MOLECULAR AND CELLULAR BIOLOGY, Oct. 1990, p. 5378–5387
0270-7306/90/105378-10$02.00/0

Vol. 10, No. 10

Z-DNA Formation in the Rat Growth Hormone Gene Promoter Region

MICHAEL J. THOMAS,[1]† THOMAS M. FREELAND,[1,2] AND JEANNINE S. STROBL[1,2]*

Department of Pharmacology and Toxicology, West Virginia University Health Sciences Center,[1] and Program in Genetics and Developmental Biology, Division of Plant and Soil Sciences,[2] West Virginia University, Morgantown, West Virginia 26506

Received 21 March 1990/Accepted 20 July 1990

The complete DNA sequence of the 1.7 kilobase pairs (kbp) 5′ of the rat growth hormone gene (rGH) has been determined and analyzed for Z-DNA-forming potential. Regions of alternating purine-pyrimidine (APP) sequences located between −1047 and −986 [$(GT)_{31}$], between −445 and −433 bp, and between −426 and −403 bp relative to the rGH RNA transcription initiation site were identified and shown to form Z-DNA in negatively supercoiled plasmids by two-dimensional gel electrophoresis. Free-energy calculations indicated that Z-DNA forms most readily in the proximal Z-DNA regions. Diethyl pyrocarbonate footprinting of physiologically supercoiled plasmid DNA confirmed the presence of Z-DNA from −444 to −404 bp spanning the two most proximal APP sequences and a short non-APP sequence in between. DNA sequence analysis also predicted a region of DNA curvature near this proximal Z-DNA region. Formation of Z-DNA in the distal Z-DNA region consisting of a $(GT)_{31}$ repeat was constrained at physiological plasmid superhelical densities. This may be related to the presence of DNA sequences (−1584 to −1559) 512 bp upstream of $(GT)_{31}$ that undergo cruciform formation and thereby utilize the available free energy. Removal of 580 bp containing the cruciform region resulted in Z-DNA formation within $(GT)_{31}$, thus demonstrating that deletion mutations can exert topological changes at a distance within the rGH 5′-flanking region. Methylation of two specific cytosines in the rGH 5′-flanking DNA that have been associated with inhibition of rGH promoter activity had no effect on Z-DNA formation. No evidence for DNA secondary structure formation within the rGH second exon-intron or 3′-flanking region was observed. We conclude that the rGH 5′-flanking region undergoes secondary-structure formation at physiological superhelical densities, thus providing a potential mechanism(s) for modulating rGH activity.

DNA is capable of adopting several secondary structures, including cruciform (stem-loop) and slipped-loop structures, bent and kinked DNA helices, and supercoiled closed circular loops (49). Formation of all but the supercoiled structure is strongly influenced by the primary DNA sequence. The DNA helix is also able to assume several right-handed helical conformations and one left-handed helical form. B-DNA, a right-handed helix, is thought to predominate in vivo. Z-DNA, a left-handed helix, is a high-energy state of the DNA helix and preferentially forms in regions of alternating purine-pyrimidine (APP) sequence (41, 47). Formation of Z-DNA has been studied by using a variety of spectroscopic (4), immunologic (28), chemical (48), electrophoretic (36, 39), and enzymatic (51) techniques. Z-DNA is stabilized by a variety of physicochemical conditions, including negative supercoiling and DNA-protein interactions which occur in vivo (21). Thus, the DNA helix can exist in various conformations, depending upon the degree of supercoiling (9).

The superhelical density (number of supercoils per helical turn of DNA) necessary for Z-DNA stabilization varies with the length and base pair composition of the APP sequence (8, 33). The tendency of an APP sequence to form Z-DNA is inversely related to the number of supercoils required in the plasmid for the B-Z transition to take place. The resolution of topoisomers of plasmid DNA representing a distribution of plasmids of varying superhelical density by two-dimensional (2-D) gel electrophoresis has been used to obtain quantitative information on the free energy of the B-Z transition ($\Delta G_{B\text{-}Z}$) (39). Formation of cruciform structures is also stabilized by increasing the plasmid negative superhelical density and gives rise to transitions in 2-D gel patterns that are indistinguishable from those of Z-DNA. Cruciform structures, however, have a primary DNA sequence requirement for an inverted repeat (49). Z-DNA and cruciform structures can also be distinguished by immunologic and chemical footprinting techniques.

The presence of widely dispersed APP sequences throughout procaryotic (20, 44) and eucaryotic (7, 15) genomic DNAs is consistent with a biological role for Z-DNA. Evidence for Z-DNA formation under negatively supercoiled conditions in vivo has been presented (3, 21, 50), and Z-DNA is likely to promote genetic recombination in vivo (5, 24, 26). The identification of DNA-binding proteins that preferentially bind to the Z conformation in *Drosophila* (37), wheat germ (29), *Escherichia coli* (30), and HeLa cells (31) has led to the speculation that Z-DNA–protein interactions may play some biological role. However, interpretation of these findings has been tempered by the observation that yolk proteins from several species also bind tightly to Z-DNA (27). No role for Z-DNA in gene expression has been established, although APP sequences in plasmid DNA have been associated with both gene activation (16) and repression (14) in mammalian cell transient expression assays. In addition, both RNA polymerase (6) and DNA polymerase I (40) activities are inhibited on Z-DNA templates in vitro,

* Corresponding author.
† Present address: Barnes Hospital, Washington University Medical Center, St. Louis, MO 63110.

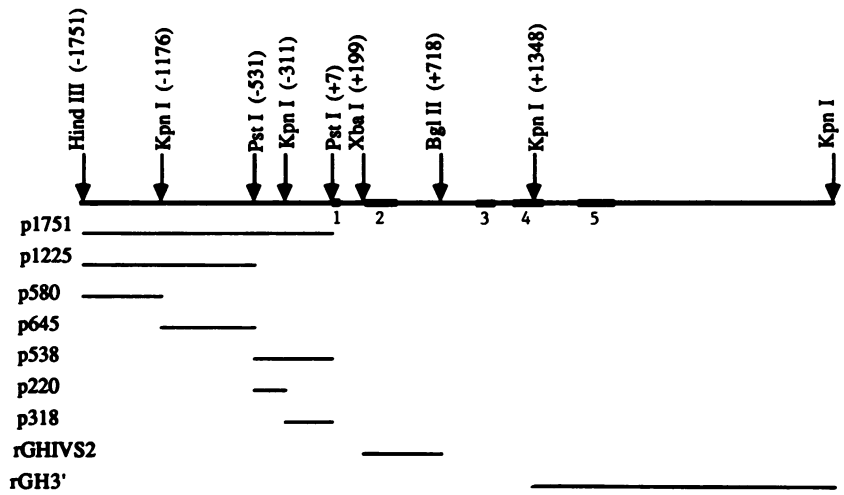

FIG. 1. Subclones of the rGH promoter. The structure of rGH is shown at the top, and the corresponding subcloned regions are labeled with plasmid names on the left. pUC8 or pUC18 plasmid vector was used. Restriction enzyme sites relative to the rGH transcription initiation site are labeled at the top.

implying that DNA transcription and replication may be affected by Z-DNA formation.

MATERIALS AND METHODS

Plasmids. Several regions of the rat growth hormone gene (rGH) 5'-flanking region were subcloned into pUC plasmids that do not contain significant stretches of APP sequence capable of Z-DNA formation (Fig. 1). The 1.7-kilobase-pair (kbp) rGH promoter and 5'-flanking region subcloned into the pUC8 polylinker is designated p1751. Additional regions of p1751 were subcloned into pUC18 as described below. Plasmid p538 contains the proximal 538 bp of the rGH 5'-flanking region extending from −531 to +7 relative to the transcription start site. Two subclones of p538, p220 (−531 to −311) and p318 (−311 to +7), represent the distal and proximal regions, respectively, of p538. Plasmid p1225 contains 1,225 bp of the 5'-flanking region immediately upstream from the p538 region. Two subclones of p1225, p580 and p645, contain the distal 580 bp and the proximal 645 bp, respectively, of the p1225 region. Two other potential regulatory regions of rGH, the second intron and the 3'-flanking region, were also subcloned into pUC18. Plasmid rGHIVS2 contains a 519-bp region of the second exon and intron. Plasmid rGH3' contains a 2.0-kbp region of the fifth exon and 3'-flanking region of the rGH gene.

Plasmid pUC18 was used as a negative control in all experiments. Plasmid pFP332, a pUC8 derivative containing a 32-bp alternating CG sequence inserted into the *Bam*HI site, was used as a positive control. Supercoiled plasmids were purified from *E. coli* HB101 by using the Triton lysis procedure and two or three bandings in ethidium bromide-cesium chloride density gradients.

DNA sequence analysis. Plasmid sequencing of p1751 was performed by using dideoxynucleotide triphosphates and T7 DNA polymerase (Sequenase; U.S. Biochemical Corp., Cleveland, Ohio) (17). In some cases, dITP was used in place of dGTP to eliminate gel compressions and verify uncertain sequences. Sequence information was analyzed by using DNAstar software and digitizer (DNAstar Inc., Madison, Wis.).

Two-dimensional gel electrophoresis. Resolution of negatively supercoiled plasmid topoisomers containing Z-DNA was achieved by using 2-D gel electrophoresis (39). Plasmid topoisomers were generated by using a modified protocol based on previously published methods (25, 42), using calf thymus topoisomerase type I (Bethesda Research Laboratories, Inc., Gaithersburg, Md.). Approximately 5 μg of pooled plasmid topoisomers was electrophoresed in a 0.7% agarose gel measuring 38 by 32 by 0.6 cm in Tris-borate-EDTA (TBE) buffer (89 mM Tris hydrochloride, 89 mM borate, 2 mM EDTA [pH 8]) at 100 V for 30 h. The gel was then soaked in 10 μM chloroquine for 10 h and electrophoresed in the second dimension at 100 V for 24 h. The plasmid DNA was stained by soaking in 10 μM ethidium bromide, photographed, and analyzed for the ability to support Z-DNA formation as detailed below (39, 48).

Quantitative analysis of the B-Z transition in rGH. The linking difference ($a - a_0$) is a value which defines the number of superhelical turns in covalently closed circular DNA (48). By using the band-counting method of Keller (25), linking numbers were assigned to each topoisomer. Transition from a B-DNA helix to a Z-DNA helix results in the loss of two helical turns of DNA for each twist of the DNA helix (48). This conformation change results in the removal of two supercoils for each helical twist of DNA which undergoes a B-Z transition. On the basis of these topological relationships, the number of base pairs undergoing a B-Z transition can be calculated (39).

Formation of Z-DNA occurs above a specific superhelical density, which is determined by the base composition and length of the APP sequence. For a Boltzmann distribution of plasmid topoisomers, the free energy of supercoiling required for stabilization of the Z conformation of the helix ($\Delta G_{B\text{-}Z}$) can be calculated by the equation $\Delta G_{B\text{-}Z} = 1{,}100\ RT(t_b2 - t_z2)/N$, in which t is the linking number based on 2-D gel analysis, N is the number of base pairs in the plasmid, and 1,100 RT is a constant, K, which is insensitive to the size of the plasmid for $N > 2{,}000$ bp and is proportional to the gas constant, R (1.987 cal/mol-°C [1 cal = 4.184 J]) and the absolute temperature, T (300 K under experimental conditions) (48).

Site specific methylation of rGH plasmids. Plasmids p1751, p538, and p318, containing the *Tha*I (CGCG) methylation sites located in the rGH 5'-flanking region, were methylated

```
         A AGCTTAGTTT CTAGTAGGAA TGAATTCTGG TTTGGTCTCT -1711
GTCACCAGAC ATCCTAAAAT GGTCCAGACT AGAGAAAAGA TCTCCAACCC -1661
CTCTGATCTC ACACTTATCC TCAGCTCGCC TCCTGGGCCT AGAGGTAGCA -1611
ATATTAACAT TCTATAAAGA CATATCCCCA GGGGTCCTCA AGGTCCCCTG -1561
GGCAACTCCT TTCTGGGGTC ACTTCAGTTC TGATATGGCA AGAATAGGTG -1511
TCAATTTCTC CCCCTCTCAG TGGGACCTTC TTTCTTAGCT CCTGTATCTT -1461
CAGTTAACTC CTGGAGCCTT CCACAGACCT AGCCCCTAAG GCCTCGCGGC -1411
CTCCTCAGAG TGCTGGATAA AACAGGGGAA GCTGCTGGGA AATGCTTTGG -1361
GGATGTTTAG AGGTTATATC TTTTGGTCCT CCATGGGGAG AGGTATAGAG -1311
GCCCCAGAAT TTACCATCGT AGGCCCCAGG GACATGATTT TTCAAAGTTG -1261
CTTGCTGACA TCACAACCAA GTACAGCTGC ATGGCAGTTC CAATGAGCTG -1211
GAATGGGTCC TGTGCTCCAC AGACCTAAGG GGTACCTTAG GGCTCCTGGG -1161
AGAGCATTGT ACGTGGCCAG TACCCTCTAG TGGTCAGTGT TAGCACTACA -1111
CCCAGCTGCC CTCCATCAGT TTATGCTGCT ATGGGAGGGA ACAAGTCTTC -1061
CTTTTCCCAA TACGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG -1011
TGTGTGTGTG TGTGTGTGTG TGTGTTCATT CATTCTCTCT CTCTGTCTCT  -961
CTTTCTCTCT CCCTCTGTCT CTCTCTTTCT CTTTCGGTTA TTTTTATTTT  -911
TATTGGTGTG TATGTGTGTG AGTCACATGT ATGACCAGAA GAGAGCACCA  -861
GATCCCCTAG AACTGTGAGC CACTTGACAT GGTGACATGG GTGCTGGGAA  -811
CCAAACTCAG GTCCTCTGAA AGAACAACAA ACTTATTGCA CAAGCCATCT  -761
CTAAATCTGG CACATGTTTG TTTGTTTTGT TTTCTTTTGT TTTCTTTTGT  -711
TTTGTTTTTT GAGGTTTAGA TTCAAGAGCA GGAGGAGGAG GAGGAGGAGC  -661
AGGAGCAGGA GCAGGAGCAG GAGCAGGAGC AGGAGCAGGA GCAGGAGCAG  -611
GAGCAGGAGC AGGAGCAGGA GCAGGAGCAG GAGGACAAAA TCAAGCCAAC  -561
AAAATGGCCA AGGAGGTGTA AGCACCTGCA GCCAAGCCTG AGTATCTGAG  -511
TTCAATCCCC AGGACCCACA AGATGGAACG ATGGAACCAA ACCCCAAAAG  -461
TTGCCTTCTG GTCCCTACAT GTACACACTT GAGGACACGC ACGAATGCAC  -411
GCACACACTC AACAAATTAA GAGGAATAAG ACAATCATGG GGAAAATACC  -361
TCCTTGGAGA GGCTCTGTTG CCCCTCGTCC CAGTGAACAA ACGATGGTAC  -311
CCTGCCAGAG TATCCTACCC CTGGATTCAA AAATACTCTC AAAAGGACAC  -261
ATTGGGTGGT CTCTGTAGCT GAGATCTTGC GTGACCATTG CCCATAAACC  -211
TGGGCAAAGG CGGCGGTGGA AAGGTAAGAT CAGGGACGTG ACCGCAGGAG  -161
AGCAGTGGGG ACGCGATGTG TGGGAGGAGC TTCTAAATTA TCCATCAGCA  -111
CAAGCTGTCA GTGGCTCCAG CCATGAATAA ATGTATAGGG GAAAAGGCAG   -61
GAGCCTTGGG GTCGAGGAAA ACAGGTAGGG TATAAAAAGG GCATGCAAGG   -11
GACCAAGTCC AGCACCCTCG A                                  +11
```

FIG. 2. Sequence of the rGH promoter from positions −1751 to +26 relative to the transcription start site. Regions with Z-DNA-forming potential are in boldface; the inverted repeat is underlined and is marked by arrows; the $(GTTT)_9$ and $(AGGAGC)_{14}$ repeats are underlined; a consensus cyclic AMP-responsive element is located at −1256 to −1247.

in vitro by using *Bsu*E DNA methyltransferase (12, 43). Complete methylation was achieved as measured by complete protection of the DNA from digestion by the cognate restriction endonuclease *Tha*I.

DEPC footprinting. The details of the procedure have been published previously (48). Supercoiled plasmid p220 was reacted with diethyl pyrocarbonate (DEPC), and then a 151-bp *Hin*dIII-*Kpn*I fragment uniquely labeled at the *Hin*dIII 3′ end with DNA polymerase I and [α-^{32}P]dATP was cleaved at DEPC-modified sites by treatment with 1 M piperidine at 90°C for 5 min. Reaction products were separated by electrophoresis in 8% polyacrylamide–8 M urea sequencing gels and exposed to X-ray film for approximately 16 h.

Nucleotide sequence accession number. The sequence reported here has been assigned GenBank accession number X12967.

RESULTS

DNA sequence of the rGH 5′-flanking region. The DNA sequence upstream of the rGH promoter from positions −1751 to +11 is shown in Fig. 2. The sequence from −1751 to −403 reported here is new; the sequence from −402 to +11 was previously determined (1). Two long stretches of APP sequence were identified between −1047 and −986 and between −426 and −403. Three shorter stretches of APP sequences were found in the 5′-flanking region, beginning at −906, −445, and −150, with one Z-Z junction occurring in the most proximal sequence. An inverted repeat sequence occurs in the far 5′-flanking sequence between −1584 and −1559. Other sequences of interest noted in Fig. 2 include an imperfect $(GTTT)_9$ repeat (a consensus topoisomerase II-binding site) beginning at −745, the $(AGGAGC)_{14}$ hexamer repeat beginning at −666, and a consensus cyclic AMP-responsive element (CTGACATCAC) at −1256.

Two-dimensional gel electrophoresis. Subclones of the rGH promoter were isolated in pUC8 or pUC18 (Fig. 1) and analyzed for Z-DNA formation by 2-D gel electrophoresis. Figure 3 compares the 2-D gels of plasmid topoisomers of pUC18 (Fig. 3A) and two regions of rGH, rGHIVS2 (Fig. 3B) and rGH3′ (Fig. 3C), subcloned into pUC18. pUC18 does not contain long stretches of APP DNA sequence and exhibited no discernible structural transition in its 2-D gel topoisomer electrophoresis pattern, indicating that supercoiled pUC18 does not form Z-DNA. Plasmids rGHIVS2 and rGH3′ also failed to exhibit any structural transitions by 2-D gel analysis, and we conclude that neither of these regions of the rGH gene is capable of Z-DNA formation. The absence of Z-DNA formation in the 41-bp APP sequence in the second intron of rGH (rGHIVS2) is consistent with the DNA sequence analysis, which revealed the presence of five Z-Z junctions and a high content of adenine and thymine residues, both of which are unfavorable for Z-DNA formation (10, 19). Nucleotide sequence data for the 3′-flanking region of the rGH gene (rGH3′) are unavailable.

Figures 4 and 5 demonstrate that structural transitions are

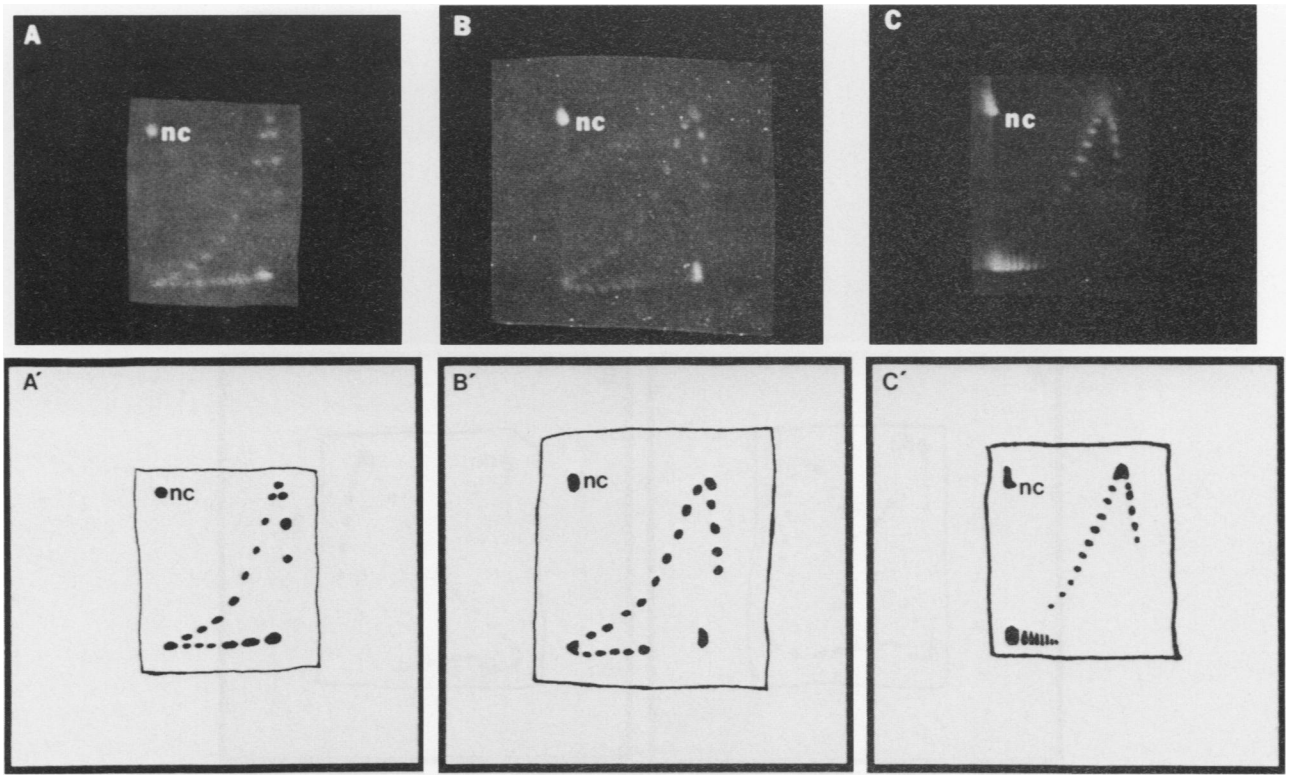

FIG. 3. Two-dimensional analysis of Z-DNA formation in plasmids containing rGH. All gels were run under conditions described in Materials and Methods. Nicked circular (nc) DNA is labeled in each panel. The first dimension of electrophoresis is from top to bottom; the second dimension is from left to right. (A and A') pUC18; (B and B') rGHIVS2; (C and C') rGH3'. Gels are representative of two to three experiments.

exhibited by sequences in the rGH 5'-flanking region. In p1751, DNA structural transitions occurred in two distinct steps (Fig. 4A). This has been interpreted previously as an indication that B-Z transitions have occurred in two distinct DNA regions (9). The 2-D gel analysis of p1225, a subclone of p1751, is shown in Fig. 4B. There are in fact two structural transitions in p1225, one due to Z-DNA formation and the other due to cruciform formation. However, only one such transition is detectable when these sequences are topologically linked. The detailed analysis of these structural transitions is presented in Fig. 5.

Plasmids p580 and p645 are subclones representing the distal and proximal portions, respectively, of p1225 (Fig. 4B), and each displayed a single structural transition in the 2-D gel pattern (Fig. 5A and B). These results indicate that two distinct regions of p1225 are capable of undergoing secondary-structure formation and are supported by the DNA sequence analysis of p1225. Plasmid p580 includes the 25-bp sequence (−1584 to −1559) capable of undergoing cruciform formation (Fig. 6), and p645 includes a 62-bp alternating GT dinucleotide [$(GT)_{31}$] sequence (−1047 to −976) that readily forms Z-DNA in vitro.

Plasmid p645, containing the $(GT)_{31}$ sequence, consistently exhibited two to three identical patterns in the 2-D gel which were offset. These probably reflect the presence of both monomer and dimer forms of p645, as Z-DNA formation has been shown to stimulate plasmid dimerization in *E. coli*, but do not affect the interpretation of the experimental results.

Plasmids p538 and p220 (Fig. 5C and D) each exhibited single structural transitions characteristic of B-Z transitions in their respective 2-D gel patterns. Plasmid p318 did not exhibit any structural transition in its 2-D gel pattern (Fig. 5E). We conclude that since p220 and p318 represent the distal and proximal sequences of p538, supercoil-dependent conformational changes consistent with Z-DNA formation occur in the rGH 5'-flanking region between 538 and 311 bp upstream of the transcription initiation site. The DNA sequence of this region of rGH contains a 24-bp APP stretch between bp −426 and −403 which is capable of adopting the Z conformation. However, consideration of only these 24 bp may underestimate the potential for Z-DNA formation in this region, since a second, 13-bp APP sequence lies only 6 bp further upstream.

The quantitative analysis of the 2-D gel electrophoresis patterns in Fig. 4 and 5 (summarized in Table 1) provides important information on the free-energy requirements of the structural transitions and the number of base pairs involved. The $\Delta G_{B\text{-}Z}$ is inversely related to the ease with which a sequence of DNA will undergo a B-Z transition on the basis of energetic considerations alone. The rGH subclone undergoing a structural transition most readily was p1225 (13.5 kcal/mol). However, most of the rGH subclones in plasmids exhibited a $\Delta G_{B\text{-}Z}$ in the range of 17 to 20 kcal/mol, which compares favorably with a plasmid, pFP332 ($\Delta G_{B\text{-}Z}$ of 19.9), bearing a highly stable Z-DNA sequence [$(CG)_{16}$]. The second structural transition of p1751 had a very high $\Delta G_{B\text{-}Z}$ of 87.9 kcal/mol, indicating that this particular transition is energetically very costly.

Approximately 24 bp are involved in the structural transitions observed with p538 and p220. This result is in agreement with the nucleotide sequence of this region, indicating

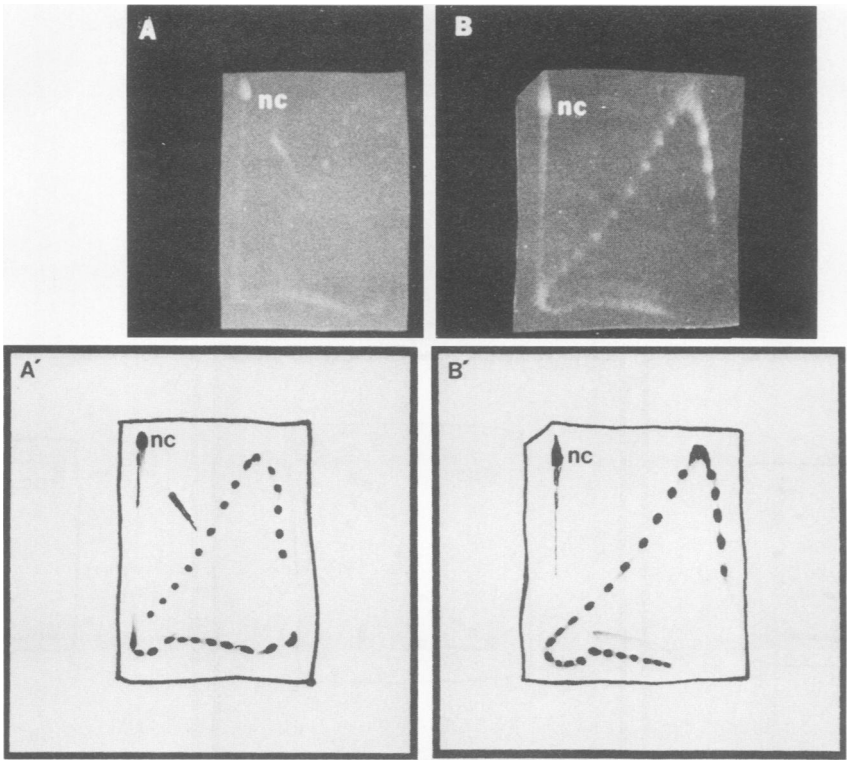

FIG. 4. Two-dimensional gel analysis of Z-DNA formation in plasmids containing the rGH 5′-flanking region. All gels were run under conditions described in Materials and Methods. Nicked circular (nc) DNA is labeled in each panel. The first dimension is from top to bottom; the second dimension is from left to right. (A and A′) p1751; (B and B′) p1225. Gels are representative of five (p1751) or three (p1225) experiments.

the presence of a 24-bp APP sequence in p220. The formation of Z-DNA is likely to occur in both p538 and p220. Approximately 22 bp are involved in the structural transitions observed with p1225 and p580. Both of these rGH subclones contain a 25-bp sequence with dyad symmetry, and cruciform formation within this 25-bp region is likely to account for the observed structural transition. Although p1225 contains the sequence $(GT)_{31}$, there is no evidence for the involvement of this large number of base pairs in the structural transition observed. In contrast, the second structural transition in p1751 involves 78 bp. The only contiguous stretch of nucleotides long enough to account for this structural transition is the 66-bp Z-DNA-forming sequence present in p1225 and p645, $ATAC(GT)_{31}$. Thus, formation of Z-DNA within the distal $(GT)_{31}$ sequence may occur in p1751.

The $\Delta G_{B\text{-}Z}$ was also estimated from the nucleotide sequence by summing the theoretical thermodynamic free-energy values of dinucleotide interactions in the Z conformation and taking into account all Z-Z and B-Z junctions (8, 19). The nucleotide sequence data for the rGH upstream region suggested that four APP DNA sequences were capable of supporting Z-DNA formation under negatively supercoiled conditions (Table 2). These theoretical calculations indicate that the distal 62-bp sequence, $(GT)_{31}$ beginning at −1047, and a proximal, 24-bp APP sequence beginning at −426, exhibiting the lowest $\Delta G_{B\text{-}Z}$ are energetically the most likely sites for Z-DNA formation in the rGH 5′-flanking region. The absence of DNA sequence dyad symmetry within either of these sequences precludes cruciform structure formation (Table 2).

DEPC footprinting. Purines in Z-DNA adopt the *-syn* conformation and are hyperreactive to N-7 alkylation by DEPC and subsequent strand cleavage by piperidine (48). This chemical footprinting technique was used to identify the base pairs involved in Z-DNA formation in the rGH sequences within p220 (Fig. 7). At the highest DEPC concentration used (lane 6), the DEPC footprint extended from −400 to −458, involving a stretch of approximately 58 bp. The extension of the footprint past the 5′ and 3′ ends of the APP sequence may reflect small differences in the degree of plasmid superhelical density or, more likely, some migration of the B-Z junctions along the DNA helix (23). All of the purines within the most proximal 24-bp APP region (−426 to −403) were hyperreactive with DEPC. In addition, the footprint extended upstream through one B-Z junction to encompass the more distal 13-bp APP sequence (−433 to −445). These results suggest that Z-DNA forms in the longer, more proximal APP sequence and extends upstream through a region of non-APP sequence at physiological superhelical densities.

Effect of DNA methylation on the B-Z transition. There are two *Tha*I restriction sites in the rGH 5′-flanking region (−144 and −1413) which are substrates for the DNA methyltransferase M · *Bsu*E (12). Plasmids p1751, p538, and p318 were methylated with M · *Bsu*E, and plasmid topoisomers were generated for 2-D gel analysis. Unmethylated plasmid topoisomers of each plasmid were electrophoresed in the same gel as a control. No discernible effects of DNA methylation on the 2-D gel electrophoresis pattern of p1751, p538, or p318 were observed (data not shown).

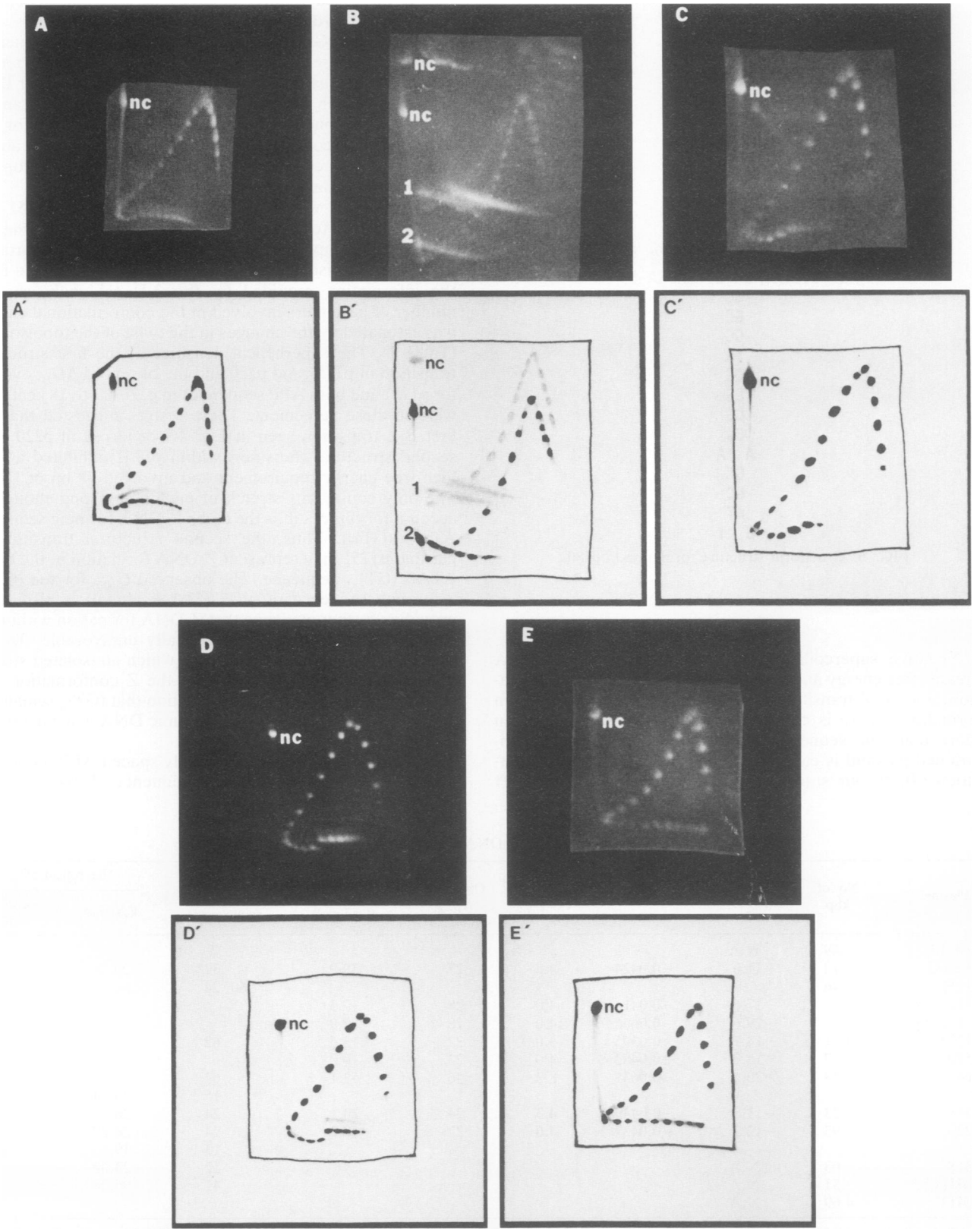

FIG. 5. Two-dimensional gel analysis of Z-DNA formation in plasmids containing the rGH 5′-flanking region. All gels were run under conditions described in Materials and Methods. Nicked circular (nc) DNA is labeled in each panel. The first dimension is from top to bottom; the second dimension is from left to right. (A and A′) p580; (B and B′) p645; (C and C′) p538; (D and D′) p220; (E and E′) p318. Gels are representative of two or three experiments.

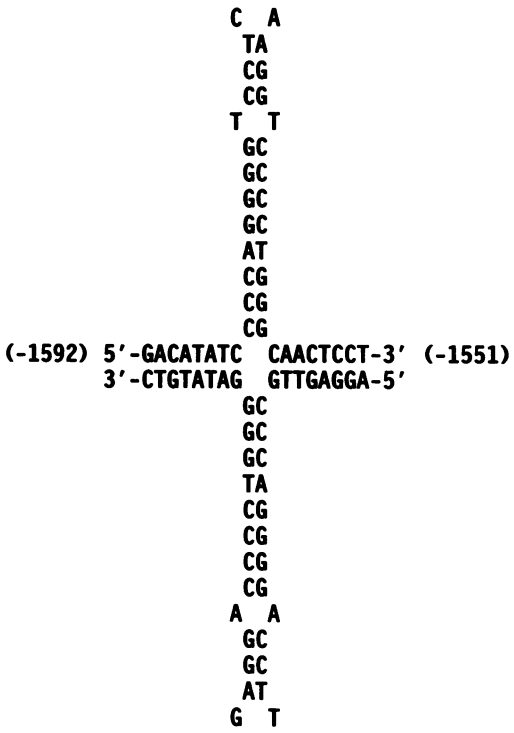

FIG. 6. Cruciform structure formation in p580.

DISCUSSION

Negative supercoiling in closed circular plasmid DNA creates free energy and supports secondary-structure formation, e.g., B-Z transitions in the DNA helix and cruciform formation. There is competition for this free energy when more than one sequence of DNA in a topologically constrained plasmid is capable of undergoing a structural transition (9). In our studies, the DNA sequence of the rGH region was used to predict areas of potential secondary structure. Two-dimensional gel analysis of various subclones of the rGH promoter and far upstream sequences indicated that secondary-structure formation occurred as predicted but was influenced by the presence of other DNA sequences within the same plasmid which were also capable of adopting secondary structures. Thus, competition for free energy of supercoiling provides an explanation for the ability of different subclones of the rGH promoter to adopt or suppress secondary-structure formation.

To determine which of the APP sequences in p1751 was most likely to form Z-DNA at physiological superhelical densities, the superhelical densities of the structural transitions and the associated free energies were calculated from the information provided by the 2-D gel analyses. The number of base pairs involved in the conformational change was estimated by the changes in the twist of the topoisomers (Table 1). The superhelical densities of the first structural transition in p1751 and p220 and the observed $\Delta G_{B\text{-}Z}$ values for p1751 and both APP sequences in p220 (17 to 18 kcal/mol) were in close agreement. These results suggested that the first B-Z transition seen in Fig. 4A occurred in p220. The second structural transition within p1751 exhibited a very high free energy requirement and involved 78 bp of DNA. The only contiguous stretch of nucleotides long enough to account for this result is the 66-bp Z-DNA-forming sequence ATAC $(GT)_{31}$. Thus, the second structural transition in plasmid p1751 may represent Z-DNA formation in the distal ATAC $(GT)_{31}$ sequence. The observed $G_{B\text{-}Z}$ for the $(GT)_{31}$ Z-forming sequence of p645 (53.1 kcal/mol) is also high, providing further evidence that Z-DNA formation within the rGH $(GT)_{31}$ sequence is energetically unfavorable. This is surprising in light of the ease with which an isolated stretch of $(GT)_{31}$ will theoretically adopt the Z conformation (19) and cautions against the generalization that $(GT)_n$ sequences within negatively supercoiled genomic DNA will exist in the Z conformation.

Plasmid p220 contains two closely spaced APP sequences separated by 6 bp of non-APP sequence (Table 2) in the

TABLE 1. Free energy of Z-DNA formation in the rGH promoter

Plasmid	No. of kbp	$a-a_0$	σ	T_w	Observed N_{B-Z}	Observed ΔG_{B-Z} (kcal/mol)	Length of APP sequence(s)	Theoretical ΔG_{B-Z}	
								kcal/mol	kcal/mol per bp
pUC18	2.69	N.A.							
pFP332	2.72	−11.0	−0.0424	4.8	27	19.9	32	20.56	
p1751	4.46						60, 24	80.21	
1st flip		−17.6	−0.0414	4.0	28	17.3			
2nd flip		−29.6	−0.0696	14.0	78	87.9			
p1225	3.91	−14.7	−0.0395	3.0	22	13.5	62		
p580	3.27	−15.0	−0.0482	4.0	22	19.0			
p645	3.33	−20.4	−0.0631	8.3	46	53.1	62	50.86	0.82
							15	20.44	1.46
p538	3.23	−13.8	−0.0449	4.3	24	20.1	24	26.27	
p220	2.93	−12.2	−0.0437	4.0	22	18.2	24	26.27	1.09
							13	19.1	1.59
p318	3.03	N.A.					12	23.50	
rGHIVS2	3.31	N.A.					41	95.28	
rGH3'	4.69	N.A.							

[a] Two-dimensional gel electrophoresis was performed, and results were confirmed two to three times for each plasmid as described in Materials and Methods. For the topoisomer at which the structural transition of supercoiled plasmid topoisomer occurs, $a-a_0$ is the linking difference (N.A. indicates that no structural transition occurred during 2-D gel electrophoresis), σ is the superhelical density, and T_w is the change in the twist. The observed $N_{B\text{-}Z}$ was determined by using the equation given in Materials and Methods and assumes helical twists of 10.5 and 12.0 per turn for B-DNA and Z-DNA helices, respectively. The length of APP sequences was determined by sequence analysis. The theoretical ΔG_{B-Z} was calculated by using the thermodynamic parameters defined by Ho et al. (19) and the highest even number of base pairs in the APP sequence.

TABLE 2. APP sequences in the rGH upstream region

Sequence[a]	Plasmid subclone	Length (bp)	No. of Z–Z junctions
−1049 ATAC$(GT)_{31}$TCATT −980	p645	66	
−910 TATTGGTGTGTATGTGTGTGAGTCA −885	p645	15	
−450 GTCCCTACATGTACACACTT −430	p220	13	
−430 GAGGACACGCACGAATGCACGCACACACTC −400	p220	24	
−155 TGGGGACGCGATGTGTGGGA −135	p318	12	1
+827 TGGTGTGTCTGAAGACAGCTACAGTGTACTTATATATAATA +868	rGHIVS2	41	5

[a] APP sequences are underlined. Numbers adjacent to sequences indicate positions relative to the transcription initiation site. Nucleotides at Z–Z junctions are printed in bold.

pattern 5′–13-bp APP–6-bp non-APP–24-bp APP–3′. Formation of Z-DNA within the entire 24-bp APP sequence and the next 13 APP bp within a physiologically supercoiled plasmid was demonstrated by DEPC footprinting. The DEPC footprint also provided direct evidence for formation of an altered helical structure within the 6 bp of non-APP sequence, TTGAGG, which separates the 24- and 13-bp APP sequences (Fig. 7). Evidence of Z-DNA formation in synthetic non-APP sequences has been presented previously (22, 46); however, the rGH promoter region may be the first native non-APP DNA sequence shown to undergo Z-DNA formation. The pattern of Z-DNA formation in p220 revealed by DEPC hyperreactivity is consistent with theoretical argument that a contiguous region of Z-DNA is energetically more favorable than two adjacent Z-DNA regions because of the high energy requirements for the maintenance of B-Z and Z-Z junctions.

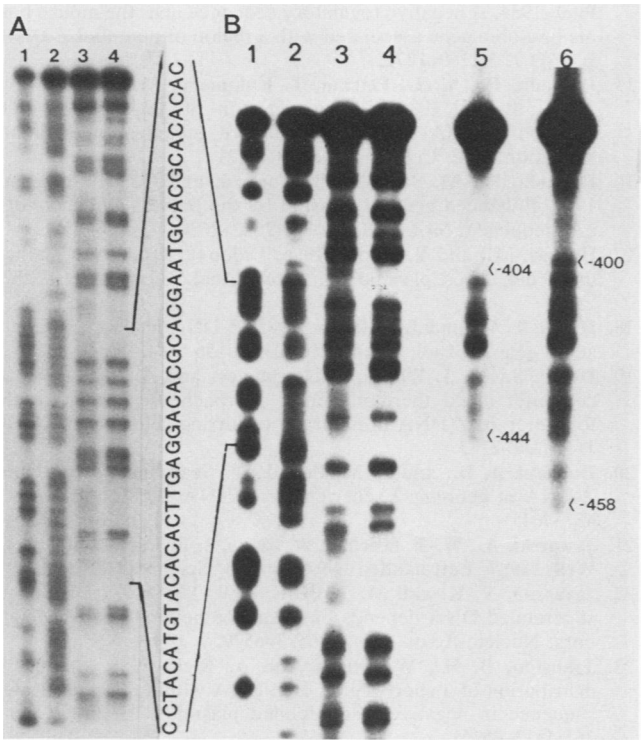

FIG. 7. DEPC footprint analysis of p220. Lanes 1 to 4 contain Maxam-Gilbert sequencing reactions (T+C, T, G+A, and G, respectively) of the 3′-end-labeled 151-bp *Hin*dIII-*Kpn*I fragment. Supercoiled p220 was reacted with DEPC at a DEPC/purine molar ratio of 1,070:1 (lane 5) or 2,140:1 (lane 6) before isolation of the *Hin*dIII-*Kpn*I fragment, end-labeling reaction, and piperidine cleavage. Electrophoresis and autoradiography were performed as described in Materials and Methods.

Structural transitions within p1751 and p1225 are complex and illustrate that interactions occur between distant DNA sequences that are linked topologically. By 2-D gel analysis, plasmid p1225 exhibits a single structural transition involving 22 bp. This structure is likely to reflect cruciform formation in the region from −1584 to −1559. McLean and Wells have shown that cruciform formation is favored over Z-DNA in 10 mM Tris–1 mM EDTA (pH 7.8) DNA storage buffer and remains stable even in the higher-salt environment of the electrophoresis running buffer (34). Formation of this stable cruciform conformation in p1225 utilizes the available free energy of supercoiling and thus prohibits the B-Z transition in the downstream $(GT)_{31}$ sequence. Subcloning p1225 into p580 and p645 effectively separated the cruciform structure and the Z-DNA sequence topologically; two structural transitions were then observed, consistent with cruciform formation in p580 and Z-DNA in p645. These results have provocative implications for the control of gene expression by mechanisms that act at a distance by suppression or induction of particular secondary structures. For example, the presence of a protein that stabilized cruciform formation would act to suppress Z-DNA formation within the proximal APP sequence in rGH; conversely, the presence of a Z-DNA-binding protein might suppress cruciform formation 500 to 600 bp upstream.

Z-DNA formation is not unique to the rGH 5′-flanking DNA. Hayes and Dixon located Z-DNA-forming sequences in the 5′-flanking region of the rat somatostatin gene (18), and the $\Delta G_{B\text{-}Z}$ required for stabilization of the left-handed helix in the rat somatostatin gene is similar to that of rGH. In addition, the upstream region of the rat prolactin gene contains at least six sequences of the type $(TG)_n$ (32). A 170-bp APP sequence derived from the rat prolactin gene 5′-flanking region is capable of inhibiting the expression of the chloramphenicol acetyltransferase gene in GH_3 rat pituitary cells (35).

We have only begun to analyze the potential biological role of Z-DNA formation in the rGH promoter region. Complete methylation of the C-5 position of cytosine in the polymer poly(dGm^{5}dC) · poly(dGm^{5}dC) has been shown to stabilize the helix in the Z conformation under physiological salt conditions (2, 11). In the study presented here, site-specific DNA methylation was carried out at two *Tha*I (CGCG) sequences in the rGH 5′-flanking region to determine whether this limited degree of DNA methylation had any effect on Z-DNA stabilization. The *Tha*I site at −144 is located within a 12-bp APP sequence with a single Z-Z junction, and it was of interest to determine whether replacement of cytosine with 5-methylcytosine could overcome the energetic constraint of the Z-Z function. Methylation of this

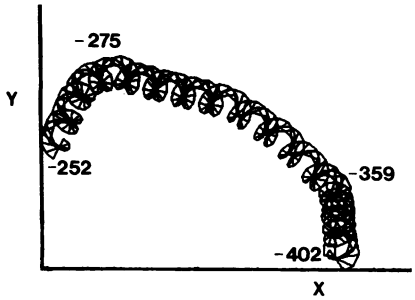

FIG. 8. Potential region of DNA bending in the rGH promoter based on computer program sequence analysis (DNAstar) using the ApA wedge model of Trifonov et al. (44, 45). The region lies between −252 and −402. The 24-bp APP sequence is between −426 and −403.

site has been associated with repression of rGH expression (13, 43). DNA methylation failed to stabilize Z-DNA formation at this site, providing conclusive evidence that methylation-induced Z-DNA formation at this site plays no role in rGH expression. Our data do not eliminate the possibility that DNA-protein interactions could stabilize Z-DNA formation at this site in vivo in either its methylated or unmethylated state, however.

Finally, the rGH DNA sequence from −252 to −434 possesses a moderate degree of bending potential as determined from computer program sequence analysis (DNAstar) using the ApA wedge model of Trifonov et al. (44, 45) (Fig. 8). It is interesting to speculate that the secondary structure of this region, containing a DNA bend and the adjacent proximal Z-DNA sequence, could be involved in the regulation of rGH expression. This hypothesis predicts that the biological activity of the rGH promoter is dependent on its topology and may vary in transient expression cell transfection assays when the superhelical density of plasmid DNA is not controlled. In their natural in vivo state, genes are associated with topological domains that are established by the attachment of chromatin to the nuclear matrix (38). Investigation of the biological role of secondary- and tertiary-structure formation in rGH may require incorporation of the gene into topologically constrained chromosome structures.

ACKNOWLEDGMENTS

We thank Patricia Earl for providing a 5.6-kbp region of genomic rGH subcloned into pBR327 and Fritz Pohl for the pUC8 derivative containing a $(GC)_{16}$ insert at *Bam*HI.

This work was supported by Public Health Service grants DK33631, T32 GM07039, and S07 RR05433-28 from the National Institutes of Health.

LITERATURE CITED

1. **Barta, A., R. I. Richards, J. D. Baxter, and J. Shine.** 1981. Primary structure and evolution of rat growth hormone gene. Proc. Natl. Acad. Sci. USA **78**:4867–4871.
2. **Behe, M., and G. Felsenfeld.** 1981. Effects of methylation on a synthetic polynucleotide: the B-Z transition in poly(dG-m^{5}dC)-poly(dG-m^{5}dC). Proc. Natl. Acad. Sci. USA **78**:1619–1623.
3. **Bhriain, N. A. M., and C. J. Dorman.** 1988. A novel genetic locus determines in vivo B-Z DNA structure transitions in E. coli. Trends Biochem. Sci. **13**:130.
4. **Brahms, S., J. Vergne, J. G. Brahms, E. DiCapua, P. Bucher, and T. Koller.** 1982. Natural DNA sequences conform left-handed helices in low-salt solutions under conditions of topological constraint. J. Mol. Biol. **162**:473–493.
5. **Bullock, P., J. Miller, and M. Kotchan.** 1986. Effects of poly [d(pGpT) d(pApCO] and poly[d(pCpG) d(pCpG)] repeats on homologous recombination in somatic cells. Mol. Cell. Biol. **6**:3948–3953.
6. **Butzow, J. J., Y. A. Shin, and G. L. Eichhorn.** 1984. Effect of template conversion from the B to the Z conformation on RNA polymerase activity. Biochemistry **23**:4837–4843.
7. **Delseny, M., M. LaRoche, and P. Penon.** 1983. Detection of sequences with Z-DNA forming potential in higher plants. Biochem. Biophys. Res. Commun. **116**:113–120.
8. **Ellison, M. J., J. Feigon, R. J. Kelleher III, A. H.-J. Wang, J. F. Habener, and A. Rich.** 1986. An assessment of the Z-DNA forming potential of alternating dA-dT stretches in supercoiled plasmids. Biochemistry **25**:3648–3655.
9. **Ellison, M. J., M. J. Fenton, P. S. Ho, and A. Rich.** 1987. Long-range interactions of multiple DNA structural transitions within a common topological domain. EMBO J. **6**:1513–1522.
10. **Ellison, M. J., R. J. Kelleher III, A. H.-J. Wang, J. F. Habener, and A. Rich.** 1985. Sequence-dependent energetics of the B-Z transition in supercoiled DNA containing nonalternating purine-pyrimidine sequences. Proc. Natl. Acad. Sci. USA **82**:8320–8324.
11. **Feuerstein, B. G., L. J. Marton, M. A. Keniry, D. L. Wade, and R. H. Shafer.** 1985. New DNA polymorphism: evidence for a low salt, left-handed form of poly(dG-m^{5}dC). Nucleic Acids Res. **13**:4133–4141.
12. **Gaido, M. L., C. P. Prostko, and J. S. Strobl.** 1988. Isolation and characterization of BsuE methyltransferase, a CGCG specific DNA methyltransferase from Bacillus subtilis. J. Biol. Chem. **263**:4832–4836.
13. **Gaido, M. L., and J. S. Strobl.** 1989. Inhibition of rat growth hormone promoter activity by site-specific methylation. Biochem. Biophys. Acta **1008**:234–242.
14. **Gilmour, R. S., D. A. Spandidos, J. K. Vass, J. W. Gow, and J. Paul.** 1984. A negative regulatory sequence near the mouse beta major globin gene associated with a region of potential Z-DNA. EMBO J. **3**:1263–1272.
15. **Hamada, H., M. G. Petrino, T. Kakunaga, M. Seidman, and B. D. Stollar.** 1984. Characterization of genomic poly(dT-dG) · poly(dC-dA) sequences: structure, organization, and conformation. Mol. Cell. Biol. **4**:2610–2621.
16. **Hamada, H., M. Seidman, B. Howard, and C. M. Goodman.** 1984. Enhanced gene expression by the poly(dT-dG) poly(dC-dA) sequence. Mol. Cell. Biol. **4**:2622–2630.
17. **Hattori, M., and Y. Sakati.** 1986. Dideoxy sequencing method using denatured plasmid templates. Anal. Biochem. **152**:232–238.
18. **Hayes, T. E., and J. E. Dixon.** 1985. Z-DNA in the rat somatostatin gene. J. Biol. Chem. **260**:8145–8156.
19. **Ho, P. S., M. J. Ellison, G. J. Quigley, and A. Rich.** 1986. A computer aided thermodynamic approach for predicting the formation of Z-DNA in naturally occurring sequences. EMBO J. **5**:2737–2745.
20. **Hoheisel, J. D., and F. M. Poh.** 1987. Searching for potential Z-DNA in genomic Escherichia coli DNA. J. Mol. Biol. **193**: 457–464.
21. **Jaworski, A., W.-T. Hsieh, J. A. Blaho, J. E. Larson, and R. D. Wells.** 1987. Left-handed DNA in vivo. Science **238**:773–777.
22. **Jayasena, V. K. and M. J. Behe.** 1989. The B-Z transition in supercoiled DNA depends on sequence beyond nearest neighbors. Nucleic Acids Res. **17**:6523–6529.
23. **Johnston, B. H., W. Ohara, and A. Rich.** 1988. Stochastic distribution of a short region of Z-DNA within a long repeated sequence in negatively supercoiled plasmids. J. Biol. Chem. **263**:4512–4515.
24. **Jong, K. I., J. Heuser, and M. M. Cox.** 1989. Enhanced recA protein binding to Z-DNA represents a kinetic perturbation of a general duplex DNA binding pathway. J. Biol. Chem. **264**: 21848–21856.
25. **Keller, W.** 1975. Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. Proc. Natl. Acad. Sci. USA **72**:4876–4880.
26. **Kmiec, E. B., and W. K. Holloman.** 1986. Homologous pairing

of DNA molecules by Ustilago rec-1 protein is promoted by sequences of Z-DNA. Cell **44**:545–554.

27. **Krishna, P., B. P. Kennedy, J. H. van de Sande, and J. D. McGhee.** 1988. Yolk proteins from nematodes, chickens, and frogs bind strongly and preferentially to left-handed Z-DNA. J. Biol. Chem. **263**:19066–19070.
28. **Lafer, E. M., A. Moller, A. Nordheim, B. D. Stollar, and A. Rich.** 1981. Antibodies specific for left-handed Z-DNA. Proc. Natl. Acad. Sci. USA **78**:3546–3550.
29. **Lafer, E. M., R. Sousa, B. Rosen, A. Hsu, and A. Rich.** 1985. Isolation of Z-DNA binding proteins from wheat germ. Biochemistry **24**:5070–5076.
30. **Lafer, E. M., R. J. Sousa, and A. Rich.** 1988. Z-DNA-binding proteins in *E. coli*: purification generation of monoclonal antibodies, and gene isolation. J. Mol. Biol. **203**:511–516.
31. **Leith, I. R., R. T. Hay, and W. C. Russell.** 1988. Detection of Z-DNA binding proteins in tissue culture cells. Nucleic Acids Res. **16**:8277–8289.
32. **McFarlane, D., I. Farrance, I. Hall, J. Morris, and R. Ivarie.** 1986. The rat prolactin gene contains at least six poly (dT-dG) poly (dC-dA) repeats. Nucleic Acids Res. **14**:7805.
33. **McLean, M. J., J. W. Lee, and R. D. Wells.** 1988. Characteristics of Z-DNA helices formed by imperfect (purine-pyrimidine) sequences in plasmids. J. Biol. Chem. **263**:7378–7385.
34. **McLean, M. J., and R. D. Wells.** 1988. The role of DNA sequence in the formation of Z-DNA versus cruciforms in plasmids. J. Biol. Chem. **263**:7370–7377.
35. **Naylor, L. H., and E. M. Clark.** 1990. $d(TG)_n \cdot d(CA)_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. Nucleic Acids Res. **18**:1595–1601.
36. **Nordheim, A., and K. Meese.** 1988. Topoisomer gel retardation: detection of anti-Z-DNA antibodies bound to Z-DNA within supercoiled DNA minicircles. Nucleic Acids Res. **16**:21–37.
37. **Nordheim, A. P., F. Tesser, F. Azorin, Y. H. Kwon, A. Moller, and A. Rich.** 1982. Isolation of *Drosophila* proteins that bind selectively to Z-DNA. Proc. Natl. Acad. Sci. USA **79**:7729–7733.
38. **Pardoll, D. M., B. Vogelstein, and D. S. Coffey.** 1980. A fixed site of DNA replication in eucaryotic cells. Cell **19**:527–536.
39. **Peck, L. J., and J. C. Wang.** 1983. Energetics of B-to-Z transition in DNA. Proc. Natl. Acad. Sci. USA **80**:6206–6210.
40. **Ramesh, N., Y. S. Shouche, and S. K. Brahmachari.** 1986. Recognition of B and Z forms of DNA by Escherichia coli DNA polymerase I. J. Mol. Biol. **190**:635–638.
41. **Rich, A., A. Nordheim, and A. H.-J. Wang.** 1984. The chemistry and biology of left-handed Z-DNA. Annu. Rev. Biochem. **53**:791–846.
42. **Singleton, C. K., J. Klysik, and R. D. Wells.** 1982. The facile generation of covalently closed, circular DNAs with defined negative superhelical density. Anal. Biochem. **122**:253–257.
43. **Strobl, J. S., P. S. Dannies, and E. B. Thompson.** 1986. Rat growth hormone gene expression is correlated with an unmethylated CGCG sequence near the transcription initiation site. Biochemistry **25**:3640–3648.
44. **Trifonov, E. N., A. K. Konopka, and T. M. Jovin.** 1985. Unusual frequencies of certain purine-pyrimidine runs in natural DNA sequences. FEBS Lett. **185**:197–202.
45. **Ulanovsky, L., M. Bodner, E N. Trifonov, and M. Chodner.** 1986. Curved DNA: design, synthesis, and circularization. Proc. Natl. Acad. Sci. USA **83**:862–866.
46. **Wang, A. H.-J., R. V. Gessner, G. A. van der Marel, J. H. van Boom, and A. Rich.** 1985. Crystal structure of Z-DNA without an alternating purine-pyrimidine sequence. Proc. Natl. Acad. Sci. USA **82**:3611–3615.
47. **Wang, A. H.-J., G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, and A. Rich.** 1979. Molecular structure of a left handed double helical DNA fragment at atomic resolution. Nature (London) **282**:680–686.
48. **Wang, J. C., L. J. Peck, and K. Becherer.** 1983. DNA supercoiling and its effect on DNA structure and function. Cold Spring Harbor Symp. Quant. Biol. **48**:85–91.
49. **Wells, R. D., and S. C. Harvey (ed.).** 1988. Unusual DNA structures. Springer-Verlag, New York.
50. **Wittig, B., T. Dorbic, and A. Rich.** 1989. The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. J. Cell Biol. **108**:755–764.
51. **Wohlrab, F., and R. D. Wells.** 1987. Enzymatic probes for left-handed Z-DNA Gene Amplif. Anal. **5**:247–256.