Volume 14 Number 24 1986 Nucleic Acids Research

Methylated DNA-binding protein from human placenta recognizes specific methylated sites on several prokaryotic DNAs

Richard Y.-H.Wang+, Xian-Yang Zhang, Rana Khan, Youwen Zhou, Lan-Hsiang Huang and Melanie Ehrlich*

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

Received 12 August 1986; Revised and Accepted 20 November 1986

ABSTRACT

Methylated DNA-binding protein (MDBP) from human placenta recognizes specific DNA sequences containing 5-methylcytosine (m^5C) residues. Comparisons of binding of various prokaryotic DNAs to MDBP indicate that m^5CpG is present in the recognition sites for this protein but is only part of the recognition sequence. Specific binding to MDBP was observed for bacteriophage XP12 DNA, which naturally contains ~1/3 of its residues as m^5C, and for Micrococcus luteus DNA, M13mp8 replicative form (RF) DNA, and pBR322 when these three DNAs were methylated at CpG sites by human DNA methyltransferase. Five DNA regions binding to MDBP have been localized by DNase I footprinting or restriction mapping in methylated pBR322 and M13mp8 RF DNAs. A comparison of their sequences reveals a common 5'-m^5CGRm^5CG-3' element or closely related sequence in which one of the m^5C residues may be replaced by a T. In addition to this motif, one upstream and one downstream m^5CpG as well as other common residues over an ~20-bp long region may be recognized by MDBP.

INTRODUCTION

Vertebrate DNA methylation is largely confined to CpG sites (1-5). However, the distribution of m^5CpG (m^5C, 5-methylcytosine) sites as compared to CpG sites is complicated. There are tissue-specific differences in the locations of m^5CpG sites as well as large variations in the levels of genomic m^5C (1,6,7). Because of the complexity of DNA methylation patterns, it has been difficult to clarify the roles of DNA methylation in cellular control mechanisms despite much evidence linking methylation of certain DNA sites to negative control of transcription (8-17).

The methylated DNA-binding protein (MDBP) which we isolated from human placental nuclei (18) provides the first example of a possible mediator of the effects of DNA methylation on macromolecular synthesis. This protein binds preferentially to human DNA or bacteriophage M13 replicative form (RF) DNA when these DNAs are substituted with m^5dCTP by nick translation or primer extension, respectively (18). Furthermore, phage XP12 DNA, which naturally contains ~1/3 of its total residues as m^5C (19,20), is an even better ligand for MDBP (18).

MDBP recognizes with high affinity a particular site in pBR322 when that site is methylated at its CpG residues (21) by human DNA methyltransferase (hDMT) (22). We show in this report that MDBP binds with different affinities to various related, m^5CpG-containing DNA sequences in methylated pBR322 and M13mp8 RF DNAs. Correspondingly, the ability of DNAs with higher sequence complexity to bind to MDBP after incorporation of m^5dCTP residues depends upon their CpG dinucleotide frequencies, which indicates the importance of the mammalian pattern of *in vivo*-like DNA methylation (m^5CpG sites) for MDBP recognition.

MATERIALS AND METHODS

MDBP and DNA methyltransferases. The procedure for isolating MDBP and hDMT from placenta was as previously described (21). Unless otherwise specified, the DEAE-cellulose peak fractions were used. To achieve a high degree of methylation of the DNA ligands used in this study, the peak fractions of hDMT from the DEAE-Sepharose CL-6B column were pooled, concentrated (to ~20 maintenance units per microliter) and desalted (21,22). Usually 400 units of enzyme were incubated with 1 µg of DNA (21). *Hha*I and *Hpa*II methylases were from New England BioLabs and other enzymes from Bethesda Research Laboratories or International Biotechnologies, Inc.

Preparation of DNA ligands for MDBP binding assays. To prepare radiolabeled m^5C-rich DNAs by nick translation, phage DNAs (λ, N7 and XP12), *Micrococcus luteus* DNA, and human placental DNA were pre-nicked with DNase I (22), extracted, precipitated, and 2.5 µg of DNA dissolved in 100 µl containing 5 units of *E. coli* DNA polymerase I, 50 µg/ml of bovine serum albumin (BSA; USBC, RIA grade), 1 mM dithiothreitol, 66 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$ and 40 µM each of m^5dCTP, dGTP and dATP plus 20 µM ^{3}H-dTTP or ^{32}P-dTTP (~40 Ci/mmole). After incubation for 4 h at 14°C, the DNA was treated with nuclease S1 and purified (18). Only DNA preparations with similar molecular-weight distributions were compared. In the case of *M. luteus* DNA, various ratios of m^5dCTP to dCTP were used during *in vitro* DNA synthesis.

Plasmid pRW751 DNA (23) was methylated with 5 units of *Hha*I methylase per µg of DNA under standard conditions. The product, which was completely resistant to *Hha*I, was digested with *Taq*I and labeled at the 3' ends with [α-^{32}P]-dCTP (21).

End-labeled XP12 DNA was prepared by digesting 1 µg of DNA with 10 units of *Mbo*I (24) followed by addition of dGTP, dTTP, [α-^{32}P]dATP, and the Klenow fragment (25). After incubation at 25°C for 30 min and then incubation at 37°C

for 1 h with 25 mM EDTA and 100 μg/ml of proteinase K, the DNA was extracted with organic solvents and precipitated with ethanol.

To obtain preparative amounts of methylated MDBP-specific fragments of pBR322 or M13mp8 RF sequences, these DNAs were digested as follows. From HinfI-cleaved pBR322 DNA (200 μg), the 298-bp fragment containing pB site 1 and the 1632-bp fragment containing pB site 2 were isolated by electrophoresis on a 5% polyacrylamide gel. To further localize pB site 2, a 603-bp/517-bp doublet band was isolated from HinfI/HindIII-digested pBR322. This doublet was digested with BamHI and the 346-bp subfragment isolated on a 7.5% polyacrylamide gel. M13 sites 1, 2, and 3 were isolated from M13mp7-β RF DNA, an M13-human β-globin recombinant (probe F; 26). This RF was digested with HinfI and the 445-, 573-, and 348-bp fragments containing M13 sites 1, 2 and 3, respectively, were purified from a 5% polyacrylamide gel. These fragments were radiolabeled by 3' end-filling as for XP12 DNA. Methylation was catalyzed by hDMT using 1 μg of the 346- or 298-bp fragment of pBR322 as a substrate or 8 μg of the M13mp7-β RF prior to restriction.

Assay of MDBP activity. Total protein was quantitated by the assay of Bradford (27) with BSA as a standard. One of the two assays used for MDBP activity was the nitrocellulose filter-binding assay (18, 28). Reaction mixtures were filtered as previously described (18) through nitrocellulose membranes (Schleicher and Schuell, BA85; 25 mm in diameter, 0.45 μm pore size) which had been previously soaked in 0.3 M NaOH for 15 min, washed five times with water, rinsed twice with 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and then with water, and finally soaked in the binding buffer (20 mM Tris-HCl, pH 7.6, 40 mM NaCl, 0.2 mM EDTA, 3 mM $MgCl_2$). The binding reaction mixture usually contained the above buffer plus MDBP, 40 μg/ml BSA, and two differentially radiolabeled DNA ligands (2-10 ng, 4-10 x 10^6 cpm/μg), for example, ^{32}P-labeled XP12 and ^{3}H-labeled M. luteus DNAs.

The second MDBP assay, the fragment selection assay, was used for analyzing the relative extent of retention of various methylated DNA fragments and was performed essentially according to the previously described methods (21). In brief, the radiolabeled restricted DNA (∿10-20 ng ∿10^7 cpm/μg) was incubated with MDBP for 15 min as in the above assay. After nitrocellulose filtration, the DNA-MDBP complexes were eluted from the nitrocellulose filter and dissociated at 37°C for 16 h in the presence of 0.2% sodium dodecyl sulfate. The eluted DNA fragments were subjected to polyacrylamide gel electrophoresis and autoradiography to determine which fragments had been enriched by binding to MDBP. One unit of methylated pBR322-binding activity is defined

as the amount of MDBP that binds 1 fmol of the 298-bp HinfI fragment of ^{32}P-labeled, hDMT-methylated pBR322 in 15 min. This was determined after 10 min of preincubation with 800 ng of MboI-digested M. luteus DNA in the above buffer followed by addition of a 3' end-labeled HinfI/HindIII digest of the methylated pBR322 (∿90% of CpG dinucleotides methylated). Specific retention of the radiolabeled, methylated 298-bp fragment was quantitated after filtration, dissociation of DNA-protein complexes, gel electrophoresis, and autoradiography by measuring Cerenkov radiation in the corresponding DNA band cut out of the electrophoresis gel. The minor amount of the analogous nonmethylated 298-bp fragment that was retained after the same treatment was subtracted to account for nonspecific binding activity. Unlike the filter-binding assay, this assay allows quantitation of MDBP activity in the crude nuclear extract and was also used on pooled chromatography fractions. However, for routine quantitation of stock samples of partially purified MDBP and for assay of multiple individual fractions from column chromatography or sedimentation analyses, the more rapid nitrocellulose-filter binding assay was used. With this assay, one unit of XP12 DNA-binding activity is defined as the amount of MDBP which binds 10 ng of ^{3}H-labeled, nick-translated XP12 DNA (∿2-10 kb) in the presence of 100 ng of unlabeled M. luteus DNA. One unit of XP12 DNA-binding activity contains a similar amount of partially purified MDBP as one unit of pBR322-binding activity.

RESULTS

Characterization of MDBP

MDBP was partially purified by extraction of nuclei with 0.3 M NaCl, and chromatography on phosphocellulose, hydroxyapatite, and DEAE-cellulose as previously described (18). The MDBP activity was determined by the fragment selection assay, a specific and highly sensitive assay for sequence-specific DNA-binding proteins. The increase in specific activity and the yield of MDBP activity are given in Table 1. Several fold higher yields are obtained when BSA is added to the pooled hydroxyapatite fractions to a final concentration of 150 µg/ml before the DEAE-cellulose chromatography. Glycerol gradient sedimentation of the partially purified peak fractions of MDBP from the hydroxyapatite column gave a major peak of MDBP activity with an $S_{20,w}$ of 6.1, which only partially overlapped the main protein peak and the main peak of nonspecific DNA-binding activity (data not shown). In addition, in some samples, a second minor peak of MDBP activity was observed with an $S_{20,w}$ of 9.5. This might represent a dimerized form of the protein in the main peak of MDBP activity.

Table 1. Partial purification of human placental MDBP

Purification step[a]	Total protein (mg)	Total activity (units)[b]	Specific activity (units/mg protein)	Recovery (%)
0.3 M NaCl extract	96	5,300	55	100
Phosphocellulose	10	1,800	180	33
Hydroxyapatite	0.47	520	1,100	10
DEAE-cellulose	0.016	110	6,900	2

[a] For this isolation, 470 g of placenta was used. The 0.3 M NaCl extract is the crude extract obtained after ultracentrifugation (18). The phosphocellulose fraction is the pooled fractions assayed after dialysis. During purification, BSA was not added to stabilize MDBP activity in the last steps because of the specific activity measurements.

[b] The units of activity are based on the fragment selection assay with ^{32}P-labeled HindIII/HinfI-digested, hDMT-methylated pBR322 as defined in Materials and Methods

To investigate various factors which can influence the binding specificity of MDBP, we employed a nitrocellulose filter-binding assay in a double-label experiment with the following ligands: nick-translated, ^{32}P-labeled, m^5C-rich XP12 DNA (34 mol% m^5C; 19,20) and ^{3}H-labeled, m^5C-deficient, human placental DNA (normally 0.76 mol% m^5C but depleted to ∿0.5 mol% m^5C by extensive nick translation with dCTP rather than m^5dCTP; 29). The effect of pH on MDBP activity was determined at pH 6.6-8.8 in buffers containing 3 mM $MgCl_2$, 25 mM NaCl, 0.2 mM EDTA and 20 µg/ml of BSA. The ratio of m^5C-rich DNA to m^5C-deficient DNA bound by MDBP was ∿6 at pH 6.6-6.8, and increased to ∿16 at pH 7.5-8.0. This ratio dropped to 4 at pH 8.8. At pH 7.6, the optimal concentration of NaCl in the reaction buffer for MDBP activity was 25-50 mM. At 100 mM NaCl, although the binding ratio was slightly increased (∿19, compared to ∿16 at 25 mM NaCl), the extent of DNA-binding activity was only ∿35% of that at 25 mM NaCl. The binding specificity of MDBP was significantly decreased in the absence of $MgCl_2$. The optimal concentration of $MgCl_2$ was 3 mM and the $MgCl_2$ could be replaced by $CaCl_2$. The half-life of MDBP-XP12 DNA complexes was ∿2 h at 25°C in 10 mM sodium phosphate, pH 7.4, 3 mM $MgCl_2$, 50 mM NaCl, 0.25 mM EDTA and 50 µg/ml BSA. This was determined by addition of a 20-fold excess of unlabeled XP12 DNA after the standard 15-min incubation followed by removal of aliquots at 10-20 min intervals for filtration through nitrocellulose membranes (30).

Table 2. Binding of MDBP to m^5C-enriched DNAs with different m^5CpG contents obtained by nick translation in the presence of m^5dCTP.

m^5C-substituted ^3H-labeled DNA[a]	% of bases as m^5C[b]	% of dinucleotides as m^5CpG[c]	Relative retention by MDBP[d]
Placenta	11	0.9	1.0
M. luteus	2.6	1.0	1.0
M. luteus	4.0	1.5	1.3
M. luteus	6.5	2.5	1.8
M. luteus	13	5.1	5.8
XP12	34	12	11
λ	11	2.9	2.5
N7	15	5.3	0.8

[a] All DNAs were prepared by nick translation as described in Materials and Methods. The extent of replacement of nucleotides by *in vitro* DNA synthesis was as follows: human placental DNA, 51%; *M. luteus* DNA, 34-37%; XP12 DNA, 32%; λ DNA, 42%; N7 DNA, 41%. The sizes of these DNAs after nuclease S1 treatment were in the range of 1-7 kb.

[b] For XP12 DNA, the mol% m^5C in the DNA was used because the m^5C content was not changed by *in vitro* DNA synthesis in the presence of m^5dCTP, dGTP, dATP, and dTTP. The data for the other samples were obtained from the following calculation: the percentage replacement of nucleotides during *in vitro* DNA synthesis multiplied by the mol% C in the DNA and, in the case of *M. luteus* DNA, by the percentage m^5dCTP divided by m^5dCTP plus dCTP in the polymerization mixture.

[c] Values were obtained from the % of CpG dinucleotides in the DNA multiplied by the mol% m^5C divided by the mol% C plus m^5C. For λ and *M. luteus* DNAs, the CpG dinucleotide frequencies are 0.069 and 0.139, respectively (41). For N7 and XP12 DNA, the CpG dinucleotide frequency was assumed to be the product of their mol% C and their mol% G.

[d] Percentage of ^3H-labeled DNA bound divided by the percentage of competing m^5C-substituted, ^{32}P-labeled human placental DNA bound by MDBP using approximately equimolar amounts of both DNAs. Binding was measured by the standard filter-binding assay.

Recognition of m^5CpG-containing sequences by MDBP

We compared the binding of MDBP to DNA in which essentially all the C residues are 5-methylated and to DNA with methylation confined to CpG sites. *M. luteus* DNA (36 mol% C, 0.3 mol% m^5C) was partially methylated with hDMT so that it contained 6% of its bases as m^5C and most or all of this m^5C in CpG sequences (21,22). It competed for binding to nick-translated, radiolabeled XP12 DNA in the nitrocellulose filter-binding assay almost half as well as did an excess of unlabeled XP12 DNA (data not shown), which has 34 mol% m^5C.

That retention of DNA by MDBP depends more on the concentration of m^5CpG sites than on the level of total m^5C residues was also seen in comparisons of DNA ligands extensively nick-translated in the presence of m^5dCTP and dCTP in various ratios (Table 2). After nick translation under comparable conditions, the concentration of m^5CpG dinucleotide sequences in human DNA was much less than in M. luteus DNA because of the ~4-fold underrepresentation of CpG sequences in mammalian DNA (1) and the higher (A + T) content of human DNA (58 vs. 28 mol%; 29). The relative binding of these DNAs indicated that the m^5CpG content correlated much better than the total m^5C content with the ability of the DNAs to be retained by MDBP (Table 2). However, m^5C-substituted N7 phage DNA was retained poorly by MDBP (Table 2) despite a moderately high concentration of m^5CpG-sequences. This DNA has a genome that is in the size range of that of λ phage (31) and can be methylated just as efficiently as M. luteus DNA by hDMT. The poor recognition of N7 DNA extensively substituted with m^5C residues is probably due to its not having the necessary sequences neighboring CpG sites (21).

We have also tested pRW751, a derivative of pBR322 containing a $(CpG)_{16}$ and a $(CpG)_{13}$ sequence (23), for MDBP binding. After methylation with HhaI methylase, restriction and end-labeling, selective retention of DNA fragments by MDBP was analyzed as described below. Little or no specific binding to MDBP was found for any of the restriction fragments from this HhaI-methylated DNA although it contained two runs of m^5CpG. Similarly, when M. luteus DNA was exhaustively methylated with both HpaII methylase and HhaI methylase, so that ~3% of its bases were m^5C residues at CpG sites (CCGG and GCGC), it did not exhibit specific binding by MDBP presumably because the CpG-containing sequences recognized by the protein were not methylated by these enzymes.

Recognition of specific methylated restriction fragments by MDBP

We next examined the binding of restriction fragments of naturally m^5C-rich XP12 DNA to determine whether some were bound by MDBP better than others in the fragment selection assay. For this assay, XP12 DNA was cleaved by MboI, one of a very small number of restriction endonucleases able to hydrolyze DNA which is fully methylated at C residues (24). After ^{32}P end-labeling, the MboI fragments of XP12 DNA were incubated with MDBP and filtered through a nitrocellulose membrane. The retained fragments were dissociated from the DNA•MDBP complexes and eluted from the membrane. They were then analyzed by polyacrylamide gel electrophoresis followed by autoradiography. Preferential retention of certain groups of DNA fragments by MDBP was observed (32 and data not shown).

Table 3. Restriction mapping of the MDBP-specific pB site 2 in CpG-methylated pBR322[a]

Type of digest		Sizes of hDMT-methylated fragments binding MDBP[b]	Sizes of hDMT-methylated fragments not binding MDBP[c]
Enzyme(s)	Substrate	(bp)	(bp)
HindIII/ HinfI	pBR322	29 * 603 632 *	1029, 506, 396, 344 221, 220, 154, 75
MspI	1632-bp HinfI fragment	170 * 217 387 *	622[d], 242, 122, 110 99, 85, 67, 35
MspI	346-bp HindIII-BamHI fragment[e]	170 205 375 *	---
HaeIII	205-bp MspI-BamHI fragment	296 79 375 *	---
TaqI	205-bp MspI-BamHI fragment	339 37 375 *	---

[a] The hDMT-treated DNA contained 80% of its CpG sites methylated. It was restricted with the indicated enzymes and then the 603-bp and 217-bp fragments were identified as containing MDBP sites by the fragment selection assay described in Materials and Methods. Analysis of the singly end-labeled MspI/BamHI, HaeIII/BamHI, and TaqI/BamHI fragments was by the nitrocellulose filter-binding assay as described in Materials and Methods and Results.

[b] The lines represent the MDBP-specific restriction fragments; these bound preferentially to MDBP only when methylated. The numbers above the lines refer to the standard map position (35,42) of the 5' base of the indicated restriction recognition sites. The numbers below the line give the distance between the two designated restriction sites, which is approximately equal to the number of base pairs in each fragment. In the case of the TaqI/BamHI fragment, the exact number of base pairs in the double-stranded region is indicated; this fragment had a 2-base overhang at one 5'-end and a 1-base overhang at the other due to partial end-filling. The radiolabeled end of each fragment is indicated by an asterisk. The first enzyme named in a double digest cut the intact DNA at the restriction site shown at the left side of the fragment and the second one cut at the restriction site on the right.

[c] The radiolabeled DNA fragments which did not bind preferentially to MDBP in the fragment selection assay (exhibiting background amounts of binding or no detectable binding) and which did not coelectrophorese with MDBP-binding fragments are listed.

[d] The 622-bp MspI fragment (derived from the 1632-bp HinfI fragment; positions 3902-161) showed weak binding which is partly attributable to its relatively large size that gives it proportionately more nonspecific binding (Fig. 1).

[e] The 205-bp MspI/BamHI fragment was radiolabeled only at its BamHI terminus because it was derived from a 346-bp HindIII/BamHI fragment (positions 29-375) that had been radiolabeled at both ends. The radiolabeling was by partial end-filling in the presence of dGTP, dTTP, and [^{32}P]dATP which generated, on the right end in the above scheme, a double-stranded region through position 378.

We had already detected one MDBP-specific site in pBR322 (21), a much smaller DNA molecule than XP12 DNA (33) and, unlike XP12 DNA, of known sequence. In that study, pBR322 was singly digested with HinfI after enrichment for m^5C residues by in vitro methylation catalyzed by hDMT (22). In the present study the plasmid was similarly methylated to give 7% of the bases as m^5C (~90% of CpG sites methylated) and then digested with HindIII and HinfI. Two end-labeled DNA fragments were observed to bind specifically to MDBP. In addition to the previously described 298-bp HinfI fragment (21), one methyl-

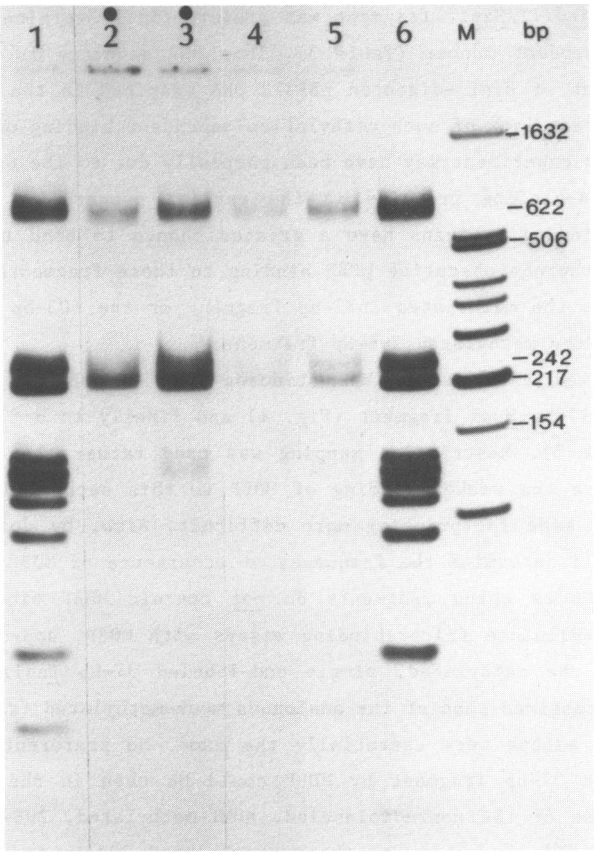

Fig. 1. Localizing pB site 2, a moderate affinity MDBP site in pBR322. Retention by MDBP of methylated subfragments from the 1632-bp fragment of HinfI-digested pBR322. The purified 1632-bp HinfI fragment was methylated, digested with MspI and end-labeled with [α-^{32}P]dCTP. This DNA had $\sim$80% of its CpG sites methylated as determined by incorporation of methyl groups from [^{3}H-methyl]S-adenosylmethionine (AdoMet) and confirmed by digestion with HpaII, HhaI, and ThaI (22, 39). Approximately 15 ng ($\sim$5 x 10^5 cpm) of the DNA fragments was incubated with 0.6 units of MDBP in the presence of 30 ng each of double-stranded and single-stranded, HinfI-digested, unlabeled λ DNA and then the nitrocellulose membrane-bound DNA was eluted and dissociated from protein. All the MDBP-retained DNA fragments eluted from a single nitrocellulose filter were electrophoresed in a 9% polyacrylamide gel and autoradiographed. The units of MDBP in this and subsequent figures refer to XP12 DNA-binding units as determined by the filter-binding assay. Lanes 1-3, methylated DNA; lanes 4-6, DNA treated with hDMT in the absence of AdoMet (mock-methylated DNA). Lanes 1 & 6, $\sim$1 ng of input DNA without incubation with proteins or passage through a nitrocellulose filter; lanes 2 & 3, DNA fragments retained by two different fractions of MDBP from the DEAE-cellulose column; lanes 4 & 5, analogous to lanes 2 & 3 except with unmethylated DNA. M, HinfI-digested pBR322 DNA serving as markers. In this and subsequent figures, the lanes with a dot on top contained methylated DNA fragments that were retained by MDBP.

ated, 603-bp HindIII/HinfI fragment was preferentially retained by MDBP in a methylation-dependent manner (Table 3). This DNA sequence is present in the 1632-bp fragment of HinfI-digested pBR322 DNA examined in the previous study (21). The apparent lack of much methylation-dependent binding of that fragment in the previous experiment may have been partially due to the presence of more nonspecific DNA-binding proteins in the earlier preparation of MDBP. Nonspecific DNA-binding proteins have a greater chance to bind to larger sized DNA fragments thereby obscuring MDBP binding to those fragments. Furthermore, MDBP binding to the methylated 1632-bp fragment or the 603-bp subfragment is weaker than to the methylated 298-bp fragment.

We localized this second MDBP binding site of pBR322 by restriction mapping to a 217-bp MspI fragment (Fig. 1) and finally to a 37-bp TaqI/BamHI fragment (Table 3). Restriction mapping was used rather than DNase I footprinting because the weaker binding of MDBP to this second site compared to the first site made footprinting more difficult. Also, by doing restriction mapping we could determine the frequency of occurrence of MDBP sites in these genomes as well as which fragments do not contain MDBP binding sites. In parallel nitrocellulose filter-binding assays with MDBP, approximately eight times more of the methylated, single end-labeled 37-bp TaqI/BamHI fragment (Table 3) was retained than of the analogous mock-methylated fragment; results from duplicate assays were essentially the same. No preferential binding of this methylated 37-bp fragment by MDBP could be seen in the presence of a five-fold excess of the nonradiolabeled, hDMT-methylated, 298-bp HinfI fragment of pBR322 DNA. In contrast, the nonmethylated 298-bp fragment under the same conditions had no detectable effect on the binding of the 37-bp fragment. This result indicates that the same protein in the MDBP preparation is responsible for recognition of the previously described high-affinity MDBP site, pB site 1, and for recognition of this moderate-affinity MDBP site, pB site 2.

Methyltransferase-methylated M13mp8 RF DNA, which contained 3.6% of its bases as m^5C (>90% of CpG sites methylated) was used in a search for additional MDBP-binding sites. In a previous report, retention of specific restriction fragments from M13mp8 RF methylated in its complementary strand by incorporation of m^5dCTP was noted (21). In the present study, after methylation with hDMT and restriction with TaqI, M13mp8 RF sequences showed preferential retention by MDBP of 927-bp and 614-bp fragments and, to a lesser extent, of a 579-bp fragment (Table 4). Such preferential binding only occurred when these fragments were methylated and was competed by the methylated (but not the unmethylated) 298-bp HinfI fragment of pBR322.

Table 4. Restriction mapping of three MDBP-specific sites in CpG-methylated M13mp8 RF

M13 site	Type of digest: Enzyme(s)	Substrate	Sizes of hDMT-methylated fragments binding MDBP[a] (bp)	Sizes of hDMT-methylated fragments not binding MDBP[b] (bp)
1	TaqI	M13mp8 RF	6248 * — 614 — * 6862	1018, 971, 791, 703, 441, 381, 239
1	EcoRI/HinfI	M13mp7-β RF[c]	6273 * — 339 — * 6612	1288,771,328,324,274,261,253, 234,232,137,96,80,63,46,45
1	Sau96I	339-bp E/H fragment	6384 * — 228 — * 6612	111
1	MspI	339-bp E/H fragment	6469 * — 143 — * 6612	178
2	TaqI	M13mp8 RF	2527 * — // — 927 — * 3454	See M13 site 1
2	HinfI	M13mp7-β RF	2844 * — 573 — * 3417	See M13 site 1
2	HaeII	573-bp HinfI fragment	2844 * — 194 — 3038	379
2	AluI	573-bp HinfI fragment	2844 * — 118 — 2962	142
3	TaqI	M13mp8 RF[d]	1948 * — 579 — * 2527	See M13 site 1
3	HinfI	M13mp7-β RF	2496 * — 348 — * 2844	See M13 site 1
3	HaeIII	348-bp HinfI fragment	2496 * — 57 — 2553	291

[a] As in Table 3, DNA was methylated with hDMT, restricted, end-labeled (indicated by the asterisk), and analyzed by the fragment selection assay (Figs. 2 and 3). The binding to MDBP of these fragments was methylation-dependent. The length in base pairs is given under the line representing the fragment. Above the line are the positions of the restriction recognition sites in the M13mp8 or the unrecombined M13mp7 map (43-45).

[b] Only those radiolabeled fragments which did not coelectropherese with MDBP-binding fragments are listed. For M13mp7-β, only the vector-derived fragments are given.

[c] This EcoRI/HinfI (E/H) fragment was obtained by first electrophoretically isolating a 445-bp HinfI fragment which contained ∿106-bp of the β-globin sequence. The vector sequences were released as a 339-bp E/H fragment. Position 6273 and 6612 refer to the M13mp7 map; the latter is position 6603 on the M13mp8 map. Positions 6469 and 6384 on the M13mp7 map are positions 6460 and 6375, respectively, on the M13mp8 map. The positions described for M13 sites 2 and 3 on the M13mp7 map are numbered identically on the M13mp8 map. The non-binding fragments listed are those from digestion with HinfI alone.

[d] The fragments in the TaqI digest were labeled by a fill-in reaction with [d-^{32}P]dCTP and dGTP. Therefore, the right end of the 579-bp fragment is actually at position 2529 after filling its 2-base 5' overhang. This provides a 33-bp overlap of double-stranded sequence (position 2497-2529) between this fragment and the partially 3' end-filled 348-bp HinfI or 58-bp HinfI/HaeIII fragments.

To further locate the MDBP-binding sites in M13 DNA by restriction analysis, we digested M13mp7-β RF recombinant DNA, which contains a human β-globin gene insert (26), instead of M13mp8 RF with HinfI. The recombinant was used because of the better separation of its MDBP-binding fragments.

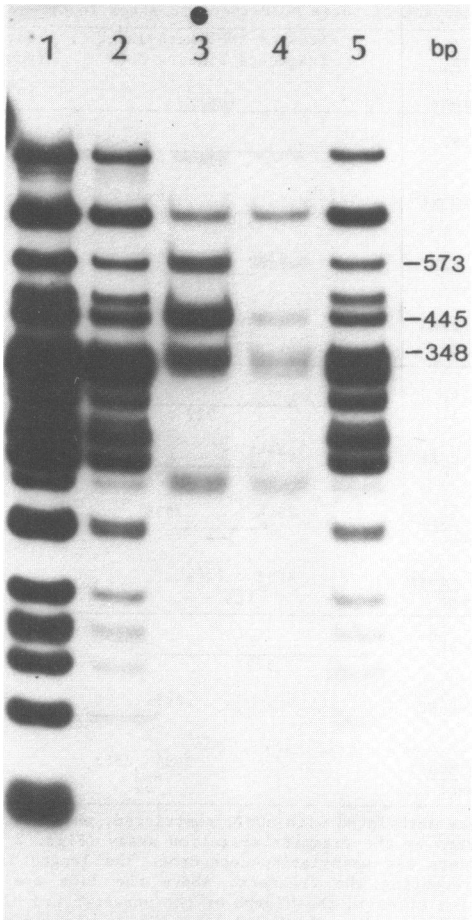

Fig. 2. Retention of hDMT-methylated M13mp7-β HinfI fragments by MDBP. About 10 ng (∿1.2 x 10^5 cpm) of DNA fragments radiolabeled at their 3' ends plus 20 ng of unlabeled, denatured HinfI-fragments of λ DNA were incubated with 0.8 units of MDBP and then analyzed as described in Fig. 1. The 348-bp band is actually a doublet of 348-bp and 345-bp fragments from the vector with an adjacent band containing a junction fragment (∿336-bp) from the β-globin gene (26; 40) and from the vector. Isolation and restriction analysis of the fragments in this band indicate that only the 348-bp vector fragment has an MDBP-specific site. The HinfI site at position 3257 of the standard M13mp7 map is missing from M13mp7-β DNA so that a 573-bp fragment instead of 413- and 160-bp fragments are present. Lanes 1-3, methylated DNA; lanes 4 & 5, mock-methylated DNA. Lane 1, input DNA (∿0.5 ng) without incubation with protein or passage through a nitrocellulose filter; lanes 2 & 5, DNA fragments retained by nonspecific DNA-binding proteins from hydroxyapatite column fractions outside the MDBP peak; lanes 3 & 4, fragments retained by the DEAE-cellulose fraction of MDBP.

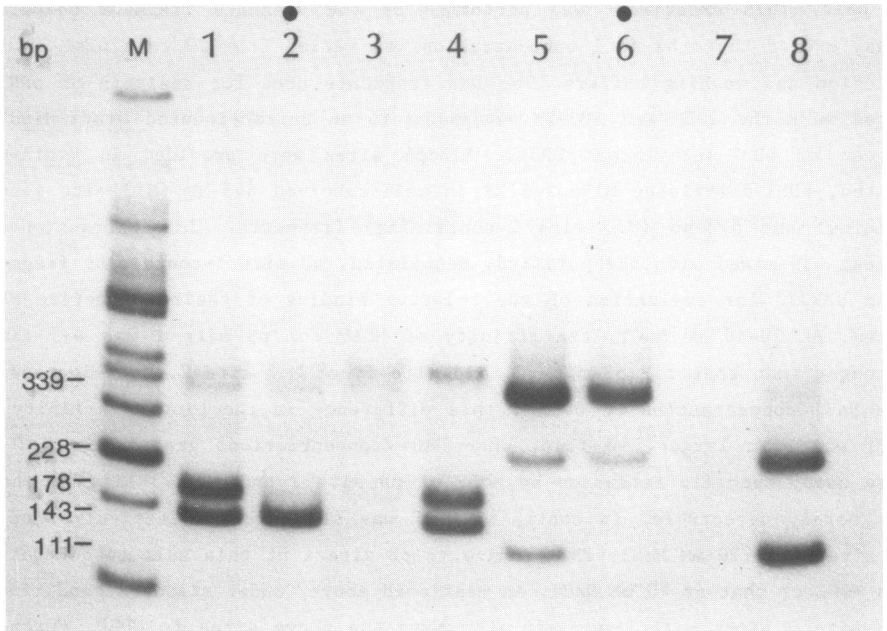

Fig. 3. Restriction mapping of the MDBP-binding site in the 339-bp fragment of M13mp7-β DNA. The hDMT-methylated or mock-methylated EcoRI/HinfI 339-bp fragment radiolabeled with [α-^{32}P]dATP at its 3'-ends was purified and further restricted with MspI or Sau96I. Approximately 3 ng (∿6 x 10^4 cpm) of radiolabeled DNA fragments plus 100 ng of unlabeled MboI fragments of M. luteus DNA were incubated with 1 unit of MDBP from the hydroxyapatite fraction and then analyzed as described in Fig. 1. Lanes 1-4, MspI-digested DNA fragments; lanes 5-8, Sau96I-digested DNA fragments. Lanes 1, 2, 5, & 6, methylated DNA; lanes 3, 4, 7, & 8, mock-methylated DNA. Lanes 1, 4, 5, & 8, input DNA without incubation with protein or passage through nitrocellulose; lanes 2, 3, 6, & 7, DNA fragments retained by MDBP. Lane M, HinfI- and HindIII-digested pBR322 DNA as molecular-weight markers. The incomplete digestion seen in lanes 5 & 6 is apparently due to the inhibition by hemimethylation at a CpG overlapping the Sau96I recognition site.

End-labeled hDMT-methylated 573-bp, 445-bp and 348-bp fragments were preferentially retained by MDBP (Fig. 2). Three MDBP binding sites, M13 sites 1, 2, and 3, all derived from vector sequences, were then localized by the fragment selection assay as shown in Table 4 and Fig. 3.

Comparison of MDBP binding to various sites

To elucidate the relationships between related sequences recognized by MDBP, the relative affinity of MDBP for DNA fragments containing pBR322 or M13mp8 MDBP sites was determined in the presence of different concentrations

of salt. This experiment was performed by the standard fragment selection assay except that the NaCl concentration was varied from 20 to 120 mM in the reaction and washing buffers. The DNA fragments used for analysis of pBR322 sites were the 298- and 603-bp fragments in an unfractionated HinfI/HindIII digest of hDMT-methylated pBR322. M13mp8 sites were provided in HinfI-digested, hDMT-methylated M13mp7-β RF in well-resolved 445-bp (M13 site 1-containing) and 573-bp (M13 site 2-containing) fragments. This unfractionated digest was mixed with the purified, methylated, pB site 1-containing fragment from pBR322 for evaluation of the relative binding of their respective MDBP sites. At 20-40 mM NaCl, the affinity of MDBP for pB site 1 was 4-5 times stronger than that for pB site 2, M13 site 1, or M13 site 2. Upon increasing the NaCl concentration to 60 mM, this difference in the binding affinity of MDBP was even larger, ∿8 fold. When NaCl concentrations greater than 80 mM were used, specific retention by MDBP of pB site 2 and of M13 sites 1 and 2 was barely detectable. In contrast, MDBP was still very selectively binding pB site 1 at 120 mM NaCl. The binding to pB site 1 at this salt concentration was ∿40% of that at 40 mM NaCl. As mentioned above, under standard conditions, M13 site 3 binds with lower affinity than the above sites to MDBP. Furthermore, yet lower-affinity MDBP sites can be found (Figs. 1 and 2).

DISCUSSION

From this study and the three preceding reports (18,21,34), it is clear that MDBP recognizes m^5C-containing DNA sequences, specifically, those with m^5CpG sites. The frequency of m^5CpG sites rather than of m^5C residues is a critical determinant of the ability of *in vitro*-methylated DNAs of moderate to high complexity to bind to MDBP (Table 2). This confirms that MDBP has a specificity for DNA methylated in the *in vivo* pattern of vertebrate DNA, that is, at CpG dinucleotides (1). However, the presence of m^5CpG dinucleotides is not sufficient for MDBP binding (21). For example, as shown in this report, MDBP will not bind to a $(m^5CpG)_{16}$ sequence in a pBR322 derivative. Having identified the sequence of a high-affinity binding site for MDBP (pB site 1) on methylated pBR322 (21), we proceeded to localize other MDBP binding sites on small prokaryotic DNAs. In addition to pB site 1, we found four other MDBP-specific sites in hDMT-methylated (CpG-methylated) pBR322 and M13mp8 RF. Like pB site 1, they showed little or no binding to MDBP when unmethylated.

We localized these four lower-affinity binding sites by sequential restriction analysis. For example, the second MDBP-specific site in CpG-methylated pBR322 (pB site 2) was localized to a 603-bp fragment, then to a

217-bp subfragment, and finally to a 37-bp subfragment (Fig. 1; Table 3). The sequence of one strand of the double-stranded portion of this CpG-methylated subfragment is as follows (with the pB site 1 region aligned underneath it for maximum homology):

5'-ACTAMGMGATCATGGMGACCACACCMGTCCTGTGGAT-3' pB site 2
5'-CTGATMGTCAMGGMGATTTATGCMG-3' pB site 1.
.....

The m^5C residues are indicated by an M and the region of pB site 1 protected against DNase I is indicated by the dotted line. In the pB site 2 sequence (positions 342-378 of the standard pBR322 map; 35), the residues showing exact homology to the DNase I-protected region of pB site 1 are underlined as is the shared 5'-CMG-3'. In pB site 1, this 5'-CMG-3' sequence occurs three bases after the apparent 3' end of the region protected by MDBP against DNase I. At that position, enhanced cleavage by DNase I was seen (21). There are four shared (M/T)pG dinucleotides at these two MDBP-specific sites if one 1-bp displacement is made on the 5' side (Table 5). Such specific binding to two related DNA sequences, one of which has an inserted base pair relative to the other, is precedented (36,37). It is interesting to note that there is considerable homology between these two regions even outside the sequence recognized by MDBP. This suggests that the areas surrounding pB sites 1 and 2 may be the relics of an ancient DNA duplication in the tet region (35).

Three sites in CpG-methylated M13mp8 RF, M13 sites 1, 2 and 3, also bind specifically to MDBP (Figs. 2 & 3; Table 4). These sites were localized by restriction mapping to regions containing 118, 143, and 33 bp, respectively. In these MDBP-specific fragments, one or, in the case of M13 site 1, three sequences had sufficient homology to pB site 1 to be considered as putative MDBP-binding sites. These methylated sequences plus those of pB sites 1 and 2 are aligned for maximum homology in Table 5. Because all of these sites in their unmethylated form are recognized very poorly or not at all by MDBP (Figs. 1-3), this protein is probably making specific contacts with the 5-methyl group of m^5C residues in these sites. Therefore, T residues with their analogous 5-methyl group in the outer major groove of the B helix (38), might be at least partially recognized when they substitute for m^5C residues. Indeed, at several of the underlined positions of the MDBP sites shown in Table 5, a T residue replaces the analogous m^5C residue in pB site 1.

A comparison of the sequences for the above sites suggests that MDBP interacts with a rather long region (~20-bp). Table 5 reveals that almost all of the putative binding sites, contain an MpG or TpG dinucleotide on the 5'

Table 5. Methylated pBR322 and M13mp8 RF DNA sequences bound by MDBP

MDBP binding site	Sequence of binding site[a]
pB site 1	A T M G T C A M G G M G A T T T A T G C M G
pB site 2	G M G A T C A T G G M G A C C A C A C C M G
M13 site 1	T G G G A A C A A A M G G M G G A T T G A C M G T - or - T G A G G G G A M G A M G A C A G T A T M G G - or - C A M G G T T A M G A T G M G C C C A T C T A
M13 site 2	M G T T A T T A T T G M G T T T C C T M G G T
M13 site 3	C T G A T T A M G G T G C T G C T A T C G

[a]The sequence of one of the two strands of each of these sites is written 5' to 3' in its fully CpG-methylated form with M representing m^5C. These sites are located at the following positions of the standard pBR322 or M13mp8 DNA sequence maps: pB site 1, 1145-1166; pB site 2, 347-368; M13 site 1, 6574-6598, 6494-6516, or 6529-6551 (the position numbers are given for the viral strand although the sequences shown for the first two sites are on the analogous positions of the complementary strand); M13 site 2, 2877-2899; and M13 site 3, 2509-2529. pB site 1 was identified by DNase I footprinting (21). The fragment to which pB site 2 was localized has only 37 bp. The sequence assignments for the M13 sites 1 and 2 are tentative because of the much larger size of the fragments to which they were localized (143 or 118 bp). M13 site 3 was localized to a 33-bp region terminating in the ATCG sequence shown at the 3' end. This CpG was unmethylated because it arose from an end-filling reaction subsequent to DNA methylation; however, its complement on the other strand was methylated. The overlined positions in pB site 1, which has the highest affinity for MDBP of all these sites, denote residues conserved in at least two of the other sites shown and include shifts of 1-3 positions at MpG dinucleotide sequences and substitutions of T for M. In pB site 2 and the M13 sites, the MpG, TpG, or CpG dinucleotide sites which are hypothesized to be recognized by MDBP are underlined.

side and a central palindrome 5'-MGGMG-3' or related pentanucleotide derivatives such as 5'-TGGMG-3', 5'-MGGTG-3', or 5'-MGAMG-3'. These features appear to be important for recognition by MDBP as determined by a comparison of the above binding sites (Table 5) as well as by oligonucleotide-directed mutagenesis and site-specific methylation of pB site 1 (34). However, this pentanucleotide motif is clearly not sufficient for MDBP binding because it is present in numerous copies in non-binding fragments of pBR322 and M13 DNAs and even the presence of an MpG dinucleotide three residues upstream did not suffice to confer specific binding. There is another partially palindromic

element in pB site 1, namely, 5'-ATMGYCRYGRMGAT-3' whose dyad symmetry appears to be important for high-affinity binding by MDBP (34). However, very little dyad symmetry in this location is seen in the other MDBP sites. This may contribute to their lower affinity for MDBP.

Also of interest are pB site 1's T and A residues which are at least partially conserved in the other sequences (Table 5). T residues on both strands as well as m^5C residues might be recognized by MDBP largely by virtue of their 5-methyl group if these residues are correctly positioned within the sequence. This hypothesis is consistent with our finding that certain MDBP recognition sites can be created by substituting a DNA with m^5C via incorporation of m^5dCTP but not by an hDMT-catalyzed reaction (21). In the former case, an m^5C residue might be able to substitute for a T residue of a variant of an MDBP consensus sequence.

Despite the homologies between pB site 1 and the three other MDBP sites described here, it is apparent that there are many differences in these sequences (Table 5). This indicates that MDBP can bind in a methylation-dependent fashion to a wide variety of related m^5CpG-containing sequences in DNA. Determination of whether only human genomic analogues of high-affinity sites (like pB site 1) or also lower-affinity sites in the human genome (like those of pB site 2 or the three M13mp8 sites) are physiologically relevant will require isolation and extensive characterization of such sites.

ACKNOWLEDGEMENTS

We thank Drs. Tim Ley and Robert Wells for recombinant DNAs. This research was supported in part by NIH grant GM33999.

*To whom correspondence should be addressed.
+Present address: Department of Transfusion Medicine, NIH Clinical Center, Bethesda, MD 20892, USA

REFERENCES

1. Ehrlich, M. and Wang, R.Y.-H. (1981) Science 212, 1350-1357.
2. Manuelidis, L. (1981) FEBS Let. 129, 25-28.
3. Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V., and Brown, D.D. (1978) Cell 13, 717-725.
4. Pech, M., Streek, R.E. and Zachau, H.G. (1979) Cell 18, 883-893.
5. Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991-1995.
6. Cooper, D.N. (1983) Hum. Genet. 64, 315-333.
7. 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.
8. Busslinger, M., Hurst, J. and Flavell, R.A. (1983) Cell 34, 197-204.
9. Lieberman, M.W., Beach, L.R. and Palmiter, R.D. (1983) Cell 35, 207-214.
10. Christman, J.K. (1984) Curr. Top. Microbiol. Immunol. 108, 49-78.
11. Darmon, M., Nicolas, J.-F. and Lamblin, D. (1984) EMBO J. 3, 961-967.

12. Blackman, M.A. and Koshland, M.E. (1985) Proc. Natl. Acad. Sci. USA 82, 3809-3813.
13. Jones, P.A. (1985) Cell 40, 485-486.
14. Thompson, M.D. and Nakhasi, H.L. (1985) Cancer Res. 45, 1291-1295.
15. Wolf, S.F. and Migeon, B.R. (1985) Nature 314, 467-469.
16. Kandala, J.C., Kistler, W.S. and Kistler, M.K. (1985) J. Biol. Chem. 260, 15959-15964.
17. Lock, L.F., Melton, D.W., Caskey, C.T. and Martin, G.R. (1986) Mol. Cell. Biol. 6, 914-924.
18. Huang, L.-H., Wang, R.Y.-H., Gama-Sosa, M.A., Shenoy, S. and Ehrlich, M. (1984) Nature 308, 293-295.
19. Kuo, T.-T., Huang, T.-C. and Teng, M.-H. (1968) J. Mol. Biol. 34, 373-375.
20. Kuo, K.C., McCune, R.A., Gehrke, C.W., Midgett, R. and Ehrlich, M. (1980) Nucleic Acids Res. 8, 4763-4776.
21. Wang, R.Y.-H., Zhang, X.-Y. and Ehrlich, M. (1986) Nucleic Acids Res. 14, 1599-1614.
22. Wang, R.Y.-H., Huang, L.H. and Ehrlich, M. (1984) Nucleic Acids Res. 12, 3473-3490.
23. Klysik, J., Stirdivant, S.M. and Wells, R.D. (1982) J. Biol. Chem. 257, 10152-10158.
24. Huang, L.-H., Farnet, C.M., Ehrlich, K.C. and Ehrlich, M. (1982) Nucleic Acids Res. 10, 1579-1591.
25. Klenow, H., Overgaard-Hansen, K. and Parker, S.S. (1971) Eur. J. Biochem. 24, 371-381.
26. Ley, T.J., Anagnou, N.P., Pepe, G. and Nienhuis, A.W. (1982) Proc. Natl. Acad. Sci. USA 79, 4775-4779.
27. Bradford, M. (1976) Anal. Biochem. 72, 248-254.
28. Riggs, A.O., Suzuki, H. and Bourgeois, S. (1970) J. Mol. Biol. 48, 67-83.
29. 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.
30. Hinkle, D.C. and Chamberlin, M.J. (1972) J. Mol. Biol. 70, 157-185.
31. Compton, S.W., Mayo, J.A., Ehrlich, M., Ackermann, H.W., Tremblay, L., Cord, C.E. and Scaletti, J.V. (1979) Can. J. Microbiol. 25, 1027-1035.
32. 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.
33. Ehrlich, M., Ehrlich, K. and Mayo, J.A. (1975) Biochim. Biophys. Acta 395, 109-119.
34. Zhang, X.-Y., Ehrlich, K.C., Wang, R. Y.-H. and Ehrlich, M. (1986) Nucleic Acids Research 14, 8387-8398.
35. Sutcliffe, J.G. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 77-90.
36. Cohen, R.B., Sheffery, M. and Kim, C.G. (1986) Mol. Cell. Biol. 6, 821-832.
37. Siebenlist, U., Simpson, R.B., and Gilbert, W. (1980) Cell 20, 269-281.
38. Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976) Proc. Natl. Acad. Sci. USA 73, 804-808.
39. Wang, R. Y.-H., Shenoy, S. and Ehrlich, M. (1984) J. Virol. 49, 674-679.
40. Lawn, R. M., Efstratiadis, O'Connell, C., and Maniatis, T. (1980) Cell 21, 647-651.
41. Josse, J., Kaiser, A.D. and Kornberg, A. (1961) J. Biol. Chem. 236, 864-875.
42. Peden, K.W.C. (1983) Gene 22, 277-280.
43. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
44. van Wezenbeek, P.M.G.F., Hulsebos, T.J.M. and Schoenmakers, J.G.G. (1980) Gene 11, 129-148.
45. Yanisch-Perron, C., Viera, J. and Messing, J. (1985) Gene 33, 103-119.