Nucleic Acids Research, 1992, Vol. 20, No. 22 ***6091–6096***

In vivo specificity of *Eco*RI DNA methyltransferase

Dean W.Smith, Scott W.Crowder and Norbert O.Reich*
Department of Chemistry, University of California, Santa Barbara, CA 93106, USA

Received March 4, 1992; Revised and Accepted October 16, 1992

ABSTRACT

The *Eco*RI adenine DNA methyltransferase forms part of a bacterial restriction/modification system; the methyltransferase modifies the second adenine within the canonical site GAATTC, thereby preventing the *Eco*RI endonuclease from cleaving this site. We show that five noncanonical *Eco*RI sites (TAATTC, CAATTC, GTATTC, GGATTC and GAGTTC) are not methylated *in vivo* under conditions when the canonical site is methylated. Only when the methyltransferase is overexpressed is partial *in vivo* methylation of the five sites detected. Our results suggest that the methyltransferase does not protect host DNA against potential endonuclease-mediated cleavage at noncanonical sites. Our related *in vitro* analysis of the methyltransferase reveals a low level of sequence-discrimination. We propose that the high *in vivo* specificity may be due to the active removal of methylated sequences by DNA repair enzymes (*J. Bacteriology* (1987), 169 3243–3250).

INTRODUCTION

The bacterial *Eco*RI restriction/modification (R/M) system consists of two enzymes that recognize the sequence 5'GAATTC3' (1,2). The endonuclease (ENase) is a magnesium-dependent dimer and cleaves both DNA strands between the guanine and the adjacent adenine. This results in the 'inactivation' of foreign (e.g., viral) DNA and thus potentially confers an advantage to the host cell over cells lacking this function. Protection of the host DNA from attack by the ENase is provided by the *Eco*RI DNA methyltransferase (Mtase), a monomeric, S-adenosylmethionine (AdoMet)-dependent enzyme which delivers a methyl group to the exocyclic amino of the internal adenine.

Type II R/M enzymes, of which the *Eco*RI system is a well studied example, provide attractive models for investigating sequence-specific DNA recognition and modification. Hundreds of such enzymes with a large variety of DNA recognition sequences (3) are an essential part of modern recombinant DNA manipulations (4); extensive efforts have been directed at the (re)design of these enzymes (5–9) and of other reagents with novel sequence-specificities (10–12). The availability of the corresponding genes for numerous Type II enzymes has facilitated detailed structural (13) and mechanistic (2, 14–16) studies. Protein sequence comparison of the *Eco*RI ENase and Mtase reveal no similarity (17); the ENases and Mtases from other Type II systems are likewise dissimiliar. Furthermore, there exists only limited protein sequence homology between various adenine DNA Mtases, or with the related cytosine C5 DNA Mtases (18). The large number of sequenced genes and the diverse recognition sites provide the basis for extensive evolutionary comparisons. Finally, the details of how the specificities of the Mtase and ENase affect each other is of interest.

Past investigations of the sequence specificity of Type II systems have focused on the restriction ENases. These enzymes show a high degree of *in vitro* specificity (14, 19), although their *in vivo* specificity remains obscure (20). In contrast, little is known about either the *in vitro* or *in vivo* specificity for any Type II DNA Mtase. Von Hippel and coworkers provided early evidence suggesting *Eco*RI DNA Mtase methylates numerous noncanonical sites *in vivo* (21). In contrast, Halford and coworkers recently demonstrated that limited methylation at noncanonical sites occurs *in vivo* with the *Eco*RV DNA Mtase and little overlap exits between noncanonical sites methylated in vivo and noncanonical sites attacked by the ENase in vitro (19).

We have re-examined the *in vivo* specificity of the *Eco*RI DNA Mtase to enable direct comparisons with our *in vitro* specificity analysis (22). Also, a full understanding of how the ENase and Mtase sequence-specificities affect each other requires information about the sequences which are preferred under *in vitro*, and ultimately *in vivo* conditions by both enzymes. Finally, although the *in vivo* specificity differences mentioned for *Eco*RI (21) and *Eco*RV (19) DNA Mtases might be explained by real differences in these systems, the different analytical methods used in the two studies may also be responsible. Here we used these previous approaches in addition to a third method to help identify sites of *in vivo* methylation.

MATERIALS

*Eco*RI DNA Mtase was isolated from *E.coli* strain 294-4 containing plasmid pPG440 (2). Oligonucleotides were prepared using a Biosearch synthesizer and phosphoramidite solid-phase chemistry and purified by reversed-phase HPLC (2). Liquiscint scintillation fluid was purchased from National Diagnostics and radioactivity determined with a Beckman scintillation counter, Model LS1700. DE81 ion exchange filter papers were from Whatman, and a Branson Sonifier (Model 450) was used to release cellular Mtase protein. Plasmid DNA was isolated by

* To whom correspondence should be addressed

Qiagen columns. En3Hance and S-[methyl^{3}H] adenosyl methionine (78.9 Ci/mM) were purchased from New England Nuclear. Isopropyl-beta-D-thiogalactopyranoside (IPTG) was purchased from Research Organics and *Eco*RI endonuclease was kindly provided by Dr. Paul Hager of University of California at San Francisco. Bovine serum albumin (BSA), dithiothreitol (DTT), and proteinase K were purchased from Sigma. X-ray film was from Kodak and Fuji.

METHODS

Mtase activity assay

The activity of the Mtase was monitored by incorporation of radioactivity (C^3H_3) into DNA (2). Samples (50 μL) contained (standard conditions) 100 mM Tris, pH 8.0, 10 mM EDTA, 0.2 mg/ml BSA, 1.0 mM DTT, 0.78 nM DNA, 0.8 nM Mtase and 1.0 μM AdoMet. Star conditions contained 100 mM Tris, pH 8.5, 0.2 mg/ml BSA, 1.0 mM DTT, 100 nM Mtase, and 20% glycerol (23). Based on our previous kinetic characterization of the Mtase, the DNA and AdoMet concentrations are in excess of their respective K_m values (2). The reactions proceeded at 37°C for the indicated times, 40 μL was removed and the reaction stopped by application onto a DE-81 filter and processed as described (2). The moles of product formed were determined from scintillation counting of the DE-81 filters. Mtase concentrations were determined using the following extinction coefficient, $E_{280nm}^{1\%} = 10.8$ (24).

Plasmid construction

Plasmid pPG440W was constructed from plasmid pPG440 (4421 basepairs) by digestion with *Hind*III endonuclease which cuts at positions 2735 and 4419, thereby removing the entire Mtase gene and most of the ENase gene. The 2737 bp fragment was isolated, treated with DNA ligase and transformed into strain D1210 (overproduces the lac repressor). Plasmid size comparisons and restriction analysis were used to confirm the plasmid construction.

DNA preparation

E.coli containing plasmid pPG440 or pPG440W were grown in YT media (4) in the presence of 50 μg/ml ampicillin. At $OD_{600nm} = 0.35$, one sample of the pPG440 containing strain was induced with 1.0 mM IPTG, and is designated pPG440$^+$. All three strains (pPG440W, pPG440$^+$, and pPG440 without induction, pPG440$^-$) were grown to stationary phase, $OD_{600nm} = 2.5$, and cells harvested by centrifugation. One portion of each sample was used for DNA isolation and another for isolation of protein (see below). Plasmid DNA was isolated by Qiagen column chromatography and subsequently digested with *Hind*III restriction endonuclease. The resulting 2737 bp fragment was purified via electro-elution, gel excision, and finally G-50 (Pharmacia) chromatography. The DNA was quantified spectrophotometrically by measuring the absorbance at 260 nm and assuming 50 μg/ml of duplex DNA is equivalent to an absorbance of 1.0 (4). Further confirmation of the DNA concentration was obtained by comparative densitometric analysis of ethidium bromide stained gels.

Protein preparation

Approximately 0.1 gram of the cell pellets described above were sonified in 1 ml of 10 mM potassium phosphate buffer (pH 7.0), 200 mM NaCl, 1 mM EDTA, 5 mM beta mercaptoethanol. Cell debris was removed by centrifugation at 13,000 g at 4°C for 60 minutes. Supernatants were assayed under standard conditions for Mtase activity using 2 μM 14mer.

Autoradiography

The 2737 bp restriction fragments isolated from plasmids pPG440W, pPG440$^+$ and pPG440$^-$ were methylated under standard or star conditions (23) and the reactions stopped at the indicated times by freezing in a dry ice/ethanol bath. Aliquots (50 μL) were then incubated at 65°C for 10 min., followed by digestion with *Mlu*I and *Dde*I restriction endonucleases for 1 hr at 37°C. The digested samples were treated with proteinase K for 30 min. at 37°C and 1 μg of predigested non-labeled fragment added to facilitate viewing of bands with ethidium bromide. The samples were applied to a 5% polyacrylamide gel and electrophoresed at 175 volts for 4 hours, the gels stained with ethidium bromide, photographed, treated with En3Hance and exposed to X-ray film for 7 days with an intensifying screen at −70°C.

*Eco*RI endonuclease assay

The three plasmids (pPG440W, pPG440$^+$ and pPG440$^-$) were submitted to digestion with *Eco*RI endonuclease under standard (supercoiled plasmids) and star (2737 bp fragment) conditions. For the standard assay a volume of 75 μL containing 50 mM NaCl, 100 mM Tris pH 7.5, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol, 100 μg/ml BSA, 41.6 nM DNA and 2 nM endonuclease was incubated at 37°C for 25 minutes and analyzed on a 1% agarose gel. The star assay conditions included: 5.7 mM NaCl, 7.7 mM $MgCl_2$, 95 mM Tris 8.5, 19% glycerol, 41.6 nM DNA and 160 nM endonuclease [23). Samples (4 μL) were removed at the indicated times and submitted to electrophoresis (4% PAGE) at 200 volts for 1.5 hours.

RESULTS AND DISCUSSION

Our goal was to determine the level of non-canonical site DNA methylation which is mediated by the *Eco*RI DNA Mtase *in vivo*. Our recent *in vitro* investigation of the Mtase showed that some sites have specificity constants only five fold lower than the canonical site (22); this high level of non-canonical site methylation might be expected to cause problems *in vivo*. The approaches used in our investigation are outlined in Figure 1B and rely on the *in vitro* treatment of a restriction fragment isolated from *E.coli* which lack the Mtase gene (pPG440W), have the Mtase gene but expression is not induced (pPG440$^-$), and have the Mtase gene and include induction (pPG440$^+$).

The subset of sites probed with the three assays includes *single* basepair substitutions within the canonical site. These are termed 'star' sites because the ENase and Mtase are known to modify such sites under altered (star) *in vitro* conditions. We refer to noncanonical sites modified under star conditions as star sites and we limit our discussion to star sites with single basepair substitutions. The essence of the three assays (Figure 1b) is that the *in vivo* specificity of the Mtase can be deduced by isolating DNA from bacteria exposed to the Mtase under *in vivo* conditions, and treating the DNA with the Mtase and ENase under star conditions. The ability of the Mtase and ENase to methylate and cleave star sites under *in vitro* conditions indicates that that site was not methylated *in vivo*. The *Hind*III fragment used throughout our study (Figure 1) contains five of the nine potential star sites (one appears twice), and therefore our assays do not

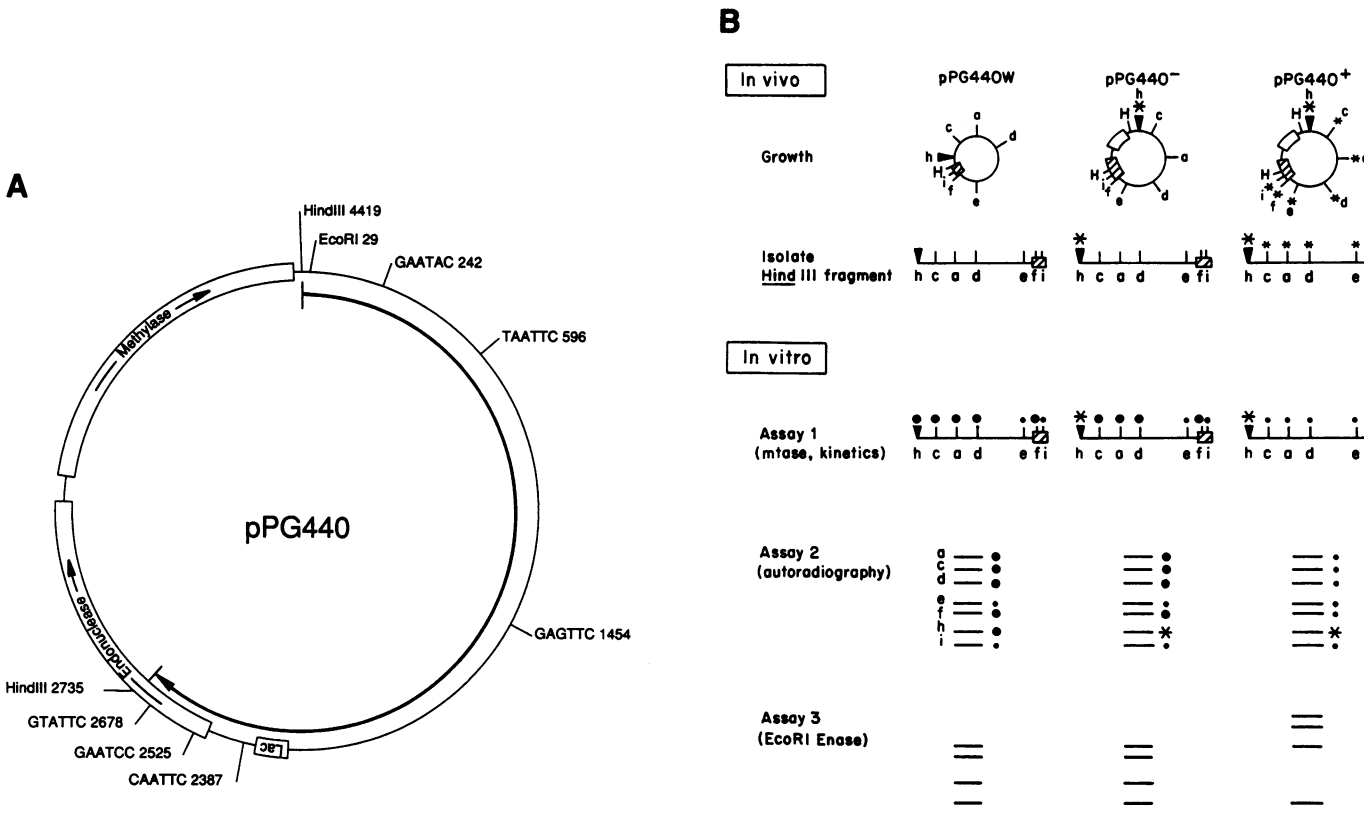

Figure 1. A. Plasmid pPG440 [4421 base pairs (26)] showing the placement of the *Eco*RI ENase and Mtase genes, the LacUV5 promoter, and the *Hind*III sites. Plasmid pPG440W was constructed from pPG440 by removing the *Hind*III fragment containing the Mtase gene (see Methods). The arrow within the plasmid shows the 2737 bp *Hind*III fragment analyzed in the three assays described in 1B. **B.** Summary of experimental strategies and results. The two plasmid constructs have the *Hind*III fragment containing the *Eco*RI endonuclease gene (cross hatched) and a portion of the *Eco*RI Mtase gene (open box) removed (pPG440W) or intact (pPG440). Bacteria containing plasmid pPG440 were grown with induction (pPG440$^+$) and produced forty fold greater Mtase activity than those without induction (pPG440$^-$). H indicates the *Hind*III sites in pPG440, and the single *Hind*III site in pPG440W (Figure 1A). All *in vitro* assays were done with the *Hind*III fragment isolated from each plasmid. Lower case c, a, d, e, f, and i refer to star sites within restriction fragments generated by digestion of the *Hind*III fragment with *Dde*I and *Mlu*I endonucleases (see Figures 3 and 4). The large asterisk in the intact plasmids (pPG440$^-$ and pPG440$^+$) refers to *in vivo* methylation at the canonical site (h, arrowhead). The first *in vitro* assay uses the *Eco*RI DNA Mtase under star conditions and tritiated cofactor to determine the extent to which each *Hind*III derived fragment can be methylated. Large and small dots indicated minor and extensive label incorporation respectively into the intact fragment at sites identified with assay 2. The pPG440W and pPG440$^-$ derived fragments show identical label incorporation except for the single canonical site (h) which was methylated *in vivo* with pPG440$^-$ (large asterisk). The lower level incorporation observed with pPG440$^+$ (smaller dots) is a consequence of the prior *in vivo* methylation (small asterisk). Assay 2 submits the *Hind*III fragments described for assay 1 to further digestion (*Dde*I and *Mlu*I) and identifies the restriction subfragments by PAGE and autoradiography. Subfragment h shows *in vitro* incorporation only with pPG440W, while all other subfragments show identical levels of incorporation for pPG440W and pPG440$^-$. The autoradiogram with pPG440$^+$ shows uniformly lower levels of incorporation in all subfragments, indicating a low level of *in vivo* methylation. The third assay uses the *Hind*III derived fragments and *Eco*RI endonuclease under star conditions: prior *in vivo* methylation of star sites results in no *in vitro* cleavage at that site. Plasmids pPG440W and pPG440$^-$ show identical cleavage patterns indicating similar *in vivo* protection of star sites, while pPG440$^+$ shows minor differences.

address *in vivo* methylation at 5'AAATTC3', 5'GCATTC3', 5'GATTTC3' and 5'GACTTC3'.

The *in vitro* kinetic results in Figure 2 compare the ability of the Mtase to methylate the three 2737 bp *Hind*III fragments under standard and reduced specificity conditions (star). Treatment of the pPG440W-derived fragment under standard conditions shows incorporation of two methyl groups per DNA molecule, consistent with the single canonical site (26) becoming doubly methylated. The lack of any detectable label incorporation into fragments isolated from pPG440$^+$ and pPG440$^-$ confirms the prior *in vivo* methylation of the single canonical site. The star conditions used in Figure 2 are known to cause the Mtase (and the ENase, see below) to modify a subset of noncanonical sites (21, 23, 25). The data in Figure 2 shows that star methylation of the pPG440W-derived *Hind*III fragment results in the incorporation of six additional methyl groups; this could come from double methylation at three noncanonical sites or partial methylation at any larger number of sites. Since the rate of label incorporation approaches zero after 120 minutes (Figure 2), a small number of noncanonical sites have most likely become completely methylated. After correction for the lack of *in vitro* canonical site methylation (see above), star methylation of the pPG440$^-$ derived fragment shows the same extent of methylation as the pPG440W derived fragment. Only when pPG440$^+$ is used, and the level of Mtase is at least forty times that found with pPG440$^-$, is decreased in vitro methylation observed (Figure 2). Even under these conditions of Mtase overexpression, incorporation of three of the six methyl groups per DNA molecule at noncanonical sites is observed. Our kinetic results differ from those of Woodbury et al (21) with the *Eco*RI DNA Mtase using a similar approach. They tested DNA grown in the absence and presence of the Mtase and found extensive

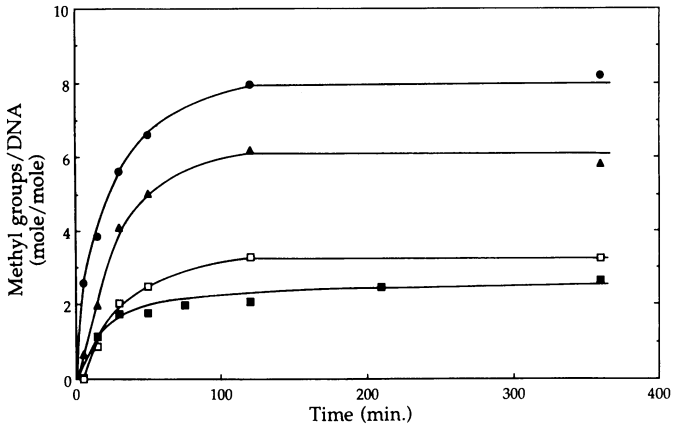

Figure 2. *In vitro* kinetic analysis shows no *in vivo* methylation of noncanonical sites under conditions when the canonical site becomes methylated. *In vitro* methylation of the 2737 bp fragment (Figure 1A,B) derived from plasmids pPG440W, pPG440$^-$ and pPG440$^+$ as detected by radiochemical filter binding assay (see Methods and Figure 1B for description of plasmids). Methylation of pPG440W-derived fragment under standard conditions (filled squares) shows that two moles of methyl groups are incorporated per mole of DNA, consistent with double methylation of the single canonical site. No incorporation into pPG440$^-$ and pPG440$^+$ derived fragments under these conditions was observed (data not shown). pPG440W treated under star conditions (see Methods) is shown by the filled circles and indicates eight moles (six noncanonical) of methyl groups per mole of DNA. pPG440$^-$ (filled triangles) under star conditions shows the same level of non-canonical site methylation (six moles per mole of DNA) as pPG440W. Some *in vivo* methylation does occur with pPG440$^+$ since the level of *in vitro* incorporation into the derived fragment (open squares) is less than that for pPG440W.

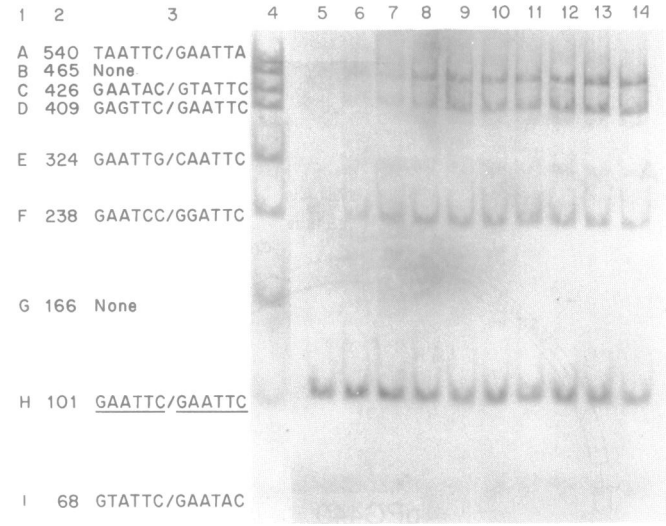

Figure 3. Identification of possible noncanonical methylation sites within pPG440W. Samples identical to those described in Figure 2 for pPG440W under star conditions (filled circles) were digested with *Mlu*I and *Dde*I restriction endonucleases, submitted to 5% PAGE analysis, stained with ethidium bromide and label incorporation detected with autoradiography (see Methods). Lane 1, the letter of each subfragment described in the text. Lane 2, the predicted subfragments and their size in base-pairs. Lane 3, the predicted noncanonical sequence within that subfragment. Lane 4, the ethidium bromide stained gel of the actual subfragments. Lanes 5–14, the autoradiogram of samples exposed to the Mtase for 5, 10, 15, 20, 30, 45, 60, 80, 100, and 150 minutes (see Methods for experimental details). Subfragment h contains the canonical site and is fully methylated within 5 minutes. Subfragments a, c, and f become methylated over the next 60 minutes while subfragments d, e and i show very small amounts of incorporation at 150 minutes. Subfragments b and g show no detectable label incorporation.

in vivo methylation. However, they compared viral, chromosomal and plasmid DNA, thus making internal comparisons and comparison with our work obscure. Moreover, several concerns arise from their analysis. Star assays are done at pH 8.5–9.0 (23) and AdoMet is known to undergo rapid hydrolytic cleavage at this pH (27). The 50 hour incubation times used by Woodbury et al (21) may therefore have led to extensive AdoMet degradation. Such long incubation times are also known to lead to Mtase inactivation, making kinetic analysis difficult (2). Von Hippel and coworkers used *E.coli* RY13 to determine the level of *in vivo* methylation; although direct comparison with the strain harboring pPG440$^-$ is not possible, RY13 does not overproduce the Mtase.

Further fragmentation of the *Hind*III fragments used in Figure 2 faciliated our autoradiographic analysis of the likely sequences which are methylated *in vitro* (and thus not methylated *in vivo*). Figure 3 shows that several fragments contain sequences which become methylated under *in vitro* star conditions. Since the sequence for the entire plasmid is known, hexanucleotides segments within the plasmid and related to the *Eco*RI sequence by single basepair changes can be identified. Thus subfragments a, c and f contain the known star sites 5′TAATTC3′, 5′GTATTC3′, and 5′GGATTC3′ respectively (Figure 3). Subfragments d, e and i (containing sites 5′GAGTTC3′, 5′CAATTC3′ and 5′GTATTC3′) show small amounts of incorporation while subfragments b and g (containing no single-basepair modified sites) show no detectable label incorporation. The fact that subfragments c and i have identical star sequences while only c becomes extensively methylated may stem from differences in the flanking sequences. Based on the intensity of the doubly methylated subfragment h which contains no star sites, label incorporation into a, c and f account for incorporation of five to six methyl groups. Thus, the level of methylation identified with pPG440W in Figure 2 is most likely due to double methylation of the star sites in subfragments a and c, hemimethylation of f and the remaining methylation distributed over d, e, and i.

Application of the strategy described in Figure 3 for pPG440W (no *in vivo* exposure to the Mtase) to pPG440$^-$ (*in vivo* Mtase exposure, no induction) shows that other than the expected canonical site methylation, no significant difference between these plasmids is observed. This provides direct evidence that under noninducing conditions (pPG440$^-$), sufficient Mtase is produced to protect the canonical site, and supports the similarity between the two plasmids observed in the kinetic assay (Figure 2). Further, some level of *in vivo* methylation of noncanonical sites is observed when the Mtase is overexpressed (pPG440$^+$, Figure 4).

The third assay to assess the level of *in vivo* methylation used the ability of the ENase to cleave noncanonical sites under star conditions (Figure 5).The star activity of the ENase has been well characterized (23, 25) and the hierarchy of preferred substitutions in the first three positions of the recognition site is as follows: first position, G>>A>T>>C; second and third position, A>>[G,C]>>T. Of the six star sites within the fragments investigated here, three are predicted to be good sites (TAATTC, position 596; GAGTTC, position 1454; GGATTC, position 2525, see Figures 1A and 1B) and three poor sites (GTATTC,

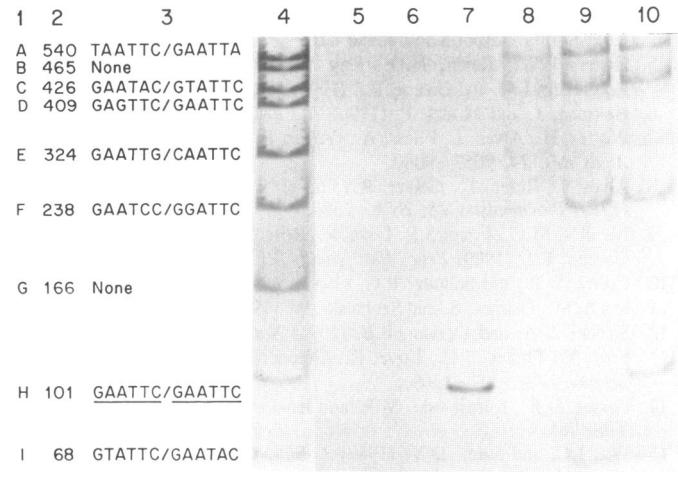

Figure 4. *In vivo* methylation at noncanonical sites analyzed by autoradiography of *in vitro* methylated 2737 bp fragments isolated from pPG440$^-$, pPG440$^+$ and pPG440W. PAGE and autoradiography analysis described in Figure 3. Lanes 1–4 are the same as in Figure 3. Lanes 5–7, the results of methylating pPG440$^+$, pPG440$^-$ and pPG440W respectively, under standard conditions for 2.5 hours (see Methods) followed by digestion with *Dde*I and *Mlu*I and analysis as described in Figure 3. Only the pPG440W-derived DNA shows label incorporation into subfragment h (canonical site) since no *in vivo* incorporation occurred with this substrate. Lanes 8–10, pPG440$^+$, pPG440$^-$ and pPG440W respectively, methylated under star conditions for 2.5 hours. Comparison of lanes 9 and 10 clearly shows that with the exception of subfragment h, these samples have identical *in vitro* methylation patterns. Some decrease in label incorporation into non-canonical sites occurs when the Mtase is overexpressed (pPG440$^+$, lane 8).

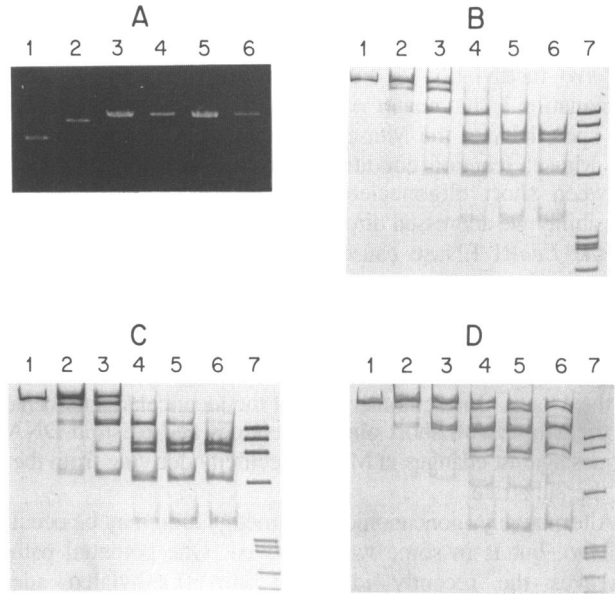

Figure 5. A. Supercoiled pPG440W, pPG440$^-$, and pPG440$^+$ were treated with *Eco*RI ENase for 25 minutes under standard conditions described in Methods. Samples were run on a 1% agarose gel, visualized with ethidium bromide staining and photographed. Lanes 1 and 2, pPG440W untreated and treated; lanes 3 and 4, pPG440$^-$ untreated and treated; lanes 5 and 6, pPG440$^+$ untreated and treated. **B** pPG440W-derivedfragment digested under star conditions and submitted to 4% PAGE analysis (see Methods). Times of digestion are 0, 10, 30, 60, 180, and 600 minutes, corresponding to lanes 1 through 6. Lane 7 shows size standards: 1353 bp, 1078 bp, 872 bp, 603 bp, 310 bp, 281 bp, 271 bp, 234 bp, 198 bp, 118 bp, 72 bp. *Eco*RI ENase fragments are approximately 930, 860, 570, and 350 bp long. This is consistent with cleavage at positions 596, 1454 and 2387 in the 2737 bp linear described in Figure 1, corresponding to sites TAATTC, GAGTTC and CAATTC respectively. No cleavage is detected at the remaining star sites, GTATTC and GGATTC. Cleavage at the canonical site is not detected with this assay since it occurs at position 29. **C.** pPG440$^-$ derived fragment digested under star conditions (see Methods). Times are as indicated in 5B. Lane 7 shows size standards. **D.** pPG440$^+$ derived fragment digested under star conditions (see Methods). Times are as indicated in 5B. Lane 7 shows size standards. The change in digestion pattern (e.g., lane 6) relative to that seen in Figure 5B is largely due to the lack of cleavage at position 2141 corresponding to the sequence TAATTC.

positions 242 and 2678; CAATTC, position 2387, see Figures 1A and 1B). Figure 5B shows that digestion of the pPG440W-derived *Hind*III fragment creates four subfragments approximately 930, 860, 570, and 350 bp in length, while cleavage at the canonical site (position 29) is not detectable with this assay. These four subfragments most likely result from cleavage at positions 596 (TAATTC), 1454 (GAGTTC), and 2387 (CAATTC) which create fragments 933, 858, 567, and 348 bp in length. Star cleavage at TAATTC and GAGTTC and the lack of cleavage at GTATTC is consistent with the known star preferences of the ENase (23,25). In contrast, the lack of cleavage at GGATTC, and the observed cleavage at CAATTC are not predicted and may be accounted for by flanking sequence differences or variations in star assay conditions.

The star ENase assay can thus be used to assess the *in vivo* methylation status of star sequences GAGTTC (position 1454) and CAATTC (position 2387) which were not possible with the autoradiography assay presented in Figures 3 and 4. This assay can also be used to confirm the *in vivo* methylation status of TAATTC (position 596), which was shown not to be methylated by our previous assay (Figures 3 and 4). The similar subfragment patterns for pPG440W and pPG440$^-$ (Figure 5B and 5C) confirm our previous results that none of the tested star sites are methylated by the Mtase *in vivo*. Based on the data with pPG440$^+$ (Figure 5D), the Mtase appears to methylate the sequences TAATTC. In a related experiment the pPG440W-derived fragment was methylated with the Mtase under star conditions followed by treatment with the ENase under star conditions; the results are identical to those shown in Figure 5D and therefore suggest that, at least with this ENase-based assay, the Mtase *in vitro* star and *in vivo* specificities are similar (data not shown).

Our results with the *Eco*RI ENase assay are similar to those reported by Halford and coworkers (19) for the *Eco*RV R/M system. Overproduction of the *Eco*RV Mtase, which modifies GATATC, also results in minimal *in vitro* protection against the *Eco*RV ENase under reduced specificity (star) conditions. *In vivo* methylation at some noncanonical sites was observed, similar to our results (Figure 5).

Our results show that the five noncanonical *Eco*RI sites (TAATTC, CAATTC, GTATTC, GGATTC and GAGTTC) probed with three *in vitro* assays are not methylated *in vivo* under conditions when the canonical site is methylated. These sites represent a majority of the noncanonical sites which the Mtase might be expected to modify. In contrast, our *in vitro* specificity analysis using synthetic oligonucleotides indicates the Mtase specificity is decreased only 5 fold to 23,000 fold as a result of single basepair substitutions within the canonical site (22): based on this level of *in vitro* discrimination a significant level of noncanonical site methylation should be occurring *in vivo*. For

example, GGATTC and CAATTC are methylated *in vitro* only 5.5 and 135 fold less respectively than the canonical site yet no *in vivo* methylation at either site was detected. A possible explanation is that our *in vitro* conditions in some way decrease the specificity of the Mtase. This could come from changes in the Mtase, reaction conditions, or conformational distinctions between short oligonucleotides and supercoiled DNA. One possibility we addressed directly was whether the *in vivo* presence of the *Eco*RI ENase caused the Mtase to be more specific. Experiments analogous to those presented in this report were performed using plasmid constructions lacking the ENase gene. The results were identical to those reported here (data not shown), suggesting that the ENase does not alter the *in vivo* specificity of the Mtase. Our demonstration of the identical kinetic behavior of the Mtase with short oligonucleotides and plasmid DNA (2) argues against changes in Mtase specificity deriving from the size of the substrate.

Alternatively, noncanonical site methylation may be occurring *in vivo*, but is in some way removed. One potential pathway involves the recently identified *mrr* (methylated adenine recognition and restriction) locus, which is proposed to encode an ENase that cleaves adenine-methylated DNA (28). Although proposed to restrict the entry of methylated DNA (28,29), this system may function in part to initiate the removal of adenine-methylated DNA involving noncanonical *Eco*RI sites. Reducing the proliferation of such methylated noncanonical sites may serve to maintain the specificity of the *Eco*RI ENase. In their initial identification of the *mrr* system, Heitman and Model (28) used a genetic screen linked to the ability of the *mrr* system to induce the SOS repair response as a result of DNA methylation. They report that expression of some DNA Mtases causes *mrr*-mediated SOS induction while others, including the *Eco*RI DNA Mtase, do not. However, these results are not inconsistent with our proposed role for the *mrr* system since their genetic screen may have missed mild induction of the SOS response (28) or other repair systems also responsive to the *mrr* system may be involved in the case of the *Eco*RI DNA Mtase.

CONCLUSIONS

The three assays described in this report clearly indicate that little if any methylation at the tested noncanonical *Eco*RI sites occurs *in vivo*. This result is at odds with our recently published analysis of the *in vitro* specificity of the Mtase (22). We excluded two potentially 'artifactual' causes of this inconsistency: the *in vivo* presence or absence of the *Eco*RI endonuclease and kinetic differences between small and large DNA substrates. The apparent discrepancy can be understood by postulating that noncanonically methylated sites are repaired, either by a system like the recently identified *mrr* system or by some as yet unidentified DNA repair process.

ACKNOWLEDGEMENT

This work was supported by National Foundation Grant DMB 89-11401 to N.O.R.

REFERENCES

1. Modrich, P., and Roberts, R.J. (1982) in *Nucleases* (Linn, S.M. and Roberts, R.J. Eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
2. Reich, N.O. and Mashhoon, N. (1991) *Biochemistry 30*, 2933–2939.
3. Wilson, G.G. (1991) *Nucl. Acids Res. 19*, 2539–2566.
4. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
5. Hager, P.W., Reich, N.O., Day, J.P., Coche, T.G., Boyer, H.W., Rosenberg, J.M. and Greene, P.J. (1990) *J. Biol. Chem. 265*, 21520–21526.
6. Heitman, J. and Model, P. (1990) *The EMBO J. 9*, 3369–3378.
7. Wolfes, H., Alves, J., Fliess, A., Geiger, R. and Pingoud, A. (1986) *Nucl. Acids Res. 14*, 9063–9080.
8. Alves, J., Ruter, T., Geiger, R., Fliess, A., Maass, G. and Pingoud, A. (1989) *Biochemistry 28*, 2678–2684.
9. Needles, M.C., Fried, S.R, Love, R., Rosenberg, J.M., Boyer, H.W., and Greene, P.J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3579–3583.
10. Corey, D.R. and Schultz, P.G. (1987) *Science 238*, 1401–1403.
11. Koob, M., Grimes, E. and Szybalski, W. (1988) *Science 241*, 1084–1086.
12. Strobel, S.A. and Dervan, P.B. (1991) *Nature 350*, 172–174.
13. Kim, Y., Grable, J.C., Love, R., Greene, P.J., Rosenberg, J.M. (1990) *Science 249*, 1307–1309.
14. Lesser, D.R., Kurpiewski, M.R. and Jen-Jacobson, L. (1990) *Science 250*, 776–785.
15. Wu, J.C. and Santi, D.V. (1987) *J. Biol. Chem. 262*, 4778–4786.
16. Pogolotti, A.L., Ono, A., Subramaniam, R. and Santi, D.V. (1988) *J. Biol. Chem. 263*, 7461–7464.
17. Newman, A.K., Rubin, R.A., Kim, S.-H. and Modrich, P. (1981) *J. Biol. Chem. 256*, 2131–2139.
18. Lauster, R. (1989) *J. Mol. Biol. 206*, 313–321.
19. Taylor, J.D., Goodall, A.J., Vermote, C.L. and Halford, S.E. (1990) *Biochemistry 29*, 10727–10733.
20. Taylor, J.D. and Halford, S.E. (1989) *Biochemistry 28*, 6198–6207.
21. Woodbury, C.P., Downey, R.L. and von Hippel, P.H. (1980) *J. Biol. Chem. 255*, 11526–11533.
22. Reich, N.O., Olsen, C., Osti, F., and Murphy, J. (1992) *J. Biol. Chem. 267* 15802–15807.
23. Polisky, B., Greene, P.J., Garfin, D.E., McCarthy, B.J., Goodman, H.M. and Boyer, H.W. (1975) *Proc. Natl. Acad. Sci. USA 72*, 33103314.
24. Rubin, R.A. and Modrich, P. (1977) *J. Biol. Chem. 252*, 7265–7272.
25. Rosenberg, J.M. and Greene, P.J. (1982) *DNA 1*, 117–123.
26. Kuhn, I., Stephenson, F.H., Boyer, H.W. and Greene, P.J. (1986) *Gene 44*, 253–263.
27. Borchardt, R.T. (1979) *J. Am. Chem. Soc. 101*,458–463.
28. Heitman, J. and Model, M. (1987) *J. Bact. 169*, 3243–3250.
29. Kretz, P.L., Kohler, S.W. and Short, J.M. (1991) *J. Bact. 173*, 4707–4716.