A human DNA-binding protein is methylation-specific and sequence-specific

Richard Y.-H. Wang, Xian-Yang Zhang and Melanie Ehrlich

Department of Biochemistry, Tulane Medical School, New Orleans, LA 70112, USA

Received 17 December 1985; Accepted 22 January 1986

ABSTRACT

A nuclear protein isolated from human placenta, methylated DNA-binding protein (MDBP), binds selectively to DNA enriched in 5-methylcytosine. We now demonstrate that MDBP is a sequence-specific, as well as methylation-specific, DNA-binding protein. From ten restriction fragments of pBR322 DNA methylated with human DNA methyltransferase, one was bound to MDBP very much more strongly than any of the others. For this preferential binding to MDBP, the DNA had to be methylated. By a DNase I protection experiment (DNase I footprinting), a 22-base sequence within this methylated restriction fragment was shown to be specifically protected by MDBP. The sequence-specificity of MDBP coupled with its dependence on DNA methylation suggests that this is one of the proteins which modulates important functions of human DNA methylation in vivo.

INTRODUCTION

5-Methylcytosine (m^5C) appears as a minor base in the DNA of many organisms including all vertebrates studied (1,2). Methylation of ~3 to 10% of the cytosine residues in vertebrate DNA occurs mostly or exclusively at CpG sequences. The role of DNA methylation in vertebrate development is poorly understood although there are large tissue-specific differences in the extent of DNA methylation and much evidence suggests that DNA methylation negatively controls transcription for a number of genes (1-8). Besides regulating transcription of certain genes, DNA methylation might play other roles, such as in modulating DNA replication, genetic rearrangements, DNA repair, or in the packing of DNA into chromosomes (1,7).

Given the evidence that 5-methylation at pyrimidine residues often has profound effects upon recognition of DNA by sequence-specific, DNA-binding proteins from prokaryotes (9-12) and the relatively exposed position of this methyl group in the major groove of the standard B-form DNA helix (13), it is very likely that some vertebrate DNA-binding proteins could strongly differentiate methylated from the analogous unmethylated DNA sequences. To help define the roles of DNA methylation in vertebrate development, we have isolated a protein from human placental nuclei that binds preferentially to m^5C -enriched

DNA sequences (14). We now report that this protein, which we refer to as methylated DNA-binding protein (MDBP), has sequence specificity in addition to its being specific for methylated DNA. Such a protein is a prime candidate for a regulatory macromolecule that mediates effects of DNA methylation on vertebrate cellular functions.

MATERIALS AND METHODS

MDBP and DNA methyltransferase. The procedure used for isolating MDBP and DNA methyltransferase from human placental nuclei involved previously described methods with some modifications (14,15). Two full-term placenta minus the chorionic plate were minced and homogenized for 10 sec at full speed in a Waring Blender and then in a Potter-Elvehjem homogenizer with six strokes at 3/4 maximum speed using 3-4 volumes of 50 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, 0.25 M sucrose, 0.1 mM phenylmethylsulfonyl fluoride, 1 μ g/ml chymostatin (Sigma). After a further 1:1 dilution with this buffer, the suspension was filtered through two layers of gauze, then through four layers, and lastly through a nylon membrane (HD3-70; Tetko, Depew, N.Y.). The filtrate was centrifuged and the crude nuclear pellet was rinsed, extracted with 0.3 M NaCl, and the extract centrifuged as previously described (14,15). The supernatant was diluted with 2.5 volumes of 25 mM potassium phosphate, pH 7.4, 0.1 mM EDTA, 0.2 mM dithiothreitol (DTT), 10% glycerol (buffer A) and applied at a flow rate of 80 ml/h to a phosphocellulose column (5 x 5 cm), which had been preequilibrated with buffer A plus 0.1 M KCl. The fractions that eluted at 0.12 - 0.24 M KCl were pooled for further purification of MDBP; those at 0.25 - 0.36 M KCl were used for isolation of DNA methyltransferase. The pooled phosphocellulose chromatography fractions were dialyzed as previously described (14). Upon dialysis, a precipitate formed which was subsequently removed by centrifugation. After column chromatography of the former fraction on hydroxyapatite and then on DEAE-cellulose (14), the peak MDBP fractions were stored in the presence of 200 μ g/ml bovine serum albumin (BSA; USBC, RIA Grade) at $-70^{\circ}C$. This partially purified MDBP did not lose its sequence-specific binding upon freezing and thawing up to five times. Unless otherwise specified, the MDBP preparations used for this study were from the DEAE-cellulose fraction.

For partial purification of the DNA methyltransferase, hydroxyapatite column chromatography was performed as previously described (15) and the active fractions were pooled and mixed with BSA at a concentration of 50 μ g/ml. After dialyzing against 50 mM Tris-HCl, pH 7.6, 0.1 mM EDTA, 0.2 mM DTT, 25 mM NaCl, 10% glycerol (buffer B), the protein was loaded onto a DEAE-Sepharose CL-6B

column (1.6 x 5 cm). The column was eluted with a linear gradient of 25 - 500 mM NaCl in buffer B. Before elution, an aliquot of BSA was added to each collection tube so that there would be a final concentration of 100 µg of BSA per ml in each fraction. The most active fractions eluted from the column were pooled, concentrated and desalted five fold in an ultrafiltration device (Centricon-30; Amicon) and stored at -70°C.

Other enzymes used in this study such as restriction endonucleases, Escherichia coli DNA polymerase, and T4 DNA ligase were obtained from IBI or Bethesda Research Laboratories unless otherwise specified.

Analysis of relative extent of retention of methylated DNA fragments by MDBP. Methylated DNA was prepared by incubating 1 µg of DNA with 360 maintenance-type units (15) of human DNA methyltransferase in 0.1 ml of 50 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.5 mM DTT, 100 µg/ml BSA, 10% glycerol, and 21 µM [³H-methyl]S-adenosylmethionine (Adomet; 0.48 Ci/mmol) at 37°C for 16 h. [³H]-AdoMet was used to allow quantitation of the extent of methylation from the amount of ³H-methyl group converted to acid-precipitable material. The mixtures were then incubated with 5 µg of proteinase K at 37°C for 3 h and extracted with organic solvents. The DNA was precipitated with ethanol, digested with 10 units of HinfI or TaqI, and the DNA fragments labeled at their 3' ends with [α-³²P]-dATP or dGTP (800 Ci/mmol) using the large fragment of E. coli DNA polymerase I (16) under standard conditions. The restricted DNA was incubated with MDBP for 15 min at room temperature and passed through nitrocellulose filters as previously (14) except that the binding buffer was 20 mM Tris-HCl, pH 7.6, 40 mM NaCl, 0.2 mM EDTA, 3 mM MgCl₂, 20 µg/ml BSA. DNA from DNA·MDBP complexes retained by nitrocellulose was eluted with 0.4 ml of 20 mM Tris-HCl, pH 8.0, 0.3 M sodium acetate, 0.2% sodium dodecyl sulfate, 5 mM EDTA, and 40 µg/ml E. coli tRNA at 37°C for 2 h, precipitated with ethanol, and subjected to gel electrophoresis followed by autoradiography.

Analysis of DNA substituted with m⁵C residues by DNA synthesis was performed similarly. In this case, the DNA was radiolabeled with [α-³²P]dATP during extension of a pentadecamer primer (Sequencing Primer; New England BioLabs) annealed to single-stranded recombinant M13mpl9 DNA (17,18) in a reaction catalyzed by E. coli DNA polymerase I. The primer extension in the presence of m⁵dCTP, dGTP, dATP, and dTTP and ligation to form covalently closed circular (CCC) DNA were performed as previously described (19). These CCC molecules were purified by exposure to pH 12.3, neutralization, filtration through a nitrocellulose membrane and then through a Centricon CM-30 ultrafilter in the presence of 30 µg of tRNA (19).

Fig. 1. A partial map of the RF from recombinants MpB-A1, MpB-A2, and MpB-A3. The inserts in MpB-A2 and MpB-A3 are identical and contain 238 bp (from pBR322 positions 1020 to 1258) in opposite orientations. The pBR322-derived inserts are shown by hatched boxes with the arrow pointing in the direction of ascending numbers of the positions on the pBR322 map (20). The M13mp19 vector sequences are indicated by a heavy black line. The numbers above the maps indicate positions of vector sequences on the standard M13mp19 vector map (17,18). Positions of the *TaqI* recognition sites are shown by the letter T. The insert in MpB-A1 is longer, 264 bp (from pBR322 positions 1020 to 1284), and includes an extra *TaqI* site giving rise to a 142-bp *TaqI* fragment containing the strong MDBP-binding site. The latter is designated by a black rectangular box under the letter M. The sizes and end points of the *TaqI* fragments containing this sequence are shown below each map. As explained in the text, partial *dam* methylation of the central *TaqI* site in viral strand of the insert in MpB-A3 is apparently responsible for the 839-bp fragment in the *TaqI* digest of the corresponding m⁵dCMP-substituted RF.

Cloning of an MDBP-specific fragment from pBR322 in M13mp19. After digestion of pBR322 with *HinfI*, the 298-base pair (bp) fragment from position 1006-1304 (20) was isolated by polyacrylamide gel electrophoresis. It was further digested (at sites 1020, 1258, and 1284) by *HpaII*. The unpurified subfragments and alkaline phosphatase-treated, *AccI*-linearized M13mp19 replicative form (RF) were incubated with T4 DNA ligase and then used to transfect JM101 (17,21). From these transfectants three recombinants were isolated, MpB-A1, MpB-A2, and MpB-A3 (Fig. 1). The insert of MpB-A1 was longer than those of MpB-A2 and MpB-A3 because of incomplete digestion by *HpaII*.

DNase I protection experiments. DNase I footprint analysis (22) of the strong MDBP binding site in the 298-bp *HinfI* fragment of pBR322 was performed as follows. This isolated fragment (1 µg) was incubated with human DNA methyltransferase as described above, extracted with organic solvents, precipitated with ethanol, redissolved and then incubated once again with DNA methyltrans-

ferase. It was radiolabeled at its 3' ends as previously described and digested with 15 units of AluI at 37°C for 90 min under standard conditions. The 214-bp subfragment was obtained by electrophoresis on a 9% polyacrylamide gel, electroelution, extraction with organic solvents, and ethanol precipitation. Then, 25 ng of this fragment was incubated at 25°C for 15 min with 40 ng of unlabeled MboI-digested Micrococcus luteus DNA, 20 mM Tris-HCl, pH 7.6, 0.2 mM EDTA, 40 mM NaCl, 1.6 mM MgCl₂, 1.6 mM CaCl₂, 57 µg/ml BSA and 8 units of MDBP (14) in a volume of 260 µl. Upon incubation with 4 ng of DNase I (electrophoresis grade; Worthington) at 25°C for 90 sec, followed by addition of EDTA, pH 8, and sodium acetate to concentrations of 25 mM and 0.3 M, respectively, and of 5 µg of E. coli tRNA, the DNA was precipitated with ethanol and electrophoresed (22,23).

RESULTS

Binding of MDBP to pBR322 fragments methylated with human DNA methyltransferase

To test whether MDBP binds only to certain m⁵C-containing DNA sequences, we methylated a small DNA molecule of known sequence, pBR322, with human placental DNA methyltransferase (15) to convert 5.2% of the bases to m⁵C (~70% of CpG sites methylated, see below). After methylation, cleavage with HinfI, and radiolabeling at their 3' ends, the DNA fragments were incubated with MDBP and filtered through a nitrocellulose membrane. The retained fragments (DNA·MDBP complexes) were eluted and then analyzed by agarose gel electrophoresis followed by autoradiography. As seen in lane 3 of Figure 2, one methylated fragment showed highly preferential retention by MDBP (24). This specific binding was demonstrated in phosphate buffer at pH 7.0 or 7.3 as well as under our standard conditions in Tris buffer at pH 7.6. It was observed with methylated DNA fragments (Fig. 2, lane 3) but not with fragments of mock-methylated DNA (Fig. 2, lane 4), which had been prepared by omission of AdoMet from the methylation mixture. Only non-specific DNA-binding was seen when a heterogeneous fraction of DNA-binding proteins not containing MDBP was incubated with the methylated DNA fragments (Fig. 2, lane 2). This latter fraction had eluted at 0.2 M potassium phosphate, pH 7.0, from the hydroxyapatite step in the partial purification of MDBP (14) and was devoid of detectable MDBP activity.

To look for this MDBP-specific site in pBR322 fragmented with a different enzyme, we digested the same preparation of methylated pBR322 with TaqI instead of HinfI. Radiolabeling, nitrocellulose membrane retention, and electrophoresis as above revealed the presence of only one fragment binding appreciably to MDBP (Fig. 3). As expected, the location of this 616-bp fragment on the standard

Fig. 2. Retention of methylated pBR322 HinfI fragments by MDBP. Human DNA methyltransferase-methylated pBR322 was digested with HinfI and end-labeled with [32 P]dATP as described in Materials and Methods. This DNA had $\sim 70\%$ of its CpG sites methylated as determined by incorporation of methyl groups from [3 H]AdoMet and confirmed by digestion with HpaII and HhaI (1). Approximately 10 ng ($\sim 10^5$ cpm) of the DNA fragments was incubated with approximately 1 unit of MDBP from the DEAE-cellulose fraction (14) and then the nitrocellulose membrane-bound DNA subjected to electrophoresis in a 2% agarose gel as described in Materials and Methods. Two of the bands seen in this autoradiogram were doublets (506/517 bp; 220/221 bp) and the smallest fragment (75-bp) could not be examined in this gel but was shown not to bind preferentially to MDBP on a polyacrylamide gel (data not shown). Half of each sample of MDBP-retained DNA fragments and one-fifth of each sample containing nonspecific protein-retained DNA fragments were electrophoresed. Lanes 1-3, methylated DNA; lanes 4-6, mock-methylated DNA. Lanes 1 & 6, ~ 0.5 ng input DNA without incubation with proteins or passage through a nitrocellulose filter; lanes 2 & 5, DNA fragments retained by non-specific DNA-binding proteins; lanes 3 & 4, fragments retained by MDBP. In this figure and in Figs. 3-6, lanes with methylated DNA fragments retained by MDBP in the absence of competing methylated DNA are designated with a dot on top.

pBR322 map (position 652-1268) overlaps that of the 298-bp HinfI fragment (position 1006-1304; 20).

We have similarly analyzed a derivative of pBR322 containing the entire SV40 genome (pM14; J.C. Cohen, unpubl. results) and another containing the human

Fig. 3. Retention of methylated pBR322 *TaqI* fragments by MDBP. Using the methods described for Fig. 2, the retention of DNA methyltransferase-methylated pBR322 [^{32}P]fragments ($\sim 10^4$ cpm/ng) by MDBP was determined except that, in this experiment, *TaqI* fragments were used and electrophoresis was in a 5% polyacrylamide gel. The appearance of a 616-bp fragment instead of 475- and 141-bp fragments seems to be due to *dam* (Gm^6ATC) methylation at a TCGATC site at position 1127; this methylation inhibits *TaqI* digestion at the TCGA sequence (12). Lanes 1-5, methylated DNA; lanes 6-10, mock-methylated DNA. Lanes 1 & 10, input DNA fragments neither incubated with proteins nor passed through a nitrocellulose filter; lanes 2 & 9, DNA fragments retained by non-specific DNA-binding proteins. Lanes 3-5, DNA fragments retained by each of three different fractions from the DEAE-cellulose chromatography peak of MDBP; lanes 6-8, analogous to lanes 3-5 except with unmethylated DNA.

α -globin gene (25). The latter insert is particularly rich in CpG sites at the 5' end of the gene (26). These DNAs were methylated with human DNA methyltransferase so that ~ 45 -50% of their CpG sites were methylated (15 and see below). Both DNAs were digested with *TaqI* and then labeled at their 3' ends. From the digest of methylated pM14 and that of the α -globin recombinant only one fragment, the 616-bp fragment from the pBR322 vector sequences, showed preferential retention by MDBP.

Binding of MDBP to hemimethylated pBR322 DNA sequences prepared by incorporation of m^5dCMP

Our identification and preliminary analysis of MDBP (14) had relied upon the use of DNA extensively substituted with m^5C residues rather than methylated just at CpG sites. In comparisons to m^5C -deficient DNAs, we had used the

following three m^5C -rich ligands displaying preferential binding of MDBP (14): the naturally m^5C -rich bacteriophage XP12 DNA (34 mol% m^5C); human DNA extensively nick-translated in the presence of m^5dCTP ; and a hemimethylated M13 RF derivative prepared by primer extension in the presence of m^5dCTP to allow incorporation of m^5dCMP into the minus strand. In order to prepare a similarly m^5C -enriched form of the MDBP-specific pBR322 site (Figs. 2 & 3), we subcloned into M13mpl9 a subfragment of the 298-bp HinfI fragment of pBR322. Three recombinant DNAs were isolated, namely MpB-A1, MpB-A2, and MpB-A3 (Fig. 1) and were then prepared as the hemimethylated RFs with all the C residues of the complementary strand replaced by m^5C residues. The first two recombinants have the same orientation of their pBR322 insert but have different sized inserts (Fig. 1). These hemimethylated DNAs were fragmented with TaqI, one of the few restriction enzymes known to be able to digest m^5C -rich DNA (11). Both MpB-A1 and MpB-A2 RF DNAs exhibited preferential retention by MDBP of their hemimethylated TaqI fragments containing part of the pBR322 insert (the left half of their inserts as depicted in Fig. 1). These MDBP-specific fragments contained 142 or 701 bp of pBR322 sequence or pBR322 plus M13mpl9 sequences (Figs. 1 & 4). Both of these fragments showed no preferential retention by MDBP when they were unmethylated (Fig. 4 and data not shown).

The sequences in the pBR322 insert not included in the above TaqI fragments should have appeared in 585-bp (MpB-A1) and 136-bp (MpB-A1 and MpB-A2) TaqI fragments (Fig. 1). Neither of these was preferentially retained by MDBP (Fig. 4), a finding consistent with there being only one specific MDBP site in the pBR322 insert. However, also a vector sequence, a 381-bp TaqI fragment from M13mpl9 positions 1126-1507 (18) was preferentially retained in both of these hemimethylated recombinants and in the unrecombined hemimethylated vector (Fig. 4). In contrast, when these recombinants had been methylated with human DNA methyltransferase instead of by primer extension with m^5dCTP , the 381-bp M13mpl9 TaqI fragment was bound to MDBP only poorly although the pBR322 insert was bound very well (data not shown). Differential MDBP binding of the hemimethylated and the methyltransferase-methylated 381-bp M13mpl9 fragment may be due to differences in binding affinities of MDBP sites depending upon whether the site is methylated at all C residues in one strand (as in the hemimethylated DNA) or is methylated at only CpG sites (as in the bifilarly methyltransferase-methylated DNA).

That the same protein in our impure MDBP preparations was binding to the methyltransferase-methylated pBR322 sequence, the hemimethylated pBR322 sequence, and to the hemimethylated M13mpl9 381-bp fragment was shown by mixing

Fig. 4. Retention of hemimethylated MpB-A1 and MpB-A2 *TaqI* fragments by MDBP. Hemimethylated RF molecules were produced by incorporation of m^5dCMP throughout the complementary strand as described in *Materials and Methods*. As in Fig. 3, *TaqI* [^{32}P] fragments retained by MDBP were analyzed by gel electrophoresis (5% polyacrylamide) and autoradiography. However, in this experiment, the hydroxyapatite fraction of MDBP (14) was used. In previous experiments no unlabeled, nonmethylated DNA was present; however, in this experiment the MDBP was preincubated with 20 ng of unlabeled *HinfI*-digested λ DNA before addition of 20 ng of the *TaqI* fragments of the RF ($\sim 10^4$ cpm/ng) and after addition of the RF fragments was further incubated for 15 min. Lanes 1 & 6, input (unincubated and unfiltered) hemimethylated DNA (~ 0.2 ng) from M13mp19 RF (lane 1) or MpB-A1 RF (lane 6). In these and other experiments with hemimethylated DNA, the fragments had been uniformly labeled; therefore, their relative intensity in the autoradiograms is proportional to their size as well as to their affinity for MDBP. Lanes 2, 3, & 5, *TaqI* fragments of hemimethylated MpB-A2, hemimethylated MpB-A1, and unmethylated MpB-A1, respectively, retained by MDBP. Lane 4, fragments retained by MDBP from *TaqI*-digested hemimethylated M13mp19 RF. The 701-bp fragment containing pBR322 insert sequences co-electrophoreses with a 703-bp M13mp19 vector fragment which showed no preferential binding to MDBP (lane 4).

experiments. The MDBP-specific, methyltransferase-methylated, 298-bp *HinfI* fragment of pBR322 competed for MDBP binding with both the cloned hemimethylated 142-bp and 701-bp *TaqI* fragments containing homologous pBR322 sequences (Fig. 5, lanes 3 & 5). Also, this methyltransferase-methylated *HinfI* fragment competed for binding to MDBP with the hemimethylated 381-bp *TaqI* fragment containing only

Fig. 5. Competition of human DNA methyltransferase-methylated 298-bp *HinfI* fragment of pBR322 with *TaqI* fragments of hemimethylated MpB-A1 and MpB-A2 RF molecules for retention by MDBP. The same preparation of MDBP and of uniformly labeled *TaqI* fragments of MpB-A1 and MpB-A2 substituted with m^5C throughout the minus strand were used and incubated together after preincubation with λ DNA as in Fig. 4. However, in this experiment the radioactively labeled *TaqI* fragments (20 ng) were incubated with MDBP in the presence of ~ 20 ng of the unlabeled, nonmethylated 298-bp *HinfI* pBR322 fragment or with this 298-bp fragment that had been methylated with human DNA methyltransferase. The DNA methyltransferase-treated *HinfI* fragment had $\sim 60\%$ of its CpG sites methylated as determined by the incorporation of the methyl group from $[^3H]AdoMet$. Retention of $[^{32}P]$ DNA fragments by MDBP was determined as in Fig. 4. Lanes 1 & 7, input *TaqI* fragments of hemimethylated MpB-A2 and MpB-A1, respectively, without incubation with proteins or passage through nitrocellulose. Lanes 2 & 4, *TaqI*-digested hemimethylated MpB-A2 or MpB-A1 $[^{32}P]$ RF as in lanes 2 & 3, respectively, of Fig. 4 but with the addition of the unlabeled nonmethylated *HinfI* fragment. Lanes 3 & 5, same as lanes 2 & 4, respectively, in this figure but with the methylated *HinfI* fragment as the unlabeled competitor instead of the nonmethylated fragment. Lane 6, *TaqI*-digested hemimethylated M13mpl9 RF as in Fig. 4, lane 4, except that, in addition, ~ 20 ng of the methylated *HinfI* fragment was present during the incubation with MDBP.

M13mpl9 sequences (Fig. 5, lanes 3, 5 & 6). Furthermore, the analogous non-methylated *HinfI* fragment of pBR322 did not compete (Fig. 5, lanes 2 & 4). Although it is not observable under the electrophoresis conditions used for these figures, another larger *TaqI* fragment from the hemimethylated M13mpl9 vector (a 927-bp fragment) also binds preferentially to MDBP (24; unpublished

Fig. 6. Retention of hemimethylated MpB-A3 TaqI fragments by MDBP. Uniformly ^{32}P -labeled TaqI fragments of hemimethylated MpB-A3 RF molecules were prepared and tested for retention by a hydroxyapatite fraction of MDBP in the presence or absence of the DNA methyltransferase-methylated 298-bp HinfI fragment from pBR322 as described for Fig. 5. Lanes 1-4, TaqI-digested hemimethylated [^{32}P]DNA; lane 5, analogous nonmethylated DNA. Lane 1, input DNA neither incubated with protein nor passed through nitrocellulose. Lanes 2-5, ^{32}P -labeled TaqI fragments retained by MDBP in the presence of the unlabeled, methylated HinfI fragment (lane 3) or the analogous nonmethylated HinfI fragment (lane 4), or in the absence of unlabeled DNA (lanes 2 & 5). The 839-bp fragment containing pBR322 insert sequences co-electrophoreses with a 790-bp M13mp19 vector fragment which showed no preferential binding to MDBP (Fig. 4, lane 4, the band above the 701-bp fragment band).

results) and is competed by the methyltransferase-methylated 298-bp HinfI fragment from pBR322 (data not shown).

In pBR322-M13mp19 recombinants MpB-A1 and MpB-A2, the pBR322 insert is in the same orientation (Fig. 1). Therefore both hemimethylated RF molecules derived from them had the same strand methylated during primer extension of the recombinant viral DNA in the presence of m^5dCTP . To test whether the MDBP-specific binding occurs when the other strand of the pBR322 insert is methylated, recombinant MpB-A3 with the opposite orientation of the insert (Fig. 1) was hemimethylated by primer extension as above. Analysis of retention by MDBP of different TaqI fragments from this RF revealed that methylation of the other

strand of the pBR322 insert also gave preferential retention of this region of pBR322 DNA. In this recombinant, complete TaqI digestion should have given only a 162-bp fragment containing the same MDBP-specific pBR322 sequences in the 701-bp vector/insert fragment from MpB-A2 (Figs. 1 & 4). The 162-bp fragment was present and selectively bound to MDBP in the TaqI digest of hemimethylated MpB-A3 (Fig. 6). In addition an overlapping (Fig. 1) 839-bp fragment was in this digest and was specifically retained by MDBP (Fig. 6). This fragment was probably due to partial digestion at the TaqI (TCGA) site in this insert. Incomplete digestion could be explained by partial methylation of A residues at the overlapping dam methylation (GATC) site since TCGm⁶A sequences are resistant to TaqI (12,27). As in the cases of MpB-A1 and MpB-A2 (Figs. 4 & 5), MpB-A3 showed preferential retention of DNA fragments only when enriched in m⁵C residues (Fig. 6, lanes 2 & 5). Similarly, the human DNA methyltransferase-methylated 298-bp HinfI fragment from pBR322 competed for the specific binding by the 839-bp, 381-bp, and 162-bp TaqI fragments of hemimethylated MpB-A3 (Fig. 6, lane 3) but the nonmethylated HinfI fragment did not compete (Fig. 6, lane 4).

Localizing the specific MDBP binding site in pBR322

To further analyze the MDBP binding site(s) on the 298-bp fragment from HinfI-digested pBR322, this fragment was purified, methylated with human DNA methyltransferase (15), and radiolabeled at its 3' ends. Digestion with AluI released an end-labeled 214-bp DNA subfragment that was assayed for MDBP binding in the standard nitrocellulose filter-binding assay. Approximately 60 times more of this methylated fragment than of the mock-methylated fragment was bound by MDBP in equivalent incubations under the conditions specified for Fig. 7, thereby confirming the presence of at least one strong binding site for MDBP in this region of pBR322.

To locate this binding site, the 214-bp subfragment was used for DNase I footprinting (22). As shown in Fig. 7, MDBP protected just one region of approximately 22 bases from digestion by DNase I. This protein-mediated protection was observed in the methylated but not in the unmethylated DNA fragment (Fig. 7). The rest of the 214-bp fragment was found to be devoid of specific MDBP binding sites in another DNase I footprinting experiment (data not shown).

This 214-bp subfragment used for DNase I footprinting was also analyzed by chemical sequencing methods that allow m⁵C sites to be distinguished from the analogous C sites by the presence of weaker bands from the former due to a much lower reactivity with hydrazine upon cytosine 5-methylation (28). At all of the C residues of the 19 CpG sites which were visualized in the methylated 214-bp subfragment, much weaker bands were seen than in the corresponding unmethylated fragment (Fig. 7B and data not shown). Furthermore, the only C bands demon-

Fig. 7. DNase I footprint analysis of the strong binding site for MDBP in pBR322 DNA. The isolated 214-bp AluI subfragment from pBR322 was end-labeled and used for DNase I footprinting as described in Materials and Methods. Electrophoresis was performed on an 8% polyacrylamide gel followed by autoradiography. Panel A: DNase I digests. Lane 1 & 3, mock-methylated DNA; lanes 2 & 4, methylated DNA. Lanes 1 & 2, no MDBP; lanes 3 & 4, MDBP present during the DNase I digestion. The dots on the left side of Panel A are adjacent to the bands which are more prominent in the methylated DNA than in the unmethylated DNA treated with DNase I in the absence of MDBP. Panel B: Left four lanes, Maxam-Gilbert type of sequencing reactions with cleavage at G & A, G, or C residues; for the C-specific reaction, methylated (m) or mock-methylated (u) DNA was used. Right four lanes, DNase I digests. Lanes 1 & 3, methylated DNA; lanes 2 & 4, mock-methylated DNA. Lanes 1 & 2, MDBP present during the DNase I digestion; lane 3 & 4, no MDBP. The dashes on the left side of the C tract in Panel B indicate the position of C residues which react much more weakly in methylated than in the unmethylated DNA (28). The brackets in Panels A and B indicate the region of methylated DNA protected by MDBP against DNase I digestion. Different MDBP preparations from DEAE-cellulose peak fractions (14) were used in panels A and B.

strating this weakened reaction upon methylation of the subfragment were those with a G residue as the 3' neighbor. Therefore, methylation by our preparation of human DNA methyltransferase was extensive and essentially confined to CpG sites.

At a few isolated sites even in the absence of MDBP, methylation increased the DNase I sensitivity of the DNA so that the corresponding bands became more prominent (Fig. 7A, lanes 1 and 2). Despite the effect of methylation alone on the cleavage by DNase, the 22-base region of methylated DNA protected by MDBP is clearly seen in a comparison of methylated DNA treated with DNase I in the presence or absence of MDBP (Fig. 7A, lanes 2 and 4; Fig. 7B, lanes 1 and 3). Furthermore, binding of MDBP resulted in enhanced cleavage at the 3' end of the protected region (Fig. 6). Locally enhanced cleavage by DNase I in the region of the protected site has been reported for several other sequence-specific DNA-binding proteins (30,31).

DISCUSSION

With recent advances in methodology, an increasing number of sequence-specific DNA-binding proteins are being isolated from vertebrate cells (32-34). In this report, we presented the first known example in vertebrates of a sequence-specific, DNA-binding protein which binds selectively to its corresponding DNA sequence only when that sequence is methylated. We determined that this protein, MDBP, is sequence-specific by comparing its ability to bind to various in vitro-methylated DNA fragments from HinfI or TaqI digests of methylated pBR322. This in vitro methylation was catalyzed by human DNA methyltransferase which methylated only or predominantly at CpG sites (Fig. 7; 15) as is the case in vivo for vertebrate DNA (1). A prokaryotic plasmid DNA was chosen for this study because it has a low sequence complexity and its DNA sequence is completely known (20). Furthermore, unlike vertebrate DNAs, it does not contain a marked deficiency in CpG sequences (35). One fragment of DNA methylated with human DNA methyltransferase was bound much more strongly than any of the others by MDBP (Figs. 2 and 3). In contrast, analogous unmethylated pBR322 DNA fragments were bound by our preparation of MDBP only poorly and with no preferential retention of certain fragments (Figs. 2 and 3). Therefore, sequence-specific binding was only observed in the methylated DNA.

This binding site for MDBP on methylated pBR322 was localized approximately to residues 1141-1162 by a DNase I protection experiment using a 214-bp subfragment. This was the only region of the 214-bp Alu I subfragment that was protected by MDBP (Fig. 7). As for certain other sequence-specific DNA-binding proteins (29,36,37), the region protected against DNase I hydrolysis by the

DNA-binding protein may be larger than the recognition site of that protein. Therefore, we hypothesize that the 22-base region in pBR322 specifically protected by MDBP, namely, 5'-CTGATm⁵CGTCAm⁵CGGm⁵CGATTATG-3' (written in its fully CpG-methylated form) contains within it a subset of sequences which constitute a highly specific MDBP recognition site. In the central region of the MDBP-protected site is the pallindromic sequence m⁵CGGm⁵CG. In addition, there is another m⁵CpG dinucleotide 5' to this sequence. Nonetheless, methylated CpG dinucleotides in close proximity are clearly not sufficient for recognition by MDBP because they occur many times in methylated pBR322 fragments and in the methylated cloned human α -globin gene, which are not preferentially retained by MDBP (Figs. 2 & 3 and data not shown). Experiments are in progress to determine how many of the three m⁵CpG sites of the pBR322 sequence protected by MDBP are necessary for specific binding.

It is interesting to note that the replacement of all the C residues on one or the other strand of this 22-bp region of pBR322 with m⁵C residues leads to the same type of preferential binding by MDBP (Figs. 4-6). This was determined with hemimethylated DNA fragments prepared by incorporation of m⁵dCMP into the complementary strand of recombinants of M13mp19 containing a pBR322 insert (Fig. 1). In this experiment, we may have created in the M13mp19 sequences additional MDBP binding sites with m⁵C residues which are specifically recognized by MDBP at positions other than m⁵CpG. Such methylated sequences may not occur in vivo because vertebrate methylation is largely confined to CpG sites (1). Also, where C residues are substituted for T residues in a variant of an MDBP recognition site, for example, one containing ApC instead of ApT, the replacement of C residues with m⁵C might increase binding by MDBP since m⁵C and T residues contain the same 5-methyl group. One or both factors may pertain to the MDBP recognition site in the 381-bp TaqI fragment from the hemimethylated M13mp19 RF (Fig. 4). Additional DNase I footprinting experiments will permit examination of several regions in this hemimethylated fragment which show partial homology to the CpG-methylated MDBP-binding site of pBR322 (Fig. 7). To elucidate the role of MDBP in modulating the effects of DNA methylation on cellular functions, it will be important to determine the consensus sequence for MDBP recognition sites, their frequency, approximate distribution, and state of methylation in vertebrate DNA.

ACKNOWLEDGEMENTS

We thank Drs. J. Craig Cohen and Arthur W. Nienhuis for the SV40-pBR322 and the α -globin-pBR322 recombinant DNAs, respectively and Dr. Stuart Linn for

insightful comments during review of the manuscript. This research was supported in part by USPHS grant CA-19942.

REFERENCES

1. Ehrlich, M. and Wang, R. Y.-H. (1981) *Science* 212, 1350-1357.
2. Gama-Sosa, M.A., Midgett, R.M., Slagel, V.A., Githens, S., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983) *Biochim. Biophys. Acta* 740, 212-219.
3. Cooper, D.N. (1983) *Hum. Genet.* 64, 315-333.
4. Ehrlich, M., Gama-Sosa, M.A., Huang, L.-H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C.W. (1982) *Nucleic Acids Res.* 10, 2709-2721.
5. Gama-Sosa, M.A., Wang, R.Y.-H., Kuo, K.C., Gehrke, C.W. and Ehrlich, M. (1983) *Nucleic Acids Res.* 11, 3087-3095.
6. Doerfler, W. (1983) *Ann. Rev. Biochem.* 52, 93-124.
7. Jaenisch, R. and Jahner, D. (1984) *Biochim. Biophys. Acta* 782, 1-9.
8. Christman, J.K. (1984) *Curr. Top Microbiol. Immunol.* 108, 49-78.
9. Fisher, E.F. and Caruthers, M.H. (1979) *Nucleic Acids Res.* 7, 401-416.
10. Hofer, B. and Koster, H. (1980) *Nucleic Acids Res.* 8, 6143-6162.
11. Huang, L.-H., Farnet, C.M., Ehrlich, K.C., and Ehrlich, M. (1982) *Nucleic Acids Res.* 10, 1579-1591.
12. McClelland, M. *Nucleic Acids Res.* 11, r169-r173.
13. Seeman, N.C., Rosenberg, J.M., and Rich, A. (1976) *Proc. Natl. Acad. Sci. USA* 73, 804-808.
14. Huang, L.-H., Wang, R.Y.-H., Gama-Sosa, M.A., Shenoy, S., and Ehrlich, M. (1984) *Nature* 308, 293-295.
15. Wang, R. Y.-H., Huang, L.-H., and Ehrlich, M. (1984) *Nucleic Acids Res.* 12, 3473-3490.
16. Klenow, H., Overgaard-Hansen, K., and Parker, S.S. (1971) *Eur. J. Biochem.* 24, 371-381.
17. Messing, J. (1983) *Meth. Enzymol.* 101, 20-78.
18. Yanisch-Perron, C., Viera, J., and Messing, J. (1985) *Gene* 33, 103-119.
19. Shenoy, S. and Ehrlich, M. (1985) *Prep. Biochem.* 14, 485-497.
20. Sutcliffe, J.G. (1979) *Cold Spring Harbor Symp. Quant. Biol.* 43, 77-90.
21. Bolivar, F. and Backman, K. (1979) *Meth. Enzymol.* 68, 245-280.
22. Galas, D. and Schmitz, A. (1978) *Nucleic Acids Res.* 5, 3157-3170.
23. Maxam, A.M. and Gilbert, W. (1980) *Meth. Enzymol.* 65, 499-560.
24. Ehrlich, M., Zhang, X.-Y., and Wang, R.Y.-H. (1985) in *Biochemistry and Biology of DNA Methylation*, G. Cantoni and A. Razin, Eds., pp. 255-269. Alan Liss, New York.
25. Humphries, R.K., Ley, T., Turner, P., Moulton, A.D. and Nienhuis, A.W. (1980) *Cell* 30, 173-183.
26. Michelson, A.M. and Orkin, S.H. (1983) *J. Biol. Chem.* 258, 15245-15254.
27. Sternberg, N. (1985) *J. Bacteriol.* 164, 490-493.
28. Ohmori, H., Tomizawa, J., and Maxam, A.M. (1978) *Nucleic Acids Res.* 5, 1479-1485.
29. Schmitz, A. and Galas, D. (1979) *Nucleic Acids Res.* 6, 111-137.
30. Leegwater, D.A.J., van Driel, W., and van der Vliet, P.C. (1985) *EMBO J.* 4, 1515-1520.
31. Aiba, H. (1983) *Cell* 32, 141-149.
32. Strauss, F. and Varshavsky, A. (1984) *Cell* 37, 889-901.
33. Emerson, R.M., Lewis, C.D., and Felsenfeld, G. (1985) *Cell* 41, 21-30.
34. Nowock, J., Borgmeyer, Uwe, Puschel, A.W., Rupp, R.A.W. and Sippell, A.E. (1985) *Nucleic Acids Res.* 13, 2045-2061.
35. Russell, G.J., Walker, P.M.B., Elton, R.A., and Subak-Sharpe, J.H. (1976) *J. Mol. Biol.* 108, 1-20.
36. Carpousis, A.J. and Gralla, J.D. (1985) *J. Mol. Biol.* 183, 165-177.
37. Ho, Y.-S., Wulff, D.L., and Rosenberg, M. (1983) *Nature* 304, 704-708.