as the results shown in Fig. 3. The line in *B* represents a summary over the data for all fragments.

resulting in the j-factor oscillation (see ref. 20 for details). There are three parameters that specify the theoretical dependence: DNA persistence length, a, the helical repeat of the double helix, γ, and DNA torsional rigidity, C. By varying the values of a and γ, we found that the best fit corresponds to a = 47 nm and γ = 10.54 nm. This value of a is in full agreement with the previous data, obtained from the cyclization experiments on longer DNA fragments, 45–49 nm (20–22, 24). The same values of a have also been obtained by other methods (reviewed in ref. 1). The value of γ is in agreement with numerous solution data (20, 22, 24, 37, 38). Thus, a theory based on the WLC with a quadratic potential describes the cyclization of these very small DNA fragments with remarkable accuracy.

Discussion

In contrast to CW (25), we found that the values of j-factors for DNA fragments in the 100-bp range are in full agreement with theoretical predictions based on the WLC. The discrepancy between our results and those of CW (25) appears to be due to the high [ligase] used in their study. To determine the value of j-factor from the ligation experiments, the rate of ligation of joined cohesive ends must be much slower than the rate of their dissociation (20). Although CW (25) showed that this condition is satisfied during cyclization of short DNA fragments, they did

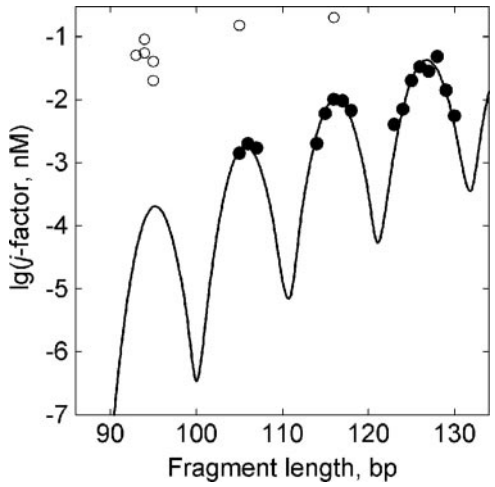

Fig. 5. *j*-factor of short DNA fragments as a function of their length. Our experimental data (filled circles) are shown together with the results obtained by CW (25) (open circles). [ligase] was equal to 20 units/ml. Our data were fitted by the theoretical values for the WLC that also account for the torsional orientation of the fragment ends (solid line). The best fit, shown here, was obtained for a DNA persistence length of 47 nm and the helical repeat of 10.54 bp per helix turn. The third parameter of the curve, DNA torsional rigidity, was taken from the previous studies (22, 24) and was equal to 2.4×10^{-19} cm·erg (1 erg = 0.1 μJ).

not see that it is not satisfied for the same cohesive ends during fragment dimerization. We found that for a [ligase] of 250 units/ml used by CW (25), the condition of slow dimer ligation is not satisfied even for the cohesive ends used in our study, with nucleotide sequence AGCT. The cohesive ends in the study of CW, GGCC, are more stable, and therefore their dissociation rate is essentially slower (see discussion in ref. 32). Indeed, the melting temperature of cohesive ends GGCC is about 20°C higher than the melting temperature of AGCT (39). Much lower [ligase] needs to be used for GGCC cohesive ends to determinate *j*-factors. Of course, using lower [ligase] complicates the *j*-factor measurements, since the enzyme makes too few circles before losing its activity. Less stable cohesive ends would not help to solve this problem, however. Although they allow performing the measurements at higher [ligase], they also slow down the ligation reaction.

The fact that the rate of cohesive end dissociation, k_{21}, is many times larger for circles than linear dimers made from short fragments is hardly surprising. Indeed, the elastic stress in the circular conformations of the fragments should accelerate the dissociation. The difference in the k_{21} values should disappear for fragments longer than two to four persistence lengths.

We conclude that the WLC with the bending potential specified by Eq. **4** gives a remarkably good description of the DNA conformational properties, even for the extreme deformations occurring during cyclization of 100-bp fragments. Neither kinks of the double helix nor nonquadratic terms of the bending potential affect the *j*-factors of these fragments. Our theoretical analysis, however, shows that relatively small changes of kink probabilities and angles would dramatically change their influence on the short fragment cyclization. One can see from Fig. 2*B* that kinks of 70°, occurring with a probability of 0.002, would increase the *j*-factor of 100-bp fragments by nearly three orders of magnitude, whereas the same kinks with probability of 0.0005 make only small contributions to the cyclization efficiency. Therefore, a good agreement between our experimental data and theoretical predictions based on WLC means that the probability of kinks by angles 70–90° does not exceed 0.0002.

We thank T. Cloutier and J. Widom for detailed experimental protocol and Mrs. J. de la Cruz for assistance with the Protein Data Bank. This work was supported by National Institutes of Health Grant R01-GM54215 (to A.V.).

1. Hagerman, P. J. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17,** 265–286.
2. Schellman, J. A. (1974) *Biopolymers* **13,** 217–226.
3. Kratky, O. & Porod, G. (1949) *J. R. Neth. Chem. Sci.* **68,** 1106–1122.
4. Hagerman, P. J. & Zimm, B. H. (1981) *Biopolymers* **20,** 1481–1502.
5. Hagerman, P. J. (1981) *Biopolymers* **20,** 1503–1535.
6. Rybenkov, V. V., Vologoskii, A. V. & Cozzarelli, N. R. (1997) *J. Mol. Biol.* **267,** 299–311.
7. Klenin, K. V., Vologodskii, A. V., Anshelevich, V. V., Dykhne, A. M. & Frank-Kamenetskii, M. D. (1988) *J. Biomol. Struct. Dyn.* **5,** 1173–1185.
8. Klenin, K. V., Vologodskii, A. V., Anshelevich, V. V., Klisko, V. Y., Dykhne, A. M. & Frank-Kamenetskii, M. D. (1989) *J. Biomol. Struct. Dyn.* **6,** 707–714.
9. Rybenkov, V. V., Cozzarelli, N. R. & Vologodskii, A. V. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5307–5311.
10. Shaw, S. Y. & Wang, J. C. (1993) *Science* **260,** 533–536.
11. Vologodskii, A. V. & Cozzarelli, N. R. (1993) *J. Mol. Biol.* **232,** 1130–1140.
12. Rybenkov, V. V., Vologodskii, A. V. & Cozzarelli, N. R. (1997) *J. Mol. Biol.* **267,** 312–323.
13. Gebe, J. A., Delrow, J. J., Heath, P. J., Fujimoto, B. S., Stewart, D. W. & Schurr, J. M. (1996) *J. Mol. Biol.* **262,** 105–128.
14. Hammermann, M., Stainmaier, C., Merlitz, H., Kapp, U., Waldeck, W., Chirico, G. & Langowski, J. (1997) *Biophys. J.* **73,** 2674–2687.
15. Klenin, K., Hammermann, M. & Langowski, J. (2000) *Macromolecules* **33,** 1459–1466.
16. Hammermann, M., Brun, N., Klenin, K. V., May, R., Toth, K. & Langowski, J. (1998) *Biophys. J.* **75,** 3057–3063.
17. Smith, S. B., Finzi, L. & Bustamante, C. (1992) *Science* **258,** 1122–1126.
18. Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. (1994) *Science* **265,** 1599–1600.
19. Vologodskii, A. V. (1994) *Macromolecules* **27,** 5623–5625.
20. Shore, D. & Baldwin, R. L. (1983) *J. Mol. Biol.* **170,** 957–981.
21. Shimada, J. & Yamakawa, H. (1984) *Macromolecules* **17,** 689–698.
22. Taylor, W. H. & Hagerman, P. J. (1990) *J. Mol. Biol.* **212,** 363–376.
23. Podtelezhnikov, A. A. & Vologodskii, A. V. (2000) *Macromolecules* **33,** 2767–2771.
24. Vologodskaia, M. & Vologodskii, A. (2002) *J. Mol. Biol.* **317,** 205–213.
25. Cloutier, T. E. & Widom, J. (2004) *Mol. Cell* **14,** 355–362.
26. Frank-Kamenetskii, M. D., Lukashin, A. V., Anshelevich, V. V. & Vologodskii, A. V. (1985) *J. Biomol. Struct. Dyn.* **2,** 1005–1012.
27. Vologodskii, A. V. & Frank-Kamenetskii, M. D. (1992) *Methods Enzymol.* **211,** 467–480.
28. Podtelezhnikov, A. A., Mao, C., Seeman, N. C. & Vologodskii, A. V. (2000) *Biophys. J.* **79,** 2692–2704.
29. Cantor, C. R. & Schimmel, P. R. (1980) *Biophysical Chemistry* (Freeman, New York).
30. Zhang, Y. & Crothers, D. M. (2003) *Biophys. J.* **84,** 136–153.
31. Hagerman, P. J. & Ramadevi, V. A. (1990) *J. Mol. Biol.* **212,** 351–362.
32. Crothers, D. M., Drak, J., Kahn, J. D. & Levene, S. D. (1992) *Methods Enzymol.* **212,** 3–29.
33. Yan, J. & Marko, J. F. (2004) *Phys. Rev. Lett.* **93,** 108108.
34. Wiggins, P. A., Phillips, R. & Nelson, P. C. (2005) *Phys. Rev. E* **71,** 021909-1–021909-19.
35. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R. & Schneider, B. (1992) *Biophys. J.* **63,** 751–759.
36. Shore, D., Langowski, J. & Baldwin, R. L. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 4833–4837.
37. Wang, J. C. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 200–203.
38. Horowitz, D. S. & Wang, J. C. (1984) *J. Mol. Biol.* **173,** 75–91.
39. SantaLucia, J., Jr., Allawi, H. T. & Seneviratne, P. A. (1995) *Biochemistry* **35,** 3555–3562.