J Mol Evol (1985) 21:317–322

Journal of Molecular Evolution

Selection Against *dam* Methylation Sites in the Genomes of DNA of Enterobacteriophages

Michael McClelland*

Department of Biochemistry, University of California, Berkeley, California 94720, USA

Summary. Postreplicative methylation of adenine in *Escherichia coli* DNA to produce G^{6m}ATC (where ^{6m}A is 6-methyladenine) has been associated with preferential daughter-strand repair and possibly regulation of replication. An analysis was undertaken to determine if these, or other, as yet unknown roles of GATC, have had an effect on the frequency of GATC in *E. coli* or bacteriophage DNA. It was first ascertained that the most accurate predictions of GATC frequency were based on the observed frequencies of GAT and ATC, which would be expected since these predictors take into account preferences in codon usage. The predicted frequencies were compared with observed GATC frequencies in all available bacterial and phage nucleotide sequences. The frequency of GATC was close to the predicted frequency in most genes of *E. coli* and its RNA bacteriophages and in the genes of nonenteric bacteria and their bacteriophages. However, for DNA enterobacteriophages the observed frequency of GATC was generally significantly lower than predicted when assessed by the chi square test. No elevation in the rate of mutation of ^{6m}A in GATC relative to other bases was found when pairs of DNA sequences from closely related phages or pairs of homologous genes from enterobacteria were compared, nor was any preferred pathway for mutation of ^{6m}A evident in the *E. coli* DNA bacteriophages. This situation contrasts with that of 5-methylcytosine, which is hypermutable, with a preferred pathway to thymine. Thus, the low level of GATC in enterobacteriophages is probably due not to ^{6m}A hypermutability, but to selection against GATC in order to bypass a GATC-mediated host function.

* *Present address:* Department of Human Genetics and Development, College of Physicians and Surgeons of Columbia University, 701 West 168th St., New York, New York 10032, USA

Key words: *dam* Methylase — 6-Methyladenine — Daughter-strand repair — Mutation — Selection

Introduction

All members of the family Enterobacteriaceae and the genus *Haemophilus* tested contain DNA sequences homologous to the *dam* gene (Brooks et al. 1983) that encodes the *Escherichia coli* G^{6m}ATC-specific methylase (Hattman et al. 1978) (^{6m}A is 6-methyladenine). This methylase is unusual in that it is not associated with a restriction modification system (Smith and Kelly 1983). Furthermore, mutations in the *dam* gene of *E. coli* are viable (Marinus et al. 1983). *dam*-Deficient mutants are, however, hyperrecombinogenic (Korba and Hays 1982) and have a mutator phenotype (Marinus and Konrad 1976). This has been attributed to a lack of strand discrimination in *dam*-methylation-dependent daughter-strand repair (Glickman and Radman 1980; Herman and Modrich 1981; Pukkila et al. 1983). *dam*-Overproducer mutants are also hypermutable (Glickman and Radman 1980; Herman and Modrich 1981), presumably because DNA methylation occurs more rapidly after replication, reducing the opportunity for methylation-mediated strand discrimination. Given this role for *dam* there is no reason to expect an unusual frequency or distribution of GATC sequences in *E. coli* chromosomal DNA. Indeed, the distribution of GATC sequences is fairly random as determined from digestions of *E. coli* total chromosomal DNA with restriction endonucleases that cut GATC (Szyf et al. 1982). However, DNA sequencing has shown that some regions, for instance, the *E. coli* origin of rep-

lication, are significantly enriched in GATC (Meijer et al. 1979; Sugimoto et al. 1979; Zyskind and Smith 1980; Zyskind et al. 1983), suggesting that $G^{6m}ATC$ may be involved in initiation of replication. There are also data that show GATC sequences to occur preferentially at the ends of Okazaki fragments (Gomez-Eichelmann 1979), which suggests a special role for GATC during replication.

The existence of viable dam-deficient mutants argues against a strict requirement for $G^{6m}ATC$ in replication, but leaves open the possibility that essential systems may be GATC specific without absolutely requiring $G^{6m}ATC$ for their function. Alternatively, since mutants made by placing insertions in the dam gene (Marinus et al. 1983) may not be completely without methylase activity, there could be sufficient residual $G^{6m}ATC$ in dam-deficient strains to enable essential $G^{6m}ATC$-dependent functions to be carried out.

Given the special role of GATC, it is probable that regions other than the *E. coli* origin also deviate from predicted GATC frequency as a consequence of their interaction with $G^{6m}ATC$-mediated functions. A survey was undertaken to determine the best predictor of GATC frequency, and the frequencies thus predicted were compared with the observed frequencies of GATC in all available bacterial and phage sequences.

Methods

Local variation in the frequency of GATC was determined by analyzing the number of GATCs in all the sequences in the data base of bacterial and phage nucleotide sequences compiled on GENBANK (Bolt, Beranek, and Newman, Boston, MA). The observed number of GATCs in each sequence was compared with the predicted number.

To determine the best predictor of tetranucleotide frequency seven different methods employing the mononucleotide, dinucleotide, and trinucleotide frequencies in 27 genes from *E. coli* were compared with respect to ability to predict the frequencies of all possible palindromic permutations of G, A, T, and C, i.e., GATC, CATG, GTAC, CTAG, TCGA, TGCA, ACGT, and AGCT. For instance, if p(i) is the frequency of the base or oligonucleotide i in a DNA sequence, equations used to predict GATC frequency were as follows: using mononucleotide frequencies, p(G) × p(A) × p(T) × p(C); using also dinucleotide frequencies, p(G) × p(AT) × p(C), and p(GA) × p(TC); and using also trinucleotide frequencies, p(GAT) × p(C), p(G) × p(ATC), and p(GAT) × p(ATC)/p(AT). Mononucleotide frequencies were a poor predictor of tetranucleotide frequencies (data not shown). The most consistently accurate predictors of all palindromes were the trinucleotide predictors. The predictor used, unless otherwise stated, was p(GAT) × p(ATC)/p(AT). The significance of the difference between observed and predicted numbers of GATC in each sequence was determined by using the chi square test with one degree of freedom.

Computer analysis systems used were GENBANK and INTELLIGENETICS (Palo Alto, CA).

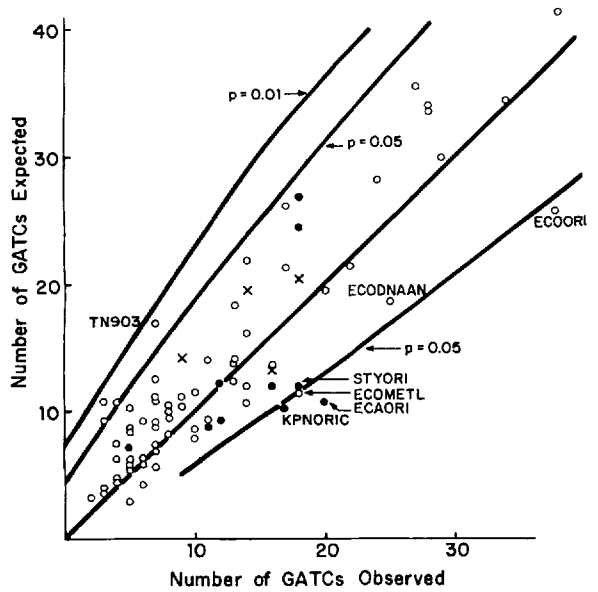

Fig. 1. Frequency of GATC in bacterial DNA. The observed numbers of GATCs in bacterial DNA sequences were compared with the numbers expected from the trinucleotide predictor p(GAT) × p(ATC)/p(AT). This figure includes DNA sequences from 29 species of bacteria. Some species are grouped into genera to increase sample size. Abbreviations are those used by GENBANK. Species for which there is no DNA homology to the *E. coli dam* gene (Hattman et al. 1978) or that do not contain *dam* (Dreiseikelmann and Wackernagel 1981) are indicated by "×"s. Species with DNA homology to the *E. coli dam* gene or that probably contain *dam* are indicated by filled circles. *E. coli* genes are indicated by open circles. Lines connect predicted and observed numbers that have significances of P = 0.01 or P = 0.05 by a chi square test

Results

Unless otherwise indicated, the predictor p(GAT) × p(ATC)/p(AT) was employed in analyses. Using trinucleotide frequencies such as p(GAT) and p(ATC) removes some of the bias due to variation in di- and trinucleotide frequency in prediction of the frequency of tetranucleotides. When di- and trinucleotide frequencies are taken into account in this manner, any residual differences between observed and predicted tetranucleotide frequencies is more likely to be due to effects at the tetranucleotide level than to lower-order effects such as codon usage.

The Frequency of GATC in Bacterial DNA

The total number of GATCs in the *E. coli* chromosomal DNA sample was 767. This was 92% of the number expected from trinucleotide predictors. Di- and mononucleotide predictors gave similar results, but with more scattering (data not shown). Significant deviations from predicted GATC frequency occurred in only a few enterobacterial genes and in no nonenteric sequences (Fig. 1). A region of

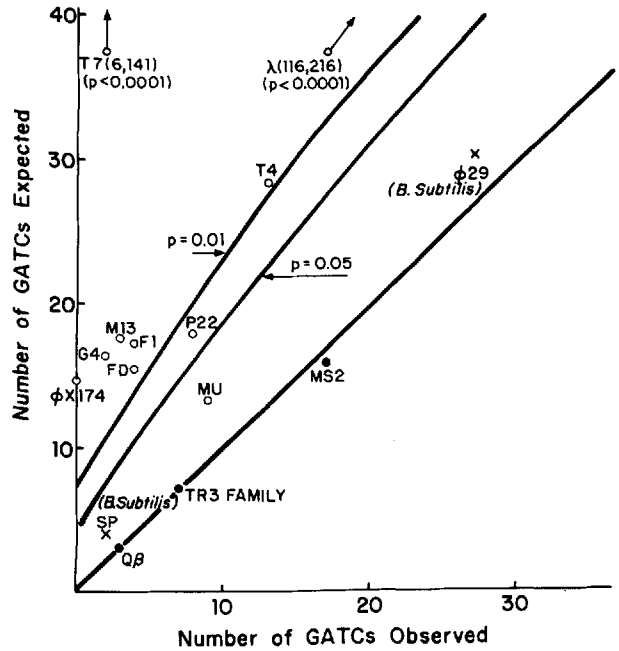

Fig. 2. Frequency of GATC in Bacteriophages. The same predictors were used as in Fig. 1. Open circles represent DNA enterobacteriophages, filled circles represent RNA bacteriophages, and "×"s represent nonenteric bacteriophages. For T4 it should be noted that 19 restriction sites containing GATC have been mapped to this phage (Gram et al. 1984). This compares with a predicted value (based on mononucleotide frequencies) of 79 such sites, giving an observed/expected ratio of about 1:4. Thus GATC is likely to be rarer in T4 than is suggested by the 1:2 observed/expected ratio calculated from the limited sequence data available

Table 1. Fate of GATC in bacteriophage T7

	Predictor				
Sequence	Di-nucl.	Tri-nucl.	XXN-XX	XNX-XX	XXX-NX
GATC	–	–	–	–	–
Transitions					
AATC	+	+	+	+	+
GGTC	+	+	+	+	+
Homotransversions					
CATC	+	+	+	+	+
GTTC	+	+	+	+	+
Heterotransversions					
TATC	+	+	0	0	+
GCTC	+	0	0	0	–

Using the sequence of T7 bacteriophage DNA (Hoffatt et al. 1983), the frequencies of GATC and of each tetranucleotide generated by a single point mutation from GATC were compared with the frequencies calculated using five different predictors. A plus sign indicates that the observed frequency for the tetranucleotide sequence in that row is greater than the frequency predicted by the method indicated at the top of that column. A minus sign represents the converse. The first two predictors were based on di- and trinucleotide frequencies (see Methods). The other three predictors were the frequencies of sequences with bases in the same order as in GATC but noncontiguous. For instance, when reading across the row for GATC, the last three columns, XXNXX, XNXXX, and XXXNX, represent comparisons of the frequency of GATC with those of GANTC, GNATC, and GATNC, respectively, where N is any base

250 base pairs at the *E. coli* origin of replication (ECOORI) contains 11 GATCs (Meijer et al. 1979; Sugimoto et al. 1979; Zyskind and Smith 1980; Zyskind et al. 1983) (Fig. 1). This is an enrichment of more than threefold relative to that predicted from trinucleotide frequencies. The origins of replication of five other related species also contain 11–14 occurrences of the *dam* site within approximately 250 base pairs, seven of which are conserved in all six species (Zyskind et al. 1983). Other palindromes were not conserved in this region. The only other *E. coli* chromosomal gene with an elevated GATC content was the DNA A gene (ECODNAA), which codes for a DNA-binding protein involved in the initiation of DNA replication (Hansen et al. 1982). Eight GATCs were found in 220 base pairs proximal to the gene.

The Frequency of GATC in Bacteriophages

If G^{6m}ATC sequences have no special role outside of Ori C-mediated replication and daughter-strand repair, then no increase or decrease in the frequency of GATC relative to that predicted by di- and trinucleotide frequencies should be found in bacteriophages. However, the frequency of GATC was significantly lower than predicted in most DNA enterobacteriophage sequences examined (Fig. 2). For instance, bacteriophage T7 (40,000 base pairs) (Dunn and Studier 1983) contains 6 GATCs, whereas the predicted number is 141. In contrast, RNA enterobacteriophages (in which the sequence of interest is GAUC) and *Bacillus* DNA phages, which are not methylated at GATC (Dreiseikelmann and Wackernagel 1981), had close to the predicted frequencies of GATC. Thus, GATC is rare only in those DNA phages that infect hosts with *dam* methylases.

The cause of the general rarity of GATC in the DNA enterobacteriophages examined is not obvious. It could be due to selection against GATC, or to some peculiarity of ^{6m}A, for instance, hypermutability. There is a precedent for this in that the other common modified base in bacterial DNA, 5-methylcytosine (^{5m}C), is hypermutable (Coulondre et al. 1978). However, ^{6m}A hypermutability seems unlikely, since the level of GATC is normal in *E. coli* chromosomal DNA. Nevertheless, to test the possibility that ^{6m}A is hypermutable and mutates by a preferred pathway, the observed and pre-

Table 2. Tetranucleotide sequence conservation in pairs of closely related enterobacterial DNAs

Sequence	Palindrome								
	GATC	GTAC	ACGT	AGCT	CATG	CTAG	TCGA	TGCA	Total
Bacteriophage sequences									
F1	1	2	1	2	1	0	2	2	11
Shared	3	16	19	16	12	3	8	10	87
Fd	1	2	1	0	2	0	2	2	10
Trp D									
E. coli	0	2	1	2	1	1	0	1	8
Shared	2	1	2	1	3	0	3	1	13
S. typ.	2	0	1	3	2	0	1	1	10
E. coli	0	2	2	1	1	1	0	0	7
Shared	2	1	1	2	3	0	3	2	14
S. dys.	0	0	0	1	1	0	1	0	3
Trp A and B									
E. coli	4	2	4	5	3	0	4	5	27
Shared	1	1	2	1	2	0	1	2	10
S. typ.	6	6	2	2	7	0	2	1	26
Ori C									
E. coli	0	0	0	0	0	0	0	0	0
Shared	13	0	0	2	0	0	0	0	15
S. typ.	0	1	0	1	0	1	0	0	3

Available bacterial and bacteriophage DNA sequences that showed a high level of homology were aligned and the numbers of occurrences of GATC and the seven other palindromes containing G, A, T, and C were determined for each sequence. If a palindrome was found at the same position in both sequences, it was placed in the "Shared" category, and not included in the count for the single sequence. If a palindrome was found in only one sequence, the mutation involved was determined (Table 3). Phages compared were F1 and Fd, and species compared were *E. coli, Salmonella typhimurium* (*S. typ.*) and *Shigella dysenteriae* (*S. dys.*)

Table 3. Mutations involved in the loss or gain of GATC

Mutation in GATC involving		Base in aligned sequence				
		A	G	T	C	Total
Internal base (A or T)	A	—	5	1	2	8
External base (G or C)	G	4	—	1	1	6

The mutations involved in the losses and gains of GATC shown in Table 2 were determined. The data are condensed such that a mutation involving C or T in GATC is represented by the G or A, respectively, of the complementary strand. For instance, the presence of a GATC in one sequence and a GAGC in the aligned sequence is counted as an A ↔ C mutation. The directions of mutation (i.e., loss or gain of GATC) could not be determined

dicted frequencies of all sequences differing from GATC by a single-base substitution were determined for bacteriophage T7. No preferential pathway for the elimination of GATC was detected (Table 1). The frequencies of all sequences differing from GATC by a single base except the heterotransversions TATC and GCTC were elevated. In particular, the increase in frequency of the sequences in which the base difference occurred in one of the outer bases (G or C) was similar to that of the sequences in which the base difference occurred in one of the central bases (A or T). Similar results were obtained for the phages phiX174, F1, and lambda (data not shown). However, the initial loss of GATC in these phages may have been quite ancient and thus obscured by multiple hits. To rule out this possibility, recently diverged sequences from the Fd bacteriophages were aligned with those of F1; other sequenced bacteriophages were not sufficiently homologous for meaningful comparison of the fates of GATC in their sequences. For each pair of aligned sequences, the numbers of GATCs and of the seven other tetranucleotide palindromes containing G, A, T, and C were determined. Where a tetranucleotide was not conserved between related sequences, the base change involved was determined. The direction of the mutation that resulted in the base change could not be determined. Since any mutation preference of ^{6m}A would be expected to occur in chromosomal as well as in bacteriophage DNA, the analysis was also performed for the known DNA sequences of the Trp A, B, and D genes and the origins of replication of enterobacteria. The data, summarized in Table 2, suggest that there is at most only a slight difference between the rate of mutation of GATC and those of other tetranucleotide palindromes. Excluding the origins of replication, there were 14 mutations of GATC in 22 occurrences, a ratio of 1:1.6, compared with 113 mutations of other palindromes in 226 occurrences, a ratio of 1:2. Fur-

thermore, 8 of 14 mutations involved the central AT of GATC (Table 3), which is not a significant bias in favor of mutation of ^{6m}A or T relative to mutation of G or C. In the origins of replication, GATCs were not only extremely abundant, but, far from being hypermutable, were in fact highly conserved, as one would expect if they have an important function.

Discussion

Evidence for Selection against GATC in DNA Enterobacteriophages

The *dam* methylation target GATC is rare in most DNA enterobacteriophages. However, evidence has been presented here that ^{6m}A is not hypermutable, eliminating a possible explanation for this rareness. It is possible that selection is occurring against the recognition sequence of a GATC-specific restriction endonuclease. Precedents for this occur in the *B. subtilis* phages phi29 and SPO1, in which the sequence GGCC is absent or rare (Ito and Roberts 1979; Reeve et al. 1980), presumably because of selection against Bsu RI, a common GGCC-specific restriction system in *B. subtilis*. Similarly, Eco RI, Eco K, and Eco B recognition sequences are two- to fourfold rarer than predicted in enterobacteriophage DNA (data not shown). However, there is no evidence for a GATC-specific restriction system in enterobacteria (Smith and Kelly 1983). Thus, it is more likely that the hypothesized selective force arises due to a need to bypass the known role of G^{6m}ATC in repair, recombination, or replication (Coulondre et al. 1978; Glickman and Radman 1980; Herman and Modrich 1981; Marinus et al. 1983; Pukkila et al. 1983). For instance, selection could be a consequence of the need for rapid (low fidelity) replication or unusual replication modes such as the rolling circle. RNA enterobacteriophages and non-enteric bacteriophages do not encounter *dam* methylation, and thus these phages have normal levels of GATC.

Some properties of certain DNA enterobacteriophages may reduce selection against GATC. For instance, the lysogenic DNA phages have higher levels of GATC than the lytic DNA phages (Fig. 2). Perhaps this is because in their quiescent state lysogenic phages are replicated and repaired many times under host control and at a rate similar to that of normal host DNA. This could require maintaining a level of GATC similar to that of the host. Another modulator may be heavy modification of bacteriophage DNA, as occurs in T4 and which renders the DNA resistant to *dam* methylation (Hattman 1970). This could reduce the resultant selection pressure against GATC by disrupting the interaction responsible for this pressure. A consequence of a lack of *dam* methylation in T4 is that this bacteriophage may not undergo daughter-strand repair (Sechaud et al. 1965; Sinha and Goodman 1983). Next, there is some evidence that phage-encoded mechanisms may inhibit *dam* methylation of phage DNA in some cases (Dreiseikelmann et al. 1979; Bauer et al. 1981; Lundahl 1982). Finally, phage mu contains a gene that is positively regulated by *dam* methylation of its promoter (Hattman 1979, 1982; Plasterk et al. 1984). Clearly, there should be selection to maintain the regulatory *dam* sites in such a phage.

It is not clear whether selection against GATC in enterobacteriophages is due to selection against G^{6m}ATC. To test this it should be possible to vary the number of GATCs in DNA enterobacteriophages by in vitro mutagenesis. Loss or gain of GATC sequences can be achieved without altering encoded amino acids because of the degeneracy of the genetic code. The possibility of selection against GATC can then be investigated by assaying the loss of GATC sequences after many infective cycles or by competing GATC-rich phage with wild-type or GATC-depleted phage by means of mixed infection. Thus, the effect of ^{6m}A could be determined by comparing phage growth and competition experiments using *dam*$^+$ cells with similar experiments using *dam*$^-$ cells.

The availability of plasmid clones of adenine-specific restriction methylases such as M.Eco RI (which produces GA^{6m}ATTC), M.Hpa I (A^{6m}AGCTT), and M.Taq I (TCG^{6m}A) (McClelland 1983) should allow the specificity of the ^{6m}A-dependent systems to be detemined. If *dam*$^-$ *E. coli* mutants are rescued or some phenotypic effects partly relieved by other sequence-specific adenine methylases, this would suggest that ^{6m}A need not be in the sequence GATC to perform at least some aspects of its function. Furthermore, selection against adenine-specific restriction methylase recognition sequences other than G^{6m}ATC may be tested in DNA enterobacteriophages grown on *dam*$^-$ restriction-methylase-positive *E. coli*.

Conclusions

G^{6m}ATC is apparently not hypermutable, and GATC occurs at predicted frequencies in most *E. coli* chromosomal genes. However, GATC is rare in some DNA enterobacteriophages. This is probably due to selection against GATC in order to circumvent a GATC- or ^{6m}A-mediated host function involved in restriction, daughter-strand repair, replication, or recombination. Partial relief from this hypothesized selection pressure against GATC may occur in lysogenic phages, the DNAs of which must more

closely resemble those of their hosts, and in bacteriophages that undergo extensive DNA modification.

Acknowledgments. I thank John Hall of the Atlanta University System for giving me access to GENBANK; Allan Wilson and Gerry Rubin for use of computer facilities; and Charles Cantor, Randy Morse, Gino Cortopassi, Russ Higuchi, Barbara Chapman, and Beth Stewart for critical reading of the manuscript. This work was carried out in Allan Wilson's lab. The author was supported by an SERC, NATO Postdoctoral Fellowship.

References

Bauer J, Krammer G, Knippers R (1981) Asymmetric repair of bacteriophage T7 heteroduplex DNA. Mol Gen Genet 181: 541–547

Brooks JE, Blumenthal RM, Gingeras TR (1983) The isolation and characterization of the *E. coli* DNA adenine methylase (Dam) gene. Nucleic Acids Res 11:837–851

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hot spots in *E. coli.* Nature 274:775–780

Dreiseikelmann B, Wackernagel W (1981) Absence in *Bacillus subtilis* and *Staphylococcus aureus* of the sequence specific DNA methylase that is conferred in *E. coli* K-12 by the *dam* and *dcm* enzymes. J Bacteriol 147:259–261

Dreiseikelmann B, Eichenlaub R, Wackernagel W (1979) The effect of differential methylation by *E. coli* of plasmid DNA and phage T7 DNA on the cleavage by restriction endonuclease *Mbo* I. Biochim Biophys Acta 562:418–428

Dunn JJ, Studier FW (1983) The complete nucleotide sequence of bacteriophage T7 DNA and the location of the T7 genetic elements. J Mol Biol 166:477–535

Glickman BW, Radman M (1980) *E. coli* mutator mutants deficient in methylation instructed DNA mismatch correction. Proc Natl Acad Sci USA 77:1063–1067

Gomez-Eichelmann MC (1979) DNA adenine and cytosine methylation in *Salmonella typhimurium* and *S. typhi.* J Bacteriol 140:574–579

Gram H, Liebig H-D, Hack A, Niggemann E, Ruger W (1984) The physical map of bacteriophage T4 including the positions of promoters and terminators recognized *in vitro.* Mol Gen Genet 194:232–240

Hansen EB, Hansen FG, Von Meyenberg K (1982) the nucleotide sequence of the DNA A gene and the first part of the DNA N gene of *E. coli.* Nucleic Acids Res. 10:7373–7385

Hattman S (1970) DNA methylation of T-even bacteriophage and of their nonglucosylated mutants—its role in P1 directed restriction. Virology 42:359–367

Hattman S (1979) Unusual modification of bacteriophage Mu DNA. J Virol 32:468–475

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

Hattman S, Brooks JE, Masurekar M (1978) Sequence specificity of the P1-modification methylase (*M.Eco* PI) and DNA methylase (*M.Eco dam*) controlled by the *E. coli dam* gene. J Mol Biol 126:367–380

Herman GE, Modrich P (1981) *E. coli* K-12 clones that overproduce *dam* methylase are hypermutable. J Bacteriol 145: 644–646

Ito J, Roberts RJ (1979) Unusual base sequence arrangement in phage phi-29 DNA. Gene 5:1–7

Korba BE, Hays JB (1982) Novel mutations of *E. coli* that produce recombinogenic lesions in DNA. 5. Recombinogenic plasmids from *Arl* mutants of *E. coli* are unusually sensitive to S1 and partially deficient in cytosine methylation at CC(A/T)GG. J Mol Biol 157:213–235

Lundahl T (1982) DNA repair enzymes. Annu Rev Biochem 51:61–87

Marinus MG, Konrad B (1976) Hyper-recombinogenic *dam* mutants of *E. coli* K-12. Mol Gen Genet 149:273–277

Marinus MG, Morris NR (1974) Biological function for 6mA residues in DNA of *E. coli* K-12. J Mol Biol 85:309–322

Marinus MG, Carraway M, Frey AZ, Brown L, Arraj JA (1983) Insertion mutations in the *dam* gene of *E. coli* K-12. Mol Gen Genet 192:288–289

McClelland M (1983) The effect of site specific methylation on restriction endonuclease cleavage (update). Nucleic Acids Res 11:r169–r173

Meijer M, Beck E, Hansen FG, Bergmans HEN, Messer W, Von Mayenburg K, Schaller H (1979) Nucleotide sequence of the origin of replication of the *E. coli* K-12 chromosome. Proc Natl Acad Sci USA 76:580–584

Plasterk RHA, Vollering M, Brinkman A, van de Putte P (1984) Analysis of the methylation-regulated mu *mom* transcript. Cell 36:189–196

Pukkila PJ, Peterson J, Herman G, Modrich P, Meselson M (1983) Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *E. coli.* Genetics 104: 571–582

Reeve JN, Amann E, Tailor R, Gunthert U, Scholz K, Trautner TA (1980) Unusual behaviour of SPO1 DNA with respect to restriction and modification enzymes recognizing the sequence GGCC. Mol Gen Genet 178:229–231

Sechaud J, Strisinger G, Emrich J, Newton H, Langford H, Reinhold H, Stahl M (1965) Chromosome structure and phage T4. II. Terminal redundancies in heterozygotes. Proc Natl Acad Sci USA 54:1333–1339

Sinha NK, Goodman MF (1983) Fidelity of DNA replication. In: Mattews CK, Kutter EM, Mosig G, Berget PB (eds) Bacteriophage T4. American Society of Microbiology, Washington DC, pp 131–137

Smith HO, Kelly SV (1983) Methylases of the type II. Restriction-modification system. In: Razin A, Cedar H, Riggs AD (eds) DNA methylation and its biological significance. Springer-Verlag, New York, pp 39–72

Sugimoto K, Oka A, Sugisaki H, Takanami M, Nishimura A, Yasuda S, Hirota A (1979) Nucleotide sequence of the *E. coli* K-12 replication origin. Proc Natl Acad Sci USA 76:575–579

Szyf M, Gruenbaum Y, Urieli-Shoval S, Razin A (1982) Studies on the biological role of DNA methylation. 5. The pattern of *E. coli* DNA methylation. Nucleic Acids Res 10:7247–7259

Zyskind JW, Smith DW (1980) Nucleotide sequence of the *Salmonella typhimurium* origin of DNA replication. Proc Natl Acad Sci USA 77:2460–2464

Zyskind JW, Cleary JM, Brusilow WS, Harding NE, Smith DW (1983) Chromosomal replication origin from the marine bacterium *Vibrio harveyi* functions in *E. coli–Ori C* consensus sequence. Proc Natl Acad Sci USA 80:1164–1168

Received November 1, 1984/Revised November 15, 1984