Volume 8 Number 8 1980 Nucleic Acids Research

Studies on the biological role of DNA methylation; IV. Mode of methylation of DNA in E. coli cells*

Aharon Razin, Simcha Urieli, Yaakov Pollack, Yosef Gruenbaum and Gad Glaser

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

Received 14 February 1980

ABSTRACT

Two pairs of restriction enzyme isoschizomers were used to study in vivo methylation of E. coli and extrachromosomal DNA. By use of the restriction enzymes MboI (which cleaves only the unmethylated GATC sequence) and its isoschizomer Sau3A (indifferent to a methylated adenine at this sequence), we found that all the GATC sites in E. coli and in extrachromosomal DNAs are symmetrically methylated on both strands. The calculated number of GATC sites in E. coli DNA can account for all its m^6Ade residues. Foreign DNA, like mouse mtDNA, which is not methylated at GATC sites became fully methylated at these sequences when introduced by transfection into E. coli cells. This experiment provides the first evidence for the operation of a de novo methylation mechanism for E. coli methylases not involved in restriction modification. When the two restriction enzyme isoschizomers, EcoRII and ApyI, were used to analyze the methylation pattern of $CC^{A}_{T}GG$ sequences in E. coli C and ϕX174 DNA, it was found that all these sites are methylated. The number of $CC^{A}_{T}GG$ sites in E. coli C DNA does not account for all m^5Cyt residues.

INTRODUCTION

Several recent studies indicate that the site methylated by the DNA adenine methylase (dam protein) of E. coli is the adenine residue in the sequence GATC (1,2,3). Studies on the distribution of m^5Cyt between pyrimidine isopliths derived from E. coli DNA suggests that the internal cytosine in the seuqnce $CC^{A}_{T}GG$ is methylated by the E. coli DNA cytosine methylase (mec protein) (4,5).

The present report describes experiments in which the mode of in vivo activity of the dam and mec methylases is investigated by the use of restriction enzymes.

METHODS

Preparation of DNA. E. coli B and E. coli C DNA was prepared by the Marmur procedure (6). Mouse L cells mtDNA was prepared according to Bogen-

haven and Clayton (7). In vitro RF was prepared using phage SS DNA as template and a synthetic oligonucleotide as primer (8). The cloning of mtDNA will be described elsewhere (Y. Pollack et al. manuscript in preparation). ϕXRFDNA, pBR322 and pKP7 plasmids were prepared as described before (9).

Restriction enzymes and restriction conditions. Restriction enzymes MboI, Sau3A, HaeIII and Bam HI were from New England Biolabs. EcoRII was a Bethesda Research Laboratories product and ApyI was prepared from Arthrobacter pyridinolis by Drs. R. DiLauro and M. Sobol (to be published) and kindly provided to us by R. DiLauro. All restriction reactions were performed at 37° for 1 hr in buffer mixtures recommended by the manufacturers. ApyI restriction conditions were identical to those of EcoRII. In all reactions, two units of enzyme were used per μg DNA. In order to rule out the possibility of endonuclease inhibition by the various DNA preparations the appropriate internal controls were run in each experiment.

Gel electrophoresis. Restricted DNA was analyzed by gel electrophoresis. Samples were run on either 1.5% agarose (Seakem) or 5% acrylamide at 200V for 2 hr. Gels were stained with 1 μg/ml ethidium bromide and photographed by a Polaroid MP4 land camera using 107 type film.

RESULTS

Methylation pattern of E. coli DNA. E. coli B DNA is methylated in around 2% of its adenine residues. E. coli C DNA is methylated in both adenine residues (2%) and cytosine residues (1%) (10). We used the pair of restriction enzymes isoschizomers MboI and Sau3A to probe for the methylation of the GATC sequence in E. coli B and E. coli C DNA. The restriction enzyme MboI cleaves DNA at the sequence GATC only when the adenine residue is not methylated (11), whereas the isoschizomer Sau3A cleaves the very same sequence irrespective of its being methylated or not. As can be seen in Fig. 1 (lanes C,E) the DNA of both E. coli B and E. coli C is completely resistant to cleavage by MboI, but is readily cleaved by Sau3A (fig. 1 D,F), indicating that most, if not all, the GATC sites in E. coli DNA are methyalted.

The sequence $^{\downarrow}CC\overset{A}{T}GG$ is recognized by the restriction enzyme EcoRII and cleaved at the arrow only when the internal cytosine of this sequence is not methylated (12). An isoschizomer of EcoRII was isolated and characterized by Drs. R. DiLauro and M. Sobol. This enzyme, ApyI, was found in our laboratory to cleave the sequence $CC^{\downarrow}\overset{A}{T}GG$, at the site indicated by the arrow (data not shown), preferentially when the internal cytosine is methylated. As can be seen in Fig. 1 (G,I), E. coli B DNA which is not methylated at this sequen-

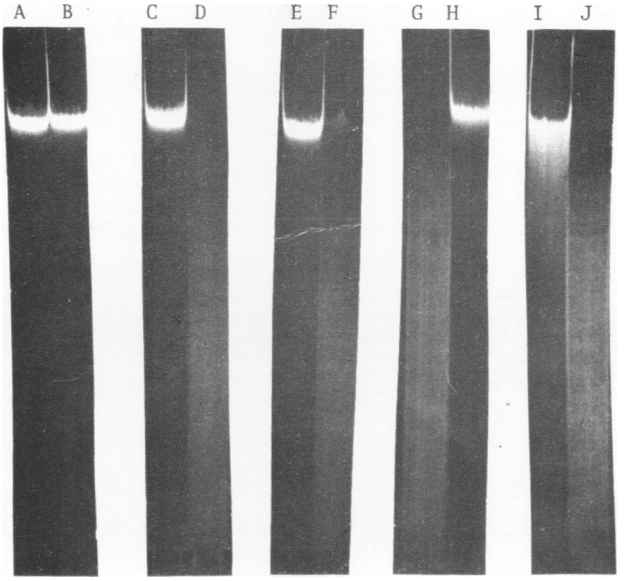

Fig 1. Restriction analysis of methylated sequences in E. coli DNA. DNAs of E. coli B and E. coli C were incubated with respective restriction enzymes and digestion products analyzed by 1.5% agarose gel electrophoresis, as described in the methods. A and B - untreated controls of E. coli B and E. coli C DNA; C and E - MboI treated E. coli B and E. coli C DNA, respectively; D and F - Sau3A treated E. coli B and E. coli C DNA; G and H - EcoRII treated E. coli B and E. coli C DNA; I and J - ApyI treated E. coli B and E. coli C DNA.

ce is extensively degraded by EcoRII, but partially degraded by ApyI, whereas E. coli C DNA which is expected to be methylated at this sequence is not cleaved by EcoRII, but readily cleaved by ApyI (fig. 1 (H,J)).

In order to calculate the extent of methylation of GATC and $CC^{A}_{T}GG$ sequences in E. coli DNA and in order to estimate whether all the methylated adenine and cytosine residues in E. coli DNA are accounted for by the frequency of these sequences, published base composition and nearest neighbour analysis data were used. As can be seen in Table 1, these calculations lead to the conclusion that all GATC and $CC^{A}_{T}GG$ sequences in E. coli DNA are methylated. However, whereas the frequency of GATC can account for all m^{6}Ade residues in E. coli DNA, there are more m^{5}Cyt residues in E. coli DNA than $CC^{A}_{T}GG$ sequences. The values presented in Table 1 suggest also that the GATC sequences are symmetrically methylated on both strands. In order to prove it the restriction enzyme DpnI, which cleaves at this sequence only

Base	Sequence	Base Frequency per 100 bases	Sequence Frequency per 100 bases
m^6Ade	GATC	0.42-0.5	0.33-0.42
m^5Cyt	$CC\frac{A}{T}GG$	0.24-0.26	0.14-0.16

Table 1. Frequency of minor bases and the sequences GATC and $CC\frac{A}{T}GG$ in E. coli DNA. The values of both base frequency and nearest neighbor analysis vary in different experiments, therefore a range of observed values is used (10,29). Sequence frequency is calculated based on published data of nearest neighbor analysis (14) and assuming 50% GC content as an average value for E. coli DNA.

when the adenine residues are methylated on both strands (13), was used. DNA of both E. coli B and E. coli C was found to be completely degraded by DpnI (data not shown).

Methylation pattern of extrachromosomal DNA. Knowing the sites methylated by E. coli methylases, we were able to study the methylation of these sequences in extrachromosomal DNA. The complete nucleotide sequence of the DNA of the bacteriophage φX174 (infects E. coli C) is well established (15). A survey of GATC and $CC\frac{A}{T}GG$ sequences in the single-stranded φX174 DNA reveals that a CCAGG sequence starts at position 881 and a CCTGG site at position 3500 of the DNA sequence, while no GATC sequences are found. One m^5Cyt residue was found previously by us in progeny single-stranded DNA (16) and localized in gene H (Z2 fragment of HaeIII digest) (17). Position 3501 in the DNA sequence is localized in Z2 fragment, therefore, the previously observed m^5Cyt residue might be localized at this position. The possibility that more than one m^5Cyt residue might exist in some of the phage molecules has been suggested by some of the early experiments (16, 17). The location of this second methylated base (17) might be at position 882. Our early attempts to analyze in vivo φX174 replicative forms (RF) for their extent of methylation failed due to the lack of an appropriate sensitive method (16). In the present study we used EcoRII and ApyI to analyze the methylation of two sites (at 882 and 3501) in RF and in progeny single-stranded phage DNA. Double-stranded φXDNA isolated from infected cells (in vivo RF) and double-stranded φXDNA made in vitro (in vitro RF) (8) were digested by the restriction enzyme HaeIII to yield Z fragments. Cleavage at position 3500 will result in cutting fragment Z2 (1078 bp) into two fragments (708 bp and 370 bp).

Cleavage at position 881 (Z5, 310 bp) will result in a 212 bp and 98 bp fragment. If a site is methylated it will resist cleavage by EcoRII, but will be cleaved by ApyI. The results indicate that positions 3501 and 882 are methylated in RF extracted from infected cells (Fig. 2 B,C). Progeny single-stranded DNA is definitely methylated at position 3501, but the results concerning the site at position 882 are not unambiguous. It seems that some of the φXDNA molecules are also methylated at position 882 (Fig. 2. D,E), as previous results suggested (16,17).

As demonstrated in Fig. 2. (lane D) it is evident that ApyI cleaves the half methylated ($\frac{C^{m}C\ T\ G\ G}{G\ G\ A\ C\ C}$) site at position 3500. Therefore a possibility exists that the methylase methylates efficiently CCTGG sequences and very slowly CCAGG. If that is the case, generally the (+) strand of <u>in vivo</u> RF will be methylated at 3501 and the (-) strand at position 882. Since the progeny φX single-stranded DNA derives from the (+) strand of <u>in vivo</u> RF during single-stranded DNA replication, it will, in most molecules, be methylated only at position 3501, and in some fraction of the DNA of progeny virus also position 882 will be methylated.

Since φXDNA was found to be devoid of GATC sites, we studied the methyl-

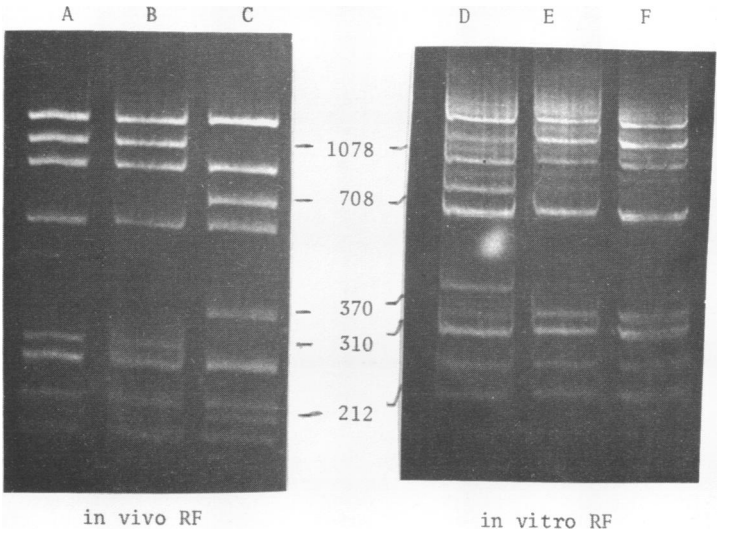

Fig. 2. Analysis of the precise location of m^{5}Cyt in φX174DNA. HaeIII fragments of <u>in vivo</u> and <u>in vitro</u> RF were treated with EcoRII and ApyI as described in the methods and analyzed by 5% acrylamide gel electrophoresis. A,F - HaeIII treated; B,E - HaeIII+EcoRII treated; C,D - HaeIII+ ApyI treated.

ation pattern of this sequence in plasmid DNA. The results of the experiment shown in Fig. 3 reveal that all 22 GATC sites in pBR322 DNA (18) are methylated in most of the molecules. The faint band corresponding to linear pBR322 (Fig. 3,C) may represent one unmethylated GATC site in a very small percentage of the pBR322 molecules. Comparable results were obtained with a number of plasmids propagated in E. coli that we carry in our laboratory (data not shown).

A "de novo" mode of methylation. Since the methylated sequences in E. coli DNA constitute palindromes, two modes of methylation can be suggested: One methylating activity could perform by a "semiconservative" mechanism

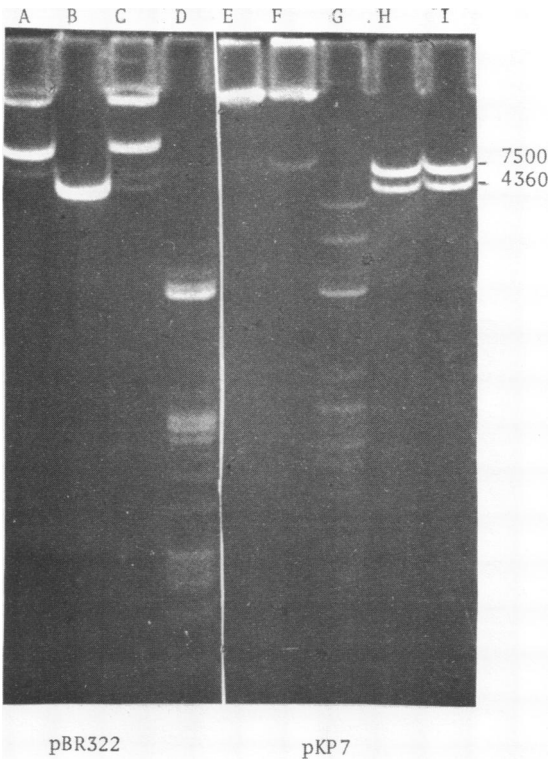

Fig. 3. Methylation pattern of GATC sites in pBR322 and pKP7 plasmids. A recombinant plasmid (pKP7) was constructed by ligation of a BamH1 7500bp fragment from mouse mtDNA and pBR322 linearized by BamHI (Pollack et al. in preparation). pBR322 and pKP7 were treated with MboI, Sau3A, and BamHI, as described in the methods. The treated DNAs were analyzed by 1.5% agarose gel electrophoresis. A - untreated control; B - BamHI digest; C - MboI treated; D - Sau3A digest; E - untreated control; F - MboI treated; G - Sau3A digest; H - BamHI digest; I - MboI + BamH1 treated.

based on methylating half methylated sequences (i.e. $\begin{matrix} G^{m}A\,T\,C \\ C\,T\,A\,G \end{matrix} \rightarrow \begin{matrix} G^{m}A\,T\,C \\ C\,T\,A\,G^{m} \end{matrix}$). Such a "maintenance enzyme" was postulated previously, to explain the inheritance of the methylation pattern in DNA (19, 20). Another methylating activity might exist in the cell, that carries out de novo methylation of specific sequences (i.e. $\begin{matrix} G\,A\,T\,C \\ C\,T\,A\,G \end{matrix} \rightarrow \begin{matrix} G^{m}A\,T\,C \\ C\,T\,A\,G \end{matrix}$). This enzyme may or may not methylate the opposite strand. The semiconservative mechanism was demonstrated in eukaryotic systems (21) and probably exists also in prokaryotes. To our knowledge, except in the case of modification enzymes, there is no evidence available to support a de novo methylation mechanism, either in prokaryotes or eukaryotes. The following experiment was designed to answer the question whether or not de novo methylation exists in E. coli. A recombinant plasmid constructed by ligation of pBR322 plasmid DNA (cut with BamHI at the single site in the tetracycline resistance region) with a BamHI fragment of mtDNA isolated from mitochondria of L cells was used to transfect E. coli cells. The progeny plasmid molecules were analyzed for their methylation in GATC sites. As can be seen in Fig. 3 (A,C), all 22 GATC sites in pBR322 are methylated. The recombinant plasmid containing the 7500 bp BamHI fragment of mtDNA is fully resistant to MboI restriction as well (see Fig. 3 F,I), suggesting complete methylation of MboI sites in the 7500 bp mtDNA fragment.

To rule out a possibility that GATC sites are methylated in the original mtDNA, we performed the control experiment presented in Fig. 4. Four BamHI fragments are obtained by digestion of mtDNA with BamHI (Fig. 4 lane C). The recombinant plasmid used in our experiment was constructed with the 7500 bp BamHI fragment. This fragment possesses at least 3 GATC sites, none of them methylated (Fig. 4 lanes A and B). With this control experiment it may be concluded that the unmethylated GATC sites in the BamHI fragment of mtDNA became fully methylated in the pKP7 plasmid by de novo methylation, which must exist in the E. coli cells.

DISCUSSION

The results of the experiments described in the present paper throw some light on the mode and function of DNA methylation. Restriction enzymes made it possible to study the state of methylation of specific sites, rather than the degree of gross methylation of the DNA (22). Our experiments provide more evidence supporting the observation that DNA adenine methylase and DNA cytosine methylase in E. coli are sequence specific. The adenine residues in all the GATC sequences of the cell genome, and of any extrachromosomal DNA propagated in the cell, are found to be fully and symmetrically methyla-

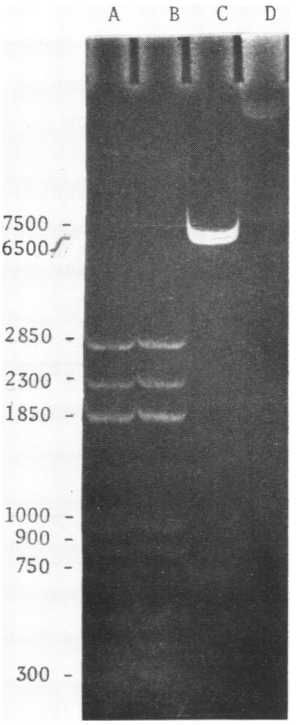

Fig. 4. GATC sites in BamHI (7500 bp) fragment of mouse mtDNA; control experiment. DNA of mouse L cell mitochondria was prepared and incubated with BamHI, MboI and BamHI+MboI, as described in the methods. The restriction fragments were analyzed on a 1.5% agarose gel. A - MboI treated mtDNA; B - BamHI+MboI digest; C - BamHI fragments; D - untreated control.

ted on both strands. With respect to the internal cytosine of CC$\frac{A}{T}$GG seuqences this conclusion probably holds for E. coli DNA, but is questionable for extrachromosomal DNA like that of the bacteriophage ϕX174. The calculated number of CC$\frac{A}{T}$GG sites in E. coli DNA cannot account for all m^5Cyt residues in this DNA. This observation implies that either both cytosine residues in this sequence are methylated, or that another, yet unknown, sequence is also methylated. Data supporting the second alternative have recently been published (23).

One important outcome of the present study is the observation of de novo methylation. Since the methylated sequences in E. coli DNA are found to have twofold rotational symmetry, methylation of DNA in E. coli could be achieved by a semiconservative type of methylation during replication. It was not clear whether or not the cell methylases are capable of methylating de novo unmethylated sites in the DNA. It should be noted that de novo methylation might still be very slow and inefficient, and the reason why we detect it in our experiment is because we select for the methylated

molecules, the only ones to survive.

Very little is known with respect to the biological function of the observed pattern of methylation of the mentioned sites. Studies with the DNA adenine methylase deficient mutants(dam$^-$) reveal that absence of methylation of GATC sequences results in high mutability and increased recombination rate. Although not lethal by themselves, lethal phenotype result when the methylase deficiency is in conjunction with non-lethal mutations in recA, recB, recC or polA (24). These strains are also deficient in a correct repair of mispairing lesions in the DNA (25). The last observation is consistant with a previous supposition that the undermethylation of newly replicated DNA provides a basis for strand discrimination between the "correct" parental strand and the "error containing" newly synthesized strand for an efficient and high fidelity editing during replication (26). Results of recent experiments in our laboratory indicate that the replication of DNA in E. coli is coupled to its methylation (manuscript in preparation). These results become even more intriguing when the nucleotide sequence of the E. coli origin of replication (27,28) is examined. The number of GATC and CC$^{A}_{T}$GG sequences found in the 422 nucleotide fragment containing the E. coli origin of replication are much more frequent than on a random basis. The sequence GATC is ten times more frequent and CC$^{A}_{T}$GG is five times more frequent than expected. Since it is evident from our results that these sequences have to be fully methylated, it implies that the origin of replication of E. coli is heavily methylated.

In order to gain closer insight into the mechanism of methylation of DNA in E. coli, in vitro experiments are required. The mec and dam proteins were isolated by us and physically separated from each other. Their characteristics will be published elsewhere (manuscript in preparation).

ACKNOWLEDGEMENTS

This study was supported by the U.S. Public Health Service, Grants No. GM 20483 and GM 25286. We thank Mr. Y. Akabi for technical assistance.

REFERENCES

* Paper III in this series is Friedman, J., Friedmann, A. and Razin, A. (1977) Nucleic Acids Res. 4, 3483-3496.

Abbreviations used: ϕXRFDNA - Replicative form DNA of bacteriophage ϕX174. mtDNA - mitochondrial DNA. m^6Ade - 6-methyladenine, m^5Cyt - 5-methylcytosine. bp - base pairs.

1. Hattman, S., Brooks, J.E. and Masurekar, M. (1978) J. Mol. Biol. 126, 367-380.
2. Dreiseikelmann, B., Eichenlaub, R. and Wackernagel, W. (1979) Biochim. Biophys. Acta 562, 418-428.
3. Geier, G.E. and Modrich, P. (1979) J. Biol. Chem. 254, 1408-1413.
4. May, M.S. and Hattmann, S. (1975) J. Bacteriol. 122, 129-138.
5. Bogdarina, I.G., Vagabova, L.M. and Buryanov, Y.I. (1976) FEBS letters 68, 177-180.
6. Marmur, J. (1961) J. Mol. Biol. 3, 208-218.
7. Bogenhaven, D. and Clayton, D.A. (1974) J. Biol. Chem. 249, 7991-7995.
8. Razin, A., Hirose, T., Itakura, K. and Riggs, A.D. (1978) Proc. Nat. Acad. Sci. USA. 75, 4268-4270.
9. Glaser, G. and Cashel, M. (1979) Cell 16, 111-121.
10. Vanyushin, B.F., Belozersky, A.N., Kokurina, N.A. and Kadirova, D.X. (1968) Nature 218, 1066-1067.
11. Gelinas, R.E., Myers, P.A. and Roberts, R.J. (1977) J. Mol. Biol. 114, 169-180.
12. Boyer, H.W., Chow, L.T., Dugaiczyk, A., Hedgpeth, J. and Goodman, H. M. (1973) Nature New Biology 244, 40-43.
13. Lacks, S. and Greenberg, W. (1975) J. Biol. Chem. 250, 4060-4072.
14. Subak-Sharpe, H., Bürk, R.R., Crawford, L.V., Morrison, J.M., Hay, J. and Keir, H.M. (1966) Cold Spring Harbor Symposium on Quantitative Biology Vol. XXXI, 737-748.
15. Sanger, F., Coulson, A.R., Friedmann, T., Air, G.M., Barrell, B.G., Brown, N.L., Fiddes, J.C., Hutchison III, C.A., Solocombe, P.M. and Smith, M. (1978) J. Mol. Biol. 125, 225-246.
16. Razin, A., Sedat, J.W. and Sinsheimer, R.L. (1970) J. Mol. Biol. 53, 251-259.
17. Lee, A.S. and Sinsheimer, R.L. (1974) J. Virol. 14, 872-877.
18. Sutcliffe, J.G. (1978) Nucleic Acids Res. 5, 2721-2728.
19. Riggs, A.D. (1975) Cytogenet Cell Genet. 14, 9-25.
20. Holliday, R. and Pugh, J.E. (1975) Science 187, 226-232.
21. Bird, A.P., (1978) J. Mol. Biol. 118, 49-60.
22. Razin, A. and Riggs, A.D. submitted for publication.
23. Nesterenko, V.F., Buryanov, Y.I. and Baev, A.A. (1979) Biokhimiya 44, 130-141.
24. Marinus, M.G. and Morris, N.R. (1975) Mutation Research 28, 15-26.
25. Glickman, B., van den Elsen, P. and Radman, M. (1978) Molec. Gen. Genet. 163, 307-312.
26. Wagner, R. Jr. and Meselson, M. (1976) Proc. Nat. Acad. Sci. USA 73, 4135-4139.
27. Meijer, M., Beck, E., Hansen, F.G., Bergmans, H.E.N., Messer, W. von Meyenburg, K. and Schaller, H. (1979) Proc. Nat. Acad. Sci. USA 76, 580-584.
28. Sugimoto, K., Oka, A., Sugisaki, H., Takanami, M., Nishimura, A., Yasuda, Y. and Hirota, Y. (1979) Proc. Nat. Acad. Sci. USA 76, 575-579.
29. Sober, H.A. ed (1970) in CRC Handbook of Biochemistry 2nd edition, page H-117.