Multimerization-Cyclization of DNA Fragments as a Method of Conformational Analysis

Alexei A. Podtelezhnikov, Chengde Mao, Nadrian C. Seeman, and Alexander Vologodskii
Department of Chemistry, New York University, New York, NY 10003 USA

ABSTRACT Ligation of short DNA fragments results in the formation of linear and circular multimers of various lengths. The distribution of products in such a reaction is often used to evaluate fragment bending caused by specific chemical modification, by bound ligands or by the presence of irregular structural elements. We have developed a more rigorous quantitative approach to the analysis of such experimental data based on determination of *j*-factors for different multimers from the distribution of the reaction products. *j*-Factors define the effective concentration of one end of a linear chain in the vicinity of the other end. To extract *j*-factors we assumed that kinetics of the reaction is described by a system of differential equations where *j*-factors appear as coefficients. The assumption was confirmed by comparison with experimental data obtained here for DNA fragments containing A-tracts. At the second step of the analysis *j*-factors are used to determine conformational parameters of DNA fragments: the equilibrium bend angle, the bending rigidity of the fragment axis, and the total twist of the fragments. This procedure is based on empirical equations that connect the conformational parameters with the set of *j*-factors. To obtain the equations, we computed *j*-factors for a large array of conformational parameters that describe model fragments. The approach was tested on both simulated and actual experimental data for DNA fragments containing A-tracts. A-tract DNA bend angle determined here is in good agreement with previously published data. We have established a set of experimental conditions necessary for the data analysis to be successful.

INTRODUCTION

Long DNA molecules can be cyclized in solution because they are flexible. The efficiency of cyclization depends on the DNA length and on the presence of bends in the double helix, either intrinsic bends or those induced by bound ligands or proteins. The experimental study of the cyclization was pioneered by Wang (Wang and Davidson, 1966a,b, 1968; Wang, 1967). A well-known application of cyclization studies to DNA flexibility was performed by Shore and Baldwin (1983), who obtained elegant evidence of the helical nature of DNA, manifested in the periodicity of the cyclization efficiency. Crothers's and Hagerman's laboratories have developed a quantitative analysis of cyclization efficiency to study DNA conformational properties. Hagerman and coworkers applied the approach to study regular DNA and obtained accurate estimates of the DNA persistence length, helical repeat, and torsional rigidity of the double helix for different ionic conditions (Taylor and Hagerman, 1990; Hagerman and Ramadevi, 1990). Crothers and coworkers investigated the intrinsic bending of DNA associated with A-tracts (Koo et al., 1990; Crothers et al., 1992) as well as bending induced by protein binding (Kahn and Crothers, 1992; Zeman et al., 1998; Kahn and Crothers, 1998). Conformational information that can be obtained by this method is hardly accessible by other biophysical methods like X-ray analysis, whose results may be affected by packing forces, or nuclear magnetic resonance (NMR), which is not sensitive enough to give information about bending angles. Thus, cyclization efficiency can serve as an efficient tool to study DNA conformational features associated with a variety of structural elements, such as branch points, mismatches, chemical modifications, and ligand or protein binding.

To evaluate a few conformational parameters the experimental data have to be obtained for different DNA lengths. The choice of these lengths is critical for the subsequent analysis. Two major approaches were suggested. In the first of them introduced by Shore and Baldwin (1983) fragment lengths are chosen to cover one helical repeat of DNA to highlight the effect of equilibrium torsional orientation of the fragment ends on the cyclization efficiency. About ten different lengths, which differ by 1–2 bp, have to be used in this case; the lengths should not exceed 400–500 bp. The approach allows very accurate and unambiguous determination of DNA persistence length, helical repeat and torsional rigidity (Shore and Baldwin, 1983; Hagerman and Ramadevi, 1990). This approach is certainly preferable to studying properties of the double helix under different conditions, but it is more difficult to apply the experimental design to the molecules with irregular structural elements, which cannot be reproduced by replication. A set of multimers of an oligonucleotide 21 bp in length, what corresponds to two turns of the double helix, is used in the second approach introduced by Ulanovsky et al. (1986). Unambiguous determination of DNA conformational parameters is much more difficult in this case, partially because DNA helical repeat is not equal to 10.5 bp exactly and the accounting for torsional misfit of DNA ends during the cy-

Received for publication 29 December 1999 and in final form 15 August 2000.

Address reprint requests to Dr. Alexander Vologodskii, New York University, Department of Chemistry, 31 Washington Place, New York, NY 10003-6630. Tel.: 212-998-3599; Fax: 212-260-7905; E-mail: alex.vologodskii@nyu.edu.

0006-3495/00/11/2692/13 $2.00

clization is not so simple. However, the approach has the important advantage of experimental simplicity for applications involving complex structural motifs, like mismatches, modified bases, Holiday junctions, etc., incorporated in the oligonucleotides. It was used to evaluate the values of DNA bends caused by A-tracts (Ulanovsky et al., 1986; Koo et al., 1990), three-way DNA junctions (Shlyakhtenko et al., 1994a,b), protein or ligand binding with DNA fragments (Lyubchenko et al., 1991, 1993; Balagurumoorthy et al., 1995; Dlakic and Harrington, 1995, 1996; Nagaich et al., 1997), or perturbation of the regular double helix resulting from a chemical modification (Xu et al., 1998). This second approach is the subject of the current paper.

There are two possible ways to obtain the data for cyclization efficiency of the multimers of an oligonucleotide. Crothers and coworkers used a straightforward solution by measuring the cyclization efficiency for different multimers individually (Koo et al., 1990). It is more attractive from the experimental point of view to extract the same data from the product distribution in a single multimerization-cyclization reaction. Both linear and circular products of multiple ligation of the oligonucleotide are obtained in the same reaction, and the formation of shorter circles correlates well with intrinsic or induced bends in the monomers. This approach introduced by Ulanovsky et al. (1986) is now widely used to detect bends in DNA fragments. However, the approach based on the multimerization-cyclization reaction has been applied as a semiquantitative method only, because there has been no regular procedure to extract conformational parameters from the experimental data. Usually only the length of the modal circular product or the length of the smallest circle has been used to draw conclusions regarding the bend angle in the monomer. These conclusions overestimate the value of the bend angle (see below and Hagerman and Ramadevi, 1990). Here we developed a procedure to extract conformational data from the distributions of linear and circular multimers obtained in the reaction. The procedure has been tested on simulated experimental data where the desired answers were known exactly, and on actual experimental data obtained for a DNA fragment containing well-studied bends associated with A-tracts. We have established a set of experimental conditions that have to be satisfied to obtain data suitable for this type of analysis.

METHODS

Synthesis and purification of DNA

The DNA molecule used in this study was synthesized on an Applied Biosystems 380B automatic DNA synthesizer (Applied Biosystems, Foster City, CA) removed from the support, and deprotected, using routine phosphoramidite procedures (Caruthers, 1985); the DNA was purified from 12% polyacrylamide denaturing gels. The molecule synthesized is shown in Fig. 1. This molecule forms a hairpin containing restriction sites that enable us to produce a DNA duplex in which each 5′ end can be guaranteed to be phosphorylated.

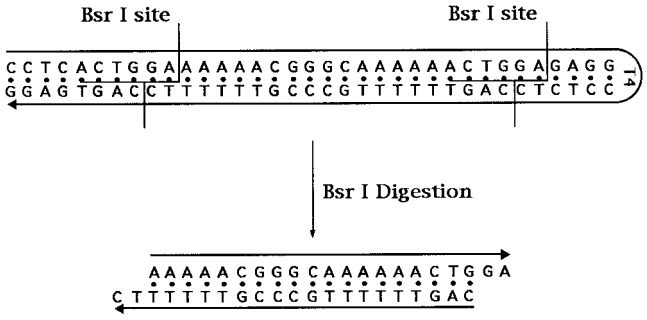

FIGURE 1 Preparation of DNA fragment for the multimerization-cyclization reaction. The drawing illustrates a DNA hairpin that has been synthesized as a single strand. The hairpin contains two Bsr I restriction sites which create non-self-complementary cohesive ends and thus guarantee identical orientation of the fragments in multimers. Digestion at these positions produces the completely phosphorylated fragment used for the experiments described here.

Preparation of the 21-bp DNA duplex

Three nanomoles of the purified hairpin molecules were treated with 500 units of corresponding restriction enzymes obtained from New England Biolabs (NEB; Beverly, MA) for 2 hours in 500 μl of NEB buffer 2 at 37°C to obtain the required fragments with cohesive ends (Fig. 1). The reaction mixture was phenol-extracted and ethanol-precipitated, followed by electrophoresis on a nondenaturing 10% polyacrylamide gel in a buffer containing 40 mM Tris-HCl, pH 8.0, 20 mM acetic acid, 2 mM EDTA, and 12.5 mM magnesium acetate (TAE/Mg) at room temperature. After staining with ethidium bromide, the band corresponding to the 21-bp DNA molecule was excised from the gel under UV illumination. The DNA molecule was eluted, butanol-extracted, ethanol-precipitated, and resuspended in a buffer containing 10 mM Tris-HCl, pH 7.6, and 1 mM EDTA (TE).

Radioactive labeling

One picomole of the DNA duplex was mixed with 1 μl of 2.2 μM γ-^{32}ATP (10 mCi/mL) and 3 units of polynucleotide kinase obtained from U.S. Biochemical (USB, Cleveland, OH) in 10 μl of kination buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, and 10 mM 2-mercaptoethanol) for 3 hours at 37°C. The reaction mixture was subjected to phenol-extraction, a MicroSpin G-25 column (Amersham Pharmacia Biotech, Piscataway, NJ), ethanol-precipitated, and resuspended in TE buffer.

Ligation time course

Ligations were performed in 20 μl of the ligation buffer used by Koo et al. (1990) (50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, and 100 μg/ml bovine serum albumin) at 22°C. The DNA concentration was 100 μM and 0.2 units/μl DNA ligase (USB) were used. From the original 20 μl reaction solution, a 1 μl solution was withdrawn and mixed with 20 μl of denaturing gel loading buffer (90% formamide, 10 mM NaOH, 50 mM EDTA, trace of xylene cyanol FF) preheated to 95°C at the following time points: 0, 30 s, 1, 2, 4, 10, 20, and 40 min, and 16 h. Five-microliter aliquots of each sample were then applied to a 7% denaturing polyacrylamide gel to determine which sample was suitable for data analysis. For the accuracy of the data extraction, the experiment must keep roughly 10% of the monomers unligated for the optimal analysis.

Two-dimensional denaturing polyacrylamide gel electrophoresis

The first dimension contained 5% polyacrylamide, and the second contained 7% acrylamide (19:1 acrylamide:bisacrylamide). These gels also contained 8.3 M urea and were run at 55°C. The running buffer consisted of 89 mM Tris-HCl, pH 8.0, 89 mM boric acid, 2 mM EDTA (TBE). After electrophoresis, they were dried onto Whatman 3MM paper (Whatman International Ltd., UK) and exposed to X-ray film overnight or quantitated using a Bio-Rad GS-525 Molecular Imager (BioRad, Hercules, CA).

Solution of kinetic differential equations

The system of differential equations describing the kinetics of the multimerization-cyclization reaction (see Eq. 2 below) was solved using the Euler difference scheme (Godunov and Riabenkii, 1987). We arbitrarily set the rate constant k_2^* equal to 1 $M^{-1}s^{-1}$ since its value scales only the absolute time of the reaction, which was not essential for the current analysis. The time step of the integration was equal to $1 \cdot 10^{-5}\ (k_2^* M_0)^{-1}$. We checked that time steps 5 times smaller improved the accuracy of calculations by less than 1%. With this time step size, the calculation of the product distribution with 1% of the monomers remaining takes less than 1 min on a Pentium II 266 MHz.

Computation of *j*-factors

The DNA model and the computational procedure used here to calculate *j*-factors for a particular set of the conformational parameters of the oligonucleotide are described in the Appendix.

Two programs used in the work for the extraction of the set of *j*-factors from the distribution of ligation products and the calculation of *j*-factors for a particular set of conformational parameters are available on the authors' web site, http://crab.chem.nyu.edu/jfactor/. The programs are written in C.

RESULTS

General description of the method

Suppose we have identical oligonucleotides with two complementary cohesive ends, so they can be ligated into multimers. These oligonucleotides can be bound by proteins or ligands, or they can contain irregular structural elements. We wish to obtain quantitative information about their conformational properties from the efficiency of cyclization of their multimers. The cyclization efficiency is specified by *j*-factors, which can be measured directly for DNA molecules (Shore et al., 1981; Shore and Baldwin, 1983). The *j*-factor is defined as the ratio of two equilibrium constants, the equilibrium constant for cyclization of a chain with cohesive ends, K_1, and the bimolecular equilibrium constant for the association of two distinct half molecules, K_2 (Jacobson and Stockmayer, 1950):

$$j = K_1/K_2. \tag{1}$$

Each of these half molecules has one cohesive end and one blunt end, which does not participate in the joining reaction. Eq. 1 means that *j*-factor equals the effective concentration of one end of the chain in the vicinity of the other end in the appropriate angular and torsional orientation. Joining the cohesive ends is a slow process, whose rate is not limited by the rate at which these ends diffuse (Wang and Davidson, 1968); therefore, the *j*-factor also can be expressed as the ratio of the corresponding kinetic constants of irreversible ligation (Shore et al., 1981; Shore and Baldwin, 1983). On the other hand, the value of the *j*-factor is completely defined by the conformational parameters of the DNA fragment: the minimum energy conformation of its axis, the bending rigidity of the fragment, its total equilibrium twist, and its torsional rigidity. The value of the *j*-factor can be computed if these parameters are known (Hagerman and Ramadevi, 1990; Koo et al., 1990). Thus, *j*-factors can be considered as a bridge between experimentally measurable properties and the basic conformational parameters of DNA fragments.

The values of *j*-factors can be measured directly for each of the considered multimers (Koo et al., 1990), but this requires substantial experimental efforts. It is more attractive to extract the values of *j*-factors for the entire set of multimers from one multimerization-cyclization experiment. Under certain conditions that must be satisfied in the experiment (Crothers et al., 1992; Livshits and Lyubchenko, 1994) the process can be described by the following system of kinetic equations:

$$\begin{aligned} \frac{dL^n}{dt} &= \frac{1}{2} k_2^* \sum_{m=1}^{n-1} L^m L^{n-m} - k_2^* L^n \sum_{m=1}^{\infty} L^m - k_1^n L^n \\ \frac{dC^n}{dt} &= k_1^n L^n \end{aligned} \tag{2}$$

with the initial conditions

$$L^1(0) = M_0, \quad L^n(0) = 0 \text{ for } n > 1, \quad C^n(0) = 0.$$

Here, $C^n(t)$ and $L^n(t)$ are the concentrations of circular and linear *n*-mers at moment *t*, k_2^* is the rate constant of joining two linear fragments of different lengths together, k_1^n is the rate constant of the *n*-mer cyclization. The expressions in Eq. 2 assume that the rate of ligating two fragments together, k_2^*, does not depend on their sizes (Crothers et al., 1992). The values of k_2^* and k_2, the rate constant of ligating two distinguishable half molecules, are related by equations

$$k_2^* = 4k_2 \tag{3a}$$

for molecules with identical cohesive ends, and

$$k_2^* = 2k_2 \tag{3b}$$

for the molecules with different cohesive ends (Shore and Baldwin, 1983; Taylor and Hagerman, 1990). Using $j^n = k_1^n/k_2$ and Eqs. 2 and 3, one can obtain the distributions of linear and circular products for any set of j^n, M_0, k_2, and reaction time *t* (Fig. 2 *A*). However, we need to solve the reverse problem, to extract *j*-factors from the measured distribution of the reaction products. This problem is the

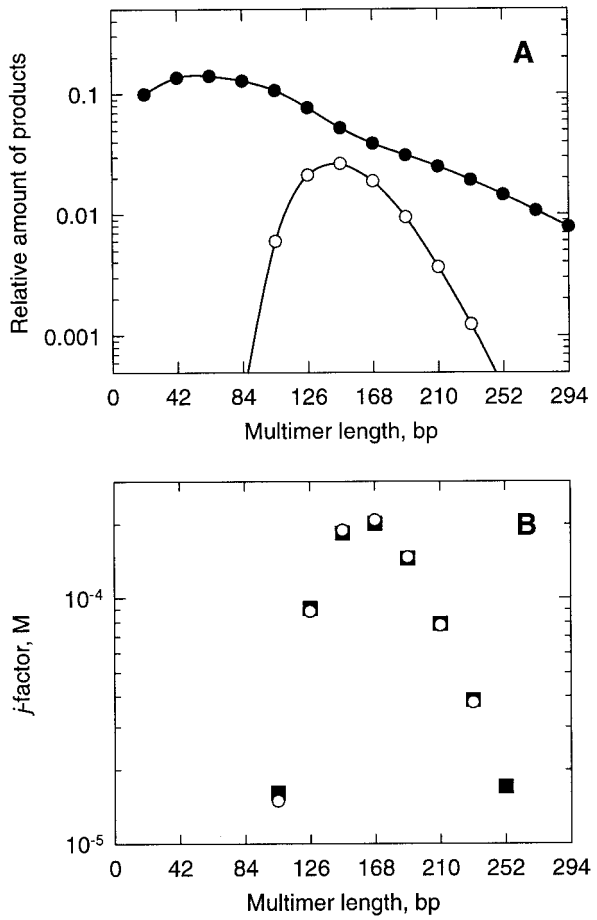

FIGURE 2 Determination of *j*-factors from the distributions of linear and circular multimers for the computer experiment. The distributions of C^n (*hollow circles*) and L^n (*filled circles*) shown in *A* were obtained by numerical solution of Eqs. 2 and 3 for $M_0 = 10^{-4}$ M, and a set of *j*-factors which was calculated for multimers of 21-bp fragment with the intrinsic bend of 36°, DNA persistence length of 50 nm, and helical repeat of 10.3 bp, and the torsional rigidity, *C*, of $2 \cdot 10^{-19}$ erg · cm. The set of *j*-factors was then reconstructed from the distributions of C^n and L^n by the procedure described in the text. The loaded (*filled squares*) and reconstructed (*hollow circles*) values of *j*-factors are shown in *B*.

subject of the first part of our analysis. A simple method to determine the conformational parameters of the monomers from the set of j^n will be described in the second part of the paper.

Extraction of *j*-factors

The first attempt to extract j^n from the product distributions was performed by Livshits and Lyubchenko (1994). They found approximate solution of Eqs. 2 for the case when the cyclization can be considered negligible with respect to multimerization. It follows from their analysis that at the very beginning of the reaction

$$j^n \cong B\frac{nC^n}{L^n}, \tag{4}$$

where B depends on M_0 and on the extent of the reaction and does not depend on n. Analyzing numerical solutions of Eqs. 2 and 3 we unexpectedly found that this proportionality is valid for a much broader time interval. Eq. 4 was verified for a wide range of the initial concentrations, M_0, and for *j*-factors corresponding to different bend angles and bending rigidities. We found that Eq. 4 holds with good accuracy while the amount of circles is less than the amount of linear multimers of the same length,

$$C^n \le L^n \tag{5}$$

for all n. The coefficient B can be found by successive approximations which minimize the difference between the experimentally measured and numerically calculated distributions of C^n and L^n. During this procedure we vary the value of B and the reaction time, t.

The first approximation for B, B_1, is chosen in such a way that the largest value of j^n obtained from Eq. 4 is larger than any reasonable expectation (we assumed that j^n cannot be larger than 10^{-2} M). The first approximation for j^n is obtained by substituting this value of B_1 and experimentally measured values of C^n and L^n into Eq. 4. For a particular set of j^n and M_0, the functions $C^n(t)$ and $L^n(t)$ depend on t only. Thus, by numerical solution of Eqs. 2 and 3, we can find t_1, which minimizes the sum $\Sigma_n(L^n(t_1) - L^n)^2$, and t_2, which minimizes the sum $\Sigma_n(C^n(t_2) - C^n)^2$. We have checked that each of the two sums has only one minimum. Comparison of t_1 and t_2 is used to choose the next approximation for B, B_{i+1}:

$$B_{i+1} = B_i - B_1/2^i \quad \text{if} \quad t_2 < t_1$$

$$B_{i+1} = B_i + B_1/2^i \quad \text{if} \quad t_2 > t_1$$

The iterations are stopped when $|B_{i+1}/B_i - 1| < 0.01$. The last condition also means that the final values of t_1 and t_2 are also very close to each other.

We tested the procedure by applying it to simulated experimental data (Fig. 2 *A*). These data, the sets of C^n and L^n, were obtained by numerical solution of Eqs. 2 and 3 for a chosen set of j^n. The advantage of the simulated data is that the desired solution, which has to be obtained in the extraction procedure, is known precisely. Fig. 2 *B* shows the result of such a test. We performed similar tests for very different sets of j^n and found that the difference between extracted and loaded values of j^n does not exceed 10% if condition 5 holds. Since the values of j^n vary by a few orders depending on the parameters of the monomer units, we consider that this accuracy is sufficient for our purposes.

Although condition 5 is the only one necessary to apply the extraction procedure, there is also a natural restriction on the dynamic range of C^n and L^n values, such that the measurements can be made with reasonable accuracy. Use of phosphoimaging allows us to extend the dynamic range to three orders of magnitude. This is sufficient for the current approach if we are careful to restrict the range of

$C^n(t)$ and $L^n(t)$ as much as possible. This can be achieved by choosing M_0 close to the largest j^n and by restricting the extent of the reaction, so that 10–20% of the monomers remains at the end of the experiment.

Determination of structural parameters from a set of *j*-factors

In the preceding part we described a method to obtain the set of *j*-factors for multimers of a short oligonucleotide. Here, we consider how to obtain conformational information from these *j*-factors. Crothers and coworkers were the first group to solve this problem (Koo et al., 1990; Crothers et al., 1992). Their approach was based on direct optimization of the model parameters to fit the measured set of *j*-factors. They used a Monte Carlo program that calculates a *j*-factor for particular parameters of DNA fragment by simulating an equilibrium set of fragment conformations. They were first who determined the bending of the DNA axis associated with A-tracts (Koo et al., 1990). This approach requires extensive computations and a good initial approximation of the conformational parameters. We have tried to find a simpler way of solving the problem. We performed computations for a large set of conformational parameters and generalized the results in the form of empirical relationships between the distribution of multimer *j*-factors and the conformational parameters of the oligonucleotide. These relationships allow one to obtain a good estimate of the conformational parameters without any extra simulations. Our DNA model used to derive the empirical equations is essentially identical to one used by others (Hagerman and Ramadevi, 1990; Koo et al., 1990); however, we designed a new Monte Carlo procedure for the efficient calculation of *j*-factor corresponding to a given set of conformational parameters that specify a DNA fragment. The model and the Monte Carlo procedure are described in the Appendix.

We assume that there is a given set of *j*-factors for different multimers of an oligonucleotide whose length corresponds to approximately two turns of the DNA double helix. In the context of our DNA model, the values of the *j*-factors are defined by four conformational parameters of the oligonucleotide: i) the equilibrium radius of curvature, ρ; ii) the DNA Kuhn statistical length, b; iii) the DNA helical repeat, h; and iv) the torsional rigidity of the double helix, C. For further analysis we will describe the set of *j*-factors, j^n, by an interpolating function of DNA length, S. Note that this function, $j(S)$, is equal to the values of *j*-factors only for DNA lengths which are integral numbers of multimer length; its intermediate values have no physical meaning. Because $j(S)$ always has a maximum for the range of S that corresponds to the product distribution of the multimerization-cyclization reaction (see Fig. 2 *B*), we can describe $j(S)$ by three parameters: the length corresponding to the maximum, S_{max}, the width of the function at half-height $\Delta S_{1/2}$, and the value at the maximum, $j(S_{max})$.

General analysis of a function of four variables usually requires enormous resources, but in our case simplifying circumstances help to handle the problem. In the following treatment we have accounted for the restricted range of possible values of the parameters, have used some approximations, and have applied a scaling treatment. In particular, we have considered only molecules with notable equilibrium curvatures.

First we studied the effect of torsional misfit of the multimer ends on $j(S)$. The simulation results show that the position of the maximum and the distribution width depend weakly on the value of h (Fig. 3). We always observed this behavior for $\rho < 0.2b$. Thus, we concluded that for molecules with notable bends, $j(S)$ can be reasonably approximated as the product of the bending and twisting components:

$$j(S) = j_0(S; \rho, b) j_{Tw}(S; h, C). \quad (8)$$

Here, $j_0(S; \rho, b)$ is the *j*-factor that does not account for the twisting alignment, and $j_{Tw}(S; h, C)$ is the contribution from the twisting alignment of the chain ends, which depends weakly on S. Equation 8 allows us to analyze $j_0(S; \rho, b)$ and $j_{Tw}(S; h, C)$ independently.

Note that $j_0(S; \rho, b)$ depends on ρ and b separately, but we can eliminate one of these variables if we measure all lengths in units of b. The new dimensionless $j_0^*(S/b, \rho/b)$ and $j_0(S; \rho, b)$ are connected by simple equation:

$$j_0^*(S/b, \rho/b) = b^3 N_A j_0(S; \rho, b), \quad (9)$$

where N_A is Avogadro's number. We simulated $j_0^*(S/b, \rho/b)$ for different values of ρ/b. Fig. 4 summarizes the simulation

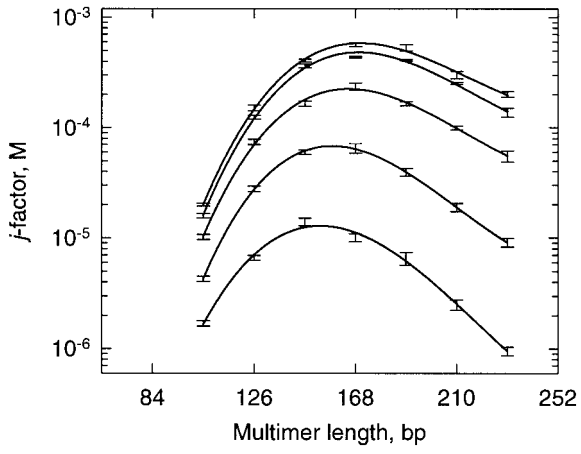

FIGURE 3 The dependence of the calculated *j*-factors on the helical repeat of the oligonucleotide. The regression lines are drawn for *j*-factors obtained for multimers of 21-bp fragments with helical repeats equal to 10.50, 10.42, 10.34, 10.25, and 10.17 bp, from top to bottom. The other conformational parameters are: $C = 2 \cdot 10^{-19}$ erg · cm, $\rho = 11.5$ nm, $b = 100$ nm.

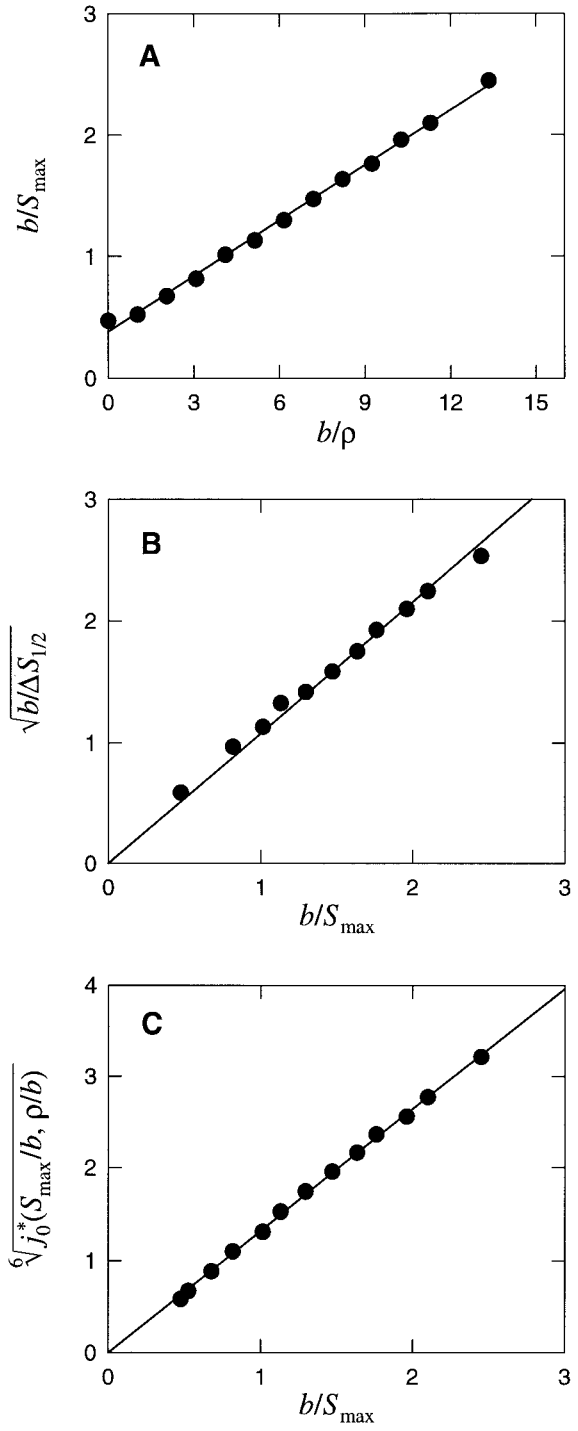

FIGURE 4 The effect of DNA curvature on the characteristics of the *j*-factor. All dependencies are shown in linearized forms. (*A*) The dependence of the maximum position of j_0^*, S_{max}/b, on the equilibrium curvature of the fragments. The DNA curvature was varied in the range $0 < b/\rho < 13.5$. (*B*) The relationship between the maximum position and the width at half-height of j_0^*. (*C*) The relationship between the maximum position and the maximum value of j_0^*.

results, which can be also expressed by empirical equations:

$$b = 0.8\,\frac{S_{max}^2}{\Delta S_{1/2}} \tag{10}$$

$$\frac{1}{S_{max}} = \frac{1}{2\pi\rho} + \frac{1}{2b} \tag{11}$$

$$j_0(S_{max};\,\rho,\,b) = \frac{5.3}{N_A}\,\frac{b^3}{S_{max}^6} \tag{12}$$

Equations 10 and 11 are sufficient for the estimation of b and ρ from the values of S_{max} and $\Delta S_{1/2}$. These two parameters of the oligonucleotide are the most interesting ones, and in many cases we can stop the analysis at this point. If we want to determine the torsional parameters, h and C, we have to compare the largest j-factor in the experimental set with the maximum value of j_0 specified by Eq. 12. Their ratio is equal to $j_{Tw}(S_{max};\,h,\,C)$. The length is measured in centimeters in Eqs. 10–13.

Using computer simulations, we found that, for relatively short bent molecules, only one topoisomer makes a notable contribution to the cyclization of a particular multimer. This topoisomer is characterized by nearly planar conformations of the circles when each monomer makes an integral number of helical turns, even if the equilibrium twist of the monomer is not equal to an integer. Since the length of the monomer is usually equal to 21 bp and the helical repeat of DNA in solution is close to 10.5 bp (Wang, 1979; Shore and Baldwin, 1983), the required torsional deformation is not too large for this topoisomer. The bending energy of this topoisomer is less than that of the others since all its equilibrium bends are directed towards the center of the circle. We tested by simulations that the bending energy of other topoisomers is much higher than the energy of the major topoisomer. Since for nearly planar conformation the writhe is very close to zero, we know the value of torsional deformation in this conformation and can estimate $j_{Tw}(S;\,h,\,C)$ from the equilibrium distribution of twist in linear DNA (Klenin et al., 1991). It gives us

$$j_{Tw}(S;\,h,\,C) = \sqrt{\frac{2\pi C}{kTS}}\,\exp\left(-\frac{CS}{2kT}\left(\frac{2\pi\{l/h\}}{l}\right)^2\right) \tag{13}$$

where l is the length of the monomer; the braces indicate the difference with the nearest integer. There are two unknowns in Eq. 13, so we cannot determine both C and h from the equation. The value of h can be obtained from Eqs. 8, 12, and 13 if we know the value of C. Although C is known relatively well for the regular double helix, it can have different value for monomers with irregular structural elements.

To test the approach described above, we simulated j-factor distributions for several sets of ρ, b, and h, and then tried to reconstruct the parameters from the simulated j-factors. In these computer experiments and all other calculations in the work we kept the value of C constant ($C = 2 \cdot$

10^{-19} erg · cm). Table 1 illustrates the accuracy of the procedure for the simulated sets of *j*-factors. One can see that for the most cases the difference between desired parameters and those obtained does not exceed 10%.

Multimerization-cyclization of DNA fragments containing A-tracts

DNA intrinsic bends associated with four or more adenines in a row are well studied, and therefore we chose DNA fragments containing these A-tracts to test our approach. The values of j^n for the multimers of these fragments were measured directly, by cyclization of the multimers of different lengths (Koo et al., 1990; Crothers et al., 1992). Analysis of the *j*-factors in conformational terms allowed Crothers and his colleagues to determine that the bending angle associated with A-tract equals 18° (Koo et al., 1990). Rivetti et al. (1998) recently studied the effect of A-tracts on DNA conformation by atomic force microscopy and found an angle of 13°.

We have performed a set of experiments with the DNA fragment shown in Fig. 1. The sequence of the fragment is similar to that used by (Koo et al., 1990). Fig. 5 shows the distribution of linear and circular products in a typical experiment. We used the measured product distributions to extract *j*-factors from the multimer distribution and its reaction time and then recalculated the distribution of C^n and L^n by substituting these values of *j*-factors and the reaction time into Eqs. 2. Comparison of the recalculated distributions with measured ones serves as an important control for the self-consistency of this analysis.

We found that the agreement between measured and recalculated C^n and L^n was always good when the extent of the reaction was sufficiently low that 10–20% of the monomers remained at the end (Fig. 5). The agreement declined when only 1–2% of the monomers remained in the reaction. Special computer experiments showed that the values of *j*-factors extracted from the experimental data are more sensitive to the errors in the measurements of C^n and L^n for

TABLE 1 Examples of reconstruction of the structural parameters from the computed *j*-factors

Actual values	S_{max}, nm	$\Delta S_{1/2}$, nm	$j(S_{max})$, M	Reconstructed values
ρ = 22.7 nm b = 100 nm h = 10.28 bp	79	47	$4.6 \cdot 10^{-7}$	ρ = 20.8 nm b = 106 nm h = 10.26 bp
ρ = 12.2 nm b = 100 nm h = 10.28 bp	56	24	$3.4 \cdot 10^{-5}$	ρ = 12.2 nm b = 106 nm h = 10.27 bp
ρ = 12.2 nm b = 100 nm h = 10.50 bp	61	39	$6.3 \cdot 10^{-4}$	ρ = 13.9 nm b = 102 nm $h \approx$ 10.5 bp
ρ = 12.2 nm b = 50 nm h = 10.28 bp	44	30	$4.2 \cdot 10^{-7}$	ρ = 12.1 nm b = 53 nm h = 10.27 bp
ρ = 6.8 nm b = 100 nm h = 10.28 bp	36	12	$2.1 \cdot 10^{-3}$	ρ = 7.3 nm b = 87 nm h = 10.31 bp

We used the torsional rigidity $C = 2 \cdot 10^{-19}$ erg · cm.

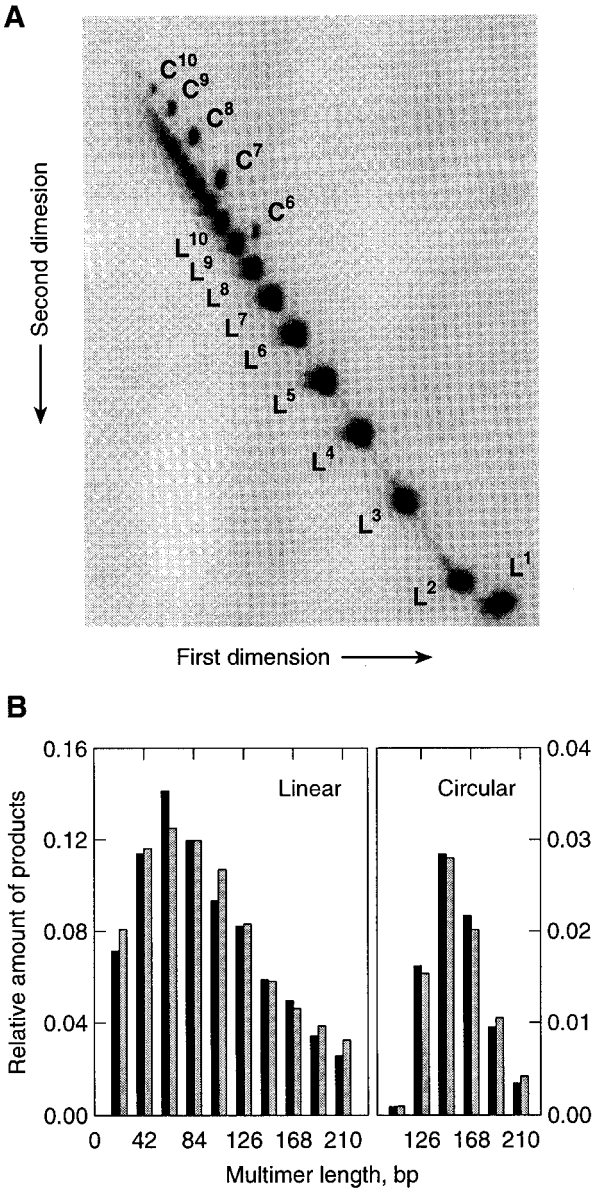

FIGURE 5 Product distribution in the multimerization-cyclization experiment with 21-bp fragment containing two A-tracts. (*A*) Two-dimensional gel-electrophoretic separation of the reaction products. The two directions of electrophoresis are indicated. The linear molecules form a diagonal on the gel, whereas the cyclic molecules form an arc above it. The cyclic products are double stranded circles; single-stranded circles are also visible on the original autoradiogram as an arc of very faint spots between the catenated and linear spots, but they are not visible in this reproduction. (*B*) The measured and reconstructed distributions of the circular and linear products. Both distributions were normalized to 1. The experimental data are shown in black and the reconstructed distributions in gray.

high extent of the reaction (data not shown). Similarly, small impurities in the samples would have a larger effect on the product distribution for deeper reactions. Thus we concluded that at least 5% of the monomers should present at the end of the reaction for successful extraction of *j*-factors. The experiments with 10–20% of the monomers remaining at the end of the reaction gave similar sets of extracted *j*-factors (Fig. 6).

We also tested that kinetics of the reaction follows Eqs. 2 over the broad time interval. We found the values of *j*-factors from the product distribution shown in Fig. 5 and substituted them into Eqs. 2 to calculate the entire kinetics. Fig. 7 shows how the fractions of selected linear and circular multimers change in the course of the reaction as a function of remaining monomers. Similar agreement was obtained for other products, which are presented in Fig. 5. The agreement declined only when less than 1% of monomers remained, probably because of some impurity of the material.

We have found that it is crucial to have nearly all 5′ ends of the nucleotides phosphorylated. This requirement is easy to understand, since multiple ligations are required for multimer formation, and a lack of phosphorylation for even a small percentage of the monomer ends can affect the distribution of the reaction products. We performed a computer simulation of the reaction with incomplete phosphorylation of the 5′ ends and found that the presence of 10% unphosphorylated 5′ ends on the monomers decreases the extracted values of *j*-factors by a factor of 1.5 to 2.5, depending on the extent of the reaction. To provide complete, close to 100%, phosphorylation we synthesized hairpin oligonucleotides of longer length, containing restriction sites near their ends. Digestion with the restriction enzyme was used to produce completely phosphorylated fragments (see Methods).

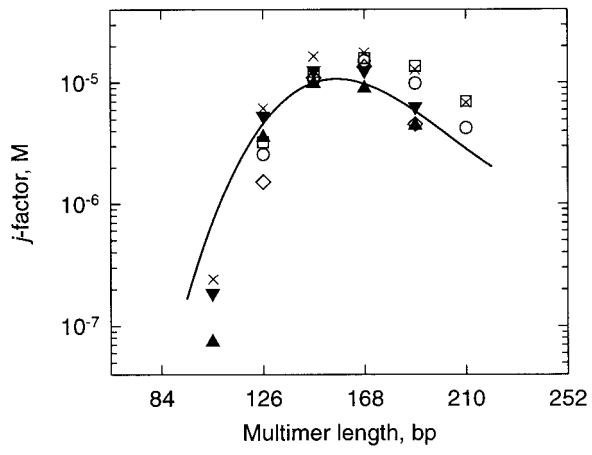

FIGURE 6 The sets of *j*-factors extracted from the multimerization-cyclization experiments with 21-bp fragment containing two A-tracts. Three independent experiments were performed (*crosses, filled symbols, and hollow symbols*). For two of them, data with different extents of ligation were used (shown by different filled and hollow symbols). From average values of these *j*-factors we found the following set of parameters: $b = 106 \pm 10$ nm, $\rho = 12.0 \pm 0.6$ nm, and $h = 10.18 \pm 0.02$ bp; the value of C was set equal to $2 \cdot 10^{-19}$ erg · cm. The *j*-factors computed for these parameters are shown by solid line.

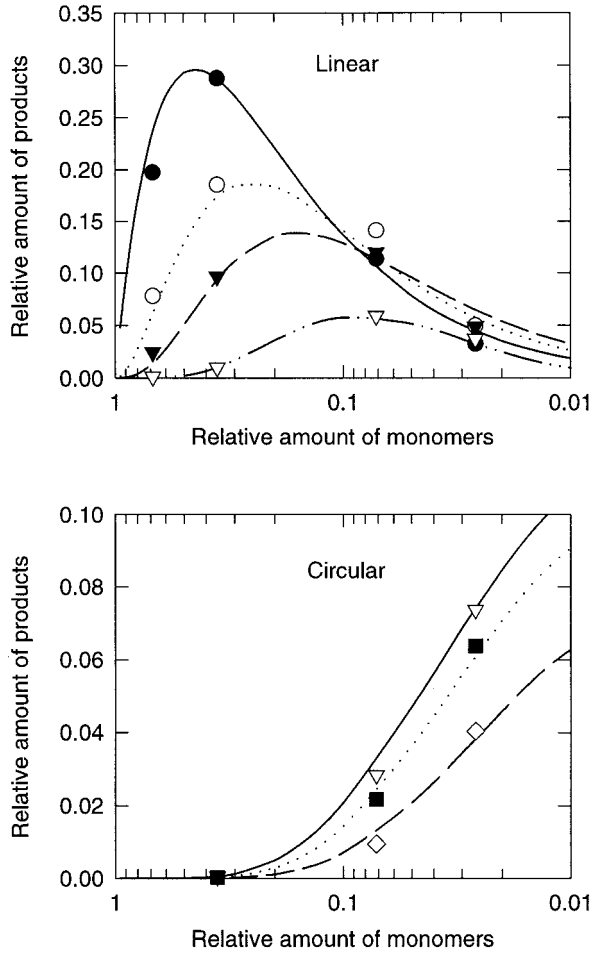

FIGURE 7 Comparison of simulated and measured kinetics of the multimerization-cyclization reaction. Change of the fractions of selected liner and circular multimers is shown as a function of the fraction remaining monomers. Experimental results are shown by symbols and corresponding theoretical dependencies are presented by lines. The data correspond to 2-mers (*filled circles*), 3-mers (*hollow circles*), 4-mers (*filled triangle*), 7-mers (*hollow triangle*), 8-mers (*filled squares*), and 9-mers (*hollow diamonds*). Similar agreement was obtained for other reaction products shown in Fig. 5.

The average set over a few extracted sets of *j*-factors was used to determine the values of ρ, b, and h using the procedure described above. The found values correspond to bend angle per A-tract of 17° ± 1°, DNA Kuhn statistical segment of 103 ± 10 nm, DNA helical repeat of 10.18 ± 0.02 bp per turn. We assumed in this analysis that DNA torsional rigidity is equal to $2 \cdot 10^{-19}$ erg · cm (see Appendix for details). Then we computed *j*-factors for different multimers specified by the calculated set of the conformational parameters. Fig. 6 shows the computed *j*-factors along with the values extracted from the experiment. The agreement between the two sets is reasonably good if we take into

account the experimental errors of the measured values of the *j*-factors. The value of the most interesting parameter, the equilibrium radius of curvature, ρ, corresponds to the equilibrium bend of 17° per A-tract, what is in good agreement with the result of Koo et al. (1990), 18° per A-tract.

DISCUSSION

We have developed and tested an approach that allows one to obtain conformational parameters of short bent DNA fragments from the distribution of linear and circular products obtained in the multimerization-cyclization reaction. The method allows one to evaluate the value of total bend angle and bending rigidity of the fragment and, with some restrictions, helical repeat and torsional rigidity. The procedure has been tested on simulated experimental data and on DNA fragments containing A-tracts, which had been studied earlier. The tests showed good accuracy of the procedure.

In the first step of our approach the values of *j*-factors for different multimers of the fragment are determined from the product distribution in the multimerization-cyclization experiment. Several experimental conditions must be satisfied to make the determination successful.

1. The amount of circular multimers should not exceed the amount of the corresponding linear multimers at the end of the reaction.

2. The extent of the reaction should not be too high, so at least 5% of the monomers remain at the end of the reaction.

3. It is important that a very high proportion of the monomers, 97% or perhaps more, contain phosphorylated sticky ends that are available for ligation.

In practice, conditions 1 and 2 can be satisfied simultaneously only if the initial concentration of monomers in the ligation reaction and maximum *j*-factor for the multimers have the same order of values. If the bend angle per fragment is very large, the corresponding *j*-factors will be very large as well, and it is difficult to work with such a high concentration of the monomers. The problem can be overcome by reducing the number of bends per original fragment. Crothers et al. (1992) have noted previously that sufficiently short cohesive ends, 1 to 2 bp in length, are another experimental requirement for the successful extraction of conformational parameters.

In the second step we analyze the distribution of *j*-factors in conformational terms. We specify the distribution of *j*-factors by only three values: the position of the maximum, the magnitude at the maximum, and the width of the distribution. Thus, we could not determine all four conformational parameters of the fragment unambiguously. It turned out, however, that these three parameters of the *j*-factor distribution uniquely define the two most interesting conformational characteristics, equilibrium bend angle and persistence length of the fragment. To determine the radius of torsion of the fragment, one must choose arbitrarily the value of DNA torsional rigidity. We used $C = 2 \cdot 10^{-19}$ erg · cm through this work, which was determined in cyclization experiments for small DNA molecules with one single-stranded nick (Shore and Baldwin, 1983; Taylor and Hagerman, 1990).

It should be noted that the angle between the monomer ends does not specify minimum energy conformation of its axis completely. Particularly, the oligonucleotide can have more than one bend and its conformation can be nonplanar. The method is equally applied to such cases and one can obtain correct value of the angle between monomer ends. However, the detailed structure of the monomer cannot be resolved by the approach since this detailed structure does not affect, practically, the distribution of *j*-factors of the multimers. This point is illustrated in Fig. 8, which compares two monomers with different detailed minimum energy conformations but the same sets of *j*-factors. Similarly, the persistence length obtained in the analysis corresponds to the average bending rigidity along the monomer, which could have a hinge at some position.

Our analysis assumes the existence of a notable intrinsic bend in the monomer unit and elasticity of its deformations that obeys Hooke's Law. Such a model was considered before by Crothers and coworkers (Koo et al., 1990; Crothers et al., 1990) and by Livshits (1996). Koo et al. (1990) solved the same problem by direct optimization of the model parameters to fit the measured set of *j*-factors. We instead found empirical equations that solve the problem without further computations. Although our solution contains some approximations, its results are rather accurate. The relatively small error related with the analysis is a

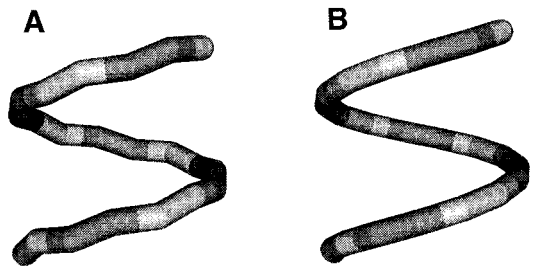

FIGURE 8 Two structures formed by decamers of 21-bp oligonucleotides with different distributions of bends along their contour. The structures correspond to minimum energy conformations. (*A*) Two bends of the first oligonucleotide, 28° each, are located in different planes, with 66° angle between them, so the minimum energy conformation of the monomer is not flat. The resulting angle between the oligonucleotide ends equals 49°. Torsional separation between adjacent bends in the helix alternates between 426° and 306°. (*B*) The bends of the second oligonucleotide, 24° each, are located in nearly parallel planes, with 6° angle between them. The resulting angle between the oligonucleotide ends equals 48°. Torsional separation between all adjacent bends in the second helix equal 326°. The helical structures formed by the oligonucleotides have nearly identical radiuses and pitches and thus *j*-factors of the corresponding multimers are nearly identical for both oligonucleotides. The bend angles between the monomer ends are nearly equal in both cases and can be determined unambiguously by the approach, although the angle does not specify a detailed structure of the monomers.

reasonable price for its simplicity, since the optimization requires extensive computations and it is difficult to estimate its accuracy. The analysis can be applied equally to DNA fragments that contain specific chemical modifications, are bound with ligands or proteins, or have irregular structural elements. We did not account, however, for excluded volume effects, which may be needed for some proteins whose size is comparable with the size of DNA monomers.

The analysis of the *j*-factors is based on the fact that they can be expressed, to a reasonable approximation, as the products of two terms, accounting separately for the bending and torsional deformations of the multimers. The possibility of performing this factorization results from the near-integral number of helical turns in the monomers considered in this work. Our simulations suggest this approximation works well if the pitch of the minimum energy helix formed by multimers (Fig. 8) is less than its diameter. The most interesting term, which specifies the bending component of the *j*-factor, is defined by Eqs. 10–12. We tested the equations for a broad range of intrinsic curvatures, including the case of intrinsically straight DNA. In the limit of small values of ρ, the equations are in quantitative agreement with the theoretical calculations of Livshits (1996), and for intrinsically straight DNA they are consistent with the results of Shimada and Yamakawa (1984).

Two restrictions should be satisfied for successful applications of the approach.

1. The fragment has to contain close to integer number of helix turns. Our analysis works properly if the number of turns differ by ± 0.1 from an integer; DNA fragments of 21 bp in length, which are usually used in the experiments, satisfy the condition.

2. Our experience shows that successful application of the approach is restricted by the condition the bend angle is in the interval from 9° to 26° per DNA helical turn. The intrinsic curvature that corresponds to smaller angles is hardly noticeable against thermal bending fluctuations, and larger values of the angle may result in very large values of the *j*-factor (see above).

The simplest way to analyze the multimerization-cyclization data assumes that the equilibrium bend angle in the DNA fragment equals $360°/n$, where n specifies the distribution maximum for circular multimers (see (Lyubchenko et al., 1993), for example). Our results show that this method essentially overestimates the bend angle. Indeed, from the distribution shown in Fig. 3 one could conclude that the equilibrium bend angle equals $360°/7 = 51°$, whereas our analysis gives 34° per fragment containing two A-tracts. We have analyzed numerous computer experiments and have concluded that such discrepancies are common, although their exact values vary. The effect has entropic origin: for shorter molecules larger fraction of conformations has juxtaposed ends, and it increases their *j*-factor. For the same reason maximal *j*-factor of intrinsically straight DNA corresponds to DNA length about 500 bp rather than infinitely long molecules. Also shorter linear multimers which serve as substrates for the corresponding circles, present at larger amounts than longer multimers, especially at the earlier stages of the reaction. It also results in larger amount of shorter circles. The issue was discussed in more detail by Hagerman and Ramadevi (1990).

APPENDIX: CALCULATION OF *j*-FACTOR FOR DNA FRAGMENT

DNA model

A DNA molecule of contour length S is modeled as a chain of N straight segments of length $l_0 = S/N$. The segment orientations are specified by orthogonal triplets of unit vectors, $(\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)$, attached to the beginning of each segment. The vector $\mathbf{z}_i$ is oriented along the axis of the segment. The orientations of the chain ends are given by the triplet $(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1)$ and the additional virtual triplet $(\mathbf{x}_{N+1}, \mathbf{y}_{N+1}, \mathbf{z}_{N+1})$ attached to the end of the last segment. These end triplets must coincide when the molecule is closed.

The equilibrium bend and twist between the segments $i-1$ and i are specified by three Euler angles, $(\psi_i, \sigma_i, \theta_i)$, so that the minimal energy orientation of the segment i for the given actual orientation of the segment $i-1$ is calculated as

$$\begin{aligned} \mathbf{x}_i^0 &= \mathbf{T}_i(\psi_i, \sigma_i, \theta_i)\mathbf{x}_{i-1} \\ \mathbf{y}_i^0 &= \mathbf{T}_i(\psi_i, \sigma_i, \theta_i)\mathbf{y}_{i-1} \\ \mathbf{z}_i^0 &= \mathbf{T}_i(\psi_i, \sigma_i, \theta_i)\mathbf{z}_{i-1}, \end{aligned} \tag{A1}$$

where $\mathbf{T}_i(\psi_i, \sigma_i, \theta_i)$ is the Euler rotation matrix (Korn and Korn, 1961) defined by the three angles; ψ_i specifies the direction of the equilibrium bend, or the angle between the axis of the bend rotation and $\mathbf{x}_i$, σ_i is the magnitude of the bend, and the sum $\psi_i + \theta_i$ defines the equilibrium twist.

The angles of bending and twisting that specify the deflection of an actual orientation of triplet i from its minimal energy orientation, $\Delta\sigma_i$ and $\Delta\psi_i + \Delta\theta_i$, are calculated as

$$\begin{aligned} \Delta\sigma_i &= \arccos(\mathbf{z}_i^0 \cdot \mathbf{z}_i), \\ \Delta\psi_i + \Delta\theta_i &= \arctan\left(\frac{\mathbf{y}_i^0 \cdot \mathbf{x}_i - \mathbf{x}_i^0 \cdot \mathbf{y}_i}{\mathbf{x}_i^0 \cdot \mathbf{x}_i + \mathbf{y}_i^0 \cdot \mathbf{y}_i}\right), \end{aligned} \tag{A2}$$

where the triplet $(\mathbf{x}_i^0, \mathbf{y}_i^0, \mathbf{z}_i^0)$ is defined by Eq. A1. Note that while $(\psi_i, \sigma_i, \theta_i)$ specify the minimal energy orientation of the segment i relative to the actual orientation of the segment $i-1$, $(\Delta\psi_i, \Delta\sigma_i, \Delta\theta_i)$ define the deflection of the actual orientation of the segment i relative to its own minimal energy orientation. In this work we assume isotropic bending properties, so the angles specify the elastic energy in the junction between segments $i-1$ and i as:

$$E_i = \frac{kTb}{4l_0}(\Delta\sigma_i)^2 + \frac{C}{2l_0}(\Delta\psi_i + \Delta\theta_i)^2, \tag{A3}$$

where kT is the Boltzmann factor, b is the Kuhn statistical length, and C is the torsional rigidity of the double helix (Shimada and Yamakawa, 1984). The total energy of the model chain is the sum of energies over all N junctions. Excluded volume interactions are not included in the model, because it is being applied here to relatively short DNA molecules.

In this work, we focus on the *j*-factors of DNA molecules obtained by the multimerization of identical oligonucleotides, so that the equilibrium conformation of large linear multimer of the oligonucleotide corresponds to a helix in large scale. Although equilibrium conformation of one

oligonucleotide only approximately corresponds to a smooth arc of the helix, we will describe the conformation by such a way. This is reasonable approximation for short oligonucleotides. We have to make this approximation since there is no chances to extract more detailed conformational description of the oligonucleotides from the experimental data. Thus we specified the conformation of the oligonucleotide axis by the radius of curvature, ρ, of a flat arc. The equilibrium bend angle σ between the segments of the model chain can be expressed as $\sigma = l_0/\rho$. The second parameter of the equilibrium conformation of the oligonucleotide is its fractional twist, ΔTw. The value of ΔTw can be calculated from the DNA helical repeat, h, as $\Delta Tw = \{l/h\}$, where the braces indicate the difference with the nearest integer. The length of the oligonucleotides, l, is commonly chosen so that it contains close to an integer number of turns. It results in small pitch of the helix in comparison with its diameter.

Note that one can eliminate ψ_i from consideration by choosing $\mathbf{x}_i$ in the direction of the equilibrium bend axis. The twist angle θ is specified by $\theta = 2\pi\Delta Tw(l_0/l)$ in this case.

Metropolis Monte Carlo procedure

We used the Metropolis Monte Carlo procedure to obtain the equilibrium set of chain conformations. Two types of displacements were applied repeatedly to randomly chosen subchains, each consisting of a random number of adjacent segments.

The first type of displacement is a rotation around a randomly directed axis of all the subchain triplets from i to $N + 1$ by an angle φ, so the subchain rotates as a rigid body. The random value of φ is uniformly distributed over the range $[-\varphi_0, \varphi_0]$. This displacement changes the mutual orientation of the triplets $i - 1$ and i, and the energy E_i given by Eq. A3. In the special cases where the axis of the rotation coincides with $\mathbf{z}_i^0$ and $\mathbf{z}_i$, the rotation only changes the angles $\Delta\psi_i$ or $\Delta\theta_i$, respectively. When the axis of rotation is perpendicular to both $\mathbf{z}_i^0$ and $\mathbf{z}_i$, the rotation only changes $\Delta\sigma_i$. In the general case, the rotation about a random axis changes all three of these angles. Note that any conformation of the model chain can be obtained using a finite number of displacements of this type only, which is a necessary requirement of Metropolis procedure.

The second type of displacement is a crankshaft rotation (Klenin et al., 1991), when the triplets from i to j are rotated around the line connecting the ends of the subchain by an angle Φ. The value of Φ is uniformly distributed over the range $[-\Phi_0, \Phi_0]$. This displacement changes the energies E_i and E_{j+1}. This displacement never changes the end-to-end distance of the chain, but it improves the biased sampling for the calculation of the j-factor (see below).

The trial rotations are accepted or rejected according to the standard rule based on the energy test (Metropolis et al., 1953). The amplitudes of the rotations, Φ_0 and φ_0, are adjusted so that approximately half of the trial rotations are accepted. The principle of microscopic reversibility of the Metropolis procedure is proved by direct calculation of the probabilities of particular trial displacements.

Calculation of the *j*-factor

The j-factor is related to the probability that the chain ends are juxtaposed and aligned. The end-to-end distance of the model chain is calculated as

$$\mathbf{r} = l_0 \sum_{i=1}^{N} \mathbf{z}_i \tag{A4}$$

The mutual alignment of chain ends is given by the bend angle, β, and the twist angle, τ, between the end triplets. These angles were calculated using equations similar to Eq. A2.

$$\beta = \arccos(\mathbf{z_1} \cdot \mathbf{z}_{N+1}),$$
$$\tau = \arccos\left(\frac{\mathbf{x}_1 \cdot \mathbf{x}_{N+1} + \mathbf{y}_1 \cdot \mathbf{y}_{N+1}}{1 + \mathbf{z}_1 \cdot \mathbf{z}_{N+1}}\right) \tag{A5}$$

By definition (Crothers et al., 1992), the j-factor is given by

$$j = \frac{10^{24}}{N_A} \lim_{\substack{r_0\to 0 \\ \beta_0\to 0 \\ \tau_0\to 0}} \left(\frac{3P(r_0)}{4\pi r_0^3} \frac{2P(\beta_0)}{1 - \cos\beta_0} \frac{\pi P(\tau_0)}{\tau_0}\right) \tag{A6}$$

where N_A is Avogadro's number, $P(r_0)$ is the probability that $r < r_0$, $P(\beta_0)$ is the probability that $\beta < \beta_0$ under the condition that $r < r_0$, $P(\tau_0)$ is the probability that $\tau < \tau_0$ under the condition that $r < r_0$ and $\beta < \beta_0$. In this expression the angular parameters are in radians, the end-to-end distance is in nanometers, and the j-factor is in moles per liter, M.

To find the j-factor we would have to calculate the fraction of the conformations with juxtaposed and aligned ends in the equilibrium set generated by the MC procedure. However, if the probability of these conformations is very small, they might not appear, even in a very large equilibrium set. To overcome the problem, previous investigators employed special approaches (Levene and Crothers, 1986; Hagerman and Ramadevi, 1990). Here, we have developed a new biased MC approach to the problem.

To calculate $P(r_0)$, we can choose a sequence of distances $r_0 < r_1 < \ldots < r_n$, where r_n is larger than or equal to the chain contour length. For each r_i we can define the probabilities, $P(r_i)$, of conformations with $r < r_i$ among all possible conformations. We can also define the conditional probability, $P(r_i|r_{i+1})$, of conformations with $r < r_i$ in the subset of conformations with $r < r_{i+1}$. Since $P(r_i) = P(r_i|r_{i+1})P(r_{i+1})$ and $P(r_n) = 1$, $P(r_0)$ can be found as

$$P(r_0) = \prod_{i=0}^{n-1} P(r_i|r_{i+1}). \tag{A7}$$

$P(\beta_0)$ can also be expanded into a product of conditional probabilities, $P(\beta_i|\beta_{i+1})$, for a chosen sequence of angles, $\beta_0 < \beta_1 < \ldots < \pi$. The same can be done with $P(\tau_0)$.

The magnitudes of the conditional probabilities can be controlled by changing the sequences of the distances and the angles. We have estimated that the efficiency of the j-factor calculation is maximized when the conditional probabilities are close to 0.2. The error in the calculations increases by about 20% if the conditional probabilities vary between 0.05 and 0.5.

Large values of $P(r_i|r_{i+1})$, $P(\beta_i|\beta_{i+1})$, and $P(\tau_i|\tau_{i+1})$ can be calculated efficiently and accurately using the following MC procedure. The simulation starts from a conformation with $r = 0$, $\beta = 0$, $\tau = 0$. First, a set of equilibrium conformations with $r < r_0$, $\beta < \beta_0$, and $\tau < \tau_1$ is obtained by rejecting any conformation with $r > r_0$, $\beta > \beta_0$, and $\tau > \tau_1$. The fraction of the conformation with $\tau < \tau_0$ in this set gives the value of $P(\tau_0|\tau_1)$. To calculate the next conditional probability, $P(\tau_1|\tau_2)$, a new set of conformations with $r < r_0$, $\beta < \beta_0$, and $\tau < \tau_2$ is generated. Then the procedure is repeated for all the conditional probabilities, gradually loosening the conformational restrictions according to the chosen sequences of angles and distances.

We tested that the limiting value of the probability density in the Eq. A6 can be accurately found using $r_0 = 1$ nm, $\beta_0 = 0.1$, $\tau_0 = 0.1$ (data not shown). The sequences of the distances and the angles were chosen so that $r_{i+1}/r_i = \sqrt{2}$, $\beta_{i+1}/\beta_i = 2$, and $\tau_{i+1}/\tau_i = 2$. We used the same sequences throughout this work.

We tested that simulations with segment length $l_0 = 3.57$ nm (10.5 bp) give nearly the same values of the j-factor as simulations with a segment

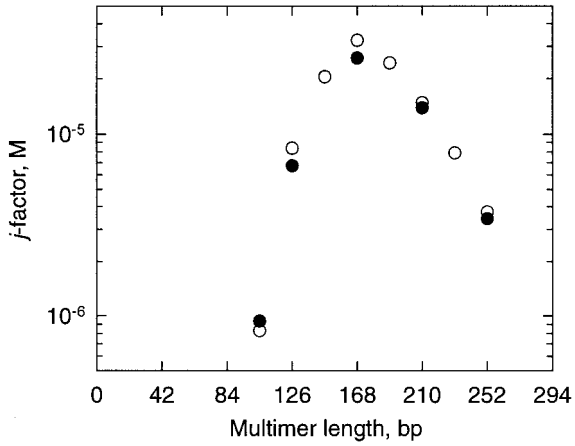

FIGURE A1 Dependence of the calculated *j*-factors on the length of the straight segment of the DNA model. The *j*-factors correspond to multimers of 21-bp oligonucleotides with two 16° bends separated by 10.5 bp (ρ = 12.8 nm). The hollow symbols represent the set of *j*-factors simulated with l_0 = 3.57 nm; the filled symbols are the *j*-factors obtained with l_0 = 0.34 nm. An identical set of parameters was used for both cases: the Kuhn statistical length of 100 nm, the helical repeat of 10.25 bp, and the torsion rigidity of $2 \cdot 10^{-19}$ erg · cm. The statistical error of the computations is about the symbol size.

length of 1 bp, which require much longer simulation times (Fig. A1). Therefore, throughout the rest of the simulations we used l_0 = 3.57 nm, half of the 21-bp fragment usually used in the experiments.

Our biased Monte Carlo approach can accelerate the computations by many orders of magnitude in comparison with unbiased Metropolis sampling. The calculation of the *j*-factor for a particular chain length with 5% standard error takes less than 1 h on a Pentium II 266 MHz processor. It seems that our approach is more efficient than the dimerization method (Hagerman and Ramadevi, 1990; Crothers et al., 1992). Similar computational efficiency can probably be achieved with the umbrella sampling method (McCammon and Harvey, 1987), but it is much easier to apply our approach.

We thank Drs. N. Geacintov, M. Livshits, and Y. Lyubchenko for helpful discussions. This research has been supported by grants GM-54215 (AVV) and GM-29554 (NCS) from the National Institute of General Medical Sciences, N00014-98-1-0093 from the Office of Naval Research (NCS), NSF-CCR-97-25021 from the National Science Foundation/DARPA (NCS) and F30602-98-C-0148 from the Information Directorate of the Air Force Research Laboratory located at Rome, NY (NCS).

REFERENCES

Balagurumoorthy, P., H. Sakamoto, M. S. Lewis, N. Zambrano, G. M. Clore, A. M. Gronenborn, E. Appella, and R. E. Harrington. 1995. Four p53 DNA-binding domain peptides bind natural p53-response elements and bend the DNA. *Proc. Natl. Acad. Sci. USA.* 92:8591–8595.

Caruthers, M. H. 1985. Gene synthesis machines: DNA chemistry and its uses. *Science.* 230:281–285.

Crothers, D. M., J. Drak, J. D. Kahn, and S. D. Levene. 1992. DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* 212:3–29.

Crothers, D. M., T. E. Haran, and J. G. Nadeau. 1990. Intrinsically bent DNA. *J. Biol. Chem.* 265:7093–7096.

Dlakic, M., and R. Harrington. 1995. Bending and torsional flexibility of G/C-rich sequences as determined by cyclization assays. *J. Biol. Chem.* 270:29945–29952.

Dlakic, M., and R. E. Harrington. 1996. The effects of sequence context on DNA curvature (published erratum appears in *Proc. Natl. Acad. Sci. USA.* 1996 Aug 6;93(16):8796). *Proc. Natl. Acad. Sci. USA.* 93: 3847–3852.

Godunov, S. K., and V. S. Riabenkii. 1987. Difference schemes: an introduction to the underlying theory. Elsevier Science Pub. Co., New York.

Hagerman, P. J., and V. A. Ramadevi. 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.

Jacobson, H., and W. H. Stockmayer. 1950. Intramolecular reaction in polycondensation. I. Theory of linear systems. *J. Chem. Phys.* 18: 1600–1606.

Kahn, J. D., and D. M. Crothers. 1992. Protein-induced bending and DNA cyclization. *Proc. Natl. Acad. Sci. USA.* 89:6343–6347.

Kahn, J. D., and D. M. Crothers. 1998. Measurement of the DNA bend angle induced by the catabolite activator protein using Monte Carlo simulation of cyclization kinetics. *J. Mol. Biol.* 276:287–309.

Kerppola, T. K. 1996. Fos and Jun bend the AP-1 site: effects of probe geometry on the detection of protein-induced DNA bending. *Proc. Natl. Acad. Sci. USA.* 93:10117–10122.

Klenin, K. V., A. V. Vologodskii, V. V. Anshelevich, A. M. Dykhne, and M. D. Frank-Kamenetskii. 1991. Computer simulation of DNA supercoiling. *J. Mol. Biol.* 217:413–419.

Koo, H. S., J. Drak, J. A. Rice, and D. M. Crothers. 1990. Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry.* 29:4227–4234.

Korn, G. A., and T. M. Korn. 1961. Mathematical handbook for scientists and engineers. McGraw-Hill, Inc., New York, Toronto, London.

Levene, S. D., and D. M. Crothers. 1986. Ring closure probabilities for DNA fragments by Monte Carlo simulation. *J. Mol. Biol.* 189:61–72.

Livshits, M. A. 1996. Calculating the probability for cyclization of a short intrinsically curved and twisted DNA fragment. *Mol. Biol.* 30:85–90.

Livshits, M. A., and Y. L. Lyubchenko. 1994. Kinetics of oligonucleotide cyclization and ligation. *Mol. Biol.* 28:687–690.

Lyubchenko, Y. L., L. S. Shlyakhtenko, E. Appella, and R. E. Harrington. 1993. CA runs increase DNA flexibility in the complex of lambda Cro protein with the OR3 site. *Biochemistry.* 32:4121–4127.

Lyubchenko, Y. L., L. S. Shlyakhtenko, B. Chernov, and R. E. Harrington. 1991. DNA bending induced by Cro protein binding as demonstrated by gel electrophoresis. *Proc. Natl. Acad. Sci. USA.* 88:5331–5334.

McCammon, J. A., and S. C. Harvey. 1987. Dynamics of proteins and nucleic acids. Cambridge University Press, Cambridge, UK.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.

Nagaich, A. K., V. B. Zhurkin, H. Sakamoto, A. A. Gorin, G. M. Clore, A. M. Gronenborn, E. Appella, and R. E. Harrington. 1997. Architectural accommodation in the complex of four p53 DNA binding domain peptides with the p21/waf1/cip1 DNA response element. *J. Biol. Chem.* 272:14830–14841.

Rivetti, C., C. Walker, and C. Bustamante. 1998. Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.* 280:41–59.

Shimada, J., and H. Yamakawa. 1984. Ring-closure probabilities for twisted wormlike chains: application to DNA. *Macromolecules.* 17: 689–698.

Shlyakhtenko, L. S., E. Appella, R. E. Harrington, I. Kutyavin, and Y. L. Lyubchenko. 1994a. Structure of three-way DNA junctions. 2. Effects of extra bases and mismatches. *J. Biomol. Struct. Dyn.* 12:131–143.

Shlyakhtenko, L. S., D. Rekesh, S. M. Lindsay, I. Kutyavin, E. Appella, R. E. Harrington, and Y. L. Lyubchenko. 1994b. Structure of three-way

DNA junctions. 1. Non-planar DNA geometry. *J. Biomol. Struct. Dynam.* 11:1175–1189.

Shore, D., and R. L. Baldwin. 1983. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* 170: 957–981.

Shore, D., J. Langowski, and R. L. Baldwin. 1981. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. USA.* 78:4833–4837.

Taylor, W. H., and P. J. Hagerman. 1990. Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* 212:363–376.

Ulanovsky, L., M. Bodner, E. N. Trifonov, and M. Choder. 1986. Curved DNA: design, synthesis, and circularization. *Proc. Natl. Acad. Sci. USA.* 83:862–866.

Wang, J. C. 1967. Cyclization of coliphage 186 DNA. *J. Mol. Biol.* 28:403–411.

Wang, J. C. 1979. Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci. USA.* 76:200–203.

Wang, J. C., and N. Davidson. 1966a. On the probability of ring closure of lambda DNA. *J. Mol. Biol.* 19:469–482.

Wang, J. C., and N. Davidson. 1966b. Thermodynamic and kinetic studies on the interconversion between the linear and circular forms of phage lambda DNA. *J. Mol. Biol.* 15:111–123.

Wang, J. C., and N. Davidson. 1968. Cyclization of phage DNAs. *Cold Spring Harbor Symp. Quant. Biol.* 33:409–415.

Xu, R., B. Mao, S. Amin, and N. E. Geacintov. 1998. Bending and circularization of site-specific and stereoisomeric carcinogen-DNA adducts. *Biochemistry.* 37:769–778.

Zeman, S. M., K. M. Depew, S. J. Danishefsky, and D. M. Crothers. 1998. Simultaneous determination of helical unwinding angles and intrinsic association constants in ligand-DNA complexes: The interaction between DNA and calichearubicin B. *Proc. Natl. Acad. Sci. USA.* 95: 4327–4332.