Studies on the biological role of DNA methylation: V. The pattern of *E. coli* DNA methylation*

Moshe Szyf, Yosef Gruenbaum, Simcha Urieli-Shoval and Aharon Razin

Department of Cellular Biochemistry, The Hebrew University-Hadassah Medical School, Jerusalem, Israel

Received 1 September 1982; Revised and Accepted 19 October 1982

ABSTRACT

The distribution of the methylatable sites GATC and CC^AGG was studied by analyzing the molecular average size of restriction fragments of *E. coli* DNA. Both sites were found to be randomly distributed, reflecting a random pattern of methylation. The methylation pattern of specific sequences such as the origin of replication and rRNA genes has been studied in wild type *E. coli* and a methylation deficient (*dam*⁻ *dcm*⁻) mutant. These sequences were found to be methylated in wild type cells and unmethylated in the mutant indicating that there is no effect of the state of methylation of these sequences on their expression. Analysis of the state of methylation of GATC sites in newly replicating DNA using the restriction enzyme Dpn I (cleaves only when both strands are methylated) revealed no detectable hemimethylated DNA suggesting that methylation occurs at the replication fork. Taking together the results presented here and previously published data (5), we arrive at the conclusion that the most likely function of *E. coli* DNA methylations is probably in preventing nuclease activity.

INTRODUCTION

Escherichia coli cells contain two modified bases in their DNA: 5-methylcytosine (m⁵Cyt) and N⁶-methyladenine (m⁶Ade) (1). While a minority of these modified bases play a role in restriction-modification systems (2,3), most of them are not produced by the type I modification enzymes (4). Two methylases are responsible for the methylation of these bases, a product of the *dam* gene methylates the adenine residues in GATC sequences and the *dcm* protein methylates the internal cytosine residue in CC^AGG sequences (4). The biological role played by these methylated sites is still obscure.

The present study attempts to gain further understanding of the role of DNA methylation in *E. coli* by elucidating the pattern of methylation. In one set of experiments the methylated sites in *E. coli* DNA were found to be randomly distributed and newly replicated DNA was found to be methylated with no detectable delay. In other experiments DNA replication and the expression of rRNA genes were not affected by the state of methylation of the corresponding

sequences. The implications of the results presented here on the understanding of the possible biological role of the methylated bases in E. coli DNA is discussed.

MATERIALS AND METHODS

The following products were purchased from New England Biolabs Co.: the restriction enzymes Msp I, Taq I, Hha I, BamH I, Mbo I, Sau 3A, BstN I, Pvu II, Pst I, E. coli DNA polymerase I, E. coli DNA polymerase I large fragment, T4 DNA ligase and 13mer primer. The enzymes EcoR II and Dpn I were obtained from Bethesda Research Laboratories. The enzymes micrococcal nuclease and spleen phosphodiesterase, non labelled nucleotides and salmon sperm DNA were products of Sigma Co. [$\alpha^{32}\text{P}$]dNTPs were obtained from New England Nuclear. [^{32}P] carrier-free phosphate was purchased from the Nuclear Research Center Negev-Israel. Agarose was obtained from Seakem Co.

Bacterial Strains and Growth Conditions. E. coli C (wild type) and E. coli F42 GM48 dam-3 dcm-6 (5) were grown in TPA medium containing per liter: 0.5g NaCl; 8g KCl; 1.1g NH_4Cl ; 0.2g $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$; 12.1g Tris base; 0.023g KH_2PO_4 ; 0.16mM Na_2SO_4 ; 0.8g sodium pyruvate; 1mg/ml amino acid mix, pH was adjusted to 7.4. The incubations were performed at 30°C or 37°C as indicated in the text.

Preparation of DNA. E. coli DNA was prepared by the Marmur procedure (6). M13 phage (mp7) single stranded (ss) DNA was prepared and used for synthesis of hemimethylated DNA, by primed repair synthesis (7).

Restriction Conditions and Determination of Average Molecular Size of Restriction Fragments. All restriction enzyme digestions of DNA were performed at 37°C in buffer mixtures recommended by the manufacturers. Restricted DNA was analyzed by electrophoresis on agarose (Seakem Co.) gels. Each gel was run with molecular size markers obtained by restriction enzyme digestion of various plasmids and phage molecules. Following electrophoresis at 20V for 16 h, the ethidium-bromide stained gels were photographed using positive-negative type 665 Polaroid film, and the film negative scanned using an Helena Quick Scan densitometer. The resulting graphs were computer digitalized and the average molecular size calculated by reference to the molecular size markers as described before (8).

Filter Hybridization. DNA fragments electrophoresed on agarose gels were transferred to nitrocellulose sheets, hybridized and washed as described (9). The plasmid pGG1 was nick translated with [$\alpha^{32}\text{P}$]dCTP to a level of $2\text{--}3 \times 10^8$ cpm (10). The plasmid pGG1 carrying the 5' end of *rrnB*, a ribosomal RNA cistron of E. coli (11) was a gift from Dr. G. Glaser. M13 Ori C26 phage, carrying

the E. coli origin of replication (a gift from Dr. D. Ray) was labelled by primed repair synthesis using [α^{32} P]dCTP (7).

RESULTS

Distribution of GATC sites in E. coli DNA. The fact that methylatable sites in E. coli DNA are fully methylated (12) allows us to use the distribution of the methylatable sites as a tool to learn the pattern of methylation. In an attempt to correlate this pattern with biochemical functions within the cell, experiments were designed to study the distribution of the methylatable sites GATC and CC^AGG in E. coli DNA sequences.

The distribution of GATC sites in the E. coli genome was studied using the restriction enzyme isoschizomers Mbo I (cleaves at GATC only when the sequence is not methylated) and Dpn I (cleaves only when GATC is methylated on both strands). DNA of wild type E. coli, cleaved with Dpn I (Fig. 1b) and DNA of a dam⁺ mutant, cleaved with Mbo I (data not shown) show the same fragment size distribution pattern (~0.3kb average fragment size) similar to that obtained by other restriction enzymes with a four nucleotide recognition site such as Msp I (0.24kb) (Fig. 1a). The distribution is random in nature and no clustering of GATC sites is observed. Non randomness is defined by deviation from the expected Gaussian shape (8). This analysis would not detect clusters of a small number of sites like the one which is present in the origin of replication (vide infra). However, if GATC sites were preferentially located at

Fig. 1. Distribution of GATC and CC^AGG sequences in E. coli DNA. E. coli DNA (5 μ g) was digested with Msp I (a), Dpn I (b) and BstN I (c). All enzymatic reactions were carried out with 2 units/ μ g DNA under the conditions recommended by the manufacturers. The digested DNAs were electrophoresed on 1.5% agarose gels, the gels stained with ethidium bromide, photographed and the negatives scanned as described in the methods. Size markers (Hind III λ phage digest and Hae III ϕ X RF digest) were electrophoresed on each gel, scanned and plotted to obtain the fragment size ordinate, and added to the DNA samples as internal markers to test complete digestion.

the ends of Okazaki fragments, as reported previously (13), it would not go undetected by our analysis.

Distribution and Frequency of CC^AGG Sites in *E. coli* DNA. The site CC^AGG is methylated, at the internal cytosine residue, in *E. coli* C and *E. coli* K12 DNA, but not in *E. coli* B DNA (4). Digestion of *E. coli* C DNA with the restriction enzyme BstN I (which cleaves in the CC^AGG site at the arrow, whether or not the internal cytosine residue is methylated) was used to study the distribution and frequency of CC^AGG sequences in *E. coli* DNA (Fig. 1c). The observed number average molecular size of the BstN I fragments was smaller than the calculated value based on published data of the nearest neighbor frequencies (12). The results presented in Table 1 indicate that the CC^AGG sites are overrepresented in the *E. coli* DNA about 70% above the expected frequencies of this site. The observed frequency of the CC^AGG sites is in agreement with the number of m⁵Cyt residues present in this DNA (12). CC^AGG is a subset of C^AG containing sequences, other subsets being CAGCTG (Pvu II) and CTGCAG (Pst I). A similar analysis using the restriction enzymes Pst I and Pvu II revealed that the sequences CTGCAG and CAGCTG as well are overrepresented in the *E. coli* C DNA. Digestion of *E. coli* C DNA with other restriction enzymes that do not contain C^AG in their recognition site, such as Msp I and BamH I, showed no overrepresentation of these sequences in *E. coli* DNA (table 1). It should be noted that in spite of the observation that C^AG sequences, in general, are overrepresented, only CC^AGG sites are methylated in *E. coli* C. The CC^AGG sites are randomly distributed in *E. coli* DNA (see Fig. 1c), thereby reflecting a random distribution of m⁵Cyt residues.

Restriction enzyme	Expected size (bp) a	Experimental size (bp) b	Over-representation % c
Msp I (CCGG)	240	240	0
BstN I (CC ^A GG)	660	380	74
Pst I (CTGCAG)	4350	2940	45
Pvu II (CAGCTG)	4350	2200	100
BamH I (GGATCC)	4900	5000	0

Table 1. Overrepresentation of C^AG sequences in *E. coli* DNA. *E. coli* C DNA (5 μ g) was digested by the restriction enzymes (2U/ μ g DNA) which are listed in the table using conditions as recommended by the enzymes manufacturers. The average molecular size was calculated as described in the methods. Percent overrepresentation (c) was calculated using the following equation: $c = \frac{a-b}{b} \times 100$.

Methylation of Specific Sequences in *E. coli* DNA. Although the experiments described above clearly demonstrate that the distribution of modified bases is random in bulk DNA, it cannot be excluded that specific sequences have a typical pattern of methylation. If DNA methylation in *E. coli* plays a role in gene activity or the biological activity of other sequences such as origin of replication, a specific methylation pattern would be expected at these sequences.

A mutant of *E. coli* K12, deficient in dam and mec methylase activities (dam-3 dcm-6) does not show any abnormality in rate of growth as compared to the wild type strain (data not shown). If methylated sites were involved in the activity of DNA sequences that play a role in cell replication, a limited number of methylated sites would be expected at strategic points on the DNA of the methylase deficient strain.

The enzyme BstN I was used to directly determine the extent of methylation of the internal cytosine in the CC^AGG sites of the mutant strain. This enzyme is not inhibited by m⁵Cyt in its cleavage site (7). It was therefore used to cleave at the 3' end of the internal cytosine CC^AAGG GGT+CC', and the resulting fragments were labelled by filling the sticky ends with [α^{32} P]dTTP or [α^{32} P]dATP using *E. coli* Pol I large fragment (14). The labelled DNA was degraded to deoxynucleoside-3'-monophosphates by micrococcal nuclease and spleen phosphodiesterase. The labelled nucleotides were chromatographed and chromatographs were autoradiographed. The results clearly indicate that while the wild type strain is fully methylated at CC^AAGG sites, the dam⁻ dcm⁻ strain is 5% methylated at these sites (Fig. 2). Although the residual methylation detected might represent a small percentage of dcm⁺ cells present in the culture, it might as well reflect specific methylated sites in the dcm⁻ genome. The extent of methylation of GATC sequences in *E. coli* dam⁻ dcm⁻ DNA was determined using the restriction enzyme Dpn I which cleaves the sequence GATC only when both strands are methylated. DNA of the dam⁻ dcm⁻ mutant was cleaved with the restriction enzyme Dpn I and size fractionated by electrophoresis on agarose gels (0.8% agarose) (Fig. 3). The gel was stained, photographed and scanned as described in Materials and Methods. Optical density scans reveal no detectable cleavage of the dam⁻ dcm⁻ DNA with Dpn I. Since under our experimental conditions fragments ≤ 30 kb could be detected (see the migration of the molecular size markers) it can be concluded that the average molecular size of Dpn I fragments of the mutant DNA is larger than 30kb. This implies $\leq 1\%$ methylation at GATC sites. (1% of methylation of GATC sequences on both strands would have resulted in an average molecular

Fig. 2. Extent of methylation of CCGG sites in *E. coli* DNA. DNA (1 μ g) of *E. coli* C (1), *E. coli* F42 GM48 (*dam*⁻ *dcm*⁻) (2) and *E. coli* B (3) were digested with 2u of BstNI as recommended by the manufacturer. The DNA fragments were labelled at their 3' end by filling the sticky ends using 20 μ c [α ³²P]dTTP (580 c/mmole) and 4 units DNA polymerase I (large fragment) for 10 min at 37°C. The labelled fragments were digested to deoxyribonucleoside-3'-monophosphates, chromatographed by two-dimensional TLC and autoradiographed (14).

size fragment of 30kb based on a frequency of the GATC sequence of one in 300 nucleotides).

The experiments described above clearly indicate that substantial methylation is not required for *E. coli* cells to carry out replication or gene expression. However, it is possible that limited methylation is sufficient and that some methylated sequences are still present in strategic positions such as the origin of replication or coding sequences even in the *dam*⁻ *dcm*⁻ strain.

Fig. 3. Extent of methylation of GATC sites in *E. coli* DNA. DNA (5 μ g) of *E. coli* F42 GM48 was digested with Dpn I, as recommended by the manufacturer. (λ DNA digested with Hind III served as molecular weight markers). Restricted DNA was separated by electrophoresis on 0.8% agarose gels, stained and photographed. (1) undigested *E. coli* DNA. (2) Dpn I treated *E. coli* DNA. (3) Dpn I treated *E. coli* DNA + λ DNA. (4) Dpn I treated λ DNA. (5) Hind III treated λ DNA.

To clarify this possibility the extent of methylation of the rRNA genes and the origin of replication of both *E. coli* C and the *dam*⁻ *dcm*⁻ strain was determined by Southern blot analysis.

The State of Methylation of rRNA Genes. The ribosomal RNA cistrons in *E. coli* do not contain BamH I sites (15). We have used this fact to visualize the seven cistrons of *E. coli* by hybridization of a BamH I digest of *E. coli* DNA to a P³² labelled pGG1 plasmid as probe (16). DNA of wild type or the *dam*⁻ *dcm*⁻ *E. coli* mutant was digested with BamH I and one of the isoschizomers BstN I or EcoR II and, in another experiment, with BamH I and one of the isoschizomers Mbo I or Dpn I. The resulting restriction fragments were blotted and hybridized to pGG1, as described in the methods. The results of this experiment (Fig. 4) clearly indicate that *E. coli* C rRNA genes are methylated at all CC^AGG and GATC sites. Digestion with EcoR II leaves the genes intact, whereas BstN I cuts the rRNA BamH I bands. Similarly, the BamH I bands are not eliminated by Mbo I but are digested with Dpn I. In contrast, *E. coli* *dam*⁻ *dcm*⁻ rRNA genes seem to be unmethylated at all GATC and CC^AGG sites judged by the same analysis (Fig. 4). Since Dpn I cleaves at GATC sites only when methylated on both strands even a single methylated GATC site could have been detected.

The State of Methylation of the Origin of Replication. The frequency of GATC sites in the 400bp stretch of the origin of replication of *E. coli* DNA is 18 fold higher than expected by the distribution of this sequence in the whole genome (17-19). If the methylation of these sites is essential to normal

Fig. 4. The state of methylation of the rRNA cistrons in *E. coli* DNA. *E. coli* C DNA (5µg) was digested with 10 units of the following restriction enzymes: 1. BamH I; 2. BamH I + Mbo I; 3. BamH I + Dpn I; 4. BamH I + EcoR II; 5. BamH I + BstN I. *E. coli* F42 GM48 (*dam*⁻ *dcm*⁻) DNA (5µg) was digested with 10 units of: 6. BamH I; 7. BamH I + Mbo I; 8. BamH I + Dpn I; 9. BamH I + EcoR II; 10. BamH I + BstN I. The digests were electrophoresed on 0.8% agarose gels blotted and hybridized to nick translated P³² labelled pGG1 and autoradiographed as described in the methods.

functioning of the origin, the sites are expected to be methylated, even in the *dam*⁻ mutant. Specific methylation of this region by residual methylase activity in the mutant cells could not be ruled out by the experiments described above (see Fig. 3). To clarify this point the following experiments were carried out: *E. coli* C and *E. coli* *dam*⁻ *dcm*⁻ DNA were digested with the restriction enzyme Hha I. Using M13-Ori C-26 ³²P labelled probe, for Southern blotting, two bands were observed: a 700bp band (containing the origin) and another band resulting from a Hha site in the neighbouring *asn* gene (~1.3kb) (17,20). DNA cut with Hha I was further digested with one of the isoschizomers Dpn I or Mbo I and EcoR II or BstN I and blotted. As can be seen in Fig. 5A, the Hha I bands derived from *E. coli* C DNA remained intact when digested with EcoR II (lane 9) or Mbo I (lane 7) and eliminated when digested with BstN I (lane 10) or Dpn I (lane 8). These results indicate that in *E. coli* C DNA the CCG^AGG and GATC sites which are present in the origin of replication are fully methylated. On the other hand all GATC sites in the

Fig. 5. The state of methylation of the origin of replication of *E. coli*.

A) *E. coli* F42 GM48 DNA (5µg) was digested with 10 units of the following restriction enzymes: 1. Hha I; 2. Hha I + Mbo I; 3. Hha I + Dpn I; 4. Hha I + EcoR II; 5. Hha I + BstN I. *E. coli* C DNA(5µg) was digested with 10 units of: 6. Hha I; 7. Hha I + Mbo I; 8. Hha I + Dpn I; 9. Hha I + EcoR II; 10. Hha I + BstN I.

B) *E. coli* F42 GM48 DNA (5µg) was digested with 10 units of the following restriction enzymes: 1. BamH I; 2. BamH I + Mbo I; 3. BamH I + Dpn I; 4. BamH I + EcoR II; 5. BamH I + BstN I. The cleaved DNA was size fractionated by agarose gel (1.5% for A and 0.8% for B), electrophoresis. The gels were blotted to nitrocellulose filters and hybridized to M13 Ori C26 ³²P labelled DNA (see methods) and autoradiographed.

origin of replication of a dam⁻ dcm⁻ mutant are not methylated as indicated by the observation that Hha I bands of this DNA remain intact after digestion with Dpn I (lane 3). It should be noted here that a single methylated site would be detected by this analysis. Also the CC^AGG sites are not methylated in this dam⁻ dcm⁻ mutant as judged by the elimination of the Hha I bands after digestion with EcoR II (lane 4).

The GATC and CC^AGG sequences are also highly represented in the flanking sequences of the origin of replication. The enzyme BamH I has three sites within the origin of replication and two sites 2.4kb and 3.5kb away from the 3' and 5' ends respectively (Fig. 5B). Thus the extent of methylation of the flanking sequences of the origin could be analyzed. The same strategy described above for Hha I digestion experiments was used for BamH I fragments. The results of these experiments suggest that the flanking sequences of the E. coli origin of replication are completely unmethylated in the dam⁻ dcm⁻ mutant DNA (see Fig. 5B).

Methylation at the Replication Fork. The state of methylation of DNA at the fork could be in one of the two alternative states: 1. If DNA methylation lags behind DNA replication the DNA at the fork will be hemimethylated. Thus pulse labelled newly replicating DNA will not be cleaved by Dpn I or Mbo I, whereas bulk DNA will be cut by Dpn I and not by Mbo I. 2. If DNA methylation occurs concomitantly with the movement of the replicating complex DNA will be fully methylated, therefore sensitive to Dpn I and undigested by Mbo I as is the case for bulk DNA. To test these alternatives the following experiments were carried out.

E. coli C cells were grown in low phosphate medium (TPA) at 30°C to 2×10^8 cells/ml. The culture was pulse-labelled with carrier-free [³²P] phosphate (100 μ Ci/ml). 10ml aliquots were withdrawn at 30 sec, 2 min and 5 min after addition of label and the pulse was terminated by pouring the samples over a frozen solution of 20mM KCN-200mM nicotinamide. (KCN is an effective quenching agent (21) and nicotinamide is a potent inhibitor of DNA methylation in vivo and in vitro (22)). Pulse-labelled DNA was purified as described in the methods and digested with the restriction enzymes Dpn I and Mbo I. These digests were subjected to fragment size analysis of both pulse-labelled and bulk DNA as described in the methods. Hemimethylated M13 phage (mp7) DNA synthesized in vitro as described in the methods was added as a control to one digestion mixture of E. coli DNA with Dpn I. Unmethylated λ phage DNA was added as control to a digestion mixture of E. coli DNA with Mbo I. The results of these experiments which are presented in Fig. 6, clearly

Fig. 6. Methylation at the replication fork. *E. coli* C cells were grown to 2×10^8 cells in TPA medium at 30°C . The culture was pulse-labelled with $100 \mu\text{C}/\text{ml}$ [^{32}P] carrier-free phosphate for 30 sec. - 5 min. The labelled DNA was purified by the Marmur procedure (6), digested with Dpn I and Mbo I under conditions recommended by the manufacturer. The cleaved DNA was electrophoresed on 1% agarose gel stained with ethidium bromide and photographed. The gels were dried under vacuum and autoradiographed.

Lanes 1-8: Ethidium bromide stained gel: 1) 30 sec pulsed DNA cleaved with Dpn I. 2) 30 sec pulsed DNA cleaved with Mbo I. 3) 2 min pulsed DNA cleaved with Dpn I. 4) 2 min pulsed DNA cleaved with Mbo I. 5) 5 min pulsed DNA cleaved with Dpn I. 6) 5 min pulsed DNA cleaved with Mbo I. 7) Hemimethylated M13 phage (mp7) DNA cleaved with Dpn I. 8) Unmethylated λ phage DNA cleaved with Mbo I.

Lanes 9-14: Autoradiogram: 9) 30 sec pulsed DNA cleaved with Dpn I. 10) 30 sec pulsed DNA cleaved with Mbo I. 11) 2 min pulsed DNA cleaved with Dpn I. 12) 2 min pulsed DNA cleaved with Mbo I. 13) 5 min pulsed DNA cleaved with Dpn I. 14) 5 min pulsed DNA cleaved with Mbo I.

demonstrate that nascent DNA synthesized for 30 sec ($\sim 15\text{kb}$) is fully methylated since it was completely digested with Dpn I (Fig. 6, lane 9). Comparing the optical density scans of the Dpn I digest of bulk DNA (Fig. 6, lane 1) with digested pulse labelled DNA (30 sec, lane 9) and DNA's labelled for longer periods (2min, lane 11; 5 min, lane 13) did not reveal any detectable hemimethylated DNA (which would have not been digested with Dpn I) even in the 30 sec pulse labelled DNA. This result allows the conclusion that DNA methylation does not lag behind the replication fork for more than 3kb (20% of the label in a 3kb band could easily be detected by our assay).

DISCUSSION

Methylated bases in DNA may affect the biological activity of specific gene sequences as suggested in eukaryotic systems (23), or alternately their function can be associated with the entire genome as in restriction-modification systems of prokaryotic cells (4). Based on the assumption that the biological significance of DNA methylation is reflected in the pattern of methy-

lation, we designed our experiments to elucidate the pattern of methylation along the entire *E. coli* chromosome and at specific sequences coding for known biological functions which are indispensable for the cell growth.

The overall pattern of methylation in *E. coli* DNA was studied by determining the distribution of the methylatable sites GATC and CC^AGG. The results of these experiments revealed a random distribution of the methylatable sites (see Fig. 1). This random distribution of GATC and CC^AGG sites along the entire chromosome suggests that the methylated bases in *E. coli* DNA play some general role rather than a local sequence specific one. For example, methylated bases could affect the tertiary structure of the DNA (26) or alternately protect the DNA from endonucleolytic cleavage. Several facts make the second possibility feasible: a mutant in DNA adenine methylation (*dam*-3) shows high mutability and high frequency of single-stranded breaks; although the *dam*⁻ mutants are viable they give rise to lethal phenotypes in conjunction with non lethal ts mutants in *recA*, *recB*, *recC* or *polA* (5). Deficiency of methylation at CC^AGG is associated with increased genetic recombination (24).

The possibility still exists, however, that methylation at defined sequences can have a local function. To test this possibility, we investigated a *dam*⁻ *dcm*⁻ mutant of *E. coli* K12 which is deficient in both DNA adenine and cytosine methylase activities. Since these mutant cells grow normally there was a strong possibility that residual methylation at strategic sequences allows the normal functioning of the cell. The overall extent of this residual methylation at the GATC and CC^AGG sites was studied and found to be $\leq 1\%$ of the GATC sites and 5% of the CC^AGG sites methylated. This residual methylation could be a result of leakiness of the mutations or reflect methylation at specific sequences. The extent of methylation of two sequences indispensable for cell growth, the rRNA genes and origin of replication, was analyzed. No methylation could be observed in these sequences (see figures 4,5), indicating that the origin of replication and the rRNA genes function normally regardless of whether these sequences are methylated or not.

The hypermutability of *dam*⁻ cells (5) led to a suggestion that the methylation of GATC sites might play a role in strand discrimination during post replication mismatch repair (25). The results of a study on the extent of methylation of nascent DNA in a *polA* lig-4 mutant (27) were considered to be supportive to the strand discrimination theory. This study revealed a reduced methylation (0.96mole %) of the nascent DNA as compared to bulk DNA (1.4mole %). Since the nascent DNA in this experiment was a 2 min pulse labelled DNA the results were interpreted to mean that methylation is

lagging behind the replication complex. In the present study we analyze in nascent DNA the extent of methylation of GATC sites rather than mole % of methylated Ade residues. This analysis revealed no detectable lag indicating that methylation takes place at the replication fork (Fig. 6). The possibility that the methylation of GATC sites plays a role in strand discrimination is very unlikely in light of this observation. The lower extent of methylation of the nascent DNA in the ligase mutant (27) can be explained by the inability of the dam methylase to methylate the sites at the ends of Okazaki fragments (14).

The fact that the methylatable sites in E. coli DNA are randomly distributed and essentially fully methylated points towards the possibility that the methylated bases protect the DNA from nuclease activity. It was suggested before that the methylation of GATC sites protects the DNA from a specific nuclease producing single stranded breaks (5). Biological processes such as genetic recombination, DNA replication and repair probably involve such a nuclease activity. It is therefore tempting to speculate that methylation of DNA in E. coli might control the extent and determine the location of these breaks in the DNA. If this were the function of DNA methylation it would be expected that a limited number of sites at specific positions remain transiently unmethylated. It is those sites which still serve as hot spots for single stranded breaks. If this model is correct the unmethylated sites on a high methylation background provide the appropriate signal for carrying out the biological function.

ACKNOWLEDGEMENTS

We acknowledge the excellent technical assistance of Mr. I. Akabi and thank Mrs. C. Gopin for her help in preparing the manuscript. We are grateful to Dr. G. Glaser for the pGG1 plasmid and help with the rRNA genes experiment, and to Dr. D. Ray for providing the Ori C26 M13 phage. This research was supported by the N.I.H. grant: GM 20483.

REFERENCES

- * Paper IV in this series is ref. 12.
- 1. Doskocil, J. and Sormova, Z. (1965) *Biochim. Biophys. Acta* 95, 513-515.
- 2. Kuhnlein, U., Linn, S. and Arber, W. (1969) *Proc. Natl. Acad. Sci. USA* 63, 556-562.
- 3. Haberman, A., Heywood, J. and Meselson, M. (1972) *Proc. Natl. Acad. Sci. USA* 69, 3138-3141.
- 4. Razin, A. and Friedman, J. (1981) *Prog. Nucleic Acids Res. and Mol. Biol.* 25, 33-52.
- 5. Marinus, M.G. and Morris, N.R. (1974) *J. Mol. Biol.* 85, 309-322.

6. Marmur, J. (1961) J. Mol. Biol. 3, 208-218.
7. Gruenbaum, Y., Cedar, H. and Razin, A. (1981) Nucleic Acids Res. 9, 2509-2515.
8. Naveh-Man, T. and Cedar, H. (1982) Mol. Cell. Biol. 2, 758-762.
9. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
10. Pollack, Y., Stein, R., Razin, A. and Cedar, H. (1980) Proc. Natl. Acad. Sci. USA 77, 6463-6467.
11. Glaser, G., Enquist, L. and Cashel, M. (1977) Gene 2, 159-172.
12. Razin, A., Urieli, S., Pollack, Y., Gruenbaum, Y. and Glaser, G. (1980) Nucleic Acids Res., 8, 1783-1791.
13. Gomez-Eichelmann, M.C. and Lark, K.G. (1977) J. Mol. Biol. 117, 621-635.
14. Shoval-Urieli, S., Gruenbaum, Y. and Razin, A. (1982) J. Bacteriol., in press.
15. Boros, I., Kiss, A. and Venetianer, P. (1979) Nucleic Acids Res. 6, 1817-1830.
16. Amikam, D., Razin, S. and Glaser, G. (1982) Nucleic Acids Res. 10, 4215-4222.
17. Sugimoto, K., Oka, A., Sugisaki, H., Takanami, M., Nishimura, A., Yasuda, Y. and Hirota, Y. (1979) Proc. Natl. Acad. Sci. USA 76, 575-579.
18. Meijer, M., Beck, E., Hansen, F.G., Bergmans, H.E.N., Messer, W., von Meyenburg, K. and Schaller, H. (1979) Proc. Natl. Acad. Sci. USA 76, 580-584.
19. Zyskind, J.W. and Smith, D.W. (1980) Proc. Natl. Acad. Sci. USA 77, 2460-2464.
20. Kaguni, J., Laverne, L.S. and Ray, D.S. (1979) Proc. Natl. Acad. Sci. USA 76, 6250-6254.
21. Okazaki, T. and Okazaki, R. (1969) Proc. Natl. Acad. Sci. USA 64, 1242-1248.
22. Razin, A., Goren, D. and Friedman, J. (1975) Nucleic Acids Res. 2, 1967-1974.
23. Razin, A. and Riggs, A.D. (1980) Science 210, 604-610.
24. Korba, B.E. and Hays, J.B. (1982) J. Mol. Biol. 157, 213-235.
25. Radman, M., van den Elsen, P. and Glickman, B. (1978) Molec. Gen. Genet. 163, 307-312.
26. Engel, J.D. and von Hippel, P.H. (1978) J. Biol. Chem. 253, 927-934.
27. Marinus, M.G. (1976) J. Bacteriol. 128, 853-854.