

Global structure and mechanical properties of a 10-bp nucleosome positioning motif

M. Roychoudhury, A. Sitlani*, J. Lapham†, and D. M. Crothers‡

Department of Chemistry, P.O. Box 208107, Yale University, New Haven, CT 06520

Contributed by D. M. Crothers, October 6, 2000

The method of DNA cyclization kinetics reveals special properties of the TATAAACGCC sequence motif found in DNA sequences that have high affinity for core histones. Replacement of 30 bp of generic DNA by three 10-bp repeats of the motif in small cyclization constructs increases cyclization rates by two orders of magnitude. We document a 13° bend in the motif and characterize the direction of curvature. The bending force constant is smaller by nearly 2-fold and there is a 35% decrease in the twist modulus, relative to generic DNA. These features are the likely source of the high affinity for bending around core histones to form nucleosomes. Our results establish a protocol for determination of the ensemble-averaged global solution structure and mechanical properties of any ≈10-bp DNA sequence element of interest, providing information complementary to that from NMR and crystallographic structural studies.

DNA in cells is highly compacted from the unfettered size it would have if free in solution. The first stage of packaging is achieved by wrapping DNA around core histones in a well-defined conformation (1), forming a beads-on-a-string structure that is further compacted by additional proteins into folded chromatin and ultimately into metaphase chromosomes. Interphase chromatin is generally less condensed, but it contains regulatory sites such as promoters at which DNA must be distorted to enable interactions between proteins that bind distant sites on linear DNA. As a relatively stiff polymer, DNA resists the bending required for packaging, which requires free energy assistance from strong interactions with core histones and other proteins.

Variations in DNA sequence can, in principle, modulate the energy cost of bending and packaging. For example, intrinsic curvature of the kind provided by A-tracts (2) should, if in the proper direction, lessen the energy required for bending. Increased local bending flexibility could reduce the energetic cost of packaging, as could an increase in twist flexibility (3), assuming that the helical phasing between distal regulatory sequence elements must be altered to form a protein complex.

The stiffness of DNA (the inverse of its flexibility) is traditionally described by the persistence length P , defined as the average projection of the end-to-end vector of a very long chain on its initial direction. Classical techniques, such as light scattering (4) and rotational diffusion (5), supplemented by DNA cyclization kinetics (6), have led to estimates of around 140–180 bp for P . However, these methods are generally unable to deconvolute rigorously the contributions of curvature and flexibility to apparent persistence length (7), unless special constructs are used (8).

Parameter sets exist for predicting DNA curvature based on roll and tilt values for dinucleotide (9–13) and tetranucleotide (14) steps, but the dinucleotide parameters vary considerably from set to set (15). Nor is there sufficient data to define unambiguously the sequence dependence of the experimental persistence length, except for repeating polymers such as alternating poly[d(A-T)] and poly[d(G-C)], which are found to be relatively more and less flexible, respectively, than genomic DNA (3, 16). Potentially predictive parameters for roll, tilt, and twist flexibilities of individual base pair steps are available from

the standard deviation of these parameters in the crystallographic database (11, 12) and accompanying computational approaches (13, 14). De Santis and colleagues (17) have shown that bending flexibility parameters based on thermal stability have predictive power for sequence-dependent nucleosome binding affinity.

The total torsional flexibility of DNA is due to the combined effects of twist and writhe fluctuations, and average values can be determined from study of closed circular molecules (18). Fluorescence polarization anisotropy studies of linear DNA molecules show little dependence of torsional flexibility on average base composition (19). However, local variations of the twist flexibility depending on sequence have not been determined in solution. In summary, there is a clear need for experimentally based parameters that describe the ensemble-averaged solution curvature and local bending and twist flexibility of specific DNA sequences. In addition to providing a test for parameter sets based on crystallography and theory, such a database would allow more stringent test of the intriguing indications that curved and flexible sequences are organized coherently in promoters and other DNA sequences (20–22) and would enable analysis of the properties of the noncoding DNA sequences that constitute ≈95% of the human genome.

We have developed a strategy, based on specialized DNA molecules for measurement of cyclization kinetics, that offers a general solution to this problem (23–28). A key feature is inclusion of phased A-tracts, whose intrinsic curvature fixes the mean rotational setting of the circle. Altering the phasing between a test sequence element and the A-tracts has a large effect on cyclization if the test sequence is bent, but not if it is simply altered in isotropic flexibility. The cyclization constructs also provide for variation of the overall DNA length, allowing determination of the bending and twist flexibility and the DNA helical repeat. The constructs are kept small (≈160 bp) as a design principle so that small changes in curvature or flexibility have a large effect on cyclization rate. The method has been used independently by Davis *et al.* (29) to show that the sequence TATAAAAG is bent (or anisotropically flexible) in a direction opposite to that induced by the TATA box binding protein.

In the work reported here we show that the distinctive curvature and mechanical properties of a small (10 bp) sequence element can be quantitatively determined by this approach. The sequence motif chosen, the TATA tetrad (5'-TATAAACGCC-3'), is found in multiple phased repeats in genomic sequences

Abbreviation: NPS, nucleosome positioning sequence.

*Present address: Department of Endocrinology and Chemical Biology, Merck & Co., Mail Code R80Y-140, P.O. Box 2000, Rahway, NJ 07065.

†Present address: Centro Nacional de Ressonancia Magnetica Nuclear de Macromoleculas, Universidade Federal do Rio de Janeiro, Brazil.

‡To whom reprint requests should be addressed at: Department of Chemistry, Yale University, P.O. Box 208104, New Haven, CT 06520-8104. E-mail: donald.crothers@yale.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.250476297. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.250476297

that are among the strongest known binders of core histones (30). We have repeated the 10-bp sequence element three times to amplify both its curvature and the influence of its flexibility, and inserted it into the cyclization constructs. Molecules in which the 30-bp test nucleosome positioning sequence (NPS) replaces 30 bp of “generic” DNA cyclize approximately 100-fold faster than the parent constructs. Recent improvements in the speed and reliability of the computer simulation of the cyclization process (J.L. and D.M.C., unpublished work) enable us to provide better quantitative interpretation of bend angle and direction, along with the flexibility parameters. The simulations reveal that the accelerated cyclization can be resolved into a combination of intrinsic (or static) curvature and increased (isotropic) bending flexibility in the NPS, relative to generic DNA.

Materials and Methods

Theoretical Background. The cyclization kinetic method requires a comparison of the rate constants for unimolecular cyclization and bimolecular ligation (6, 23). The basic assumption is that the rate of each of these processes is limited by the concentration of the substrate for DNA ligase, which consists of molecules in which the complementary DNA ends are transiently hydrogen bonded. It is further required that the ends dissociate rapidly compared with the rate of the ligation step. Under these conditions, the rate constants k_1 and k_2 for cyclization and bimolecular ligation, respectively, are proportional to the corresponding equilibrium constants K_1 and K_2 for transient pairing of the substrate ends. The ratio J is given by

$$J = \frac{K_1}{K_2} = \frac{k_1}{k_2}.$$

The experiments consist of determination of J values for the different DNA constructs by measurement of the reaction kinetics, using gel electrophoresis of ^{32}P -labeled DNA to quantify the extent of reaction.

Design of the Cyclization Constructs. The design of our constructs, shown schematically in Fig. 1, is based on earlier work reported from our laboratory (23–28). The molecules contain a set of six phased A-tracts separated from the 30-bp NPS by linkers with variable lengths of 12, 16, and 20 bp. The adaptor region enables variation of the overall length of the constructs.

DNA Sequences. These are derived from the 17A9 molecule described by Sitlani and Crothers (26), with the 30-bp NPS site replacing 30 bp of sequence around the AP-1 site. All molecules have 2-bp 5′ overhanging CG ends, and share the A-tract region of the sequence, which is: (1)CGATGAATTCCTGTACGGA-TCCGGAAAAACGGGCAAAAAACGGCAAAAAACGGGCAAAAAACGG-ATC(90). This is followed in the 156-bp B-DNA molecule by (91)CGAATTCTAGACCTAGGTGGATGACTCATTTTTTTT-TTGCTCGAGCTACAACGTGCTGCCATGGAAT(156) and, in the variable phasing linker sequences, by

12_156

CGAAT**TATAACGCCTATAAACGCCTATAAACGCC-**
TTGCTCGAGCTACAACGTGCTGCCATGGAAT

16_156

CGAAT**TTTGCTATAAACGCCTATAAACGCCTATAAA-**
CGCCTCGAGCTACAACGTGCTGCCATGGAAT

20_156

CGAAT**TTTGCTCGATATAAACGCCTATAAACGCCTA-**
TAAACGCCGCTACAACGTGCTGCCATGGAAT

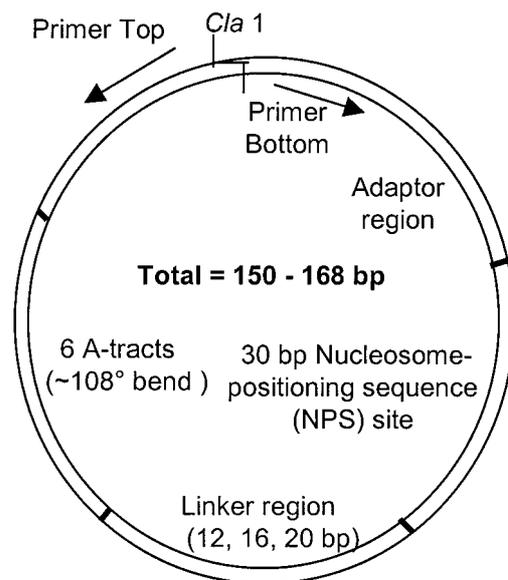


Fig. 1. A-tract–NPS minicircles. The six phased A-tracts and the 30-bp NPS region are separated by linker regions with variable lengths of 12, 16, of 20 bp. The adaptor region enables variation of the overall length of the constructs. The first number in the nomenclature of our constructs, e.g., 12.154, indicates the length of the phasing linker, and the second number indicates the overall length of the construct.

in which the 30-bp segment is indicated in boldface, and the sequence elements that differ in the phasing linkers are underlined. The molecules with different overall lengths were constructed by using variable bottom-strand PCR primers taken from the set

```

150 5'-GCAGATATCGATGCAGCACGTTGTAGC
151   GCAGATATCGATCGCAGCACGTTGTAGC
152   GCAGATATCGATTCCAGCACGTTGTAGC
154   GCAGATATCGATTCCAGCACGTTGTAGC
155   GCAGATATCGATTCCATGCAGCACGTTGTAGC
156   GCAGATATCGATTCCATGGCAGCACGTTGTAGC
157   GCAGATATCGATTCCATGAGCACGTTGTAGC
158   GCAGATATCGATTCCATGGCAAAGCACGTTGTAGC
160   GCAGATATCGATTCTTGGACATGGCAGCACGTTGTAGC
161   GCAGATATCGATTCTTGGACATGGCAGCACGTTGTAGC
162   GCAGATATCGATTCTTGGACATGGCAGCACGTTGTAGC
164   GCAGATATCGATTCTTCTGGACATGGCAGCACGTT-
      GTAGC
166   GCAGATATCGATGGTTCTCTGGACATGGCAGCACG-
      TTGTAGC
168   GCAGATATCGATCTGGTTCTCTGGACATGGCAGCA-
      CGTTGTAGC

```

where the variable sequences are underlined. The *ClaI* site, indicated in boldface, is used to create the lower strand 2-bp 5′ CG overhang in all molecules. The upper strand PCR primer, 5′-CGCCATGGAATCGATGAATTC, also contained a *ClaI* site. The molecules were assembled by standard DNA synthesis, ligation, and PCR methods (31). Cyclization substrates were prepared by PCR amplification in the presence of [α - ^{32}P]dATP, purified by using the Qiagen (Chatsworth, CA) PCR purification kit, and digested at 37°C overnight with 2.5 units of *ClaI* (New England Biolabs) per 100 μl . After purification on nondenaturing 10% polyacrylamide gels and electroelution, the radioactivity was monitored by scintillation counting and the concentration was determined by UV absorbance.

Cyclization Kinetics. Rate constants were measured as described previously (23, 24). Briefly, T4 DNA ligase (New England Biolabs) at 500 units·ml⁻¹ was added to a preincubated solution of 2–10 nM DNA in a reaction buffer of 10 mM Tris·HCl (pH 7.5), 5 mM MgCl₂, 50 mM NaCl, 1 mM EDTA, 50 μg·ml⁻¹ BSA, 5% sucrose, 5% (vol/vol) glycerol, 5 mM DTT, and 0.05–0.1% Nonidet P-40 at 21°C. Samples (8 μl) of the reaction mixture were taken at various time points from 0 to 8 min and added to a quenching mixture of 4 μl of 45 mM EDTA, 3 mg·ml⁻¹ proteinase K, 15% glycerol, and electrophoresis dyes. Quenched reactions were incubated at 55°C for 10 min and loaded on a native 6% polyacrylamide gel that was prerun for 1 h. The gels were analyzed with a FUJIX Bio-imaging Analyzer System using MACBAS software. In most cases the cyclization rate constant k_1 and bimolecular rate constant k_2 could be obtained directly from the time dependence of the concentration of circular products and linear and circular dimers, respectively (24), related to the approach of Taylor and Hagerman (32). In the case of rapid cyclizers we accessed the bimolecular rate constant either by increasing the DNA concentration to 25 nM or by determining k_2 for a standard molecule (17A9) in a parallel reaction.

Computer Simulations. Monte Carlo simulation of chain configurations was carried out as described (23, 33). Bends at A-tracts were introduced by specific roll and tilt angles at the junctions of the A-tract with adjacent DNA (34). The chain is assumed to be inextensible, and it is assumed that roll, tilt, and twist fluctuations are uncorrelated. Any curvature that results from anisotropic flexibility (29, 35) is represented in our model by a static bend and isotropic flexibility. The program generates libraries of first and second half chains, and screens whole-chain combinations of subsets of these for end-to-end distances below a cut-off value. Chains passing this criterion are further examined for colinear alignment of the helix ends and the twist angle between base pairs at the two ends to calculate J . The original computer algorithm for simulation of DNA cyclization (33) was further developed by Rice (34) and Kahn (36). Most recently, we (J.L. and D.M.C., unpublished work) found that the main source of statistical fluctuations in the results was the excessive importance of a small number of highly bent half chains, particularly in the non-A-tract portion of the molecule. These extreme fluctuations arise from in-phase bend fluctuations, and are less important in the A-tract half, for which the static curvature makes highly bent states more accessible. The new version of the program, available from the authors on request, generates a larger and more diverse population of chains. Rather than combining all first half chains with all second half chains, the program generates a larger (adjustable) number of half chains and combines these in subsets. This approach enables statistical analysis of the J values that result from the individual subsets. The typical standard deviation of simulated values of J ranges from less than 5% for J values greater than 100 nM to about 8% for J values of ≈ 1 nM. These numbers are significantly smaller than the experimental standard deviation, which is typically 10–25%. The difference between simulated and experimental data sets was characterized by the “absolute error” (36), which is a sum of squares of the difference between theoretical and experimental values of $\log_{10} J$ for all molecules in the set under consideration.

Results and Discussion

Variation of J Values with Phasing Linker Length. We determined J values for the three different phasing linker lengths of 12, 16, and 20 bp for three different overall lengths, 154, 156, and 158 bp. Fig. 2 shows typical data for the kinetic measurements. The rate of the bimolecular ligation step was determined from the time dependence of the sum of the concentrations of linear dimers and circular multimers (24). As is evident from Fig. 2, the 16-bp linker at 156-bp overall length cyclizes more rapidly than the

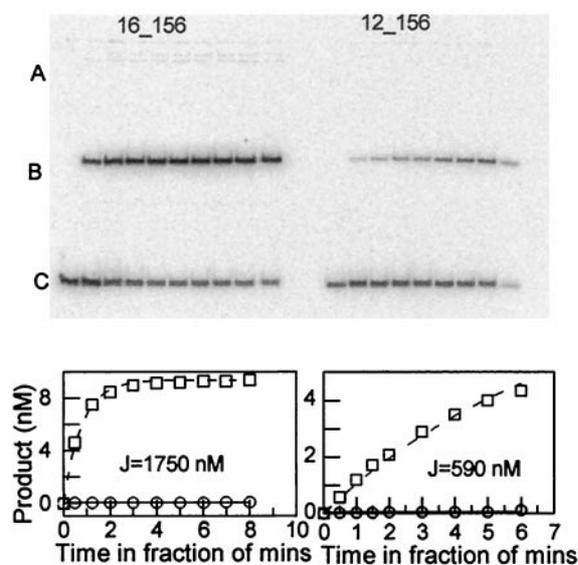


Fig. 2. Cyclization kinetic measurements on 16_156 and 12_156 constructs. (Upper) Cyclization reactions of 16_156 and 12_156 constructs. A, products of bimolecular ligation; B, cyclization product; and C, linear starting material. (Lower) The corresponding determination of cyclization and bimolecular association rate constants and J factors. □, Circular monomer; ○, products of bimolecular ligation.

12-bp linker. The results are summarized in Table 1, which shows that the highest measured J value is found for the 16-bp linker. Note that whereas the 16-bp linker has an optimum cyclization rate at 156 bp, the rate for the 12-bp linker is greater at 154- than at 156-bp length. Two noncoplanar bends contribute writhe to the molecules, depending on the angle between the bend planes. As a consequence, the optimum number of base pairs needed to yield an integral number of helical turns in the circles depends on the phasing linker length. This rather complicated problem is dealt with automatically by the computer simulations described below, from which we deduce that the ideal linker length required for optimal phasing of the A-tract and test bends would be about 14.5 bp (31). The closest experimental length is 16, which is therefore the most nearly planar molecule among our set.

Variation of J Values with Overall Length. The 16-bp linker was chosen for study of the dependence on overall length. These results are required for a full determination of the bending and twist flexibility and the helical repeat. The constructs differed by 1 or 2 bp in length, for both the NPS-containing molecules and the parent set containing “generic” DNA in the 30-bp test site.

Table 1. Variation of J factors with linker length for the NPS-containing molecules

DNA length, bp	Linker length, bp	J , nM
154	12	960 (± 120)
154	16	650 (± 75)
154	20	45 (± 5)
156	12	650 (± 80)
156	16	1,650 (± 150)
156	20	30 (± 5)
158	12	70 (± 10)
158	16	500 (± 120)
158	20	25 (± 5)

Numbers in parentheses are errors.

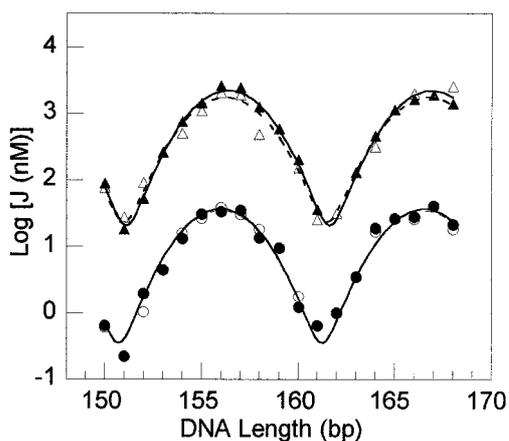


Fig. 3. Comparisons of experimental and theoretical $\log J$ versus DNA length for constructs with and without the NPS site. The upper sets of curves correspond to constructs having the NPS site with a fixed linker length of 16 bp. The experimental $\log_{10} J$ values are shown as \triangle and the corresponding curve fit is the dashed line. The simulation values are indicated by \blacktriangle and their curve fit is the solid line. The lower sets of curves correspond to the B-DNA molecules containing a generic 30-bp sequence in place of the NPS site. The experimental $\log_{10} J$ values are the \circ with the dashed line indicating the curve fit; the simulation values are the \bullet with solid line indicating the curve fit. The data are fitted to double Gaussian equations (25).

[Earlier work showed that there is no appreciable bend in this sequence (26, 27).] The results are summarized in Fig. 3. The expected (18) modulation of the cyclization rate with the ≈ 10 -bp DNA helical periodicity is evident, and there is a striking increase of about two orders of magnitude in the cyclization rate for the NPS-containing molecules compared with the parent set. This dramatic effect must be due at least in part to the curvature of the NPS segment detected in the variation of rate with phasing linker length. However, flexibility changes may also contribute, and it requires a simulation of the process to deconvolute these two variables, as described below.

Verification of the Properties of DNA Lacking the 30-bp Sequence. In the simulations to determine the properties of the NPS, it is necessary to assign flexibility and helical repeat values to the remainder of the molecule. The values for the A-tract segment come from earlier experiments (34). The experiments for the parent set of molecules reported in Fig. 3 allow determination of average parameters for the rest of the molecule, including the 30-bp segment replacing the NPS site, which we took to be generic B-DNA (27). However, we were concerned that this “generic” 30-bp segment might have properties different from the average for the molecule, so we carried out cyclization kinetic experiments for test molecules in which the 30-bp segment was

Table 2. Comparison of experimental and simulated J values for molecules lacking the 30-bp insert

DNA length, bp	J factor, nM	
	Experiment	Simulation*
124	19.8	20.8
126	16.0	14.8
128	1.2	1.2
134	21.5	21.0
136	22.5	23.4
138	4.0	5.4

*The parameters assumed for the simulation are those used for the non-NPS sections of the molecules shown in Fig. 3 and summarized in Table 3.

deleted entirely. Table 2 shows the J values measured, compared with simulation, based on the flexibility and helical repeat parameters determined for the longer B-DNA molecules in Fig. 3. The good correspondence between measured and simulated values for the shortened molecules confirms that the DNA outside the NPS sequence is accurately described by the parameter set. We also infer that the deleted 30-bp segment has properties indistinguishable from those of the rest of the B-DNA portion of the molecule.

Global Bend Direction Determined by Computer Simulation. Determination of the properties of DNA molecules by matching simulated and experimental J values for a set of 20 cyclization constructs is a multivariable minimization problem, which we have solved by an iterative series of refinements of the parameters. We took advantage of the dependence, to first order, of some of the molecular properties on single parameters. For example, the variation of J with phasing linker length is primarily a function of the direction of curvature in the NPS sequence, assuming that at least three different phasing linkers have been measured. The distance between successive maxima in the periodic variation of J with overall length in Fig. 3 is determined by the helical repeat, and the peak-to-trough amplitude of the curves is primarily a function of the twist flexibility. However, the overall 100-fold upward shift of the J values for the NPS sequence relative to B-DNA depends on a combination of increased bending flexibility and the magnitude of the curvature, and it must be treated to first order as a two-dimensional minimization problem in which the entire data set, including all phasing linkers, is included.

The cyclization kinetic method can be used to determine the global direction of curvature, but not its structural origin or actual locus. For simplicity we have modeled the bend by roll between the base pairs at either a single base pair step in each NPS 10-bp element, or between adjacent base pair steps (to improve the resolution of global bend direction). The simulations were used to find the position at which a positive roll angle minimizes the disagreement between experiment and calculated $\log_{10} J$ values for the set of phasing linkers. The actual curvature can be due to any combination of roll and tilt angles in the NPS that has the same resultant as our arbitrary roll-bend.

Fig. 4a shows the “absolute error” (the sum of squares of the difference between simulated and experimental values of $\log_{10} J$) for three different linker lengths at a total molecular length of 154 bp. All three curves share a common minimum for a total roll angle, taken here to be 13° for reasons documented below, distributed between nucleotides 8, 9, and 10, in each of the three GCC trinucleotides in the 30-bp sequence. (The kinks at the GC and CC steps were taken to be 3.25° and 9.75° , respectively, so that the scalar sum is 13° .) The same bend position was found from fitting the data for the set of 156-bp and 158-bp molecules. If a negative roll angle, indicating a compression of the minor groove, is assumed, the optimum position is shifted by half a helical turn and partitioned over the steps between nucleotides 3, 4, and 5, corresponding to the TAA trinucleotide in the NPS. Taken together, the results are consistent with some combination of a bend toward the minor groove in the A-T-rich section

Table 3. Best-fit parameters derived from the simulations

DNA	Roll and tilt fluctuations, $^\circ$	Torsional modulus, 10^{-19} erg-cm	Kink angle, $^\circ$	Helical repeat, bp
B-DNA	4.68	2.4	0	10.45
A-tract	4.68	2.4	*	10.33
NPS site	6.5	1.5	13	10.30

*The A-tracts are bent by roll and tilt kinks at the junctions (34).

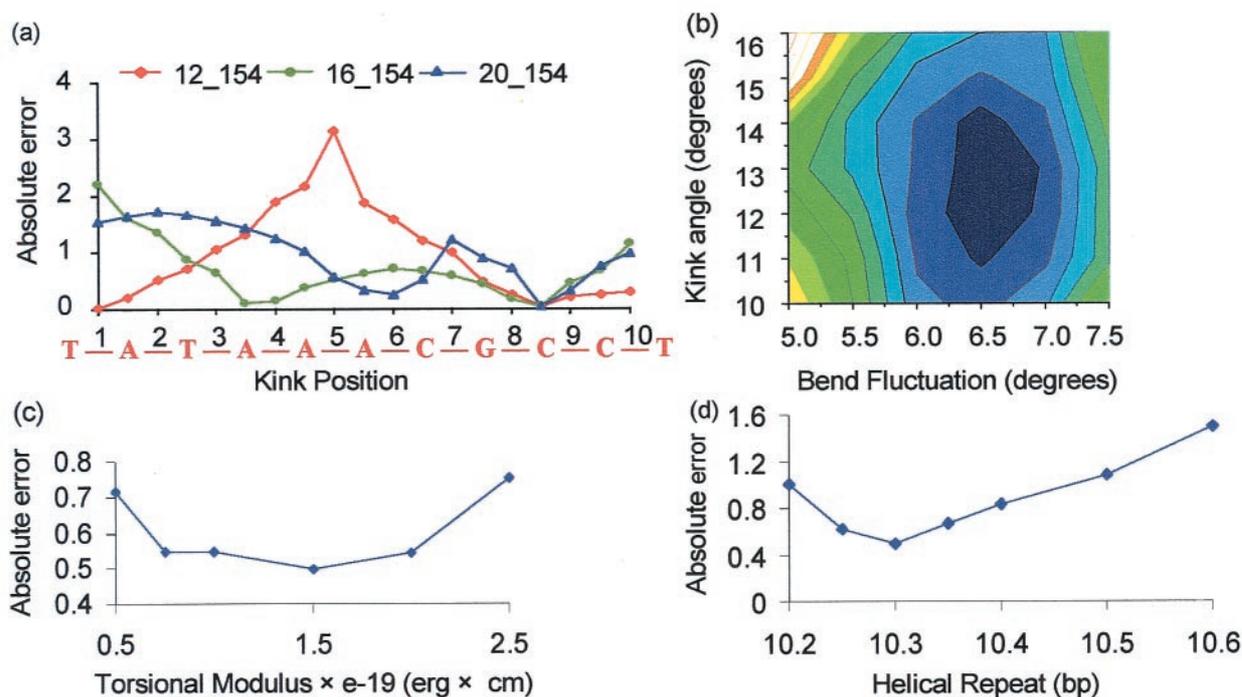


Fig. 4. Refinement of geometric and flexibility parameters against experiment. A lower value for the absolute error corresponds to a better agreement between experimental and theoretical $\log_{10} J$ factors. (a) Determination of the kink position for molecules with a fixed overall length of 154 bp and varied linker lengths of 12, 16, and 20 bp. Kink angles are centered on a base pair—for example, the second C in CCG has one-fourth the total (scalar) kink angle at the CC step, and three-fourths at CG. (b) Projection of a three-dimensional plot on a two-dimensional xy plane, showing absolute error as a function of kink angle and rms angle fluctuations. The black center corresponds to minimum error (0.4–0.64), with subsequent contours increasing by 0.26. The figure indicates that a kink angle between 11.5° and 14° and roll and tilt fluctuations ranging from 6.25° to 6.75° at each base pair step in the 30-bp NPS site provide the best fit between experimental and simulation values of the $\log_{10} J$ factors. (c) Absolute errors for varying the twist modulus of the 30-bp NPS site. (d) Absolute errors for varying the helical repeat of the 30-bp NPS site.

and toward the major groove in the G-C-rich segment of the NPS. This conclusion is consistent with the observed direction of curvature of the NPS when incorporated into nucleosomes, as shown by Widlund *et al.* (37). Those authors also reported an increase in cyclization rate for molecules containing the NPS, but they did not deconvolute the effects of curvature and flexibility nor establish the direction of curvature of the DNA in the absence of histones.

Determination of Flexibility and Helical Repeat. The dependence of the absolute error on the twisting modulus and the helical repeat is shown in Fig. 4 *c* and *d*. In these simulations, the kink angle and bending flexibility for the NPS segment were set at the optimum values of 13° and rms roll and tilt fluctuations of 6.5° (equivalent to a persistence length of 78 bp), as documented in Fig. 4*b*. The B-DNA segment has roll and tilt fluctuations of 4.68°, equivalent to a persistence length of 150 bp. The process of iterative refinement of the NPS properties by minimization of the absolute error leads to the parameter set summarized in Table 3. The overall curvature of the 30-bp sequence is about 40° and the average bending flexibility (proportional to the square of the rms angle fluctuations, or to the inverse of the persistence length or bending force constant) is increased by nearly 2-fold from the value for the remainder of the molecule. In independent simulations, we showed that either increased flexibility with zero curvature, or 40° curvature with normal B-DNA flexibility, accounts for about a factor of 10 in the 100-fold increase of J

values for the NPS molecules relative to the parent set. There is also a significant increase in twist flexibility, corresponding to the reduction of the twisting modulus from 2.4 to about 1.5×10^{-19} erg \cdot cm.

Significance. Our results establish a protocol for determination of the global structure and flexibility parameters of defined DNA sequence elements containing an approximately integral number of helical turns. The technique provides a view of overall DNA structure for an ensemble of molecules which is complementary to that available from crystallographic and NMR structural studies. The high sensitivity of J values to DNA properties should make it possible to characterize curvature of a few degrees and flexibility changes of $\approx 10\%$; examination of a large number of sequences can enable construction of a database that is predictive for the curvature and flexibility of any arbitrary sequence.

At the present time we can specify only the average flexibility and curvature parameters for the NPS, but future mutagenic studies on the sequence should identify specific sequence determinants of the curvature and enhanced flexibility. It seems likely that the bend, enhanced twist and bending flexibility, and reduced helical repeat all contribute to the strong binding affinity of the NPS for core histones.

We thank Zippora Shakked, Wilma Olson, and Jason Kahn for their comments on the manuscript. This work was supported by Grant GM 21966 from the National Institutes of Health.

- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997) *Nature (London)* **389**, 251–260.
- Wu, H.-M. & Crothers, D. M. (1984) *Nature (London)* **308**, 509–513.

- Hogan, M. E. & Austin, R. H. (1987) *Nature (London)* **329**, 263–266.
- Harpst, J. A. (1980) *Biophys. Chem.* **11**, 295–302.
- Hagerman, P. J. (1981) *Biopolymers* **20**, 1503–1535.

6. Shore, D., Langowski, J. & Baldwin, R. L. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4833–4837.
7. Trifonov, E. N. (1985) *CRC Crit. Rev. Biochem.* **19**, 89–106.
8. Levene, S. D., Wu, H.-M. & Crothers, D. M. (1986) *Biochemistry* **25**, 3988–3995.
9. Bolshoy, A., McNamara, P., Harrington, R. E. & Trifonov, E. N. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2312–2316.
10. De Santis, P., Palleschi, A., Savino, M. & Scipioni, A. (1990) *Biochemistry* **29**, 9269–9273.
11. Gorin, A. A., Zhurkin, V. B. & Olson, W. K. (1995) *J. Mol. Biol.* **247**, 34–48.
12. Olson, W. K., Babcock, M. S., Gorin, A., Liu, G., Markey, N. L., Martino, J. A., Pederson, S. C., Srinivasan, A. R., Tobias, I. & Westcott, T. P. (1995) *Biophys. Chem.* **55**, 7–29.
13. Packer, M. J., Dauncey, M. P. & Hunter, C. A. (2000) *J. Mol. Biol.* **295**, 71–83.
14. Packer, M. J., Dauncey, M. P. & Hunter, C. A. (2000) *J. Mol. Biol.* **295**, 85–103.
15. Crothers, D. M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 15163–15165.
16. Thomas, T. J. & Bloomfield, V. A. (1983) *Nucleic Acids Res.* **11**, 1919–1930.
17. Anselmi, C., Bocchinfuso, G., De Santis, P., Savino, M. & Scipioni, A. (1999) *J. Mol. Biol.* **286**, 1293–1301.
18. Shore, D. L. & Baldwin, R. L. (1983) *J. Mol. Biol.* **170**, 957–981.
19. Fujimoto, B. S. & Schurr, J. M. (1990) *Nature (London)* **344**, 175–177.
20. Wada-Kiyama, Y., Kuwabara, K., Sakuma, Y., Onishi, Y., Trifonov, E. N. & Kiyama, R. (1999) *FEBS Lett.* **444**, 117–124.
21. Ozoline, O. N., Deev, A. A. & Trifonov, E. N. (1999) *J. Biomol. Struct. Dyn.* **16**, 825–831.
22. Pederson, A. G., Jensen, L. J., Brunak, S., Stærfeldt, H.-H. & Ussery, D. W. (2000) *J. Mol. Biol.* **299**, 907–930.
23. Crothers, D. M., Drak, J., Kahn, J. D. & Levene, S. D. (1992) *Methods Enzymol.* **212B**, 3–29.
24. Kahn, J. D. & Crothers, D. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6343–6347.
25. Kahn, J. D., Yun, E. & Crothers, D. M. (1994) (1994) *Nature (London)* **368**, 163–166.
26. Sitlani, A. & Crothers, D. M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 3248–3252.
27. Hockings, S. C., Kahn, J. D. & Crothers, D. M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 1410–1415.
28. Sitlani, A. & Crothers, D. M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 1404–1409.
29. Davis, N. A., Majee, S. S. & Kahn, J. D. (1999) *J. Mol. Biol.* **291**, 249–265.
30. Widlund, H. R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P. E., Kahn, J. D., Crothers, D. M. & Kubista, M. (1997) *J. Mol. Biol.* **267**, 807–817.
31. Roychoudhury, M. (2000) Ph.D. thesis (Yale Univ., New Haven, CT).
32. Taylor, W. H. & Hagerman, P. J. (1990) *J. Mol. Biol.* **212**, 363–376.
33. Levene, S. D. & Crothers, D. M. (1983) *J. Biomol. Struct. Dyn.* **1**, 429–435.
34. Koo, H.-S., Drak, J., Rice, J. A. & Crothers, D. M. (1990) *Biochemistry* **29**, 4227–4234.
35. Olson, W. K. & Zhurkin V. B. (2000) *Curr. Opin. Struct. Biol.* **10**, 286–297.
36. Kahn, J. D. & Crothers, D. M. (1998) *J. Mol. Biol.* **276**, 287–309.
37. Widlund, H. R., Kuduvalli, P. N., Bengtsson, M., Cao, H., Tullius, T. D. & Kubista, M. (1999) *J. Biol. Chem.* **274**, 31847–31852.