JOURNAL OF BACTERIOLOGY, Mar. 1987, p. 939–943
0021-9193/87/030939-05$02.00/0

Vol. 169, No. 3

N^4-Methylcytosine as a Minor Base in Bacterial DNA

MELANIE EHRLICH,[1]* GEOFFREY G. WILSON,[2] KENNETH C. KUO,[3] AND CHARLES W. GEHRKE[3]

Department of Biochemistry, Tulane Medical School, New Orleans, Louisiana 70112[1]; New England BioLabs, Beverly, Massachusetts 01915[2]; and Department of Biochemistry, Experiment Station Chemical Laboratories, University of Missouri, Columbia, Missouri 65201[3]

Received 16 September 1986/Accepted 2 December 1986

The DNA base composition, including the minor base content, of 26 strains of bacteria was determined. The studied bacteria are sources of widely used restriction endonucleases. Approximately 35% of the bacterial DNAs contained N^4-methylcytosine, about 60% contained 5-methylcytosine, and about 90% had N^6-methyladenine.

Both N^6-methyladenine (m^6A) and 5-methylcytosine (m^5C) have been long known to be minor bases in bacterial DNA (7, 9, 16). One or both of these methylated bases are present in most bacterial DNAs examined (4, 11, 33). Recently, a third minor base, N^4-methylcytosine (m^4C), has been found in DNA from eight types of thermophilic bacteria (11) and in that from one type of mesophilic bacteria (19, 20). Previously, m^4C residues had been detected only in the RNA of the small ribosomal subunit of bacteria and of mammalian and insect mitochondria (8, 13, 40). This modified base is not present in a variety of eucaryotic DNAs (13; Gehrke et al., unpublished results).

The minor base composition of bacterial DNA is partially determined by restriction-modification systems (30). The level of minor bases should be consistent with the known specificities of the host restriction and modification enzymes. Also, it can reveal the presence of new DNA methyltransferases. For example, if the known restriction enzymes in a bacterium have recognition sites containing only G · C base pairs and if m^6A as well as m^5C is found in this cellular DNA, then one or more other unrelated modification pathways must be present. These could be involved in previously unidentified restriction-modification systems in the bacterium or in the control of various DNA functions (3, 15, 29, 35).

The present study demonstrates that genomic m^4C is not mostly limited to the DNA of thermophilic bacteria. Rather, it is present as a minor base in the DNA of many bacterial mesophiles. Of the 26 bacterial species examined in this study, 9 contained m^4C in their genomes. Most of the examined bacteria are mesophiles, and all of them are sources of commercially available restriction endonucleases. The prevalence of m^4C as a minor base in these bacterial DNAs indicates that many restriction endonucleases may be inhibited by N^4-methylation of cytosine residues in their recognition sites in vivo and that bacterial DNA cytosine methyltransferases must be carefully checked to determine whether they catalyze the formation of m^4C.

MATERIALS AND METHODS

Strains. All bacteria came from the New England BioLabs strain collection except *Xanthomonas campestris* pv. oryzae, which was from the collection of M. Ehrlich. The strains are listed in Table 1 with the sources from which they were originally obtained and the temperature at which the cultures were grown.

Purification of bacterial DNA. DNA was prepared from 10 g of wet packed cells suspended in 20 ml of 25% sucrose–50 mM Tris hydrochloride (pH 8.0), and then 10 ml of 0.25 M disodium EDTA (pH 8.0) and 6 ml of 10-mg/ml lysozyme in 0.25 M Tris (pH 8.0) were added. After 2 h at 0°C, 24 ml of 1% Triton X-100 in 50 mM Tris hydrochloride–67 mM disodium EDTA (pH 8.0) and 5 ml of 10% sodium dodecyl sulfate were added. The solutions were shaken gently for several minutes to achieve complete mixing and cell lysis. They then were extracted with phenol and chloroform, dialyzed, and treated with RNase I, and the DNA was precipitated by standard techniques (26).

Digestion of DNA and chromatography of deoxynucleosides. DNA was digested to nucleosides with nuclease P1 and *Escherichia coli* alkaline phosphatase (14). The resulting deoxynucleosides from 20 to 90 μg of DNA were separated by high-performance liquid chromatography. A previously described elution system (11, 14) was improved to increase the resolution, sensitivity, and speed of the chromatography. We used a specially developed reversed-phase column (Supelcosil LC-18S; 250 by 4.6 mm; Supelco); a two-buffer, single-ramp elution gradient; a 1.0-ml/min flow rate; and a 26°C isothermal column temperature. Buffer A was 2.5% methanol in 0.05 M potassium phosphate (pH 4.5), and buffer B was 20% methanol in 0.05 M potassium phosphate (pH 4.0). To assure consistency, all of the columns were pretested by using a mixture of nucleoside standards. The gradient consisted of 100% buffer A isocratic from 0 to 5 min, a linear ramp from 100% buffer A to 100% buffer B from 5 to 20 min, and 100% buffer B isocratic from 20 to 30 min. The column was washed with 70% methanol in water for 5 min and equilibrated with buffer A for 15 min before the next run. The ramp rate, pH, and column temperature must be rigorously maintained to ensure separation of m^5dCyd and m^4dCyd. A Hewlett-Packard model 1090A liquid chromatography instrument equipped with DR-5 solvent delivery system, automatic injector, automatic sampler, diode-array detector, HP-80B controller, HP-7470A plotter, HP-9133 hard disk drive, HP Think Jet printer, and data process unit multichannel integrator was used. Absorption of light was measured at 254 ± 2 nm and 280 ± 2 nm. Given a sample size of >30 μg, a methylated deoxynucleoside representing 0.01 mol% of the total deoxynucleosides could be detected.

* Corresponding author.

TABLE 1. Sources of bacteria

Bacterial strain	Original source	Growth temp (°C)
Arthrobacter luteus	ATCC 21606	32
Bacillus amyloliquefaciens H	F. Young	37
Bacillus aneurinolyticus	IAM 1077	30
Bacillus caldolyticus	A. Atkinson	60
Bacillus globigii	C. Duncan and G. Wilson	35
Fusobacterium nucleatum D	M. Smith	37
Haemophilus aegyptius	ATCC 11116	37
Haemophilus gallinarum	ATCC 14385	37
Haemophilus haemolyticus	ATCC 10014	37
Haemophilus influenzae Rd	H. Smith	37
Haemophilus influenzae Rf	C. Hutchinson III	37
Haemophilus parainfluenzae	J. Setlow	37
Klebsiella pneumoniae OK8	J. Davies	37
Moraxella bovis	ATCC 10900	35
Moraxella nonliquefaciens	ATCC 17953	35
Nocardia aerocolonigenes	ATCC 23870	30
Proteus vulgaris	ATCC 13315	35
Rhodopseudomonas sphaeroides	S. Kaplan	30
Serratia marcescens	S. Brenner	37
Streptococcus pneumoniae 641	S. Lacks	37
Streptomyces achromogenes	ATCC 12767	30
Thermus aquaticus YT1	J. Harris	65
Xanthomonas campestris pv. badrii	ATCC 11672	30
Xanthomonas campestris pv. holcicola	ATCC 13461	30
Xanthomonas campestris pv. malvacearum	ATCC 9924	30
Xanthomonas campestris pv. oryzae 507	T.-T. Kuo	30

RESULTS AND DISCUSSION

In a previous study, m^4C was found in the genomes of approximately one-half of the 17 types of examined thermophilic bacteria (11). In that study, 15 types of mesophilic bacteria were also examined; most were anaerobes, like the thermophiles. None of these contained m^4C in their DNA. In the present study, we looked at a greater variety of bacteria, almost all of which are mesophiles (Table 1). Because of their similar chemical nature, specially designed chromatography systems (5, 11) have to be used to resolve m^4dCyd and m^5dCyd (or the corresponding bases), which are difficult to separate. With a high-performance liquid chromatography system that resolves m^4dCyd and m^5dCyd, we have quantitated these deoxynucleosides in DNA digests. Besides using the different retention times of these deoxynucleosides to identify the m^4dCyd and m^5dCyd peaks, we also rely on their different UV-light absorbance ratios. In the eluting buffer, A_{280} relative to A_{254} is 3.0 for m^5dCyd and 1.5 for m^4dCyd.

The base composition of the different bacterial DNAs is given in Tables 2 and 3. There were no irregularities in their major base content, and these data are in agreement with previously reported values available for some of these species (33). The similarity of the major base compositions of the DNA of the five examined *Haemophilus* species (Tables 2 and 3) reflects their genetic relatedness. In contrast, in duplicate determinations, the major base content of the DNA from *Bacillus caldolyticus* differed considerably from that of the other *Bacillus* species (Tables 2 and 3). This result suggests that *B. caldolyticus* is genetically quite divergent from the rest of the examined species of its genus.

All of the bacteria studied contain at least one modified base (m^6A, m^5C, or m^4C) as a minor component. Of the 26 types of bacteria examined in this study, 9 had m^4C in their DNA (Tables 2 and 3). m^5C was present in the genomes of 16 of the DNAs studied (Tables 2 and 3). This is in good agreement with the previous findings of m^5C in the DNA of about 60% of 44 examined strains of mesophilic bacteria (10, 11, 21, 32, 33).

As was observed in previous studies (11, 33), most (88%) of the types of DNA studied had m^6A as a minor base. Often it was present at rather high levels (>0.3 mol%) (Tables 2 and 3). These bacteria may have a *dam*-type (5′-Gm^6ATC-3′) methylation system directing mismatch repair (15, 23, 29). The *dam* methylase of *E. coli* is unaccompanied by a restriction enzyme with corresponding specificity for unmethylated 5′-GATC-3′ sites. Consistent with their moderately high m^6A content (Tables 2 and 3), *Haemophilus gallinarum*, *Haemophilus parainfluenzae*, *Proteus vulgaris*, *Klebsiella pneumoniae*, and *Serratia marcescens* have *E. coli*-like *dam* methylation as determined by restriction and DNA hybridization analysis (3).

The 11 bacterial strains containing <0.2 mol% m^6A in their DNA (Tables 2 and 3) may lack the *dam*-type modification pathway, which has been implicated in the regulation of transcription, of transposition, and of initiation of DNA replication as well as in directing mismatch repair in *E. coli* (15, 29, 35). Indeed, DNAs from *Bacillus globigii* and *X. campestris* pathovars holcicola, malvacearum, and oryzae, which have <0.2 mol% m^6A (Table 2), were previously shown not to possess *dam*-type methylation (3). The m^6A content of *Moraxella bovis* DNA can be accounted for by its methylation at the A residues of 5′-GATC-3′ sites as part of a restriction-modification system (4). *B. caldolyticus* has only one known restriction endonuclease, *Bcl*I, which recognizes a 6-base-pair sequence (30). The relatively high level of m^6A in its DNA and the previously reported absence in this bacterium of sequences that hybridize with those of the *E. coli dam* gene (3) suggest that *B. caldolyticus* harbors either another restriction-modification system probably recognizing a 4-base-pair sequence or a DNA (adenine-N^6)methyltransferase other than a *dam* methylase or restriction-associated enzyme.

More than one modified base was present in most of the bacteria examined (Tables 2 and 3). In some cases, as for *H. parainfluenzae*, this may reflect the presence in one bacterium of *dam* methylation (3) as well as several restriction-modification systems involving different modified bases (25, 41). In contrast, *Haemophilus influenzae* Rd, whose only modified base in its DNA is m^6A (Table 3), has multiple DNA methyltransferases, all of which methylate adenine residues (31). Also, *B. caldolyticus*, *B. globigii*, *Streptococcus pneumoniae*, *H. influenzae* Rf, *Streptomyces achromogenes*, and *X. campestris* pv. oryzae contained only one detectable modified base in their DNA (Tables 2 and 3). From their major base compositions and the sequences of their restriction recognition sites, the genomic frequency of modified bases resulting from methylation at a given class of restriction sites can be estimated by assuming a random sequence of bases in the genome and one modified base on each strand per recognition site. For example, in the cases of *Streptomyces achromogenes* (*Sac*I, GAGCTC; *Sac*II, CCGCGG [30]) and *X. campestris* pv. oryzae (*Xor*I,

TABLE 2. Minor and major base composition of DNAs from bacteria containing restriction endonucleases that had been tested for sensitivity to cytosine methylation

Bacterial strain	Restriction enzyme(s)[a]	Inhibition by cytosine 5-methylation[b]	mol%			
			m^4C	m^5C	m^6A	A+T
A. luteus	*Alu*I	+	<0.01	0.38	0.03	25
B. amyloliquefaciens	*Bam*HI	+	<0.01	0.29	0.01	53
B. caldolyticus	*Bcl*I	−	<0.02	<0.02	0.44	47
B. globigii	*Bgl*I, *Bgl*II	+	0.05	<0.01	<0.01	56
H. aegyptius	*Hae*II, *Hae*III	+	<0.01	0.07	0.55	60
H. haemolyticus	*Hha*I	+	<0.02	0.07	0.61	59
H. influenzae Rf	*Hin*fI	±	<0.01	<0.01	0.60	61
H. parainfluenzae	*Hpa*II	+	<0.01	0.13	0.37	60
M. bovis	*Mbo*I	±	0.09	<0.01	0.40	55
P. vulgaris	*Pvu*I, *Pvu*II	+	0.02	<0.01	0.32	62
S. marcescens	*Sma*I	+	<0.01	0.18	0.89	40
T. aquaticus	*Taq*I	−	0.11	<0.01	0.38	32
X. campestris pv. holcicola	*Xho*I, *Xho*II	+	0.14	<0.01	0.10	36
X. campestris pv. malvacearum	*Xma*I, *Xma*III	+	0.09	0.19	0.18	36
X. campestris pv. oryzae	*Xor*II	+	<0.01	0.09	<0.01	35

[a] Only the commonly used restriction endonucleases found in the given bacterial strain (Table 1) are listed. See reference 30 for a listing of all known restriction enzymes and their recognition sequences in these bacteria.

[b] The bacteria contain restriction endonucleases at least one of which was tesed for inhibition by site-specific cytosine 5-methylation. *Alu*I, *Bam*HI, *Bgl*II, *Hae*III, *Hha*I, *Hpa*II, *Pvu*I, *Xho*II, *Xma*I, and *Xor*II were inhibited by methylation of their DNA substrates catalyzed by at least one bacterial DNA (cytosine-5)methyltransferase, and *Bcl*I was not inhibited by such methylation (4, 17, 25, 27, 34). In the cases of *Alu*I, *Bam*HI, *Hae*II, *Hha*I, and *Hpa*II, this inhibition was catalyzed by a DNA (cytosine-5)methyltransferase isolated from the host bacterium and, hence, implicated in that restriction-modification system. *Hpa*II, *Hha*I, *Bgl*I, *Hae*III, *Sma*I, and *Xho*I were inhibited by the vertebrate-type DNA methylation, namely, cytosine 5-methylation at CpG sites within or overlapping their recognition sites (2, 36, 37, 42). *Mbo*I and *Hin*fI are able to cleave DNA in which virtually all of the cytosine residues are methylated, although the rate of catalysis is unusually slow (18); *Hin*fI was also partially inhibited by methylation of a DNA substrate at its CpG sites by a human DNA methyltranferase (R. Y.-H. Wang and M. Ehrlich, unpublished results). Catalysis by *Taq*I is unusual in being unaffected by methylation of all the cytosine residues of a DNA substrate (18). Although almost all restriction endonucleases are inhibited by complete substitution of cytosine residues in their DNA by m^5C residues (18), that is not considered in this table because the extensive nature of such substitution might have indirect effects on enzyme-DNA interactions.

CTGCAG; *Xor*II, CGATCG [30]), the predicted frequency of methylated base resulting from any one restriction-modification system (~0.03 mol% each for *Sac*I, *Xor*I, and *Xor*II and ~0.1 mol% for *Sac*II) is much more than the detection limits for m^4C and m^6A (<0.01 mol% m^4C or m^6A; Tables 2 and 3). Therefore, *Streptomyces achromogenes* with 0.23 mol% m^5C in its DNA and *X. campestris* pv. oryzae with 0.09 mol% m^5C probably possess only DNA (cytosine-5)methyltransferases.

Evidence for unexpectedly low frequencies of restriction recognition sites was observed for *Haemophilus haemolyticus* and *Haemophilus aegyptius*. They both have DNA (cytosine-5)methyltransferases that recognize a 4-base-pair site containing only C · G base pairs (30). From the major base compositions of *H. haemolyticus* and *H. aegyptius* DNAs (Table 2), the methylation of the *Hha*I or *Hae*III recognition sequences would be expected to yield ~0.15 mol% m^5C in the genome if a random distribution of bases is assumed. The actual level of m^5C in these genomes was 0.07 mol%, implying at least a twofold underrepresentation of the recognition sites. A similar underrepresentation of about fourfold in 5′-GATC-3′ sites was previously observed for *Thermobacteroides acetoethylicus* (11). An even greater difference between the predicted and the observed minor base compositions was seen for *Fusobacterium nucleatum* D. This bacterium harbors three restriction endonucleases, all of which recognize 4-base-pair sequences containing only C · G base pairs, namely, 5′-GGCC-3′, 5′-CGCG-3′, and 5′-GCGC-3′ (24). *F. nucleatum* has only 28 mol% G+C in its genome (Table 3). If the above tetranucleotide sequences occurred at the frequency expected for a random distribution of bases in this DNA, then there should have been ~0.12 mol% methylated cytosine to confer resistance at the recognition sequences to the host restriction endonucleases. However, in duplicate determinations, no m^4C was detected in this DNA, and only 0.023 mol% m^5C was found (Table 3). As expected, DNA from this organism was resistant to digestion by *Fnu*DII and the *Fnu*DI and *Fnu*DIII isochizomers *Hae*III and *Hha*I. This underrepresentation of the modified base involves thousands of sites per bacterial genome. It might be due to selective pressure against too many potential restriction sites or too many modification sites which could be subject to interactions with sequence-specific DNA-binding proteins (35, 39).

TABLE 3. Minor and major base composition of some bacterial DNAs[a]

Bacterial strain	Restriction enzyme(s)[b]	mol%			
		m^4C	m^5C	m^6A	A+T
B. aneurinolyticus	*Ban*I, *Ban*II	0.06	0.09	0.35	56
F. nucleatum	*Fnu*DII	<0.01	0.02	0.22	72
H. gallinarum	*Hga*I	<0.01	0.14	0.46	58
H. influenzae Rd	*Hin*dII, *Hin*dIII	<0.02	<0.02	0.38	61
K. pneumoniae	*Kpn*I	<0.02	0.13	0.63	43
M. nonliquefaciens	*Mnl*I	<0.01	0.18	0.53	57
N. aerocolonigenes	*Nae*I	<0.01	0.18	0.04	32
R. sphaeroides	*Rsa*I	0.16	<0.01	0.12	34
S. pneumoniae	*Dpn*I	<0.01	<0.01	0.14	58
S. achromogenes	*Sac*I, *Sac*II	<0.01	0.23	<0.01	36
X. campestris pv. badrii	*Xba*I	0.04	0.05	0.01	42

[a] These bacteria, unlike those listed in Table 2, have restriction endonucleases that have not been tested for inhibition by site-specific cytosine methylation.

[b] Only the commonly used restriction endonucleases found in the given bacterial strain (Table 1) are listed. See reference 30 for a listing of all known restriction enzymes and their recognition sequences in these bacteria.

Alternatively, it might result from selective pressures generating nonrandom dinucleotide and trinucleotide frequencies such as selective codon usage (1, 28). On the other hand, methylated bases in some bacteria, such as *Bacillus amyloliquefaciens* (Table 2) (17), are present at much higher concentrations than expected based on their identified restriction endonuclease recognition sites.

The strain of *Streptococcus pneumoniae* used for this study has a rare, although not unique (I. Schildkraut, unpublished results), methylation-dependent restriction system. The single known restriction enzyme of this strain, *Dpn*I, cleaves DNA only if the genome is methylated at 5′-GATC-3′ sites to yield bifilarly modified 5′-Gm^{6}ATC-3′ sequences (22). This appears to be a strategy to restrict foreign *dam*-methylated DNA, especially from bacteriophages. The DNA of this strain is unmethylated at its 5′-GATC-3′ sites and is thereby protected from such restriction (22). Because this strain contains 0.14 mol% m^{6}A in its DNA (Table 3), we predict that it will be found to possess another restriction system of the more conventional type that requires adenine methylation for inhibiting the restriction enzyme from hydrolyzing the host DNA. Alternatively, it might have a DNA adenine methyltransferase with functions other than protection against restriction (15, 23, 29, 35).

Methylated cytosine (m^{5}C or m^{4}C) residues were found in the DNA of *Serratia marcescens*, *Nocardia aerocolonigenes*, *Bacillus aneurinolyticus*, and *X. campestris* pv. malvacearum (Tables 2 and 3), which possess at least one restriction endonuclease (30) able to cleave recognition sites containing only C · G base pairs. The first two genomes contain m^{5}C and no detectable m^{4}C. Because the last two species contain genomic m^{4}C as well as m^{5}C, either DNA (cytosine-N^4)methyltransferases or 5-methyltransferases might methylate the corresponding *Xma*I (CCCGGG), *Xma*III (CGGCCG), *Ban*I (GGYRCC), or *Ban*II (GRGCYC) recognition sites in these bacteria.

In this study, we have found that m^{4}C is frequently present in mesophilic bacteria as a minor genomic base, just as previously found in thermophilic bacteria (11) and in the mesophile *Bacillus centrosporus* (19). m^{5}C was not found in the DNA of any of 15 examined thermophiles which grow optimally at ≥60°C (11). Similarly, the two extreme thermophiles examined in this study, *B. caldolyticus* and *Thermus aquaticus*, contained no detectable genomic m^{5}C (Table 2). Also, *Thermoplasma acidophilum* 122-1B2, which is grown optimally at 59°C (and contained 54 mol% A+T in its genome), had m^{4}C (0.21 mol%), m^{6}A (2.05 mol%), and no m^{5}C (<0.01 mol%) in its DNA (D. Swinton, S. Hattman, D. Searcy, and C. Gehrke, unpublished results). These results are consistent with the hypothesis that bacteria which grow at high temperatures avoid m^{5}C in their genomes because of the propensity of this base to heat-induced deamination and because of the inefficient excision of T from the resulting T · G mismatched base pairs (12, 38; S. Shenoy, K. Erlich, and M. Erlich, submitted for publication).

DNA (cytosine-N^4)methyltransferases, like the analogous 5-methyltransferases, can participate in restriction-modification systems. Janulaitis and co-workers (6, 19, 20) have demonstrated that site-specific DNA methylases from *B. centrosporus*, *Micrococcus varians*, and *Citrobacter freundii* catalyze the formation of m^{4}C residues within sequences recognized by the restriction endonucleases of these cells. However, some cytosine methyltransferases might control DNA repair, expression, replication, or transposition like the adenine-specific *dam* methylase in *E. coli* (15, 23, 29, 35).

ACKNOWLEDGMENTS

We thank Ira Schildkraut of New England BioLabs for suggesting this study and for critical reading of the manuscript. We are also very grateful to Dat Phan for his excellent technical assistance in the chromatography.

This research was supported in part by National Science Foundation grant PCM-8219000.

LITERATURE CITED

1. **Adams, J., and E. D. Rothman.** 1982. Estimation of phylogenetic relationships from DNA restriction patterns and selection of endonuclease cleavage sites. Proc. Natl. Acad. Sci. USA **79**:3560–3564.
2. **Bird, A. P., and E. M. Southern.** 1978. Use of restriction enzymes to study eukaryotic DNA methylation. J. Mol. Biol. **118**:27–47.
3. **Brooks, J. E., R. M. Blumenthal, and T. R. Gingeras.** 1983. The isolation and characterization of the *Escherichia coli* DNA adenine methylase (*dam*) gene. Nucleic Acids Res. **11**:837–851.
4. **Brooks, J. E., and R. J. Roberts.** 1982. Modification profiles of bacterial genomes. Nucleic Acids Res. **10**:913–924.
5. **Butkus, V. V., S. J. Klimasauskas, and A. A. Janulaitis.** 1985. Analysis of products of DNA modification by methylases: a procedure for the determination of 5- and N^4-methylcytosines in DNA. Anal. Biochem. **148**:194–198.
6. **Butkus, V. V., S. J. Klimasauskas, D. Kersulyte, D. Vaitkevicius, A. Lebionka, and A. Janulaitis.** 1985. Investigation of restriction-modification enzymes from *M. varians* RFL19 with a new type of specificity toward modification of substrate. Nucleic Acids Res. **13**:5727–5746.
7. **Doskocil, J., and Z. Sormova.** 1965. The occurrence of 5-methylcytosine in bacterial deoxyribonucleic acids. Biochim. Biophys. Acta **95**:513–515.
8. **Dubin, D. T., and C. C. HsuChen.** 1983. The 3′-terminal region of mosquito mitochondrial small ribosomal subunit RNA: sequence and localization of methylated residues. Plasmid **9**:307–320.
9. **Dunn, D. B., and J. D. Smith.** 1958. The occurrence of 6-methylaminopurine in deoxyribonucleic acids. Biochem. J. **68**:627–636.
10. **Dybvig, K., D. Swinton, J. Maniloff, and S. Hattman.** 1982. Cytosine methylation of the sequence GATC in a mycoplasma. J. Bacteriol. **151**:1420–1424.
11. **Ehrlich, M., M. Gama-Sosa, L. Carreira, L. Ljungdahl, K. Kuo, and C. Gehrke.** 1985. DNA methylation in thermophilic bacteria: N^4-methylcytosine, 5-methylcytosine, and N^6-methyladenine. Nucleic Acids Res. **13**:1399–1412.
12. **Ehrlich, M., K. F. Norris, R. Y.-H. Wang, K. C. Kuo, and C. W. Gehrke.** 1986. DNA cytosine methylation and heat-induced deamination. Biosci. Rep. **6**:387–393.
13. **Fellner, P.** 1969. Nucleotide sequences from specific areas of the 16S and 23S ribosomal RNAs of *E. coli*. Eur. J. Biochem. **11**:12–27.
14. **Gehrke, C. W., R. A. McCune, M. A. Gama-Sosa, M. Ehrlich, and K. C. Kuo.** 1984. Quantitative reversed-phase high-performance liquid chromatography of major and modified nucleosides in DNA. J. Chromatogr. **301**:199–219.
15. **Glickman, B. W., and M. Radman.** 1980. *E. coli* mutator mutants deficient in methylation-instructed DNA mismatch correction. Proc. Natl. Acad. Sci. USA **77**:1063–1067.
16. **Hattman, S.** 1981. DNA methylation, p. 517–548. *In* P. D. Boyer (ed.), The Enzymes, vol. 14. Academic Press, Inc., New York.
17. **Hattman, S., T. Keister, and A. Gottehrer.** 1978. Sequence specificity of DNA methylases from *Bacillus amyloliquefaciens* and *Bacillus brevis*. J. Mol. Biol. **124**:701–711.
18. **Huang, L.-H., C. M. Farnet, K. C. Ehrlich, and M. Ehrlich.** 1982. Digestion of highly modified bacteriophage DNA by restriction endonucleases. Nucleic Acids Res. **10**:1579–1591.
19. **Janulaitis, A., S. Klimasauskas, M. Petrusyte, and V. Butkus.** 1983. Cytosine modification in DNA by *Bcn*I methylase yields N^4-methylcytosine. FEBS Lett. **161**:131–134.
20. **Janulaitis, A. A., P. S. Stakenas, M. P. Pyatrushite, Y. B. Bitinaite, S. I. Klimashauskas, and V. V. Butkus.** 1984. Speci-

ficity of new restrictases and methylases. Unusual modification of cytosine in position 4. Mol. Biol. (Engl. Transl. Mol. Biol. [Moscow]) **18**:93–105.
21. **Korch, C., P. Hagblom, and S. Normark.** 1983. Sequence-specific DNA modification in *Neisseria gonorrhoeae*. J. Bacteriol. **155**:1324–1332.
22. **Lacks, S., and B. Greenberg.** 1977. Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. J. Mol. Biol. **114**:153–168.
23. **Lu, A.-L., S. Clark, and P. Modrich.** 1983. Methyl-directed repair of DNA base-pair mismatches *in vitro*. Proc. Natl. Acad. Sci. USA **80**:4639–4643.
24. **Lui, A. C. P., B. C. McBride, G. F. Vovis, and M. Smith.** 1979. Site specific endonuclease from *Fusobacterium nucleatum*. Nucleic Acids Res. **6**:1–15.
25. **Mann, M. B., and H. O. Smith.** 1977. Specificity of Hpa II DNA methylation. Nucleic Acids Res. **4**:4211–4221.
26. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. **3**:208–218.
27. **McClelland, M.** 1983. The effect of site specific methylation on restriction endonuclease cleavage (update). Nucleic Acids Res. **11**:r169–r173.
28. **Nussinov, R.** 1984. Doublet frequencies in evolutionary distinct groups. Nucleic Acids Res. **12**:1749–1763.
29. **Pukkila, P. J., J. Peterson, G. Herman, P. Modrich, and M. Meselson.** 1983. Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *E. coli*. Genetics **104**:571–582.
30. **Roberts, R. J.** 1985. Restriction and modification enzymes and their recognition sequences. Nucleic Acids Res. **13**:r165–r200.
31. **Roy, P. H., and H. O. Smith.** 1973. DNA methylases of *Haemophilus influenzae* Rd. II. Partial recognition site base sequences. J. Mol. Biol. **81**:445–459.
32. **Schein, A, B. J. Berdahl, M. Low, and E. Borek.** 1972. DNA methylases of *Haemophilus influenzae* Rd. Biochim. Biophys. Acta **272**:481–485.
33. **Shapiro, H. S.** 1976. Distribution of purines and pyrimidines in deoxyribonucleic acids, p. 242–283. *In* G. D. Fasman (ed.), Handbook of biochemistry and molecular biology, nucleic acids section, vol. 2. CRC Press, Inc., Cleveland.
34. **Smith, H. O., and S. V. Kelly.** 1984. The methylases of type II restriction and modification systems, p. 39–71. *In* H. Cedar, A. P. Riggs, and A. Razin (ed.), DNA methylation. Springer-Verlag, New York.
35. **Sternberg, N.** 1985. Evidence that adenine methylation influences DNA-protein interactions in *Escherichia coli*. J. Bacteriol. **164**:490–493.
36. **Ulrich, A., T. J. Dull, A. Gray, J. A. Phillips, and S. Peter.** 1982. Variation in the sequence and modification state of the human insulin gene flanking regions. Nucleic Acids Res. **10**:2225–2240.
37. **van der Ploeg, L. H. T., and R. A. Flavell.** 1980. DNA methylation in the human γδβ-globin locus in erythroid and nonerythroid tissues. Cell **19**:947–958.
38. **Wang, R. Y.-H., K. C. Kuo, C. W. Gehrke, L.-H. Huang, and M. Ehrlich.** 1982. Heat and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. Biochem. Biophys. Acta **697**:371–377.
39. **Wang, R. Y.-H., X.-Y. Zhang, and M. Ehrlich.** 1986. A human DNA-binding protein is methylation-specific and sequence-specific. Nucleic Acids Res. **14**:1599–1614.
40. **Woese, C. R., G. E. Fox, L. Zablen, T. Uchida, L. Bonen, K. Pechman, B. J. Lewis, and D. Stahl.** 1975. Conservation of primary structure in 16S ribosomal RNA. Nature (London) **254**:83–86.
41. **Yoo, O. J., P. Dwyer-Hallquist, and K. L. Agarwal.** 1982. Purification and properties of the Hpa I methylase. Nucleic Acids Res. **20**:6511–6519.
42. **Youssofian, H., and C. Mulder.** 1981. Detection of methylated DNA sequences in eukaryotic DNA with the restriction endonucleases *Sma*I and *Xma*I. J. Mol. Biol. **150**:133–136.