

## Site-Specific Methylation of Adenine in the Nuclear Genome of a Eucaryote, *Tetrahymena thermophila*

GAIL S. HARRISON,<sup>1†</sup> R. CRAIG FINDLY,<sup>2</sup> AND KATHLEEN M. KARRER<sup>1\*</sup>

Department of Biology, Brandeis University, Waltham, Massachusetts 02254,<sup>1</sup> and Department of Genetics, University of Georgia, Athens, Georgia 30609<sup>2</sup>

Received 9 December 1985/Accepted 3 April 1986

**DNA in the polyploid macronucleus of the ciliated protozoan *Tetrahymena thermophila* contains the modified base N<sup>6</sup>-methyladenine. We identified two GATC sites which are methylated in most or all of the 45 copies of the macronuclear genome. One site is 2 kilobases 5' to the histone H4-I gene, and the other is 5 kilobases 3' to the 73-kilodalton heat shock protein gene. These sites are de novo methylated between 10 and 16 h after initiation of conjugation, during macronuclear anlage development. The methylation states of these two GATC sites and four other unmethylated GATC sites do not change in the DNA of cells cultured under conditions which change the activity of the genes, including logarithmic growth, starvation, and heat shock.**

The nuclear DNAs of most eucaryotes contain the modified base 5-methylcytosine (5MeC) as a minor component. In many cases, site-specific methylation of cytosine, particularly near the 5' ends of genes, has been correlated with gene inactivity (for reviews see references 1, 17, and 45). Therefore, it has been suggested that 5MeC may be involved in transcriptional regulation.

N<sup>6</sup>-methyladenine (N<sup>6</sup>MeA) is found in the nuclear DNAs of several unicellular eucaryotes, either in combination with other modifications or as the only modified base. The DNAs of the green algae *Chlamydomonas reinhardi* and *Chlorella* spp. have both methylcytosine and methyladenine (28, 46), while only methylcytosine has been found in the DNA of another green alga, *Euglena gracilis* (40). In the ciliates *Tetrahymena* (23), *Paramecium* (16), *Oxytricha* (39), and *Stylonychia* (5) N<sup>6</sup>MeA is the sole modified base. The function of N<sup>6</sup>MeA in eucaryotic DNA is unknown.

We examined methylation at specific sites in the genome of *Tetrahymena thermophila*. *Tetrahymena* cells contain two nuclei, a diploid micronucleus and a polyploid macronucleus. The macronucleus is responsible for most, if not all, transcriptional activity during vegetative growth. During sexual reproduction (conjugation), the macronucleus is destroyed, and mitotic products of the zygotic micronucleus develop into a new micronucleus and a new macronucleus. The developing macronucleus, called the anlage, undergoes many structural and morphological changes, including DNA replication to a final DNA content of 45c (49), elimination and rearrangement of germ line DNA sequences (50), and de novo methylation (9, 26).

Although the micronuclear DNA of *Tetrahymena* cells does not contain modified bases, 0.8% of the adenine residues in the transcriptionally active macronucleus are modified to N<sup>6</sup>MeA (23, 28). A partially purified methylase activity was isolated from macronuclei (10). A nearest-neighbor analysis of in vivo methylated DNA showed that all four bases are found 5' to the methylated adenine, but only thymidine is found 3'. Some of the methyladenine residues in *Tetrahymena* macronuclear DNA occur in the sequence

5'-G<sup>Me</sup>ATC-3' and can be assayed by digestion with restriction enzyme *DpnI*, which cleaves at the sequence GATC only if the adenine is methylated on both strands (31).

Since the micronucleus, which is largely transcriptionally inactive, does not contain N<sup>6</sup>MeA, one possible hypothesis is that N<sup>6</sup>MeA is required for transcriptional activation. Methylation would then be expected to precede or correlate with the onset of transcription in the developing macronuclear anlagen. De novo methylation of bulk anlage DNA at GATC sites is first detectable between 13.5 and 15 h after initiation of conjugation, approximately midway through anlage development (9, 26).

Estimates for the time of transcriptional activation in the macronuclear anlagen vary somewhat depending on the type of analysis and the particular gene studied. Wenkert and Allis (48) performed light microscope autoradiography of cells conjugated in the presence of [<sup>3</sup>H]uridine. These authors reported the presence of silver grains over the macronuclear anlagen at 8 h after initiation of conjugation, 5 h before de novo methylation can be detected. In another study, the rate of incorporation of [<sup>3</sup>H]uridine into rRNA was found to be low until after 13 h of conjugation, when it rose sharply (33). The results of electron microscope autoradiography experiments by Weiske-Benner and Eckert (47) were in agreement with the results of earlier studies, showing incorporation of [<sup>3</sup>H]uridine over macronuclear anlagen at 8 to 12 h after initiation of conjugation and incorporation over the nucleoli beginning at about 12 h.

The onset of transcriptional activity in developing macronuclei has also been monitored by using genetic markers. Bruns and Brussard (13) described methods for the construction of functional heterokaryons which carry mutations for drug resistance in the micronucleus, but express drug sensitivity in the macronucleus. Genetic activity of the developing macronucleus in the progeny of these cells can be assayed in terms of acquisition of drug resistance. This occurs at 12 h after initiation of conjugation for a mutation conferring resistance to 6-methylpurine (33) and at 15.5 h for two mutations which confer cycloheximide resistance (13). Thus, de novo methylation of the DNA in the macronuclear anlagen is correlated in a general way with transcriptional activation.

\* Corresponding author.

† Present address: Eleanor Roosevelt Cancer Institute, Denver, CO 80262.

If methylation has a role in regulation of transcription in *Tetrahymena*, then the state of methylation might change with the transcriptional activity of nearby genes. Blackburn et al. (9) investigated the distribution of methylated GATC sequences in the extrachromosomal ribosomal DNA molecules of *Tetrahymena*. These authors found that 90% of the molecules did not contain any GATC sites methylated on both strands. The remaining 10% of the molecules were heterogeneous with respect to methylation. Methylation at GATC sites in the ribosomal DNA molecules did not change in DNAs from cells in different physiological states. Similarly, although the mRNA content in starved cells decreases fourfold relative to growing cells (7), the overall levels of DNA methylation do not change detectably (23).

A more sensitive way to test for involvement of N<sup>6</sup>MeA in transcriptional regulation would be to examine the methylation state of specific sites in the genome. With this in mind we looked for sites of methylation near genes whose transcriptional activity could be manipulated by changing the cell culture conditions. The histone genes and the heat shock protein genes are good candidates for this kind of analysis because the amounts of their messages vary dramatically with the physiological state of the cells.

In this study we demonstrated the existence of uniformly methylated GATC sites near the genes coding for a histone H4 gene and a heat shock-induced protein, hsp73, and determined the timing of de novo methylation of these sites in the DNA of the developing macronuclear anlagen. We found that methylation of the sites does not change in DNAs from cells in different physiological states in which the transcriptional activity of the genes changes.

## MATERIALS AND METHODS

**Cell culture.** *T. thermophila* BVII, CU399 (mating type VI), and CU401 (mating type VII) were kindly provided by P. Bruns. Stocks were maintained in 1% PPYS, and cultures were grown in 2% PPYS as previously described (30).

To induce conjugation, vegetative cells were grown to a density of approximately  $3 \times 10^5$  cells per ml. The cells were pelleted by centrifugation at  $100 \times g$  for 1.5 min and then washed with 10 mM Tris (pH 7.4) and starved overnight at a concentration of  $2 \times 10^5$  cells per ml. Equal numbers of cells of the two mating types were mixed and placed at 29°C without shaking to initiate conjugation. These methods have been described by other workers previously (12, 35).

To heat shock cells, a logarithmic culture of strain CU399 was placed in a water bath at 55°C. The temperature of the culture was monitored until it reached 38°C, at which time the culture was transferred to a water bath at 40°C. The cells were heat shocked at 40°C for 40 min. This procedure induces maximal production of heat shock proteins.

Stationary cells were obtained by growing a vegetative culture of strain CU399 to a density of approximately  $1.6 \times 10^6$  cells per ml.

**Nuclear isolation and DNA purification.** Micro- and macronuclei were isolated from exponentially growing vegetative cells by using the method of Gorovsky et al. (24), with the modifications suggested by Howard and Blackburn (29).

Macronuclear anlagen were separated from micro- and macronuclei by centrifugation to equilibrium in Percoll (Pharmacia) gradients by using a modification of the methods of Allis and Dennison (4) and Allen et al. (3). Conjugating cells were washed and lysed by homogenization in a Waring blender as described by Howard and Blackburn (29). The

total nuclei were pelleted by centrifugation at  $16,300 \times g$  for 10 min, suspended in 40 ml of 0.05 M sucrose–4% gum arabic–0.05% spermidine trihydrochloride–1 mM MgCl<sub>2</sub> (pH 6.75), and pelleted by centrifugation at  $2,000 \times g$  for 10 min. Nuclei from  $1.2 \times 10^8$  cells were suspended in 4 ml of the buffer described above by vigorous agitation with a Pasteur pipette, and Percoll was added to a concentration of 30%. Two Percoll gradients were formed in siliconized test tubes by centrifuging a solution of 57% Percoll ( $n = 1.3455$ ) in the gradient buffer of Allis and Dennison (4) at  $16 \times 10^3$  rpm in type SS34 rotor for 20 min. One-half of the nuclei were layered on top of each gradient and centrifuged to equilibrium at  $7.5 \times 10^3$  rpm in a type HB4 rotor for 20 min. The top band, containing macronuclear anlagen, was collected, and the nuclei were diluted with 10 volumes of the buffer described above and pelleted by centrifugation at  $5,860 \times g$  for 10 min. The anlagen were suspended in the nuclear isolation medium of Howard and Blackburn (29). The nuclei were stained with methyl green, and the proportion of anlagen was estimated by counting nuclei under both phase-contrast and bright-field optics as described by Allis and Dennison (4). The purity of macronuclear anlage DNA was estimated in the preparations from cells at different times after initiation of conjugation as follows: 10 h, 72%; 14 h, 86%; 16 h, 92%. The majority of the contamination was macronuclear DNA; the amount of micronuclear DNA was less than 1%. DNA was purified from nuclei as described by Gall (21).

Whole-cell DNA was isolated from starved, heat-shocked, and stationary cells by suspending a cell pellet in 10 mM Tris–1 mM EDTA–1% Sarcosyl–2% Nonidet P-40. The lysate was digested with pancreatic RNase (Worthington Diagnostics), T<sub>1</sub> RNase (Boehringer Mannheim Biochemicals), and pronase (Sigma Chemical Co.), as previously described (30). DNA was phenol extracted and ethanol precipitated twice.

**Probes.** Plasmids p508.8, containing the H4-I gene, was generously provided by G. Bannon and M. Gorovsky. Plasmid pA.3.2.1 is a 1-kilobase (kb) *Hind*III fragment of *Tetrahymena* DNA containing the 5' end of the 73-kilodalton heat shock gene cloned into the *Hind*III site of pBR322. Plasmid DNA was prepared by alkaline hydrolysis and polyethylene glycol precipitation (8; L. Holland, personal communication). The insertion was separated from vector sequences by digestion with the appropriate restriction enzymes, agarose gel electrophoresis, and isolation on a type NA-45 DEAE membrane according to the recommendations of the manufacturer (Schleicher & Schuell, Inc.).

**Southern analysis.** Restriction digestion of plasmid and genomic DNAs was carried out as recommended by the supplier of the enzymes. All of the enzymes except *Dpn*I were obtained from Boehringer Mannheim Biochemicals; *Dpn*I was obtained from International Biotechnologies, Inc.

Between 1 and 12 U of enzyme was added per  $\mu$ g of DNA; digestion was at 37°C for 3 to 6 h. To stop digestion, EDTA was added to a concentration of 6 mM along with sucrose and bromophenol blue (final concentrations, 7.5 and 0.025%, respectively). DNA samples were loaded onto 0.65 to 0.8% agarose gels. The gels were run for 400 to 600 V-h in 40 mM Tris–20 mM sodium acetate–2 mM EDTA–18 mM sodium chloride (pH 8.5).

DNA was transferred to nitrocellulose filters (Schleicher & Schuell) by the method of Southern (44), but the reservoir consisted of a sheet of Whatman 3MM paper saturated with  $20 \times$  SSC ( $1 \times$  SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and transfer was for 2.5 h.

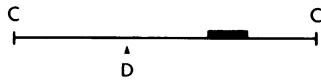


FIG. 1. General scheme for identification of  $C^{Me}ATC$  sites. Genomic DNA was cut with restriction enzyme *Clal* to generate large restriction fragments. The fragments were detected by short probes, indicated by the bar. A methylated GATC site at any position on the *Clal* fragment was detectable by subsequent digestion of the DNA with *DpnI*, which decreased the size of the fragment detected by the probe.

Probes were nick translated as described by Maniatis et al. (34), except that the nick-translated product was separated from unincorporated counts by precipitating the DNA twice with 95% ethanol and washing with 70% ethanol.

The nitrocellulose filters were prehybridized, hybridized, and washed as previously described (30); there was an additional wash in  $0.2\times$  SSC at  $65^{\circ}C$ .

## RESULTS

**Uniformly methylated GATC sites detected by probes for the histone H4-I gene and the 73-kilodalton heat shock protein gene.** We employed a strategy that enabled us to search for methylated GATC sites in long stretches of genomic DNA surrounding a short region for which a probe is available. As illustrated in Fig. 1, macronuclear DNA was digested with restriction enzyme *Clal*, which cuts *Tetrahymena* DNA infrequently and generates fragments of relatively high molecular weight. A short probe can be used to detect methylation at any position on the *Clal* fragment by comparing DNA digested with *Clal* with DNA digested with *Clal* plus *DpnI*. If the *Clal* fragment contains a methylated GATC site, then double digestion with *Clal* plus *DpnI* should generate a shorter fragment than digestion with *Clal* alone.

Using this method, we located uniformly methylated GATC sites approximately 2 kb 5' to the histone H4-I gene and 5 kb 3' to the gene coding for heat shock-induced protein hsp73. The probes used for these experiments are shown in Fig. 2. p508.8 is a 4.3-kb insertion in the pBR322 derivative RVIIΔ7 (Fig. 2A); it is bordered by *HindIII* and *EcoRI* sites and contains the gene coding for one of the two histone H4 genes, H4-I (6). We mapped the *Clal* sites in genomic DNA where indicated, outside the cloned region.

Fig. 3 shows hybridization of probe p508.8 to macronuclear DNA digested with *Clal* (lane 1), *Clal* plus *DpnI* (lane 2), *HindIII* plus *EcoRI* (lane 3), *HindIII* plus *EcoRI* plus *DpnI* (lane 4), and *EcoRI* plus *DpnI* (lane 5). In lanes 1

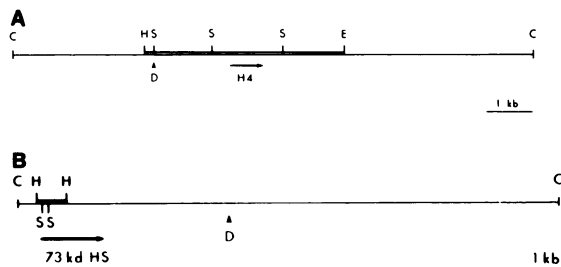


FIG. 2. Restriction maps of macronuclear DNA in the vicinity of the H4-I gene (a) and hsp73 gene (b) of *Tetrahymena*. The arrows indicate transcribed regions of the DNA. The bar indicates the probe. Restriction enzyme sites are indicated as follows: C, *Clal*; D, *DpnI*; E, *EcoRI*; H, *HindIII*; S, *Sau3A*. kd, Kilodalton.

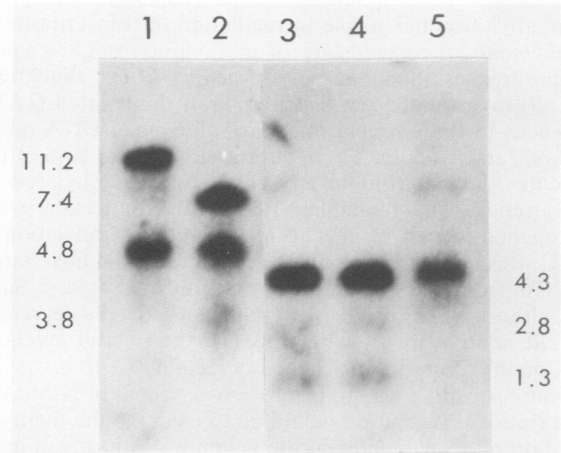


FIG. 3. Mapping of the *DpnI* site 5' to the H4-I gene. Macronuclear DNA from strain BVII was digested with various restriction enzymes and probed with pGB508.8. The following enzymes were used: lane 1, *Clal*; lane 2, *Clal* plus *DpnI*; lane 3, *HindIII* plus *EcoRI*; lane 4, *HindIII* plus *EcoRI* plus *DpnI*; lane 5, *EcoRI* plus *DpnI*. The sizes (in kilobases) for the fragments in lanes 1 and 2 are indicated by the numbers on the left, and the sizes for the fragments in lanes 3 through 5 are indicated on the right. The 4.8-kb band in lanes 1 and 2 is the result of cross-hybridization of the probe to the H4-II gene.

and 2, the higher-molecular-weight band corresponds to the H4-I gene. The lower-molecular-weight band, at 4.8 kb, is the result of the cross-hybridization of the H4-I probe to the H4-II gene (6).

Plasmid p508.8 detects an 11.2-kb fragment in macronuclear DNA digested with *Clal*. Hybridization to a fragment of this size was barely detectable in DNA digested with *Clal* plus *DpnI* (lane 2). Instead, we found a major band at 7.4 kb and very faint hybridization to a band at 3.8 kb. (The 3.8-kb band can be seen more clearly in Fig. 4). This strongly suggests that every 11.2-kb *Clal* fragment with homology to the histone H4-I gene contains at least one *DpnI* site.

The detection of the two new bands in DNA digested with *Clal* and *DpnI* suggested that the *DpnI* site lies within the cloned fragment of DNA. Bannon et al. (6) mapped three

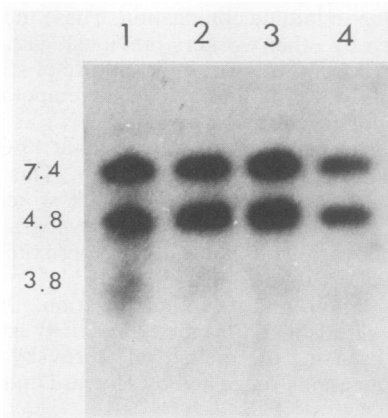


FIG. 4. Constancy of methylation in DNAs from cells in different physiological states. Macronuclear DNAs from vegetative cells (lane 1), starved cells (lane 2), stationary-phase cells (lane 3), and heat-shocked cells (lane 4) were digested with *Clal* plus *DpnI* and hybridized with pGB508.8. The numbers on the left indicate the sizes of the fragments (in kilobases).

*Sau3A* sites within the fragment of *Tetrahymena* DNA in p508.8 (Fig. 2A). *Sau3A* is an isoschizomer of *DpnI*; it cleaves at the GATC sequence regardless of whether the adenine is methylated. All *DpnI* sites are also *Sau3A* sites, but *Sau3A* sites are *DpnI* sites only if the adenines in both strands are methylated (32). The data shown in Fig. 3 suggest that the leftmost *Sau3A* site mapped by Bannon et al. (6) on p508.8 corresponds to the uniformly methylated *DpnI* site which we detected.

The *Tetrahymena* H4-I gene was originally cloned as an 8.6-kb *EcoRI* fragment, which corresponds to an 8.6-kb band seen on Southern blots of genomic DNA cut with *EcoRI* and probed with a yeast H4 probe (6). Digestion of the 8.6-kb *EcoRI* fragment with *HindIII* generated the 4.3-kb *HindIII-EcoRI* fragment of p508.8. Genomic DNA digested with *HindIII* plus *EcoRI* (Fig. 3, lane 3) and probed with p508.8 had a major band at 4.3 kb. DNA digested with *DpnI* and *EcoRI* (Fig. 3, lane 5) contained a major band of approximately the same size, as would be expected if the left *Sau3A* site in Fig. 2A were methylated. The same band was present in DNA digested with *HindIII* plus *DpnI* plus *EcoRI* (Fig. 3, lane 4). In similar gels which were run for longer times, it was possible to show that the *DpnI-EcoRI* fragment was slightly smaller than the *HindIII-EcoRI* fragment (data not shown). Hybridization of the small region between the left *HindIII* site and the *DpnI* site (Fig. 2A) would account for the faint 3.8-kb band in DNA digested with *ClaI* plus *DpnI*.

The 4.8-kb *ClaI* fragment in lanes 1 and 2 of Fig. 3 represents cross-hybridization of the probe to the H4-II gene, which apparently does not have any sites recognized by *DpnI*. We suspect that the two minor bands in lanes 3 and 4 also represent cross-hybridization of the H4-II gene with the H4-I probe, although this has not been directly demonstrated. The *EcoRI* fragment containing the H4-II gene is about 4.0 kb long (6). Digestion of the *EcoRI* fragment with *HindIII* may generate the two fragments at 1.3 and 2.8 kb in lanes 3 and 4. Although no *HindIII* site is present in the H4-II gene of strain CU399 (S. Horowitz and M. A. Gorovskiy, personal communication), it is possible that one exists in the DNA of strain BVII, which was used in this experiment. In all other respects, the restriction maps derived from our data confirm those previously published by Bannon et al. (6). The 4.0-kb *EcoRI* fragment containing the H4-II gene may comigrate with the H4-I *DpnI-EcoRI* fragment in lane 5.

Taken together, the data in Fig. 3 suggest that there is a GATC site approximately 2 kb 5' to the H4-I gene which is methylated in most or all copies of the macronuclear genome. Additionally, the *Sau3A* sites located approximately 900 base pairs upstream and downstream from the H4-I gene are not methylated on both strands.

A second region of the genome investigated for the presence of methylated GATC sites was the region in the vicinity of the 73-kilodalton heat shock protein gene. Probe pA.3.2.1 is a 1.0-kb *HindIII* fragment of *Tetrahymena* DNA containing coding sequences for heat shock-induced protein hsp73. A partial restriction map of the DNA in this region of the genome is shown in Fig. 2B. pA.3.2.1 hybridized to a 17-kb *ClaI* fragment in macronuclear DNA (Fig. 5, lane 1). The 17.0-kb *ClaI* fragment was digested with *DpnI* to generate an 8.0-kb *ClaI-DpnI* fragment (Fig. 5, lane 2). Therefore, a uniformly methylated GATC site is located 8 kb from the left *ClaI* site; this places it about 5 kb 3' to the 73-kilodalton heat shock protein gene.

We searched for the 5' *DpnI* site nearest the hsp73 gene by hybridizing pA.3.2.1 to macronuclear DNA digested with

*DpnI* alone. This was expected to reveal a fragment bounded on the 3' side of the gene by the *DpnI* site shown in Fig. 2B and by the most proximal *DpnI* site on the 5' side of the gene. pA.3.2.1 hybridized to a 22-kb *DpnI* band in macronuclear DNA (data not shown). Therefore, the nearest 5' *DpnI* site is 22 kb from the first *DpnI* site which we mapped, or approximately 14 kb 5' to the gene.

**Methylation of GATC sites near the H4-I and hsp73 genes does not change in DNAs from cells in different physiological states.** The amounts of transcripts from the *Tetrahymena* histone genes and heat shock-induced genes are known to vary in different physiological states. In the case of the hsp73 gene, it has been demonstrated that mRNA abundance in heat-shocked cells is dramatically increased compared with the low level of message present in vegetatively growing cells (25). In *Drosophila* tissue culture cells, similar changes in the abundance of heat shock mRNAs have been shown to be due to changes in transcriptional activity (19). It is reasonable to suppose that transcriptional activation is a major factor for regulation of heat shock genes in *Tetrahymena*, as it is in *Drosophila*.

Figure 5 shows the results of hybridization of pA.3.2.1 to *ClaI*-digested DNAs (lanes 1, 3, and 5) and to *ClaI-DpnI*-digested DNAs (lanes 2, 4, and 6) from vegetative cells (lanes 1 and 2), starved cells (lanes 3 and 4), and heat-shocked cells (lanes 5 and 6). The fragments generated by digestion with *ClaI* or *ClaI* plus *DpnI* were identical in all physiological states. The pattern was also the same for DNA from stationary cells (data not shown). Thus, methylation of the GATC site 5 kb 3' to the hsp73 gene did not change with mRNA abundance, nor did we observe any methylation of the two *Sau3A* sites near the 5' end of the gene.

We also examined methylation of the site 5' to the H4-I gene in order to determine whether methylation changed with gene activity. *Tetrahymena* cells are viable over periods of several days of starvation in 10 mM Tris, a treatment which is routinely used to induce mating (12). There is little or no DNA synthesis in starved cells (26), and the level of histone mRNA is reduced 30- to 40-fold, compared with a 4-fold reduction in poly(A)<sup>+</sup> mRNA levels (7). The decrease

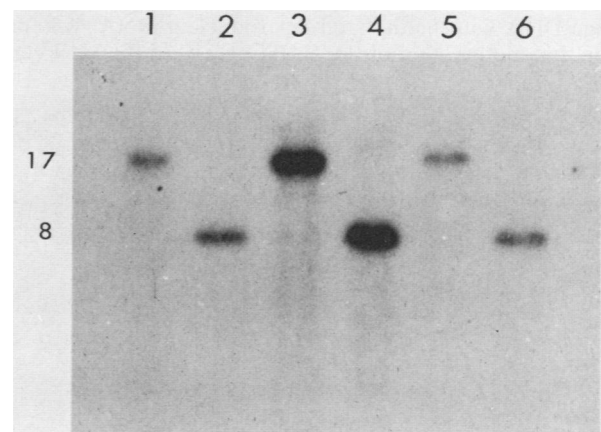


FIG. 5. Mapping of the *DpnI* site 3' to the hsp73 gene and constancy of methylation in DNAs from cells in various physiological states. Macronuclear DNAs were digested with *ClaI* (lanes 1, 3, and 5) or *ClaI* plus *DpnI* (lanes 2, 4, and 6) and probed with the 1.0-kb *HindIII* fragment of pA.3.2.1. The DNAs were from cells in vegetative growth (lanes 1 and 2), starved cells (lanes 3 and 4), and heat-shocked cells (lanes 5 and 6). Lanes 3 and 4 contained more DNA than the other lanes.

in mRNA levels is due in part to a fivefold decrease in transcriptional activity (D. S. Pederson, K. Shupe, G. A. Bannon, and M. A. Gorovsky, submitted for publication). Figure 4 shows the results of hybridization of p508.8 to *ClaI-DpnI*-digested DNAs from macronuclei of vegetative cells (lane 1), starved cells (lane 2), stationary cells (lane 3), and heat-shocked cells (lane 4). For all physiological states, the bands generated by digestion with *ClaI* plus *DpnI* were identical. Thus, as was the case with the gene for *hsp73*, methylation of the GATC site 2 kb 5' to the H4-I gene did not change with changes in transcriptional activity of the H4-I gene. Additionally, no *Sau3A* sites other than the one located nearest the *HindIII* site were methylated in any physiological state.

**De novo methylation of the GATC site 2 kb 5' to the histone H4-I gene during conjugation.** We have shown previously that newly synthesized DNA in the anlagen becomes detectably methylated at GATC sites between 13.5 and 15 h after initiation of conjugation (26). This was determined by measuring when newly synthesized DNA that was labeled with [<sup>3</sup>H]thymidine was first detectably digested by *DpnI*. The methylase activity appeared to act rapidly because within 90 min (between 13.5 and 15 h after initiation of conjugation) newly synthesized DNA was methylated at GATC sites to a level similar to that seen in vegetative cells.

Analysis of de novo methylation of the GATC site near the H4-I gene confirmed the timing of methylation previously indicated in the studies with bulk DNA. Anlage DNA was isolated from conjugating cells at 10, 14, and 16 h after initiation of conjugation. We estimated that 75 to 85% of the DNA was from the anlagen, with the major contaminant being DNA from the old macronucleus of paired and unpaired cells. Figure 6 shows p508.8 hybridized to *ClaI-DpnI*-digested anlage DNA from conjugants at 10 h (lane 1), 14 h (lane 2), and 16 h (lane 3). The band at 11.2 kb corresponds to the unmethylated *ClaI* fragment, demonstrating that a substantial portion of the H4-I gene was unmethylated at 10 h after initiation of conjugation. Densitometric analysis of the film indicated that 24% of the hybridization to the H4-I gene was to the 7.4-kb band representing the methylated form of the gene. Since this could be fully accounted for by the estimated 28% contamination of the anlage DNA with methylated macronuclear DNA, we conclude that methylation of the GATC site 5' to the H4-I gene

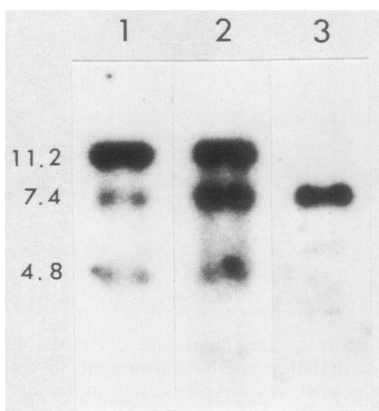


FIG. 6. Methylation occurs during macronuclear anlage development. Macronuclear anlage DNAs obtained from cells at 10 h (lane 1), 14 h (lane 2), and 16 h (lane 3) after initiation of conjugation were digested with *ClaI* plus *DpnI* and hybridized with pGB508.8.

had not commenced in the developing macronuclear anlagen by 10 h after initiation of conjugation.

By 14 h after initiation of conjugation, approximately 52% of the anlage DNA molecules were methylated at this site, as judged from the relative intensity of the 11.2- and 7.4-kb bands corresponding to the unmethylated and methylated *ClaI* fragments, respectively. By 16 h no 11.2-kb fragment was detectable, suggesting that methylation was complete.

Although the results of preliminary experiments suggested that the region of the genome near the *hsp73* gene also becomes methylated between 10 and 16 h after initiation of conjugation, the analysis is complicated by the presence of a genome rearrangement in that same region (data not shown).

## DISCUSSION

The micronuclear DNA of *Tetrahymena* cells is unmethylated. During development of the macronucleus there is de novo methylation of 0.8% of the adenine residues to N<sup>6</sup>MeA. Studies on total nuclear DNA have shown that the methylation occurs between 13.5 and 15 h after initiation of conjugation (9, 26). Thus, methylation is correlated with both DNA rearrangement and transcriptional activation of the macronuclear genome.

In order to further investigate the role of methylation, we identified several GATC sites which are methylated at adenine residues in most or all of the copies of macronuclear DNA in *Tetrahymena* cells. We found that two of these sites are methylated between 10 and 16 h after initiation of conjugation, which is in good agreement with the results of studies on bulk DNA.

We examined the methylation state of specific GATC sites in the DNA of cells where the transcriptional activity of nearby genes is known to vary. We detected uniformly methylated GATC sites 2 kb 5' to the histone H4-I gene and 5 kb 3' to the gene coding for heat shock-induced protein *hsp73*. Methylation at these sites remained constant in DNAs from cells in various physiological states. While these sites are relatively far from the 5' end of the genes which are most likely to be the regions involved in transcriptional regulation, we never detected any methylation of the two GATC sites which are very near the 5' end of the *hsp73* gene or of the two sites 900 base pairs upstream and downstream from the H4-I gene. Thus, it appears that methylation of GATC sites is not involved in transcriptional control in *Tetrahymena* cells, at least for the genes coding for H4-I and *hsp73*. This is consistent with the findings of other workers that methylation levels do not change in different physiological states, both for total DNA (23) and for ribosomal DNA (9).

These findings do not rule out a role for methylation in the control of transcriptional activity for several reasons. It is possible that methylation at sites other than GATC is involved in transcriptional regulation. Digestion with *DpnI* only allows us to assay for a small percentage of the methylation events in *Tetrahymena* DNA. Bromberg et al. (10) have shown that all four bases occur 5' to the <sup>Me</sup>AT in *in vivo* methylated DNA. Thus, there are methylated sites in the genome other than those which occur at the sequence GATC. Judging from the very large size of *DpnI* fragments (10 to 20 kb) (26), we estimate that fully methylated GATC sites represent fewer than 5% of the methylation events in macronuclear DNA.

It is likely that sites with sequences other than GATC are uniformly methylated in *Tetrahymena*. Martindale et al. (36) identified an *EcoRI* site which is cut in micronuclear DNA

but not in macronuclear DNA. Since *EcoRI* is sensitive to methylation (18) and analysis with other restriction enzymes gave no evidence for genome rearrangements in that region, the simplest explanation of this result is that the *EcoRI* site is uniformly methylated in the macronuclear DNA. Thus, it would be premature to conclude that methylation does not vary with the physiology of the cells until such time as methods become available to assay a larger percentage of the methylated sites.

On the other hand, the sites which we have been able to assay may be representative of the methylated sites with respect to their constancy in the DNAs of cells in different physiological states. The methylation state of specific sites may be determined during macronuclear development and remain constant thereafter. In this view, development of the macronucleus would be analogous to the differentiation of various tissues in metazoa. During tissue differentiation hypomethylation of specific DNA sequences seems to be a developmental event that is correlated with transcriptional potential (as opposed to transcriptional activity) in a specific cell type (reviewed in references 27 and 42).

One possible effect of methylation of macronuclear DNA might be a remodeling of the chromatin structure in the macronucleus. Pratt and Hattman (38) reported that methyladenine residues in *Tetrahymena* chromatin, but not in naked DNA, are preferentially digested by both staphylococcal nuclease and DNase I. These authors concluded that most of the methyladenine is in linker DNA. The converse seems to be the case with 5MeC. Cedar and colleagues (41, 43) found that 5MeC is preferentially localized in the DNA which is relatively protected upon digestion of native chromatin with micrococcal nuclease. These workers concluded that the bulk of the 5MeC is in nucleosomal core particles.

In the *Tetrahymena* genome, which has an adenine-thymine content of 75% (2), 0.8% methyladenine would represent one methylated adenine residue per 165 base pairs. This is slightly more than would be required for methylation of the entire genome once per nucleosome, since the internucleosomal distance has been estimated to be about 200 base pairs by both nuclease digestion and cross-linking experiments (15, 20, 22). It would be of interest to know whether methylated adenines are evenly distributed in the *Tetrahymena* genome.

Uniformly methylated sites may occur in the genomes of other ciliates. Cartinhour and Herrick (14) have reported the existence of a *ClaI* site in a cloned fragment of *Oxytricha* DNA which is refractory to digestion by *ClaI* in the genomic DNA. One possible explanation of this observation is that the macronuclear *ClaI* site is methylated, since *ClaI* does not digest DNA when the internal adenine in the recognition site is methylated (37).

Whether N<sup>6</sup>MeA plays a role in the transcriptional activation of the macronucleus or not, other possible functions should be considered. In procaryotes, N<sup>6</sup>MeA functions in mismatch repair (11) and in restriction-modification systems (for a review, see reference 27). We have argued elsewhere (26) that since de novo methylation does not occur in the macronuclear anlagen until there have been several rounds of DNA synthesis, it cannot be involved in mismatch repair, at least in these early stages. Another possibility, as first suggested by Blackburn et al. (9), is that methylation functions in DNA rearrangement, perhaps to distinguish sequences which are retained in the macronucleus from the micronucleus-specific sequences which are degraded. We have looked for methylation sites in the vicinity of genome

rearrangements and have identified a cluster of methylated GATC sites which occur near a site of genomic rearrangement (data not shown). Experiments are in progress to determine the relative timing of methylation and DNA sequence rearrangement.

#### ACKNOWLEDGMENTS

We are indebted to G. Bannon for generously providing the clone for the H4-I gene and to L. Wangh and J. Haber for stimulating discussions.

This work was supported by Public Health Service grants GM 32989 and BRSG S07 RR07044 to K.M.K. from the National Institutes of Health and by grant DCB-8504116 to R.C.F. from the National Science Foundation. G.S.H. was supported in part by a Gillette predoctoral fellowship.

#### LITERATURE CITED

1. Adams, R. L. P., and R. H. Burdon. 1982. DNA methylation in eukaryotes. *Crit. Rev. Biochem.* **13**:349-384.
2. Allen, S. L., and I. Gibson. 1973. Genetics of *Tetrahymena*, p. 307-373. In A. M. Elliott (ed.), *Biology of Tetrahymena*. Dowden, Hutchinson and Roth, Stroudsburg, Pa.
3. Allen, S. L., T. C. White, J. P. Langmore, and M. A. Swancutt. 1983. Highly purified micro- and macronuclei from *Tetrahymena thermophila* isolated by percoll gradients. *J. Protozool.* **30**:21-30.
4. Allis, C. D., and D. K. Dennison. 1982. Identification and purification of young macronuclear anlagen from conjugating cells of *Tetrahymena thermophila*. *Dev. Biol.* **93**:519-533.
5. Ammerman, D., and G. Steinbruck. 1981. Methylated bases in the DNA of the ciliate *Stylonychia mytilus*. *Eur. J. Cell Biol.* **24**:154-156.
6. Bannon, G. A., J. K. Bowen, M.-C. Yao, and M. A. Gorovskiy. 1984. *Tetrahymena* H4 genes: structure, evolution and organization in macro- and micronuclei. *Nucleic Acids Res.* **12**:1961-1975.
7. Bannon, G. A., F. J. Calzone, J. K. Bowen, C. D. Allis, and M. A. Gorovskiy. 1983. Multiple, independently regulated, polyadenylated messages for histone H3 and H4 in *Tetrahymena*. *Nucleic Acids Res.* **11**:3903-3917.
8. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.
9. Blackburn, E. H., W.-C. Pan, and C. C. Johnson. 1983. Methylation of ribosomal RNA genes in the macronucleus of *Tetrahymena thermophila*. *Nucleic Acids Res.* **11**:5131-5145.
10. Bromberg, S., K. Pratt, and S. Hattman. 1982. Sequence specificity of the DNA-adenine methylase in the protozoan *Tetrahymena thermophila*. *J. Bacteriol.* **150**:993-996.
11. Brooks, J. E., R. M. Blumenthal, and T. R. Gingeras. 1983. The isolation and characterization of *Escherichia coli* DNA adenine methylase (dam) gene. *Nucleic Acids Res.* **11**:837-851.
12. Bruns, P. J., and T. B. Brussard. 1974. Pair formation in *Tetrahymena pyriformis*, an inducible developmental system. *J. Exp. Zool.* **188**:337-344.
13. Bruns, P. J., and T. B. Brussard. 1974. Positive selection for mating with functional heterokaryons in *Tetrahymena pyriformis*. *Genetics* **78**:831-841.
14. Cartinhour, S. W., and G. A. Herrick. 1984. Three different macronuclear DNAs in *Oxytricha fallax* share a common sequence block. *Mol. Cell. Biol.* **4**:931-938.
15. Cech, T. R., and K. M. Karrer. 1980. Chromatin structure of the ribosomal RNA genes of *Tetrahymena thermophila* as analyzed by trimethylpsoralen crosslinking *in vivo*. *J. Mol. Biol.* **136**:395-416.
16. Cummings, D. J., A. Tait, and J. M. Goodard. 1974. Methylated bases in DNA from *Paramecium aurelia*. *Biochim. Biophys. Acta* **374**:1-11.
17. Doerfler, W. 1983. Methylation and gene activity. *Annu. Rev. Biochem.* **52**:93-124.
18. Dugaiczky, A., J. Hedgpeth, H. W. Boyer, and H. M. Goodman. 1974. Physical identity of the SV40 deoxyribonucleic acid



- sequence recognized by the Eco RI restriction endonuclease and modification methylase. *Biochemistry* **13**:503-512.
19. Findly, R. C., and T. Pederson. 1981. Regulated transcription of the genes for actin and heat-shock proteins in cultured *Drosophila* cells. *J. Cell Biol.* **88**:323-328.
  20. Frado, L.-L., A. T. Annunziato, and C. L. F. Woodcock. 1977. Structural repeating units in chromatin III. A comparison of chromatin subunits from vertebrate, ciliate and angiosperm species. *Biochim. Biophys. Acta* **475**:514-520.
  21. Gall, J. G. 1974. Free ribosomal RNA genes in the macronucleus of *Tetrahymena*. *Proc. Natl. Acad. Sci. USA* **71**:3078-3081.
  22. Gorovsky, M. A., C. Glover, C. A. Johmann, J. B. Keevert, D. J. Mathis, and M. Samuelson. 1977. Histones and chromatin structure in *Tetrahymena* macro- and micronuclei. *Cold Spring Harbor Symp. Quant. Biol.* **42**:493-503.
  23. Gorovsky, M. A., S. Hattman, and G. L. Pleger. 1973. [<sup>14</sup>N]methyl adenine in the nuclear DNA of a eucaryote, *Tetrahymena pyriformis*. *J. Cell Biol.* **56**:697-701.
  24. Gorovsky, M. A., M.-C. Yao, J. B. Keevert, and G. L. Pleger. 1975. Isolation of micro- and macronuclei of *Tetrahymena pyriformis*. *Methods Cell Biol.* **9**:311-327.
  25. Hallberg, R. L., K. W. Kraus, and R. C. Findly. 1984. Starved *Tetrahymena thermophila* cells that are unable to mount an effective heat shock response selectively degrade their rRNA. *Mol. Cell. Biol.* **4**:2170-2179.
  26. Harrison, G. S., and K. M. Karrer. 1985. DNA synthesis, methylation and degradation during conjugation in *Tetrahymena thermophila*. *Nucleic Acids Res.* **13**:73-87.
  27. Hattman, S. 1970. DNA Methylation, p. 517-548. In P. D. Boyer (ed.), *The enzymes*, vol. 14. Academic Press, Inc., New York.
  28. Hattman, S., C. Kenny, L. Berger, and K. Pratt. 1978. Comparative study of DNA methylation in three unicellular eucaryotes. *J. Bacteriol.* **135**:1156-1157.
  29. Howard, E. A., and E. H. Blackburn. 1985. Reproducible and variable genomic rearrangements occur in the developing somatic nucleus of the ciliate *Tetrahymena thermophila*. *Mol. Cell Biol.* **5**:2039-2050.
  30. Karrer, K. M. 1983. Germ line specific DNA sequences are present on all five micronuclear chromosomes in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **3**:1909-1919.
  31. Lacks, S., and B. Greenberg. 1975. A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA. *J. Biol. Chem.* **250**:4060-4066.
  32. 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.
  33. Leick, V., B. Bro, and P. Bruns. 1979. Macronuclear development: a model system to study specific gene expression in nuclear differentiation, 342-350. In J. Engberg, H. Klenow, and V. Leick (ed.), *Specific eukaryotic genes*. Alfred Benzon Symposium 13. Munksgaard, Copenhagen.
  34. Maniatis, T., A. Jeffrey, and D. G. Kleid. 1975. Nucleotide sequence of the rightward operator of phage  $\lambda$ . *Proc. Natl. Acad. Sci. USA* **72**:1184-1188.
  35. Martindale, D. W., C. D. Allis, and P. J. Bruns. 1982. Conjugation in *Tetrahymena thermophila*: a temporal analysis of cytological stages. *Exp. Cell Res.* **140**:227-236.
  36. Martindale, D. W., H. M. Martindale, and P. J. Bruns. 1986. *Tetrahymena* conjugation-induced genes: structure and organization in macro- and micronuclei. *Nucleic Acids Res.* **14**:1341-1354.
  37. McClelland, M. 1981. Purification and characterization of two new modification methylases; MclI from *Caryophanon latum* L and MtaqI from *Thermus aquaticus* YTI. *Nucleic Acids Res.* **9**:6795-6804.
  38. Pratt, K., and S. Hattman. 1981. Deoxyribonucleic acid methylation and chromatin organization in *Tetrahymena thermophila*. *Mol. Cell Biol.* **1**:600-608.
  39. Rae, P. M. M., and B. B. Spear. 1978. Macronuclear DNA of the hypotrichous ciliate *Oxytricha fallax*. *Proc. Natl. Acad. Sci. USA* **75**:4992-4996.
  40. Ray, D. S., and P. C. Hanawalt. 1964. Properties of the satellite DNA associated with the chloroplasts of *Euglena gracilis*. *J. Mol. Biol.* **9**:812-824.
  41. Razin, A., and H. Cedar. 1977. Distribution of 5-methylcytosine in chromatin. *Proc. Natl. Acad. Sci. USA* **74**:2725-2728.
  42. Razin, A., and J. Friedman. 1981. DNA methylation and its possible biological roles. *Prog. Nucleic Acid Res. Mol. Biol.* **25**:33-52.
  43. Solage, A., and H. Cedar. 1978. Organization of 5-methylcytosine in chromosomal DNA. *Biochemistry* **17**:2934-2938.
  44. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503-517.
  45. Taylor, J. H. 1984. DNA methylation and cellular differentiation. *Cell Biol. Monogr.* **11**:1-135.
  46. Van Etten, J. L., A. M. Schuster, L. Girton, D. E. Burbank, D. Swinton, and S. Hattman. 1985. DNA methylation of viruses infecting a eukaryotic *Chlorella*-like green alga. *Nucleic Acids Res.* **13**:3471-3478.
  47. Weiske-Benner, A., and W. A. Eckert. 1985. Differentiation of nuclear structure during the sexual cycle in *Tetrahymena thermophila*. I. Development and transcriptional activity of macronuclear anlagen. *Differentiation* **28**:225-236.
  48. Wenkert, D., and C. D. Allis. 1984. Timing of the appearance of macronuclear-specific histone variant hv 1 and gene expression in developing new macronuclei of *Tetrahymena thermophila*. *J. Cell Biol.* **98**:2107-2117.
  49. Woodard, J., M. Gorovsky, and E. Kaneshiro. 1968. Macronuclear subnuclei in *Tetrahymena*. *J. Cell Biol.* **39**:182a.
  50. Yao, M.-C., and M. Gorovsky. 1974. Comparison of the sequences of macronuclear and micronuclear DNA of *Tetrahymena pyriformis*. *Chromosoma (Berl.)* **48**:1-18.