Helical repeat and chirality effects on DNA gel electrophoretic mobility

JACQUELINE DRAK AND DONALD M. CROTHERS

Department of Chemistry, Yale University, New Haven, CT 06511

Contribution by Donald M. Crothers, December 19, 1990

ABSTRACT We determined a value of 10.34 ± 0.04 base pairs (bp) per turn for the helical repeat of bent DNA sequences of the form A_{n}N_{2}-A_{n}N_{2} by estimating the sequence repeat required to produce a planar curve, as judged from the maximum in the electrophoretic mobility anomaly of multimers containing different sequence repeats (10.00, 10.33, 10.50, 10.67, and 11.00 bp per turn). This result provides the basis for a method to evaluate the helical repeat of any DNA segment by comparative electrophoresis measurements. The sequence of interest is placed between two A-tract bends and the phasing is varied over an entire helical turn. Knowledge of the number of base pairs between the bends in the cis isomer, which has the lowest electrophoretic mobility, allows calculation of the average helical repeat of the inserted sequence. In the course of these experiments we observed an unexpected dependence of electrophoretic mobility on the shape of DNA molecules: in high-percentage polyacrylamide gels, those bent molecules for which we deduced a right-handed superhelical form are less retarded than their homologous left-handed isomers. To explain this finding we propose that superhelical chirality influences the choice of DNA migration pathway, leading to rotation of the DNA molecule relative to the local coordinate frame in the gel. High-percentage gels have sufficiently close contact with the right-handed DNA helical twist to differentiate the frictional consequences of right- and left-handed twisting motions.

The fact that curved DNA sequences have drastically reduced mobilities in polyacrylamide gels has been widely used to gain understanding of the origin and nature of DNA bending. Gel electrophoresis has been the method of choice to detect and localize bent sequences (1–5) and also to determine the extent and magnitude of a bend, both for intrinsically bent sequences and for protein-induced bends (6–11). Despite the wealth of experimental data, there is still no quantitative theory relating the electrophoretic mobility with a set of molecular parameters. Our current understanding of the electrophoresis of linear DNA is largely based on the reptation model proposed by De Gennes (12) to explain the dynamic properties of concentrated polymers. In the application of his model to DNA (13–16), it was proposed that DNA chains migrate in a snakelike fashion through a "tube" in the gel under the influence of the electric field. The tube is determined by constraints that the gel fibers impose on the longitudinal motion of the DNA. The configuration of the tube changes only when the leading segment of the chain makes an electric field-biased (but otherwise random) step among the gel fibers. Ultimately, the orientation of the leading segment determines the path of the entire chain.

Under the set of simplifying assumptions of Lumpkin and Zimm (15), one can derive the equation of translational motion of the DNA along the "tube" and obtain the following expression for the chain's mobility $\mu$:

$$\mu = \frac{Q}{\zeta} \left( \frac{h_r^2}{L^2} \right),$$

where $Q$ is the total effective charge of the chain, $L$ is the contour length of the tube, $\zeta$ is the friction coefficient for motion along the tube, and $h_r$ represents the component of the tube's end-to-end vector in the field direction. The angle brackets denote the average over an ensemble of conformations. However, this simple reptation model cannot explain (17), among other things, why the effect of a bend is more pronounced when it is located at the center of the molecule rather than at the end: the conformation of the tube should not be sensitive to the conformation of any of the internal segments. Nonetheless, despite its deficiencies, this model provides a starting point to understand experiments in which electrophoresis is used as a tool to extract structural information pertaining to bent molecules.

We have previously determined the magnitude of the bend induced by an A tract by comparing the experimentally determined ring closure probabilities for systematically bent molecules with the calculated values from Monte Carlo simulations (18). As the calculations proved to be very sensitive to the value of the helical repeat, we determined it by the following experiment. We synthesized A-tract-containing oligonucleotides such that the sequence repeat was 10.00, 10.33, 10.50, 10.66, and 11.00 base pairs (bp)/turn in five different monomers (see Fig. 1) and separated the corresponding ligation ladders on a gel. According to Eq. 1, we expect the most retarded molecule of a given length to have the shortest mean-square end-to-end distance, which results from a close match between sequence repeat and helical repeat, yielding an essentially planar curved molecule.

In an experiment designed to determine the helical repeat of normal B DNA in solution, we constructed two sets of isomeric molecules, each containing two separate large bends (lbDNA) or a large and a small bend (sbDNA) produced by runs of adenines repeated every 10.5 bp (Fig. 2). The phasing between the two bends was varied over one helical turn in steps of 2 bp by the insertion of sequences $n$, $n + 2$, ..., $n + 10$ bp long between them (8). The end-to-end distance of these molecules varies from a minimum value when the two bends are perfectly in phase (cis isomer) to a maximum for the molecule in which the bends are 180° out of phase (trans isomer). As previously (8), we expect the cis and trans isomers to be the slowest- and fastest-moving fragments, respectively, of the isomeric set. For the bends to be perfectly phased in the cis isomer, the distance between their centers must be a multiple of the average helical repeat. However, these experiments revealed that the determination of which isomer is cis depends on the polyacrylamide gel.

Abbreviations: lbDNA and sbDNA, large-bend and small-bend DNA, respectively; CAP, catabolite gene activator protein.
Fig. 1. Sequences of the oligonucleotides used in the cloning of the isomeric sets and of the top strands of the molecules used in the ligation ladder experiment. The numbers in parentheses indicate the corresponding sequence repeat. Bottom strands have the same length as their complementary top strand but their ends are shifted by 3 bp, so that the duplexes could have a 5' end overhang for the ligation reaction.

percentage, an effect that is more pronounced for molecules with larger bends. We traced this phenomenon to the ability of high-percentage gels to discriminate between DNA molecules having right- or left-handed intrinsically superhelical shapes.

**MATERIAL AND METHODS**

**Preparation of the Multimers of Synthetic Duplexed Oligonucleotides.** The oligonucleotides in Fig. 1 were made on a DNA synthesizer (Applied Biosystems) and purified on either 15% or 20% polyacrylamide gels in the presence of 50% (wt/vol) urea. Each purified oligonucleotide (8 μg) was 5'-labeled with [γ-32P]ATP (1 μCi/μl; 1 Ci = 37 GBq) in 70 mM Tris-HCl, pH 8.0/10 mM MgCl₂/5 mM dithiothreitol and T4 polynucleotide kinase (0.7 unit/μl). After 15 min at 37°C, 5–10 more units of kinase were added, the solution was made 0.5 M in unlabeled ATP, and the reaction was allowed to proceed for 30 min more at 37°C. Complementary oligonucleotides were annealed by heating to 75°C and slowly

cooling to room temperature. Ten to 20 μl of annealed solution was self-ligated overnight at 4°C by using 800 units of DNA ligase and 2.5 μl of 10 mM ATP in the same buffer as the kinase reaction. Ligase was quenched by addition of EDTA (pH 8.0) to 25 mM final concentration.

**Ligation Ladder Experiment: Measurement of Gel Mobilities and Gel Calibration.** Aliquots of the ligation mixtures (1.5–2 μl) were loaded on 5%, 8%, 12%, and 16% nondenaturing polyacrylamide gels (39:1 acrylamide/bisacrylamide) in 1× TBE buffer (89 mM Tris-borate/2 mM EDTA) and subjected to electrophoresis at 10–12 V/cm for 8–32 hr. The gels were then dried and autoradiographed. It was necessary to run size markers along with the bent DNA ligation ladders to help calibrate the gel. For the 5% and 8% gels, we used two different ligation ladders, one derived from BamHI linkers and the other from a 21-bp oligonucleotide of mixed sequence that we took as "unbent" DNA (see Fig. 1). For the high-percentage gels we also digested the vector pUC19 (Promega) with the enzyme Dde I or Dde I and EcoRI and end-labeled the digestion products to obtain high molecular weight markers. For each gel percentage, a plot of log(bp) vs. migration distance for the markers run on that gel proved to be linear, and an equation was fit to that calibration plot. The retardation coefficient Rₗ obtained for each fragment as the ratio of apparent length to true length was plotted vs. sequence repeat length for the multimers that contained either 12 or 15 A tracts. The position of the maximum of the plot was determined by fitting the data to a second- or third-degree polynomial.

**Cloning Vectors.** The parent constructs (kindly provided by S. Zinkel, Harvard University) consisted of a set of isomeric molecules containing a sequence-directed bend (produced by 6 A tracts phased every 10.5 bp) phased against a catabolite gene activator protein (CAP) binding site and one helical turn in steps of 2 bp. The CAP binding site was removed by digesting the vectors with BamHI and the ends were dephosphorylated by treatment with calf intestinal alkaline phosphatase. The dephosphorylated vectors were separated from the BamHI insert on a 5% polyacrylamide gel, visualized by UV shadowing, excised from the gel, electroeluted and precipitated with ethanol. To prepare the bent DNA insert, monomer A was first dimerized by ligation (Fig. 1); the dimer was isolated on an 8% gel and further purified. After the ends had been filled in it was ligated to BamHI linkers (New England Biolabs) and digested overnight with BamHI. The insert was precipitated, ligated onto the six previously dephosphorylated vectors, and cloned to produce the lbDNA set. The DNA was sequenced and purified by centrifugation in CsCl using standard procedures (19, 20). The sbDNA set of bent molecules was obtained from lbDNA by digesting with EcoRI, dephosphorylating, and separating the long fragment from the A-tract-containing insert on a 5% gel. The new set of dephosphorylated vectors was visualized, purified, and ligated to monomer B. Cloning, DNA sequencing, and plasmid preparation were as described above.

**Gel Electrophoresis.** DNA molecules from either set of isomers were digested with Rsa I and Pvu II or with Rsa I and Psi I and electrophoresed on polyacrylamide/TBE gels of various percentages. The acrylamide-to-bisacrylamide ratio was always 39:1. The gels were run in a constant-temperature apparatus (Hoefer) at 22°C. Gels were stained with ethidium bromide and photographed.

**Calculation of the DNA Helix Trajectory.** We used the junction bending model to simulate the DNA trajectories using the same computer program and bend angles as in ref. 18. The helical repeat of the sequence A₆N₄-A₆N₅ having been determined from the ligation ladder experiment, the only parameter in the simulation was the helical twist of the linker region between the two bends, hₗ (pending further experiments we are assuming that the helical repeat is
independent of the nature of N). We varied $h_B$ in steps of 0.02 bp/turn and calculated the mean-square end-to-end distance ($h^2$) for all the isomers of the set. We compared the variation of the normalized mean-square end-to-end distance with linker length (obtained by dividing ($h^2)/L^2$ by the average value for the set) to the experimental mobilities on a 5% gel for the different constructs. A narrow range of $h_B$ values was found to give good agreement between experiment and theory.

RESULTS

Determination of the Helical Repeat of Bent DNA from A-Tract Ligation Ladders. Fig. 3 shows that a plot of $R_L$ against sequence repeat for a given multimer has a well-defined maximum; we associate the corresponding sequence repeat with the helical repeat since exact match between sequence and helical repeats produces planar molecules of minimum gel mobility. The position of the maximum is better defined in the higher-percentage gel experiments, although for reasons that will be discussed shortly, it might be preferable to rely on the low- and intermediate-percentage gel data. Our results indicate that the helical repeat of sequences of the type $A_N-N_A-A_N$ is $10.34 \pm 0.04$ bp/turn and that within experimental error this value does not change with gel percentage.

Electrophoretic Properties of cis/trans Isomers. The result of digesting the lbDNA set (containing large bends) with the enzymes Rsa I and Pvu II and running it on a gel is illustrated Fig. 4. In a 5% gel the mobilities of the A-tract-containing fragment did not vary in a sinusoidal fashion as a function of linker length. Furthermore, when the gel percentage was increased to 12%, the slowest and fastest moving isomers were 6 bp longer and 2 bp shorter than their counterparts on the lower-percentage gel. We thought that the extremely large bend in these molecules could be hindering their reptation motion through the gel matrix, and we constructed a new set of molecules (sbDNA) containing two bends of about 60° and 110°, respectively, instead of the original 130° and 110° (lbDNA set). The mobilities of sbDNA on a 5% gel varied as expected (Fig. 5A). Having identified the apparent cis isomer, we simulated the DNA trajectories and calculated the mean square end-to-end distance of the bent fragments. If $h_{A6}$ is taken to be 10.35 bp/turn, a value of $h_B = 10.49 \pm 0.05$ bp/turn, which is within the expected range, gives the best fit to the experimental results.

Mobility Anomalies That Depend on Gel Percentage. We investigated the properties of the sbDNA restriction digest when run on gels of different percentage. To our surprise, the mobilities on an 8% gel showed the same irregularities as in the lbDNA isometric set, as evidenced in Fig. 5B. The mobility changes for the molecules on both sides of the trans isomer are particularly striking: one of them moves faster (relative to the average mobility of the set), whereas the other

![Fig. 3. Results of the ligation ladder experiment. The retardation coefficient ($R_L$) of multimers containing 12 A tracts (A) or 15 A tracts (B) was plotted against sequence repeat for different gel percentages (●, 5%; □, 8%; ○, 12% polyacrylamide gel). The experimental results were fit to a second- or third-degree polynomial. The arrows indicate the position of the maximum of each plot; the abscissa at that point corresponds to the helical repeat of the ligated sequences.](image)

![Fig. 4. Rsa I–Pvu II digest of the first isomeric set (lbDNA) run on two different polyacrylamide gels (A, 5%; and B, 12%) at room temperature (22°C). The numbers on top indicate the linker length. The arrows point to the DNA fragment containing the bent sequences. The slowest- (●) and fastest- (○) moving isomers are indicated.](image)

![Fig. 5. Rsa I–Pst I digest of the second isomeric set (sbDNA) run on three different percentage polyacrylamide gels (A, 5%; B, 8%; and C, 16%) at room temperature. The notation is the same as in Fig. 4.](image)
DISCUSSION

We have described two different experiments designed to determine the helical repeat of DNA sequences. Comparative electrophoresis of ligated multimers of DNA at variable sequence phasing provides a convenient way to estimate the helical repeat of bent sequences of the form $A_N^2A_N^2A_N^2N_1$, which we found to be $h_{A6} = 10.34 \pm 0.04$ bp/turn. We varied $h_B$ to try to fit the experimental results that we obtained at high gel percentage. To account for the change in cis isomer when the gel concentration is increased from 5% to 16%, the helical repeat of the B-DNA has to decrease from $\approx 10.50$ bp/turn to $9.79 \pm 0.05$ bp/turn. This value is far outside the expected range, from which we conclude that the helical repeat should be determined from the limiting properties at low gel percentage. We address below the possible causes of the anomalous mobilities in high gel percentages, which may also be related to the anomalous behavior observed by Zinkel (21) for similar isomeric molecules containing a CAP-induced bend.

We then stimulated the DNA trajectories of the isomers of the sbDNA isomeric set assuming $h_{A6} = 10.34 \pm 0.04$ bp/turn. We varied $h_B$ to try to fit the experimental results that we obtained at high gel percentage. To account for the change in cis isomer when the gel concentration is increased from 5% to 16%, the helical repeat of the B-DNA has to decrease from $\approx 10.50$ bp/turn to $9.79 \pm 0.05$ bp/turn. This value is far outside the expected range, from which we conclude that the helical repeat should be determined from the limiting properties at low gel percentage. We address below the possible causes of the anomalous mobilities in high gel percentages, which may also be related to the anomalous behavior observed by Zinkel (21) for similar isomeric molecules containing a CAP-induced bend.

In the lower gel percentage experiments, the observed mobilities for the isomeric sets of bent DNAs are in good qualitative agreement with the calculated mean-square end-to-end distances, as predicted by the Lumpkin-Zimm equation, with a helical repeat $h_A$ in the expected range (Fig. 7). However, anomalies became evident at increased gel percentage. One possible explanation for this effect, that as the gel matrix becomes tighter the molecules are distorted in such a way that their helical repeat varies as a function of gel percentage, is inconsistent with the results of the ligation ladder experiment at different gel percentages.

With respect to the changes in mobility observed for the two molecules on each side of the plane of the trans isomer, it is apparent that the helix axes of these molecules describe superhelices of opposite hands. The molecule whose mobility increases with gel percentage (relative to the average mobility of the set) is right-handed, whereas the other one is left-handed. It is possible that if the twisting modulus of these DNA molecules were nonconstant and asymmetrical, right- and left-handed supercoils would have different distributions of deformation as they move through the gel. To our knowledge, there is only one report of a second-order correction to the quadratic dependence of the free energy of supercoiling on the superhelical winding number (22). That correction is
there is no coupling of frictional effects with the handedness of the rotation of the DNA helix, and some "global" parameter that describes the molecule's three-dimensional solution conformation (for example the mean-square end-to-end distance) can be used to predict its relative electrophoretic mobility.

Our results at different gel percentages suggest that in all cases there is some memory of the original solution conformation of the bent molecules. At low gel percentages this is reflected in the qualitative agreement between the experimental results and Eq. 1, although the fractional variation in experimental mobilities is smaller than predicted from the solution conformations. At high gel percentages this "memory" is demonstrated by the strong modulation of the mobilities of molecules that differ only in the handedness of their superhelical shape.

Memory of solution conformation within the gel matrix is contrary to the simple view that the molecule's path through a rigid gel depends only on the leading end of the chain, which is not curved in the molecules used throughout these studies (Fig. 2). A possible modification of the model incorporates flexibility in the gel, which allows the molecule to tend to adopt a shape which reflects its minimal energy conformation. Levene and Zimm (17) recently analyzed the electrophoretic mobility of bent DNA molecules in terms of the elastic modulus for the DNA-gel combination; our results indicate that gel flexibility makes a significant contribution to this effect.

We thank Dr. S. Zinkel for constructs and for stimulating discussions. This work was supported by Grant GM21966 from the National Institutes of Health.