PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

# The Significance of DNA Methylation Patterns: Promoter Inhibition by Sequence-Specific Methylation is One Functional Consequence

W. Doerfler

*Phil. Trans. R. Soc. Lond. B* 1990 **326**, 253-265
doi: 10.1098/rstb.1990.0009

*Phil. Trans. R. Soc. Lond.* B **326**, 253–265 (1990)
*Printed in Great Britain*



# The significance of DNA methylation patterns: promoter inhibition by sequence-specific methylation is one functional consequence

By W. Doerfler

*Institute of Genetics, Universität zu Köln, D-5000 Köln, F.R.G.*

This paper presents a review of previously published results from my laboratory on the inactivating or inhibiting function of sequence-specific methylation on promoter activity. In this study, viral promoters, mostly those from adenovirus type 2 (Ad2) or type 12 (Ad12), have been used. It has also been shown that the transcriptional block of these methylated viral promoters can, at least partly, be reversed by a transactivating protein or by the presence of a strong enhancer. We have also adduced evidence that the methylation of the late E2A promoter of Ad2 DNA at positions +6 and +24 from the cap site of this promoter interferes with the binding of one or several proteins at these particular sites, at least when 50 or 73 base-pair long fragments of this promoter have been used for studies on protein binding. With a 377 base-pair fragment, binding differences between the unmethylated and the 5′-CCGG-5′ methylated late E2A promoter are not obvious. By applying the genomic sequencing technique developed by Church & Gilbert (1984), the patterns of methylation in all 5′-CG-3′ dinucleotides in the late E2A promoter in the active or inactive state in different Ad2-transformed cell lines have been determined. It has been found that 5-methyldeoxycytidine residues introduced into foreign DNA, which is then integrated into the mammalian cell genome, can lead to the spreading of methylation starting from the point of initial methylation. We have begun to investigate whether certain patterns of methylation in mammalian DNA can also influence biological processes other than promoter activity. We have developed a cell-free system using nuclear extracts from hamster cells to study recombination between Ad12 DNA and hamster pre-insertion sites into which Ad12 DNA had previously integrated. The DNAs used in recombination experiments are in the unmethylated or the methylated form.

Some speculative aspects have also been discussed in this review. Could existing patterns of methylation in mammalian (human) DNA represent composites of several interdigitating patterns each one of which might have a different signal value? Can a 5-methyldeoxycytidine group in DNA modulate DNA–protein interactions in a positive or negative way, for proteins which could have positively or negatively regulating functions? Patterns of methylation appear to be relatively stable over many years for cell lines propagated in culture. Are patterns of methylation stable also in different parts of the human chromosome? To what extent are these patterns inheritable?

## 1. Introduction

Molecular biology has enabled many biological phenomena to be described in biochemical terms. Among these reactions, the recognition of specific DNA sequences by proteins has assumed an important role. The interaction of regulatory proteins with sequence motifs in prokaryotic and, in particular, in eukaryotic promoters lies at the core of the regulation of gene expression. Many complex biological phenomena, for example, differentiation and embryonic development, the regulation of the immune system and the regulated expression of viral genomes can be explained by an understanding of integrated patterns of differential gene

expression. The realm of molecular pathology will provide a basis for clinical medicine in the future. It is a reasonable assumption that many chronic diseases, including malignancies, may be caused, or are certainly accompanied, by severe disregulations of gene expression in different parts of the genome.

DNA or RNA motif recognition by specialized polypeptides has a central function in the regulation of gene expression, in DNA replication, in recombination, in the processing of RNA and in many other biological mechanisms. Although the primary genetic code, which relates nucleotide sequence to amino acid sequence in protein structure, has been deciphered in the classical period of molecular biology, the genetic codes determining motif recognition by proteins are not well understood. Once a sequence motif has been established in evolution, it seems to be meticulously conserved. One of the very few possibilities for modulating an existing sequence motif is the post-replicational attachment of a methyl group to deoxycytidine in eukaryotes and to deoxycytidine or deoxyadenosines in prokaryotes. In this study, I investigated the role of DNA methylation exclusively in mammalian systems. The methyl group in 5-methyldeoxycytidine (5-mC) presents a physically conspicuous protuberance in the major groove of the double helical DNA molecule, and it must be that protuberance, or unknown disturbances generated by it, that might signal altered interactions of polypeptides with a known sequence motif. To what extent such a signal will be able to interfere with the binding of a functionally essential protein or proteins, cannot *a priori* be predicted. The functional consequences probably depend on the nature of the specific interaction and have to be determined experimentally. The presence of a methyl group might not alter the binding of a protein A to DNA directly, but only that of functionally essential protein B, which is able to recognize protein A after the attachment to a DNA sequence motif. The modulatory effect of a methyl group could be functionally reversible, either by the physical elimination of 5-mC, if such a mechanism existed, or in more subtle ways, by transactivators or enhancers that could lead to the elimination of the modifying effect that a 5-mC signal might exert.

In analysing the overall biological implications of DNA methylation, it has been useful to take a critical, reductionist, approach and deduce the effects on reactions that can be measured in a reliable way. In that sense, work on the inhibitory function of sequence-specific promoter methylations offers opportunities to contribute to the understanding of gene regulation. It has been our aim to decipher this inhibitory mechanism in as much detail as possible. After having studied gene regulation in detail, one can begin to consider more complicated consequences of DNA methylation patterns, for example, the involvement in genomic imprinting (Reik *et al.* 1987; Surani, this symposium), development and differentiation, or X-chromosome inactivation (see Holliday, Monk and Rigg, this symposium). The study of these complex biological problems is based partly on the evidence for an inactivating function of sequence-specific methylations in eukaryotic gene expression. Patterns of methylation seem to be inherited at least in cell lines carried in culture, but the mode of inheritance is poorly understood. Many genomic segments in mammalian DNA will have to be investigated under a variety of conditions, for example, endogenous DNA compared with foreign DNA, recently added artificially to a genome, before newly and valid generalized conclusions can be drawn. There are reasons to question the notion that DNA artificially added to a genome can be considered to be identical in structure and organization to endogenous cellular DNA. The expression of inserted DNA might not be regulated in a way identical to that of endogenous genes. The type of inheritance of a given pattern of methylation may depend on the geography




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

of the mammalian genome. One of the important unresolved problems in DNA methylation work is to determine how a given pattern of methylation is generated, how it is maintained, and whether it is subject to fluctuations.

Ideally, one would first analyse the effects of methylated sequences on the interference with the binding of regulatory proteins, then proceed to the inhibition of promoter activity, and eventually to a study of more complex consequences. The best studied and technically exploited influence of a methylated nucleotide on specific DNA–protein interactions is the inhibition of restriction endonucleases by sequence-specific methylations. In general, a methylated sequence inhibits restriction endonucleases, that is, it has a negative effect on this particular DNA–protein interaction. There are rare examples of positive effects, such as that of the restriction endonuclease *Dpn*I, which requires for its activity an *N*-6-methyldeoxyadenosine residue in the sequence 5′-GATC-3′. Hence, examples are documented for negative and positive modulations of the activity of restriction endonucleases. Methylated nucleotides might, therefore, be expected to exert either effect also in the more complex mammalian systems.

In eukaryotes, the influence of DNA methylation on functions other than gene regulation has been studied to a limited extent. Considering the complexities in patterns of methylation in well-studied segments of the mammalian chromosome, it has been suggested that the observed patterns might be composed of several interdigitating patterns, each of which might have a different function. Thus a cluster of 5-mC residues in a DNA sequence might influence several different mechanisms. It will not be a simple task to determine the role that individual methylated nucleosides play in different reactions, particularly because these effects could be exerted via changes in the chromatin structure of individual genome segments.

## 2. The experimental system

Most of our work on the biological effects of DNA methylation has utilized the adenovirus system. Adenovirus-transformed or adenovirus type 12 (Ad12)-induced tumour cells (for review of cell lines, see Doerfler (1982)) or adenovirus promoters have been used, depending on whether the patterns of methylation or the details of the promoter-inhibiting effect, respectively have been investigated. Integrated viral genomes can be analysed under different conditions and in cell lines from different species. Individual viral promoters can be isolated, cloned and their activity can be studied after reintroduction into cells or in cell-free transcription experiments. In this way, it has been possible to imitate methylation patterns of active or inactive viral promoters that were previously determined in living cells in culture. The activity levels of the viral promoters that had been left unmethylated or had been methylated *in vitro* were then assessed under a variety of experimental conditions (see table 1). The majority of our experiments were performed with the late E2A (E2AL) promoter of adenovirus type 2 (Ad2) DNA (Langner *et al.* 1984, 1986; Müller & Doerfler 1987; Dobrzanski *et al.* 1988; Knebel-Mörsdorf *et al.* 1988; Weisshaar *et al.* 1988; Knust *et al.* 1989). In some experiments, the E1A promoter of Ad12 DNA (Kruczek & Doerfler 1983; Knebel & Doerfler 1986) or the major late promoter (MLP) of Ad2 DNA (Dobrzanski *et al.* 1988) were also investigated.

As it is still questionable whether 5-mC can occur in *Drosophila* cells or in cells of the lepidopteran cell line *Spodoptera frugiperda*, it has been interesting to determine the consequences of the sequence-specific methylation of an insect virus promoter in these cells. In the




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

bacaulovirus genome of *Autographa californica* nuclear polyhedrosis virus (AcNPV), the p10 gene encoding a 10000 relative molecular mass polypeptide of unknown function was abundantly expressed late in the viral infection cycle. In *Spodoptera frugiperda* cells, the p10 gene promoter was inactivated when one 5′-CCGG-3′ sequence in the promoter and two such sequences in the coding part of the gene were methylated (Knebel *et al.* 1985). These results indicate that insect cells can recognize this genetic signal as an inactivating module in a viral promoter. Thus it appears necessary to apply more sensitive techniques in the search for 5-mC or *N*-6-mA in insect-cell DNA, before final conclusions about the absence of modified nucleosides are reached.

In further experiments, we demonstrated that a viral *trans*-activating protein (Langner *et al.* 1986; Weisshaar *et al.* 1988; Knust *et al.* 1989) or a strong viral enhancer (Knebel-Mörsdorf *et al.* 1988), could at least partly reverse the promoter-inhibiting or inactivating effects of sequence-specific methylations. We will continue to investigate viral promoters or viral gene products and genetic elements to gain improved insights into the functional consequences of DNA methylation, but will also investigate cellular promoters in this respect. The focus of our work has so far been on the inactivation by sequence-specific methylation of eukaryotic promoters and on the study of the emergence and the maintenance of methylation patterns. As will be discussed in more detail, the finding that patterns of methylation in an endogenous viral promoter are by far more extensive and complex than would appear to be essential for promoter inactivation, has raised the question whether a given methylation pattern is composed of several, functionally different patterns that are superimposed on each other. Another important problem requires clarification, that is, to what extent could DNA added artificially to cells – in culture or in transgenic animals – truly reflect the functional equivalent of an endogenous gene. This question relates to much experimental work in contemporary molecular biology. These problems are not understood at present, but point to necessary experimental alternatives in future work.

### 3. Patterns of methylation in the late E2A promoter of Ad2 DNA

The map of the late E2A 5′ region of Ad2 DNA in figure 1 presents all the 5′-CG-3′ dinucleotide sequences present in this segment of DNA. Only three out of a total of 39 of these sequences are located in the *Hpa*II recognition sequence (5′-CCGG-3′), but none in the *Hha*I recognition sequence (5′-GCGC-3′). It was, therefore, necessary to determine in the Ad2-transformed hamster cell lines HE1 and HE2 the status of methylation in as many of the 5′-CG-3′ dinucleotide sequences as possible in the E2AL promoter of Ad2 DNA. The genomic sequencing method (Church & Gilbert 1984) was the only technique available to survey all 5′-CG-3′ dinucleotides and all other cytidine residues in a sequence for the presence of 5-mC. The E2AL promoter segment from nucleotides +24 to −160, in relation to the +1 site of transcriptional initiation, the cap site of the promoter, was analysed by using the genomic sequencing procedure (Toth *et al.* 1989).

We have previously demonstrated an inverse correlation between methylation at the three 5′-CCGG-3′ sequences in positions +24, +6, and −215 in relation to the cap site of the late E2A promoter of adenovirus type 2 (Ad2) DNA and its activity (Vardimon *et al.* 1980). The patterns of methylation in all 5′-CG-3′ sequences over a region of about 180 base pairs (b.p.) required for gene activity in the late E2A promoter of integrated Ad2 DNA were determined




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

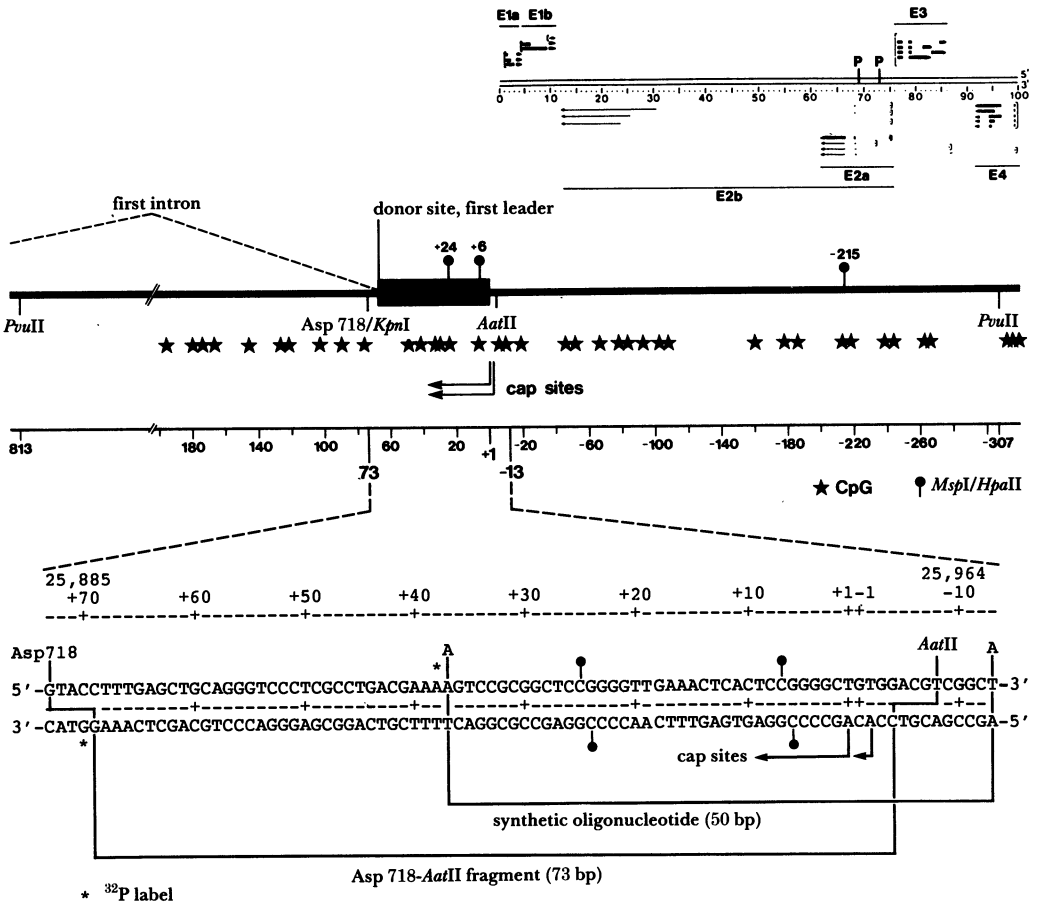


FIGURE 1. Functional map of the late E2A promoter of Ad2 DNA and sequence of the synthetic oligonucleotide and the restriction fragment used in protein binding. The map details important restriction sites, the cap site of the promoter and the nucleotide sequence of the synthetic 50 bp oligonucleotide and the 73 bp restriction endonuclease fragment. These oligonucleotides contained either C or 5-mC in the +24 and +6 positions relative to the cap site (+1) of the promoter. The asterisks show all the 5′-CpG-3′ sequences. For reference, the map locations of the early Ad2 functions are presented on top. The map indicates the *Pvu*II (P) fragment that carries the late E2A promoter (Hermann *et al.* 1989).

in cell lines that carried this promoter in an active or inactive state. In the cell line HE1, the late E2A promoter was active and all thirteen 5′-CG-3′ sequences between positions +24 and −160 were unmethylated. In cell line HE2, the same promoter was permanently shut off and all 5′-CG-3′ sequences were methylated in both strands. Thus the inverse correlation was perfect in these cell lines over a region of about 180 b.p. in the late E2A promoter (Toth *et al.* 1989).

The same promoter segment was analysed in cell lines mc23 and mc40 in which an E2AL promoter–chloramphenicol acetyltransferase (CAT) gene construct had been genomically fixed after transfection and before *in vitro* 5′-CCGG-3′ methylation (Müller & Doerfler 1987). In cell line mc23, the pre-imposed methylation pattern was stable and the cells contained an inactive CAT gene. Genomic sequencing confirmed the presence of 5-mC at the 5′-CCGG-3′ sequences and revealed the spreading of methylation to neighbouring 5′-CG-3′ sequences along the entire promoter (Toth *et al.* 1989). In cell line mc40, several of the 120 integrated copies were demethylated in positions +24 and +6, but the promoter was methylated

upstream of position −50. The cell line expressed the CAT gene. In cell line uc2, the unmethylated E2AL promoter-CAT gene construct had been genomically fixed. The results of the genomic sequencing analysis revealed that all 5′-CG-3′ sequences remained unmethylated (Toth *et al.* 1989).

As it was shown that the E2AL promoter of Ad2 DNA was inactivated by the methylation of the 5′-CCGG-3′ sequences at positions +24, +6 and −215 (for review, see Doerfler (1989)), and that at least thirteen 5′-CG-3′ dinucleotides in the same promoter were methylated when it persisted in the inactive state (Toth *et al.* 1989), for example, in cell line HE2 (Vardimon *et al.* 1980; Vardimon & Doerfler 1981), the excess of 5-mC residues in that promoter sequence posed an additional problem. What was the function of these 5-mC residues? A number of possibilities can be considered.

1. The introduction of an excess of 5-mC groups may be due to the spreading effect of DNA methylation that has been described above.

2. The additional 5-mC residues serve as a safeguard for long-term promoter inactivation (Doerfler 1983).

3. A high level of DNA methylation may be dependent upon and a characteristic of the chromosomal domain into which the viral (foreign) DNA has been integrated.

4. A given pattern of methylation may have more than one functional meaning, that is, the established pattern could be composed of different interdigitating patterns, each having a different or complementing functional significance, for example, on promoter function and chromatin structure. It will be a challenge to investigate these possibilities further.

We have shown that foreign (adenoviral) DNA, which integrated only recently into a mammalian genome, did not assume its final pattern of methylation immediately after insertion. Ad12 DNA that had become integrated into a newly induced hamster tumour exhibited a very low degree of methylation (Kuhlmann & Doerfler 1982; Kuhlmann *et al.* 1982). It took a considerable number of cell generations, before a distinct, possibly final, pattern of methylation could be distinguished in the integrated viral DNA (Kuhlmann & Doerfler 1983). Usually, the inactive, late viral genes were highly methylated and the active early genes showed hypomethylation (Sutter & Doerfler 1979, 1980). Thus the mechanisms determining the extent and patterns of DNA methylation in DNA segments that had been recently added to an existing mammalian genome were dependent upon time and probably upon continuous cell replication, to eventually generate a definite pattern with all its assigned functional consequences. This observed gradual establishment of a final pattern of methylation was consistent with the reported spreading effect of DNA methylation. I consider it one of the most interesting questions in DNA methylation to unravel these interdependencies.

## 4. Sequence-specific promoter methylations and gene inactivation: a synopsis of previously published work

For reasons of limited space, I shall restrict this review to work done on adenovirus promoters. The authors of other papers in this symposium will summarize results on other eukaryotic systems. In addition, I will refrain from repeating in detail results that have been published and reviewed before (Doerfler 1981, 1983, 1989; Doerfler *et al.* 1985, 1988). Table 1 shows a summary of the results obtained from experiments with the E2AL promoter of




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 1. Inactivation of the late E2A promoter of Ad2 DNA: survey of test systems

1. Inverse correlations in Ad12- or in Ad2-transformed hamster cell lines between DNA methylation and gene activity (Sutter & Doerfler 1979, 1980; Vardimon *et al.* 1980)

| | | | |
|---|---|---|---|
| HE1 | E2A active | T637 | E3 active |
| HE2 | E2A inactive | HA12/7 | E3 inactive |
| HE3 | E2A inactive | | |

2. *In vitro* 5′-CCGG-3′ methylation of late E2A promoter transient expression:
   (i) microinjection into *Xenopus laevis* oocytes (E2A as indicator gene) (Langner *et al.* 1984)
   (ii) transfection into mammalian cells (chloramphenicol acetyltransferase gene as indicator) (Langner *et al.* 1986)
3. Cell-free transcription system (HeLa nuclear extracts) (Dobrzanski *et al.* 1988) essentials:
   (i) circular template
   (ii) high protein concentration
   (iii) pre-test of DNA–protein ratio
4. Genomic fixation of the unmethylated or the 5′-CCGG-3′ methylated E2AL promoter-CAT gene construct in hamster cells (Müller & Doerfler 1987)

Ad2 DNA. These projects were started a decade ago to determine and to imitate 5′-CCGG-3′ DNA methylation patterns found for the E1A promoter of Ad12 DNA or for the E2AL promoter of Ad2 DNA in the active or in the inactive configuration in several Ad12- or Ad2-transformed hamster cell lines (Sutter & Doerfler 1979, 1980; Vardimon *et al.* 1980, 1982; Kruczek & Doerfler 1982). The results of *in vitro* methylation studies showed that it was sufficient to methylate three 5′-CCGG-3′ sequences in the 5′ located and promoter regions of the E2AL promoter and leave eleven of these 3′ located sequences unmethylated and still effect inactivation or severe inhibition of the E2AL promoter. This observation was documented in transient expression assays in *Xenopus laevis* oocytes (Langner *et al.* 1984) or in mammalian cells (Langner *et al.* 1986), in an *in vitro* cell-free transcription system (Dobrzanski *et al.* 1988), or after genomic fixation of the E2AL promoter–indicator gene construct in mammalian cells (Müller & Doerfler 1987). Similar, though less extensive, experiments were performed with the E1A promoter of Ad12 DNA (Kruczek & Doerfler 1982, 1983). These results demonstrated that sequence-specific methylations of eukaryotic promoters caused their inactivation, at least in the test systems employed. At present, it cannot be predicted for a given promoter which one of the methylated cytidine residues effects its inactivation. The sensitive sites will have to be experimentally determined for each promoter.

It is an interesting and unresolved problem of how sequence-specific methylations lead to promoter inactivation or inhibition. For promoter activity, the interaction of several proteins with specific sequence-motifs in the promoter are required. In addition, the promoter may have to assume a highly specific conformation. Hence, it is essential to investigate the modulation of these interactions by the presence of a 5-mC group in each of these motifs. As has already been pointed out, 5-mC may have a positive or a negative effect, or no influence at all, on the promoter binding of positively or negatively regulating proteins. Also, the exerted effects need not be all or none, but may be responsible for subtle gradations of promoter activity. Hence,





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

the occurrence of a 5-mC residue in a promoter sequence may have several functional implications, depending on the site and on the type of promoter.

Obviously, the gene inactivating or inhibiting functions of promoter methylations have to be viewed in the wider context of all the factors contributing to the regulation of promoter activity and as such, of gene expression. It is also apparent from the influence of DNA methylation on the regulation of gene expression that promoter methylation can strongly influence or perhaps have dominant effects on several important processes, such as development and differentiation, the immune response or the alteration of gene regulation in malignant cells.

## 5. Methylation of promoter motifs and the binding of specific proteins

One possible explanation for the inhibitory effect of three 5-mC residues in a promoter sequence was the positive or negative modulation of the sequence-specific binding of proteins. A synthetic oligonucleotide of 50 b.p. or a restriction fragment of 73 b.p., which both the +24 and +6 5′-CCGG-3′ sequences of the late E2A promoter, were methylated at these two sites or were left unmethylated and were then incubated with a partially purified nuclear extract from human HeLa cells. Protein binding was detected by retardation of electrophoretic migration of the $^{32}$P-labelled oligonucleotide or the restriction fragment on polyacrylamide gels. The formation of one of the DNA–protein complexes was abrogated when the 5′-CCGG-3′ methylated oligonucleotide or fragment was used in the binding assays. The same complex could also be destroyed by adding the same (non-methylated) oligonucleotide as competitor to the reaction mixture. The results showed that the *in vitro* methylation of sequences in the E2AL promoter not only inhibited promoter function, but also prevented the binding of specific proteins (Hermann *et al.* 1989). In contrast, when a 377 bp fragment of the late E2A promoter was used for studies on the binding of specific nuclear proteins, no difference was detectable between the unmethylated and the 5′-CCGG-3′ methylated promoter (Hoeveler & Doerfler 1987). Thus the results on protein binding with short fragments might not truly reflect the full complexity of possible DNA–protein interactions in the E2AL promoter region.

The effect (or lack of it) that methylation in promoter motifs has on the binding of proteins has been examined in other eukaryotic promoters. In the globin promoter, interference with protein binding resulting from DNA methylation was not observed (G. Felsenfeld, cited by Keshet *et al.* 1986). In the early E2A promoter of Ad2 DNA, the methylation of one *Hha*I site abolished protein binding and promoter function (Kovesdi *et al.* 1987). In contrast, the methylation of an Sp1 binding site in several eukaryotic promoters did not affect protein binding (Harrington *et al.* 1988; Höller *et al.* 1988). Methylation of the major late promoter of Ad2 DNA prevents the binding of specific proteins (Watt & Molloy 1988; Molloy & Watt, this symposium). Extensive further work on well analysed promoters will be required, before the code governing DNA–protein interactions and the specific interference by methylated nucleosides can be deciphered.

## 6. The reversal of the inhibitory effect of promoter methylations

One of the requirements for a genetic signal, such as DNA methylation, that has been shown to involve the long-term inactivation of a promoter, is its reversibility. Demethylation by DNA replication and concomitant suppression of the mechanisms responsible for maintenance




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

methylation is one formal possibility in altering an existing pattern of methylation. This pathway of demethylation is dependent upon cell replication. Alternatively, pathways leading to transient demethylation, for example, a type of repair function in which a methylated nucleotide would be excised and replaced by its non-methylated homologue was proposed (Razin *et al.* 1986) and remains to be proven.

For the reversal of the inactivation of the methylated E2AL promoter of Ad2 DNA, we also considered more subtle processes that were capable of reactivating the promoter, although it had remained fully methylated. The E1A adenovirus gene-encoded 289 amino acid protein (Langner *et al.* 1986; Weisshaar *et al.* 1988; Knust *et al.* 1989), a classical transactivator (for review, see Nevins *et al.* (1987)) and a strong enhancer from an immediate early gene in the human cytomegalovirus (Knebel-Mörsdorf *et al.* 1988) were both shown to reactivate the methylated E2AL promoter (table 2).

Table 2. Reversal of methylation-inhibition of late E2A promoter

1. demethylation, for example, by 5-azacytidine (Knust *et al.* 1989)
2. E1A transactivating 289 amino acid protein (Ad2 DNA) (Langner *et al.* 1986; Weisshaar *et al.* 1988; Knust *et al.* 1989)
3. strong enhancer (human cytomegalovirus) (Knebel-Mörsdorf *et al.* 1989)

The finding that promoter methylation did not involve inactivation unconditionally underscored current concepts on the high level of complexity in the regulation of promoter activity. The presence of 5-mC even in decisive and sensitive promoter sequences did not *eo ipso* guarantee promoter inactivation. As outlined above, a series of other pre-conditions had to be fulfilled to render promoter methylations consistent with inactivation. Conversely, and as pointed out previously (Kuhlmann & Doerfler 1982), absence of promoter methylation at the decisive sites was only one of the pre-conditions for promoter functionality.

## 7. N-6-methyl-deoxyadenosine in mammalian DNA?

There is at present no evidence for the occurrence of *N*-6-methyl-deoxyadenosine (*N*-6-mA) in mammalian DNA. For that reason, it is all the more remarkable that the enzymatic methylation of two adenosine residues in the thymidine kinase gene of herpes simplex virus (Waechter & Baserga 1982) or the introduction of one *N*-6-mA group into the E1A promoter of Ad12 DNA (Knebel & Doerfler 1986) leads to the inactivation or the inhibition of these promoters. Apparently, factors in mammalian cells are capable of recognizing an *N*-6-mA residue and interfering with promoter function. Because of these data, it will be necessary to develop more sensitive techniques, before accepting the commonly held idea that *N*-6-mA does not occur in mammalian DNA.

## 8. Effects of foreign DNA integration on the methylation of flanking cellular sequences

In mammalian cells, the apparently non-methylated DNA of human adenovirions (Günthert *et al.* 1976) was subject to *de novo* methylation after integration into the host hamster genome (Sutter *et al.* 1978). The establishment of these specific patterns of methylation in the integrated adenovirus sequences (Sutter & Doerfler 1979, 1980) required a considerable number of cell




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

divisions after integration of foreign DNA (Kuhlmann & Doerfler 1982, 1983). Recently, we reported on the results of an analysis of the site of linkage between the left terminus of Ad12 DNA and unique hamster DNA in the Ad12-induced tumour T1111(2) (Lichtenberg *et al.* 1987). In what way, if any, was the methylation pattern of the adjacent cellular DNA affected by this insertion of unmethylated foreign (Ad12) DNA? In normal hamster kidney and spleen DNA and in several Ad12-transformed hamster cell lines, this pre-insertion sequence was completely methylated at the 5′-CCGG-3′ (*Hpa*II) and 5′-GCGC-3′ (*Hha*I) sequences. The same pre-insertion sequences in the DNA of cell line BHK21 and on the non-occupied chromosome in the tumour cell line H1111(2) were almost completely methylated. In contrast, the same sequence on the chromosome, that carried the integrated Ad12 DNA sequence in the tumour T1111(2), was unmethylated at the 5′-CCGG-3′ and 5′-GCGC-3′ sequences, as were the abutting Ad12 DNA sequences. Thus the insertion of unmethylated foreign DNA could result in the hypomethylation of the flanking cellular DNA in the target sequences (Lichtenberg *et al.* 1988).

In contrast, for the proviral genome of mouse murine leukaemia virus it was shown that its insertion into the cellular genome involved an enhancement of DNA methylation in the flanking cellular sequences (Jähner & Jaenisch 1985). Effects on the methylation of neighbouring cellular sequences might depend on the inserted foreign DNA, on the site of insertion or on a combination of these parameters.

## 9. What other biological functions could be affected by DNA methylation?

Many possibilities for research exist in this context. One project has been initiated in our laboratory and will be briefly described. We developed a cell-free system from nuclear extracts of BHK21 cells to study the *in vitro* recombination between pre-insertion sites from hamster DNA and Ad12 DNA (Doerfler *et al.* 1987; Jessberger *et al.* 1989). These pre-insertion sequences had previously served as sites for the integration of Ad12 DNA, as they had been analysed in the DNA from Ad12-induced tumour cells (Stabel & Doerfler 1982; Lichtenberg *et al.* 1987). More than 100 recombinants between two different hamster pre-insertion sequences and Ad12 DNA were isolated and partly characterized. At comparable frequencies, randomly cloned hamster DNA sequences did not recombine with Ad12 DNA in this system (Jessberger *et al.* 1989). Hamster pre-insertion sequences were chosen (Doerfler *et al.* 1987), because these sequences had previously recombined with Ad12 DNA and gave rise to Ad12-induced tumour cells. Apparently, these sequences fulfilled some of the pre-conditions for a recombinationally active DNA segment. We have set out to study whether the methylation of these hamster or of the Ad12 DNA sequences will influence the recombination reaction.

## 10. Patterns of methylation as markers for genetic polymorphisms in human populations

Once a certain pattern of DNA methylation is established, it remains very stable, at least in cell lines that are carried in culture (Sutter *et al.* 1978; Sutter & Doerfler 1980; Kuhlmann & Doerfler 1983). How permanent and by what factors are patterns of DNA methylation determined in different segments of the human chromosomes? There is evidence that patterns




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

of DNA methylation in human populations may be genetically determined and inherited, at least for a limited set of chromosomal DNA segments (Silva & White 1988).

We have started to use randomly selected, chromosomally assigned probes of human DNA, which have been derived from human chromosomal DNA libraries, on DNA from human lymphocytes. A number of families have been identified and individuals have been screened for family background. These families had been living in a remote area of Europe with intermarriages restricted to a limited geographical region. The patterns of methylation in three to four generations of these families in a series of DNA segments located on different human chromosomes are being determined with a number of methylation-sensitive restriction endonucleases. In the selection of these families, particular care and emphasis was placed on the inclusion of families who had been followed for some time for the familiar accumulation of certain human malignancies. Families free of malignancies, or at least of their clustering, were chosen as controls.

I am indebted to Petra Böhm for her expert editorial work. Research in the author's laboratory has been supported over many years by the Deutsche Forschungsgemeinschaft through SFB74-C1; the Universität zu Köln; the State Ministry for Science and Research Nordrhein-Westfalen and the Fonds der Chemischen Industrie.

## References

Church, G. M. & Gilbert, W. 1984 Genomic sequencing. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991–1995.

Dobrzanski, P., Hoeveler, A. & Doerfler, W. 1988 Inactivation by sequence-specific methylations of adenovirus promoters in a cell-free transcription system. *J. Virol.* **62**, 3941–3946.

Doerfler, W. 1981 DNA methylation – a regulatory signal in eukaryotic gene expression. *J. gen Virol.* **57**, 1–20.

Doerfler, W. 1982 Uptake, fixation, and expression of foreign DNA in mammalian cells: the organization of integrated adenovirus DNA sequences. *Curr. Top. Microbiol. Immun.* **101**, 127–194.

Doerfler, W. 1983 DNA methylation and gene activity. *A. Rev. Biochem.* **52**, 93–124.

Doerfler, W. 1989 Complexities in gene regulation by promoter methylation. In *Nucleic acids and molecular biology* (ed. F. Eckstein & D. M. Lilley), vol. 3, pp. 92–119, Berlin, Heidelberg, New York, London, Paris and Tokyo: Springer Verlag.

Doerfler, W., Langner, K.-D., Knebel, D., Weyer, U., Dobrzanski, P. & Knust-Kron, B. 1985 Site-specific promoter methylations and gene inactivation. In *Biochemistry and biology of DNA methylation* (ed. G. L. Cantoni & A. Razin), pp. 133–155. New York: Alan R. Liss Inc.

Doerfler, W., Spies, A., Jessberger, R., Lichtenberg, U., Zock, C. & Rosahl, T. 1987 Recombination of foreign (viral) DNA with the host genome. Studies *in vivo* and in a cell free system. In *38. Colloquium Mosbach 1987. Molecular basis of viral and microbial pathogenesis* (ed. R. Rott & W. Goebel), pp. 60–72. Berlin and Heidelberg: Springer-Verlag.

Doerfler, W., Langner, K.-D., Knebel, D., Hoeveler, A., Müller, U., Lichtenberg, U., Weisshaar, B. & Renz, D. 1988 Eukaryotic gene inactivation by sequence-specific promoter methylation and the release of the transcription block. In *Architecture of eukaryotic genes* (ed. G. Kahl), pp. 409–417, Weinheim: VCH.

Günthert, U., Schweiger, M., Stupp, M. & Doerfler, W. 1976 DNA methylation in adenovirus, adenovirus-transformed cells, and host cells. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3923–3927.

Harrington, M. A., Jones, P. A., Imagawa, M. & Karin, M. 1988 Cytosine methylation does not affect binding of transcription factor Spl. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2066–2070.

Hermann, R., Hoeveler, A. & Doerfler, W. 1989 Sequence-specific methylation in a downstream region of the late E2A promoter of adenovirus type 2 DNA prevents protein binding. *J. molec. Biol.* **210**.

Höller, M., Westin, G., Jiricny, J. & Schaffner, J. 1988 Spl transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Devel.* **2**, 1127–1135.

Hoeveler, A. & Doerfler, W. 1987 Specific factors binding to the E2A late promoter region of adenovirus type 2 DNA: no apparent effects of 5′-CCGG-3′ methylation. *DNA* **6**, 449–460.

Jähner, D. & Jaenisch, R. 1985 Retrovirus-induced *de novo* methylation of flanking host sequences correlates with gene inactivity. *Nature, Lond.* **315**, 594–597.





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Jessberger, R., Heuss, D. & Doerfler, W. 1989 Recombination in hamster cell nuclear extracts between adenovirus type 12 DNA and two hamster preinsertion sequences. *EMBO J.* **8**, 869–878.
Keshet, I., Lieman-Hurwitz, J. & Cedar, H. 1986 DNA methylation affects the formation of active chromatin. *Cell* **44**, 535–543.
Knebel, D., Lübbert, H. & Doerfler, W. 1985 The promoter of the late p10 gene in the insect nuclear polyhedrosis virus *Autographa californica*: activation by viral gene products and sensitivity to DNA methylation. *EMBO J.* **4**, 1301–1306.
Knebel, D. & Doerfler, W. 1986 *N*-6-Methyldeoxyadenosine residues at specific sites decrease the activity of the E1A promoter of adenovirus type 12 DNA. *J. molec. Biol.* **189**, 371–375.
Knebel-Mörsdorf, D., Achten, S., Langner, K.-D., Rüger, R., Fleckenstein, B. & Doerfler, W. 1988 Reactivation of the methylation-inhibited late E2A promoter of adenovirus type 2 DNA by a strong enhancer of human cytomegalovirus. *Virology* **166**, 166–174.
Knust, B., Brüggemann, U. & Doerfler, W. 1989 Reactivation of a methylation-silenced gene in adenovirus-transformed cells by 5-azacytidine or by E1A *trans* activation. *J. Virol.* **63**, 3519–3524.
Kovesdi, I., Reichel, R. & Nevins, J. R. 1987 Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2180–2184.
Kruczek, I. & Doerfler, W. 1982 The unmethylated state of the promoter/leader and 5′-regions of integrated adenovirus genes correlates with gene expression. *EMBO J.* **1**, 409–414.
Kruczek, I. & Doerfler, W. 1983 Expression of the chloramphenicol acetyltransferase gene in mammalian cells under the control of adenovirus type 12 promoters: Effect of promoter methylation on gene expression. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7586–7590.
Kuhlmann, I. & Doerfler, W. 1982 Shifts in the extent and patterns of DNA methylation upon explantation and subcultivation of adenovirus type 12-induced hamster tumor cells. *Virology* **118**, 169–180.
Kuhlmann, I., Achten, S., Rudolph, R. & Doerfler, W. 1982 Tumor induction by human adenovirus type 12 in hamsters: loss of the viral genome from adenovirus type 12-induced tumor cells is compatible with tumor formation. *EMBO J.* **1**, 79–86.
Kuhlmann, I. & Doerfler, W. 1983 Loss of viral genomes from hamster tumor cells and non-random alterations in patterns of methylation of integrated adenovirus type 12 DNA. *J. Virol.* **47**, 631–636.
Langner, K.-D., Vardimon, L., Renz, D. & Doerfler, W. 1984 DNA methylations of three 5′C-C-G-G3′ sites in the promoter and 5′ region inactivate the E2a gene of adenovirus type 2. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2950–2954.
Langner, K.-D., Weyer, U. & Doerfler, W. 1986 *Trans* effect of the E1 region of adenoviruses on the expression of a prokaryotic gene in mammalian cells: Resistance to 5′-CCGG-3′ methylation. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1598–1602.
Lichtenberg, U., Zock, C. & Doerfler, W. 1987 Insertion of adenovirus type 12 DNA in the vicinity of an intracisternal A particle genome in Syrian hamster tumor cells. *J. Virol.* **61**, 2719–2726.
Lichtenberg, U., Zock, C. & Doerfler, W. 1988 Integration of foreign DNA into mammalian genome can be associated with hypomethylation at site of insertion. *Virus Res.* **11**, 335–342.
Müller, U. & Doerfler, W. 1987 Fixation of the unmethylated or the 5′-CCGG-3′ methylated adenovirus late E2A promoter-*cat* gene construct in the genome of hamster cells: gene expression and stability of methylation patterns. *J. Virol.* **61**, 3710–3720.
Nevins, J. R. 1987 Regulation of early adenovirus gene expression. *Microbiol. Rev.* **51**, 419–430.
Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. & Cantoni, G. L. 1986 Replacement of 5-methylcytosine by cytosine: a possible mechanism for transient DNA demethylation during differentiation. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2827–2831.
Reik, W., Collick, A., Norris, M. L., Barton, S. C. & Surani, M. A. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature, Lond.* **328**, 248–251.
Silva, A. J. & White, R. 1988 Inheritance of allelic blueprints for methylation patterns. *Cell.* **54**, 145–152.
Stabel, S. & Doerfler, W. 1982 Nucleotide sequence at the site of junction between adenovirus type 12 DNA and repetitive hamster cell DNA in transformed cell line CLAC1. *Nucl. Acids Res.* **10**, 8007–8023.
Sutter, D., Westphal, M. & Doerfler, W. 1978 Patterns of integration of viral DNA sequences in the genomes of adenovirus type 12-transformed hamster cells. *Cell.* **14**, 569–585.
Sutter, D. & Doerfler, W. 1979 Methylation of integrated viral DNA sequences in hamster cells transformed by adenovirus 12. *Cold Spring Harb. Symp. quant. Biol.* **44**, 565–568.
Sutter, D. & Doerfler, W. 1980 Methylation of integrated adenovirus type 12 DNA sequences in transformed cells is inversely correlated with viral gene expression. *Proc. natn. Acad. Sci. U.S.A.* **77**, 253–256.
Toth, M., Lichtenberg, U. and Doerfler, W. 1989 Genomic sequencing reveals a 5-methylcytosine-free domain in active promoters and the spreading of preimposed methylation patterns. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3728–3732.
Vardimon, L., Neumann, R., Kuhlmann, I., Sutter, D. & Doerfler, W. 1980 DNA methylation and viral gene expression in adenovirus-transformed and -infected cells. *Nucl. Acids. Res.* **8**, 2461–2473.
Vardimon, L. & Doerfler, W. 1981 Patterns of integration of viral DNA in adenovirus type 2-transformed hamster cells. *J. molec. Biol.* **147**, 227–246.




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M. & Doerfler, W. 1982 Expression of a cloned adenovirus gene is inhibited by *in vitro* methylation. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1073–1077.

Waechter, D. E. & Baserga, R. 1982 Effect of methylation on expression of microinjected genes. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1106–1110.

Watt, F. & Molloy, P. L. 1988 Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus late promoter. *Genes Devel.* **2**, 1136–1143.

Weisshaar, B., Langner, K.-D., Jüttermann, R., Müller, U., Zock, C., Klimkait, T. & Doerfler, W. 1988 Reactivation of the methylation-inactivated late E2A promoter of adenovirus type 2 by E1A (13S) functions. *J. molec. Biol.* **202**, 255–270.





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES