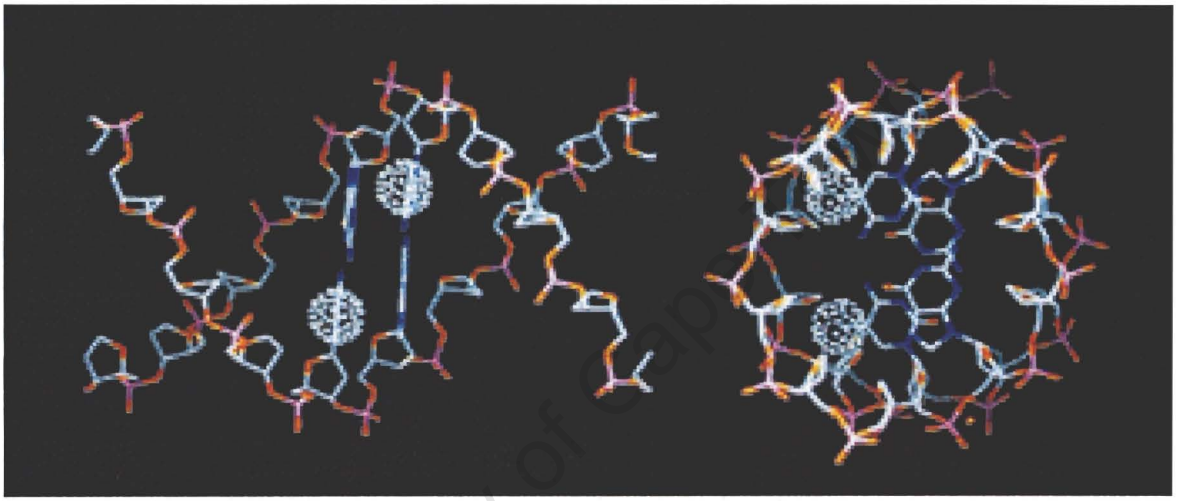
THE ROLE OF DNA METHYLATION IN TRANSCRIPTIONAL REGULATION OF THE HUMAN TYPE 1 ALPHA 2 COLLAGEN (COL1A2) GENE.



BY 'MATLADI N. NDLOVU

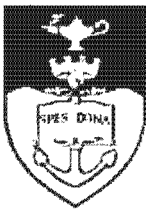
SUPERVISED BY PROF. M.I PARKER

Thesis submitted in fulfillment of the requirement for the degree of

Master of Science

Department of Medical Biochemistry, Faculty of Health Sciences.

UNIVERSITY OF CAPE TOWN.



THE ROLE OF DNA METHYLATION IN TRANSCRIPTIONAL REGULATION OF THE HUMAN TYPE 1 ALPHA 2 COLLAGEN (COL1A2) GENE

by

'MATLADI N. NDLOVU

Thesis submitted in fulfillment of the requirement for the degree of

Master of Science

Department of Medical Biochemistry, Faculty of Health Sciences,
University of Cape Town.

SUPERVISED BY PROF. M.I PARKER

This thesis is dedicated to the loving memory of:

BaFanwell Chirwa

Rakhali Nambitha Ndaba

Joel Sonopo

Limpho Ts`ita

whose unconditional love and support was unwavering until their untimely death. Your memory lives on.

ACKNOWLEDGEMENTS

Not to us, Oh Lord, not to us, but to your name be the glory, because of your steadfast love and faithfulness!

Psalm 115:1.

The writing of this thesis has been an eventful journey, both figuratively and literally. Having started in Cape Town, where most of the work was done under the meticulous supervision of Professor Parker, it passed through Guy's Hospital at the University of London, assisted by Dr. Stöger, and was finally completed at the University of Ghent in Belgium. The financial support for my research came mainly from the Medical Research Council (MRC) of South Africa and Waddell, Baxter and Bensara Scholarships.

Like most journeys, it contained several obstacles, yet I was blessed to meet a number of people who helped me to overcome them. It is therefore with deep gratitude that I would like to thank each one of them for their relentless support.

Thank you very much Prof M.Iqbal Parker for taking me under your wings and opening up a world of scientific opportunities for me. Iqbal, your attentiveness to every detail of the experiments has made me more meticulous as a scientist. Moreover, your style of management in lab 105 enabled open and synergetic discussions. As such I would like to sincerely thank Agatha, Verna, Yolanda, Fundi, Emma, Howard, Fred, Erin, Hajira and Robea for your meaningful discussions and assistance with my experiments; my experiences in lab 105 would never have been the same without all of you.

I would also like to thank the department of Medical Biochemistry at UCT for establishing an environment that encouraged rich discussions and creativity. In particular I would like to thank Prof. Lutz Thilo, Dr. Cynthia Sikakana and Roshan Ebrahim for their invaluable input in this thesis. Thank you Dr Cyril Donninger for your assistance with the HPLC technique.

My gratitude also goes to the DNA Methylation Society, which in addition to being a great resource for information, created a forum to troubleshoot with experts in the field and to gain hands-on experimental assistance. It was through this forum that I was co-supervised by Dr Reinhard Stöger of the University of London. Thank you Reinhard for graciously opening your lab to me and for sharing your time and expertise on the bisulphite technique. Thank you to the technical team and to Mark for assistance with analyzing the sequencing results. I thank you Dilys and Duncan Threshie for offering me a home during my stay in London and for making me part of your family.

I wish to offer special thanks to Prof. Guy Haegeman and all the members of the unit 4 at the University of Gent who have allowed me time to finish off the writing of this thesis even though its subject is somewhat removed from my present research.

This journey would never have been possible without the nurturing, steadfast love and support of my family. Ma le Ntate enkosi kakhulu for creating an environment that allowed me to blossom into who I am today and to follow my dreams. Kea leboha Ts`eli, Nontso, Ntsiki, Ginger le Nkhono and the rest of my family for your prayers and faith in me.

No man is an island. Never have I experienced this more than on this journey. I discovered the true meaning of friendship. Dr Nzeku, Siphiwe, Raesibe, Refiloe, Kathy, Wijou, Deb, Zola, Linda, Aunt Esther, 'Mamolefi, Mark, Natasha, Ann, Mary, Dixie, Daniel, 'Mamoabi, Mots`elisi, Maliehe, Komang, Mowbray Baptist Church, Vuyelwa, Dr Tony & Hazel Mbewu, Ntshekhe, Jenny, Peter and many other friends from around the world; you were all my pillars of support. Over and above the prayers, the laughter and exhortations you consistently sent to lift me up, you also assisted in the proofreading and editing of this manuscript.

I truly believe that each one that crossed my path is a precious gift from God as you have not only helped me with this thesis but have also equipped me for an even greater

journey: life. Your names will be forever engraved in my heart. I pray that your lives may be touched as you have touched mine.

Signed by candidate

'Matladi N. Ndlovu

February 2002.

ABSTRACT

Type I collagen is the most abundant collagen molecule in vertebrate connective tissue and it consists of a heterotrimer of two alpha 1 (COL1A1) and one alpha 2 (COL1A2) chains. Reduced collagen gene expression is almost always correlated with pathological conditions and cellular transformation. Numerous studies have suggested that methylation of the cytosines in CpG dinucleotides is inversely correlated with transcriptional activity and plays a critical role in differential gene expression.

Previous studies have shown that simian virus 40 (SV 40) transformation of the WI-38 human lung fibroblast cell line results in decreased alpha 2 (I) procollagen gene expression. The aim of this study was to investigate the role of DNA methylation in transcriptional regulation of the COL1A2 gene. The main objectives of this study were to:

1. Analyse the methylation status of the 350bp proximal COL1A2 promoter in genomic DNA from collagen-producing and non-collagen producing embryonic lung fibroblasts in order to determine the significance of DNA methylation in the transcriptional regulation of the gene.
2. Investigate the inhibitory effect of *in vitro* methylated CG dinucleotides on COL1A2 promoter activity using COL1A2CAT promoter constructs in transient transfection assays.

Genomic DNA was modified with bisulphite, to convert all cytosines to uracil while 5-methylcytosine remains unmethylated. Treated DNA was amplified using the polymerase chain reaction (PCR). PCR fragments were cloned into TA vectors pGEM®-T Easy and pCR®2.1 TOPO for sequence analysis. Complete deamination of a cytosine to uracil was confirmed by cation exchange high performance liquid chromatography (HPLC) of hydrolysed DNA. The results showed that all the cytosine residues were unchanged in the 350bp proximal promoter of the α2(I) collagen gene in collagen expressing fibroblasts (CT-1). In contrast, 11 of the 17 CpG dinucleotides within the α2(I) collagen promoter were methylated in SV40-transformed non-collagen producing fibroblasts.

In vitro methylation of the CpG dinucleotides in the COL1A2-CAT constructs was achieved with the bacterial methyltransferase M.SssI that mimics the mammalian DNA methyltransferase. The methylated promoter constructs of the COL1A2 gene cloned upstream of the chloroamphenicol acetyl transferase reporter gene, were transiently transfected into cultured human embryonic lung fibroblasts and assayed for promoter activity. The results demonstrated a significant reduction in the promoter activity of methylated constructs compared to their unmethylated controls. This study clearly implicates DNA methylation in repression of the human COL1A2 gene in SV40-transformed WI-38 fibroblasts.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION: REGULATION OF EUKARYOTIC GENE EXPRESSION BY DNA METHYLATION

1.1 REGULATION OF EUKARYOTIC GENE EXPRESSION

Elucidation of the mechanisms controlling gene expression in eukaryotes is important to both the biological and medical sciences since it is the key to understanding the process of development and disease progression. Unlike prokaryotes, multicellular organisms are characterised by the presence of specialised cells with dramatic differences in their structure and function. Although these phenotypic differences arise as a result of differential gene expression, there are some rare cases, where specialised cells have lost their pluripotency (an example being the mammalian lymphocyte and neurone). It is evident therefore that differentiated eukaryotic cells possess a remarkable capacity for the selective expression of specific genes.

Compounding the complexity of differential gene expression in eukaryotes is the presence of many processes common to all cells. These processes include the synthesis of a large number of proteins that are essential for cell survival and maintenance, such as the major structural proteins of the cytoskeleton, chromosomes, ribosomal proteins, endoplasmic reticulum, golgi membrane proteins, to mention but a few. Furthermore, the ability of most eukaryotic cells to alter the expression of genes in response to external stimuli emphasises the complexity in regulation of eukaryotic gene expression.

Initial studies (Ullmann *et al.*, 1965) on the control of genes involved in lactose metabolism in *E.coli* mutants postulated a discrete set of regulated steps by which control of gene expression is achieved (transcription, translation, mRNA localisation and protein activity). Transcription was subsequently identified as the primary level at which control of eukaryotic gene expression is exerted (Hofer and Darnell, 1981). Recently, an interesting dimension, the apparent presence of a significant overlap between sets of proteins involved in gene expression and substantial cross talk between transcriptional and post transcriptional events, has been added to the process of gene expression (reviewed by Ladomery, 1997). These findings indicate that our knowledge of the complex network of regulatory processes in the eukaryotic gene expression is still limited despite the enormous progress during recent years. This review will discuss the factors that are essential for transcriptional regulation, with an emphasis on the *cis*-acting mechanisms associated with DNA methylation (the negative mechanisms regulating transcription).

1.2 TRANSCRIPTIONAL CONTROL

In order to understand and appreciate the role of DNA methylation in gene expression, it is important to first look at the processes involved in transcriptional control. Three key enzymes essential for transcription of eukaryotic genes are RNA polymerase I, II, and III. These enzymes though, cannot initiate transcription in the absence of at least six auxiliary transcription factors. Each enzyme requires a different set of transcription factors, one of which is distinguished by including a subunit common to all three enzymes, the TATA-binding protein (TBP). Transcription initiation involves the assembly of the transcription apparatus and the positioning of the RNA polymerase at the specific site on the promoter. The individual proteins that form this complex aggregate at the proximal promoter in a highly regulated and defined order making transcription initiation the key control point of eukaryotic gene transcription. The processes that orchestrate and facilitate the assembly of general transcription factors (GTFs) are therefore of great significance in transcriptional control. The GTFs, TFIID, TFIIB, TFIIA, TFIIE, TFIIF, TFIIIH, have been shown to bind DNA in a sequence specific manner (Buratowski, 1994, Cosma, *et al.* 1999).

The model for the assembly of the RNA polymerase core enzyme and the GTFs into a preinitiation complex at a typical eukaryotic promoter is illustrated in Figure 1 (Ogbourne and Antalis, 1998). For most genes such as the procollagen gene, the core promoter elements recognised by the GTFs are the TATA-element and the pyrimidine-rich initiator region (Inr)

located between –45 and –25 bp (Smale, *et al.* 1990; Usheva, *et al.* 1992). The TATA-binding protein (TBP) and TBP-associated factors (TAFs) make up the TFIID complex. This is the only GTF capable of sequence specific contact with a core promoter element and subsequently nucleate the formation of the pre-initiation complex (Buratowski, *et al.* 1989; Van Dyke, *et al.* 1989; Roy, *et al.* 1993).

A series of crystal structures of the TBP-TATA element interaction have demonstrated that this complex is conserved throughout the course of evolution (Patikoglou, *et al.* 1999). Studies on the TATA-binding proteins and TBP-associated factors that make up the TFIID complex have revealed that the TAFII250 has histone acetyltransferase activity which may be involved in TFIID accessing transcriptionally repressed chromatin (Mizzen, *et al.* 1996). TFIIA, the second complex recruited to the promoter in the assembly of the GTF, has recently been shown to play a key role in stabilising the interaction between the TFIID and the promoter; in addition to facilitating the binding of TFIIB and the remaining GTFs (Coleman, *et al.* 1999). These reports and others have established that the primary mechanism in the regulation of eukaryotic transcription involves localised chromatin remodelling and modification and the assembly of the initiation complex is but one step in the transcription reaction. It is also becoming apparent that differential gene expression involves variations on the standard mechanisms of transcription (Kuhlman, *et al.* 1999, Mittal, *et al.* 1999).

Early studies proposed that only positively acting transcription factors were involved in transcriptional control, but it is now apparent that negative regulation of transcription is equally significant (see review by Goodbourn, 1990). For most, if not all genes, the preinitiation complex on the core promoter will ensure their constitutive expression at a minimal rate, unless switched off by negative regulatory proteins (silencers) or increased by positive regulatory proteins (activators and enhancers). These negative and positive regulators of transcription play a critical role in ensuring the constitutive expression of housekeeping genes, modulating gene expression in response to changes in the micro-environment of the cell and for complete inactivation of gene expression in specific tissues. Due to their earlier discovery, positive regulatory factors are much better understood in comparison to repressors.

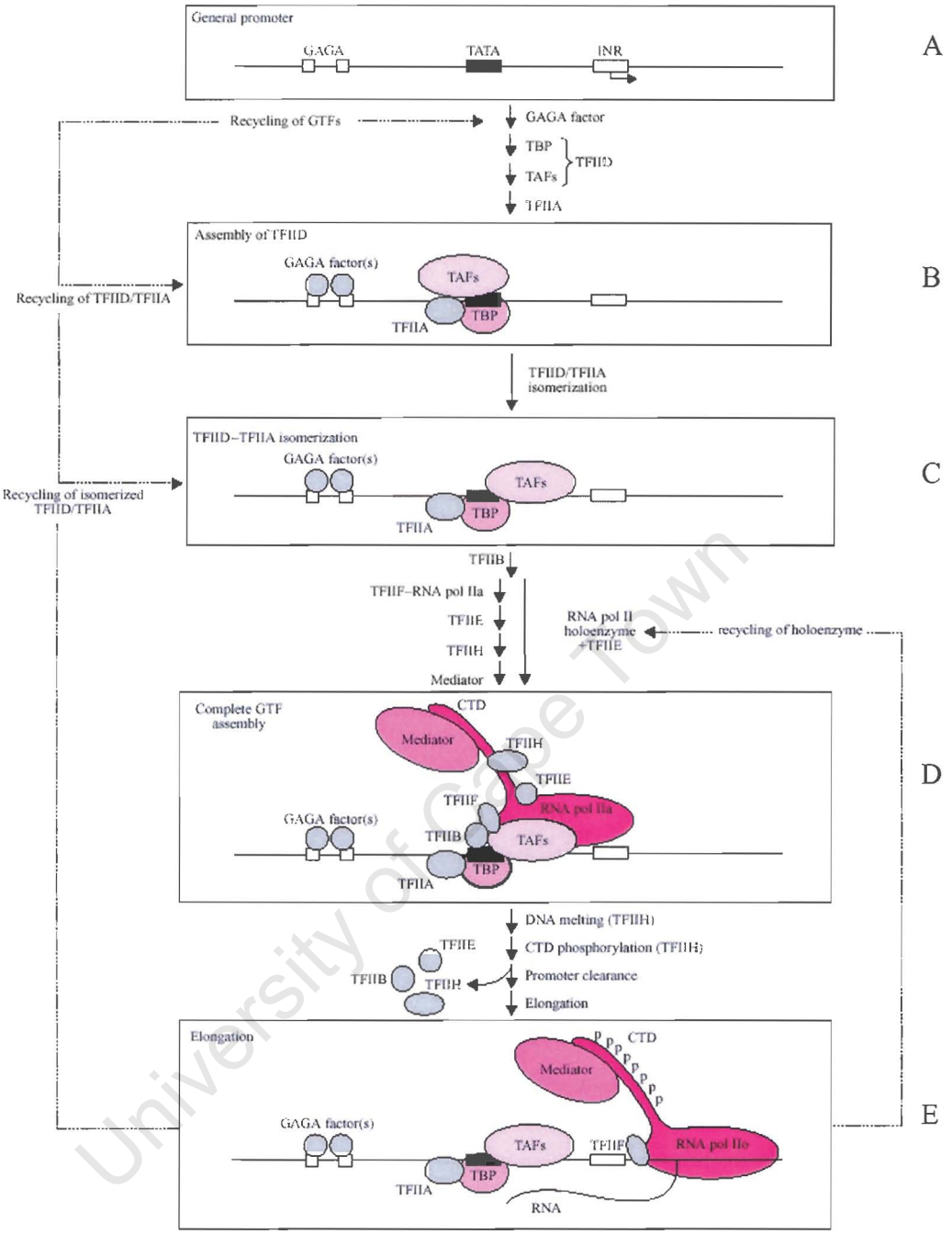


Figure 1.1 Model for assembly of the general transcription factors into a preinitiation complex on a typical eukaryotic promoter. **(A)** Depicts a general eukaryotic proximal promoter containing a TATA box (TATA), an initiator motif (INR) that overlaps with the transcription start site (arrow), and two GAGA elements (clear boxes). **(B)** The recently identified GAGA factors (Weber, *et al.* 1997) are proposed to bind one or more of the GAGA elements thus priming the chromatin to enable the interaction of the TATA-binding protein (TBP) (pink). Several TBP-associated factors-TAFs- (pink) then bind to TBP and making up the TFIID complex, binding of TFIIA then follows. (**C**) Isomerization of the TFIIA-TFIID complex must occur to allow further stepwise binding of TFIIB, RNA pol IIa (red)-TFIIF dimer, TFIIE, TFIIH, and the mediator (dark pink) complex (**D**). The frequency of initiation can be increased by interaction with RNA pol II holoenzyme. In order to allow elongation to take place (**E**), DNA melting and CTD phosphorylation, which together stimulate promoter clearance, must occur. The recycling of various parts of the spent GTFs (dotted arrows) provides mechanisms to further increase the rate of re-initiation (adapted from Ogbourne and Antalis 1998).

1.3 DNA METHYLATION IN EUKARYOTES

Most eukaryotic genomes are characterised by the methylation of the DNA at the 5-carbon of cytosine residues residues in CpG dinucleotides. Wyatt (1950) discovered this modification of DNA in the eukaryotic genome exactly 5 decades ago. However, it is only recently that there has been a rapid expansion in our knowledge of the molecular mechanisms of DNA methylation in the regulation of eukaryotic gene expression. The role of DNA methylation remained controversial for several decades due to the absence of reliable detection techniques. The tide changed somewhat in the 1970's following the discovery of methylation-sensitive restriction enzymes that enabled mapping of 5-methylcytosine in animal DNA. Using these methylation sensitive enzymes, the first insight into the systematic distribution of 5mC in eukaryotes and their heritable traits were shown. Predictions that gene expression during mammalian development could well be regulated by programmed methylation and demethylation of genes were also made (Holliday and Pugh 1975; Riggs 1975; Bird 1978; Bird and Southern 1978; Bird *et al.* 1979). The screening of a large number of vertebrate genes for the presence of CpG islands and their position relative to the transcription unit of the associated gene was significant in unlocking the mechanism(s) of DNA methylation in eukaryotes. Disputes on the significance of DNA methylation in normal development and tissue specific gene expression, sparked by the fact that DNA methylation is absent in some eukaryotes such as *Drosophila melanogaster* (Urieli-Shoval *et al.*, 1982) and *Saccharomyces cerevisiae* (Proffitt *et al.*, 1984) however, slowed down progress somewhat.

Interest in DNA methylation as a regulatory element during normal development was rekindled by studies that reported the lethality of reduced DNA methyltransferase activity in mouse embryonic development (Li *et al.*, 1992; Li *et al.*, 1993). Tate (1996) has also reported recently that the ubiquitous methyl-C-binding protein 2 (MeCP2) is essential for mouse embryonic development.

Moreover, a large body of evidence correlating altered DNA methylation patterns with carcinogenesis has accumulated (reviewed by Baylin, 1992, Counts and Goodman, 1995; Laird and Jaenisch, 1994 & 1996; Jones, 1996, Schmutte and Jones, 1998; Bergman and Mostoslavsky, 1998; Bird and Wolffe, 1999; Jones, 1999). Over the past 18 months there has been an explosion in the data on new advances in the mechanisms and identification of crucial proteins in the DNA methylation-mediated transcriptional repression.

It is becoming clear that the mechanisms through which 5methylcytosine (5mC) affects oncogenesis are complex, but recent reports have shown that they are centred on both the genetic and epigenetic properties of methylated DNA. Many genetic mutations can be attributed to the deamination of 5mC residues and the high DNA methyltransferase activity levels observed in cancer cells (Laird and Jaenisch, 1994; Laird and Jaenisch, 1996). The hydrolytic deamination of mC→ T is difficult to detect and to repair by the DNA repair enzymes since T (thymine) is a naturally occurring base in DNA. Consequently these transition mutations have been found at a 12-fold higher frequency at CpG sites (Cooper and Krawczak, 1990). The p53 tumour suppressor gene, which has been documented as the most frequently mutated gene in human tumours, displays the most CpG → TpG transition mutations (Hainaut, *et al.* 1997; Hainaut and Hollstein 2000).

Presently, a number of mechanisms have been suggested for these mutation hotspots (reviewed in Gonzalgo and Jones, 1997; Schmutte and Jones, 1998). DNA methyltransferase 1 (DNMT1) is the most abundant methyltransferase in mammalian cells (Yoder *et al.*, 1997) and was the first methyltransferase discovered (Bestor *et al.*, 1988) and as such it has enjoyed a much more comprehensive analysis than the recently discovered methyltransferases. Enzyme-mediated deamination of cytosines to thymine through covalent interaction of the DNMT1 with its target cytosine were first reported in prokaryotes (Shen *et al.*, 1992; Shen *et al.*, 1995) and subsequently shown to create mutational hotspots in eukaryotes (Yebra and Bhagwat, 1995). High levels of enzyme occur in human tumours. In colorectal tumours both 15 to 200 fold-increase (el-Deiry *et al.*, 1991, Schmutte *et al.*, 1996) and in tumour cells in culture (Kautiainen and Jones 1985; Kautiainen and Jones 1986, Robertson *et al.*, 2000). Moreover, DNMT1 expression is upregulated by various oncogenic signals including activated ras (MacLeod *et al.*, 1995; Rouleau *et al.*, 1995), fos (Bakin and Curran, 1999) and SV40 T-antigen (Slack *et al.*, 1999).

Overexpression of the recently identified de novo methyltransferases, DNMT3a and 3b, has been reported (Okano *et al.*, 1998, Robertson *et al.*, 1999; Xie *et al.*, 1999) and the mRNA levels of all three DNMTs (DNMT1, 3a and 3b) have been shown to differ during the cell cycle in both normal and tumour cells (Szyf *et al.*, 1985; Szyf *et al.*, 1991; Robertson *et al.*, 2000). DNMT3b and DNMT1 levels were significantly decreased in G_0/G_1 compared to DNMT3a, and the ratio of DNMT3a to DNMT1 at G_1 was increased 2-3 fold in tumour cells

compared to normal fibroblasts. These results suggest that deregulation of these enzymes that control DNA methylation may lead to tumourigenesis.

The epigenetic effects of 5mC on carcinogenesis have been reported in a number of genes including oncogenes, tumour suppressor genes, and imprinted genes, some of which are indicated in Table 1. Studies on these genes and others have conclusively demonstrated that methylated cytosine residues in the promoter of a gene play a crucial role in suppressing gene activity. A reciprocal relationship between the density of methylated cytosine residues and the transcriptional activity of a gene has been widely observed and documented for a large number of genes (Eden and Cedar, 1994, Herman *et al.*, 1995; Nagatake *et al.*, 1996; Gonzalgo and Jones, 1997; Stirzaker *et al.*, 1997; Gonzalgo *et al.*, 1998; Jones, 1999).

The mechanisms through which DNA methylation mediate transcriptional inactivation are still not completely understood, but recent advances indicate modulation of chromatin structure to be the focal point (Bestor, 1998; Razin, 1998; Wolffe, 1998). Although Keshet *et al.*, (1986) proposed this as the major pathway for transcriptional repression by DNA methylation, evidence of the recruitment of histone deacetylases by methyl-C-binding protein 2 (MeCP2) together with a multiprotein repression complex have been presented only fairly recently (Heinzel *et al.*, 1997; Nan *et al.*, 1998). The discovery of the two new mammalian cytosine DNA methyltransferases, DNMT3a and 3b (Okano *et al.*, 1998; Yoder and Bestor 1998), and four new methyl-CpG-binding proteins (Cross *et al.*, 1997; Hendrich and Bird 1998) in the past few years has shed further light on the role of DNA methylation on transcriptional silencing in eukaryotes. Studies by Cameron *et al.*, (1999) on the synergistic effect of DNA methylation and histone deacetylation on gene silencing have proposed that 5-methylcyosine is the critical factor in the stable maintenance of transcriptional repression of densely methylated endogenous promoters.

Much of the focus of research amongst DNA methylation investigators currently is to determine whether DNA methylation is responsible for gene inactivation or if it is a consequence of gene inactivation. There is strong support for the idea that DNA methylation is a consequence of transcriptional inactivation and that cytosine methylation is more restricted to highly specialised processes such as allele-specific gene inactivation and transcriptional silencing of transposons and viral parasites (Walsh and Bestor, 1999). Martienssen (1998) attempted to reconcile such views by comparing the role of DNA

methylation and transposable elements in plant gene regulation with that in mammals. He argued that the ability to suppress transposons is not restricted to organisms with methylated DNA since such suppression is also observed in *Drosophila* and yeast that do not contain methylated cytosines. In these organisms, it is the modulation of chromatin structure rather than methylation that mediate transposon suppression.

Table 1: Genes with altered DNA methylation patterns in human tumours (Adapted from (Schmutte and Jones, 1998)

Gene	Reference
A. **Tumour suppressor genes**	
pRB	(Ohtani-Fujita et la., 1993)
p16	(Gonzalez-Zulueta *et al.*, 1995; Herman *et al.*, 1995; Merlo *et al.*, 1995; Swafford *et al.*, 1997)
p15	(Hernan *et al.*, 1996)
B. Imprinted genes	
H19/IGF2	(Bartolomei *et al.*, 1991; Elson and Bartolomei, 1997; Forné *et al.*, 1997)
IGF2r	(Wutz *et al.*, 1997)
VHL	(Herman *et al.*, 1994)
$p57^{KIP2}$	(Hatada *et al.*, 1996)
WT1	(Jinno *et al.*, 1994; Mitsuya *et al.*, 1997)
HTR2	(Kato *et al.*, 1996)
Ube3a	(Albrecht *et al.*, 1997; Kishino *et al.*, 1997)
SNRPN	(Glenn *et al.*, 1996)
ZNF-127	(Glenn *et al.*, 1997)
Xist	(Beard *et al.*, 1996)

Recent studies supporting the hypothesis that DNA methylation is responsible for gene inactivation have uncovered that proteins involved in DNA methylation are directly responsible for various human genetic disorders. The *de novo* DNA methyltransferase DNMT3b, located on the long arm of chromosome 20, has been shown to be the candidate gene responsible for the ICF (immunodeficiency centromere instability and facial anomalies) syndrome (Robertson *et al.*, 1999). Nine different mutations in the DNMT3b gene, including

one that produces a protein with severely impaired *de novo* methyltransferase activity, have been detected in eight different ICF families (Hansen *et al.*, 1999; Okano *et al.*, 1999; Xu *et al.*, 1999). Furthermore, Amir *et al.*, (1999) and Wan *et al.*, (1999) have shown that spontaneous and familial mutations in the gene encoding the transcriptional repressor protein, MeCP2 affect neuronal development in humans (Rett syndrome). Hung *et al.*, (1999) have recently presented immunoprecipitation and immunocytochemistry data showing the presence of two *Drosophila* proteins, DmMTR1 and DmMT2, that exhibit similar structural and cytological characteristics to those of the mammalian DNMT1 and DNMT2 respectively, yet their functional role has yet to be determined since Drosophila is devoid of any cytosine methylation in its genome. DmMTR1 was shown to interact with proliferating cellular nuclear antigen (PCNA) *in vivo*, similar to the mammalian DNMT1, suggesting its involvement in both DNA replication and repair. These authors also postulated that DmMTR1 might play a role in chromosome condensation similar to mammalian 5mCpG/MeCP complexes. Interestingly, the DNMT1 homolog expressed in *Drosophila,* DmMTR1, is also located in the cytoplasm during interphase as is DNMT1 in the mouse blastocyst (Mertineit *et al.*, 1998; Cardoso and Leonhardt 1999).

The globin gene family was amongst the first differentiated genes reported to be regulated by DNA promoter methylation (Busslinger *et al.*, 1983). To date, a large number of genes that are regulated by DNA methylation have been reported, including the procollagen genes. This study will focus on the role of DNA methylation in regulation of the α2 (I) procollagen gene. Type I collagen is the major fibrillar constituent of the connective tissue and is abundantly expressed in bones, tendons, skin, ligaments, cornea and teeth. Stringent control mechanisms are necessary for this tissue- and cell- specific expression of these genes and for the co-ordinate expression of α1 and α2 chains in a 2:1 ratio in order to form the triple helical fibrils characteristic of type I collagen. Previous studies have implicated DNA methylation as playing a major role in the regulation of several procollagen genes (Rhodes *et al.*, 1994; Kopp *et al.*, 1997). The α1(I) gene of type I collagen has been studied extensively in order to determine the regulatory elements involved in controlling gene activity (Rossert *et al.*, 1996). DNA methylation was shown to regulate expression of the α1(I) chain independent of methyl-CpG- binding proteins (Kass *et al.*, 1997). Recently DNA methylation has been shown to play a key role in transcriptional regulation of the mouse α2(I) collagen gene (Sengupta and Smith 1998; Sengupta *et al.*, 1999).

1.4 DNA METHYLATION AND GENE EXPRESSION

There is overwhelming evidence for an inverse relationship between methylation of CpG islands found in gene promoters and transcriptional inactivity (see review Jones 1999). Mouse knockout experiments have conclusively shown that DNA methylation is essential for mammalian development (Li *et al.* 1992). Altered DNA methylation patterns in the promoters of tumour suppressor genes and DNA repair genes are among the most common markers of cancer development. The most widely recognised role of DNA methylation in gene expression is gene silencing. However, the presence of CpG islands in exons and downstream of transcription initiation sites has been associated with gene expression and this has therefore introduced a paradox in the role of DNA methylation in gene expression (Jones 1999). A number of studies on the relationship between transcription and DNA methylation are therefore underway in an attempt to resolve this paradox.

1.4.1 DISTRIBUTION OF CG DINUCLEOTIDES

The CpG dinucleotide has been progressively depleted from the vertebrate genome, over evolutionary time, to less than 10% of the predicted frequency and this is correlated with an unexpected increase in TpG and CpA dinucleotides (Bird, 1980; Adams *et al.*, 1987; Gardiner-Garden and Frommer, 1987). Recent reports, however, have indicated the presence of methylated cytosines at dinucleotides sequences other than CG, these include CC, CT, CA and CCG (Parker and Gevers 1984; Parker *et al.* 1986, Woodcock, *et al.* 1987; Woodcock, *et al.* 1988; Toth, *et al.* 1990; Tasheva and Roufa 1994; Woodcock, *et al.* 1997; Woodcock, *et al.* 1999). The plant genome on the other hand, has long been shown to have methylation in CNG trinucleotides and in bacterial and some lower eukaryotes in addition to 5mC, methylation of adenine is also observed (Hattman, *et al.* 1978; Razin and Szyf 1984; Van Etten, *et al.* 1985; Theiss, *et al.* 1987; Meyer, *et al.* 1994; Pelissier, *et al.* 1999). Kay, *et al.* (1994) have also reported the presence of N6-adenine methylation in mammals and Tasheva and Roufa (1994) have suggested that these unusual methylation sites could play a major role in DNA replication as these have been shown to be concentrated at origins of replication. Analysis of eukaryotic DNA methylation at sites other than CpG dinucleotides is beyond the scope of this thesis, but their role in gene expression is currently being investigated (Woodcock, *et al.* 1999).

DNA methylation is a post-replicative modification of DNA. A methyl group is transferred from the methyl donor S-adenosine methionine (SAM) to a potential methylation site in the newly synthesised DNA strand by maintenance methyltransferase (DNMT1) following replication of a parental strand containing 5mC (Gruenbaum, *et al.* 1981). Only about 5% of cytosine residues or more specifically less than 1% of all DNA bases in vertebrates, on average, are methylated (Ehrlich and Wang, 1981). There are, however, regions of DNA containing a high G+C content and a normal density of non-methylated CpG dinucleotides, termed CpG islands (Bird, 1986; Gardiner-Garden and Frommer, 1987; Antequera and Bird, 1993). These CpG islands are at least 200bp in length and could be as long as 1kb. Most CpG islands are associated with the transcription start site of the vast majority of human genes (Antequera and Bird, 1993) and they have been shown to be rigorously protected from methylation in the germline. Studies on the Epstein-Barr virus latency C promoter (Robertson et al. 1995) have shown that methylation of a single CpG site within the promoter can strongly suppress its transcriptional activity. Hence the association of most human cancers with the methylation of CpG islands.

Recently, it has been reported that CpG islands are able to escape methylation in the germline due to their colocalization with DNA replication origins in mammals (Tommasi and Pfeifer 1995; Waltz, *et al.* 1996; Delgado, *et al.* 1998; Di Matteo, *et al.* 1998). A number of factors that bind to GC-rich DNA have been implicated in the general maintenance of methylation-free CpG islands, ultimately ensuring correct expression of the gene. Sp1-like elements for example, play a key role in protecting the CpG islands in the adenine phosphoribosyltransferase gene from *de novo* methylation (Marin, *et al.* 1997; Kudo 1998). New speculations on the origins and maintenance of methylation-free CpG dinucleotides are based on the recent reports on the relationship between promoter CpG islands and replication origins. It is proposed that CpG islands are a result of active promoters which exclude methylation by recruiting proteins that initiate DNA replication during early embryonic development and their methylation status would then be imprinted and maintained by the maintenance methyltransferase (reviewed in Antequera and Bird 1999).

Furthermore, Sp1 has been shown to regulate the expression of a methyl-CpG-binding protein, MeCP2, which is thought to repress transcription *in vivo* via a highly basic transcriptional repression domain (Tate and Bird, 1993). Expression of the MeCP2 gene, and regulation thereof by Sp1, seems to be essential for normal embryonic development, but not

for growth and differentiation (Ng and Bird, 1999). There are approximately 50 000 CpG islands in the haploid human genome (Jones and Laird, 1999), however not all of them show an inverse relationship characteristic of methylation-induced genes. All housekeeping genes and widely expressed genes such as c-Jun and Ha-Ras have CpG islands 5′ to the transcription start site and extending into the one or more exons while CpG islands of most tissue-specific genes are not biased toward the 5′ end of the transcription unit (Gardiner-Garden and Frommer 1987; Larsen, *et al.* 1992). It is proposed that DNA methylation could exert its repressive influence on housekeeping genes through methylation of the coding region (Stein, *et al.* 1983).

Imprinted genes have provided the best evidence for the causal link between DNA methylation and transcription. The human small nuclear ribonucleoprotein polypeptide N (SNRPN) (Glenn, *et al.* 1996), the human insulin-like growth factor II (IGF2) (Issa, *et al.* 1996) and the mouse insulin-like growth factor 2 receptor (Igf2r) (Wutz, *et al.* 1997) have all shown a clear correlation between allele-restricted RNA expression and allele-specific DNA methylation. Jaenisch, *et al.* (1998) have also highlighted the significance of CpG methylation in genes subjected to X inactivation. Although there is overwhelming evidence to date on the important role of DNA methylation in development and tissue-specific expression of genes, there are several authors who have shown that the evidence is not conclusive enough (Walsh and Bestor, 1999). These authors suggest that for methylation to be central in regulating tissue-specific genes, the promoters of these genes should be densely methylated in non-expressing tissues and expression should always be activated by demethylation. Using methylation sensitive enzymes, they examined the methylation status of the 5′ regions of a group of tissue-specific genes and could not correlate their expression with their methylation status in tissues of foetal and newborn mice. A study of the expression of the tissue specific genes, such as α1 type 1 procollagen, skeletal muscle α-actin and β-globin, in DNMT1 mutant mouse embryos of 10.5 days postcoitum showed that demethylation in these embryos caused biallelic expression of imprinted genes and activation of endogenous retroviruses, however no precocious expression was observed for the tissue-specific genes (Li, *et al.* 1992).

Methylation changes observed in tissue-specific genes during transformation therefore would be as a result of and not the cause of transcriptional activation. A large body of evidence accumulated in recent years however, has provided insight into the mechanisms of

transcriptional regulation by DNA methylation and demonstrated that DNA methylation is tightly associated with chromatin structural modification (reviewed by Ng and Bird 1999).

1.4.2 ESTABLISHMENT OF DNA METHYLATION PATTERNS

Early experiments in mouse embryos have demonstrated that DNA methylation patterns change considerably during embryogenesis (Razin and Cedar, 1993). Most CpG sites are methylated in the mature oocyte and sperm DNA during oogenesis and spermatogenesis and are involved in controlling expression of the imprinted genes (Sanford, *et al.* 1987; Silva and White 1988; Kafri, *et al.* 1992; Brandeis, *et al.* 1993; Constancia, *et al.* 1998). After fertilization, global demethylation occurs prior to the 16 cell morula stage and remains through blastulation. An extensive wave of *de novo* methylation occurs at the time of implantation, with the exception of housekeeping genes, which are supposedly protected by Sp1 protein. After gastrulation, developmental gene activation occurs concomitantly with selective demethylation at tissue-specific genes in the different somatic lineages (Turker and Bestor, 1997). The adult genome therefore exhibits a bimodal pattern of genome methylation; a high frequency of unmethylated CpG-rich sites associated with promoters and methylated CpG dinucleotides within the coding region and body of genes (Bird, 1986). The DNA methyltransferase enzyme (Dnmt1) then faithfully propagates the established methylation patterns after every cycle of replication. The mechanisms employed by mammalian cells to establish methylation patterns are currently unknown. However, the newly identified mammalian DNA methyltransferases, DNMT3a and 3b, that are responsible for the wave of *de novo* methylation observed during embryogenesis and essential for embryonic development will undoubtedly provide new insights (Jones *et al.*, 1998; Okano *et al.*, 1998b; Cameron *et al.*, 1999; Okano *et al.*, 1999). *Cis*- and *trans*- acting factors have also been proposed as role players in the formation of mammalian DNA methylation patterns. This was based on observations by various investigators of specific methylation pattern in all somatic tissues in the region upstream of the human tumour necrosis factor α and β genes and mouse *aprt* gene (reviewed by (Turker and Bestor, 1997). Experimental evidence to substantiate these postulations is still lacking, however.

1.4.3 PROTEINS INVOLVED IN DNA METHYLATION

1.4.3.1 DNA-(Cytosine-5) Methyltransferases

Until recently, our knowledge on the DNA methyltransferase enzymes was very limited. Riggs (1975) and Holliday and Pugh (1975) were the first to predict the existence of a maintenance methylase and proposed that this enzyme preferentially methylates hemi-methylated DNA over unmethylated DNA. The studies by Wigler, *et al.* (1981) and Stein, *et al.* (1982) supported this hypothesis by showing clonal inheritance of methylation patterns in mammalian cells. Bestor, *et al.* (1988) recently cloned and sequenced the first DNA methyltransferase gene (Dnmt1) from mouse. Dnmt1 consists of two important domains, a regulatory amino terminal and a catalytic COOH-terminal domain. X-ray crystallographic studies have revealed a striking structural similarity of the C-terminal catalytic domain to that of bacterial N6-methyltransferase (Malone, *et al.* 1995). The N-terminal domain, however, is absent in bacterial enzymes and this could explain why the mammalian DNA methyltransferase is able to discriminate between unmethylated and hemimethylated DNA unlike its bacterial counterpart. In vitro studies using purified Dnmt1 protein confirmed that Dnmt1 methylates DNA with hemi-methylated CpG dinucleotides much more efficiently than unmethylated DNA (Bestor, 1992). A detailed crystal structure of the M. HhaI methyltransferase active site during the transfer of the methyl group from the donor, S-adenosylmethionine, was subsequently determined (Klimasauskas, *et al.* 1994). M.HhaI was shown to distort the DNA helix by flipping the target cytosine base completely out of helix, into the cavity of the enzyme where catalysis occurs. The possibility that maintenance methylation is coupled to DNA replication has been proposed following studies that identified amino acids and peptide sequences involved in targeting the Dnmt1 protein to DNA replication foci (Leonhardt *et al.*, 1992; Liu *et al.*, 1998).

Two isoforms of the Dnmt 1 have subsequently been described; the ~190kD somatic and ~170kD oocyte form (Gaudet, *et al.* 1998) It is proposed that an alternative promoter and a downstream translation start site produces the shorter oocyte isoform. During embryonic development, when methylation patterns are being established, the Dnmt1 levels are very high in the oocyte and preimplantation embryos and immunoflourescence studies located the Dnmt1 in the cytoplasm in these stages (Carlson, *et al.* 1992). The recent work of Cardoso and Leonhardt *et al.* (1999) supports these findings by mapping a region in the regulatory domain of the Dnmt1 that is responsible for cytoplasmic localisation during early

development. They postulated a mechanism involving interaction with factors expressed in early developmental stages that retains the Dnmt1 in the cytoplasm. The cytoplamic retention of Dnmt1 is therefore closely correlated with the observed global demethylation during embryogenesis. The phenotype of Dnmt1 deficient embryos reported by (Li, *et al.* 1992; Lei, *et al.* 1996) also fits with this concept of diminished Dnmt1 activity during pre-implantation. Murine embryos homozygous for the Dnmt1 mutation are normal during pre-implantation, but abnormal development and subsequent death is observed at 12.5 days post-coitus. These studies emphasize the importance of Dnmt1 in embryonic development. Furthermore, the unexpected observation of low levels of DNA methylation for more than 20 cell divisions in null mutant embryonic stem cells provided the first evidence for the existence of a *de novo* mammalian DNA methyltransferase distinct from Dnmt1 (Lei, *et al.* 1996). It has now been established that Dnmt1 is a maintenance methylase.

These findings triggered the search in expressed-sequence-tag (EST) databases for genes that could encode protein motifs common to DNA (cytosine5-) methyltransferases. The first methyltransferase to be reported was Dnmt2 (Yoder and Bestor, 1998). Subsequently, two other mammalian DNA (cytosine-5-) methyltransferases were identified and cloned, Dnmt3a and Dnmt3b (Okano, *et al.* 1998). Although all three genes were reported to be candidates for the de novo methyltransferase, only Dnmt3a and 3b have been reported to function as de novo methyltransferases (Okano, *et al.* 1998; Okano, *et al.* 1999). DNA methyltransferase activity has not yet been demonstrated for Dnmt2. A series of experiments including deletion of the gene in ES cells did not show any reduction in the amount of global DNA methylation or in de novo methylation (Okano, *et al.* 1998; Yoder and Bestor 1998). Interestingly, the methyltransferase-like proteins found in Drosophila (DmMT2) (Hung, *et al.* 1999) and fission yeast (pmt1) (Wilkinson, *et al.* 1995) exhibit a high sequence homology to the mammalian Dnmt2. As mentioned earlier, both Drosophila melanogaster and fission yeast do not have detectable amounts of cytosine methylation in their genomes. The function of these Dnmt2-like methyltransferases therefore is currently unknown. On the contrary, Dnmt3a and 3b enzymes have been shown to methylate unmethylated DNA *de novo* both in vivo and in vitro and do not show preference for unmethylated or hemi-methylated DNA (Okano, *et al.* 1998; Lyko, *et al.* 1999; Okano, *et al.* 1999). Although the subcellular localisation of these proteins has not been characterised, both genes are highly expressed in undifferentiated ES cells, but show low levels of expression in differentiated embryoid bodies that mimics that in somatic cells. De novo methylase activity is not detected in ES cells with double knockouts of both

Dnmt3a and Dnmt3b but only when one gene was mutated, that is Dnmt3a$^{-/-}$ or Dnmt3b$^{-/-}$ ES cells still exhibited de novo methylase activity. Dnmt3a$^{-/-}$ or Dnmt3b$^{-/-}$ in mouse embryos exhibit different phenotypes, with a specific role for Dnmt3b in methylation of centromeric minor satellite repeat. These results, therefore, are indicative of distinct functions of these enzymes during development (Okano, *et al.* 1999). The pre*cis*e role of Dnmt3b in methylation of centromeric satellite repeats was further substantiated by the association of mutations in the human Dnmt3b gene with the immunodeficiency, centromeric instability and facial anomalies (ICF) syndrome (Hansen, *et al.* 1999; Okano, *et al.* 1999; Xu, *et al.* 1999).

1.4.3.2 DNA Demethylase

The mechanism involved in the selective demethylation of tissue-specific genes in the cell type in which they are expressed still remains elusive. The direct removal of methyl groups from 5-methylcytosines has been presumed energetically unlikely until recently. Enzymatic demethylation and replication-dependent demethylation have been described extensively for a number of genes (Razin, *et al.* 1986; Saluz, *et al.* 1986; Paroush, *et al.* 1990). Although the mechanism of demethylation is still unknown, results from a new in vitro system that monitors demethylation in whole cell extracts have postulated that enzymatic demethylation involves the exchange of DNA nucleotides, is partly mediated by RNA (Weiss, *et al.* 1996). Although further purification of the demethylase activity revealed that it is not sensitive to RNase treatment, involvement of RNA could not be completely discarded (Swisher, *et al.* 1998). Fremont, *et al.* (1997) and Jost, *et al.* (1997) have proposed that the active demethylation reaction could involve RNA in association with proteins such as 5-methylcytosine DNA glycosylase. They showed that a highly purified glycosylase removes the 5-methylcytosine in conjunction with RNA, which recognise methylated cytosine residues in the DNA. Other reports suggest that the presence of specific *cis* upstream regulatory elements provides signals that induce demethylation as in the case in immunoglobulin κ (Igκ) light chain genes (Bergman and Mostoslavsky, 1998). The matrix attachment region (MAR) has been shown to play a critical role in directing the demethylation machinery to the Igκ gene.

Demethylation and gene activation are closely associated during development in tissue-specific genes. Chromatin remodelling through a multi-protein complex that includes MeCP2 and histone deacetylases (HDACs) is a prerequisite for gene activity and the question of

which process occurs first has still not been resolved. Studies on the Ig genes (Kelley, *et al.* 1988; Burger and Radbruch 1992) have suggested that changes in chromatin precede demethylation. It is possible that removal of MeCPs may destabilise the transcriptionally repressed chromatin and induce hypomethylation, resulting in accessibility by *trans*-acting factors to genes and consequently transcriptional activation (Bergman and Mostoslavsky, 1998). The first demonstration for an enzyme that catalyses demethylation by removal of methyl groups from 5-methyl-cytosine residues in DNA came from studies by Bhattacharya, *et al.* (1999) and Cervoni, *et al.* (1999). Following the identification and characterisation of three new mammalian methyl-CpG-binding proteins Hendrich and Bird (1998), Bhattacharya, *et al.* (1999) used the methyl-CpG binding domain (MBD) shared by these proteins to search EST database for sequences homologous to this domain on the assumption that demethylases should be able to recognise methylated CpG dinucleotides. The protein encoded by one of the complementary DNAs (cDNA) revealed explicit demethylase activity in vitro and an identical demethylase that converted ^{32}P-labeled 5-methylcytosine nucleotides in DNA to cytosine without disrupting the DNA backbone was also identified and cloned from A549 non-small-cell human lung carcinoma cells. The cDNA was found to be the same as that encoding methylated-DNA-binding protein 2 (MBD2) previously identified and characterised by Hendrich and Bird (1998).

The discovery, that MBD2 exhibited demethylase activity, has not been endorsed in more recent studies (Ng, *et al.* 1999; Wade, *et al.* 1999; Zhang, *et al.* 1999) that demonstrated that MBD2 and MBD3, which are 71% identical at the amino acid level, associate with transcriptional repressor complexes containing histone deactylase activity resembling the fully characterised MeCP2 (Nan, *et al.* 1998), however all three proteins interacted with different complexes, indicative of their different functions in vivo. To solve this apparent disagreement, it needs to be verified that the cDNA cloned by (Bhattacharya, *et al.* 1999) and found to be MDB2, does indeed encode the activity purified from A549 cells. Figure 1.3 depicts the proposed mechanism for enzymatic demethylation of 5-methylctytosine revealed by the two-dimensional thin-layer chromatography studies in Bhattacharya, *et al.* (1999) and Cedar and Verdine 1999).

1.4.4 DNA METHYLATION PATTERNS IN TRANSFORMED CELLS

Tumour and transformed cells are characterized by both changes in 5-methylcytosine levels and modified DNA methylation patterns within genes. Both widespread DNA hypomethylation and regional hypermethylation, as well as increased Dnmt1 levels are associated with cellular transformation. Recently, overexpression of DNMT3a and 3b has also been reported (Robertson, *et al.* 1999; Xie, *et al.* 1999). The fact that these changes are prevalent in cancer cell is indicative of the significant role they play and mechanisms employed by methylation in carcinogenesis. Deamination-driven mutations and altered CpG island methylation are the two major mechanisms that account for loss of gene function associated with DNA methylation in transformed cells. The possibility that these two processes are intertwined such that one can predispose to the other in driving tumour progression is becoming more evident (Jones 1996; Baylin, *et al.* 1998) and has led to the revision of the Knudson's two-hit hypothesis by Jones and Laird (1999) in (Figure 1.4).

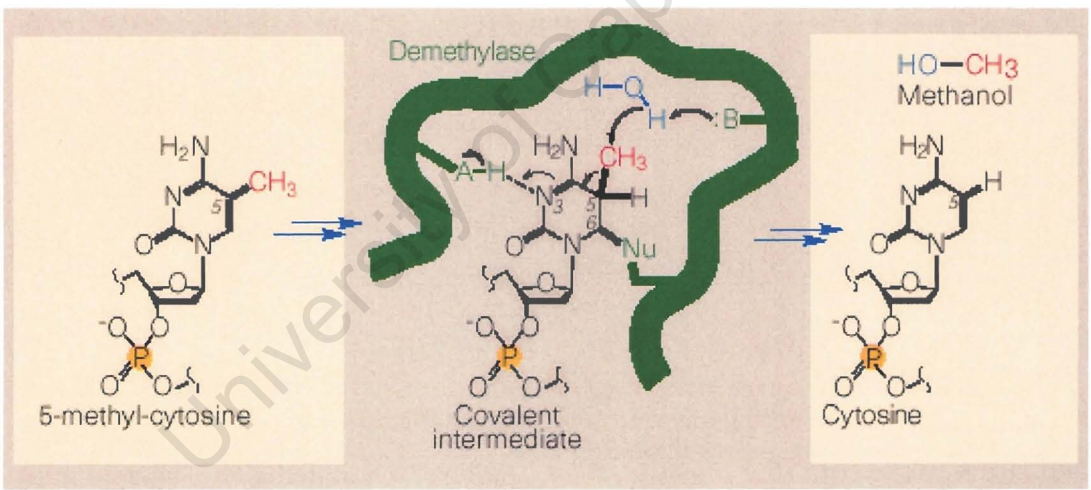


Figure1.3 Enzymatic demethylation of the 5-methyl-cytosine. The reactants of the above reaction are proposed to be water and methyl-dCpdG-bearing DNA. The demethylase (green) forms a covalent intermediate, comparable to that found during enzymatic methylation, through the addition of an enzymatic nucleophile (Nu-H) across the 5,6 double bond, assisted by the proton shuffling at N3. In situ activation of water generates a hydroxide that is then attacked by this intermediate with the ultimate products as non-methylated cytosine and methanol. The double arrows in blue indicate that the reaction involves two steps but the intermediates are not shown. This reaction is said to be thermodynamically favourable, even though it involves cleavage of a carbon-carbon bond, because its products are methanol and non-methylated cytosine (redrawn from Bhattacharya, *et al.* (1999) and Cedar and Verdine 1999).

The mechanisms by which these altered methylation patterns could be involved in transformation are through the inactivation of tumour suppressor genes and the activation of oncogenes (Herman, *et al.* 1994; Graff, *et al.* 1995; Herman, *et al.* 1995; Fang, *et al.* 1996; Herman, *et al.* 1996; Hsieh, *et al.* 1998; Klump, *et al.* 1998). Regional increases in DNA methylation at CpG sites that are usually never methylated yet found methylated in tumour cells results in gene inactivation (Counts and Goodman, 1995; Nagatake *et al.*, 1996; Gonzalgo and Jones, 1997; Gonzalgo *et al.*, 1998; Deng *et al.*, 1999).

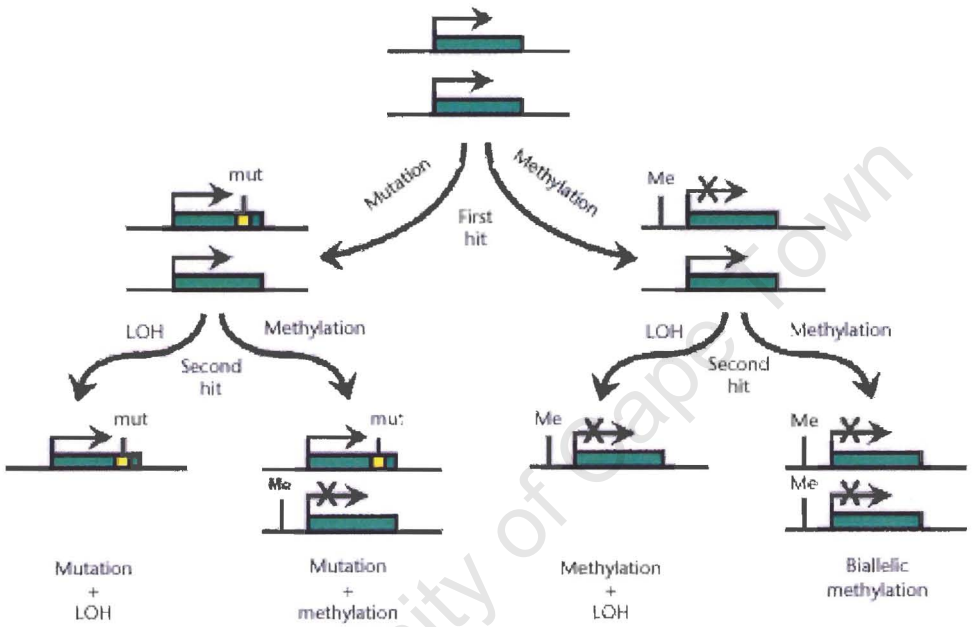


Figure 1.4. The revised Knudson's two-hit hypothesis. Jones and Laird (1999) have revised the Knudson's hypothesis of tumour progression to include the role of DNA methylation. They follow the fate of two active alleles of a tumour suppressor gene (green boxes) and show gene inactivation in two steps referred to as hits in the diagram. The first hit, representing the first step of inactivation, could either be due to a localised mutation shown by a yellow box in the top allele on the left panel of the diagram or by transcriptional repression represented by a cross (×) in the top allele on the right panel. The final step, or second hit, leading to repression could either be by loss of heterozygosity (LOH) or transcriptional silencing or a combination of the two (redrawn from Jones and Laird 1999).

Approximately half of the tumour suppressor genes causing familial cancer through germline mutations have been found to show inactivation associated with promoter hypermethylation in sporadic cancers. In some of these genes, hypermethylation was the only event that led to loss of expression (Baylin, *et al.* 1998). Examples of such tumour suppressor genes include von Hippel-Lindau (VHL) which promotes angiogenesis (Herman, *et al.* 1994), p16 ^{INK4a} a

cyclin dependent kinase inhibitor critical for maintaining Rb protein in its active, non-phosphorylated state (Herman, *et al.* 1995; Baylin, *et al.* 1998); and E-cadherin a cell adhesion protein that plays a key role in maintaining the epithelial phenotype (Graff, *et al.* 1995; Yoshiura, *et al.* 1995; Graff, *et al.* 1998). Transient transfection of specifically hypermethylated Rb gene constructs in sporadic retinoblastoma results in reduced expression compared to its unmethylated control. These results therefore suggest that hypermethylation inactivates tumour suppressor genes giving a selective growth advantage.

Reduced DNA methylation levels are also associated with activation of most proto-oncogenes such as the c-fos, c-myc, bcl-2, Ha-ras and Ki-ras genes (Hanada *et al.*, 1993; Counts and Goodman, 1994; Rainier and Feinberg, 1994; Fang *et al.*, 1996). Previous studies on the role of hypomethylation in the progression of colorectal cancer suggested an early role for DNA methylation in tumour progression (Rainier and Feinberg, 1988). These changes in methylation patterns are associated with increased DNA methyltransferase enzyme levels and could be modulated by association of DNA with histones (Kautiainen and Jones 1985; Kautiainen and Jones 1986; Issa, *et al.* 1993). Recent studies on the relationship between mismatch repair genes have implicated the changed methylation patterns in chromosomal imbalance (Lengauer, *et al.* 1997).

1.5 MECHANISMS OF TRANSCRIPTION INHIBITION BY DNA METHYLATION

Mechanisms by which DNA methylation could effect transcriptional inhibition have been the subject of many studies and only recently has more insight into the molecular details involved in this repression been described. The three best documented mechanisms invoke alterations of both DNA-protein and protein-protein interactions and are described briefly below.

1.5.1 DIRECT INHIBITION OF TRANSCRIPTION FACTOR BINDING.

One of the mechanisms by which DNA methylation interferes with transcription is through hindrance of the basal transcription machinery or ubiquitous transcription factors that require contact with cytosine in the major groove of the double helix. The methyl group on the C-5 position of cytosine residue in a CpG dinucleotide protrudes into the major groove (Klimasauskas, *et al.* 1994) where most sequence-specific interactions with DNA-binding proteins occur. The methyl group in that region thus directly affects the binding affinity of some sequence-specific proteins, such as AP-2 and c-myc, to their recognition DNA sequence

(Comb and Goodman, 1990; Prendergast, *et al.* 1991). CpG methylation of the cyclic AMP responsive element (CRE; 5'-TGACGTCA), a transcription factor-binding site associated with promoters of cell-type specific genes, has also been shown to result in loss of specific factor binding, accompanied by loss of transcription both in vivo and in vitro (Iguchi-Ariga and Schaffner, 1989). Furthermore, the binding of the preinitiation complex at the TATA-box and around transcription start site have also been reported to be directly excluded by the presence of the methyl-CpG in the vicinity of the TATA box (Meehan, *et al.* 1989; Boyes and Bird 1991; Bird 1992; Levine, *et al.* 1992). The presence of methyl-CpG dinucleotide on a promoter may deny access to the transcription machinery by influencing nucleosome stability or positioning.

Davey, *et al.* (1997) have also reported exclusion of the histone octamer from interacting with a high affinity positioning sequence when the promoter region of the chicken adult β-globin gene is methylated. Histone H1 is well established as a key element in the formation and stability of chromatin fibres and preferential binding to methylated DNA would be a clear indication of the role that methylation plays in nucleosome stability. The role of DNA methylation on the stability of the nucleosome is, however, still a contentious issue because of differences in experimental results investigating the relationship between DNA methylation and the linker histone H1 (Higurashi and Cole 1991; Levine, *et al.* 1993; Campoy, *et al.* 1995; Nightingale and Wolffe 1995; McArthur and Thomas 1996). This difference in results has been attributed to the choice of naked or nucleosomal templates, specific variants of H1 and the particular promoters used (Bird and Wolffe, 1999). Direct gene inactivation by DNA methylation is said to generally play a minor role in the methylation-dependent repression if it is not accompanied by the binding of methyl-CpG binding proteins (Bird, 1992).

1.5.2 DNA METHYLATION AND CHROMATIN STRUCTURE

Chromatin has long been viewed as a dynamic structure capable of compacting and decompacting DNA and ensuring a higher magnitude of regulation of gene expression in eukaryotes (Udvardy, 1999). This is achieved by remodelling of the chromatin structure with the reversible acetylation of lysines on the amino-terminal tails of nucleosomal histones (review in Pollard and Peterson, 1998). Post-translational modification of the core histones through acetylation has been shown to destabilise the nucleosomes, thereby relieving chromatin induced transcriptional repression and allowing access to recognition elements by

transcription factors. Deacetylation of histones on the other hand, stabilises the repressed state and excludes transcription regulatory proteins from binding DNA and in some cases even impede elongation (Workman and Kingston 1992; Ura, *et al.* 1997). Proteins capable of remodelling chromatin are key players in regulating transcription and consequently chromatin remodelling through modification of its core histones is a prerequisite for transcription.

The recently identified proteins containing the 80 amino acid methyl-CpG-binding domain (MBD) has the ability to interact with a co-repressor complex containing histone deacetylases HDAC1 and 2 and are therefore key players in DNA methylation mediated-transcriptional repression (Hendrich and Bird 1998; Jones, *et al.* 1998; Nan, *et al.* 1998; Ng, *et al.* 1999). These proteins are ubiquitously expressed in somatic cells and consequently good candidates for the global effect of DNA methylation on transcriptional repression in eukaryotes (Nan, *et al.* 1998). Further characterisation of these proteins, particularly methylcytosine binding protein 2 (MeCP2), is proposed to have a common mechanistic pathway in transcriptional repression by DNA methylation and histone deacetylation (Eden, *et al.* 1998; Jones, *et al.* 1998; Nan, *et al.* 1998; Cameron, *et al.* 1999). The details of this process are still unclear, however, a model that links the two important mechanisms of gene regulation, DNA methylation and histone deacetylation, has recently been demonstrated with methylcytosine binding proteins at the forefront (Jones, *et al.* 1998; Nan, *et al.* 1998; Ng, *et al.* 1999; Wade, *et al.* 1999; Zhang, *et al.* 1999).

A surprising recent discovery is the association of the deacetylase activity with the maintenance methyltransferase DNMT1 in vivo (Fuks, *et al.* 2000). Based on the sequence similarity observed between a region of Dnmt1 and the repressor domain of the trithorax-related protein HRX (Cross, *et al.* 1997), transient transfection in U2OS osteosarcoma cells with fusion proteins containing the Dnmt1 region were carried out to show this association. These authors suggest that this deacetylase activity domain in DNMT1 may be necessary for remodelling chromatin to enable DNA methylation to take place. These reports have recently been substantiated by findings from Robertson, *et al.* (2000) and Rountree, *et al.* (2000). The same non-catalytic region in Dnmt1 described in Fuks, *et al.* (2000), was shown to bind HDAC2 and a new co-repressor protein DMAP1 (DNMT1 Associated Protein) that can mediate transcriptional repression. Further, DNMT1 was shown to co-fractionate with eight polypeptides, five of which were identified through western-blot analysis as HDAC1, the retinoblastoma (Rb) tumour suppressor gene product RbAp46 and RbAp48, and proliferating

cellular nuclear antigen (PCNA). The association of RbAp46/48 and PCNA with DNMT1 has been previously described (Chuang, *et al.* 1997; Fuks, *et al.* 2000) and these proteins were found to dissociate from the DNMT1, HDAC1 and E2F complex (Robertson, *et al.* 2000). Further characterisation of the interactions of the DNMT1-HDAC1/E2F complex revealed that DNMT1 also associates with the A/B pocket of Rb found to be frequently mutated in human tumours (Grana, *et al.* 1998). Taken together, these results establish a definite link between DNA methylation, chromatin and transcriptional repression. It remains to be shown how these processes are linked in eukaryotic gene regulation.

1.5.3 CHROMATIN REMODELLING BY MeCP AND HISTONE DEACETYLASE

Recently (Eden and Cedar 1994; Eden, *et al.* 1998; Jones, *et al.* 1998; Nan, *et al.* 1998) have provided evidence that MeCP2 is one of the key players in methylation-dependent transcriptional repression. Chromatin structure modification mediated by the methylcytosine-binding protein2 (MeCP2) inhibits transcription. Immunoprecipitation and stable transfection experiments have shown that MeCP2 is capable of repressing transcription through its ability to bind methylated CpG residues and to recruit and interact with the mSin3A co-repressor complex and histone deacetylases-HDAC 1 and HDAC2 in mammalian and Xenopus cells (Jones, *et al.* 1998; Nan, *et al.* 1998). This multiprotein complex then induces chromatin remodelling by altering the charge on the core histone H3 and H4 tails resulting in stronger interactions between the positively charged lysine residues and DNA phosphate backbone. The subsequent chromatin structure therefore most probably precludes transcription by allowing tighter nucleosomal packaging. Subsequently the promoter of a gene is rendered inaccessible to the transcription machinery thus either transcription initiation or elongation or both are greatly reduced (Barry, *et al.* 1993; Hsieh 1997; Kass, *et al.* 1997). Transcriptional reactivation of hypermethylated tumour suppressor genes MLH1, TIMP3, CDKN2B (INK4B, p15), CDKN2A (INK4, p16) and the human β-globin promoter is consistent with the above model. These genes could not be reactivated through inhibition of histone deacetylation using trichostatin A (TSA) alone, they also required minimal demethylation with 5-aza-2 deoxycytidine (5aza-dC) (Razin 1998; Cameron, *et al.* 1999).

Another post-synthetic modification that has recently been correlated with DNA methylation poly(ADP-ribosyl)ation and could be involved in chromatin structural modulation (Zardo, *et*

al. 1997; de Capoa, *et al.* 1999; Zardo, *et al.* 1999). Treatment of cells with the well-known inhibitor of poly(ADP-ribose) polymerase, 3-aminobenzamide, resulted in increased CpG methylation in the promoter region of the Htf9 housekeeping gene and was proposed to implicate poly(ADP-ribosyl)ation in protecting CpG islands from complete methylation. These authors provided cytological evidence to support this hypothesis and also found that reduced levels of poly(ADP-ribosyl)ation not only resulted in DNA hypermethylation but also chromatin compaction.

Of the four novel genes encoding proteins containing an MBD-like motif (MBD1-MBD4) recently characterised (Hendrich and Bird, 1998), only three were capable of binding methylated DNA in vitro and in vivo. MBD3 protein, despite having 71% amino acid identity with MBD2, did not show a significant affinity for methylated DNA sequences (Hendrich and Bird, 1998, Zhang, *et al.*1999). MBD2 on the other hand has been reported to readily bind methylated DNA and both MBD2 and 3 appear to interact with different histone deacetylase complexes. MBD3 associates with the Mi/NRD histone deacetylase complex in both human (Zhang, *et al.* 1999) and Xenopus cells (Wade, *et al.* 1999).

A recent report describing an association between the MBD2-MBD3 complex and DNMT1 (Tatematsu, *et al.* 2000) is very interesting in light of the data describing DNMT1 association with histone deacetylases and co-repressor complexes explained in section 1.5.2 above. MBD2 and MBD3 were found to form homo- and hetero-dimers that are able to recognise hemi-methylated DNA in vivo and in vitro. The MBD2-MBD3 complex was further shown to co-localise with DNMT1 at late S phase. Based on these results, MBD3 was proposed to play an important role in the recognition of hemi-methylated cytosines during replication and recruitment of DNMT1-HDAC complexes to the replication foci in order to establish/maintain transcriptionally repressed chromatin. A somewhat controversial finding is that MBD2 is a transcriptional repressor and component of multi-subunit histone deacetylase complexes of MeCP1, HDAC1, HDAC2, and RbAp46 and RbAp48 (Ng, *et al.* 1999) and DNMT1 (Tatematsu, *et al.* 2000). Contrary to the demethylating role described above, (Ng, *et al.* 1999), these authors have used Gal4-MBD2 fusion protein, in MeCP2 experiments, to show that MBD2 represses promoters with the nearby Gal4-binding sites. TSA was also shown to reduce repression due to this protein. In vitro and in vivo experiments Apart from the characterised methylated-binding domain, the functions of MBD1 and MBD4 are not yet fully understood.

Proteins may also preferentially bind methylated DNA sequences, thereby preventing transcription factors from accessing their binding sites. These proteins include the ubiquitous methyl-DNA-binding protein 1 (MDBP-1) (Zhang et al., 1990), methyl-cytosine-binding protein 1 (MeCP1) (Meehan et al., 1989 & Boyes and Bird, 1991), MeCP2 (Meehan, *et al.* 1989), and MDBP-2 (Jorst and Hofsteenge, 1992). The latter protein was identified in the chicken vitellogenin II gene as a member of the histone H1 family in that it has two peptide motifs that are similar to those found in histone H1. However, the complete amino acid sequence has not yet been documented. Knockout experiments of the methyl-C binding protein 2 (MeCP2) (Tate et. al, 1996) resulted in lethal embryonic development similar to DNMT1 deficiency (Li, *et al.* 1992). However, in contrast to DNMT1, MeCP2 deficiency was found to be compatible with somatic cell viability indicating that MeCP2 is not required for somatic cell viability. Furthermore, imprinted genes such as Xist gene on the X chromosome, were not expressed in DNMT1 yet were expressed normally in MeCP2-deficient cells.

The above mechanisms, however, are not mutually exclusive. Experimental evidence to date does not allow preference over any particular one although the repression by MeCP2 seems to be mainly associated with heterochromatin. With the results of the new MBD that are associated with repressor complexes containing histone deacetylases, these results collectively reveal chromatin modification as a critical feature of the methylation-mediated silencing mechanism. A general overview of how methylation-mediated repression could be effected in eukaryotes is shown in Figures, 1.5 and 1.6, taken from recent reviews (Jones and Laird 1999; Newell-Price, *et al.* 2000).

1.6 DNA METHYLATION AND α2(I) COLLAGEN GENE EXPRESSION

The structural integrity of vertebrate organs is maintained to a large extent by the collagen proteins. A number of developmental and physiological processes are also greatly influenced by the fibril-forming collagen proteins. These include cell adhesion, cell migration, tissue remodelling, inflammation and wound healing (Sage and Bornstein, 1991). Although the collagens are structurally and functionally related, they remain genetically distinct. A series of complex interactions between *cis*- and trans- acting regulatory elements is involved in the expression of these proteins. Type I collagen is the most abundant protein of all the collagens,

it is present in virtually all tissues with significant prevalence in bone, skin, teeth, blood vessels, tendons and ligaments (Karsenty and de Crombrugghe, 1991). The unique triple helical structure of type I collagen is achieved by a heterotrimer of two α1(I) chains and one α2(I) chain. Transcription of the α1(I) and α2(I) procollagen genes is tightly controlled, with the mRNA steady-state levels maintained at a relative stoichiometry of 2:1 (Vuust *et al.*, 1985).

The type I collagen genes are located on different chromosomes, the α1(I) gene is on chromosome 17q21-22 and the α2(I) gene is on chromosome 7q21-22 (Kotze *et al.* 1986; Zengerling *et al.*, 1987, Lathrop, *et al.* 1988). Akai, *et al.* (1999) have proposed that the observed coordinate expression of the genes in normal conditions is a result complex mechanisms that involves similar trans-acting factors binding to common *cis*-regulatory elements. Previous studies have identified similarities in the *cis*- and *trans*- acting factors involved in the regulation of these genes (Vuorio and de Crombrugghe, 1990). A number of studies have further indicated that the proximal promoter element (350bp) contains all the necessary sequences for directing the tissue-specific expression of the gene (Boast, *et al.* 1990, Niederreither, *et al.*1992, D'Souza, *et al.* 1993; Bou-Gharios, *et al.* 1996).

Moreover, regulatory and responsive elements that can be influenced by a variety of cytokines and growth factors have been identified in the proximal promoter (Parker, *et al.* 1992, Inagaki, *et al.* 1994, Inagaki, *et al.*1995, Tamaki, *et al.* 1995; Chung, *et al.* 1996; Collins, *et al.* 1997). The TGF-β responsive element (TbRE), which mediates the stimulatory effect of TGF-β through the action of two *cis*-elements located within nucleotides –378 and –255, has also been described (Inagaki, *et al.* 1994, Inagaki, *et al.*1995, Chung, *et al.* 1996). Although the identity of the factors interacting with these *cis*-elements is still unknown, there is currently evidence for the involvement of both the AP1 or Sp1 factor in the TGF-β response (Chung, *et al.* 1996, Greenwel, *et al.* 1997). The TNF-α responsive element (TaRE), located between nucleotides –271 and –235, has been identified and shown to involve NF-κB1 and Rel1A in mediating COL1A2 transcription inhibition in response to TNF-α (Inagaki, *et al.* 1995, Kouba, *et al.* 1999). Another proximal promoter element that mediates COL1A2 transcriptional repression was found to be the interferon-gamma response element (IgRE) is located between nucleotides –161 and -125 (Higashi, *et al.* 1998). IgRE was shown to be distinct from the TaRE, as they are located in different regions of the promoter. Moreover an additive inhibitory effect of the two cytokines has been shown (Higashi *et al.*, 1998).

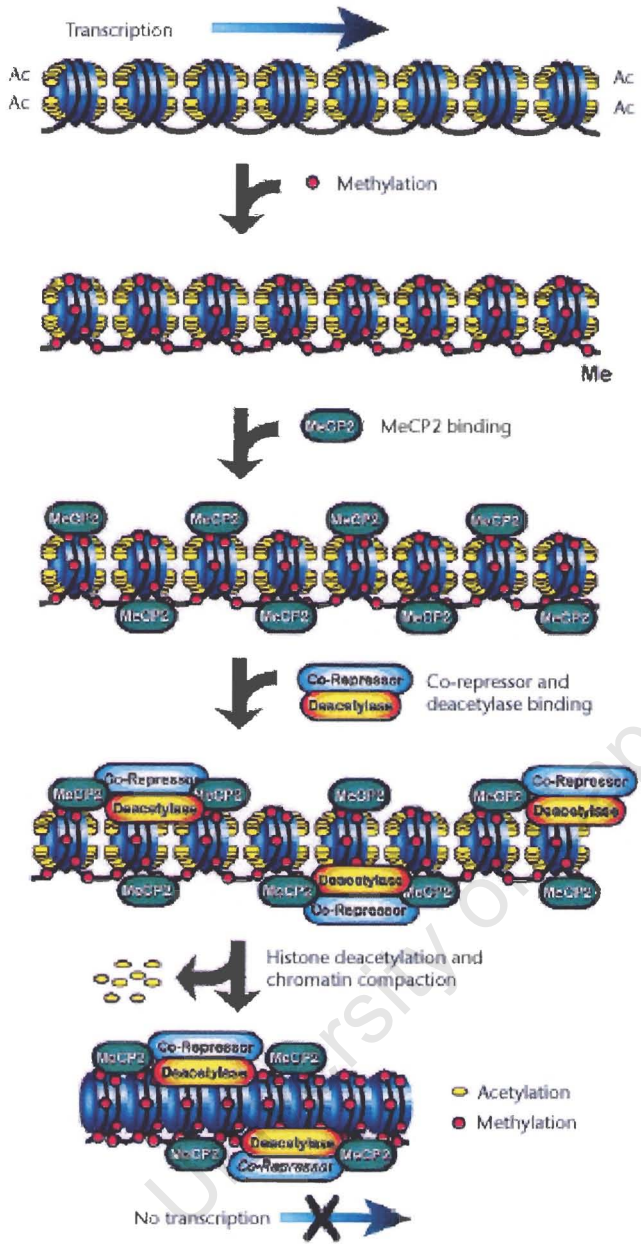


Figure 1.5 Schematic representation of transcriptional repression due to chromatin remodelling effected by DNA methylation. The top diagram depicts the structure of a transcriptionally active gene (blue arrow) where the nucleosomes (blue cylinders), representing the core histones that complex with DNA have acetylated histone tails (yellow ovals). When the gene is methylated (mCpG shown as red circles), binding of the MeCP2 (or other MBD proteins) is facilitated which then recruits transcriptional co-repressors (yellow with red edges) and histone deacetylases (light blue). This results in a compact structure that is inaccessible to trans-acting factors and the transcription machinery and thus transcriptional repression. Adapted from Jones and Laird (1999).

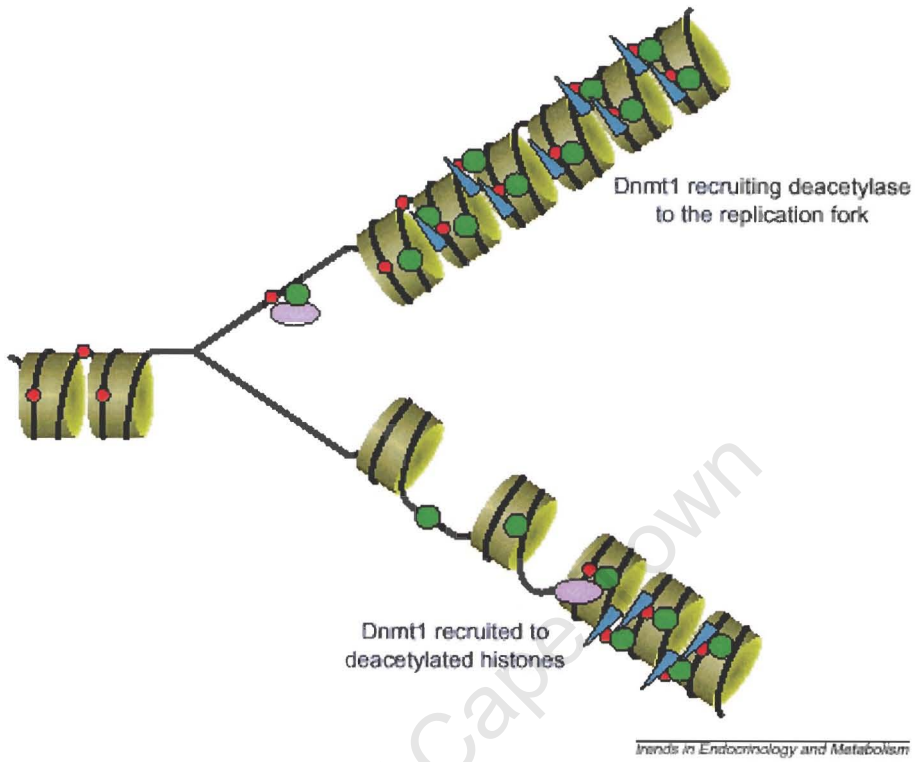


Figure 1.6 The hypothetical interplay between Dnmt1 and histone deacetylation. The upper strand postulates that Dnmt1 (pink ovals) methylates the daughter strands after replication (symmetrically methylated cytosines shown as red circles) thus attracting proteins with a methyl-C-binding domain (MBD-blue triangles) that are in complexes with histone deacetylases (HDACs-green circles), to ensure nucleosome assembly of deacetylated histones and subsequently condensed chromatin. An alternative pathway is shown in the bottom strand. Here it is proposed that Dnmt1 is recruited to the replication foci, by the MBD2-MBD3 hetero-dimer and its associated HDACs, after histone deacetylation has occurred. This pathway therefore suggests that efficient maintenance methylation of DNA might require compact chromatin. Adapted from Newell-Price, *et al.* (2000).

The human α2(I) collagen gene is downregulated in transformed human lung embryonic fibroblasts (Parker, *et al.* 1989). Also, several investigators have demonstrated the proximal promoter of the COL1A2 gene to be sufficient for transcriptional regulation of the gene (Boast, *et al.* 1990). Thus the *cis*-acting and trans- acting factors responsible for COL1A2 expression are the focus of a variety of studies. The proximal promoter of the COL1A2 gene is CG- rich and as such is a good candidate gene for methylation-induced repression.

Previous studies have shown that the rat COL1A2 gene is hypermethylated in the promoter and first exon after chemical transformation, and transfection of the rat COL1A2 promoter/first exon is sensitive to DNA methylation (Smith & Marsilo, 1988). Recently Sengupta & Smith (1998) using in vitro assays have reported an inhibition of transcription when CpG sites in the first exon were methylated.

Previous data on the role of DNA methylation are inconsistent and this is primarily due to the absence of good methylation detection techniques at the time. The methods used did not enable identification of all the methylated cytosine residues. In this study, the methylation status of the human COL1A2 promoter was studied using the recently described bisulphite genomic sequencing technique, which enables a clear distinction of all methylated cytosines (Frommer, *et al.* 1992; Clark, *et al.* 1994). Two cell lines, normal human lung embryonic fibroblasts and their SV40 transformed counterparts, were used as a model as they show the expected transformation associated changes in collagen gene expression. The results demonstrated that none of the CpG dinucleotides in the proximal promoter of the collagen expressing CT-1 cell line were methylated whereas 9 out of 17 CpG dinucleotides were methylated in the SV40-transformed cells. Moreover, these methylated sites were observed predominantly in the first exon around the region where the preinitiation complex forms. The sensitivity of the human α2(I) collagen gene promoter to in vitro methylation was tested using the bacterial M. SssI methylase, an enzyme that methylates CG dinucleotides. Expression constructs of the COL1A2 promoter were methylated *in vitro* and transiently transfected into lung fibroblasts. These studies showed that COL1A2 promoter activity was sensitive to DNA methylation since the methylated constructs had drastically reduced promoter activity.

CHAPTER 2

METHYLATION STATUS OF THE HUMAN α2(I) PROCOLLAGEN GENE PROXIMAL PROMOTER

2.1 INTRODUCTION

De novo methylation of CpG islands in the promoter region has conclusively demonstrated a reciprocal relationship between methylation status and transcriptional repression of the genes (Kass *et al*, 1997; Razin *et al*, 1991). Type I collagen is a transformation sensitive protein, and as such could be hypermethylated in transformed cells as the gene is inactivated in many transformed cells. The α2(I) procollagen proximal promoter is characterised by a high GC content and more potential methylation sites than the pro-α1(I) gene. Progress in the analysis the methylation status of the COL1A2 proximal promoter has been somewhat controversial due to limitations in techniques that were available for DNA methylation analysis. Recently Sengupta & Smith (1998) have demonstrated that methylated CpG sites in the first exon play a crucial role in repressing transcription of the rat α2(I) procollagen gene.

A number of studies have also demonstrated that DNA methylation plays a pivotal role in transcriptional regulation of various collagen genes. Rhodes *et al* (1992 & 1994) have studied the molecular mechanisms that regulate the stage- and tissue- specific expression of the murine α1(I) collagen gene and have shown that hypermethylation of the regulatory elements in the 5` region of the gene is involved in repressing promoter activity. The α2(I) collagen gene in chemically transformed rat liver epithelial cells is transcriptionally repressed and is associated with increased methylation of the α2(I) promoter.

Transient transfection experiments with murine and human promoter constructs have shown that the 351bp proximal promoter of the α2(I) procollagen gene is sufficient to direct tissue-specific expression (Boast *et al*, 1990) and that both positive and negative regulation of these genes is achieved through proximal promoter elements (Karsenty & de Crombrugghe, 1990 & 1991).

These findings have been supported by studies showing that the key control features in eukaryotic gene expression reside primarily at the transcriptional level (Simkevich *et al*, 1992; Slack *et al*, 1993). Furthermore, it has been proposed that the repressor mechanism of a given promoter is determined by the DNA sequences and their context within that given promoter and not exclusively by the DNA-binding proteins (Ogbourne and Antalis, 1998). Therefore, identification of modified bases within the COL1A2 promoter and their role in transcriptional regulation of this gene is important in understanding its temporal and/or spatial expression patterns observed during development and in pathological conditions.

Some of the major regulatory elements in the COL1A2 proximal promoter include the TATA box, inverted CCAAT box, nuclear factor-1-like site (NF-1), inhibitory factor (IF-1) site, Sp1/ AP-2 site and the more recently identified initiation region (Inr). These sequences recruit transcription factors (TFs) to the promoter and modification of these sites or sites in close proximity could result in inhibition of TF binding and consequently inhibit transcription. DNA methylation is the most widely documented modification of DNA that has been shown to affect chromatin structure and binding of TFs.

The use of methylation sensitive restriction endonucleases to detect the DNA methylation status has a limited scope because not all methylated sites are located within restriction endonuclease sites. The bisulphite modification technique described by Frommer *et al* 1992 and Clark *et al* 1994 overcame this limitation. This technique enables positive identification of methylated cytosine residues and consequently the methylation status of individual DNA strands. PCR amplification of the DNA sequence of interest with strand-specific primers yields DNA fragments in which all cytosine residues have been replaced by thymine and only the 5mC remains as cytosine.

The bisulphite reaction is based on the fact that oxidation reagents such as bisulphite, form an adduct across the 5-6 double bond of cytosine (Figure1).

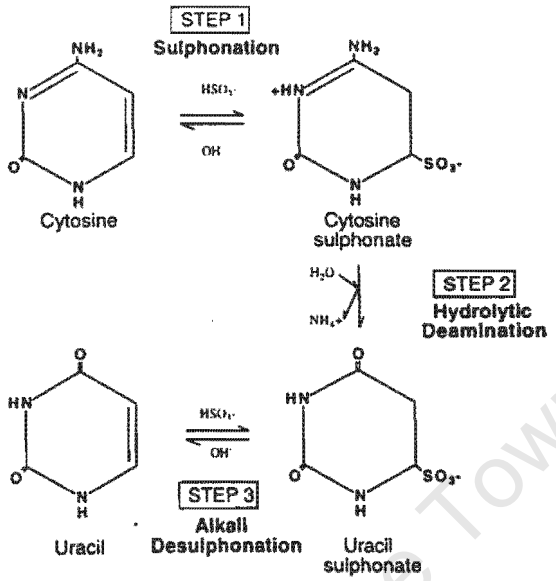


Figure 1. Schematic representation of the steps involved in the bisulphite conversion reaction mechanism. The addition of the bisulphite ion across the carbon 5-6 bond of cytosine (step1) and the removal of the sulphonate group (step3) are reversible reactions and as such require stringent pH, temperature and bisulphite concentration conditions to favour the forward reaction. (Adapted from Clark *et al.*, 1994).

Once cytosine has been successfully deaminated to uracil, PCR amplification is carried out with primers designed specifically for the deaminated template. Parker *et al* (1986), have shown that a clone of simian virus transformed human embryonic lung fibroblasts (SVWI-38) does not produce the α2(I) chain although the gene is still intact. The proximal 351bp COL1A2 promoter that is CpG rich was used to study the influence of DNA methylation on COL1A2 expression.

2.2 RESULTS

2.2.1 Confirmation of C⟶U conversion in bisulphite-modified DNA by HPLC analysis.

Genomic DNA from collagen expressing and non-collagen expressing cells was subjected to bisulphite treatment in order to determine the extent of cytosine methylation. The bisulphite technique depends on the preferential deamination of

unmethylated cytosine to uracil. Successful deamination of cytosine to uracil was confirmed by base composition analysis by ion-exchange high performance liquid chromatography (HPLC).

Bisulphite treated DNA and untreated DNA was hydrolysed in formic acid (88%, v/v) for 30 min at 180°C to its constituent bases as described in section 4.2. The hydrolysate was then analysed by HPLC on a Beckman Ultracil CX cation exchange column. Base standards eluted in the following order: uracil, thymine, guanine, cytosine, adenine and 5-methylcytosine. Although DNA from the different cell types was analysed by HPLC, only the SVWI-38 elution profiles of untreated and bisulphite treated DNA are shown since a similar pattern was observed for all the DNA samples. Comparison of the HPLC profiles of the untreated (Fig 2.2 a) and bisulphite treated DNA (Fig 2.2b), showed the expected uracil peak in the former with the corresponding disappearance of cytosine peak. These results therefore confirmed that sodium bisulphite treatment achieved deamination of cytosine to uracil.

2.2.2 Nucleotide sequence of the human proximal α2(1) procollagen promoter.

The GC-rich promoter region -220 to +130 was analysed for CpG methylation. This sequence constitutes the minimal promoter region required for cell-type specific expression of the human α2(1) procollagen gene (Figure 2.3).

2.2.3 PCR Amplification of the bisulphite-modified human proximal α2(1) procollagen promoter.

Bisulphite treated DNA was PCR-amplified using the primers shown in figure 2.3 and the PCR fragments fractionated on 2% agarose gels. Plasmid pUCT-1 containing the human α2(1) procollagen promoter was used as a positive control. Two rounds of PCR amplification were performed to increase specificity and yield. The expected 351-bp PCR product was successfully amplified in all three samples (Figure 2.4).

a)

b)

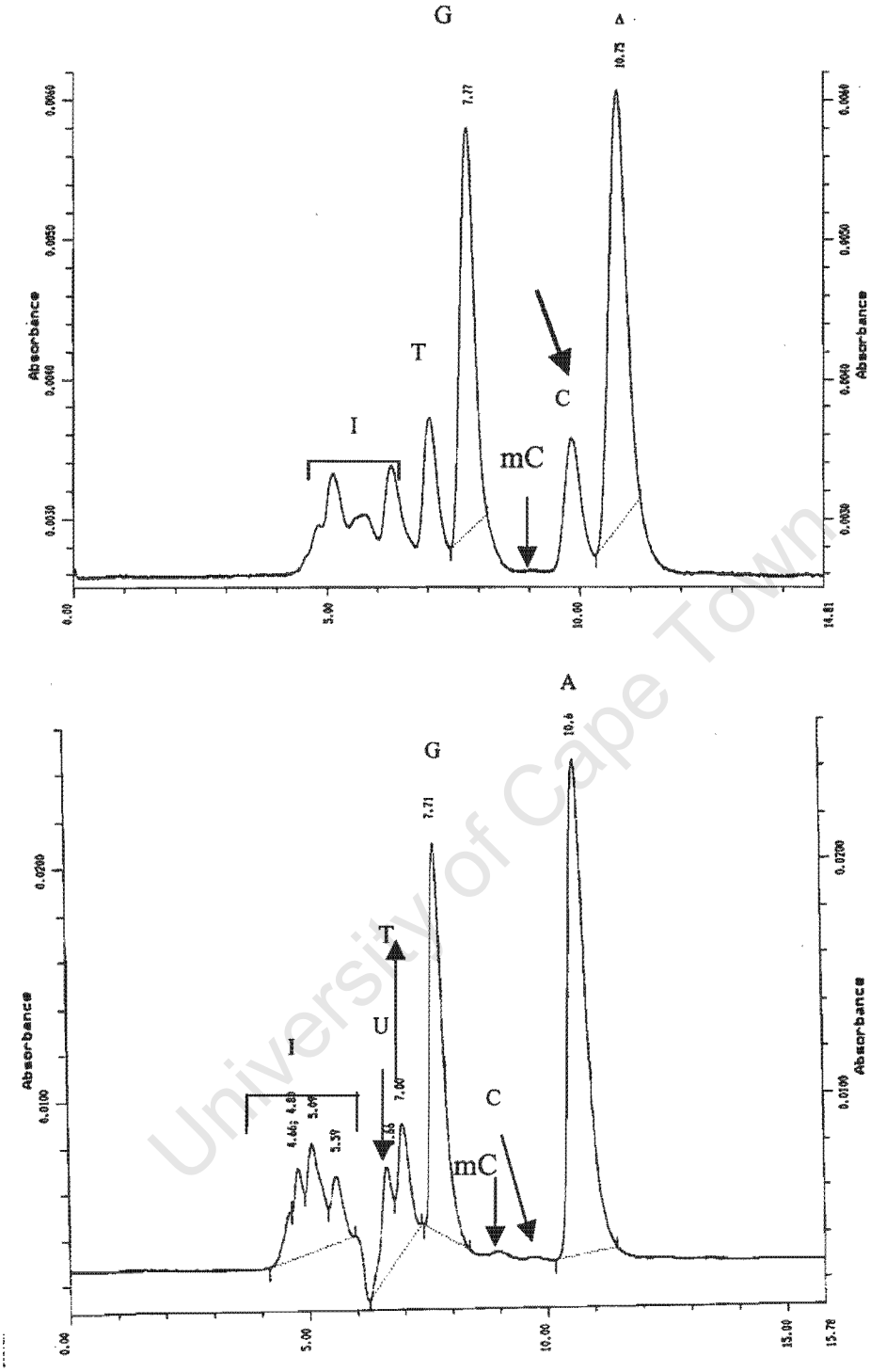


Figure 2.2 HPLC Profile of untreated and bisulphite treated DNA: Genomic DNA from SV40-WI38 fibroblasts was treated with bisulphite as described in Materials and Methods (section 4.2), hydrolysed in 88% formic acid and the bases separated on a cation-exchange HPLC column (Beckman Ultrasil CX). The bases were eluted with 100mM sodium acetate and 5% methanol at a flow rate of 1ml/min. Peak I represents uncharacterised hydrolysis products. The elution profiles are of untreated (a) and bisulphite treated (b) DNA. The position of A, G, T, U and mC are indicated.

-350 TGCAGAGCAC TCCGACGTGT CCCATAGTGT TTCCAAACTT GAAAGGGCG
NF-1 S_p1/AP-2

-300 GGGGAGGGCG GGAGGATGCG GAGGGCGGAG GTATGCAGAC AACGAGTCAG
AP-2 S_p1 CAGA BOX
*

-250 AGTTTCCCCT TGAAAGCCTC AAAAGTGTCC ACGTCCTCAA AAAGAATGGA
(FORWARD PRIMER) TTTTAA AAAGAATGGA
* * *

-200 ACCAATTTAA GAAGCCAGCC CCGTGGCCAC GTCCCTTCCC CCATTCGGGC
ATTAATTTAA GAAG **AP-2**
* * *

-150 CCTCCTCTGC GCCCCCGCAG GCTCCTCCCA GCTGTGGCTG CCCGGGCCCC
AP-2

-100 CAGCCCCAGC CCTCCCATTG GTGGAGGCCC TTTTGGAGGC ACCCTAGGGC
G/CBE CME
* *

-50 CAGGGAAACT TTTGCCGTAT AAATAGGGCA GATCCGGGCT TTATTATTTT
TATA BOX
* * *

+1 ***AGCACCACGG CAGCAGGAGG TTTCGGCTAA GTTGGAGGTA CTGGCCACGA***
* * * * *

+70 ***CTGCATGCCC GCGCCCGCCA GGTGATACCT CCGCCGGTGA CCCAGGGGCT***

CTGCGACACA AGGAGTCTGC ATGTCTAAGTGCTA
CTATAT TCCTCAAACA TACAAATTCACAAT (REVERSE PRIMER)

Figure 2.3 The nucleotide sequence of the -350bp minimal promoter of the α2(I) procollagen gene. The methylation status of the CpG dinucleotides indicated by an asterix was determined with the bisulphite genomic sequencing method. Exon 1 is shown in italics. Transcription factor binding sites identified as important for α2(I) procollagen gene regulation in the mouse are underlined. The CCAAT box element (G/CBE), the collagen-modulating element (CME) and the TATA binding element are doubly underlined. The primers used for PCR amplification of the bisulphite modified DNA are also indicated.

PCR amplification is the most critical step in the bisulphite genomic analysis protocol. The primers should be specific for bisulphite converted DNA in order to enable analysis of the methylation profile of each CpG dinucleotide. A number of primer sets were used for optimisation until a discreet PCR product of 351pb was obtained. The absence of a product in untreated DNA further confirmed the specificity of the primer set used to amplify bisulphite treated DNA. During PCR amplification, the deaminated cytosine is amplified as thymine and the 5-methylcytosine as cytosine.

The following properties were suggested for successful amplification (Frommer *et al*, 1992; Clark *et al*, 1994):

a) Primers should be between 25 and 30 bases in length,
b) Primers should anneal to a region of the DNA that has little or no CpG sites,
c) Primers should be designed to G-rich regions,
d) Primers should have limited internal homology or complementarity to each other,
e) Primers should contain a C at the 3′ end.

Nested primers have also been suggested to improve efficiency of amplification, but in this study, two rounds of PCR with the same primer pair proved to be sufficient. A number of investigators have still found a problem of PCR bias regardless of the adherence to the above conditions. Warnecke *et al* (1997) have shown that methylated and unmethylated DNA are not always amplified proportionately, especially in CpG rich DNA. To circumvent this problem of bias, Voss *et al* (1998) have suggested the inclusion of betaine in the PCR reactions. In our experiment, untreated genomic DNA was used as control for PCR amplification bias. The primers were found to efficiently amplify converted DNA, as there was no PCR product obtained with untreated DNA.

The 351bp PCR product was subcloned into pCR2.1® TA vector or pGEM-T-Easy vector for automated or manual sequencing respectively. PCR amplification with M13 primers was used to determine clones with inserts before isolation of plasmid DNA for further sequencing (Figure 2.5).

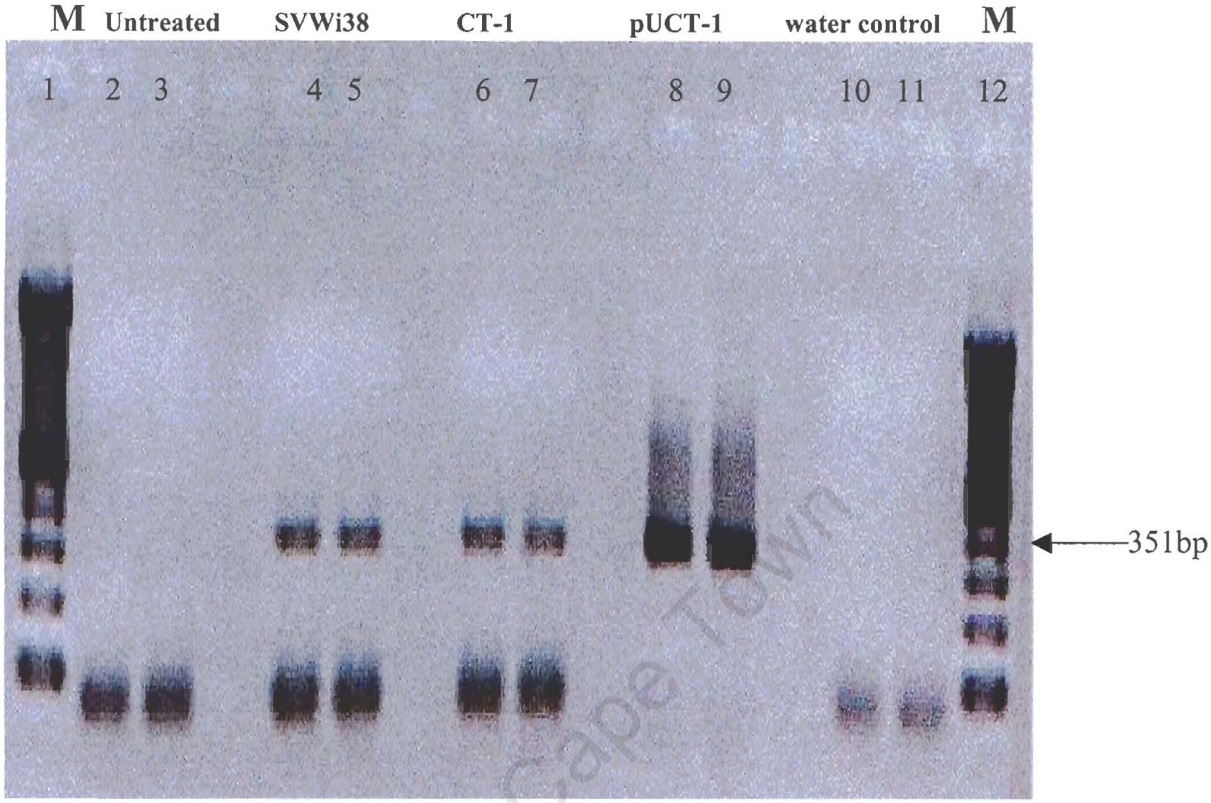


Figure 2.4 PCR Amplification of the bisulphite treated proximal human COL1A2 promoter. 500ng DNA isolated from SVWI38 and CT-1 fibroblasts, or the cloned α2(I) promoter (pUCT-1) was treated with bisulphite and subjected to two rounds of PCR amplification (94°C/2min×1cycle, (94°C/20sec, 53°C/20sec, 72 °C/1min) × 29 cycles, 72°C/5min×1cycle). Primers specific for converted DNA were designed with C or G at their 3′ or 5 end to increase stability of hybrids, (Forward primer 5′-TTT TAA AAA GAA TGG AAT TAA TTT AAG AAG-3′ and reverse primer 5′-TAA CAC TTA AAC ATA CAA ACT CCT TAT ATC-3′). The PCR products were analysed on 2% agarose gels, stained with ethidium bromide, and viewed under UV-light. Lanes 2 & 3 are untreated DNA and the lanes 4,5,6,7,8,9, are bisulphite treated DNA. A 1000bp DNA ladder was used as marker and the 351bp is shown by the arrow. Lanes in between the different samples were left empty to avoid spill over.

2.2.4 Subcloning of the PCR product into TA vectors.

The PCR products obtained from the bisulphite treated DNA were subcloned into linearized TA vectors for analysis of the methylation profile. Both manual and automated sequencing were performed thus two TA vectors were used, pGEM®-T Easy and pCR®2.1-TOPO. These vectors are highly efficient systems for the cloning of PCR products as they provide a direct insertion of Taq polymerase-amplified PCR products into the single 3′-T overhangs at the insertion site. A number of thermostable polymerases, including Taq polymerase, are characterised by their ability to add a single deoxyadenosine (A) to the 3′ end of amplified fragments in a template-independent terminal transferase activity. Thus the linearized vectors with 3′ deoxythymidine (T) overhangs of the vectors provide compatible termini allowing efficient ligation.

Positive selection of clones with inserts was carried out using the blue/white colour screening using β-galactosidase. When an insert is cloned into the vector, β-galactosidase is destroyed and unable to cleave Xgal hence the appearance of white colonies. The white colonies were then selected, plasmid DNA isolated and digested with EcoRI. Figure 2.5 shows the clones used in subsequent sequencing experiments.

2.2.5 Methylation pattern of the α2(I) collagen promoter.

Sequence analysis of the plasmids containing the cloned PCR products revealed a striking difference between the collagen producing (CT-1) and non-collagen producing (SVWI-38) fibroblasts. The 351bp proximal promoter was completely unmethylated in the former whereas 11 of the 17 CpG dinucleotides (64%) were found to be methylated in the latter. Figures 2.6 and 2.7 depict representative results obtained following sequencing of bisulphite treated α2(I) proximal promoter from SVWI-38 and CT-1 cells respectively. Figure 2.8 summarises these results and compares the methylation status of the CpG dinucleotides between the two cell lines. The published sequence is included for reference.

It is significant that methylated CpG dinucleotides were found in the vicinity of the TATA box where the assembly of the transcription machinery is anchored through specific interactions with the TATA binding proteins (TBP). Furthermore, the methylated cytosine residues were found predominantly in the first exon and recent findings have reported that the pre-initiation complex extends to nucleotide +30 for

stability and proper positioning of the RNA polymerase II (Buratowski *et al.*, 1989; Roy *et al.*, 1993). These results therefore suggest that DNA methylation could play a critical role in α2(I) procollagen gene transcription by inhibiting the assembly of the pre-initiation complex.

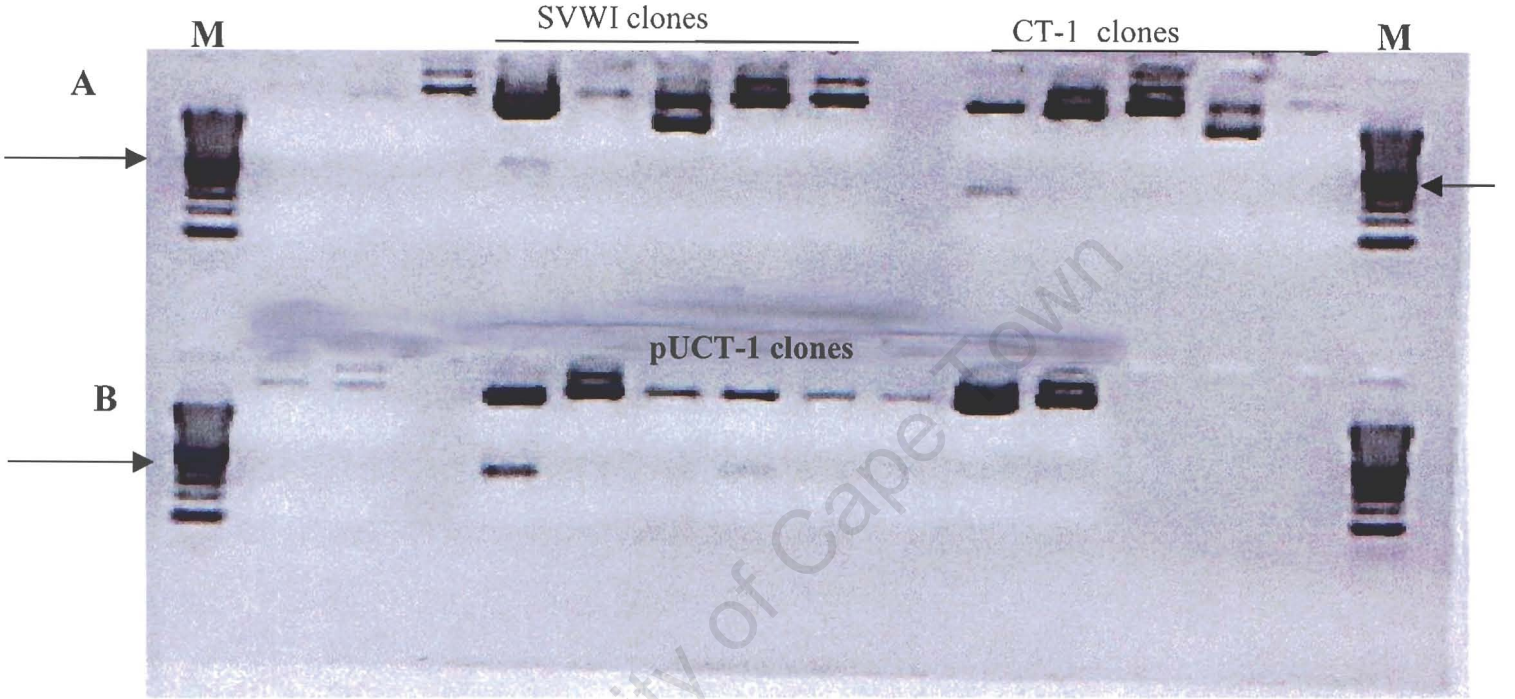


Figure 2.5 Analysis of cloned PCR products on a 2% agarose gel. Using blue/white selection of colonies, clones with inserts were picked and digested with EcoRI in order to release the inserts. Digested products were analysed on a 2% agarose gel. The arrows indicate 351bp inserts from (**A**) SVWI and CT-1 and (**B**) pUCT-1 clones. M represents the pBR322/HhaI marker. Clones with inserts were then prepared for DNA sequence analysis.

An interesting feature was the methylation of cytosines not present within CpG dinucleotides, although most of these residues were found in the methylated CpG clusters. Taking these additional methylated residues into account, the total methylation observed in SVWI-38 fibroblasts was even higher. However, these methylated non CpG sites could be artefacts (see Discussion). In an effort to understand mechanisms through which DNA methylation could be involved in COL1A2 transcriptional silencing, transient-transfection assays using methylated and unmethylated promoter constructs, described in chapter 3, were carried out. These experiments further supported the hypothesis that transcriptional silencing of the α2(I) procollagen gene

may be achieved via DNA methylation. For genes whose expression is influenced by DNA methylation, the main mechanisms for silencing have been reported to be mainly through changes in chromatin structure.

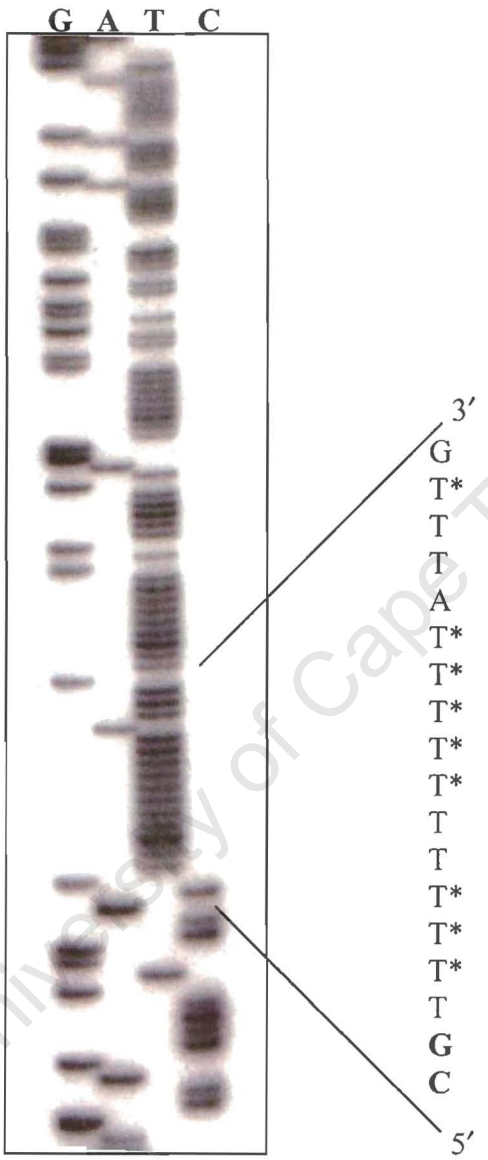


Figure 2.6 Methylation analysis of the bisulphite treated α2(I) proximal promoter in SVWI-38 fibroblasts. Unmethylated cytosines in the human α2(I) proximal promoter were deaminated to uracil using the sodium bisulphite procedure (section 4.3). The mutagenised promoter was then PCR amplified with primers specific for the bisulphite modified DNA, subcloned into pGEM-T-Easy plasmid and sequenced using M13 forward primers as described in section 4.3.4.3. Bisulphite converted cytosines were then amplified as thymine (shown above with *) while methylated cytosines remained unchanged (i.e. amplified as cytosine as illustrated with the representative CpG in bold). Some of the unchanged cytosines, which did not occur within any CpG dinucleotides, were found to be part of the vector sequence.

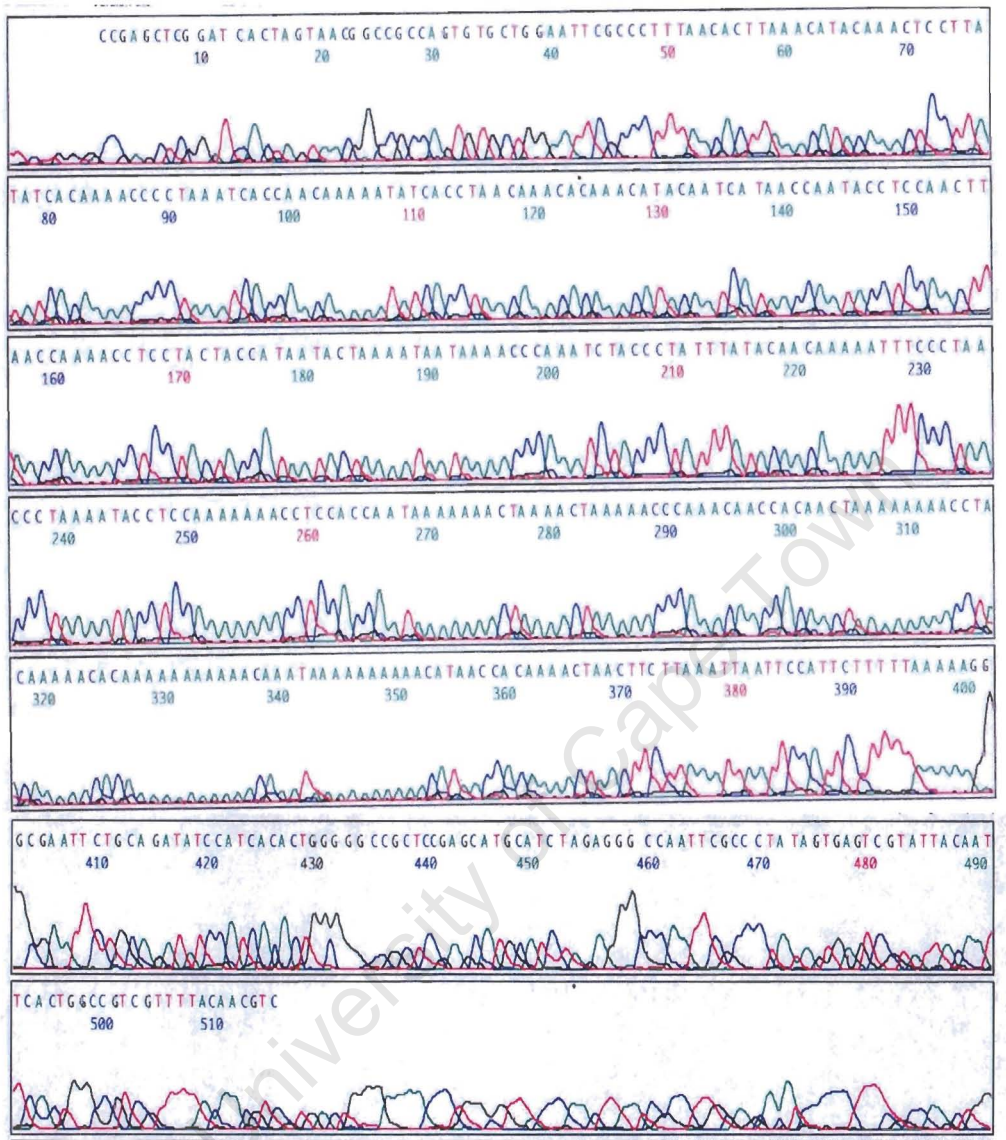


Figure 2.7 DNA sequence of the α2(I) procollagen gene proximal promoter from CT-1 cells (collagen expressing cells). DNA was extracted from CT-1 cells, treated with bisulphite and PCR amplified as explained in sections 4.2 and 4.3. The amplified fragment was then subcloned into pCR®2.1-TOPO vector and sequenced with an automated sequencer ABI PRISM Model 377 DNA Sequencer. The antisense strand is represented above with all cytosines converted to thymines. The 17 CG dinucleotides examined were all unmethylated.

2.3 DISCUSSION

Several studies have implicated DNA methylation in the regulation of type I collagen gene expression (Parker *et al.*, 1982; Smith & Marsilo, 1988; Thompson *et al.*, 1991; Chan *et al.*, 1991; Guenette *et al.*, 1992; Rhodes *et al.*, 1994; Sengupta & Smith 1998;). Most of these studies have clearly correlated the methylation status of regulatory elements with transcriptional repression of the gene. Methylated DNA-binding proteins have been implicated to play a central role in mediating this repression (Sengupta *et al* 1999). The mechanism(s) by which DNA methylation affects expression of most genes, including the type I collagen genes, are still only partly understood despite the recent evidence on gene silencing and chromatin structure modulation by DNA methylation (reviewed by Razin, 1998; Wolffe 1998 Baylin & Herman, 2000; Szyf, 2001). The fact that DNA methylation does not affect all genes in the same manner highlights the complex role it plays in gene inactivation and disease progression.

Initial studies on the involvement of DNA methylation in the regulation of the type I procollagen genes (Parker *et al.* 1982, Smith *et al.*, 1988, and Thompson *et al.*, 1991) have pointed towards promoter methylation as a possible mechanism for the regulation of this gene. Furthermore, Sengupta *et al.*, (1998) have suggested that methylation sites in the initiation region and the first exon could be crucial for inhibiting transcription of the murine α2(I) collagen promoter.

Transcription of a number of procollagen genes is inhibited by DNA methylation, these include the procollagen IV genes (Burbelo *et al.*, 1990, Kopp *et al.*, 1997) and rat α1(I), (Waye *et al.*, 1989). Earlier studies, however, could not establish any relationship between methylation patterns and procollagen gene expression, more specifically the α2(I) procollagen gene. The methylation status of the chicken α2(I) procollagen gene in transformed cells (Mckeon *et al.*, 1982; 1984), and the human α 2(I) and α1(III) (Chandler *et al.*, 1986) procollagen genes did not show any correlation with gene expression. This apparent discrepancy in the role of DNA methylation in transcriptional regulation of the procollagen genes could be attributed mainly to the limitations in the techniques employed in the early investigations (reviewed by Harrison *et al.*, (1998).

Figure 2.8: **Summary of the methylation profile of the human α2(I) procollagen gene**. Cytosine residues in CpG dinucleotides within the COL1A2 proximal promoter from collagen expressing (CT-1) and non-expressing (SVWI38) cells were compared. The published nucleotide sequence is in bold and the CpG dinucleotides studied are represented in pink. The cytosine residues found methylated in the cell lines are depicted in red while the unmethylated cytosine residues are shown in green. The arrow indicates the transcription start site.

The recently developed bisulphite technique overcomes most of the drawbacks of the earlier procedures used in the identification of 5-methylcytosine residues (5mC) (Frommer *et al*, 1992 and Clark *et al* 1994). This method reveals the methylation status of every cytosine residue in the gene of interest. In light of recent studies that have demonstrated the relationship between the density of methylated cytosines in promoter regions and methyl binding proteins (both sequence specific and non-sequence specific) in inhibiting transcription (reviews by Tate & Bird, 1993; Kass *et al.*, 1997) it is essential to determine the methylation status of every cytosine residue in the gene. Although several procollagen genes have been shown to be regulated by DNA methylation, detailed methylation analysis of the procollagen promoters has not been reported.

This study was therefore designed to evaluate the methylation pattern of the human α2(I) procollagen gene (COL1A2) proximal promoter in collagen-producing and non-collagen producing fibroblasts. This study presents the first detailed genomic methylation analysis of COL1A2 proximal promoter in fibroblasts that express the α2(I) procollagen gene (CT-1) and its SV40-transformed counterpart in which the gene is not expressed. The striking difference observed in the methylation patterns of the promoter region in the two cell lines is noteworthy. Of the 9 methylated CpG dinucleotides within the SVWI-38 α2(I) procollagen proximal promoter, 6 occurred mainly around the transcription initiation site. These results are in agreement with recently published results for the mouse α2(I) procollagen promoter by Sengupta and Smith (1998). These authors have used *in vitro* methylated and/or mutated CpG sites at positions +7 and +23 in the first exon, in *in vitro* transcription assays to show that the methylation sites in the vicinity of the transcriptional start sites in the promoter and the first exon play a critical role in inhibiting transcription. Recent reports have also revealed that not all methylation sites in the promoter region are critical for transcriptional regulation, those occurring in the vicinity of transcriptional start sites appear to be more important (Graessmann *et al.*, 1994; Archey *et al.*, 1999).

Hypermethylation of the COL1A2 promoter in SV40-transformed WI-38 fibroblasts has been reported in previous studies using the HpaII and MSp I isoschizomer pair to cleave at unmethylated and methylated CCGG sequences respectively (Parker *et al* 1982). The extent of CpG methylation could also be sufficient to interfere with the assembly of the transcription apparatus. The methylated CpG sites in the present study were located at positions –179, -171, -36, -17, +23, +48, +59, +61, +65, +81 and +84, with the majority clustered around the TATA box. It is highly likely that these methylated CpG sites interfere with the formation of the preinitiation complex and consequently lead to reduced transcription of the COL1A2 gene (Parker *et al*, 1989, 1992).

The only transcription factor capable of sequence-specific binding to the TATA box is the transcription factor IID (TFIID) complex, through its TATA binding protein (TBP) and TBP-associated factors (TAFs) (Dynlacht *et al*, 1991 and Burley & Roeder, 1996). Binding of TFIID facilitates the step-wise addition of the other factors, TFIIA, TFIIB, TFIIF-RNA polymerase II, TFIIE and TFIIH. This assembly involves a conformational

change and priming of chromatin to allow proper binding of the preinitiation complex, both of which may be regulated by acetylation or methylation of nucleosomal cores (Turner *et al*, 1992; Ramirez *et al*, 1995; Weber *et al* 1997; Horikoshi *et al*, 1988; Chi & Carey 1996). In addition, the preinitiation complex of most genes extends to nucleotide +30 to achieve the stability it requires for transcription initiation. Thus methylation of CpG sites within this region could interfere with preinitiation complex formation and/or stability and consequently decreased transcription initiation. This hypothesis is also supported by previous reports that showed that transcriptional repression of COL1A2 is not due to mutation, or deletion of the gene (Parker *et al*., 1989). Methylation has been shown to interfere with transcription factor binding to sequences of interest in a number of genes. In addition to inhibiting transcription factor binding by steric hindrance, as was suggested above, other mechanisms that could explain methylation-induced inhibition of COL1A2 transcription are:

a) Binding of a methyl-C-binding protein, such as MeCP2, enables recruitment of histone deacetylases and mSin3 transcription repressor complex to promoters (Nan *et al*, 1998, Jones *et al*, 1998) which deacetylate the histone protein tails resulting in a compact chromatin that is inaccessible to transcription factors and preventing the formation of the preinitiation complex.
b) The presence of a collagen specific repressor whose binding is enhanced by de novo methylation of the CG dinucleotides in promoter regions of transformed cells, consequently impeding formation of the transcriptionally active preinitiation complex (Sengupta & Smith, 1999). This could be the initial step in inactivation of transcription through chromatin condensation.
c) Exclusion of transcription activating protein complexes with intrinsic acetylase activity from the promoter region by the density of methylated CpG dinucleotides (Baylin & Herman 1999).

The above-mentioned possibilities are not necessarily mutually exclusive; instead they could complement each other, thereby providing a molecular lock as suggested by Siegfried and Cedar (1997). Although there is currently no direct evidence for the methylation-mediated transcriptional repression suggested by Sengupta & Smith (1998), studies on the role of transcription factors in COL1A2 gene repression have reported an increased binding of a repressor protein in cells that do not express COL1A2 (Collins *et*

al 1997). Recent reports seem to indicate that DNA methylation alone does not achieve gene repression, but that there exists substantial cross talk between DNA methylation and other transcriptional processes and post synthetic modifications such as acetylation (Nan *et al* 1998, Jones and Wolffe 1999). Further work that links the studies on the presence of increased repressor protein binding in transformed fibroblasts and the observed de novo methylation in the same cells could unravel yet another mechanism through which DNA methylation could mediate tissue-specific expression.

Another interesting feature observed in this study is the presence of methylated cytosines at non-CpG sites. A number of studies using nearest neighbour and restriction enzyme analysis of genomic DNA have previously suggested the presence of methylation at non CpG sites (Woodcock *et al*, 1987, Grafstrom *et al*, 1985, Nyce *et al*, 1986). A considerable amount of studies have recently reported non-CpG methylation in plants, animals and fungi using the bisulphite genomic sequencing technique. Several studies have reported non-CpG methylation sites in cancer cells and in repeated DNA sequences (Stirzaker *et al*, 1997 and Clark *et al*, 1997). The significance of this finding is still unclear, although it is conceivable that the DNA methyltransferase (DNAMTase) or a novel DNA methyltransferase is responsible for non-CpG methylation. Another explanation could be that these sites are artefacts. Harrison *et al* (1998) and Rother *et al* (1995) have reported that cytosines adjacent to methylated CpG sites can be partially resistant to bisulphite conversion, thus resulting in methylation artefacts. The latter seems more likely as most of the non-CpG methylation sites observed are adjacent to methylated CpG sites. Based on these findings one can speculate that the non CpG methylation sites observed in SVWI-38 cells are essentially due to the "nearest neighbour" influence during the bisulphite conversion reaction.

CHAPTER 3

THE EFFECT OF IN-VITRO C_PG METHYLATION ON THE ACTIVITY OF THE HUMAN α2(I) PROCOLLAGEN PROMOTER.

3.1 INTRODUCTION

Several studies have conclusively established a reciprocal relationship between CpG methylation in promoter regions and transcriptional activity of the gene (for reviews see (Doerfler, 1983; Bird, 1984, Chandler and Jones, 1985; Bird, 1986; Kass *et al.*, 1997; Jones, 1999; Jones and Laird, 1999). These studies suggest that the inactivation of genes is mainly due to a change in the sequence-specific DNA methylation patterns in their promoters. The change in the DNA methylation status of promoter regions is also increasingly being recognised as one of the most frequent mechanisms of loss of gene function in human cancers. Inactivation of tumour-suppressor genes due to increase in methylation of the promoter region is an excellent example of such a change (reviewed by Jones (1996), and Baylin *et al.*, (1998). Clusters of CpG dinucleotides, also known as CpG islands, are associated with the promoters of many human genes (Antequera and Bird, 1993). Consequently CpG islands have been frequently used as markers for the presence of genes in uncharacterised genomic DNA (Lindsay and Bird, 1987). Furthermore, these clusters have provided evidence for the role of methyl-CpG in repressing gene expression. With the exception of housekeeping genes, hypomethylated CpG islands in promoter regions are associated with gene activation while hypermethylated CpG's are associated with inactive genes (reviewed by Doerfler, 1993).

Presently, no direct evidence for the mechanism(s) that might allow the prevalence of unmethylated promoter CpG islands on autosomal chromosomes in normal cells and altered patterns in transformed cells is available. However, recently Antequera and Bird, (1999), have proposed a new model to explain the origin and the maintenance of these

CpG islands in mammalian cells. Their hypothesis is based on the fact that promoters associated with CpG islands are highly active in the germ line (Macleod *et al.*, 1998) and colocalise with DNA replication origins (Delgado *et al.*, 1998) and suggests that CpG islands are formed only during early embryogenesis by origin-associated promoters. That is, in totipotent cells, only active promoters can recruit proteins that initiate DNA replication, thereby excluding *de novo* methyltransferases and resulting in the formation of methylation-free footprints that coextend with the initiation loop intermediate. Once these stretches of methylation-free DNA are formed, they are then transmitted through somatic cell divisions by a maintenance methyltransferase.

Most CpG island containing genes that are expressed in highly differentiated cells such as the human α-globin and the mouse major histocompatibility complex class II I-Aβ genes (Macleod *et al.*, 1998) are expressed in embryonic cells, while the non-CpG island genes i.e. human β-globin (Kitsberg *et al.*, 1993) and rat aldolase B gene (Zhao *et al.*, 1994) are silent in germ cells. These results therefore are compatible with the model proposed above, but experimental tests are still required to further validate it. This model and other results have underscored transcriptional activity as the key control point necessary for regulating gene expression. The fact that transcription initiation, activation, and repression involve specific interaction(s) of a wide spectrum of different *cis*- and *trans*- acting elements further indicates that interference with the recognition sequences and affinity of the factors to each other is paramount in repressing gene activity.

Transfection experiments of promoter constructs linked to a reporter gene, and lately those of minichromosomes (Hsieh, 1997), have enabled better understanding of the possible role of DNA methylation in transcriptional repression. These experiments have shown that in addition to the effect of chromatin structure, the extent of transcriptional repression on a given promoter is largely dependent on the proximity of methyl-CpGs to the promoter, CpG methylation density and the strength of the promoter (Boyes and Bird, 1992; Nan *et al.*, 1997). Recently, more specialised studies have pointed out that the effect of the location of methylated CpG dinucleotides may be even more pronounced than the CpG frequency on methylation-mediated transcriptional repression. Evidently the main determining factor in methylation-dependent transcriptional repression is the presence of methylated CpG dinucleotides in the promoter (Davey *et al.*, 1997).

The human α2(I) procollagen gene (COL1A2) is expressed in a tissue-specific manner and is also very G/C rich, containing several CpG islands. These features are generally characteristic of genes regulated by DNA methylation, and COL1A2 would be a good candidate for transcriptional regulation by DNA methylation. Previous studies have shown that collagen synthesis is greatly reduced during transformation, and DNA methylation has been implicated in the loss of expression of various procollagen genes (Thompson *et al.*, (1991), Chan *et al.*, (1991), Guenette *et al.*, (1992), Rhodes *et al.*, (1994). Studies by (Parker *et al.*, 1982) have reported the loss of COL1A2 gene expression, in Simian virus40 (SV40)-transformed fibroblasts, and have further reported that this loss in expression of the α2(I) collagen gene was neither due to mutations nor deletions in the promoter region, but due to differential transcription factor binding activity (Parker *et al.*, 1989).

Having established the methylation profile of COL1A2 proximal promoter (chapter 2), we determined the sensitivity of promoter activity to CpG methylation. The present study examined the *in vitro* effect of DNA methylation on the activity of the COL1A2 promoter as a possible mechanism for transcriptional repression of the gene. To this end, COL1A2 promoter activity in methylated and mock-methylated constructs was analysed by transient transfection assays. COL1A2 promoter constructs driving a chloramphenicol acetyl transferase (CAT) reporter gene were methylated with M.SssI methylase, which methylates all CpG dinucleotides. Completeness of methylation was tested with the restriction enzyme HpaII whose recognition sequence CCGG is sensitive to methylation of the internal cytosine. The density of CpG dinucleotides varied greatly among the COL1A2 constructs as shown in Figure 3.1

Our results show that the human α2(I) promoter is sensitive to DNA methylation and that COL1A2 transcriptional repression is methylation-mediated. The methylation effect was greater more in the larger constructs, implying that the extent of repression depended on the density or amount of methyl-CpG's.

3.2 RESULTS

The direct effect of *in vitro* methylation on transcription of the human COL1A2 gene was determined by promoter activity in transient transfection experiments. A series of constructs with different sized inserts of methylated and unmethylated COL1A2

promoter cloned upstream of the bacterial chloroamphenicol acetyl transferase (CAT) reporter gene (Figure 3.1), were transiently transfected into gamma-radiation-transformed human embryonic lung WI-38 fibroblasts (CT-1). CT-1 was used in place of wild type WI-38 fibroblasts because they are much easier to transfect and the expression of the α2(I) collagen gene is similar in both cell lines. In addition, the methylation status of the COL1A2 promoter had also been determined (chapter 2).

As stated above, some of the parameters that determine the extent of transcriptional repression by DNA methylation are the density of methyl CpGs on the promoter, and the location of the methylated sites relative to the transcription start site. For this purpose, four COL1A2 promoter constructs, namely COLCAT -1700/+54, COLCAT -343/+54, COLCAT -120/+54 and COLCAT -107/+54 were used in analysing the effect of DNA methylation on the COL1A2 promoter. These constructs harboured different amounts of CpG dinucleotides although the frequency with which the recognition sequence for the methylation sensitive restriction enzymes HpaII occurred did not differ dramatically. Thus the result of the restriction digestion of mock methylated constructs does not show any drastic differences between the constructs. However, the transient transfection results clearly indicate the significance of the differences in the total CpG content. These results therefore, further underscore the limitations of the use of methylation-sensitive enzymes in determining the presence of methylated cytosine residues. Table 3.1 below summarises the lengths of the α2(I) procollagen promoter CAT constructs and the total number of CpG dinucleotides in each fragment. 11 of these CpG sites occurred within transcription factor binding sites.

Table 3.1 The nucleotide position of promoter fragments relative to the transcription start site and their total CpG content. These constructs were cloned upstream of a CAT reporter gene.

Plasmid constructs	Nucleotide positions	Total CpG dinucleotides
-1700COLCAT	-1700 to +54	72
-343COLCAT	-343 to +54	17
-120COLCAT	-120 to +54	9
-107COLCAT	-107 to +54	5

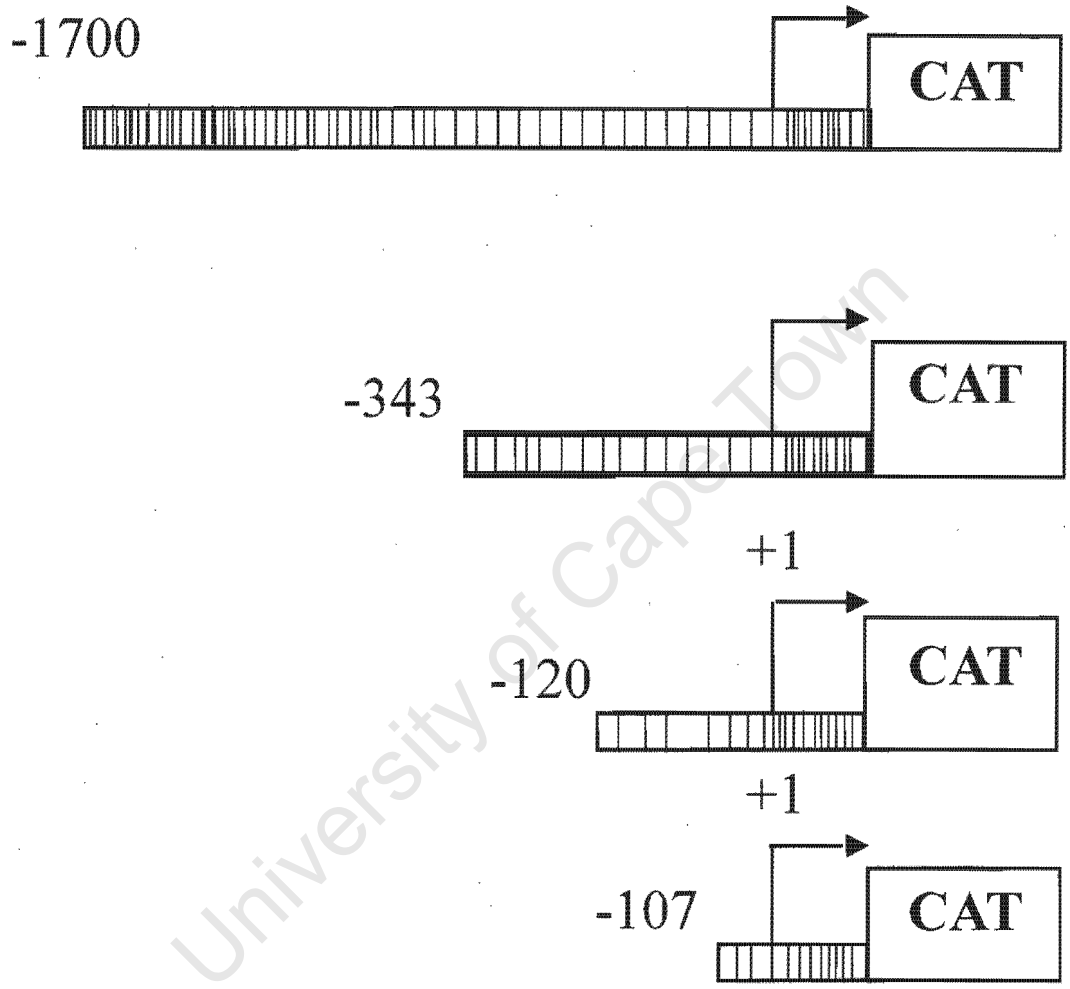


Figure 3.1: Schematic map of the human COL1A2 promoter constructs. Fragments of the COL1A2 promoter extending to -1700, -343, -120, or -107 were cloned upstream of chloroamphenicol acetyl transferase (CAT) reporter gene. The vertical lines indicate the CpG dinucleotides for each construct. There are 72 for –1700; 17 for –343; 9 for –120 and 5 for –107. The arrow indicates the transcription start site.

3.2.1 In vitro methylation

The COL1A2 promoter constructs were methylated *in vitro* using the bacterial CpG methylase, M.SssI that methylates all cytosine residues within the double stranded recognition sequence 5′...CpG...3′. However, unlike the mammalian enzyme, all cytosine residues in the dinucleotide recognition sequence are efficiently methylated in both unmethylated and hemimethylated substrates. Completeness of methylation was determined by the protection of the fragment from digestion with the methylation sensitive restriction enzyme HpaII, which is sensitive to methylation of the internal cytosine in the recognition sequence CCGG. The methylated constructs were found to be 100% resistant to HpaII digestion compared to the unmethylated constructs (Figure 3.2).

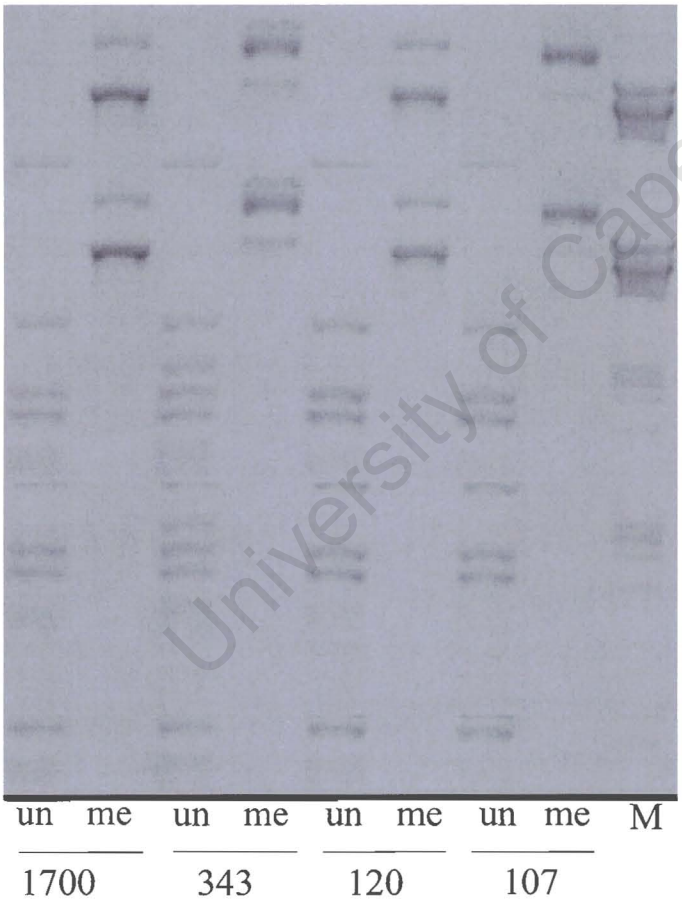


Figure 3.2: Methylation of the promoter constructs. The -1700bp, -343bp, -120bp, & -107bp COL1A2/CAT promoter constructs were methylated with M SssI -as described in section 4.3.9. Methylated (me) and mock methylated (un) constructs were digested with the methylation sensitive restriction enzyme, HpaII that recognizes the CCGG sequence and analysed on a 2% agarose gel. The methylated constructs are fully protected from digestion while the unmethylated constructs were not. PvuII digested λ-DNA marker is shown in lane M.

3.2.2 DNA methylation and COL1A2 promoter activity

Methylated and mock-methylated constructs were transiently transfected into CT-1 cells. 48 hours after transfection, the promoter activity of the methylated and mock-methylated promoter constructs was determined by measuring the CAT activity. The pCMVβGal plasmid was cotransfected with the promoter constructs and used to normalise the CAT activity for transfection efficiency. Methylation of all the promoter constructs resulted in a significant reduction in promoter activity (Figure 3.3).

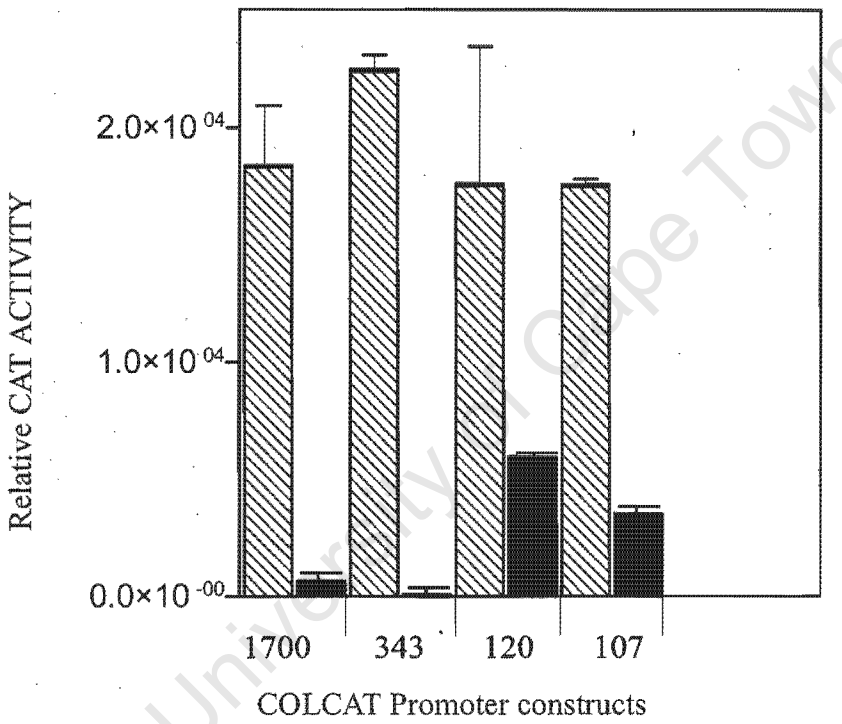


Figure 3.3 Effect of CpG methylation on promoter activity. Methylated and mock methylated –1700/+54, -343/+54, -120/+54, and –107/+54 COL1A2/CAT promoter constructs depicted in figure 3.1 were transiently transfected into CT-1 fibroblasts. pCMV β-gal was contransfected with the promoter constructs to normalise for transfection efficiency. Each experiment was carried out in triplicate, the relative CAT activity represents the mean values and standard deviations. Methylated constructs are represented by the black bars (■) and the unmethylated constructs by the shaded bars (▨). The CAT activity in methylated promoter constructs was greatly reduced, suggesting that the COL1A2 promoter is sensitive to methylation.

The larger methylated constructs (COLCAT-1700/+54 and COLCAT-343/+54) exhibited minimal activity compared to the smaller constructs (COLCAT -120/+54 and COLCAT-107/-+54). The mock-methylated constructs on the other hand showed much higher promoter activity, demonstrating that the human COL1A2 promoter is sensitive to DNA methylation, as has been shown for the other collagen genes (Rhodes and Breindl, 1992), (Kopp *et al.*, 1997). Furthermore, the extent of COL1A2 transcriptional repression is dependent on the density of methyl-CpGs. Both the COLCAT –1700/+54 and COLCAT –343/+54 promoter constructs have a higher CpG density and transcription is inhibited to a greater extent than in the constructs COLCAT-120/+54 and COLCAT-107/+54 with a low CpG density. However, all the constructs were found to be sensitive to methylation, including the smaller constructs. Moreover, the CpG sites of the smallest promoter construct (COLACT-107/+54) are not within any of the known transcription factor binding sites, thus CpG methylation *per se* was not expected to have any significant effect on promoter activity. These observations could therefore indicate a synergistic effect between the methylated CpGs and other methylation-dependent transcriptional regulators.

Interestingly, recent reports have shown that the COL1A2 transcription start site, from nucleotides –1 to +13, contains a sequence that is homologous to the 14-bp consensus sequence of a family of closely related proteins known as methylated DNA-binding protein (MDBP) or regulatory factor for X box 1-4 (RFX) (Huang *et al.*, 1984; Zhang *et al.*, 1993, Reith, 1994; Emery *et al.*, 1996; Sengupta *et al.*, 1999). These MDBP/RFX proteins are distinct from most methylated DNA-binding proteins including MeCP1, MeCP2, MDBP-2-H1, DBP-m (Lewis *et al.*, 1992; Schwarz *et al.*, 1997; Nan *et al.*, 1998) in that they bind methylated DNA with high affinity in a sequence-specific manner within the consensus sequence 5′-RT(m^5C/T)RYYA(m^5C/T)RG(m^5C/T)RAY-3′ where m^5C/T represents 5-methylcytosine or T, R represents G or A and Y represents C or T (Zhang *et al.*, 1989). Using competition electrophoretic mobility shift assays (Sengupta *et al.*, 1999) have demonstrated that the MDBP/RFX binds its consensus sequence in the COL1A2 5′ region in a methylation-dependent manner. These authors also showed that the CpG site critical for methylation-dependent high affinity binding of the MDBP/RFX to its consensus sequence occurred in the same position in COL1A2 as in those of previously described methylation-dependent MDBP/RFX sites (Zhang *et al.*, 1990).

These results therefore underscore the findings in chapter 2, where using the sodium bisulphite genomic sequencing technique, we have mapped the methylated cytosine residues in the endogenous gene to nucleotides between –40 and + 90 of the proximal promoter in transformed cell lines where collagen production is abolished. Evidently, methylation of CpG dinucleotides within the first exon of human collagen α2(I) gene plays a significant role in mediating its transcriptional repression.

3.3 DISCUSSION

Several studies have shown loss of α2(I) collagen gene expression to be associated with cellular transformation and carcinogenesis. Furthermore, transformed cells and a variety of human cancers are characterized by regional hypermethylation and global hypomethylation in genes critical for cancer progression (Baylin & Herman 2000). Drastic changes in gene expression result from these altered DNA methylation patterns. In most cases, tumour-suppressor genes are silenced due to abnormal hypermethylation of the 5′ promoter regions, while oncogenes are unmethylated and over-expressed. Documented examples of tumour suppressor genes and oncogenes include the c-myc, c-fos, p16, Rb, globin genes-which were among the first to be reported- with abnormal methylation patterns (McGhee & Ginder, 1979, Shen & Maniatis, 1980, Singal *et al.*, 1997). Thus a number of studies have sought to determine whether there is abnormal DNA methylation in the 5′ region of COL1A2 and whether it might play a role in down-regulating the expression of this gene in transformed cells.

Transient transfection assays of *in vitro* methylated expression plasmids have been used extensively in establishing the role of DNA methylation in various collagen genes (Thompson, *et al.*, 1991, Guenette *et al.*, 1992, Rhodes *et al.*, 1994, Sandberg & Schalling, 1997, Kopp *et al.*, 1997 Umezawa, *et al.* 1997,Kudo & Fukuda, 1995, and Kuramasu *et al.*, 1998). These assays have also been used in studies on the estrogen and progesterone receptors (Lapidus *et al.*, 1996 & 1998), retinoblastoma protein (Ohtani-Fujita *et al.*, 1993), E-cadherin (Graff *et al.*, 1995), BRCA1 (Mancini *et al.*, 1998), keratin 18 gene (Umezawa, 1997), the MDR1 gene (Nakayama *et al.*, 1998), and others. The above-mentioned studies and many others have reported that the promoter-driven expression of a reporter gene construct is completely abolished when *in vitro* methylated expression plasmids are transiently transfected into cultured cells. This silencing of expression by *in vitro* methylation is not reversed by treatment with Trichostatin A,

implicating repression by means other than recruitment of histone deacetylase activity (Yoshida M *et al* 1995).

Although transient transfections do not enable chromatin assembly, rapid repression of reporter genes, similar to that of microinjected chromatin (Buschhausen *et al*, 1985; 1987) is observed. These seemingly conflicting results indicate therefore that DNA methylation mediates repression through several routes (discussed in detail in Chapter 1). The current proposed models of repression include all the key components in the control of transcription, namely the transcription machinery and structural proteins such as histones and DNA. A "shuttle-system" or three-way connection that links DNA methylation, chromatin structure and gene activity explains the repressive effects of in vitro- and in vivo-mediated DNA methylation on transcription. For most tissue-specific genes, methylation of their CpG islands in promoter regions has been conclusively correlated with transcriptional repression of these genes. For instance, the mismatch repair (MMR) genes show an inverse relationship between promoter methylation and gene expression. Of the two major MMR genes known to date, the loss of hMLH1 expression frequently observed in sporadic colon cancer tissues and cell lines has been correlated with cytosine methylation in its promoter (Kane *et al.*, 1997, Cunningham *et al.*, 1998; Herman *et al.*, 1998; Deng *et al.*, 1999). DNA methylation is therefore one of the most powerful mechanisms for suppressing gene activity.

Methylation-mediated transcriptional repression of the promoter is mainly due to alterations in interactions between cis- and trans- acting elements and to some extent also between trans- elements where complex formation is necessary for transcription. The presence of the methyl group either sterically hinders binding of transcription factors or perhaps the methyl-C-binding of proteins inhibit binding of transcription factors or it induces changes in chromatin structure. The use of naked DNA (in vitro and transient assays) therefore is predominantly when the effect of the methyl-CpG location and density with respect to the promoter is being investigated (Keshet *et al*, 1985, Murray & Grosveld, 1987, Kass *et al*, 1993).

However, not all tissue-specific genes are repressed when methylated. For some genes such as the H-2K gene, hyper-methylation of the promoter region is necessary for gene

activation (Tanaka *et al.*, 1983). Graessmann *et al.*, (1994) have thus classified genes into three categories based on their response to DNA methylation. These are:

i. The methylation-sensitive genes in which hypermethylation represses activity
ii. The methylation insensitive genes where methylation does not block expression, and
iii. The methylation dependent genes in which hypermethylation is associated with activation.

From our results we can therefore conclude that the α2(I) procollagen gene is a methylation sensitive gene and that DNA methylation plays an important role in transcriptional repression of this gene. The experimental data described above, taken together with results described in chapter 2, clearly indicate that both the density of methylation and the site of methylation are crucial in methylation-dependent transcriptional repression of COL1A2. Furthermore, the most critical CpG sites have been found in close proximity to the transcription start site. This conclusion has been confirmed by other studies (Sengupta and Smith, 1998; Sengupta *et al.*, 1999). These authors have shown conclusively that the CpG sites between positions –25 to +30 of the α2(I) promoter bind sequence-specific nuclear proteins when methylated. They also reported that it is the CpG site at +7 relative to the transcription start site, which is critical for this methylation-mediated repression. Moreover, this CpG falls within the consensus sequence of the MDBP/RFX, which confers methylation-dependent binding.

Previous findings have proposed that the protein that binds to the Collagen-Modulating Element (CME) (indicated in Figure 2.3) is a repressor found in non-collagen producing cells (Learner, 1998; Masemola, 1999). It is possible that DNA methylation is involved in repression of the COL1A2 gene via methylated DNA-binding proteins described by Sengupta *et al.*, (1999). Future work should investigate the mechanisms through which this repression could be mediated. Immunoprecipitation and glutathione S-transferase (GST) pulldown assays may establish if the MDBP/RFX co-precipitates with a CME-containing corepressor complex. It is also possible that modulation of chromatin structure is involved in COL1A2 transcriptional repression. Perhaps the changed chromatin structure induced by DNA methylation facilitates interactions of the CME repressor proteins with other potent transcription repressors or it may be shielded from modifications necessary to reduce its activity.

There is still a lot of uncertainty regarding whether DNA methylation plays a causal or consequential role in tumourigenesis. The findings in this thesis further complicate the role of DNA methylation during cellular transformation as the two cell lines studied were both transformed and yet they differentially express COL1A2. These results would therefore argue for a consequential role for DNA methylation in cellular transformation as it is evident that transformation and the methylation status of the COL1A2 gene are not linked. However, transcriptional activity and DNA methylation are linked. Although significant progress has been made in understanding the molecular mechanisms by which DNA methylation mediates transcriptional repression of the α2(I) procollagen gene and other genes (see Razin, 1998; Wolffe, 1998 for reviews), it is evident that more work still has to be done to further unravel the link between DNA methylation and cellular transformation.

CHAPTER 4

MATERIALS AND METHODS

4.1 CELL CULTURE

Normal human embryonic lung fibroblasts (WI-38), simian virus 40- (SVWI-38) and γ-radiation (CT-1) transformed cells were grown in Dullbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat inactivated fetal calf serum (FCS), 100µg/ml streptomycin and 100u/ml penicillin. The cells were maintained in a humidified incubator at 37°C in 5% CO_2 and 95% air. The medium was replaced every two days until confluency was reached.

For harvesting, the culture medium was removed, the cell layer rinsed with PBS and incubated with 10ml of 0.5% trypsin and 0.2% EDTA in Ca^{2+}/Mg^{2+} -free phosphate buffered saline (PBS) at 37°C. Cell detachment was monitored microscopically, trypsinisation halted by the addition of an equal volume of complete medium, the cells transferred to a centrifuge tube and centrifuged in a bench-top centrifuge for 1 minute. Cells were resuspended in 5ml of DMEM and an aliquot counted in a Coulter counter (Coulter Electronics). 1×10^8 cells were re-seeded into tissue culture dishes for further propagation, DMSO was added to a final concentration of 10% (v/v) to the remaining cells and stored in liquid nitrogen.

4.2 DNA EXTRACTION

Genomic DNA was extracted from cells using the proteinase K/SDS DNA isolation procedure described by (Gross-Bellard *et al.*, 1973) and modified by (Ramsay *et al.*, 1982) and (Laird *et al.*, 1991). Trypsin/EDTA was used to detach the cells as described in section 4.1, the supernatant was discarded and the cells resuspended in 1vol (about 0.5ml) digestion buffer (100mM Tris-HCl pH 8.5, 5mM EDTA, 0.2% SDS, 200mM NaCl,). 100µg Proteinase K was added and the samples incubated for 12 to 18 hours at 50°C with constant agitation in tightly sealed tubes. Rnase A was added to a final concentration of 10µg/ml and the sample mix incubated for a further 30 minutes at 50°C. One volume of isopropanol was added to the viscous lysate and mixed thoroughly to precipitate the DNA.

To recover the DNA, the tip of a pasteur pipette was used to spool the DNA from the solution, the excess isopropanol was dabbed off and the DNA transferred to a sterile 1.5ml eppendorf tube. The precipitate was rinsed with 70% ethanol and resuspended in 500 to 1000µl sterile distilled water. DNA was quantitated spectrophotometrically and then used to verify the purity and integrity of isolated DNA on a 1% agarose gel.

4.3 DNA METHYLATION ANALYSIS

The bisulphite conversion protocol initially described by (Frommer *et al.*, 1992) and (Clark *et al.*, 1994) and modified by (Stoger *et al.*, 1997) was used in the bisulphite conversion of the human α2(I) procollagen promoter.

The following illustration summarises the bisulphite genomic sequencing reaction:

a) 5′CGATTGTACTC^{m}**CG**AT3′

b) 3′GCTAACATGAG**GC**$_{m}$TA5′

↓ Denaturation

a) 5′CGATTGTACTC^{m}**CG**AT

and

b) 3′GCTAACATGAG**GC**$_{m}$TA

↓ Bisulphite treatment

a) 5′UGATTGTAUTU^{m}CGAT

and

b) 3′GUTAAUATGAG**GC**$_{m}$TA

↓ PCR amplification with strand specific primers

a) 5′TGATTGTATTT**CG**AT

3′ACTAACATAAA**GC**TA

and

b) 5′GTTAATATGAG**GC**TA

3′CAATTATACTC**CG**AT

↓

Direct Sequencing and/or sequencing of cloned products

The outline above shows double stranded DNA, with the coding and noncoding strands represented by a) and b). Denatured DNA is crucial for efficient bisulphite deamination of the unmethylated cytosine residues to uracil. Methylated cytosine residues remain non-reactive (as 5mC) under the reaction conditions. Bisulphite modified strands are therefore PCR amplified separately with specific primers where the uracil is replaced by thymine and leaving 5-methylcytosine as cytosine. Sequencing of the PCR products (directly or after cloning) will result in only the methylated cytosines in the cytosine lane.

4.3.1 DNA denaturation using sodium hydroxide

DNA concentrations of 1-500ng to a maximum of 1µg have been recommended to ensure complete deamination (Gardiner-Garden and Frommer, 1994), (Razin and Shemer, 1995). As mentioned in section 2.1, the bisulphite reaction is highly single-strand specific thus the initial denaturation is critical to ensure complete deamination and high PCR yields. Freshly prepared 3M NaOH was added to a final concentration of 0.3M to 500ng double-stranded genomic DNA in a final volume of 10µl, incubated at 42°C for 30 minutes and transferred to 100°C for 1-2 minutes. The solution was thoroughly mixed, and overlaid with mineral oil (about 150µl). The α2(I) procollagen promoter has a high GC content and a high melting temperature making complete denaturation even more important.

4.3.2 Bisulphite treatment of DNA

Freshly prepared bisulphite solution (90 µl of a solution of 3.4M sodium bisulphite [Sigma], 1mM hydroquinone [Sigma], and adjusted to pH5.0 with 10M NaOH) was added and gently mixed by pipetting under the mineral oil. The reaction was allowed to proceed at 55°C for 6 hours in the dark. The modified DNA was then purified using MicroSpin S-200HR columns as described below.

4.3.3 Desalting & desulphonation of DNA

The MicroSpin columns are pre-packed with Sephacryl S-200 HR resin and equilibrated with Tris-EDTA (TE) buffer, pH 7.6. Prior to purification, the columns were vortexed to resuspend the resin and prepared as recommended by the manufacturer (Pharmacia Biotech). The bisulphite-treated DNA (100µl) was recovered from under the oil, loaded onto the columns and centrifuged at 735 × g for 2 minutes in a bench-top microcentrifuge to separate the excess bisulphite from the DNA. The purified DNA was collected in a clean eppendorf tube. Desulphonation of the reaction intermediate was carried out by the addition of freshly prepared 3M NaOH (to a final concentration of 0.3M) and incubated at 37°C for 20 minutes. The NaOH was removed by passing the DNA through a second MicroSpin S-200 HR column and 100µl of the flow-through containing the purified converted DNA was stored in aliquots for use as PCR templates.

4.3.4 HPLC cation exchange analysis

4.3.4.1 Preparation of standards for HPLC

In order to determine the elution times of the different bases, 1mg/ml base standard solutions of guanine, adenine, thymine, cytosine, uracil and 5-methylcytosine (Sigma) were dissolved in 0.1N HCl. A cocktail containing equal volume of each standard was also prepared.

4.3.4.2 Chemical hydrolysis of DNA for HPLC Analysis

50µg of bisulphite treated DNA and untreated genomic DNA from SVWI-38, CT-1 and WI-38 cells and cloned promoter (pUCT-1) was frozen in liquid nitrogen, dried under vacuum and dissolved in 100µl of 88% formic acid. 50µl was taken up in capillary

tubes, sealed and hydrolysed at 180°C for 25 minutes. To reduce the possibility of explosion upon opening the capillaries, the tubes were allowed to cool to room temperature and frozen in liquid nitrogen. The contents were emptied into eppendorf tubes, lyophilised and redissolved in distilled water. The dried bases were redissolved in 100µl water and lyophilised in order to ensure complete removal of the formic acid. The dried hydrolysis products were resuspended in 50µl 0.1N HCl for HPLC analysis.

4.3.4.3 Base composition analysis by HPLC

Samples were analysed on an Ultracil CX cation exchange column (4.6mm × 25cm) using a Beckman System Gold Model 166 detector. The column was equilibrated with filter-sterilised sodium acetate buffer (100mM sodium acetate, 5% methanol, adjusted to pH 3.35 with glacial acetic acid). The base standards were first loaded individually and then as a cocktail onto the column to determine the elution times of the different bases. The hydrolysate was then loaded and eluted with the sodium acetate buffer at a flow rate of 1ml/min; all bases were eluted within 20minutes. The effluent was monitored at 254nm. The base standards were also loaded separately to confirm the elution times.

4.3.5 Polymerase chain reaction (PCR) amplification

A pair of strand-specific primers designed to amplify bisulphite treated DNA of the COL1A2 promoter had the following sequence: forward primer (nt. positions -216 to -187) 5′-TTTTAAAAAGAATGGAATTAATTTAAGAAG-3′ and the reverse primer (nt. positions +104 to +133) 5′-TAACACTTAAACATACAAACTCCTTATATC-3′. Two rounds of PCR were performed in a Hybaid Omn-E thermal cycler to amplify the 351bp fragment of interest. The following cycle conditions were used: 94°C/2min for first cycle only, 94°C/20sec, 53°C/20sec, 72°C/1min for 29 cycles; 72°C for 5 min for the last cycle. Each PCR amplification was carried out in a 25µl reaction mixture containing 5µl of the bisulphite-treated DNA, 200µM dNTPs, 10mM Tris-HCl, pH 8.3, 1.5mM $MgCl_2$, 50mM KCl, 50µM primers and 1.25U *Taq* polymerase. 0.2ml MicroAmp PCR tubes were used with a hot-lid PCR machine. For the second round PCR reaction, 5µl from the first PCR was used under the same conditions. The two-round PCR is extremely sensitive and to detect contamination, two DNA-negative PCR controls were included

with the amplifications and extra precautions were taken with the sodium bisulphite conversion reaction.

4.3.5.1 Analysis of PCR products

The PCR products were analysed on a 2% agarose gel in TBE buffer (90mM Tris, 90mM Boric acid, 2.5mM EDTA, pH 8.3) containing 0.5µg/ml ethidium bromide. To the samples were added 2µl of DNA loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, and 25% glycerol). The samples were heated at 65°C for 5 minutes, loaded onto the gels and electrophoresed in 1×TBE buffer at 75 volts for 45 minutes. DNA was visualised under UV light. A 1000bp DNA ladder was used as a size marker.

4.3.5.2 QIAquick Gel Extraction of PCR products from Agarose Gels

The expected 351bp PCR product was excised from the gel and the DNA extracted using the QIAquick Gel Extraction Kit (Qiagen). 3-volumes of solubilization buffer (QC) was added to a 200mg-gel slice and incubated at 50°C for 10 min with vortexing every 2 min. Once the gel was completely solubilized, 1 vol. of isopropanol was added and the 800µl mixture loaded onto a QIAquick column, centrifuged for 1 minute at 13 000rpm, washed with buffer PE and the DNA eluted with 50µl distilled water. This purification step ensured that the PCR products were free of primers, *Taq* polymerase and dNTPs that would interfere with the subsequent cloning and sequencing reactions.

4.3.5.3 Subcloning of PCR products into plasmid vectors with 3′ overhangs.

Most thermostable polymerases with a non-template-dependent terminal transferase activity, such as *Taq* polymerase, add a single deoxyadenosine (A) to the 3′-ends of PCR products. Recently plasmid vectors with a single 3′ T overhang (pCR® II-TOPO or pCR® 2.1 TOPO and pGEM®-T & pGEM® -T Easy) that significantly increase the ligation efficiency of such products have been developed and are commercially available (Shuman, 1994). In this study, two different sequencing systems were used. The TOPO™ TA Cloning Kit is a highly efficient system for the cloning of PCR products. The linearized plasmid vector has a single 3′-T overhang and is covalently bound to topoisomerase thereby taking advantage of the topoisomerase ligation activity. The

resulting vector is referred to as an "activated" vector that enables direct ligation of the PCR product in the absence of ligase. 4µl gel purified PCR product was ligated into 1µl plasmid pCR2.1-TOPO for 4 minutes at room temperature, centrifuged briefly and placed on ice for 1 minute. After 5 minutes of incubation, the ligation product was transformed into highly competent 1×10^9 TOP 10F cells as recommended by the manufacturer (TA cloning kit, Invitrogen). The transformation and cloning efficiencies have been reported to significantly decrease if the ligation reaction is left for more than five minutes. Blue/white selection was used to pick recombinant clones from LB/ampicillin/100mM IPTG/2%X-gal plates. Plasmid DNA was extracted with the QIAprep 8 Miniprep kit (Qiagen) and sequenced using the ABI PRISM dye terminator cycle sequencing reaction described in 4.3.6.1.

The pGEM-T Easy Vector systems are plasmid-cloning vectors with improved ligation efficiency of PCR products in the presence of T4-DNA ligase. Promega's pGEM vectors pGEM-5Zf (+) and pGEM-T Easy vectors were linearized with EcoRV followed by addition of a 3′-T (thymine) overhang at the insertion sites. The optimum molar insert: vector ratio of 3:1 was used in these experiments. The ligation reaction with 2×Rapid Ligation Buffer, 50ng pGEM-T Easy, 3µl PCR product, 3U T4DNA ligase was incubated at 4°C overnight to increase transformation efficiency. The ligation reaction was used to transform 1×10^8 competent *E.coli* XL-1 cells and blue/white selection was used to select transformants containing an insert. The rapid plasmid prep protocol, described in 4.3.6, was used to isolate DNA for subsequent use in sequencing reactions.

4.3.6 DNA sequencing

4.3.6.1 ABI PRISM dye terminator Cycle Sequencing reaction

The plasmid DNA was digested with EcoRI and analysed on a 2% agarose gel to determine the presence of inserts before performing the sequencing reactions. The cycle sequencing protocol described by (McCombie *et al.*, 1992) was performed as described by the manufacturer (Perkin Elmer). The 10µl sequencing reaction mixture was prepared as follows: 500ng-1µg plasmid DNA (the amount of template used varied according to template sample, i.e. SVWI-38, pUCT-1, and CT-1) was added to 2pmol M13 reverse primer, and distilled water. 4µl terminator/Big Dye was added last and the following cycling conditions were used: 96°C for 1 min for the first cycle, followed by 96°C for 30

sec, 50°C for 15 sec, 60°C for 4 min for 25cycles. Dye was then removed by the addition of 1µl of 3M sodium acetate (NaOAc) pH 4.6 and 25µl 95% ethanol (EtOH). The tubes were vortexed and placed on ice for 10 minutes to facilitate DNA precipitation, centrifuged at ≥10,000 × g in a table-top microcentrifuge for 15 min and the supernatant carefully aspirated and discarded. The pellet was washed with 250µl of 70% EtOH, vortexed briefly and centrifuged for 5 minutes at ≥10,000 × g in a table-top microcentrifuge. The supernatant was carefully discarded, the pellet air-dried for 10 minutes and resuspended in 2µl loading buffer.

Automated analysis was performed on an ABI PRISM Model 377 DNA Sequencer. A 4.25% denaturing polyacrylamide gel was prepared as follows: 18g urea, 5.3ml 40% acrylamide, 25ml MilliQ water, 1g of amberlite mix and 5ml 10× TBE buffer stirred with warming until crystals dissolved, passed through Nalgene filter and then degassed for about 5 minutes. The solution was made up to 50ml with water and 250µl 10% APS and 35µl TEMED were added. The gel injection procedure was used to pour the gel and allowed to set for 2 hours before electrophoresis. 1.5µl of the samples was loaded and a 7-hour run was performed in 1× TBE buffer. Sequence analysis was carried out with computer programs of the ABI 377 DNA Sequencer Package and SIGNAL SCAN.

4.3.6.2 Direct Sequencing of PCR products

The T7 Sequenase version 2.0 reaction was performed as described by the manufacturer. The PCR products were treated with Shrimp alkaline phosphatase and exonuclease I to remove excess primers and dNTPs. 10pmol forward primer was added to the treated DNA and heat denatured for 3 min at 100°C. The mixture was cooled for 5 minutes on ice and 0.1M DTT, 0.5µl α^{32}P-dCTP, 3µl labelling mix, T7 Sequenase buffer and DNA polymerase were added by gentle pipetting to start the reaction. The reaction was incubated at room temperature for 5 minutes. To terminate the labelling reaction 4.5µl of the remix was transferred into four separate tubes each containing a specific chain-terminating dideoxynucleotide (ddATP, ddTTP, ddGTP, ddCTP). 4µl of stop buffer was added and the samples heated for 2 minutes at 80°C before loading on a 6% denaturing polyacrylamide gel.

4.4 PROMOTER CONSTRUCTS:

Previous investigators in the laboratory have constructed the promoter constructs depicted in Figure 3.1. The lengths of the α2(I) procollagen promoter CAT constructs and the total number of CpG dinucleotides in each fragment are summarized in Table 3.1.

4.4.1 Rapid plasmid DNA preparation

The presence of the correct-sized promoter inserts in the glycerol stocks was first determined on small amounts of DNA isolated by using the modified alkaline method (Birnboim, 1983). Small quantities of the plasmid constructs were prepared by inoculating 50µl of the glycerol stocks into 5ml Luria-Broth supplemented with 50µg/ml ampicillin and incubated at 37°C with shaking (250 rpm) for 12-18 hours (Alter and Subramanian, 1989),(Liszewski *et al.*, 1989). 2ml of the overnight culture was transferred to 1.5 ml eppendorf tubes, centrifuged at maximum speed ≥10,000 × g in a microcentrifuge at 4°C for 2 minutes and the pellet resuspended in 100µl solution 1 (25mM Tris-HCl, pH 8.0, 10mM EDTA and 50mM glucose). Bacteria were lysed by addition of 200µl freshly prepared solution 2 (0.25M NaOH and 1% (w/v) SDS) and mixed by inverting the tube until solution was viscous (about 6 times). 200µl of solution 3 (3M KAc, pH 4.8) was added to the above solution and centrifuged for 5 minutes in a microcentrifuge at 4°C. The supernatant was transferred to a clean tube, mixed with an equal volume of 100% (v/v) isopropanol and centrifuged for 10 minutes at 4°C to pellet the plasmid DNA. The pellet was then washed in 70% (v/v) ethanol, vacuum dried and resuspended in 50µl TE. 2.5 to 5µl of the resuspended DNA was used for restriction endonuclease digestion to ensure the presence of an insert.

4.4.2 Large scale plasmid DNA purification

For large-scale plasmid DNA preparation, 250ml L-Broth supplemented with 50µg/ml ampicillin were inoculated with 50µl glycerol stock and allowed to grow for about 16 hours at 37°C in shaking incubator (250 rpm) (Birnboim and Doly, 1979). Cells were harvested by centrifugation at 6 000rpm in a Beckman JA-10 rotor for 15minutes, and QAIGEN Maxiprep columns were used to isolate and purify plasmid DNA as recommended by the manufacturer (Diagen, Hilden, Germany).

4.5 *IN VITRO* METHYLATION OF DNA

Plasmid DNA constructs were in vitro methylated with the bacterial enzyme M.SssI methylase in the presence of the methyl donor S-adenosylmethionine (SAM) as recommended by the suppliers (New England Biolabs). The substrate specificity of M.SssI closely resembles that of the mammalian enzyme, except that it methylates every CpG dinculeotide in the fragment of interest (Nur *et al.*, 1985). 2U of enzyme were used per μg of DNA in a total volume of 20μl in the presence of NEBuffer 2 (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT) and 160μM SAM. The methylation reaction was allowed to proceed for 2-4 hrs at 37°C. In a parallel control reaction, plasmid DNA was incubated in the absence of M.SssI methylase (mock methylated) but in the same buffer, for comparative studies.

4.5.1 Assessment of the completion of methylation

The extent of methylation was determined by digestion with the methylation sensitive restriction enzymes HpaII (whose recognition sequence is CCGG) and SmaI (which recognises CCCGGG). These enzymes do not cleave their recognition sequences respectively when the internal cytosine is methylated (Sharp *et al.*, 1973). The digests were examined by agarose gel electrophoresis and uncut plasmid DNA was included as a control. The methylated and mock-methylated plasmids were extracted with phenol and ethanol precipitated prior to transfection.

4.5.2 Transient transfections and reporter gene activity

The methylated and mock methylated promoter constructs were transiently transfected into CT-1 fibroblasts using the calcium phosphate precipitation method. The cells were grown to 40% confluency in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum. 20μg reporter DNA and 5μg pCMVβ-galactosidase were cotransfected per 10cm diameter petri dish (MacGregor and Caskey, 1989). Every transfection was performed in triplicate. The medium was changed 16-24 hours after transfection and the cells were harvested 48 hours post transfection. The medium was aspirated; the cells washed with phosphate buffered saline (PBS) and scraped into eppendorf tubes. The cells were harvested by centrifugation at 13000rpm for one minute, resuspended in 150μl of 0.25M Tris-HCl, pH 8.0 and lysed using the

freeze-thaw protocol. Cells in suspension were frozen in liquid nitrogen, immediately transferred to a 37°C water-bath until thawed then quickly vortexed. This freeze-thaw cycle was repeated two more times; the nuclear debris was pelleted by centrifugation and the extracts transferred to a clean tube.

The β-gal activity was determined using the substrate o-nitrophenyl-β-D-galactoside (ONPG) as described by Sambrook *et al.*, (1986). Of the two products that result from the cleavage of ONPG by β-galactosidase, o-nitrophenol formation can be determined spectrophotometrically as its yellow colour can be measured at 420nm.

CAT activity was measured using the method described by Seed & Sheen (1988). The conversion of ^{14}C-chloramphenicol (CAM) to acetylated CAM. 50µl cell extract was incubated at 37°C for between 30 minutes to 2 hours with 2.5µl ^{14}C Chloroamphenicol, 5µl n-butyryl CoA (5mg/ml), and the volume adjusted to 125µl with 0.25M Tris/HCl pH 8.0. The butyrylated chloroamphenicol was extracted with 300µl xylene by vortexing and centrifugation, transferred to a clean tube where it was re-extracted with 100µl 0.25M Tris/HCl, pH 8.0. 200µl of the xylene was added to a scintillation vial and the transfer of the n-butyryl moiety to chloramphenicol by the CAT enzyme, was determined using a Packard Tri-carb 1900CA liquid scintillation counter. The CAT activity was normalised for transfection efficiency using β-galactosidase as an internal standard.

REFERENCES

Adams, R.L., Davis, T., Rinaldi, A. and Eason, R. (1987) CpG deficiency, dinucleotide distributions and nucleosome positioning. *Eur J Biochem*, **165**, 107-15.

Akai, J., Kimura, A. and Hata, R.I. (1999) Transcriptional regulation of the human type I collagen alpha2 (COL1A2) gene by the combination of two dinucleotide repeats. *Gene*, **239**, 65-73.

Alter, D.C. and Subramanian, K.N. (1989) A one step, quick step, mini prep. *Biotechniques*, **7**, 456, 458.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. and Zoghbi, H.Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2 [see comments]. *Nat Genet*, **23**, 185-8.

Antequera, F. and Bird, A. (1993) Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A*, **90**, 11995-9.

Antequera, F. and Bird, A. (1999) CpG islands as genomic footprints of promoters that are associated with replication origins. *Curr Biol*, **9**, R661-7.

Archey, W.B., Sweet, M.P., Alig, G.C. and Arrick, B.A. (1999) Methylation of CpGs as a determinant of transcriptional activation at alternative promoters for transforming growth factor-beta3. *Cancer Res*, **59**, 2292-6.

Bakin, A.V. and Curran, T. (1999) Role of DNA 5-methylcytosine transferase in cell transformation by fos. *Science*, **283**, 387-90.

Barry, C., Faugeron, G. and Rossignol, J.L. (1993) Methylation induced premeiotically in Ascobolus: coextension with DNA repeat lengths and effect on transcript elongation. *Proc Natl Acad Sci U S A*, **90**, 4557-61.

Baylin, S.B. (1992) Abnormal regional hypermethylation in cancer cells. *AIDS Res Hum Retroviruses*, **8**, 811-20.

Baylin, S.B., Herman, J.G., Graff, J.R., Vertino, P.M. and Issa, J.P. (1998) Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res*, **72**, 141-96.

Baylin, S.B. and Herman, J.G. (2000) DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet*, **16**, 168-74.

Bergman, Y. and Mostoslavsky, R. (1998) DNA demethylation: turning genes on. *Biol Chem*, **379**, 401-7.

Bestor, T., Laudano, A., Mattaliano, R. and Ingram, V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol*, **203**, 971-83.

Bestor, T.H. (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *Embo J*, **11**, 2611-7.

Bestor, T.H. (1998) Gene silencing. Methylation meets acetylation [news; comment]. *Nature*, **393**, 311-2.

Bhattacharya, S.K., Ramchandani, S., Cervoni, N. and Szyf, M. (1999) A mammalian protein with specific demethylase activity for mCpG DNA [see comments]. *Nature*, **397**, 579-83.

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

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

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

Bird, A.P. (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res*, **8**, 1499-504.

Bird, A.P. (1984) DNA methylation versus gene expression. *J Embryol Exp Morphol*, **83**, 31-40.

Bird, A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209-13.

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

Bird, A.P. and Wolffe, A.P. (1999) Methylation-induced repression--belts, braces, and chromatin. *Cell*, **99**, 451-4.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*, **7**, 1513-23.

Birnboim, H.C. (1983) A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol*, **100**, 243-55.

Boast, S., Su, M.W., Ramirez, F., Sanchez, M. and Avvedimento, E.V. (1990) Functional analysis of cis-acting DNA sequences controlling transcription of the human type I collagen genes. *J Biol Chem*, **265**, 13351-6.

Bou-Gharios, G., Garrett, L.A., Rossert, J., Niederreither, K., Eberspaecher, H., Smith, C., Black, C. and Crombrugghe, B. (1996) A potent far-upstream enhancer in the mouse pro alpha 2(I) collagen gene regulates expression of reporter genes in transgenic mice. *J Cell Biol*, **134**, 1333-44.

Boyes, J. and Bird, A. (1991) DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*, **64**, 1123-34.

Boyes, J. and Bird, A. (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**, 327-33.

Brandeis, M., Kafri, T., Ariel, M., Chaillet, J.R., McCarrey, J., Razin, A. and Cedar, H. (1993) The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *Embo J*, **12**, 3669-77.

Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*, **56**, 549-61.

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

Burbelo, P.D., Horikoshi, S. and Yamada, Y. (1990) DNA methylation and collagen IV gene expression in F9 teratocarcinoma cells. *J Biol Chem*, **265**, 4839-43.

Burger, C. and Radbruch, A. (1992) Demethylation of the constant region genes of immunoglobulins reflects the differentiation state of the B cell. *Mol Immunol*, **29**, 1105-12.

Burley, S.K. and Roeder, R.G. (1996) Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem*, **65**, 769-99.

Buschhausen, G., Graessmann, M. and Graessmann, A. (1985) Inhibition of herpes simplex thymidine kinase gene expression by DNA methylation is an indirect effect. *Nucleic Acids Res*, **13**, 5503-13.

Buschhausen, G., Wittig, B., Graessmann, M. and Graessmann, A. (1987) Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc Natl Acad Sci U S A*, **84**, 1177-81.

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

Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G. and Baylin, S.B. (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat Genet*, **21**, 103-7.

Campoy, F.J., Meehan, R.R., McKay, S., Nixon, J. and Bird, A. (1995) Binding of histone H1 to DNA is indifferent to methylation at CpG sequences. *J Biol Chem*, **270**, 26473-81.

Cardoso, M.C. and Leonhardt, H. (1999) DNA methyltransferase is actively retained in the cytoplasm during early development. *J Cell Biol*, **147**, 25-32.

Carlson, L.L., Page, A.W. and Bestor, T.H. (1992) Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev*, **6**, 2536-41.

Cedar, H. and Verdine, G.L. (1999) Gene expression. The amazing demethylase [news; comment]. *Nature*, **397**, 568-9.

Cervoni, N., Bhattacharya, S. and Szyf, M. (1999) DNA demethylase is a processive enzyme. *J Biol Chem*, **274**, 8363-6.

Chan, H., Hartung, S. and Breindl, M. (1991) Retrovirus-induced interference with collagen I gene expression in Mov13 fibroblasts is maintained in the absence of DNA methylation. *Mol Cell Biol*, **11**, 47-54.

Chandler, L.A. and Jones, P.A. (1985) Hypomethylation of DNA in the regulation of gene expression. *Dev Biol (N Y*, **5**, 335-49.

Chandler, L.A., DeClerck, Y.A., Bogenmann, E. and Jones, P.A. (1986) Patterns of DNA methylation and gene expression in human tumor cell lines. *Cancer Res*, **46**, 2944-9.

Chi, T. and Carey, M. (1996) Assembly of the isomerized TFIIA--TFIID--TATA ternary complex is necessary and sufficient for gene activation. *Genes Dev*, **10**, 2540-50.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G. and Li, B.F. (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science*, **277**, 1996-2000.

Chung, K.Y., Agarwal, A., Uitto, J. and Mauviel, A. (1996) An AP-1 binding sequence is essential for regulation of the human alpha2(I) collagen (COL1A2) promoter activity by transforming growth factor-beta. *J Biol Chem*, **271**, 3272-8.

Clark, S.J., Harrison, J., Paul, C.L. and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*, **22**, 2990-7.

Clark, S.J., Harrison, J. and Molloy, P.L. (1997) Sp1 binding is inhibited by (m)Cp(m)CpG methylation. *Gene*, **195**, 67-71.

Coleman, R.A., Taggart, A.K., Burma, S., Chicca, J.J., 2nd and Pugh, B.F. (1999) TFIIA regulates TBP and TFIID dimers. *Mol Cell*, **4**, 451-7.

Collins, M., Leaner, V.D., Madikizela, M. and Parker, M.I. (1997) Regulation of the human alpha 2(1) procollagen gene by sequences adjacent to the CCAAT box. *Biochem J*, **322**, 199-206.

Comb, M. and Goodman, H.M. (1990) CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Res*, **18**, 3975-82.

Constancia, M., Pickard, B., Kelsey, G. and Reik, W. (1998) Imprinting mechanisms. *Genome Res*, **8**, 881-900.

Cooper, D.N. and Krawczak, M. (1990) The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions. *Hum Genet*, **85**, 55-74.

Cosma, M.P., Tanaka, T. and Nasmyth, K. (1999) Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell*, **97**, 299-311.

Counts, J.L. and Goodman, J.I. (1994) Hypomethylation of DNA: an epigenetic mechanism involved in tumor promotion. *Mol Carcinog*, **11**, 185-8.

Counts, J.L. and Goodman, J.I. (1995) Alterations in DNA methylation may play a variety of roles in carcinogenesis. *Cell*, **83**, 13-5.

Cross, S.H., Meehan, R.R., Nan, X. and Bird, A. (1997) A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat Genet*, **16**, 256-9.

Cunningham, J.M., Christensen, E.R., Tester, D.J., Kim, C.Y., Roche, P.C., Burgart, L.J. and Thibodeau, S.N. (1998) Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res*, **58**, 3455-60.

D'Souza, R.N., Niederreither, K. and de Crombrugghe, B. (1993) Osteoblast-specific expression of the alpha 2(I) collagen promoter in transgenic mice: correlation with the distribution of TGF-beta 1. *J Bone Miner Res*, **8**, 1127-36.

Davey, C., Pennings, S. and Allan, J. (1997) CpG methylation remodels chromatin structure in vitro. *J Mol Biol*, **267**, 276-88.

de Capoa, A., Febbo, F.R., Giovannelli, F., Niveleau, A., Zardo, G., Marenzi, S. and Caiafa, P. (1999) Reduced levels of poly(ADP-ribosyl)ation result in chromatin compaction and hypermethylation as shown by cell-by-cell computer-assisted quantitative analysis. *Faseb J*, **13**, 89-93.

Delgado, S., Gomez, M., Bird, A. and Antequera, F. (1998) Initiation of DNA replication at CpG islands in mammalian chromosomes. *Embo J*, **17**, 2426-35.

Deng, G., Chen, A., Hong, J., Chae, H.S. and Kim, Y.S. (1999) Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. *Cancer Res*, **59**, 2029-33.

Di Matteo, G., Salerno, M., Guarguaglini, G., Di Fiore, B., Palitti, F. and Lavia, P. (1998) Interactions with Single-stranded and Double-stranded DNA-binding Factors and Alternative Promoter Conformation upon Transcriptional Activation of the Htf9-a/RanBP1 and Htf9-c Genes. *J. Biol. Chem.*, **273**, 495-505.

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

Doerfler, W. (1993) Patterns of de novo DNA methylation and promoter inhibition: studies on the adenovirus and the human genomes. *Exs*, **64**, 262-99.

Dynlacht, B.D., Hoey, T. and Tjian, R. (1991) Isolation of coactivators associated with the ATA-binding protein that mediate transcriptional activation. *Cell*, **66**, 563-76.

Eden, S. and Cedar, H. (1994) Role of DNA methylation in the regulation of transcription. *Curr Opin Genet Dev*, **4**, 255-9.

Eden, S., Hashimshony, T., Keshet, I., Cedar, H. and Thorne, A.W. (1998) DNA methylation models histone acetylation [letter]. *Nature*, **394**, 842.

Ehrlich, M. and Wang, R.Y. (1981) 5-Methylcytosine in eukaryotic DNA. *Science*, **212**, 1350-7.

el-Deiry, W.S., Nelkin, B.D., Celano, P., Yen, R.W., Falco, J.P., Hamilton, S.R. and Baylin, S.B. (1991) High expression of the DNA methyltransferase gene characterizes human neoplastic cells and progression stages of colon cancer. *Proc Natl Acad Sci U S A*, **88**, 3470-4.

Emery, P., Durand, B., Mach, B. and Reith, W. (1996) RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Res*, **24**, 803-7.

Fang, J., Zhu, S., Xiao, S., Shi, Y., Jiang, S., Zhou, X. and Qian, L. (1996) Alterations of level of total genomic DNA methylation and pattern of c-myc, c-Ha-ras oncogene methylation in human gastric carcinogenesis. *Chin Med J (Engl)*, **109**, 787-91.

Fremont, M., Siegmann, M., Gaulis, S., Matthies, R., Hess, D. and Jost, J.P. (1997) Demethylation of DNA by purified chick embryo 5-methylcytosine-DNA glycosylase requires both protein and RNA. *Nucleic Acids Res*, **25**, 2375-80.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul, C.L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A*, **89**, 1827-31.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L. and Kouzarides, T. (2000) DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet*, **24**, 88-91.

Gardiner-Garden, M. and Frommer, M. (1987) CpG islands in vertebrate genomes. *J Mol Biol*, **196**, 261-82.

Gardiner-Garden, M. and Frommer, M. (1994) Transcripts and CpG islands associated with the pro-opiomelanocortin gene and other neurally expressed genes. *J Mol Endocrinol*, **12**, 365-82.

Gaudet, F., Talbot, D., Leonhardt, H. and Jaenisch, R. (1998) A short DNA methyltransferase isoform restores methylation in vivo. *J Biol Chem*, **273**, 32725-9.

Glenn, C.C., Saitoh, S., Jong, M.T., Filbrandt, M.M., Surti, U., Driscoll, D.J. and Nicholls, R.D. (1996) Gene structure, DNA methylation, and imprinted expression of the human SNRPN gene. *Am J Hum Genet*, **58**, 335-46.

Gonzalgo, M.L., Bender, C.M., You, E.H., Glendening, J.M., Flores, J.F., Walker, G.J., Hayward, N.K., Jones, P.A. and Fountain, J.W. (1997) Low frequency of p16/CDKN2A methylation in sporadic melanoma: comparative approaches for methylation analysis of primary tumors. *Cancer Res*, **57**, 5336-47.

Gonzalgo, M.L. and Jones, P.A. (1997) Mutagenic and epigenetic effects of DNA methylation [see comments]. *Mutat Res*, **386**, 107-18.

Gonzalgo, M.L., Hayashida, T., Bender, C.M., Pao, M.M., Tsai, Y.C., Gonzales, F.A., Nguyen, H.D., Nguyen, T.T. and Jones, P.A. (1998) The role of DNA methylation in expression of the p19/p16 locus in human bladder cancer cell lines. *Cancer Res*, **58**, 1245-52.

Goodbourn, S. (1990) Negative regulation of transcriptional initiation in eukaryotes. *Biochim Biophys Acta*, **1032**, 53-77.

Graessmann, A., Sandberg, G., Guhl, E. and Graessmann, M. (1994) Methylation of single sites within the herpes simplex virus tk coding region and the simian virus 40 T-antigen intron causes gene inactivation. *Mol Cell Biol*, **14**, 2004-10.

Graff, J.R., Herman, J.G., Lapidus, R.G., Chopra, H., Xu, R., Jarrard, D.F., Isaacs, W.B., Pitha, P.M., Davidson, N.E. and Baylin, S.B. (1995) E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Res*, **55**, 5195-9.

Graff, J.R., Greenberg, V.E., Herman, J.G., Westra, W.H., Boghaert, E.R., Ain, K.B., Saji, M., Zeiger, M.A., Zimmer, S.G. and Baylin, S.B. (1998) Distinct patterns of E-cadherin CpG island methylation in papillary, follicular, Hurthle's cell, and poorly differentiated human thyroid carcinoma. *Cancer Res*, **58**, 2063-6.

Grafstrom, R.H., Yuan, R. and Hamilton, D.L. (1985) The characteristics of DNA methylation in an in vitro DNA synthesizing system from mouse fibroblasts. *Nucleic Acids Res*, **13**, 2827-42.

Grana, X., Garriga, J. and Mayol, X. (1998) Role of the retinoblastoma protein family, pRB, p107 and p130 in the negative control of cell growth. *Oncogene*, **17**, 3365-83.

Greenwel, P., Inagaki, Y., Hu, W., Walsh, M. and Ramirez, F. (1997) Sp1 is required for the early response of alpha2(I) collagen to transforming growth factor-beta1. *J Biol Chem*, **272**, 19738-45.

Gross-Bellard, M., Oudet, P. and Chambon, P. (1973) Isolation of high-molecular-weight DNA from mammalian cells. *Eur J Biochem*, **36**, 32-8.

Gruenbaum, Y., Naveh-Many, T., Cedar, H. and Razin, A. (1981) Sequence specificity of methylation in higher plant DNA. *Nature*, **292**, 860-2.

Guenette, D.K., Ritzenthaler, J.D., Foley, J., Jackson, J.D. and Smith, B.D. (1992) DNA methylation inhibits transcription of procollagen alpha 2(I) promoters. *Biochem J*, **283**, 699-703.

Hainaut, P., Soussi, T., Shomer, B., Hollstein, M., Greenblatt, M., Hovig, E., Harris, C.C. and Montesano, R. (1997) Database of p53 gene somatic mutations in human tumors and cell lines: updated compilation and future prospects. *Nucleic Acids Res*, **25**, 151-7.

Hainaut, P. and Hollstein, M. (2000) p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res*, **77**, 81-137.

Hanada, M., Delia, D., Aiello, A., Stadtmauer, E. and Reed, J.C. (1993) bcl-2 gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia. *Blood*, **82**, 1820-8.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M. and Gartler, S.M. (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A*, **96**, 14412-7.

Harrison, J., Stirzaker, C. and Clark, S.J. (1998) Cytosines adjacent to methylated CpG sites can be partially resistant to conversion in genomic bisulfite sequencing leading to methylation artifacts. *Anal Biochem*, **264**, 129-32.

Hattman, S., Kenny, C., Berger, L. and Pratt, K. (1978) Comparative study of DNA methylation in three unicellular eucaryotes. *J Bacteriol*, **135**, 1156-7.

Heinzel, T., Lavinsky, R.M., Mullen, T.M., Soderstrom, M., Laherty, C.D., Torchia, J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R., Seto, E., Eisenman, R.N., Rose, D.W., Glass, C.K. and Rosenfeld, M.G. (1997) A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression [see comments]. *Nature*, **387**, 43-8.

Hendrich, B. and Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol*, **18**, 6538-47.

Herman, J.G., Latif, F., Weng, Y., Lerman, M.I., Zbar, B., Liu, S., Samid, D., Duan, D.S., Gnarra, J.R., Linehan, W.M. and et al. (1994) Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A*, **91**, 9700-4.

Herman, J.G., Merlo, A., Mao, L., Lapidus, R.G., Issa, J.P., Davidson, N.E., Sidransky, D. and Baylin, S.B. (1995) Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res*, **55**, 4525-30.

Herman, J.G., Jen, J., Merlo, A. and Baylin, S.B. (1996) Hypermethylation-associated inactivation indicates a tumor suppressor role for p15INK4B. *Cancer Res*, **56**, 722-7.

Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K., Hamilton, S.R., Kinzler, K.W., Kane, M.F., Kolodner, R.D., Vogelstein, B., Kunkel, T.A. and Baylin, S.B. (1998) Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A*, **95**, 6870-5.

Herman, J.G. and Baylin, S.B. (2000) Promoter-region hypermethylation and gene silencing in human cancer. *Curr Top Microbiol Immunol*, **249**, 35-54.

Higashi, K., Kouba, D.J., Song, Y.J., Uitto, J. and Mauviel, A. (1998) A proximal element within the human alpha 2(I) collagen (COL1A2) promoter, distinct from the tumor necrosis factor-alpha response element, mediates transcriptional repression by interferon-gamma. *Matrix Biol*, **16**, 447-56.

Higurashi, M. and Cole, R.D. (1991) The combination of DNA methylation and H1 histone binding inhibits the action of a restriction nuclease on plasmid DNA. *J Biol Chem*, **266**, 8619-25.

Hofer, E. and Darnell, J.E., Jr. (1981) The primary transcription unit of the mouse beta-major globin gene. *Cell*, **23**, 585-93.

Holliday, R. and Pugh, J.E. (1975) DNA modification mechanisms and gene activity during development. *Science*, **187**, 226-32.

Horikoshi, M., Hai, T., Lin, Y.S., Green, M.R. and Roeder, R.G. (1988) Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell*, **54**, 1033-42.

Hsieh, C.L. (1994) Dependence of transcriptional repression on CpG methylation density. *Mol Cell Biol*, **14**, 5487-94.

Hsieh, C.L. (1997) Stability of patch methylation and its impact in regions of transcriptional initiation and elongation. *Mol Cell Biol*, **17**, 5897-904.

Hsieh, C.J., Klump, B., Holzmann, K., Borchard, F., Gregor, M. and Porschen, R. (1998) Hypermethylation of the p16INK4a promoter in colectomy specimens of patients with long-standing and extensive ulcerative colitis. *Cancer Res*, **58**, 3942-5.

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

Hung, M.S., Karthikeyan, N., Huang, B., Koo, H.C., Kiger, J. and Shen, C.J. (1999) Drosophila proteins related to vertebrate DNA (5-cytosine) methyltransferases. *Proc Natl Acad Sci U S A*, **96**, 11940-5.

Iguchi-Ariga, S.M. and Schaffner, W. (1989) CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev*, **3**, 612-9.

Inagaki, Y., Truter, S. and Ramirez, F. (1994) Transforming growth factor-beta stimulates alpha 2(I) collagen gene expression through a cis-acting element that contains an Sp1-binding site. *J Biol Chem*, **269**, 14828-34.

Inagaki, Y., Truter, S., Tanaka, S., Di Liberto, M. and Ramirez, F. (1995) Overlapping pathways mediate the opposing actions of tumor necrosis factor-alpha and transforming growth factor-beta on alpha 2(I) collagen gene transcription. *J Biol Chem*, **270**, 3353-8.

Issa, J.P., Vertino, P.M., Wu, J., Sazawal, S., Celano, P., Nelkin, B.D., Hamilton, S.R. and Baylin, S.B. (1993) Increased cytosine DNA-methyltransferase activity during colon cancer progression. *J Natl Cancer Inst*, **85**, 1235-40.

Issa, J.P., Vertino, P.M., Boehm, C.D., Newsham, I.F. and Baylin, S.B. (1996) Switch from monoallelic to biallelic human IGF2 promoter methylation during aging and carcinogenesis. *Proc Natl Acad Sci U S A*, **93**, 11757-62.

Jaenisch, R., Beard, C., Lee, J., Marahrens, Y. and Panning, B. (1998) Mammalian X chromosome inactivation. *Novartis Found Symp*, **214**, 200-9; discussion 209-13, 228-32.

Jones, P.A. (1996) DNA methylation errors and cancer. *Cancer Res*, **56**, 2463-7.

Jones, P.A. and Gonzalgo, M.L. (1997) Altered DNA methylation and genome instability: a new pathway to cancer? [comment]. *Proc Natl Acad Sci U S A*, **94**, 2103-5.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet*, **19**, 187-91.

Jones, P.A. and Laird, P.W. (1999) Cancer epigenetics comes of age. *Nat Genet*, **21**, 163-7.

Jones, P.A. (1999) The DNA methylation paradox. *Trends Genet*, **15**, 34-7.

Jones, P.L. and Wolffe, A.P. (1999) Relationships between chromatin organization and DNA methylation in determining gene expression. *Semin Cancer Biol*, **9**, 339-47.

Jones, P.A.a.V., P.K. (2000) *DNA methylation and Cancer (Current Topics in Microbiolgoy and Immunology 249)*. Springer-Verlag, Berlin Heidelberg.

Jost, J.P. and Hofsteenge, J. (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 U S A*, **89**, 9499-503.

Jost, J.P., Fremont, M., Siegmann, M. and Hofsteenge, J. (1997) The RNA moiety of chick embryo 5-methylcytosine- DNA glycosylase targets DNA demethylation. *Nucleic Acids Res*, **25**, 4545-50.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. and Razin, A. (1992) Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev*, **6**, 705-14.

Kane, M.F., Loda, M., Gaida, G.M., Lipman, J., Mishra, R., Goldman, H., Jessup, J.M. and Kolodner, R. (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res*, **57**, 808-11.

Karsenty, G. and de Crombrugghe, B. (1990) Two different negative and one positive regulatory factors interact with a short promoter segment of the alpha 1 (I) collagen gene. *J Biol Chem*, **265**, 9934-42.

Karsenty, G. and de Crombrugghe, B. (1991) Conservation of binding sites for regulatory factors in the coordinately expressed alpha 1 (I) and alpha 2 (I) collagen promoters. *Biochem Biophys Res Commun*, **177**, 538-44.

Kass, S.U., Goddard, J.P. and Adams, R.L. (1993) Specific methylation of vector sequences inhibits transcription from the SV40 early promoter. *Biochem Soc Trans*, **21**, 9S.

Kass, S.U., Landsberger, N. and Wolffe, A.P. (1997) DNA methylation directs a time-dependent repression of transcription initiation. *Curr Biol*, **7**, 157-65.

Kass, S.U., Pruss, D. and Wolffe, A.P. (1997) How does DNA methylation repress transcription? *Trends Genet*, **13**, 444-9.

Kautiainen, T.L. and Jones, P.A. (1985) DNA methylation in mammalian nuclei. *Biochemistry*, **24**, 5575-81.

Kautiainen, T.L. and Jones, P.A. (1985) Effects of DNA binding proteins on DNA methylation in vitro. *Biochemistry*, **24**, 1193-6.

Kautiainen, T.L. and Jones, P.A. (1986) DNA methyltransferase levels in tumorigenic and nontumorigenic cells in culture. *J Biol Chem*, **261**, 1594-8.

Kay, P.H., Pereira, E., Marlow, S.A., Turbett, G., Mitchell, C.A., Jacobsen, P.F., Holliday, R. and Papadimitriou, J.M. (1994) Evidence for adenine methylation within the mouse myogenic gene Myo-D1. *Gene*, **151**, 89-95.

Kelley, D.E., Pollok, B.A., Atchison, M.L. and Perry, R.P. (1988) The coupling between enhancer activity and hypomethylation of kappa immunoglobulin genes is developmentally regulated. *Mol Cell Biol*, **8**, 930-7.

Keshet, I., Yisraeli, J. and Cedar, H. (1985) Effect of regional DNA methylation on gene expression. *Proc Natl Acad Sci U S A*, **82**, 2560-4.

Keshet, I., Lieman-Hurwitz, J. and Cedar, H. (1986) DNA methylation affects the formation of active chromatin. *Cell*, **44**, 535-43.

Kitsberg, D., Selig, S., Keshet, I. and Cedar, H. (1993) Replication structure of the human beta-globin gene domain. *Nature*, **366**, 588-90.

Klimasauskas, S., Kumar, S., Roberts, R.J. and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix [see comments]. *Cell*, **76**, 357-69.

Klump, B., Hsieh, C.J., Holzmann, K., Gregor, M. and Porschen, R. (1998) Hypermethylation of the CDKN2/p16 promoter during neoplastic progression in Barrett's esophagus. *Gastroenterology*, **115**, 1381-6.

Kopp, M.U., Winterhalter, K.H. and Trueb, B. (1997) DNA methylation accounts for the inhibition of collagen VI expression in transformed fibroblasts. *Eur J Biochem*, **249**, 489-96.

Kouba, D.J., Chung, K.Y., Nishiyama, T., Vindevoghel, L., Kon, A., Klement, J.F., Uitto, J. and Mauviel, A. (1999) Nuclear factor-kappa B mediates TNF-alpha inhibitory effect on alpha 2(I) collagen (COL1A2) gene transcription in human dermal fibroblasts. *J Immunol*, **162**, 4226-34.

Kudo, S. and Fukuda, M. (1995) Tissue-specific transcriptional regulation of human leukosialin (CD43) gene is achieved by DNA methylation. *J Biol Chem*, **270**, 13298-302.

Kudo, S. (1998) Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated. *Mol Cell Biol*, **18**, 5492-9.

Kuhlman, T.C., Cho, H., Reinberg, D. and Hernandez, N. (1999) The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. *Mol Cell Biol*, **19**, 2130-41.

Kuramasu, A., Saito, H., Suzuki, S., Watanabe, T. and Ohtsu, H. (1998) Mast cell-/basophil-specific transcriptional regulation of human L-histidine decarboxylase gene by CpG methylation in the promoter region. *J Biol Chem*, **273**, 31607-14.

Ladomery, M. (1997) Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. *Bioessays*, **19**, 903-9.

Laird, P.W., Zijderveld, A., Linders, K., Rudnicki, M.A., Jaenisch, R. and Berns, A. (1991) Simplified mammalian DNA isolation procedure. *Nucleic Acids Res*, **19**, 4293.

Laird, P.W. and Jaenisch, R. (1994) DNA methylation and cancer. *Hum Mol Genet*, **3**, 1487-95.

Laird, P.W. and Jaenisch, R. (1996) The role of DNA methylation in cancer genetic and epigenetics. *Annu Rev Genet*, **30**, 441-64.

Lapidus, R.G., Ferguson, A.T., Ottaviano, Y.L., Parl, F.F., Smith, H.S., Weitzman, S.A., Baylin, S.B., Issa, J.P. and Davidson, N.E. (1996) Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone receptor gene expression in breast tumors. *Clin Cancer Res*, **2**, 805-10.

Lapidus, R.G., Nass, S.J., Butash, K.A., Parl, F.F., Weitzman, S.A., Graff, J.G., Herman, J.G. and Davidson, N.E. (1998) Mapping of ER gene CpG island methylation-specific polymerase chain reaction. *Cancer Res*, **58**, 2515-9.

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992) CpG islands as gene markers in the human genome. *Genomics*, **13**, 1095-107.

Lathrop, G.M., Farrall, M., O'Connell, P., Wainwright, B., Leppert, M., Nakamura, Y., Lench, N., Kruyer, H., Dean, M., Park, M. and et al. (1988) Refined linkage map of chromosome 7 in the region of the cystic fibrosis gene. *Am J Hum Genet*, **42**, 38-44.

Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenisch, R. and Li, E. (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*, **122**, 3195-205.

Lengauer, C., Kinzler, K.W. and Vogelstein, B. (1997) DNA methylation and genetic instability in colorectal cancer cells [see comments]. *Proc Natl Acad Sci U S A*, **94**, 2545-50.

Leonhardt, H., Page, A.W., Weier, H.U. and Bestor, T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*, **71**, 865-73.

Leonhardt, H., Rahn, H.P. and Cardoso, M.C. (1999) Functional links between nuclear structure, gene expression, DNA replication, and methylation. *Crit Rev Eukaryot Gene Expr*, **9**, 345-51.

Levine, A., Cantoni, G.L. and Razin, A. (1991) Inhibition of promoter activity by methylation: possible involvement of protein mediators. *Proc Natl Acad Sci U S A*, **88**, 6515-8.

Levine, A., Cantoni, G.L. and Razin, A. (1992) Methylation in the preinitiation domain suppresses gene transcription by an indirect mechanism. *Proc Natl Acad Sci U S A*, **89**, 10119-23.

Levine, A., Yeivin, A., Ben-Asher, E., Aloni, Y. and Razin, A. (1993) Histone H1-mediated inhibition of transcription initiation of methylated templates in vitro. *J Biol Chem*, **268**, 21754-9.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905-14.

Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915-26.

Li, E., Beard, C. and Jaenisch, R. (1993) Role for DNA methylation in genomic imprinting [see comments]. *Nature*, **366**, 362-5.

Lindsay, S. and Bird, A.P. (1987) Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature*, **327**, 336-8.

Liszewski, M.K., Kumar, V. and Atkinson, J.P. (1989) "Midi-prep" isolation of plasmid DNA in less than two hours for sequencing, subcloning and hybridizations. *Biotechniques*, **7**, 1079-81.

Liu, Y., Oakeley, E.J., Sun, L. and Jost, J.P. (1998) Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic Acids Res*, **26**, 1038-45.

Lyko, F., Ramsahoye, B.H., Kashevsky, H., Tudor, M., Mastrangelo, M.A., Orr-Weaver, T.L. and Jaenisch, R. (1999) Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. *Nat Genet*, **23**, 363-6.

MacGregor, G.R. and Caskey, C.T. (1989) Construction of plasmids that express E. coli beta-galactosidase in mammalian cells. *Nucleic Acids Res*, **17**, 2365.

MacLeod, A.R., Rouleau, J. and Szyf, M. (1995) Regulation of DNA methylation by the Ras signaling pathway. *J Biol Chem*, **270**, 11327-37.

Macleod, D., Ali, R.R. and Bird, A. (1998) An alternative promoter in the mouse major histocompatibility complex class II I-Abeta gene: implications for the origin of CpG islands. *Mol Cell Biol*, **18**, 4433-43.

Malone, T., Blumenthal, R.M. and Cheng, X. (1995) Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. *J Mol Biol*, **253**, 618-32.

Mancini, D.N., Rodenhiser, D.I., Ainsworth, P.J., O'Malley, F.P., Singh, S.M., Xing, W. and Archer, T.K. (1998) CpG methylation within the 5' regulatory region of the BRCA1 gene is tumor specific and includes a putative CREB binding site. *Oncogene*, **16**, 1161-9.

Marin, M., Karis, A., Visser, P., Grosveld, F. and Philipsen, S. (1997) Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. *Cell*, **89**, 619-28.

Martienssen, R.A. and Richards, E.J. (1995) DNA methylation in eukaryotes. *Curr Opin Genet Dev*, **5**, 234-42.

Martienssen, R. (1998) Transposons, DNA methylation and gene control [letter; comment]. *Trends Genet*, **14**, 263-4.

McArthur, M. and Thomas, J.O. (1996) A preference of histone H1 for methylated DNA. *Embo J*, **15**, 1705-14.

McCombie, W.R., Heiner, C., Kelley, J.M., Fitzgerald, M.G. and Gocayne, J.D. (1992) Rapid and reliable fluorescent cycle sequencing of double-stranded templates. *DNA Seq*, **2**, 289-96.

McGhee, J.D. and Ginder, G.D. (1979) Specific DNA methylation sites in the vicinity of the chicken beta-globin genes. *Nature*, **280**, 419-20.

McKeon, C., Ohkubo, H., Pastan, I. and de Crombrugghe, B. (1982) Unusual methylation pattern of the alpha 2 (I) collagen gene. *Cell*, **29**, 203-10.

McKeon, C., Pastan, I. and de Crombrugghe, B. (1984) DNase I sensitivity of the alpha 2(I) collagen gene: correlation with its expression but not with its methylation pattern. *Nucleic Acids Res*, **12**, 3491-502.

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

Mertineit, C., Yoder, J.A., Taketo, T., Laird, D.W., Trasler, J.M. and Bestor, T.H. (1998) Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development*, **125**, 889-97.

Meyer, P., Niedenhof, I. and ten Lohuis, M. (1994) Evidence for cytosine methylation of non-symmetrical sequences in transgenic Petunia hybrida. *Embo J*, **13**, 2084-8.

Mittal, V., Ma, B. and Hernandez, N. (1999) SNAP(c): a core promoter factor with a built-in DNA-binding damper that is deactivated by the Oct-1 POU domain. *Genes Dev*, **13**, 1807-21.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y. and Allis, C.D. (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell*, **87**, 1261-70.

Murray, E.J. and Grosveld, F. (1987) Site specific demethylation in the promoter of human gamma-globin gene does not alleviate methylation mediated suppression. *Embo J*, **6**, 2329-35.

Nagatake, M., Osada, H., Kondo, M., Uchida, K., Nishio, M., Shimokata, K. and Takahashi, T. (1996) Aberrant hypermethylation at the bcl-2 locus at 18q21 in human lung cancers. *Cancer Res*, **56**, 1886-91.

Nakayama, M., Wada, M., Harada, T., Nagayama, J., Kusaba, H., Ohshima, K., Kozuru, M., Komatsu, H., Ueda, R. and Kuwano, M. (1998) Hypomethylation status of CpG sites at the promoter region and overexpression of the human MDR1 gene in acute myeloid leukemias. *Blood*, **92**, 4296-307.

Nan, X., Campoy, F.J. and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471-81.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386-9.

Newell-Price, J., Clark, A.J. and King, P. (2000) DNA Methylation and Silencing of Gene Expression. *Trends Endocrinol Metab*, **11**, 142-148.

Ng, H.H. and Bird, A. (1999) DNA methylation and chromatin modification. *Curr Opin Genet Dev*, **9**, 158-63.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. and Bird, A. (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex [see comments]. *Nat Genet*, **23**, 58-61.

Niederreither, K., D'Souza, R.N. and de Crombrugghe, B. (1992) Minimal DNA sequences that control the cell lineage-specific expression of the pro alpha 2(I) collagen promoter in transgenic mice. *J Cell Biol*, **119**, 1361-70.

Nightingale, K. and Wolffe, A.P. (1995) Methylation at CpG sequences does not influence histone H1 binding to a nucleosome including a Xenopus borealis 5 S rRNA gene. *J Biol Chem*, **270**, 4197-200.

Nur, I., Szyf, M., Razin, A., Glaser, G., Rottem, S. and Razin, S. (1985) Procaryotic and eucaryotic traits of DNA methylation in spiroplasmas (mycoplasmas). *J Bacteriol*, **164**, 19-24.

Nyce, J., Liu, L. and Jones, P.A. (1986) Variable effects of DNA-synthesis inhibitors upon DNA methylation in mammalian cells. *Nucleic Acids Res*, **14**, 4353-67.

Ogbourne, S. and Antalis, T.M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J*, **331**, 1-14.

Ohtani-Fujita, N., Fujita, T., Aoike, A., Osifchin, N.E., Robbins, P.D. and Sakai, T. (1993) CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene. *Oncogene*, **8**, 1063-7.

Okano, M., Xie, S. and Li, E. (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases [letter]. *Nat Genet*, **19**, 219-20.

Okano, M., Xie, S. and Li, E. (1998) Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res*, **26**, 2536-40.

Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247-57.

Parker, M.I., Judge, K. and Gevers, W. (1982) Loss of type I procollagen gene expression in SV40-transformed human fibroblasts is accompanied by hypermethylation of these genes. *Nucleic Acids Res*, **10**, 5879-91.

Parker, M.I. and Gevers, W. (1984) Demethylation of the type I procollagen genes in transformed fibroblasts treated with 5-azacytidine. *Biochem Biophys Res Commun*, **124**, 236-43.

Parker, M.I., de Haan, J.B. and Gevers, W. (1986) DNA hypermethylation in sodium butyrate-treated WI-38 fibroblasts. *J Biol Chem*, **261**, 2786-90.

Parker, M.I., Smith, A.A. and Gevers, W. (1989) Absence of alpha 2(1) procollagen synthesis in a clone of SV40-transformed WI-38 human fibroblasts. *J Biol Chem*, **264**, 7147-52.

Parker, M.I., Smith, A.A., Mundell, K., Collins, M., Boast, S. and Ramirez, F. (1992) The abolition of collagen gene expression in SV40-transformed fibroblasts is associated with trans-acting factor switching. *Nucleic Acids Res*, **20**, 5825-30.

Paroush, Z., Keshet, I., Yisraeli, J. and Cedar, H. (1990) Dynamics of demethylation and activation of the alpha-actin gene in myoblasts. *Cell*, **63**, 1229-37.

Patikoglou, G.A., Kim, J.L., Sun, L., Yang, S.H., Kodadek, T. and Burley, S.K. (1999) TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev*, **13**, 3217-30.

Pelissier, T., Thalmeir, S., Kempe, D., Sanger, H.L. and Wassenegger, M. (1999) Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. *Nucleic Acids Res*, **27**, 1625-34.

Pollard, K.J. and Peterson, C.L. (1998) Chromatin remodeling: a marriage between two families? *Bioessays*, **20**, 771-80.

Prendergast, G.C., Lawe, D. and Ziff, E.B. (1991) Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. *Cell*, **65**, 395-407.

Proffitt, J.H., Davie, J.R., Swinton, D. and Hattman, S. (1984) 5-Methylcytosine is not detectable in Saccharomyces cerevisiae DNA. *Mol Cell Biol*, **4**, 985-8.

Rainier, S. and Feinberg, A.P. (1988) Capture and characterization of 5-aza-2'-deoxycytidine-treated C3H/10T1/2 cells prior to transformation. *Proc Natl Acad Sci U S A*, **85**, 6384-8.

Rainier, S. and Feinberg, A.P. (1994) Genomic imprinting, DNA methylation, and cancer [published erratum appears in J Natl Cancer Inst 1994 Jul 6;86(13):1028]. *J Natl Cancer Inst*, **86**, 753-9.

Ramsay, G.M., Enrietto, P.J., Graf, T. and Hayman, M.J. (1982) Recovery of myc-specific sequences by a partially transformation-defective mutant of avian myelocytomatosis virus, MC29, correlates with the restoration of transforming activity. *Proc Natl Acad Sci U S A*, **79**, 6885-9.

Ramirez, J.C., Santaren, J.F. and Almendral, J.M. (1995) Transcriptional inhibition of the parvovirus minute virus of mice by constitutive expression of an antisense RNA targeted against the NS-1 transactivator protein. *Virology*, **206**, 57-68.

Razin, A. and Szyf, M. (1984) DNA methylation patterns. Formation and function. *Biochim Biophys Acta*, **782**, 331-42.

Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. and Cantoni, G.L. (1986) Replacement of 5-methylcytosine by cytosine: a possible mechanism for transient DNA demethylation during differentiation. *Proc Natl Acad Sci U S A*, **83**, 2827-31.

Razin, S.V., Vassetzky, Y.S., Jr., Kvartskhava, A.I., Grinenko, N.F. and Georgiev, G.P. (1991) Transcriptional enhancer in the vicinity of a replication origin within the 5' region of the chicken alpha-globin gene domain. *J Mol Biol*, **217**, 595-8.

Razin, A. and Cedar, H. (1993) DNA methylation and embryogenesis. *Exs*, **64**, 343-57.

Razin, A. and Shemer, R. (1995) DNA methylation in early development. *Hum Mol Genet*, **4**, 1751-5.

Razin, A. (1998) CpG methylation, chromatin structure and gene silencing-a three-way connection. *Embo J*, **17**, 4905-8.

Reith, W., Ucla, C., Barras, E., Gaud, A., Durand, B., Herrero-Sanchez, C., Kobr, M. and Mach, B. (1994) RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. *Mol Cell Biol*, **14**, 1230-44.

Rhodes, K. and Breindl, M. (1992) Developmental changes in the methylation status of regulatory elements in the murine alpha 1(I) collagen gene. *Gene Expr*, **2**, 59-69.

Rhodes, K., Rippe, R.A., Umezawa, A., Nehls, M., Brenner, D.A. and Breindl, M. (1994) DNA methylation represses the murine alpha 1(I) collagen promoter by an indirect mechanism. *Mol Cell Biol*, **14**, 5950-60.

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

Roberts, R.J. (1995) On base flipping. *Cell*, **82**, 9-12.

Robertson, K.D., Hayward, S.D., Ling, P.D., Samid, D. and Ambinder, R.F. (1995) Transcriptional activation of the Epstein-Barr virus latency C promoter after 5-azacytidine treatment: evidence that demethylation at a single CpG site is crucial. *Mol Cell Biol*, **15**, 6150-9.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A. and Jones, P.A. (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res*, **27**, 2291-8.

Robertson, K.D., Keyomarsi, K., Gonzales, F.A., Velicescu, M. and Jones, P.A. (2000) Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. *Nucleic Acids Res*, **28**, 2108-13.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L. and Wolffe, A.P. (2000) DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet*, **25**, 338-42.

Rossert, J.A., Chen, S.S., Eberspaecher, H., Smith, C.N. and de Crombrugghe, B. (1996) Identification of a minimal sequence of the mouse pro-alpha 1(I) collagen promoter that confers high-level osteoblast expression in transgenic mice and that binds a protein selectively present in osteoblasts. *Proc Natl Acad Sci U S A*, **93**, 1027-31.

Rother, K.I., Silke, J., Georgiev, O., Schaffner, W. and Matsuo, K. (1995) Influence of DNA sequence and methylation status on bisulfite conversion of cytosine residues. *Anal Biochem*, **231**, 263-5.

Rouleau, J., MacLeod, A.R. and Szyf, M. (1995) Regulation of the DNA methyltransferase by the Ras-AP-1 signaling pathway. *J Biol Chem*, **270**, 1595-601.

Rountree, M.R., Bachman, K.E. and Baylin, S.B. (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet*, **25**, 269-77.

Roy, A.L., Malik, S., Meisterernst, M. and Roeder, R.G. (1993) An alternative pathway for transcription initiation involving TFII-I. *Nature*, **365**, 355-9.

Sage, E.H. and Bornstein, P. (1991) Extracellular proteins that modulate cell-matrix interactions. SPARC, tenascin, and thrombospondin. *J Biol Chem*, **266**, 14831-4.

Saluz, H.P., Jiricny, J. and Jost, J.P. (1986) Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucocorticoid-receptor binding sites and the rate of avian vitellogenin mRNA synthesis. *Proc Natl Acad Sci U S A*, **83**, 7167-71.

Sambrook, J., Hanahan, D., Rodgers, L. and Gething, M.J. (1986) Expression of human tissue-type plasminogen activator from lytic viral vectors and in established cell lines. *Mol Biol Med*, **3**, 459-81.

Sandberg, G. and Schalling, M. (1997) Effect of in vitro promoter methylation and CGG repeat expansion on FMR-1 expression. *Nucleic Acids Res*, **25**, 2883-7.

Sanford, J.P., Clark, H.J., Chapman, V.M. and Rossant, J. (1987) Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. *Genes Dev*, **1**, 1039-46.

Schmidt, A., Rossi, P. and de Crombrugghe, B. (1986) Transcriptional control of the mouse alpha 2(I) collagen gene: functional deletion analysis of the promoter and evidence for cell-specific expression. *Mol Cell Biol*, **6**, 347-54.

Schmutte, C., Yang, A.S., Nguyen, T.T., Beart, R.W. and Jones, P.A. (1996) Mechanisms for the involvement of DNA methylation in colon carcinogenesis. *Cancer Res*, **56**, 2375-81.

Schmutte, C. and Jones, P.A. (1998) Involvement of DNA methylation in human carcinogenesis. *Biol Chem*, **379**, 377-88.

Schwarz, S., Hess, D. and Jost, J.P. (1997) The methylated DNA binding protein-2-H1 (MDBP-2-H1) consists of histone H1 subtypes which are truncated at the C-terminus. *Nucleic Acids Res*, **25**, 5052-6.

Seed, B. and Sheen, J.Y. (1988) A simple phase-extraction assay for chloramphenicol acyltransferase activity. *Gene*, **67**, 271-7.

Sengupta, P.K. and Smith, B.D. (1998) Methylation in the initiation region of the first exon suppresses collagen pro-alpha2(I) gene transcription. *Biochim Biophys Acta*, **1443**, 75-89.

Sengupta, P.K., Ehrlich, M. and Smith, B.D. (1999) A methylation-responsive MDBP/RFX site is in the first exon of the collagen alpha2(I) promoter. *J Biol Chem*, **274**, 36649-55.

Sharp, P.A., Sugden, B. and Sambrook, J. (1973) Detection of two restriction endonuclease activities in Haemophilus parainfluenzae using analytical agarose--ethidium bromide electrophoresis. *Biochemistry*, **12**, 3055-63.

Shen, C.K. and Maniatis, T. (1980) Tissue-specific DNA methylation in a cluster of rabbit beta-like globin genes. *Proc Natl Acad Sci U S A*, **77**, 6634-8.

Shen, J.C., Rideout, W.M.d. and Jones, P.A. (1992) High frequency mutagenesis by a DNA methyltransferase. *Cell*, **71**, 1073-80.

Shen, J.C., Zingg, J.M., Yang, A.S., Schmutte, C. and Jones, P.A. (1995) A mutant HpaII methyltransferase functions as a mutator enzyme. *Nucleic Acids Res*, **23**, 4275-82.

Shuman, S. (1994) Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *Journal of Biological Chemistry*, **269**, 32678-32684.

Siegfried, Z. and Cedar, H. (1997) DNA methylation: a molecular lock. *Curr Biol*, **7**, R305-7.

Silva, A.J. and White, R. (1988) Inheritance of allelic blueprints for methylation patterns. *Cell*, **54**, 145-52.

Simkevich, C.P., Thompson, J.P., Poppleton, H. and Raghow, R. (1992) The transcriptional tissue specificity of the human pro alpha 1 (I) collagen gene is determined by a negative cis-regulatory element in the promoter. *Biochem J*, **286**, 179-85.

Singal, R., Ferris, R., Little, J.A., Wang, S.Z. and Ginder, G.D. (1997) Methylation of the minimal promoter of an embryonic globin gene silences transcription in primary erythroid cells. *Proc Natl Acad Sci U S A*, **94**, 13724-9.

Slack, J.L., Liska, D.J. and Bornstein, P. (1993) Regulation of expression of the type I collagen genes. *Am J Med Genet*, **45**, 140-51.

Slack, A., Cervoni, N., Pinard, M. and Szyf, M. (1999) DNA methyltransferase is a downstream effector of cellular transformation triggered by simian virus 40 large T antigen. *J Biol Chem*, **274**, 10105-12.

Smale, S.T., Schmidt, M.C., Berk, A.J. and Baltimore, D. (1990) Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc Natl Acad Sci U S A*, **87**, 4509-13.

Smith, B.D. and Marsilio, E. (1988) Methylation of the alpha 2(I) collagen gene in chemically transformed rat liver epithelial cells. *Biochem J*, **253**, 269-73.

Smith, L.T., Sakai, L.Y., Burgeson, R.E. and Holbrook, K.A. (1988) Ontogeny of structural components at the dermal-epidermal junction in human embryonic and fetal skin: the appearance of anchoring fibrils and type VII collagen. *J Invest Dermatol*, **90**, 480-5.

Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. and Cedar, H. (1982) Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proc Natl Acad Sci U S A*, **79**, 61-5.

Stein, R., Sciaky-Gallili, N., Razin, A. and Cedar, H. (1983) Pattern of methylation of two genes coding for housekeeping functions. *Proc Natl Acad Sci U S A*, **80**, 2422-6.

Stirzaker, C., Millar, D.S., Paul, C.L., Warnecke, P.M., Harrison, J., Vincent, P.C., Frommer, M. and Clark, S.J. (1997) Extensive DNA methylation spanning the Rb promoter in retinoblastoma tumors. *Cancer Res*, **57**, 2229-37.

Stoger, R., Kajimura, T.M., Brown, W.T. and Laird, C.D. (1997) Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. *Hum Mol Genet*, **6**, 1791-801.

Swisher, J.F., Rand, E., Cedar, H. and Marie Pyle, A. (1998) Analysis of putative RNase sensitivity and protease insensitivity of demethylation activity in extracts from rat myoblasts. *Nucleic Acids Res*, **26**, 5573-80.

Szyf, M., Kaplan, F., Mann, V., Giloh, H., Kedar, E. and Razin, A. (1985) Cell cycle-dependent regulation of eukaryotic DNA methylase level. *J Biol Chem*, **260**, 8653-6.

Szyf, M., Bozovic, V. and Tanigawa, G. (1991) Growth regulation of mouse DNA methyltransferase gene expression. *J Biol Chem*, **266**, 10027-30.

Szyf, M. (2001) Towards a pharmacology of DNA methylation. *Trends Pharmacol Sci*, **22**, 350-4.

Tamaki, T., Ohnishi, K., Hartl, C., LeRoy, E.C. and Trojanowska, M. (1995) Characterization of a GC-rich region containing Sp1 binding site(s) as a constitutive responsive element of the alpha 2(I) collagen gene in human fibroblasts. *J Biol Chem*, **270**, 4299-304.

Tanaka, K., Appella, E. and Jay, G. (1983) Developmental activation of the H-2K gene is correlated with an increase in DNA methylation. *Cell*, **35**, 457-65.

Tasheva, E.S. and Roufa, D.J. (1994) Densely methylated DNA islands in mammalian chromosomal replication origins [published erratum appears in Mol Cell Biol 1995 Dec;15(12):7161]. *Mol Cell Biol*, **14**, 5636-44.

Tate, P.H. and Bird, A.P. (1993) Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr Opin Genet Dev*, **3**, 226-31.

Tate, P., Skarnes, W. and Bird, A. (1996) The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nat Genet*, **12**, 205-8.

Tatematsu, K.I., Yamazaki, T. and Ishikawa, F. (2000) MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. *Genes Cells*, **5**, 677-688.

Theiss, G., Schleicher, R., Schimpff-Weiland, G. and Follmann, H. (1987) DNA methylation in wheat. Purification and properties of DNA methyltransferase. *Eur J Biochem*, **167**, 89-96.

Thompson, J.P., Simkevich, C.P., Holness, M.A., Kang, A.H. and Raghow, R. (1991) In vitro methylation of the promoter and enhancer of Pro alpha 1(I) collagen gene leads to its transcriptional inactivation. *J Biol Chem*, **266**, 2549-56.

Tommasi, S. and Pfeifer, G.P. (1995) In vivo structure of the human cdc2 promoter: release of a p130-E2F-4 complex from sequences immediately upstream of the transcription initiation site coincides with induction of cdc2 expression. *Mol Cell Biol*, **15**, 6901-13.

Toth, M., Muller, U. and Doerfler, W. (1990) Establishment of de novo DNA methylation patterns. Transcription factor binding and deoxycytidine methylation at CpG and non-CpG sequences in an integrated adenovirus promoter. *J Mol Biol*, **214**, 673-83.

Turker, M.S. and Bestor, T.H. (1997) Formation of methylation patterns in the mammalian genome [see comments]. *Mutat Res*, **386**, 119-30.

Turner, D.L., Duncan, R. and Lee, P.W. (1992) Site-directed mutagenesis of the C-terminal portion of reovirus protein sigma 1: evidence for a conformation-dependent receptor binding domain. *Virology*, **186**, 219-27.

Udvardy, A. (1999) Dividing the empire: boundary chromatin elements delimit the territory of enhancers. *Embo J*, **18**, 1-8.

Ullmann, A., Perrin, D., Jacob, F. and Monod, J. (1965) [Identification, by in vitro complementation and purification, of a peptide fraction of Escherichia coli beta-galactosidase]. *J Mol Biol*, **12**, 918-23.

Umezawa, A., Yamamoto, H., Rhodes, K., Klemsz, M.J., Maki, R.A. and Oshima, R.G. (1997) Methylation of an ETS site in the intron enhancer of the keratin 18 gene participates in tissue-specific repression. *Mol Cell Biol*, **17**, 4885-94.

Ura, K., Kurumizaka, H., Dimitrov, S., Almouzni, G. and Wolffe, A.P. (1997) Histone acetylation: influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression. *Embo J*, **16**, 2096-107.

Urieli-Shoval, S., Gruenbaum, Y., Sedat, J. and Razin, A. (1982) The absence of detectable methylated bases in Drosophila melanogaster DNA. *FEBS Lett*, **146**, 148-52.

Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D. and Aloni, Y. (1992) Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell*, **69**, 871-81.

Van Dyke, M.W., Sawadogo, M. and Roeder, R.G. (1989) Stability of transcription complexes on class II genes. *Mol Cell Biol*, **9**, 342-4.

Van Etten, J.L., Schuster, A.M., Girton, L., Burbank, D.E., Swinton, D. and Hattman, S. (1985) DNA methylation of viruses infecting a eukaryotic Chlorella-like green alga. *Nucleic Acids Res*, **13**, 3471-8.

Voet, D.V.a.J. (1990) *Biochemistry*. John Wiley & Sons, Philadelphia.

Voss, A., Pfaller, M.A., Hollis, R.J., Melchers, W.J. and Meis, J.F. (1998) Evaluation of the discriminatory power of pulsed-field gel electrophoresis and PCR fingerprinting for epidemiologic typing of Candida species. *Clin Microbiol Infect*, **4**, 82-87.

Vuorio, E. and de Crombrugghe, B. (1990) The family of collagen genes. *Annu Rev Biochem*, **59**, 837-72.

Vuust, J., Sobel, M.E. and Martin, G.R. (1985) Regulation of type I collagen synthesis. Total pro alpha 1(I) and pro alpha 2(I) mRNAs are maintained in a 2:1 ratio under varying rates of collagen synthesis. *Eur J Biochem*, **151**, 449-53.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F. and Wolffe, A.P. (1999) Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation [see comments]. *Nat Genet*, **23**, 62-6.

Walsh, C.P. and Bestor, T.H. (1999) Cytosine methylation and mammalian development. *Genes Dev*, **13**, 26-34.

Waltz, S.E., Trivedi, A.A. and Leffak, M. (1996) DNA replication initiates non-randomly at multiple sites near the c-myc gene in HeLa cells. *Nucleic Acids Res*, **24**, 1887-94.

Wan, M., Lee, S.S., Zhang, X., Houwink-Manville, I., Song, H.R., Amir, R.E., Budden, S., Naidu, S., Pereira, J.L., Lo, I.F., Zoghbi, H.Y., Schanen, N.C. and Francke, U. (1999) Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots. *Am J Hum Genet*, **65**, 1520-9.

Warnecke, P.M., Stirzaker, C., Melki, J.R., Millar, D.S., Paul, C.L. and Clark, S.J. (1997) Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res*, **25**, 4422-6.

Waye, M.M., Robinson, R., Orfanides, A.G. and Aubin, J.E. (1989) Loss of alpha I type I collagen gene expression in rat clonal bone cell lines is accompanied by DNA methylation. *Biochem Biophys Res Commun*, **162**, 1446-52.

Weber, J.A., Taxman, D.J., Lu, Q. and Gilmour, D.S. (1997) Molecular architecture of the hsp70 promoter after deletion of the TATA box or the upstream regulation region. *Mol Cell Biol*, **17**, 3799-808.

Weiss, A., Keshet, I., Razin, A. and Cedar, H. (1996) DNA demethylation in vitro: involvement of RNA [published erratum appears in Cell 1998 Nov 13;95(4):following 573]. *Cell*, **86**, 709-18.

Weitzel, J.M., Buhrmester, H. and Stratling, W.H. (1997) Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. *Mol Cell Biol*, **17**, 5656-66.

Wigler, M., Levy, D. and Perucho, M. (1981) The somatic replication of DNA methylation. *Cell*, **24**, 33-40.

Wilkinson, C.R., Bartlett, R., Nurse, P. and Bird, A.P. (1995) The fission yeast gene pmt1+ encodes a DNA methyltransferase homologue. *Nucleic Acids Res*, **23**, 203-10.

Wolffe, A.P. (1997) Transcriptional control. Sinful repression [news; comment]. *Nature*, **387**, 16-7.

Wolffe, A.P. (1998) Packaging principle: how DNA methylation and histone acetylation control the transcriptional activity of chromatin. *J Exp Zool*, **282**, 239-44.

Woodcock, D.M., Crowther, P.J. and Diver, W.P. (1987) The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide. *Biochem Biophys Res Commun*, **145**, 888-94.

Woodcock, D.M., Crowther, P.J., Jefferson, S. and Diver, W.P. (1988) Methylation at dinucleotides other than CpG: implications for human maintenance methylation. *Gene*, **74**, 151-2.

Woodcock, D.M., Lawler, C.B., Linsenmeyer, M.E., Doherty, J.P. and Warren, W.D. (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J Biol Chem*, **272**, 7810-6.

Woodcock, D.M., Linsenmeyer, M.E., Doherty, J.P. and Warren, W.D. (1999) DNA methylation in the promoter region of the p16 (CDKN2/MTS-1/INK4A) gene in human breast tumours. *Br J Cancer*, **79**, 251-6.

Workman, J.L. and Kingston, R.E. (1992) Nucleosome core displacement in vitro via a metastable transcription factor-nucleosome complex. *Science*, **258**, 1780-4.

Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F. and Barlow, D.P. (1997) Imprinted expression of the Igf2r gene depends on an intronic CpG island [see comments]. *Nature*, **389**, 745-9.

Wyatt, G.R. (1950) Occurence of 5-methylcytosine in nucleic acids. *Nature*, **166**, 237-238.

Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W.W., Okumura, K. and Li, E. (1999) Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene*, **236**, 87-95.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J. and Viegas-Pequignot, E. (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, **402**, 187-91.

Yang, A.S., Shen, J.C., Zingg, J.M., Mi, S. and Jones, P.A. (1995) HhaI and HpaII DNA methyltransferases bind DNA mismatches, methylate uracil and block DNA repair. *Nucleic Acids Res*, **23**, 1380-7.

Yebra, M.J. and Bhagwat, A.S. (1995) A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine. *Biochemistry*, **34**, 14752-7.

Yoder, J.A., Soman, N.S., Verdine, G.L. and Bestor, T.H. (1997) DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J Mol Biol*, **270**, 385-95.

Yoder, J.A. and Bestor, T.H. (1998) A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. *Hum Mol Genet*, **7**, 279-84.

Yoshida, M., Horinouchi, S. and Beppu, T. (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays*, **17**, 423-30.

Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T. and Hirohashi, S. (1995) Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc Natl Acad Sci U S A*, **92**, 7416-9.

Zardo, G., D'Erme, M., Reale, A., Strom, R., Perilli, M. and Caiafa, P. (1997) Does poly(ADP-ribosyl)ation regulate the DNA methylation pattern? *Biochemistry*, **36**, 7937-43.

Zardo, G., Marenzi, S., Perilli, M. and Caiafa, P. (1999) Inhibition of poly(ADP-ribosyl)ation introduces an anomalous methylation pattern in transfected foreign DNA. *Faseb J*, **13**, 1518-22.

Zengerling, S., Tsui, L.C., Grzeschik, K.H., Olek, K., Riordan, J.R. and Buchwald, M. (1987) Mapping of DNA markers linked to the cystic fibrosis locus on the long arm of chromosome 7 [published erratum appears in Am J Hum Genet 1987 Aug;41(2):330]. *Am J Hum Genet*, **40**, 228-36.

Zhang, X.Y., Supakar, P.C., Khan, R., Ehrlich, K.C. and Ehrlich, M. (1989) Related sites in human and herpesvirus DNA recognized by methylated DNA-binding protein from human placenta. *Nucleic Acids Res*, **17**, 1459-74.

Zhang, X.Y., Asiedu, C.K., Supakar, P.C., Khan, R., Ehrlich, K.C. and Ehrlich, M. (1990) Binding sites in mammalian genes and viral gene regulatory regions recognized by methylated DNA-binding protein. *Nucleic Acids Res*, **18**, 6253-60.

Zhang, X.Y., Jabrane-Ferrat, N., Asiedu, C.K., Samac, S., Peterlin, B.M. and Ehrlich, M. (1993) The major histocompatibility complex class II promoter-binding protein RFX (NF-X) is a methylated DNA-binding protein. *Mol Cell Biol*, **13**, 6810-8.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A. and Reinberg, D. (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev*, **13**, 1924-35.

Zhao, Y., Tsutsumi, R., Yamaki, M., Nagatsuka, Y., Ejiri, S. and Tsutsumi, K. (1994) Initiation zone of DNA replication at the aldolase B locus encompasses transcription promoter region. *Nucleic Acids Res*, **22**, 5385-90.