JOURNAL OF BACTERIOLOGY, Nov. 1988, p. 5177–5184
0021-9193/88/115177-08$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 170, No. 11

Molecular Cloning, Sequencing, and Mapping of the Bacteriophage T2 *dam* Gene

ZOE MINER AND STANLEY HATTMAN*

Department of Biology, University of Rochester, Rochester, New York 14627

Received 30 June 1988/Accepted 10 August 1988

Bacteriophage T2 codes for a DNA-(adenine-N^6)methyltransferase (Dam), which is able to methylate both cytosine- and hydroxymethylcytosine-containing DNAs to a greater extent than the corresponding methyltransferase encoded by bacteriophage T4. We have cloned and sequenced the T2 *dam* gene and compared it with the T4 *dam* gene. In the Dam coding region, there are 22 nucleotide differences, 4 of which result in three coding differences (2 are in the same codon). Two of the amino acid alterations are located in a region of homology that is shared by T2 and T4 Dam, *Escherichia coli* Dam, and the modification enzyme of *Streptococcus pneumoniae*, all of which methylate the sequence 5′ GATC 3′. The T2 *dam* and T4 *dam* promoters are not identical and appear to have slightly different efficiencies; when fused to the *E. coli lacZ* gene, the T4 promoter produces about twofold more β-galactosidase activity than does the T2 promoter. In our first attempt to isolate T2 *dam*, a truncated gene was cloned on a 1.67-kilobase *Xba*I fragment. This construct produces a chimeric protein composed of the first 163 amino acids of T2 Dam followed by 83 amino acids coded by the pUC18 vector. Surprisingly, the chimera has Dam activity, but only on cytosine-containing DNA. Genetic and physical analyses place the T2 *dam* gene at the same respective map location as the T4 *dam* gene. However, relative to T4, T2 contains an insertion of 536 base pairs 5′ to the *dam* gene. Southern blot hybridization and computer analysis failed to reveal any homology between this insert and either T4 or *E. coli* DNA.

The *dam* [DNA-(adenine-N^6)methyltransferase (Dam)] genes of bacteriophages T2 and T4 code for a protein that methylates adenine in the sequence 5′ GATC 3′ or 5′ GAThmC 3′ (12, 32, 33) by using *S*-adenosylmethionine as the methyl donor (14). In addition, both methyltransferases recognize additional sequences belonging to a subset of GA(T/C) (3, 5, 12; S. Schlagman, Z. Miner, Z. Fehér and S. Hattman, in press). However, T2 Dam appears to have a more relaxed specificity than T4 Dam because it attains higher levels of methylation than T4, both on its normal substrate, hydroxymethylcytosine (hmC)-containing DNA (hmC-DNA), and on cytosine-containing DNA (C-DNA) (10; S. Schlagman, Z. Miner, and S. Hattman, manuscript in preparation). Mutants (dam^h) exist in both T2 and T4 which hypermethylate DNA (3, 10, 16); however, T2 Dam^h methylates virion DNA to a greater extent than does the T4 Dam^h (10).

The T4 *dam* gene has been cloned (32) and sequenced (22; Z. Miner, S. Schlagman, and S. Hattman, manuscript in preparation). The nucleotide sequence data revealed two overlapping, in-phase open reading frames encoding polypeptides of 26 and 30 kilodaltons, respectively. Recent results have identified the latter as the Dam polypeptide (Schlagman et al., in press). This is in contrast to the M_r of about 14,000 determined for T2 Dam by sedimentation velocity centrifugation in a glycerol gradient (13). In addition, the T2 *dam* gene had been genetically mapped between genes 49 and rI (2); this is significantly different from the site where the T4 *dam* gene was genetically and physically mapped (22, 32; M. Myers, S. Schlagman, and S. Hattman, unpublished observations). Therefore, it was of considerable interest to be able to compare the nucleotide and encoded amino acid sequences of the T4 *dam* and T2 *dam* genes. In this paper we report the cloning and sequencing of the T2 *dam* gene and show that there are only three amino acid differences between T2 and T4 Dam. Two of these differences are located in a region of amino acid sequence homology shared by T4 and several other Dam proteins (13, 25); namely, the Dam of *Escherichia coli* and the Dam of *Streptococcus pneumoniae*, both of which also recognize the sequence GATC (8, 12, 26), and the M · *Eco*RV modification methyltransferase.

We also present genetic and physical evidence that places the T2 *dam* gene at the same relative location in its genome as T4 *dam*; i.e., adjacent to the overlapping genes 56 and 69 (G. Mosig, personal communication). However, there is a 536-base-pair (bp) insertion 5′ to the T2 *dam* gene. This insertion appears to have no homology to either T4 DNA or *E. coli* DNA, as shown by Southern hybridization. A computer search comparing this nucleotide sequence with T4 and *E. coli* sequences contained in GenBank failed to reveal any significant homology. This insert could account for a region of nonhomology near the *dam* gene as revealed by electron microscope examination of heteroduplexes between T2 and T4 DNAs (19, 35).

MATERIALS AND METHODS

Bacteria, bacteriophage, and plasmids. The following *E. coli* strains were used: K704 (F^- *rglA rglB met gal supII* r_K^+ m_K^+); SK1036 *dam-4* (18); GM119 *dam-3* (1); F^+1100 (29); JM105 (34); GM2971 [F^- *mrr hsdS20* r_B^- m_B^- *ara-14 proA2 lacYl galK2 rpsL20* (*strR*) *xyl-5 mtl-1 supE44 dam-13*::Tn9 Cm], from M. G. Marinus; Bc48 *rgl*, from our collection; MC1061 (28). The following bacteriophages were used: T2 αgt_1; T4 αgt_1 βgt_{27}; T6 αgt_2; T2 αgt_1 dam^h *am*61 (gene 56), made in our laboratory by genetic crosses; T2L *am*61, from J. Wiberg; T4 *am*N55X5 (gene 42) *amE*51 (gene 56) *nd*28 (gene *denA*) D2a2 (gene *denB*) *alc*10 (T4c819), from J. Wiberg. Plasmid pMAC19 was from G. Mosig (22), and plasmid pNM480 was from N. Minton (28).

Chemicals and media. TAE, TE, nick translation buffer, LB medium, and M9 minimal medium were described pre-

* Corresponding author.

viously (24). The M9 minimal medium was supplemented with 0.2% glucose and 5 μg of thiamine hydrochloride per liter. Ampicillin was added to media to 50 μg/ml, and 2-aminopurine was added to 400 μg/ml. Restriction enzymes, T4 DNA ligase, bacterial alkaline phosphatase, and the DNA polymerase I Klenow fragment were from Bethesda Research Laboratories, Inc., New England BioLabs, Inc., or Boehringer Mannheim Biochemicals. Agarose was ultra pure DNA grade from Bio-Rad Laboratories. Nitrocellulose was from Schleicher & Schuell Co. RNase A, DNase I, and 2-aminopurine were from Sigma Chemical Co. Nick translation kit 5000 and radioactive labeled nucleotides were from Amersham Corp. The specific activity of [α-^{35}S]dATP was 600 Ci/mmol. The specific activity of [α-^{32}P]dCTP was 400 Ci/mmol. Bethesda Research Laboratories was the supplier of the dideoxy sequencing materials. Phage buffer was 10 mM Tris hydrochloride (pH 7.5), 0.01% (wt/vol) gelatin, 0.1 M NaCl, and 5 mM $MgSO_4$.

General methods. Restriction enzyme digestions and bacterial alkaline phosphatase reactions were performed to the specifications of the suppliers. DNA ligation reactions were carried out overnight at room temperature in a 20-μl volume. Competent cells for transformation were made by using the $CaCl_2$ method (23), except that cells were used the same day as they were made competent. Isolation of DNA fragments was from low-gelling-point agarose (SeaPlaque; FPC Biochemicals) followed by chromatography with an NACS Prepac column (Bethesda Research Laboratories) or Gene Clean (Bio 101) as recommended by the manufacturers.

Plasmid DNA purification. Large-scale preparations were performed as described by Maniatis et al. (24). Lysis was by the alkali method, and closed circular DNA was purified with ethidium bromide-cesium chloride gradients. Ethidium bromide was removed from DNA by extraction with water-saturated 1-butanol, and DNA was dialyzed extensively against TE.

Small-scale plasmid DNA purification was by the alkali lysis method as previously described (24).

Phage DNA isolation. *E. coli* K704 was grown to 2×10^8 cells per ml in LB medium and infected with T2 αgt, T4 αgt βgt, or T6 αgt phage at a multiplicity of infection of <1. The infected culture was incubated with vigorous aeration at 37°C for about 5 h while the culture was monitored for cell lysis. At this time $CHCl_3$ was added, and the lysate was centrifuged at 5,900 × *g* for 10 min at 4°C. The supernatant was removed and NaCl (0.5 M) and polyethylene glycol 8000 (7%, wt/vol) were added. After stirring at 4°C overnight, the phages were collected by centrifugation at 12,000 × *g* for 15 min. The pellet was suspended in phage buffer in 1/50 the volume of the original culture. An equal volume of $CHCl_3$ was added, and the culture was centrifuged at 12,000 × *g* for 15 min. RNase A and DNase I (15 μg/ml) were added to the aqueous layer, and the culture was incubated at 37°C for 1 h. After centrifugation for 15 min at 12,000 × *g*, the supernatant was collected and centrifuged at 31,000 × *g*. The pellet was suspended gently in phage buffer in 1/100 the original volume. The low- and high-speed spins were repeated. Deproteinization was by three water-saturated phenol extractions followed by one chloroform-isoamyl alcohol (24:1) extraction. Finally, the phage DNA was dialyzed overnight against 10 mM Tris hydrochloride (pH 8.0)–0.1 mM disodium EDTA (pH 8.0). The A_{260}/A_{280} ratio was 1.8.

Southern hybridization. Restriction digests of phage DNA were electrophoresed overnight at 60 V in a 0.8% agarose gel in TAE buffer. The gel was photographed and then soaked sequentially for 30 min at room temperature in 250 ml of (i) 0.25 M HCl, (ii) 0.5 M NaOH–1 M NaCl, and (iii) 0.5 M Tris hydrochloride (pH 7.5)–3 M NaCl. The treated gel was then blotted overnight by capillary action onto nitrocellulose. The blot was washed in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) for 10 min, air dried, and baked for 2 h at 80°C in a vacuum oven. Prehybridization and hybridization were done by the methods of Maniatis et al. (24).

Colony hybridization. Transformant colonies were suspended in 100 μl of 0.85% saline in microdilution wells and stamped onto a nitrocellulose filter that had been placed on an LB-ampicillin plate. The master plate was made by stamping onto LB-ampicillin agar. The procedure was continued as described by Maniatis et al. (24).

Nick translation. The instructions of the supplier were followed to achieve a specific activity of 10^8 dpm/μg of probe DNA with [α-^{32}P]dCTP as the labeled nucleotide (incorporation was 60 to 70%).

Dideoxy sequencing. *E. coli* JM105 was routinely maintained in supplemented M9 minimal medium. Preparation of single-stranded DNA templates, sequencing reactions, and gel electrophoresis were essentially as described by Sanger et al. (30) and as described in the M13 cloning-dideoxy sequencing instruction manual from Bethesda Research Laboratories. [α-^{35}S]dATP was used as the labeled nucleotide. Sequence computer analysis was carried out by using the sequence analysis software from the University of Wisconsin Genetics Computer Group and DNA Inspector II (Textco).

Isolation of ^{3}H-labeled phage DNA. ^{3}H-labeled phage DNA was isolated as described previously (10), except that *E. coli* GM2971 containing a plasmid to be tested for its ability to methylate hmC-DNA was grown to 3×10^8 cells per ml and infected with T6 αgt at a multiplicity of infection of 5 to 10.

Isolation of ^{3}H-labeled *E. coli* DNA. *E. coli* SK1036 containing a plasmid to be tested for its ability to methylate C-DNA was grown to 2×10^8 cells per ml. A 5-ml sample of this culture was transferred to a flask containing 0.1 mCi of [2-^{3}H]adenine and grown for 2 h at 37°C. Then 25 ml of ice-cold 0.85% saline was added, and the mixture was centrifuged for 10 min at 12,000 × *g*. The pellet was suspended in 3 ml of 0.85% saline, harvested at 12,000 × *g* for 10 min, and suspended in 2.5 ml of 1× SSC and 0.05 ml of 20% (wt/vol) Sarkosyl (CIBA-GEIGY Corp.). Lysis was completed at 37°C for 15 min. The DNA was isolated by phenol extraction, ethanol precipitation, and low-speed centrifugation. The DNA pellet was dissolved in 2.1 ml of 0.2 M NaOH and placed at 37°C for about 20 h. Then 1 ml of 3 M HCl and 0.3 ml of 70% perchloric acid were added, and the reaction was placed on ice for 30 min. After centrifugation at 17,000 × *g* for 30 min, the pellet was washed three times with 7% perchloric acid, dried, and dissolved in 1.5 ml of 1 M HCl.

Depurination of ^{3}H-labeled DNA and descending paper chromatography. The procedures for depurination of ^{3}H-labeled DNA and descending paper chromatography were those of Hattman (10).

β-Galactosidase assays. β-Galactosidase was assayed by the method of Miller (27).

Two-factor genetic cross: T2 αgt dam^+ 56^+ × T2 αgt dam^h 56. *E. coli* K704 was grown to 2×10^8 cells per ml and infected with each parental phage at an input ratio of 5. The culture was incubated with agitation at 30°C for 5 min and then centrifuged at 5,900 × *g* for 5 min to remove unadsorbed phage. The pellet was suspended in LB medium and agitated vigorously at 30°C for 90 min. Chloroform was added, and the lysed culture was centrifuged at 5,900 × *g* for

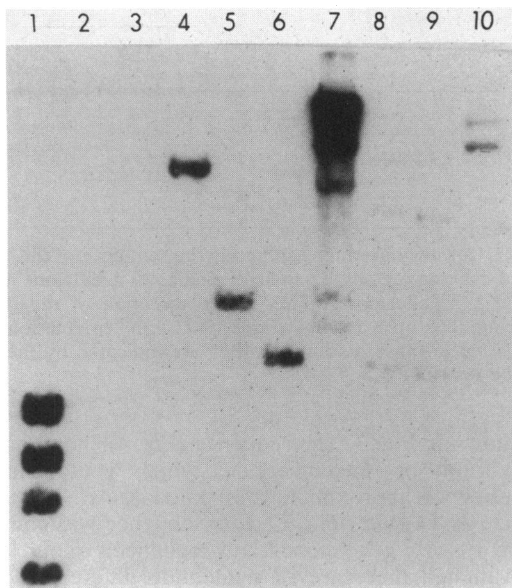

FIG. 1. Southern blot analysis of T2 αgt, T4 αgt βgt, and T6 αgt genomic digests hybridized with a T4 *dam* probe. Lanes (digests): 1, labeled PhiX174 *Hae*III markers; 2, T6 *Xba*I; 3, T6 *Eco*RI; 4, T4 *Eco*RI; 5, T4 *Xba*I; 6, T4 *Eco*RI-*Xba*I; 7, labeled *Hin*dIII-digested lambda markers; 8, T2 *Eco*RI-*Xba*I; 9, T2 *Xba*I; 10, T2 *Eco*RI.

5 min. The supernatant was diluted in 0.85% saline and plated on *E. coli* K704. Individual plaques were picked with toothpicks, suspended in 0.85% saline in microdilution wells, and stamped onto plates containing *E. coli* K704, K704(P1), or Bc48 as the indicator host. All phages grow on K704, only phage with the dam^h allele grow on K704(P1), and *56* phage are unable to grow on Bc48. Thus, all the genotypes could be deduced from the growth pattern on these strains. Both parental phages were αgt so that we could distinguish between dam^+ and dam^h. The *56* defect was confirmed by dCTPase assays of extracts made from sup^0 or *supE* cells infected with T2 αgt dam^h *56*.

RESULTS

Identification of the T2 *dam* gene by Southern hybridization. The T4 *dam* gene was cloned on a *Hin*dIII fragment (32) obtained from a multiple T4 mutant that produces C-DNA. Because the corresponding mutant does not exist in T2, it was necessary to clone from hmC-DNA. Both *Eco*RI and *Xba*I are able to cleave the nonglucosylated hmC-DNA produced by T2 αgt mutants; these two enzymes are also advantageous in that an intact T4 *dam* gene is included on a 2.1-kilobase (kb) *Xba*I fragment, as well as on a 1.8-kb *Xba*I-*Eco*RI fragment (22, 32). On the chance that the T2 *dam* and T4 *dam* genes share some homology, we first screened T2 αgt genomic digests with a T4 *dam* probe. The T4 *dam* probe was homologous to sequences in T2 (Fig. 1, lanes 8 and 9), but the hybridization pattern was quite different than that with T4 genomic digests (lanes 4 and 5). Although digestion was incomplete, it was still evident that the T2 *dam* gene was located on a 6.6-kb *Eco*RI fragment; the *Xba*I and *Xba*I-*Eco*RI digestions show two fragments, both 4.2 and 1.7 kb, hybridizing to the T4 *dam* probe. These results suggest that, unlike T4, the T2 *dam* gene contains an internal *Xba*I site; furthermore, either T2 *dam* is at a different map location than T4 *dam*, or mutations have altered flanking restriction sites.

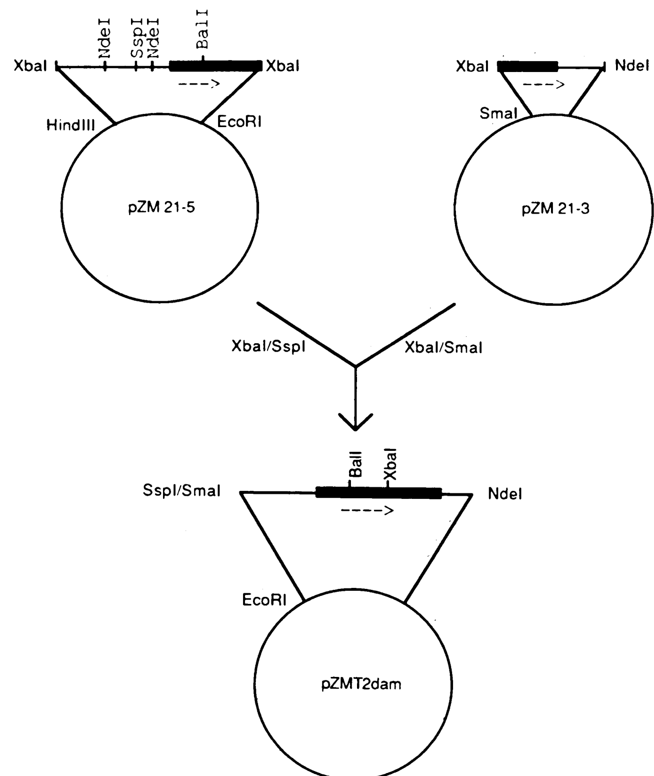

FIG. 2. Construction of pZMT2dam, containing the entire T2 *dam* gene, from pZM21-5 and pZM21-3, containing the 5′ and 3′ regions, respectively, of the gene. Plasmid pZM21-3 was digested with *Xba*I and *Sma*I, and the *Ssp*I-*Xba*I fragment of pZM21-5 containing the 5′ end of the gene was ligated in place. The resulting plasmid, pZMT2dam, produced a Dam that was active on both C-DNA and hmC-DNA.

It should be noted that the T4 *dam* probe did not hybridize to a genomic digest of T6 αgt DNA (Fig. 1, lanes 2 and 3), even at low stringency. This is consistent with the fact that T6 does not produce any Dam activity (7) and suggests that the gene either is missing in T6 or has diverged extensively from T4 *dam*. We have also observed that there is no nucleotide sequence homology between the phage T1 *dam* and T4 *dam* genes (data not shown). However, T1 does encode a Dam which recognizes the sequence GATC (31).

Cloning the T2 *dam* gene. After agarose gel electrophoresis of an *Xba*I T2 αgt genomic digest, the 1.7-kb fragment was purified (see Materials and Methods) and ligated into the pUC18 *Xba*I site. This construct is designated pZM21-5 (Fig. 2). When transformed into *E. coli* SK1036 *dam*, which is unable to grow on media containing 2-aminopurine because of the *dam* mutation (9), ampicillin-resistant transformant cells also acquired resistance to 2-aminopurine. Plasmid pZM21-5 DNA isolated from SK1036 was cleaved by *Dpn*I but not by *Mbo*I (data not shown); this is diagnostic of DNA containing N^6-methyladenine (m^6A) in the sequence GATC. Furthermore, in the presence of pZM21-5, *E. coli* SK1036 cellular DNA contained 1.9 mol% m^6A (Table 1). This level of methylation is similar to what has been observed with the cloned T4 dam^+ gene (32). However, pZM21-5 is unable to methylate the hmC-DNA of T6 αgt phage grown in these cells. These unexpected observations are now understood in the light of the DNA sequence analysis (see below); pZM21-5 contains a truncated *dam* gene corresponding to the coding

TABLE 1. Moles percent m6A in cellular DNA or T6 αgt DNA after propagation in *E. coli dam* cells containing various plasmids[a]

Plasmid[b]	*dam* gene[c]	m6A (mol%) in: Cellular C-DNA	m6A (mol%) in: Virion hmC-DNA
pZM21-5	T2-pUC chimera	1.9	0.02
pZMT2dam	T2 *dam*+	4.0	1.1
pMAC19	T4 *dam*+	2.6	0.33
pBR322	None	0.14	0.03

[a] *E. coli* GM2971 or GM119 cells containing various plasmids were grown and infected with T6 αgt (or not infected) as described in Materials and Methods. Cellular DNA from uninfected cells or T6 αgt virion DNA from infected cells was labeled and purified, and the moles percent m6A was determined as described in Materials and Methods.

[b] With the exception of pBR322, the plasmids are pUC derivatives.

[c] The origin of the *dam* gene carried on the resident plasmid is indicated.

region for the N-terminal 163 amino acids of T2 Dam. The reading frame continues across the *dam*-pUC18 vector junction and encodes an additional 83 amino acids. Thus, this hybrid reading frame encodes a chimeric protein with only partial Dam activity; i.e., it methylates only C-DNA, probably in GATC sequences only. Although the amino acid sequence of the remaining carboxy-terminal portion of T4 Dam is quite different than that of the chimeric protein, computer analysis shows similar hydropathy profiles for the two (data not shown).

Plasmid pZM21-5 was linearized with *Eco*RI, and the ends were filled in and religated. This procedure introduced a nonsense codon 9 amino acids downstream of the *Xba*I site; the resulting plasmid did not produce an active Dam (data not shown). Therefore, the carboxy-terminal region encoded by pUC18 is essential for activity of the chimera. In addition, when the orientation of the *Xba*I fragment in pZM21-5 was reversed, so that the chimera now contained a different set of amino acids at the carboxy terminus, Dam activity was lost.

Numerous attempts were made to clone the 4.2-kb *Xba*I fragment of T2, but none proved successful. We thought that some lethal function may be encoded by this fragment of DNA. Therefore, to retrieve the 3′ end of the T2 *dam* gene, a second strategy was employed. Southern hybridization to an *Nde*I T2 genomic digest with a T4 *dam* probe showed that the entire gene appeared to be contained on a fragment of approximately 900 bp (data not shown). From the known sequence of pZM21-5 (see Fig. 4), we deduced that the 3′ end of the gene should be contained on an *Xba*I-*Nde*I fragment of approximately 350 bp. After such a double digestion of T2 αgt genomic DNA, all fragments less than 500 bp were excised and eluted from an agarose gel and then ligated into the pUC18 *Xba*I-*Nde*I fragment. Transformants were screened for their insert by colony hybridization by using a probe that contained only the 3′ end of the T4 *dam* gene; three clones gave a positive signal. One clone was analyzed and shown to contain the 3′ end of the T2 *dam* gene. The plasmid was designated pZM21-3 (Fig. 2).

To reconstruct the entire T2 *dam*+ gene, pZM21-3 was cut with *Xba*I and *Sma*I, and an 800-bp *Ssp*I-*Xba*I fragment from pZM21-5 was ligated in place. *E. coli* GM2971 *mrr* was used for transformation to prevent restriction due to high levels of adenine methylation (17). This construction, designated pZM-T2dam (Fig. 2), specifies T2 Dam activity. It should be noted that the levels of m6A in C-DNA and hmC-DNA due to pZM-T2dam were much greater than the levels achieved with the cloned T4 *dam*+ gene (Table 1).

T2 *dam* gene sequence. The strategy employed for sequencing is shown in Fig. 3. The sequence (Fig. 4) shows an open reading frame of 780 nucleotides which would encode a polypeptide of approximately 30 kilodaltons. This is the same size as T4 Dam. The T2 *dam* sequence was compared with the T4 *dam* gene cloned and sequenced in our laboratory. Although there are 22 nucleotide differences in the reading frame, only four alter coding and two are in the same codon. Two of the amino acid alterations are at residues 20 and 26; namely, T2 (Pro-20) versus T4 (Ser-20) and T2 (Asp-26) versus T4 (Asn-26). Both of these residues are located in a region of homology shared by two other procaryotic DNA-adenine methyltransferases, *E. coli* Dam and the modification enzyme of *S. pneumoniae*, both of which recognize the sequence GATC (13). The third amino acid difference is at residue 188: T2 (Glu-188) versus T4 (Asp-188). This residue is located outside any region of homology.

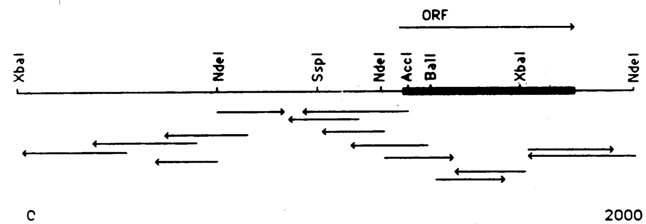

FIG. 3. Strategy used for generating the sequence of the T2 *dam* gene and flanking regions. Each arrow indicates a subclone made in M13mp18/19. The arrow points in the direction of the dideoxy sequencing. The open reading frame (ORF) and orientation of the gene relative to these restrictions sites are indicated by the arrow above the restriction map.

It should be noted that our T4 *dam*+ sequence has five nucleotide differences and three coding differences from the T4 *dam*+ gene analyzed by Macdonald and Mosig (22). Both plasmids produce an enzyme capable of methylating C-DNA and hmC-DNA (data not shown), and the three amino acid differences occur in the nonconserved regions (Z. Miner, S. Schlagman, and S. Hattman, manuscript in preparation). We presume that the sequence differences are due to phage strain differences. In any event, the T2 *dam*+ sequence is more closely related to the T4 *dam*+ clone from our laboratory.

Promoter fusions to *E. coli lacZ*. Beginning near the −35 site of the promoter, the 5′-flanking region sequence of T2 *dam* diverges from T4 *dam*. As a result, the two genes have identical −10 sites but different −35 sequences. To test the possibility that increased methylation levels by T2 Dam are due simply to increased gene expression from a more efficient promoter, fusions to *E. coli lacZ* were constructed. The T2 *dam* fusion contained 1,185 nucleotides of upstream flanking region plus the first 6 codons of Dam fused to codon 8 of *lacZ* in the vector pNM480 (Fig. 5). The T4 *dam* fusion was constructed in a similar manner, except that the upstream flanking region contained 753 nucleotides. The T4-*lacZ* fusion produced 1.9 times greater amounts of β-galactosidase than did the T2-*lacZ* fusion (Table 2). This implies that if there is a difference in promoter strength, it is the T4 promoter that is more efficient. Therefore, the higher level of methylation by T2 Dam appears to be due to some difference in the intrinsic methylation ability of the enzyme.

Physical mapping of T2 *dam*. The 3′-flanking region of T2 shows only three nucleotide differences from T4 in the intergenic region between *dam* and gene 56/69. The T2 sequence is identical to the 56/69 open reading frame for the first 17 amino acids. However, amino acids 18 through 20

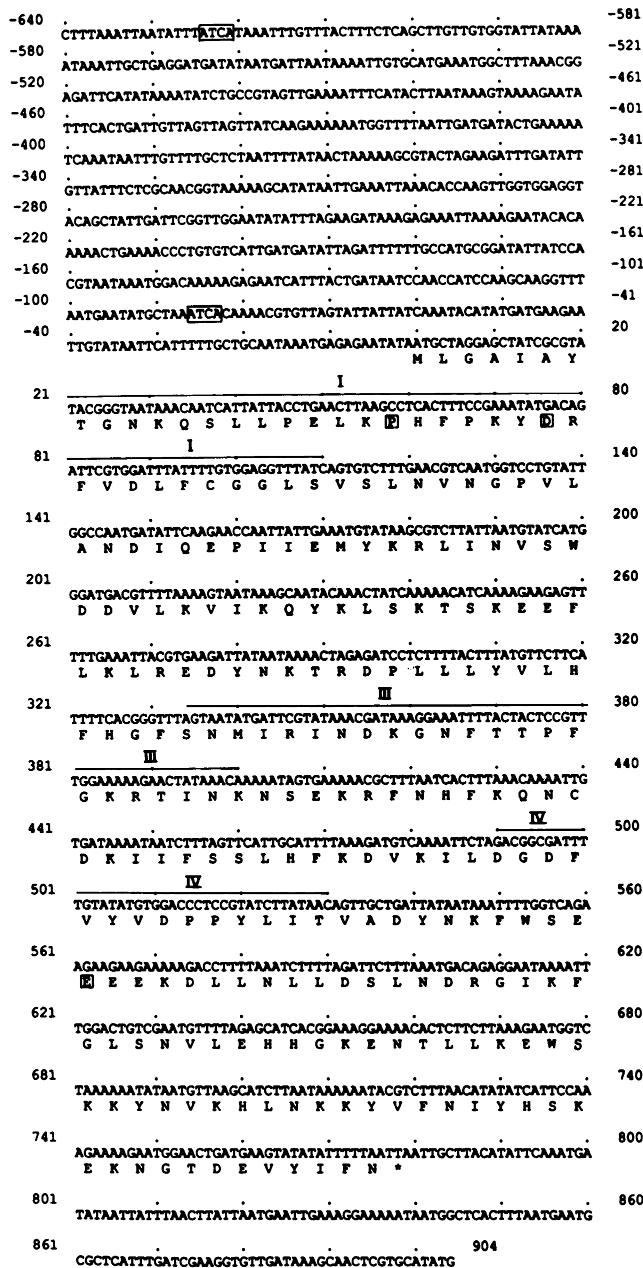

FIG. 4. The nucleotide and encoded amino acid sequences of the T2 *dam* gene and flanking regions. Some of the intergenic region between *dam* gene 58-61 is shown (positions -640 to −624). The 536-bp DNA insert is flanked by a 4-base ATCA direct repeat (enclosed in rectangular boxes). The amino acid differences between T2 Dam and T4 Dam are enclosed in squares. Regions of homology shared by T4 and T2 Dam, *E. coli* Dam, and the modification enzyme of *S. pneumoniae* are overlined and labeled with Roman numerals. The ATG of gene 56/69 is at position 841.

differ from the published sequence (22), and amino acid 21 is the same as that in T4. Amino acid 22 is likely to be different, since the first nucleotide in the codon (the last nucleotide in our clone) is different. This is consistent with T2 gene 56/69 and T4 gene 56/69 being slightly different but located at the same position in their respective genomes. However, it could not be ruled out that T2 *dam* and a portion of gene 56/69 had translocated to some other region.

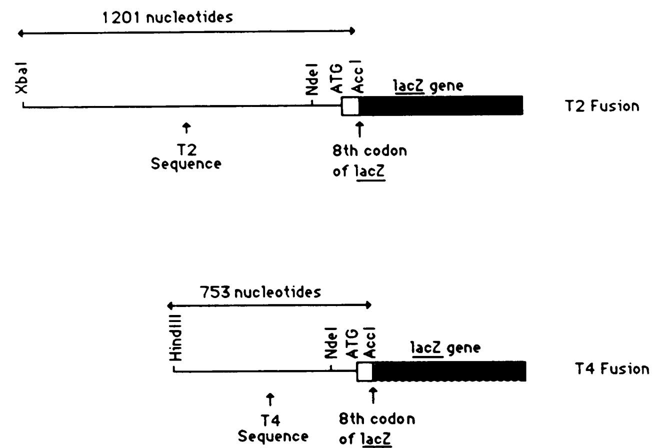

FIG. 5. T2 and T4 promoter fusions to *E. coli lacZ* were constructed by using the unique *AccI* site within the *dam* gene for cloning into the *AccI* site in the vector pNM480. This fuses the upstream flanking region, as well as the first 6 codons of the *dam* gene, in frame to the eighth codon of *lacZ*. The constructions were verified by sequencing and used in β-galactosidase assays (Table 2).

Therefore, in an attempt to physically map T2 *dam*, an *NdeI* fragment from pZM21-5 containing the 5′-flanking region of *dam* (Fig. 6) was used as a probe (probe 2) of T4 genomic digests. Surprisingly, this fragment did not hybridize to T4, although it did hybridize to the fragment of expected size in the T2 genomic digest (Fig. 6). However, an *NdeI-XbaI* fragment (probe 1) from a region farther upstream did hybridize to T4 (Fig. 6). The fragments from various restriction digests of T4 genomic DNA that hybridized to the T2 probe were analyzed by using the known restriction map of T4 (20). This analysis placed T2 *dam* near gene 56/69, the same location as T4 *dam*. Sequence data confirmed that homology between T2 and T4 resumed upstream of the *NdeI* fragment in probe 2; i.e., 536 bp upstream of the −35 region, the T2 and T4 segments are once again identical (Fig. 4), both in the intergenic region and the coding region of gene 58-61 that is present upstream of the *dam* gene.

Probe 2 did not hybridize to a Southern blot containing *E. coli* genomic digests, even under low-stringency conditions (data not shown). Moreover, no significant homology was found when *E. coli* and T4 sequences in GenBank were compared with the sequence in probe 2. The origin and function (if any) of this DNA insert in T2 have not been determined.

The sequence of the 536-bp DNA insert has been analyzed by computer for inverted and direct repeats. There is a 4-bp

TABLE 2. β-Galactosidase activity produced by T2 and T4 promoter fusions to *E. coli lacZ*[a]

Plasmid	β-Galactosidase sp act[b]	
	I	II
T2-*lacZ* fusion	704	535
T4-*lacZ* fusion	1,309	1,013
pNM480	8	ND

[a] *E. coli* MC1066 containing either the vector or one of the fusion plasmids (see Fig. 5) was grown to 2×10^8 cells per ml. Enzyme assays were as described by Miller (27).

[b] Two separate experiments are indicated by I and II. β-Galactosidase specific activities are expressed in Miller units. ND, Not done.

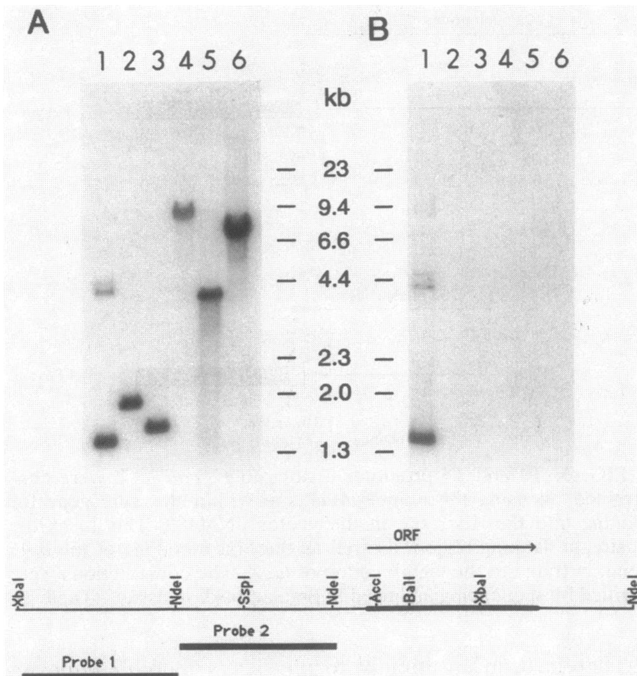

FIG. 6. Southern blot analysis of T2 αgt and T4c819 genomic digests hybridized with probes shown. Blot A was probed with the *Nde*I-*Xba*I fragment (probe 1). Lanes (digests): 1, T2 *Xba*I; 2, T4 *Xba*I; 3, T4 *Hin*dIII; 4, T4 *Eco*RV; 5, T4 *Eco*RI; 6, T4 *Cla*I. Blot B was probed with the *Nde*I fragment (probe 2). Lanes (digests): 1, T2 *Xba*I; 2, T4 *Xba*I; 3, T4 *Hin*dIII; 4, T4 *Eco*RV; 5, T4 *Eco*RI; 6, T4 *Cla*I. The multiple bands in lane 1 of blots A and B are attributed to incomplete digestion of the T2 αgt DNA.

direct repeat present at the insertion borders (ATCA), but it is not followed by an inverted repeat which would be indicative of an insertion element. Since T2 is quite AT rich, many other repeats, both direct and inverted, are present. Their significance is unknown. A search for open reading frames within the insert was made; the longest, beginning within the insert and terminating near the ATG of the *dam* gene, is 61 amino acids long. It begins with a methionine but does not have a Shine-Dalgarno sequence.

Genetic mapping of T2 *dam*. If T2 *dam* were physically located in the same region as T4 *dam*, as indicated above, then it should also be genetically linked to gene 56/69. To test this, a two-factor cross was performed with parental phages T2 αgt dam^+ 56^+ and T2 αgt dam^h *56*. The dam^h mutation allows *gt* phage to form plaques on *E. coli*(P1) lysogens. The *56* mutation (*am*61) restricts plaque formation only to hosts that contain an appropriate suppressor. The results of two separate crosses gave recombinant frequencies of 8.8 and 6.7% between the dam^h and *56* mutations (Table 2). It should be noted that, although cells were infected with equal amounts of parental phages, the yield of parental genotypes was skewed in favor of the 56^+ parent; this is also evident for the recombinant genotypes. Nevertheless, the data are consistent with T2 *dam* being in the same map location as T4 *dam*, and they are in contrast to results reported earlier (2).

DISCUSSION

Phages T2 and T4 each produce an enzyme, Dam, that methylates specific adenine residues in viral DNA as a postreplicative modification. However, T2 virion DNA has a higher m^{6}A content than T4 virion DNA (10). The two *dam* gene promoters are not identical; barring plasmid or post-transcriptional effects, T4 *dam* appears to have a slightly higher promoter strength (Table 2). This would rule out the trivial possibility that T2 DNA methylation is higher due to a relatively greater concentration of Dam in T2-infected cells. It is interesting to note that both enzymes appear to methylate GAThmC sites (their canonical sequence) to the full extent (Schlagman et al., in preparation). However, both enzymes also methylate noncanonical sequences, some subset of GA(T/C) (3, 5, 12; Schlagman et al., in preparation). In this regard, T2 Dam appears to methylate more of these secondary sites than does T4 Dam (5). For example, we show here that a cloned T2 Dam is able to methylate both C-DNA and hmC-DNA in vivo to a twofold higher level than T4 Dam (Table 1). We propose that T2 Dam recognizes a greater variety of sequences than T4 Dam and/or that it recognizes and transfers methyl groups to the noncanonical sites at higher rates than T4 Dam.

If this notion were correct, then we would expect that the two Dam proteins would not be identical. To test this, we determined the sequence of the T2 dam^+ gene and compared it with T4 dam^+. We found that the T2 *dam* open reading frame has three amino acid coding differences compared with T4 *dam*; in addition to these, there are 18 silent nucleotide substitutions (one of which produces the *Xba*I site present in T2 *dam*). Two of the amino acid differences occur at residues 20 and 26: T2 (Pro-20) versus T4 (Ser-20) and T2 (Asp-26) versus T4 (Asn-26). These residues are located in homology region I common to Dam proteins from *E. coli* and *S. pneumoniae* (13); these Dam proteins also recognize the sequence GATC. This implicates homology region I, and possibly these amino acids, in sequence recognition. The third amino acid difference is 15 residues downstream of homology region IV (13) and is a conservative change (T2 [Glu-188] versus T4 [Asp-188]). This does not rule out the possibility that this residue participates in sequence recognition; conservative amino acid changes have been implicated in alteration of gene activity (6, 15). It will be interesting to construct T2-T4 hybrid genes or mutant T4 genes which specify proteins containing one or more of these differences to determine which one(s) is responsible for the alteration in enzyme specificity.

The location of the T4 *dam* gene has been established by mapping and DNA sequencing; it is between gene 58-61 and gene 56/69 (22). In contrast, T2 *dam* was genetically mapped to a different region (2). However, from the sequence, 3′ to the T2 *dam* open reading frame we note the start of an open reading frame that appears to correspond to gene 56/69 (Fig. 4). Thus, the first 17 amino acids encoded by T2 are identical to those in T4 gene 56/69; although the next three T2 residues are different, residue 21 is again identical to that of T4. Since this represents the extent of our T2 clone, we do not know how the two phage genes compare in the remainder of their gene 56/69 sequences; but, since both T4 and T2 produce active dCTPase (56 gene product), these three amino acid changes appear to represent silent mutational alterations.

At the 5′ end of T2 *dam*, there is a 536-bp segment, starting in the −35 region of the promoter, that is missing in T4 (Fig. 4). This insert does not have all the characteristics indicative of an insertion sequence element. Although it has a 4-bp direct repeat at the borders of the insert, it lacks an inverted repeat present inside these borders, and it is smaller than known insertion sequence elements. Southern hybridization and computer analysis showed that the insert sequence is not found in *E. coli* DNA or anywhere else in T2

TABLE 3. Analysis of progeny from a two-factor cross between T2 αgt dam^+ 56^+ and T2 αgt dam^h 56 (am61)[a]

Expt	Genotype	No. of progeny	% Recombinants
1	dam^+ 56^+	149	
	dam^h 56	70	
	dam^+ 56	6	
	dam^h 56^+	15	8.8
2	dam^+ 56^+	133	
	dam^h 56	91	
	dam^+ 56	6	
	dam^h 56^+	10	6.7

[a] See Materials and Methods.

or T4 DNA. It is possible that it originated in a different enterobacterial strain that serves as a T2 host.

Farther upstream of the insert, the T2 and T4 sequences are once again identical. Our T2 sequence includes the 58-61–*dam* intercistronic divide and the 3′ end of the 58-61 open reading frame (only a portion is shown in Fig. 4). These results, taken together with the genetic linkage of T2 genes 56 and *dam* (Table 3), indicate that the T2 *dam* gene is in the same relative genomic location as T4 *dam*. We do not understand why the earlier T2 mapping results indicated otherwise (2).

In the first attempt to clone T2 *dam* on an *Xba*I fragment, we constructed a plasmid, pZM21-5, which contains a truncated gene fused in frame to a vector sequence that contributes additional amino acids. It is surprising that this chimeric protein has methylation activity. The N-terminal two-thirds of the polypeptide (163 residues) is encoded by T2 *dam*; it lacks one of the homology regions (IV) containing a Asp-Pro-Pro-Tyr motif, which is conserved in several Dam proteins (4, 21). However, although the carboxyl-terminal region encoded by pUC18 is smaller (83 amino acids) than the missing T2 Dam portion (96 amino acids), it contains a Pro-Pro-Phe motif and hydropathy profile similar to those of the missing T2 Dam portion (data not shown). The location of the Pro-Pro-Phe in the T2-pUC chimera is 25 amino acid residues farther away than the T2 Dam Asp-Pro-Pro-Tyr. This indicates a tolerance in spacing between regions of homology III and IV. Although the chimera is capable of methylating C-DNA, it is inactive on hmC-DNA (Table 2). These data suggest that two (or more) domains of the Dam are involved in sequence recognition and interaction. Recent work carried out in our laboratory has directly implicated homology region III in sequence recognition in both C-DNA and hmC-DNA (Miner et al., in preparation). Further experiments are in progress to delineate the methyltransferase domains responsible for protein-DNA and protein-*S*-adenosylmethionine interactions.

ACKNOWLEDGMENTS

We thank S. Schlagman for the many helpful discussions and for performing some of the m^6A content assays. Additional thanks are due to G. Mosig for sharing her unpublished data.

This work was supported by Public Health Service grant GM29227 from the National Institute for General Medical Sciences.

LITERATURE CITED

1. **Arraj, J. A., and M. G. Marinus.** 1983. Phenotypic reversal in *dam* mutants of *Escherichia coli*. Gene **25**:165–178.
2. **Brooks, J., and S. Hattman.** 1973. Location of the DNA-adenine methylase gene on the genetic map of phage T2. Virology **55**: 285–288.
3. **Brooks, J. E., and S. Hattman.** 1978. In vitro methylation of bacteriophage λ DNA by wild-type (dam^+) and mutant (dam^h) forms of phage T2 DNA adenine methylase. J. Mol. Biol. **126**: 381–394.
4. **Chandrasegaran, S., and H. O. Smith.** 1987. Amino acid sequence homologies among twenty-five restriction endonucleases and methylases, p. 149–156. *In* R. H. Sarma and M. H. Sarma (ed.), Structure and expression, volume 1: from proteins to ribosomes. Adenine Press, Schenectady, N.Y.
5. **Doolittle, M. M., and K. Sirotkin.** 1988. Bacteriophage T2 and T4, dam^+ and dam^h and *Eco* dam^+ methylation: preference at different sites. Biochim. Biophys. Acta **949**:240–246.
6. **Douglas, C. M., and R. J. Collier.** 1987. Exotoxin A of *Pseudomonas aeruginosa*: substitution of glutamic acid 553 with aspartic acid drastically reduces toxicity and enzymatic activity. J. Bacteriol. **169**:4967–4971.
7. **Gefter, M., R. Hausmann, M. Gold, and J. Hurwitz.** 1966. The enzymatic methylation of ribonucleic acid and deoxyribonucleic acid. X. Bacteriophage T3-induced S-adenosyl-methionine cleavage. J. Biol. Chem. **241**:1995–2006.
8. **Geier, G. E., and P. Modrich.** 1979. Recognition sequence of the *dam* methylase of *Escherichia coli* K12 and mode of cleavage of *Dpn*I endonuclease. J. Biol. Chem. **254**:1408–1413.
9. **Glickman, B. W., P. van den Elsen, and M. Radman.** 1978. Induced mutagenesis in dam^- mutants of *Escherichia coli*, a role of 6-methyladenine residues in mutation avoidance. Mol. Gen. Genet. **163**:307–312.
10. **Hattman, S.** 1970. DNA methylation of T-even bacteriophages and of their non-glucosylated mutants: its role in P1-directed restriction. Virology **42**:359–370.
11. **Hattman, S., J. E. Brooks, and M. Masurekar.** 1978. Sequence specificity of the P1 modification methylase (M · *Eco* P1) and the DNA methylase (M · *Eco dam*) controlled by the *E. coli dam* gene. J. Mol. Biol. **126**:367–380.
12. **Hattman, S., H. van Ormondt, and A. deWaard.** 1978. Sequence specificity of the wild-type (dam^+) and mutant (dam^h) forms of bacteriophage T2 DNA adenine methylase. J. Mol. Biol. **119**: 361–376.
13. **Hattman, S., J. Wilkinson, D. Swinton, S. Schlagman, P. M. Macdonald, and G. Mosig.** 1985. Common evolutionary origin of the phage T4 *dam* and host *Escherichia coli dam* DNA-adenine methyltransferase genes. J. Bacteriol. **164**:932–937.
14. **Hausmann, R., and M. Gold.** 1966. The enzymatic methylation of ribonucleic and deoxyribonucleic acid. IV. Deoxyribonucleic acid methylase in bacteriophage-infected *Escherichia coli*. J. Biol. Chem. **241**:1985–1994.
15. **Heaphy, S., M. Singh, and M. J. Gait.** 1987. Effect of single amino acid changes in the region of the adenylylation site of T4 RNA ligase. Biochemistry **26**:1688–1696.
16. **Hehlmann, R., and S. Hattman.** 1972. Mutants of bacteriophage T2 *gt* with altered DNA methylase activity. J. Mol. Biol. **67**: 351–360.
17. **Heitman, J., and P. Model.** 1987. Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. J. Bacteriol. **169**:3243–3250.
18. **Herman, G. E., and P. Modrich.** 1981. *Escherichia coli* K-12 clones that overproduce *dam* methylase are hypermutable. J. Bacteriol. **145**:644–646.
19. **Kim, J. S., and N. Davidson.** 1974. Electron microscope heteroduplex study of sequence relations of T2, T4, and T6 bacteriophage DNAs. Virology **57**:93–111.
20. **Kutter, E., and W. Rüger.** 1983. Structure, organization, and manipulation of the genome, p. 286. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), Bacteriophage T4. American Society for Microbiology. Washington, D.C.
21. **Lauster, R., A. Kriebardis, and W. Guschlbauer.** 1987. The GATATC-modification enzyme *Eco*RV is closely related to the GATC-recognizing methyltransferases *Dpn*II and *dam* from *E. coli* and phage T4. FEBS Lett. **220**:167–175.
22. **Macdonald, P. M., and G. Mosig.** 1984. Regulation of a new bacteriophage T4 gene, 69, that spans an origin of DNA replication. EMBO J. **3**:2863–2871.
23. **Mandel, M., and A. Higa.** 1970. Calcium dependent bacterio-

phage DNA infection. J. Mol. Biol. **53**:159–162.

24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
25. **Mannarelli, B. M., T. S. Balganesh, B. Greenberg, S. S. Springhorn, and S. A. Lacks.** 1985. Nucleotide sequence of the *Dpn*II DNA methylase gene of *Streptococcus pneumoniae* and its relationship to the *dam* gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82**:4468–4472.
26. **Marinus, M. G., and N. R. Morris.** 1973. Isolation of DNA methylase mutants from *Escherichia coli* K-12. J. Bacteriol. **114**:1143–1150.
27. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Minton, N. P.** 1984. Improved plasmid vectors for the isolation of translation *lac* gene fusions. Gene **31**:269–273.
29. **Revel, H. R.** 1967. Restriction of nonglucosylated T-even bacteriophage: properties of permissive mutants of *Escherichia coli* B and K12. Virology **32**:688–701.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
31. **Scherzer, E., B. Auer, and M. Schweiger.** 1987. Identification, purification, and characterization of *Escherichia coli* virus T1 DNA methyltransferase. J. Biol. Chem. **262**:15225–15231.
32. **Schlagman, S. L., and S. Hattman.** 1983. Molecular cloning of a functional *dam*$^+$ gene coding for phage T4 DNA adenine methylase. Gene **22**:139–156.
33. **van Ormondt, H., J. Gorter, K. J. Havelaar, and A. deWaard.** 1975. Specificity of a deoxyribonucleic acid transmethylase induced by bacteriophage T2. I. Nucleotide sequence isolated from *Micrococcus luteus* DNA methylated *in vitro*. Nucleic Acids Res. **2**:1391–1400.
34. **Yanisch-Perron, C., J. Viera, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp19 and pUC19 vectors. Gene **33**:103–119.
35. **Yee, J. K., and R. C. Marsh.** 1981. Alignment of a restriction map with the genetic map of bacteriophage T4. J. Virol. **38**:115–124.