Probing the Protein−DNA Interface of the *Eco*RV Modification Methyltransferase Bound to Its Recognition Sequence, GATATC†

Mark D. Szczelkun,‡,§ Helen Jones,‡ and Bernard A. Connolly*,‖

Department of Biochemistry, The University of Southampton, Southampton SO9 3TU, U.K., and Department of Biochemistry and Genetics, The University of Newcastle, Newcastle upon Tyne NE2 4HH, U.K.

Received February 15, 1995; Revised Manuscript Received May 31, 1995⊗

ABSTRACT: The DNA contacts produced between the *Eco*RV modification methyltransferase and its recognition sequence, GATATC, have been determined. The enzyme's general location in a methylase/DNA/sinefungin ternary complex was evaluated by protection from exonuclease III digestion. Important phosphate contacts were resolved using *N*-ethyl-*N*-nitrosourea ethylation interference footprinting. Methylation protection and interference using dimethyl sulfate were employed to assess significant contacts to purinic bases. The protein−DNA interface was further probed using oligodeoxynucleotides containing base analogues within the GATATC sequence. Most of the experiments were carried out using hemimethylated sequences, i.e., having 6-methyladenosine at the methylation site in one of the strands. The monomeric methylase was found to bind to the DNA in two different orientations for the methylation of each strand. The enzyme approaches the DNA, predominantly from one "side", and makes most of its contacts in the major groove. In either of the two binding events contacts are made to the four phosphates NpNpNpGpA and the three bases GAT (where GAT represents the 5′ half of the GATATC site) on both DNA strands. The phosphates and bases in the 3′ ATC half are much less important. Although the contacts made to the equivalent locations on each strand are similar, they display a slight but consistent change dependent on which strand contains the 6-methyldeoxyadenosine. This strand variation shows completely reciprocal behavior, switching around exactly, depending entirely on the methylated deoxyadenosine location. It is this that provides evidence for the two binding modes. The results obtained are discussed in terms of possible models for the protein−DNA interface.

The elucidation of the important individual nucleoside and phosphate contacts made during protein−DNA association is an invaluable step toward rationalizing specific DNA sequence recognition. The definitive method is X-ray crystallography of protein−DNA complexes and a growing number of such structures are available. These include the *Hha*I DNA dC methylase (Klimasauskas *et al.,* 1994) and the *Eco*RV endonuclease (Winkler *et al.,* 1993) as well as several repressors and activators [see reviews by Harrison (1991) and Pabo and Sauer (1992)]. Other solution spectroscopic techniques, especially multidimensional NMR, have also been used, e.g., antennapedia homeodomain (Billeter *et al.,* 1993) and *Escherichia coli lac* repressor (Chuprina *et al.,* 1993). However, many DNA binding proteins are difficult to crystallize or have low solubility and so are not readily amenable to study by these two methods. The *Eco*RV modification methylase falls into this class, and in these cases other techniques must be used. Perhaps the most useful approach, in these instances, involves the elegant set of methods first described by Galas and Schmitz (1978) as DNA footprinting. The technique relies on the premise that a tightly bound protein will alter the sensitivity of nucleotides to chemical or enzymatic modification or that a prior modification will directly interfere with protein association. An alternative solution approach involves the use of analogues of the four normal bases (Aiken & Gumport, 1991). Here the kinetic or binding properties of a protein are evaluated with an oligonucleotide containing a modified base and the parameters seen compared to those found with unsubstitued DNA. Examples have included studies with the *Eco*RI endonuclease (Brennan *et al.,* 1986a; McLaughlin *et al.,* 1987; Lesser *et al.,* 1990, 1993), *Taq*I endonuclease (Zebala *et al.,* 1992), and the *Eco*RV endonuclease (Newman *et al.,* 1990a,b; Waters & Connolly, 1994). There have been a few studies with methylases, e.g., *Eco*RI (Brennan *et al.,* 1986b; Reich & Danzitz, 1991, 1992) and *Eco*RV (Newman *et al.,* 1990a). It is generally agreed that base analogues are most useful when additional structural data are available and on its own the method is not fully reliable at predicting protein−DNA contacts.

The *Eco*RV methyltransferase forms part of a prokaryotic restriction modification system and in conjunction with its partner endonuclease serves to protect bacteria from viral attack (Wilson & Murray, 1991; Wilson, 1992; Bickle & Krüger, 1993). The 33 kDa methylase recognizes the sequence d(GATATC), adding a CH_3 group to the 6-NH_2 group of the first dA (Nwosu *et al.,* 1988) and thereby making host nucleic acid resistant to hydrolysis by the cognate endonuclease. In the cell the actual substrate for the methylase is hemimethylated DNA that transiently arises following DNA replication and cell division. Addition of CH_3 groups to such molecules gives the fully dimethylated

† This work was supported by the Molecular Recognition Initiative of the SERC. M.D.S. and H.J. were SERC-supported Ph.D. students. B.A.C. is a fellow of the Lister Institute of Preventive Medicine.

* Address correspondence to this author.

‡ The University of Southampton.

§ Current address: Department of Biochemistry, School of Medical Sciences, University Walk, Bristol BS8 1TD, U.K.

‖ The University of Newcastle.

⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1995.

0006-2960/95/0434-10734$09.00/0

product and so restores and maintains the methylation pattern. However, the *Eco*RV methylase is also capable of *de-novo* methylation of unmethylated DNA (Nwosu *et al.*, 1988; Newman *et al.*, 1990a) as indeed are many DNA methylases such as *Eco*RI (Rubin & Modrich, 1977) and *Eco* dam (Herman & Modrich, 1982). The majority of DNA methylases, for example, *Eco*RI (Rubin & Modrich, 1977), *Hpa*II (Yoo & Agarwal, 1980), *Eco* dam (Herman & Modrich, 1982), and *Fok*I (Sugisaki, *et al.*, 1989), are active as monomers, and this includes *Eco*RV (Nwosu *et al.*, 1988) [for a review see Wilson (1992)]. It is generally accepted that methylases transfer a CH_3 group to one strand of duplex DNA per binding event, with dissociation of the enzyme from the DNA prior to any subsequent methylation (Rubin & Modrich, 1977; Herman & Modrich, 1982). Thus, as first pointed out in these two papers from Modrich's group, one of the most interesting features of this class of enzyme is that they normally recognize an asymmetric sequence (due to hemimethylation). Furthermore, the binding of a monomeric enzyme to duplex DNA that is either asymmetric (hemimethylated) or even fully symmetric (unmethylated) would in itself be expected to generate asymmetry. This is in contrast with the more commonly studied DNA sequence-specific proteins such as restriction endonucleases or repressors where a homodimeric protein interacts with palindromic DNA (or near palindromic DNA) giving a fully symmetrical interaction. It is thus expected that the methylation of each of the strands will proceed from a different unique binding orientation. Support for this postulate has been demonstrated for the *Eco*RI methylase using a pre-steady-state kinetic analysis (Reich & Mashhoon, 1993). Only one methylase–DNA crystal structure has been reported, that for *Hha*I (Klimasauskas *et al.*, 1994), which shows that the enzyme makes contacts to the target bases and phosphates of both strands. The interactions with the phosphates on each of the strands are very different, and the target dC bases are also handled differently with the one to be methylated "flipped" out of the helix. This certainly agrees with the idea of unique binding orientations. Recently, asymmetry on binding to palindromic DNA has also been observed for the dC methylase *Msp*I (Dubey & Roberts, 1992) using DNase I footprinting. However, it is not clear how much of what has been observed with the dC methylases will apply to the dA class, and structural information for the latter group is lacking. Recently (Labahn *et al.*, 1994), the structure of the dA methylase *Taq*I bound to AdoMet was reported. The distance from the cofactor to the putative DNA binding site was large, and this suggested that these methylases might also "flip" out their target base. However, confirmation of this awaits ternary complex structure determination.

To help resolve the question of how dA methylases interact with their target sites and to provide evidence for or against unique binding orientations, we have carried out a detailed footprinting analysis of the *Eco*RV methylase. This is possible as our previous paper had shown specific binding to GATATC sites. Most of the work has been carried out using ternary complexes produced from the enzyme, hemimethylated DNA, and sinefungin, a nonmethylating analogue of the essential cofactor AdoMet. Both enzymatic methods using exonuclease III and chemical approaches with ethylation and methylation interference and protection have been employed. The results give some of the first structural data for a dA methylase bound to a DNA substrate. While this paper was being considered for publication, a methylation interference study of the type I dA methylase *Eco*KI appeared (Powell & Murray, 1995). The results found indicate that the *Eco*RV methylase binds to DNA in two similar, but definitely distinct, orientations for CH_3 addition to each of the strands.

MATERIALS AND METHODS

*The purification of the Eco*RV *methylase and the synthesis* and purification of the oligonucleotide substrates were described fully in the preceding paper. Oligonucleotides were either radioactively labeled at their 5′-ends using polynucleotide kinase and [γ-^{32}P]ATP or at their 3′-ends with terminal deoxynucleotidyl transferase and [α-^{32}P]ddATP.

Electrophoresis Protocol. All footprinting experiments were analyzed by an identical electrophoretic method. Lyophilized DNA pellets were dissolved in deionized formamide to a final concentration of about 10 counts/s^{-1} μL^{-1}. The fragments were analyzed by urea denaturing gel electrophoresis [15 × 20 × 0.03 cm, 50% (w/v) urea, using TBE (89 mM Tris–borate, pH 8.0, 1 mM EDTA) as the running buffer]. The gel was prerun for 1 h at 20 W constant power. The DNA solutions were boiled for 5 min and cooled on ice, and then 5 μL aliquots were loaded. Maxam and Gilbert G reactions (see below) were run alongside the footprinting lanes, and bromophenol blue was used as a size marker. The gels were run for ≈3 h at 20 W constant power or until the bromophenol blue band was two-thirds down the gel. Gels were fixed in 10% (v/v) acetic acid/10% (v/v) methanol for 15 min and dried using a gel dryer. Dry gels were exposed to X-ray film (Fuji-RX) with an intensifying screen (HI Speed X) at −70 °C overnight. The amount of material in each band was analyzed by densitometry using a Joyce-Lobel ChromoScan fitted with a 530 nm filter and a 1.0 mm aperture width. Peak heights were taken as indicative of the amount of material in each band. The bands were internally normalized to take into account different loading and cleavages by selecting a band whose density was unaffected by methylase binding (Lu *et al.*, 1981). For phosphate ethylation the standards used for this normalization were the phosphate between dG23 and dC24 for the 30-mer and that between dG12 and dC13 with the 33-mer. For methylation protection the dG residues at positions 22 and 12 were used as the normalization standards with the 30- and 33-mers, respectively. Bands were identified by alignment with comigrating fragments produced by Maxam–Gilbert dG reactions.

Footprinting Experiments. Oligonucleotides labeled at the 5′-terminus were used for exonuclease III digestions and methylation protection. For ethylation and methylation interference 3′-labeling was used. The limited methylation of dG residues to produce standard dG ladders and dG tracks in methylation protection and interference experiments was based on the protocol of Maxam and Gilbert (1980) as described in detail by Sambrook *et al.* (1989). Methylation protection and interference using dimethyl sulfate were carried out as described (Siebenlist & Gilbert, 1980; Shaw & Stewart, 1994). For the protection experiments 0.1 pmol of 5′-^{32}P-labeled oligonucleotide and 10 pmol of methylase were incubated in 40 μL of 50 mM Hepes, pH 7.5, containing 100 mM NaCl, 10 mM EDTA, 5 mM DTT, 2 μg of acetylated bovine serum albumin, and 10 mM sinefungin for

30 min at 20 °C. Subsequently, 2 μL of 5% dimethyl sulfate was added and incubation continued at 37 °C for 5 min. For the interference experiments 0.1 pmol of 3′-^{32}P-labeled oligonucleotides modified at dG residues was incubated with 10 pmol of the methylase in the above buffer for 20 min at room temperature. The resulting enzyme−DNA complexes and free DNA were separated by band-shift electrophoresis as detailed in the preceding paper. Ethylation interference with *N*-ethyl-*N*-nitrosourea was performed as described (Siebenlist & Gilbert, 1980; Manfield & Stockley, 1994). 3′-^{32}P-labeled oligonucleotides (0.1 pmol) modified at phosphates were incubated with 10 pmol of methylase in the above buffer and free and bound DNA separated by gel-shift electrophoresis. Protocols for exonuclease III protection were as detailed in Metzger and Heumann (1994). For a typical reaction, 0.2 pmol of 5′-^{32}P-labeled oligonucleotide and 16 pmol of the methylase were incubated in 80 μL of 50 mM Hepes, pH 7.5, 5 mM DTT, 5 mM $MgCl_2$, 85 mM NaCl, 4 μg of acetylated bovine serum albumin, and 10 mM sinefungin for 30 min at 20 °C. DNA was digested with 1 μL (3 units) of exonuclease III at 20 °C, and 20 μL samples were removed at 1, 5, 10, and 15 min. The reactions were stopped by addition of 480 μL of 50 mM EDTA−Tris, pH 6.0, containing 25 μg of sonicated calf thymus DNA.

Methylation of Oligonucleotides Containing Base Analogues. The preparation of self-complementary 12-mers, based on the parent sequence GACGATATCGTC, containing the dC analogues 5-methyldeoxycytidine (d^{5Me}C) and 2-pyrimidinone 1-β-2′-deoxyriboside (d^{4H}C) and the dG analogues 7-deazadeoxyguanosine (d^{7C}G), 6-thiodeoxyguanosine (d^{6S}G), 2-aminopurine 1-β-D-2′-deoxyriboside (d^{am}P), deoxyinosine (dI), and 3-deazadeoxyguanosine (d^{3C}G) at the underlined positions has been described (Waters & Connolly, 1994). The synthesis of the complementary oligonucleotides AGGG^{6Me}ATATCGGA and TCCGATATCCCT containing the combinations of d^{7C}G and dU, outlined in the Results section, was exactly analogous. The conditions used for the evaluation of the rate of methylation of the self-complementary base analogue substituted 12-mers by an HPLC-based assay have also been published (Newman *et al.*, 1991a). Rates with the AGGG^{6Me}ATATCGGA/TCCGATATCCCT-based duplexes were evaluated similarly using HPLC, but the triethylammonium acetate buffer used previously was replaced by ammonium acetate which gave better resolution. A gradient consisting of 2−25% CH_3CN over 25 min in 0.1 M ammonium acetate was used. The order of elution observed (and approximate elution times) was AGGG^{6Me}-ATATCGGA (and its derivatives) (14 min), TCCGATATC-CCT (and its derivatives) (15.5 min), and TCCCG^{6Me}-ATATCCCT (and its derivatives) (16 min).

RESULTS AND DISCUSSION

Oligonucleotides Used. In the accompanying publication it was demonstrated that a GATATC-containing duplex, formed from a complementary 30- and 33-mer, was specifically bound by the *Eco*RV methylase. The introduction of 6-methyldeoxyadenosine into one of the strands to give a hemimethylated duplex resulted in the strongest recognition by the enzyme. DNA containing a d^{6Me}A base on one strand only is thought to be the physiological substrate for the methylase, in agreement with its tight binding, and in view of this we have carried out most of our footprinting experiments with hemimethylated duplexes. The use of these

Table 1: Oligodeoxynucleotides Used in This Study[a]

STANDARD		*1 10 Me 20 30*
30-MER/	5′	AAAGTCTGTGGATATCCAAGTGGCTACCGT 3′
33-MER	3′	TCAGACACCTATAGGTTCACCGATGGCACCCCC 5′
		30 20 10 1
STANDARD		*1 10 20 30*
30-MER/	5′	AAAGTCTGTGGATATCCAAGTGGCTACCGT 3′
33-MER	3′	TCAGACACCTATAGGTTCACCGATGGCACCCCC 5′
		30 Me 10 1
		20
STANDARD		*1 10 20 30*
30-MER/	5′	AAAGTCTGTGGATATCCAAGTGGCTACCGT 3′
33-MER	3′	TCAGACACCTATAGGTTCACCGATGGCACCCCC 5′
		30 20 10 1
EXTRA dG (5′)		*1 10 Me 20 30*
30-MER/	5′	AAAGTCTGGGGATATCCCCGTGGCTACCGT 3′
33-MER	3′	TCAGACCCCTATAGGGGCACCGATGGCACCCCC 5′
		30 20 10 1
EXTRA dG (3′)		*1 10 Me 20 30*
30-MER/	5′	AAAGTCTCCCGATATCGGGGTGGCTACCGT 3′
33-MER	3′	TCAGAGGGCTATAGCCCCACCGATGGCACCCCC 5′
		30 20 10 1
CONTROL		*1 10 20 30*
30-MER/	5′	AAAGTCTGTGTACGTACAAGTGGCTACCGT 3′
33-MER	3′	TCAGACACATGCATGTTCACCGATGGCACCCCC 5′
		30 20 10 1

[a] The standard 30/33-mers are complementary and form the duplexes shown which contain a GATATC *Eco*RV site. A 6-methyldeoxyadenosine residue could be placed at the methylation site in either strand to give a hemimethylated substrate. The two extra dG-containing oligonucleotides allow methylation interference and protection experiments to probe the environment of the enzyme-bound oligonucleotide at regions immediately flanking the GATATC site. The control oligonucleotide has the GATATC site replaced with one containing TACGTA.

oligonucleotides also allows us to address questions associated with the possibility of two unique binding orientations of the enzyme facilitating the addition of CH_3 groups to the target dA bases on each strand. Of practical importance is that the two strands can be separately labeled with ^{32}P, allowing the interactions of the methylase with both to be evaluated. The oligonucleotides we have used are shown in Table 1 and consist of the two 30/33-mer duplexes with the d^{6Me}A in either the 30 or the 33 base strand. In addition, we have utilized alternative 30/33-mers which contain extra dG residues to either the 3′ or the 5′ direction of the GATATC sequence. These only contain d^{6Me}A in the 30-mer strand. This allows any influences that flanking bases have on methylase binding to be probed using dimethyl sulfate methylation intreference. Finally, some experiments have been carried out with the fully unmethylated standard 30/33-mer and a control oligonucleotide in which the GATATC *Eco*RV site has been replaced by TACGTA.

Exonuclease III Protection. Exonuclease III digestion was utilized to approximate the binding locus of the methylase (Shalloway *et al.*, 1980; Metzger & Heumann, 1994). This enzyme's 3′→5′ exonuclease activity processively degrades double-stranded DNA simultaneously from both ends. Tightly bound proteins block the passage of the nuclease, resulting in a resistant core which delineates the protein's binding site. The ideal exonuclease III concentration was determined empirically by digesting the hemimethylated duplexes formed from the standard 30/33-mers in the absence of the methylase. Conditions were selected such that only digestion-resistant species remained after 15 min incubation (Figure 1). These bands correspond to approximately half the length

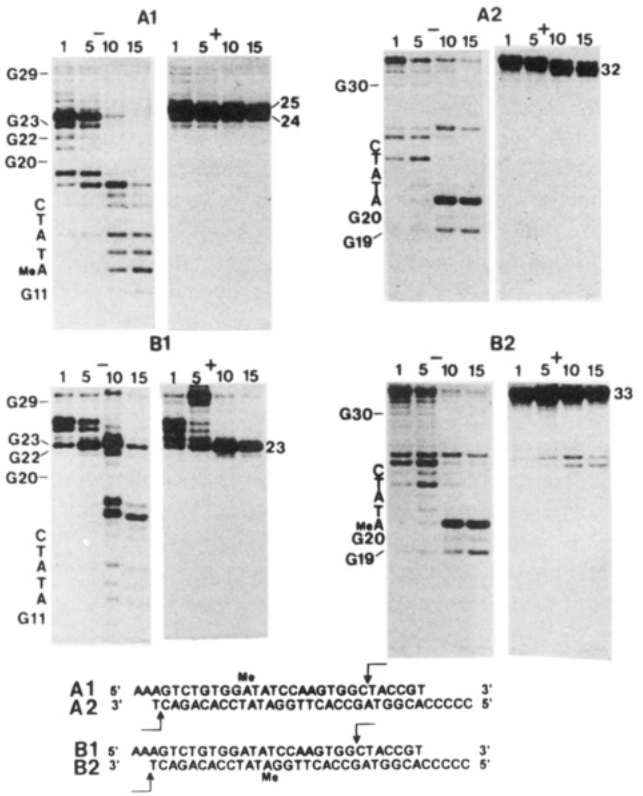

FIGURE 1: Exonuclease III protection assay. Effect of digesting various combinations of oligodeoxynucleotides, either alone (−) or in the presence of the *Eco*RV methylase and sinefungin (+), for 1, 5, 10, and 15 min with exonuclease III: (A) AAAGTCTGTG-G^{6Me}ATATCCAAGTGGCTACCGT (A1) plus CCCCCACGG-TAGCCACTTGGATATCCACAGACT (A2); (B) AAAGTCT-GTGGATATCCAAGTGGCTACCGT (B1) plus CCCCCACGG-TAGCCACTTGG^{6Me}ATATCCACAGACT (B2). Individual bands were located by coelectrophoresis with Maxam−Gilbert dG sequencing reactions (not shown). The dG bases so identified are illustrated on the right-hand side of the gels as are the GATATC *Eco*RV sites. The main stop sites in the presence of the methylase are shown on the right-hand side of the gels. The footprint determined in each case is also given.

of the oligonucleotide, at which point the DNA becomes single stranded or there is a reciprocal steric blockage of the two nuclease molecules coming from opposite ends of the DNA (Shalloway *et al.*, 1980). An unusually large number of other "stops" were produced as a result of the breakdown of processive cleavage, probably due to the relatively small size of the oligonucleotide substrate. The results of digesting a DNA/methylase/sinefungin complex are also shown in Figure 1, and clear protection of the hemimethylated duplexes from exonuclease III hydrolysis is observed. This is visualized as discrete and stable stops that are distinct from the "half-site" stops seen in the absence of methylase. Digestion of the protected DNA for a further 15 min did not affect the position of the stop bands at the concentration of exonuclease III used. When this experiment was repeated with the control oligonucleotide, TACTGTA/TACGTA, the cleavage ladders produced were identical regardless of whether or not the methylase was present (results not shown). The stability of the exonuclease footprint and its location to the GATATC recognition site strongly suggest that the band shifts observed in the preceding paper do indeed represent specific binding of the *Eco*RV methylase to its target site. The lack of a footprint seen in the controls agrees with the absence of any binding to a nonspecific oligonucleotide, inferred by the band-shift assay in the previous paper. This also confirms that the results obtained with the GATATC sequences do indeed represent specific binding. Although the footprint is centered on the GATATC sequence, it encompasses almost the entire length of the oligonucleotide and is particularly close to the 3′-end of the 33-mer strand in each substrate. For the methylated strand of the 30-mer/methylated 33-mer (B1) the protected region is actually at the end of the strand. It is thus possible that the methylase projects beyond this base, but given the smaller footprint seen by chemical methods, we consider this unlikely. The protected site is symmetrically arranged around the GATATC sequence. This compares to the asymmetry of the 18 base pair DNase I protection site for the *Msp*I methylase (Dubey & Roberts, 1992), where the recognition sequence is displaced by several nucleotides 5′ to the median of the footprint. The protected region is the same size for G^{6Me}ATATC/GATATC and GATATC/G^{6Me}ATATC (i.e., d^{6Me}A in the 30- and 33-mer, respectively) and 21 nucleotides in length. It should not be concluded that the methylase actually spans this large number of bases as footprinting using nucleases usually overestimates the size of binding sites due to the bulk of the two proteins involved. Thus with the *Eco*RI endonuclease a footprint of 17 base pairs was seen using nucleases (Fox, 1988), 12 base pairs with hydroxyl radicals (Dixon *et al.*, 1991), and 8 base pairs with phosphate ethylation interference (Lu *et al.*, 1981; Lesser *et al.*, 1990). In comparison, crystallography shows 8 base pairs in close proximity to the protein (Kim *et al.*, 1990). Of greatest interest in this study is that the protected region shifts by one base pair dependent on the orientation of the methylated half-site. Thus when the d^{6Me}A is in the 30-mer strand, the 21 nucleotide protected region is displaced one base pair to the right as compared to the situation seen when the methylated base is in the 33-mer strand (Figure 1). This provides the first indication that there are two distinct binding orientations of the methylase, for the addition of CH_3 groups to each strand.

Ethylation Interference. Further detailed information on the interaction of the *Eco*RV methylase with DNA was obtained using ethylation interference with *N*-ethyl-*N*-nitrosourea to ethylate phosphates (Siebenlist & Gilbert, 1980; Manfield & Stockley, 1994). The results obtained for each strand of the hemimethylated duplexes formed from the standard 30/33-mer and with the d^{6Me}A in both orientations are shown in Figure 2. With fragments >15 bp in length, the 5′-hydroxyl and 5′-ethyl phosphate doublets resulting from cleavage of the modified phosphate are not resolved and run as a single band. However, with fragments of 15 bp or less, the doublets resolve into two closely running bands. Increased activity at dG and dA residues (Sun & Singer, 1975; Affolter *et al.*, 1990), independent of phosphate ethylation, was also observed (e.g., at position dG8 for the enzyme-bound pool in Figure 2, gel A). As can be seen from Figure 2 the strongest inhibition to binding arises when one of four contiguous phosphates on each strand are ethylated. This is much more easily observed following quantitation of the bands in the gels by scanning as shown in Figure 3. This figure clearly shows that modification at two symmetrical sets of four phosphates prevents binding. These are located at the 5′-ends (NpNpNpGpA, where N = bases to the 5′ of the GATATC site) of both strands of the duplex. The interference pattern is independent of the

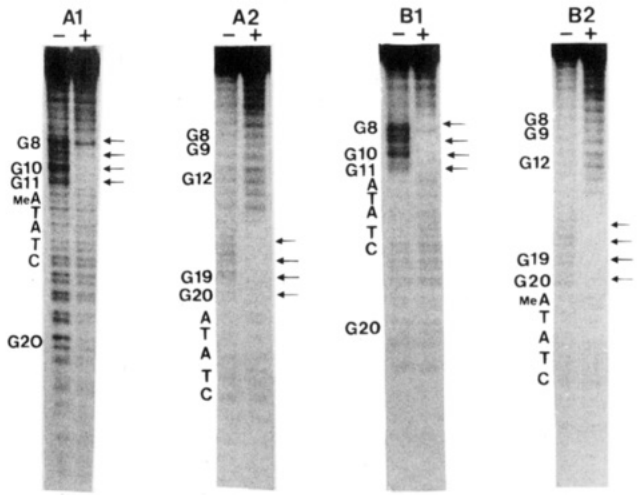

FIGURE 2: Phosphate ethylation interference analysis of the various *Eco*RV/sinefungin/DNA complexes. Free and bound DNA isolated from gel-shift analysis are labeled (−) and (+) respectively. (A) AAAGTCTGTGG^{6Me}ATATCCAAGTGGCTACCGT (A1) plus CCCCCACGGTAGCCACTTGGATATCCACAGACT (A2); (B) AAAGTCTGTGGATATCCAAGTGGCTACCGT (B1) plus CCCCCACGGTAGCCACTTGG^{6Me}ATATCCACAGACT (B2). Individual phosphates were located by coelectrophoresis with Maxam–Gilbert dG sequencing reactions (not shown). The phosphate bands migrate slightly more slowly than their corresponding dG bands. Doublets are clearly visible for the smaller fragments due to cleavage both 3′ and 5′ to the modified phosphate. The phosphates located by dG sequencing are indicated as are the GATATC sites. Strongly interfering phosphates are arrowed.

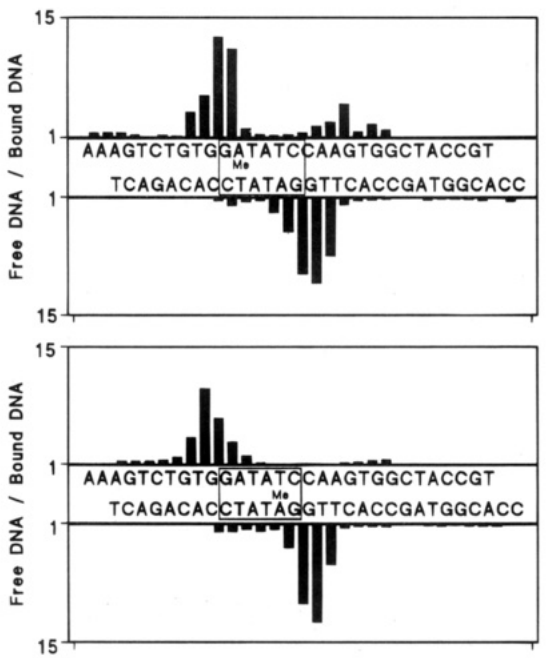

FIGURE 3: Quantitation of the gels presented in Figure 2. The phosphates are not indicated, but the bars are located between bases to represent the result seen for the phosphate between the two bases. The last three dC bases of the 33-mer strand have been omitted from the diagram (ratios of 1 were obtained here).

orientation of the hemimethylated site unlike the situation observed for exonuclease III protection. Large amounts (easily enough to fully complex unmodified DNA) of methylase were used in the reactions with the modified DNA, and so the resulting footprint suggests that alkylation of a single one of the above phosphates is an absolute block to binding. The absence of interference at the other phosphate positions within the *Eco*RV recognition sequence supports the view that the actual contacts are probably functionally significant and not due merely to steric hindrance. However, the possibility that interference is due to alkylation-induced conformational transitions of the DNA helix cannot be absolutely ruled out (Lu *et al.*, 1981; Gartenberg *et al.*, 1990). Smaller effects were also observed at other phosphates, e.g., at positions 19 and 20 in the d^{6Me}A strand for the methylated 30-mer/unmethylated 33-mer duplex (figure 3). However, the exact significance of these rather marginal results is unclear, and so they are not discussed further. The two sets of four, strongly interfering, phosphates give a DNA footprint 11 bases in length. The smaller size of the footprint relative to the exonuclease III data is typical, and the chemical as opposed to the enzymological results are probably a better reflection of the true binding site. It is also interesting that some of the important contacts are actually outside the GATATC *Eco*RV site, but this has been observed before with, for example, the *Eco*RI endonuclease (Lu *et al.*, 1981; Lesser *et al.*, 1990).

Methylation Interference. Methylation interference experiments, using dimethyl sulfate to modify the N7 and N3 sites of dG and dA, respectively (Siebenlist & Gilbert, 1980; Shaw & Stewart, 1994), gave the results shown in Figure 4 with the hemimethylated 30/33-mers. The results seen with dA bases, which react much more poorly with dimethyl sulfate than dG, were not reproducible and are not discussed. It proved very difficult to determine the ratios of free DNA/bound DNA by densitometric analysis of the gels shown in Figure 4. This arises because the methylation of the dG, within the GATATC site, completely abolishes protein binding and thus gives no band to scan in the + (enzyme-bound) lanes. This indicates that this dG is critical for methylase binding. Furthermore, the behavior of this particular dG is identical irrespective of whether it is on the same or the opposite strand to the d^{6Me}A. These oligonucleotides also contain an additional dG residue, immediately 5′ to the dG in the recognition hexamer, and methylation here also influences enzyme binding to the nucleic acid. Interference here is less pronounced than for the dG within the recognition sequence but is sensitive to the location of the d^{6Me}A. Thus, as can be seen by a visual inspection of Figure 4, modification of the flanking dG in GGA sequences diminishes methylase binding rather strongly, whereas its reaction in GG^{6Me}A runs has a much smaller effect. The interference patterns seen with each hemimethylated substrate are related by rotational symmetry around the dyad axis, and again this might be indicative of two different binding orientations of the enzyme to permit CH_3 addition to each strand. Support for this comes from the results seen with the unmethylated 30/33-mer, which are also given in figure 4. This shows that the interference at the dG residues, 5′ external to the *Eco*RV site, are now similar for both strands due to the absence of ^{6Me}A. The strong interference with the dG in the GATATC site could indicate a direct interaction, i.e., a hydrogen bond, between the N7 of this base and the protein or alternatively a close approach subject to steric hindrance. The flanking dG, which gives weaker interference, is unlikely to directly hydrogen bond to the enzyme as in this case the methylase would recognize the sequence GGATATCC. This would lead to methylation of only a subset of *Eco*RV sites and incomplete protection from the

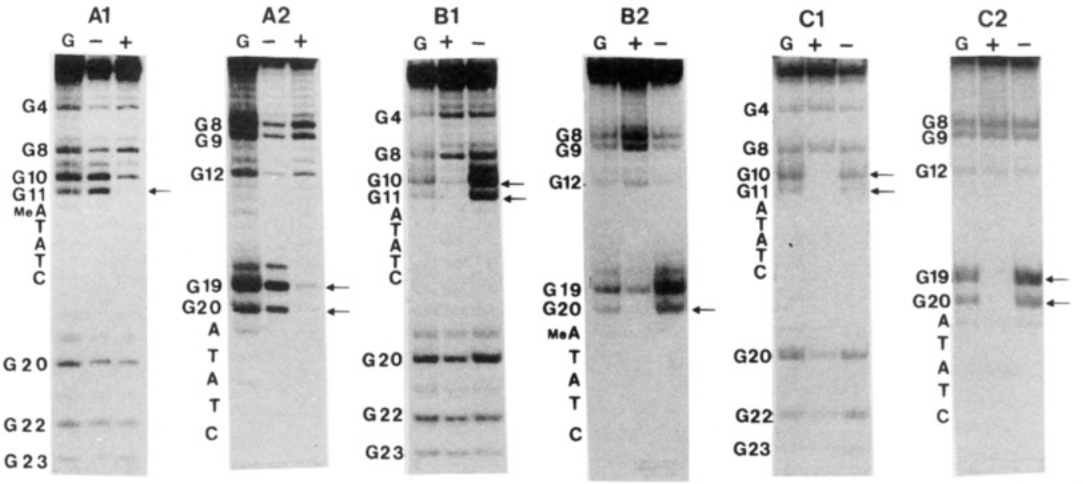

FIGURE 4: Methylation interference analysis of dG bases. Tracks labeled G are Maxam−Gilbert dG sequencing reactions which serve to identify bands and as a control (they are the same samples as those subsequently subjected to gel-shift separation). (−) represents the free DNA pool obtained from gel-shift analysis, and (+) represents the DNA pool bound to the methylase in the presence of sinefungin obtained from gel-shift analysis. (A) AAAGTCTGTGG^{6Me}ATATCCAAGTGGCTACCGT (A1) plus CCCCCACGGTAGCCACTTGGATATCCACAGACT (A2); (B) AAAGTCTGTGGATATCCAAGTGGCTACCGT (B1) plus CCCCCACGGTAGCCACTTGG^{6Me}ATATCCACAGACT (B2); (C) AAAGTCTGTGGATATCCAAGTGGCTACCGT (C1) plus CCCCCACGGTAGCCACTTGGATATCCACAGACT (C2). The individual dG bases are indicated as are the GATATC sites. Strongly interfering dG bases are arrowed.

endonuclease. Thus in this case we favor the close approach/steric hindrance explanation.

In view of the fact that both phosphate ethylation and dG methylation interference indicated that regions 5′ to the GATATC site are important for DNA recognition by the methylase, we have carried out experiments with hemimethylated duplexes containing three dG residues to the 5′ to the *Eco*RV site. For completion studies with three dG's to the 3′-end of GATATC have also been undertaken (see Table 1 for sequences). The phosphate ethylation interference patterns seen with both of these extra dG oligonucleotides were almost identical to those presented in Figures 2 and 3 (not shown). This confirms the importance of the phosphates NpNpNpGpA, reported above, and also shows that the purinic region introduced into the sequence does not affect overall binding or significantly alter DNA structure. The corresponding dG methylation interference results were also very similar to those seen for the standard 30/33-mer given in Figures 4 and 5 (not shown). This indicates that the methylase does not closely approach bases further out than one from the 5′-end of the GATATC site. Bases immediately 3′ of the target hexamer are also not in the proximity of the protein.

Methylation Protection. Using methylation protection, in which a protein is bound to a nucleic acid prior to reaction with dimethyl sulfate and so protects proximal purines from reaction with the reagent, gave the results shown in Figure 5. In contrast to the interference experiments, it proved possible to scan these gels to give the free DNA/bound DNA ratios illustrated in Figure 6. The methylation protection results strongly support those seen in the complementary interference experiments. Both Figure 5 and Figure 6 confirm the importance of the dG, within the GATATC site, for the interaction of the methylase with DNA. Unfortunately, rather weak bands were produced for this dG in this set of experiments, even when no methylase was present, and so the role of this base is better assessed using the interference data. However, the protection experiments gave excellent data for the 5′-flanking dG. Figures 5 and 6 clearly show that this base is protected when it is in the unmethylated

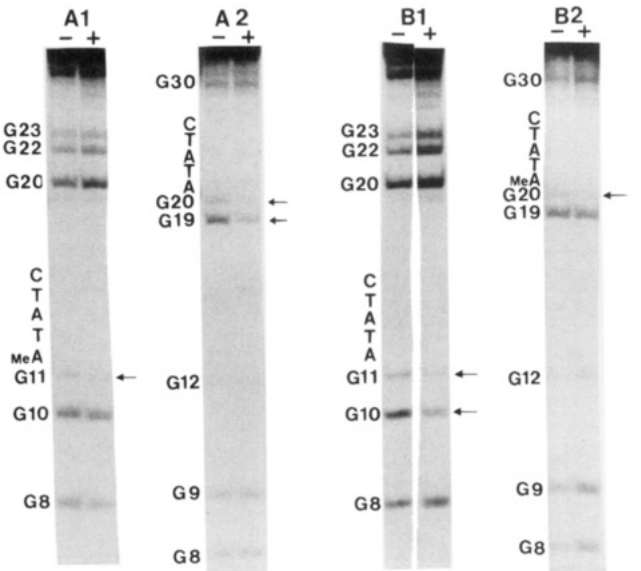

FIGURE 5: Methylation protection analysis of dG bases with dimethyl sulfate. (−) and (+) represent the reaction of dimethyl sulfate with DNA in the absence and the presence of the *Eco*RV methylase/sinefungin, respectively. (A) AAAGTCTGTGG^{6Me}ATATCCAAGTGGCTACCGT (A1) plus CCCCCACGGTAGCCACTTGGATATCCACAGACT (A2); (B) AAAGTCTGTGGATATCCAAGTGGCTACCGT (B1) plus CCCCCACGGTAGCCACTTGG^{6Me}ATATCCACAGACT (B2); (C) AAAGTCTGTGGATATCCAAGTGGCTACCGT.

(GGA) strand but not when present in the methylated (GG^{6Me}A) strand. This confirms the idea of two binding orientations for CH_3 addition to each strand. Methylation protection with the control oligonucleotide lacking a GATATC site gave identical patterns whether or not protein was present (not shown), confirming the binding specificity of the methylase.

Base Analogues. The interfaces between both the *Eco*RV methylase and endonuclease have previously been probed using the self-complementary 12-mer GACGATATCGTC containing dA and T base analogues within the GATATC sequence (Newman *et al.*, 1990a,b). This approach was extended to dG and dC analogues with the endonuclease only

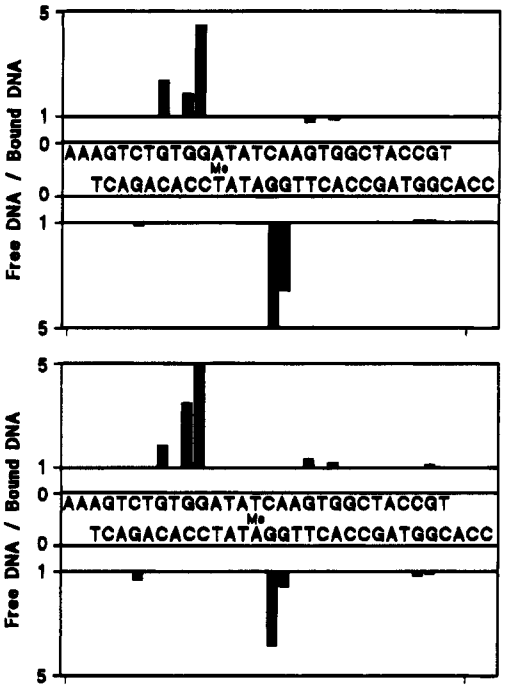

FIGURE 6: Quantitation of the gels shown in Figure 4. The ratio free DNA/bound DNA represents the intensity of a particular dG band found in the absence of DNA/sinefungin divided by the intensity found in the presence of these components.

Table 2: Rates Observed with the Methylase and the Oligonucleotide Combinations Shown at the Single Oligonucleotide Concentration of 20 μM[a]

oligonucleotide (unmethylated, base analogue in both strands)	rate, nmol min^{-1} mg^{-1} (%)
GACGATATCGTC	9.5 (100)
GAC^{7C}GATATCGTC	0 (0)
GAC^{6S}GATATCGTC	24.9 (100)
GACIATATCGTC	1.3 (13.7)
GAC^{3C}GATATCGTC	25.2 (265)
GACGATAT^{4H}CGTC	13.3 (140)
GACGATAT^{5Me}CGTC	0.74 (7.8)
GTCGACGATATCGTCGAC	4.5 (100)
GTCGAC^{2amP}ATATCGTCGAC	5.9 (131)

oligonucleotide (hemimethylated, base analogue in strand indicated)	rate, nmol min^{-1} mg^{-1} (%)
AGGG^{6Me}ATATCGGA/TCCGATATCCCT	16.4 (100)
AGG^{7C}G^{6Me}ATATCGGA/TCCGATATCCCT	0.1 (0.6)
AGGG^{6Me}ATATCGGA/TCC^{7C}GATATCCCT	<0.015 (<0.1)
AGG^{7C}G^{6Me}ATATCGGA/TCC^{7C}GATATCCCT	<0.015 (<0.1)
AGGG^{6Me}AUATCGGA/TCCGATATCCCT	1.9 (12)
AGGG^{6Me}ATATCGGA/TCCGAUATCCCT	11.2 (68)
AGGG^{6Me}AUATCGGA/TCCGAUATCCCT	0.3 (2)

[a] The top set of 12-mers are self-complementary and so form duplexes with the analogue present in both strands. These 12-mers containing d^{2amP} do not form stable duplexes at the assay temperature, and so self-complementary 18-mers must be used in this case (Waters & Connolly, 1994). The second set of 12-mers are not self-complementary, but the two strands given will form a duplex. This allows a hemimethylated oligonucleotide to be used and the base analogue to be placed separately on either strand. For the exact assay conditions and rate determination, see Newman *et al.* (1991a). All strands are written in the 5′ → 3′ direction.

(Waters & Connolly, 1994). However, as yet, the effects of incorporating dG and dC bases analogues into the GATATC sequence on the methylase activity have not been reported. These data are presented in Table 2. One of the most striking features seen in Table 2 is the complete inhibition of the methylase when the dG in the recognition sequence is replaced by d^{7C}G, an analogue in which the 7N is replaced with a CH. This confirms the importance of this site inferred from dimethyl sulfate interference and supports the view that a direct enzyme−DNA interaction takes place at this site rather than mere close proximity. All the other dG analogues had much smaller effects as did the dC-modified bases. The faster rates than control, which are sometimes seen with analogues, are most probably due to product release being the rate-limiting step with this enzyme with unmethylated substrates (this is currently under investigation). Previously (Newman *et al.*, 1990a), we showed that modifications to the first dA and T bases in the GATATC sequence strongly inhibited methylation and that in the particular instance of GAUATC no turnover was observed. Changes to the second dA and T had, with a very few exceptions, little effect. In order to commence a more precise study with a modified base present on only one strand, we have measured rates using the analogues d^{7C}G and dU present in only one strand of a hemimethylated duplex (Table 2). A single d^{7C}G on either the unmethylated or methylated strand is very inhibitory with the more pronounced effect being observed when the analogue was present on the unmethylated strand. As expected, the presence of the base on both strands resulted in no turnover. Thus the 7N locations of both the symmetrically related dG bases within the *Eco*RV site are important for the methylase. With dU double substitution gave a very low rate (2%) as opposed to the lack of activity seen with the complementary 12-mer previously. However, introduction of a single dU into the methylated strand has a much larger effect than placing this analogue in the unmethylated sequence. Caution is needed in interpreting the analogue data as an extensive kinetic analysis has not been carried out and only a steady-state rate at a single concentration of 20 μM is reported. Additionally, most of the work was performed with the self-complementary oligonucleotide, and so the modification will occur on both strands. As mentioned earlier product release is likely to be rate limiting in the case of unmethylated oligonucleotides (which initially generate the tightest bound hemimethylated substrate rather than the dimethylated product), and this places limits on the interpretation of the results. Nevertheless, the analogue data agree extremely well with the footprinting results and can be used to reinforce the conclusions drawn from these experiments. Both methods show the importance of the N7 of the dG. The phosphates and bases that interfere with binding are all at and beyond the 5′-end of the GATATC. Similarly, introducing base analogues into the 5′ GAT half-site is much more detrimental to methylase turnover than changes to the 3′ ATC bases.

CONCLUSIONS

The model for *Eco*RV methylase binding to its target site that we feel most agrees with our data is presented in Figure 7. The enzyme binds to two symmetrically disposed continuous runs of four phosphates (NpNpNpGpA) on both DNA strands. The displacement of phosphate contacts along the DNA backbone identifies the groove to which the protein binds. The phosphates on a pair of opposite strands which approach each other most closely are displaced by 3−7 base pairs in the 5′ direction when spanning the major groove and by 2−4 base pairs in the 3′ direction when spanning the minor groove. The four phosphate contacts on one strand

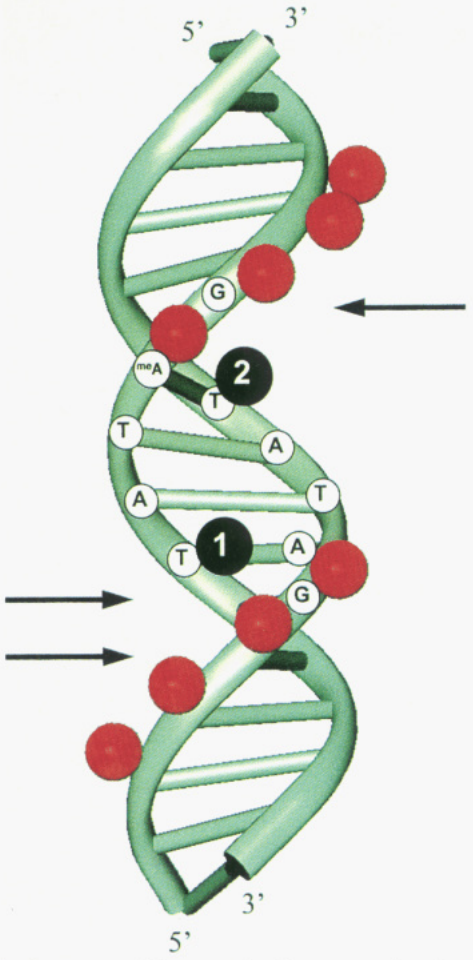

FIGURE 7: Summary of the postulated contact points between the *Eco*RV methylase and DNA. Red spheres on the backbone represent the van der Waals radii of phosphates at which ethylation interferes with binding. These divide into two sets of four contiguous phosphates, whose modification strongly inhibits binding, and are thus thought to contact the protein. The GATATC *Eco*RV recognition sequence is shown (the dC bases are obscured, mA is 6-methyldeoxyadenosine). Spheres 1 and 2 give the positions of the 6-NH_2 groups of the unmethylated and methylated dA bases, respectively. The arrows point toward the major groove of dG bases at which methylation protection and interference produces an effect. The use of base analogues additionally showed that the major groove determinants of the GAT sequence within the GATATC recognition site were important in recognition by the methylase. Only one methylated orientation is shown. Switching the strand of methylation gives a second unique set of contacts that are the reciprocal to those presented above (see text). In the absence of any other information we have made the most simple assumption that DNA bound to the methylase is in an "ideal" B-DNA form (see text).

are displaced by seven nucleotides in a 5′ direction from the contacts on the opposite strand, indicating a major groove contact. Furthermore, these phosphates span about 180° and so are located on one side of the nucleic acid, consistent with the methylase binding to one face of the DNA. As shown in Figure 7 the contacted phosphates are not exactly opposite each other across the major groove but are displaced somewhat. Nevertheless all these data can be accommodated by positioning the methylase obliquely across the major grove, and these results also suggest that the predominant areas of contact will be to this groove. Although the interference of these two sets of phosphates does not vary with the strand location of the $d^{6Me}A$ and so is consistent with only a single binding orientation, all the other results suggest asymmetry. Methylation protection and interference studies indicate three important dG residues, and these data, which agrees well with the phosphate ethylation, are also summarized in Figure 7. Not surprisingly, both dG residues within the *Eco*RV site are important. Additionally, the 5′-flanking nucleotide appears to be near the enzyme but only when it is on the unmethylated strand. The enzyme interaction with this flanking base exactly reflects the position of the methylated base, and both switch strands in complete synchrony, providing clear evidence of two binding modes. This result is confirmed by the exonuclease III digestions which shift one base outward in the 5′ direction for an unmethylated as compared to a methylated but otherwise identical strand. The analogue data which show that the N7 of dG, the N7 of dA, and the O4 and 5-CH_3 of T in the first half GAT are important for catalysis, using symmetrical oligonucleotides with the modified base in both strands, agree with the model in Figure 7 as all these functions are in the major groove. With these doubly substituted oligonucleotides it is impossible to decide what proportion of the inhibition arises from contacts at each of the two GAT half-sites. However, the initial experiments with hemimethylated oligonucleotides and a modified base in either strand have clearly shown that the N7 positions on both dG bases are important. In contrast, the 5-CH_3 group of the first T on the methylated strand makes a much larger contribution to recognition than does this function on the unmethylated sequence. The strong inhibition seen with some minor groove determinants of these three bases (N3 of dA and O2 of T) may be due to either secondary conformational effects or to direct interactions. Unfortunately, the irreproducible data obtained for the methylation of dA, which takes place at N3 in the minor groove, do not allow us to draw any conclusions concerning interaction at this location. In general, changes to the second half-site, ATC, were far less inhibitory to catalysis by the methylase. In summary, we propose that two regions of the methylase interact with the two NpNpNpGpApT sequences on each strand. The area binding to the unmethylated strand will contain the catalytic functions and presumably amino acids responsible for interaction with AdoMet. The other methylase region is optimized for recognizing the methylated strand and has no catalytic role. A two-domain structure (comprising a catalytic and target recognition domain) has been resolved for the dC methylase *Hha*I (Cheng *et al.*, 1993; Klimasauskas *et al.*) and *Taq*I methylase (Labahn *et al.*, 1994) and inferred for the dA methylase *Eco*RI (Reich *et al.*, 1991). It has been further proposed that all DNA methylases may share this architecture (Schluckebier *et al.*, 1995). The two domains form a cleft into which the DNA binds, and contacts to the nucleic acid are made from both regions. It is tempting to speculate that one half NpNpNpGpApT site interacts with one of the domains and the second with the other. However, the validity of this remains to be established. The two binding orientations which are observed are clearly distinct even though the footprinting data suggest they have several elements of similarity.

Two caveats are needed when the above mode of interaction are considered. We have assumed that the bound DNA is undistorted and in an idealized B form. It is recognized that almost all DNA binding proteins distort DNA on binding, sometimes drastically. Preliminary unpublished observations (S. J. Cal, P. Heslop, and B. A. Connolly)

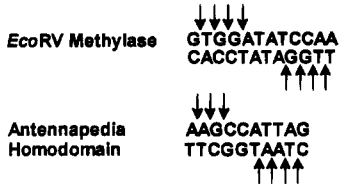

FIGURE 8: Comparison of the phosphate contacts seen for the monomeric *Eco*RV modification methylase and the antennapedia homodomain (only strongly interfering phosphates are indicated). The interference patterns are rather similar, and the asymmetric protein−DNA interaction gives rise to a very symmetric phosphate footprint in both cases. For a compilation of phosphate ethylation patterns seen with dimeric proteins, see Nye and Graves (1990).

suggest that the methylase severely bends the DNA and that this may serve to bring the two contacted DNA regions closer together. These data will serve to refine, rather than invalidate, the above model. It is also impossible to completely exclude an alternative scenario in which the enzyme does not bind across the major groove but only contacts the NpNpNpGpApT region of one strand. With the hemimethylated sequences a mixture consisting of an enzyme−substrate complex (binding to the unmethylated strand) and an enzyme−product complex (binding to the methylated strand) with slightly different conformations could conceivably give the footprinting and analogue data reported. However, the ethylation or methylation of sensitive sites results in a nearly complete absence of bands for the positions corresponding to the two half-sites on both DNA strands. Thus, for example, binding to the NpNpNpGp^{6Me}-ApT half-site only should give bands for the noncontacted NpNpNpGpApT site and vice versa, resulting in a reduced rather than the near absolute footprints observed. Furthermore, the mixture model agrees rather poorly with that in the previous paper. Binding to dimethylated (product) DNA is weak and gives a "slow" complex on gels whereas binding to unmethylated (substrate) sequences is much tighter and gives a "fast" complex. If reaction with hemimethylated species was just a mixture of substrate and product complexes, one should see a slow plus fast complex. The single fast complex with tighter binding than the unmethylated DNA can only be easily explained by interaction with both half-sites as in Figure 7. The importance of the two 7N dG sites also suggests that both bases are simultaneously contacted. Recognition of the two GAT sites also serves to fix the recognition sequence as GATATC (i.e., the *Eco*RV sequence) and build the d^{6Me}A into the recognition process. Interaction with one GAT site only would suggest a much broader methylation specificity than for GATATC for which there is little evidence.

Finally, we have observed that the methylase phosphate interference pattern is very similar to that seen with the antennapedia homeodomain (Affolter *et al.*, 1990). This is illustrated in Figure 8. Interestingly, both the methylase and the homeodomain are active as monomers and bind to nonsymmetrical DNA sequences (if hemimethylation is taken into account for the methylase). This contrasts to the more common homodimeric proteins (prokaryotic repressors and restriction endonucleases) that bind symmetrically to palindromes. Thus these two monomeric proteins which interact with DNA in an asymmetric manner both produce highly symmetric phosphate ethylation patterns. This is presumably because binding to phosphates on one face of the DNA across the major groove is favored as it allows access to the bases via the major groove. Both the *Hha*I (Klimasauskas *et al.*, 1994) and the *Eco*K (Powell & Murray, 1995) methylases predominantly contact their recognition bases through the major groove. It is in this region that the bases show their greatest variability (Seeman *et al.*, 1977), and so binding here can be used to generate specificity with an economical number of base pairs. Binding to phosphates across the major groove is rather common and also occurs for many dimeric DNA binding proteins, e.g., C/EDP (Nye & Graves, 1990), λ repressor (Tullius & Dombroski, 1988), and *Eco*RI endonuclease (Lu *et al.*, 1981; Lesser *et al.*, 1990). However, the exact interference patterns in these cases differ from the two monomeric proteins, suggesting that the fine details of the protein−DNA interactions vary between the two classes.

ACKNOWLEDGMENT

We thank I. B. Vipond for help with Figure 7 and S. E. Halford for fruitful discussions.

REFERENCES

Affotler, M., Percival-Smith, A., Müller, M., Leupin, W., & Gehring, W. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4093−4097.

Aiken, C. R., & Gumport, R. I. (1991) *Methods Enzymol. 208*, 433−468.

Bickle, T. A., & Krüger, D. H. (1993) *Microbiol. Rev. 57*, 434−450.

Billiter, M., Qian, Y. Q., Otting, G., Müller, M., Gehring, W., & Wüthrich, K. (1993) *J. Mol. Biol. 234*, 1084−1097.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986a) *J. Biol. Chem. 261*, 7270−7278.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I (1986b) *J. Biol. Chem. 261*, 7279−7286.

Cheng, X., Kumar, S., Pósfai, J., Pflugrath, J. W., & Roberts, R. J. (1993) *Cell 74*, 299−307.

Chuprina, V. P., Rullman, J. A. C., Lamerichs, R. M. J. N., Van Boom, J. H., Boelens, R., & Kaptein, R. (1993) *J. Mol. Biol. 234*, 446−462.

Dixon, W. J., Hayes, S. S., Levin, S. R., Weidner, M. F., Dombroski, B. A., & Tullius, T. D. (1991) *Methods Enzymol. 208*, 380−413.

Dubey, A., & Roberts, R. J. (1992) *Nucleic Acids Res. 20*, 3167−3173.

Fox, K. R. (1988) *Biochem. Biophys. Res. Commun. 155*, 779−785.

Galas, D. S., & Schmitz, A. (1987) *Nucleic Acids Res. 5*, 3157−3170.

Garner, M. M., & Revzin, A. (1986) *Trends Biochem. Sci. 11*, 395−396.

Gartenberg, M. R., Ampe, C., Steitz, T. A., & Crothers, D. M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6034−6038.

Harrison, S. C. (1991) *Nature 353*, 715−719.

Herman, G. E., & Modrich, P. (1982) *J. Biol. Chem. 257*, 2605−2612.

Kim, Y., Grable, J. C., Love, R., Greene, P. J., & Rosenberg, J. M. (1990) *Science 249*, 1307−1309.

Klimasauskas, S., Kumar, S., Roberts, R. J., & Cheng, X. (1994) *Cell 76*, 357−369.

Labahn, J., Granzin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I., & Saenger, W. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10957−10961.

Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776−786.

Lesser, D. R., Kurpiewski, M. R., Waters, T. R., Connolly, B. A., & Jen-Jacobson, L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7548−7552.

Lu, A. -L., Jack, W. E., & Modrich, P. (1981) *J. Biol. Chem. 256*, 13200−13206.

Manfield, I., & Stockley, P. G. (1994) in *Methods in Molecular Biology, Vol. 30, DNA−Protein Interactions, Principles and Protocols* (Kneale, G. G., Ed.) pp 125−139, Humana Press, Totowa, NJ.
Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499−560.
McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtossek, S. (1987) *Biochemistry 26*, 7238−7245.
Metzger, W., & Heumann, H. (1994) in *Methods in Molecular Biology, vol. 30, DNA-Protein Interactions, Principles and Protocols* (Kneale, G. G., Ed.) pp 11−20, Humana Press, Totowa, NJ.
Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990a) *Biochemistry 29*, 9891−9901.
Newman, P. C., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990b) *Biochemistry 29*, 9902−9910.
Nwosu, V. U., Connolly, B. A., Halford, S. E., & Garnet, J. (1988) *Nucleic Acids Res. 16*, 3705−3720.
Nye, S. A., & Graves, B. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3992−3996.
Pabo, C. O., & Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.
Powell, L. M., & Murray, N. (1995) *Nucleic Acids Res. 23*, 967−974.
Reich, N. O., & Danzitz, M. J., Jr. (1991) *Nucleic Acids Res. 19*, 6587−6594.
Reich, N. O., & Danzitz, M. J., Jr. (1992) *Biochemistry 31*, 1937−1945.
Reich, N. O., & Mashhoon, N. (1993) *J. Biol. Chem. 268*, 9191−9193.
Reich, N. O., Maegley, K. A., Shoemaker, D. D., & Everett, E. (1991) *Biochemistry 30*, 2940−2946.
Rubin, R. A., & Modrich, P. (1977) *J. Biol. Chem. 252*, 7265−7272.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) in *Molecular Cloning, a Laboratory Manual*, 2nd ed., pp 11.35−11.36, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schluckebier, G., O'Gara, M., Saenger, W., & Cheng, X. (1995) *J. Mol. Biol. 247*, 16−20.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804−808.
Shalloway, D., Kleinberger, T., & Livingston, D. M., (1980) *Cell 20*, 411−422.
Shaw, P. E., & Stewart, A. F. (1994) in *Methods in Molecular Biology, vol. 30, DNA-Protein Interactions, Principles and Protocols* (Kneale, G. G., Ed.) pp 79−87, Humana Press, Totowa NJ.
Siebenlist, U., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 122−126.
Sugisaki, H., Kita, K., & Takanami, M. (1989) *J. Biol. Chem. 255*, 5757−5761.
Sun, L., & Singer, B. (1975) *Biochemistry 14*, 1795−1802.
Tullius, T. D., & Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469−5473.
Waters, T. R., & Connolly, B. A. (1994) *Biochemistry 33*, 1812−1819.
Wilson, G. G. (1992) *Methods Enzymol. 216*, 259−279.
Wilson, G. G., & Murray, N. E. (1991) *Annu. Rev. Genet. 25*, 585−627.
Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglu, D. J., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., & Wilson, K. S. (1993) *EMBO J.* **12**, 1781−1795.
Yoo, O. J., & Agarwal, K. L. (1980) *J. Biol. Chem. 255*, 6445−6449.
Zebala, J. A., Choi, J., Trainor, G. C., & Barany, F. (1992) *J. Biol. Chem. 267*, 8106−8116.

BI950347S