DNA METHYLATION IN EUKARYOTES

Authors: **R. L. P. Adams**
R. H. Burdon
Department of Biochemistry
University of Glasgow
Glasgow, Scotland

Referee: J. Herbert Taylor
Institute of Molecular Biophysics
Florida State University
Talahassee, Florida

I. INTRODUCTION

A. Background

In addition to the four bases incorporated into DNA during semiconservative replication there is a variable number of other bases present in DNA in small amounts. These arise by postsynthetic modifications to some of the original four bases. In this review we shall consider primarily 5-methylcytosine (m^5C) and to a lesser extent 6-methylaminopurine (6-methyladenine, m^6A) as reports of other minor bases have either not been substantiated or indicate extremely low levels of the minor base. A number of other reviews have been published recently.[222-227]

Minor bases were discovered many years ago using the then new techniques of paper chromatography and electrophoresis. Table 1 gives some selected figures which indicate that while some species lack either m^5C or m^6A some plants have up to a third of their cytosines methylated. Indeed these figures, by giving the level of m^5C in total cellular DNA, hide the fact that certain fractions of the DNA may have levels of m^5C well above the average for the total DNA (see Section IV.B).

It is the very variable levels of base modifications which have made the study of DNA methylation so frustrating and have led to the conclusion that several different functions may be served by minor bases. Either these functions are dispensable in some species or are fulfilled by alternative mechanisms.

B. The Effect of Methylation on DNA Structure

As with thymine, the methyl group in m^5C is located in the major groove and does not change the base pairing characteristics of cytosine.[10] Neither does it produce a point of weakness in DNA as the presence of the methyl group reduces the rate at which reagents are able to be added across the 5.6 double bond. There is no evidence for a demethylation reaction[12] and indeed such a reaction which involves breaking a carbon:carbon bond is unlikely and is unknown for thymine. Deamination of m^5C can occur in vivo[13] probably by nonenzymic mechanisms and does occur in vitro especially in the presence of hot acid.[14,15] In vivo the deamination reaction occurs only rarely and leads to the production of thymine. This base substitution is the basis of some mutational hotspots which have been investigated in *Escherichia coli*[13] and is probably the cause of the unexpectedly low frequency of the dinucleotide CG in vertebrate DNA[16] (see Section V.C). In contrast, deamination of cytosine leads to formation of uracil which is recognized as an incorrect base by DNA repair enzymes[17] and removed.

Recently evidence has become available to show that DNA, in solution, may assume configurations other than the traditional B form. Thus a left-handed helical form

Table 1
METHYLCYTOSINE CONTENT OF DNA FROM SELECTED ORGANISMS

Organism	$\frac{m^5C \times 100}{C + m^5C}$	G+C (mol %)	Ref.
E. coli B	0	52.2	1
E. coli C	0.95	52.2	1
M. luteus	0	72.0	1
Tetrahymena pyriformis	<0.02	30	2
C. reinhardi	0.70	65	2
S. cerevisiae	0.3—1.0[a]	40	2
Microplasma arginini	0	26	3
Mosquito	0.17	42	4
Sea Urchin	6.5	35	5
Herring	8.9	45.2	6
X. laevis	7.2	41	7
Chick	4.9	44.2	6
Mouse	3.6	42	8
Calf thymus	5.4	44	8
Bull sperm	2.5	44	8
Wheat	31	41	9
Bracken	33	38	9

[a] S. Hattman now reports the absence of m^5C in yeast DNA (personal communication).

(Z DNA) is the preferred structure for DNA having alternating residues of C or more particularly m^5C and G.[18] Even the presence of a short string of four alternating C and G residues may cause a DNA duplex to change to the Z form over that region in high salt.[19] The low frequency of CG dinucleotides in vertebrate DNA would be expected to mitigate against their occurring in strings (but see below) but the use of specific antibodies has indicated the presence of m^5C and Z DNA in the polytene chromosomes of *Drosophila*.[39,212]

Although there is little evidence for clustering there are 6 CG dinucleotides in the Alu region preceding the human globin gene[192] and in the GC-rich ribosomal gene region of *Xenopus* there is a high incidence of CG dinucleotides, some of which occur in clusters.[193–196] Thus toward the end of the 18S rRNA gene of *X. laevis* there is the sequence CGCGCG,[194,195] and in the spacer between the 18S and 5.8S rRNA gene in *X. borealis* there is the sequence CGCGCGCG.[196] As most of such regions are methylated in the chromosomal genes[189] there may be a tendency for such regions to take up the Z conformation. It will be interesting to see whether DNA from the phage XP12 (in which m^5C completely replaces cytosine[20]) and from higher plants exists, at least partly, in the Z form.

The presence of m^5C in DNA affects both the buoyant density (lowered) and melting temperature (raised) of the DNA.[21–23] This may result from a slight increase in the energy of the H-bonds holding the DNA duplex together, or more likely an increase in base stacking interactions coupled with a change in structure.

The methyl group in m^6A is attached to the amino group involved in base pairing to thymine and this leads to a destabilization of these H-bonds.[24]

II. MEASUREMENT OF 5-METHYLCYTOSINE

A. Base Analysis

The amount of m^5C in a given DNA sample can be assayed by fractionation of the constituent mononucleotides, nucleosides, or purine and pyrimidine bases following DNA hydrolysis. The composition is then expressed as moles % m^5C or, as this fails to take into account the varying G + C content of different DNAs, a more convenient way of expressing the results is as the proportion of cytosines methylated, i.e., $\frac{mC}{C + mC} \times 100$.

Enzymic hydrolysis of DNA produces less deamination of m^5C than does hydrolysis in hot acid.[15] The deoxyribonucleotides can be separated by two-dimensional thin-layer chromatography,[25] but a column of Nucleosil 5C18 (Mackery and Nagel Co.)[23] gives a more convenient separation of deoxyribonucleosides.

Chemical hydrolysis is achieved using perchloric, formic, or trifluoracetic acid at elevated temperatures and many different methods have been used for base separation. Paper and thin-layer chromatographic methods are time consuming and often give separations inferior to those using high-performance liquid chromatography columns (see Figure 1). (For a summary of methods of separation see Reference 26.)

Although base analysis can be done using the optical absorption of the individual bases (usually at 260 nm), a more convenient and sensitive method involves the prelabeling of the DNA. The aim is to label the cytosine residues and then to measure the proportion of label recovered in m^5C (see Figure 5).[27] ^{14}C-labeled deoxycytidine is the most obvious precursor but is very expensive and available only at low specific activity. A more suitable precursor is [6-^3H]uridine. Although this is also incorporated into RNA, prior treatment of the DNA with alkali removes the contaminant. However, we have obtained lower figures for the proportion of cytosines methylated when using 6(^3H) uridine compared with ^{14}C deoxycytidine even in double label experiments and it is probable that the hydrogen (or tritium) in the 6 position undergoes increased exchange during methylation of the 5 position. A similar "NIH shift" has been noted when 5(^3H) uridine is converted to thymidine.[28,29] An alternative method of base analysis is to use mass spectroscopy which gives the ratio of m^5C to thymine from which it differs by only one mass unit.[30]

B. Use of Specific Antibody

Antibodies can be raised to m^5C and linked to a fluorescent label.[31] Such fluorescent antibodies have been used to show the unequal distribution of m^5C along and between chromosomes. Thus the centromeric heterochromatic regions of the chromosomes of the kangaroo rat stain intensely and these are the regions where the HS-β satellite is found.[32]

C. Restriction Enzyme Treatment

Several prokaryotic restriction enzymes will only act on DNA which is unmethylated at their site of cleavage. Certain enzymes, of which the most useful is HpaII, have a CG dinucleotide as part of their recognition sequence and will only cut DNA if the C is unmethylated. As most methylcytosine occurs in the CG dinucleotide, these enzymes can be used to investigate eukaryotic DNA. Thus HpaII cuts the sequence[33]

—C C G G—
—G G C C—

but will not cut

—C_mC G G—
—G G C_mC—

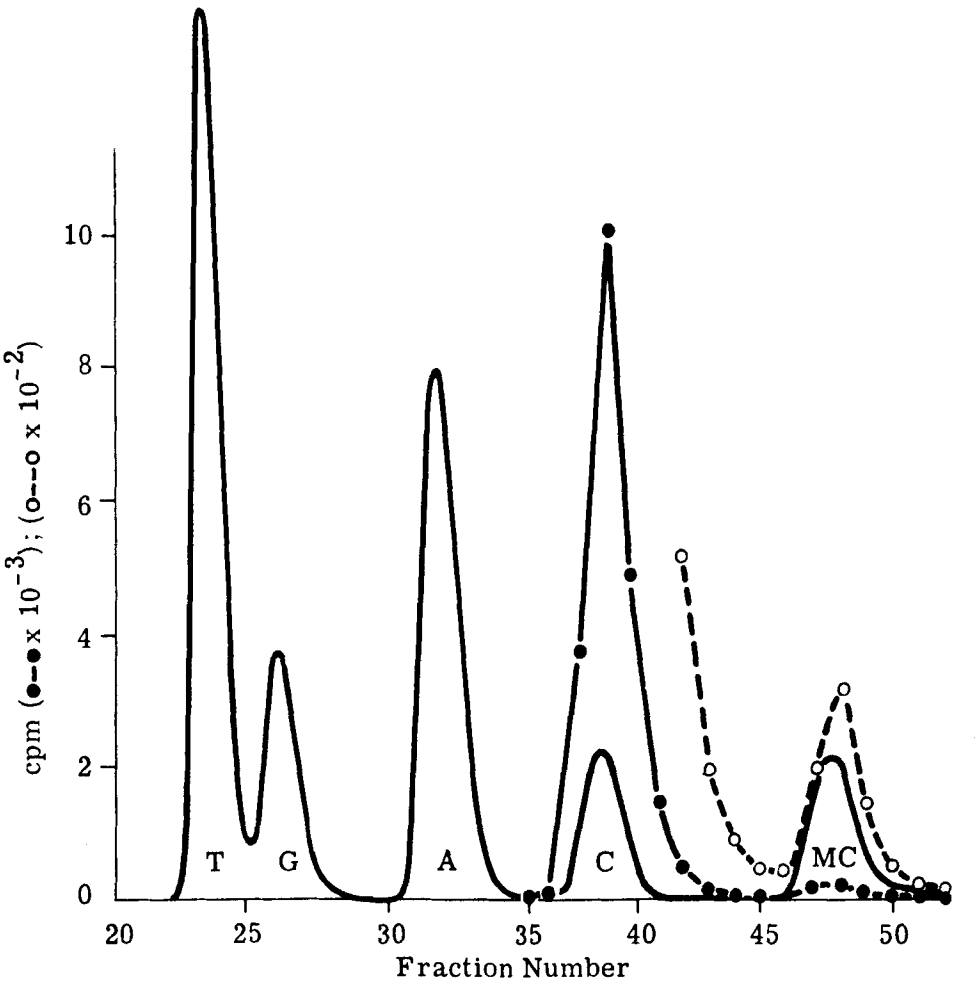


FIGURE 1. DNA base separation on Aminex A6. DNA isolated from mouse L929 cells labeled with [6 ^{3}H] uridine for 24 hr was hydrolyzed with formic acid and separated on a column of Aminex A6. The continuous line indicates the elution of marker bases (A: adenine, G: guanine, T: thymine, C: cytosine, mC: 5-methylcytosine). Fractions were collected and assayed for tritium radioactivity over the region of elution of cytosine and methylcytosine.

In contrast both these sequences are cut by MspI,[34,35] which will not, however, cut

$$\begin{array}{l} -{}_{m}C\ C\ G\ G- \\ -\ G\ G\ C\ C_{m}- \end{array}$$

The use of this pair of enzymes has enabled the extent of methylation of a DNA preparation to be established. The method relies on the assumption that CG dinucleotides in the sequence CCGG— are a random sample of all —CG— dinucleotides and there is no reason to doubt this.

When a preparation of DNA is treated with HpaII and then subjected to electrophoresis on an agarose gel, methylated vertebrate DNA is scarcely reduced in size. Parallel treatment with MspI releases fragments from mouse DNA of average size about 4.5 kbp which is little more than that predicted from a random distribution of MspI sites (3.8 kbp). The actual sizes can be obtained either by densitometric scanning of a

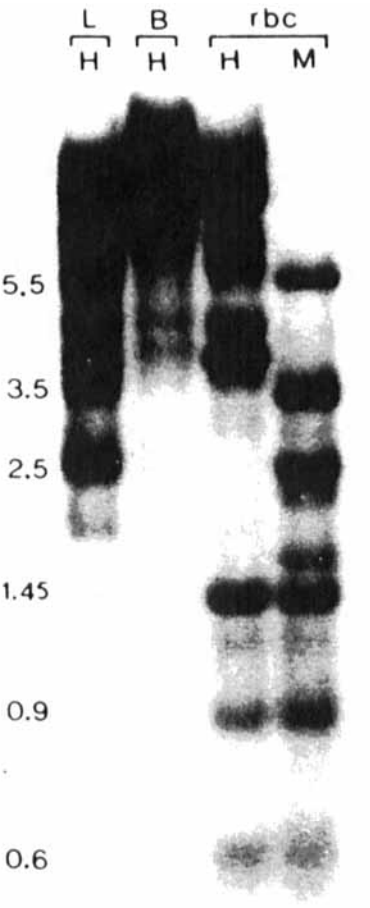


FIGURE 2. Digestion of chicken DNA with MspI or HpaII. DNA from 14-day definitive red blood cells (rbc), brain (B), or liver (L) was digested to completion with either MspI (M) or HpaII (H) and the restricted DNA fractionated by agarose gel electrophoresis. After transfer to a nitrocellulose sheet the DNA was hybridized to a radioactive probe containing the chick α globin genes (see Figure 8). This is the resulting autoradiogram. (From Weintraub, H., Larsen, A., and Groudine, M., *Cell*, 24, 333, 1981. With permission.)

photograph of the ethidium bromide stained gel[36] or by slicing a gel of radioactive DNA.

Restriction enzyme treatment can be coupled with "staining" with the m^5C antibody indirectly labeled with ^{125}I to demonstrate the distribution of m^5C in fragments of different size.[37]

Treatment of bulk preparations of animal cell DNA yield a smear on a gel following restriction enzyme digestion. This is a result of superposition of thousands of overlapping patterns. If, following electrophoresis, the DNA is transferred to a nitrocellulose sheet (Southern blot), it can be hybridized with a radioactive probe for a specific gene. The MspI digest of chicken DNA, when probed with radioactive globin DNA, yields a typical banding pattern[38,178] (Figure 2). The different pattern obtained following HpaII digestion

indicates methylation of specific sites in the globin gene (see Figure 2). This technique has been used to follow methylation of specific genes during development or between tissues (see Section VI.A.).

5′ end labeling of a DNA preparation from calf thymus digested with MspI showed that most of the terminal nucleotide was 5-methyl-cytidylate.[25] This indicates the symmetrical methylation of CCGG sequences in this DNA and the low level of unmethylated CG dinucleotides.

III. METHYLATION IN PROKARYOTES

For comparative purposes a brief review of bacterial methylation follows.

A. Restriction:Modification

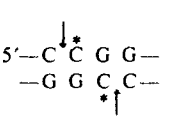
Bacteriophage have a specified host range. When a 'phage is transferred from growth on one host to a different host its efficiency is reduced by several orders of magnitude. However, those 'phage which do survive can now reinfect the second host with high efficiency.

The initial poor growth is caused by the action of bacterial endonucleases which attack the 'phage DNA. Host DNA is modified and this makes it resistant to these endonucleases. The modification usually involves the methylation of a cytosine or adenine at the site on the DNA recognized by the endonucleases.

The endonucleases are of two types. Class I restriction endonucleases are high molecular weight, multisubunit enzymes which show DNA methylating as well as endonuclease activity and require ATP and S-adenosylmethionine (AdoMet) for activity.[40-43] Class II restriction endonucleases have a single polypeptide chain of relatively low molecular weight and require neither ATP nor AdoMet for activity.[44-46] It is believed that for each class II endonuclease there is a corresponding but separate DNA methylase but this has been proved in only a few cases.

The class II endonucleases and methylases recognize the same symmetrical site on the DNA. The site usually consists of four or six bases which exist as a short palindrome. Thus the EcoRI endonuclease and methylase recognize

and the HpaII enzymes recognize



where the asterisk indicates the base methylated and the arrow the site cleaved by the restriction endonuclease.[46-48] If the sequence is not methylated, the endonuclease will cut the DNA as indicated by the arrows. If the DNA is methylated on one or both strands, no nuclease action occurs.

The Class I enzymes recognize a more complicated site on the DNA.[49-51] Again, however, it is a site where methylated bases may be present on both strands of the DNA. If both strands are methylated, the enzyme fails to interact with the recognition site. If one strand is methylated, the enzyme methylates the second strand. If neither strand is

methylated, the enzyme enters its restriction mode. Without releasing the recognition site the enzyme binds to a second nearby site on the DNA and, in the presence of ATP, a translocation is induced, resulting in the formation of a supercoiled loop which is then cleaved at one of a number of potential sites.

Thus in the case of the bacterial restriction phenomenon it is clear that methylation of DNA renders it resistant to endonuclease action. However a more general lesson to be learned from such findings is that enzymes can recognize and interact with sequences of nucleotides as short as four base pairs in length and that this recognition is profoundly affected by the presence of methyl groups. Endonucleases may be only one of many proteins whose binding is affected by DNA methylation which could thereby have a function in many other areas of DNA biochemistry. That any DNA:protein interaction may be complex is shown by the partial analysis of the structure of the EcoRI endonuclease and methylase. These enzymes, coded for by adjacent genes, have little sequence or structural similarity and yet interact with the same 6 base pair segment of DNA.[52,53]

B. The *dam* and dcm Methylases

E. coli DNA usually has about 0.5 mol% m^6A and 0.25 mol% m^5C which is very much more than can be accounted for by the restriction modification methylases of Class I or II.[1,54] Thus the EcoB methylase (the only one present in plasmid-free *E. coli* B) accounts for only 0.08 mol% m^6A and mutants have been obtained which have only this low level of m^6A.[55,56] These mutants (*dam* mutants) lack an enzyme that methylates the sequence GATC and the calculated number of GATC sites can account for all the m^6A present in wild-type DNA.[57-59] *E. coli* B lacks m^5C but an enzyme in *E. coli* C (the dcm protein) methylates the sequence $CC^{\hat{A}}_{\hat{T}}GG$ (and probably some other sequences).[58,60]

Although the RII plasmid codes for a restriction enzyme which recognizes the same site as the dcm methylase it is believed for a variety of reasons that the dam and dcm proteins do not form part of a restriction modification system but that the methyl groups introduced by these enzymes possibly play a role in recombination events.[54,56] However, the *dam*-3 mutant is more sensitive than the wild type to ultraviolet radiation and mitomycin which suggests that these methyl groups may play a role in excision repair though no details are available of what this might be.[56]

In contrast, evidence has become available recently that methylation of bacterial DNA may direct the repair endonuclease to the incorrect strand in mismatch repair.[61-63] The problem here is that when an incorrect but normal base is inserted into DNA during replication this is normally removed by proofreading. If this fails, some arrangement is required to indicate which is the incorrect daughter strand of the duplex DNA. This may be brought about by a repair mechanism recognizing the daughter strand in the few seconds after synthesis, before methylation is complete (see next section). *dam* cells show high mutation rates and the *dam* gene is one of a series of mutant loci involved in mismatch correction. This has been shown using genetic markers in mismatched, hemimethylated lambda DNA when the base in the unmethylated strand is the one which is repaired.

Gomez-Eichelmann and Lark[64] have suggested that methyladenine residues are not located randomly on the DNA but that they occur toward the ends of Okazaki pieces. However, as *dam* mutants replicate normally it may be that the sequence GATC (methylated or unmethylated) plays some role in the termination of synthesis of Okazaki pieces.

C. Timing of Methylation

In 1968 Billen showed that in *E. coli* methylation of DNA occurs only on the nascent

DNA close to the growing replication fork.[65] This is confirmed by the finding that bacterial methylases fail to add methyl groups to homologous DNA, implying that this DNA is already fully methylated.[40]

Although methylation occurs rapidly it probably is delayed until the Okazaki pieces have been ligated, as in a ligase-deficient mutant the level of methylation in the Okazaki pieces which accumulate is considerably lower than normal.[66]

Such rapid methylation can be prevented by growing methionine auxotrophs in the absence of methionine.[67,68] The ongoing round of replication is completed and DNA lacking methyl groups in the daughter strand is formed. New rounds of DNA synthesis are only reinitiated on addition of methionine when the unmethylated strand is rapidly modified.

IV. METHYLATION IN EUKARYOTES

A. Species and Tissue Variation

The level of m^5C present in the total nuclear DNA of higher eukaryotes varies from the extremely low values found in insects where only 1 cytosine in 600 is methylated to the situation in higher plants where one third of cytosines may be methylated (Table 1). Vertebrates tend to lie in the middle of the range and have about 1 m^5C for every 30 cytosines in their DNA while lower eukaryotes either lack methylcytosine or have less than 1 cytosine in 100 methylated. Some lower eukaryotes also have a low level of m^6A which is also present in trace amounts in insect DNA.

These very variable levels of m^5C in DNA pose a problem when searching for a function for DNA methylation. Obviously if methylation is to play an important part in regulating gene expression in vertebrates, then this process must be regulated in a very different manner in insects and in plants.

Careful analysis of the levels of m^5C in the DNA isolated from different tissues of one animal have revealed minor variations which have been claimed as highly significant. Thus the group headed by Vanyushin has shown differences in the following situations:

1. A difference between the total level of m^5C of DNA from several tissues. This was shown for bull, rabbit, mouse, pig, and sheep as well as several other vertebrate and invertebrate species[69,70]
2. An effect of disease on the levels of m^5C in cotton[71]
3. An effect of hormones on the levels of m^5C in liver, etc.[73-75]
4. An effect of aging on the levels of m^5C in cows[73-75]

Minor changes have also been reported in the level of m^5C during differentiation of teratocarcinoma cells[76,77] and the induction of globin synthesis in Friend cells may be accompanied by lower levels of DNA methylation.[78,79] Kappler[80] has also shown a variation in the levels of m^5C in the DNA of various cultured chick cells and there are a number of reports on the effects of transformation on the levels of m^5C in the DNA of cells in culture.[81,82] Results with cultured cells, and especially those which have been transformed, may tell us no more than that the chromosomes present are different in the different cell lines. Thus if one chromosome has a higher than average content of m^5C, a significant reduction in the overall value will accompany the loss of this chromosome or part of it. No significant changes have been found in the level of m^5C during development of sea urchins,[83,85] *Xenopus*,[84] or mouse[76] but during early rabbit development the extensively methylated zygote after 9 to 10 divisions has differentiated into a totipotent embryoblast with the same level of methylation as the zygote and a unipotent

trophoblast (95% of the cells) which has lost m^5C and now has a level typical of various adult tissues.[86]

It has been difficult to assess the significance of the variations found by Vanyushin's group especially as other groups have not always obtained corroborating results.[79,83,84,87] In most cases the differences reported are small but were shown to be reproducible by very careful work. In other cases large variations have been found and, for instance, the dramatically lower level of m^5C found in bull sperm as compared to other bovine tissue has been confirmed by several other laboratories.[88-90]

Recently a new technique has been used to estimate the extent of methylation of DNA. This involves the use of restriction enzymes whose action is blocked by DNA methylation. Such enzymes can be used to ask whether or not their recognition site is methylated in eukaryotic DNA. The use of such enzymes and in particular of HpaII has confirmed the results obtained by chemical analysis and has indicated a number of further complications.

Thus the genome of the sea urchin is made up of two fractions.[91] One fraction is insensitive to HpaII and is presumed to be heavily methylated while the other fraction is degrated to fragments ranging in size from 1 to 2 kbp. This is the expected result for DNA containing randomly positioned unmethylated HpaII sites.

Bird and Taggart[92] have divided animals into three groups: those with most of their CCGG sites methylated (vertebrates), those with few CCGG sites methylated (arthropods), and an intermediate group which shows a methylated and an unmethylated compartment similar to the sea urchin (invertebrates other than arthropods). There is a correlation between the extent of methylation and the deficiency in the CG dinucleotide in the genome of the three groups[93] which supports the suggestion that the deficiency arises as a result over millions of years of deamination of m^5C in highly methylated DNA of vertebrates. One must conclude that those m^5C dinucleotides which remain must be essential; otherwise they would have been inexorably lost by deamination.

Apart from these gross differences, the use of restriction enzymes coupled with probes for specific genes has shown that the pattern of methylation of such genes varies from tissue to tissue. (For more details see Section VI.A.)

The cumulative effect over the whole genome of a myriad of specific variations may have no significant effect on the total level of methylation of nuclear DNA. However in some cases small differences may be detected with careful technique.

B. Methylation of DNA Fractions of Different Sequence Complexity

The amount of m^5C in DNA is related to some extent to the frequency of occurrence of the dinucleotide CG. Thus certain simple DNA sequences have very high levels of methylation, e.g., the HS-β satellite DNA of the kangaroo rat has a basic 10 base pair repeat containing one methylcytosine[94] and the basic 34 base pair satellite DNA from the plant Scilla has more methylcytosine than cytosine.[95] Some of these methylcytosines are in mCT and mCA dinucleotides. Using a fluorescent antibody to m^5C it has been shown that the centromeric heterochromatic regions of the kangaroo rat chromosomes are enriched in m^5C and this is the location of the HS-β satellite.[32] That both AT- and GC-rich satellite DNAs often have high levels of m^5C[97-99] may point to a special function of the minor base in satellite DNA. Considerable tissue variation exists in the extent of methylation of some satellite DNAs. Thus in bovine sperm, satellites are not methylated (at HpaII sites) whereas the same satellites are methylated in thymus, skin, and thyroid.[88-90]

In addition to satellite DNA, foldback DNA is enriched in m^5C.[74,100-102] This DNA contains adjacent regions of inverted repeats whose function is as yet not known. The

high levels of m^5C are recovered in the repeats following S1 treatment of the isolated foldback DNA suggesting that, like the higher levels of m^5C found in some satellite DNAs, this merely reflects the sequence of the repeat. Moreover, the Scilla satellite referred to above[95] is palindromic, i.e., it is also foldback DNA. The proportion of DNA recovered in foldback and highly reiterated fractions is increased in certain tumors and transformed cells which therefore have a higher overall level of methylation than their normal counterpart.[81,82,102,103] The level of methylation of the foldback DNA fraction from bovine tissues falls with the age of the cow.

Schneiderman and Billen[100] showed that rapidly reannealing DNA from Chinese hamster ovary cells is enriched two- to fourfold in m^5C and this was particularly true for DNA made early in S-phase. They suggested that the origins of replicons may be enriched in such sequences. This enrichment of rapidly reannealing DNA in m^5C is not affected by the presence in the DNA of high levels of bromouracil[105] or azacytosine.[198]

These variations in the level of m^5C in DNA fractions of different complexity are probably the cause of the variation of methylation of DNA fractions of different buoyant density or melting temperature. Thus DNA fractions rich in G + C are more highly methylated in human and hamster cells whereas it is the A + T rich fraction which is enriched in methyl groups in mouse DNA.[103] Fractions of mouse DNA which melt at lower temperature are enriched threefold in methyl groups relative to bulk DNA.[198]

The intermediate repetitive and unique classes of DNA are usually either all highly methylated (vertebrates) or unmethylated (arthropods). However, in sea urchins[91] and the slime mould *P. polycephalum,*[104] the genome contains two compartments: M+ is highly methylated and M− is only slightly methylated. The two compartments can be separated by electrophoresis following incubation with HpaII which cleaves M− DNA into low molecular weight fragments. Repetitive DNA sequences are found in both compartments but primarily in the M+ compartment in sea urchins. Certain multiple copy genes (e.g., 5S and histone) are found in the M− compartment of sea urchins DNA.

C. Methylation of Chromatin

Adams et al.[107] could find no significant difference between the m^5C content of nucleosomal core DNA and total DNA of Chinese hamster cells. However, Cedar and co-workers[108,109] showed a preferential localization of m^5C in micrococcal nuclease-resistant DNA. It may be that in some species which have highly methylated satellite DNA located in heterochromatic regions such chromatin is more resistant to micrococcal nuclease digestion and this results in an apparent enrichment of residual DNA in m^5C.

What the effect of a nucleosomal structure might be on DNA methylation in vivo is unknown but in vitro we have shown with mouse nuclei and mouse DNA methylase that methyl groups are added predominantly to micrococcal nuclease-sensitive DNA[106] This is presumed to be linker DNA but must also include certain regions of core DNA. Similar results using bacterial DNA methylases have been reported,[110,111] and Tosi and Scarano have shown that treatment of sea urchin nuclei with trypsin considerably increases the ability of their DNA to accept methyl groups in vitro.[112] These results indicate that the presence of nucleosomes may prevent the access or function of a DNA methylase in vivo.

As nucleosomes (or at least H3 and H4 histones) do not leave the leading strand of DNA at the replication fork,[113,114] it is likely that the daughter strand of DNA becomes associated with the parental nucleosomes before methylation has occurred (see Section V.B.). This would lead to hemimethylated DNA in protected regions which in the next generation would produce a distinct pattern of methylated and unmethylated —CG— dinucleotides in one quarter of the cells (see Section V.D and Figure 3). Such a pattern would be lost if the nucleosomes were allowed to move along the DNA but

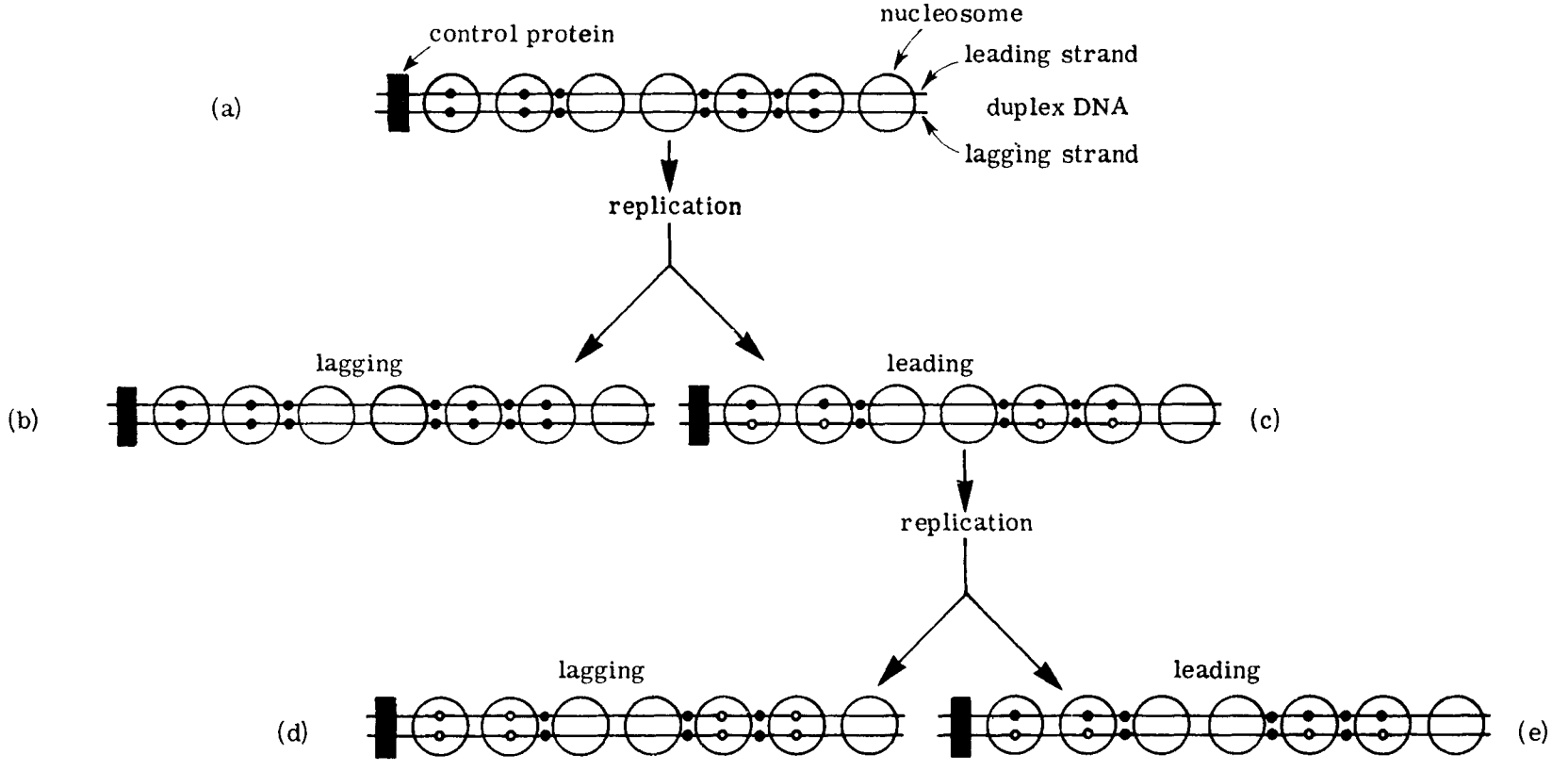


FIGURE 3. The possible effect of nucleosomes on DNA methylation following DNA replication (compare Figure 7). (●) A methylated site. (○) A potential, but unmethylated site. (a) Nucleosomes are phased on transcribing DNA by the presence of a control protein. (b) Lagging strand: following replication the methylation pattern is maintained before new nucleosomes are formed. (c) Leading strand: the daughter strand cannot be methylated where covered by conserved nucleosomes. (d) Lagging strand: as the parental strand lacks methyl groups at some sites these will not be methylated in the daughter strand. (e) Leading strand: the pattern of c will again be produced.

would be maintained for as long as the nucleosomes remained "phased". Rather than leading to a preferential location of m^5C in core DNA this argument would predict the presence of unmethylated —CG— only in core DNA.

Brief treatment of nuclei with micrococcal nuclease releases nucleosomes which are believed to be enriched in transcribing DNA.[115] Similarly such DNA is preferentially solubilized by treatment with DNase I.[116] In both these cases the transcribing DNA is found by some methods to be somewhat enriched in m^5C.[107,117] This finding is not actually inconsistent with the undermethylation of certain CG dinucleotides in transcribing DNA (see below and Section VI.A). Taken together with the argument that in vertebrates there has been a loss of m^5C dinucleotides during evolution this result may simply reflect the retention of a greater proportion of essential (methylated) CG dinucleotides in coding regions. Indeed late replicating, heterochromatic DNA is deficient in m^5C.[136,140]

By using DNase I to preferentially nick DNA in transcribing chromatin Naveh-Many and Cedar[118] have used radioactive triphosphates and *E. coli* DNA polymerase to nick translate these regions. By this method they have shown that not only do transcribing regions have a smaller proportion of CCGG sequences methylated but also only about one quarter of the CG dinucleotides are methylated.

The use of fluorescent antibodies to m^5C has shown a concentration of this base in C-band regions of heterochromatin of mouse and human chromosomes where satellite DNA occurs.[119] This is consistent with the finding of high levels of m^5C in certain satellite DNAs (Section IV.B). However, a fairly even, low intensity reaction is seen over most of the chromosomes, consistent with the lower level of methylation found in unique and intermediate repetitive DNA.

On brief incubation of nuclei with micrococcal nuclease the bulk of the DNA is solubilized in the form of oligonucleosomes.[107] Much protein and DNA methylase activity is also solubilized. On further incubation with micrococcal nuclease nucleohistone is again precipitated but most of the DNA methylase activity remains soluble. This might indicate a preferential localization of the enzyme on linker DNA.[120]

D. Methylation of Extrachromosomal DNA

1. Mitochondrial DNA

Vanyushin and Kirnos[121,122] have reported that in several vertebrates and protozoa the mitochondrial DNA has a higher proportion of m^5C than does the corresponding nuclear DNA. Nass[123] found the opposite result with several animal mitochondrial DNAs studied and David[124] was not able to detect m^5C in mitochondrial DNA from HeLa or *Xenopus* cells.

Positive results have been criticized on the basis that the mitochondrial DNA preparations used may have been contaminated with nuclear DNA and Groot and Kroon[125] failed to detect m^5C in the CCGG sequence in a variety of mitochondrial DNAs. The "contamination" argument cannot, however, explain the results of Vanyushin and Kirnos (more m^5C in mitochondrial DNA) and the possibility remains that mitochondrial DNA contains m^5C in sequences other than —CG—, thus reflecting the postulated prokaryotic origin of mitochondria.

2. Chloroplast DNA

Chloroplasts of *Chlamydomonas* in the vegetative stage have no m^5C in their DNA but chloroplast DNA in female gametes and in zygotes is methylated at the sequence —CCGG— among others.[126-129] Chloroplast DNA in male gametes is not methylated and is degraded in zygotes showing that a system akin to the bacterial restriction modification system is operating.

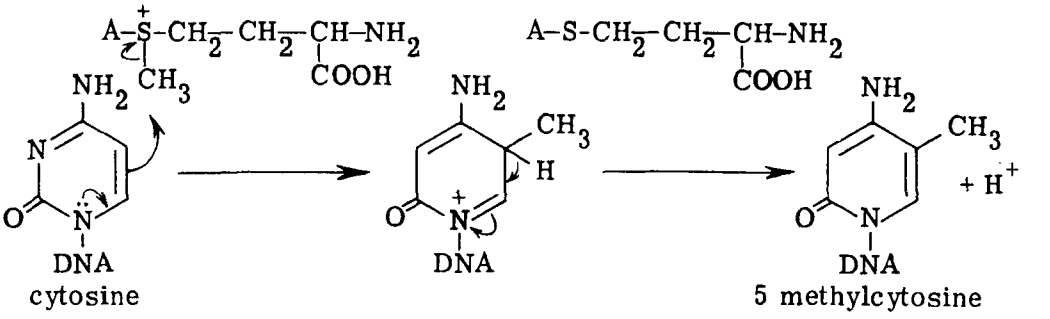


FIGURE 4. DNA cytosine methylation. The reaction catalyzed by DNA cytosine methylase (A–adenosine).

A mutation (*mat* 1) leads to methylation of male chloroplast DNA suggesting that in wild-type mt⁻(male) cells the expression or function of the methylase gene is repressed by a factor closely linked to the mating-type locus.[128]

3. Viral DNA

There is no evidence for methylation of animal virus DNA when it has an extrachromosomal location.[15,130–132] However, viral DNA can be methylated in vitro and is methylated at least partially when it is integrated into the chromosomal DNA of transformed cells.[131] This problem is considered further in Section VI.A.4.

V. SYNTHESIS OF m^5C

A. The In Vivo Reaction

m^5C is produced by the transfer of a methyl group from S-adenosyl-L-methionine to the 5 position on the pyrimidine ring of certain cytosines in DNA[133] (see Figure 4). The reaction is catalyzed by a DNA methylase or DNA (cytosine-5-)S-adenosyl-L-methionine: methyl transferase (E.C.2.1.1.37).

The S-adenosyl methionine itself arises from methionine and ATP and so in vivo labeled methyl groups on methionine can be transferred to DNA. This is one of the approaches used to follow DNA methylation.

A second approach to following methylation in vivo is to measure the proportion of radiolabeled cytosine in DNA which is converted to methylcytosine.[27] This, perforce, requires a method of base separation (see Section II) but is free from the criticism that methyl-labeled methionine can transfer radioactivity to bases other than cytosine leading to the production, in particular, of radioactive thymidine.[4] (The incorporation of radioactivity from methionine into bases other than methylcytosine can largely be prevented by incubating in the presence of (1) 20 m*M* Na formate to swamp the 1 C pool, (2) aminopterin to block folic acid metabolism and hence thymidylate synthesis, and (3) thymidine to circumvent the need for endogenous thymidylate synthesis and to dilute out any radioactive thymidylate which may be produced by failure of (1) or (2) to be 100% effective).[12]

The cytosine in DNA can be rendered radioactive most effectively by labeling with ^{14}C deoxycytidine or 6-^3H uridine and the former was first exploited by Kappler[27] who showed that in cultured mouse cells methylation of DNA occurred very shortly following synthesis (see below and Section II.A).

B. Methylation and DNA Synthesis

Following administration of uniformly ^{14}C-labeled deoxycytidine to exponentially

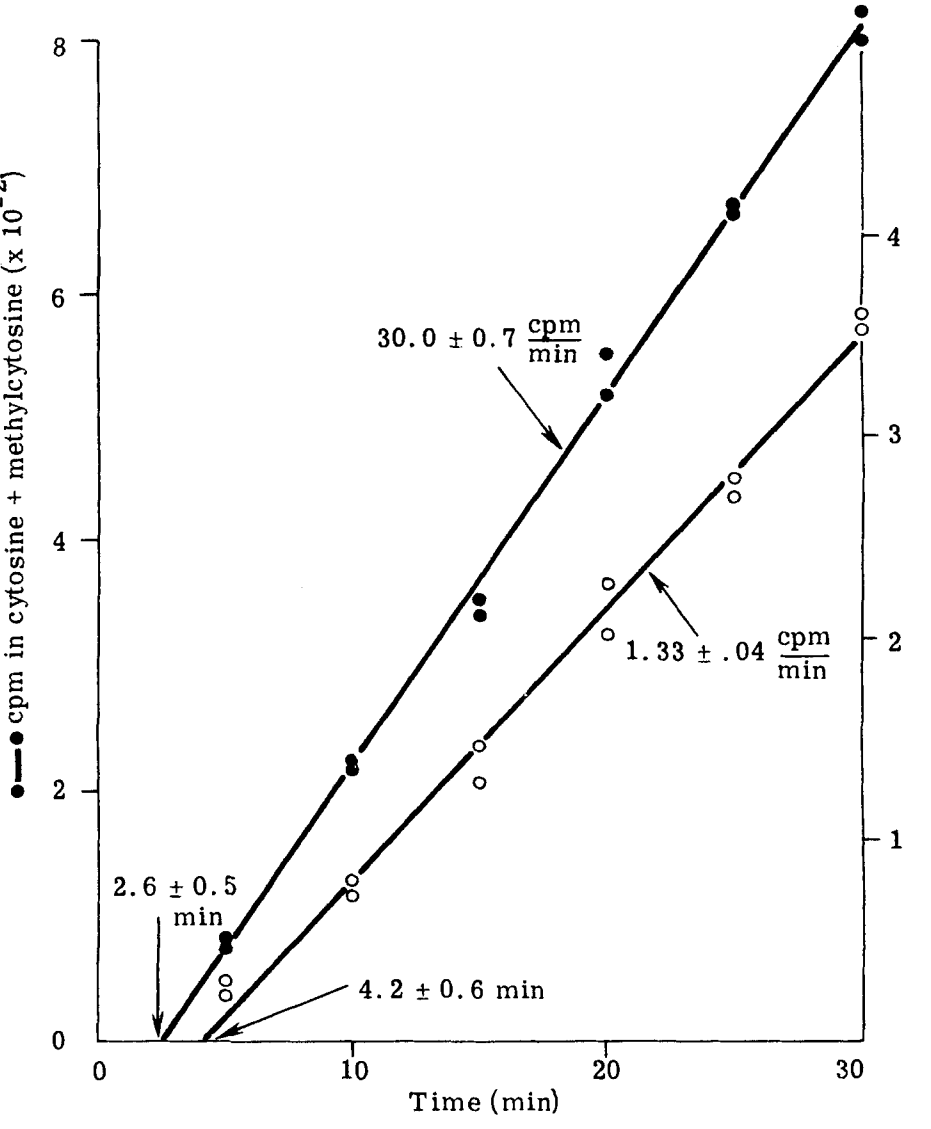


FIGURE 5. [^{14}C] deoxycytidine used to measure the rate and extent of DNA methylation. At zero time [^{14}C] deoxycytidine was added to a series of cultures of mouse adrenal cells. Duplicate cultures were harvested at the indicated times; DNA was isolated and hydrolyzed in perchloric acid and the bases separated by chromatography. The value cpm in methylcytosine: cpm in cytosine + methylcytosine gives a measure of the proportion of cytosines methylated. The best straight lines intersect the time axis at 2.6 and 4.2 min indicating the delay in incorporation of exogenous deoxycytidine into DNA and the methylation of DNA cytosine, respectively. (From Kappler, J., *J. Cell Physiol.*, 75, 21, 1970. With permission.)

growing cultured mouse cells, Kappler[27] showed a linear rate of incorporation of radioactivity into DNA cytosine and methylcytosine. Extrapolating back to zero radioactivity, his interpretation was that the DNA became fully methylated 1.6 min after synthesis (see Figure 5). One problem with short incubation times especially with ^{14}C-labeled precursor is that the amount of incorporated radioactivity is necessarily small and other workers have shown a considerable delay between the time of synthesis of DNA and its eventual complete methylation.[12,134-139] This is not to say that the bulk of methylation

does not occur very shortly after synthesis but that methylation is not complete for several hours. Such conclusions have been obtained by following DNA synthesis and methylation in synchronized cells when the peak rate of methylation follows the peak rate of synthesis by up to 1 hr (Figure 6). Furthermore, using pulses of bromodeoxyuridine and radioactive methionine it can be shown that in vivo some methylation is occurring on DNA which was synthesized before the methionine was added.[136] Similarly, methylation continues when DNA synthesis is inhibited by hydroxyurea or aminopterin.[12,27,136]

Kappler[27] argued that some of the apparent lag observed in synchronized cells may arise if higher levels of methylation were a characteristic of DNA made later in S-phase but in fact the opposite seems to be true[136,140] (see also Sections IV.B and C). In addition, the fact that homologous DNA can act as an acceptor of methyl groups[26,141] in a DNA methylase assay points to the conclusion that DNA isolated from exponentially growing cells is not fully methylated. DNA isolated from nongrowing cells is, in general, a poor acceptor of methyl groups in vitro[141] and we could find no evidence for incorporation of methyl groups in vivo into DNA of stationary mouse cells or lymphocytes prior to stimulation of DNA synthesis with phytohemagglutinin.[136]

In certain instances it appears that methylation continues well beyond the end of S-phase[134,138,142,144] but in general it is supposed that methylation would be complete prior to initiation of a new round of DNA synthesis. If this were not so, the consequences would be a change in the pattern of methylation (see Section V.D).

Although initiation of methylation occurs very shortly after synthesis of DNA, it does appear that the Okazaki pieces typical of nascent DNA are ligated prior to methylation[135,145,146] (although Kiryanov et al.[139] did detect some methylation of Okazaki pieces). It is also probable that the nucleosomal structure is reestablished prior to methylation[113,114] but the greater sensitivity to nuclease of newly synthesized DNA is not lost for about 30 min,[147] by which time methylation is well underway.

Incubation of nuclei with S-adenosyl [^{3}H Me]methionine leads to incorporation of radioactivity into methylcytosine.[137,148] The DNA which is methylated in isolated nuclei is predominantly that made shortly before nuclear isolation, but a considerable number of methyl groups are added to DNA made several hours earlier.[148] No methyl groups were added to the parental DNA strand made in the previous cell generation, showing that methyl groups are being added to hemimethylated sites unfilled in vivo (see Section V.D). This contrasts the finding with isolated DNA where methyl groups can be added to both strands by a purified DNA methylase.[26]

From the foregoing discussion it is clear that the results reported are consistent with an addition in vivo of methyl groups to hemimethylated sites produced on DNA replication (see Figure 7) although there does appear to be a lag before methylation is complete. This conclusion is substantiated by results from experiments where a cloned chick thymidine kinase gene, methylated with HpaII methylase, was introduced into tk$^-$ cells. The pattern of methylation was maintained for 25 cell generations but not with 100% fidelity.[149]

C. Nucleotide Sequences Surrounding m^{5}C

Early work involving digestion of DNA with DNase I showed that methylcytosine was recovered predominantly in the dinucleotide-m^{5}CG—.[150] This method, however, results in variable recovery of dinucleotides and the presence of methylcytosine in other sequences could not be ruled out. The dinucleotide —CG— is the same as its complement and to this extent the location of methyl groups in eukaryotes is similar to in prokaryotes. This symmetrical arrangement of methylated bases is believed to be very important especially in the replication of methylated DNA (see Figure 7 and Section V.D).

A comparison of pyrimidine isostics leads to the conclusion that methylcytosine occurs

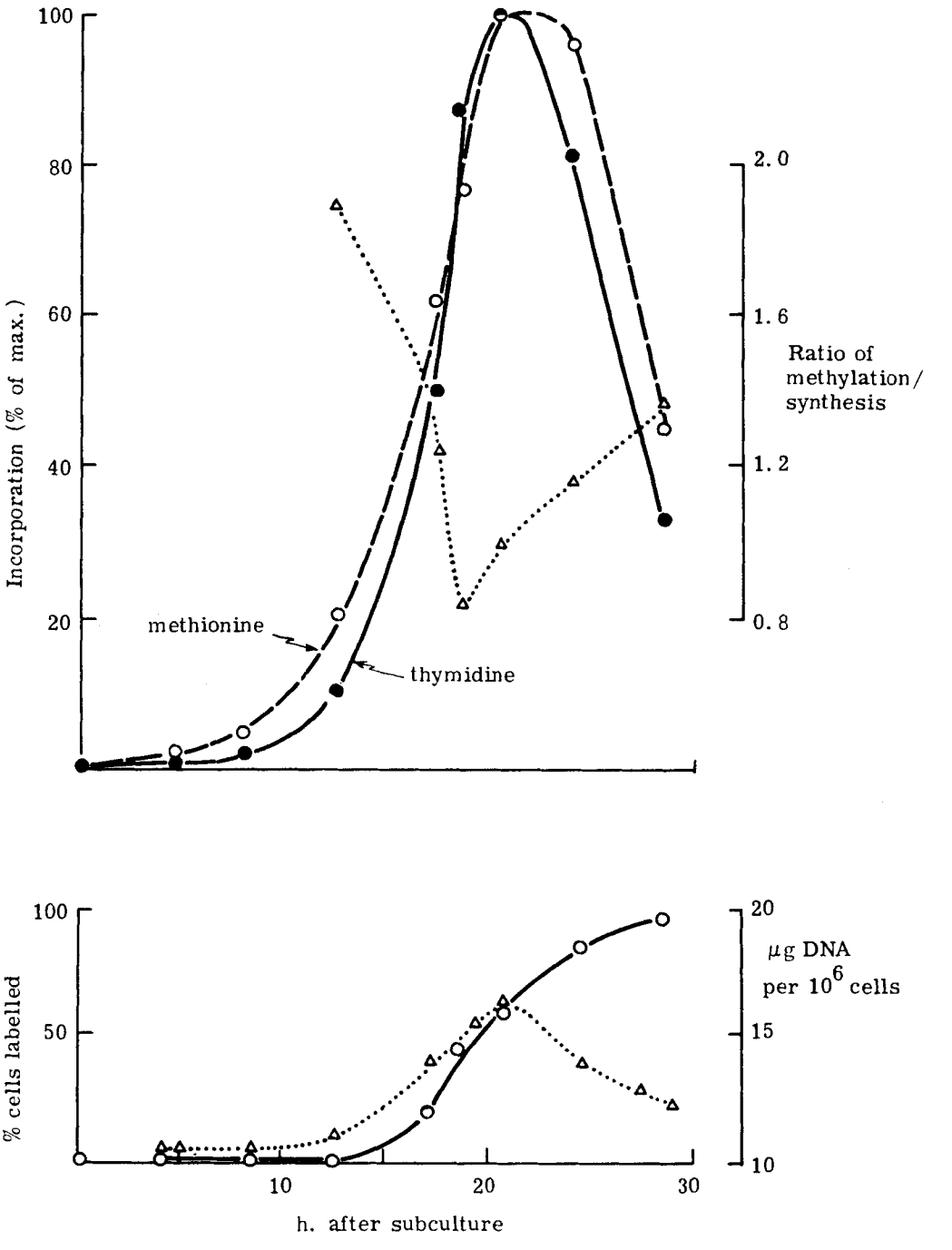


FIGURE 6. The lag period between DNA synthesis and methylation. Mouse L929 cells were released from stationary phase at 0 hr and the rate DNA synthesis and methylation estimated by measuring the incorporation of radioactivity into the appropriate base following a 1-hr incubation with [2 ^{14}C] thymidine and [Me ^{3}H] methionine in the presence of aminopterin (0.2 μM), adenosine (60 μM), glycine (80 μM), deoxycytidine (20 μM), and sodium formate (20 μM). The triangles indicate the ratio of methylation to DNA synthesis and show that early replicated DNA is enriched in mC but that the peak rate of methylation lags behind the peak rate of DNA synthesis. (From Adams, R.L.P. and Hogarth, C., *Biochim. Biophys. Acta*, 331, 214, 1973. With permission.)

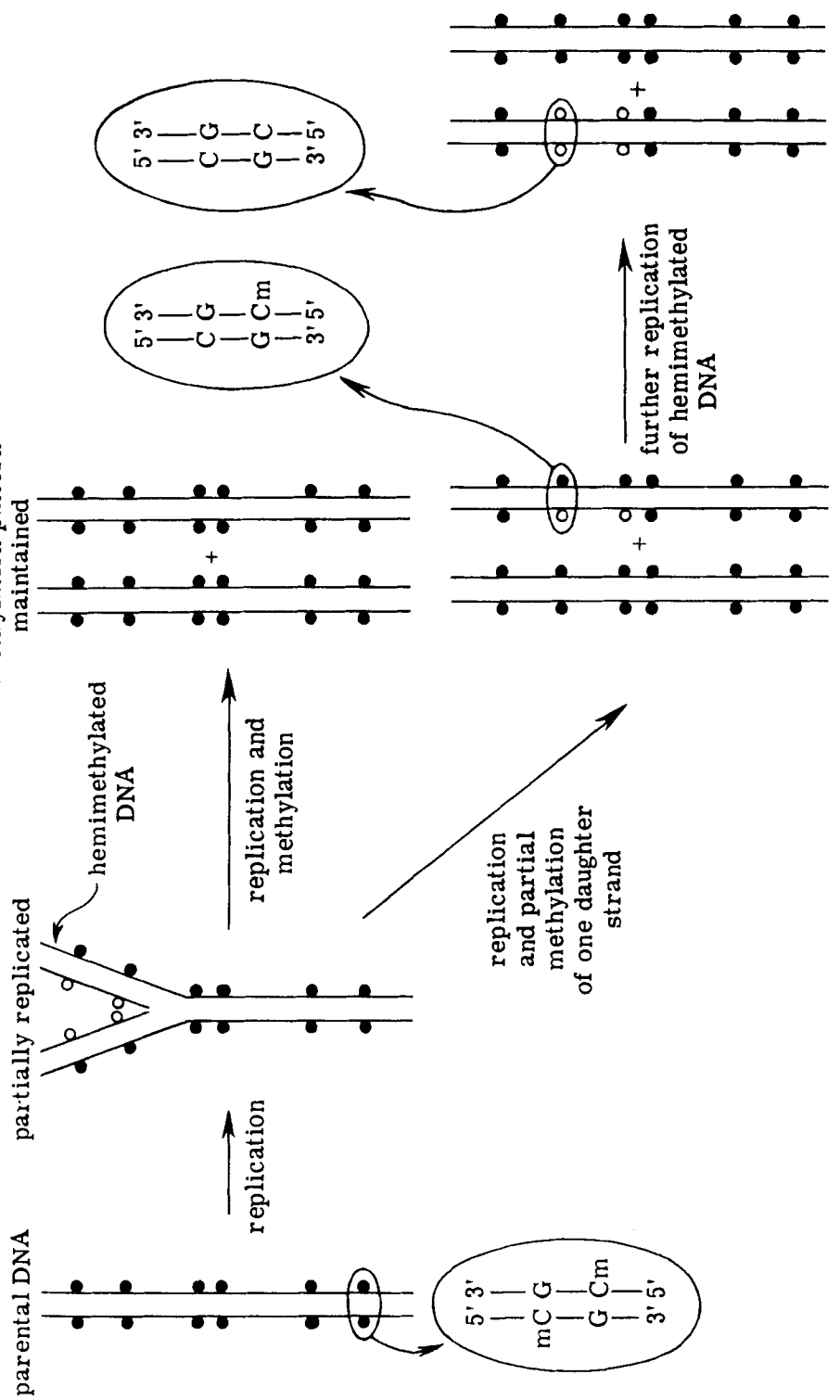


FIGURE 7. Replication and methylation of DNA. The pattern of methylation of the parental strand is conserved in the daughter DNA by the action of a maintenance methylase acting shortly after DNA synthesis. Alteration of the pattern of methylation can occur by replication in the absence of methylation. ● represents a methyl group on a CG dinucleotide. ○ represents a potential but unmethylated site on a CG dinucleotide.

predominantly in the sequence —NmCR— (presumably largely —NmCG).[151-155] The nature of N is what might be predicted from a random choice of bases which suggests that no further sequence specificity is involved. This conclusion is supported by two further observations:

1. When stretches of DNA containing m^5C have been sequenced no obvious neighbors or limitations on location of m^5C are readily apparent.
2. Several restriction enzymes which have a CG in their recognition sequence are inhibited by the methyl groups in eukaryotic DNA (see Section II.C).

In contradistinction however the action of MspI is blocked in some cases and this has been shown to be caused by methylation of the 5'C in the sequence —CCGG.[35,154]

In higher plants methyl groups are found in the symmetrical sequence mC.N.G. as well as in CG, CC, and CA dinucleotides.[95,155]

Although most methylcytosine is found in the mCG dinucleotide it must not be forgotten when considering the maintenance of methylation patterns that some methylcytosine also occurs in dinucleotides such as mCA, mCC, and mCT.[35,151,156-158]

It is clear, however, that not all cytosines in the CG dinucleotide are methylated in animal cell DNA. Selective methylation of cytosines could modulate the efficiency with which certain DNA sequences interact with specific regulating molecules. For instance conversion of an AT base pair in the lac operator to GmC does not affect repressor binding but conversion to GC reduces the affinity of the repressor several fold.[159]

Thus methylation can increase the information content of DNA. If *all* CG dinucleotides were methylated there would be no such increase in the information content of the DNA.

It has been shown that cytosine and m^5C can be deaminated both in vivo and in vitro.[13-15] Although deamination of cytosine is more probable, the product, uracil, is recognized as an abnormal base in DNA and a specific repair mechanism exists to remove it.[17] The deamination product of m^5C is thymine, a normal base, and hence this leads to a base substitution mutation. Such a deamination is the cause of the mutational hotspots found in *E. coli*.[13]

The long-term result of deamination of methylcytosine is the conversion of the —CG— dinucleotide into TG and CA and vertebrate DNA has only about one fifth of the expected content of CG dinucleotides and is enriched in TG and CA.[15,140]

This deamination reaction will thus have a greater effect the higher the initial concentration of m^5C in DNA. Vertebrate DNA has the greatest deficiency in CG dinucleotides and the highest proportion of methylcytosine while the reverse is true of arthopod DNA[16] (see Section IV.A). If mCG dinucleotides occur in noncoding regions of DNA, their mutation might be expected to have no deleterious effect on the cell. This may be true to the extent that late replicating DNA present in heterochromatin is found to be undermethylated (see Section V.B) but the finding of high levels of m^5C in satellite DNA points to an essential role of this base in the function of such DNA.

D. Maintenance and Alteration of Methylation Patterns

Methylcytosines are largely found in the symmetrical dinucleotide m^5CG. Thus immediately following DNA replication an unmethylated —CG in the daughter strand is paired with a m^5CG in the parental strands (Figure 7). It is believed that such hemimethylated DNA is the natural substrate for DNA methylases and indeed such a substrate is a very suitable acceptor of methyl groups in vitro.[26,141,1650,197] The efficient action of DNA methylase in vivo would therefore be expected to maintain the parental distribution of m^5C in the daughter DNA molecules and this appears to be the case.[149,162,164]

In order to change the pattern of methylation, methyl groups must either be added to totally unmethylated DNA or removed from DNA. The former reaction can occur, in

vitro and must complement any removal of methyl groups if they are not to be lost altogether with time.

No demethylating enzyme activity has been found but DNA deficient in methyl groups would be produced by replication in the absence of methylation (Figure 7). Two rounds of replication are required to completely remove a pair of methyl groups and the DNA present after only one round of replication is hemimethylated and hence particularly susceptible to methylation.

Early attempts to alter experimentally the level of methylation of DNA in a cell have been largely unsuccessful as the agents employed have not been selective. Thus treatment with ethionine or cycloleucine reduces the level of methylation of DNA, RNA, and protein as does withholding of methionine from cultured cells.[141,161,163] Such procedures quickly lead to dramatic effects on cell growth and metabolism.

5-Azacytidine and more particularly 5-azadeoxycytidine are incorporated into DNA in place of deoxycytidine and this analogue cannot be methylated.[161] DNA containing azacytosine has a disproprotionately low level of methylcytosine suggesting that the analogue in addition to its not being a substrate for DNA methylase may actively block enzyme action.[161,198] This has also been shown to be true in vitro.

The effects of incorporation of azacytosine into DNA are drastic and cell growth almost ceases after a few days. However cells arise following treatment which have a much lower level of methylcytosine in their DNA. In some cases these "mutants" can express differentiated functions absent from the initial cells (see Section VI.6). Indeed the induction of metalothionein in thymoma cells occurs following only one round of DNA synthesis in the presence of the drug, implying that the substitution in only one strand of duplex DNA is sufficient to alter its expression. Such findings indicate that azacytosine not only interferes with DNA methylase action but probably exerts its effect on gene expression by directly blocking the interaction of other (controlling) proteins with DNA.[198]

E. DNA Methylases

Enzymes which catalyze the transfer of a methyl group from Ado Met to cytosines on DNA have been isolated and partially purified from a number of vertebrate cells and also from *Chlamydomonas* (see Reference 26 for a recent review). The enzymes are largely nuclear but can readily be extracted with dilute salt solutions. We have evidence, however, of some activity which remains firmly bound in nuclei following extraction with 2 *M* NaCl and this activity may be bound in a complex near the replication fork.[166]

The DNA methylases which have been partially purified from eukaryotes are high molecular weight enzymes and they have no cofactor requirements. They can add methyl groups to unmethylated duplex DNA but the best substrate in hemimethylated DNA. Thus they are capable of maintaining or changing the pattern of methylation of their substrate DNA.

The sole product of vertebrate DNA methylase action is m^5C and on isostic analysis labeled m^5C is found in pyrimidine runs in proportions consistent with a location in Nm^5CG.[153] However, there is evidence from dinucleotide analysis that a considerable amount of m^5CA is also produced.[34]

One of the enzymes from *Chlamydomomas* apparently resembles the vertebrate enzyme in methylating cytosines in CG dinucleotides but another enzyme methylates cytosines in the sequences TCA, TCG, and TCC,[95,126,129] and in higher plants a significant amount of methylcytosine occurs in partly symmetrical sequence mCNG.[155]

F. Methylation Following Repair of DNA

Damage to DNA occurs when cells are synchronized by treatment with excess thymidine and the level of methylation in the repaired regions is greater than normal.[167] This is

unlikely to be a result of the selective repair of highly methylated regions as Drahovsky et al have shown that methylation occurs in both unique and repetitive DNA sequences of human lymphocytes when these are repairing damage caused by nitrogen mustard.[168] However, following repair the repetitive DNA fraction was no longer more highly methylated than unique DNA and this may be the result of overmethylation of the unique DNA.

VI. FUNCTION OF DNA METHYLATION

A. Gene Expression

1. Introduction

For many years it has been proclaimed that all the cells of an organism throughout its life cycle have an identical complement of DNA. A few variants were acknowledged in the unicellular eukaryotes, and dipteran salivary glands were found to have polytene chromosomes where most of the genetic information is multiply represented. In some other special instances specific DNA sequences are amplified or lost but the fact that nuclei from differentiated cells can be introduced into enucleated zygotes and result in normal embryogenisis[169] supports our long-held beliefs, as do the studies on limb regeneration.[170] Thus if a change in the base sequence of DNA occurs on differentiation it must be a reversible change. This conclusion is supported by the fact that gametes are also differentiated cells whose DNA must be totipotent.

For methylation to have effect on the transcription of DNA the pattern of methylation must differ from tissue to tissue in an organism. Section IV.A has already reviewed the gross differences in the levels of methylation of DNA from different tissues and below we shall review the variations found for particular genes.

In addition, it should be mentioned that there is a correlation between undermethylation and heterochromatin[140] and that tumor cells often show a level of DNA methylation different from the corresponding normal cell.[81,82,171]

That variations are found shows that the maintenance of a methylation pattern cannot be the only action of a DNA methylase. This must be true even if the differentiation event itself is a demethylation for the original pattern must be reestablished in the zygote. No demethylase has been described and it is assumed that methyl groups can only be lost from DNA following replication in the absence of methylation (see Figure 7).

In order to examine the methylation of a specific gene in different tissues it is essential to have a probe for that gene. Such a probe is a piece of DNA carrying the gene in question. Usually the gene has been cloned in *E. coli.* The cloned gene is then made radioactive by nick translation using ^{32}P deoxyribonucleotides.

The DNA from the tissue under test is treated with a restriction enzyme whose action is sensitive to methylation of its recognition sites. Usually the enzyme employed is HpaII but MspI, HhaI, and several others can be used.

The restricted DNA is fractionated by electrophoresis on an agarose gel and Southern blotted onto nitrocellulose or DBN paper. A stained photograph of the gel simply shows a smear of digested DNA but if the nitrocellulose paper is now hybridized with the radioactive probe then, on autoradiography, bands show up corresponding to the various restriction fragments of the gene (see Figure 2). A comparison with the expected pattern indicates which sites are methylated.

Results of a more general nature have been reported by Naveh-Many and Cedar[118] who used DNase I treatment of chick erythrocyte chromatin to introduce nicks into transcriptionally active genes. Following nick translation with radioactive deoxyribonucleoside triphosphates and cleavage of the DNA with HpaII or MspI they concluded that only 43% of CCGG sites were methylated in active regions compared with 70% in total DNA. When the only triphosphate present was $\alpha^{32}P$ dGTP a modified nearest

neighbor analysis showed that 30% of all CG dinucleotides were methylated in active regions compared with 66% in total DNA. Furthermore, if mouse liver chromatin nick translated with α^{32}PdGTP is hybridized to poly A containing mRNA isolated from mouse liver, then the percentage of CG dinucleotides methylated in the hybrid is only 25% as compared with 65% in the unhybridized DNA. All these results confirm the positive correlation between transcriptional activity and undermethylation (but see Section IV.C).

2. Globin Genes

Using the isoschizomeric pair of restriction enzymes, HapII and MspI, Waalwijk and Flavell[172] showed that a particular site in the large intron of the rabbit β globin gene was methylated to different extents in different tissues. Sperm (100%) and brain (80%) DNA were most heavily methylated while most somatic (erythroid and nonerythroid) tissue DNA was about 50% methylated.

Shen and Maniatis[173] considered 13 CCGG sites in a region of the rabbit genome covering 4 β-type globin genes. Just three of these sites (two between and one within a coding region) showed changes in the extent of methylation which could be inversely correlated with transcription. Again brain DNA was found to be heavily methylated but even here some sites are only 50% methylated (i.e., methylated in half the brain cells).

The initial result of Waalwijk and Flavell was followed by a more extensive study covering 17 sites in the human $\gamma\ \delta\ \beta$ globin region.[174] Although sperm DNA is always fully methylated there is no obvious correlation between methylation and gene expression. Thus other cells (placenta, HeLa) which do not express globin genes show little methylation. The region around the γ globin genes is only slightly methylated in fetal liver but the rest of the region is heavily methylated. However bone marrow cells show partial undermethylation at 8 of 16 sites scattered over the globin region. This study also indicated that some CCGG sites were resistant to MspI showing that the 5′ cytosine can be methylated at least occasionally in the dinucleotide CC.

With the chicken globin gene McGhee and Ginder[175] showed that many CCGG sites are only partly methylated. In particular, sites near the ends of the gene sequence were completely unmethylated in cells which are expressing or have expressed the adult β globin gene. Marcaud et al. have shown a correlation between methylation of the α^{D}, β, and β-like genes in chicken erythroblasts and transformation of the cells with avian erythroblastosic virus which blocks transcription of the β-globin gene.[176]

Weintraub's group has shown a correlation between the DNase I sensitivity, the presence of HMG proteins 14 and 17 and acetylated histones, a deficiency of methyl cytosine at certain sites, and the expression of globin genes during the development of the chick.[177,178] They considered four stages of development. In 20- to 22-hr embryos the area opaca has 250,000 cells which develop into erythroid cells two generations later but which at 20 hr do not make globin mRNA and whose globin gene chromatin shows a low sensitivity to DNaseI. The two α globin genes and the fetal and adult β globin genes are fully methylated at this stage. At 5 days the α globin and the embryonic β globin genes are active (together with the α-type U globin gene) and their chromatin shows regions hypersensitive to DNase I. From this stage on the α globin genes are not methylated at several CCGG sites and the embryonic and adult β globin genes are also undermethylated. By 14 days when adult β globin synthesis predominates the embryonic β globin and U globin genes are less sensitive to DNase I and more heavily methylated than in 5-day cells.

The chick α-globin region has been carefully mapped by Weintraub et al.[178] who have shown that in red blood cells from 14-day embryos and adults that a region stretching from about 0.5 kb 5′ to the α^{D} globin gene to 1.5 kb 3′ to the α^{A} globin gene is highly sensitive to DNase I and devoid of methylcytosine in —CCGG— sequences (see Fig-

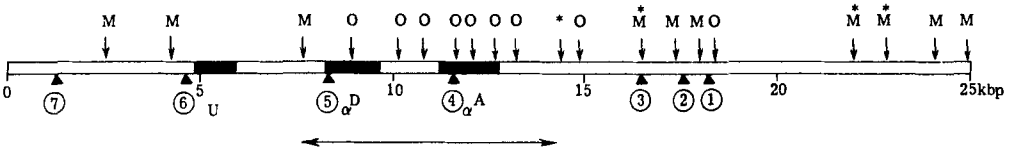


FIGURE 8. The location of CCGG regions and their methylation in the chick α-globin gene region. The position of the genes is marked with the thick bar and CCGG sites with arrows (↓). Sites not methylated in 14-day red blood cell DNA have an O while those methylated at the internal cytosine have an M. Those with an asterisk are probably methylated at the 5′ cytosine. The region marked with the double-headed arrow is highly sensitive to DNase I in 14-day chromatin. The solid triangles indicate DNase I hypersensitive regions — sites 1 to 5 are present in 5- and 14-day red blood cells chromatin but not in brain chromatin; site 6 is only sensitive at 5 days when the CCGG site between U and α^D is unmethylated; site 7 is also present in brain chromatin. (From Weintraub, H., Larsen, A., and Groudine, M., *Cell*, 24, 333, 1981. With permission.)

ure 8). This region is believed to be transcribed over its entire length and the chromatin has two DNase I-hypersensitive regions near the 5′ end of the regions coding for α^D and α^A globin. Outside this region most —CCGG— sites are methylated (some are methylated twice, i.e., mCmCGG). However, about 6 kb 3′ to the α^A gene is an unmethylated —CCGG— and three DNase I-hypersensitive sites whose relationship to the α-globin genes is unclear.

In contrast to this reduced methylation in transcribing cells, results reported by Anderson and Diacumakos[179] imply a specific modification occurs in erythroid precursor cells which is necessary to enable subsequent globin gene expression. This modification is apparently lost when the gene is cloned in *E. coli* as no globin synthesis occurs on transfection of the gene into mammalian cells.

3. Ovalbumin Gene

In a manner similar to the globin gene the chick ovalbumin gene region has several —CCGG— sites some of which are methylated in all tissues studied (m^+), some never methylated, (m^-) and some show variable degrees of methylation (m^v).[38,180] in the oviduct the variable sites have a low level of methylation which correlates with the expression of the ovalbumin gene. However, as the oviduct is a tissue made up of several cell types Kuo et al.[180] treated oviduct chromatin with DNase I to degrade the actively transcribed genes and concluded the sites which show low methylcytosine levels in the whole organ are totally unmethylated in the active duct cells. The correlation does not work so well in the reverse manner, however, as the same sites are only partially methylated in erythrocyte nuclei which do not express ovalbumin or conalbumin.

4. Viral Genes

The DNA of adenovirus types 2 or 12 is not methylated in virions or when free or integrated in productively infected cells.[181,182] In contrast the majority of CCGG sequences in this DNA are methylated in transformed cells. In such cells the sequences which are not fully methylated are those which are transcribed. In an adenovirus 12 induced rat brain tumor a late gene coding for a 72,000 mol wt DNA binding protein is expressed and is undermethylated at —CCGG— sites in contrast to its complete methylation in cells where it is not expressed.

A similar story has been shown for *Herpesvirus saimiri* DNA which is not methylated in virions or in lymphoid cells which yield virus but in nonproducing lymphoid cell lines more than 80% of the HpaII sites together with SmaI (CCCGGG) and SacII (CCGCGG) are methylated.[183]

Mouse mammary tumor virus is a retrovirus which can be transferred genetically

following integration of proviral DNA into the germ line when the —CCGG— sites are fully methylated and the provirus is not transcribed.[184] An alternative method of transfer is via the milk when the provirus can be present in normal or in transformed mammary cells. In either case it is partially undermethylated. In some cases both fully methylated (genetically acquired) and partially methylated (milk borne) provirus can be found in the same tissue, suggesting that the milk borne DNA was not exposed to a methylating event which rendered it transcriptionally inert in the germ cells or early embryo.

5. Other Genes

One HpaII site on a rat hepatoma protein (hp22) gene is methylated in the Morris hepatoma where it is not expressed in contrast to in normal liver where it is only partially methylated and expressed.[185]

In another hepatoma cell line the ribosomal RNA genes are amplified about tenfold. As is the case with amplified rRNA genes in certain human families[186] the amplified hepatoma genes are mostly transcriptionally inactive and are highly methylated.[187] This contrasts to the situation in *Xenopus* oocytes where the amplified ribosomal RNA genes are expressed and are not methylated.[188] Here the chromosomal ribosomal RNA genes are methylated at HpaII and HhaI (GCGC) sites except for two sites which are only partially methylated.[189] Yet these genes are expressed at least in somatic cells and in the sea urchin the ribosomal RNA genes together with the tRNA and histone genes are found in the unmethylated compartment irrespective of whether they are active.[91] Other genes which show site-specific hypomethylation on transcription are those for lactalbumin[190] and δ crystallin.[191]

6. The Effect of 5-azacytosine

The nucleoside analogues 5-azacytosine and 5-azadeoxycytidine are incorporated into DNA but cannot be methylated. Their incorporation even in trace amounts drastically reduces the level of DNA methylation probably by inhibiting DNA methylase activity.[197,198] (A similar effect may be caused by the alkylating agent acetylaminofluorene.[199]) Cell growth is severely affected following incorporation of the analogue but from such cultures we have selected variants of mouse L929 cells with only one third of the amount of methylcytosine present in the parental cell DNA.[198]

Incorporation of azacytosine brings about changes in differentiation of cells but it is not clear whether this is as a result of inhibition of methylation. Thus mouse embryo cells can be induced to differentiate into muscle cells[200] and 3T3 cells can be induced to differentiate into muscle cells, chondrocytes, and adipocytes.[201] The differentiation is observed in about 1% of cells several days after exposure to the analogue whose concentration needs to be carefully controlled to prevent extreme toxicity. Similarly, treatment of chick cells carrying an inactive retrovirus provirus (ev-1) with 5-azacytidine leads to activation of ev-1 accompanied by hypomethylation and the acquisition of a nuclease-hypersensitive region within the ev-1 chromatin region.[202]

The above events are observed several generations after exposure to the drug which is removed after a 24-hr treatment. However Compere and Palmiter[165] observed that a mouse thymoma cell on treatment with 5-azacytidine for as little as 4 hr became sensitive to induction of the metallothionein gene by cadmium or glucocorticoids, i.e., only one round of DNA synthesis is required to produce the change which is thus measured in cells whose DNA contains one normal parental strand and one strand deficient in methylcytosine and containing azacytosine. Again high levels of toxicity are reported but populations of cells which survive retain about one fifth of their maximal ability to induce metallothionein.

Although these experiments do show a correlation between undermethylation and differentiative changes, the high levels of toxicity and the presence of an unusual base analogue in the DNA must also be considered as possible inducers of the differentiation. The analogue obviously affects the process of methylation directly (not only as a consequence of the inability of azacytosine to be methylated) and may equally affect proteins other than DNA methylase which interact with the DNA.

Furthermore, it was shown in 1968 and 1969 that while azadeoxycytidine is an effective antileukemic agent, azacytidine is ineffective in this respect but does inhibit the *de novo* synthesis of pyrimidines and, following its incorporation into RNA, blocks enzyme induction and protein synthesis.[203]

Ethionine — an inhibitor of methylation in general — is one of many inducers of globin synthesis in Friend cells, though again it is difficult to prove the effect is mediated via an inhibition of methylation of the globin gene region.[78] A recent paper[204] has shown that in cells treated with 5-azacytidine a previously inactive heterochromatic human X chromosome can be activated at least in part (see Section VI.B).

7. The Effect of Methylation on Cloned DNA Introduced into Animal Cells

Most experiments have been done with a thymidine kinase (tk) gene introduced into a tk^- cell line, which is then selected in HAT medium for expression of the tk gene. When the source of the gene is DNA from a tk^+ cell then the gene expression in the recipient appears to be under normal cell cycle control, but the herpes virus tk gene is constitutively expressed.[205-207] The herpesvirus tk gene (and plasmid and cloned growth hormone genes introduced at the same time) remains unmodified at CCGG sites in 90% of the clones investigated[207] and this remains true in most tk^- revertants (in these experiments in an exceptional tk^- variant there was extensive methylation of the tk gene which on reverting again to tk^+ became unmodified.[208]

Wigler et al. have shown that when a cloned chicken tk gene is methylated in vitro using HpaII methylase and then introduced into cells the methylation pattern is maintained with very high (but not 100%) fidelity at all but one site.[149] However Pollack et al.[207] showed that fewer than 10% of similar clones maintained the methylcytosines in —CCGG— sites in herpesvirus tk gene and none gained methylcytosines in —GCGC— or TCGA sites. In these experiments there was a positive selection for cells expressing the tk gene and this may have selected for those cells which had lost their modification at particular —CCGG— sites.

The question as to whether cells are able to modify the pattern of methylation of their DNA or only maintain existing patterns at first sight appears unresolved in light of such contradicting evidence. It is clear, however, that methyl groups can be lost from DNA (presumably by replication in the absence of methylation) and can be added to completely unmethylated DNA (e.g., the unexpressed viral DNA in transformed cells). This, together with all the data reviewed in this section, indicates that modulation of methylation patterns occurs and must be taken into account when formulating possible functions of DNA methylation.

When cloned globin genes are introduced into mouse tk^- cells along with the herpesvirus tk gene the globin genes are not expressed despite the fact that they are not methylated,[179,218] i.e., undermethylation is not, by itself, able to ensure transcription. Enhancing sequences which can be remote from the globin genes may normally control transcription. Such enhancing sequences may be lost on gene cloning but can be substituted for by a 72 bp repeated sequence element from the SV40 genome.[219-221]

8. Conclusion

It is clear that particular genes show different patterns of methylation in different

tissues and that azacytosine incorporation into DNA inhibits methylation and causes changes in gene expression. However, undermethylation is not a sufficient stimulus for transcription and it is not possible, therefore, to conclude from this circumstantial evidence that methylcytosine affects gene expression. It is at least as probable that changes in gene expression bring about changes in chromatin structure and thereby affect methylation (see Section IV.C).

It is also clear that where a correlation has been shown between undermethylation and transcriptional activity the correlation usually exists for only some of a subset of methylatable sites, e.g., the m^v CCGG sites of the ovalbumin gene region. It is possible that m^v sites affect, or are affected by, gene expression but that the m^+ and m^- sites have some other function which necessitates a permanent level of methylation (see below). We feel that the most likely explanation is that m^+ and m^- sites are permanently located in a chromosomal structure which either allows (m^+) or prevents (m^-) DNA methylation. This could arise as a result of nucleosomes being phased over certain regions of the gene. m^v sites would then reflect the different nucleosome phasing found in active and inactive chromatin or the presence of other DNA binding proteins which interfere with DNA methylase action in transcribed regions.

Nucleosome phasing itself may be controlled by DNA methylation which could then indirectly affect transcription (see Section IV.C). Unfortunately no compelling evidence exists at present to allow a definite conclusion to be drawn.

B. Chromosome Inactivation

Sager and Kitchen[209] have proposed that methylation of DNA may be linked with restriction (in its widest sense) to explain the elimination or inactivation of chromosomes or parts of chromosomes. Theirs is a development of the idea put forward by Riggs[210] to explain the inactivation of one X chromosome in female mammals. However Sager and Kitchen draw examples in addition from several insect species and also from a study of chloroplast inheritance and chromosome elimination in cell hybrids.

Their proposal is that in the sperm of some species some or all of the DNA fails to become methylated and so becomes sensitive to inactivation in the developing egg. The inactivation initially takes the form of condensation to form heterochromatic regions and in some cases this is followed by elimination of the DNA from the cell presumably following digestion with nucleases. Two pieces of evidence may bear indirectly on this idea.

Mohandas et al.[204] have obtained a mouse human hybrid cell where the only copy of the hypoxanthine phosphoribosyltransferase (HPRT) gene is on an inactive human X chromosome. Such cells fail to grow in HAT medium (medium containing hypoxanthine, aminopterin, and thymidine) and the rate of appearance of variants expressing the human HPRT gene is very low. However treatment of these cells with 5-azacytidine, a reagent which among other actions interferes with DNA methylation (see Section VI.A.6.), increases the rate of appearance of variants 1000-fold.

Satellite (and possibly other) DNA which is normally heavily methylated is partially or totally unmethylated in the sperm of several mammals.[88,89] This may reflect a lack of DNA methylase during synthesis of this DNA. As the rest of the sperm DNA is heavily methylated a more likely explanation is that the chromosomal location of the DNA precludes methylation. It is possible that any DNA in a similar location will fail to be methylated in sperm but will subsequently become subject to inactivation by heterochromatization following or followed by methylation in the zygote. Both satellite DNA and the extra X chromosome may have a function in the sexual process which would be inconvenient if expressed in somatic cells. Thus it is believed that satellite DNA may be involved in recombination events which are very much more frequent in gametogenesis.

Satellite DNA is also found associated with amplified genes which probably arise by recombination.[211] Amplification is an uncommon occurrence in somatic cells although it is a normal occurrence in certain oocytes.

Whether methylation would interfere with some of the functions of satellite DNA or whether developmentally regulated methylation provides a signal which leads to the binding of proteins prior to inactivation of such DNA in somatic cells is unclear. (Alternatively methylation may follow heterochromatization.) However, there may be a link with the proposed function of methylation in gene expression, i.e., methylation leads to chromosome condensation, inactivation, and possible elimination.

This is also supported by the finding of increased levels of m^5C in the polytene chromosomes of dipteran salivary glands which are largely transcriptionally inert[212] and similarly inert amplified ribosomal RNA genes of vertebrates.[186,187,213] Inactivation of DNA may involve a methylcytosine-induced condensation of chromatin but the action may be more direct. Thus the presence of m^5C in DNA increases both the melting temperature and the likelihood of formation of Z DNA (see Section I.B). However, as methylcytosines are seldom clustered — indeed they are in general very widely spaced — such effects on physical structure are unlikely to be important. Of more significance is the alteration of base sequence with its effect on DNA-protein interactions as exemplified by the failure of the lac repressor to bind to its operator when this contains a cytosine rather than methylcytosine or thymine.[159]

In the chloroplasts of *Chlamydomonas* the function of DNA methylation is much clearer.[126-129] As in insects the level of DNA methylation in chloroplast DNA is very low or zero in vegetative cells and male (mt^-) gametes. However in the gametic mt^+ (female) cells chloroplast DNA is heavily methylated. This methylation is brought about by an enzyme present only in female gametes and in zygotes and protects chloroplast DNA of maternal origin from digestion by a restriction enzyme. Chloroplast DNA of paternal origin is unmethylated and is lost from the zygote. The *mat-1* mutation carried by males leads to methylation of chloroplast DNA in mt^- gametes and to biparental inheritance and has been associated with the assembly of a high molecular weight form of DNA methylase in male gametes.

Thus a role for methylation in the control of maternal inheritance has been established. As in bacteria the methyl group presumably prevents the action of a corresponding endonuclease. As the mating process in *mat-1* mutants is not affected, however, the reason for maternal inheritance is unclear.

C. Aging

Scarano[214,215] originally proposed that m^5C may be deaminated to form thymine and that if this were to happen in a programmed manner throughout development it would lead to permanent differentiation. Thus the base sequence could be changed in a two-step reaction:

1. Methylation of cytosine to m^5C
2. Deamination of m^5C to thymine

Although he claimed to detect the formation of thymine by this route it is very difficult to rule out nonenzymic deamination. Such deamination has been shown to occur during base analysis in vitro,[14,15] and also in vivo in bacteria where mutational hotspots can be related to the presence of methylcytosine.[13] In these instances the deamination is a random process rather than a step in a programmed differentiation.

Evidence for similar mutational events occurring in mammalian tissues may include the differences in the levels of m^5C in the DNA of cows of different ages.[74,75] The random accumulation of thymine arising from m^5C may lead to cell death by a so-called error

catastrophe. Differences in the amount of m^5C in the DNA of related tumors may also arise by steady accumulation of deaminated products and indeed the cause of the tumor may be a base substitution (m^5C to T). However a more probable explanation of enhanced or reduced levels of m^5C may be found in the altered karyotype of the tumor or aging cell.

Holliday and Pugh[216] envisaged a mechanism whereby progressive methylation of a repetitive DNA sequence would act as a biological clock counting cell divisions. Starting from a methylcytosine at one end of a repeated sequence a "clock" methylase may modify a complementary pair of cytosines in the next frame of the repeat at each round of DNA synthesis. Depending on the number of repeats the last one (and the adjacent control region) will be methylated after a fixed number of divisions. Methylation of the control region would lead to a change in gene expression a fixed number of generations after the initial methyl group was added. This change may signal the differentiation for example of the humerus of the chick limb or alternatively it may signal the death of the cell. However, no evidence has been produced in support of such a mechanism.

D. Mismatch Repair

When an incorrect deoxyribonucleoside triphosphate is incorporated into the growing DNA chain by DNA polymerase, proofreading mechanisms lead to its removal and the substitution of the correct nucleotide. In prokaryotes the DNA polymerases have associated 3′ ⟶ 5′ exonuclease activity believed to be primarily involved in proofreading though other mechanisms have also been considered. Eukaryote DNA polymerases lack the 3′ ⟶ 5′ exonuclease activity but replication shows a similar fidelity in both eukaryotes and prokaryotes.

Should an incorrect (but normal) base be incorporated and escape the proofreading enzymes acting at the point of replication, a mismatch is introduced into the DNA. The mismatch is easily recognized but because both bases are normal, there is apparently no way of distinguishing which is the incorrect base.

The incorrect base is the one in the daughter strand. As this strand does not become methylated until a finite time after replication a correction mechanism may exist which recognizes the unmethylated daughter strand and replaces the incorrect base. Some evidence exists for the occurrence of such a mechanism in bacteria and Radman's group are looking for parallel evidence in eukaryotes.[61-63] It must be borne in mind however that such a function must be absent in those species or in those DNA regions lacking methylcytosine.

VII. CONCLUSIONS

Throughout vertebrate evolution it appears that 80% of mCG dinucleotides have been converted to TG and CA by deamination of m^5C. Thus much of the m^5C originally present must have become dispensable. It may have been in redundant sequences, spacer DNA, or introns which are often AT rich.[72,96] Moreover arthropods appear to manage with very low levels of m^5C as do animal viruses. Such observations suggest some functions for m^5C which are not always required or which can be taken over by some other mechanism when m^5C is absent or at low levels.

If a function of m^5C is to switch off gene transcription, then it could be argued that in certain organisms all, or part, of the genome must be permanently active by virtue of a low level of m^5C. This may apply to some animal viruses but is unlikely to be true of metazoa. Transcription control may be exerted through proteins which interact with specific base sequences and modification of such sequences (which may be widely spaced) may provide a level of control which is not universal. Thus it may be simply that only a very small number of m^5C residues are actually involved in transcriptional regulation.

However, in mammalian DNA there is, on average, 1 methylcytosine every 100 bases and so a more plausible alternative is that lack of methylation of particular regions of the genome leads to an overall alteration of chromatin structure. This change may be induced by a sequence-recognizing protein and may be a preliminary to transcription. Although the correlation exists between undermethylation and transcription there is no evidence which supports this order of events.

We believe it is more likely that alterations in chromatin structure which occur in the region of transcribed genes may simply interfere with the action of DNA methylase leading to localized undermethylation of DNA. This undermethylation may occur to different extents along the DNA depending on the relative location of CG dinucleotides and protected regions. It is clear that in isolated nuclei methyl groups are added preferentially to DNA which is susceptible to micrococcal nuclease. This would imply that nucleosomes and other DNA binding proteins can interfere with DNA methylation. On transcribed regions a control protein may dictate the location of nucleosomes over certain regions of DNA ("phasing") thus exposing some regions to methylase action and blocking others. This would lead to a particular pattern of undermethylation of transcribed DNA typical of that found in transcribed regions. On transcriptionally inactive DNA the lack of control proteins bound to particular sequences may allow more freedom of movement to the nucleosomes and thus all regions of DNA would at some time or other be accessible to DNA methylase.

Leaving aside the possible effect of DNA methylation on transcription, the presence of particular unmethylated sequences in DNA may make that DNA susceptible to interaction with proteins involved in recombination or repair. The finding that methylated cloned DNA is not expressed in recipient cells as often as unmethylated DNA[207] may have nothing to do with its transcription but with its integration into the host genome. Recombination events occur with high frequency during meiosis and are believed to involve regions of tandemly repeating DNA (e.g., satellite DNA).[161,217] Following meiosis such DNA is found in heterochromatic regions where it is relatively inactive in recombination. Satellite DNA is somatic cells (whether AT or GC rich) has a high concentration of m^5C compared with unique DNA and also compared with satellite DNA in sperm and methylation may represent the inactivation of a recognition signal no longer desirable in somatic cells. Thus the bulk of m^5C may have a role in protecting desirable gene sequence arrangements by specifically inhibiting deleterious recombination events or by facilitating chromatin condensation which would minimize recombination. An investigation of the effect of DNA methylation on gene rearrangements involved in lymphocyte differentiation and gene amplification may be very rewarding.

ACKNOWLEDGMENTS

The authors would like to thank the University of Glasgow and in particular Professor R.M.S. Smellie for support over ther years and also the Medical Research Council and the Cancer Research Campaign for funds.

REFERENCES

1. **Vanyushin, B. F., Belozersky, A. N., Kokurina, N. A., and Kadirova, D. X.,** 5 Methylcytosine and 6 methylaminopurine in bacterial DNA, *Nature (London),* 218, 1066, 1968.
2. **Hattman, S., Kenny, C., Berger, L., and Pratt, K.,** Comparative study of DNA methylation in three unicellular eukaryotes, *J. Bacteriol.,* 135, 1156, 1978.
3. **Razin, A. and Razin, S.,** Methylated bases in mycoplasmal DNA, *Nucl. Acids Res.,* 8, 1383, 1980.
4. **Adams, R. L. P., McKay, E. L., Craig, L. M., and Burdon, R. H.,** Methylation of mosquito DNA, *Biochim. Biophys. Acta,* 563, 72, 1979.

5. **Chargaff, E., Lipshitz, R., and Green, C.,** Composition of the deoxypentose nucleic acids of four genera of sea urchin, *J. Biol. Chem.*, 195, 155, 1952.
6. **Vanyushin, B. F., Mazin, A. L., Vasilyev, V. K., and Belozosky, A. N.,** The content of 5 methylcytosine in animal DNA: the species and tissue specificity, *Biochim. Biophys. Acta*, 299, 397, 1973.
7. **Dawid, I. B.,** DNA in amphibian eggs, *J. Mol. Biol.*, 12, 581, 1965.
8. **Adams, R. L. P.,** unpublished.
9. **Thomas, A. J. and Sherratt, M. S. A.,** The isolation of nucleic acid fractions from plant leaves and their purine and pyrimidine composition, *Biochem. J.*, 62, 1, 1956.
10. **Stahl, S. J. and Chamberlin, M. J.,** Transcription of T7 DNA containing modified nucleotides by bacteriophage T7 specific RNA polymerase, *J. Biol. Chem.*, 453, 4951, 1978.
11. **Wang, R. Y-H., Gehrke, C. W., and Ehrlich, M.,** Comparison of bisulphate modification of 5 methyl deoxycytidine and deoxycytidine residues, *Nucl. Acids Res.*, 8, 4777, 1980.
12. **Burdon, R. H. and Adams, R. L. P.,** The *in vivo* methylation of DNA in mouse fibroblasts, *Biochim. Biophys. Acta*, 174, 322, 1969.
13. **Coulandre, C., Miller, J. H., Farabough, P. J., and Gilbert, W.,** Molecular basis of base substitution hotspots in *E. coli, Nature (London)*, 274, 775, 1978.
14. **Sneider, T. W.,** On the source of "minor thymine" in DNA from a Novikoff rat hepatoma cell line, *J. Mol. Biol.*, 79, 731, 1973.
15. **Ford, J. P., Coca-Prados, M., and Hsu, M-T.,** Enzymatic analysis of 5 methylcytosine content in eukaryotic DNA, *J. Biol. Chem.*, 255, 7544, 1980.
16. **Bird, A. P.,** DNA methylation and the frequency of CpG in animal DNA, *Nucl. Acids Res.*, 8, 1499, 1980.
17. **Gates, F. T. and Linn, S.,** Endonuclease V of *E. coli, J. Biol. Chem.*, 252, 1647, 1977.
18. **Behe, M. and Felsenfeld, G.,** Effects of methylation on a synthetic polynucleotide: the B-Z transition in poly(dG-m^{5}dC) poly(dG-m^{5}dC), *Proc. Natl. Acad. Sci. U.S.A.*, 78, 1619, 1981.
19. **Quadrifoglio, F., Manzini, G., Vasser, M., Dinkelspiel, K., and Crea, R.,** Conformational stability of alternating d(CG) oligomers in high salt solution, *Nucl. Acids Res.*, 9, 2195, 1981.
20. **Ehrlich, M., Ehrlich, K., and Mayo, J. A.,** Unusual properties of the DNA from *Xanthomonas* phage XP-12 in which 5 methylcytosine completely replaces cytosine, *Biochim. Biophys. Acta*, 395, 109, 1975.
21. **Szer, W. and Shugar, D.,** The structure of poly-5-methylcytidylic acid and its twin-stranded complex with poly inosinic acid, *J. Mol. Biol.*, 17, 174, 1966.
22. **Gill, J. E., Mazrimas, J. A., and Bishop, C. C.,** Physical studies on synthetic DNAs containing 5 methylcytosine, *Biochim. Biophys. Acta*, 335, 330, 1974.
23. **Wagner, I. and Capesius, I.,** Determination of 5 methylcytosine from plant DNA by high performance liquid chromatography *Biochim. Biophys. Acta*, 654, 52, 1981.
24. **Engel, J. D. and von Hippel, P. H.,** Effects of methylation on the stability of nucleic acid conformation, *J. Biol. Chem.*, 253, 927, 1978.
25. **Cedar H., Solage, A., Glaser, G., and Razin, A.,** Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI, *Nucl, Acids Res.*, 6, 2125, 1979.
26. **Adams, R. L. P. and Burdon, R. H.,** DNA methylases in *Enzymes of Nucleic Acid Synthesis and Processing*, Vol. 2, Hnilica, L., Ed., CRC Press, Boca Raton, Fla., 1982.
27. **Kappler, J.,** The kinetics of DNA methylation in cultures of a mouse adrenal cell line, *J. Cell Physiol.*, 75, 21, 1970.
28. **Oldham, K. G.,** Spurious isotope effects in enzymatic reactions with tritium labelled substrates, *J. Labelled Compd.*, IV, 127, 1968.
29. **Adams, R. L. P.,** Incorporation of [5−^{3}H] uridine into DNA, *FEBS Lett.*, 2, 91, 1968.
30. **Razin, A. and Cedar, H.,** Distribution of 5 methylcytosine in chromatin, *Proc. Natl. Acad. Sci. U.S.A*, 74, 2725, 1977.
31. **Lubit, B. W., Pham, T. D., Miller, O. J., and Erlanger, B. F.,** Localisation of 5 methylcytosine in human metaphase chromosomes by immunoelectron microscopy, *Cell*, 9, 503, 1976.
32. **Schreck, R. R., Erlanger, B. F., and Miller, O. J.,** Binding of antinucleoside antibodies reveals different classes of DNA in the chromosomes of the kangaroo rat *(Dipodomys ordii), Exp. Cell Res.*, 108, 403, 1977.
33. **Sharp, P. A., Sugden, B., and Sambrook, J.,** Detection of the restriction endonuclease activities in *H. parainfluenzae* using analytical agarose ethidium electrophoresis, *Biochemistry*, 12, 3055, 1973.
34. **Waalwijk, C. and Flavell, R. A.,** MspI, an isoschizomer of HpaII which cleaves both unmethylated and methylated HpaII sites, *Nucl. Acids Res.*, 5, 3231, 1978.
35. **Sneider, T. W.,** The 5′ cytosine in CCGG is methylated in two eukaryotic DNAs and MspI is sensitive to methylation at this site, *Nucl. Acids Res.*, 8, 3829, 1980.
36. **Singer, J., Robert-Ems, J., and Riggs, A. D.,** Methylation of mouse liver DNA studied by means of the restriction enzymes MspI and HpaII, *Science*, 203, 1019, 1979.

37. **Sano, H., Royer, H-D., and Sager, R.,** Identification of 5 methylcytosine in DNA fragments immobilised on nitrocellulose paper, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 3581, 1980.
38. **Mandel, J. L. and Chambon, P.,** DNA methylation: organ specific variations in the methylation pattern within and around ovalbumin and other chicken genes, *Nucl. Acids Res.*, 7, 2081, 1979.
39. **Nordheim, A., Pardue, M. L., Lafer, E. M., Möller, A., Stollar, B. D., and Rich A.,** Antibodies to left-handed Z DNA bind to interband regions of *Drosophila* polytene chromosomes, *Nature (London)*, 294, 417, 1981.
40. **Lautenberger, J. A. and Linn, S.,** The DNA modification and restriction enzymes of *E. coli* B. I. Purification, subunit structure and catalytic properties of the modification methylase, *J. Biol. Chem.*, 247, 6167, 1972.
41. **Eskin, B. and Linn, S.,** The DNA modification and restriction enzymes of *E. coli* B. II. Purification, subunit structure and catalytic properties of the restriction endonuclease, *J. Biol. Chem.*, 247, 6183, 1972.
42. **Eskin, B. and Linn, S.,** The DNA modification and restriction enzymes of *E. coli* B. III. Studies of the restriction ATPase, *J. Biol. Chem.*, 247, 6192, 1972.
43. **Reiser, J. and Yuan, R.,** Purification and properties of the P15 specific restriction endonuclease from *E. coli*, *J. Biol. Chem.*, 252, 451, 1977.
44. **Roberts, R. J.,** Restriction endonucleases, *CRC Crit. Rev. Biochem.*, 4, 123, 1976.
45. **Bingham, A. H. A. and Atkinson, T.,** Restriction endonucleases and modification methylases in bacteria, *Trans. Biochem. Soc.*, 6, 315, 1978.
46. **Roberts, R. J.,** Restriction and modification enzymes and their recognition sequences, *Nucl. Acids Res.*, 8, r63, 1980.
47. **Dugaiczyk, A., Hedgepeth, J., Boyer, H. W., and Goodman, H. M.,** Physical identity of the SV40 sequence recognised by the EcoRI restriction endonuclease and modification methylase, *Biochemistry*, 13, 503, 1974.
48. **Mann, M. B. and Smith, H. D.,** Specificity of HpaII and HaeIII DNA methylases, *Nucl. Acids Res.*, 4, 4211, 1978.
49. **Ravetch, J. V., Horiuchi, K., and Zinder, N. D.,** Nucleotide sequence of the recognition site for the restriction modification enzyme of *E. coli* B, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 2266, 1978.
50. **Lautenberger, J. A., Kan, N. C., Lackey, D., Linn, S., Edgell, M. H., and Hutchison, C. A.,** Recognition site of the *E. coli* B. restriction enzyme on ØX SB1 and SV 40 DNAs: an interrupted sequence, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 2271, 1978.
51. **Kan, N. C., Lautenberger, J. A., Edgell, M. H., and Hutchison, C. A.,** The nucleotide sequence recognised by the *E. coli* K12 restriction and modification enzymes, *J. Mol. Biol.*, 130, 191, 1979.
52. **Newman, A. K., Rubin, R. A., Kim, S-H., and Modrich, P.,** DNA sequences of structural genes for EcoRI DNA restriction and modification enzymes, *J. Biol. Chem.*, 256, 2131, 1981.
53. **Greene, P. J., Gupta, M., Boyer, H. W., Brown, W. E., and Rosenberg, J. M.,** Sequence anaylsis of the DNA encoding the EcoRI endonuclease and methylase, *J. Biol. Chem.*, 256, 2143, 1981.
54. **Mamelak, L. and Boyer, H. W.,** Genetic control of the secondary modification of DNA in *E. coli*, *J. Bact.*, 104, 57, 1970.
55. **Marinus, M. G. and Morris, N. R.,** Isolation of DNA methylase mutants of *E. coli* K 12, *J. Bacteriol.*, 114, 1143, 1973.
56. **Marinus, M. G. and Morris, N. R.,** Biological function of 6 methyladenine residues in the DNA of *E. coli* K12, *J. Mol. Biol.*, 85, 309, 1974.
57. **Hattman, S., Brooks, J. E., and Masurekar, M.,** Sequence specificity of the P_1 modification methylase (M.*Eco*P1) and the DNA methylase (M.*Eco*.dam) controlled by the *E. coli* dam gene, *J. Mol. Biol.*, 126, 367, 1978.
58. **Razin, A., Urieli, S., Pollack, Y., Gruenbaum, Y., and Glaser, G.,** Studies on the biological role of DNA methylation. IV. Mode of methylation of DNA in *E. coli* cells, *Nucl. Acids Res.*, 8, 1783, 1980.
59. **Geier, G. E. and Modrich, P.,** Recognition sequence of the dam methylase of *E. coli* K12 and the mode of cleavage of DpnI endonuclease, *J. Biol. Chem.*, 254, 1408, 1979.
60. **Schlagman, S., Hattman, S., May, M. S., and Berger, L.,** *In vivo* methylation by *E. coli* K12 mec$^+$ DNA cytosine methylase protects against *in vitro* cleavage by the RII restriction endonuclease (R-EcoRII), *J. Bacteriol.*, 126, 990, 1976.
61. **Radman, M., Villani, G., Boiteux, S., Kinsella, A. R., Glickman, B. W., and Spadari, S.,** Replication fidelity: mechanisms of mutation avoidance and mutation fixation, *Cold Spring Harbor Symp. Quant. Biol.*, 43, 937, 1978.
62. **Glickman, B. W. and Radman, M.,** *E. coli* mutator mutants deficient in methylation — instructed DNA mismatch correction, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 1063, 1980.
63. **Radman, M.,** DNA repair processes and genetic stability, *Biochem. Soc. Trans.*, 9, 78p, 1981.
64. **Gomez-Eichelmann, M. C. and Lark, K. G.,** Endo R. DpnI restriction of *E. coli* DNA synthesised *in vitro*, *J. Mol. Biol.*, 117, 621, 1977.

65. **Billen, D.,** Methylation of the bacterial chromosome: an event at the replication point?, *J. Mol. Biol.*, 31, 477, 1968.
66. **Marinus, M. G.,** Adenine methylation of Okazaki fragments in *E. coli, J. Bacteriol.*, 128, 853, 1976.
67. **Lark, C.,** Studies on the *in vivo* methylation of DNA in *E. coli* 15T⁻, *J. Mol. Biol.*, 31, 389, 1968.
68. **Lark, C.,** Effect of methionine analogs, ethionine and norleucine, on DNA synthesis in *E. coli* 15T⁻, *J. Mol. Biol.*, 31, 401, 1968.
69. **Vanyushin, B. E., Mazin, A. L., Vasilyev, V. K., and Belozersky, A. N.,** The content of 5-methylcytosine in animal DNA: the species and tissue specificity, *Biochim. Biophys. Acta,* 299, 397, 1973.
70. **Vanyushin, B. F., Tkacheva, S. G., and Belozersky, A. N.,** Rare bases in animal DNA, *Nature (London),* 225, 948, 1970.
71. **Guseinov, V. A., Kiryanov, G. I., and Vanyushin, B. F.,** Intragenome distribution of 5 methylcytosine in DNA of healthy and wilt infected cotton plants, *Mol. Biol. Rep.*, 2, 59, 1975.
72. **Portmann, B., Schaffner, W., and Birnstiel, M.,** Partial denaturation mapping of cloned histone DNA from the sea urchin *Psammechinus miliaris, Nature (London),* 264, 31, 1976.
73. **Vanyushin, B. F., Nemirovsky, L. E., Klinmenko, V. V., Vasiliev, U. K., and Belozersky, A. N.,** The 5-methylcytosine in DNA in rats, *Gerontologia,* 19, 138, 1973.
74. **Romanov, C. A. and Vanyushin, B. F.,** Intragenome specificity of DNA methylation in animals, tissue specificity and variation in the methylation of repeated sequences due to aging, carcinogenisis and hormonal induction, *Mol. Biol.*, 14, 279, 1980.
75. **Romanov, C. A. and Vanyushin, B. F.,** Methylation of reiterated sequences in mammalian DNAs: effects of tissue type, age, malignancy and hormonal induction, *Biochim. Biophys. Acta,* 653, 204, 1981.
76. **Singer, J., Robert-Ems, J., Luthardt, F. W., and Riggs, A. D.,** Methylation of DNA in mouse early embryos, teratocarcinoma cells and adult tissues of mouse and rabbit, *Nucl. Acids Res.*, 7, 2369, 1979.
77. **Fabricant, J. D., Wagner, E. F., Auer, B., and Schweiger, M.,** 5 Methylcytosine content of DNA during differentation in mouse teratocarcinoma cells, *Exp. Cell Res.*, 124, 25, 1979.
78. **Christman, J. K., Price, P., Pedrinau, L., and Acs, G.,** Correlation between hypomethylation of DNA and expression of globin genes in Friend erythroleukaemia cells, *Eur. J. Biochem.*, 81, 53, 1977.
79. **Baur, R., Wohlert, H., Kluge, N., and Kroger, H.,** 5-mcyt of DNA in hormone treated rat liver in DMSO induced erythroleukaemia cells and in embryos of *Psammechinus miliaris,* 12th FEBS Meet. Dresden, Abstr. 642, 1978.
80. **Kappler, J. W.,** The 5 methylcytosine content of DNA: tissue specificity, *J. Cell Physiol.*, 78, 33, 1971.
81. **Rubery, E. D. and Newton, A. A.,** DNA methylation in normal and tumour virus transformed cells in tissue culture, *Biochim. Biophys. Acta,* 324, 24, 1973.
82. **Cato, A. C. B., Adams, R. L. P., and Burdon, R. H.,** Genome modification in two lines of baby hamster kidney fibroblasts (BHK-21), *Biochim. Biophys. Acta,* 521, 397, 1978.
83. **Baur R., Wohlert, H., and Kroger, H.,** Methylation of DNA in developing embryos of the sea urchin *Psammechinus miliaris, Ges. Biol. Chem.*, 359, 254, 1978.
84. **Baur, R. and Kroger, H.,** Methylation in eukaryotes, 10th Int. Congr. Biochemistry, Hamburg, Abstr. 01-1-085, 1976.
85. **Pollock, J. M., Swihart, M., and Taylor, J. H.,** Methylation of DNA in early development: 5 methylcytosine content of DNA in sea urchin sperm and embryos, *Nucl. Acids Res.*, 5, 4855, 1978.
86. **Manes, C. and Menzel, P.,** Demethylation of CpG sites in DNA of early rabbit trophoblast, *Nature (London),* 293, 589, 1981.
87. **Turkington, R. W. and Spielvogel, R. L.,** Methylation of DNA during hormonal stimulation of mammary cells *in vitro, J. Biol. Chem.*, 246, 3835, 1971.
88. **Adams, R. L. P.,** Methylation of satellite DNA in vertebrate sperm, manuscript in preparation.
89. **Kaput, J. and Sneider, T. W.,** Methylation of somatic vs. germ cells DNA analysed by restriction endonuclease digestions, *Nucl. Acids Res.*, 7, 2303, 1979.
90. **Sturm, K. S. and Taylor, J. H.,** Distribution of 5 methylcytosine in the DNA of somatic and germline cells from bovine tissues, *Nucl. Acids Res.*, 9, 4537, 1981.
91. **Bird, A. P., Taggart, M. H., and Smith, B. A.,** Methylated and unmethylated DNA compartments in the sea urchin genome, *Cell,* 17, 889, 1979.
92. **Bird, A. P. and Taggart, M. H.,** Variable patterns of total DNA and rDNA methylation in animals, *Nucl. Acids Res.*, 8, 1485, 1980.
93. **Bird, A. P.,** DNA methylation and the frequency of CpG in animal DNA, *Nucl. Acids Res.*, 8, 1499, 1980.
94. **Fry, K., Poon, R., Whitcome, P., Idriss, J., Salser, W., Mazrimas, J., and Hatch, F.,** Nucleotide sequence of HS-β satellite DNA from kangaroo rat, *Dipodomys ordii, Proc. Natl. Acad. Sci. U.S.A.*, 70, 2642, 1973.

95. **Deumling, B.,** Sequence arrangement of a highly methylated satellite DNA of a plant, *Scilla;* a tandemly repeated inverted repeat, *Proc. Natl. Acad. Sci. U.S.A.,* 78, 338, 1981.
96. **Wozney, J., Hanahan, D., Tate, V., Boedtker, H., and Doby, P.,** Structure of the pro α 2(I) collagen gene, *Nature (London),* 294, 129, 1982.
97. **Horz, W. and Altenburger, W.,** Nucleotide sequence of mouse satellite DNA, *Nucl. Acids Res.,* 9, 683, 1981.
98. **Saloman, R., Kaye, A. M., and Herzberg, M.,** Mouse nuclear satellite DNA: 5 methylcytosine content, primidine isoplith distribution and electron microscopic appearance, *J. Mol. Biol.,* 43, 581, 1969.
99. **Gautier, F., Bunemann, H., and Grotjahn, L.,** Analysis of calf thymus satellite DNA: evidence for specific methylation of cytosine in C-G sequences, *Eur. J. Biochem.,* 80, 175, 1977.
100. **Schneiderman, M. H. and Billen, D.,** Methylation of rapidly reannealing DNA during the cell cycle of Chinese hamster cells, *Biochim. Biophys. Acta,* 308, 352, 1973.
101. **Drahovsky, D., Boehm, T. L. J., and Kreis, W.,** Distribution pattern and enzymic hypermethylation of inverted repetitive DNA sequences in P815 mastocytoma cells, *Biochim. Biophys. Acta,* 563, 28, 1979.
102. **Adams, R. L. P., Ulanovsky, M., Kirk, D., Qureshi, M., Gibb, S., and Burdon, R. H.,** Methylation of DNA of different renaturation classes isolated from various normal, transformed and tumour cells, *Eur. J. Cell Biol.,* 22, 34, 1980.
103. **Browne, M. J., Cato, A. C. B., and Burdon, R. H.,** The distribution of modified and non-modified C-G doublets in BHK-21 cell DNA, *FEBS Lett.,* 91, 69, 1978.
104. **Whittaker, P. A., McLachlan, A., and Hardman, N.,** Sequence organisation in nuclear DNA from *Physarum polycephalum:* methylation of repetitive sequences, *Nucl. Acids Res.,* 9, 801, 1981.
105. **Singer, J., Stellwagen, R. H., Robert-Ems, J., and Riggs, A. D.,** 5 Methylcytosine content of rat hepatoma DNA substituted with bromodeoxyuridine, *J. Biol. Chem.,* 252, 5509, 1977.
106. **Adams, R. L. P. and Davis, T.,** *In vitro* methylation of chromatin, in preparation.
107. **Adams, R. L. P., McKay, E. L., Douglas, J. T., and Burdon, R. H.,** Methylation of nucleosomal and nuclease sensitive DNA, *Nucl. Acids Res.,* 4, 3097, 1977.
108. **Razin, A. and Cedar, H.,** Distribution of 5 methylcytosine in chromatin, *Proc. Natl. Acad. Sci. U.S.A.,* 74, 2725, 1977.
109. **Solage, A. and Cedar, H.,** Organisation of 5 methylcytosine in chromosomal DNA, *Biochemistry,* 17, 2934, 1978.
110. **Bloch, S. and Cedar, H.,** Methylation of chromatin DNA, *Nucl. Acids Res.,* 3, 1507, 1976.
111. **Doenecke, D.,** Modification of DNA in chromatin with methyltransferase from *Haemophilus influenzae* Rd., *Eur. J. Biochem.,* 93, 481, 1979.
112. **Tosi, L. and Scarano, E.,** Effect of trypsin on DNA methylation in isolated nuclei from developing sea urchin embroys, *Biochem. Biophys. Res. Commun.,* 55, 470, 1973.
113. **Seidman, M. M., Levine, A. J., and Weintraub, H.,** The asymmetric segregation of parental nucleosomes during chromosome replication, *Cell,* 18, 439, 1979.
114. **Jackson, V. and Chalkley, R.,** A new method for the isolation of replicative chromatin: selective deposition of histone on both new and old DNA, *Cell,* 23, 121, 1981.
115. **Goodwin, G. H., Wright, C. A., and Johns, E. W.,** The characterisation of 1SF monomer nucleosomes from hen oviduct and the partial characterisation of a third HMG14/17 like protein in such nucleosomes, *Nucl. Acids Res.,* 9, 2761, 1981.
116. **Stalder, J., Larson, A., Engel, J. D., Dolan, M., Groudine, M., and Weintraub, H.,** Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNase I, *Cell,* 20, 451, 1980.
117. **Remington, A. and Ranjekar, P. K.,** DNA methylation in chromatin fractions of chick embryo cells, *Experientia,* 36, 178, 1980.
118. **Naveh-Many, T. and Cedar, H.,** Active gene sequences are undermethylated *Proc. Natl. Acad. Sci. U.S.A.,* 78, 4246, 1981.
119. **Miller, O. J., Schnedl, W., Allen, J., and Erlanger, B. F.,** 5 Methyl cytosine localised in mammalian constitutive heterochromatin, *Nature (London),* 251, 636, 1974.
120. **Creusot, F. and Christman, J. K.,** Localisation of DNA methyltransferase in the chromatin of Friend erythroleukaemia cells, *Nucl. Acids Res.,* 9, 5359, 1981.
121. **Vanyushin, B. F. and Kirnos, M. D.,** The nucleotide composition and pyrimidine clusters in DNA from beef heart mitochondria, *FEBS Lett.,* 39, 195, 1974.
122. **Vanyushin, B. F. and Kirnos, M. D.,** Structure of animal mitochondrial DNA (base composition, pyrimidine clusters, character of methylation), *Biochim. Biophys. Acta,* 475, 323, 1977.
123. **Nass, M. M. K.,** Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells, *in vivo* and *in vitro* methylation, *J. Mol. Biol.* 80, 155, 1973.
124. **David, I. B.,** 5 Methycytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells, *Science,* 184, 80, 1974.

125. **Groot, G. S. P. and Kroon, A. M.,** Mitochrondial DNA from various organisms does not contain internally methylated cytosine in —CCGG— sequences, *Biochim. Biophys. Acta,* 564, 355, 1979.
126. **Royer, H-D. and Sager, R.,** Methylation of chloroplast DNAs in the life cycle of *Chlamydomonas, Proc. Natl. Acad. Sci. U.S.A.,* 76, 5794, 1979.
127. **Sano, H., Grabowy, C., and Sager, R.,** Differential activity of DNA methyltransferase in the life cycle of *Chlamydomonas reinhardi, Proc. Natl. Acad. Sci. U.S.A.,* 78, 3118, 1981.
128. **Sager, R., Grabowy, C., and Sano, H.,** The *mat-1* gene in *Chlamydomonas* regulates DNA methylation during gametogenisis, *Cell,* 24, 41, 1981.
129. **Sano, H. and Sager, R.,** DNA methyltransferase from the eukaryote, *Chlamydomonas reinhardi, Eur. J. Biochem.,* 105, 471, 1980.
130. **Sharma, S. and Biswal, N.,** Studies on the *in vivo* methylation of replicating herpes simplex virus type 1 DNA, *Virology,* 82, 265, 1977.
131. **Gunthert, U., Schweiger, M., Stupp, M., and Doerfler, W.,** DNA methylation in adenovirus, adeno-transformed cells, and host cells, *Proc. Natl. Acad. Sci. U.S.A.,* 73, 3923, 1976.
132. **Kaye, A. M. and Winocour, E.,** On the 5 methylcytosine found in the DNA extracted from polyoma virus, *J. Mol. Biol.,* 24, 475, 1967.
133. **Borek, E. and Strinivasan, P. R.,** The methylation of nucleic acids, *Annu. Rev. Biochem.,* 35, 275, 1966.
134. **Evans, H. H., Evans, T. E., and Littman, A.,** Methylation of parental and progeny DNA strands in *Physarum polycephalum, J. Mol. Biol.,* 74, 563, 1973.
135. **Adams, R. L. P.,** Newly synthesised DNA is not methylated, *Biochim. Biophys. Acta,* 335, 365, 1974.
136. **Adams, R. L. P.,** The relationship between synthesis and methylation of DNA in mouse fibroblasts, *Biochim. Biophys. Acta,* 254, 205, 1971.
137. **Geraci, D., Eremenko, T., Cocchiara, R., Granieri, A., Scarano, E., and Volpe, P.,** Correlation between synthesis and methylation of DNA in HeLa cells, *Biochim. Biophys. Res. Commun.,* 57, 353, 1974.
138. **Theiss, G. and Follmann, H.,** 5 Methylcytosine formation in wheat DNA, 12th FEBS Meet. Dresden, Abstr. 0641, 1978.
139. **Kiryanov, G. I., Kirnos, M. D., Demidkina, N. P., Alexandrushkina, N. I., and Vanyushin, B. F.,** Methylation of DNA in cells on replication, *FEBS Lett.,* 112, 225, 1980.
140. **Cumings, D.,** Methylation of euchromatic and heterochromatic DNA, *Exp. Cell Res.,* 74, 383, 1972.
141. **Turnbull, J. F. and Adams, R. L. P.,** DNA methylase: purification from ascites cells and the effect of various DNA substrate on its activity, *Nucl. Acids Res.,* 3, 677, 1976.
142. **Adams, R. L. P.,** Delayed methylation of DNA in developing sea urchin embryos, *Nature New Biol.,* 244, 27, 1973.
143. **Magnavel, R., Bertaux, O., and Valencia, R.,** Methylation of nucleic acids in chromatin fractions during the cell cycle of *Euglena gracilis,* 12th FEBS Meet. Dresden, Abstr. 0638, 1978.
144. **Bugler, B., Valencia, R., and Bertaux, O.,** Mitotic overmethylation of DNA in synchronous BHK21 fibroblasts -HS 5 transformed 12th FEBS Meet., Dresden, Abstr. 0637, 1978.
145. **Hotta, Y. and Hecht, N.,** Methylation of *Lilium* DNA during the meiotic cycle, *Biochim. Biophys. Acta,* 238, 50, 1971.
146. **Drahovsky, D. and Walker, A.,** Enzymatic methylation of replication-DNA intermediates in Ehrlich ascites tumour, *Naturwissenschaften,* 62, 189, 1975.
147. **Seale, R. L.,** Assembly of DNA and protein during replication in HeLa cells, *Nature (London),* 255, 247, 1975.
148. **Adams, R. L. P. and Hogarth, C.,** DNA methylation in isolated nuclei: old and new DNA is methylated, *Biochim. Biophys. Acta,* 331, 214, 1973.
149. **Wigler, M., Levy, D., and Perucho, M.,** The somatic replication of DNA methylation, *Cell,* 24, 33, 198.
150. **Sinsheimer, R. L.,** The action of pancreatic DNase:II isomeric dinucleotides, *J. Biol. Chem.,* 215, 579, 1955.
151. **Doskocil, J. and Sorm, F.,** Distribution of 5 methylcytosine in pyrimidine sequences of DNA, *Biochim. Biophys. Acta,* 55, 953, 1962.
152. **Browne, M. J. and Burdon, R. H.,** The sequence specificity of vertebrate DNA methylation, *Nucl. Acids Res.,* 4, 1025, 1977.
153. **Browne, M. J., Turnbull, J. F., McKay, E. L., Adams, R. L. P., and Burdon, R. H.,** The sequence specificity of a mammalian DNA methylase, *Nucl. Acids Res.,* 4, 1039, 1977.
154. **van der Ploeg, L. H. T., Groffen, J., and Flavell, R. A.,** A novel type of secondary modification of two CCGG residues in the human γ δ β-globin gene locus, *Nucl. Acids Res.,* 8, 4563, 1980.
155. **Gruenbaum, Y., Naveh-Many, T., Cedar, H., and Razin, A.,** Sequence specificity of methylation in higher plant DNA, *Nature (London),* 292, 860, 1981.

156. **Gruenbaum, Y., Stein, R., Cedar, H., and Razin, A.**, Methylation of CpG sequences in eukaryotic DNA, *FEBS Lett.*, 124, 67, 1981.
157. **Sneider, T. W.**, Methylation of mammalian DNA, *J. Biol. Chem.*, 247, 2872, 1972.
158. **Simon, D., Grunert, F., Kroger, H., and Grassman, A.**, *In vitro* methylation of SV 40 DNA by a methylase from rat liver, *Eur. J. Cell Biol.*, 22, 33, 1980.
159. **Fisher, E. F. and Caruthers, M. H.**, Studies on gene control regions. XII. The functional significance of a lac operator constitutive mutant, *Nucl. Acids Res.*, 7, 401, 1979.
160. **Adams, R. L. P., McKay, E. L., Craig, L. M., and Burdon, R. H.**, Mouse DNA methylase: methylation of native DNA, *Biochim. Biophys. Acta*, 561, 345, 1979.
161. **Yamamoto, M. and Miklos, G. L. G.**, Genetic studies on heterochromatin in *D. melanogaster* and their implications for the functions of satellite DNA, *Chromosoma*, 66, 71, 1978.
162. **Bird, A. P.**, Use of restriction enzymes to study eukaryotic DNA methylation. II. The symmetry of methylated site supports a semi-conservative copying of the methylation pattern, *J. Mol. Biol.*, 118, 49, 1978.
163. **Caboche, M. and Bachellerie, J-P.**, RNA methylation and control of eukaryotic RNA biosynthesis, *Eur. J. Biochem.*, 74, 19, 1977.
164. **Wigler, M. H.**, The inheritance of methylation patterns in vertebratae, *Cell*, 24, 285, 1981.
165. **Compere, S. J. and Palmiter, R. D.**, DNA methylation controls the inducibility of the mouse metallothionein-I gene in lymphoid cells, *Cell*, 25, 233, 1981.
166. **Qureshi, M., Burdon, R. H., and Adams, R. L. P.**, A DNA methylase tightly bound to the chromatin scaffold, manuscript in preparation.
167. **Hilliard, J. K. and Sneider, T. W.**, Repair methylation of parental DNA in synchronised cultures of Novikoff hepatoma cells, *Nucl. Acids Res.*, 2, 809, 1975.
168. **Drahovsky, D., Lacko, I., and Wacker, A.**, Enzymatic DNA methylation during repair synthesis in non-proliferating human peripheral lymphocytes, *Biochim. Biophys. Acta*, 447, 139, 1976.
169. **Gurdon, J. B.**, Adult frogs derived from the nuclei of single somatic cells, *Dev. Biol.*, 4, 256, 1962.
170. **Slack, J. M. W.**, The source of cells for regeneration, *Nature (London)*, 286, 760, 1980.
171. **Lapeyre, J-N. and Becker, F. F.**, 5 Methylcytosine content of nuclear DNA during chemical hepatocarcinogenisis and in carcinomas which result, *Biochem. Biophys. Res. Commun.*, 87, 698, 1979.
172. **Waalwijk, C. and Flavell, R. A.**, DNA methylation at a CCGG sequence in the large intron of the rabbit β-globin gene: tissue specific variations, *Nucl. Acids Res.*, 5, 4631, 1978.
173. **Shen, C-K. J. and Maniatis, T.**, Tissue specific DNA methylation in a cluster of rabbit β-like globin gene; *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6634, 1980.
174. **van der Ploeg, L. H. T. and Flavell, R. A.**, DNA methylation in the human γ δ β-globin locus in erythroid and nonerythroid tissues, *Cell*, 19, 947, 1980.
175. **McGhee, J. D. and Ginder, G. D.**, Specific DNA methylation sites in the vicinity of the chicken β-globin genes, *Nature (London)*, 280, 419, 1979.
176. **Marcaud, L., Reynaud, C-A., Therwath, A., and Sherrer, K.**, Modification of the methylation pattern in the vicinity of the chicken globin genes in avian erythroblastosis virus transformed cells, *Nucl. Acids Res.*, 9, 1841, 1981.
177. **Groudine, M. and Weintraub, H.**, Activation of globin genes during chicken development, *Cell*, 24, 393, 1981.
178. **Weintraub, H., Larsen, A., and Groudine, M.**, α-Globin gene switching during the development of chicken embryos: expression and chromosome structure, *Cell*, 24, 333, 1981.
179. **Anderson, W. F. and Diacumakos, E. G.**, Genetic engineering in mammalian cells. *Sci. Am.*, 245(1), 60, 1981.
180. **Kuo, M. T., Mandel, J. L., and Chambon, P.**, DNA methylation: correlation with DNase I sensitivity of chicken ovalbumin and conalbumin chromatin, *Nucl. Acids Res.*, 7, 2105, 1979.
181. **Vardimon, L., Neumann, R., Kuhlmann, I., Sutter, D., and Doerfler, W.**, DNA methylation and viral gene expression in adenovirus-transformed and infected cells, *Nucl. Acids Res.*, 8, 2461, 1980.
182. **Sutter, D. and Doerfler, W.**, Methylation of integrated adenovirus type 12 DNA sequences in transformed cells is inversely correlated with viral gene expression, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 253, 1980.
183. **Desrosiers, R. C., Mulder, C., and Fleckenstein, B.**, Methylation of *Herpesvirus saimiri* DNA in lymphoid tumour cell lines, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 3839, 1979.
184. **Cohen, J. C.**, Methylation of milk-borne and genetically transmitted mouse mammary tumourvirus proviral DNA, *Cell*, 19, 653, 1980.
185. **Nakhasi, H. L., Lynch, K. R., Dolan, K. P., Unterman, R. D., and Feigelson, P.**, Covalent modification and repressed transcription of a gene in hepatoma cells, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 834, 1981.
186. **Tantravahi, U., Breg, W. R., Wertelecki, V., Erlanger, B. F., and Miller, O. J.**, Evidence for methylation of inactive human rRNA genes in amplified regions, *Hum. Genet.*, 56, 319, 1981.

187. **Tantravahi, U., Guntaka, R. V., Erlanger, B. F., and Miller, O. J.,** Amplified ribosomal RNA genes in a rat hepatoma cell line are enriched in 5 methylcytosine, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 489, 1981.
188. **Dawid, I. B., Brown, D. D., and Reeder, R. H.,** Composition and structure of chromosomal and amplified ribosomal DNAs of *Xenopus laevis, J. Mol. Biol.*, 51, 341, 1970.
189. **Bird, A. P. and Southern, E. M.,** Use of restriction enzymes to study eukaryotic DNA methylation. I. The methylation pattern in ribosomal DNA from *Xenopus laevis, J. Mol. Biol.*, 118, 27, 1978.
190. **Dandekar, A. M. and Qasba, P. K.,** Inverse relationship between expression of α-lactabumin and methylation of its gene sequences in rat mammary gland and tumours, *Eur. J. Cell Biol.*, 22, 10, 1980.
191. **Jones, R. E., DeFao, D., and Piatigorsky, J.,** Transcription and site specific hypomethylation of the δ-crystallin genes in the embryonic chicken lens, *J. Biol. Chem.*, 256, 8172, 1981.
192. **Baralle, F. E., Shoulders, C. C., Goodbourn, S., Jeffreys, A., and Proudfoot, N. J.,** The 5′ flanking regions of human ϵ-globin gene, *Nucl. Acids Res.*, 8, 4393, 1980.
193. **Moss, T., Boseley, P. C., and Birnstiel, M. L.,** More ribosomal spacer sequences from *Xenopus laevis, Nucl. Acids Res.*, 8, 467, 1980.
194. **Hall, L. M. C. and Maden, B. E. H.,** Nucleotide sequence through the 18S-28S intergene regions of a vertebrate ribosomal transcription unit, *Nucl. Acids Res.*, 8, 5993, 1980.
195. **Maden, B. E. H., Hall, L. M. C., and Salim, M.,** Ribosome formation in the eukaryotic nucleolus: recent advances from sequence analysis, in *The Nucleolus,* Jordan, E. G. and Cullis, C.A., Eds., S.E.B. symposium in press.
196. **Furlong, J. and Maden, B. E. H.,** personal communication.
197. **Jones, P. A. and Taylor, S. M.,** Hemimethylated duplex DNAs prepared from 5 azacytidine-treated cells, *Nucl. Acids Res.*, 9, 2933, 1981.
198. **Adams, R. L. P., Fulton, J., and Kirk, D.,** The effect of 5 azadeoxycytidine on cell growth and DNA methylation, *Biochim. Biophys. Acta,* 697, 286, 1982.
199. **Pfohl-Leszkowicz, A., Salas, C., Fuchs, R. R. P., and Dirtheines,** Mechanism of inhitition of enzymatic DNA methylation by 2-(acetylamino) fluorene bound to DNA, *Biochemistry,* 20, 3020, 1981.
200. **Jones, R. A. and Taylor, S. M.,** Cellular differentiation, cytidine analogs and DNA methylation, *Cell,* 20, 85, 1980.
201. **Taylor, S. M. and Jones, P. A.,** Multiple new phenotypes induced in 10T½ and 3T3 cells treated with 5 azacytidine, *Cell,* 17, 771, 1979.
202. **Groudine, M., Eisenman, R., and Weintraub, H.,** Chromatin structure and endogenous retroviral genes and activation by an inhibitor of DNA methylation, *Nature (London),* 292, 311, 1981.
203. **Vesely, J., Cihak, A., and Sorm, F.,** Metabolic alterations in leukaemic mouse cells following 5-aza-2′ deoxycytidine, *Cell Czech. Chem. Commun.*, 34, 901, 1969.
204. **Mohandas, T., Sparkes, R. S., and Shapiro, L. J.,** Reactivation of an inactive human X -chromosome — evidence for X inactivation by DNA methylation, *Science,* 211, 393, 1981.
205. **Schlosser, C. A., Steglich, C., de Wet, J. R., and Scheffler, I. E.,** Cell cycle dependent regulation of thymidine kinase activity introduced into mouse LHTK$^-$ cells by DNA and chromatin-mediated gene transfer, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 1119, 1981.
206. **Lin, S. S. and Munyon, W.,** Expression of the viral thymidine kinase gene in herpes simplex virus-transformed L cells, *J. Virol.*, 14, 1199, 1974.
207. **Pollack, Y., Stein, R., Razin, A., and Cedar, H.,** Methylation of foreign DNA sequences in eukaryotic cells, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6463, 1980.
208. **Pellicer, A., Robins, D., Wold, B., Sweet, R., Jackson, J., Lowry, I., Roberts, J. M., Sim, G. K., Silverston, S., and Axel, R.,** Altering genotype and phenotype by DNA-mediated gene transfer, *Science,* 209, 1414, 1980.
209. **Sager, R. and Kitchen, R.,** Selective silencing of eukaryotic DNA, *Science,* 189, 426, 1975.
210. **Riggs, A. D.,** X inactivation, differentiation, and DNA methylation, *Cytogenet. Cell Genet.*, 14, 9, 1975.
211. **Bostock, C. J. and Clark, E. M.,** Satellite DNA in large marker chromosomes of methotrexate-resistant mouse cells, *Cell,* 19, 709, 1980.
212. **Eastman, E. M., Goodman, R. M., Erlanger, B. F., and Miller, O. J.,** 5 Methylcytosine in the DNA of the polytene chromosomes of the diptera *Sciara caprophila, Drosophila melanogaster and D. persimilis, Chromosoma,* 79, 225, 1980.
213. **Miller, O. J., Tantravahi, U., Katz, R., Erlanger, B. F., and Guntaka, R. V.,** Amplification of mammalian ribosomal RNA genes and their regulation by methylation in *Genes, Chromosomes and Neoplasia,* Arrighi, F. E., Rao, P. N., and Stubblefield, E., Eds., Raven Press, New York, 1981.
214. **Scarano, E.,** Enzymatic modifications of DNA and embryonic differentiation, *Ann. Embryol. Morphogen.*, 1(Suppl.), 51, 1969.
215. **Tosi, L., Granieri, A., and Scarano, E.,** Enzymatic DNA modifications in isolated nuclei from developing sea urchin embryos, *Exp. Cell. Res.*, 72, 257, 1972.

216. **Holliday, R. and Pugh, J. E.,** DNA modification mechanisms and gene activity during development, *Science,* 187, 226, 1975.
217. **Cooper, K. W.,** Meiotic conjunctive elements not involving chiasmata, *Proc. Natl. Acad. Sci. U.S.A.,* 52, 1248, 1964.
218. **Chen, M-J. and Nienhuis, A. W.,** Structure and expression of human globin genes introduced into mouse fibroblasts, *J. Biol. Chem.,* 256, 9680, 1981.
219. **Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. P., and Chambon, P.,** The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants, *Nucl. Acids, Res.,* 9, 6047, 1981.
220. **Mellon, P., Parker, V., Gluzman, Y., and Maniatis, T.,** Identification of DNA sequences required for transcription of the human α1-globin gene in a new SV40 host-vector system, *Cell* 27, 279, 1981.
221. **Banerji, J., Rusconi, S., and Schaffner, W.,** Expression of a β-globin gene is enhanced by remote SV40 DNA sequences, *Cell,* 27, 299, 1981.
222. **Taylor, J. H.,** Enzymatic methylation of DNA: pattern and possible regulatory roles. Chromosome structure, in *Molecular Genetics,* Vol. 3, Taylor, J. H., Ed., Academic Press, New York, 1979, 89.
223. **Burdon, R. H. and Adams, R. L. P.,** Eukaryotic DNA methylation, *Trends Biochem. Sci.,* 5, 294, 1980.
224. **Drahovsky, D. and Boehm, T. L. J.,** Enzymatic DNA methylation in higher eukaryotes, *Int. J. Biochem.,* 12, 523, 1980.
225. **Razin, A. and Riggs, A. D.,** DNA methylation and gene function, *Science,* 210, 604, 1980.
226. **Razin, A. and Friedman, J.,** DNA methylation and its possible biological roles, *Prog. Nucl. Acid Res. Mol. Biol.,* 25, 33, 1981.
227. **Ehrlich, M. and Wang, R. Y-H.,** 5 Methylcytosine in eukaryotic DNA, *Science,* 212, 1350, 1981.