The Effect of DNA Methylation on Transcription from the SV40 Early Promoter

By

Margaret Bryans

Thesis submitted to the University of Glasgow
for the degree of Doctor of Philosophy

Department of Biochemistry,
The University of Glasgow,
Glasgow G12 8QQ.   August, 1989.

ProQuest Number: 10970928

ProQuest 10970928

Acknowledgements

I would like to express my gratitude to the late Professor R.M.S. Smellie and Professor

M.D. Housley for making available the facilities of the Department of Biochemistry for this

research. I would also like to thank the following people: Dr. Roger Adams for advice and

for useful critical discussions during the writing of this thesis and Dr. John Goddard and

my colleagues in lab. C30 for their good humoured friendship and many helpful suggestions.

Special thanks go to my friends Margaret McGarvey and Peter White for encouragement and

many interesting scientific discussions and Margaret White for helping me type this

thesis.

Abbreviations

The abbreviations used in this thesis are in agreement with the recomendations of the editors of the Biochemical Journal (Biochem. J (1989) **257**, 1-21) with the following additions:

m^{5}C	5-methylcytosine
m^{6}A	N^6-methyladenine
CAT	chloramphenicol acetyltransferase
dNTP	deoxyribonucleoside-5'-triphosphate
DTT	dithiothreitol
EtBr	ethidium bromide
IPTG	isopropyl-β-D-thio-galactopyranoside
PBS	phosphate buffered saline
PCA	perchloric acid
PMSF	phenylmethylsulphonylflouride
SAM (AdoMet)	S-adenosylmethionine
SDS	sodium dodecyl sulphate
TCA	trichloroacetic acid
tk	thymidine kinase
X-gal	5-bromo-4-chloro-3-indolyl-β-galactoside

Contents

Chapter 1:   Introduction

2.2. Methods

2.2.2. *In Vitro* Methylation of DNA with Mouse Ascites DNA Methylase

2.2.3. Chloramphenicol Acyl Transferase Transient Expression Assay

2.2.4. Gel Retardation Assays

Chapter 4:   The Effect of Methylation on Transcription

from the SV40 Early Promoter in a Transient

Expression Assay

Chapter 5:   The Effect of Methylation on the Binding of Sp1

Chapter 6:   Additional Methods of Studying the Effect of

Methylation on Transcription From the SV40

Promoter

List of Figures

(b)

Summary

The effect of DNA methylation on transcription from the SV40 early promoter was studied. The plasmid pVHC1 containing the SV40 promoter in the early direction linked to the chloramphenicol acetyltransferase (CAT) gene was methylated *in vitro* with mouse Krebs II ascites cell DNA methylase. This resulted in methylation of 10-25% of CpG dinucleotides. Comparison of this DNA with unmethylated plasmid in a transient expression assay indicated a 30-40% inhibition of CAT expression by methylation. At this level of methylation an average of only 1 to 2 of the CpGs contained within the promoter region will be methylated and the enhancer sequence which does not contain any CpG dinucleotides will remain unmodified. Therefore it is unlikely that this effect is caused by the direct inhibition of transcription factor binding, it is most likely due to the formation of inactive chromatin structures induced by the presence of methyl groups throughout the plasmid. This theory is supported by the observation that binding of the transcription factor Sp1 to GC box motifs is unaffected by methylation. In experiments where mouse cells were cotransfected with pVHC1 and oligonucleotides containing methylated or unmethylated GC boxes, the level of CAT expression was greatly reduced regardless of the methylation state of the GC box, indicating that methylation had no effect on Sp1 binding. Similarly in gel retardation assays, methylated and unmethylated GC box-containing oligonucleotides competed equally well with a labelled restriction fragment containing the SV40 promoter for Sp1 binding. These findings indicate that the SV40 promoter is methylation sensitive in a methylcytosine rich environment where these residues are most likely acting in an indirect manner, causing the formation of inactive chromatin and hence limiting transcription.

Chapter 1. Introduction

1.1. The Presence of 5-Methylcytosine in DNA

1.1.1   5-Methylcytosine Distribution

The presence of the modified bases 5-methylcytosine (m^5C) and N6-methyladenine (m^6A)(fig. 1.1) in DNA was first observed by Wyatt (1951). Prokaryotes contain both modified bases and in some cases N4-methylcytosine (m^4C), the product of the *Bcn* I methylase reaction, also (Janulaitis *et al.*, 1984). Lower eukaryotes often contain both m^5C and m^6A, whereas higher eukaryotes contain only m^5C. The content of this base in eukaryotic DNA varies greatly with species. Lower eukaryotes such as yeast and nematodes contain negligible levels. This is also true of insects, for example, less than 0.02% of cytosines are methylated in *Drosophila* (Pollack *et al.*, 1984). In higher eukaryotes, such as mammals, approximately 4-7% of cytosines are methylated whereas some plants contain up to 30% of methylated cytosines.

5-methylcytosine   N6-methyladenine
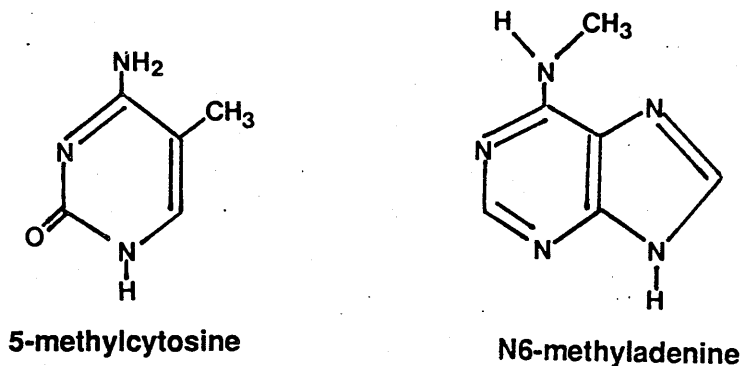
Figure 1.1.   5-methylcytosine and N6-methyladenine

The m^5C content of viral DNA is very variable. Despite their host cell genomes containing demonstrable levels of m^5C, the DNAs of herpes saimiri virus (Desrosiers *et al.*,1979), adenoviruses (Vardimon *et al.*,1980; Gunthert *et al.*, 1976), polyoma virus (Kaye and Winocour,1967) and simian virus 40 (Ford *et al.*,1980) have

no detectable m^5C. However, there are exceptions, Willis and Granoff (1980) observed that 20% of cytosine residues in frog virus 3 were methylated and Diala and Hoffman (1983) found that 14.5% were methylated in Epstein-Barr virus. In contrast to the majority of these animal viruses, many types of bacteriophage have m^5C or m^6A in their genomes at levels comparable to that of their bacterial host chromosomes. It has been shown (Desrosiers *et al*.,1979, Vardimon *et al.*, 1980) that the DNA of viruses infecting cells may be extensively methylated, for example when viral episomes are formed or when viral DNA is integrated into the host chromosome. This methylation is associated with relatively low levels of transcription from the viral genome or with the production of little or no virus.

Interest in the possible role of methylcytosine was initiated when Chargaff *et al*.,(1953,1955) observed that the distribution of m^5C residues in calf thymus DNA was non random, implying some functional significance. In mammalian DNA m^5C residues exist almost exclusively (>90%) in the dinucleotide CpG. Plants however also contain m^5C in the sequence CpNpG, where N is any nucleotide. On the entire genome level, the extent of CpG methylation varies from tissue to tissue. Somatic tissues are relatively highly methylated, whereas extra embryonic tissues appear to be hypomethylated. For example, in mouse spleen cells 82% of CpGs are methylated compared with 49% in placental DNA. However, at the gene level, most tissue specific genes are highly methylated in embryonic tissue and unmethylated in somatic cells. Satellite DNA is heavily methylated, the extent of methylation may be determined by the frequency of CpG sequences. Whereas in unique sequences the pattern of methylation varies from one tissue to another.

1.1.2 Detection of 5-Methylcytosine

There are several methods available for the identification and quantification of m^{5}C in DNA including chromatographic methods (Adams and Burdon, 1983), high resolution mass spectrometry (Razin and Cedar, 1977) and high performance liquid chromatography (Ford *et al*.,1980). In addition specific antibodies against m^{5}C have been used to detect the presence of the base in intact chromosomes (Schreck *et al*.,1977). Much information regarding the methylation patterns of specific genes has been obtained by the use of bacterial restriction endonucleases. For example the isochizomers *Hpa* II and *Msp* I, both recognise the sequence CCGG, however the former will cleave the DNA only if the inner C residue is unmethylated. Waalwyk and Flavell (1978), developed an experiment whereby genomic DNA is cleaved by either of the enzymes, the DNA fragments are then separated according to size by agarose gel electrophoresis and transferred by blotting to nitrocellulose (Southern, 1975). The DNA fragments in or near the gene of interest are then visualised by hybridisation with a labelled probe. A comparison of the bands observed for the *Hpa* II digestion with those observed for the *Msp* I digested DNA shows the location of the methylated CCGG sites.

However detection of methylcytosines is limited to those contained within methylsensitive restriction enzyme sites. This limitation has been overcome by a method introduced by Church and Gilbert (1984) and recently updated by Saluz and Jost (1989), known as genomic sequencing. This technique is based on the observation that m^{5}C residues could be detected in DNA by chemical sequencing (Maxam and Gilbert,1980). The presence of a CH_3 group in the pyrimidine ring of cytosine interferes with its reaction with hydrazine and results in a gap in the sequencing ladder. Based on this principle,

genomic sequencing enables the study of the methylation state of any gene in a strand and sequence -specific manner. This approach was used by Saluz *et al* .(1986) to study the *in vitro* methylation pattern of two CpG sites in the promoter region of the avian vitellogenin gene before and after hormone stimulus. *In vivo* protein-DNA interactions were also studied by combining this technique with the reaction *in vitro* of unprotected guanosine with dimethyl sulphate, known as genomic footprinting (Saluz *et al.*,1988). These methods will allow detailed investigation of the methylation patterns of inactive and activated genes and the relationship between methylation and protein DNA interactions in these genes.

1.1.3 Inheritance of Methylation Patterns.

The distribution of m^5C residues in DNA i.e. the methylation pattern can be stably and accurately inherited from one generation to the next. This is achieved by a maintenance methylation mechanism. Following replication an unmethylated CpG in the daughter strand of DNA will be paired with the mCpG in the parental strand. This site will then be recognised by DNA methylase which will add a methyl group to the daughter strand resulting in a fully methylated site. Hemimethylated duplex DNA has been shown by several groups to be the preferred substrate for purified DNA methylase. Totally unmethylated duplex DNA is a poor substrate for the enzyme, this ensures that *in vivo* the existing methylation pattern is copied during DNA replication (figure 1.2).

This maintenance methylation was demonstrated in transfection experiments performed by Wigler *et al* .(1981), where a cloned chicken thymidine kinase gene was

methylated at CCGG sequences using *Hpa* II methylase and introduced into TK⁻ mouse cells.

Twenty five generations later the methylation of the gene was shown to have been

maintained with almost 100% fidelity.

De novo methylation i.e. methylation of totally unmethylated DNA, is thought

only to occur during early development in non differentiated cells, where most genes

require to be silenced.
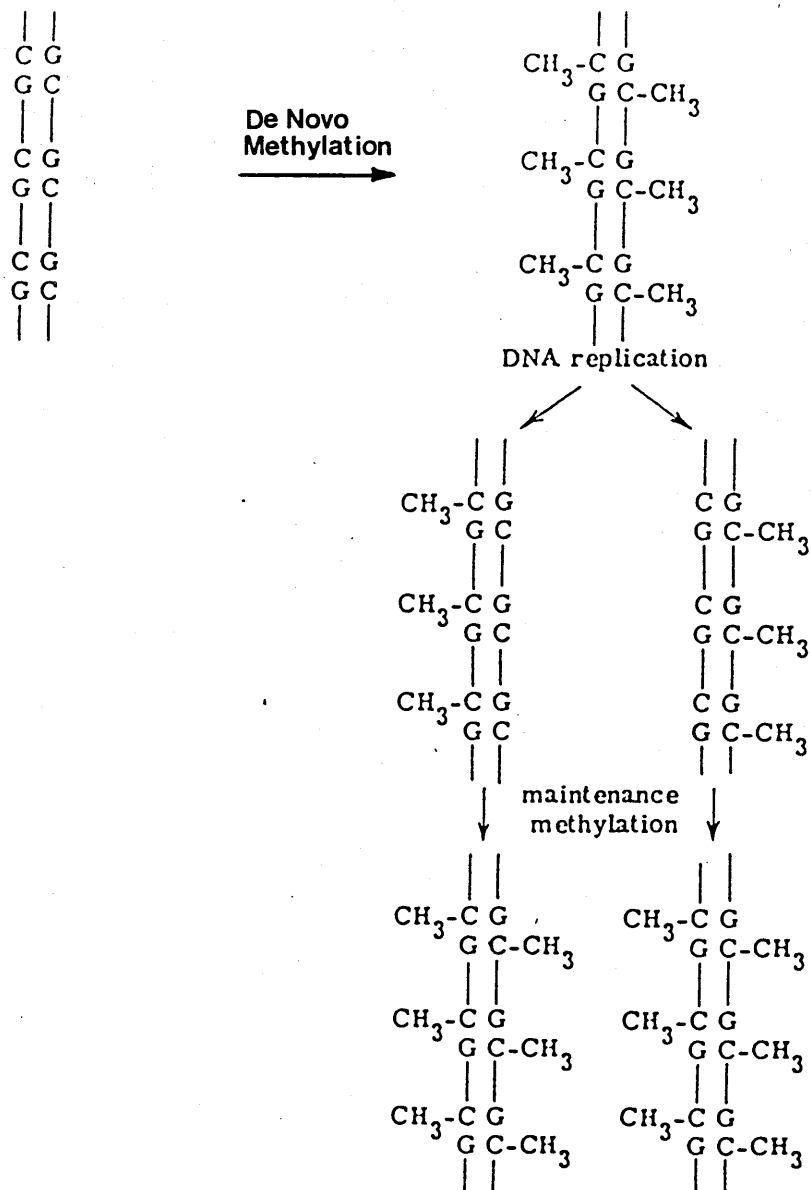
Figure 1.2. *De Novo* and Maintenance Methylation

1.1.4. DNA Methyltransferases

The addition of a methyl group to adenine and cytosine residues in DNA is a post

replicative event which is catalysed by an enzyme known as DNA methyltransferase (DNA

methylase, EC 2.1.1.3.7). The methyl group donor in these reactions is S-adenosyl

methionine. The methylation of cytosine residues is shown in fig 1.3 and probably

proceeds *via* the formation of an enzyme-DNA covalent link (Santi *et al.*,1983,1984).

The products of the reaction are 5-methylcytosine and S adenosylhomocysteine, a potent

inhibitor of the reaction.
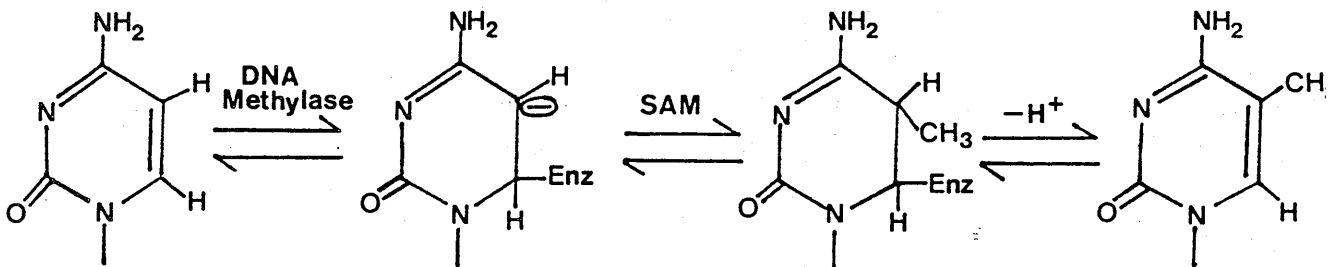
Figure 1.3   Methylation of Cytosine: Proposed Mechanism of Action

In prokaryotes, DNA methylases (with the exception of the dam and dcm types)

are associated with endonucleases as part of restriction modification systems.   In contrast,

eukaryote DNA methylases have no associated endonuclease and also unlike their prokaryote

counterparts have no co-factor requirements and are not stimulated by ATP or Mg^{2+} .

Mammalian DNA methylase has been isolated from a number of tissues including rat liver

(Simon *et al.*,1978), mouse ascites cells (Turnbull and Adams,1976), murine

erythroleukemia cells (Bestor and Ingram, 1983) and Hela cells (Roy and

Weissbach,1975). These enzymes have now also been purified from plants; Theiss *et al.* (1987) isolated DNA methylase from wheat embryo cells and Yesufu *et al* .(1989) from pea seedlings.

The molecular weight of these enzymes lie between 120,000 and 280,000 and may differ between different vertebrate tissues. *In vitro* the enzyme aggregates in low salt but it is only under these conditions that it interacts with native DNA. It is therefore not clear whether the active form of the enzyme *in vivo* is the monomer or some multimeric form.

The precise mechanism of action of DNA methylase is still unknown. In studies using the rat liver enzyme, Drahovsky and Morris (1971) proposed that the enzyme processed along the DNA in a "random walk" manner, adding methyl groups as it came to unmethylated DNA. This one dimensional diffusion or sliding would greatly increase the chances of the methylase finding its recognition sequence and hence would be more efficient than an association-dissociation mechanism. In support of this processive mechanism Taylor and Jones (1982) found that the mouse spleen enzyme bound to duplex DNA and remained associated after the addition of S-adenosylmethionine and hemi-methylated duplex DNA. No methylation of the hemi-methylated DNA was observed, indicating that once bound the enzyme remained associated with the DNA. Also, treatment of cells with low levels of 5-azacytidine resulted in an unexpectedly high inhibition of methylation. Substitution of 5% of cytosine residues with 5-azacytidine resulted in an 80% drop in the level of methylation (Jones and Taylor,1980). This can be explained if the occurrence of a 5-azacytidine residue at a methylation site impedes the progress of the enzyme as it travels down the newly synthesised DNA strand at DNA replication.

All mammalian enzymes purified so far show a strong preference for hemimethylated DNA as a substrate. Since these enzymes are purified from fully differentiated cells their primary function *in vivo* would be the maintenance of established methylation patterns. There is no evidence for the presence of a separate enzyme in non-differentiated cells. Hence it is possible that an altered form of this enzyme performs the required *de novo* methylation during early development. Adams and Burdon (1985) proposed that DNA methylase may have a regulatory domain that recognises a methyl cytosine on one strand of the hemimethylated DNA and only then will the active site catalyse the transfer of a methyl group to the unmethylated strand. Such a mechanism could explain the increase in activity on unmethylated DNA following limited proteolysis of the enzyme (Adams *et al.*,1983). Removal of the domain that recognises the hemimethylated site may relieve a constraint, thereby allowing the enzyme to work on totally unmethylated DNA.

Insight into the mechanism of action of DNA methylase may be gained from sequence analysis of the recently cloned mouse methyltransferase gene (Bestor *et al.*, 1988). The inferred protein sequence revealed a carboxy terminal region which showed homology with the bacterial type II cytosine methylase. Bestor postulated that the amino-terminus portion of the molecule may be involved in regulating the activity of the enzyme.

1.1.5 Possible Roles For DNA Methylation

The major part of the evidence arising from the studies on m^{5}C in eukaryotic cells indicates that methylation is involved in the control of gene expression and this will be discussed in detail later. However, since m^{5}C residues are present in sequences not associated with the control of transcription, other roles for cytosine methylation in eukaryotes have been postulated. These include;

(a) a restriction modification system similar to that present in prokaryotes. However, sequence specific endonucleases in mammalian cells have not been detected. The only evidence in support of such a system comes from studies on the chloroplast inheritance in *Chlamydomonas* (Burton *et al.*, 1979; Royer and Sager,1979). In gametogenesis methylation of the chloroplast DNA of the female gamete (mt$^+$) protects it against degradation, whereas the unmethylated chloroplast DNA of the male gamete (mt$^-$) is digested after fertilisation. A general role for DNA methylation in a possible eukaryotic restriction modification system is however unlikely, since eukaryotic genomes are only partially methylated and much of the DNA would therefore be unprotected.

(b) It has also been postulated that cytosine methylation plays a role in the mismatch repair mechanism in eukaryotes. It has been shown that m^{6}A has a role in guiding strand selection in base mismatch repair in prokaryotes (Jones *et al*.,1987). Cytosine methylation in post replicative hemimethylated DNA would enable discrimination between the methylated parental strand and the unmethylated daughter strand. Hare and Taylor (1985) found that mismatch repair occurred only on the unmethylated strand of hemimethylated SV40 DNA after transfection into host cells. Such a system would allow for the correction of mismatches overlooked by the proofreading enzymes present at

replication and the ^{mCG}/GT mismatch resulting from m^{5}C deamination at CpG dinucleotides.

Brown and Jiricny, (1987) have reported that mammalian cells possess a specific repair pathway , which counteracts the mutagenic effects of this deamination by correcting GT mismatches almost exclusively to GC pairs. Wiebauer and Jiricny (1989) have recently shown that this repair action is present in Hela cell nuclear extracts and involves the excision of the aberrant thymidine monophosphate residue, followed by gap filling to generate a CG pair.

(c) DNA methylation has also been postulated to be involved in DNA replication, however, although evidence exists suggesting that methylation is involved in replication in prokaryotes (Razin, 1978; Billin, 1968), there is little evidence supporting the role of methylation in the control of replication in eukaryotes. Cooney *et al.*, (1988) studied the role of methylation in the control of replication in the rDNA of the slime mould *Physarum polycephalum* and found that of the four origins of replication present the single active origin was hypomethylated. They propose that hypomethylation has a role in selecting the active origin.

1.2. The Relationship Between DNA Methylation and Transcription

1.2.1 DNA Methylation and X Chromosome Inactivation

After fertilisation, female mammalian cells have two X chromosomes. In early embryogenesis, genetic inactivation of one of these chromosomes occurs to ensure that

these females are equivalent to XY males with respect to X chromosome dosage. This

process of inactivation is thought to occur in at least three steps : the initial event in which

one of the two active X chromosomes in each cell is chosen for inactivation, the spread of

inactivation along the length of the chromosome and events that result in the heritable

maintenance of repression of the genes on the inactivated X. This chromosome appears as

heterochromatic and condensed, it is late replicating and is known as a Barr body.

The role of methylation in the inactivation process was first proposed by Riggs

(1975) and several lines of experimental evidence supporting this idea followed. For

example, Liskay and Evans(1980) found that in transcription studies the *hprt*

(hypoxanthine phosphoribosyl transferase) gene isolated from the inactive X chromosome

could not transform HPRT$^-$ recipient cells to a positive phenotype; whereas the same gene

from the active X could. This indicates that inactive and active X chromosome DNA

fragments are differentially modified. Strong evidence for the involvement of methylation

in this inactivation process comes from studies using 5 azacytidine. This compound is a

strong inhibitor of DNA methylase when incorporated in DNA and results in the loss of 5mC

and altered methylation patterns. Mohandas *et al* ., (1981) exposed a mouse-human

somatic cell hybrid, deficient in HPRT and containing a structurally normal inactive human

X chromosome, to this drug , and tested for the reactivation of the human X linked *hprt*

gene. They found that the frequency of HPRT positive clones after 5-azacytidine treatment

was 1,000 times greater that obtained in untreated cells. Later Jones *et al* .,(1982)

showed that this reactivation required that the hybrid cells be treated with the drug in the

latter part of S phase of the cell cycle, when the inactive X chromosome is replicating,

suggesting that it is involved in the alteration of methylation patterns.

Reactivation is thought to be due to a demethylation of a CpG island upstream from these genes. Hansen *et al*.,(1988) observed that demethylation of specific sites within the CpG cluster of the X· linked phosphoglycerate kinase gene were important for DNAase I sensitivity and transcriptional activity of the gene.

From these experiments and others, it is generally accepted that DNA methylation is involved in the inactivation of the X chromosome. However, it is unknown at what stage in the inactivation process it is involved and how important a role it plays in relation to the other factors involved.

Monk (1986) proposed a model whereby after the initial event which determines the inactive X chromosome, keysite sequences along the length of the chromosome are methylated resulting in inactivation of the entire chromosome. A conformational event demarcating the chromosome into segments of inactivity then occurs. Evidence for this step comes from studies indicating that X chromosome reactivation occurs in a segmented fashion. Finally a further methylation of the inactive genes on the inactive X chromosome in somatic tissue occurs. It would be this second methylation step that causes decreased transcription activity of the *hprt* gene since, if the same gene is isolated from the inactive X chromosome of extraembryonic tissue no decrease in activity is observed (Kratzer *et al.*,1982).

However, the initial role of methylation in inactivation presented in this model, is disputed by Lock *et al.*, (1987). They studied the relationship between DNA methylation of the *hprt* gene and the inactivation of this gene on the X chromosome in teratocarcinoma cells and postimplantation mouse embryos. They found that the sequences studied are unmethylated prior to X inactivation and do not become methylated on the

inactive X chromosome until several days postinactivation. This suggests that methylation does not play any role in the primary events of X inactivation, but may function as a secondary mechanism for maintaining the inactive state.

This secondary or 'locking' mechanism, first introduced by Razin and Riggs (1980), may function in a localised manner either in combination with or in some cases replacing the primary mechanism by which X linked genes are maintained in the repressed state.

1.2.2 DNA Methylation and Cellular Differentiation

The ordered expression of tissue specific genes during development is controlled by numerous cellular factors. The knowledge that DNA modification by methylation could be inherited from one generation to the next led Riggs (1975) to propose that DNA methylation was one of the factors involved in this complex process. Strong support for the involvement of DNA methylation in differentiation came from reports that agents that inhibit the methylation reaction, such as ethionine, butyrate, dimethylsulphoxide and 5-azacytidine, induce differentiation in cultured cells. Christman et al .,(1977) observed that when Friend erythroleukaemia cells are grown in the presence of L-ethionine, globin mRNA accumulates and 25-30% of the cells produce haemoglobin 4-5 days after initiation of treatment. The DNA isolated from these cells was significantly hypomethylated compared to that of uninduced cells. Also Jones and Taylor (1980) found that 5-azacytidine and other cytosine analogues could have profound effects upon the differentiation state of cultured mouse embryo cells. These cells developed into

contractile muscle cells and biochemically differentiated adipocytes and chondrocytes

several days or weeks after exposure to the analogue. These experiments suggest an

important role for methylation in the maintenance and acquisition of differentiated

phenotypes.

Further support for the role of methylation came from the observation

that tissue specific genes are highly methylated in germline DNA and undermethylated in

the tissue in which the gene is expressed. This led to the theory that *de novo* methylation

occurs in the early embryo and tissue specific methylation patterns are produced by the

loss or gain of critical methylated CpG dinucleotides during the process of differentiation.

Cedar *et al*., (1983) found that teratocarcinoma cells derived from mouse embryos

carried out *de novo* methylation of a transfected β-globin gene. The same DNA introduced

into either L cells or Friend erythroleukaemia cells remained unmethylated.

Riggs and Jones (1983) proposed various mechanisms by which tissue

specific methylation patterns might be set up during development. Specific demethylation

at critical control sites during development could be directed by sequence specific proteins

that sterically inhibit the relatively non-specific DNA methylase. Alternatively the

pattern of methyl groups could be altered by the action of a demethylase that acts in a

sequence and developmental specific manner. This event may occur before or after the

activation of genes. Gjerset and Martin (1982), observed the presence of a m^5C

demethylase activity in murine erythroleukaemia cells, where this event was thought to

occur *via* the removal of the methyl group from the DNA. Also Razin *et al*., (1986)

observed a genome wide transient demethylation in differentiating erythroleukaemia cells.

The kinetics of this event suggested that it did not depend on DNA synthesis indicating that

demethylation was active rather than passive i.e not via DNA replication. This induced

demethylation is thought to be due to the more feasible specific replacement of 5-methylcytosines with cytosine residues. Studies on such enzyme activities are still in their infancy and have still to be confirmed by other groups. The third possibility for changing methylation patterns is to have some *de novo* methylation occurring at all developmental stages. Stewart *et al*.,(1982) suggested that *de novo* methylation is most active in the pre-implantation embryo. However, gene transfer experiments performed by Doerfler, (1981), have shown that fully differentiated cells can perform *de novo* methylation.

In order to understand how methylation influences the tissue specific expression of genes during differentiation it is useful to consider the two classes of genes involved, that is, housekeeping genes and tissue specific genes. Housekeeping genes are constitutively expressed in all somatic tissues. Many of these genes contain CpG islands usually in their 5' flanking sequences. First identified by Bird *et al*.,(1985), these islands are regions of DNA, 500-2000bp long, which are C+G rich and contain a high proportion of unmethylated CpG dinucleotides. As cells move from germline to committed cells *en route* to differentiation these regions would remain free from the repressive effects of methylation, possibly by the aid of steric hindrance from permanently bound factors. CpG islands have also been found in the 3' or 5' flanking regions of some tissue specific genes and may in the distant past have been a characteristic of all genes. During early development these tissue specific genes would become methylated and in some cases the CpG rich character of 5'sequences would be lost due to high rates of mutation of mCpG. As differentiation proceeds tissue specific and stage specific factors may then alter the methylation patterns of these genes *via* one of the mechanisms outlined above.

1.2.3 Evidence for an Inverse Relationship between Methylation and

Transcription

The role of methylation in the regulation of transcription has been

extensively studied. However the evidence accumulated so far does not provide a concrete

theory for how or to what extent DNA methylation controls the expression of genes in

eukaryotes. Evidence supporting the inverse relationship between methylation and

transcription has come from studies on X-chromosome inactivation and cellular

differentiation, discussed in sections 1.2.1. and 1.2.2 respectively. However, analysis of

the relationship between extent of methylation and transcription of a number of genes

provided more ambiguous conclusions. There are many known examples of genes in which

methylation of specific sequences in the 5' regulatory region causes inactivation. Using

the technique introduced by Waalwijk and Flavell (1978) involving the use of restriction

enzyme isochizomers (outlined in section 1.1.2) the methylation patterns of several cloned

genes have been studied. For example, studies on the methylation pattern of the active and

inactive mouse metallothionein (MT-1) gene by Compere and Palmiter (1981) revealed

that all *Hpa* II (CCGG) sites in the vicinity of the MT gene are methylated in cells in which

the gene is never expressed but unmethylated in cells in which the gene can be induced by

cadmium or glucocorticoids. In addition, pretreatment of non-expressing cells with

5-azacytidine allowed the gene to be induced by these agents. Also, Weintraub *et al.*,

(1981) observed an undermethylation of *Hpa* II sites in the DNAase I sensitive domain 5'

to the α-globin gene cluster in developing chicken embryos. A similar observation was

made with the rabbit β-like globin genes (Shen and Maniatis, 1980) where analysis of 13

sites flanking the gene cluster revealed that certain sites surrounding the embryonic and

adult genes are relatively undermethylated in DNA from embryonic and adult erythroid tissue respectively. Other examples include the chicken ovalbumin gene (Mandel and Chambon, 1979), where certain CpG's are unmethylated only in the oviduct of laying hens, the chicken β globin genes (McGhee and Ginder, 1979) and the mouse (Sellem et al., 1985), and rat (Ott et al., 1982) albumin genes.

Further support for this inverse correlation between DNA methylation and gene expression came from DNA mediated gene transfer experiments, using both viral and cellular genes. For example, Stein et al.,(1982) methylated the cloned hamster adenine phosphoribosyl transferase (aprt) gene at all CCGG sequences using Hpa II methylase. The ability of this DNA to transform APRT⁻ cells to a positive phenotype was compared to that of unmethylated DNA. Although both types of DNA became integrated into the host genome, only the unmethylated gene conferred an APRT⁺ phenotype, suggesting that expression is inhibited by methylation.

Christy and Scangos (1982) found that introducing a plasmid containing the herpes simplex virus thymidine kinase (tk) gene into mouse TK⁻ cells resulted in TK⁺ cell lines. A derivative of these cells lines that retained but no longer expressed the tk gene was found to be hypermethylated in the gene sequence compared to the expressing cells. Furthermore the inactive gene could be activated by exposing the cells to 5-azacytidine, suggesting a causative relationship between DNA methylation and transcription. Further experiments using this gene have confirmed its methylation sensitivity. Graessman and Graessman (1988) methylated this gene with Hpa II methylase and studied its expression on microinjection into the nuclei of TK⁻ cells. They found that methylation could block transcription of this gene. The inhibitory effect was not immediate, suggesting that it is mediated via the formation of chromatin structures.

DNA viruses have provided useful model systems for studying the effect of methylation on transcription. Despite naturally possessing an unmethylated genome, in many cases viral gene expression can be inhibited by methylation of specific sites in promoter regions. These include the herpes saimiri virus (Desrosiers *et al* .,1979) , hepatitis B virus (Miller and Robinson 1983) and Moloney leukemia virus (Stahlmann *et al* .,1981). Also, the methylation of the single *Hpa* II site in SV40 located near the late transcription start site inhibited transcription of late genes but had no effect on early transcription (Fradin *et al.,* 1982). By far the most extensively studied virus, with respect to the effect of methylation on transcription, is the adenovirus system. Doerfler and co-workers (Sutter and Doerfler 1979 , Vardiman *et al* .,1980 ,) have shown that although virion DNA and free viral DNA in infected cells is not detectably methylated, integrated viral DNA sequences are extensively methylated. Analysis of the methylation status of integrated adenovirus 12 (Ad 12) sequences in transformed cells revealed an inverse correlation between methylation of CCGG sites and transcription (Kruczek and Doerfler,1983).

Similar results were also obtained for the Ad 2 system, where early adenovirus genes were undermethylated in cell lines in which they were expressed; late viral genes not expressed in transformed cells were heavily methylated at CCGG sites. Also, they found that *in vitro* methylation of the Ad 12 E1a promoter at *Hpa* II and *Hpa* I sites abolished the transcription of a linked chloramphenicol acetyl transferase gene. However, methylation of sites downstream of the promoter of the Ad 12 protein 1x gene and greater than 300bp upstream of the TATA box had no effect on transcription. This experiment indicates that the effect of methylation is site specific and usually involves CpG's in or near promoter regions.

In contrast to this, studies on the frog virus 3(FV3) which has a highly

methylated genome (20% m^{5}C , Willis and Granoff, 1980) revealed that methylation in

the promoter region of an immediate early gene had no effect on transcription (Thompson *et*

al.,1988). Also *in vitro* methylation of *Hpa* II sites on polyoma virus DNA had no effect

on T antigen expression (Subramanian, 1982).

These contrasting results are typical of the data so far accumulated on the

effect of methylation on transcription. The experiments discussed so far provide good

evidence for an inverse relationship. However, a proportion of the studies carried out to

date have revealed no obvious relationship between methylation and transcription in a

number of genes suggesting that the control mechanisms are totally independent of

methylation. These include the *Xenopus* vitellogenin genes A1 and A2 which are expressed

in hepatocytes upon oestrogen treatment. Gerber-Huber *et al* .,(1983) found that these

genes are heavily methylated in all adult tissues, irrespective of expression. In contrast

the *Xenopus* β1-globin gene and albumin gene are hypomethylated only in expressing cells,

i.e. erythrocytes and hepatocytes . Also McKeon *et al* .,(1982) studied the methylation

pattern of the α2 (type 1) collagen gene in DNA from different cell types with varying

rates of collagen synthesis. Using methylation sensitive restriction enzymes they found

that DNA around the start site of transcription was unmethylated whereas the DNA from the

central and 3' region of the gene is methylated. This methylation pattern was the same for

collagen expressing and non expressing cells indicating that the level of expression of this

gene is independent of methylation. Using the genomic sequencing technique Nick *et al* .,

(1986) found that a 900bp fragment upstream of the non expressed maize alcohol

dehydrogenase gene was unmethylated despite being rich in methylation sites indicating that

methylation is not involved in repression. This group of apparently methylation insensitive genes also includes the rat (Cate *et al*.,1983) and human (Ullrich *et al*., 1982) insulin genes and the rat albumin and α fetoprotein genes (Kunnath and Locker, 1983), where both genes are active in the liver of the 18 day old foetus despite being heavily methylated.

From these studies we can conclude that site specific methylation of CpG's in a large number of (but not all) viral, developmentally active and housekeeping genes has a direct effect on expression. It appears that methylation exerts its effect only at critical CpG's, the location of these sites differs for different genes. In some genes methylation in the body of the gene inhibits expression, whereas in the housekeeping *aprt* and *dhfr* genes the active gene is heavily methylated throughout except for the 5' regulatory region (Stein *et al*.,1983). It is possible that in these housekeeping genes regulation is not so rigorous and activity depends only on the upstream sequences, containing a CpG island, to be methylation free allowing the necessary ubiquitous transcription factors to bind. On the other hand , in tissue specific genes transcription control is more complex probably involving tissue specific factors as well as ubiquitous factors. Hence critical methylation sites within these genes would differ between genes depending on the control systems involved and where in the gene sequence the tissue specific factors bind.

Although these arguments indicate that methylation is involved in gene regulation, one cannot conclude whether this modification is a cause or an effect of transcriptional inactivation. It is possible that methylation or hypomethylation provides a means of stabilising the structure of a gene in an inactive or active state.

1.3. Events Involved in the Transcription of Eukaryotic

Genes

1.3.1. Transcriptional Activation in Cellular Genes

Transcription in eukaryotes is thought to be controlled by cis and trans acting

mechanisms and by the accessibility of the template DNA to protein factors which initiate

transcription. Transcriptionally active genes are known to be preferentially digested by

pancreatic DNAase I. This sensitivity is not limited to the transcribed portion of a gene,

but extends both 5' and 3' to define an active domain. For example, the ovalbumin domain

in hen oviduct nuclei contains three genes and extends over a 100kbp of DNA, and the active

β globin domain extends 6-7kbp 5' and 8kbp 3' of the gene. These active or open

chromatin domains often contain sites of nuclease hypersensitivity immediately preceding a

transcribing gene. (Larson and Weintraub,1982). The dissociation of nucleosomes in

these regions would greatly increase the accessibility of the DNA to RNA polymerase and

transcription factors.

1.3.1.1 Common Control Elements in Eukaryotic Promoters

The upstream control region of many genes studied to date share several

sequence elements. These include the TATA box, which has a consensus sequence TATAAA

and is located 25-30bp upstream from the transcription start site. Mutation in this

region results in 5' heterogeneity in the transcripts produced, although the overall level of

RNA synthesis remains similar. Further upstream is a region containing one or more

sequence elements of 8-12 base pairs, known as upstream promoter elements (UPEs). A

number of different UPEs have been identified and some of these, for example, the CCAAT

box and GC box have been found in a large number of promoters. Others such as the

metallothionein metal regulatory element (Karin et al., 1984) or the heat shock

regulatory element (Pelham,1982), have a more specialised role. All of these sequence

elements have been shown to bind specific cellular factors, indicating that cells contain

both 'general' and cell type-specific transcription factors.

As well as the transcription factors that bind to the CCAAT and GC boxes,

CTF and Sp1 respectively, a number of other ubiquitous transcription factors are also

involved in transcription initiation. Roeder and co-workers found that purified RNA

polymerase II lacks the ability to recognise promoters in an in vitro reaction. However,

the addition of factors purified from crude cell extracts stimulated polymerase activity and

allowed in vitro initiation of transcription. These factors have been identified as TFIIA,

TFIID, TFIIE and TFIIB (Reinberg and Roeder,1987;Reinberg et al.,1987). It is thought

that TFIIA and TFIID act first to form a stable preinitiation complex on the -45 to +30

region of the promoter, commiting the template to transcription. TFIIB interacts with

RNA polymerase and TFIIE and this complex then interacts with the TFIID-TFIIA complex

bound to the selector region and the gene is transcribed.

The activity of many promoters is modulated by an enhancer sequence. This

is a separate cis-acting regulating element which can exert its effect at a distance of up to

several Kbp from the promoter. These sequences are active in either orientation and may

be located either upstream or downstream from the transcription start site. Enhancers

usually bind one or more sequence specific factors. It is the interactions of these factors

with those bound to the UPE's of the promoter that is thought to regulate transcription.

These factor-factor interactions and possible factor-polymerase interaction can be

envisaged if the intervening DNA sequences 'loop-out' as the interaction occurs.

1.3.1.2 Tissue Specific Expression of Genes

In higher eukaryotes there are many examples of genes that are under tight

developmental control and are transcribed only in specific differentiated cell types. The

mechanisms by which the transcription of these genes is controlled are still largely

unknown. Studies on two genes - the immunoglobulin gene and the albumin gene have

provided some insight into how tissue specific expression is achieved. In transfection

studies where the introduced DNA is believed to assume an accessible chromatin structure

(Weintraub,1983), both these genes were shown to be active only in cell types that

expressed the endogenous gene i.e. lymphocytes and hepatocytes respectively. This

indicated that cell type specific transacting factors are required for their activity.

In the albumin gene transcriptional activity is dependent on sequences contained

within the region of -32 to -164. Lichsteiner *et al*., (1987) tested the ability of

nuclear extracts from several tissues to transcribe the albumin gene in an *in vitro* assay.

They found that only the liver cell extract could direct efficient transcription and spleen

and brain extracts were much less efficient. Footprinting analysis revealed at least six

specific protein binding sites, four of which were occupied by nuclear proteins that are

specifically more abundant in liver than in the other tissues tested.

Similarly lymphocyte specific factors appear to be directly involved in mediating tissue-specific immunoglobulin transcription. Immunoglobulin heavy and light chain genes possess a conserved DNA sequence octomer (ATGCAAAT) shortly upstream from the transcription start site which is thought to be involved in tissue specific expression. However there appear to be two species of factor that interact with this site. One factor is present in all cells and the other only in B and T lymphocytes (Landolfi *et al*.,1986 , Staudt *et al*., 1986). It has been suggested that one factor may be a modified form of the other differing for example in phosphorylation state. However, there is evidence to suggest that two entirely different proteins may exist. Staudt *et al*.,(1986) found that the lymphocyte specific factor is inducible in a pre B cell line by lipopolysaccharides and that this induction could be blocked by the protein synthesis inhibitor cycloheximide. However it is unlikely that all inducible or tissue specific gene expression involves the synthesis of new proteins. Activation of pre existing transcription factors could be achieved by covalent modification of the protein or an allosteric change in its structure. It has been shown (Imbra and Karin,1987) that treatment of cells with phorbolesters can dramatically increase the activities of the SV40 and polyoma virus enhancers. These compounds stimulate protein kinase C activity, suggesting that enhancer binding factors may be modified by phosphorylation. Phosphorylation has also been proposed as a regulatory mechanism for transacting factor E4F (Raychandhuri *et al*.,1989) and for the heat shock transcription factor (HSTF)(Sorger *et al*.,1987). Other regulatory mechanisms proposed are glycosylation (Jackson and Tjian,1988), induction by cAMP (Karin, 1989) and in the case of some *Drosophila* factors , self regulation (Serfling,1989).

In many tissue specific and developmentally regulated genes, control of

transcription involves the action of steroid hormones. Hormone regulatory elements

(HRE's) have been identified in a number of these genes including the chicken vitellogenin

gene (Klein-Hitpaβ *et al.*,1986) , the tyrosine aminotransferase gene (Jantzen *et*

al.,1987) , and the mouse mammary tumour virus long terminal repeat (MMTV-LTR)

(Payvar *et al.*, 1983; Scheidereit *et al.*,1983). These elements share many properties of

the enhancer sequence; interactions with the hormone receptor molecule is followed by

protein-protein interactions with transcription factors bound to promoter sequences and

results in stimulation of transcription.

1.3.2. Transcriptional Activation in SV40

The SV40 DNA tumour virus is an ideal eukaryotic model system to study.

Its small (~5.2Kbp) and well characterised genome has provided much information

regarding the processes of DNA transcription and replication. The regulatory sequences

for both these processes are contained within a region of the genome of approximately

300bp. Detailed studies on this region revealed several interesting sequence motifs.

These include a 72bp repeat enhancer sequence, a bidirectional promoter sequence and the

origin of replication including specific binding sites for T antigen.

The SV40 promoter region consists of two perfect 21bp repeats and a

degenerate 22bp repeat. Each of these repeats contain two hexanucleotide motifs known as

GC boxes, which have been shown by several groups to be crucial for both early and late

transcription (Brady *et al.*,1984; Baty *et al.*, 1983; Benoist and Chambon,1981).

Mutation studies on each motif from GC box I to VI lowered early transcription to 4.8%,20%,15%,42%,33% and 56% of wild type respectively (Wildeman, 1985). This is in agreement with Gidoni et al .,(1984) who observed that GC boxes I, II, and III are the most important for transcription in vitro in the early direction and boxes III,IV and VI are most important for transcription in the late direction.

Studies on these sequences revealed that they interact with a transcription factor known as Sp1. This protein has since been purified to 95% homogeneity and studied extensively by Tjian and co-workers (Kadonaga and Tjian,1986; Kadonaga et al.,1986) who observed that it has a molecular weight of 95-105K and is capable of enhancing transcription 10-50 fold from an apparently unrelated group of promoters. Analysis of the conferred amino acid sequence from the cloned human Sp1 gene revealed that it interacts with its recognition sequence via three contiguous zinc finger motifs (Kadonaga et al .,1987). These metalloprotein structures were first identified by Miller et al ., (1985) in transcription factor IIIA and are now known to be common to many DNA binding proteins. In a complementation assay using Sp1-deficient Drosophila tissue culture cells, Courey and Tjian (1988) were able to identify the protein domains that confer transcriptional activation. The two most active domains have a high glutamine content, a property shared by several Drosophila transcription factors and which is thought to represent a novel structure for transcriptional activation.

It appears that Sp1 does not bind to all GC boxes with equal affinity, and that the flanking nucleotides may play some part in binding. Analysis of Sp1 interaction with several sites resulted in a decanucleotide consensus sequence (5' $^G/_T$GGGCGG $^{GGC}/_{AAT}$), (Kadonaga et al .,1986). In SV40 boxes III and V are high affinity sequences whereas GC box I, shown by mutation studies to be most important for early transcription, is a low

affinity Sp1 binding sequence.   This surprising observation can be explained if one considers that *in vivo* there are other determinants of a site's activity such as, the position of the binding site relative to the RNA start site and to other factor binding sequences.

Sp1 is now regarded as an ubiquitous transcription factor since it appears to be involved in the regulation of a number of cellular and viral genes.   These include the mouse dihydrofolate reductase (*dhfr*) gene (Dynan *et al.*, 1986),the human metallothionein gene (Lee *et al* .,1987) and the HSV immediate early gene (Jones and Tjian,1985).   Since this group of Sp1 responsive genes contains several housekeeping genes it has been suggested that Sp1 interacts with the methylation-free CpG islands which are often present upstream of these genes and allows constitutive expression in all cell types.   However, Sp1 is also involved in tissue specific gene expression as indicated by an experiment performed by Hayashi *et al* .,(1986).   Using the lens specific chicken δ–crystallin gene 1, they found that co-injection of this gene with GCbox-containing DNA fragments from δ-crystallin, SV40 early, or HSV *tk* promoters effectively suppressed expression in lens cells, whereas coinjection with DNA fragments not containing a GC box did not.

Sp1 sequences are often found near binding sites for other transcription factors, such as CTF (Jones *et al.*,1985) and AP-1 (Lee *et al.*,1987) which suggests that these may act in conjunction with each other to modulate transcription.   In SV40 it is thought that Sp1 interacts with enhancer binding factors as well as other promoter binding factors to stimulate both early and late transcription.   The SV40 enhancer sequence has been extensively studied and several transcription factor binding sites have been identified. The functional domains of this enhancer appear to have little stimulatory activity by themselves but act synergistically to give high levels of activity (Fromental,1988; Zenke

et al.,1986). Individual domains are also found associated with other viral and cellular promoters such as the human metallothionein IIA gene and the human growth hormone gene (Imagawa *et al.*, 1987).

Several proteins have been identified as binding to the SV40 enhancer motifs, the best characterised being the activator proteins (AP) 1,2,3 and 4. Mermod *et al.*,(1988) demonstrated that AP-1 and AP-4 act coordinately to activate late transcription *in vitro*. However, with AP-2 and AP-3, although both factors activate transcription , footprinting analysis revealed that their recognition sequences overlapped and it appears that these factors act in a mutually exclusive manner (Mercurio and Karin,1989). It has now become clear from several sources that multiple factors can recognise the same sequence of DNA (Schaffner,1989; Graves *et al.*,1986; Dorn *et al.*, 1987; Santoro *et al.*, 1988; Clark and Hay,1989), indicating that the stimulatory or repressive nature of such a sequence may depend in the factors present in a given cell type.

Several other factors have been shown to interact with the SV40 enhancer, these include c/EPB (Johnson *et al.*, 1987) TEF-1 and TEF-2 (Davidson *et al.*, 1988), EBP1 (Clark *et al.*, 1988) and NFkB (Kawakami,1988).

Tjian *et al.*, (1987) proposed that it was the interplay between these factors with those that interact with the SV40 promoter region, such Sp1 and AP-2 (Mitchell *et al.*, 1987) that results in the high levels of both early and late transcription in SV40.

1.4 How Methylation Exerts its Effect on Transcription

1.4.1 The Effect of Methylation on DNA Structure and Stability

It appears that the presence of m^5C in DNA molecules can induce changes in the conformation of the molecule. For example, methylation of DNA is known to enhance the formation of left - handed Z DNA from the traditional right handed anti-parallel form. Behe and Felsenfeld (1981) found that the transition of poly(dG - mdC) from B to Z forms occurs at close to physiological salt concentrations, whereas conversion of the unmethylated polymer requires very high salt concentrations. Fox (1986) studied changes in the precise local DNA structure of methylated DNA by examining its sensitivity to digestion by DNAase I. He found that the presence of a methyl group renders the bond on its 5' side more susceptible to cleavage, suggesting that it causes small changes in the local DNA phosphate orientation.

Cytosine methylation also decreases the buoyant density of DNA (Dawid *et al.*, 1970), probably due to an increase in volume due to the protrusion of the methyl group into the major groove, and increases the thermal stability. This increased helix stability is similar to that found in double stranded molecules containing thymine in place of uracil (Cassidy *et al.*,1965). This property has been demonstrated using denaturing gradient gel electrophoresis of methylated and unmethylated linear DNA fragments (Collins and Myers,1987). Murchie and Lilley (1989) examined the effect of base methylation of inverted repeats on the rate of cruciform extrusion; the results obtained are consistent

with the above, in that increased stability caused by cytosine methylation inhibited the

initial opening of the DNA.

Although it appears that some properties of DNA are influenced by

methylation, the change in the overall structure of the molecule would be less than that

induced by a point mutation.

1.4.2 The Effect of Methylation on Chromatin Structure

DNA methylation probably inhibits gene expression by affecting the

protein DNA interactions required for transcription. This theory was first proposed by

Riggs (1975 and Razin and Riggs,1980) in light of evidence indicating that the binding to

DNA of proteins such as the lac repressor, histones and hormone receptors were affected

by changes in the major groove (Lin and Riggs,1972; Lin et al .,1976). For example,

changing a thymine residue to a uracil or cytosine in the major groove of the lac repressor

recognition sequence resulted in a decrease in affinity of the protein for the DNA.

Inserting m^5C at this site restores affinity, indicating that the repressor only senses the

presence or absence of a methyl group at this position (Fisher and Caruthers,1979).

A clear example of methylation affecting the binding of proteins comes from

the bacterial restriction modification system (Yuan and Meselson,1970). Some restriction

endonucleases will not bind or cleave their recognition sequence on DNA in the presence of

methylated adenine or cytosine.

What is not clear about this theory is whether methylation directly inhibits

the transcriptional machinery, i.e. transcription factors and RNA polymerase, or whether

it creates altered chromatin states rendering the DNA inaccessible to these factors and hence indirectly inhibiting their interaction with DNA. Over the past few years much study has been directed to answering this question. Results to date indicate that there is evidence for both mechanisms and that it appears that the mechanism involved depends on the gene and possibly the tissue in question.

Insight into the altered chromatin mode of action can be gained by examining the chromatin structure of genes introduced into cells by DNA mediated gene transfer. Keshet *et al.*, (1986) found that unmethylated DNA adopts a DNAase I sensitive structure characteristic of active genes, whereas fully methylated sequences are resistant to DNAase I and are structurally similar to the inactive genes of the cell. Further support for this theory came from studies on the Herpes virus *tk* gene where experiments show that the inhibitory effect of methylation is most likely due to an altered chromatin state. Keshet *et al.*, (1986) showed that methylation in the body of this gene, thought to be devoid of sequence specific protein binding, can inhibit expression even when the clearly defined promoter region is unmethylated. This suggests that methylation is altering the level of expression by mechanisms that do not involve the direct alteration of recognition sequences for RNA polymerase. Further studies on this gene by Buschhausen *et al.*, (1987) showed that inhibition of this gene by methylation occurred with a latency period of 8 hours after microinjection into mouse LTK⁻ cells, indicating that chromatin formation is necessary for inactivation. Ben-Hatter and Jiricny (1988) looked at the effect of promoter methylation on the expression of this gene. They observed that methylation (on both strands) of any of the four CpGs present resulted in a 20 fold reduction in the expression of the *tk* gene in *Xenopus* oocytes. This dramatic effect is unlikely to be due to the inhibition of transcription factor binding since in the same study

it was shown that the binding of both Sp1 and CTF is unaffected by methylation. Hence it is more likely that the presence of these methyl groups is resulting in the formation of altered chromatin structures which are not conducive to active transcription. This also appears to be the case for the γ-globin genes, where a minimum length of methylation -free DNA (100bp) in the 5' region is required for expression of these genes (Murray and Grosveld,1987).

The precise structure of this altered chromatin or how it is brought about is still unknown. However, it is thought that methylation prevents the formation of "activated" chromatin in the regulatory regions of genes. Activated chromatin containing DNAase I hypersensitive sites allows access of transcription factors and RNA polymerase to the DNA. This has been demonstrated *in vitro*, where the ability of the transacting factor TFIIIA to generate a transcriptionally active 5S gene template depends on the addition of the factor prior to or during nucleosome formation, but not after (Kmiec and Worcel,1985). However there appears to be no direct evidence to indicate how methylation might interfere with this process. It is possible that DNA methylation increases the local stability of the DNA, preventing cis acting DNA sequences from forming DNAase I hypersensitive sites. Alternatively a GC-rich region containing several methylated sequences could form an altered DNA structure, similar to Z DNA which may be resistant to packaging into nucleosomes and hence the activation process may not occur. Binding of a methyl cytosine specific factor such as the MDBP (methylated DNA binding protein), identified by Huang *et al* .,(1984) may amplify these effects of methylation. In the presence of such a factor the DNA would not be free to form single stranded regions of DNA in preparation for transcription.

1.4.3 The Effect of Methylation on Specific Protein-DNA Interactions

There are several known examples where methylation at sequence specific sites in promoter regions cause gene inactivation. Examples include the adenovirus E1a and E2a genes (Kruczek and Doerfler,1983;Langer *et al.*, 1984;1986), and the hamster adenine phosphoribosyl transferase (*aprt*) gene (Keshet *et al.*,1985). Hence it has been suggested that methylation may not only alter the overall structure of a gene but may also interfere with the binding of protein factors at discrete cis acting loci. Studies on the effect of methylation on the binding of several transacting factors (these are discussed in detail in section 5.10.), revealed the presence of both methylation sensitive and insensitive factors.

However it can be argued that the effect of methylation on binding *in vitro* is irrelevant to the situation *in vivo*. Methylation sensitivity *in vitro* may solely be determined by the stereochemistry of a given factor's interaction with its recognition sequence or the choice of key interactive bases within the sequence. Whereas *in vivo*, the methylation sensitivity of a gene may be determined by the DNA sequence, if the gene is methylation sensitive it will assume an inactive conformation inhibiting the access of transcription factors. The methylation sensitivity or insensitivity of these factors would hence be irrelevant.

However, Yisreali *et al* .,(1986) demonstrated that tissue specific factors can overcome the repressive effects of methylation. The methylated α- actin gene was inactive when introduced into fibroblasts but active when introduced into myoblasts. This can be explained if one considers the presence of a tissue specific transcription factor. Binding of this factor to sequences upstream of the gene may result in the formation of a

transcriptionally active form of chromatin. In the absence of this factor the presence of methyl groups throughout the gene may be sufficient to maintain an inactive form of chromatin. Such a situation may exist *in vivo* during differentiation, where the binding of such factors may be determined by the methylation state of their recognition sequences. This would be important in switching genes to tissue specific expression . After DNA replication, the binding of a transcription factor would, *via* the formation of active chromatin, render a gene active . However, in genes to be silenced in that cell type , the binding of such a factor could be inhibited by methylation, hence preventing the formation of active chromatin and transcription of the gene.

1.4.4 Reversing the Inhibitory Effect of Methylation

Several lines of evidence now suggest that promoter inhibition by methylation in certain genes is a reversible process. In addition to the active demethylation of certain sequences by proposed demethylase enzymes (Razin *et al* .,1986; Hughes *et al.,* 1989), more subtle mechanisms such as the partial release of the inhibitory action of methylation by a transactivating protein or by the juxtaposition of a strong enhancer have been proposed by Doerfler *et al.*, (1988). In the first instance it was observed that the methylation sensitive adenovirus 2 E2a gene could be partially reactivated in the presence of the adenovirus 5 E1a gene product (Langer *et al.*,1986; Weisshaar *et al.*,1988). In a similar set of experiments the activity of transfected methylated Ad 12 E1a and E2a promoters was restored by subsequently infecting the cells with frog virus 3 (FV3)(Thomson *et al.*,1986). Also, insertion of the strong enhancer of

the human cytomegalovirus (HCMV) into a construct containing the E2a gene resulted in high levels of expression in transfection studies irrespective of methylation state of the gene (Knebel-Mordorf, *et al.*,1988).

The alleviation of the repressive effects of methylation was also observed in a mammalian gene. As described in the previous section, Yisreali *et al.*, (1986) observed that the methylated α-actin gene is activated in myoblasts but not fibroblasts. Only the former expresses the endogenous gene. In myoblasts demethylation of specific sites occurred in the exogenous gene resulting in an identical pattern of methylation to the endogenous gene. This is consistent with the observation that tissue specific genes are highly methylated in germline DNA and during very early development, then undergo demethylation in a tissue specific manner, probably involving tissue specific determinator proteins (Yisreali and Szyf, 1984).

1.5. Aims of Project

It is generally accepted that there exists an inverse relationship between DNA methylation and transcription. However, it is still unclear how methylation interferes with transcription or how important a role methylation plays in the complex control processes of eukaryotic gene expression.

Recent advances in the identification and purification of cellular factors that stimulate gene expression have provided some insight into answering these questions. Two mechanisms of transcriptional inactivation by methylation have been proposed: (a) an

indirect effect whereby the presence of methyl groups in and around a gene results in the formation of an inactive form of chromatin, rendering the gene inaccessible to the transcriptional machinery. (b) A direct effect whereby methyl groups in the promoter regions of genes inhibit the binding of sequence specific transcription factors, hence limiting transcription.

The aim of the present study was to look at the effect of DNA methylation on eukaryotic transcription using the well characterised SV40 promoter as a model system. In previous studies using this promoter, all but one of the experiments involved the use of bacterial DNA methylases (Fradin et al., 1982; Waechter et al., 1982). No effect on early transcription was observed. The disadvantage of using these enzymes is that they only allow modification of cytosines in the enzyme recognition sequences, such sites do not always lie in the regions of genes that control transcription. In the SV40 promoter, despite having a high CG content, containing a cluster of six GC boxes, there are no bacterial methylase recognition sites. Since this region is known to be extremely important for transcription in SV40, use of these enzymes to study the effect of methylation on transcription would not be conclusive.

A more relevant experiment was performed by Graessman et al.,(1983), who methylated the CpGs in the SV40 promoter using mammalian DNA methylase purified from rat liver. No effect in the level of early or late gene expression was observed upon microinjection of this DNA into CV1 cells. However, the SV40 genome is CpG deficient, containing only 54 dinucleotides in 5243bp, hence in this experiment the SV40 promoter is contained in methylation deficient surroundings. It was the objective of the present study to examine the effect of methylation on transcription from the SV40 promoter in a CpG rich environment. This can be achieved by linking the promoter to a reporter gene,

the chloramphenicol acetyltransferase (CAT) gene contained within a bacterial plasmid.

Unlike eukaryote DNA prokaryote DNA is not CpG deficient. This plasmid DNA can then be

methylated *in vitro* using DNA methylase isolated from mouse Krebs II ascites tumour

cells and used in transient expression assays. Such an experiment would provide

information on the effect of regional methylation on transcription.

In addition, studying the effect of methylation on the binding of the

transcription factor Sp1 to DNA would provide a possible mechanism for transcriptional

inactivation by DNA methylation. This can be achieved by the use of synthetic

oligonucleotides containing m^{5}C. The use of these DNA molecules in gel retardation assays

allows the effect of methylation on transcription factor-promoter interactions to be

analysed *in vitro*. Since Sp1 is also known to bind to several cellular promoters such an

experiment would provide information on the possible role of DNA methylation *in vivo* .

Chapter 2. Material and Methods

2.1. Materials

2.1.1. List of Suppliers

Cell Growth Medium

Eagle's Minimal Essential Medium (EMEM) and the Glasgow modification of Eagle's medium (GMEM) were supplied as 10x concentrates by Gibco, Paisley, U.K. Newborn calf serum, sodium bicarbonate (7.5%), L-glutamine (200mM), penicillin (10,000U/ml) / streptomycin (10,000μg/ml) and non-essential amino acids (100x) were also obtained from Gibco.

Bacterial Growth Media

Bactotryptone, yeast extract and Bactoagar were obtained from Difco, Detroit, U.S.A.

Fine Chemicals

Amberlite MB3, acrylamide, ammonium persulphate, ammonium sulphate (enzyme grade), bromophenol blue, caesium chloride, EDTA, 99% formic acid (Analar), 8-hydroxyquinoline, NN' methylene bis/acrylamide, perchloric acid (PCA), polyethyleneglycol 6000, NNN',N'-tetramethylene diamine (TEMED), trichloroacetic acid

(TCA), SDS and xylene cyanol, were obtained from BDH Chemicals, Poole, Dorset, U.K.

Acetyl- CoA, agarose, ampicillin, 5-bromo-4-chloro-3-indolyl-B-galactosidase

(X-gal), butyryl-CoA, chloramphenicol, deoxynucleotide triphosphates (dNTPs) ethidium

bromide, isopropyl-B-D-thiogalactoside (IPTG), Low Melting Point (LMP) agarose, *M.*

luteus DNA, nucleotide triphosphates (NTPs), phenol (Ultrapure),

phenylmethylsulphonylfluoride (PMSF), salmon testis DNA and urea (Ultrapure) were

obtained from Sigma Chemical Co., Poole, Dorset,U.K.

Chloroform, ethylacetate, glycerol and xylenes were obtained from May & Baker,

Manchester, U.K.

Deoxynucleotide triphosphates (dNTPs) and dithiothreitol (DTT) were supplied by

Boehringer Mannheim, East Sussex, U.K.

Salicylic acid and tetramethylpentadecane (TMPD) were supplied by Aldrich Chemical Co.,

Gillingham, Dorset, U.K.

All other reagents were of the highest grade available. All solutions were prepared, where

appropriate with de-ionised, distilled water (dH$_2$0).

Chromatographic Media

DEAE-Sepharose, CNBr activated Sepharose and Heparin Sepharose were obtained from

Pharmacia (GB) Ltd., London, U.K.

Enzymes

Proteinase K and T4 polynucleotide kinase were supplied by Boehringer Mannheim

Bacterial restriction endonucleases, T4 DNA ligase, Klenow fragment of *E.coli* DNA

polymerase I were supplied by Gibco BRL Ltd., Paisley, Scotland, U.K.

Lysozyme and trypsin were supplied by Sigma.

Restriction Enzyme Digestion Buffers

Restriction enzyme digests were performed in the appropriate "REact" buffer. These

buffers were supplied with the restriction enzymes from Gibco BRL Ltd.

Radiochemicals

[γ-^{32}P]-adenosine triphosphate (6000Ci/mmol)

[α-^{32}P]-deoxyadenosine triphosphate (3000Ci/mmol)

[α-^{32}P]-uridine triphosphate (400Ci/mmol)

[α-^{35}S]- deoxyadenosine triphosphate (600Ci/mmol)

Chloramphenicol D- threo-[dichloroacetyl-1,2-^{14}C] (54Ci/mmol)

S-adenosyl-[^{3}H-CH$_3$]-methionine (0.5Ci/mmol, 15Ci/mmol)

were obtained from Amersham International, plc, Amersham U.K.

Bacterial Strains

E. coli HB101 (Boyer and Roulland, 1977) was the host used for the growth of all plasmid

DNA. It has the following genotype:F⁻,*pro, leu, thi, lac* Y, *hsd* R, *end* A,*rec* A, *rsp* L20, *ara*

14, *gal* K2, *xyl* 5, *mt* 11, *sup* E⁴⁴·

E.coli TGl cells (T. Gibson, unpublished) were the host used for propagation of M13 DNA.

It has the following genotype: Δ(*lac, pro*), *thi, sup* E, F'*tra* Δ36, *pro* AB, *lac* I𝑞, *lac* Z

ΔM15, *hsd* Δ5 (*Eco* Kr-m-).

Both these strains are used routinely in the Department of Biochemistry.

Continuous Cell Lines

Hela S3 cells (Gey *et al.*, 1952) which are human, cervical carcinoma cells were used in

the preparation of whole cell and nuclear extracts for *in vitro* transcription and gel

retardation assays respectively.

L929 cells (Sanford *et al.*,1948) which are mouse fibroblast cells were used in transient

expression assays.

CVI cells (Sweet and Hilleman, 1960), a monkey kidney cell line was used for the

preparation of SV40 DNA.

Plasmids

pVHCl and pSVCl were constructed as described in text (figures 3.3. and 4.2.) pTK3CAT, p200,

pLW4 and pLW2 were a gift from Dr. J. Lang.

2.1.2. Media

Cell Culture Media

Glasgow modified Eagle's minimal essential medium (GMEM) was stored as a 10X concentrated solution. For use, the concentrate was diluted with sterile distilled water to 500ml and supplemented with 50ml newborn calf serum, 20ml of 7.5% sodium bicarbonate, 5ml of (200mM) L-glutamine and 5ml of penicillin/strepomycin (10,000U/ml / 10,000μg/ml).

Eagles minimal essential medium (EMEM) was made up in a similar manner but supplemented with 50ml newborn calf serum, 5ml (200mM) L-glutamine, 5ml (100x) non essential amino acids and 5ml of penicillin - streptomycin (10,000U/ml / 10,000μg/ml). The pH was adjusted to 7.1 by the addition of approximatetly 250μl of 5M NaOH.

All freshly made media were checked for contamination by a 2-3 day incubation at 37°C before storage at 4°C.

The complete composition of EMEM and GMEM is presented in table 2.1

Versene

A stock solution of versene (diaminoethanetetraacetic acid) 0.2g/l was made up in PBS A containing 1% phenyl red and autoclaved before use.

Non Essential Amino Acids

Eagle's non essential amino acids were supplied as a 100x concentrate which contained:

	g/l
L-Alanine	0.98
L-Asparginine-H_2O	1.50
L-Aspartic acid	1.33
L-Glutamic acid	1.47
Glycine	0.75
L-Proline	1.15
L-Serine	1.05

Table 2.1   Eagle's media formulations (amino acid and vitamin components)

	EMEM	GMEM
Amino acids (mg/l)		
L-arginine-HCl	126.0	42.0
L-cysteine	24.0	24.0
L-glutamine	292.0	292.0
L-histidine HCl. H_2O	42.0	21.0
L-isoleucine	52.0	52.4
L-leucine	52.0	52.4
L-lysine-HCl	72.5	73.1
L-methionine	15.0	15.0
L-phenylalanine	32.0	33.0
L-threonine	48.0	47.6
L-tryptophan	10.0	8.0
L-tyrosine	36.0	36.2
L-valine	46.0	46.8
Vitamins (mg/l)		
D-Ca pantothenate	1.0	1.0
choline chloride	1.0	1.0
folic acid	1.0	1.0
i- Inositol	2.0	2.0
nicotinamide	1.0	1.0
pyridoxal-HCl	1.0	1.0
riboflavin	0.1	0.1
thiamin-HCl	1.0	1.0

Bacterial Growth Medium

Media (liquid)	Constituents per litre
L- Broth (LB)	10g bactotryptone
	10g NaCl
	5g yeast extract
	(adjusted to pH 7.2 with NaOH)

2 X TY	16g bactotryptone
	5g NaCl
	10g yeast extract

made up to 1 litre with dH$_2$O.

Media (solid)	Constituents per litre
L - Agar	as LB + 15g agar

H - Agar	10g bactotryptone
	8g NaCl
	12g agar

H - top agar	10g bactotryptone
	8g NaCl
	8g agar

made up to 1 litre with dH$_2$O.

2.1.3. Solutions

Table 2.2 Commonly used solutions

During the course of this work a number of solutions were used repeatedly the composition of these solutions are presented here.

Buffer M

50mM Tris-HCl,pH7.8,1mM EDTA, 1mM DTT,

0.02%NaN$_3$, 10% glycerol

M$^+$ indicates that PMSF(60µg/ml) has been added just before use

PBS

A- 10mM Na$_2$HPO$_4$, 1.84mM KH$_2$HPO$_4$,

0.172mMNaCl, 3.35mM KCl, pH7.2

B- 5mM CaCl$_2$.6H2O

C- 4.6mM MgCl$_2$.6H2O

A,B and C were autoclaved separately, when appropriate, and mixed in ratio of 8:1:1 before use.

T.E.

10mM Tris-HCl, pH 8.0, 1mM EDTA

TBE

89mM Tris, 89mM boric acid, 25mM EDTA, pH8.3

Sterilising solutions

All solutions used in the preparation or manipulation of DNA or involvimg the use of tissue cultured cells were sterilised either by autoclaving, 15psi for 20 minutes, or by passing through a Millipore sterilising filter unit (0.22μm).

2.2 Methods

2.2.1 General Procedures

During the course of this work a number of procedures were frequently used. The following section describes these general procedures.

2.2.1.1 Culture of Continuous Cell Lines

Continuous cell lines such as mouse L929 cells and CV1 cells were grown in the Glasgow modification of Eagle's minimal essential medium (GMEM) as monolayers in $80cm^2$ tissue culture flasks. At confluency the cells were subcultured as follows: the medium was removed and the cells washed with 5ml of prewarmed trypsin-versene (0.25% trypsin : versene, 1:4) solution at $37^{o}C$. This was immediately removed and replaced with 5ml of fresh trypsin -versene. This was then incubated at $37^{o}C$ for 3-5 minutes while the cells became released into suspension. To prevent excess trypsin action 20ml of GMEM containing serum was added and the appropriate volume of cell suspension transferred to new flasks and fresh medium added to 25ml. The cells were gassed with 5% CO_2 in air (to create a pH of 7.3-7.5) and incubated at $37^{o}C$. Generally the cells were split 1:3 every 6-7 days.

Hela cells were grown in suspension with constant stirring in Eagle's minimal essential medium (EMEM) at $37^{o}C$. Cells were fed by adding an equal volume of medium every second day.

2.2.1.2. Infection of CV1 cells with SV40

CV1 cells were grown to approximately 80% confluency in six 175cm^2 tissue culture flasks in GMEM. The medium was removed and the cells washed with 10ml of serum free medium. SV40 infected cell sonicate was diluted 1:100 in serum free medium and 10ml of this was added to each flask creating a viral concentration of approximately 1 p.f.u. per 10 cells. The flasks were incubated at 37°C for one hour. Fresh GMEM containing 10% newborn calf serum was added and the cells incubated at 37°C for one week.

2.2.1.3 Isolation of SV40 DNA

SV40 DNA was isolated from infected CV1 cells according to the method of Hirt (1967). Six 175cm^2 tissue culture flasks of infected CV1 cells were harvested and SV40 DNA prepared as follows. The cells were washed in PBS A then lysed by adding 5ml of 0.6% SDS, 20mM EDTA, pH 8, to each flask and incubating at room temperature for 20 minutes. 1ml of 5M NaCl was added and the viscous lysate scraped into Corex tubes. After incubation at 4°C overnight the lysate was centrifuged at 8000g for 30 minutes at 4°C. The supernatant containing the viral DNA was carefully removed and extracted with an equal volume of phenol, then chloroform. The DNA was precipitated with two volumes of cold ethanol at -20°C overnight. The precipitate was collected by centrifugation and washed with 70% ethanol at -20°C, dried by lyophilization, and resuspended in T.E. buffer (300μl). The DNA was then subjected to CsCl/EtBr centrifugation, to remove

contaminating cellular DNA, as follows. The DNA was transferred to two 5ml VTi65 heat sealable centrifuge tubes and made up to 5ml with a solution containing CsCl (0.96g/ml) and EtBr (0.6mg/ml) in T.E. buffer. The density of the solution was checked using a refractometer and adjusted to a density of 1.59g/ml by the addition of solid CsCl when necessary. The tubes were sealed and centrifuged at 350000g for 16 hours in a Beckman Ultracentrifuge. Two bands were visible under normal light, the upper band represents cellular DNA and nicked circular SV40 DNA. The lower band represents supercoiled SV40 DNA. The latter was removed from the CsCl./EtBr gradient and the EtBr removed by extracting at least four times with two volumes of isopropanol. The DNA was then dialysed against several changes of 100 volumes of T.E. buffer at 4°C. The concentration and purity of the DNA was checked both spectrophotometrically and by agarose gel electrophoresis.

2.2.1.4. Storage of Bacterial Stocks

Long term storage was as a 50% (v/v) glycerol/LB solution at -20°C. Short term storage was on a suitable sealed agar plate at 4°C.

2.2.1.5. Transformation of Bacterial cells with DNA

Competent E.coli cells were prepared by a modified version of the method of Dagert and Erlich (1979). 0.5ml of an overnight 10ml LB culture of HB101 was used to inoculate 50ml of LB medium. The cells were grown at 37°C, with shaking for 4-5 hours

(to an optical density of 0.4 at 650nm). The cells were cooled on ice for 10 minutes then harvested by centrifugation at 2500 rpm ,4°C, for 10 minutes in a Beckman Benchtop centrifuge. The cells were gently resuspended in 20ml of ice cold sterile 2M $CaCl_2$ and left on ice for 20 minutes. The cells were pelleted as before and resuspended in 1ml sterile 50mM $CaCl_2$. Transformations were carried out in plastic, 5ml, sterile bijoux tubes. DNA (10 -20ng) was added to 100μl of competent cells and the mixture incubated on ice for 10 minutes, then heat shocked at 42°C for 5 minutes. 2ml of prewarmed (37 °C) LB medium was added and the cells incubated at 37°C for 1.5 hours before plating on a suitable, antibiotic-containing plate. Plates were incubated at 37°C overnight allowing colony formation.

2.2.1.6. Small Scale Isolation of Plasmid DNA

A modification of the method of Holmes and Quigley (1981) was used for small scale plasmid preparations.

A 5ml culture of plasmid transformed cells in L. Broth was grown up overnight. The cells were pelleted by centrifugation at 3000 rpm, 4°C, for 10 minutes in a Beckman Benchtop centrifuge. The supernatant was carefully removed and the cells resuspended in 1ml of dH_2O and transferred to a microfuge tube. The cells were pelleted by centrifugation for 2 minutes in a microfuge. The supernatant was removed and the cells resuspended in a buffer containing 50mM Tris-HCl, pH8.0, 8% Sucrose, 5% Triton X-100, 50mM EDTA. A fresh solution of lysozyme was prepared (5mg/ml) and 30μl of this added to the

resuspended cells and the solutions mixed by a brief vortex. Denaturation was achieved by a

90 second incubation in a boiling water bath. After centrifugation in a microfuge the

supernatant was removed to a fresh tube and the plasmid DNA precipitated by addition of

200µl 5M NH_4 acetate and 1ml of isopropanol followed by a 10 minute incubation at

-20°C. Plasmid DNA was recovered by centrifugation at 4°C for 10 minutes. The DNA

pellet was washed with 1ml of cold ethanol at 4°C and resuspended in 25µl of T.E. buffer.

2.2.1.7. Large Scale Isolation of Plasmid DNA

The alkaline lysis method of Birmboin and Daly (1979) was used to prepare

milligram amounts of pure plasmid DNA.

10ml of L-Broth containing ampicillin (100µg/ml) was inoculated with a single colony of

bacteria (usually HB101) transformed with the appropriate plasmid and incubated

overnight at 37°C. This culture was used to inoculate one litre of L-Broth. The bacteria

were grown at 37°C with vigorous shaking to an optical density of 0.8 at 630nm.

Chloramphenicol was added to a final concentration of 170µg/ml. This increases plasmid

copy number per cell by preventing the bacteria multiplying while still allowing plasmid

DNA replication. The culture was incubated for a further 16 hours at 37°C with vigorous

shaking.

The bacteria were pelleted by centrifugation at 4000g for 10 minutes at 4°C.

The supernatant was removed and the cells resuspended in 6ml of cold alkaline lysis 1

buffer(50mM glucose,10mM EDTA, 25mM Tris-HCl, pH 8.0).   The bacterial

suspension was then distributed between two 30ml Ti60 centrifuge tubes on ice. To each

was added 150µl of lysozyme (40mg/ml in the above buffer). After mixing the tubes were

incubated on ice for 30 minutes before adding 6ml of alkaline lysis 2 buffer (0.2M NaOH,

1% SDS) and a further 5 minute incubation on ice. 3M Na acetate, pH 4.8 (7.5ml per

litre culture) was added and incubation on ice continued for a further 60 minutes.

Centrifugation was then performed at 65,000g ,20°C, for 30 minutes in a Beckman

ultracentrifuge. The supernatant containing the plasmid DNA was retained, the volume

measured and isopropanol, 0.6x the total volume added. The DNA was precipitated by

incubation at room temperature for 10 minutes and pelleted by centrifugation 7,000g,

20°C, for 15 minutes in a Sorval SS-34 rotor. The DNA precipitate was resuspended in

T.E. buffer (30ml) and subjected to CsCl / EtBr centrifugation as follows. CsCl (28.6g)

and EtBr (1.8ml of a 10mg/ml stock) were added and the density of the solution checked

using a refractometer and adjusted to a density of 1.59 g/ml by the addition of solid CsCl if

necessary. The solution was transferred to VTi50 heat -sealable tubes and centrifuged at

200000g ,20°C, for 16 hours in a Beckman LH-8 ultracentrifuge. Long wave U.V. light

was used to visualise the plasmid DNA band which migrates ahead of the host chromosomal

DNA in the caesium gradient. Plasmid DNA was removed by side puncture of the tube using

a 21 gauge needle and a 5ml syringe to withdraw the band. Further purification was

achieved by a second centrifugation in a VTi65 rotor, in heat sealable centrifuge tubes at

350,000g, 20°C, for 16 hours. The lower band was collected as before and the EtBr was

removed by four extractions with two volumes of isoamylalcohol. The lower aqueous phase

was collected and made up to a final volume of 1ml with T.E. buffer and dialysed against

several changes of 100 volumes of T.E. buffer at 4°C. The DNA concentration was

determined spectrophotometrically and the purity examined by agarose gel electrophoresis.

2.2.1.8. Quantitation of Nucleic Acids by Optical Density (O.D.)

Measurement

The absorption at 260nm of various dilutions of DNA samples was measured using quartz 1ml cuvettes in a Cecil linear readout U.V. spectrophotometer. The concentration was estimated assuming that O.D.=1.0 is equivalent to 50μg/ml of double stranded DNA.

2.2.1.9. Extraction of DNA with Phenol/Chloroform and Precipitation with

Ethanol

DNA was routinely purified free of protein by extraction with phenol/chloroform and precipitated with ethanol. Solid phenol was melted and saturated with T.E. buffer. It was then mixed with an equal volume of chloroform:isoamylalcohol (24:1) and stored at 4°C.

The volume of the sample to be extracted was adjusted to 100μl with T.E., when necessary. 100μl of the phenol mixture was added and the solution was vortexed for 1-2 minutes, then centrifuged for 4 minutes in a microfuge. The upper aqueous layer was removed to a fresh microfuge tube and the extraction procedure repeated. The sample was then twice extracted with ether saturated with water to remove the residual phenol.

The DNA was precipitated by adding one tenth volume of 3M sodium acetate pH 6.0 and 2.5 volumes of cold ethanol. This was vortexed and placed at -20°C overnight or at -70°C for 1 hour.

2.2.1.10. Restriction Enzyme Digests of DNA

DNA was incubated at 37°C for at least 3 hours with the appropriate restriction enzyme (10U of enzyme per µg of DNA) in the presence of the appropriate "REact" (BRL) buffer in a total volume such that the DNA is at a final concentration of 0.1µg/µl. For digestion with two enzymes the reactions were carried out simultaneously if the enzymes operated under the same salt conditions. If not the two digestions were carried out separately. After the first digestion the DNA was phenol/chloroform extracted and ethanol precipitated. The DNA was pelleted by centrifugation and resuspended in a small volume of buffer and the second digestion performed.

2.2.1.11. Agarose Gel Electrophoresis

Electrophoresis on 1%(w/v) agarose gels was used routinely to visualise and quantitate plasmid and SV40 DNA and to check the result of restriction enzyme digestions. The agarose was dissolved in TBE buffer containing 0.5µg/ml EtBr by boiling for 5 minutes. It was then cooled to approximately 50°C before pouring. One tenth volume of BPB-containing sample buffer (15% Ficoll, 100mM EDTA, 0.1% BPB) was added to the DNA samples before being applied to the gel. Electrophoresis was carried out at ambient temperature at 70 volts, 30mA for 2-3 hours in TBE buffer containing 0.5 µg/ml EtBr.

2.2.1.12. Purification of DNA from Low Melting Point Agarose

Restriction fragment DNA from pVHC1 and pSVC1 to be used for cloning into

M13 vectors or for gel retardation assays was isolated from 1-1.5% low-melting point

agarose gels. The gels were prepared and run as described in section 2.2.1.11. The

separated bands were visualised under longwave U.V. light and the appropriate band excised

using a sterile scalpel. The gel slice was transferred to a sterile 1.5ml microfuge tube

and an equal volume of T.E. buffer added . The agarose was melted at 65°C for 5 minutes.

The DNA was extracted first with phenol/T.E. then phenol,as described in section 2.2.1.9.

The aqueous phase, containing the DNA was then subjected to three extractions with 1ml of

ether to ensure all phenol was removed. The DNA was precipitated at -20°C with 2.5

volumes of cold ethanol in the presence of 0.3M Na acetate for 16-20 hours.

2.2.1.13. Non-denaturing Polyacrylamide Gel Electrophoresis

A stock solution of 30%(w/v) acrylamide:1%(w/v) NN'-methylene

bisacrylamide was prepared in dH_2O. The solution was deionised using "Amberlite" mono

bed resin MB-1 and filtered through Whatman No.1 filter paper, then stored at 4°C.

50ml of the appropriate % gel mix was prepared using the following components:

5ml TBE x 10

400µl 10% ammonium persulphate

appropriate volume of stock acrylamide

dH_2O to 50ml

16µl TEMED

TEMED was added last and the mixture poured immediately. Electrophoresis was carried

out in TBE buffer at 30mA for 2 hours. Gels were dried under vacuum at 80°C for 2 hours

using a Bio-rad model 1125 gel drier.

2.2.1.14. Denaturing Polyacrylamide Gel Electrophoresis

This method was used to analyse synthetic oligonucleotides and to purify single

stranded restriction fragments for use in gel retardation assays. The polyacrylamide stock

solution was prepared as above. The components of the gel were in the same concentration

as for non-denaturing gels with the addition of urea to a final concentration of 7M.

Samples were heated to 95-100°C for 5-10 minutes before loading.

2.2.1.15. Photography of Gels

DNA was visualised by ethidium bromide (EtBr) fluorescence on a trans -

illuminator (U.V Products Inc.). Gels were photographed with a Polaroid CU-5 camera and

type 665 positive/negative film.

2.2.1.16 Autoradiography

Dried polyacrylamide or agarose gels were exposed to a sheet of Kodak-X-Omat

H-film with or without an intensifying screen at -70°C. The exposure times varied

between 1 and 10 days depending on the type and amount of radioactivity involved.

2.2.1.17. Fluorography

Agarose gels containing *Hpa* II digests of methylated pVHC1 were subjected to fluorography allowing the detection of tritium incorporation into the restriction fragments. This was performed according to the method of Chamberlain (1979). The gel was soaked for two hours in a solution of sodium salicylate made from equimolar amounts of salicylic acid and NaOH adjusted to pH 5-7. The gel was then dried as described previously and autoradiographed.

2.2.1.18. Estimation of Protein Concentration

Protein concentration was estimated by the method of Bradford (1976)

2.2.2. *In Vitro* Methylation of DNA with Mouse Ascites DNA Methylase

2.2.2.1. Preparation of DNA Methylase

Mouse Krebs II ascites tumour cells were obtained from mice 7 days after inoculation. Nuclei were prepared as described in Turnbull and Adams (1976). Cells were washed twice with PBS, each time the cells were pelleted by centrifugation at

2000rpm for 10 minutes in a Beckman Benchtop centrifuge. The cells were then swollen

in dH_2O containing PMSF (60µg/ml), pelleted (3000 rpm, 5 minutes), then resuspended

in 1% Tween 80 containing PMSF (60µg/ml). The cells were lysed by two passages

through a Yamato LSC LH-21 homogenizer and checked microscopically for cell lysis. The

nuclei were pelleted by centrifugation at 2000 rpm, 4°C, for 10 minutes, as before. The

supernatant was carefully decanted and the nuclei stored at -70°C until required.

DNA methylase activity was extracted from ascites nuclei by stirring them

with an equal volume of buffer M^+ containing 0.4M NaCl for 20 minutes. The nuclei were

pelleted by centrifugation at 3000 rpm, 4°C, for 15 minutes in a Beckman Benchtop

centrifuge and the supernatant labelled A and retained. The extraction process was

repeated by adding an equal volume of buffer M^+ containing 0.2M NaCl, the nuclei were

pelleted as before and the supernatant labelled B and retained. Supernatants A and B were

combined to give a nuclear extract which was then absorbed onto phosphocellulose and

eluted batchwise with buffer M^+ containing 0.5M NaCl. This material was made 50%

saturated with respect to ammonium sulphate [   ] and the precipitated material

collected by centrifugation (8000g, 20 minutes,4°C). The pellet was redissolved in a

minimum volume of buffer M^+ and desalted on a small column of Sephadex G-50. The

non-retarded fraction from this column was bound directly to an 8ml column of tRNA

Sephadex (made from *E.coli* tRNA and CNBr-Sephadex according to the method of Arnt-

Jovin *et al.*,1975). The tRNA column was washed with at least a column volume of buffer

M^+ and the enzyme eluted with buffer M^+ containing 0.18M NaCl. The 1ml fractions

collected were assayed for protein by measuring absorbance at 280nm and the protein peak

dialysed overnight against buffer M^+ containing 50% v/v glycerol at -20°C. Each

fraction was then assayed for enzyme activity.

2.2.2.2. DNA Methylase Assay

The enzyme was assayed in a reaction mixture containing: 5µg heat denatured

M.luteus DNA, S-Adenosyl [Me-^{3}H] methionine (3.3µM;1.08µCi), BSA (1.4mg/ml) and

5µl of enzyme, made up to 70µl with buffer M$^+$ containing 0.1M NaCl. After incubation at

37°C for one hour the reaction was stopped by adding 500µl of stopper solution (1% SDS,

2mM EDTA, 3% aminosalicylic acid (PAS), 5% n-butanol, 0.5M NaCl and salmon testis

DNA (0.5mg/ml) to act as a carrier). Protein was removed by adding an equal volume of

phenol mix (88% phenol/12% m-cresol/0.1% 8 hydroxyquinoline). After

centrifugation (microfuge, 20°C, 4minutes), the upper aqueous layer was removed to a

fresh tube and the DNA precipitated by the addition of two volumes of ethanol. The DNA was

pelleted by centrifugation and redissolved in 0.3M NaOH and incubated at 37°C for one hour

to remove RNA. The DNA was precipitated onto filter paper (Whatman 3MM, 2.5cm) by

washing 5 times at 4°C with 5% TCA. The DNA was washed with ethanol and ether and

dried under a stream of air then transferred to scintillation vials prior to solubilisation in

0.5M perchloric acid at 60°C for 20 minutes. "Ecoscint" scintillator (5ml) was added and

the radioactivity counted on a liquid scintillation counter.

2.2.2.3. Methylation of Plasmid DNA

1-10µg of plasmid DNA was methylated in a reaction containing 10µM

S-Adenosyl-(Me-^{3}H)-methionine (0.5Ci/mmol)*, 0.1mg/ml BSA, DNA methylase at a

concentration of 1U/pmol CpG present in buffer M$^+$. The final reaction volume varied

such that the plasmid concentration was at a concentration of $\boxed{0.1\mu g/\mu l}$. The reaction was

performed at 37°C for 4 hours.

*except when the methylated DNA was to be subjected to fluorography when

S-adenosyl-(Me-^{3}H)-methionine with a specific activity of 4.6Ci/mmol was used.

2.2.3. Chloramphenicol Acetyl Transferase Transient Expression Assay

2.2.3.1. Transfection of Cells with Plasmid DNA

All solutions used in these transfection experiments were sterilised either

by autoclaving at 15psi for 20 minutes or by passage through a 0.22μM Millipore filter

before use. All manipulations were performed in a flow hood to prevent microbial

contamination. The reactions were carried out in sterile 5ml plastic bijoux bottles. The

quantities specified are for transfection of approximately 10^6 cells grown in 25cm^2 tissue

culture flasks. Each reaction was carried out in duplicate.

10μg of plasmid DNA, 125μl 2M CaCl$_2$ and 30μg of salmon testis DNA (used as carrier

DNA) was made up to 1ml with sterile dH$_2$O and mixed thoroughly by inversion or

pipetting. The mixture was then added dropwise, with mixing, to a tube containing 1ml 2 x

HBS buffer (50mM Hepes, 280mM NaCl, 1.5mM Na_2HPO_4, 1.5mM NaH_2PO_4, pH7.12), to

ensure the formation of a fine precipitate.   It was then incubated at room temperature for

30 minutes, without agitation, to allow the formation of the $Ca_3(PO_4)_2$-DNA precipitate.

1ml of this precipitate was added dropwise into the medium of duplicate flasks

containing mouse L929 cells. It is important that these cells are in log phase when

transfected. For this reason they were set up the previous day at the appropriate density

(10^6/flask) from a stock flask.   The cells were incubated with the precipitate at 37°C for

12 to 16 hours . The medium was then removed and the cells washed twice with

prewarmed PBS before fresh prewarmed growth medium was added.   Incubation was

continued for a further 24 hours (unless otherwise specified).   The cells were then

harvested as follows. The medium was removed and the cell monolayer washed twice with

PBS.   1ml of TEN buffer (0.04M Tris-HCl, pH7.8, 0.01M EDTA, 0.12M NaCl) was then

added and the cells incubated for 5 minutes at room temperature before being scraped off

the dish using a rubber policeman.   The cells were pelleted by centrifugation at 1,000

rpm in a microfuge for 5 minutes at 4°C and resuspended, by pipetting, in 90µl of 0.25M

Tris-HCl, pH7.8 on ice.   The cells were lysed by three cycles of freeze / thawing

consisting of 5 minutes incubation on dry ice (-70°C) followed by a 5 minute incubation at

37°C. The cell debris was pelleted by centrifugation at 5,000 rpm for 5 minutes in a

microfuge.   The supernatant was removed to a fresh tube.   Aliquots of the supernatant

were used for protein estimation and for assaying CAT activity and the remainder stored at

-20°C.

2.2.3.2. Chloramphenicol Acyl Transferase Assay

(a) According to the method of Cohen et al. (1980)

A fresh solution of 4mM acetyl CoA was used for each set of assays. 10μl of this was added to an microfuge tube containing 100μM ^{14}C chloramphenicol (54Ci/mmol), 30-60μl cell extract, in 0.1M Tris-HCl, pH 7.8, in a final volume of 90μl. The reaction was performed at 37°C for 1 hour, then stopped by the addition of 400μl of ethyl acetate. The tubes were then shaken for 15 seconds. The organic and aqueous layers were separated by centrifugation at high speed in a microfuge for 30 seconds. The top, organic layer was removed to a fresh tube and lyophilised. The chloramphenicol products were resuspended in 30μl of ethyl acetate and spotted onto silica thin-layer chromatography (TLC) plates. Ascending chromatography was in 95% chloroform and 5% methanol. The TLC plates were allowed to dry and then subjected to autoradiography with an exposure time of 1-2 days. The radioactive spots were marked, cut out and removed to scintillation vials. 5ml of "Ecoscint" was added and the radioactivity counted on a liquid scintillation counter.

(b) According to the method of Seed and Sheen (1988)

30-60 μl of cell extract was added to a microfuge tube containing a buffered solution of substrates constituted to give a final volume of 100μl, and final concentrations of 100mM Tris-HCl, pH 8.0, 100μM [^{14}C] chloramphenicol (54Ci/mmol) and 250μM butyryl CoA (fresh). After incubation at 37°C for 1 hour, the reaction was terminated by

addition of two volumes of a 2:1 mixture of TMPD(tetramethylpentadecane): xylenes, and

mixed vigorously by vortexing. After centrifugation for 1 minute in a microfuge,90% of

the upper, organic phase was removed to a fresh tube and extracted with an 100μl of

100mM Tris-HCl, pH8.0. The solutions were mixed as before and the organic layer

removed to a scintillation vial and counted.

2.2.4. Gel Retardation Assays

2.2.4.1. Preparation of Hela Cell Nuclear Extract

Hela cell nuclear extracts were prepared according to the method of Dignam

et al., (1983).

Hela cells were grown in stirring flasks at 37°C in EMEM containing 10% calf serum.

They were grown to a density of 3 to 5 x 10^5 cells per ml prior to harvesting, 4 to 7 x10^8

cells were used for a typical extract preparation.

Hela cells were harvested from cell culture media by centrifugation (at room temperature)

for 10 minutes at 2000g . Pelleted cells were then suspended in five volumes of PBS

(4°C) and collected by centrifugation as above; all subsequent steps were performed at

4°C. The cells were resuspended in five packed cell volumes of buffer A (10mM Hepes (pH

7.9), 1.5mM $MgCl_2$, 10mM KCl, 0.5mM DTT, 0.5mM PMSF) and incubated on ice for 10

minutes. The cells were collected by centrifugation as before and resuspended in two

packed cell volumes of buffer A and lysed by homogenisation (10 strokes). The homogenate

was checked microscopically for cell lysis and centrifuged at 2000g, 4°C, for 10 minutes

to pellet nuclei. The supernatant was removed and the nuclei subjected to a second

centrifugation for 20 minutes at 25000g, to remove residual cytoplasmic material. The

nuclei were resuspended in 3ml of buffer C (20mM Hepes, pH 7.9,25% (v/v) glycerol,

0.42M NaCl, 1.5mM MgCl$_2$, 0.2mM EDTA, 0.5mM PMSF and 0.5mM DTT) per 10^9 cells,

using ten strokes of homogenisation. The resulting suspension was stirred gently with a

magnetic stirring bar for 30 minutes at 4°C, then centrifuged for 30 minutes at 25000g.

The resulting clear supernatant was dialysed against 50 volumes of buffer D (20mM

Hepes,pH7.9,20% (v/v) glycerol, 0.1MKCl,0.2mM EDTA, 0.5mM PMSF and 0.5mM DTT)

at 4°C. The dialysate was centrifuged at 25000g for 20 minutes as before, the

supernatant, designated the nuclear extract, carefully removed and stored at -70°C in

aliquots.

2.2.4.2. Fractionation of Hela Cell Nuclear Extract

(a) Using Heparin Sepharose Affinity Chromatography

Nuclear extracts were prepared as in section 2.2.4.1. After dialysis a final

volume of 1.4ml (2.5-3mg/ml) of nuclear extract was obtained. A 200μl aliquot was

saved as nuclear extract, the remainder was applied to a 2ml heparin agarose column

equilibrated in TM buffer (50mM Tris-HCl, pH 7.9, 20% glycerol, 12.5mM MgCl$_2$, 1mM

EDTA, 1mM DTT) containing 0.1M KCl. The column was then washed with 1.5 column

volumes of the same buffer and the "run through" collected. The bound protein was then

eluted stepwise as follows; the column was washed with 1.5 column volumes of TM

containing 0.2M KCl, then 2 column volumes of TM containing 0.4M KCl. Finally the

column was regenerated by washing with TM plus 0.1M KCl. 0.2ml fractions were

collected throughout, the protein content was estimated as described in section 2.2.1.18.

Approximately 5µg of protein from each fraction was used in a gel retardation assay.

(b) Using DEAE Sepharose Ion -Exchange Chromatography

Nuclear extract was prepared as described in section 2.2.4.1. 1.6ml(4mg)

was loaded onto a 1.5ml DEAE Sepharose column pre-equilibrated in buffer D (20mM

Hepes, pH7.9, 20% (v/v) glycerol, 0.1M KCl, 0.2mM EDTA, 0.5mM PMSF and 0.5mM

DTT). The column was washed with 1.5ml of buffer D and 100µl fractions were collected

at 4°C. Approximately 5µg of protein from each fraction was used in a gel retardation

assay.

2.2.4.3.   End Labelling of Restriction Fragments

Restriction fragments of DNA were purified from low-melting point agarose

(section 2.2.1.12.) and end labelled using the Klenow fragment of DNA polymerase and

$[\alpha- ^{32}P]$-dATP, and used in gel retardation assays.

After ethanol precipitation the DNA was dried by lyophilisation.The following reaction

components were added to the dried DNA:-

3.75µl 10X medium salt buffer (10mM Tris-HCl, pH7.5, 50mM NaCl,10mM $MgCl_2$)

3µl (30µCi) $[\alpha -^{32}P]$-dATP (3,000 Ci/mmol)

1µl 100µM dTTP,dCTP, dGTP

1U Klenow fragment of DNA Polymerase I

dH_20 was added to give a final reaction volume of 25µl

The reaction was performed at 25°C for 30 minutes. To remove unincorporated

deoxyribonucleotide the DNA was extracted twice with ethanol, as follows. 90μl 2.5M NH₄

Acetate and 360μl of cold ethanol were added and the DNA precipitated by incubation on dry

ice (-70°C) for 10 minutes. The labelled DNA was pelleted by centrifugation in a

microfuge for 10 minutes at 4°C, then resuspended in 100μl 0.3M Na acetate, pH6.0, and

300μl of cold ethanol added. The DNA was precipitated for 15 minutes on dry ice, pelleted

as before and lyophilised until dry. It was then resuspended in the appropriate volume of

dH₂O and used in gel retardation assays.

2.2.4.4. Synthesis of Oligonucleotides

Oligonucleotides were synthesised by the phosphoamidite method on an

Applied Biosystems model 381A DNA Synthesizer by Dr. V. Math, Department of

Biochemistry. This involved successive diester bond formation between the 5'-OH group

of one nucleotide derivative and the 3'-OP of another (the nucleotide derivative in this case

being dialkyl phosphoamidite) resulting in a phosphite triester which must be oxidised

(e.g. with aqueous iodine) to a stable phosphotriester before the next step. The potentially

reactive groups elsewhere in the molecules are "protected" by reversible chemical

modification. The oligonucleotide chain is chemically synthesised in a 3' to 5' direction.

The first nucleotide is directly linked to a solid support (e.g. silica gel) packed in a column,

allowing the excess reagents to be washed off the column after each step. At the end of the

synthesis the oligonucleotide must be chemically released from the column and deprotected.

The DNA was then purified by ethanol precipitation.

2.2.4.5. Annealing and Ligation of Oligonucleotides

Synthetic oligonucleotides were annealed as described in Kadonaga and Tjian

(1986). 10μg of each of the two complementary oligonucleotides were annealed in a

reaction containing 67mM Tris-HCl,pH7.6,13mM MgCl$_2$, 6.7mM DTT, 1.3mM spermidine

and1.3mM EDTA in a final volume of 20μl. The annealing was performed in at 88°C for 2

minutes, 65°C for 10 minutes, 37°C for10 minutes, followed by 5 minutes at room

temperature. An aliquot was analysed by electrophoresis on a 2% agarose gel and the

remainder stored at -20°C, for later use, or ligated as follows. The DNA was precipitated

by ethanol, dried and resuspended in a total volume of 40μl containing 66mM Tris-HCl,

pH7.6,6.6mM MgCl$_2$, 0.5mM ATP,10mM DTT and 10U of T4 DNA ligase. The mixture was

incubated at 16°C for16-24 hours. The DNA was phenol extracted and ethanol precipitated

and an aliquot analysed on a 2% agarose gel.

2.2.4.6. End Labelling of Oligonucleotides

Both single and double standed oligonucleotides were end labelled using

polynucleotide kinase and [γ-^{32}P]- ATP and used in gel retardation assays. 100ngof

synthetic oligonucleotide in a total volume of 11μl containing 0.05M Tris-HCl, pH8.0,

0.01M MgCl$_2$, 5mM DTT, 12pmoles (60μCi) [γ-^{32}P] ATP and 10 units T4 polynucleotide

kinase was incubated at 37°C for 1 hour. The labelled oligonucleotide DNA was recovered

and separated from unincorporated radioactivity by two ethanol precipitations, as follows.

40μl 2.5M NH₄acetate and 160μl of cold ethanol were added to the incubation mix and the DNA was precipitated at -70°C for 15minutes. The DNA was pelleted by centrifugation at 4°C for 10minutes in a microfuge. The pellet was resuspended in 100μl 0.3M Na acetate, 300μl of cold ethanol was added and the DNA precipitated and collected as before. The labelled oligonucleotides were dried by lyophilisation for 2-3 minutes.

2.2.4.7. Base Composition Analysis of Oligonucleotides

5μg of oligonucleotide DNA was hydrolysed using 8% formic acid at 180°C for 90 minutes. The formic acid was evaporated and the bases dissolved in ammonium carbonate buffer (20mM, pH 10.0, at room temperature) and applied to a column of Aminex A6 (BioRad; 27cm x 12cm). The column was then eluted with the same buffer at 50°C with a flow rate of approximately 1ml/minute. Bases were quantified by measuring absorbance at 260nm. This analysis was performed by Dr. R. Adams.

2.2.4.8. Gel Retardation Assay

A 91bp Sph I/Nco I restriction fragment of pSVC1 was ^{32}P end labelled as described in section 2.2.3.4. 3-5ng of this was incubated in the presence of 2-10μg of Hela cell nuclear extract in a total volume of 10μl containing 20mM Hepes (pH7.9), 20%v/v glycerol, 0.1MKCl, 0.2mM EDTA, 5mM DTT and 5μg of poly(dl-dC) DNA. The assay mixture was incubated at 25°C for 30 minutes. 10μl of sample buffer (15% Ficoll,

100mM EDTA, 0.1% bromophenol blue) was then added and the samples loaded onto a 4%

polyacrylamide gel. The gel was prerun for 30minutes at $20Vcm^{-1}$, then run at

$10Vcm^{-1}$ for 2.5-3 hours after the samples were applied. Gels were dried as described

and autoradiographed without an intensifying screen.

2.2.5. *In Vitro* Transcription Assays

2.2.5.1. Preparation of Hela Cell Extract

Soluble whole cell extracts were prepared from Hela cells by the method of

Manley *et al* .,(1980) for use in *in vitro* transcription assays.

Hela cells were grown in suspension culture in Eagle's minimal essential medium

supplemented with 10% newborn calf serum to a density of approximately $8x10^5$

cells/ml. Generally 1-1.5 l of cells were used for each preparation. All steps were

performed between 0 and 4^oC. Cells were pelleted by centrifugation at 3000 rpm,4^oC,

for 10 minutes in a Beckman Benchtop centrifuge. The pelleted cells were washed in PBS

by resuspending them in 3-5 ml of cold PBS followed by centrifugation as before. The cell

pellet was resuspended in 4 packed cell volumes of buffer I (0.01M Tris-HCl, pH9.0,

0.001M EDTA, 0.005M DTT). After a 20 minute incubation on ice, the cells were lysed by

8 strokes of homogenisation. Four packed cell volumes of buffer II (0.05M Tris-HCl,pH

7.9, 0.01M $MgCl_2$,0.002M DTT,25% sucrose,50% (v/v) glycerol) were then added and

the mixture stirred gently on ice. To this suspension one packed cell volume of saturated

$(NH_4)_2SO_4$ was added dropwise, the lysate was gently stirred for an additional 20minutes.

The extract was then centrifuged at 180000g for three hours in a Beckman Ti60 rotor.

The supernatant was carefully decanted to a fresh tube and precipitated by the addition of

solid $(NH_4)_2SO_4$ (0.33g/ml of suspension), slowly with stirring. After the $(NH_4)_2SO_4$

had completely dissolved, 0.01ml of 1M NaOH per 10g $(NH_4)_2SO_4$ was added and the

suspension stirred on ice for 20 minutes. The precipitate was collected by centrifugation

at 15000g, 4°C, for 20 minutes and resuspended in 1/10th the volume of high speed

supernatant of buffer III (50mM Tris-HCl,pH7.9,6mM $MgCl_2$, 40mM $(NH_4)_2SO_4$,0.2mM

EDTA, 1mM DTT, 15% glycerol). This suspension was then dialysed for 4-8 hours

against two changes of 100 volumes of the same buffer. The diasylate was centrifuged at

10000g, 4°C, for 10 minutes and the supernatant aliquotted into small samples and stored

at -70°C.

2.2.5.2. *In Vitro* Transcription Assays

In vitro transcription assays were performed in a final volume of 20μl

containing 10μl(20-25μg) of Hela cell extract, 250μM CTP, GTP,ATP, 25μM (10μCi)

[α-^{32}P]-UTP (400Ci/mmol) and 100-200ng of DNA template. The reaction was

performed at 30°C for 1 hour. The reaction was terminated by the addition of 50μl of

proteinase K mix (proteinase K, 1mg/ml, 0.05% SDS, 0.5mg/ml *E.coli* tRNA) and a

further 20 minute incubation at 37°C. 90μl of 1M NH_4 acetate was added and the protein

removed by the addition of 200μl of phenol/ chloroform followed by a 1 minute vortex and

centrifugation for 4 minutes in a microfuge. The upper aqueous layer was removed to a

fresh tube and the phenol back-extracted by addition of 200μl of 1M NH$_4$ acetate, mixed

and centrifugated as before, the upper aqueous layer was then removed and mixed with the

previous one. The nucleic acids were then precipitated by the addition of 2.5 volumes of

cold ethanol, followed by incubation on dry ice (-70°C) for 15 minutes. The RNA was

dried by lyophilisation and redissolved in 12μl of dH$_2$O. 8μl of formamide dye mix

(deionised formamide containing 0.3% BPB, 20mM EDTA) was added and the samples loaded

onto a 5% denaturing polyacrylamide gel or a 1.4% agarose gel and electrophoresed for

3hours at 250V.

2.2.6. Cloning Into M13 Vectors

The protocols used to generate recombinant M13 molecules were as described in

the " M13 Cloning and Sequencing Handbook" (Amersham International Plc).

2.2.6.1. Large Scale Preparation of M13 RF DNA

5ml of a 1:100 diluted overnight culture of TG1 cells was inoculated with a

single plaque of virus-containing cells. Cells were grown with vigorous shaking for 4

hours. The cells were then pelleted by centrifugation at 3000 rpm, 4°C, for 10 minutes

in a Beckman Benchtop centrifuge and the virus-containing supernatent used to inoculate a

500ml culture of exponentially growing (with an Optical density of 0.2-0.3 at 600nm)

TG1 cells. Cells were grown for a further 4-5 hours at 37°C with shaking, then harvested

by centrifugation at 4000g, 4°C, for 10 minutes . Viral RF DNA was then isolated by the ·

alkaline lysis method described in section 2.2.1.7.

2.2.6.3. Preparation of M13mp19 RF DNA for Cloning

1μg of RF DNA was digested with the restriction enzymes *Eco* RI and *Kpn* I as

described in section 2.2.1.10. The extent of digestion was checked by agarose gel

electrophoresis. The DNA was phenol/chloroform purified free of enzyme and ethanol

precipitated, as described in section 2.2.1.9.

2.2.6.3. Preparation of Insert DNA

5μg of pVHC1 plasmid DNA was digested with 50U of *Kpn* I in low salt buffer

(10mM Tris-HCl, pH 7.5, 10mm $MgCl_2$, 1mM DTT) in conditions described in section

2.2.1.10. NaCl conditions were adjusted to final concentration of 100mM and the DNA

subjected to partial digestion with *Eco* RI as follows. 25U of *Eco* RI was added to the

incubation mix at 37°C and aliquots were removed after 2,5,10,20,30 and 60 minutes. The

DNA fragments were separated by electrophoresis on a 1% low-melting point agarose gel

alongside DNA size markers (as in figure 2.1.). The 601bp fragment containing the SV40

promoter-enhancer sequence and the 2107bp fragment containing the

promoter-enhancer-CAT gene and terminater sequence were excised from the gel and

purified.
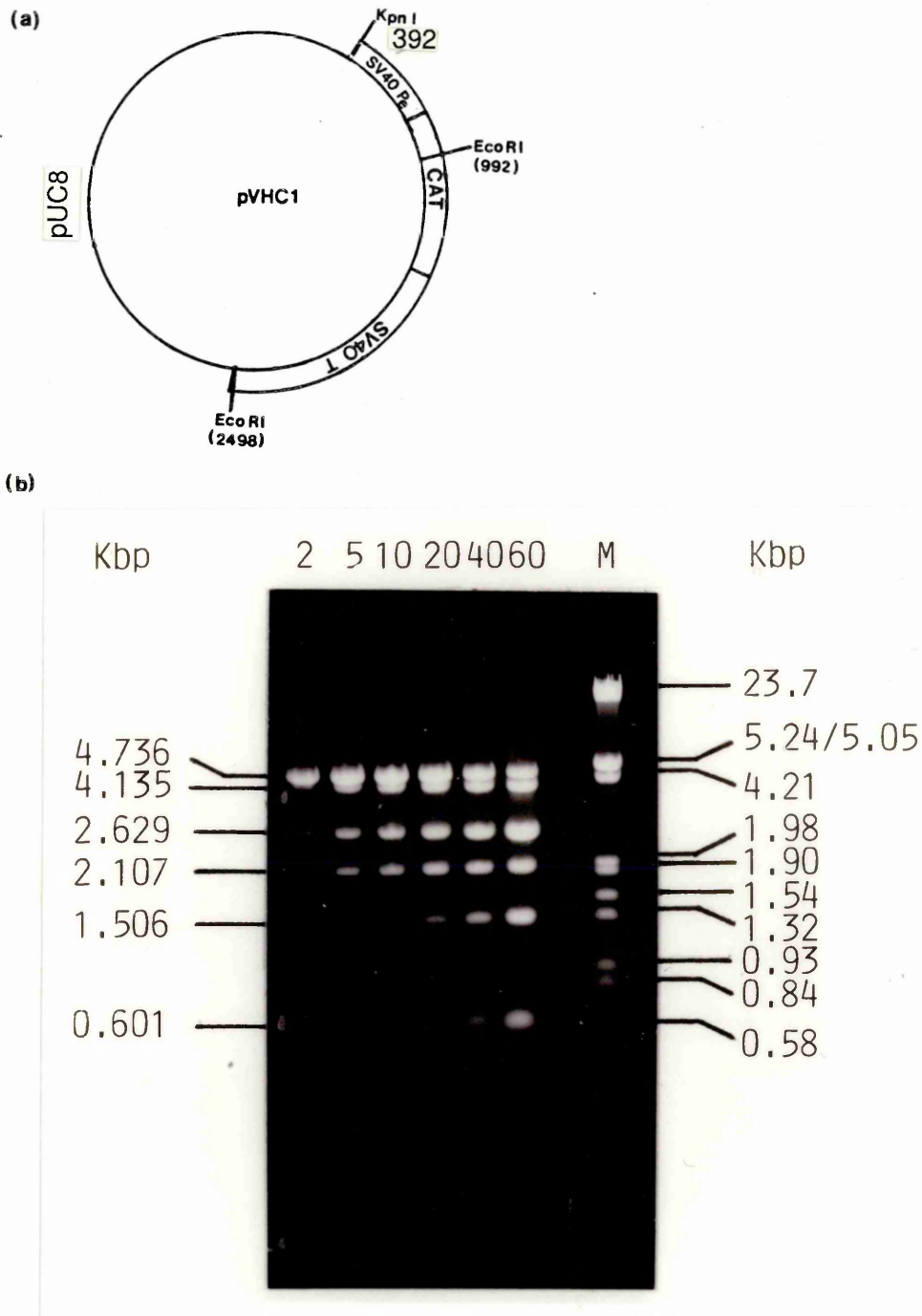
Figure 2.1. *Kpn* I / *Eco* RI Digestion of pVHC1

(a) *Kpn* I / *Eco* RI Restriction Map of pVHC1 (see figure 3.3.)

(b) *Kpn* I / *Eco* RI digestion of pVHC1

Kpn I digested DNA was treated with 25U of *Eco* RI for 2,5,10,20,30, or 60 minutes (lanes 1-6), as described in section 2.2.6.3., and subjected to electrophoresis on a 1% agarose gel.   Lane M contains λ DNA digested with *Hind* III and *Eco* RI, as marker DNA.

2.2.6.4. Ligations

Ligations were performed using bacteriophage T4 DNA ligase, in a final volume of 20µl containing 66mM Tris-HCl, pH7.6, 6.6mM $MgCl_2$, 0.5mM ATP, 10mM DTT, 100ng vector DNA and 20ng of insert DNA. 1U of DNA ligase was added and the reaction carried out at 15°C for 16hours.

2.2.6.5. Transformation of E. coli (TG1) cells with M13 DNA

TG1 cells were made competent for DNA uptake as described in section 2.2.1.5. The entire ligation mix was added to 300µl of competent cells and the mixture incubated on ice for at least 1 hour. The cells were then heat shocked at 42°C for 3 minutes and returned to ice. The transformation mix was then added to 3ml of molten top agar containing 40µl of 100mM IPTG, 40µl of 2% (w/v) X-gal (in dimethyl formamide) and 200µl of fresh (exponentially growing) TG1 cells 100mM IPTG, at 45°C. The contents were mixed and plated out on a prewarmed (37°C) H plate. Plates were incubated at 37°C overnight.

2.2.6.6. Preparation of Single Stranded Template DNA

The M13 cloning system allows the easy identification of recombinant phage by the inactivation of the β-galactosidase gene. A simple plaque assay can be used to

differentiate recombinants (colourless plaques) from non-recombinants (blue plaques).

A 10ml 2TY overnight culture of *E. coli* was used to inoculate fresh 2TY medium at a ratio of 1ml:100ml. A single plaque was lifted using a sterile Eppendorf pipette tip and used to inoculate 1.5 ml of the low density *E. coli* TG1 culture. The culture was shaken at 37^oC for 5 hours then centrifuged for 5 minutes in a microfuge. The supernatant was carefully removed to a fresh tube and recentrifuged to remove any residual cells. The supernatant was added to 200μl PEG/NaCl (20%w/v polyethylene glycol 6000 / 2.5M NaCl), mixed thoroughly and left for 15 minutes at room temperature. Precipitated phage were harvested by centrifugation in a microfuge for 5 minutes and residual PEG/NaCl removed by a second centrifugation and capillary action of a drawn out Pasteur pipette. The viral pellet was resuspended in 100μl T.E. buffer, 50μl of phenol/T.E. was added , the mixture vortexed for 20 seconds and left at room temperature for 15 minutes. The upper aqueous layer was transferred to a fresh microfuge tube and extracted with 500μl of chloroform. Na acetate (pH6.0) was added to a final concentration of 0.3M, 250μl of ice cold ethanol was added and the DNA precipitated at -20^oC overnight. The DNA was recovered by centrifugation at 4^oC (10 minutes, microfuge) and the template DNA resuspended in 20μl of T.E. buffer and stored at -20^oC. The purity of the DNA template was examined by electrophoresis on a 1% agarose gel.

2.2.6.7. T-Track Analysis

Potential recombinant clones were screened for insert by T-tracking and comparison with vector DNA.

For each ten recombinants to be analysed the following priming mix was prepared : 4µl

M13 primer(25µg/ml), 6µl 10x Klenow buffer(100mM Tris-HCl, pH7.8, 100mM

MgCl$_2$) and 12µl dH$_2$O. 2µl of this priming mix was annealed to 2µl of each template at

60°C, for 1.5-2 hours.

The M13 sequencing primer is a 17mer with the sequence 5' - GTA AAA CGA CGG CCA GT

-3'.

3µl (30µCi) of [α-^{35}S] dATP (600 Ci/mmol) and 16µl 0.5mM dTTP/0.5mM ddTTP mix[*]

and 2U of Klenow fragment of *E. coli* DNA polymerase I were mixed and 2µl of mix added to

each annealed primer/template. After a 20 minute incubation at room temperature, 1µl of

a 0.5mM uniform chase solution of all four dNTPs was added and the reaction continued at

room temperature for a further 15 minutes. Reactions were stopped by the addition of 1µl

of formamide dye mix (96% w/v deionised formamide, 0.03% w/v xylene cyanol FF,

10mM EDTA). The samples were incubated at 95-100°C for 3 minutes and the primer

extended DNA fragments resolved by denaturing polyacrylamide gel electrophoresis

(20x40x0.4cm). The gels were then fixed in 10% v/v acetic acid, 10% v/v methanol for

20 minutes. Gels were dried onto a sheet of Whatman 3MM paper using a Bio-rad model

1125 gel drier. Autoradiography was carried out as described in section 2.2.1.16 with

exposure times of 24-48 hours. Figure 2.2. shows identification of recombinant M13

molecules by comparison with M13mp18 and mp19 in a T-Tracking experiment.

---

* The dTTP/ ddTTP mix used also contains dCTP and dGTP at a concentration of 0.08mM.

Figure 2.2. T-Track Analysis of M13 Recombinants

T-Tracking was performed as described in section 2.2.6.7. Lanes 3,4,5,10,11 and 12 correspond to the sequence 5' $GAATTC_{992}$ _____ $CTTAA_{900}$ of pVHC1 indicating the presence of the Kpn I (392)/ Eco RI (992) insert. Lanes 1,2,8 and 9 are probably artefactual.
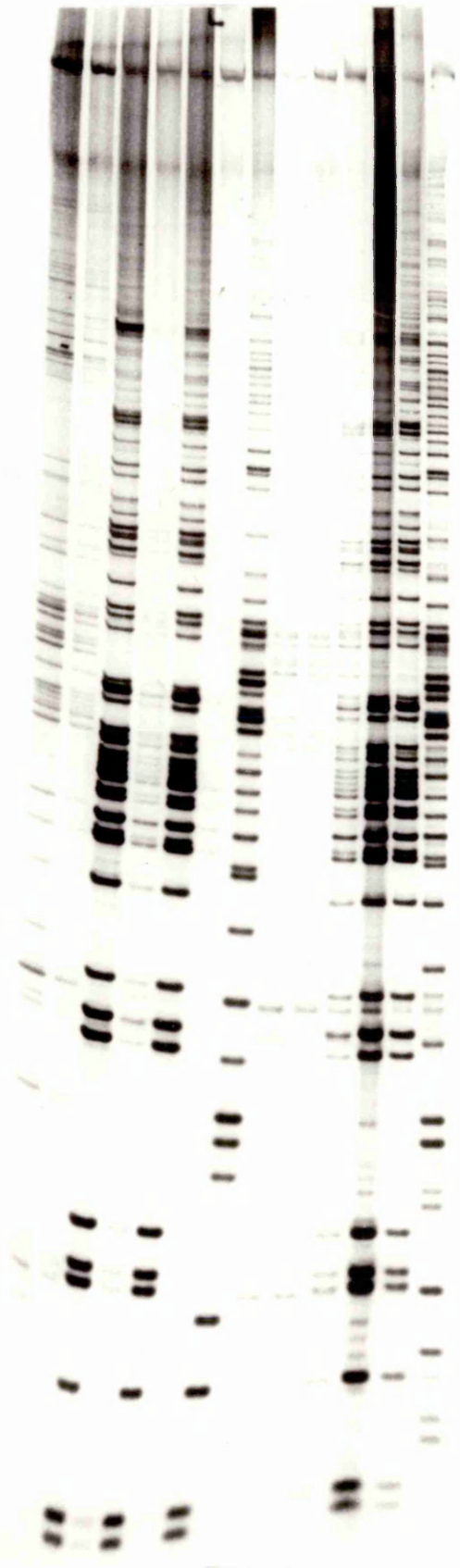
2.2.6.8. Synthesis of second strand

M13 single stranded recombinants were used as templates for preparation of double

stranded molecules containing methyl dCTP in one strand.

The universal primer was annealed as for T-Tracking (section 2.7.7.), using 5µl (1µg)

of single stranded recombinant DNA and 4µl of primer DNA (25µg/ml) and 2 µl of Klenow

buffer (100mM Tris-HCl ,pH 7.8, 100mM $MgCl_2$).

Primer extension was carried out as follows.   To the 11µl of the above annealing mixture

the following was added: 3µl dNTP cocktail containing 2.5mM dCTP, dTTP, dGTP, 3µl of

15µM dATP, 30µCi [α–^{32}P]-dATP (3000Ci/mmol) and 10U of Klenow fragment of DNA

polymerase in a final volume of 25µl.   The contents were mixed and incubated at room

temperature for 10 minutes before the addition of 3µl of 2.5mM dATP and the incubation

continued for a further 20 minutes.   The reaction was stopped by incubating at 68°C for

10 minutes.   The reaction mix was cooled and restriction enzyme digested by adding 2µl of

the appropriate "REact" buffer x10 and 10U of restriction enzyme.   The DNA was digested

at 37°C for three hours.   Protein was removed by phenol/chloroform extraction and the

DNA precipitated with ethanol. The second digest was then performed.   The reaction was

stopped by incubation at 68°C for 10 minutes.   The restriction fragment (usually Sph

I/Nco I) was separated from the remainder of the DNA by electrophoresis on a 7%

non-denaturing polyacrylamide gel.   After autoradiography of the gel the desired

restriction fragment was identified and cut out of the gel using a sterile scalpel blade.

500µl of sterile dH_2O was added and the acrylamide broken up by several passings through

a 1ml syringe. The DNA was extracted by incubating this at 65°C for 1-2 hours. The mixture was centrifuged for 10 minutes in a microfuge to pellet the acrylamide. The supernatent was removed and the radioactivity estimated by Cherenkov counting in a liquid scintillation counter. This provided an estimation of percentage recovery of the DNA. The labelled restriction fragment DNA was then precipitated with ethanol at -70°C and pelleted by centrifugation at 4°C for 10 minutes in a microfuge. It was then resuspended in the appropriate volume of dH_2O and used in a gel retardation assay.

Chapter 3. *In Vitro* Methylation of DNA
with Mouse Krebs II Ascites Tumour
Cell DNA Methylase

Introduction

In the following experiments DNA methylase purified from mouse Krebs II ascites tumour cells was used to methylate SV40 and plasmid DNA *in vitro*. Unmethylated, double stranded DNA is a poor substrate for the enzyme, requiring it to perform *de novo* methylation. Hence, conditions were optimised to obtain maximum incorporation of CH_3 groups into the DNA.

3.1. Purification of DNA Methyltransferase from Mouse Krebs II Ascites Tumour Cells.

Since tumour cells are rapidly dividing and therefore have a high level of DNA synthesis, they provide a good starting material for the purification of enzymes such as DNA methylase. Purification of DNA methylase from mouse Krebs II ascites tumour cells was first performed by Turnbull and Adams (1976). Improvements to the original method have also been published (Adams *et al.* 1979). DNA methylase was prepared as described in section 2.2.2.1.

A typical purification table is as follows:

FRACTION	TOTAL PROTEIN (mg)	TOTAL ACTIVITY (units)	SPECIFIC ACTIVITY (units/mg)
Nuclear Extract	1352	36600	27
Phosphocellulose	84	16100	192
tRNA-Sepharose	5.1	10100	1980

(1 unit = 1 pmol CH_3 /h)

Table 3.1.   Purification of DNA Methylase from Mouse Krebs II Ascites Tumour Cells

Preparation of the low salt nuclear extract from the original cell suspension generally gives a 30 fold purification.   Hence the overall purification is approximately 2000 fold and the enzyme is around 10% pure at this stage.

Preparation of DNA methylase for use in some of the later experiments was performed by M. McGarvey and in these purifications the tRNA Sepharose step was replaced by DNA Sepharose.   However, the level of purity remained similar.

DNA methylase has been purified to homogeneity in our laboratory (Adams *et al.*, 1986) by adding further affinity chromatography and gel filtration steps to the procedure.   However, these preparations were unstable as DNA methylase loses activity at low protein concentrations, and produced low yields of enzyme.   For these reasons, partially purified enzyme was used in these *in vitro* methylation experiments.

3.2. Stimulation of DNA Methylase Activity by Partial Proteolysis.

Adams *et al*.,(1983) reported that *de novo* methylase activity could be stimulated 3-30 fold by limited proteolysis of the enzyme. The native molecular weight of mouse DNA methylase is 185 KDa (Adams *et al.*, 1979). On treatment with trypsin, at an enzyme to protease ratio of 200, several peptides are generated. The M_r of the major species being 170, 100 and 50KDa. The stimulation is restricted largely to activity on native, unmethylated DNA substrates, such as plasmid DNA, and is much less obvious with a denatured or hemi-methylated substrate. Most mammalian DNA methylases exhibit a preference for denatured DNA over native DNA. For the mouse enzyme the ratio is approximately 5:2, this is reduced to 5:4 or 1:1 on partial proteolysis with trypsin. Hence to increase methyl group incorporation most plasmid methylation experiments were performed using pre-trypsinised methylase.

In general, the enzyme was treated with trypsin at a ratio of methylase to protease of 200 for 1-5 minutes at 37°C. The reaction was then stopped by adding a five fold excess of trypsin inhibitor.

3.3. Methylation of SV40 DNA

Duplex SV40 DNA contains 54 potential methylation sites (i.e. CpG dinucleotides), 70% of these exist in a 400 bp region containing the overlapping early and late promoters and the origin of replication (see figure 3.1).

SV40 DNA was methylated *in vitro* by incubating 8μg of DNA with 20U of DNA methylase and S-adenosyl (Me-^{3}H) methionine (3.3μM , 1.08μCi) in buffer M$^+$ in a total volume of 70μl. The reaction was performed at 37°C and 2μl (0.23μg) were removed at 0,1,2,4 and 6 and 24 hours. An additional 10U of methylase was added at 6 hours. Aliquots were assayed for tritium incorporation as described in section 2.2.2.2, except that toluene/PPO scintillant (efficiency 40%) was used.

Figure 3.2 shows that the reaction proceeds in a linear manner over this period of time with a maximum of 51% incorporation achieved at 24 hours.

Due to the lack of sufficient DNA methylase some of the methylation experiments presented in this chapter were not performed in duplicate. However, these experiments were always repeated more than once and the results of each experiment followed the same trend.
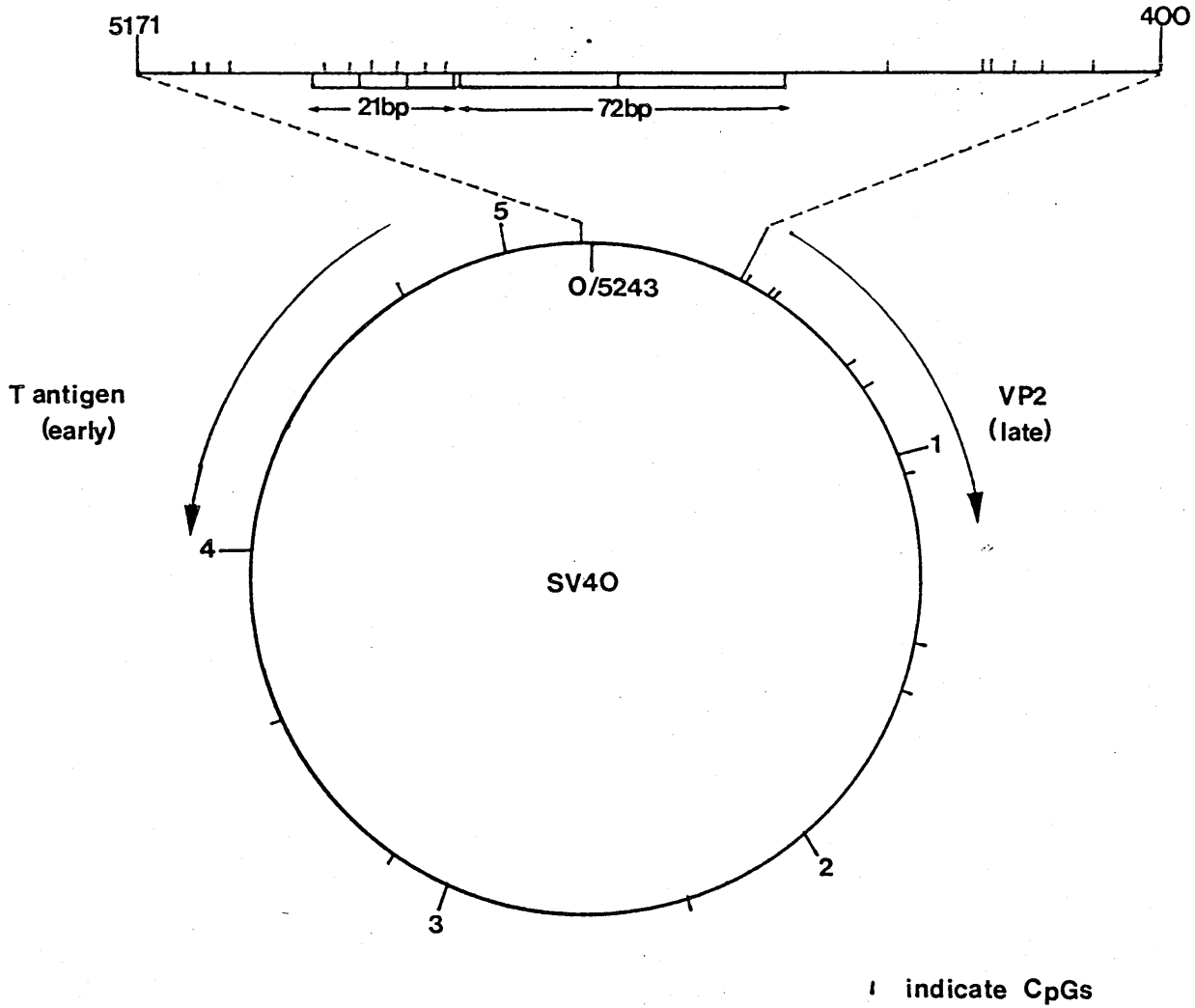
Figure 3.1.   Map of the SV40 genome showing the distribution of CpG
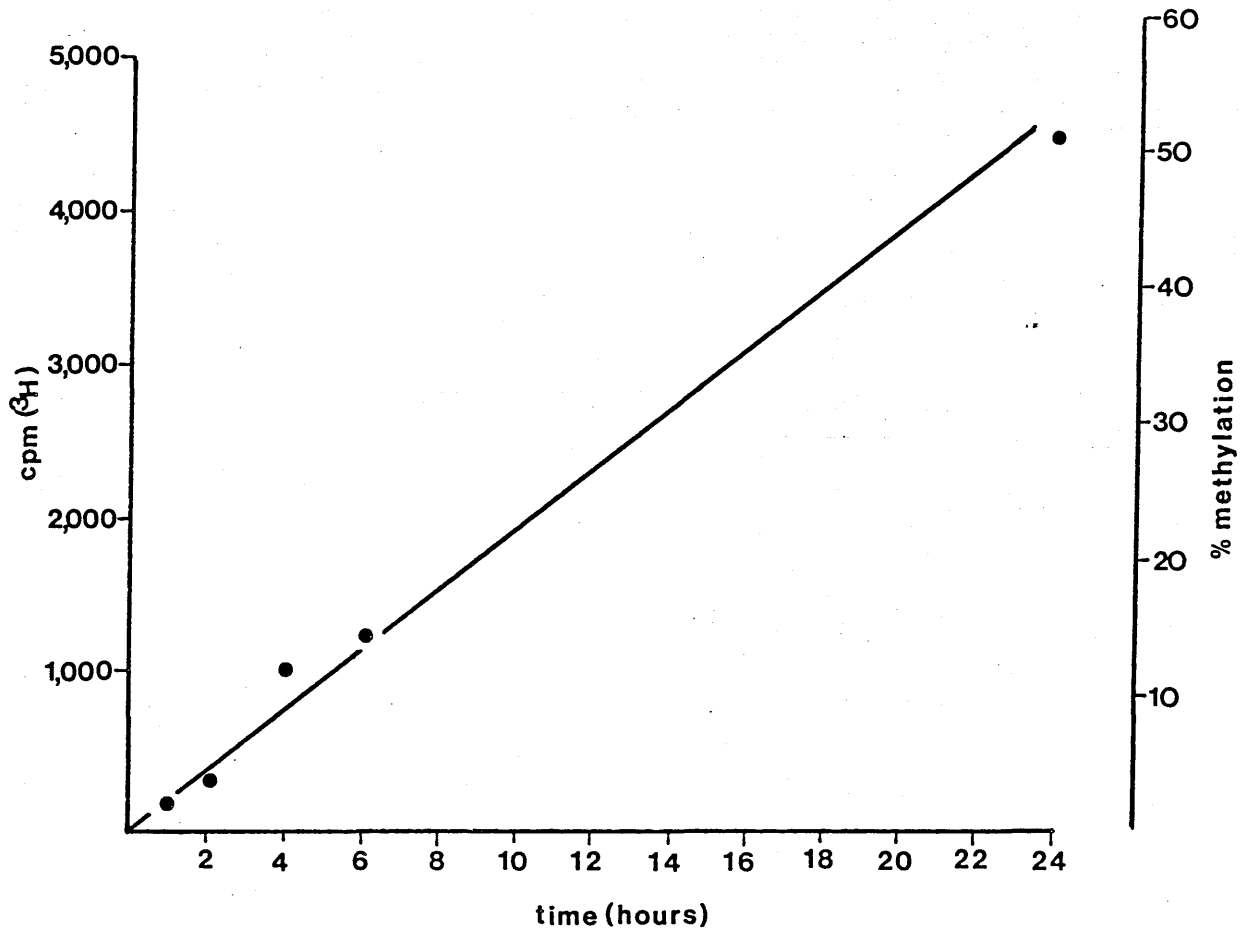
dinucleotides

Figure 3.2.   Methylation of SV40 DNA

SV40 DNA was methylated for 0-24 h as described in section 3.3.   The level of methylation is expressed as $cpm(^3H)$ incorporated and percentage of total CpG dinucleotides methylated.

3.4. *In Vitro* Methylation of Plasmid DNA

3.4.1. pVHCl

pVHCl is a pUC8 derivative containing the SV40 promoter region in the early direction linked to the bacterial chloramphenicol acetyl transferase (CAT) gene. Hence, the effect of modifications in the promoter will be reflected in the level of CAT expression when the plasmid is introduced into mouse cells. The plasmid is 4736bp in size, it retains the ampicillin resistance gene present in pUC8 and also contains a 950 bp SV40 fragment, containing a termination sequence, linked to the CAT gene.

pVHCl was constructed by inserting a 342bp *Pvu* II / *Hind* III fragment, containing the SV40 promoter region isolated from pSV2CAT (Gorman *et al.*,1982a) into a plasmid known as p200. p200 is a derivative of pTK3CAT minus the *tk* (Thymidine Kinase) promoter but containing a multiple cloning sequence isolated from a plasmid known as pIC20H (Marsh *et al.*, 1984) linked to the CAT gene and SV40 terminator (see figure 3.3). pTKCAT3 and p200 were a gift from Dr. J. Lang.

This plasmid was constructed by Sandra Hislop previously of our laboratory.
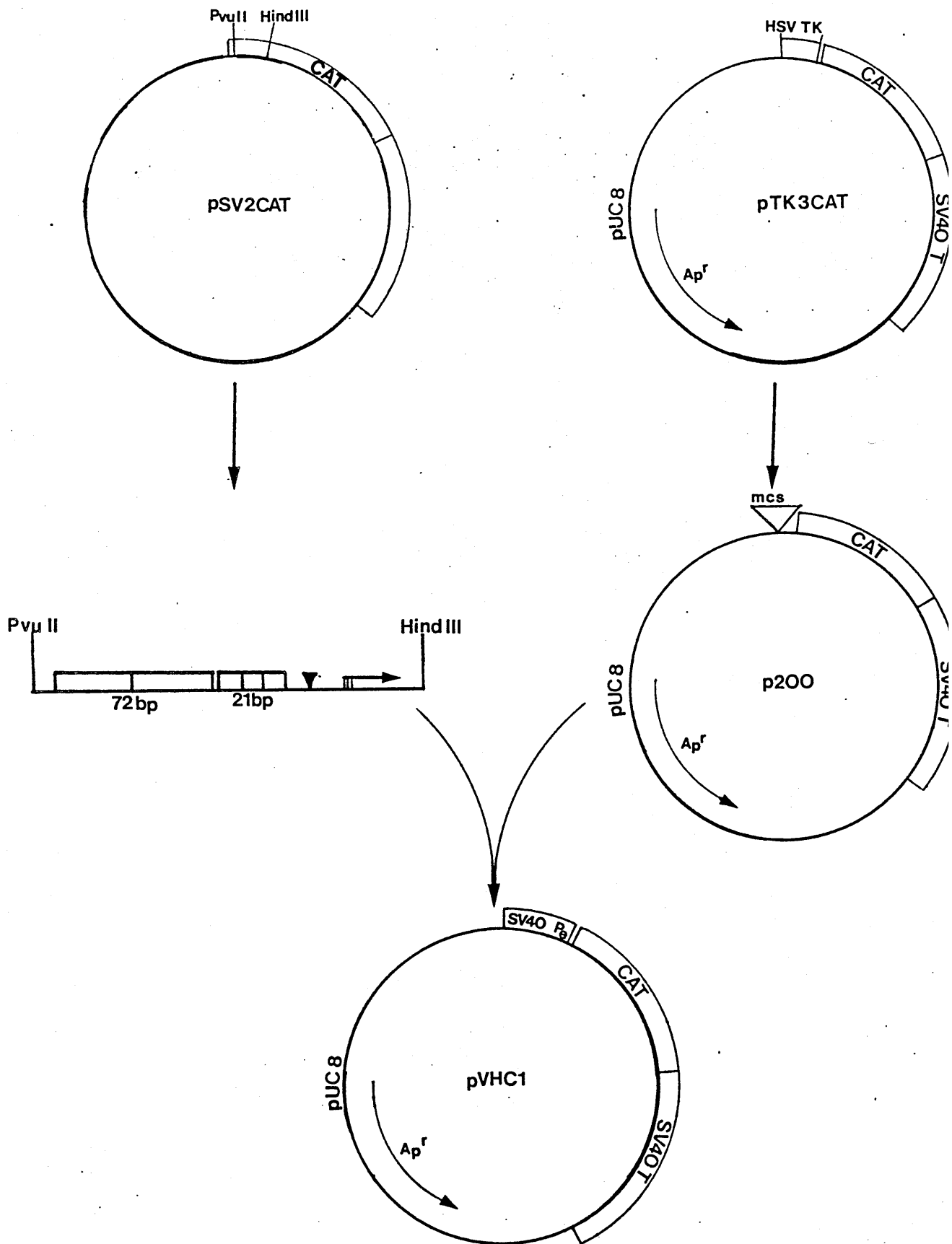
Figure 3.3.   Construction of pVHC1

Details are described in the text in section 3.4.1.

3.4.2 *In vitro* methylation of pVHCl

The following experiments were performed to achieve optimum conditions

for maximum incorporation of CH_3 groups at CpG dinucleotides.

3.4.2.1 DNA Methylase concentration curve

pVHC1 (0.2µg) was incubated in a total volume of 100µl containing S-adenosyl

(Me-^{3}H) methionine(3.3µM, 1.08µCi)and the specified concentration of enzyme in buffer

M$^+$ at 37°C for 4 hours. The reaction was stopped, the DNA recovered and the

incorporated tritium estimated as described in section 2.2.2.2.

Figure 3.4 shows that incorporation of methyl groups increases with increasing enzyme

concentration and plateaus at 1U of enzyme per pmol CpG.
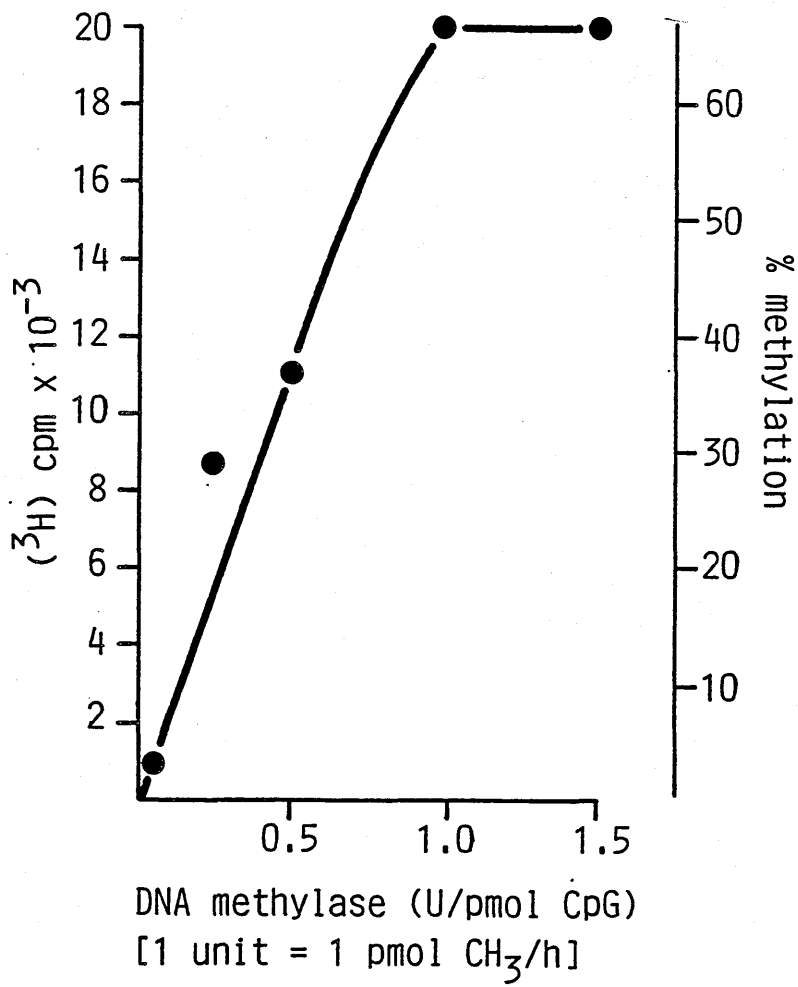
Figure 3.4. Methylation of pVHC1: Enzyme Concentration Curve

pVHCl was methylated using 0.5-1.5 U enzyme / pmol CpG as described in section 3.4.2.1.
The extent of methyl group incorporation is expressed as cpm of ^{3}H incorporated and
percentage of total CpG dinucleotides methylated.

3.4.2.2 Time Course Methylation pVHCl

1μg of pVHCl was incubated at 37°C for 0,2,4 or 6 hours in a total volume of

100μl containing S-adenosyl (Me-^{3}H) methionine (20 μM ,2μCi) , 140U of DNA

methylase and bovine serum albumin (1mg/ml) in buffer M$^+$.

After the specified incubation time, a 0.1μg DNA aliquot was removed and assayed for ^{3}H

incorporation as described in section 2.2.2.2. The remainder of the DNA was isolated by

stopping the reaction at 68°C for 10 minutes followed by phenol/chloroform extraction

and ethanol precipitation of the DNA. The level of methylation in this DNA can be analysed

by its resistance to cleavage by restriction enzyme *Hpa* II (Section 3.4.2.5).

Figure 3.5 shows that the reaction proceeds linearly over this time span, with a maximum

of 61% incorporation at six hours.

However, it was found that increasing the incubation time beyond six hours led

to a decrease in CH_3 group incorporation (this was also occasionally observed at 6 hours),

which meant that saturation was never achieved. This phenomenon was also observed by

Simon *et al*. (personal communication) with the rat liver enzyme and in both cases it was

due to a reduction in the level of recovery of the DNA. This was most likely due to the

formation of tight protein-DNA complexes which fail to dissociate at 68°C. This leads to a

loss of DNA in the subsequent phenol extraction - the tightly bound DNA remaining

associated with the protein at the interphase.

In more recent experiments the problem has been overcome by the inclusion of a

post-methylation trypsinisation step, causing complete proteolysis of the enzyme and
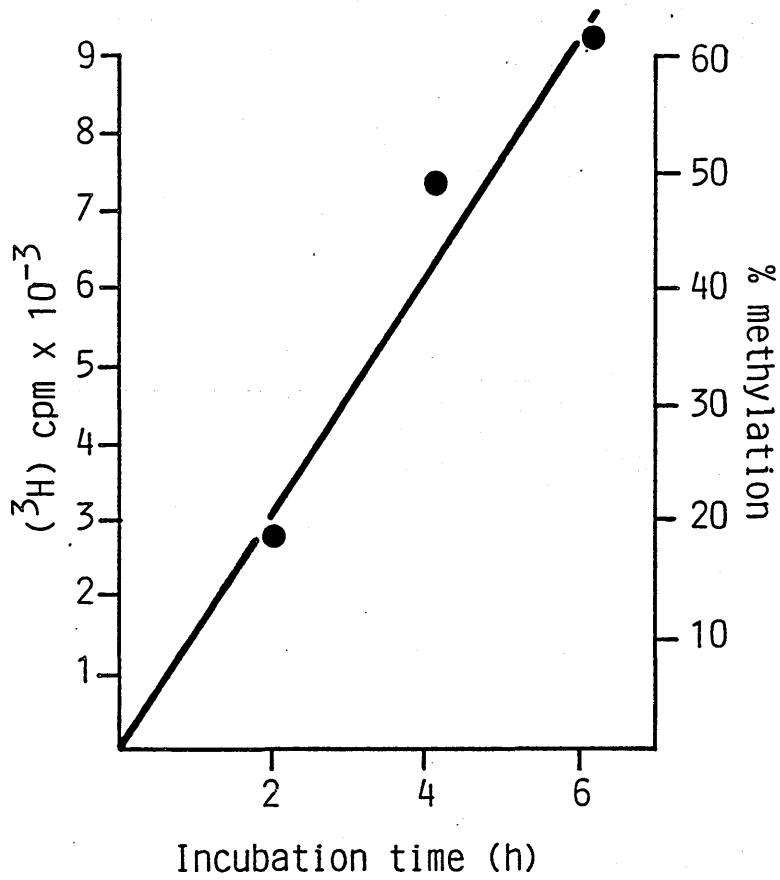
release of the DNA.

Figure 3.5. Time Course Methylation of pVHC1

1μg of pVHC1 DNA was methylated for 0,2,4, or 6h. The level of methylation is expressed as cpm of (^{3}H) incorporated per 0.5μg of DNA and percentage total CpGs methylated.

3.4.2.3. Effect of Temperature on DNA Methylase Activity

Duplex native DNA i.e. plasmid DNA, is not an ideal substrate for DNA methylase. The enzyme will only interact under conditions favourable for "DNA breathing" i.e. low salt and high temperatures. Drahovsky *et al* .,(1971) in studies using rat liver DNA methylase found that the enzyme may interact with DNA in a different manner depending on the temperature of the reaction. Interaction at 0^oC led to the formation of a loosely bound enzyme - DNA complex, dissociable by 0.2M salt. However, interaction at 37^oC resulted in a tightly bound salt resistant complex. In the latter case methylation proceeded in a linear fashion in the presence of salt, suggesting that this tight complex has formed and remains attached to the DNA between successive methyl transfers. Also, at higher temperatures DNA 'breathing' is encouraged which may allow a greater number of stable enzyme - DNA interactions to occur and hence increase the the number of CH_3 groups incorporated.

In an attempt to improve the efficiency of pVHCl methylation the effect of temperature on enzyme binding was examined.

0.5μg of pVHCl was incubated with S-Adenosyl (Me-^{3}H) methionine (20μM,1μCi) in a total volume of 40μl in buffer M* at 0, 20, 37, 40 or 45^oC for five minutes. 80U of DNA methylase was added to each tube and the incubation continued for a further five minutes before all samples were transferred to 37^oC for 4 hours.

Figure 3.6 shows that the optimum temperature for enzyme DNA interaction is 20^oC, and preincubation at higher temperatures decreased CH_3 group incorporation.
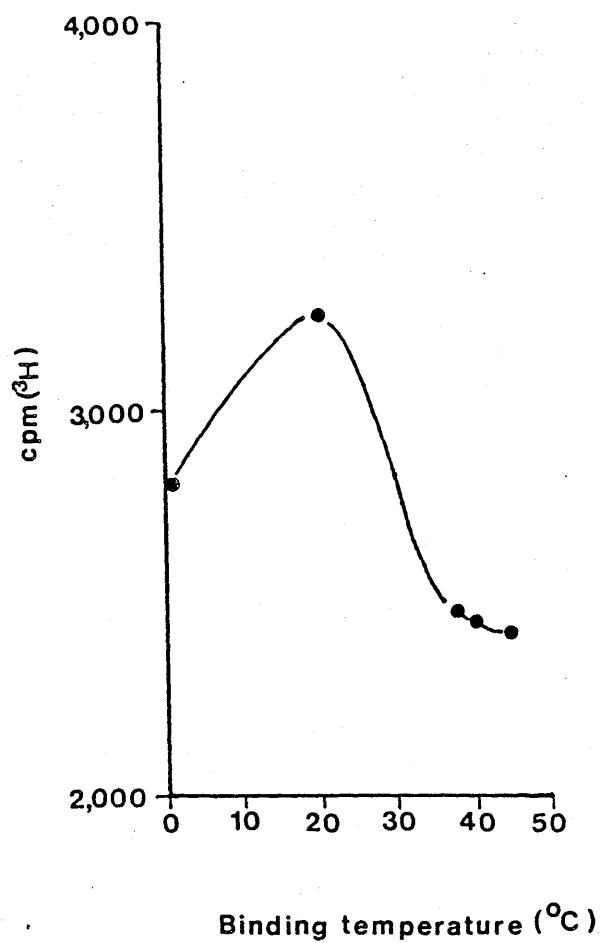
Figure 3.6. The Effect of Temperature on DNA Methylase Activity

0.5µg of pVHC1 was preincubated with 80U of DNA methylase at 0,20,37,40 or 45°C for 5 minutes. ^{3}H -SAM was added to each tube and the incubation continued at 37°C for 4 hours. The resulting levels of methylation are expressed as cpm of ^{3}H incorporated.

3.4.2.4. Estimation of Level of Tritium Incorporation

Since the sequence of the plasmid pVHCl is known the percentage of methylated CpG dinucleotides can be estimated as follows :

Specific Activity of S-Adenosyl (Me -^{3}H) Methionine is 5mCi /μmol i.e. 1pmol/10,000 dpm.

The efficiency of counting using liquid scintillant is 20%* and the estimated recovery is 50%.*

(*-see over)

Therefore 1pmol CH$_3$ = 1000cpm

The number of pmols of CpG present in any reaction is as follows :

$$\frac{\text{amount DNA (μg)} \times \text{No. of CpG dinucleotides/plasmid}}{\text{M}_r \text{ nucleotide (pair)} \times \text{No. bp/plasmid}} \times 10^6 = \text{pmols CpG present}$$

Therefore, percentage incorporation can be calculated :-

$$\frac{\text{cpm/assay}}{\text{No. pmols CpG} \times 1000} \times 100$$

*   The efficiency of counting for tritium was estimated from the H number obtained.

The H number is the method used in L-8000 series liquid scintillation systems to

determine the amount of quench in each sample and provides a means for correcting all the

data to some common point of reference.

From the published calibration curves (The H number concept , D.L. Hurrocks,Beckman

Ltd) the typical H number obtained (100-120) was consistent with an efficiency of

counting for tritium of 20-25%.

(The H number measures the response produced in any liquid scintillation sample by the

same electron energy and hence is a measure of the decrease in the light produced by that

amount of energy due to the quenching that is present in the sample.)

The estimation of % recovery was calculated from data obtained from experiments where

labelled ^{14}C DNA was included in the methylase assay.

i.e. 20ng of ^{14}C labelled DNA (41μCi/mg) was included in a methylase assay containing

5μg of *M. luteus* DNA.

Assuming an efficiency of counting for ^{14}C in channel 2 of 65% (from the H number

obtained), the expected cpm was 1172. 595 and 611cpm were obtained for duplicate

assays indicating 53% recovery of the DNA.

3.4.2.5. Resistance of Methylated pVHCl to Cleavage by *Hpa* II

Methylation sensitive restriction enzymes have proved to be a very useful tool in studies on DNA methylation. The methylation status of a CpG dinucleotide, in an enzyme recognition sequence can be assessed from its resistance or susceptibility to cleavage by that enzyme. The restriction enzyme *Hpa* II recognises and cleaves DNA at the sequence CCGG only if the inner cytosine is unmethylated.

Computer analysis of pVHCl revealed 14 *Hpa* II sites, and this was confirmed by a restriction digest of the DNA (see fig. 3.7). Treatment of methylated pVHCl with the enzyme allows the analysis of the extent of methylation of the DNA and some insight of how the methyl groups are distributed.

Fig. 3.7(a) shows that pVHCl DNA methylated for 0, 2, 4, or 6 hours becomes increasingly resistant to HpaII although even after 6 hours there are some molecules still completely susceptible to *Hpa* II cleavage and none completely resistant. The corresponding fluorograph of the gel is represented by lanes 7-10. Figure 3.7(b) shows an *Hpa* II digest of pVHCl DNA methylated to approximately 75% saturation, and the corresponding fluorograph of the gel. This pattern of *Hpa* II inhibition indicates that not all molecules are methylated to the same degree and the estimated level of incorporation is a mean value. From the fluorograph it is evident that those molecules most resistant to cleavage contain the highest level of tritium incorporation.

Msp I digestion of methylated DNA resulted in a few DNA molecules being resistant to cleavage at a small number of sites. This suggests that at these sites cytosine in the dinucleotide CpC has been methylated. It has been reported by Graessman *et al.*, (1983) that the rat liver methylase will methylate cytosines in the sequences CpA, CpT as

well as CpG.   When DNA methylated by the ascites enzyme was subjected to the nearest

neighbour analysis technique (Urieli-Shovel *et al.*,1982), Adams *et al* .,(1988) found

that untrypsinized DNA methylase methylated only cytosines in CpG dinucleotides.

However, a small proportion of methylcytosines in CpC and CpT dinucleotides were found

in DNA methylated by pre-trypsinized enzyme, indicating that some trypsinized enzyme

molecules have reduced specificity.
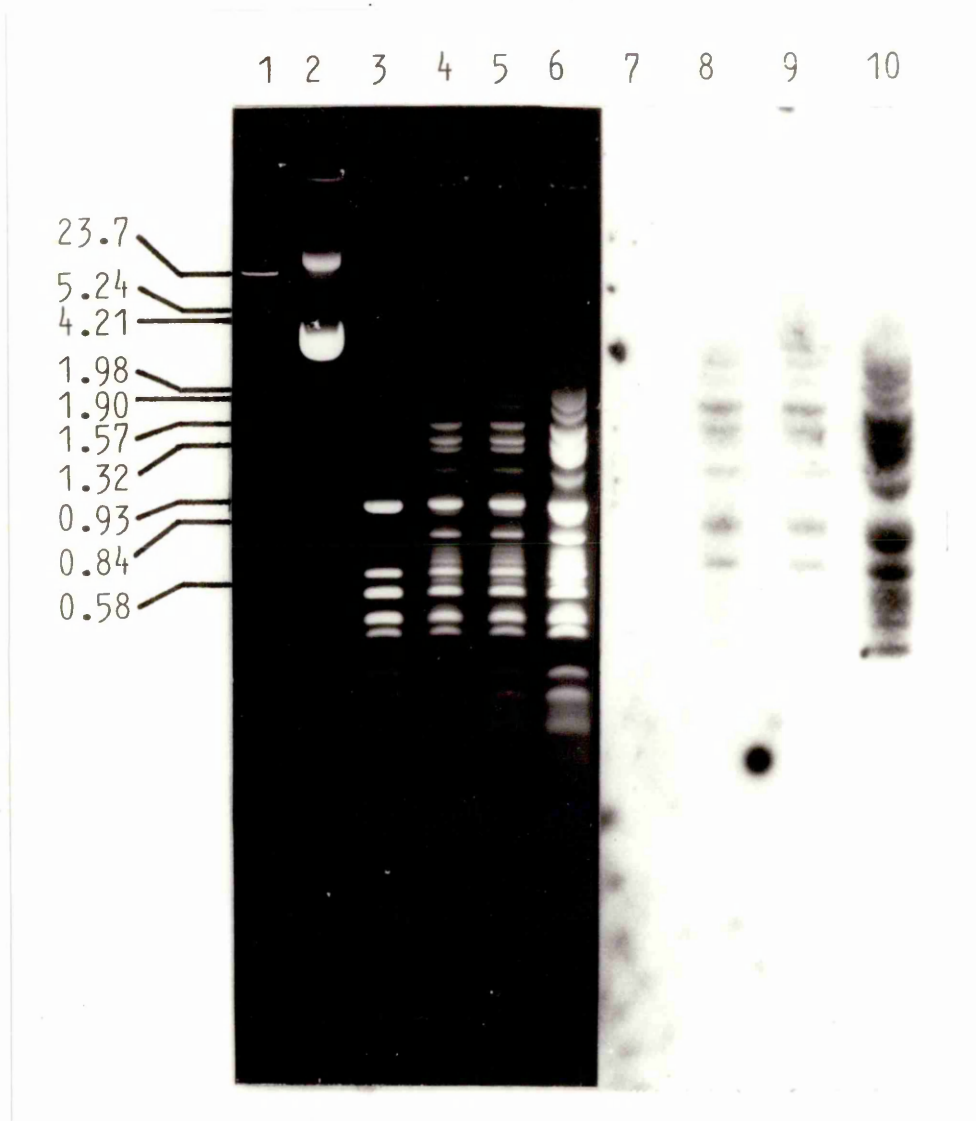
Figure 3.7.   Resistance of Methylated pVHC1 to Cleavage by *Hpa* II

(a)   *Hpa* II Digestion of a Time Course Methylation of pVHC1

1µg of pVHC1 was methylated for 0,2,4,or 6h (in conditions described in section 2.2.2.3.) then digested with *Hpa* II (section 2.2.1.10.), shown in lanes 3-6.   The corresponding fluorograph is shown in lanes 7-10.   Lane 1 contains *Hind* III and *Eco* RI digested λ DNA as markers.   Lane 2 contains untreated pVHC1.

Kbp   1 2 3   4

23.7
5.24/5.05
4.21

1.98
1.90
1.57
1.32
0.93
0.84
0.58
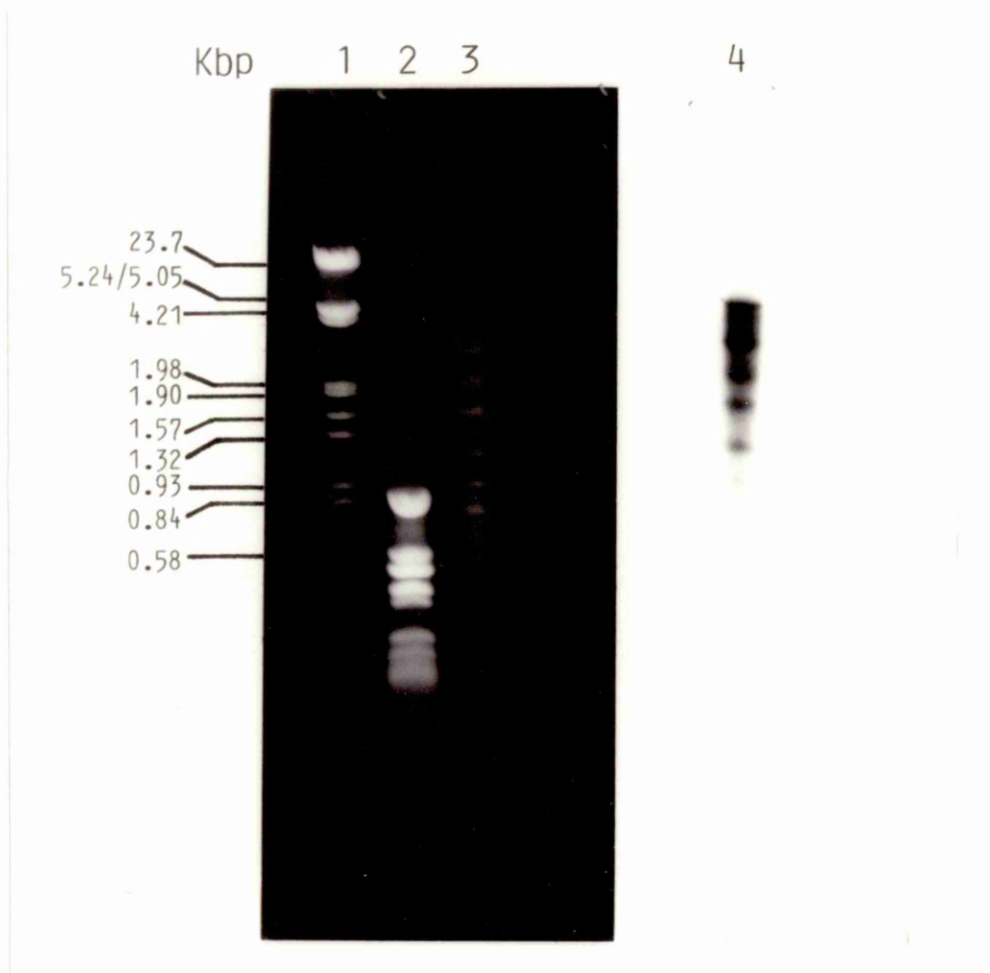
Figure 3.7. Resistance of Methylated pVHC1 to Cleavage by *Hpa* II

(b) *Hpa* II Digestion of Methylated pVHC1

1µg of pVHC1 was methylated to approximately 75% saturation.   0.5µg of this was digested with *Hpa* II and the fragments visualised on a 1% agarose gel (lane 3).   The corresponding fluorograph is represented by lane 4.   Lane 1 contains *Hind* III and *Eco* RI digested λ DNA as markers.   Lane 2 contains *Hpa* II digested mock methylated pVHC1 (1µg).

3.4.2.6. CpG Distribution in pVHCl

Computer analysis of the pVHCl sequence provided a CpG distribution map. From figure 3.8 it is evident that pVHCl unlike SV40 DNA, has an even distribution of CpG's apart from the CpG deficient SV40 terminator fragment. Also, unlike SV40, pVHCl is not CpG deficient, being largely made up of prokaryotic DNA. It contains 430 CpG dinucleotides in its 4736bp, i.e. 215 CpG per strand.
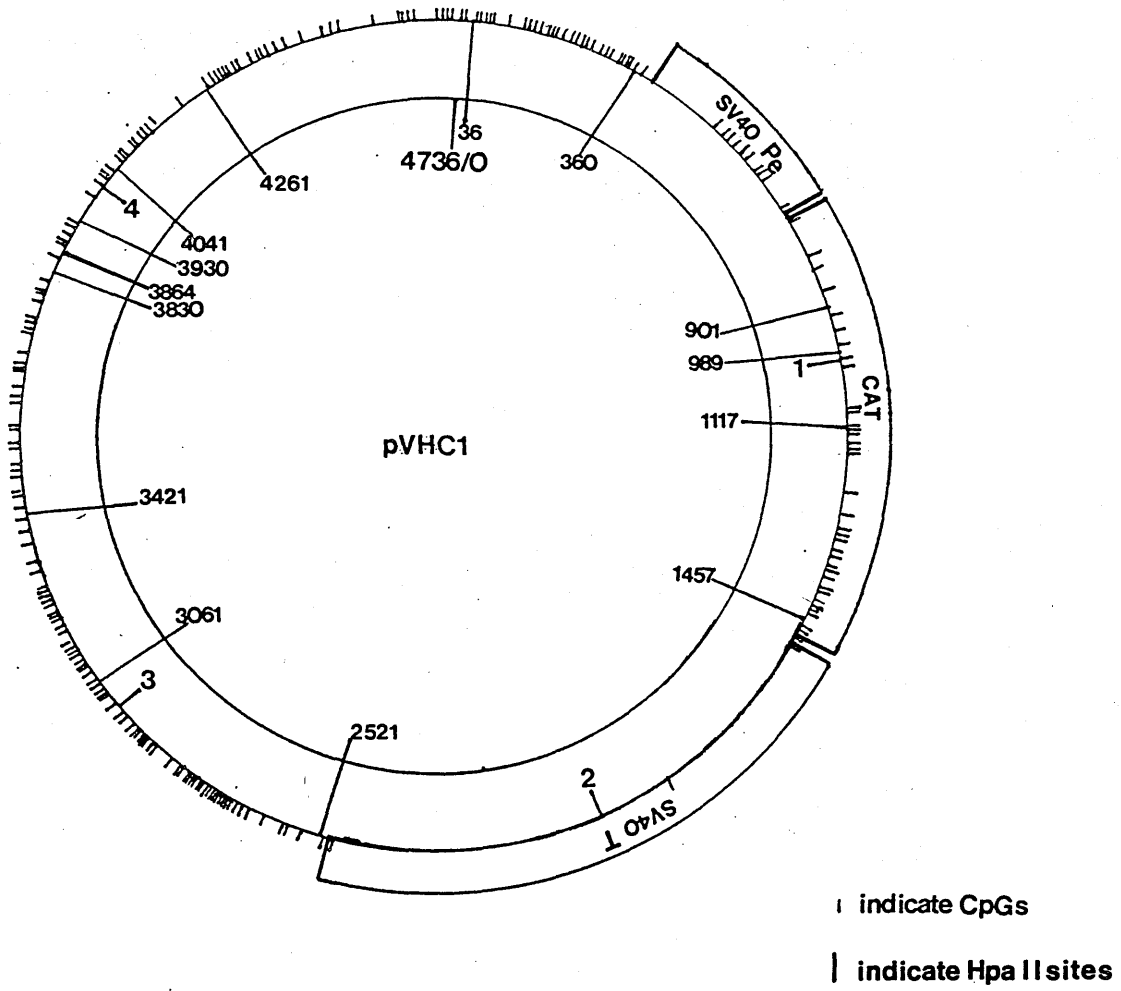
Figure 3.8. CpG Distribution in pVHC1

3.5. Discussion

The aim of the experiments presented in this chapter was to optimise conditions for maximum methylation of native DNA *in vitro*. Double stranded, unmethylated DNA is a poor substrate for mammalian DNA methylase. DNA methylase isolated from fully differentiated cells has a primary *in vivo* function of maintaining established methylation patterns. Although the enzyme is capable of both maintenance and *de novo* methylation *in vitro*, it exhibits a strong preference for hemimethylated DNA. Also, most mammalian enzymes prefer denatured DNA to native, indicating that the enzyme requires single stranded regions of DNA with which to interact. Such sites also become available when duplex DNA "breathes"- a situation that is encouraged at low salt and high temperatures. Adams *et al* .,(1979) reported that the ascites methylase was strongly inhibited by salt, retaining only 50% activity in 50mM NaCl. This was also observed when methylating SV40 DNA. In addition, the enzyme preferred to methylate supercoiled SV40 DNA, a conformation conducive to the formation of single stranded protrusions. There was a ten fold drop in incorporation if the DNA was linearised or nicked with *N. crassa* nuclease before methylation (results not shown).

Several groups have reported difficulty or failure in methylating SV40 DNA *in vitro*. Simon *et al.*, (1978) and Roy and Weissbach (1975) have reported resistance of SV40 DNA to methylation by the rat liver and Hela cell methylases respectively. It is also a relatively poor substrate for the mouse ascites enzyme as shown in section 3.3. Even in conditions of zero salt and long incubation periods, saturation is never achieved. This resistance to methylation is probably due to the uneven distribution of CpG dinucleotides in the molecule. SV40 DNA is CpG deficient and 70% of the 54 dinucleotides

present are clustered in a region of 400bp.   Using synthetic oligonucleotides, Bolden *et al*

.,(1986) found that DNA substrates containing more than one CpG in close proximity were

better substrates for the enzyme than those containing only a single site.   This would

suggest that the enzyme once bound can methylate more than one CpG before dissociating.

The idea of a "walking" mechanism was first introduced by Drahovsky and Morris (1971)

who found that the enzyme and DNA formed a tightly bound complex which was resistant to

salt and where the addition of methyl groups was linear for 30 minutes.   This "walking"

theory may predict that SV40 DNA would be a good substrate for the enzyme, since if the

enzyme bound near the CpG cluster it could methylate a large proportion of the sites before

dissociating.   However, DNA methylase binds to DNA in a fairly non-specific manner and

the point of attachment may depend on the ability of a particular sequence to "melt" or it

could simply be which part of the DNA molecule collides with the enzyme.   Hence, the

point of attachment in most SV40 DNA-enzyme interactions will not be adjacent to the CpG

cluster and so most enzyme molecules will bind, scan and dissociate without adding a CH_3

group.   This is reflected in the results shown in section 3.4. were the presence of one unit

of enzyme for every [six] pmol of CpG results in only 2% incorporation in the initial hour of

the reaction.   Indeed, the addition of a further 10units of enzyme and a longer incubation

time (24h) resulted in a maximum of only 50% incorporation.   We can compare this to

the methylation of pVHC1, section 3.4.   Being largely composed of prokaryotic DNA,

pVHC1 is not CpG deficient, having 460 CpGs in 4736bp and these CpGs are evenly

distributed, with the exception of the 950bp CpG deficient SV40 terminator fragment.

This makes it a better substrate for DNA methylase than SV40 DNA.   1U of enzyme per

pmol of CpG in pVHC1 resulted in 60% methylation after 4 hours (figure 3.5) compared

to 10% for SV40 DNA .   This difference is most likely due to the even distribution of CpGs

in pVHC1 (figure 3.9). Once bound the enzyme in most cases does not have to migrate far

to reach a CpG. So with pVHC1 most enzyme-DNA interactions will be "successful" with

the enzyme adding a CH_3 group at one or more sites before dissociating. However, figure

3.4 shows that optimum methylase concentration is 1U per CpG present. This is

consistent with the enzyme having a turnover of one and the addition of methyl groups

being distributive rather than processive. Simon *et al* .,(1978) found that the rat liver

DNA methylase has a turnover of 7, this is consistent with the estimated turnover of the

ascites enzyme. The requirement of a high enzyme-DNA ratio can be explained if one

considers the possibility that in zero salt conditions the enzyme is not operating in a

monomeric form. The DNA maybe interacting with a high molecular weight enzyme

aggregate, with only the enzyme molecule directly in contact with DNA transferring a CH_3

group. It would appear that once bound these aggregates do not readily dissociate from the

DNA (see section 3.4.2.2.). This would limit the access of "new" enzyme molecules from

binding to the DNA. Also, since the enzyme preparations used were generally only

10-20% pure it is likely that other DNA binding proteins such as transcription factors

and histones are also present. In the methylation reaction these proteins would bind to the

DNA preventing the binding of DNA methylase and hence limiting the efficiency of the

reaction. This phenomenon was also observed by Kautiainen and Jones (1985).

The precise mechanism of action of DNA methylase is still unknown and needs

to be studied in greater detail to understand fully both the *de novo* and maintenance

reactions. However *in vitro* methylation studies are currently hindered by the lack of

procedures for obtaining sufficient amounts of pure and stable preparations of mammalian

DNA methyltransferases.

When methylating any type of DNA it is important to keep the S-adenosyl

methionine concentration at or above the Km, which is around 2μM.   This is due to the fact

that the by-product of the reaction S-adenosylhomocysteine is a potent inhibitor of the

reaction.   It has been shown (Adams and Burdon,1985) that when present in equimolar

concentration with Ado Met this inhibitor reduces methyl transfer by 84% and even a

ratio of Ado Met: S adenosylhomocysteine of 10:1 will reduce methylation by one third.

In summary, the optimised conditions for plasmid methylation are as follows:

DNA concentration of 1μg per 100μl total volume, DNA methylase concentration of 1unit

per pmol of CpG present, S-adenosylmethionine at 20μM in buffer M^+ with a 4 hour

incubation period at $37^\circ C$.   With active and stable preparations of DNA methylase this led

to a 50-70% saturation of the plasmid DNA.   However, the efficiency of methylation was

somewhat reduced when the reaction was scaled up.   Methylation of 5-10μg of pVHC1 led

to only 20-30% saturation.   This problem could not be overcome even when the enzyme

was stabilised by the addition of 1mg/ml BSA or when the final reaction volume was

increased or decreased.

Chapter 4. Effect of Methylation on Transcription from the SV40 Early Promoter

4.1. Introduction

In the following experiments the effect of DNA methylation on transcription from the SV40 early promoter was studied using a transient expression assay. The plasmid pVHCl (described in section 3.4.1) containing the SV40 promoter in the early direction linked to the bacterial chloramphenicol acetyl transferase gene, was methylated using mouse ascites DNA methylase in conditions described in section 3.5. This methylated DNA was then introduced into mouse L929 cells by the calcium phosphate - DNA co-precipitation method (section 2.2.3.1.). The effect of methylation could then be analysed by measuring the cellular levels of chloramphenicol acetyl transferase and comparing them to levels obtained with methylated DNA.

The Chloramphenicol acetyl transferase assay

The chloramphenicol acetyl transferase expression assay was first introduced by Gorman et al., (1982a) and has been used to study a number of viral and cellular promoters (Gorman et al., 1982(a) , Gorman et al., 1982b , Levison et al., 1982). This assay exploits the fact that mammalian cells do not possess the gene for chloramphenicol acetyl transferase, hence any activity found in transfected cells is derived from the exogenous DNA.

There are several ways of introducing foreign DNA into tissue cultured cells, the method used here was the calcium phosphate - DNA co-precipitation method (see section 2.2.3.1 for details). Once the DNA has entered the nucleus of the cells it associates

with histones to form chromatin (Gilmour *et al* .,(1982) and CAT activity has been detected twelve hours after transfection. In the experiments described here CAT activity was normally measured 48 hours after transfection, a time shown by Gorman *et al.*, 1982(a) to be optimum for transient expression assays.

Chloramphenicol acetyl transferase interacts with chloramphenicol by adding acetyl groups to the hydroxyl groups on C1 and C3 of the molecule (see figure 4.1). CAT activity can be measured by quantifying the level of ^{14}C labelled chloramphenicol acetylated in a given time.

Initially the method designed by Cohen *et al.* (1980) was adopted in which the acetylation of chloramphenicol is measured by silica gel thin layer chromatography. However, a new method introduced by Seed and Sheen. (1988) provided a simpler and more sensitive assay for CAT activity. This increased sensitivity is advantageous when studying changes in activity induced by DNA methylation. This assay exploits the relatively low specificity of CAT for the acyl donor, allowing butyryl CoA to be substituted for acetyl CoA. The greater hydrophobicity of butyrylated chloramphenicol allows an effective discrimination between free and acylated chloramphenicol by a simple phase extraction assay. Details of both assays are presented in section 2.2.3.2(a) and (b).
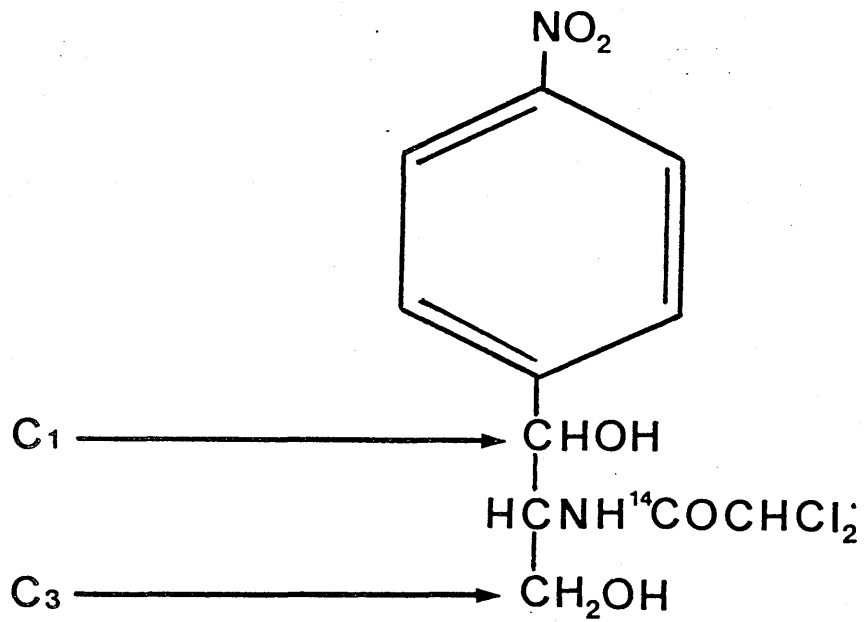
Figure 4.1. Structure of [^{14}C] chloramphenicol and positions of acyl group additions.

4.2. Comparison of Different Promoter Activities in

Expressing the CAT Gene

The ability of the SV40 early promoter in pVHCI to initiate expression of the CAT gene was compared to that of the HSV2 immediate early (IE) promoter. Two plasmids, pLW2 and pLW4 containing the IE promoter linked to the CAT gene were used in this experiment (Fig 4.2(a)). These plasmids were constructed by Lindsay Whitton (Gaffney et al.,1985) at the Institute of Virology , Glasgow . In addition to the immediate early promoter, PLW4 contains an activator sequence which has been shown to increase transcriptions 2-3 fold in transfection assays. Figure 4.2(b) shows that both these plasmids are more active than pVHC1 in initiating CAT activity. These plasmids were used as positive controls in the following transfection experiments. The plasmid p200 (see section 3.4.1) which is identical to pVHCI minus the SV40 promoter shows negligible levels of activity and was used as a negative control. The stimulatory effect of the SV40 enhancer is evident from reduced activity obtained on transfection with the plasmid pSVCI. This plasmid was constructed by inserting the Sph I-Xho I fragment of pVHCI into p200 (see figure 3.3) immediately preceding the CAT gene. Hence in this construct the promoter consisted of the 21bp repeat region and only 22bp of the first 72bp repeat of the enhancer (Fig 4.2(a)). Fig 4.2(b) shows that deletion of the enhancer sequence significantly reduces CAT expression. This observation is in agreement with Gorman et al. (1984), although the effect is not as pronounced as the 10 fold inhibition they observed using a pSV2CAT deletion.
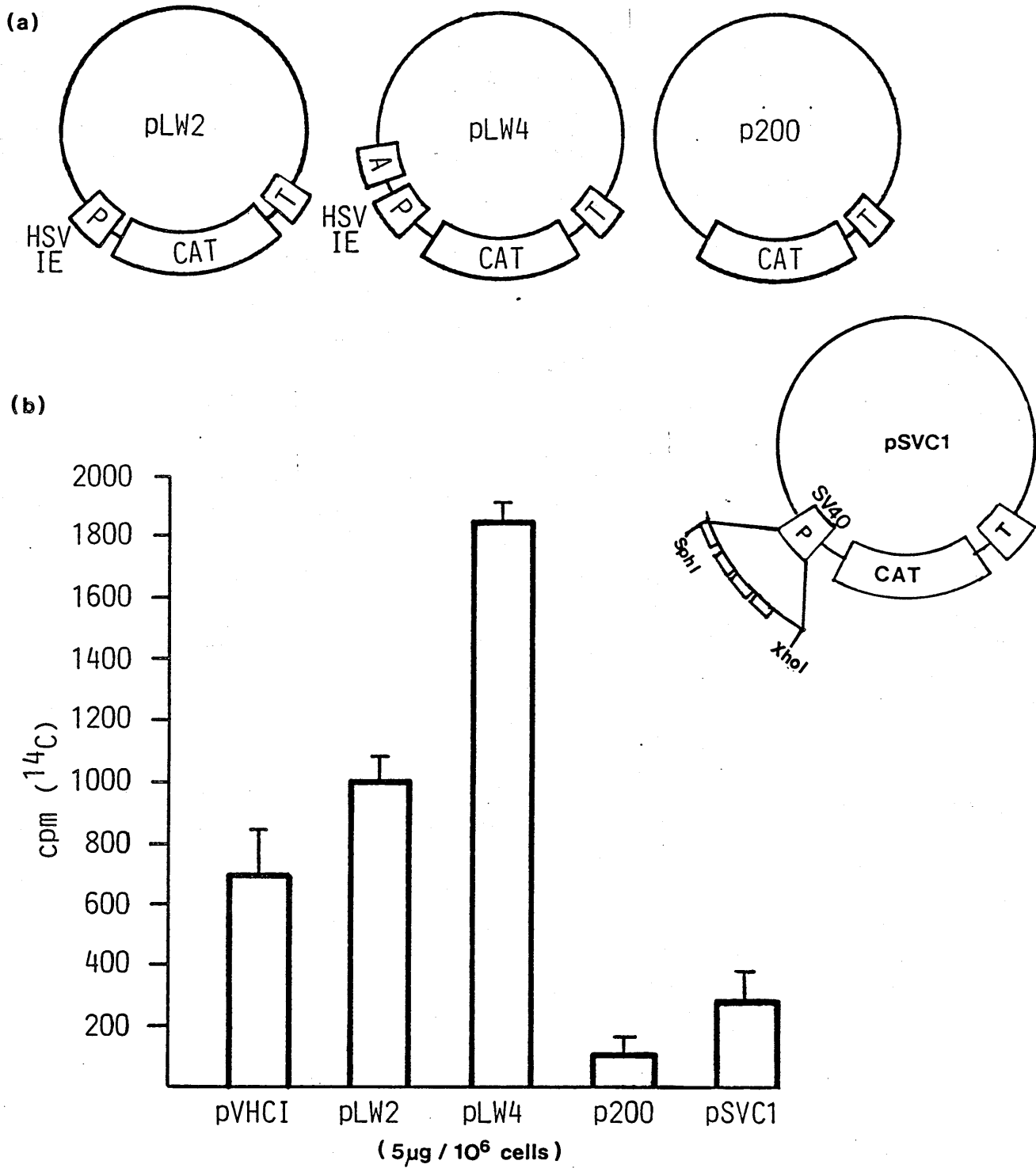
Figure 4.2.   Comparison of Different Promoter Activities in Expressing
                the CAT Gene

(a)   structures of plasmids used

(b)   Mouse L929 cells were transfected with 5μg of plasmid DNA and the resulting CAT
       activity quantified by the method of Cohen *et al.*, (1980)and expressed as cpm of
       acetylated ^{14}C -chloramphenicol.

4.3. DNA Concentration Curve

Tissue cultured mouse L929 cells (approximately 10^6 per 25cm^2 flask) were transfected with 0-5μg of pVHCl in the presence of carrier DNA (total DNA 20μg). Figure 4.3 shows that CAT production increases with increasing DNA over this range of concentrations. Transfecting with higher concentrations of DNA did not lead to a reproducible increase in CAT production using this number of cells (results not shown). Since purified DNA methylase was not available in large quantities it was necessary to limit the amount of DNA used in these transfection experiments. For these reasons, in the transfection experiments that follow 3-5μg of pVHCl was used per 10^6 cells.

Kruczek *et al*., (1983) showed that the amount of CAT expressing plasmid could be reduced to 0.1-1μg by cotransfecting with a promoterless construct (in this case pSVOCAT). This reduced amount still yielded efficient expression of the CAT gene. Transfecting with smaller amounts of DNA would overcome the restrictions of the methylation step. However significant levels of CAT activity were not obtained when 0.5-2μg of pVHCl were cotransfected with 5μg p200. Hence this approach was discontinued.
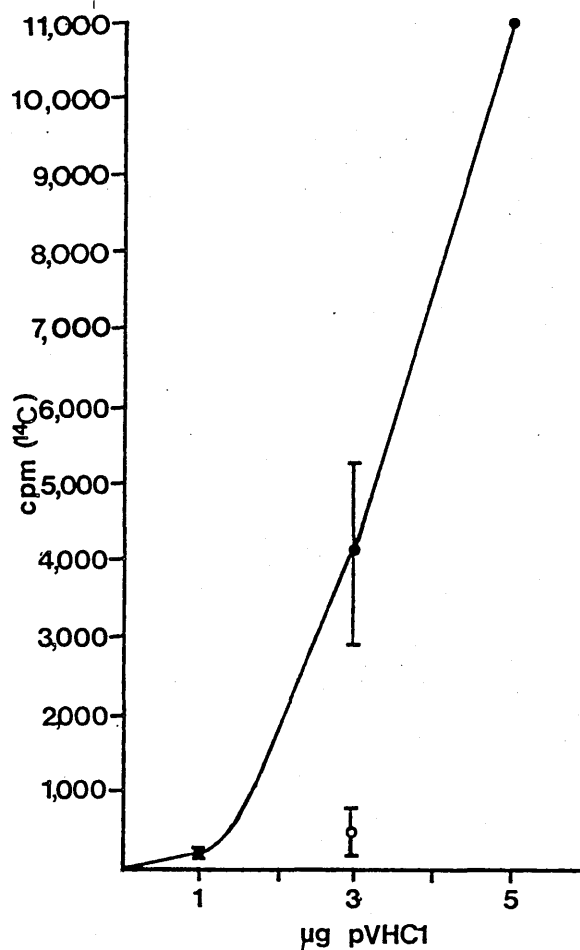
Figure 4.3. Chloramphenicol Acetyltransferase Assay: DNA Concentration

Curve / Effect of Methylation

(a) Mouse L929 cells were transfected with 0,3 or 5μg of pVHCl in conditions described

in section 2.2.3.1. CAT activity was measured 48 hours after transfection by the

method of Cohen *et al.*, (1980) and expressed as cpm ^{14}C labelled acetyl

chloramphenicol.

The open circle shows the activity obtained on transfecting 3μg pVHCl methylated to

8-10% saturation (see section 4.4).
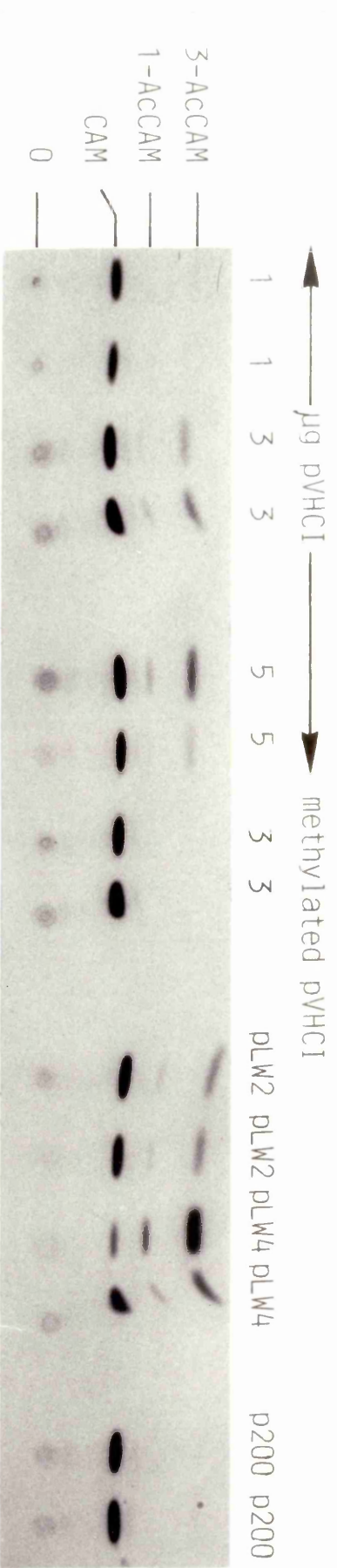
Figure 4.3. Chloramphenicol Acetyltransferase Assay: DNA Concentration Curve / Effect of Methylation

(b) Acetylated and non acetylated ^{14}C- Chloramphenicol were separated by thin layer chromatography and visualised by autoradiography.

4.4.   The Effect of DNA Methylation on CAT Activity

pVHCl DNA was methylated *in vitro* using mouse ascites DNA methylase as described in section 3.5. This resulted in 8-25% of CpG dinucleotides being methylated. The following results show that even these low levels of methylation cause a significant inhibition of CAT expression. In initial experiments CAT expression was almost totally abolished by 8-10% methylation of the DNA.   Figures 4.3(a) and (b) show that transfecting with 3µg of methylated pVHC1 results in only 10% of the activity obtained with unmethylated DNA.   CAT activity was measured by quantifying the level of acetylated chloramphenicol by silica gel thin layer chromatography.   However, in these experiments the control unmethylated DNA was not 'mock' methylated before transfection.   This involves incubating the control DNA with DNA methylase, in the absence of S-adenosyl methionine, in conditions identical to the methylation experiment.   This eliminates the possibility of reduced CAT activity caused by loss of DNA , possibly due to poor recovery or nuclease digestion. Hence, in these experiments it is possible that the effect of methylation may have been exaggerated by one of the above.   However, the methylated pVHC1(1µg) was analysed by agarose gel electrophoresis before transfection (results not shown) indicating that DNA was present at the stated concentration.   Also, pVHC1 activity was always compared to that of PLW4.   The ratio of pLW4 to pVHC1 activity was comparable whether pVHC1 was mock methylated or not.

This almost complete inhibitory effect of DNA methylation was not observed when these experiments were later repeated.   4µg of pVHCl was methylated or 'mock' methylated as described above and introduced into mouse L929 cells.   Figure 4.4 shows

the effect of 15-20% methylation of pVHCl on CAT activity.   In these experiments only a

30-40% drop in activity was observed.

It is evident from the *Hpa* II digestion of methylated DNA (Figure 3.7) that

not all DNA molecules are methylated to the same extent.   There are some   molecules that

remain completely susceptible to *Hpa* II.   Hence one would always expect to see CAT

activity from these DNA molecules and from those in which the transcriptionally

important CpG dinucleotides have not been methylated.

When using DNA mediated gene transfer experiments to study the effects of

methylation one must consider the possibility that methylated DNA behaves differently

from unmethylated DNA when used to transfect cultured cells.   Although there is evidence

that methylated DNA is less efficient at transforming cells compared to unmethylated DNA

(Wigler *et al* ., (1981) , Stein *et al* .,(1982)) , there is no evidence to indicate that

methylated DNA is less efficiently taken up by the cell, or slowly packaged and poorly

integrated into the host chromosome.   In order to try to allow for some inconsistency in

transfection, some experiments were performed involving cotransfection with a marker

plasmid that encodes a separate enzyme activity.   For this a plasmid containing the

β-galactosidase gene linked to the HSV-2 IE4 promoter, pHSVβgal was used.   However,

although CAT activity was regularly present in the extracts little β–galactosidase activity

was found and any comparisons attempted were worthless.
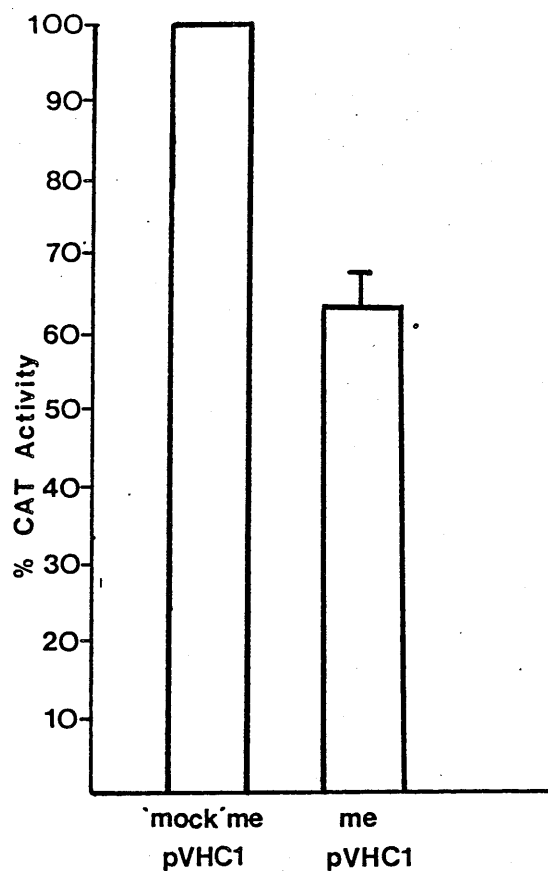
Figure 4.4 **Effect of DNA methylation on CAT expression**

CAT activity was assayed by the method of Seed and Sheen. (1988)(section2.2.3.2(b)) and calculated as pmol acylated chloramphenicol / μg/hour, presented here as a percentage of 'mock' methylated DNA activity. Methylated pVHCl activity is presented as a mean value for two separate experiments +/- standard deviation.

4.5. The Effect of Methylation at Various Times After

Transfection

From these transfection experiments it is clear that DNA methylation has an

inhibitory effect on transcription from the SV40 early promoter. This effect is most

likely due to the presence of CH_3 groups interfering with protein DNA interactions, and it

is possible that this is mediated *via* the formation of altered chromatin structures when

the DNA is introduced into host cells. This was observed in microinjection studies using

the cloned HSV thymidine kinase gene (Buschhausen *et al.* ,1987) where the inhibitory

effect of methylation was only evident after the formation of chromatin in the recipient

cell i.e. 8 hours after injection.

In an attempt to determine whether chromatin formation is involved in SV40

promoter inhibition , mouse L cells were transfected with methylated (10-20%

saturation) or unmethylated pVHCI and CAT activity measured 12-48 hours after

transfection. Figure 4.5 shows the results obtained. Since the level of CAT activity is

initially very low it is difficult to compare activities for the early time points. However

is appears that the inhibitory effect of methylation is at its strongest (approximately

50%) 48 hours after transfection suggesting that chromatin formation may be involved in

inhibition. This experiment would have provided more information if longer time points

had been included , increased inhibition at say 72 hours would have strengthened the
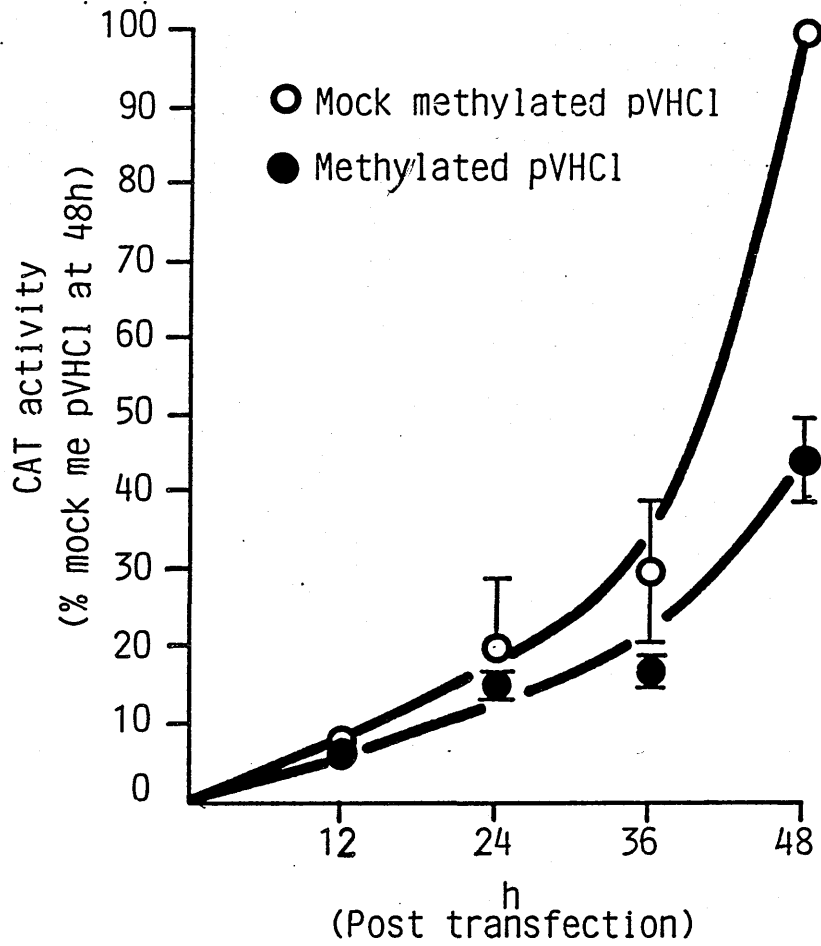
argument.

Figure 4.5.   Analysis of CAT Activity 0-48 Hours After Transfection

Mouse L929 cells were transfected with methylated or 'mock' methylated pVHCl (5μg).
CAT activity was measured 12, 24, 36 and 48 hours after transfection and expressed as a
percentage of control CAT activity after 48 hours, +/- standard deviation.

4.6. Co-transfection of pVHCl With Methylated and Unmethylated Oligonucleotides

The experiments presented in sections 4.4 and 4.5 indicate that methylation has an inhibitory effect on CAT expression in this system. However it remains to be established how this effect is mediated. The SV40 early promoter contains several functional motifs essential for transcription. (see section 1.3.2). The 21bp repeat region has been shown by several groups to be crucial for efficient transcription (Brady et al., 1984; Benoist and Chambon,1985). To establish the importance of Sp1 binding and the effect DNA methylation may have on it mouse L929 cells were co-transfected with pVHCl and oligonucleotides containing Sp1 binding sequences in the methylated or unmethylated state. The oligonucleotides used are shown in figure 4.7(a) and are described in detail in section 5.5. All three oligonucleotides bind Sp1 *in vitro* (section 5.5). Figure 4.7(b) shows that cotransfection with either methylated or unmethylated oligonucleotides reduces CAT expression to approximately 20% of that obtained for pVHCl alone. This suggests that the presence of these oligonucleotides is eliminating the stimulatory effect of the 21bp repeat region and/or the binding of Sp1. It is possible that the pool of Sp1 and other protein factors that bind to this region is being depleted by these oligonucleotides which are in a 30-40 fold molar excess compared with the SV40 promoter sequence. This depleted pool of transacting factors would be less efficient at initiating transcription from the SV40 promoter. This effect occurs irrespective of the methylation state of the binding site, indicating that methylation has no effect on Sp1- DNA interactions.

| 72bp | 72bp | 21pb | 21bp | 22bp | A/T |

0/5243

ATTCTCCGCCCCATGGCTGACT 22mer

ATTCTCC̲G̲CCCCATGGC̲T̲GACT m22mer

CTCCGCCCAGTTCCGCCCCATTCTCCGCCCCATGGCTGACT m40mer
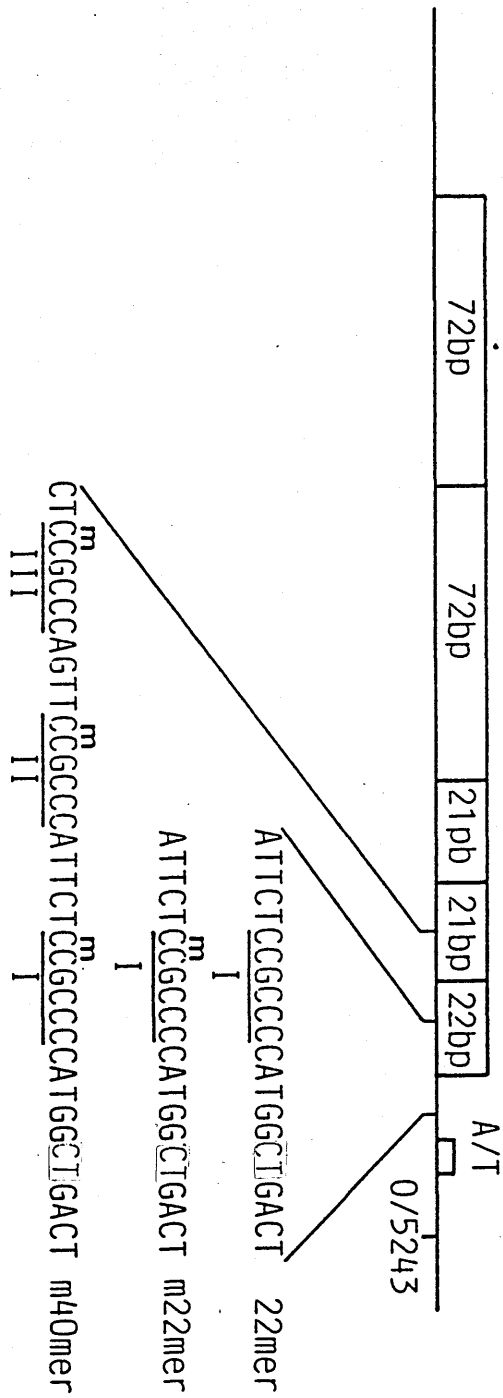
Figure 4.6.   Cotransfection of pVHCI With Methylated and Unmethylated

Oligonucleotides

(a) Diagram of oligonucleotides used.

Oligonucleotides were annealed to their complementary sequences, 22a,m22a and

m40a, before transfection.   The sequence shown corresponds to the early RNA sense strand.
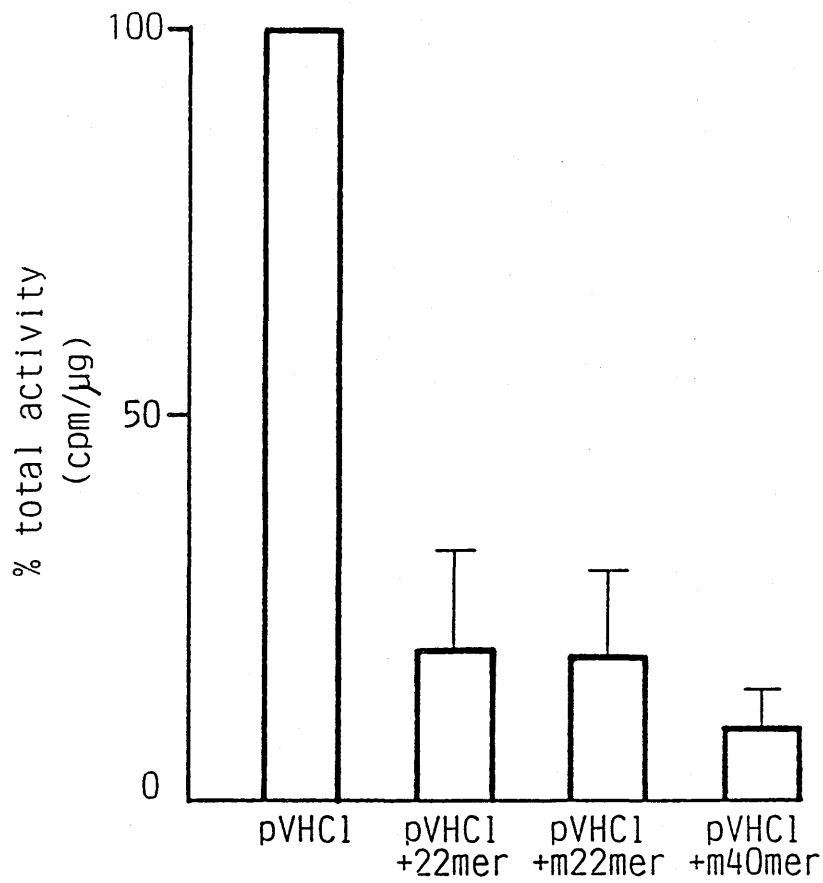
Figure 4.6. Cotransfection of pVHC1 with Methylated and Unmethylated Oligonucleotides

(b) Mouse L929 cells were transfected with 5μg of pVHC1 either alone or with 5μg each of the oligonucleotides. The results show the effect of the oligonucleotides expressed as a percentage of control +/- standard deviation for 7 separate estimations.

Activity was assayed 48h after transfection.

To rule out the possibility that these small double stranded DNA molecules are

in some way inhibiting DNA uptake by the cells or are limiting expression in a

non-specific manner, L929 cells were co-transfected with small DNA fragments obtained

from a *Hae* III restriction digest of the plasmid pAT153. (This produces fragment sizes of

approximately 500-50bp).

Table 4.1 shows that co-transfection with this DNA had no effect on CAT expression.   The

possibility that the inhibiting effect was due to inhibition by a chemical contaminant,

remaining from oligonucleotide synthesis, was eliminated by co-transfecting cells with an

unrelated oligonucleotide which had been prepared in the same manner.   Again no

inhibitory effect was observed .

Table 4.1

DNA TRANSFECTED	CAT ACTIVITY (% control)
pVHCl	100
pVHCl +oligo 22	19 (+/- 13.7)
pVHCl +oligo m22	18.25 (+/- 10.8)
pVHCl +oligo m40	9 (+/- 4.7)
pVHCl +pAT153/*Hae* III	108 (+/- 3.5)
pVHCl +oligo 62	116 (+/-15.5)

Oligo 62 was prepared as described in section 2.2.4.4. and has the sequence:-

CAA GCT TGG CGT AAT CA.

4.7. Discussion

The experiments presented in this chapter indicate that transcription from the SV40 early promoter is methylation sensitive. This conclusion contradicts previous studies involving the methylation of this promoter. However, since pure mammalian methylases are difficult to obtain, all but one of these experiments have involved the use of bacterial methyltransferases. These enzymes are not ideal since they only methylate a small proportion of CpG dinucleotides which may not always occur in important control regions. The SV40 promoter region, although GC-rich is completely deficient in methyl sensitive restriction enzyme sites.

Fradin et al. (1982) found that complete methylation at the single Hpa II site, located near the 5'end of the late coding region in SV40, caused a marked reduction in the synthesis of the viral capsid protein VP-1, but had no effect on the production of early proteins (the large and small tumour antigens). Weachter et al. (1982) used Eco R1 methylase to methylate the SV40 large T antigen gene at adenine residues in 24 sites and found that the expression of this gene remained unchanged. A more relevant experiment was performed by Graessman et al. (1983), who used rat liver DNA methylase to methylate the cytosines of all CpG dinulcleotides in SV40 before microinjection into CV1 cells. They found no change in either early or late expression caused by methylation.

In contrast, the experiments reported here have shown that relatively low levels of methylation caused inhibition of transcription from the SV40 early promoter. These observations can be explained if one considers the mechanism by which methylation is thought to inhibit gene expression. There are two favoured theories of how methylation interferes with gene expression : (a) a direct effect, whereby the presence of a CH_3 group

inhibits the binding of sequence specific DNA binding proteins, such as transcription factors , hence reducing the level of transcription. (b) An indirect effect, where DNA methylation by altering the binding of chromatin forming proteins, such as histones or HMG proteins, causing the formation of an inactive, DNase I insensitive form of chromatin. Evidence suggests that both (a) and (b) exist and the mechanism of inactivation of gene expression by methylation is dependent on the gene involved.

An example of the former was provided by Doerfler and co workers. Kruczek and Deorfler,(1983) showed that methylation of certain CpG dinucleotides in the promoter of the adenovirus type 12 E1a gene causes transcriptional inactivation, whereas methylation of the coding region has no effect. This is also true for the Ad2 E2a gene promoter (Langer et al.,1984). This suggests that inhibition is due to methylation interfering with sequence specific protein DNA interactions in the promoter region. This also appears to be the case for the rat growth hormone gene promoter. A recent study demonstrated that methylation of two Bsu E (CGCG) sites in this promoter inhibited expression of a linked CAT gene by 60% in a transient expression assay. Methylation of plasmid sequences outside the promoter and CAT gene had no effect on expression (Gaido and Strobl,1989). Further support for the direct method of inhibition comes from the recent identification of methylation sensitive transcription factors , for example the adenovirus major late transcription factor (Watt & Molloy, 1988), see section 5.7.

In support of the indirect method of inhibition, direct evidence for chromatin disruption by DNA methylation comes from DNA mediated gene transfer experiments performed by Keshet et al., (1986). They found that although both fully methylated and unmethylated M13 constructs could integrate into the host genome only the latter formed

potentially active DNaseI sensitive , chromatin structures.   This indicates that DNA

methylation directs DNA into an inactive type of structure.   Buschhausen *et al.*, (1987)

in microinjection experiments found that transcriptional inactivation of the HSV

thymidine kinase gene requires chromatin formation in the recipient cells.   Also

reconstitution of minichromosomes *in vitro* showed that methylated chromatin was always

inactive after microinjection whereas 'mock' methylated chromatin was fully active.

The effect of methylation observed in the present experiments involving the

SV40 promoter is most likely due to an indirect inhibition of transcription.   The DNA used

in these experiments was methylated using mouse ascites DNA methylase, resulting in

methylation of 10-25% of CpG dinucleotides.   From the *Hpa* II restriction digest of this

DNA it appears that methyl group incorporation at CpGs is random with some molecules

more resistant to *Hpa* II than others and some remaining totally susceptible to HpaII

cleavage, although these fragments are methylated at other sites (see figure 3.8(a)).   At

this level of methylation , on average only 1-2 CpGs of the 10 present in and around the

promoter will be methylated.   Hence it is unlikely that inhibition is due to interference

with sequence specific protein-DNA interactions in the promoter.   There are no CpG

dinucleotides in the enhancer and 6 clustered in the 21bp repeat region.   Fig. 4.6 shows

that oligonucleotide competition is unaffected by methylation suggesting that Sp1-DNA

interactions are not directly affected.   It is more likely that the presence of methyl groups

throughout the entire molecule causes a disruption of chromatin structure when

introduced into cells, hence reducing the transcriptional efficiency of the promoter.   This

is particularly relevant as pVHCl is rich in CpG's, containing 215 CpG dinucleotide pairs

in 4736b.

A similar situation exists in the herpes simplex thymidine kinase gene, where

the presence of a single methylated CpG dinucleotide in the promoter region was sufficient for transcriptional inactivation of the gene (Ben Hatter and Jiricny ; 1988). This methylation could be at any one of the four CpG's within a region containing the Sp1 and CTF (CAAT-binding transcription factor) binding sites. Since it has been shown that Sp1 binding is not affected by methylation (see chapter 5) it is likely that this effect is due to the alteration of chromatin structures. This is supported by the fact that methylation of the coding sequences of the *tk* gene inhibits expression even when the promoter is unmethylated (Keshet *et al*., 1985). This region specific effect of methylation was also observed by Murray and Grosveld (1987) in studies involving the β-globin gene , where although methylation in the body of the gene did not affect transcription , a minimum length of methylation free DNA in the promoter was required for expression. Also, there appeared to be a correlation between the level of expression and the extent of the methylation free area around the promoter. An area of 1500bp upstream of the gene was required to be unmethylated to achieve a level of expression similar to totally unmethylated DNA.

Hence it is likely that the inhibitory effect of methylation observed in the present transfection studies is due to the presence of methyl groups in the CAT gene and pUC8 sequences as well as the promoter region. It is possible that there is a threshold level of methyl group content above which the rate of transcription is severely reduced. This region-specific inhibitory effect of methylation is presently being investigated in our laboratory using "gap-duplex" molecules (see section 6.3) .

Chapter 5. The Effect of Methylation on the Binding of Sp1 to DNA
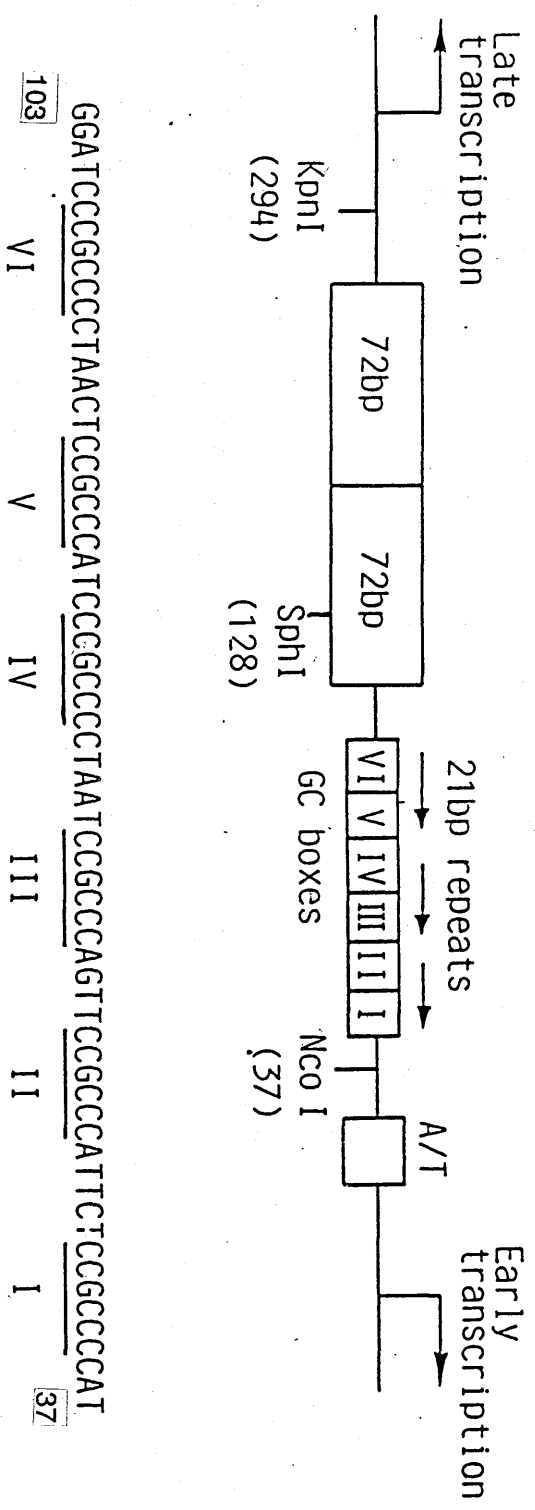
5.1. Introduction

In chapter 4 the effect of DNA methylation from the SV40 early promoter was studied using a transient expression assay. The results suggested that transcription from this promoter is methylation sensitive. It is possible that this effect is mediated by the direct inhibition of transcription factors binding to the promoter region. Since the CH_3 group of 5-methyl cytosine protrudes into the major groove of the DNA helix, the site of protein - DNA interaction, it is possible that the presence of a methyl group could stereochemically inhibit protein binding.

To study this hypothesis the effect of cytosine methylation on the binding of the transacting factor Sp1 to its recognition sites in the 21bp repeat region of the SV40 promoter was investigated. Although Sp1 was first identified as binding to this region of SV40 and stimulating early and late transcription, Sp1 recognition sequences, GC boxes, have also been located upstream of several other viral genes and many cellular genes (Dynan and Tjian, 1983; Dynan et al., 1985; 1986, see section 1.3.2 for details). Although no shared regulatory feature is evident among the diverse promoters in which Sp1 binding sites have been found, the occurrence of these sites in such a wide range of eukaryotic genes makes Sp1 binding a good model system to study.

The 21bp repeat region of SV40 has been shown in deletion experiments (Wildeman et al ,1984) to be crucial for SV40 transcription in vitro. It contains 6 GC box motifs (figure 5.1) which are known to bind Sp1 with varying affinity (Gidoni et al ,1984). These are interesting from a DNA methylation viewpoint since each GC box contains a CpG dinucleotide, i.e. a potential in vivo methylation site and a substrate for purified DNA methylase in vitro.

The effect of methylation on Sp1 binding was studied using a gel retardation assay

(Fried and Crothers, 1981). In this assay an end labelled DNA fragment is incubated with an Sp1-containing Hela cell extract. Free DNA is separated from protein-DNA complexes by differences in their electrophoretic mobility in polyacrylamide gels. Assay conditions are described in section 2.2.4.8.

Late
transcription

KpnI
(294)

72bp

72bp

SphI
(128)

GC boxes

21bp repeats

VI V IV III II I

NcoI
(37)

A/T

Early
transcription

GGATCCGCCCCTAACTCCGCCCATCCGCGCCCTAATCCGCCCAGTTCCGCCCATTCTCCGCCCAT

103   VI   V   IV   III   II   I   37

VI   V   IV   III   II   I

Figure 5.1 The SV40 Promoter Sequence

Numbered according to the SV40 sequence.

5.2. Preparation of Hela Cell Nuclear Extract

Nuclear extract was prepared from Hela S3 cells according to the method of Dignam *et al.*, 1983, as described in section 2.2.4.1. A typical protein concentration for the extract was 2.5-4mg/ml. It has been shown that these extracts can be fractionated by DEAE ion-exchange chromatography (Dynan and Tjian,1983) or heparin Sepharose affinity chromatography (Harrington *et al.*,1988), to give a partially purified Sp1 fraction. However in this study, gel retardation assays using nuclear extract fractionated by either method (see section 2.2.4.2.(a) and (b)) showed patterns of retardation identical to those obtained using unfractionated extract (results not shown).

Sp1 has been purified to homogeneity and studied extensively by Tjian and co-workers (Kadonaga *et al.*,1986; Kadonaga and Tjian, 1986). Since transcription factors in general are low abundance proteins (there are only 5,000-10,000 Sp1 molecules per cell i.e. 0.001% of total cell protein), purification procedures are lengthy and require large amounts of starting material. The one attempt made at attaining pure Sp1 following the fractionation method of Dynan and Tjian (1983) was unsuccessful. The procedure involves several chromatographic steps and the assay for Sp1 relied on its stimulatory activity in an *in vitro* transcription assay. The result was a very dilute product and activity was barely detectable using this system. However a technique introduced by Kadonaga and Tjian (1986), known as recognition site affinity chromatography has since made the purification of low abundance DNA binding proteins, like Sp1, much easier. Unfortunately time did not allow this to be pursued. Therefore all the experiments presented in this chapter were performed using Hela cell nuclear extracts. The advantage of purifying Sp1 for these experiments is that from its already defined

properties (Kadonaga *et al.*,1985) e.g. molecular weight, ability to stimulate early

- transcription, etc., one can be sure that it is Sp1 that is being studied.

However, the advantage of using of using nuclear extracts rather than pure

protein is that it mimics more closely the *in vivo* situation. If one considers the presence

of a methyl cytosine binding protein (Huang *et al.*, 1984, see section 1.4.2), Sp1 binding

may only be blocked by the interference of another protein at its binding site and not

merely by a CH_3 group. The use of nuclear extracts containing both proteins would allow

for the effect of methylation to be observed whereas the use of pure Sp1 in these binding

reactions would not.

5.3. Choice of Restriction Fragment for Use in Gel

Retardation Assays

The region of approximately 350bp constituting the SV40 control region

contains several functional elements. For example, the 72bp repeat enhancer sequence,

the early and late transcription initiation sites, the TATA box like sequence, the origin of

replication and the 21bp repeat sequence. All of these features are known to interact with

sequence specific binding proteins (Lee *et al.*,1987;Fromental *et al.*,1988;Yamaguchi and

DePamphilis,1986; Reinberg *et al.*,1987; Reinberg and Roeder, 1987). The enhancer

sequence itself is known to bind several transcription factors (see chapter 1.3.2 for

details). All of these factors will be present in a HeLa cell nuclear extract, hence to avoid

a very complex picture in the gel retardation assays it is important to use the smallest

possible restriction fragment containing the 21bp repeat region. This is evident in figure

5.2 (b) where restriction fragments containing different amounts of the promoter region

were compared in a gel retardation assay. The SV40 control region is deficient in

restriction enzyme recognition sites limiting the choice of restriction fragments available.

Figure 5.2 (b) shows that the Sph I / Nco I fragment isolated from pSVC1 (see section 4.2

) is most suitable for analysing Sp1 binding by this method since it results in the

formation of discrete retarded bands rather than the large multi-protein complexes

observed for the other fragments. Hence this restriction fragment was chosen for use in

the following gel retardation assays.
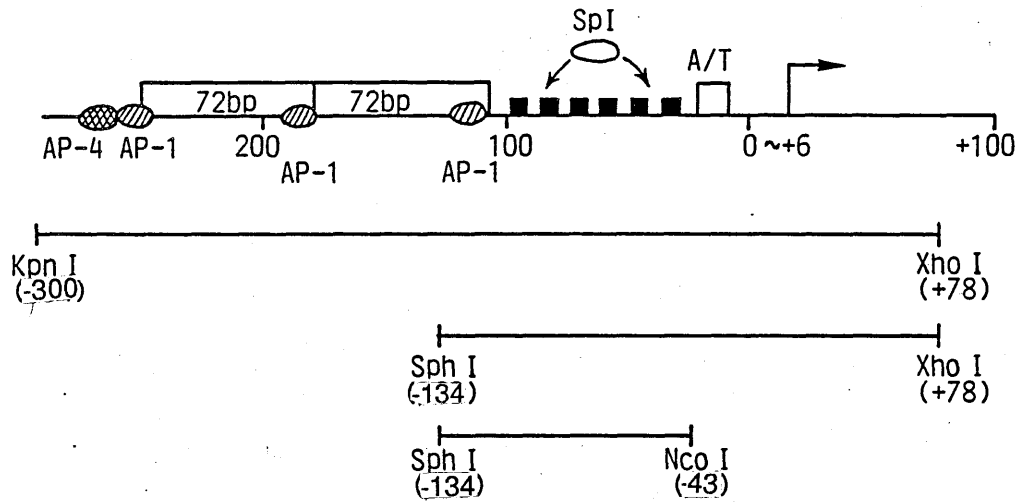
Map numbered according to standard SV40 nomenclature.

Fragments numbered with reference to the start of early transcription.

Figure 5.2 Choice of Restriction Fragment for Use in Gel Retardation

             Assays

(a) SV40 promoter region showing Sp1 binding region and the binding sites of
some enhancer binding proteins.   The restriction fragments considered were: the *Kpn* I
/*Xho* 1 fragment isolated from pVHC1; the *Sph* I/*Xho* I fragment and the *Sph* I/*Nco* I
fragment isolated from pSVC1 (section 4.2).

(b) Comparison of these fragments in a gel retardation assay; 5ng of labelled
DNA (section 2.2.4.3.) was incubated with 0,1,5 or 10µg of nuclear extract in conditions
described in section 2.2.4.7.   The products were analysed on a 4% polyacrylamide gel and
visualised by autoradiography.

139a
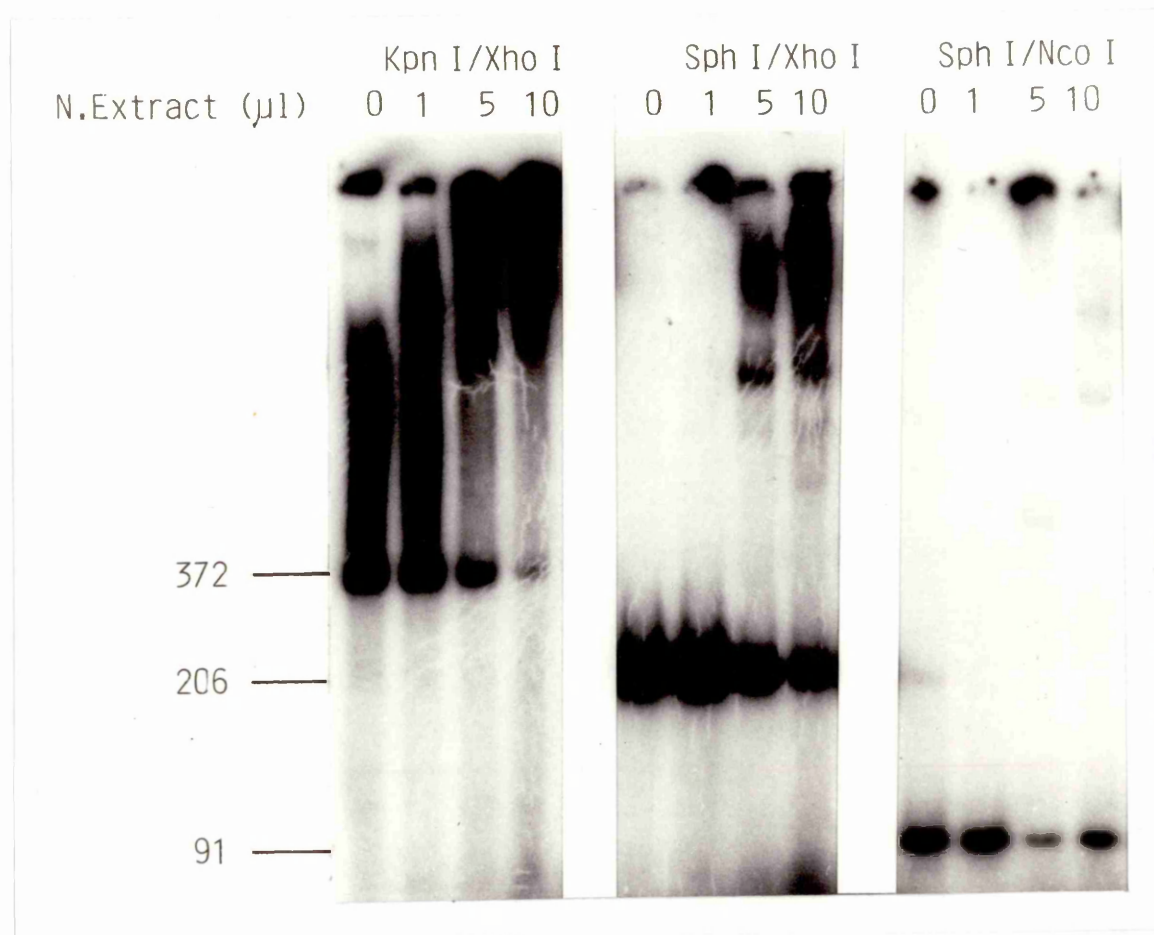
5.4. The Effect of Increasing Protein Concentration

When nuclear extract is incubated with a DNA fragment containing the 21bp repeat region three protein-DNA complexes with differing electrophoretic mobilities are observed. Figure 5.3 shows that the intensity of all three retarded bands increases with increasing protein concentration. The *Sph* I / *Nco* I fragment of pSVC1 contains only the 21bp repeat region and approximately 20bp of the enhancer sequence and all three retarded bands are sensitive to competition with GC box containing oligonucleotides (see section 5.5.2.), indicating that the protein(s) involved are binding to the GC box motifs.

These retarded bands could represent different numbers of Sp1 molecules binding to the DNA. Fried and Crothers (1981), found that when titrating the lac repressor in gel retardation assays, the mobility of the DNA decreased as a function of the number of protein molecules bound. This region of SV40 contains six potential Sp1 binding sites each with differing affinity for the protein. Dimethyl sulphate footprinting and experiments involving the mutagenesis of the GC box motifs indicated the presence of up to 5 independently bound Sp1 molecules in this region (Gidoni *et al* ., (1984). However it is thought that Sp1 binds to DNA as a monomer, the molecular weight of which is 95-105KDa. It is unclear whether the stereochemistry of the binding reaction will allow 5 molecules to bind simultaneously to a region of only 70bp i.e. 7 turns of the DNA helix. It is more likely that under the non-denaturing conditions of the gel retardation assay only one Sp1 molecule will bind. Once bound the DNA may wrap round the protein preventing others binding. Westin *et al.*, (1988) observed only one major retarded Sp1-DNA complex when using an oligonucleotide containing two strong Sp1 binding sites in a gel retardation assay. (Probing the gel with anti-Sp1 antiserum would indicate which of the retarded bands represent Sp1-DNA complexes).

Alternatively these retarded bands may represent different sequence specific proteins binding to the same DNA. Kim *et al*.,(1987) identified a second transacting factor that interacts with the 21bp repeat region of SV40. The late SV40 transcription factor (LSF) is a cellular factor that stimulates late transcription. It is known to interact with the 21bp repeat region but its precise binding site has not yet been identified. When partially purified LSF was used in gel retardation assays, two major protein - DNA complexes were formed both of which migrated more rapidly than the Sp1-DNA complex. When both Sp1 and LSF (distinguished by their ability to stimulate early and late transcription respectively) were present a pattern of retardation similar to that observed in figure 5.3 was observed. It is more probable then that in these experiments only the slowest migrating complexes are due to Sp1-DNA interactions.
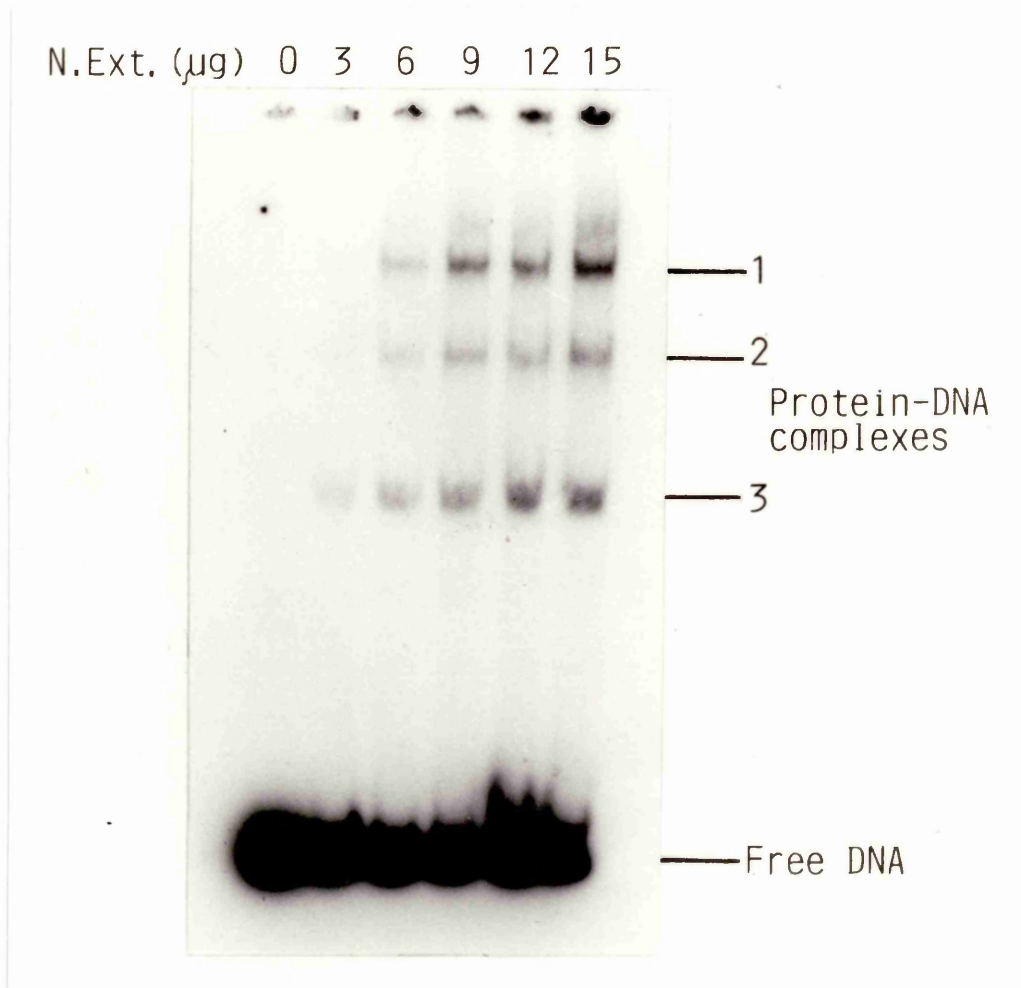
Figure 5.3. The Effect of Increasing Protein Concentration

5ng of ^{32}P labelled *Sph* I/*Nco* I fragment was incubated with 0, 3,6,9,12 or 15µg of nuclear extract in conditions described in section 2.2.4.8. Protein-DNA complexes were separated from free DNA on a 4% polyacrylamide gel and visualised by autoradiography (overnight exposure).

5.5. Competition with Methylated and Unmethylated

Oligonucleotides

The original strategy for studying the effect of methylation on Sp1 binding was to isolate a restriction fragment containing the 21bp repeat region, methylate it using mouse ascites DNA methylase and study the effect on Sp1 binding.

However this approach was not feasible for two reasons: (a) DNA methylase does not methylate small DNA fragments with high efficiency and recovery of these small methylated fragments is poor. (b) Since the SV40 promoter region does not possess any methylation sensitive restriction enzyme sites, it is difficult to assess the extent of methylation and the position of the methyl groups added , the only possibility being Maxam and Gilbert sequencing (1980) which requires µg amounts of methylated fragment. For these reasons the approach was changed to using methylated oligodeoxynucleotides. Synthetic oligonucleotides can be produced with a 5-methylcytosine replacing a cytosine residue at the desired site. Hence the methyl groups added can be limited to the sites of interest and these DNA molecules are always 100% methylated at these sites.

Figure 5.4 shows the three oligonucleotide used. Oligonucleotide 22 contains GC box I and 3' flanking sequence. Oligonucleotide m22 is identical but with a 5-methyl cytosine residue replacing the cytosine immediately preceding the guanine in the GC box. Oligonucleotide m40 contains GC boxes I, II and III all in the methylated state. The complementary oligonucleotides of all three were also synthesised.

GC box I was chosen because initially these oligonucleotides were used in a separate set of experiments (see section 6.3). These experiments involved the use of these methylated and unmethylated oligonucleotides as primers for complementary strand synthesis by DNA polymerase 1 Klenow fragment. The 3' GC box 1 flanking sequence was included to ensure

that binding of the oligonucleotide to the template DNA was limited to GC box 1 only.

Ideally a strong Sp1 binding site such as GC box III would be used in these gel retardation

experiments and figure 5.6. shows that oligonucleotide m40 containing GC boxes I, II and III

has greater affinity for Sp1. (However, in the absence of a strong binding site, Sp1 will

bind to GC box I).
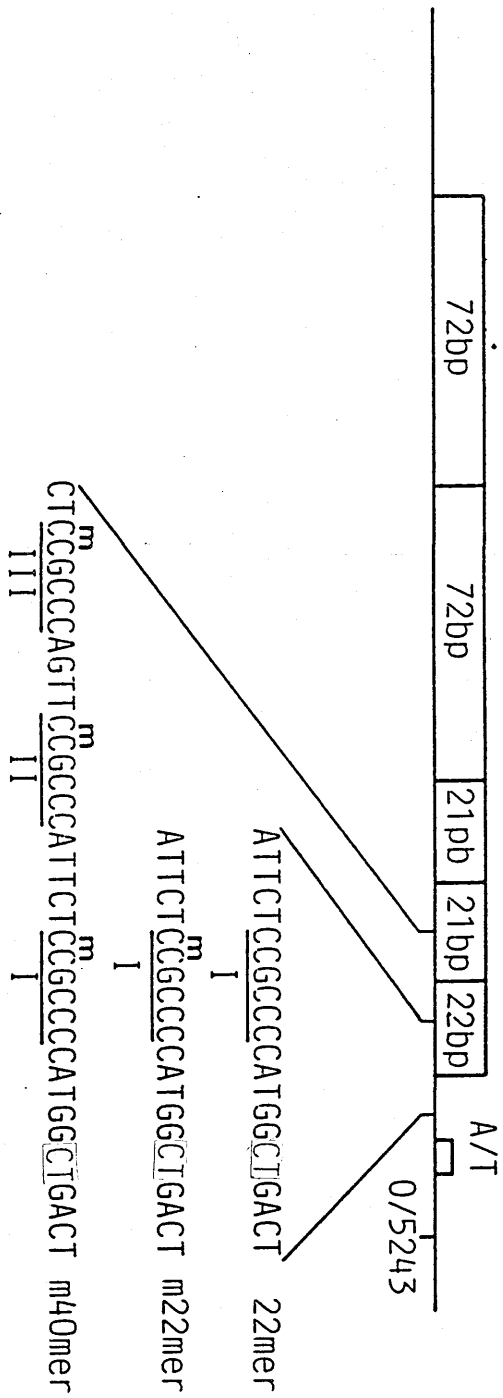
Figure 5.4 Oligonucleotides 22, m22 and m40

(shown here is the C-rich (late coding) strand)

5.5.1. Annealing and Ligation of Oligonucleotides

Since GC boxes exist as multiple copies in SV40 and in other Sp1 sensitive promoters it is reasonable to assume that oligonucleotides containing more than one GC box would compete more efficiently for Sp1 binding than those containing only a single GC box. Hence, oligonucleotides were annealed, phosphorylated and ligated in conditions described in section 2.2.4.5. Figure 5.5.(a) shows 100% annealing of oligonucleotide 22 to its complementary strand. Figure 5.5(b) shows that double stranded oligonucleotides 22 and m22 were successfully ligated forming repeats of 1-25 units. Although, oligonucleotide m40 annealed to its complementary sequence, ligation was very poor. This may have been due to a contaminant such as phenol, inhibiting the ligation reaction. However, extraction with ether and reprecipitation of the oligonucleotide DNA with ethanol did not improve the ligation efficiency.
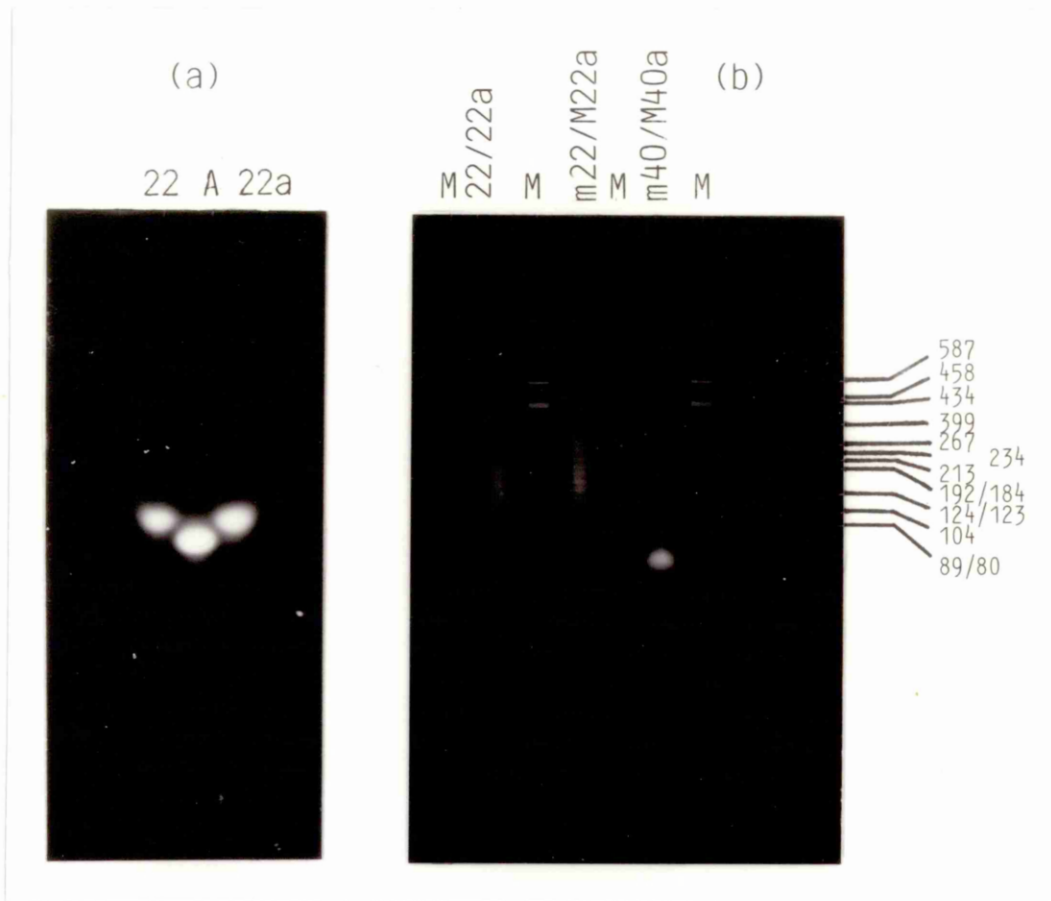
Figure 5.5. Annealing and Ligation of Oligonucleotides

(a) 2μg of annealed oligonucleotide 22/22a (lane 2) were analysed on a 2% agarose gel with 2μg of oligonucleotide 22 and 22a (lanes 1 and 3)

(b) 1μg of each annealed /ligated oligonucleotide was analysed on a 2% agarose gel (lanes 2, 4 and 6) with *Hae* III digested pAT153 (Twigg and Sherratt, 1980) DNA as marker DNA (M).

The gels were stained with ethidium bromide and the DNA visualised with U.V. irradiation.

5.5.2.The Effect of Oligonucleotide Competition in a Gel Retardation Assay

The affinity of specifically bound protein for methylated and unmethylated GC boxes was compared by studying the ability of methylated and unmethylated oligonucleotides to compete with the labelled *Sph* I/ *Nco* I fragment from pSVC1 (see figure 5.2(a)) for protein binding. Annealed or ligated oligonucleotides were used in a competition gel retardation assay. Nuclear extract (5µg) was preincubated with of each of the oligonucleotides (100ng) at 25°C for 10 minutes. 5ng of labelled DNA fragment was then added and the incubation continued for a further 20 minutes before electrophoresis on a 4% polyacrylamide gel. The effect of methylation can be determined by comparing the remaining protein pool (by its ability to bind to labelled DNA fragment) after oligonucleotide binding.

Unfortunately competition with annealed 22mer was unsuccessful, but otherwise figure 5.6 (a) and (b) show that methylated and unmethylated oligonucleotides compete equally well for specific protein binding, indicating that cytosine methylation has no effect on these protein - DNA interactions. It is important to explain the origin of the retarded bands seen in these figures . It is clear from lane 1 that the DNA is heterogeneous and the slower moving band could represent some single stranded DNA which may have arisen as a result of excessive drying of the DNA following end labelling. Retardation of this material on incubation with nuclear extract probably gives rise to the diffuse bands as a result of nonspecific protein DNA interactions. The fact that bands 1 and 2 are competed out with cold 22mer suggests that the protein(s) involved specifically bind the DNA sequence contained in these molecules. There is one proviso, however, since following the practice of other workers (Landolfi *et al.*(1986) and Holler *et al.*(1988)), the only other competitor DNA was a large excess of poly dl-dC which may not adopt the required conformation for optimal protein binding.
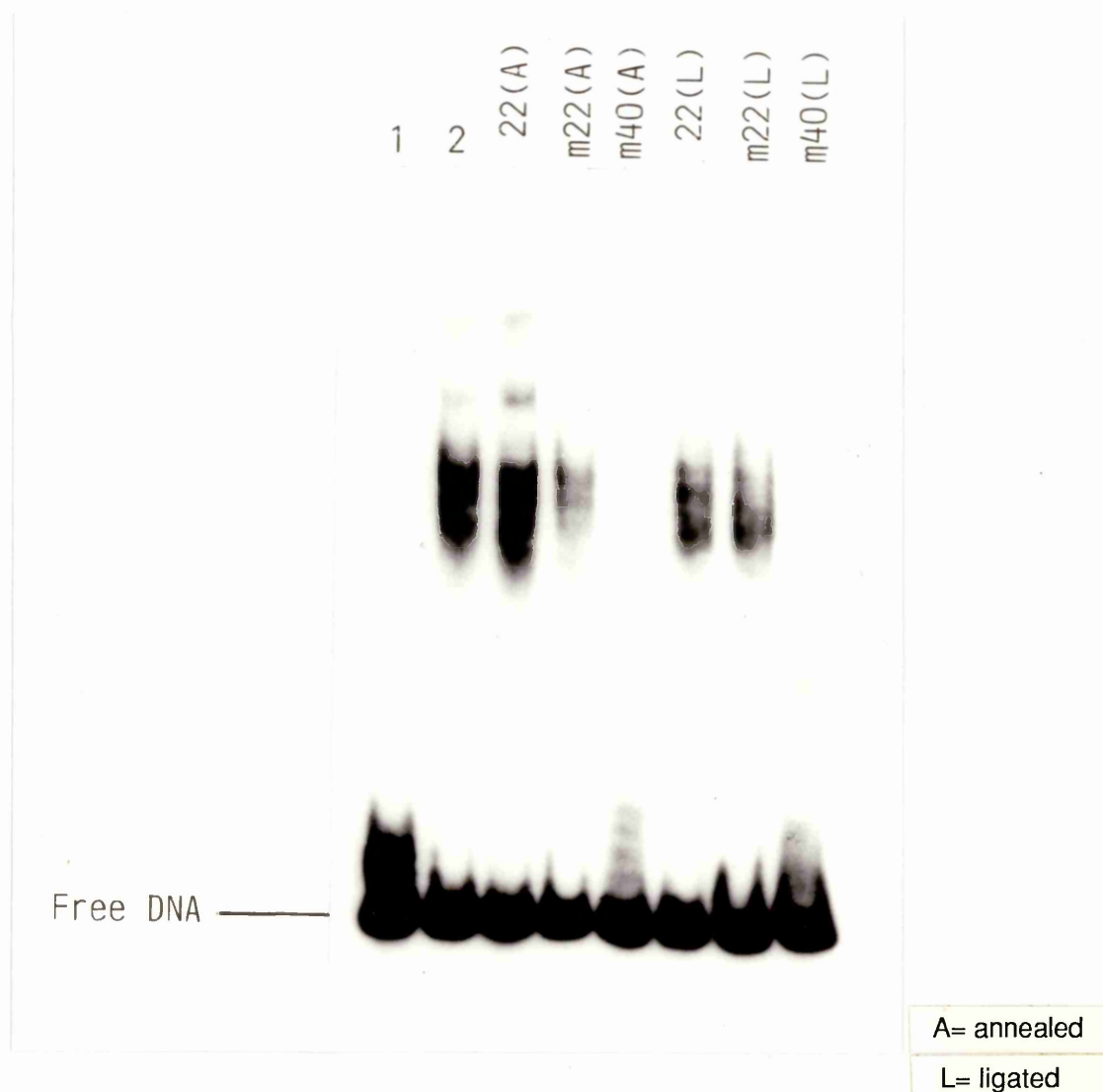
Figure 5.6.   Oligonucleotide Competition

(a) 5ng of ^{32}P labelled *Sph* I/*Nco* I fragment of pSVC1cwas incubated with 5µg of nuclear

extract (with the exception of lane 1) in the presence, lanes 3-8, or absence,lane2, of

competing oligonucleotide in conditions described in section 2.2.4.8..   The products were

separated on a 4% polyacrylamide gel and visualised by autoradiography.
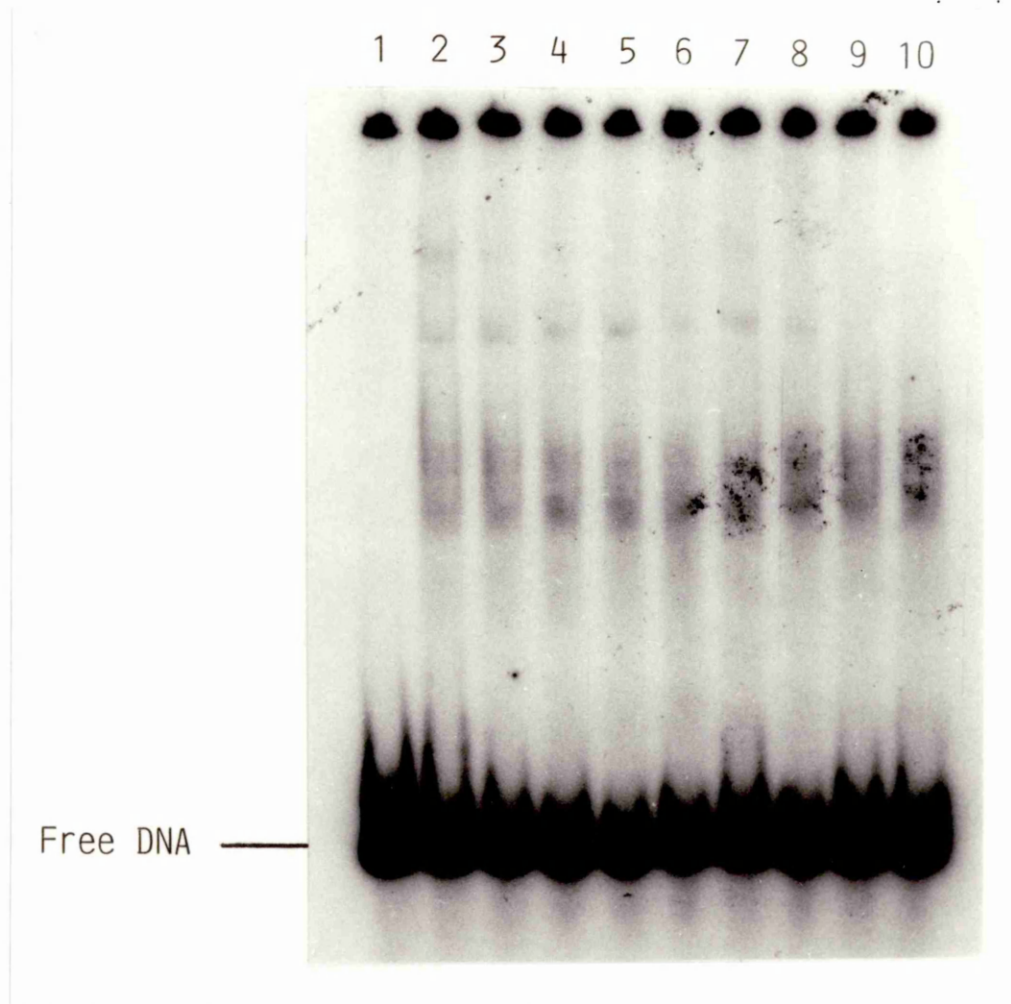
Figure 5.6. Oligonucleotide Competition

(b)   5ng of labelled *Sph* I/*Nco* I fragment of pSVC1 was incubated with 5μg of nuclear extract, in the presence of 0.025,0.1,0.3,and 1.0μg of ligated oligonucleotide 22/22a (lanes 3-6) or oligonucleotide m22/m22a (lanes 6-10).   The products were separated on a 4% polyacrylamide gel and visualised by autoradiography.   Lane 1 contains labelled DNA in identical conditions minus nuclear extract and oligonucleotides. Lane 2 contains nuclear extract but no competing oligonucleotides.

5.6. The Use of labelled Oligonucleotides in a Gel

Retardation Assay

Oligonucleotides 22a and 22b or m22a and m22b were annealed as described in section 2.2.4.5. and the double stranded oligonucleotides gel purified after end labelling with [γ-^{32}P]-ATP and T4 polynucleotide kinase as described in section 2.2.4.6. Protein binding was analysed in a gel retardation assay with increasing amounts of nuclear extract.

Figure 5.7 shows that both unmethylated (lanes 1-5) and methylated (lanes 6-10) oligonucleotides bind protein with similar affinity. This suggests that cytosine methylation has no inhibitory effect on protein binding.

In order to determine whether the retarded bands represent the binding of sequence specific binding proteins to this region various types of DNA were used as competitor in a gel retardation assay. Figure 5.8 shows as expected the cold 22mer methylated or unmethylated is the most effective competitor in the reaction. p200 (described in section 3.4.1.) does not contain a GC box motif and, in the presence of poly dI-dC, it can compete for the protein only when in 50 fold excess. This shows that the retarded band contains a protein with a higher affinity for the SV40 promoter region than for other non-specific DNA. Lanes 14,15 and 16 represent the use of single stranded 22mer (oligo 22a, the late coding C-rich strand) at different concentrations as competitor. This DNA is a very poor competitor indicating that the retarded material is not due to a single stranded DNA binding protein.
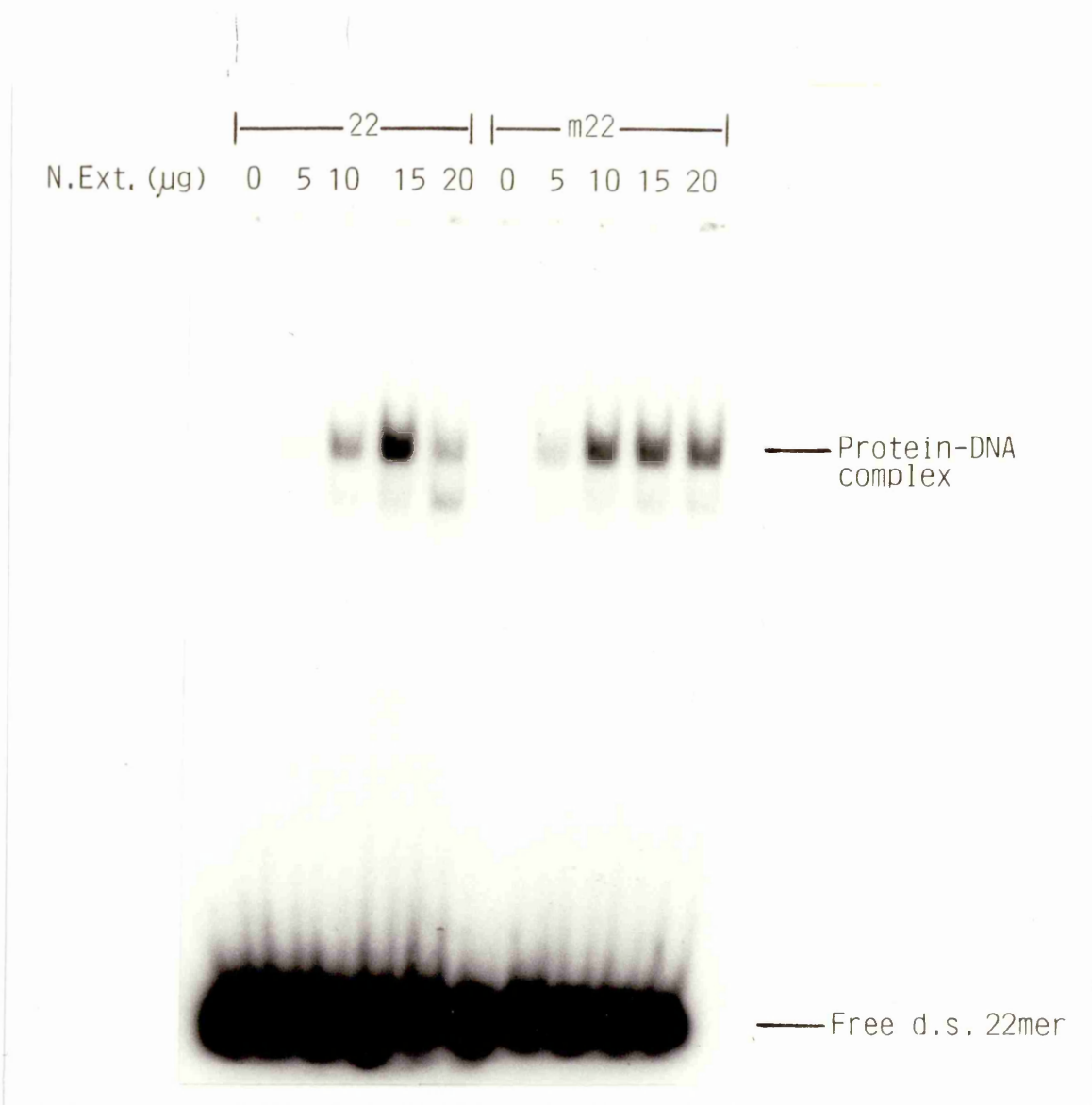
Figure 5.7.   Binding of Nuclear Protein to Oligonucleotides

2ng of labelled double stranded oligonucleotide 22 or m22 were incubated with 0-20 µg
of nuclear extract in conditions described in section 2.2.4.8.   The products were
separated on a 7% polyacrylamide gel. The gel was dried and autoradiographed (section
2.2.1.16).
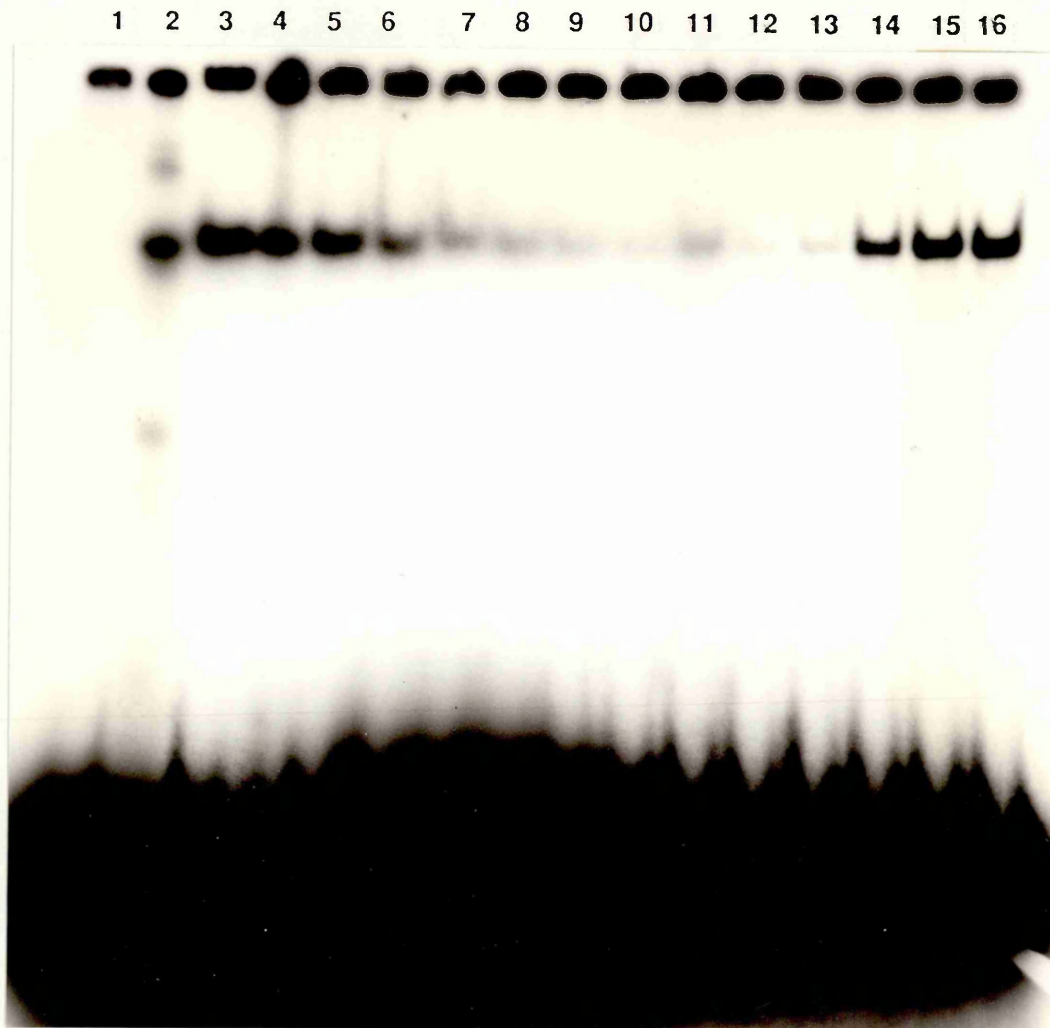
Figure 5.8. Comparison of the Competitive Effect of Various DNAs in a Gel Retardation Assay

Gel retardation assays were performed using 20,000 cpm(~8ng) of end labelled 22mer, 10μg of Hela cell nuclear extract and 1μg of poly dI-dC in conditions described in section 2.2.4.8., in the presence or absence of competitor as indicated below. After a 30 minute incubation at 25°C 10μl of sample buffer containing 30% glycerol, 0.1% bromophenol blue in 0.5x TBE was added and the samples electrophoresed on a 7% polyacrylamide gel. Lane 1, no extract; lane 2, no TBE in added sample buffer; lane 3, standard assay; lane 4, poly dI-dC replaced by 1μg p200 with no TBE in sample buffer. Lanes 5-13 also contain double stranded competitor DNA as follows: lanes 5-7, p200 at 100,300 and 500ng; lanes 8-10, oligo 22 at 25,50 and 100ng; lanes 11-13, oligo m22 at 25,50 and 100ng Lanes 14-16, the single stranded oligo 22a (the C-rich strand of oligo22) at 25,50 and 100ng.

There are three proteins which have been reported to bind to sequences at least partly contained in the 22mer: Sp1(Dynan and Tjian, 1983), LSF(Kim *et al.*, 1987) and AP1(Mermod *et al.*, 1988).   In addition proteins such as NFI and NFIII bind to promoter regions of other genes and such proteins may show an intermediate affinity for the SV40 promoter region even though there is no obvious sequence homology between their specific binding sequences.   Oligonucleotides containing binding sites for NFI(de Vries *et al.*,1987) and NFIII (Pruijn *et al.*,1986)may therefore act as competitors intermediate between the GC box I of the 22mer and model or plasmid DNA.   Furthermore an oligonucleotide containing GCbox III should be a better competitor than the 22mer containing GC box I.

In order to further test the binding specificity, oligonucleotides containing the binding sites of NFI and NFIII were used as competitors in a gel retardation assay.   Figures 5.9 (a) and (b) show that both these oligonucleotides compete with the labelled 22mers for protein binding.   The competitive effect of these oligonucleotides was compared to that of the cold 22mers by excising the retarded bands from the gel and estimating the amount of radioactivity present by Cherenkov counting.   In table 5.1. this is expressed as percentage competition for four separate estimations plus or minus the standard deviations.   The results are presented graphically in figure 5.10.   From these figures it appears that both the NFI and NFIII binding sequences compete for protein binding, although less effectively than the 22mer.   Oligonucleotide m40 which contains GC box III, a strong Sp1 binding site appears to be the best competitor.
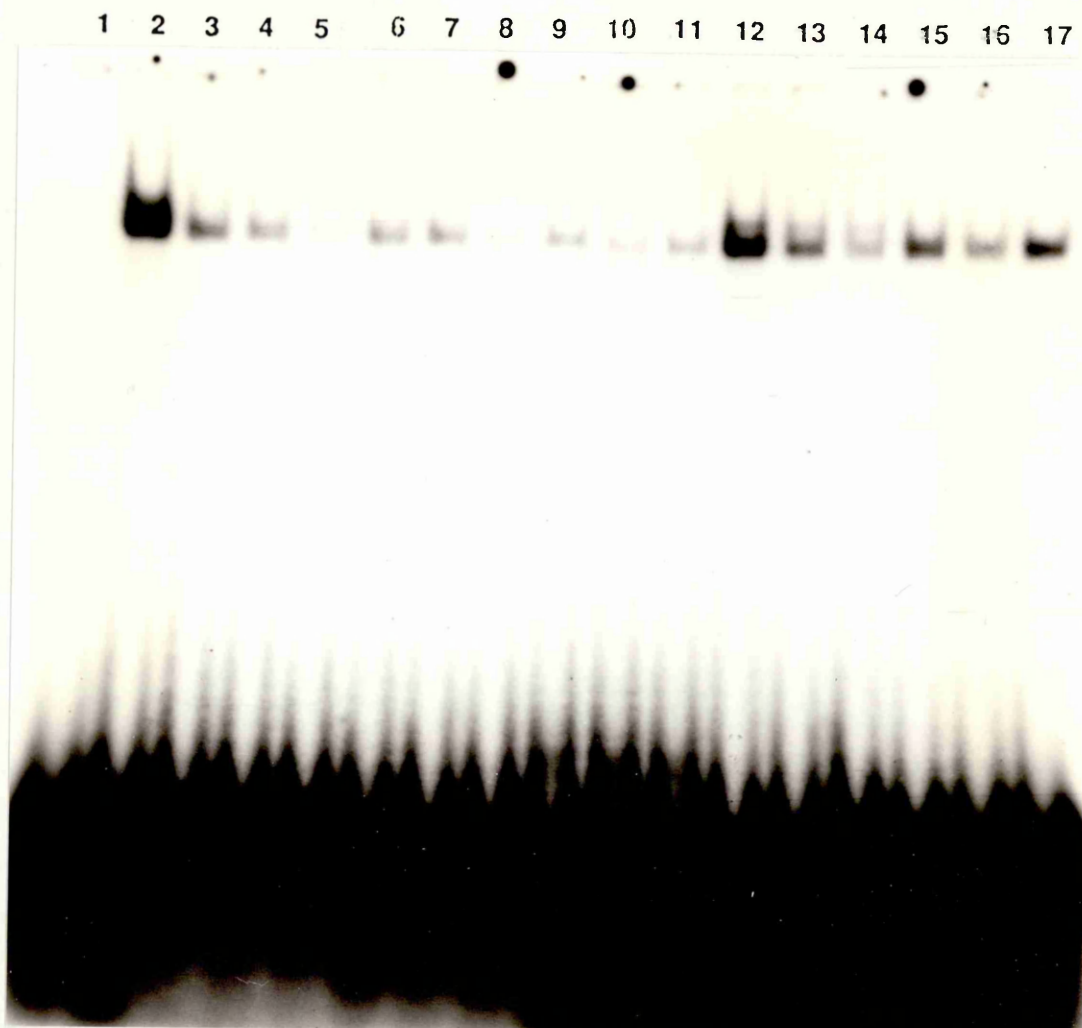
Figure 5.9 The Competitive Effect of the NFI and NFIII Binding Sites

(a) Gel retardation assays were performed using 20,000 cpm (~8ng) of end labelled, double stranded oligo22,10μg of Hela cell nuclear extract and 1μg of poly dl-dC in conditions described in section 2.2.4.8. The reaction was stopped by the addition of a buffer containing 30% glycerol, 0.1% bromophenol blue in 0.5xTBE and electrophoresis on a 7% polyacrylamide gel. Lane 1 contains labelled DNA minus nuclear extract , lane 2 contains labelled DNA plus nuclear extract and lanes 3-17 also contain unlabelled double stranded competitor DNA mixed with the labelled oligo prior to addition of the extract:- lanes 3,4,and 5 contain oligo 22 at 50,100 and 200ng respectively, lanes 6,7 and 8 contain olig n22 at 50,100 and 200ng; lanes 9,10 and 11 contain oligo m40 at 50,100 and 200ng, lanes12,13 and14 contain oligo NFI and lanes 15,16 and17 contain oligo NFIII also at 50,100 and 200ng. (Oligonucleotides NFI and NFIII were provided by Dr. R. Hay , University of St. Andrews)
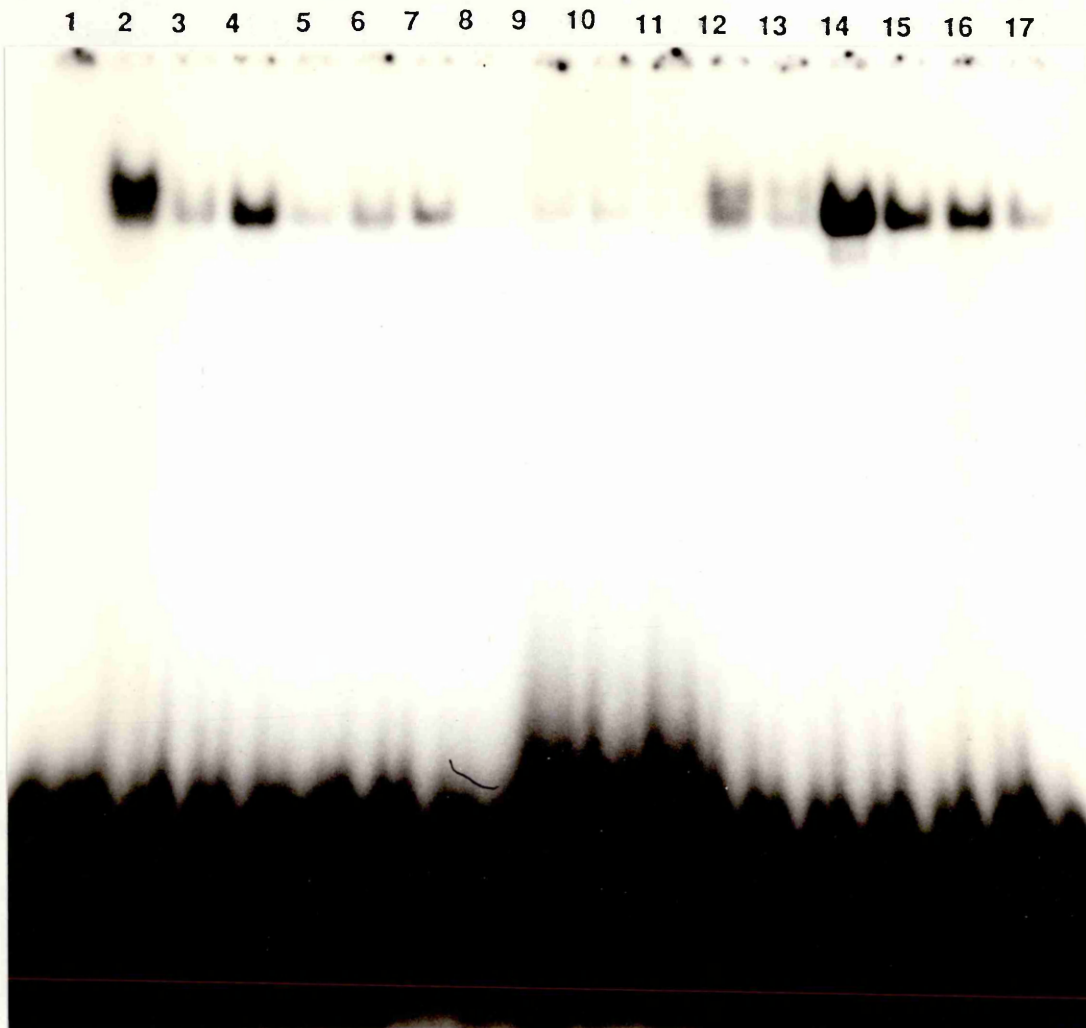
(b) Conditions are identical to those desribed in figure 5.9 (a) except that labelled oligo
m22 was used in the assays. The apparently anomalous result seen in lanes 4 and 14
was obtained when the experiment was repeated independently.

Table 5.1 Relative Affinity of Bound Protein for Different Promoters

These figures represent the amount of retarded oligo, methylated or unmethylated, expressed as % of control in the absence of specific competing oligo. Standard deviations for 4 separate estimations are shown in parentheses.

Competing oligonucleotide

ng competitor DNA	22	m22	m40	NFI	NFIII	p200
10	46.7 (+/-10.6)					
50	25.2 (+/-7.0)	17.0 (+/-7.7)	15.7 (+/-4.1)	42.7 (+/-11.2)	46 (+/-21.4)	
100	20.5 (+/-6.4)	14.7 (+/-6.3)	13.0 (+/-4.5)	28.0 (+/-8.1)	41.2 (+/-17.2)	
200	14.5 (+/-4.8)	8.5 (+/-1.3)	13.7 (+/-2.1)	16.5 (+/-2.1)	23.0 (+/-12.1)	~70
amount(ng) required to give 50% competition (from figure 5.10)	8	6	3	31	40	~500
relative affinity	100 (37.5)	133 (50)	270 (100)	26 (9.7)	20 (7.5)	1.6 (0.6)

Approximately 8ng (20,000cpm) of labelled oligonucleotide were used per assay.

Although these were qualitatively the expected results the differential binding is much less than expected. This could mean that under the binding conditions used Sp1 is less specific than in the presence of 12.5 mM $MgCl_2$ used by Kadonaga and Tjian (1986) for Sp1 -DNA interactions. The alternative possibility exists that the retarded band is only partly (or not at all) the result of Sp1 binding. If two different proteins are binding to the 22mer one may have an affinity for any promoter region and may be competed for equally well by the 22mer ,NFI oligo or NFIII oligo. The other protein would be specific for the GCbox containing 22mer.

It seems highly improbable that Sp1 binding is not being observed because Sp1 retarded bands have been reported by several authors when similar GC box containing oligonucleotides are incubated with these Hela extracts. None of these authors however used promoter containing oligonucleotides in competition experiments, so these are novel results.
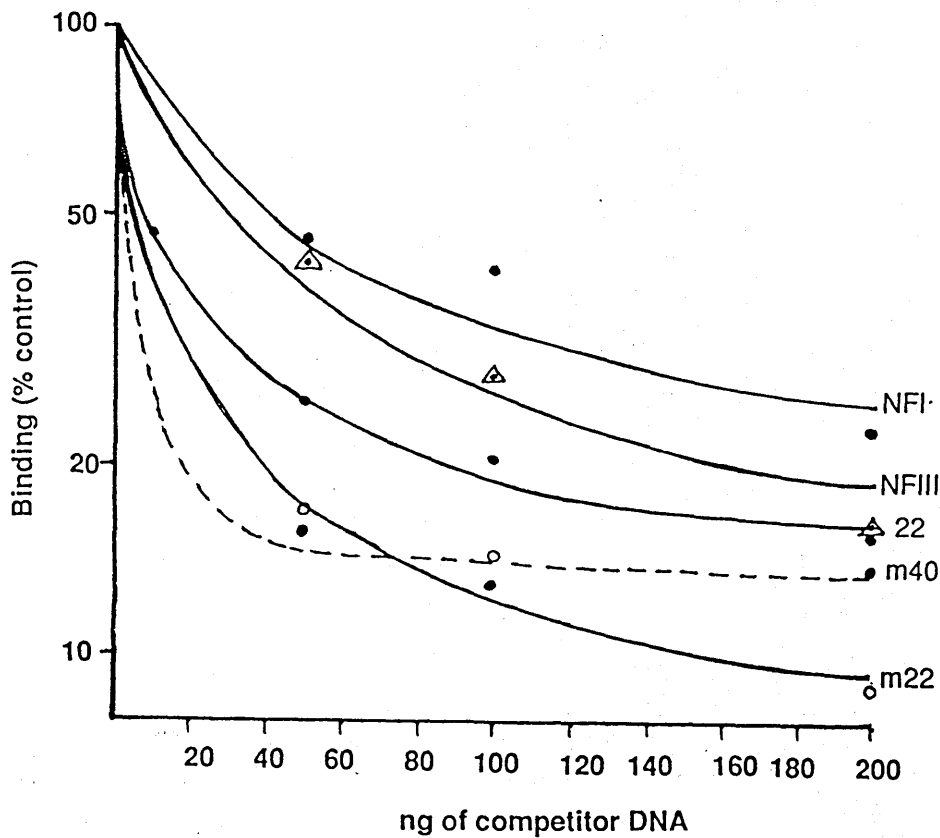

Figure 5.10. Competition by different promoters for nuclear protein binding to oligo 22

Decrease in the amount of of retarded oligo22 in the presence of potential competitor oligonucleotides (NFI ———— • ———— ;NFIII ————△———— ; 22 ———— • ———— ; m22 ———— o ———— ; m40 ———— • ————) as shown in table 5.1.

5.7. Discussion

In this chapter Sp1-containing Hela cell nuclear extracts were used to study the effect of methylation on protein - DNA interactions in a gel retardation assay. Although such extracts have been used by other workers to study Sp1-DNA interactions (Holler *et al.*,1988; Harrington *et al.*,1988)they omitted certain controls. Thus binding to the DNA of other "specific" or "non specific" binding proteins must be eliminated. Since specific and non specific are not absolute terms for DNA-protein interactions but measures of relative affininty of different DNAs for protein(s), varoius types of DNA ((i) model or random DNA , (ii) single stranded DNA, (iii) promoter containing DNA) were used to distinguish them.

From figure 5.8 cold oligo 22 was a far more effective competitor thanp200 or single stranded DNA. Both methylated and unmethylated double stranded oligo competed equally well indicating that cytosine methylation has no inhibitory effect on protein binding. Table 5.1 shows that the addition of 10ng of cold 22mer resulted in approximately 50% competition. Since the assay contains approximately 8ng of labelled 22mer (estimated independently from the specific activity of the γ-^{32}P ATP used and the efficiency of Cherenkov counting) this implies that, cold 22mer had approximately the same affinity for the protein as the labelled material. However 100% competition was never achieved, even after the addition of 200ng of competing 22mer. The only difference between the labelled oligo and the competing DNA is the terminal phosphate. The possibility that this could in some way alter the affinity of the protein, or bind a separate protein was eliminated by competing with cold phosphorylated oligo. No difference in affinity was observed (results not shown). It is thus not known why competition by the 22mer had reduced efficiency when in large excess .

The use of p200 as competitor DNA resulted in some affinity for the binding protein, even in the presence of 1μg poly dI-dC. However, this was approximately

70-100 fold weaker than that observed for the 22mer. This competing effect of p200 was not observed in the absence of poly dl-dC and EDTA, where the presence of 1μg of p200 had no effect on complex formation, but acted to "mop up" non specific DNA binding proteins. The presence of 12.5mM EDTA in the sample buffer eliminates some of the non specific binding (compare lanes 2 and 3 , figure 5.8), presumably by those proteins requiring metal ions as cofactors. These results show that the proteins bind with at least 100 fold greater affinity for the 22mer than for the sequences contained in non promoter DNA.

When single stranded oligonucleotide 22 was used as competitor DNA, no competition was observed. This eliminates the possibility that the retarded material was due to the binding of a single strand specific DNA binding protein such as that identified by Gaillard et al., (1988).

However in these competition experiments double stranded oligos containing other promoter sites e.g. the binding site for nuclear factor I or III had an affinity for the protein only 2-3 fold weaker than did the 22mer. It is possible that zinc finger containing proteins such as Sp1 will bind to these recognition sequences in a sequence non specific manner and it is also possible that the conditions used for protein binding are not ideal for maximum specificity. This experiment should be repeated under varying conditions for example [Mg] and EDTA. Alternatively the retarded complex observed could represent the binding of two separate factors, only one of which is specific for the GC box motif. This could explain why less than the expected degree of specificity was obtained. The fact that oligo m40, containing three GC box motifs appears to be the most effective competitor implies that a GC box specific factor (most likely Sp1) is involved. Sp1 is now regarded as a ubiquitous factor and hence is likely to be more abundant than LSF (Kim et al., 1988) the other factor known to bind this motif.

From the gel retardation studies presented in this chapter it appears that cytosine methylation has no inhibitory effect on protein binding to the SV40 promoter. The oligonucleotides used in this study, 22 and m22, contain GC box I. Gidoni et al.,

(1984) have shown that GC box I is a weak Sp1 binding site compared to interaction at boxes III, V and VI. This weaker interaction may be more susceptible to methylation interference, whereas a strong Sp1-DNA interaction may override methylation interference. However, this was not observed, methylation of GC box I had no inhibitory effect on what is interpreted to be Sp1 binding.

These findings are in accordance with those of Harrington *et al*.,(1988), who found that both methylated and unmethylated oligonucleotides compete equally well with the human metallothionine IIA promoter for binding to Sp1 in DNase I protection experiments. Also, Holler *et al.*,(1988) using a 32bp double stranded oligonucleotide containing two strong Sp1 binding sites found that Sp1 bound tightly regardless of whether the central CpGs were methylated or not. In both these cases no attempt was made to show that Sp1 was the protein involved.

The effect of methylation on the binding of several transcription factors has been studied to date. Hoeveler and Doerfler (1987) found that a cellular factor that interacts upstream of the methylation sensitive Adenovirus late E2A gene binds with the same affinity to methylated and unmethylated promoter DNA.

In contrast, there are several examples of transacting factors that are sensitive to methylation. For example, Watt and Molloy (1988) in studies involving the Adenovirus major late promoter found that binding of the major late transcription factor was strongly inhibited by methylation at a CpG centrally located within the binding site. Whereas, methylation on only one strand caused partial inhibition of binding, the effect being greater when the non-coding strand was methylated. Also, Becker *et al.*,(1987) found, using genomic sequencing, that certain CpG dinucleotides upstream of the rat tyrosine amino transferase gene are methylated in non-expressing cells and unmethylated in expressing cells. DNase I footprinting analysis revealed that *in vitro* methylation of DNA from expressing cells prevented the binding of a cellular factor to its recognition sequence.

These findings suggest that DNA methylation exerts its effects by different mechanisms in different promoters and that not all methylation sensitive promoters bind methylation sensitive transcription factors.   Indeed, there are at least two examples of transacting factors that can overcome the inhibitory effect of methylation. Thomson *et al.*, (1988) have shown that a frog virus 3 induced cellular transacting factor allows transcription from Adenovirus promoters inactivated by methylation. Dobranski *et al.*,(1988) have shown that the inhibitory effect of methylation in the Adenovirus late E2A promoter can be overcome by a 289-amino acid transacting protein that is encoded in the E1A region of Adenovirus.

The inhibitory effect of methylation was also shown to be overcome in myoblast cells.   Yisreali *et al* .,(1986), in transfection experiments found that the muscle specific α– actin gene is inhibited by methylation when introduced into fibroblasts, but not when introduced into myoblasts. This indicates that myoblast cells contain the information required to recognise and activate the α– actin gene despite the presence of methyl groups.

It is important to remember when drawing conclusions on the effect of methylation on protein-DNA interactions that the experiments discussed all involve the use of *in vitro* systems.   We must consider the possibility that DNA methylation may affect the overall chromatin structure and hence interfere with factor binding in an indirect manner (Keshet *et al.*, (1986);Murray and Grosveld (1987)).

Chapter 6. Alternative methods of Studying the Effect of Methylation on Transcription from the SV40 Promoter

6.1. *In Vitro* Transcription Assays

6.1.1. Introduction

The *in vitro* transcription assay provides a useful tool for studying the regulation of transcription of cloned genes. To date these assays have been used to study the control mechanism of a variety of genes and have been particularly useful in establishing the transcriptionally important motifs in promoter regions, mainly through deletion and mutation studies (e.g. Brady *et al.*, 1984; Wildeman *et al.*,1985). Although it is generally believed that an inverse relationship exists between DNA methylation and expression in a large number of genes, most of the evidence for this has come from DNA mediated gene transfer experiments (e.g. Stein *et al* .,1982,1983). There have been relatively few experiments involving the use of *in vitro* methylated genes in *in vitro* transcription assays. One such experiment was performed by Jove *et al.*, (1984) who cloned long segments of poly(dC-dG) adjacent to the transcription control regions of SV40 and Adenovirus 2. After *in vitro* methylation with *Hha* I methylase, transcription was assessed in a soluble *in vitro* system. The presence of these methyl cytosines had no affect on transcription. Neither did methylation at the unique *Hpa* II site downstream of the SV40 late promoter. This contradicts the findings of Fradin *et al* ., (1982), who found that methylation at this *Hpa* II site specifically inhibited late gene expression after microinjection into mouse oocytes. This may suggest that the soluble system is lacking in components that confer sensitivity to DNA methylation and that such components are distinct from those that are required for accurate transcription initiation *in vitro*. Alternatively, methylation may be mediating its effect by altering chromatin formation and hence limiting the accessibility of target genes to the transcriptional apparatus.

Exogenous DNA templates are packaged into chromatin in frog oocyte nuclei but not in the

soluble *in vitro* system. In contrast to these findings with the SV40 promoter ,

Dobranski *et al* .,(1988) found that methylation of the adenovirus late E2A and major late

promoters resulted in almost complete inhibition of transcription in an *in vitro* system.

These findings suggest that methylation may exert its inhibiting effect on different

promoters in a different way. Therefore, the use of *in vitro* transcription assays may be

important in discovering the mechanism by which methylation imposes its inhibitory

effect.

6.1.2. *In Vitro* Transcription of pVHCl

The aim of this set of experiments was to methylate plasmid DNA *in vitro* using

mouse ascites DNA methylase and study the effect this has on transcription in a soluble *in*

vitro system. In chapter 4, the effect of methylation on transcription from the SV40

promoter was studied by transfection experiments. However, transfection experiments

require relatively large amounts of plasmid DNA (10μg). Using partially purified

mammalian methylases it is difficult to methylate such a quantity of DNA to a high degree of

saturation. However *in vitro* transcription assays require only 0.1-0.5μg of template

DNA, a quantity that can be efficiently methylated in vitro (see section 3.3).

Using a cell free extract from Hela cells , prepared according to the method of

Manley *et al* .,(1980) (see section 2.2.5.1), the plasmid pVHCl was used as a DNA template

in a *in vitro* transcription assays. Transcripts were labelled by the incorporation of

[α32P] -UTP and separated by agarose gel electrophoresis. Fig. 6.1 shows the resulting

mRNA transcripts. Three separate transcripts were observed, the intensity of which

increased with increasing template concentration. Cold RNA size markers were run

alongside the transcripts, these were visualised by ethidium staining and U.V. irradiation.

The size of the transcripts were approximately 350, 500 and 650 nucleotides. All three

are smaller than the expected 770 nucleotide CAT gene transcript. Unfortunately it

appears that the same three transcripts are also produced when the promoterless p200

construct was used template DNA. This indicates that transcription is not being initiated at

the SV40 promoter but at some endogenous site in the pUC8 vector or within the CAT gene.

It is obvious then, that this vector system is not suitable for this type of experiment.

Hence this line of research was discontinued.
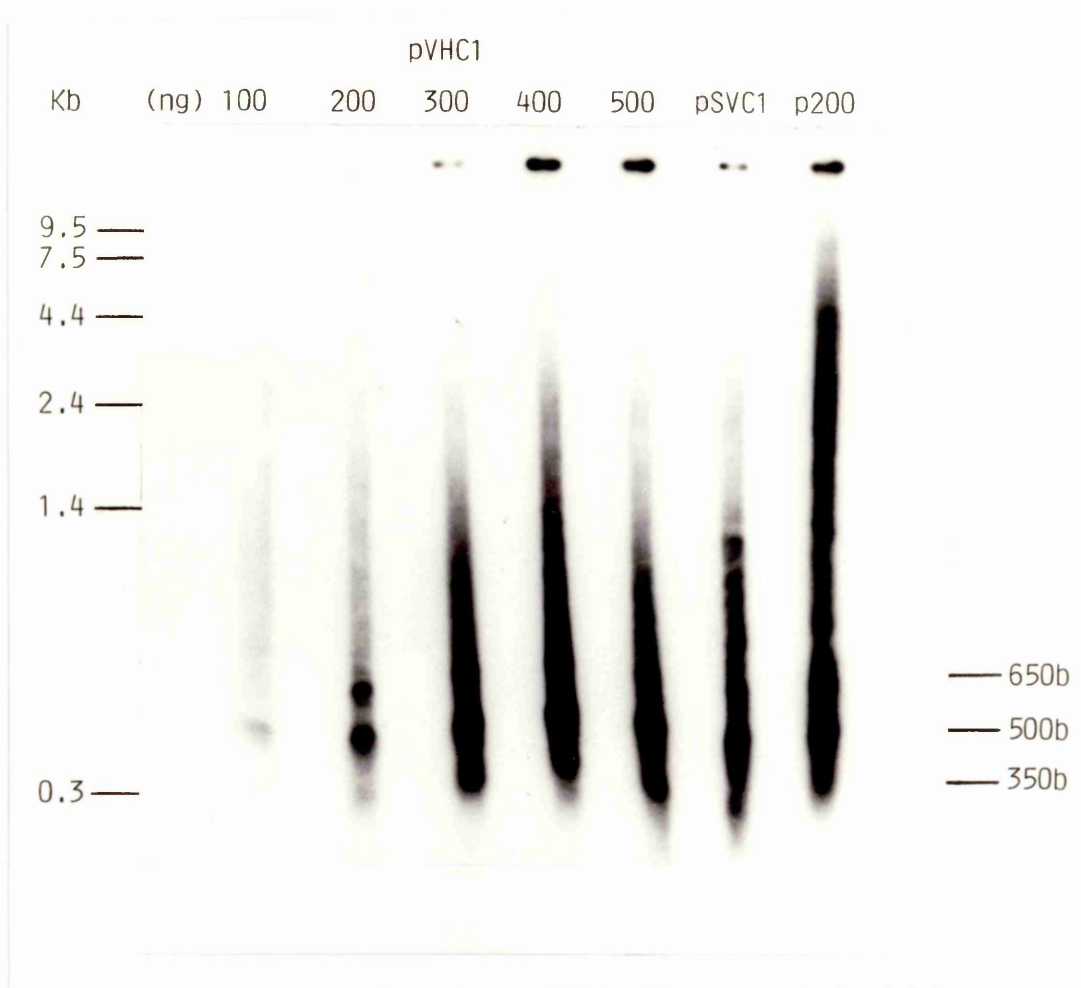
Figure 6.1. *In vitro* transcription of pVHCl

pVHCl (100-500ng) lanes 1-5 was used in an *in vitro* transcription assay as described in section 2.2.5.2. In lane 6 pVHCl was replaced with pSVC1(200ng) and in lane 7, p200 (200ng). The three major radioactive bands are present in all lanes. An incubation omitting the DNA produced no radioactive material on electrophoresis(results not shown).

6.2. Cloning into M13

Another way to produce methylated plasmid for use in transcription assays, and a way which might circumvent the adventitious promotion observed with pVHC1 is to use DNA cloned into M13. Since the M13 bacteriophage has the ability to produce single stranded DNA in addition to replicative form, it has proved to be a useful research tool in a wide range if areas. These include site directed mutagenesis (Zoller and Smith, 1983) and DNA sequencing (Sanger et al., 1977). Keshet et al .,(1986) found that single stranded recombinant M13 molecules provided a convenient way of generating hemimethylated DNA . Using single stranded M13 DNA as a template molecule, a short single stranded primer could be annealed and the complementary strand of DNA synthesised by the Klenow fragment of DNA polymerase I. Replacing deoxycytidinetriphosphate (dCTP) with deoxy- methylcytidine triphosphate (mdCTP) meant that the DNA was hemimethylated at every cytosine in the sequence. The unmethylated strand could then be methylated using mammalian methylases either *in vitro* or *in vivo* after introduction into tissue cultured cells (since hemimethylated DNA is the preferred substrate for these enzymes), and used in transcription studies.

In order to use this approach in studying the SV40 promoter, the 601bp *Kpn* I / *Eco* RI fragment of pVHCl containing the SV40 enhancer and 21bp repeat region, and the 2107bp *Kpn* I / *Eco* RI fragment of pVHCl containing the SV40 enhancer and promoter, CAT gene and SV40 terminator sequence were cloned into M13 mp19. Recombinant M13 molecules were analysed for the presence of insert by restriction digestion (results no shown), and by T. tracking (section 2.2.6.7.). The complementary strand of these single stranded recombinant molecules could then be synthesised as described above (details in

section 2.2.6.8) in the presence or absence of mdCTP. This DNA could then be used in *in vitro* transcription and transfection studies.

It was also intended to use such constructs to study the effect of methylation on transcription factor binding in gel retardation assays by using the approach outlined in Fig. 6.2. The M13 sequencing primer was annealed to the template just upstream of the insert, the second strand was synthesised as described above, in the presence of dCTP or mdCTP and [α–32P]-dATP. The SV40 promoter region was then isolated by digestion with the appropriate restriction enzymes and the fragments separated by polyacrylamide gel electrophoresis (see section 2.2.6.7.). However a major problem encountered was the failure of restriction enzymes to cleave DNA with a high methyl cytosine content (fig. 6.3). The high cytosine content of the SV40 promoter meant that restriction enzymes such as *Nco* I cleaved the hemimethylated DNA with very poor efficiency, which led to very low yields of methylcytosine containing restriction fragment. This problem could be partially alleviated by 'filling in' with a 1:1 or 2:1 ratio of dCTP:mdCTP. Preliminary studies indicated that protein DNA interactions still occurred in the presence of 50% methyl cytosine.

Synthetic oligonucleotides containing 5-methylcytosine were synthesised for use as primers. The sequence of these oligonucleotides were complementary to GC box I, , or I,II and III of the SV40 promoter. Substitution of these molecules for the universal primer in the above experiment produced molecules methylated at these sites only. However, since the use of these oligonucleotides in gel retardation assays involving competition with the labelled *Sph* I/*Nco* I fragment of pSVCI revealed that the presence of 5-methylcytosine had no effect on binding (presented in chapter 5), these 'fill in' experiments were discontinued.

6.2.1. The Use of M13 Recombinant Molecules in Transfection Studies

Unfortunately the synthesis of sufficient double stranded DNA for transfection assays (10μg) proved to be very difficult and this approach was abandoned and replaced by the synthesis of gap duplex molecules. This involves the annealing of two single stranded molecules: the M13 recombinant and a linearised non recombinant molecule. This produces a duplex molecule with a single stranded "window" at the position of the insert. The second strand of this region can be synthesised in the presence or absence of mdCTP. Alternatively, this single stranded region can be methylated using mammalian methylase (Adams *et al*.,(1989) found that only single stranded DNA is methylated in the presence of 0.1M NaCl.). It can then be made double stranded (using Klenow fragment), creating a hemimethylated region, which itself is the preferred substrate for DNA methylase. This gap duplex approach allows for region specific methylation of the SV40 promoter, the transcription initiation sites, and the CAT gene, and is important for determining which CpG's or CpG containing regions are required to be methylation free for efficient transcription to occur. These experiments are currently being continued in our laboratory.
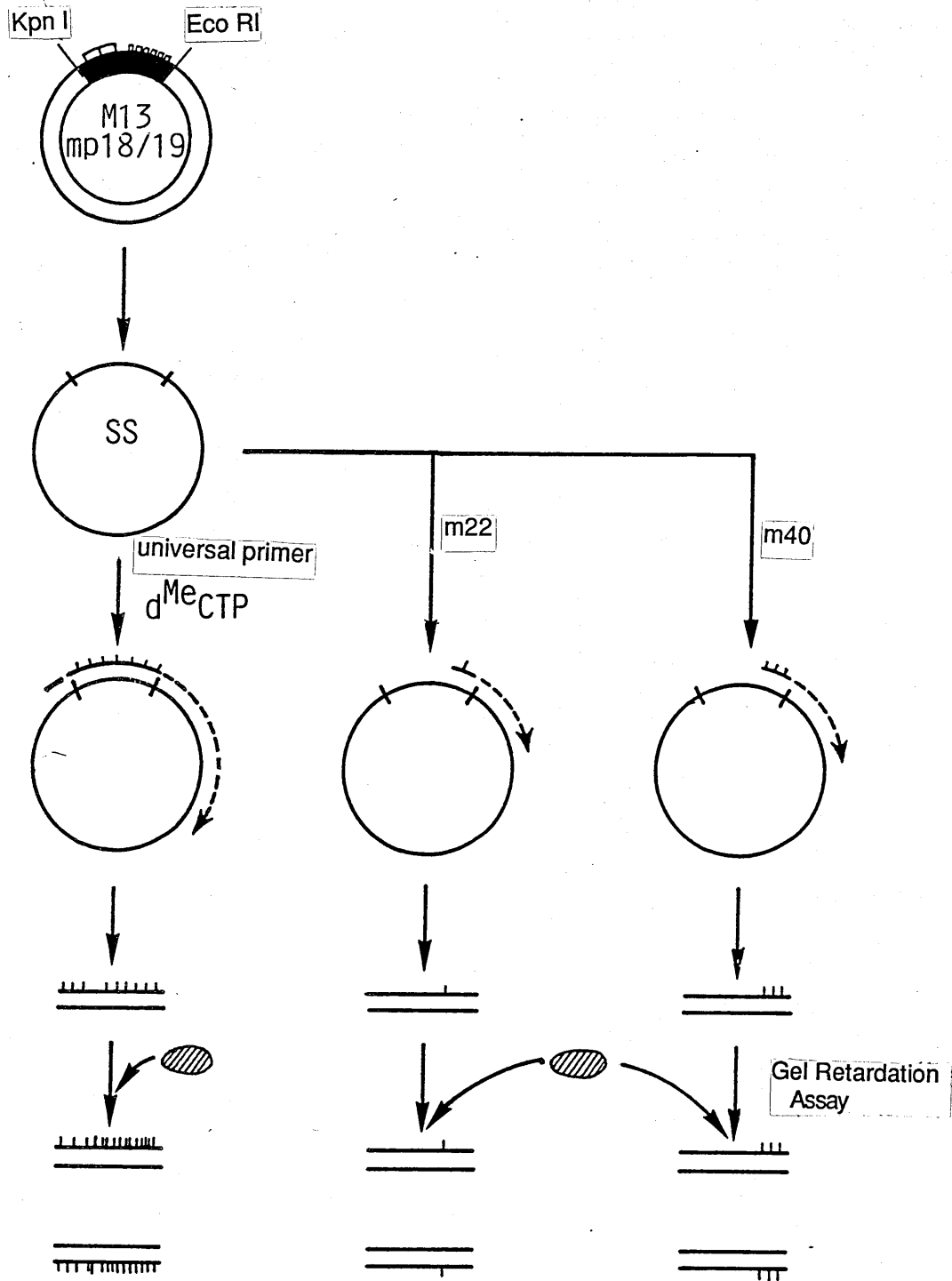
Figure 6.2. Use of M13 Recombinants in Gel Retardation Assays
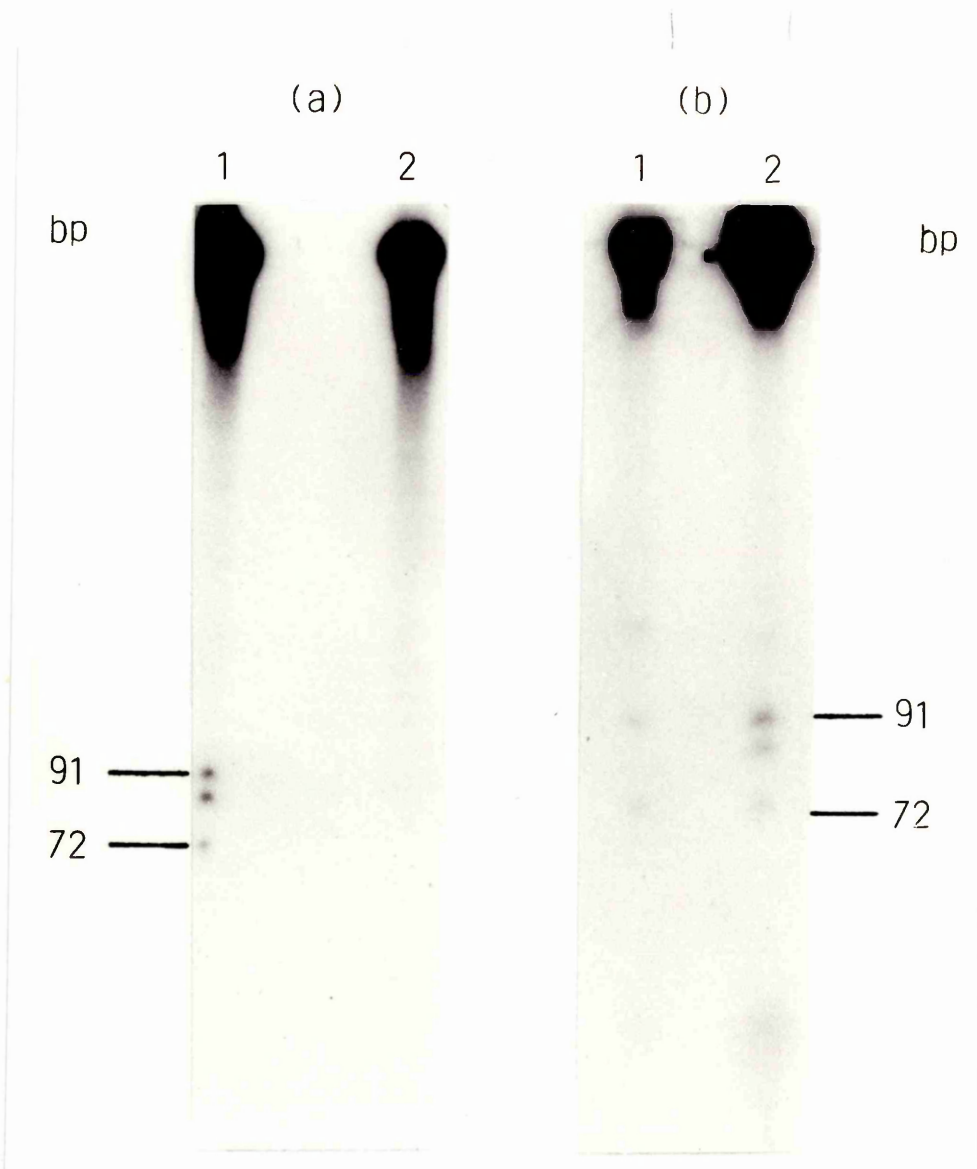
Figure 6.3. Isolation of Hemimethylated Restriction Fragments

Labelled restriction fragments were separated on a 8% polyacrylamide gel

(a) lane 1 *Sph* I / *Nco* I digest of DNA filled-in with dCTP

lane 2 *Sph* I/ *Nco* I digest of DNA filled-in with mdCTP

(b) lane 1 *Sph* I/ *Nco* I digest of DNA filled-in with a 1:1 ratio of dCTP:mdCTP

lane 2 *Sph* I/*Nco* I digest of DNA filled-in with dCTP

Both gels were autoradiographed with a 4 minute exposure time.   In gel (a) the radioactive

bands in lane 2 were visible after a longer exposure (several hours)

Chapter 7. Discussion

The results presented in Chapter 4 show that relatively low levels (10-25%) of methylation lead to inhibition of transcription from the SV40 early promoter. Assuming an even distribution of methyl groups this level of saturation would result in only one to two of the ten or so CpG dinucleotides in the promoter region being methylated. Since there are six Sp1 binding sites in this region, each containing a CpG, on average at least four of these would remain unmodified. Hence it is unlikely that the inhibitory effect of methylation is due to inhibition of transcription factor binding to the GC boxes or indeed to any other site. It is more likely that it is due to methylation -induced formation of an altered chromatin state after the plasmid DNA has been introduced into mouse cells. This would suggest that the SV40 promoter is one of a group including the γ^A-globin (Murray and Grosveld) and HSV thymidine kinase (Keshet *et al.*,1986 ,Ben Hatter and Jiricny,1988) genes, where it has been shown that the down regulatory effect of methylation is mediated *via* formation of inactive chromatin.

In order to establish which regions in the SV40 promoter containing construct are important for this inhibitory effect it is necessary to devise a system for carrying out region specific methylation . Murray and Grosveld (1987) investigated which of the six CpG s in the γ^A-globin promoter were important for inhibition by methylation. Using mutation studies they found that both the three sites upstream of the cap site and the three sites downstream were required to be unmethylated for transcription i.e the region between residues +92 and -210 in the promoter was required to be methylation free for expression of this gene by ubiquitous factors. It would be instructive to carry out such a study on the SV40 promoter region but since none of the CpGs lie in sequences methylated by prokaryotic methylases it is not possible to perform site directed methylation by this method. However methylation of specific regions of DNA

can be achieved using gap duplex molecules. The preparation of these molecules is outlined

in chapter 6 and involves cloning the SV40 promoter-CAT gene restriction fragment into

M13 and annealing linearised vector DNA to the single stranded form of this recombinant,

thus creating a single stranded window where the insert lies. If, in place of linearised

vector, recombinant molecules containing inserts with specific deletions are used a range

of molecules with different positions and lengths of single stranded window can be

generated. The single stranded regions can then be preferentially methylated using mouse

ascites cell DNA methylase. They could then be converted to the double stranded form

using the Klenow fragment of DNA polymerase I, and used in transfection studies. Using

this method one can study the importance of the CpGs in the promoter, transcription

initiation site, CAT gene and vector sequences in conferring this transcription inhibition by

methylation.

Consistent with the mechanism of transcription inactivation proposed above

the results presented in chapter 5 indicate that the binding of Sp1 is unaffected by

methylation. This has also been observed by Harrington *et al.*,(1988)and Holler *et

al.*,(1988). However studies on the effect of methylation on the binding of several other

transcription factors revealed that methylation can interfere with binding. Such factors

include the adenovirus major late transcription factor (MTLF) (Watt and Molloy,1988)

and a cellular factor that binds upstream of the tyrosine aminotransferase gene promoter

(Becker *et al.*,1987). Studying the effect of methylation on protein-DNA interactions in

a range of genes may reveal a relationship between those genes that bind methylation

sensitive factors and those that bind methylation insensitive factors. It is possible that, *in

vivo*, ubiquitous factors that are involved in the transcription of many genes, such as Sp1,

will be methylation insensitive; whereas tissue specific factors will be methylation

sensitive, allowing only those genes with the appropriate methylation pattern to be transcribed.

Iguchi-Ariga and Schaffner (1989) compared the effect of methylation on two factors -the ubiquitous Sp1 and a factor that binds to the cAMP response element (CRE). This sequence motif has been identified in the 5' flanking regions of various cell-type specific genes, including hormone coding and viral genes. They found that this sequence element was methylation sensitive both "*in vivo*", in a transient expression assay, and *in vitro* in a cell free transcription system. Using gel retardation assays they found that the protein factor(s) that bind to this region were methylation sensitive. These findings indicate that the inhibitory effect of methylation in this cell type specific gene is not due to the disruption of chromatin since this effect is present in *in vitro* transcription assays which do not involve the formation of chromatin structures prior to transcription. It appears that, in these genes, methylation inhibits transcription by preventing the binding of cAMP induced transcription factor(s).

Another example of this type of inhibition is found in the adenovirus 2 late E2A gene promoter and major late promoter where methylation of *Hpa* II and *Hha* I sites in the promoters inhibited transcription in a cell free transcription system as well as in transfection assays. However footprinting analysis on the E2A promoter revealed that binding of specific factors was not inhibited by methylation. However it is conceivable that even though these host proteins bound to methylated DNA their activity may be altered in the presence of methyl groups.

In contrast Holler *et al.*, (1988) found that methylation of a GC box upstream of the β-globin gene had no effect on transcription in a cell free system. This is consistent with the argument that the inactivation by methylation in SV40 is due to

altered chromatin formation as opposed to the direct inhibition of transcription factor binding.

One could propose, that house keeping genes are not under the control of methylation *in vivo*. Although methylation of several of these genes *in vitro* results in transcriptional inhibition, *in vivo* these genes would be unmethylated and constitutively expressed in all cells. This constitutive expression would be regulated by CpG islands upstream of these genes which are characteristically methylation free. It has been proposed that during development these sequences would be maintained in an unmethylated state by the binding of Sp1(Holler *et al.*,1988) which would prevent *de novo* methylation of these sites. Alternatively, since these islands can be up to 2 kilobases in size methylation may be discouraged by formation of an altered DNA structure induced by the high G+C content of the DNA or by the binding of a stage specific factor. On the other hand, tissue-specific expression of genes requires further fine tuning of the transcription control mechanisms which involve the establishment of site specific methylation patterns. Each methylcytosine in promoter regions may act independently by inhibiting the binding of sequence specific factors or a group of methylcytosines may act together to create an inactive stretch of chromatin.

An alternative theory is that methylation acts as a secondary mechanism of inactivation. It is possible that genes are initially silenced by other mechanisms which then recruit methylation as a back up system to ensure that the gene remains in an inactive state possibly by the formation regions of condensed inactive chromatin.

From the experiments discussed here it is clear that it is beneficial to study the effect of methylation on the activity of a promoter using more that one method. DNA mediated gene transfer experiments such as transfection assays and microinjection into

cell nuclei result in the exogenous DNA being packaged into chromatin before expression. These systems mimic the in vivo system most closely and allow the effect of methylation on chromatin formation to be studied. Comparing these experiments with cell free in vitro transcription assays which do not involve the formation of chromatin prior to transcription can provide insight into the mechanism of methylation inactivation in a given gene. For example, Fradin et al., (1982) showed that in oocyte microinjection experiments, methylation of the single Hpa II site in SV40, located near the 5' region of the late gene specifically repressed late gene expression. However this effect was not observed in an in vitro transcription assay using Hela cell extracts (Jove et al.,1984). Similarly, while Holler et al.,(1988) found that in an in vitro transcription assay, methylation of a GC box inserted upstream of the β-globin gene had no effect on expression, when the same sequence was inserted upstream of the HSVtk gene and used in Xenopus oocyte injection experiments, a 20 fold inhibition was observed (Ben-Hatter and Jiricny,1988). These results indicate that the down regulatory effect of methylation is brought about by a change in the chromatin assembly, rather than through a direct interference with the binding of transcription factors. This latter possibility has been studied in several systems using gel retardation or DNAase I footprinting assays on cloned genes. These assays provide a convenient method for analysing protein - DNA interactions and the possible effects of methylation. However,in these systems one is observing an all or nothing effect and the possibility that methylation results in a partial inhibition of binding cannot be accurately studied. Similarly the possibility that methylation allows factor binding but not activity, as proposed for the adenovirus E2a gene (Hoeveler and Doerfler,1987),would not be apparent from these studies.

When performing the above types of experiments to study the effect of

methylation on transcription of a cloned gene it must be remembered that the results

obtained *in vitro* may not necessarily mimic the *in vivo* situation. This is also true for

the correlation of restriction enzyme-derived methylation patterns of cloned genes and

their expression. However this problem has been largely overcome by the use of genomic

sequencing. Using this technique all methylcytosines within a gene can be located and the

involvement of these sites in specific factor binding can be studied using *in vivo*

footprinting. These experiments should in time provide definitive evidence for the

involvement of methylation in the transcription of eukaryotic genes and the role of

methylation in specific and non specific protein binding.

Most of the *in vitro* methylation experiments performed to date have

involved the use of bacterial methylases, such as *Hpa* II and *Hha* I methylase. Their major

disadvantage is that their use is restricted solely to those genes which contain the

methylase recognition sites in their 5' flanking regions. The binding sequences for several

important transcription factors such as Sp1, CTF, AP-1 and AP-2, do not contain such

sites. In recent years synthetic oligonucleotides containing m^5C have provided a

convenient way of studying the effect of methylation on the binding of these factors.

Ideally, all methylation sites within the control and coding regions of a gene should be

studied *in vitro*, however this could prove difficult and very expensive using synthetic

oligonucleotides and will only be achieved on the purification of sufficient quantities of a

pure and stable mammalian DNA methyltransferase.

Although the effect of methylation on the expression of many genes has been

studied to date, it is still not known where methylation fits into the complex multilevel

control system of gene regulation. This understanding may be aided by further studies on

the effect of methylation on the interaction with DNA of the large number of specific factors

that have recently been identified as being involved in the transcription of eukaryotic

genes.

References

Adams, R.L.P., Bryans, M., Rinaldi, A., Smart, A., and Yesufu, M. (1989) *Proc. Royal Soc.* In press.

Adams, R.L.P. and Burdon, R.H. (1985) *Molecular Biology of DNA Methylation,* Springer Verlag. New York.

Adams, R.L.P. and Burdon, R.H. (1983) In *Enzymes of Nucleic Acid Synthesis and Modification,* vol 1, 119-144, Jacob,S.T.,ed. CRC Press Inc., Boca Raton, FL.

Adams, R.L.P., Burdon, R.H., and Fulton, J. (1983a) *Biochim. Biophys. Res. Comm.* **113,** 695-702.

Adams, R.L.P., Burdon,R.H., McKinnon, K., and Rinaldi, A. (1983b) *FEBS. Lett.* **163** (2), 194-198.

Adams, R.L.P., Gardiner, K., Rinaldi, A., Bryans, M., McGarvey,M., and Burdon, R.H. (1986)*Biochim. Biophys. Acta.* **868,** 9-16.

Adams, R.L.P., McKay, E.L., Craig,L.M., and Burdon,R.H. (1979) *Biochim. Biophys. Acta* **561,** 345-357.

Adams, R.L.P., Rinaldi, A., McGarvey, M., Bryans, M., and Ball, K. (1988) *Gene* **74** (1), 125-128.

Arnt-Jovin, J.D., Robert-Nicoud, M., Zarling, D.A., Greider, C., Weiner, E., and Jovin, J.M. (1983) *Proc. Natl. Acad. Sci. USA.* **80,** 4344-4348.

Baty, D., Barrera-Sardana,H., Everett, R., Vigneron, M., and Chambon, P. (1983) *Nucleic Acids Res.* **12,** 915-932.

Becker, P.B., Ruppert , S. and Schutz, G. (1987) *Cell* **51,** 435-443.

Ben-Hatter, J. and Jiricny, J. (1988) *Gene* **65,** 219-227.

Behe, M. and Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. USA.* **78,** 1619-1623.

Benoist, C. and Chambon, P. (1981) *Nature* **290,** 304-310.

Bestor, T.H. and Ingram, V.M. (1983) *Proc. Natl. Acad. Sci. USA.* **80**, 5559-5563.

Bestor, T., Landano, A., Mattaliano, R., and Ingram, V. (1988) *J. Mol. Biol.* **203**, 971-983.

Billin, D. (1968) *J. Mol. Biol.* **31**, 477-486.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and MacLeod, D. (1985) *Cell* **40**, 91-99.

Birmboin, H.C. and Daly, J. (1979) *Nucleic Acids Res.* **7**, 1513-1523.

Bolden, A.H., Ward, C.A., Nalin, C.M., and Weissbach, A. (1986) *Prog. Nucleic Acids Res.* **33**, 231-251.

Boyer, H.W. and Roulland-Dussoix, D. (1969) *J. Mol. Biol.* **44**, 459-472.

Bradford, M.M., (1976) *Anal. Biochem.* **72**, 248-254.

Brady, J., Radonovich, M., Thoren, M., Das, G. and Salzman, N.P. (1984) *Mol. Cell. Biol.* **4** (1), 133-141.

Brown, T.C. and Jiricny, J. (1987) *Cell* **50**, 945-960.

Burton, W.G., Grabowy, C.T., Sager, R. (1979) *Proc. Natl. Acad. Sci. USA.* **76**, 1390-1394.

Buschhausen, G., Wittig, M., Graessmann, M.,and Graessmann, A. (1987) *Proc. Natl. Acad. Sci. USA.* **84**,1177-1181.

Cate, R.L., Cluck, W., and Gilbert, W. (1983) *J. Biol. Chem.* **258**, 6645-6652.

Cassidy, P., Kahan, F. and Algeria, A.A. (1965) *Fed. Proc.* **24**, (2I) 226.

Cedar, H., Stein, R., Gruenbaum, Y., Naven-Many, T. ,Sciaky-Gallili, N. and Razin, A. (1983)*Cold Spring Harbor Symp. Quant. Biol.* **47**, 605-609.

Chamberlain, J.P. (1979) *Analytical Biochem* . **98**, 132-135.

Chargaff, E. (1955) In: *The Nucleic Acids* , Chargaff, E. and Davidson, J.N., Eds. Vol. I;
Chapter 10. pp 307-371, Academic, New York.

Chargaff, E., Crampton, C.F., and Lipshitz, R. (1953) *Nature* **172**, 289-292.

Christman, J.K., Price, P., Pedrinan, L., and Acs, G. (1977) *Eur. J. Biochem.* **81**,53-61.

Christy, B. and Scangos, G. (1982) *Proc. Natl. Acad. Sci. USA.* **79**, 6299-6303.

Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA.* **81**, 1991-1995.

Clark, L. and Hay, R. (1989) *Nucleic Acids Res.* **17**(2), 499-516.

Clark, L., Pollock, R.M., and Hay R.T. (1988) *Genes & Development* **2**, 991-1002.

Cohen, J., Eccleshall, T., Needleman, R., Federoff, H., Buchferer, B., and Marmur, J.
(1980) *Proc. Natl. Acad. Sci. USA.* **77**, 1078-1082.

Collins, M. and Myers, R.M. (1987) *J. Mol. Biol.* **198**, 737-744.

Compere, S.J. and Palmiter, R.D. (1981) *Cell* **25**, 233-240.

Cooney, C.A., Eykholt, R.A. and Bradbury, E.M. (1988) *J. Mol. Biol.* **204**, 889-901.

Courey, A.J. and Tjian, R. (1988) *Cell* **55**,887-898.

Dagert, M. and Erlich, S.D. (1979) *Gene* **6**, 23-28.

Davidson, I., Xiao, J.H., Rosales, R., Staub, A., and Chambon, P. (1988) *Cell* **54**,
931-942.

Dawid, I.B., Brown, D.D., and Reeder, P.H. (1970) *J. Mol. Biol.* **51**, 341-360.

Desrosiers, R.C., Mulder, C., and Fleckenstein, B. (1979) *Proc. Natl. Acad. Sci. USA.* **76**,
3839-3843.

Diala, E.S. and Hoffman, R.M. (1983) *J. Virol.* **45**, 482-483.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**(5),
1475-1489.

Dobranski, P., Hoevler, A., and Doerfler, W. (1988) *J. Virol.* **62** (II), 4941-3946.

Doerfler, W., Weisshaar, B., Hoeveler, A., Kneber, D., Muller, U., Dobranski, P.,
Lichtenberg, U., Achten, S. and Hermann, R. (1988) *Gene* **74**, 129-133.

Doerfler, W. (1981) *J. Gen. Virol.* **57**, 1-20.

Dorn, A., Bollekeus, J., Stauls, A., Benoist, C. and Mathis, D. (1987) *Cell* **50**, 863-872.

Drahovsky, D. and Morris, N.R. (1971) *J. Mol. Biol.* **57**, 475-489.

Dynan, W.S., Saffer, J., Lee, W.S., and Tjian, R. (1985) *Proc. Natl. Acad. Sci. USA.* **82**,
4915-4919.

Dynan, W.S., Sazer, S., Tjian, R., and Schimke, R.T. (1986) *Nature* **319**, 246-248.

Dynan, W.S. and Tjian, R. (1983) *Cell* **32** , 669-680.

Dynan, W.S. and Tjian, R. (1985) *Nature* **316**, 774-778.

Fisher, E.F. and Caruthers, M.H. (1979) *Nucleic Acids Res.* **7**, 401-416.

Ford, J.P., Coca-Prados, M., and Hsu, M. (1980) *Biol. Chem.* **255**, 7544-7547.

Fox, K.R. (1986) *Biochem J.* **234**, 213-216.

Fradum, A., Manley, J.L., and Prives, C.L. (1982) *Proc. Natl. Acad. Sci. USA.* **79**,
5142-5146.

Fried, M. and Crothers, D.M. (1981) *Nucleic Acids Res.* **9** (23), 6505-6525.

Fromental, C., Kanno, M., Nomiyama, H., and Chambon, P. (1988) *Cell* **54**, 943-953.

Gaffney, D.F., McLaughlan, J., Whitton, J.L., and Clements, J.B. (1985) *Nucleic Acids Res.*
13(21), 7847-7863.

Gaido, M.L. and Strobl, J.S. (1989) *Biochim. Biophys. Acta* **1008**, 234-242.

Gaillard, C., Weber, M., and Strauss, F. (1988) *J. Virol* . **62**, 2380-2385.

Garner, M.M. and Revzin, A. (1981) *Nucleic Acids Res.* **9**(13), 3047-3060.

Gerber-Huber, S., May, F.E.B., Westley, B.R., Felber, B.K., Hosbach, H.A., Andres, A.C.,
and Ryfell, G.H. (1983) *Cell* **33**, 43-51.

Gey, G.O., Coffman, W.D. and Kubicek, M.T. (1952) *Cancer Res.* **12**, 264-265.

Gidoni, D., Dyaan, W.S. and Tjian, R. (1984) *Nature* **312**, 409-413.

Gilmour, R.S., Gow, J.W., and Spandidos, D.A. (1982) *Biosci. Rep.* **2**, 1031.

Gjerset, R.A. and Martin, D.W. (1982) *J. Biol. Chem.* **257**, 8581-8583.

Gorman, C., Laimons, L., Merlino, G.T., Gruss, P., Khoury, G., and Howard, B. (1984) In :

    Eukaryote Gene Expression ed. A. Kumour. G. Wash. Univ. Med. Centre Spring

    Symp. pp. 129-139.

Gorman, C., Merlino, G., Willingham, M., Pastan, I., and Howard, B. (1982b) *Proc. Natl.*

    Acad. Sci. USA. **77**, 6777-6781.

Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982a) *Mol. Cell. Biol.* **2**(9),

    1044-1051.

Graessman, A. and Graessman, M. (1988) *Gene* **74**, 135-137.

Graessman, M., Graessman, A., Wagner, H., Werner, E., and Simon, D. (1983) *Proc. Natl.*

    Acad. Sci. USA. **80**, 6470-6474.

Graves, B.J., Johnson, P.F., and McKnight, S.L. (1986) *Cell* **44**, 565-576.

Gunthert, U., Schweiger, M., Stupp, M., and Doerfler, W. (1976) *Proc. Natl. Acad. Sci.*

    USA. **73**, 3923-2927.

Haashoot, C.A.G. and Altona, C. (1979) *Nucleic Acids Res.* **6**, 1135-1149.

Hansen, R.S., Ellis, N.A. and Garther, S.M. (1988) *Mol. Cell. Biol.* **8**(11), 4692-4699.

Hare, J.T. and Taylor, J.H. (1985) *Proc. Natl. Acad. Sci. USA.* **82**, 7250-7354.

Harrington, M.A., Jones, P.A., Imagawa, M. and Karin, M. (1988) *Proc. Natl. Acad. Sci.*

    USA. **85**, 2066-2070.

Hayashi, S. and Kondon, H. (1986) *Mol. Cell Biol.* **6**(II), 4130-4132.

Hirt, B. (1967) *J. Mol. Biol.* **26**, 365-369.

Hoeveler, A. and Doerfler, W. (1987) *DNA* 6(5), 449-460.

Holler, M., Westin, G., Jiricny, J., and Schaffner, W. (1988) *Genes & Development* **2**,

1127-1135.

Holmes, D.S. and Quigley, M. (1981) *Analytical Biochem* . **114**, 193-197.

Huang, L-H., Wang, R., Gama-Sosa, M.A., Shenoy, S., and Erlich, M. (1984) *Nature* **308**,

293-295.

Hughes, M.J., Liang, H., Jiricny, J., and Jost, J.P. (1989) *Biochemistry* in press.

Iguchi-Ariga, S.M.M. and Schaffner, W. (1989) *Genes & Development* **3**, 612-619.

Imagawa, M., Chin, R., and Karin, M. (1987) *Cell* **51**, 251-260.

Imbra, R.J. and Karin, M. (1987) *Nature* **323**, 555-558.

Jackson, S.P. and Tjian, R. (1988) *Cell* **55**, 125-133.

Jantzen, H-M., Strahle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R.,

and Schutz, G (1987) *Cell* **489**, 29-38.

Janulaitis, A., Klimasauskas, S., Petrusyte, M., and Butkus, V. (1983) *FEBS Letts.* **161**,

131-134.

Johnson, P.F., Landshulz, W.H., Graves, B.J., McKnight, S.L. (1987) *Genes & Develoment*

1, 133-143.

Jones, K.A. and Tjian, R (1985) *Nature* **317**, 179-182.

Jones, K.A., Yamamoto, K.R. and Tjian, R. (1985) *Cell* **42**, 559-572.

Jones, M., Wagner, R., and Radman, M. (1987) *Cell* **50**, 621-626.

Jones, P.A. and Taylor, S.M. (1980) *Cell* **20**, 85-93.

Jones, P.A., Taylor, S.M., Mohandas, T., and Shapiro, L.J. (1982) *Proc. Natl. Acad. Sci.*

USA. **79**, 1215-1219.

Jove, R., Sperber, D.E., and Manley, J.L. (1984) *Nucleic Acids Res.* **12** (II),

4715-4730.

Kadonaga, J.T., Carner, K.R., Masiarz, F.R., and Tjian, R. (1987) *Cell* **51**, 1079-1090.

Kadonaga, J.T., Jones, K.A., and Tjian, R. (1986) *Trends in Biochem* **II**, 20-23.

Kadonaga, J.T. and Tjian, R. (1986) *Proc. Natl. Acad. Sci. USA.* **83**, 5889-5893.

Karin, M. (1989) *Trends in Genetics* **5** (3), 65-67.

Karin, M., Haslinger, A., Holtgreve, H., Richards, R.I., Kranter, P., Westphal, H.M., and

Beato, M. (1984) *Nature* **308**, 513-519.

Kautiainen, T.L. and Jones, P.A. (1985) *Biochemistry* **24**,1193-1196.

Kawakami, K., Scheidereit, C., and Roeder, R.G. (1988) *Proc. Natl. Acad. Sci. USA.* **85**

4700-4704.

Kaye, A.M. and Winocour, E. (1967) *J. Mol. Biol.* **24**, 75-81.

Kesher, I., Lieman-Hurwitz, J. and Cedar, H. (1986) *Cell* **4** , 535-543.

Keshet, I., Yisreali, J., and Cedar, H. (1985) *Proc. Natl. Acad. Sci. USA.* **82**, 2560-2564.

Kim, C.H., Heath, C., Bertuch, A., and Hansen, U. (1987) *Proc. Natl. Acad. Sci. USA.* **84** ,

6025-6029.

Klein-Hitpaβ, L., Schorpp, M., Wagner, U., and Ryffel, G.U. (1986) *Cell* **46**,

1053-1061.

Kmiec, E.G. and Worcel, A. (1985) *Cell* **41**, 945-953.

Knebel-Morsdorf, D., Achten, S., Langner, K-D., Ringer, R., Fleckenstein, B., and

Doerfler, W. (1988) *Virology* **166**, 166-174.

Kratzer, W., Schughart, K., and Fritz, H-J. (1982) *Nucleic Acids Res.* **10**, 6475-7485.

Kruczek, I. and Doerfler, W. (1982) *EMBO J.* **1**, 409-414.

Kruczek, I. and Doerfler, W. (1983) *Proc. Natl. Acad. Sci. USA.* **80**, 7586-7590.

Kunnath, L. and Locker, J. (1983) *EMBO J.* **2**, 317-324.

Landolfi, N.F., Capra, D.J., and Tucker, P.W. (1986) *Nature* **323**, 548-551.

Langer, K.D., Vardimon, L., Renz, D., and Doerfler, W. (1984) *Proc. Natl. Acad. Sci. USA.* **91**, 2950-2954.

Langer, K.D., Weyer, U., and Doerfler, W. (1986) *Proc. Natl. Acad. Sci. USA.* **83**, 1598-1602.

Larson, A. and Weintraub, H. (1982) *Cell* **29**, 609-622.

Lee, W., Haslinger, A., Karin, M., and Tjian, R. (1987) *Nature* **325**, 368-372.

Levinson, B., Khoury, G., Vandewonde, G., and Gruss, P. (1982) *Nature* **295**, 568-572.

Lichtsteiner, S., Wuarin, J., and Schibler, U. (1987) *Cell* **51**, 963-973.

Lin, S., Lin, D., and Riggs, A.D. (1976) *Nucleic Acids Res.* **3**, 2183-2191.

Lin, S. and Riggs, A.D. (1972) *Proc. Natl. Acad. Sci. Wash.* **69**, 2574-2576.

Liskay, R.M. and Evans, R.J. (1980) *Proc. Natl. Acad Sci. USA.* **77**, 4895-4898.

Lock, L.F., Takagi, N., and Martin, G.R. (1987) *Cell* **48**, 39-46.

Mandel, J.L. and Chambon, P. (1979) *Nucleic Acids Res.* **7**, 2081-2103.

Manley, J.L., Fire, A., Cano, A., Sharp, P.A., and Gefter, M.L. (1980) *Proc. Natl. Acad. Sci. USA.* **77**, 3855-3859.

Marsh, J.L., Erfle, M., and Wykes, E.J. (1984) *Gene* **32**, 481-485.

Maxam, A. and Gilbert, W. (1980) *Methods enzymol.* **65**, 499-560.

McGhee, J.D. and Ginder, G.D. (1979) *Nature* **280**, 419-420.

McKeon, C., Ohkubo, H., Pastan, I., and deCrombrugghe, B. (1982) *Cell* **29**, 203-210.

Mermod, N., Williams, T.J., and Tjian, R. (1988) *Nature* **332**, 557-561.

Mercurio, F. and Karin, M. (1989) *EMBO J.* **8** (5), 1455-1460.

Miller, J., McLachlan, A.D., and Klug, A. (1985) *EMBO J.* **4**, 1609-1614.

Miller, R.H. and Robinson, W.S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2534-2538.

Mitchell, P.J., Wang, C., and Tjian, R. (1987) *Cell* **50**, 847-861.

Mohandas, T., Sparkes, R.S., and Shapiro, L.J. (1981) *Science* **211**, 393-396.

Monk, M. (1986) *Bio Essays* **4** (5), 204-208.

Muller, U. and Doerfler, W. (1987) *J. Virol.* **61**(12), 3710-3720.

Murchie, A.I.H. and Lilley, D.M.J. (1989) *J.Mol. Biol.* **205**, 593-602.

Murray, E.J. and Grosveld, F. (1987) *EMBO J.* **6** (8), 2329-2335.

Nick, H., Bowen, B., Ferl, R.J., and Gilbert, W. (1986) *Nature* **319**, 243-246.

Ott, M-O., Sperling, L., Cassio, D., Levilliers, J., Sala-Trepat, J., and Weiss, M.C. (1982)
Cell **30**, 825-833.

Payvar, F., DeFranco, D., Firestone, G.L., Edgar, B., Wrange, O., Okret, S., Gustafsson,
J.A., and Yamamoto, K.R. (1983) *Cell* **35**, 381-392.

Pelham, H.R.B. (1982) *Cell* **30**, 517-528.

Pollack, Y., Kasir, J., Shemer, T., Metzger, S., and Szyf, M. (1984) *Nucleic Acids Res.*
12, 4811-4824.

Raychaudhuri, P., Bagchi, S. and Nevins, J.R. (1989) *Genes & Development.* **3**, 620-627.

Razin, A. (1978) In: *The Single Stranded Phages* Denhardt,D.T., Dressler, D. and Ray D.S.,
Eds. (Cold Spring Harbour Laboratory, Cold Spring Harbour. N.Y.)

Razin, A. and Cedar, H. (1977) *Proc. Natl. Acad. Sci. USA* . **74**, 2725-2728.

Razin, A. and Riggs, A.D. (1980) *Science* **210**, 604-609.

Razin, A., Szyf, M., Kafri, T., Giloh, H., Scarpa, S., Carotti, D., and Cantoni, G.L. (1986)
Proc. Natl. Acad. Sci. USA. **83**, 2827-2831.

Reinberg, D. and Roeder, R.G. (1987) *J. Biol. Chem.* **262** (7), 3310-3321.

Reinberg, D., Horikoshi, M., and Roeder, R.G. (1987) *J. Biol. Chem.* **262** (7),
3322-3330.

Riggs, A.D. (1975) *Cytogenet. Cell. Genet.* **14**, 9-25.

Riggs, A.D. and Jones, P.A. (1983) *Adv. Cancer Res.* **40**, 1-30.

Roy, P.H. and Weissbach, A. (1975) *Nucleic Acids Res.* **2** (10), 1669-1684.

Royer, H-D. and Sager, R. (1979) *Proc. Natl. Acad. Sci. USA.* **76**, 5794-5798.

Saluz, H.P., Feavers, I.M., Jiricny, J., and Jost, J.P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6697-6700.

Saluz, H.P., Jiricny, J. and Jost, J.P. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7167-7171.

Saluz, H-P. and Jost. J.P. (1989) *Nature* **338**, 277.

Sanford, K.K., Earle, W.R., and likely, G.D. (1984) *J. Natl. Cancer Inst.* **9**, 229-246.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.

Santi, D.V., Garrett, C.E., and Barr, P.J. (1983) *Cell* **33**, 9-10.

Santi, D.V., Norment, A., and Garrett, C.E. (1984) *Proc. Natl. Acad. Sci. USA.* **81**, 6993-6997.

Santoro, C., Mermod, N., Andrews, P.C., and Tjian,R. (1988) *Nature* **334**, 318-224.

Schaffner, W. (1989) *Trends in Genetics* **5** (2), 37-39.

Scheidereit, C., Geisse, S., Westphal, H.M., and Beato, M. (1983) *Nature* **304**, 749-752.

Schreck, R.R., Erlanger, B.F., and Muller, O.J. (1977) *Exp. Cell. Res.* **108**, 403-411.

Seed, B. and Sheen, J-Y. (1988) *Gene* **67**, 271-277.

Sellem, C.H., Weiss, M.C., and Cassio, D. (1985) *J. Mol. Biol.* **181**, 363-371.

Serfling, E. (1989) *Trends in Genetics* **5** (5), 131-133.

Shen, C-K.J. and Maniatis, T. (1980) *Proc. Natl. Acad. Sci. USA.* **77**, 6634-6638.

Shimada, T., Inokuchi, K., and Nienhuis, A.W. (1987) *Mol. Cell. Biol.* **7** (8), 2830-2837.

Simon, D., Grunert, F., Acken, U.V., Doring, H.P., and Kroger, H. (1978), *Nucleic Acids Res.* **5** (6), 2153-2167.

Southern, E.M. (1975) *J. Mol. Biol.* **98**, 503-517.

Sorger, P.K., Lewis, M.J., and Pelham, H.R.B. (1987) *Nature* **329**, 81-84.

Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A., and Baltimore, D. (1986) *Nature*
        323, 640-643.

Stein, R., Scaiky-Gallili, N., Razin, A., and Cedar, H. (1983) *Proc. Natl. Acad. Sci. USA*
        80, 2422-2426.

Stein, R., Razin, A., and Cedar, H. (1982) *Proc. Natl.Acad. Sci. USA* . **79**, 3418-3422.

Stewart, C.L., Stahlman, H., Jahner,D., and Jaenisch, R. (1982) *Proc. Natl. Acad. Sci. USA*
        79, 4098-4102.

Stuhmann, H., Jahner, D., and Jaenisch, R. (1981) *Cell* **26**, 221-232.

Subramanian, K.N. (1982) *Nucleic Acids Res.* **10**, 3475-3486.

Sutter, D. and Doerfler, W. (1980) *Proc. Natl. Acad. Sci. USA.* **77**, 253-256.

Sweet, B.H. and Hilleman, M.R. (1960) *Proc. Soc. Exp. Boil. Med.* **105**, 420-425.

Taylor, S.M. and Jones, P.A. (1982) *J. Mol. Biol.* **162**, 679-692.

Theiss, G., Schleicher, R., and Schimpff-Weiland, G., and Follman, H. (1987) *Eur. J.*
        Biochem. **167**, 89-96.

Thompson, J.P., Granoff, A., and Willis, D.B. (1986) *Proc. Natl. Acad. Sci. USA* **83**,
        7688-7692.

Thompson, J.P., Granoff, A., and Wilis, D.B. (1988) *J. Virol.* **62** (12),4680-4685.

Tjian, R., Briggs, M., Mitchell,P., Lee, W., Kadonaga, J., and Jones, K. (1987) In the 19th
        Miami Winter Symp. on Adv. in Gene Technology: *The Mol. Biol. Devel.* Vol **4**

Turnbull, J.F. and Adams, R.L.P. (1976) *Nucleic Acids Res.* **3** (3), 677-695

Twigg, A.J. and Sherratt, D. (1980) *Nature* **283**, 261-218.

Ullrich, A., Dull, T.J., Gray, A., Philips, J.A., and Peter, S. (1982) *Nucleic Acid Res.* **10**,
        2225-2240.

Urieli-Shovel, S., Gruenbuam, Y., Sedat, J., and Razin, A. (1982) *FEBS Letts.* **146,** 148-152.

Vardimon, L., Neumann, R., Kuhlman, I., Sutter, D., and Doerfler, W. (1980) *Nucleic Acids Res.* **8,** 2461-2473.

Waalwijk, C. and Flavell, R.A. (1978) *Nucleic Acids Res.* **5,** 4631-4641.

Waechter, D.E. and Baserga, R. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1106-1110.

Watt, F. and Molloy, P.L. (1988) *Genes & Development* **2,** 1136-1143.

Weintraub, H., Larson, A., and Groudine, M (1981) *Cell* **24,** 333-334.

Weisshaar, B., Langer, K-D., Jutterman, r., Muller, U., Zock, C., Klimkait, T., and Doerfler, W. (1988) *J. Mol. Biol.* **202,** 225-270.

Westin, G. and Schaffner, W. (1988), *Nucleic Acids Res.* **16** (13), 5771-5782.

Wiebauer, K. and Jiricny, J. (1989) *Nature* **339,** 234-236.

Wigler, M., Levy, D., and Perucho, M. (1981) *Cell* **24,** 33-40.

Wildeman, A.G., Zenke, M., Schatz, C., Takahashi, K., Barrera-Saldana, H.,Vigneron, M., and Chambon, P. (1985) In: *Current Communications in Molecular Biology, Eukaryotic Transcription,* Ed: Y. Gluzman. pp.19-26

Willis, D.B. and Granoff, A. (1980) *Virology* **107,** 250-257.

Wyatt, G.R. (1951) *Biochem. J.* **48,** 584-590.

Yamaguchi, M. and DePamphilis, M.L. (1986) *Proc. Natl. Acad. Sci. USA .* **83,** 1646-1650.

Yesufu, H.M.I., Rinaldi, A., and Adams, R.L.P. (1989) *Eur. J. Biochem.* In Press.

Yisreali, J., Adelstein, R.S., Melloni, D., Nudel,U., Yaffe, D., and Cedar, H. (1986) *Cell* **46,** 409-415.

Yisraeli, J. and Szyf, M. (1984). In: *DNA Methylation* S. Razin, H. Cedar and A.D. Riggs eds (New York : Springer-Verlag), pp 352-378.

Yuan, R. and Meselson, M. (1970) *Proc. Natl. Acad. Sci. USA.* **65**, 357-362.

Zenke, M., Grundstrom, T., Matthes, H., Winzerith, M., Schatz, C., Wildeman, A., and

Chambon, P. (1986) *EMBO J.* **5**, 387-397.

Zoller, V. and Smith, M. (1983) *Meths. Enzym.* **100**, 468-500.