Critical Reviews in Biochemistry and Molecular Biology, 2010; 45(2): 125–145

informa
healthcare

REVIEW ARTICLE

Diversity of DNA methyltransferases that recognize asymmetric target sequences

Urulangodi Kunhiraman Madhusoodanan, and Desirazu N Rao

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

Abstract

DNA methyltransferases (MTases) are a group of enzymes that catalyze the methyl group transfer from S-adenosyl-L-methionine in a sequence-specific manner. Orthodox Type II DNA MTases usually recognize palindromic DNA sequences and add a methyl group to the target base (either adenine or cytosine) on both strands. However, there are a number of MTases that recognize asymmetric target sequences and differ in their subunit organization. In a bacterial cell, after each round of replication, the substrate for any MTase is hemimethylated DNA, and it therefore needs only a single methylation event to restore the fully methylated state. This is in consistent with the fact that most of the DNA MTases studied exist as monomers in solution. Multiple lines of evidence suggest that some DNA MTases function as dimers. Further, functional analysis of many restriction-modification systems showed the presence of more than one or fused MTase genes. It was proposed that presence of two MTases responsible for the recognition and methylation of asymmetric sequences would protect the nascent strands generated during DNA replication from cognate restriction endonuclease. In this review, MTases recognizing asymmetric sequences have been grouped into different subgroups based on their unique properties. Detailed characterization of these unusual MTases would help in better understanding of their specific biological roles and mechanisms of action. The rapid progress made by the genome sequencing of bacteria and archaea may accelerate the identification and study of species- and strain-specific MTases of host-adapted bacteria and their roles in pathogenic mechanisms.

Keywords: *Asymmetric MTases; base flipping; DNA MTases; DNA–protein interactions; oligomeric MTases; R-M system; sinefungin; S-adenosyl-L- methionine*

Introduction

Restriction endonucleases (ENases) occur ubiquitously among bacteria, archaea (Bickle and Krüger, 1993; Raleigh and Brooks, 1998; Roberts *et al.*, 2003; Sistla and Rao, 2004; Pingoud *et al.*, 2005; Roberts *et al.*, 2007), and in viruses of certain unicellular algae (Van Etten, 2003), and they are usually accompanied by a modification enzyme of identical specificity. The two activities together form a restriction-modification (R-M) system, the prokaryotic equivalent of an immune system. Their main function is to protect the host against invading foreign DNA. The incoming DNA, that is not modified, would be cleaved by the ENase. The host DNA is resistant to cleavage as the sites are modified by a cognate methyltransferase (MTase). It was proposed that the R-M systems have functions other than host protection, including maintenance of species identity among bacteria (Jeltsch, 2002; 2003), generation of genetic variation (Arber, 2000; 2002), recombination and transposition (Heitman, 1993; McKane and Milkman, 1995; Carlson and Kosturko, 1998). Many R-M systems have been considered as selfish DNA elements (Naito *et al.*, 1995; Kobayashi, 2004).

Post-replicative base methylation is the most common DNA modification in bacteria. C^5-methylcytosine and N^6-methyladenine are found in the genomes of many fungi, bacteria and protists, whereas N^4-methylcytosine is found only in bacteria (Cheng, 1995). In addition, N^6-methyladenine is present in archaeal DNA (Barbeyron *et al.*, 1984; Cheng, 1995). Two classes of DNA MTase perform base modifications in bacterial genomes: those associated with R-M systems (Bickle and Kruger, 1993),

Address for Correspondence: Desirazu N. Rao, Department of Biochemistry, Indian Institute of Science, Bangalore 560 012 India. Tel: +91-80-22932538. Fax: +91-80-23600814. E-mail: dnrao@biochem.iisc.ernet.in

(Received 11 November 2009; revised 11 January 2010; accepted 15 January 2010)

ISSN 1040-9238 print/ISSN 1549-7798 online
DOI: 10.3109/10409231003628007

and solitary MTases that do not have a restriction enzyme counterpart. Solitary MTases include the N^6-adenine MTases Dam and CcrM (cell cycle-regulated MTase), and the C^5-cytosine MTase Dcm (Low *et al.*, 2001; Løbner-Olesen *et al.*, 2005; Wion and Casadesús, 2006; Marinus and Casadesus, 2009).

Based on the chemistry of the methylation reaction catalyzed, DNA MTases have been classified as C^5-enzymes (endocyclic MTases), which transfer the methyl group of *S*-adenosyl-L-methionine (AdoMet) to the C^5 position of cytosine (C^5-methylcytosine), and N^6- and N^4- enzymes (exocyclic amino MTases), which transfer the methyl group to the exocyclic amino group of adenine (N^6-methyladenine) or cytosine (N^4-methylcytosine), respectively (Bheemanaik *et al.*, 2006a). DNA MTases of all three types contain conserved regions, which are responsible for catalysis and AdoMet binding, and variable regions known as target recognition domain (TRD), which determine the substrate specificity of a particular enzyme. Ten conserved amino acid motifs (I–X) are found in C^5 MTases, the motif order being almost constant (Posfai *et al.*, 1989; Kumar *et al.*, 1994). Exocyclic DNA MTases are subdivided further into six groups (namely α, β, γ, ζ, δ and ε), according to the linear arrangement of three conserved motifs (Figure 1), the AdoMet-binding domain (FXGXG), the TRD and the catalytic domain (D/N/S)PP(Y/F) motifs (Malone *et al.*, 1995; Bheemanaik *et al.*, 2006a). A structure-guided sequence comparison of exocyclic MTases, N^6-adenine and N^4-cytosine MTases, revealed that they are closely related to one another. Both N^6-adenine and N^4-cytosine MTases contain nine conserved motifs, motifs I-VIII and X (Malone *et al.*, 1995; Bujnicki and Radlinska, 1999).

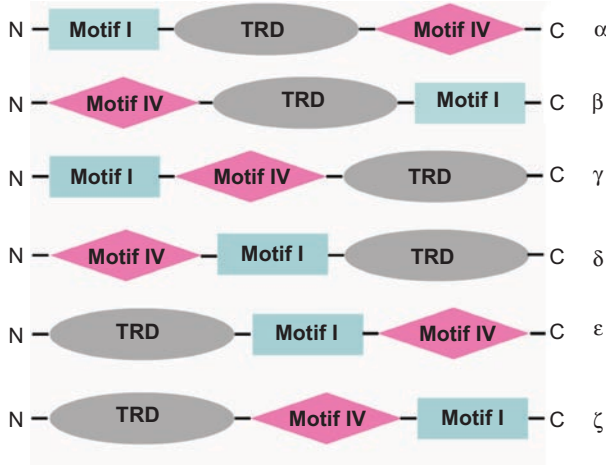

Figure 1. Schematic diagram representing the arrangement of conserved motifs in the primary structure of DNA MTases. The linear arrangements of three conserved motifs, AdoMet-binding motif (motif I), the TRD (target recognition domain) and the catalytic motif (motif IV) are shown.

Base flipping was first reported when the co-crystal structure of M.HhaI, a C^5-cytosine MTase bound to DNA in the presence of *S*-adenosyl-L-homocysteine (AdoHcy), was determined (Klimasauskas *et al.*, 1994). The crystal structure of the M.HhaI-DNA-AdoHcy complex revealed that the target base had been rotated 180° out of the DNA helix via the minor groove into a catalytic domain (motif IV). Base flipping has been proposed as a general mechanism used by MTases for the AdoMet-dependent methylation of DNA (Roberts and Cheng, 1998). Various methods have been used to demonstrate the flipping of the target base out of the DNA duplex, X-ray analysis being the most reliable. Substitution of the target base with its fluorescent analog and detection of a change in fluorescence upon substrate interaction with an enzyme has been used to demonstrate the base flipping mechanism indirectly (Allan and Reich, 1996; Holz *et al.*, 1998; Reddy and Rao, 2000). 2-Aminopurine is widely used as a fluorescent analog to substitute for the target base, be it adenine or cytosine (Vilkaitis *et al.*, 2000). In addition, chemical probing has been used to demonstrate the flipping of the target cytosine by M.HhaI (Serva *et al.*, 1998; Neely *et al.*, 2005).

The X-ray analysis of the ternary complex of the N^6-adenine MTase, M.TaqI-DNA-AdoHcy revealed the flipping of the target adenine out of the DNA double helix (Goedecke *et al.*, 2001). Target base flipping was postulated for other N^6-adenine MTases and supported by biochemical and biophysical evidences and by computer modeling with the X-ray data (Reddy and Rao, 2000; Goedecke *et al.*, 2001; Jeltsch, 2002; Gromova and Khoroshaev, 2003). Both 2-aminopurine and chemical probing were used to demonstrate the base flipping mechanism in N^6-adenine MTase EcoP15I (Reddy and Rao, 2000). It was also observed that base flipping is utilized by numerous other DNA-enzymes (uracil-DNA glycosylase, T4 endonuclease V) (Roberts and Cheng, 1998; Cheng and Roberts, 2001), suggesting its great biological significance.

Monomeric versus dimeric DNA MTases

Orthodox Type II DNA MTases typically recognize palindromic DNA sequences of 4–6 bp, and add a methyl group to a target base on both strands. In a bacterial cell, after each round of replication, the substrate for any MTase is hemimethylated DNA and it, therefore, needs only a single methylation event to restore its fully methylated state. This is in agreement with the fact that most of the DNA MTases studied exist as monomers in solution. These findings established the paradigm for the oligomeric state of the Type II R-M system as dimeric ENases and monomeric MTases (Modrich, 1982; Reich and Mashhoon, 1993). However, it was found that M.BamHI occurs as a

dimer in solution, but dissociates into monomers upon the addition of specific DNA duplex (Malygin *et al.*, 2001). DNA MTases, T4Dam and CcrM, which are not a part of an R-M system, are known to occur as dimers in solution (Shier *et al.*, 2001; Malygin *et al.*, 2004). However, surface plasmon resonance experiments in the presence of AdoHcy showed that CcrM binds as a monomer to DNA. On the other hand, it was shown that EcoDam dimerizes in the presence of a 20-mer oligonucleotide duplex containing the GATC recognition site (Tuzikov *et al.*, 1986; Rechkunova *et al.*, 1987). M.DpnII (de la Campa *et al.*, 1987), M.HaeIV (Piekarowicz *et al.*, 1999), M.LlaCI (Mruk *et al.*, 2003), and M.MboIIA (Osipiuk *et al.*, 2003) have been shown to exist as dimers in solution. In the absence of DNA substrate, M.HhaI (Cheng *et al.*, 1993; Dong *et al.*, 2004), M.TaqI (Labahn *et al.*, 1994), M.PvuII (Gong *et al.*, 1997), and M.MboIIA (Osipiuk *et al.*, 2003) were crystallized as protein dimers. However, further biochemical and structural studies are required to confirm the relevance of this dimerization. The extent of subunit contacts observed in the crystal structure of M.MboIIA suggested that the two molecules in the asymmetric unit represent a biologically relevant dimer. In M.AhdI, two copies of the AhdIS subunit combine with two AhdIM subunits to form a tetrameric MTase (M_2S_2). The tetrameric M.AhdI dissociates into heterodimers (M_1S_1) at low protein concentrations. However, only the tetrameric form binds to DNA with high affinity (Marks *et al.*, 2003). M.KpnI has been shown to exist as a dimer in solution (Bheemanaik *et al*, 2003). Mutation at the subunit interface (I146) resulted in an inactive, monomeric M.KpnI with reduced DNA and cofactor binding (Bheemanaik, *et al.*, 2006b). M.RsrI (Kaszubska *et al.*, 1992) and M.MspI (Dubey and Roberts, 1992) have been shown to dimerize at high protein concentrations. However, M.MspI is functional as a monomer, but forms a dimer and higher aggregates as the enzyme concentration increases. Site-directed mutagenesis (S124D) of M.RsrI resulted in a monomeric protein that showed reduced methylation activity relative to that of the wild type (Thomas and Gumport, 2006). Therefore, it was proposed that the dimer observed in the crystal structure of M.RsrI may be the functional unit of M.RsrI and may bind to two DNA molecules simultaneously, one along each groove formed by the monomer–monomer interface (Thomas and Gumport, 2006). In contrast, M.MboIIA, which recognizes an asymmetric sequence and methylates only one DNA strand, functions as a dimer. Thus, it appears that there is no correlation between target site (a)symmetry and MTase dimerization (Malygin *et al.*, 2009).

A putative adenine specific DNA MTase TTHA0409 from *Thermus thermophilus* HB8 was crystallized as a dimer (Morita *et al.*, 2008). The dimer interface area was calculated to be 2400 Å^2. This is larger than those in M.MboIIA (1900 Å^2) and M.RsrI (1500 Å^2). The presence of such an extensive dimer interface strengthens the possibility that M.TTHA0409 functions as a dimer. Gel filtration analysis yielded an apparent molecular weight of 82 kDa, suggesting that M.TTHA0409 exists in a dimeric form in the absence of DNA. Although its structure has been determined, the recognition sequence and methylation activity of M.TTHA0409 is yet to be determined. *E. coli* 23S rRNA C^5-DNA MTase, RlmI was found to exist as a dimer in solution (Sunita *et al.*, 2008). Interestingly, dimerization has also been shown for vertebrate MTases Dnmt3a and Dnmt3L (Jia *et al.*, 2007) and human placental DNA-C^5-cytosine MTase (Yoo *et al.*, 1987). The disruption of the Dnmt3a dimer by specific point mutations (R881 and D872) resulted in loss of catalytic activity (Jia *et al.*, 2007). Recently, Dnmt1 was shown to exist as a stable dimer in solution (Fellinger *et al.*, 2009). The increasing number of dimeric MTases raises the question as to whether dimerization is important for catalysis. However, further studies are required to assess the possible biological relevance of the dimerization of DNA MTases. There are an increasing number of DNA MTases being characterized that recognize asymmetric sequences with unique target site specificities and structural organization (Figure 2). In this review, DNA MTases have been grouped into different subgroups based on their organization and unique properties.

M1-M2 DNA MTases: Two independent MTases, each modify bases on the opposite DNA strands of an asymmetric recognition sequence

Typically, all MTases of this group recognize short (4–6 bp) asymmetric sequences and consist of two MTases (Table 1). Each MTase recognizes and methylates bases on the opposite strands of the same recognition sequence, making the modified DNA resistant to the restriction enzyme cleavage. The modification can be C^5-methylcytosine, N^4-methylcytosine, or N^6-methyladenine.

MboII, NcuI, BcnI, MnlI, BcoKI, BfiI, HgaI, Eco31I, BspLU11III and BstF5I DNA MTases

All these MTases recognize asymmetric sequences and belong to the Type IIS R-M system. The MboII target sequence (5′-GAAGA-3′/3′-CTTCT-5′) displays a peculiar kind of asymmetry in that one strand is composed of only purines and the other strand contains only pyrimidine nucleotides. Initially, McClelland *et al.* (1985) reported that the M.MboII methylates the 3′ adenine, producing 5′-GAAG^{m6}A-3′ (McClelland *et al.*, 1985; Bocklage *et al.*, 1991). After replication, however, one daughter DNA molecule is unmethylated, and

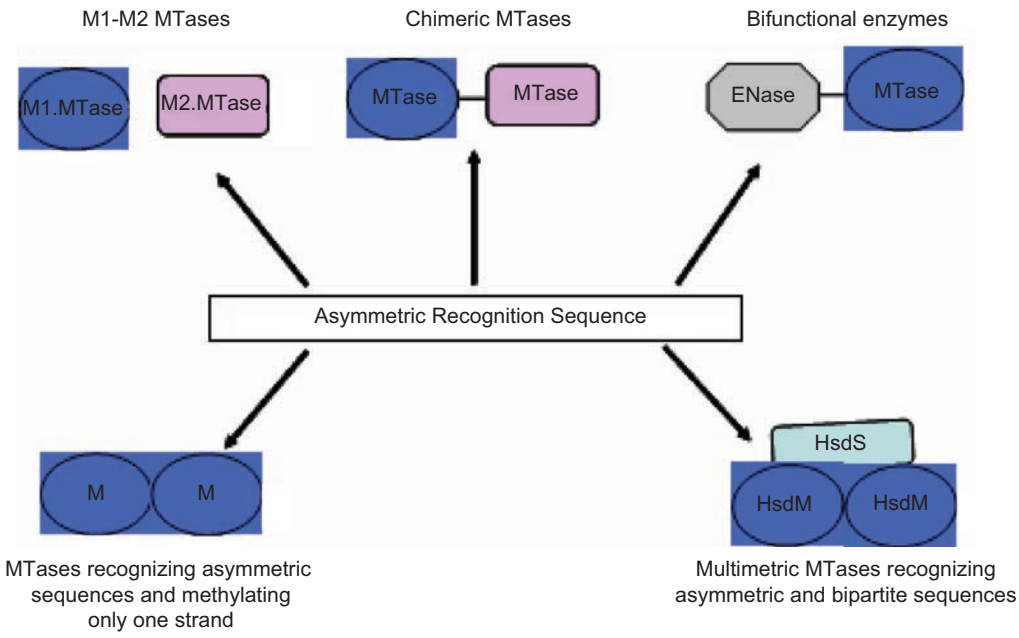

Figure 2. Schematic diagram representing the organization of DNA MTases recognizing asymmetric target sequence.

Table 1. M1-M2 DNA MTases: one MTase for each strand in the asymmetric recognition sequence.

Enzyme	Type of methylation	Recognition sequence and base modified
M1.MboII	N^6-methyladenosine	5′-GAAGA^{m6}-3′
M2.MboII	N^4-methylcytosine	3′-CTTC^{m4}T-5′
M1.MnlI	C^5-methylcytosine	5′-CC^{m5}TC-3′
M2.MnlI	N^6-methyladenosine	3′-GGA^{m6}G-5′
M1.BcoKI	N^4-methylcytosine	5′-C^{m4}TCTTC-3′
M2.BcoKI	N^6-methyladenosine	3′-GA^{m6}GAAG-5′
M1.BfiI	N^4-methylcytosine	5′-TGACC^{m4}C-3′
M2.BfiI	N^4-methylcytosine	3′-AC^{m4}TGGG-5′
M1.HgaI	C^5-methylcytosine	5′-CTGC^{m5}G-3′
M2.HgaI	C^5-methylcytosine	3′-GAC^{m5}GC-5′
M1.BspLU11III	N^6-methyladenosine	5′-GGGA^{m6}C-3′
M2.BspLU11III	C^5-methylcytosine	3′-CCC^{m5}TG-5′
M1.Eco31I	C^5-methylcytosine	5′-GGTC^{m5}TC-3′
M2.Eco31I	N^6-methyladenosine	3′-CCA^{m6}GAC-5′
M1.BstF5I	N^6-methyladenosine	5′-GGA^{m6}TG-3′
		3′-CCTA^{m6}C-5′
M2.BstF5I	N^6-methyladenosine	5′-CA^{m6}TCC-3′
		3′-GTA^{m6}GG-5′
M3.BstF5I	N^6-methyladenosine	5′-GGA^{m6}TG-3′
		3′-CCTA^{m6}C-5′
M4.BstF5I	N^6-methyladenosine	5′-CA^{m6}TCC-3′
		3′-GTA^{m6}GG-5′

therefore sensitive to restriction. This is in contrast to MTases with palindromic targets, where both daughter molecules are hemimethylated, thus providing sufficient protection against the cognate ENase. Since there is no adenine residue in the lower strand of the recognition sequence, it was proposed that the cytosine in the bottom strand might be modified by an alternative activity of the M.MboII or a completely different MTase. Indeed, a second methylase, M2.MboII, was cloned, and it was demonstrated that it modifies the internal cytosine in the bottom strand, yielding N^4-methylcytosine (Figure 3A) (Furmanek-Blaszk *et al.*, 2009). Further, functional analysis showed that the complete MboII R-M system consists of two MTases genes and the R.MboII gene. M1.MboII alone is sufficient to support the growth of *E.coli* cells carrying the R.MboII gene. However, this could not be demonstrated for M2.MboII. Taken together with their dissimilar primary structures, this disparity indirectly suggests that the two MboII MTases may have distinct functional roles. The presence of two MTases responsible for the recognition and methylation of strand-specific sequences would be advantageous for bacterial cells, because nascent strands generated during DNA replication could be simultaneously protected from attack by site-specific ENases. It was found that Mn^{2+} and Zn^{2+} acted as strong inhibitors of the enzyme (IC_{50} – 0.4 mM), while Ca^{2+} and Mg^{2+} inhibited the enzymes less efficiently (IC_{50} – 4 mM and 5 mM, respectively) (Furmanek-Blaszk *et al.*, 2009). The two separately existing MTases differ in their properties. A comparison of the sensitivity of M1.MboII and M2.MboII to Mg^{2+} showed that the latter is more tolerant to Mg^{2+}. Most interestingly, M2.MboII proved to be equally efficient on the single-stranded or double-stranded DNA. The specific activity of M2.MboII for double- and single-stranded DNA is almost the same.

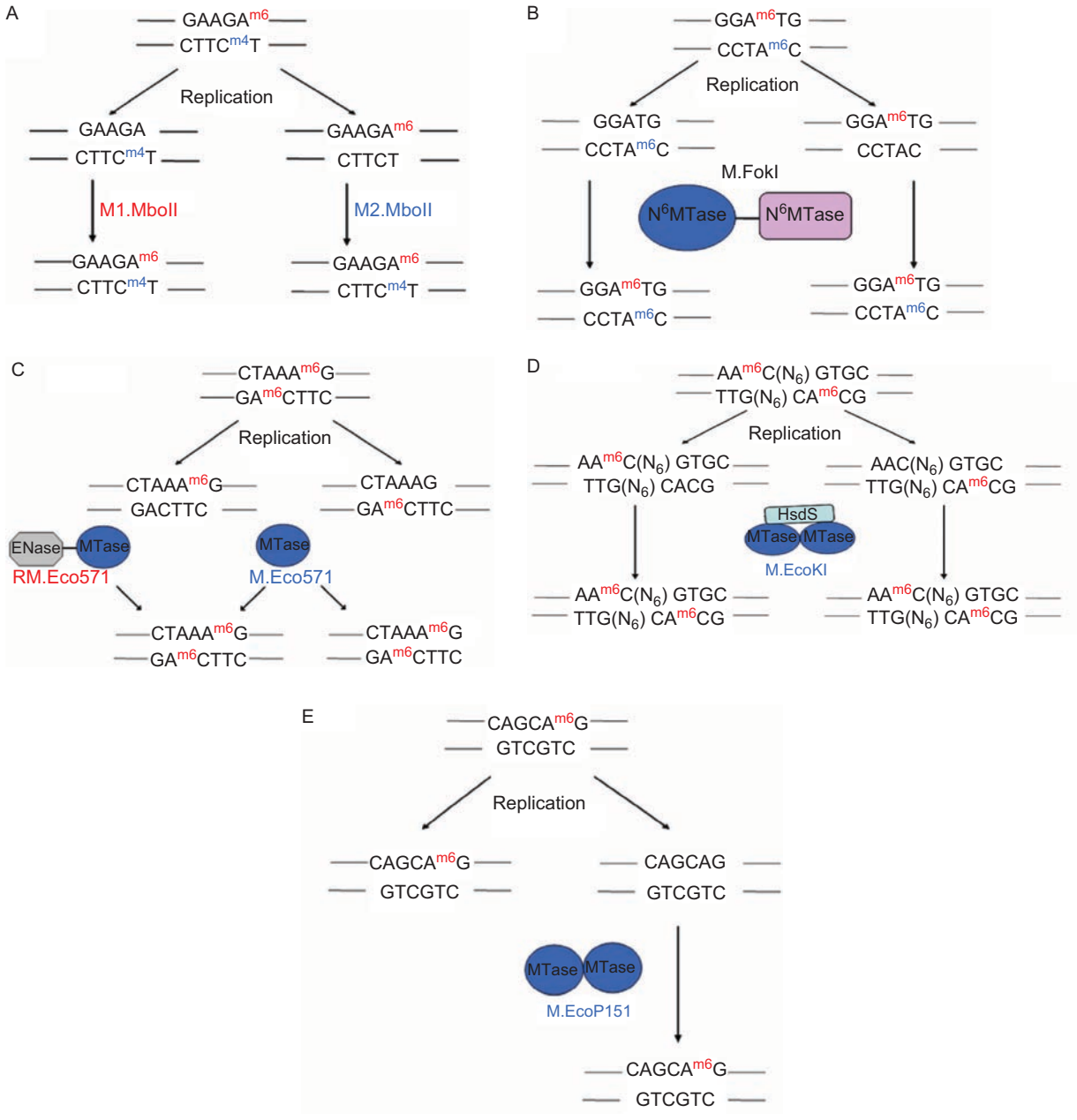

Figure 3. Schematic diagram representing the post-replicative base modifications by MTases recognizing asymmetric sequence. (A) MboII MTases (M1-M2 MTase), (B) FokI MTase (chimeric MTase), (C) Eco571 MTase (bifunctional enzyme), (D) EcoKI MTase (belonging to the Type I R-M system) and (E) EcoP15I MTase (belonging to the Type III R-M system).

In contrast, no detectable activity on the single-stranded substrate was observed with M1.MboII.

M1.MboII showed a significant level of identity with M1.NcuI (90%) and M1.HpyAII (55%) (Furmanek *et al.*, 2007). The M2.MboII protein showed 88% and 62% identities with the M2.NcuI and M2.HpyAII of *N. cuniculi* and *H. pylori*, respectively (Furmanek-Blaszk *et al.*, 2009). The isomeric NcuI system is also composed of two (N^6-methyladenine and N^4-methylcytosine) MTases and a single ENase. Both methyltransferases of the NcuI R-M system showed similar properties and sequence similarity (Furmanek-Blaszk *et al.*, 2009), indicating that these proteins have emerged from a common ancestor.

Similar results were obtained with the BcnI system. However, both BcnI MTases are isospecific on double-stranded DNA, modifying the second cytosine residue in the opposite strands of the recognition sequence 5′-CC(C/G)GG-3′ (Vilkaitis *et al.*, 2002). However, M1.BcnI can also methylate the target sites in single-stranded DNA. It was shown that M2.BcnI alone is sufficient to support the growth of *E. coli* cells carrying the *R.BcnI* gene. This could not be demonstrated for M1.BcnI. It was

proposed that single-strand specific DNA MTases may be required to enhance plasmid transmission to the cells during transformation, which are known to proceed via a single-stranded intermediate (Bickle and Kruger, 1993; Merkiene *et al.*, 1998). During plasmid transformation, methylation of the incoming strand by M1.BcnI would protect the subsequently reconstituted plasmid from R.BcnI cleavage (Vilkaits *et al.*, 2002).

M.MnlI is similar to the MboII R-M system with respect to the asymmetric DNA recognition site and the base modified. M1.MnlI recognizes 5′-CCTC-3′ and methylates the second cytosine, yielding C^5-methylcytosine. M2.MnlI modifies the unique adenine residue in the bottom strand of the target sequence, yielding N^6-methyladenosine (Kriukiene *et al.*, 2005).

The BcoKI R-M system consists of one restriction enzyme and two different MTases, one for each strand of its asymmetrical recognition site. M1.BcoKI recognizes asymmetric sequence 5′-CTCTTC-3′, and methylates the first cytosine residue, resulting in N^4-methylcytosine. M2.BcoKI recognizes the same sequence, but modifies the first adenine residue in the bottom strand producing N^6-methyladenosine (Svadbina *et al.*, 2005).

The BfiI R-M system comprises two MTases and a single restriction enzyme. Each MTases modifies cytosines on opposite strands of the recognition sequence, yielding N^4-methylcytosine (Sapranauskas *et al.*, 2000). The R.BfiI shows the unusual property of cleaving DNA in the absence of Mg^{2+} ions.

The HgaI R-M system consists of two cytosine MTase genes responsible for the modification of different strands in the target DNA. M1.HgaI methylates the internal cytosine residue in the bottom strand, 3′-CTGCG-5′, whereas M2.HgaI methylates the internal cytosine reside in the top strand, 5′-GACGC-3′ (Sugisaki *et al.*, 1991). The sequences of the two HgaI MTases have been well conserved in overall regions, suggesting that the two MTases have evolved from a common ancestor by gene duplication (Sugisaki *et al.*, 1991). The two HgaI MTases present a unique opportunity to understand the structure–function relationships of the C^5-methylcytosines for two main reasons. First, the genes encoding the two MTases are very similar throughout the entire sequence. However, a comparison of the two MTase domains of another Type IIS enzyme, M.FokI, showed no sequence similarity except for tetra-amino-acid sequences common to adenine-specific MTases (discussed below). Because both enzymes are hemimethyltransferases, kinetic studies of their reaction mechanisms are free from the complications arising in the analysis of the reaction catalyzed by palindromic DNA MTases, where the product of the first methyl-transfer step becomes a substrate for the second methyl-transfer event (Baldwin *et al.*, 1994).

Eco31I MTase recognizes hexanucleotide 5′-GGTCTC/3′-CCAGAG-5′. Two separate genes encode N^6-adenine MTase and C^5-cytosine MTase in the Eco31I R-M system. This R-M system shares a common pentanucleotide 5′-GTCTC-3′ recognition sequence with two other Type IIS MTases, Alw26I (5′-GTCTC-3′), and Esp3I (5′-CGTCTC-3′) (Butkus *et al.*, 1985; Bitinaite *et al.*, 1991). However, Alw26I and Esp3I belong to the chimeric MTase family because both M.Alw26I and M.Esp3I are composed of an N^6-adenine MTase domain fused to a C^5-cytosine MTase domain (discussed below). It was shown that the DNA modified by M.Alw26I is resistant to cleavage by R.Eco31I or R.Esp3I (Bitinaite *et al.*, 1992), which supports the idea that they are evolutionarily related.

The BspLU11III R-M system consists of two MTases and one ENase. An interesting feature of this R-M system is that the R.BspLU11III has both ENase and MTase activity in a single polypeptide chain (Chernov *et al.*, 1996). Unlike other Type II ENases, R.BspLU11III posesses N^6-adenine MTase activity, resulting in the methylation of adenine residue in the upper strand of the recognition sequence. Hence, it is a member of the Type IIG system. Type IIG enzymes have both R and M domains fused to form single polypeptides and these are stimulated or inhibited by AdoMet. Therefore, the BspLU11III R-M system is similar to bifunctional MTases (see below). DNA cleavage by bifunctional enzymes is positively affected by low concentrations of AdoMet but ATP has no effect on DNA cleavage (Jurenaite-Urbanaviciene *et al.*, 2001). For all Type IIG enzymes it has been shown that the ENase associated MTase activity is N^6-adenine specific.

Methylation by R.BspLU11III of the only A in the upper strand of the recognition sequence 5′-GGGAC-3′ renders the site resistant to the restriction activity of this enzyme. However, the semiconservative mode of DNA replication would expose one of the unmethylated daughter strands to restriction. As expected, another cytosine-specific MTase was isolated, which recognizes the same sequence but methylates the third cytosine in the bottom strand (M2.BspLU11III). It was also shown that the BspLU11III R-M operon codes for two MTases. M1.BspLU11III was found to be an N^6-adenine specific MTase, which methylates the adenine residue in the top strand, and is similar to the MTase activity present in the ENase protein. The upstream gene encodes the N^6-adenine MTase (ORF2) and overlaps the downstream C^5-cytosine MTase (ORF3) gene by 1 bp. ORF 4 encodes the R.BspLU11III, which has both ENase and N^6-adenine MTase activity. All the three genes have potential promoter regions, indicating that they are not organized in a single operon. Based on this, it was assumed that this particular R-M system might have evolved by convergent evolution. The presence of a transposase gene (ORF1) indicates that the whole locus containing the BspLU11III R-M system is part of a transposon (Lepikhov *et al.*, 2001).

Thus, each strand within the asymmetric recognition sequence is methylated by a different enzyme. However,

the biological significance of the ability of ENase to methylate the same adenine residue within the same recognition sequence as the corresponding MTase is not clear. It was shown that both MTase and ENase methylate the same residue in the recognition sequence for the enzymes Eco57I and BseMII (see below). However, in contrast to Eco57I and BseMII, two MTases are present to modify both strands of the BspLU11III recognition site. It was proposed that the modification of both DNA strands within the BspLU11III recognition site is crucial for *Bacillus sp*. LU11 to maintain the integrity of the chromosomal DNA after replication. Although the BspLU11III system is grouped under Type IIG, it was proposed that this group of the R-M system might be an intermediate between Type IIS and Type III systems.

The BstF5I R-M system from *Bacillus stearothermophilus* F5 includes four site-specific DNA MTases, thus differing from all known R-M systems. This system consists of two pairs of MTases with identical substrate specificities (M1.BstF5I and M3.BstF5I, which recognize 5′-GGATG-3′, and M2.BstF5I and M4.BstF5I, which recognize 5′-CATCC-3′) that modify adenine residues on both DNA strands. There is a single restriction enzyme R.BstF5I that recognizes the same specific sequence, 5′-GGATG-3′/5′-CATCC-3′, as the restriction ENases FokI and StsI, but cleaves DNA at a different position away from the site in position 2/0 (Abdurashitov *et al*., 1996). BstF5I-2 and BstF5I-3 are homologous to the C- and N-terminal domains of the enzymes-isoschizomers FokI and StsI (Degtyarev *et al*., 1997). The presence of two sets of MTases with identical specificities could be related to some peculiarities of functioning of these enzymes in *B. stearothermophilus* F5. These MTases prefer unmodified substrates to the hemimethylated one. However, M2.BstF5I showed several-fold higher enzymatic activity than M4.BstF5I due to the higher affinity of M2.BstF5I to the substrate DNA and AdoMet, and due to its high specificity constant (k_{cat}/K_M). Since M2.BstF5I has a broader pH optimum, a greater catalytic constant, and more affinity to DNA in comparison to M4.BstF5I, it appears that the second MTase would be sufficient for methylation of the unmethylated host DNA in the cell (Chernukhin *et al*., 2003). Therefore, it was suggested that an additional two pairs of MTases would be necessary for the methylation of another substrate type; hemimethylated double-stranded DNA and/or a single-stranded substrate.

The genome analysis of the two sequenced strains of *Helicobacter pylori* (26695 and J99) showed that the bacterium possesses an unusually high number of strain-specific R-M genes (Tomb *et al*., 1997; Alm *et al*., 1999). It was proposed that the presence of numerous R-M systems might be related to its unique adaptability to acidic environments (Xu *et al*., 2000). The genome of the *H. pylori* strain 26695 contains 25 open reading frames encoding putative DNA MTases (Vitkute *et al*., 2001). It was surprising to see the presence of seven M1-M2 putative methylase ORFs in *H. pylori* (Table 2). It will be interesting to analyze their role because all of them recognize the identical asymmetric sequence 5′-CCATC-3′ (REBASE, http://rebase.neb.com). It has been shown that the HpyC1I R-M system is involved in the adherence of *H. pylori* to gastric epithelial cells, a crucial step in bacterial pathogenesis. The putative MTases, M1.HpyC1I and M2.HpyC1I, are located upstream of the gene encoding R.HpyC1I, suggesting that these three ORFs form an operon corresponding to a complete R-M system (Lin *et al*., 2004). A knockout of this ORF showed cell elongation and decreased adherence to gastric epithelial cells, suggesting that this R-M system could be of major importance for the bacteria (Lehours *et al*., 2007).

In general, MTases belonging to Type II R-M systems that recognize a symmetric sequence, contain a single functional domain for methylation, and methylate the target base (either C or A) of both strands in a rotationally symmetric manner (McClelland, 1981; Lauster *et al*., 1989; Klimasauskas *et al*., 1989; Smith, 1996). However, a few R-M systems that recognize palindromic DNA have been found to contain two MTases (M1-M2). The best studied among them is DpnII, which recognizes the sequence 5′-GATC-3′. The two DNA adenine MTases encoded by the DpnII R-M system, M1.DpnII and M2.DpnII, methylate double-stranded substrates, but M1.DpnII can also modify single-stranded DNA (Cerritelli *et al*., 1989). A few other GATC-specific R-M systems having M1-M2 MTases include LlaDCHI, MboI, LlaAI, and SsuDAT1I. However, the specific roles of dual MTases have been studied in detail only in the case of DpnII. The main biological role of M1.DpnI was suggested to be facilitation of the natural transfer of plasmid DNA between strains by the transformation pathway in which the incoming plasmid DNA is in a single-stranded form (Lacks *et al*., 2000).

Table 2. Putative M1-M2 DNA MTases of *H. pylori.*

ORF	Recognition sequence
M1.HpyB34ORFAP	5′-CCATC-3′
M2.HpyB34ORFAP	3′-GGTAG-5′
M1.HpyC1I	5′-CCATC-3′
M2.HpyC1I	3′-GGTAG-5′
M1.HpyC1IMORFAP	5′-CCATC-3′
M2.HpyC1IMORFAP	3′-GGTAG-5′
M1.HpyGORF1422P	5′-CCATC-3′
M2.HpyGORF1422P	3′-GGTAG-5′
M1.HpyHORF1413P	5′-CCATC-3′
M2.HpyHORF1413P	3′-GGTAG-5′
M2.HpyM38ORFAP	5′-CCATC-3′
	3′-GGTAG-5′
M1.Hpy166ORFDP	5′-CCATC-3′
M2.Hpy166ORFDP	3′-GGTAG-5′

The biological meaning of the coexistence of an ENase and two individual MTases, recognizing the same DNA sequence, remains to be elucidated. The requirement for two MTases can be rationalized; first, by the asymmetry of the recognition site, and second by the fact that each chain of such a site contains different types of bases that are to be methylated. Thus, for methylation in both strands, cytosine-specific and adenine-specific MTases are required. In addition, the presence of single-strand specific DNA MTases (M1.BcnI and M2.MboII) may be required to enhance plasmid transmission to the cells during transformation. It has been shown that substrate specificity and the level of activity for two separately existing MTases may be different, even though they modify either the same or different nucleotides of the target DNA (Cerritelli *et al.*, 1989; Merkiene *et al.*, 1998; Kriukiene *et al.*, 2005).

Chimeric DNA MTases

The MTases of this group are represented by proteins of unusual structural and functional organization. All the MTases of this group recognize asymmetric sequences and belong to the Type IIS R-M system. Here two MTases are fused to each other forming one enzyme in tandem. They are two independent domains recognizing the same sequence, which modify the same type (N^6-methyladenine) or different types (N^6-methyladenine and C^5-methylcytosine) of nucleotides in the recognition sequence (Table 3).

FokI DNA MTase

FokI MTase (M.FokI), from *Flavobacterium okeanokoites*, is a fusion protein that consists of two domains, each of which contains all the amino-acid sequence motifs characteristic of N^6-adenine MTases (Landry *et al.*, 1989; Sugisaki *et al.*, 1989; Malone *et al.*, 1995). M.FokI modifies both adenine residues within its asymmetric recognition sequence 5′-GGATG/CATCC-3′. N-terminal (amino acids 1–367) and C-terminal fragments (amino acids 335–647) are catalytically active (Leismann *et al.*, 1998), indicating that both domains are independent of each other. The N-terminal domain methylates adenine residue in the GGATG-strand of the recognition sequence, whereas the C-terminal domain modifies adenine in the CATCC-strand (Sugisaki *et al.*, 1989; Looney *et al.*, 1989; Leismann *et al.*, 1998) (Figure 3B). It was found that the C-terminal domain of M.FokI also methylates the upper strand of the recognition sequence (GGATG), albeit less efficiently (Leismann *et al.*, 1998).

Various studies have shown that the interactions of N- and C-terminal domains of M.FokI with DNA are mutually exclusive, and neither domain influences the other functionally. The N-terminal domain displays a very high specificity for the GGATG site (Lesser *et al.*, 1990; Thielking *et al.*, 1990; Alves *et al.*, 1995). In contrast, the C-terminal domain is not very specific for CATCC, because it also recognizes most sequences differing in 1 bp from the recognition site. Moreover, it accepts substrates differing in two positions from the 5-bp consensus sequence. The N-terminal domain strongly prefers hemimethylated (5′-GGATG/C^{m6}ATCC-3′) over unmethylated substrates. In contrast, the C-terminal domain prefers unmethylated DNA substrates (Leismann *et al.*, 1998). These results clearly show that MTases are a heterogeneous group of enzymes, some being very specific while others not (Friedrich *et al.*, 2000).

M.FokI was shown to be strongly inhibited by divalent cations (Mg^{2+}, Ca^{2+}, Ba^{2+}, Mn^{2+}, and Zn^{2+}). The intracellular concentration of Mg^{2+} ions in *E. coli* cells is estimated to be 4 mM (Lusk *et al.*, 1968), and it is very likely that this value is similar for other bacterial species. This means that in *F. okeanokoites* cells, the activity of M.FokI is inhibited (IC_{50} at 0.4 mM of Mg^{2+}) when the activity of the cognate restriction endonuclease (R.FokI) is stimulated. It is significant that the two M.FokI truncated derivatives show different tolerances to the Mg^{2+} ions. M.FokIC (CTD) is very sensitive to Mg^{2+} (IC_{50} at 0.4 mM), while the IC_{50} for M.FokIN (NTD) is 10-fold higher (4.3 mM) (Kaczorowski *et al.*, 1999). These results confirm the existence of two catalytic centers within a single molecule of M.FokI, and strongly support the idea that M.FokI could be the result of fusion between two enzymes whose genes evolved from independent ancestors. This evolutionary trait has to be different from those that resulted in the M.MboII and M.HgaI, and possibly other members of Type IIS enzymes (McClelland *et al.*, 1985; Sugisaki *et al.*, 1991).

Surprisingly, despite their catalytic efficiency and their DNA sequence specificity, it was shown that M.FokI do not have a pronounced specificity for their target base, because these enzymes also catalyze N^4-methylation of cytosine residues. The rates of cytosine methylation by both domains of M.FokI are reduced by only 1–2 orders of magnitude, demonstrating that these enzymes have a low specificity with respect to the target base (Jeltsch *et al.*, 1999). Studies on M.FokI provide interesting implications on the molecular evolution of Type IIS R-M

Table 3. Chimeric DNA MTases: two MTase activities in a single polypeptide and each methylates one strand.

Enzyme	Recognition sequence and base modified[a]
M.FokI	5′-GGA^{m6}TG-3′(NTD)
	3′-CCTA^{m6}C-5′(CTD)
M.Alw26I	5′-GTC^{m5}TC-3′(CTD)
	3′-CA^{m6}GAG-5′(NTD)
M.Esp3I	5′-CGTC^{m5}TC-3′
	3′-GCA^{m6}GAG-5′

[a]NTD: N-terminal domain; CTD: C-terminal domain.

systems, which is difficult to rationalize, because three enzymes (two MTases and an ENase) of matching specificities are required to establish a functional R-M system that provides an evolutionary advantage. However, it is very unlikely that such a gene cluster could have formed by chance. To overcome this barrier, an MTase would be required that could methylate both strands of the asymmetric recognition sequence of the ENase. It was suggested that the C-terminal domain of M.FokI could have been the precursor MTase, while the N-terminal enzyme could have been added later either by gene transfer or gene duplication. After acquisition of this second enzyme, the original MTase was no longer under evolutionary pressure to recognize GGATG sequences, and hence was free to specialize for CATCC. It was further proposed that M.FokI (C-terminal domain) might be an evolutionary intermediate that has currently changed its specificity from SSATSS (S = G or C) to CATCC.

Alw26I and Esp3I DNA MTases

Both M.Alw26I and M.Esp3I are represented by a single chimeric protein composed of an N^6-adenine MTase domain fused to a C^5-cytosine MTase domain. They methylate both strands of their recognition sequences and show different strand-specificity, i.e. a cytosine residue is methylated in the top strand and an adenine residue in the bottom strand. The nine conserved motifs characteristic of N^6-adenine MTases are localized in the N-terminal half of the polypeptide. The C-terminal part of the polypeptide contains 10 conserved motifs characteristic of C^5-cytosine MTases. Alw26I MTase recognizes a 5-base pair sequence 5′-GTCTC-3′/3′-CAGAG-5′, whereas M.Esp3I recognizes 5′-CGTCTC-3′/3′-GCAGAG-5′ (Bitinaite *et al.*, 1992). Both the MTases modify the internal cytosine in the top strand and the first adenine residue in the bottom strand (Table 3). The Esp3I recognition sequences include a common pentanucleotide, 5′-GTCTC-3′, which is an Alw26I recognition site. Among the known Type IIS MTases, M.Alw26I and M.Esp3I represent unique examples in which the γ N^6-adenine MTase is fused to the C^5-cytosine MTase. It would be interesting to investigate further the structural and functional organization of these unusual chimeric MTases.

Bifunctional R-M systems

These groups of enzymes have both cleavage and modification domains within one polypeptide. They recognize asymmetric bipartite sequences, and cleave DNA either at one side or at both sides of the recognition sequence to produce a ~30–35 bp fragment with an intact recognition sequence (Type IIB). However, a few members of this group recognize an asymmetric sequence and cleave only one side of the recognition sequence. They have been found to contain an additional MTase that methylates both strands (Type IIG). Many of their properties resemble those of Type I R-M systems. The MTases of this group methylate an adenine residue on one or both strands, yielding N^6-methyladenine (Table 4).

BcgI, HaeIV, and AloI R-M systems

All three MTases recognize asymmetric bipartite sequences and belong to the Type IIB R-M system. BcgI is a multi-subunit R-M system that differs from all the other type of R-M systems in its genetic and functional organization. Like the Type I R-M system, it recognizes a bipartite sequence, 5′-GCA(N_6)TCG-3′. The BcgI R-M system consists of two subunits, A and B. It is a bifunctional protein complex which can cleave or methylate DNA (RM.BcgI). Both subunits are required for ENase and MTase activities. BcgI cleaves DNA on both sides of the recognition sequence, generating a 34-bp fragment. Following cleavage, BcgI remains bound to the 34-bp fragment and methylates it at an extremely slow rate (Kong *et al.*, 1993). The regulation of two competing activities of the same complex is determined by DNA substrates and cofactors. BcgI is an active ENase and a poor MTase on unmodified DNA substrates. In contrast, BcgI is an active MTase and an inactive ENase on hemimethylated DNA substrates (Kong and Smith, 1997). As a result, any unmethylated exogenous DNA, such as invading phage DNA, will be preferentially cleaved. When hemimethylated DNA substrates are present, BcgI is an active MTase and an inactive ENase, so that its host endogenous DNA can be methylated efficiently following replication.

The cleavage and methylation reactions share cofactors. While BcgI requires Mg^{2+} and AdoMet for DNA

Table 4. Bifunctional R-M systems: MTase and ENase activities present in a single polypeptide.

Enzyme	Recognition sequence and base modified[a]
RM.AloI	5′-GAA^{m6}C(N_6)TCC-3′
	3′-CTTG(N_6)A^{m6}GG-5′
RM.BcgI	5′-CGA^{m6}(N_6)TGC-3′
	3′-CT(N_6)A^{m6}CG-5′
RM.HaeIV	5′-GA^{m6}Y(N_5)RTC-3′
	3′-CTR(N_5)YA^{m6}G-5′
RM.Eco57I	5′-CTGAA^{m6}G-3′
	3′-GACTTC-5′
M.Eco571	5′-CTGAA^{m6}G-3′
	3′-GA^{m6}CTTC-5′
RM.MmeI	5′-TCCRA^{m6}C-3′
	3′-AGGYTG-5′
RM. TspGWI	5′-ACGGA-3′ (ND)
	3′-TGCCT-5′

[a]R = G or A; Y = C or T; ND = not determined.

cleavage, its methylation reaction requires only AdoMet, and yet is significantly stimulated by Mg^{2+} (Kong *et al.*, 1993). The BcgI A subunit is the R-M subunit, which contains the restriction (R) domain in the amino-terminal half and a methylation (M) domain in the carboxyl-terminal half. Although the A subunit contains the two conserved MTase motifs, it lacks any methylation activity by itself in the absence of the B subunit. The dimerization surface of the R-M subunit is proposed to be located near the M domain in the carboxyl-terminal half. The BcgI B subunit exhibits a repeat structure similar to that of the specificity subunit (HsdS) of the Type I R-M system. The Type IS subunit contains two large variable regions separated by two repeated regions. It has been shown that the variable regions are target recognition domains (TRDs), each recognizing one component of the bipartite recognition sequence (Gann *et al.*, 1987; Bickle and Kruger, 1993). The variable regions found in BcgI could be TRDs, and each could recognize one of the half sites of the BcgI recognition sequence. It was proposed that two R-M subunits associate with one S subunit via interaction between the dimerization domains in two R-M subunits, and two conserved domains in a single S subunit. This would allow the active A_2B complex to bind to its target DNA (Kong *et al.*, 1994).

The dependence of BcgI methylation on AdoMet concentration is a typical sigmoidal curve, indicating the presence of cooperativity of AdoMet in the DNA methylation reaction. Kinetic studies of BcgI methylation suggest that AdoMet acts both as a methyl donor and an allosteric effector. The DNA methylation and cleavage activities of the BcgI enzyme are very interesting in that the two competing activities share the same cofactors, AdoMet and Mg^{2+}, the divalent cation required for endonucleolytic digestion, which can greatly stimulate the methylation activity of BcgI. Therefore, the cofactor requirement of BcgI is similar to that of Type III R-M enzymes. It has been shown that AdoMet and Mg^{2+} are required for both DNA cleavage and methylation function in EcoP15I R-M system (Bist *et al.*, 2001).

It is speculated that the BcgI R-M system might have been derived from a precursor to a Type I R-M system, based on the following reasons: (a) both BcgI and Type I R-M systems contain a specificity subunit with two variable regions and two conserved regions; (b) the recognition sequences of both BcgI and Type I R-M systems are asymmetric bipartite sites, with specific segments separated by a nonspecific spacer of defined length; (c) AdoMet serves as an allosteric activator for both BcgI and Type I restriction enzymes such as EcoKI (Burckhardt *et al.*, 1981a; Burckhardt *et al.*, 1981b); and (d) both BcgI and Type I restriction enzymes are N^6-adenine MTases, and both BcgI and some of the Type I restriction enzymes prefer hemimethylated DNA substrates in methylation reactions (Suri *et al.*, 1984; Kong and Smith, 1997). It is possible that the R and M genes that are separated in Type I R-M systems could be fused together to form the BcgI R-M gene. Without the ATPase domain, the BcgI R-M system no longer requires ATP as its cofactor, nor can it translocate DNA. Instead, it cleaves DNA at a fixed distance on both sides of its recognition sequence (Kong, 1998). This makes BcgI distinct from Type I R-M systems. Another enzyme with similar properties has been reported in *H. pylori* (Tomb *et al.*, 1997). It was predicted that more BcgI like R-M systems may exist in different *H. pylori* strains.

HaeIV R-M system

This enzyme recognizes a symmetric bipartite sequence, 5′-GAPy(N_5)PuTC-3′, and cleaves double-stranded DNA on both strands upstream and downstream of the recognition sequence, releasing a ~33 bp fragment (Piekarowicz *et al.*, 1999). It is composed of a single polypeptide having both MTase and ENase activities, and the active form of the enzyme is a dimer. RM.HaeIV methylates adenine residues on both strands within the recognition sequence. The MTase domain of RM.HaeIV is located in the central part of the protein. This part has significant homology to the central part of the RM.Eco57I enzyme. However, R.Eco57I cleaves DNA only on one side of its asymmetric recognition sequence, needs AdoMet for restriction activity and methylates only one strand. Although BcgI and HaeIV cleave both sides of their recognition sequence, they differ in that BcgI needs AdoMet for cleavage, and recognizes the symmetrical recognition sequence.

AloI R-M system

AloI is a bifunctional enzyme that possesses both DNA cleavage and methylation activities (RM.AloI). The viability of *E. coli* cells expressing the AloI ENase indicates that the methylation activity of the bifunctional enzyme, modifying adenine residues in both DNA strands, is able to protect the host DNA against R.AloI. Thus, no separate cognate MTase, the necessary component of classical R-M systems, is required, and the two activities of the AloI R-M system are contained within a single polypeptide (Cesnaviciene *et al.*, 2001). The structure of the AloI C-terminal part reveals similarities to specificity (HsdS) subunits of Type I R-M systems (Cesnaviciene *et al.*, 2001). The structure of the bipartite, asymmetric target GGAN_6GTTC recognized by AloI resembles DNA targets typical for Type I R-M systems. It was predicted that the C-terminal regions of AloI-like Restriction-endonuclease (REases), just like HsdS subunits, have two TRDs, each recognizing an individual part of the bipartite DNA target (Kong and Smith, 1997).

The properties of AloI and the HaeIV R-M system indicate that they are related. Their possible evolutionary

scenario might have employed the same mechanism, i.e. fusion of HsdS, HsdM, and an ENase. In addition, AloI reveals properties of maintenance MTase, e.g. Type I restriction enzymes. R.AloI activity is not affected, in contrast to Type I restriction enzymes, by ATP. For ENase activity, the enzyme requires only the presence of Mg^{2+}. AdoMet, unlike other bifunctional ENases, does not stimulate the endonucleolytic activity of AloI, while DNA methylation, occurring in conjunction with DNA cleavage, protects some of the target sites. Methylation by AloI is stimulated by metal ions (Ca^{2+}) (Cesnaviciene *et al.*, 2001). These properties closely resemble those of the HaeIV R-M system. On the other hand, AloI and BcgI differ in enzymatic properties. BcgI needs AdoMet not only for methylation but also for cleavage, while, in the presence of AdoMet, AloI yields incomplete DNA fragmentation presumably due to a competing methylation reaction. Sinefungin, which effectively substitutes for AdoMet as a necessary cofactor or activator of the DNA cleavage reaction in other bifunctional restriction endonucleases such as BcgI, BplI, BseMII, or Eco57I, has no effect on the ENase activity of AloI.

Eco57I, MmeI and TspGWI R-M systems

All these MTases belong to the Type IIG R-M system. The Eco57I R-M system consists of an ENase–MTase hybrid (RM.Eco57I) that cleaves both DNA strands on one side of an asymmetric recognition sequence, and also methylates one of the DNA strands. In addition to the hybrid enzyme (bifunctional enzyme), the Eco57I R-M system consists of a separate MTase, which recognizes the same sequence as the hybrid enzyme and methylates both DNA strands (Janulaitis *et al.*, 1992a; 1992b). However, in contrast to BspLU11III, only one MTase is present to modify both strands of the Eco57I recognition site.

Both the RM.Eco57I and M.Eco57I enzymes modify the outer adenine residue in the target sequence 5′-CTGAAG-3′/3′-CTTCAG-5′, yielding N^6-methyladenine. M.Eco57I modifies both strands of the substrate while R.Eco57I modifies only one (Figure 3C). It was found that M.Eco57I enzyme activity is stimulated by Ca^{2+} and Mg^{2+} by a 20-fold magnitude. ATP has no influence on either activity of the enzymes. However, differences were found in the influence of pH, and ionic strength, on the enzymatic properties of both MTases (Janulaitis *et al.*, 1992a). M.Eco57I catalyzes DNA modification in the absence of Mg^{2+}, although such catalysis is stimulated by Mg^{2+}.

RM.Eco57I, possessing both DNA methylation and cleavage functions, very closely resembles a hypothetical Mod-Res fusion protein. Interestingly, R.Eco57I is able to methylate only one of the two strands of the target sequence, which might be due to steric restrictions imposed by the asymmetric location of the cleavage site. This methylation, most likely, is a relic of progenitor activity, since it is not by itself sufficient to protect the host DNA *in vivo*, and therefore, an additional copy of the "free" methylase is needed. The properties of RM.Eco57I indicate that it may be regarded as an intermediate type of R-M system reflecting the evolutionary link between Type III and IIG R-M systems.

The MTases Eco57I, BcgI, HaeIV, AloI, and all Type I restriction enzymes share a common property in that they methylate adenine residues, and that their MTases belong to the γ group of N^6-adenine MTases. It was therefore, speculated that all these R-M systems might be derived from a Type I R-M system. The fusion of all three genes together, with some kind of rearrangement, could produce R-M systems composed of a single protein (such as Eco57I, HaeIV, and AloI) or two proteins, such as BcgI (Piekarowicz *et al.*, 1999). Therefore, such rearrangements could lead to the observed differences between these enzymes in subunit composition (single protein versus two subunits), cofactor requirements, and the type of the recognition sequence.

MmeI R-M system

The MmeI R-M system consists of an ENase-MTase hybrid that cleaves DNA, and also methylates the adenine residue present in the asymmetric recognition sequence, 5′-TCCRAC-3′ (R indicates G or A). Initially it was found that MmeI R-M system consists of a separate MTase, which recognizes the same sequence and methylates adenine residues on both DNA strands (Tucholski *et al.*, 1998). However, RM.MmeI alone can be cloned and expressed even though it produces only single-strand modification, in contrast to Eco57I, which cannot be expressed without a second M.Eco57I (Morgan *et al.*, 2008). DNA sequencing of the *MmeI* gene showed no additional MTase gene on either flanking regions. Consistent with this, it was found that single-strand methylation by RM.MmeI prevents the ENase activity of RM.MmeI. In contrast to the previous results, it was shown that MmeI is not sensitive to bottom strand adenine methylation. Further, it was shown that even after cleaving the DNA RM.MmeI could still modify the intact recognition site. The ENase activity of RM.MmeI is relatively fast compared to its MTase activity, indicating some kind of regulation of competing activities in a single polypeptide (Morgan *et al.*, 2008). Many properties of the MmeI R-M system – stimulation of its restriction activity by AdoMet, and the specificity of methylation activity – indicate its similarity to the Eco57I R-M system.

The availability of complete genome sequences of bacteria led to the identification of 20 putative genes highly similar to RM.MmeI (Morgan *et al.*, 2009). No additional DNA MTase genes were observed flanking either side of these MmeI homologs and all of them employ single-strand methylation for host protection. The RM.MmeI

family of enzymes requires two unmethylated recognition sites for efficient cleavage. This high frequency would suggest that they might have an important biological function. Therefore, the diversity of R-M systems continues to grow as Type IIG systems have evolved in a second line, whereas a fused ENase-MTase requires only single-strand methylation for host protection. This high capacity of R-M systems to evolve may confer a selective advantage for organisms facing a rapidly evolving phage challenge.

Three functional domains in the structure of the MmeI enzyme include: (i) an N-terminal portion containing the endonucleolytic domain with the catalytic Mg^{2+}-binding motif D_{70}-X_9-EXK_{82},characteristic of the PD-(D/E)XK superfamily of nucleases; (ii) a central portion (amino acids 310–610) containing nine sequence motifs conserved among N^6-adenine γ–class DNA MTases; and (iii) a C-terminal portion (amino acids 610–919) containing a putative target recognition domain. All three domains show close similarity to the corresponding elements of Type I R-M enzymes rather than to classical Type II R-M enzymes (Nakonieczna *et al.*, 2009).

TspGWI R-M system

TspGWI is a bifunctional protein that recognizes the 5′-ACGGA-3′ sequence and cleaves 11/9 nucleotides downstream. Structural analysis suggests that RM.TspGWI is a fusion protein comprising a tandem arrangement of Type I-like domains: a PD-(D/E)XK nuclease domain related to the domains in HsdR subunits, a HsdM-like domain comprising a highly conserved AdoMet binding catalytic MTase domain, and a C-terminal extension, responsible for DNA sequence recognition (Zylicz-Stachula *et al.*, 2009). On the other hand, TspGWI lacks other domains characteristic of Type I R-M enzymes, including the ATP-dependent translocase domain and the C-terminal domains of either HsdM or HsdR, which are important for interactions between Type I R-M subunits. This domain organization resembles the structure of RM.MmeI (Nakonieczna *et al.*, 2009). The MTase activity of RM.TspGWI is stimulated by Mg^{2+}, Mn^{2+}, Fe^{2+}, Ca^{2+}, and Co^{2+}. It has been shown that the MTase and ENase activities of TspGWI are independent and can be uncoupled (Zylicz-Stachula *et al.*, 2009).

Multimeric DNA MTases recognizing asymmetric bipartite sequence

The MTases of this group are multimeric proteins that recognize asymmetric bipartite sequences and belong to the Type I R-M system. Type I R-M systems are complex, oligomeric proteins with both restriction and methylation activities carried out by the same enzyme. They usually consist of three subunits, HsdS, HsdM, and HsdR. The HsdS subunit serves as a core subunit for both restriction and modification to which the other subunits bind, and specifies the target sequence. It consists of two TRDs. The N-terminal TRD recognizes the 5′-part of the bipartite recognition sequence (two half-sites separated by a non specific spacer of 6–8 bp), and the C-terminal TRD recognizes the 3′-part of the target (Fuller-Pace and Murray, 1986; Gubler *et al.*, 1992; Sistla and Rao, 2004). HsdM subunits carry out methylation of adenine residues. The target specificity is provided by the HsdS subunit, and the stoichimetry for methylation is M_2S_1 (Taylor, *et al.*, 1992; Dryden *et al.*, 1993; Janscak and Bickle, 1998). For all the target sequences recognized, one adenine residue targeted for methylation is located within the top strand of the DNA at the 5′ end of the target, and the second adenine is located within the 3′ part of the target. Hemimethylated DNA is the preferred target for methylation (Suri and Bickle, 1985; Fuller-Pace and Murray, 1986). All the MTases of Type I R-M system belongs to the γ- subgroup and are N^6-adenine MTases (Table 5).

EcoKI DNA MTase (M.EcoKI) recognizes the bipartite DNA sequence, 5′-AACNNNNNNGTGC-3′, where N is any nucleotide, and methylates the N^6-position of

Table 5. Multimeric DNA MTases recognizing asymmetric bipartite sequences.

Enzyme	Recognition sequence and base modified[a]
M.KpnBI	5′-CAAA^{m6}(N_6)RTCA-3′
	3′-GTTT(N_6)YA^{m6}GT-5′
M.EcoAI	5′-GA^{m6}G(N_7)GTCA-3′
	3′-CTC(N_7)CA^{m6}GT-5′
M.EcoBI	5′-TGA^{m6}(N_8)TGCT-3′
	3′-ACT(N_8)A^{m6}CGA-5′
M.EcoKI	5′-AA^{m6}C(N_6)GTGC-3′
	3′-TTGC(N_6)A^{m6}CG-5′
M.EcoR124I	5′-GAA^{m6}(N_6)RTCG-3′
	3′-CTT(N_6)YA^{m6}GC-5′
M.EcoR124II	5′-GAA^{m6}(N_7)RTCG-3′
	3′-CTT(N_7)YA^{m6}GC-5′
M.KpnAI	5′-GAA^{m6}(N_6)TGCC-3′
	3′-CTT(N_6)A^{m6}CGG-5′
M.StyLTIII	5′-GA^{m6}G(N_6)RTAYG-3′
	3′-CTC(N_6)YA^{m6}TRC-5′
M.StySEAI	5′-ACA^{m6}(N_6)TYCA-3′
	3′-TGT(N_6)A^{m6}RGT-5′
M.StySENI	5′-CGA^{m6}(N_6)TACC-3′
	3′-GCT(N_6)A^{m6}TGG-5′
M.StySGI	5′-TAA^{m6}(N_6)RTCG-3′
	3′-ATT(N_6)YA^{m6}GC-5′
M.StySPI	5′-AA^{m6}C(N_6)GTRC-3′
	3′-TTG(N_6)CA^{m6}YG-5′
M.HindI	5′-CA^{m6}CN-3′
	3′-GTGN-5′

[a]N = A or C, G or T; R = G or A; Y = C or T.

adenine within both parts of the target sequence (Table 5, Figure 3D). M.EcoKI comprises two methylation (M) subunits, and one specificity (S) subunit, with a combined molecular weight of 169,000 Da (Dryden *et al.*, 1993). The addition of two restriction subunits forms the complete Type I restriction enzyme R.EcoKI, capable of performing both methylation and restriction. Mutation in motif I (first G in the FXGXG motif to D) completely abolished AdoMet binding without altering enzyme structure and DNA target recognition. Substitution of the N with D, or F with either G or C, in motif II (NPPF), abolished enzyme activity, but left AdoMet and DNA binding unaltered (Willcock *et al.*, 1994). M.EcoKI displays a strong preference (50-fold difference) for methylating hemimethylated targets, produced after each round of host DNA replication, over unmethylated targets (Dryden *et al.*, 1993; Dryden, 1999).

The preference of M.EcoKI for methylating hemimethylated target sequences requires a strong degree of internal communication between the parts of the enzyme surrounding each half of the bipartite target sequence. The communication of the methylation status from one adenine target to the other presumably occurs via conformational changes induced by the recognition of the methyl groups on the adenines by the enzyme, via steric clashes between the enzyme–AdoMet complex, and the DNA (Burckhardt *et al.*, 1981a). The recognition of adenine methylation requires flipping of the target base out of the DNA helix into the enzyme methylation pocket, where it can interact with AdoMet (Su *et al.*, 2004). By using 2-aminopurine fluorescence, it was shown that no base flipping occurred upon the introduction of a single-stranded gap in the spacer region. M.EcoKI bends the non-specific spacer, and the energy stored in a double-stranded bend is utilized to flip out the bases (Su *et al.*, 2005).

The C-terminal region of the HsdM subunit of EcoKI is susceptible to cleavage by the protease elastase, and leaves a 55 kDa fragment, which is resistant to further proteolysis by elastase when AdoMet is present (Cooper and Dryden, 1994). The fragment fails to interact with the other subunits, even though it still possesses secondary and tertiary structure, and the ability to bind the AdoMet cofactor. This M_1S_1 protein can bind DNA and AdoMet but is unable to methylate DNA in the absence of a second HsdM subunit (Kelleher *et al.*, 1991; Dryden *et al.*, 1993; 1995; 1997; Powell *et al.*, 1998; 2003). The hexachlorofluorescein anisotropy results indicate that M_1S_1 has poorer affinity for the EcoKI DNA target than the complete MTase M_2S_1. The presence of the second M subunit in the M_2S_1 MTase is very important for DNA sequence recognition, even though the M subunit by itself does not bind DNA, and the HsdS subunit is solely responsible for sequence recognition (Powell *et al.*, 1998). Methylation interference results indicate that although M_1S_1 binds AdoMet with an affinity similar to that of M_2S_1, the cofactor is unable to exert the same influence at the protein–DNA interface. Both DNase I and exonuclease III footprinting experiments showed weaker footprints for the M_1S_1 complex when compared to the M_2S_1, particularly on the complementary bottom strand where the exonuclease III nuclease was able to completely displace M_1S_1 from the DNA (Powell *et al.*, 1998). The presence of the cofactor AdoMet has a striking effect on the interference pattern for unmodified DNA, while it does not affect the interference pattern for hemimethylated or fully modified DNA. These results demonstrate that AdoMet is important for discrimination between unmodified and modified DNA by M.EcoKI (Powell and Murray, 1995).

Recently, negative-stain electron microscopy structures, combined with the computational model, of the M.EcoKI core complex (M_2S_1) bound to Ocr (a DNA mimic, T7 phage antirestriction protein) have been reported (Kennaway *et al.*, 2009). The interface between the two HsdM subunits was formed by their N-terminal domains (amino acids 1–153). The N-terminal domain of each HsdM contains the m* mutations, so called because they change the strong preference of M.EcoKI for methylating the hemimethylated DNA target sequences, in addition to the methylation of unmodified targets (Kelleher *et al.*, 1991; Dryden *et al.*, 1993). Interactions between HsdM and HsdS occur via the C-terminal regions of each HsdM (amino acids 470–529). This region was disordered in the HsdM crystal structures suggesting that this region might be flexible in the absence of HsdS (Burckhardt *et al.*, 1981a; Kneale, 1994; Dryden *et al.*, 1995; Obarska *et al.*, 2006).

EcoR124I DNA MTase (M.EcoR124I) consists of two HsdM subunits (each 58 kDa) and one HsdS subunit (46 kDa), to form a trimeric enzyme (162 kDa) with a subunit stoichiometry of M_2S_1 (Taylor *et al.*, 1992; 1993). M.EcoR124I recognizes the bipartite sequence 5′-GAANNNNNNRTCG-3′ and methylates the N^6 position of adenine residues on both strands (Table 5). It was shown that M.EcoR124I with bound DNA confers considerable protection from proteolysis (Webb *et al.*, 1995). Non-sequence-specific interactions between HsdM and DNA have been shown by chemical modification experiments, where lysine residues in HsdM as well as HsdS were implicated in contacts with DNA (Taylor *et al.*, 1996). Chemical modification experiments demonstrated that a large fraction of the lysine residues on the surface of the protein are inaccessible in the DNA–protein complex, including a number of sites in regions implicated in inter-subunit contacts, as well as likely DNA binding residues (Taylor *et al.*, 1996). These results indicated that the MTase undergoes a significant conformational change when bound to DNA.

X-ray scattering and DNA footprinting (Mernagh and Kneale, 1996) indicated that the MTase completely

encloses the DNA. The circular dichroism spectrum showed that the structural transition in the enzyme due to DNA binding is accompanied by considerable distortion of the DNA structure in the complex. The N-terminal region of HsdS is in close proximity to the C-terminal region, with the two held in place by interactions with one of the HsdM subunits (Kneale, 1994; Mernagh *et al.*, 1997). It was demonstrated that both HsdM and HsdS are required for effective DNA binding, and thus HsdM must play a structural role, in addition to its catalytic role. HsdM is also required to maintain the solubility of HsdS (Patel *et al.*, 1992). Thus, the role of HsdM in promoting DNA binding could be by maintaining solubility of the HsdS subunit (Webb *et al.*, 1995).

It was shown that the rate of methylation of each adenine by M.EcoRl241 is increased at least 100-fold by prior methylation at the other site. However, this is accompanied by a significant decrease in the affinity of the MTase for these substrates as shown by competitive gel retardation assays. In the case of M.EcoKI, unmethylated sites are modified very slowly, whereas if the site is hemimethylated, the second adenine is rapidly modified (Suri *et al.*, 1984; Dryden *et al.*, 1993). M.EcoR124I showed a higher reaction rate for hemimethylated DNA than for the unmodified substrate. Modification of either half-site greatly affects the rate of methylation at the other, implying a significant communication between the two active sites on the MTase. This could occur through a conformational change in the DNA, or through a structural change in the EcoRl241 MTase itself (Taylor *et al.*, 1993). The large increase in the rate of reaction for hemimethylated substrates is unlikely to be the result of an increase in the affinity of the MTase for the DNA substrate. It was demonstrated by gel retardation competition assays that the affinity of the MTase for either hemimethylated duplex is indeed lower than for the unmodified form.

DNA MTases recognizing asymmetric sequences and methylate only one strand

In general, MTases belonging to Type III R-M systems recognize short (5–6 bp) asymmetric recognition sequences, and methylate only one strand of the DNA (hemi-methyltransferases). Interestingly, all members of this group are the β-subclass of MTases and methylate adenine residues, yielding N^6-methyladenosine. There are six known Type III R-M systems: EcoP15I, EcoP1I, StyLTI, HinfIII, HindVI and PstII (Table 6).

Type III R-M systems consist of two subunits, Mod (M), which is a functional DNA MTase, able to recognize and modify the target sequences, and Res (R), which is responsible for DNA cleavage, but functions only in a complex with the Mod subunit (Sistla and Rao, 2004). Therefore, Type III restriction enzymes are heterotetramers, which use the Mod subunit for sequence recognition and methylation, and the Res subunit for DNA cleavage. Cleavage always occurs in the vicinity of rather than within the asymmetric target sequence, so that both subunits can be accommodated. The modification subunit of Type III restriction enzymes functions as a separate MTase (Hadi *et al.*, 1983) because it contains both TRD and MTase domains (Ahmad and Rao, 1996a). However, for the Type III R-M system, it was shown that the modification of only one strand is sufficient to protect the DNA cleavage by the cognate ENase (Meisel *et al.*, 1992; 1995; Kruger *et al.*, 1995). EcoP15I and EcoP1I are two well-characterized members of the Type III R-M system (Rao *et al.*, 2000; Dryden *et al.*, 2001; McCleland and Sczcelkun, 2004). It was shown that the Mod subunit stabilizes the Res protein during expression (Redaschi and Bickle, 1996).

EcoP15I DNA MTase

EcoP15I DNA MTase (M.EcoP15I) recognizes the asymmetric recognition sequence 5′-CAGCAG-3′ and catalyses the transfer of a methyl group from AdoMet to the second adenine (Figure 3E) in the presence of magnesium (Meisel *et al.*, 1991). M.EcoP15I follows a random mechanism of substrate binding, in which either DNA or AdoMet can bind to the free enzyme. After the methyl-transfer reaction, the product release (methylated DNA and AdoHcy) also follows a random order. DNA-dependent substrate inhibition by AdoMet suggested that the release of methylated DNA might be the rate-limiting step (Rao *et al.*, 1989). Studies using gel filtration and chemical cross-linking demonstrated that M.EcoP15I exists as a dimer in solution (Ahmad *et al.*, 1995).

M.EcoP15I is an N^6-adenine MTase, and belongs to the β-group of amino MTases. M.EcoP15I, like all N^6-adenine MTases, contains nine characteristic motifs, including two highly conserved sequences, FXGXG (motif I), involved in AdoMet binding, and (DPPY) (motif IV) involved in catalysis. Methylation protection analysis demonstrated

Table 6. DNA MTases recognizing asymmetric sequences and methylating only one strand.

Enzyme	Recognition sequence and base modified
M.StyLTI	5′-CAGA^{m6}G-3′
	3′-GTCTC-5′
M.EcoP15I	5′-CAGCA^{m6}G-3′
	3′-GTCGTC-5′
M.EcoP1I	5′-AG^{m6}ACC-3′
	3′-TCTGG-5′
M.HinfIII	5′-CGAA^{m6}T-3′
	3′-GCTTA-5′
M.HindVI	5′-CGAA^{m6}T-3′
	3′-GCTTA-5′
M.PstII	5′-CTGATG-3′
	3′-GACTA^{m6}C-5′

that M.EcoP15I makes contacts in the major groove of its substrate DNA. The results of the partial proteolysis of the wild-type M.EcoP15I-AdoMet complex (Ahmad and Rao, 1994), and the enhanced UV crosslinking of AdoMet to the mutant Y128W enzyme, suggested that the N-terminal half of the enzyme might be involved in AdoMet binding. However, mutations in motif I (which lies in the C-terminal half) of the protein abolish AdoMet binding. Similarly, the partial proteolysis of the M.EcoP15I-DNA adduct indicated that the 50 kDa C-terminal domain was also involved in DNA recognition. Therefore, it was proposed that DNA binds inside a cleft created by the two domains (Ahmad and Rao, 1996b).

M.EcoP15I can bind DNA in the absence of AdoMet, but the presence of AdoMet or its analogs enhanced the binding. However, in the presence of ATP, M.EcoP15I was able to discriminate between specific and non-specific DNA significantly (Ahmad and Rao, 1994). When the M.EcoP15I DNA bound to DNA containing 2-aminopurine substitutions within the cognate sequence, an eight to 10-fold fluorescent enhancement resulting from enzymatic flipping of the target adenine base was observed (Reddy and Rao, 2000). It has been demonstrated, using chemical modification and site-directed mutagenesis, that cysteine 344 and histidine 335 play an essential role in the DNA methylation reaction catalyzed by M.EcoP15I (Reddy and Rao, 1998; Jois *et al.*, 2008). Both these amino acids are present in the putative TRD region of M.EcoP15I.

An interesting aspect of M.EcoP15I is that the methylation of the 5′-CAGCAG-3′ sequence requires magnesium ions. It has been shown that the PD(X)$_n$(D/E)XK-like motif present in M.EcoP15I is involved in magnesium binding (Bist and Rao, 2003). This motif is characteristic of the majority of Type I, II, III, and IV restriction enzymes as well as other nucleases. It was surprising to find such a Mg^{2+}-binding **M**D(X)$_n$(D/E)XK insertion domain in the M.EcoP15I. Although Mn^{2+} enhances the binding of M.EcoP15I to DNA, methylation activity was found to be less than 2% in the presence of Mn^{2+} when compared with Mg^{2+} (Ahmad and Rao, 1994). It was shown that substitution of methionine by conserved proline converts M.EcoP15I to an Mg^{2+}-dependent site-specific endonuclease (Bist *et al.*, 2007). The cryptic endonuclease activity in EcoP15I DNA MTase perhaps provides an evolutionary link between restriction ENase and the MTase protein families.

EcoP1I DNA MTase

M.EcoP1I DNA MTase (M.EcoP1I), an N^6-adenine MTase from bacteriophage P1, belongs to the Type III R-M system. It recognizes the sequence 5′-AGACC-3′ and methylates the internal adenine at the N^6 position. It requires Mg^{2+} for the methyl-transfer reaction. Deletion mutants of M.EcoP1I revealed that the AdoMet binding domain is probably located at the N-terminal half of the protein (Krishnamurthy and Rao, 1994). It was also shown that dimeric M.EcoP1I could bind and methylate both single-stranded and double-stranded DNA with similar K_M and k_{cat} values (Sistla *et al.*, 2004). While M.EcoP1I bound to ssDNA in the absence of Mg^{2+}, the presence of Mg^{2+} greatly enhanced the single-strand DNA binding. M.EcoP1I showed a greater affinity of binding to single-stranded DNA than double-stranded DNA. However, the specific activity of M.EcoP1I, for both double-stranded pUC19 DNA as well as single-stranded M13mp18 DNA, was almost the same (Sistla *et al.*, 2004). The interesting property of binding and methylation of single-stranded DNA by M.EcoP1I might have an important role in the biology of single-stranded phages (Sistla *et al.*, 2004).

Phase variable Type III R-M systems

Type III R-M systems represent an atypical class of phase variable genes. Phase variable Type III R-M systems are present in a variety of pathogenic bacteria. The presence of repeats within a gene is a strong indicator that the gene is phase variable at a frequency determined by the hypermutability of its microsatellite. This has been experimentally proven in organisms such as *Haemophilus influenzae* (De Bolle *et al.*, 2000; Srikhanta *et al.*, 2005) and *H. pylori* (de Vries *et al.*, 2002). From the sequence analysis, it was predicted that Type III MTases would undergo phase variation in pathogenic organisms such as *Pasteurella haemolytica* (Ryan and Lo, 1999), *Neisseria meningitidis* (Fox *et al.*, 2007a), *Neisseria gonorrhoeae* (Saunders *et al.*, 2000) and *Moraxella catarrhalis* (Seib *et al.*, 2002). It has been shown that within a single strain of human pathogens, *N. meningitidis*, *N. gonorrhoeae*, *H. pylori* and *M. catarrhalis*, multiple phase variable Type III *mod* genes are present. *H. influenzae*, which has only a single *modA* gene, displays significant sequence diversity among strains, with 15 distinct *mod* alleles being identified (Bayliss *et al.*, 2006; Fox *et al.*, 2007b). Many of these *mod* genes showed distinct sequence specificity (Fox *et al.*, 2007b). The multiple *mod* genes (*modA* and *modB*) of pathogenic *Neisseria* also display distinct *mod* alleles (Fox *et al.*, 2007a). A recent report suggests that multiple phasevarions exist within the pathogenic *Neisseria*, each regulating a different set of genes (Srikhanta *et al.*, 2009). Both *N. meningitidis* and *N. gonorrhoeae* have two distinct *mod* genes, *modA* and *modB*, and these genes switch independently. There are also distinct alleles of *modA* (major alleles include *modA11*, *12* and *13*, and minor alleles include *modA4*, *15* and *18*) and *modB* (*modB1*, *2*). These alleles differ only in their DNA recognition domain. *modA11* was only found in *N. meningitidis* and *modA13* only in *N. gonorrhoeae*. When expressed, ModA13 recognizes and methylates all

5′-AGAAA-3′ sites in the genome, and thereby controls gene expression. It was shown that two strains with the same DNA recognition domain (*mod*A13 allele) regulated the same set of genes, while *N. meningitides modA11* and *modA12* were found to regulate the expression of different sets of genes, consistent with differences in their DNA recognition domain (Srikhanta *et al*., 2009).

Conclusions

Since most bacteria and archaea possess one or more R-M systems, the recent availability of complete genome sequences have provided quite a number of putative R-M systems. Consistent with this, the understanding of the diversity of MTases that recognize asymmetric sequences continues to grow as more and more variations on the organization and general theme of methylation are being characterized. Though highly conserved, the MTase family contains several members with variant structural features. While some MTases are homodimeric or even tetrameric, most are monomers. These differences suggest structural flexibility but it seems that MTases have variations on a single basic theme. An important unanswered question is how does quaternary structure of the MTase influence its catalytic activity? It is quite evident that there is no correlation between target site asymmetry and the oligomeric nature of these MTases. Recent evidences suggest that R-M systems are rapidly evolving to compete with the constantly evolving phage challenge. If a phage escapes from cleavage, the phage can infect the cell and its progeny carries the same methylation patterns as the host cell. Then other bacteria containing the same R-M system are no longer protected against infection. Therefore, it is advantageous that different bacteria carry R-M systems with different recognition sequences, and as a result a large diversity of R-M systems with different recognition sequences has developed through the billions of years of bacteria or phage coevolution. Two R-M systems of the same specificity are unable to enjoy stabilization simultaneously. This type of "incompatibility" implies competition for specific sequences by R-M systems, which would result in each selfish-gene unit specializing in only one specific sequence among diverse sequences. Similarly, DNA MTases not paired with a restriction endonuclease can alleviate post-segregational killing by an R-M gene complex. Thus, solitary MTases may act as live vaccines to defend the genome against pathogenic R-M complexes (Kobayashi, 2004). How the cells coordinate these two contrasting enzyme activities and phage challenges is a current area of research.

The ability to covalently add methyl groups to the target adenine or cytosine residues in specific DNA sequences, without any other changes in the primary DNA structure, is a remarkable feature of DNA MTases. This property of DNA MTases has been used as a practical tool in recombinant DNA technology, to determine the site-specific level of methylation of eukaryotic DNA, and to probe the chromatin structure (Buryanova and Shevchuk, 2005; Jeltsch *et al*., 2007). However, the use of prokaryotic DNA MTases in basic and applied research is still in its infancy. Recent evidence suggests that in addition to protecting the bacterial genome from external DNA, R-M systems have other biological functions. The genome analysis of the first two sequenced strains of *H. pylori* (26695 and J99) showed that this bacterium possesses an unusually high number of strain-specific R-M genes. *H. pylori* is well adapted to the gastric environment, and the acquisition of numerous R-M systems might be related to its unique adaptability to an acidic environment. Phase variable Type III R-M systems are present in a variety of pathogenic bacteria. The widespread distribution of phase variable R-M systems in host-adapted pathogenic bacteria suggests that this regulated random switching of multiple genes may be a commonly used strategy for bacterial pathogens. The identification and study of both species-specific and strain-specific MTases of pathogenic bacteria will certainly improve the understanding of their pathogenic mechanisms.

Acknowledgements

Work from our laboratory has been supported by grants from the DST, CSIR, DBT, and DAE, Government of India. UKM acknowledges CSIR for a senior research fellowship. We thank Ruchika Sharma, Arun Banerjee and Ritesh Kumar for carefully reading the manuscript and Shivakumar for helping us with the figures. DNR acknowledges J.C. Bose Fellowship from DST.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Abdurashitov MA, Kileva EV, Shinkarenko NM, Shevchenko AV, Dedkov VS and Degtyarev S. 1996. BstF5I, an unusual isoschizomer of FokI. Gene 172: 49–51.

Ahmad I and Rao DN. 1994. Interaction of EcoP15I DNA methyltransferase with oligonucleotides containing the asymmetric sequence 5′-CAGCAG-3′. J Mol Biol 242:378–388.

Ahmad I and Rao DN. 1996a. Chemistry and biology of DNA methyltransferases. Crit Rev Biochem Mol Biol 31:361–380.

Ahmad I and Rao DN. 1996b. Functional analysis of conserved motifs in EcoP15I DNA methyltransferase. J Mol Biol 259: 229–40.

Ahmad I, Krishnamurthy V, Rao DN. 1995. DNA recognition by the EcoP15I and EcoPI modification methyltransferases. Gene 157:143–147.

Allan BW and Reich NO. 1996. Targeted base stacking disruption by the EcoRI DNA methyltransferase. Biochemistry 35:14757–14762.

Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, Trust TJ. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature 397:176–180.

Alves J, Selent U and Wolfes H. 1995. Accuracy of the EcoRV restriction endonuclease: binding and cleavage studies with oligodeoxynucleotide substrates containing degenerate recognition sequences. Biochemistry 34:11191–11197.

Arber W. 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol Rev 24:1–7.

Arber W. 2002. Evolution of prokaryotic genomes. Curr Top Microbiol Immunol 264:1–14.

Baldwin GS, Kelly SM, Price NC, Wilson GW, Connolly BA, Artymiuk PJ and Hornby DP. 1994. Ligand-induced conformational states of the cytosine-specific DNA methyltransferase M.HgaI-2. J Mol Biol 235:545–553.

Barbeyron T, Kean K and Forterre P. 1984. DNA adenine methylation of GATC sequences appeared recently in the *Escherichia coli* lineage. J Bacteriol 160:586–590.

Bayliss CD, Callaghan MJ and Moxon ER. 2006. High allelic diversity in the methyltransferase gene of a phase variable type III restriction-modification system has implications for the fitness of *Haemophilus influenzae*. Nucleic Acids Res 34: 4046–4059.

Bheemanaik S, Chandrashekaran S, Nagaraja V and Rao DN. 2003. Kinetic and catalytic properties of dimeric KpnI DNA methyltransferase. J Biol Chem 278:7863–7874.

Bheemanaik S, Reddy YV and Rao DN. 2006a. Structure, function and mechanism of exocyclic DNA methyltransferases. Biochem J 399:177–190.

Bheemanaik S, Bujnicki JM, Nagaraja V and Rao DN. 2006b. Functional analysis of amino acid residues at the dimerisation interface of KpnI DNA methyltransferase. Biol Chem 387:515–523.

Bickle TA and Kruger DH. 1993. Biology of DNA restriction. Microbiol Rev 57:434–450.

Bist P and Rao DN. 2003. Identification and mutational analysis of Mg^{2+} binding site in *Eco*P15I DNA methyltransferase: involvement in target base eversion. J Biol Chem 278:41837–41848.

Bist P, Sistla S, Krishnamurthy V, Acharya A, Chandrakala B and Rao DN. 2001. S-adenosyl-L-methionine is required for DNA cleavage by type III restriction enzymes. J Mol Biol 310:93–109.

Bist P, Madhusoodanan UK and Rao DN. 2007. A mutation in the Mod subunit of EcoP15I restriction enzyme converts the DNA methyltransferase to a site-specific endonuclease. J Biol Chem 282:3520–3530.

Bitinaite J, Grigaite R, Maneliene Z, Butkus V and Janulaitis A. 1991. Esp3I – a novel type IIs restriction endonuclease from *Hafnia alvei* that recognizes the sequence 5′-CGTCTC(N)1/5-3′. Nucleic Acids Res 19:5076.

Bitinaite J, Maneliene Z, Menkevicius S, Klimasauskas S, Butkus V and Janulaitis A. 1992. Alw26I, Eco31I and Esp3I – type IIs methyltransferases modifying cytosine and adenine in complementary strands of the target DNA. Nucleic Acids Res 20:4981–4985.

Bocklage H, Heeger K and Muller-Hill B. 1991. Cloning and characterization of the MboII restriction-modification system. Nucleic Acids Res 19:1007–1013.

Bujnicki JM and Radlinska M. 1999. Molecular evolution of DNA-(cytosine-N4) methyltransferases: evidence for their polyphyletic origin. Nucleic Acids Res 27:4501–4509.

Burckhardt J, Weisemann J, Yuan R. 1981a. Characterization of the DNA MTase activity of the restriction enzyme from *Escherichia coli* K. J Biol Chem 256:4024–4032.

Burckhardt J, Weisemann J, Hamilton DL and Yuan R. 1981b. Complexes formed between the restriction endonuclease EcoK and heteroduplex DNA. J Mol Biol 153:425–440.

Buryanova Y and Shevchuk T. 2005. The use of prokaryotic DNA methyltransferases as experimental and analytical tools in modern biology. Anal Biochem 338:1–11.

Butkus V, Bitinaite J, Kersulyte D and Janulaitis A. 1985. A new restriction endonuclease Eco31I recognizing a non-palindromic sequence. Biochim Biophys Acta 826:208–212.

Carlson K and Kosturko LD. 1998. Endonuclease II of coliphage T4: a recombinase disguised as a restriction endonuclease? Mol Microbiol 27:671–676.

Cerritelli S, Springhorn SS and Lacks SA. 1989. DpnA, a MTase for single-strand DNA in the Dpn II restriction system, and its biological function. Proc Natl Acad Sci USA 86:9223–9227.

Cesnaviciene E, Petrusyte M, Kazlauskiene R, Maneliene Z, Timinskas A, Lubys A and Janulaitis A. 2001. Characterization of AloI, a restriction-modification system of a new type. J Mol Biol 314:205–216.

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

Cheng X and Roberts RJ. 2001. AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res 29:3784–3795.

Cheng X, Kumar S, Posfai J, Pflugrath JW and Roberts RJ. 1993. Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine. Cell 74:299–307.

Chernov AV, Matvienko NN, Zheleznaya LA and Matvienko NI. 1996. BspLU11III, a bifunctional restriction and modification enzyme from a thermophilic strain *Bacillus* species LU11. Nucleic Acids Res 23:1213–1214.

Chernukhin VA, Golikova LN, Gonchar DA, Abdurashitov MA, Kashirina YG, Netesova NA and Degtyarev S. 2003. M.*Bst*F5I-2 and M.*Bst*F5I-4 DNA methyltransferases from BstF5I restriction-modification system of *Bacillus stearothermophilus* F5. Biochemistry (Moscow). 68:967–975.

Cooper LP and Dryden DT. 1994. The domains of a type I DNA methyltransferase. Interactions and role in recognition of DNA methylation. J Mol Biol 236:1011–1021.

De Bolle X, Bayliss CD, Field D, van de Ven T, Saunders NJ, Hood DW and Moxon ER. 2000. The length of a tetranucleotide repeat tract in *Haemophilus influenzae* determines the phase variation rate of a gene with homology to type III DNA methyltransferases. Mol Microbiol 35:211–222.

Degtyarev S, Netesova NA, Abdurashitov MA and Shevchenko AV. 1997. Primary structure and strand specificity of *Bst*F5I-1 DNA methyltransferase which recognizes 5′-GGATG-3′. Gene 187:217–219.

de la Campa AG, Kale P, Springhorn SS and Lacks SA. 1987. Proteins encoded by the DpnII restriction gene cassette. Two methylases and an endonuclease. J Mol Biol 196:457–469.

de Vries N, Duinsbergen D, Kuipers EJ, Pot RG, Wiesenekker P, Penn CW, van Vliet AH, Vandenbroucke-Grauls CM and Kusters JG. 2002. Transcriptional phase variation of a type III restriction-modification system in *Helicobacter pylori*. J Bacteriol 184:6615–6623.

Dong A, Zhou L, Zhang X, Stickel S, Roberts RJ and Cheng X. 2004. Structure of the Q237W mutant of HhaI DNA methyltransferase: an insight into protein–protein interactions. Biol Chem 385:373–379.

Dryden DT. 1999. Bacterial DNA methyltransferases, pp. 283–340. In: Cheng, X and Blumenthal, RM, eds, S-Adenosylmethionine-Dependent Methyltransferases: Structures and Functions. Singapore: World Scientific Publishing.

Dryden DT, Cooper LP and Murray NE. 1993. Purification and characterization of the methyltransferase from the type 1 restriction and modification system of *Escherichia coli* K12. J Biol Chem 268:13228–13236.

Dryden DT, Sturrock SS and Winter M. 1995. Structural modelling of a type I DNA methyltransferase. Nat Struct Biol 2:632–635.

Dryden DT, Cooper LP, Thorpe PH and Byron O. 1997. The *in vitro* assembly of the EcoKI type I DNA restriction/modification enzyme and its *in vivo* implications. Biochemistry 36:1065–1076.

Dryden DT, Murray NE and Rao DN. 2001. Nucleoside triphosphate-dependent restriction enzymes. Nucleic Acids Res 29:3728–3741.

Dubey AK and Roberts RJ. 1992. Sequence-specific DNA binding by the MspI DNA methyltransferase. Nucleic Acids Res 20:3167–3173.

Fellinger K, Rothbauer U, Felle M, Langst G and Leonhardt H. 2009. Dimerization of DNA methyltransferase 1 is mediated by its regulatory domain. J Cell Biochem 106:521–528.

Fox KL, Dowideit SJ, Erwin AL, Srikhanta YN, Smith AL and Jennings MP. 2007a. *Haemophilus influenzae* phasevarions have evolved from type III DNA restriction systems into epigenetic regulators of gene expression. Nucleic Acids Res 35:5242–5252.

Fox KL, Srikhanta YN and Jennings MP. 2007b. Phase variable type III restriction-modification systems of host-adapted bacterial pathogens. Mol Microbiol 65:1375–1379.

Friedrich T, Fatemi M, Gowhar H, Leishmann O and Jeltsch A. 2000. Specificity of DNA binding and methylation by the M.FokI DNA Methyltransferase. Biochem. Biophys Acta 1480:145–159.

Fuller-Pace FV and Murray NE. 1986. Two DNA recognition domains of the specificity polypeptides of a family of type I restriction enzymes. Proc Natl Acad Sci USA 83:9368–9372.

Furmanek B, Sektas M, Wons E and Kaczorowski T. 2007. Molecular characterization of the DNA methyltransferase M1.NcuI from *Neisseria cuniculi* ATCC 14688. Res Microbiol 158:164–174.

Furmanek-Blaszk B, Boratynski R, Zolcinska N and Sektas M. 2009. M1.MboII and M2.MboII type IIS methyltransferases: different specificities, the same target. Microbiology 155:1111–1121.

Gann AA, Campbell AJ, Collins JF, Coulson AF and Murray NE. 1987. Reassortment of DNA recognition domains and the evolution of new specificities. Mol Microbiol 1:13–22.

Goedecke K, Pignot M, Goody RS, Scheidig AJ and Weinhold E. 2001. Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. Nat Struct Biol 8:121–125.

Gong W, O'Gara M, Blumenthal RM and Cheng X. 1997. Structure of Pvu II DNA-(cytosine N4) methyltransferase, an example of domain permutation and protein fold assignment. Nucleic Acids Res 25:2702–2715.

Gromova ES and Khoroshaev AV. 2003. Prokaryotic DNA methyltransferases: the structure and the mechanism of interaction with DNA. Mol Biol (Moscow) 37:300–314.

Gubler M, Braguglia, D, Meyer J, Piekarowicz A and Bickle TA. 1992. Recombination of constant and variable modules alters DNA sequence recognition by type IC restriction-modification enzymes. EMBO J 11:233–240.

Hadi SM, Bachi B, Iida S and Bickle TA. 1983. DNA restriciton – modification enzymes of phage P1 and plasmid P15B. Subunit functions and structural homologies. J Mol Biol 165:19–34.

Heitman J. 1993. On the origins, structures and functions of restriction-modification enzymes. Genet Eng (NY) 15:57–108.

Holz B, Klimasauskas S, Serva S and Weinhold E. 1998. 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases. Nucleic Acids Res 26:1076–1083.

Janscak P and Bickle TA. 1998. The DNA recognition subunit of the type IB restriction-modification enzyme EcoAI tolerates circular permutions of its polypeptide chain. J Mol Biol 284:937–948.

Janulaitis A, Petrusyte M, Maneliene Z, Klimasauskas S and Butkus V. 1992a. Purification and properties of the Eco57I restriction endonuclease and methylase-prototypes of a new class (type IV). Nucleic Acids Res 20:6043–6049.

Janulaitis A, Vaisvila R, Timinskas A, Klimasauskas S and Butkus V. 1992b. Cloning and sequence analysis of the genes coding for Eco57I type IV restriction- modification enzymes. Nucleic Acids Res 20:6051–6056.

Jeltsch A. 2002. Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem 3:274–293.

Jeltsch A. 2003. Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? Gene 317:13–16.

Jeltsch A, Christ F, Fatemi M and Roth M. 1999. On the substrate specificity of DNA methyltransferases. adenine-N6 DNA methyltransferases also modify cytosine residues at position N4. J Biol Chem 274:19538–19544.

Jeltsch A, Jurkowska RZ, Jurkowski TP, Liebert K, Rathert P and Schlickenrieder M. 2007. Application of DNA methyltransferases in targeted DNA methylation. Appl Microbiol Biotechnol 75:1233–1240.

Jia D, Jurkowska RZ, Zhang X, Jeltsch A and Cheng X. 2007. Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. Nature 449:248–251.

Jois PS, Madhu N and Rao DN. 2008. Role of histidine residues in EcoP15I DNA methyltransferase activity as probed by chemical modification and site-directed mutagenesis. Biochem J 410:543–553.

Jurenaite-Urbanaviciene S, Kazlauskiene R, Urbelyte V, Maneliene Z, Petrusyte M, Lubys A and Janulaitis A. 2001. Characterization of BseMII, a new type IV restriction-modification system, which recognizes the pentanucleotide sequence 5′-CTCAG(N)(10/8). Nucleic Acids Res 29:895–903.

Kaczorowski T, Sektas M, Skowron P and Podhajska AJ. 1999. The FokI methyltransferase from *Flavobacterium okeanokoites*. Purification and characterization of the enzyme and its truncated derivatives. Mol Biotechnol 13:1–15.

Kaszubska W, Webb HK and Gumport RI. 1992. Purification and characterization of the M.RsrI DNA methyltransferase from *Escherichia coli.* Gene 118:5–11.

Kelleher JE, Daniel AS and Murray NE. 1991. Mutations that confer *de novo* activity upon a maintenance methyltransferase. J Mol Biol 221:431–440.

Kennaway CK, Obarska-Kosinska A, White JH, Tuszynska I, Cooper LP, Bujnicki JM, Trinick J and Dryden DT. 2009. The structure of M.EcoKI Type I DNA methyltransferase with a DNA mimic antirestriction protein. Nucleic Acids Res 37:762–770.

Klimasauskas S, Timinskas A, Menkevicius S, Butkiene D, Butkus V and Janulaitis A. 1989. Sequence motifs characteristic of DNA[cytosine-N4]methyltransferases: similarity to adenine and cytosine-C5 DNA-methylases. Nucleic Acids Res 17:9823–9832.

Klimasauskas S, Kumar S, Roberts RJ and Cheng X. 1994. HhaI methyltransferase flips its target base out of the DNA helix. Cell 76:357–369.

Kneale GG. 1994. A symmetrical model for the domain structure of type I DNA methyltransferases. J Mol Biol 243:1–5.

Kobayashi I. 2004. Restriction-modification systems as minimal forms of life, pp. 19–62. In: Pingoud A., ed., Restriction Endonucleases Berlin: Springer.

Kong H. 1998. Analyzing the functional organization of a novel restriction modification system, the BcgI system. J Mol Biol 279:823–832.

Kong H and Smith CL. 1997. Substrate DNA and cofactor regulate the activities of a multi-functional restriction-modification enzyme, BcgI. Nucleic Acids Res 25:3687–3692.

Kong H, Morgan RD, Maunus RE and Schildkraut I. 1993. A unique restriction endonuclease, BcgI, from *Bacillus coagulans*. Nucleic Acids Res 21:987–991.

Kong H, Roemer SE, Waite-Rees PA, Benner JS, Wilson GG and Nwankwo DO. 1994. Characterization of BcgI, a new kind of restriction-modification system. J Biol Chem 269:683–690.

Krishnamurthy V and Rao DN. 1994. Interaction of EcoP1 modification methylase with S-adenosyl-L-methionine: a UV-crosslinking study. Biochem Mol Biol Int 32:623–632.

Kriukiene E, Lubiene J, Lagunavicius A and Lubys A. 2005. MnlI – The member of H-N-H subtype of Type IIS restriction endonucleases. Biochim Biophys Acta 1751:194–204.

Kruger DH, Kupper D, Meisel A, Reuter M and Schroeder C. 1995. The significance of distance and orientation of restriction endonuclease recognition sites in viral DNA genomes. FEMS Microbiol Rev 17:177–184.

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

Labahn J, Granzin J, Schluckebier G, Robinson DP, Jack WE, Schildkraut I and Saenger W. 1994. Three-dimensional structure of the adenine-specific DNA methyltransferase M.Taq I in complex with the cofactor S-adenosylmethionine. Proc Natl Acad Sci USA 91:10957–10961.

Lacks SA, Ayalew S, de la Campa AG and Greenberg B. 2000. Regulation of competence for genetic transformation in *Streptococcus pneumonia*: expression of *dpn*A, a late competence gene encoding a DNA methyltransferase of the DpnII restriction system. Mol Microbiol 35:1089–1098.

Landry D, Looney MC, Feehery GR, Slatko BE, Jack WE, Schildkraut I and Wilson GG. 1989. M.FokI methylates adenine in both strands of its asymmetric recognition sequence. Gene 77:1–10.

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

Lehours P, Dupouy S, Chaineux J, Ruskone-Fourmestraux A, Delchier JC, Morgner A, Megraud F and Menard A. 2007. Genetic diversity of the HpyC1I restriction modification system in *Helicobacter pylori*. Res Microbiol 158:265–271.

Leismann O, Roth M, Friedrich T, Wende W and Jeltsch A. 1998. The *Flavobacterium okeanokoites* adenine-N6-specific DNA-methyltransferase M.FokI is a tandem enzyme of two independent domains with very different kinetic properties. Eur J Biochem 251:899–906.

Lepikhov K, Tchernov A, Zheleznaja L, Matvienko N, Walter J and Trautner TA. 2001. Characterization of the type IV restriction modification system BspLU11III from *Bacillus sp*. LU11. Nucleic Acids Res 29:4691–4698.

Lesser DR, Kurpiewski MR and Jen-Jacobson L. 1990. The energetic basis of specificity in the Eco RI endonuclease–DNA interaction. Science 250:776–786.

Lin TL, Shun CT, Chang KC and Wang JT. 2004. Isolation and characterization of a HpyC1I restriction-modification system in *Helicobacter pylori*. J Biol Chem 279:11156–11162.

Løbner-Olesen A, Skovgaard O and Marinus MG. 2005. Dam methylation: coordinating cellular processes. Curr Opin Microbiol 8:154–160.

Looney MC, Moran LS, Jack WE, Feehery GR, Benner JS, Slatko BE and Wilson GG. 1989. Nucleotide sequence of the FokI restriction-modification system: separate strand-specificity domains in the methyltransferase. Gene 80:193–208.

Low DA, Weyand NJ and Mahan MJ. 2001. Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. Infect Immun 69:7197–7204.

Lusk JE, Williams RJ and Kennedy EP. 1968. Magnesium and the growth of *Escherichia coli*. J Biol Chem 243:2618–2624.

Malone T, Blumenthal RM 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.

Malygin EG, Evdokimov AA, Zinoviev VV, Ovechkina LG, Lindstrom WM, Reich NO, Schlagman SL and Hattman S. 2001. A dual role for substrate S-adenosyl-L-methionine in the methylation reaction with bacteriophage T4 Dam DNA-[N6-adenine]-methyltransferase. Nucleic Acids Res 29:2361–2369.

Malygin EG, Sclavi B, Zinoviev VV, Evdokimov AA, Hattman S and Buckle M. 2004. Bacteriophage T4Dam DNA-(adenine-N(6)-methyltransferase. Comparison of pre-steady state and single turnover methylation of 40-mer duplexes containing two (un) modified target sites. J Biol Chem 279:50012–50018.

Malygin EG, Evdokimov AA and Hattman S. 2009. Dimeric/oligomeric DNA methyltransferases: an unfinished story. Biol Chem 390:835–844.

Marinus MG and Casadesus J. 2009. Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS Microbiol Rev 33:488–503.

Marks P, McGeehan J, Wilson G, Errington N and Kneale G. 2003. Purification and characterisation of a novel DNA methyltransferase, M.AhdI. Nucleic Acids Res 31:2803–2810.

McCleland SE and Sczcelkun MD. 2004. The type I and III restriction endonucleases: structural elements in molecular motors that process DNA, pp. 111–135. In: Pingoud A., ed., Restriction Endonucleases, Berlin: Springer.

McClelland M. 1981. The effect of sequence specific DNA methylation on restriction endonuclease cleavage. Nucleic Acids Res 9:5859–5866.

McClelland M, Nelson M and Cantor CR. 1985. Purification of *Mbo*II methylase (GAAGmA) from *Moraxella bovis*: site specific cleavage of DNA at nine and ten base pair sequences. Nucleic Acids Res 13:7171–7182.

McKane M and Milkman R. 1995. Transduction, restriction and recombination patterns in *Escherichia coli*. Genetics 139:35–43.

Meisel A, Kruger DH and Bickle TA. 1991. M.EcoP15 methylates the second adenine in its recognition sequence. Nucleic Acids Res 19:3997.

Meisel A, Bickle TA, Kruger DH and Schroeder C. 1992. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. Nature 355:467–469.

Meisel A, Mackeldanz P, Bickle TA, Kruger DH and Schroeder C. 1995. Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. EMBO J 14:2958–2966.

Merkiene E, Vilkaitis G and Klimasauskas S. 1998. A pair of single-strand and double-strand DNA cytosine-N4 methyltransferases from *Bacillus centrosporus*. Biol Chem 379:569–571.

Mernagh DR and Kneale GG. 1996. High resolution footprinting of a type I methyltransferase reveals a large structural distortion within the DNA recognition site. Nucleic Acids Res 24:4853–4858.

Mernagh DR, Reynolds LA and Kneale GG. 1997. DNA binding and subunit interactions in the type I methyltransferase M.EcoR124I. Nucleic Acids Res 25:987–991.

Modrich P. 1982. Studies on sequence recognition by type II restriction and modification enzymes. CRC Crit Rev Biochem 13:287–323.

Morgan RD, Bhatia TK, Lovasco L and Davis TB. 2008. MmeI: a minimal Type II restriction-modification system that only modifies one DNA strand for host protection. Nucleic Acids Res 36:6558–6570.

Morgan RD, Dwinell EA, Bhatia TK, Lang EM and Luyten YA. 2009. The MmeI family: type II restriction-modification enzymes that employ single-strand modification for host protection. Nucleic Acids Res 37:5208–5221.

Morita R, Ishikawa H, Nakagawa N, Kuramitsu S and Masui R. 2008. Crystal structure of a putative DNA methylase TTHA0409 from *Thermus thermophilus* HB8. Proteins 73:259–264.

Mruk I, Cichowicz M and Kaczorowski T. 2003. Characterization of the LlaCI methyltransferase from *Lactococcus lactis* subsp. cremoris W15 provides new insights into the biology of type II restriction-modification systems. Microbiology 149:3331–3341.

Naito T, Kusano K and Kobayashi I. 1995. Selfish behavior of restriction-modification systems. Science 267:897–899.

Nakonieczna J, Kaczorowski T, Obarska-Kosinska A and Bujnicki JM. 2009. Functional analysis of MmeI from methanol utilizer *Methylophilus methylotrophus*, a subtype IIC restriction-modification enzyme related to type I enzymes. Appl Environ Microbiol 75:212–223.

Neely RK, Daujotyte D, Grazulis S, Magennis SW, Dryden DT, Klimasauskas S and Jones AC. 2005. Time-resolved fluorescence of 2-aminopurine as a probe of base flipping in M.HhaI-DNA complexes. Nucleic Acids Res 33:6953–6960.

Obarska A, Blundell A, Feder M, Vejsadova S, Sisakova E, Weiserova M, Bujnicki JM and Firman K. 2006. Structural model for the multisubunit Type IC restriction-modification DNA methyltransferase M.EcoR124I in complex with DNA. Nucleic Acids Res 34:1992–2005.

Osipiuk J, Walsh MA and Joachimiak A. 2003. Crystal structure of MboIIA methyltransferase. Nucleic Acids Res 31:5440–5448.

Patel J, Taylor I, Dutta CF, Kneale G and Firman K. 1992. High-level expression of the cloned genes encoding the subunits of and intact DNA methyltransferase, M.EcoR124. Gene 112:21–27.

Piekarowicz A, Golaszewska M, Sunday AO, Siwinska M and Stein DC. 1999. The HaeIV restriction modification system of *Haemophilus aegyptius* is encoded by a single polypeptide. J Mol Biol 293:1055–1065.

Pingoud A, Fuxreiter M, Pingoud V and Wende W. 2005. Type II restriction endonucleases: structure and mechanism. Cell Mol Life Sci 62:685–707.

Posfai J, Bhagwat AS, Posfai G and Roberts RJ. 1989. Predictive motifs derived from cytosine methyltransferases. Nucleic Acids Res 17:2421–2435.

Powell LM and Murray NE. 1995. S-adenosyl methionine alters the DNA contacts of the EcoKI methyltransferase. Nucleic Acids Res 23:967–974.

Powell LM, Connolly BA and Dryden DT. 1998. The DNA binding characteristics of the trimeric EcoKI methyltransferase and its partially assembled dimeric form determined by fluorescence polarisation and DNA footprinting. J Mol Biol 283:947–961.

Powell LM, Lejeune E, Hussain FS, Cronshaw AD, Kelly SM, Price NC and Dryden DT. 2003. Assembly of EcoKI DNA methyltransferase requires the C-terminal region of the HsdM modification subunit. Biophys Chem 103:129–137.

Raleigh EA and Brooks JE. 1998. Restriction modification systems: where they are and what they do. p. 78–92. In F. J. De Bruijn, J. R. Lupski, and G. M. Weinstock (ed.), Bacterial genomes. Chapman & Hall, New York, N.Y.

Rao DN, Page MG and Bickle TA. 1989. Cloning, over-expression and the catalytic properties of the EcoP15 modification methylase from *Escherichia coli*. J Mol Biol 209:599–606.

Rao DN, Saha S and Krishnamurthy V. 2000. ATP-dependent restriction enzymes. Prog Nucleic Acid Res Mol Biol 64:1–63.

Rechkunova NI, Zinoviev VV, Malygin EG, Gorbunov, Yu A, Popov SG, Nesterenko VF, Buryanov Ya I and Bayev AA. 1987. Dimerization of Eco dam methylase induced by the ODN substrate. Biopolymers Cell (Russian) 3:152–154.

Redaschi N and Bickle TA. 1996. Posttranscriptional regulation of EcoP1I and EcoP15I restriction activity. J Mol Biol 257:790–803.

Reddy YV and Rao DN. 1998. Probing the role of cysteine residues in the EcoP15I DNA methyltransferase. J Biol Chem 273:23866–23876.

Reddy YV and Rao DN. 2000. Binding of *Eco*P15I DNA methyltransferase to DNA reveals a large structural distortion within the recognition sequence. J Mol Biol 298:597–610.

Reich NO and Mashhoon N. 1993. Presteady state kinetics of an S-adenosylmethionine-dependent enzyme. Evidence for a unique binding orientation requirement for EcoRI DNA methyltransferase. J Biol Chem 268:9191–9193.

Roberts RJ and Cheng X. 1998. Base flipping. Annu Rev Biochem 67:181–198.

Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev S, Dryden DT, Dybvig K, Firman K, Gromova ES, Gumport RI, Halford SE, Hattman S, Heitman J, Hornby DP, Janulaitis A, Jeltsch A, Josephsen J, Kiss A, Klaenhammer TR, Kobayashi I, Kong H, Kruger DH, Lacks S, Marinus MG, Miyahara M, Morgan RD, Murray NE, Nagaraja V, Piekarowicz A, Pingoud A, Raleigh E, Rao DN, Reich N, Repin VE, Selker EU, Shaw PC, Stein DC, Stoddard BL, Szybalski W, Trautner TA, Van Etten JL, Vitor JM, Wilson GG, Xu SY. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic Acids Res 31:1805–1812.

Roberts RJ, Vincze T, Posfai J and Macelis D. 2007. REBASE – enzymes and genes for DNA restriction and modification. Nucleic Acids Res 35:D269–270.

Ryan KA and Lo RY. 1999. Characterization of a CACAG pentanucleotide repeat in *Pasteurella haemolytica* and its possible role in modulation of a novel type III restriction-modification system. Nucleic Acids Res 27:1505–1511.

Sapranauskas R, Sasnauskas G, Lagunavicius A, Vilkaitis G, Lubys A and Siksnys V. 2000. Novel subtype of type IIs restriction enzymes. *Bfi*I endonuclease exhibits similarities to the EDTA-resistant nuclease Nuc of *Salmonella typhimurium*. J Biol Chem 275:30878–30885.

Saunders NJ, Jeffries AC, Peden JF, Hood DW, Tettelin H, Rappuoli R and Moxon ER. 2000. Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. Mol Microbiol 37:207–215.

Seib KL, Peak IR and Jennings MP. 2002. Phase variable restriction-modification systems in *Moraxella catarrhalis*. FEMS Immunol Med Microbiol 32:159–165.

Serva S, Weinhold E, Roberts RJ and Klimasauskas S. 1998. Chemical display of thymine residues flipped out by DNA methyltransferases. Nucleic Acids Res 26:3473–3479.

Shier VK, Hancey CJ and Benkovic SJ. 2001. Identification of the active oligomeric state of an essential adenine DNA methyltransferase from *Caulobacter crescentus*. J Biol Chem 276:14744–14751.

Sistla S and Rao DN. 2004. S-Adenosyl-L-methionine-dependent restriction enzymes. Crit Rev Biochem Mol Biol 39:1–19.

Sistla S, Krishnamurthy V, Rao, DN. 2004. Single-stranded DNA binding and methylation by EcoP1I DNA methyltransferase. Biochem Biophys Res Commun 314:159–165.

Smith DR. 1996. Restriction endonuclease digestion of DNA. In : Basic DNA and RNA protocols. A.J. Harwood (ed) p. 11-15. Humana Press, Totowa, NJ.

Srikhanta YN, Maguire TL, Stacey KJ, Grimmond SM and Jennings MP. 2005. The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. Proc Natl Acad Sci USA 102:5547–5551.

Srikhanta YN, Dowideit SJ, Edwards JL, Falsetta ML, Wu HJ, Harrison OB, Fox KL, Seib KL, Maguire TL, Wang AH, Maiden MC, Grimmond SM, Apicella MA and Jennings MP. 2009. Phasevarions mediate random switching of gene expression in pathogenic *Neisseria*. PLoS Pathog 5:1–22.

Su TJ, Connolly BA, Darlington C, Mallin R and Dryden DT. 2004. Unusual 2-aminopurine fluorescence from a complex of DNA and the EcoKI methyltransferase. Nucleic Acids Res 32:2223–2230.

Su TJ, Tock MR, Egelhaaf SU, Poon WC and Dryden DT. 2005. DNA bending by M.EcoKI methyltransferase is coupled to nucleotide flipping. Nucleic Acids Res 33:3235–3244.

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

Sugisaki H, Yamamoto K and Takanami M. 1991. The *Hga*I restriction-modification system contains two cytosine methylase genes responsible for modification of different DNA strands. J Biol Chem 266:13952–13957.

Sunita S, Tkaczuk KL, Purta E, Kasprzak JM, Douthwaite S, Bujnicki JM and Sivaraman J. 2008. Crystal structure of the *Escherichia coli* 23S rRNA:m5C methyltransferase RlmI (YccW) reveals evolutionary links between RNA modification enzymes. J Mol Biol 383:652–666.

Suri B and Bickle TA. 1985. EcoA: the first member of a new family of type I restriction modification systems. Gene organization and enzymatic activities. J Mol Biol 186:77–85.

Suri B, Nagaraja V and Bickle TA. 1984. Bacterial DNA modification. Curr Top Microbiol Immunol 108:1–9.

Svadbina IV, Matvienko NN, Zheleznaya LA and Matvienko NI. 2005. Location of the bases modified by M.BcoKIA and M.BcoKIB methylases in the sequence 5-CTCTTC-3/5-GAAGAG-3. Biochemistry (Moscow) 70:1126–1128.

Taylor I, Patel J, Firman K and Kneale G. 1992. Purification and biochemical characterisation of the EcoR124 type I modification methylase. Nucleic Acids Res 20:179–186.

Taylor I, Watts D and Kneale G. 1993. Substrate recognition and selectivity in the type IC DNA modification methylase M.EcoR124I. Nucleic Acids Res 21:4929–4935.

Taylor IA, Webb M and Kneale GG. 1996. Surface labelling of the type I methyltransferase M.EcoR124I reveals lysine residues critical for DNA binding. J Mol Biol 258:62–73.

Thielking V, Alves J, Fliess A, Maass G and Pingoud A. 1990. Accuracy of the EcoRI restriction endonuclease: binding and cleavage studies with oligodeoxynucleotide substrates containing degenerate recognition sequences. Biochemistry 29:4682–4691.

Thomas CB and Gumport RI. 2006. Dimerization of the bacterial RsrI N6-adenine DNA methyltransferase. Nucleic Acids Res 34:806–815.

Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD,

Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature 388:539–547.

Tucholski J, Zmijewski JW and Podhajska AJ. 1998. Two intertwined methylation activities of the *Mme*I restriction-modification class-IIS system from *Methylophilus methylotrophus*. Gene 223:293–302.

Tuzikov FV, Zinoviev VV, Jashina LN, Vavilin VI, Gorbunov Yu A, Popov SG, Buryanov Ya I and Bayev AA. 1986. Application of the small-angle X-ray scattering for the investigation of substrate-induced changes of methylase state. Mol Biol (Moscow) 20:1002–1007.

Van Etten JL. 2003. Unusual life style of giant chlorella viruses. Annu Rev Genet 37:153–195.

Vilkaitis G, Dong A, Weinhold E, Cheng X and Klimasauskas S. 2000. Functional roles of the conserved threonine 250 in the target recognition domain of HhaI DNA methyltransferase. J Biol Chem 275:38722–38730.

Vilkaitis G, Lubys A, Merkiene E, Timinskas A, Janulaitis A and Klimasauskas S. 2002. Circular permutation of DNA cytosine-N4 methyltransferases: in vivo coexistence in the BcnI system and in vitro probing by hybrid formation. Nucleic Acids Res 30:1547–1557.

Vitkute J, Stankevicius K, Tamulaitiene G, Maneliene Z, Timinskas A, Berg DE and Janulaitis A. 2001. Specificities of eleven different DNA methyltransferases of *Helicobacter pylori* strain 26695. J Bacteriol 183:443–450.

Webb M, Taylor IA, Firman K and Kneale GG. 1995. Probing the domain structure of the type IC DNA methyltransferase M.EcoR124I by limited proteolysis. J Mol Biol 250:181–190.

Willcock DF, Dryden DT and Murray NE. 1994. A mutational analysis of the two motifs common to adenine methyltransferases. EMBO J 13:3902–3908.

Wion D and Casadesus J. 2006. N6-methyl-adenine: an epigenetic signal for DNA–protein interactions. Nat Rev Microbiol 4:183–192.

Xu Q, Morgan RD, Roberts RJ and Blaser MJ. 2000. Identification of type II restriction and modification systems in *Helicobacter pylori* reveals their substantial diversity among strains. Proc Natl Acad Sci USA 97:9671–9676.

Yoo HY, Noshari J and Lapeyre JN. 1987. Subunit and functional size of human placental DNA methyltransferase involved in de novo and maintenance methylation. J Biol Chem 262:8066–8070.

Zylicz-Stachula A, Bujnicki JM and Skowron PM. 2009. Cloning and analysis of a bifunctional methyltransferase/restriction endonuclease TspGWI, the prototype of a *Thermus* sp. enzyme family. BMC Mol Biol 10:52.

Editor: Michael M. Cox