Critical Reviews in Biochemistry and Molecular Biology, 31(5&6):361–380 (1996)

Chemistry and Biology of DNA Methyltransferases

*Ishtiyaque Ahmad and Desirazu N. Rao**

Department of Biochemistry, Indian Institute of Science, Bangalore 560 012, India

* Address correspondence to Dr. D. N. Rao, Department of Biochemistry, Indian Institute of Science, Bangalore 560 012, India.

Referee: Dr. Noreen E. Murray, Institute of Cell and Molecular Biology, Univ. of Edinburgh, Scotland

ABSTRACT: Recognition of a specific DNA sequence by a protein is probably the best example of macromolecular interactions leading to various events. It is a prerequisite to understanding the basis of protein-DNA interactions to obtain a better insight into fundamental processes such as transcription, replication, repair, and recombination. DNA methyltransferases with varying sequence specificities provide an excellent model system for understanding the molecular mechanism of specific DNA recognition. Sequence comparison of cloned genes, along with mutational analyses and recent crystallographic studies, have clearly defined the functions of various conserved motifs. These enzymes access their target base in an elegant manner by flipping it out of the DNA double helix. The drastic protein-induced DNA distortion, first reported for *HhaI* DNA methyltransferase, appears to be a common mechanism employed by various proteins that need to act on bases. A remarkable feature of the catalytic mechanism of DNA (cytosine-5) methyltransferases is the ability of these enzymes to induce deamination of the target cytosine in the absence of *S*-adenosyl-L-methionine or its analogs. The enzyme-catalyzed deamination reaction is postulated to be the major cause of mutational hotspots at CpG islands responsible for various human genetic disorders. Methylation of adenine residues in *Escherichia coli* is known to regulate various processes such as transcription, replication, repair, recombination, transposition, and phage packaging.

KEY WORDS: DNA-protein interactions, DNA methyltransferase, sequence comparison, base flipping, catalytic mechanism

I. INTRODUCTION

Enzymatic transfer of a methyl group from *S*-adenosyl-L-methionine (AdoMet) to certain nucleotides in DNA is the most common form of biological DNA modification. The modification of DNA by methylation occurs in organisms as diverse as bacteria, plants, and mammals. The C5 or N4 position of cytosines and N6 position of adenines are the natural modification sites (Anderson, 1993). Methylation at the C5 position of cytosines is the only methylation reported in higher eukaryotes (Adams, 1990) and is involved in many regulatory processes (Martienssen and Richards, 1995). Methylation of cytosine and adenine residues in

1040-9238/96/$.50

prokaryotes is primarily involved in restriction-modification (R-M) systems that serve as immune responses to phage infection (Wilson and Murray, 1991). Adenine methylation in prokaryotes is also involved in regulating the initiation of DNA replication (Boye and Lobner-Olesen, 1990; Campbell and Kleckner, 1990; Slater et al., 1995) and in targeting the correction of errors in DNA replication (Modrich, 1989). R-M systems have served as an excellent model system for studying the structural aspects of protein-DNA interactions. Based on the type of methylation catalyzed and amino acid sequences, DNA methyltransferases (MTases, or M) can be divided into those that methylate the endocyclic carbon-5 of cytosines (m5C-MTases or DNA (cytosine-5 methyltransferase) and those that methylate the exocyclic amino group of adenines and cytosines (m4C- and m6A-MTases or DNA (cytosine-N4) MTases and DNA (adenine-N6) MTases, respectively]). While m5C-MTases have been characterized to a large extent, very little information is available regarding the other two classes of MTases. Several excellent review articles have appeared in recent times, but they addressed special aspects of DNA MTases or DNA methylation (Smith, 1994; Bestor and Verdine, 1994; Roberts, 1995; Cheng, 1995). In addition, the outburst of information in the area demanded a collective revision of the subject. In this article, we have tried to present a brief and updated account of DNA MTases.

II. DNA RECOGNITION

Sequence-specific recognition of DNA by proteins plays a major role in regulating many processes. Interaction of classical DNA-binding proteins such as transcription factors and repressors has been investigated in great detail and two motifs, helix-turn-helix and zinc finger, have been documented to mediate tight DNA binding (Harrison, 1991). These classical DNA-binding proteins need to bind to specific DNA sequences tightly and facilitate the assembly of other factors. In contrast, proteins with enzymatic activities bind to a specific DNA sequence, catalyze the reaction, and dissociate from DNA. Such proteins bind to DNA with moderate affinity, and this would suggest the difference in the mode of DNA recognition by these two diverse group of DNA-binding proteins (Dubey and Roberts, 1992).

DNA MTases with varying sequence specificities provide an excellent model system for investigating the molecular mechanisms of sequence-specific DNA binding. Most DNA MTases have been reported to bind to their cognate DNA sequence more strongly compared with noncognate sequences in the presence of AdoMet or its analogs (Bergerat and Guschlbauer, 1990; Dubey and Roberts, 1992; Reich et al., 1992; Powell et al., 1993; Szczelkun and Connolly, 1995). However, M.*Eco*R124I binds to its cognate sequence at least 50-fold more strongly compared with noncognate sites even in the absence of any cofactor (Taylor et al., 1992). We found that M.*Eco*P15I bound to its cognate site approximately threefold more tightly compared with noncognate sites in the presence of sinefungin. Interestingly, in the presence of ATP, M.*Eco*P15I exhibited a significant level of sequence selectivity (Ahmad and Rao, 1994b).

Studies so far have suggested that DNA MTases bind to their target sequence in an asymmetric fashion. In order to define the binding-site size for M.*Msp*I by DNaseI footprinting, Dubey and Roberts (1992) observed an asymmetric mode of DNA recognition. An extensive footprinting study by Szczelkun et al. (1995) suggested that M.*Eco*RV binds to its target site via the

major groove and in an asymmetric fashion. We have made similar observations with M.*Eco*P15I (Ahmad and Rao, 1996). The asymmetry in DNA recognition by DNA MTases is not surprising, as the natural substrate for these enzymes is a hemimethylated DNA and thus asymmetric in nature. This could be one reason why most of the DNA MTases characterized so far recognize DNA as a monomer. In contrast, M.*Eco*P15I exists as a dimer in solution, possibly due to the nature of its target site, which is unmethylated DNA (Ahmad et al., 1995). A few other reports in the literature describe the dimeric nature of M.*Dpn*A and M.*Rsr*I (de la Campa et al., 1987; Kaszubska et al., 1992).

A remarkable feature of the MTase-DNA complex is flipping out of the target base from the DNA double helix (Figure 1). The phenomenon of base flipping has been directly demonstrated in at least two m5C-MTases, M.*Hha*I and M.*Hae*III, and one DNA repair enzyme (Klimasauskas et al., 1994; Reinisch et al., 1995; Vassylyev et al., 1995). The target-base flipping could be a widespread phenomenon and may not be restricted to one particular class of enzymes. Indeed, model-building studies have suggested that m6A-MTases as well as DNA repair enzymes might also flip out the base to catalyze the reaction (Labahn et al., 1994; Dryden et al., 1995; Roberts, 1995). The binding of M.*Eco*P15I to its cognate site (CAGCAG) resulted in hypermethylation of the second guanine, which is adjacent to the target adenine. We proposed that the hypermethylation of the second guanine could be due to target-base flipping (Ahmad and Rao, 1996). It may not be premature to generalize that most DNA MTases bind to their cognate site in a sequence-specific manner in the presence of AdoMet or its analogs and flip the target base out of the DNA double helix to catalyze the base methylation.

III. DNA (CYTOSINE-5) METHYLTRANSFERASES

m5C-MTases are a class of enzymes that catalyze the transfer of a methyl group to position 5 of cytosine in DNA. From an evolutionary point, m5C-MTases appear to be unique and can be found in eukaryotes and prokaryotes. They share a set of well-conserved sequence motifs that simplify their identification from primary sequence data and serve as a natural target for functional studies. These enzymes have been well characterized, and their functional significance is reasonably well understood.

A. Primary Structure

More than 50 genes of m5C-MTases have been cloned and sequenced from various sources. Sequence comparisons among the members of this group have revealed ten well-conserved blocks of amino acids (Posfai et al., 1989; Lauster et al., 1989; Kumar et al., 1994). This reflects a common architecture and reaction mechanism for these enzymes. Of the ten conserved motifs, motif I (FxGxG) or a close relative has been identified in all the AdoMet-dependent methyltransferases. Mutational analysis showed that replacement of glycine residues by other amino acids in the *Bacillus subtilis* phage SPR MTase destroyed the general methylation capacity and therefore was postulated to be involved in AdoMet binding (Wilke et al., 1988). This was confirmed from the crystal structure of M.*Hha*I and M.*Taq*I (an adenine MTase) (Cheng et al., 1993; Klimasauskas et al., 1994; Labahn et al., 1994). Motif IV possesses an invariant Pro-Cys dipeptide that is known to be involved in methyl group transfer (Wyszynski et al., 1993; Mi and Roberts, 1993; Chen et al.,

1993). The region between motif VIII and motif IX was found to be highly diverse and was termed a variable region. The variable region was shown to be involved in target DNA recognition (Balganesh et al., 1987; Klimasauskas et al., 1991). Hybrid swap experiments in these MTases established that both the variable region and the choice of the base to be methylated within the target sequence determined the sequence-specificity (Klimasauskas et al., 1991; Mi and Roberts, 1992). In M.*Aqu*I, an unusually large variable region was shown to be partly dispensable. Deletion of 30% of the variable region did not affect enzyme activity or target selectivity (Zhang et al., 1993). A detailed analysis of the variable region of multispecific m5C-MTases defined the amino acid residues involved in specific DNA recognition (Trautner et al., 1996). Hybrid swap experiments demonstrated that the amino terminal part of the variable region was essential for binding to the 5′ region of the recognition sequence (Lange et al., 1996). The situation here is similar to type I DNA MTases, where the two variable regions are interrupted by a linker sequence in protein required to recognize a bipartite sequence (Nagaraja et al., 1985). While characterizing the variable region of multispecific m5C-MTase M.*Bss*HII, Schumann et al. (1996) assigned one of the target-recognizing domains to the enzyme core instead of the variable region.

B. Reaction Mechanism

Wu and Santi (1985) proposed that the catalytic mechanism of m5C-MTases could be analogous to the mechanism of thymidylate synthase. The methyl group transfer from AdoMet to the C5 position of the target cytosine in the cognate sequence by these enzymes involves a covalent intermediate between the C6 position of the target cytosine and the catalytic thiol group of the protein. The reaction is initiated by nucleophilic attack at the C6 position of the target cytosine by the free thiol group of the catalytic cysteine of motif IV. This results in the formation of a covalent Michael adduct. In this intermediate, the C5 position of the target cytosine becomes sufficiently nucleophilic to accept a methyl group from AdoMet. A proton is eliminated from the C5 position, and a conjugate elimination step releases the product. Under experimental conditions, enzyme-catalyzed reactions are so fast that it is difficult to observe these intermediate structures. To demonstrate that the mechanism proposed was valid, Santi and colleagues synthesized a cytosine analog, 5-fluoro-2-deoxycytidine (FdC), and incorporated it into DNA at the target cytosine position. Once the methyl group is transferred to this substrate, abstraction of F^+ is not possible, resulting in a stable covalent adduct formation. The mechanism-based inhibitor FdC inhibits the β-elimination step, and hence the covalent intermediate can be trapped in the reaction. Using DNA containing this analog, intermediates formed by several bacterial enzymes have been trapped (Osterman et al., 1988; Friedman and Ansari, 1992; Smith et al., 1992; Hanck et al., 1993). Earlier, it was argued that this mechanism could only be attained at the expense of substantial distortion in DNA structure. This was because the required trajectory for attack of the thiolate on the C6 position was blocked by the DNA backbone and the path for methyl group delivery at the C5 position was blocked by the neighboring DNA bases (Chen et al., 1991). It was later argued that protonation of the N3 position of the cytosine by an enzyme-derived acid would require strand separation. Verdine and colleagues synthesized a DNA substrate where both strands were cross-linked by a disulfide bond. When

this substrate was used, M.*Hae*III catalyzed the methyl group transfer (Erlanson et al., 1993).

The cysteine residue of motif IV (PCQ) has been shown to be involved in covalent complex formation. This has been directly demonstrated by isolating and sequencing the modified peptide (Chen et al., 1991; Friedman and Ansari, 1992). Mutation of this cysteine in many MTases abolishes the enzyme activity without affecting DNA recognition and cofactor binding (Wilke et al., 1988; Wyszynski et al., 1993; Mi and Roberts, 1993). Replacement of the catalytic cysteine with a serine in M.*Eco*RII resulted in a partially active enzyme with catalytic efficiency 10,000-fold less than the wild-type enzyme (Gabbara et al., 1995), suggesting that serine could also act as a nucleophile. It was postulated that a histidine residue activates the serine residue in the mutant enzyme and possibly the catalytic cysteine in the wild-type enzyme.

All studies so far clearly suggest that for type II DNA MTases, a single polypeptide is essential and sufficient to catalyze the reaction. However, a complexity was added to this myth when two peptides were identified to be essential for transfer of a methyl group in the case of M.*Aqu*I and M.*Eco*HK31I (Karreman and Waard, 1990; Lee et al., 1995, 1996). Functional M.*Aqu*I is composed of two polypeptides wherein the larger polypeptide consists of motif I-VIII, but motif IX and X are present in the smaller polypeptide. It is logical to assume that the target recognition domain (between motifs VIII and IX) would be formed once the two polypeptides associate together. This situation is analogous to the binding of DNA between two domains, as revealed by the crystal structures of M.*Hha*I and M.*Hae*III (Klimasauskas et al., 1994; Reinisch et al., 1995). In M.*Eco*HK31I, apart from motif IX, which is contributed by a smaller polypeptide, the other nine conserved motifs are present in the larger polypeptide (Lee et al., 1995).

C. Three-Dimensional Structure

M.*Hha*I from *Haemophilus haemolyticus* and M.*Hae*III from *H. aegyptius* are the only m5C-MTases for which three-dimensional structures have been solved. The structure of M.*Hha*I was solved in a binary complex with AdoMet at 2.5 Å resolution and in a covalent ternary complex at 2.8 Å resolution (Cheng et al., 1993; Klimasauskas et al., 1994). The crystal structure of a binary complex of M.*Hha*I with AdoMet revealed the relationship between the tertiary structure and the ten highly conserved sequence motifs. The structure does not show any recognizable DNA-binding motifs. The molecule is folded into two domains: a large catalytic domain containing the catalytic and cofactor-binding sites and a small DNA recognition domain implicated in sequence-specific DNA recognition. Residues from all five of the amino-terminal motifs (I-V) and C-terminal motif X contribute to the AdoMet-binding pocket. Motif I (FxGxG) forms a tight loop in the first turn of a β1-αA-β2 structural unit. The conserved glycine residues appear important in allowing a tight turn to form in the structure. The conserved phenylalanine interacts with the adenosyl moiety of the cofactor. The combined region of motifs I to III around the AdoMet-binding pocket bears a strong resemblance to the Rossmann fold of the NAD-binding motif. The structure of the ternary complex revealed that the DNA substrate is bound in a cleft formed by the two domains of the protein (Klimasauskas et al., 1994). One of the most interesting observations made from this structure was that the target cytosine flipped out of the DNA double helix by 180° into the pocket of the

enzyme (Figure 1), where the chemistry of methylation takes place. As mentioned earlier, the stereochemistry of the methylation reaction dictates the spatial relationship between the reacting groups. In a normal B DNA, the C5 of the target cytosine is too deeply buried in the helix to allow this reaction to proceed, and it was therefore argued that distortion of the helix was necessary. The distortion that occurs in m5C-MTases is the flipping of the target cytosine out of the helix and into the catalytic site, without severely distorting the rest of the DNA helix. Two major changes take place upon DNA binding: the complete disruption of the target base-pair and a conformational change in the active-site loop. Glutamine 237 of M.*Hha*I occupies the gap left by the target cytosine after it has flipped out of the helix and makes contact with the orphan guanine through hydrogen bonds. Subsequent closing of the active-site loop on the target brings an additional residue (Ser 87) into the DNA helix to stabilize the conformation of the glutamine loop. The two loops restore the structure in the DNA. The glutamine residue is critical for protein-DNA complex stabilization but not for the methyl group transfer, as revealed by mutational analysis (Mi et al., 1995). The new position of the active-site loop is maintained by covalent bond formation between Cys 81 and the C6 atom of the target cytosine (Klimasauskas et al., 1994).

The crystal structure of M.*Hae*III, bound covalently to its cognate DNA site (GGCC), is very similar to the structure of M.*Hha*I in many respects (Reinisch et al., 1995). In addition to target-base flipping, the DNA double helix undergoes further rearrangement, resulting in the loss of the characteristic B form of DNA (Figure 1). The guanine opposite the target cytosine base pairs with the last cytosine in the cognate

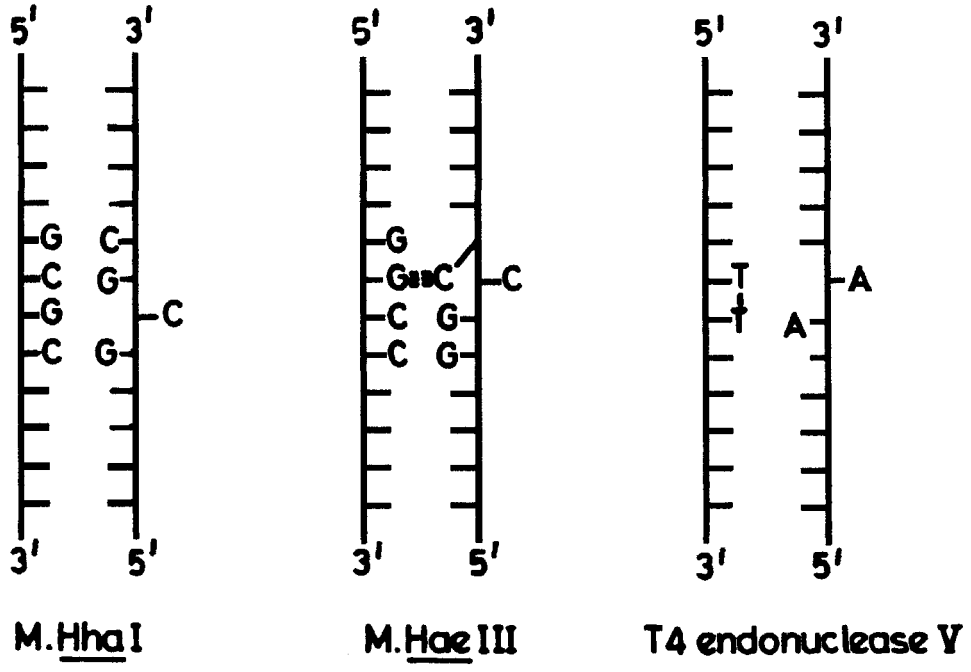

FIGURE 1. Isolated DNA structures from complexes of M.*Hha*I, M.*Hae*III, or T4 endo V with their target sites. The schematic diagram highlights only a special feature (i.e., base flipping) and does not truly represent the DNA structures in complex with the respective enzymes. In the diagram, only the target sequence is represented.

sequence, thus making the first guanine an orphan base. M.*Hae*III differs from M.*Hha*I in that the former promotes both the cytosine extrusion and the reorganization of bases. As a result, the bases are unstacked and a gap 8 Å long opens in the DNA (Reinisch et al., 1995). One of the remarkable features of the structure, base flipping by an enzyme, could occur in other classes of MTases as well as DNA repair enzymes (Roberts, 1995). Indeed, the recent crystallographic analysis of the T4 endonuclease V-DNA complex revealed base flipping by the enzyme (Figure 1) (Vassylyev et al., 1995). However, there are interesting differences between the structure of the DNA complexed to m5C-MTases and that to the repair enzyme. In the case of m5C-MTases, the target cytosine base flips out of the DNA double helix into a cavity of the enzyme where the chemistry of methylation takes place (Klimasauskas et al., 1994; Reinisch et al., 1995). Thus, base flipping by MTases seems to be a specific phenomenon. In contrast, base flipping by T4 endonuclease V seems to be nonspecific, as it is the base opposite to the target base that flips out into a cavity of the enzyme. Here, the chemistry of the reaction does not occur in the cavity of the enzyme, but in the cavity of the DNA formed due to base flipping (Vassylyev et al., 1995). There is strong evidence to support a base-flipping model in the case of other DNA repair enzymes such as uracil *N*-glycosylase (Mol et al., 1995) and Ada protein (Moore et al., 1994), but direct visualization is still awaited.

D. Other Reactions Catalyzed by m5C-MTases

It was earlier proposed by Selker (1990) that the enzymes that methylate cytosine could also deaminate it to uracil in the absence of AdoMet. This suggestion was based on the similarity of an intermediate in the bisulfite-catalyzed deamination of cytosine to a product formed by a side reaction of the m5C-MTases. Using a combination of genetic and biochemical assays, it was shown that M.*Hpa*II drastically increased the rate of target cytosine deamination in the absence of cofactor (Shen et al., 1992). As expected, AdoMet inhibited the deamination reaction. Interestingly, AdoHcy also inhibited the rate of deamination. The role of cofactor in the deamination reaction catalyzed by m5C-MTases was further probed by mutational analysis of motif I in M.*Hpa*II. The mutant enzymes did not bind the cofactor and exhibited a mutator phenotype. These results further substantiated the earlier observation that the binding of cofactor inhibits the enzyme-catalyzed deamination reaction (Shen et al., 1995). A possible explanation given was that a water molecule essential for deamination was excluded from the enzyme active site in the presence of either AdoMet or AdoHcy. An alternative explanation could be that in a binary MTase-DNA complex, the target cytosine is more prone to deamination. The binding of AdoMet or its analogs would lead to a different conformation in the enzyme that may not favor the deamination reaction anymore. Enzyme-catalyzed cytosine deamination is postulated to be a major cause of mutational hotspots in the genome (Figure 2). However, the highly efficient uracil *N*-glycosylase would remove any of the uracil formed. It was therefore argued that enzyme-catalyzed deamination could not be a major cause of mutational hotspots at CpG islands in cells (Wyszynski et al., 1994). Using a set of duplex oligonucleotides containing a mismatch at the target base-pair, it was shown that M.*Hha*I and M.*Hpa*II bind to mismatch oligonucleotides much better than normal recognition sequences. Further, both these enzymes bind strongly to U:G

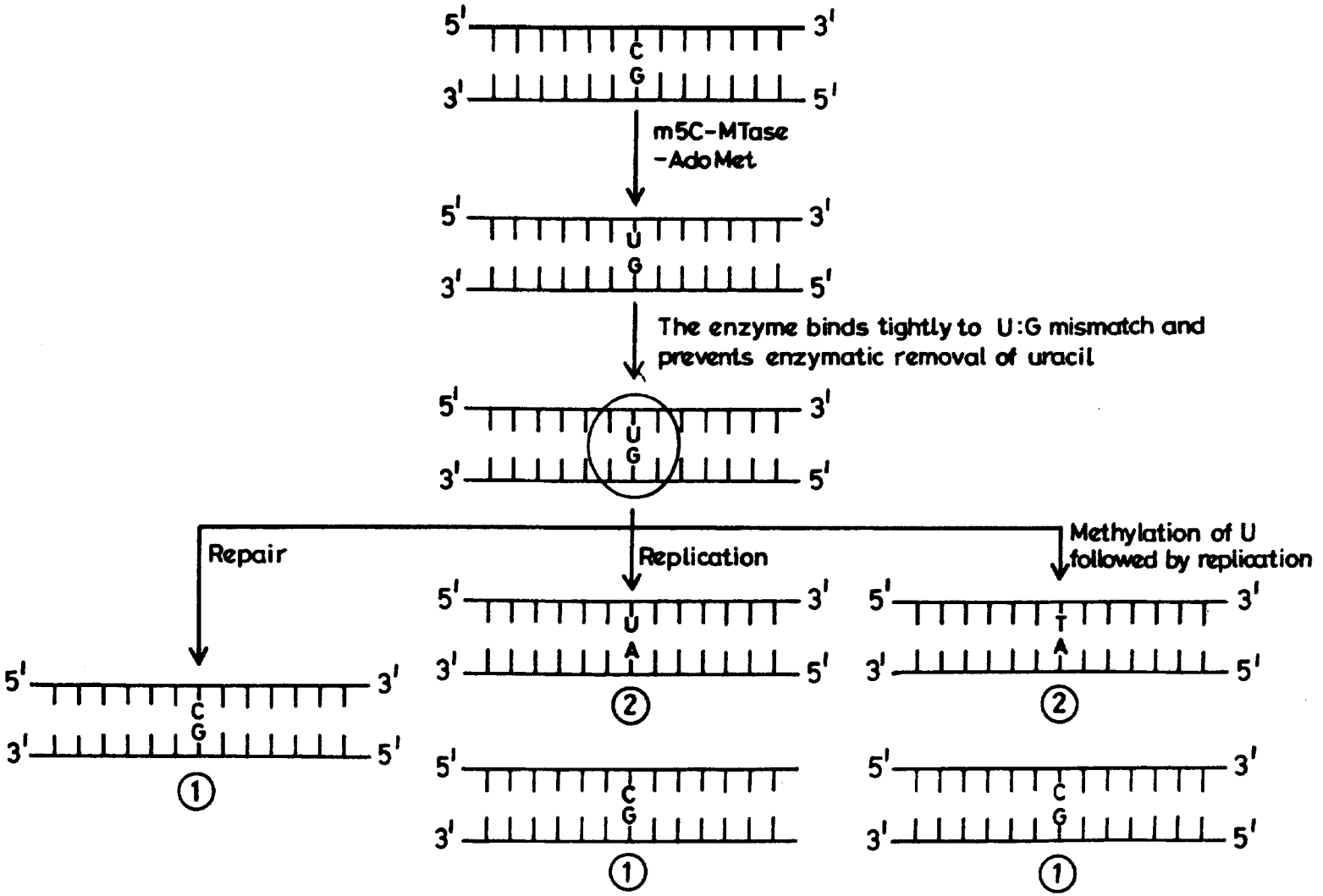

FIGURE 2. A model to explain the consequences of enzymatic deamination of the target cytosine. Only the target basepair is shown. As a result of deamination, DNA can follow any of the three pathways: (1) enzymatic repair of the DNA, (2) DNA replication resulting in two genetically distinct species, or (3) enzymatic methylation of uracil to thymine, followed by DNA replication. Products (1) represent the DNA where correct genetic information is restored, whereas products (2) are genetically altered molecules.

mismatch, also a product of deamination of target cytosine, and slowly methylated the uracil to thymine (Klimasauskas and Roberts, 1995; Yang et al., 1995). *In vivo,* at low AdoMet concentrations, these MTases would deaminate the target cytosine to uracil and bind to U:G mismatch strongly, thus preventing the enzymatic removal of uracil (Figure 2). This would probably contribute to the hypermutatable phenotypes at CpG islands. We propose a model to explain the phenomenon of enzyme-catalyzed deamination that agrees with the available biochemical and genetic data (Figure 2). It is known that most of the MTases have evolved as part of R-M systems. At low AdoMet concentrations, the target base would not be methylated, making it susceptible to corresponding restriction endonucleases. This would lead to degradation of the host genome and self-destruction of the host. To avert this suicidal phenomenon, MTases incorporate mutations in the cognate site by deaminating the target cytosine. This results in a C:G to T:A transition, which cannot serve as a substrate for the corresponding endonuclease (products 2 in Figure 2). All these activities — deamination of target cytosine, binding to mismatch DNA, and methylation of uracil to thymine — are related to the reaction mechanisms as well as to the evolution of these enzymes. As suggested by Roberts (1995), these enzymes might have evolved from mismatch DNA-binding proteins.

Another activity exhibited by at least one of the cytosine MTases, M.*Sss*I, is a topoisomerase activity in the presence of Mg^{2+} (Matsuo et al., 1994). Supportive evidence for this activity comes from the sequence comparison of M.*Sss*I and several topoisomerases. The enzyme possesses the catalytic tyrosine residue essential for the topoisomerase activity (Matsuo et al., 1994).

E. Functional Significance

One of the best understood and undisputed roles of DNA methylation in prokaryotes is that of protecting the host DNA from degradation by the corresponding endonucleases (Wilson and Murray, 1991). ENases (endonucleases) and MTases usually occur together as a component of R-M systems. Less frequently they are identified in a solitary condition that may or may not be relevant to restriction and modification. Some viruses encode MTases of the same specificity as that of the R-M system encoded by their host, which helps them to evade the host. An unaccompanied m5C-MTase, Dcm, has been identified in *Escherichia coli* that recognizes and modifies the same DNA sequence as that of M.*Eco*RII (Schlagaman et al., 1976). The enzyme has been shown to function to the disadvantage of the host by increasing the mutagenic potential of the target site. Methylation of the target cytosine followed by spontaneous or enzyme-catalyzed deamination would lead to a T:G mispair that, if left uncorrected, would alter the genetic information. However, the host counteracts this potentially dangerous act in the form of a very short patch repair (*vsr*) gene. The *vsr* gene encodes a protein product that corrects any T:G mismatch created in the genome (Sohail et al., 1990).

Although the exact function of DNA methylation in eukaryotic cells is not fully understood, it has been implicated in the control of a number of cellular processes, including transcription, genomic imprinting, developmental regulation, mutagenesis, transposition, DNA repair, X-inactivation, and chromatin organization (Boyes and Bird, 1991; Lewis and Bird, 1991; Pfeifer et al., 1990; Leonhardt et al., 1992). Abnormal methylation is associated with various human genetic disorders such as cancer and

fragile X syndrome (Jones et al., 1992; Pieretti et al., 1991).

IV. DNA (CYTOSINE-N4) METHYLTRANSFERASES

These enzymes have been reported only from prokaryotes and are part of R-M systems (Wilson, 1992). They modify the target cytosine to 4-methylcytosine. Only a few members are known in this group, of which typical examples are M.*Sma*I and M.*Bam*HI (Heidman et al., 1989; Brooks et al., 1991). Although these enzymes methylate cytosine residues in their cognate sequence, they do not share any sequence motifs with the m5C-MTases except for motif I. It is believed that the motif I in these enzymes is involved in AdoMet binding because these enzymes also use AdoMet as a methyl group donor. These enzymes harbor a common motif similar to motif IV of the adenine MTases. Motif IV (SPPY in m4C- and D/NPPY/F in m6A-MTases) has been implicated in catalysis because these enzymes catalyze similar kind of reactions (Wilson, 1992). These enzymes have been further classified into two groups based on the relative positions of motif I and IV. Enzymes of the α-subclass have motif I followed by motif IV in the primary amino acid sequence, whereas β-subclass enzymes have motif IV followed by motif I (Wilson, 1992).

V. DNA (ADENINE-N6) METHYLTRANSFERASES

m6A-MTases catalyze the transfer of a methyl group from AdoMet to the N6 position of a specific adenine in their cognate sequence. These enzymes are important components of R-M and mismatch repair systems in prokaryotes (Marinus, 1984). Although much is known about the biochemistry and biology of DNA adenine methylation, little is known about the catalytic mechanism of the methylation reaction. A few enzymes from this class have been well characterized.

A. Primary Structure

Sequence comparisons among the members of this group have revealed nine conserved motifs, of which motif I and motif IV are highly conserved (Malone et al., 1995). Analysis of amino acid sequences reveals that the exocyclic amino MTases belong to three groups distinguished by differences in the linear orders of the conserved motifs in their primary sequences (Figure 3) (Wilson, 1992). In the α subclass of enzymes, the motif I (FxGxG) is followed by motif IV (DPPY) in the primary sequence. Enzymes of the β-subclass have these motifs in reverse order i.e., motif IV (DPPY) followed by motif I (FxGSG) whereas in the γ-subclass, motif I (GxG) is followed by motif IV (NPPY). Between subclasses and within a class there are certain variations in amino acid sequences, but these are mainly conservative changes. The functional significance of these motifs has been probed by site-directed mutagenesis. It was first observed in M.*Fok*I that the replacement of an asparagine residue in the NPPY motif by either glycine or alanine destroyed catalytic activity (Sugisaki et al., 1989). Mutation of any of these residues in three m6A-MTases — T4 Dam MTase, M.*Eco* Dam, and M.*Eco*KI — resulted in distinct phenotypes (Kossykh et al., 1993; Guyot et al., 1993; Willcock et al., 1994). Altering motif IV of T4 Dam MTase from DPPY to DAPY or DTPY substantially in-

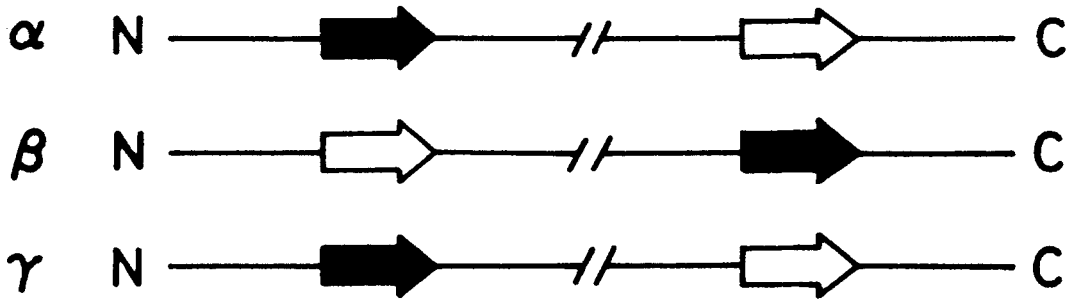

FIGURE 3. Arrangement of conserved motifs I (FxGxG) and IV [(D/N/S)PP(Y/F)] in the primary structure of m6A-MTases. Filled arrows indicate motif I, and open arrows indicate motif IV.

creased the K_m for AdoMet but had a smaller effect on the K_{cat} and K_a for AdoMet and K_m for DNA (Kossykh et al., 1993). In M.*Eco* Dam, changing DPPY to DGPY, DVPY, DPGY, DPRY, DPQY, DPEY, or DPVY abolished activity (Guyot et al., 1993). We have shown recently that replacement of the tyrosine residue in motif IV (DPPY) with tryptophan in M.*Eco*P15I resulted in enhanced cross-linking of the mutant enzyme to AdoMet and DNA (Ahmad and Rao, 1996). Our results provide direct evidence that motif IV is close to AdoMet as well as the substrate DNA-binding site. All these results are consistent with the suggested role for motif IV in catalysis. This motif has been proposed to be analogous to the catalytic PCQ motif of m5C-MTases. Recently, it has been shown that the conserved NPPY motif in M.*Taq*I, an m6A-MTase, overlaps with the PCQ motif of M.*Hha*I, a m5C-MTase, almost exactly when the two structures are superimposed (Schluckebier et al., 1995).

In m6A-MTases, the role of motif I has not been probed in detail. It was shown in M.*Eco*KI that substitution of aspartic acid for the first glycine in motif I completely abolished AdoMet binding but did not alter the DNA-binding properties of the enzyme (Willcock et al., 1994). Substitution of the second glycine residue in motif I of M.*Eco*P15I with arginine or serine abolished AdoMet binding, leading to a loss of enzyme activity (Ahmad and Rao, 1996). In an attempt to identify the amino acid residues involved in AdoMet binding in M.*Eco*RI, Reich and Everett (1990) photolabeled the enzyme with 8-azido AdoMet. However, on sequencing the photolabeled peptide, it became clear that the amino acid(s) that were cross-linked did not represent any of the conserved motifs present in m6A-MTases. Similar results were obtained with M.*Eco* Dam (Wenzel and Guschlbauer, 1993). Irradiation of M.*Eco*P15I with short-wave UV light in the presence of [methyl-^{3}H]AdoMet resulted in covalent adduct formation. Limited proteolysis of the adduct revealed the role of the amino-terminal domain in AdoMet binding (Ahmad and Rao, 1994a). In a study to identify amino acid residues involved in DNA binding in m6A-MTases, photochemical cross-linking has been done with M.*Eco*P15I and M.*Eco*KI (Ahmad and Rao, 1994b; Chen et al., 1995; Ahmad and Rao, 1996). In the case of M.*Eco*KI, tyrosine 27 of the specificity subunit was cross-linked to bromouracil-substituted DNA (Chen et al., 1995). Photochemical cross-linking and mutational analysis of M.*Eco*P15I indicated the role of the cleft formed by the two domains in DNA recognition (Ahmad and Rao, 1996). These results are in agreement with the recent crystallographic analysis of

M.*Hha*I, M.*Hae*III, and M.*Taq*I, where the hinge region has been shown or implicated in DNA recognition (Klimasauskas et al., 1994; Labahn et al., 1994; Reinisch et al., 1995). In an attempt to probe the role of lysine residues in DNA recognition by M.*Eco*R124I, Taylor et al. (1996) identified six lysine residues of the specificity subunit critical in DNA binding.

B. Reaction Mechanism

Chemical methylation of an adenine involves the transfer of a methyl group to the N1 position of the adenine. The adduct undergoes N1-C2 bond cleavage, followed by C5-C6 bond rotation and dehydration to close the ring, which results in the formation of 6-methyladenine. This sequence of reactions is called the Dimroth rearrangement (Haines et al., 1964). The fact that the N1 position of an adenine is more basic and nucleophilic than the N6 position suggests that the Dimroth rearrangement might be a more likely mechanism than direct methylation. Using a dodecamer substrate for M.*Eco*RI with [6-^{15}N]adenine as the target base, Pogolotti et al. (1988) showed that M.*Eco*RI catalyzed the direct transfer of the methyl group from AdoMet to the N6 position of the target adenine. This mechanism was further confirmed by the observation that there was an inversion of the configuration of the methyl group when it was transferred from AdoMet to the adenine. This suggested that the reaction follows a S_N2 mechanism (Ho et al., 1991). These results argue against a ping-pong kinetic mechanism. In order to obtain better insight into the kinetic mechanism of M.*Eco*RI, Surby and Reich (1996b) determined the kinetic parameters of the enzyme using DNA substrates ranging in length from 14 bp to 4.36 kb. They observed a fourfold increase in the specificity constant of the enzyme when the length of the DNA substrate was increased from 14 bp to 4.36 kb. Further, they determined the affinity constant of the enzyme for various DNA substrates (Surby and Reich, 1996a). From such a study, it was concluded that the enzyme binds to nonspecific DNA sequences and searches for the specific DNA sequence by facilitated diffusion in a one-dimensional search.

Three reports in the literature describe a detailed kinetic mechanism for the m6A-MTases. M.*Eco*P15I shows a random bi-bi steady-state mechanisms, where AdoMet or DNA can bind first (Rao et al., 1989). In contrast, M.*Eco*RI follows an ordered bi-bi mechanism where AdoMet binding is followed by DNA binding (Reich and Mashhoon, 1991). However, in both cases, AdoHcy (*S*-adenosyl-L-homocysteine) acts as a competitive inhibitor with respect to AdoMet (Rao et al., 1989; Reich and Mashhoon, 1991). M.*Eca*I catalyzes the transfer of a methyl group from AdoMet to the adenine residue of the GGTNACC sequence with a random rapid equilibrium mechanism (Szilak et al., 1993).

C. Three-Dimensional Structure

M.*Taq*I is the only m6A-MTase for which a crystal structure is available. The enzyme recognizes the sequence 5′-TCGA-3′ and transfers a methyl group from AdoMet to the N6 position of adenine. The crystal structure has been solved for the MTase complexed with AdoMet (Labahn et al., 1994). The structure of this enzyme is similar to the structure of M.*Hha*I (Cheng et al., 1993). The enzyme folds into two domains of equal sizes. The amino-terminal domain, which carries both the essential motifs, binds to AdoMet. However, in this case, only the second glycine is the conserved residue

(ACAHG). The carboxy terminal domain is less structured than the amino terminal domain. The second glycine residue in motif I makes a sharp turn between α-helix B and β-strand 1. AdoMet is inserted into a cavity in the amino-terminal domain formed by the two conserved motifs I and IV. Motif I is at the carboxy terminal of β1 and is part of a classical Rossmann fold (β1-αB-β2) (Labahn et al., 1994). The shape of the cavity can accommodate the AdoMet in an extended conformation. Interestingly, motif IV (NPPY) also makes contacts with the methyl moiety of AdoMet. This explains why some of the earlier mutants in motif IV were unable to bind to AdoMet (Kossykh et al., 1993; Guyot et al., 1993) or showed enhanced photolabeling with AdoMet (Ahmad and Rao, 1996). The cleft formed between the two domains is 21 Å wide, which is equivalent to the diameter of a B DNA. Model-building studies suggest that the DNA binds in the cleft. In a normal B DNA, the target adenine is located 15 Å away from the methyl moiety of AdoMet. Therefore, a direct methyl group transfer onto the TCGA adenine residue in undistorted M.*Taq*I is not possible. Hence, it has been proposed that if the target adenine flips out of the DNA double helix, it will align the methyl group donor and acceptor, resulting in a direct methyl-group transfer.

D. Functional Significance

Like all DNA MTases that are a part of R-M systems, these MTases also methylate self-DNA to protect it from the corresponding endonucleases (Wilson and Murray, 1991). A solitary adenine methyltransferase from *E. coli*, M.*Eco* Dam, regulates many molecular events. M.*Eco* Dam recognizes the sequence 5′-GATC-3′ and transfers a methyl group from AdoMet to the N6 position of adenine (Lacks and Greenberg, 1977). Immediately after replication, these sites are transiently hemimethylated, and several events occur during that phase. If any of the misincorporations are left unnoticed by the proofreading activity of the DNA polymerase, the methyl-directed mismatch repair pathway operates to correct the mismatch. DNA adenine methylation determines strand selectivity for the methyl-directed mismatch repair system (Modrich, 1989).

Dam methylation plays an important role in chromosome replication and segregation in *E. coli* (Boye and Lobner-Olesen, 1990). There are eleven GATC sequences in the 245-bp origin of replication (*ori*C) region (Oka et al., 1980). Methylation of *ori*C is required for efficient initiation of chromosomal replication. The promoter element of the *dna*A gene also contains GATC sites. These GATC sites are transiently hemimethylated after the start of replication. Hemimethylated *ori*C and *dna*A are sequestered into the membrane by certain components of the membrane that have a high affinity for hemimethylated GATC sites (Ogden et al., 1988; Campbell and Kleckner, 1990; Slater et al., 1995). In addition, hemimethylated GATC sites in the promoter element of the *dna*A gene downregulate its own expression. The product of the *dna*A gene is essential for initiation of replication (Bramhill and Kornberg, 1988). The sequestration of *ori*C and *dna*A ensures single initiation of chromosome replication per cell division (reviewed by Crooke, 1995). In addition, the association of hemimethylated GATC sites has been implicated in proper chromosome segregation following cell division in *E. coli*.

There are a number of promoters whose methylation status seems to play a role in regulating the expression of genes. Methylation of GATC sites is reported to regulate

the expression of the *mom* gene of Mu phage (Hattman and Ives, 1984). The level of *mom* RNA is decreased at least 20-fold in dam$^-$ strains compared with dam$^+$ strains (Hattman, 1982). Similarly, Dam methylation regulates the Pap pili phase variation by regulating the expression of genes responsible for phase variation (Blyn et al., 1990; Braaten et al., 1994). A *dam* site overlaps with the -10 region of the inward promoter of Tn10, the product of which is a transposase. In *dam* mutants, transposase expression is increased about tenfold. In addition, the transposase cleaves the target site only if the GATC site in the target sequence is hemimethylated (Roberts et al., 1985; Kleckner, 1989). There are several *E. coli* genes whose expression is increased in *dam* mutants (Noyer-Widner and Trautner, 1993). It should be pointed out that no essential genes can be regulated solely via DNA methylation because *dam* mutants are viable.

Packaging of bacteriophage P1 requires the methylation of GATC sites in the *pac* region, which carries seven sites in a 160-bp fragment (Sternberg and Coulby, 1990). Methylation of these sites by either the host M.*Eco*Dam or P1 phase M.Dam ensures the cleavage of the *pac* site by the enzyme pacase after headful packaging of DNA into the capsid.

VI. CONCLUSIONS

In the past 2 to 3 years, our understanding of DNA MTases has grown exponentially. Three crystal structures have truly enlightened us on the mechanism of specific DNA sequence recognition and, more importantly, the mechanism of target-base modification. A novel mode of DNA recognition, the target-base flipping first reported for a m5C-MTase, seems to be a widespread phenomenon. Although it is a drastic protein-induced DNA distortion, base flipping is an elegant mechanism, especially if buried bases are to be modified.

Future research will definitely focus on m6A-MTases. It will be interesting to see how related and distinct are these two groups of DNA MTases. It is well understood that the catalytic mechanisms of m5C-MTases and m6A-MTases are different, as reflected in the primary structures of these enzymes. Base flipping by m6A-MTases is likely, but direct visualization is still awaited. MTases promise to be a rich source of information that can be unraveled by structural, biochemical, and genetic studies.

ACKNOWLEDGMENT*

The work in our laboratory was supported by DST, DAE, and CSIR (Government of India) grants to D.N.R. I.A acknowledges a CSIR Senior Research Fellowship, and D.N.R. acknowledges the Rockefeller Foundation Biotechnology Career Fellowship.

REFERENCES

Adams, R. L. P. 1990. DNA methylation. The effect of minor bases on DNA-protein interactions. *Biochem. J.* **265:** 309–320.

Ahmad, I. and Rao, D. N. 1994a. Photolabeling of *Eco*P15I DNA methyltransferase with *S*-adenosyl-L-methionine. *Gene* **142:** 67–71.

Ahmad, I. and Rao, D. N. 1994b. Interaction of *Eco*P15I DNA methyltransferase with oligonucleotides containing the asymmetric se-

* The author has asked to submit a Note in Proof that reflects recent information affecting this article. Please see after the references for this note.

quence 5′-CAGCAG-3′. *J. Mol. Biol.* **242:** 378–388.

Ahmad, I. and Rao, D. N. 1996. Functional analysis of conserved motifs in *Eco*P15I DNA methyltransferase. *J. Mol. Biol.* **259:** 229–240.

Ahmad, I., Krishnamurthy, V., and Rao, D. N. 1995. DNA recognition by the *Eco*P15I and *Eco*PI modification methyltransferases. *Gene* **157:** 143–147.

Anderson, J. E. 1993. Restriction endonucleases and modification methylases. *Curr. Opin. Struct. Biol.* **3:** 24–30.

Balganesh, T. S., Reiners, L., Lauster, R., Noyer-Weidner, M., Wilke, K., and Trautner, T. A. 1987. Construction and use of chimeric SPR/o3T DNA methyltransferases in the definition of sequence recognizing enzyme regions. *EMBO J.* **6:** 3543–3549.

Bergerat, A. and Guschlbauer, W. 1990. The double role of methyl donor and allosteric effector of S-adenosylmethionine for Dam methylase of *E. coli. Nucleic Acids Res.* **18:** 4369–4375.

Bestor, T. H. and Verdine, G. L. 1994. DNA methyltransferases. *Curr. Opin. Cell Biol.* **6:** 380–389.

Blyn, L. B., Braaten, B. A., and Low, D. A. 1990. Regulation of *pap* pilin phase variation by a mechanism involving differential Dam methylation states. *EMBO J.* **9:** 4045–4054.

Boye, E. and Lobner-Olesen, A. 1990. The role of *dam* methyltransferase in the control of DNA replication in *E. coli. Cell* **62:** 981–989.

Boyes, J. and Bird, A. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64:** 1123–1134.

Braaten, B. A., Nou, X., Kaltenbach, L. S., and Low, D. A. 1994. Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli. Cell* **76:** 577–588.

Bramhill, D. and Kornberg, A. 1988. A model for initiation at origins of DNA replication. *Cell* **54:** 915–918.

Brooks, J. E., Nathan, P. D., Landry, D., Sznyter, P., Waite-Rees, A., Ives, C. L., Moran, L. S., Slatko, B. E., and Benner, J. S. 1991. Characterization of the cloned *Bam*HI restriction modification system: its nucleotide sequence, properties of the methylase, and expression in heterologous hosts. *Nucleic Acids Res.* **19:** 841–850.

Campbell, J. L. and Kleckner, N. 1990. *E. coli oriC* and the *dnaA* gene promoter are sequestered from *dam* methyltransferase following the passage of the chromosomal replication fork. *Cell* **62:** 967–979.

Chen, L., MacMillan, A. M., Chang, W., Ezaz-Nikpay, K., Lane, W. S., and Verdine, G. L. 1991. Direct identification of the active site nucleophile in a DNA (cytosine-5) methyltransferase. *Biochemistry* **30:** 11018–11025.

Chen, L., MacMillan, A. M., and Verdine, G. L. 1993. Mutational separation of DNA binding from catalysis in a DNA cytosine methyltransferase. *J. Am. Chem. Soc.* **115:** 5318–5319.

Chen, A., Powell, L. M., Dryden, D. T. F., Murray, N. E., and Brown, T. 1995. Tyrosine 27 of the specificity polypeptide of *Eco*KI can be UV crosslinked to a bromodeoxyuridine-substituted DNA target sequence. *Nucleic Acids Res.* **23:** 1177–1183.

Cheng, X. 1995. Structure and function of DNA methyltransferases. *Annu. Rev. Biophys. Biomol. Struct.* **24:** 293–318.

Cheng, X., Kumar, S., Posfai, J., Pflugrath, W., and Roberts, R. J. 1993. Crystal structure of the *Hha*I DNA methyltransferase complexed with *S*-adenosyl-L-methionine. *Cell* **74:** 299–307.

Crooke, E. 1995. Regulation of chromosomal replication in *E. coli:* sequestration and beyond. *Cell* **82:** 877–880.

de la Campa, A. G., Purushottam, K., Springhorn, S. S., and Lacks, S. A. 1987. Proteins encoded by the *Dpn*II restriction gene cassette. Two methylases and an endonuclease. *J. Mol. Biol.* **196:** 457–469.

Dryden, D. T. F., Sturrock, S. S., and Winter, M. 1995. Structural modelling of a type I DNA methyltransferase. *Nature Struct. Biol.* **2:** 632–635.

Dubey, A. K. and Roberts, R. J. 1992. Sequence specific DNA binding by the *Msp*I DNA methyltransferase. *Nucleic Acids Res.* **20:** 3167–3173.

Erlanson, D. A., Chen, L., and Verdine, G. L. 1993. DNA methylation through a locally unpaired intermediate. *J. Am. Chem. Soc.* **115:** 12583–12584.

Friedman, S. and Ansari, N. 1992. Binding of the *Eco*RII methyltransferase to 5-fluorocytosine containing DNA. Isolation of a bound peptide. *Nucleic Acids Res.* **20:** 3241–3248.

Gabbara, S., Sheluho, D., and Bhagwat, A. S. 1995. Cytosine methyltransferase from *Escherichia coli* in which active site cysteine is replaced with serine is partially active. *Biochemistry* **34:** 8914–8923.

Guyot, J. B., Grassi, J., Hahn, U., and Guschlbauer, W. 1993. The role of preserved sequences of Dam methylase. *Nucleic Acids Res.* **21:** 3183–3190.

Haines, J. A., Reese, C. B., and Todd, L. 1964. The methylation of nucleosides and mononucleotides with diazamethane. *J. Chem. Soc.* 1406–1412.

Hanck, T., Schmid, S., and Fritz, H. J. 1993. Sequence-specific and mechanism-based cross-linking of Dcm DNA cytosine-5 methyltransferase of *E. coli* K-12 to synthetic oligonucleotides containing 5-fluoro-2′-deoxycytidine. *Nucleic Acids Res.* **21:** 303–309.

Harrison, S. C. 1991. A structural taxonomy of DNA binding domains. *Nature (London)* **353:** 715–719.

Hattman, S. 1982. DNA methyltransferase-dependent transcription of the phage Mu *mom* gene. *Proc. Natl. Acad. Sci. U.S.A.* **79:** 5518–5521.

Hattman, S. and Ives, J. 1984. S1 nuclease mapping of the phage Mu *mom* gene promoter: a model for the regulation of *mom* expression. *Gene* **29:** 185–198.

Heidman, S., Seifert, W., Kessler, C., and Domdey, H. 1989. Cloning, characterization and heterologous expression of the *Sma*I restriction-modification system. *Nucleic Acids Res.* **17:** 9783–9796.

Ho, D. K., Wu, J. C., Santi, D. V., and Floss, H. G. 1991. Stereochemical studies of the C-methylation of deoxycytidine catalyzed by *Hha*I methylase and the N-methylation of deoxyadenosine catalyzed by *Eco*RI methylase. *Arch. Biochem. Biophys.* **284:** 264–269.

Jones, P. A., Rideout, M. W. III, Shen, J.-C., Spruck, C. M., and Tsai, Y. C. 1992. Methylation, mutation and cancer. *BioEssays* **14:** 33–36.

Karreman, C. and Waard, A. D. 1990. *Agmenellum quadruplicatum* M.*Aqu*I, a novel modification methylase. *J. Bacteriol.* **172:** 266–272.

Kaszubska, W., Webb, H. K., and Gumport, R. I. 1992. Purification and characterization of the M.*Rsr*I DNA methyltransferase from *Escherichia coli. Gene* **118:** 5–11.

Kleckner, N. 1989. Transposon Tn10. In: *Mobile DNA,* pp. 229–268. (Berg, D. D. and Howe, M. M., Eds.) Washington, D.C.: American Society for Microbiology.

Klimasauskas, S. and Roberts, R. J. 1995. M.*Hha*I binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res.* **23:** 1388–1395.

Klimasauskas, S., Nelson, J. L., and Roberts, R. J. 1991. The sequence specificity domain of cytosine-C5 methylases. *Nucleic Acids Res.* **19:** 6183–6190.

Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. 1994. *Hha*I methyltransferase flips its target base out of the DNA helix. *Cell* **76:** 357–369.

Kossykh, V. G., Schlagaman, S. L., and Hattman, S. 1993. Conserved sequence motif DPPY in region IV of the phage T4 Dam DNA-[N6-adenine]-methyltransferase is important for *S*-adenosyl-L-methionine binding. *Nucleic Acids Res.* **21:** 3563–3566.

Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J., and Wilson, G. G. 1994. The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.* **22:** 1–10.

Labahn, J., Grazin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I., and Saenger, W. 1994. Three-dimensional structure of the adenine-specific DNA methyltransferase M.*Taq*I in complex with the cofactor S-adeno-

sylmethionine. *Proc. Natl. Acad. Sci. U.S.A.* **91:** 10957–10961.

Lacks, S. A. and Greenberg, B. 1977. Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. *J. Mol. Biol.* **114:** 153–168.

Lange, C., Wild, C., and Trautner, T. A. 1996. Identification of a subdomain within DNA-(cytosine-C5)-methyltransferases responsible for the recognition of the 5′ part of their DNA target. *EMBO J.* **15:** 1443–1450.

Lauster, R., Trautner, T. A., and Noyer-Weidner, M. 1989. Cytosine-specific methyltransferases, a conserved enzyme core with variable target recognizing domains. *J. Mol. Biol.* **206:** 305–312.

Lee, K.-F., Kam, K.-M., and Shaw, P.-C. 1995. Bacterial methyltransferase M.*Eco*HK31I requires two proteins for *in vitro* methylation. *Nucleic Acids Res.* **23:** 103–108.

Lee, K.-F., Liaw, Y.-C., and Shaw, P.-C. 1996. Overproduction, purification and characterization of M.*Eco*HK31I, a bacterial methyltransferase with two polypeptides. *Biochem. J.* **314:** 321–326.

Leonhardt, H., Page, A. W., Weier, H.-U., and Bestor, T. H. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865–873.

Lewis, J. and Bird, A. 1991. DNA methylation and chromatin structure. *FEBS Lett.* **285:** 155–159.

Malone, T., Blumenthal, R. M., and Cheng, X. 1995. Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. *J. Mol. Biol.* **253:** 618–632.

Marinus, M. G. 1984. Methylation of prokaryotic DNA. In: *DNA Methylation,* pp. 81–109. (Razin, A., Cedar, H., and Riggs, A. D., Eds.) New York: Springer-Verlag.

Martienssen, R. A. and Richards, E. J. 1995. DNA methylation in eukaryotes. *Curr. Biol.* **5:** 234–242.

Matsuo, K., Silke, J., Gramatikoff, K., and Schaffner, W. 1994. The CpG-specific methylase *Sss*I has topoisomerase activity in the presence of Mg^{2+}. *Nucleic Acids Res.* **22:** 5354–5359.

Mi, S. and Roberts, R. J. 1992. How M.*Msp*I and M.*Hpa*II decide which base to methylate. *Nucleic Acids Res.* **20:** 4811–4816.

Mi, S. and Roberts, R. J. 1993. The DNA binding affinity of *Hha*I methylase is increased by a single amino acid substitution in the catalytic center. *Nucleic Acids Res.* **21:** 2459–2464.

Mi, S., Alonso, D., and Roberts, R. J. 1995. Functional analysis of Glu-237 mutants of *Hha*I methyltransferase. *Nucleic Acids Res.* **23:** 620–627.

Modrich, P. 1989. Methyl-directed DNA mismatch correction. *J. Biol. Chem.* **264:** 6597–6600.

Mol, C. D., Arvai, A. S., Slupphaugh, G., Kavli, B., Alseth, I., Krokan, H. E., and Tainer, J. A. 1995. Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell* **80:** 869–878.

Moore, M. H., Gulbis, J. M., Dodson, E. J., Demple, B., and Moody, P. C. E. 1994. Crystal structure of a suicidal DNA repair protein: the Ada O6-methylguanine-DNA methyltransferase from *E. coli. EMBO J.* **13:** 1495–1501.

Nagaraja, V., Shepherd, J. C. W., and Bickle, T. A. 1985. A hybrid recognition sequence in a recombinant restriction enzyme and the evolution of DNA sequence specificity. *Nature (London)* **316:** 371–372.

Noyer-Weidner, M. and Trautner, T. A. 1993. Methylation of DNA in prokaryotes. In: *DNA Methylation: Molecular Biology and Biological Significance,* pp. 39–108. (Jost, J. P. and Saluz, H. P., Eds.)

Ogden, G. B., Pratt, M. J., and Schaechter, M. 1988. The replicative origin of the *E. coli* chromosome binds to cell membrane only when hemimethylated. *Cell* **54:** 127–135.

Oka, A., Sugimoto, K., Takanami, M., and Hirota, Y. 1980. Replication origin of *Escherichia coli* K-12 chromosome: the size and structure

of the minimum DNA segment carrying the information for autonomous replication. *Mol. Gen. Genet.* **178:** 9–20.

Osterman, D. G., DePillis, G. D., Wu, J. C., Matsuda, A., and Santi, D. V. 1988. 5-Fluorocytosine in DNA is a mechanism based inhibitor of *Hha*I methylase. *Biochemistry* **27:** 5204–5210.

Pfeifer, G. P., Steigerwald, S. D., Hansen, R. S., Gartler, S. M., and Riggs, A. D. 1990. Polymerase chain reaction-aided genomic sequencing of an X- chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of active state stability. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 8252–8256.

Pieretti, M., Zhang, F., Fu, Y.-H., Warren, S. T., Oostra, B. A., Caskey, C. T., and Nelson, D. L. 1991. Absence of expression of the *FMR-1* gene in fragile X syndrome. *Cell* **66:** 817–822.

Pogolotti, A. L., Ono, A., Subramaniam, R., and Santi, D. V. 1988. On the mechanism of DNA adenine methylase. *J. Biol. Chem.* **263:** 7461–7464.

Posfai, J., Bhagwat, A. S., Posfai, G., and Roberts, R. J. 1989. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res.* **17:** 2421–2435.

Powell, L. M., Dryden, D. T. F., Willcock, D. F., Pain, R. H., and Murray, N. E. 1993. DNA recognition by the *Eco*KI methyltransferase: the influence of DNA methylation and the cofactor *S*-adenosyl-L-methionine. *J. Mol. Biol.* **234:** 60–71.

Rao, D. N., Page, M. G., and Bickle, T. A. 1989. Cloning, overexpression and the catalytic properties of the *Eco*P15I modification methylase from *Escherichia coli. J. Mol. Biol.* **209:** 599–606.

Reich, N. O. and Everett, E. A. 1990. Identification of peptides involved in *S*-adenosylmethionine binding in the *Eco*RI DNA methylase. *J. Biol. Chem.* **265:** 8929–8934.

Reich, N. O. and Mashhoon, N. 1991. Kinetic mechanism of the *Eco*RI DNA methyltransferase. *Biochemistry* **30:** 2933–2939.

Reich, N. O., Olsen, C., Osti, F., and Murphy, J. 1992. *In vitro* specificity of *Eco*RI DNA methyltransferase. *J. Biol. Chem.* **267:** 15802–15807.

Reinisch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. 1995. Crystal structure of *Hae*III methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell* **82:** 143–153.

Roberts, R. J. 1995. On base flipping. *Cell* **82:** 9–12.

Roberts, D., Hoopes, B. C., McClure, W. R., and Kleckner, N. 1985. *IS*10 transposition is regulated by DNA adenine methylation. *Cell* **43:** 117–130.

Schluckebier, G., O'Gara, M., Saenger, W., and Cheng, X. 1995. Universal catalytic domain structure of AdoMet-dependent methyltransferases. *J. Mol. Biol.* **247:** 16–20.

Schlagaman, S., Hattman, S., May, M. S., and Berger, L. 1976. *In vivo* methylation by *Escherichia coli* K-12 mec$^+$ deoxyribonucleic acid cytosine methylase protects against *in vitro* cleavage by RII restriction endonuclease (R.*Eco*RII). *J. Bacteriol.* **126:** 990–996.

Schumann, J., Walter, J., Willert, J., Wild, C., Koch, D., and Trautner, T. A. 1996. M.*Bss*HII, a multispecific cytosine-C5-DNA-methyltransferase with unusual target recognizing properties. *J. Mol. Biol.* **257:** 949–959.

Selker, E. V. 1990. Premeiotic instability and repeated sequences in *Neurospora crasa. Annu. Rev. Genet.* **24:** 579–613.

Shen, J.-C., Rideout, M. W., and Jones, P. A. 1992. High frequency mutagenesis by a DNA methyltransferase. *Cell* **71:** 1073–1080.

Shen, J.-C., Zingg, J.-M., Yang, A. S., Schmutte, C., and Jones, P. A. 1995. A mutant *Hpa*II methyltransferase functions as a mutator en-

zyme. *Nucleic Acids Res.* **23:** 4275–4282.

Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K., and Kleckner, N. 1995. *E. coli* seqA protein binds *ori*C in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequesteration. *Cell* **82:** 927–936.

Smith, S. S. 1994. Biological implications of the mechanism of action of human DNA (cytosine-5) methyltransferase. *Progr. Nucleic Acids Res. Mol. Biol.* **49:** 65–111.

Smith, S. S., Kaplan, B. E., Sowers, L. C., and Newman, E. M. 1992. Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation. *Proc. Natl. Acad. Sci. U.S.A.* **89:** 4744–4748.

Sohail, A., Lieb, M., Dar, M., and Bhagwat, A. S. 1990. A gene required for very short patch repair in *Escherichia coli* is adjacent to the DNA cytosine methylase gene. *J. Bacteriol.* **172:** 4214–4221.

Sternberg, N. and Coulby, J. 1990. Cleavage of the bacteriophage P1 packaging site (*pac*) is regulated by adenine methylation. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 8070–8074.

Sugisaki, H., Kita, K., and Takanami, M. 1989. The *Fok*I restriction modification system. Presence of two domains in *Fok*I methylase responsible for modification of different DNA strands. *J. Biol. Chem.* **264:** 5757–5761.

Surby, M. A. and Reich, N. O. 1996a. Facilitated diffusion of the *Eco*RI DNA methyltransferase is described by a novel mechanism. *Biochemistry* **35:** 2209–2217.

Surby, M. A. and Reich, N. O. 1996b. Contribution of facilitated diffusion and processive catalysis to enzyme efficiency: implications for the *Eco*RI restriction-modification system. *Biochemistry* **35:** 2201–2209.

Szczelkun, M. D. and Connolly, B. A. 1995. Sequence-specific binding of DNA by the *Eco*RV restriction and modification enzymes with nucleic acid and cofactor analogues. *Biochemistry* **34:** 10724–10733.

Szczelkun, M. D., Jones, H., and Connolly, B. A. 1995. Probing the protein-DNA interface of the *Eco*RV modification methyltransferases bound to its recognition sequence, GATATC. *Biochemistry* **34:** 10734–10743.

Szilak, L., Der, A., Deak, F., and Venetianer, P. 1993. Kinetic characterization of the *Eca*I methyltransferase. *Eur. J. Biochem.* **218:** 727–733.

Taylor, I., Patel, J., Firman, K., and Kneale, G. 1992. Purification and biochemical characterization of the *Eco*R124I type I modification methylase. *Nucleic Acids Res.* **20:** 179–186.

Taylor, I. A., Webb, M., and Kneale, G. G. 1996. Surface labelling of the type I methyltransferase M.*Eco*R124I reveals lysine residues critical for DNA binding. *J. Mol. Biol.* **258:** 62–73.

Trautner, T. A., Pawlek, B., Behrens, B., and Willert, J. 1996. Exact size and organization of DNA target recognizing domains of multispecific DNA-(cytosine-C5)-methyltransferases. *EMBO J.* **15:** 1434–1442.

Vassylyev, D. G., Kashiwagi, T., Mikami, Y., Ariyoshi, M., Iwai, S., Ohtsuka, E., and Morikawa, K. 1995. Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: Structural basis for damaged DNA recognition. *Cell* **83:** 773–782.

Wenzel, C. and Guschlbauer, W. 1993. Dam methyltransferase from *Escherichia coli:* sequence of a peptide segment involved in S-adenosyl-methionine binding. *Nucleic Acids Res.* **21:** 4604–4609.

Wilke, K., Rauhut, E., Noyer-Weidner, M., Lauster, R., Pawlek, B., Behrens, B., and Trautner, T. A. 1988. Sequential order of target-recognizing domains in multispecific DNA methyltransferases. *EMBO J.* **7:** 2601–2609.

Willcock, D. F., Dryden, D. T. F., and Murray, N. E. 1994. A mutational analysis of the two motifs common to adenine methyltransferases. *EMBO J.* **13:** 3902–3908.

Wilson, G. G. 1992. Amino acid sequence arrangements of DNA methyltrans ferases. *Methods Enzymol.* **216:** 259–279.

Wilson, G. G. and Murray, N. E. 1991. Restriction and modification systems. *Annu. Rev. Genet.* **25:** 585–627.

Wu, J. C. and Santi, D. V. 1985. On the mechanism and inhibition of DNA cytosine methyltransferases. In: *Biochemistry and Biology of DNA Methylation,* pp. 119–129. (Cantoni, G. I. and Razin, A., Eds.) New York: Alan R. Liss.

Wyszynski, M. W., Gabbara, S., Kurareva, E. A., Romanova, E. A., Oretskaya, T. S., Gromova, J. S., Shabarova, Z. A., and Bhagwat, A. S. 1993. The cysteine conserved among DNA cytosine methylases is required for methyl transfer, but not for specific DNA binding. *Nucleic Acids Res.* **21:** 295–301.

Wyszynski, M. W., Gabbara, S., and Bhagwat, A. S. 1994. Cytosine deaminations catalyzed by DNA cytosine methyltransferases are unlikely to be the major cause of mutational hot spots at sites of cytosine methylation in *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A.* **91:** 1574–1578.

Yang, A. S., Shen, J.-C., Zingg, J.-M., Mi, S., and Jones, P. A. 1995. *Hha*I and *Hpa*II DNA methyltransferases bind DNA mismatches, methylate uracil and block DNA repair. *Nucleic Acids Res.* **23:** 1380–1387.

Zhang, B., Tao, T., Wilson, G. G., and Blumenthal, R. M. 1993. The M.*Alu*I DNA-(cytosine-C5)-methyltransferase has an unusually large, partially dispensable, variable region. *Nucleic Acids Res.* **21:** 905–911.

NOTE IN PROOF

Temporal control of DNA methylation state in Caulocacter crescentus plays an important role in the development of this organism. Recently, it has been shown that CcrM adenine DNA MTase, which specifically modifies GANTC sequences, is necessary for viability in this organism (Stephens et al., 1996). In studies performed on the small plant *Arabidopsis thaliana,* it has been clearly demonstrated that DNA cytosine methylation is important in establishing or maintaining epigenetic developmental states in the meristem (Ronemus et al., 1996).

More recently, O'Gara et al. (1996) have reported four structures of complexes involving MhaI, three ternary complexes and one binary complex that represent individual steps along the reaction pathway. In the ternary complexes, the target cytosine flips out from the DNA helix and into the target nucleotide binding pocket, close to the cofactor.

References

O'Gara, M., et al. (1996) *J. Mol. Biol.* **261:** 634–645.
Ronemus, M. J., et al. (1996) *Science* **273:** 654–657.
Stephens, C., et al. (1996) Proc Natl. Acad Sci. U.S.A. **93:** 1210–1214.