CA/TG Sequence at the 5' End of Oligo(A)-tracts Strongly Modulates DNA Curvature*

(Received for publication, June 23, 1993, and in revised form, November 12, 1993)

Akhilesh K. Nagaich‡‡, Dhananjay Bhattacharyya†, Samir K. Brahmacari‡‡, and Manju Bansal†¶
From the ‡Molecular Biophysics Unit and the ¶Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012, India

An analysis of the base pair doublet geometries in available crystal structures indicates that the often reported intrinsic curvature of DNA containing oligo(dA)-d(T) tracts may also depend on the nature of the flanking sequences. The presence of CA/TG doublet in particular at the 5' end of these tracts is expected to enhance their intrinsic bending property. To test this proposition, three oligonucleotides, d(GAAAAAC-CCCTCC), d(CCCCCCAAAAAG), and d(GAAAAATTTC), and their complementary sequences were synthesized to study the effect of various flanking sequences, at the 5' and 3' ends of the A-tracts, on the curvature of DNA in solution. An analysis of the polyacrylamide gel electrophoretic mobilities of these sequences under different conditions of salts and temperatures (below their melting points) clearly showed that the oligomer with CA/TG sequence in the center was always more retarded than the oligomer with AC/GT sequence, as well as the oligomer with AT/AT sequence. Hydroxy radical probing of the sequences with AC/GT and CA/TG doublet junctions gives a similar cutting pattern in the A-tracts, which is quite different from that in the C-tracts, indicating that the oligo(A)-tracts have similar structures in the two oligomers. KMnO4 probing shows that the oligomer with a CA/TG doublet junction forms a kink that is responsible for its inherent curvature and unusual electrophoretic mobility. UV melting shows a reduced thermal stability of the duplex with CA/TG doublet junction, and circular dichroism (CD) studies indicate that a premelting transition occurs in the oligomer with CA/TG doublet step before global melting but not in the oligomer with AC/GT doublet step, which may correspond to thermally induced unbinding of the oligomer. These observations indicate that the CA/TG doublet junction at the 5' end of the oligo(A)-tract has a crucial role in modulating the overall curvature in DNA.

The anomalous electrophoretic mobility of A-tract-containing DNA has become closely identified with DNA curvature (1–6). There are several reports of DNA curvature in a wide variety of biological systems; in many cases, the loci of curvature are found in regions of functional importance, i.e. origin of replication, promoters, etc., suggesting the functional importance of curvature per se (7–8). The effect of oligo(A)-tracts has been attributed to the intrinsic features of the AA doublet sequence, namely the "wedge model" (9), as well as to discontinuities at the junction between random sequence DNA and oligo(A)-tracts, the so-called "junction model" (10–12). Theoretical Monte Carlo simulations of oligo(A)-tract-containing polymers also indicate that these tracts show features that are in agreement with both these models (13, 14). Thus, while it is now generally accepted that the presence of A-tracts, repeating in a phased manner, coinciding with the helical repeat in DNA causes the helix axis to curve, the actual physical nature of the bending locus is not yet clearly elucidated. Recently it has been suggested that non-(A).-tract containing sequences can also have slight curvature (15, 16), and presence of CA/TG dinucleotide steps has been shown to be important in DNA curvature (17). Since the oligonucleotide crystal structures give the most accurate details of DNA structure, we have analyzed the local geometries of the various base pair doublets from all available crystal structures and used this data to predict the curvature of some well characterized DNA sequences (18, 19). Our analysis indicated that the sequences flanking the A-tract are also quite crucial in determining the overall curvature of DNA. In particular the presence of a CA/TG doublet at the 5' end can considerably enhance the intrinsic bending traits of an oligo(A)-tract. It is interesting to note that the CA/TG step is present in both the naturally occurring best DNA from kinetoplasts and the synthetic sequences studied by Crothers and co-workers that are characterized as being intrinsically curved (10). In addition, the synthetic promoter sequences that have been reported to have a bending locus and in turn a strong influence on transcription activation are also preceded by a C residue at the 5' end of A-tracts (20). A detailed sequence analysis of the natural DNA indicates that oligo(A)-tracts are preferentially flanked by the complementary base T, but these sequences have not generally been characterized as curved (21). Most of the gel mobility studies have been carried out using polymers with a variety of sequences in the non-A-tracts and by their very nature show the cumulative effects of oligo(A)-tracts, as well as the junctions at both 5' and 3' ends. An investigation of multimers of d(C2A2C2)n, d(C2A2G2)n, and d(G2A2G2)n shows a decreasing order of anomaly in gel mobility, indicating the importance of the sequences flanking the A-tract in inducing the overall curvature (22). However, in this case also it is difficult to differentiate the effects due to various junctions. An oligomeric sequence with an oligo(A)-tract, juxtaposed on one side only with other sequences, can help in understanding the role of doublet sequences at the 5' and 3' ends. An earlier study (23) has already shown that even small oligomers such as the decamers d(GA4TnC) and d(GT4AnC) have different electrophoretic mobilities, a feature attributed to the differences in curvature between the two oligomers, arising due to AT and TA sequences at the center. We therefore decided to investigate the effect of CA/TG doublet sequence on the mobility of short oli-
gonucleotides containing oligo(A)-tracts and to compare it with the related AC/ GT and the earlier characterized AT/AT sequences. We chose dodemercan fragments to increase the length to diameter ratio for the molecules and, hence, their anisotropic shape. We find that oligomers d(GA\textsubscript{10}C\textsubscript{10}) and d(C\textsubscript{10}A\textsubscript{10}G\textsubscript{10}), which have identical base composition and A-tract length, show different electrophoretic mobilities, as well as differences in the geometry at their central junction doublet as indicated by hydroxyl radical and K\textsubscript{MnO\textsubscript{4}} probing. This sequence-dependent structural distortion in oligomer d(C\textsubscript{10}A\textsubscript{10}G\textsubscript{10}) is also reflected in its thermodynamic behavior as revealed by UV melting and CD analysis. These observations confirm that the origin of greater bending in the oligomer d(C\textsubscript{10}A\textsubscript{10}G\textsubscript{10}) as compared to the oligomer d(GA\textsubscript{10}C\textsubscript{10}) lies in the unique geometry of its central CA/TG junction doublet.

**MATERIALS AND METHODS**

**Model Building of Sequence-dependent Polymeric DNA Molecules**—In an earlier analysis, we have shown that the mean values of local doublet parameters (24), obtained by considering the B-DNA crystal structures data, available in the January 1991 release of the Brookhaven Protein Data Bank (25), can provide a structural basis for the sequence-dependent electrophoretic migration of some well-characterized genomic sequences (19). Similar analysis has now been carried out for synthetic polynucleotide sequences, using the data from the June 1993 update of the Nucleic Acid Data Base (26), which includes crystal structure data for several new oligonucleotides. The mean values of the local doublet parameters for the 10 unique doublet sequences (considering parameters for CC, TT, TG, TC, GT, and CT doublets as equivalent to their complementary base sequences, i.e., GG, AA, CA, GA, AC, and AG respectively) are given in Table I. It is found that inclusion of data from protein-DNA complexes does not lead to any significant differences in the mean parameter values, indicating that the local parameters are quite well defined unless the DNA undergoes a large protein-induced distortion. Hence the protein-DNA complex data have been included to obtain the parameters for the sequences AC and AG, for which only a few data points are available even in the current oligonucleotide data base. Model structures of several polynucleotide sequences have been generated following the procedure described earlier (27).

The generated structures have been characterized by two different parameters. (i) The ratio of the end-to-end distance to the actual path length \(l_{\text{end}}\) was calculated; (ii) the three principle moments of inertia parameters. (i) The ratio of the end-to-end distance to the actual path distance for the molecules and, hence, their anisotropic diameter ratio for the molecules and, hence, their anisotropic structure was calculated. (ii) Four principle moments of inertia were calculated based on temperature-dependent \(W\) absorption data assuming a two-state model of helix to coil transition reproducible.

**Preparation and Purification of Oligonucleotides**—Two pairs of complementary oligonucleotides d(GA\textsubscript{10}C\textsubscript{10}) and d(C\textsubscript{10}A\textsubscript{10}G\textsubscript{10}) and, as well as the self-complementary sequence d(GA\textsubscript{10}C\textsubscript{10}) were synthesized on an Applied Biosystems 381A DNA synthesizer equipped with a 96-well plate handler. B-DNA crys-

**Polyacrylamide Gel Electrophoresis**—Equimolar amounts of complementary oligonucleotides d(GA\textsubscript{10}C\textsubscript{10}) and d(C\textsubscript{10}A\textsubscript{10}G\textsubscript{10}) were radio labeled using \([\gamma^32P]dATP\) and polyadenylate kinase. The phosphorylated oligomers were run on a 20% polyacrylamide gel containing 8 \(\mu\) urea, eluted in water, and de- salted as described previously. Labeled oligonucleotides of appropriate radioactive were mixed with their complementary oligonucleotides, phosphorylated with cold dATP and dialyzed in 2 \(m\) sodium cacodylate, 0.1 \(m\) EDTA, and 50 \(m\) NaCl, pH 7.2 (35 \(m\)), heated at 90 °C for 3 min, and slowly cooled to room temperature to achieve efficient annealing. The oligomers were preincubated at 4 and 20 °C for 15 min before reaction. The hydroxyl radical cleavage reactions were initiated by adding 15 \(\mu\)l of a mixture containing 0.2 \(m\) FeSO\textsubscript{4}. \(N\)\textsubscript{2}H\textsubscript{4}SO\textsubscript{4}.H\textsubscript{2}O and 0.4 \(m\) EDTA (5 \(m\)), 0.6% hydrogen peroxide (5 \(m\)), and 20 \(m\) sodium salt of L-ascorbic acid (5 \(m\)). The reactions were carried out at appropriate temperatures for 5 min and terminated by adding 0.1 \(m\) thiourea (20 \(m\)). The DNA was precipitated using 3 \(m\) sodium acetate (25 \(m\)), ethanol (75 \(m\)), and tRNA (10 \(mg\)) at -70 °C. The DNA was pelleted and washed three times with 70% ethanol, dried, and suspended in DNA sequencing gel loading buffer containing (80% formamide in 1 x TBE, bromphenol blue, and xylene cyanol dye). A part of the reaction mixture was loaded on a 20% denaturing polyacrylamide gel containing 8 \(\mu\) urea. The electrophoresis was performed at 40 V/cm, and the gels were autoradiographed. Each hydroxyl radical probing experiment was repeated three times to ensure reproducibility of the results.

**UV Melting Analysis**—The temperature-dependent absorption spectra were measured at 260 \(nm\) using a Beckman DU-8B spectrophotometer equipped with a Peltier temperature-controlled and programmed cell holder that accommodates five cuvettes. Melting profiles of oligonu-

**Circular Dichroism**—CD spectra were recorded in the spectral range of 220 to 320 \(nm\) on a JASCO J-500A automatic spectropolarimeter equipped with a DP 501 data processor and thermostatically controlled cuvette holder. The samples were prepared by heating at 95 °C for 3 min in 5 x 10\(^{-4}\) m strand concentration followed by gradual cooling to room temperature. Prior to scan, samples were incubated to give a strand con-

**DISCUSSION**

**Theoretical Analysis of DNA Bending**—Model structures were generated (using the set of base pair step parameters listed in Table I) for polynucleotide sequences, with decamer.
The mean local doublet parameters (18, 24) obtained from analysis of all available crystal structures of B-form DNA for the 10 unique dinucleotide sequences studied are listed in Table I. The CA/TG sequence modulates DNA curvature.

### Table I

<table>
<thead>
<tr>
<th>Doublet sequence</th>
<th>Tilt (°)</th>
<th>Roll (°)</th>
<th>Twist (°)</th>
<th>Shift (D₁)</th>
<th>Slide (D₂)</th>
<th>Rise (D₃)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG (72)</td>
<td>0.1</td>
<td>3.2</td>
<td>35.1</td>
<td>0.1</td>
<td>0.5</td>
<td>3.53</td>
</tr>
<tr>
<td>GC (39)</td>
<td>-0.2</td>
<td>-6.5</td>
<td>38.6</td>
<td>-0.2</td>
<td>0.5</td>
<td>3.34</td>
</tr>
<tr>
<td>AA (49)</td>
<td>-1.7</td>
<td>2.6</td>
<td>35.6</td>
<td>-0.1</td>
<td>0.2</td>
<td>3.23</td>
</tr>
<tr>
<td>GA (35)</td>
<td>-0.9</td>
<td>0.3</td>
<td>38.5</td>
<td>-0.0</td>
<td>-0.1</td>
<td>3.10</td>
</tr>
<tr>
<td>AT (30)</td>
<td>0.2</td>
<td>-0.3</td>
<td>32.5</td>
<td>0.1</td>
<td>-0.4</td>
<td>3.29</td>
</tr>
<tr>
<td>GG (20)</td>
<td>-1.8</td>
<td>5.8</td>
<td>32.4</td>
<td>0.1</td>
<td>0.7</td>
<td>3.47</td>
</tr>
<tr>
<td>CA (10)</td>
<td>0.6</td>
<td>-7.5</td>
<td>47.9</td>
<td>0.1</td>
<td>2.9</td>
<td>3.40</td>
</tr>
<tr>
<td>AC (15)</td>
<td>-0.2</td>
<td>-0.7</td>
<td>32.3</td>
<td>-0.2</td>
<td>-0.1</td>
<td>3.55</td>
</tr>
<tr>
<td>AG (13)</td>
<td>-0.5</td>
<td>7.1</td>
<td>39.6</td>
<td>-0.1</td>
<td>0.4</td>
<td>3.48</td>
</tr>
<tr>
<td>TA (07)</td>
<td>0.0</td>
<td>0.5</td>
<td>39.0</td>
<td>-0.2</td>
<td>0.2</td>
<td>3.22</td>
</tr>
<tr>
<td>Average</td>
<td>-0.5</td>
<td>0.5</td>
<td>36.1</td>
<td>-0.1</td>
<td>0.5</td>
<td>3.35</td>
</tr>
</tbody>
</table>

Thus, the mean local doublet geometries, as obtained from all the available B-DNA crystal structures, can explain the anomalous gel migration data of polynucleotides of repetitive sequences, as well as that of the well characterized genomic sequences. In view of the static model considered by us, the correlation between the theoretical models and the gel retardation data is quite good, indicating that the average molecular shape is correctly predicted. The extensive Monte Carlo simulation studies also show only a qualitative agreement with the experimental data, with the theoretical models for the few sequences studied, showing a much smaller variation than expected from the gel mobility data (13). The values of local parameters reported by other workers were deduced from theoretical consideration or by empirical fitting of cyclization and gel mobility data. It is interesting to note that the roll and twist values for dinucleotide steps other than AA, CA, and GG are found to be very similar in crystal structures and from the gel data (11, 12). However, there is a basic assumption that the repeats containing oligo(A)-tracts of varying length. The values obtained for the ratios of end-to-end distance to the actual path length (d/\text{end}) are shown in Table II. The calculated values of the ratio of the largest to smallest principal moments (\text{I}_{\text{max}}/\text{I}_{\text{min}}) also show a similar relation, i.e., both d/\text{end} and \text{I}_{\text{max}}/\text{I}_{\text{min}} values are relatively large for those sequences that are straight (with \text{R}_{\text{C}} values close to 1.0) and are smaller for the curved fragments.

Model structures were also generated for the restriction fragments of kinetoplast DNA from Crithidia fasciculata (I, II, and III) and Leishmania tarentolae (IV) and are shown in Fig. 1, while the bending parameters are included in Table II. It is clearly seen that fragments II and IV, which show abnormally retarded mobility, are predicted to be curved, while fragments I and III are straight as expected from their \text{R}_{\text{C}} values (29, 30). A few models of random nucleotide sequences with different percentages of AT content have also been built for comparison and the relevant parameters for a few representative sequences are given in Table II. It is clearly seen that random sequence polynucleotide models have large values for d/\text{end} (i.e., in the range of 0.93 to 1.0) as well as I_{\text{max}}/I_{\text{min}} (\approx 30), indicating that these model structures are nearly straight but can occasionally be slightly bent. This is in agreement with the deduction from their gel migration data (31) that the random sequence structures can also show a slight amount of curvature.

Thus, the mean local doublet geometries, as obtained from all the available B-DNA crystal structures, can explain the anomalous gel migration data of polynucleotides of repetitive sequences, as well as that of the well characterized genomic sequences. In view of the static model considered by us, the correlation between the theoretical models and the gel retardation data is quite good, indicating that the average molecular shape is correctly predicted. The extensive Monte Carlo simulation studies also show only a qualitative agreement with the experimental data, with the theoretical models for the few sequences studied, showing a much smaller variation than expected from the gel mobility data (13). The values of local parameters reported by other workers were deduced from theoretical consideration or by empirical fitting of cyclization and gel mobility data. It is interesting to note that the roll and twist values for dinucleotide steps other than AA, CA, and GG are found to be very similar in crystal structures and from the gel data (11, 12). However, there is a basic assumption that the

### Table II

The structural parameters describing the bending of the predicted structural models of 150-base pair fragments for a few synthetic sequences with decamer repeats and some genomic DNA sequences are listed, along with the experimentally observed relative mobility (R_{C}) of the fragments in polyacrylamide gel.

<table>
<thead>
<tr>
<th>Synthetic polymers</th>
<th>Percentage AT content</th>
<th>\text{R}_{\text{C}}</th>
<th>d/\text{end}</th>
<th>I_{\text{max}}/I_{\text{min}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCGAAACCCG</td>
<td>30</td>
<td>1.23</td>
<td>0.45</td>
<td>3.9</td>
</tr>
<tr>
<td>GCGCAAAACCG</td>
<td>40</td>
<td>1.60</td>
<td>0.52</td>
<td>4.4</td>
</tr>
<tr>
<td>GCGCAAAACG</td>
<td>50</td>
<td>2.00</td>
<td>0.42</td>
<td>3.6</td>
</tr>
<tr>
<td>GCAAAAAAAC</td>
<td>60</td>
<td>2.31</td>
<td>0.38</td>
<td>3.2</td>
</tr>
<tr>
<td>GCAAAAAAAA</td>
<td>80</td>
<td>2.21</td>
<td>0.35</td>
<td>3.1</td>
</tr>
<tr>
<td>CAAAAAAA</td>
<td>90</td>
<td>1.73</td>
<td>0.57</td>
<td>4.8</td>
</tr>
<tr>
<td>CAATTGTG</td>
<td>80</td>
<td>2.23</td>
<td>0.50</td>
<td>2.4</td>
</tr>
<tr>
<td>CTTTTAAGG</td>
<td>80</td>
<td>1.04</td>
<td>0.78</td>
<td>9.8</td>
</tr>
</tbody>
</table>

**Genomic sequences**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Percentage AT content</th>
<th>\text{R}_{\text{C}}</th>
<th>d/\text{end}</th>
<th>I_{\text{max}}/I_{\text{min}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>45</td>
<td>1.09</td>
<td>0.92</td>
<td>55.8</td>
</tr>
<tr>
<td>II</td>
<td>63</td>
<td>1.18</td>
<td>0.89</td>
<td>16.3</td>
</tr>
<tr>
<td>III</td>
<td>52</td>
<td>1.00</td>
<td>0.97</td>
<td>90.7</td>
</tr>
<tr>
<td>IV</td>
<td>57</td>
<td>3.78</td>
<td>0.77</td>
<td>8.0</td>
</tr>
</tbody>
</table>

**Computer-generated 150-mer sequences**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Percentage AT content</th>
<th>\text{R}_{\text{C}}</th>
<th>d/\text{end}</th>
<th>I_{\text{max}}/I_{\text{min}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCCCCAAAA</td>
<td>50</td>
<td>0.32</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Random 1</td>
<td>20</td>
<td>0.93</td>
<td>32.7</td>
<td></td>
</tr>
<tr>
<td>Random 2</td>
<td>50</td>
<td>0.97</td>
<td>59.7</td>
<td></td>
</tr>
<tr>
<td>Random 3</td>
<td>80</td>
<td>0.99</td>
<td>139.5</td>
<td></td>
</tr>
</tbody>
</table>

* The genomic sequences are the restriction fragments from kinetoplast DNA of C. fasciculata (I, II, and III) (29) and L. tarentolae (IV) (39) of length 274, 211, 150, and 414 base pairs, respectively.

AA step has a negative roll angle (10–13) and the CA/TG step has a roll angle opposite to that of AA/TT step, so that a CA step occurring out of phase with the center of the oligo(A)-tract implicitly enhances the curvature, although it was not explicitly mentioned by the authors of these studies. On the contrary, the mean roll angle for the 49 AA/TT steps observed in the oligonucleotide crystal structures, solved using high resolution data, has a small positive mean value (+2.6°) with a standard deviation of 3.7° and hence a better agreement with the gel data is obtained if the accompanying CA steps are taken to have the negative roll, high twist combination (18, 32) that is
seen in several crystal structures. A small number of CA steps have also been observed with positive roll (bending into major groove) and small twist, as suggested from Monte-Carlo simulations.

In our analysis, it is seen that a major determinant for polynucleotide curvature is the kink induced at the CA/TG dinucleotide, which is present in almost all the curved sequences, both synthetic and natural. To further characterize the unusual features of the AC, AT, and CA doublets, models of three related oligomers, namely C5A5, A5C5, and A5T5, were generated. The C5A5 model has a much larger kink or bending angle as compared to A5C5 and A5T5 decamer sequences as seen in Fig. 2, where the base pairs of the A5C5 and C5A5 model structures are shown along with vectors indicating directions of the local helix axes, positioned at each base pair center. The intrinsic curvature of the C5A5 sequence is even more clearly seen when the decamer sequence is repeated to generate a 150-mer structure (Fig. 3A). The ratio d/lnax for this polymer is calculated to be 0.32, while that for the related sequence (GCCGCCGG), studied by Crothers and coworkers, is 0.42 (Fig. 3B). Both these sequences have a CA step at the 5' end of the A-tract and an AC step at the 3' end; however, whereas (C5A5), has only CC steps in the non-A-tract, the Crothers sequence has GG, GC, and CG steps. The large bending predicted for the polymer of C5A5 sequence is in agreement with the high gel retardation reported for the similar sequence (C5A5)9 (22). To examine whether the unusually large kink predicted for the C5A5 oligomer on the basis of our theoretical analysis is actually observed in solution, three oligonucleotides (d(GA5C6), d(GA5T5C), and d(Cd5G)) were synthesized and with the high gel retardation reported for the similar sequence d(GA5T5C) (lane 6). These results clearly suggest that the oligomer containing a CA/TG doublet junction at the center is considerably more retarded and presumably more curved than the duplexes containing AC/GT and AT/AT doublet steps.

In order to check whether these differences in mobility reflect differences in the relative stability of duplexes, leading to a process of strand dissociation during migration or differential binding of counter ions by different sequences, the gel migration was studied at various temperatures and salt concentrations. Heating to 15 °C does not substantially alter the differential mobilities of the duplexes (data not shown). However, at higher temperatures the bands diffuse during electrophoresis making it difficult to distinguish the differential gel mobility of the oligomers. Single-stranded oligomers under identical conditions show much higher electrophoretic mobility than duplexes. In presence of 12 mM MgCl2 (Fig. 4B), the oligomer d(CGA5G) (lane 2) is significantly retarded, as compared to d(GA5C6) (lane 1) and d(GA5T5C) (lane 3), indicating that the anomalous gel migration of the three oligomers is an intrinsic property of the duplexes and is not due to strand dissociation or differential interaction with counter ions. Since duplexes d(GA5C6) and d(Cd5G) have identical molecular weight, base composition, and charge distribution, the difference in their relative gel mobility can only be rationalized on the basis of their structural differences.

Hydroxyl Radical Probing—The global measure of bending monitored by gel mobility of DNA molecules offers no direct information on the structural anomaly at the bend loci. The observation that similar sequences that differ only in the relative position of adenine tracts have different electrophoretic mobilities prompted us to investigate in detail the structural origin for the difference in their relative curvature. Taking a cue from the hydroxyl radical probing studies done on kinetoplast DNA and the polymer sequences d(CGAT5T5)3 and d(CGTA5G)3 (33, 34) (where a sinusoidal cutting pattern arising, presumably, from progressive narrowing of minor groove in the A-tracts has been linked to a curved DNA, and an even cutting pattern has been linked to a straight DNA), we investigated whether such a feature is observed in the case of short oligomers.

The frequency of hydroxyl radical cleavage at each nucleotide of the oligomers d(GA5C6), d(Cd5G), and d(GA5T5C) at two different temperatures, 4 and 20 °C, is shown in Fig. 5 (A and B). The densitometric scans of the autoradiograms are
shown in Fig. 5 (C–E). The A-tracks in both the oligomers d(C6A5C) and d(GA6C6) show reduced cleavage at 4 °C (lanes 1 and 2, respectively, in panel A), as compared to the C-tracks, indicating the unique nature of the oligo(A)-tracts in both the sequences, particularly at low temperature. The central CA step in the oligomer d(C6A5C) shows a remarkably high susceptibility to cleavage at 4 °C as compared to any other doublet step, which may indicate its unusual structure (Fig. 5C). At 20 °C the A-tracks and the C-tracks of both the oligomers show nearly uniform hydroxyl radical cleavage (Fig. 5A, lanes 3 and 4), indicating the presence of sequence-independent structures.

The oligomer d(GA6T5C) shows a marginally reduced cutting around the central region at 4 °C (lane 1) and a fairly even cutting frequency at 20 °C (Fig. 5B, lane 2). It is interesting to note that this temperature-induced transition in the A-tracks of all the oligomers is observed well below their respective Tm values under identical NaCl concentration. A premelting transition in A-tracks has been reported earlier by chemical probing studies in the case of longer synthetic sequences indicating that oligo(A)-tracts are capable of forming more than one structure before global melting (35, 36). However, we do not observe a progressive decrease in the intensity of bands in the A-tracks, similar to that reported previously (33, 34) for polymeric sequences and attributed to the narrowing of minor groove in the region. In a separate hydroxyl radical probing study of 30-mers containing three oligo(A)-tracts of varying lengths, phased with the helical repeat, we have found a progressive decrease in band intensity in the A-tracks. Since the oligomers reported here are smaller and oligo(A)-tracts are flanked only on one side, the groove width in their A-tracks may not be severely constrained.

**KMnO4 Probing**—KMnO4 has been used as a probe to study the conformational flexibility of short adenine tracts present in DNA (35, 36). An in vivo probing of B-Z junction sequences having a T residue at the junction has also been reported using KMnO4 as a probe (37). Recently, oligonucleotides d(G16T2G2) (n = 2–4), adopting a hairpin G-quartet structure by dimerization of hairpin loops, have also been probed in our laboratory using KMnO4 (38). Since KMnO4 reacts with the DNA primarily via oxidation of the C5-C6 double bond of pyrimidines (T>>C), stacking of the bases causes protection from this reaction (39–40). When the base stacking is disrupted, the susceptiblity of thymines, to react with KMnO4 increases significantly (41, 42).

We investigated the geometry of the dinucleotide junction steps AC/CT and CA/TG present in oligomers d(GA6C6) and d(C6A5C) using KMnO4 as a probe. The oligomer d(GA6C6) does not react with KMnO4 at 4 °C, while in the oligomer d(C6A5C), the T at the CA/TG junction shows a remarkable reactivity at this temperature (Fig. 6A, lanes 1 and 2). At 20 °C the T-tracks in the oligomers show reactivity with KMnO4 (Fig. 6A, lanes 3 and 4). Similarly, in the duplex d(GA6T5C), all the T residues remain totally unreactive at 4 °C, but at 12 °C some reactivity is observed, which is considerably enhanced at 20 °C (Fig. 6B, lanes 1, 2, and 3). All the T residues in the single stranded oligonucleotides show very high reactivity under similar conditions. This clearly indicates that, at low temperatures, the sequence d(C6A5C) has a conformation having a more pronounced kink at CA/TG junction compared to that of the AC/CT junction in the oligomer d(GA6C6) or the AT/AT junction in oligomer d(GA6T5C). In order to see whether the sequence-dependent structural distortion of the oligomers leading to different electrophoretic mobilities, as well as the difference in the geometry of central junction doublet, as revealed by KMnO4 and hydroxyl radical probing, is reflected in their global conformation and thermodynamic properties, we analyzed their melting behavior, as well as CD spectra.

**UV Melting Analysis**—The UV melting profiles of the three dodecamer sequences d(GA6C6), d(C6A5C), and d(GA6T5C) at 260 nm, in 50 mM NaCl are shown in Fig. 7A. This absorbance versus temperature data was converted into a fractional helicity (a) versus T melting curve (Fig. 7B) and the transition midpoint (Tm) of each oligomer was determined by d(a/dT)−1 versus temperature (°C) plot (Fig. 7C). The van’t Hoff transition enthalpy was calculated from the general form of the van’t Hoff equation.

**FIG. 5.** A, autoradiograph of the gel comparing the hydroxyl radical cleavage pattern of the oligomers d(CGA_G) (lane 1) and d(GAC_G) (lane 2) at 4°C and d(CGA_G) (lane 3) and d(GAC_G) (lane 4) at 20°C. The arrow indicates the enhanced cleavage at the CA/TG junction of the duplex d(CGA_G) at 4°C. The absolute intensities of the bands in different lanes cannot be compared, as they result from two separate reactions. The relative intensities of the bands within the lane can be compared, since they result from a single reaction (33, 34). B, autoradiograph of the gel comparing the hydroxyl radical cleavage pattern of the duplex d(GAT_TG) at 4°C (lane 1) and at 20°C (lane 2), respectively. C, densitometric scans of the hydroxyl radical cleavage pattern of the duplex d(CGA_G) at 4°C and 20°C, i.e. lanes 1 and 3 of panel A. D, densitometric scans of the hydroxyl radical cleavage pattern of the duplex d(GAC_G) at 4°C and 20°C, i.e. lanes 2 and 4 of panel A. The A-tract shows much reduced cleavage as compared to C-tract at 4°C. E, densitometric scans of the hydroxyl radical cleavage pattern of the duplex d(GAT_TG) at 4°C and 20°C, i.e. lanes 1 and 2 of panel B.
CA/TG Sequence Modulates DNA Curvature

\[
\Delta H = (2 + 2n) \frac{RT}{\alpha^2} \left( \frac{\alpha}{\alpha} \right)_{T_m}
\]

where \( n \) is the molecularity and \( T_m \) is the transition midpoint (43). The values of \( \Delta G \) were calculated by determining \( K_{eq} \) at \( T_m \), extrapolating it to 25 °C and substituting this value in the standard equation \( \Delta G = -RT \ln K_{eq} \). Assuming \( \Delta C_p = 0 \), these values were substituted in the equation \( \Delta G = \Delta H - T \Delta S \) to calculate entropy and the values obtained are listed in Table III. It is clear from Fig. 7A that helix to coil transition in the oligomer d(Cd5G) is broader and less cooperative than in oligomers d(GA5Cs) and d(GA5T5C). The curved oligomer d(Cd5G) also displays thermal activity in the premelting domain, as is obvious in Fig. 7C. The duplex d(Cd5G) has lower \( T_m \) (40.3 °C) as compared to duplex d(GA5Cs) (44.8 °C), while the duplex d(GA5T5C) has a much lower \( T_m \) value (34 °C), as expected from its higher AT content. The oligomer d(GA5Cs) also shows large changes in enthalpy (\( \Delta H = -55.7 \text{ kcal/mol} \)) as well as free energy (\( \Delta G = -11.7 \text{ kcal/mol} \)) due to the highly cooperative helix to coil transition observed for this oligomer (Fig. 7A). Since the oligomers d(GA5C6) and d(Cd5G) differ only in the sequence at the central junction step, a difference of about 20 kcal/mol in \( \Delta H \) can only be correlated to the loss of stacking energy at the CmG step in the oligomer d(Cd5G).

Circular Dichroism Studies—We have employed circular dichroism measurements to study the equilibrium melting and temperature-dependent conformational change in all three decamers. The CD spectra of the three duplexes have been recorded in 50 mM NaCl in cacodylate buffer at different temperatures. The spectra at a few selected temperatures have been plotted in Fig. 8 (A–C). At low temperature (5 °C), the curved duplex d(Cd5G), with a CA/TG doublet at the junction, displays a spectrum characteristic of B-DNA, with a positive peak at 265 nm and a negative peak at 242 nm (Fig. 8A). Increasing the temperature up to 35 °C causes an increase in positive ellipticity accompanied by a small decrease in the negative ellipticity. Further increase in temperature up to 75 °C causes a decrease in positive, as well as negative, ellipticity, which indicates melting of the duplex structure. The spectrum of the relatively straight oligomer d(GA5Cs), with AC/TG junction, has a broad positive peak with maxima at 260 nm and a shoulder at 280 nm and a negative peak at 245 nm at 5 °C (Fig. 8B). Increasing temperature up to 35 °C does not cause any substantial change in the positive ellipticity, and further increase in the temperature causes global melting of the duplex. Considering the transition midpoint of both these oligomers under identical salt concentration, it can be postulated that the duplex d(Cd5G) undergoes a premelting transition before global melting. The self complementary duplex d(GA5T5C) shows some of the features of poly(d(A))-poly(d(T)) and exhibits a positive peak at 282 nm, a shoulder at 262 nm, and a negative peak at 247 nm at 5 °C (Fig. 8C). The shoulder at 262 nm was also observed earlier for the polymer d(GA5T5C) and has been attributed to the base pairs at the AT/AT junction having large propeller twist (44). The disappearance of this shoulder in the duplex d(GA5T5C) as the temperature increases indicates a thermally induced rearrangement in the oligomer before global melting. It is interesting to note that the oligomer d(Cd5G), which is kinked, displays...
spectral features of normal B-DNA, as revealed by CD spectra, suggesting that there are only marginal differences in the conformation of oligo(A)-tracts and C-tracts as compared to random sequence DNA and that the bending in the oligomer arises mainly due to the characteristic geometry at the central junction step, which is not necessarily reflected in the CD spectra.

The oligomers d(CoA5G) and d(GA5C6) also do not undergo a B → A transition on lowering the water activity up to 60% ethanol, despite the presence of oligo C-tracts (data not shown). It has been proposed earlier that temperature-dependent change in negative ellipticity in the DNA usually monitors duplex to single-strand transition, whereas change in the positive ellipticity with temperature gives an indication about the intramolecular helix to helix transition (45). In order to facilitate a comparison in the premelting behavior of all the three oligomers, the change in positive ellipticity at [6]max was carefully measured. Results of four such measurements along with their mean deviations are given in Table IV. The increase in positive ellipticity from 8.1 × 10^3 degrees cm^2 dmol^-1 to 9.7 × 10^3 degrees cm^2 dmol^-1 in the duplex d(CoA5G) at 265 nm, with increase in the temperature from 5 °C to 35 °C, clearly shows a premelting transition before global melting. Our results are in partial agreement with the earlier spectroscopic and calorimetric studies done on poly(d(A))poly(d(T)), kinetoplast DNA, and bent decamer d(GA5T6C), which show a premelting behavior in the oligo A-tracts prior to the global melting (45, 46). A comparison in the melting behavior of sequences d(GA5T6C) and d(GT5A6C) had shown that the former had greater thermal stability than the latter, a feature that was attributed to the spine of hydration associated with the A-tract in this sequence. However, in our study the bent oligomer d(CoA5G) shows a premelting behavior and lower thermal stability as compared to the straight oligomer d(GA5C6). The trends seen in the temperature-dependent gel mobility and the change in the ellipticity of the positive peak in the spectra of all the three dodecamers suggest that the structural features responsible for the unusual electrophoretic migration of duplexes d(CoA5G) and d(GA5T6C) also account for their premelting behavior prior to global melting and strand dissociation. This lends further support to our gel mobility data that the oligomers during the electrophoretic migration are not the kinetically trapped species but thermodynamically stable structures.

Role of CA/TG Sequence in DNA Bending—The predictions of relative bending of DNA oligomers on the basis of observed local geometries of various doublet sequences in crystal structures are corroborated by our studies on the dodecamer sequences. In particular, they highlight the important role of the CA/TG sequence at the 5’ end of an oligo(A)-tract. This doublet seems to have structural features that are different from the other two pyrimidine-purine doublets TA and CG. It has been suggested from empirical data analysis (12) that both CA/TG and AC/GT doublets play a passive role by allowing large trans-
CA/TG Sequence Modulates DNA Curvature

**Table IV**

<table>
<thead>
<tr>
<th>Temperature</th>
<th>d(GA4C6)</th>
<th>d(C6A4G)</th>
<th>d(GA5T6C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>[θ]_200 x 10^{-3}</td>
<td>[θ]_200 x 10^{-3}</td>
<td>[θ]_200 x 10^{-3}</td>
</tr>
<tr>
<td>5</td>
<td>9.9 ± 0.1</td>
<td>8.1 ± 0.1</td>
<td>8.2 ± 0.1</td>
</tr>
<tr>
<td>15</td>
<td>10.2 ± 0.05</td>
<td>9.2 ± 0.05</td>
<td>8.9 ± 0.05</td>
</tr>
<tr>
<td>25</td>
<td>10.2 ± 0.05</td>
<td>9.6 ± 0.05</td>
<td>8.2 ± 0.05</td>
</tr>
<tr>
<td>35</td>
<td>10.2 ± 0.05</td>
<td>9.7 ± 0.05</td>
<td>7.5 ± 0.1</td>
</tr>
<tr>
<td>55</td>
<td>8.3 ± 0.2</td>
<td>8.8 ± 0.1</td>
<td>5.4 ± 0.1</td>
</tr>
<tr>
<td>75</td>
<td>6.8 ± 0.2</td>
<td>7.2 ± 0.1</td>
<td>3.8 ± 0.1</td>
</tr>
</tbody>
</table>

While the observation that isolated CA steps in random sequence DNA modulate DNA bending has been made by Bolshoy et al. (12), they in fact suggested that it is a locus for a flexible hinge and should reduce bending. A recent gel circularization study by Harrington et al. (48) on point mutations and single base mismatches in the O93 site of λ phage and its complex with the Cro protein also indicates that oligonucleotides with CA, CAC, and CACA sequence elements are anisotropically flexible and facilitate protein-induced DNA bending. Circular dichroism studies carried out by these authors show that the middle of the O93 operator region, which has a CAA sequence element, is sensitive to overwinding upon Cro binding. NMR studies carried out by Petel et al. (49) suggest an unusual structure for CAC/GTG triplet, possibly involving a partial unstacking in one strand. Our KMnO₄ probing results indicate large destacking at CA/TG doublet step in the oligomer d(C6A4G), making the T residue at the junction accessible to the probe. The destacking seems to be enhanced by the presence of a neighboring A-tract, but it remains to be seen to what extent it is affected by the length of the adjoining A-tract. We believe that CA/TG step in the sequence CA'A can readily adopt the dinucleotide step geometry involving negative roll, large positive slide, and very large twist. This conformation with a
large twist can explain Harrington's observation that the CAA sequence in the middle of the 5'3' sequence facilitates overwinding of the DNA upon binding with Cro protein, as well as our observation that a CA sequence at the 5'end of A-tracts causes a static kink with base destacking.

Acknowledgments—We thank the oligonucleotide synthesis facility at the Indian Institute of Science. We thank Dr. M. S. Shailla and P. Balagurumoorthy for careful editing of the manuscript.

REFERENCES