Nucleic Acids Research, 1999, Vol. 27, No. 21 **4135–4142**

Improved quantitation of DNA curvature using ligation ladders

Eric D. Ross, Robert B. Den, Philip R. Hardwidge and L. James Maher III*

Department of Biochemistry and Molecular Biology, Mayo Foundation, Rochester, MN 55905, USA

Received August 6, 1999; Revised and Accepted September 16, 1999

ABSTRACT

It is often desirable to estimate accurately the local shape of DNA molecules. Such measurements are useful in understanding the intrinsic contribution of DNA sequence to curvature, as well as in assessing the effects of chemical modifications. We have been investigating the effects of asymmetric phosphate neutralization on DNA shape using the well-characterized ligation ladder approach developed by Crothers and co-workers [D.M.Crothers and J.Drak (1992) *Meth. Enzymol.*, 212, 46–71]. This technique is remarkably sensitive to differences in DNA shape. We now report a general quantitative assay of DNA curvature that we have validated using a set of phased A_5 tract standards. This approach allows simultaneous estimation of helix axis deflection magnitude and direction when a test sequence is monitored in at least three phasings relative to a reference $A_{5–6}$ tract in short DNA duplexes. Analysis using this improved approach confirms our published data on DNA curvature due to electrostatic effects.

INTRODUCTION

DNA curvature and bending play critical roles in biology. DNA curvature refers to changes in the helix axis that are intrinsic to the DNA sequence in a given environment of solvent and counterions. DNA bending involves changes in the helix axis due to the binding of ligands such as drugs and proteins. The DNA double helix is rather rigid locally as measured by persistence length, a descriptor of the tendency of segments within a polymer to remain aligned. The persistence length of DNA, *P*, is ~150 bp under physiological conditions (1). Over DNA lengths on the order of *P*/2 (~75 bp), DNA can be modeled as a rigid rod with elastic resilience (2). Because many compact nucleoprotein complexes important for DNA function involve interactions over this length range, factors that change DNA shape and/or flexibility have important biological consequences.

Perhaps the most obvious contexts in which DNA curvature and bending are significant are packaging and transcriptional regulation. The immense length of DNA in mammalian genomes requires extreme volume compaction in the cell nucleus, a challenge met by packaging DNA in a highly spooled complex with histone proteins. In this complex, a persistence length of DNA is bent through almost two full circles. In transcriptional control, information specified by the patterns of protein occupancy of multiple DNA sequences is sensed through protein–protein interactions among these tethered factors. The properties of the intervening DNA are expected to influence the mechanism by which this signal integration results in recruitment of RNA polymerase (3). DNA shape has also been implicated as an important factor in genetic recombination (4).

A diverse array of factors can induce DNA bending. Various classes of proteins induce DNA bending (5), including molecules such as 'architectural proteins' for which DNA bending may be their primary function (6–8). In addition, certain chemical cross-links (9,10) and artificial ligands (11,12) have been shown to distort DNA.

Of particular interest in the present work are DNA modifications (sequence changes, chemical adducts, etc.) that induce DNA curvature, perhaps by mimicking the consequences of protein binding. For example, we have studied DNA curvature induced by alterations in the charge balance along the double helix (13–17). There is growing interest in the physical basis for both intrinsic DNA curvature, and for curvature induced by chemical modifications. Recent proposals include the concept that DNA electrostatics dominate intrinsic curvature, revealing sequence-dependent heterogeneities in the counterion atmosphere (18). Such reasoning has been applied to explain sequences such as $A_{5–6}$ tracts that appear to be strongly curved in solution experiments (19–22).

DNA curvature and bending can be analyzed using a number of methods. High resolution approaches include X-ray diffraction and NMR spectroscopy. Both require relatively large amounts of material. There is concern that intrinsic DNA curvature may be obscured in X-ray crystallography by solvent and ion conditions and crystal packing forces (23). NMR spectroscopy is limited in its ability to define the shape of elongated helices. Other spectroscopic tools such as FRET (24), LRET (25) and TEB (26) offer rich opportunities in analysis of DNA curvature, bending and flexibility. Traditional hydrodynamic approaches such as analytical ultracentrifugation have also been applied. Other clever strategies have been devised to measure the shape of trace quantities of radiolabeled DNA using assays of enzyme-mediated cyclization (27,28), or differential protein binding to preformed minicircles (29–31).

With respect to convenience, economy, simplicity and speed, gel-based electrophoretic assays of DNA shape are extremely attractive. Such methods (reviewed in 32) include

*To whom correspondence should be addressed. Tel: +1 507 284 9041; Fax: +1 507 284 2053; Email: maher@mayo.edu

circular permutation assays, phasing assays and ligation ladder experiments. These approaches suffer from the absence of a rigorous physical theory of native gel electrophoresis through media such as polyacrylamide. However, through the use of insightful calibrations and empirical quantitative relationships, DNA shape can be measured with remarkable sensitivity in such assays.

The present paper involves the analysis of DNA curvature using ligation ladders, a semiquantitative technique pioneered by Crothers and co-workers (reviewed in 32). This group showed that an electrophoretic retardation anomaly typifies the mobility of curved DNA sequences. It was further observed that the anomaly increases as the square of the degree of curvature. Comparative electrophoresis experiments were shown to allow estimation of the magnitude and direction of DNA bends induced by sequence, structure and/or chemical modifications (32). In this approach, DNA duplexes comprising two turns of the double helix are chemically synthesized with cohesive molecular termini for unidirectional ligation in the presence of T4 DNA ligase. Each DNA duplex contains a reference shape [typically one or more A_{5-6} tracts that are intrinsically curved by ~18° in an orientation that has been established (32)]. Upon ligation the A_{5-6} tracts in the resulting ligated multimers are aligned nearly in a plane. Thus, multimers of increasing length are increasingly curved and display corresponding length-dependent mobility retardation anomalies in native polyacrylamide gels.

The ligation ladder technique is limited in certain ways. Ligation ladder experiments require synthetic DNA duplexes susceptible to radiolabeling by T4 polynucleotide kinase, and that retain their activity as substrates for T4 DNA ligase. In addition, DNA shape is measured under the solvent and dilute ionic conditions required for electrophoresis. In spite of these limitations, ligation ladder experiments have been profitably applied to studies of curvature due to DNA sequence, unusual DNA structures such as bulges (33) and chemical cross-links (9,10). DNA curvature induced by metal ions has also been studied using this technique (34–36). The ligation ladder approach offers unprecedented sensitivity and simplicity in its ability to measure reproducible changes of DNA helix axis of a few degrees.

Our own work has utilized the ligation ladder technique to measure changes in DNA shape upon alteration of DNA electrostatics by neutralizing phosphates (13,16,17,37) or appending ionic functions (14,15). After calibrating an empirical curvature equation with A_{5-6} tract standards, we estimated angles of induced curvature in three steps. First, helical repeat parameters were determined. Second, apparent relative curvatures were assigned to each ligation ladder. Third, curvature induced by the charge perturbation was estimated by simultaneously fitting relative curvature data from three phasings to a crude periodic function based on the assumption that curvature was proportional to the vector sum of projections of DNA segments perpendicular to the long axis of the ligated DNA oligomers (13). The physical basis for this assumption is not rigorous and the previous analysis only qualitatively assigned the direction of DNA curvature. Thus, we have sought a rigorous and general approach to deduce both the direction and magnitude of induced DNA curvature from ligation ladder experiments. Based on data from experiments using a calibration standard with two phased A_{5-6} tracts, we now present such an approach and apply it to examples of our previous data. In all cases the results of the refined analysis support our previous conclusions, but in more quantitative detail.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were prepared by standard phosphoramidite chemistry. All oligomers were purified by denaturing 20% polyacrylamide gel electrophoresis, eluted overnight from diced gel slices and desalted using C_{18} reverse-phase cartridges. Oligonucleotide concentrations were determined at 260 nm using appropriate nearest neighbor molar extinction coefficients (M^{-1} cm^{-1}) as described (38). In some cases oligonucleotides were characterized by electrospray ionization mass spectrometry.

Ligation ladders

Purified oligonucleotides (400 pmol) were radiolabeled using polynucleotide kinase and [γ-^{32}P]ATP (10 μCi, 3000 Ci/mmol) in 10 μl reactions containing 8 U T4 polynucleotide kinase in the buffer recommended by the manufacturer (New England Biolabs). Incubation was for 45 min at 37°C, followed by addition of unlabeled ATP to 4 mM and further incubation for 20 min. Equal amounts (169 pmol each) of labeled complementary oligomers were then mixed and adjusted to 10 μl with H_2O, heated to 80°C and cooled gradually to 4°C. The annealed oligonucleotides (169 pmol) were then assembled into ligation reactions (20 μl) containing 800 U T4 DNA ligase in the buffer recommended by the manufacturer (New England Biolabs). A small fraction of unligated duplex was retained as a marker for gel studies. Ligation reactions were incubated at 16°C for 30 min, and then terminated by addition of EDTA to a final concentration of 30 mM.

Exonuclease treatment

In cases where it was necessary to distinguish circular versus linear ligation products, *Bal*31 exonuclease treatment was used to eliminate linear products while sparing DNA circles. *Bal*31 treatment was performed by adding 9 μl 2× *Bal*31 exonuclease reaction buffer and 1 U *Bal*31 exonuclease (New England Biolabs) to 8 μl ligation reaction, and incubating at 30°C for 30 min prior to electrophoresis.

Electrophoretic analysis

Ligation ladders were analyzed on 5% native polyacrylamide (1:29 bisacrylamide:acrylamide) prepared in 1× TBE buffer (39). Electrophoresis was at 10 V/cm at 22°C. Molecular weight markers were created by labeling a 100 bp ladder (Gibco BRL) using T4 DNA polymerase (New England Biolabs) and [α-^{32}P]dATP as recommended by the ladder manufacturer. Gels were dried and exposed to storage phosphor screens for analysis using a Molecular Dynamics Storm 840 phosphorimager. Curve fitting was performed using Kaleidagraph™ software running on a power Macintosh computer.

RESULTS AND DISCUSSION

Experimental design

We sought to optimize a method for estimating the magnitude and direction of DNA curvature in short oligonucleotides using a phasing approach in ligation ladder experiments. To validate our analysis, we studied a test case in which both the reference curvature and the test curvature were caused by A_5 tracts, of which the curve magnitude and direction in solution are well established (reviewed in 32).

This analysis requires synthesis of oligonucleotide duplexes with appropriate 3′ extensions for unidirectional ligation. Duplexes for the test case are shown in Figure 1A and B. Such duplexes are typically designed to represent two helical turns of DNA, and hence will often require 21 nt oligomers. The appropriate duplex length depends on accurate measurement of the helical repeat, as described below. An oligonucleotide standard duplex with a sequence composition comparable to the test sequence (e.g. duplex **1**, Fig. 1A) was designed to contain one $A_{5–6}$ tract per two helical turns. This sequence forms the polymer $(G_2CG_2CG_2CGCG_2CGT_5C)_n$. Other possible standards include ligated duplexes that form the polymers $(G_3CA_6CG_2CA_6C)_n$ (32) or $(G_2CG_2CG_2CGCG_2CGT_6)_n$ (13). Three experimental test duplexes are ultimately synthesized with $A_{5–6}$ tracts phased at three distances from the curvature locus of interest (‘*cis*’, ‘*ortho*’ and ‘*trans*’, e.g. duplexes **4**–**6** in Fig. 1B). However, a preliminary set of test duplexes (e.g. duplexes **2**–**4**, Fig. 1B) is first designed in lengths of 20, 21 and 22 bp for helical repeat estimation. In these duplexes, the $A_{5–6}$ tract and test sequence are placed on the same helical face. Lengths are changed by adding single base pairs to a single location to preserve overall sequence composition as much as possible. Typically, length changes are not made between the $A_{5–6}$ tract reference and the test sequence. However, this choice is irrelevant in duplexes **2**–**4** (Fig. 1) where A_5 tracts serve as both reference and test sequences. These duplexes are then used to estimate accurately the helical repeat of the sequence under study. It should be noted that even more precise estimates of the helical repeat can be made, if necessary, by analysis of 30, 31, 32 and 33 bp duplexes (40).

Gel analysis

After electrophoresis as described in Materials and Methods, data such as those shown in Figure 2 (lanes 1–6) are obtained. The distances migrated are measured manually for all radiolabeled bands in standard and experimental lanes. An exponential equation is fit by a least-squares method to a plot of distance migrated (x) as a function of length (L) for 100 bp ladder markers (e.g. Fig. 2, lanes labeled M):

$$L = \alpha e^{(\beta x)} \qquad \mathbf{1}$$

where L is length, x is distance migrated (typically measured in mm), and α and β are fit parameters. Once determined, parameters α and β are fixed and substituted back into equation **1**, and the resulting function is used to calculate the apparent length (L_{app}) for all bands in experimental ligation ladders. Electrophoresis of unligated samples allows assignment of the monomer duplex band, and this allows assignment of an actual length (L_{act}) to each gel band. Because of the sharp curvature of duplexes such as those shown in Figure 1A and B, circular ligation products may be observed. In extreme cases (e.g. Fig. 2, lanes 4 and 5) such products may predominate and obscure the linear products whose mobilities must be measured. To distinguish linear and circular ligation products, half of the ligation products were treated with *Bal*31 exonuclease to eliminate linear molecules (Fig. 2, compare lanes 1–6 with 7–12). After

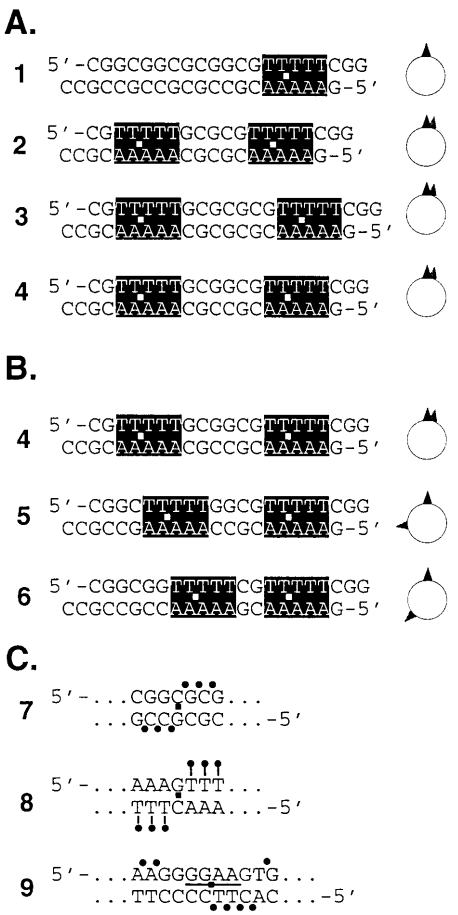

Figure 1. Experimental DNA duplexes. (**A**) Curvature standard (duplex **1**) and duplexes **2** (20 bp), **3** (22 bp) and **4** (21 bp) used to establish the helical repeat of the test sequence. A_5 tracks are in black boxes. The center of A_5 tract curvature (toward the minor groove in a reference frame 0.5 bp 3′ to the center of the A_5 tract) is indicated by a white square. End views are shown at the right: if each duplex is viewed from the right end, the black arrowhead indicates the direction of curvature due to the 3′ A_5 tract, and the gray arrowhead indicates the direction of curvature due to the 5′ A_5 tract. (**B**) Duplexes (21 bp) used to validate the analysis. Duplexes **4**–**6** correspond to *cis*, *ortho* and *trans* cases, respectively. (**C**) Sequences representing previous experimental designs re-analyzed in this work. Assigned centers of the test sequences are indicated by black squares. As noted in the text, the assigned center of the test sequence does not necessarily correspond to the actual center of curvature. This discrepancy is reflected in parameter *D*. Sequence **7** contains six methylphosphonate substitutions (dots) lining the DNA minor groove as described (13). Sequence **8** was studied in unmodified form, or with six appended propylamines or acetylated propylamines at position 5 of the indicated thymine residues (15). Sequence **9** contains a binding site for transcription factor PU.1 and was studied in unmodified form, or with methylphosphonate internucleoside substitutions at the positions indicated by dots (17).

Table 1. Parameter definitions

Parameter	Definition	Units
h	helical repeat	bp/turn
r	relative curvature	A_{5-6} tract equivalents per helical turn
r′	curvature of test duplex	degrees per two helical turns
s	distance closer to the reference A_{5-6} tract that the test sequence would need to be (in the experimental *cis* case) to achieve a 'true' *cis* geometry	bp
c	curvature direction angle	degrees
D	discrepancy between the actual center of test sequence curvature and the center of curvature assigned before the analysis	bp[a]
θ	rotary angle between the assigned center of curvature at the test sequence and the reference frame of the A_{5-6} reference sequence	degrees[b]
x	actual spacing between the center of test sequence curvature as initially assigned, and the center of A_{5-6} tract curvature in the experimental *cis* case	bp

[a]If $D > 0$, the true center of curvature of the test sequence is toward the minor groove in a reference frame *D* bp closer to the A_{5-6} tract than initially assigned. If $D > 0$, the true center of curvature of the test sequence is toward the minor groove in a reference frame *D* bp further from the A_{5-6} tract than initially assigned.
[b]Arbitrarily set to 0 for the experimental *cis* case

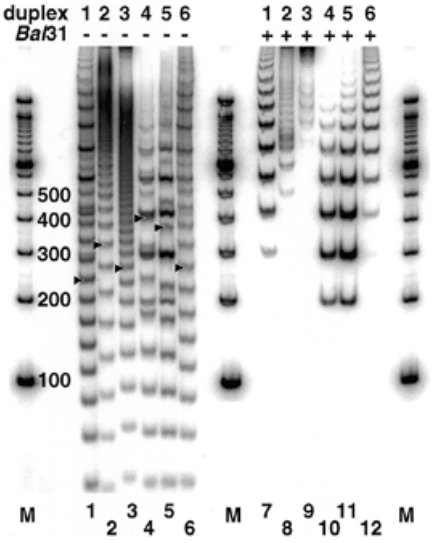

Figure 2. Ligation ladders. Image obtained after native gel electrophoresis of 100 bp marker ladders (M) or ligated duplexes **1**–**6** without (lanes 1–6) or with (lanes 7–12) exonuclease treatment to define circular molecules. Arrows to the left of lanes 1–6 indicate molecules containing nine duplexes ligated end-to-end.

circular products were identified, only linear products were analyzed. Mobility differences are apparent: the bands indicated by arrows in lanes 1–6 of Figure 2 correspond to molecules involving nine duplexes ligated end-to-end.

Helical repeat determination

Relative mobility, R_L, is defined as:

$$R_L = L_{app}/L_{act} \quad \mathbf{2}$$

Values of R_L are then plotted versus L_{act} for ligation ladders of 20, 21 and 22 bp duplexes. For duplexes **1**–**4** (Fig. 1A), the corresponding data are shown in Figure 3A. Duplex **4** (length 21 bp) consistently displays the greatest electrophoretic retardation anomaly, suggesting that the A_5 tracts remain maximally aligned on one helical face for these molecules. To determine more accurately the helical repeat, data points corresponding to three different multimers of each duplex length were replotted against $L_{duplex}/2$ as shown in Figure 3B. The data are fit to a parabolic function:

$$R_L = i(L_{duplex}/2)^2 + j(L_{duplex}/2) + k \quad \mathbf{3}$$

finding the best values for *i* and *j*. The helical repeat, *h* (bp/turn: see Table 1), is then given by:

$$h = -\frac{j}{2i} \quad \mathbf{4}$$

which corresponds to the value of $L_{duplex}/2$ at the maximum of the parabolic fit through the data. The value of *h* is typically near 10.5. If *h* is closer to 10 or 11, then 20 bp or 22 bp, respectively, test duplexes will be appropriate. Determination of helical repeat is performed for all unmodified and modified sequences before phasing studies are performed. In the case of demonstration duplexes **2**–**4** (Fig. 1), the electrophoretic data in lanes 2–4 of Figure 2, and analysis in Figure 3 yield a helical repeat estimate of 10.4 bp/helical turn (Table 2, column 4). Therefore, 21 bp phasing duplexes **4**–**6** (Fig. 1) were synthesized.

Curvature analysis

R_L values are calculated for all gel bands according to equation **2**. A plot of these data against L_{act} is displayed in Figure 4A for duplexes **1**–**4**. It is evident from inspection of the data that the electrophoretic mobility anomaly is as expected: the duplex with two A_5 tracts *in cis* (duplex **4**) displays a greater mobility anomaly than the *ortho* (duplex **5**) or *trans* (duplex **6**) cases.

Using the method of Crothers and co-workers (reviewed in 32), the data are then transformed and $R_L - 1$ is plotted versus $(L_{act})^2$. This result is shown in Figure 4B for duplexes **4**–**6**. Crothers and co-workers demonstrated that this transformation allows fitting of plots to an empirical equation in terms of $(L_{act})^2$ and (relative curvature)^2, where relative curvature (*r*) is

Table 2. DNA bending data

Test sequence[a]	Experiment	Reference	Helical repeat (h)[b]	Bend angle at test sequence (t)[c]	Bend position (D)[h]
4–6	A_5 tract	This work	10.4	18.9	0
7	Six neutral phosphates (minor groove)	(13)	10.5	17.8	0.77
8	Six propylamines (minor groove)	(15)	10.5	9.6[d]	0.27[d]
				21.4[e]	0.45[e]
				0.5[f]	0.15[f]
9	Seven neutral phosphates simulating PU.1	(17)	10.5	5.0[d]	4.0[d]
	(major groove)			25.5[g]	–5.1[g]

[a]See Figure 1.
[b]In units of bp/helical turn; standard deviation <1% of value.
[c]Degrees deflection of helix axis; standard deviation <5% of value.
[d]Unmodified.
[e]Six tethered propyl amines.
[f]Six tethered acetylated propyl amines.
[g]Seven methylphosphonates simulating PU.1.
[h]Base pair shift from the specified center of the test sequence toward the reference $A_{5–6}$ tract, required to define the frame in which the test bend is directed toward the minor groove; standard deviation <0.1 bp.

measured in units of $A_{5–6}$ tract equivalents per helical turn of DNA:

$$R_L - 1 = (pL^2 - q)(r^2) \quad \mathbf{5}$$

Data are examined for ligated duplexes in the range 105 bp < L_{act} < 189 bp (32). First, equation **5** is used to obtain the best least-squares fits for parameters p and q, by setting r = 0.5 $A_{5–6}$ tract equivalents per helical turn (for the case of standard duplex **1**). In general, the value of r for a standard duplex is the number of phased $A_{5–6}$ tracts (one for duplex **1**) divided by the number of helical turns of DNA in the duplex (two for duplex **1**). Values for p and q are ~9.6×10^{-5} and 0.47, respectively, under these electrophoresis conditions (32), but must be determined in each experiment.

Once values of p and q have been determined using equation **5** and a standard duplex such as **1**, these values for p and q are substituted back into equation **5**, and the resulting expression is used to obtain least-squares estimates for r for the *cis*, *ortho* and *trans* duplexes (duplexes **4**–**6**, Fig. 4B). The curvature (in degrees) of each test duplex (two helical turns) is r':

$$r' = 36r \quad \mathbf{6}$$

The rotary angle (θ) between the reference frame of the $A_{5–6}$ tract and the center of the test sequence is arbitrarily assigned a value of zero for the DNA duplex representing the *cis* case (duplex **2**, Fig. 1B). Values of θ for the *ortho* and *trans* cases are assigned by counting the number of base pairs from the center of $A_{5–6}$ tract curvature (0.5 bp 3′ of the center of the $A_{5–6}$ tract) to the assigned center of the test sequence, and multiplying this spacing by 360°/h. A plot of test duplex curvature (r') versus radial angle (θ) is then prepared (Fig. 4C). These data are then fit by a least-squares method to a general cosine function:

$$r' = a + b\cos(\theta - c) \quad \mathbf{7}$$

If b < 0 from the curve fit, 180° is added to (or subtracted from) c to invert the sign of b. The value of c can then be adjusted by adding or subtracting 360° such that $b \geq 0$ and $-90 < c < 90$. The curve fit for the particular data set shown in Figure 4C gave the values a = 25.88, b = –11.16, c = 202.63, which, after the conversions described above, yielded a = 25.88, b = 11.16, c = 22.63. Because equation **7** involves fitting a cosine function with three adjustable parameters to three data points, the fits will be perfect in all cases. Where possible, therefore, it is desirable to study more than three phasings to improve estimates of the fitting parameters.

The interpretation of Figure 4C is as follows. As the phasing experiment alters the relationship of the two loci of curvature in the test duplex from *cis* to *ortho* to *trans*, these loci contribute to different overall test duplex curvature values (r') measured in the electrophoretic experiments. When the cosine function reaches a maximum, this corresponds to the two loci of curvature occurring in the same plane causing maximal electrophoretic retardation (the 'true' *cis* case). The value of r' at its maximum (which occurs when θ = c) is given by $a + b$. This value also represents the linear sum of the angles of curvature of the two independent loci of curvature. Because the degree of curvature of the reference $A_{5–6}$ tract has a known value of 18°, the curvature of the test sequence, t, is simply:

$$t° = a + b - 18° \quad \mathbf{8}$$

While least-squares values of a and b allow determination of t, the value of c (converted to the range $-90 < c < 90$ as described above) provides information about the direction of the test curvature. The value of c is first converted to base pair units, s:

$$s = c \cdot h/360 \quad \mathbf{9}$$

where s gives the number of base pairs closer to the reference $A_{5–6}$ tract the test sequence would have to be (relative to the experimental *cis* case) in order to obtain the 'true' *cis* case. If s > 5 bp, then $h - s$ can be interpreted as the number of base pairs further from the reference $A_{5–6}$ tract the test sequence would have to be (relative to the experimental *cis* case) in order to

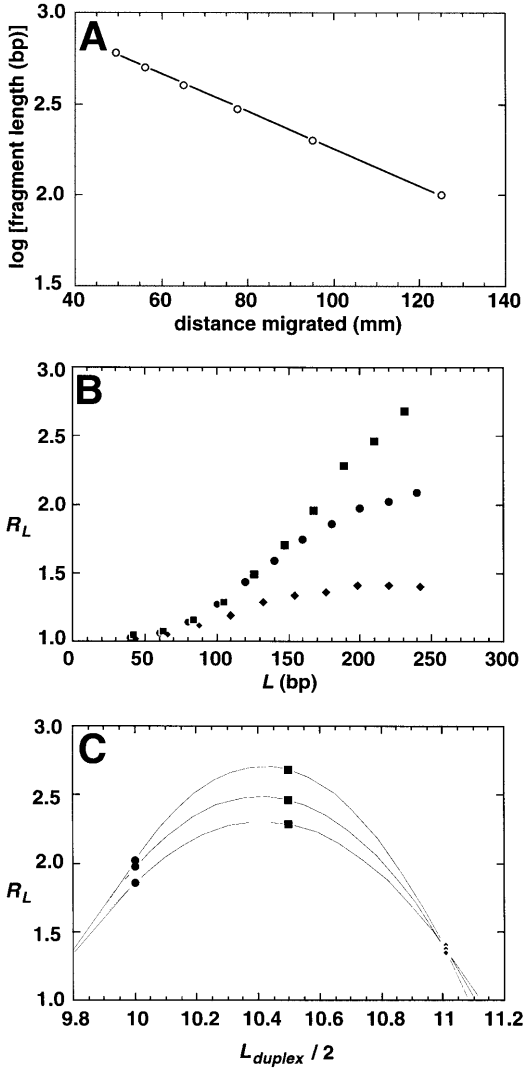

Figure 3. Helical repeat data analysis. (**A**) Fitting of equation **1** to uncurved mobility standard data. (**B**) Electrophoretic anomaly as a function of duplex length: circles, 20 bp; squares, 21 bp; diamonds, 22 bp. (**C**) Parabolic curve fits through mobility anomaly data for molecules composed of 9 (lower curve), 10 (middle curve) or 11 (upper curve) ligated duplexes. Symbols are as in (B).

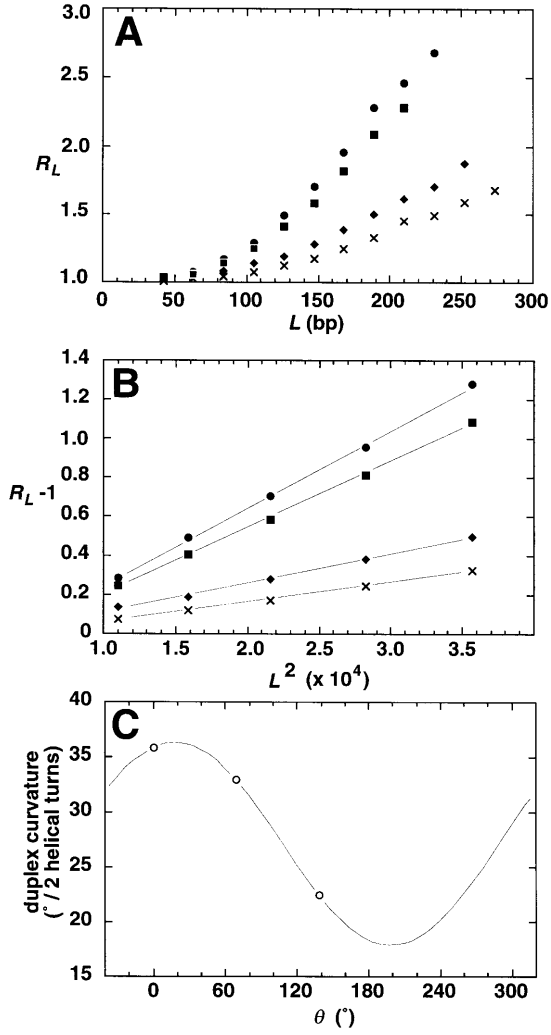

Figure 4. Curvature analysis for DNA duplexes **4**–**6** containing two phased A_5 tracts. (**A**) Electrophoretic mobility data for standard duplex **1** (crosses), *cis* duplex **4** (circles), *ortho* duplex **5** (squares) and *trans* duplex **6** (diamonds). (**B**) Data and curve fits to equation **5** for standard duplex **1** (crosses), *cis* duplex **4** (circles), *ortho* duplex **5** (squares) and *trans* duplex **6** (diamonds). (**C**) Fit of relative curvature data deduced from (B) to equation **7**.

obtain the 'true' *cis* case. At this point it is useful to change the reference frame in order to express the true bend direction relative to the assigned center of the test sequence. This transformation is accomplished by defining the spacing correction, *D*:

$$D = h - x + s \qquad \mathbf{10}$$

where *h* is the helical repeat, *x* is the actual spacing between the assigned center of the test sequence and the center of curvature of the $A_{5–6}$ tract of the *cis* experimental case, and *s* is as defined in equation **9**. *D* is a useful parameter that refers to the test sequence independent of any particular molecular spacing in the *cis*, *ortho* or *trans* cases. *D* describes the discrepancy (if any) between the actual center of bending of the test bend, and the center of bending assigned for the analysis. In particular, if $D > 0$, *D* gives the number of base pairs that the assigned center of the test sequence must be moved closer to the $A_{5–6}$ tract in order that the direction of the test curvature is toward the minor groove in that reference frame. If $D < 0$, the correct frame of reference is *D* bp further from the $A_{5–6}$ tract. Note that curvature toward the minor groove in a reference frame *D* bp from the assigned center of the test sequence is equivalent to curvature toward the major groove $D \pm h/2$ bp from the assigned center of the test sequence. Table 1 summarizes the definitions of the variables described above.

Confirmation with test duplexes

In duplexes **4**–**6** (Fig. 1B) both loci of curvature are caused by $A_{5–6}$ tracts, for which the magnitude and direction of curvature is known. The resulting data (Table 2, row 1) therefore provide a powerful confirmation of the validity of the quantitative

approach described above. The average value of *t* from three experiments is 18.9° (Table 2), in excellent agreement with the literature value of 18 ± 3° (32). The standard deviation of *t* was <5% of its value. Moreover, the average value of *D* from three experiments is 0 (standard deviation <0.1 bp). Because the center of A_5 tract curvature was assigned at the theoretical position (0.5 bp 3′ to the center of the A_5 tract), this value for *D* confirms that the assigned reference frame for curvature toward the minor groove agrees exactly with the experimental result.

Application of the optimized method to previous curvature data

The optimized curve-fitting approach described above provides estimates of both the magnitude and direction of curvature. In our previous analyses (13–17,37), the direction of curvature was only approximated by observing the ranking of the electrophoretic anomaly between *cis*, *ortho* and *trans* cases. In light of the improved analysis, we selected three previous data sets relevant to asymmetric charge neutralization in DNA and subjected them to analysis. The examples are indicated as duplexes **7**–**9** (Fig. 1), with corresponding data included in Table 2. Duplex **7** (Fig. 1C) refers to data obtained for DNA containing six racemic methylphosphonate substitutions flanking the minor groove (13). Duplex **8** (Fig. 1C) refers to the case of six propylammonium (cationic) or acetylated propylamine (neutral) modifications at the five positions of thymine bases flanking the minor groove (15). Duplex **9** (Fig. 1C) depicts the case of seven racemic methylphosphonate substitutions simulating the electrostatic consequences of transcription factor PU.1 binding to DNA (17).

Analysis of these prior data sets confirms the qualitative and quantitative results previously reported while providing higher accuracy and precise curvature direction information. In the case of the original study of asymmetric phosphate neutralization (duplex **7**, Table 2), the refined curvature magnitude is 17.8°, close to the originally reported value of 21° (13). The direction of curvature is now shown to be toward the minor groove in a reference frame ~0.8 bp (i.e. rotated ~27°) 3′ from the assigned center of the neutralized face. This direction of curvature is similar, but not identical, to that assumed in the original analysis, providing a more detailed view of the manner in which the double helix responds to the charge alteration. It is interesting that the bend direction induced at the neutralized surface is not toward the precise center of the modified sequence. Because the methylphosphonate substitutions are positioned within an asymmetric sequence, details of the local sequence context must influence the overall direction of curvature.

When data for modified pyrimidine bases is considered, the new conclusions again support previous interpretations, with additional accuracy now possible. The unmodified test sequence appears slightly curved toward the minor groove [9.6° versus 9° as originally calculated (15)], in a reference frame 0.27 bp (9° rotation) from the center of the modified sequence. When six cationic groups are appended, the angle of curvature increases (21.4° versus 17° as originally calculated) and it is now possible to define the direction of curvature as toward the minor groove 0.45 bp (15° rotation) 3′ of the center of the sequence. When the cationic groups were neutralized by acetylation, curvature was relaxed as previously observed. Curvature was 10.5° toward the minor groove in a reference frame 0.15 bp (5° rotation) 3′ of the center of the sequence. The previous analysis deduced 8° of curvature toward the minor groove at the center of the sequence under these conditions.

Data from our previous analysis of the electrostatic consequences of PU.1 binding to its DNA site (duplex **9**, Fig. 1) are also shown in Table 2. The curvature of the unmodified sequence had been reported to be 6.5° toward the major groove in a reference frame 1.5 bp 3′ from the center of the sequence (17). The improved analysis refines the magnitude of curvature to 5.0°, in similar reference frame 1.0 bp (34° rotation) 3′ from the center of the test sequence. Note that in this case the experimental duplexes were constructed with A_5 tracts upstream of the test sequence, so *D* defines a correction in the 5′ direction from the assigned sequence center, and the *D* value of 4.0 bp 5′ from the sequence center for curvature toward the minor groove (Table 2) corresponds to curvature toward the major groove *h*/2 bp 3′ from this location, or 1.25 bp (43° rotation) 3′ from the assigned center of the sequence. Consideration of data for the neutralized PU.1 test sequence yields a curvature of 25.5° [compared to the original estimate of 27.6° (17)]. The least-squares estimate for *D* is –5.1 (Table 2), demonstrating that the direction of curvature is toward the minor groove 5.1 bp 3′ from the center of the test sequence. This corresponds to curvature toward the major groove *h*/2 = 5.25 bp upstream of that position, i.e. 0.15 bp 5′ to the assigned center of the test sequence. This reference frame is nearly identical to the previously estimated curvature direction toward the major groove at exactly the assigned center of the test sequence (17).

CONCLUSION

We have presented an approach to quantitating DNA curvature in ligation ladder experiments which we believe to be superior to our previous adaptation (13) of the method of Crothers *et al.* (41). In our method, the magnitude of curvature was estimated based on the general concept that molecular cross-section reflects the vector sum of projections of curved DNA segments orthogonal the end-to-end distance vector of a central segment. Though qualitatively correct, this relationship could not be derived simply from first principles and was therefore approximate. In addition, the previous analysis did not allow us to estimate the exact reference frame of the measured curvature. It was necessary to assume that the curvature was directed toward the center of the test sequence.

The improved approach described here is superior for several reasons. The application of a cosine fitting function (equation **7**) to deconvolute curvature data is superficially reminiscent of our previous approach (13). However, the present function has a stronger theoretical basis. Moreover, directions of curvature are directly estimated by this approach. We show that when applied to a test sequence whose curvature is due to combinations of two A_5 tracts, estimates for curvature magnitudes and directions are consistent with established values. We further re-examine examples of our previous DNA curvature data obtained in studies of asymmetric charge neutralization. We show that the modified method provides superior interpretation of these data. The revised data in Table 2 may be treated with greater confidence, but show qualitative and quantitative similarity to previous estimates. The degree of correspondence between the two methods is surprising in view of the superior theoretical approach inherent in the new procedure.

Application of the analytical procedure described here will provide improved structural information relevant to understanding DNA structure and its perturbation by sequence and chemical modifications. This analysis is mathematically accessible and complements the exquisite sensitivity of the ligation ladder technique.

ACKNOWLEDGEMENTS

We thank M. Doerge in the Mayo Foundation Molecular Biology Core Facility for providing excellent oligonucleotide synthesis services, and L. Benson in the Mayo Biomedical Mass Spectrometry Core Facility for analytical assistance. This work was supported by the Mayo Foundation and NIH grant GM54411 to L.J.M. R.B.D. was a recipient of a Summer Undergraduate Research Fellowship from Mayo Graduate School.

REFERENCES

1. Hagerman,P.J. (1988) *Annu Rev. Biophys. Biophys. Chem.*, **17**, 265–286.
2. Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*. W.H. Freeman and Co., New York, NY.
3. Rippe,K., von Hippel,P.H. and Langowski,J. (1995) *Trends Biol. Sci.*, **20**, 500–506.
4. Goodman,S.D. and Nash,H.A. (1989) *Nature*, **341**, 251–254.
5. Maher,L.J. (1998) *Curr. Opin. Chem. Biol.*, **2**, 688–694.
6. Crothers,D. (1993) *Curr. Biol.*, **3**, 675–676.
7. Werner,M.H., Gronenborn,A.M. and Clore,G.M. (1996) *Science*, **271**, 778–784.
8. Werner,M.H. and Burley,S.K. (1997) *Cell*, **88**, 733–736.
9. Rice,J.A., Crothers,D.M., Pinto,A.L. and Lippard,S.J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 4158–4161.
10. Rink,S.M. and Hopkins,P.B. (1995) *Biochemistry*, **34**, 1439–1445.
11. Liberles,D.A. and Dervan,P.B. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 9510–9514.
12. Akiyama,T. and Hogan,M.E. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 12122–12127.
13. Strauss,J.K. and Maher,L.J. (1994) *Science*, **266**, 1829–1834.
14. Strauss,J.K., Roberts,C., Nelson,M.G., Switzer,C. and Maher,L.J. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 9515–9520.
15. Strauss,J.K., Prakash,T.P., Roberts,C., Switzer,C. and Maher,L.J. (1996) *Chem. Biol.*, **3**, 671–678.
16. Strauss-Soukup,J.K. and Maher,L.J. (1997) *Biochemistry*, **36**, 8692–8698.
17. Strauss-Soukup,J.K. and Maher,L.J. (1997) *J. Biol. Chem.*, **272**, 31570–31575.
18. Shui,X.Q., Sines,C.C., McFail-Isom,L., VanDerveer,D. and Williams,L.D. (1998) *Biochemistry*, **37**, 16877–16887.
19. Hud,N.V., Schultze,P. and Feigon,J. (1998) *J. Am. Chem. Soc.*, **120**, 6403–6404.
20. Hud,N.V., Sklenar,V. and Feigon,J. (1999) *J. Mol. Biol.*, **286**, 651–660.
21. Young,M.A., Jayaram,B. and Beveridge,D.L. (1997) *J. Am. Chem. Soc.*, **119**, 59–69.
22. Young,M.A. and Beveridge,D.L. (1998) *J. Mol. Biol.*, **281**, 675–687.
23. Dlakic,M., Park,K., Griffith,J.D., Harvey,S.C. and Harrington,R.E. (1996) *J. Biol. Chem.*, **271**, 17911–17919.
24. Parkhurst,K.M., Brenowitz,M. and Parkhurst,L.J. (1996) *Biochemistry*, **35**, 7459–7465.
25. Heyduk,E., Heyduk,T., Claus,P. and Wisniewski,J.R. (1997) *J. Biol. Chem.*, **272**, 19763–19770.
26. Vacano,E. and Hagerman,P.J. (1997) *Biophysical J.*, **73**, 306–317.
27. Crothers,D.M., Drak,J., Kahn,J.D. and Levene,S.D. (1992) *Methods Enzymol.*, **212**, 3–29.
28. Kahn,J.D. and Crothers,D.M. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 6343–6347.
29. Parvin,J.D., McCormick,R.J., Sharp,P.A. and Fisher,D.E. (1995) *Nature*, **373**, 724–727.
30. Sitlani,A. and Crothers,D.M. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 3248–3252.
31. Sitlani,A. and Crothers,D.M. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 1404–1409.
32. Crothers,D.M. and Drak,J. (1992) *Methods Enzymol.*, **212**, 46–71.
33. Rice,J.A. and Crothers,D.M. (1989) *Biochemistry*, **28**, 4512–4526.
34. Brukner,I., Susic,S., Dlakic,M., Savic,A. and Pongor,S. (1994) *J. Mol. Biol.*, **236**, 26–32.
35. Han,W., Dlakic,M., Zhu,Y.J., Lindsay,S.M. and Harrington,R.E. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 10565–10570.
36. Ussery,D.W., Higgins,C.F. and Bolshoy,A. (1999) *J. Biomol. Struct. Dyn.*, **16**, 811–823.
37. Tomky,L.A., Strauss-Soukup,J.K. and Maher,L.J. (1998) *Nucleic Acids Res.*, **26**, 2298–2305.
38. Puglisi,J.D. and Tinoco,I.,Jr (1989) *Methods Enzymol.*, **180**, 304–325.
39. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
40. Drak,J. and Crothers,D.M. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 3074–3078.
41. Koo,H.S. and Crothers,D.M. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 1763–1767.