Volume 17 Number 21 1989 Nucleic Acids Research

How different DNA sequences are recognized by a DNA-binding protein: effects of partial proteolysis

Prakash C.Supakar, Xian-Yang Zhang, Sherwood Githens[1], Rana Khan, Kenneth C.Ehrlich[2] and Melanie Ehrlich

Department of Biochemistry, Tulane Medical School, New Orleans, LA 70112, [1]Department of Biological Sciences, University of New Orleans, New Orleans, LA 70148 and [2]Southern Regional Research Center, US Department of Agriculture, New Orleans, LA 70179, USA

Received July 5, 1989; Revised and Accepted September 22, 1989

ABSTRACT

MDBP is a sequence-specific DNA-binding protein from mammals that recognizes a variety of DNA sequences, all of which show much homology to a partially palindromic 14 base-pair consensus sequence. MDBP subjected to limited proteolysis and then incubated with various specific oligonucleotide duplexes yielded two types of complexes. The relative concentrations of these complexes varied greatly depending on how closely the MDBP site matched the consensus sequence. No such DNA sequence-specific differences in the types of complexes formed were seen with intact MDBP. Partial proteolysis also changed the relative affinity of MDBP for several of its binding sites. The nature of the two types of complexes formed from fragmented MDBP and DNA was studied by DNA competition assays, protein titration, site-directed mutagenesis, and dimethyl sulfate and missing base interference assays. The results suggest that, for some specific DNA sequences, half-site interactions with one MDBP subunit predominate and for others, strong interaction of two subunits with both half-sites readily occur.

INTRODUCTION

Methylated DNA-binding protein (MDBP) is a ubiquitous mammalian protein originally identified by its specific recognition of DNA sequences methylated at cytosine residues (1−3). It recognizes certain DNA sequences devoid of 5-methylcytosine (m^5C) but only if they contain TpG or TpA dinucleotides replacing m^5CpG's of the methylation-dependent binding sites (4, 5). MDBP may play a role in transcription control as evidenced by its recognition of two sites in a cytomegalovirus enhancer (Zhang *et al.*, unpublished data).

The 14 base-pair (bp) consensus sequence recognized by MDBP, 5′-RTm^5YRYYAm^5-YRGm^5YRAY-3′ on one strand (R, A or G; Y, T or C; m^5Y, m^5C or T), exhibits partial dyad symmetry (6). It contains much degeneracy although many different single-site mutations abolish detectable MDBP binding (5,7). This degeneracy can be seen, for example, in comparing MDBP recognition sites in *Escherichia coli* plasmid pBR322 (pB site 1), the thymidine kinase gene of herpes simplex virus type 1 (HSV site 2), and bacteriophage λ DNA (λ site 1). Only five out of fourteen base-pairs are conserved from pB site 1 to HSV site 2 (6). λ site 1 differs from previously described MDBP sites by having, in effect, a deletion of the central A·T base-pair conserved in the other eleven MDBP sites (3, 6). Despite their differences, HSV site 2, pB site 1, and λ site 1, efficiently compete with each other for MDBP, exhibit similar DNase I footprint patterns, and show much homology to the consensus sequence (6), all of which suggests that they are recognized by the same DNA-binding domain of MDBP.

To investigate how MDBP interacts with considerably different sites, we tested the effect of partial proteolysis on its recognition of a variety of MDBP-specific DNA sequences.

Two types of complexes of different mobilities were formed between this fragmented MDBP and oligonucleotide duplexes containing an MDBP recognition site. The relative amounts of these MDBP-specific complexes obtained from a given duplex varied greatly depending on the sequence of the DNA recognition site. The evidence suggests that this depends on whether a predominately half-site interaction with a single subunit of MDBP occurs or two monomeric fragments bind to two half-sites.

MATERIALS AND METHODS

Purification of MDBP

Unless otherwise noted, MDBP was partially purified from human placental nuclei through the hydroxylapatite chromatography step (2). As indicated, for one set of experiments, the hydroxylapatite fraction was further purified on a DNA-affinity column that had been prepared by coupling annealed synthetic oligonucleotides containing pB site 1 to cyanogen bromide-activated Sepharose CL-4B (Pharmacia, Piscataway, New Jersey). In one experiment, the DNA·protein complexes formed from oligonucleotide duplexes and partially proteolyzed rat lung MDBP or mouse L cell MDBP were compared to those from analogously treated human placental MDBP. The rodent samples were ammonium sulfate (30% saturation) fractions of nuclear extracts prepared as previously described (4).

Preparation of labeled oligonucleotide duplexes

Oligonucleotides were synthesized and purified and one strand of each pair of complementary oligonucleotides was radiolabeled with [γ-^{32}P]ATP at its 5′ terminus as described previously (5, 7). After removal of the unincorporated triphosphate, the radiolabeled strand was annealed with a five-fold excess of the complementary unlabeled strand. In some cases, 5′ overhangs were present that were filled in as described previously (7) and the 22 to 30-bp duplexes were purified by chromatography (Nensorb, Dupont, Wilmington, Delaware). Unless otherwise specified, sequences containing m^5C were symmetrically (bifilarly) methylated at CpG dinucleotides (7).

Protease treatment and band shift (gel retardation) assays

MDBP (0.5 units per μl) was digested with 1 ng/μl of proteinase K, 100 ng/μl of elastase, or 1 ng/μl of chymotrypsin, subtilisin or papain (Boehringer Mannheim, Indianapolis, Indiana) for 10 min at room temperature. One unit of MDBP, assayed before proteolysis, complexes 1 fmol of the 22-bp pB site 1 duplex (Table 1) under the standard conditions described below. Unless otherwise noted, 3 units of MDBP were used per binding reaction. Aliquots of the partially proteolyzed MDBP were preincubated for 5 min with 400 ng of poly(dI-dC)·poly(dI-dC) (Pharmacia) as a nonspecific competitor. The protease-pretreated MDBP was then incubated for 15 min with 10 fmol of ^{32}P-labeled oligonucleotide duplex. Incubations were in 30 μl of 10 mM Tris-HCl, pH 7.6, 80 mM NaCl, 10% glycerol, 0.2 mM EDTA, 5 mM $MgCl_2$, 2 mM dithiothreitol. The samples were immediately electrophoresed on a 5% polyacrylamide gel as described previously (4). Gels were dried and autoradiographed and complex formation quantitated by cutting out bands and determining their Cerenkov radiation. These bands were the high-mobility (H2-type) complexes, the low-mobility (pB-type complexes) or, in the case of intact MDBP, the very-low-mobility complexes. When specific competitor DNAs were used, they were added at the same time as the poly(dI-dC)·poly(dI-dC). For determination of the half-time for dissociation of the DNAprotein complexes, the ^{32}P-labeled DNA ligand was incubated with intact or partially proteolyzed MDBP as described above except that, after a 60-min

incubation, 1 pmol of unlabeled pB site 1 duplex (for intact MDBP) or HSV site 2 duplex (for partially proteolyzed MDBP) was added. Samples were removed for band shift assays at various time intervals after further incubation at room temperature (8).

Recovery of partially proteolyzed MDBP fragments from MDBP·DNA complexes

MDBP (10 units) was digested with 10 ng of proteinase K for 20 min followed by addition of phenylmethylsulfonyl fluoride to a concentration of 2 mM. The MDBP was then incubated with the ^{32}P-labeled pB site 1 duplex. A band shift assay was performed as described above and a gel slice containing the main complex of proteolyzed MDBP and pB site 1 (pB-type complex) was obtained. The fragmented MDBP in the gel slice was electroeluted and dissociated in 3 M NaCl (Extraphor; LKB, Gaithersburg, Maryland). The eluted MDBP fragments were dialyzed, incubated with 40 fmol of the ^{32}P-labeled HSV site 2 duplex, and complex formation determined again by band shift assay.

Missing base and guanine methylation (dimethyl sulfate) interference assays

Missing base reactions to detect the effect of very limited depurination or depyrimidination of oligonucleotide duplexes on complex formation were performed as described by Brunelle and Schlief (9) except that specific unlabeled competitor DNA was added, just before electrophoresis, only for the reactions with intact MDBP. The effect on complex formation of methylation of G residues was determined by the standard dimethyl sulfate interference reactions as previously described (7). When MDBP partially proteolyzed with proteinase K was used for the binding reactions, the well-resolved H2- and pB-type complexes were separately electroeluted for isolation of their DNA prior to piperidine treatment (9) and analysis on a DNA sequencing gel.

RESULTS

Partially proteolyzed MDBP gave mostly a lower-mobility complex with pB site 1 and a higher-mobility complex with HSV site 2

MDBP was partially digested with proteinase K, a nonspecific protease, for 10 min prior to addition of a specific 22-bp DNA ligand containing pB site 1, HSV site 2, or λ site 1 (Table 1). Upon gel electrophoresis, discrete bands of DNA·protein complexes containing the partially digested MDBP were seen (Fig. 1). The same results were obtained whether or not phenylmethylsulfonyl fluoride (2 mM) was added to the proteinase K-treated MDBP before the DNA. Also, no change in complex formation was seen when the concentration of proteinase K was increased ten fold. Therefore, MDBP has a relatively protease-resistant DNA-binding domain. Much more extensive digestion abolished activity.

Different patterns of complex formation were obtained when the partially fragmented protein was incubated with these three different DNA recognition sites. Such DNA ligand-dependent differences were not seen with the intact protein (Fig. 1A). With the HSV site 2 duplex, one band of a rapidly electrophoresing complex (H2-type complex) was almost the only radiolabeled product using our standard conditions of 3 units of MDBP and 10 fmol of oligonucleotide duplex per 30 μl reaction (Fig. 1A, Table 1). In contrast, when the pB site 1 ligand was used under the same conditions, a more slowly migrating complex (pB-type complex) predominated (Fig. 1A). With a duplex containing λ site 1, mostly the pB-type complex formed although more of the H2-type complex was seen than with pB site 1 as the ligand (Fig. 1A). The same complexes were obtained whether we used the hydroxylapatite fraction of MDBP (1) or MDBP further purified by affinity chromatography on a pB site 1-containing column. Similar results were obtained when

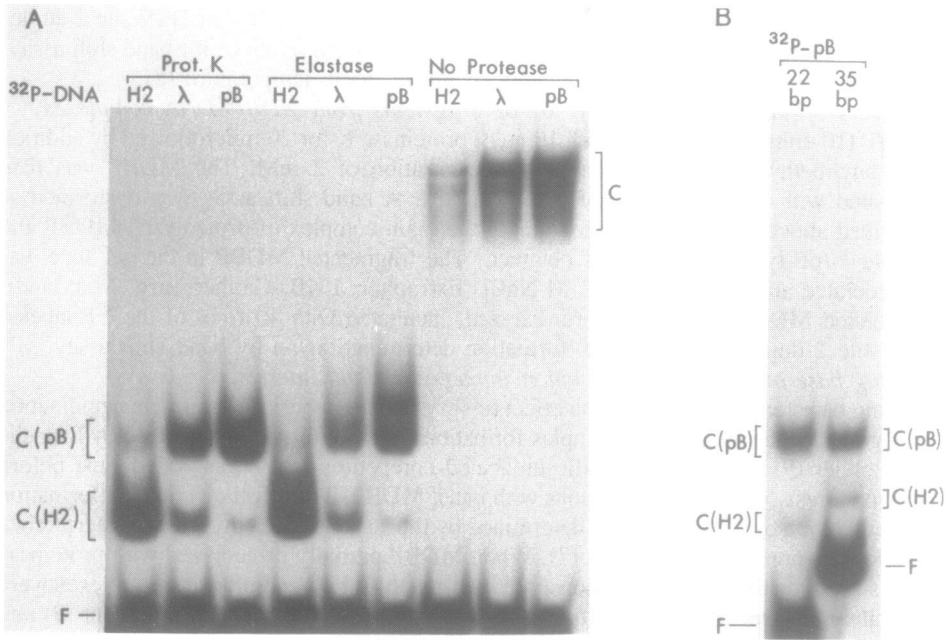


Fig. 1. Two types of complexes formed by partially proteolyzed MDBP and various DNA ligands. MDBP was partially digested with either proteinase K or elastase and incubated with 10 fmol (~ 20,000 cpm) of the indicated ^{32}P-labeled duplex (pB, pB site 1; λ, λ site 1; H2, HSV site 2; Table 1). For partially proteolyzed samples, 5 units of MDBP was used and for native MDBP, 2.5 units were used. (A), only 22-bp duplexes were used. (B), complex formation with the 22-bp pB site 1 duplex was compared to that with a 35-bp duplex W9/C9 (7) that contains pB site 1 in a longer pBR322-derived sequence. C(pB) and C(H2), the two predominant types of complexes formed with partially proteolyzed MDBP and the pB site 1 or HSV site 2 ligands, respectively. C, the position of the family of complexes containing intact MDBP; F, the free DNA band.

elastase, chymotrypsin, papain, or subtilisin replaced proteinase K (Fig. 1 and data not shown). Increasing the size of the pB site 1 ligand from 22 bps to 35 bps only decreased slightly the mobility of the complexes (Fig. 1B).

We also tested two rodent sources of MDBP for formation of H2- and pB-type complexes after limited treatment with protease K. As for human placental MDBP, these rat lung and mouse L cell samples gave only H2-type and pB-type complexes upon proteinase K treatment (data not shown). Also, mostly H2-type complexes were formed with the HSV site 2 duplex and mostly pB-type complexes with pB site 1.

Partial proteolysis of MDBP increased the extent of complex formation by HSV site 2 and decreased it for pB site 1

Not only did partial proteolysis of MDBP result in DNA sequence-dependent differences in the electrophoretic mobility of MDBP·DNA complexes, but also, the relative extent of binding of HSV site 2 and pB site 1 changed greatly as a result of the protein fragmentation. When native MDBP was used, the HSV site 2 duplex was bound only 35% as well as the analogous pB site 1 duplex (Fig. 1). In contrast, after partial proteolysis with proteinase K, papain, elastase, chymotrypsin, or subtilisin, two to five times more binding to HSV site 2 than to pB site 1 was observed (Fig. 2B and data not shown). Quantitation of binding before and after elastase treatment showed that complex formation

from HSV site 2 quadrupled upon partial digestion of MDBP with elastase although complex formation with pB site 1 decreased about three fold.

This reversal in relative binding affinities for HSV site 2 and pB site 1 was also seen in competition experiments. To largely inhibit the binding of fragmented MDBP to [^{32}P]HSV site 2, much more unlabeled pB site 1 was needed than for a comparable reduction in binding to [^{32}P]pB site 1 (Fig. 2A). In contrast, with intact MDBP, half as much unlabeled pB site 1 is needed for 50% competition of [^{32}P]HSV site 2 as for 50% competition of [^{32}P]pB site 1 (6). As expected from studies of intact MDBP (2), unmethylated pB site 1 competed poorly for formation of all complexes (Fig. 2A).

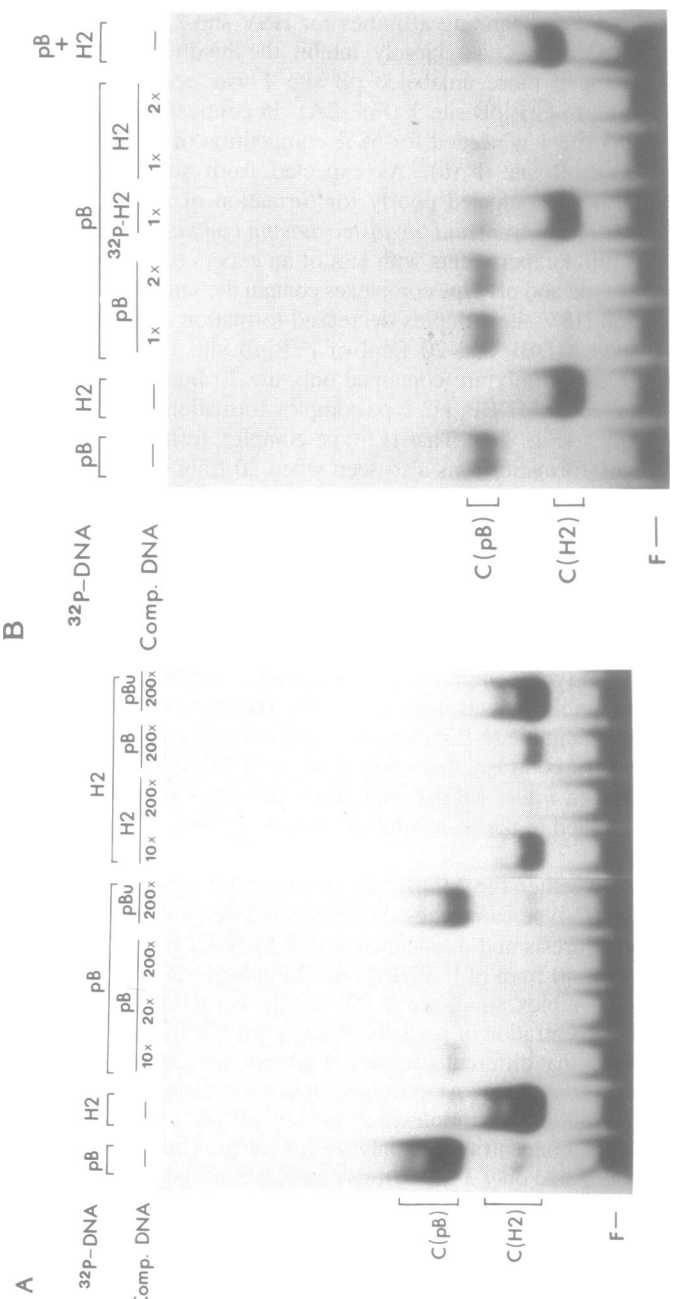
Analysis of whether the slower- and the faster-moving complexes contain the same protein

Additional competition experiments with less of an excess of competitor were conducted to determine if H2-type and pB-type complexes contain the same MDBP fragments. Twenty fmoles of unlabeled HSV site 2 duplex decreased formation of the pB-type complex from proteinase K-treated MDBP and 20 fmol of [^{32}P]pB site 1 duplex by 78% (Fig. 2B). However, if the reaction mixture contained only the 20 fmol of [^{32}P]HSV site 2 duplex and the same fragmented MDBP, H2-type complex formation predominated with less than 5% pB-type complex (Fig. 2B). That H2-type complex formation occurs *at the expense* of pB-type complex formation was also seen when 20 fmol each of [^{32}P]HSV site 2 and [^{32}P]pB site 1 were simultaneously incubated with the fragmented MDBP (Fig. 2B).

Further evidence that the same protein fragments are involved in H2- and pB-type complex formation was seen in many band shift assays in which partially proteolyzed MDBP gave similar doublet bands of H2-type complexes or pB-type complexes, rather than single bands, when complexed to either pB site 1 or HSV site 2 (Fig. 2A). When a partial digest of MDBP gave only one band of H2-type complexes with HSV site 2 as a ligand, also only one band of pB-type complexes was obtained with pB site 1. Similar behavior of pB- and H2-type complexes was also seen in the complete resistance of their formation to 10 mM N-ethylmaleimide (NEM) in the incubation buffer even in the absence of added sulfhydryl reagent. In contrast, this concentration of NEM largely abolished formation of complexes between *intact* MDBP and either pB site 1 or HSV site 2 unless 0.1 M dithiothreitol was added to the incubation mixture to counteract the effects of NEM (data not shown).

Also, we tested whether the MDBP fragments in pB-type complexes are capable of efficiently forming H2-type complexes. We recovered the protein from pB-type complexes isolated by electrophoresis and dissociated it in 3 M NaCl from the pB site 1 ligand. As expected, addition of 40 fmol of [^{32}P]HSV site 2 duplex to the dialyzed protein recovered from the pB site 1 complex now gave >90% of the complexes of the H2-type (Fig. 3).

By varying the concentration of partially proteolyzed MDBP used for the binding assays, we obtained evidence that different numbers of protein molecules are involved in formation of H2- and pB-type complexes. At extremely low concentrations of partially proteolyzed MDBP, as much as half of the complexes formed by pB site 1 were of the H2 type although at moderate and high concentrations, only 4−8% of the complexes were of the H2 type (Fig. 4A). HSV site 2 had only 3−5% of its complexes as the pB type at low and moderate concentrations of partially proteolyzed MDBP whereas up to 21% of the complexes were of the pB type at high concentrations (Fig. 4B). When the highest studied concentration of proteolyzed MDBP was incubated with 10-fold less HSV site 2 duplex to further increase the protein-to-DNA ratio, pB-type complexes constituted 25% of the total complexes. The protein titration curve for pB-type complex formation with pB site 1 showed a much lower


A
32P-DNA
Comp. DNA
pB
H2
pB
pB
10×
20×
200×
pBu
200×
H2
H2
10×
200×
pB
200×
pBu
200×
C(pB)
C(H2)
F—
B
32P-DNA
Comp. DNA
pB
H2
pB
pB
1×
2×
32P-H2
1×
H2
1×
2×
pB
+
H2
C(pB)
C(H2)
F —

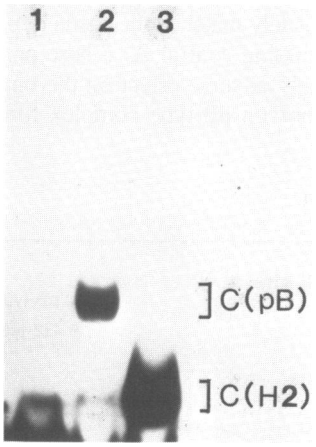


Fig. 3. Ability of the protein fragments from pB-type complexes to form H2-type complexes. Lane 1, fragmented MDBP from the pB-type complex with pB site 1 was recovered as described in *Materials and Methods* and incubated with 40 fmol of [^{32}P]HSV site 2, and electrophoresed in lane 1. Lanes 2 and 3, analogously proteolyzed MDBP directly incubated with 10 fmol of [^{32}P]pB site 1 or HSV site 2, respectively, to serve as markers. The lower activity of the MDBP fragments in lane 1 is due to losses during its recovery. The same single band of H2-type complex was seen in two repetitions of this experiment and in a protein sample recovered from H2-type complexes and then re-challenged with ^{32}P-labeled HSV site 2. The top of the bands of uncomplexed oligonucleotide duplex is seen at the bottom of the autoradiogram.

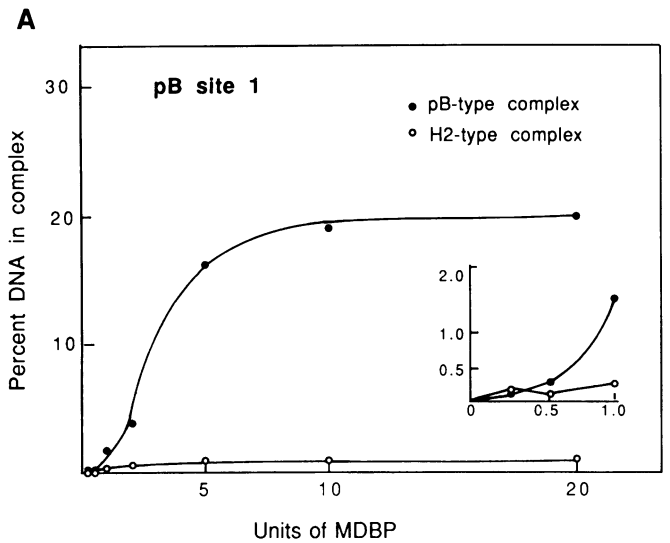
plateau value for the extent of complex formation than did that for H2-type complex formation from HSV site 2 (Fig. 4A and B), which is consistent with the hypothesis that more molecules of fragmented MDBP are necessary for pB-type than for H2-type complex formation.

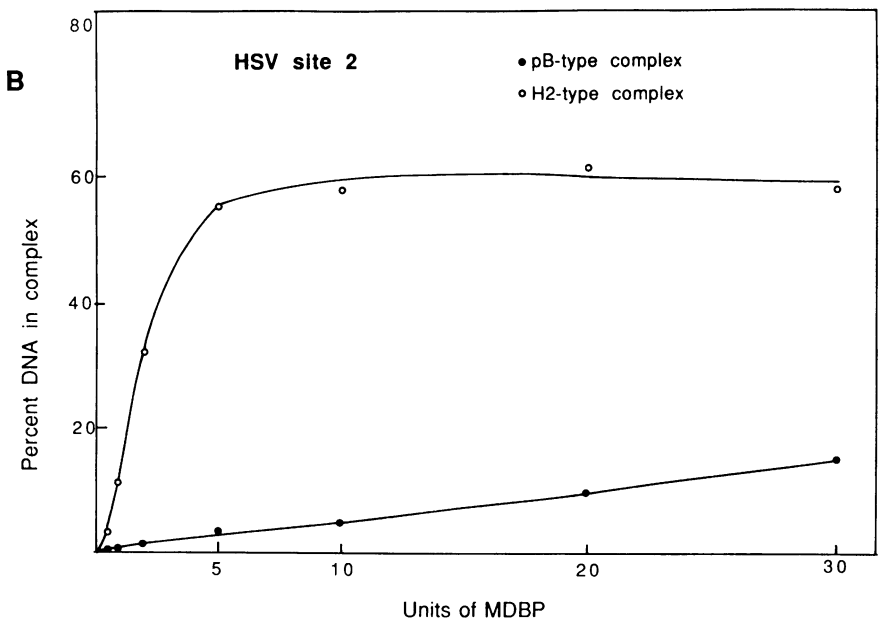
Complex formation between fragmented MDBP and other DNA sites

We tested other MDBP-specific oligonucleotide duplexes for binding to proteinase K-treated MDBP (Table 1). All of these ligands, including hu site 1, an anonymous human site (6), and CMV sites 1 and 2 (Zhang *et al.*, unpublished data) in an enhancer of human cytomegalovirus (11, 15), formed only pB-type and H2-type complexes (Fig. 5; Table 1). Using constant concentrations of fragmented MDBP and DNA duplex, we found that the relative amounts of these two types of complexes varied from ligand to ligand. In general, more pB-type complex formation with fragmented MDBP was seen with recognition sites

Fig. 2. Specific competition for formation of the two types of complexes. MDBP was fragmented with proteinase K and assayed as in Fig. 1 except that specific DNA competitors were present. These 22-bp competitors were unlabeled unless otherwise specified. The indicated competitor was used for a 5-minute preincubation with fragmented MDBP before the addition of radiolabeled ligand. A form of the 22-bp pB site 1 duplex (Table 1) in which none of C residues were methylated (pBu) and which, therefore, is not recognized by MDBP (2), was used as a competitor for some of the samples. The two types of complexes formed are indicated as in Fig. 1. (A), specific unlabeled competitors, in a 10- to 200-fold molecular excess over the 10 fmol of radiolabeled pB site 1 or HSV site 2 (Table 1), were present in all but the first two lanes. (B), 20 fmol of [^{32}P]pB site 1 was competed with twice as much (2×) or equal amounts (1×) of the indicated unlabeled duplex. In the fifth lane, 20 fmol of [^{32}P]HSV site 2 duplex (~20,000 cpm) was incubated with fragmented MDBP for 5 min before the addition of 20 fmol of [^{32}P]pB site 1 duplex of equal specific activity. In the last lane, 20 fmol of the two [^{32}P]duplexes were added simultaneously at the beginning of the 15-min incubation period and in the first two lanes they were added separately.

that completely or almost completely matched the partially palindromic consensus sequence (Table 1). Surprisingly, the missing central A·T base-pairs in λ site 1 and HSV site 4 (Table 1) did not have a strongly adverse effect on the binding of these sites to intact or partially proteolyzed MDBP nor on pB-type complex formation. However, deletion of





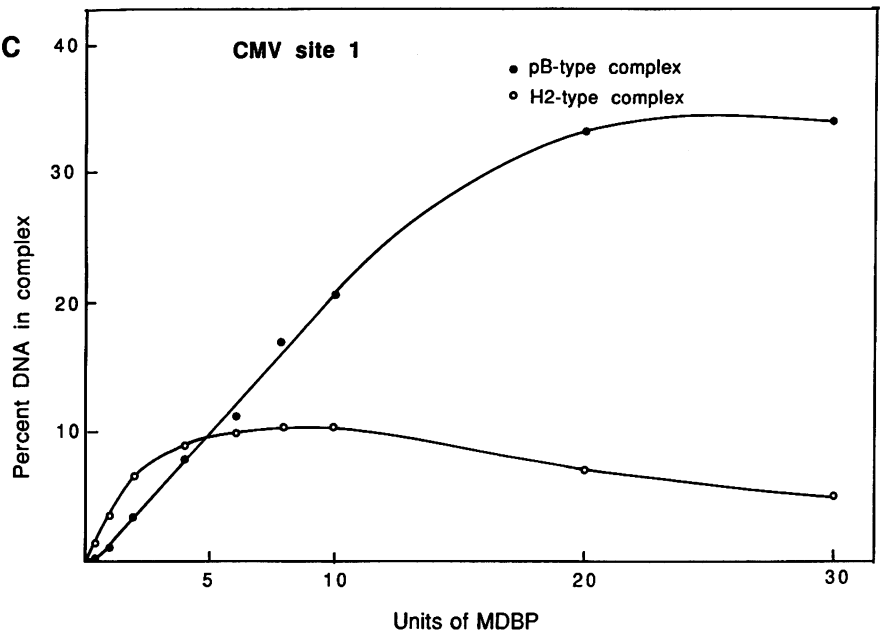


Fig. 4. Effect of varying the concentration of partially proteolyzed MDBP on formation of pB- and H2-type complexes. The percentage of oligonucleotide duplex which formed pB-type complexes or H2-type complexes was quantitated by band shift assay with 10 fmol each (~70,000 cpm) of ^{32}P-labeled pB site 1 (A), HSV site 2 (B), or CMV site 1 (C) and proteinase K-treated MDBP. Note the change of scale in the ordinate for panel B. The insert in panel A shows an enlargement of the first part of the curves.

this A·T base-pair is tolerated only in certain limited sequence contexts because deletion of the same base-pair from pB site 1 abolishes binding of MDBP (6).

In addition to the degree of homology to one strand of the consensus sequence, optimal methylation of pyrimidine residues at the three critical m^{5}YpR dinucleotides (m^{5}Y, m^{5}C or T) on each strand was important in determining the relative amounts of the two types of complexes formed. The DNA sequence which gave the largest percentage of its complex of the pB type was pB site 1 (bifilarly methylated). This sequence was the only one listed in Table 1 with 100% homology to the consensus sequence and three m^{5}YpR dinucleotides per strand in the 14-bp recognition sequence. When only the three CpG dinucleotide pairs in the bottom strand of pB site 1 were methylated, the percentage of complexes that were of the H2-type increased from 8% to 52% due to a decrease in pB-type complex formation with almost no change in the extent of H2-type complex formation. Consistent with this conclusion was the finding that two other MDBP sites with 100% homology to the consensus sequence and 3 out of 3 or 2 out of 3 of the conserved bifilar m^{5}YpR dinucleotides (5′-ATMGCCAMGGMGAT-3′ and 5′-GTTGTTATAGTAAC-3′) gave almost as high a percentage of their complexes of the pB type as did pB site 1 (data not shown).

The nature of the H2- and pB-type complexes containing CMV site 1 was examined by DNA competition, protein titration, and dissociation experiments. With 1 pmol of unlabeled pB site 1 as a competitor, the formation of pB-type complexes by 10 fmol of [^{32}P]CMV site 1 decreased to less than 5% the level without competitor while H2-type

Table 1. Complex formation between different MDBP binding sites and partially proteolyzed MDBP

DNA site (source)	Sequence[1]	Binding to intact MDBP	% of total complex formed[2] H2-type	pB-type
	1 5 10 14			
HSV site 2 (TK gene)	gtcg G T T G C T A•T G G C C G C gaga	++	96	4
pB site 1 (pBR322)	ctag A T M G T C A•M G G M G A T ctag	+++	8	92
CMV site 2 (enhancer)	cgct A T T A C C A•T G G T G A T gcgg	+++	20	80
λ site 1 (λ phage)	ctc G T T G C C -•C G G T A A C aacagc	+++	30	70
HSV site 5 (DBP gene)	cagg G G M G C C A•M G G M G A C gtcg	++	36	64
hTGF-β site 1 (TGF gene)	tggagtgt G T T G C T A•T G G T G A C tggagtgt	+++	37	63
CMV site 1 (enhancer)	tggc G T T A C T A•T G G G A A C atac	+++	45	55
hu site 1 (human DNA)	tgtg G T T A C C A•T G G G A A T gaca	+++	59	41
HSV site 4 (exonuc. gene)	acccg G T T T C C -•C G G T T A C cagt	+	65	35
HSV site 3 (α-TIF gene)	tcgac G M G G C C A•M G G M G A C tcga	+	71	29

[1]Sequences are written 5' to 3' for one strand of the oligonucleotide duplex. The MDBP recognition sequence is capitalized and its three conserved positions containing m^5CpG or TpA are underlined. Bases deviating from the consensus sequence are indicated with an "x"; a hyphen is inserted when necessary for optimal alignment. A dot denotes the dyad axis. Except for pB site 1 and hTGF-β site 1, the bases adjacent to the recognition site are those found in the naturally occurring sequence. pB site 1, HSV sites 2 and 3, λ site 1, and hu site 1 were described but for two of these the opposite strand is shown (2,6). hTGF-β site 1 is present at position +2392 relative to the beginning of a human transforming growth factor β1 cDNA (10). CMV sites 1 and 2 in an enhancer of the human cytomegalovirus and HSV sites 4 and 5 of the HSV type 1 alkaline exonuclease and major DNA binding protein genes are at positions -430, -227, -230, and +415, respectively, relative to the transcription initiation site (11-14).

[2]Duplexes (10 fmol) were tested in band shift assays with 3 units of MDBP partially hydrolyzed by proteinase K. + and ++, ~7-25 and ~25-60%, respectively, as much binding of unproteolyzed MDBP to the duplex as seen for the sequences designated +++. The relative extents of complex formation with partially proteolyzed MDBP were similar except for HSV site 2 and pB site 1 as described in Results. H2-type complexes are of high electrophoretic mobility and pB-type of lower mobility.

complex formation from the same ligand was still 22% as much as without competitor (Fig. 5) even though the identical DNA ligand was involved in formation of both kinds of complexes. Similar results were obtained with ^{32}P-labeled λ site 1 or hu site 1. When the concentration of proteinase K-treated MDBP was varied from the standard 3 units per 30 μl in assays with CMV site 1, mostly pB-type complexes were formed at the high protein concentrations and mostly H2-type complexes at very low concentrations (Fig. 4C). As determined by dissociation experiments (data not shown), the half-life for pB- and H2-type complexes of partially proteolyzed MDBP and CMV site 1, pB site 1 and HSV site 2 were all <1 min. This is to be compared to half-lives of ~ 100 min, 90 min, and 5 min for complexes of *intact* MDBP with the respective sites.

Missing base and G methylation analysis of the interaction of CMV site 1 with fragmented MDBP

Because CMV site 1 under standard assay conditions formed nearly equal amounts of H2- and pB-type complexes with fragmented MDBP (Table 1), it was used as a ligand to determine which of its bases make important contacts with the protein in the formation of the two types of complexes. MDBP partially proteolyzed with proteinase K was incubated with a CMV site 1-containing duplex that had been pretreated so as to remove a very low level of purines or pyrimidines (missing base analysis; 9) or just to methylate a very small percentage of G residues (16). If loss of a specific base or N7 methylation of a specific

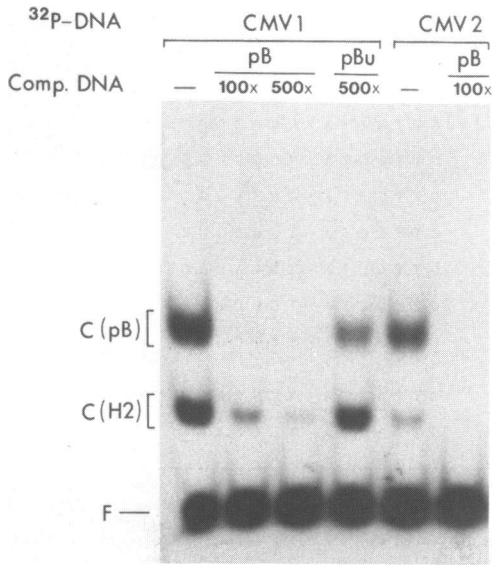


Fig. 5. Binding of sites from a cytomegalovirus enhancer to partially proteolyzed MDBP in the presence or absence of specific competitors. Proteinase K-treated MDBP was incubated with 10 fmol of ^{32}P-labeled 22-bp duplexes containing either of two cytomegalovirus enhancer sites for MDBP (CMV site 1 and CMV site 2, Table 1) with or without specific competitors described in the legend to Figure 2. In a separate gel, the pB-type and H2-type complexes formed from these CMV sites were shown to coelectrophorese with those from pB site 1 and HSV site 2.

G residue in this DNA sequence interferes with binding because that base makes a critical contact with the DNA-binding domain of the protein, then there will be little or no pB-type or H2-type complex formation from that modified DNA molecule. This will be revealed by the loss of a corresponding band in a DNA sequencing gel containing DNA recovered from a given type of DNAprotein complex isolated by the band shift assay.

The results from these analyses of CMV site 1 suggest that seven identical bases on the upper strand and up to eleven on the lower strand contact fragmented MDBP for both H2- and pB-type complex formation (Figs. 6 and 7). However, in the pB-type complex, five additional bases on the lower strand and six on the upper strand appear to be involved in complexation with fragmented MDBP. With the exception of two bases in the lower strand, all of the bases whose removal decreased complex formation with fragmented MDBP are within the previously assigned 14-bp recognition sequence (6). When complex formation with intact MDBP was subjected to these analyses, the results obtained were similar to those for pB-type complex formation with partially proteolyzed MDBP (data not shown).

Involvement of different bases of HSV site 2 and pB site 1 in complex formation

Missing base analysis was also performed on the upper strand of a 22-bp ligand containing HSV site 2. This analysis indicated that removal of the following underlined bases largely inhibited H2-type complex formation with proteinase K-treated MDBP: 5′-gtcgGTTGCTA·TGGCCGCgaga-3′ (data not shown). Loss of the bases with a dot underneath gave only ~40−60% decreases in complex formation. The dot in the middle of the sequence demarcates the axis of hyphenated dyad symmetry. The bases in lower

Table 2. Effect of mutations at HSV site 2 and CMV site 1 on complex formation with MDBP

Name of site	Sequence[1]	% Complex formation[2] Intact MDBP	Fragmented MDBP H2-type	Fragmented MDBP pB-type[3]
HSV site 2 (wild type)	G T T G C T A · T G G C C G C (x x x)	100	100	-
HSV site 2/mutant 1	G T T G C T A · T G G C C G G	77	118	-
HSV site 2/mutant 2	G T T G C T A · T G T C C G C	42	56	-
HSV site 2/mutant 3	G T T G C T A · T G G C G G C	256	80	-
HSV site 2/mutant 4	G T T G C T A · T G G G C G C	164	99	-
HSV site 2/mutant 5	G C T G C T A · T G G C C G C	10	7	-
HSV site 2/mutant 6	C T T G C T A · T G G C C G C	3	4	-
CMV site 1 (wild type)	G T T A C T A · T G G G A A C (x)	100	100	100
CMV site 1/mutant 1	G T T A C T A · T G G G A A G	29	146	17

[1]Oligonucleotide duplexes containing 22 bps (Table 1) were used as ligands in band shift assays. The sequence of only one strand of the 14-bp recognition site region is shown here. A dot denotes the axis of hypenated dyad symmetry. The bases of the wild-type sequences differing from the consensus sequence are indicated by "x's." The base-pairs that were mutated for this experiment by synthesis of variant oligonucleotide duplexes are underlined.

[2]The percent complex formation under standard conditions (Materials and Methods) was compared for mutant and the corresponding wild type duplexes.

[3]From the HSV site 2 and its variants, the slight extent of pB-type complex formation (Table 1) was not quantitated. However in the autoradiograms, it was apparent that more pB-type complex formed from HSV mutant 3, than from wild type HSV site 2 or the other variants. Wild type CMV site 1 formed similar amounts of pB- and H2-type complexes under these standard assay conditions (Table 1).

case letters are outside the region of the 14-bp consensus sequence. G-methylation (16) and hydroxyl radical (17) interference analyses also indicated that the left half of this 14-bp MDBP site is much more important than the right for H2-type complex formation except that, in the latter assays, a decrease in complex formation was seen when the sugar-phosphate backbone in the beginning of the right half of this site (5′-TGG-3′) was attacked (data not shown). Missing base and G-methylation interference assays on pB site 1 showed that its formation of pB-type complexes was decreased by alteration of the following underlined positions in the lower strand of the 14-bp recognition site: 3′-TAGMAGT·GMCGMTA-5′ (data not shown). Positions whose alteration largely blocked complex formation are underlined; less heavily implicated positions have a dot beneath them. For technical reasons, the involvement of the 3′ T is unclear. It can be seen that both halves of pB site 1 are of much importance for pB-type complex formation in accord with both halves of this site completely matching the consensus sequence. In contrast, only the left half of HSV site 2, the one which is completely homologous to the consensus sequence, was implicated in making specific base contacts during formation of H2-type complexes.

To further study the importance of the two halves of HSV site 2, mutant sites were used as ligands for binding to fragmented or intact MDBP. Single transversions at three different positions (mutants 1, 3, 4) in the 7 bps of the right half of HSV site 2 resulted in little or no decrease in complex formation with either intact or partially proteolyzed MDBP. The change at one of these positions (mutant 3) made the sequence closer to the consensus sequence and resulted in a large increase in binding of intact MDBP and in the appearance of a small amount of pB-type complex formation with partially proteolyzed MDBP (Table 2). In the same right half, a change at a position which was closer to the

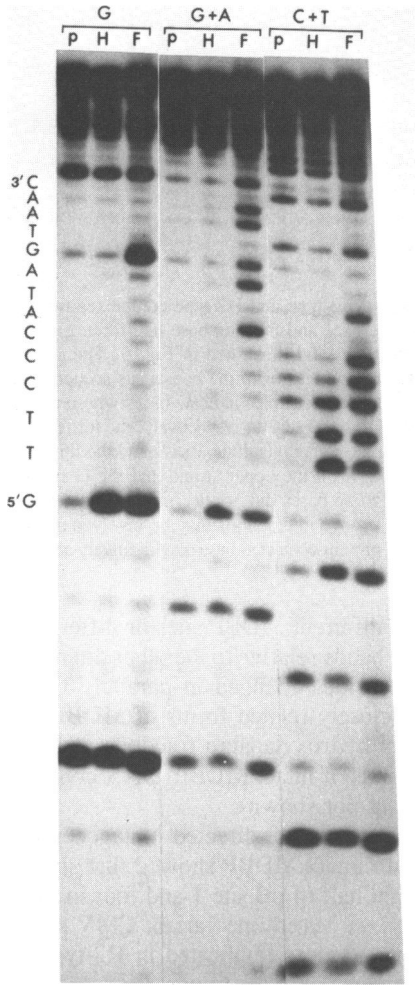


Fig. 6. Dimethyl sulfate interference analysis and missing base interference analysis of the interaction of CMV site 1 with fragmented MDBP. Very limited depurination (G + A lanes), depyrimidination (C + T lanes), or G methylation (G lanes) reactions were performed on a 30-bp oligonucleotide duplex (see Fig. 7) whose lower strand was labeled at its 5'-phosphate terminus. As described in *Materials and Methods*, the reacted duplex was incubated with the proteinase K-treated MDBP and either pB-type (p) or H2-type (H) complexes isolated prior to electrophoresis of the extracted DNA on a sequencing gel. As a control, reacted DNA was electrophoresed on the sequencing gel without being subject to complex formation (F, free DNA). The sequence of the MDBP recognition site is indicated next to the corresponding bands.

dyad axis (mutant 2) approximately halved complex formation with intact or partially proteolyzed MDBP (Table 2). In contrast to the mutations in the right half, each of two mutations in the left half of HSV site 2 drastically reduced binding of both intact and partially proteolyzed MDBP (Table 2).

With intact MDBP a family of a few bands of complexes are usually seen (Fig. 1A).

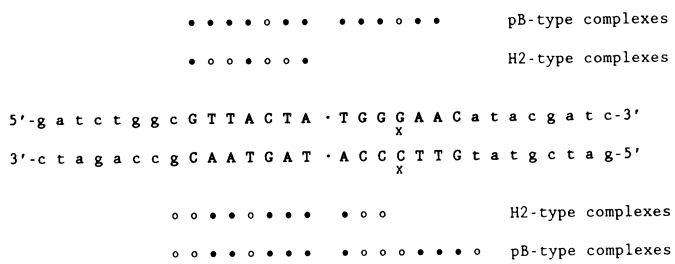


Fig. 7. Base contacts for formation of pB-type and H2-type complexes from partially proteolyzed MDBP and CMV site 1. Dimethyl sulfate interference and missing base interference assays were conducted on CMV site 1 as described in *Materials and Methods* and illustrated in Fig. 6. The canonical 14-bp recognition sequence is capitalized and its single base-pair deviating from the consensus sequence denoted by 'x's.' Its dyad axis is shown by two dots. Filled-in circles and open circles indicate bases whose removal or, in the case of G residues, N7 methylation, decreased by $>80\%$ or $\sim 40-60\%$, respectively, the formation of pB-type complexes or H2-type complexes. Results for G residues in the dimethylsulfate assays and the missing base assays were similar except for the G residue with an 'x' beneath it on the upper strand and the G residue 3' to the capitalized sequence on the lower strand. Methylation of these bases did not decrease pB-type complex formation but their removal did.Minor groove rather than major groove interactions may be involved at these positions or, alternatively, the loss of a non-contacted base next to a protein-contacted base may indirectly affect binding of the protein by changing the DNA conformation.

Just as previously seen with different MDBP sites or different mutant MDBP sites (6,7), the intensity of each of these bands relative to the other did not vary with the above mutant ligands but rather these intensities changed in parallel. This suggests that each of the complexes is derived from closely related forms of MDBP with identical DNA-binding sites. Furthermore, different hydroxylapatite fractions of MDBP, which differed in the relative intensity of these bands of intact MDBP·DNA complexes, gave the same results upon partial proteolysis (data not shown).

Previous studies of the effects of site-directed mutagenesis and methylation of pB site 1 on complex formation with intact MDBP showed that single base-pair changes at five different positions in the right half of pB site 1 and four in the left half strongly inhibited binding (5, 7). In this study, we tested one variant CMV site 1 having a mutation at the far end of its right side, the side less implicated in H2-type complex formation than in pB-type complex formation (Fig. 7). This transversion mutant showed a large reduction in binding intact MDBP compared to the analogous non-mutant duplex although considerable binding was still seen. However, with proteinase K-treated MDBP, there was no decrease in H2-type complex formation compared to the nonmutant site although pB-type complex formation was largely inhibited (Table 2).

DISCUSSION

MDBP, a sequence-specific DNA-binding protein novel in its dependence on cytosine methylation for recognition of certain sites, has a high affinity for a remarkable variety of related 14-bp sites (Table 1; 3). All of these have much homology to the degenerate consensus sequence 5'-RTm^{5}YRYYAm^{5}YRGm^{5}YRAY-3' (m^{5}Y, m^{5}C or T; 6). We found that MDBP formed complexes with a much higher electrophoretic mobility when isolated from rat or bovine tissues than when it is extracted from human placenta or cultured human or rodent cells (4). Nonetheless, the specificity of all of these samples with respect to DNA sequence and DNA methylation was indistinguishable leading us to suspect that the higher

mobility resulted from limited proteolysis during extract preparation from rat and bovine tissues without much loss of activity (4,18). In the present study, we showed that rat lung MDBP partially digested with proteinase K still bound specifically to oligonucleotide duplexes containing pB site 1, an m^5C-containing pBR322 sequence, or HSV site 2, a herpes virus DNA sequence, and gave complexes indistinguishable from those of human placental MDBP. Therefore, this proteolytic clipping band-shift assay (19) supports our hypothesis that there were not inherent differences in rodent tissue MDBP and human placental MDBP.

Surprisingly, this study revealed major differences in the mobility of the predominant DNAprotein complex when MDBP partially proteolyzed *in vitro* was incubated with different specific oligonucleotide duplexes although no such differences were observed when intact MDBP was used (Table 1; Figs. 1 and 5). This was seen most dramatically in a comparison of pB site 1 and HSV site 2, which form complexes predominantly of the pB (low-mobility) type and H2 (high-mobility) type, respectively. That HSV site 2 formed pB-type complexes so poorly and yet competed very well even at low concentrations for pB-type complex formation (Fig. 2) indicates that H2-type complexes contain an MDBP fragment necessary for pB-type complexes. Furthermore, fragmented MDBP recovered from pB-type complexes containing pB site 1 could form H2-type complexes with HSV site 2 (Fig. 3).

We propose that *in vitro* proteolysis of MDBP at some highly exposed region has separated the DNA-binding domain from a subunit junction domain (20–23) and that the low-mobility, pB-type complexes contain two truncated subunits and the higher-mobility, H2-type complexes only one (Fig. 8). Preliminary experiments involving gel filtration and UV-crosslinking of MDBP and pB site 1 followed by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate suggest that intact MDBP, if globular, has a molecular weight in the range of ~250–300 kDa and consists of two subunits of ~130–150 kDa (data not shown). The model of 2-subunit or 1-subunit complex formation by partially proteolyzed MDBP is consistent with higher concentrations of fragmented MDBP giving a much higher percentage of total complexes that are of the pB (2-subunit) type (Figs. 4A and 4C). Also, the shape of the curves for complex formation with CMV site 1 and different amounts of fragmented MDBP (Fig. 4C) suggests that H2-type complexes are intermediates on the pathway to pB-type complexes (Fig. 8). This model explains why a given concentration of specific competitor decreases pB-type (2-subunit) complex formation to a greater extent than H2-type (1-subunit) complex formation (Figs. 2 and 5) although *intact* MDBP has a much higher affinity for pB site 1 than for HSV site 2. Also consistent with the model is the finding that most of the bases in both halves of the partly palindromic CMV site 1 or pB site 1 were important for pB-type complex formation (Figs. 7 and 8 and *Results*). In contrast, for formation of H2-type complexes, mostly just those on the left half of CMV site 1 or HSV site 2 were implicated. Analogously, only the left half of HSV site 2 conforms well to the consensus sequence whereas the two partially symmetrical halves of pB site 1 exactly match the consensus sequence (Table 1). CMV site 1, which shows an intermediate behavior to that of HSV site 2 and pB site 1 as a ligand for partially proteolyzed MDBP, matches the consensus sequence completely in the left half and at 6 bps out of 7 in the right half (Table 1). As predicted by our model (Fig. 8), a transversion mutation in the right half of CMV site 1, to give a total of two non-consensus base-pairs in this half, largely inhibited formation of pB-type complexes without decreasing formation of H2-type complexes (Table 2).

pB site 1
(bifilarly methylated)
fragmented MDBP
very unstable
1-subunit complex
(H2-type complex)
2-subunit complex
(pB-type complex)
HSV site 2
XXX
fragmented MDBP
1-subunit complex
(H2-type complex)
CMV site 1
X
fragmented MDBP
1-subunit complex
(H2-type complex)
2-subunit complex
(pB-type complex)
any MDBP
recognition site
intact MDBP
intact dimer
complex
Intact MDBP·DNA complexes are
much more stable than the analogous
fragmented MDBP·DNA complexes.

We propose that a single truncated MDBP subunit bound to pB site 1 is much less stable than that bound to HSV site 2 but that two such separated, fragmented subunits can cooperatively form a rather stable complex with pB site 1 (Fig. 8). High cooperativity in binding of two subunits to pB site 1 would explain the difficulty in obtaining H2-type as opposed to pB-type complexes from this site (24) and the fact that there is only a modest decrease in the extent of complex formation with pB site 1 upon partial proteolysis of MDBP. This cooperativity is apparently dependent on bifilar methylation of pB site 1 as seen in the sharp decline in formation of pB-type complexes, but not of H2-type complexes when pB site 1 was methylated on only one strand instead of on both strands. Our model also explains why, at high concentrations, pB site 1 was a much worse competitor for complexing limiting amounts of fragmented MDBP than was HSV site 2, hTGF-β site 1, or CMV site 1 (Fig. 2A and data not shown), all of which can form H2-type (1-subunit) complexes better than pB site 1 can.

With *intact* MDBP, the non-consensus half of HSV site 2 may play a role in binding although less so than the consensus half. This is evidenced by the matching of four bases in this half to the consensus sequence (Table 1) and the finding that methylation of two G residues in this half interfered with binding of intact MDBP (6). Furthermore, more factors than just the one subunitone half-site interaction appear to be involved in complex formation between intact MDBP and HSV site 2. The half-life of complexes of partially proteolyzed MDBP and HSV site 2 is less than 1 min whereas that of the corresponding complexes containing intact MDBP is $\sim$5 min. That the half-life of complexes of *intact* MDBP with pB site 1 is almost 20 times longer than those with HSV site 2 and that the extent of complex formation of *intact* MDBP with pB site 1 is greater than with HSV site 2 suggests the importance of a strong interaction between each of the two DNA half-sites and the two putative subunits of native MDBP. This interaction could be stabilized by a tight association of the subunits (Fig. 8).

A number of sequence-specific DNA-binding proteins have been shown to bind to rather disparate DNA sequences (25–27). Based on this study, we propose an explanation for how pB site 1 and HSV site 2 are both recognized specifically by MDBP although they share only 5 bps out of 14 in common. One half-site of HSV site 2 is hypothesized to

Fig. 8. A simplifed model for the interaction of partially proteolyzed or intact MDBP with its DNA recognition sites. The partly palindromic MDBP recognition sequences are indicated by two rectangles each representing a 7-bp half-site. Base-pairs within the half-site that deviate from the consensus sequence are symbolized by 'x's.' The MDBP fragment containing the DNA-binding domain released from the rest of the protein by partial proteolysis is drawn as a circle. This figure shows the predominant complexes formed at moderate ratios of oligonucleotide duplex to fragmented MDBP (Table 1). Intact MDBP is depicted as a dimer whose subunits are non-covalently joined outside the DNA-binding domain; however, there may be some heterogeneity in the subunits outside the DNA-binding domain because a family of several bands rather than a single band is seen in band shift assays with intact MDBP (Fig. 1). According to the model, predominant formation of H2-type complexes with fragmented MDBP would generally indicate a weaker interaction of an MDBP recognition site with intact MDBP due to one subunit of the intact protein interacting much less than the other with the partially palindromic recognition site.

interact strongly with one of two subunits of intact MDBP whereas, although each half-site of pB site 1, in comparison, is poor by itself, together they may establish highly favorable cooperative interactions with intact MDBP (Fig. 8). This makes pB site 1 a better ligand for intact MDBP and HSV site 2 a better ligand for partially proteolyzed MDBP, which appears to have been converted to separated truncated subunits as a result of the proteolysis. A similar strategy of relying predominantly on a one half-site interaction with certain DNA sequences and on nearly equal interaction with both half-sites of others may be used by other sequence-specific DNA-binding proteins recognizing fairly long, divergent DNA sequences. Cloning such a protein from an expression library with a DNA ligand probe (28,29) that shows a strong half-site interaction with the protein could obviate the need for dimerization of the cloned, truncated recombinant protein subunits.

ACKNOWLEDGEMENT

This research was supported in part by USPHS grant GM33999.

REFERENCES

1. Huang, L.-H., Wang, R., Gama-Sosa, M.A., Shenoy, S., and Ehrlich, M. (1984) Nature *308*, 293–295.
2. Wang, R.Y.-H., Zhang, X.-Y., and Ehrlich, M. (1986) Nucleic Acids Res. *14*, 1599–1614.
3. Wang, R.Y.-H., Zhang, X.-Y., Khan, R., Zhou, Y., Huang, L-H., and Ehrlich, M. (1986) Nucleic Acids Res. *14*, 9843–9860.
4. Supakar, P.C., Weist, D., Zhang, D., Inamdar, N., Zhang, X.-Y., Khan, R., Ehrlich, K.C., and Ehrlich, M. (1988) Nucleic Acids Res. *16*, 8029–8046.
5. Zhang, X.-Y., Ehrlich, K.C., Wang, R.Y.-H., and Ehrlich, M. (1986) Nucleic Acids Res. *14*, 8387–8397.
6. Zhang, X.-Y., Supakar, P.C., Khan, R., Ehrlich, K.C., and Ehrlich, M. (1989) Nucleic Acids Res. *17*, 1459–1474.
7. Khan, R., Zhang, X.-Y., Supakar, P.C., Ehrlich, K.C., and Ehrlich, M. (1988) J. Biol. Chem. *263*, 14374–14383.
8. Riggs, A.D., Bourgeois, S., and Choa, M. (1970) J. Mol. Biol. *53*, 401–417.
9. Brunelle, A. and Schleif, R.F. (1987) Proc. Natl. Acad. Sci. USA *84*, 6673–6676.
10. Derynck, R., Jarrett, J.A., Chen, E.Y., Eaton, D.H., Bell, J.R., Assoian, R.K., Roberts, A.B., Sporn, M.B., and Goeddel, D.V. (1985) Nature *316*, 701–704.
11. Boshart, M., Weber, F., Jahn, G., Dorsch-Hasler, K., Fleckenstein, B., and Schaffner, W. (1985) Cell *41*, 521–530.
12. Costa, R., Draper, K., Banks, L., Powell, K., Cohen, G., Eisenberg, R., and Wagner, E. (1983) J. Virol. *48*, 591–603.
13. Draper, K.G., Devi-Rao, G., Costa, R.H., Blair, E.D., Thompson, R.L., and Wagner, E.K. (1986) J. Virol. *57*, 1023–1036.
14. Quinn, J.P. and McGeoch, D.J. (1985) Nucleic Acids Res. *13*, 8143–8163.
15. Ghazal, P., Lubon, H., and Henninghausen, L. (1988) Molec. Cell. Biol. *8*, 1809–1811.
16. Seibenlist, U. and Gilbert, W. (1980) Proc. Natl. Acad. Sci. USA *77*, 122–126.
17. Chalepakis, G. and Beato, M. (1989) Nucleic Acids Res. *17*, 1783.
18. Mather, E.L. (1988) Molec. Cell. Biol. *8*, 1812–1815.
19. Schreiber, E., Matthias, P., Muller, M.M., and Schaffner, W. (1988) EMBO J. *7*, 4221–4229.
20. Pabo, C.O., Sauer, R.T., Sturtevant, J.M., and Ptashne, M. (1979) Proc. Natl. Acad. Sci. USA *76*, 1608–1612.
21. Sauer, R.T., Pabo, C.O., Meyer, B.J., Ptashne, M., and Backman, K.C. (1979) Nature *279*, 396–400.
22. Geisler, N. and Weber, K. (1977) Biochemistry *16*, 938–943.
23. Tsai, S., Caristedt-Duke, J., Weigel, N.L., Dahlman, K., Gustafsson, J.-A., Tsai, M.-J., and O'Malley, B.W. (1988) Cell *55* 361–369.
24. Tsai, S., Tsai, M.-J., and O'Malley, B.W. (1989) Cell *57*, 443–448.
25. Pfeifer, K., Prezant, T., and Guarente, L. (1987) Cell *49*, 19–27.
26. Hwung, Y.-P., Wang, L.-H., Tsai, S.Y., and Tsai, M.-J. (1988) J. Biol. Chem. *263*, 13470–13474.
27. Costa, R.H., Grayson, D.R., Xanthopoulos, K.G., and Darnell, J.E. (1988) Proc. Natl. Acad. Sci. USA *85*, 3840–3844.

28. Singh, H., LeBowitz, J.H., Baldwin, A.S., and Sharp, P.A. (1988) Cell *52*, 415–423.
29. Vinson, C.R., LaMarco, K.L., Johnson, P.F., Lanschulz, W.H., and McKnight, S.L. (1988) Genes Develop. *2* 801–806.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.