Mung bean nuclease cleavage of a dA + dT-rich sequence or an inverted repeat sequence in supercoiled PM2 DNA depends on ionic environment

Lowell G. Sheflin and David Kowalski

Roswell Park Memorial Institute, Department of Cell and Tumor Biology, Buffalo, NY 14263, USA

Received 25 June 1984; Revised and Accepted 28 August 1984

ABSTRACT

We have determined the nucleotide sequences around two alternative sites cleaved in supercoiled PM2 DNA by single-strand-specific mung bean nuclease in different ionic environments. In 10 mM Tris-HCl (pH 7.0, 37°C), the major site is a dA+dT-rich sequence which maps with a known early denaturation region at 0.75 map units. About 30 cleavages occurred in a 135 bp region. Cleavages were largely excluded at (dA)_n.(dT)_n (n=3-7) sequences. Cleavage patterns of this type have not been previously observed in dA+dT-rich sequences. With the addition of 0.1 M NaCl the major alternative site occurred in a hyphenated inverted repeat sequence 500 bp away (0.70 map units) and did not map to an early denaturation region. One major and 4 minor cleavages occurred in the region between the repeats, suggesting that a hairpin containing at most a 12 bp stem and 10 base loop is recognized. The basis for nuclease recognition of the dA+dT-rich sequence is not clear. The differences in the sequences and cleavage patterns at the alternative sites indicate that their secondary structures differ.

INTRODUCTION

Negative supercoiling favors untwisting of DNA strands which can be detected using single-strand-specific endonucleases. The locations of these sites were originally found to correlate with early denaturation regions (1-3). Such regions are enriched in dA+dT sequences (4,5), which may undergo transient untwisting or breathing under non-denaturing conditions (6,7). However, subsequent studies have shown that not all early denaturation regions are cleaved and sites other than early denaturation regions can be cleaved (8,9). The application of fine mapping and DNA sequencing technologies has revealed that single-strand-specific endonucleases can cleave between inverted repeat sequences (potential hairpins or cruciforms with non-base-paired loops) (10-12), regions flanking (dGdC)_n and (dGdT)_n (unwound regions between Z- and B-DNA) (13-15), regions within (dGdA)_n (16), (dG)_n (17-19) or (AGAGG)₃₂ (20) (potential slippage sites) or regions between (dGdA)_n repeats (looped-out sequences between interacting direct repeats) (21). Of these sequences, only the inverted repeat can be dA+dT-rich. Although some inverted repeats are associated with

dA+dT-rich sequences (12), dA+dT-richness per se has not been found to be required for cleavage. Therefore, although single-strand-specific endonuclease cleavage sites were originally correlated with dA+dT-rich sequences which might breathe in supercoiled DNA, cleavage within such sequences for reasons other than hairpin or cruciform formation have not been directly demonstrated.

Our laboratory has mapped a major mung bean nuclease site, located at 0.75 map units, within one of eight early denaturation regions in supercoiled PM2 DNA (9). An interesting feature of this site is that its detection depends on ionic environment (Kowalski, the preceding paper). The site is cleaved in 10 mM Tris-HCl (pH 7.0, 37°C). Addition of 0.1 M NaCl nearly eliminates cleavage at this site and alternative sites are cleaved. A major alternative site occurs at 0.70 map units and does not correlate with an early denaturation region. Since the sequence of PM2 DNA is not known, the nature of the DNA sequences in the vicinity of these two sites could not be inferred from their locations.

In this paper we report the assignment of these two mung bean nuclease sites to two different restriction fragments of PM2 DNA, the nucleotide sequences of these fragments, and the location of mung bean nuclease cleavages within these sequences. The site at 0.75 map units is dA+dT-rich and contains about 30 cleavage sites distributed non-randomly over a 135 bp region. These findings, while consistent with A-T breathing at preferred sites in this region, are subject to other possible interpretations which will be discussed. The region around 0.70 map units is not dA+dT-rich and contains one major cleavage site in a hyphenated inverted repeat sequence. These findings are consistent with recognition of a hairpin, two types of which will be discussed. Overall, our results show that the nature of the DNA sequences cleaved by mung bean nuclease, and the altered secondary structures which can be detected, vary with ionic environment. Both the nucleotide sequences and the altered secondary structures described here are particularly interesting since they occur in regions of potential importance to PM2 DNA replication and transcription.

MATERIALS AND METHODS

Enzymes

Mung Bean nuclease was isolated and purified as described by Kowalski et al. (22). Restriction endonucleases were purchased from New England BioLabs. Bacterial alkaline phosphatase was obtained from Worthington Biochemicals. Polynucleotide kinase was obtained from P-L Biochemicals.

DNA

Bacteriophage PM2 DNA was purified as previously described (23). The DNA

was >90% supercoiled as determined by electrophoresis on 1% agarose gels. The superhelical density of the DNA was -0.084 (Kowalski, the preceding paper).

Cleavage of Supercoiled PM2 DNA with Mung Bean Nuclease

Supercoiled PM2 DNA was cleaved in the absence and presence of NaCl. In the absence of NaCl the reaction mixture contained 10 mM Tris-HCl (pH 7.0), 32 µg of PM2 DNA in a volume of 288 µl. In the presence of NaCl the reaction mixture contained the preceding components plus 0.1 M NaCl in the same volume. After preincubation for 15 min at 37°C, mung bean nuclease (32 µl, diluted in 10 mM Tris-HCl, pH 7.0, 0.005% Triton X-100) was added and the mixture incubated at 37°C. The mung bean nuclease concentration in the reaction mixture was 0.1 units per ml in the absence of NaCl and 2 units per ml in the presence of NaCl. Reaction at 2 units per ml in the absence of NaCl gave the same site specificity as 0.1 units per ml (see preceding paper). After 30 min incubation, which was sufficient to cleave all the supercoiled DNA, 0.1 M Na₂EDTA (35 µl) was added and the mixture was cooled to 0°C. The enzyme was removed by extraction with phenol (saturated with 10 mM Tris-HCl (pH 7.4) 1 mM Na₂EDTA).

End Labeling of DNA

Restriction endonuclease digestion of PM2 DNA was carried out using conditions recommended by New England BioLabs. Restriction endonuclease fragments of DNA were labeled at the 5' ends according to the procedure of Woodworth-Gutai (24). First, 5' terminal phosphates were removed from DNA fragments (5-50 µg per ml) by treatment with bacterial alkaline phosphatase (2.4 µg per ml), in 50 mM NaCl, 50 mM Tris-HCl (pH 7.4) at 37°C for 30 min. The reaction was stopped by adding Na₂EDTA to a final concentration of 50 mM. After 30 min at 20-22°C, the reaction mixture was extracted twice with phenol (saturated with 10 mM Tris-HCl (pH 7.4), 1 mM Na₂EDTA) and the aqueous phase dialyzed in 10 mM Tris-HCl (pH 7.4), 1 mM Na₂EDTA. DNA fragments were end-labeled using polynucleotide kinase (4 units) in a 100 µl reaction mixture containing 50 mM Tris-HCl (pH 9.5), 10 mM MgCl₂, 5 mM dithiothreitol, 5% (v per v) glycerol, 0.1 mM spermidine and gamma-[³²P]-ATP. Thirty µCi gamma-[³²P]-ATP (3000 Ci per mmol) were used for mapping restriction enzyme sites and 100 µCi were used for DNA sequencing and mapping mung bean nuclease sites. After 30 min at 37°C the reaction mixture was dialyzed for 18 h in 1 M NaCl, 5 mM Na₂EDTA, 10 mM Tris-HCl (pH 7.4) and 3 h in 10 mM Tris-HCl (pH 7.4).

Purification of End-Labeled DNA Fragments

End-labeled restriction endonuclease fragments of PM2 DNA were electrophoresed in vertical 4% (top half)-8% (bottom half) polyacrylamide gels

(1.5 mm thick, 18 cm wide, and 40 cm long) at 150 volts for 18–24 h in electrophoresis buffer containing 40 mM Tris (pH 7.8), 20 mM sodium acetate, and 2 mM Na₂EDTA. End-labeled DNA fragments were identified by exposure of the gel to X-Ray film (XRP, Eastman Kodak) at 20–22°C. Portions of the gel containing fragments of interest were diced with a scalpel, placed in elution buffer (0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.1 mM Na₂EDTA, 0.1% SDS), and incubated at 37–42°C for 24–48 h. After removal of gel fragments by filtration through glass wool, the DNA was precipitated by adding 2 volumes of ethanol (–70°C for 30 min), resuspended in 0.3 M sodium acetate (pH 6.0), reprecipitated using ethanol, and dried under vacuum. DNA was redissolved in H₂O.

DNA Sequencing

DNA fragments containing a single 5' end label were generated by cleavage of end-labeled DNA with restriction enzymes, and were isolated and purified according to the procedures described above. DNA was sequenced using the Maxam and Gilbert (25) procedure. Products of the sequencing reactions were denatured by heating at 90°C for 2 min in a solution containing 1 mg per ml xylene cyanole, 1 mg per ml bromophenol blue, 98% formamide and 1 mM Na₂EDTA and electrophoresed in 10% or 20% polyacrylamide–8 M urea slab gels (0.4 mm thick, 18 cm wide and 40 cm long) at 1400 volts for 6 h (10% gel) or 24 h (20% gel). Each set of samples was loaded at two different times on each gel (10%:0 h and 3 h or 20%:0 h and 16 h) in order to examine different regions of the same DNA sequence. Mung bean nuclease cleavage products were denatured and electrophoresed in the same way. The gels were wrapped in Saran wrap and exposed to X-ray film both with and without intensifying screen (Cronex, Dupont) at –70°C for appropriate time periods.

To read the locations of the mung bean nuclease cleavages within the DNA sequence, standard correction factors used for S1 nuclease mapping were applied. The basis for these correction factors is as follows. A DNA fragment obtained from the Maxam–Gilbert sequencing reactions moves one nucleotide faster than a fragment cut at the same nucleotide by mung bean nuclease since the nucleotide at the 3' end of the Maxam–Gilbert fragment is destroyed (25). Also, since the Maxam–Gilbert fragment has a 3'P while the mung bean nuclease fragment has a 3'OH, the Maxam–Gilbert fragment moves an additional one nucleotide or less faster than the mung bean nuclease fragment, depending upon the fragment length. Nucleotide increments for given fragment lengths determined by Tapper and Clayton (47) are: 1 for <30 bases, 0.5 for 30–50 bases, 0.2 for 50–100 bases, and 0 for >100 bases.

RESULTS

Assignment of Major Mung Bean Nuclease Cleavage Sites to Specific Hind III Fragments of PM2 DNA

Although mung bean nuclease cleavage sites have been mapped on the PM2 genome (9; Kowalski, the preceding paper), these sites have not been located in restriction fragments of suitable size for nucleotide sequence analysis. Given an error of ± 0.01 map units, it was predicted that the mung bean nuclease site at 0.75 map units would be located in Hind III fragment 4 (0.72-0.76 map units) and the site at 0.70 would be located in Hind III fragment 3 (0.63-0.72 map units). Confirmation of these predictions was necessary before attempting nucleotide sequence analysis. Supercoiled PM2 DNA was cleaved once per molecule by mung bean nuclease in the absence and presence of NaCl (see Materials and Methods). The products, mainly nicked circular and some linear DNA, were digested by restriction endonuclease Hind III and the resultant fragments ^{32}P -labeled at their 5' ends. The end-labeled DNA fragments were then denatured and electrophoresed in a 10% polyacrylamide-8 M urea gel. Shown in Figure 1 are two autoradiograms representing two exposures of this gel to X-ray film. After reaction with mung bean nuclease in the absence of NaCl (lane 2), Hind III fragment 4 is significantly decreased in intensity compared to fragment 5. However, in the presence of 0.1 M NaCl the intensity of these two fragments is similar (lane 3). These results show that in the absence of NaCl, mung bean nuclease cleavages occur in Hind III fragment 4. Since other fragments are not comparably reduced in intensity, fragment 4 contains the major mung bean nuclease site previously mapped at 0.75 (Kowalski, the preceding paper). In the presence of 0.1 M NaCl, the major mung bean nuclease cleavage sites do not occur in Hind III fragment 4 since fragments 4 and 5 are similar in intensity (lane 3). The major cleavage has shifted to Hind III fragment 3 (lane 3), since this fragment is decreased in intensity relative to the other fragments. Therefore, in the presence of NaCl, Hind III fragment 3 contains the major mung bean nuclease cleavage site previously mapped at 0.70 (Kowalski, the preceding paper).

The number and size distribution of mung bean nuclease cleavage products from these restriction endonuclease fragments depends on the concentration of NaCl in the mung bean nuclease reaction. In the absence of NaCl, a large number of cleavage products mainly <420 bases to >80 bases are produced (Figure 1, lane 5). These data suggest that a broad region in fragment 4 is cleaved. In the presence of 0.1 M NaCl, however, only two major cleavage products of 750 and 300 bases are generated (Figure 1, lanes 3 and 6). These data suggest that in the

Figure 1. Autoradiogram showing the Hind III fragments of PM2 DNA with and without prior cleavage of supercoiled DNA by mung bean nuclease. Supercoiled PM2 DNA was nicked once per molecule with mung bean nuclease in the absence and presence of 0.1 M NaCl (see Materials and Methods). The nicked-circular DNA was then digested with restriction endonuclease Hind III and the resultant fragments 5' end-labeled with ^{32}P (See Materials and Methods). As a control, supercoiled PM2 DNA was cut with Hind III and the fragments end-labeled. The end-labeled fragments were heat denatured and electrophoresed in a 10% polyacrylamide-8 M urea gel (25). The same gel was exposed to different X-ray films for 36 h (lanes 1-3) and 96 h (lanes 4-6). Lanes 1, 4: control without mung bean nuclease treatment. Lanes 2, 5: mung bean nuclease treatment in the absence of NaCl. Lanes 3, 6: mung bean nuclease treatment in the presence of 0.1 M NaCl. Numbers along the left side of the autoradiogram refer to Hind III fragments 1-7, in order of decreasing size (5350, 2150, 924, 402, 371, 273, and 80 bases (46)). The nucleotide sequence of fragment 4 shows that the actual size is 469 bp. Our estimates for the sizes of fragments 3, 5, 6 and 7 are 1050, 420, 200 and 80 bases, respectively (unpublished results). The band below fragment 5 is a partial digestion product consisting of fragments 6 + 7. The numbers along the right side of the autoradiogram are the number of bases in major fragments produced by mung bean nuclease in the presence of 0.1 M NaCl.

Figure 2. Restriction endonuclease maps and nucleotide sequencing strategies for (A) Hind III fragment 4 and (B) the right end of Hind III fragment 3. The numbers represent map units (ratio of the fragment size in base pairs to 9550 bp (46) in PM2 DNA) using the Hind III site at 0.72 map units (45) as a reference point. The arrows show the regions of the nucleotide sequences determined. The tail of each arrow is at the 5' end of each fragment. The sequence of each fragment was determined at least 3 times.

presence of 0.1 M NaCl there may be only one major cleavage in Hind III fragment 3 since the sum of the two mung bean nuclease cleavage products is approximately equal to the size of Hind III fragment 3 (see Figure 1 legend for Hind III fragment sizes). The dependence of the number and size distribution of the mung bean nuclease cleavage products on NaCl concentration suggests that the two regions recognized by mung bean nuclease differ in single-strand character and thus in the nature of their sequences.

Restriction Endonuclease Maps and Nucleotide Sequences of Hind III Fragments 3 and 4

To facilitate nucleotide sequence analyses, restriction endonuclease maps of more detail than those available for the two regions of interest in PM2 DNA were determined. These results are summarized in Figure 2. Figure 2A shows the restriction endonuclease map of Hind III fragment 4. Figure 2B shows a restriction endonuclease map of the right end the Hind III fragment 3. Also shown in Figure 2 are the regions of PM2 DNA whose sequences were determined (arrows). Nucleotide sequences determined for both of the Hind III fragments are shown in Figure 3. A sequence for Hind III fragment 4 has been reported by Urquiaga et al. (26). Their sequence contains 437 bp instead of 469 bp as shown in Figure 3A. In addition, after allowing for deletions and insertions, small portions of the two sequences differ. It is possible that the sequence

differences reflect differences in bacteriophage PM2 strains. Alternatively, since Urquiaga et al. (26) determined the sequence after cloning PM2 DNA in pBR322, while we have determined the sequence of PM2 DNA directly, it is possible that these differences may be due to a cloning artifact. The sequence of Hind III fragment 4 is enriched in dA+dT since the dA+dT content is 64% compared to 57% for the PM2 genome (27). Also, this sequence is slightly enriched in dA+dT compared to the 376 bp sequence at the right end of the Hind III fragment 3 (61% dA+dT).

Nucleotide Sequences Containing Mung Bean Nuclease Nicks

We wanted to determine the nucleotide sequences at the initial sites of mung bean nuclease nicking in supercoiled PM2 DNA. To accomplish this, we used the following procedure. Supercoiled PM2 DNA was cleaved with mung bean nuclease in the absence and presence of 0.1 M NaCl. The DNA was digested with restriction enzymes (Figure 2A and B), and the 5' ends of the restriction sites labeled with ^{32}P . End-labeled DNA of interest was purified and digested with a second restriction endonuclease in order to obtain singly end-labeled DNA. This DNA was denatured and electrophoresed along side the products of the 4 Maxam-Gilbert sequencing reactions performed on the same singly end-labeled restriction enzyme fragment of PM2 DNA without mung bean nuclease treatment. This procedure selects for restriction enzyme fragments of DNA containing mung bean nuclease nicks and excludes those fragments containing double-strand breaks. The procedure permits the determination of the location and relative frequency of the mung bean nuclease nicks within the DNA sequence. Using this approach, only one region of mung bean nuclease cleavages was found in each of the two nucleotide sequences shown in Figure 3A and B and these results are summarized below.

The result of mung bean nuclease cleavage in the absence of NaCl is shown in Figure 4. Figure 4A shows two loadings of sequencing reactions which reveal the nucleotide sequence in the bottom strand to the left of the Hind III site at 0.7692 map units (see Figure 2A). The lanes labeled MBN display the mung bean nuclease cleavage products. About 38 cleavage products are seen. The intensity of the products varies, indicating that the frequency of cleavages at different sites within this region also varies. The nucleotide sequence containing the mung bean nuclease cleavages is shown in Figure 4B (bottom strand). The 38 cleavages are distributed over a 102 base region from 0.7484-0.7591 map units. Also shown in Figure 4B are the results of the sequence analysis of mung bean nuclease cleavages within the complementary (top) strand (to the right of the Hinc II site at 0.7403 map units, see Figure 2A). The distribution of mung bean

A

```

1
AGCTTACTGG TAGGTGAACA CCCGAAACCC TTAAGAGGCA TGCTTTGCT
51
TGAAGCTCAC AAGGCGAAA GAAATGAAA TGAAGAACAG GCTAAATTGT
101
GGCGAAAAAG CGTTGCTTAA GTAATTAACA CCCCGTCAA AATGTGTTG
151
TTGTTTGGG TAAAGTAGG TAAATAAGG GCTGTAGGCT GTTGACGTGC
201
TTACGATTTA TATTATGTTA AATAACTATA AATCGTCAT TCTGTGCGCT
251
TTCCCTGTT GGTGTGTTAA ATCCTATCTA ACTGTTATTT ATATAAGTTT
301
TCTATAATTT ACCATTAATC AATTTTATTA ATTGAACGCG TTATGCAATC
351
TAAAAAAATC ATCGTTAAAT CTCGCGGCT GGTAAATAA CTATGTCAAT
401
GGAACTCAA GGAATTAAT GACTGATTC AAACAATTAC TTTTAAAGGC
451
TGGCTTTATG AATTTTGCA

```

B

```

1
TCACCTTGGC GATTAAAAAA CTGTTTCACT CTGGCGTCAT AAAAAAGCGG
51
TTACTGTATA TAAACCACT AACGCGCCAG CCAAGATTA GCAGCGTTG
101
CCCTTGCTAA AAGTGATGAT AGTTCAITGC TCOCGACTAT GTCAAGCGCT
151
CGCACTAATG CAATAGCTGT TGACTAGGTC TATAATCAA TTAATTGAGA
201
CGGTTACGA GTACAAAACCT TATGAAAAA ACAACTTGA TTGAAGTCT
251
TGATTGGTTT AAAAGCGTTG TAGATATTGA CTCTGACTAT ATGGTTAGTA
301
AGTTAACAGG CATCCCAAAA CAAACACTTA GCAGTGTAAG AACTGGAAC
351
TCTGAATTTT CAGACTACAC AGCTTA

```

Figure 3. Nucleotide sequence of (A) Hind III fragment 4 and (B) the right end of Hind III fragment 3. The sequence of Hind III fragment 4 begins at 0.7200 map units and ends at 0.7692 map units (Figure 2A). The sequence of the right end of Hind III fragment 3 begins at the Hph I recognition site (TCACC) at 0.6809 map units and ends at the Hind III site at 0.7200 map units (Figure 2B). The Hph I cleavage site is 7 bp to the left of the first nucleotide. One-hundred base pairs is equal to 0.0105 map units in PM2 DNA.

nuclease cleavages in the top strand resembles that observed in the bottom strand. These results confirm the earlier finding that mung bean nuclease cleavage of this region is not strand specific (9). A prominent feature of this sequence is its dA+dT content. The 135 bp region spanning the cleavages in both strands has a dA+dT content of 74% which is significantly higher than that of 57% for the PM2 genome (27).

The result of mung bean nuclease cleavage in the presence of 0.1 M NaCl is shown in Figure 5. The autoradiogram in Figure 5A reveals the nucleotide sequence in the bottom strand to the left of the Hinc II site at 0.6984 map units. The lanes labeled MBN display the mung bean nuclease cleavage products. One major and four minor cleavage products are seen. The nucleotide sequences containing these cleavages are depicted along the left side of Figure 5A. While the cleavages fall within a 10 base dA+dT-rich region, the surrounding sequence is not enriched for dA+dT. The sequences flanking the cleavages are inverted repeats of 10 bases. Allowing G-T pairing (11), this sequence can form a hairpin with a stem of 12 and a loop of 10. The mung bean nuclease cleavages occur within the 10 base loop (see Figure 5B).

DISCUSSION

The preceding paper has shown that the location of mung bean nuclease cleavage sites in supercoiled PM2 DNA can change with ionic environment and temperature. This report shows that the nature of the sequences cleaved by mung bean nuclease in supercoiled PM2 DNA change with ionic environment. In the absence of NaCl, a long dA+dT-rich sequence is cleaved. Previous studies mapped a major cleavage site at 0.75 map units (Kowalski, the preceding paper) or at 0.75 and 0.76 map units (9) depending upon the resolution of the agarose gel. The fine mapping results presented here show that there are actually multiple cleavages spanning 135 bp from 0.7451–0.7591 map units. The dA+dT-content of this region is enriched (74%) compared to that of the PM2 genome (57%). Addition of 0.1 M NaCl results in the appearance of alternative cleavage sites, the major of which is at 0.70 map units (Kowalski, the preceding paper). Fine mapping places this site at 0.6868 map units. The cleavages fall in the center of a hyphenated inverted repeat sequence. This sequence is 30 bp long containing a 10 bp inverted repeat separated by a 10 bp nonrepeated sequence. The dA+dT content of the region, ± 67 bp from the center of the hyphenated inverted repeat sequence (chosen for comparison to the 135 bp sequence discussed above) is similar (59%) compared to that of the PM2 genome (57%).

The cleavage pattern observed in the dA+dT-rich sequence is complex (Figure 4B). The top strand is cleaved at 29 positions and the bottom strand is cleaved at 38 positions. These cleavages are derived from a population of DNA molecules

Figure 4. Nucleotide sequences containing major mung bean nuclease cleavage sites in the absence of NaCl. (A) Autoradiogram of cleavage products derived from mung bean nuclease (MBN) or Maxam-Gilbert sequencing reactions (A>C, G, C, C+T) after electrophoresis in a 10% polyacrylamide-8 M urea gel. Two different electrophoresis times for the same set of cleavage products are shown. The strategy for mapping the mung bean nuclease cleavages in the DNA sequence is described in the Results. The sequence was obtained from ³²P end-labeled Hind III fragment 4 cleaved with Hae III. The 5' end-label is at the Hind III site at 0.7692 map units. The 3' end of the fragment is at the Hae III site at 0.7268 map units (see Figure 2A). A similar restriction enzyme fragment isolated from PM2 DNA nicked with mung bean nuclease in the absence of NaCl (see Materials and Methods) was used to examine the mung bean nuclease cleavages. The bracketed area spans the mung bean nuclease cleavage products from 0.7484–0.7591 map units. Since the mung bean nuclease fragments in Figure 4A are >100 bases long, no correction was required for the absence of a 3'P (see Materials and Methods). (B) Nucleotide sequence. The sequence determined in Figure 4A is shown in the bottom strand. Sequence analysis of the top strand was performed independently (autoradiogram not shown). The sequence begins at 0.7427 map units and ends at 0.7615 map units. Vertical lines represent mung bean nuclease cleavages. Increasing height of the lines indicates increasing relative cleavage frequencies within one strand determined by inspection of band intensities. A 13 bp direct repeat sequence which flanks the cleavages is enclosed in boxes.

Figure 5. Nucleotide sequences containing major mung bean nuclease cleavage sites in 0.1 M NaCl. (A) Autoradiogram of cleavage products derived from mung bean nuclease (MBN) or Maxam-Gilbert sequencing reactions (A>C, G, C, C+T) after electrophoresis in a 10% polyacrylamide-8 M urea gel. The strategy for mapping the mung bean nuclease cleavages within the DNA sequence is described in the Results. The sequence was obtained from a DNA fragment which was ³²P end-labeled at the Hinc II site at 0.6984 map units. The 3' end of the fragment is at the Hinf I site at 0.6600 map units. A similar restriction enzyme fragment isolated from PM2 DNA nicked with mung bean nuclease in the presence of 0.1 M NaCl (see Materials and Methods) was used to examine the mung bean nuclease cleavages. The horizontal arrows mark the locations of the five mung bean nuclease cleavages within the nucleotide sequence. Increasing length of the arrows indicates increasing relative cleavage frequencies determined by inspection of band intensities. Since the mung bean nuclease fragments in Figure 5A are >100 bases long, no correction was required for the absence of a 3'P (see Materials and Methods). The vertical arrows indicate a 10 base inverted repeat sequence. The 5' end of the sequence shown is at T. The sequence is the reverse complement of part of the sequence shown in Figure 3B.

(B) Possible hairpin formed by the inverted repeat sequence. The inverted repeat sequence in Figure 5A is shown as a hairpin with a stem of 12 bases (permitting two GT pairs (11)) and a loop of 10 bases. Increasing line lengths indicate increasing cleavage frequencies. The major cleavage occurs at 0.6868 map units. A double-stranded hairpin is a possible alternative to this structure (see Discussion).

and do not occur in one DNA molecule. A single nick relieves the torsional strain of negative supercoiling rendering the molecule resistant to further nicking. Multiple cleavages in the same strand are not the result of nibbling at the ends of one or more initial sites of cleavage since our procedure selects for nicked-circular and not linear DNA molecules. These DNA molecules contain mainly nicks and not gaps since 1) the bulk of the cleavages have been shown to be ligatable (9) and 2) gaps are rapidly cleaved by mung bean nuclease (28). In contrast to the cleavage pattern in the dA+dT-rich sequence, the cleavage pattern observed in the hyphenated inverted repeat sequence is very simple (Figure 5B). Only 5 cleavages were detected. One major and 4 minor cleavages are distributed over 7 bases. Thus, the complexity of the cleavage patterns at the two sequences differ.

The complex cleavage pattern observed in the dA+dT-rich sequence is different than patterns of S1 nuclease cleavage in sequences which can form hairpins (cruciforms), Z-DNA, or loop outs resulting from slippage (see Introduction). The two longest inverted repeat sequences (ATTAAAT and TCAATT) are cleaved both between and within the repeats. Cleavages within the repeats indicate that structures other than hairpins are recognized. Also, many other sequences which are not inverted repeats are cleaved. This is the first evidence that a dA+dT-rich sequence in supercoiled DNA is cleaved by a single-strand-specific nuclease for reasons other than hairpin formation. An alternating Py-Pu stretch of 10 bases with one base out of alternation occurs at the left end of the dA+dT-rich sequence cleaved. Minor cleavages flank this sequence and one minor cleavage occurs within it. Z-DNA formation could account for the flanking minor cleavages but not for the majority of the cleavages throughout the dA+dT-rich region. This region does not contain (dGdA)_n or (dG)_n sequences which are cleaved by S1 nuclease within but not outside the sequences, possibly as a result of slippage. Cleavage of the dA+dT-rich sequence occurs mainly outside but not within simple repeats of (dA)_n and (dT)_n (n=3-4), suggesting slippage alone cannot account for the cleavage pattern. S1 nuclease cleaves between (dCdT)₃ repeats separated by 1 to 13 bp (21). It was suggested that the latter sequences can slip out of register, interact with an upstream or downstream repeat and loop-out sequences between the repeats. The (dT)_n

stretches, although separated by 6 to 14 bp differ from the $(dCdT)_n$ sequences in length and base composition. If the $(dT)_n$ sequences could slip out of register and interact with each other, this might explain the complex cleavage pattern. An alternative possible explanation which also takes into account the exclusion of cleavages in $(dA)_n.(dT)_n$ is considered later.

The dA+dT-rich sequence cleaved is flanked by a 13 bp direct repeat (Figure 4B). However, it is not simple and therefore it would not be expected to undergo slippage. Also, interaction of the 13 bp repeats would create an enormous looped-out sequence of 165 bp. Since every base of this 165 base region is not cleaved it is necessary to postulate additional interactions between or within the two strands. In addition to the 13 bp repeats, this region also contains many smaller complex repeats of 3 to 7 bp. These repeats would not be expected to undergo slippage. No clear pattern of cleavages with respect to the smaller complex repeated sequences can be deduced. Thus, the interaction between complex repeated sequences is an unlikely explanation for the cleavage pattern.

Since the cleavages occur in a long dA+dT-rich sequence that maps with an early denaturation region in PM2 DNA (9), it might be inferred that this sequence is undergoing A-T breathing. The present study can not distinguish whether the dA+dT-rich sequence is breathing (transiently untwisted), stably untwisted or both. Also, the minimum degree of untwisting required for single-strand-specific nuclease recognition is not known. Clearly, the entire sequence is not single-stranded, since if it were it is expected that nearly every base would be cleaved. The nonrandom nature of the cleavages suggest that the region is breathing or stably untwisted at preferred sites. No preferred sequences can be identified at these sites. Instead, cleavages appear to be excluded at $(dA)_n.(dT)_n$ ($n=3-7$ bp), possibly because they are less susceptible to breathing or are more twisted. Consistent with these possibilities, $(dA)_n.(dT)_n$ sequences have been shown to have a higher twist (fewer bp per turn) (30-32) and a higher T_m (33) than other dA+dT sequences in relaxed DNA. Exclusion of cleavages is not likely the result of inaccessibility because many other cleavages occur throughout the entire region. Additionally, exclusion of cleavages in $(dA)_3$ to $(dA)_7$ sequences can not be the result of preference for other bases because many of the $(dA)_2$ sequences present are cleaved. Thus, the complexity of the cleavage pattern could be the result of the high dA+dT content, length of the sequence and the presence of $(dA)_n.(dT)_n$ repeats. As an alternative to the possibility of simple direct repeat interactions, the basis for nuclease recognition could be breathing or stable untwisting of the duplex

at sequences other than $(dA)_n.(dT)_n$ within this dA+dT-rich region.

It is interesting to note that mung bean nuclease cleavages are not detected at seven other early denaturation regions present in PM2 DNA (9). Features that might be responsible for the inability to detect cleavages at these other regions include dA+dT content, length, nucleotide sequence and accessibility. Until the entire sequence of PM2 is determined we can not address the first three possibilities. The possibility that all seven early denaturation regions would be totally inaccessible to the enzyme seems unlikely.

Cleavage of the dA+dT-rich region is nearly eliminated in the presence of 0.1 M NaCl (Kowalski, the preceding paper) suggesting that the altered secondary structure either switches to duplex DNA or becomes inaccessible to the enzyme. It is unlikely that competition from newly-formed, more favorable sites explains the decrease in cleavage frequency of this site because addition of NaCl decreases, rather than increases, the initial rate of cleavage (data not shown). Since detection of this site depends on the ionic environment and temperature but is independent of three different single-strand-specific nuclease probes (Kowalski, the preceding paper), it is likely that similar cleavage of a dA+dT-rich region has not been observed using S1 nuclease because the environmental conditions differed from those used in our studies. Cleavage of a dA+dT-rich region is not unique to PM2 DNA since, under the same environmental conditions, pBR322 DNA is also cleaved in a single dA+dT-rich region by mung bean and P1 nucleases (unpublished results from this laboratory).

The simple mung bean nuclease cleavage pattern observed in the hyphenated inverted repeat sequence is similar to previously reported S1 nuclease cleavage patterns for inverted repeat sequences with the potential to form hairpins or cruciforms (11-12). The non-base-paired loop of potential cruciforms are cleaved while the base-paired stems are not. Therefore, it is possible that the major site cleaved by mung bean nuclease in this ionic environment is a hairpin or cruciform.

The 10 base loop of the potential cruciform recognized in the presence of NaCl is unusually large compared to the 3-6 base loops found in cruciforms detected in other supercoiled DNAs (10-12). The 12 bp stem is similar in length to the 9-13 bp stems reported. Potential cruciforms with loops greater than 6 bases have not been detected in other DNAs by S1 nuclease at acid pH (10-12) or T7 gene 3 endonuclease at neutral pH (12). However, the superhelical density of PM2 DNA (-0.084 for our preparation) is greater than that for many other DNAs studied (-0.05 to -0.07). Using published data to estimate the free energy of hairpin formation (48) and to calculate the free energy of supercoiling in the

absence and presence of the hairpin (49,50), the hairpin structure shown in Figure 5B is predicted to be thermodynamically stable in our preparation of PM2 DNA in 0.1 M NaCl at 37°C.

As an alternative to a cruciform or hairpin, a four-stranded helix has been proposed for inverted repeat sequences (34). This structure could be described as a double-stranded hairpin. Formation of this structure requires bending the DNA between the inverted repeat sequences allowing them to interact and form a 4-stranded helix. Local untwisting is expected at the bend and might be recognized by mung bean nuclease. The resultant cleavage pattern is expected to be indistinguishable from that of a cruciform or single-stranded hairpin. Thus, the altered secondary structure at this site could be associated with a single- or double-stranded hairpin. The basis for cleavage of this altered secondary structure in the presence but not in the absence of 0.1 M NaCl may involve either the formation of this structure or differences in the ability to detect this structure which exists under both conditions (Kowalski, the preceding paper).

The data presented here provide information on the primary and secondary structures at or near possible regulatory sites in supercoiled PM2 DNA. In the absence of NaCl, the major mung bean nuclease cleavages (0.7451-0.7591 map units) occur in sequences at or near the initiation of PM2 DNA replication (26) (within Hind III fragment 4, 0.7200-0.7692 map units) and near a strong RNA polymerase binding site (35) (Hinc II site at 0.7403 map units). In the presence of 0.1 M NaCl, the major mung bean nuclease cleavage (0.6868 map units) occurs near a second strong RNA polymerase binding site (35) (Hinc II site at 0.6984 map units). Thus, altered secondary structures can occur at or near these potential regulatory sites. Others have demonstrated S1 nuclease cleavage within the lac promoter (18), the adenovirus major late promoter (36), and in the 5' flanking region of globin genes (17-19,37-39) in supercoiled DNA. We and others (37,40) have shown that detection of altered secondary structures can change with ionic environment. Changes in ionic environment are known to affect RNA polymerase binding to some promoters in linear DNA (41-43) as well as the specificity of inhibition of *in vitro* transcription by chemicals which react with single-stranded regions of supercoiled DNA (44). Our findings suggest that one way the ionic environment may affect these processes is by altering the secondary structure in supercoiled DNA. It remains to be determined whether these altered secondary structures are involved in the regulation of supercoiled PM2 DNA transcription and replication *in vitro* or *in vivo*.

ACKNOWLEDGEMENT

This research was supported by National Institutes of Health Grants GM24950 and GM30614.

REFERENCES

1. Beard, P., Morrow, J.F. and Berg, P. (1973) *J. Virol.* 12, 1303-1313.
2. Dasgupta, S., Allison, D.P., Snyder, C.E. and Mitra, S. (1977) *J. Biol. Chem.* 252, 5916-5923.
3. Germond, J.E., Vogt, V.M. and Hirt, B. (1974) *Eur. J. Biochem.* 43, 591-600.
4. Funnell, B.E. and Inman, R.B. (1979) *J. Mol. Biol.* 131, 331-340.
5. Tong, B.Y. and Battersby, S.J. (1979) *Nucleic Acids Res.* 6, 1073-1079.
6. von Hippel, P.H. and Felsenfeld, G. (1964) *Biochemistry* 3, 27-39.
7. Wingert, L. and von Hippel, P.H. (1968) *Biochem. Biophys. Acta.* 157, 114-126.
8. Pritchard, A.E. and Laskowski, M., Sr. (1978) *J. Biol. Chem.* 253, 6606-6613.
9. Kowalski, D. and Sanford, J.P. (1982) *J. Biol. Chem.* 257, 7820-7825.
10. Lilley, D.M.J. (1980) *Proc. Natl. Acad. Sci. USA* 77, 6468-6472.
11. Lilley, D.M.J. (1981) *Nucleic Acids Res.* 9, 1271-1289.
12. Panayotatos, N. and Wells, R.D. (1981) *Nature* 289, 466-470.
13. Singleton, C.K., Klysik, J., Stirdivant, S.M. and Wells, R.D. (1982) *Nature* 299, 312-316.
14. Singleton, C.K., Kilpatrick, M.W. and Wells, R.D. (1984) *J. Biol. Chem.* 259, 1963-1967.
15. Kilpatrick, M.W., Wei, C.-F., Gray, H.B., Jr. and Wells, R.D. (1983) *Nucleic Acid Res.* 11, 3811-3822.
16. Hentschel, C.C. (1982) *Nature* 295, 714-716.
17. Nickol, J.M. and Felsenfeld, G. (1983) *Cell* 35, 467-477.
18. Schon, E., Evans, T., Welsh, J. and Efstratiadis, A. (1983) *Cell* 35, 837-848.
19. Shen, C.-K. J. (1983) *Nucleic Acid Res.* 11, 7899-7910.
20. Dybvig, K., Clark, C.D., Aliperti, G. and Schlesinger, M.J. (1983) *Nucleic Acid Res.* 11, 8495-8508.
21. Mace, H.A.F., Pelham, H.R.B. and Travers, A.A. (1983) *Nature* 304, 555-557.
22. Kowalski, D., Kroeker, W.D. and Laskowski, M., Sr. (1976) *Biochemistry* 15, 4457-4463.
23. Kowalski, D. (1979) *Anal. Biochem.* 93, 346-354.
24. Woodworth-Gutai, M. (1981) *Virology* 109, 344-352.
25. Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560-564.
26. Urquiza, I., Venegas, A. and Yudelevich, H. (1982) *Arch. Biol. Med. Exp.* 15, 407-416.
27. Espejo, R.T., Canelo, E.S. and Sinsheimer, R.L. (1969) *Proc. Natl. Acad. Sci. USA* 63, 1164-1168.
28. Kroeker, W.D. and Kowalski, D. (1978) *Biochemistry* 17, 3236-3243.
29. Brack, C., Bickle, T.A. and Yuan, R. (1975) *J. Mol. Biol.* 96, 693-702.
30. Peck, L.J. and Wang, J.C. (1981) *Nature* 292, 375-378.
31. Rhodes, D. and Klug, A. (1981) *Nature* 292, 378-380.
32. Strauss, F., Gaillard, C. and Prunell, A. (1981) *Eur. J. Biochem.* 118, 215-222.
33. Wells, R.D., Larsen, J.E., Grant, R.C., Shortle, B.E. and Cantor, C.R. (1970) *J. Mol. Biol.* 54, 465-497.
34. Lim, V.I. and Mazanov, A.L. (1978) *FEBS Letters* 88, 118-123.
35. Streeck, R.E. and Gebhart, C. (1979) *Hoppe Seyler's Z. Physiol. Chem.* 360, 529-539.
36. Goding, C.R. and Russell, W.C. (1983) *Nucleic Acid Res.* 11, 21-36.

37. Larsen, A. and Weintraub, H. (1982) *Cell* 29, 609-622.
38. Weintraub, H. (1983) *Cell* 32, 1191-1203.
39. Cockerill, P.N. and Goodwin, G.H. (1983) *Biochem. Biophys. Res. Comm.* 112, 547-554.
40. Kohwi-Shigematsu, T., Gelinas, R. and Weintraub, H. (1983) *Proc. Natl. Acad. Sci. USA* 80, 4389-4393.
41. West, R.W., Jr. and Rodriguez, R.L. (1980) *Gene* 9, 175-193.
42. Strauss, H.S., Burgess, R.R. and Record, M.T., Jr. (1980) *Biochemistry* 19, 3496-3504.
43. Kupper, H., Contreras, R., Khorana, H.G. and Landy, A. (1976) in *RNA Polymerase*, Lasids, R. and Chamberlin, M., Eds., pp.473-484, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
44. Hale, P., Woodward, R.S. and Lebowitz, J. (1983) *J. Biol. Chem.* 258, 7828-7839.
45. Brack, C., Eberle, H., Bickle, T.A. and Yuan, R. (1976) *J. Mol. Biol.* 104, 305-309.
46. Parker, R.C., Watson, R.M. and Vinograd, J. (1977) *Proc. Natl. Acad. Sci. USA* 74, 851-855.
47. Tapper, D.P. and Clayton, D.A. (1981) *Nucleic Acid Res.* 9, 6787-6794.
48. Hsieh, T.-S. and Wang, J.C. (1975) *Biochemistry* 14, 527-535.
49. Depew, R.E. and Wang, J.C. (1975) *Proc. Natl. Acad. Sci. USA* 72, 4275-4279.
50. Pulleyblank, D.E., Shure, M., Tang, D., Vinograd, J. and Vosberg, H.-P. (1975) *Proc. Natl. Acad. Sci. USA* 72, 4280-4284.