Pyrimidine 5-Methyl Groups Influence the Magnitude of DNA Curvature†

Paul J. Hagerman

Department of Biochemistry, Biophysics, and Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262

Received November 15, 1989; Revised Manuscript Received December 20, 1989

ABSTRACT: DNA containing short sequences of the form $(dA)_n$·$(dT)_n$ can exhibit pronounced degrees of stable curvature of the helix axis, provided that these homooligomeric stretches are approximately in phase with the helix repeat. However, the precise origin of this effect is unknown. We have observed that pyrimidine 5-methyl groups can have a significant effect on the degree of curvature, depending on their locations within the homooligomeric sequences. Such effects are observed in both $(dA)_n$·$(dT/dU)_n$ and $(dI)_n$·$(dC/d5^{me}C)_n$ sequence motifs, arguing for a general structural perturbation due to the methyl group. The current observations suggest that pyrimidine methyl groups could influence protein–DNA interactions not only through direct protein–methyl group contacts but also by methyl group induced alterations in local DNA structure.

Sequence-directed curvature of DNA has become established as a nearly ubiquitous phenomenon in both prokaryotic and eukaryotic genomes [for a review, see Hagerman (1990)]. A hallmark of curved DNA molecules is their tendency to display reduced electrophoretic mobilities on polyacrylamide gels compared to noncurved molecules of the same length (Marini et al., 1977; Simpson, 1979; Challberg & Englund, 1980; Stellwagen, 1983; Hagerman, 1984; Wu & Crothers, 1984). Systematic investigations of DNA molecules of defined sequence, both naturally and synthetically derived, have identified sequence blocks of the form $(dA)_n$·$(dT)_n$ ("A-tracts") as being the principal determinant of curvature (Wu & Crothers, 1984; Hagerman, 1984, 1985, 1986, 1988; Diekmann, 1986, 1987; Koo et al., 1986; Koo & Crothers, 1987); however, knowledge of the precise structural basis for curvature remains elusive.

On the basis of their computational studies of A-tract DNA structure, Jernigan et al. (1986, 1987) proposed that the close approach of the thymine (T) methyl groups in the major groove, facilitated by favorable hydrophobic interactions, would tend to bend the helix axis through a narrowing of the major groove. An immediate prediction of their model was that removal of the 5-methyl groups should result in the straightening of the helix axis, with concomitant normalization of electrophoretic mobility. Using a single, U-substituted A-tract sequence, Koo and Crothers (1987) failed to observe any significant change in apparent curvature, leading them to conclude that 5-methyl residues play no role in determining curvature. However, since those authors only examined one sequence, their conclusion may not be a general one. This last issue is important, since current conformational energy computations (Srinivasan et al., 1987; Olson et al., 1989) also fail to display a significant effect of the methyl groups on curvature.

In order to more carefully examine the possible influence of the 5-methyl group on curvature, a series of synthetic DNA molecules were constructed in which the positions of the methyl groups were varied in a systematic fashion for both A-tract and I-tract sequence motifs. The basic conclusion of the current investigation is that methyl groups do significantly influence curvature and that such effects are position-dependent.

† This research was supported by Grant GM28293 from the National Institutes of Health.

Materials and Methods

Synthesis and Purification of Oligonucleotides. All oligodeoxyribonucleotide decamers employed in the current study were either synthesized manually as described previously (Hagerman, 1985) or produced with a Biosearch Model 8600 automated DNA synthesizer. T-to-U substitutions were carried out by selective replacement of the deoxythymidine phosphoramidite by the deoxyuridine phosphoramidite. Similarly, C-to-5^{me}C replacement utilized the 5-methyldeoxycytidine phosphoramidite. All phosphoramidites were purchased from American BioNuclear. Cleavage of the oligonucleotides from the CPG supports, base deprotection, and subsequent purifications were also performed as described previously (Hagerman, 1985). The DNA oligomers were phosphorylated with polynucleotide kinase (Bethesda Research Labs, BRL) at 37 °C in KB buffer (0.5 mM ATP, 50 mM Tris-HCl, pH 8, and 10 mM $MgCl_2$).

Purification of the 53 Base Pair Reporter Arm. The reporter arm used in the current study comprises the following sequence (one strand shown):

5′-agctTGCGCAGCTGCAATTCCACATGTGGAATTG
TGAGCGGATAACAATTTGTGG

where the upper-case letters refer to paired bases and the lower-case letters refer to a *Hin*dIII-generated overhang. The above fragment was derived from a 159-bp *Eco*RI/*Hin*dIII fragment (Taylor & Hagerman, 1990) by cleavage with *Hae*III, followed by preparative purification as described elsewhere (Cooper & Hagerman, 1987).

Construction of DNA Molecules Containing n Contiguous Decamers Flanked by 53 Base Pair Reporter Arms. DNA molecules of the form N_{53}-$(decamer)_n$-N_{53}, were constructed by performing mixed-ligation reactions in which the 53-mer and the decamer of interest (blunt-ended duplex) were present in various molar ratios. Ligation reactions were allowed to proceed overnight at room temperature and were terminated by extraction with redistilled, buffered phenol, followed by extraction with diethyl ether. Since the 53-mer possesses a single blunt end, its ligation to the decamers is directional, with the *Hin*dIII end facing outward. Therefore, following ligation, the ether-extracted reaction mixes were diluted 5-fold with Core buffer (BRL) and digested with *Hin*dIII (BRL). The resulting digests were phenol/ether-extracted, followed by precipitation with ethanol overnight at −20 °C.

0006-2960/90/0429-1980$02.50/0

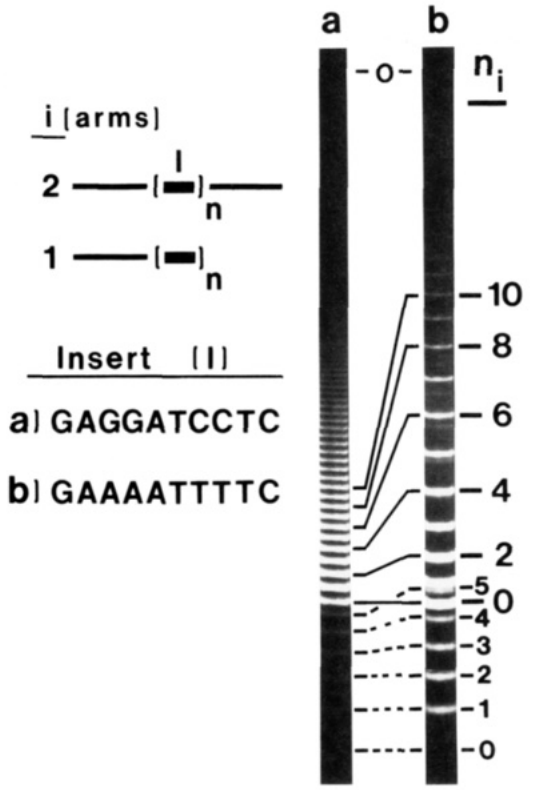

FIGURE 1: Outline of the current experimental approach. A variable number, n, of copies of a decamer insert are placed between 53-bp reporter arms by directed ligation, and the resultant molecules are run on polyacrylamide gels. The small-print numbers to the right of lane b refer to species containing one reporter arm and the large-print numbers to two-arm species. (O) refers to the origin of electrophoresis. The two lanes displayed in this figure are taken from a single, 9% polyacrylamide gel. All members of the ladder in lane a display normal electrophoretic mobilities.

Gel Electrophoresis. Polyacrylamide gels were poured and run at room temperature with buffer recirculation. The gel dimensions were 0.8 mm (thickness) × 14 cm (width) × 24 cm (length). Running currents were kept below 20 mA to avoid significant heating of the gels. The gel running buffer was 40 mM Tris–acetate, 20 mM sodium acetate, and 1 mM NaEDTA, pH 7.9. For the A (T/U) series of decamers, 9% polyacrylamide gels were used; for the I ($C/^{me}C$) series, 12% gels were used in order to enhance the electrophoretic anomalies associated with the I·C-containing molecules. The gels were always poured with a 37:1 monomer to bis ratio.

Results

Construction of DNA Molecules Having Sets of Curved Decamers Nested between Electrophoretically Normal Reporter Arms. Most of the information pertaining to the sequence dependence of curvature has derived from the study of the electrophoretic behavior of sets of DNA molecules produced by partial ligation of double-stranded oligomers (Hagerman, 1985, 1986; Koo et al., 1986; Koo & Crothers, 1987; Diekmann et al., 1987). These "ligation ladders", each comprising a set of molecules containing sequences of the form $(\text{oligomer})_n$, all display normal electrophoretic behavior for members of the set $(\text{oligomer})_n$ having $n \leq 5$ and maximal electrophoretic retardation for $n \geq 15$. Comparisons of electrophoretic mobilities among various oligomer sequence motifs are therefore usually made for members of each ladder having $n = 15$–20. Differences in relative mobility between two molecules having the same number of base pairs are almost invariably attributed to differences in axial curvature; however,

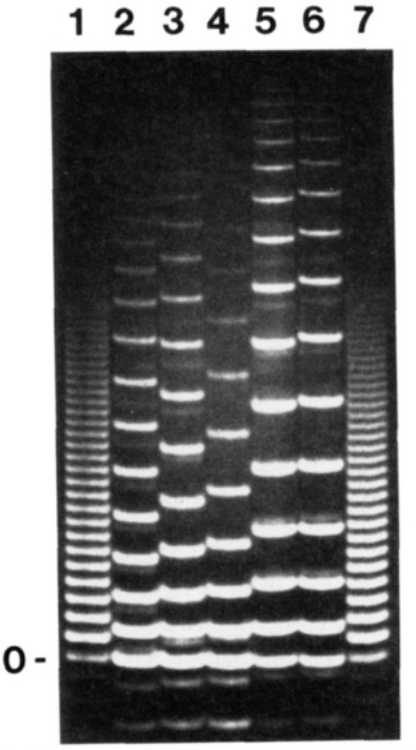

FIGURE 2: Electrophoretic behavior of A-tract-containing molecules of the form N_{53}-$(\text{decamer})_n$-N_{53}, containing various T to U substitutions (Table I). Lanes 1 and 7, GAG_2ATC_2TC control; lane 2, U_1; lane 3, T (reference); lane 4, U_{23}; lane 5, U_4; lane 6, U_{1-4}. (0) 106-bp molecules containing zero copies of the inserts.

for $n = 15$–20, small differences in helical repeat can lead to large differences in global conformation. As a consequence, differences in the maximal degree of retardation would be expected for two molecules having the same local axial curvature, but with slightly different helical repeats. Conversely, two sequences having both different local curvatures and different helical repeats could give rise to identical electrophoretic mobilities.

In order to reduce the dependence of the electrophoretic mobilities on helical repeat within the test sequences, the oligomers of interest have been embedded between 53-bp, electrophoretically normal "reporter arms". The current approach follows that published previously for in vitro constructions (Hagerman, 1988) and is analogous to that employed by Diekmann (1986), although the latter author used asymmetric inserts and unequal reporter arms. An example of the current constructions is displayed in Figure 1. With this approach, a slight degree of electrophoretic retardation is evident for the molecule containing just one GA_4T_4C decamer insert, whereas the mobilities of molecules containing $(GAG_2ATC_2TC)_n$ inserts are all normal, thus arguing that the failure to see gel retardation in the latter instance is not due to altered phasing by the GC residues.

Curvature of the Helix Axis within $(dA)_n$·$(dT)_n$ Tracts Is Modulated by Thymidine 5-Methyl Groups in a Position-Dependent Manner. Using the approach described above, in which a small number of decamers is placed between electrophoretically normal reporter arms, we have examined the consequences of selective removal of the 5-methyl groups from the thymine bases within the GA_4T_4C decamer. The results of a series of such replacements are displayed in Figures 2 and 3. For $n = 5$, the apparent sizes of the species containing various 5-methyl deletions differ by as much as 30%, with mobility differences clearly detectable by $n = 2$. In addition, there is a clear position dependence of the influence of the

Table I: Sequences of the Synthetic Oligodeoxynucleotides Used in This Study[a]

designation	sequence[b]
A-tract	
T	GAAAATTTTC
U_1	GAAAAUTTTC
U_{23}	GAAAATUUTC
U_4	GAAAATTTUC
U_{1-4}	GAAAAUUUUC
I-tract	
C	GIIIICCCCC
$^{me}C_1$	GIIII^{me}CCCC
$^{me}C_{1-4}$	GIIII^{me}C^{me}C^{me}C^{me}CC
$^{me}C_4$	GIIIICCC^{me}CC

[a] Only one strand is shown; however, all oligomers form self-complimentary duplexes. The duplexes therefore contain twice the number of methyl groups represented by the single strands. [b] All sequences are written in a 5′ to 3′ direction; U refers to deoxyuridine, I refers to deoxyinosine, and ^{me}C refers to 5-methyl deoxycytidine.

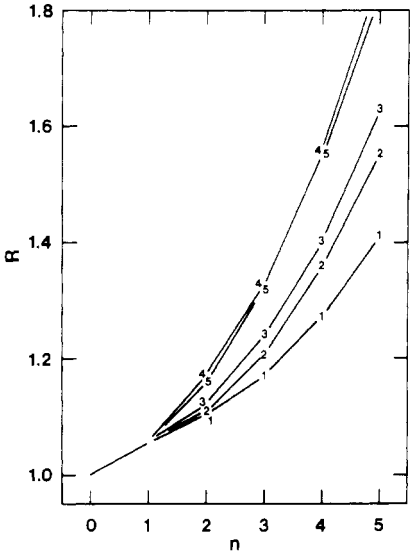

FIGURE 3: Plot of the gel data of Figure 2, plotted as R (apparent number of base pairs/actual number of base pairs) versus the number of inserts, n. (1) U_1; (2) T; (3) U_{23}; (4) U_4; (5) U_{1-4}.

5-methyl groups, with removal of a single 5-methyl group (per strand) at the 5′ end of the T-run resulting in a reduction in apparent curvature and with removal of the methyl groups at the 3′ ends of the T-runs resulting in an increase in apparent curvature. Two additional features of the data should be noted. First, removal of a single methyl group at the 3′ end of the T-run has nearly the same effect as removal of all methyl groups, suggesting that the 3′ T methyl group plays a significant role in modulating curvature. Second, removal of the 5-methyl groups from the two central T residues in the T-tract results in an effect that is intermediate between the 3′ and 5′ effects. In aggregate, these results demonstrate that the thymine 5-methyl groups do play a significant modulatory role in the curvature generated by A-tracts.

Methylation of C Residues in $(dI)_n$·$(dC)_n$ Tracts Mirrors the Effects Observed in $(dA)_n$·$(dT)_n$ Tracts. It has been demonstrated by Diekmann et al. (1987) and by Koo and Crothers (1987) that, whereas replacement of A·T base pairs by G-C base pairs in an A-tract results in the elimination of curvature (presumably due to steric effects of the G 2-NH_2 group in the minor groove), the corresponding replacement by I-C base pairs results in molecules that are still curved, although the mobilities of the I-C-substituted species are always greater (smaller apparent curvature) than those of the parent A-tract molecules. Despite the reduced degree of electrophoretic abnormality associated with the I·C-containing species, a common model for curvature predicts that methylation of the 5-position of C residues should modulate the apparent curvature in a manner analogous to that accompanying methylation of the 5-position of U residues in A-tract species. We have therefore constructed several I-tract molecules in which selected C residues have been methylated at the 5-position. The results of the mobility studies are displayed in Figure 4. Clearly, methylation of the central four C residues has the same effect as the methylation of the U residues in the A-tracts (U_4 to T_4). Moreover, methylation of the 3′ C residue results in reduced mobility, consistent with the 3′ U to T conversion in the A-tract species. Furthermore, methylation of the 3′ C residue has an effect that is as pronounced as that of the C_{1-4} to $5^{me}C_{1-4}$ conversion and is again in the same direction as the corresponding methylations in the A-tract species. These results argue for a common role for pyrimidine methyl groups in modulating curvature, irrespective of the underlying tract sequence [e.g., $(dA)_n$·$(dT)_n$ vs $(dI)_n$·$(dC)_n$].

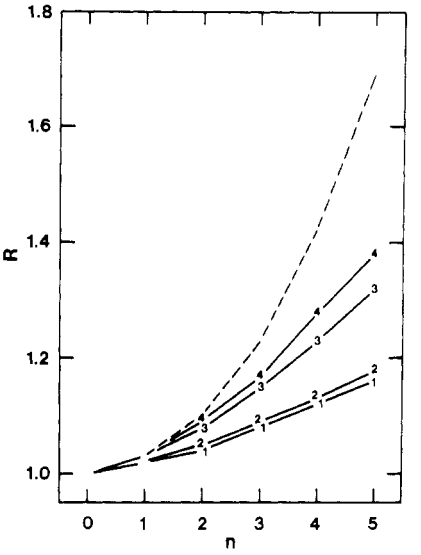

FIGURE 4: Plot of the electrophoretic behavior of I-tract sequences. (1) $^{me}C_{1-4}$; (2) $^{me}C_4$; (3) C; (4) $^{me}C_1$; (--) T reference ladder, run on the same gel.

Discussion

The central conclusion of the current work is that the 5-methyl substituent of pyrimidines is capable of exerting a significant modulatory effect on DNA curvature. Moreover, the 5-methyl group exerts its influence in a position-dependent manner. The current observation that the conversion of the sequence $(GA_4T_4C)_n$ to its unmethylated counterpart $(GA_4U_4C)_n$ is associated with increased curvature argues against the model of Jernigan et al. (1986, 1987), which held that curvature is generated by attractive interactions within tracts of methyl groups. Koo and Crothers (1987) have demonstrated previously that the 5-methyl groups on T residues are not essential for the generation of curvature, in agreement with the current results. However, those authors examined only a single unmethylated motif (AAAAAcgggc/gcccgTTTTT to AAAAAcgggc/gcccgUUUUU) and failed to observe any significant difference in electrophoretic mobility between the two species, leading them to conclude that the thymine methyl group exerts very little influence on the magnitude of DNA curvature. The current observations (Figure 2) thus demonstrate that the conclusion of Koo and Crothers is not a general one. Moreover, our results argue for the further refinement of theoretical/com-

putational models for DNA conformation, since the latter currently do not demonstrate any significant dependence on the presence of the 5-methyl substituent (Srinivasan et al., 1987; Olson et al., 1989).

Although the DNA constructs used in the current study argue strongly for the direct influence of the methyl substituents on local curvature, as opposed to simple changes in torsional phasing (twist), the precise basis of the methyl group effect is unclear. In light of the systematic investigations of the influence of major groove substituents on curvature by Diekmann and co-workers (Diekmann, 1987; Diekmann et al., 1987; Diekmann & McLaughlin, 1988), it is reasonable to suppose that the pyrimidine methyl groups are simply modulating the width of the major groove in a manner that depends on steric interactions between the 5-methyl group and the base to the 5′ side of the T (or $5^{me}C$) residue, although this suggestion remains to be proven.

Finally, although it is clear that pyrimidine methylation plays an important role in the regulation of gene expression (Stein et al., 1983; Goelz et al., 1985; Wolf & Migeon, 1985), the influence of such methylation events is generally attributed to direct interactions between the methyl groups and sequence-specific DNA binding proteins. The current observations suggest that such a model may be too restrictive; in particular, methyl groups may also regulate by altering the conformation of the DNA within the binding domain of the protein. Such an effect has been proposed by Diekmann and McLaughlin (1988) for N6-methylation.

References

Behe, M., & Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1619–1623.

Challberg, S. S., & Englund, P. T. (1980) *J. Mol. Biol. 138*, 447–472.

Cooper, J. P., & Hagerman, P. J. (1987) *J. Mol. Biol. 198*, 711–719.

Diekmann, S. (1986) *FEBS Lett. 195*, 53–56.

Diekmann, S. (1987) *EMBO J. 6*, 4213–4217.

Diekmann, S., von Kitzing, E., McLaughlin, L., Ott, J., & Eckstein, F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8257–8261.

Diekmann, S., & McLaughlin, L. W. (1988) *J. Mol. Biol. 202*, 823–834.

Goelz, S. E., Vogelstein, B., Hamilton, S. R., & Feinberg, A. P. (1985) *Science 228*, 187–190.

Hagerman, P. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4632–4636.

Hagerman, P. J. (1985) *Biochemistry 24*, 7033–7037.

Hagerman, P. J. (1986) *Nature 321*, 449–450.

Hagerman, P. J. (1988) *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) pp 224–236, Springer-Verlag, New York.

Hagerman, P. J. (1990) *Annu. Rev. Biochem.* (in press).

Jernigan, R. L., Sarai, A., Ting, K.-L., & Nussinov, R. (1986) *J. Biomol. Struct. Dyn. 4*, 41–48.

Jernigan, R. L., Sarai, A., Shapiro, B., & Nussinov, R. (1987) *J. Biomol. Struct. Dyn. 4*, 561–567.

Koo, H.-S., & Crothers, D. M. (1987) *Biochemistry 26*, 3745–3748.

Koo, H.-S., Wu, H.-M., & Crothers, D. M. (1986) *Nature 320*, 501–506.

Marini, J. C., Weisberg, R., & Landy, A. (1977) *Virology 83*, 254–270.

Olson, W. K., Srinivasan, A. R., Marky, N. L., Maroun, R. C., Babcock, M. S., et al. (1989) *Theoretical Chemistry and Molecular Biophysics: A Comprehensive Survey* (Beveridge, D. L., & Lavery, R., Eds.) Adenine, Guilderland, NY (in press).

Simpson, L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1585–1588.

Srinivasan, A. R., Torres, R., Clark, W., & Olson, W. K. (1987) *J. Biomol. Struct. Dyn. 5*, 459–496.

Stein, R., Sciaky-Gallili, N., Razin, A., & Cedar, H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2422–2426.

Stellwagen, N. C. (1983) *Biochemistry 22*, 6186–6193.

Taylor, W. H., & Hagerman, P. J. (1990) *J. Mol. Biol.* (in press).

Wolf, S. F., & Migeon, B. R. (1985) *Nature 314*, 467–469.

Wu. H.-M., & Crothers, D. M. (1984) *Nature 308*, 509–513.

Wu, H.-Y., & Behe, M. J. (1985) *Biochemistry 24*, 5499–5502.