# Structural and Functional Analysis of *Eco*RI DNA Methyltransferase by Proteolysis

Norbert O. Reich,* Karen A. Maegley, Dan D. Shoemaker, and E. Everett‡
*Department of Chemistry, University of California, Santa Barbara, California 93106*
*Received August 9, 1990; Revised Manuscript Received December 3, 1990*

ABSTRACT: Native *Eco*RI DNA methyltransferase (Mtase, $M_r$ 38 050) is proteolyzed by trypsin to generate an intermediate 36-kDa fragment (p36) followed by the formation of two polypeptides of $M_r$ 23 000 and 13 000 (p23 and p13, respectively). Protein sequence analysis of the tryptic fragments indicates that p36 results from removal of the first 14 or 16 amino acids, p23 spans residues 15–216, and p13 spans residues 217–325. The relative resistance to further degradation of p23 and p13 suggests stable domain structures. This is further supported by the generation of similar fragments with SV8 endoprotease which has entirely different peptide specificities. Our results suggest the Mtase is a two-domain protein connected by a highly flexible interdomain hinge. The putative hinge region encompasses previously identified peptides implicated in AdoMet binding [Reich, N. O., & Everett, E. (1990) *J. Biol. Chem. 265*, 8929–8934] and catalysis [Everett et al. (1990) *J. Biol. Chem. 265*, 17713–17719]. Protection studies with DNA, *S*-adenosylmethionine (AdoMet), *S*-adenosylhomocysteine (AdoHcy), and sinefungin (AdoMet analogue) show that the Mtase undergoes significant conformational changes upon ligand binding. Trypsinolysis of the AdoMet-bound form of the Mtase generates different fragments, and the AdoMet-bound form is over 800 times more stable than unbound Mtase. The sequence-specific ternary complex (Mtase–DNA–sinefungin) is 2000 times more resistant to degradation by trypsin; cleavage eventually generates 26- and 12-kDa fragments which span residues 104–325 and 1–103, respectively (p26 and p12). The first 14 or 16 amino acids of the Mtase are not essential since p36 retains activity. Activity analysis of the p26 and p12 mixture also indicates retention of activity. Therefore, either p26 is catalytically active or the two fragments remain associated to create a functional enzyme. The former rationale is supported by the fact that for *Eco*RI Mtase, all peptide regions implicated in DNA binding, AdoMet binding, and catalysis reside in the p26 fragment.

*S*-Adenosylmethionine (AdoMet) is a methyl donor in a wide range of critical biological reactions, and organisms expend considerable energy in maintaining intracellular levels of this cofactor. AdoMet-dependent methyltransferases (Mtases) modify nucleic acids, proteins, lipids, and various small molecules (e.g., neurotransmitters), and although extensively investigated, Mtases remain poorly understood (Usdin et al., 1979; Razin et al., 1984). Thus, no tertiary structure is available for any AdoMet-dependent enzyme, assignment of amino acids involved in cofactor and substrate binding is limited, and details of the catalytic mechanisms remain obscure.

DNA methylation has biological consequences ranging from the protection of host DNA against bacterial restriction endonucleases to regulation of gene expression in eukaryotes (Razin et al., 1984; Modrich & Roberts, 1982). Type II DNA Mtases are part of bacterial restriction–modification systems; hundreds of these enzymes are known, each with a unique recognition site (Modrich & Roberts, 1982). Most DNA Mtases function as monomers and recognize a palindromic canonical DNA sequence. *Eco*RI Mtase methylates the second adenine at the N6 position in the DNA substrate 5′GAATTC3′, thereby rendering the site resistant to the *Eco*RI endonuclease (Rubin & Modrich, 1977). The gene for the $M_r$ 38 050 monomeric Mtase has been cloned and the enzyme overexpressed, the amino acid sequence is known (Greene et al., 1981), and X-ray crystallographic and $^1$H NMR derived structures are available for a DNA substrate (Wing et al., 1980; Nerdal et al., 1989) as well as the methylated product (Frederick et al., 1988). The Mtase is extremely efficient, having a specificity constant for plasmid DNA of over $10^8$ $s^{-1}$ $M^{-1}$ (Reich & Mashhoon, 1991). Kinetic data obtained with small (14 bp) and large (plasmid) substrates are consistent with an ordered steady-state mechanism in which AdoMet binds prior to DNA binding; however, Mtase binding randomly to AdoMet and noncanonical DNA is suggested, with formation of the functional ternary complex (Mtase–DNA$^c$–AdoMet, DNA$^c$ = canonical DNA) requiring AdoMet to be bound.

DNA Mtases have so far proven refractory to crystallization, and we have therefore undertaken several structure–function studies of the *Eco*RI Mtase. The 325 amino acid polypeptide chain contains seven cysteines, and we recently demonstrated that Cys$^{223}$ is critical for function: reaction of the Mtase with *N*-ethylmaleimide causes modification of Cys$^{223}$ and complete loss of activity (Everett et al., 1990). Cys$^{223}$ is adjacent to a peptide segment confirmed by photoaffinity experiments to form part of the AdoMet site in *Eco*RI Mtase (Gly$^{205}$ to Pro$^{220}$) (Reich & Everett, 1990).

Given the lack of knowledge of how *Eco*RI Mtase, or any other Mtase, is structurally organized, we sought to answer several related questions: (1) What is the domain organization of the enzyme? (2) What regions are involved in AdoMet and DNA binding? (3) Do significant conformational changes occur upon substrate and/or cofactor binding? The approach described in this report uses controlled proteolysis of the Mtase with proteases having diverse peptide cleavage preferences. Mtase peptide fragments isolated at intermediate stages of degradation were sequenced to identify the site of proteolysis. Protection against proteolysis afforded by various combinations

* To whom correspondence should be addressed.
‡ Present address: Protein Chemistry Department, Diagnostic Products Corp., 5700 W. 96th, Los Angeles, CA 90045.

0006-2960/91/0430-2940$02.50/0 

of substrates, products, and inhibitors was used to assess ligand-induced conformational changes. Finally, activity analysis was used to implicate protein segments essential for catalysis.

## Experimental Procedures

### *Materials*

Methionine, bovine serum albumin (BSA), trypsin, *S*-adenosylhomocysteine (AdoHcy), and sinefungin were obtained from Sigma Chemicals; *S*-adenosylmethionine (AdoMet) was purchased from Sigma and purified as described (Reich & Mashhoon, 1990); *S*-[*methyl*-$^3H$]adenosylmethionine (78.9 Ci/mmol) was purchased from New England Nuclear; endoproteinase glutamate C (SV8) was from Boehringer Mannheim; *Eco*RI Mtase was isolated from *Escherichia coli* strain 294 carrying plasmid pPG440 (Greene et al., 1978). The following DNA substrates were prepared by using a Biosearch DNA synthesizer and phosphoramidite solid-phase chemistry (top strand shown only): 5′CGCGAATTCGCG3′ (dodecamer) and 5′GGCGGAATTCGCGG3′ (14-mer). The methylated 14-mer ($N^6$-methyladenine at the second adenine position of both strands) was synthesized with the corresponding modified phosphoramidite (Pharmacia). Oligonucleotides were purified on a reverse-phase C-18 HPLC system, first with the 5′-trityl group intact and then with the trityl group removed (Becker et al., 1985). Calf thymus DNA was from Sigma. Liquiscint scintillation fluid (National Diagnostics) and a Beckman Model LS1700 scintillation counter were used to measure radioactivity. DE81 ion-exchange filter papers were from Whatman.

### *Methods*

*Mtase Activity Assay.* The activity of the Mtase was monitored by measuring tritium incorporation into the DNA substrate from [methyl-$^3H$]AdoMet (Reich & Mashhoon, 1990, 1991). Samples of 50 μL contained 100 mM Tris, pH 8.0, 10 mM EDTA, 0.2 mg/mL BSA, and 1.0 mM DTT with the indicated enzyme, DNA, and AdoMet concentrations. A 40-μL aliquot of each sample was placed on DE81 filters after incubation at 37 °C for the indicated time. Filters were prepared as described (Reich & Mashhoon, 1990, 1991). The moles of product formed were determined from scintillation counting of the DE81 filters corrected for quenching from the filters. Mtase concentrations were obtained either by direct spectrophotometric determination ($E^{1\%}_{280_{nm}} = 10.8$) (Rubin & Modrich, 1977) or with the Bio-rad protein assay and a Mtase standard curve.

*Gel Electrophoresis and Densitometry.* Protein samples were analyzed by SDS–PAGE with 12% and 4% acrylamide in the running and stacking gels, respectively (Laemmli, 1970). Gels were stained with Coomassie dye, destained, and photographed. Gel scanning was carried out with a Hoeffer 350 densitometer. A linear dependence of Mtase loaded and the amount determined by densitometry was confirmed.

*Proteolysis (Mtase Alone).* Mtase at 5 μM in 3.3 mM potassium phosphate, pH 7.0, 1.7 mM DTT, 0.33 mM EDTA, and 3.3 mM NaCl (digestion buffer, DB) was treated with trypsin [5% or 1.5% (w/w) as indicated] or SV8 (1%) for the indicated length of time. Incubation temperature was 37 °C unless otherwise indicated. Digestion was stopped with trypsin inhibitor [15% (w/w of Mtase)] or phenylmethanesulfonyl fluoride (PMSF) at a final concentration of 0.5 mM. Samples were subjected to the SDS–PAGE analysis described above.

*Trypsinolysis (Substrate Titration).* Unless indicated, Mtase at 5 μM was preincubated at room temperature for 5 min with various substrates individually [DNA (dodecamer), AdoMet, AdoHcy, or sinefungin] in DB. The following concentrations were used: DNA, 0–300 μM; sinefungin, 0–500 μM; AdoHcy, 0–1.1 mM; and AdoMet, 0–110 μM. Trypsin was added to 5% (w/w), and samples were heated to 37 °C for 1 (DNA, sinefungin), 2 (AdoHcy), and 2.5 (AdoMet) min. The digests were stopped as above, samples were analyzed by SDS–PAGE as described earlier, and bands were quantified by densitometry. The dissociation constant for each ligand ($K_d$) is defined as that concentration of ligand providing 50% protection of the full-length Mtase (p38). Control experiments included trypsinolysis of BSA in the presence and absence of ligands.

*Trypsinolysis (Substrate Protection).* Mtase (5 μM unless indicated) in DB was preincubated with saturating concentrations of individual ligands (or 10 μM sinefungin and 10 μM DNA, 14-mer, or methylated 14-mer) for 3–5 min on ice and then treated with 5% trypsin for the indicated times at 37 °C. Samples were combined with trypsin inhibitor and kept on ice until all time points were taken. Analysis was by SDS–PAGE as described above and quantified by using densitometry. The half-life for unprotected and protected p38 was determined by plotting percent p38 remaining as a function of time.

*Mtase Activity as a Function of Trypsinolysis.* Mtase (3 μM) was incubated with 10 μM DNA (dodecamer) and 10 μM sinefungin in DB and digested with 5% (w/w) trypsin. Aliquots were removed at 0, 15, 30, 60, 120, 240, and 360 min, combined with trypsin inhibitor, and analyzed by SDS–PAGE and densitometry. Duplicate aliquots were diluted in 20 mM potassium phosphate, pH 7.4, 200 mM NaCl, 0.2 mM EDTA, 0.2 mg/mL BSA, 2 mM DTT, and 10% glycerol to a Mtase concentration of 2.0 nM. These samples were assayed for Mtase activity at 37 °C for 15 min with the following conditions: 100 mM Tris, pH 8.0, 10 mM EDTA, 0.2 mg/mL BSA, 1.0 mM DTT, 0.2 nM Mtase, 3 μM AdoMet, and 0.1 mg/mL calf thymus DNA.

*Peptide Sequencing.* After SDS–PAGE separation, peptides were electroeluted onto poly(vinylidene difluoride) (PVDF) membranes as described (LeGendre & Matsudaira, 1988). The desired bands were sequenced on a Applied Biosystems Model 477 liquid-phase sequencer with on-line reverse-phase chromatography.

*Protein Secondary Structure Analysis.* Sequence analysis was performed with the University of Wisconsin Genetics Computer Group programs (Devereux et al., 1984) run on a Digital Equipment Corp. VAX computer.

## Results

*Digestion of Substrate-Free Mtase.* As shown in Figure 1, limited proteolysis of the *Eco*RI DNA Mtase by two residue-specific endopeptidases [trypsin (1.5% w/w) and SV8] generates polypeptide fragments of sizes 22–23 and 13–15 kDa. A transient 36-kDa (p36) fragment is visible with trypsin. Trypsin cleaves on the carboxy side of arginines and lysines and is predicted to generate over 50 peptide fragments of the Mtase; SV8 cleaves on the carboxy side of glutamate and aspartate side chains and is predicted to generate over 45 fragments. Each of the trypsin-generated fragments (p36, p23, p22.5, and p13) was transferred onto a PVDF membrane by electroblotting and sequenced by automated methods (LeGendre & Matsudaira, 1988). Peptide sequence analysis indicated that p36 has the first 14 residues removed to reveal serine-15 at the amino terminus. Although the carboxy-terminal residue was not determined, removal of few if any carboxy-terminal residues is predicted based on the apparent molecular weight as determined by comparison with size standards (Figure 1). Sequencing of p22.5 and p23 showed that they result from cleavage after lysine-16 and lysine-14,

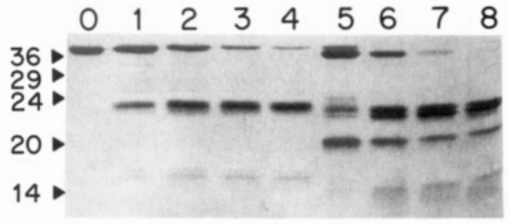


FIGURE 1: Digestion of substrate-free Mtase. Mtase (5 $\mu$M) in DB was treated with SV8 protease (1%) (lanes 1–4) or trypsin (1.5%) (lanes 5–8) at 37 °C; the digests were stopped with PMSF or trypsin inhibitor at the indicated times and analyzed by SDS–PAGE as described (Methods). Molecular weights ($\times 10^{-3}$) as determined by size standards shown at left. Lane 0, no digestion; lane 1, 5 min; lane 2, 15 min; lane 3, 30 min; lane 4, 45 min; lane 5, 5 min; lane 6, 15 min; lane 7, 30 min; lane 8, 45 min. The 20-kDa band in lanes 5–8 is trypsin inhibitor.

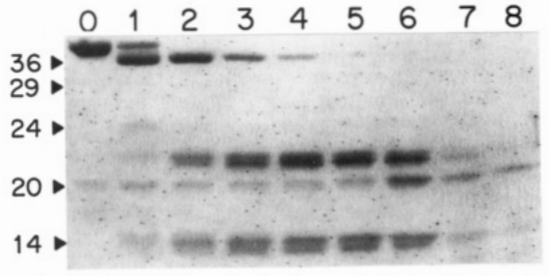


FIGURE 2: SDS–PAGE analysis of trypsinolysis of substrate-free Mtase. As a Figure 1 but 5% trypsin was used and the digest was stopped at the indicated times: lane 0, no digestion; lane 1, 30 s; lane 2, 1 min; lane 3, 1.5 min; lane 4, 2 min; lane 5, 3 min; lane 6, 5 min; lane 7, 30 min; lane 8, 1 h. Molecular weight ($\times 10^{-3}$) size markers as in Figure 1.

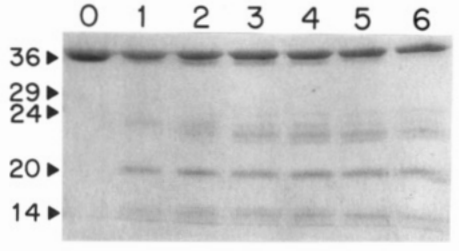


FIGURE 3: Trypsinolysis of AdoMet-bound Mtase. Mtase (5 $\mu$M) was preincubated with 20 $\mu$M AdoMet in DB at 37 °C and treated with trypsin, aliquots were stopped with trypsin inhibitor and analyzed at the indicated times: lane 0, no digestion; lane 1, 30 s; lane 2, 1 min; lane 3, 2 min; lane 4, 5 min; lane 5, 30 min; lane 6, 3 h.

respectively. Furthermore, sequencing of p13 indicates the amino terminus begins at isoleucine-217, and that based on size, this is where p23 and p22.5 are predicted to end. Thus, on the basis of N-terminal sequencing and relative size, p36 contains $Ser^{15}$ to $Lys^{325}$, p23 contains $Ser^{15}$ to $Arg^{216}$, p22.5 contains $Ser^{17}$ to $Arg^{216}$, and p13 contains $Ile^{217}$ to $Lys^{325}$.

*Substrate-Bound Mtase Digestion with Trypsin.* Figure 2 shows the time course of trypsin cleavage (5% w/w) of the Mtase at 37 °C over 3 h. Figures 3–6 show the same experiments done in the presence of various components of the Mtase reaction, all at saturating ligand concentrations as previously determined by trypsinolysis (data not shown). The sinefungin concentration was based on our previous determination of the $K_d$ (43 $\mu$M) (Reich & Mashhoon, 1990). With the exception of AdoMet (Figure 3), the same fragmentation as that found with the unprotected (Figure 2 and Figure 1) pattern was observed. For AdoMet, no detectable p36 was formed. Table I shows the half-life of the full-length Mtase (p38) for each experiment, as calculated by quantification from densiometric scanning of the corresponding SDS–PAGE gels. These alterations in half-lives are specific

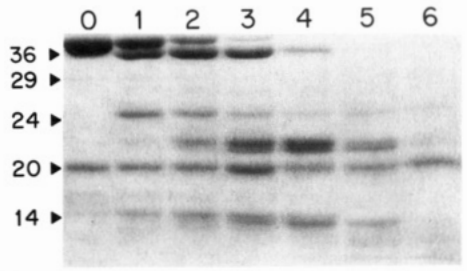


FIGURE 4: Trypsinolysis of the AdoHcy-bound Mtase. Mtase (5 $\mu$M) was preincubated with 300 $\mu$M AdoHcy in DB at 37 °C and treated with trypsin; aliquots were stopped with trypsin inhibitor and analyzed at the indicated times: lane 0, no digestion; lane 1, 30 s; lane 2, 1 min; lane 3, 2 min; lane 4, 5 min; lane 5, 30 min; lane 6, 3 h.

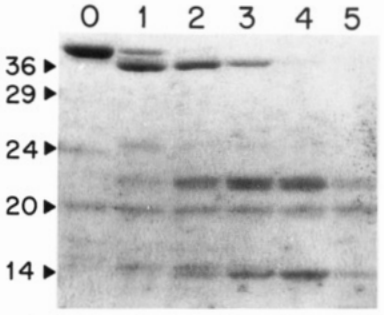


FIGURE 5: Trypsinolysis of sinefungin-bound Mtase. Mtase (5 $\mu$M) was preincubated with 500 $\mu$M sinefungin in DB at 37 °C and treated with trypsin, aliquots were stopped with trypsin inhibitor and analyzed at the indicated times: lane 0, no digestion; lane 1, 30 s; lane 2, 1 min; lane 3, 2 min; lane 4, 5 min; lane 5, 30 min.

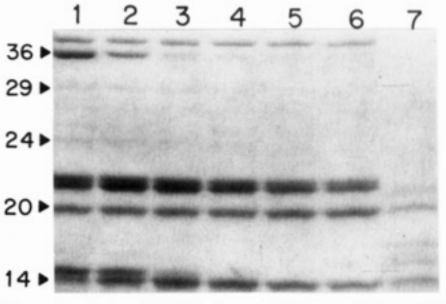


FIGURE 6: Trypsinolysis of DNA-bound Mtase. Mtase (5 $\mu$M) was preincubated with 50 $\mu$M DNA (dodecamer) in DB at 37 °C and treated with trypsin; aliquots were stopped with trypsin inhibitor and analyzed at the indicated times: lane 1, 30 s; lane 2, 1 min; lane 3, 2 min; lane 4, 5 min; lane 5, 30 min; lane 6, 1 h; lane 7, 3 h.

Table I: Rate Constants for p38 Degradation and Equilibrium Dissociation Constants for Mtase–Ligand Interactions

| ligand | $k$ (min$^{-1}$) | $K_{dp}$ ($\mu$M) | $K_d$* ($\mu$M) |
|---|---|---|---|
| unprotected | 2.3 | ND | ND |
| AdoMet | 0.0028 | 13 | 30 |
| AdoHcy | 1.1 | 215 | 225 |
| DNA | ND[b] | ND | ND |
| sinefungin | 2.3 | ND | 43 |
| DNA/sinefungin | 0.0029–0.0012 | ND | ND |

[a] The amount of p38 remaining at different time points in the experiments shown in Figures 2–7 was quantitated by densitometry and used to determine the degradation rate constants. The dissociation constants were obtained by using proteolysis (Methods) or by pre-steady-state competition methods (asterisk) (Reich & Mashhoon, 1990). [b] Not done, ND.

for the Mtase since no such changes were observed in identical experiments using BSA in place of Mtase (data not shown). The patterns of fragmentation were not altered when the proteolysis was done at 4 and 25 °C.

In contrast to the results with individual components (Figures 3–6), protection in the presence of DNA ans sine-

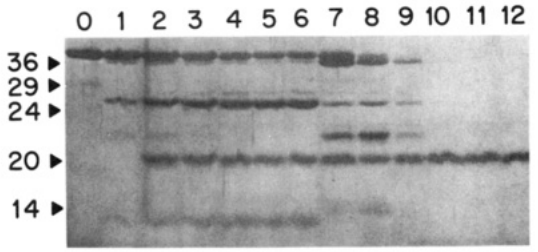


FIGURE 7: Trypsinolysis of Mtase in ternary complex (Mtase–DNA–sinefungin). Mtase (5 $\mu$M) was incubated with 10 $\mu$M DNA (14-mer, lanes 0–6, or methylated 14-mer lanes 7–12) and 10 $\mu$M sinefungin in DB at 37 °C and digested with 5% trypsin. Samples were analyzed as described under Methods at the following times: lane 0, no digestion; lane 1, 30 min; lane 2, 2 h; lane 3, 4 h; lane 4, 10 h; lane 5, 12 h; lane 6, 24 h; lane 7, 30 min; lane 8, 2 h; lane 9, 4 h; lane 10, 10 h; lane 11, 12 h; lane 12, 24 h.

fungin significantly alters both the half-life of p38 (Table I) and the pattern of fragmentation (Figure 7). The conditions of the experiment in Figure 7 are slightly altered from those described for Figures 2–6; however, the same pattern and extended half-life were observed under all conditions when both sinefungin and DNA were present. Sinefungin is a potent inhibitor of the Mtase in which the methyl sulfonium of AdoMet is replaced with an aminomethine functionality. This analogue stabilizes the sequence-specific ternary complex (Mtase–DNA–sinefungin) (Reich & Mashhoon, 1990). Two peptide fragments with relative molecular weights of 26K (p26) and 12K (p12) are formed. Amino-terminal peptide sequencing of the approximately 26-kDa fragment revealed that cleavage occurred between $Lys^{103}$ and $Lys^{104}$. Thus, based on the relative size, p26 comprises $Lys^{104}$ to $Lys^{325}$. Not shown in Table I, but apparent from Figure 7, is that p26 is resistant to further degradation. The 12-kDa fragment is probably the N-terminal fragment ($Ala^{1}$ to $Lys^{103}$). The dependence on an intact canonical site for protection is apparent when the doubly methylated 14-mer was used (with sinefungin), shown in Figure 7. Under these conditions, p26, p22.5, and p23 are formed, and p38 is degraded much more rapidly than with the intact site.

*Functional Analysis of Tripsinolysis Fragments.* Figure 8 shows the loss of p38 and Mtase activity resulting from Mtase trypsinolysis in the presence of DNA and sinefungin. The data clearly demonstrate that more Mtase activity is present after 360 min than can be accounted for by the amount of p38 remaining. Control experiments show that less than 10% Mtase activity is lost due to the reaction conditions (data not shown).

## Discussion

*Structural Organization.* Digestion of *Eco*RI DNA Mtase with trypsin (peptide cleavage occurring after arginine or lysine) generates two large proteolytic fragments of molecular masses 22.5 kDa (and 23 kDa) and 13 kDa (p22.5, p23, and p13, respectively). On the basis of a combination of peptide sequence analysis and relative fragment sizes (Figures 1 and 2), we have assigned p23, p22.5, and p13 to the peptide sequences shown in Figure 9. Digestion of the Mtase with SV8 (peptide cleavage after glutamate or aspartate) generates similar-sized fragments, p23.5 and p15 (Figure 1). Therefore, although the SV8 fragments were not sequenced, they most likely derive from peptide cleavages related to the trypsin fragments. For the SV8 digest, the lack of a doublet in the 23-kDa range can be explained by the lack of any acidic amino acids in the first 17 amino acids. No intermediate p36 fragment is generated, and the two SV8-generated fragments

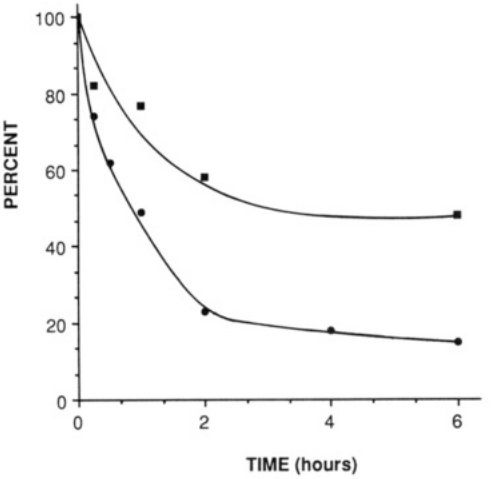


FIGURE 8: Percentage of remaining Mtase activity (■) and p38 polypeptide (●) as a function of trypsinolysis of ternary complex (Mtase–DNA–sinefungin). Mtase (3 $\mu$M) was preincubated with 10 $\mu$M dodecamer and 10 $\mu$M sinefungin in DB and digested with trypsin (5%). Aliquots were analyzed for residual Mtase activity, and residual p38 was quantitated as described under Methods. The amount of enzyme activity loss over the same period of time in the absence of proteolysis is less than 10%.

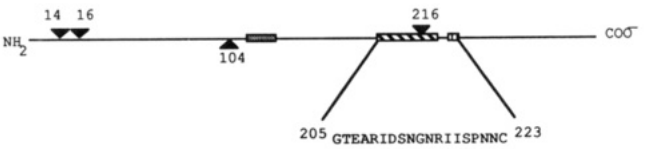


FIGURE 9: Schematic representation of *Eco*RI Mtase polypeptide. Boxes identify regions implicated in DNA recognition (dots, $^{139}NPPF^{142}$) (Chandrasegaran & Smith, 1987), AdoMet binding (diagonals, $G^{205}$ to $P^{220}$) (Reich & Everett, 1990), and catalysis (vertical lines, $N^{222}C^{223}$) (Everett et al., 1990). Sites of trypsinolysis identified by peptide sequencing are identified by solid triangles: cleavages occurring without ligand protection and in the presence of AdoHcy, sinefungin, or DNA above the line; cleavages occurring in the presence of DNA and sinefungin below the line. Peptide fragments described in the text: p36 ($S^{15}$ to $K^{325}$), p26 ($K^{104}$ to $K^{325}$), p23 ($S^{15}$ to $R^{216}$), p22.5 ($S^{17}$ to $R^{217}$), p13 ($I^{217}$ to $K^{325}$).

(p23.5 and p15) are slightly larger than the related trypsin fragments. Since trypsin and SV8 have very different cleavage specificities and both are predicted to have a large number of cleavage sites within the Mtase, generation of similar fragments suggests that the targeted region is unusually vulnerable. Although numerous structural features can contribute to this greater susceptibility, a likely contributor is a high degree of peptide flexibility in this region. For example, based on analysis of crystallographically determined temperature factors, peptide segments particularly susceptible to cleavage were shown to be highly flexible and frequently located on the protein surface (Fontana et al., 1986). The comparative resistance to further degradation of the p23, p22.5, and p13 fragments (Figures 1 and 2) suggests that these fragments either form stable tertiary structures (domains) connected by a flexible "hinge" or "tether" region or that a single domain has a flexible loop region which undergoes order–disorder transitions upon ligand binding. Although our data cannot be used to distinguish between these alternatives, comparison with similar results found for other proteins shows the multidomain structures to be quite common (Horowitz & Bowman, 1987; Galakatos & Walsh, 1987, 1989). Further support for the enhanced flexibility of the peptide segment near $Arg^{216}$ comes from analysis of the secondary structure of the entire Mtase polypeptide (Devereux et al., 1984). The larger N-terminal segment most likely does not include the first 14

or 16 amino acids since trypsinolysis rapidly releases these N-terminal residues (formation of p36, Figure 1).

Hinge regions connecting domains have been identified in numerous proteins including citrate synthetase (Remington et al., 1982), aspartate aminotransferase (Ford et al., 1980), alcohol dehydrogenase (Eklund et al., 1981), hexokinase (Bennett & Steitz, 1980), and adenylate kinase (Egner et al., 1987). In many cases, ligand binding was shown to cause conformational changes involving domain motion. Figure 9 shows the extent to which the previously identified AdoMet binding site (Reich & Everett, 1990) and the putative hinge region implicated in this study overlap. Further evidence for such overlap (hinge and AdoMet binding sites) comes from protection studies described below using AdoMet. Our implication of $Cys^{223}$ as being critical for catalysis (Everett et al., 1990) also suggests that this putative hinge region makes up part of the active site.

*Substrate Protection.* To further probe the structural organization of the Mtase, we determined the extent to which the two substrates (DNA and AdoMet), the two products (methylated DNA and S-adenosylhomocysteine, AdoHcy), and the AdoMet analogue sinefungin altered both the pattern of fragments generated by trypsinolysis and also the stability of the full-length Mtase (p38). Sinefungin is a potent Mtase inhibitor which stabilizes formation of a sequence-specific Mtase–DNA–sinefungin complex (Reich & Mashhoon, 1990). The appropriate concentration of each ligand was determined by measuring the level of p38 "protection" at various concentrations of ligand (data not shown); the data were used to determine dissociation constants ($K_{dp}$, Table I), and all subsequent protection experiments were done with ligand concentrations in excess of these constants. Since sinefungin alone did not afford any p38 protection, we used the dissociation constant for the Mtase–sinefungin complex obtained by pre-steady-state competition methods (Reich & Mashhoon, 1990). The patterns of Mtase trypsinolysis in the presence of each ligand are shown in Figures 2–6, and the half-life of p38 under these conditions is shown in Table I.

The $K_{dp}$ value for AdoMet is similar to the dissociation constant for the Mtase–AdoMet complex measured by equilibrium dialysis (Reich & Everett, 1990) and pre-steady-state competition methods (Reich & Mashhoon, 1990) (Table I), suggesting involvement of the same Mtase–AdoMet complex. Protection against trypsinolysis at high AdoMet concentrations shows that both the cleavage pattern and rate of p38 degradation are altered (Figures 2 and 3 and Table I). The p38 is over 830-fold more stable in the presence of AdoMet, and in contrast to the unprotected pattern, little p36 is formed. Since formation of p36 comes from cleavage at $Lys^{14}$ and $Lys^{16}$, the AdoMet-bound form of the Mtase must render these sites less prone to cleavage. This could either result from physical obstruction of the labile peptide bond due to direct contact with AdoMet, or conformational changes induced by AdoMet binding. The only other peptide segment implicated in AdoMet binding, amino acids $Gly^{205}$ to $Pro^{220}$, is considerably removed from this N-terminal region (Reich & Everett, 1990). Thus, either several noncontiguous peptide segments make up the AdoMet site, or AdoMet binding causes sufficient conformational change within the Mtase to alter proteolysis in regions removed from the AdoMet binding site. Since formation of p23, p22.5, and p13 results from cleavage in the peptide segment implicated in AdoMet binding, the significant stabilization of p38 (lack of p23, p22.5, and p13 formation) by AdoMet further supports this segment in binding AdoMet.

As with AdoMet, the dissociation constants for AdoHcy obtained by using proteolysis ($K_{dp}$) and other methods are similar (Table I), suggesting that the same Mtase–AdoHcy complex is involved. However, in this case, a negligible increase in p38 half-life is provided by formation of the complex. Although potentially accountable (in part) by the 16-fold difference in $K_{dp}$, the 380-fold difference in half-lives between the Mtase–AdoMet and Mtase–AdoHcy complexes must also derive from differential susceptibilities to proteolysis of the two binary complexes. This difference is particularly interesting since protection with AdoHcy generates significant quantities of the p36 intermediate, further supporting conformational differences in the two complexes (Figure 4). The fragmentation pattern with AdoHcy protection is actually quite complex since although formation of p36, p23, p22.5, and p13 is evident, a larger fragment seen with AdoMet protection (and possibly DNA and sinefungin, see below) is also apparent. This suggests two concurrent pathways of degradation since in no case have we observed the progression p38 to p36 and then to another intermediate prior to formation of p23 and p22.5. The data with AdoMet and AdoHcy protection clearly show the corresponding binary complexes to be conformationally distinct.

Figure 5 shows that saturating levels of sinefungin have no detectable effect on the fragmentation pattern of the Mtase nor the stability of p38 (Table I). This is a surprising result given the similar affinities of sinefungin and AdoMet for the Mtase (Table I), the fact that sinefungin is a competitive inhibitor with respect to AdoMet (Reich & Mashhoon, 1990), and the structural similarity of AdoMet and sinefungin. As with AdoHcy, these data suggest conformationally distinct binary complexes resulting from relatively minor changes in the cofactor at the sulfonium center.

Protection with DNA (dodecamer, Figure 6) results in a biphasic pattern of p38 degradation which precluded a definitive assignment of the p38 half-life. The first time point in Figure 6 is after 30 s of proteolysis and shows the majority of p38 is converted to p36, p23, p22.5, and p13. However, a small but reproducible amount of p38 is extremely resistant to degradation. As seen in Figures 1, 2, 4, and 5, trypsinolysis results in complete degradation of p38. One potential explanation for the biphasic pattern is that two or more DNA conformers exist and Mtase binding to the different forms results in altered degradation kinetics. This is inconsistent with our finding that the pattern and rates are insensitive to DNA concentration and to our observing similar biphasic kinetics with the 14-mer DNA (data not shown). Alternatively, two conformations of the Mtase may be present, which bind DNA in distinct ways and therefore have different susceptibilities to trypsinolysis. The only data relating to this explanation are our determination that the Mtase is completely active by active-site titration measurements, precluding the possibility that some fraction of the Mtase is inactive (Reich & Mashhoon, 1991).

In the presence of sinefungin and DNA, the Mtase forms sequence-specific ternary complexes with high affinity [10–40 nM (Reich & Mashhoon, 1990)]. As shown in Figure 7 and Table I, this complex is between 800 and 2000 times more stable to trypsinolysis than the unbound Mtase. The fragmentation pattern seen under these conditions (p26 and p12) is only found in the presence of both DNA and sinefungin. Peptide sequencing shows that p26 starts with $Lys^{104}$ and, based on relative sizes, continues to the end of the Mtase protein (see Figure 9). Given our assignment of the possible domain structure described earlier, the relative stability of p26

and p12 is at first difficult to reconcile. However, the stability of both fragments (p26 and p12) in the presence of DNA and sinefungin suggests that maintenance of the ternary complex subsequent to cleavage results in protection of both fragments. Cleavage at the $Lys^{104}$ position shows that this region is conformationally different in the ternary complex than in the unprotected, AdoMet-bound, or AdoHcy-bound forms of the enzyme. One feature of the fragmentation pattern similar to that observed with AdoMet protection is the lack of p36 formation. Thus, the conformation of the N-terminal region up to $Lys^{16}$ must be different in the ternary complex in comparison to the sinefungin-bound form of the Mtase.

Further evidence that formation of p26 and p12 requires a sequence-specific ternary complex with a canonical site (Mtase–$DNA^c$–sinefungin) is provided by the lack of protection (and limited formation of p26 and p12) with the methylated 14-mer (Figure 7). Although the dissociation constant for the Mtase–methylated 14-mer–sinefungin complex is not known, the $K_i$ for methylated DNA which derives from the Mtase–methylated DNA–AdoHcy complex is 41 nM (Reich & Mashhoon, 1991). The different fragmentation pattern and relative stabilities observed with the methylated 14-mer therefore was most likely due to a conformationally distinct ternary complex, not the inherent instability of this complex relative to that involving the unmethylated site.

*Activity Analysis of Peptide Fragments.* Mtase activity analysis of early trypsinolysis time points under unprotected conditions showed that formation of p36 did not detectably alter enzyme activity (data not shown). Thus, the first 14–16 amino acids of the Mtase are not essential for activity, although determination of true kinetic parameters (e.g., $K_m^{DNA}$, $K_m^{AdoMet}$, and $k_{cat}$) was not performed. Further correlation of activity to structure was obtained by characterizing Mtase activity as a function of p26 formation, shown in Figure 8. The fact that after 3 h of proteolysis at 4 °C only 15% of the p38 remains while 50% of the Mtase activity is retained strongly suggests that p26 retains enzymic activity. Alternatively, both p26 and p12 fragments may be required for activity and remain associated under these conditions.

Although only limited information is available about critical residues within the *Eco*RI Mtase (or any other Mtase), all residues previously implicated in function reside in the p26 fragment. The highly conserved Asp(Asn)ProProTyr(Phe) tetrapeptide found in the majority of adenine Mtases (Chandrasegaran & Smith, 1987), the Gly-X-Gly tripeptide found in all AdoMet-dependent enzymes (Ingrosso et al., 1989), the region implicated in AdoMet binding ($Gly^{205}$ to $Pro^{220}$) (Reich & Everett, 1990), and the Asn-Cys dipeptide found in most adenine Mtases and implicated in the catalysis (Everett et al., 1990) are shown in Figure 9 to occur within the p26 fragment.

*Relationship of Our Results to Prior Characterization of EcoRI Mtase and other DNA Mtases.* The *Eco*RI Mtase is a highly efficient enzyme with a $k_{cat}/K_m$ for plasmid DNA greater than $4 \times 10^8$ $s^{-1}$ $M^{-1}$ (Reich & Mashhoon, 1991), while at the same time showing a high degree of substrate specificity.[1] The molecular details for the discriminating mechanism(s) are unknown, but results from our proteolysis experiments may provide some insight. Our data suggest the Mtase contains either two domains connected by a flexible hinge region making up all or part of the active site or a single domain protein containing a flexible loop. Binding of AdoMet and DNA causes significant conformational changes that enable (1) the reorientation of the polypeptide-containing elements critical for DNA recognition ($Asn^{139}ProProPhe^{142}$), AdoMet binding ($Gly^{205}$ to $Pro^{220}$), and catalysis ($Asn^{222}Cys^{223}$), and (2) the hinge/loop region to become less flexible. Although the precise conformational changes are unknown, the dramatic (up to 2000-fold) changes in susceptibility to trypsinolysis and the unique fragmentation patterns under various conditions suggest significant structural changes. Many proteins which interact with DNA in a sequence-specific fashion are known to undergo limited conformational changes upon DNA binding (Jordan & Pabo, 1988; Rafferty et al., 1989), although the tryptophan repressor is known to undergo ligand-induced structural changes (Lawson et al., 1988; Otwinowski et al., 1988) and recent evidence for yeast transcriptional activator (Weiss et al., 1990), the protein products of the c-*fos* and c-*jun* protooncogenes (Patel et al., 1990), bacteriophage Cro protein (Brennan et al., 1990), and *lac* repressor (Pace et al., 1990) all suggest these proteins undergo significant ligand-induced conformational changes. One interesting possibility is that such conformational changes in the *Eco*RI DNA Mtase system may be critical for sequence recognition, since the Mtase appears to have distinct conformations when bound to the methylated versus native canonical sites (Figure 7).

Previous structure–function analyses involving DNA Mtases have relied on comparisons of primary sequences (Lauster, 1989; Posfai et al., 1989; Lauster et al., 1987; Chandrasegaran & Smith, 1987). These studies identified relatively small peptide regions which occur more frequently and with higher degrees of homology in the cytosine C5 Mtases than in the adenine N6 Mtases. In all cases involving adenine Mtases, the identified peptide regions occur within the p26 fragment of *Eco*RI Mtase, supporting the possibility that this fragment is functional. Genetic manipulation of closely related cytosine Mtases showed that entire regions could be "switched" while retaining (some) enzymic function (Wilke et al., 1988). However, since there is so little sequence similarity between cytosine and adenine Mtases, these results have limited application to the adenine Mtases.

[1] N. O. Reich, J. Murphy, and F. Osti, unpublished observations.

## References


Becker, C. R., Efcavitch, J. W., Heiner, C. R., & Kaiser, N. F. (1985) *J. Chromatogr. 326*, 293–299.

Bennett, W. S., & Steitz, T. A. (1980) *J. Mol. Biol. 140*, 211.

Brennan, R. G., Roderick, S. L., Takeda, Y., & Matthews, B. W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8165–8169.

Chandrasegaran, S., & Smith, H. O. (1987) in *Structure and Expression: From Proteins to Ribosomes* (Sarma, R. H., & Sarma, M. H., Eds.) Vol. I, pp 149–156, Adenine Press, Schenectady, NY.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Egner, U., Tamasselli, A. G., & Schulz, G. E. (1987) *J. Mol. Biol. 195*, 649.

Eklund, H., Samama, J. P., Wallen, L., Branden, C. I., Akeson, A., & Jones, T. A. (1981) *J. Mol. Biol. 146*, 561.

Everett, E., Falick, A. M., & Reich, N. O. (1990) *J. Biol. Chem. 265*, 17713–17719.

Fontana, A., Fassina, G., Claudio, V., Dalzoppo, D., Zamai, M., & Zambonin, M. (1986) *Biochemistry 25*, 1847–1851.

Ford, G. C., Eichele, G., & Jansonius, J. N. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2559.

Frederick, C. A., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., & Rich, A. (1988) *J. Biol. Chem. 263*, 17872–17879.

Galaktos, N. G., & Walsh, C. T. (1987) *Biochemistry 26*, 8475–8480.

Galaktos, N. G., & Walsh, C. T. (1989) *Biochemistry 28*, 8167–8174.

Greene, P. J., Heynecker, H. L., Bolivar, F., Rodriquez, R. L., Betlack, M. C., Covarrubias, A. A., Backman, K., Russel, D. J., Tait, R., & Boyer, H. W. (1978) *Nucleic Acids Res. 5*, 2373–2380.

Greene, P. J., Gupta, M., Boyer, H. W., Brown, W. E., & Rosenberg, J. M. (1981) *J. Biol. Chem. 256*, 2143–2153.

Horowitz, P. M., & Bowman, S. (1987) *J. Biol. Chem. 262*, 14544–14548.

Ingrosso, D., Fowler, A. V., Bleibaum, J., & Clarke, S. (1989) *J. Biol. Chem. 264*, 20131–20139.

Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.

Laemmli, V. K. (1970) *Nature 227*, 680–685.

Lauster, R. (1989) *J. Mol. Biol. 206*, 313–321.

Lauster, R., Kriebardis, A., & Guschlbauer, W. (1987) *FEBS Lett. 220*, 167–176.

Lawson, C. L., Zhang, R.-G., Schevitz, R. W., Otwinoswki, Z., Joachimiak, A., & Sigler, P. B. (1988) *Proteins: Struct., Funct., Genet. 3*, 1831.

LeGendre, N., & Matsudaira, P. (1988) *BioTechniques 6*, 154–159.

McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science 234*, 1526–1541.

Modrich, P., & Roberts, R. J. (1982) in *Nucleases* (Linn, S. M., & Roberts, R. J., Eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321–329.

Pace, H. C., Lu, P., & Lewis, M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1870–1873.

Patel, L., Abate, C., & Curran, T. (1990) *Nature 347*, 572–574.

Posfai, J., Bhagwat, A. S., Posfai, G., & Roberts, R. J. (1989) *Nucleic Acids Res. 17*, 2421–2435.

Rafferty, J. B., Somers, W. S., Saint-Girons, I., & Philips, S. E. V. (1989) *Nature 341*, 705–710.

Razin, A., Cedar, H., & Riggs, A. D., Eds. (1984) *DNA Methylation: Biochemistry and Biological Significance*, Springer-Verlag, New York.

Reich, N. O., & Everett, E. (1990) *J. Biol. Chem. 265*, 8929–8934.

Reich, N. O., & Mashhoon, N. (1990) *J. Biol. Chem. 265*, 8966–8970.

Reich, N. O., & Mashhoon, N. (1991) *Biochemistry* (preceding paper in this issue).

Remington, G., Wiegand, G., & Huber, R. (1982) *J. Mol. Biol. 158*, 111.

Rubin, R. A., & Modrich, P. (1977) *J. Biol. Chem. 252*, 7265–7272.

Usdin, E., Borchardt, R. T., & Creveling, C. R., Eds. (1979) *Transmethylation*, Vol. 5, Elsevier/North-Holland, New York.

Weiss, M. A., Ellenberger T., Wobbe, C. R., Lee, J. P., Harrison, S. C., & Struhl, K. (1990) *Nature 347*, 575–578.

Wilke, K., Rauhut, E., Noyer-Weidner, M., Lauster, R., Pawlek, B., Behrens, B., & Trautner, T. A. (1988) *EMBO J. 7*, 2601–2609.

Wing, R., Drew, H. R., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 287*, 755–758.