METHYLATION AND CHROMATIN CONFORMATION OF ADENOVIRUS TYPE 12 DNA SEQUENCES IN TRANSFORMED CELLS

By

© ROSEMARY JEANNE REDFIELD, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

September, 1980

VIRAL DNA METHYLATION IN

AD 12-TRANSFORMED CELLS

MASTER OF SCIENCE (1980) McMASTER UNIVERSITY

(Biochemistry) Hamilton, Ontario

TITLE: Methylation and Chromatin Conformation of Adenovirus Type 12 DNA Sequences in Transformed Cells.

AUTHOR: Rosemary Jeanne Redfield, B.Sc. (Monash University, Australia)

SUPERVISOR: Professor S. Mak

NUMBER OF PAGES: ix, 98.

ABSTRACT

The DNA blot hybridization technique was used to study the integrated adenovirus type 12 (Ad 12) DNA sequences in three cell lines established by transformation of baby rat kidney cells with the leftmost 16.5% of the Ad 12 genome. The cell lines 702-C1, 702-C2, and 702-C3 contained respectively 3 - 4, 9, and 4 copies of the DNA used for transformation, integrated at 1, 9, and 4 distinct sites in cellular DNA. Partial copies of the transforming fragment were detected at additional integration sites. Methylation of CCGG sequences at the viral integration sites was investigated, using the methylation-sensitive restriction endonuclease Hpa II and its methylation-insensitive isoschizomer Msp I. In lines 702-C2 and 702-C3 some integrated copies of the viral DNA were methylated at all CCGG sequences while other copies were methylated at only some of these sequences. In line 702-C1 no viral CCGG sequences were methylated. Limited DNase I digestions of transformed cell nuclei were used to examine the chromatin conformation of the methylated and unmethylated viral DNA. A correlation was found between the absence of methylation at viral CCGG sequences in lines 702-C2 and 702-C3 and preferential cleavages by DNase I of the chromatin around these sequences. Results reported here suggest that these preferential DNase I cleavages occur at sites which are hypersensitive to DNase I. If so, the cleavages may be unrelated to the DNase I-sensitive nucleosome structure which has been associated with transcriptionally active genes.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. S. Mak, for his advice and assistance, Dr. S. Bayley and Dr. J. Smiley for critically reading this manuscript, and Mr. Sheldon Girvitz for help with the photography.

I gratefully acknowledge the financial support of the Medical Research Council of Canada.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS (cont'd)

LIST OF FIGURES

LIST OF FIGURES (cont'd)

INTRODUCTION

I. Cellular Differentiation

The processes of cellular differentiation in higher eukaryotes give rise to populations of cells which show distinctive patterns of gene expression. In each differentiated cell type a specific subset of the genome is expressed as proteins, while the proteins associated with other cell types are absent. These altered patterns of gene expression are stable, in that they persist in the absence of the stimuli which gave rise to them, and heritable, in that the altered phenotype persists in the descendants of the cell which originally underwent differentiation.

The mechanisms which bring about and then maintain the specific patterns of gene expression are not known. There is abundant evidence that this does not result from selective degradation of mRNAs or proteins, but from altered patterns of transcription. Nor does differentiation result from deletion of those genes which are not expressed in the differentiated cells. Evidence for this comes from nuclear transplantation experiments which showed that nuclei of differentiated Xenopus cells, when placed into Xenopus oocytes, allow some of the oocytes to differentiate normally into mature frogs (Gurdon and Woodland, 1968). The presence of inactive genes in differentiated tissues has also been demonstrated directly by nucleic acid hybridization experiments. Globin genes are present in oviduct DNA (Weintraub and Groudine, 1976) and ovalbumin genes in

erythrocyte DNA (Garel and-Axel, 1976).

Differentiated cells might inactivate genes which they do not need by minor changes in base sequences (i.e. at control regions), or by gene rearrangements. Gurdon's experiments suggest that this is not the case, but such specific changes might be reversible, and proof must await the sequencing of genes and control regions cloned from differentiated cells.

Direct alterations of DNA base sequence would ensure stable inheritance of differentiated phenotypes. However, prokaryotes modulate gene expression through regulatory proteins which bind to specific DNA sequences and prevent or permit transcription. Similar sequence-specific interactions between DNA and protein may directly control the expression of vertebrate genes; as yet none have been identified.

In eucaryotic nuclei the DNA is compacted with proteins into the ordered structures of chromatin. The organization of chromatin is extremely complex, and is understood only at the most basic level. The chromatin structure of eucaryotic genomes is likely to play a major role in the control of gene expression in differentiated cells.

II. A. Chromatin

At the simplest level of chromatin structure, eukaryote DNA is organized into regular repeating nucleoprotein subnits called nucleosomes. Each nucleosome contains about 200 base pairs of DNA; 140 base pairs of this is tightly coiled around a core histone octamer (two each of histones H2A, H2B, H3, and H4), while the remaining variable length of DNA, loosely bound to histone H1, acts as a linker between adjacent cores (reviewed by

Felsenfeld, 1978). The overall nucleosome repeat length can be both species-specific and tissue-specific. Histone variants have been identified, as have post-synthetic histone modifications such as acetylation and phosphorylation. Although some of these show tissue specificity, in general the histones appear too invariant to be responsible for tissue-specific patterns of gene expression.

There are many non-histone chromosomal proteins; high-resolution analysis of HeLa chromatin revealed 450 distinct polypeptides (Peterson and McConkey, 1976). The functions of these proteins are not known. Some may play a role in gene regulation, and others must be involved in the higher-order structures of chromatin.

The nucleosome is only the first level of chromatin organization. Linearly compacted nucleosomes may form the 100Å-wide "thin filaments" (Worcel, 1977), and solenoidal supercoils of these may make up the 300Å "thick fibres" (Carpenter et al, 1976). The thick fibre is the major form of chromatin found in interphase nuclei, and probably represents the native conformation of inactive DNA sequences (Chambon, 1977).

The chromatin fibre is known to be further organized into loops or domains. Each loop contains 50,000 to 100,000 base pairs of DNA. When histone-depleted mammalian chromosomes are examined by electron microscopy, all the DNA is seen projecting as radial loops from a protein chromosome scaffold (Laemmli et al, 1977). Laemmli has proposed a model for the organization of mammalian chromosomes, in which a chromosomal scaffold organizes the DNA into loops, and histones compact the loops of DNA into nucleosomes, filaments, and fibres, with the degree of compaction of the DNA (and perhaps of the scaffold) depending on the transcriptional status

of the sequences in the loop, and on the stage of the cell cycle.

II. B. Chromatin Structure of Active Genes

Nucleosomes or similar structures have been detected in association with transcriptionally active chromatin both by electron microscopic visualization and by their characteristic pattern of nuclease sensitivity. This has been demonstrated for both transcribed sequences in general and for the globin and ovalbumin genes in particular (Lacy and Axel, 1975; Gottesfeld and Melton, 1978). There is some E.M. evidence that very active transcription units, in particular the rRNA genes, may have more widely spaced nucleosomes or may lack them altogether (Franke and Scheer, 1978). The apparent absence of a beaded morphology during transcription may result from an opening out of the nucleosome octamer, with the histones remaining bound to the DNA and causing it to retain the patterns of nuclease sensitivity which are characteristic of nucleosomes (Weintraub et al, 1977).

Attempts have been made to fractionate chromatin preparations into transcriptionally active and inactive components. Fractionation of chromatin after digestion with micrococcal nuclease (Bloom and Anderson, 1978; Levy-Wilson and Dixon, 1979) or DNase II (Gottesfeld, 1978) can give up to 7-fold enrichment in sequences complementary to cytoplasmic polyadenylated RNAs and to tissue-specific mRNAs. These procedures have the advantage that the active chromatin fraction can be recovered for further analysis, but few results have been obtained to date.

The majority of the information on the chromatin structure of active genes has come from studies which utilize the sensitivity of active

chromatin to digestion by DNase I. In the original experiment, (Weintraub and Groudine, 1976) chick erythrocyte nuclei were incubated with pancreatic deoxyribonuclease (DNase I) until 10 - 20% of the DNA was degraded to acid soluble fragments. The remaining DNA was purified, and analysed for globin sequences by solution hybridization to a globin cDNA probe. The DNase I-resistant DNA was found to be depleted in globin sequences. The globin gene chromatin did not show this DNase I sensitivity when erythrocyte nuclei were replaced with nuclei from fibroblasts or brain cells, in which the globin genes are not transcribed. Ovalbumin chromatin was similarly resistant in erythrocyte nuclei. Other experiments have confirmed that preferential DNase I digestion is a general feature of transcriptionally active chromatin, independent of the frequency of transcription; both extremely active and rarely transcribed genes appear to have similar nuclease sensitivities (Weintraub and Groudine, 1976).

Miller et al (1978) studied the DNase I sensitivity of globin genes in mouse erythroleukemia cells (Friend cells), which can be induced to produce large amounts of globin mRNA. The globin genes were in a highly DNase I-sensitive chromatin conformation in uninduced as well as in induced cells, in a non-inducible cell line, and in a lymphoblast cell line. This result suggests that the DNase I-sensitive conformation of the globin genes is acquired early in erythroid differentiation and is indicative of the potential for transcription, while the actual rate of transcription is dependent on subsequent differentiation steps. This conclusion is supported by experiments by Weintraub and Groudine (1976) on globin genes in mature erythrocytes, and by Palmiter et al (1977) on ovalbumin sequences in the hormone-withdrawn oviduct. In both cases transcription of the gene under

study has ceased, but the once-active sequences retain the DNase I-sensitive chromatin conformation.

Inactive genes and non-coding DNA sequences may normally be compacted into supra-nucleosomal structures and consequently shielded from the action of nucleases. It might be thought that the basis for the preferential DNase I digestion of active sequences was the absence of higher-order packaging. However, when nucleosome monomers prepared from erythrocyte chromatin are assayed by limited DNase I digestion, globin-containing nucleosomes retain their relative DNase I sensitivity in comparison to the bulk of nucleosomal DNA (Weintraub and Groudine, 1976). This indicates that the DNase I-sensitive conformation is a property of the nucleosomes and not of spacer region or higher-order structures.

Further experiments (Weisbrod and Weintraub, 1979) have shown that while the DNase I-sensitive conformation is dependent on the presence of elutable non-histone proteins (high mobility group (HMG) 14 and 17), the tissue specificity does not reside with the elutable proteins but with the depleted nucleosome core. They concluded that the nucleosome cores of active chromatin must contain a unique feature (not HMG 14 or 17) which distinguishes them from inactive chromatin.

The structure of each active nucleosome is not likely to be directly determined by the DNA sequence it contains, as this might severely restrict the DNA's coding potential. There are modifications of DNA which do not interfere with coding, and which might be important in determining tissue-specific patterns of gene expression. The most interesting of these modifications is cytosine methylation, which is discussed in the following section.

III. DNA Methylation

The modified base 5-methylcytosine is present as a minor component in the DNA of higher organisms. In vertebrates 2 - 4% of the total cytosines are methylated (Vanyushin et al, 1973); in plants up to 30% of the cytosines may be modified in this way (Pivec et al, 1974). Methylation takes place at the polymer level, and is normally complete within a few minutes after incorporation of cytosine into newly synthesized DNA (Kappler, 1970). Once established, cytosine methylation is stable; there is no evidence that methyl groups are ever removed from DNA (Burdon and Adams, 1969).

Almost all the methylated cytosines occur at the 5' side of guanidine (5'...^{m}CpG...3') (Doskocil and Sorm, 1962; Grippo et al, 1968). This CpG dinucleotide is under-represented in vertebrate and other methylated DNAs. For example, in human DNA CpG doublets occur at a frequency of 0.01 (Swartz et al, 1962) rather than the 0.04 predicted from the base composition of human DNA (40% G+C). This deficiency is thought to be due to mutational loss of the ^{m}CpG dinucleotide by deamination of 5-methylcytosine to thymidine (Bird, 1980). The proportion of the CpG dinucleotides which are methylated can be estimated from the nearest neighbour and base composition data. Although in vertebrate tissues only a few percent of cytosines are methylated, the restriction of this methylation to the scarce CpG dinucleotide means that between 40% and 80% of the CpG dinucleotides are probably methylated.

In bacterial restriction/modification systems, bacterial DNA is protected from restriction endonuclease cleavage by methylation of a

specific base (6-methyladenine or 5-methylcytosine) in the DNA sequence recognized by the endonuclease. A fortuitous consequence of this is that a number of restriction endonucleases which have the CpG dinucleotide in their recognition sequence ("CpG enzymes") are unable to cleave methylated sites in animal DNAs. This property can be exploited in the investigation of a) overall levels and arrangements of CpG methylation, and b) the extent of CpG methylation at specific restriction endonucleases cleavage sites in and around selected genes.

The most useful of these CpG enzymes are Hpa II and its isoschizomer Msp I. Both enzymes recognize and cleave at the same DNA sequence, CCGG; however Hpa II will not cleave the DNA if the internal cytosine is methylated ($C^{m}CGG$). Cleavage by Msp I is not inhibited by methylation of this CpG doublet (Waalwijk and Flavell, 1978a). Comparison of patterns produced by the two enzymes in parallel digestions will demonstrate the presence or absence of methylation at CCGG sites in the DNA under investigation.

Methyl groups are not randomly distributed in the genome; satellite DNA and heterochromatic regions are heavily methylated (Gautier et al, 1977; Kahana et al, 1977). Using Hpa II, Msp I, and other CpG enzymes, Bird et al (1979) showed that in Echinus DNA methylation of CpG sites occurs mainly in long tracts (15 - 50 kb or more) in which all the recognition sites for three CpG enzymes are blocked. These tracts take up 40% - 50% of the Echinus genome. Vertebrate DNA is more highly methylated than Echinus, and the methylated stretches apparently occupy a higher proportion (about 70%) of the vertebrate genome (Bird and Taggart, 1980).

The study of unique genes in higher eucaryote genomes has been hindered by the low specific concentrations of these sequences in cellular

DNA (about one part in 10^6 for an average mammalian gene). The blot-hybridization method developed by Southern (1975) allows the restriction endonuclease digestion patterns of individual genes to be studied. In this technique, total cellular DNA is digested with a restriction endonuclease and the fragments are separated by agarose gel electrophoresis. The gel is then soaked in alkali to denature the DNA, neutralized, and the DNA is transferred directly onto a nitrocellulose sheet by blotting. After immobilization of the DNA on the sheet, specific cellular DNA fragments can be located by hybridization of the bound DNA to a high-specific activity radioactive DNA or RNA probe for the sequences of interest, followed by autoradiography of the sheet.

The patterns of CpG methylation of sites in and around specific genes have been determined, using methylation-sensitive CpG enzymes and the blot-hybridization technique described above. These studies have shown that DNA methylation patterns can be both sequence specific and tissue specific.

McGhee and Ginder (1979) compared the Hpa II and Msp I patterns of chicken adult β-globin sequences in various tissues. Some CCGG sites were found to be completely methylated in all tissues. Other sites were unmethylated in erythrocytes and reticulocytes (which express these sequences), but partially or completely methylated in cells not expressing globin genes. In a similar study, Waalwijk and Flavell (1978b) reported tissue-specific differences in the extent of methylation at a CCGG site in the rabbit β-globin intron.

Bird and Southern (1978) analysed the methylation pattern of Xenopus genes which encode rRNA. The amplified rDNA genes found in Xenopus oocytes

are not methylated (Dawid et al, 1970). In erythrocytes nearly all CpG dinucleotides in rDNA were found to be methylated; one site was identified which was 30 - 60 times less methylated than the others (Bird and Southern, 1978). In their study of the methylated and non-methylated compartments of Echinus, Bird et al (1979) found that there was no methylation of the active histone genes, ribosomal RNA genes, and 5S RNA genes in any tissues.

These observations suggest that DNA methylation may be related to gene activity. Riggs (1975) and Holliday and Pugh (1975) have proposed elegant mechanisms whereby DNA methylation patterns may be stably transmitted from parental cells to progeny cells. If so, DNA methylation may be an important factor in the initiation and maintenance of differentiated patterns of gene expression.

Both models invoke two cellular methyltransferase activities. The first class of enzymes would be responsible for the initial methylation of previously unmethylated CpG sites, giving rise to altered patterns of transcription. Enzymes of this class would be under strict developmental regulation, and might have precise sequence specificities.

Enzymes of the second class would be responsible for maintenance of methylation patterns in subsequent generations of cells, as follows. The major methyl-acceptor sequence is a palindrome. If both strands of the palindrome are methylated, ($^{5'...\overset{*}{C}pG...3'}_{3'...G\underset{*}{p}C...5'}$), DNA replication will produce two molecules, each methylated on only one strand of the methyl-acceptor sequence. The second class of methyltransferases is postulated to recognize all such half-methylated CpG sequences, and to complete the methylation of each daughter strand. This will restore the parental DNA

methylation pattern in the progeny cells. In this way the complex and subtle controls required for the establishment of the differentiated methylation pattern would not be necessary for maintenance and inheritance of the pattern.

Enzymes of the first class need act only transiently in development, and might be expected to be rare. None have been identified to date. Enzymes of the second class are predicted to be ubiquitous in all dividing cells, and cellular methyltransferases with the substrate specificities expected of these maintenance enzymes have been characterized in vivo and in vitro. These enzymes rapidly methylate DNA after DNA replication (Kappler, 1970), and show high activity only on partly undermethylated double stranded DNAs, as predicted for an enzyme whose normal substrate is half-methylated DNA (Adams et al, 1979). Bird (1978) has shown a) that methylation is symmetric at the palindromic sites, and b) that only the newly-synthesized strands receive new methyl groups after DNA replication. All of these findings support the Riggs and Holliday and Pugh models of semi-conservative replication of the methylation pattern.

In summary, there are three major findings which, taken together, strongly support the hypothesis that DNA methylation plays an important role in the initiation and maintenance of differentiated patterns of gene expression.

1. Methylation patterns can be gene-specific.

2. Gene-specific methylation patterns can be tissue-specific.

3. Methylation patterns can be stably inherited, in the absence of the enzymes which originated them.

IV. Viral Genes in Transformed Cells

When viral DNA is introduced into cells under conditions which prevent the establishment of a lytic infection, viral DNA may become covalently integrated into the genome of viable cells. In such cells, expression of appropriate viral sequences will give rise, by poorly-understood mechanisms, to the abnormal properties characteristic of transformed cells. These cells can be selected out of a population of non-transformed cells by their less stringent growth requirements.

The regions of the viral genome responsible for transformation have been identified by transformation experiments using mutant virus strains or purified restriction endonuclease fragments of viral DNA, and by analysis of viral DNA and RNA sequences present in transformed cells. In all cases, only those regions of the genome which are expressed early in lytic infection (before replication of the viral DNA) are required for transformation. In the adenoviruses only the leftmost of the early transcription units (early region 1) is consistently present and expressed in transformed cells (Flint et al, 1976; Mak et al, 1979), and DNA fragments containing the leftmost 7 - 12% of the genome contain sufficient information to bring about cell transformation (Graham et al, 1974; Shiroki et al, 1977).

Little is known about the mechanisms which lead to integration of the viral sequences. Integration may occur by illegitimate recombination or as a consequence of cellular repair mechanisms. Blot-hybridization analysis of viral sequences in transformed cells has not shed any light on the mechanism of integration; there is no evidence for specific or preferred integration sites (Ketner and Kelly, 1976; Sutter et al, 1978).

The viral DNA used for transformation may be present in the transformed cell in single or multiple copies, which may be separately or tandemly integrated at single or multiple sites in the cellular genome. When the viral DNA sequences in the transformed cells are compared with the viral DNA used in the transformation experiment, fragmented and internally rearranged forms are frequently detected (see reviews by Topp et al, 1980, and Flint, 1980).

Adenovirus Type 12 Transformed Cells:

Adenovirus type 12 (Ad 12) belongs to the highly oncogenic subgroup A adenoviruses. The leftmost 2.2 kb (6.8%) of the 33 kb Ad 12 genome is sufficient to induce oncogenic transformation of rodent cells (Shiroki et al, 1977). The early transcription units located in this region, (Subisaki et al, 1980) are expressed as mRNA in all Ad 12-transformed cells (Ortin et al, 1976; Mak et al, 1979). Cell lines which contain other early region sequences often express them, while late viral-specific mRNAs are only occasionally expressed even when the entire viral genome is integrated (Sutter et al, 1978).

The numbers of copies of the transforming DNA in cells transformed with Ad 12 virions and with Ad 12 DNA fragments have been determined by analysis of the reassociation kinetics of labelled viral DNA probes in the presence of transformed cell DNAs (Sharp et al, 1974). Virion-transformed hamster cells may contain 30 or more copies of some regions of the Ad 12 genome (Fanning and Doerfler, 1976). Rat cells transformed with the Eco RI-C fragment of Ad 12 DNA (leftmost 16%) contained between 4 and 12 copies

of most of the fragment per diploid amount of cell DNA (Sawada et al, 1979).

Additional results have been obtained by a complementary method. Analysis of the viral-specific restriction fragments of transformed cell DNA by gel electrophoresis and blot hybridization provides information about the arrangement of the viral sequences as well as giving an indication of the amount of viral DNA present in the cells. Ad 12 virion-transformed hamster cells analysed by Sutter et al (1978) contained large numbers of apparently intact viral genomes, integrated at multiple sites in the cellular DNA, as well as some fragmented genomes. The large amounts of viral-specific DNA present in individual high molecular weight fragments suggested that multiple copies of viral DNA might be present at each of several integration sites. The sensitivity of these experiments was not high enough to detect bands containing only single copies of parts of the Ad 12 genome. Sawada et al (1979) have reported on the viral integration patterns in cells transformed with the Ad 12 Eco-RI C fragment. These cells contain only a few copies of the transforming fragment, each apparently integrated at a separate site in the cellular DNA.

DNA extracted from tumor virus virions contains little or no 5-methylcytosine. The trace amounts that have been reported for Ad 12 viral DNA preparations (Gunthert et al, 1978) are almost certainly due to contaminating cellular DNA (Tjia et al, 1977).

Before this research was begun, there were indications that viral DNA sequences integrated in transformed cells might become methylated. In cells transformed by Herpesvirus saimiri, episomal viral DNA is heavily methylated in cell lines which do not produce virus (Desrosiers et al,

1979). In Ad 12-transformed hamster cells, Sutter et al (1978) found that much of the integrated viral DNA is resistant to cleavage by Hpa II and Sma I; both of these enzymes have CpG dinucleotides in their recognition sites and are blocked by cytosine methylation.

These observations suggest that in transformed cells viral DNA can be a substrate for cellular methylation processes. Characterization of the viral methylation patterns may provide information about the mechanisms and regulation of cellular DNA methylation. It may also be possible to relate the viral methylation patterns to the activity of viral genes, and thus derive information about the cellular functions of DNA methylation.

In this report I describe the patterns of integration and methylation of the Ad 12 DNA sequences present in three lines of cells which were transformed with the leftmost 5.5 kb (Eco RI-C fragment) of the Ad 12 genome. I have also investigated the chromatin conformations of the integrated viral DNA, using limited digestions of nuclei with DNase I.

MATERIALS AND METHODS

I. Cell Lines

Cell lines 702-C1, 702-C2, and 702-C3 are baby rat kidney cells which have been transformed with the Eco RI-C DNA fragment (leftmost 5.5 kb, 16.5% of the genome) of adenovirus type 12 (Ad 12) strain Huie (Mak et al, 1979). These cells were grown as monolayers in plastic dishes (Falcon, Lux) in Joklik's modified minimum essential medium supplemented with 5% horse serum and 1% calf serum. Cells were subcultured twice weekly by trypsinization.

II. Virus and Viral DNA

Adenovirus type 12 (Huie strain) was propagated in KB cells in suspension culture, and labelled by addition of ^{3}H thymidine (0.1 μCi/ml) at 12 hours post infection. Purification of the virus has been described (Green and Pina, 1963). Briefly, cells were pelleted 72 hours post infection, sonicated, and homogenized with Freon 113. The virus was concentrated and purified by banding in CsCl density gradients. Viral DNA was extracted from dialysed virions by Pronase digestion, three extractions with phenol, and exhaustive dialysis.

Ad 12 DNA (100 μg) was digested with Eco RI and the viral restriction fragments were isolated by electroelution of the digest from

a preparative 1% agarose gel using a homemade apparatus designed by F. L. Graham. Fractions containing viral DNA were identified by liquid scintilation counting, pooled, and concentrated by precipitating the DNA with two volumes of ethanol. The Eco RI-C fragment was labelled with ^{32}P as described below and used as a probe in blot hybridization experiments. This fragment was slightly contaminated (less than 5%) with the adjacent Eco RI-D fragment.

III. In Vitro Labelling of Viral DNA with ^{32}P

DNA was labelled by nick translation using the method of Maniatis (Weinstock et al, 1978). Two modified procedures were used.

1. Complete nick translation: 0.25 - 0.5 µg of DNA was incubated for 10 minutes at 37°C with 0.25 ng of DNase I, in 50 µl of 50 mM Tris pH 7.9, 5 mM $MgCl_2$, 10 mM β- mercaptoethanol, 100 µg/ml BSA, and 10 µM dATP, dTTP, and dGTP. The mixture was then combined with 100 µCi of ^{32}P α-dCTP (air dried to remove ethanol) (NEN, specific activity 300 - 600 Ci/mmol) and 6 units of E. coli DNA polymerase I, and incubated for 1 hour at 12°C. The reaction was stopped with EDTA (25 mM), protein was removed by phenol extraction, and the DNA was isolated by gel filtration through a 7 ml column of Sephadex G-50. Incorporation of label was usually 40 - 50%, giving a product with specific activity of 2 - 4 x 10^8 cpm/µg.

2. Limited nick translation: DNA labelled by the above method was extensively nicked and fragmented, and not suitable for restriction endonuclease

analysis. The following modifications of procedure I gave a product of high integrity but lower specific activity ($1 - 5 \times 10^7$ cpm/μg). The preincubation with DNase I was omitted. 1 - 2 μg of Ad 12 DNA was labelled with 10 - 20 μCi of ^{32}P α-dCTP.

IV. Isolation and Limited DNase I Digestion of Nuclei

Monolayers of transformed cells were prelabelled with ^3H thymidine (about 0.1 μCi/ml for 24 hours) to a specific activity of approximately 10^3 cpm/μg DNA, in order that the extent of digestion by DNase I could be determined. Confluent monolayers of cells were washed twice with 10 ml of cold phosphate buffered saline without Ca^{++} or Mg^{++} (PBS) and overlaid with 10 ml of cold lysis buffer (0.2% NP-40 in reticulocyte standard buffer plus calcium (RSB (Ca^{++}); 0.01M Tris, pH 7.4, 0.01M NaCl, 3 mM $MgCl_2$, 0.2 mM $CaCl_2$)). After 5 minutes, the lysis buffer was gently removed, replaced with 10 ml of RSB (Ca^{++}), and the nuclei were shaken loose. The nuclei were gently pelleted (700 rpm for 7 minutes at 0^oC) and resuspended in RSB (Ca^{++}) at a concentration equivalent to 0.5 - 1 μg DNA/ml (about 10^8 nuclei/ml).

Nuclei were warmed to 37^oC, pancreatic deoxyribonuclease (DNase I, Worthington) was added at 0 - 5 μg/ml, and the nuclei were digested for 5 minutes. Reactions were stopped by addition of EDTA to 10 mM. The extent of digestion of the nuclear DNA by DNase I was determined as follows: At the end of each digestion 2 aliquots (50 μl each) of the digest were taken into tubes containing 0.5 ml of cold trichloroacetic acid (TCA). One

sample was boiled for 10 minutes to liberate total DNA from the nuclei, and the other was pelleted to remove acid-insoluble DNA. The soluble radioactivity in each sample was determined by liquid scintillation counting. In these experiments, the release of acid-soluble radioactivity from control incubations without DNase I was not significantly higher than background. DNA was prepared from DNase I-digested and control nuclei using the method for cell DNA.

V. Preparation of DNA from Transformed Cells

Confluent monolayers of cells were washed twice with cold PBS, lysed with 10 mM Tris pH 7.4, 10 mM EDTA, 200 µg/ml Pronase, 0.4% SDS, and the lysate was incubated for 3 - 12 hours at 37°C. SDS was then increased to 1% and the protein was extracted twice with phenol, twice with phenol:chloroform:isoamyl alcohol (25:24:1), and twice with chloroform:isoamyl alcohol (24:1). Pancreatic ribonuclease (preincubated at 90°C for 10 minutes to inactivate DNases) was added to 100 µg/ml, RNA was digested for 30 minutes at 37°C, and the organic extractions were repeated. The final aqueous phase was dialysed against 5 changes of 10 mM Tris, 1 mM EDTA, pH 7.4, at 4°C for 48 hours. DNA preparations were concentrated (when necessary) either in dialysis bags using Aquacide II-A (Calbiochem), or by adjusting the NaCl to 0.15M and precipitating the DNA with two volumes of ethanol.

VI. Restriction Endonuclease Digestions and Agarose Gel Electrophoresis

Restriction endonucleases were obtained from Boehringer Mannheim, Miles, New England Biolabs, and Bethesda Research Laboratories. For each enzyme the digestion buffer recommended by the supplier was used. Digestions of cell DNA for analysis by blot hybridization usually were prepared with 12 or 22 µg of cell DNA in a volume of 24 - 88 µl, with 0.5 - 2 units of restriction endonuclease used per microgram of DNA.

Digests were analysed by electrophoresis in vertical or horizontal agarose slab gels, with a Tris-acetate buffer (40 mM Tris pH 7.8, 5 mM Na acetate, 1 mM EDTA). A thorough discussion of this technique is given by Southern (1980). Agarose concentration, electrophoresis voltage, and running time were varied to maximize resolution; specific conditions are given in the figure legends.

Controls for complete digestion:

It is difficult to determine whether a restriction enzyme digest of cellular DNA has gone to completion, particularly if the enzyme is one whose cleavage is blocked by cytosine methylation. In these experiments, complete digestions were essential to the analysis of integration and methylation patterns. Accordingly, the completeness of each cell DNA digest was checked before the digest was analysed by blot hybridization.

The transformed rat cell lines studied all showed one or more conspicuous repetitive-DNA bands after the DNA was digested with methylation-insensitive restriction enzymes and analysed in gels. Before blotting, all gels were stained with ethidium bromide and photographed under UV

illumination, and these repetitive-DNA bands were examined for the complete digestion pattern.

A parallel control digestion ensured that digestion conditions were sufficient for complete digestion of the cell DNA. Before each digestion mixture (DNA, enzyme and buffer) was incubated at 37°C an aliquot was removed (2 µg of DNA in 4 - 8 µl) and mixed with several nanograms of ^{32}P-labelled Ad 12 DNA (labelled by limited nick translation). This control digest was incubated in parallel with the main digest, and analysed by electrophoresis in a 1% agarose gel. Complete digestion of the ^{32}P-labelled viral DNA was determined for all digests by autoradiography of this control gel after drying. Complete digestion of the cellular DNA in the control gel was also checked by inspection of the ethidium bromide-stained repetitive-DNA bands before the gel was dried.

VII. Blot Hybridizations

DNA was transferred from gels to nitrocellulose sheets using a published modification of the "Southern blot" method (Southern, 1975; Wahl et al, 1979). Briefly, after electrophoresis agarose gels were soaked in 0.25 M HCl to partially depurinate the DNA, in 0.5 M NaOH + 1.0 M NaCl to denature and fragment the DNA, and in 0.5 M Tris pH 7.4 + 3 M NaCl to neutralize the gel. DNA was then transferred from the gel to a nitrocellulose sheet by blotting (usually overnight). The sheet was rinsed in 4 x SSC (SSC is 0.15 M NaCl, 0.015 M NaCitrate) and baked 2 hours at 80°C.

Hybridizations were also done using the modifications suggested by Wahl et al (1979). Nitrocellulose sheets were preincubated in sealed plastic bags with 5 x Denhardt's reagent (Denhardt's reagent is 0.02% (wt/vol) each of bovine serum albumin, polyvinyl pyrrolidone, and Ficoll (Mr 400,000)) and 250 µg/ml salmon sperm DNA to reduce non-specific binding of the probe to the nitrocellulose. Glycine was omitted from the pretreatment, as it is only required with the diazobenzyloxymethyl-paper blots described by Wahl et al (1979). Reagent grade formamide was purified with mixed-bed ion-exchange resin before use. Incubations used 1 - 3 x 10^6 cpm of probe (nick translated Eco RI-C fragment of Ad 12 DNA) per millilitre of hybridization solution (50% formamide), and were carried out at 42^{o}C for 16 - 40 hours. Dextran sulphate was included (10%) to increase the efficiency of hybridization.

After hybridization, sheets were washed at room temperature and at 50^{o}C (Wahl et al, 1979), and exposed to Kodak XR X-ray film using intensifying screens at -50^{o}C for 1 - 14 days. The intensities of the bands in the resulting autoradiograms were estimated by microdensitometry with a Joyce Loebl microdensitometer.

RESULTS

I. Viral Integration Sites in 702-C1 and 702-C3 Cells

Before Hpa II and Msp I digestions could be used to analyse cytosine methylation of Ad 12 DNA in transformed cells it was necessary to determine the arrangement of the viral sequences in the cellular DNA. Reassociation kinetics analysis of viral DNA sequences in cells transformed by Ad 12 or its restriction fragments has established that multiple copies of the transforming DNA are usually present per diploid amount of cell DNA (Fanning and Doerfler, 1976; Sawada et al, 1979). I have used blot hybridization analysis to determine the integration patterns of the viral DNA in lines of rat cells transformed with the Eco-RI-C fragment of Ad 12 DNA (Mak et al, 1979).

Analysis of integration sites and methylation patterns in cell lines 702-C1 and 702-C3 are presented first. I have characterized viral DNA sequences in line 702-C2 in less detail and these results are described in a later section (Section IV-C).

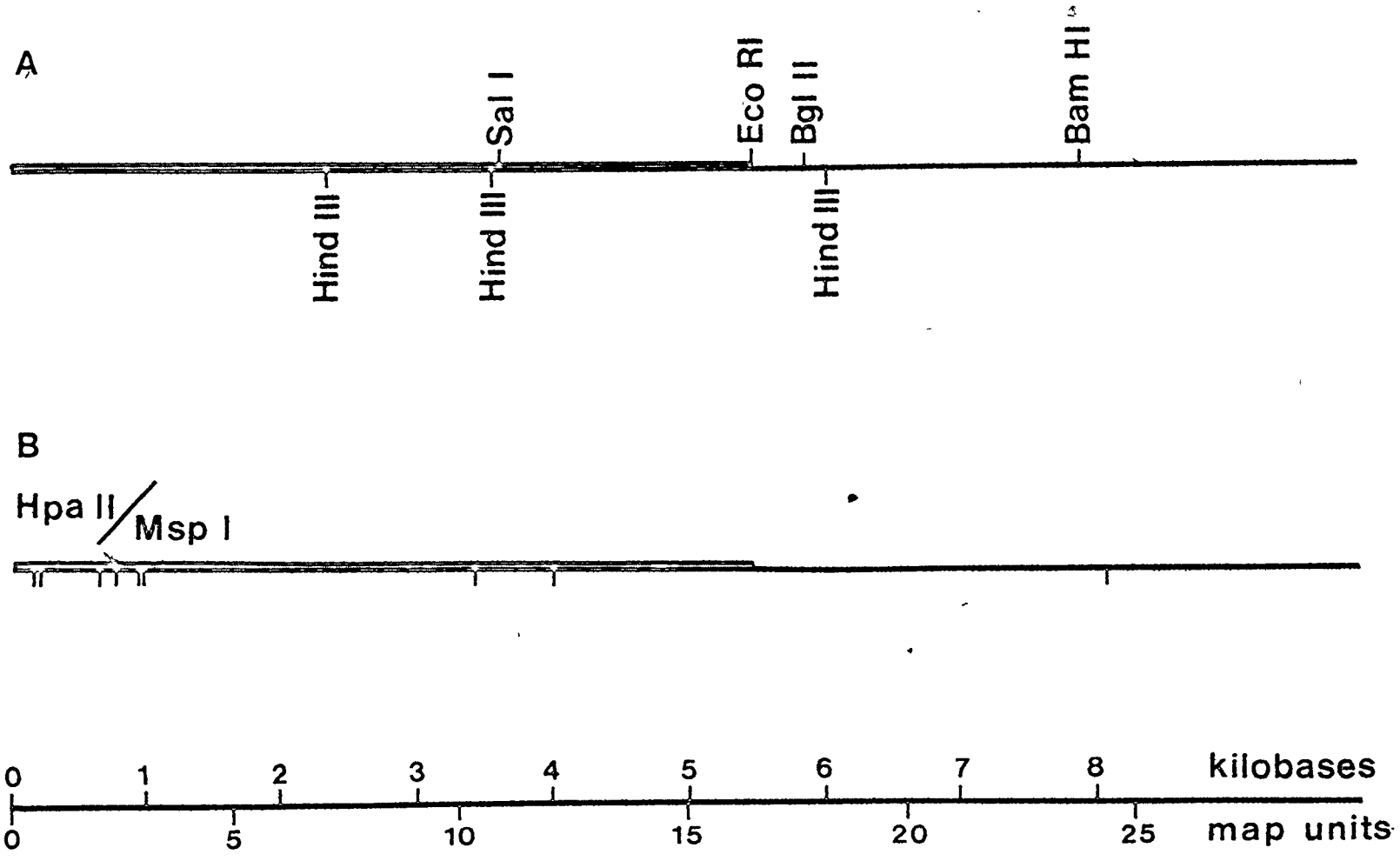
Recognition sites for a number of restriction endonucleases have been located on the Ad 12 genome (Tooze, 1980). Some of these sites are shown in the map of the left region of Ad 12 DNA in Figure 1a. The enzymes Eco RI, Bam HI, and Bgl II are particularly useful in the identification of integration patterns in these cell lines because they do not cut within the Eco RI-C fragment, and are not affected by cytosine methylation.

Figure 1. The transforming region of Adenovirus 12 DNA

Vertical lines represent restriction endonuclease cleavage sites. The thick line indicates the DNA fragment used to transform the 702 cell lines. Mapping of the Hpa II/Msp I sites at 3.35, 3.95, and 8.05 kb is described in section III. All other sites are taken from Tooze, Appendix F (1980).

A. Cleavages sites for Bam HI, Bgl II, Eco RI, Hind III and Sal I.

B. Cleavage sites for Hpa II/Msp I.


A
Hind III
Sal I
Hind III
Eco RI
Bgl II
Hind III
Bam HI
B
Hpa II
Msp I
0
1
2
3
4
5
6
7
8
kilobases
0
5
10
15
20
25
map units

These enzymes will cleave the transformed cell DNA only in cellular sequences, and thus each viral integration site will be contained within a single DNA fragment, in which the viral sequences are flanked on both sides by cellular sequences. Consequently, each band detected by blot hybridization with a viral probe will represent a complete viral DNA integration site. Because unrelated fragments of similar sizes may co-migrate, the number of bands seen in this analysis provides a minimum estimate of the number of integration sites.

Because the efficiencies of DNA transfer and hybridization are variable, band intensities in separate blot hybridization experiments cannot be compared. Accordingly, internal control lanes containing known amounts of Ad 12 DNA were included in each agarose gel electrophoresis used for blot hybridization analysis. Microdensitometry was used to compare the intensities of the bands in the transformed cell lanes with these control bands. The control DNA fragment bands also served as internal size standards for each experiment.

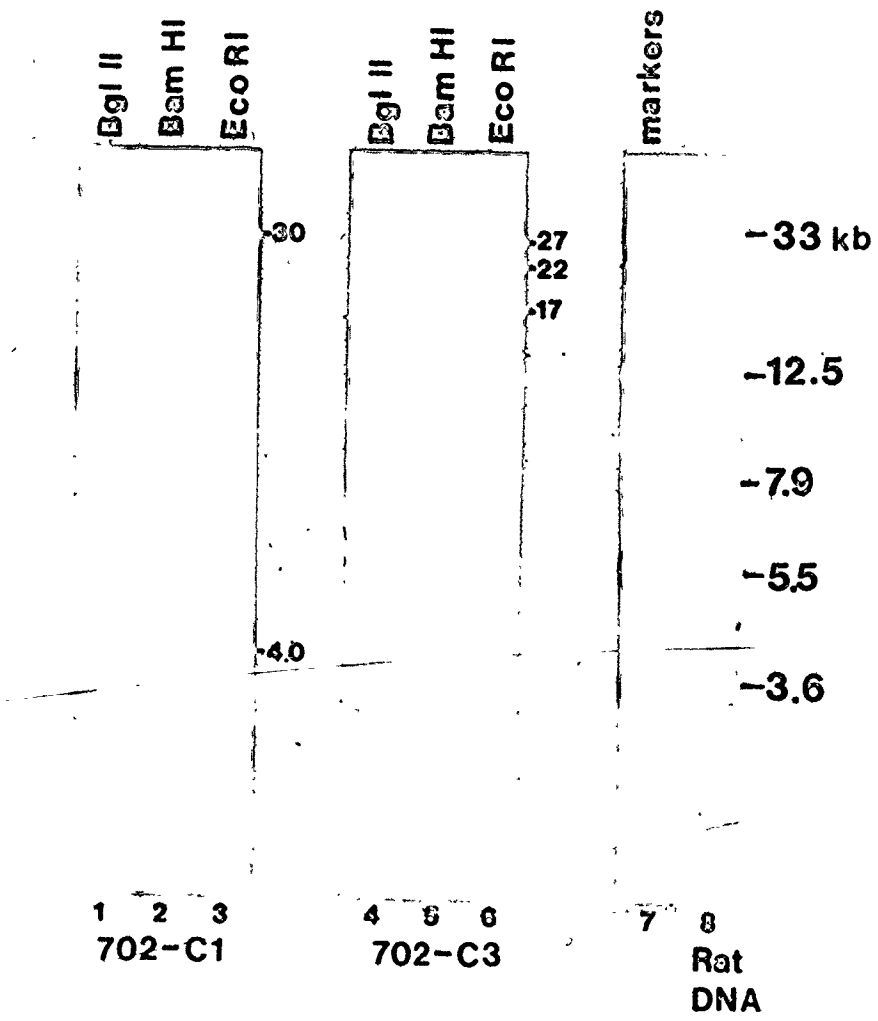
DNA from the cell lines 702-C1 and 702-C3 was digested with the restriction endonucleases Eco RI, Bam HI, and Bgl II. The fragments were separated by agarose gel electrophoresis, transferred onto a nitrocellulose sheet by blotting, and hybridized to an Ad 12 Eco RI-C probe, ^{32}P-labelled by nick translation to a specific activity of 4 x 10^8 cpm/ug. The results are shown in Figure 2.

A. Line 702-C1: (Figure 2, lanes 1 - 3)

In line 702-C1 (lanes 1-3), all three enzymes yielded two bands,

Figure 2. 702-C1 and 702-C3 DNA fragments containing integrated Ad 12 sequences.

DNA from transformed cells (702-C1 lanes 1 - 3, 702-C3 lanes 4 - 6) was digested with Bgl II, Bam HI, and Eco RI (lanes 1 and 4, 2 and 5, 3 and 6 respectively), separated by electrophoresis in a 0.5% agarose gel at 0.5 V/cm for 72 hours, blotted onto nitrocellulose, and hybridized to a ^{32}P-labelled Ad 12 Eco RI-C fragment probe. Ad 12-specific fragments were visualized by autoradiography. Lanes 1, 2, 3, 5, and 6 contained 10 µg DNA; lane 4 contained 5 µg. The marker lane 7 contained 10 µg of Hind III-digested normal rat DNA, plus 60 pg of each of the following Ad 12 digests; undigested, Sal I, Bam HI, and Eco RI; giving the following Ad 12 bands: undigested, 33 kb; Sal I-B, 12.5 kb; Bam HI-A, 7.9 kb; Eco RI-C, 5.5 kb; and Sal I-C, 3.6 kb. Lane 8 contained 10 µg of Hind II-digested normal rat DNA.


Bgl II
Bam HI
Eco RI
Bgl II
Bam HI
Eco RI
markers
30
4.0
27
22
17
-33 kb
-12.5
-7.9
-5.5
-3.6
1 2 3
702-C1
4 5 6
702-C3
7 8
Rat
DNA

one faint and one dark. Microdensitometry was used to compare the intensities of the bands in the transformed cell digestions with the marker bands in lane 7. This marker lane contained 60 pg of each of several Ad 12 restriction enzyme digests in 10 µg of untransformed cell DNA. This amount is equivalent to about 1.3 copies of each fragment per diploid amount of cell DNA (8×10^{-12} g/cell).

The intensities of the dark bands in lanes 1, 2, and 3 of Figure 2 are about 2.5 - 3 times those of the control bands in lane 7, suggesting that in each digest of 702-C1 DNA the dark band represents 3 - 4 copies of the transforming fragment per cell.

Similarly, in all three digests the intensities of the faint band are slightly less than that expected for a single complete copy of the Eco RI-C fragment per diploid amount of cell DNA. In the Eco RI digest in lane 3, the faint band is a fragment of only 4 kb, while the fragment used to transform the cells was 5.5 kb in size. This 4 kb band therefore, only contains sequences from part of the transforming fragment. This suggests that in all three digests the faint band represents a site at which a partial copy of the Eco RI-C fragment is integrated.

B. Line 702-C3: (Figures 2, lanes 4 - 6)

Lanes 4 - 6 of Figure 2 show Bgl II, Bam HI, and Eco RI viral-specific digestion patterns of DNA from 702-C3 cells. (The Bgl II digest in lane 4 contained only 5 µg of cell DNA, and is therefore fainter than the other digests which contained 10 µg of DNA.) The Bgl II

digest shows four major bands of equal intensities, each corresponding to about one copy of the transforming fragment per cell. The Bam HI digest in lane 5 also shows four major bands of equal intensities; again each is of an intensity slightly less than the marker bands and probably contains one copy of the Eco RI-C fragment per diploid amount of cell DNA.

The Eco RI digest of 702-C3 DNA shown in lane 6 has only three major bands, rather than the four bands predicted from the Bgl II and Bam HI analyses. However, the intensity of the largest major Eco RI band (27 kb) corresponds to about two copies of the transforming DNA, while the other two major Eco RI bands (22 kb and 17 kb) are of single-copy intensity. I conclude that the largest band is probably composed of two co-migrating fragments.

Three minor bands can also be detected in each of the 702-C3 digestions (lanes 4 - 6). These may result from the presence of fragmented portions of the viral DNA integrated into additional sites in cellular DNA.

C. Controls and Assumptions:

DNA purified from normal rat liver and digested with Hind III was analysed in the same experiment as the 702-C1 and 702-C3 DNAs described above (Figure 2, lane 8). The Ad 12 Eco RI-C probe did not hybridize at all to this DNA, showing that normal rat DNA does not contain sequences homologous to the transforming region of Ad 12 DNA.

From the Eco RI, Bam HI, and Bgl II digestions described above, I

have concluded that line 702-C1 contains viral DNA equivalent to 3 or 4 copies of the Eco RI-C transforming fragment at a single integration site, with a partial copy present at a second site. Similarly, I conclude that line 702-C3 contains single copies of the transforming fragment at 4 distinct integration sites, and partial copies at 3 other sites. I have not characterized the viral sequences present at each integration site in any detail, but have made the simplifying assumption that most of the Eco RI-C fragment is present intact at each major (non-partial) integration site. The band intensities seen in the Msp I and Hind III digestions which will be described in Section IV support this assumption.

II. Cytosine Methylation in Transformed Cell DNA and in Ad 12 Virion DNA.

Before investigating the methylation of the integrated viral DNA sequences in the transformed cells, it was desirable to determine a) the overall level of DNA methylation in the transformed cells, and b) the extent, if any, of methylation of viral DNA extracted from Ad 12 virions. This was done by comparision of digestion patterns obtained using the methylation-sensitive restriction endonuclease Hpa II and its methylation-insensitive isoschizomer Msp I.

The ethidium bromide-stained gel shown in Figure 3a shows that total cellular DNA is extensively methylated in lines 702-C1 and 702-C3. Undigested DNA samples are in lanes 1 and 4, Hpa II-digested samples in lanes 2 and 5, and Msp I-digested samples in lanes 3 and 6. In both cell lines the major part of the cellular DNA is in fragments of 15 - 50 kb

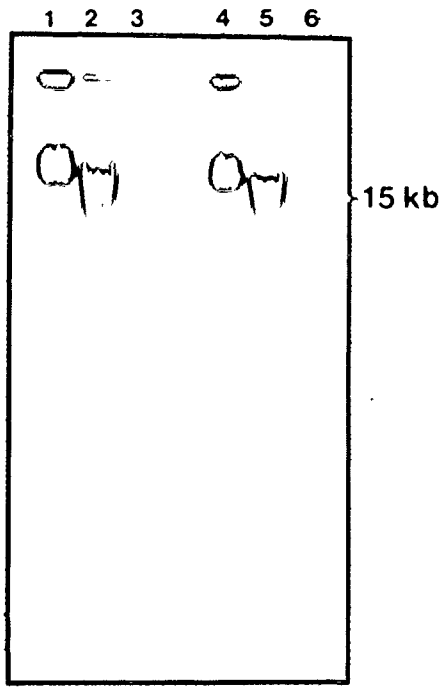
Figure 3. Hpa II and Msp I digestion patterns of transformed cell DNAs and of Ad 12 DNA.

5 µg samples of DNA from 702-C1 cells (lanes 1 - 3) and 702-C3 cells (lanes 4 - 6) were combined with 2 ng of ^{32}P-labelled Ad 12 DNA and digested with Hpa II (lanes 2 and 5) or Msp I (lanes 3 and 6). Lanes 1 and 4 contained undigested DNAs. Fragments were separated by electrophoresis in a 1% agarose gel at 4 V/cm for 4 hours.

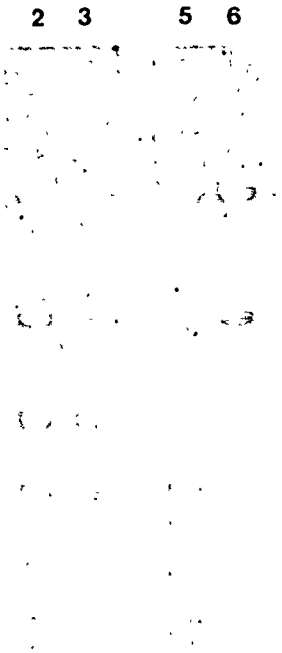
A. Photograph of the ethidium bromide-stained gel.

B. Autoradiogram of the same gel.

A



B



after Hpa II digestion, although only a trace of Msp I-digested DNA is in this size class. Most of the cellular CCGG sites (Hpa II and Msp I recognition sites) appear to be blocked for Hpa II cleavage, presumably by methylation of the internal cytosine (Waalwijk and Flavell, 1978a). Thus these transformed rat cell DNAs must be extensively methylated, as has been reported for Ad 12-transformed hamster cells (Gunthert et al, 1977).

Trace amounts (about 2 ng) of ^{32}P-labelled Ad 12 DNA (extracted from virions and labelled by limited nick translation) were included in the digests shown in Figure 3a, and after being stained and photographed the gel was dried and exposed to X-ray film (Figure 3b). This autoradiogram served two purposes. First, no differences can be seen between the Hpa II and Msp I digestion patterns of the Ad 12 DNA (compare lanes 2 and 5 with lanes 3 and 6 in Figure 3b). This observation has been repeatedly confirmed by analysis of the control digests described in the Methods section. I conclude that none of the viral CCGG sites in virion DNA are blocked by cytosine methylation. A second function for the ^{32}P-labelled viral DNA was to act as internal controls for complete digestion of the cell DNAs by Hpa II and Msp I. This is particularly necessary for the Hpa II digests, which resemble the digestion patterns seen after incomplete digestion by other, methylation-insensitive, enzymes.

III. Hpa II/Msp I Recognition Sites in Ad 12 DNA.

In order to use Hpa II and Msp I digestions to analyse patterns of methylation of Ad 12 DNA sequences integrated into transformed cells,

it was necessary to locate recognition/cleavage sites for these enzymes in the viral DNA. No restriction cleavage map is available for the 25 or more CCGG sites in Ad 12 DNA, although sites in the leftmost 2.3 kb (Hind III-G fragment) are known from the published sequences (Sugisaki et al, 1980). I have used Hpa II and Hpa II + Eco RI digests of Ad 12 virion DNA, and blot hybridizations of these digests to an Ad 12 Eco RI-C probe, to extend the Hpa II/Msp I map to 8.0 kb (about 25 map units) (Figure 4).

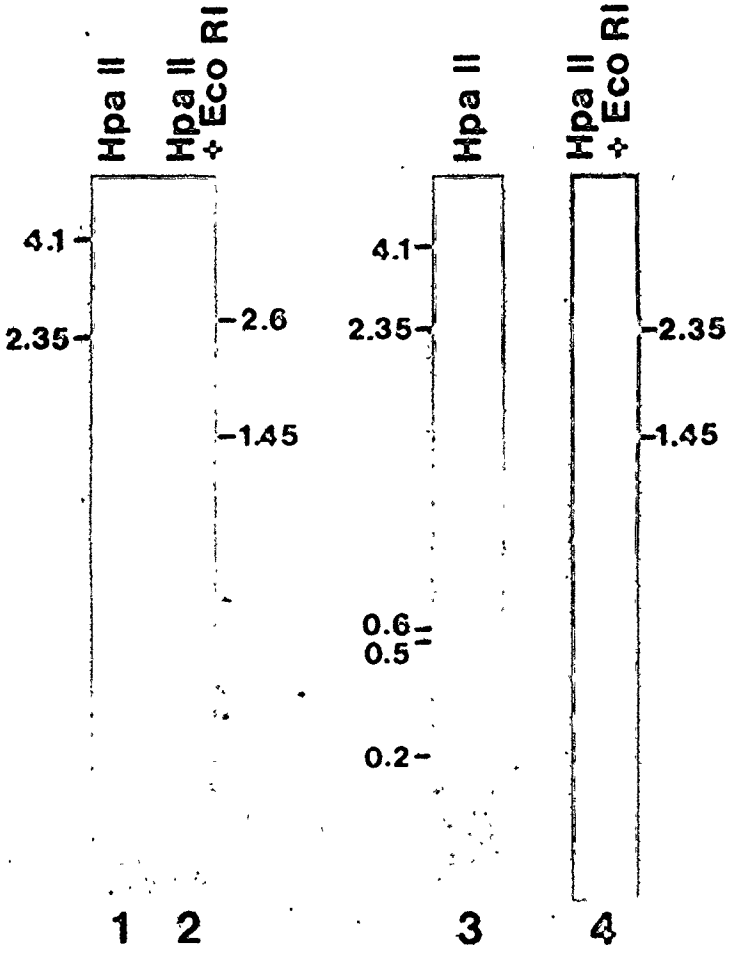
The Eco RI-C probe hybridizes to the two largest Ad 12 Hpa II fragments of 4.1 and 2.35 kb (more strongly to the smaller fragment), and to fragments of about 600, 500, and (faintly) 200 base pairs (bp) (Figure 4, lanes 1 and 3). Any smaller fragments from this region would not be detected, as they do not bind well to the nitrocellulose sheets used for the blot hybridization. The 4.1 kb fragment is cut by Eco RI (lane 2) to fragments of 2.6 and 1.45 kb. Only the 1.45 kb fragment hybridizes to the Eco RI-C probe (lane 4).

There are 6 CCGG sites in the sequenced region, at 144, 153, 657, 740, 966, and 997 bp from the left end, with no sites between 997 bp and the end of the sequenced fragment at 2320 bp. Hpa II digestion of these sites will give fragments of about 150, 9, 500, 100, 200, and 30 bp, and a fragment larger than 1300 bp. I have interpreted these data as follows (Figure 1B). The 500 and 200 bp fragments which hybridize to the Eco RI-C probe are predicted by the published sequence. The 2.35 kb fragment must be bounded by the CCGG site at 997 bp, and by a site at 3.35 kb from the left end. The 4.1 kb fragment must be bounded by sites at 3.95 and 8.05 kb, because of the sizes of its Eco RI cleavage fragments. The 600

Figure 4. Hpa II cleavage patterns of Ad 12 DNA

Lanes 1 and 2; Hpa II and Hpa II + Eco RI digestion patterns of Ad 12 DNA (^{32}P-labelled by limited nick translation).

Lanes 3 and 4; Blot hybridizations of a ^{32}P-labelled Ad 12 Eco RI-C probe to Hpa II and Hpa II + Eco RI digests of Ad 12 DNA.


Hpa II
Hpa II + Eco RI
Hpa II
Hpa II + Eco RI
4.1
2.35
2.6
1.45
4.1
2.35
0.6
0.5
0.2
2.35
1.45
1
2
3
4

bp fragment is now predicted by the sites mapped at 3.35 and 3.95 kb from the left end.

Although the presence of undetected CCGG cleavage sites very close to the mapped sites cannot be excluded without fine-structure analysis, the data presented in Figure 4 and the sequence data are sufficient to establish the positions I have mapped, within an accuracy of about 150 bp. Double digestions have been used to confirm these positions. For example, the maps in Figure 1 predict that, of the Hpa II fragments in this region, only the 600 bp fragment (from 3.35 - 3.95 kb) will contain a Sal I cleavage site (at 3.6 kb, Figure 1a). Blot hybridization analyses of Hpa II and Hpa II + Sal I digests of Ad 12 DNA were compared, and as predicted, only the 600 bp Hpa II fragment was cut by Sal I (data not shown). Similar double digestions with Hind III, Acc I, and Kpn I have provided additional confirmation of the mapped Hpa II sites. The location of a CCGG Hpa II site at 8.05 kb is further confirmed by the presence of a Sma I site (CCCGGG) at this position, mapped independently (Tooze, Appendix F, 1980).

IV. Methylation of Viral Sequences in Ad 12-Transformed Cells.

A. Hpa II, Msp I, and Sal I Blot Hybridization Analysis

Adenovirus DNA does not become methylated during lytic replication. However, when viral sequences become integrated into cellular chromosomes as a consequence of cell transformation, they fall under the influence of cellular regulatory mechanisms. DNA methylation may be such a regulatory mechanism, so it was of interest to determine whether the systems which specifically initiate and maintain methylation of cell DNA would also act

on the exogenously-acquired Ad 12 sequences in lines 702-C1 and 702-C3.

To investigate whether any of the viral sequences integrated in the transformed cells had become methylated at CCGG sites, DNA samples from 702-C1 and 702-C3 cells were cut with Hpa II and Msp I and analysed by blotting and hybridization to the ^{32}P-labelled viral Eco RI-C probe. The transforming fragment contains 8 Hpa II recognition sites (Figure 1) and, if the viral DNA has remained unmethylated, Hpa II cleavage of each integrated copy should yield a 2.35 kb fragment, several fragments 0.6 kb and smaller, and 2 fragments linked to cellular sequences.

i. Line 702-C1:

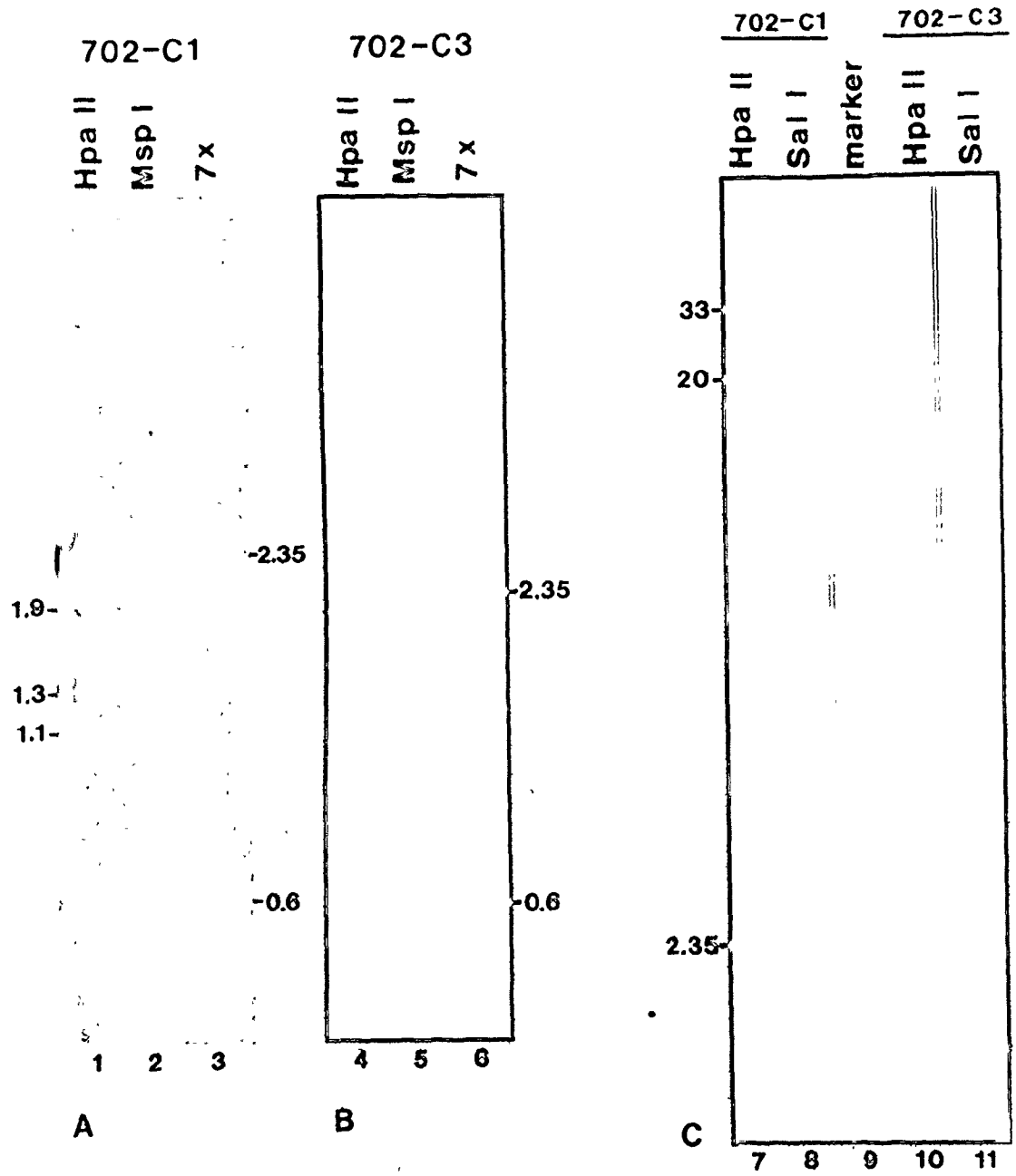
Digests of 702-C1 DNA analysed by blot hybridization are shown in Figure 5A. The Hpa II and Msp I digests (lanes 1 and 2) are almost indistinguishable. The 2.35 and 0.6 kb bands which were predicted to be cut from unmethylated viral DNA are seen in both the Hpa II and Msp I digests. (Compare with the Hpa II digest of Ad 12 DNA in the 7x marker in lane 3.) In addition in both lanes there are three bands (1.9, 1.3, and 1.1 kb) which must contain viral DNA linked to cell sequences, as homologous bands are not present in the marker lane.

The intensities of the 2.35 kb band in lanes 1 and 2 are equivalent to 3 - 5 copies of this fragment per cell, by comparision with the marker band. It is not possible to determine the copy numbers of the linker fragments, as they consist of an unknown proportion of cellular sequences. A rough calculation of the amount of viral DNA in each band (by relative intensity) gives values greater than the sizes of the band fragments (i.e. the 1.9 kb band contains 2.5 kb of viral DNA, the 1.3 band 2.2 kb, and the

Figure 5. Hpa II, Msp I, and Sal I digestion patterns of integrated Ad 12 DNA in 702-C1 and 702-C3 cells.

10 µg samples of transformed cell DNAs (702-C1, lanes 1, 2, 7, and 8; 702-C3, lanes 4, 5, 10, and 11) were digested with Hpa II, Msp I, or Sal I as indicated. The 7X markers (lanes 3 and 6) contained 0.3 ng of Ad 12 DNA digested with Hpa II. The lane 9 marker contained 60 pg of Ad 12 DNA and of the Sal I, Bam HI, and Eco RI Ad 12 digests described in the legend to Figure 2.

Electrophoresis conditions: A and B; 1% agarose, 2 V/cm, 12 hr. C; 0.5% agarose, 0.5 V/cm, 96 hr. Viral fragments were detected by blot hybridization and autoradiography, using the ^{32}P Eco RI-C probe.


702-C1
Hpa II
Msp I
7x
1.9-
1.3-
1.1-
1 2 3
A
702-C3
Hpa II
Msp I
7x
2.35
0.6
4 5 6
B
702-C1
702-C3
Hpa II
Sal I
marker
Hpa II
Sal I
33-
20-
2.35-
7 8 9 10 11
C

1.1 band 1.7 kb) suggesting that these fragments are present at more than one copy per cell.

There is a small amount of viral DNA at the top of the Hpa II digest in lane 1. This DNA is similar in mobility to the heavily-methylated cell DNA which was seen by ethidium bromide staining of this gel before blotting (not shown, but see the Hpa II digest of 702-C1 DNA in Figure 3A, lane 2). This high-molecular weight DNA is not present in the Msp I digest in lane 2; instead there appears to be a corresponding increase in the intensity of the 1.9 kb linker band.

These results show that there is little or no methylation of CCGG sites in the viral sequences in this cell line. The trace of high-molecular weight DNA in the Hpa II digest most likely consists of a copy of the 1.9 kb linker fragment attached to cellular sequences by a methylated cellular CCGG site. If this is the case, there is no evidence that any of the viral CCGG sites are methylated.

ii. Line 702-C3:

Blot hybridization analysis of Hpa II and Msp I digests of DNA from line 702-C3 is shown in Figure 5B, lanes 4 and 5. Comparison of these two digests indicates that there is extensive methylation of viral sequences in this cell line.

The Msp I digest analysed in lane 5 shows, as expected, the Ad 12-specific 2.35 and 0.6 kb fragments, with intensities corresponding to about 4 copies of each per cell. In addition at least 8 cell DNA + viral DNA linker bands can be distinguished.

The Hpa II digest in lane 4 indicates that most of the viral CCGG

sites in this cell line cannot be cut by Hpa II. The 2.35 and 0.6 kb bands are present, but are reduced in intensity relative to the Msp I digest. Most of the Msp I linker bands are completely absent from the Hpa II digest. Instead, more than half of the viral-specific DNA in lane 4 is in the same high-molecular weight Hpa II-resistant size class which was seen for ethidium bromide stained 702-C3 cell DNA in Figure 3a, lane 5.

I conclude that at least half of the viral CCGG sites in line 702-C3 have become methylated at the internal cytosine, as such methylation is the only modification known to block Hpa II cleavage but permit Msp I cleavage. The molecular weight of much of the Hpa II-resistant DNA is much larger than that of the transforming Eco RI-C fragment (20 - 50 kb compared to 5.5 kb) which suggests that the viral DNA sequences are linked to cellular sequences whose CCGG sites are also blocked by methylation.

iii. High Resolution Analysis:

In the gels shown in Figure 5A and 5B, and in other gels, viral-specific bands could not be resolved in the high-molecular weight, Hpa II resistant DNA seen by blot hybridization. This was true for the trace of viral-specific Hpa II resistant DNA seen in 702-C1 DNA, as well as for the large amount of methylated viral DNA in line 702-C3.

The failure to resolve distinct bands in this region of the gels might be due to poor resolution of very large DNA fragments in the gel systems used (1% agarose vertical gels run for 12 h at 2 V/cm and 0.5% agarose horizontal gels run for 40 h at 1 V/cm), or might be due to real

heterogeneity of methylation at the CCGG sites bounding the methylated viral DNA regions. To distinguish between these possibilities, Hpa II digests of 702-C1 and 702-C3 DNAs were analysed in a 0.5% agarose horizontal gel run for 96 h at 0.5 V/cm, conditions which should increase the resolution of large DNA fragments (Figure 5C).

For line 702-C1, this high-resolution analysis showed that the faint smear of high-molecular weight DNA seen in the low-resolution Hpa II analysis (Figure 5, lane 1) was actually a faint band of 20 kb (Figure 5C, lane 7). This band was not seen in a parallel Msp I digest (not shown, but compare to lane 2 of Figure 5).

Similar high-resolution Hpa II analysis of 702-C3 DNA is shown in lane 10 of Figure 5C. Two faint bands can be seen at 35 and 40 kb, superimposed on a background of unresolved viral fragments. The adjacent markers (lane 9) show that poor resolution is not simply a consequence of fragment size, as the 33 kb marker fragment forms a sharp band.

When Hpa II digests of cell DNA are analysed by gel electrophoresis most of the DNA is concentrated in a small area near the top of the gel (see Figure 3A, lanes 2 and 5). It is possible that this local high concentration exceeded the capacity of the agarose in this region of the gel, and that the local overloading caused a loss of resolution of high-molecular weight DNA fragments. To test this, a lane was analysed (in the same gel) which contained only 3 ug of Hpa II-digested 702-C3 DNA, but no improvement of resolution was seen. (The autoradiograph was too faint for photographic reproduction, but was otherwise identical to the one shown in lane 10 of Figure 5.)

Although this high resolution gel eliminated the apparent hetero-

geneity of the high-molecular weight Hpa II-resistant viral DNA in line 702-C1, it did not eliminate similar heterogeneity of the 702-C3 viral Hpa II-resistant DNA. Therefore, the integrated copies of viral DNA in this cell line may be flanked by cell DNA whose methylation patterns are heterogeneous in the 702-C3 cell population.

iv. Sal I Analysis:

Digestions with other restriction endonucleases which are sensitive to CpG methylation can also be used to study vertebrate DNA methylation patterns. Cleavage at the Sal I recognition site (CTCGAC) is blocked by methylation of the internal cytosine (van der Ploeg and Flavell, 1980). There is one recognition site for this enzyme in the Ad 12 Eco RI-C fragment used for transformation, at 3.6 kb from the left end (Figure 1A), and cleavage of the integrated DNA at this site should give two cell DNA-linked fragments from each integrated copy of the viral DNA.

Sal I digests of 702-C1 and 702-C3 DNAs were included in the high-resolution gel shown in Figure 5C (lanes 8 and 11). Sal I digestion of 702-C1 DNA produces two distinct bands, at 33 and 5 kb (lane 8). The 702-C3 digest contains only unresolved viral fragments averaging 40 - 50 kb in size, suggestive of extensive methylation (lane 11). Although not in themselves conclusive (no methylation-insensitive isoschizomer of Sal I is available), the results of these digests agree with the major conclusions drawn from the Hpa II/Msp I analysis.

B. Methylation of Viral DNA at Different Integration Sites:

Analysis of the Hpa II and Msp I digestion patterns of viral sequences in lines 702-C1 and 702-C3 suggests that virus-specific CCGG sites may occur in both methylated and unmethylated forms. One general observation leading to this conclusion is that while all of the viral CCGG sites have remained unmethylated in line 702-C1, some of the same sites are methylated in line 702-C3. A more specific example is seen in line 702-C3. Although the 2.35 kb fragment (produced by cleavage at the CCGG sites at 1.0 and 3.35 kb on the Ad 12 genome) is present at about 4 copies per cell in the Msp I digest of 702-C3 DNA (Figure 5, lane 4), the same fragment is reduced in Hpa II digests of the same DNA (1 - 3 copies/cell; Figure 5, lane 5). This implies that the CCGG sites at 1.0 and 3.35 kb are present in both unmethylated and methylated forms in this cell line. The same argument can be applied to the site at 3.95 kb on the viral genome, which gives rise to the 0.6 kb fragment.

The partial methylation of these viral sites does not reflect non-specific partial methylation of all CCGG sites associated with the viral sequences (i.e. each site having a 50% probability of being methylated in any one cell). Most of the viral linker bands seen in the Msp I digest (Figure 5, lane 5) are absent from the parallel Hpa II digest. This implies that the sites which give rise to these fragments are methylated in all cells in the population.

These observations suggest that copies of the viral DNA which are integrated at different sites may have different patterns of methylation. I have used double digests to investigate the relationship between

methylated and unmethylated CCGG sites in the viral and adjacent cellular DNA.

As discussed in Section I, digestions with the enzymes Eco RI and Bam HI will give fragments each containing an integrated copy of the viral DNA with its adjacent cellular sequences. To determine if the CCGG sites in these fragments were methylated, the susceptibility of the Eco RI and Bam HI viral fragments to Hpa II was determined. Double digestions using Eco RI + Hpa II and Bam HI + Hpa II were compared to the Eco RI and Bam HI digestions which were shown in Figure 2. (The comparisons are shown in Figure 6.)

i. Line 702-C1: (Figure 6, lanes 1 - 5)

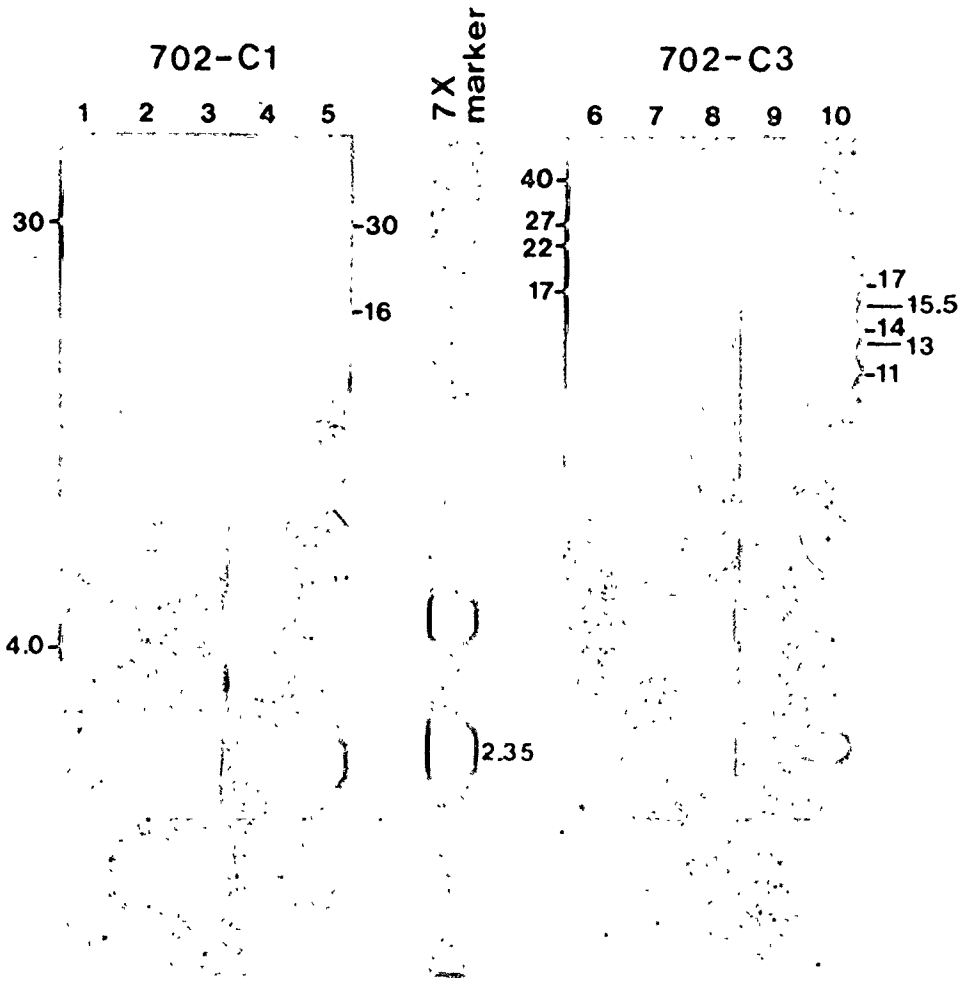
Hpa II and Msp I digestions have already shown that the viral CCGG sites in this cell line are not methylated (Section II, part A). Consistent with this, the two viral-specific bands seen in both Eco RI digests (30 kb and 4 kb) and Bam HI digests (30 kb and 16 kb) are cut by Hpa II (compare lane 1 with lane 2, and lane 4 with lane 5).

ii. Line 702-C3: (Figure 6, lanes 6 - 10)

Comparisons of the 702-C3 double digests in lanes 7 and 10 with the Eco RI and Bam HI digests in lanes 6 and 9 confirm the hypothesis that the viral sequences at different integration sites have different patterns of methylation. The most striking observation is that some Eco RI and Bam HI viral fragments are totally resistant to cleavage by Hpa II. Fragments which contain a complete copy of the Ad 12 Eco RI-C fragment will contain 8 CCGG sites, and the cellular sequences in these large Eco

Figure 6. Hpa II and Msp I cleavages of the Eco RI and Bam HI viral integration fragments of 702-C1 and 702-C3 DNAs.

10 µg samples of 702-C1 DNA (lanes 1 - 5) and 702-C3 DNA (lanes 6 - 10) were digested as indicated below, separated by electrophoresis in a 0.5% agarose gel at 0.5 V/cm for 70 hr, blotted, and hybridized to a ^{32}P-labelled Eco RI-C probe. Lanes 1 and 6, Eco RI; lanes 2 and 7, Eco RI + Hpa II; lanes 3 and 8, Eco RI + Msp I; lanes 4 and 9, Bam HI; lanes 5 and 10, Bam HI + Hpa II. The 7X marker lane in the centre contained 0.3 ng of Hpa II-digested Ad 12 DNA plus 10 µg of Hind III-digested normal rat DNA.


702-C1
1 2 3 4 5
30
4.0
30
16
7X marker
2.35
702-C3
6 7 8 9 10
40
27
22
17
17
15.5
14
13
11

RI and Bam HI fragments are likely to include several additional CCGG sites. The total resistance of some of the Eco RI and Bam HI fragments to cleavage by Hpa II means that all of these CCGG sites (viral and cellular) are blocked by cytosine methylation.

The 11 kb Bam HI fragment is totally resistant to Hpa II. The 13 kb fragment is partly resistant (lanes 9 and 10). The two larger Bam HI fragments (20 and 18 kb) are cut by Hpa II, giving three faint fragments of 17, 15.5, and 14 kb; fragments this large are not seen in any Msp I digestions of 702-C3 DNA, (Figure 5, lane 4) so these fragments must contain methylated CCGG sites.

Similar results are seen in the Eco RI + Hpa II digest in lane 7. The faint 40 kb fragment and the double-intensity 27 kb fragment are both highly resistant to Hpa II which implies that all internal CCGG sites in these fragments are methylated. The 17 kb band appears partly resistant, while the 22 kb band is entirely absent after Hpa II + Eco RI digestion. I have observed some variability in the relative intensities of some of the 702-C3 viral bands after Eco RI and Eco RI + Hpa II digestions. Although the Hpa II resistance of the 40 kb fragment and the Hpa II sensitivity of the 22 kb fragment are reproducible both the 27 kb and 17 kb fragments have varied from about 50% to 100% Hpa II resistant in different experiments. These variations are seen both within and between different DNA preparations, and probably result from variability inherent in the experimental techniques.

Both Bam HI + Hpa II and Eco RI + Hpa II digestion patterns are consistent with total methylation (at CCGG sites) of the viral DNA at one integration site in all cells, total methylation of a second viral

integration site in 50% or more of the cells, and the presence of unmethylated CCGG sites in the viral DNA at the other major integration sites in all cells.

C. Line 702-C2: (Figure 7)

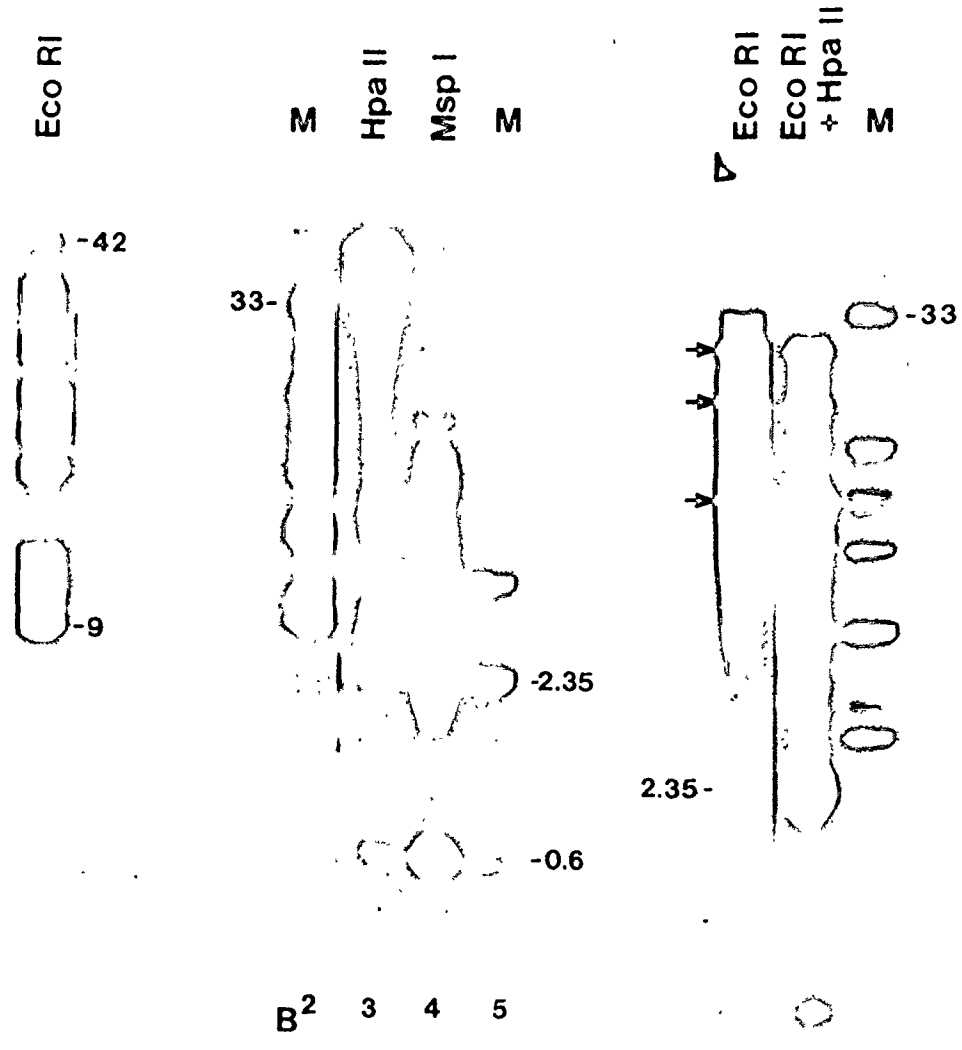
I have characterized the viral integration sites in a sister cell line, 702-C2, in less detail. This cell line was found to contain about 9 copies of the Ad 12 Eco RI-C fragment per diploid amount of cell DNA by analysis of reassociation kinetics data (Mak et al, 1979). I have resolved 9 major and 2 minor Ad 12-specific Eco RI fragments in DNA from these cells, using the high-resolution gel electrophoresis described above (Figure 7a). These fragments range in size from 9 to 42 kb.

Blot hybridization analysis of viral-specific Hpa II and Msp I fragments (Figure 7b) shows patterns similar to those of line 702-C3. More than half of the viral DNA remains in the 15 - 50 kb region of the gel after Hpa II digestion; Msp I fragments are mostly between 2 and 4 kb (lanes 3 and 4). This implies that at least half of the CCGG sites associated with viral sequences are methylated. The Ad 12-specific Hpa II fragments (2.35 and 0.6 kb) are faint in the Hpa II digest (2 - 4 copies per cell), but much darker in the Msp I digest, so most of the CCGG sites which define these two fragments must be methylated.

Comparison of Eco RI + Hpa II digests with parallel Eco RI digests (Figure 7c, lanes 6 and 7) shows three bands of Hpa II-resistant viral DNA (indicated by arrows beside lane 6) all of which co-migrate with Eco RI bands and are of comparable intensities. This is a low resolution gel

Figure 7. Restriction endonuclease digestion patterns of viral sequences in 702-C2 cells.

10 μg samples of DNA from line 702-C2 were digested with Eco RI, Hpa II, Msp I, or Eco RI + Hpa II as indicated, separated by electrophoresis, blotted, and hybridized to a ^{32}P Ad 12 Eco RI-C probe. Lane 1 is from one gel (0.5% agarose, 0.5 V/cm, 96 hr) lanes 2 - 5 and 6 - 8 are from two additional gels (0.5% agarose, 1 V/cm, 65 hr). The high molecular weight markers (lanes 2 and 8) contained 120 pg of the same Ad 12 digests as lane 7 of Figure 2 (the major bands are 33, 12.5, 7.9, 5.5, and 3.6 kb). The Hpa II marker (lane 5) contained 120 pg of Hpa II-digested Ad 12 DNA.


Eco RI
M
Hpa II
Msp I
M
Eco RI
Eco RI
+ Hpa II
M
-42
-9
33-
-2.35
-0.6
2.35-
-33
A 1
B 2 3 4 5
C 6 7 8

702-C2

compared to the one in lane 1, and the lowest of the Hpa II-resistant bands in lanes 6 and 7 is probably two unresolved fragments.

I conclude that this cell line, like 702-C3, contains single copies of the transforming Ad 12 fragment integrated at 9 or more sites in cellular DNA. At three or four of these sites, all of the CCGG sites in viral and flanking cellular sequences are methylated.

D. Methylation Patterns of Viral Hind III Fragments:

To allow analysis of viral methylation patterns in the three cell lines at a finer scale, Hind III digestions were used. Hind III cuts at two sites in the Eco RI-C transforming fragment (Figure 1a) yielding two cell DNA-linked fragments and one internal viral fragment of 1.2 kb from each separately-integrated copy. The Hpa II map in Figure 1b shows that all of these Hind III fragments will contain viral CCGG sites, and thus should be amenable to Hpa II/Msp I analysis.

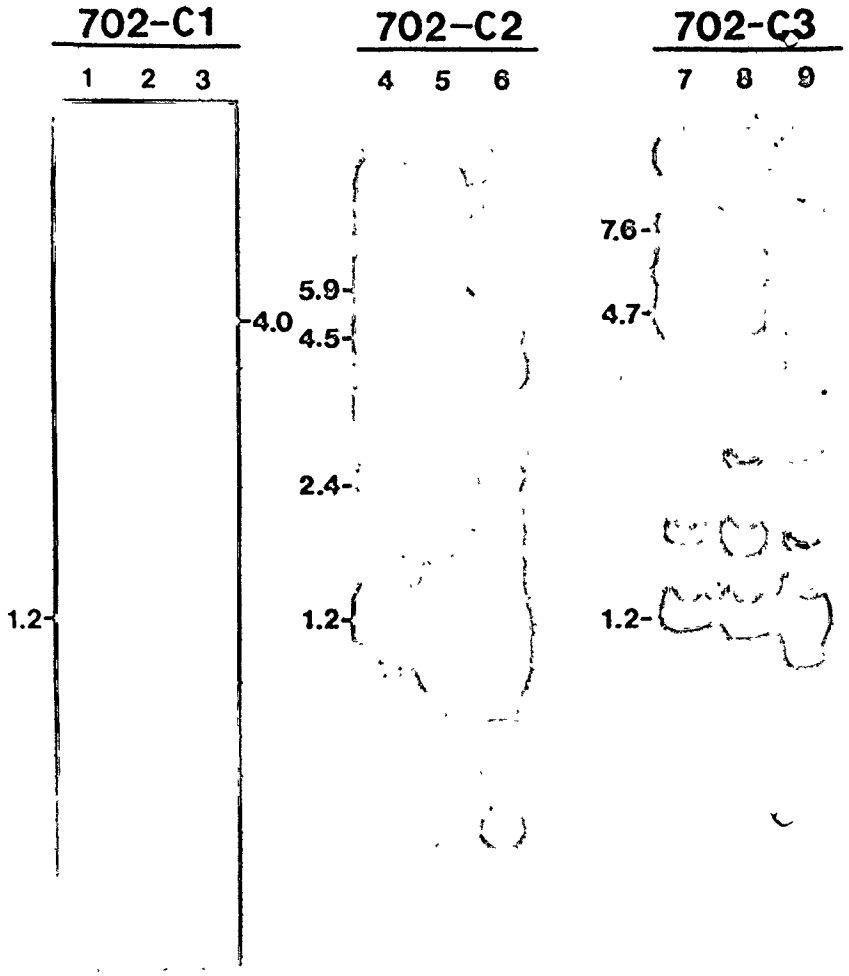
Figure 8 shows Hind III, Hind III + Hpa II, and Hind III + Msp I digestion patterns for viral-specific DNA from all three cell lines. Comparisons of Hind III and Hind III + Msp I digestion patterns show that most of the Hind III fragments in each cell line are cut by Msp I as predicted, and therefore can be analysed for methylation of CCGG sites by Hpa II digestions. (Compare lane 1 with lane 3, lane 4 with lane 6, and lane 7 with lane 9.)

The 1.2 kb Hind III fragment is of course present in multiple copies in digests from all three cell lines. This fragment is cut by Msp I to 1.03 kb, but its disappearance is obscured by the appearance

Figure 8. Hpa II and Msp I cleavage patterns of the Hind III viral integration fragments of transformed cell DNAs.

10 µg samples of transformed cell DNAs (left panel 702-C1, centre panel 702-C2, right panel 702-C3) were digested as indicated below, separated by electrophoresis in 1% agarose gels at 2 V/cm for 12 hr, blotted, and hybridized to a ^{32}P Ad 12 Eco RI-C probe.

Hind III lanes 1, 4, and 7; Hind III + Hpa II lanes 2, 5, and 8; Hind III + Msp I lanes 3, 6, and 9.


702-C1
1 2 3
4.0
1.2
702-C2
4 5 6
5.9
4.5
2.4
1.2
702-C3
7 8 9
7.6
4.7
1.2

of a new, Hind III + Msp I fragment of 1.2 kb (predicted by the cleavage maps in Figure 1).

i. Line 702-C1 (Figure 8, left panel)

In addition to the Ad 12-specific 1.2 kb Hind III fragment, line 702-C1 has three Hind III cell DNA-viral DNA linker bands (Figure 8, lane 1). This is in agreement with the three Hpa II linker bands seen in Figure 5, lane 1.

All of the Hind III viral fragments in line 702-C1 are cut by Hpa II (compare lanes 1 and 2). This is in agreement with the earlier conclusion that there is no methylation of viral sequences in this cell line. One faint linker fragment of 4.0 kb, produced by Hind III + Hpa II digestion, is further cut by Hind III + Msp I digestion, apparently to a fragment which co-migrates with another linker fragment at 1.9 kb. This agrees with the earlier observation in this cell line of a Hpa II fragment, containing viral and cellular sequences, which was cut by Msp I to co-migrate with a 1.9 kb linker fragment (Figure 5). All other bands are identical between the Hind III + Hpa II and Hind III + Msp I digests.

ii. Line 702-C2: (Figure 8, centre panel)

In the blot hybridization of 702-C2 Hind III digests shown in lane 4, 14 viral bands are resolved. Some of these bands are probably composed of co-migrating but unrelated fragments, as 9 integration sites should give rise to 18 linker bands. The intensity of the 1.2 kb viral

band is commensurate with the large number of copies of the transforming fragment present in this cell line.

The 702-C2 Hind III fragment of 2.0 kb may contain a CCGG site, but possible cleavage is obscured by Hind III + Hpa II and Hind III + Msp I fragments of similar intensities and mobilities. The other twelve of the thirteen Hind III linker fragments are seen to be cut by Msp I; of these only three are cut by Hpa II (those at 5.9, 4.5, and 2.4 kb; compare lanes 4 and 5). The 4.5 kb band is only partly sensitive; about 55% of the viral DNA in this band remains uncut by Hpa II. The intensity of this 4.5 kb band is double the maximum predicted for a single Hind III linker fragment, and it probably results from two or more co-migrating fragments.

The Hpa II and Eco RI + Hpa II digests of 702-C2 DNA seen in Figure 7 showed that the viral sequences in this cell line are extensively methylated. This analysis of Hind III fragments confirms and extends the previous results for this cell line, showing that most of the viral DNA-cell DNA linker fragments contain methylated CCGG sites but no unmethylated sites. The 9½ totally methylated Hind III bands cannot be accounted for solely by the number of totally methylated copies of the transforming fragment which were identified in Figure 7. (The 3 (or 4) totally methylated copies in line 702-C2 would give 6 (or 8) Hind III linker bands which contain only methylated CCGG sites.) The additional totally methylated Hind III linker bands must come from integrated copies which also contain some unmethylated CCGG sites.

iii. Line 702-C3: (Figure 8, right panel)

Like line 702-C2, this line has been shown in previous sections to be heavily methylated at viral CCGG sites. There are 9 Hind III linker bands in the 702-C3 digest; as expected from the maps in Figure 1 all of these are cut by Msp I (compare lanes 7 and 9). Only two of these Hind III bands are cut by Hpa II (lane 8), one at 7.6 kb and one at 4.7 kb. (The 4.7 kb band is poorly separated from a band at 4.5 kb which is not cut by Hpa II.)

As in 702-C2 DNA, a high proportion of the CCGG sites in the Hind III linker fragments are methylated (7 resistant bands out of 9), and these cannot be accounted for by the small number of copies which were estimated to be completely methylated in Section IV-B. Integrated copies of the viral DNA which contain unmethylated CCGG sites must give rise to some of the totally methylated Hind III linker fragments.

V. DNase I Digestions.

Other researchers have described correlations between the methylation of specific sites in genes and the transcription of those genes. In particular, Waalwijk and Flavell (1978b) and McGhee and Ginder (1979) have identified CCGG sites in globin genes which are methylated only in tissues which do not transcribe the globin gene sequences. The Ad 12-transformed cell lines 702-C2 and 702-C3 contain copies of the Ad 12-transforming fragment which are methylated at all CCGG sequences, and some copies which are unmethylated at some CCGG sequences. Line 702-C1 contains the same viral genes in a totally unmethylated state. Therefore, I have

tried to test the hypothesis that the totally methylated integrated copies of the transforming viral sequences are not transcriptionally active.

Transcriptional activity of a DNA sequence is normally determined by assaying the cells for complementary RNA. It has already been determined that these transformed cells do contain RNA complementary to the transforming Eco RI-C fragment (Mak et al, 1979). To determine which of the integrated copies are transcriptionally active, I have used limited DNase I digestions of the transformed cell chromatin. DNase I will discriminate between the chromatin conformations of active and inactive DNA sequences, preferentially digesting the former (Weintraub and Groudine, 1976).

The preferential degradation of active DNA sequences by DNase I digestion of chromatin has usually been assayed by reassociation kinetics analysis of the gene sequences present in DNA from control and DNase I-treated nuclei, with probes for inactive gene sequences used as negative controls. The sensitivity of specific DNA sequences to DNase I can also be determined by comparison of blot hybridization analyses of control and DNase treated DNAs. This technique is less cumbersome than analysis of reassociation kinetics, and is the only method suitable for the identification of transcriptionally inactive integrated copies of the viral DNA in the transformed cell lines.

No probes were available for genes known to be totally active or totally inactive in the transformed cells, which could have served as positive and negative controls for the DNase I digestions. Accordingly, the classification of a viral-specific band as DNase I-sensitive or

resistant was done by comparison with the extent of degradation of total cellular DNA and of other viral bands in the blot hybridizations.

Interpretation of DNase I digestions analysed by blot hybridization is made more complex by the background of non-specific DNase I cleavages; large DNA fragments are more rapidly degraded by this random fragmentation than are small fragments, because of their larger target size. Resistance or sensitivity of any one band must be assessed by comparison with the relative survival of adjacent bands of similar sizes.

Nuclei were prepared from transformed cells by hypotonic lysis with NP-40 and were digested lightly with DNase I (0 - 5 μg/ml for 5 minutes). The extent of digestion was such that greater than 95% of the DNA remained acid precipitable. The DNase I-resistant DNA was then purified and digested with Eco RI. The fragments were separated by gel electrophoresis, blotted, and hybridized to the ^{32}P-labelled Ad 12 Eco RI-C probe.

Results of one such experiment are shown in Figure 9. There are two interesting findings. The first is the appearance of new DNase I-specific bands in blot hybridizations of all three cell lines. The second is that no clear-cut differences can be seen in the nuclease sensitivities of the various Eco RI bands in each cell line.

A. Sequence-specific Cleavages by DNase I:

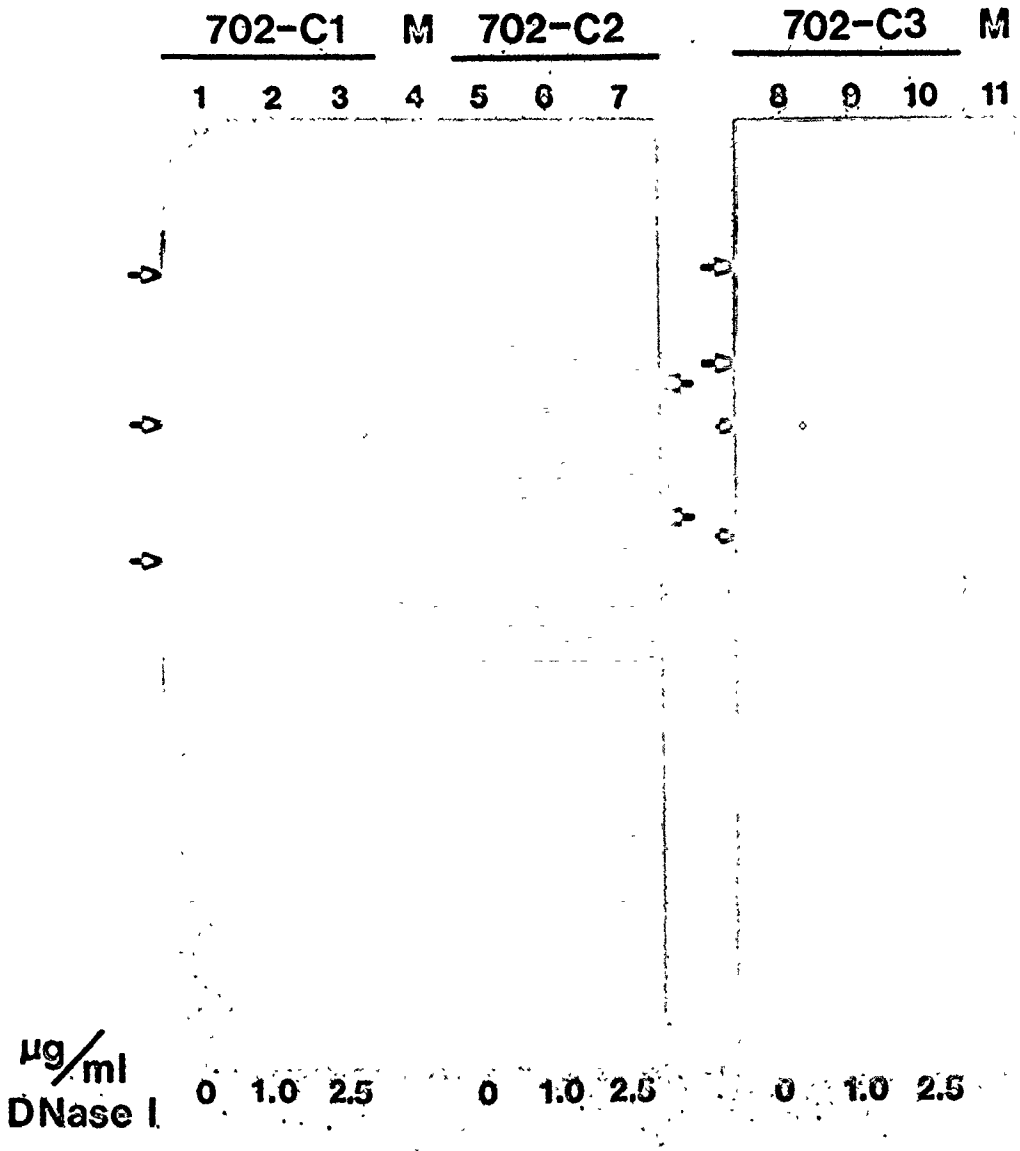
In all three cell lines, DNase I digestion of nuclei produced new viral-specific bands in Eco RI digests of the purified DNAs. These bands can be seen in Figure 9 when the DNase I-treated lanes 2, 6, and 9 are compared to the adjacent control DNA lanes 1, 5, and 8.

Figure 9. DNase I sensitivity of viral Eco RI fragments in nuclei from lines 702-C1, 702-C2, and 702-C3.

Nuclei were prepared from transformed cells and treated with DNase I at the indicated concentrations for 5'. DNA was then purified, digested with Eco RI, separated in a 0.5% agarose gel at 1 V/cm for 65 hrs, and analysed by blotting and hybridization to a ^{32}P Eco RI-C probe.

Lanes 1 - 3, 702-C1; lanes 5 - 7, 702-C2; lanes 8 -10, 702-C3. Lanes 1 - 7 are from a single gel; lanes 8 - 11 are from a shorter exposure of a second gel.

The marker lanes 4 and 11 contained intact Ad 12 DNA and the Sal I, Bam HI, and Eco RI digests of Ad 12 DNA which were described for Figure 2 (33, 12.5, 7.9, 5.5, and 3.6 kb). Lane 4, 120 pg of each digest; lane 11, 60 pg of each.


702-C1
M
702-C2
702-C3
M
1 2 3 4 5 6 7 8 9 10 11
µg/ml DNase I
0 1.0 2.5 0 1.0 2.5 0 1.0 2.5

i. Line 702-C1: (Figure 9, lanes 1, 2, and 3)

Three new viral-specific bands are seen when Eco RI-digested DNA from DNase I-treated 702-C1 nuclei is analysed by blot hybridization to the Ad 12 probe (compare lanes 2 and 3 to lane 1). These new bands are about 25, 11, and 7 kb in size; their positions in Figure 9 are indicated by the arrows beside lane 1. The new bands in lanes 2 and 3 are a product of DNase I digestion, as they are not present in the control 702-C1 DNA in lane 1.

The DNase I-treated DNA samples analysed by Eco RI digestion and blot hybridization in lanes 2 and 3 have also been analysed by electrophoresis and blot hybridization without prior restriction endonuclease digestion (Figure 10, lane 2). No viral-specific bands are seen in this autoradiograph, so each of the DNase I-specific bands seen in the Eco RI digests (Figure 9, lanes 2 and 3) must be bounded by an Eco RI cleavage site at one end and a site or region which is extremely sensitive to DNase I at the other end. (I shall refer to these as "DNase I-sensitive sites".)

ii. Line 702-C2: (Figure 9, lanes 5, 6, and 7)

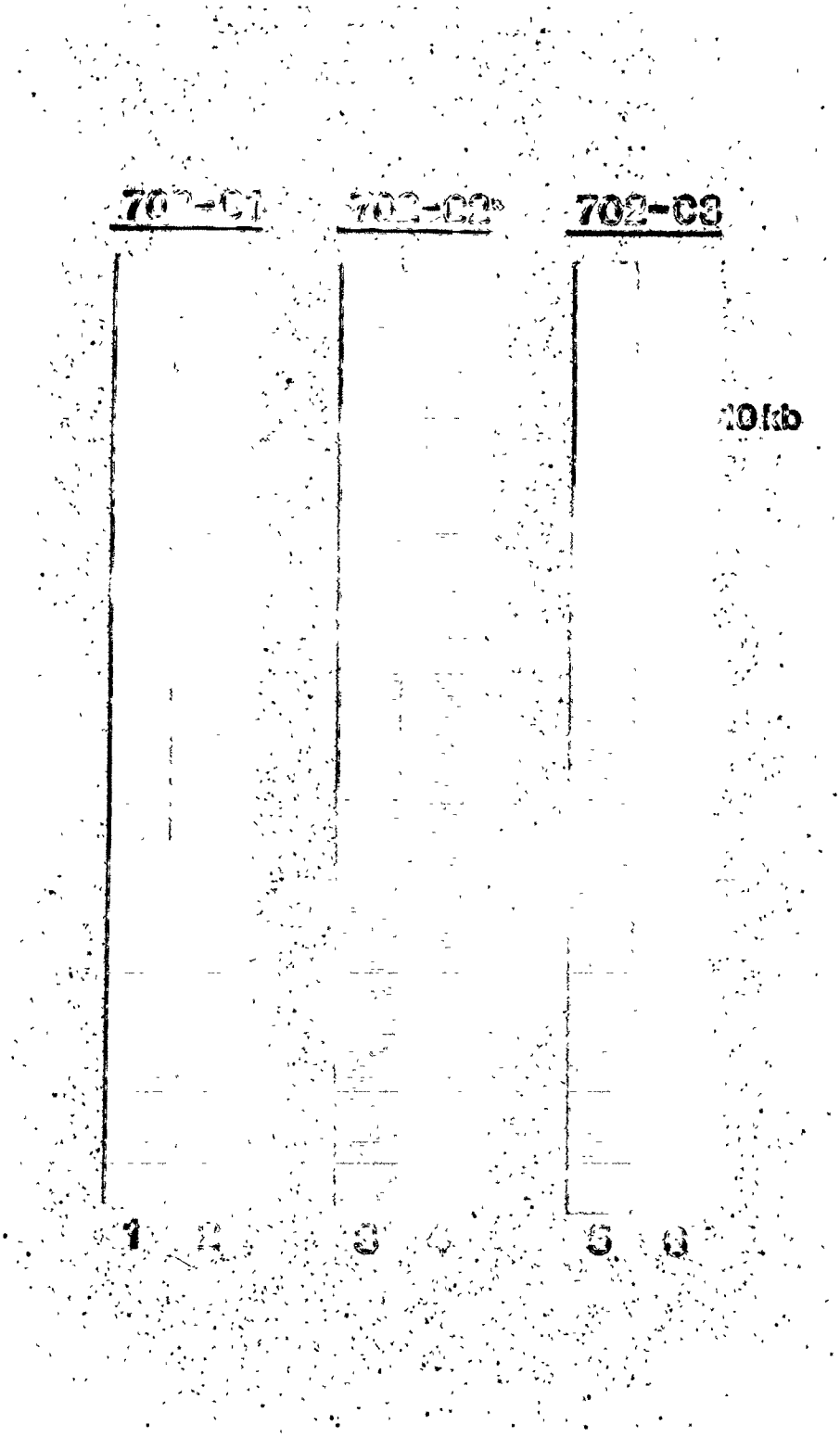
Line 702-C2 contains many large Ad 12-specific Eco-RI fragments (seen in Figure 7a) which are only poorly resolved in Figure 9 (lane 5). Resolution in lanes 6 and 7 is further hampered by the high background produced by DNase I digestion. In spite of these difficulties, the Eco RI digests of DNA from DNase I-treated nuclei in lanes 6 and 7 can be seen to contain new bands at about 12 kb and 8 kb, not present in the digest of control DNA in lane 5 (indicated by the arrows beside lane 7). As with

Figure 10. Relative DNase I sensitivities of total cellular DNA and of integrated viral sequences in transformed cell nuclei.

Nuclei were prepared from transformed cells and treated with DNase I for 5 minutes (702-C1, lanes 1 and 2, 2.5 µg DNase I/ml; 702-C2, lanes 3 and 4, 5 µg DNase I/ml; 702-C3, lanes 5 and 6, 5 µg DNase I/ml). The fragments were separated by electrophoresis in 1% agarose gels at 2 V/cm for 12 hr.

Lanes 1, 3, and 5; photographs of ethidium bromide-stained lanes.

Lanes 2, 4, and 6; autoradiograms of the same lanes after blotting and hybridization to a ^{32}P Eco RI-C probe.


702-01
702-02
702-03
10kb
1
2
3
5
6

line 702-C1, there were no well-defined viral bands when gels of unrestricted DNase I-treated DNA were analysed by blot hybridization (Figure 10, lane 4).

iii. Line 702-C3: (Figure 9, lanes 8, 9, and 10)

In line 702-C3, DNase I digestion produced sharp new bands of 24 and 13 kb, as well as several diffuse smaller bands, in the Eco RI digests in lanes 9 and 10 (indicated by arrows beside lane 8). Diffuse bands have also been seen in Bam HI digests of DNA from DNase I-digested 702-C3 nuclei, at about 7 and 9 kb (not shown).

Although DNA from cell lines 702-C1 and 702-C2 DNase I-digestion experiments showed only diffuse regions of greater hybridization when analysed by blot hybridization without restriction endonuclease digestion (Figure 10, lanes 2 and 4), DNA from DNase I-treated 702-C3 nuclei showed a well-defined band (Figure 10, lane 6). This band must result from sequence-specific cleavages by DNase I at two of the DNase I-sensitive sites, and its size, 10 kb, gives one value for the distances between such sites. Because the 10 kb fragment is larger than the 5.5 kb transforming fragment, at least one of the sensitive sites must be in cellular-sequence chromatin.

iv. Controls:

The presence of DNase I-specific bands in autoradiographs of Eco RI-digested DNAs implies that the DNase I-sensitive sites are sequence specific, and that the new viral bands seen in DNA from DNase I-digested nuclei do not result simply from cleavage of a repeating higher-order

chromatin structure, in a manner analogous to the production of nucleosome-multimer fragments by limited micrococcal nuclease digestion of chromatin (Noll, 1974). This is confirmed by the absence of bands in ethidium bromide-stained gels of unrestricted DNA from DNase I-digested nuclei (Figure 10, lanes 1, 3, and 5). Similarly, Eco RI and Hind III digestions of DNase I-treated DNAs showed only the same repetitive-DNA bands which are seen in digests of control DNAs (not shown). Thus, I conclude that the new viral bands seen in the DNase I digests in Figure 9 occur because DNase I preferentially cuts at sequence-specific positions in the chromatin at or near the viral integration sites.

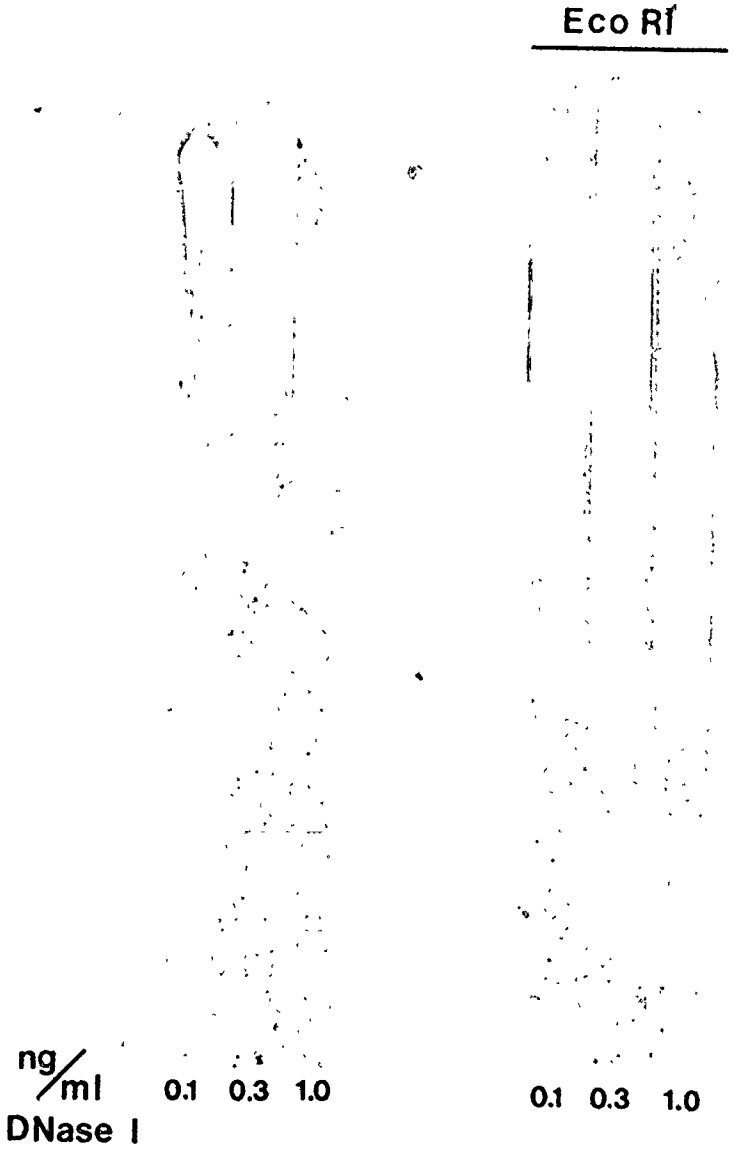
These sequence-specific cleavages by DNase I could occur because the enzyme preferentially binds to and cuts at particular DNA sequences, or because the DNA sequence determines a nucleosome or chromatin conformation which leaves the sequence exposed to the nuclease. To distinguish between these explanations, DNA was purified from 702-C3 cells and the naked DNA was digested very lightly with DNase I (0.1 - 1.0 ng DNase I/ml for 5 minutes). No new bands are seen in either Eco RI-digested or Eco RI-undigested samples (Figure 11) showing that the specific cleavages by DNase I are dependent on chromatin structure.

B. DNase I Sensitivities of Viral Eco RI Fragments:

Overall, the chromatin regions which contain Ad 12 sequences are only slightly more sensitive to DNase I than is the bulk of cellular DNA. This conclusion comes from comparisons of the size distributions of

Figure 11. DNase I sensitivities of viral sequences in purified 702-C3 DNA.

Purified DNA from 702-C3 cells was digested with DNase I for 5 minutes at the concentrations indicated in the figure and re-purified. Samples were analysed (with and without Eco RI digestion, right and left panels respectively) by electrophoresis in a 0.5% agarose gel at 0.5 V/cm for 75 hr, blotting, and hybridization to a ^{32}P Eco RI-C probe.


Eco RI
ng/ml
DNase I
0.1
0.3
1.0
0.1
0.3
1.0

a) total cell DNA (not digested with a restriction endonuclease) from DNase I-digested nuclei, in agarose gels stained with ethidium bromide (Figure 10, lanes 1, 3, and 5), and b) the viral sequences in the same gels visualized by blot hybridization and autoradiography (Figure 10, lanes 2, 4, and 6). This is consistent with the possibility that only some of the integrated viral sequences are active.

The DNase I-Eco RI blot hybridization experiment in Figure 9 was intended to show which of the integrated copies of the Eco RI-C fragment had the DNase I sensitivity characteristic of active chromatin, and which were transcriptionally inactive. Unexpectedly, the individual viral-specific Eco RI fragments seen in Figure 9 cannot be neatly classified into two categories on the basis of sensitivity or resistance to DNase I.

None of the viral Eco RI fragments (those bands which are present in the control Eco RI digests in Figure 9) appear to be highly sensitive to DNase I. Instead, the relative rates of disappearance of these bands resemble what would be produced by random fragmentation of chromatin; the high molecular weight fragments at the tops of the lanes disappear after very light digestion because of their large target sizes, while the smaller fragments persist to greater extents of digestion.

There may be several factors in these blot hybridization experiments which would tend to blur the distinction between active and inactive chromatin conformations. DNA fragments larger than 10 kb are difficult to analyse by electrophoresis and blotting, and subtle distinctions are often lost. Large DNA fragments may also be heterogeneous in chromatin/confor-

mation (e.g. a few kilobases of active chromatin flanked by inactive regions, or inactive chromatin containing a DNase I-sensitive site).

Both of these problems can be minimized by confining the blot hybridization analysis to the smaller viral DNA fragments produced by Hind III digestion. Nearly all of these fragments are between 1 and 10 kb in size. These fragments can be more reproducibly transferred to nitrocellulose, and are less likely to be heterogeneous in chromatin conformation.

C. Hind III Analysis of DNase I Sensitivities:

When Ad 12-specific Hind III fragments are displayed by blot hybridization, each fragment does not contain an entire integrated copy of the transforming DNA. Instead each integrated copy gives rise to two cell DNA-linked fragments plus a 1.2 kb internal fragment. Thus it will not be possible to determine the DNase I sensitivity or resistance of any individual copy of the Ad 12 DNA by Hind III analysis.

The extent of methylation of CCGG sites in the Hind III linker bands was established in Section IV, part D. Unmethylated CCGG sites are present in all three 702-C1 Hind III linker bands, three of the thirteen 702-C2 bands, and two of the nine 702-C3 bands. The remaining 702-C2 and 702-C3 Hind III linker bands contain only methylated CCGG sites. It may be possible to test the hypothesis that methylated DNA is transcriptionally inactive DNA, by determining whether those bands which contain unmethylated CCGG sites have a DNase I-sensitive chromatin conformation while the bands methylated at all CCGG sites are DNase I resistant.

Because the 1.2 kb band will contain DNA derived from all of the integrated copies, it must be excluded from the analysis of DNase I sensitivities. Similarly, because the methylation status of the 2.0 kb band from 702-C2 DNA is unknown it is excluded.

Nuclei from the three cell lines were given limited DNase I digestions and the remaining DNase I-resistant DNA was purified and digested with Hind III or Hind III + Hpa II. The fragments were separated by electrophoresis in vertical gels of 1% agarose, blotted, and hybridized to the ^{32}P-labelled Ad 12 Eco RI-C probe. Representative autoradiograms are shown in Figures 12a and 12b. For each cell line Hind III and Hind III + Hpa II digests of both control and DNase I-digested preparations are shown.

i. Line 702-C1: (Figure 12a, lanes 1 - 4)

In Section IV, part D, I showed that all of the Ad 12-specific Hind III fragments of line 702-C1 contain unmethylated CCGG sites (i.e. sites which can be cut by Hpa II). This is confirmed by the digests shown in lanes 2 and 1, which contain respectively Hind III and Hind III + Hpa II digests of 702-C1 control DNA. Lanes 3 and 4 contain comparable digests of DNA from nuclei digested with 2.5 µg DNase I.

The 702-C1 viral Hind III bands are widely separated, so their target sizes are not comparable; they are also of very different intensities. Because of this, comparison of relative band intensities in the control and DNase I-treated lanes does not give a very good indication of the relative DNase I sensitivities of the 702-C1 DNA fragments.

Figure 12. DNase I sensitivity of viral Hind III and Hind III + Hap II fragments in nuclei from lines 702-C1, 702-C2, and 702-C3.

Nuclei were digested for 5 minutes with DNase I at the concentrations indicated, and the DNA was purified and digested with the restriction enzymes specified below. Fragments were separated in 1% agarose gels at 2 V/cm for 12 hours, blotted, and hybridized to a ^{32}P Eco RI-C probe.

A. Left panel, 702-C1; right panel, 702-C2. Lanes 1, 4, 5, and 8, Hind III + Hpa II; lanes 2, 3, 6, and 7, Hind III. Lanes 3 and 4 contain 20 µg DNA, other lanes contain 10 µg.

B. 702-C3 DNA. Lanes 1 and 4, Hind III + Hpa II; lanes 2, 3, and 5 - 8 Hind III. Lanes 3 and 4 contain 20 µg DNA, the other lanes contain 10 µg.

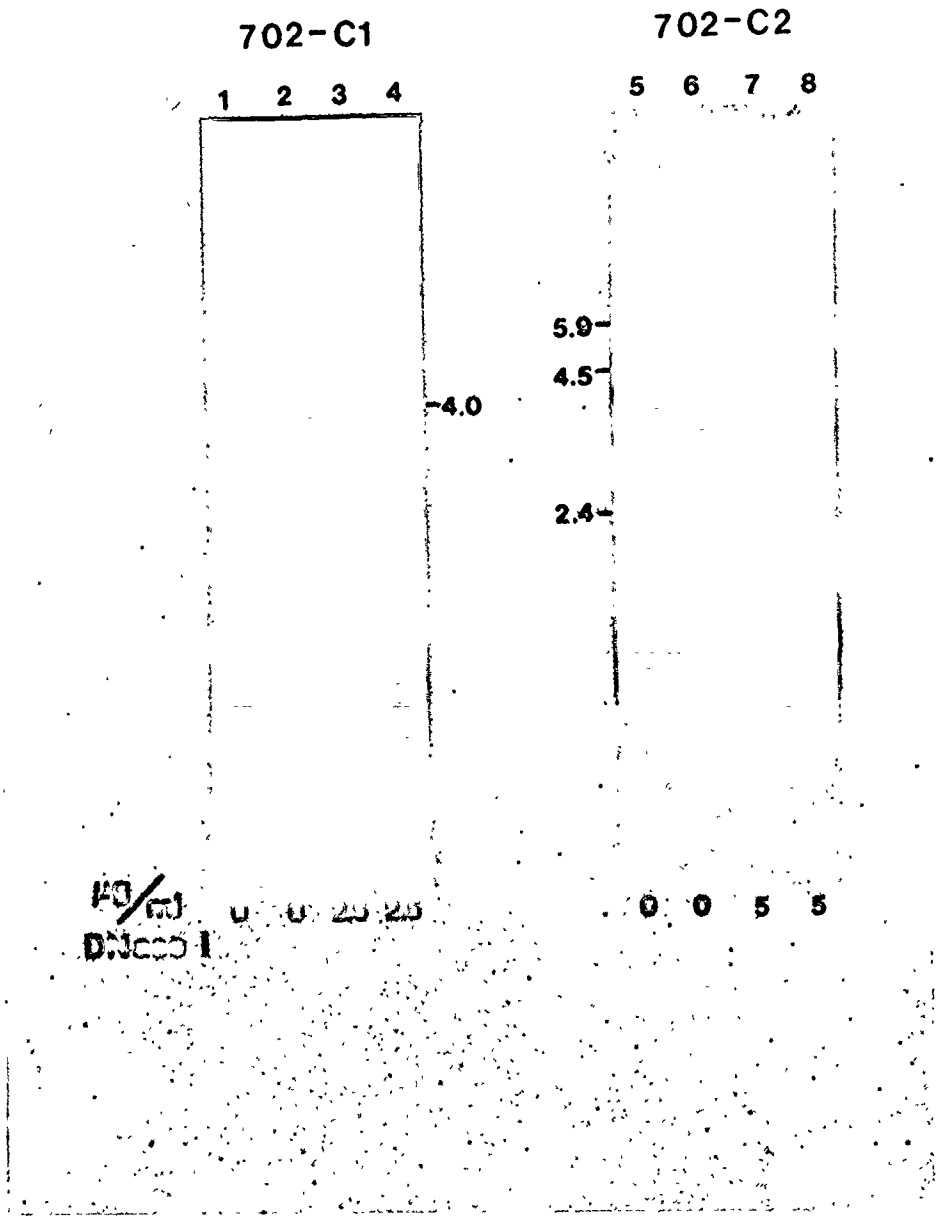


Figure 12A

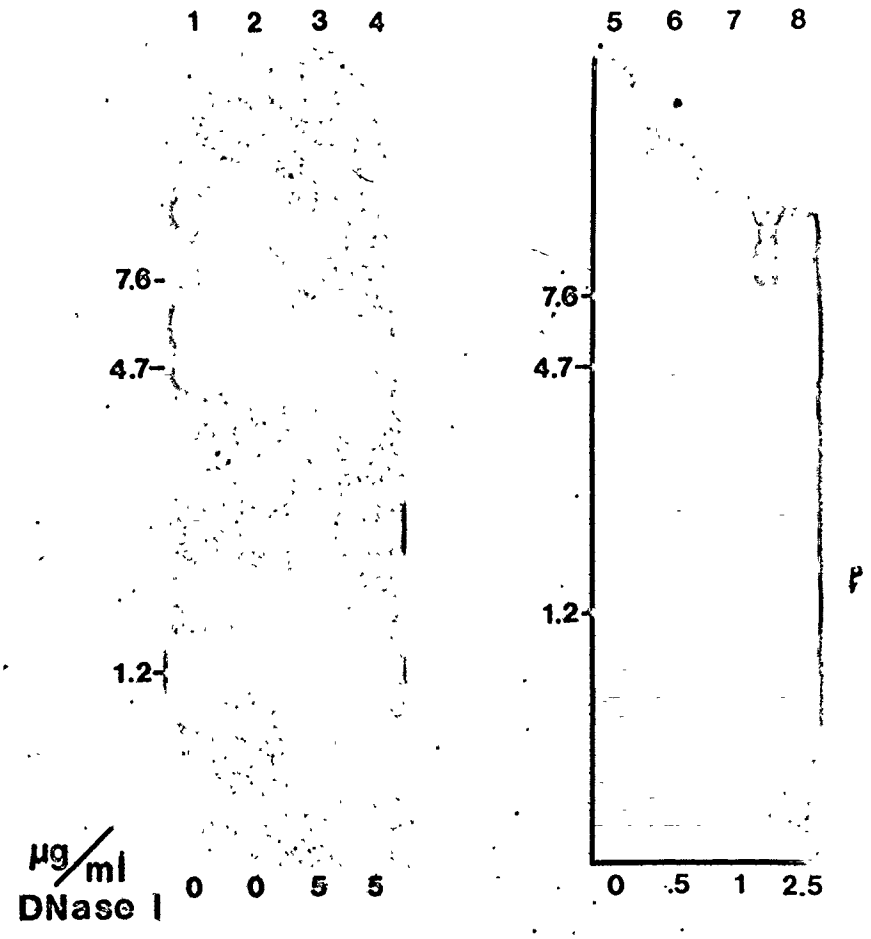


Figure 12B

There are no obvious differences in the sensitivities of the individual Ad 12-specific bands to DNase I. The only fragment that contains a methylated CCGG site is the faint 4.0 kb linker band seen in the Hind III + Hpa II digests in lanes 1 and 4. (This band is absent from the Hind III + Msp I digest shown in lane 3 of Figure 8.) This band does not show a disproportionate decrease in intensity after DNase I digestion (compare lanes 1 and 4), and I conclude that it is not significantly more DNase I resistant than the other bands.

ii. Line 702-C2: (Figure 12a, lanes 5 - 8)

I have earlier shown that most of the Ad 12-specific Hind III fragments in this cell line contain only methylated CCGG sites (i.e. no unmethylated CCGG sites); only the bands at 2.4, 4.5, and 5.9 kb are cut by Hpa II (lanes 5 and 6 of Figure 12a, see also Figure 8). In the Hind III digest of DNA from DNase I-treated 702-C2 nuclei shown in lane 7, the bands larger than 7 kb have all become very faint, as a consequence of their large target size. In less extensively DNase I-digested samples these bands show no preferential sensitivity to DNase I.

Of the bands smaller than 8 kb, only the three Hpa II-sensitive bands described above appear to be sensitive to digestion of nuclei with DNase I. The sensitivities are established by the following criteria:

2.4 kb band - In the control digest in lane 6 this band is darker than the bands immediately above and below it, and it is completely cleaved by Hpa II (lane 5). In the Hind III digest of the DNase I-digested sample, the same 2.4 kb band has become very faint, more so than the bands above

and below it. The decrease relative to the higher molecular weight band above it is very significant, because the 2.4 kb band has the smaller target size and will suffer less random cleavage by DNase I. Therefore its increased cleavage relative to the larger band must be due to preferential cleavage by DNase I of the 2.4 kb band.

5.9 kb band - In the control (lane 6) this band is equal in intensity to the band directly above it; this band is completely cut by Hpa II in lane 5. After DNase I digestion the 5.9 kb band has become much fainter than the band above it (lane 7).

4.5 kb band - The 4.5 kb band is very dark in lane 6, and is probably a double band; the intensity of this band was found to be reduced about 45% by Hpa II co-digestion (Figure 8, lanes 5 and 6). The intensity of this band appears to decrease substantially more on DNase I digestion than the intensity of the methylated band directly below it, and more than the bands above it, although this is not as marked as for the 2.4 and 5.9 kb bands. The extent of methylation of the DNase I-resistant DNA in this band in lane 7 was determined by digestion with Hind III + Hpa II (Lane 8). The band is only slightly less sensitive to Hpa II than the control band (40% reduction in intensity), so the DNase I-resistant sequences in the band are only slightly enriched in methylated sequences. This suggests that both the methylated and unmethylated components of this band may be moderately sensitive to DNase I. However, this is based on only one Hpa II analysis of the 4.5 kb band, so no firm conclusions can be drawn. (There are two dark bands (about 3.5 and 3.0 kb) in the Hind III digest in lane 6, and, in the DNase I-treated lanes 7 and 8, the lower of these

bands appears to be decreased in intensity relative to the upper band. The apparent DNase I-sensitivity of this 3.0 kb band is an artefact, seen because the autoradiogram was overexposed to bring out the faint bands in the DNase I-treated lanes.)

iii. Line 702-C3: (Figure 12b)

Hind III and Hind III + Hpa II digests of DNA from DNase I-digested 702-C3 nuclei are shown in Figure 12b, lanes 3 and 4, with similar digests of control DNA in lanes 2 and 1 for comparison. The DNase I-treated DNA analysed in lanes 3 and 4 is extensively digested, and most bands have become very faint. A range of DNase I digestions from the same experiment are shown in lanes 5 - 8.

Like line 702-C2, the viral sequences in this cell line are highly methylated. Only two of the Hind III linker bands contain unmethylated CCGG sites (4.7 and 7.6 kb; lanes 1 and 2). Only the 4.7 band is preferentially digested by DNase I. This band is poorly separated in these gels from a methylated band of 4.5 kb (compare lanes 1 and 2). Two observations suggest that the decrease in intensity of the broad 4.5 - 4.7 kb band in DNA from DNase I-digested nuclei is due to preferential digestion of the 4.7 kb band. 1) The shape and mobility of the doublet band are shifted, from the broad band averaging 4.6 kb in control digests (lanes 2 and 5) to a sharp 4,5 kb band in DNase I-treated digests (lanes 3 and 8). 2) The Hind III + Hpa II co-digestion in lane 4 shows that all of the DNase I-resistant sequences in this band are also Hpa II resistant and therefore methylated.

None of the other 702-C3 Hind III bands have shown significant and reproducible sensitivities to DNase I. In particular, although the unmethylated 7.6 kb band has occasionally appeared to be disproportionately reduced in digests of DNA from DNase I-digested nuclei, this decrease has not been reproducibly significant.

iv. DNase I-specific bands in Hind III digestions:

The DNase I-digested DNAs used for the Hind III analyses shown in Figures 12a and 12b have also been analysed by Eco RI digestion. Some of these Eco RI digests were shown in Figure 9. Although new DNase I-specific bands were seen in the Eco RI blot hybridizations of all three cell lines, new viral bands were never seen in the Hind III analyses (Figure 12).

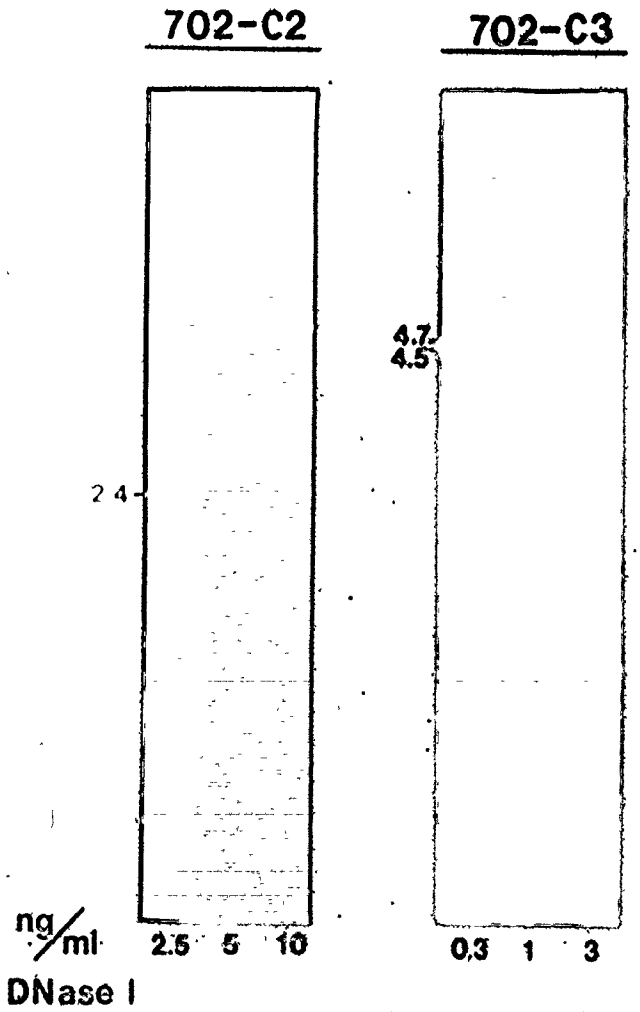
v. Control digestions:

To confirm that the DNase I sensitivities which were identified for some Hind III bands in lines 702-C2 and 702-C3 (Figure 12) were due to their chromatin conformation, purified DNA from these cell lines was digested with DNase I and analysed by Hind III digestion and blot hybridization (Figure 13). No preferential digestion of any Hind III fragments was seen in DNA from either cell line.

In particular it can be seen for line 702-C2 (left panel) that the 2.4 kb fragment remains darker than the flanking bands, while in DNA from DNase I-digested nuclei it became fainter than the adjacent bands (compare with Figure 12a, lane 7). In line 702-C3 (right panel) there is no shift in the width or average mobility of the 4.5 & 4.7 kb double band

Figure 13. DNase I sensitivities of viral Hind III integration fragments in purified DNA from 702-C3 and 702-C3 cells.

Purified DNA from 702-C2 cells (left panel) and 702-C3 cells (right panel) was digested for 5 minutes with DNase I at the concentrations indicated on the figure, re-purified, and digested with Hind III. The fragments were separated by electrophotesis in 1% agarose gels for 14 hr (left panel 1.5 V/cm, right panel 2 V/cm), blotted, and hybridized to a ^{32}P Eco RI-C probe.


702-C2
702-C3
4.7
4.5
2.4
ng/ml
2.5
5
10
0.3
1
3
DNase I

on increasing DNase I digestion, and it remains substantially darker than the bands above it. This should be compared to lanes 3 and 8 of Figure 12b, where the 4.7 kb band has been preferentially cleaved by DNase I.

In summary, DNase I digestions have shown that there are differences between the chromatin conformations of methylated and unmethylated viral integration fragments in cell lines 702-C2 and 702-C3. Of the total of five Hind III bands containing unmethylated CCGG sites which were seen in the two cell lines, four were clearly sensitive to DNase I (the 2.4, 4.5, and 5.9 kb bands in line 702-C2, and the 4.7 kb band in line 702-C3). Of the sixteen Hind III bands totally methylated at CCGG sites, fifteen were DNase I resistant. (The DNase I sensitivity of the two non-conforming bands (4.5 kb in 702-C2 and 7.6 kb in 702-C3) was equivocal, and the methylation of the 2.0 kb band in line 702-C2 could not be determined because of co-migrating fragments.) The implications of this correlation will be considered in the Discussion.

DISCUSSION

The experiments described in the preceeding section were designed to investigate the methylation of viral DNA in Ad 12-transformed cells. Before the patterns of viral DNA methylation could be determined, it was necessary to find out how the viral DNA sequences were arranged in the cellular DNA.

I. Patterns of Integration

Blot hybridization analysis of DNA from line 702-C3 with restriction endonucleases which do not cut within the viral DNA (Eco RI, Bam HI, Bgl II) yielded four bands with intensities corresponding to one copy each of the transforming DNA (Figure 2, lanes 4 - 6). Similar analysis of line 702-C2 DNA with Eco RI gave nine such bands (Figure 7, lane 1). In the following discussion I have made the working assumption that each of these bands contains a more-or-less complete copy of the Ad 12 Eco RI-C fragment used in the original transformation experiment (Mak et al, 1979). The intensities of the Ad 12 Hind III and Msp I internal bands (those at 1.2 and 2.35 kb in the appropriate digests in Figures 5 - 8) support this hypothesis, and the conclusions I have drawn are not critically dependent on it.

In line 702-C1, viral DNA equivalent to three or four copies of the transforming fragment is present in a single band after digestion with

each of three restriction enzymes (Figure 2, lanes 1 - 3). Chance co-migration of three or four unrelated viral integration fragments in each case is unlikely. Two tandemly integrated copies might be present in a single integration fragment, but the Bgl II multiple copy band (13 kb) is not large enough to contain in a single DNA fragment three tandem copies of the 5.5 kb transforming fragment (16.5 kb of viral DNA). An alternative explanation is that a single copy of the viral DNA originally became integrated at this site and that the region containing the viral DNA (or the entire chromosome) was subsequently amplified. This has been suggested for Ad 5 sequences in transformed rat cells (Dorsch-Hasler et al, 1980), and for Ad 12 DNA in transformed hamster cells (Sutter et al, 1978).

In each cell line there are also bands whose intensity is less than that expected for a complete copy of the transforming DNA. The minor band in line 702-C1 must contain only a partial copy, as it is smaller than the 5.5 kb Eco RI-C fragment. This is likely the case for the minor bands in the other two lines as well. In most cases I have not considered these minor integration sites in the ensuing analysis. They may be responsible for some of the cell DNA-viral DNA linker bands seen on Hind III and Msp I digestion, but this will not affect the conclusions I have drawn.

There is no evidence in the experiments I have described for preferential or specific integration sites. In this regard, and in the identification of multiple copies and of fragmented copies, my findings agree with other studies of adenovirus-transformed cells and papova virus transformed cells (Sutter et al, 1978; Sawada et al, 1979).

II. Patterns of Methylation

I have used blot hybridizations of DNA digested with the methylation-sensitive restriction endonuclease Hpa II to detect methylated CCGG sites in and around the viral sequences. Although Hpa II can sample only a small subset of the potential CpG methylation sites, the availability of its methylation-insensitive isoschizomer Msp I provides a powerful tool for the analysis of methylation at specific sites in the genome.

By comparison of Hpa II and Msp I digestion patterns of viral DNA I have confirmed that none of the CCGG sites are methylated in Ad 12 DNA extracted from virions. Thus all of the methylation detected at viral CCGG sites in the transformed cells can be attributed to cellular mechanisms which have initiated methylation at previously unmethylated sites.

Comparisons of Hpa II and Msp I digestion patterns (singly and in double digestions with Eco RI, Bam HI, and Hind III) of the integrated viral sequences in cell lines 702-C1, 702-C2 and 702-C3 leads to the following conclusions:

1. No viral CCGG sites are methylated in line 702-C1, because the Hpa II and Msp I patterns are almost identical (Results IV-A-i; Figure 5 lanes 1 and 2).

2. More than half of the viral CCGG sites are methylated in lines 702-C2 and 702-C3, because more than half of the viral DNA cannot be cut by Hpa II, although it is cut by Msp I (Results IV-C, IV-A-ii; Figure 7, lanes 3 and 4; Figure 5, lanes 4 and 5).

3. At some integration sites in lines 702-C2 and 702-C3 all of the viral CCGG sites are methylated, because there are copies of the viral transforming fragment (Eco RI and Bam HI integration fragments) which cannot be cut by Hpa II but are cut by Msp I (Results IV-C, IV-B-ii; Figure 7, lanes 6 and 7; Figure 6, lanes 6 and 7).

4. At other integration sites in lines 702-C2 and 702-C3 the integrated viral DNA must contain both unmethylated and methylated CCGG sites. This information comes from the Hind III analysis presented in Results IV-D (Figure 8). The large number of Hpa II-resistant Hind III fragments in these cell lines implies that those integrated copies of the viral DNA which contain some unmethylated CCGG sites (Eco RI and Bam HI integration fragments which are cut by Hpa II) must give rise on Hind III cleavage to cell DNA-linked fragments which contain only methylated CCGG sites.

The observation of this variety of methylation patterns allows evaluation of some of the regulatory mechanisms which might be postulated to control methylation of particular DNA sequences. I will list some of these hypotheses and then discuss them in the light of the patterns described above.

1. Methylation might be determined by the primary DNA sequences at the methylation site.

2. Methylation might depend on the presence or absence of either sequence-specific or non-specific cellular methylases, which might in turn depend on the differentiated state of the cell.

3. Methylation of integrated viral DNA might be a direct or indirect consequence of cellular sequences near the viral integration site.

4. Transcription might prevent methylation.

The hypothesis that methylation of integrated viral DNA is determined solely by the DNA sequence at the methylation site is easily disproved. All of the integrated viral CCGG sequences have remained unmethylated in line 702-C1. The identical CCGG sequences have become methylated at some viral integration sites in lines 702-C2 and 702-C3 (for example, the copy of the transforming fragment integrated in the 11 kb Bam HI fragment of line 702-C3 (Figure 6, lane 10) is methylated at all viral CCGG sequences, as it is entirely resistant to Hpa II). Therefore methylation at viral CCGG sequences cannot be determined solely by the DNA sequence at the methylation site.

The patterns of methylation I have described cannot be completely explained by the presence of different cellular methylases in the different cell lines, as identical CCGG sites are methylated at some integration sites, and unmethylated at other integration sites, within a single cell line. For example, each of the nine major integration fragments in line 702-C2 includes a viral DNA region defined by CCGG sites at 1 kb and 3.35 kb from the left end of the Ad 12 genome (Figure 7, lanes 1 and 6). In the Eco RI + Hpa II digest shown in Figure 7, lane 7, some of these CCGG sites have been cut by Hpa II to give the 2.35 kb fragment whereas in three (or four) of the integration fragments (arrows) the same CCGG are blocked by methylation. Thus, although the cell is apparently capable of methylating all CCGG sequences in the transforming fragment, some of these are

not methylated at some integration sites.

Copies of the Ad 12 DNA at different integration sites (i.e. separate bands in Bam HI and Eco RI digests) are differently methylated. This suggests that DNA sequences and/or chromatin structures in the cellular DNA around the integration sites may influence the methylation of nearby CpG sequences in integrated viral DNA, as suggested in the third hypothesis above.

The double digestion and high resolution experiments described in Results sections IV-A, -B, and -C provide some information about methylation of the cellular sequences around the viral integration sites. In Hpa II digests, most of the Hpa II-resistant viral-associated DNA (those Hpa II DNA fragments which hybridize to the probe and which can be further cut by Msp I) has molecular weights of 20 - 50 kb. The CCGG sites should occur in cellular DNA on average about once in every 3 kb, so the 20 - 50 kb Hpa II fragments probably contain 5 - 15 methylated CCGG sites. The distribution of CCGG sites about once every 3 kb in cellular DNA is confirmed by the sizes of the Msp I linker fragments, which are mostly between 1 and 6 kb (Figures 5 and 7). Thus the large sizes of the Hpa II-resistant viral-associated DNA fragments indicates that the viral DNA is covalently linked to extensively methylated cellular sequences.

In particular, the totally methylated copies of the transforming fragment must be flanked by methylated cellular DNA, by the following reasoning. If totally methylated viral DNA copies were flanked by unmethylated cellular DNA, Hpa II codigestion of the Eco RI and Bam HI integration fragments would produce new fragments, each containing a complete copy of the Eco RI-C transforming fragment. However, none of the DNA

fragments created by Hpa II cleavage of the Eco RI and Bam HI integration fragments of line 702-C3 (Figure 6, lanes 6 - 10), or by cleavage of the Eco RI fragments of line 702-C2 (Figure 7, lanes 6 and 7) are large enough (bigger than 5.5 kb) and intense enough (by comparison with control digests) to contain a complete copy of the transforming fragment. Therefore the totally methylated integrated copies of the viral DNA in lines 702-C2 and 702-C3 are probably flanked by methylated cellular DNA.

In most cases it is not possible to determine whether those copies of the viral DNA which contain unmethylated CCGG sites (the Eco RI and Bam HI fragments cut by Hpa II) are flanked by methylated or unmethylated cellular sequences. For line 702-C1 it is likely that the only methylated site associated with the integrated viral DNA is in a cellular sequence, since the loss of intensity of the methylated 20 kb band on Msp I digestion is about equal to the increase in intensity of the 1.9 kb band. For lines 702-C2 and 702-C3 there are Hpa II linker fragments which are not further cut by Msp I, implying that they are flanked by an unmethylated cellular CCGG site. There are other linker fragments, for example the 17, 15.5, and 14 kb Bam HI + Hpa II fragments seen in Figure 6, lane 10, which must contain methylated cellular CCGG sites.

Bird et al (1979) have described for Echinus regions of heavily methylated DNA, resistant to the CpG enzymes Hpa II, Hha I, and Ava I, which make up 40% of the total cellular DNA and are from 15 kb to more than 50 kb in length. More recently they have extended this analysis and suggest that most of the Hpa II-resistant DNA of vertebrates (50 - 80% of total DNA) also consists of regions where most or all CpGs are methylated, regardless of flanking sequences (Bird and Taggart, 1980).

The heavily methylated Hpa II-resistant copies of the viral fragment may have integrated into such regions. The integrated viral DNA might then have come under the influence of a cellular mechanism which determines the methylation of a region of DNA, independent of at least some of the sequences which actually become methylated. There may be signals for this methylation in the DNA sequences which delineate the regions, or there may be determinants of higher-order chromatin structure which indirectly influence methylation of regions of DNA.

Integration of transforming viral DNA into cellular methylation domains could explain the presence of totally methylated copies of the viral DNA flanked by regions of methylated cellular DNA. The totally unmethylated copies in line 702-C1 could also be explained in this way, by being integrated outside of such domains. However this model cannot account for the presence of partly-methylated copies of the transforming fragment, which are seen in lines 702-C2 and 702-C3. These partly-methylated copies may be best explained by the hypothesis (number 4 above) that transcription prevents methylation.

Correlations between transcriptional activity and the absence of DNA methylation have been reported for both integrated viral sequences and cellular genes. Sutter and Doerfler (1980) have investigated methylation patterns of viral DNA in Ad 12-transformed hamster and rat cells. They report a general inverse correlation between the levels of methylation of specific DNA segments and the extent to which these segments are expressed as mRNAs. The sequences complementary to the Ad 12 Eco RI-C fragment were much less methylated than sequences hybridizing to Eco RI-A and -B fragment probes, which detect only late transcription units. In two lines of trans-

formed rat brain cells which expressed some "late" regions, the Eco RI-B sequences containing this late region were undermethylated. The cell lines all contained 10 - 30 copies of the viral genome per cell (Fanning and Doerfler, 1976) and no attempt was made to discriminate between active and inactive copies.

Methylation of viral DNA in retrovirus-transformed cells has also been reported. Guntaka et al (1980) reported that avian sarcoma provirus sequences are extensively methylated in non-producer cells, but under-methylated in cells which transcribe high levels of these sequences. Similarly, Cohen (1980) described extensive methylation of Hpa II and Hha I sites in endogenous (inactive) mouse mammary tumor virus provirus DNA, but no methylation of identical but active exogenously acquired provirus. For exogenously acquired viral sequences there appears to be a strong inverse correlation between DNA methylation and gene expression. The relationship between methylation and expression of cellular genes may be more complex. Studies of the tissue-specificity of methylation patterns have been hampered by the presence of multiple cell types in most tissues.

A detailed study of cytosine methylation at 17 CpG sites in the human γδβ-globin locus found only low-level methylation of transcribed globin sites in tissues expressing the globin genes, while flanking non-transcribed regions were methylated more extensively (van der Ploeg and Flavell, 1980). Levels of methylation of most sites were higher in tissues which do not express these genes. Similar results have been obtained for Hha I and Hpa II sites in the ovalbumin gene cluster (Mandel and Chambon, 1979). Both groups were also able to identify a few sites which were un-methylated or under-methylated in tissues which did not express these genes

at all. A contrary finding is that, although the active ribosomal genes are predominantly unmethylated in all tissues of most vertebrates, they are highly methylated in the tissues of fish and amphibia (Bird and Taggart, 1980).

The studies described above show a number of correlations between transcriptional activity and the absence of DNA methylation, but they do not clarify which is the cause and which the effect. The possibility that methylation is a cellular mechanism leading to gene inactivation is very appealing, but it does not explain what causes methylation. The alternate hypothesis, that only inactive sequences are substrates for cellular methylases, sheds no light on how genes may be turned on or off.

A causal relationship between transcription and lack of methylation is compatible with most but not all of the viral methylation patterns I have found. This hypothesis could explain both the partly-methylated copies of the transforming fragment and the methylation of flanking cellular sequences in lines 702-C2 and 702-C3. It is also possible to account for the totally methylated viral copies in these cell lines. However the hypothesis is not able to explain why there is no methylation of viral CCGG sites in line 702-C1. I give these arguments in some detail below.

Transcription in the 702 transformed cell lines has been investigated (Mak et al, 1979). Lines 702-C2 and 702-C3 contain RNA transcripts complementary to the leftmost 3.5 kb of the Ad 12 Eco RI-C fragment (Ad 12 Hind III-G and -I fragments) which encode the Ad 12 early transcription units E1a and E1b (Sugisaki et al, 1980). These transcripts correspond to the leftward Hind III cell DNA-viral DNA linker fragments and to the internal Ad 12 1.2 kb Hind III fragments produced by digestion of integrated Ad 12 DNA in

these cell lines.

Given these transcription patterns and the reported correlations between expressed regions and lack of methylation, it is tempting to speculate that in the partly-methylated viral copies integrated in lines 702-C2 and 702-C3 it is the left hand Hind III linker fragments and the 1.2 kb Hind III fragment, which contain the transforming genes, which are active and unmethylated, while the rightward Hind III linker fragment (which is inactive) is methylated. This could be tested by hybridization of a probe for these right hand linkers (Hind III-F) to Hind III and Hind III + Hpa II digests of 702-C2 and 702-C3 DNAs, similar to those shown in Figure 8. The above model predicts that this probe will not hybridize to any of the unmethylated Hind III linker fragments in these cell lines.

Methylation of all inactive DNA sequences can also explain the presence of totally methylated copies of the integrated viral fragment, if these copies have become integrated in such a manner that their transcription is not possible. For example, the viral DNA at some integration sites might have undergone deletions or rearrangements during the integration process or subsequently. For example, a deletion of 300 - 500 base pairs at the left end of the Eco RI-C fragment would eliminate the promoters for the transformation genes in region Ela (Sugisaki et al, 1980), and a deletion of this size would not have been detected in the experiments I have done. If sequences which could not be transcribed were automatically methylated by some cellular enzyme, such a deletion might cause the complete methylation of the viral sequences at that integration site. Other copies which had retained the promoters would be unmethylated in the transcribed region.

I have not attempted detailed mapping of the viral sequences present at each integration site, since the presence of multiple copies in each cell line makes such analysis very difficult. It would be possible to determine whether the Ela promoter region is present at each site, by hybridizing a probe for the extreme left end of Ad 12 (e.g. the Kpn I-G fragment, leftmost 500 bp) to nitrocellulose blots of transformed cell DNA digested with Eco RI, Bam HI, or other enzymes which do not cut within the transforming fragment. A great deal more information would be obtained by cloning the integrated copies with the flanking cellular sequences and determining the precise sequences and arrangements present at each site.

Although the methylation patterns of lines 702-C2 and 702-C3 are compatible with methylation of all inactive viral and cellular sequences, there is no evidence from the integration and methylation patterns described that the methylated sequences are not in fact transcriptionally active. In addition, it is particularly difficult to reconcile the transcription model to the lack of methylation seen in line 702-C1.

In line 702-C1 only the Ela genes in the viral Hind III-G fragment (leftmost 2.3 kb of the transforming fragment) are transcriptionally active (Mak et al, 1979). If all non-transcribed sequences were methylated, as predicted by hypothesis 4, the viral Hpa II sites at 3.35 and 3.95 kb would be methylated. This is clearly not the case, as the 2.4 kb and 0.6 kb viral fragments defined by these cleavage sites are produced by Hpa II digestion of 702-C1 DNA (Results IV-A-i; Figure 5, lane 1).

Some of the viral sequences in line 702-C1 are both unmethylated

and transcriptionally inactive. It might be that these sequences are only transcribed at a very low frequency, such that the transcripts cannot be detected by RNA:DNA hybridization experiments. It is also possible that the methylation patterns which have been described here and in the literature reflect not transcription per se, but a pre-transcriptional activation characteristic of potentially active genes. To clarity these possibilities, I have investigated the chromatin conformations of the integrated viral fragments by limited digestions with DNase I.

III. DNase I Analysis

The results show a strong correlation between enhanced sensitivity to DNase I and the absence of methylation in the viral-specific Hind III fragments of lines 702-C2 and 702-C3. Unfortunately the phenomenon of site-specific cleavages by DNase I may prevent this DNase I sensitivity being used as an indicator of transcriptional activity of the DNA therein.

DNase I discriminates between the conformations of the nucleosomes of active and inactive chromatin; the preferential degradation of active DNA sequences in limited DNase I digestions of nuclei has been well documented (Weintraub and Groudine, 1976; Weisbrod and Weintraub, 1979). As described above, viral RNA transcripts have been characterized in all three cell lines (Mak et al, 1979). As active integrated viral genes have the same DNase I sensitivity as active cellular genes (Flint and Weintraub, 1977), at least one copy of these sequences is expected to be in a DNase I-sensitive chromatin conformation in each of the cell lines.

The chromatin conformations of each of the integrated copies could

not be determined from Eco RI blot hybridization analysis of DNA from DNase I-digested nuclei. Distinctions between "sensitive" DNA fragments and "resistant" fragments could not be established. There may be a number of reasons for this. Technical barriers include the difficulties of resolving very large DNA fragments in agarose gels, and the vagaries of the blot hybridization technique. There are other non-technical factors which may blur or block the distinctions between different chromatin conformations. One again arises from the large sizes of the Eco RI fragments which were analysed. In their study of the DNase I sensitivity of integrated Ad 5 sequences, Flint and Weintraub (1977) found that the DNase I-sensitive chromatin structure was restricted to the transcribed DNA, and did not extend into flanking viral or cellular regions. Similarly in two investigations the high DNase I sensitivity of the globin genes was found to be limited to the coding regions, with flanking sequences having only low-to-moderate sensitivity (Stalder et al, 1980; Zasloff and Camerini-Otero, 1980). In the cells I have studied it is quite likely that a 30 kb viral-specific Eco RI fragment might consist of only 3 kb of DNase I-sensitive chromatin (transcribed Ad 12 sequences), and 27 kb of DNase I-resistant cellular and viral chromatin. Unless the level of discrimination by DNase I is very high (early studies using reassociation kinetics analysis reported only a 6-fold depletion of the active sequences), detection of short DNase I-sensitive regions in large DNA fragments will not be possible, because most of the fragment will be DNase I resistant.

Blot hybridization analysis of the DNase I sensitivity of the smaller viral Hind III fragments was able to circumvent these problems and reveal a correlation between the absence of methylation at CCGG sites

in Hind III bands and sensitivity of the bands to DNase I (Results V-C). The correlation is not absolute. The viral-specific 7.6 kb band in line 702-C3 contains one or more unmethylated CCGG sites, but the intensity of this band does not appear to be significantly reduced in DNA from DNase I-digested nuclei (Figure 12b). On the other hand, the Hpa II-resistant 4.5 kb viral band in Hind III-digested 702-C2 DNA may be moderately sensitive to DNase I (Figure 12a). The correlation in the two cell lines is nonetheless strong. Of the five Hind III bands containing unmethylated sites, four are clearly differentially sensitive to DNase I. Of the sixteen methylated bands, only one may be sensitive.

The finding of new DNase I-specific sub-bands in Eco RI digests of DNase I-treated DNA from all three cell lines makes it difficult to determine the significance of the preferential disappearance of unmethylated Hind III bands.

The new bands did not arise from cleavages by DNase I at regularly repeating sites in chromatin; no bands were seen in total ethidium bromide stained DNA. The cleavages by DNase I therefore must occur at specific DNA sequences in the chromatin. DNase I did not preferentially cut at these sequences when purified DNA was the substrate. The preferential cleavages consequently must arise because the DNA sequences determine chromatin structures in which these sites can be preferentially cut by DNase I.

The nature of the DNase I cleavage sites which give rise to the Eco RI-DNase I bands is not known. One possibility is that these "sites" are the boundaries between active DNase I-sensitive regions and inactive DNase I-resistant regions. If so, the new bands which appear in Eco RI

digests and in unrestricted DNA consist solely of inactive DNase I-resistant sequences, and the disappearance of specific fragments in Hind III digests can be attributed to transcriptional activity of the encoded genes. This explanation is supported by the finding that the DNase I sensitivity of integrated Ad.5 DNA was limited to the transcribed region with a precision of only a few nucleosomes or about 500 base pairs (Flint and Weintraub, 1977).

Although it is possible that the DNase I-specific new bands seen in Figure 9 result from preferential degradation of active nucleosomes by DNase I, there are several reports of sequence-specific cleavages of chromatin by DNase I which do not result from discrimination between active and inactive nucleosomes (Wu et al, 1979a, 1979b; Stalder et al, 1980).

Wu and coworkers have reported that very limited DNase I digestion of Drosophila chromatin produce new bands in restriction enzyme digests analysed by blot hybridization with a number of cloned probes from a Drosophila genome library (Wu et al, 1979b). Some bands were also seen when unrestricted DNA was analysed with the same probes. Well-defined bands were seen when an inactive heat shock gene was analysed, suggesting that the sequence-specific cleavages can occur in inactive chromatin (Wu et al, 1979b).

Site-specific cleavages of chromatin by DNase I have been mapped around the globin coding regions (Stalder et al, 1980). These sites are not the boundaries of DNase I-sensitive active regions. A similar DNase I-hypersensitive site was identified by Kuo et al (1979) in oviduct chromatin, several kb downstream from the 3' end of the conalbumin gene.

The nature of these DNase I-hypersensitive sites is entirely open to speculation. Varshavsky and coworkers have identified a region of the

SV40 minichromosome which lacks nucleosomes and has a high sensitivity to DNase I (Varshavsky et al, 1979; Saragosti et al, 1980). This region includes both the origin of viral DNA replication and the promoter for the late viral transcription unit. The DNase I cleavage sites around the globin genes showed some tissue specificity (Stalder et al, 1980); some sites were not present in chromatin of brain cells or of a chicken cell line, and some sites differed between embryonic and adult red cell chromatin. The possibility that some DNase I-hypersensitive sites may be tissue-specific origins of replication ties in nicely with a study of the segregation of nucleosomes which suggested that a relationship might exist between active transcription units and the direction of DNA replication (Siedman et al, 1979).

The kinetics of appearance and disappearance of the DNase I bands reported by Stalder et al (1980) resemble those of the new bands I have described. In both cases the bands appeared after very light DNase I digestion, and were slightly more resistant to further DNase I degradation than the major (restriction endonuclease specific) bands.

If such site-specific cleavages are responsible for the new bands in the Eco RI analysis, they may also be the cause of the preferential disappearance of some or all of the DNase I-sensitive Hind III fragments. As the cleavages reported in the literature may occur outside of coding regions and in transcriptionally inactive DNA, the sensitivities of the Hind III fragments may be unrelated to transcription.

The experiments I have described in this report do not allow me to distinguish between the alternative explanations for the appearance of new DNA bands after DNase I cleavage of chromatin. Therefore it is not possible to unequivocally attribute the DNase I cleavages of the unmethylated Hind

III fragments to an "active" chromatin conformation.

The correlation I have observed between DNase I sensitivity and unmethylated CCGG sites may be due to an active chromatin conformation associated with unmethylated DNA, or due to a relationship between DNase I-hypersensitive cleavage sites and unmethylated DNA. It may not be coincidence that the DNase I hypersensitive site which has been identified in the conalbumin region is closely linked to unmethylated Hha I and Hpa II sites (Kuo et al, 1979).

IV. Summary

I have used hybridizations of a ^{32}P-labelled Ad 12 DNA fragment probe to "Southern blots" of restriction endonuclease digested DNA from Ad 12 transformed cell lines to study the methylation of the integrated viral DNA in these cells. All three cell lines were found to contain multiple copies of the transforming fragment (702-C1, 3 - 4; 702-C2, 9; 702-C3, 4) at different sites in the cell genomes, with partial copies at additional integration sites.

The methylation-sensitive restriction endonuclease Hpa II and its methylation-insensitive isoschizomer Msp I were used to investigate methylation of viral CCGG sequences at the various viral integration sites, with the following results. 1) No viral CCGG sequences are methylated in line 702-C1. 2) More than half of the viral CCGG sequences in lines 702-C2 and 702-C3 are methylated. 3) At some integration sites in lines 702-C2 and 702-C3 all of the viral CCGG sequences are methylated. 4) At other integration sites in lines 702-C2 and 702-C3, the integrated viral DNA contains

both methylated and unmethylated CCGG sequences.

The significance of these methylation patterns is not clear. I have considered several possible ways in which cellular processes could determine viral methylation patterns; none is compatible with all of the patterns I have seen. In particular, methylation is not determined solely by the DNA sequence at the methylation site, nor by the presence or absence of cellular methylases. It is difficult to see how cellular DNA sequences flanking the viral integration sites could give rise to the variety of methylation patterns seen. The total absence of viral DNA methylation in line 702-C1 shows that not all inactive sequences are methylated.

Nonetheless, the DNase I analysis of chromatin conformations shows that unmethylated regions of DNA have chromatin structures which are more sensitive to DNase I than is the chromatin of methylated sequences. In addition, the published reports of the methylation patterns of transcriptionally active and inactive DNA sequences show strong correlations between DNA transcription and the absence of methylation.

It is possible that the observed patterns of viral DNA methylation in the 702 cell lines do not result from cellular regulatory mechanisms at all. Integration of viral DNA may temporarily disrupt controls around the integration site, so methylation of unmethylated sequences occurs at random. If extensive methylation leads to gene inactivation, only those cells which reestablish regulation of methylation before all the transforming genes are inactivated will give rise to transformed cell lines such as the ones I have studied.

This explanation can account for all of the methylation patterns in the three cell lines, as follows. In line 702-C1 integration of the

transforming fragment did not disrupt control mechanisms, so no methylation of the viral DNA took place. In the cells which gave rise to lines 702-C2 and 702-C3, methylation of the integrated viral and flanking cellular sequences occurred at random until all of the DNA at some integration sites was methylated, while some or all of the sequences at other sites escaped methylation until normal regulation of methylation was reestablished. The cells in which one or more copies of the transforming genes remained unmethylated gave rise to the transformed cell lines.

Inactivation of exogenously acquired DNA by random methylation is compatible with most of the published reports on methylation of viral DNA in transformed cells (Sutter and Doerfler, 1980; Guntaka et al, 1980), as selection for the transformed phenotype would have ensured that unmethylated active copies of the viral genes were present.

This explanation gives DNA methylation a central role in determing patterns of gene inactivation. Unfortunately it does not shed any light on the cellular mechanisms which determine the patterns of methylation.

REFERENCES

Adams, R.L.P., E.L. McKay, L.M. Criag and R.H. Burdon (1979). Mouse DNA methylase: methylation of native DNA. Biochem. Biophys. Acta. 561, 345-357.

Bird, A.P. (1978). Use of restriction enzymes to study eukaryotic DNA methylation: II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. J. Mol. Biol. 118, 49-60.

Bird, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. Nucl. Acids Res. 8, 1499-1504.

Bird, A.P. and E.M. Southern (1978). Use of restriction enzymes to study eukaryotic DNA methylation: I. The methylation pattern in ribosomal DNA from Xenopus laevis. J. Mol. Biol. 118, 27-47.

Bird, A.P. and M.H. Taggart (1980). Variable patterns of total DNA and rDNA methylation in animals. Nucl. Acids Res. 8, 1485-1497.

Bird, A.P., M.H. Taggart and B.A. Smith (1979). Methylated and unmethylated DNA compartments in the sea urchin genome. Cell 17, 889-901.

Bloom, K.S. and J. Anderson (1978). Fractionation of hen oviduct chromatin into transcriptionally active and inactive regions after selective micrococcal nuclease digestion. Cell 15, 141-150.

Burdon, R.H. and R.L.P. Adams (1969). The in vivo methylation of DNA in mouse fibroblasts. Biochim. Biophys. Acta. 174, 322-329.

Carpenter, B.G., J. Baldwin, E. Bradbury and K. Ibel (1976). Organisation of subunits in chromatin. Nucleic Acids Res. 3, 1739-1746.

Chambon, P. (1977). Summary: The molecular biology of the eukaryotic genome is coming of age. Cold Spring Harbour Symp. Quant. Biol. 42, 1209-1234.

Cohen, J. Craig (1980). Methylation of milk-borne and genetically transmitted mouse mammary tumor virus proviral DNA. Cell 19, 653-662.

Dawid, I.B., D. Brown and R. Reeder (1970). Composition and structure of chromosomal and amplified ribosomal DNAs of Xenopus laevis. J. Mol. Biol. 51, 341-360.

Desrosiers, R.C., C. Mulder and B. Fleckenstein (1979). Methylation of Herpesvirus saimiri DNA in lymphoid tumor cell lines. Proc. Natl. Acad. Sci. U.S.A. 76, 3839-3843.

Dorsch-Hasler, K., P.B. Fischer, I.B. Weinstein and H.S. Ginsberg (1980). Patterns of viral DNA integration in cells transformed by wild type of DNA-binding protein mutants of adenovirus type 5 and affect of chemical carcinogens on integration. J. Virol. 34, 305-314.

Doskocil, J. and F. Sorm (1962). Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. Biochem. Biophys. Acta. 55, 953-959.

Fanning, E., and W. Doerfler (1976). Intracellular forms of Adenovirus DNA. J. Virol. 20, 373-383.

Felsenfeld, G. (1978). Chromatin. Nature 271, 115-122.

Flint, S.J. and H. Weintraub (1977). An altered subunit configuration associated with the actively transcribed DNA of integrated adenovirus genes. Cell 12, 783-794.

Flint, S.J. Transformation by Adenoviruses in DNA Tumor Viruses. IN Cold Spring Harbour Monographs, 1980 (J. Tooze, ed.) pp. 547-575.

Flint, S.J., J. Sambrook, J. Williams and P. Sharp (1976). Viral nucleic acid sequences in transformed cells: IV. A study of the sequences of adenovirus 5 DNA and RNA in four lines of adenovirus 5-transformed rodent cells using specific fragments of the viral genome. Virology 72, 456-470.

Franke, W.W. and U. Scheer (1978). Morphology of transcriptional units at different states of activity. Phil. Trans. R. Soc. Lond. B. 283, 333-342.

Garel, A. and R. Axel (1976). Selective digestion of transcriptionally active ovalbumin genes from oviduct nuclei. Proc. Nat. Acad. Sci. U.S.A. 73, 1102-1106.

Gautier, F., H. Bunemann and L. Grotjahn (1977). Analysis of calf-thymus satellite DNA: evidence for specific methylation of cytosine in C-G sequences. Eur. J. Biochem. 80, 175-183.

Gottesfeld, J.M. (1978). Organization of transcribed regions of chromatin. Phil. Trans. R. Soc. Lond. B. 283, 343-357.

Graham, F.L. A. van der Eb and H. Heijneker (1974). Size and location of the transforming region in human adenovirus type 5 DNA. Nature 251, 687-691.

Green, M. and M. Pina (1963). Biochemical studies on adenovirus multiplication: IV. Isolation, purification, and chemical analysis of adenovirus. Virology 20, 199.

Grippo, P., M. Iaccarino, E. Parisi and E. Scarino (1968). Methylation of DNA in developing sea urchin embryos. U. Mol. Biol. 36, 195-208.

Guntaka, R.V., P.Y. Rao, S.A. Mitsialis and R. Katz (1980). Modification of avian sarcoma provirus DNA sequences in nonpermissive SC cells but not in permissive chicken cells. J. Virol. 34, 569-572.

Gurdon, J. and H.R. Woodland (1968). Cytoplasmic control of nuclear activity in animal development. Biology Reviews 43, 233-260.

Gunthert, U., M. Schweiger, M. Stupp and W. Doerfler (1977). DNA methylation in adenovirus, adenovirus-transformed cells, and host cells. Proc. Natl. Acad. Sci. U.S.A. 73, 3923-2927.

Holliday, R. and J.E. Pugh (1975). DNA modification mechanisms and gene activity during development. Science 187, 226-232.

Kahana, Z.E., O.J. Miller and B.F. Erlanger (1977). Immunochemical studies on the 5-methylcytosine content of African green monkey satellite DNA. Cold Spring Harbour Symp. Quant. Biol. 42, 397-400.

Kappler, J.W. (1970). The kinetics of DNA methylation in cultures of a mouse adrenal cell line. J. Cell Physiol. 75, 21-32.

Ketner, G. and T.J. Kelly (1976). Integrated simian virus 40 sequences in transformed cell DNA: Analysis using restruction endonucleases. Proc. Nat. Acad. Sci. U.S.A. 73, 1102-1105.

Kuo, M.T., J.L. Mandel and P. Chambon (1979). DNA methylation: correlation with DNase I sensitivity of chicken ovalbumin and conalbumin chromatin. Nucl. Acids Res. 7, 2105-2113.

Lacy, E. and R. Axel (1975). Analysis of DNA of isolated chromatin subunits. Proc. Nat. Acad. Sci. U.S.A. 72, 3978-3982.

Laemmli, U.K., S.M. Cheng, K. Adolph, J. Paulson, J. Brown and C. Schmid (1977). Metaphase chromosome structures: the role of nonhistone proteins. Cold Spring Harbour Symp. Quant. Biol. 42, 351-360.

Levy-Wilson, B. and G. Dixon (1979). Limited action of micrococcal nuclease on trout testis nuclei generates two mononucleosome subsets enriched in transcribed DNA sequences. Proc. Nat. Acad. Sci. U.S.A. 76, 1682-1686.

Mak, S., I. Mak, J.R. Smiley and F.L. Graham (1979). Tumorigenicity and viral gene expression in rat cells transformed by Ad 12 virions or by the Eco RI-C fragment of Ad 12 DNA. Virology 98, 456-460.

Mandel, J.L. and P. Chambon (1979). DNA methylation: Organ specific variations in the methylation pattern within and around ovalbumin and other chicken genes. Nucl. Acids. Res. 7, 2081-2103.

Miller, D.M., P. Turner, A.W. Nienhuis, .E. Axelrod and T.V. Gopalakrishnan (1978). Active conformation of the globun genes in uninduced and induced mouse erythroleukemia cells. Cell 14, 511-521.

McGhee, J.D. and G.D. Ginder (1979). Specific DNA methylation sites in the vicinity of the chicken β-globin genes. Nature 280, 419-420.

Noll, M. (1974). Subunit structure of chromatin. Nature 251, 249-251.

Ortin, J., K.-H. Scheidtmann, R. Greenberg, M. Westphal and D. Doerfler (1976). Transcription of the genome of adenovirus type 12: Maps of stable RNA from productively infected human cells and abortively infected and transformed hamster cells. J. Virol. 20, 355-372.

Palmiter, R., S. McKnight, E. Mulvihill and A. Senear (1977). Regulation of gene expression in the chick oviduct by steroid hormones. Cold Spring Harbour Symp. Quant. Biol. 42, 639-647.

Peterson, J.L. and E.H. McConkey (1976). Non-histone chromosomal proteins from HeLa cells. A survey by high resolution two-dimensional electrophoresis. J. Biol. Chem. 251, 548-554.

Pivec, L., K. Horska, A. Vitek and J. Doskocil (1974). Plurimodal distribution of base composition in DNA of some higher plants. Biochim. Biophys. Acta. 340, 199-206.

Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. Cytogenet. Cell Genet. 14, 9-25.

Saragosti, S., G. Moyne and M. Yaniv (1980). Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. Cell 20, 65-73.

Sawada, Y., S. Ojima, H. Shimojo, K. Shiroki, and K. Fujinaga (1979). Transforming DNA sequences in rat cells transformed by DNA fragments of highly oncogenic human adenovirus type 12. J. Virol. 32, 379-385.

Seidman, M.M., A.J. Levine and H. Weintraub (1979). The asymmetric segregation of parental nucleosomes during chromosome replication. Cell 18, 439-449.

Sharp, P.A., U. Pettersson and J. Sambrook (1974). Viral DNA in transformed cells: I. A study of Adenovirus 2 DNA in a line of transformed rat cells using specific fragments of the viral genome. J. Mol. Biol. 86, 709-726.

Shiroki, K., H. Handa, H. Shimojo, S. Yano, S. Ojima and K. Fujinaga (1977). Establishment and characterization of rat cell lines transformed by restriction endonuclease fragments of adenovirus 2 DNA. Virology 82, 462-471.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Southern, E.M. (1980). Agarose Gel Electrophoresis. IN Methods in Enzymology, Vol. 68; Recombinant DNA (R. Wu, ed.) Academic Press, New York.

Stalder, J., A. Larsen, J. Engel, M. Dolan, M. Groudine and H. Weintraub (1980). Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNase I. Cell 20, 451-460.

Sugisaki, H., K. Sugimoto, M. Takanami, K. Shiroki, I. Saito, H. Shimojo, Y. Sawada, Y. Uemizu, S. Uesugi and K. Fujinaga. Structure and gene organization in the transforming Hind III-G fragment of Ad 12. Cell 20, 777-786.

Sutter, D. and W. Doerfler (1980). Methylation of integrated adenovirus type 12 DNA sequences in transformed cells is inversely correlated with gene expression. Proc. Natl. Acad. Sci. U.S.A. 77, 253-256.

Sutter, D., M. Westphaland, W.Doerfler (1978). Patterns of integration of viral DNA sequences in the genomes of adenovirus type 12-transformed hamster cells. Cell 14, 569-585.

Swartz, N.M., T.A. Trautner and A. Kornberg (1962). Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbour base sequences in deoxyribonucleic acids. J. Biol. Chem. 237, 1961-1967.

Tjia, S., E. Fanning, J. Schick and W. Doerfler (1977). Incomplete particles of adenovirus type 2. III. Viral and cellular DNA sequences in incomplete particles. Virology 76, 365-379.

Tooze, J. (1980). DNA Tumor Viruses. Cold Spring Harbour Monographs.

Topp, W.C., D. Lane and R. Pollack (1980). Transformation by SV40 and polyoma virus. IN DNA Tumor Viruses, Cold Spring Harbour Monographs (J. [illegible] 205

Van der Ploeg, L. and R. Flavell (1980). DNA methylation in the human γδβ-globin locus in erythroid and nonerythroid tissues. Cell 19, 947-958.

Vanyushin, B.F., A. Mazin, V. Vasilyev and A. Belozersky (1973). The content of 5-methylcytosine in animal DNA: the species and tissue specificity. Biochim. Biophys. Acta. 299, 397-403.

Varshavsky, A.J., O. Sundin and M. Bohn (1979). A stretch of "late" SV40 viral DNA about 400 bp long which includes the origin of replication is specifically exposed in SV40 minichromosomes. Cell 16, 453-466.

Waalwijk, C. and R.A. Flavell (1978a). Msp I, an isoschizomer of Hpa II which cleaves both unmethylated and methylated Hpa II sites. Nucleic Acids Res. 5, 3231-3236.

Waalwijk, C. and R.A. Flavell (1978b). DNA methylation at a CCGG sequence in the large intron of the rabbit β-globin gene: tissue-specific variations. Nucl. Acids Res. 5, 4631-4641.

Wahl, G.M., M. Stern and G.R. Stark (1979). Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization by using dextran sulphate. Proc. Natl. Acad. Sci. U.S.A. 76, 3683-3687.

Weinstock, R., R. Sweet, M. Weiss, H. Cedar and R. Axel (1978). Intragenic spacers interrupt the ovalbumin gene. Proc. Nat. Acad. Sci. U.S.A. 75, 1299-1303.

Weintraub, H. and M. Groudine (1976). Chromosomal subunits in active genes have an altered conformation. Science 193, 848-856.

Weintraub, H., A. Worcel and B. Alberts (1977). A model for chromatin based upon two symmetrically paired half-nucleosomes. Cell 9, 400-417.

Weisbrod, S. and H. Weintraub (1979). Isolation of a subclass of nuclear proteins responsible for conferring a DNase I-sensitive structure on globin chromatin. Proc. Natl. Acad. Sci. U.S.A. 76, 630-634.

Worcel, A. Molecular architecture of the chromatin fiber. Cold Spring Harbour Symp. Quant. Biol. 42, 313-323.

Wu, C., P.M. Bingham, K.J. Livak, R. Holmgren and S.C.R. Elgin (1979a). The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. Cell 16, 797-806.

Wu, C., Y-C Wong and S.C.R. Elgin (1979b). The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. Cell 16, 807-814.

Zasloff, M. and R.D. Camerini-Otero (1980). Limited DNase I nicking as a probe of gene conformation. Proc. Natl. Acad. Sci. U.S.A. 77, 1907-1911.