DNA Methylation, Transcription and Chromatin Assembly *In Vitro*

Colin Anfimov Johnson

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

November 1995

Division of Biochemistry and Molecular Biology
Institute of Biomedical and Life Sciences
The University of Glasgow
Glasgow, G12 8QQ
Scotland
United Kingdom

Abstract

Histone H1 is implicated in the establishment of stable and tissue-specific gene repression. It is presumed to repress transcription by binding to the internucleosomal, linker DNA leading to chromatin compaction and the formation of the 30 nm chromatin fibre. Histone H1 is abundant in heterochromatin and is associated with nucleosomes containing 5-methylcytosine. Conversely, it is absent from CpG island chromatin which is characteristic of active autosomal housekeeping genes and it is depleted in chromatin that contains active genes. DNA methylation correlates with the inactivity of many genes in vertebrates. It has been proposed that methylation may direct the formation of an inactive chromatin structure that is inaccessible to transcription factors, in a process requiring the participation of proteins that bind preferentially to methylated DNA. This study is an attempt to understand the molecular mechanisms that underlie the repression of gene activity by a combination of DNA methylation and chromatin. *In vitro* systems for transcription and the formation of chromatin are used as simplified models.

The extent of *in vitro* transcription from unmethylated or methylated template is assayed in the presence of varied levels of total histone H1. Two templates are used: plasmid pArg/Leu contains two tRNA genes, which are transcribed by RNA polymerase III (pol III), and plasmid pVHCk contains the SV40 promoter linked to a reporter gene, which is transcribed by RNA polymerase II (pol II). A nuclear extract of HeLa cells is used for *in vitro* transcription assays of both types of template. Histone H1 forms characterised complexes on DNA, under the conditions used for these studies. Histone H1-DNA complexes are presumed to be a valid model of inactive chromatin.

Transcriptional inactivation by histone H1 is effective at lower levels with methylated templates, in comparison with unmethylated templates. Complete inactivation of all types of template is obtained with a further increase in histone H1 levels. Different somatic variants of histone H1 show differing degrees of preferential inhibition. Furthermore, histone H1 is a contaminant of the nuclear extract. The extract can be depleted of endogenous histone H1 either by the addition of competitor DNA or by fractionation of the extract with ammonium sulphate. Both treatments increase the level of

transcription from the methylated pol III template to that of the unmethylated template. The effect is reversed by the addition of exogenous histone H1 to the pol III template. The preferential inhibition of transcription from methylated templates by histone H1 does not appear to be due to a greater binding affinity of the protein to methylated DNA, in comparison to unmethylated DNA. Instead, the conformation of the complex between histone H1 and methylated DNA is changed, which prevents the formation of initiation and elongation transcription complexes on the methylated pol III template.

Endogenous histone H1 in the nuclear extract therefore prevents fully methylated pol III and pol II templates from being transcribed as efficiently as the unmethylated templates. This effect is most obvious when only the promoter region of the pol II template is methylated. Fully methylated DNA is intrinsically resistant to limited digestion with the restriction enzyme *Msp* I, in comparison to unmethylated DNA. This differential effect of DNA methylation is also observed when these templates are reconstituted as chromatin using *Xenopus* S150 egg extract. Chromatin reconstituted on fully methylated or regionally-methylated DNA shows a greater resistance to digestion with *Msp*I than chromatin reconstituted on unmethylated DNA. The preferential resistance to *Msp*I, which is enhanced by the addition of histone H1 during the chromatin reconstitution, occurs even on regions of unmethylated DNA if another region of that DNA is patch-methylated prior to chromatin reconstitution. This is consistent with DNA methylation acting as a focus for the formation of inactive chromatin. The transcriptional activity of unmethylated, patch-methylated and fully methylated pol II templates supports these observations.

INDEX

Acknowledgements

I would like to express my gratitude to Drs. Roger Adams and John Goddard for their supervision of this work, and their consistently helpful and friendly advice throughout the three years of this study. More recently, their critical advice and encouragement were invaluable during the writing of this thesis.

The technical assistance and wit of Tom Carr and Heather Lindsay are greatly appreciated. Toads were kindly provided by Dr. Gwyn Gould, and both he and other members of his lab gave helpful advice on handling toads. Histone H1 variants and several oligonucleotides were kindly provided by Prof. Paola Caiafa and colleagues.

Past and present members of the lab are all thanked for their many helpful suggestions and for making the workplace so friendly. These people are: Ali Alloueche, Bill Caspary, Gavin Collett, Michele Cummings, Stef Kass, Margo Murphy and Sriharsa Pradhan.

I wish to thank the following friends in Glasgow and elsewhere for their many instances of kindness, good humour and entertaining lapses into sadness: Paul Allcock, Diane Alldrit, Florence Banales, Belén Calvo, Ritchard Cook, Nils Eckhardt, Will Goodwin, Saiqa Khan, Tino Krell, Heidrun Krüger, Callum Livingstone, Jo Long, Sally Martin, Tim Martin, Alirio and Anneka Mélendez, Gillian Muir, Alastair McCullough, Lisa Porter, Lipika and Mihika Pradhan, Phil Raines, Ian Ramsey, Tim Sawbridge, Graeme Thompson, Amanda Walters, Phillip Walters, Ken and Paula Welch.

In particular, I would like to thank my parents for their support and encouragement over the years. This thesis is dedicated to them, for this only the beginning of the thanks that I shall always give them.

I acknowledge financial support from the Medical Research Council.

Цветок

Цветок засохший безуханный,
Забытый в книге вижу я;
И вот уже мечтою странной
Душа наполнилась моя:

Где цвел? когда? какой весною?
И долго ль цвел? и сорван кем,
Чужой, знакомой ли рукою?
И положен сюда зачем?

На память нежного ль свиданья,
Или разлуки роковой,
Иль одинокого гуляпья
В тиши полей, в тени лесной?

И жив ли тот, и та жива ли?
И нынче где их уголок?
Или уже они увяли,
Как сей неведомый цветок?

А. С. Пушкин

CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

Abbreviations

The abbreviations used in this thesis are in agreement with the recommendations of the editors of the Biochemical Journal (*Biochem. J.* (1995) **305**, 1-15) with the following additions:

aprt	adenine phosphoribosyltransferase
CAT	chloramphenicol acetyltransferase
DEPC	diethylpyrocarbonate
dsDNA	double-stranded DNA
5-azaC	5-azacytidine
5-mC	5-methylcytosine
g6pd	glucose-6-phosphate dehydrogenase
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
hprt	hypoxanthine-guanine phosphoribosyltransferase
MDBP	methylated DNA-binding protein
MeCP	methyl CpG-binding protein
MTase	methyltransferase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pgk	phosphoglycerate kinase
PMSF	phenylmethylsulphonylfluoride
pol II	RNA polymerase II
pol III	RNA polymerase III
SAM	*S* -adenosyl-L-methionine
ssDNA	single-stranded DNA
SV40	simian virus 40
tk	thymidine kinase

CHAPTER ONE

Introduction

1.1 General introduction

The massive size of the eukaryotic genome requires the DNA in the nucleus to be compacted into a stable but readily accessible form (for a recent general review see Paranjape et al. 1994), so that the activities of appropriate genes are coordinated both spatially and temporally. The human genome of 3×10^9 bp contains at least 50,000 genes and yet is compacted into a nucleus of only 10^{-5} m diameter. This compaction is achieved by the association of DNA into the nucleosome, the fundamental repeating unit of chromatin, and supranucleosomal structures. The most widely accepted model of chromosome structure is that DNA is wrapped around a core histone octamer to form a 10 nm chain of nucleosomes, which is then coiled into a 30 nm chromatin fibre or solenoid (see Section 1.8.5). The chromatin fibre is then folded into 75 kbp average size loops tethered to a protein scaffold. The folded chromatin fibre is about 250 nm wide and is further coiled into the 700 nm chromatid of a metaphase chromosome, which is the most tightly compacted chromosome state (reviewed in Wolffe 1992). Genes in chromatin tend be to repressed or inactive because the closed, tightly-packed structure of the chromatin will limit the accessibility of diffusible *trans* -acting factors to these genes. *Cis* -acting elements adjacent to genes, such as promoter and enhancer elements, can therefore influence the role of *trans* -acting factors. An additional *cis* -acting element may be the covalent modification of cytosine residues in DNA to 5-methylcytosine (Adams 1990b). In mammals, recent evidence has suggested that DNA methylation has an important role in the processes of gene regulation, imprinting, oncogenesis and inherited diseases such as fragile X syndrome (see Sections 1.4, 1.7–1.11).

In particular, current opinion is that methylation can act as a signal for the global repression of gene activity. Patterns of DNA methylation at cytosine residues have been proposed to be a *cis*-acting element of vertebrate gene regulation during cell differentiation (Eden and Cedar 1994). Methylation correlates with the inactivity of many tissue-specific genes (Busslinger et al. 1983; Doerfler 1983; Hergersberg 1991; Razin and Cedar 1991; Bird 1992) and exerts this inhibitory *cis* -acting effect on *trans* -acting factors. Methyl groups in a promoter can interfere with the binding of transcription factors (Watt and Molloy 1988; Comb and Goodman 1990; Ehrlich and Ehrlich 1993), but in addition, DNA methylation can also direct the formation of an inactive chromatin structure that is inaccessible to transcription factors (Lewis and Bird 1991; Graessmann and Graessmann 1993). An attractive hypothesis is that DNA methylation can inactivate chromatin by providing targets for the recently characterised methyl-CpG binding proteins and methylated DNA binding proteins. The importance of DNA methylation is obvious from the dramatic demonstration that mouse DNA methyltransferase knockout mutations are lethal during early development (Li et al. 1992). The biological significance of DNA methylation will be discussed in subsequent sections.

1.2 DNA methylation

1.2.1 Introduction

In addition to the four major bases, the DNA of most organisms contains minor bases that are formed by the covalent modification of the major ones. The most common minor base of higher eukaryotes is 5-methylcytosine (Fig. 1.1), which is formed by the activity of the enzyme DNA-(cytosine-5) methyltransferase. The DNA of prokaryotes contains the minor bases N^6-methyladenine (6-mA) and N^4-methylcytosine (4-mC), in addition to 5-methylcytosine (5-mC), in varying amounts (Doerfler 1983; Barras and Marinus 1989 and references therein). Since 6-mA has been reported in few eukaryotes, and 4-mC in none, this study will be concerned only with 5-methylcytosine. The term DNA methylation will therefore refer exclusively to C5 cytosine methylation, unless otherwise stated.

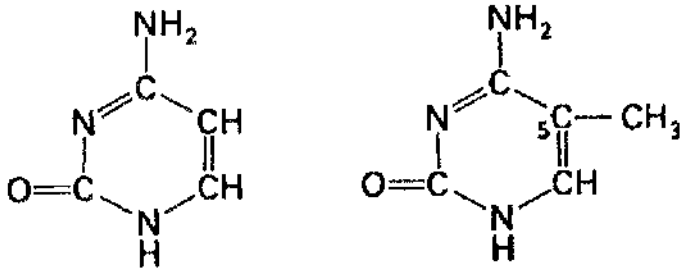


Figure 1.1

Structure formulae of cytosine and 5-methylcytosine

1.2.2 Methods to study DNA methylation

The base 5-methylcytosine was first identified as a component of bacterial nucleic acid in 1925 (Johnson and Coghill 1925). It was the first modified base to be discovered in DNA, by using the then newly developed technique of paper chromatography (Hotchkiss 1948). In the intervening decades, the molecular biology of nucleic acids has made remarkable progress as new techniques and equipment have been created and refined. In particular, the ideas and data on DNA methylation have been driven by the available techniques, which have lead to break-throughs in the understanding of bacterial restriction-modification systems. However, the techniques that are applicable to eukaryotic systems have, in general, not been sufficient to offer clear insights. The role of methylation in the control of gene expression, genomic imprinting, replication and development remains unclear. This is due, in part, to the use of sequence-unspecific methods, some of which use methylation-sensitive restriction enzymes. These techniques can provide a gross correlation between methylation status and biological effect, but cannot locate individual modified bases within a sequence. In addition, the flux of methylation status during development cannot be studied. This problem is overcome by using sequence-specific methods. Some of these techniques will be discussed in the following sections.

1.2.3 Sequence-unspecific methods for studying 5-methylcytosine distribution in DNA

The 5-mC content of DNA has been quantified after hydrolysis by chemical or enzymatic means. Fractionation procedures such as electrophoresis, thin-layer chromatography (TLC) and high performance liquid chromatography (HPLC) have been used to separate 5-mC after total acid hydrolysis of DNA (Adams and Burdon 1985; Saluz and Jost 1993), or after degradation by enzymatic hydrolysis. Enzymatic cleavage products are mononucleotides with either 5' phosphates (in the case of digestion with pancreatic DNaseI) or 3' phosphorylated mononucleotides (with micrococcal nuclease). A modification of nearest-neighbour analysis has allowed the detection of 5-mC (Adams and Burdon 1985). Nicks were introduced into the DNA using DNaseI, the nucleotide 3' to the nicked base was labelled using [α-^{32}P]dNTP and the DNA was hydrolysed to 3' monophosphates which were fractionated by thin-layer chromatography. This analysis quantifies the relative proportions of methylation at each of the four CpN dinucleotides (Gruenbaum et al. 1981; Pollack et al. 1984).

The close relationship between DNA methylation and DNA restriction (reviewed in Noyer-Weidner and Trautner 1993) has made restriction endonucleases important tools in the analysis of methylation patterns (Bird and Southern 1978). Several isoschizomeric restriction endonucleases are available. Isoschizomeric enzymes recognise the same sequence, but have differential sensitivity to the presence of methylated bases within that recognition sequence. For example, *Hpa* II will cleave the unmethylated recognition sequence CCGG, but is refractory to C^{m}CGG. The isoschizomer *Msp* I is insensitive to the methylation of the internal cytosine (it will cleave both CCGG and C^{m}CGG) but is refractory to ^{m}CCGG. Comparison of the digestion patterns with isoschizomeric enzymes can therefore reveal the methylation status of genomic DNA (reviewed by Adams and Burdon 1985; Saluz and Jost 1993). For example, the distribution of CpG dinucleotides has been estimated in DNA from mouse liver, as well as the proportion of methylated CpGs (Singer et al. 1979).

1.2.4 Sequence-specific methods

Sequencing methodologies can locate all the 5-methylcytosines on uncloned or genomic DNA in a sequence- and strand-specific manner (Saluz and Jost 1993; Grigg and Clark 1994). This has provided valuable insights into *de novo* methylation during embryogenesis and subsequent demethylation during cell differentiation (reviewed in Jost and Saluz 1993; discussed in Section 1.9). The original method as devised by Church and Gilbert (1984) involves digestion of total genomic DNA with restriction enzymes which provides fragments of defined length containing the target sequence. The DNA is then subject to chemical sequencing reactions (Maxam and Gilbert 1980), and the cleavage products are then resolved on a sequencing gel and transferred onto a nylon filter. Southern hybridisation using the appropriate probe reveals the sequencing ladder of the target sequence. Methylated cytosines are represented as gaps in the ladder as hydrazine does not react with 5-mC in the C specific reaction. One improvement on the original procedure has been the linear PCR amplification of the cleavage products with labelled primers (Saluz and Jost 1989). A new method for genomic sequencing is the bisulphite method (Frommer et al. 1992), in which bisulphite is used to specifically deaminate cytosine, but not 5-methylcytosine, to uracil. The target sequence is then amplified by PCR and subject to dideoxy (Sanger) sequencing. Upon sequencing of the amplified DNA, all the thymine and uracil residues become detectable as thymine and only 5-mC as cytosine.

1.3 Distribution of 5-methylcytosine in higher eukaryotes

The percentage of 5-mC in eukaryotic DNA varies over a wide range (reviewed in Doerfler 1983; Adams and Burdon 1985), and it has been suggested that the genome size of an organism is an important factor that determines the level of cytosine methylation (Antequera and Bird 1993). Higher eukaryotes with large, complex genomes will retain DNA methylation as a universal regulatory mechanism, despite selective pressure for it to

be removed (see Section 1.5). In vertebrates 3–8% of DNA cytosine is methylated, which represents about 5 x 10^7 5-mC residues per diploid nucleus. In higher plants as many as 30% of the total cytosines are methylated. By contrast, 5-mC is present at low levels, or is absent, in the DNA of lower eukaryotes with smaller genomes. The DNAs of the fruit-fly *Drosophila*, the nematode *Caenorhabditis* and the yeast *Saccharomyces* are exceptional in that they do not contain any detectable 5-methylcytosine (reviewed in Doerfler 1983). In fungi (Rothnie et al. 1991) or echinodermata (Bird et al. 1979), only repetitive satellite sequences appear to be methylated. An interesting parallel observation is that repetitive DNA in mammals is also enriched in 5-mC. For example, major satellite DNA contains over 40% of all 5-mC in the mouse genome (Miller et al. 1974). However, DNA methylation also occurs in the non-repetitive DNA of vertebrates and higher plants, where it can be a *cis* -acting element on gene expression.

In the vertebrate genome, almost all 5-mC residues are restricted to the symmetrical dinucleotide CpG. In mammals 60–70% of all CpGs are methylated (Antequera and Bird 1993). In higher plants, 5-mC occurs at both CpG dinucleotides and CpNpG trinucleotides, where N is any nucleotide, so that the methylation again occurs at a symmetrical sequence (Gruenbaum et al. 1981; Belanger and Hepburn 1990). The CG and the CNG methyltransferase activities of the pea *Pisum sativum* have been fractionated into two distinct species (Pradhan and Adams 1995), but the existence of CpNpG methylation in mammalian genomes remains controversial. Genomic sequencing fails to detect significant levels of CpNpG methylation, although a nearest-neighbour analysis has suggested that 5-mC was not located exclusively at CpG dinucleotides (Woodcock et al. 1987). CNG methyltransferase activity has been observed only by using the bisulphite sequencing of plasmid DNA that has been transfected and stably integrated into mammalian cells (Clark et al. 1995).

1.4 DNA methylation, mutation and oncogenesis

Cytosine methylation is inherently unstable and mutagenic, as first demonstrated in bacteria (Coulondre et al. 1978). This is because 5-methylcytosine can be converted to

thymine by spontaneous oxidative deamination. A methyl-CpG dinucleotide therefore mutates into TpG, which gives rise to a T-G mismatch of base-pairs. An enzymatic repair system that corrects the mismatch in favour of the original cytosine does exist, but appears not to be totally efficient (Wiebauer et al. 1993). The failure of the repair mechanism results in C-G to T-A transition mutations, and 5-methylcytosines represent mutational hot spots in the genome. It has been calculated that there are about 8 deamination events per day per human genome (Jones et al. 1992). Methylated cytosines are therefore a mutational load for higher eukaryotes, but this disadvantage appears to be outweighed by the advantages that methylation confers on the organism. The primary advantage may be that methylation can separate the genome into compartments, each of which has autonomous control over gene expression (discussed in Section 1.8).

Over the course of millions of years, selective pressure has tended to delete CpG sequences from vertebrate genomes. CpGs are therefore comparatively rare and occur at one-fifth of the expected frequency in bulk vertebrate DNA (Bird 1986). The suppression of CpGs is matched by a proportional excess of TpG and CpA dinucleotides. In contrast, GpC dinucleotides are not suppressed and are distributed evenly throughout the genome. C-G to T-A transition mutations account for 30–40% of all point mutations (Cooper and Youssoufian 1988) and restriction fragment length polymorphisms (Barker et al. 1984). In addition, these mutations occur preferentially at CpG dinucleotides. Point mutations in the *p53* tumour suppressor gene have been found in over half of all human tumours, which suggests that such mutations have a role in oncogenesis (Jones et al. 1992). Transition mutations of C-G base-pairs at CpG sequences predominate in colon cancer, and it has been directly demonstrated that some of these CpGs are methylated by genomic sequencing (Rideout et al. 1990).

1.5 DNA Methyltransferases

DNA is methylated in an early post-replicative step in eukaryotes, with the newly synthesised strand becoming methylated. 5-methylcytosine is formed by the activity of the enzyme DNA-(cytosine-5) methyltransferase (MTase) on the 5-carbon of the pyrimidine

ring of cytosine. The DNA of prokaryotes contains, in addition to 5-methylcytosine (5-mC), the minor bases *N* ^{6}-methyladenine (6-mA) and *N* ^{4}-methylcytosine (4-mC) that are formed by the modification of exocyclic nitrogens. These two classes of exocyclic methyltransferases are closely related and have been isolated only from prokaryotes. The third class of MTase, DNA-(cytosine-5) methyltransferase, is found in both eukaryotes and prokaryotes. The methyl group donor in the catalytic process is *S* -adenosyl-L-methionine (SAM; also known as AdoMet in some literature) for all classes of enzyme.

1.5.1 Prokaryotic DNA methyltransferases

In prokaryotes, DNA methylation is involved in restriction-modification systems, but it also plays a role in regulating the initiation of DNA replication and the correction of errors immediately following replication (Modrich 1991; Noyer-Weidner and Trautner 1993). In restriction-modification systems, the modification by a methyltransferase prevents endonuclease attack by the cognate restriction enzyme (reviewed in Wilson and Murray 1991). These systems serve as protection against bacteriophage infection. There are three classes of MTase involved in these systems: types I, II, and III. Types I and III are multi-subunit enzymes in which the restriction and modification activities reside in different subunits and, in type I enzymes, yet another subunit is responsible for determining target sequence specificity. All members of types I and III are 6-mA methyltransferases. The type II MTases are separate proteins from their cognate restriction endonucleases. Members of this class can be 5-mC, 4-mC or 6-mA MTases. An unusual prokaryotic enzyme is *Sss* I methyltransferase (M.*Sss* I), which was detected in the mycoplasma *Spiroplasma sp.* (Nur et al. 1985). It methylates CpG dinucleotides and is therefore an isomethylomer of eukaryotic methyltransferases. (An isomethylomer is analogous to an isoschizomer in that it is an enzyme from a different source, but with the same specificity).

The mechanism of methyl transfer has been determined for the type II *Hha* I MTase (Klimasauskas et al. 1994). The enzyme catalyses a dramatic alteration in the structure of the DNA helix at the site of the target cytosine. The C-G base pair is disrupted and the cytosine is flipped completely out of the double helix, which represents a hitherto

unknown mode of protein-DNA interaction (reviewed in Cheng 1995). The cytosine is then positioned in the active site of the enzyme, where a covalent bond is formed between carbon-6 of the pyrimidine substrate of cytosine and a cysteine of a conserved Pro-Cys dipeptide (Wu and Santi 1987). This activates the 5-position of the cytosine, which now accepts a methyl group from the AdoMet cofactor. This catalytic mechanism is presumed to be common to all DNA (cytosine-5) methyltransferases, including eukaryotic MTases (reviewed in Cheng 1995), since the Pro-Cys dipeptide is highly conserved. In addition, all the known eukaryotic MTases (see Section 1.5.2) have a carboxyl-terminal domain that is closely related to the much smaller prokaryotic enzymes.

1.5.2 Eukaryotic methyltransferases

Unlike their prokaryotic counterparts, which do not discriminate between unmethylated and hemi-methylated targets, the eukaryotic DNA (cytosine-5) MTases preferentially methylate hemi-methylated substrates (Gruenbaum et al. 1982). This reflects the different biological role that MTases play in eukaryotes. There appears to be only one MTase in mammals, which can catalyse both maintenance and *de novo* methylation (Leonhardt and Bestor 1993). The maintenance activity predominates because of the preference for hemi-methylated sites over unmethylated sites. The cDNAs for the murine (Bestor et al. 1988), human (Yen et al. 1992), *Arabidopsis* (Finnegan and Dennis 1993) and *Pisum sativum* (Pradhan and Cummings, 1995; personal communication) enzymes have been cloned. The open reading frames for the vertebrate enzymes indicate a mass for the primary translation product of 170 kDa, although the apparent mass of MTase in normal somatic tissues and proliferating cell types is about 190 kDa (Adams 1990b). In higher plants both CpG and CpNpG sequences are methylated, which has lead to speculation that there are separate CG and CNG methyltransferases. Two such activities have been fractionated from the pea *Pisum sativum* , with masses of 140 kDa for the CG-specific enzyme and 110 kDa for the CNG enzyme (Pradhan and Adams 1995). Both enzymes are believed to arise by proteolytic processing of the initial translation product.

The primary role of mammalian DNA MTase appears to be the maintenance of tissue-specific methylation patterns after DNA replication. The semi-conservative nature

of DNA replication results in the parental strand remaining methylated while the newly synthesised daughter strand is unmethylated. In an early post-replicative step (Burdon and Adams 1980), hemi-methylated CpG sites are recognised and methylated symmetrically on the daughter strand by MTase. This mechanism would enable methylation patterns to be transmitted by clonal inheritance in somatic tissues (Holliday and Pugh 1975). This speculation was verified by the observation that patterns of *in vitro* methylation on transfected DNA could be maintained for many cell generations (Wigler et al. 1981). More recent work has used DNA constructs that are integrated into the genome of transgenic mice. A preimposed pattern of methylation on the transgenic construct could be maintained for up to four generations beyond the founder animal (Lettmann et al. 1991). DNA methyltransferases have been demonstrated to associate specifically with replication foci in mammalian S phase nuclei (Leonhardt et al. 1992). The clonal inheritance of methylation patterns therefore requires the close coordination of replication and methylation, as well as the preference of eukaryotic MTase for hemi-methylated DNA.

New patterns of tissue-specific methylation are established during embryogenesis (Monk 1990) by *de novo* methylation of previously unmethylated DNA (reviewed by Adams et al. 1993). It is assumed that maintenance and *de novo* activities reside on the same protein, but the possibility remains that there are different specific *de novo* MTases expressed during embryogenesis, although there is no evidence for the presence of a gene family in mammals (Leonhardt and Bestor 1993). *De novo* methylation is facilitated by the high levels of MTase present in the egg and early embryo (Monk et al. 1991). The role of *de novo* methylation during embryogenesis will be discussed in more detail in Section 1.9. *De novo* methylation is not just confined to cells of the germline or the early embryo, but can occur in somatic cells. Patterns of methylation can be restored after treatment with 5-azacytidine (see below), an inducer of demethylation (Jones 1984). Many cell lines have *de novo* methylation imposed at the promoters of certain tissue-specific genes, that are not essential under the condition of cell culture (Antequera et al. 1990; Jones et al. 1990). This epimutation (Holliday 1987) leads to the inactivation of these genes, which could explain the loss of cell type-specific functions in culture. This topic is discussed in more detail in subsequent sections.

1.5.3 Inhibition and disruption of eukaryotic methyltransferase function and expression

The importance of DNA methylation in mammals is demonstrated by the use of MTase inhibitors and, more recently, by the disruption of the MTase gene in knock-out experiments. Several drugs that inhibit methylation both *in vivo* and *in vitro* have been identified (reviewed in Adams and Burdon 1985). The most widely used of these are the base analogues 5-azacytidine (5-azaC) and 5-azadeoxycytidine. On incorporation into DNA, 5-azacytosine is a powerful inducer of demethylation (Jones 1984). This is because MTase binds irreversibly to DNA containing 5-azacytosine, which prevents the maintenance of methylation patterns. The demethylating activity of 5-azacytidine has been correlated with selective gene activation for vertebrate cells in culture (reviewed in Razin and Cedar 1991). For example, an inactive endogenous retroviral gene in a chicken cell line is activated when the cells are exposed to 5-azaC (Groudine et al. 1981). Silent genes on the inactive X chromosome (see Section 1.10) can also be activated by the same treatment (Mohandas et al. 1981). However, 5-azaC is toxic at concentrations over 1 μM and leads to a marked change in the metabolism and development of the exposed mammalian cells (Jones 1984). In addition, the genes which are activated by 5-azaC in these studies are in cultured cells, which themselves have aberrant patterns of methylation (Antequera et al. 1990). For example, the *MyoD* gene contains a CpG island (see Section 1.6) that is normally unmethylated at all stages of development in tissues. The gene can be activated by 5-azaC treatment of 10T1/2 cells, but the CpG island is modified by *de novo* methylation (Jones et al. 1990). The significance of experiments that use 5-azaC must be assessed carefully (Bird 1992), since it cannot be proved that the effects of this drug are only exerted by inhibiting MTase activity.

The biological importance of methylation in mammals has been demonstrated by the technique of gene targeting in murine embryonic stem (ES) cells, which allows predetermined mutations in any mouse gene for which clones are available. This technique has been used to construct targeted deletion mutations that disrupt alleles of mouse DNA methyltransferase in ES cells (Li et al. 1992). The modified ES cells have a

third of the wild-type 5-mC level and only 5% of the MTase activity in an *in vitro* assay. Embryos that are homozygous for the mutation complete gastrulation, but are both stunted and developmentally retarded, and die at mid-gestation. The inactivation of the gene is far from complete in the embryo, since 30% of methyltransferase activity remains in mutant mice, but even this is sufficient to perturb patterns of methylation during embryogenesis. An attractive hypothesis to explain this mutant phenotype is that genes that are normally inactivated by DNA methylation are inappropriately activated in the mutant mice. Over-expression of the MTase gene in mouse fibroblasts causes the tumourgenic transformation of these cells (Wu et al. 1993). This provides further evidence that the balanced expression of the MTase gene is a prerequisite for normal embryogenesis.

1.6 CpG islands

As discussed previously, most if not all methylation in the vertebrate genome occurs at the CpG dinucleotide, which has brought about the suppression of this dinucleotide as discussed in Section 1.4. CpGs tend to occur in clusters, called CpG islands, where suppression is absent. CpG islands were identified by the analysis of the cleavage patterns of genomic DNA by methyl-sensitive restriction enzymes (Cooper et al. 1983). These studies showed that CpG islands consisted of about 1–2% of the total genome, and were enriched in clustered, unmethylated *Hpa* II sites which gave rise to so-called *Hpa* II Tiny Fragments (HTFs). The remaining 98% of the genome is the ocean in which the islands are landmarks, and this contains dispersed methylated CpG sites. Vertebrate genomes are therefore divided into compartments of unmethylated islands and methylated oceans. In summary, the characteristics of CpG islands are as follows: they are about 0.5–3 kb in length, lack CpG suppression, have a high G+C content of >60% and are invariably unmethylated, with the certain exceptions discussed below (Antequera and Bird 1993).

All human CpG islands identified so far are associated with genes, and all housekeeping genes have been found to have a CpG island starting 5' of the transcription

unit and covering one or more exons (reviewed in Cross and Bird 1995). In general, CpG islands contain multiple binding sites for transcription factors (Somma et al. 1991). However, some CpG islands have been found to be downstream of the promoter, and even at the 3' end of genes (Shemer et al. 1990). Several highly tissue-specific genes show no evidence of associated CpG islands and remain methylated across the whole length of the gene in non-expressing cells. Other tissue-specific genes can be associated with CpG islands that remain unmethylated even in non-expressing tissues. Examples include the human α-globin gene and the mouse MyoD gene (Antequera and Bird 1993). A strategy to purify genomic restriction fragments that contain unmethylated CpG islands has been devised (Cross et al. 1994), which will allow the isolation of genes that are associated with the islands. The procedure fractionates DNA on the basis of methylation status, by using a methylated DNA binding column.

CpG islands are also distinctive at the level of chromatin (Tazi and Bird 1990). Most noticeable is the presence of nucleosome-free gaps within most CpG islands. The nucleosomes that are formed at a CpG island tend to contain histones H3 and H4 in highly acetylated forms, compared to bulk chromatin (refer to Section 1.8.6). In addition, the levels of histone H1 in CpG island chromatin are reduced. These properties are characteristic of active chromatin (Turner 1991), which reflects the fact that most CpG islands include the promoter of a housekeeping gene.

Although most CpG islands are unmethylated at all times in all tissues, there are some exceptions. One exception is the islands of some genes that have their activity determined by the epigenetic process of genomic imprinting (see Section 1.11). A second exception is the islands of the inactive X chromosome (Riggs and Pfeifer 1992). As mentioned previously and discussed in Section 1.10, the X chromosome of eutherian mammals is inactivated at an early stage of development. The characteristics of the so-called "facultative" heterochromatin that forms are hypermethylation of the DNA and a lack of hyperacetylated histones (Grant and Chapman 1988; Jeppesen and Turner 1993). All the known CpG islands on the inactive X chromosome show a strict correlation between transcriptional inactivity and DNA methylation (Tribioli et al. 1992). Thirdly, CpG islands are *de novo* methylated in cell-lines (Antequera et al. 1990). As has been discussed previously, these aberrant patterns of methylation silence certain tissue-specific

genes that are not essential under the condition of cell culture. At the chromatin level, methylated CpG islands are inactive and closed, as defined by an increase in the resistance to nuclease digestion (Antequera et al. 1990). In some cases, treatment of the cells with 5-azacytidine can activate these genes and lead to the demethylation of the CpG island.

Methylation of a CpG island may be a factor in causing fragile X syndrome, which is the most common heritable form of mental retardation in males (Sutherland and Richards 1995). In affected individuals the expansion of a CCG trinucleotide repeat in an island, and subsequent methylation, may trigger a specific break of the X chromosome near to the methylated CpG island (Oberlé et al. 1991; Ritchie et al. 1994). Three such fragile sites, *FRAXA* , *FRAXE* and *FRAXF* , have been found on the distal long arm of the X chromosome (Sutherland and Richards 1995).

1.7 DNA methylation and transcriptional activity

It is now well established that DNA methylation correlates with inhibition of transcriptional activity of many tissue-specific and housekeeping genes (reviewed in Doerfler 1983; Hergersberg 1991; Razin and Cedar 1991; Eden and Cedar 1994). The same observation has been made both *in vivo*, following the transfection of methylated reporter genes into cells, and in studies that have investigated the *in vitro* transcription of methylated genes. Two mechanisms have been proposed to mediate gene inactivation *in vivo*. In both mechanisms the protrusion of a methyl moiety into the major groove of the DNA helix can alter protein-DNA interactions.

1.7.1 DNA methylation prevents binding of transcription factors

Early studies used gene constructs with regional patterns of methylation (Busslinger et al. 1983; Keshet et al. 1985). to map the sites at which DNA methylation can affect transcription. These sites were shown to be the regulatory regions upstream of the

reporter genes, which suggests that methyl groups at a binding site can prevent the initiation of transcription by blocking the binding of transcription factors because the base sequence is changed (Watt and Molloy 1988). Several transcription factors have been found that do not bind to methylated target sequences (reviewed by Ehrlich and Ehrlich 1993; Tate and Bird 1993). Examples of known proteins include E2F (Kovesdi et al. 1987), the cAMP responsive element binding protein (CREB; Iguchi-Ariga and Schaffner 1989), AP-2 (Comb and Goodman 1990), c-Myc/Myn (Prendergast et al. 1991) and NF-κB (Bednarik et al. 1991), all of which recognise sequences that contain CpG residues. Most of these studies demonstrate a general correlation between DNA methylation, the prevention of specific factor binding and transcriptional inhibition, but their weakness is that the binding studies were performed *in vitro* . All of the factors that appear to be affected by DNA methylation are not specific to one cell type, but methylation may be the *cis*-acting element that gives a degree of cell type-specific gene expression by a direct mechanism of action. It may also suppress basal transcription of tissue-specific genes in non-expressing cells, even though the factors that are required for expression are ubiquitous in many cell types (Bird 1992). This mechanism was first suggested by studies of the liver-specific tyrosine aminotransferase (TAT) gene in rat. Becker et al. (1987) used *in vivo* genomic footprinting techniques to show that part of the TAT promoter can bind ubiquitous factors in hepatocytes, but is unoccupied in non-expressing tissues where it would be expected to bind the same ubiquitous factors. The promoter is unmethylated in expressing cells but is methylated in non-expressing cells, such as fibroblasts. Methylation is implicated in the establishment of cell type-specific gene expression because *in vitro* methylation of the promoter eliminates the *in vitro* footprint. Similar results have been shown for the testis-specific H2B histone gene (Choi and Chae 1991).

In the case of some factors that are part of the transcriptional machinery, methylation does not affect their binding to DNA. Common transcription factors such as Sp1 (Höller et al. 1988; Bryans et al. 1992) and CTF (Ben Hattar et al. 1989) are insensitive to the presence of methyl CpG in their target sites. However, transcription from the promoters that bind these factors can be inhibited by methylation (Ben Hattar et al. 1989; Bryans et al. 1992). Although there is some evidence that, under certain

conditions, DNA methylation may inhibit the processive activity of RNA polymerase (Kobagashi et al. 1990), this does not appear to be the explanation for the inhibition of transcription. A second mechanism that mediates the inactivation of genes by methylation is described in the next section.

1.7.2 Specific binding of proteins to methylated DNA

In addition to a direct effect on factor binding, methylation may cause chromatin to adopt an inactive or closed conformation, which can therefore influence gene accessibility by an indirect effect. The inactivation is presumed to be mediated by the specific binding of repressor proteins to methylated DNA which, in turn, have been proposed to influence the structure of chromatin. Several such factors have been identified in mammals (reviewed in Ehrlich and Ehrlich 1993; Tate and Bird 1993) and are referred to as methylated DNA-binding proteins (MDBPs) or methyl CpG-binding proteins (MeCPs). These proteins include MDBP-1 (Huang et al. 1984), which is a ubiquitous vertebrate protein that binds in a sequence-specific manner (Wang et al. 1986; Ehrlich and Ehrlich 1993), and MDBP-2 (Jost et al. 1991). MDBP-2 requires 30 nucleotides with a single methylated CpG, to which it binds in a sequence non-specific manner (Zhang et al. 1990). MDBP-2 is a repressor of *in vitro* transcription of the avian vitellogenin gene (Pawlak et al. 1991) and, in addition, shares significant sequence homologies with histone H1 (Jost and Hofsteenge 1992). MeCP1, previously known as MeCP (Meehan et al. 1989), has been shown to inactivate methylated genes *in vitro* and *in vivo* , and binds to a minimum of 15 methylated CpGs in >100 bp of DNA (Boyes and Bird 1991; Boyes and Bird 1992). MeCP2 requires only one methylated CpG for binding and is associated with pericentromeric heterochromatin (Lewis et al. 1992; Nan et al. 1993). In one study, MeCP2 does not appear to preferentially inhibit *in vitro* transcription from methylated genes (Meehan et al. 1992), but a recent communication from the same group has reported that MeCP2 can inhibit a methylated gene that is transiently expressed in HeLa cells (Lin et al. 1995). In view of the characteristics of MeCP1 and 2, Bird and co-workers have proposed a model whereby MeCP1 competes with transcription factors to

bind to methylated DNA and guides the DNA to form a heterochromatin structure which, in turn, is maintained by MeCP2 (Meehan et al. 1992).

Other proteins have been characterised that bind preferentially to methylated DNA. The first methylation-specific, sequence non-specific DNA-binding protein to be identified was the plant protein DBP-m (Zhang et al. 1989). A protein that binds to part of the promoter of a number of major histocompatibility class II genes has been shown to have the same DNA-binding specificity as MDBP-1 (Zhang et al. 1993). Although methylation correlates with gene inactivation, a role in activation is possible either by attracting an activator that binds preferentially to methylated DNA or by repelling a methylation-sensitive repressor (Barlow 1993). Factors that fulfil this regulatory role have yet to be identified, but their involvement can explain the surprising observation that regions of the active allele of some imprinted genes is methylated (Efstratiadis 1994; see Section 1.11).

1.8 Chromatin and DNA methylation

There is growing evidence that DNA methylation in organisms with large genomes (vertebrates and higher plants) is accompanied by a change in chromatin conformation (reviewed in Lewis and Bird 1991; Graessmann and Graessmann 1993). This may be a means of dividing the genome into compartments (Lewis and Bird 1991). A small, unmethylated compartment would be accessible to diffusible *trans* -acting factors whereas a large, methylated compartment would be associated with condensed chromatin and be transcriptionally inert. Compartmentalisation can reduce the time required for a *trans*-acting factor to locate a target site in a large genome.

1.8.1 Active and inactive chromatin

Chromatin can exist in two general conformations that affect the degree of accessibility of genes. The open conformation is transcriptionally active in contrast to chromatin in the closed conformation, which is transcriptionally inert (reviewed in Gross and Garrard

1988). It is thought that the open chromatin conformation is a prerequisite for the initiation of transcription, because genes are the most accessible to *trans* -acting factors in this structure (Adams and Workman 1993). In contrast, closed chromatin forms large distinctive segments on a chromosome (Pardue and Hennig 1990). These distinctive regions are called heterochromatin, and are characterised by being transcriptionally inactive, highly condensed and replicated late in S-phase. The paradigm of this structure is the inactive X chromosome of eutherian mammals, which involves the formation of facultative heterochromatin on one copy of the chromosome. This condenses at interphase to form the cytologically distinctive Barr body (Grant and Chapman 1988). The inactive X chromosome is almost completely silent and remains heterochromatic throughout metaphase, which is a characteristic of regions of the genome that play a structural rather than a coding role. Regions that are adjacent to the telomere and the centromere are transcriptionally inert and are associated with so-called constitutive heterochromatin. It remains unknown if the mechanisms for the formation of facultative and constitutive heterochromatin are the same.

The heterogeneous structure of nuclear chromatin can be characterised at the molecular level. Differential sensitivity to a nuclease, such as DNase I, has been used to classify the genome into compartments (reviewed in Gross and Garrard 1988; Wolffe 1992). Regions of the genome that contain active genes show a general sensitivity to nucleases, whereas inactive chromatin is comparatively resistant. The structural changes that underlie these alterations in the conformation of chromatin remain unknown.

It is generally believed that chromatin in a simple closed conformation can inactivate genes by limiting the accessibility of diffusible *trans* -acting factors, such as transcription factors. However, in recent years it has been shown that the *trans* -acting factors can themselves alter the structure of chromatin, and that chromatin can play an additional active role in the regulation of gene expression. The generalised repression of gene activity by closed chromatin can be relieved by *trans* -acting factors, such as enhancer- and promoter-binding factors (Winston and Carlson 1992; Paranjape et al. 1994) which open the chromatin structure. This phenomenon is referred to as "anti-repression", and is distinct from the "true activation" of a gene by specific transcription factors which increase the rate of transcription. It is now obvious that the regulation of

eukaryotic gene expression can only be understood in the context of chromatin, since stretches of naked DNA are probably never found *in vivo*.

1.8.2 Methylation alters the conformation of chromatin

Razin and Cedar (1977) were the first to demonstrate that methylated DNA is in a specific chromatin structure. Mouse L cell nuclei were digested with micrococcal nuclease, and analysis showed that 5-methylcytosine in chromatin was relatively resistant to degradation. After extensive digestion, only 50% of the DNA was solubilised and over 70% of the total 5-mC content of the DNA remained in the unsolubilised fraction. A subsequent study demonstrated that the DNA solubilised at early times of digestion had a reduced 5-mC content, which indicates that linker DNA is hypomethylated compared to nucleosomal core DNA (Solage and Cedar 1978). Both of these studies are weakened by the observation that 5-mC is preferentially associated with regions that are resistant to digestion by staphlococcal nuclease not only in chromatin from human fibroblast nuclei, but also in purified naked fibroblast DNA (Barr et al. 1985). The authors suggest that methylated naked DNA has an intrinsic resistance to digestion by this nuclease, which is of a sufficient magnitude to account for the nuclease resistance of chromatin that contains methylated DNA. This point is discussed further in Chapter 6, which describes the resistance of naked DNA and chromatin to the nuclease *Msp* I.

In the mammalian genome, the regions that are enriched in 5-mC are generally heterochromatic, transcriptionally inert and contain repetitive DNA. Pericentromeric heterochromatin, which contains major satellite DNA in mouse, is preferentially visualised by using an antibody raised against 5-mC (Miller et al. 1974). In addition, a number of studies have shown that methylated CpG islands on the inactive X chromosome are insensitive to *Msp* I in intact nuclei (Wolf and Migeon 1985; Hansen et al. 1988; Antequera et al. 1989). In contrast, CpG islands on the active X chromosome are comparatively nuclease sensitive (Kerem et al. 1983).

The first direct evidence that methylation has an effect on chromatin conformation came from transfection studies of methylated constructs that were integrated into the genome of L cells (Keshet et al. 1986). All unmethylated DNA inserted into the mouse

genome was found to be in a DNase I-sensitive conformation, whereas methylated constructs were nuclease resistant, which is characteristic for closed inactive chromatin. The accessibility of sites in nuclei for CpG enzymes, such as *Msp* I or *Tth* I, depends on their methylation status (Antequera et al. 1989; Levine et al. 1991). After the treatment of bulk chromatin with nuclease, under conditions of complete digestion, it was shown that *Msp* I or *Tth* I were preferentially blocked from cutting methylated sites at specific locations. Unmethylated CpG sites, and restriction sites that did not contain a CpG, were comparatively nuclease sensitive. The formation of closed, nuclease-resistant chromatin correlated with transcriptional inactivation (Levine et al. 1991).

1.8.3 Methylation, chromatin structure and the regulation of gene expression

Two studies have used reporter gene constructs, that contained either the γ-globin gene (Busslinger et al. 1983) or the β-globin gene (Yisraeli et al. 1988), to show that *in vitro* methylation of the promoter regions abolishes expression of the gene after transfection into cells. In the case of the γ-globin gene, the inhibition is not dependant on the methylation of specific sites in the promoter, since the presence of three 5-mC residues in different regions of the promoter was sufficient to inhibit transcription (Murray and Grosveld 1987). Reporter gene constructs have also been used in microinjection experiments, in which the construct is methylated *in vitro* , and after transfer into recipient cells is found to be inactive (reviewed in Doerfler 1983). Two examples from more recent literature describe the transcription of methylated constructs that contained either a tRNA gene (Besser et al. 1990) or SV40 late genes (Götz et al. 1990). After the microinjection of the constructs into *Xenopus* oocytes, the transcription of genes was inhibited. These studies provide only circumstantial evidence that chromatin is involved in the mechanism of gene inactivation, since the formation of minichromosomes on construct DNA was implied but not proven.

Further evidence has been provided in a study by Buschausen et al. (1987), in which a construct containing the herpes simplex virus thymidine kinase (TK) gene was microinjected into TK^- tissue culture cells. Expression of the methylated reporter gene,

as assayed by the number of TK^+ cells, dropped sharply only 8 hours after microinjection. This time coincided with the assembly of the DNA into chromatin. In contrast, mock-methylated minichromosomes remained active 72 hours after injection. When methylated constructs were assembled into chromatin *in vitro* prior to injection, gene inactivation was immediate. An extension of this study has shown that a hemimethylated TK reporter gene construct was sufficient to block expression (Deobagkar et al. 1990). However, the authors did not investigate the methylation status of the construct after transfection, but it can be presumed that the construct became fully methylated due to maintenance methylation. DNA methylation of a reporter gene construct has been shown to inhibit the activity of the murine α1(I) collagen promoter by an indirect mechanism (Rhodes et al. 1994), after the transfection of the construct into tissue culture cells. An indirect mechanism requires both chromatin and methyl-CpG binding proteins to cooperatively mediate the inhibitory effect of methylation. The degree of transcriptional inactivation may be dependent on the density of CpG methylation, as shown for an episomal system that can be stably maintained in human cells. Plasmids containing the Epstein-Barr virus replication origin were methylated to four different levels of density, and transcriptional inactivation of minichromosomes was shown to be graded to an increase in methylation density (Hsieh 1994).

1.8.4 The effect of DNA methylation on chromatin structure

Although it is well established that methylation alters the conformation of chromatin, it is a much more complex problem to identify the proteins that cause the structural changes at the molecular level. This problem also brings us back to the more fundamental question of what constitutes the difference between active and inactive chromatin. So far, not many reports have dealt with the special structure of chromatin assembled on methylated DNA. In an early study it was reported that a methylated CpG polymer has a two-fold greater affinity for core histones than the unmethylated polymer, but that the core particles deposited on the substrate DNA were identical on both types of polymer (Felsenfeld et al. 1982). The study by Buschausen et al. (1987), which is described above, demonstrates an obvious difference in function between the chromatin of

unmethylated and methylated construct DNA. However, a structural difference was not observed, as tested by electron microscopy and micrococcal nuclease digestion (Buschausen et al. 1987; Graessmann and Graessmann 1993). In a recent important study core histone octamers and tetramers were assembled *in vitro* on unmethylated or methylated templates that contained human Alu repetitive elements (Englander et al. 1993). Alu elements possess internal RNA pol III promoters that are strongly transcribed *in vitro*, but are almost silent in somatic cells. The authors suggest that chromatin plays a role in silencing Alu elements, since the assembly of nucleosomes using histone octamers results in the equal inactivation of *in vitro* transcription for both unmethylated and methylated templates. The unexpected result with assembly of the H3/H4 tetramer was that the *in vitro* transcription of the methylated template was preferentially inhibited. The significance of these results may be that transcription activators compete with the H2A and H2B dimers for association with the histone H3/H4 tetramer. If the DNA that is bound to the tetramer is methylated then this may favour the completion of the histone octamer, with repression of transcription as a consequence.

CpG islands usually contain the promoters of housekeeping genes. Since these genes are by definition transcriptionally active, the chromatin at island regions should be in the open conformation. This has been shown to be the case with two X-linked genes, those for hypoxanthine phosphoribosyltransferase (*HPRT*) and glucose 6-phosphate dehydrogenase (*G6PD*), whose CpG islands are accessible to nucleases such as *Msp* I, DNase I and S1 nuclease on the active X chromosome (Kerem et al. 1983). This observation has been used to isolate short oligonucleosomes derived from CpG island chromatin (Tazi and Bird 1990). Nuclei were digested with CpG enzymes, which released a highly enriched fraction of transcriptionally active chromatin, but left bulk chromatin intact. The analysis of the protein composition of these two fractions showed significant differences between CpG island chromatin and bulk chromatin. The island chromatin contained hyperacetylated histones H3 and H4 (see Section 1.8.6), compared to bulk chromatin, which is now known to be a characteristic of transcriptionally active chromatin (reviewed in Turner 1991; Turner 1993). Nucleosomes that were obtained from CpG island chromatin were deficient in histone H1 (approximately 10% of the level observed in bulk chromatin), and one region of the island chromatin contained a DNase I

hypersensitive site, which coincided with the absence of a nucleosome. These features are all characteristic of active chromatin (Kamakaka and Thomas 1990; Hebbes et al. 1994; reviewed in Wolffe 1992).

1.8.5 The effect of histone H1 on gene expression

The special so-called "linker histone" in vertebrates, typically histone H1, may be the protein that can account for differences between chromatin of methylated and unmethylated DNA. Histone H1 is implicated in the formation of supranucleosomal structures, such as the 30 nm chromatin fibre (Thoma et al. 1979). Several models have been proposed for the structure of the 30 nm chromatin fibre, based on electron microscopy studies, but the strongest evidence supports the original solenoidal model (Felsenfeld and McGhee 1986) . This is in the form of neutron scattering data, that shows histone H1 to be located in the interior of the 30 nm chromatin fibre (Graziano et al. 1994). Supranucleosomal structures would be refractory to nuclease attack, which may account for the nuclease resistance of transcriptionally inert chromatin. Treatment of intact cells and isolated nuclei with the oligopeptide distamycin results in a decrease in nuclease resistance of chromatin (Käs et al. 1993). The increased accessibility is thought to result from the unfolding of the chromatin fibre, because distamycin binds to AT-rich sequences in the minor groove of DNA which would displace histone H1.

In an early study by Ball et al. (1983) it was demonstrated that at least 80% of 5-methylcytosine is localised in nucleosomes which contain histone H1, which is a characteristic of inactive chromatin (Weintraub 1984). Several studies have provided direct evidence that histone H1 mediates stable and tissue-specific gene inactivation in chromatin (reviewed in Zlatanova 1990), for genes transcribed by either RNA polymerase II (pol II) or RNA polymerase III (pol III). Two studies have used the developmentally-regulated oocyte 5S rRNA genes of *Xenopus laevis*, which are inactive in somatic tissues, as a model system for pol III transcription (Schlissel and Brown 1984; Wolffe 1989). In both cases, *in vitro* transcription from sperm chromatin that contains the oocyte-specific 5S rRNA genes could be repressed by the addition of histone H1. *Xenopus* sperm chromatin is naturally deficient in histone H1, whereas the removal

of histone H1 from somatic cell chromatin, such that regular nucleosomal spacing was maintained, brought about the activation of the oocyte 5S rRNA genes (Schlissel and Brown 1984). These studies implicate changes of chromatin structure in the developmental control of gene expression during embryogenesis. A recent study has shown that histone H1 mediated the selective repression of the oocyte 5S rRNA genes *in vivo* , but did not affect the expression of somatic 5S rRNA genes or endogenous tRNA genes (Bouvet et al. 1994; see Section 1.8.7).

The activity of the *Xenopus* somatic 5S rRNA gene has also been studied in an *in vitro* system. Chromatin with regularly spaced nucleosomes was assembled on the DNA *in vitro* by using a *Xenopus* oocyte extract (Shimamura et al. 1988; Gottesfeld 1989) and subsequent addition of histone H1 brought about transcriptional repression (Shimamura et al. 1989). Histone H1 could be removed from the chromatin without the disruption of the nucleosomal array by using an ion-exchange resin. This treatment restored the activity of the 5S rRNA gene. A second type of *in vitro* system has used purified core histone octamers to reconstitute chromatin on DNA in the absence or presence of histone H1 (Laybourn and Kadonaga 1991). The analysis of *in vitro* transcription from a pol II reporter gene again demonstrated gene repression by histone H1, which could be counteracted by the "anti-repressor" function of transcriptional activators. The accessibility of transcription factors to DNA could also be affected by the mobility of nucleosomes (Meersseman et al. 1992). In this model, transcriptionally active chromatin is repressed by linker histones that restrict the mobility of nucleosomes (Pennings et al. 1994). An *in vitro* system, that used only octamers as a model for chromatin, has shown that linker histones directly inhibit both nucleosome mobility and transcription (Ura et al. 1995).

The distribution of histone H1 in the nuclei of certain tissues has shown that most genes in transcriptionally active chromatin are depleted of histone H1, compared to genes that are inactive in this tissue (Kamakaka and Thomas 1990). A similar series of experiments have used the mouse mammary tumour virus (MMTV) promoter, that is present on a construct that has integrated stably into the genome of tissue culture cells. UV-crosslinking and immunoadsorption studies have shown that the transcriptionally-

active promoter is depleted of histone H1 (Bresnick et al. 1991). The inactive promoter binds histone H1, but treatment of the cells with glucocorticoid results in the induction of gene expression and depletion of histone H1 at the promoter. This important result, obtained in a system that is relevant physiologically, suggests that histone H1 has a fundamental role in the control of gene expression. However, it must be borne in mind that transcriptionally active chromatin is not completely depleted of histone H1 (reviewed in Garrard 1991), so the correlation between gene repression and the presence of histone H1 in chromatin is rather general. Kamakaka and Thomas (1990) suggest that the change from inactive to active genes involves an increase in the stoichiometry of histone H1 in chromatin.

Histone H1 has been demonstrated to be a type of methylated-DNA binding protein (Levine et al. 1993), at least under the *in vitro* conditions that were used in this study. This supports the data of Jost and Hofsteenge (1992), who reported that the methylated-DNA binding protein MDBP-2 shares sequence homologies with histone H1. Several studies have used histone H1 bound to DNA in orderly complexes as a model for inactive chromatin (reviewed in Paranjape et al. 1994). The assumptions of this model will not be discussed at this stage, since it is a major aspect of the work presented in Chapter 5. A full discussion can be found in Section 5.4. When histone H1 is associated with DNA *in vitro* , digestion with *Msp*I, but not with other restriction enzymes is specifically inhibited on the methylated DNA substrate (Higurashi and Cole 1991). This difference could be explained if the binding of histone H1 to methylated DNA changed the conformation of the complex, thus rendering the DNA more resistant to digestion by *Msp* I. This study by Higurashi and Cole (1991) did not demonstrate any preferential binding of histone H1 to methylated DNA, an observation that was repeated in a more recent study (Nightingale and Wolffe 1995). In contrast, a third study has shown that histone H1 can bind to methylated DNA with approximately two-fold greater affinity than to unmethylated DNA (Levine et al. 1993), which supports the role of histone H1 as an MDBP (Jost and Hofsteenge 1992). In addition, Levine et al. (1993) demonstrated that histone H1 not only inhibits *in vitro* transcription from unmethylated DNA (Jerzmanowski and Cole 1990; Croston et al. 1991), but that it inhibits transcription from methylated templates with even more efficiency. This observation has been extended with my own work (Johnson et al. 1995), which is presented in Chapter 5.

1.8.6 Histone modifications and changes in chromatin structure

Although it is well established that the nucleosome is the fundamental structural unit of eukaryotic chromatin, it is still unclear why the histone proteins, that are the highly conserved structural components of the nucleosome, exist in so many different forms. Some of this variation arises because each of the five classes of histone protein (H1, H2A, H2B, H3 and H4) exists as sub-types or isoforms that differ in a few amino acid residues. Further complexity arises from various post-translational modifications (reviewed in Wu et al. 1984; Wolffe 1992) which include acetylation, phosphorylation, ADP-ribosylation and ubiquitination of the parent histone molecule. Electrophoresis techniques have been used to identify and characterise these histone sub-types and modified forms (Lennox and Cohen 1989). Reversible modifications of histones at the molecular level are presumed to be the basis for the changes in chromatin structure, that could lead to the condensation and decondensation of chromosomes during the cell-cycle. For example, five specialised histone H2A variants and at least six variants of histone H1 (see Section 1.8.7) are present during sea urchin embryogenesis (Wolffe 1991). Each variant has a different expression profile for the cleavage, blastula and gastrula stages of embryogenesis (reviewed in Khochbin and Wolffe 1994). Core histones and linker histones are also subject to reversible post-translational modifications that correlate with stages of the cell-cycle.

Histone acetylation is the only post-translational modification that occurs consistently in transcriptionally active chromatin, compared to bulk chromatin that is transcriptionally inert (reviewed in Turner 1991; Turner 1993), and is the most extensively studied histone modification. Acetylation of the core histones is present in all the animal and plant species that have been studied (Csordas 1990), and occurs at specific lysine residues, all of which occur in the amino-terminal domains of the core histones. High-resolution proton magnetic resonance has shown that the amino-terminal domains of core histones are either fully mobile or interact weakly with core DNA in the nucleosome (Cary et al. 1978). A chemical modification procedure has also been used to

investigate core histone-DNA interactions (Lambert and Thomas 1986; Thomas 1989), and a later study from the same group has shown that the amino-terminal tail of histone H2A does not bind to core DNA in sea urchin sperm chromatin (Hill and Thomas 1990). Mobile amino-terminal "tails" of core histones could therefore interact with linker DNA, linker histones such as H1 or other non-histone proteins to mediate the formation of supranucleosomal structures. This model is supported by the observation that amino-terminal tails are required for the *in vitro* assembly of a solenoidal structure (Allan et al. 1982).

Each acetate group added to a lysine group reduces the net positive charge of the histone by one. This is presumed to reduce the electrostatic attraction between the basic core histones and the acid-phosphate backbone of the DNA, so that the core DNA is partially unwound or loosened around the core octamer (reviewed in Turner 1991). An electrostatic mechanism could therefore explain the open conformation and hyperacetylated status of core histones in transcriptionally active chromatin, because transcription factors are allowed access to regulatory elements in chromatin (Lee et al. 1993). Although a description of this mechanism at the molecular level is far from complete, the correlation between acetylation and transcriptional activity is well established. For example, experiments using Hg-agarose affinity chromatography have shown that the DNA of transcriptionally active genes was copurified with hyperacetylated core histones (Allegra et al. 1987). Tazi and Bird (1990) showed that chromatin fractions enriched in sequences characteristic of CpG islands contained hyperacetylated histones (see Section 1.8.4). Immunoprecipitation of mononucleosomes with an antibody recognising acetylated histones has shown the modification to be present at the β-globin locus in chicken embryo erythrocytes (Hebbes et al. 1994). The region of histone acetylation in this locus corresponded to a generalised DNase I sensitivity of the β-globin chromosomal domain. Recently it has been confirmed that inactive chromatin does not carry acetylated histone H4. (Jeppesen and Turner 1993) have shown that the human female inactive X chromosome (see Section 1.10) fails to stain with antibodies for the acetylated isoforms of histone H4, in contrast to all the other chromosomes. Autosomes were strongly labelled at the R-bands, which are the regions enriched in coding DNA, but were unlabelled at those regions adjacent to centromeres. This provides strong

evidence that a lack of histone acetylation may define both facultative and constitutive heterochromatin.

A second well-studied post-translational modification of chromatin is the reversible phosphorylation of linker histones, including histone H1 (reviewed in Wolffe 1992). Phosphorylation levels of histone H1 vary during the mitotic cell cycle of eukaryotes, with the levels increasing during S phase and mitosis, and peaking during metaphase when chromosomes are most highly condensed. Histone H1 is modified by a growth-associated histone H1 kinase, which is homologous to p34/cdc2 kinase. This kinase regulates many other nuclear processes that are driven by phosphorylation during the cell cycle (reviewed in Draetta 1990). Phosphorylation of serine residues in the highly basic amino-terminal domain of histone H1 reduces the net positive charge of the molecule. This would be expected to weaken the electrostatic interactions between linker histone and the DNA, in much the same way as core histone acetylation reduces histone-DNA interactions. Changes in the level of phosphorylation of sea urchin sperm-specific histone H1 have been studied during spermatogenesis and fertilisation (Hill et al. 1990), which are associated with changes in the condensation of chromatin in the nucleus. Phosphorylation of six sites in the C -terminal tail of histone H1 could abolish binding of the tail to DNA in both histone H1-DNA complexes and in sea urchin spermatid chromatin (Hill et al. 1991). This is expected from the weakened electrostatic attraction between a phosphorylated linker histone and the DNA. In view of this, it is a paradox that the phosphorylated chromatin in the spermatid becomes condensed and inert at this stage of spermatogenesis. The functional significance of linker histone phosphorylation remains unresolved.

1.8.7 Histone H1 variants

Linker histones exist in chromatin as a family of sub-types or isoforms, consisting of histones H1, H1°, H1t and H5, each of which has a set of variants (Lennox 1984; Wu et al. 1984). The diversity arises because the number of variants and the amount of each variant can change during cell differentiation. They also differ from tissue to tissue and for a given tissue even differ from one species to another (reviewed in Khochbin and

Wolffe 1994). Histone H1t is a testis-specific linker histone (Wu et al. 1984) and H1° tends to occur in quiescent cells such as neurones (Castiglia et al. 1993), but the functional significance of these variants is not known. Histone H5 is discussed in more detail below. The differences between linker histones are located in the highly basic amino- and carboxyl-terminal "tail" domains of linker histones, rather than in the central globular domain. The sequence of this region is highly conserved because it binds to the dyad of DNA in the nucleosome. This has lead to the proposal that the developmental regulation of linker histones can mediate changes in chromatin structure and transcriptional activity during gametogenesis, embryogenesis and within somatic cells (Bouvet et al. 1994; Pruss et al. 1995).

Histone H1 is the normal isoform found in somatic cells, but H5 is a special form that is expressed during the development of chicken erythroid cells. During the final stages of chicken erythrocyte development, the nucleus is condensed into inactive heterochromatin due in part to the appearance of histone H5. Histone H5 can reconstitute a supranucleosomal structure *in vitro* when it is added to chromatin that is depleted in linker histones (Graziano et al. 1988), but histone H5 both binds more tightly and confers greater stability to chromatin than does histone H1 (Sun et al. 1990). Both of these linker histones have a non-random distribution in chicken erythrocytes (Muyldermans et al. 1994). Telomeric nucleosomes are preferentially associated with histone H1 in this system, compared to bulk chromatin in which the most abundant sub-type is histone H5. Another study has shown that the chromatin associated with the β-globin gene in chicken erythrocytes is transcriptionally active and depleted in histones H1 and H5 (Postnikov et al. 1991), but that inactive chromatin is relatively enriched in these linker histones.

The changes of linker histone type in chromatin have also been studied during embryogenesis in two systems. In the first, the expression of the linker histone H1-β gene during sea urchin embryogenesis occurs in the late blastula-stage embryo, and subsequently in adult tissues (Lai and Childs 1988). The second system is the developmental regulation of chromatin-associated proteins during the early development of *Xenopus laevis* . Histone H1 is not found in *Xenopus* oocytes, sperm or eggs and the role of linker histone is taken by the B4 protein. B4 protein is replaced by histone H1

before the end of gastrulation (Dimitrov et al. 1993). Bouvet et al. (1994) have shown that *Xenopus* 5S rRNA gene transcription *in vivo* can be repressed by raising the level of histone H1 during gastrulation (see Section 1.8.5), after micro-injection of *in vitro* -synthesised mRNA into eggs. The over-expression of histone H1 specifically repressed oocyte 5S rRNA gene transcription, but did not affect the expression of somatic 5S rRNA genes or endogenous tRNA genes. Specific repression was mediated by the association of the *Xenopus* histone H1C variant with chromatin, despite adequate levels of the transcription factor TFIIIA which is required for the transcription of 5S rRNA genes. By contrast, the H1A variant did not appear to cause any preferential repression of the oocyte genes compared to the somatic genes.

The somatic variants of histone H1 have been isolated and characterised by reverse-phase HPLC (Cole 1989; Santoro et al. 1995) and the mouse variants consist of H1a, H1b, H1c, H1d and H1e (Lennox 1984). The number and relative amounts of these variants differ in various tissues and species, as well as during cell differentiation (reviewed in Khochbin and Wolffe 1994). The expression patterns of the variants is different in mouse lung carcinoma tissue, compared to normal lung tissue (Giancotti et al. 1993), and the relative proportions of histone H1 variants are changed during the differentiation of murine erythroleukaemic cells (Helliger et al. 1992). The variants may play different roles in the organisation of chromatin structure, with a non-random distribution of particular forms (see the discussion on histone H5, above). For example, one variant in larval tissue of the midge *Chironomus thummi* is localised in specific regions of polytene interphase chromosomes, whereas other variants have a uniform distribution (Schulze et al. 1993). A recent study has suggested that the histone H1e–c variants can specifically inhibit the *in vitro* methylation of a double-stranded DNA substrate (Santoro et al. 1995). A similar inhibition of the *in vitro* methylation of linker DNA is observed on the addition of histone H1 to H1-depleted oligonucleosomes (Caiafa et al. 1995). The addition of histone H1 correlates with the condensation of the oligonucleosomes, as revealed by distortions in the circular dichroism spectra of the particles (D'Erme et al. 1993). Caiafa et al. (1995) suggest that the specific H1e–c variants could maintain linker DNA in a hypomethylated form. This may be important in preserving the unmethylated state of CpG islands, since H1e–c are the only variants that are capable of binding to CpG-rich DNA (Santoro et al. 1995).

1.9 DNA methylation during development and cell differentiation

New patterns of tissue-specific methylation are established during embryogenesis by *de novo* methylation (reviewed in Razin and Cedar 1993). This is because the patterns of methylation derived from the gametes are erased in the morula and early blastula, and need to be re-established during the implantation of the embryo (Kafri et al. 1992). Adult levels of 5-mC are attained only after the completion of gastrulation (Monk 1990). The wave of *de novo* methylation during implantation remodels the genome into separate methylated and unmethylated compartments. Unmethylated regions of the genome contain CpG islands, which are associated with active and housekeeping genes, even in tissues where the gene is not expressed. In exceptional cases, CpG islands are methylated (see Section 1.6), but even in these cases the CpG islands appear to be unmethylated in the germ line (Brandeis et al. 1993). This is, of course, expected since methylation in a CpG island would lead to deamination and the erosion of the island. Methylation patterns of particular human genes, in lymphocytes and other blood cells, are often indistinguishable amongst different individuals even of different genetic backgrounds (Behn-Krappa et al. 1991). In addition, this pattern does not change when the lymphocytes are stimulated to divide by cytokines. These highly cell type-specific patterns of *de novo* methylation cannot be imposed just by the specificity of MTase for CG sequences, and must involve other *cis* -acting regulatory elements. Another example is the allele-specific methylation of one X chromosome of eutherian mammals.

Demethylation of modified CpG residues appears to be a necessary stage of early embryogenesis. The biological significance of demethylation is unknown, but demodification is a mechanism for removing differences in specific gene methylation patterns that emerge from the male and female germ lines. Monk et al. (1987) demonstrated that extensive demethylation of the genome takes place in the early mouse embryo, between the 8-cell and blastula stage. This is paralleled by a decrease in methylase activity (Monk et al. 1991). By the blastocyst stage all inherited patterns of methylation are lost, which results in hypomethylation of both extra-embryonic and

embryonic tissues (Kafri et al. 1992). This study also showed that *de novo* methylation that occurs during implantation lead to the remodification of non-island DNA, whereas CpG islands remain unmethylated.

An important development in this field has been the discovery that the presence of Sp1 sites in close proximity to CpG islands protects these islands from *de novo* methylation *in vivo* (Brandeis et al. 1994; Macleod et al. 1994). A similar result has been seen for the human X-linked *HPRT* gene, with the inactive allele completely methylated except at putative transcription factor binding sites (Hornstra and Yang 1994; see Section 1.10). This evidence suggests that steric hindrance by chromatin-associated factors, including transcription factors, may regulate *de novo* methylation in a tissue and allele-specific manner (Adams et al. 1993). Protein factors may therefore prevent the methylation of CpG islands by blocking the methyltransferase. However, this model cannot explain how even inactive genes are associated with unmethylated CpG islands. A previous study has shown that an *in vitro* methylated transgene, introduced into a fertilised mouse oocyte, can be specifically demethylated at a CpG island (Frank et al. 1991), and that the CpG island was protected from the *de novo* methylation of flanking regions.

Transfection experiments *in vitro* have suggested that demethylation is an active process (Frank et al. 1991), that is specific for islands. Findings in other systems are consistent with the view that demethylation is an active process (Razin et al. 1986; Paroush et al. 1990; Shemer et al. 1991). Nuclear extracts of chicken embryos can promote the active demethylation of DNA by an excision repair mechanism that is specific for 5-methyldeoxycytidine (Jost 1993). Excision is catalysed by a 5-mCpG endonuclease, which has been purified, and is found in developing mouse and chicken embryos as well as differentiating mouse myoblasts (Jost and Jost 1994). A 5-mC glycosylase activity, that has been detected in crude HeLa cell extracts, may also demethylate DNA by an active, enzymatic process (Vairapandi and Duker 1993).

Tissue-specific and gene-specific methylation patterns are established in late embryonic development. The relevant genes specific for the cell type in the differentiated tissue are demethylated, and this appears to be an integral part of the activation process. (Yisraeli et al. 1986) demonstrated that an *in vitro* methylated α-actin construct

undergoes demethylation and activation in myoblasts, but not in fibroblasts. The muscle-like cells therefore have the ability to recognise and activate the tissue-specific gene by demethylation. The demethylation appears to be dependent on *cis* -acting sequences and is independent of DNA replication (Paroush et al. 1990), that indicates that it is not a passive mechanism in which DNA is replicated in the absence of maintenance methylation. In conclusion, it is clear that vertebrate methylation patterns are dynamic and subject to genetic and developmental control.

1.10 The inactive X chromosome

The X chromosome contains 5% of the genes of the haploid mammalian genome, comprising of several thousand genes. Dosage equivalence for these genes between XY males and XX females is achieved by the random inactivation of one of the X chromosomes in each somatic cell of the female (reviewed in Grant and Chapman 1988; Rastan 1994). The silencing of transcription is achieved by the formation of facultative heterochromatin, which becomes highly condensed during interphase and is replicated late in S-phase. In metaphase spreads of female chromosomes the inactive X can be distinguished as the densely staining Barr body (Grant and Chapman 1988). In addition, the CpG islands of some genes on the inactive X are *de novo* methylated (Antequera and Bird 1993), and there is a decreased level of histone H4 acetylation (Jeppesen and Turner 1993). The latter observation has led to the proposal that histone H4 hyperacetylation defines active chromatin, since it is absent in both constitutive and facultative heterochromatin (see Section 1.8.6).

Primordial germ cells and oogonia also have one inactive X, but this becomes functional during early embryogenesis in a unique process of reactivation. Both X chromosomes are therefore active in oocytes and in cleavage-stage female embryos, but one X is again inactivated in somatic cells. Once X inactivation has occurred, it is stable and heritable in all subsequent cell divisions. The inactivation process is thought to occur in three stages: initiation of inactivation, the spreading of inactivation along the length of the chromosome and the maintenance of the inactive state of the chromosome.

A role for DNA methylation has been proposed for the maintenance of X chromosome inactivation, since methylation does not appear to be the primary mechanism that initiates X inactivation (Lyon 1993). Several studies have shown that the formation of facultative heterochromatin and transcriptional silencing precede *de novo* methylation of genes on the inactive X by several days (reviewed in Riggs and Pfeifer 1992). For example, transcriptional silencing of the mouse X-linked *hprt* gene takes place some days before methylation of the CpG island (Lock et al. 1987). Earlier studies have provided a general correlation between *de novo* methylation of CpG islands on the inactive X with nuclease insensitivity and transcriptional inactivation of X-linked genes, such as the glucose 6-phosphate dehydrogenase (*G6PD*) gene (Wolf and Migeon 1985; Toniolo et al. 1988) and the hypoxanthine phosphoribosyl-transferase (*HPRT*) gene (Sasaki et al. 1992). Reactivation of these genes on the inactive X coincided with the appearance of nuclease sensitivity in the CpG island chromatin after treatment with the demethylating drug 5-azacytidine. In a series of studies by Pfeifer and co-workers on the human X-linked phosphoglycerate kinase (*PGK*) gene, it was demonstrated that the 5' CpG island on the active X was completely unmethylated, free of nucleosomes and showed genomic footprints for putative transcription factors. In contrast, the same CpG island sequence on the inactive X was found to be hypermethylated, wrapped around two nucleosomes and did not reveal any footprints of transcription factor binding (Pfeifer et al. 1990; Pfeifer and Riggs 1991; Riggs and Pfeifer 1992). A similar study has compared the methylation status of a 5' region of the *HPRT* gene, which contains both a CpG island and several GC boxes, on the active and inactive X chromosomes (Hornstra and Yang 1994). Methylation analysis using genomic sequencing showed that the active allele was unmethylated and the inactive allele was completely methylated at all CpG sites, except at the GC boxes. However, the GC boxes exhibited an *in vivo* footprint of a putative transcription factor only on the active X chromosome.

The formation of facultative heterochromatin on the X chromosome is initiated at a locus known in humans as the X chromosome inactivating centre (XIC), after which the inactivating signal spreads in *cis* from the XIC in both directions along the X chromosome. In humans, the *XIST* gene (Brown et al. 1991) has been mapped to the region Xq13, which also contains the XIC locus. *XIST* , for X inactive specific

transcript, has the unique property of being transcriptionally active only on the inactive X chromosome (Brown et al. 1991). The *XIST* RNA transcript is 17 kb in size, includes several tandem repeats and does not appear to code for a protein. It has been suggested that either the RNA has a structural role, or that the transcript is irrelevant (Rastan 1994), and only the act of transcription through the XIC locus is sufficient for the initiation of the inactivating signal. The state of inactivation is then maintained by DNA methylation. Recent work has shown that a CpG-rich region at the 5' end of the mouse *Xist* gene is totally unmethylated on the inactive X in somatic tissues, in which *Xist* is expressed, but is completely methylated on the active X, in which *Xist* is silent (Norris et al. 1994). In addition, for tissues that undergo imprinted paternal X inactivation (see below), the paternal *Xist* allele is always unmethylated and the maternal allele is totally methylated. In the male germ line, demethylation of the CpG-rich region occurs during spermatogenesis, an event that precedes reactivation of *Xist* in the testis. Methylation is therefore implicated in the establishment of an imprinting signal (Ariel et al. 1995; Zuccotti and Monk 1995), as has been proposed for other imprinted genes (Peterson and Sapienza 1993). This is discussed in the next section.

1.11 Genomic imprinting

It has been known for many years that some loci in animals and plants are expressed in an epigenetic fashion that is dependent on their gamete of origin. In other words, when an allele at the locus passes through gametogenesis in one sex it remains active, but becomes silent when it passes through gametogenesis in the opposite sex, despite the identical genetic constitutions of the two parental loci (reviewed in Peterson and Sapienza 1993; Efstratiadis 1994; Holliday 1994). The silent locus is said to be "imprinted" in the germline by an epigenetic modification (Holliday 1989) that distinguished the parental origin of the two loci. The retention of epigenetic information in eukaryotes is required for phenomena other than parental imprinting, such as position-effect variegation in *Drosophila* , the telomere position effect and X chromosome inactivation (reviewed in Riggs and Pfeifer 1992; Bestor et al. 1994; Karpen 1994). Since the patterns of active

and inactive compartments of chromatin are faithfully reproduced each cell generation, it has been proposed that clonally-inherited conformations of chromatin may underlie all examples of epigenetic phenomena in eukaryotic cells (Sapienza 1990).

Several alternative models have been proposed to explain imprinting of mammalian genes at the molecular level and, although some models cannot be formally rejected (Efstratiadis 1994), there is now considerable evidence that methylation is essential (Li et al. 1993a; Razin and Cedar 1994). Four characteristics have been proposed to be essential for the imprinting mechanism, all of which can be fulfilled by methylation (Surani 1993; Efstratiadis 1994; Fundele and Surani 1994). First, the primary imprint must be established during gametogenesis, as this is the only stage during which male and female genome are separated and can be differentially imprinted. Second, the imprint must be clonally inherited from one cell generation to the other which requires the imprint to be reimposed following DNA replication. Third, the mechanism must be reversible as reprogramming is possible after passage through the germ line of the opposite sex. Finally, the imprint must affect the expression of the gene, in a direct or indirect way, to account for parental-dependent expression.

The first example of genomic imprinting to be discovered in mammals was the inactivation of the X chromosome during the early development of female mice. Extra-embryonic tissues, such as the trophectoderm and the primitive endoderm, displayed a non-random inactivation of the paternally-derived X chromosome (Takagi and Sasaki 1975). Several days later, the X chromosome was inactivated at random in the epiblast lineage, which provides all the cells of the embryo. Evidence for the involvement of DNA methylation came from the differential methylation observed during X inactivation (see Section 1.10), and recent studies on the *Xist* gene (Norris et al. 1994).

The molecular basis of imprinting on autosomes has been investigated following the discovery of endogenous imprinted mouse genes. These include the *Insulin-like growth factor II* (*Igf2*) gene, *Igf2 receptor* (*Igf2r*), *Insulin2* (*Ins2*), *Small nuclear ribonucleoprotein polypeptide N* (*Snrpn*) and the *H19* gene, which has no known function (reviewed in.Peterson and Sapienza 1993; Surani 1994). Restriction landmark genomic scanning (RLGS) has used the methylation imprints themselves to detect other imprinted loci, regardless of where or when they are expressed (Hatada et al. 1993). This

approach has shown that there are 100 imprinted loci throughout the genome, and has enabled the *U2afbp-rs* (U2af binding protein related sequence) gene to be cloned and characterised in detail (Hayashizaki et al. 1994). *Igf2* (Sasaki et al. 1992; Brandeis et al. 1993), *U2afbp-rs* and *Snrpn* are expressed exclusively from the paternal allele, whereas *Igf2r* (Stöger et al. 1993) and *H19* (Bartolomei et al. 1993; Brandeis et al. 1993; Ferguson-Smith et al. 1993) are transcribed only from the maternal copy in somatic cells. These studies also describe the allele-specific differences in methylation patterns for these imprinted genes (reviewed in Razin and Cedar 1994). In summary, the *H19* gene is methylated exclusively on the inactive paternal allele, whereas a small upstream region of the *Igf2* gene is methylated on the active paternal allele. This result is surprising, not least because *Igf2* and *H19* are closely linked in the genome and yet are imprinted in opposite directions. The *Igf2r* gene has an alternating pattern of parental methylation with methylation in the expected manner on the inactive paternal allele in one region, but with the unexpected methylation of the active allele in a second region.

The important role for DNA methylation in monoallelic expression of imprinted genes has been elegantly demonstrated in mice deficient in DNA methyltransferase activity (Li et al. 1992; Li et al. 1993b; see Section 1.5.3). The authors analysed the methylation status and expression of three imprinted genes in embryos that were homozygous for the targeted deletions in the MTase gene. As expected, the *H19* paternal allele, which was methylated and inactive in the wild-type, became active in the mutant embryos. Surprisingly, the lack of methylation silenced both the normally active paternal allele of *Igf2* and the normally active maternal allele of *Igf2r*. This observation suggests that methylation on the active allele can either attract an activator protein that binds preferentially to methylated DNA, or repel a methylation-sensitive transcriptional repressor. Such factors have yet to be characterised. The effect of methylation on individual imprinted genes may provide insights into regulatory mechanisms, but conclusions cannot be made from the study described above because the entire mouse genome is hypomethylated.

Chromatin, and chromatin-associated proteins such as transcription factors, may act as an epigenetic mechanism that can preserve the phenotype and determined state of a cell from one generation to the next. Some models of this phenomenon of "cell memory"

invoke stable transcription complexes that remain bound to DNA from one DNA replication to the next (Brown 1984; Wolffe 1994). Heritable chromatin states and DNA methylation may therefore act together to transfer epigenetic information to progeny cells. A recent study has examined if possible epigenetic chromatin structures remain in metaphase chromosomes, which represent the most compact chromosome state (Hershkovitz and Riggs 1995). The authors examined the promoter region of the X-linked human *PGK1* gene on the active and inactive X chromosome, and showed that transcription factors present on the active allele in interphase chromatin are absent in metaphase chromatin. This first study suggests that transcription complexes must form *de novo* each cell generation, at least for the *PGK1* gene. However, the mitotic inheritance of chromatin states in *Drosophila* are implicated in the heritable phenomenon of position-effect variegation (Karpen 1994). *Drosophila* cannot, of course, utilise clonally inherited patterns of DNA methylation to establish cell memory, so an alternative mechanism may well involve the clonal inheritance of chromosomal proteins. For example, mitotic inheritance regulates homeotic gene expression in *Drosophila* by the assembly of *Polycomb* group (Pc-G) and *trithorax* group (trx-G) proteins on or around the homeotic genes (Paro 1993).

1.12 Aims of the project

The complexity of gene regulation in eukaryotes has hampered attempts to analyse and describe mechanisms, compared to the systems in prokaryotes. The exact roles of chromatin and DNA methylation in eukaryotic systems remain elusive, so that their importance for the control of gene expression, genomic imprinting, replication and development is implied rather than proven. As described above, DNA methylation plays a role in these fundamental biological processes, most of which centre around the vital question of how DNA methylation exerts its effects on gene expression. Evidence has been gathered for three different mechanisms of transcriptional inactivation by DNA methylation:

(a) The presence of methyl groups in the promoter region can inhibit transcription factor
binding and thus reduces transcriptional activation.
(b) Proteins which bind specifically to methylated DNA can act as repressors, restricting access of the transcriptional machinery to promoter regions.
(c) the presence of methyl groups in a gene sequence results in the formation of an inactive chromatin structure, rendering the gene inaccessible for transcription.

In this study I have investigated the correlations between DNA methylation, inactive chromatin formation and the repression of gene expression. To simplify matters, I chose to model the *in vivo* situation, in which genes are always expressed in the context of chromatin, by using *in vitro* systems for transcription (chapter 3) and chromatin assembly (chapter 4). The assembly of chromatin was further simplified in a series of experiments that used only complexes of histone H1 and DNA as a model of inactive chromatin (chapter 5). Kass et al. (1993) have shown that regional DNA methylation on a plasmid transfected into mammalian cells acts as a focus for the formation of inactive chromatin. I have investigated this further by the study of *in vitro* chromatin formation and transcription on a variety of regionally-methylated plasmid constructs (chapter 6).

CHAPTER TWO

Materials and methods

2.1 Materials

2.1.1 List of suppliers

Unless otherwise stated all the chemicals were Analar grade, supplied by BDH Chemicals, Poole, Dorset, or Fisons Scientific, Loughborough, Leics. Radiochemicals were purchased from Amersham International PLC, Aylesbury, Bucks. Growth media for bacteria were supplied by Difco Laboratories, Detroit, USA, and for the cultivation of mammalian cells by GIBCO/BRL Ltd., Paisley, Scotland. Prokaryotic DNA methylases were purchased from New England Biolabs, Beverley, MA, USA. Restriction enzymes and other DNA modifying enzymes were obtained from Promega Ltd., Southampton; Pharmacia Ltd., Milton Keynes and Boehringer Mannheim Ltd., Lewes, East Sussex, if not otherwise specified. Where special chemicals, reagents or equipment were obtained from other sources, this is indicated in the text.

2.2. Bacterial cell culture

2.2.1 Bacterial strain

Escherichia coli XL1-Blue MRF' (Stratagene Ltd., Cambridge) was the host strain used for the growth of all plasmid DNA as well as for the isolation of single-stranded phagemid DNA. It has the following genotype: Δ*(mcrA)183*, Δ*(mcrCB-hsdSMR-mrr)173*, *endA1*, *supE44*, *thi-1*, *recA*, *gyrA96*, *relA1*, *lac*, λ-, (F', *proAB*, *lacI*^{q}*ZΔM15*, Tn*10*,(tet^{r})).

2.2.2. Bacterial growth media

LB medium
10 g bactotryptone
10 g NaCl
5 g yeast extract
adjusted to pH7.0 with 5M NaOH and made up to 1 litre with dH_2O

2x YT medium
16 g bactotryptone
5 g NaCl
10 g yeast extract
adjusted to pH7.0 with 5M NaOH and made up to 1 litre with dH_2O

SOB medium
20 g bactotryptone
0.5 g NaCl
5 g yeast extract
10 ml of 250 mM KCl
5 ml of 2M $MgCl_2$
made up to 1 litre with dH_20;
sterile $MgCl_2$ added after autoclaving

The media were supplemented with 15 g bactoagar/litre for growth on plates. All media and solutions used for the growth of micro-organisms were sterilised as appropriate by autoclaving for 20 min at 15 psi or by filter sterilisation. Antibiotics were added as indicated in Sections 2.7.1.5 to 2.7.1.9.

2.3. Mammalian cell culture

2.3.1. Mammalian cell line

HeLa S3 cells (Gey et al. 1952), a human, cervical carcinoma cell line, were used for the preparation of nuclear extracts for *in vitro* transcription assays.

2.3.2. Cell culture media

HeLa S3 cells were grown in Eagle's minimum essential medium (EMEM) which contained 450 ml dH_2O, 50 ml 10x EMEM, 50 ml new-born calf serum, 30 ml of 7.5% sodium bicarbonate, 5 ml of 200 mM L-glutamine, 5 ml non-essential amino acids and 5 ml of penicillin/streptomycin (10,000 units/ml and 10,000 µg/ml, respectively). Refer to Adams (1990a) for further details.

All components were obtained sterile and freshly made media were checked for bacterial and fungal contamination 2–3 days before use.

2.4. Buffers and solutions

All solutions used in the routine handling of nucleic acids were sterilised, as appropriate, by autoclaving for 20 min at 15 psi, or otherwise by filter sterilisation. Buffers that were used for more specialised procedures are detailed elsewhere in the text, in appropriate sections.

Phosphate buffered saline (PBS)

Buffer A (pH 7.2)

NaCl	0.17 M
KCl	3.35 mM
Na_2HPO_4	10 mM
KH_2PO_4	1.84 mM

Buffer B

$CaCl_2 \cdot 6H_2O$	6.8 mM

Buffer C

$MgCl_2 \cdot 6H_2O$	4.9 mM

buffers A, B and C were autoclaved separately and mixed in a ratio of 8:1:1 before use

TE buffer

Tris-HCl (pH 7.9)	10 mM
EDTA	1 mM

TBE buffer, 1x (pH 8.3)

Tris base	89 mM
boric acid	89 mM
EDTA	2 mM

Tris-glycine buffer, 1x (pH 8.3)

Tris base	25 mM
glycine	250 mM
SDS	0.1% (w/v)

PAGE/DNA elution buffer

ammonium acetate	0.5 M
magnesium acetate	10 mM
EDTA	1 mM
SDS	0.1% (w/v)

SSC buffer, 20x

NaCl	3 M
tri -sodium citrate	0.3 M

adjusted to pH 7.5 with HCl

50x Denhardt's reagent

Ficoll (Type 400, Pharmacia)	1% (w/v)
BSA (fraction V, Sigma)	1% (w/v)
polyvinylpyrrolidone	1% (w/v)

Hybridisation buffer

SSC	5x
Denhardt's reagent	5x
SDS	0.5% (w/v)

FSB buffer

potassium acetate	10 mM
$MnCl_2 \cdot 4H_2O$	45 mM
$CaCl_2 \cdot 2H_2O$	10 mM
KCl	100 mM
$(Co[NH_3]_6)Cl_3$	3 mM
glycerol	10% (v/v)

10x T4 polynucleotide kinase buffer

Tris-HCl (pH 7.6)	0.5 M
$MgCl_2$	0.1 M
DTT	50 mM
spermidine	1 mM
EDTA	1 mM

10x S1 nuclease buffer

sodium acetate (pH 4.5)	300 mM
NaCl	2 M
$ZnSO_4$	20 mM
glycerol	5% (v/v)

10x T4 ligase buffer

Tris-HCl (pH 7.4)	200 mM
$MgCl_2$	50 mM
DTT	50 mM
BSA	500 µg/ml

10x CIP buffer

Tris-HCl (pH 8.4)	500 mM
$MgCl_2$	10 mM
$ZnCl_2$	10 mM
spermidine	10 mM

10x Klenow buffer

Tris-HCl (pH 7.6)	500 mM
$MgCl_2$	100 mM
DTT	1 mM

Solutions for plasmid preparations

Solution I:

glucose	50 mM
Tris-HCl (pH 8.0)	25 mM
EDTA	10 mM

Solution II:

NaOH	0.2 M
SDS	1% (w/v)

Solution III:

K-acetate	60 ml
glacial acetic acid	11.5 ml
dH_2O	28.5 ml

DNA loading buffer IV

sucrose	40% (w/v)
EDTA	5 mM
bromophenol blue	0.1% (w/v)
xylene cyanol FF	0.1% (w/v)
NaN_3	0.02% (w/v)

Orange G loading buffer

glycerol	50% (v/v)
EDTA	5 mM
orange G	0.1% (w/v)
NaN_3	0.02% (w/v)

Denaturing gel loading buffer

formamide (deionised)	98% (v/v)
EDTA	10 mM
bromophenol blue	0.1% (w/v)
xylene cyanol FF	0.1% (w/v)

2.5. DNA vectors, recombinants and oligonucleotides

DNA vectors, recombinants and oligonucleotides were routinely stored in TE buffer (see Section 2.4) at –20°C. DNA stored in this way remains stable for several years. Long-term storage of DNA was under 100% ethanol at –70°C. Supercoiled plasmid, prepared by caesium chloride gradients (see Section 2.7.1.8), was stored in TE buffer at 4°C.

2.5.1. DNA vectors and recombinants

Plasmid pArg/Leu is a derivative of pUC19 (Sambrook et al. 1989). A 2.8 kb *Bam*HI-*Eco*RI restriction fragment, which contains a human $tRNA^{Arg}$ and a $tRNA^{Leu}$ gene, was sub-cloned from a recombinant plasmid that has been described previously (Bourn et al. 1994). This plasmid contained a cluster of four human tRNA genes: $tRNA^{Lys}$, $tRNA^{Gln}$, $tRNA^{Leu}$ and $tRNA^{Arg}$, as well as a GC-rich region which was also sub-cloned into pArg/Leu (Figure 2.1). The GC-rich region has prevented the complete sequencing of the insert, so the size of the insert and the positions of some restriction sites, such as those for *Sma* I, have been estimated from restriction mapping (Bourn et al. 1994). The positions of *Bgl* I sites in pArg/Leu are shown in Figure 5.17A, some of which were determined by restriction mapping (discussed in Section 5.3.7).

Figure 2.1 (see overleaf)

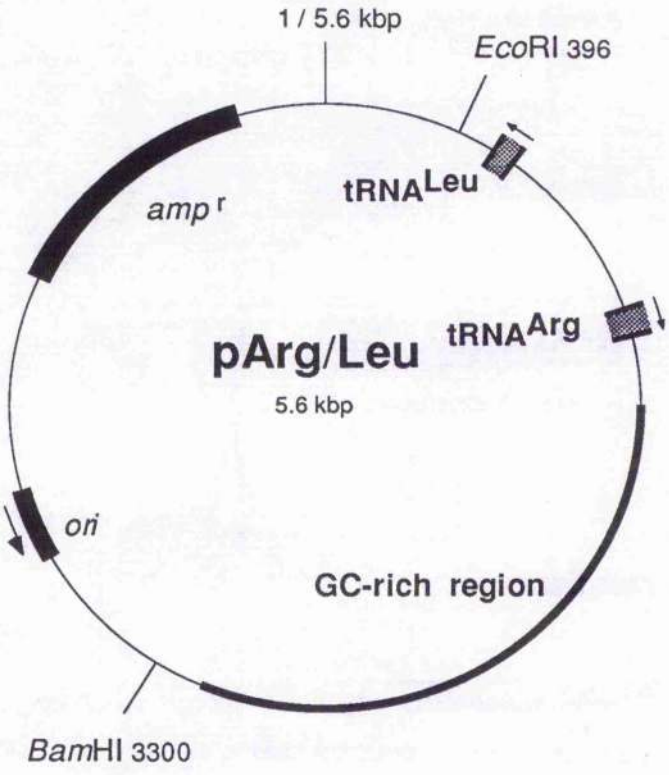
Maps of plasmid pArg/Leu:

A: The diagram shows a simplified circular map of pArg/Leu, comprising of two human tRNA genes ($tRNA^{Arg}$ and $tRNA^{Leu}$) and a GC-rich region inserted into pUC19 at the *Bam* HI and *Eco* RI sites. Regions in pUC19 are: amp^r (ampicillin resistance gene); *ori* (origin of replication). The size of regions are to scale.

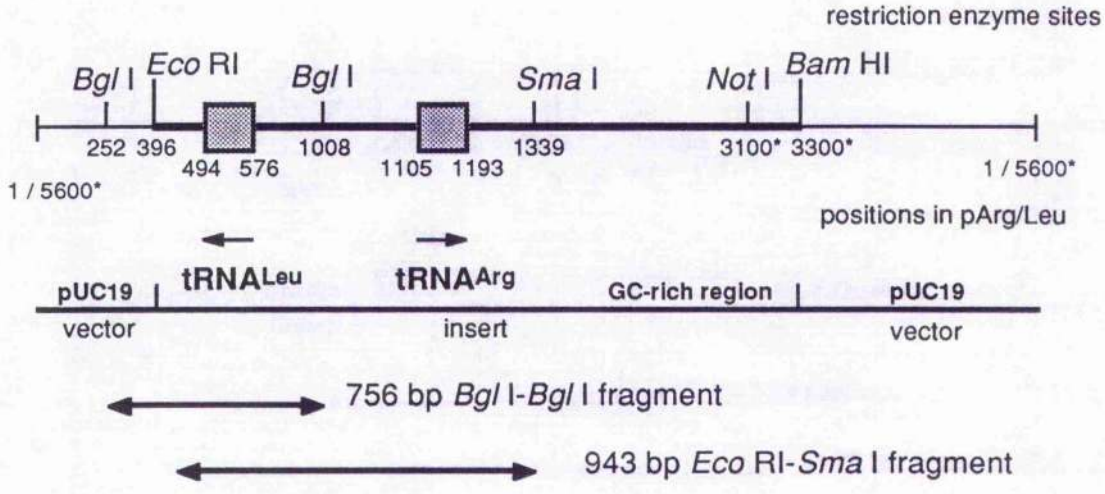
B: Linear map of the sub-cloned insert in pArg/Leu. The position of the tRNA genes is known from sequencing data (see Section 2.5.1), but the position of some restriction sites has only been determined by restriction mapping, indicated by asterisks (positions to the nearest 100 bp). The 756 bp fragment is a *Bgl* I-*Bgl* I restriction fragment that contains the $tRNA^{Leu}$ gene, and is used in subsequent experiments (see Section 5.3.7). The *Eco* RI-*Sma* I fragment is 943 bp in size, and is used in the experiments described in Section 5.3.4.

Figure 2.1

A


1 / 5.6 kbp
EcoRI 396
tRNALeu
amp r
tRNAArg
pArg/Leu
5.6 kbp
ori
GC-rich region
BamHI 3300

B


restriction enzyme sites
Bgl I
Eco RI
Bgl I
Sma I
Not I
Bam HI
252 396
494 576
1008
1105 1193
1339
3100* 3300*
1 / 5600*
1 / 5600*
positions in pArg/Leu
pUC19
tRNALeu
tRNAArg
GC-rich region
pUC19
vector
insert
vector
756 bp Bgl I-Bgl I fragment
943 bp Eco RI-Sma I fragment

Plasmid pVHCk is a construct based on pBluescript II KS- (Stratagene, Cambridge), which contains an f1 origin of replication and hence allows the isolation of single-stranded DNA. It contains the SV40 early promoter/enhancer region linked to the CAT reporter gene and the SV40 terminator region as shown in Figure 2.2. The construction of pVHCk is described in Kass et al. (1993), and is based on an earlier plasmid pVHC1 used for transfection studies (Bryans et al. 1992).

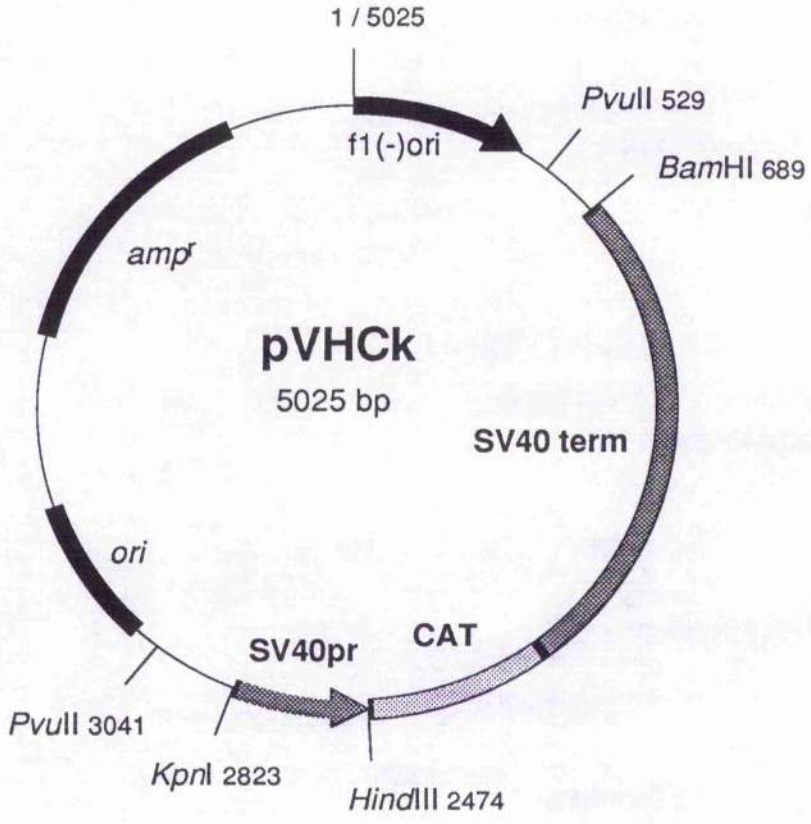


Figure 2.2

Map of plasmid pVHCk:
A circular map of pVHCk (see Section 2.5.1) The map shows essential regions, and their positions in the plasmid, to scale. The regions include: *amp*^{r} (ampicillin resistance gene); *ori* (origin of replication); SV40pr (promoter region of SV40); CAT (chloramphenicol acetyltransferase reporter gene); SV40term (terminator region of SV40); f1(-)ori (f1 origin of replication).

2.5.2 Synthetic oligonucleotides

Synthetic oligonucleotides for primer extension assays were made in the Department of Biochemistry by Dr. V. Math, using an Applied Biosystems 381A DNA synthesiser for the phosphite-triester method.

2.6. DNA size markers

The following DNA standards were used as size markers for the analysis of nucleic acids during gel electrophoresis. Figure legends refer to markers by the following labels:

λ^{H}: **bacteriophage λ DNA**, digested with ***Hin*** **dIII**: 23.130, 9.416, 6.557, 4.361, 2.322, 2.027, 0.564, 0.125 (size in kb).

Φ **: bacteriophage *Φ* X174** (RF form dsDNA), digested with ***Hinf*** **I**: 726, 713, 553, 500, 427, 417, 413, 311, 249, 200, 151, 140, 118, 100, 81, 66, 48, 42, 40, 24 (size in bp).

100_B: 100 bp marker: 2200, 1500, 1400, 1300, 1200, 1100, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100 (size in bp). The 600 bp fragment is approximately three times more intense than other bands. (The marker was supplied by GIBCO/BRL).

2.7. Methods

2.7.1 General methods used in molecular biology

The following methods are standard procedures in molecular biology and essentially follow protocols described by Sambrook et al. (1989), adapted to the equipment available in the lab.

2.7.1.1 Phenol/chloroform extraction

Phenol, equilibrated with TE buffer and containing 8-hydroxyquinoline at 0.1% (w/v) was mixed with chloroform/isoamylalcohol (24:1) at a ratio of 1:1. Extraction of an aqueous DNA solution was performed as follows: an equal volume of phenol/chloroform was added to the DNA solution, vortexed and centrifuged for 5 min in a microfuge. The upper aqueous phase was transferred to a fresh tube and the extraction repeated if necessary.

2.7.1.2 Ethanol precipitation

A salt solution was added to the DNA solution (sodium acetate, adjusted to pH5.2 with glacial acetic acid, to 0.3 M; NaCl to 0.2 M; ammonium acetate to 2.0 M) and the DNA precipitated with 2.5 volumes of ethanol (98% v/v) at –20°C for 15 min. The DNA was recovered by centrifugation at 12,000 g for 15 min at 4°C. Subsequent washes with cold 70% (v/v) ethanol removed any salt which was coprecipitated with the DNA.

2.7.1.3 Quantification of nucleic acids

Spectrophotometry

The absorbance of various dilutions of DNA samples was measured using quartz 1.0 ml cuvettes in a spectrophotometer at 260 nm. An absorbance of 1.0 corresponds to approximately 50 µg/ml for double-stranded DNA, 40 µg/ml for single-stranded DNA and 20 µg/ml for oligonucleotides.

Fluorimetry

Accurate measurements of DNA concentrations, down to 10 ng/ml, were determined using a Hoefer TKO 100 DNA minifluorometer. DNA solutions were added to 2.0 ml of a working dye solution. Aliquots of the working dye solution, of final volume 100 ml, were replaced daily and stored in the dark at room temperature. Each aliquot consisted of: 90 ml double-distilled, filter sterilised water (to remove particulate contaminants), 10 ml TNE, x10 (100 mM Tris-HCl pH8.0, 10 mM EDTA, 1.0 M NaCl; adjusted to pH7.4 with conc. HCl) and Hoerst H33258 dye solution to a final concentration of 100ng/ml. Readings were performed as described in the manufacturer's instructions. Serial dilutions of both calf thymus DNA and pUC19 plasmid DNA, from 10–500 ng/ml final concentration, were used to construct standard curves from which the DNA concentration of a fixed volume of sample DNA was determined.

2.7.1.4 Preparation of competent cells

For the preparation of competent cells a single colony of *E. coli* XL1-Blue was transferred to 20 ml of LB medium containing tetracycline (12.5 µg/ml) and incubated overnight at 37°C. 500 µl of this suspension were used to inoculate 100 ml of LB medium containing tetracycline (12.5 µg/ml). Incubation was for 2 to 3 hours until an OD^{600} (optical density at 600 nm) of 0.45–0.55 was reached. The cells were then cooled down on ice for 10 min and harvested by centrifugation at 4,000 rpm for 10 min at 4°C using a Beckman J2-21 centrifuge and JA14 rotor. The pellet was carefully resuspended in 40 ml of ice cold FSB buffer (Section 2.4) and stored for 10 min on ice. The cells were recovered as above and resuspended in 6 ml of FSB buffer. 200 µl of DMSO were then added. Cells were dispensed into aliquots of 150 µl, snap-frozen in liquid nitrogen and stored at -70°C.

2.7.1.5 Transformation of bacteria

Transformation-competent cells (150 µl, section 2.7.1.4) were thawed on ice and plasmid DNA (no more than 50 ng) was added. After a 30 min incubation on ice the

tubes were transferred to a 42°C water-bath, incubated for 90 seconds and rapidly transferred to an ice bath. After 5 min, 800 µl of LB medium was added and incubated for 45 min at 37°C to allow the bacteria to express the antibiotic resistance. Dilutions of the bacterial suspension were made and plated out on LB plates containing ampicillin (100 µg/ml) and tetracycline (12.5 µg/ml).

2.7.1.6 Small scale preparation of plasmid DNA

A 3 ml overnight culture was prepared from a single colony of transformed bacteria in LB medium, supplemented with ampicillin (100 µg/ml). 1.5 ml of the bacterial suspension was transferred to an Eppendorf tube and the bacteria pelleted at 5,000 rpm (Eppendorf 5415 microcentrifuge) for 5 min. The pellet was suspended in 100 µl solution I (see Section 2.4). 200 µl solution II was added, vortexed and incubated for 5 min on ice. Subsequently, 150 µl of solution III was added, vortexed and incubated again for 5 min on ice. The bacterial debris was pelleted at 12,000 rpm for 10 min at 4°C. To the supernatant was added 460 µl phenol/chloroform mix; the mixture was vortexed and the phases separated at 10,000 rpm for 10 min at 4°C. To the aqueous phase was added 1 ml ethanol and the sample was incubated for 15 min at –70°C. The DNA was pelleted at 12,000 rpm for 10 min at 4°C and the pellet resuspended in 40 µl of TE buffer. 10 µl were used for a restriction digest which included RNase A at 1 µg/µl.

2.7.1.7 Large scale plasmid preparation

Transformed cells were grown overnight in 200 ml LB medium supplemented with ampicillin (100 µg/ml). The cells were harvested by centrifugation for 15 min at 6,000 rpm in a Beckman J2-21 centrifuge using a JA 14 rotor. The pellet was resuspended in 10 ml solution I (Section 2.4.). Lysis of the cells was achieved by addition of 10 ml solution II and incubation on ice for 5 min. Addition of 7.5 ml solution III precipitated out the cell debris, which was sedimented by centrifuging at 10,000 rpm at 4°C (Beckman J2-21 centrifuge, JA 20 rotor) for 15 min. The supernatant was filtered through Whatman 3MM paper and 0.6 volumes of *iso*-propanol were added to precipitate

the nucleic acids for 15 min at room temperature. The nucleic acids were pelleted at 3,000 rpm for 15 min at 4°C using a Beckman CS-6R benchtop centrifuge and resuspended in 2 ml TE buffer. 2 ml 5 M LiCl/50 mM Tris-HCl (pH 8.0) was added and incubated 10 min on ice to precipitate RNA. RNA was sedimented at 3,000 rpm for 5 min at 4°C (Beckman benchtop centrifuge) and the DNA in the supernatant precipitated with 2 volumes of ethanol for 15 min on ice. The plasmid DNA was recovered by centrifugation for 15 min at 3,000 rpm at 4°C (Beckman benchtop centrifuge) and resuspended in 500 µl TE buffer. The DNA was then ethanol precipitated a second time after a phenol, a phenol/chloroform and a chloroform extraction and finally dissolved in 500 µl TE buffer.

2.7.1.8 Preparation of supercoiled plasmid DNA

Supercoiled plasmid DNA for *in vitro* transcription assays was prepared by isopycnic centrifugation in a caesium chloride gradient. Recombinant plasmid was prepared as described in Section 2.7.1.7, and diluted to 30 ml with TE buffer to which were added 28.9 g CsCl (BRL, UK; ultrapure optical grade) and 1.8 ml ethidium bromide stock solution (at 10 mg/ml). This gave a solution of density 1.55 g/ml and a refractive index of 1.3860, as measured with an Abbe refractometer (Bellingham and Stanley Ltd., London) for the sodium D_1 line (wavelength 589.6 nm), as recommended by the manufacturer.

The solution was clarified by centrifugation at 10,000 rpm for 15 min at 4°C (JA20 rotor), after which the clear red supernatant was carefully aspirated to avoid the surface scum. The solution was transferred to Beckman VTi50 QuickSeal tubes, and the tubes filled with light mineral oil if necessary. The samples were centrifuged at 50,000 rpm at 20°C overnight in a Beckman VTi50 rotor. After equilibrium was reached two bands of DNA were seen. The upper band consisted of linear bacterial DNA and nicked plasmid DNA, whilst the lower band consisted of supercoiled DNA. The latter was collected using an 18 gauge hypodermic needle, which pierced the tube just below the band. Ethidium bromide was removed by repeated extraction with an equal volume of 3-methylbutan-1-ol, followed by the addition of 4.0 vol. TE buffer. DNA was precipitated

by adding 2.0 (total) vol. ethanol and centrifuging at 15,000 rpm for 20 min at 4°C (JA20 rotor). The DNA pellet was resuspended in TE and the quality of the preparation was assessed by agarose gel electrophoresis (Section 2.7.1.13). This protocol gave preparations of at least 95% supercoiled plasmid.

2.7.1.9 Isolation of single-stranded DNA from phagemids

Single-stranded DNA from pBluescript phagemids was isolated as described by Blondel and Thillet (1991). A colony of *E. coli* XL1-Blue, harbouring the phagemid of interest and grown on LB plates supplemented with ampicillin (100 µg/ml) for selection of the phagemid and tetracycline (12.5 µg/ml) for selection of the F' episome, was suspended in 20 µl 2x YT medium (Section 2.4). The bacterial suspension was then infected with 5 µl of helper phage (M13 K07, Stratagene, Cambridge), with a titre of at least 10^{10} PFU/ml, and incubated for 15 min at room temperature. 500 µl of 2x YT medium with antibiotics (ampicillin at 100 µg/ml and tetracycline at 12.5 µg/ml) was added and incubated for 1 hour at 37°C to let the helper phage express antibiotic resistance. 120 µl of the incubated cells were then added to 3 ml 2x YT medium supplemented with ampicillin (100 µg/ml), tetracycline (12.5 µg/ml) and kanamycin (75 µg/ml). The culture was then incubated with vigorous shaking for 18 to 20 hours at 37°C. Cultures were then pooled (typically into a total volume of 30 ml) and the bacteria pelleted by centrifugation in a Beckman JA20 rotor at 6.5k rpm for 7 min. Phage particles were precipitated by incubating with 2.5% (w/v) PEG 6000 and 0.35 M NaCl for 1 hr. at room temperature, and recovered by centrifugation (12k rpm, 5 min in a JA20 rotor). The phage pellet was dissolved in 1000 µl TE buffer. After two phenol extractions and one chloroform extraction the single-stranded DNA was ethanol precipitated and dissolved in 300 µl TE buffer. The preparation of ssDNA was then checked on an agarose gel and the concentration determined by spectrophotometry.

2.7.1.10 Restriction enzyme digests of DNA

The DNA was incubated with the appropriate restriction enzyme in the presence of the appropriate buffer (supplied together with the restriction enzyme as 10x buffer) for at least 2 hours at 37°C, with a DNA concentration not higher than 1 µg/10 µl and an enzyme concentration of about 10 units/µg DNA. The volume of the restriction enzyme did not exceed a 1/10 of the total volume, as glycerol can affect the performance of the enzyme.

2.7.1.11 Ligations

Routinely, 200 ng of linearised vector DNA were used in a ligation in a total volume of 10 µl. The vector:insert molar ratio was 1:3. 1 µl of 10x T4 ligase buffer (section 2.4) and 1 µl of 10 mM ATP were added to the DNA and the ligation initiated by addition of 1 Weiss unit of bacteriophage T4 DNA ligase. Incubation was for 8 to 12 hours at 16°C.

2.7.1.12 Dephosphorylation of plasmid DNA

The 5´-phosphate groups of a linear plasmid were removed by calf intestinal alkaline phosphatase (CIP) to avoid self-ligation of the vector. 1 µg of linearised vector DNA was incubated in CIP buffer (Section 2.4) with 1 unit of calf intestinal alkaline phosphatase (Boehringer Mannheim GmbH) in a total volume of 50 µl at 37°C for 30 min. SDS and EDTA were added to final concentrations of 0.5% and 5 mM, respectively. Proteinase K (Boehringer Mannheim GmbH) was added to a final concentration of 100 µg/ml and incubated for 30 min at 55°C. The reaction was subjected to a phenol and phenol/chloroform extraction and the DNA was recovered by ethanol precipitation.

2.7.1.13 Agarose gel electrophoresis of DNA

Agarose gels were routinely used to separate DNA fragments of size range 0.2–20 kbp. Agarose (GIBCO/BRL) was dissolved at the desired concentration (0.8 - 2.0% w/v) in TBE 0.5x buffer (Section 2.4) containing 0.5 µg/ml ethidium bromide by heating in a microwave oven until boiling. A well-forming comb was inserted and the gel allowed to set. One half volume of DNA loading buffer IV (Section 2.4) or orange G loading buffer (Section 2.4) was added to the DNA samples before being applied to the gel. Electrophoresis was carried out at 50 mA in TBE 0.5x buffer containing 0.5 µg/ml ethidium bromide. The DNA fragments were visualised by using a long wavelength UV transilluminator.

2.7.1.14 Isolation of DNA fragments from LMP agarose gels

Low melting point (LMP) agarose (GIBCO/BRL) gels were prepared exactly as agarose gels, but were run in a cold room. The DNA band that was required was visualised using a hand-held long wavelength UV lamp and excised from the gel using a sterile scalpel blade. The gel slice was transferred to an Eppendorf tube and heated to 65°C for 15 min to melt the agarose. To isolate DNA fragments of size range 0.4–14 kbp, to yield amounts up to about 10μg, a Geneclean II kit (Bio 101, CA, USA) was used as recommended by the manufacturer.

Higher yields of restriction fragments of plasmid DNA (10–50 μg) were required for the generation of patch-methylated constructs (Section 2.7.3.2), so a different strategy was chosen to isolate sufficient amounts of the DNA fragment. Depending on the size of the restriction fragment to be isolated, 50 to 100 μg of plasmid DNA were digested with the desired restriction enzyme(s). Complete digestion was checked on a minigel and the bulk of the restriction digest was separated on a preparative LMP agarose gel, where the sample was applied in one large slot. The desired band was identified and the fragment cut out of the gel. The plunger of a 1 ml (or 2 ml) disposable syringe was removed and the syringe plugged with sterile siliconised glass wool. The agarose slice was placed in the syringe, and the DNA was recovered by centrifugation of the syringe for 20 min at 25°C in a Beckman CS-6R benchtop centrifuge. 500 μl of boiling TE buffer (section 2.4) was added to the agarose that was trapped by the glass wool, and the procedure of centrifugation was repeated to extract any remaining DNA. The DNA solution was chloroform extracted and the DNA recovered by ethanol precipitation.

2.7.1.15 Polyacrylamide gel electrophoresis (PAGE) of DNA

Polyacrylamide gels were routinely used to separate DNA fragments of size range 10–300 bp. They were prepared by mixing a suitable volume of Easigel acrylamide/bis-acrylamide stock solution (Scotlab, Paisley, UK), containing 30% [w/v] acrylamide and 1.034% [w/v] N,N'-methylene-bisacrylamide, with water and an appropriate volume of

TBE x10 buffer to give the desired acrylamide concentration (5–20% [w/v]) in TBEx1.0. The acrylamide was polymerised by the addition of 0.006 vol. of 10% ammonium persulphate and 0.001% TEMED at 42°C for 30 min. DNA samples were mixed with one half volume of DNA loading buffer IV, and gels were run at 50 mA in TBE x1.0 electrophoresis buffer. Gels were stained in TBE x1.0 containing 0.5µg/ml ethidium bromide for 30 min, and after rinsing in distilled water were visualised as for agarose gels.

2.7.1.16 Isolation of DNA fragments from PAGE gels

The desired band was identified and the DNA fragment cut out of the gel. The DNA was recovered by a modification of the "crush and soak" method (Sambrook et al. 1989). Gel slices, each weighing approximately 100 mg, were transferred to Eppendorf tubes and crushed using a yellow pipette tip. To each tube was added 300 µl PAGE/DNA elution buffer (Section 2.4) and the mixture was incubated for 16 hr at 37°C with vigorous shaking. The gel pieces were pelleted by centrifugation at 12,000 rpm for 10 min in a microcentrifuge. The supernatants were pooled, ethanol precipitated and dissolved in 50 µl TE buffer. Finally, to remove any remaining gel particles the DNA fragments were purified using a ChromaSpin-10 TE spin column (Section 2.7.1.20).

2.7.1.17 Purification of oligonucleotides

60 µg (approx. 2 OD^{260}) of the oligonucleotide were lyophilised to dryness in a rotary vacuum lyophiliser to remove the aqueous ammonia. The pellet was then dissolved in 20 µl H_2O and 20 µl formamide was added. Before loading on a gel the DNA was heat denatured for 5 min at 55°C. Separation was on 1 mm thick, 19% denaturing polyacrylamide (acrylamide : N,N'-methylene-bisacrylamide 19:1) gels in 1x TBE buffer (Section 2.4). The gel was pre-run for 30 min at 30 mA before the oligonucleotide was loaded. To monitor the electrophoresis, 5 µl of denaturing gel loading buffer (Section 2.4) was run alongside the oligonucleotide samples. After completion of the electrophoresis the apparatus was dismantled and the gel transferred onto a fluorescent

surface. The oligonucleotide was visualised by illuminating with a hand-held long wavelength UV lamp. The oligonucleotide absorbed the UV radiation and appeared as dark bands on the fluorescing background. The top bands of the oligonucleotide (containing the full size oligonucleotide) were marked, excised from the gel and isolated as described in Section 2.7.1.16.

2.7.1.18 End labelling of oligonucleotides

Synthetic oligonucleotides were end-labelled by transfer of the γ-^{32}P from [γ-^{32}P]ATP using bacteriophage T4 polynucleotide kinase (PNK). 10 pmol of oligonucleotide were incubated in T4 PNK buffer (Section 2.4) with 5 µl of [γ-^{32}P]ATP (5,000 Ci/mmol, 10 mCi/ml) and 10 units of T4 PNK in a total volume of 20 µl. Unincorporated radiolabel was removed using ChromaSpin-10 columns (Section 2.7.1.20).

2.7.1.19 Random-primed radiolabelling of DNA fragments

The Megaprime DNA labelling system from Amersham (RPN 1606) was used for the random-primed radioactive labelling of probes for Southern hybridisations. 5 µl of primer solution (containing random nonamer primers) and water to a final reaction volume of 50 µl were added to 100 ng of purified fragment DNA. The DNA was denatured for 5 min at 95°C in a boiling water-bath. Subsequently 10 µl of Megaprime reaction buffer (containing dATP, dGTP, dTTP, $MgCl_2$, 2-mercaptoethanol and Tris-HCl buffer [pH 7.5]; the exact concentrations are not given in the company literature), 5 µl [α-^{32}P]dCTP (3,000 Ci/mmol, 10 mCi/ml), 2 µl Klenow enzyme (1 unit/µl) and water to a final reaction volume of 100 µl were added and the reaction incubated for 30 min at 37°C. The reaction was stopped by the addition of 5 µl 0.2 M EDTA and unincorporated radionucleotide was removed by using ChromaSpin-10 columns (Section 2.7.1.20).

2.7.1.20 Removal of unincorporated radionucleotides from radiolabelled DNA

ChromaSpin-10 TE spin columns (Clontech Laboratories, Palo Alto, CA) were used to remove unincorporated radionucleotide from radioactively labelled DNA. The column was drained and packed by centrifugation at 2,200 rpm for 5 min at 4°C in a Beckman CS-6R bench-top centrifuge. The sample (maximum volume 100 µl) was carefully applied onto the centre of the column and centrifuged as above. The radioactively labelled sample was collected in an Eppendorf tube at the bottom of the column, whereas the unincorporated nucleotides remained in the column. The specific activity of the radiolabelled sample was calculated by comparison of the radioactivity before and after removal of the free nucleotides.

2.7.1.21 Southern blotting and hybridisation

A combined vacuum/alkaline blot method was used to transfer DNA from agarose gels to nylon membranes. A vacuum blotting apparatus from Hybaid (Teddington, Middx., UK) was used to perform the alkaline Southern blot. A piece of Whatman 3MM filter paper and a piece of nylon membrane (Hybond N+, Amersham International PLC, Amersham, UK) were cut slightly larger than the gel to be blotted and pre-wetted in 0.4 M NaOH. The paper was placed on top of the porous plate of the blotting apparatus, followed by the nylon membrane and a rubber gasket, which was cut with an aperture slightly smaller than the gel to be blotted. The agarose gel was briefly soaked in 0.4 M NaOH and then carefully placed on the nylon membrane, to avoid trapping any air bubbles between gel and membrane. The vacuum was then applied and the transfer buffer (0.4 M NaOH) poured in the chamber so that the gel was completely immersed. The vacuum was reduced by a valve to 80 cm of water and the transfer time was 30 min. After the transfer was completed the nylon membrane was neutralised by brief washing in 5x SSC buffer. For some experiments, the DNA fragments was blotted by the traditional technique of capillary transfer (Sambrook et al. 1989). Diffusion of fragments

in the agarose during transfer reduced the resolution of bands, as compared to the vacuum transfer procedure, so the latter was the preferred method for the experiments described in Chapter 6.

Southern blots were prehybridised in hybridisation buffer (section 2.4) containing 100 µg/ml heat denatured sonicated salmon sperm DNA for 1–2 hours at 65°C. The heat denatured probe (1–5 pg/ml hybridisation buffer, 10^8–10^9 cpm/µg) was added and incubated for 12–16 hours at 65°C in a shaking water bath. The filter was washed twice in 2x SSC, 0.1% SDS for 15 min at room temperature and subsequently once in 1x SSC, 0.1% SDS at 65°C for 15 min and twice in 0.1x SSC, 0.1% SDS at 65°C for 15 min. For experiments with relatively high amounts of target DNA (>50 ng) and a long homologous probe (>400 bp.) the stringency conditions for hybridisation were increased. (The majority of *Msp* I fade-out assays that are described in chapter 6 are experiments of this type). The conditions for higher stringency washes were: 2x SSC, 0.1% SDS for 15 min at 68°C and subsequently once in 1x SSC, 0.1% SDS at 68°C for 15 min and twice in 0.1x SSC, 0.1% SDS at 68°C for 15 min. (and longer if required). The membrane was then autoradiographed for the desired time.

If the blot was subjected to a hybridisation with a second probe the membrane was never allowed to dry during or after hybridisation and washing. The blot was stripped of the first probe by pouring a boiling solution of 0.5% SDS on the membrane and allowing it to cool to 50°C. The complete removal of the probe was checked by autoradiography, after which the blot was hybridised to the second probe, which was prepared to have a higher specific activity than the first.

2.7.2 Methods for quantification and analysis of proteins

2.7.2.1 Quantification of protein concentrations

The concentration of proteins in solution was assayed using the method of Bradford (1976). Water was added to an aliquot of the protein sample to a total volume of 100 µl, followed by 1.0 ml of Bradford's reagent (8.5% [v/v] phosphoric acid, 5% [v/v] ethanol, 0.01% [w/v] Coomassie Brilliant Blue G). The A^{595} of the mixture was

measured, and the protein concentration estimated from a calibration curve of standard concentrations (0–40 µg per assay) of bovine serum albumin for routine determinations. The concentrations of histone proteins were determined by comparison with standards of purified core histones (Section 2.7.6), to compensate for the anomalous effect of basic proteins on colorimetric assays.

2.7.2.2 Analysis of proteins by SDS-PAGE

SDS-polyacrylamide gel electrophoresis was used to separate proteins of size range between 10–100 kDa. SDS-PAGE resolving gels were prepared by mixing suitable volumes of Easigel acrylamide/bisacrylamide stock solution (Scotlab, Paisley, UK), containing 30% [w/v] acrylamide and 1.034% [w/v] N,N'-methylene-bisacrylamide, with 1.5 M Tris-HCl [pH8.8], 10% SDS and water to give the desired acrylamide concentration (10–15% [w/v]), 0.375 M Tris buffer and 0.1% SDS. The acrylamide was supplemented by the addition of 0.006 vol. of 10% ammonium persulphate and 0.001% TEMED, poured between a pair of vertical glass plates, overlaid with *iso*-butanol and allowed to polymerise at 42°C for 30 min. The overlay was then poured off, the top of the resolving gel was washed with water, and SDS-PAGE stacking gel solution was poured directly onto the surface of the resolving gel. Stacking gels were prepared by mixing suitable volumes of 30% acrylamide/bisacrylamide stock solution, 1.0 M Tris-HCl [pH6.8], 10% SDS and water to give final concentrations of 5% acrylamide, 0.125 M Tris buffer and 0.1% SDS. The stacking gel was polymerised as described above.

Samples containing approximately 10 µg protein were denatured by heating to 100°C for 3 min in 1x SDS gel-loading buffer (50 mM Tris-HCl [pH6.8], 100 mM DTT, 2% SDS, 0.1% [w/v] bromophenol blue, 10% glycerol) and loaded onto the gel. Electrophoresis was initially at 8V/cm, until the dye front had moved into the resolving gel, after which the voltage was increased to 15V/cm. Gels were fixed and stained in a solution containing 10% [v/v] methanol, 10% [v/v] glacial acetic acid and 0.25% [w/v] Coomassie Brilliant Blue R250 for 4 hr at room temperature. The gel was then thoroughly destained by soaking in several changes of destain solution containing 10% methanol and 10% glacial acetic acid.

2.7.3 Methylation of plasmid DNA

2.7.3.1 Methylation of plasmid DNA *in vitro* using prokaryotic methylases

Prokaryotic methylases M.*Sss* I, M.*Hpa* II and M.*Hha* I and their assay buffers were purchased from New England Biolabs (Beverley, MA, USA). Routinely, 20 µg plasmid DNA were incubated with 20 units of methylase (60 units in the case of M.*Hpa* II) in the appropriate assay buffer for 16 hours at 37°C in a total volume of 100 µl. The concentration of the methyl-group donor S-adenosyl methionine (SAM) was 80 µM for M.*Hpa* II and M.*Hha* I and 160 µM for M.*Sss* I, respectively. For inactivation of the methylases SDS and EDTA were added to final concentrations of 0.5% and 5 mM, respectively. Proteinase K (Boehringer Mannheim GmbH) was added to a final concentration of 100 µg/ml and incubated for 30 min at 55°C. The reaction mixture was subjected to a phenol and phenol/chloroform extraction and the DNA recovered by an ethanol precipitation. For every methylation assay, a mock-methylated control was treated in the same way as described above, but omitting the methylase in the methylation assay.

2.7.3.2 Patch-methylation of pVHCk plasmid DNA

Restriction fragments were gel purified and annealed to ssDNA at a molar ratio of 3:1 (fragment DNA:ssDNA) in 20 mM Tris-HCl (pH7.5); 10 mM $MgCl_2$; 50 mM NaCl; 1 mM DTT by heating for 2 min at 95°C and allowing to cool down from 70 to 30°C in 1 hr. To prevent binding to and methylation of the single-stranded region by methylase *Sss*I (M.*Sss* I), T4 gene 32 protein (Boehringer Mannheim Gmbh) was added (10 µg/µg of ssDNA), and the reaction mixture was incubated at 37°C for 15 min. The double-stranded DNA patch was methylated for 16 h, using M.*Sss* I (New England Biolabs) under conditions recommended by the manufacturer. After proteinase K treatment and phenol/chloroform extraction, the unmethylated gap was filled in and ligated by standard procedures (Sambrook et al. 1989). For every patch-methylated construct, a

mock-methylated control was made in the same way but omitting M.*Sss* I. Constructs were separated from unannealed restriction fragment by gel purification. Precise concentrations of purified DNA were determined by using a Hoefer TKO 100 DNA minifluorometer (see Section 2.7.1.3). Calf thymus or pBluescript KSII- DNA was used as a standard. To confirm the location of the methylated region, constructs were digested with the appropriate restriction enzymes to release the methylated region and subjected to *Hpa* II digestion. The DNA was then separated on a 1.3% agarose gel, transferred to a nylon membrane (Hybond N+; Amersham), and hybridised to the appropriate labelled fragment (Section 2.7.1.21). Membranes were reprobed after stripping with boiling 0.5% SDS solution. The results of these analyses are presented in Figure 6.6 A–C and discussed in Section 6.3.1.

2.7.4. *In vitro* transcription assays

All solutions for the preparation of nuclear extracts and for the *in vitro* transcription assay were either treated with diethylpyrocarbonate (DEPC) or made up with DEPC-treated water, and were autoclaved when possible. DEPC was added to a final concentration of 0.1% (v/v) and incubated at 37°C for at least 4 hr with vigorous agitation. Glassware was treated with 3% (v/v) H_2O_2 for 16 hr and then rinsed with DEPC-treated water. Disposable plasticware was either subjected to multiple autoclaving or rinsed briefly with chloroform to reduce the risk of RNase contamination. *In vitro* transcription is discussed in more detail in Section 3.1.

2.7.4.1 Preparation of nuclear extracts from HeLa cells

Nuclear extracts for *in vitro* transcription assays were prepared according to Dignam et al. (1983). HeLa S3 cells were grown in siliconised spinner flasks at 37°C in EMEM medium (Section 2.3) to a density of 4 to 6 x 10^5 cell/ml. A total of about 10^9 cells were used for a typical extract preparation. Cells were harvested by centrifugation for 10 min at 2,000 rpm and room temperature (Beckman J2-21 centrifuge, JA 14 rotor). Pelleted

cells were suspended in 5 volumes of PBS (4°C) and collected by centrifugation as above. Subsequent steps were performed at 4°C. The cells were suspended in 5 packed cell volumes of buffer A (10 mM HEPES-KOH [pH 7.9], 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT) and incubated on ice for 10 min. Cells were collected by centrifugation as before and resuspended in two packed cell volumes of buffer A and lysed by 10 strokes of a Kontes all glass Dounce homogeniser, using the B type pestle. The homogenate was centrifuged as before to pellet the nuclei. The nuclear pellet was subjected to a second centrifugation for 20 min at 15,000 rpm (Beckman J2-21 centrifuge, JA 20 rotor) at 4°C to remove any residual cytoplasmic material. These crude nuclei were resuspended in 3 ml buffer C (20 mM HEPES-KOH [pH 7.9], 25% v/v glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT) per 10^9 cells and homogenised as above. The resulting suspension was stirred gently for 30 min at 4°C and centrifuged for 30 min at 15,000 rpm at 4°C. The resulting clear supernatant was dialysed against 50 volumes of buffer D (20 mM HEPES-KOH [pH 7.9], 20% v/v glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT) for 5 hours and the dialysate centrifuged at 15,000 rpm for 20 min at 4°C. The supernatant was snap frozen as 50 µl aliquots in liquid nitrogen and stored at -70°C.

In a modification to the protocol, histone H1 was depleted from the dialysed extract by the selective precipitation of proteins by the addition of 1.25 vol of 4 M ammonium sulphate (Croston et al. 1991) buffered with 50 mM HEPES-KOH pH7.9, in a protocol described elsewhere (Shapiro, Sharp et al. 1988). The proteins were pelleted by centrifugation at 100,000 g for 30 min at 4^oC, and were resuspended in buffer D (Dignam et al. 1983).

2.7.4.2 *In vitro* transcription assay using direct internal radiolabelling

This method was used to assay the transcripts from the human $tRNA^{Leu}$ and $tRNA^{Arg}$ genes, contained in the recombinant pArg/Leu (Section 2.5). These are class III genes that are transcribed by RNA polymerase III (refer to Section 3.2). The conditions for transcription of tRNA genes by pol III have been described previously (Gonos and

Goddard 1990), and with subsequent modifications (Bourn et al. 1994; Johnson et al. 1995). The protocol is discussed in greater detail in Sections 3.2.2 and 3.2.3. The nuclear extract used for this pol III system was the one prepared by the Dignam protocol (Dignam et al. 1983), so the extract contained some contaminating, endogenous histone H1. A reaction mixture, of final volume 20 µl, contained 5 µl extract (7.0 µg protein/µl), 40 U RNasin (Promega) and final concentrations of 14 mM HEPES-KOH pH7.9, 80 mM KCl, 20 mM NaCl, 3.5 mM $MgCl_2$, 10% (v/v) glycerol, 0.3 mM DTT and 0.3 mM PMSF. The reaction mixture included the desired amount of template DNA (usually 150 ng), with pUC19 DNA as a non-specific carrier, if required, to a total of 150 ng DNA per assay. The mixture was preincubated at 30°C for 20 min (see Section 3.2.3), after which transcription was initiated by the addition of unlabelled ribonucleoside triphosphates at pH7.0 (600 µM ATP, CTP and GTP; 25 µM UTP), 2.5 µCi of [α-^{32}P] UTP (3,000 Ci/mmol, 10 mCi/ml) and 10 mM creatine phosphate. The transcription reaction was allowed to proceed at 30°C for 60 min and terminated by the addition of 10% (w/v) SDS to a final concentration of 0.5%, 100 µg proteinase K (Boehringer Mannheim GmbH) and 40 µg *E. coli* crude tRNA as carrier. The mixture was incubated at 37°C for 20 min and then an equal volume 1.0 M ammonium acetate was added. The mixture was extracted with phenol-chloroform (Section 2.7.1.1), which was then back-extracted with an equal volume of 0.5 M ammonium acetate. The RNA in the pooled aqueous phases was ethanol precipitated (Section 2.7.1.2), washed with 70% ethanol, dried thoroughly and resuspended in denaturing gel loading buffer (Section 2.4). Samples were analysed by electrophoresis on a 12% polyacrylamide, 4 M urea denaturing slab gel, followed by autoradiography. Transcription was quantified either by scintillation counting or by using a Fuji BAS1000 phosphoimager.

2.7.4.3 *In vitro* transcription assay using primer extension

In vitro transcription assays of the pVHCk construct (Section 2.5.2) were quantified using a primer extension approach (Leonard and Patient 1991; Mason et al. 1993). Transcription from the SV40 promoter requires RNA polymerase II and preliminary

experiments established that this transcription system required an optimal $MgCl_2$ concentration of 4 mM. These optimisation experiments and an extensive discussion of this protocol can be found in Sections 3.3.1 to 3.3.4. The nuclear extract that was used in this pol II system was prepared according to the Shapiro protocol (Shapiro et al. 1988), so the extract was depleted of histone H1. This simplified the interpretation of results for the pol II system to the effects of exogenous histone H1 (see Section 6.2.3). A standard *in vitro* transcription assay was carried out as follows:

HeLa nuclear extract	27 µl
$MgCl_2$ (50 mM)	3.2 µl
ATP-CTP-GTP-UTP-mix (pH 7.0, each at 10 mM)	4.0 µl
RNasin (Promega, 40 U)	1.0 µl
supercoiled pVHCk plasmid DNA (50 - 500 ng)	5 µl
total reaction volume:	40 µl

The reaction mixture was incubated for 45 min at 30°C, after which 100 µl of DNase digestion buffer (20 mM HEPES-NaOH [pH 7.6], 50 mM NaCl, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 10 mM DTT) and 1.0 U DNase I (RQ1 RNase-free DNase, Promega) was added. The reaction was incubated at 37°C for 10 min and was stopped by the addition of 150 µl DNase stop buffer x2 (20 mM Tris-HCl [pH 8.0], 0.3 M NaCl, 1.0% SDS, 50 mM EDTA, 40 µg tRNA, 20 µg proteinase K) and incubation at 37°C for a further 30 min. After phenol/chloroform extraction and back extraction, the aqueous phase was heated to 65°C for 10 min to inactivate any residual DNase. The RNA was ethanol precipitated, air-dried and dissolved in 10 µl of dH_2O.

To assay the level of the specific mRNA, first strand cDNA was synthesised from the transcripts using an 5' end-labelled primer complementary to the target RNA. The amount of cDNA obtained was a measure of the level of transcription of that particular gene. An oligonucleotide (oligo PE;1 shown in Figure 3.5), complementary to the mRNA of the CAT gene was end-labelled (Section 2.7.1.18) and diluted to 50 fmol/µl. 1.0 µl of primer (10^4 cpm) was mixed with 10 µl RNA and 4.0 µl 5x reverse transcriptase buffer (250 mM Tris-HCl [pH 8.3], 250 mM KCl, 50 mM $MgCl_2$, 50 mM DTT, 5 mM spermidine) and incubated for 20 min at 65°C, followed by 10 min at 42°C.

To this was added 1 µl dNTP-mix, each at 10 mM (Promega), and 10 U AMV reverse transcriptase (Promega, 12 U/µl). After incubation for 30 min at 42°C, the nucleic acids were directly ethanol precipitated, dissolved in 8 µl denaturing gel loading buffer (Section 2.4) and separated on a 10% denaturing polyacrylamide gel.

A necessary preliminary experiment was to determine that a significant proportion of the primer formed a hybrid with the mRNA. This was achieved by treating hybridisation reactions (see above) with an excess of nuclease S1 (a suitable protocol is described in Mason et al. 1993). In brief, a 1.0 µl aliquot of the hybridisation reaction was added to 200 µl of nuclease S1 buffer x1(Section 2.4). Four 50 µl aliquots were removed, and 300 U nuclease S1 (Boehringer Mannheim GmbH) were added to only two of the aliquots. At the concentration used, the nuclease S1 degraded all single-stranded DNA, including unannealed primer, but did not digest duplexes of primer hybridised to RNA. The digestion reactions were then spotted onto a disc of DE81 paper (Whatman). The discs were allowed to air-dry and were then washed extensively with 5% (w/v) Na_2HPO_4, and once each with water, and then ethanol. The filters were air-dried and the radioactivity assayed by scintillation counting. The proportion of the primer which is resistant to nuclease S1 is determined by averaging duplicate values, and expressing the radioactivity of the samples with nuclease S1 as a percentage of the radioactivity of the samples with no nuclease S1. See Section 3.3.2 for further details.

2.7.5 Histone H1 preparation and complex formation with DNA

2.7.5.1 Renaturation of histone H1

Total acid-extracted, lyophilised calf thymus histone H1 (Boehringer Mannheim GmbH) was used in initial experiments. Lyophilised protein was dissolved in a buffer containing 10 mM Tris-HCl [pH7.5], 1.0 mM DTT, 0.4 mM PMSF, 5 mM EDTA, 0.01% (v/v) Nonidet-P40, 10% (v/v) glycerol. Nonidet-P40 was included in all buffers containing histone H1 to prevent non-specific adsorption to plastic or glass (Croston et al. 1991). The presence of EDTA facilitated histone renaturation (Caiafa et al. 1991). Samples were

renatured by step-dialysis (Hentzen and Bekhor 1985) to the histone H1 buffer, starting from buffer containing 2 M NaCl, 5 M urea, 10 mM TRIS-HCl [pH7.5], 1.0 mM DTT, 0.4 mM PMSF, 5 mM EDTA. The subsequent buffers used in the dialysis had the following decreasing amounts of urea and NaCl: 1 M NaCl, 5 M urea; 0.8 M NaCl, 5 M urea; 0.6 M NaCl, 5 M urea; 0.4 M NaCl, 5 M urea; 0.4 M NaCl. Protein concentrations were determined using the method of Bradford, with purified core histones as standards to compensate for the anomalous effect of histones on colorimetric assays (Section 2.7.2.1). Histone H1 preparations were snap-frozen in liquid nitrogen and stored frozen in histone H1 buffer at –70^{O}C.

Somatic histone H1 variants H1a-e were partially purified by reversed-phase HPLC chromatography (Quesada et al. 1989), as modified by (Santoro et al. 1993). Protein fractions were lyophilised and renatured as described above. Histone H1 variants are discussed in detail in Section 1.8.6, and experiments that use histone H1 variants are described in Section 5.3.5.

2.7.5.2. Formation of histone H1-DNA complexes

Total histone H1 was allowed to bind to 100 ng supercoiled template DNA under conditions that facilitated the formation of a "slow" complex, as defined by Clark and Thomas (1986). "Slow" complexes are presumed to form by the non-cooperative and reversible binding of H1 to DNA. Histone H1 dissolved in histone H1 buffer was mixed, in various proportions, with DNA in 20 mM HEPES-KOH [pH7.0] and 2.0 mM $MgCl_2$ in the presence of 15 mM NaCl. Mixtures were incubated at 27^{O}C for 60 min, after which the complexes were used in subsequent procedures. "Fast" complexes, as defined by Clark and Thomas (1986), were presumed to be aggregates formed by the irreversible binding of histone H1 to the DNA. The conditions used in this study were optimised for slow complex formation and minimised the formation of fast complexes. The formation of histone H1-DNA complexes, under these conditions, was assayed by agarose (1.0%) gel electrophoresis in 0.25x TBE buffer and 20% (v/v) glycerol (see Section 5.2.1). H1-DNA complexes were retarded in comparison to naked supercoiled DNA, whereas aggregates of histone H1 and DNA did not migrate out of the wells of the agarose gel (Figure 5.2).

2.7.6. Purification of core histones

The method used was a consensus of the methods of Stein and Mitchell (1988) and Workman et al. (1991). Core histones were prepared from dispersed chromatin by salt elution on hydroxyapatite. All steps in the procedure were carried out at 4°C. Nuclear material, derived from the procedure for making HeLa nuclear extracts (Section 2.7.4.1), was stored at –70°C in chromatin storage buffer (CSB; 10 mM TRIS-HCl [pH7.4], 0.1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 0.4 mM PMSF, 10% [v/v] glycerol) until required. The dispersed chromatin was pelleted by centrifugation at 6k rpm for 20 min (Beckman JA 20 rotor) and resuspended in 20 pellet volumes of lysis buffer (10 mM TRIS-HCl [pH8.0], 0.25 M sucrose, 3 mM $MgCl_2$, 0.5 mM PMSF, 1.0% [v/v] Nonidet-P40). Nuclear material was washed twice in lysis buffer and twice in rinse buffer (the same as lysis buffer but omitting Nonidet-P40) by centrifugation as described above. The concentration of DNA was determined by resuspending the nuclear material in an aliquot of 2 M NaCl and measuring the A^{260}. Nuclear material containing approximately 15 mg DNA (300 A^{260} units) was used in subsequent steps.

Pelleted nuclear material was resuspended in 10 ml low salt buffer (LSB; 0.4 M NaCl, 10 mM TRIS-HCl [pH8.0], 1 mM EDTA, 0.5 mM PMSF) by gentle homogenisation, stirred for 15 min on ice, pelleted by centrifugation as above and washed once more in LSB. Pelleted material was then resuspended in 10 ml medium salt buffer (MSB; 0.6 mM NaCl, 50 mM Na_2PO_4 [pH6.8], 0.5 mM PMSF) and stirred for 10 min on ice. To the suspension was added 5.0 g dry DNA-grade Bio-Gel HTP hydroxyapatite (Bio-Rad) with slow stirring, to allow the resin to swell to a paste. Dispersed chromatin was immobilised by the adsorption of the DNA onto the hydroxyapatite matrix. The paste was resuspended in a minimum volume of MSB and poured into a wide (10 x 2.5 cm) glass barrel chromatography column. The suspended hydroxyapatite was allowed to settle under gravity and the column volume estimated to be 10 ml. The column was washed with 10 column volumes of MSB and core histones were eluted, in the form of natural core complexes, with a step of high salt buffer (HSB; 2.5 M NaCl, 50 mM Na_2PO_4 [pH6.8], 0.5 mM PMSF). Fractions of 1.5 ml were

collected and the protein concentration of each one was determined by measuring A^{230}. Peak fractions were pooled and the core histones were concentrated to 2–4 mg/ml by using low-molecular weight cut-off Centriprep YM10 concentrators (Amicon), as recommended by the manufacturers. Concentrated core histones were dialysed against CSB for 3 hr, snap-frozen as small aliquots in liquid nitrogen and stored at -70^{o}C. The purity of core histone preparations was assessed by SDS-PAGE (Section 2.7.2.2) on 15% acrylamide gels (Figure 5.1).

2.7.7 *In vitro* reconstitution of chromatin

Chromatin was reconstituted on both plasmid dsDNA and ssDNA prepared from phagemid (Section 2.7.1.9) using a crude cytoplasmic fraction derived from *Xenopus* eggs. This S150 extract was enriched in cytoplasmic components, such as histones and the DNA replication machinery, but depleted in nuclear membrane, mitochondria and ribosomes. The S150 extract assembled nucleosomes on both ssDNA and dsDNA with a physiological spacing of approximately 180 bp, which is typical for this system (Wolffe and Schild 1991).

2.7.7.1 Preparation of S150 extract from *Xenopus* eggs

The protocol is essentially that of Wolffe and Schild (1991). A female *Xenopus laevis* was primed in the morning by injection of 250 U human chorionic gonadotropin into the dorsal lymph sac. In the evening, a further 500–1000 U were injected and the eggs were collected overnight at room temperature in high salt Barth's solution (110 mM NaCl, 15 mM Tris-HCl [pH7.4], 2 mM KCl, 1 mM $MgSO_4$, 0.5 mM Na_2HPO_4, 2 mM $NaHCO_3$). The eggs were removed using a wide-mouth 25 ml pipette and stored in ice-cold high salt Barth's solution.

The Barth's saline was decanted from the eggs and the jelly coat surrounding the eggs was removed by treatment at room temperature with 20% modified Barth's solution (18 mM NaCl, 2 mM HEPES-NaOH [pH7.5], 0.2 mM KCl, 0.5 mM $NaHCO_3$, 0.15

mM $MgSO_4$, 50 µM $Ca(NO_3)_2$, 0.1 mM $CaCl_2$), supplemented with 2% [w/v] cysteine adjusted to pH8.0 with 5 M NaOH immediately before use. The eggs were allowed to sit in the cysteine solution for 5 min with periodic gentle swirling so that all the eggs were resuspended in the liquid. The dissociation of the jelly coats was followed by closer packing of the eggs, at which point the eggs were washed thoroughly with ice-cold 20% modified Barth's/cysteine solution and damaged or discoloured eggs removed. All further procedures were carried out at 4^{o}C. Eggs were subsequently washed with extraction buffer (50 mM HEPES-KOH [pH7.4], 50 mM KCl, 5 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 0.5 mM PMSF) and excess buffer was removed. Eggs were then carefully packed into 5.1 ml tubes for a Beckman SW 55 rotor and centrifuged at 9k rpm for 30 min to remove yolk platelets. Following centrifugation the central brown cytoplasmic layer was removed and clarified by recentrifuging at 9k rpm for 30 min. The supernatant was then centrifuged at 40k rpm (150k x*g*) for 60 min to sediment nuclear membranes, mitochondria and ribosomes. The clear, yellow upper fraction (the S150 egg extract) was carefully aspirated, snap-frozen in liquid nitrogen and stored at -70^{o}C.

2.7.7.2 *In vitro* reconstitution of chromatin

Single-stranded phagemid DNA or double-stranded plasmid DNA was reconstituted with chromatin, consisting of nucleosomes with physiological spacing, at a concentration of 2.5 µg/µl, routinely in a final reaction volume 100µl, following a protocol described by Rodríguez-Campos et al. (1989). The S150 *Xenopus* egg extract did not exceed 60% of the final reaction volume, otherwise the reconstitution of chromatin was inefficient.

A reaction mixture of final volume 100µl contained 250 ng ssDNA or dsDNA, 10 µl assembly salts x10 (200 mM HEPES-NaOH [pH7.0], 10 mM $MgCl_2$), 3 mM ATP, 40 mM creatine phosphate and water to a final volume of 100µl. DNA replication on the ssDNA template could be monitored by supplementing the reaction with 10 µCi of [α-^{32}P] dATP (3,000 Ci/mmol, 10 mCi/ml), in addition to 0.5 mM for each of the unlabelled dNTPS, so that the newly synthesised DNA was internally labelled. The reaction mixture was incubated at 27^{o}C for 4 hr (sometimes 2 hr), after which the extent of chromatin reconstitution was assayed immediately by one of the methods described below (Section 2.7.7.3).

2.7.7.3 Assays for chromatin reconstitution

The structure of chromatin reconstituted on plasmid DNA was established by treatment of the chromatin assembly reaction mixture with either staphylococcal nuclease or *Msp* I:

Digestion with Staphylococcal nuclease

Chromatin reconstitution reactions were supplemented with $CaCl_2$ to a final concentration of 3 mM, and the digest was initiated by the addition of *Staphylococcus aureus* nuclease (Boehringer Mannheim GmbH) at a concentration of 20 U/µg DNA, diluted in buffer (10 mM Tris-HCl [pH7.4], 1.25 mM $CaCl_2$). The DNA was digested at 37°C and aliquots were removed at suitable time points (typically 0, 2, 5, 15 and 30 min). The reaction was stopped by the addition of 0.25 vol. of stop solution (2.5% [w/v] *N*–lauryl sarcosine, 100 mM EDTA) and the RNA digested with 100 µg RNase A at 37°C for 30 min. Subsequently, 0.2% SDS and 200 µg proteinase K were added and the reaction incubated at 55°C for 30 min. After phenol/chloroform extraction and ethanol precipitation, the DNA was separated on a 1.3% agarose gel, with orange G loading buffer (Section 2.4) as the marker. The gel was run until the orange G was two-thirds of the way down the gel.

Digestion with *Msp* I restriction enzyme

For *Msp* I digestion, an aliquot of the chromatin reconstitution reaction (typically containing 250 ng DNA) was digested with the appropriate restriction enzymes (typically 50 units/µg DNA) to linearise the DNA or to release the methylated patch (refer to Section 6.3.3). The sample was then diluted with *Msp* I buffer (50 mM NaCl; 10 mM Tris-HCl [pH 7.4]; 10 mM $MgCl_2$; 1 mM DTE) to a final concentration of 1 ng DNA/µl and digested with *Msp* I (Promega), 10 units/µg DNA) at 37°C. Aliquots, typically of volume 40 µl, were removed at suitable time points (typically 0, 2, 5, 15, 30 and 60 min). Samples were then processed as described above.

CHAPTER THREE

In vitro transcription of class III and class II genes

3.1 Introduction

The work presented in this thesis investigates the inhibitory action of DNA methylation and chromatin on gene transcription. Two systems to assay *in vitro* transcription are used in subsequent sections. In the first system, the *in vitro* transcription of two tRNA genes is presented in Chapter 5 (see Section 2.7.4.2 for the protocol of this assay). These genes are class III genes, transcribed by RNA polymerase III (pol III; reviewed in White 1994). The genes are for $tRNA^{Arg}$ and $tRNA^{Leu}$, which are part of a characterised cluster of tRNA genes (Bourn et al. 1994). The pArg/Leu plasmid template was constructed by subcloning the $tRNA^{Arg}$ and $tRNA^{Leu}$ genes into the pUC19 vector, as described in Section 2.5.1 (also refer to Figure 2.1). The second system examines transcriptional initiation from the SV40 promoter by using the technique of primer extension (Section 2.7.4.3). The SV40 promoter is part of the pVHCk plasmid, which contains a chloramphenicol acetyltransferase (CAT) reporter gene (refer to Figure 2.2), transcribed by the RNA polymerase II (pol II) transcriptional machinery (reviewed in Buratowski 1994). This chapter discusses the two systems, and the modifications to existing protocols that had to be made to optimise *in vitro* transcription. These optimisations had to be made before the transcription from the two different templates, pArg/Leu and pVHCk, could be examined. In addition, the structure and function of tRNA gene promoters (Section 3.2.1) and the SV40 promoter (Section 3.3.1) are discussed briefly.

Much of the understanding of eukaryotic transcription has been built on data from *in vitro* transcription experiments. The biochemical analysis of the function and complex structure of RNA polymerases has involved the addition or removal of highly purified or recombinant transcription factors to transcription assays. *In vitro* transcription of cloned genes in cell-free extracts can provide some information on the interactions between factors (reviewed in Roeder 1991), but this approach needs to be complemented by molecular biology techniques. These can be used to identify and purify regulatory factors, initiation factors and polymerase sub-units, and to clone the genes that encode these *trans* -acting factors and therefore provide the means to manipulate the factors by reverse genetics (Jackson 1993). Several methods have been described in the literature for the preparation of transcription-competent nuclear extracts from cells in tissue culture (Weil et al. 1979; Manley et al. 1980; Dignam et al. 1983; Shapiro et al. 1988). Tissue-specific *in vitro* transcription can be studied if transcriptionally-competent nuclear extracts are prepared from highly-differentiated tissues (reviewed in Sierra et al. 1993). A highly active pol II transcription system has also been developed that uses *Drosophila* embryo extracts (for example refer to Kadonaga 1990). The nuclear extracts that have been used for the *in vitro* transcription experiments presented in this thesis were prepared by the protocol of Dignam et al. (1983); refer Section 2.7.4.1. These extracts were depleted of histone H1 before they were used for the transcription of the pVHCk template (Chapter 6).

3.2 *In vitro* transcription of class III tRNA genes

The pol III transcription system that is the basis of experiments in Chapter 5 is discussed in detail in the following section. In addition, certain relevant aspects of the formation of initiation and elongation transcription complexes on a methylated template are discussed in Section 5.3.6. The transcription of eukaryotic tRNA genes has been reviewed elsewhere (Sprague 1995). The other genes that are transcribed by pol III all encode a variety of small RNA molecules (reviewed in White 1994). These include 5S rRNA, which is a component of the large ribosomal subunit, and small pol III transcripts that are

encoded viral class III genes. For example, adenovirus encodes the VA_I and VA_{II} transcripts, which subvert the translational machinery of an infected cell to the efficient production of viral proteins.

3.2.1 tRNA genes, pol III transcription and DNA methylation

Transfer RNA molecules are 70 to 90 nucleotides in length. They function as the adaptor molecules that translate the genetic information carried by mRNA into a particular order of amino acid residues in a protein. A tRNA molecule is able to recognise a particular amino acid and match this to the correct codon in the message. There are 50 to 100 distinct tRNA species in most eukaryotic cells, with the relative abundance of a species correlating to the usage frequency of the codon (reviewed in Sharp et al. 1984). The haploid human genome contains 10 to 20 copies of each of about 60 different tRNA genes, to give a total of approximately 1300 genes. The genes that encode tRNA often occur in complex multigene families dispersed throughout the eukaryotic genome. In higher eukaryotes, tRNA genes are organised into clusters (Rosenthal and Doering 1983). Some human tRNA genes occur in clusters of two, three or four genes (for example, refer to Shortridge et al. 1989). Bourn et al. (1994) have isolated and characterised a genomic clone from human placenta that contains the tRNA genes for lysine, glutamine, leucine and arginine in a cluster of 2 kbp, as well as a gene for $tRNA^{Gly}$ that is within 4 kbp. I have sub-cloned these $tRNA^{Arg}$ and $tRNA^{Leu}$ genes into the pUC19 vector, to construct the pArg/Leu template for transcription (see Section 2.5.1).

The promoters of most class III genes include intragenic elements, termed internal control regions (ICRs). Transfer RNA genes, as well as most class III genes, have type II ICRs that consist of two essential and highly conserved domains. The A-block (also known as the 5' ICR) has the consensus sequence TGGCNNAGTGG and encodes the D-loop region of the tRNA gene product. The B-block (or 3' ICR) has the consensus sequence GGTTCGANNCC and encodes the TΨC-loop. The high degree of conservation of these blocks therefore reflects their dual role in promoter function and tRNA structure. The sequence of the $tRNA^{Arg}$ and $tRNA^{Leu}$ genes is shown in Figure 3.1,

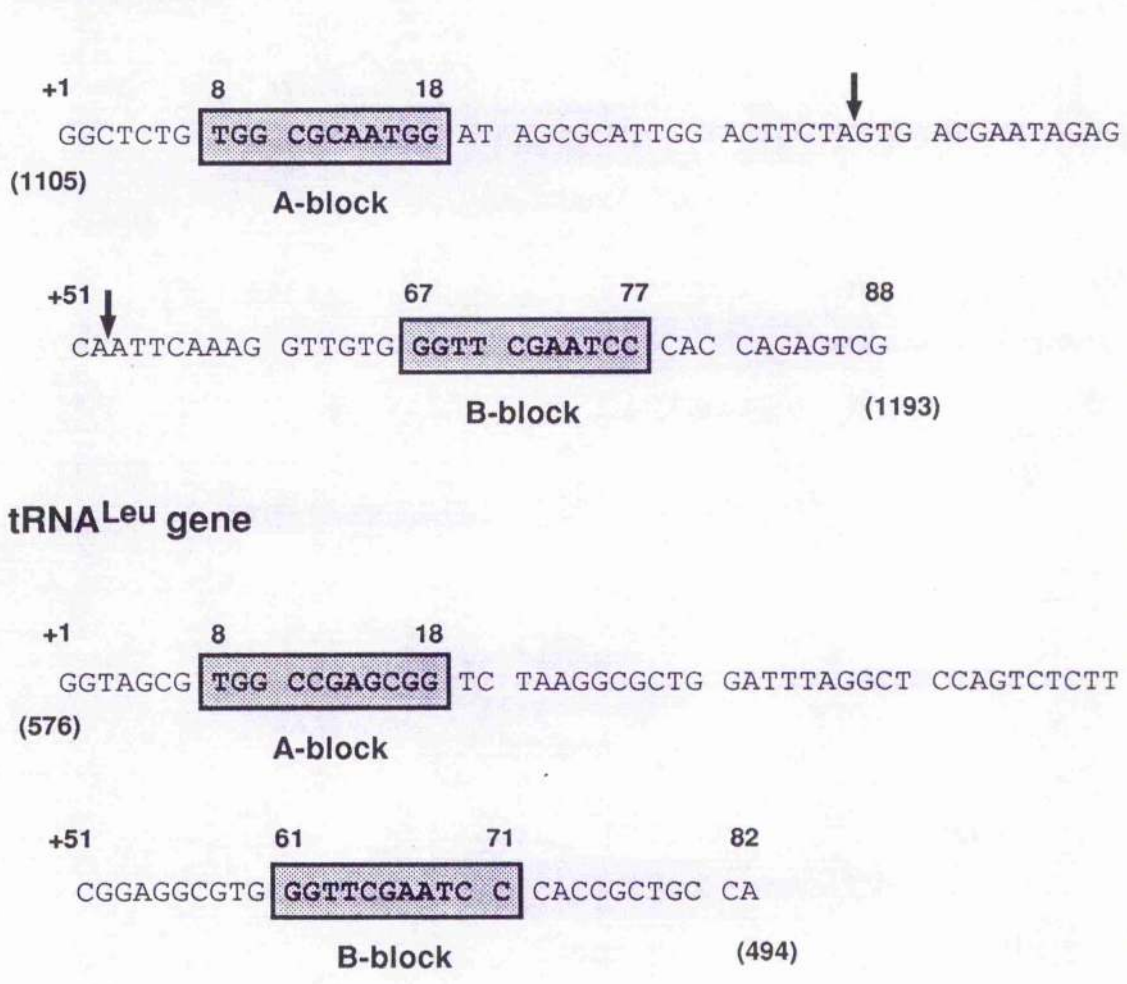


Figure 3.1

Sequences of the tRNA^{Arg} and tRNA^{Leu} genes

Coding strand sequences of the two tRNA genes that were transcribed in pol III transcription assays are shown. The extent of the sequences correspond to the nascent transcript, beginning at the +1 position and finishing at the site of termination, as indicated by the numbers above the sequences. The positions of the A-block (consensus sequence TGGCNNAGTGG) and B-block (consensus sequence GGTTCGANNCC) are shown for each gene. The expected splice sites for the tRNA^{Arg} precursor are indicated in the DNA sequence by the two arrows. Refer to Bourn et al. (1994) for further details. The positions of these coding regions in pArg/Leu (see Section 2.5.1 and Figure 2.1) are indicated by the numbers in brackets under the sequence.

with the positions of A-blocks and B-blocks for each gene highlighted by the boxes. The separation between A- and B-blocks can vary from 30 to 60 bp without affecting the efficiency of transcription. This is necessary because extra arms in tRNA molecules can vary in length and some tRNA genes contain introns (reviewed in Westaway and Abelson 1995). For example, the $tRNA^{Arg}$ gene shown in Figure 3.1 contains an intron of 15 bp (Bourn et al. 1994). Introns have been found in the tRNA genes of *Archaebacteria* , fungi and many eukaryotic organisms, although the best characterised system is that of the yeast, *Saccharomyces cerevisiae* (Westaway and Abelson 1995). The splicing mechanisms that remove introns are diverse, but in all cases the intervening sequences are removed from the precursor or nascent tRNA (pre-tRNA) to form the functional product. The expected splice sites for the $tRNA^{Arg}$ gene in the figure are indicated by arrows, and the splicing of the gene product is discussed further in Section 3.2.2. Many other steps are involved in the maturation of pre-tRNAs, including the removal of the 5' leader and 3' trailing sequences (Deutscher 1984), tRNA base modification (Björk et al. 1987) and the addition of the CCA sequence to the 3' end of the tRNA .

The ICR binds the polypeptides that make up the TFIIIC transcription factor fraction (Gabrielsen and Sentenac 1991). TFIIIC is one of the three traditional fractions that are obtained after purification of transcription factors from yeast, that together are sufficient to reconstitute tRNA transcription *in vitro* . The other two fractions are TFIIIB and polymerase. Regulation of tRNA genes is necessary for the adaptation of the tRNA population to the different frequencies of codon usage and amino acid utilisation in different cell types. Variations in the 5' flanking sequences of tRNA genes from yeast, silk-worms, mice and humans (for example, refer to Gonos and Goddard 1990) have been shown to affect the *in vitro* transcription of these genes (reviewed in Sprague 1995). A recent study has shown that the 5' flanking sequence negatively modulates the *in vivo* expression of a human tRNA gene (Tapping et al. 1993). The protein components of the TFIIIB fraction, including the TATA-binding protein (TBP), are presumed to bind to 5' flanking promoter elements which would allow tissue-specific and developmental regulation of expression (Sprague 1995).

DNA methylation can inhibit the expression of class III genes (see Sections 1.7 and 1.8.3). Methylation has been shown to inhibit the transcription of a methylated $tRNA^{Leu}$ gene, after the micro-injection of a construct containing this gene into *Xenopus* oocytes (Besser et al. 1990). Inhibition was only observed if the CpG dinucleotides in the B-block were methylated, but not if the methylated sites were elsewhere. In contrast, an oocyte 5S rRNA gene was not affected by methylation. The expression of the adenovirus type 2 VA_I gene is inhibited by methylation at CpG sites, either after transfection into mammalian cells or in a pol III *in vitro* transcription system (Jüttermann et al. 1991). Transfer RNA genes are perfect examples of housekeeping genes and, in general, they are associated with CpG islands (see Section 1.6). For example, the genomic clone that was isolated and characterised by Bourn et al. (1994) contained an extensive GC-rich region associated with the cluster of tRNA genes (see Figure 2.1B). Genomic tRNA sequences in human leukocytes lack CpG suppression and are unmethylated at certain sites (Schorderet and Gartler 1990).

Since these studies were published, the effect of methylation on class III gene expression has been seldom investigated. I decided to investigate the expression of tRNA genes for this reason, but also because the pol III transcription assay (described in Section 2.7.4.2) has many advantages as a system. Most importantly, it is relatively straight-forward to perform, gives labelled transcription products of a high specific activity and is quite resilient to RNase contamination.

3.2.2 *In vitro* transcription of the pArg/Leu template

The protocol for the pol III *in vitro* transcription assay (Section 2.7.4.2) has been developed and optimised for tRNA genes by Gonos and Goddard (1990), and in a later study by Bourn et al. (1994). The important points of this protocol are that the optimal Mg^{2+} concentration for this transcription system is 3.5 mM, creatine phosphate is included in the reaction mixture to allow the regeneration of ribonucleoside triphosphate levels and the saturating level of template is approximately 150 ng, although transcription can be detected with as little as 10 ng of template. Sierra et al. (1993) describe a similar protocol that uses the G-free cassette to analyse pol II transcription. My own

improvements on the original pol III transcription assay are as follows. Firstly, the nuclear extract was preincubated with the template for 20 min before transcription was initiated by the addition of ribonucleoside triphosphates (discussed in more detail in Section 3.2.3) and an RNase inhibitor was added to the reaction mixture (Section 2.7.4.2). The preincubation of template with extract stimulates transcription (Section 3.2.3). Secondly, a nuclear extract depleted of histone H1 was used for some transcription assays. The preparation of this extract is described in Section 2.7.4.1, and transcription assays that use it are described in Sections 5.3.3, 6.2.3 and 6.3.4.

The transcription products for the pArg/Leu template are shown in Figure 3.2, together with the transcription products from plasmid constructs containing only the $tRNA^{Arg}$ gene or only the $tRNA^{Leu}$ gene. Unlike the $tRNA^{Leu}$ gene, the $tRNA^{Arg}$ gene contains an intron (see Section 3.2.1). Nascent RNA, that is transcribed from the $tRNA^{Arg}$ gene, is processed by the nuclear extract from an original size of approximately 100 nucleotides, to a mature tRNA of size 85 nucleotides. During processing, half-molecule intermediates of size 35–40 nt are formed and subsequently joined together by an RNA ligase (Westaway and Abelson 1995). An assay time of 60 min was chosen because the yield of mature tRNA was greatest after this time (results not shown). Shorter times of incubation maximised the yield of the processing intermediates. Nascent RNA that is transcribed from the $tRNA^{Leu}$ gene is also processed to a mature tRNA of size 75 nucleotides. There is also a small amount of longer (approximately 150 nucleotides) transcript, which could be caused by read-through at the usual termination signal following the gene, and subsequent termination at a T-rich sequence 70 nt downstream (Bourn et al. 1994). To simplify the interpretation of subsequent results, only the major products of transcription (i.e. nascent and processed RNA) were used for quantitative analyses and for subsequent figures. Preliminary experiments showed that the degree of processing was highly reproducible between samples, and that processing was not affected by the methylation status of the template (results not shown). However, the pattern of nascent and processed products did vary between different batches of nuclear extract (results not shown). For this reason, a series of transcription assays was always performed with a particular nuclear extract.

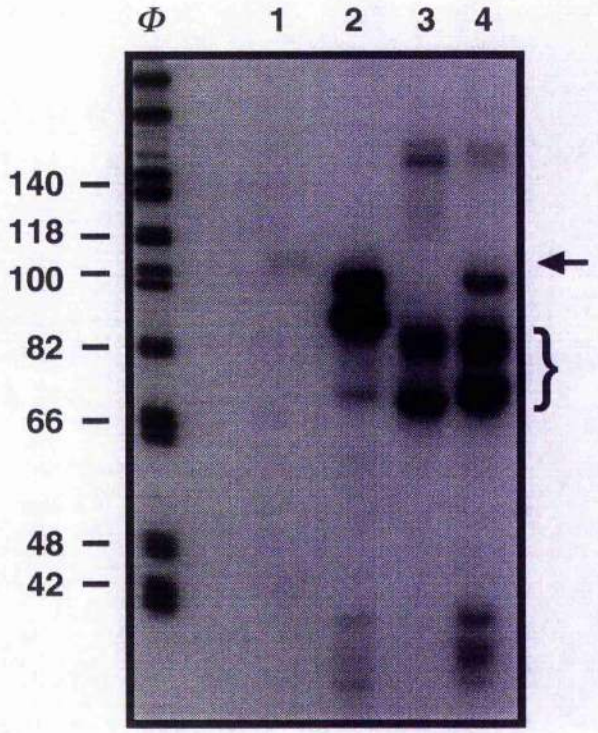


Figure 3.2

***In vitro* transcription of a tRNA^{Arg} gene, a tRNA^{Leu} gene and both of these genes in the pArg/Leu template:**

Transcription assays were performed to determine the transcription products from the pArg/Leu template. **1:** transcription in the absence of added DNA template, showing a small amount of an unknown endogenous product (indicated by the arrow); **2:** transcription from a plasmid construct containing only the tRNA^{Arg} gene; **3:** transcription of tRNA^{Leu} gene; **4:** transcription of pArg/Leu plasmid template. Refer to Section 3.2.2 for the identification of the transcription products. The transcription products used for subsequent figures and quantitative analyses are indicated by the brace. The marker Φ is denatured, end-labelled ΦX174 RF form dsDNA digested with *Hin*fI, with the approximate size of fragments indicated in nucleotides (refer to Section 2.6 for further details).

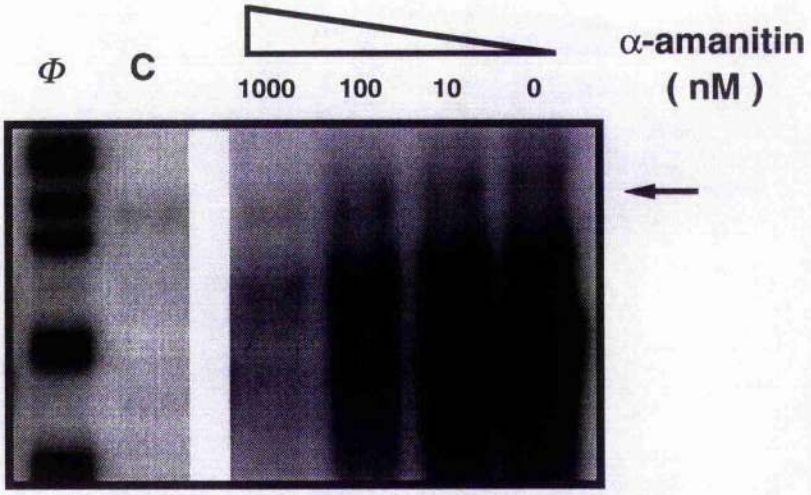


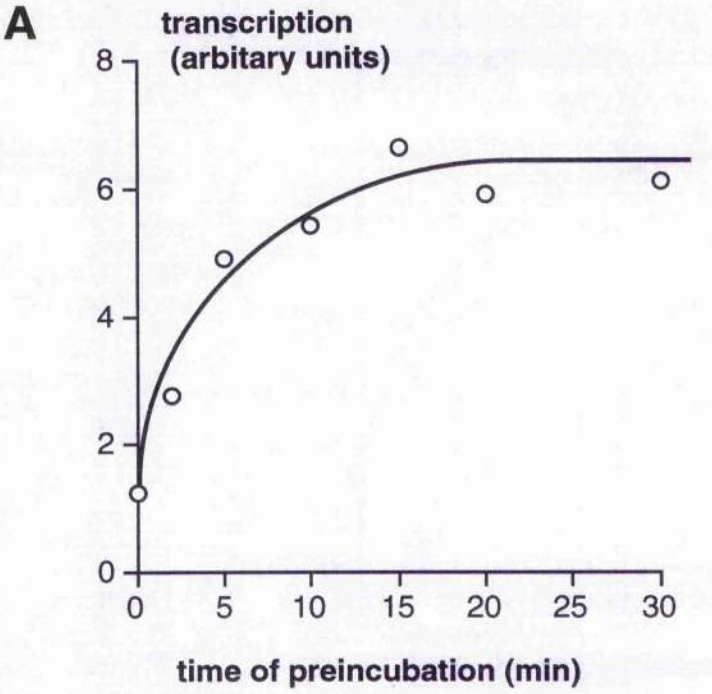
Figure 3.3

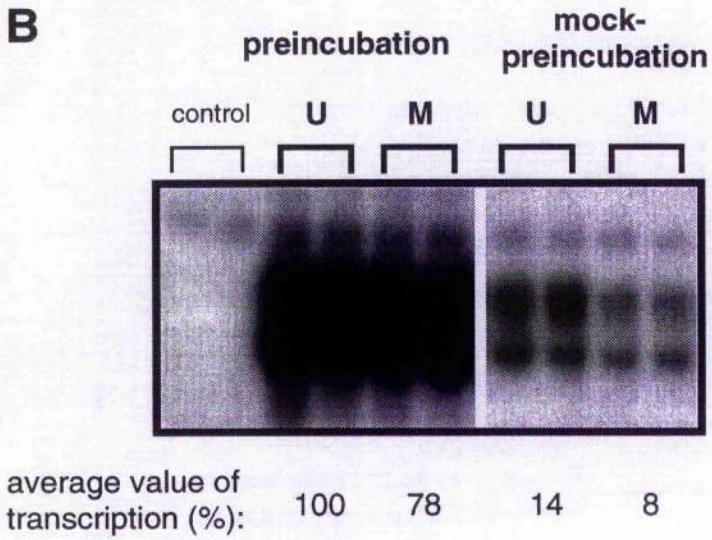
Inhibition of pol III transcription by α-amanitin:
In vitro transcription of the unmethylated pArg/Leu template was assayed at the indicated concentrations (from 1000 to 0 nM) of the RNA polymerase III inhibitor α-amanitin. The control lane C is an assay performed in the absence of DNA template. The arrow indicates an artifactual band that is present in all lanes, irrespective of the α-amanitin concentration. The transcription product bands are smeared due to contamination by RNase. See Section 3.2.2 for further details.

In the absence of added DNA template, the nuclear extract also supports the formation of a small amount of unknown product (Figures 3.2 and 3.3). Artefacts of this kind tend to occur if the radio-label for the transcription assay is UTP (Sierra et al. 1993). The problem can be alleviated by using GTP as the radio-label. However, this possibility was not investigated because the artefact was minor and consistent for all the samples of an experiment, irrespective of the type or amount of template. The enzymatic activity is not connected with transcription, because neither low nor high concentrations of α-amanitin inhibit the formation of the endogenous product (Figure 3.3). High concentrations of α-amanitin (1 μM) are known to inhibit both RNA polymerase II and III (Schwartz et al. 1974), which is clearly seen for the pol III assay in Figure 3.3. Low concentrations of α-amanitin (10 nM) inhibit pol II, but do not affect pol III or the formation of the endogenous product.

3.2.3 Preincubation of the template with nuclear extract enhances transcription

Preincubation of the nuclear extract with the template was found to be a necessary step in the protocol for optimal *in vitro* transcription (Figure 3.4A). The template was preincubated with extract at 30°C for times varying from 0 to 30 min. The reaction mixture of all samples was exposed to 30°C for exactly 30 min, irrespective of the time of addition of the template. This ensures that any inhibitory effect on the pol III transcriptional machinery because of exposure to 30°C is consistent between samples, so that any inhibitory effect does not mask the stimulatory effect of preincubation. After the period of preincubation, transcription was initiated by the addition of ribonucleotides to the complete reaction mixture. The preincubation time of 20 min is consistent with the known kinetics of initiation complex assembly from other systems e.g. the *Drosophila* transcription system (Kadonaga 1990). The graph of the time-course experiment (Figure 3.4A) shows that a preincubation time of 20 min is sufficient for optimal levels of transcription, for this particular system. Preincubation stimulates transcription by about six-fold, compared to the level of transcription in the absence of any preincubation before the initiation of transcription. Some transcription is possible in the absence of





average value of transcription (%):	100	78	14	8

Figure 3.4

The effect on transcription of pArg/Leu of preincubation of nuclear extract with DNA template :

A: the unmethylated pArg/Leu template was preincubated with nuclear extract for various times (from 0 to 30 min), followed by the initiation of transcription by the addition of NTPs (Section 2.7.4.2). Transcription for each sample was performed in duplicate, and the level quantified by using a phosphoimager. Units of transcription are arbitary units of phosphostimulable luminesence (PSL).

B: a sample gel, showing the effect of preincubation of nuclear extract with 100 ng unmethylated (U) and methylated (M) pArg/Leu template. Samples were either preincubated or mock-preincubated (see Section 3.2.3 for details) with nuclear extract for the optimal time of 20 min, as indicated. The level of transcription was quantified for each assay by using a phosphoimager, and the average value for each set of duplicates is shown. The standard 100% value is the average transcription level of unmethylated preincubated template.

preincubation because initiation complex formation can form during the elongation period of the transcription assay, after the NTPs have been added to the reaction mixture. The elongation period for this system is 60 min.

The effect of preincubation is also seen in Figure 3.4B. Unmethylated and methylated pArg/Leu template were preincubated with nuclear extract for the optimal time of 20 min. Mock-preincubations for the same templates were also carried out, for which the reaction mixtures were exposed to 30°C for 20 min in the absence of template. The template was then added with the NTPs at the initiation of transcription. The preincubation treatment stimulates transcription six to ten-fold, but is less effective for the methylated template. The final level of transcription from the methylated template is also lower, being 80% of that from the unmethylated template. This preferential inhibition of the methylated template is discussed further in Chapter 5; for example, refer to Figure 5.5.

3.3 *In vitro* transcription of the class II SV40 promoter

The pol II transcription system that is the basis of several experiments in Chapter 6 is discussed in the following section. The transcription of eukaryotic class II genes has been reviewed in detail elsewhere (Buratowski 1994), and will not be discussed except in general terms. The initiation of transcription from the SV40 promoter in the pVHCk plasmid template (Figure 3.5) was assayed by using a primer extension approach (see Section 2.7.4.3 for a description of the protocol). The principle of primer extension involves the hybridisation of an end-labelled oligonucleotide to RNA, that has been transcribed from the reporter gene, followed by extension of the oligonucleotide primer using the enzyme reverse transcriptase. The extension reaction terminates at the extreme 5' end of the RNA, and can therefore be used to determine the start point of transcription of an mRNA sequence. Protocols for primer extension can be found elsewhere (Sambrook et al. 1989; Leonard and Patient 1991; Mason et al. 1993). The SV40 promoter and pol II transcription are discussed in Section 3.3.1. Preliminary experiments to optimise this pol II transcription system are described in Sections 3.3.2 and 3.3.3. Further optimisations of the primer extension protocol are described in Section 3.3.4, and are concerned with the optimal amounts of template, end-labelled primer and the concentration of Mg^{2+} that is required for *in vitro* transcription in this system.

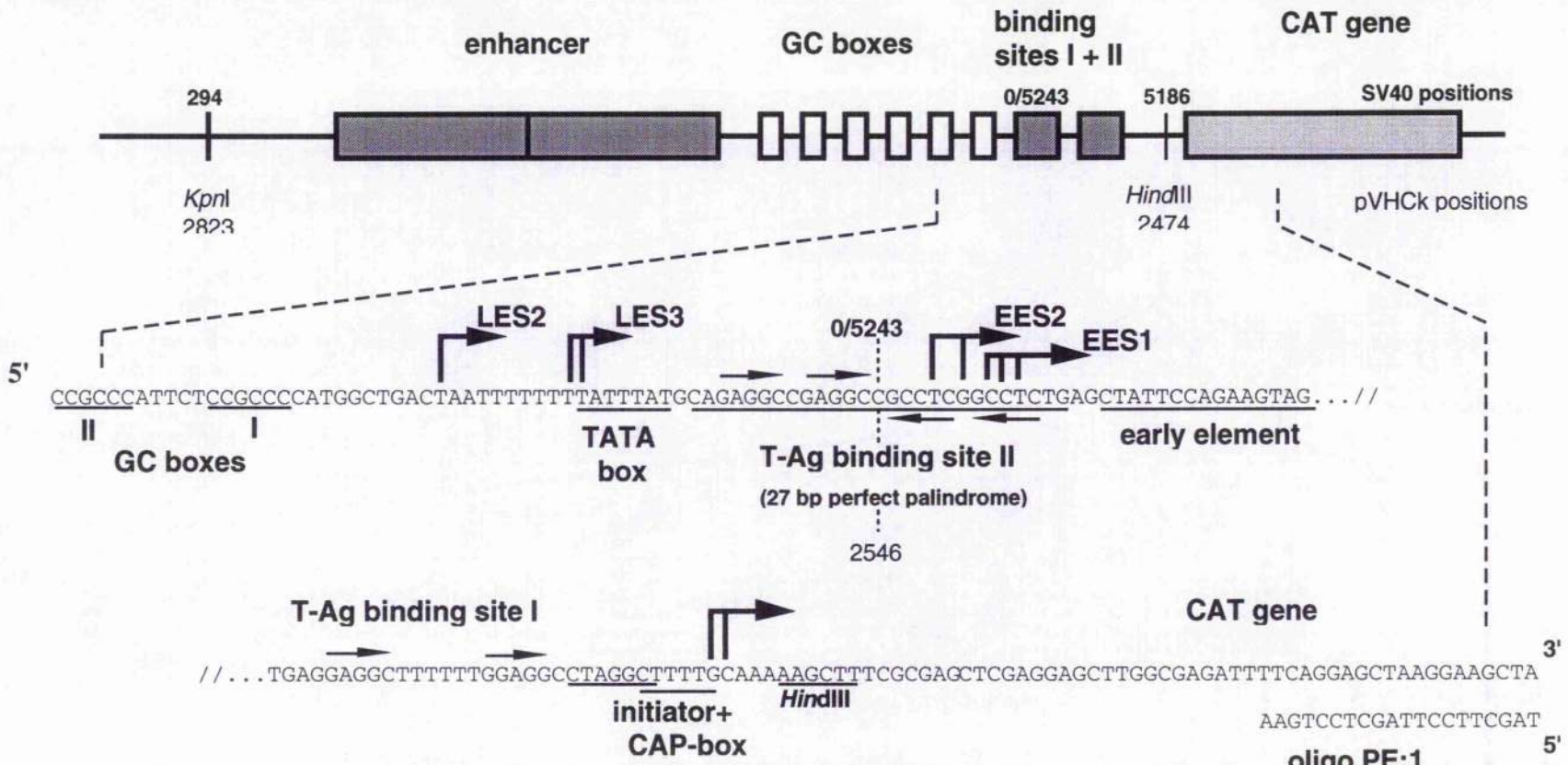


Figure 3.5 Map of the SV40 promoter in the template pVHCk:

The SV40 early promoter and functional elements are shown, including positions of the promoter in plasmid pVHCk. The major start sites of transcription (thick arrows) are marked, which include immediate early (EES) and late early (LES) start sites. Recognition pentanucleotides in T-antigen binding sites I and II are shown by thin arrows. Oligo PE;1 is a primer used for primer extension (see Sections 2.7.4.3 and 3.3.4 for further details). It was used to detect transcipts from the EE and LE start sites, as well as those that appeared to initiate further downstream, from an initiator and a CAP-box sequence. The sizes, in nucleotides, of primer extension products are shown below the oligonucleotide.

3.3.1 The SV40 promoter and pol II transcription

The structure of the SV40 promoter in pVHCk is shown in Figure 3.5, and a complete circular map of this template is shown in Figure 2.2 for reference. Figure 3.5 is based on figures in several articles on the structure and function of the SV40 promoter (Barrera-Saldana et al. 1986; Zenke et al. 1986; Fanning and Knippers 1992). The sequence of part of the SV40 promoter and the adjacent chloramphenicol acetyltransferase (CAT) reporter gene in pVHCk has been confirmed (see Section 3.3.4 and Figure 3.6 for further details), and is incorporated into Figure 3.5. The figure shows the location of the late early start sites and immediate early start sites of transcription, large T-antigen binding sites and a TATA box.

The plasmid pVHCk was constructed from previous mammalian plasmid expression vectors (Bryans et al. 1992; Kass et al. 1993; also refer to Section 2.5.1). These were developed to study the transient expression of the CAT reporter gene after transfection of the vector into mammalian cells in culture (reviewed in Docherty and Clark 1993). This thesis describes the use of this same vector as a template for *in vitro* transcription. The technique of nuclease S1 protection was used by Chambon and colleagues to determine the transcription start sites of SV40, following *in vitro* transcription with HeLa nuclear extract (for example, refer to Barrera-Saldana et al. 1986). Although these studies were qualitative, both nuclease S1 protection and primer extension can be used in quantitative transcription assays (reviewed in Mason et al. 1993). Kass, previously of our laboratory, has quantitated transcripts from pVHCk by using primer extension (unpublished results), and I have modified his original protocol for this study (see Section 3.3.4).

The inhibition of pol II transcription by DNA methylation, through either a direct or an indirect mechanism in a variety of systems, has already been discussed in detail (see Sections 1.7 and 1.8). However, one study is of particular relevance: Götz et al. (1990) describe the inhibition by methylation of transcription from immediate early sites of SV40. A reporter construct was used that contained a thymidine kinase (TK) reporter gene, driven by the SV40 promoter. The construct was micro-injected into *Xenopus*

oocytes, and transcriptional inhibition due to methylation was assayed by primer extension. The pattern of major primer extension products is consistent with transcription from both the immediate early and late early sites, but minor products suggest that the initiation of transcription is not totally specific for these sites. A similar pattern of major and minor products is seen for *in vitro* transcription of pVHCk, followed by primer extension of transcripts specific for the CAT reporter gene (see Figure 3.6).

3.3.2 Nuclease S1 protection assay

A necessary preliminary experiment before quantitative primer extension can be used as an assay of transcription is to determine that a significant proportion of the primer forms a hybrid with the mRNA (see Section 2.7.4.3). This is achieved by treating hybridisation reactions with an excess of nuclease S1. The proportion of the primer which is resistant to nuclease S1 is determined by expressing the radioactivity of the samples treated with nuclease S1 as a percentage of the radioactivity of the control samples (with no nuclease S1). Table 3.1 shows the proportion of resistant primer annealed to the transcripts from either unmethylated or methylated pVHCk. The control value is for the proportion of resistant primer observed when 500 ng unmethylated pVHCk template was mock-transcribed (i.e. the template was taken through all the steps of a normal transcription assay, except that nuclear extract was not included during *in vitro* transcription). This value may indicate that some of the DNA template is not degraded by the DNase I treatment after *in vitro* transcription, and could therefore anneal to the primer. However, this interpretation is unlikely because an excess of DNase I is used and a large margin of error is introduced during the washing of filters (see Section 2.7.4.3).

Since 50 fmol of end-labelled primer (equivalent to 10^4 cpm) is used in the primer extension reaction, the data from the protection assay for 500 ng of unmethylated pVHCk template shows that 1.8 fmol target RNA is produced during the transcription assay. The amount of the template is 151 fmol, so transcription is initiated on 1.2% of the templates. This calculation assumes that *in vitro* transcription is limited to one round of initiation and elongation, although it is known that two rounds of transcription can be completed in HeLa nuclear extract (Hawley and Roeder 1987). A previous study by Kadonaga (1990)

has shown that transcription is initiated on 3% of the templates, when RNA synthesis was limited to one round in the presence of Sarkosyl (see Section 5.3.6 for further details of the effect of Sarkosyl on the formation of initiation complexes).

template:	background	control	unmethylated pVHCk (500 ng)	methylated pVHCk (500 ng)
proportion of annealed primer (%):	0.2	0.6	3.8	2.7

Table 3.1

Proportions of primer that anneal to target RNA:
The end-labelled oligonucleotide that is used for primer extension was annealed to the transcripts formed during *in vitro* transcription of unmethylated or methylated pVHCk template, as indicated. The proportion of primer that annealed to the target RNA was determined by a nuclease S1 protection assay, as described in Section 3.3.2. The control was 500 ng unmethylated pVHCk template taken through all the steps of a normal transcription assay, except that *in vitro* transcription was performed in the absence of nuclear extract. All protection assays were performed in duplicate. The background level is for the protection seen for assays performed in the absence of pVHCk template. The sample variance for protection assays that were not digested with nuclease S1 is a measure of the error introduced during the washing of filters (see Section 2.7.4.3); this value is 0.4%.

3.3.3 Assay for RNase contamination

Primer extension is sensitive to RNase contamination, since cleavage of the mRNA will cause the premature termination of the cDNA product. A second preliminary experiment was therefore to determine that RNase levels were kept at a minimum throughout the transcription assay. Precautions against RNase contamination were taken (Section 2.7.4) and nuclear extracts, which are the major source of RNase, were supplemented with RNasin (Promega). End-labelled RNA was prepared by treating marker RNA (0.28–6.58 kb, Promega) with calf intestinal alkaline phosphatase (Section 2.7.1.12), followed by [γ-^{32}P] ATP and T4 polynucleotide kinase, as described in Section 2.7.1.18 for the end-labelling of oligonucleotides. The labelled RNA was diluted to 100 ng/μl, with a

final specific activity of 10^3 cpm/µl. The integrity of the labelled marker was checked on a denaturing gel (Section 2.7.4.2), and the bulk of the RNA was treated as described below.

Each reaction mixture was spotted onto discs of DE81 paper and washed as described in Section 2.7.4.3. Intact, labelled RNA remained bound to the filters. However, 5' exonuclease or endonuclease activity degraded the RNA into small enough fragments so that the label could not bind to the filters, and was removed during the washing procedure. This assay can test for RNase contamination, but it is only a qualitative indicator of this activity because a large margin of error is introduced during the washing procedure. The results are tabulated in Table 3.2.

	control	RNase	RNase + inhibitor	nuclear extract	extract + inhibitor
proportion of intact RNA (%):	100	9	86	35	95

Table 3.2

Assay for contamination of nuclear extract by RNase:
RNase contamination was assayed as described in Section 3.3.3. The proportion of intact end-labelled RNA was measured after treatment with RNase A or nuclear extract, as indicated. RNase activity could be inhibited by the addition of RNasin inhibitor to the reaction mixture (see Section 3.3.3). The control for the 100% standard is the radioactivity retained on filters after the labelled RNA has not been exposed to RNase activity. The sample variance for the control is 13%.

Aliquots of labelled RNA (1.0 µl) were treated either with 10 pg RNase A diluted in 9 µl RNase-free buffer D (the buffer used for the dialysis of nuclear extracts; see Section 2.7.4.1), or with RNase supplemented with 40 U RNasin to test the efficiency of this inhibitor. The definition of one unit activity of RNasin is the inhibition of 5 ng RNase by 50%. Reaction mixtures were incubated at 42°C for 30 min, and performed in duplicate. Table 3.2 shows that this amount of RNasin can inhibit the action of RNase, so that minimal degradation occurs. Labelled RNA was also treated either with 9 µl

nuclear extract, or with extract supplemented with 40 U RNasin. Although the extract does degrade the RNA, this appears to be prevented by RNasin. The control for the 100% standard is the radioactivity retained on filters after the labelled RNA has been incubated with buffer D alone. Eight control assays were performed, with a sample range of 13%.

3.3.4 Primer extension of transcripts from the *in vitro* transcription of pVHCk

Typical gels of primer extension products, for the transcripts formed during *in vitro* transcription of pVHCk, are shown in Figure 3.6 A and B. The products which assay transcription from the SV40 immediate early start sites (EES1 and EES2) are marked by arrows. In the sequencing gel (Figure 3.6 A) these products are not resolved clearly, but are major products of the expected size of 110 and 116 nt in Figure 3.6 B. Transcription from the late early start sites (LES2 and LES3) is assayed by the levels of a minor product of size approximately 140 nt for both types of gel. A product of size 56-58 nt. may arise from non-specific initiation of transcription at an initiator sequence and consensus sequence for a CAP-box (see Figure 3.5). The consensus sequence of the CAP-box is CTTYTG (Baralle and Brownlee 1978), and is often found just downstream of an initiation and capping site in genes that lack a TATA box (Locker 1993). A possible initiator element may be CCTAGGCT (refer to Figure 3.5), which has a weak resemblance to the initiator consensus sequence of CTCANTCT (Smale et al. 1990). A minor product of size approximately 78 nt is also seen in these gels, and a later experiment (see Figure 5.11). Other minor products probably arise from RNase degradation of the target mRNA, which would cause the premature termination of primer extension. Transcription was therefore quantitated by a summation of the levels of all primer extension products, by using a Fuji BAS1000 phosphoimager. This is justified because all of the products are labelled at the 5' end, irrespective of their size, because an end-labelled primer is used for the extension procedure. However, figures of primer extension products (Figures 5.11 and 6.11) show only those of size 56-58 nt, since these were the major products in most transcription assays.

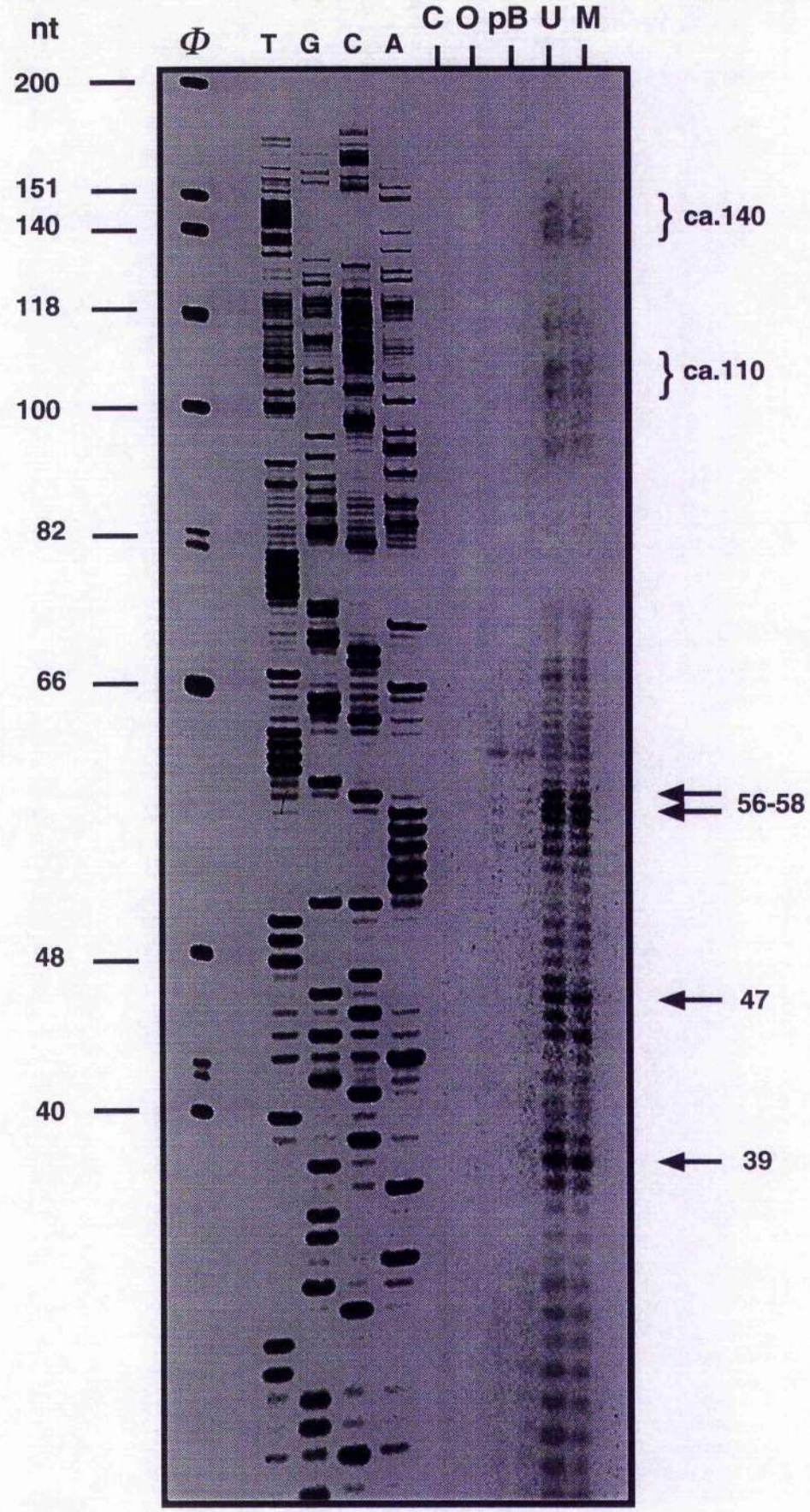


Figure 3.6 A
Primer extension products from transcripts of pVHCk:
see overleaf for figure legend

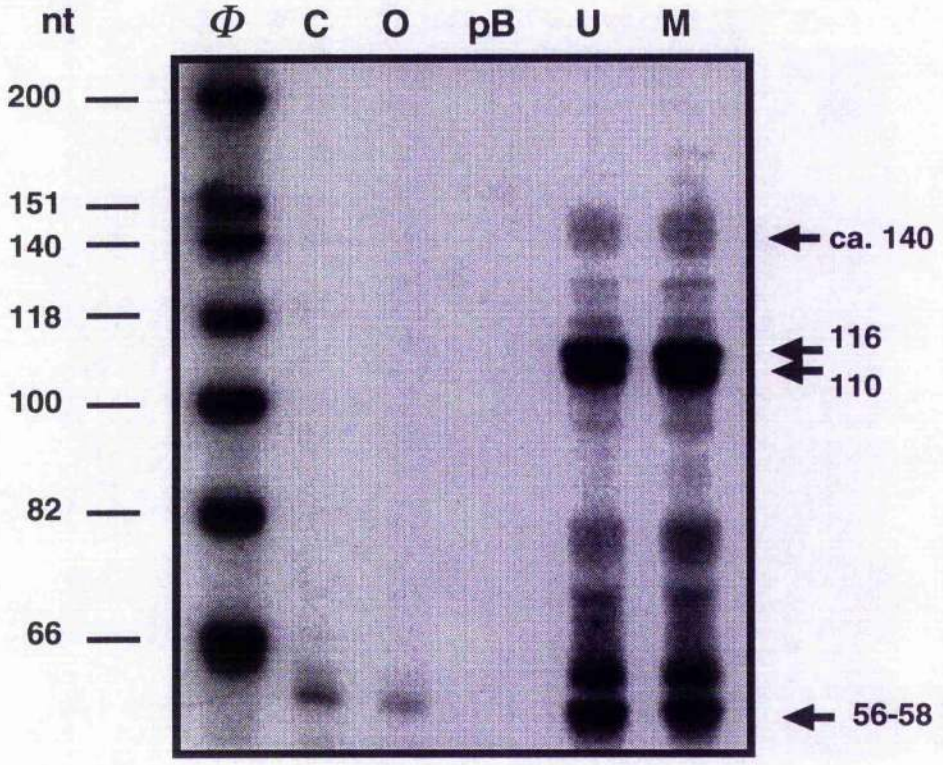


Figure 3.6 A and B:

The use of primer extension products to assay levels of *in vitro* transcription from mock-methylated (U) and methylated (M) pVHCk plasmid template:

A: Transcripts from pVHCk are quantitated by primer extension, which gives rise to major extension products (marked by arrows) and several minor products. The products were resolved on a 0.4 mm denaturing, 12% polyacrylamide sequencing gel.The size of products is determined by comparison with: *Φ* (denatured, end-labelled *Φ*X174/HinfI marker; the approximate size of denatured fragments is given in nucleotides under **nt**) and sequencing tracts **T**, **G**, **C** and **A.** The sequence was obtained by using end-labelled primer extension oligonucleotide (oligo PE;1) and double-stranded pVHCk as template. Extension products: **C**, primer only; **O**, no DNA; **pB**, 150 ng pBluescript KS-; **U**, 150 ng mock-methylated pVHCk; **M**, 150 ng methylated pVHCk.

B: the same legend as for Figure 3.6 A, except that sequencing tracts were not run on this gel. The gel was a denaturing, 12% PAGE gel of 1.5 mm in thickness, so does not resolve the primer extension products as well as the sequencing gel. The autoradiograph of this transcription assay was kindly provided by Dr. S U Kass, and is used with his permission in this thesis.

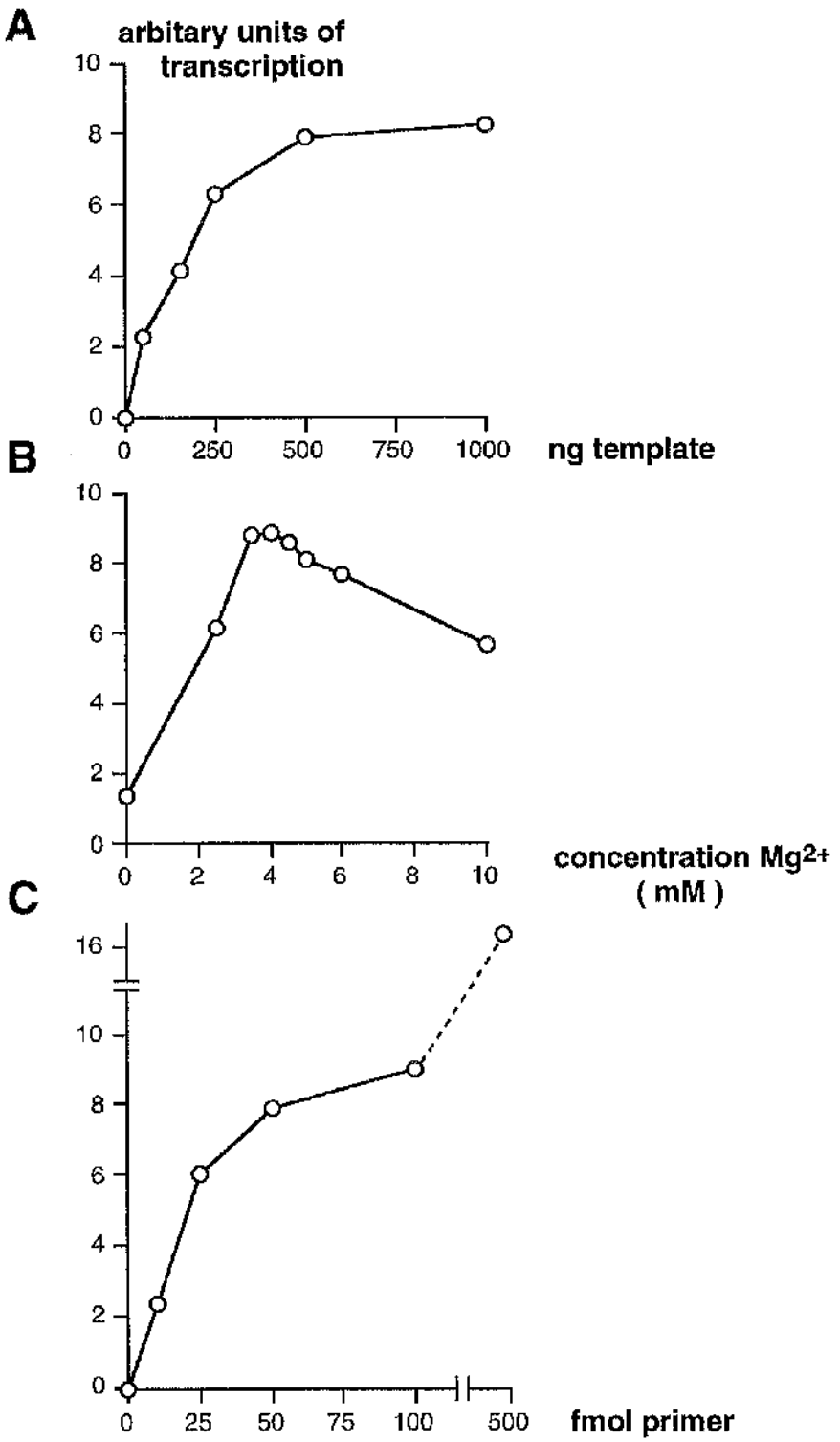


Figure 3.7

Optimisation of *in vitro* transcription from pVHCk template:
A series of experiments were performed to optimise the conditions for the *in vitro* transcription of unmethylated pVHCk template, as described in Section 3.3.4. **A:** The amount of template was varied, from 50 to 1000 ng; the optimal amount is 500 ng. **B:** The concentration of Mg^{2+} was varied; the optimal concentration is 4 mM. **C:** The amount of end-labelled primer was varied from 10 to 500 fmol, and the optimal amount is 50 fmol per primer extension reaction.

The pol II transcription assay was optimised as described below, and summarised in Figure 3.7. The optimal amount of template for this system was found to be 500 ng (graph A) and the optimal concentration of Mg^{2+} was 4.0 mM (graph B). Quantitative primer extension of transcription products requires that the primer is in an excess (at least ten-fold) molar ratio to the amount of target RNA (Mason et al. 1993). The titration experiment that is summarised in graph C varies the amount of primer from 10 fmol to 500 fmol. Nuclease S1 protection (see Section 3.3.2) has shown that 500 ng of template can produce 1.8 fmol of RNA during the course of the transcription assay. The plateau in graph C shows that all the target RNA becomes annealed to primer, at moderate levels (25–100 fmol) of the primer. At the highest level of primer (500 fmol), the apparent increase in transcription is caused by non-specific annealing during the extension reaction.

CHAPTER FOUR

Chromatin assembly *in vitro*

4.1 Introduction

A current challenge for molecular and developmental biologists is to understand the molecular mechanisms through which chromatin and chromosome structure influence gene activity. A number of methods have been developed to reconstitute chromatin *in vitro*. These have been coupled to assays that measure either access of the DNA to *trans*-acting factors or transcription *in vitro*. These studies have allowed major insights to be made into the structure and function of chromatin.

High salt concentrations dissociate nucleosome core particles into their components of DNA and histones. The first methods to assembly chromatin *in vitro* reversed this process, and purified core histones were reconstituted onto DNA by dialysis from high salt (reviewed in Dimitrov and Wolffe 1995). A typical method is described in Rhodes and Laskey (1989). Further refinements of this method of chromatin assembly have been the exchange of octamers from donor chromatin to recipient DNA at high salt (Tatchell and van Holde 1977; Drew and Travers 1985) and the use of a negatively charged third component as an assembly factor. The use of poly(glutamic acid) as a nucleosome and chromatin assembly factor has been reviewed elsewhere (Stein 1989). These methods have a serious drawback: the nucleosomes are not positioned with physiological spacing. Instead of the correct spacing of 180–200 bp found in native chromatin, the histone octamers that are reconstituted onto DNA by these methods pack together as closely as possible, with each octamer occupying 150–160 bp of DNA (reviewed in Dimitrov and Wolffe 1995). Since the octamers are closely packed together there is insufficient linker DNA for the correct binding of histone H1. Nucleosomal

arrays that are reconstituted with histone H1 by the high salt or exchange methods result in the formation of aggregates and the precipitation of DNA.

A second method of chromatin assembly was therefore chosen for the studies in this thesis, to allow chromatin to be reconstituted with histone H1. Chromatin was reconstituted on dsDNA using a crude cytoplasmic fraction derived from *Xenopus* eggs. These extracts (prepared as described in Section 2.7.7.1) were used to reconstitute chromatin *in vitro* on plasmid DNA (see Section 2.7.7.2). A typical method is described in Wolffe and Schild (1991) and the technique is discussed in Section 4.2.1. In some experiments the reaction mixture was supplemented with histone H1, added at a 0.6 w/w ratio with respect to the DNA (see Section 4.2.2). Chromatin assembled in this way is used in the studies described in Sections 6.2.3 and 6.3.4, that investigate the effect of DNA methylation and chromatin on *in vitro* transcription.

Nucleoplasmin is an acidic thermostable protein from *Xenopus* eggs that has been used as an assembly factor during chromatin reconstitution with purified octamers (Sealy et al. 1989). Two other acidic proteins called N1/N2 have been purified from eggs (Kleinschmidt et al. 1990). N1/N2 appear to act with nucleoplasmin as molecular chaperones that ensure the correct spacing of nucleosomes. *Xenopus* oocytes have been used to prepare chromatin assembly extracts (for example, refer to Shimamura et al. 1989), but the quality of oocytes is more difficult to assess as compared to the quality of eggs. Extracts from mammalian cells (Stillman 1986; Gruss et al. 1990) and *Drosophila* embryos (Kamakaka et al. 1993) have also been used for *in vitro* chromatin assembly.

Chromatin assembly *in vitro* with *Xenopus* egg extracts was chosen because the quality of eggs was easy to assess, the preparation of extracts was straight-forward (see Section 2.7.7.1) and chromatin was reconstituted on plasmid DNA with high efficiency. The egg extract was not competent for the transcription of either tRNA genes or the initiation of transcription from the SV40 promoter (refer to Sections 6.2.3 and 6.3.4). Transcription could therefore be assayed from templates reconstituted with chromatin by the use of HeLa nuclear extract, without endogenous transcription due to the egg extract to complicate the interpretation of the data.

4.2 Chromatin assembly on double-stranded DNA

Chromatin was reconstituted on pVHCk plasmid, by using *Xenopus* egg extract. In some experiments the reaction mixture was supplemented with histone H1, added at a 0.6 w/w ratio with respect to the DNA. Section 4.2.2 presents evidence that this amount of histone H1 is incorporated into the assembled chromatin at physiological levels. This observation is necessary before it can be argued that the chromatin formed on methylated DNA, in the presence of histone H1, is a valid model of heterochromatin. The case for this argument is presented in Section 6.4. The structure of chromatin reconstituted on plasmid DNA was established by treatment of the chromatin assembly reaction mixture with either staphylococcal nuclease or *Msp* I, as described in Section 2.7.7.3. The more elaborate experiments described in Chapter 6 investigate the structure and transcriptional activity of chromatin reconstituted on either fully methylated DNA (Sections 6.2.2 and 6.2.3) or DNA methylated in defined regions (Sections 6.3.3 and 6.3.4). These sections in Chapter 6 show that the chromatin reconstituted *in vitro* on methylated DNA is more resistant to digestion by *Msp*I, than is that formed on unmethylated DNA. This has been shown previously to be true for methylated DNA in nuclei and for methylated DNA transfected into mammalian cells (see Section 1.8.2), but not for an *in vitro* system. This section will discuss the digestion of chromatin with staphylococcal nuclease, and will not discuss *Msp* I digestion any further.

The digestion of assembled chromatin with staphylococcal nuclease results in oligonucleosome fragments, which were deproteinised and the DNA resolved on agarose gels (see Figures 4.1 and 4.2). The DNA was visualised by a Southern blotting procedure (see Section 2.7.1.21), in all respects identical to the procedures used for the series of *Msp* I fade-out assays described in Chapter 6. The DNA was hybridised to the *Pvu* II-*Pvu* II restriction fragment labelled as fragment d in Figure 6.1. Nucleosomal ladders were easier to see on hybridised filters than on agarose gels stained with ethidium, because the input amount of DNA was kept quite low to conserve the stocks of *Xenopus* egg extract.

4.2.1 Chromatin assembly on unmethylated or methylated pVHCk

Chromatin reconstituted on mock-methylated or fully-methylated pVHCk shows no difference in the size of the spacing between nucleosomes (Figure 4.1). In both cases, the spacing is 180 bp, decreasing to the expected 146 bp for nucleosome core DNA after extensive nuclease digestion. The spacing value was determined by scanning the lanes in the autoradiograph of Figure 4.1 in a densitometer, and comparing the positions of oligonucleosomes with those of marker fragments in a semi-log plot (results not shown). The same average spacing is also obtained from the densitometry traces in Figure 4.3 B. This repeat length is typical for the chromatin formed with *Xenopus* egg extracts (Dimitrov and Wolffe 1995), and for the 180–200 bp repeat length seen in the native chromatin of normal somatic cells (Finch et al. 1975; Wolffe 1992).

4.2.2 Chromatin assembly on unmethylated or methylated pVHCk, in the presence of histone H1

Chromatin was also reconstituted on unmethylated or methylated pVHCk in the presence of histone H1 (Figure 4.2). The level of histone H1 was chosen after a series of empirical experiments as that which increased the nucleosomal spacing to the greatest extent. This level is an histone H1:DNA w/w ratio of 0.6 (i.e. 150 ng of histone H1 was added to a reaction mixture containing 250 ng DNA), which approximates the physiological level of histone H1 in chromatin. The physiological level is calculated by assuming that 25 nucleosomes can form on pVHCk, which is 5025 bp in size. Each nucleosome associates with one molecule of histone H1, which is indicated as a molar ratio of histone H1 to nucleosome of 1 in Figure 4.2. For convenience, the molar ratio is represented as H1/n. An H1/n molar ratio of 1 corresponds to an histone H1:DNA w/w ratio of 0.22, if the molecular mass of histone H1 is taken to be 26 kDa. A w/w ratio of 0.60 therefore is equivalent to a molar ratio of 2.7 molecules of histone H1 per nucleosome. A previous study has shown that the incorporation of histone H1 into

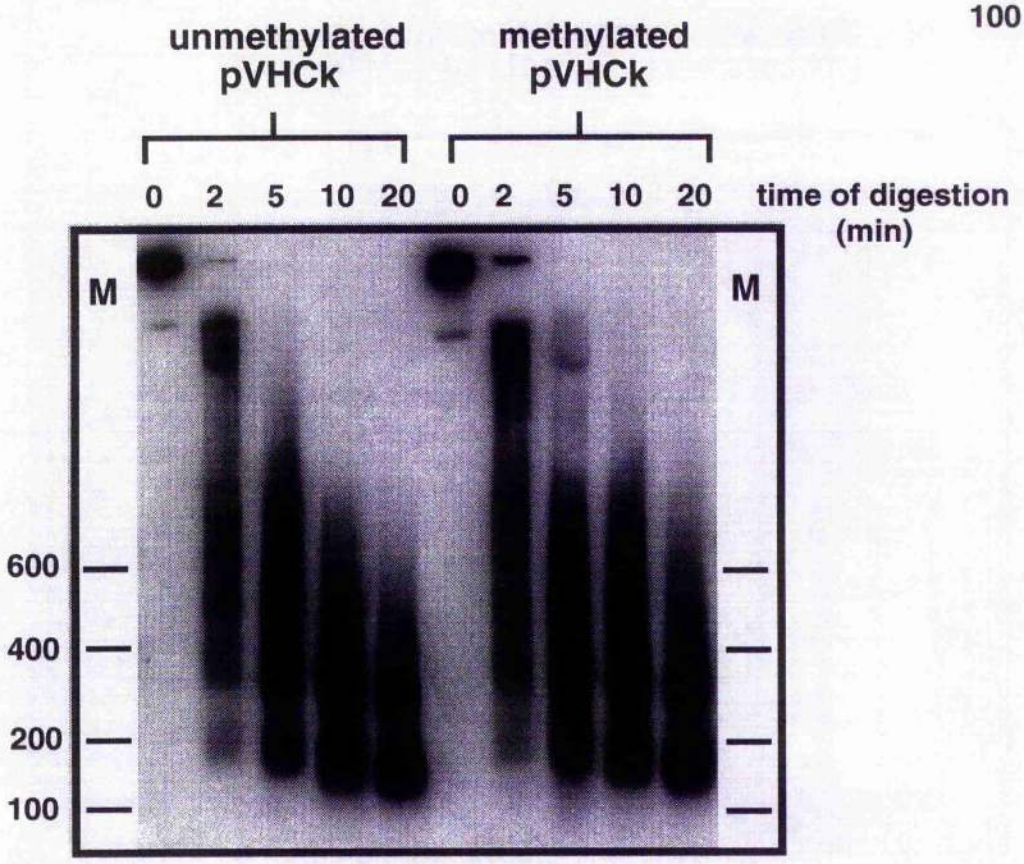


Figure 4.1

Chromatin reconstituted on unmethylated and methylated pVHCk and digested with staphylococcal nuclease:
Chromatin was reconstituted on unmethylated and methylated pVHCk using *Xenopus* S150 egg extract (see Section 2.7.7.2 for further details). The formation of nucleosomes was assayed by limited digestion of the chromatin by staphlococcal nuclease (20 U/µg DNA; see Section 2.7.7.3) for the indicated times. Oligonucleosome fragments were resolved on a 1.3% agarose gel, transferred to a nylon membrane and hybridised to restriction fragment d from pVHCk (refer to Figure 6.1 for a map of pVHCk, and the position of restriction fragments). The figure above shows that the nucleosomes are spaced with an average distance of 180 bp. The marker M indictes the position of 100, 200, 400 and 600 bp fragments from a 100 bp marker (Section 2.6).

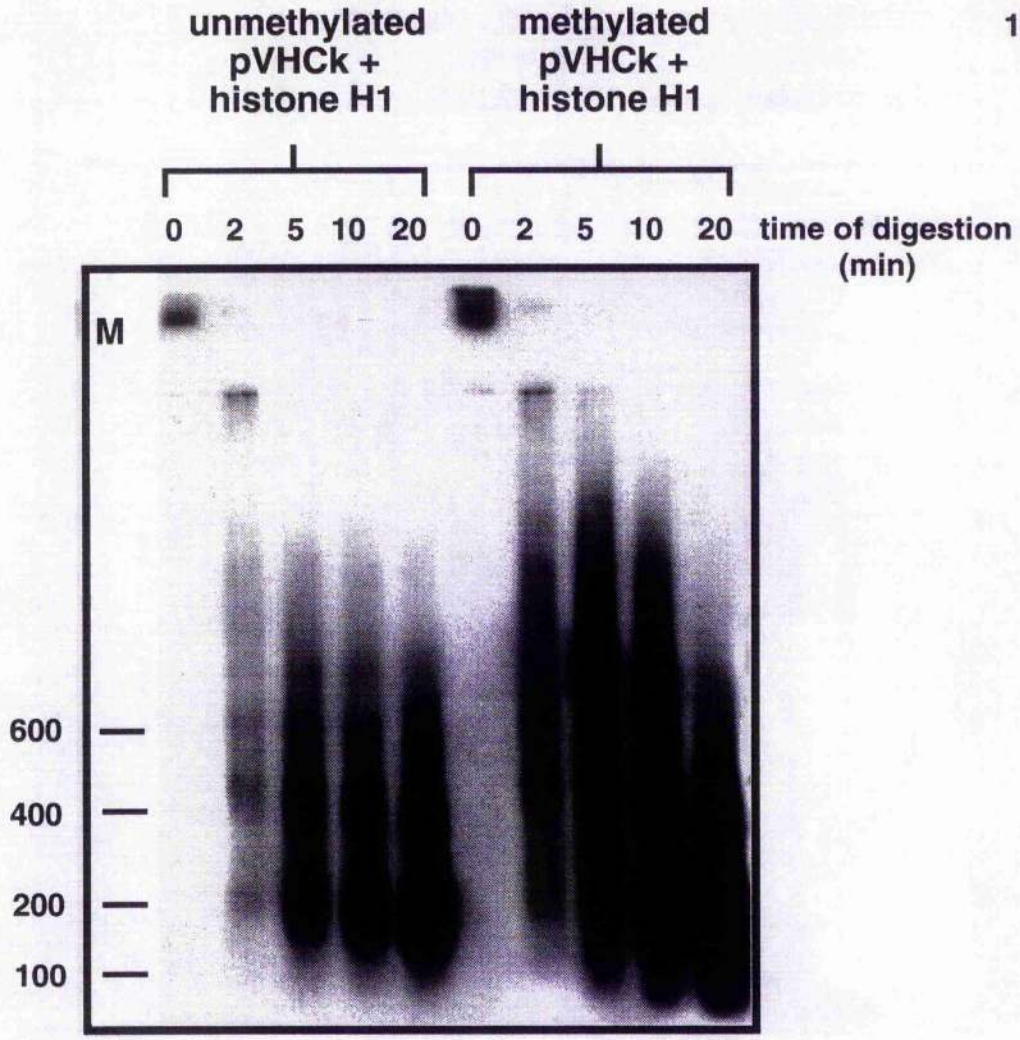


Figure 4.2

Chromatin and histone H1 reconstituted on unmethylated and methylated pVHCk and digested with staphylococcal nuclease: Chromatin was reconstituted on unmethylated and methylated pVHCk using *Xenopus* S150 egg extract, prepared as described in Figure 4.1, and supplemented with histone H1 at an histone H1:DNA w/w ratio of 0.6. The position of the oligonucleosome fragments indicate that the nucleosomes are spaced with an average distance of 210 bp in this system. The marker M indictes the position of 100, 200, 400 and 600 bp fragments from a 100 bp marker (Section 2.6).

chromatin assembled *in vitro* using *Xenopus* oocyte extracts is most efficient at an H1/n molar ratio of between 2 and 5 (Rodríguez-Campos et al. 1989). Furthermore, this input level of histone H1 is required for a physiological level of histone H1 in the assembled chromatin. For example, an input ratio of H1/n of 3 gives rise to a ratio of 0.8 of total recoverable histone H1 from assembled chromatin. The level of histone H1 that is used in the chromatin assembly experiments described in this thesis is therefore consistent with the study by Rodríguez-Campos et al. (1989). These levels approximate the physiological level of one histone H1 molecule per nucleosome.

Histone H1 that supplements a chromatin assembly reaction at a histone H1:DNA w/w ratio of 0.6 increases the average nucleosome spacing from 180 to 210 bp (Figure 4.2), which is also consistent with the study of Rodríguez-Campos et al. (1989). The increase in spacing can be used to confirm the incorporation of histone H1 into chromatin. This is clearly seen in the autoradiograph of Figure 4.3A, and the corresponding densitometry plots shown in the Figure 4.3B. There is an equal increase in repeat length for both unmethylated and methylated pVHCk DNA. Very high levels of input histone H1 (H1/n=10 to 50) either aggregate with the DNA, which prevents the formation of chromatin and subsequent digestion with staphylococcal nuclease, or cause a closed chromatin structure to form that is resistant to the concentration of nuclease used in this experiment. Early studies on the properties of chromatin showed that chromatin that contained histone H1 was five- to ten-fold more resistant to micrococcal nuclease than chromatin that did not contain histone H1 (Noll and Kornberg 1977; Allan et al. 1981). Hydrodynamic data from these studies suggested that the increase in resistance was a steric effect of histone H1 occupying regions of linker DNA, which is the target of the nuclease in native chromatin, rather than a result of aggregation or compaction of the chromatin.

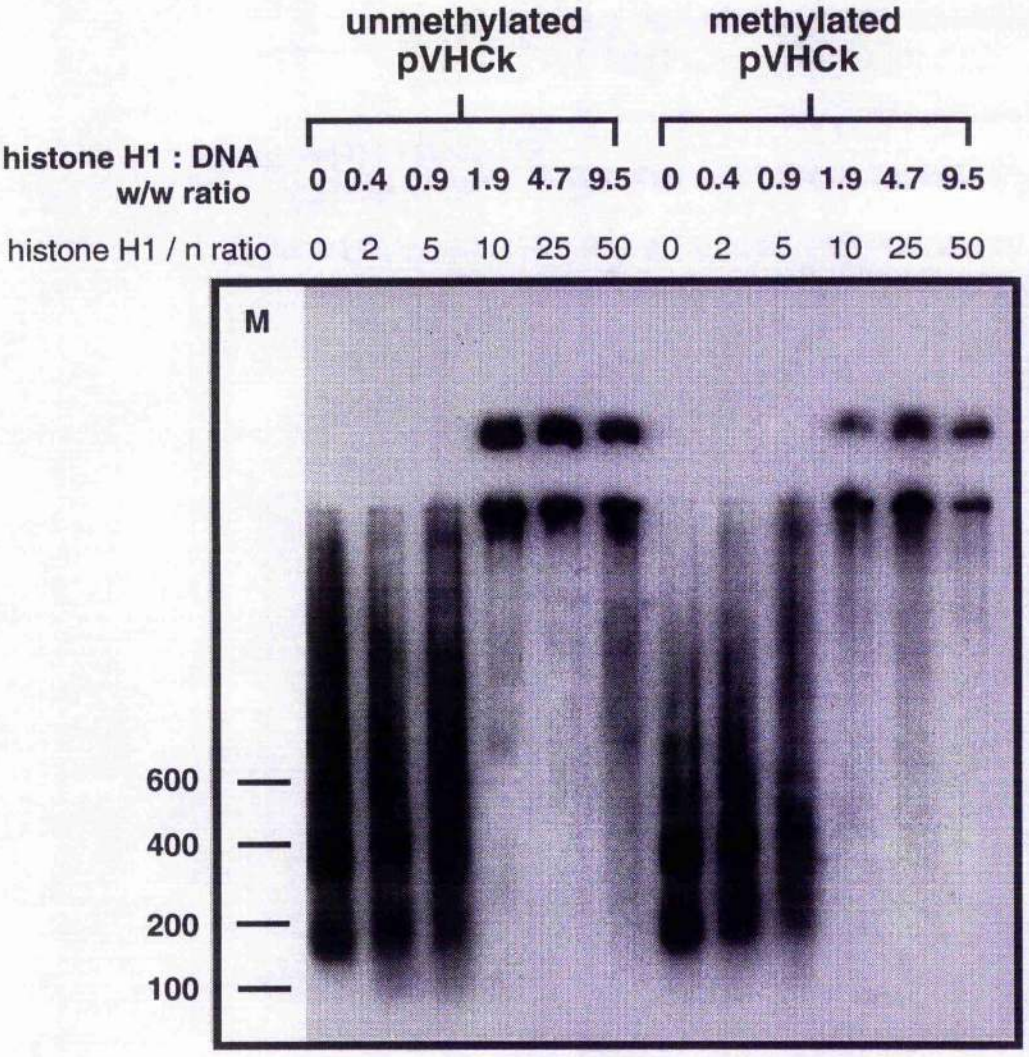


Figure 4.3 A

The effect of increasing levels of histone H1 on nucleosomal spacing of chromatin reconstituted on unmethylated and methylated pVHCk plasmid DNA:

Chromatin was reconstituted on unmethylated and methylated pVHCk plasmid DNA in the presence of increasing levels of histone H1 and subjected to limited digestion with staphlococcal nuclease. Digestions were for 5 min at a nuclease concentration of 20 U/μg DNA. Oligonucleosomal fragments were then processed as described in Figure 4.1. Histone H1 was added at an increasing histone H1: nucleosome molar ratio (H1 / n ratio), which corresponds to the indicated histone H1:DNA w/w ratios. Increasing levels of histone H1 increase the nucleosomal spacing from 180 to 210 bp (refer to Figure 4.3 B for further details). The marker M indicates the positions of 100, 200, 400 and 600 bp fragments of a 100 bp marker (Section 2.6).

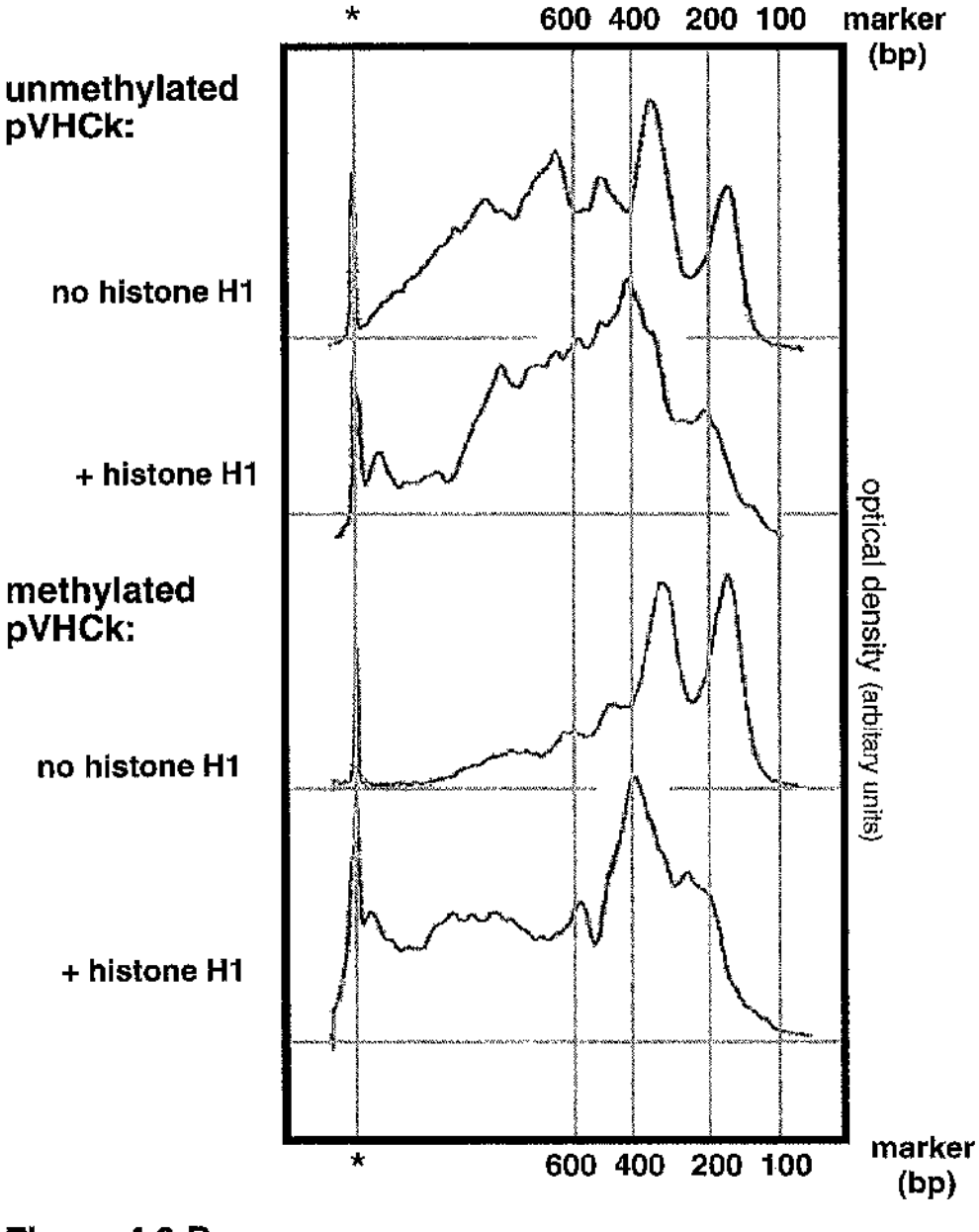


Figure 4.3 B

Nucleosomal spacing of chromatin reconstituted on unmethylated or methylated pVHCk, in the absense or presence of histone H1:

Lanes in the autoradiograph shown in Figure 4.3A were scanned in a densitometer and the plots of optical density are shown in this figure. Nucleosomal spacing increases from 180 bp to 210 bp when histone H1 is present during chromatin reconstitution. Plots for chromatin reconstituted on unmethylated or methylated pVHCk are compared in the absence of histone H1 (no histone H1) and in the presence of a histone H1:DNA w/w ratio of 0.4 (+histone H1), as indicated. The positions of 600, 400, 200, and 100 bp fragments of a 100 bp marker are indicated by the vertical grey lines. The horizontal grey lines indicate the baseline for each scan. The asterisk (*) marks the position of a calibration mark on the autoradiograph to allow the alignment of each scan.

4.3 Chromatin assembly on single-stranded phagemid DNA

Chromatin can be assembled on replicating DNA *in vitro*. *Xenopus* egg extract can use single-stranded DNA (ssDNA) as a template for complementary strand synthesis, with concomitant nucleosome assembly on the resulting double-stranded DNA (dsDNA) (Almouzni and Méchali 1988; Almouzni et al. 1990). Although chromatin assembly does not occur on ssDNA in preference to dsDNA, the process of replication appears to enhance chromatin assembly relative to non-replicating dsDNA.

Chromatin was assembled on pVHCk (+) single-stranded DNA, that was prepared from phagemid (see Section 2.7.1.9 for the protocol). Chromatin assembly on ssDNA is described in Section 2.7.7.2, with the incorporation of [α-^{32}P] dATP into the replicated strand used to monitor the progress of DNA synthesis (Figure 4.4A). The labelled products were deproteinised and resolved on an agarose gel. The products include double-stranded nicked DNA and relaxed covalently-closed circular DNA which cannot be resolved from the nicked form. The relaxed molecules undergo negative supercoiling during nucleosome assembly, to form the fast-moving supercoiled form seen in Figure 4.4A. This is due to the activity of a topoisomerase in the extract. Intermediate topoisomers should form in between the relaxed and supercoiled molecules, but are not formed because the assembly of chromatin is presumed to be very efficient under these conditions (Almouzni et al. 1990; Wolffe and Schild 1991). Reconstitution of chromatin on double-stranded pVHCk was included as a control (Figure 4.4A), and as a marker for supercoiled dsDNA. The plasmid also became labelled during the incubation, but this is presumed to be due to the incorporation of the dATP label into nicks of relaxed molecules that were then supercoiled. That chromatin was assembled on replicated DNA is confirmed by the formation of a nucleosomal ladder after digestion with staphylococcal nuclease (Figure 4.4B).

This is the most efficient method for the assembly of chromatin *in vitro* , but was not chosen for the studies in this thesis because the patch-methylated constructs of Chapter 6 (Section 6.3.1, in particular) could only be analysed as double-stranded molecules.

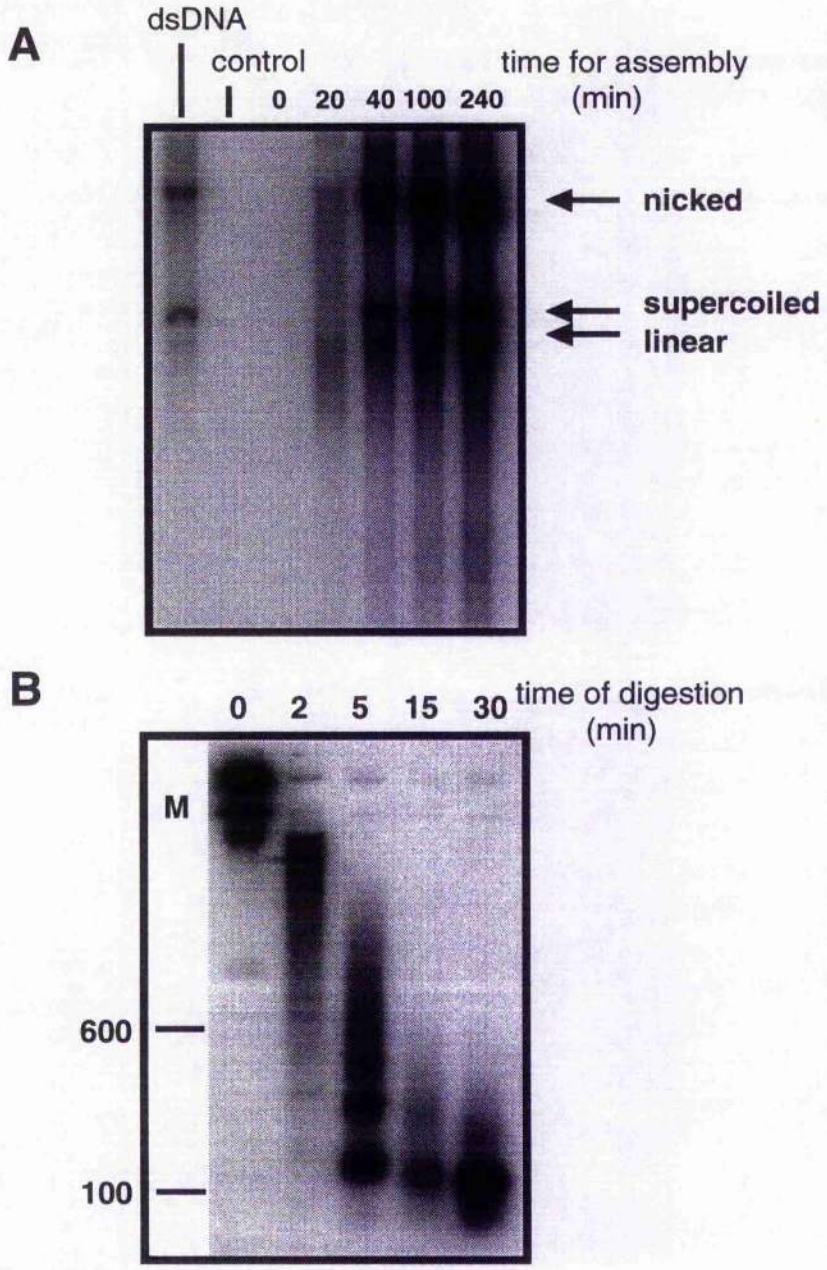


Figure 4.4

Chromatin reconstitution on single-stranded phagemid DNA:
Chromatin was reconstituted on pVHCk (+) single-strand DNA, prepared from phagemid as described in Section 2.7.1.9. The reconstitution reaction was performed in the presence of [α-^{32}P] dATP, as described in Section 2.7.7.2. See Section 4.3 for further details.

A: Chromatin was reconstituted on 50 ng ssDNA for the indicated times. The DNA was then deproteinised and resolved on a 1% agarose gel before autoradiography. The positions of nicked dsDNA, supercoiled dsDNA and linear dsDNA are indicated. Double-stranded pVHCk (50 ng) was incubated in a reconstitution reaction for 240 min (dsDNA), and the control assay was a reaction in the absence of DNA.

B: Approximately 100 ng of dsDNA, reconstituted with chromatin as described above, was digested with staphylococcal nuclease at a concentration of 20 U/μg DNA for the indicated times. Oligonucleosomal fragments were resolved on a 1.3% agarose gel. The positions of 600 and 100 bp fragments of a 100 bp marker (lane M) are indicated by the marks.

CHAPTER FIVE

The effect of histone H1 and DNA methylation on *in vitro* transcription

5.1. Introduction

The following experiments describe the formation of complexes of histone H1 on unmethylated and methylated DNA, and the effect of these complexes on *in vitro* transcription of genes transcribed by RNA polymerase III or RNA polymerase II. The aim of this study was, firstly, to characterise an experimental system that modelled inactive chromatin and then, secondly, to use this system to determine the effects of DNA methylation, increasing levels of histone H1 and depletion of histone H1 on the transcription of the reporter genes. Complexes of H1 with either unmethylated or methylated DNA were used as models for inactive chromatin, since the interactions between histone H1 and either DNA alone or DNA in chromatin appear to be similar (see Discussion, Section 5.4).

5.2. The formation and analysis of histone H1 and DNA complexes

Histone H1:DNA complexes are an attractive system to work with because the starting materials for complex formation are inexpensive and readily obtained, and the complexes are easily formed under suitable ionic conditions (see Section 2.7.5.2). Acid-extracted preparations of histone H1 from calf thymus were renatured as described in Section 2.7.5.1, and were then assayed for purity and the ability to form complexes on naked DNA. Histone H1 is commercially available as relatively pure preparations (Figure 5.1),

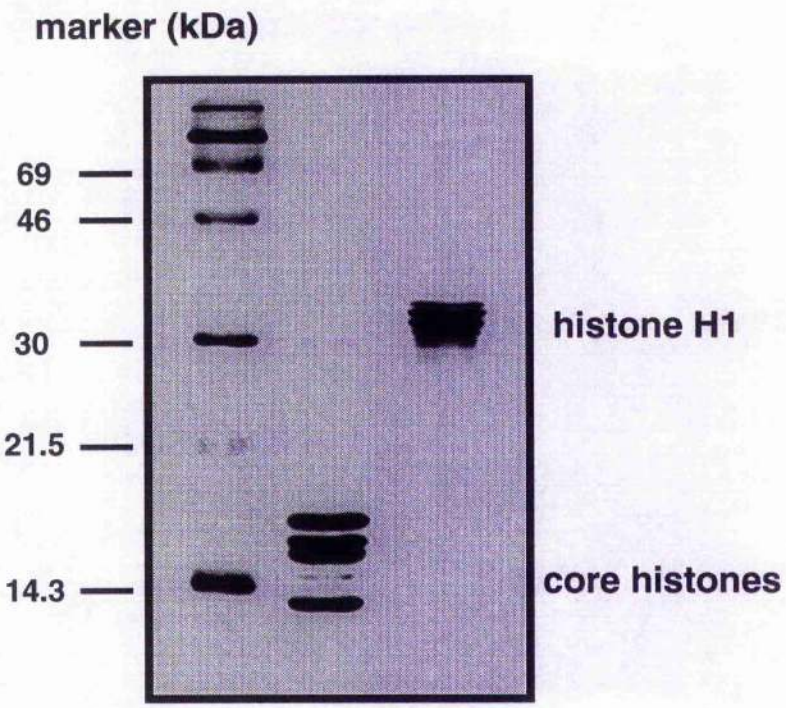


Figure 5.1

SDS-PAGE of core histones and histone H1:
The purity of preparations of core histones (see Section 2.7.6) and acid-extracted total histone H1 from calf thymus (see Section 2.7.5) were assessed by SDS-PAGE, using Coomassie Blue R250 (Section 2.7.2.2) as the stain. The actual molecular masses of proteins are 11–16 kDa for core histones and 26 kDa for this preparation of histone H1. The somatic variants of histone H1 (see Section 1.7.7) are resolved on this gel, as can be seen from the multiple bands in the histone H1 lane. The protein standard is Rainbow Marker (Amersham).

although the acid extraction of the protein from tissue (Cole 1989) is quite harsh and causes some protein degradation. Commercial preparations contain several histone H1 variants (see Section 1.8.7 for further details), that can be resolved by SDS-PAGE (Figure 5.1).

5.2.1. Gel retardation analyses of histone H1-DNA complexes

The complexes on both unmethylated and methylated DNA were analysed by agarose gel electrophoresis. Complexes were formed with increasing levels of histone H1 (0.25–2.0 w/w ratio) on unmethylated supercoiled pVHCk plasmid. Aliquots were then run under electrophoresis conditions that favoured the maintenance of protein-DNA complexes. These included the use of an electrophoresis buffer of low ionic strength (0.25 x TBE), supplemented with 20% v/v glycerol, and a low running voltage. Figure 5.2 shows the progressive retardation of complexes with an increase in histone H1, which suggests that all the protein binds to the DNA under these conditions. At high levels of histone H1 (2.0 w/w ratio) there is some aggregation of histone H1 with the DNA, since some material does not migrate from the gel well. This aggregated material is presumed to be the "fast" complexes described by Clark and Thomas (1986), which are sedimented quickly in sucrose density gradients. However, the majority of the material are complexes that migrate freely through the agarose gel, and are presumed to be "slow" complexes (Clark and Thomas 1986) that are sedimented relatively slowly in sucrose density gradients. The ionic conditions for complex formation (20 mM HEPES-KOH pH7.0; 15 mM NaCl; 2.0 mM $MgCl_2$) were chosen as those that favour the formation of non-aggregated, or "slow" complexes. They were a consensus from several studies (Clark and Thomas 1986; Clark and Thomas 1988; Jerzmanowski and Cole 1990; Higurashi and Cole 1991).

Gel retardation experiments were also carried out on both unmethylated and methylated plasmid DNA, in assays identical to those described above. Previously, other workers have shown that the binding affinity of histone H1 for DNA is similar, for both unmethylated and methylated DNA (Higurashi and Cole 1991). Similar patterns of gel

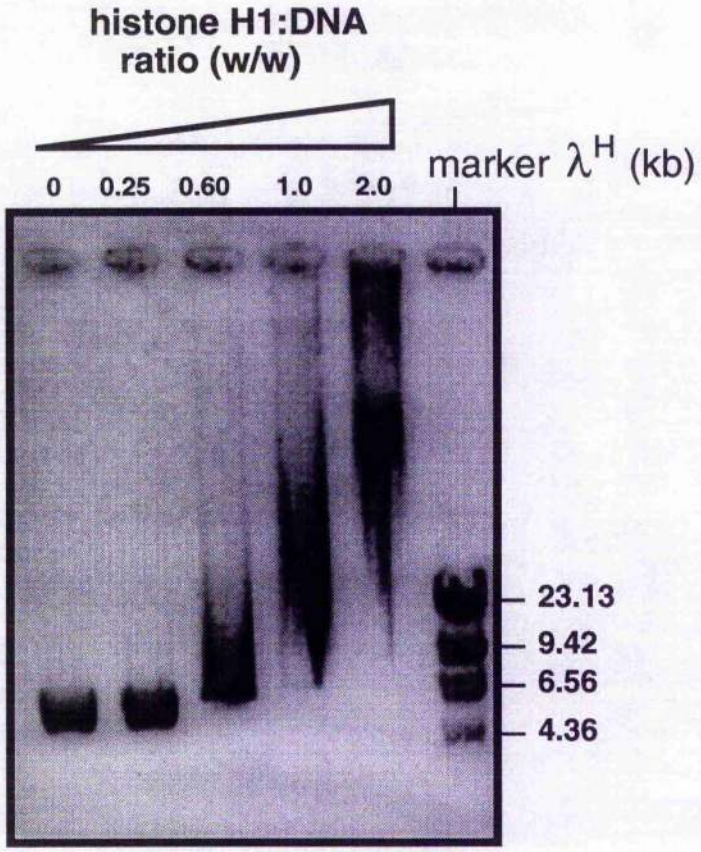


Figure 5.2

Gel retardation analysis of histone H1-DNA complexes: Complexes were formed on unmethylated pVHCk plasmid DNA and analysed by agarose gel electrophoresis (see Section 5.2.1). The size of the plasmid is 5025 bp. With an increase of the histone H1:DNA w/w ratio (0–2.0), as indicated by the wedge, the complexes are retarded. At a high ratio (2.0 w/w) there is significant aggregation of histone H1 and DNA, so that the aggregate remains in the well of the agarose gel.

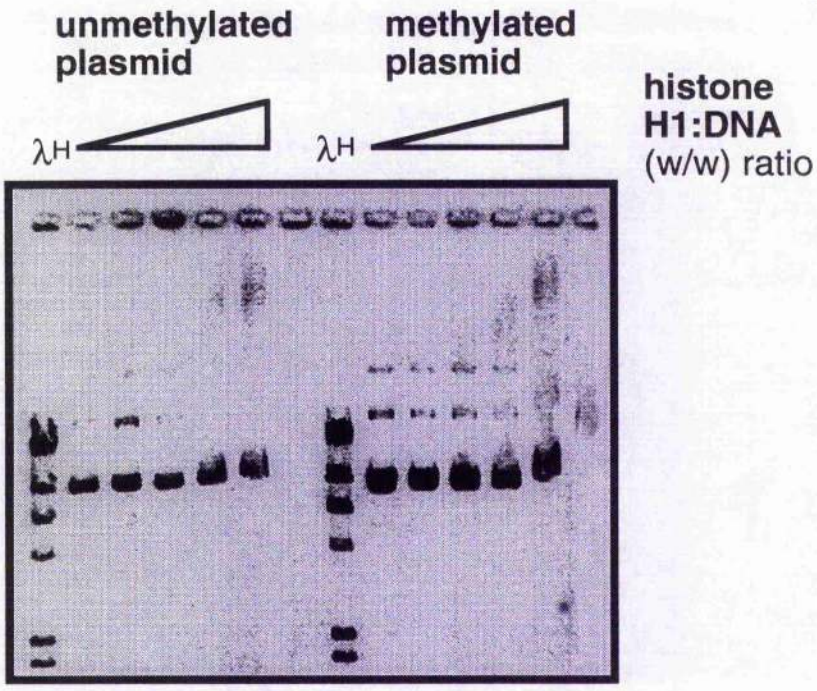


Figure 5.3

Gel retardation of histone H1-DNA complexes on unmethylated and methylated plasmid DNA:
Unmethylated and methylated plasmid DNA are retarded to similar extents with an increase in histone H1:DNA w/w ratio (0, 0.25, 0.60, 1.0, 2.0). The increases in the w/w ratios are indicated by the wedges, and the marker λ^H is *Hin* dIII digest of λ DNA. The size of the plasmid is approx. 8 kb. and aggregates are seen in the upper part of the gel at high w/w ratios. See Section 5.2.1. for details.

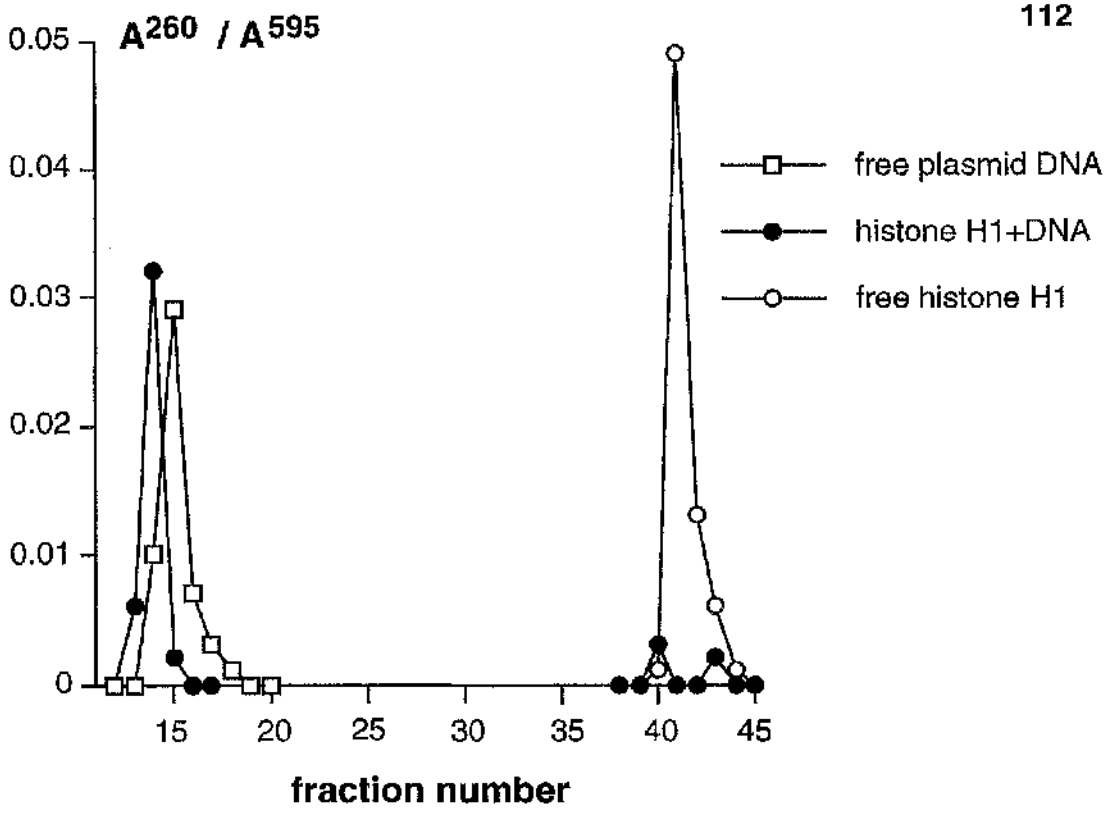


Figure 5.4

Gel filtration of histone H1-DNA complexes:
Histone H1-DNA complexes were loaded on a gel filtration column, as described in Section 5.2.2, to determine the extent that the input histone H1 binds to free plasmid DNA. Complexes (filled circles) and free plasmid DNA (squares) were eluted in the void volume, which was determined to be 13 fraction units of volume. Free histone H1 was eluted in fractions 41-42. Free histone H1 was not detected in the presence of DNA, which shows that all the available histone H1 forms complexes with the DNA. Histone H1 was detected by a Bradford assay of fractions, and DNA was detected by direct spectrophotometry of fractions at A^{260}.

retardation occur for both unmethylated and methylated DNA (Nightingale and Wolffe 1995), a finding which is confirmed in our own studies (Figure 5.3).

5.2.2. Gel filtration of histone H1-DNA complexes

Gel retardation studies confirmed that histone H1-DNA aggregates were not formed under the ionic conditions used, so aggregation could not be the facile explanation of transcriptional inactivation in this system. However, these studies did not confirm that all the input histone H1 actually bound to the DNA, so that there was no free histone H1 present in the reaction mixture after a suitable time of incubation. To study this gel filtration analyses were carried out on free plasmid DNA, free histone H1 and putative histone H1-DNA complexes, at an H1:DNA ratio of 1.0 w/w. A Sephadex G-50 (Pharmacia) gel filtration column (length 10 cm x diameter 0.5 cm) was packed under gravity and the void volume determined using Dextran Blue 2000 (Pharmacia). The column was then used to determine the extent of binding of histone H1 to DNA. DNA was detected by spectrometry of fractions at A^{260}, and protein by Bradford assay of fractions. The DNA and the complexes both passed through the column in the void volume (Figure 5.4), whereas free histone H1 was significantly retarded. No free histone H1 was present in the "complex" sample, which implies that all the histone H1 binds to the plasmid DNA under these conditions.

5.3. The effect of histone H1 and DNA methylation on transcription

This section details the inhibitory effect of histone H1-DNA complex formation on templates that were transcribed *in vitro*. The plasmid template used is pArg/Leu (see Section 2.5.1), which contains a $tRNA^{Arg}$ and a $tRNA^{Leu}$ gene that are transcribed by RNA polymerase III. The template is either unmethylated or methylated. Transcription is also assayed from the pVHCk plasmid template (see Section 2.5.1) which is transcribed by RNA polymerase II. However, the latter system is technically more difficult to assay

(Section 2.7.4.3), so most experiments were carried out on the pol III system. Unless it is explicitly stated otherwise, the template used can be assumed to be pArg/Leu. The HeLa nuclear extract that was used for this series of experiments on the pol III system was prepared according to the Dignam protocol (Dignam et al. 1983; see Section 2.7.4.1). This meant that the Dignam extract is contaminated by endogenous histone H1 which is, of course, fortunate because it allows the preferential inhibition of methylated templates to be observed in a simple titration experiment (Section 5.3.1).

The extent of transcription from unmethylated or fully methylated templates was assayed in the presence of various levels of histone H1. The transcriptional activity of both unmethylated and methylated templates are inhibited by increasing amounts of histone H1, although inhibition of the methylated template occurs at a lower H1:DNA ratio. The H1c variant shows the greatest preferential inhibition of the methylated template. Clearly, histone H1 complexed to DNA is one of the factors that inhibits transcription. Histone H1 inhibits transcription by preventing the formation of initiation complexes, particularly on methylated template (see Section 5.3.6), rather than by the formation of disordered H1-DNA aggregates.

5.3.1. Titration of unmethylated and methylated templates

This preliminary experiment investigated the effect of increasing amounts of template DNA on *in vitro* transcription. Transcription was preferentially inhibited from methylated templates, as compared to unmethylated templates, in transcription assays that used nuclear extract prepared according to the Dignam protocol (Dignam et al. 1983; see Section 2.7.3.2). At low levels of template (2–20 ng DNA/assay) little transcription is observed from the methylated pArg/Leu in comparison with the unmethylated control. This difference is less pronounced at high levels of template (>50 ng/assay; Figure 5.5), since a 20% reduction in transcription from the methylated template is observed. The extent of this reduction was rather variable, and depended on the particular batch of nuclear extract. (Transcription from high levels of methylated template shows little reduction for the batch of extract used in the experiment for Figure 5.5). These results indicate the presence of a limiting amount of a selective inhibitor in the nuclear extract.

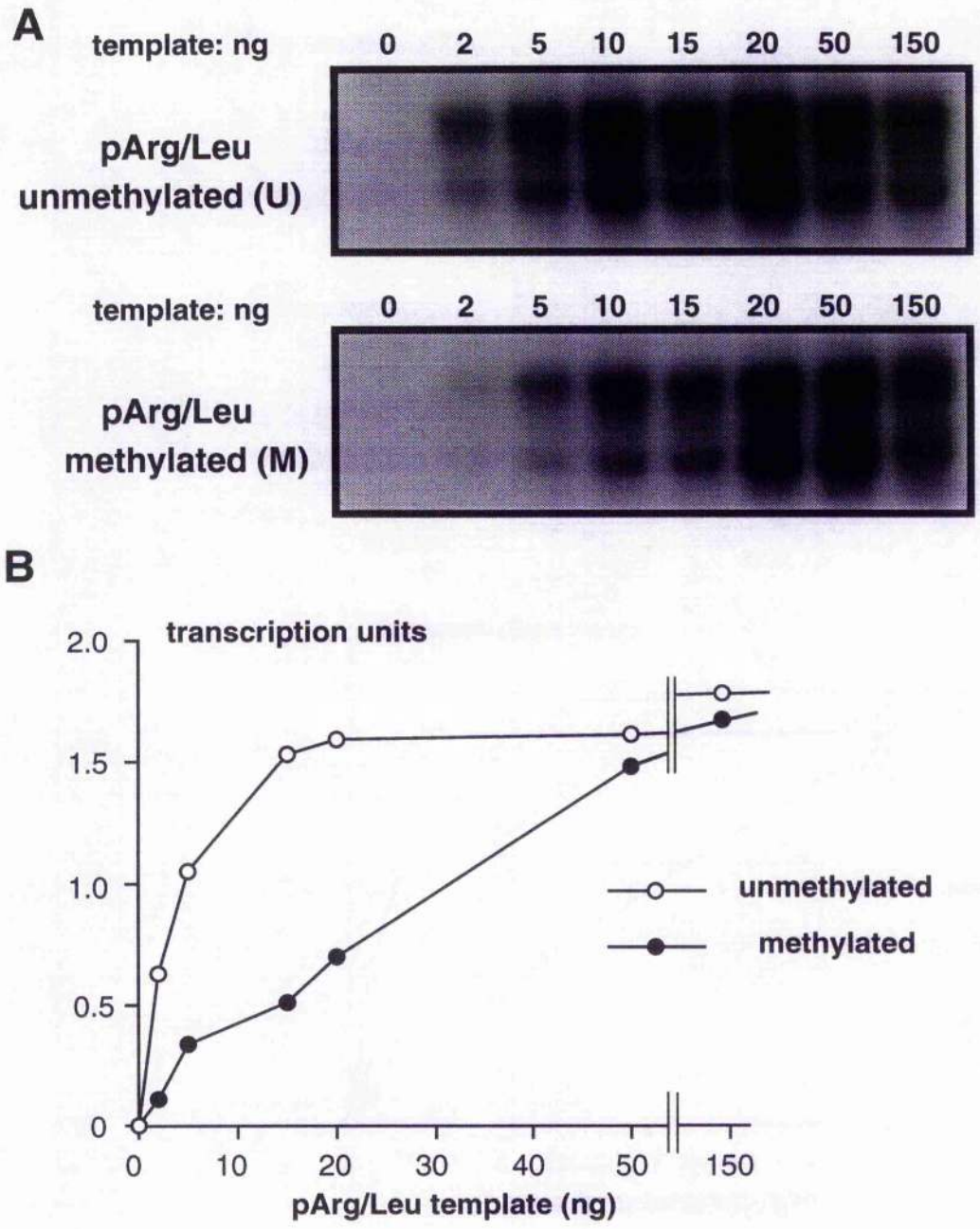


Figure 5.5

Titration of unmethylated and methylated template:
A: The effect of methylation on *in vitro* transcription of the pArg/Leu template was assayed at low levels of the unmethylated (U) or methylated (M) template (2–50 ng of DNA per assay) and a constant concentration of extract (see Section 5.3.1). At saturating levels of unmethylated and methylated template (150 ng of DNA per assay), the transcriptional activity of the methylated template is at least 80% that of the unmethylated template.
B: Transcription from both genes was quantified using a phosphoimager. The graph shows the preferential inhibition of methylated template (filled circles) compared with the unmethylated template (open circles). Units of transcription are arbitary units of PSL (phosphostimulable luminesence).

Similar results have been described for transcription of pol II genes by Boyes and Bird (1991), who suggested that MeCP-1 was the limiting protein.

Titrations were also carried out on increasing levels (50–500 ng) of unmethylated and methylated pVHCk plasmid template, using the standard *in vitro* transcription assay for this pol II system (Section 2.7.3.3). The results (not shown) were similar to those seen for the pol III system: a selective inhibition of the methylated template, at low levels (50 ng) of template, which largely disappeared at high levels (500 ng) of template.

5.3.2.Titrations with unmethylated and methylated competitor DNA

Transcription was carried out in the presence of increasing levels of unmethylated or methylated competitor DNA, in the form of either pUC19 plasmid DNA or a double-stranded oligonucleotide (see sequence below). It was supposed that the inhibitory effect of a methyl-CpG binding protein (MeCP), or a similar protein, on low levels of template would only be reversed by the addition of methylated competitor plasmid DNA. This is because MeCPs bind selectively to methylated DNA, but have little affinity for identical sequences that are unmethylated (Boyes and Bird 1991; see Section 1.7.2). Were the putative inhibitor to be competed out by both unmethylated and methylated competitor plasmid DNA, this would indicate that the inhibitor does not have the binding characteristics of MeCPs, since it can bind to both unmethylated and methylated DNA. A methylated double-stranded oligonucleotide would not be expected to bind MeCP-1 because it contains too few methylcytosines for the sequence binding affinity of MeCP-1 (Meehan et al. 1989). The oligonucleotide duplex could, however, act to compete out other inhibitory factors. Titration studies with both plasmid and oligonucleotide DNA could therefore determine if the putative inhibitor in the extract binds to methylated DNA exclusively, or if it binds to both unmethylated and methylated DNA.

5.3.2.1 Titration with unmethylated and methylated plasmid DNA

In the presence of 10–240 ng unmethylated or methylated pUC19 competitor DNA, transcription from 10 ng template was increased considerably (Figure 5.6). Nevertheless, transcription remains less effective from the methylated template, reaching only about 80% of the level obtained from the unmethylated template even at the highest levels of competitor tested. Somewhat greater stimulation of transcription is observed

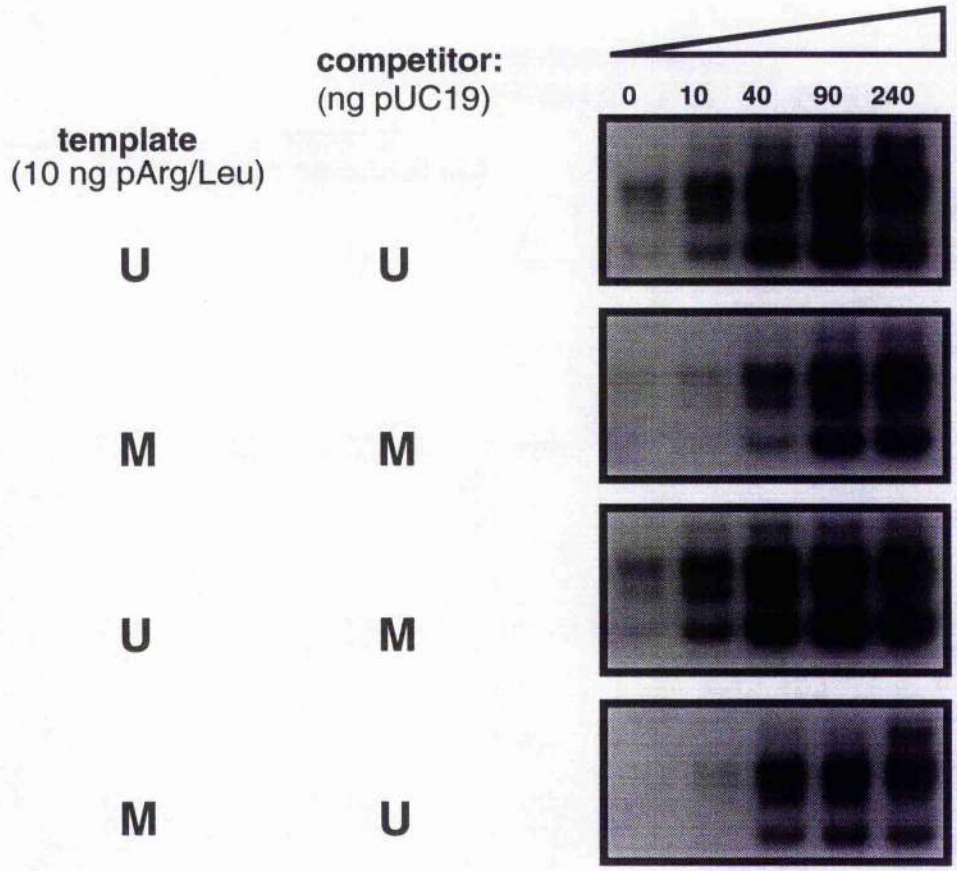


Figure 5.6

Removal of inhibitors from the nuclear extract with competitor DNA:
Transcription from 10 ng of unmethylated (U) or methylated (M) pArg/Leu template is increased in the presence of increasing levels (0–240 ng; indicated by the wedge) of unmethylated or methylated pUC19 competitor DNA.

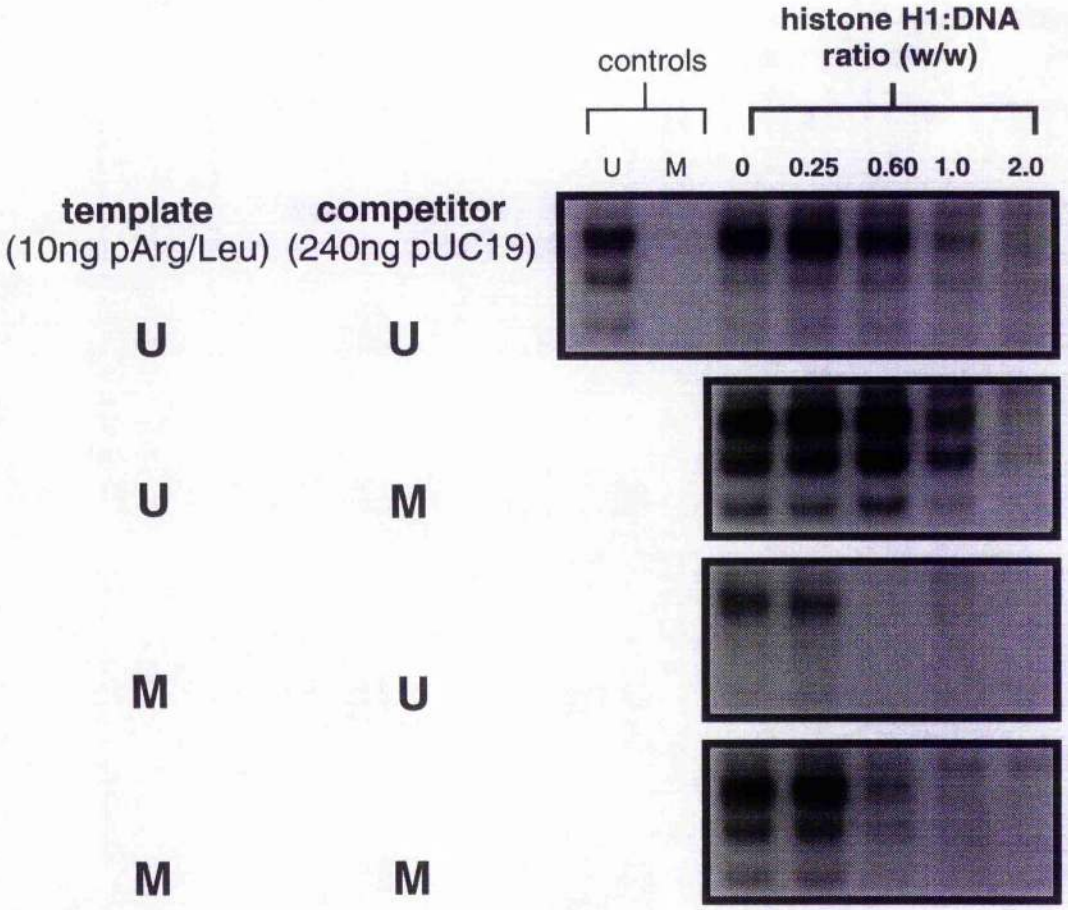


Figure 5.7

Reversal of enhanced transcription by addition of histone H1: Transcription from 10 ng of unmethylated (U) or methylated (M) pArg/Leu template, in the presence of 240 ng of unmethylated or methylated pUC19 competitor DNA, is inhibited by increasing levels of histone H1 (0–2.0 w/w ratio, with respect to template DNA). The two controls are transcription assays, in the absence of competitor, from 10 ng of unmethylated or methylated template, as indicated by the small letters U and M above the figure.

with methylated competitor but, in both cases, transcription was increased at least ten-fold. It is assumed that the stimulation of transcription was caused by the binding to the competitor DNA of inhibitory proteins present in the nuclear extract. As the inhibitory factors bind to both unmethylated and methylated competitor DNA, this observation is not consistent with the limiting inhibitor being MeCP-1, or a similar protein that binds preferentially to methylated DNA. It also appears to rule out the involvement of MeCPs in the general lowered template activity of the methylated template as this is still observed in the presence of excess methylated competitor. It is, however, consistent with a role for histone H1 as the limiting component responsible for preferentially inhibiting transcription from a methylated template. Histone H1 is known to be present in low amounts in the HeLa nuclear extracts (Croston et al. 1991; Paranjape et al. 1994) and will be removed by binding to competitor DNA, be it methylated or not.

In a second experiment, it is seen that preincubation of the template DNA (either unmethylated or methylated) with histone H1 leads to an inhibition of the enhanced transcription seen on addition of competitor DNA (Figure 5.7). This inhibition of transcription by histone H1 occurs preferentially with methylated templates, whether or not the competitor DNA is methylated. In other words, this is evidence for histone H1 alone being essential (and limiting) for the preferential inhibition of transcription from the methylated template. These titration experiments do not, by themselves, provide evidence for the involvement of a second inhibitory factor (e.g. MeCP-1) that could mediate the preferential inhibition of transcription from methylated templates. However, these experiments cannot rule out the involvement of MeCPs in the general lowered transcriptional activity of the methylated template. This activity is consistently no more than 80% of that from the unmethylated template.

5.3.2.2. Titration with double-stranded oligonucleotides

A second type of DNA competitor was also used, in an attempt to compete out inhibitory factors from the nuclear extract. MeCP-1 requires 15 methylated CpGs within a short region of DNA (see Section 1.7.2), and so would be expected to bind to the methylated plasmid competitor (used in the titration experiments of Section 5.3.2.1) but not to a short double-stranded unmethylated or methylated oligonucleotides (shown in Figure 5.8A).

A

unmethylated (U) dsDNA 44-mer oligonucleotide, containing six CpGs:

5' GTCAA**CG**AGG GAGC**CG**ACTG C**CG**A**CG**TG**CG** CTC**CG**GAGGC TTGC
CAGTT**GC**TCC CTCG**GC**TGAC G**GC**T**GC**AC**GC** GAG**GC**CTCCG AACG 5'

methylated (M) dsDNA 44-mer oligonucleotide, containing six mCpGs:

5' GTCAA**MG**AGG GAGC**MG**ACTG C**MG**A**MG**TG**MG** CTC**MG**GAGGC TTGC
CAGTT**GM**TCC CTCG**GM**TGAC G**GM**T**GM**AC**GM** GAG**GM**CTCCG AACG 5'

B

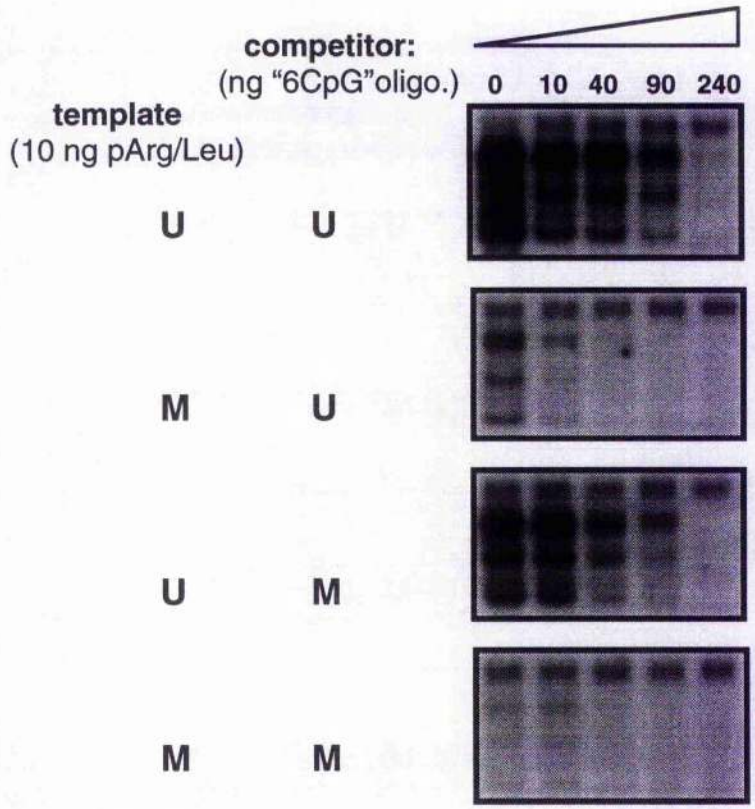


Figure 5.8

Possible removal of transcription factors from the nuclear extract with unmethylated or methylated dsDNA oligonucleotide:

A: The unmethylated (U) 44-mer dsDNA oligonucleotide contains six CpGs (shown in bold font). In the methylated (M) oligonucleotide the cytosine of each CpG is replaced by 5-mC (indicated by the bold M). **B:** Double-stranded oligonucleotides, either unmethylated (U) or methylated (M), were used as competitors in transcription assays with 10 ng unmethylated (U) or methylated (M) pArg/Leu template. The competitor was added in increasing amounts (0–240 ng, indicated by the wedge) to each series of assays. See Section 5.3.2.2 for further details.

The unmethylated 44-mer dsDNA oligonucleotide contains six CpG dinucleotides (shown in bold font in Figure 5.8A). In the methylated oligonucleotide the cytosine of each CpG dinucleotide is replaced by 5-mC (indicated by the bold M). These oligonucleotides have been used to investigate the preferential binding of histone H1 variants to a series of different oligonucleotides in gel-retardation assays (Santoro et al. 1995) . Refer to Sections 5.3.5 and 5.4 for a discussion about histone H1 variants. The oligonucleotides used in this experiment were kindly provided by Prof. P. Caiafa. MeCP-1 would not be expected to bind to these oligonucleotides because there are fewer than 15 closely spaced methylated CpGs, but they could compete out other inhibitory factors that bind to both methylated and unmethylated DNA.

With increasing amounts of double-stranded unmethylated or methylated oligonucleotide, transcription from 10 ng template is correspondingly reduced (Figure 5.8B). This can be explained by the removal of DNA-binding proteins, including basal transcription factors, from the nuclear extract by the binding of the proteins to the competitor DNA. Figure 5.8B suggests that transcription factors can be competed out equally well from either the unmethylated template or the methylated template, in the presence of increasing levels of competitor. However, the data does not provide evidence that both the unmethylated and methylated oligonucleotide compete out transcription factors to the same extent. These observations suggest that the preferential inhibition of transcription from methylated templates (see Figure 5.5) could be explained by prevention of transcription factor binding to methylated DNA (see Section 1.7.1), for this pol III *in vitro* transcription system. An identical competition effect is seen with a different unmethylated or methylated 40-mer dsDNA oligonucleotide that forms part of the SV40 promoter, and which contains three GC boxes and three CpG dinucleotides (results not shown).

5.3.3. Transcription with histone H1-depleted nuclear extract

Another way to remove histone H1 is to subject the nuclear extract to ammonium sulphate fractionation (see Section 2.7.3.1). This is a standard technique to remove contaminants, such as histone H1, from crude nuclear extracts such as that prepared by the Dignam

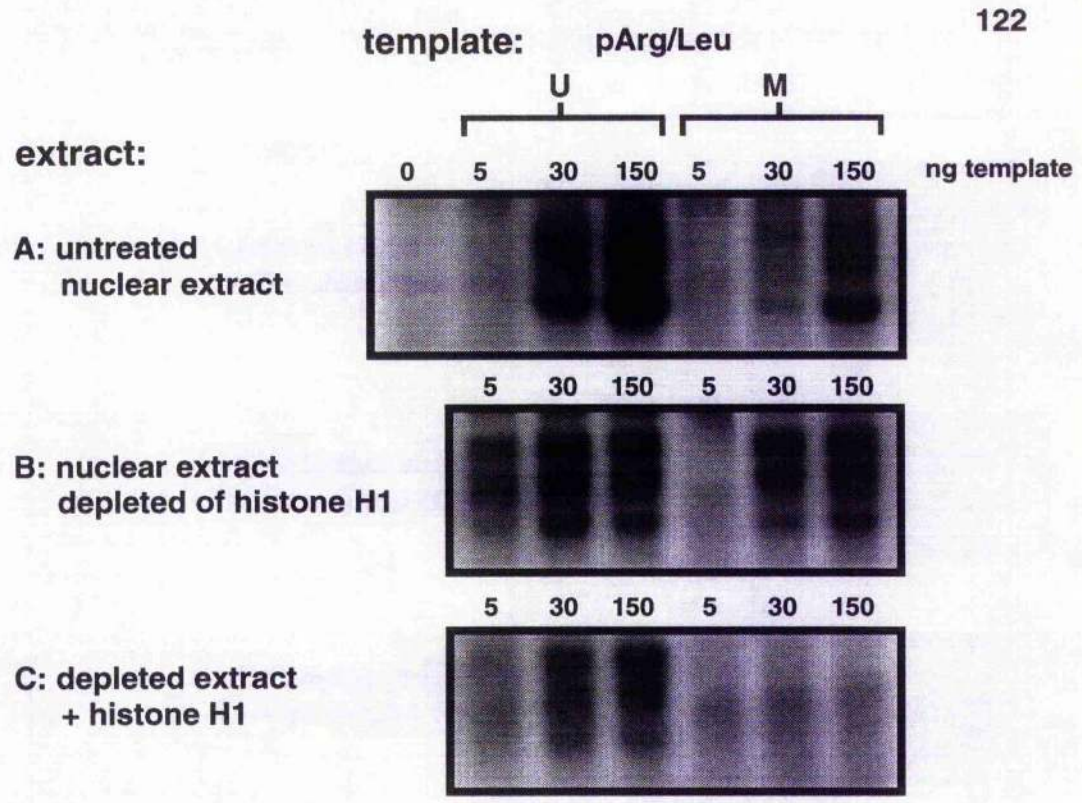


Figure 5.9

Transcription with histone H1-depleted extract

Depletion of nuclear extract of histone H1 (see Section 5.3.3) increases the level of transcription from the methylated template compared to the unmethylated template (U). The figure shows the results of *in vitro* transcription assays performed in the presence of increasing amounts (0–150 ng) of unmethylated (U) or methylated (M) template. **A:** untreated extract (results comparable to those of Figure 5.5); **B:** histone H1-depleted extract; **C:** same as B, but with the addition of exogenous total histone H1 at an H1:DNA ratio of 0.60 w/w.

Table 5.1

extract:	template (30 ng pArg/Leu)	
	unmethylated	methylated
A: untreated nuclear extract	**6.9**	**3.0**
B: histone H1-depleted nuclear extract	**6.4**	**5.4**
C: depleted extract + histone H1	**3.8**	**0.8**

units of transcription (PSL)

Loss of preferential inhibition of methylated template

The results of Figure 5.9 were quantitated for 30 ng of unmethylated and methylated pArg/Leu template using a phosphoimager. Units of transcription are arbitary units of PSL (phosphostimulable luminesence).

protocol (Dignam et al. 1983), and is described in more detail elsewhere (Shapiro et al. 1988). The fractionation procedure precipitates all the proteins involved in the transcriptional machinery, but allows highly basic proteins such as histone H1 to remain in the supernatant.

In comparison to the crude nuclear extract that was used in the studies described above, a nuclear extract that is depleted of histone H1 shows increased transcription with low levels (5–30 ng) of methylated or unmethylated template (Figure 5.9). Table 5.1 tabulates the level of transcription from 30 ng of unmethylated or methylated template, which shows that the nuclear extract depleted of histone H1 reduces the degree of preferential inhibition of transcription from the methylated template. The addition of histone H1 (histone H1:DNA w/w ratio of 0.6) to the depleted extract increases the degree of the preferential inhibition, with transcription from the methylated template almost completely inhibited. Inhibition is presumed to be due to the formation, during the preincubation stage of the transcription assay (Section 2.7.3.2), of H1-DNA complexes that inhibit transcription from methylated and, to a lesser extent, unmethylated constructs.

5.3.4. Histone H1 preferentially inhibits transcription from methylated templates

Since low levels of methylated template can be preferentially inhibited by the endogenous histone H1 present in crude nuclear extracts, it seemed reasonable to try to inhibit transcription from high levels of template by the addition of exogenous histone H1. With amounts of template in excess of 50 ng per assay, the difference between transcription from methylated and unmethylated templates is small (Figure 5.5B), in the absence of exogenous histone H1. However, at high levels of template DNA, transcription is inhibited from both templates (Figure 5.10) onto which "slow" complexes of total calf thymus histone H1 have been deposited (Section 2.7.5.2). Transcriptional inhibition occurs at a lower H1:DNA ratio for methylated templates, in comparison to unmethylated or mock-methylated templates. Complete inhibition is obtained at high levels of H1 (1.0–2.0 ratio), for both types of template. Previous data (Section 5.2.2) suggests that all the exogenous histone H1 that is added binds to the DNA, so H1:DNA w/w ratios of 0.25,

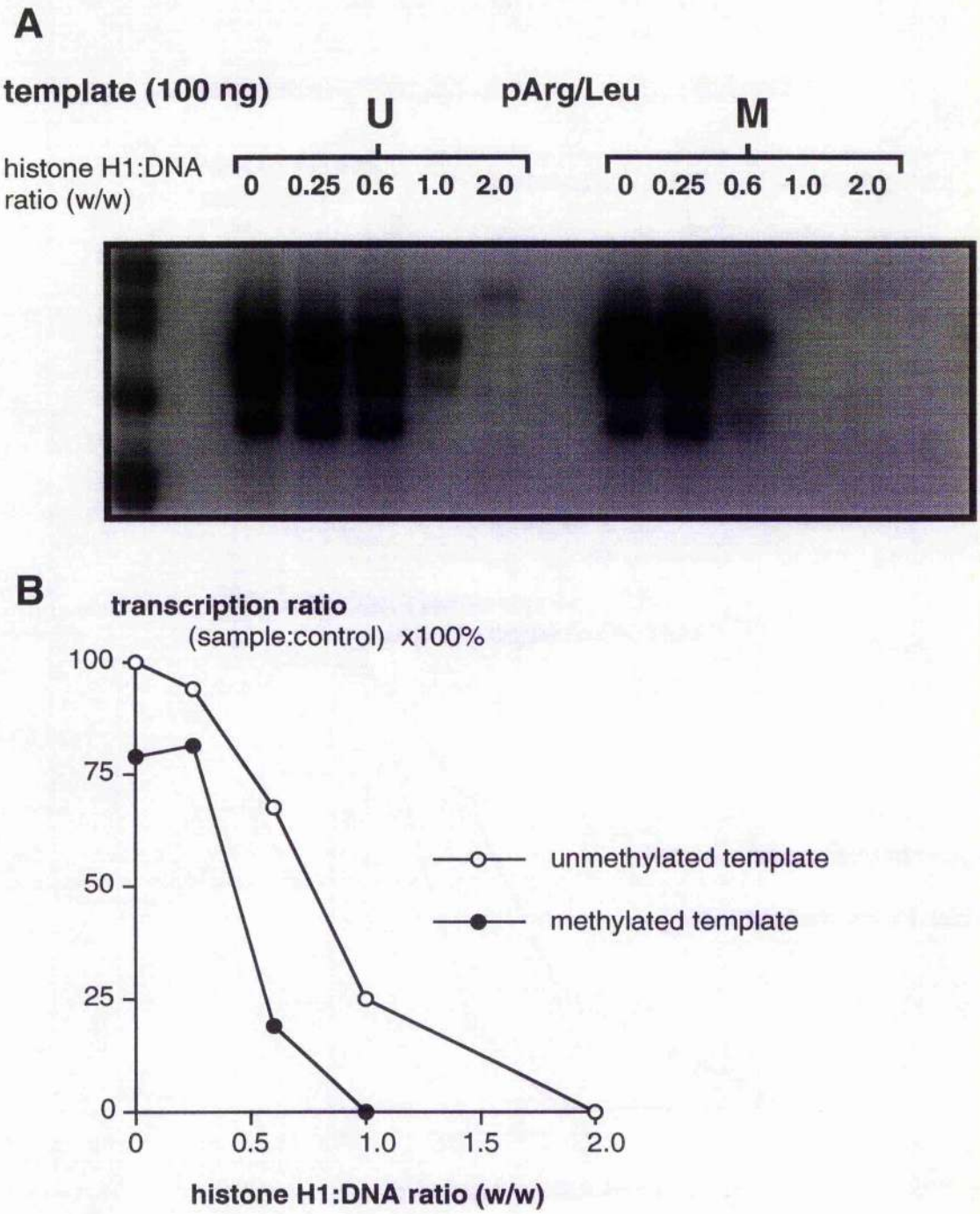


Figure 5.10

Histone H1 preferentially inhibits transcription from a methylated pol III template, pArg/Leu:

A: The extent of transcription from 100 ng of unmethylated (U) or methylated (M) pol III template was assayed in the presence of increasing levels (0–2.0 w/w ratio, as indicated) of total histone H1 (see Section 5.3.4).

B: Transcription for both the unmethylated template (open circles) and the methylated template (filled circles) was quantified using a phosphoimager. The graph shows the preferential inhibition of transcription from the methylated template by histone H1. The level for unmethylated pArg/Leu template, in the absense of histone H1, is taken as the standard 100 % value, against which other samples are compared. Units of transcription are arbitary units of PSL (phosphostimulable luminesence).

0.60 and 1.0 correspond to one histone H1 molecule binding every 125, 55 and 39 bp of DNA respectively. Control experiments using core histones (purified as described in Section 2.7.6) over the same range of protein:DNA ratios, showed no inhibition with either template (see below, Figure 5.14 and Section 5.3.5).

An identical series of assays was also carried out on 500 ng of unmethylated and methylated pVHCk plasmid template, in the presence of increasing levels of histone H1 (H1:DNA ratio 0–2.0 w/w). The standard *in vitro* transcription assay for this pol II system was used, as described in Section 2.7.4.3. The results (Figure 5.11) are similar to those seen for the pol III system: a preferential inhibition of the methylated template at low levels of histone H1 (0.4–0.6 w/w), and inhibition of both unmethylated and methylated template at higher levels of histone H1 (0.6–2.0 w/w). However, preferential inhibition occurs at higher levels of histone H1 and is not as pronounced as the inhibition seen for the pol III system.

The preferential inhibition of transcription from methylated templates by histone H1 is also seen for covalently-closed circular (form II topoisomer) and linear pArg/Leu template (Figure 5.12). Covalently-closed circular pArg/Leu plasmid was prepared by treatment of the DNA with topoisomerase I (Promega), under conditions recommended by the manufacturer, and linear template was obtained by digestion of pArg/Leu with *Not*I (see Section 2.5.1). There is a decrease in the absolute level of transcription by two-fold and five-fold for covalently-closed circular and linear templates respectively, in comparison to the supercoiled template (results not shown). This result is consistent with the previous observation that linearisation of negatively supercoiled recombinant tRNA genes reduces their *in vitro* transcriptional activity two to three-fold (Shapiro et al. 1988). These results suggest that the overall supercoiled conformation of the template does not affect the inhibitory activity of histone H1 nor the differential transcription from the two templates. To test if the number or density of methylated CpGs has an effect, a 943 bp *Eco* RI-*Sma*I fragment (indicated in Figure 2.1B) containing both tRNA genes but lacking the G+C rich region was used as a template. Preferential inhibition of transcription from the methylated template is still observed, despite a reduced methyl CpG density (Figure 5.12).

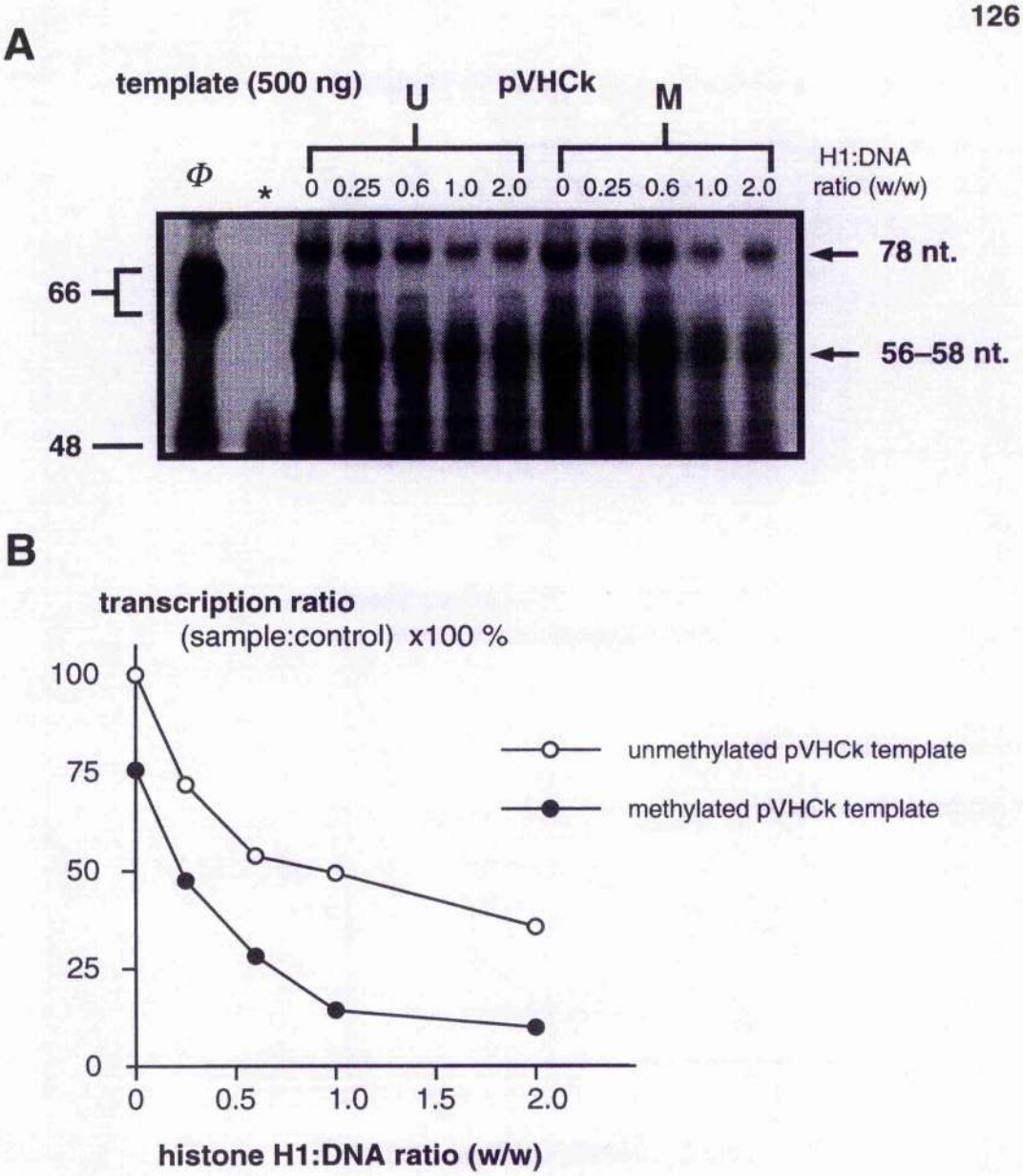


Figure 5.11

Histone H1 preferentially inhibits transcription from methylated pol II template, pVHCk:

A: The extent of transcription from 500 ng of unmethylated (U) or methylated (M) pol II template was assayed in the presence of increasing levels (0–2.0 w/w ratio, as indicated) of total histone H1. Primer extension products of size 56–58 and 78 nt. are indicated (see Section 3.3.4 and Figure 3.7 for details). The asterisk (*) marks the negative control assay, which has no template DNA. The marker Φ is end-labelled ΦX174 DNA digested with *Hin*fI.

B: Transcription was quantified by measuring the levels of all the primer extension products with a phosphoimager. The level for unmethylated pVHCk template, in the absence of histone H1, is taken as the standard 100% value, against which other samples are compared.

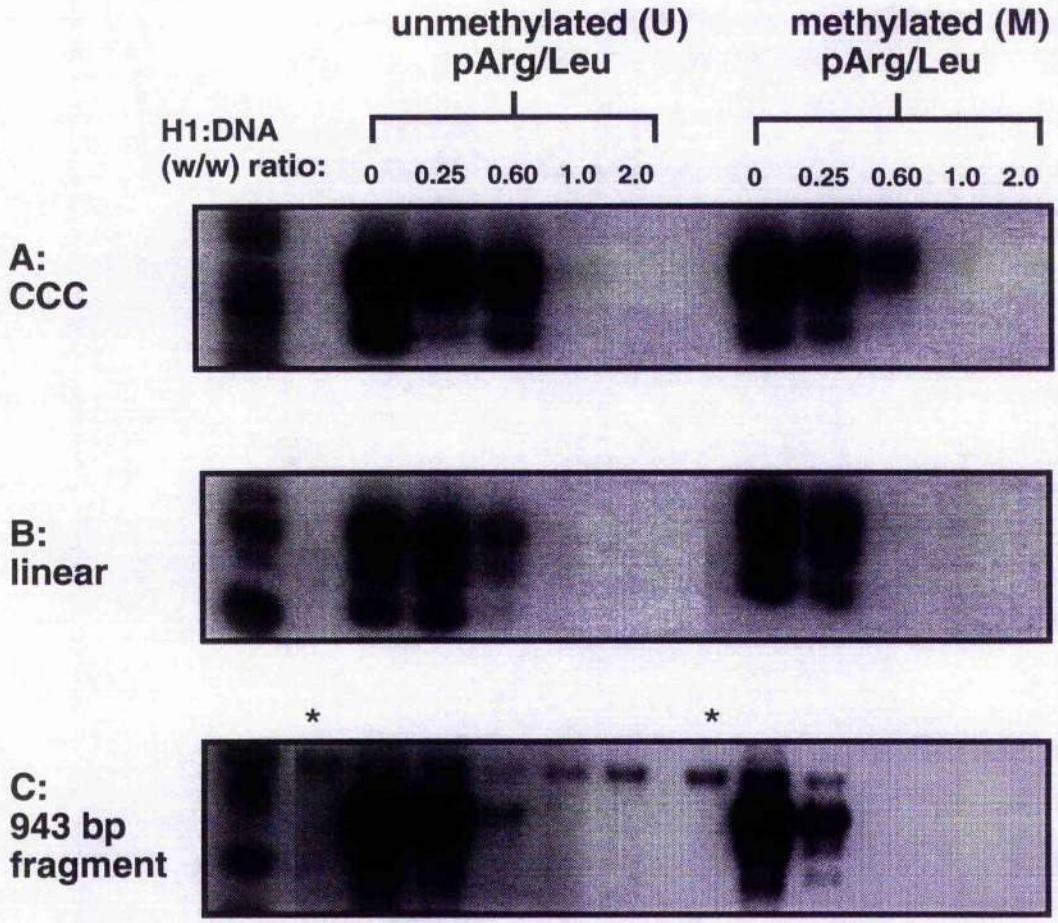


Figure 5.12

Effect of different conformations of methylated template on preferential inhibition of transcription by histone H1:
Preferential inhibition by increasing levels of histone H1 (0–2.0 w/w ratio) of the transcription from different conformations of 100 ng methylated pArg/Leu template were assayed (see Section 5.3.4). **A:** covalently-closed circular (CCC) pArg/Leu; **B:** linear pArg/Leu; and **C:** a 943 bp restriction fragment containing both tRNA genes. The asterisk (*) indicates negative control assays, which were performed in the absence of DNA template. The marker on the extreme left is the usual end-labelled ΦX174/*Hin*fI marker.

In summary, histone H1 appears to preferentially inhibit transcription from methylated templates over a critical range of histone H1:DNA w/w ratios (0.25 - 0.60). Preferential inhibition of transcription from methylated templates is observed irrespective of the supercoiled status or of the methyl CpG density of the templates.

5.3.5. Different variants of histone H1 inhibit transcription to unequal extents

The contribution of different histone H1 variants to the formation of inactive chromatin is not known but it has been suggested that they may act at different sites to regulate the stability of the 30 nm chromatin fibre (Wolffe 1992; Dimitrov et al. 1993; Pruss et al. 1995). This topic is discussed in detail in Section 1.8.7. I decided to investigate any differences between the inhibitory action of somatic histone H1 variants and the preparation of total calf thymus histone H1 that had been used in the experiments described above. Any differences *in vitro* could reflect the *in vivo* role of variants during the formation of the 30 nm chromatin fibre.

The different variants of histone H1 were partially purified by reverse-phase HPLC (Santoro et al. 1995) using a method described elsewhere by (Quesada et al. 1989). The variants were kindly provided by Prof. P. Calafa. Protein fractions were lyophilised and renatured as described in Section 2.7.5.1. Total histone H1 and H1 variants were analysed and characterised by SDS-polyacrylamide (12%) gel electrophoresis and a spectrometric trace of the column fractions is shown in Figure 5.13. Variant preparations are not homogenous as a result of incomplete resolution obtained by reversed-phase HPLC chromatography. Variants were characterised as previously described (Santoro et al. 1995), from which it could be concluded that preparations of H1a and H1d were 95% homogenous, whereas the H1c preparations (fraction 2 and 3) were contaminated with H1e. The H1e preparation (fraction 1) was also contaminated with H1c, and H1b was not resolved using this method of chromatography. Protein concentrations were determined by Bradford's procedure (Section 2.7.2.1), after correction for the anomalous effect of histones on colorimetric assays.

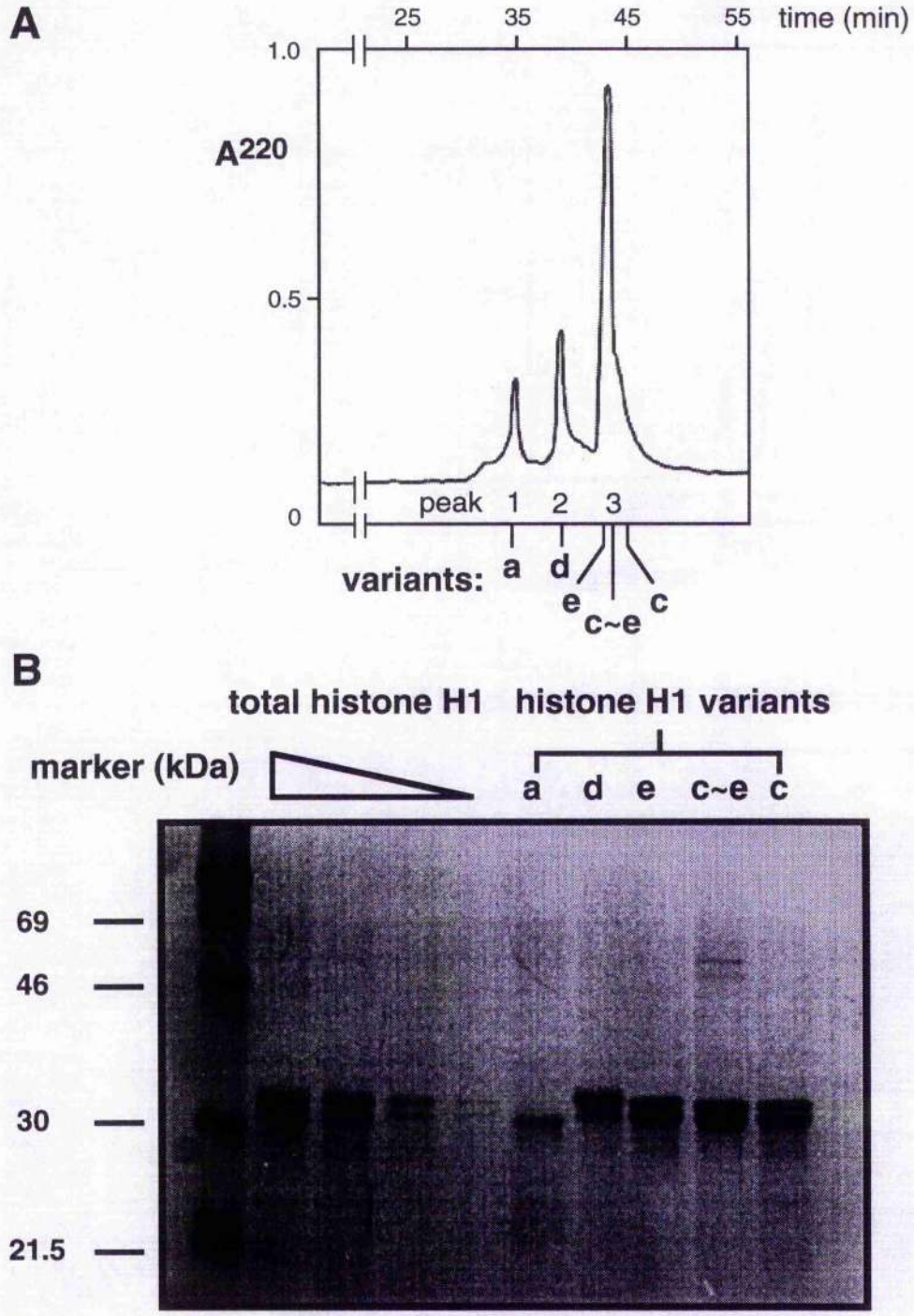


Figure 5.13

Elution profile and characterisation of histone H1 variants from reverse-phase HPLC column:

A: Histone H1 variants were separated by reverse-phase HPLC as described in Section 5.3.5. The elution profile shown is taken from Santoro et al. (1995), which describes the partial purification of the indicated variants.

B: SDS-PAGE was used to characterise each variant preparation. The preparations are compared to decreasing levels (indicated by the wedge) of total acid-extracted calf thymus histone H1.

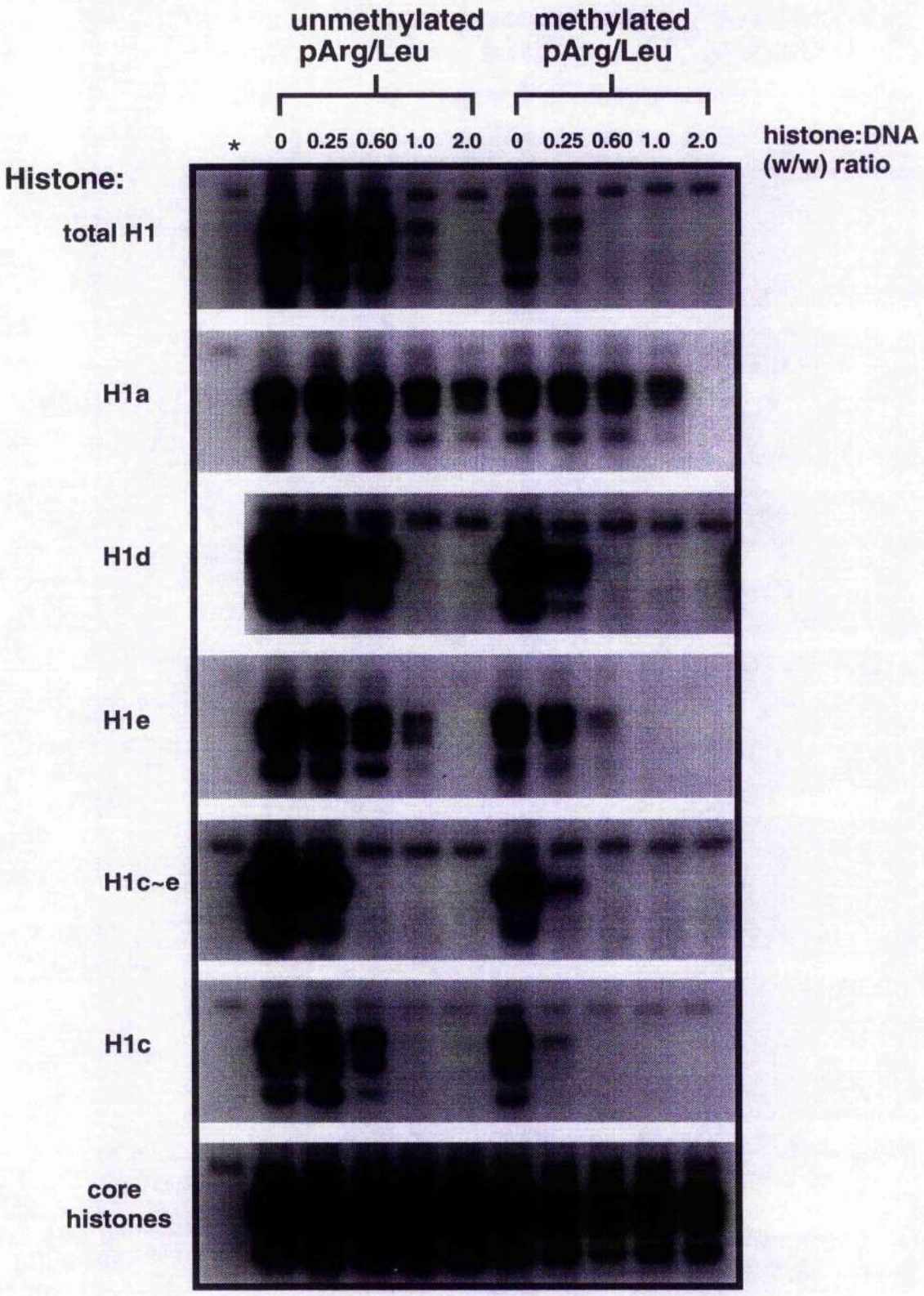


Figure 5.14

Effect of histone H1 variants on preferential inhibition of transcription from methylated templates:
Each preparation of a histone H1 variant (indicated on the left) and core histones were assayed for their ability to preferentially inhibit the methylated pArg/Leu template, compared to the unmethylated template. See Section 5.3.5 for details. The asterisk (*) indicates negative control assays that were performed in the absense of DNA template.

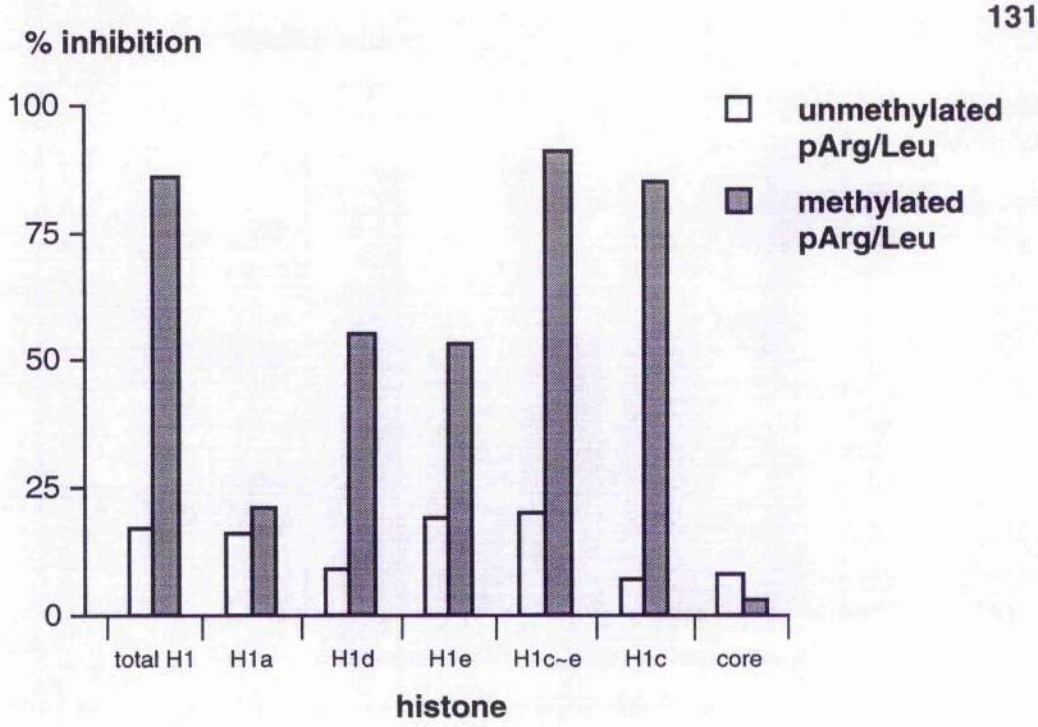


Figure 5.15

Effect of different histone H1 variants on transcription:
The transcription assays for each histone H1 variants shown in Figure 5.14 were quantified using a phosphoimager. The transcription from unmethylated pArg/Leu template (open columns) and methylated template (filled columns) at a histone H1:DNA w/w ratio of 0.25 is compared with the level of transcription in the absence of histone H1 (control assays). The difference for each template was then expressed as a %inhibition, compared to the control. Most variants inhibit the methylated template preferentially (Section 5.3.5).

Each variant was complexed with methylated or unmethylated (control) DNA templates, at increasing histone H1:DNA w/w ratios, and used in transcription assays as described above (Figure 5.14). Assays for the total histone H1 preparation and a preparation of core histones as control are also shown in Figure 5.14. Preparations containing predominantly H1c show inhibition at H1:DNA levels greater than those shown by total histone H1 (Figure 5.14). The unmethylated template is inhibited at H1:DNA ratios of 0.60-1.0 (w/w), whereas the methylated template is inhibited at the 0.25 (w/w) ratio. The preparations which contain predominantly H1e or H1d have a reduced, preferential inhibitory activity, whereas H1a has little preferential inhibitory activity (Figure 5.15). Core histones do not have any inhibitory activity in this system.

5.3.6. Histone H1 prevents the formation of initiation and elongation transcription complexes preferentially on methylated templates

It has previously been shown that histone H1 inhibits RNA polymerase II transcription by preventing the assembly of initiation complexes on template DNA (Croston et al. 1991). The results in this section show that a similar interaction occurs with the general RNA polymerase III transcriptional machinery, but that this process appears to occur preferentially on methylated templates.

The results of Figure 5.16 show that preincubation of the template with histone H1 at a H1:DNA ratio of 1.0 (w/w) prevents the assembly of initiation complexes on both unmethylated (lane 2) and methylated (lane 9) template. Histone H1 also inhibits the conversion of preassembled initiation complexes into elongation complexes on the addition of ribonucleoside triphosphates, although this inhibition is less effective on unmethylated template (lane 3) than on methylated template (lane 10). However, the presence of H1 does not inhibit elongation from either template (lanes 4 and 11), although transcription is limited to a single round in both cases, as shown by comparison with transcription assays containing 0.025% (w/v) Sarkosyl (*N* -laurylsarcosine) (lanes 7 and 14). Sarkosyl prevents the formation of initiation complexes (lanes 5 and 12) and the conversion of preassembled initiation complexes into elongation complexes (lanes 6 and

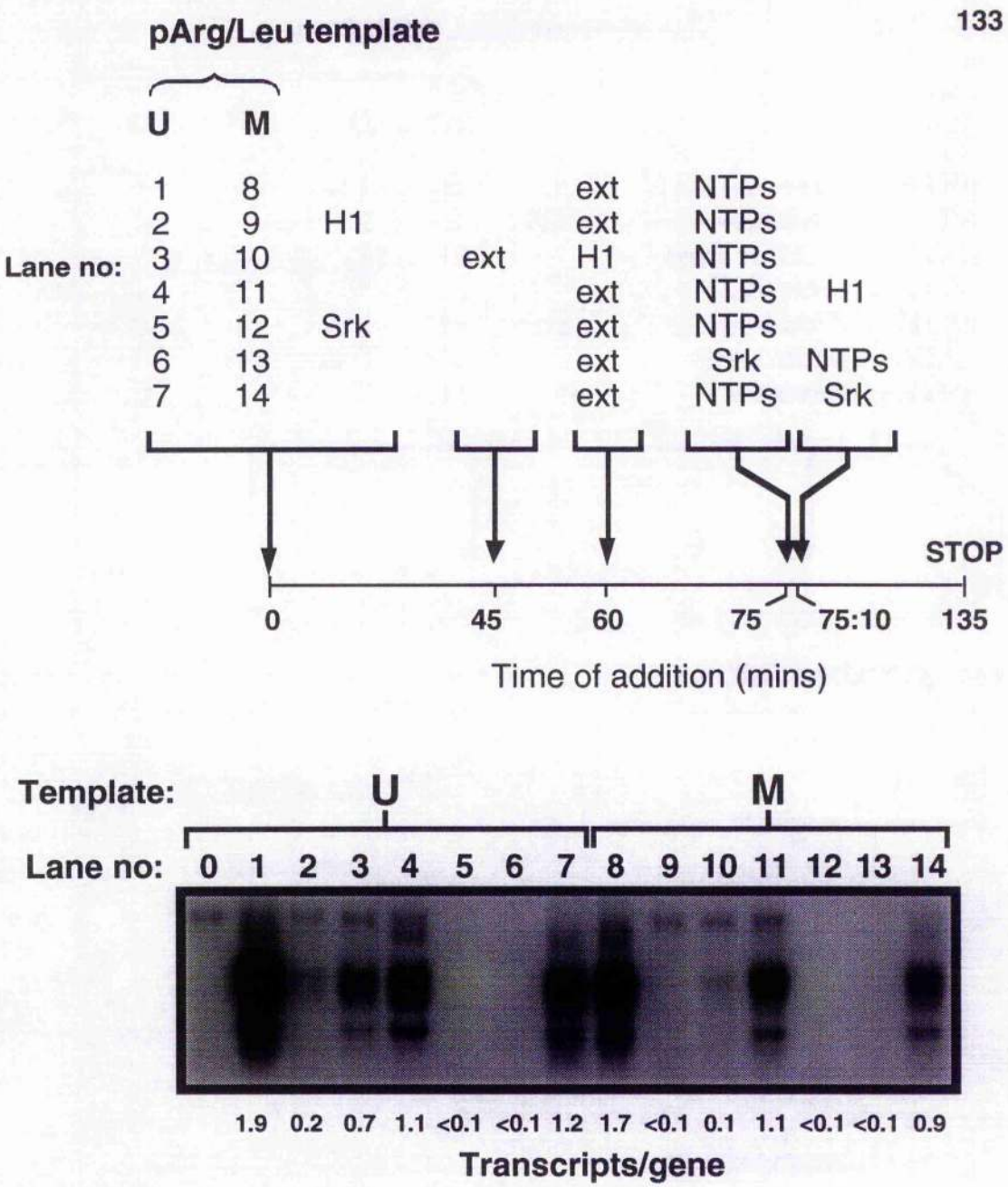


Figure 5.16

Initiation of transcription on methylated template:
Assays were performed with 100 ng of unmethylated (U) or methylated (M) pArg/Leu, as outlined in in the scheme above (see Section 5.3.6 for details). Histone H1 and nuclear extract (ext) were added at the indicated times to either unmethylated (lanes 1–7) or methylated (lanes 8–14) template. Lane 0 is a control assay and is as lanes 1 and 8, but performed in the absense of template. The upper part of the figure shows the order of addition of the various components to the transcription assys. Transcription was initiated with ribonucleoside triphosphates (NTPs). Initiation or elongation complex assembly was inhibited by the addition of 0.025% Sarkosyl (Srk) at the indicated times. Sections of gels corresponding to the transcripts for one particular gene were excised and quantified by scintillation counting. Transcriptional efficiency is expressed as transcripts per gene.

13) but, like H1, does not stop the elongation by the polymerase once ribonucleoside triphosphates have been added (Hawley and Roeder 1987). As a consequence, the presence of Sarkosyl prevents reinitiation of transcription, limiting the transcription to one round. This was confirmed by excising gel slices containing the labelled RNA transcripts for one particular gene, counting in a liquid scintillation counter and determining the transcriptional efficiency in terms of transcripts per gene (Figure 5.16).

In summary, it can be concluded that histone H1, like Sarkosyl, acts by preventing the formation of initiation complexes, particularly on methylated templates, rather than by preventing elongation which proceeds at similar rates from unmethylated and methylated templates.

5.3.7. Gene methylation is not essential for H1-mediated inhibition of transcription

Patch-methylated constructs of pArg/Leu were made by ligating together different combinations of unmethylated or methylated *Bgl* I-*Bgl* I fragments. This was achieved by methylating the restriction fragments obtained by digestion of pArg/Leu with *Bgl* I, separating and purifying the fragments by standard procedures and ligating suitable combinations of methylated fragments with unmethylated fragments. *Bgl* I was chosen because it cleaves the sequence GCCNNNN/NGGC at the site indicated by the stroke. In principle, the *Bgl* I fragments could be religated to form a covalently-closed circular molecule, if all the *Bgl* I sites had unique sequences. This prevents the formation of a variety of linear molecules and concatamers, which would otherwise occur if different fragments had compatible sticky ends. On average, a restriction fragment of unknown sequence will have a 1 in 64 chance of having compatible sticky ends with a second fragment. Plasmid pArg/Leu has five *Bgl* I sites, which give rise to five fragments on cleavage with the enzyme. Two sites are of unknown sequence (Figure 5.17A), but their approximate location is known from restriction mapping. The chance that one of these sites has the same sticky end as the other four is therefore 1 in 16, but since the sequence of two sites is unknown the chance is reduced to 1 in 32. This means that the probability

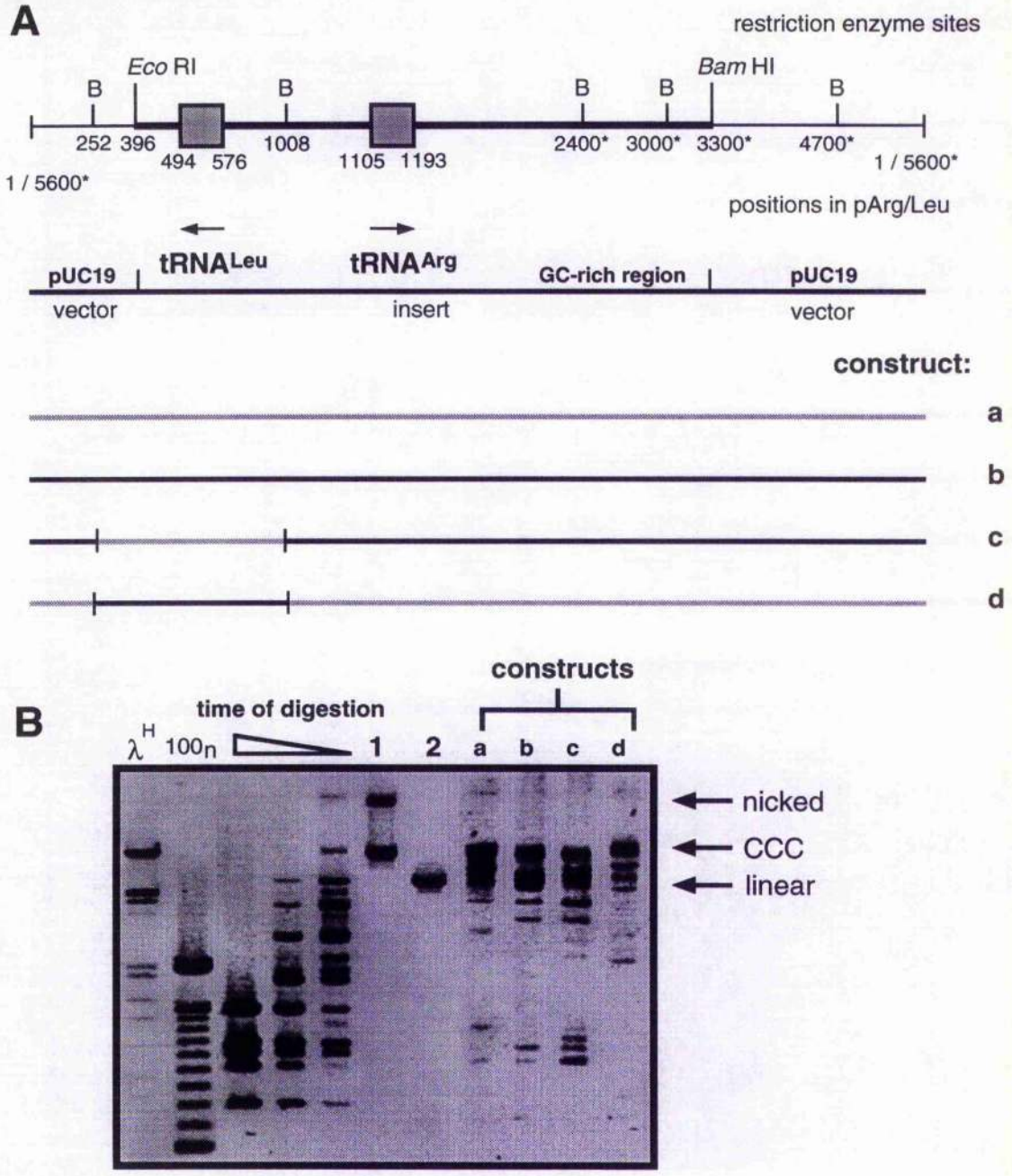


Figure 5.17

Patch-methylated constructs of the pArg/Leu template:
A: Patch-methylated constructs of pArg/Leu were constructed as described in Section 5.3.7. The top of the diagram shows the positions of restriction enzyme sites (see Section 2.5.1), including *Bgl*I sites (indicated by B), and the two tRNA genes within the insert of pArg/Leu (indicated by the bold line). The position of sites marked with an asterisk (*) were determined by restriction mapping. Methylated regions of constructs b–d are indicated by a thick line in the lower part of the diagram. The dashed line indicates unmethylated regions in constructs a, c and d.
B: Religated covalently-closed circular (CCC) constructs a–d were gel purified. The products of the ligation reaction are compared to DNA samples of pArg/Leu that are either in the CCC and nicked forms (lane 1) or linear (lane 2). These samples were prepared by digesting pArg/Leu with DNA topoisomerase or *Not* I. Plasmid pArg/Leu was also partially and fully digested with *Bgl*I during a time-course experiment, as indicated by the wedge.

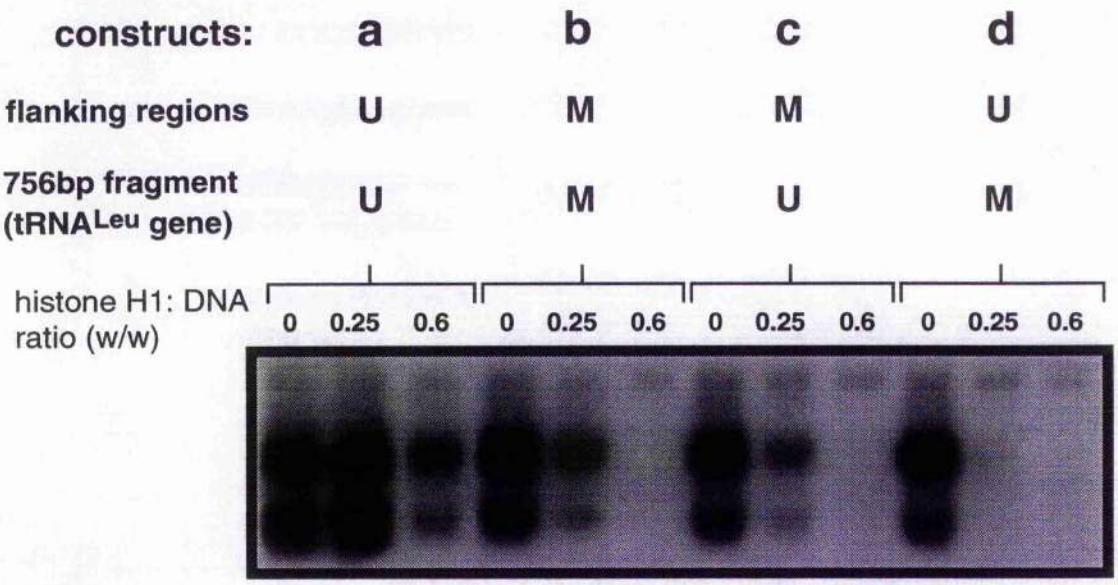


Figure 5.18

Flanking region methylation causes preferential inhibition of transcription of the tRNA ^{Leu} gene by histone H1:

Patch-methylated constructs of the pArg/Leu template were constructed as described in Section 5.3.7 and Figures 5.17A and B. The unmethylated construct (construct a) is inhibited only at higher levels of histone H1 (>0.6 w/w), compared with the fully-methylated construct b or the regionally-methylated constructs c and d. Constructs b–d are strongly inhibited at an H1:DNA ratio of 0.25 w/w.

of forming circular constructs is 31/32, compared to that for linear concatamers which is 1/32.

Covalently-closed circular patch-methylated templates were constructed which were methylated either in all regions of plasmid pArg/Leu except a *Bgl* I-*Bgl* I fragment of size 756 bp. (Figure 5.17A) which contained the $tRNA^{Leu}$ gene, or methylated only in this 756 bp. fragment. The 756 bp. fragment which contained the $tRNA^{Leu}$ gene was either unmethylated, with all the flanking regions methylated (construct c; refer to Figure 5.17A), or was itself methylated with all other regions unmethylated (construct d). Control templates were either fully methylated (construct b; refer to Figure 5.17A) or unmethylated (construct a), and were constructed by ligating together only methylated or unmethylated *Bgl* I fragments. Ligations were carried out at 16°C for 16 h using T4 DNA ligase (Promega) at a concentration of 3 Weiss U/μg DNA (refer to Section 2.7.1.11). Religated covalently-closed circular plasmid (Figure 5.17B) was separated and purified from unligated fragments by gel electrophoresis in 0.7% LMP agarose (Gibco/BRL).

It is clear that methylation of the flanking region can preferentially reduce transcription of the unmethylated $tRNA^{Leu}$ gene in the presence of increasing levels of histone H1 (Figure 5.18), as well as the transcription from the methylated $tRNA^{Arg}$ gene in the flanking region. Construct c which contains an unmethylated $tRNA^{Leu}$ gene in a methylated vector is inactivated at similar H1:DNA ratios, as is the fully methylated construct b. Construct d which contains the methylated $tRNA^{Leu}$ is inactivated to the greatest extent. Inactivation appears to spread to include the adjacent unmethylated $tRNA^{Arg}$ gene. It is impossible to be certain that the template used in these experiments is covalently-closed circular, because the sequence of two *Bgl* I sites is unknown. However, irrespective of the nature of the template, be it circular or linear, it is clear that methylated flanking regions can influence the transcription from an adjacent unmethylated region. This observation is supported by the experiment described in Section 5.3.4 and Figure 5.12: histone H1 can preferentially inhibit transcription from either covalently-closed circular or linear methylated template.

This method of making patch-methylated constructs appears to be quite promising, but a plasmid with known, unique *Bgl* I sites should be used. The yield of construct is poor because the efficiency of fragment religation is low. The yield is much better for the protocol described in Section 2.7.3.2, which forms the basis of the studies presented in Chapter 6, and is therefore the preferred method of patch methylation.

5.4. Discussion

H1-DNA complexes as a model for inactive chromatin

Complexes of histone H1 and DNA have been used in several studies to model inactive chromatin. This approach is valid if it is assumed that in chromatin the histone H1 molecule has minimal interaction with the core histones or the dyad of the chromatosome, but interacts mainly with the linker DNA. Histone H1 molecules are then presumed to form an ordered array along the length of the naked DNA. This suggests that the presence of a chromatosome is not required for the repression of genes by histone H1, at least for the *in vitro* systems studied in this thesis. However, under certain conditions, such as a high H1:DNA ratio or high ionic strength (Clark and Thomas 1986), histone H1 and DNA form an aggregate. The aggregation is presumably mediated by cooperative hydrophobic interactions between histone H1 molecules. Thomas et al. (1992) demonstrate that neighbouring globular domains of histone H1 cooperatively bind to DNA, but this can occur at protein:DNA w/w ratios of 0.1–0.6 and at 5 mM NaCl with the formation of "fast" and "slow" complexes (refer to Section 5.2.1 and Clark and Thomas 1986). The linker histone constrains two adjacent double helices, and additional linker histones stack between the helices with cooperative interactions. The globular domains bridge "tramlines" of DNA and therefore bind to two different segments of DNA, which is presumed to also occur in nucleosomes. Aggregation is the facile explanation for the repression of genes *in vitro* , but the conditions that were used in this series of experiments have precluded the formation of histone H1-DNA aggregates. For example, the repression mediated by histone H1 is reversible on addition of excess competitor DNA (Figure 5.6). In addition, a template that is complexed with histone H1 can still be readily transcribed (Figures 5.7 and 5.16), albeit at a reduced efficiency, which would be impossible if the template was in the form of an aggregate. Thomas et al. (1992) use conditions similar to those described in Chapter 5 (essentially, protein:DNA ratios of 0.25–0.6 and a NaCl concentration of 15 mM; refer to Section 2.7.5.2), to study an histone H1-DNA complex that can be characterised by using biochemical techniques and is not an aggregate.

In fact, the interactions between histone H1 and naked DNA or DNA in chromatin appear to be similar (Croston et al. 1991). For example, histone H1 has a similar salt dependence when it binds to either DNA or to chromatin (Clark and Thomas 1986). Cooperative binding of H1 to DNA occurs over the same range of salt concentrations as the compaction of nucleosomes into the 30 nm chromatin fibre (Clark and Kimura 1990). H1 appears to interact mostly with linker DNA in chromatin, rather than with core histones (Thoma et al. 1979), although the globular domain of the molecule can be cross-linked to H3, H2A and H2B in the core octamer (Pruss et al. 1995). The crystal structure of the globular domain of histone H5 (a linker histone that resembles histone H1; see Section 1.8.7) has been solved (Ramakrishnan et al. 1993), which imposes several constraints on previous models of chromatosome structure (reviewed in Ramakrishnan 1994). None of the new structure data suggest that the globular domain has a strong interaction with the central turn of DNA that wraps round the core octamer. Instead, it has been proposed that the globular domain interacts with both DNA linkers at some distance from the core octamer. However, this model appears to be incompatible with neutron scattering and cross-linking data (Graziano et al. 1994; Boulikas et al. 1980).

Jerzmanowski and Cole (1990) describe the transcriptional inhibition of an unmethylated plasmid at H1:DNA ratios 0.4–0.6 w/w. This corresponds well with the values of 0.25–0.6 w/w ratio for inhibition of the plasmid templates used in this study. These H1:DNA ratios correspond to one molecule of H1 per 125–55 bp DNA. It is generally accepted that the nucleosome core particle contains 146 bp of DNA (Paranjape et al. 1994), and that there is 39 bp of linker DNA per nucleosome that is accessible to H1 in the chromatin of mammalian cells. The observation that the spacing of histone H1 molecules is quite similar in both H1-DNA complexes and chromatin suggests that the interactions of H1 with DNA are similar in naked DNA and in chromatin (Croston et al. 1991). In addition, the reversible binding of H1 to DNA *in vitro* under certain conditions, is comparable to the exchange of H1 within chromatin *in vivo* (Louters and Chalkley 1985). There have also been two recent studies of H1-mediated inactivation of class II genes on naked DNA templates (Croston et al. 1991; Levine et al. 1993). In Section 5.2 I have shown that histone H1 forms ordered complexes with DNA under

certain concentrations and ionic conditions which precluded the formation of aggregates. At these concentrations of DNA and histone H1 all the input histone H1 forms ordered complexes with the DNA. No difference is seen in the ability of histone H1 to form complexes on either unmethylated or methylated DNA, as assayed by retardation in agarose gels.

In summary, ordered complexes of histone H1 and DNA were chosen as a simplified model system for inactive chromatin because the interaction of histone H1 with naked DNA in such complexes resembles the interaction with chromatin.

Histone H1 preferentially inhibits transcription from methylated templates.

The initial observation which prompted the studies reported in this chapter was that *in vitro* transcription from low levels of a methylated template was less efficient than from low levels of an unmethylated template (Section 5.3.1). This suggested the presence of an inhibitor in the nuclear extract that is specific for methylated DNA. There is no evidence that the binding of transcription factors to pol III genes is affected by DNA methylation, but HeLa nuclear extract does contain low levels of histone H1 (Croston et al. 1991; Paranjape et al. 1994), and would also be expected to contain the methyl CpG binding protein, MeCP-1. Two methods were used to remove histone H1 from nuclear extracts.

Addition of competitor DNA, either methylated or unmethylated, led to enhanced transcription from both templates implying that an inhibitor was present that bound to both methylated and unmethylated DNA (Section 5.3.2 and Figure 5.6). This inhibitor cannot, therefore, be MeCP-1 (that binds only to methylated DNA) but could be histone H1 that has a similar affinity for methylated and unmethylated DNA. The differential inhibition from the methylated template was considerably reduced in the presence of high levels of competitor DNA which is consistent with the structure of the complex between histone H1 and methylated DNA being different from that formed with unmethylated DNA (Higurashi and Cole 1991). Addition of histone H1 alone is able to restore the strong preferential inhibition of transcription from the methylated template (Figure 5.7). Presumably, the methylated competitor has removed methyl CpG binding proteins as

well as histone H1 and so these cannot be involved in the selective inhibition exerted by histone H1 on transcription from the methylated template.

Depletion of the nuclear extract of histone H1 using ammonium sulphate precipitation also abolished the preferential inhibition from the methylated template (Section 5.3.3 and Figure 5.9). This differential effect could be restored by addition of histone H1 alone, again indicating that this protein is able to preferentially inhibit *in vitro* transcription from the methylated template. In this case, it is presumed that MeCP-1 remains in the H1-depleted nuclear extract and it alone is insufficient to inhibit transcription from methylated templates. Previous evidence and our own data suggests that histone H1 can act as a methylated-DNA binding protein (MDBP; see Section 1.7.2), which is supported by the observation that MDBP-2 has sequence homologies with histone H1 (Jost and Hofsteenge 1992). The inhibition of transcription by an MDBP-like activity of histone H1 could therefore supplement the generalised repression of chromatin by histone H1, presumably by the formation of supranucleosomal structures (see Section 1.8.5).

Histone H1 functions as a repressor of transcription for both pol II transcribed genes (Bresnick et al. 1991; Laybourn and Kadonaga 1991) and pol III transcribed genes (Schlissel and Brown 1984; Shimamura et al. 1989; Wolffe 1989) in the context of chromatin or nucleosomes (see Section 1.8.5). I show here that this inhibition is more efficient for methylated pol III templates. However, histone H1 does not prevent RNA chain elongation, which suggests that the movement of the transcription complex is not impeded by the H1-DNA complexes. The results in Figure 5.16 show that histone H1 exerts its differential effect on transcription of methylated DNA during the formation of the initiation complex. This observation is supported by a study on the expression of the herpes simplex virus (HSV) thymidine kinase (*tk*) gene, after transfection of a construct containing the gene into mammalian cells (Levine et al. 1992). Promoter methylation appeared to affect the formation of the initiation complex, rather than the rate of transcription.

The binding of histone H1 to methylated DNA may lead to a change in nucleoprotein conformation, thereby rendering the promoters inaccessible to transcription factors. Higurashi and Cole (1991) have shown that histone H1-DNA complexes are resistant to digestion by *Msp* I (see Section 1.8.5), and have also suggested that this can be explained by a local change in the conformation of the nucleoprotein complex. These

findings are confirmed by my own work, presented in Section 6.2.2. Figure 5.18 shows that methylation of flanking regions alone is able to bring about H1-mediated inhibition of transcription of an unmethylated gene. This suggests that the conformational change can spread, either by the cooperative binding of histone H1 molecules or by a change in the double helix. This effect appears to only occur over short distances (the $tRNA^{Leu}$ gene is 242 and 432 bp from the start of the flanking regions; refer to Figure 5.17A). Histone H1-DNA complexes do not appear to mediate the spread of inactivation over greater distances: there is no evidence of this effect being mediated by histone H1 on the patch-methylated constructs for the pol II system, described in Sections 6.3.2 and 6.3.4.

The possible roles of histone H1 variants

Different histone H1 variants occur in different species, cell types and developmental stages (see Section 1.8.7), which has implications for chromatin structure and folding (Pruss et al. 1995). In some vertebrates, the presence of certain histone H1 variants has been correlated with distinct developmental stages, that involve changes in chromatin structure prior to changes in gene transcription. For example, the special linker histone H5 of chicken erythrocytes compacts and inactivates the entire erythrocyte nucleus by the formation of heterochromatin and in the sea urchin there are well-characterised changes of histone H1 variants during embryogenesis (see Section 1.8.7).

I have shown that different somatic variants of calf thymus histone H1 inhibit transcription to different extents (Figures 5.14 and 5.15). Recently, a mixture of two variants, H1c and H1e, has been shown to specifically inhibit *in vitro* enzymatic DNA methylation (Santoro et al. 1995) and it has been proposed that these variants act at CpG islands to inhibit their methylation by DNA methyltransferase, thereby maintaining the islands in an unmethylated state. This follows from the finding that H1c is the only variant that can bind to DNA that is rich in CpG dinucleotides (6 CpG dinucleotides per 44 bp). Here I demonstrate that H1c is particularly effective at inhibiting *in vitro* transcription from methylated templates where the overall CpG concentration is approximately 400 per 5600 bp (i.e. about three 5-mCpG dinucleotides per 44 bp). Island DNA undergoing maintenance methylation during DNA replication could thus be packaged into inactive chromatin by H1c, whereas unmethylated genes would remain active.

CHAPTER SIX

The effect of methylation on *in vitro* chromatin formation and transcription

6.1 Introduction

This chapter describes a series of *in vitro* studies that investigate the inhibition of transcription by methylation and chromatin. The experiments use as template the plasmid pVHCk (see Section 2.5.1 for a map of the plasmid), which is transcribed by RNA polymerase II (pol II). One experiment (presented in Section 6.2.3) uses pArg/Leu as the template. Plasmid pArg/Leu (see Section 2.5.1 for a map) is transcribed by RNA polymerase III (polIII). The protocol for *in vitro* transcription of the pol II system has been described in Section 2.7.4.3, and experiments that detail the optimisation of this protocol are presented in Section 3.3.4. The method for the *in vitro* reconstitution of chromatin on plasmid DNA, by using *Xenopus* S150 egg extract, has been described in Section 2.7.7 and several experiments that were based on this system were presented and discussed in Chapter 4. The data in this chapter have been obtained by using these systems of *in vitro* transcription and chromatin assembly to investigate the inhibition of pol II transcription by methylation.

The results presented in Section 6.2.3 demonstrate the preferential inhibition of transcription from methylated chromatin templates (i.e. chromatin that was *in vitro* reconstituted on methylated plasmid DNA, that was subsequently used as the template for *in vitro* transcription assays). This is caused by methylation directing the formation of an inactive chromatin structure that is inaccessible to transcription factors (Lewis and Bird 1991; Graessmann and Graessmann 1993), as has been discussed in Section 1.8.2. Inhibition in the context of chromatin is shown for both the pol III system (that uses pArg/Leu as the template) and the pol II system (that uses pVHCk), as discussed in Section 6.2.2. This observation has been made in previous studies (see Section 1.8.3),

and serves as a necessary control for subsequent experiments. A similar preferential inhibition of methylated pol III and pol II templates by histone H1 has been discussed at some length in Chapter 5 (see Section 5.3.4. in particular). Inhibition of transcription by histone H1 supports the argument that histone H1-DNA complexes are a useful model of chromatin for use in transcription studies. In view of the results in Chapter 5, the protocol for pol II transcription (see Section 2.7.4.3) uses a nuclear extract that is depleted of histone H1 (prepared according to Shapiro et al. 1988), which simplifies the interpretation of these transcription experiments. In addition, chromatin has been reconstituted on methylated DNA in the presence of exogenous histone H1 (see Section 4.2.2), for which preferential inhibition of transcription is also observed (Section 6.2.3). I argue that this is a valid model of heterochromatin in Section 6.4. Methylated DNA in chromatin is shown to be inaccessible to *Msp* I (Section 6.2.2), which examines the accessibility of the chromatin at specific locations. By comparison, unmethylated DNA in chromatin is relatively hypersensitive to this and other restriction enzymes and, by analogy, to transcription factors (Graessmann and Graessmann 1993; also see Section 1.8.2). This would explain the preferential inhibition of transcription from methylated chromatin templates.

The second series of experiments, described in Section 6.3, extends previous observations made in this laboratory. Kass et al. (1993) have shown that regional DNA methylation on plasmid pVHCk acts as a focus for the formation of inactive chromatin, when the regionally-methylated plasmid is transfected into mouse L929 fibroblasts. The 1993 study showed that regional methylation of the CpG-rich prokaryotic pBluescript KS- vector DNA, which resembles a CpG island, could inactivate transcription even when the promoter was unmethylated. The regional methylation could also seed the spread of an *Msp* I-inaccessible chromatin conformation into adjacent unmethylated regions. However, the 1993 study did not examine the chromatin structure of transfected plasmid, as assayed by limited digestion of cell nuclei with staphylococcal or micrococcal nuclease. The experiments that did examine the fate of transfected plasmid could not provide the necessary evidence for the formation of nucleosomes on the minichromosomes. The *Msp* I fade-out assays that were presented in Kass et al. (1993) provide somewhat weaker evidence for the spread of an *Msp* I-inaccessible chromatin

conformation. Another limitation in the 1993 study is that the transcription assay in this system analysed CAT activity in cell extracts that were prepared 48 hr. post-transfection, which does not provide a direct measure of transcriptional initiation at the SV40 promoter.

I have therefore investigated the spread of inactive chromatin by using *in vitro* chromatin formation and transcription on the pVHCk plasmid, using several patch-methylated plasmid constructs. The presence of nucleosomes is assessed by the formation of a nucleosomal ladder on digestion of chromatin by staphylococcal nuclease (see Section 4.2.1). That the initiation of transcription occurs at the expected sites was confirmed by using the primer-extension protocol described in Section 2.7.4.3 and discussed in Chapter 3. The use of *in vitro* systems extends the transfection studies of Kass et al. (1993). However, because the nuclear and egg extracts contain many components other than the transcription and chromatin assembly machinery, the identities of the proteins that mediate the spread of inactive chromatin remain unknown. This *in vitro* approach is therefore purely phenomenological, in contrast to the rather more analytical experiments with histone H1-DNA complexes in Chapter 5. It could be rendered more informative by using purified histones in chromatin assembly (refer to the discussion in Section 4.1).

6.2 *In vitro* chromatin formation and transcription on fully methylated plasmid templates

6.2.1 Assembly and structure of chromatin templates

Chromatin was reconstituted on dsDNA using a crude cytoplasmic fraction derived from *Xenopus* eggs, designated S150 extract. This was prepared essentially as described by Wolffe and Schild (1991) and chromatin was assembled by following a protocol described by Rodríguez-Campos et al. (1989), and is described in detail in Section 2.7.7. In some experiments the reaction mixture was supplemented with histone H1, added at a 0.6 w/w ratio with respect to the DNA. Renatured histone H1 was prepared from lyophilised total acid-extracted calf thymus histone H1 (Boehringer Mannheim GmbH) as

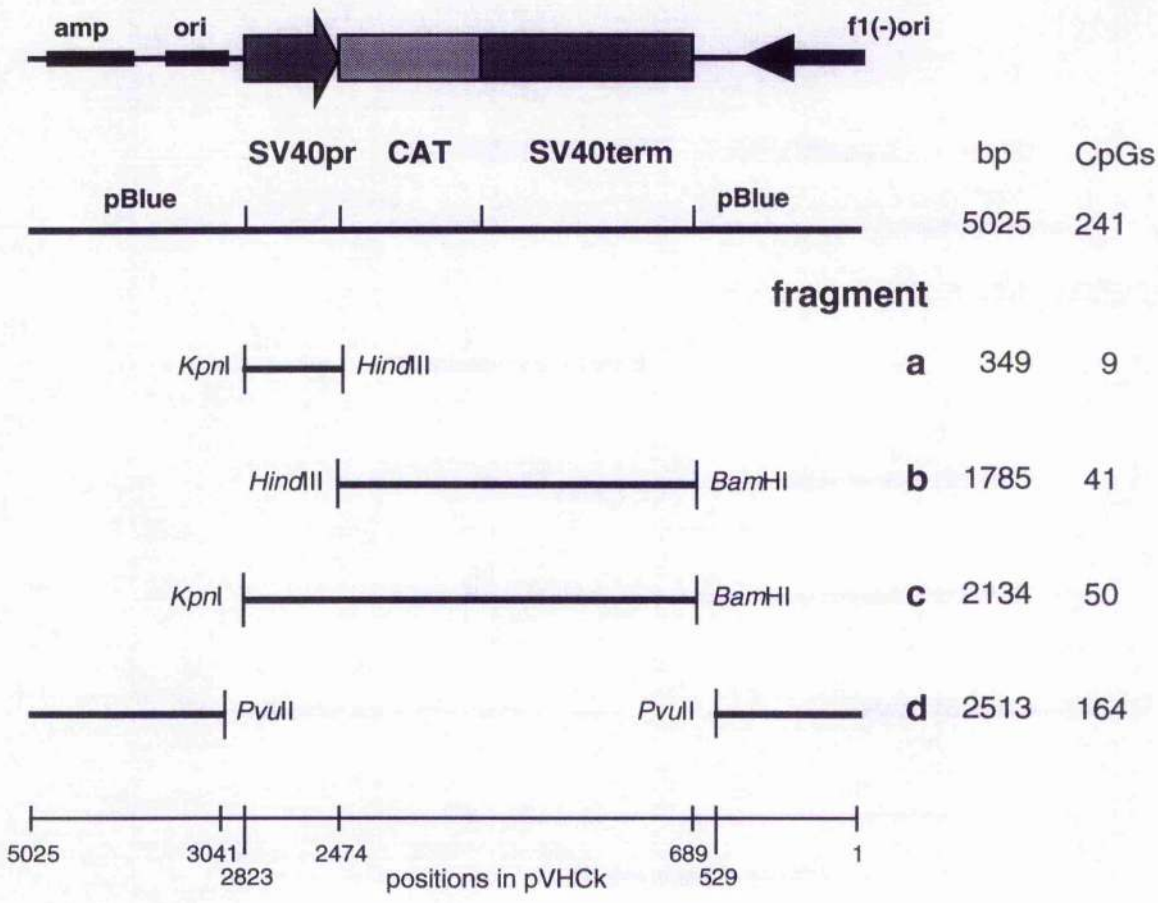


Figure 6.1

Linear map of plasmid pVHCk showing location of patches:
Linear representation of the circular plasmid pVHCk and the location of regional methylation in patch-methylated constructs. The plasmid contains the chloramphenicol acetyltransferase gene (CAT) under the control of the SV40 early promoter (SV40pr) and the SV40 terminator region (SV40term) inserted into pBluescript KS- (pBlue). Horizontal bold lines show the restriction fragments (fragments a–d) used for the regional methylation and Southern blot analysis, with the location of relevant restriction sites and their positions in pVHCk indicated. See Section 6.3.1 for further details. A map of pVHCk is also shown in Figure 2.2.

described previously (Johnson et al. 1995; also refer Section 2.7.5). If the chromatin template was to be used subsequently for *in vitro* transcription assays, then the appropriate amount of HeLa nuclear extract was also included (see Section 2.7.4). The reaction mixture was incubated at 30°C for 2 hr, after which the extent of chromatin reconstitution was assayed immediately by one of the methods described below. The structure of chromatin reconstituted on plasmid DNA was established by treatment of the chromatin assembly reaction mixture with either staphylococcal nuclease or *Msp* I, as described in Section 2.7.7.3. The results of staphylococcal nuclease digestion are presented in separate sections, depending on the methylation status of the DNA. Section 4.2.1 deals with the digestion of fully methylated plasmid, and Section 6.3.3 describes the structure of chromatin reconstituted on patch-methylated constructs. The data from time courses of *Msp* I digestion are also presented below. Digested DNA fragments are separated on an agarose gel, Southern blotted and hybridised to the desired fragment (see Section 2.7.1.21). The fragments that are used in this series of experiments are shown in Figure 6.1. The degree of digestion of the topmost intact band is quantitated on the blot in so-called fade-out assays.

6.2.2 Methylation affects *Msp* I sensitivity of DNA, histone H1-DNA complexes and chromatin

The sensitivity of control or fully-methylated supercoiled pVHCk plasmid to limited digestion with *Msp* I was used to assay the accessibility of CCGG sites in four different substrates: naked plasmid DNA, histone H1-DNA complexes, DNA reconstituted with chromatin and DNA reconstituted with chromatin in the presence of histone H1. The procedure used to digest the first two types of substrate was essentially that described in Section 2.7.7.3, except that ten-fold less enzyme was used (1.0 unit *Msp* I/μg DNA) to prevent over-digestion of the DNA substrates. Each substrate was assayed in triplicate, except the chromatin with histone H1 substrate which was assayed in duplicate. The digestion products were hybridised to fragment d (refer to Figure 6.1), and typical autoradiographs for all four types of substrate are shown in Figure 6.2. The disappearance of the labelled full-length band was quantitated using a phosphoimager.

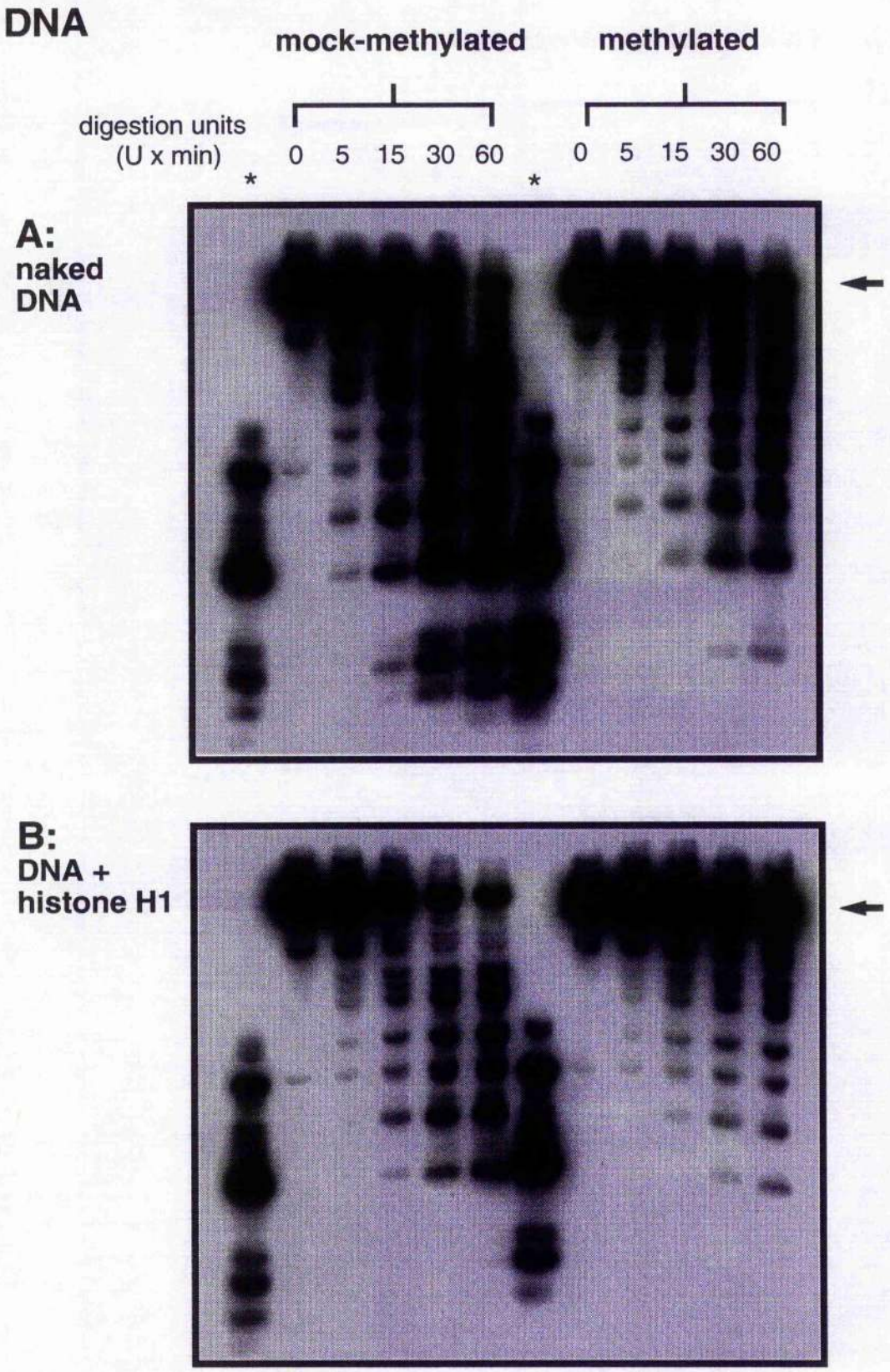


Figure 6.2 A and B

***Msp* I fade-out assay of mock-methylated and methylated DNA, in the absence and presence of histone H1:**
see overleaf for figure legends

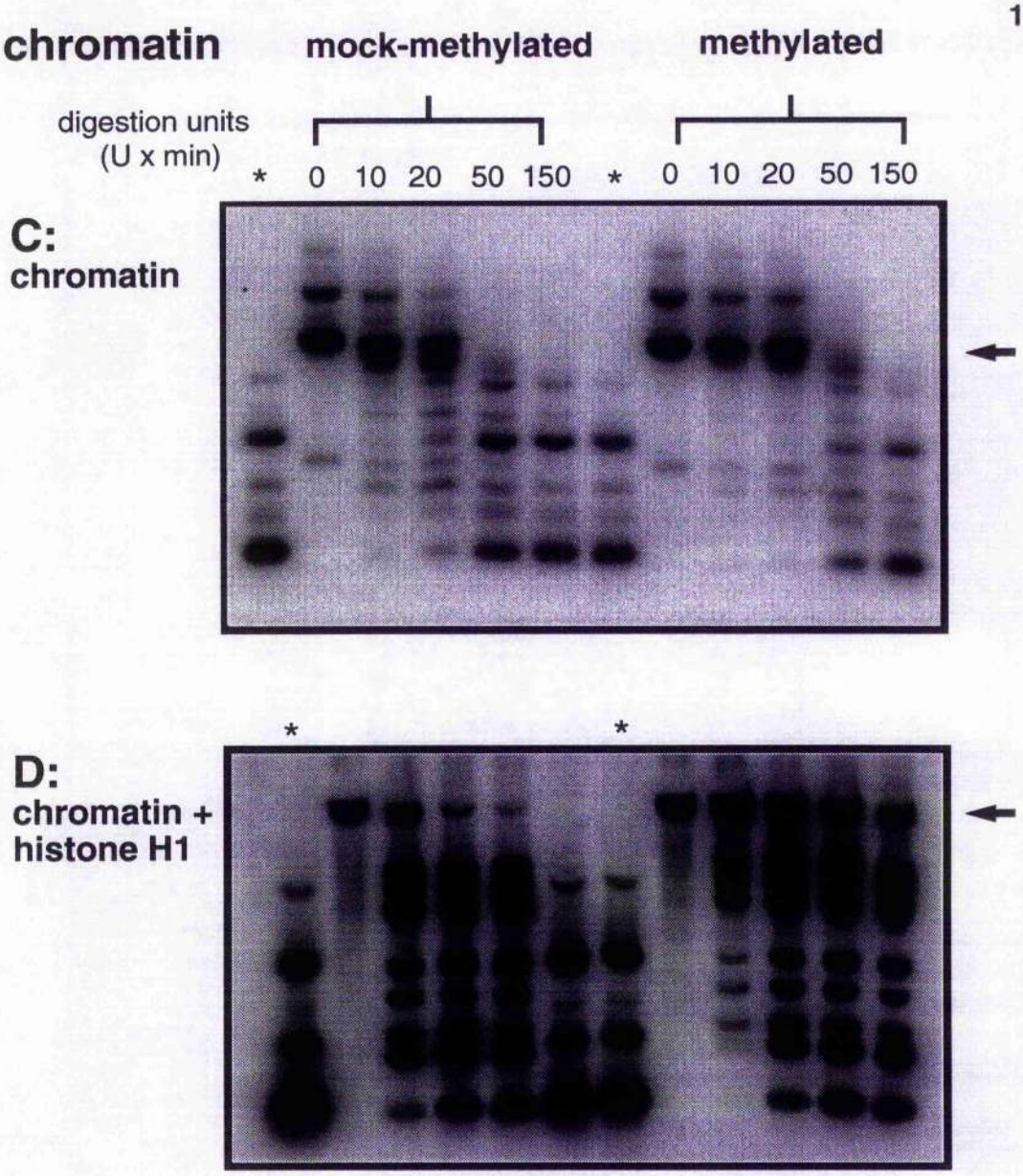


Figure 6.2 C and D

***Msp* I fade-out assay of mock-methylated and methylated chromatin, in the absence and presence of histone H1:**

The sensitivity of control mock-methylated or fully methylated supercoiled pVHCk plasmid to limited digestion with *Msp* I was used to assay the accessibility of CCGG sites in four types of substrate (see Section 6.2.2 for further details):

Figure 6.2 A: naked DNA; **B:** histone H1-DNA complexes.

Figure 6.2 C: chromatin, in the absence of histone H1; **D:** chromatin, in the presence of histone H1.

The digestion products were hybridised to fragment d (see Figure 6.1) and typical autorads for each substrate are shown. Complete digestion of pVHCk by *Msp* I is indicated by the asterisk (*), and the intact band (corresponding to fragment b) that was used in quantifications is indicated by the arrow. Refer to Section 6.2.2 for an explanation of digestion units.

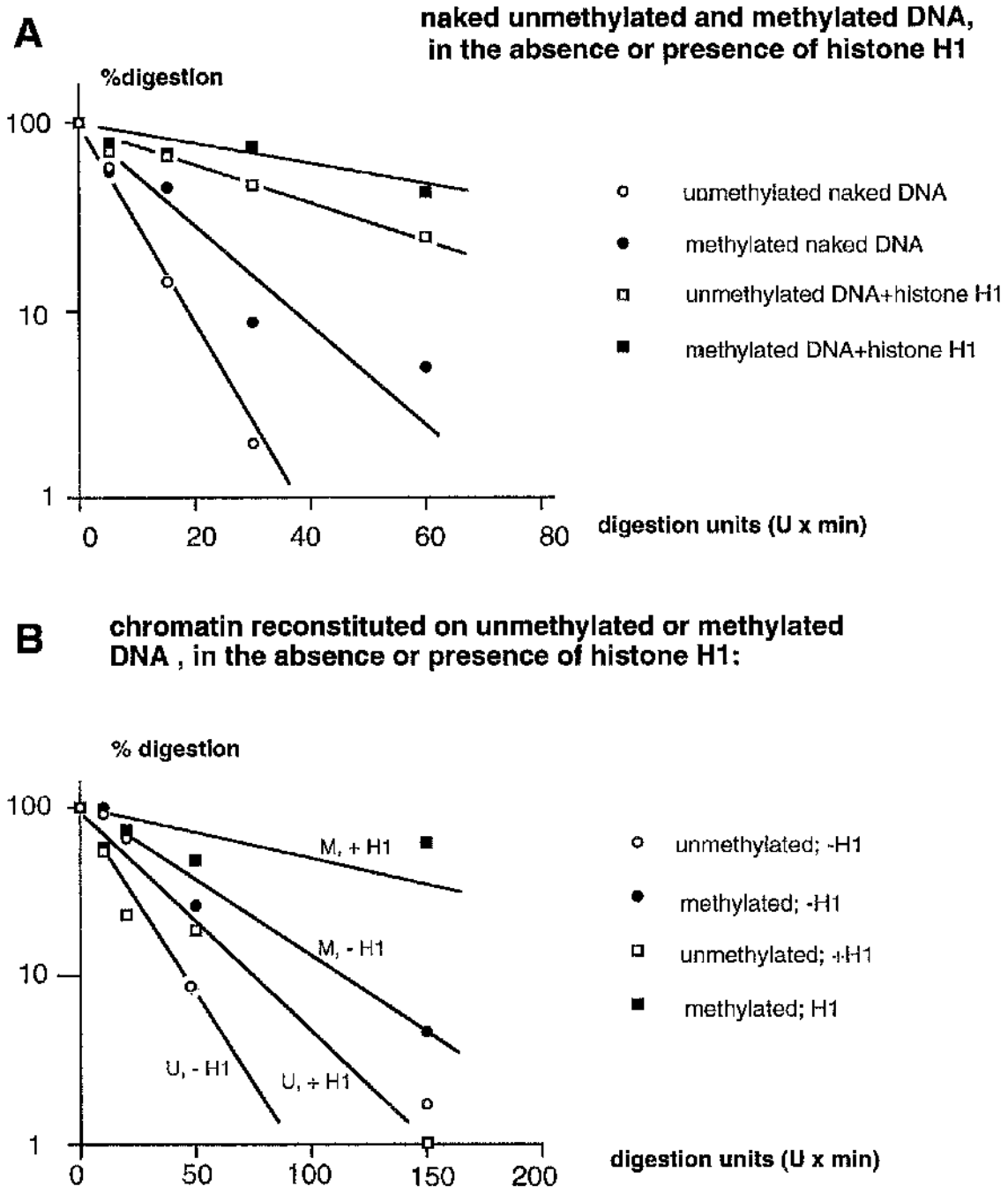


Figure 6.3

Rate of *Msp* I digestion of DNA, histone H1-DNA complexes and chromatin in the absence and presence of histone H1:
The disappearance of the labelled intact band in fade-out assays was quantified for each type of substrate on either unmethylated or fully methylated pVHCk template. These results are shown in semi-log plots, with average values plotted as a percentage of the zero-time value. See Section 6.2.2 for further details.
A: Plots for naked unmethylated (U) or methylated (M) DNA, and for histone H1-DNA complexes, with points shown as an average of values in triplicate.
B: Plots for chromatin, in the absence and presence of histone H1, with points shown as an average of values in triplicate or in duplicate, respectively.

These results are shown in semi-log graphs in Figure 6.3, with the average value for separate assays plotted as a percentage of the zero-time value. Lines in these graphs are biased to consider the points for high % digestion values, because the line of best fit did not appear to be close to the points for the lowest % digestion values on a logarithmic scale. However, it was confirmed that the line of best fit did pass through the error bar for these lowest values (error bars not shown). The abscissa is expressed as the product of units of enzyme and time of digestion in minutes, to allow the comparison of substrates of different sensitivities. For convenience, this product is called the digestion unit.

In a separate series of graphs all points for a digestion experiment were considered, rather than the average value as shown in the graphs of Figure 6.3 (results not shown). The possible range of lines was then drawn, which provides an estimate of sample deviation for each type of substrate. The digestion units required for 50% loss of the full-length band was used as a measure of the gradient of each line. The average value of the gradients for each type of substrate are shown in Table 6.1. The sample deviation of each set of gradients is expressed as a percentage value of the average gradient.

Table 6.1

Accessibility of different substrates to *Msp* I:

digestion units (U x min) required for 50% digestion:

sample deviation (%)	naked DNA ±12 %	DNA+ histone H1 ±14 %	chromatin ±24 %	chromatin+ histone H1 *
unmethylated	4	19	22	21
methylated	7	37	42	60

* insufficent samples

Methylated plasmid DNA is somewhat less sensitive to *Msp* I digestion than is mock-methylated (control) DNA (Figure 6.3A). This experiment, which was carried out in triplicate, may be a consequence of the altered conformation of the methylated plasmid,

or of the reduced rate of cleavage at $C^{m}CGG$ sequences. The rate of digestion of histone H1-DNA complexes is considerably reduced relative to that of naked DNA, but the substrate containing methylated DNA is again less sensitive to limited *Msp* I digestion than is that containing mock-methylated DNA (Figure 6.3). This differential sensitivity is approximately the same as that seen with naked DNA substrate (Table 6.1). When these templates are reconstituted as chromatin using *Xenopus* S150 egg extract the rates of digestion by *Msp* I are very similar to those observed in the presence of histone H1 and the differential effect of DNA methylation is still observed (Figures 6.2C and 6.3B; Table 6.1). In the presence of histone H1 the chromatin reconstituted on methylated DNA is even more resistant to *Msp* I digestion (Figures 6.2D and 6.3B; Table 6.1).

6.2.3 Methylation reduces transcriptional activity of DNA, histone H1-DNA complexes and chromatin

The effect of methylation on the activity of the SV40 promoter in pVHCk was assayed by *in vitro* transcription of mock-methylated or fully methylated pVHCk, in the form of the four types of substrates described above. The transcription assays for these substrates are presented in Section 6.3.4, because the assays for one type of substrate were done at the same time as those for patch-methylated constructs. This allows the direct comparison of fully methylated plasmid and patch-methylated constructs by minimising the variability during the formation of the substrate. This is particularly true for chromatin substrates. However, in this section the analysed data for mock-methylated and fully methylated substrates is presented in Figure 6.4. Transcription of control (mock-methylated) templates decreases in the order: naked DNA (100%), histone H1-DNA complexes (80%), chromatin (70%) and chromatin with histone H1 (40%). In all cases, transcription from the fully methylated DNA template is lower when compared to that from the control mock-methylated templates.

It has been shown that histone H1-DNA complexes preferentially inhibit the *in vitro* transcription of pol III genes compared to complexes formed on unmethylated DNA (see Section 5.3.4 and Johnson et al. 1995). Similar results for the pol II system

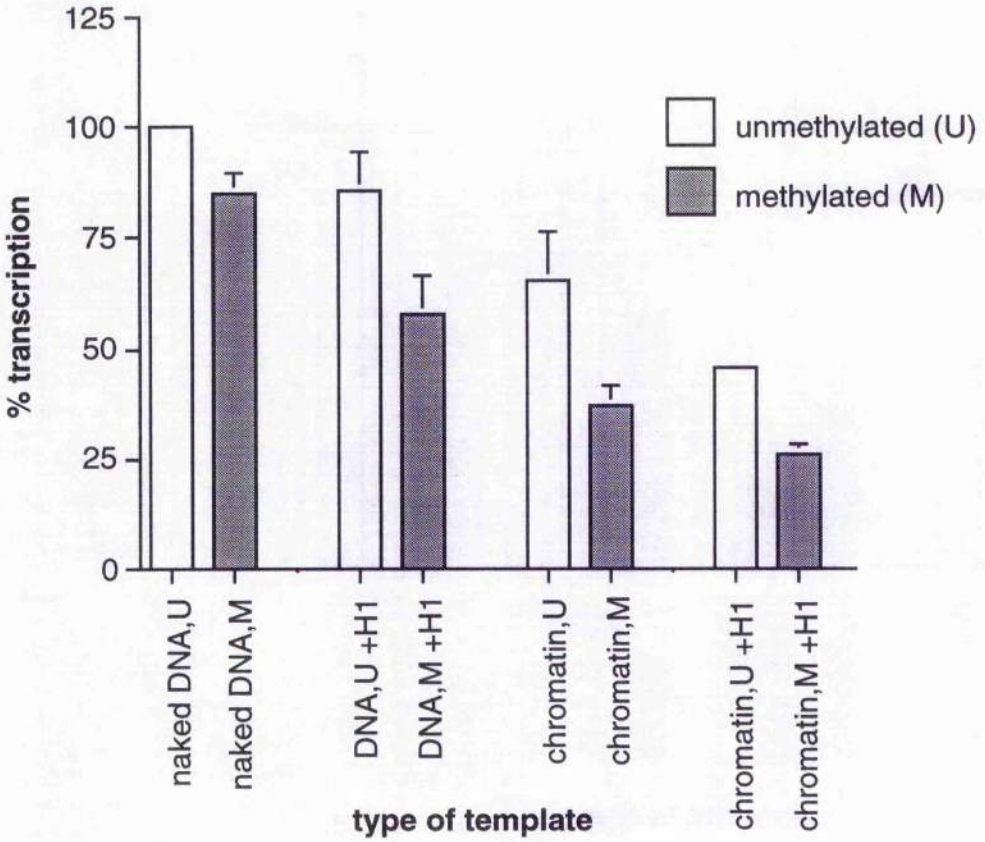


Figure 6.4

***In vitro* transcription of pVHCk, using different types of mock-methylated or fully methylated templates:**
The effect of methylation on the activity of the SV40 early promoter was assayed by *in vitro* transcription using four substrates: naked pVHCk plasmid, histone H1-DNA complexes, DNA reconstituted with chromatin and DNA reconstituted with chromatin in the presence of histone H1, as indicated on the abscissa of the histogram. See Section 6.2.3 for further details, and Figure 6.11 for gels of the original transcription assays. In all cases, transcription is reduced from the methylated pVHCk template (M), as compared with the control mock-methylated pVHCk template (U).

are seen in Figure 6.4, at the same level of histone H1 that was used for the pol III study (a histone H1:DNA w/w ratio of 0.6), although the degree of preferential inhibition is not as pronounced. This confirms a previous experiment on the pol II system, also described in Section 5.3.4. The *in vitro* transcription of naked methylated DNA is also preferentially inhibited compared to naked unmethylated DNA (Figure 6.4). Since the HeLa nuclear extract that is used for the transcription studies is depleted of endogenous histone H1 (Section 2.7.4.3), the preferential inhibition of naked methylated DNA indicates the presence of a limiting amount of a selective inhibitor in the nuclear extract, as described previously (Johnson et al. 1995). Similar results have been described for transcription of pol II genes by Boyes and Bird (1991), who suggested that MeCP1 was the limiting protein.

Chromatin was reconstituted on methylated DNA in the absence and in the presence of exogenous histone H1. Transcription from both of these substrates is also preferentially inhibited (Figure 6.4), although unmethylated templates are inhibited when compared to naked DNA templates. Other studies have described the inhibition of transcription by nucleosomes and histone H1 in more detail (Section 1.8.5). It is known that both full methylation of pVHC1 (Bryans et al. 1992) and regional methylation of pVHCk (Kass et al. 1993) can inactivate the SV40 early promoter after transfection into mammalian cells, although viral SV40 early gene expression is insensitive towards methylation (Graessmann et al. 1983; Graessmann and Graessmann 1993). In this, the results are similar to those observed for histone H1-DNA complexes. However, histone H1-DNA complexes cannot mediate the spread of inactivation from a region of methylation (Section 6.3.4) that can be mediated by chromatin. The effect of histone H1 in the context of methylated chromatin is to inhibit transcription still further (Figure 6.4), which is reflected in an increase in the inaccessibility to *Msp* I (Table 6.1), as compared to unmethylated chromatin. By contrast, histone H1 does not appear to inhibit unmethylated chromatin to as great an extent.

The effect of methylation on tRNA gene expression was also assayed by *in vitro* transcription of mock-methylated or fully methylated pArg/Leu template, which has been reconstituted with chromatin (Figure 6.5). Assays were performed in duplicate. These templates were reconstituted with chromatin and preincubated with nuclear extract

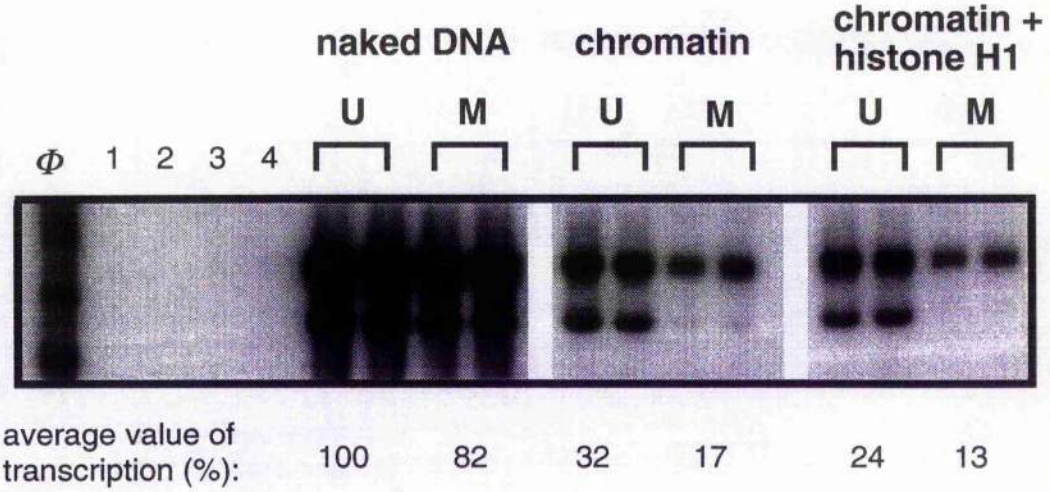


Figure 6.5

In vitro transcription of pArg/Leu, using different types of mock-methylated or fully methylated templates:
The effect of methylation on the activity of the tRNA genes in pArg/Leu was assayed in duplicate by *in vitro* transcription using three substrates, as indicated. In all cases, transcription is reduced from the methylated pArg/Leu template (M), as compared to the control mock-methylated template (U).The level of transcription was quantified for each assay by using a phosphoimager, and the average value for each set of duplicates is shown. The standard 100% value is the average transcription level of naked unmethylated template. The numbers indicate control experiments **1:** unmethylated and **2:** methylated templates, in the absence of HeLa nuclear extract or *Xenopus* egg extract; **3:** unmethylated and **4:** methylated templates, treated only with the egg extract to assay the extent of endogenous transcription in this extract.

simultaneously as described in Section 6.2.1. Control experiments confirmed that the *Xenopus* egg extract did not support endogenous transcription of the template. The effect of chromatin reconstitution was to preferentially inhibit transcription from the methylated template, as shown in Figure 6.5. The addition of histone H1 to the reaction mixture appears to have a further inhibitory effect on both unmethylated and methylated chromatin. These observations support the data presented in Figure 6.4.

6.3 *In vitro* chromatin formation and transcription on patch-methylated pVHCk

This section describes experiments, similar to those in Section 6.2, that were performed on patch-methylated constructs. This allows the direct comparison of the effect of regional methylation on transcription with control experiments that use the fully methylated pVHCk template. The preparation of patch-methylated pVHCk constructs is described in Section 2.7.3.2. These patch-methylated constructs for the pol II system must not be confused with the constructs for the pol III system, which were prepared by the ligation of *Bgl* I fragments of pArg/Leu (see Section 5.3.7).

6.3.1 Analysis of patch-methylated constructs

The preparation of patch-methylated pVHCk constructs yielded double-stranded plasmid constructs that were methylated at CpG dinucleotides only in specific regions, so-called "patches", with an efficiency of 60-80% as reported by Kass et al. (1993). Mock-methylated plasmids and non-patch regions are not methylated. Four different restriction fragments, ranging in size from 349 to 2513 bp (fragments a-d in Figure 6.1) were used to make the patch-methylated constructs of the pVHCk plasmid. The location of methylated patches in all constructs were verified by restriction enzyme digestion with a methyl-sensitive enzyme (typically *Hpa* II) and, if possible, by Southern blot analysis. The construct with fragment. a could not be analysed in this way because the SV40 promoter lacks *Hpa* II sites, and indeed sites for any other methylation-sensitive restriction enzymes. The analyses of constructs b to d are shown in Figure 6.6A–C, with further explanation of the analyses in the figure legends.

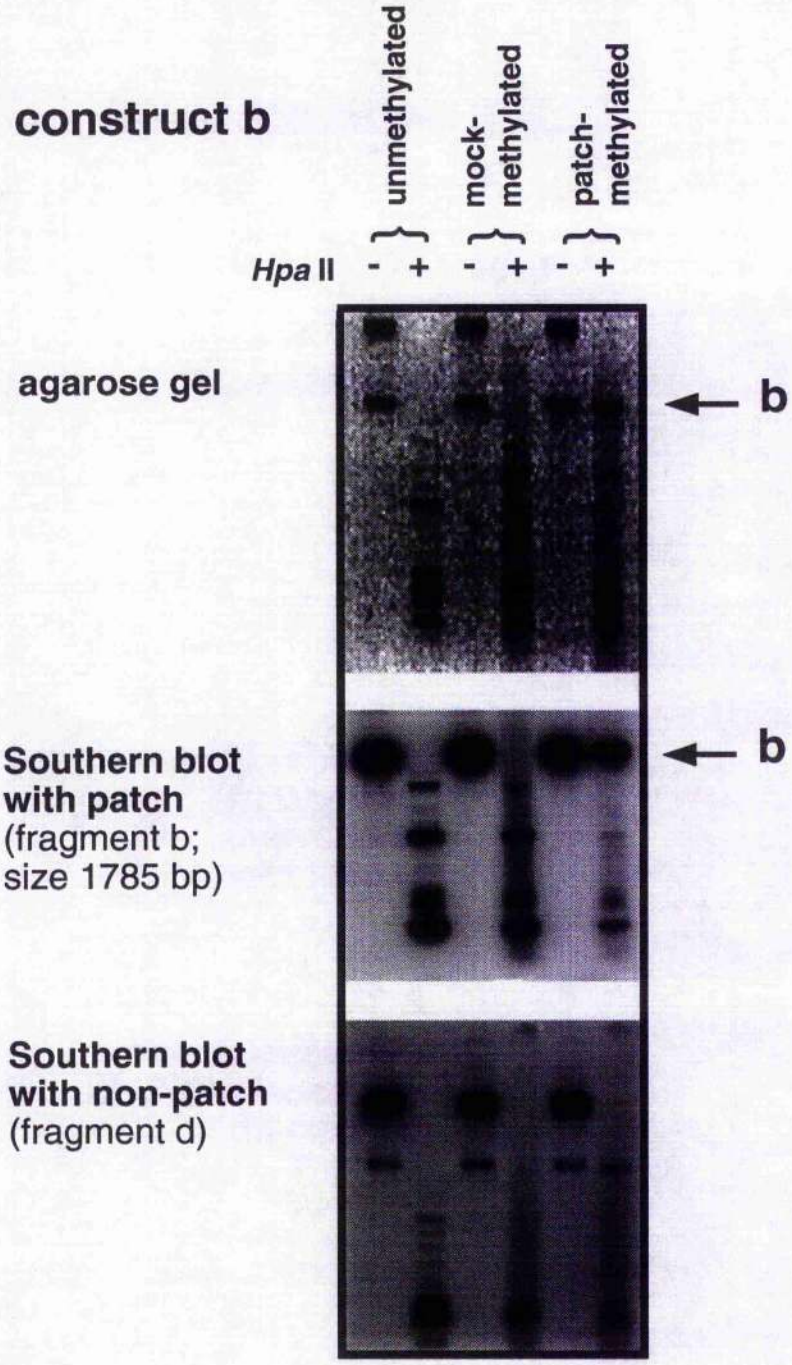


Figure 6.6 A

Analysis of patch-methylated construct b

The patch-methylated construct b was analysed by digestion with *Hin*dIII and *Bam*HI which release the methylated patch, followed by digestion with *Hpa*II, as indicated by +*Hpa*II. The methylated patch is protected against *Hpa*II digestion after the restriction digests were run on an agarose gel (top panel). Fragment b is marked by an arrow. The DNA was then transfered to a nylon membrane and probed with the patch fragment b (middle panel). Fragment b is absent from digests of the two controls (unmethylated pVHCk plasmid and the mock-methylated construct). The blot was then stripped of the probe and re-probed with a non-patch fragment (fragment d; bottom panel), in which the patch fragment is detected at minimal levels in all *Hpa*II digests. See Section 6.3.1 for further details.

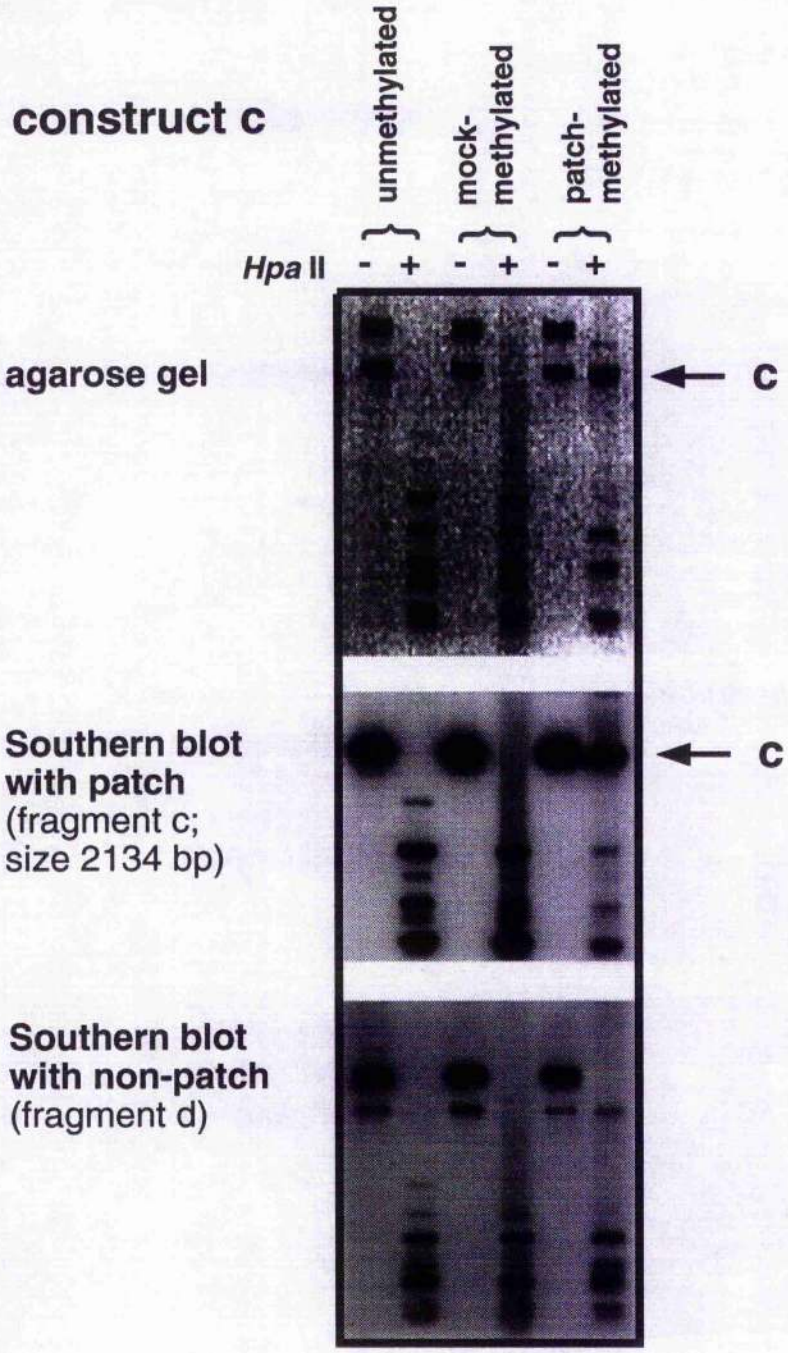


Figure 6.6 B

Analysis of patch-methylated construct c:
The patch-methylated construct c was analysed by digestion with *Kpn*I and *Bam*HI which release the methylated patch, followed by digestion with *Hpa*II. The methylated patch is protected against *Hpa*II digestion, so it is detected when the restriction digests are probed with fragment b (middle panel). Fragment c is marked by an arrow. This band is absent from digests of the two controls (unmethylated pVHCk plasmid and the mock-methylated construct). The blot was then stripped of the probe and re-probed with a non-patch fragment (fragment d; bottom panel), in which the patch fragment is absent in all *Hpa*II digests.

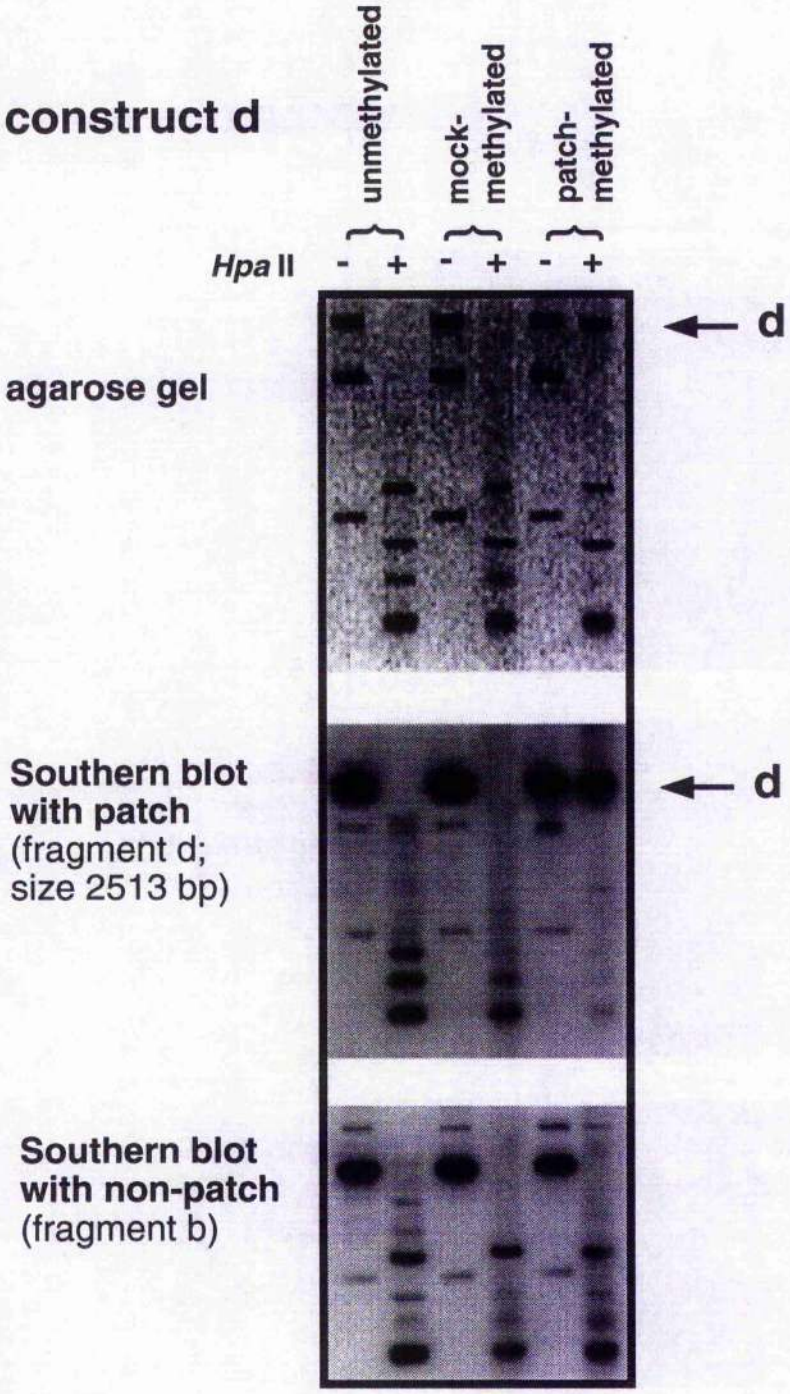
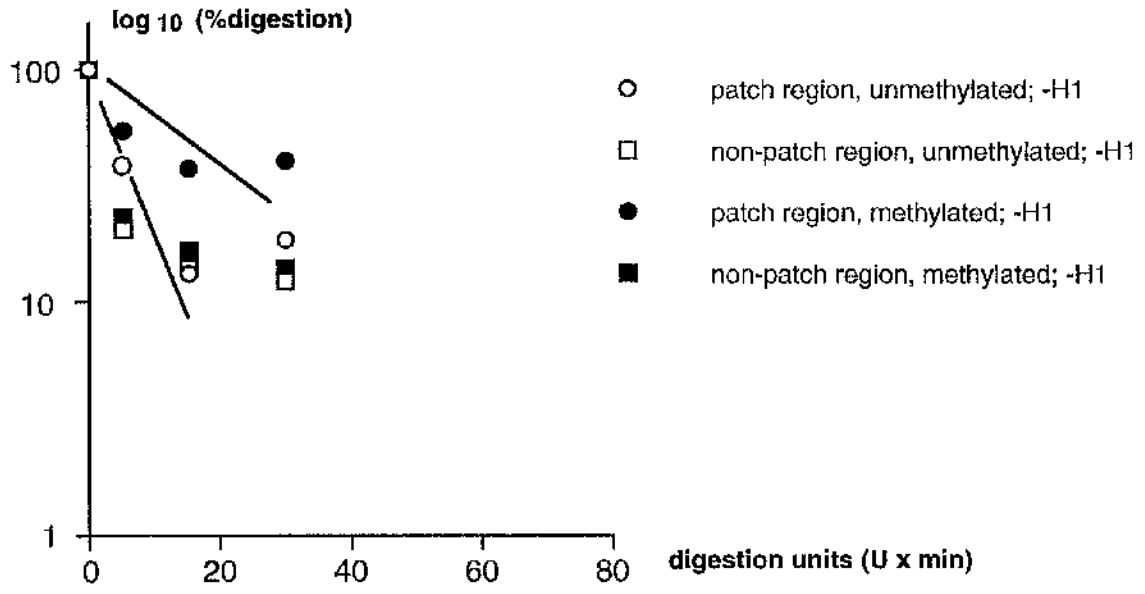


Figure 6.6 C

Analysis of patch-methylated construct d

The patch-methylated construct d was analysed by digestion with *Pvu*II which releases the methylated patch and two other restriction fragments, followed by digestion with *Hpa*II. A blot of the agarose gel (top panel) was probed with the patch fragment d (middle panel) and subsequently with a the non-patch fragment b (bottom panel), as described in Figures 6.6 A and 6.6 B. See Section 6.3.1 for further details.

A **naked mock-methylated and patch-methylated DNA, in the absence of histone H1**



B **naked mock-methylated and patch-methylated DNA, in the presence of histone H1**

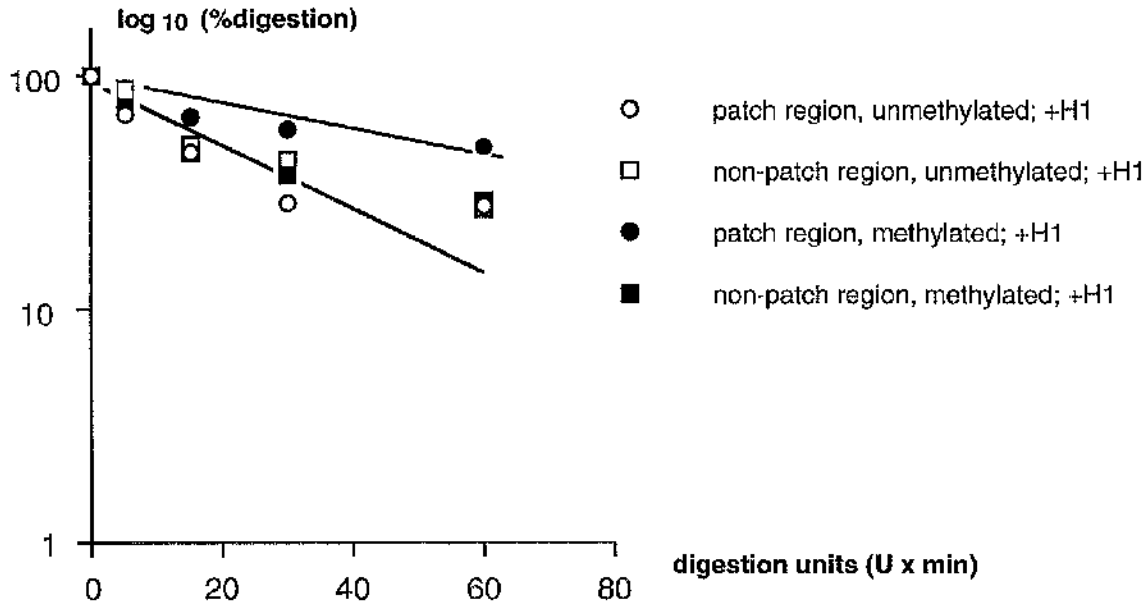


Figure 6.7

Methylation does not reduce *Msp* I sensitivity of unmethylated regions of naked DNA or histone H1-DNA complexes:
The rate of *Msp* I digestion was quantified in fade-out assays for the methylated patch region and the unmethylated non-patch region for construct d. The two substrates used were **A:** naked DNA construct, and **B:** histone H1-DNA complexes formed on the construct. The control assays were performed on mock-methylated constructs. The results show that there is no spreading of resistance from the methylated to the unmethylated region for these templates. The points that are plotted are single values from one experiment. Refer to Section 6.2.3 and Table 6.2 for further details.

6.3.2 Methylation does not reduce *Msp* I sensitivity of unmethylated regions of DNA or histone H1-DNA complexes

Experiments using naked patch-methylated constructs show that methylated regions are only very slightly more resistant to *Msp* I than the corresponding regions on the unmethylated control constructs and, in contrast to the results with chromatin (Section 6.3.3), there is no spreading of resistance from the methylated to the unmethylated region (Figure 6.7A and Table 6.2). Similar results are seen for patch-methylated constructs complexed to histone H1, although these templates are more resistant to *Msp* I digestion than are naked DNA templates (Figure 6.7B and Table 6.2). The graphs in Figure 6.7 are plotted with points that are the result of quantifications of single values, so there is some ambiguity in the gradient of the line of best fit. Nonetheless, it is clear that methylated regions are significantly less sensitive to *Msp* I digestion than unmethylated regions.

Table 6.2

Accessibility of patch or non-patch regions to *Msp* I, for naked mock-methylated and patch-methylated constructs, in the absence or presence of histone H1:

digestion units (U x min) required for 50% digestion:

	naked DNA		DNA + histone H1	
	patch	non-patch	patch	non-patch
unmethylated	5	4	17	19
methylated	9	4	39	19

6.3.3 Methylation affects the chromatin structure of unmethylated regions

Chromatin was reconstituted on patch-methylated constructs and subjected to limited digestions with either staphylococcal nuclease or *Msp* I (10 unit *Msp* I/μg DNA; see Section 2.7.7.3 for further details). The digestion products from these treatments were hybridised to restriction fragments corresponding to either the region that had been patch-methylated, or to the adjacent unmethylated region.

The results for staphylococcal digestion are shown in Figure 6.8. As expected, there is no difference in the nucleosome spacing between the chromatin at the methylated patch and the chromatin in the same region of the mock-methylated construct (see Section 4.2.1). There is also no apparent difference between the non-patch regions. The relative rate of digestion for these regions was not determined, as it was for the *Msp* I digestions, because the quantification of the diffuse bands that are produced in this assay would be inaccurate. Instead, this assay was used to ensure that chromatin had been reconstituted on construct DNA, before it was used in the *Msp* I digestion experiments.

The results for *Msp* I digestions are shown in Figures 6.9 and 6.10. Typical autoradiographs for chromatin, or chromatin in the presence of histone H1, reconstituted on a patch-methylated construct are shown in Figure 6.9. The assays for each substrate were performed in triplicate, and average values are shown in the graphs of Figure 6.10. (This is the same procedure as that adopted for the graphs of Figure 6.3; refer to Section 6.2.2 for further details). The results in Figure 6.9A show clearly that the chromatin reconstituted on either the methylated region or the unmethylated region of each patch-methylated construct is less accessible to *Msp* I digestion compared to that formed on the mock patch-methylated control construct (Table 6.3). Similar results are observed when chromatin reconstitution takes place in the presence of histone H1 (Figure 6.9B and Table 6.3), but the methylated construct is even more resistant to nuclease digestion of both the patched and the contiguous regions. The sample variance for the average value of each gradient in Table 6.3 was calculated as described in Section 6.2.2. The sample variance is expressed as a percentage value of the average gradient.

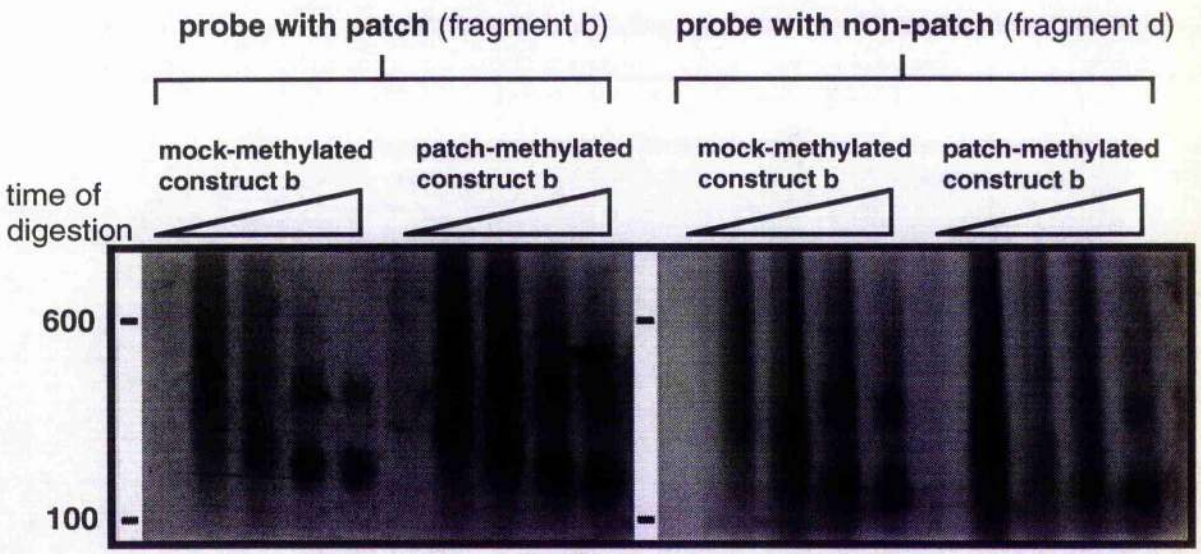


Figure 6.8

Staphylococcal nuclease digestion of chromatin reconstituted on mock- and patch-methylated constructs:
Chromatin was reconstituted on the patch-methylated construct b, and on the control mock-methylated construct. Aliquots from the reaction mixtures, each containing 100 ng DNA, were digested with 1.5 U staphylococcal nuclease (see Section 6.3.3) for 0, 2, 5, 15 and 30 min. as indicated by the wedges. Digestion products were transfered to a nylon membrane and probed with the fragment b (left-hand panel). The blot was then stripped of the probe, and re-probed with fragment d (right-hand panel). The calibration marks at the left of each panel indicate the positions of 100 bp and 600 bp marker fragments after agarose gel electrophoresis of the digestion products.

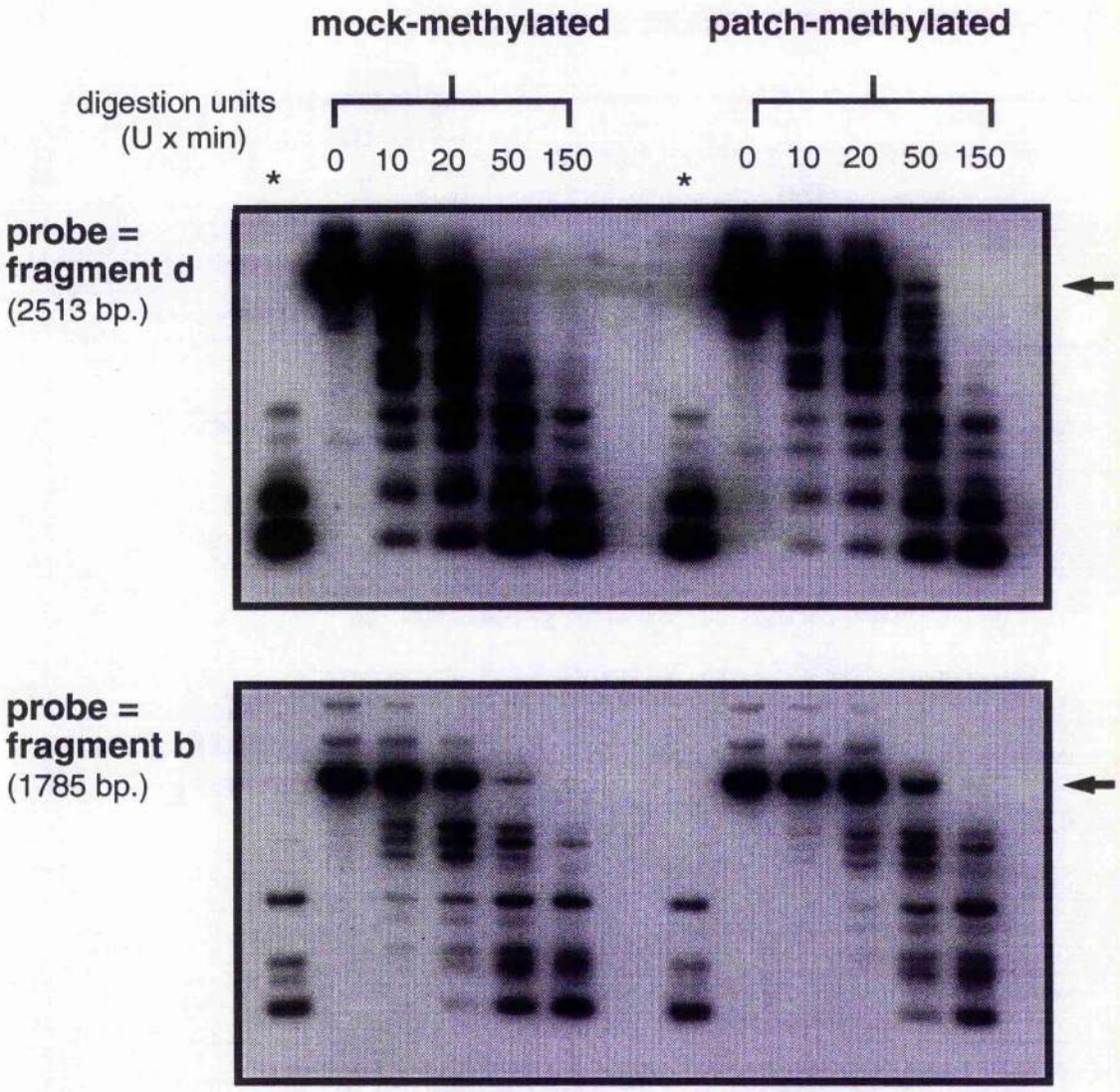


Figure 6.9 A

Msp I fade-out assay of patch-methylated construct d, reconstituted with chromatin:

The sensitivity of patch-methylated construct d to digestion with *Msp* I was used to assay the accessibility of DNA reconstituted with chromatin, in both the patch regions and the unmethylated non-patch regions. The mock-methylated construct was the control. The digestion products were hybridised to restriction fragments corresponding to either the region that had been patch-methylated (fragment d; top panel), or to the adjacent unmethylated region (fragment b; bottom panel). The arrows indicate the intact band that was used to quantify the rates of digstion (see Figure 6.9 A). The asterisk (*) indicates the complete digestion of the construct by *Msp* I. Refer to Section 6.3.3 for further details.

construct d + chromatin and histone H1:

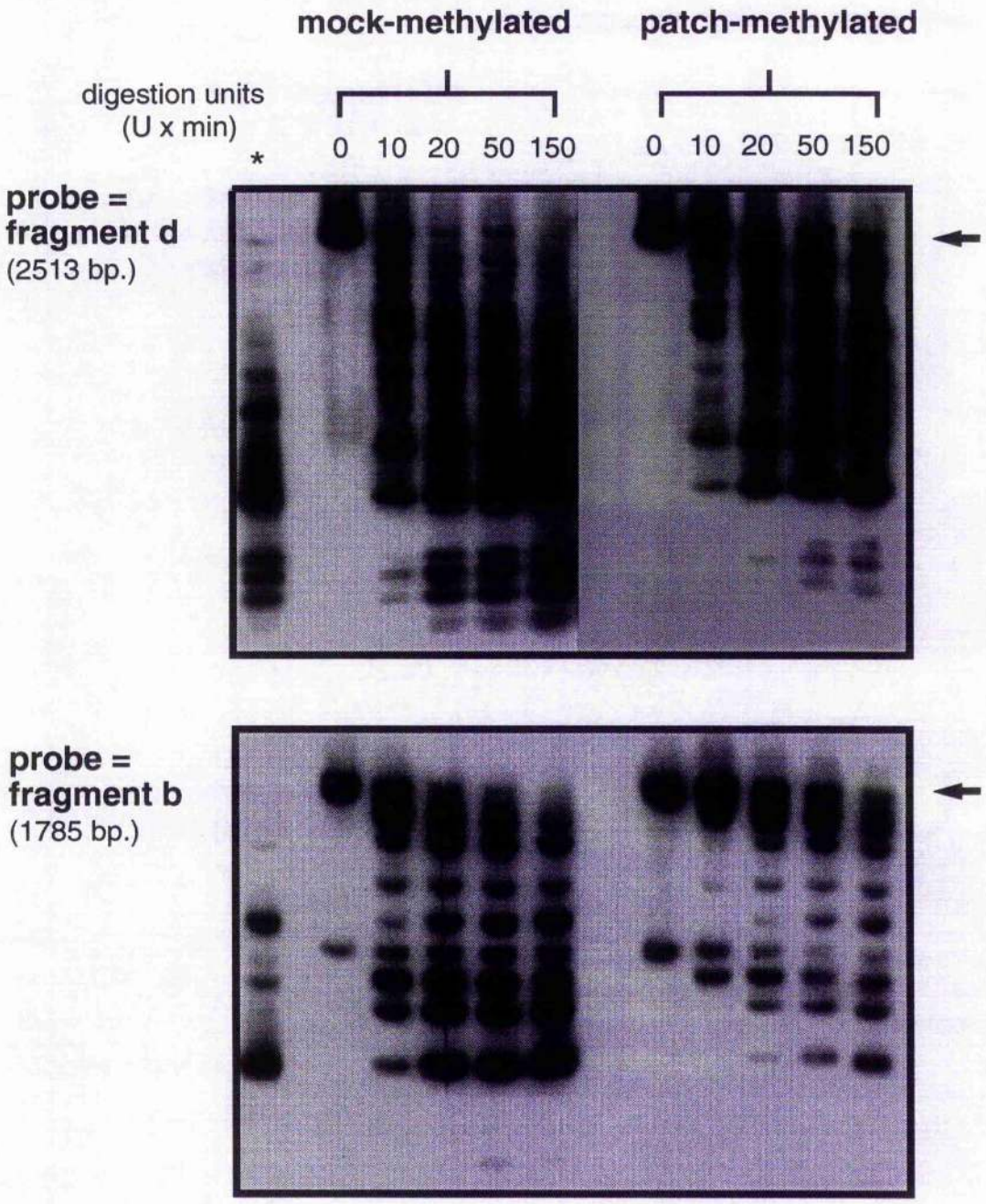


Figure 6.9 B

***Msp* I fade-out assay of patch-methylated construct d, reconstituted with chromatin and histone H1:**
The sensitivity of patch-methylated construct d to digestion with *Msp* I was used to assay the accessibility of DNA reconstituted with chromatin and histone H1, in both the patch regions and the unmethylated non-patch regions. Figure legend is identical to Figure 6.8 A.

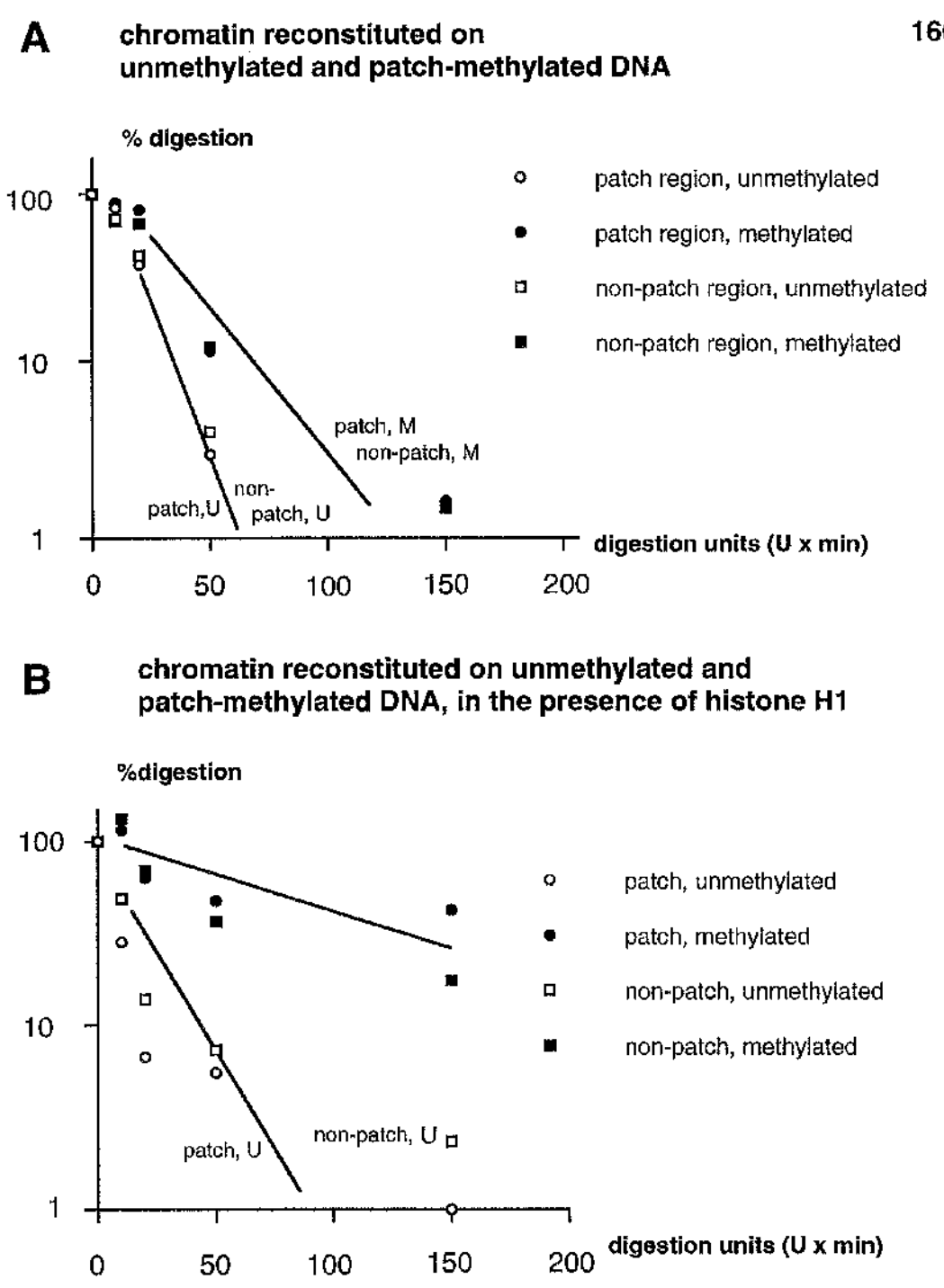


Figure 6.10

Rate of *Msp* I digestion of patch-methylated construct d reconstituted with chromatin: methylation affects the chromatin structure of unmethylated regions

The rate of *Msp* I digestion was quantified for the fade-out assays in Figure 6.8, for the methylated patch region and the unmethylated non-patch region of construct d. The two substrates that were used in these assays are **A:** chromatin, and **B:** chromatin and histone H1. The results are shown in semi-log plots, with points shown as an average of values in triplicate. The plots show that there is a spread of resistance from the methylated to the unmethylated region for these templates. Refer to Section 6.3.3 and Table 6.3 for further details.

Table 6.3

Accessibility of patch or non-patch regions to *Msp* I, for mock-methylated and patch-methylated constructs reconstituted with chromatin, in the absence or presence of histone H1:

digestion units (U x min) required for 50% digestion:

	chromatin		chromatin + histone H1	
sample deviation (%)	patch ±17 %	non-patch ±7 %	patch ±24 %	non-patch ±27 %
unmethylated	17	18	14	17
methylated	35	31	54	44

6.3.4 *In vitro* transcription of patch-methylated constructs

The levels of *in vitro* transcription from control mock-methylated or patch-methylated constructs were used to assay the activity of the SV40 promoter in four substrates: naked plasmid DNA, histone H1-DNA complexes, DNA reconstituted with chromatin and DNA reconstituted with chromatin in the presence of histone H1. Four constructs were used (constructs a–d), with different sizes and locations of regional methylation, as shown in Figure 6.1. In addition, mock-methylated and fully methylated pVHCk plasmid were also assayed at the same time, for each type of substrate, to act as controls. The transcription assays for all of these constructs, for each type of substrate, are shown in Figure 6.11. Only the product band of size 56–58 nt. is shown, but other primer extension products were formed that are not shown. Complete pictures of all the products are shown in Figure 3.6A and B, and the transcription assay is discussed in Section 3.3.4.

The primer extension products of Figure 6.11 were quantified by using a phosphoimager, analysed and presented in the series of graphs in Figure 6.12A–D for each one of the four types of substrate. Levels of transcription from patch-methylated constructs were compared with the average level from all mock-methylated constructs,

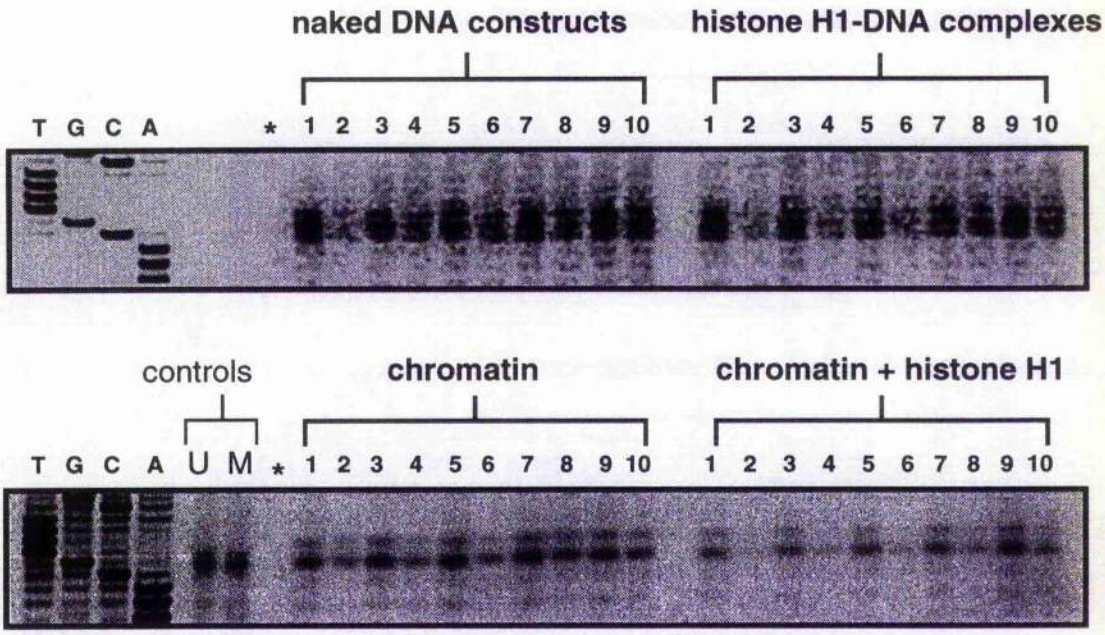


Figure 6.11

***In vitro* transcription of different types of mock-methylated, patch-methylated and fully methylated templates:**
Patch-methylated constructs a, b, c and d (see Section 6.3.1) and the control mock-methylated constructs were transcribed *in vitro* for the following substrates: naked DNA, histone H1-DNA complexes, chromatin reconstituted on construct DNA and chromatin reconstituted in the presence of histone H1. The primer extension products shown are 56-58 nt. in size, as determined by comparison with the sequencing tracts (T, G, C, A), and were the major products in this experiment; see Figure 3.7 and Section 3.3.4 for further details. For each substrate, odd-numbered lanes are assays of mock-methylated constructs and even-numbered lanes are for the corresponding patch-methylated construct: **1,2** construct a; **3,4** construct b; **5,6** construct c and **7,8** construct d. In addition, the activity of these substartes was assayed for mock-methylated (lane **9**) and fully methylated (lane **10**) pVHCk plasmid DNA as the template for transcription. Quantifications for lanes 9 and 10 are used in Figure 6.4 (see Section 6.2.3 for further details). The two control assays are for mock-methylated (U) and fully methylated (M) naked pVHCk, which allows the direct comparison of transcriptional activity between the assays in the top panel and those in the bottom panel. The asterisk (*) indicates transcription assays performed in the absence of template DNA. See Section 6.3.4 for further details and a discussion of this experiment, and Figure 6.12 for graphs that interpret the data from these experiments.

which did not vary significantly with size or location of mock-methylated patch. These results show that fully methylated naked DNA template is transcribed less efficiently than unmethylated template. This inhibition is most obvious when only the promoter of the reporter gene was methylated (Figure 6.12A). The degree of inhibition decreases with increasing patch size (compare transcription from construct a and c) and there is no inhibition of construct d. This is consistent with the HeLa nuclear extract containing limiting levels of an inhibitory factor (e.g. methyl CpG binding proteins). The density of this factor bound to the promoter of the otherwise naked plasmid could then dictate the activity of the promoter. Similar results are observed for constructs complexed with histone H1 (Figure 6.12B), although the absolute levels of transcription are only 80% of those with naked DNA constructs, as is consistent with the results in Section 6.2.3 and Figure 6.4. Therefore, with naked DNA in the absence or presence of histone H1, no spreading of inactivation is observed and inhibition is dependant on a high density of promoter methylation. However, in the presence of chromatin (Figure 6.12C), or chromatin with histone H1 (Figure 6.12D), all patch-methylated constructs are equally inhibited, including construct d, in which the patch of methylation is covers only vector sequences. In these cases a methylated patch in the vector DNA is able to seed the formation of inactive chromatin that leads to the inhibition of transcription from the distant promoter. This is reflected in the *Msp* I resistance of the non-patch regions of chromatin-associated plasmids (see Section 6.3.3).

Transcription of all the patch-methylated constructs (constructs a-d) reconstituted with chromatin, in the absence or presence of histone H1, is inhibited to a common level that is 40-50% of the unmethylated constructs (Figure 6.12C and D). Transcription from naked DNA, or histone H1-DNA complexes, is at various levels depending on the size and position of the methylated region. The methylated CpG-rich prokaryotic vector DNA in construct d does not inhibit transcription from the distant promoter for these templates (Table 6.2), as discussed above. Transcription from construct c, in which the patch covers both promoter and reporter gene, is inhibited to 60-80% of the unmethylated constructs which is consistent with the results for fully-methylated templates (Figure 6.12A). However, transcription from construct a, in which only the promoter is methylated, is inhibited still further to about 40%. This result could be explained if a

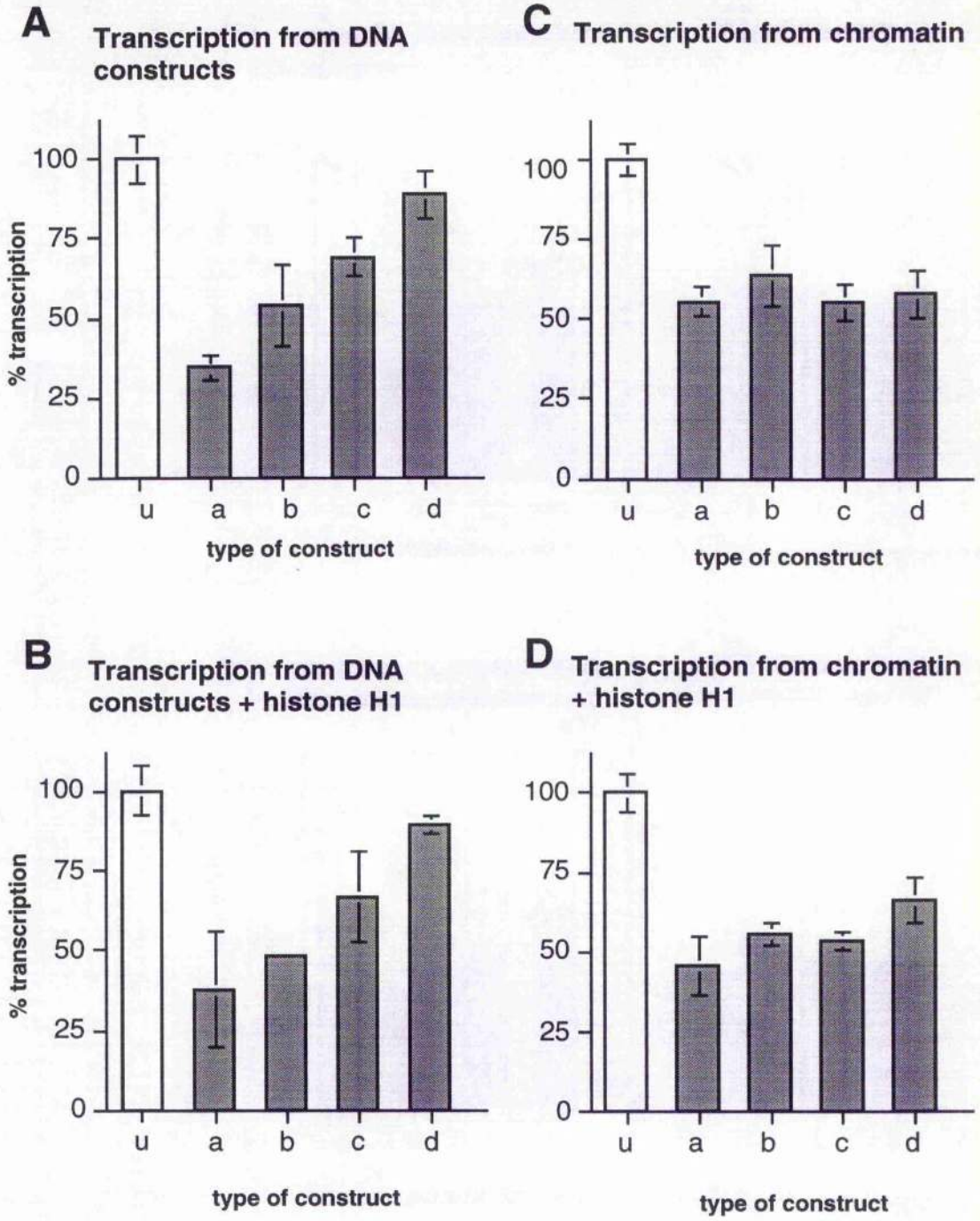


Figure 6.12

Comparison of *in vitro* transcription from different types of mock-methylated and patch-methylated templates:
The graphs interpret the results of Figure 6.11. The effect of different size and location of patches of methylation on *in vitro* transcription is shown for the substrates: naked DNA, histone H1-DNA complexes, DNA reconstituted with chromatin and DNA reconstituted with chromatin in the presence of histone H1. The type of patch-methylated construct that was used for each experiment is indicated (letters a–d; filled boxes), with each error bar showing the sample range for results in duplicate. The letter u (open boxes) indicates the average value for all mock-methylated constructs for one particular type of substrate. The value of each box labelled u is therefore an average of eight results, which is taken as the 100% standard value.

large patch (c, containing 50 methylated CpGs) is able to compete out limiting amounts of a selective inhibitor (such as MeCP1) that is present in the nuclear extract, whereas a small patch (a, containing 9 methylated CpGs) is unable to compete out the inhibitor. Construct b is also inhibited, even though the promoter is unmethylated, which suggests that the putative MeCP can mediate an indirect mechanism of promoter inhibition even in the absence of chromatin. Kass et al. (1993) have shown that a patch at the promoter (as in construct a) can reduce the transcriptional activity of the methylated construct to 69% of the mock-methylated control, after transient expression of the constructs.

6.4 Discussion

Summary

This chapter investigates the spread of inactive chromatin from a region of methylation by using *in vitro* systems of chromatin assembly and transcription. The presence of inactive chromatin was assayed by resistance to digestion by *Msp* I. However, fully methylated naked plasmid DNA is also less sensitive to digestion with *Msp* I, compared to naked unmethylated DNA. This reduced sensitivity is confined to the region of methylation in a construct and does not affect contiguous regions, even when histone H1 is complexed to the DNA. Using HeLa nuclear extract, fully methylated naked DNA template is transcribed less efficiently than is unmethylated template. This effect is most obvious when only the promoter of the reporter gene is methylated, a finding that is indicative of limiting levels of an inhibitor in the nuclear extract. Histone H1-DNA complexes on unmethylated and fully-methylated pVHCk were also used as templates, and the results support the previous observation that histone H1 preferentially represses *in vitro* transcription from fully-methylated template DNA (Johnson et al. 1995; and Section 5.3.4). When chromatin was reconstituted on these templates using *Xenopus* S150 egg extract, the lower sensitivity to *Msp*I digestion of the methylated DNA is maintained and this reduced sensitivity is now able to spread from the methylated patch to a contiguous unmethylated region. Methylated chromatin templates retain a reduced transcriptional activity compared with unmethylated chromatin templates, and this is also the case when

any region of the plasmid is unmethylated. These effects are further enhanced by the addition of histone H1 during the chromatin reconstitution.

These results indicate that on naked DNA templates, in the absence or presence of histone H1, promoter methylation is important to inhibit transcriptional activity and only regions of the plasmid that are actually methylated show reduced sensitivity to nuclease digestion. However, with chromatin templates patch-methylation affects transcription and nuclease sensitivity of a distant promoter.

Msp I sensitivity of naked DNA and histone H1-DNA complexes

The accessibility of nucleases to chromatin has been used as an assay of chromatin conformation (Paranjape et al. 1994) and, in particular, *Msp* I has been used to assay differences between unmethylated and methylated chromatin (Keshet et al. 1986; Antequera et al. 1989). The results discussed in Section 6.2.2 show that *in vitro* chromatin reconstitution onto methylated DNA is also more resistant to digestion by *Msp*I than unmethylated chromatin (Figure 6.3), using a fade-out assay for intact restriction fragments as described above. Histone H1-DNA complexes on methylated DNA are also more resistant than unmethylated DNA (Figure 6.3 and Table 6.1), and are results that confirm a previous study by Higurashi and Cole (1991). This study also stated that there was no detectable difference in *Msp* I sensitivity between naked unmethylated and methylated DNA, after simple agarose gel electrophoesis of the restriction digests.

In the present study a difference between naked unmethylated and methylated DNA was observed (Figure 6.3 and Table 6.1). This was achieved after the precise quantification of the extent of hybridisation of a radioactive probe to a Southern blot of pVHCk digested with *Msp* I. A similar sensitive technique has been used to show differences in accessibility of naked methylated and unmethylated DNA to DNase I (Kochanek et al. 1993). This effect of methylation may be caused both by local distortions in the DNA structure and by the steric effects of a methyl group in the major groove of the DNA (Hodges-Garcia and Hagerman 1992; Kochanek et al. 1993), which would both modify direct protein-DNA interactions. An earlier study by Barr et al. (1985) has shown that methylated DNA has an intrinsic resistance to digestion by

staphylococcal nuclease, both in native chromatin from human fibroblasts and in naked DNA that was purified from the nuclei of these cells. The authors of this study suggested that the intrinsic nuclease resistance of methylated DNA could be of sufficient magnitude to explain the nuclease resistance of chromatin that contained methylated DNA. This point is supported by the results in Section 6.2.2 and Table 6.1. The rate of digestion of chromatin is considerably reduced relative to that of naked DNA, but the substrate containing methylated DNA is less sensitive to *Msp* I digestion than is that containing mock-methylated DNA. However, the differential sensitivity is approximately two-fold irrespective of the type of substrate. The results of Figure 6.8 would also be expected to show that chromatin reconstituted on methylated DNA has an intrinsic resistance to digestion by staphylococcal nuclease, but no such difference is evident. However, the bands in this assay were not quantified (see Section 6.3.3), so a difference could exist. The original paper by Barr et al. (1985) used labelling with ^{14}C and HPLC fractionation of staphylococcal nuclease fragments as a very sensitive assay of 5-mC distribution.

A local distortion of DNA structure in a methylated region could propagate as a change in the topology of the plasmid to adjacent unmethylated regions. However, there is no evidence for such a spread in patch-methylated constructs (Figure 6.7A) since the methylated patch and the unmethylated non-patch have differences in accessibility (Table 6.2) which resemble the values for fully-methylated and fully-unmethylated naked DNA (Table 6.1). Similar results are observed for histone H1-DNA complexes on patch-methylated constructs (Figure 6.7B and Table 6.2), implying that the increase in resistance of complexes on methylated DNA is mediated by the structure and the interaction between histone H1 and the DNA (Higurashi and Cole 1991), rather than a cooperative binding of histone H1 that could spread into adjacent unmethylated regions. These conclusions are drawn from a single set of data points (in the graphs of Figure 6.7), so the gradients of the lines of best fit are ambiguous. Although the gradients have been quantified and tabulated in Table 6.2, it is safer to assume that a qualitative difference does exist between DNA in the methylated patch and DNA in unmethylated regions, but that the limited data does not allow the difference to be quantified.

Evidence for the spread of inactive chromatin

Several studies that precede the one by Kass et al. (1993) show that the *in vitro* methylation of sequences that are distant to an unmethylated promoter can inhibit gene expression. *In vitro* methylation of sequences adjacent to an unmethylated promoter could inhibit expression of the human β-globin gene (Yisraeli et al. 1988) and the herpes simplex virus (HSV) thymidine kinase (*tk*) gene (Graessmann and Graessmann 1993), after transfection or microinjection of the DNA into mammalian cells. The study by Graessmann and Graessmann (1993) showed that methylation of the coding region of the *tk* gene caused inactivation. In addition, the methylation of single *Hpa*II sites, of the several located from 122 to 571 bp downstream of the promoter, was sufficient for *tk* inactivation. Methylation of one such site within the promoter did not bring about any inactivation. It has also been reported that the methylation of flanking sequences does not inhibit the expression of the hamster *aprt* gene following transfection into mammalian cells, but that methylation of the 3' structural region of the HSV *tk* gene does inhibit expression (Keshet et al. 1985). These conflicting observations may be explained by the non-physiological pattern of hemimethylation imposed on the reporter gene constructs, since *in vitro* methylation used restriction fragment primer-directed second-strand synthesis with 2'-deoxy-5-methylcytidine triphosphate. The protocol for regional methylation described in this study overcomes this problem by using the prokaryotic M.*Sss* I methyltransferase. Levine et al. (1992) have used several methyltransferases, both separately and in combination, to methylate constructs that had small inserts within the preinitiation domain of the SV40 promoter. This resulted in transcriptional inhibition, after the constructs were transfected into mouse L cells. The methylation of inserts further upstream or downstream of the TATA box with individual methyltransferases caused no inhibition, but methylation with a combination of the enzymes was able to mediate a weak inhibition.

The spread of inactive chromatin, in the absence of DNA methylation, has been proposed to underlie the phenomenon of position-effect variegation in *Drosophila* and yeast (reviewed in Karpen 1994). Heterochromatin in these systems does not contain genes that can be detected by conventional genetic assays of mutation and recombination (Pardue and Hennig 1990), but it does have other significant genetic effects. Position-effect variegation describes the repression of genes that are transferred into, or adjacent

to, heterochromatic regions by chromosome rearrangement. Cytological observations have shown that euchromatin juxtaposed with heterochromatin can itself become heterochromatic. Several models, such as the chromatin assembly model, have suggested that the spread of repressed gene activity is driven by the diffusion of a repressor protein, or multimeric complex of proteins, from their normal location in heterochromatin (Reuter and Spierer 1992). The gene for HP1 (heterochromatin protein 1) has been identified in *Drosophila* because mutations in the *HP1* gene suppress position effects on gene expression (Elgin 1990). The HP1 protein tends to associate with the heterochromatin regions of polytene chromosomes and does not appear to bind to DNA. The so-called "chromo domain" in HP1, and the homologous protein Polycomb (Pc; reviewed in Paro 1993), has been proposed to enhance the protein-protein interactions that are required for higher-order chromatin structures. However, it is still unclear if the chromo domain functions to compact chromatin, with gene repression as a direct consequence of this interaction. In a similar fashion, repressor proteins could mediate the spread of inactive chromatin from regions of methylation in the mammalian cell systems described above, or in the chromatin reconstituted with *Xenopus* egg extracts, as demonstrated by my own results. Candidate proteins include MDBPs and MeCPs (see Section 1.7.2), but the mechanism by which these proteins repress gene expression in the context of chromatin remains unknown.

This study has extended the observations of Kass et al. (1993) for transfected cells to an *in vitro* system. This approach cannot yet yield information on the proteins other than histones that are required for methylation-mediated chromatin inactivation, since these components in *Xenopus* egg extracts remain unidentified and unpurified. An *in vitro* system using purified components may yield less ambiguous results. The significance for gene expression *in vivo* of inactive chromatin spreading from methylated regions may be connected with the *de novo* methylation of CpG sites during early embryogenesis. Methylation could mediate long-range organisation of chromatin and a global regulation of gene expression within nuclear domains. However, a housekeeping gene would remain active because an adjacent CpG island could act as an insulator to the spread of inactive chromatin, because CpG islands are unmethylated throughout development.

CHAPTER SEVEN

Discussion

The original aim of the experiments described in previous sections was to investigate the inhibition by methylation of gene expression. The first series of experiments, described in Chapter 5, investigated the *in vitro* transcription of tRNA genes, since it was not known whether the expression of these genes was inhibited by methylation (see Section 3.2.1). The transcription assays for this system used a HeLa nuclear extract to provide the basal transcription factors for transcription by pol III. The first investigation of this system showed that transcription from a methylated template was consistently lower than from an unmethylated template, and that the degree of preferential inhibition was even greater at low levels of methylated template (Section 5.3.1). This observation was entirely consistent with a proposal by Boyes and Bird (1991) that MeCP-1 could mediate the preferential inhibition by methylation of pol II transcription. The surprising observation was then made that both unmethylated and methylated competitor DNA could remove the putative inhibitor or inhibitors from the nuclear extract used in these transcription studies (Section 5.3.2.1). The conclusion is clear: preferential inhibition of a methylated template could not be mediated just by a protein that binds specifically to methylated DNA, although a role for such a protein is not ruled out.

Several lines of evidence pointed to histone H1 as an inhibitor of transcription (discussed in Section 5.4). Two studies of particular relevance have been published recently (Jost and Hofsteenge 1992; Levine et al. 1993). Jost and Hofsteenge (1992) showed that the methylated DNA binding protein MDBP-2 has certain sequence homologies with histone H1. Levine et al. (1993) showed that histone H1 can preferentially inhibit *in vitro* pol II transcription of methylated templates, which is supported by my own study of histone H1 and pol III templates (Johnson et al. 1995) The crucial observation in my study is that *in vitro* transcription from defined histone H1-DNA complexes (Section 5.2) is sensitive to the methylation status of the DNA

(Section 5.3.4). Histone H1 appears to prevent the formation of initiation and elongation transcription complexes preferentially on methylated templates (Section 5.3.6).

MDBP-2 has characteristics similar to histone H1. It preferentially inhibits transcription from methylated DNA, but in addition it appears to bind preferentially to methylated DNA, both *in vitro* and *in vivo* . If histone H1 does indeed have an MDBP-2-like activity, then studies of the behaviour of histone H1 in the absence of chromatin may be valid, or at least complementary to those of histone H1 and chromatin. However, a gel-retardation experiment of histone H1 complexed to unmethylated or methylated DNA (Section 5.2.1) does not demonstrate that histone H1 has any greater affinity to methylated DNA, which is an observation supported by previous studies (Higurashi and Cole 1991; Nightingale and Wolffe 1995). In contrast, Levine et al. (1993) have shown that histone H1 has a two-fold greater binding affinity to methylated DNA. Although MDBP-2 and histone H1 share sequence homologies, it remains unclear if these two proteins are similar in function. Histone H1-DNA complexes that are formed on methylated DNA are resistant to digestion by *Msp* I, compared with complexes formed on unmethylated DNA (Higurashi and Cole 1991; also refer to Section 6.2.2). The resistance of methylated CCGG sites does not appear to arise by a greater binding affinity of histone H1 to methylated DNA. Instead, the resistance could be due to either a steric effect of methyl groups at CCGG sequences or because the conformation of the methylated DNA changes at these sites.

In summary, the current opinion is that histone H1 has certain characteristics of a methylated-DNA binding protein, which is reasonable since it tends to be located in quiescent, methylated regions of the mammalian genome. However, the current literature is undecided over whether histone H1 binds preferentially to methylated DNA. Inactive chromatin could be formed in the following scenario: histone H1 binds to methylated linker DNA which causes inactivation of the gene, firstly by the MDBP-like repressor activity of histone H1 and, secondly, by the formation of condensed supranucleosomal structures.

The assumption that underlies the studies described in Chapter 5 is that histone H1-DNA complexes are a valid model of inactive chromatin. This assumption is discussed in detail elsewhere (Croston et al. 1991; also refer to Section 5.4) but, suffice

it to say, that the model has obvious weaknesses. The exact nature of the interaction between a histone H1 molecule and the dyad of DNA that wraps around the octamer is not known (Ramakrishnan 1994; Pruss et al. 1995). In particular, it is not known how closely the globular domain of histone H1 interacts with the octamer (discussed in Section 5.4). One model of the chromatosome suggests that the tails of histone H1 interact with linker DNA at some distance from the nucleosome, with minimal contacts between the globular domain and octamer (Ramakrishnan 1994). A second model, based on cross-linking and neutron scattering data (Boulikas et al. 1980; Graziano et al. 1994), suggests that there are additional interactions of the globular domain with the nucleosome. The first model supports the assumption that histone H1 locks the DNA into an inactive, inaccessible structure and that nucleosomes are just structural components to keep the DNA neat and safe. Recent evidence refutes this: nucleosomes are not just passive packages of nucleoprotein but have active roles to play in gene activation. In particular, nucleosomal arrays can be disrupted by a variety of transcriptional activators (Carlson and Laurent 1994; Paranjape et al. 1994) in a process called anti-repression (Section 1.8.1), and hyperacetylation of core histones is presumed to reduce the interaction of the DNA with the octamer (Section 1.8.6). In view of this, it is not safe to draw conclusions about the behaviour of chromatin from the *in vitro* system of histone H1-DNA complexes. For example, different somatic variants of histone H1 preferentially inhibit the *in vitro* transcription from a methylated template to different extents (Sections 5.3.5 and 5.4). However, it is impossible to say if these differences reflect defined physiological roles for the different variants during development and cell differentiation (refer to Section 1.8.7).

A complete understanding of gene expression can only be achieved by studying transcription in the context of chromatin. This is a difficult area of study, with technical limitations on the experiments that are possible. A necessary simplification has been the *in vitro* transcription of naked DNA templates, which is an approach that has identified many of the components of transcription complexes (discussed in Section 3.1). However, DNA *in vivo* is almost certainly never naked, and is bound at all times to a variety of proteins such as histones, non-histone proteins, transcription factors and high mobility group (HMG) proteins (Wolffe 1992).

The logical progression for this study was to investigate the inhibition of transcription by DNA methylation in the context of a substrate that resembles native chromatin. The system that was chosen assembled chromatin on double-stranded DNA (Section 4.2) using *Xenopus* egg extracts as a source of core histones, nucleosome assembly factors and topoisomerases (Dimitrov and Wolffe 1995). The assembly was straight-forward, and allows histone H1 to be incorporated into the chromatin at physiological levels (Section 4.2.2). The average nucleosome spacing of the assembled chromatin was determined by digestion with staphylococcal nuclease and found to be 180 bp. This spacing is the same as that for native *Xenopus* chromatin (Almouzni and Wolffe 1993). In this assay there is no discernible difference between the chromatin assembled on unmethylated and methylated DNA, confirming the results of Buschausen et al. (1987). A separate study by Englander et al. (1993) reports that the *in vitro* assembly of histone octamers onto DNA was unaffected by methylation (discussed in Section 1.8.4).The addition of histone H1 to the chromatin reconstitution reaction increases the average nucleosome spacing of both unmethylated and methylated chromatin from 180 to 210 bp (Section 4.2.2). Both types of assembled chromatin appear to be equally sensitive to staphylococcal nuclease digestion.

It was therefore an unexpected observation that the unmethylated or methylated chromatin was differentially sensitive to digestion by *Msp* I (Section 6.2.2). Methylated DNA in nuclei, or in the chromatin structures formed when methylated constructs are transiently transfected into mammalian cells, is known to be preferentially resistant to *Msp* I (Section 1.8.2). The experiments in Chapter 6 extend this observation to an *in vitro* system. The components of the egg extract that are responsible for the observed differences in *Msp* I sensitivity of methylated and unmethylated DNA are unknown. It is reasonable, however, to assume that interactions between ubiquitous nuclear proteins, such as histones and HMG proteins and proteins that bind specifically to methylated DNA are responsible. This is discussed below in more detail. Buschausen et al. (1987) demonstrated that chromatin formation was required for transcriptional inactivation of a methylated *tk* gene, and there is considerable circumstantial evidence for a role for MeCPs in the formation of inactive chromatin (Section 1.7.2). The intrinsic resistance of methylated DNA to digestion by *Msp*I (Section 6.2.2) may also contribute to the

observed difference between methylated and unmethylated chromatin. Barr et al. (1985) originally proposed that the nuclease resistance of naked methylated DNA could be of a sufficient magnitude to explain the nuclease resistance of chromatin that contained methylated DNA.

The most surprising observation is that unmethylated regions in regionally methylated constructs of plasmid pVHCk are resistant to *Msp* I compared to mock-methylated constructs, when these constructs are reconstituted with chromatin (Section 6.3.3). This indicates that methylation in one region can influence the nuclease sensitivity in an unmethylated region on the same plasmid. These results support the study by Kass et al. (1993), which reports the same finding for patch-methylated constructs of plasmid pVHCk transiently transfected into L cells. In addition, the same study demonstrated that methylated regions of between 349 to 2513 bp in size could repress transcription from an SV40 promoter at least 1 kbp away, with the degree of inactivation dependant on the size of the patch. The spreading of an inactive chromatin structure was previously observed for integrated sequences, but it was concluded that the range of spreading is limited to a few hundred nucleotides (Keshet et al. 1986). A similar conclusion was reached for methylated constructs that were transiently transfected into L cells (Levine et al. 1992), but in contrast to the study by Kass et al. (1993) the region of methylation was limited to 38 bp in size and contained 9 methylated sites. Histone H1-DNA complexes also appear to mediate a spread of inactivation from methylated sequences over a few hundred nucleotides (Section 5.3.7), but these experiments should be repeated to obtain a conclusive result. Cooperative interactions between histone H1 molecules could explain the spread of inactivation. Cooperative interactions between the globular domains occur at low salt concentrations and low input histone H1:DNA ratios (5 mM NaCl and 0.1–0.6 w/w; Thomas et al. 1992), which are similar to the conditions described for the formation of histone H1-DNA complexes in this thesis (Section 2.7.5.2). In contrast, chromatin appears to mediate an activation that can travel for longer distances: the results in Section 6.3.3 suggest that this distance is at least 1 kbp. The data from *Msp* I fade-out assays correlates with that from *in vitro* transcription studies of patch-methylated constructs of plasmid pVHCk (Section 6.3.4), which assay the initiation of transcription from the SV40 promoter. Transcriptionally inactive chromatin forms on methylated

DNA, but can also form on unmethylated DNA if adjacent regions are methylated. There is no evidence that methylation can prevent transcription from the SV40 promoter through a direct mechanism (Bryans et al. 1992; refer to Section 1.7.1), so the inhibition of transcription is due to the formation of chromatin.

This study has not addressed the obvious questions of how far the inactivation can spread and if this distance is dependant on the size of the patch. For example, the ability of the small promoter patch in construct a (refer to Figure 6.1) to inactivate chromatin at different regions of the vector (fragment d in Figure 6.1) was not studied. If suitable small fragments of DNA from the vector region were used as probes then the degree of inactivation could be assayed beside the patch at region a, and then at progressively further distances away from the patch. Kass et al. (1993) have shown that a patch at the promoter (as in construct a) can reduce the transcriptional activity of the methylated construct to 69% of the mock-methylated control, after transient expression of the constructs. Studies such as these could determine if there is a function between distance from the methylated patch, the size of the patch and the degree of inactivation. Alternatively, the size of the patch could stay the same, but progressively longer stretches of "buffer" DNA could be inserted in between the patch and the region that is to be probed. Bird (1992) suggests a model whereby promoter activity is dependent on the proximity of methyl-CpGs to the promoter, the strength of the promoter and the density of methyl-CpGs. The CpG density of the patch-methylated constructs is abnormally high and this might be able to override the distance effect normally associated with methylation-mediated gene inactivation.

Another question that could be addressed is if the spread can be affected by the nature of the intervening sequence. The vector sequences of plasmid pVHCk are rich in CpG dinucleotides, compared to the reporter gene sequences (Figure 6.1) and, in particular, the SV40 terminator region. In other words, 80% of the sequence of pVHCk is CpG-rich and could therefore represent a rather large CpG island of 4 kbp in size. It might be argued that methylation of the vector sequences (in, for example, construct d) does not reflect the *in vivo* situation, because CpG islands are almost invariably unmethylated at all times of development (refer to Section 1.6). The exception is the methylation of islands on the inactive X chromosome (reviewed in Riggs and Pfeifer

1992; Rastan 1994), which is thought to play a role in transcriptional inactivation. It has been demonstrated that DNA methylation follows X-chromosome inactivation (Lock et al. 1987), indicating that methylation is not a prerequisite for the formation of inactive chromatin. Methylation could support the formation of inactive chromatin, presumably through the specific binding of proteins to methylcytosine. In addition, methylation could maintain this inactive state, which then can spread to adjacent regions. Alternatively, a CpG island could act as an insulator to the spread of inactive chromatin, because islands are unmethylated in autosomes. This would ensure that inactive chromatin could not form in close proximity to a housekeeping gene. This hypothesis implies that the formation of inactive chromatin is the default pathway during the maturation of chromatin after replication. Repression would therefore be the default state for gene expression, which could only be overcome by "anti-repression" with transcriptional activator proteins. This is reasonable since large proportions of the chromosome contain repetitive DNA sequences that become associated with constitutive heterochromatin (Pardue and Hennig 1990; also refer to Section 1.8.1). CpG islands could be distinguished as targets for transcriptional activator proteins from the rest of the ocean if the chromatin at islands could not become inactive and closed.

An analysis of the protein composition of unmethylated and methylated chromatin has not been performed for this study. In preliminary experiments the assembled chromatin could not be isolated on sucrose gradients from the other protein components of *Xenopus* egg extract. The use of sucrose gradients to purify chromatin has been reviewed in (Noll and Noll 1989). The nuclear proteins that mediate the formation of inactive chromatin on methylated DNA, and those that mediate the spread of inactivation remain unknown. Proteins that bind to methylated DNA or methylcytosine (reviewed by Ehrlich and Ehrlich 1993; Tate and Bird 1993) are ideal candidates to mediate the effects of methylation on chromatin structure and transcription (Section 1.7.2). Both MeCP1 and MeCP2 have been reported to preferentially inhibit *in vitro* transcription from methylated pol II genes (Boyes and Bird 1991; Lin et al. 1995), and the presence of such a factor in HeLa nuclear extracts is required to explain the consistent preferential inhibition of methylated genes in both pArg/Leu and pVHCk. The level of transcription from fully methylated templates is about 80% of that from unmethylated templates, for

both the pol II and the pol III system. However, if the density of methylCpGs is high and localised in a small region (fragment a, the promoter region of pVHCk; see Figure 6.1) then transcription is inhibited even further (Section 6.3.4 and Figure 6.12A). This clearly indicates that limiting amounts of the inhibitor are present in the volume of nuclear extract that is used for the pol II transcription assays, but the identity of the inhibitor cannot be determined.

Although *Xenopus* egg extract does not contain histone H1, it does contain the B4 protein which appears to be the embryonic homologue (Dimitrov et al. 1993). This homologue may be a component of chromatin assembled *in vitro* , and could mediate the spreading of inactive chromatin by the formation of supranucleosomal structures, even in the absence of histone H1 (Section 6.3.3 and Figure 6.9A). Exogenous histone H1 that is added during chromatin reconstitution may enhance the spread of inactivation. High mobility group (HMG) proteins have been proposed to have an architectural role in assembling supranucleosomal chromatin structures (reviewed in Grosschedl et al. 1994). For example, HMG1 and 2 can facilitate nucleosome assembly (Bonne-Andrea et al. 1984) and HMG-D, a *Drosophila* homologue of the vertebrate HMG1 protein, can condense chromatin structures during embryogenesis (Ner et al. 1993). These properties are also characteristic of histone H1 but, in addition, the proteins of the HMG 1/2 group from vertebrates can induce and bind sharp bends in DNA.

In conclusion, this thesis has been an attempt to understand the molecular mechanisms that underlie the repression of gene activity by a combination of DNA methylation and chromatin. *In vitro* systems for transcription and chromatin assembly have been used to model the correct *in vivo* regulation of genes in a natural chromosomal context.

REFERENCES

Adams, C.C. and J.L. Workman (1993). "Nucleosome displacement in transcription." Cell **72**: 305-308.

Adams, R.L.P. (1990a). *Cell culture for biochemists*. Amsterdam, Elsevier Science Publishers.

Adams, R.L.P. (1990b). "DNA methylation: the effect of minor bases on DNA-protein interactions." Biochem. J. **265**: 309-320.

Adams, R.L.P. and R.H. Burdon (1985). *Molecular Biology of DNA Methylation*. New York, Springer Verlag.

Adams, R.L.P., H. Lindsay, A. Reale, C. Seivwright, S. Kass, et al. (1993). "Regulation of de novo methylation." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 120-144.

Allan, J., G.J. Cowling, N. Harborne, P. Cattim, R. Craigie, et al. (1981). "Regulation of the higher-order structure of chromatin by histones H1 and H5." J. Cell Biol. **90**: 279-288.

Allan, J., N. Harborne, D.C. Rau and H. Gould (1982). "Participation of core histone 'tails' in the stabilisation of the chromatin solenoid." J. Cell. Biol. **93**: 285-297.

Allegra, P., R. Sterner, D.F. Clayton and V.G. Allfrey (1987). "Affinity chromatographic purification of nucleosomes containing transcriptionally active DNA sequences." J. Mol. Biol. **196**: 379-388.

Almouzni, G., D.J. Clark, M. Méchali and A.P. Wolffe (1990). "Chromatin assembly on replicating DNA in vitro." Nucleic Acids Res. **18**: 5767-5774.

Almouzni, G. and M. Méchali (1988). " Assembly of spaced chromatin promoter by DNA synthesis in extracts from Xenopus eggs." EMBO J. **7**: 665-672.

Almouzni, G. and A.P. Wolffe (1993). "Nuclear assembly, structure, and function: the use of *Xenopus* in vitro systems." Exp. Cell Res. **205**: 1-15.

Antequera, F. and A. Bird (1993). "CpG islands." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 169-187.

Antequera, F., J. Boyes and A. Bird (1990). "High level of de novo methylation and altered chromatin structure at CpG islands in cell lines." Cell **59**: 503-514.

Antequera, F., D. Macleod and A.P. Bird (1989). "Specific protection of methylated CpGs in mammalian nuclei." Cell **58**: 509-517.

Ariel, M., E. Robinson, J.R. McCarrey and H. Cedar (1995). "Gamete-specific methylation correlates with imprinting of the murine *Xist* gene." Nature Genet. **9**: 312-315.

Ball, D.J., D.S. Gross and W.T. Garrard (1983). "5-Methylcytosine is localised in nucleosomes that contain histone H1." Proc. Natl. Acad. Sci. USA **80**: 5490-5494.

Baralle, F.E. and G.G. Brownlee (1978). "AUG is the only recognisable signal sequence in the 5' noncoding regions of eukaryotic mRNA." Nature **274**: 84-87.

Barker, D., M. Schafer and R. White (1984). "Restriction sites containing CpG show a higher frequency of polmorphism in human DNA." Cell **36**: 131-138.

Barlow, D.P. (1993). "Methylation and imprinting: from host defence to gene regulation?" Science **260**: 309-310

Barr, F.G., M.B. Kastan and M.W. Lieberman (1985). "Disribution of 5-methyldeoxy-cytidine in products of staphylococcal nuclease digestion of nuclei and purified DNA." Biochem. **24**: 1424-1428.

Barras, F. and M.G. Marinus (1989). "The great GATC: DNA methylation in *E. coli*." Trends Genet. **5**: 139-143.

Barrera-Saldana, H., K. Takahashi, M. Vigneron, A. Wileman, I. Davidson, et al. (1986). "All six GC-motifs of the SV40 early upstream element contribute to promoter activity in vivo and in vitro." EMBO J. **5**: 3839-3849.

Bartolomei, M.S., A.L. Webber, M.E. Brunkow and S.M. Tilghman (1993). "Epigenetic mechanisms underlying the imprinting of the mouse H19 gene." Genes Devel. **7**: 1663-1673.

Becker, P.B., S. Ruppert and G. Schutz (1987). "Genomic footprinting reveals cell type-specific DNA binding of ubiquitious factors." Cell **51**: 435-443.

Bednarik, D.P., C. Duckett, S.U. Kim, V.L. Perez, K. Griffis, et al. (1991). "DNA CpG methylation inhibits binding of NF-κ B proteins to the HIV-1 long terminal repeat cognate DNA motifs." New Biol. **3**: 969-976.

Behn-Krappa, A., I. Hölker, U. Sandaradura de Silva and W. Doerfler (1991). "Patterns of DNA methylation are indistinguishable in different individuals over a wide range of human DNA sequences." Genomics **11**: 1-7.

Belanger, F.C. and A.G. Hepburn (1990). "The evolution of CpNpG methylation in plants." J. Mol. Evol. **30**: 26-35.

Ben Hattar, J., P. Beard and J. Jiricny (1989). "Cytosine methylation in CTF and Sp1 recognition sites of an HSV tk promoter: effects on transcription in vivo and on factor binding in vitro." Nucleic Acids Res. **17**: 10179-10190.

Besser, D., F. Götz, K. Schulze-Forster, H. Wagner, H. Kröger, et al. (1990). "DNA methylation inhibits transcription by RNA polymerase III of a tRNA gene, but not of a 5S rRNA gene." FEBS Lett. **269**: 258-362.

Bestor, T., A. Laudano, R. Mattaliano and V. Ingham (1988). "Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells." J. Mol. Biol. **203**: 971-983.

Bestor, T.H., V.L. Chandler and A.P. Feinberg (1994). "Epigenetic effects in eukaryotic gene expression." Develop. Genet. **15**: 458-462.

Bird, A. (1992). "The essentials of DNA methylation." Cell **70**: 5-8.

Bird, A.P. (1986). "CpG-rich islands and the function of DNA methylation." Nature **321**: 209-213.

Bird, A.P. and E.M. Southern (1978). "Use of restriction enzymes to study eukaryotic DNA methylation." J. Mol. Biol. **118**: 27-47.

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

Björk, G.R., J.U. Ericson, C.E.D. Gustafsson, T.G. Hagervall, Y.H. Jönsson, et al. (1987). "Transfer RNA modification." Ann. Rev. Biochem. **56**: 263-287.

Blondel, A. and J. Thillet (1991). "A fast and convenient way to produce single stranded DNA from a phagemid." Nucleic Acids Res. **19**: 181.

Bonne-Andrea, C., F. Harper, J. Sobczak and A.-M. DeRecondo (1984). "Rat liver HMG1: a physiological nucleosome assembly factor." EMBO J. **3**: 1193-1199.

Boulikas, T., J.M. Wiseman and W.T. Garrard (1980). "Points of contact between histone H1 and the chromatin octamer." Proc. Natl. Acad. Sci. USA **77**: 127-131.

Bourn, D., T. Carr, D. Livingstone, A. McLaren and J.P. Goddard (1994). "An intron-containing $tRNA^{Arg}$ gene within a large cluster of human tRNA genes." DNA Sequence **5**: 83-92.

Bouvet, P., S. Dimitrov and A.P. Wolffe (1994). "Specific regulation of *Xenopus* chromosomal 5S rRNA gene transcription in vivo by histone H1." Genes Dev. **8**: 1147-1159.

Boyes, J. and A. Bird (1992). "Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein." EMBO J. **11**: 327-333.

Boyes, J. and A.P. Bird (1991). "DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein." Cell **64**: 1123-1134.

Bradford, M.M. (1976). "A rapid and sensitive method for quantitation of microgramme quantities of protein utilising the principle of protein dye binding." Anal. Biochem. **72**: 248-254.

Brandeis, M., D. Frank, I. Keshet, Z. Siegfried, M. Mendelsohn, et al. (1994). "Sp1 elements protect a CpG island from de novo methylation." Nature **371**: 435-438.

Brandeis, M., T. Kafri, M. Ariel, J.R. Chaillet, J. McCarrey, et al. (1993). "The ontogeny of allele-specific methylation associated with imprinted genes in the mouse." EMBO J. **12**: 3669-3677.

Bresnick, E.H., M. Bustin, V. Marsaud, H. Richard-Foy and G.L. Hager (1991). "The transcriptionally-active MMTV promoter is depleted of histone H1." Nucleic Acids Res. **20**: 273-278.

Brown, C.J., A. Ballabio, J.L. Rupert, R.G. Lafreniere, M. Grompe, et al. (1991). "A gene from the region of the human X-inactivation centre is expressed exclusively from the inactive X chromosome." Nature **349**: 38-44.

Brown, C.J., R.G. Lafreniere, V.E. Powers, G. Sebastio, A. Ballabio, et al. (1991). "Localisation of the X Inactivation Centre on the human X chromosome in Xq13." Nature **349**: 82-84.

Brown, D.D. (1984). "The role of stable complexes that repress and activate eukaryotic genes." Cell **37**: 359-365.

Bryans, M., S. Kass, C. Seivwright and R.L.P. Adams (1992). "Vector methylation inhibits transcription from the SV40 early promoter." FEBS Lett. **309**: 97-102.

Buratowski, S. (1994). "The basics of basal transcription by RNA polmerase II." Cell **77**: 1-3.

Burdon, R.H. and R.L.P. Adams (1980). "Eukaryotic DNA methylation." Trends Biochem. Sci. **5**: 294-297.

Buschausen, G., B. Wittig, M. Graessmann and A. Graessmann (1987). "Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene." Proc. Natl. Acad. Sci. USA **84**: 1177-1181.

Busslinger, M., J. Hurst and R.A. Flavell (1983). "DNA methylation and the regulation of globin gene expression." Cell **34**: 197-206.

Caiafa, P., A. Reale, P. Allegra, M. Rispoli, M. D'Erme, et al. (1991). "Histones and DNA methylation in mammalian chromatin : differential inhibition by histone H1." Biochem. Biophys. Acta **1090**: 38-42.

Caiafa, P., A. Reale, R. Santoro, M. D'Erme, S. Marenzi, et al. (1995). "Does hypo-methylation of linker DNA play a role in chromatin condensation?" Gene **157**: 247-251.

Carlson, M. and B.C. Laurent (1994). " The SNF/SWI family of global transcriptional activators." Curr. Opin. Cell Biol. **6**: 396-402.

Cary, P.D., T. Moss and E.M. Bradbury (1978). "High-resolution proton-magnetic-resonance studies of chromatin core particles." Eur. J. Biochem. **89**: 475-482.

Castiglia, D., R. Gristina, M. Scaturro and I. Di Liegro (1993). "Cloning and analysis of cDNA for rat histone H1°." Nucleic Acids Res. **21**: 1674.

Cheng, X. (1995). "Structure and function of DNA methyltransferases." Ann. Rev. Biophys. Biomol. Struct. **24**: 293-318.

Choi, Y.C. and C.B. Chae (1991). "DNA hypomethylation and germ cell-specific expression of testis-specific H2B histone gene." J. Biol. Chem. **266**: 20504-20511.

Church, G.M. and W. Gilbert (1984). "Genomic sequencing." Proc. Natl. Acad. Sci. USA **81**: 1991-1995.

Clark, D.J. and T. Kimura (1990). "Electrostatic mechanism of chromatin folding." J. Mol. Biol. **211**: 883-896.

Clark, D.J. and J.O. Thomas (1986). "Salt-dependent co-operative interaction of histone H1 with linear DNA." J. Mol. Biol. **187**: 569-580.

Clark, D.J. and J.O. Thomas (1988). "Differences in the binding of H1 variants to DNA." Eur. J. Biochem. **178**: 225-233.

Clark, S.J., J. Harrison and M. Frommer (1995). "CpNpG methylation in mammalian cells." Nature Genet. **10**: 20-27.

Cole, R.D. (1989). "Purification and analysis of H1 histones." *Nucleosomes.* P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170:** 524-536.

Comb, M. and H.M. Goodman (1990). "CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2." Nucleic Acids Res. **18**: 3975-3982.

Cooper, D.N., M.H. Taggart and A.P. Bird (1983). "Unmethylated domains in vertebrate DNA." Nucleic Acid Res. **11**: 647-658.

Cooper, D.N. and H. Youssoufian (1988). "The CpG dinucleotide and human genetic disease." Hum. Genet. **78**: 151-155.

Coulondre, C., J.H. Miller, P.J. Farabaugh and W. Gilbert (1978). "Molecular basis of base substitution hotspots in *Escherichia coli.* " Nature **274**: 775-780.

Cross, S.H. and A.P. Bird (1995). "CpG islands and genes." Curr. Opin. Genet. Devel. **5**: 309-314.

Cross, S.H., J.A. Charlton, X. Nan and A.P. Bird (1994). "Purification of CpG islands using a methylated DNA binding column." Nature Genet. **6**: 236-244.

Croston, G.E., L.A. Kerrigan, L.M. Lira, D.R. Marshak and J.T. Kadonaga (1991). "Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription." Science **251**: 643-649.

Csordas, A. (1990). "On the biological role of histone acetylation." Biochem. J. **265**: 23-38.

D'Erme, M., R. Santoro, P. Allegra, A. Reale, S. Marenzi, et al. (1993). "Inhibition of CpG methylation in linker DNA by H1 histone." Biochem. Biophys. Acta **1173**: 209-216.

Deobagkar, D.D., M. Liebler, M. Graessmann and A. Graessmann (1990). "Hemi-methylation of DNA prevents chromatin expression." Proc. Natl. Acad. Sci. USA **87**: 1691-1695.

Deutscher, M.P. (1984). "Processing of tRNA in prokaryotes and eukaryotes." Crit. Rev. Biochem. **17**: 45-71.

Dignam, J.D., R.M. Lebovitz and R.G. Roeder (1983). "Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei." Nucleic Acids Res. **11**:1475-1486.

Dimitrov, S., G. Almouzni, M. Dasso and A.P. Wolffe (1993). "Chromatin transitions during early *Xenopus* embryogenesis: changes in histone H4 acetylation and in l inker histone type." Develop. Biol. **160**: 214-227.

Dimitrov, S. and A.P. Wolffe (1995). "Chromatin and nuclear assembly: experimental approaches towards the reconstitution of transcriptionally active and silent states." Biochem. Biophys. Acta **1260**: 1-13.

Docherty, K. and A.R. Clark (1993). "Transcription of exogenous genes in mammalian cells." *Gene Transcription: a Practical Approach*. B.D. Hames and S.J. Higgins. Oxford, IRL Press: 65-123.

Doerfler, W. (1983). "DNA methylation and gene activity." Ann. Rev. Biochem. **52**: 93-124.

Draetta, G. (1990). "Cell cycle control in eukaryotes: molecular mechanisms of cdc2 activation." Trends Biochem. Sci. **15**: 378-383.

Drew, H.R. and A.A. Travers (1985). "DNA bending and its relation to nucleosome positioning." J. Mol. Biol. **186**: 773-790.

Eden, S. and H. Cedar (1994). "Role of DNA methylation in the regulation of transcription." Curr. Opin. Genet. Dev. **4**: 255-259.

Efstratiadis, A. (1994). "Parental imprinting of autosomal mammalian genes." Curr. Opin. Genet. Dev. **4**: 265-280.

Ehrlich, M. and K.C. Ehrlich (1993). "Effect of DNA methylation on the binding of vertebrate and plant proteins to DNA." *DNA Methtlation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 145-168.

Elgin, S.C.R. (1990). "Chromatin structure and gene activity." Curr. Opin. Cell. Biol. **2**: 437-445.

Englander, E.W., A.P. Wolffe and B.H. Howard (1993). "Nucleosome interactions with a human Alu element; transcriptional repression and effects of template methylation." J. Biol. Chem. **268**: 19565-19573.

Fanning, E. and R. Knippers (1992). "Simian virus 40 large 'I' antigen." Ann. Rev. Biochem. **61**: 55-85.

Felsenfeld, G. and J.D. McGhee (1986). "Structure of the 30 nm chromatin fiber." Cell **44**: 375-377.

Felsenfeld, G., J. Nickol, M. Behe, J. McGhee and D. Jackson (1982). "Methylation and chromatin structure." Cold Spring Harbor Symp. Quant. Biol. **47**: 577-584.

Ferguson-Smith, A.C., H. Sasaki, B.M. Cattanach and M.A. Surani (1993). "Parental-origin-specific epigenetic modification of the mouse *H19* gene." Nature **362**: 751-754.

Finch, J.T., M. Noll and R.D. Kornberg (1975). "Electron microscopy of defined lengths of chromatin." Proc. Natl. Acad. Sci. USA **72**: 3320-3322.

Finnegan, E.J. and E.S. Dennis (1993). "Isolation and identification by sequence homology of a putative cytosine methyltransferase from *Arabidopsis thaliana*." Nucleic Acids Res. **21**: 2383-2388.

Frank, D., I. Keshet, M. Shani, A. Levine, A. Razin, et al. (1991). "Demethylation of CpG islands in embryonic cells." Nature **351**: 239-241.

Frommer, M., L.E. McDonald, D.S. Millar, C.M. Collis, F. Watt, et al. (1992). "A genomic sequencing protocol which yields a positive display of 5-methylcytosine residues in individual DNA strands." Proc. Natl. Acad. Sci. USA **89**: 1827-1831.

Fundele, R.H. and M.A. Surani (1994). "Experimental embryological analysis of genetic imprinting in mouse development." Dev. Genet. **15**: 515-522.

Gabrielsen, O.S. and A. Sentenac (1991). "RNA polymerase III (C) and its transcription factors." Trends Biochem. Sci. **16**: 412-416.

Garrard, W.T. (1991). "Histone H1 and the conformation of transcriptionally active chromatin." BioEssays **13**: 87-88.

Gey, G.O., W.D. Coffman and M.T. Kubicek (1952). "Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium." Cancer Res. **12**: 264-265.

Giancotti, V., A. Bandiera, L. Ciani, D. Santoro, C. Crane-Robinson, et al. (1993). "High-mobility-group (HMG) proteins and histone H1 subtypes expression in normal and tumour tissues of mouse." Eur. J. Biochem. **213**: 825-832.

Gonos, E.S. and J.P. Goddard (1990). "The role of the 5'-flanking sequence of a human $tRNA^{Glu}$ gene in modulation of its transcriptional activity in vitro." Biochem. J. **272**: 797-803.

Gottesfeld, J.M. (1989). "Analysis of RNA polymerase III transcription in vitro using chromatin and cloned gene templates." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170**: 347-359.

Götz, F., K. Schulze-Forster, H. Wagner, H. Kröger and D. Simon (1990). "Transcription inhibition of SV40 by in vitro DNA methylation." Biochim. Biophys. Acta **1087**: 323-329.

Graessmann, M. and A. Graessmann (1993). "DNA methylation, chromatin structure and the regulation of gene expression." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 404-424.

Graessmann, M., A. Graessmann, H. Wagner, E. Werner and D. Simon (1983). "Complete DNA methylation does not prevent polyoma and simian virus 40 virus early gene expression." Proc. Natl. Acad. Sci. USA **80**: 6470-6474.

Grant, S.G. and V.M. Chapman (1988). "Mechanisms of X-chromosome inactivation." Ann. Rev. Genet. **22**: 199-233.

Graziano, V., S.E. Gerchman and V. Ramakrishnan (1988). "Reconstitution of chromatin higher-order structure from histone H5 and depleted chromatin." J. Mol. Biol. **203**: 997-1007.

Graziano, V., S.E. Gerchman, D.K. Schneider and V. Ramakrishnan (1994). "Histone H1 is located in the interior of the chromatin 30-nm filament." Nature **368**: 351-354.

Grigg, G. and S. Clark (1994). "Sequencing 5-methylcytosine residues in genomic DNA." BioEssays **16**: 431-438.

Gross, D.S. and W.T. Garrard (1988). "Nuclease hypersensitive sites in chromatin." Ann. Rev. Biochem. **57**: 159-197.

Grosschedl, R., K. Giese and J. Pagel (1994). "HMG domain proteins: architectural elements in the assembly of nucleoprotein structures." Trends Genet. **10**: 94-100.

Groudine, M., R. Eisenman and H. Weintraub (1981). "Chromatin structure of endogenous retroviral genes and activation by an inhibitor of of DNA methylation." Nature **292**: 311-317.

Gruenbaum, Y., H. Cedar and A. Razin (1982). "Substrate and sequence specificity of a eukaryotic DNA methylase." Nature **295**: 620-622.

Gruenbaum, Y., T. Naveh-Many, H. Cedar and A. Razin (1981). "Sequence specificity of methylation in higher plant DNA." Nature **292**: 860-862.

Gruss, C., C. Gutierrez, W.C. Burhans, M.L. DePamphilis, T. Koller, et al. (1990). "Nucleosome assembly in mammalian cell extracts before and after DNA replication." EMBO J. **9**: 2911-2922.

Hansen, R., N. Ellis and S. Gartler (1988). "Demethylation of specific sites in the 5' r egion of the inactive X-linked human phosphoglycerate kinase gene correlates with the appearance of nuclease sensitivity and gene expression." Mol. Cell. Biol. **8**: 4692-4699.

Hatada, I., T. Sugama and T. Mukai (1993). "A new imprinted gene cloned by a methylation-sensitive genome scanning method." Nucleic Acids Res. **21**: 5577-5582.

Hawley, D.K. and R.G. Roeder (1987). "Functional steps in transcription inhibition and reinitiation from the major late promoter in a HeLa nuclear extract." J. Biol. Chem. **262**: 3452-3461.

Hayashizaki, Y., H. Shibata, S. Hirotsune, H. Sugino, Y. Okazaki, et al. (1994). "Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method." Nature Genet. **6**: 33-40.

Hebbes, T.R., A.L. Clayton, A.W. Thorne and C. Crane-Robinson (1994). "Core histone hyperacetylation co-maps with generalised DNase I sensitivity in the chicken β-globin chromosomal domain." EMBO J. **13**: 1823-1830.

Helliger, W., H. Lindner, O. Grübl-Knosp and B. Puschendorf (1992). "Alteration in proportions of histone H1 variants during the differentiation of murine erythroleukaemic cells." Biochem. J. **288**: 747-751.

Hentzen, P.C. and I. Bekhor (1985). "Characterisation of the 2 M NaCl-resistant chromatin fraction from chicken erythroid cells." *Progress in Non-Histone Research*. I. Bekhor. Boca Raton FA, CRC Press. **1**: 75-101.

Hergersberg, M. (1991). "Biological aspects of cytosine methylation in eukaryotic cells." Experientia **47**: 1171-1185.

Hershkovitz, M. and A.D. Riggs (1995). "Metaphase chromosome analysis by ligation-mediated PCR: heritable chromatin structure and a comparison of active and inactive X chromosomes." Proc. Natl. Acad. Sci. USA **92**: 2379-2383.

Higurashi, M. and R.D. Cole (1991). "The combination of DNA methylation and H1 histone binding inhibits the action of a restriction nuclease on plasmid DNA." J. Biol. Chem. **266**: 8619-8625.

Hill, C.S., L.C. Packman and J.O. Thomas (1990). "Phosphorylation at clustered -Ser-Pro-X-Lys/Arg- motifs in sperm-specific histones H1 and H2B." EMBO J. **9**: 805-813.

Hill, C.S., J.M. Rimmer, B.N. Green, J.T. Finch and J.O. Thomas (1991). "Histone-DNA interactions and their modulation by phosphorylation of -Ser-Pro-X-Lys/Arg- motifs." EMBO J. **10**: 1939-1948.

Hill, C.S. and J.O. Thomas (1990). "Core histone-DNA interactions in sea urchin sperm chromatin." Eur. J. Biochem. **187**: 145-153.

Hodges-Garcia, Y. and P.J. Hagerman (1992). "Cytosine methylation can induce local distortions in the structure of duplex DNA." Biochem. **31**: 7595-7599.

Höller, M., G. Westin, J. Jiricny and W. Schaffner (1988). "Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated." Genes Dev. **2**: 1127-1135.

Holliday, R. (1987). "The inheritance of epigenetic defects." Science **238**: 163-170.

Holliday, R. (1989). "A different kind of inheritance". Sci. Amer. **266**: 40-48.

Holliday, R. (1994). "Epigenetics: an overview." Dev. Genet. **15**: 453-457.

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

Hornstra, I.K. and T.P. Yang (1994). "High-resolution methylation analysis of the human hypoxanthine phosphoribosyltransferase gene 5' region on active and inactive X chromosomes: correlation with binding site for transcription factors." Mol. Cell. Biol. **14**: 1419-1430.

Hotchkiss, R.D. (1948). "The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography." J. Biol. Chem. **168**: 315-332.

Hsieh, C.-L. (1994). "Dependence of transcriptional repression on CpG methylation density." Mol. Cell. Biol. **14**: 5487-5494.

Huang, L.-H., R. Wang, M.A. Gama-Sosa, S. Shenoy and M. Ehrlich (1984). "A protein from human placental nuclei binds preferentially to 5-methylcytosine-rich DNA." Nature **308**: 293-295.

Iguchi-Ariga, S.M.M. and W. Schaffner (1989). "CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptioal activation." Genes Dev. **3**: 612-619.

Jackson, S.P. (1993). "Identification and characterisation of eukaryotic and transcription factors." *Gene Transcription: A Practical Approach.* B.D. Hames and S.J. Higgins. Oxford, IRL Press: 189-242.

Jeppesen, P. and B.M. Turner (1993). "The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression." Cell **74**: 281-289.

Jerzmanowski, A. and R.D. Cole (1990). "Flanking sequences of *Xenopus* 5S RNA genes determine differential inhibition of transcription by H1 histone in vitro." J. Biol. Chem. **265**: 10726-10732.

Johnson, C.A., J.P. Goddard and R.L.P. Adams (1995). "The effect of histone H1 and DNA methylation on transcription." Biochem. J. **305**: 791-798.

Johnson, T.B. and R.D. Coghill (1925). "The discovery of 5-methylcytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus." J. Am. Chem. Soc. **47**: 2838-2844.

Jones, P. (1984). "Gene activation by 5-azacytidine." *DNA Methylation: Biochemistry and Biological Significance*. A. Razin, H. Cedar and A.D. Riggs. New York, Springer-Verlag: 165-187.

Jones, P.A., W.M. Rideout III, J.-C. Shen, C.H. Spruck and Y.C. Tsai (1992). "Methylation, mutation and cancer." BioEssays **14**: 33-36.

Jones, P.A., M.J. Wolkowicz, W.M. Rideout, F.A. Gonzales, C.M. Marziasz, et al. (1990). "De novo methylation of the *MyoD1* CpG island during the establishment of immortal cell lines." Proc. Natl. Acad. Sci. USA **87**: 6117-6121.

Jost, J.-P. and J. Hofsteenge (1992). "The repressor MDBP-2 is a member of the histone H1 family that binds preferentially in vitro and in vivo to methylated nonspecific DNA sequences." Proc. Natl. Acad. Sci. USA **89**: 9499-9503.

Jost, J.-P., H.-P. Saluz and A. Pawlak (1991). "Estradiol down regulates the binding activity of an avian vitellogenin gene repressor (MDBP-2) and triggers a gradual demethylation of the mCpG pair of its DNA binding site." Nucleic Acids Res. **19**: 5771-5775.

Jost, J.P. (1993). "Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxtcytidine." Proc. Natl. Acad. Sci. USA **90**: 4684-4688.

Jost, J.P. and Y.C. Jost (1994). "Transient DNA demethylation in differentiating mouse myoblasts correlates with higher activity of 5-methyldeoxycytidine excision r epair." J. Biol. Chem. **269**: 10040-10043.

Jost, J.P. and H.P. Saluz (1993). "Steroid hormone-dependent changes in DNA methylation and its significance for the activation or silencing of specific genes." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 425-451.

Jüttermann, R., K. Hosokawa, S. Kochanek and W. Doerfler (1991). "Adenovirus type 2 VAI RNA transcription by polymerase III is blocked by sequence-specific methylation." J. Virol. **65**: 1735-1742.

Kadonaga, J.T. (1990). "Assembly and disassembly of the *Drosophila* RNA polymerase II complex during transcription." J. Biol. Chem. **265**: 2624-2631.

Kafri, T., M. Ariel, M. Brandeis, R. Shemer, L. Urven, et al. (1992). "Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line." Genes Dev. **6**: 705-714.

Kamakaka, R.T., M. Bulger and J.T. Kadonaga (1993). "Potentiation of RNA polymerase II transcription by Gal4-VP16 during but not after DNA replication and chromatin assembly." Genes Dev. **7**: 1779-1795.

Kamakaka, R.T. and J.O. Thomas (1990). "Chromatin structure of transcriptionally competent and repressed genes." EMBO J. **9**: 3997-4006.

Karpen, G.H. (1994). "Position-effect varigation and the new biology of heterochromatin." Curr. Opin. Genet. Dev. **4**: 281-291.

Käs, E., L. Poljak, Y. Adachi and U.K. Laemmli (1993). "A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin." EMBO J. **12**: 115-126.

Kass, S.U., J.P. Goddard and R.L.P. Adams (1993). "Inactive chromatin spreads from a focus of methylation." Mol. Cell. Biol. **13**: 7372-7379.

Kerem, B.S., R. Goitein, C. Richler, M. Marcus and H. Cedar (1983). "In situ nick-translation distinguishes between active and inactive X-chromosomes." Nature **304**: 88-90.

Keshet, I., J. Lieman-Hurwitz and H. Cedar (1986). "DNA methylation affects the formation of active chromatin." Cell **44**: 535-543.

Keshet, I., J. Yisraeli and H. Cedar (1985). "Effect of regional DNA methylation on gene expression." Proc. Natl. Acad. Sci. USA **82**: 2560-2564.

Khochbin, S. and A.P. Wolffe (1994). "Developmentally regulated expression of linker-histone variants in vertebrates." Eur. J. Biochem. **225**: 501-510.

Kleinschmidt, J.A., A. Seiter and H. Zentgraf (1990). "Nucleosome assembly in vitro: separate histone transfer and synergistic interaction of native histone complexes purified from nuclei of *Xenopus laevis* oocytes." EMBO J. **9**: 1309-1318.

Klimasauskas, S., S. Kumar, R.R. Roberts and X. Cheng (1994). "*Hha* I methyltransferase flips its target base out of the DNA helix." Cell **76**: 357-369.

Kobagashi, H., J. Ngernprasirtsir and T. Akazawa (1990). "Transcriptional regulation and DNA methylation in plastids during transitional conversion of chloroplasts to chromoplasts." EMBO J. **9**: 307-313.

Kochanek, S., D. Renz and W. Doerfler (1993). "Differences in the accessibility of methylated and unmethylated DNA to DNase I." Nucleic Acids Res. **21**: 5843-5845.

Kovesdi, I., R. Reichel and J.R. Nevins (1987). "Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control." Proc. Natl. Acad. Sci. USA **84**: 2180-2184.

Lai, Z.-C. and G. Childs (1988). "Characterisation of the structure and transcriptional patterns of the gene encoding the late histone subtype H1-β of the sea urchin *Strongylocentrotus purpuratus*." Mol. Cell. Biol. **8**: 1842-1844.

Lambert, S.F. and J.O. Thomas (1986). "Lysine-containing DNA-binding regions on the surface of the histone octamer in the nucleosome core particle." Eur. J. Biochem. **160**: 191-201.

Laybourn, P.J. and J.T. Kadonaga (1991). "Role of nucloesomal cores and histone H1 in regulation of transcription by RNA polymerase II." Science **254**: 238-245.

Lee, D.Y., J.J. Hayes, D. Pruss and A.P. Wolffe (1993). "A positive role for histone acetylation in transcription factor access to nucleosomal DNA." Cell **72**: 73-84.

Lennox, R.W. (1984). "Differences in evolutionary stability among mammalian H1 subtypes." J. Biol. Chem. **259**: 669-672.

Lennox, R.W. and L.H. Cohen (1989). "Analysis of histone subtypes and their modified forms by polyacrylamide gel electrophoresis." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170:** 532-549.

Leonard, M.W. and R.K. Patient (1991). "Primer extension analysis of mRNA isolated from transfected cell lines." *Methods in Molecular Biology*. E.J. Murray. Clifton NJ, The Humana Press Inc. **7:** 297-306.

Leonhardt, H. and T.H. Bestor (1993). "Structure, function and regulation of mammalian DNA methyltransferase." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 109-119.

Leonhardt, H., A.W. Page, H.-U. Weier and T.H. Bestor (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei." Cell **71**: 865-873.

Lettmann, C., B. Schmitz and W. Doerfler (1991). "Persistence or loss of preimposed methylation patterns and de novo methylation of foreign DNA integrated in transgenic mice." Nucleic Acids Res. **19**: 7131-7137.

Levine, A., G.L. Cantoni and A. Razin (1991). "Inhibition of promoter activity by methylation: possible involvment of protein mediators." Proc. Natl. Acad. Sci. USA **88**: 6515-6518.

Levine, A., G.L. Cantoni and A. Razin (1992). "Methylation in the preinitiation domain suppresses gene transcription by an indirect mechanism." Proc. Natl Acad. Sci. USA **89**: 10119-10123.

Levine, A., A. Yeivin, E. Ben-Asher, Y. Aloni and A. Razin (1993). "Histone H1-mediated inhibition of transcription initiation of methylated templates in vitro." J. Biol. Chem. **268**: 21754-21759.

Lewis, J. and A. Bird (1991). "DNA methylation and chromatin structure." FEBS Lett. **285**: 155-159.

Lewis, J.D., R.R. Meehan, W.J. Henzel, I. Maurer-Fogy, P. Jeppesen, et al. (1992). "Purification, sequence and cellular localisation of a novel chromosomal protein that binds to methylated DNA." Cell **69**: 905-914.

Li, E., C. Beard, A.C. Forster, T.H. Bestor and R. Jaenisch (1993a). "DNA methylation, genomic imprinting, and mammalian development." Cold Spring Harbor Symp. Quant. Biol. **58**: 297-305.

Li, E., C. Beard and R. Jaenisch (1993b). "Role for DNA methylation in genomic imprinting." Nature **366**: 362-365.

Li, E., T.H. Bestor and R. Jaenisch (1992). "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." Cell **69**: 915-926.

Lin, X.H., L.J. Gu, A.P. Bird and T.F. Denel (1995). "A methyl CpG binding protein MeCP-2 selectively suppresses the transcription of methylated gene and activates non-methylated gene." J. Cell. Biochem. suppl. 19A (S1A): 68.

Lock, L., N. Takagi and G. Martin (1987). "Methylation of the HPRT gene on the inactive X occurs after chromosome inactivation." Cell **48**: 39-46.

Locker, J. (1993). "Transcription controls: *cis* -elements and *trans* -factors." *Gene Transcription: a Practical Approach.* B.D. Hames and S.J. Higgins. Oxford, IRL Press: 321-345.

Louters, L. and R. Chalkley (1985). "Exchange of histone H1, H2A, and H2B in vivo." Biochem. **24**: 3080-3085.

Lyon, M.F. (1993). "Epigenetic inheritance in mammals." Trends Genet. **9**: 123-128.

Macleod, D., J. Charlton, J. Mullins and A.P. Bird (1994). "Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island." Genes Devel. **8**: 2282-2292.

Manley, J.L., A. Fire, A. Cano, P.A. Sharp and M.L. Gefter (1980). "DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract." Proc. Natl. Acad. Sci. USA **77**: 3855-3859.

Mason, P.J., T. Enver, D. Wilkinson and J.G. Williams (1993). "Assay of gene transcription in vitro." *Gene Transcription, a practical approach.* B.D. Hames and S.J. Higgins. Oxford, IRL Press.

Maxam, A.M. and W. Gilbert (1980). "Sequencing end-labeled DNA with base-specific chemical cleavages." Meth. Enzymol. **65**: 499-560.

Meehan, R.R., J.D. Lewis and A.P. Bird (1992). "Characterisation of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA." Nucleic Acids Res. **20**: 5085-5092.

Meehan, R.R., J.D. Lewis, S. McKay, E.L. Kleiner and A.P. Bird (1989). "Identification of a mammalian protein that binds specifically to DNA containing methylated DNA." Cell **58**: 499-507.

Meersseman, G., S. Pennings and E.M. Bradbury (1992). "Mobile nucleosomes: a general behavior." EMBO J. **11**: 2951-2959.

Miller, L.L., W. Schnoedl, J. Allen and B.F. Erlanger (1974). "5-methylcytosine localised in mammalian constitutive heterochromatin." Nature **251**: 636-637.

Modrich, P. (1991). "Mechanisms and biological effects of mismatch repair." Ann. Rev. Genet. **25**: 229-253.

Mohandas, T., R.S. Sparkes and L.J. Shapiro (1981). "Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation." Science **211**: 393-396.

Monk, M. (1990). "Changes in DNA methylation during mouse embryonic development in relation to X-chromosome activity and imprinting." Phil. Trans. Roy. Soc. Lond. **B 326**: 179-187.

Monk, M., R.L.P. Adams and A. Rinaldi (1991). "Decrease in DNA methylase activity during preimplantation development in the mouse." Development **112**: 189-192.

Monk, M., M. Boubelik and S. Lehnert (1987). "Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development." Development **99**: 371-382.

Murray, E.J. and F. Grosveld (1987). "Site specific demethylation in the promoter of human γ-globin gene does not alleviate methylation mediated suppression." EMBO J. **6**: 2329-2335.

Muyldermans, S., J. De Jonge, L. Wyns and A.A. Travers (1994). "Differential association of linker histones H1 and H5 with telomeric nucleosomes in chicked erythrocytes." Nucleic Acids Res. **22**: 5635-5639.

Nan, X., R.R. Meehan and A. Bird (1993). "Dissection of the methyl-CpG binding domain from chromosomal protein MeCP2." Nucleic Acids Res. **21**: 4886-4892.

Ner, S.S., M.E.A. Churchill, M.A. Searles and A.A. Travers (1993). "dHMG-Z, a second HMG-1-related protein in *Drosophila melanogaster*." Nucleic Acids Res. **21**: 4369-4371.

Nightingale, K. and A.P. Wolffe (1995). "Methylation at CpG sequences does not influence histone H1 binding to a nucleosome including a Xenopus borealis 5S rRNA gene." J. Biol. Chem. **270**: 4197-4200.

Noll, H. and M. Noll (1989). "Sucrose gradient techniques and applications to nucleosome structure." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170**: 57-116.

Noll, M. and K. Kornberg (1977). "Action of micrococcal nuclease on chromatin and the location of histone H1." J. Mol. Biol. **109**: 393-404.

Norris, D.P.N., D. Patel, G.F. Kay, G.D. Penny, N. Brockdorff, et al. (1994). "Evidence that random and imprinted *Xist* expression is controlled by preemptive methylation." Cell **77**: 41-51.

Noyer-Weidner, M. and T.A. Trautner (1993). "Methylation of DNA in prokaryotes." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 39-109.

Nur, I., M. Szyf, A. Razin, G. Glaser, S. Rottem, et al. (1985). "Procaryotic and eucaryotic traits of DNA methylation in *Spiroplasmas* (Mycoplasmas)." J. Bacteriol. **164**: 19-24.

Oberlé, I., F. Rousseau, D. Heitz, C. Kretz, D. Devys, et al. (1991). "Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome." Science **252**: 1097-1102.

Paranjape, S.M., R.T. Kamakaka and J.T. Kadonaga (1994). "Role of chromatin structure in the regulation of transcription by RNA polymerase II." Ann. Rev. Biochem. **63**: 265-297.

Pardue, M.L. and W. Hennig (1990). "Heterochromatin: junk or collector's item?" Chromosoma **100**: 3-7.

Paro, R. (1993). "Mechanisms of heritable gene repression during development of *Drosophila*." Curr. Opin. Cell. Biol. **5**: 999-1005.

Paroush, Z., I. Keshet, J. Yisraeli and H. Cedar (1990). "Dynamics of demethylation and activation of the α-actin gene in myoblasts." Cell **63**: 1229-1237.

Pawlak, A., M. Bryans and J.P. Jost (1991). "An avian 40 kDa nucleoprotein binds preferentially to a promoter sequence containing one single pair of methylayed CpG." Nucleic Acids Res. **19**: 1029-1034.

Pennings, S., G. Meersseman and E.M. Bradbury (1994). "Linker histones H1 and H5 prevent the mobility of positioned nucleosomes." Proc. Natl. Acad. Sci. USA **91**: 10275-10279.

Peterson, K. and C. Sapienza (1993). "Imprinting the genome: imprinted genes, imprinting genes, and a hypothesis for their interaction." Ann. Rev. Genet. **27**: 7-31.

Pfeifer, G.P. and A.D. Riggs (1991). "Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilised cells using DNase I and ligation-mediated PCR." Genes Devel. **5**: 1102-1113.

Pfeifer, G.P., R. Tanguay, S. Steigerwald and A. Riggs (1990). "In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1." Genes Devel. 4: 1277-1287.

Pollack, Y., J. Kasir, T. Shemer, S. Metzger and M. Szyf (1984). "Methylation pattern of mouse mitochondrial DNA." Nucleic Acids Res. **12**: 4811-4824.

Postnikov, Y.V., V.V. Shick, A.V. Belyavsky, K.R. Khrapko, K.L. Brodolin, et al. (1991). "Distribution of high mobility group proteins 1,2E and 14/17 and linker histones H1 and H5 on transcribed and non-transcribed regions of chicken erythrocyte chromatin." Nucleic Acids Res. **19**: 717-725.

Pradhan, S. and R.L.P. Adams (1995). "Distinct CG and CNG methyltransferases in *Pisum sativum*." Plant J. **7**: 471-481.

Prendergast, G.C., D. Lawe and E.B. Ziff (1991). "Association of Myn, the murine homolog of Max, with c-myc stimulates methylation-sensitive DNA binding and *ras* cotrans-formation." Cell **65**: 395-407.

Pruss, D., J.J. Hayes and A.P. Wolffe (1995). "Nucleosomal anatomy - where are the histones?" BioEssays **17**: 161-170.

Quesada, P., B. Farina and R. Jones (1989). "Poly(ADP-ribosylation) of nuclear proteins in rat testis correlates with active spermatogenesis." Biochim. Biophys. Acta **1007**: 167-175.

Ramakrishnan, V. (1994). "Histone Structure." Curr. Opin. Struct. Biol. **4**: 44-50.

Ramakrishnan, V., J.T. Finch, V. Graziano, P.L. Lee and R.M. Sweet (1993). "Crystal structure of globular domain of histone H5 and its implications for nucleosome binding." Nature **362**: 219-224.

Rastan, S. (1994). "X chromosome inactivation and the Xist gene." Curr. Opin. Genet. Dev. **4**: 292-297.

Razin, A. and H. Cedar (1977). "Distribution of 5-methylcytosine in chromatin." Proc. Natl. Acad. Sci. USA **74**: 2725-2728.

Razin, A. and H. Cedar (1991). "DNA methylation and gene expression." Microbiol. Rev. **55**: 451-458.

Razin, A. and H. Cedar (1993). "DNA methylation and embryogenesis." *DNA Methylation: Molecular Biology and Biological Significance.* J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 343-357.

Razin, A. and H. Cedar (1994). "DNA methylation and genomic imprinting." Cell **77**: 473-476.

Razin, A., M. Szyf, T. Kafri, M. Roll, H. Giloh, et al. (1986). "Replacement of 5-methylcytosine by cytosine: a possible mechanism for transient DNA demethylation during differentiation." Proc. Natl. Acad. Sci. USA **83**: 2827-2831.

Reuter, G. and P. Spierer (1992). "Position effect varigation and chromatin proteins." BioEssays **14**: 605-612.

Rhodes, D. and R.A. Laskey (1989). "Assembly of nucleosomes and chromatin in vitro." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170**: 575-585.

Rhodes, K., R.A. Rippe, A. Umezawa, M. Nehls, D.A. Brenner, et al. (1994). "DNA methylation represses the murine α1(I) collagen promoter by an indirect mechanism." Mol. Cell. Biol. **14**: 5950-5960.

Rideout, W.M., G.A. Coetzee, A.F. Olumi and P.A. Jones (1990). "5-methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes." Science **249**: 1288-1290.

Riggs, A.D. and G.P. Pfeifer (1992). "X-chromosome inactivation and cell memory." Trends Genet. **8**: 169-174.

Ritchie, R.J., S.J.L. Knight, M.C. Hirst, P.K. Grewal, M. Bobrow, et al. (1994). "The cloning of *FRAXF* : trinucleotide repeat expansion and methylation at a third fragile site in distal Xqter." Hum. Mol. Genet. **3**: 2115-2121.

Rodríguez-Campos, A., A. Shimamura and A. Worcel (1989). "Assembly and properties of chromatin containing histone H1." J. Mol. Biol. **209**: 135-150.

Roeder, R.G. (1991). "The complexities of eukaryotic trancription initiation: regulation of preinitiation complex assembly." Trends Biochem. Science **16**: 402-408.

Rosenthal, D.S. and J.L. Doering (1983). "The genomic organisation of dispersed tRNA and 5S RNA genes in *Xenopus laevis*." J. Biol. Chem. **258**: 7402-7410.

Rothnie, H.M., K.J. McCurrach, L.A. Glover and N. Hardman (1991). "Retrotransposon-like nature of Tp1 elements: implications for the organisation of highly repetitive, hypermethylated DNA in the genome of *Physarum polycephalum*." Nucleic Acids Res. **19**: 279-286.

Saluz, H.P. and J.P. Jost (1989). "Genomic footprinting with Taq polymerase." Proc. Natl. Acad. Sci. USA **86**: 2602-2606.

Saluz, H.P. and J.P. Jost (1993). "Major techniques to study DNA methylation." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser Verlag: 11-26.

Sambrook, J., E.F. Fritsch and T. Maniatis (1989). *Molecular cloning; a laboratory manual*. New York, Cold Spring Harbour Press.

Santoro, R., M. D'Erme, S. Mastrantonio, A. Reale, S. Marenzi, et al. (1995). "Binding of histone H1e-c variants to CpG-rich DNA correlates with the inhibitory effect on enzymic DNA methylation." Biochem. J. **305**: 739-744.

Santoro, R., M. D'Erme, A. Reale, R. Strom and P. Caiafa (1993). "Effect of H1 histone isoforms on the methylation of single- or double-stranded DNA." Biochem. Biophys. Res. Comm. **190**: 86-91.

Sapienza, C. (1990). "Parental imprinting of genes." Sci. Amer. **263**: 52-60.

Sasaki, T., R.S. Hansen and S.M. Gartler (1992). "Hemimethylation and hypersensitivity are early events in transcriptional reactivation of a human inactive X-linked genes in a hamster x human somatic cell hybrid." Mol. Cell. Biol. **12**: 3819-3826.

Schlissel, M.S. and D.D. Brown (1984). " The transcriptional regulation of *Xenopus* 5S RNA genes in chromatin: the roles of active stable transcription complexes and histone H1." Cell **37**: 903-913.

Schorderet, D.F. and S.M. Gartler (1990). "Absence of methylation at *Hpa* II sites in three human genomic tRNA sequences." Nucleic Acids Res. **18**: 6965-6969.

Schulze, E., L. Trieschmann, B. Schulze, E.R. Schmidt, S. Pitzel, et al. (1993). "Structural and functional differences between histone H1 sequence variants with differential intranuclear distribution." Proc. Natl. Acad. Sci. USA **90**: 2481-2485.

Schwartz, L.B., V.E.F. Sklar, S.J. Jachning, R. Weinmann and R.G. Roeder (1974). "Isolation and partial characterisation of the multiple forms of deoxyribonucleic acid-dependent ribonucleic acid polymerase in mouse myeloma MOPC." J. Biol. Chem. **249**: 5889-5897.

Sealy, L., R.R. Burgess, M. Cotten and R. Chalkley (1989). "Purification of *Xenopus* egg nucleoplasmin and its use in chromatin assembly in vitro." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170**: 612-630.

Shapiro, D.L., P.A. Sharp, W.W. Wahli and M.J. Keller (1988). "A high efficiency HeLa cell nuclear transcription extract." DNA **7**: 47-55.

Sharp, S.J., J. Schaack, L. Cooley, D.J. Burke and D. Soll (1984). "Structure and transcription of eukaryotic tRNA genes." CRC Crit. Rev. Biochem. **19**: 107-144.

Shemer, R., T. Kafri, A. O'Connel, S. Eisenberg, J.L. Breslow, et al. (1991). "Methylation changes in the apolipoprotein A1 gene during embryonic development of the mouse." Proc. Natl. Acad. Sci. USA **88**: 11300-11304.

Shemer, R., A. Walsh, S. Eisenberg, J. Breslow and A. Razin (1990). "Tissue-specific methylation patterns and expression of the human apolipoprotein A1 gene." J. Biol. Chem. **265**: 1010-1015.

Shimamura, A., B. Jessee and A. Worcel (1989). "Assembly of chromatin with oocyte extracts." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170**: 603-615

Shimamura, A., M. Sapp, A. Rodriguez-Campos and A. Worcel (1989). "Histone H1 represses transcription from minichromosomes assembled in vitro." Mol. Cell. Biol. **9**: 5573-5584.

Shimamura, A., D. Tremethick and A. Worcel (1988). "Characterisation of the repressed 5S DNA minichromosomes assembled in vitro with a high-speed supernatent of *Xenopus laevis* oocytes." Mol. Cell. Biol. **8**: 4257-4269.

Shortridge, R.D., G.D. Johnson, L.C. Craig, I.M. Pirtle and R.M. Pirtle (1989). "A human tRNA gene hetero-cluster encoding threonine, proline and valine tRNAs." Gene **79**: 309-324.

Sierra, F., J.-M. Tian and U. Schibler (1993). "In vitro transcription with nuclear extracts from differentiated tissues." *Gene Transcription: A Practical Approach*. B.D. Hames and S.J. Higgins. Oxford, IRL Press: 125-152.

Singer, J., J. Robert-Ems and A.D. Riggs (1979). "Methylation of mouse liver DNA studied by means of the restriction enzymes *Msp* I and *Hpa* II." Science **203**: 1019-1021.

Smale, S., M.C. Schmidt, A.J. Berk and D. Baltimore (1990). "Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID." Proc. Natl. Acad. Sci. USA **87**:4509-4513.

Solage, A. and H. Cedar (1978). "Organisation of 5-methylcytosine in chromosomal DNA." Biochem. **17**: 2934-2938.

Somma, P., C. Pisano and P. Lavia (1991). "The housekeeping promoter from the mouse CpG island *HTF9* contains multiple protein-binding elements that are functionally redundant." Nucleic Acids Res. **19**: 2817-2824.

Sprague, K.U. (1995). "Transcription of eukaryotic tRNA genes." *tRNA: Structure, Biosynthesis and Fuction*. D. Söll and U.L. RajBhandary. Washington, DC, ASM Press: 31-50.

Stein, A. (1989). "Reconstitution of chromatin from purified components." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170**: 585-603

Stein, A. and M. Mitchell (1988). "Generation of different nucleosome spacing periodicities in vitro: possible origin of cell type specificity." J. Mol. Biol. **203**: 1029-1043.

Stillman, B. (1986). "Chromatin assembly during SV40 DNA replication in vitro." Cell **45**: 555-565.

Stöger, R., P. Kubicka, C.-G. Liu, T. Kafri, A. Razin, et al. (1993). "Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal." Cell **73**: 61-71.

Sun, J.-M., Z. Ali, R. Lurz and A. Ruiz-Carrillo (1990). "Replacement of histone H1 by H5 in vivo does not change the nucleosome repeat length of chromatin but increases stability." EMBO J. **9**: 1651-1658.

Surani, M.A. (1993). "Silence of the genes." Nature **366**: 302-303.

Surani, M.A. (1994). "Genomic imprinting: control of gene expression by epigenetic inheritance." Curr. Opin. Cell Biol. **6**: 390-395.

Sutherland, G.R. and R.I. Richards (1995). "The molecular basis of fragile sites in human chromosomes." Curr. Opin. Genet. Devel. **5**: 323-327.

Takagi, N. and M. Sasaki (1975). "Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse." Nature **256**: 640-642.

Tapping, R.I., D.E. Syroid, P.T. Bilan and J.P. Capone (1993). "The 5' flanking sequence negatively modulates the in vivo expression and in vitro transcription of a human tRNA gene." Nucleic Acids Res. **21**: 4476-4482.

Tatchell, K. and K.E. van Holde (1977). "Reconstitution of chromatin core particles." Biochem. **16**: 5295-5303.

Tate, P.H. and A.P. Bird (1993). "Effects of DNA methylation on DNA-binding proteins and gene expression." Curr. Opin. Genet. Dev. **3**: 226-231.

Tazi, J. and A. Bird (1990). "Alternative chromatin structure at CpG islands." Cell **60**: 909-920.

Thoma, F., T. Koller and A. Klug (1979). "Involvment of histone H1 in the organisation of the nucleosome and of the salt dependant superstructures of chromatin." J. Cell Biol. **83**: 407-427.

Thomas, J.O. (1989). "Chemical radiolabelling of lysines that interact strongly with DNA in chromatin." *Nucleosomes*. P.M. Wassarman and R.D. Kornberg. San Diego CA, Academic Press Inc. **170:** 369-385.

Thomas, J.O., C. Rees and J.T. Finch (1992). "Cooperative binding of the globular domains of histones H1 and H5 to DNA." Nucleic Acids Res. **20**: 187-194.

Toniolo, D., G. Martini, B. Migeon and R. Dono (1988). "Expression of G6PD locus on the human X chromosome is associated with demethylation of three CpG islands within 100 kb of DNA." EMBO J. **7**: 401-406.

Tribioli, C., F. Tamanini, C. Patrosso, L. Milanesi, A. Villa, et al. (1992). "Methylation and sequence analysis around *Eag* I sites: identification of 28 new CpG islands in XQ24-Xq28." Nucleic Acids Res. **20**: 727-733.

Turner, B.M. (1991). "Histone acetylation and control of gene expression." J. Cell Sci. **99**: 13-20.

Turner, B.M. (1993). " Decoding the nucleosome." Cell **75**: 5-8.

Ura, K., J.J. Hayes and A.P. Wolffe (1995). "A positive role for nucleosome mobility in the transcriptional activity of chromatin templates: restriction by linker histones." EMBO J. **14**: 3752-3765.

Vairapandi, M. and N.J. Duker (1993). "Enzymatic removal of 5-methylcytosine from DNA by a human DNA-glycosylase." Nucleic Acids Res. **21**: 5323-5327.

Wang, R.-H., X.-Y. Zhang, R. Khan, Y. Zhou, L.-H. Huang, et al. (1986). "Methylated DNA-binding protein from human placenta recognises specific methylated sites on several prokaryotic DNAs." Nucleic Acids Res. **14**: 9843-9860.

Watt, F. and P.L. Molloy (1988). "Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter." Genes Dev. **2**: 1136-1143.

Weil, P.A., D.S. Luse, J. Segall and R.G. Roeder (1979). "Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA." Cell **18**: 469-484.

Weintraub, H. (1984). "Histone-H1-dependant chromatin superstructures and the suppression of gene activity." Cell **38**: 17-27.

Westaway, S.K. and J. Abelson (1995). "Splicing of tRNA precursors." *tRNA: Structure, Biosynthesis and Function*. D. Söll and U. RajBhandary. Washington, DC, ASM Press: 79-92.

White, R.J. (1994). *RNA polymerase III transcription*. Austin, TX, R.G. Landes Co.

Wiebauer, K., P. Nedderman, M. Hughes and J. Jiricny (1993). "The repair of 5-methylcytosine deamination damage." *DNA Methylation: Molecular Biology and Biological Significance*. J.P. Jost and H.P. Saluz. Basel, Birkhäuser-Verlag: 510-522.

Wigler, M., D. Levy and M. Perucho (1981). "The somatic replication of DNA methylation." Cell **24**: 33-40.

Wilson, G.G. and N.E. Murray (1991). "Restriction and modification systems." Annu. Rev. Genet. **25**: 585-627.

Winston, F. and M. Carlson (1992). "Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection." Trends Genet. **8**: 387-391.

Wolf, S. and B. Migeon (1985). "Clusters of CpG dinucleotides implicated by nuclease hypersensitivity as control elements of housekeeping genes." Nature **314**: 467-469.

Wolffe, A.P. (1989). "Dominant and specific repression of Xenopus oocyte 5S RNA genes and satellite I DNA by histone H1." EMBO J. **8**: 527-537.

Wolffe, A.P. (1991). "Developmental regulation of chromatin structure and function." Trends Cell. Biol. **1**: 61-66.

Wolffe, A.P. (1992). *Chromatin: Structure and Function*. San Diego, CA, Academic Press Inc.

Wolffe, A.P. (1994). "Inheritance of chromatin states." Dev. Genet. **15**: 463-470.

Wolffe, A.P. and C. Schild (1991). "Chromatin assembly." Methods Cell Biol. **36**: 541-559.

Woodcock, D.M., P.J. Crowther and W.P. Diver (1987). "The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide." Biochem. Biophys. Res. Comm. **145**: 888-894.

Workman, J.L., I.C.A. Taylor, R.E. Kingston and R.G. Roeder (1991). "Control of class II gene transcription during in vitro nucleosome assembly." *Functional Orgainisation of the Nucleus: a laboratory guide*. B.A. Hamlako and S.C.R. Elgin. San Diego CA, Academic Press Inc. **35**: 419-447

Wu, J., J.-P. Issa, J. Herman, D.E. Basset Jr, B.D. Nelkin, et al. (1993). "Expression of an exogenous eukaryotic DNA methyltransferase gene induces transformation of NIH 3T3 cells." Proc. Natl. Acad. Sci. USA **90**: 8891-8895.

Wu, J.C. and D.V. Santi (1987). "Kinetic and catalytic mechanism of *Hha* I methyltransferase." J. Biol. Chem. **262**: 4778-4786.

Wu, R.S., H.T. Panusz, C.L. Hatch and W.M. Bonner (1984). "Histones and their modifications." CRC Crit. Rev. Biochem. **20**: 201-263.

Yen, R.-W.C., P.M. Vertino, B.D. Nelkin, J.J. Yu, W. El-Deiry, et al. (1992). "Isolation and characterisation of the cDNA encoding human DNA methyltransferase." Nucleic Acids Res. **20**: 2287-2291.

Yisraeli, J., R.S. Adelstein, D. Melloul, U. Nudel, D. Yaffe, et al. (1986). "Muscle-specific activation of a methylated chimeric actin gene." Cell **46**: 409-416.

Yisraeli, J., D. Frank, A. Razin and H. Cedar (1988). "Effect of in vitro DNA methylation on β-globin gene expression." Proc. Natl. Acad. Sci. USA **85**: 4638-4642.

Zenke, M., T. Grundström, H. Matthes, M. Wintzerith, C. Schatz, et al. (1986). "Multiple sequence motifs are invovled in SV40 enhancer function." EMBO J. **5**: 387-397.

Zhang, X.-Y., C.K. Asiedu, P.C. Supakar, R. Khan, K.C. Ehrlich, et al. (1990). "Binding sites in mammalian genes and viral gene regulatory regions recognised by methylated DNA-binding protein." Nucleic Acids Res. **18**: 6253-6260.

Zhang, X.-Y., K.C. Ehrlich, P.C. Supakar and M. Ehrlich (1989). "A plant DNA-binding protein that recognises 5-methylcytosine residues." Mol. Cell. Biol. **9**: 1351-1356.

Zhang, X.-Y., N. Jabrane-Ferrat, C.K. Asiedu, S. Samac, B.M. Peterlin, et al. (1993). "The Major Histocompatibility Complex class II promoter binding protein RFX (NF-X) is a methylated DNA-binding protein." Mol. Cell. Biol. **13**: 6810-6818.

Zlatanova, J. (1990). "Histone H1 and the regulation of eukaryotic genes." Trends Biochem. Sci. **15**: 273-276.

Zuccotti, M. and M. Monk (1995). "Methylation of the mouse *Xist* gene in sperm and eggs correlates with imprinted *Xist* expression and paternal X-inactivation." Nature Genet. **9**: 316-320.