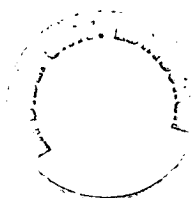


# Transcriptional repression by methyl-CpG binding proteins

Huck Hui Ng

Thesis presented for the Degree of Doctor of Philosophy  
at the University of Edinburgh

1999



## **Declaration**

I declare that this thesis was composed by myself and that the research presented is my own. Due acknowledgement is made within the text for the assistance of others.



## **Dedication**

To Wong Woon Yeng for her patience, understanding, encouragement and support.

To my family for their encouragement and support.

## **Acknowledgements**

I am most grateful to Adrian for replying my first E-mail so fast, for taking me as a PhD student, for giving me the freedom and support to pursue my many interests, and for making my PhD as wonderful experience in Science. I would like to thank the Darwin Trust of Edinburgh for sponsoring my PhD. Last but not least, I would like to thank Aileen, Joan, Jillian and other members of the laboratory for help and assistance. I would grateful to my collaborators, Xinsheng Nan, Colin Johnson, Peter Jeppesen, and Yi Zhang for their contribution to the project. I would also like to thank Bernie Ramsahoye for reading a draft of this thesis.

## Table of Contents

Title	i
Declaration	ii
Dedication	iii
Acknowledgements	iv
Table of Contents	v
Abbreviations	x
Abstract	xii

## Chapter 1 Introduction

1.1	Covalent modification of DNA	1
1.2	DNA methylation in the prokaryotes	1
1.3	DNA methylation in the eukaryotes	3
1.3.1	Inheritance of DNA methylation	3
1.3.2	Distribution of cytosine methylation in eukaryotes	6
1.4	Functions of DNA methylation in vertebrates	7
1.4.1	Profile of genomic methylation during mouse development	7
1.4.2	Mammalian DNA cytosine methyltransferases	8
1.4.3	Role of DNA methylation in development	10
1.5	Dynamics of DNA methylation	10
1.6	5-meC is a mutagenic hotspot	14
1.7	DNA methylation and gene regulation	16
1.7.1	Evidence that DNA methylation is associated with transcriptional repression	16
1.7.2	DNA methylation: a cause or consequence of gene silencing	18
1.8	DNA methylation and chromatin modification	18
1.8.1	Methylated and non-methylated DNA adopt fundamentally different chromatin structures	18
1.9	How does DNA methylation repress transcription?	19
1.10	Post-translational modification of histones: lysine acetylation	24
1.10.1	The mSin3 and NuRD histone deacetylase complexes	26

## Chapter 2      Materials and Methods

2.1	General buffers	30
2.2	Growth and maintenance of <i>Escherichia coli</i> cells	30
2.3	General methods	30
2.3.1	Agarose gel electrophoresis	30
2.3.2	Recovery of DNA fragment	30
2.3.3	Purification of DNA fragment	31
2.3.4	Restriction endonuclease digestion, dephosphorylation, and ligation	31
2.4	Plasmid miniprep	31
2.5	Large scale plasmid preparation	32
2.6	DNA sequencing	33
2.7	Radioactive labeling of DNA	33
2.8	PCR	34
2.9	Preparation of electro-competent cells	34
2.10	Transformation	34
2.11	Cloning	34
2.12	Protein quantitation	35
2.13	SDS polyacrylamide gel electrophoresis	35
2.14	Western blotting	36
2.15	Expression and affinity purification of recombinant proteins	36
2.16	Antibody production	38
2.17	Affinity purification of antibodies	38
2.18	Maintenance of cell cultures	38
2.19	Making frozen stocks of cultures	38
2.20	Thawing of frozen stock	39
2.21	Transfection of mammalian cells	39
2.22	Cell staining	39
2.23	Reporter assays	40
2.24	Isolation of nuclei from cells	40
2.25	Preparation of nuclear extracts	40
2.26	Bandshift	41
2.27	Immunoprecipitation	41
2.28	Coupled <i>in vitro</i> transcription and translation in TNT reticulocyte lysates	42
2.29	GST pulldown assays	42
2.30	<i>In vitro</i> transcription assay	43

## **Chapter 3      Mechanism of transcriptional repression by methyl-CpG binding protein 2, MeCP2**

3.1	Introduction	44
3.2	Results and discussion	44
3.2.1	MeCP2 can binds to mSin3A and HDACs: mapping of co-repressor interacting region	44
3.2.2	MeCP2 is associated with mSin3A and histone deacetylases	45
3.2.3	MeCP2 can bind to <i>in vitro</i> synthesised mSin3A but not HDACs	48
3.2.4	Mapping of MeCP2 interacting regions within mSin3A	48
3.2.5	Ternary complex between MeCP2/ mSin3A/ HDAC2 cannot be assembled <i>in vitro</i>	51
3.2.6	Repression by GAL4-MeCP2 TRD is sensitive to histone deacetylase inhibitor, TSA	51
3.2.7	HeLa cells is deficient in MeCP2, but can repress methylated reporter construct in a deacetylation dependent manner	54
3.3	Conclusion	57

## **Chapter 4      MBD2 is a component of the transcriptional repressor MeCP1**

4.1	Introduction	61
4.2	Results and discussion	62
4.2.1	Characterisation of antibodies against MBD2	62
4.2.1.1	S752 antibody	62
4.2.1.2	S923 and R593 antibodies	68
4.2.2	MBD2 is a component of MeCP1	71
4.2.3	F9 embryonic carcinoma cells are deficient in MBD2	75
4.2.4	MBD2 is associated with histone deacetylases	77
4.2.5	Interaction between MBD2 and other components of HDAC1/2 complex	82
4.2.6	MBD2 exists in a high molecular weight complex	82
4.2.7	MBD2 can repress transcription <i>in vitro</i>	82
4.2.8	MBD2 can repress transcription <i>in vivo</i>	85
4.2.9	Trichostatin A (TSA) can relieve repression on the human DNA polymerase $\beta$ promoter	88
4.2.10	Mapping of a repression domain?	88

4.2.11	Is MBD3 a methyl-CpG binding protein?	92
4.3	Conclusion	95

## **Chapter 5      MBD1 is a transcriptional repressor**

5.1	Introduction	102
5.2	Results and discussion	102
5.2.1	Full-length MBD1 binds selectively to methylated DNA <i>in vitro</i>	102
5.2.2	The N-terminal of MBD1 is sufficient for binding to methylated DNA	104
5.2.3	Characterisation of an antibody raised against MBD1	106
5.2.4	MBD1 is not a component of MeCP1	111
5.2.5	Biochemical characterisation of MBD1	115
5.2.6	MBD1 can repress transcription in a DNA methylation dependent manner <i>in vivo</i>	118
5.2.7	MBD1 contains a potent transcriptional repression domain	122
5.2.8	Characterisation of the C-terminal TRD (mutagenesis, long distance repression, TSA sensitivity)	124
5.2.9	MBD1 is a chromosomal protein	135
5.2.10	MBD1 has an affinity for core histones	139
5.3	Conclusion	147

## **Chapter 6      General discussion**

6.1	Outstanding questions	152
6.2	MeCPs: proteins with no footprint	155
6.3	The mojo of MBD2	155
6.4	Is MBD3 a methyl-CpG binding protein?	157
6.5	Repression by DNA methylation: more to chromatin than meet the eye	157
6.6	Functions of DNA methylation?	159

<b>References</b>	161
-------------------	-----

<b>Publication</b>	185
--------------------	-----

## Abbreviations:

5-meC	5-methylcytosine
5-aza-dC	5-aza-2'-deoxycytidine
5-aza-C	5-azacytidine
°C	degree centigrade
aa	amino-acids
Ab	antibody
ATP	adenosine triphosphate
bp	basepair
BSA	bovine serum albumin
Ci	Curie
CIR	co-repressor interacting region
C-terminal	carboxy-terminal
CpG	cytosine-guanine dinucleotide
cpm	counts per minute
DAPI	4,6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle medium
DMSO	dimethylsulphoxide
DNA	deoxyribose nucleic acid
DNMT	DNA cytosine methyltransferase
dNTPs	2'deoynucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetic acids, disodium salt
EC	embryonic carcinoma
<i>E.coli.</i>	<i>Escherichia coli</i>
ER	estrogen receptor
ES	embryonic stem
EST	expressed sequence tagged
GSH	glutathione
GST	glutathione S-transferase
HAT	histone acetyltransferase
HDAC	histone deacetylase
HEPES	2-[4(2-hydroxyethyl)-1-piperaziny]-ethanesulphonic acid
hr	hour
IPTG	isopropylthio-β-D-galactose
kb	kilobase
kDa	kiloDalton



KCl	potassium chloride
M	molar
M-	non-methylated
M+	methylated
MBD	methyl-CpG binding domain
MBD1	methyl-CpG binding domain protein 1
MBD2	methyl-CpG binding domain protein 2
MBD3	methyl-CpG binding domain protein 3
MBD4	methyl-CpG binding domain protein 4
MeCP	methyl-CpG binding protein
min	minute
NaCl	sodium chloride
NMR	Nuclear Magnetic Resonance
N-terminal	amino-terminal
nt	nucleotides
O.D.	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
rGSH	reduced glutathione
RNA	ribonucleic acid
RNase	ribonuclease
RNasin	RNase inhibitor
rpm	revolution per minute
RT-PCT	reverse transcription - polymerase chain reaction
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
TAE	Tris-acetate EDTA
TBE	Tris-borate EDTA
TE	Tris-EDTA
TEMED	N,N,N',N'-tetra-methylethylenediamine
TRD	transcriptional repression domain
TSA	Trichostatin A
tk	thymidine kinase
Tris	(hydroxymethyl) aminomethane
tRNA	transfer RNA
uv	ultraviolet
WT	wild-type

## Abstract

One approach to understand the underlying mechanisms for transcriptional repression by DNA methylation is through the study of mammalian methyl-CpG binding proteins. MeCP1 and MeCP2 are two methyl-CpG binding activities previously characterised in this laboratory. The molecular nature of MeCP1 is unknown, but the protein responsible for MeCP2 has been identified. Both *in vitro* and *in vivo* data strongly suggest that MeCP1 is a methyl-CpG dependent transcriptional repressor. MeCP2 is also a transcriptional repressor and contains a transcriptional repression domain. Recent efforts to identify other novel methyl-CpG binding proteins by searching the EST databases for proteins with methyl-CpG DNA binding domain of MeCP2 were successful. MBD1, MBD2 and MBD4 have been shown to bind specifically to methylated DNA *in vitro*. MBD4 is a DNA repair protein and therefore unlikely to be involved in transcriptional repression.

The thesis describes collective studies of these methyl-CpG binding proteins (MeCP2, MBD1, and MBD2). MeCP2 is associated with the mSin3 corepressor complex which contains histone deacetylase subunits. Repression by the transcriptional repression domain of MeCP2 is sensitive to the histone deacetylase inhibitor, Trichostatin A (TSA), indicating that deacetylation is a critical component of the repression mechanism of MeCP2. Interestingly, MBD2 was found to be the DNA binding component of the long sought MeCP1 complex by several experimental criteria. MBD2 is also associated with histone deacetylases, and can repress transcription when tethered to the DNA. Repression of certain promoters by MBD2 can be relieved by treatment with TSA. MBD1 can also repress transcription and contains a potent repression domain. The repression by this novel domain was found to be sensitive to TSA, suggesting that deacetylation may again be involved. Altogether, these studies provide a molecular link to account for the long known relationship between DNA methylation, transcriptional repression and chromatin modification.

# **Chapter 1: Introduction**

## **1.1 Covalent modification of DNA**

5-methylcytosine (5-meC) is a modified base with a methyl group attached to the carbon 5 position of cytosine and can be found in both plants and animals (Hotchkiss, 1948; Wyatt, 1951). There are however, several eukaryotes such as the yeast *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans* with no detectable of 5-meC in the genome (Proffitt et al, 1984; Urieli-Shoval et al, 1982; Simpson et al, 1986). Other minor bases such as N6-methyladenine or N4-methylcytosine are also found in the DNA of some prokaryotes (Dunn and Smith, 1955). The identification of a DNA methyltransferase in *Escherichia coli* and the detection of DNA methyltransferase activity in mammalian extracts established that this covalent modification is mediated by enzymes, termed DNA methyltransferases (Figure 1.1, Gold and Hurwitz, 1964; Sheid, 1968). DNA methylation at cytosine and adenine does not affect the base pairing of cytosine to guanine or adenine to thymine, therefore the genetic coding system based on the four major bases (guanine, adenine, thymine and cytosine) is not altered. However, methyl modification may serve to increase the coding potential of the genetic material by providing additional information.

## **1.2 DNA methylation in the prokaryotes**

The inception of the recombinant DNA era is partly the result of intensive research on bacteria and phage molecular biology. It is therefore not surprising that the biological functions of DNA methylation are well studied in prokaryotes. Both 5-methylcytosine and N6-methyladenine have been detected in the DNA of *E.coli* (Dunn and Smith, 1955). These modified bases are the products of Hsd (host specificity), Dcm (DNA cytosine methylation) and Dam (DNA adenine methylation) methyltransferases (reviewed in Noyer-Weidner and Trautner, 1993). Two well-documented functions of DNA methylation in *E. coli* are: the host defence restriction/modification system involving N6-methyladenine methylation by Hsd methyltransferase and the methyl-directed mismatch repair system involving Dam methyltransferase. The biological function of Dcm methylation is however less clear.

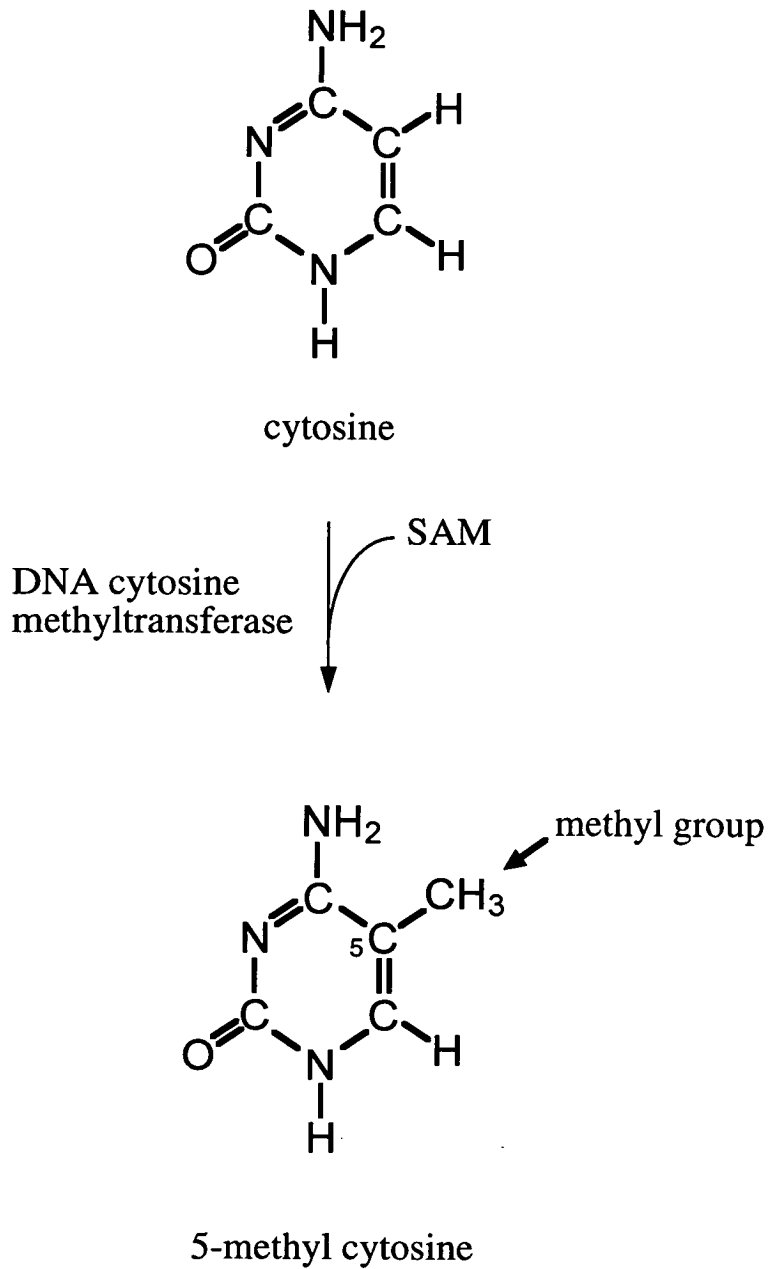


Figure 1.1 Enzymatic methylation of cytosine by DNA cytosine methyltransferase.

S-adenosyl methionine (SAM) is the methyl group donor for the reaction. Carbon 5 position of the cytosine is indicated.

The restriction/modification (R/M) system serves as a mechanism to protect the genetic information from foreign phage DNA invasion (reviewed in Noyer-Weidner and Trautner, 1993). The R/M system consists of two components: 1) a methylation-sensitive restriction endonuclease that cleaves DNA at specific recognition site and 2) a modification methyltransferase that recognises the cognate restriction site and methylates a specific base (either adenine or cytosine depending on the bacteria). Methylation of target sites in the chromosomal DNA of the bacterium protects against “self” restriction. On the other hand, the unmodified DNA of invading phage will be degraded by the host restriction endonuclease. Therefore, the bacterial R/M system is clearly involved in host defence.

The Dam methylation (methylation at GATC sites) is involved in directing strand specific DNA mismatch repair (Figure 1.2, reviewed in Modrich, 1986). Replication errors that arise during DNA replication can escape the proof-reading activity of DNA polymerase, and faithful correction of the mismatch relies on the recognition of parental DNA strand. The Dam methyltransferase methylates the N6 position of adenine at all the GATC sites. During a transient period after replication and before methylation by Dam methyltransferase, the hemi-methylated GATC sites serve as signals for the mismatch repair system to direct correction of mistakes on newly synthesised strand. The proteins involved in this postreplicative mismatch repair are MutS, MutL and MutH. Recognition of mismatch is mediated by MutS. Upon binding to a mispair, MutS recruits other components to the DNA. MutH binds to the hemimethylated GATC site and generates a nick immediately 5' to the GATC site on the unmethylated strand. MutL is an ATPase, but the role of ATP hydrolysis in mismatch repair is unclear (Worth et al, 1998). In addition, MutL has been shown to stimulate the binding of MutS to mispair and can activate the endonuclease activity of MutH (Drotschmann et al, 1998; Hall and Matson, 1999). Cleavage at the hemi-methylated DNA initiates exonucleolytic removal of the unmethylated strand. *De novo* DNA synthesis, followed by ligation and remethylation complete the DNA repair process and correct the replication associated mistake.

### **1.3 DNA methylation in the eukaryotes**

Cytosine methylation in animals occurs predominantly at the CpG dinucleotides. The plants are unique in that cytosine methylation can be found at both CpG and CpNpG sites where N is any nucleotide. Although the target for cytosine methylation is reduced to two (CpG) or three (CpNpG) basepairs, the eukaryotes differ from prokaryotes in that not all

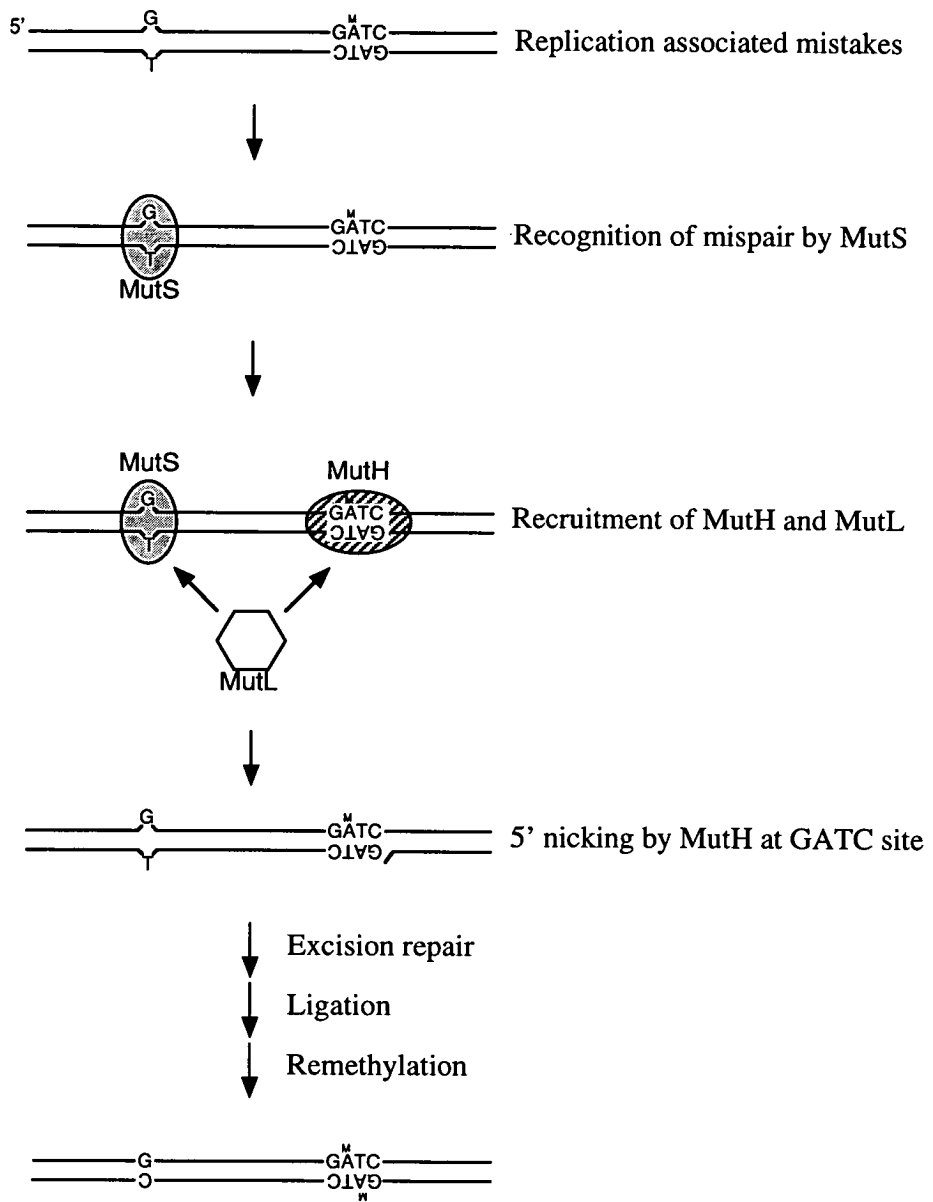


Figure 1.2. Methyl-directed repair system in *E. coli*.

methylatable sites are methylated. However, non-symmetrical methylation also exists in plants and certain fungi.

### 1.3.1 Inheritance of DNA methylation

Two methyltransferase activities can be distinguished biochemically. The *de novo* methyltransferase activity adds a methyl group to previously unmodified DNA, whereas the maintenance methyltransferase activity transfers the methyl group to the non-methylated cytosine at a hemi-methylated site (Figure 1.3.1). During semiconservative DNA replication, the symmetrically methylated CpG or CpNpG will become hemi-methylated due to the incorporation of unmodified nucleotides at the newly synthesised strand. Thus, post-replicative restoration of hemi-methylated sites by the maintenance methyltransferase activity is proposed to ensure faithful duplication of the methylation pattern in the daughter cells (Riggs, 1975; Holliday and Pugh, 1975).

The mammalian DNA cytosine methyltransferase 1 (DNMT1) has both activities, but maintenance methyltransferase activity is the predominant catalytic activity as hemi-methylated CpG sites are the preferred substrates *in vitro* (Bestor and Ingram, 1983). Localisation studies showed that DNMT1 is associated with replication foci during S phase (Leonhardt et al, 1992; Chuang et al, 1997). Therefore, targeting of a maintenance methyltransferase to the sites of replication can potentially explain the rapid remethylation of newly synthesised DNA (Gruenbaum et al, 1983).

### 1.3.2 Distribution of cytosine methylation in eukaryotes:

#### Fungi

Lower eukaryotes such as fungi *Neurospora crassa* and *Ascobolus immersus* have non-CpG (asymmetrical) methylation (Selker and Stevens, 1985; Selker et al, 1993). Such non-CpG methylation is associated with *de novo* methylation of duplicated or repeated DNA sequences. In *Neurospora*, the phenomenon is known as repeat-induced point mutation (RIP), and is characterised by repeated DNA sequences trigger G:C to A:T mutations. The mutated sequences become methylated at cytosine residues through *de novo* methylation. A very similar process in which duplicated DNA segments are *de novo* methylated is also observed in *Ascobolus*, and is termed MIP (methylation induced premeiotically). Unlike RIP, the methylation and inactivation of duplicated genes in *Ascobolus* is reversible and not associated with mutations (reviewed in Rossignol and Faugeron, 1995).

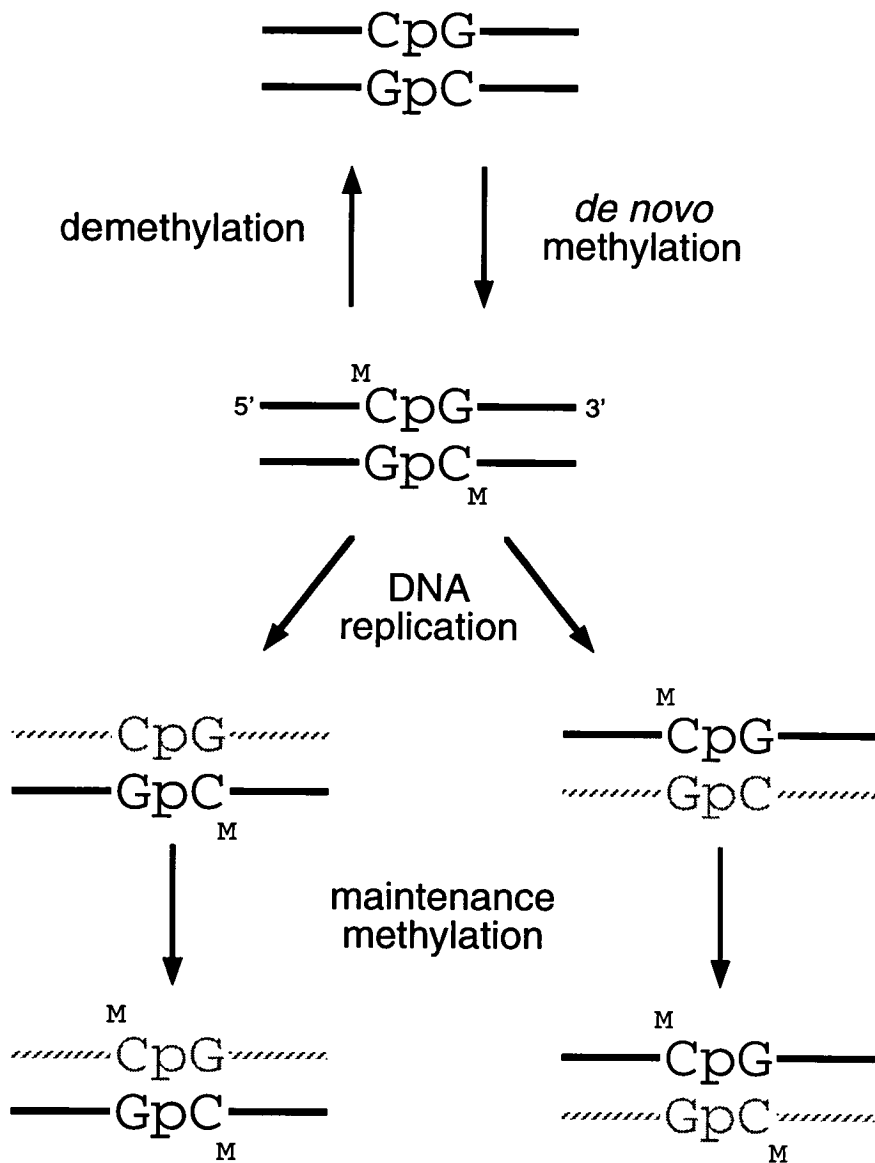


Figure 1.3.1 Different states of DNA methylation.

*De novo* methylation by *de novo* methyltransferase results in the addition of methyl groups at non-modified CpG sites. Previously methylated DNA can lose methylation through DNA demethylation. After DNA replication, the newly synthesised strands (dotted lines) contain un-modified cytosine residues (shaded in gray). Maintenance methylation by maintenance methyltransferase restores a fully methylated site by methylating the un-modified cytosine residues at hemi-methylated site.



## **Invertebrates and Vertebrates**

The genomes of invertebrates, apart from those without detectable 5-meC, contain a majority of non-methylated DNA and a minor fraction of methylated DNA (Bird et al, 1979; Bird and Taggart, 1980; Tweedie et al, 1997). However, the genomes of vertebrates are globally methylated, i.e. DNA methylation is distributed almost everywhere. Digestion of DNA from vertebrates with methylation sensitive restriction enzyme HpaII (CCGG) results in the release of methylation free, CG-rich tiny fragments, termed HpaII tiny fragments (HTF) (Cooper et al, 1983). These HTFs are derived from CpG islands, which are 500 bp to 1 kb in length and commonly associated with promoters or 5' regions of housekeeping genes. Essentially, the vertebrate genomes exhibit a characteristic bimodal distribution of DNA methylation, with methylation-free CpG islands and methylation of bulk genomic DNA.

## **Plants**

In contrast to the vertebrate genomes with 3-8 % of the cytosine residues methylated, the level of methylated cytosine residues in plants is highly variable and can be as high as 30 % (Shapiro, 1975). The genomes of some higher plants also contain CG rich clusters resembling CpG islands of vertebrates, as revealed by HpaII restriction digestion (Antequera and Bird, 1988).

Based on the diverse patterns of DNA methylation in eukaryotes, it is likely that DNA modification plays different, but perhaps not mutually exclusive, roles in different organisms.

## **1.4 Functions of DNA methylation in vertebrates**

### **1.4.1 Profile of genomic methylation during mouse development**

The level and pattern of DNA methylation during mammalian development are not static (Monk et al., 1987). During early embryonic development at the blastula stage, most of the genome undergoes global demethylation. The erasure is then followed by genome wide *de novo* methylation by the time of implantation. The restoration of methylation pattern is presumably catalysed by a *de novo* methyltransferase. The role of dynamic demethylation and remethylation of genomic DNA during mouse development remains unclear. Interestingly, dynamic fluctuation in methylation level appears to be absent in certain vertebrates such as the zebrafish (Macleod et al, 1999). CpG islands which are

enriched in CpGs and commonly associated with 5' end of genes, are however immune to the remethylation process. It is found that for both mouse and rat *Aprt* genes, Sp1 sites within the *Aprt* CpG island are necessary for keeping the island methylation-free (Macleod et al, 1994; Brandeis et al, 1994). Deletion or mutagenesis of the Sp1 sites results in *de novo* methylation of the island. It is proposed that transcriptional activity of the promoter is required to maintain methylation-free CpG islands.

#### **1.4.2 Mammalian DNA cytosine methyltransferases**

Several DNA cytosine-5 methyltransferases (DNMT1 to 3) exist in mammals (Figure 1.4.2). DNMT1 is a 190 kDa protein with a C-terminal catalytic domain related to bacterial cytosine methyltransferases and a novel N-terminal regulatory domain not found in the bacterial enzymes (Bestor et al, 1988). As hemimethylated substrates are highly preferred over non-methylated substrates, DNMT1 is an attractive maintenance methyltransferase candidate for restoring the DNA methylation pattern after DNA synthesis. Most importantly, DNMT1 is found associated with sites of DNA replication, and replication foci targeting is mediated through two distinct regions at the N-terminal domain (Leonhard et al, 1992; Chuang et al, 1997; Liu et al, 1998). DNMT2 was discovered by searching the EST databases for proteins with homology to the bacterial cytosine methyltransferases (Okano et al, 1998a; Yoder and Bestor, 1998; Van den Wyngaert et al, 1998). Unlike DNMT1, DNMT2 does not have a N-terminal domain, and is more related to the putative yeast cytosine methyltransferase, PMT1 (Wilkinson et al, 1995). Several lines of evidence indicate that DNMT2 is not an active cytosine methyltransferase: 1) purified recombinant protein from *E.coli* does not transfer methyl group to different DNA substrates (Van den Wyngaert et al, 1998), 2) DNMT2 expressed in insect cells or mammalian cells is not active (Van den Wyngaert et al, 1998; Okano et al, 1998a), 3) inactivation of the DNMT2 gene in mouse cells by gene targeting does not lead to demethylation of endogenous viral DNA (Okano et al, 1998a). The most recently discovered methyltransferases in mammals are the DNMT3 family, comprising DNMT3 $\alpha$  and DNMT3 $\beta$ . These proteins were also identified by EST database screening (Okano et al, 1998b). Both DNMT3 $\alpha$  and DNMT3 $\beta$  have a novel N-terminal extension before the C-terminal catalytic domain. Northern analysis reveals that they are highly expressed in embryonic stem (ES) cells, and not abundant in differentiated embryoid bodies or somatic tissues. Intriguingly, DNMT3 $\alpha$  and DNMT3 $\beta$  expressed in insect cells have comparable *de novo* and maintenance methyltransferase activities *in vitro*, i.e. there is no preference for hemi-methylated DNA over non-methylated DNA as the methyl group acceptor. It should however be noted that the reported methyltransferase activity of DNMT3 $\alpha/\beta$  is

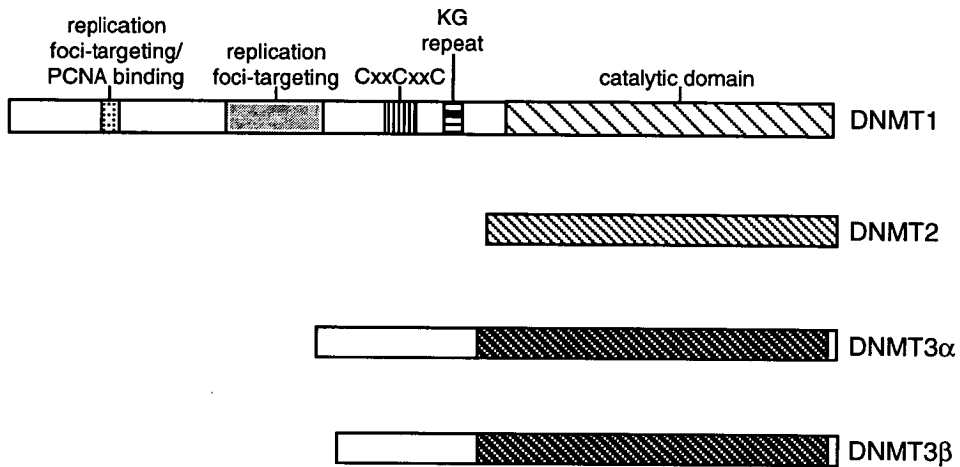


Figure 1.4.2. Mammalian DNA cytosine methyltransferases (DNMTs).

DNMT1 is the largest member of the cytosine methyltransferase family. Two regions in DNMT1 can target the protein to replication foci during S-phase. The catalytic domains are shaded in hatch. DNMT2 has not been shown to catalyse the transfer of methyl group to cytosine. DNMT3 $\alpha$  and DNMT3 $\beta$  are related proteins based on sequence homology but are encoded by two different genes. Both DNMT3 $\alpha$  and DNMT3 $\beta$  have comparable *de novo* and maintenance methyltransferase activities.

very low *in vitro* (Okano et al, 1998b). Thus, there are two major differences between DNMT1 and DNMT3 $\alpha/\beta$ . Firstly, DNMT1 is ubiquitously expressed in a wide variety of cell-types and tissues (Bestor et al, 1988). Secondly, hemi-methylated DNA is the preferred substrate for DNMT1. It is well established that extracts from cell-lines and tissues contain predominantly maintenance methyltransferase activity, and this observation correlates with the expression profile of DNMT1 and DNMT3 $\alpha/\beta$ . As certain cell-types such as ES cells have relatively high level of *de novo* methylation activity (Stewart et al, 1982; Lei et al, 1996), it is tempting to speculate that DNMT3 $\alpha$  and DNMT3 $\beta$  which are abundant in these cell-types, are responsible for the observed activity.

### **1.4.3 Role of DNA methylation in development**

The role of DNA methylation in mammals can be critically assessed by studying mutants with reduced levels of genomic methylation. The gene encoding a major player of the mammalian DNA methylation system, DNMT1 has been disrupted by gene targeting (Li et al., 1992). Homozygous mutant embryos suffer from marked reduction in both genomic methylation and DNA methyltransferase activity, and do not survive beyond mid-gestation. This result indicates that DNA methylation is essential for mouse development. How and why reduced genomic methylation leads to embryonic lethality is largely unknown. Transgenic plants with reduced genomic DNA methylation have been generated by using antisense strategy to target the *Arabidopsis* methyltransferase MET1 (Finnegan et al, 1996; Ronemus et al, 1996). Interestingly, these plants also exhibit developmental and phenotypic abnormalities, which include changes in developmental timing with prolonged vegetative and reproductive phases, smaller plant size, and reduced fertility.

## **1.5 Dynamics of DNA methylation**

As mentioned in above section, the DNA methylation pattern during mouse development is not static. Although DNA methylation is somatically heritable (Shmookler Reis and Goldstein, 1982; Stein et al, 1982), it is also known that inheritance of methylation can be stochastic and does not occur with 100 % fidelity (Wigler et al, 1981; reviewed in Riggs et al, 1998). This is likely to give rise to interclonal methylation heterogeneity observed by some investigators (Fitzpatrick et al, 1998; Zhu et al, 1999). Fluctuation of methylation could be accounted for by a demethylation process that results in the removal of methyl groups from methylated DNA (Figure 1.5).

Several models can be proposed for DNA demethylation:

1) Dilution model (Passive demethylation)

Anti-5-meC staining profiles of mouse metaphase chromosomes from the one-cell stage to blastocyst stage have been analysed (illustrated in Figure 1.5 A, Rougier et al, 1998). At the one cell stage, sister chromatids are symmetrically labeled, indicating a similar distribution of 5-meC on both chromatids. Curiously, only one of the two chromatids of each chromosomes is labeled by the antibody at the 2-cell stage. The asymmetrical labeling indicates that the overall distribution of genomic methylation between sister chromatids is not reproduced after chromosome duplication. This is opposite to that which would be expected if maintenance methylation restores methylation after each round of DNA replication (Riggs, 1975; Holliday and Pugh, 1975). More importantly, the number of asymmetrically stained chromosomes is halved after each chromosomal division and the intensity of labeling decreases progressively, indicating that methylation is lost through cell divisions. One plausible mechanism to explain this finding is that as DNMT1 is not localised in the nuclei before the 8-cell stage, maintenance methylation does not occur at this time (Carlson et al, 1992). The result supports the model that DNA methylation is lost through passive demethylation during early development due to failure of maintenance methylation.

Two recent studies also lend support to a passive mode of DNA demethylation, which is dependent on DNA replication (Figure 1.5 B, Matsuo et al, 1998; Hsieh, 1999). Binding of nuclear factors (either transcription factors or EBNA-1) to the DNA was shown to be necessary for the loss of DNA methylation after DNA replication. It is hypothesized that the nuclear factors may block the access of maintenance DNA methyltransferase to the hemi-methylated sites that are formed during DNA replication. Consequently, fully-methylated sites cannot be restored, and the remaining hemi-methylated sites are either lost by dilution through further rounds of DNA replication, or may be substrate for active DNA demethylation (see below).

An alternative way to prevent re-methylation of hemi-methylated sites is by preventing the maintenance methyltransferase from localising to replication foci. For example, the mitotic inhibitor, p21 (WAF1/Cip1) is shown to compete for binding of DNMT1 to replication auxiliary factor, PCNA *in vitro* (Chuang et al, 1997). If such displacement of DNMT1 from PCNA can occur *in vivo*, hemi-methylated sites can escape re-methylation by DNMT1 after DNA replication. This hypothesis remains to be tested.

A)

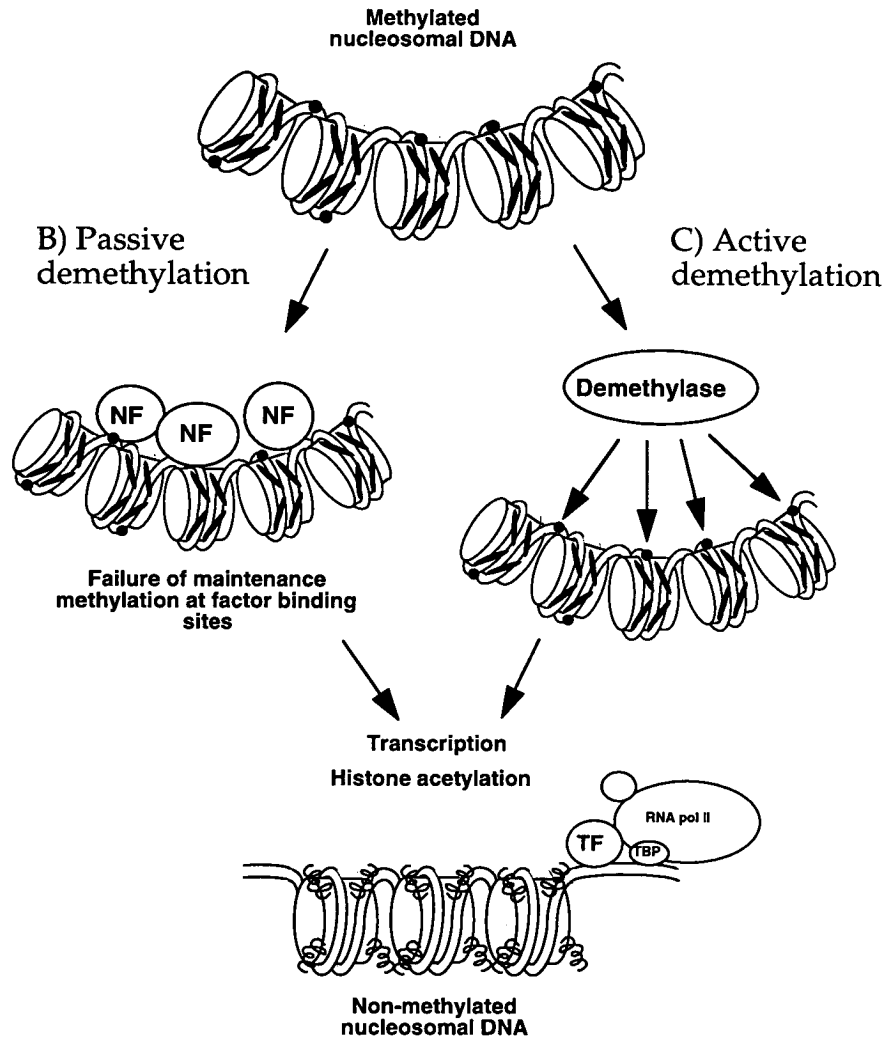
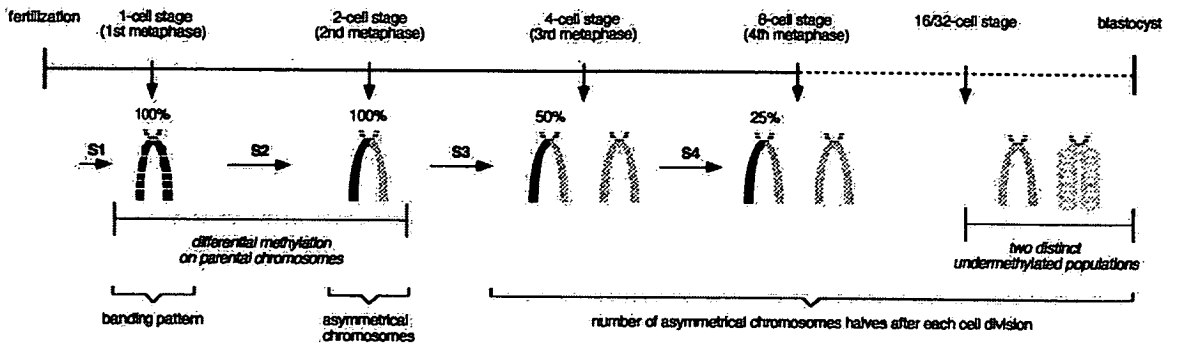


Figure 1.5 Mechanisms of DNA demethylation

A) Passive demethylation during mammalian preimplantation development. S1 to S4: S phases. (adapted from Rougier et al, 1998. *Genes Dev.* 12, 2108-2113.)

B) Passive demethylation through binding of nuclear factors (NFs) at methylated sites. TF, transcription factor.

C) Active demethylation catalysed by demethylase.

## 2) Active demethylation

A wide variety of biological processes utilise reversible modification to achieve regulation. Examples include the phosphorylation and dephosphorylation of proteins by kinases and phosphatases in many signal transduction pathways. Therefore, it is conceivable that methylated DNA can be demethylated through a process of active enzymatic removal of methyl groups by a demethylase. Indeed, demethylase activities have been reported from extracts of vertebrate in origin.

A methylated  $\alpha$ -actin gene becomes demethylated and transcriptionally active upon transfection into rat L8 myoblasts (Paroush et al, 1990). As demethylation can occur in the absence of DNA replication, loss of 5-meC (passive demethylation) through rounds of replication is unlikely. The demethylase activity in this experimental system has been characterised (Weiss et al, 1996). Demethylation appears to be carried out by a nucleotide exchange reaction which may involve RNA molecules as revealed by the sensitivity of the reaction towards RNase treatment. Unfortunately, it was later realised that the RNase sensitivity is an artifact due to usage of large amount of RNase (100  $\mu$ g/ml), which is likely to “coat” the methylated DNA substrates and inhibit the demethylase activity (Swisher et al, 1998). The demethylation reaction under study is no longer thought to involve RNA cofactors or ribozyme components.

Jost (1993) discovered an activity in the nuclear extracts of chicken embryos that mediate the removal of 5-meC through a nucleotide excision repair pathway. The reaction is initiated by 5' nicking from 5-methyldeoxycytidine, and is followed by excision repair to replace the 5-meC with cytosine. Subsequent work showed that a 5-methylcytosine-DNA glycosylase is responsible for this activity. Intriguingly, this enzyme also contains a mismatch-specific thymine DNA glycosylase activity. A recurring theme in the demethylation studies is the involvement of RNA. RNA was found to copurify with the 5-meC-DNA glycosylase activity through all chromatographic steps (Fremont et al, 1997). The copurified RNA can be recovered at exactly the same position of the enzyme on a polyacrylamide-SDS gel. Reconstitution experiments demonstrate that both the RNA and protein components are required for the 5-meC-DNA glycosylase activity (Jost et al, 1999).

Recently, a demethylase activity has also been studied in human cell extracts (Ramchandani et al, 1999; Cervoni et al, 1999). The demethylase can hydrolyse 5-methyl

cytosine to cytosine and methanol (Ramchandani et al, 1999). This reaction would involve the breaking of a carbon-carbon bond between the methyl group and cytosine, and grafting of free methyl group to a water molecule. This novel demethylase activity from human cells has exceptional biochemical and unprecedented catalytic properties. Intriguingly, a truncated version of a methyl-CpG binding protein, MBD2 (MBD2b) has been reported to contain a novel 5-methyl cytosine demethylase activity (Bhattacharya et al., 1999). MBD2b is shown to catalyse the conversion of methyl-CpG to CpG by the direct removal of methyl-group from methylated DNA. It is however not determined if MBD2b is responsible for the demethylase activity reported in Ramchandani et al (1999) and Cervoni et al (1999). The identification of a mCpG demethylase marks a new chapter in the field of DNA methylation. Antagonizing action between the DNA cytosine methyltransferases and demethylase may potentially be involved in intricate regulation of gene expression during development and tumorigenesis. It is also tempting to speculate if the MBD2 demethylase is involved in demethylation of mouse genome before implantation. Characterisation of MBD2 as part of the thesis work however does not support the finding that MBD2b is a demethylase (see Chapter 4 and Ng et al, 1999).

## **1.6 5-meC is a mutagenic hotspot**

Both cytosine and 5-meC are inherently unstable and can deaminate spontaneously *in vitro* and *in vivo* to give rise to uracil and thymine, respectively (Duncan and Miller, 1980; Wang et al, 1982; Frederico et al, 1990; Shen et al, 1994). The mutation rate of 5-meC is however 10- to 40- fold higher than cytosine (Rideout III et al, 1990). With the assumption that the half-life of 5-meC is 30 000 years, it is estimated that eight 5-meC deamination events occur per day per diploid human genome (Jones et al, 1992). Rideout III et al (1990) showed that mutational hotspots for CpG transitions at codon 175, 273, and 282 of the p53 tumour suppressor gene are methylated in white blood cells, sperm and urothelial cell DNA, suggesting that 5-meC may be responsible for causing elevated mutations at these sites. It is also evident that conversion of 5-meCpG to TpG is a relatively frequent event, as CpG deficiency of various animal genomes is matched with a proportional TpG/CpA excess (Bird, 1980). Therefore, spontaneous loss of 5-meC is likely to contribute to CpG suppression in organisms with heavily methylated genomes. Methylation-free CpG islands are however less prone to mutagenic replacement from 5-meC to T and retain their characteristic CpG richness.



The question is: how does the cell cope with the endogenous mutagen 5-meC? Mammalian cells have two known proteins that can directly repair the G/T mismatch which arises when 5-meCpG/GpmeC is deaminated to TpG/GpmeC. The first identified protein is the thymine DNA glycosylase (TDG) which can excise the mispaired thymine (Wiebauer and Jiricny, 1989; 1990). The apyrimidinic (AP) site will subsequently be replaced with cytosine through a repair pathway that involves AP endonuclease, DNA polymerase  $\beta$  and DNA ligase. The second mammalian protein with similar activity is the *methyl-CpG binding domain protein 4*, MBD4 (Hendrich and Bird, 1998; Hendrich et al, 1999). MBD4 contains a glycosylase domain which is distinct from TDG. The glycosylase domain is shown to specifically remove thymine residue from a G/T mismatch, but has no activity towards C/C or G/G mismatches. Interestingly, MBD4 also contains a functional methyl-CpG binding domain, and ectopically expressed protein is localised to methylated DNA *in vivo*. The exact role of the DNA binding domain in relation to the thymine glycosylase activity is not clear. It is also shown that the methyl-CpG binding domain of MBD4 can bind to a T/G mismatch in addition to a mCpG pair. The involvement of MBD4 in DNA repair is also strengthened by the finding that MBD4 can interact with hMLH1, a component of the mammalian MutSL mismatch repair system (Bellacosa et al, 1999). Based on a weak endonuclease activity detected in MBD4, it is speculated that MBD4 may be the long-sought MutH homolog of the bacterial methyl-directed mismatch repair system (see Figure 1.2, Bellacosa et al, 1999). Both TDG and MBD4 can also cleave uracil from a G/U mismatch *in vitro*. Based on the specificity of these two proteins towards excision of thymine/uracil from mismatches that can arise through deamination of CpG or 5-meCpG, it is likely that they are involved in counteracting the mutagenic potential of 5-meC. Beside TDG and MBD4, the mammalian cells contain a novel G/T binding protein (GTBP, also known as hMSH6 and p160) which is implicated in mediating G/T mismatch repair (Palombo et al, 1996; Papadopoulos et al, 1995; Drummond et al, 1995). GTBP interacts with hMSH2 to form a heterodimer hMutS $\alpha$  that can recognise various abnormal structures such as base-base mismatches and small insertion/ deletion loops. Unlike TDG or MBD4, GTBP is not a glycosylase and is likely to play a role in detection of G/T mismatches. Recognition of G/T mismatch by GTBP presumably leads to recruitment of downstream DNA repair proteins and correction of the defect. Therefore, the mammalian cells may have at least three distinct pathways to target G/T mismatches.

## **1.7 DNA methylation and gene regulation**

### **1.7.1 Evidence that DNA methylation is associated with transcriptional repression**

One known fact about DNA methylation in vertebrates is its incompatibility with transcriptional activity. Work carried out over the last 20 years has established an inverse correlation between DNA methylation and gene expression (Razin and Riggs, 1980). For example, CpG island-containing genes on the X chromosomes become methylated following X chromosome inactivation in eutherian mammals (Grant and Chapman, 1988). In addition, several lines of evidence indicate that DNA methylation is involved in gene silencing. Firstly, DNA demethylation, induced by an inhibitor of DNA cytosine methyltransferase, 5-azacytidine, causes reactivation of previously methylated, silenced autosomal or X-linked genes (Jones and Taylor, 1980; Mohandas et al, 1981) and retroviruses (Groudine et al, 1981). Similarly, removal of methylation through propagation of provirus in bacteria restores infectivity (Harbers et al, 1981). More direct evidence that DNA methylation prevents gene expression has been obtained by methylating specific genes *in vitro* and testing them either in *in vitro* transcription systems (Shen, 1984) or transfecting them into mammalian cells (Busslinger et al, 1983).

The availability of a DNMT1 null mouse model with reduced genomic methylation allows one to study the effect of demethylation on endogenous gene expression (Li et al, 1992). Imprinted genes are differentially methylated and silenced depending on their parental origin, and in the mutant embryos these genes are mis-expressed (Li et al., 1993). This is the first demonstration that DNA methylation is essential for regulating differential expression of the paternal and maternal copies of imprinted genes.

In the mammalian female, one of the two X-chromosomes is selectively inactivated to achieve dosage compensation for having two X chromosomes (Lyon, 1999). *Xist* encodes a untranslated nuclear RNA which is transcribed from and colocalised to the inactive X chromosome. In female somatic cells, inactive *Xist* on the active X chromosome is highly methylated, while active *Xist* on the inactive X chromosome is hypomethylated, suggesting that DNA methylation may be involved in allele-specific expression of *Xist* (Norris et al, 1994; Beard et al, 1995). Indeed, in male DNMT1 null mouse embryos, *Xist* becomes demethylated and expressed. *Xist* is also aberrantly expressed in female DNMT1 null mouse embryos, and *Xist* RNA coats both X

chromosomes (Beard et al, 1995; Panning and Jaenisch, 1996). Therefore, loss of methylation can lead to reactivation of *Xist* in somatic cells.

Intracisternal A particle (IAP) retroviruses are normally methylated at the LTR regions and transcriptionally silent. In DNMT1- mutants, it is found that transcription of IAP retroviruses is elevated and mRNA transcripts are detected in most tissues of the embryos (Walsh et al., 1998). In the wildtype embryos, these elements are silenced, presumably by DNA methylation. Curiously, northern analysis of 3 tissue-specific genes ( $I\alpha 1$  collagen, skeletal muscle  $\alpha$ -actin, and  $\beta$ -globin) did not reveal an up-regulation of transcripts in DNMT1- embryos even when these genes are undermethylated (Walsh and Bestor, 1999). Hence, it is suggested that DNA methylation does not play an important role in the regulation of tissue-specific genes. With the advent of DNA microarray technology, it is now possible to analyse changes in gene expression with greater precision (for example: see Iyer et al., 1999). This technology should help to clarify any mis-expression of tissue specific genes in DNMT -/- embryonic stem cells or embryos.

Recently, an elegant animal model system was used to study the consequence of inducible methylation of regulatory elements on transcription *in vivo* (Siegfried et al., 1999). Sp1 sites contained within the promoter region of both the hamster and mouse *Aprt* genes have been shown to protect the *Aprt* CpG island from *de novo* methylation during development (Brandeis et al., 1994; Macleod et al., 1994). By flanking Sp1 sites-containing island elements (IE) with *loxP* sites and using Cre-mediated excision, it has been possible to conditionally remove the IE and bring about *de novo* methylation of a promoter. Transgenic mice harboring a *loxP*-flanked IE reporter construct have been generated. Introduction of Cre recombinase before implantation by genetic crossing induces the removal of IE. The promoter region downstream of the *loxP* site becomes methylated by the wave of *de novo* methylation during implantation and transcription from the reporter is diminished. Removal of the IE after implantation does not lead to methylation of the reporter construct and transcription activity is not significantly altered as compared to the founder animal with an intact IE. These results indicate that DNA methylation represses transcription in animal. However, questions arise as to whether the result can be extrapolated to endogenous genes and what mechanism of repression is conferred by methylation using this artificial system.

### **1.7.2 DNA methylation: a cause or consequence of gene silencing?**

“Is methylation dancing to the piper or calling the tune?”. This question still awaits more experimental investigation (Bird, 1992). Several studies support DNA methylation as a secondary event that follows transcriptional silencing: a) the X-linked mouse *Hprt* genes remain unmethylated for several days after X-inactivation (Lock et al, 1987), b) proviral transcriptional inactivation in embryonal carcinoma cells occurs before *de novo* methylation (Gautsch and Wilson, 1983; Niwa et al, 1983), c) inactivation of  $\gamma$ -globin gene takes place before methylation (Enver et al, 1988). Why are transcriptionally inactive genes susceptible to *de novo* methylation? What are the hallmarks of inactive genes that attract *de novo* methylation? It is likely that investigation along this line will uncover the relationship between transcriptional inactivation and *de novo* methylation. It should also be emphasised that DNA methylation is not necessarily an overriding factor over gene expression. For example, it was found that developmental activation of the H-2K gene correlates with an increase in methylation (Tanaka et al, 1983).

## **1.8 DNA methylation and chromatin modification**

In eukaryotic cells, DNA is packaged along with other chromosomal proteins into chromatin (van Holde, 1989). This allows for the efficient compaction of genomic DNA within a nucleus. The repeating units of chromatin are nucleosomes. Each nucleosome consists of 147 bp of DNA wrapped round a histone octamer. Since the *in vivo* substrate for transcription is chromatin, it is not surprising that many transcriptional regulators impact directly on chromatin infrastructure to exert an influence on transcription.

### **1.8.1 Methylated and non-methylated DNA adopt fundamentally different chromatin structures**

The mammalian genome is characterised by methylation-free CpG islands interspersed within the methylated bulk chromatin. Early studies indicated that chromatin containing high levels of 5-meC is resistant to micrococcal nuclease digestion in chicken nuclei (Razin and Cedar, 1977). On the other hand, actively transcribed sequences are found to be undermethylated compared with bulk DNA (Naveh-Manly and Cedar, 1981). Upon transfection and integration of methylated DNA into the mouse L cell genome, the DNA becomes resistant to DNase I and restriction endonuclease digestion (Keshet et al, 1986). Upon integration, the non-methylated version remains sensitive to nuclease digestion. These studies provide direct evidence that chromatin containing methylated DNA has a

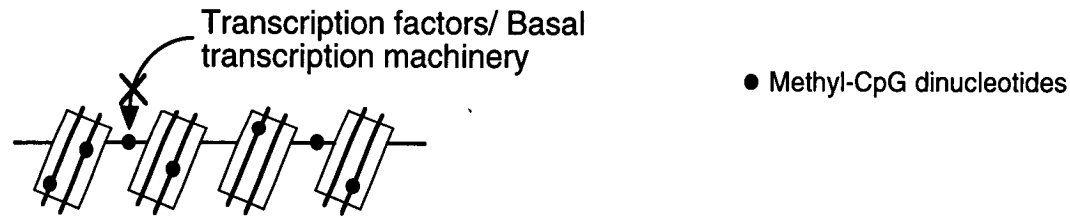
distinctive structure which is resistant to nuclease digestion. Using restriction enzymes that recognise CpG sites but are not sensitive to cytosine methylation, it is found that accessibility towards methylated sites in the nuclei is blocked (Antequera et al, 1989). Under similar conditions, bulk chromatin can be extensively cleaved by restriction enzymes that do not contain CpG in their recognition sites. The CpG restriction enzymes do however release oligonucleosomes that are exclusively derived from CpG islands. Biochemical analysis of CpG oligonucleosomes reveals that the chromatin is enriched in highly acetylated histones H3 and H4, but is deficient in histone H1 (Tazi and Bird, 1990).

Is the inactive chromatin structure associated with DNA methylation important for the loss of transcriptional activity? This question is critically addressed by experiments analyzing the kinetics of transcriptional arrest from methylated templates (Buschhausen et al, 1987). Both methylated and non-methylated versions of the herpes simplex virus thymidine kinase (HSV TK) gene are equally active during the first 8 hours after microinjection into rodent cells. It is only after longer time intervals (more than 8 hours) that the inhibitory effect of DNA methylation was exerted on the methylated reporter construct, as measured by the accumulation of HSV TK RNA. It is hypothesised that the injected DNA is not fully assembled into chromatin during the early stages, and that DNA methylation status has no effect on transcriptional activity in the absence of chromatin. Indeed, when an *in vitro* reconstituted chromatinised methylated reporter is introduced into the cells, it becomes transcriptionally inert. Subsequently Kass *et al* showed clear evidence for a time-dependent repression of methylated template following injection into *Xenopus* oocytes (Kass et al, 1997). Loss of transcriptional activity from the methylated template coincides with the appearance of a nucleosomal array in the vicinity of the promoter, and disappearance of engaged RNA polymerases. The above studies demonstrate that DNA methylation can lead to alteration of chromatin structure, and that in some cases, DNA methylation *per se* is not sufficient for transcriptional repression.

### **1.9 How does DNA methylation repress transcription?**

A most simplistic model to explain why DNA methylation is inhibitory to transcription is that mCpG prevents transcriptional initiation by interfering with binding of transcription factors to their cognate recognition sites (Figure 1.9A). Indeed, the two basepair CG can be found in many transcription factor recognition sites, and DNA methylation has been shown to inhibit the binding activity of various proteins such as E2F, NF $\kappa$ B, etc

A) Direct interference of transcription by methyl-CpGs



B) MeCPs mediated transcriptional repression (Indirect repression)

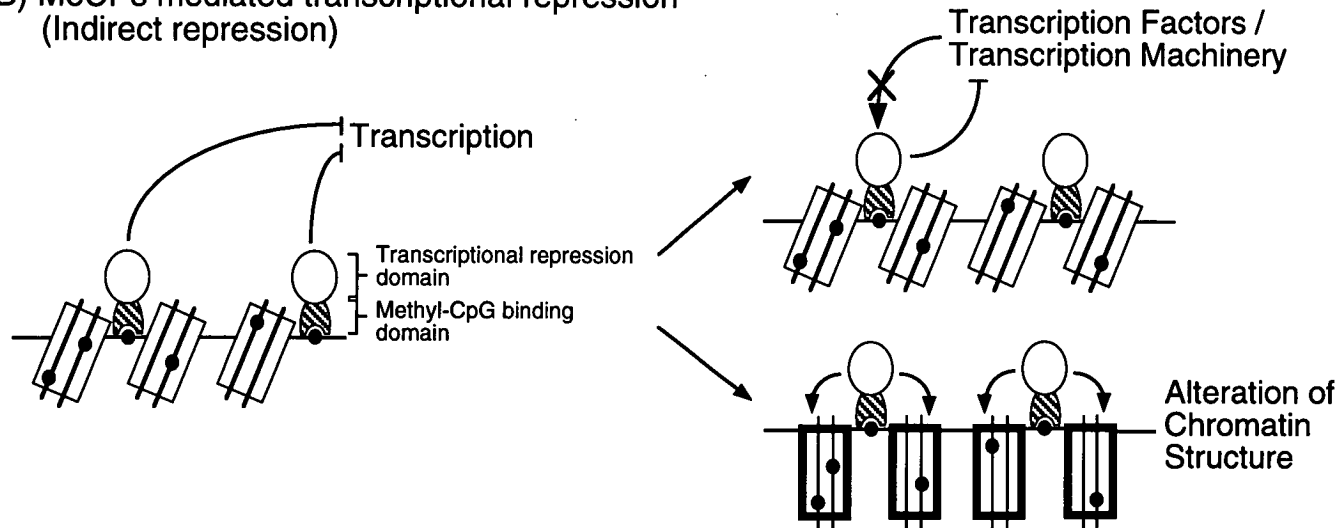


Figure 1.9. How does DNA methylation repress transcription?

A) 5-meC inhibits the binding of transcription factors or basal transcription machinery to DNA.

B) Indirect repression by MeCPs.

MeCPs are proposed to contain two functional domains (transcriptional repression domain and methyl-CpG binding domain). MeCPs can either communicate with components of transcription apparatus or alter chromatin structure to silence gene expression.

(reviewed in Tate and Bird, 1993). In addition, DNA methylation can also influence the translational positioning of nucleosomes on certain DNA sequences *in vitro* (Davey et al, 1997). Methylation induced occlusion of essential regulatory elements by nucleosomes can in theory provide a way to repress transcription. It remains to be tested whether this mechanism actually operates *in vivo*. Repression through such mechanisms is also referred to as “direct” inhibition of transcription.

Certain transcription factors such as Sp1 are however not affected by methylation within their recognition sites, and not all sequence-specific transcription factor sites contain CpG (Harrington et al, 1988; Holler et al, 1988). Therefore, a second hypothesis needs to be invoked to explain how DNA methylation leads to transcriptional repression as well as protection of mCpG in mammalian nuclei (Antequera et al, 1990). Proteins with affinity for methylated DNA are attractive candidates for mediating indirect repression by DNA methylation (Figure 1.9 B).

MeCP1 (methyl-CpG binding protein 1) was one of the first methyl-CpG binding activity detected in various mammalian nuclear extracts and can bind to 12 or more symmetrically methylated CpGs in any sequence context (Meehan et al, 1989). The molecular nature of MeCP1 is not defined, but the functional significance of this activity has been studied both *in vitro* and *in vivo* (Boyes and Bird, 1991). The extent of repression on methylated reporters in living cells correlates strongly with the density of methylation as well as *in vitro* affinity of MeCP1 for the methylated reporters. Repression on the methylated reporters can also be rescued by co-transfection of methylated ligands for MeCP1. More importantly, cells and extracts deficient in MeCP1 failed to repress the methylated reporters (Boyes and Bird, 1991; Levine et al, 1991). These studies implicate MeCP1 as a methylation-dependent transcriptional repressor.

MeCP2 is another protein with an affinity for methylated DNA, but differs from MeCP1 in that it can bind to a single methylated CpG pair (Lewis et al, 1992; Meehan et al, 1992; Nan et al, 1993). There is significant variation in the abundance of MeCP2. Murine brain nuclear extracts contain the highest levels of MeCP2 as revealed by Southwestern assay and immunoblotting with anti-MeCP2 antibodies (Meehan et al, 1992; Nan et al, 1997). MeCP2 is also a chromosomal protein that localised to mCpG-riched heterochromatin in mouse cells (Lewis et al, 1992; Nan et al, 1996). Deletion studies of MeCP2 led to the mapping of a 85 amino-acids methyl-CpG binding domain (MBD) (Nan et al, 1993). That the MBD of MeCP2 is also required for methyl-CpG binding *in vivo* was critically

addressed by a later study to look at the *in vivo* localisation of MeCP2 (Nan et al, 1996). In mouse cells, ectopically expressed MeCP2-LacZ fusion protein was targeted to heterochromatin which is enriched with 5-methylcytosine. Deletion within the MBD region of MeCP2-LacZ fusion protein lead to inefficient targeting to heterochromatin. The results strongly suggest that the MBD of MeCP2 is required for preferential association with methyl-CpG *in vivo*. Beside binding to naked DNA, MeCP2 has also been shown to interact specifically with methylated chromatin (Nan et al, 1997). However, it is not certain if MeCP2 binds to the methylated, exposed linker DNA or interacts directly with methyl-CpGs on the surface of nucleosomes. More recently, MeCP2 was shown to footprint to specific methyl-CpG pairs on nucleosomal DNA in a highly defined system (Chandler et al, 1999). The results further strengthen the idea that MeCP2 is an integral component of chromatin and has abundant binding sites in the mammalian genome (Nan et al, 1997).

Several other proteins have also been shown to bind methylated DNA. A methylated DNA binding protein (MDBP1) has been detected in nuclear extract from human placenta (Huang et al, 1984). Further characterisation of this activity reveals that in addition to methylation specificity, the protein also binds to specific fragments within the prokaryotic plasmid vectors, pBR322 and M13mp8 (Wang et al, 1986). The consensus binding sequence can be found in the mammalian genome and viral DNA, and methylation in these sequences enhances DNA binding but is not an absolute requirement. The 5-meC can also be substituted by T (5-methyluracil) for MDBP1 binding (Zhang et al, 1989). As MDBP1 is a sequence specific DNA binding protein, and is present in undifferentiated embryonal carcinoma cells which do not efficiently repress methylated reporters (Boyes and Bird, 1991; Levine et al, 1991), it is unlikely to play a role in the global repression of methylated sequences (Supakar et al, 1988).

MDBP2 is an avian protein that binds both *in vitro* and *in vivo* to the methylated promoter of avian vitellogenin II gene (Pawlak et al, 1991; Jost and Hofsteenge, 1992). The binding of MDBP2 is not sequence-specific and requires a minimum length of 30 bp with a symmetrically methylated CpG. Interestingly, MDBP2 turns out to be a histone H1 subtype (Jost and Hofsteenge, 1992). Phosphorylation of MDBP2 is crucial for its affinity towards methylated DNA, and more importantly, the phosphorylated form is shown to repress transcription of the vitellogenin II promoter *in vitro* (Bruhat and Jost, 1996). However, it is not clear if MDBP2 is an active repressor *in vivo*. The demonstration that an avian histone H1 subtype can bind specifically to methylated DNA



prompted an investigation of the methylated DNA binding property of mammalian histone H1. The answer is still unclear and the findings are controversial. While two groups find that histone H1 had no preference for methylated nucleosomal or naked DNA (Nightingale and Wolffe, 1995; Campoy et al, 1995), some groups observed an effect of histone H1 on naked methylated DNA templates (Levine et al, 1993; Johnson et al, 1995; McArthur and Thomas, 1996).

An EST databases searching strategy was used to find novel proteins with homology to the methyl-CpG binding domain (MBD) of MeCP2. The rationale is that the MBD is likely to be a structurally conserved domain for recognition of methyl-CpG pairs. A family of novel mammalian proteins with homology to the MBD (MBD1 to 4), has been identified using this approach (Cross et al, 1997; Hendrich and Bird, 1998). Northern analysis reveals that MBD1 and MBD2 are expressed in a variety of tissues but the transcripts are deficient in ES cells. Transcripts for MBD3 and MBD4 can however be detected in ES cells. Using bandshift assays, MBD1, MBD2 and MBD4 are shown to bind specifically to mCpG pairs in any DNA sequence context. The heterochromatin in mouse cells is enriched in heavily methylated major satellite DNA which is preferentially stained by DAPI. Ectopic expression of MBD1, MBD2 and MBD4 results in the localisation of these proteins to the DAPI bright spots, and suggests that these proteins bind to methylated DNA *in vivo* (Hendrich and Bird, 1998). Ectopically expressed MBD2 and MBD4 failed to localise to heterochromatic foci in DNMT1 *-/-* ES cells with low levels of genomic methylation, indicating that methylation is an essential parameter for their localisation in the cells. With the same analysis, MBD1 remains localised to DAPI bright spots, suggesting that protein factors may be involved in *in vivo* localisation, or MBD1 is targeted by residual methylated sites. Although MBD2 and MBD3 are related as revealed by the high level of sequence conservation at the C-terminal region (including the MBD), MBD3 does not bind to methylated DNA specifically either *in vitro* and *in vivo* using the same set of binding assays for MBD1, MBD2 and MBD4. Therefore, among the mammalian methylated DNA binding proteins, only MeCP1, MeCP2, MBD1, MBD2 and MBD4 bind to methylated DNA in a non-sequence specific manner, and exhibit strict requirement for mCpG.

Until recently, the molecular mechanism by which DNA methylation represses transcription remained largely unknown. However, a few reports hinted at a connection between DNA methylation and histone deacetylation. In one instance, an inhibitor of histone deacetylase, sodium butyrate was shown to elevate the expression of a methylated

episomal reporter gene (Hsieh, 1994). It was also reported that both sodium butyrate and trichostatin A could substitute for the DNA demethylating agent, 5-aza-2'-deoxycytidine (5-aza-dC) in restoring transcription from previously methylated and silenced plant ribosomal RNA genes (Chen and Pikaard, 1997). The interchangeable effect of TSA and a DNA demethylating agent prompted the notion that DNA methylation and histone deacetylation may operate along a common mechanistic pathway to silence transcription.

### **1.10 Post-translational modification of histones: lysine acetylation**

It is known that core histones can be subjected to several form of post-translational modification, e.g. methylation, phosphorylation, ubiquitination, and acetylation. Histone acetylation has attracted much attention based on an early observation that histone acetylation is linked to RNA synthesis, and it is proposed that one way to regulate transcription of chromatin templates is through acetylation of histones (Pogo et al, 1966). Using an antibody against acetylated histone H4, it was demonstrated that transcriptionally active genes are associated with acetylated core histones (Hebbes et al, 1988). The link between histone acetylation and transcriptional activation is further substantiated by the finding that *Tetrahymena* histone acetyltransferase A is homologous to the yeast transcriptional activator, Gcn5p (Brownell et al, 1996). Both *Tetrahymena* histone acetyltransferase A and Gcn5p possess a novel histone acetyltransferase (HAT) activity that can acetylate core histones *in vitro*. Gcn5p is also shown to acetylate specific lysines of histones H3 and H4 *in vivo* (Kuo et al, 1996). Point mutations within the HAT domain abolish the capability of Gcn5p to direct histone acetylation and mediate transcriptional activation *in vivo* (Kuo et al, 1998).

Many biological regulatory pathways make use of antagonising processes to achieve different states of activity, histone modification by acetylation is no exception. The removal of acetyl groups from core histones is catalysed by histone deacetylases (Libby, 1970; Boffa et al, 1971; Krieger et al, 1974). Trapoxin is a fungal product that can inhibit histone deacetylase activity *in vitro* and can cause accumulation of hyperacetylated core histones *in vivo* (Kijima et al, 1993). A trapoxin affinity matrix was used to purify mammalian nuclear proteins with affinity for trapoxin. Intriguingly, one of the two proteins that binds to the trapoxin, HD1 is related to the yeast transcriptional regulator, Rpd3 (Taunton et al, 1996). Deletion of yeast Rpd3 leads to an increase in acetylation of core histones H3 and H4 *in vivo* (Rundlett et al, 1996). Rpd3 interacts with another transcriptional corepressor Sin3 *in vivo* (Wang and Stillman, 1993; Kadosh and Struhl,

1997). Targeted recruitment of both Rpd3 and Sin3 to promoters leads to transcriptional repression, and repression is reduced or abolished in yeast strains with Rpd3 or Sin3 mutations (Kadosh and Struhl, 1997). The yeast DNA binding repressor Ume6 is found to interact with the Rpd3/Sin3 complex, and repression of Ume6-regulated genes is sensitive to both Rpd3 and Sin3 mutation. Mutations within the putative catalytic motif of Rpd3 destroyed histone deacetylase activity *in vitro* and repression of Ume6 regulated genes *in vivo* (Kadosh and Struhl, 1998a). Therefore, histone deacetylase activity is essential for repression by Rpd3 *in vivo*. In addition to Ume6, a variety of mammalian DNA binding repressors are shown to interact with the mammalian mSin3/ HDAC1/ HDAC2 complex. mSin3 is the mammalian homologue of yeast Sin3 corepressor protein, while HDAC1 and HDAC2 are the mammalian homologues of yeast histone deacetylase RPD3. The proteins that interact with the mammalian mSin3 corepressor complex include the Mad/Max and Mxi1/Max heterodimers (Ayer et al, 1995; Schreiber-Agus et al, 1995; Laherty et al, 1997; Alland et al, 1997; Heinzl et al, 1997; Hassig et al, 1997; Zhang et al, 1997), unliganded nuclear hormone receptors (Heinzl et al, 1997; Nagy et al, 1997), AML1-ETO fusion (Gelmetti et al, 1998; Lutterbach et al, 1998; Wang et al, 1998), PML-RAR $\alpha$  and PLZF-RAR $\alpha$  fusion (Grignani et al, 1998; He et al, 1998; Lin et al, 1998), Aiolos and Ikaros (Koipally et al, 1999), and neural-specific repressor, REST (Huang et al, 1999).

The proteins mentioned in the above list have been shown to interact with the mSin3/ HDAC complex and in many cases repression can be relieved by histone deacetylase inhibitors (trichostatin A or trapoxin). However, only two reports demonstrate that tethering of the repressors (Rb, Aiolos and Ikaros) to the reporters leads to reduction in acetylated histones as revealed by chromatin immunoprecipitation experiments (Luo et al, 1998; Koipally et al, 1999). As these experiments were performed by transient transfection of reporters into mammalian cells, it is not certain if the transfected DNA is assembled into proper nucleosomal arrays (Wu, 1997). Two independent studies using the yeast system show that targeting of Rpd3 to a promoter leads to localised deacetylated nucleosomal domain that spans one to two nucleosomes (Kadosh and Struhl, 1998b; Rundlett *et al*, 1998). This result suggests that short range deacetylation of chromatin is sufficient to repress transcription, and begs the question as to how histone deacetylation represses transcription.

### 1.10.1 The mSin3 and NuRD histone deacetylase complexes

Biochemical fractionation of *Xenopus* oocyte and mammalian nuclear extracts revealed two major histone deacetylase complexes, namely mSin3 and NuRD complexes (Figure 1.10.1, Wade et al, 1998; Wade et al, 1999; Zhang et al, 1998a). The mSin3 complex as defined by immunoprecipitation with antibodies against different subunits, contains at least 7 subunits mSin3, HDAC1, HDAC2, RbAp48, RbAp46, SAP30, and SAP18 (Zhang et al, 1997; Zhang et al, 1998b). The mammalian NuRD (Nucleosome Remodeling and Deacetylation) complex contains 7 subunits: Mi-2, MTA2, HDAC1, HDAC2, RbAp48, RbAp46, and MBD3 (Zhang et al, 1998a; Zhang et al, 1999). HDAC1 and HDAC2 are the catalytic subunits found in both complexes. Co-immunoprecipitation experiments established that HDAC1 and HDAC2 are associated in mammalian cells (Hassig et al, 1997; Zhang et al, 1997; Hassig et al, 1998). The two other common subunits RbAp46 and RbAp48, have been shown to bind directly to helix 1 of histone H4 which is not accessible when H4 is incorporated into nucleosomes (Verreault et al, 1998; Vermaak et al, 1999). This is consistent with the finding that the mSin3 complex is active in deacetylating core histone octamers, but unable to deacetylate nucleosomal histones (Zhang et al, 1998b). The probable explanation is that the histone binding subunits RbAp46 and RbAp48 are unable to access nucleosomal histones.

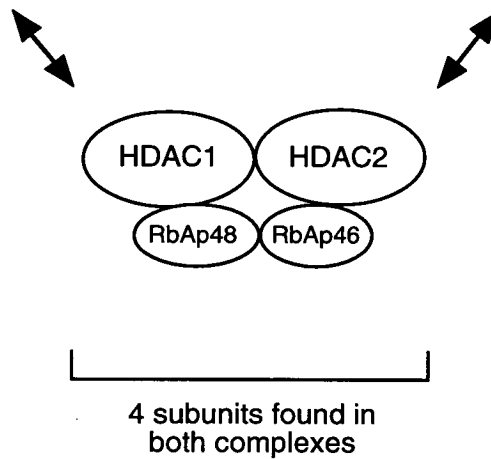
The NuRD complex differs from the mSin3 complex by possessing a novel ATP-dependent nucleosome remodeling activity, and the activity can be attributed to the presence of Mi2, a member of the SWI/ SNF2 helicase/ ATPase family (Zhang et al, 1998a; Tong et al, 1998; Xue et al, 1998; Wade et al, 1998). Importantly, it is shown that ATP stimulates deacetylation of oligonucleosomal histones *in vitro*, suggesting that hydrolysis of ATP leads to nucleosome disruption and aids the deacetylases to access its substrates. Several proteins are known to interact with either the NuRD complex or a subunit of the complex. For example, the E7 oncoprotein of human papilloma virus type 16 has been shown to interact with the Mi2 subunit of the NuRD complex using yeast two hybrid assay (Brehm et al, 1999). Ikaros, a lymphoid lineage-determining factor is also found associated with the NuRD complex in T cells as revealed by co-fractionation and co-immunoprecipitation (Kim et al, 1999). Upon T cell activation, both Ikaros and NuRD complex are co-localised to centromeric heterochromatin, suggesting that Ikaros brings about the silencing of inactive genes and centromeric heterochromatin in T cells through histone deacetylation. The same group also reported the interaction between Ikaros and mSin3 based on a yeast two hybrid assay and co-immunoprecipitation (Koipally et al, 1999). The potentially confusing findings may be reconciled by the proposal that Ikaros

## mSin3 complex

mSin3A

SAP30

SAP18



## NuRD complex (Nucleosome Remodeling and Deacetylation)

Mi2 (contains SWI2/SNF2 helicase/ ATPase domain)

MTA2 (Metastasis associated protein 2)

MBD3 (contains MBD-like region, related to MBD2)

Figure 1.10.1 mSin3 and NuRD histone deacetylase complexes.

Both complexes share a core histone deacetylase complex (HDAC1, HDAC2, RbAp46 and RbAp48). Unique subunits of each complex are shown.

may recruit distinct histone deacetylase complexes to specific promoters during lymphocyte development.

## **Thesis aims:**

To determine whether methyl-CpG binding proteins repress transcription and to elucidate the mechanism of repression.

The thesis centres on the functional studies of three methyl-CpG binding proteins, MeCP2, MBD2 and MBD1. Chapter 3 describes the mechanism of repression by MeCP2. Chapter 4 describes the identification of components of MeCP1 and how MBD2 which is the methyl-CpG binding factor in MeCP1 represses transcription. Chapter 5 describes the functional characterisation of MBD1.

## **Chapter 2: Materials and Methods**

### **2.1 General buffers**

TE

10 mM Tris.HCl (pH 8.0), 1 mM EDTA (pH 8.0)

PBS

8mM Na<sub>2</sub>HPO<sub>4</sub>, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, 147 mM NaCl, 3 mM KCl

TAE

40 mM Tris.acetate, 1 mM EDTA, pH 7.2

TBE

90 mM Tris base, 90 mM boric acid, 2 mM EDTA

### **2.2 Growth and maintenance of *Escherichia coli* cells**

*E. coli*. cells were grown in Luria-Bertani (LB) medium (10 g bacto-tryptone, 5 g bacto-yeast extract, 10 g NaCl per litre). The *E.coli* strain XL1 Blue (*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac*, [F' *proAB*, *lacI<sup>q</sup>ZΔM15*, Tn10 (tet<sup>r</sup>)]). )was routinely used for cloning and propagation of plasmids. BL21 (DE3) pLysS *E. coli* B F' *dcm ompT hsdS* (r<sub>B</sub>. m<sub>B</sub>.) *gal* λDE3 [pLysS Cam<sup>r</sup>].

### **2.3 General methods**

#### **2.3.1 Agarose gel electrophoresis**

Horizontal agarose gel electrophoresis was used for analysing DNA molecules and restriction digests. DNA was loaded with 10 % volume of 10 x loading buffer (0.1 % orange G, 15 % ficoll type 400 in TE buffer) and separated at 100 V on agarose gel (0.5 to 1.8 % depending on the size of DNA) in 1 X TAE buffer containing 0.5 µg/ ml ethidium bromide. The resolved DNA was visualised using a UV-transilluminator.

#### **2.3.2 Recovery of DNA fragments**

After separation of DNA fragments (either PCR products or restriction fragments) by electrophoresis, the band of interest was located with a long-wavelength UV lamp, and excised from the gel using a razor blade. 3 volumes of QG buffer (Qiagen) was added to the gel slice, and heated to 50 °C to dissolve the gel. The mixture was then passed



through a Qiaquick Spin Column (containing a silica-gel based membrane for binding DNA under high concentration of chaotropic salt and low pH conditions). Loading of the column 2 more times was found to increase the amount of DNA bound to the column. The column was washed once with QG buffer and twice with PE buffer (Qiagen). DNA was eluted with either TE or 10 mM Tris.HCl pH 8.5.

### **2.3.3 Purification of DNA fragment**

Cloning procedures routinely require the exchange of buffers and removal of restriction enzymes or other enzymes such as Klenow or calf intestinal phosphatase. Five volumes of PB buffer (Qiagen) was added to the DNA solution and passed through the Qiaquick Spin Column twice. The column was washed once with PB buffer and twice with PE buffer. DNA was eluted with either TE or 10 mM Tris.HCl pH 8.5.

### **2.3.4 Restriction endonuclease digestion, dephosphorylation, and ligation**

Endonuclease digestion of DNA were performed under conditions and buffers recommended by the manufacturers. Typically, 1 µg of DNA was digested with 5 units of restriction enzyme in a 20 µl reaction mixture. Calf intestinal alkaline phosphatase (CIP) was used to remove phosphate group from the 5' end of linearised vector DNA to prevent self-ligation. Typically two rounds of dephosphorylation were carried out to ensure minimal background resulting from self-ligation of vector DNA. Ligation was carried out with buffer supplied by the manufacturer. For sticky end ligation, the reaction was performed at 25 °C for 4 hours before transforming competent bacterial cells. For blunt end ligation, the reaction was incubated at 16 °C for 16 hours.

## **2.4 Plasmid minipreparation**

A bacterial colony was picked and inoculated into 5 ml of LB (with the appropriate antibiotics for selection) for 16 hours at 37 °C. The bacterial cells were pelleted by centrifugation. The pellet was resuspended in 250 µl of P1 buffer (100 µg/ml RNaseA, 50 mM Tris.HCl, 10 mM EDTA, pH 8.0). 250 µl of P2 buffer (0.2 M NaOH, 1 % SDS) was added to lyse the bacteria, and this was followed addition of N3 buffer (Qiagen) to neutralise the lysate and precipitate genomic DNA and proteins. The lysate was clarified by centrifugation at 14 000 rpm for 10 min at room temperature. The supernatant was then loaded onto Qiagen Miniprep column. The column was washed once with PB buffer

and twice with PE buffer. The plasmid DNA was eluted with either TE or 10 mM Tris.HCl pH 8.5.

## **2.5 Large scale plasmid preparation**

Method 1 (ion-exchange column):

A bacterial colony was picked and inoculated into 200 ml of LB and incubated overnight at 37 °C. The bacterial cells were harvested by centrifugation and resuspended in 10 ml of P1 buffer. 10 ml of P2 buffer was added and mixed gently. After 5 min of incubation at room temperature, 10 ml of ice-cold P3 buffer (2.55 M sodium acetate, pH 4.8) was added. The lysate was clarified by filtration and supernatant was loaded onto a pre-equilibrated Qiagen 500 column. The column was washed twice with QC buffer (1 M NaCl, 50 mM MOPS pH 7.0, 15 % ethanol). The bound plasmid DNA was eluted with 15 ml of QF buffer (1.25 M NaCl, 50 mM MOPS pH 8.2, 15 % ethanol). The DNA was precipitated by addition of 10.5 ml of isopropanol. After centrifugation and 70 % ethanol wash, the dried DNA pellet was resuspended in TE buffer.

Method 2 (CsCl gradient):

Bacterial cells from an overnight 500 ml culture was pelleted, and resuspended in 10 ml of solution I (50 mM glucose, 25 mM Tris.HCl, pH 8.0, 10 mM EDTA). Chicken egg white lysozyme (Boehringer) was added to a final concentration of 40 mg/ml. Lysis was allowed proceed for 10 min at room temperature. 20 ml of solution II (0.2 M NaOH, 1 % SDS) was added and the mixture was incubated for 5 min at room temperature. 15 ml of ice-cold solution III (3M sodium acetate, pH 5.2) was added and the mixture was left on ice for 10 min. Supernatant removed after centrifugation at 5000 rpm, was precipitated with 0.6 volume of isopropanol. The precipitate was recovered by centrifugation (5000 rpm for 15 min) and washed once with 70 % ethanol. The air-dried pellet was resuspended in 10 ml of TE. Cesium chloride and ethidium bromide were added to the solubilised pellet at final concentration of 1 g/ml and 0.8 mg/ml respectively. The solution was poured into two 5.1 ml Quick Seal polyallomer tubes (Beckman) which were subsequently sealed by heat. The tubes were spun in a NVT 90 rotor (Beckman) at 20 °C according to the following program:

- 90 000 rpm for 90 min
- 87 000 rpm for 15 min
- 83 000 rpm for 15 min
- 81 000 rpm for 30 min

80 000 rpm for 30 min

The plasmid band was isolated from the tube by a sterile syringe and needle. The DNA solution was extracted for a few rounds with 10 ml of isoamyl alcohol. TE was sometimes added when the CsCl started to precipitate due to excessive dehydration by isoamyl alcohol. The DNA was precipitated by 3 volumes of 70 % ethanol at room temperature. The DNA pellet was washed in 70 % ethanol and resuspended in TE. A second round of ethanol precipitation was carried out with 2.5 volume of 100 % ethanol in the presence of 0.3 M sodium acetate pH 5.2. The pellet was washed again in 70 % ethanol, dried, and stored in TE.

## **2.6 DNA sequencing**

Fluorescent dye-primer sequencing reactions were performed using a Perkin Elmer/ Applied Biosystems ABI PRISM Dye Primer Cycle Sequencing Kit with AmpliTaq DNA polymerase. Each reaction contained 1  $\mu$ l primer (30 ng/ $\mu$ l), 4  $\mu$ l sequencing premix, and 5  $\mu$ l of DNA solution (500 ng of template). The reactions were subjected to 25 sequencing cycles of 98 °C for 60 s, 55 °C for 30 s, and 60 °C for 4 min. The mixture (10  $\mu$ l) was then precipitated by addition of 1  $\mu$ l of 3M sodium acetate (pH 4.8) and 25  $\mu$ l of 95 % ethanol. After centrifugation at 14 000 rpm for 20 min, the DNA pellet was washed once with 70 % ethanol. The DNA samples were analysed using an automated 373 DNA sequencer (Applied Biosystems Perkin-Elmer). The data collected were analysed using the GeneJockey II software (Biosoft Cambridge).

## **2.7 Radioactive labeling of DNA**

Klenow end filling method:

Typically about 100 ng of CG11 sticky end fragment was labeled using 2 U of Klenow, in the presence of 2  $\mu$ l 3000 Ci / mmol [ $\alpha$   $^{32}$ P] dATP (20  $\mu$ Ci), 2  $\mu$ l 3000 Ci / mmol [ $\alpha$   $^{32}$ P] dCTP (20  $\mu$ Ci), 2  $\mu$ l 3000 Ci / mmol [ $\alpha$   $^{32}$ P] dTTP (20  $\mu$ Ci), and 0.5 mM cold dGTP at 25 °C for 1 hour. The reaction buffer was 50 mM Tris.HCl (pH 7.5), 10 mM MgSO<sub>4</sub>, 0.1 mM DTT in a total volume of 40  $\mu$ l. The reaction was stopped by adding 5  $\mu$ l of 0.5 M EDTA, and DNA was ethanol precipitated in the presence of 2  $\mu$ g of tRNA. The DNA/RNA pellet was washed two times with 70 % ethanol to remove unincorporated dNTPs. After drying at room temperature for 1 hour, the pellet was resuspended in TE buffer. 2  $\mu$ l of the resuspended probes were aliquoted into scintillation vials and counted in a Beckman scintillation counter. The specific activity of the probe was usually 5 x 10<sup>7</sup> cpm/ $\mu$ g.

## **2.8 PCR**

Reactions were performed using 200 ng of template DNA with specific primers at 0.1 nmol in a 100  $\mu$ l volume. Pfu DNA polymerase was used for amplification of inserts for cloning because of its proofreading activity that leads to reduced errors. Each reaction consisted of 200  $\mu$ M of each dNTPs, 1.5 mM MgCl<sub>2</sub>, and 5 U of Pfu polymerase. The samples were overlaid with mineral oil (Sigma) and placed in a Techne PHC2 thermoblock. 22 amplification cycles with parameters of 94 °C/ 45 seconds, 55 °C/ 45 seconds, 72 °C/ 10 seconds to 2 minutes (depending on the length of insert to amplify) per cycle was used.

## **2.9 Preparation of electro-competent cells**

A bacterial colony was picked and inoculated into 50 ml of LB for 16 hours at 37 °C. The culture was expanded to 1 litre with addition of fresh LB. The cells were allowed to grow at 37 °C, until mid-logarithmic phase (OD<sub>595 nm</sub> of 0.5). The flasks were chilled on ice for 10 min and cells were harvested by centrifugation at 4000 rpm for 15 min at 2 °C. The cell pellet was washed twice with 500 ml of ice cold sterile water. The cells were also washed twice with 25 ml of ice cold sterile 10 % glycerol. Finally the cells were resuspended in ice cold 10 % glycerol to a final volume of 2 ml. The resuspended cells were snap-frozen in liquid nitrogen in aliquots of 40  $\mu$ l.

## **2.10 Transformation**

Typically 1  $\mu$ l of ligation mixture was added to a thawed tube of electro-competent cells (40  $\mu$ l). The mixture was transferred to a chilled 0.2 cm electroporation curvette and subjected to a 1.7 kilovolt pulse (25  $\mu$ F, 200  $\Omega$ ) using a Gene Pulser apparatus (BioRad). The time constant was typically 4.5 to 4.6 ms. 500  $\mu$ l prewarmed SOC medium (2 % Bacto-tryptone, 0.5 % Bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose) was added immediately to the cells after electroporation. The cells were removed from the curvette and allowed to recover for 1 to 1.5 hr at 37 °C. The cells suspension (100 to 200  $\mu$ l) was finally plated on LB agar supplemented with antibiotics.

## **2.11 Cloning**

For eukaryotic expression of GAL4-fusion proteins:

PCR products (after cutting with Bgl II or BamHI) were cloned into pCMV-Gal4 BamHI cut vector. Positive clones were screened by PCR and integrity of inserts were determined

by sequencing. All clones used for transfection assays have insert fused in-frame with GAL4 DNA binding domain ORF.

For making mutation within the MBD:

Primers harboring mutated nucleotides were used to amplify inserts for cloning.

For prokaryotic expression of His-tagged or GST-fusion proteins:

The desired fragments were amplified by PCR with primers with the appropriate restriction sites. PCR products were digested and cloned into pET6H, pCT2, pGEX2T, pGEX3x vectors. pET6H vector was used for expression of N-terminal 6xHis tagged protein (Cross et al, 1997). pCT2 vector was used for expression of C-terminal 6xHis tagged protein. pGEX2T and pGEX3x vectors (Pharmacia) were used for expression of GST-fusion proteins (Smith and Johnson, 1988). Positive clones were screened by PCR and integrity of inserts were determined by sequencing.

For *in vitro* transcription and translation of <sup>35</sup>S methionine labeled protein:

The desired fragments were amplified by PCR with primers with the appropriate restriction sites. PCR products were digested and cloned into pCS2+MT vector (Laherty et al, 1997). This vector contains a SP6 polymerase binding site. Positive clones were screened by PCR and integrity of inserts were determined by sequencing. The vector can also be used for expression in mammalian cells, as the insert can be expressed by the CMV promoter upstream of the SP6 site.

## **2.12 Protein quantitation**

Bradford assay: All reagents required for this assay were found in the Bradford assay kit (Biorad). Protein samples (1 - 10  $\mu$ l) and BSA standards (1 to 10  $\mu$ g) were added to diluted Bradford reagent (1 in 10 dilution of a stock solution), and incubated for 5 minutes at room temperature. The mixtures were then transferred to plastic curvettes and OD<sub>595 nm</sub> measurements were taken. The blank was diluted Bradford reagent without protein. The BSA standard calibration was included for every protein concentration measurement.

## **2.13 SDS polyacrylamide gel electrophoresis (PAGE)**

SDS-polyacrylamide gels (Laemmli, 1970) were cast using Bio-Rad protein mini-gel II system. The resolving gel typically contained 7%, 10%, or 12 % acrylamide/bisacrylamide mixture (37.5:1), 375 mM Tris.HCl pH 8.8, and 0.1 % SDS. The stacking

gel contained 4 % acrylamide/ bisacrylamide mixture, 125 mM Tris.HCl pH 6.8, and 0.1 % SDS. The resolving gel was poured into the space between the two glass plates. The gel was layered with water and allowed to polymerise for 1 hour. The water was decanted and stacking gel was layered on top of the resolving gel. Protein samples were adjusted to 31 mM Tris.HCl pH 6.8, 1% SDS, 0.1 M  $\beta$ -mercaptoethanol and 5% glycerol. The samples were subjected to 98 °C for 3 minutes before loading into the protein gel. The gels were electrophoresed in 25 mM Tris, 250 mM glycine, and 0.1 % SDS buffer at 200 V for 45 minutes. After electrophoresis, the gel was either staining in Coomassie blue solution (0.1 % Coomassie brilliant blue dye, 45 % methanol, 10 % acetic acid) or blotted using Bio-Rad TransBlot Submarine apparatus.

## **2.14 Western blotting**

After SDS-PAGE, the protein gel was electroblotted to nitrocellulose membrane (20  $\mu$ m pore size, Schleicher and Schuell) using Bio-Rad TransBlot submarine transfer apparatus (2 hr at 200 mA). Pre-stained molecular weight markers (New England Biolabs) were used to monitor the transfer of proteins onto the nitrocellulose membrane. The membranes were blocked overnight at 4 °C in TBST buffer containing 5 % milk powder (w/v). The membranes were then rinsed in TBST buffer, and incubated with primary antibodies (at 1:500 to 1: 5000) in binding buffer (TBST with 1 % milk powder) for 1 hr at room temperature. After three 15 min washes in the same binding buffer, the membranes were incubated with secondary antibodies (1:2000 to 1:15 000) for 45 min. The secondary antibodies were either affinity purified anti-rabbit IgG-HRPO (Amersham), or monoclonal GT-34 anti-goat/ sheep IgG peroxidase conjugate (Sigma). The membranes were then washed three time in binding buffer, before detection with Enhanced Chemiluminescence (ECL) as described by the manufacturer (Amersham).

## **2.15 Expression and affinity purification of recombinant proteins**

### **Expression and purification of GST-fusion proteins**

Single colony was picked and cultured overnight at 37 °C. The overnight culture was diluted 1 in 100 in 500 ml of fresh LB. The bacteria were allowed to grow to an OD<sub>595 nm</sub> of 0.5, and expression was induced for 3 to 5 hours by addition of 0.2 mM IPTG (final concentration). The bacterial cells were harvested by centrifugation and washed once with ice-cold PBS. The pellet were resuspended in 15 ml of lysis buffer (containing 50 mM Tris.HCl (pH 8.0), 0.3 M NaCl, 0.1 % Triton X-100, 1 mM DTT, 10 % glycerol and protease inhibitors (0.5 mM PMSF, pepstatin A, leupeptin, chymotrypsin, and antipain at

5 µg/ml each) and lysed by ultrasound using BRANSON sonicator (set at power 4, 30 % cycle for 6 min) on ice. The lysate was clarified by centrifugation on a SS34 rotor at 18 000 rpm at 0 °C for 45 min. The supernatant was then incubated with 1 ml packed volume of rGSH-sepharose (Pharmacia) for 2 hr at 4 °C on a rotating platform. The sepharose beads were pelleted by centrifugation at 500 rpm for 3 min. The beads were washed three times in lysis buffer. Bound proteins were eluted in 15 mM rGSH in lysis buffer. Typically, three rounds of elution with rGSH were performed. The eluted fractions were analysed by SDS-PAGE. Purified proteins were stored in aliquots at -70 °C.

### **Expression and purification of His-tagged MBD1 (MBD1 with C-terminal 6 His tag)**

Clone CP1 (pCT2-MBD1 in BL21 DE3 pLyS) was cultured overnight at 37 °C. The culture was then grown in fresh media (under selection with 100 µg/µl ampicillin and 34 µg/µl chloramphenicol) to an OD<sub>595 nm</sub> of 0.5, and expression was induced for 3 hours by addition of 0.4 mM IPTG (final concentration). The bacterial cells were harvested by centrifugation and washed once with ice-cold PBS. The pellet were resuspended in lysis buffer (containing 50 mM HEPES.KOH (pH 7.9), 1 M NaCl, 0.1 % Triton X-100, 10 mM β-mercaptoethanol, 10 % glycerol and protease inhibitors (0.5 mM PMSF, pepstatin A, leupeptin, chymotrypsin, and antipain at 5 µg/ml each) and lysed by ultrasound using BRANSON sonicator (set at power 4, 30 % cycle for 6 min) on ice. The lysate was clarified by centrifugation on a SS34 rotor at 18 000 rpm at 0 °C for 45 min. The supernatant was supplemented with imidazole (final concentration of 25 mM) and loaded onto Ni<sup>2+</sup>-NTA superflow (Qiagen) column (pre-equilibrated with lysis buffer plus 25 mM imidazole). The column was washed stepwise with lysis buffers containing (0.8 M, 0.5 M, 0.3 M NaCl and 25 mM imidazole). The bound proteins were eluted in 50 mM HEPES.KOH (pH 7.9), 0.25 M imidazole, 0.3 M NaCl, 0.1 % Triton X-100, 10 mM β-mercaptoethanol, and 10 % glycerol. The eluted sample which contained four major proteins, was loaded directly onto Fractogel EMD S0<sub>3</sub><sup>e</sup>-650 (M) (Merck) strong cation-exchange column pre-equilibrated with 50 mM HEPES.KOH (pH 7.9), 0.3 M NaCl, 0.1 % Triton X-100, 10 mM β-mercaptoethanol, and 10 % glycerol. At this concentration of salt, only the full-length MBD1 protein preferentially bound to the column. The column was washed sequentially with the same buffer containing 0.3 M NaCl and 0.5 M NaCl. The full-length protein was eluted with 1 M NaCl, and dialysed sequentially against dialysis buffer (50 mM HEPES.KOH (pH 7.9), 0.1 % Triton X-100, 10 mM β-mercaptoethanol, and 10 % glycerol) containing 0.75 M NaCl, 0.5 M NaCl, and 0.25 M NaCl. The final concentration of protein was typically 50 ng/ µl. The optimal amount of

cation-exchanger was crucial for selective adsorption of full-length product. A bacterial protein of about the same size as full-length MBD1 was found to co-purify with MBD1. The purification procedure was monitored by analysing all the samples by SDS-PAGE. Purified proteins were stored in aliquots at -70 °C.

### **2.16 Antibody production**

Recombinant proteins purified from *E.coli* were used for immunisation of animals (either rabbits or sheep). Only protein preparations that were judged to be 98 % pure were used. Approximately 150 to 200 µg of proteins were used for each immunisation (either for priming or subsequent boost). The animals were primed by injection of antigen mixed with complete Freund's adjuvant. Boost injections were carried out at 3 weeks intervals with antigen and incomplete Freund's adjuvant. The animals were bled on the 7th day after each injection for serum donation. Injection of proteins into animals and processing of sera were carried out by Scottish Antibody Production Unit (SAPU).

### **2.17 Affinity purification of antibodies**

When sufficient antigens were available (more than 5 mg of purified proteins), affinity purification of the antisera was carried out. The choice of matrix (Affi-10 or Affi-15, Bio-Rad) depends on the iso-electric point of the protein of interest. Acidic proteins (as predicted based on amino acid composition) were coupled to Affi-15 activated matrix, while basic proteins were coupled to Affi-10 activated matrix. Typically, 8 mg of protein was coupled to 500 µl of matrix.

### **2.18 Maintenance of cell cultures**

HeLa cells, F9 EC cells, mouse L929 fibroblast cells were routinely cultured in Dulbecco's Minimal Eagles medium (DMEM, Gibco-BRL), supplemented with 10 % bovine calf serum, 100 U/ml penicillin and 100 µg/ml streptomycin at 37 °C in an atmosphere of 5 % CO<sub>2</sub>.

### **2.19 Making frozen stocks of cultures**

70-90 % confluent adherent cells were trypsinised, resuspended in fresh medium (with bovine calf serum to quench the activity of trypsin), pelleted by centrifugation at 1000 rpm for 3 minutes. The cell pellet was resuspended in 10 % DMSO/ DMEM with 10 % bovine calf serum and frozen at - 80 °C for 24 hours before storage in liquid nitrogen.



## **2.20 Thawing of frozen stock**

Frozen stocks were thawed rapidly at 37 °C, and the contents were transferred into a 15 ml Falcon tube with 12 ml of fresh medium (prewarmed at 37 °C). The cells were collected by centrifugation at 1000 rpm for 5 minutes. The cell pellet was then resuspended in prewarmed fresh medium and seeded in culture flasks at high density.

## **2.21 Transfection of mammalian cells**

Liposome-mediated transfection:

Transfection of adherent HeLa cells and F9 EC cells was carried out using Lipofectamine (Gibco-BRL) according to the manufacturer's instructions. 1.5 µl lipofectamine and 2 µg of plasmid DNA were used for transfecting a 35 mm dish of HeLa cells at 50 % confluence. 5 µg of plasmid DNA and 2 µl lipofectamine were used for transfecting a 60 mm dish of F9 EC cells at 30 % confluence.

DEAE-dextran mediated transfection:

Mouse L929 fibroblast cells were seeded at a density of  $5 \times 10^5$  per 60 mm (diameter) tissue culture dishes a day before transfection. Reporter plasmid (2 µg) were cotransfected with different amounts of effector plasmids. The total amount of plasmids per transfection was kept at 6 µg by including pCMV-BamNeo plasmid as carrier.

## **2.22 Cell staining**

HeLa cells were seeded onto 20 x 20 mm coverslips in 6-well tissue culture plate and allowed to grow to 80 % confluence overnight. The cells were rinsed in PBS before fixation with 3.7 % paraformaldehyde (Fluka)/ PBS pH 7.0 for 10 minutes at room temperature. After three washes with PBS, the cells were permeabilised with PBS supplemented with 0.2 % Triton X-100 for 12 minutes. Non-specific sites were blocked by incubation in 10 % rabbit serum in PBS for 10 minutes at room temperature. The coverslips were incubated with primary serum (anti-MBD1 or preimmune) at 1:200 dilution in PBS with 10 % rabbit serum for 1 hour at room temperature. After primary antibody binding step, the coverslips were washed three times in PBS. The coverslips were then incubated with secondary anti-sheep-FITC antibodies (Vector Laboratory) in 1:100 dilution in PBS with 10 % rabbit serum for 45 minutes at room temperature. After three washes with PBS, the coverslips were mounted with Vectashield media supplemented with DAPI (Vector Laboratory). Immunofluorescence was visualised and photographed on a Zeiss Axioplan POI fitted with Photometrics cooled CCD camera.

### **2.23 Reporter assays**

The cells were harvested 60 hours post-transfection.  $\beta$ -galactosidase activity or luciferase activity was measured using kits from Promega. For  $\beta$ -galactosidase assay, the substrate was o-nitrophenyl- $\beta$ -D-galactopyranoside, and the value for each  $\beta$ -galactosidase measurement was subtracted against the  $\beta$ -galactosidase reading for lysate obtained from mocked transfected cells. For the luciferase assay, the substrate was luciferin, and the value for each luciferase measurement was subtracted against the background light reading for lysate obtained from mocked transfected cells.

### **2.24 Isolation of nuclei from cells**

Cells grown as suspension culture were spun down at 1500 rpm for 5 minutes in 50 ml Falcon tubes. The cell pellets were washed two times with ice-cold PBS. The cell pellets were then resuspended lysis buffer (10 mM Tris.HCl pH 7.5, 10 mM NaCl, 2 mM  $MgCl_2$ ). 1 % NP40 (80  $\mu$ l/ ml) was then added to the resuspension to initiate cell lysis. Cell lysis was monitored under phase contrast microscopy to check for the release of nuclei. The nuclei were spun down at 3000 rpm at 0 °C in swing-out rotor JS 13.1 in 30 ml Corex tube for 5 minutes. The supernatant was then removed by vacuum suction. The nuclear pellet was resuspended in lysis buffer and layered onto 15 ml of 1.2 M sucrose in lysis buffer. The nuclei were pelleted by centrifugation at 5000 rpm at 0 °C for 5 minutes. The procedure was repeated again to purify the nuclei from detergent. The nuclear pellet was finally resuspended in nuclear freezing buffer (50 mM Tris.HCl pH 8.3, 5 mM  $MgCl_2$ , 0.1 mM EDTA pH 8.0, 42 % glycerol).

### **2.25 Preparation of nuclear extracts**

Nuclear extracts of mammalian cells were prepared according to the method by Dignam et al (1983). The procedure was carried out at 0 °C to 4 °C in the cold room to prevent protein denaturation and degradation. Freshly prepared or frozen nuclei were thawed on ice water. The nuclei were then pelleted by centrifugation at 1000 rpm for 10 minutes at 0 °C. The nuclear pellet was resuspended in 1.5 x packed volume of ice-cold extraction buffer (20 mM HEPES.KOH pH 7.9, 1.5 mM  $MgCl_2$ , 0.2 mM EDTA, 0.5 mM DTT, 25 % glycerol, 0.42 M KCl, 0.5 mM PMSF) by Dounce homogenizer (loose fit, 15 strokes). Nuclear proteins were extracted for 30 minutes at 0 °C with constant mixing. Chromatin and non-soluble proteins were pelleted by centrifugation for 45 minutes at 17 500 rpm at 0 °C in a SS34 rotor (Beckman). The supernatant was dialysed against 3

changes of 1 litre of dialysis buffer (20 mM HEPES.KOH pH 7.9, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 0.5 mM DTT, 20 % glycerol, 0.1 M KCl, 0.5 mM PMSF). The dialysed extract was cleared again by centrifugation for 30 minutes at 17 500 rpm at 0 °C in a SS34 rotor. The supernatant was aliquoted into small volume, snap-frozen in liquid nitrogen and stored at - 80 °C. The concentration of the proteins was typically 5 mg/ml using BSA as standard.

## **2.26 Bandshift**

Bandshift reactions with nuclear extracts were carried out with 4 to 6 µg of extracts in binding buffer (final volume of 30 µl; 20 mM HEPES.KOH pH 7.9, 1 mM EDTA, 3 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol, 4 % glycerol, 0.05 % bromophenol blue, 0.05 % xylene cyanol, and 1 % digitonin) containing 1 to 2 µg of sonicated *E.coli* competitor DNA (Cross et al, 1997). <sup>32</sup>P-labeled probe (approximately 1 fmol or 0.1 ng) was added to the extracts. The mixtures were then incubated on ice for 30 minutes before electrophoresis in 1.5 % agarose gel at 4 °C and 120 V in 0.5 x TBE buffer. The gel was dried onto Whatman DE81 paper before exposure onto X-ray film at -70 °C. The same protocol was used for bandshift assays with recombinant proteins, except that the amount of competitor DNA was varied. For antibody supershift experiments, 0.5 to 1 µl of antiserum was added to each reaction and incubated for 2 hours on ice before addition of the labeled probe. In certain experiments, 5 % polyacrylamide gel was used instead of 1.5 % agarose gel to resolve the DNA/ protein complex.

## **2.27 Immunoprecipitation**

Antibodies (monoclonal or affinity purified polyclonal antibodies) or serum was bound first to 25 µl of protein G sepharose (Pharmacia) at 4 °C for 1 hour on a rotating platform. Unbound antibodies or serum proteins were removed by washing the beads three times with PBS supplemented with 0.1 % Triton X-100. The beads were then incubated with HeLa nuclear extracts (100 µg to 200 µg) 4 °C for 5 hour on a rotating platform. The beads were then washed 5 times with wash buffer (50 mM Hepes.KOH pH 7.9, 150 mM NaCl, 10 % glycerol, 0.5 mM EDTA, 0.1 % Triton X-100, 1 mM DTT). The bound proteins were eluted by adding 12 µl of 2 x SDS loading buffer and heated at 98 °C for 5 minutes. The supernatant was then loaded onto SDS protein gel for subsequent Western blot analysis.

## **2.28 Coupled *in vitro* transcription and translation in TNT reticulocyte**

### **lysates (Promega)**

*In vitro* protein synthesis was carried out in 25  $\mu$ l volume using components supplied by the Promega TNT kit:

8  $\mu$ l nuclease free water

12.5  $\mu$ l TNT reticulocyte lysates

1  $\mu$ l TNT buffer

0.5  $\mu$ l SP6 or T7 RNA polymerase

0.5  $\mu$ l amino-acids mixture (1 mM, without methionine)

1  $\mu$ l  $^{35}$ S methionine (Amersham, *in vitro* translation grade, 1000 Ci/ mmol)

0.5  $\mu$ l RNasin (40 U/ $\mu$ l)

1.5  $\mu$ l (1  $\mu$ g) DNA

The reaction was incubated for 1 to 1.5 hours at 30 °C. The labeled mixture can be stored at - 70 °C.

## **2.29 GST pulldown assays**

Pulldown of nuclear proteins:

GST-fusion or GST protein (6  $\mu$ g) was first immobilised onto GSH-sepharose beads (Pharmacia). HeLa nuclear extract (60  $\mu$ g) was then added and allowed to mix with the beads for 4 hours at 4 °C on a rotating platform. Unbound proteins were removed by washing the beads for five times with wash buffer (20 mM HEPES.KOH pH 7.9, 150 mM NaCl, 10 % glycerol, 0.5 mM EDTA, 0.1 % Triton X-100, 1 mM DTT). The bound proteins were eluted by adding 12  $\mu$ l of 2 x SDS loading buffer and heated at 98 °C for 5 minutes. The supernatant was then loaded onto SDS protein gel for subsequent Western blot analysis.

Pulldown of *in vitro* translated  $^{35}$ S labeled proteins:

GST-fusion or GST protein (3  $\mu$ g) was immobilised onto GSH-sepharose beads (Pharmacia). The slurry was then incubated with 1  $\mu$ l of *in vitro* translated mixture in 100  $\mu$ l of pulldown buffer (50 mM HEPES.KOH pH 7.9, 250 mM NaCl, 0.1 % Triton X-100, 1 mM DTT, 0.5 mM EDTA, 10 % glycerol, 0.5  $\mu$ g/  $\mu$ l of BSA) for 2 hours at 4 °C on a rotating platform. After three washes with pulldown buffer, the bound proteins were eluted by adding 12  $\mu$ l of 2 x SDS loading buffer, fractionated by SDS-PAGE, and detected by autoradiography.

Pulldown of core histones:

GST-fusion (4 µg) or GST protein (6 µg) was immobilised onto GSH-sepharose beads and blocked in binding buffer (50 mM HEPES.KOH pH 7.9, 200 mM NaCl, 0.1 % Triton X-100, 2 mM DTT, 0.5 mM EDTA, 10 % glycerol, 0.5 µg/ µl of BSA, 0.5 mM PMSF). The beads were then incubated with 12 µg of calf thymus core histones (or 30 µg of lysozyme in control experiments) in a 120 µl volume for 3 hours at 4 °C on a rotating platform. The core histones (H2A, H2B, H3 and H4) were purchased from Boehringer. After four washes with binding buffer (the last wash did not contain BSA), the bound proteins were eluted by adding 12 µl of 2 x SDS loading buffer, fractionated on a 16.5 % SDS protein gel.

### **2.30 *In vitro* transcription assay**

Bacterially expressed GST-MBD2a or GST-MBD2b fusion proteins were incubated at room temperature for 30 mins in 19.5 µg of buffer containing 8.2 mM HEPES.KOH (pH 7.9), 41 mM KCl, 3 mM MgCl<sub>2</sub>, 0.08 mM EDTA, 0.2 mM DTT, and 8.2 % glycerol in the presence of 100 ng AdML promoter constructs (pAdomal either mock methylated or HhaI methylated, and pAdBXN as internal control). The amount of DNA was normalised to 1 µg by using pGEX3x as carrier. 8 units of HeLa nuclear extract (Promega) was added and incubated for 10 mins. Transcription was initiated by adding 2.5 µl of NTP mix (5 mM each of ATP, CTP, GTP, and UTP) and further incubated for 1 hr at 30 °C. The reactions were terminated by adding 100 µl of stop solution (20 mM EDTA pH 8.0, 0.2 M NaCl, 1 % SDS, and 0.25 µg/µl glycogen) and 300 µl of 0.3 M sodium acetate (pH 4.8). The samples were deproteinised by phenol/chloroform extraction and ethanol precipitated. The samples were dissolved in 5 µl of water. The amount of transcripts from the two promoter constructs were quantified by primer extension assay (Promega). The primer used was TAGCGCAGAAGTCATGCCCGCTTT. The <sup>32</sup>P end-labeled primer was allowed to anneal with the transcripts at 52-54 °C. Reverse transcription was then carried out using Superscript reverse transcriptase (Gibco, BRL). Transcription from the pAdomal construct gives a primer extension product of 87 bases in length, while transcription from the pAdBXN construct gives a primer extension product of 103 bases in length. Primer extension products were analysed on denaturing 7% polyacrylamide gel. Quantification of transcription products was determined by Phosphor Imager (Molecular Dynamics) using ImageQuant ver 3.3 software.

## **Chapter 3: Mechanism of transcriptional repression by methyl-CpG binding protein 2, MeCP2**

### **3.1 Introduction**

In addition to the methyl-CpG binding activity MeCP1 (Meehan et al, 1989), a second methyl-CpG binding activity (MeCP2) was also detected in many mammalian nuclear extracts (Lewis et al, 1992). This activity was purified, and the cDNA encoding the protein was subsequently cloned (Lewis et al, 1992). MeCP2 contains two important functional domains expected of a methylated DNA binding transcriptional repressor (Nan et al, 1993; Nan et al, 1997).

The methyl-CpG binding domain (MBD) of MeCP2 was mapped by methylated DNA binding assays using a panel of truncated recombinant MeCP2 proteins (Nan et al, 1993). A minimal region of 85 amino acids was shown to bind specifically to a single methyl-CpG pair *in vitro*. That the MBD of MeCP2 is also required for methyl-CpG binding *in vivo* was critically addressed by a later study to look at the *in vivo* localisation of MeCP2 (Nan et al, 1996). In mouse cells, ectopically expressed MeCP2-LacZ fusion protein was targeted to heterochromatin which is enriched with 5-methylcytosine. Deletion within the MBD region of MeCP2-LacZ fusion protein lead to inefficient targeting to heterochromatin. The localisation studies strongly suggests that the MBD of MeCP2 is required for preferential association with methyl-CpG *in vivo*.

The second functional domain is the transcriptional repression domain (TRD). Recruitment of the TRD to a reporter gene via a GAL4-DNA binding domain leads to transcriptional repression (Nan et al, 1997). The TRD is highly basic and not homologous to other known nuclear proteins. The contents in this chapter deal with the studies on the mechanism of repression by MeCP2.

### **3.2 Results and discussion**

#### **3.2.1 MeCP2 can bind to mSin3A and HDACs: mapping of co-repressor interacting region**

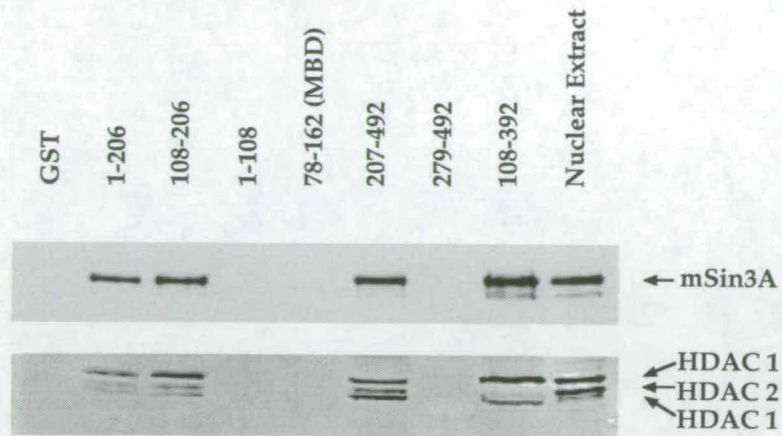
Previous work carried out in this laboratory defined a transcriptional repression domain in MeCP2 that can repress transcription of a reporter when tethered to the DNA via a GAL4-

DNA binding domain. As MeCP2 is a chromosomal protein, it is speculated that MeCP2 may make use of chromatin modifying repressor to repress transcription. An attractive candidate for MeCP2 interacting repressor is the mSin3A co-repressor complex required for a variety of DNA binding repressors (e.g. Yeast Ume6, Mad-Max heterodimer, unliganded nuclear hormone receptors). An *in vitro* GST-pulldown assay was used to test this hypothesis. GST and GST-MeCP2 fusion proteins were first coated onto GSH-sepharose beads. GST protein alone served as a negative control. The beads were then incubated with HeLa nuclear extract for more than 2 hours at 4°C to allow interaction between GST-fusion proteins with other nuclear proteins. After extensive washes of the GST fusion-sepharose beads, bound proteins were analysed by Western blotting using antibodies against mSin3A, HDAC1 and HDAC2. mSin3A, HDAC1 and HDAC2 were found to interact with GST-MeCP2 (Figure 3.2.1A). GST control did not bind to these co-repressor proteins. Using a panel of different GST-MeCP2 fragments, a co-repressor interacting region (CIR) was delineated. The co-repressor interacting region overlaps substantially with the TRD mapped using the GAL4-fusion *in vivo* repression assay (Figure 3.2.1B).

### **3.2.2 MeCP2 is associated with mSin3A and histone deacetylases**

The immunoprecipitation and histone deacetylase experiments described in this section were carried out by Dr Xinsheng Nan (University of Edinburgh) and Dr Colin Johnson (University of Birmingham), respectively. Rat brain nuclear extract which is enriched with MeCP2 was used for both assays. It was found that anti-MeCP2 antibodies (670 and 674) can both co-immunoprecipitate mSin3A protein, in addition to MeCP2 itself (Figure 3.2.2A). Preimmune antibodies did not immunoprecipitate mSin3A or MeCP2 proteins. Conversely, reciprocal immunoprecipitation experiment with different anti-mSin3 antibodies also revealed that MeCP2 and mSin3 proteins exist in a complex. In addition, it was found that anti-MeCP2 antibodies can also immunoprecipitate histone deacetylase activity (Figure 3.2.2B). The substrate for the histone deacetylase assay is a synthetic peptide corresponding to the 18 N-terminal residues of histone H4, chemically acetylated using <sup>3</sup>H-labeled acetic anhydride. Treatment of the immunoprecipitates with a histone deacetylase inhibitor Trichostatin (TSA), abolished the activity. The result indicates that MeCP2 is associated with a catalytically active histone deacetylase complex.

**A**



**B**

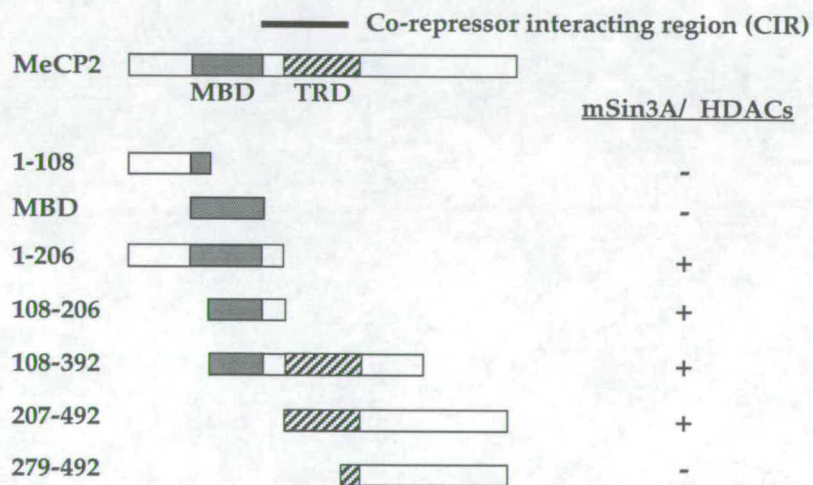


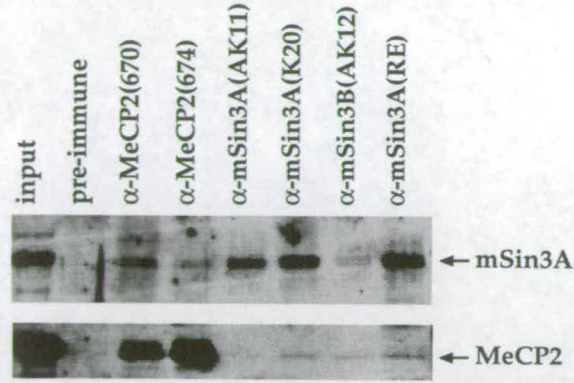
Figure 3.2.1. Interaction between mSin3A/HDAC1/HDAC2 and MeCP2

A) Western blot analysis of proteins that interact with different fragments of MeCP2. The numbers shown correspond to amino-acid position in the protein. The HDAC2 band in lane 108-392 migrates aberrantly due to co-migration of GST-fusion protein with HDAC2. mSin3A antibody was K20 (Santa Cruz), HDAC1 antibody was raised against a peptide corresponding to the C-terminal domain of HDAC1 (obtained from Dr. Bryan Turner), HDAC2 antibody was obtained from Dr. Ed Seto.

B) Summary of MeCP2 / mSin3A co-repressor complex interaction  
Black line corresponds to the co-repressor interacting region



A



B

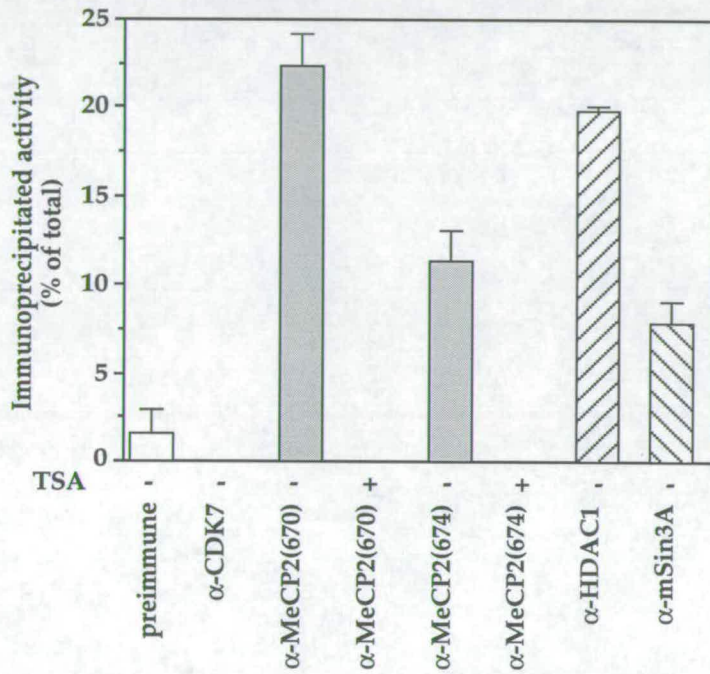


Figure 3.2.2. Co-immunoprecipitation of MeCP2, mSin3, and histone deacetylase activity from rat brain nuclear extracts

A) Western blot proteins immunoprecipitated by anti-MeCP2, and anti-mSin3 antibodies.

The antibodies used for immunoprecipitation were indicated in the brackets. AK11, K20, AK12 were obtained from Santa Cruz. RE was obtained from Dr. Robert Eisenman. The blot was probed with mSin3A antibody (RE) and anti-MeCP2 (670 and 674).

B) Immunoprecipitation of histone deacetylase activity with MeCP2, HDAC1, and mSin3A antibodies. Preimmune and CDK7 antibodies were controls. + indicates pre-incubation of the immunoprecipitates in the presence of 5 ng/ml TSA.

### 3.2.3 MeCP2 can bind to *in vitro* synthesised mSin3A but not HDACs

The mSin3A corepressor complex contains at least 7 subunits: mSin3A, HDAC1, HDAC2, RbAp46, RbAp48, SAP30 and SAP18 (Zhang et al, 1997; 1998). From *in vitro* protein-protein interaction studies, the largest known subunit mSin3A appears to serve as a scaffold protein which other subunits interact (Zhang et al, 1998; Laherty et al, 1998). The next question to ask is which subunit does MeCP2 interact with. To address this question, mSin3A, HDAC1 and HDAC2 were *in vitro* synthesised and labeled with <sup>35</sup>S-methionine using rabbit reticulocyte lysates. The labeled proteins were then tested for the affinity with GST-CIR (amino acids 108-392 of rat MeCP2). GST-CIR interacts strongly with mSin3A, as more than 15 % of the mSin3A input material was found to be retained by GST-MeCP2 (Figure 3.2.3). The interaction with HDAC1 or HDAC2 was weak but significant. The interaction between GST-MeCP2 and mSin3A is specific because GST control did not bind to mSin3A under the same experimental condition.

### 3.2.4 Mapping of MeCP2 interacting regions within mSin3A

Further work was carried out to map the region in mSin3A that can interact with GST-CIR *in vitro*. A panel of truncated mSin3A protein was synthesised using rabbit reticulocyte lysate. GST-pulldown assays revealed that two regions in mSin3A appears to mediate binding to MeCP2 (Figure 3.2.4). The first region is the N-terminal half of the histone deacetylase interaction domain (HID, Laherty et al, 1997). The binding of mSin3A to GST-CIR was drastically reduced upon removal of C-terminal 200 amino acids of mSin3A, it is inferred that the C-terminal of mSin3A contains a second MeCP2 interacting site. Furthermore, removal of the HID region did not destroy the interaction between MeCP2 and mSin3A. The interaction between CIR of MeCP2 and HID of mSin3A suggests MeCP2 may block the interaction between mSin3A and histone deacetylase. This possibility is unlikely based on the result of co-immunoprecipitation experiments which indicate that MeCP2 is associated with both mSin3A and histone deacetylases.

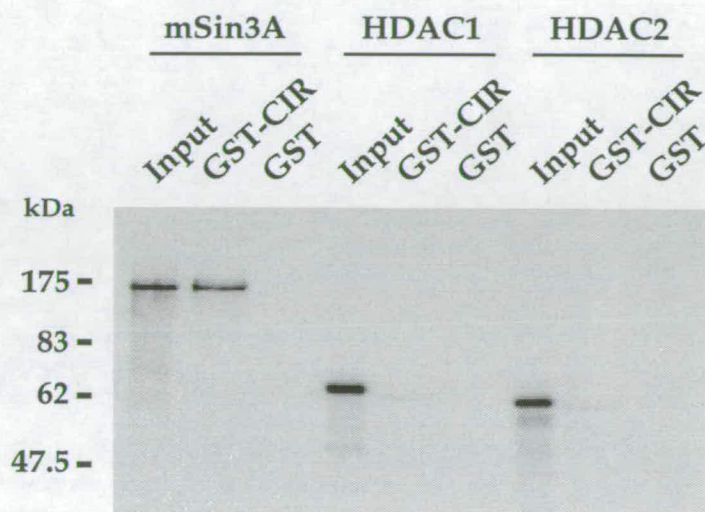


Figure 3.2.3. CIR interacts strongly with mSin3A, but weakly with HDAC1 and HDAC2.

mSin3A, HDAC1, and HDAC2 were *in vitro* translated and labelled with  $^{35}\text{S}$  methionine. The labeled proteins were incubated with either GST or GST-CIR. Input is 20 % of total.



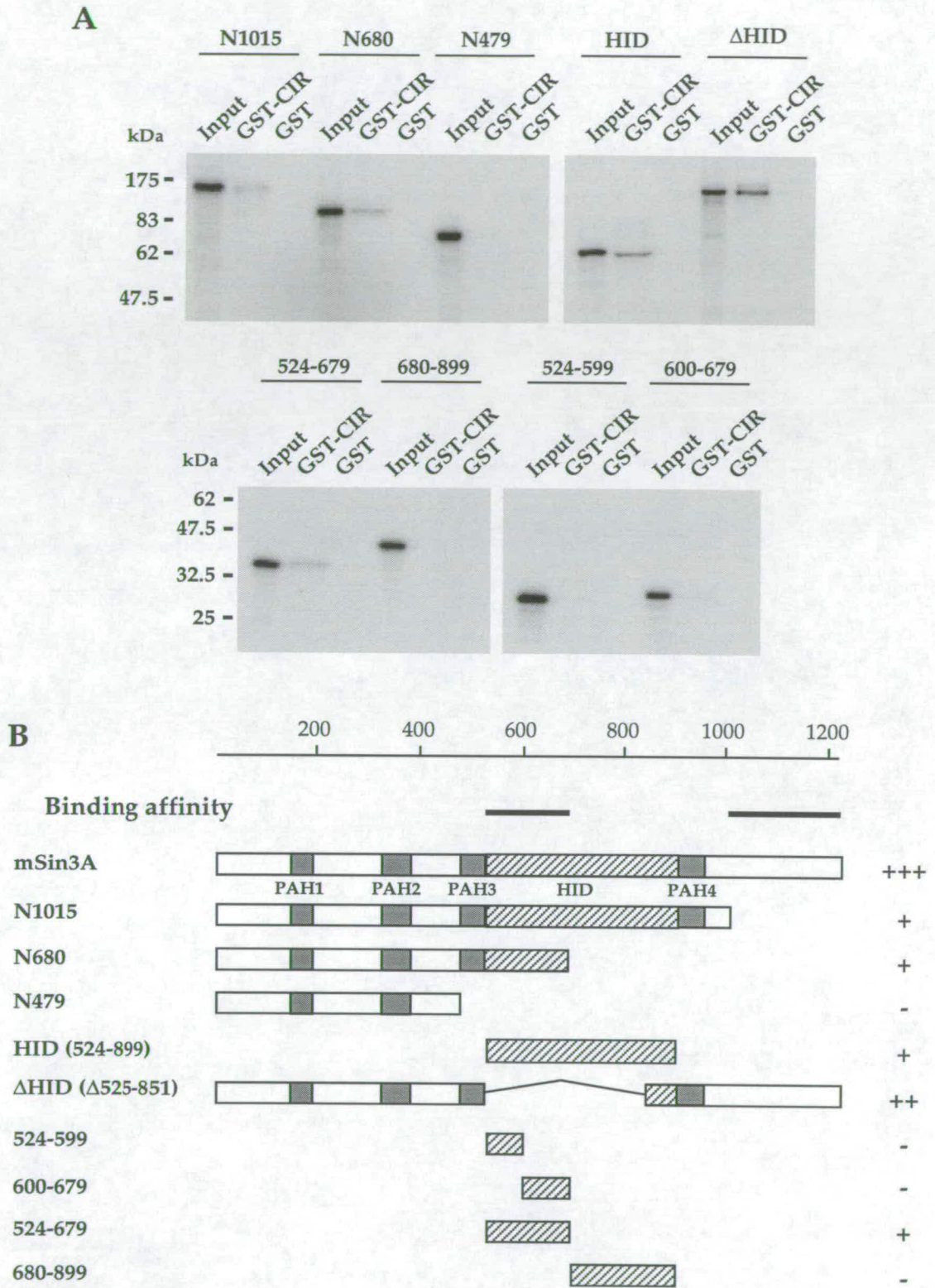


Figure 3.2.4. Interaction between CIR and mSin3A.

A) Different fragments of mSin3A were synthesised and labeled with  $^{35}\text{S}$  methionine. The labeled proteins were incubated with either GST or GST-CIR. Input is 10 % of total.

B) Maps of different mSin3A deletion constructs. The thick black lines indicates the regions of mSin3A that bind strongly to CIR.

### **3.2.5 Ternary complex between MeCP2/ mSin3A/ HDAC2 cannot be assembled *in vitro***

As the CIR of MeCP2 interacts *in vitro* with previously identified HID (Laherty et al, 1997), this raises the question of whether mSin3A can bind simultaneously to MeCP2 and HDAC1 or HDAC2. mSin3A, HDAC2, luciferase or MeCP2 were *in vitro* co-translated and labeled with <sup>35</sup>S-methionine. mSin3A was then immunoprecipitated using anti-myc 9e10 monoclonal antibody that recognises five myc epitopes that the N-terminal of mSin3A. Binding of *in vitro* translated proteins to anti-myc antibody/protein G sepharose was dependent on the presence of myc epitopes as HDAC2 and MeCP2 did not bind to the antibody-protein G sepharose beads (Figure 3.2.5A). It is found that HDAC2 can co-immunoprecipitate with mSin3A. The interaction between mSin3A and HDAC2 is specific because the luciferase in the same reaction did not co-immunoprecipitate with mSin3A and HDAC2 (Figure 3.2.5B). MeCP2 was absent in the immunoprecipitates. Co-translation with increased amount of MeCP2 construct to yield more MeCP2 did not result in co-immunoprecipitation of MeCP2 with mSin3A/ HDAC2. Therefore, no ternary complex between MeCP2/ mSin3A/ HDAC2 can be formed *in vitro*. It is plausible that a co-factor for mediating the ternary interaction was missing in the reaction, or *in vitro* synthesised MeCP2 was not folded properly.

### **3.2.6 Repression by GAL4-MeCP2 TRD is sensitive to histone deacetylase inhibitor, TSA**

Biochemical evidences described in section 3.2.1 to 3.2.3 established an association between MeCP2 and mSin3A/ histone deacetylases complex. An important question is whether repression by TRD is dependent on histone deacetylation. Experiments were carried out to test whether repression by TRD is sensitive to a specific inhibitor of histone deacetylases, trichostatin A (TSA). The relative reporter activity (G5/G0) was used as a measure of the specific transcriptional repression by GAL4-TRD on the reporter construct with Gal4 binding sites. Each G5/G0 value is the ratio of  $\beta$ -galactosidase activity from separate transfection of cells with either G5 reporter (containing 5 upstream Gal4 binding sites) or G0 reporter (with no Gal4 binding site). The relative reporter activity in the absence of Gal4-MeCP2 remains close to 1 as the Gal4-binding sites do not affect the  $\beta$ -actin promoter activity. The presence of GAL4-TRD reduced the relative reporter activity to 0.2. However, when the cells were treated with TSA for 24 hours prior to harvest, the

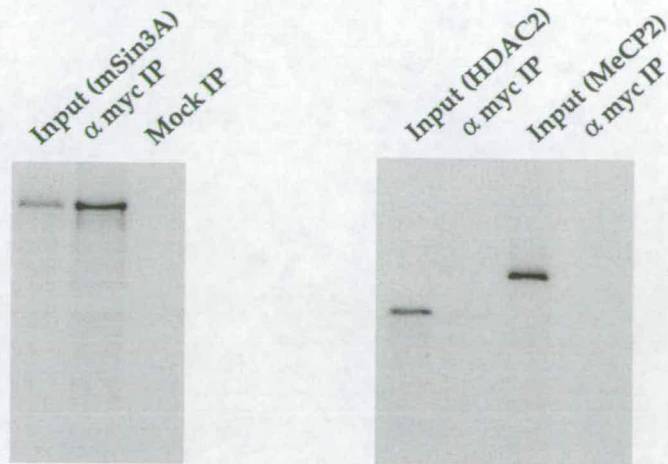
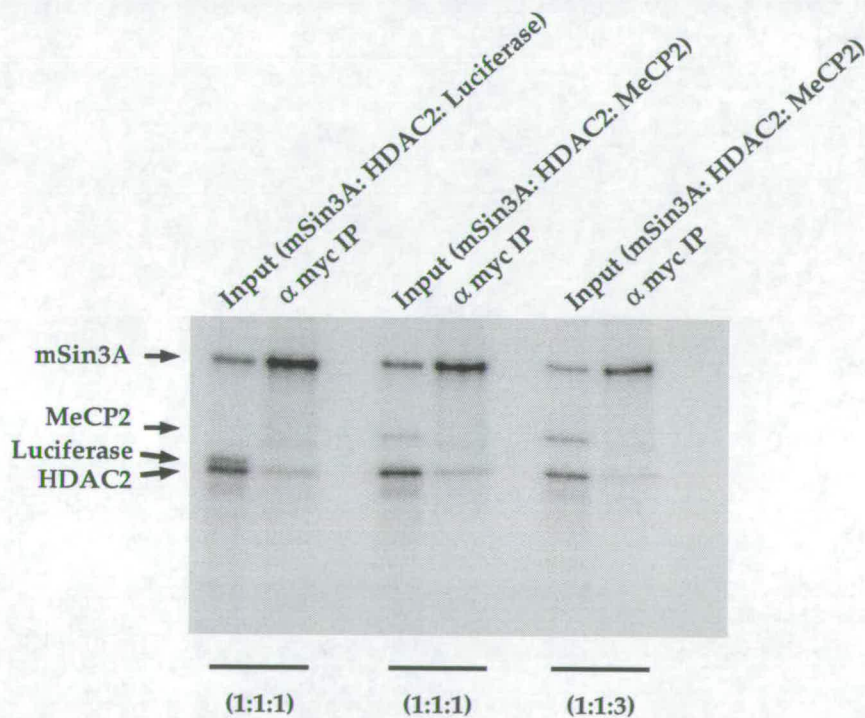
**A****B**

Figure 3.2.5. Failure to reconstitute a ternary mSin3A/ HDAC2 / MeCP2 complex *in vitro*

A) Anti-myc antibody immunoprecipitates only myc-tagged mSin3A, and not untagged HDAC2 and MeCP2. Mock immunoprecipitation does not contain antibody.

B) Immunoprecipitation of *in vitro* co-translated mSin3A/Luciferase/HDAC2 or mSin3A/HDAC2/ MeCP2. 1:1:1 indicates the ratio of plasmids used for *in vitro* translating the three different proteins. In 1:1:3 mixture, the amount of MeCP2 construct was three times more than that for mSin3A and HDAC2 constructs. Input is 16 % of total.

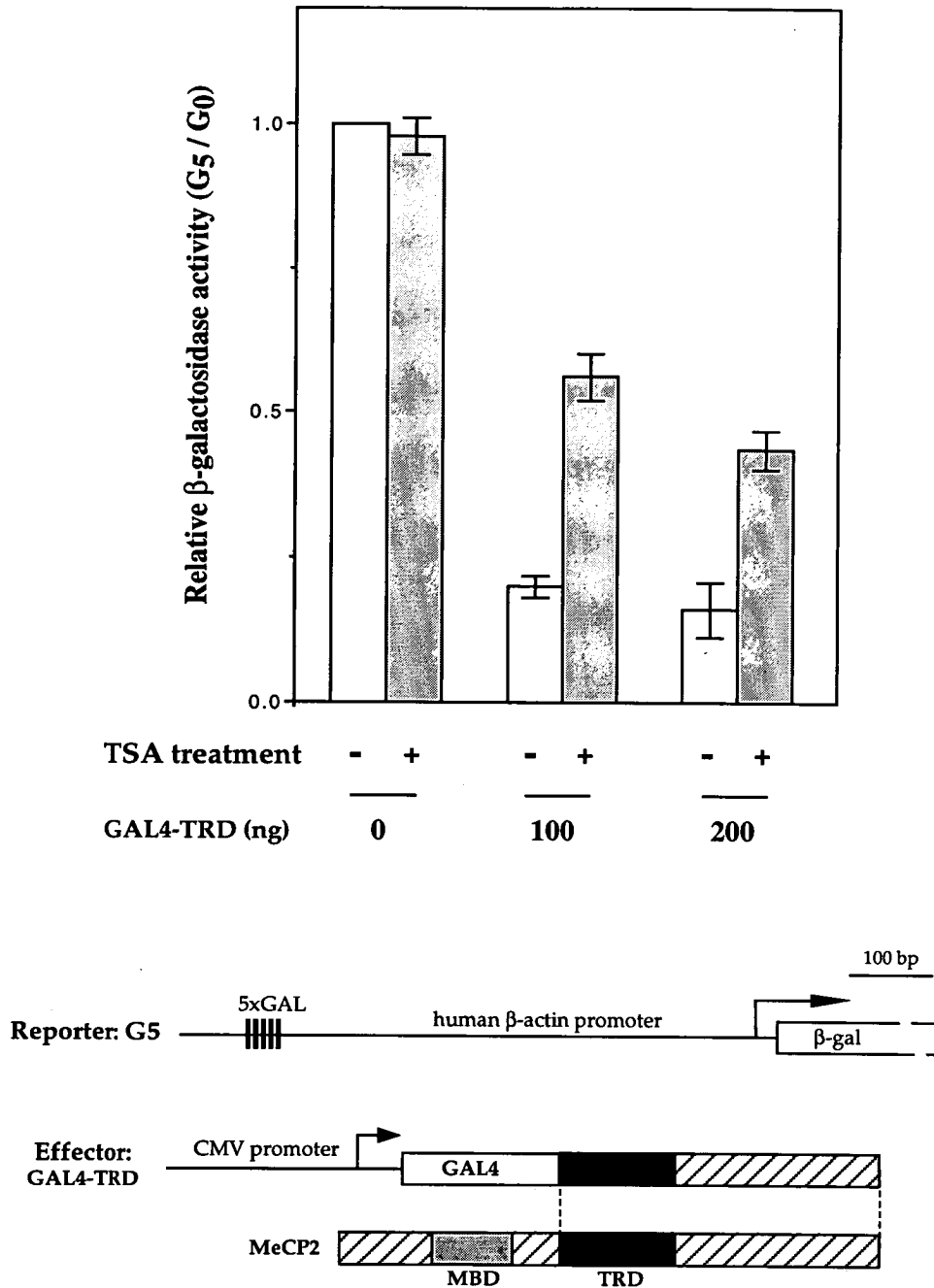


Figure 3.2.6. Repression by TRD is sensitive to histone deacetylase inhibitor, TSA

Mouse L cells were transiently transfected with GAL4-TRD (207-492 amino-acids of MeCP2) and reporter with or without 5 Gal4 binding sites (G5 or G0). Relative  $\beta$ -gal activity was used to determine the specific repression by TRD. Transfected cells were treated with 100 ng/ml TSA for 24 hours before harvest. The result shown are based on three independent experiments. The G5 reporter and GAL4-TRD effector constructs used for the experiment were shown.

relative reporter activity was increased by 2 to 3 fold (Figure 3.2.6). Similar TSA treatment did not affect the relative reporter activity in the absence of GAL4-TRD. It is concluded that GAL4-TRD can repress transcription in a deacetylation dependent manner.

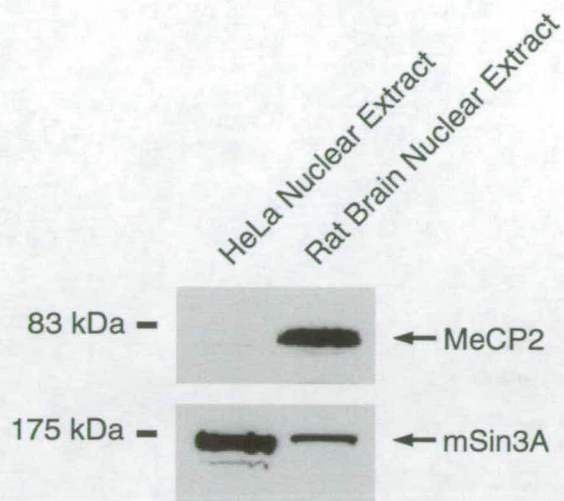
### **3.2.7 HeLa cells is deficient in MeCP2, but can repress methylated reporter construct in a deacetylation dependent manner**

Repression by DNA methylation is well-studied for HeLa cells (Boyes and Bird, 1991). Transfection of methylated genes into this cell-line result in transcriptional silencing, and MeCP1 is implicated as a mediator of methylation dependent repression. However, the finding that MeCP2 is also a transcriptional repressor with an affinity for methyl-CpGs evokes the possibility that MeCP2 may also be responsible for repression of methylated DNA in HeLa cells. A western blot of HeLa nuclear extract with anti-MeCP2 antibody (R670) failed to detect MeCP2 (Figure 3.2.7A). Control experiment with anti-mSin3A antibody showed that mSin3A proteins were present in both HeLa and rat brain nuclear extracts. As the anti-MeCP2 antibodies was raised against GST-rat MeCP2, it is possible that the antibodies did not recognise human MeCP2 due to sequence variation between species. The same anti-MeCP2 antibody was tested on human liver and kidney extracts, and it detected proteins of 80 kDa, which is likely to be MeCP2 (Figure 3.2.7B). The blot was also probed with anti-MBD1 (S751) antibody as a control for loading of extracts. The R670 anti-MeCP2 antibody was raised against rat MeCP2 spanning amino-acids 207-492, and the similar region in human MeCP2 shares a 93% identity over 286 amino-acids compared to the rat MeCP2. The result indicates that MeCP2 is below the level of detection, and is either absent or not abundant in HeLa cells. Therefore, MeCP2 is unlikely to mediate repression of methylated genes in this cell-line, and other methyl-CpG binding proteins are implicated.

The next question to ask is whether the repression of methylated reporter in the absence of MeCP2 is dependent on histone deacetylation. It has previously been shown that repression of methylated reporters in HeLa cells is dependent on DNA methylation density, and the repression is very likely to be mediated by another methyl-CpG binding activity, MeCP1 (Boyes and Bird, 1991; 1992). Therefore a reporter construct, pGL2-CG11 (Figure 3.2.8B, constructed by Xinsheng Nan) was used to address this question. CG11 is a 150 bp CpG rich probe used for detecting MeCP1 bandshift activity (Meehan et al, 1989). The 6 kb parent plasmid pGL2 consists of a SV40 promoter driving a luciferase reporter gene, a downstream SV40 enhancer, and a bacterial plasmid backbone



**A**



**B**



Figure 3.2.7. HeLa cells are deficient in MeCP2

A) Western blot analysis of nuclear extracts from HeLa or rat brain nuclei. The blot were probed with anti-MeCP2 (R670) or anti-mSin3A (K20) antibodies.

B) Anti-MeCP2 (R670) antibody can detect MeCP2 in human liver and kidney extracts. The blot was also probed with anti-MBD1 (S751) antibody.

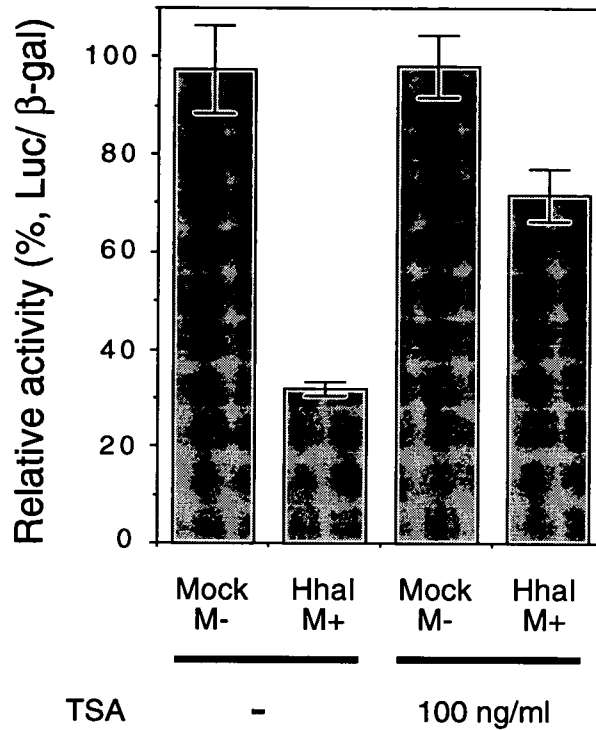
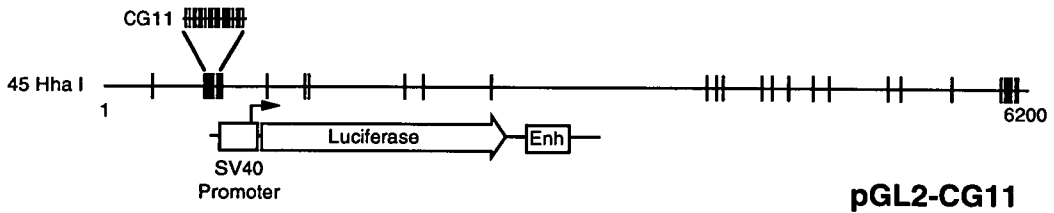
**A****B**

Figure 3.2.8. HeLa cells can repress methylated reporter in a deacetylation dependent manner.

A) HeLa cells were transfected with either mock-methylated (M-) or *HhaI*-methylated (M+) pGL2-CG11 luciferase reporter. Transfected cells were either treated with 100 ng/ml TSA for 24 hours or left untreated. Relative activity of the luciferase activity was normalised with β-gal activity from a co-transfected β-gal reporter driven by SV40 promoter. The result shown are based on three independent experiments.

B) Map of pGL2-CG11 construct. The positions of 45 *HhaI* sites are shown.

with an ampicillin resistance gene. In pGL2-CG11 construct, a CG11 cluster is located upstream of the SV40 promoter. The CG11 which is methylatable at 27 methyl-CpGs, is designed to “attract” endogenous MeCP1 to the vicinity of the promoter. HpaII methylated pGL2 construct (at 25 “CCGG” sites) was repressed to a greater extent than HhaI methylated pGL2 construct (at 25 “GCGC” sites) in mouse L fibroblasts (Xinsheng Nan, personal communication). However, the difference between mock-methylated pGL2 construct and HhaI methylated pGL2 construct was minimal (Xinsheng Nan, personal communication). The result indicates that HpaII methylation caused direct inhibition of transcription, and it is very probable that the binding of certain transcription factors to their cognate recognition sites were affected by HpaII methylation. It is also suspected that methylation of the pGL2 construct at all CpGs (including HpaII and HhaI sites) by SssI methyltransferase will also bring about direct effect of repression. In order to study the genuine effect of “tethered” MeCPs on transcription of a promoter, I decided on using HhaI methylated pGL2-CG11 construct. By methylating the construct with HhaI methyltransferase, a densely methylated region of 20 methyl-CpGs over 150 bp was created. The rest of the 6 kb plasmid is only methylatable at 25 methyl-CpG sites.

The HhaI methylated pGL2-CG11 luciferase reporter was repressed to 30 % of the mock methylated reporter in HeLa cells (Figure 3.2.8A). These values were obtained after correcting for transfection efficiency by co-transfection of a non-methylated pSV40-LacZ reporter (SV40 promoter driving LacZ reporter gene). TSA treatment of transfected HeLa cells restored the relative activity (Luciferase/  $\beta$ -galactosidase activity) of HhaI methylated pGL2-CG reporter to approximately 70 % of the mock methylated reporter. The result indicates that DNA methylation dependent repression in HeLa cells is dependent on deacetylation. Therefore, TSA sensitive repression of methylated reporter can occur in the absence of the known methyl-CpG binding repressor MeCP2, and MeCP1 is implicated for this effect.

### **3.3 Conclusion**

MeCP2 has previously been shown to repress transcription *in vitro* and *in vivo*, and repression activity has been mapped to a region central in MeCP2 (Nan et al, 1997; Kudo, 1998). Using GST-pulldown assays, a region in MeCP2 that was found to interact with components of the mSin3A co-repressor complex such as mSin3A, HDAC1 and HDAC2. Importantly, the corepressor interacting region (CIR) outlined by the *in vitro* pulldown assay overlaps substantially with the previously identified transcriptional

repression domain delineated using an *in vivo* repression assay. The interaction between MeCP2, mSin3A and histone deacetylases was further confirmed by reciprocal immunoprecipitation using anti-MeCP2 and anti-mSin antibodies. Anti-MeCP2 can immunoprecipitate both mSin3A and histone deacetylase activity from rat brain nuclear extract, and anti-mSin3A antibodies can immunoprecipitate MeCP2. GST-CIR interacts preferentially with *in vitro* synthesised mSin3A, and poorly with *in vitro* synthesised HDAC1 and HDAC2. The result suggests that MeCP2 binds directly to mSin3A which in turn interacts with HDAC1 and HDAC2 (Laherty et al, 1997; Zhang et al, 1998). GST-CIR binds to two regions within mSin3A. Attempt to assemble a ternary complex between MeCP2, mSin3A and HDAC2 was unsuccessful, suggesting that additional co-factor may be required for reconstitution of the MeCP2 corepressor complex. Most importantly, repression by GAL4-TRD can be partially relieved by histone deacetylase inhibitor Trichostatin A, indicating that deacetylation contributes to *in vivo* repression by TRD of MeCP2.

The interaction between MeCP2 and the Sin3A histone deacetylase complex was independently verified by Wolffe and colleagues (Jones et al, 1998). *Xenopus* MeCP2 was found to co-purify with the *Xenopus* Sin3A, HDAC1 and HDAC2. Co-immunoprecipitation experiments confirmed that they exist in a complex. Importantly, the transcriptional repression domain of *Xenopus* MeCP2 can repress transcription in a deacetylation dependent manner as revealed by TSA relief experiment.

The finding that MeCP2 is associated with the chromatin modifying complex, Sin3/HDAC co-repressor complex is intriguing, as the interaction can potentially provide the molecular mechanism for the long-standing relationships between DNA methylation, transcriptional repression and histone deacetylation. Collectively, the data suggest a model in which MeCP2 interacts with the mSin3A histone deacetylase complex, deacetylates core histone tails and leads to transcriptional repression (Figure 3.2.9).

Is MeCP2 the only player in mediating transcriptional repression by DNA methylation? In certain cell-lines such as HeLa cells, MeCP2 is below the limit of detection by Western blot analysis. Methylated reporters can however be efficiently repressed when transfected into HeLa cells (Boyes and Bird, 1991). In addition, repression of methylated reporter is sensitive to TSA, indicating that deacetylation dependent repression of methylated genes can occur in the apparent absence of MeCP2. Interestingly, HeLa cells contain methyl-CpG binding proteins such as MBD1 and MBD2, which are attractive candidates for

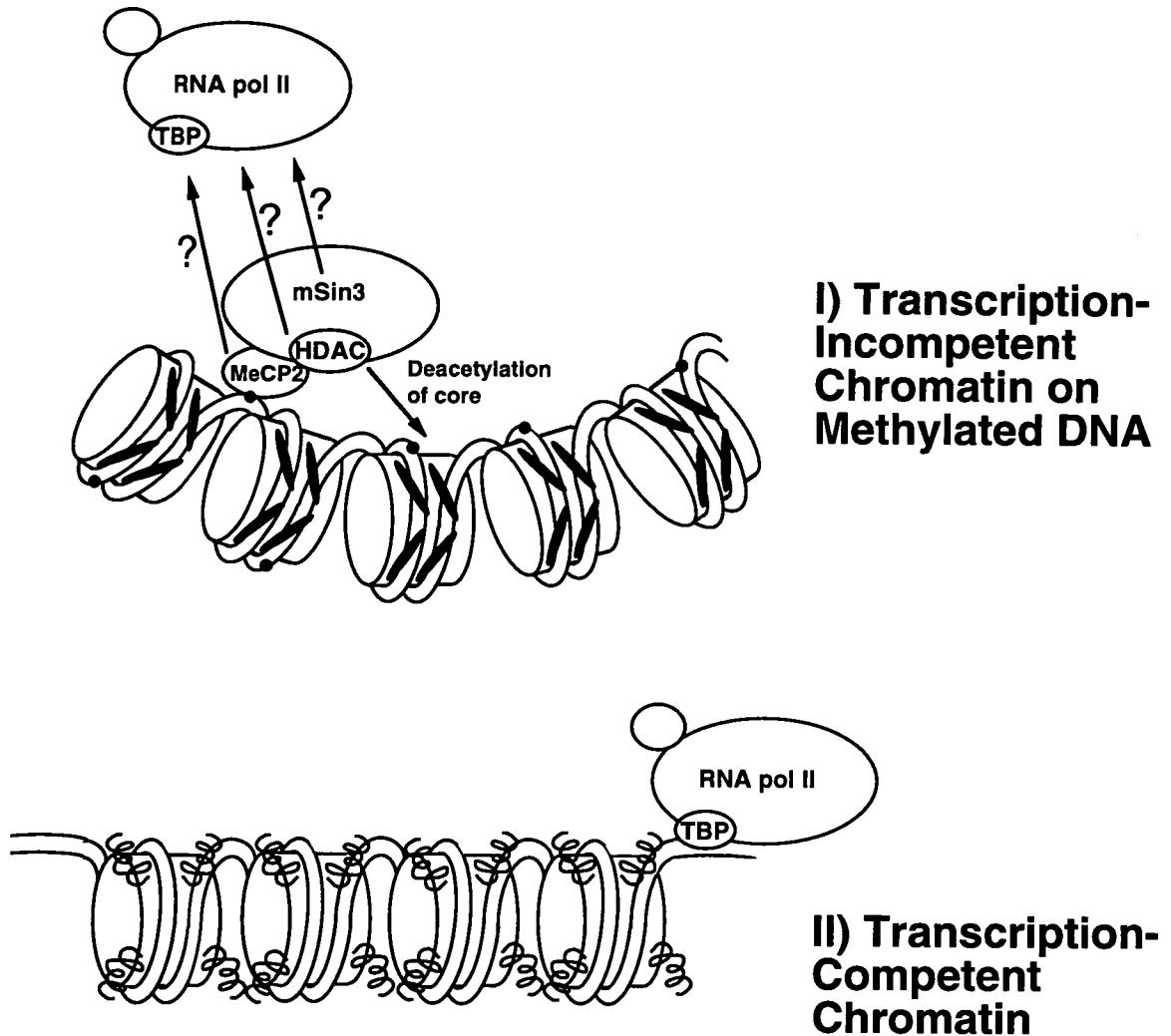


Figure 3.2.9. A model for transcriptional repression by MeCP2.

I) The mSin3A/ HDAC co-repressor complex is targeted to methylated DNA by MeCP2. It is proposed that this complex deacetylates core histone tails at the vicinity of chromatin bound MeCP2. Question marks represent unknown interaction between the MeCP2/ mSin3A/ HDAC complex with the basal transcription machinery. Deacetylation of core histone tails presumably lead to transcriptional silencing.

II) Chromatin associated with non-methylated DNA is however associated with hyperacetylated histones and transcriptionally active.

mediating repression. MBD2 is the subject of the work described in Chapter 4, and functional studies of MBD1 are presented in Chapter 5.

## **Chapter 4: MBD2 is a component of the transcriptional repressor MeCP1**

### **4.1 Introduction**

The work described in Chapter 3 indicates that MeCP2 can repress transcription via an interaction with the mSin3/ histone deacetylases complex. Therefore, the transcriptional repression mechanism by MeCP2 seems sufficient for the explanation of the longstanding relationship between DNA methylation and chromatin modification (Bestor, 1998; Razin, 1998). However, certain cell-types such as HeLa cells are deficient in the known repressor MeCP2 (see Chapter 3), and can still silence methylated genes (Boyes and Bird, 1991). It is clear that other methyl-CpG binding proteins are responsible for mediating repression of methylated sequences in the absence of MeCP2.

The purification of MeCP1 activity has been met with severe technical difficulties. With the rationale that component(s) of MeCP1 may have a methyl-CpG binding region homologous to the MBD of MeCP2, the EST database was searched to identify proteins with similar domain. The database search resulted in the identification of four novel mammalian proteins (MBD1, MBD2, MBD3, and MBD4) with MBD-like region (Cross et al, 1997; Hendrich and Bird, 1998). Interestingly, MBD1 contains two cysteines-rich regions which are also found in the DNA cytosine methyltransferase 1 (DNMT1) and the mammalian trithorax-like protein ALL1 (Cross et al, 1997). Studies on the characterisation of MBD1 are presented in Chapter 5. MBD2 is highly similar to MBD3 at the C-terminal region, and possesses a basic N-terminal extension with a run of glycine-arginine repeats. MBD4 has no significant homology to MBD1, MBD2 and MBD3 apart from the MBD-like region. Interestingly, MBD4 has a region similar to the glycosylase domains of various DNA repair proteins. Functional studies of MBD4 uncover a novel G/T mismatch glycosylase activity within the glycosylase domain of the protein (Hendrich et al, 1999). Thus, MBD4 is unlikely to be the mediator of transcriptional repression by DNA methylation. MBD1, MBD2 and MBD4 can bind specifically to methylated DNA *in vitro* (Cross et al, 1997; Hendrich and Bird, 1998). In mouse cells, ectopically expressed MBD1, MBD2 and MBD4 were targeted to DAPI-bright spots which are enriched in heavily methylated major satellite DNA. Localisation of MBD2 and MBD4 to these DAPI-bright spots were disrupted in mouse cells with low level of 5-methyl cytosines. These studies demonstrated that MBD2 and MBD4 can bind to methylated DNA *in vivo*. MBD3

failed to bind specifically to methylated DNA *in vitro*, and did not localise to DAPI-bright foci *in vivo*.

These other methyl-CpG binding proteins (MBD1 and MBD2) are therefore attractive candidates for MeCP1. This chapter concerns the characterisation of MBD2.

## **4.2 Results and discussion**

### **4.2.1 Characterisation of antibodies raised against MBD2**

In order to facilitate the biochemical studies of MBD2, three antibodies were raised against different MBD2 fragments.

#### **4.2.1.1 S752 antibody**

S752 antiserum (S specifies the host animal which is a sheep) was raised against an internal fragment of human MBD2 (corresponding to amino acids 232 to 394 of murine MBD2a). The immune antiserum detected a doublet band of approximately 47.5 kDa and a much weaker 28 kDa band of same mobility as MBD2b, a shorter version of full-length MBD2 (Figure 4.2.1.1 A and C). Preimmune serum did not light up these bands. Attempts to express full-length MBD2 as a His-tagged protein were unsuccessful, therefore MBD2b was used as a size marker as well as a positive control for Western analysis.

An alternative assay was used to test the specificity of this antibody. The immune antiserum was incubated with either GST-sepharose or GST-MBD2(antigen)-sepharose. This step is to deplete the antiserum of anti-GST and anti-MBD2 antibodies. The treated antisera were then used for Western probing. Pre-incubation with GST-MBD2(antigen)-sepharose removed both the 47.5 kDa and 28 kDa bands, while incubation with GST-sepharose alone had no effect (Figure 4.2.1.1 B). This result indicates that the 47.5 kDa and 28 kDa proteins were bound by antibodies against MBD2 antigen.

The MBD2 cDNA can potentially encode 2 ORFs starting from 2 different ATG sites (Figure 4.2.1.1 C). The shorter ORF, MBD2b was originally deduced from various cDNAs found in EST databases. Therefore it is not certain if MBD2a or MBD2b is the correct product of MBD2 gene. *In vitro* transcription and translation of the MBD2 cDNA



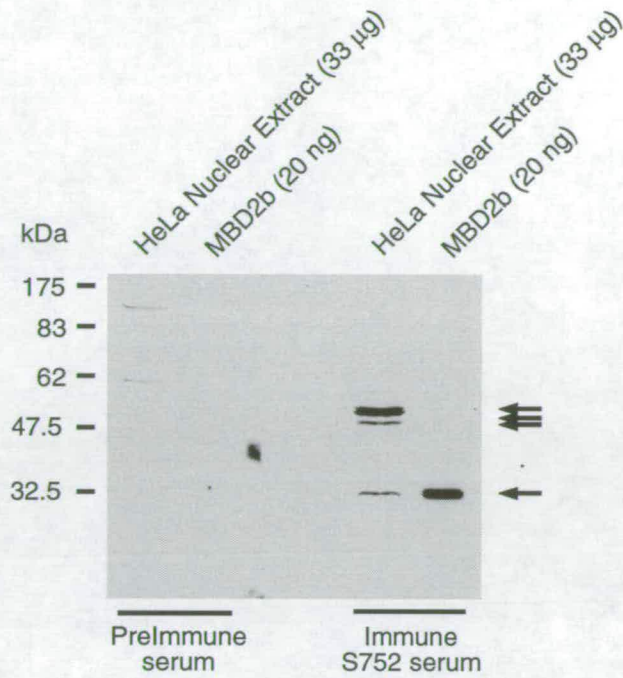
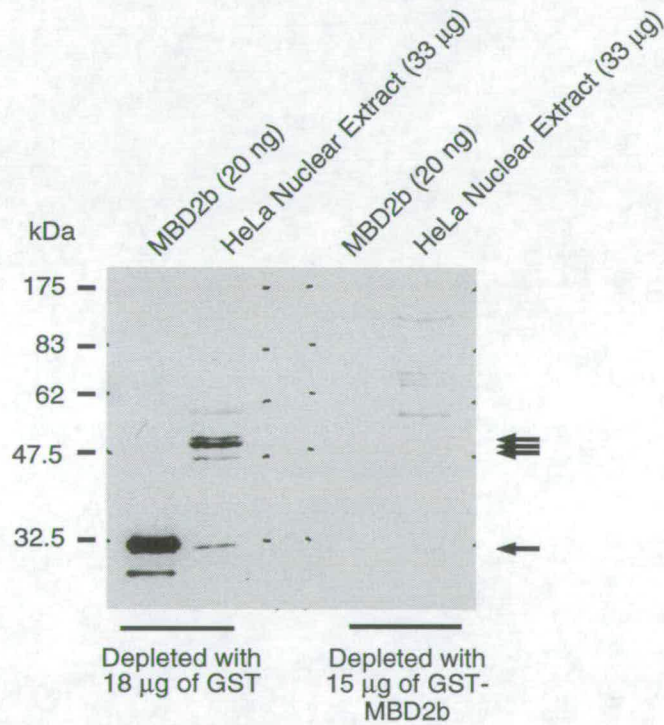
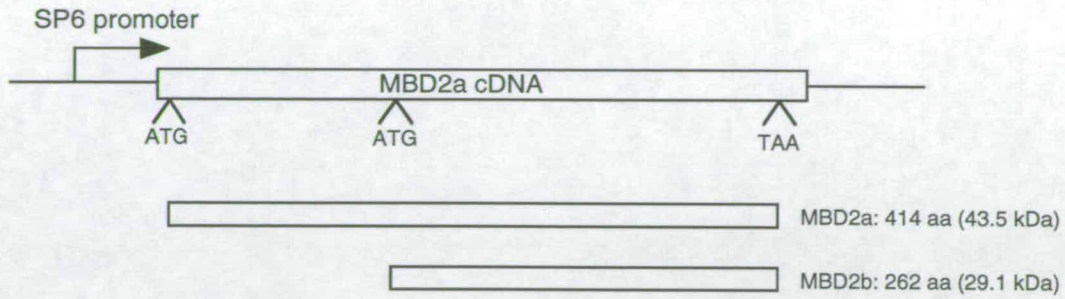
**A****B**

Figure 4.2.1.1. Characterisation of anti-MBD2 S752 antibody

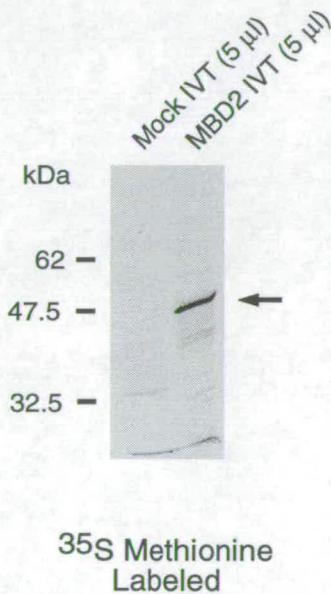
A) S752 antiserum detected proteins of 47.5 kDa and 28 kDa (indicated by arrows) on a Western blot of HeLa nuclear extract. Preimmune serum did not cross-react with similar bands on a Western blot.

B) The 47.5 kDa and 28 kDa proteins are MBD2-specific. Left, the S752 antiserum was preincubated with GST-sepharose before Western blot analysis. Right, the S752 antiserum was depleted of anti-GST and anti-MBD2 antibodies by pre-incubation with GST-MBD2 antigen. The depleted signals are indicated by arrows.

C



D



E

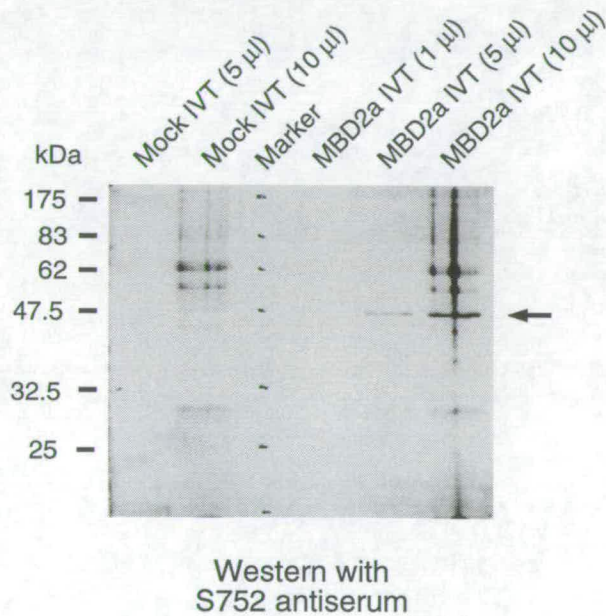


Figure 4.2.1.1. Characterisation of anti-MBD2 S752 antibody

C) Construct used for *in vitro* transcription and translation of MBD2.

The MBD2 cDNA can potentially generate MBD2a and MBD2b proteins through two ATG start sites.

D) A 47 kDa protein (indicated by arrow) was generated from the construct shown in C.

The MBD2 cDNA was transcribed using SP6 RNA polymerase, and followed by translation in the presence of <sup>35</sup>S-methionine to label the protein. Both mock and MBD2 translation mixtures were resolved by SDS-PAGE and the labeled proteins were detected by autoradiography. IVT is *in vitro* transcription/translation mixture.

E) The 47 kDa protein (indicated by arrow) reacted with anti-MBD2 antibody. Both mock and MBD2 translation mixtures were probed with anti-MBD2 (S752) antiserum. The 47 kDa protein was the only product derived from full-length MBD2 cDNA, and the 28 kDa (MBD2b version) protein was not detected by *in vitro* translation of MBD2 RNA.

gave a major <sup>35</sup>S labeled product of 47.5 kDa (Figure 4.2.1.1 D). The S752 antibodies also detected a protein of 47.5 kDa in the same *in vitro* translated mixture (Figure 4.2.1.1 E). This protein was not found in a mock *in vitro* translated mixture. It is concluded that the major protein product that can be produced from the MBD2 cDNA is MBD2a. Furthermore, the *in vitro* translated MBD2a is of the same mobility as the 47.5 kDa band on Western blot of HeLa nuclear extract, indicating that the anti-MBD2 antibody can detect MBD2-specific protein.

The S752 antibodies were affinity purified using a His-tagged MBD2b - AffiGel 10 column. The affinity purified antibodies detected the 47.5 kDa doublet bands as well as the 28 kDa band on western blot of HeLa nuclear extract (Figure 4.2.1.1 F).

The specificity of the antiserum was also tested for its ability to differentiate between recombinant MBD2b, MBD1, and MBD4. The antiserum did not recognise MBD1 and MBD4 (Figure 4.2.1.1 G). Based on these series of experiments, it is concluded that the 47.5 kDa protein is MBD2a. It is not certain whether the 28 kDa band (denoted as MBD2b?) is a protein product (MBD2b version) of the MBD2 gene, a degraded product of MBD2a or a MBD2b related protein.

The S752 antibody was also tested for its ability to recognise native MBD2 proteins in an extract by immunodepletion assay. Preimmune or immune antibodies were immobilised on protein G sepharose beads. The coated beads were then incubated with HeLa nuclear extracts. Western blot analysis of the unbound extracts revealed that the antibodies did not deplete MBD2 (Figure 4.2.1.1 H). It was not possible to analyse the immunoprecipitate for MBD2 because the co-eluted IgG migrates with the same mobility as MBD2a on the protein gel. It is inferred that this antibody does not recognise native MBD2.

The effect of anti-MBD2 antibody on MeCP1 bandshift activity was also tested in an attempt to address whether MBD2 is a component of MeCP1. The methylated CG11 probe was used for this experiment. The CG11 fragment can be released from the parent plasmid pCG11 by cutting with restriction enzymes Asp718 and BamHI. The 5' overhangs of both ends allow radiolabeling of the probe by end-filling with Klenow. 27 symmetrically methylated CpG can be created by methylating the CG11 fragment with Sss1 CpG cytosine methyl-transferase. The antibodies obtained from different bleeds did not affect the mobility of MeCP1 bandshift activity (Figure 4.2.1.1 I).



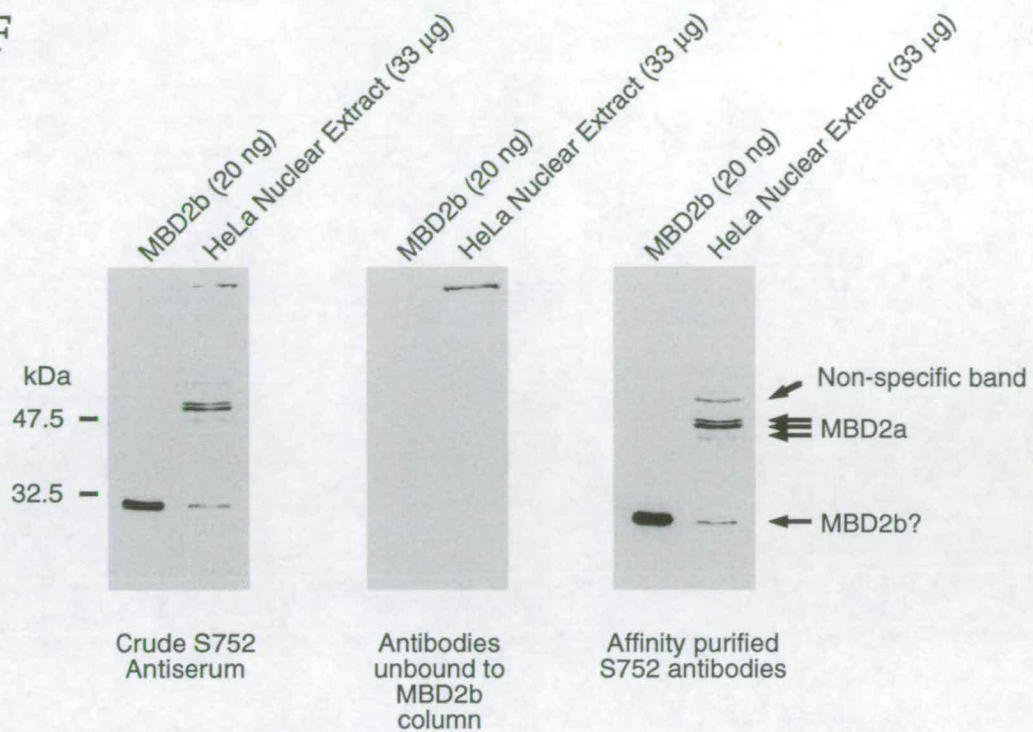
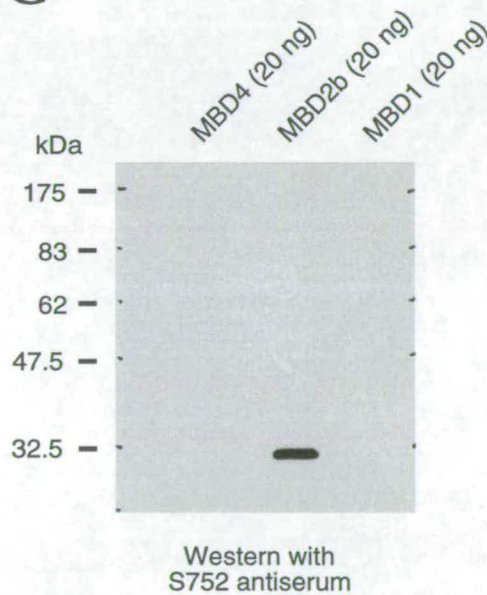
**F****G**

Figure 4.2.1.1. Characterisation of anti-MBD2 S752 antibody

F) Affinity purified anti-MBD2 antibodies detected both the 47.5 kDa and 28 kDa protein.

Left panel, Western blot of HeLa nuclear extract probed with crude S752 antiserum.

Center panel, Western blot of HeLa nuclear extract probed with serum after adsorption on MBD2b affinity column.

Right panel, Western blot of HeLa nuclear extract probed with affinity purified anti-MBD2 antibodies. A non MBD2-specific band is indicated, this is likely due to cross-reactivity of the secondary anti-sheep antibodies.

G) The S752 antiserum only detected recombinant MBD2b, but not recombinant MBD1 or MBD4.

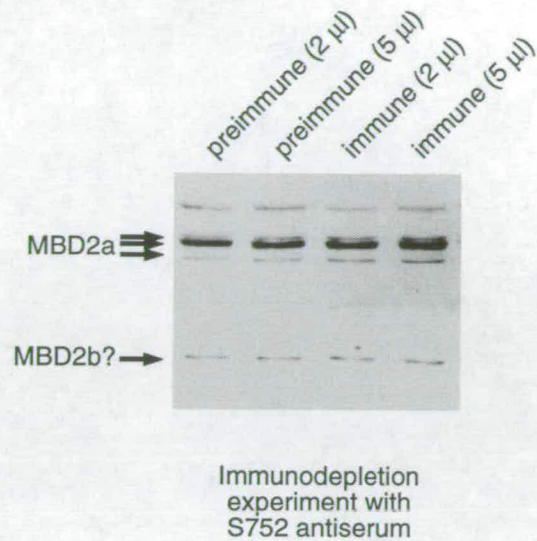
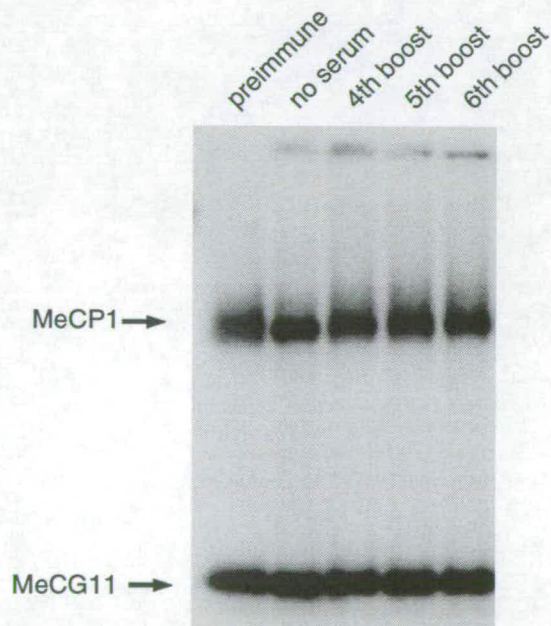
**H****I**

Figure 4.2.1.1. Characterisation of anti-MBD2 S752 antibody

H) S752 antibody cannot immunodeplete MBD2 from HeLa nuclear extracts. Different amounts of preimmune or immune S752 antibodies (2 µl or 5 µl of serum) were immobilised on protein G-sepharose beads. The beads were then incubated with HeLa nuclear extracts, and the supernatants were analysed by Western blotting using anti-MBD2 (S752) antiserum.

I) S752 antiserum did not supershift MeCP1 bandshift activity. 1 µl of serum (preimmune serum or immune sera from different bleeds) was added to HeLa nuclear extracts. The mixtures were then incubated with <sup>32</sup>P-labeled MeCG11 probe (a 135 bp probe with 27 methyl-CpGs). The DNA/protein complexes were resolved on 1.5 % agarose gel.

#### 4.2.1.2 S923 and R593 antibodies

As S752 antibody did not recognise native MBD2, it cannot be used for immunoprecipitation experiments. Further attempts to make anti-MBD2 antibody focused on raising antibodies against larger regions of MBD2.

S923 antiserum was raised against full-length MBD2 (MBD2a form). The immunogen used was purified GST-MBD2a fusion protein. This preparation contained mostly C-terminal truncation products (Figure 4.2.1.2 A, left panel). Based on the size of the degraded products, it is estimated that the site of truncation occurred before the MBD region (Figure 4.2.1.2 A, right panel). Therefore, majority of the antigen will not contain MBD2b sequence.

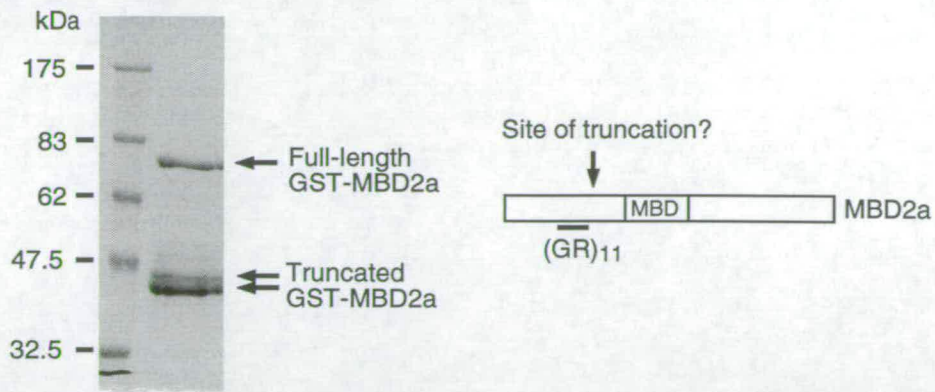
S923 antibody recognised MBD2a on a Western blot of HeLa nuclear extract (Figure 4.2.1.2 B, left panel). The reactivity of this antibody towards recombinant MBD2b and the 28 kDa band of HeLa nuclear extract is weak compared to another anti-MBD2 antibody R593 (see below). The S923 antibody did not cross-react with MBD1 and MBD4 (Figure 4.2.1.2 B, right panel).

R593 (R specifies the host animal which is a rabbit) was raised against MBD2b. This antibody detected both MBD2a and the 28 kDa protein on a Western blot of HeLa nuclear extract (Figure 4.2.1.2 C, left panel). Under the same intensity of MBD2a bands, the R593 antibody recognised recombinant MBD2b and the 28 kDa protein more efficiently. The R593 antibody did not cross-react with MBD1 and MBD4 (Figure 4.2.1.2 C, right panel). The 70 kDa band in the MBD1 lane is a bacterial protein that co-purifies with several recombinant proteins purified from *E.coli*. The additional bands greater than 47.5 kDa on the Western blots detected with S923 and R593 antibodies are unlikely to be MBD2 specific because they were not detected with the S752 antibody.

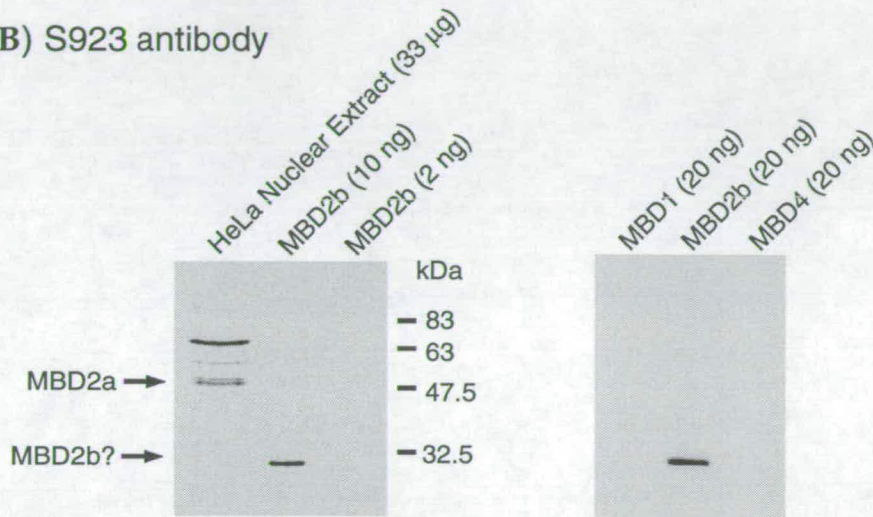
Both S923 and R593 antisera were tested for the ability to differentiate between two highly related proteins, MBD2b and MBD3 (71 % identity). 30 ng of MBD2b and MBD3 were loaded for Western analysis. Under identical condition, S923 antiserum did not detect recombinant MBD3 while R593 antiserum reacted with recombinant MBD3 (Figure 4.2.1.2 D). No band was detected in the MBD3 lane with S923 antiserum even after prolonged exposure. It is however not certain if the S923 antibodies can recognise larger



A) Antigen for S923 antibody



B) S923 antibody



C) R593 antibody

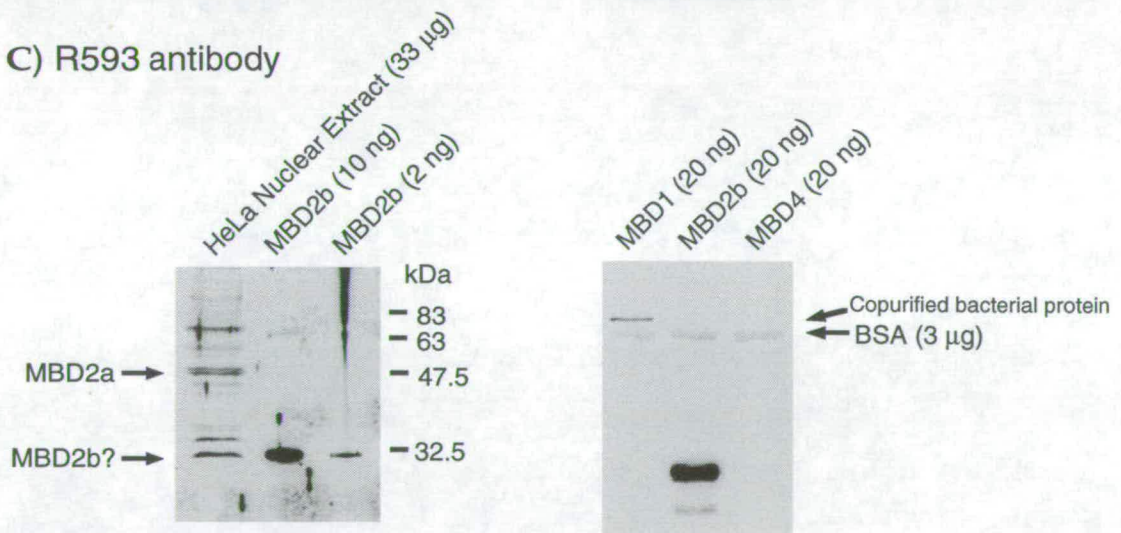


Figure 4.2.1.2. Characterisation of anti-MBD2 S923 and R593 antibodies

A) Antigen for raising S923 antibody was enriched with N-terminal MBD2a fragments.

The approximate site of truncation based on the sizes of GST-MBD2a products is before the MBD-like region.

B) S923 antiserum detected the 47.5 kDa MBD2a, and cross-reacted weakly with the 28 kDa protein (left panel). S923 antiserum did not cross-react with MBD1 or MBD4 (right panel).

C) R593 antiserum detected both the 47.5 kDa MBD2a and 28 kDa protein (left panel). R593 antiserum did not cross-react with MBD1 or MBD4 (right panel). The antiserum also detected a co-purified bacterial protein in the MBD1 preparation. A weak reaction with BSA (3 μg) which was present in all lanes was observed.

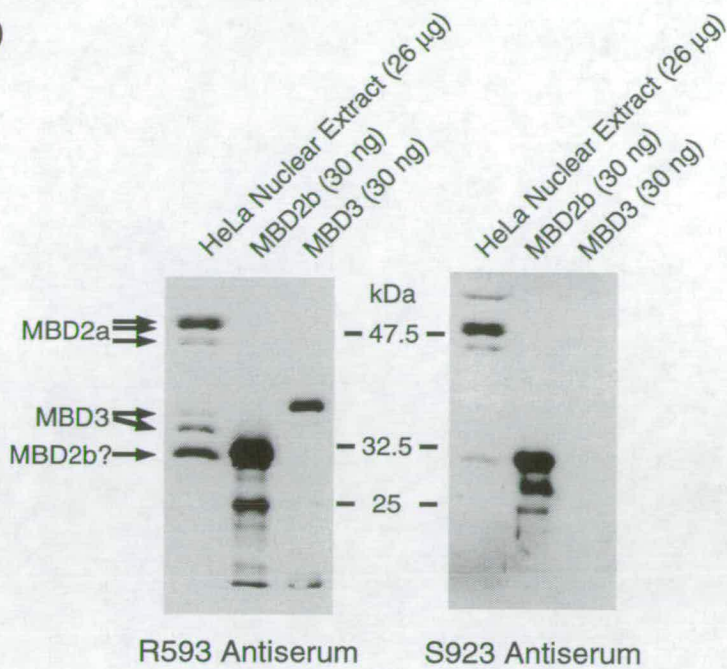
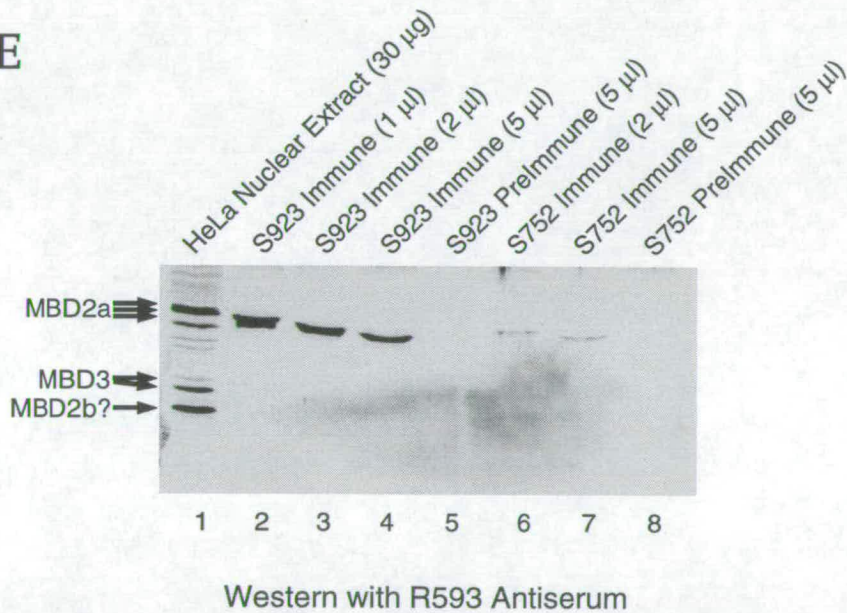
**D****E**

Figure 4.2.1.2. Characterisation of anti-MBD2 S923 and R593 antibodies

D) S923 antiserum (right panel) had minimal cross-reactivity with MBD3 (a protein with 71% identity compared to MBD2b). R593 antiserum (left panel) cross-reacted with MBD3. In addition, R593 antiserum also detected two proteins larger than 28 kDa on Western blot of HeLa nuclear extracts. As these two bands were not detected by S923 antiserum, they are likely to be MBD3.

E) S923 antiserum can immunoprecipitate MBD2a. Different amount of antibodies (S923 immune, S923 preimmune, S752 immune, and S752 preimmune) were used to immunoprecipitate nuclear proteins from HeLa nuclear extracts. The immunoprecipitates were analysed by Western blot using R593 antiserum. Only MBD2a was immunoprecipitated by S923 antibodies. S752 antibodies bound very low level of MBD2a.



amount of recombinant MBD3. Both antisera recognised the 28 kDa “MBD2b?” band, but with different efficiency. It is likely that the additional bands greater than 28 kDa are MBD3 proteins (Figure 4.2.1.2 D). As R593 antibody cross-reacts with MBD3, it is not characterised further and not used for experiments addressing MBD2.

Two anti-MBD2 antibodies (S923 and S752) were tested for their ability to immunoprecipitate native MBD2 from HeLa nuclear extracts. The immunoprecipitates were probed with R593 antiserum which bound both MBD2 and MBD3. It is found that the S923 antibodies can efficiently bind MBD2a, while S752 antibodies immunoprecipitated very low level of MBD2a (Figure 4.2.1.2 E). In addition, the S923 antibodies only immunoprecipitate MBD2a, but not the other smaller proteins detected by the R593 antiserum. It is likely that the epitopes recognised by S752 antibodies were not exposed in the HeLa nuclear extract, and can account for the different ability of the two antibodies in immunoprecipitating MBD2a. The MBD2 bands were displaced downward in certain lanes (e.g. lane 3,4,7) due to co-migrating IgG. It is concluded that S923 antibodies can recognise native MBD2.

The properties of these three antibodies are summarised in Figure 4.2.1.2 F.

#### **4.2.2 MBD2 is a component of MeCP1**

It is of interest to test whether MBD2 is responsible for MeCP1 activity. As shown in the previous section, S923 antiserum was most efficient in immunoprecipitating native MBD2 from HeLa nuclear extracts. Therefore, S923 antiserum was used for MeCP1 supershift experiment.

The S923 antiserum can quantitatively supershift HeLa MeCP1 activity, while the preimmune antiserum had no effect on the mobility of MeCP1 (Figure 4.2.2 A). Both the immune and preimmune antisera did not affect the DNA/protein complex on the non-methylated version of CG11 probe. The nature of the DNA/protein complex on non-methylated CG11 is not known, and this activity has also been observed previously by Boyes and Bird (1991).

Another antibody S752 raised against an internal fragment of MBD2 (corresponding to amino acids 232 to 394 of murine MBD2a) was not effective in immunoprecipitating native MBD2, and failed to supershift the MeCP1 activity (see Figure 4.2.1.1 I). Thus,

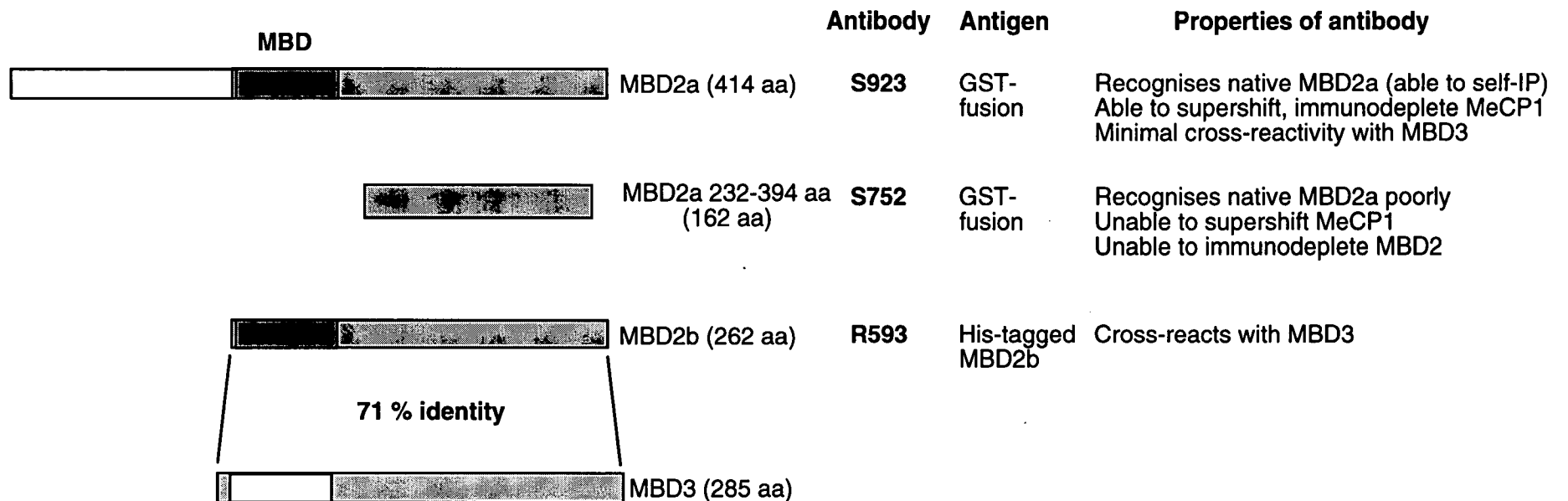


Figure 4.2.1.2.F Summary of anti-MBD2 S752, S923 and R593 antibodies.

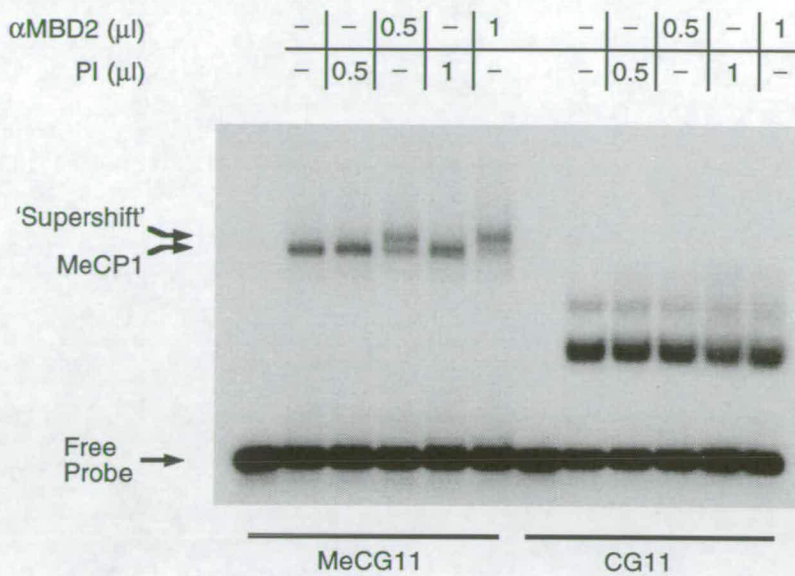
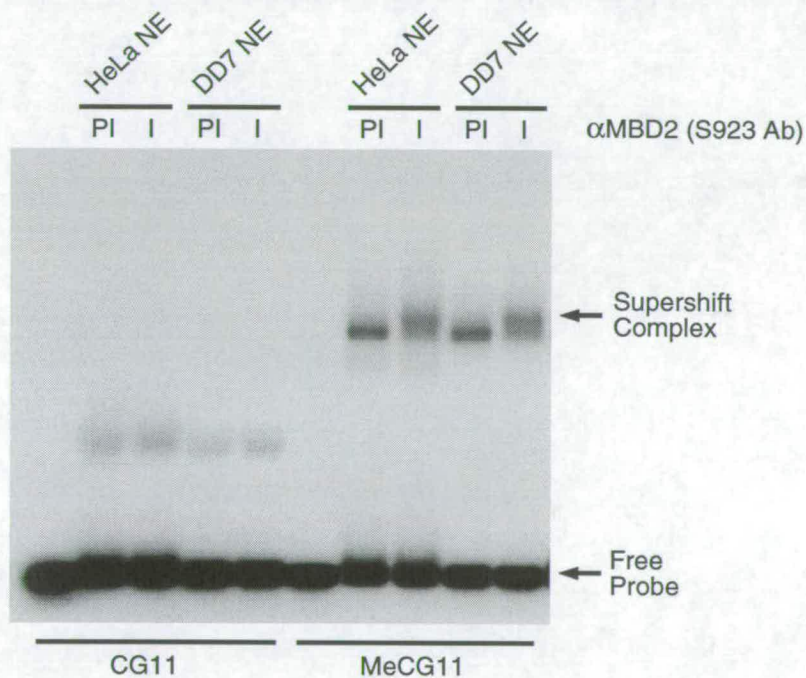
**A****B**

Figure 4.2.2. MBD2 is a component of MeCP1.

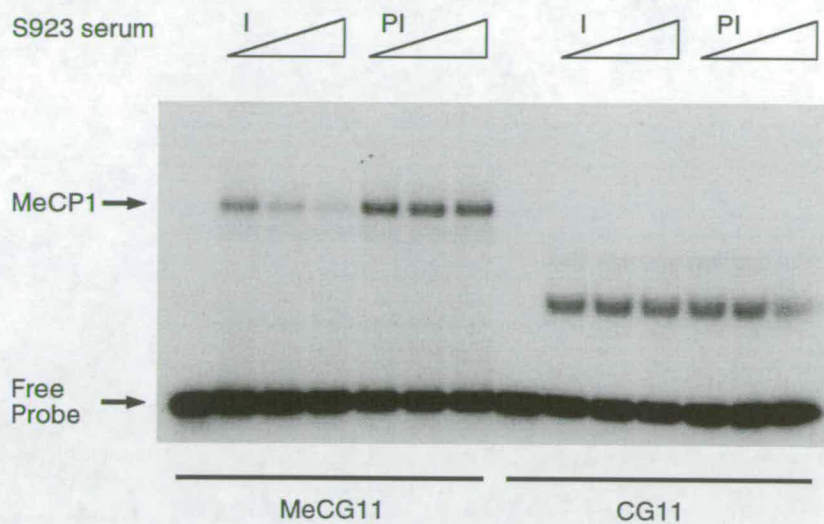
A) S923 antiserum can supershift HeLa MeCP1 activity.

Different amount of serum (preimmune (PI) or immune) were added to HeLa nuclear extracts before subsequent incubation with methylated (MeCG11) or non-methylated (CG11) probes. S923 antiserum specifically supershifted the MeCP1 activity, but not the DNA/protein complex on the non-methylated probe. The preimmune serum did not affect the mobility of MeCP1 or the DNA/protein complex on the non-methylated probe.

B) S923 antiserum can also supershift MeCP1 activity of DD7 nuclear extract.

1  $\mu$ l of preimmune (PI) or immune (I) serum was added to HeLa nuclear extracts (NE) or DD7 nuclear extracts (NE) before bandshift assays.

C



D

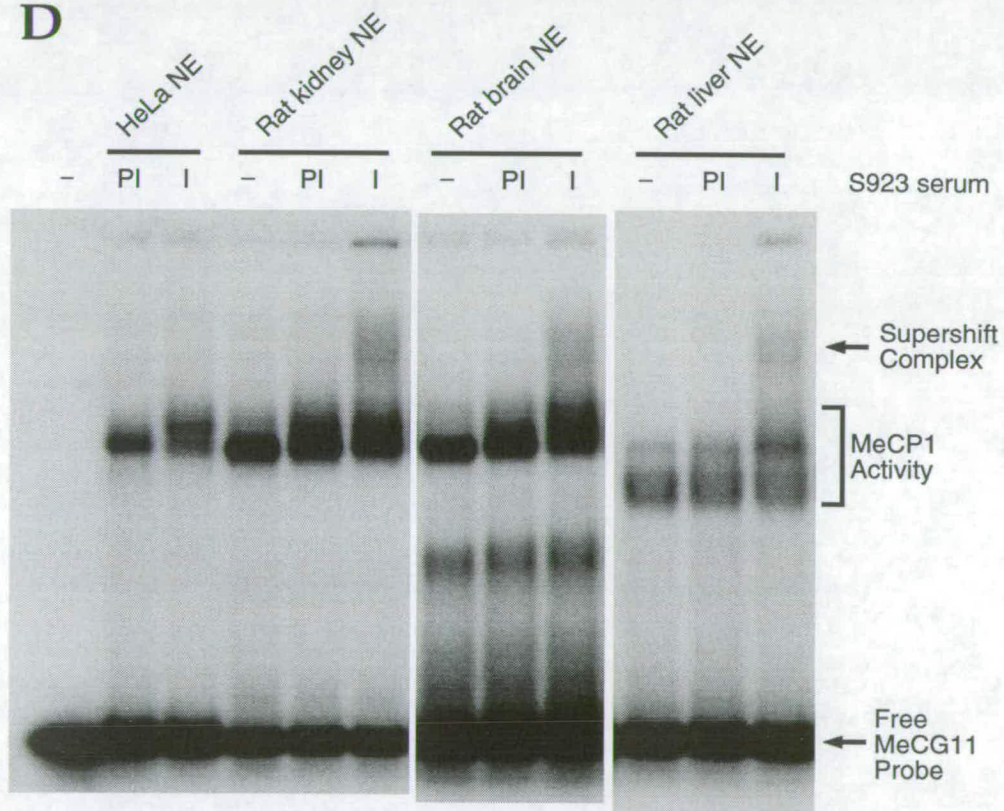


Figure 4.2.2. MBD2 is a component of MeCP1.

C) Immunodepletion of MBD2 from HeLa nuclear extracts with S923 antiserum (2.5, 5, and 10  $\mu$ l) resulted in the loss of MeCP1 bandshift activity. Preimmune sera did not deplete MeCP1 activity. The DNA/protein complexes of the non-methylated probe were not affected by immune (I) or preimmune (PI) antibodies depletion.

D) S923 antiserum can also supershift MeCP1 activity of tissue nuclear extracts. Nuclear extracts were made from rat kidney, rat brain and rat liver. 1  $\mu$ l of preimmune (PI) or immune (I) S923 serum was added to each bandshift reaction.



there is a correlation between the ability of the antibodies to bind native MBD2 and supershift MeCP1 activity.

The S923 antiserum was also tested for its ability to supershift MeCP1 activity derived from another human cell-line, DD7. DD7 is a human lymphoblastoid line containing a balanced X/9 translocation. Again, the immune antiserum can supershift DD7 MeCP1 activity, while the preimmune serum did not affect the mobility (Figure 4.2.2 B).

Immunodepletion experiments were also carried out to determine whether the antiserum S923 can deplete HeLa MeCP1 activity. Different amounts of the immune and preimmune IgG were first bound to protein G sepharose. The beads were then mixed with HeLa nuclear extracts. Supernatants were subsequently separated from the beads, and assayed for MeCP1 bandshift activity. Immune IgG were found to deplete MeCP1 in a quantitative manner (Figure 4.2.2 C). Preimmune IgG had no effect on MeCP1. Both immune and preimmune IgG did not deplete the non-specific complex on the non-methylated CG11 probe.

This series of experiments leads to the conclusion that MBD2 is a component of MeCP1 in HeLa cells.

Interestingly, S923 antiserum can also supershift MeCP1 activity of nuclear extracts derived from rat brain and kidney tissues (Figure 4.2.2 D). Preimmune antiserum did not affect the mobility of MeCP1 activity in the same bandshift assay. It is notable that the supershift by the immune serum was different from HeLa MeCP1 supershift. The retarded 'supershifted' complexes were of much lower mobility. In addition, the immune antiserum did not shift a remaining portion of the MeCP1 activity. The result suggests that other methyl-CpG binding proteins may potentially contribute to the composite MeCP1 activity derived from somatic tissue extracts.

### **4.2.3 F9 embryonic carcinoma (EC) cells are deficient in MBD2**

Although DNMT1 is crucial for the mouse embryogenesis, it is not required for the survival and proliferation of embryonic stem cells, and these cells can also tolerate very low levels of DNA methylation (Li et al, 1992). Interestingly, ES and EC cells are deficient in MeCP1 activity (Meehan et al, 1989; Boyes and Bird, 1991). The lack of MeCP1 in F9 embryonic carcinoma cell-line correlates with the reduced ability of these



Figure 4.2.3. F9 embryonic carcinoma cells which lack MeCP1 activity are deficient in MBD2.

F9 EC nuclear extract (25  $\mu$ g), rat liver nuclear extract (25  $\mu$ g), HeLa nuclear extract (30  $\mu$ g), and rat brain nuclear extract (22  $\mu$ g) were analysed by Western blot using S923 antiserum. The position of MBD2a is indicated. The blot was also probed with anti-mSin3A (K20) antibody for loading control.

cells to silence methylated reporter constructs. This data strongly suggests that MeCP1 is the mediator for transcriptional repression by DNA methylation (Boyes and Bird, 1991). The level of MBD2 mRNA was also found to be drastically reduced in ES cells (Hendrich and Bird, 1998).

Western blot analysis using anti-MBD2 antibody was carried out on nuclear extracts made from rat liver and F9 embryonic carcinoma cells. It was found that there is no detectable MBD2 in F9 nuclear extract (Figure 4.2.3). However, these extracts contained comparable levels of the mSin3A corepressor protein, thus the lack of MBD2 in F9 nuclear extract was not due to unequal loading of proteins. It is also apparent that rat brain nuclear extract which is enriched in MeCP2 has lower level of MBD2 compared to rat liver nuclear extract, this suggests that there may be a tissue-specific variation in the abundance of these MeCPs.

#### **4.2.4 MBD2 is associated with histone deacetylases**

MeCP2 had been shown to interact with the mSin3A/ histone deacetylases complex. It is of interest to see if MBD2 is also associated with histone deacetylases. S923 antiserum was used to immunoprecipitate nuclear proteins. This antibody was chosen because of its ability to recognise native MBD2. The immunoprecipitated materials were separated by SDS-PAGE, and followed by Western blotting analysis to detect co-immunoprecipitated proteins. HDAC1, HDAC2 and RbAp48 were immunoprecipitated by antibody S923 (Figure 4.2.4 A). Control preimmune antibody did not immunoprecipitate these proteins. The western blot was also probed with anti-retinoblastoma and anti-p150 (the largest subunit of chromatin assembly factor 1) antibodies. Retinoblastoma protein (Rb) is known to interact with histone deacetylases (HDAC1 and HDAC2), but was not immunoprecipitated by anti-MBD2 antibody (Brehm et al, 1998; Luo et al, 1998; Magnaghi-Jaulin et al, 1998). Chromatin assembly factor 1 complex contains the core histone binding protein, RbAp48 (Verreault et al, 1996; 1998). Although RbAp48 protein was found in anti-MBD2 immunoprecipitate, p150 subunit of CAF1 was not detected. The absence of Rb and p150 in anti-MBD2 immunoprecipitates indicates that the antibody can specifically bringing down MBD2 associated proteins.

Does the anti-MBD2 immunoprecipitate contain histone deacetylase activity? This question was answered by experiments carried out by our collaborator Dr Colin Johnson (University of Birmingham). He found that approximately 27 % of total histone

A

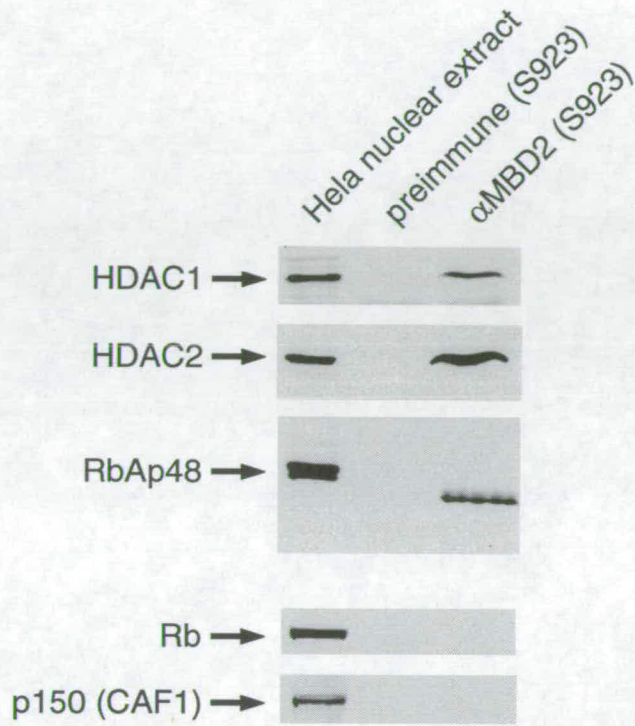


Figure 4.2.4. MBD2 is associated with histone deacetylases.

A) Anti-MBD2 antibody can immunoprecipitate histone deacetylases HDAC1, HDAC2, and histone-binding protein RbAp48 from HeLa nuclear extracts. Preimmune antibodies did not immunoprecipitate these proteins. Anti-MBD2 antibody did not immunoprecipitate HDAC1 associated protein, Rb, and RbAp48 associated protein, p150 of CAF1.



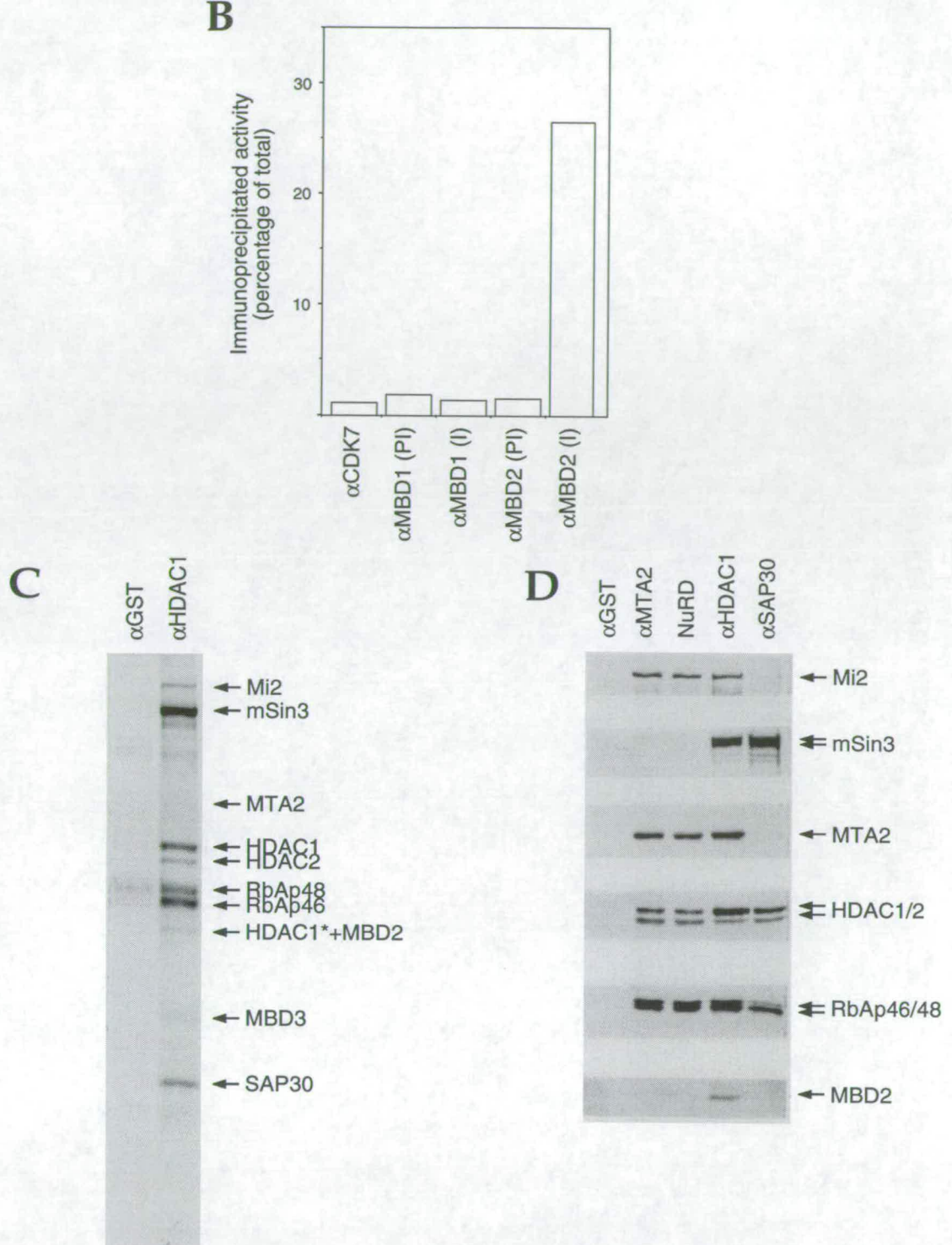


Figure 4.2.4. MBD2 is associated with histone deacetylases.

B) Anti-MBD2 antibodies immunoprecipitated histone deacetylase activity from HeLa crude cell extracts, but preimmune, anti-CDK7 and anti-MBD1 antibodies did not.

C) MBD2 is present in affinity-purified HDAC1 complexes.

A silver-stained gel showing the proteins retained on HDAC1 ( $\alpha$ HDAC1) and GST ( $\alpha$ GST) antibody columns. Mass spectrometric and sequence analysis of the 43 kDa band identified a peptide contained in HDAC1, as well as a 39 amino-acids peptide that matches MBD2 sequence. Known components of the complexes are indicated on the right.

D) MBD2 is not present in previously characterised NuRD or mSin3 complexes.

Affinity purified ( $\alpha$ MTA2,  $\alpha$ HDAC1,  $\alpha$ SAP30) or native (NuRD) complexes were analysed by Western blotting using anti-Mi2, anti-mSin3, anti-MTA2, anti-HDAC1, anti-HDAC2, anti-RbAp16, anti-RbAp48, and anti-MBD2 antibodies.

deacetylase activity can be immunoprecipitated by anti-MBD2 antibody (Figure 4.2.4 B). Control antibodies such as preimmune, anti-CDK7, or anti-MBD1 antibodies did not bring down histone deacetylase activity.

Although anti-MBD2 antibody can immunoprecipitate histone deacetylases and deacetylase activity, it is not certain if this is due to cross-reactivity of the antibody towards components of histone deacetylase complexes. Therefore, reciprocal immunoprecipitation experiments are essential to confirm the interaction between MBD2 and histone deacetylases. Our collaborator Dr Yi Zhang (New Jersey) found that affinity purified HDAC1-containing complexes contain MBD2 (Figure 4.2.4 C). This was determined by microsequencing of a 43 kDa band from the silver-stained gel. A derived peptide sequence of 39 amino acids (GLQGVGPGSNDETLLSAVASALHTSSAPITGQVSAÁVEK) was identical to sequence of MBD2. Microsequencing also showed that the 43 kDa silver-stained band contained sequences derived from HDAC1. Western blotting of the affinity purified HDAC1 complexes with anti-MBD2 antibody independently corroborated the interaction between MBD2 and HDAC1 (Figure 4.2.4 D). Interestingly, MBD2 was not found in affinity purified NuRD (anti-MTA2 lane) and mSin3 (SAP30 lane) complexes. This suggests that MBD2 is found in a novel HDAC1-containing complex.

One prediction based on the MBD2/ HDAC1 association is that anti-HDAC1 antibody should immunodeplete MeCP1 activity. Indeed, anti-HDAC1 antibody could quantitatively reduce the level of both MeCP1 and MBD2 (Figure 4.2.4 E). Control anti-SAP30 antibody had no significant effect on either MeCP1 or MBD2. The two non-specific bands (indicated by asterisks) bound by S923 anti-MBD2 antiserum were not depleted by HDAC1 antibodies, and thus serves as internal negative controls for the specific depletion of MeCP1 and MBD2 using anti-HDAC1 antibodies. This result further substantiates the links between MBD2, MeCP1 and HDAC1.

Curiously, a panel of antibodies directed against components of the HDAC1/2 complexes (anti-HDAC1, anti-HDAC2, and anti-RbAp48) did not supershift the MeCP1 bandshift activity (Figure 4.2.4 F). Given the observation that anti-HDAC1 antibodies can immunodeplete both MBD2 and MeCP1 activity, it is proposed that the MeCP1 bandshift activity contains only the DNA binding component which is MBD2. It is conceivable that the MBD2/ HDAC complex may have dissociated during the process of electrophoresis.



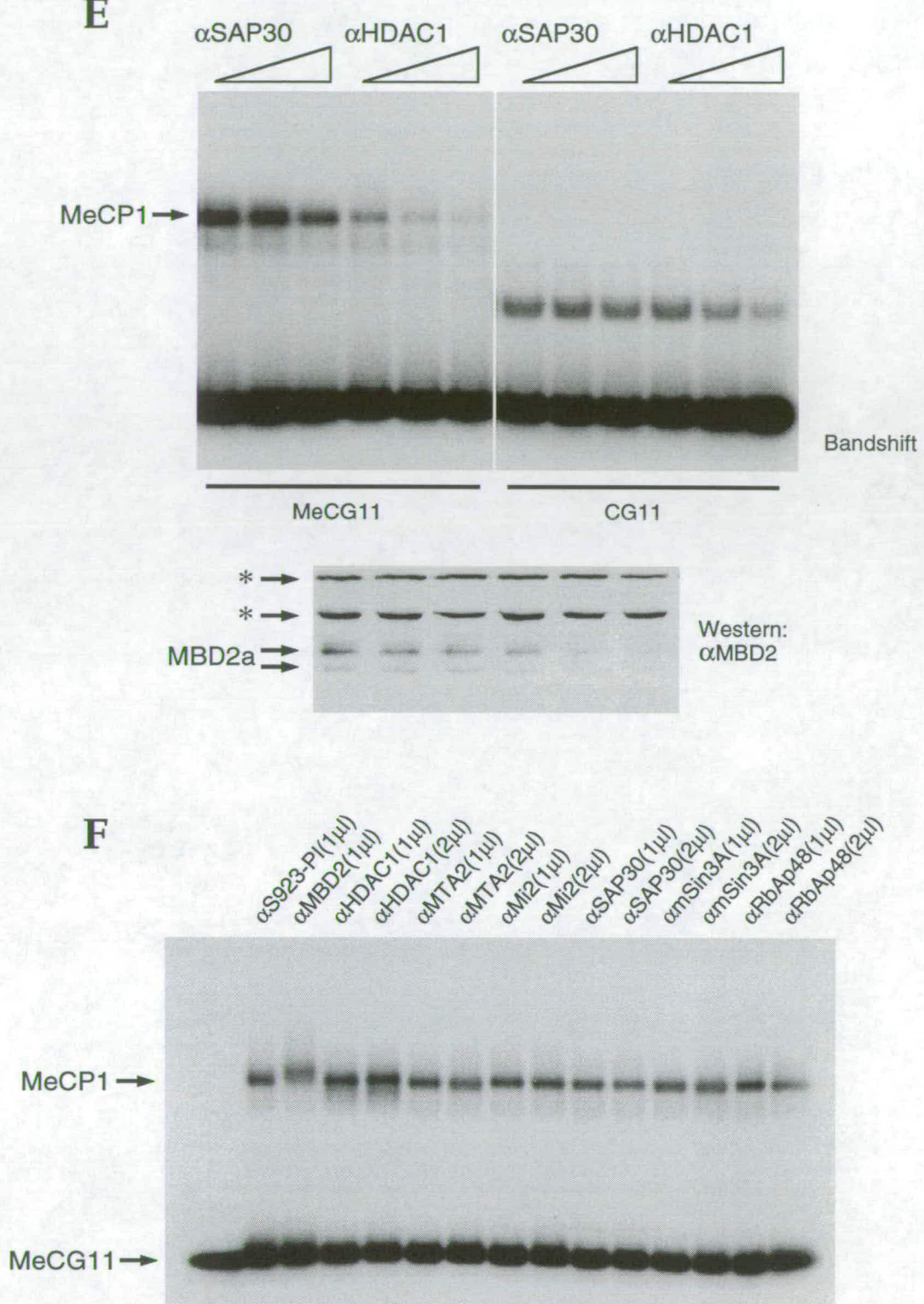


Figure 4.2.4. MBD2 is associated with histone deacetylases.

E) Anti-HDAC1 antibodies (10  $\mu$ l, 20  $\mu$ l, 30  $\mu$ l) immunodepleted both MBD2 and MeCP1 activity. Control anti-SAP30 did not deplete MBD2 or MeCP1 activity. The asterisks indicate two non-MBD2 specific bands which are bound by anti-MBD2 S923 antiserum. These bands were not depleted by HDAC1 or SAP30 antibodies.

F) Effect of anti-MBD2, anti-HDAC1, anti-MTA2, anti-Mi2, anti-SAP30, anti-mSin3A, and anti-RbAp48 antibodies on MeCP1 bandshift activity. The amount of serum added to the bandshift reactions is indicated.

Alternatively, the conditions for the bandshift assay may lead to the dissociation of MBD2 / HDAC complex.

#### **4.2.5 Interaction between MBD2 and other components of HDAC1/2 complex**

HDAC1, HDAC2, RbAp48 and MBD2a were synthesised *in vitro*, labeled with <sup>35</sup>S-methionine and tested for interaction with GST-MBD2a. It is found that GST-MBD2a can interact very weakly (less than 5 % of input was bound) with HDAC1, HDAC2 and RbAp48 (Figure 4.2.5). On the other hand, MBD2a bound efficiently to GST-MBD2a (more than 10 % of input was bound). These interaction appear to be specific, as the GST control did not bind to the *in vitro* translated proteins. The self-interaction of MBD2a is interesting because it may facilitate the binding of MBD2a complex (MeCP1) to densely methylated DNA. It is perhaps not surprising that MBD2a can interact with at least four proteins *in vitro*, because a highly related protein, MBD3, can bind directly to four proteins, namely MTA2, HDAC1, RbAp46 and RbAp48 (Zhang et al, 1999). However, the possibility that a yet unknown protein(s) in the *in vitro* translation mixture mediates the interaction cannot be excluded with the GST-pulldown assay.

#### **4.2.6 MBD2 exists in a high molecular weight complex**

Previous work had demonstrated that MeCP1 migrates with a large molecular weight (from 400 to 800 kDa) on a gel-filtration column. When native HeLa extract was fractionated on a Superose 6 gel-filtration column, MBD2 eluted in high molecular weight fractions (Figure 4.2.6). It is estimated that MBD2 migrates at about 800 kDa. The result is consistent with the finding that MBD2 can interact with multiple proteins *in vitro* and is found in a multiprotein HDAC1/2 complex. Unfortunately, MeCP1 activity could not be detected in all the fractions, therefore in this experiment the relationship between MBD2 and MeCP1 cannot be critically determined.

#### **4.2.7 MBD2 can repress transcription *in vitro***

Previously, MeCP2 was shown to repress transcription of a methylated template specifically *in vitro* (Nan et al, 1997). The same assay was used to test whether MBD2 can repress transcription in a cell-free system. Recombinant GST-MBD2a or GST-

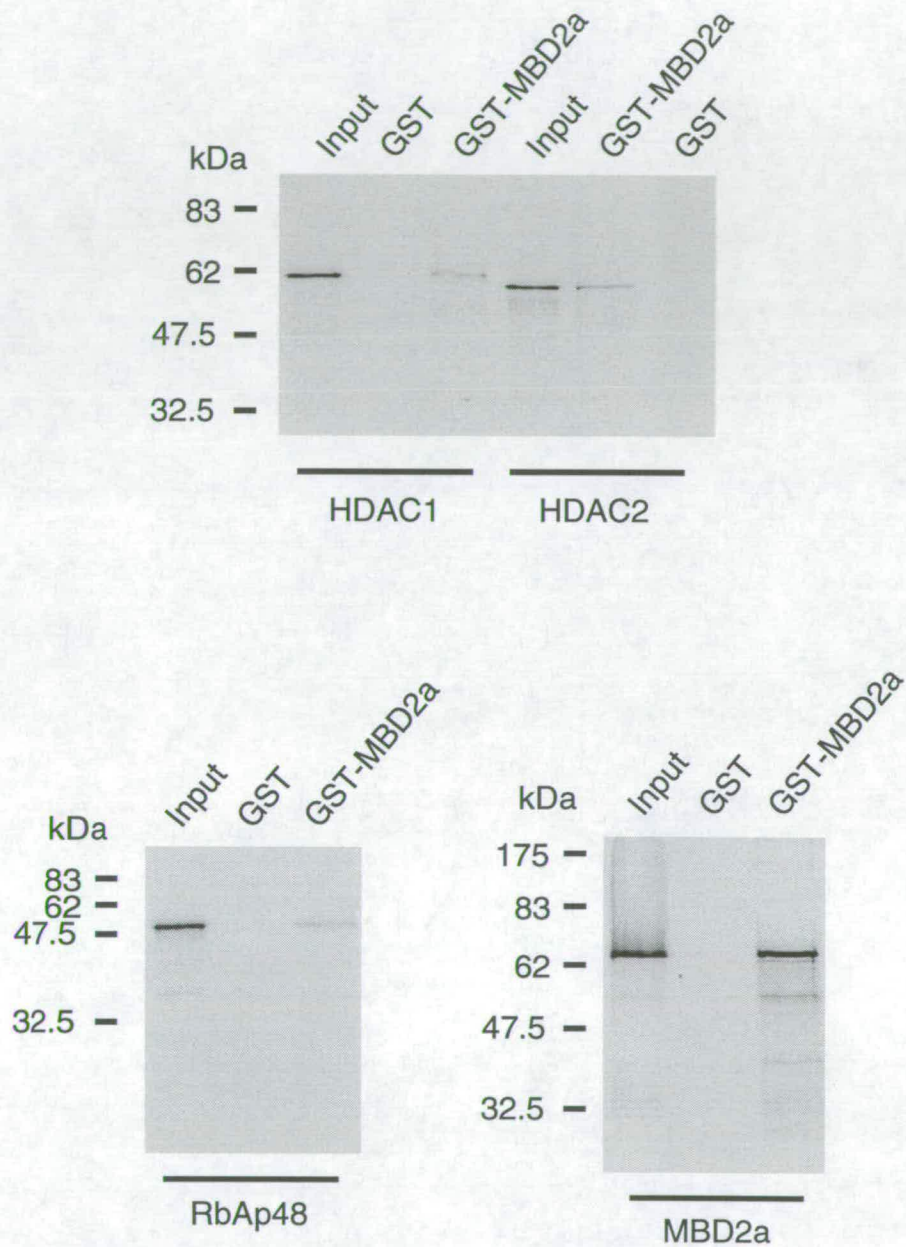


Figure 4.2.5. MBD2 interacts weakly with HDAC1, HDAC2, RbAp48, but strongly with itself.

HDAC1, HDAC2, RbAp48, and MBD2a were *in vitro* translated and labeled with  $^{35}\text{S}$  methionine. The labeled products were incubated with GST or GST-MBD2a. Bound proteins were eluted, separated by SDS-PAGE, and detected by autoradiography. *In vitro* translated MBD2a contained 5 myc epitopes, and migrated with a lower mobility compared to untagged proteins.



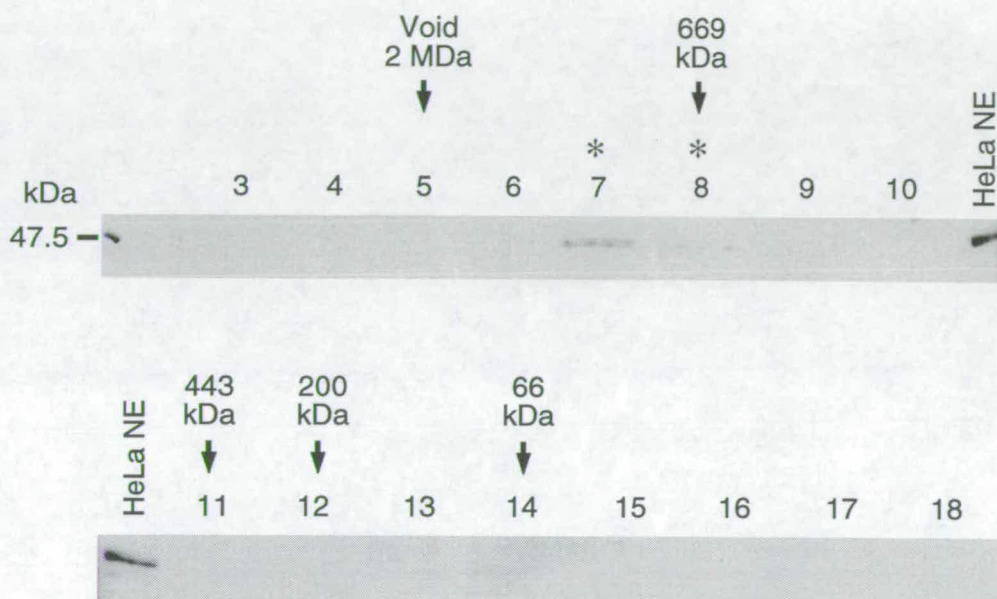


Figure 4.2.6. MBD2 exists in a high molecular weight complex.

HeLa nuclear extract was fractionated by Superose 6 gel-filtration column. The fractions were analysed by Western blotting using anti-MBD2 S923 antiserum. MBD2 was found in fractions 7 and 8 (indicated by asterisks) corresponding to approximately 800 kDa. The column was calibrated with known molecular weight markers. Identical result was obtained using anti-MBD2 S752 antiserum (data not shown).

MBD2b was incubated with pAdomal template (construct driven by adenovirus major late promoter) methylated at 44 HhaI sites (Figure 4.2.7). HeLa nuclear extracts and NTPs were then added to the mixture to initiate transcription from the reporter template. A mock-methylated pAdBXN template (pAdomal with 18 bp insertion downstream of transcription start site) was included as an internal positive control in these assays. Due to the extra 18 bp insertion in pAdBXN reporter, the RNA product is 18 nucleotides longer than that transcribed from the pAdomal reporter. The RNA transcripts derived from the pAdomal and pAdBXN reporters were determined by primer extension assay. In the absence of recombinant GST-MBD2 protein, the relative transcription of the HhaI methylated template (ratio of M-/ M+) was close to 1. Addition of 1  $\mu$ g and 2  $\mu$ g of GST-MBD2a resulted in 2 fold and 4 fold repression of the methylated template respectively. GST-MBD2b was found to abolish transcription from the methylated template. The preparation of GST-MBD2a contained mostly C-terminal truncation products, and therefore deficient in products with the MBD region. This is likely to account for the weaker repression observed using the full-length MBD2.

It is however not possible to distinguish whether MBD2 is repressing the methylated template by sterically blocking the access of transcription factors or through an active repression mechanism. Therefore the next experiment is designed to address if MBD2 can function as a transcriptional repressor *in vivo* when tethered to specific sites of a reporter.

#### **4.2.8 MBD2 can repress transcription *in vivo***

GAL4-MBD2a fusion construct was co-transfected with a reporter gene containing Gal4 binding sites upstream of the promoter (G5). The GAL4-MBD2a fusion protein is tethered to the reporter construct *in vivo*, and the influence of MBD2 on the promoter can be assayed. A reporter without Gal4 binding sites (G0) was used as a control in separate co-transfection experiments. GAL4-MBD2a repressed the G5 reporter to a greater extent compared to the G0 reporter (Figure 4.2.8 A, top graph). After normalising the  $\beta$ -galactosidase activity of the G5 or G0 reporters with the luciferase activity from an internal pGL2 reporter, it is clear that the repression of GAL4-MBD2a depends on the presence of Gal4 binding sites (Figure 4.2.8 A, bottom graph). Strangely, increasing amounts of co-transfected effector did not reduce the relative activity of the G5 reporter further. Western blot analysis of the lysates from transfected cells with anti-GAL4 antibodies indicated that expression of GAL4-MBD2a from the effector construct was not saturated (Figure 4.2.8 C).

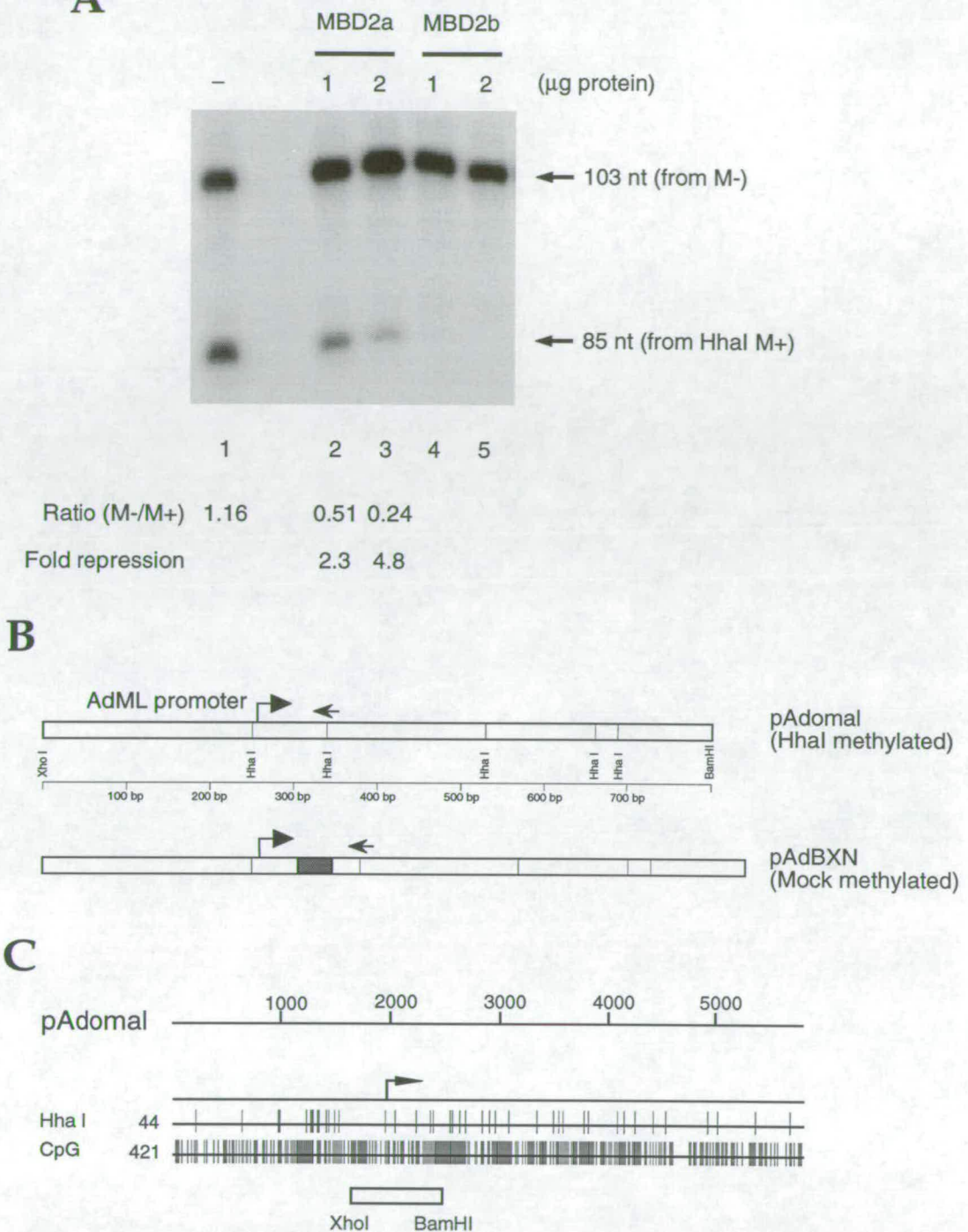


Figure 4.2.7. Methylation dependent repression by recombinant MBD2.

A) Recombinant MBD2a and MBD2b specifically repress the methylated pAdomal promoter *in vitro*.

Different amounts of recombinant MBD2a and MBD2b (1  $\mu\text{g}$  or 2  $\mu\text{g}$ ) were incubated with mock methylated pAdBXN or HhaI methylated pAdomal reporters. HeLa nuclear extracts and NTPs were then added to the mixtures to initiate transcription from the two reporters. The level of transcription was quantitated by primer extension assays. No recombinant protein was added in lane 1. The ratio of M-/M+ was calculated for lane 1 to 3 using values obtained by PhosphoImager. The values for fold repression on the methylated reporter in the presence of MBD2a are shown.

B) Constructs used in A. The location of transcription start site, HhaI sites and primer are shown. The 18 bp insertion is shown as a gray box.

C) HhaI and CpG sites for pAdomal reporter are shown. Regions outside the XhoI/BamHI fragment correspond to plasmid sequences.



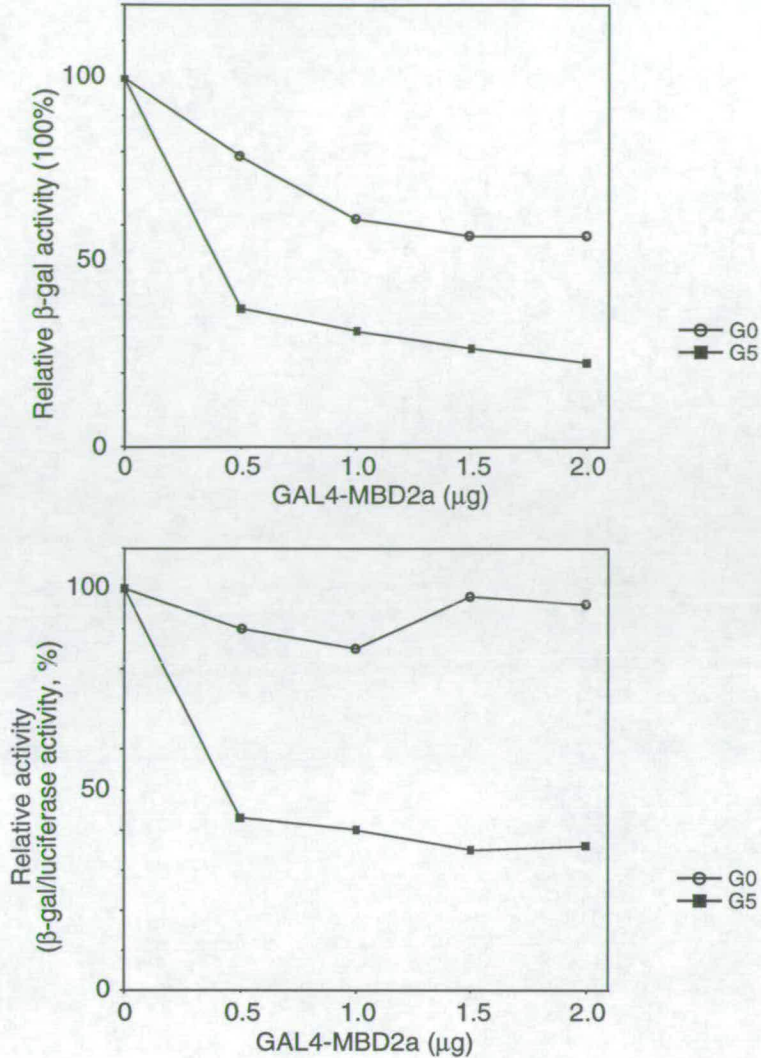
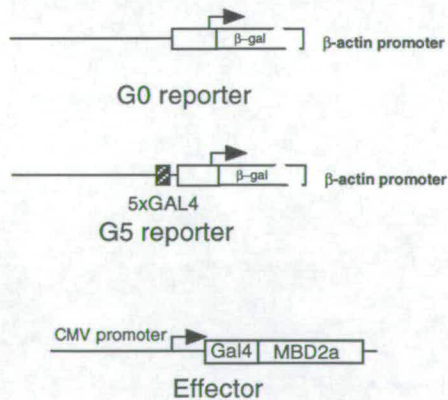
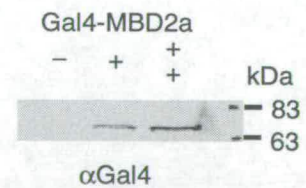
**A****B****C**

Figure 4.2.8. Transcriptional repression by MBD2.

A) GAL4-MBD2a specifically represses a reporter with Gal4 binding sites. Increasing amounts of GAL4-MBD2a effector was co-transfected with G0 or G5 reporters (human  $\beta$  actin promoter driving lacZ gene) in the presence of luciferase reporter as an internal control. Top graph,  $\beta$ -gal reporter activities are shown. Bottom graph, normalised reporter activity ( $\beta$ -gal/luciferase) are shown.

B) Constructs used in A.

C) The effector construct expressed a GAL4-MBD2a fusion protein of the correct size as revealed by Western blot analysis using anti-GAL4 antibodies. ++ (2  $\mu$ g) contains double amount of effector used compared to + (1  $\mu$ g).

The effect of GAL4-MBD2a was also tested on another reporter driven by the human DNA polymerase  $\beta$  promoter. The repression of the human DNA polymerase  $\beta$  promoter was weaker compared to that of the human  $\beta$ -actin promoter (Figure 4.2.8 D). Saturation of repression on the G5 reporter was again observed. Different preparation of DNA (cesium chloride or Qiagen ion exchange) did not affect the repression activity of GAL4-MBD2a (data not shown). The reason for the modest repression on both the human  $\beta$ -actin promoter and human DNA polymerase  $\beta$  promoter is not known. It is plausible that the presence of two DNA binding domains (methyl-CpG binding domain and GAL4 DNA binding domain) may lead to mistargeting of the GAL4-MBD2a fusion protein. Alternatively, full repression activity of MBD2a may require chromatinised templates, and such chromatin conformation may be absent in transiently transfected DNA templates (Wu, 1997).

#### **4.2.9 Trichostatin A (TSA) can relieve repression on the human DNA polymerase $\beta$ promoter**

As MBD2a was found associated with histone deacetylases and can repress transcription in a transient transfection assay, an inhibitor relief experiment will allow one to address whether repression is dependent on deacetylation *in vivo*. Transfected cells were incubated in media with 100 ng/ml TSA or 200 ng/ml TSA for 24 hours before collection for reporter assays. It is found that repression on the human  $\beta$ -actin promoter was not sensitive to TSA (Figure 4.2.9, left panel), while TSA abolished repression on the human DNA polymerase  $\beta$  promoter (Figure 4.2.9, right panel). Promoter-specific derepression upon TSA treatment has previously been documented (Luo et al, 1998; Koipally et al, 1999).

#### **4.2.10 Mapping of a repression domain?**

It is of interest to map the region in MBD2a that is responsible for the repression observed in section 4.2.8. A panel of GAL4-MBD2 fusion constructs was generated (Figure 4.2.10). The GAL4 fusion protein with amino-acids 215 to 414 aa of MBD2 repressed transcription of both G0 and G5 reporters, and showed no apparent specificity towards the G5 reporter. All GAL4 fusion proteins did not exhibit the repression profile of the

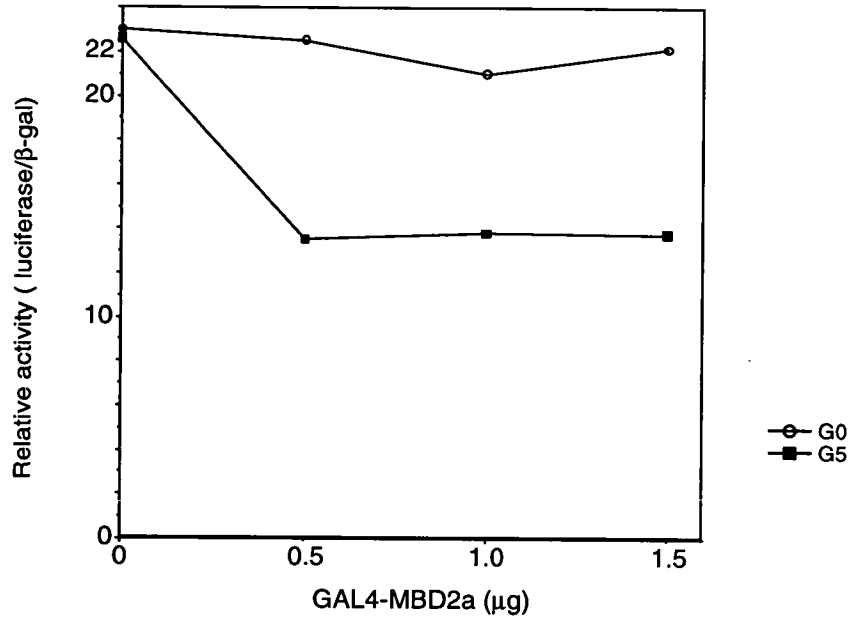
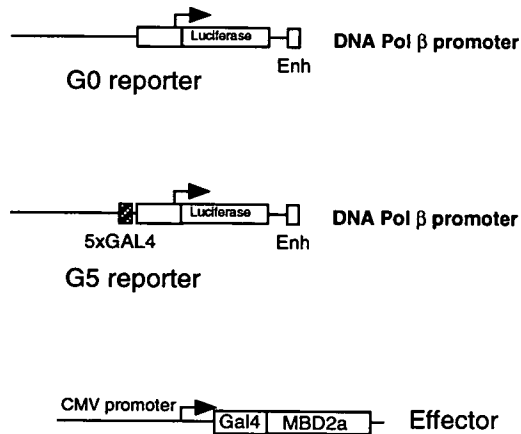
**D****E**

Figure 4.2.8. Transcriptional repression by MBD2.

D) GAL4-MBD2a specifically represses a reporter with Gal4 binding sites. Increasing amounts of GAL4-MBD2a effector was co-transfected with G0 or G5 reporters (human DNA polymerase β promoter driving luciferase gene) in the presence of lacZ reporter as an internal control. Normalised reporter activity (luciferase/β-gal) are shown.

E) Constructs used in D.

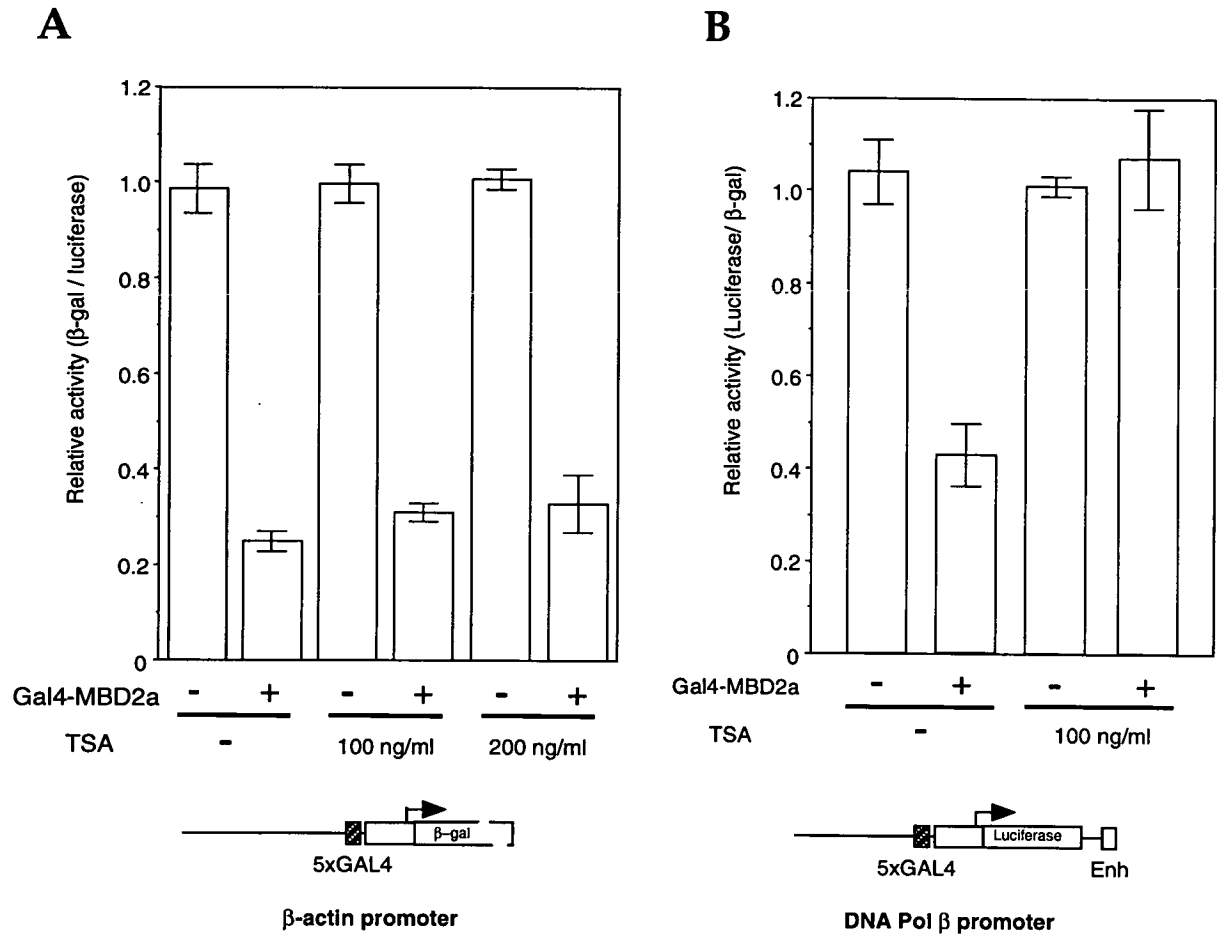


Figure 4.2.9. Effect of histone deacetylase inhibitor, TSA on repression by GAL4-MBD2a.

A) Repression on the β actin promoter by GAL4-MBD2a is not sensitive to TSA.

B) Repression on the DNA polymerase β promoter by GAL4-MBD2a is sensitive to TSA.

The results shown are based on three independent transfection.

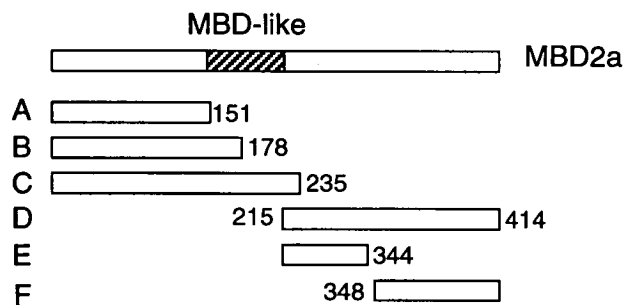
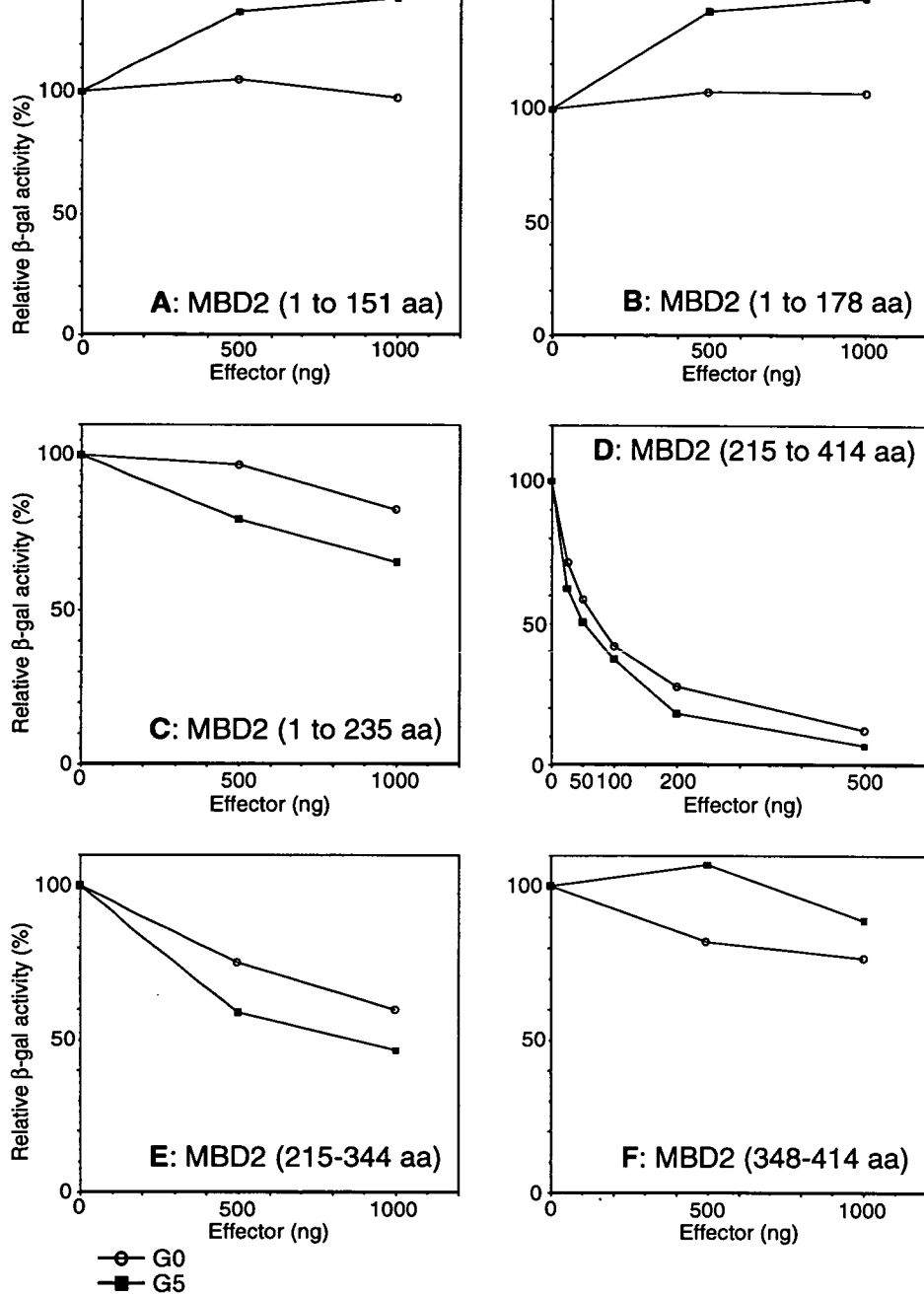


Figure 4.2.10. Mapping of a repression domain in MBD2a.

The effect of different GAL4-MBD2 fusions (A to F) was tested on G0 and G5 reporters (human  $\beta$  actin promoter driving lacZ gene). Repression profiles on G0 and G5 reporters are shown.

full-length MBD2. It is likely that the repression property of MBD2 is sensitive to deletion, as in the case of yeast histone deacetylase, RPD3 (Kadosh and Struhl, 1998a).

#### **4.2.11 Is MBD3 a methyl-CpG binding protein?**

The remarkable similarity between MBD2 and MBD3 prompted the reinvestigation of whether MBD3 is a methyl-CpG binding protein. MBD3 was expressed in *E.coli* as a His-tagged protein. The purified protein was used for bandshift experiments using methylated and non-methylated version of CG11 (0.1 ng of labeled probe). A titration experiment with increasing concentration of protein in the absence of competitor DNA demonstrated that MBD3 has a preference for methylated DNA (Figure 4.2.11 A). The mobility of the DNA/protein complex at a particular concentration was different between the methylated and non-methylated probes. A change in detergent used in the bandshift assays from digitonin (1 %) to Triton X-100 (0.1 %) did not alter the DNA binding property of MBD3 (data not shown).

Increasing amounts of non-methylated *E.coli* competitor DNA included in the bandshift assay also lead to the preferential formation of complex with the methylated probe (Figure 4.2.11 B). Under certain conditions (e.g. 100 ng and 200 ng of competitor DNA), only the methylated probe was shifted by MBD3. A similar conclusion is reached that MBD3 has a preference for methylated DNA *in vitro*.

In collaboration with Dr. Danny Reinberg (New Jersey), we studied the DNA binding property of MBD3 expressed and purified from baculovirus infected insect cells. Two versions of MBD3 (a and b) were tested for binding to methylated DNA (Figure 4.2.11 C). MBD3a contains the full MBD-like region, and MBD3b is a splice variant with only half of the MBD-like region (Zhang et al, 1999). Initial bandshift experiment was carried out with increasing amounts of protein and in the absence of competitor DNA (Figure 4.2.11 D). At certain protein concentration, MBD3a bound both the methylated and non-methylated CG11 probes. However, similar to the result obtained from *E.coli* expressed protein, MBD3a exhibited a preference for methylated DNA. For example, at 100 ng of MBD3a, only the methylated CG11 probe was shifted.

These series of experiments showed that MBD3 has a preference for methylated DNA *in vitro*. In addition, the MBD-like region is important for selectivity as MBD3b, which does not have the full MBD-like region did not bind to DNA.



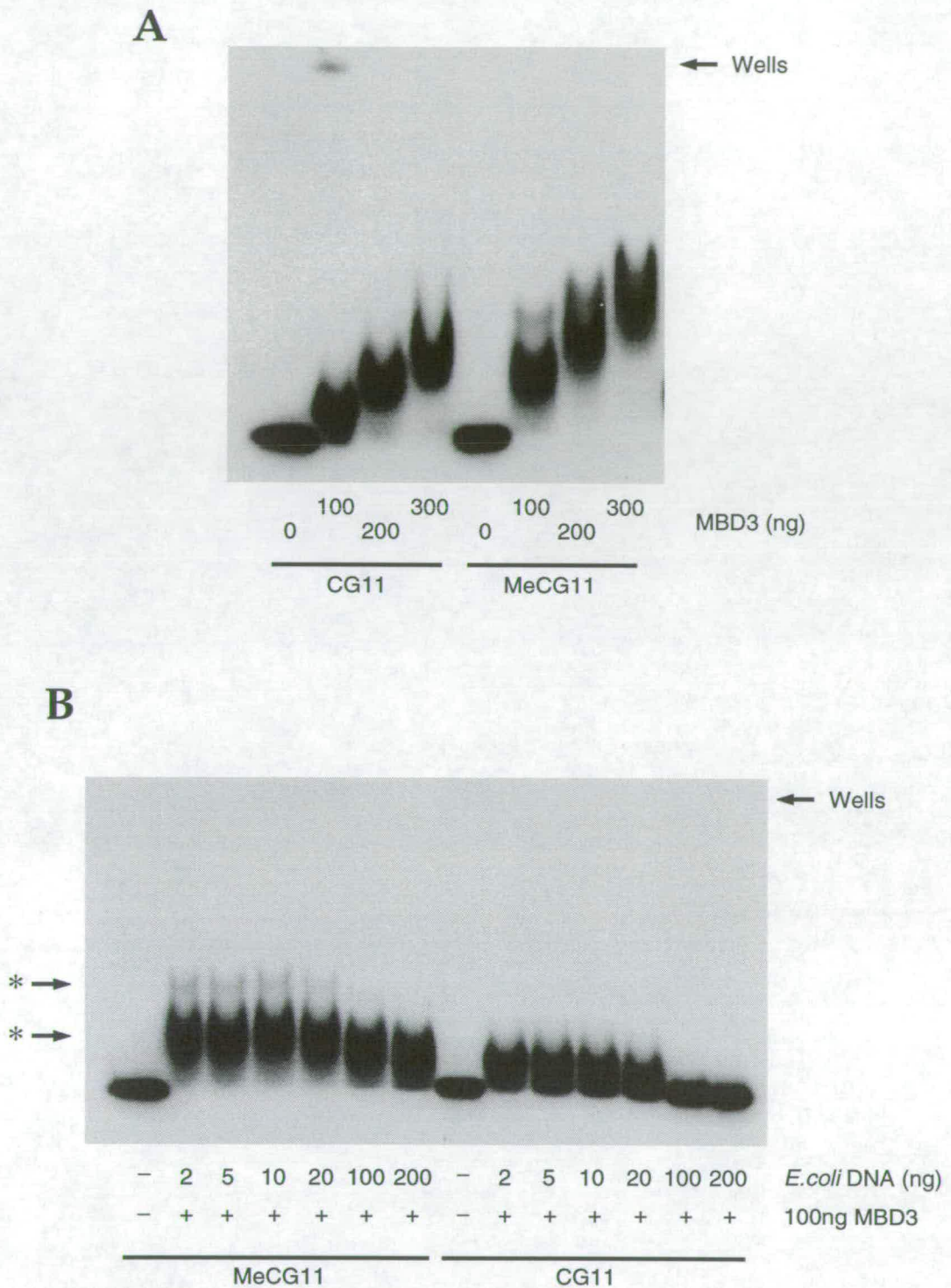
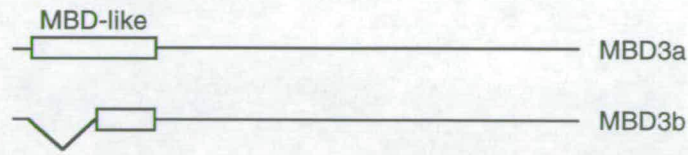


Figure 4.2.11. Methyl-CpG binding property of MBD3

A) Increasing amounts of recombinant MBD3 (expressed in *E.coli*) were incubated with either non-methylated or methylated CG11 probes in the absence of competitor DNA. The DNA/protein complexes were resolved on 1.5 % agarose gel. The position of the wells is indicated.

B) Increasing amount of *E.coli* DNA as added to 100 ng of MBD3 before incubation with either non-methylated or methylated CG11 probes. The asterisks indicate DNA/MBD3 complexes.

C



D

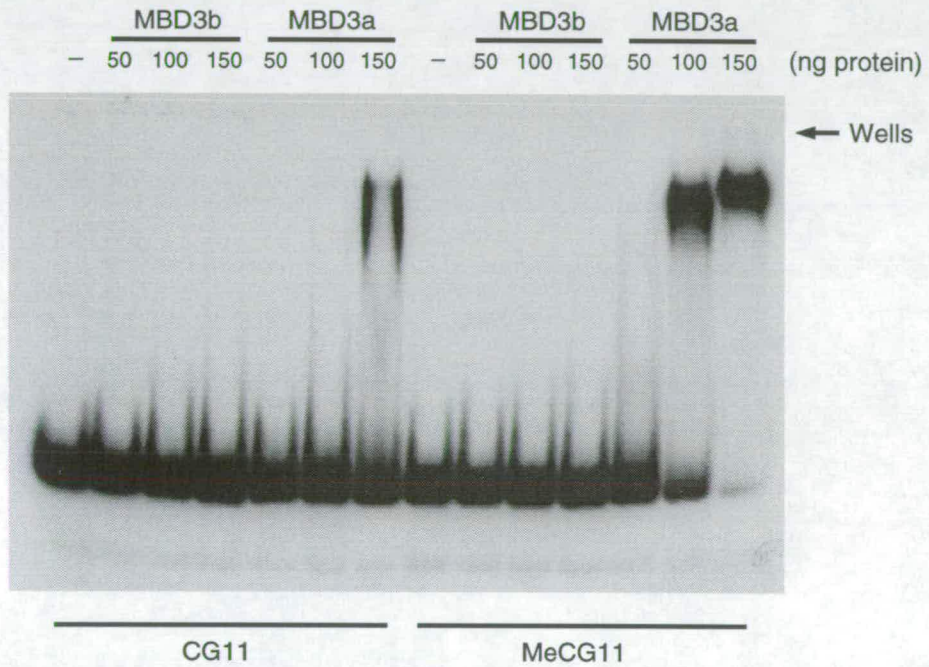


Figure 4.2.11. Methyl-CpG binding property of MBD3

C) Two variants of MBD3 used in bandshift experiments shown in D. MBD3a contains a MBD-like region shown as a box. MBD3b contains C-terminal half of the MBD-like region.

D) Increasing amount of MBD3a or MBD3b (expressing and purified from baculovirus infected insect cells) was incubated with either non-methylated or methylated CG11 probes in the absence of competitor DNA. The DNA/protein complexes were resolved on 5 % polyacrylamide gel. The position of the wells is indicated.



The specificity of the MBD3a/ MeCG11 complex was next tested by competition experiments with methylated and non-methylated competitor DNA. It was found that the MBD3a/ MeCG11 complex is very sensitive to competition by non-methylated DNA, as the DNA binding activity was completely abolished by the addition of 100 ng of *E.coli* DNA (Figure 4.2.11 E, compare lane 1 and lane 3). When the bandshift reactions were challenged with equal amount of non-methylated or methylated DNA, no clear difference was evident (Figure 4.2.11 E, compare lanes 4 &5, 6&7, 8 & 9, 10 & 11, 12 & 13, 14 & 15). It is concluded that the specificity of MBD3 towards methylated DNA is weak.

As MBD3 is identified as a component of the Nucleosome Remodelling and histone Deacetylation complex (NuRD) (Zhang et al, 1999; Wade et al, 1999). A more relevant question is whether the native NuRD complex has an affinity for methylated DNA. Under conditions with no non-methylated competitor DNA, the NuRD complex bound both the methylated and non-methylated DNA (Figure 4.2.11 F, lanes 3 and 6). Addition of non-methylated competitor DNA (300 ng) abolished the DNA/protein complexes (see Figure 4.2.11 H, lane 4). Therefore, no methylated DNA binding specificity was detected for the NuRD complex. In addition to MBD3, Wade et al (1999) showed that a MTA1-like protein in the *Xenopus* NuRD complex has an affinity for methylated DNA using a Southwestern assay. The human homolog of MTA1-like protein, MTA2 bound both the non-methylated and methylated probe, indicating that MTA2 binds non-specifically to DNA independent of methylation (Figure 4.2.11 G).

Although MBD2 is not a component of the NuRD complex, it can however interact with the NuRD complex *in vitro* as revealed by GST pulldown assay (Zhang et al, 1999). Incubation of GST-MBD2a with the purified NuRD complex (300 ng) also resulted in the formation of a DNA/ protein complex with lower mobility (Figure 4.2.11 H, compare lanes 2 and 3). Under the same assay condition, the addition of 3 µg of BSA (10 fold more than the amount of NuRD used) did not alter the mobility of the DNA/ protein complex (Figure 4.2.11 H, compare lanes 6 and 7).

### **4.3 Conclusion**

Several antibodies were raised against MBD2. Antibody R593 cross-reacts with a related protein MBD3, which did not show a marked preference for methylated DNA *in vitro* and *in vivo* (Hendrich and Bird, 1998). As MBD3 is a component of the NuRD complex, the

E

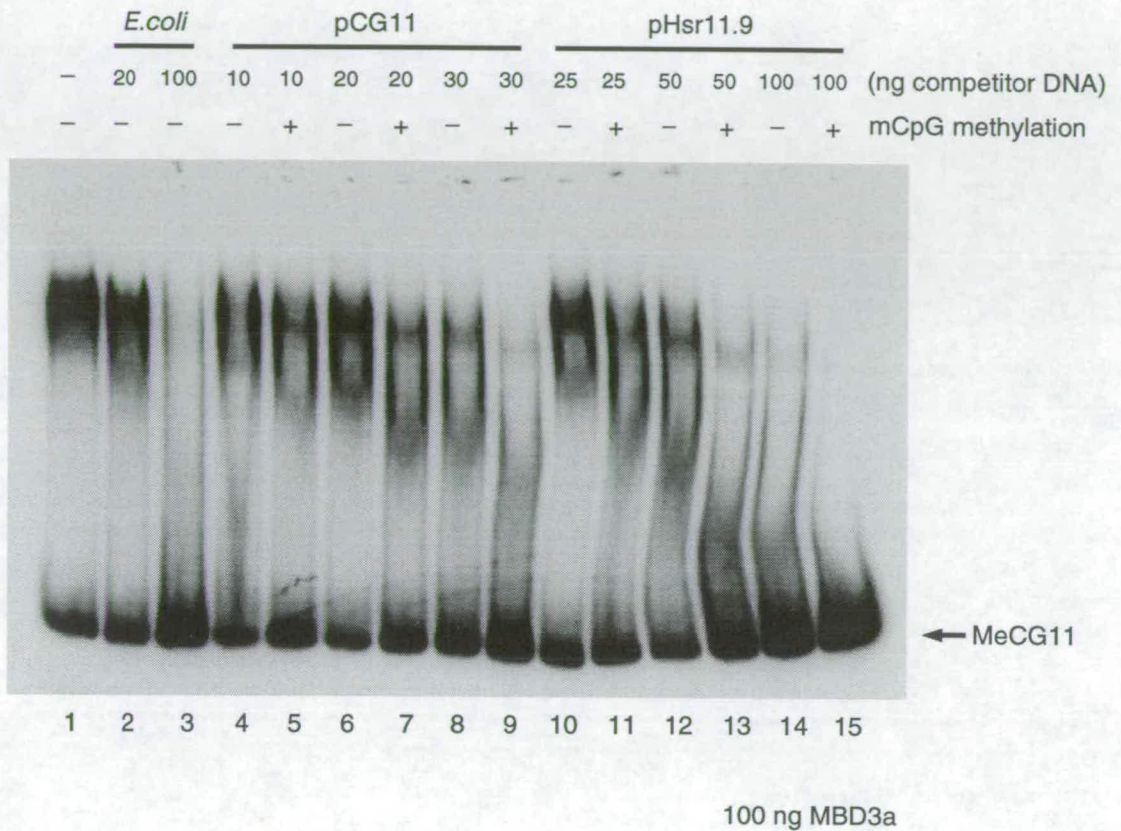


Figure 4.2.11. Methyl-CpG binding property of MBD3

E) The MeCG11/MBD3 complex was challenged with non-methylated or methylated competitor DNA (lanes 2 to 15). pCG11 is a pUC19 plasmid containing a CG11 insert. pHsr11.9 is a pUC9 with a 11.9 kb human rDNA EcoRI fragment. The plasmids were methylated at all CpG sites using SssI methylase. The DNA/protein complexes were resolved on 5 % polyacrylamide gel.

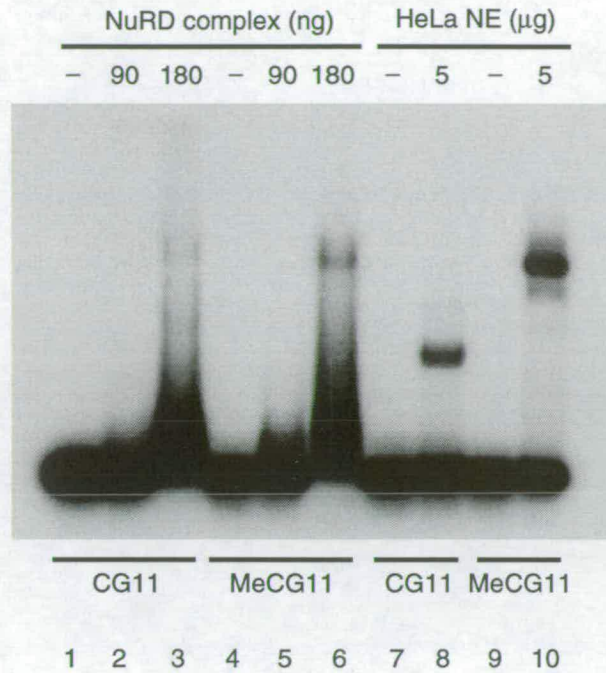
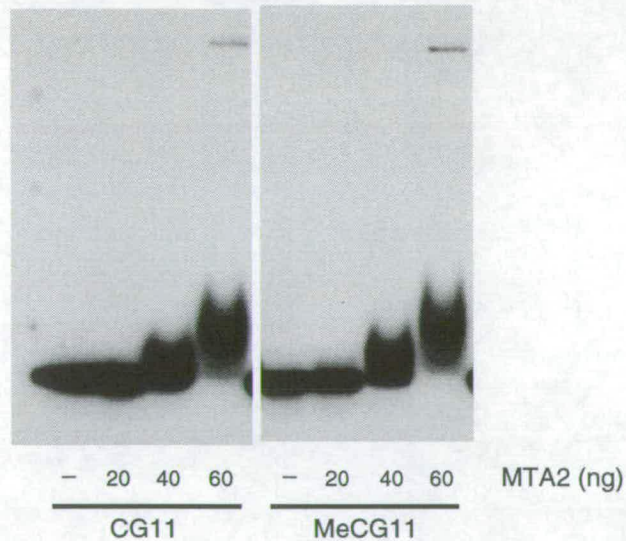
**F****G**

Figure 4.2.11. Methyl-CpG binding property of NuRD complex and MTA2

F) Increasing amounts of purified NuRD complex were incubated with non-methylated or methylated CG11 probes in the absence of competitor DNA. The DNA/protein complexes were resolved on 1.5 % agarose gel. HeLa nuclear extracts were included as controls.

G) Increasing amounts of recombinant MTA2 (expressed in *E.coli.*) were incubated with either non-methylated or methylated CG11 probes in the absence of competitor DNA. The DNA/protein complexes were resolved on 1.5 % agarose gel.





R593 antibody was not used in studies addressing MBD2 interacting proteins. Although the two other antibodies (S752 and S923) can discriminate between MBD1, MBD2, and MBD4, only antibody S923 can efficiently recognise native MBD2 from HeLa nuclear extracts as revealed by immunoprecipitation assay. Antibody S752 recognised native MBD2 poorly, suggesting that the C-terminal epitopes are not exposed. These two antibodies were tested for the ability to interact with the methyl-CpG binding activity in HeLa cells, MeCP1. Antibody S923 can supershift and immunodeplete MeCP1 activity in HeLa cells. Antibody S752 did not affect MeCP1 activity under similar assays. A correlation between the ability of these antibodies to recognise native MBD2 and interact with MeCP1 was observed, and the result suggests that MBD2 is a component of MeCP1. In addition, S923 antibody was able to supershift MeCP1 activity from nuclear extracts derived from rat brain, kidney, liver and a human lymphoblastoid cell-line, DD7.

Co-immunoprecipitation experiments using S923 antibody indicate that MBD2 is associated with HDAC1, HDAC2 and RbAp48. These proteins are common subunits of histone deacetylase complexes found in mammalian cells (Hassig et al, 1997; Laherty et al, 1997; Zhang et al, 1997; Zhang et al, 1998). In addition, S923 antibody can immunoprecipitate histone deacetylase activity from HeLa cell extracts. Reciprocal immunodepletion and immunoprecipitation experiments using anti-HDAC1 antibody confirmed the interaction. Anti-HDAC1 antibodies were able to immunodeplete both MBD2 and MeCP1. Affinity purified HDAC1 containing complexes also contained MBD2. Based on the work with our collaborators (Drs Yi Zhang and Danny Reinberg), MBD2 is not found in previously characterised mSin3 or Mi2/ NuRD histone deacetylase complexes. Using GST-pulldown assay, MBD2 was shown to interact weakly with HDAC1, HDAC2 and RbAp48. Interestingly, MBD2 can also interact with itself under the same assay. Homodimerisation may provide the MBD2 complex with extra binding affinity for methylated DNA. Consistent with the finding that MBD2 exists in a multiprotein complex, the native size of MBD2 as revealed by gel filtration chromatography is in the range of 800 kDa.

GAL4-MBD2 can repress transcription from reporter constructs containing Gal4 binding sites upstream of human DNA polymerase  $\beta$  promoter or human  $\beta$ -actin promoter. Repression on the human DNA polymerase  $\beta$  promoter is sensitive to TSA, indicating that deacetylation is involved in repression of this promoter. Repression on the human  $\beta$ -actin promoter was not relieved by TSA, suggesting that MBD2 can repress transcription through deacetylation independent pathway. It is notable that repression on the two

promoters was not strong. It is possible that chromatinised template is required for full repression activity.

Interestingly, it is reported that a truncated version of MBD2 (MBD2b) contains a 5-meC demethylase activity (Bhattacharya et al., 1999). MBD2b was shown to catalyse the removal of methyl group from 5-meC. Unfortunately, this novel activity cannot be reproduced by two laboratories (Ng et al, 1999; Wade et al, 1999).

As MBD3 is identified as a component of the *Xenopus* and mammalian NuRD complex (Zhang et al, 1999; Wade et al, 1999), the issue of whether MBD3 is a methyl-CpG binding protein was re-investigated. Bandshift experiments with MBD3 expressed from *E.coli* using a densely methylated probe showed that MBD3 has a preference for methylated DNA. The bandshift experiments were also repeated using affinity purified MBD3 derived from baculovirus infected insect cells. Full-length MBD3 (MBD3a) bound preferentially to methylated probe in the absence of competitor DNA. However, the non-methylated probe was also shifted in the presence of more protein. A shorter version of MBD3 with half the MBD-like domain deleted did not bind to either the non-methylated or methylated probes. Both the full-length MBD3 expressed in *E.coli* and insect cells showed preference for methylated DNA. It should however be emphasised that the methylated DNA/ protein complex is very sensitive to competition with non-methylated DNA. Furthermore, addition of non-methylated or methylated DNA did not yield a specific DNA/ protein complex.

The DNA binding activity of mammalian NuRD complex was also studied. The native NuRD complex shifted both the non-methylated and methylated probe in the absence of competitor DNA. Addition of non-methylated competitor DNA abolished the DNA /protein complex. A component of the NuRD complex, MTA2 also bound to DNA in a methylation independent manner. Therefore, unlike the *Xenopus* MBD3 and MTA1-like protein, the mammalian homologues do not bind specifically to methylated DNA.

GST-MBD2 can pull-down all the subunits of the NuRD complex (Zhang et al, 1999). In addition, the NuRD complex can 'supershift' the MBD2/ DNA complex. The result indicate that MBD2 can interact with the Mi2 containing NuRD complex *in vitro*. It should however be emphasised that there is currently no evidence that the interaction between MBD2 and NuRD can occur *in vivo*. The result can also be explained by the possibility

that MBD2 interacts directly with certain subunits of the NuRD complex, e.g. HDAC1, HDAC2 or RbAp48.

## **Chapter 5: MBD1 is a transcriptional repressor**

### **5.1 Introduction**

MBD1 (formerly known as PCM1) was originally identified by searching the EST databases for proteins with methyl CpG binding domain of MeCP2. The N-terminal of MBD1 contains a MBD-like region (Cross et al, 1997). In addition, this protein has two cysteine-rich regions (termed the CxxC motifs) also found in DNA methyltransferase protein 1 (DNMT1) and mammalian trithorax-like protein HRX. The *HRX*, also known as *MLL* (mixed lineage leukemia) and *ALL1* (acute lymphoblastic leukemia) is frequently translocated in acute leukemia. A region encompassing the CxxC motif in DNMT1 has been shown to bind zinc (Bestor, 1992; Chuang et al, 1996).

An antibody raised against MBD1 was able to supershift the MeCP1 activity, this suggests that MBD1 is a component of the long sought MeCP1 (Cross et al, 1997). When an *in vitro* transcription extract was supplemented with recombinant MBD1, the transcription from a methylated template was preferentially repressed. The result from the *in vitro* experiment suggests that MBD1 is a mediator of transcriptional repression by DNA methylation.

The work described in this chapter concerns the functional characterisation of the methyl-CpG binding protein, MBD1.

### **5.2 Results and discussion**

#### **5.2.1 Full-length MBD1 binds selectively to methylated DNA *in vitro***

MBD1 is susceptible to degradation when expressed in *E.coli* (Cross et al, 1997; Hendrich and Bird, 1998). When the N-terminal His-tagged MBD1 was purified over a Ni<sup>2+</sup>-NTA column, the majority of the eluted proteins contained only the N-terminal fragments. It was found that a combination of affinity purification of C-terminal His-tagged MBD1 and strong cation-exchange chromatography can preferentially enrich for full-length MBD1 (Figure 5.2.1 A). The first step of purification over a Ni<sup>2+</sup>-NTA column yielded three major products. The two smaller proteins of approximately 70 kDa are likely to be degraded MBD1 lacking the basic N-terminal parts. When the eluate was applied to a strong cation-exchange Fractogel EMD S0<sub>3</sub><sup>c</sup>-650 (M) column at 0.3 M NaCl,



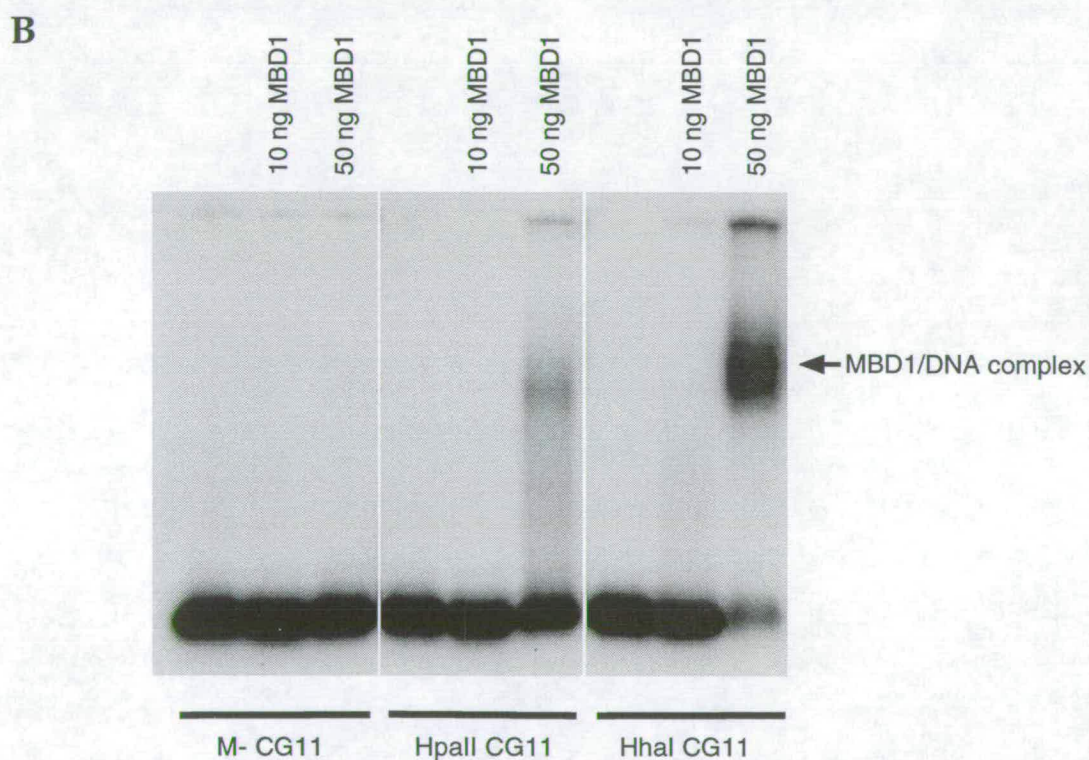
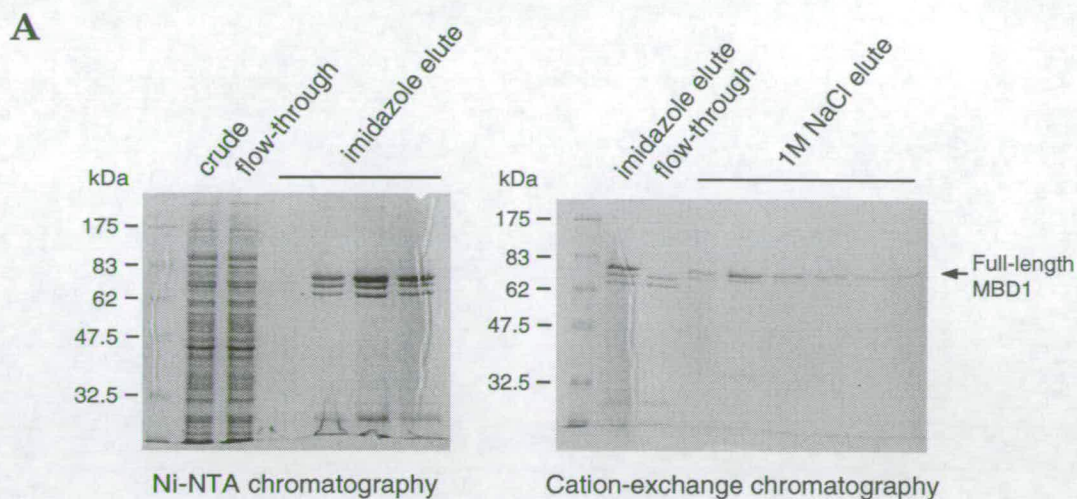


Figure 5.2.1 Affinity of intact MBD1 for methylated DNA *in vitro*

A) Purification of C-terminal tagged MBD1 over Ni<sup>2+</sup>-NTA and cation exchange column.

B) Full-length MBD1 binds specifically to methylated probes. HpaII methylated CG11 probe contains 7 methyl-CpG pairs. HhaI methylated CG11 probe contains 20 methyl-CpG pairs. M- probe was mock-methylated.

only the full-length protein was retained. The eluted protein was judged to be 95 % pure by Coomassie blue staining and Western blotting (Figure 5.2.1 A and 5.2.3 A).

Bandshift assays were carried out using the full-length MBD1. MBD1 bound to the HpaII methylated and HhaI methylated CG11 probes, but not the non-methylated probe (Figure 5.2.1 B).

## **5.2.2 The N-terminal of MBD1 is sufficient for binding to methylated DNA**

The MBD-like region in MBD1 shares a 43.6 % identity with MBD of MeCP2 over a 55 amino-acids sequence (Figure 5.2.2 A). Since the minimal region in MeCP2 required for specific interaction with methyl-CpG pair spans 85 amino-acids (Nan et al, 1993), it is curious to test whether this homologous region is sufficient for specific binding to methylated DNA. The first 67 amino-acids of MBD1 was expressed as a GST-fusion protein. The purified recombinant protein bound specifically to 17 mer AB17 probe with a single methyl-CpG pair, and not to the control non-methylated or hemi-methylated probes (Figure 5.2.2 B).

Proline at the 38 position is conserved between the MBD of different methylated DNA binding proteins (MeCP2, MBD1, MBD2, and MBD4). Using PCR-directed mutagenesis, the codon coding for proline was mutated to encode alanine or lysine. The purified GST-MBD1 (1-67 aa) with these amino-acid substitutions failed to bind methylated DNA as demonstrated using bandshift assay (Figure 5.2.2 C). The result indicates that the N-terminal 67 amino-acids of MBD1 is a methyl-CpG binding domain. The NMR structure of the MBD of MeCP2 has been solved, and the corresponding proline in MBD of MeCP2 was found at a sharp turn between two strands of beta sheets (Wakefield et al, 1999). The amino acids around this proline did not undergo a chemical shift upon DNA binding, suggesting that it is not in contact with DNA. Conversion of the proline to alanine or lysine is likely to perturb the local structure of the MBD by removing the unique chain bending property of proline.



**A**

MBD1 7 DCPALGPGWKRREVFRKSGATCGRSDTYYS**P**TGDRIRSKVELTRYLGPACDLTL----FDF 64  
 D P L G W R+ R K S G + G+ D Y +**P** G R S K V E L Y D + L F D F  
 MeCP2 96 DDPTLPEGWTRKRLKQRKSGRSAGKYDVYLIN**P**Q GKAFRSKVELIAYFEKVGDTSLDPNDFDF 157

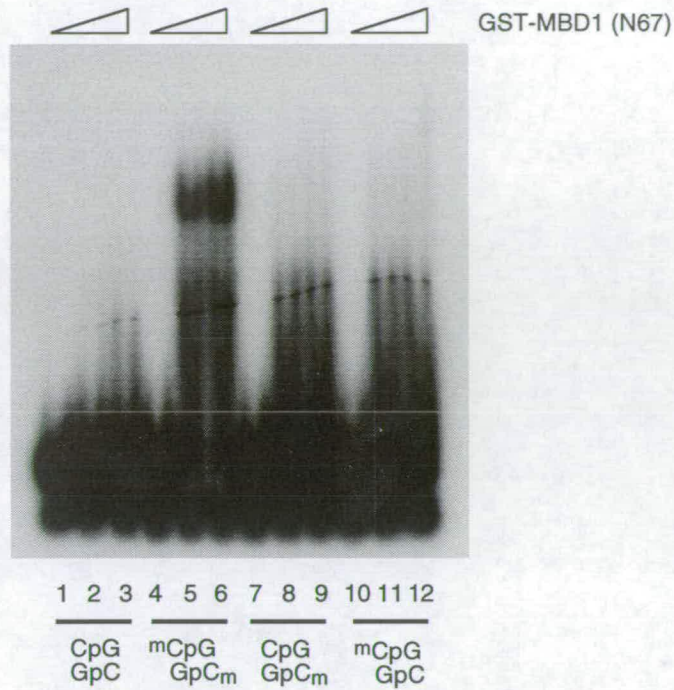
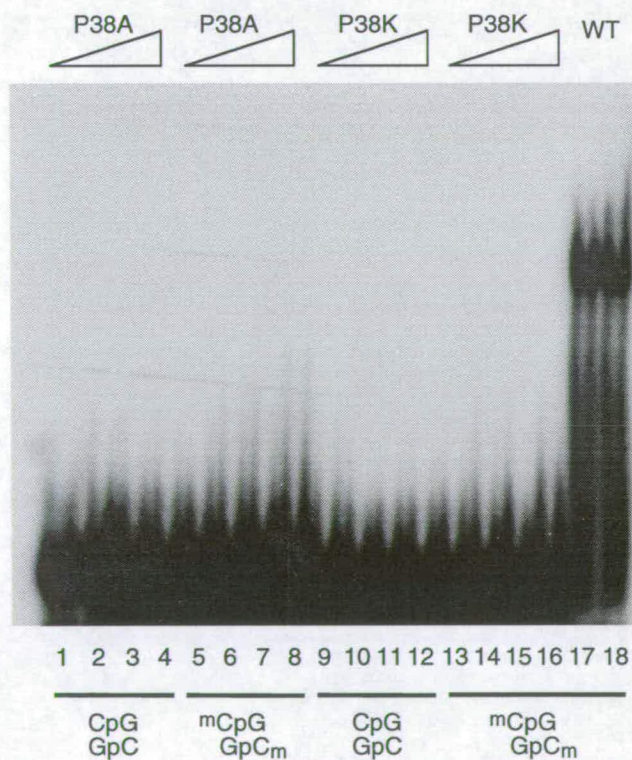
**B****C**

Figure 5.2.2. Methyl-CpG binding domain of MBD1

A) Sequence comparison between MBD of MeCP2 and MBD-like region in MBD1. A conserved proline is highlighted in bold.

B) The N-terminal 67 amino-acids of MBD1 binds specifically to a symmetrical methyl-CpG pair.

C) Mutation in proline 38 abolished methyl-CpG binding.

### 5.2.3 Characterisation of an antibody raised against MBD1

To facilitate the functional studies of MBD1, an antibody was raised against a GST-fusion of MBD1 (spanning amino acids 351 to 556) in sheep. The C-terminal fragment of MBD1 was chosen to minimise cross-reactivity towards other MBD or CxxC domain-containing proteins. The S571 immune antiserum is specific for MBD1 and does not cross-react with MBD2b or MBD4. When a Western blot of HeLa nuclear extracts was probed with this antibody, two bands were detected (the bands correspond to 83 and 75 kDa; see Figure 5.2.3 A, left panel). Two lines of evidence indicate that the two bands are specific for MBD1. First, the preimmune serum did not detect these proteins (data not shown). Second, the specific antibodies in the immune antiserum can be depleted by pre-incubation with GST-MBD1 coupled to GSH-sepharose beads (Figure 5.2.3 A, left panel). GST-GSH-sepharose beads did not deplete the MBD1 signal (Figure 5.2.3 A, right panel). The two proteins of different molecular weight are possibly spliced variants of MBD1 (see Figure 5.2.7 A for two spliced versions: MBD1 and MBD1\*). It is very likely that the 83 kDa MBD1 protein contains 3 CxxC motifs while the 75 kDa protein contains only 2 CxxC motifs. An additional CxxC motif of about 68 amino-acids can account for an increase of 5 kDa in size. mRNA encoding for 2 or 3 CxxC motifs can be detected in HeLa cells by RT-PCR (data not shown). Affinity purified antibody also detected the same bands as the crude antiserum, therefore the 83 and 75 kDa bands are MBD1-specific (Figure 5.2.3 B). Furthermore, the antiserum only recognised MBD1 but not MBD2b or MBD4 (Figure 5.2.3 C).

Beside specificity, the antibody was also examined for sensitivity. Different amounts of His-tagged antigen (amino acids 351 to 556 of MBD1) were blotted and challenged with either preimmune or immune antisera (Figure 5.2.3 D). The immune antiserum detected even 500 picograms of antigen.

Next, the S751 antibody was tested for its ability to recognise native MBD1. Anti-MBD1 antibodies, together with control preimmune antibodies, was used to immunoprecipitate nuclear proteins from HeLa nuclear extracts. The presence of MBD1 in the anti-MBD1 immunoprecipitate indicates that the antibody can recognise native MBD1 (Figure 5.2.3 E).

The S751 antiserum was also tested on rat liver nuclear extract (containing both MeCP1 and MeCP2), and mouse F9 embryonic carcinoma nuclear extract (deficient in MeCP

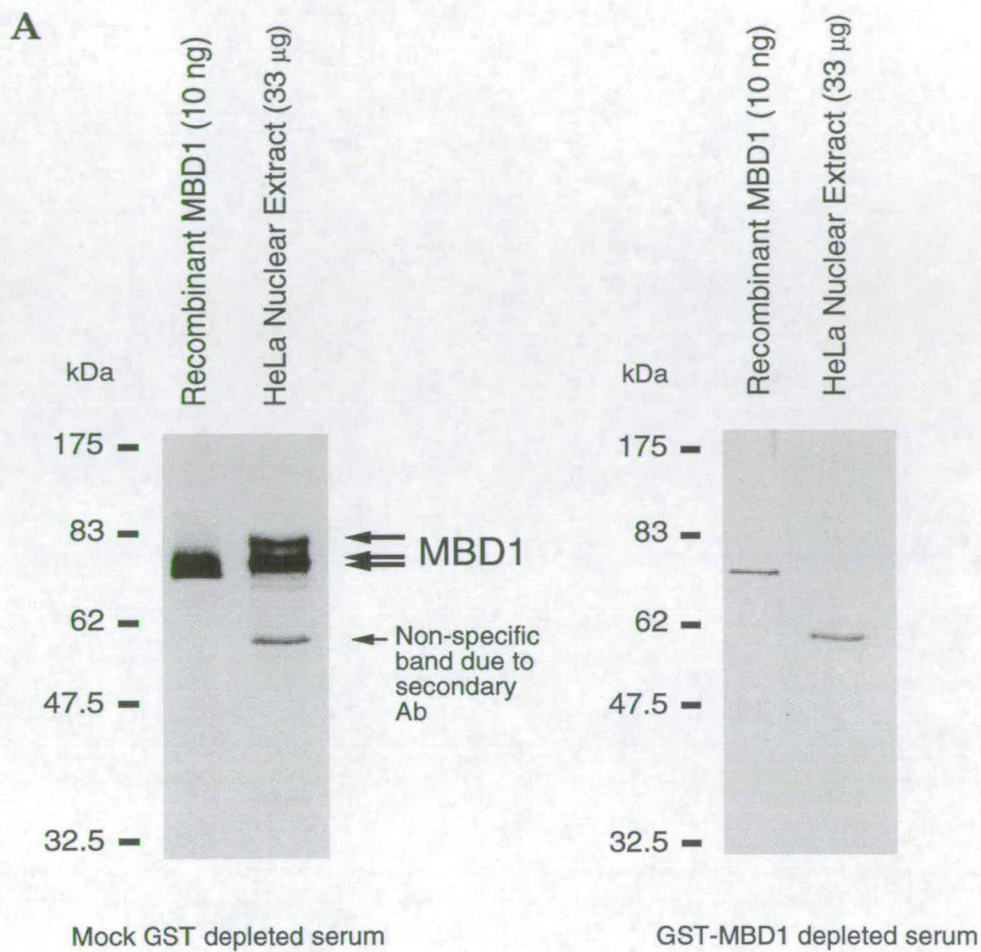


Figure 5.2.3.A. Western blot of HeLa nuclear extract with anti-MBD1 S751 antiserum.

Left panel: the antiserum was preincubated with GST. MBD1 specific signals are indicated. The non-specific band was due to cross-reactivity with secondary anti-sheep antibody.

Right panel: the antiserum was preincubated with GST-MBD1 to deplete MBD1 specific antibodies. The residual band in recombinant MBD1 lane is probably a co-purified bacterial protein.



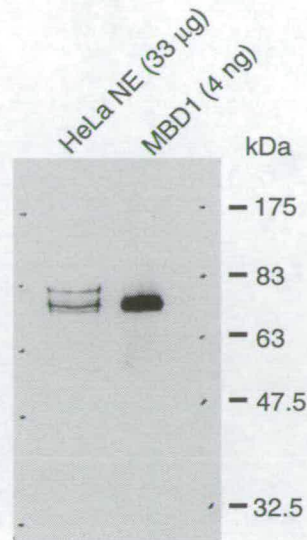
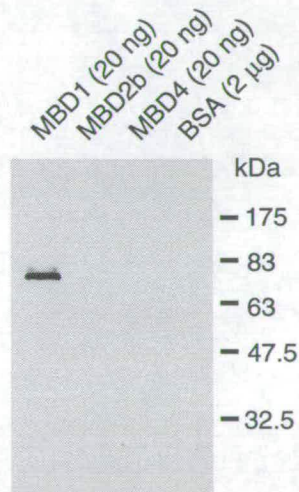
**B****C**

Figure 5.2.3. Specificity of S751 antiserum

B) Affinity purified anti-MBD1 antibody detected both 83 and 75 kDa nuclear proteins. HeLa NE is HeLa nuclear extract.

C) S751 antiserum reacted only with recombinant MBD1, but not MBD2b, MBD4 or BSA.

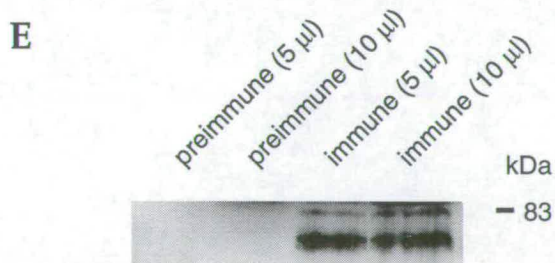
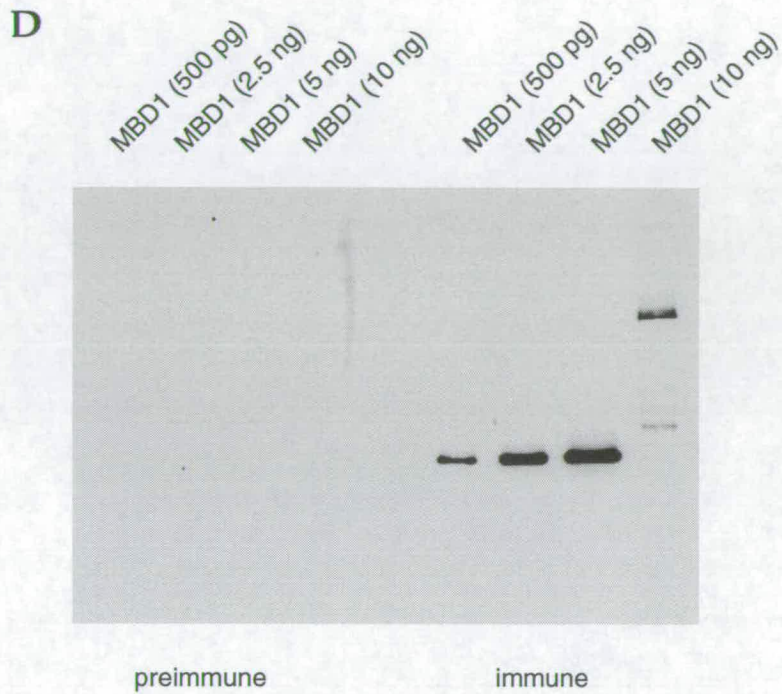


Figure 5.2.3. S751 antiserum

D) S751 antiserum can detect picograms of antigen. Preimmune serum was used as a control.

E) S751 antiserum can recognise native MBD1 from a HeLa nuclear extract. Both preimmune and immune S751 sera were used to immunoprecipitate nuclear proteins from HeLa nuclear extracts. Western blot of the immunoprecipitates were probed with S751 antiserum.

**F**

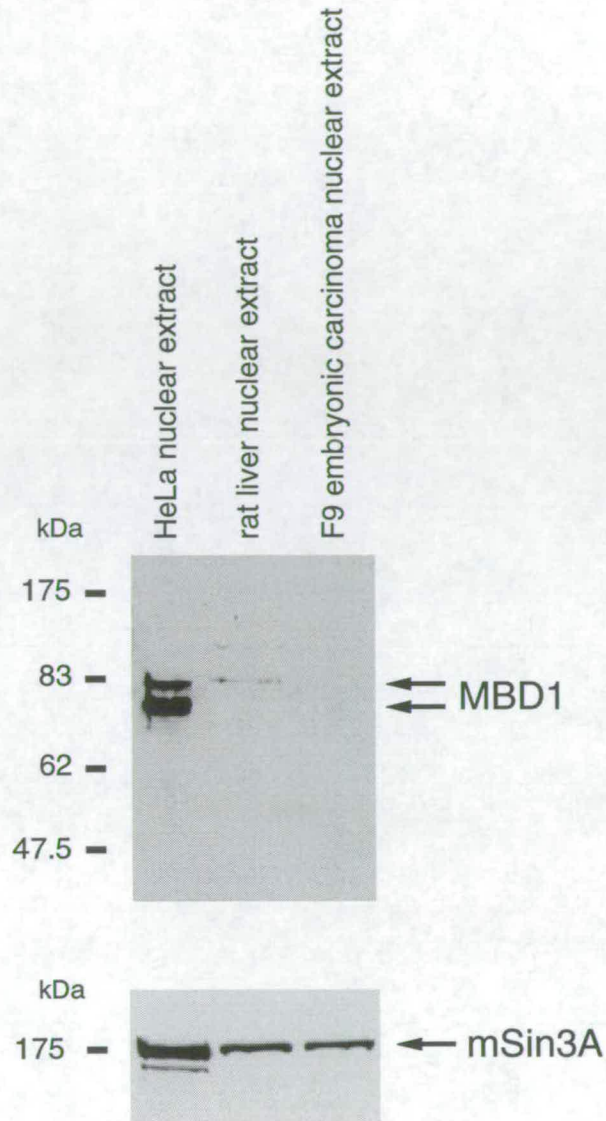


Figure 5.2.3. S751 antiserum

F) Anti-MBD1 antibody cross-reacts poorly with murine MBD1.

Western blot of HeLa, rat liver, mouse F9 EC nuclear extracts were probed with S751 antiserum. The blot was also probed with anti-mSin3A (K20) antibodies as a control for loading.



bandshift activity). The antiserum detected a band of 83 kDa for rat liver nuclear extract, albeit very weakly (Figure 5.2.3 F). Since the conservation in amino acid sequences between the human and mouse MBD1 is not high (55 % identity and 61 % similarity over the region of 206 amino-acids used to raise the S751 antibody), it is suspected that the antibody was not efficient in recognising the murine protein, and the weak signal is unlikely due to the lack of MBD1 in the murine tissue extract.

#### **5.2.4 MBD1 is not a component of MeCP1**

The discovery of three novel methyl-CpG binding proteins (MBD1, MBD2 and MBD4) raises the possibility that these proteins may contribute to the MeCP1 activity. Previous work using an antibody raised against full-length MBD1 was able to supershift HeLa MeCP1 activity, suggesting that MBD1 is a component of MeCP1 (Cross et al, 1997). With a specific antibody against MBD1, the relationship between MBD1 and MeCP1 is studied.

Immunodepletion experiments described in Chapter 4 (section 4.2.2) was carried out. The HeLa nuclear extract after immuno-absorption with MBD1 antibodies had reduced level of MBD1, but the MeCP1 bandshift activity was not affected (Figure 5.2.4 A and B). The non-specific complex on the non-methylated probe was not affected by similar treatment.

Antibody “supershift” experiments were also carried out. Addition of preimmune or immune sera (1  $\mu$ l) to HeLa nuclear extracts slightly altered the mobility of MeCP1 complex as compared to bandshift reaction with no added serum (Figure 5.2.4 C, compare lanes 8, 9 and 11). When half the amount of serum (0.5  $\mu$ l) was added to the bandshift reactions, no alteration in mobility of MeCP1 was observed (Figure 5.2.4 C, compare lanes 8, 10, 12). No difference between the effects of preimmune and immune sera on MeCP1 was detected.

MBD2 was demonstrated to be a component of MeCP1 (based on work described in Chapter 4). Therefore, if MBD1 is also a component of MeCP1, MBD1 and MBD2 should co-exist in a same complex. Immunoprecipitation experiments revealed that MBD1 is not associated with MBD2, as MBD2 was not detected in MBD1 immunoprecipitate and MBD2 immunoprecipitate did not contain MBD1 (Figure 5.2.4 D).

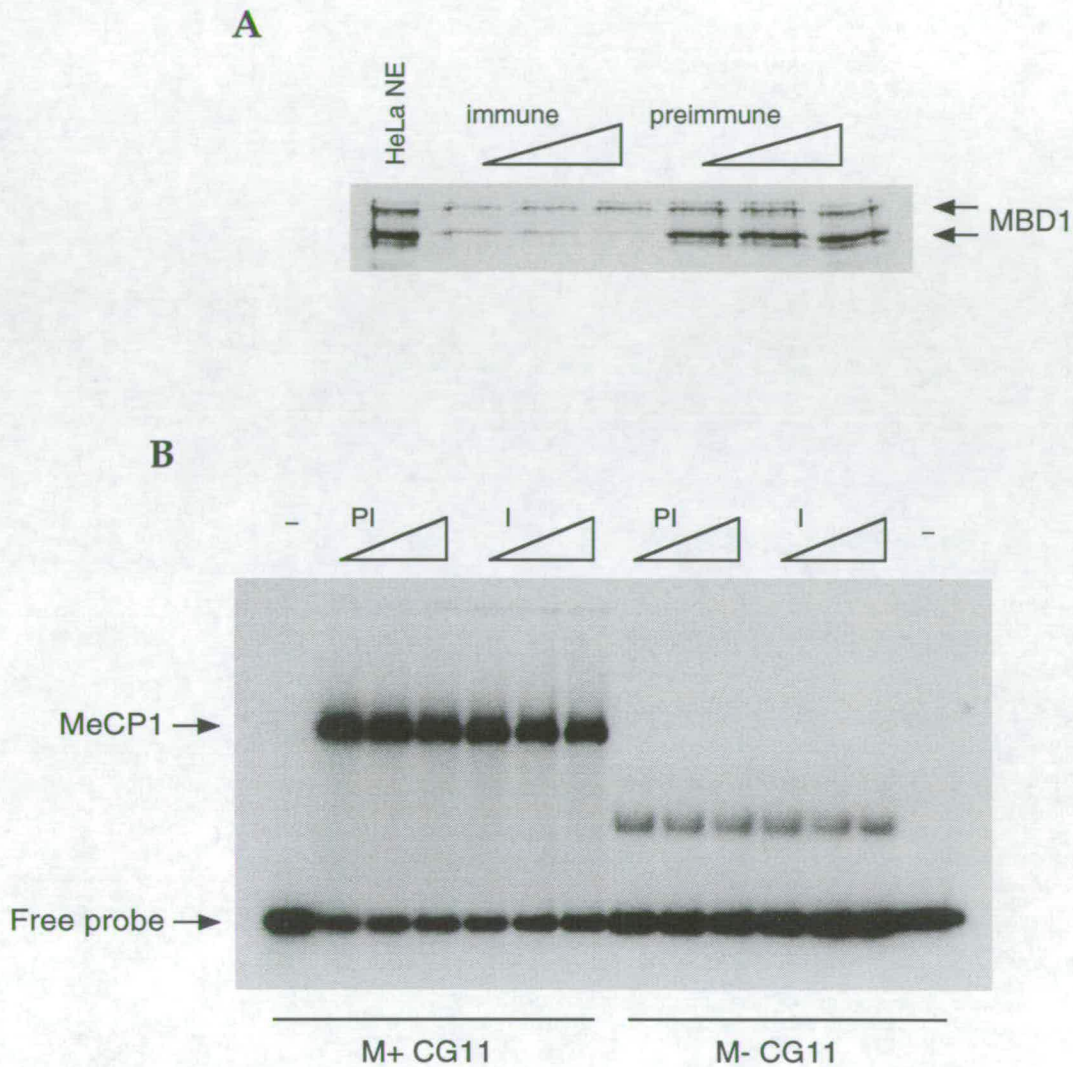


Figure 5.2.4. MBD1 and MeCP1

A) Immunodepletion of MBD1 with anti-MBD1 antibodies.

HeLa nuclear extracts were incubated with either preimmune or immune S751 sera. The unbound extracts were separated from the antibodies and analysed on a Western blot to detect MBD1. HeLa NE is non-treated HeLa nuclear extract. 1  $\mu$ l, 5  $\mu$ l, and 10  $\mu$ l of sera were used. Depletion of the 83 kDa MBD1 was not efficient.

B) Reduction in MBD1 does not lead to loss of MeCP1.

The immunodepleted extracts were analysed for MeCP1 bandshift activity. PI is preimmune antibodies depleted extract. I is anti-MBD1 antibodies depleted extract. M+ CG11 is CG11 probe methylated at 27 methyl-CpGs. M- CG11 is mock-methylated probe. No nuclear extract was added in the first and last lanes (lane -).

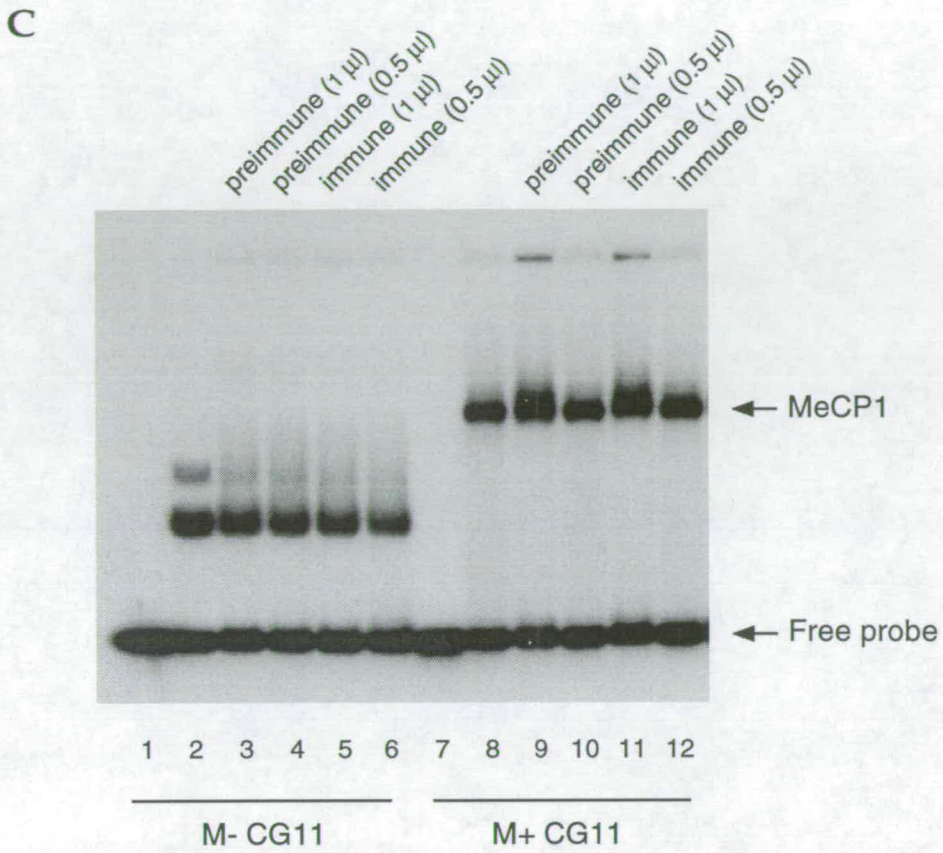


Figure 5.2.4. MBD1 and MeCP1

C) Anti-MBD1 antibodies do not supershift MeCP1 activity.

Different amount of preimmune and immune S751 sera were added to bandshift reactions as indicated. No serum was added to lanes 2 and 8. No nuclear extract was added to lanes 1 and 7. CG11 is CG11 probe methylated at 27 methyl-CpGs. M- CG11 is mock-methylated probe.



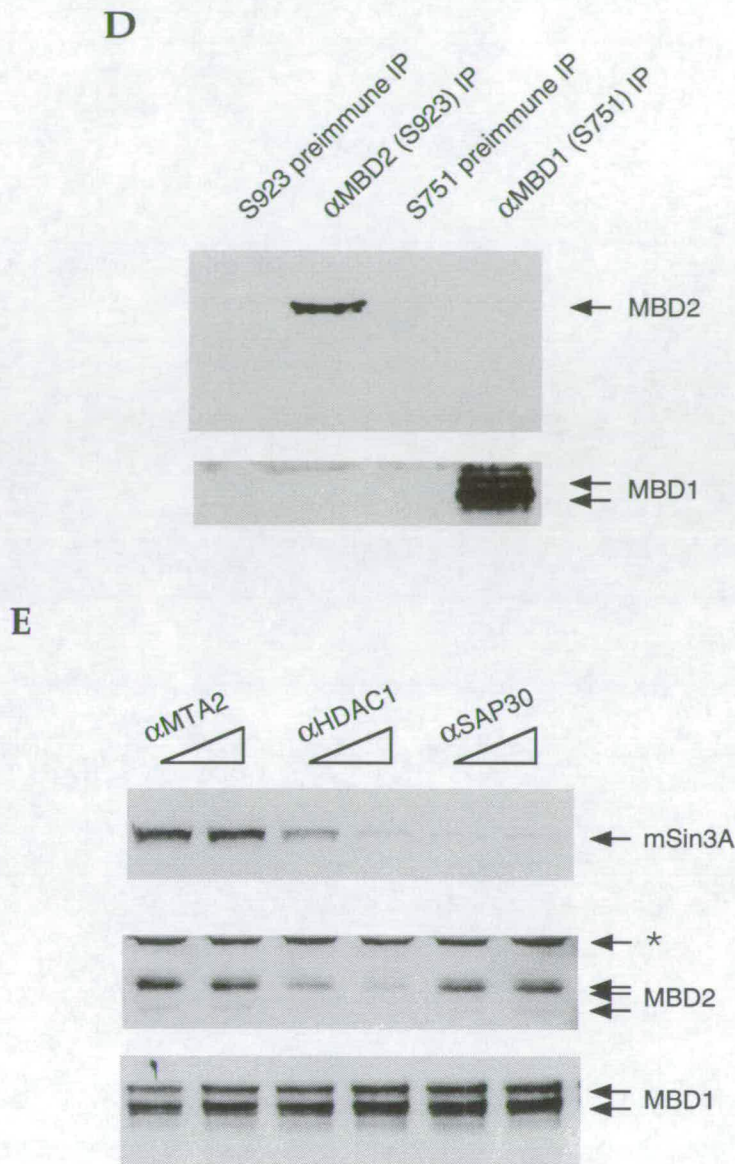


Figure 5.2.4. MBD1 and MeCP1

D) MBD1 and MBD2 are not associated in the same complex.

Immunoprecipitates of S923 (preimmune or immune anti-MBD2) and S751 (preimmune or immune anti-MBD1) were probed with anti-MBD1 or anti-MBD2 antibodies. MBD1 was not immunoprecipitated by anti-MBD2 antibodies, and vice versa.

E) MBD1 is not associated with HDAC1 complexes.

Nuclear extracts were immunodepleted with anti-MTA2 (component of NuRD complex), anti-HDAC1, and anti-SAP30 (component of mSin3 complex) antibodies. Anti-MTA2 antibodies did not deplete mSin3A, MBD1 or MBD2. Anti-HDAC1 antibodies depleted both mSin3A and MBD2. Anti-SAP30 antibodies depleted mSin3A but not MBD2 or MBD1. Asterisk indicates a non-MBD2 specific band which was not affected by immunodepletion with antibodies.

As MeCP2 and MBD2 are both associated with HDAC1 complexes, the relationship between MBD1 and HDAC1 was also examined. It was found that under conditions in which MBD2 and mSin3A were specifically depleted by anti-HDAC1 antibodies (Figure 5.2.4 E), the level of MBD1 in the HDAC1-depleted extracts was unaffected. Neither anti-MTA2 nor SAP30 antibodies depleted MBD2, consistent with the finding that MBD2 is not associated with NuRD or mSin3 complexes (Figure 4.2.4 C). Collectively, the data indicate that MBD1 is not associated with MBD2/ MeCP1 complex or other HDAC1-containing complexes.

### **5.2.5 Biochemical characterisation of MBD1**

A HeLa nuclear extract was passed over a Superose 6 gel-filtration column to determine the native size of MBD1. Western blot analysis of the fractions that elute from the column revealed that MBD1 migrates at the range of 200 kDa (Figure 5.2.5 A). The calculated molecular weight for MBD1 is 66 kDa, therefore it is likely that MBD1 interacts with itself or other nuclear proteins to form a higher molecular weight complex. MBD1 did not co-migrate with MBD2, as MBD2 eluted in fractions 7 and 8 under the same conditions (see Figure 4.2.6).

The abundance of MBD1 in HeLa nuclei was estimated by Western blotting of known amount of recombinant MBD1 and nuclei (Figure 5.2.5 B). The MBD1 signal from 4 ng of recombinant protein and  $1.4 \times 10^7$  nuclei was judged to be roughly equal (compare lanes 2 and 7). Based on this estimation, each nucleus is calculated to contain 2500-5000 molecules of MBD1. Therefore, MBD1 is not abundant in HeLa cells.

Experiments were also carried out to determine how tightly bound MBD1 was in HeLa nuclei. A typical 0.42 M KCl extraction was sufficient to elute MBD1 from the nuclei (Figure 5.2.5 C, lane 3). However, it was found that a significant amount of MBD1 was still retained in the insoluble chromatin pellet (lane 4). Dialysis of the 0.42 M KCl extract resulted in the precipitation of a fraction of 0.42 M KCl soluble MBD1 (lane 1), indicating that MBD1 is unstable under low salt condition. Thus, MBD1 exhibits character of chromatin-associated proteins such as MeCP2, which require high salt for efficient extraction from the nuclei (Meehan et al, 1992). The result suggests that MBD1 is tightly bound to chromatin (see also sections 5.2.9 and 5.2.10).

A

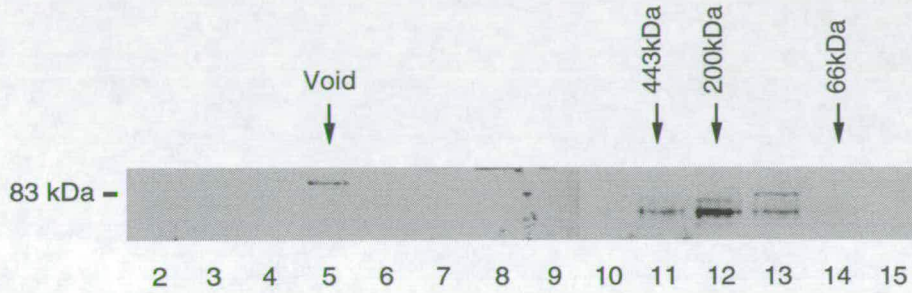


Figure 5.2.5 Biochemical characterisation of MBD1

A) MBD1 migrates at a molecular weight of 200-400 kDa.

HeLa nuclear extract was fractionated on a Superose 6 gel-filtration column. The eluted fractions were analysed by Western blotting to detect MBD1. The column was calibrated with known molecular weight markers.



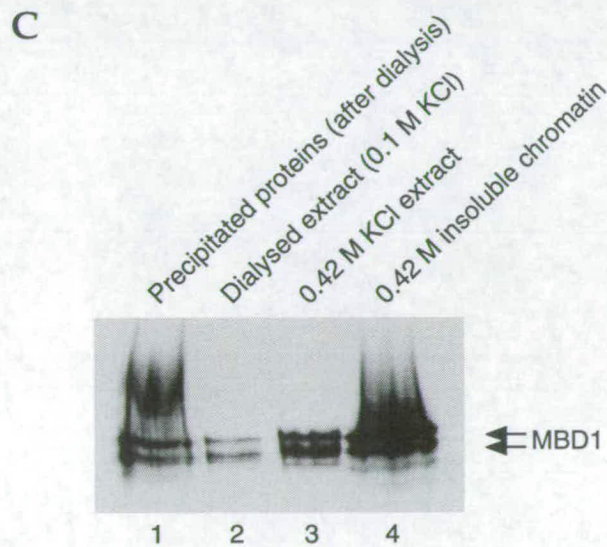
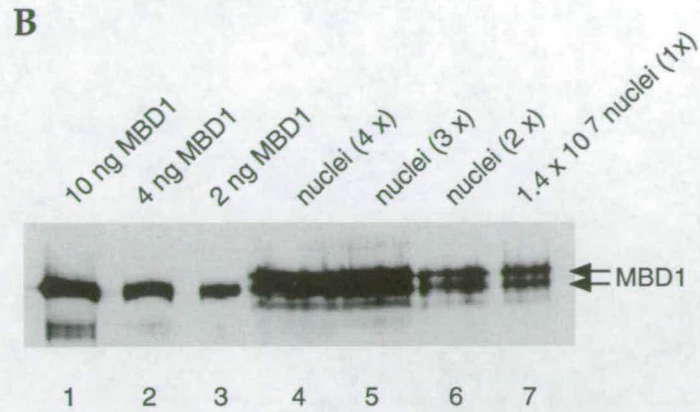


Figure 5.2.5. Biochemical characterisation of MBD1

B) MBD1 is not abundant in HeLa cells.

HeLa nuclei were solubilised in SDS loading buffer and analysed by Western blotting using anti-MBD1 antibody (lane 7 contained  $1.4 \times 10^7$  nuclei). Lanes 6, 5, and 4 contained 2x, 3x, and 4x the number of nuclei in lane 7, respectively. Known amounts of recombinant MBD1 were loaded for estimating the abundance of MBD1.

C) MBD1 is tightly bound in the nuclei.

HeLa nuclei were extract with 0.42 M KCl (lane 3, 26.5  $\mu$ g of protein). 0.42 M KCl insoluble chromatin fraction was resuspended in 0.9 M KCl and sonicated before SDS-PAGE (lane 4, 23.5  $\mu$ g of protein). MBD1 can be found in 0.1 M KCl insoluble fraction after dialysis (lanes 1 and 2; 23  $\mu$ g of protein was loaded in lane 2).



## 5.2.6 MBD1 can repress transcription in a DNA methylation dependent manner *in vivo*

Although MBD1 is not a component of the MeCP1 repressor complex, previous work from this laboratory showed that recombinant MBD1 can repress transcription of a methylated template in an *in vitro* system. It is therefore of interest to determine whether MBD1 can repress transcription of methylated reporter *in vivo*.

Methylated reporters are inevitably silenced when transfected into many mammalian cell-lines (Boyes and Bird, 1991; Levine et al, 1991). Two main parameters that dictate the transcriptional fate of the methylated reporters are: direct repression due to exclusion of transcription factor(s) and indirect repression due to methyl-CpG binding repressors. The presence of endogenous methyl-CpG binding repressors imposes a major hindrance to analysis of repression on methylated reporters by a protein of interest. It is also known that embryonic stem (ES) cells or embryonic carcinoma (EC) cell-lines are deficient in methyl-CpG binding proteins, e.g. MeCP1 (MBD2) and MeCP2 (Meehan et al, 1989; Boyes and Bird, 1991; Tate, 1996). The RNA transcripts of MBD1 and MBD2 are not detected in ES cells (Hendrich and Bird, 1998). The lack of methyl-CpG binding activity in F9 EC cells also correlates with the failure of these cells in repressing methylated reporters (Boyes and Bird, 1991; Levine et al, 1991). Therefore, co-transfection of a construct expressing MBD1 and differentially methylated reporters into F9 EC cells should allow one to test whether MBD1 can selectively repress the methylated reporter.

Consistent with previous studies, the HhaI methylated p $\beta$ G5BglII reporter (human  $\beta$ -actin promoter driving the  $\beta$ -galactosidase gene) was not significantly repressed compared to the mock methylated version (see Figure 5.2.6 A). In contrast, the same HhaI methylated reporter retained less than 20 % activity (compared to the non-methylated version) upon transfection into mouse L fibroblasts (Nan et al, 1997). The result suggests that repression on HhaI methylated reporter can be attributed to methyl-CpG binding repressors in mouse L cells.

Two concentrations of effector (construct for expression of MBD1) were tested for repression of reporters with different levels of methylation. At 0.25  $\mu$ g of effector concentration, similar level of expression from mock-methylated and HhaI-methylated reporters was observed (Figure 5.2.6 A, compare bars 1 and 3). When MBD1 was co-transfected along with the mock-methylated reporter, the  $\beta$ -galactosidase activity was

A

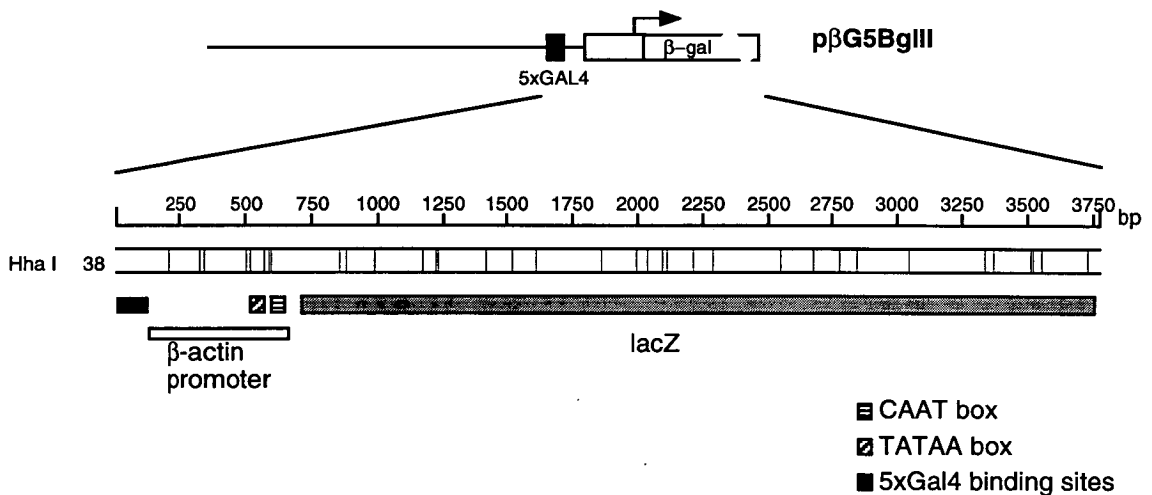
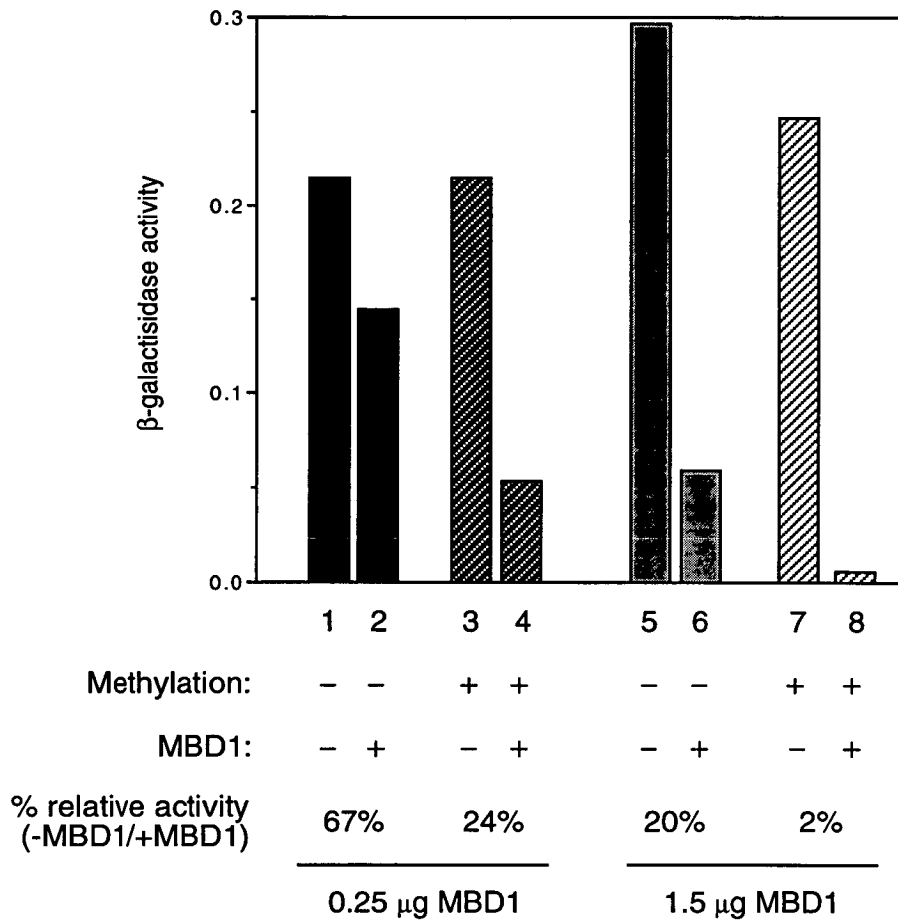


Figure 5.2.6.A Methylation dependent repression by MBD1

MBD1 (0.25 µg or 1.5 µg) was co-transfected either mock-methylated or methylated reporter. Methylation was introduced at all HhaI sites. Relative activity is the ratio of β-gal activity in the absence of co-transfected MBD1/ β-gal activity in the presence of MBD1 (e.g. value of column 1/ value of column 2). Reporter construct and HhaI methylation sites at the promoter and reporter gene are shown.

reduced to 67 % (Figure 5.2.6 A, bar 2). The  $\beta$ -galactosidase activity of the methylated reporter in the presence of MBD1 was repressed to 24 % (Figure 5.2.6 A, bar 4). As MBD1 binds specifically to methylated DNA *in vitro* (see Figure 5.2.1 B), it is suspected that the reduction in activity of the non-methylated reporter upon co-transfection of MBD1 expressing construct is due to a toxic effect of MBD1. Therefore, the 2.8 fold difference in the relative reporter activity between non-methylated and methylated reporters can be attributed to the only difference in one parameter: DNA methylation. Increasing the amount of effector construct from 0.25  $\mu$ g to 1.5  $\mu$ g, resulted in even higher level of repression on non-methylated reporter, and the  $\beta$ -galactosidase activity was reduced to 20 % relative to the  $\beta$ -galactosidase activity when no effector was co-transfected (Figure 5.2.6 A, bar 6). With 1.5  $\mu$ g of transfected effector, the relative reporter activity of the HhaI methylated reporter was 2 % (Figure 5.2.6 A, bar 8). More than 20 fold difference in relative reporter activity between the non-methylated and methylated reporter was observed. Therefore, the result showed that MBD1 can preferentially repress the methylated reporter. Furthermore, MBD1 repressed the methylated reporter in a dose-dependent manner (2.8 fold at 0.25  $\mu$ g of effector and 20 fold at 1.5  $\mu$ g of effector).

It is of interest to test whether MBD1 can still repress transcription in a methylation dependent manner when the promoter is activated by potent transcriptional activators such as the VP16, E2F1, Sp1 activation domain (Blau et al, 1996; Emili et al, 1994; Sadowski et al, 1988; Triezenberg et al, 1998). The p $\beta$ G5BglII reporter (mock methylated or HhaI methylated) was co-transfected into F9 cells along with constructs for expression of GAL4-VP16 and MBD1. The relative reporter activity of the mock methylated reporter was 72.4 %, and that of the HhaI methylated reporter was 21.9 % (Figure 5.2.6 B, bars 1 to 4). Again, a toxic effect on the non-methylated reporter in the presence of MBD1 was observed. There is a 3 fold difference in the relative reporter activity between non-methylated and methylated reporters. The level of repression on the methylated reporter (based on fold difference) is surprisingly very similar to the result when the reporters were not transactivated by GAL4-VP16 (Figure 5.2.6 A, 0.25  $\mu$ g of effector). As the Gal4 binding sites (5 x CGGAAGACTCTCCTCCG) do not have HhaI methylation (GCGC) sites (Giniger et al, 1985; Lin et al), it is not likely that HhaI methylation prevents binding of GAL4-VP16 to the Gal4 binding sites and leads to reduced reporter activity. In addition, the HhaI methylated reporter was activated to similar extent as the mock methylated reporter by GAL4-VP16 (see also below). The effect of two other activators (E2F1 and Sp1) on repression by MBD1 was also tested. Similar to the GAL4-VP16 experiment, GAL4-Sp1 or GAL4-E2F were co-transfected along with the reporter

**B**

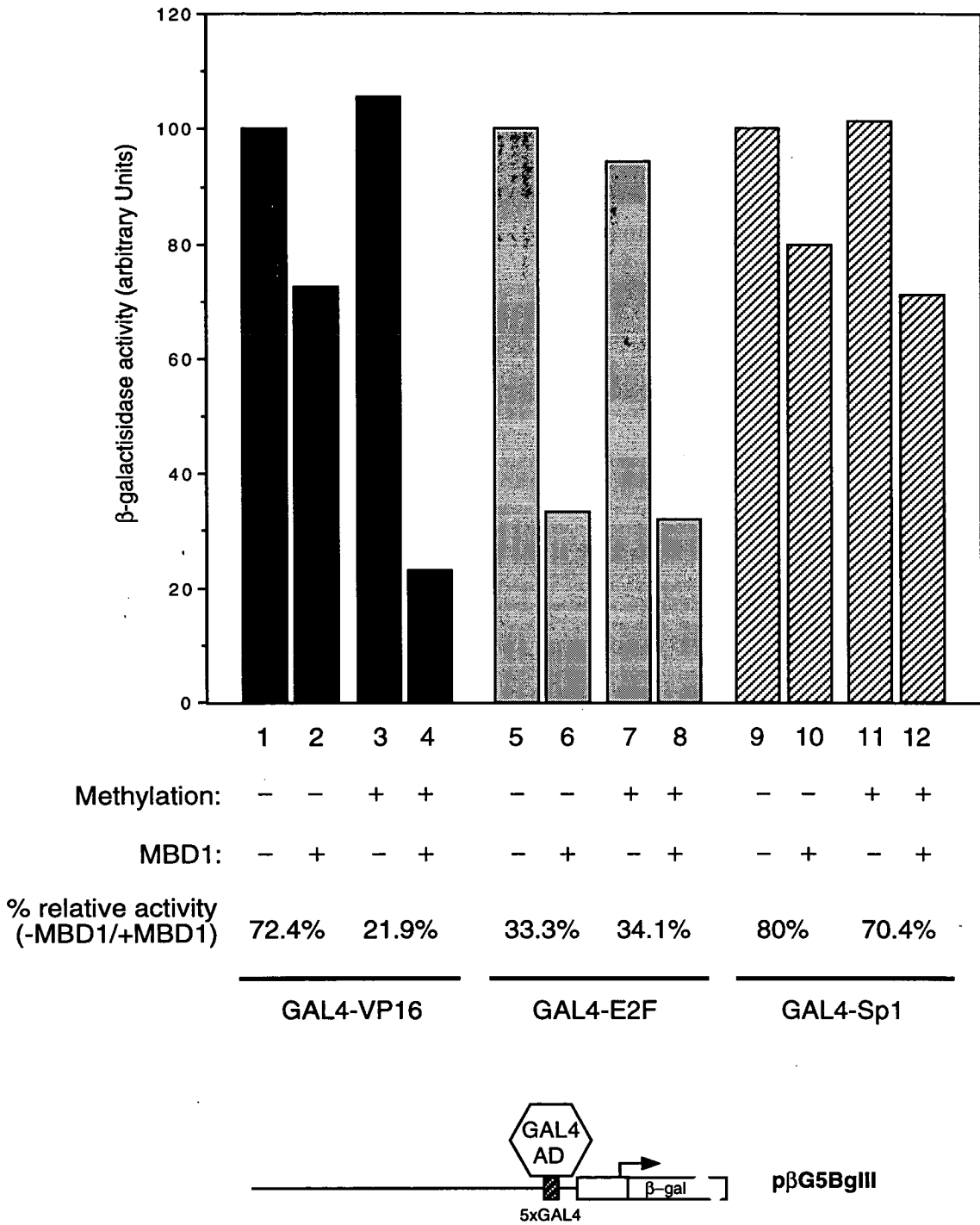


Figure 5.2.6.B. Effect of MBD1 on reporters activated by different activation domains.

Methylation was introduced at all HhaI sites. The results are derived from average of duplicates. Relative activity is the ratio of β-gal activity in the absence of co-transfected MBD1 / β-gal activity in the presence of MBD1 (e.g. value of column 1 / value of column 2). GAL4-AD is GAL4-Activation domain fusion.

(either mock-methylated or HhaI-methylated) and effector (for expression of MBD1). Strangely, the HhaI methylated reporter was not selectively repressed in the presence of GAL4- E2F1 or GAL4-Sp1 (Figure 5.2.6 B, bars 5 to 12). The result suggests that repression by MBD1 is only targeted towards certain classes of activators.

During the course of this work, Fujita et al (1999) also showed that MBD1 can repress transcription in a methylation-dependent manner in COS cells and *Drosophila* cells.

### **5.2.7 MBD1 contains a potent transcriptional repression domain**

As the HhaI methylation sites occurred throughout the whole p $\beta$ G5BglIII reporter plasmid (including promoter and lacZ gene), it is not certain if repression due to MBD1 was the results of steric blocking of transcriptional initiation and/ or elongation (see Figure 5.2.6 A). Therefore, a more informative assay is required to ask if MBD1 can actively repress transcription when tethered to sites remote from the promoter region. An inkling that MBD1 may repress transcription comes from functional studies of the HRX protein (Figure 5.2.7 A). A CxxC containing region in HRX has been reported to repress transcription *in vivo* using a GAL4 fusion assay (Prasad et al, 1995; Zeleznik-Le et al, 1994), therefore it is of interest to test whether the similar CxxC domains in MBD1 can also repress transcription.

A GAL4 fusion repression assay was used to address this possibility. Constructs with different fragments of MBD1 fused to cDNA encoding the GAL4 DNA binding domain were generated. To avoid mis-targeting of the GAL4 fusion proteins to methylated sites in the genome, the first 43 amino-acids of MBD1, corresponding to half of the MBD, was omitted from all constructs. These effector constructs were separately co-transfected into mouse L929 fibroblasts with a  $\beta$ -galactosidase reporter (with or without 5 Gal4 binding sites upstream of the human  $\beta$ -actin promoter; G0 or G5). Approximately 60 hours (two and a half days) after transfection, the cells were harvested and extracts were assayed for  $\beta$ -galactosidase activity. For each construct, titration of at least 3 concentrations of effector was carried out.

As a pilot experiment to address whether the CxxC domains can repress transcription, three GAL4-MBD1 constructs (81-556 aa, 261-556 aa, 350-556 aa) were tested for transcriptional repression activity. If the repression property of CxxC domain is also retained in MBD1, the first two constructs with CxxC domain(s) (81-556 aa and 261-556

A

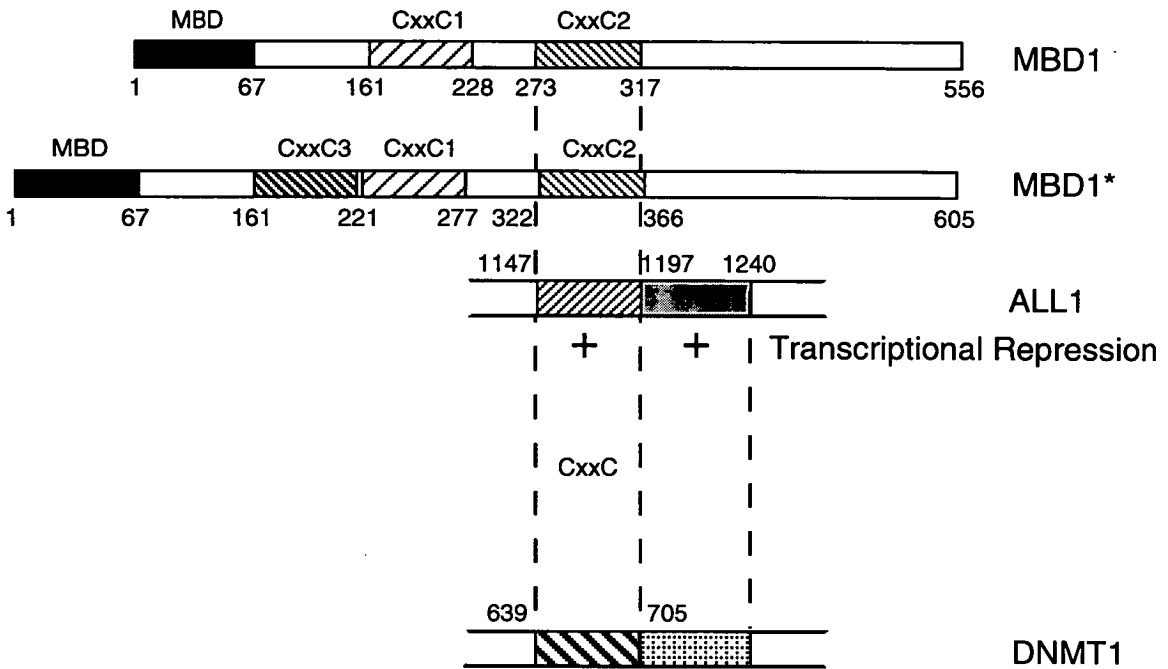


Figure 5.2.7 A. Is MBD1 a transcriptional repressor?

The cysteine-rich (CxxC domain) transcriptional repression domain of ALL1 can also be found in DNMT1 and MBD1. Two forms of MBD1 that can be expressed by the human gene: MBD1 and MBD1\* (Cross et al, 1997; Hendrich et al, 1999a). MBD1 contains two CxxC domains, while MBD1\* has three CxxC domains. The two shaded regions in ALL1 has been shown to contain transcriptional repression activity (indicated by +; Prasad et al, 1995). The CxxC domains are shaded in hatch.

aa) should repress transcription of a reporter with Gal4 binding sites (G5). Unexpectedly, all three effectors, including the construct encoding MBD1 without CxxC domains (350-556 aa), repressed the G5 reporter (Figure 5.2.7 B). The two constructs containing the CxxC domain also repressed the reporter without Gal4 binding sites (G0). It is concluded that the C-terminal of MBD1 can repress transcription and repression is dependent on interaction with the promoter. It is not clear whether the CxxC-containing effectors can repress transcription specifically or if their expression is toxic to the cells leading to loss of reporter activity.

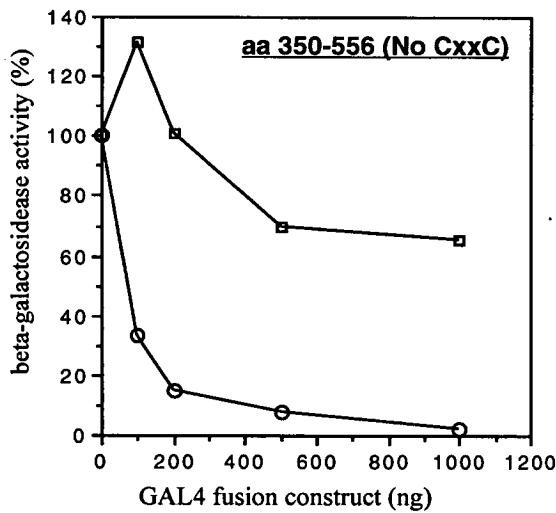
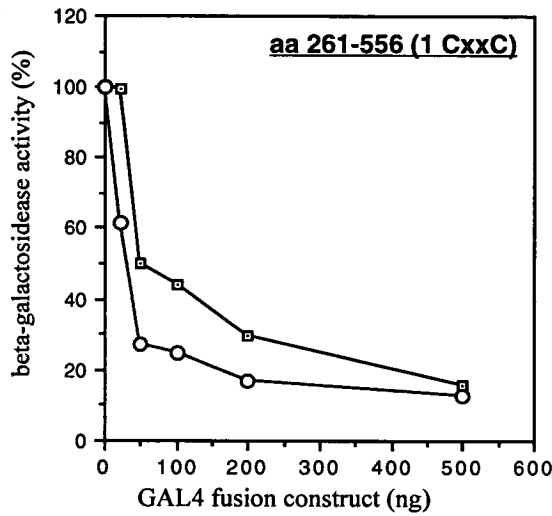
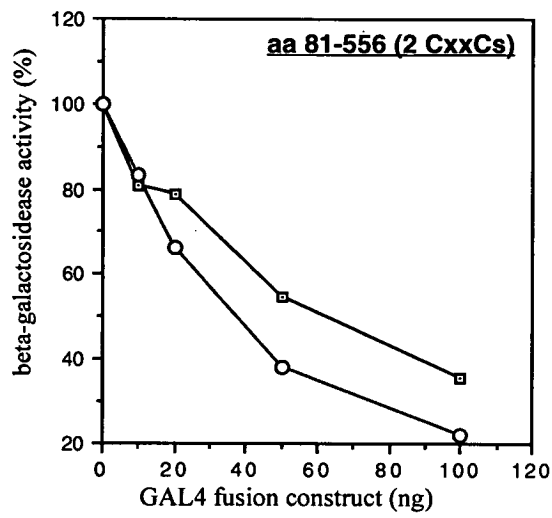
In the next experiment, an internal control reporter (SV40 promoter driving luciferase gene) was included in the transfection assay for normalising the  $\beta$ -galactosidase activity. Again, the 81-556 effector repressed the G0 reporter (Figure 5.2.7 C). It should also be noted that the G5 reporter was consistently repressed to a greater extent than the G0 reporter by 81-556 aa effectors. The C-terminal fragment of MBD1 (350-556 aa) however specifically repressed the G5 reporter.

A detailed mapping experiment was performed to delineate the C-terminal transcriptional repression domain, TRD (Figure 5.2.7 D, E, F). As the human and mouse MBD1 have two or three CxxC domains similar to that of ALL1, it is of particular interest to know if they can all repress transcription (see Figure 5.2.7 A and G). A 33 amino-acids region at the C-terminal of MBD1 was identified as the TRD. It is notable that only the second CxxC domain (273-340 aa) can repress transcription (Figure 5.2.7 H). Related N-terminal CxxC domains (161-228 aa of MBD1 and 161-221 aa of MBD1\*) did not repress the reporters. The CxxC region of the DNA cytosine 5-methyl transferase also did not repress. Western blotting using anti-GAL4 antibodies indicates that the GAL4-fusion proteins were expressed, although GAL4-CxxC1 (161 to 221 aa of MBD1\*) was poorly expressed (Figure 5.2.7 I). The results suggest that, despite of the significant sequence conservation among the CxxC domains, their functions in different proteins may differ.

### **5.2.8 Characterisation of the C-terminal TRD (Mutagenesis; Long distance repression; TSA sensitivity)**

Comparison between the human and mouse MBD1 revealed that the TRD is conserved with 66.7 % identity and 78.6 % similarity over a 33 amino-acids region. The TRD is enriched in hydrophobic amino acids (alanine-A, isoleucine-I, leucine-L, phenylalanine-



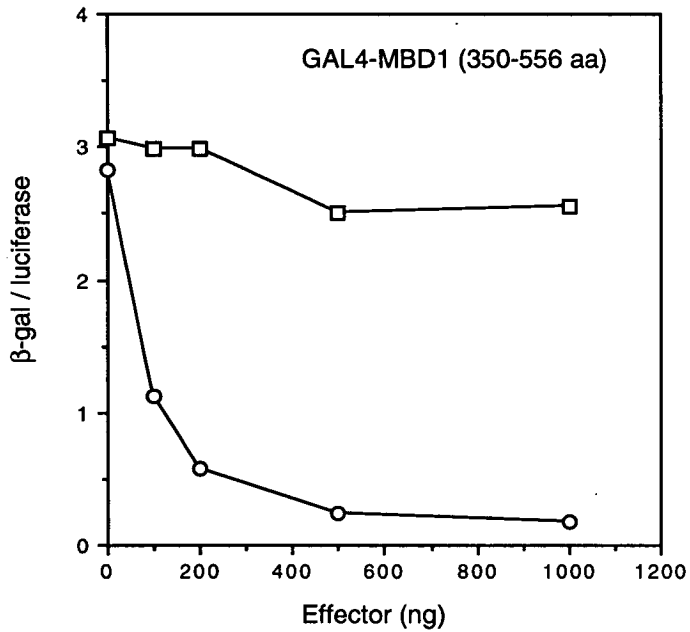
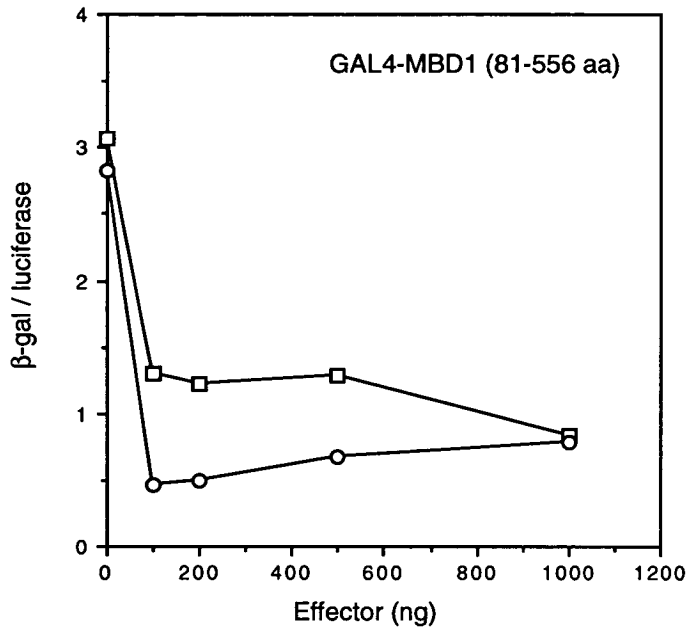
**D**

- G5 reporter (with Gal4 binding sites)
- G0 control reporter (no Gal4 binding sites)

Figure 5.2.7. MBD1 is a transcriptional repressor

B) An effector lacking the CxxC domains can specifically repress transcription of a reporter with Gal4 binding sites. The lacZ reporter is driven by human  $\beta$ -actin promoter.

C



- G5 reporter (with Gal4 binding sites)
- G0 control reporter (no Gal4 binding sites)

Figure 5.2.7. MBD1 is a transcriptional repressor

C) GAL4-MBD1 (81 to 556 aa) repressed both the G0 and G5 reporters. The  $\beta$ -gal activity was normalised with luciferase activity from a co-transfected luciferase reporter. GAL4-MBD1 (350-556 aa) repressed only the G5 reporter. The lacZ reporter is driven by human  $\beta$ -actin promoter, while the luciferase reporter is driven by SV40 promoter.

D

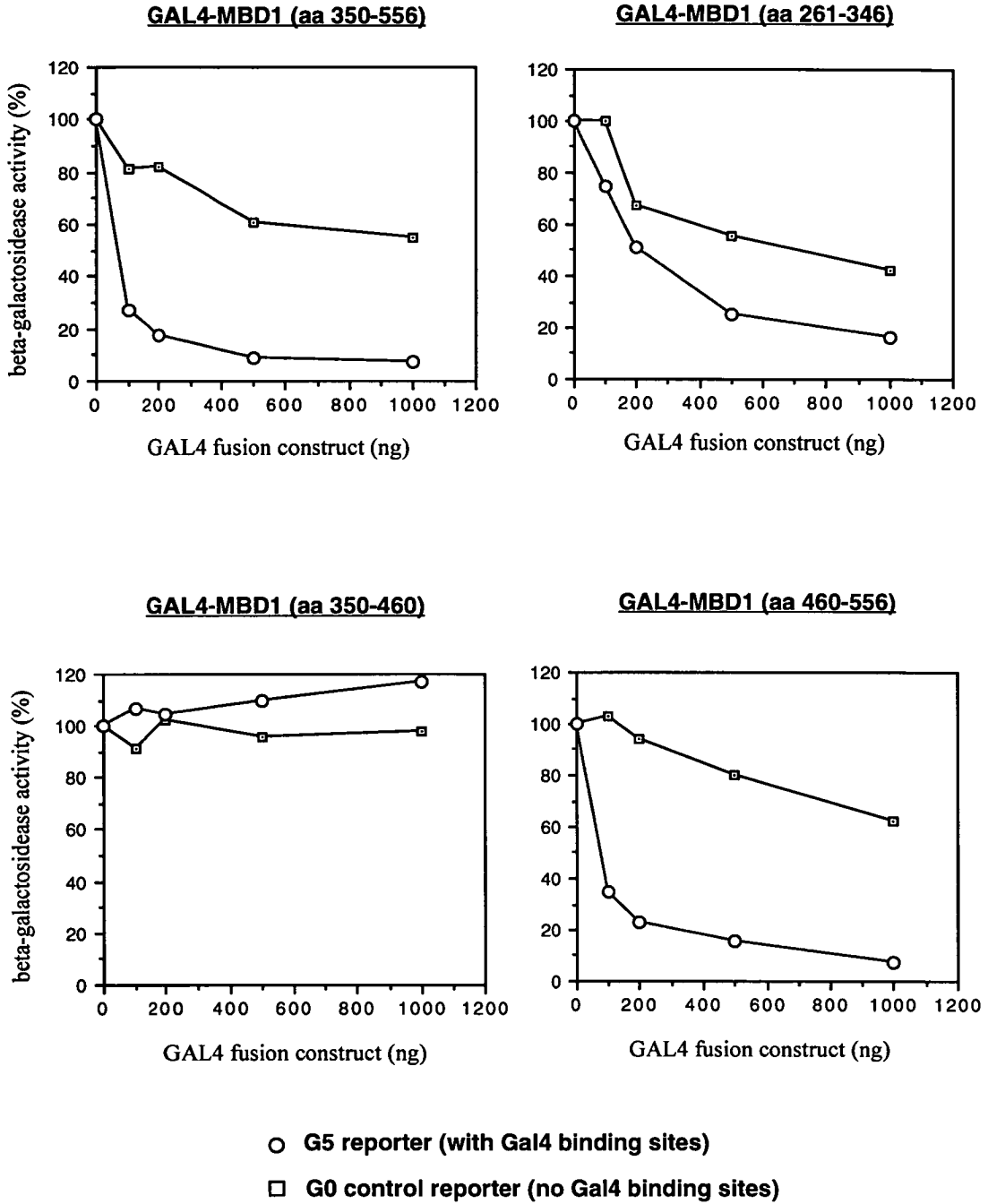


Figure 5.2.7. MBD1 is a transcriptional repressor

D) Examples of repression profile on G5 and G0 reporters by various GAL4-MBD1 fragments.

**E**

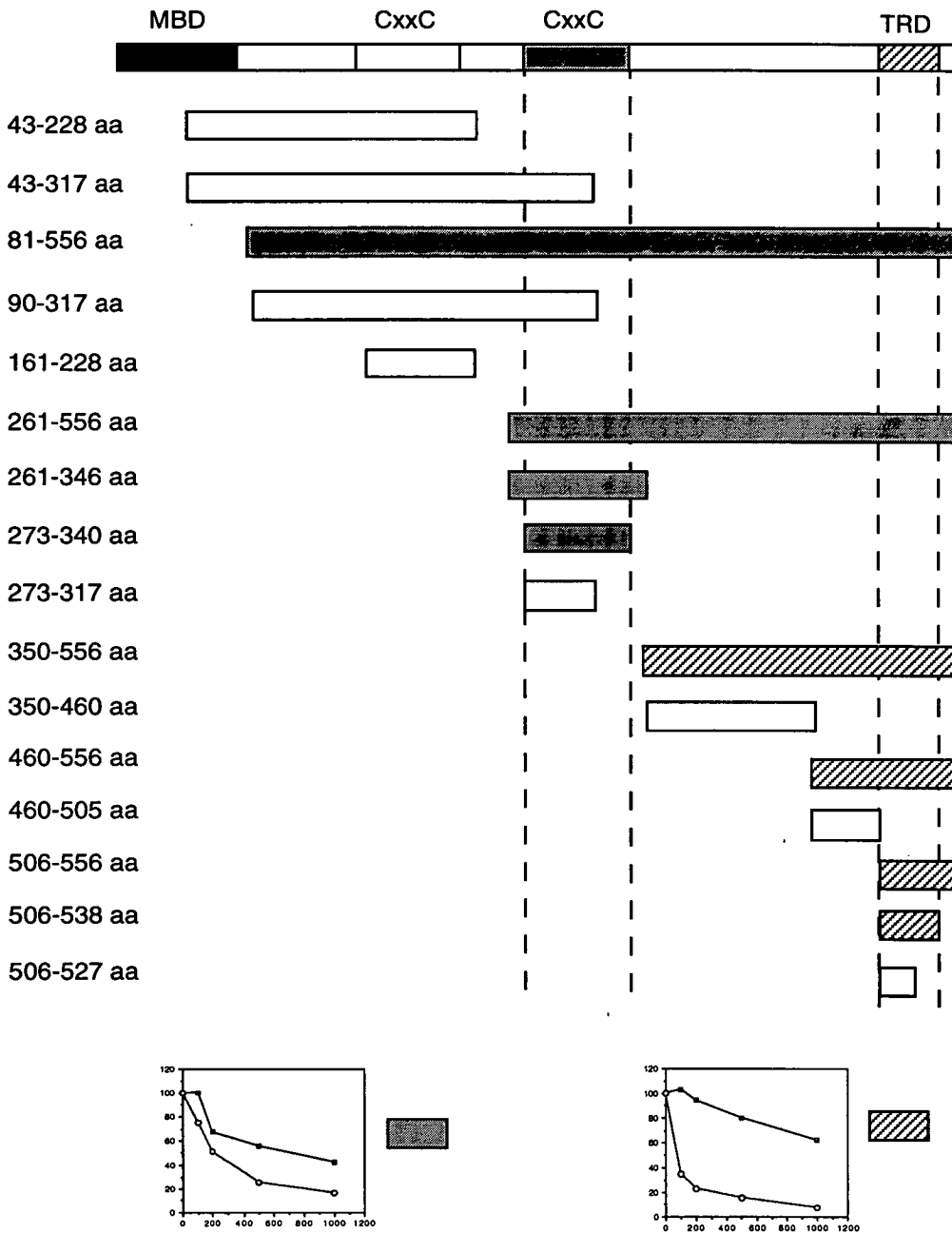


Figure 5.2.7. MBD1 is a transcriptional repressor

E) Mapping of fragments of MBD1 with transcriptional repression activities. Two regions with distinct repression profile on G0 and G5 reporters are identified (gray, and hatched). The minimal hatched region repressed the G5 reporter specifically, and is denoted as the transcriptional repression domain (TRD). A CxxC domain repressed both the G0 and G5 reporters (gray).

F

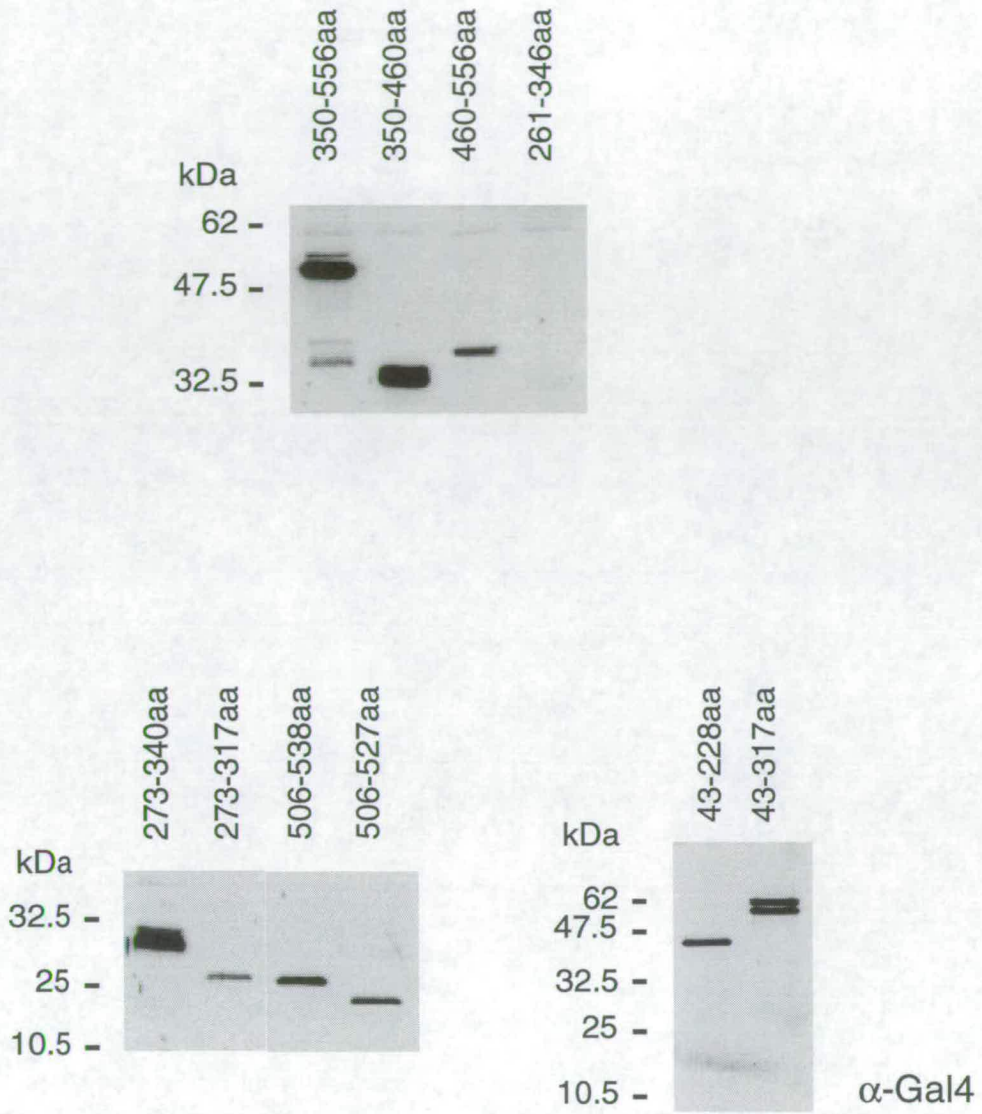


Figure 5.2.7. MBD1 is a transcriptional repressor

F) Expression of different GAL4-MBD1 fusion proteins as detected by anti-GAL4 antibodies. Certain constructs e.g. 261-346 aa were poorly expressed. All constructs expressed GAL4-fusions of the correct expected sizes.

G

```

      .10      .20      .30      .40      .50      .60      .70
MBD1* (161 to 221) .RIAFNREQRMFRRVGCCECAACQVTEDCGAGSTCTLLQLPHDVASGLFCERRRCLRIVER.....
MBD1 (161 to 228) .RIAFNREQRMFRRSGCGVCRCCQTOEDCGEPCICL..RPPRFGLRRQWKCVRRCLRGKHARRKGGCDSK
MBD1 (273-340)   .DEDELQPYTHRRQNRRCGACAAQLRRMDCCRCDFCCDKPKFPGGSNPKRQKCRWRCCLQFAMRLLPSV...
MBD1 (273-317)   .DEDELQPYTHRRQNRRCGACAAQLRRMDCCRCDFCCDKPKFPGGSN.....
DNMT1 (639-705)  .EYDKEDKENAMRRRCGVCEVCQQP.ECGRCKACNDMVKFGGTCGRSKQACLKRRCPNLAVKADDE...
HRX (1147-1197)  .....KKGRRSRRRCQCPCQVPEDCGVCTNCLDKPKFPGGRNPKKRCCKMRRRCQN.....
  
```

H

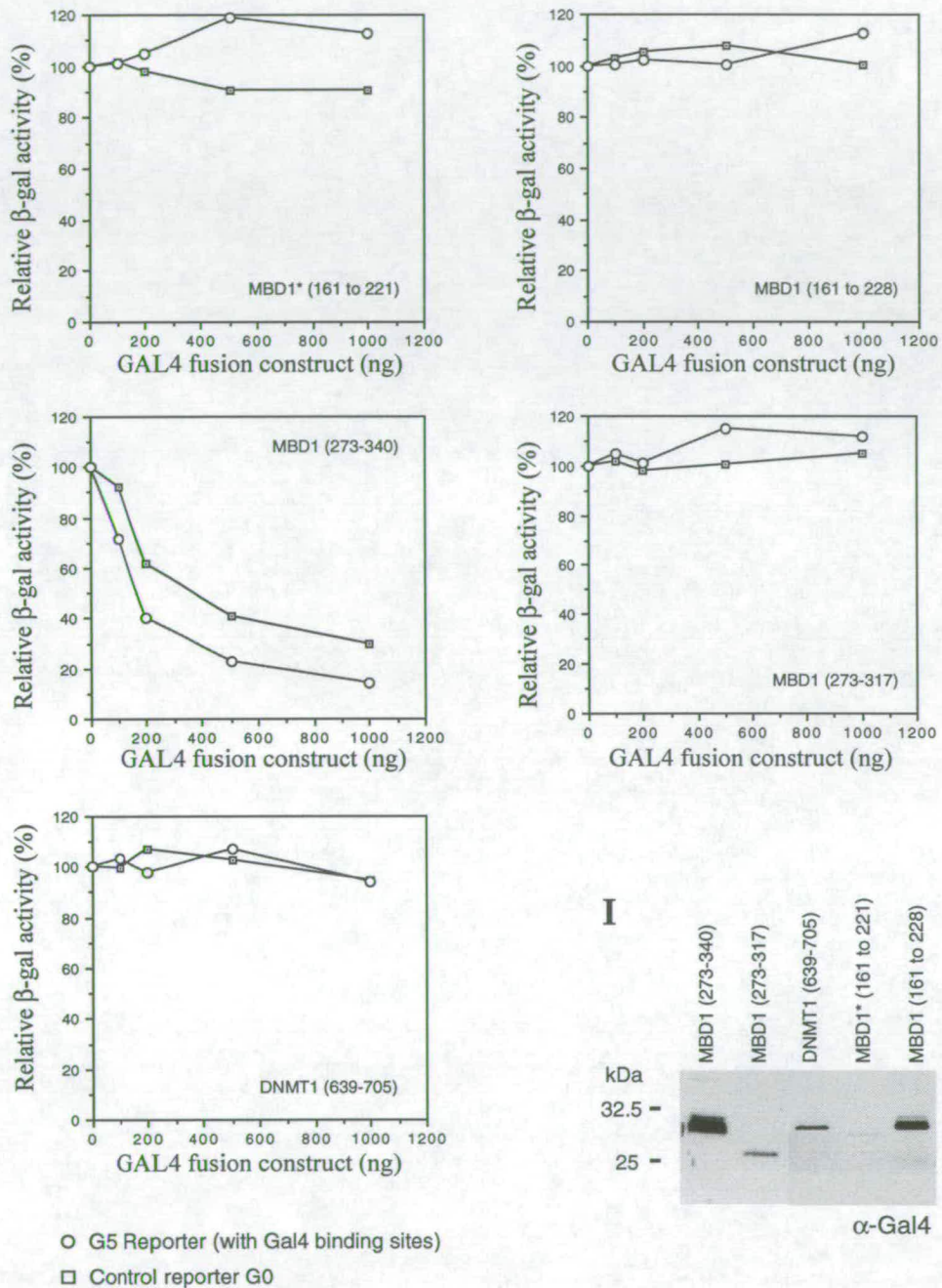


Figure 5.2.7. MBD1 is a transcriptional repressor

G) Sequence alignment of CxxC domains from MBD1, DNMT1, and HRX (also known as ALL1).

H) Repression profile on the G5 and G0 reporters by various CxxC domains.

I) Expression of GAL4-CxxC domains as detected by anti-GAL4 antibodies.



F, valine-V): 33 % and glycine: 15.1 % (Figure 5.2.8 A). Database searching with the 33 amino-acids TRD did not uncover similarity with known transcriptional regulators.

To test the importance of the conserved amino acids, three TRD mutations with single amino-acid substitution were generated (Figure 5.2.8 A). Conversion of hydrophobic residues I527 and L530 to basic residue arginine (R) abolished the repression activity of TRD. Substitution of polar residue T525 to lysine (K) lead to partial loss of repression activity. Western blot analysis using anti-GAL4 antibodies indicated that the expression of these GAL4-TRD mutants was comparable to the wild-type TRD fusion protein.

The next question to ask is whether the TRD can repress transcription from a distance. Two reporters (p $\beta$ -887-G5 and p $\beta$ -1774-G5) with the 5 Gal4 binding sites located 1.3 kb and 2.1 kb away from the translation start site were used to address this question (Figure 5.2.8 B). The TRD repressed the activity of p $\beta$ G5BglII reporter (G5 reporter with 5 Gal4 binding sites 400 bp upstream of the translation start site) to about 10 %. Under the same effector concentration (250 ng), the activity of p $\beta$  $\beta$ geoN/B reporter (G0 reporter without Gal4 binding site) was not affected. The GAL4-TRD also efficiently repressed both the p $\beta$ -887-G5 and p $\beta$ -1774-G5 reporters to 12 %. The result indicates that TRD can mediate long distance repression. This finding contrasts the significant reduction with distance in the efficiencies of activation by GAL4-VP16 and of repression by the GAL4-TRD of MeCP2 in the same assay.

The association between methylated DNA, transcriptional repression and hypoacetylated histones raised the possibility that MBD1 may also repress transcription via interaction with histone deacetylases. TSA treatment restored the relative reporter activity by approximately two to three fold, suggesting that deacetylation may be involved (Figure 5.2.8 C). The repression by the CxxC domain was however not sensitive to TSA (data not shown). MBD1 was not depleted by HDAC1, SAP30, or MTA2 antibodies under conditions where MBD2 was specifically depleted by HDAC1 antibodies (see Figure 5.2.4 E). Therefore, it is unlikely that MBD1 is associated with HDAC1 complexes. Previous experiments showed that anti-MBD1 antibody did not immunoprecipitate histone deacetylase activity from crude HeLa cell extract (Figure 4.2.4 A). As the extract was prepared by lysing the cells with NP40 detergent in the presence of 150 mM NaCl, it is not certain if MBD1 was effectively extracted under the low salt condition (see Figure 5.2.5 C). Thus, the low abundance of MBD1 in the low salt extract may help to explain the lack of histone deacetylase activity immunoprecipitated by anti-MBD1 antibody.

**A**

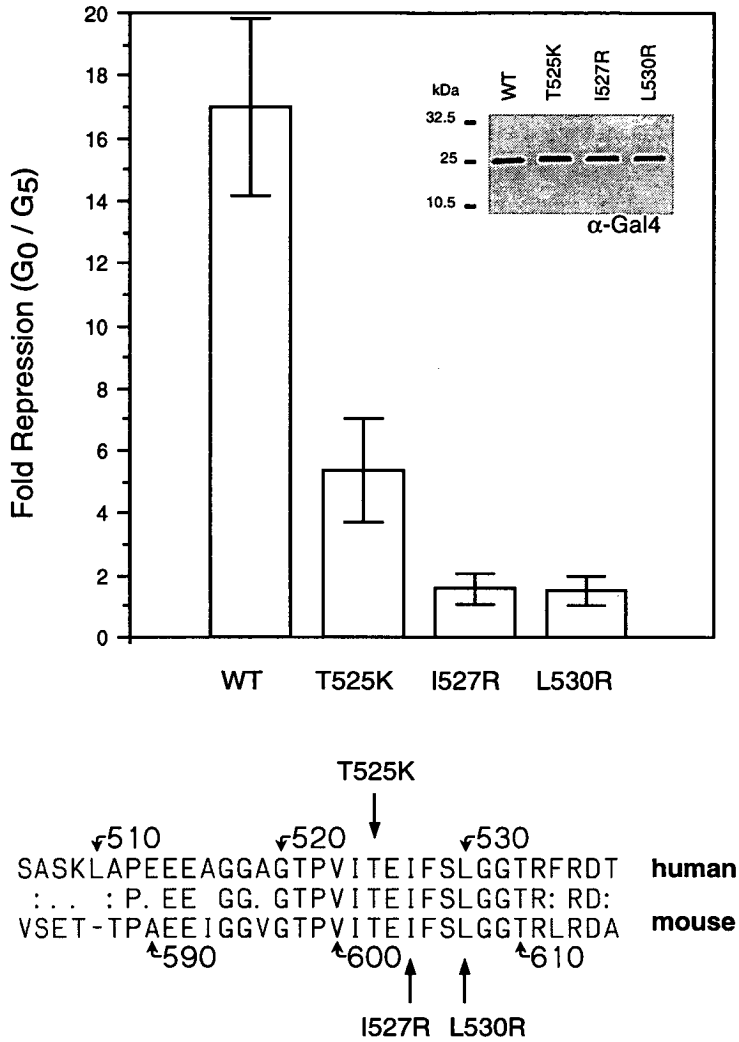


Figure 5.2.8. Characterisation of TRD

A) The TRD is conserved between the human and mouse MBD1. Mutation of two hydrophobic residues (I527 and L530) abolished transcriptional repression property of the TRD. Substitution of T525 residue only lead to partial loss of activity. The result shown are based on three independent transfection.

Inset: expression of all GAL4-TRD are comparable.

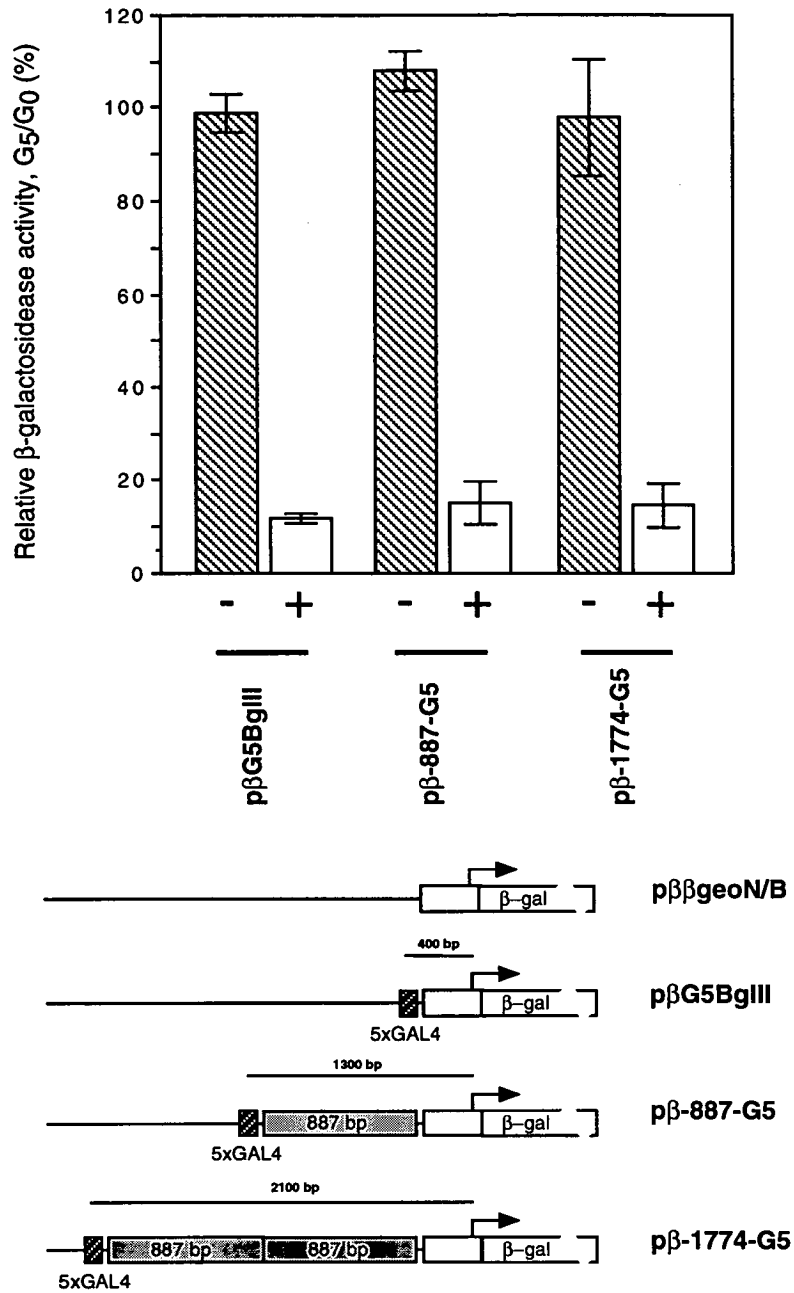
**B**

Figure 5.2.8. Characterisation of TRD

B) The TRD can repress transcription from a distance. 250 ng of effector (GAL4-TRD) was used. The reporter constructs and the location of Gal4 binding sites relative to the transcription start site are shown. The result shown are based on three independent transfection.

C

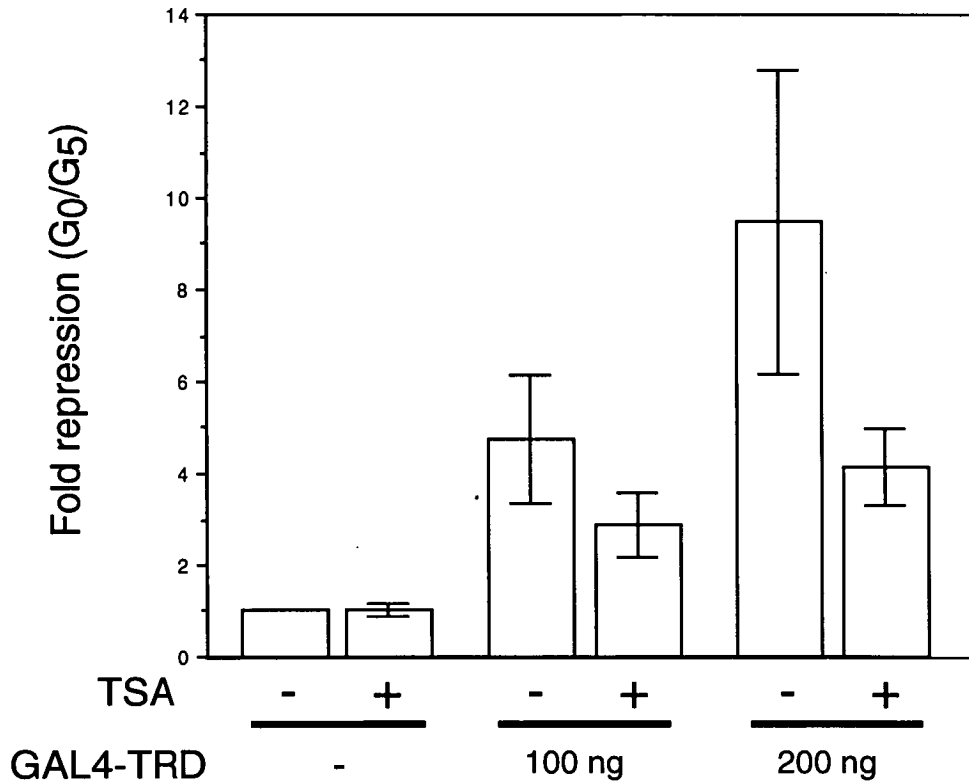


Figure 5.2.8. Characterisation of TRD

C) The TRD repressed transcription in a deacetylation dependent manner. Transfected cells were incubated with TSA (100 ng/ml) for 24 hours before harvest. Two concentration of effector (100 ng and 200 ng of GAL4-TRD) were transfected. The result shown are based on three independent transfection. TSA did not affect the relative activity of the reporters (G0/G5) in the absence of GAL4-TRD.

### 5.2.9 MBD1 is a chromosomal protein

As the S751 antibody can recognise native MBD1 in a nuclear extract, indirect immunofluorescence with this antibody was used to look at localisation of endogenous MBD1 in HeLa cells. Nucleoplasmic staining with the exclusion of nucleoli was detected (Figure 5.2.9 A). Preimmune antibodies gave only background staining (data not shown). Affinity purified anti-MBD1 antibodies gave the same staining pattern as crude immune serum (data not shown). The staining of MBD1 on mitotic chromosomes was weak and it was not possible to analyse the chromosomal localisation of MBD1 on aneuploid HeLa cells.

In collaboration with Dr. Peter Jeppesen (Western General Hospital, Edinburgh), we used this antibody to look at localisation of MBD1 on human metaphase chromosomes prepared from human DD7 lymphoblastoid cell-line (Figure 5.2.9 B). The major sites of labeling are centromeric heterochromatin on chromosomes 1, 9, 15 and 16. In addition, MBD1 was also detected at the satellites on acrocentrics (rDNA), and some telomeres. The staining pattern is specific because locus specificity was observed (staining on both chromatids). Preimmune serum gave only background staining (data not shown). Affinity purified MBD1 antibodies also gave the same staining pattern as the immune serum. The localisation of MBD1 is also consistent with the staining pattern of sites of methylation using anti-5-meC antibody (Barbin et al, 1994).

To further analyse the relationship between MBD1 and constitutive heterochromatin or facultative heterochromatin (e.g. inactive X-chromosome), metaphase chromosomes of a primary female fibroblast line HF19 (Jeppesen and Turner, 1993) were co-stained with anti-MBD1 and anti-acetylated H4 sera (Figure 5.2.9 C, performed by Dr. Peter Jeppesen). MBD1 was found to be concentrated at centric constitutive heterochromatin which is deficient in acetylated histones (Jeppesen et al, 1992). The localisation of MBD1 on both active and inactive X chromosomes was however not significantly different, though the inactive X chromosome was hypoacetylated. The likely explanation is that the overall methylation status of the X chromosomes are not very different (Barbin et al, 1994). In addition, low-level and non-uniform labeling of euchromatic chromosome arms was observed. The result suggests that the primary factor governing the chromosomal localisation of MBD1 is methylation.

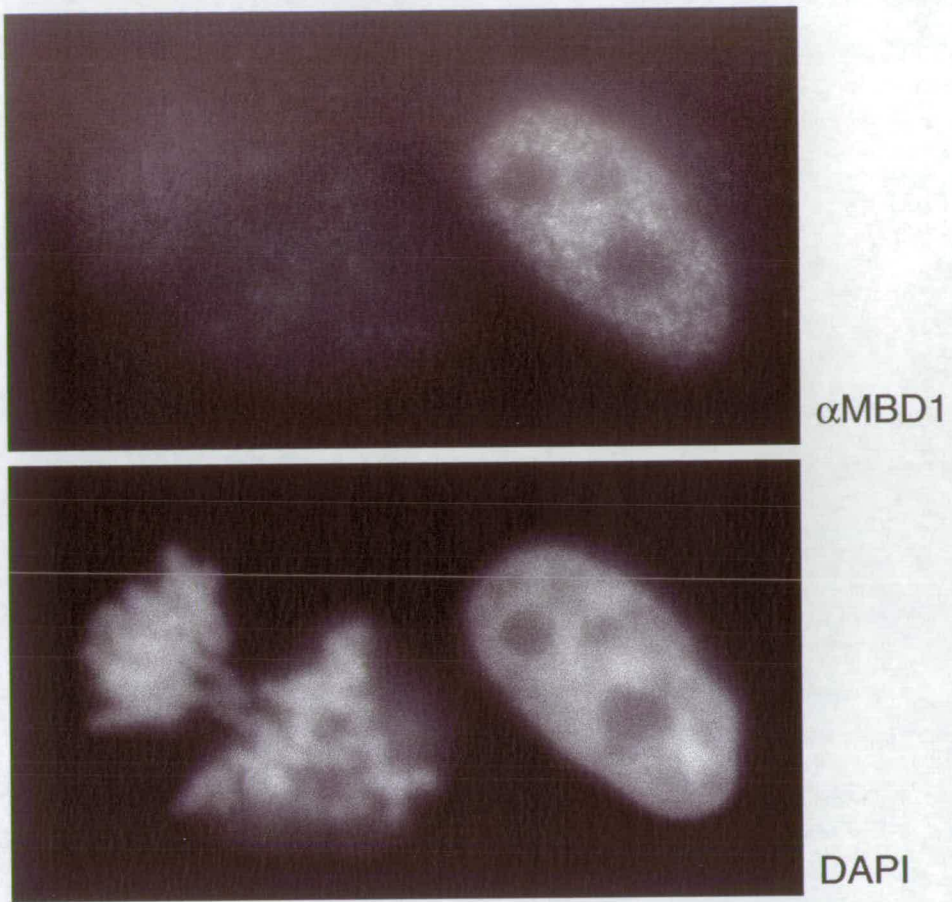
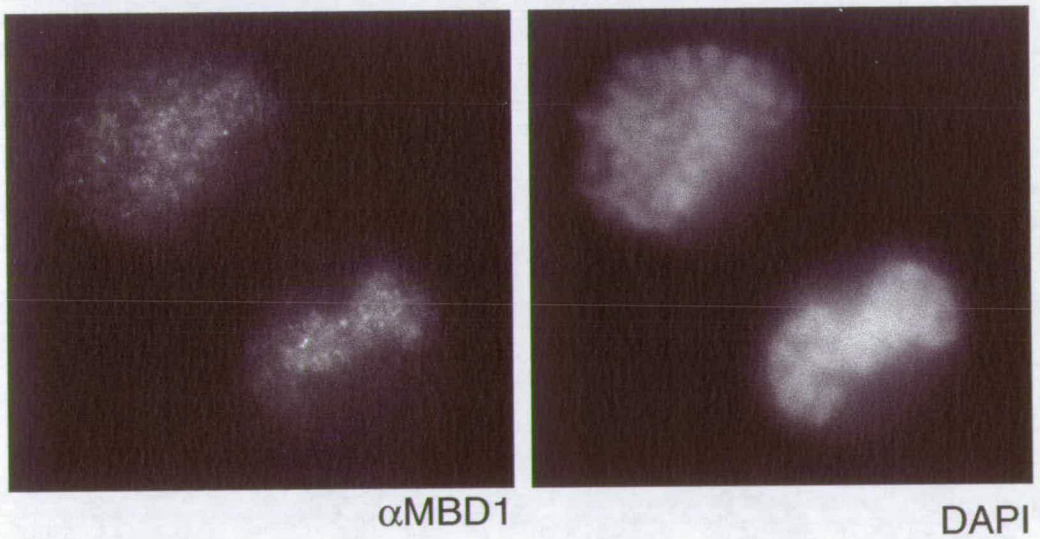
**A****B**

Figure 5.2.9. Localisation of MBD1

A) MBD1 is a nuclear protein.  
HeLa cells were stained with anti-MBD1 S751 antiserum.

B) MBD1 can be found on mitotic chromosomes.



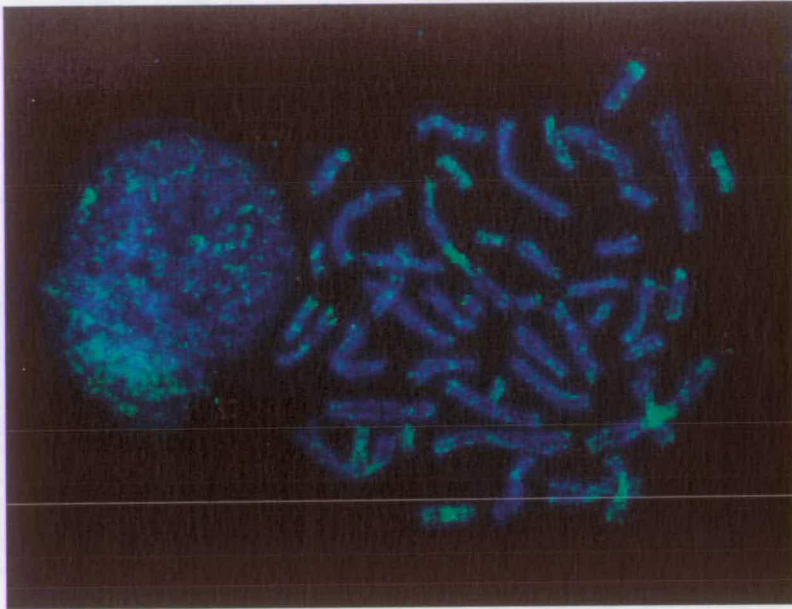


Figure 5.2.9. Localisation of MBD1.

D) MBD1 is a chromosomal protein.

Metaphase chromosomes of DD7 lymphoblastoid cells were stained with affinity purified anti-MBD1 antibodies (green). The spread was counterstained with DAPI for DNA (blue). The major sites of MBD1 are centric heterochromatin of chromosomes 1, 9, 15, and 16. MBD1 was also detected at the satellites on acrocentrics (rDNA), and some telomeres.

This experiment was performed by Dr. Peter Jeppesen.

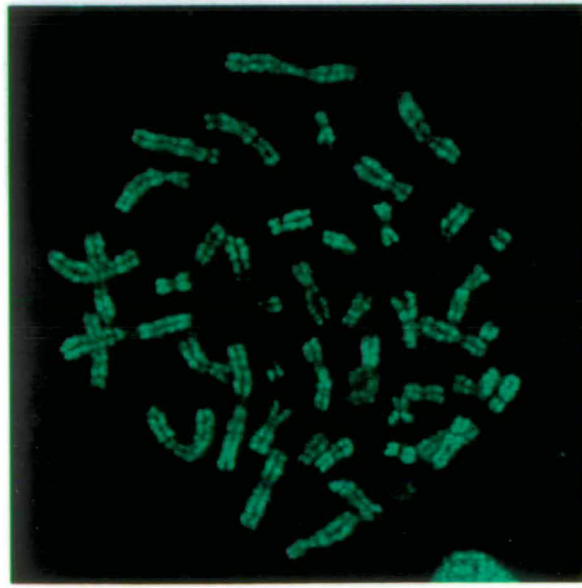
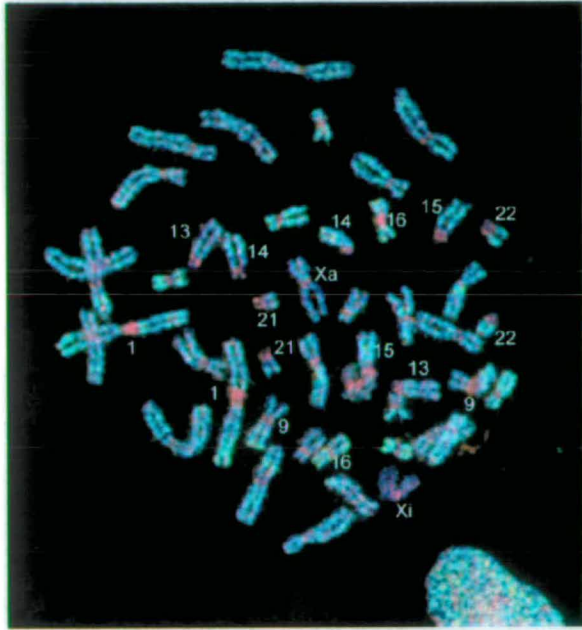
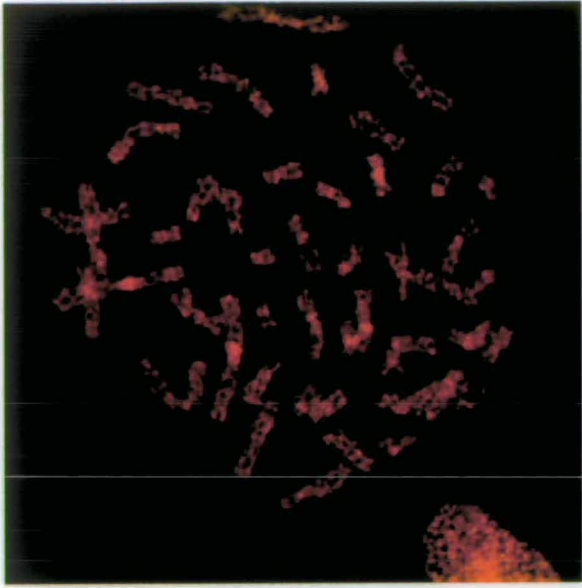


Figure 5.2.9. Localisation of MBD1.

E) MBD1 is a chromosomal protein.

Top panel:

Metaphase chromosomes of female human HF19 cell-line were stained with anti-MBD1 antiserum (red).

Middle panel:

Merged image of top and bottom panels. The spread was counterstained with Hoechst 33258 for DNA (blue). Several chromosomes were identified based on Hoechst and H4 acetylation patterns. Xi is inactive X chromosome, while Xa is active X chromosome.

Bottom panel:

The chromosomes were simultaneously stained with anti-acetyl lysine 12 H4 antiserum (green).

MBD1 is concentrated at pericentromeric heterochromatin of chromosomes 1, 9, 15, and 16. In addition, significant staining of MBD1 was observed at euchromatic chromosome arms. This experiment was performed by Dr. Peter Jeppesen.

### 5.2.10 MBD1 has an affinity for core histones

In order to probe the mechanism of repression by MBD1, candidate interacting proteins were identified. Different regions of MBD1 were tested for their affinity for core histones for two reasons. Firstly, the transcriptional repression domain of yeast Ssn6/Tup1 global repressor has been shown to interact directly with histone H3 and H4 tails (Edmondson et al, 1996). Secondly, MBD1 is a chromosomal protein *in vivo* (see section 5.2.7), and it is therefore of interest to examine its affinity with components of chromatin.

A GST-pull down assay was used to address whether the TRD of MBD1 has affinity for core histones in solution. In this assay, immobilised GST-MBD1 (different regions) fusion was incubated with highly purified calf thymus core histones in the presence of BSA (to reduce non-specific interaction). After extensive washing of the GST fusion-sepharose beads, bound proteins were eluted by boiling in SDS loading buffer, and analysed by SDS-PAGE. It was found that although different C-terminal fragments containing TRD did not bind to core histones, a fusion containing the CxxC domain can bind directly to core histones (predominantly histone H3 and H4, see Figure 5.2.10 A). A mapping study using more fragments of MBD1 was performed. Two regions of MBD1 can interact directly with core histones (Figure 5.2.10 B and C). It was notable that when a strong binding was detected, all four core histones were bound. Only CxxC2 domain of MBD1 interacted with core histones, the CxxC1 domain of MBD1 and the CxxC domain of DNMT1 did not bind to core histones. The interaction is specific and not likely due to highly basic nature of core histones, because GST alone did not interact with histones, and different GST-MBD1 fusions did not interact with a highly basic protein lysozyme under the same condition (Figure 5.2.10 D). However, it is curious that full-length MBD1 did not interact with the histones (Figure 5.2.10 E). One plausible explanation is that the full-length protein is not folded in the correct conformation. Attempts to determine which core histones are the direct binding partners failed because certain histones bind non-specifically to GST protein (data not shown).

In order to test for interaction between full-length MBD1 and core histones, a HeLa nuclear extract was incubated with calf thymus core histones column. The column was washed extensively to remove unbound proteins and bound proteins were eluted in SDS loading buffer. Immunoblotting with anti-MBD1 antiserum showed that MBD1 was retained on the core histone column (Figure 5.2.10 F). MBD1 did not bind to control



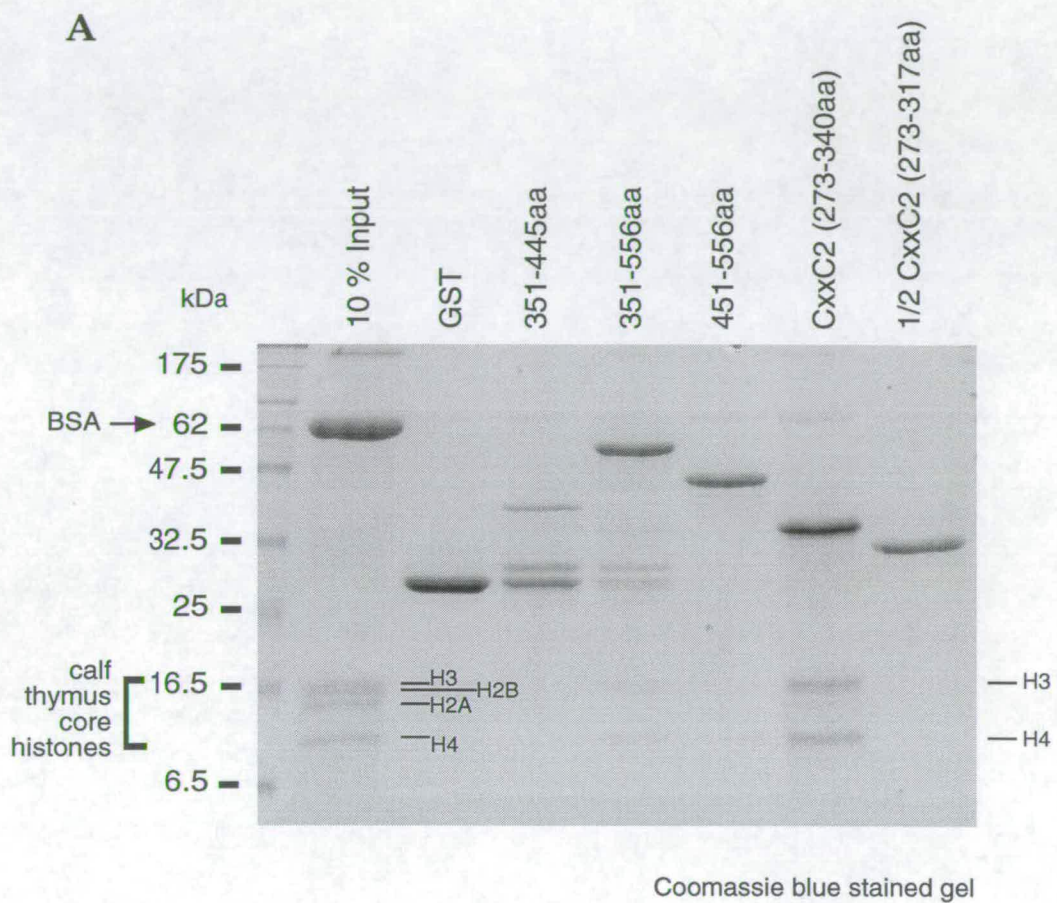


Figure 5.2.10. MBD1 can bind to core histones

A) CxxC2 domain of MBD1 (273-340 aa) interacts preferentially to histones H3 and H4 *in vitro*, with more than 10 % of input bound. GST control and several C-terminal fragments of MBD1 did not bind core histones.

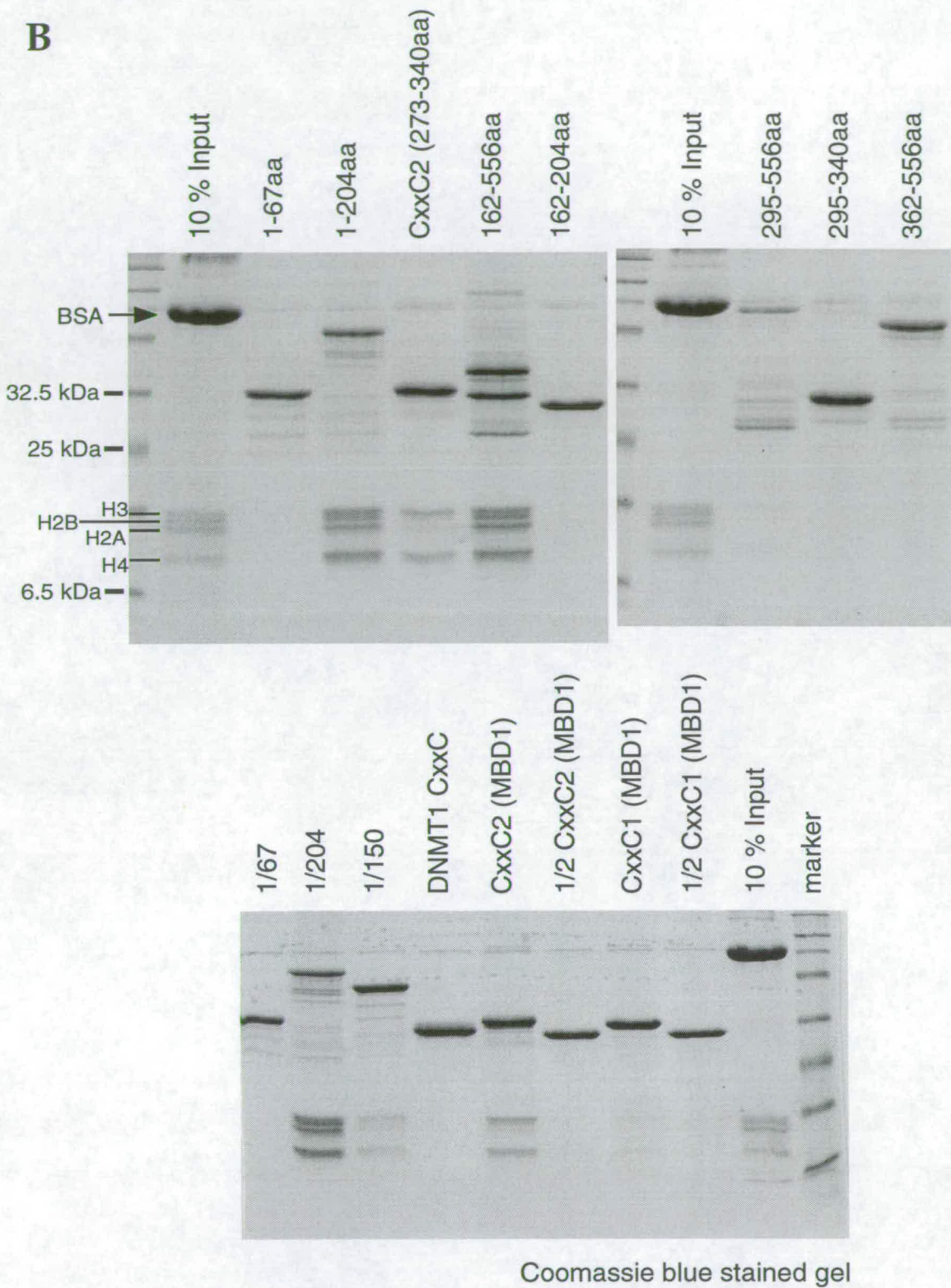
**B**

Figure 5.2.10. MBD1 binds to core histones

B) Mapping of core histones binding domains using different GST-MBD1 fusions. Only strong interaction (with recovery of more than 10 % of input) were scored as positive. A weak interaction (less than 5 % of input) was detected for GST-295-556aa and GST-CxxC1 fusion.



C

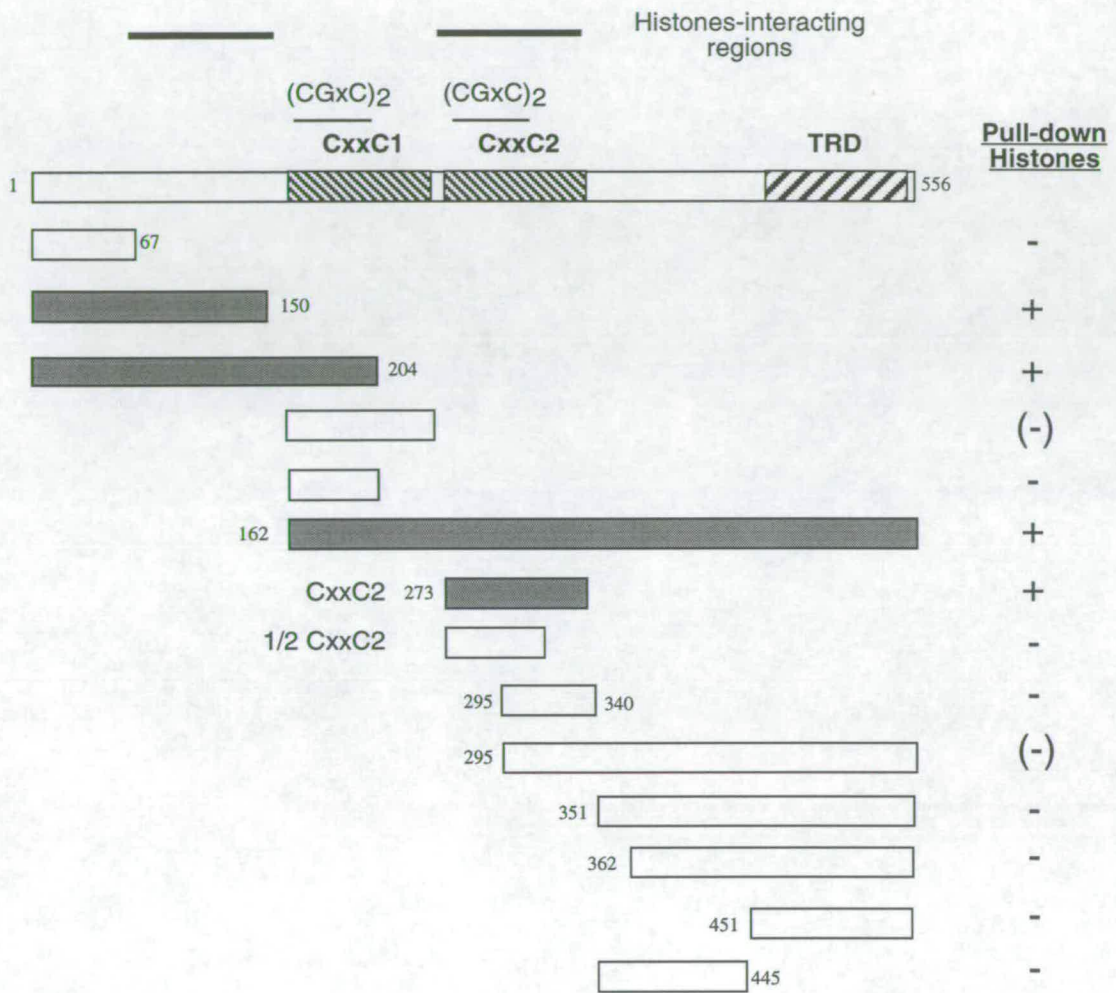
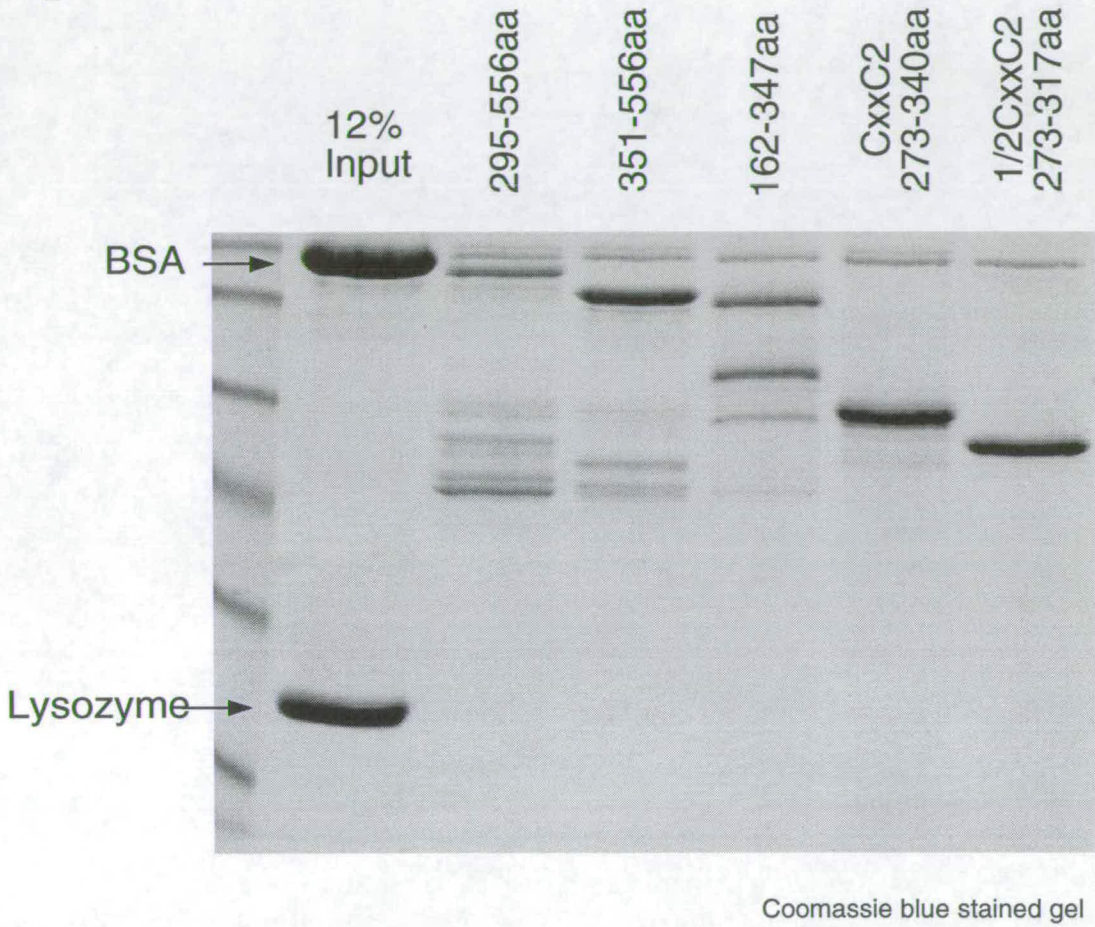


Figure 5.2.10. MBD1 binds to core histones

C) Summary of the interaction between core histones and different GST-MBD1 fusions. Only strong interaction (with recovery of more than 10 % of input) was scored as positive. Fragments of MBD1 that bound more than 10 % of input were highlighted in gray. (-) indicates weak binding. GST alone did not interact with core histones. The black lines above the diagram indicate the histone-interacting regions.

D



Coomassie blue stained gel

Figure 5.2.10. MBD1 binds to core histones

D) GST-MBD1 fusions did not interact with lysozyme (a 14.4 kDa highly basic protein from hen egg white).

E

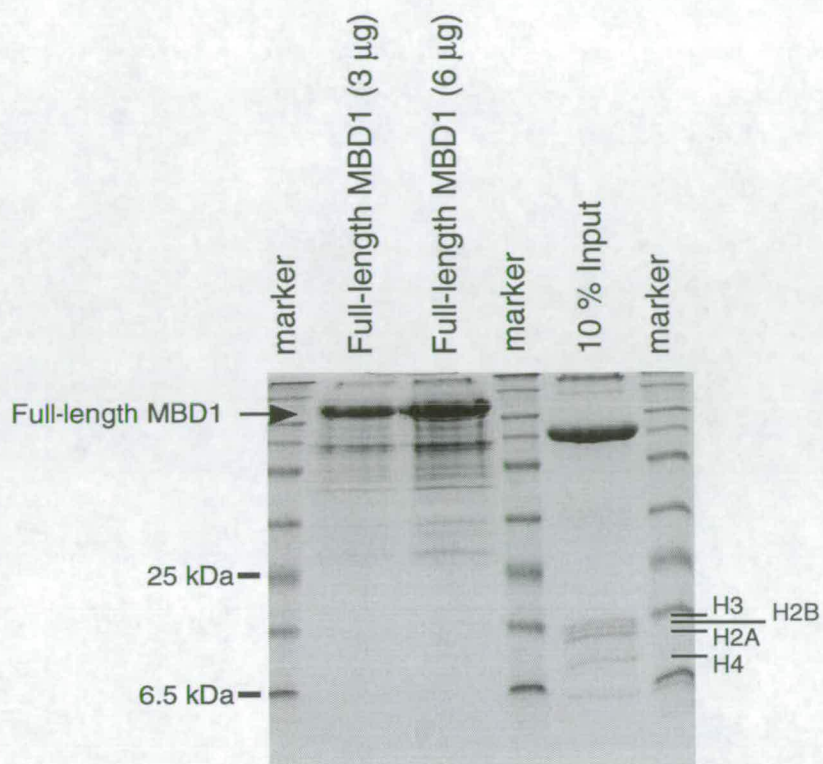


Figure 5.2.10. MBD1 binds to core histones

E) Core histones did not bind to recombinant full-length MBD1.



G

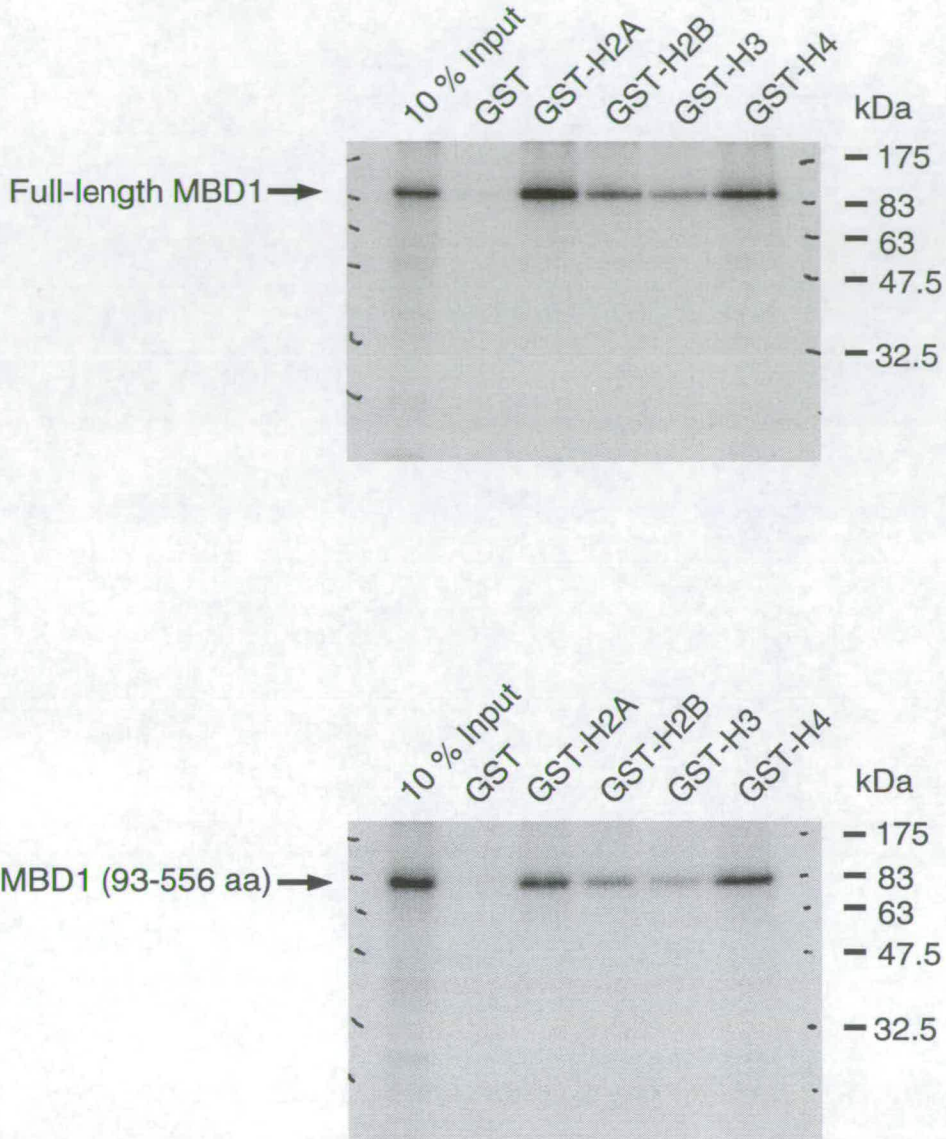


Figure 5.2.10. MBD1 binds to core histones

G) *In vitro* translated MBD1 binds to GST-core histones. Full-length MBD1 (top panel) or a truncated MBD1 without the methyl-CpG binding domain (bottom panel) were *in vitro* translated and incubated with GST-H2A, GST-H2B, GST-H3 or GST-H4. Bound proteins were eluted, fractionated on SDS-PAGE and detected by autoradiography.

GSH-sepharose column. The core histone column not incubated with HeLa nuclear extract did not contain MBD1.

Full-length MBD1 was synthesised *in vitro* and labeled with <sup>35</sup>S methionine. The labeled protein was then incubated with different histone GST-fusion proteins. MBD1 bound to all core histones but not to GST control (Figure 5.2.10 G). MBD1 without methyl-CpG binding domain (93-556 aa) exhibited similar binding activity towards all core histones, thereby excluding the possibility that the interaction is mediated by nucleic acids. However, it is possible that the interaction between MBD1 and histones may be indirect, and bridged by proteins present in the reticulocyte lysate.

These results established an *in vitro* interaction between MBD1 and core histones.

### **5.3 Conclusion**

An antibody highly specific towards MBD1 was used to evaluate the relationship between MBD1 and MeCP1. Anti-MBD1 antibody did not supershift or immunodeplete MeCP1 activity, suggesting that MBD1 does not contribute to MeCP1 bandshift activity derived from HeLa cells. The work described in Chapter 4 showed that another methyl-CpG binding protein, MBD2 is a component of MeCP1. MBD1 is however not associated with MBD2/ HDAC complex as revealed by co-immunoprecipitation experiments. In addition, MBD1 and MBD2 do not co-migrate on a gel-filtration column.

Although MBD1 is not a component of MeCP1, it can nevertheless repress transcription in a methylation dependent manner in F9 embryonic carcinoma cells. It is however possible that MBD1 is repressing transcription indirectly by blocking access of transcription factors to methylated sites. A GAL4 fusion repression assay was used to address this question. Fragments of MBD1 are fused to GAL4 DNA binding domain, and their effect on transcription can be assayed with reporter harboring Gal4 binding sites upstream of the promoter. A CxxC domain in MBD1 repressed the reporter with or without Gal4 binding sites, suggesting that this domain is toxic to transcription in a non-specific manner. A novel transcriptional repression domain (TRD) was also mapped to the C-terminal of MBD1. The TRD preferentially repress transcription of a reporter with Gal4 binding sites. Repression by the TRD can occur at a distance and several hydrophobic residues are important for the repression activity. The repression is sensitive to histone deacetylase inhibitor TSA, suggesting that histone deacetylase may be involved.

Immunolocalisation studies indicate that MBD1 is bound at sites enriched with 5-methylcytosine on human chromosomes (Barbin et al, 1994). The major sites are centric heterochromatin on chromosomes 1, 9, 15 and 16. Using GST-pulldown assay, an interaction between two regions of MBD1 and core histones was detected. MBD1 was also found to bind to core histone column. The result suggests a mechanism in which MBD1 might be anchored tightly onto chromatin.

In summary, functional analysis of MBD1 delineated a methyl-CpG binding domain, a CxxC domain which represses transcription independent of binding to the promoter and a C-terminal transcriptional repression domain. Like MeCP2, MBD1 is also a chromosomal proteins that is concentrated at centric heterochromatin.



## **Chapter 6: General discussion**

One well known consequence of CpG methylation in vertebrates is transcriptional silencing. As the vertebrate genomes are heavily methylated, 5-meC can be found in transcription factor recognition sites and impair the recruitment of transcription factors or basal transcription machinery to regulatory elements. However, many transcription factor recognition sites do not contain CpG and some transcription factors are not sensitive to methylation at their cognate binding sites, therefore indirect repression by methyl-CpG binding proteins is invoked to explain methylation induced repression. Indeed, methyl-CpG binding activities have been detected in mammalian extracts (MDBP1, Huang et al, 1984; MeCP1, Meehan et al, 1989; MDBP2, Pawlak et al, 1991; MeCP2, Lewis et al, 1992).

Apart from its preference for methylation, MDBP1 binds DNA in a sequence specific manner, with recognition site of more than two basepairs. Furthermore, its binding activity does not exhibit strict requirement for 5-meC. Therefore, MDBP1 is unlikely to mediate global repression by methylation. MDBP2 is an avian histone H1 subtype, and binding to DNA requires a minimum length of 30 bp with a symmetrically methylated CpG. However, whether mammalian histone H1 has similar affinity for methylated DNA is still unclear in light of opposing findings (Nightingale and Wolffe, 1995; Campoy et al, 1995; Levine et al, 1993; Johnson et al, 1995; McArthur and Thomas, 1996). On the other hand, MeCP1 is shown to bind specifically to methylated DNA, and binding affinity is dependent on methylation density (Meehan et al, 1989). The protein(s) responsible for this activity was not known until recently (see Chapter 4 of this thesis). MeCP2 is a second DNA binding activity discovered after MeCP1 (Lewis et al, 1992). Cloning of the cDNA for MeCP2 provides a headway for the characterisation of this methyl-CpG binding protein and understanding of the biological consequences of DNA methylation. Gene targeting experiments demonstrate that MeCP2 is essential for embryonic development (Tate et al, 1996). Chimaeric embryos derived from ES cells with MeCP2 gene disrupted, failed to develop normally and did not survive beyond weaning stage. Functional analysis of MeCP2 resulted in the mapping of a methyl-CpG binding domain (MBD) and a transcriptional repression domain (Nan et al, 1993; Nan et al, 1997).

The EST database screening effort has been fruitful in identifying four novel mammalian proteins (MBD1 to 4) with MBD-like region (Cross et al, 1997; Hendrich and Bird,

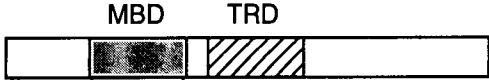
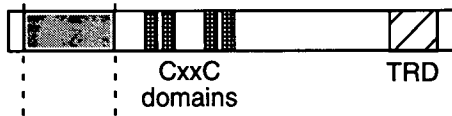


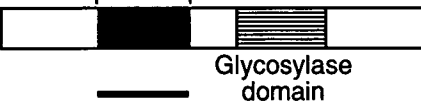
1998). Previous work from this laboratory showed that MBD1, MBD2 and MBD4 bind specifically to methylated DNA.

This thesis addresses the fundamental question of how DNA methylation represses transcription. The work described in this thesis provide insight into the mechanisms of repression of three of these methyl-CpG binding proteins (MeCP2, MBD2 and MBD1). MBD4 is shown to be a T/G glycosylase and therefore unlikely to be involved in mediating silencing by DNA methylation (Hendrich et al, 1999).

Table 6.1 summarises the current understanding of how methyl-CpG binding proteins repress transcription.

MeCP2, MBD2 and MBD1 can actively repress transcription when recruited to sites upstream of the promoter. MeCP2 interacts with the HDAC1/2 containing mSin3 co-repressor complex, and MBD2/ MeCP1 associates with a HDAC1/2 containing complex which is distinct from the mSin3 co-repressor complex and the Mi2 containing NuRD complex. On the other hand, MBD1 is not associated with a HDAC1/2 containing complex. However, TSA sensitivity of repression by TRD of MBD1 suggests that deacetylation may be involved. Collectively, these studies indicate that DNA methylation attracts a group of methyl-CpG binding transcriptional repressors, two of which are shown to associate with histone deacetylases.

A causal link between DNA methylation and histone deacetylation is further strengthened by studies that show parallel effects of DNA demethylating agents and TSA in reactivating transcription of sequences repressed by DNA methylation. In one study, mouse fibroblast L cells are stably transfected with either a methylated or a non-methylated tk gene (Eden et al, 1998). Subsequent immunoprecipitation of chromatin with an antibody against acetylated histone H4 revealed that, in contrast to the non-methylated tk gene, the methylated tk gene was not enriched for acetylated histones. TSA treatment results in an increased expression of the methylated tk gene. In addition, TSA treatment can increase the accessibility of the methylated tk gene locus to nucleases, suggesting that acetylation of histones re-establishes features of transcriptionally active chromatin despite the presence of DNA methylation. Similar results are obtained with integrated transgenes that are subjected to spontaneous epigenetic silencing with the concomitant appearance of DNA methylation and loss of histone hyperacetylation (Pikaart et al, 1998). Again, both butyrate/TSA and 5-aza-C are able to restore transgene expression.

		<sup>3</sup> H CpG Binding Hendrich & Bird(1998)	Comments	References
MeCP2		+	Transcriptional repressor Interacts with mSin3/HDAC complex	Nan/Ng et al (1998) Jones et al (1998)
MBD1		+	Represses transcription <i>in vitro</i> and <i>in vivo</i> Contains a transcriptional repression domain	Cross et al (1997) Fujita et al (1999) Ng et al
MBD2		+	Contains demethylase activity Failure to reproduce demethylase activity  Transcriptional repressor Associates with HDACs DNA binding component of MeCP1	Bhattacharya et al (1999) Ng et al (1999) Wade et al (1999)  Ng et al (1999)
MBD3		-	Integral component of NuRD complex	Zhang et al (1999) Wade et al (1999)
MBD4		+	Endonuclease, interacts with MLH1 Mismatch-specific T/U DNA glycosylase	Bellacosa et al (1999) Hendrich et al (1999)


 MBD-like region

Table 6.1. Functional properties of MBD/MeCP proteins

## **6.1 Outstanding questions:**

### **1) How do the MeCP/ HDAC complexes repress transcription?**

Work on MeCP2 and MBD2/ MeCP1 has provided insight as to how the covalent marking on cytosine can trigger further modification of the chromatin, and presumably lead to transcriptional silencing. Several mechanistic questions need to be addressed: How does histone deacetylation lead to transcriptional silencing? Can MeCPs repress transcription through HDAC independent pathways?

Studies of RPD3 (yeast homolog of HDAC1/2) established that recruitment of histone deacetylases by Ume6 transcriptional repressor to specific sequences leads to local deacetylation of histone tails *in vivo* (Rundlett et al., 1998; Kadosh and Struhl, 1998). This short range modification of chromatin seems sufficient to extinguish transcription. It remains to be demonstrated whether MeCP/ HDAC complexes exert a similar effect on a methylated nucleosomal template. MeCP2 has been shown to exert long distance repression on a promoter in a transient transfection assay; is the interaction with histone deacetylase sufficient to account for the repression? The repression by TRD of MeCP2 was not completely relieved by TSA. In the case of MBD2, repression on certain promoters is not sensitive to TSA. Therefore, it is very likely that these MeCPs can also repress transcription in a deacetylation-independent manner. Isolation of novel nuclear proteins that interact with these MeCPs is likely to shed light on this issue.

### **2) What are the biological roles of these methyl-CpG binding repressors?**

Question also arises as to whether MBD1, MBD2/MeCP1 and MeCP2 have redundant functions in the repression of methylated sequences. There are differences between these MeCPs. MBD1 and MeCP2 are both chromosomal proteins, whereas MBD2/ MeCP1 seems to be loosely bound in the nucleus (based on low salt extractability from the nucleus; Meehan et al, 1992). MeCP1 has a strong preference for high methyl-CpG density *in vitro*, therefore an untested possibility is whether MBD2 binds cooperatively to multiple methyl-CpGs or the MeCP1 complex contains multiple MBD2 molecules resulting in high affinity for densely methylated DNA. Is there tissue-specific expression of these MeCPs? Preliminary data suggests that this may be the case. For example, rat

brain nuclear extracts have high levels of MeCP2 but low levels of MBD2 (see Figure 4.2.3).

It has been shown that under some circumstances, methylation *per se* is not sufficient for transcriptional silencing, and repression on methylated templates is concomitant with packaging of the naked DNA into nucleosomes (Buschhausen et al., 1987; Kass et al., 1996). It is uncertain which MeCP is involved in the chromatin assembly coupled repression. Hence, the physiological substrates for each of the MeCPs will be crucial to distinguish their functional differences and biological role.

The phenotype of the mouse knockout of MeCP2 resembles that of DNMT1, and indicates that other MeCPs/ MBDs are not sufficient to substitute for MeCP2 (Tate et al., 1996). It is hopeful that gene targeting of the other member of the MBD family will provide further insights into the biological roles of these proteins. Important questions that can be answered through gene targeting experiments include:

- a) Are these proteins essential for mammalian development?
- b) Is gene expression mis-regulated in the absence of MeCP?
- c) Which are the target genes of these MeCPs?

Recently, it is found that the Rett syndrome, a progressive neurodevelopmental disorder, is caused by mutations within the MECP2 gene (Amir et al, 1999). These patients suffer from gradual loss of speech, and develop microcephaly, seizures, autism and ataxia. The finding underlines the importance of MeCP2, and deregulation is likely to bring about disease state.

The gene for MBD2 has been disrupted in mouse (B. Hendrich, personal communication). Interestingly, cell-lines derived from the knockout mice failed to repress transfected methylated constructs efficiently, and did not contain MeCP1 activity. These findings are consistent with the result showing that MBD2 is a major determinant of MeCP1 and is a methyl-CpG binding transcriptional repressor. Such a mouse model is invaluable for dissecting the biological role of MBD2.

An experimental tool that is currently available to study the *in vivo* substrates of these MeCPs is the chromatin immunoprecipitation assay (Aparicio et al, 1997; Tanaka et al, 1997; Alberts et al, 1998; Kadosh and Struhl, 1998b; Rundlett et al, 1998; Tanaka et al, 1998; Cosma et al, 1999). This technique involves fixation of nuclear proteins on the

DNA using formaldehyde, sonication of the chromatin to small fragments, and finally enrichment of chromatin bound to specific protein by immunoprecipitation. As MeCP2 and MBD1 are shown to be stably bound onto mitotic chromosomes, one can isolate and analyse the sequences that are bound by MeCP2 or MBD1 *in vivo*.

Transiently transfected methylated constructs in mammalian cells are presumably repressed by endogenous MeCPs. Using the chromatin immunoprecipitation assay, it is possible to address the following questions:

- a) which MeCP(s) is associated with the methylated DNA?
- b) are HDACs recruited to the methylated DNA?
- c) are deacetylated histones associated with the methylated DNA?
- d) is the promoter occupied by basal transcriptional machinery under the repressed state?
- e) if MBD2 is bound to the transfected methylated DNA, is the NuRD complex associated with MBD2 on the DNA?
- f) in MBD2 null cell-lines, can the NuRD complex be targeted to the transfected methylated DNA?

### **3) How can the inactive state of methylated nucleosomal DNA be reversed?**

It is of interest to understand how the transcriptional machinery gains access to specific loci (e.g. tissue-specific genes) amidst of the bulk of the methylated genome. A recent report suggests how a densely methylated CpG island can be activated (Ferguson et al, 1998). The progesterone receptor (PR) gene is an estrogen-responsive gene and in certain human breast cancers, DNA methylation at the CpG island of this gene can lead to transcriptional inactivity. Treatment with 5-aza-dC can reactivate the PR gene through DNA demethylation and activation of the ER gene. Intriguingly, it is shown that the ectopic expression of ligand-bound estrogen receptor (ER) can reactivate the PR gene even in the presence of DNA methylation. Most importantly, induction of the silenced methylated PR gene by ER was blocked by co-expression of a dominant negative form of SRC-1A. SRC-1 is a coactivator with an intrinsic histone acetyltransferase activity, and can potentiate the activity of ligand-bound ER (McInerney et al, 1996; Spencer et al, 1997). This result demonstrates that the association between ER and coactivator is crucial in mediating the induction of the methylated PR gene. Therefore, a possible mechanism to



overcome the effect of DNA methylation is to recruit proteins with histone acetyltransferase activity that can reverse the state of histone acetylation (Figure 6.1).

## **6.2 MeCPs: proteins with no footprint**

It has been reported that the CpG island of human PGK-1 gene on the inactive X-chromosome is heavily methylated but no footprint was detected over the methylated region (Pfeifer et al, 1990). This suggests that MeCPs are not bound to methylated sequences *in vivo*. A possible explanation for this observation is that the occupancy of MeCP at a particular methylated site is not saturated, i.e. 100 %. Indeed, there are more methyl-CpGs in the genome ( $2 \times 10^7$ ) compared to the number of molecules of MeCPs ( $5 \times 10^6$  MeCP2 in a rat brain cell, and  $2 \times 10^4$  MBD1 in a HeLa cell). Alternatively, the binding of MeCPs on chromatin is transient, as suggested by the loose interaction of MeCP1 with chromatin based on the ability to extract MeCP1 from the nucleus using low salt (Meehan et al, 1992). This is unlikely to be the case for MeCP2 and MBD1 which are stably bound to chromosomes. Both MeCP2 and MBD1 have potent transcriptional repression domains that can mediate long range repression, therefore such mechanism may suffice to effectively repress methylated sequences in the genome.

## **6.3 The mojo of MBD2**

An interesting twist to the MBD2 story comes from a report claiming that MBD2 contains 5-methyl cytosine demethylase activity (Bhattacharya et al., 1999). MBD2b was shown to catalyse the conversion of methyl-CpG to CpG by the direct removal of methyl-group. The same group also studied the end products of the demethylation process using a purified DNA demethylase from human cells (Ramchandani et al, 1999). Although the authors did not show that the purified DNA demethylase is MBD2, they presented evidence that the demethylase hydrolyses 5-methyl cytosine to cytosine and methanol. As pointed out by Wolffe et al. (1999), the newly discovered MBD2 demethylase and a yet unknown demethylase from human cells have remarkable biochemical and unprecedented catalytic properties. However it is unfortunate that the MBD2b demethylase activity cannot be reproduced by two laboratories, namely this laboratory and Dr. Wolffe's laboratory (Ng et al, 1999; Wade et al, 1999).

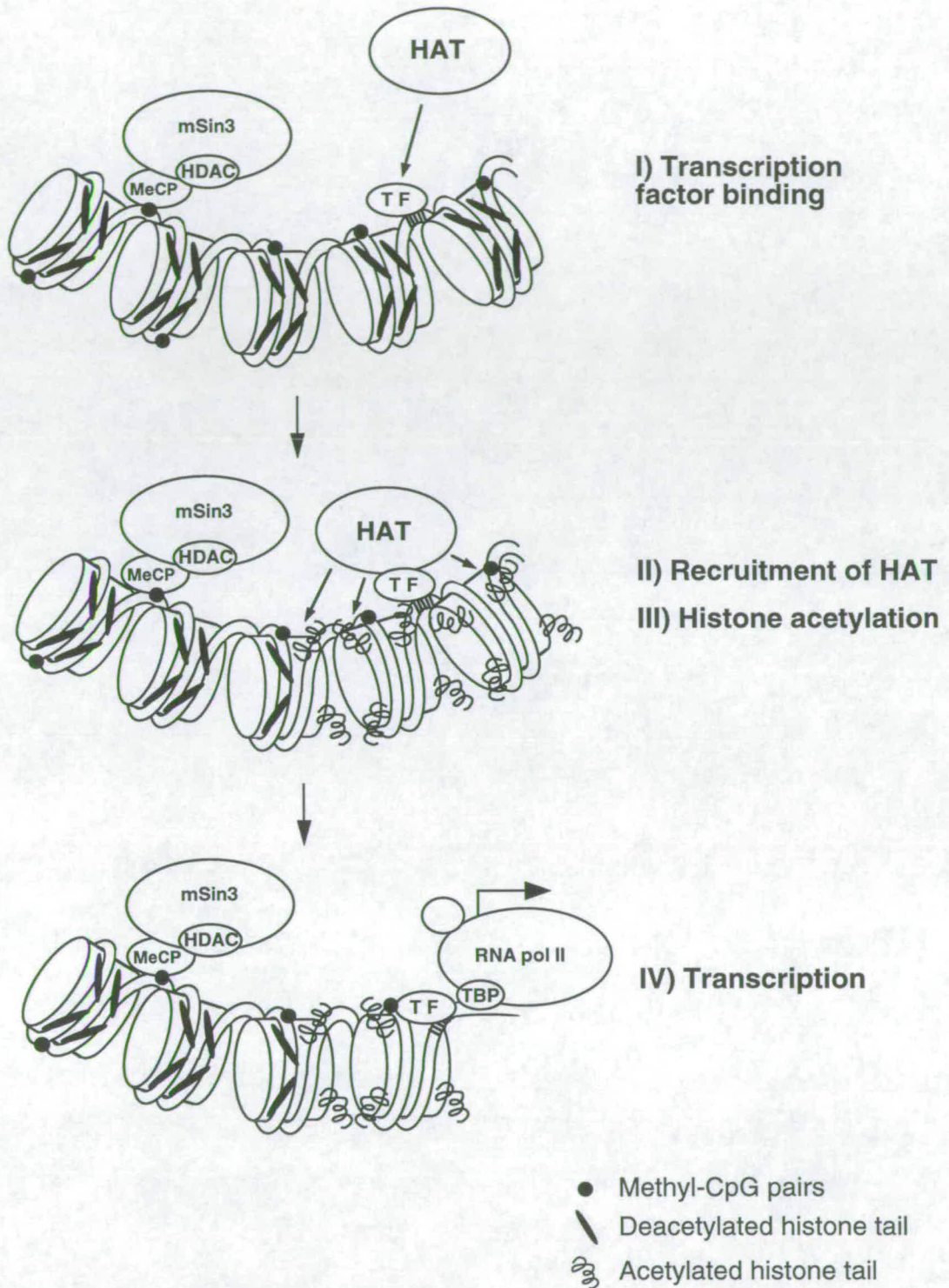


Figure 6.1. How can the inactive state of methylated nucleosomal DNA be reversed?

HAT, histone acetyltransferase. TF, transcription factor.

#### **6.4 Is MBD3 a methyl-CpG binding protein?**

MBD2 and MBD3 are related proteins with highly similar C-terminus regions. Despite the similarity between these two proteins, only MBD2 can bind specifically to methylated DNA both *in vitro* and *in vivo* (Hendrich and Bird, 1998). It is likely that differences within the MBD-like region contribute to the disparity in their DNA binding property. The finding that xMBD3 which is a component of the *Xenopus* NuRD complex, binds specifically to methylated DNA instigates the reinvestigation of whether the mammalian MBD3 is a methyl-CpG binding protein. Both MBD3 expressed in *E.coli* and insect cells showed a clear preference for methylated DNA using bandshift assays. This binding activity was also dependent on the MBD-like region, as a truncated version of MBD3 with half the MBD-like region did not bind methylated DNA. However, competition experiments using non-methylated and methylated competitor DNA failed to reveal a clear answer regarding the specificity of the DNA/ protein complex. In the light of equivocal data, additional assays need to be developed to address this question critically.

#### **6.5 Repression by DNA methylation: more to chromatin than meet the eye**

In the cell, the likely situation is that different genes are regulated differently, i.e. not all genes will be repressed by methylation. It is also likely that methylation-sensitive genes are controlled through both direct and indirect repression by DNA methylation.

A good example is the mouse M-lysozyme gene. The myeloid-specific mouse M-lysozyme gene is methylated and silenced in non-myeloid cells (Figure 6.2). The myeloid-specific expression is primarily regulated by a downstream enhancer (MLDE, Klages et al, 1992). The MLDE contains the binding site for ubiquitous GABP heterotetrameric transcription factor. The binding of GABP to the recognition site is inhibited by methylation at a single HpaII site, thus in non-myeloid cells, DNA methylation inhibits the tissue-specific enhancer activity (Nickel et al, 1995). During myeloid differentiation, the MLDE becomes demethylated and M-lysozyme expression is activated (Schmitz et al, 1997). Treatment of non-expressing cells such as the mouse fibroblasts, lymphocytic T-cells and myeloid precursor cells with demethylating agent 5-aza-dC induces demethylation and restores the *in vivo* footprint at GABP site. Interestingly, when these cells are treated with TSA, the expression of M-lysozyme gene is induced even though it remains methylated and GABP is not bound to MLDE *in vivo*

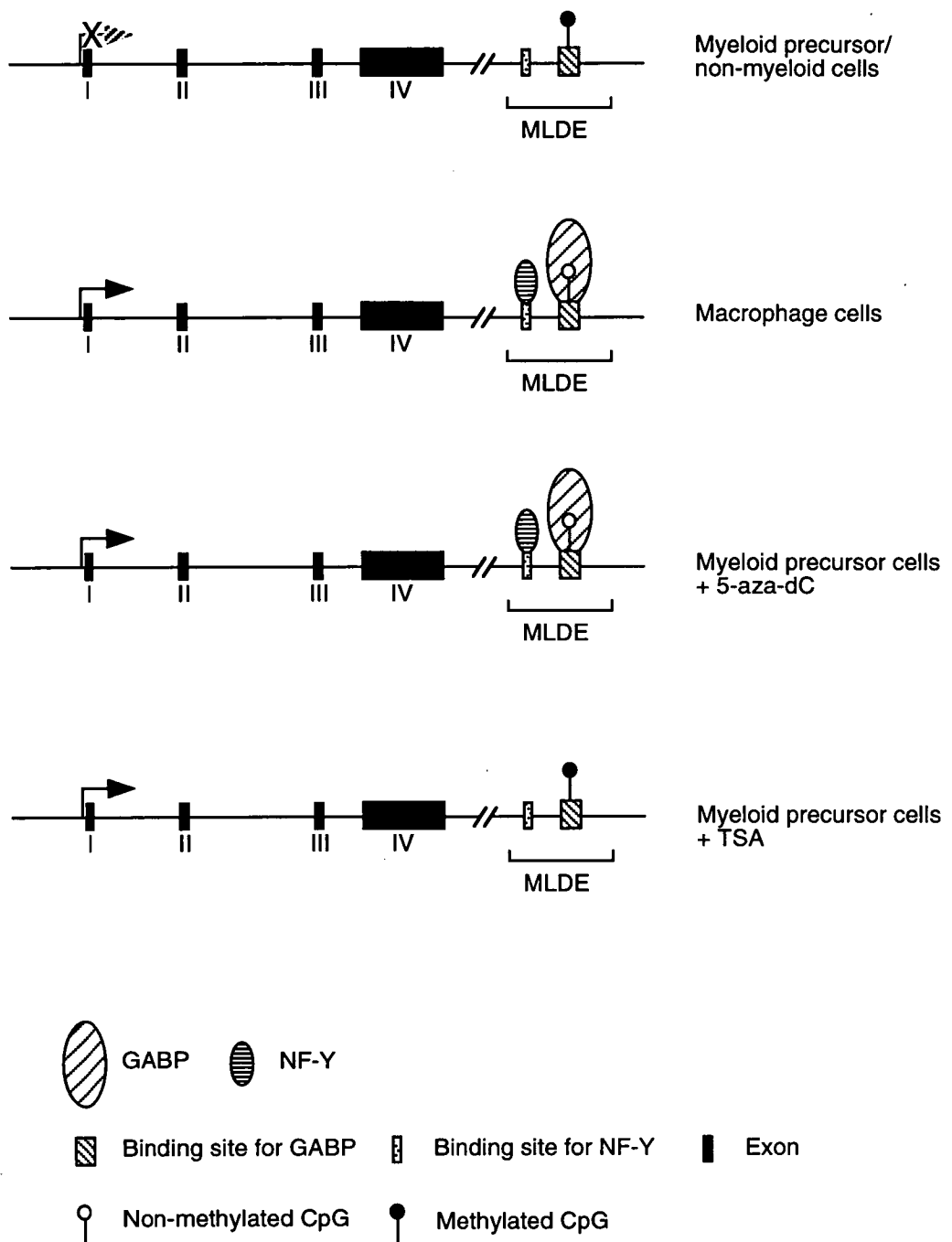


Figure 6.2. Regulation of macrophage-specific (M-lysozyme) gene expression by DNA methylation: direct and indirect repression.

MLDE, M-lysozyme downstream enhancer. Only a single methylated site within the MLDE is shown.

(Ammerpohl et al, 1998). The level of induction of M-lysozyme gene expression was comparable with that achieved by 5-aza-dC. The results strongly suggest that both direct and indirect mechanisms are involved in repressing expression of M-lysozyme gene.

In two other reports, TSA failed to reactivate hypermethylated, transcriptionally silenced tumour suppressor genes in tumour cells, and expanded CGG repeat of *FMR1* in fragile X patients (Cameron et al, 1999; Coffee et al, 1999). When the tumour cells were treated with 5-aza-dC in conjunction with TSA, a synergistic and robust reactivation of the hypermethylated genes was observed. In the case of methylated *FMR1*, TSA was shown to reverse the acetylation status of histone H4, but transcription was not restored. However, when the fragile X cells were treated with 5-aza-dC, histones H3 and H4 associated with *FMR1* became acetylated and transcription was restored. The result strongly suggest that both direct (methylation induced steric occlusion of regulatory factor binding) and indirect (MeCPs or chromatin mediated repression) effects were operating to repression transcription of methylated sequences.

## **6.6 Functions of DNA methylation?**

*In vitro* methylated genes transfected into mammalian cells become transcriptionally silenced. The correlation between methylation of genes and silencing are well documented in the mammalian system. 5-aza-dC or 5-aza-C treatment can reactivate previously methylated and silenced genes or proviruses. Several of the known mammalian methyl-CpG binding proteins (MeCP2, MBD2 and MBD1) are active transcriptional repressors (Nan et al, 1997; Nan et al, 1998; Ng et al, 1998; Ng et al). It leaves little room for doubt that DNA methylation in vertebrates is involved in repression of methylated sequences. Curiously, it is found that the genome of an invertebrate (*Ciona intestinalis*) is composed of methylated genes and non-methylated transposable elements (Simmen et al, 1999). Are the methylated genes repressed by DNA methylation? What is the function of DNA methylation in this organism?

Organisms with 5-meC have to bear with the mutagenic burden of using DNA methylation system. 5-meC can spontaneously deaminate to form thymine. As thymine is a naturally occurring base, a T/G mismatch is obviously a difficult lesion to be corrected by DNA repair system(s). An attractive candidate repair protein to deal with such mutagenic event is the DNA thymine glycosylase (Wiebauer and Jiricny, 1989; 1990). TDG preferentially excises the thymine base from a T/G mismatch *in vitro*. Intriguingly,

mammals also have a methyl-CpG binding protein with a T/G glycosylase activity (MBD4; Hendrich et al, 1999). So, the question is: what is the role of methyl-CpG binding in MBD4?

The studies of methyl-CpG binding proteins are pivotal in understanding the DNA methylation system in vertebrates and complementary molecular, biochemical and genetics tools will inevitably provide new exciting insights into the functions of DNA methylation.



## **References:**

- Alberts AS, Geneste O, Treisman R. (1998). Activation of SRF-regulated chromosomal templates by Rho-family GTPases requires a signal that also induces H4 hyperacetylation. *Cell*. **92**, 475-487.
- Alland L, Muhle R, Hou H Jr, Potes J, Chin L, Schreiber-Agus N, DePinho RA. (1997). Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. *Nature*. **387**, 49-55.
- Amir RE, Van Den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23**, 185-188.
- Antequera F, Bird AP. (1988). Unmethylated CpG islands associated with genes in higher plant DNA. *EMBO J.* **7**, 2295-2299.
- Antequera F, Macleod D, Bird AP. (1989). Specific protection of methylated CpGs in mammalian nuclei. *Cell*. **58**, 509-517.
- Aparicio OM, Weinstein DM, Bell SP. (1997). Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell*. **91**, 59-69.
- Ayer DE, Lawrence QA, Eisenman RN. (1995). Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell*. **80**, 767-776.
- Barbin A, Montpellier C, Kokalj-Vokac N, Gibaud A, Niveleau A, Malfoy B, Dutrillaux B, Bourgeois CA. (1994). New sites of methylcytosine-rich DNA detected on metaphase chromosomes. *Hum. Genet.* **94**, 684-692.
- Beard C, Li E, Jaenisch R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev.* **9**, 2325-2334.

- Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golemis EA, Genuardi M, Neri G. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proc. Natl. Acad. Sci. USA*. **96**, 3969-3974.
- Bestor TH, Ingram VM. (1983). Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl. Acad. Sci. USA*. **80**, 5559-5563.
- Bestor TH. (1988). Cloning of a mammalian DNA methyltransferase. *Gene*. **74**, 9-12.
- Bestor TH. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J*. **11**, 2611-2617.
- Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. *Nature*. **397**, 579-583.
- Bird AP, Taggart MH, Smith BA. (1979). Methylated and unmethylated DNA compartments in the sea urchin genome. *Cell*. **17**, 889-901.
- Bird AP. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucl. Acids Res*. **8**, 1499-1504.
- Bird A. (1992). The essentials of DNA methylation. *Cell*. **70**, 5-8.
- Bird AP, Taggart MH. (1980). Variable patterns of total DNA and rDNA methylation in animals. *Nucl. Acids Res*. **8**, 1485-1497.
- Blau J, Xiao H, McCracken S, O'Hare P, Greenblatt J, Bentley D. (1996). Three functional classes of transcriptional activation domain. *Mol. Cell. Biol*. **16**, 2044-2055.
- Boffa LC, Gershey EL, Vidali G. (1971). Changes of the histone deacetylase activity during chick embryo muscle development. *Biochim Biophys Acta*. **254**, 135-143.
- Boyes J, Bird A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*. **64**, 1123-1134.

- Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H. (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature*. **371**, 435-438.
- Brehm A, Miska EA, McCance DJ, Reid JL, Bannister AJ, and Kouzarides T. (1998). Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature*. **391**, 597-601.
- Brehm A, Nielsen SJ, Miska EA, McCance DJ, Reid JL, Bannister AJ, Kouzarides T. (1999). The E7 oncoprotein associates with Mi2 and histone deacetylase activity to promote cell growth. *EMBO J*. **18**, 2449-2458.
- Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*. **84**, 843-851.
- Bruhat A, Jost JP. (1996). Phosphorylation/dephosphorylation of the repressor MDBP-2-H1 selectively affects the level of transcription from a methylated promoter *in vitro*. *Nucl. Acids Res*. **24**, 1816-1821.
- Busslinger M, Hurst J, Flavell RA. (1983). DNA methylation and globin gene expression. *Cell*. **34**, 197-206.
- Buschhausen G, Wittig B, Graessmann, Graessmann A. (1987). Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc Natl Acad Sci USA*. **84**, 1177-1181.
- Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB. (1999). Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat. Genet*. **21**, 103-107.
- Campoy FJ, Meehan RR, McKay S, Nixon J, Bird A. (1995). Binding of histone H1 to DNA is indifferent to methylation at CpG sequences. *J. Biol. Chem*. **270**, 26473-26481.

Carlson LL, Page AW, Bestor TH. (1992). Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* **6**, 2536-2541.

Cervoni N, Bhattacharya S, Szyf M. (1999). DNA demethylase is a processive enzyme. *J. Biol. Chem.* **274**, 8363-8366.

Chandler SP, Guschin D, Landsberger N, Wolffe AP. (1999). The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. *Biochemistry.* **38**, 7008-7018.

Chuang LS, Ng HH, Chia JN, Li BF. (1996). Characterisation of independent DNA and multiple Zn-binding domains at the N terminus of human DNA-(cytosine-5) methyltransferase: modulating the property of a DNA-binding domain by contiguous Zn-binding motifs. *J. Mol. Biol.* **257**, 935-948.

Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science.* **277**, 1996-2000.

Chen ZJ, Pikaard CS. (1997). Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance. *Genes Dev.* **11**, 2124-2136

Coffee B, Zhang F, Warren ST, Reines D. (1999). Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. *Nat. Genet.* **22**, 98-101.

Cooper DN, Taggart MH, Bird AP. (1983). Unmethylated domains in vertebrate DNA. *Nucl. Acids Res.* **11**, 647-658.

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

Cross SH, Meehan RR, Nan X, Bird A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat. Genet.* **16**, 256-259.

Davey C., Pennings S., Allan J. (1997). CpG methylation remodels chromatin structure in vitro. *J. Mol. Biol.* **267**, 276-288.

Drotschmann K, Aronshtam A, Fritz HJ, Marinus MG. (1998). The Escherichia coli MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucl. Acids. Res.* **26**, 948-953.

Drummond JT, Li GM, Longley MJ, Modrich P. (1995). Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science.* **268**, 1909-1912.

Duncan BK, Miller JH. (1980). Mutagenic deamination of cytosine residues in DNA. *Nature.* **287**, 560-561.

Dunn DB, Smith JD. (1955). The occurrence of 6-methylcytosine in deoxyribonucleic acids. *Biochem J.* **68**, 627-636.

Eden S, Hashimshony T, Keshet I, Cedar H, Thorne AW. (1998). DNA methylation models histone acetylation. *Nature.* **394**, 842.

Edmondson DG, Smith MM, Roth SY. (1996). Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev.* **10**, 1247-1259.

Emili A, Greenblatt J, Ingles CJ. (1994). Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. *Mol. Cell. Biol.* **14**, 1582-1593.

Enver T, Zhang JW, Papayannopoulou T, Stamatoyannopoulos G. (1988). DNA methylation: a secondary event in globin gene switching? *Genes Dev.* **2**, 698-706.

Ferguson AT, Lapidus RG, Davidson NE. (1998). Demethylation of the progesterone receptor CpG island is not required for progesterone receptor gene expression. *Oncogene.* **17**, 577-583.

Finnegan EJ, Peacock WJ, Dennis ES. (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8449-8454.

Fitzpatrick, D. R., Shirley, K. M., McDonald, L. E., Bielefeldt-Ohmann, H., Graham, F. K. & Kelso, A. (1998). Distinct methylation of the Interferon  $\gamma$  (IFN- $\gamma$ ) and Interleukin 3 (IL-3) genes in newly activated primary CD8+ T lymphocytes: regional IFN- $\gamma$  promoter demethylation and mRNA Expression are heritable in CD44<sup>high</sup>CD8+ T cells. *J. Exp. Med.* **188**, 103-118.

Frederico LA, Kunkel TA, Shaw BR. (1990). A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry.* **29**, 2532-2537.

Fremont M, Siegmann M, Gaulis S, Matthies R, Hess D, Jost JP. (1997). Demethylation of DNA by purified chick embryo 5-methylcytosine-DNA glycosylase requires both protein and RNA. *Nucl. Acids Res.* **25**, 2375-2380.

Fujita N, Takebayashi Si, Okumura K, Kudo S, Chiba T, Saya H, Nakao M. (1999). Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. *Mol. Cell. Biol.* **19**, 6415-6426.

Gautsch JW, Wilson MC. (1983). Delayed de novo methylation in teratocarcinoma suggests additional tissue-specific mechanisms for controlling gene expression. *Nature.* **301**, 32-37.

Gelmetti V, Zhang J, Fanelli M, Minucci S, Pelicci PG, Lazar MA. (1998) Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Mol. Cell. Biol.* **18**, 7185-7191.

Giniger E, Varnum SM, Ptashne M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell.* **40**, 767-774.

Gold M, Hurwitz J. (1964). The enzymatic methylation of ribonucleic acid and deoxyribonucleic acid. V. purification and properties of the deoxyribonucleic acid-methylating activity in *Escherichia coli*. *J. Biol. Chem.* **239**, 3858-3865.



Grant SG, Chapman VM. (1988). Mechanisms of X-chromosome regulation. *Annu. Rev. Genet.* **22**, 199-233.

Grignani F, De Matteis S, Nervi C, Tomassoni L, Gelmetti V, Ciocce M, Fanelli M, Ruthardt M, Ferrara FF, Zamir I, Seiser C, Grignani F, Lazar MA, Minucci S, and Pelicci PG (1998). Fusion proteins of the retinoic acid receptor- $\alpha$  recruit histone deacetylase in promyelocytic leukemia. *Nature.* **391**, 815-818.

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

Gruenbaum Y, Szyf M, Cedar H, Razin A. (1983). Methylation of replicating and post-replicated mouse L-cell DNA. *Proc. Natl. Acad. Sci. USA.* **80**, 4919-4921.

Hall MC, Matson SW. (1999). The Escherichia coli MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity. *J. Biol. Chem.* **274**, 1306-1312.

Harbers K, Schnieke A, Stuhlmann H, Jahner D, Jaenisch R. (1981). DNA methylation and gene expression: endogenous retroviral genome becomes infectious after molecular cloning. *Proc. Natl. Acad. Sci. USA.* **78**, 7609-7613.

Harrington MA, Jones PA, Imagawa M, Karin M. (1988). Cytosine methylation does not affect binding of transcription factor Sp1. *Proc. Natl. Acad. Sci. USA.* **85**, 2066-2070.

Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE. (1997). Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell.* **89**, 341-347.

Hassig CA, Tong JK, Fleischer TC, Owa T, Grable PG, Ayer DE, Schreiber SL. (1998). A role for histone deacetylase activity in HDAC1-mediated transcriptional repression. *Proc. Natl. Acad. Sci. USA.* **95**, 3519-3524.

- He LZ, Guidez F, Tribioli C, Peruzzi D, Ruthardt M, Zelent A, and Pandolfi PP. (1998). Distinct interactions of PML-RARalpha and PLZF-RARalpha with co-repressors determine differential responses to RA in APL. *Nat. Genet.* **18**, 126-135.
- Hebbes TR, Thorne AW, Crane-Robinson C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* **7**, 1395-1402
- Heinzel T, Lavinsky RM, Mullen TM, Soderstrom M, Laherty CD, Torchia J, Yang WM, Brard G, Ngo SD, Davie JR, Seto E, Eisenman RN, Rose DW, Glass CK, Rosenfeld MG. (1997). A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. *Nature.* **387**, 43-48.
- Hendrich B, Bird A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18**, 6538-6547.
- Hendrich B, Abbott C, McQueen H, Chambers D, Cross S, Bird A. (1999). Genomic structure and chromosomal mapping of the murine and human mbd1, mbd2, mbd3, and mbd4 genes. *Mamm. Genome.* **10**, 906-912.
- Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature.* **401**, 301-304.
- Holler M, Westin G, Jiricny J, Schaffner W. (1988). Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.* **2**, 1127-1135.
- Holliday R, Pugh JE. (1975). DNA modification mechanisms and gene activity during development. *Science.* **187**, 226-232.
- Honda BM, Candido PM, Dixon GH. (1975). Histone methylation. Its occurrence in different cell types and relation to histone H4 metabolism in developing trout testis. *J. Biol. Chem.* **250**, 8686-8689.
- Hotchkiss RD. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J. Biol. Chem.* **168**, 315-332.

- Hsieh CL. (1994). Dependence of transcriptional repression on CpG methylation density. *Mol. Cell. Biol.* **14**, 5487-5494.
- Hsieh CL. (1999). Evidence that protein binding specifies sites of DNA demethylation. *Mol. Cell. Biol.* **19**, 46-56.
- Huang LH, Wang R, Gama-Sosa MA, Shenoy S, Ehrlich M. (1984). A protein from human placental nuclei binds preferentially to 5-methylcytosine-rich DNA. *Nature.* **308**, 293-295.
- Huang Y, Myers SJ, Dingledine R. (1999). Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nat. Neurosci.* **2**, 867-872.
- Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JCF, Trent JM, Staudt LM, Hudson J Jr, Boguski MS, Lashkari D, Shalon D, Botstein D, Brown PO. (1999). The transcriptional program in the response of human fibroblasts to serum. *Science.* **283**, 83-87.
- Jeppesen P, Mitchell A, Turner B, Perry P. (1992). Antibodies to defined histone epitopes reveal variations in chromatin conformation and underacetylation of centric heterochromatin in human metaphase chromosomes. *Chromosoma.* **101**, 322-332.
- Jeppesen P, Turner BM. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell.* **74**, 281-289.
- Johnson CA, Goddard JP, Adams RL. (1995). The effect of histone H1 and DNA methylation on transcription. *Biochem. J.* **305**, 791-798.
- Jones PA, Taylor SM (1980). Cellular differentiation, cytidine analogs and DNA methylation. *Cell.* **20**, 85-93.
- Jones PA, Rideout WM 3d, Shen JC, Spruck CH, Tsai YC. (1992). Methylation, mutation and cancer. *Bioessays.* **14**, 33-36.

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**, 187-191.

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

Jost JP, 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. USA.* **89**, 9499-9503.

Jost JP, Siegmann M, Thiry S, Jost YC, Benjamin D, Schwarz S. (1999). A re-investigation of the ribonuclease sensitivity of a DNA demethylation reaction in chicken embryo and G8 mouse myoblasts. *FEBS Lett.* **449**, 251-254.

Kadosh D, Struhl K. (1997). Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell.* **89**, 365-71.

Kadosh D, Struhl K. (1998a). Histone deacetylase activity of Rpd3 is important for transcriptional repression *in vivo*. *Genes Dev.* **12**, 797-805.

Kadosh D, Struhl K. (1998b). Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin *in vivo*. *Mol Cell Biol.* **18**, 5121-5227.

Kass SU, Landsberger N, Wolffe AP. (1997). DNA methylation directs a time dependent repression of transcription initiation. *Curr. Biol.* **7**, 157-165.

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

Kijima M, Yoshida M, Sugita K, Horinouchi S, Beppu T. (1993) Trapoxin, an antitumor cyclic tetrapeptide, is an irreversible inhibitor of mammalian histone deacetylase. *J. Biol. Chem.* **268**, 22429-22435.

Kim J, Sif S, Jones B, Jackson A, Koipally J, Heller E, Winandy S, Viel A, Sawyer A, Ikeda T, Kingston R, Georgopoulos K. (1999). Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity*. **10**, 345-355.

Klages S, Mollers B, Renkawitz R. (1992). The involvement of demethylation in the myeloid-specific function of the mouse M lysozyme gene downstream enhancer. *Nucl. Acids Res.* **20**, 1925-1932.

Koipally J, Renold A, Kim J, Georgopoulos K. (1999). Repression by ikaros and aiolos is mediated through histone deacetylase complexes. *EMBO J.* **18**, 3090-3100.

Krieger DE, Levine R, Merrifield RB, Vidali G, Allfrey VG. (1974). Chemical studies of histone acetylation. Substrate specificity of a histone deacetylase from calf thymus nuclei. *J. Biol. Chem.* **249**, 332-334.

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**, 5492-5499.

Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, Roth SY, Allis CD. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*. **383**, 269-272.

Kuo MH, Zhou J, Jambeck P, Churchill ME, Allis CD. (1998). Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes *in vivo*. *Genes Dev.* **12**, 627-639.

Laemmli UK. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**, 680-685.

Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, Eisenman RN. (1997). Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell*. **89**, 349-356.

Laherty CD, Billin AN, Lavinsky RM, Yochum GS, Bush AC, Sun JM, Mullen TM, Davie JR, Rose DW, Glass CK, Rosenfeld MG, Ayer DE, Eisenman RN. (1998). SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. *Mol. Cell*. **2**, 33-42.

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*. **122**, 3195-3205.

Leonhardt H, Page AW, Weier HU, Bestor TH. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*. **71**, 865-873.

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

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

Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, Bird A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*. **69**, 905-914.

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

Libby PR. (1970). Activity of histone deacetylase in rat liver and Novikoff hepatoma. *Biochim Biophys Acta*. **213**, 234-236.

Lin RJ, Nagy L, Inoue S, Shao W, Miller WH Jr, and Evans RM. (1998). Role of the histone deacetylase complex in acute promyelocytic leukemia. *Nature*. **391**, 811-814.



Lin YS, Carey MF, Ptashne M, Green MR. (1988). GAL4 derivatives function alone and synergistically with mammalian activators *in vitro*. *Cell*. **54**, 659-664.

Liu Y, Oakeley EJ, Sun L, Jost JP. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucl. Acids Res.* **26**, 1038-1045.

Lock LF, Takagi N, Martin GR. (1987). Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell*. **48**, 39-46.

Luo RX, Postigo, AA, and Dean DC. (1998). Rb interacts with histone deacetylase to repress transcription. *Cell*. **92**, 463-473.

Lutterbach B, Westendorf JJ, Linggi B, Patten A, Moniwa M, Davie JR, Huynh KD, Bardwell VJ, Lavinsky RM, Rosenfeld MG, Glass C, Seto E, Hiebert SW. (1998). ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Mol. Cell. Biol.* **18**, 7176-7184.

Lyon MF. (1999). X-chromosome inactivation. *Curr. Biol.* **9**, R235-R237.

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

Macleod D, Clark VH, Bird A. (1999). Absence of genome-wide changes in DNA methylation during development of the zebrafish. *Nat. Genet.* **23**, 139-140.

Magnaghi-Jaulin L, Groisman R, Naguibneva I, Robin P, Lorain S, Le Villain JP, Troalen F, Trouche D, and Harel-Bellan A. (1998). Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature*. **391**, 601-605.

Martinez-Balbas MA, Dey A, Rabindran SK, Ozato K, Wu C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell*. **83**, 29-38.

- Matsuo K, Silke J, Georgiev O, Marti P, Giovannini N, Rungger D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. *EMBO J.* **17**, 1446-1453.
- McArthur M, Thomas JO. (1996). A preference of histone H1 for methylated DNA. *EMBO J.* **15**, 1705-1714.
- McInerney EM, Tsai MJ, O'Malley BW, Katzenellenbogen BS. (1996). Analysis of estrogen receptor transcriptional enhancement by a nuclear hormone receptor coactivator. *Proc. Natl. Acad. Sci. USA.* **93**, 10069-10073.
- Meehan RR, Lewis JD, McKay S, Kleiner EL, Bird AP. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell.* **58**, 499-507.
- Meehan RR, Lewis JD, Bird AP. (1992). Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucl. Acids Res.* **20**, 5085-5092.
- Modrich P. (1986). DNA mismatch correction. *Annu. Rev. Biochem.* **56**, 435-466.
- Mohandas T, Sparkes RS, Shapiro LJ. (1981). Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science.* **211**, 393-396.
- Monk M, Boubelik M, Lehnert S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development.* **99**, 371-382.
- Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL, Evans RM. (1997). Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. *Cell.* **89**, 373-380.
- Nan X, Meehan RR, Bird A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucl. Acids Res.* **21**, 4886-4892.
- Nan X, Tate P, Li E, Bird A. (1996). DNA methylation specifies chromosomal localization of MeCP2. *Mol. Cell. Biol.* **16**, 414-421.

Nan X, Campoy FJ, Bird A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*. **88**, 471-481.

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. **393**, 386-389.

Naveh-Many T, Cedar H. (1981). Active gene sequences are undermethylated. *Proc. Natl. Acad. Sci. USA*. **78**, 4246-4250.

Nickel J, Short ML, Schmitz A, Eggert M, Renkawitz R. (1995). Methylation of the mouse M-lysozyme downstream enhancer inhibits heterotetrameric GABP binding. *Nucl. Acids Res*. **23**, 4785-4792.

Nightingale K, Wolffe AP. (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**, 4197-4200.

Ng HH, Bird A. (1999). DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev*. **9**, 158-163.

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

Ng HH, Jeppesen P, Bird A. Active repression of methylated promoters by chromosomal protein MBD1. (submitted).

Nightingale K, Wolffe AP. (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**, 4197-2000.

Niwa O, Yokota Y, Ishida H, Sugahara T. (1983). Independent mechanisms involved in suppression of the Moloney leukemia virus genome during differentiation of murine teratocarcinoma cells. *Cell*. **32**, 1105-1113.

Norris DP, Patel D, Kay GF, Penny GD, Brockdorff N, Sheardown SA, Rastan S. (1994). Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell*. **77**, 41-51.

Noyer-Weidner M, Trautner TA (1993). Methylation of DNA in prokaryotes. p39-108. In DNA methylation: molecular biology and biological significance. Edited by: Jost JP and Saluz HP.

Okano M, Xie S, Li E. (1998a). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucl. Acids Res.* **26**, 2536-2540.

Okano M, Xie S, Li E. (1998b). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19**, 219-220.

Palombo F, Gallinari P, Iaccarino I, Lettieri T, Hughes M, D'Arrigo A, Truong O, Hsuan JJ, Jiricny J. (1995). GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science*. **268**, 1912-1914.

Panning B, Jaenisch R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev.* **10**, 1991-2002.

Papadopoulos N, Nicolaidis NC, Liu B, Parsons R, Lengauer C, Palombo F, D'Arrigo A, Markowitz S, Willson JK, Kinzler KW, et al. (1995). Mutations of GTBP in genetically unstable cells. *Science*. **268**, 1915-1917.

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

Pawlak A, Bryans M, Jost JP. (1991). An avian 40 KDa nucleoprotein binds preferentially to a promoter sequence containing one single pair of methylated CpG. *Nucl. Acids Res.* **19**, 1029-1034.

Pfeifer GP, Tanguay RL, Steigerwald SD, Riggs AD. (1990). In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and

inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev.* **4**, 1277-1287.

Pikaart MJ, Recillas-Targa F, Felsenfeld G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* **12**, 2852-2862.

Pogo BG, Allfrey VG, and Mirsky AE. (1966). RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proc. Natl. Acad. Sci. USA.* **55**, 805-812.

Prasad R, Yano T, Sorio C, Nakamura T, Rallapalli R, Gu Y, Leshkowitz D, Croce CM, Canaani E. (1995). Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. *Proc. Natl. Acad. Sci. USA.* **92**, 12160-12164.

Proffitt JH, Davie JR, Swinton D, Hattman S. (1984). 5-Methylcytosine is not detectable in *Saccharomyces cerevisiae* DNA. *Mol. Cell. Biol.* **4**, 985-988.

Ramchandani S, Bhattacharya SK, Cervoni N, Szyf M. (1999). DNA methylation is a reversible biological signal. *Proc. Natl. Acad. Sci. USA.* **96**, 6107-6112.

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

Razin A, Riggs AD. (1980). DNA methylation and gene function. *Science.* **210**, 604-610.

Rideout WM 3d, Coetzee GA, Olumi AF, Jones PA. (1990). 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science.* **249**, 1288-1290.

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

- Riggs AD, Xiong Z, Wang L, LeBon JM. (1998). Methylation dynamics, epigenetic fidelity and X chromosome structure. *Novartis Found Symp* 1998; 214:214-25.
- Ronemus MJ, Galbiati M, Ticknor C, Chen J, Dellaporta SL. (1996). Demethylation-induced developmental pleiotropy in *Arabidopsis*. *Science*. **273**, 654-657.
- Rossignol JL, Faugeron G. (1995). MIP: an epigenetic gene silencing process in *Ascobolus immersus*. *Curr. Top. Microbiol. Immunol.* **197**, 179-191.
- Rougier N, Bourc'his D, Gomes DM, Niveleau A, Plachot M, Paldi A, and Viegas-Pequignot E. (1998). Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108-2113.
- Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M. (1996). HDA1 and Rpd3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *Proc. Natl. Acad. Sci. USA.* **93**, 14503-14508.
- Rundlett SE, Carmen AA, Suka N, Turner BM, Grunstein M. (1998). Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3. *Nature.* **392**, 831-835.
- Sadowski I, Ma J, Triezenberg S, Ptashne M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. *Nature.* **335**, 563-564.
- Scanlan MJ, Chen YT, Williamson B, Gure AO, Stockert E, Gordan JD, Tureci O, Sahin U, Pfreundschuh M, Old LJ. (1998). Characterization of human colon cancer antigens recognized by autologous antibodies. *Int. J. Cancer.* **76**, 652-658.
- Schmitz A, Short M, Ammerpohl O, Asbrand C, Nickel J, Renkawitz R. (1997). Cis-elements required for the demethylation of the mouse M-lysozyme downstream enhancer. *J. Biol. Chem.* **272**, 20850-20856.
- Schreiber-Agus N, Chin L, Chen K, Torres R, Rao G, Guida P, Skoultschi AI, DePinho RA. (1995). An amino-terminal domain of Mx11 mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor SIN3. *Cell.* **80**, 777-786.



Selker EU, Fritz DY, Singer MJ. (1993). Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in *Neurospora*. *Science*. **262**, 1724-1728.

Selker EU, Stevens JN. (1983). DNA methylation at asymmetric sites is associated with numerous transition mutations. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8114-8118.

Shapiro HS. (1975). Content of 6-methylaminopurine and 5-methylcytosine in DNA. In *Handbook of biochemistry, selected data for molecular biology: nucleic acids*. (ed. G.D. Fasman). CRC Press, pp258-262.

Sheid B, Srinivasan PR, Borek E. (1968). Deoxyribonucleic acid methylase of mammalian tissues. *Biochemistry*. **7**, 280-285.

Shen, CKJ. (1984). DNA methylation and developmental regulation of eukaryotic globin gene transcription. In *DNA methylation: Biochemistry and biological significance* (ed A Razin, H Cedar, AD Riggs), pp 249-268. Springer-Verlag, New York.

Shen JC, Rideout WM 3rd, Jones PA. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucl. Acids. Res.* **22**, 972-976.

Shmookler Reis, R. J. & Goldstein, S. (1982) Interclonal variation in methylation patterns for expressed and non-expressed genes. *Nucl. Acids Res.* **10**, 4293-4300.

Siegfried Z., Eden S., Mendelsohn M., Feng X., Tsuberi BZ., Cedar H. (1999). DNA methylation represses transcription in vivo. *Nat. Genet.* **22**, 203-206.

Simmen MW, Leitgeb S, Charlton J, Jones SJ, Harris BR, Clark VH, Bird A. (1999). Nonmethylated transposable elements and methylated genes in a chordate genome. *Science*. **283**, 1164-1167.

Simpson VJ, Johnson TE, Hammen RF. (1986). *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucl. Acids Res.* **14**, 6711-6719.

- Smith DB, Johnson KS. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*. **67**, 31-40.
- Spencer TE, Jenster G, Burcin MM, Allis CD, Zhou J, Mizzen CA, McKenna NJ, Onate SA, Tsai SY, Tsai MJ, O'Malley BW. (1997). Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature*. **389**, 194-198.
- Stein R, Gruenbaum Y, Pollack Y, Razin A, Cedar H. (1982). Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proc. Natl. Acad. Sci. USA*. **79**, 61-65.
- Stewart CL, Stuhlmann H, Jahner D, Jaenisch R. (1982). De novo methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. *Proc. Natl. Acad. Sci. USA*. **79**, 4098-4102.
- Supakar PC, Weist D, Zhang DL, Inamdar N, Zhang XY, Khan R, Ehrlich KC, Ehrlich M. (1988). Methylated DNA-binding protein is present in various mammalian cell types. *Nucl. Acids Res.* **16**, 8029-8044.
- Swisher JF, Rand E, Cedar H, Marie Pyle A. (1998). Analysis of putative RNase sensitivity and protease insensitivity of demethylation activity in extracts from rat myoblasts. *Nucl. Acids. Res.* **26**, 5573-5580.
- Tanaka K, Appella E, Jay G. (1983). Developmental activation of the H-2K gene is correlated with an increase in DNA methylation. *Cell*. **35**, 457-465.
- Tanaka T, Knapp D, Nasmyth K. (1997). Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. *Cell*. **90**, 649-660.
- Tanaka T, Nasmyth K. (1998). Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases. *EMBO J.* **17**, 5182-5191.
- Tate P. (1996). MeCP2: a methyl-CpG binding protein essential for embryonic development in the mouse. PhD thesis. University of Edinburgh.

- Tate PH, Bird AP. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr. Opin. Genet. Dev.* **3**, 226-231.
- Tate P, Skarnes W, Bird A. (1996). The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nat. Genet.* **12**, 205-208.
- Taunton J, Hassig CA, Schreiber SL. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science.* **272**, 408-411.
- Tazi J, Bird A. (1990). Alternative chromatin structure at CpG islands. *Cell.* **60**, 909-920.
- Tong JK, Hassig CA, Schnitzler GR, Kingston RE, Schreiber SL. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature.* **395**, 917-921.
- Triezenberg SJ, Kingsbury RC, McKnight SL. (1988). Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* **2**, 718-729.
- Tweedie S, Charlton J, Clark V, Bird A. (1997). Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol. Cell. Biol.* **17**, 1469-1475.
- Urieli-Shoval S, Gruenbaum Y, Sedat J, Razin A. (1982). The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *FEBS Lett.* **146**, 148-152.
- van Holde. (1988). Chromatin. Springer Series in Molecular Biology. New York: Springer Verlag.
- Van den Wyngaert I, Sprengel J, Kass SU, Luyten WH. (1998). Cloning and analysis of a novel human putative DNA methyltransferase. *FEBS Lett.* **426**, 283-289.
- Vermaak D, Wade PA, Jones PL, Shi YB, Wolffe AP. (1999). Functional analysis of the SIN3-histone deacetylase RPD3-RbAp48-histone H4 connection in the *Xenopus* oocyte. *Mol. Cell. Biol.* **19**, 5847-5860.

- Verreault A, Kaufman PD, Kobayashi R, Stillman B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell*. **87**, 95-104.
- Verreault A, Kaufman PD, Kobayashi R, Stillman B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr. Biol.* **8**, 96-108.
- Wade PA, Jones PL, Vermaak D, Wolffe AP. (1998). A multiple subunit Mi-2 histone deacetylase from *Xenopus laevis* cofractionates with an associated Snf2 superfamily ATPase. *Curr. Biol.* **8**, 843-846.
- Wade PA, Geggion A, Jones PL, Ballestar E, Aubry F, Wolffe AP. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat. Genet.* **23**, 62-66.
- Walsh CP, Chaillet JR, Bestor TH. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116-117.
- Walsh CP, Bestor TH. (1999). Cytosine methylation and mammalian development. *Genes Dev.* **13**, 26-34.
- Wang H, Stillman DJ. (1993). Transcriptional repression in *Saccharomyces cerevisiae* by a SIN3-LexA fusion protein. *Mol. Cell. Biol.* **13**, 1805-1814.
- Wang J, Hoshino T, Redner RL, Kajigaya S, Liu JM. (1998). ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. *Proc. Natl. Acad. Sci. USA.* **95**, 10860-10865.
- Wang RY, Kuo KC, Gehrke CW, Huang LH, Ehrlich M. (1982). Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. *Biochim. Biophys. Acta.* **697**, 371-377.
- Wang RY, Zhang XY, Ehrlich M. (1986). A human DNA-binding protein is methylation-specific and sequence-specific. *Nucl. Acids Res.* **14**, 1599-1614.

- Weiss A, Keshet I, Razin A, Cedar H. (1996). DNA demethylation in vitro: involvement of RNA. *Cell*. **86**, 709-718.
- Wiebauer K, Jiricny J. (1989). In vitro correction of G.T mispairs to G.C pairs in nuclear extracts from human cells. *Nature*. **339**, 234-236.
- Wiebauer K, Jiricny J. (1990). Mismatch-specific thymine DNA glycosylase and DNA polymerase beta mediate the correction of G.T mispairs in nuclear extracts from human cells. *Proc. Natl. Acad. Sci. USA*. **87**, 5842-5845.
- Wigler M, Levy D, Perucho M. (1981). The somatic replication of DNA methylation. *Cell*. **24**, 33-40.
- Wilkinson CR, Bartlett R, Nurse P, Bird AP. (1995). The fission yeast gene *pmt1+* encodes a DNA methyltransferase homologue. *Nucl. Acids Res*. **23**, 203-210.
- Wolffe AP, Jones PL, Wade PA. (1999). DNA demethylation. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5894-5896.
- Worth L Jr, Bader T, Yang J, Clark S. (1998). Role of MutS ATPase activity in MutS,L-dependent block of in vitro strand transfer. *J. Biol. Chem*. **273**, 23176-23182.
- Wu C. (1997). Chromatin remodeling and the control of gene expression. *J. Biol. Chem*. **272**, 2871-2874.
- Wyatt GR. (1951). Recognition and estimation of 5-methylcytosine in nucleic acids. *Biochem J*. **48**, 581-584.
- Xue Y, Wong J, Moreno GT, Young MK, Cote J, Wang W. (1998). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol. Cell*. **2**, 851-861.
- Yoder JA, Bestor TH. (1998). A candidate mammalian DNA methyltransferase related to *pmt1p* of fission yeast. *Hum. Mol. Genet*. **7**, 279-284.

Zeleznik-Le NJ, Harden AM, Rowley JD. (1994). 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proc. Natl. Acad. Sci. USA.* **91**, 10610-10614.

Zhang XY, Supakar PC, Khan R, Ehrlich KC, Ehrlich M. (1989). Related sites in human and herpesvirus DNA recognized by methylated DNA-binding protein from human placenta. *Nucl. Acids Res.* **17**, 1459-1474.

Zhang Y, Iratni R, Erdjument-Bromage H, Tempst P, Reinberg D. (1997) Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell.* **89**, 357-364.

Zhang Y, LeRoy G, Seelig HP, Lane WS, Reinberg D. (1998a). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell.* **95**, 279-289.

Zhang Y, Sun ZW, Iratni R, Erdjument-Bromage H, Tempst P, Hampsey M, Reinberg D. (1998b). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell.* **1**, 1021-1031.

Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* **13**, 1924-1935.

Zhu, X., Deng, C., Kuick, R., Yung, R., Lamb, B., Neel, J. V., Richardson, B., Hanash, S. (1999). Analysis of human peripheral blood T cells and single-cell-derived T cell clones uncovers extensive clonal CpG island methylation heterogeneity throughout the genome. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 8058-8063.

## **Publication:**

Nan X\*, Ng **HH\***, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. (1999). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. 393, 386-389.

\* Joint first authorship

Ng **HH**, Bird A. (1999). DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev.* 9, 158-163.

Zhang Y, Ng **HH**, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* 13, 1924-1935.

Ng **HH**, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature Genet.* 23, 58-61.

Hendrich B, Hardeland U, Ng **HH**, Jiricny J, Bird A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*. 401, 301-304.

Ng **HH**, Jeppesen P, Bird A. (2000). Active repression of methylated genes by chromosomal protein MBD1. *Mol. Cell. Biol.* (In press)



# Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex

Xinsheng Nan<sup>\*,†</sup>, Huck-Hui Ng<sup>\*,†</sup>, Colin A. Johnson<sup>‡</sup>, Carol D. Laherty<sup>§</sup>, Bryan M. Turner<sup>‡</sup>, Robert N. Eisenman<sup>†</sup> & Adrian Bird<sup>\*</sup>

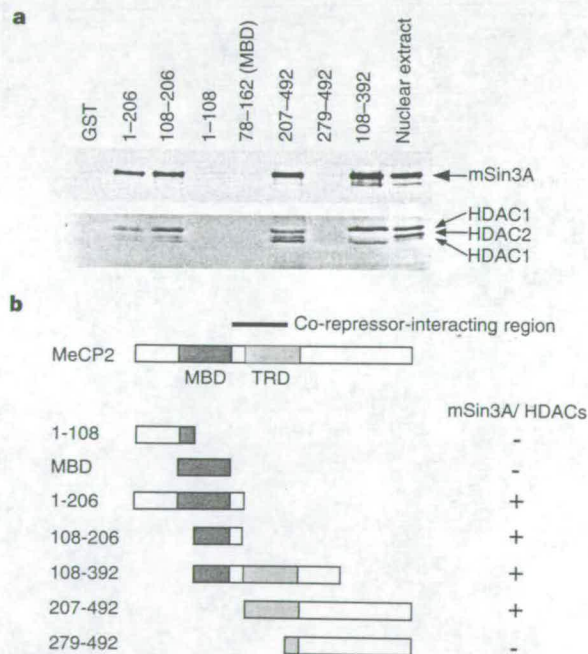
<sup>\*</sup> Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JR, UK

<sup>‡</sup> Department of Anatomy, The Medical School, University of Birmingham, Birmingham B15 2TT, UK

<sup>§</sup> Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, USA

<sup>†</sup> These authors contributed equally to this work.

Cytosine residues in the sequence 5' CpG (cytosine-guanine) are often postsynthetically methylated in animal genomes. CpG methylation is involved in long-term silencing of certain genes during mammalian development<sup>1,2</sup> and in repression of viral genomes<sup>3,4</sup>. The methyl-CpG-binding proteins MeCP1 (ref. 5) and MeCP2 (ref. 6) interact specifically with methylated DNA and mediate transcriptional repression<sup>7-9</sup>. Here we study the mechanism of repression by MeCP2, an abundant nuclear protein that is essential for mouse embryogenesis<sup>10</sup>. MeCP2 binds tightly to chromosomes in a methylation-dependent manner<sup>11,12</sup>. It contains a transcriptional-repression domain (TRD) that can

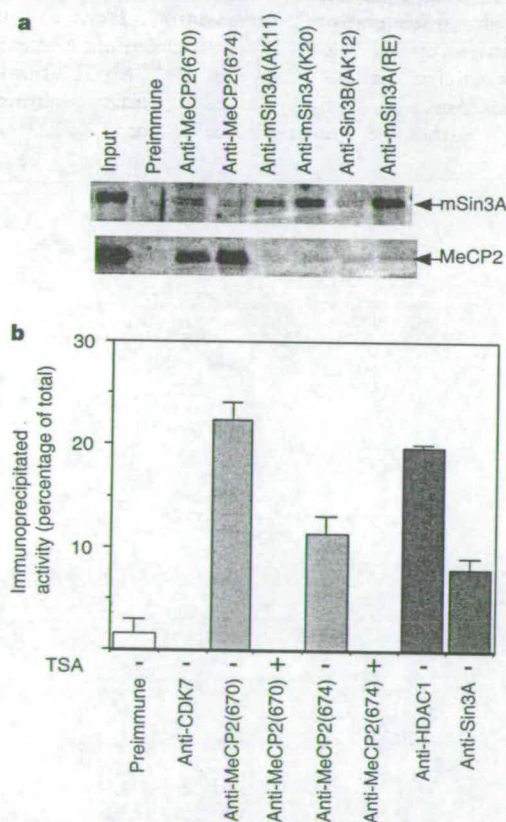


**Figure 1** Interaction of MeCP2 with mSin3A and histone deacetylases. **a**, Western blot analysis of HeLa nuclear proteins pulled down by immobilized GST or by GST-fusion proteins that included different regions of MeCP2 (see **b**). Numbers correspond to amino-acid positions in the protein. Blots were probed with antibodies against mSin3A (antibody K-20), or against HDAC1 and HDAC2. The lower HDAC1 band probably corresponds to a degradation product of HDAC1. The HDAC2 band in lane 108-392 is displaced upwards because a large amount of fusion protein migrates with the HDAC2. **b**, Map of regions of MeCP2 that do (+) or do not (-) pull down mSin3A and HDACs. The thick black line corresponds to the corepressor-interacting region defined by these experiments. MBD, methyl-CpG-binding domain<sup>11</sup>.



function at a distance *in vitro* and *in vivo*<sup>9</sup>. We show that a region of MeCP2 that localizes with the TRD associates with a corepressor complex containing the transcriptional repressor mSin3A and histone deacetylases<sup>13-19</sup>. Transcriptional repression *in vivo* is relieved by the deacetylase inhibitor trichostatin A<sup>20</sup>, indicating that deacetylation of histones (and/or of other proteins) is an essential component of this repression mechanism. The data suggest that two global mechanisms of gene regulation, DNA methylation and histone deacetylation, can be linked by MeCP2.

DNA methylation is associated with altered chromatin structures<sup>21-23</sup>. As MeCP2 binds tightly to chromatin in a methylation-dependent manner, it seemed possible that transcriptional repression may be due to recruitment by MeCP2 of a chromatin-modifying corepressor. A corepressor complex containing mSin3 and histone deacetylases (HDACs) is thought to exert its effects by modulating chromatin structure<sup>13-19</sup>. We therefore asked whether regions of MeCP2 could 'pull down' the mSin3-containing corepressor complex from nuclear extracts. We used proteins containing glutathione S-transferase (GST) fused to different regions of MeCP2 in 'pull-down' reactions, and probed proteins bound to these complexes with antibodies against mSin3A, HDAC1 and HDAC2. A region in the centre of MeCP2 could bind all three proteins (Fig. 1a). GST alone and regions outside the interacting region of MeCP2 did not pull down any of the corepressor proteins (Fig. 1a).

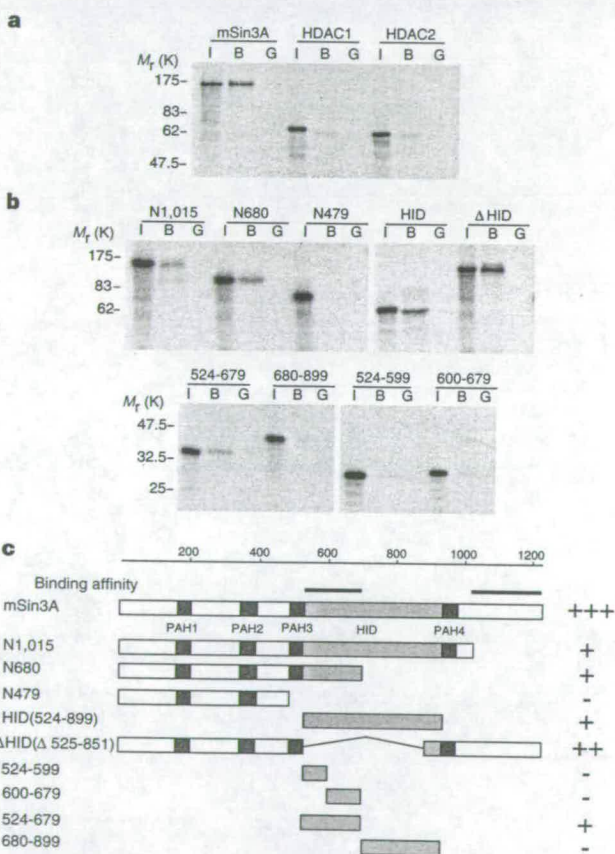


**Figure 2** Co-immunoprecipitation of mSin3, MeCP2 and histone deacetylase activity from rat brain nuclear extracts. **a**, Western blot of immunoprecipitates prepared using antibodies against MeCP2 (antibodies 670 and 674) or mSin3 antibodies AK11, AK12, K-20 and RE) or using preimmune serum. Probes were anti-mSin3A (RE) and anti-MeCP2 (670 and 674) antibodies. **b**, Immunoprecipitation of histone deacetylase activity from rat brain nuclear extracts with antibodies against MeCP2 (670 and 674), HDAC1, mSin3A or cyclin-dependent kinase 7 (CDK7) or with preimmune serum, in the presence or absence of 5 ng ml<sup>-1</sup> TSA. Precipitated activities from duplicate parallel experiments are shown as percentage total activity added to each tube (19,200 d.p.m. per 2 h in the experiment shown).

Control nuclear proteins (proliferating cell nuclear antigen and TATA-binding protein) were not bound by MeCP2 (data not shown). The region of MeCP2 that interacted with the corepressor substantially overlaps the TRD, as defined *in vivo* (Fig. 1b)<sup>9</sup>.

To test for a possible *in vivo* interaction between native MeCP2 and the mSin3 corepressor complex, we used antibodies against MeCP2 and mSin3 to immunoprecipitate proteins from rat brain nuclear extracts (Fig. 2a). Both anti-MeCP2 antibodies precipitated mSin3A as well as MeCP2. Four different anti-mSin3 antibodies immunoprecipitated MeCP2 as well as mSin3 (the anti-Sin3 antibody AK11 precipitated less MeCP2 than the other antibodies). Immunoprecipitates of transiently expressed HDAC2 also contained MeCP2 (data not shown). These results led us to expect that MeCP2 would be associated with a catalytically active deacetylase complex. This was confirmed by the finding that both anti-MeCP2 antibodies precipitated deacetylation activity, which was abolished by the highly specific deacetylase inhibitor trichostatin A (TSA)<sup>20</sup> (Fig. 2b). Antibodies against components of the corepressor complex, mSin3A and HDAC1, also immunoprecipitated deacetylase activity, as expected, whereas control antibodies (anti-CDK7 antibody and preimmune serum) gave background activity (Fig. 2b).

Does MeCP2 interact with mSin3 itself or with another component of the complex? To answer this question, we assayed the ability



**Figure 3** MeCP2 interacts with mSin3A, but not with HDACs. **a**, *In vitro*-translated <sup>35</sup>S-labelled mSin3A and HDACs 1 and 2 were incubated with an immobilized fusion protein containing GST and amino acids 108-392 of MeCP2 (Fig. 1b). Input protein (I) (20% of total), protein bound by the 108-392 fusion protein (B) and protein bound by GST alone (G) were analysed by SDS-PAGE. **b**, Deletion constructs of mSin3A were transcribed and translated *in vitro* to give <sup>35</sup>S-labelled products and processed as in **a**. Input (I) corresponds to 10% of total. **c**, Maps of the mSin3A deletion constructs<sup>13</sup>, indicating truncated proteins that did (+) or did not (-) bind to MeCP2. Regions of strong binding to MeCP2 are indicated by the thick black lines (top).



of the corepressor-interacting region of MeCP2 (residues 108–392) to bind *in vitro*-translated mSin3A, HDAC1 and HDAC2. mSin3A is the preferred binding partner, as the HDACs have a much weaker affinity for MeCP2 (Fig. 3a). There is probably a direct interaction between MeCP2 and mSin3A, although we cannot exclude the possibility that an unknown component of the translation lysate mediates indirect binding. We mapped the sites of interaction between MeCP2 and mSin3A by translating *in vitro* a series of mSin3A proteins with deleted portions (ref. 13) and testing their ability to bind to amino acids 108–392 of MeCP2 (Fig. 3b, c). The amino-terminal half of the HDAC-interaction domain (HID)<sup>13</sup> of mSin3A was common to most fragments that were able to bind MeCP2, but a deletion lacking all but the carboxy-terminal 48 amino acids of HID ( $\Delta$ HID) also bound efficiently, indicating a strong binding site within the C-terminal 200 amino acids of mSin3A (Fig. 3). We infer that the interaction between MeCP2 and mSin3A involves both the N-terminal half of the HID of mSin3A and the C-terminus of mSin3A. MeCP2 and deacetylases seem to bind to distinct sites within the 375-residue HID region, as both can bind simultaneously (Figs 1a and 2a, b).

To determine whether deacetylation is important for the effects of the TRD on transcription *in vivo*, we attempted to alleviate

transcriptional repression with TSA<sup>20</sup>. We transiently transfected mouse L929 fibroblasts with a reporter gene, which either did (G<sub>5</sub>) or did not (G<sub>0</sub>) have GAL4-binding sites near its promoter, in the presence or absence of a GAL4-TRD fusion construct<sup>9</sup>. In the absence of repressor, the ratio of G<sub>5</sub> expression to control (G<sub>0</sub>) expression was unaffected by TSA, remaining close to unity (Fig. 4). In the presence of the GAL4-TRD repressor, the G<sub>5</sub>:G<sub>0</sub> expression ratio was about 0.2, but was increased 2.8-fold by TSA. Incubations with TSA for 24 h (Fig. 4) or 9 h (data not shown) gave essentially the same result, eliminating the possibility that TSA exerted its effects by arresting the cells at a stage of the cell cycle where repression was minimal. Repression by GAL4-TRD was not completely alleviated by TSA in any of the experiments (Fig. 4, data not shown). The results indicate that a component of repression by the TRD may be deacetylase-independent, consistent with the observation that mSin3A retains some ability to repress transcription even in the absence of associated HDACs<sup>13</sup>.

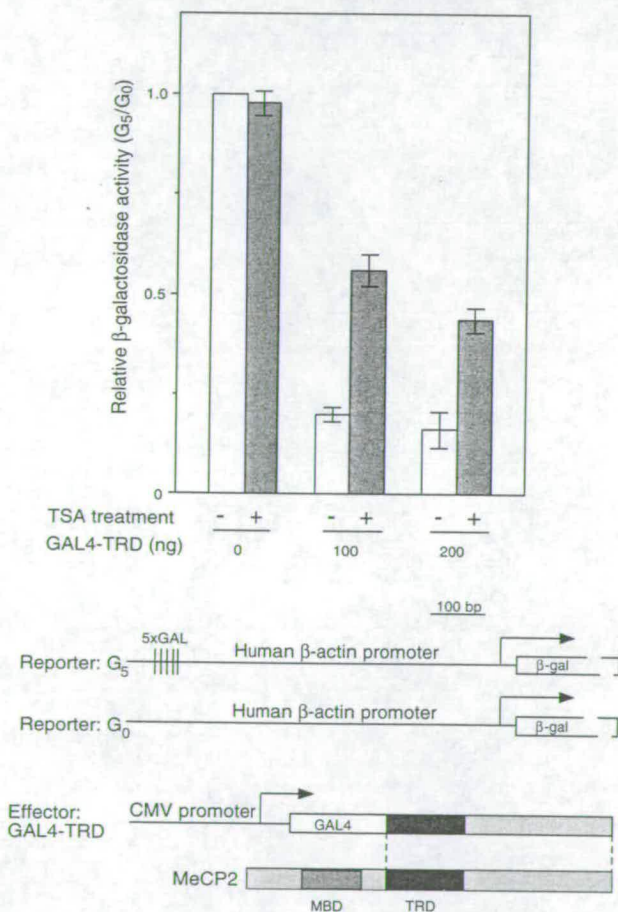
Our results indicate that repression by the TRD of MeCP2 relies to a significant extent, on histone deacetylation. Evidence for a functional relationship between DNA methylation and chromatin has come from studies showing that methylated DNA packaged as chromatin is incapable of transcription, whereas methylated DNA alone or chromatin alone remains transcriptionally active<sup>22,23</sup>. Given the strong correlation between deacetylation of histone protein tails and transcriptional repression<sup>24</sup>, the idea that methylated DNA provokes deacetylation is attractive<sup>25</sup>, and is supported by evidence that TSA can substitute for the demethylating drug 5-azacytidine to derepress methylated ribosomal RNA genes in plants<sup>26</sup>. Our results indicate that MeCP2 may provide a mechanistic bridge between DNA methylation and histone deacetylation. Recruitment of the corepressor complex by chromatin-bound MeCP2 may lead to local deacetylation of core histones (and perhaps of other proteins involved in transcription<sup>27</sup>), with consequent elimination of transcription.

#### Methods

**GST pulldown assay.** GST and GST fusion proteins<sup>11</sup> (6  $\mu$ g) were first immobilized on glutathione-Sepharose beads. The coated beads were then incubated with 60  $\mu$ g HeLa cell nuclear extract<sup>28</sup> in buffer A (20 mM HEPES pH 7.9, 150 mM NaCl, 0.5 mM EDTA, 1 mM dithiothreitol, 10% glycerol, 0.1% Triton X-100) for 2 h at 4°C. After five washes with buffer A, the proteins bound to the beads were eluted in Laemmli buffer and analysed by western blotting nearly as described<sup>13</sup>. The primary antibodies used were K20 (anti-mSin3A, Santa Cruz Biotech. Inc.), anti-HDAC1, and anti-HDAC2 (ref. 13). Nuclear extract (6  $\mu$ g protein) was loaded as a positive control. Polyclonal antiserum against mammalian HDAC1 was raised against a synthetic peptide corresponding to the C-terminal domain of HDAC1 (EEKPEAKGVKKEVKLA) as described<sup>29</sup>.

**Immunoprecipitation.** Antibodies (4  $\mu$ l serum, or 2  $\mu$ g purified IgG) were immobilized on protein A-Sepharose beads and washed with 20 mM HEPES pH 7.9, 0.1 M NaCl, 10% glycerol, 0.2 mM EDTA and 0.01% Triton X-100. The antisera were: anti-MeCP2 (670) raised against the amino acids 207–492 of MeCP2; anti-MeCP2 (674) raised against amino acids 1–392 of MeCP2; anti-mSin3A (RE)<sup>13</sup>; and anti-mSin3A/B antibodies AK-11, K-20 and AK12 (Santa Cruz Biotech. Inc.). Preimmune serum was from the rabbit that was subsequently immunized with MeCP2 amino acids 207–492. Bound antibodies were incubated with 20  $\mu$ g brain nuclear extract<sup>6,28</sup> in 240  $\mu$ l wash buffer plus 1 mg ml<sup>-1</sup> BSA for 1.5 h. Beads were washed three times and protein was eluted in Laemmli buffer. Immunoprecipitated proteins were detected by western blotting.

**Immunoprecipitation of histone deacetylase activity.** Aliquots (100  $\mu$ g protein) of rat brain nuclear extract were incubated for 16 h at 4°C, with 5  $\mu$ g affinity-purified antibodies against HDAC1 or mSin3A (Santa Cruz Biotech. Inc.) or with whole antiserum against MeCP2 (670 or 674) or CDK7 (a gift from J. Shuttleworth) or with preimmune serum. The final volume was 500  $\mu$ l in 20 mM HEPES pH 7.9, 150 mM NaCl, 0.1% Nonidet P40, 0.25% gelatin, 1 mM EDTA, 0.5 mM EGTA, 0.2% NaN<sub>3</sub>, 0.1 mM PMSF, and 1  $\mu$ g ml<sup>-1</sup> each of



**Figure 4** Relief of TRD-mediated repression of transcription by the deacetylase inhibitor TSA. Mouse fibroblasts (L cells) were transfected with reporter genes that did or did not contain GAL4 DNA-binding sites (G<sub>5</sub> or G<sub>0</sub>, respectively), in the presence or absence of a fusion between the GAL4 DNA-binding domain and amino acids 207–492 of MeCP2. After 24 h, transfected cells were treated with TSA (shaded bars) or left untreated (white bars) and were incubated for a further 24 h. Columns represent the mean ratio between  $\beta$ -galactosidase activities of G<sub>5</sub> and G<sub>0</sub> in three independent experiments ( $\pm$  standard deviation). Relative  $\beta$ -galactosidase activity in the absence of TSA and GAL4-TRD was normalized to 1.0. CMV, cytomegalovirus; bp, base pairs.



protinin, leupeptin and pepstatin A. Antibody-bound material was pelleted with protein A-Sepharose (Pharmacia), washed and assayed for histone deacetylase activity as described<sup>30</sup>. The substrate was a synthetic peptide corresponding to the 18 N-terminal residues of histone H4, chemically acetylated using <sup>3</sup>H-labelled acetic anhydride (Amersham) and added at 3 × 10<sup>5</sup> d.p.m. per assay.

**In vitro binding assay.** Intact mSin3A, HDAC1 (a gift from T. Kouzarides), HDAC2 and fragments of mSin3A (ref. 13) were labelled with [<sup>35</sup>S]methionine using a coupled transcription-translation TNT system (Promega). GST-MeCP2 fusion proteins or GST (3 μg) were first bound to glutathione-Sepharose beads. The slurry was then incubated with the labelled products (1 μl) in 100 μl buffer B (50 mM HEPES pH 7.9, 250 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 10% glycerol, 0.1% Triton X-100, 0.5 mg ml<sup>-1</sup> BSA) at 4 °C for 2 h. After washing the beads for three times with buffer B, the bound proteins were eluted in Laemmli loading buffer, fractionated by SDS-PAGE, and detected by autoradiography.

**Transfection and β-galactosidase assay.** Mouse fibroblasts (L cells) were transfected with reporters G<sub>0</sub> or G<sub>5</sub> in the presence or absence of GAL4-TRD(207-492). Transfection and β-galactosidase assays were performed nearly as described<sup>9</sup>. Fresh medium with or without 100 ng ml<sup>-1</sup> TSA was added to the mouse L929 fibroblast cells 24 h after transfection and the cells were collected after a further 24 h. In some experiments, treatment with TSA was reduced to 9 h.

Received 9 February; accepted 19 March 1998.

1. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362-365 (1993).
2. Beard, C., Li, E. & Jaenisch, R. Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev.* **9**, 2325-2334 (1995).
3. Jahner, D. *et al.* De novo methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* **298**, 623-628 (1982).
4. Groudine, M., Eisenman, R. & Weintraub, H. Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. *Nature* **292**, 311-317 (1981).
5. Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L. & Bird, A. P. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**, 499-507 (1989).
6. Lewis, J. D. *et al.* Purification, sequence and cellular localisation of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905-914 (1992).
7. Boyes, J. & Bird, A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**, 1123-1134 (1991).
8. Cross, S. H., Meehan, R. R., Nan, X. & Bird, A. A component of the transcriptional repressor MeCP2 is related to mammalian DNA methyltransferase and throrax-like protein. *Nature Genet.* **16**, 256-259 (1997).
9. Nan, X., Campoy, J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471-481 (1997).
10. Tate, P., Skarnes, W. & Bird, A. The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nature Genet.* **12**, 205-208 (1996).
11. Nan, X., Meehan, R. R. & Bird, A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**, 4886-4892 (1993).
12. Nan, X., Tate, P., Li, E. & Bird, A. P. DNA methylation specifies chromosomal localization of MeCP2. *Mol. Cell Biol.* **16**, 414-421 (1996).
13. Laherty, C. D. *et al.* Histone deacetylases associated with the mSin3 corepressor mediate transcriptional repression. *Cell* **89**, 849-856 (1997).
14. Alland, L. *et al.* Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. *Nature* **387**, 49-55 (1997).
15. Hassig, C. A., Fleischner, T. C., Billin, A. N., Schreiber, S. L. & Ayer, D. E. Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell* **89**, 341-347 (1997).
16. Nagy, L. *et al.* Nuclear receptor repression mediated by a complex containing SMRT, mSINA, and histone deacetylase. *Cell* **89**, 373-380 (1997).
17. Heinzel, T. *et al.* A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. *Nature* **387**, 43-48 (1997).
18. Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell* **89**, 357-364 (1997).
19. Kadosh, D. & Struhl, K. Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* **89**, 365-371 (1997).
20. Yoshida, M., Horinouchi, S. & Beppu, T. Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays* **17**, 423-430 (1995).
21. Lewis, J. & Bird, A. DNA methylation and chromatin structure. *FEBS Lett.* **285**, 155-159 (1991).
22. Kass, S. U., Landsberger, N. & Wolffe, A. P. DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.* **7**, 157-165 (1997).
23. Buschhausen, G., Witting, B., Graessmann, M. & Graessmann, A. Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc. Natl Acad. Sci. USA* **84**, 1177-1181 (1987).
24. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349-352 (1997).
25. Kass, S. U., Pruss, D. & Wolffe, A. P. How does DNA methylation repress transcription? *Trends Genet.* **13**, 444-449 (1997).
26. Chen, Z. J. & Pikaard, C. S. Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation histone modification in nucleolar dominance. *Genes Dev.* **11**, 2124-2136 (1997).
27. Pazin, M. J. & Kadonaga, J. T. What's up and down with histone deacetylation and transcription? *Cell* **89**, 325-328 (1997).
28. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **5**, 1475-1489 (1983).
29. Turner, B. M. & Fellows, G. Specific antibodies reveal ordered and cell cycle-related use of histone H4 acetylation sites in mammalian cells. *Eur. J. Biochem.* **179**, 131-139 (1989).

30. Li, W., Chen, H. Y. & Davie, J. R. Properties of chicken erythrocyte histone deacetylase associated with the nuclear matrix. *Biochem. J.* **314**, 631-637 (1996).

**Acknowledgements.** We thank J. Davidson and A. Greig for technical assistance, E. Seto for anti-HDAC2 antiserum and T. Kouzarides for the HDAC1 expression plasmid. The work was supported by a grant from the N.I.H. (to R.N.E.) and by grants from The Wellcome Trust (to A.B. and B.M.T.). H.-H.N. is a Darwin Trust scholar; C.D.L. is a Leukemia Society of America Special Fellow; and R.N.E. is an American Cancer Society Research Professor.

Correspondence and requests for materials should be addressed to A.B. (e-mail: A.Bird@ed.ac.uk).



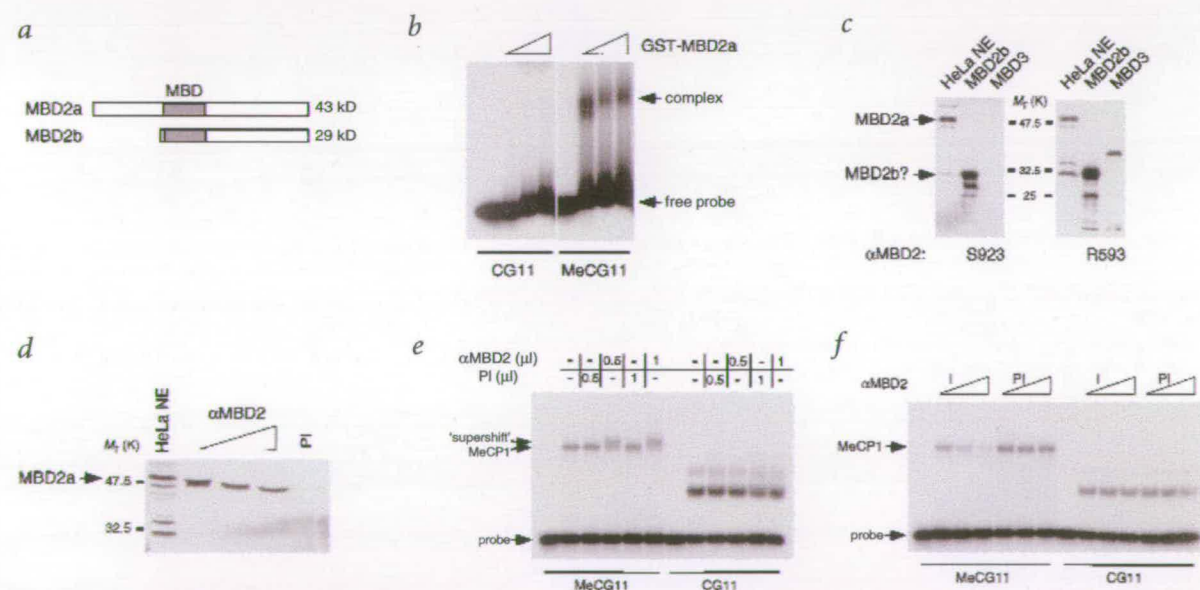
# MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex

Huck-Hui Ng<sup>1</sup>, Yi Zhang<sup>2</sup>, Brian Hendrich<sup>1</sup>, Colin A. Johnson<sup>3</sup>, Bryan M. Turner<sup>3</sup>, Hediye Erdjument-Bromage<sup>4</sup>, Paul Tempst<sup>4</sup>, Danny Reinberg<sup>2</sup> & Adrian Bird<sup>1</sup>

Mammalian DNA is methylated at many CpG dinucleotides. The biological consequences of methylation are mediated by a family of methyl-CpG binding proteins<sup>1-4</sup>. The best characterized family member is MeCP2, a transcriptional repressor that recruits histone deacetylases<sup>5-7</sup>. Our report concerns MBD2, which can bind methylated DNA *in vivo* and *in vitro*<sup>4</sup> and has been reported to actively demethylate DNA (ref. 8). As DNA methylation causes gene silencing, the MBD2 demethylase is a candidate transcriptional activator. Using specific antibodies, however, we find here that MBD2 in HeLa cells is associated with histone deacetylase (HDAC) in the MeCP1 repressor complex<sup>1,9</sup>. An affinity-purified HDAC1 corepressor complex<sup>10,11</sup> also contains MBD2, suggesting that MeCP1 corresponds to a fraction of this complex. Exogenous MBD2 represses transcription in a transient assay, and repression can be relieved by the deacetylase inhibitor trichostatin A (TSA; ref. 12). In our hands, MBD2 does not demethylate DNA. Our

data suggest that HeLa cells, which lack the known methylation-dependent repressor MeCP2, use an alternative pathway involving MBD2 to silence methylated genes.

Previous work showed that Mbd2b, a truncated form of mouse Mbd2 (Fig. 1a), binds DNA *in vitro* in a methylation-dependent manner<sup>4</sup>. Full-length recombinant mouse MBD2a (Fig. 1a) also formed a complex with a densely methylated probe in a bandshift assay, but did not bind to non-methylated counterpart (Fig. 1b). To search for MBD2-associated proteins, we first raised polyclonal antisera against recombinant MBD2. Antibodies were tested on western blots of HeLa nuclear extract and assessed for their ability to discriminate between MBD2 (in particular the MBD3-like MBD2b form) and its relative, MBD3 (ref. 4). Antiserum S923 bound to recombinant MBD2b, but not to MBD3 (Fig. 1c). Antiserum R593 recognized MBD2b and cross-reacted with MBD3 (Fig. 1c). In HeLa extract, S923 antibodies detected an approxi-

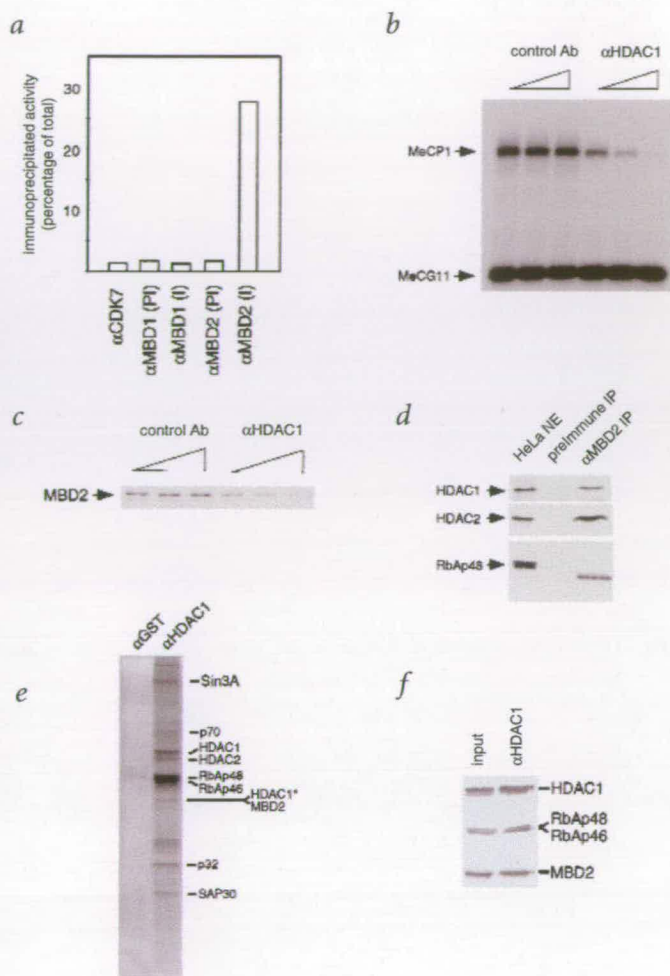


**Fig. 1** MBD2a binds methylated DNA specifically and is part of the methyl-CpG binding complex, MeCP1. **a**, Maps of full-length MBD2a and a hypothetical truncated form, MBD2b, that initiates at a downstream methionine<sup>4</sup>. **b**, Recombinant MBD2a forms a complex with probe MeCG11 (26 methyl-CpGs), but not with the non-methylated form CG11. **c**, Anti-MBD2 antisera S923 and R593 probed against western blots of HeLa nuclear extract (26  $\mu$ g protein; HeLa NE), recombinant MBD2b (30 ng; ref. 4) and recombinant MBD3 (30 ng; ref. 4). **d**, Immunoprecipitates with anti-MBD2 antiserum S923 ( $\alpha$ MBD2) contained only the 48-kD MBD2a protein. MBD2a always resolves as a doublet band for unknown reasons; the doublet resolves less well at higher concentrations of antibody due to the increased presence of closely migrating IgG. The control was pre-immune serum (PI). **e**, S923 anti-MBD2 antibodies supershift the bandshift complex between MeCP1 and the methylated DNA probe MeCG11. No supershift was seen with pre-immune serum or of the non-specific CG11 complex (PI). **f**, Depletion of MBD2 from HeLa cell nuclear extracts with S923 antiserum (2.5, 5 and 10  $\mu$ l) causes loss of MeCP1 complex in a bandshift assay. Pre-immune serum (PI) has no effect.

<sup>1</sup>Institute of Cell and Molecular Biology, University of Edinburgh, The King's Buildings, Edinburgh EH9 3JR, UK. <sup>2</sup>Howard Hughes Medical Institute, Division of Nucleic Acids Enzymology, Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, USA. <sup>3</sup>Department of Anatomy, The Medical School, University of Birmingham, Birmingham B15 2TT, UK. <sup>4</sup>Molecular Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA. Correspondence should be addressed to A.B. (e-mail: A.Bird@ed.ac.uk).



**Fig. 2** MBD2 is associated with histone deacetylases in the MeCP1 complex. **a**, Anti-MBD2 antibodies co-immunoprecipitate HDAC activity from HeLa cell extracts, but pre-immune serum (PI) and control antibodies against MBD1 and CDK7 do not. **b,c**, Anti-HDAC1 antibodies (10, 20 and 30  $\mu$ l) deplete MeCP1 bandshift activity (**b**) and MBD2 (**c**) from a HeLa nuclear extract. Control non-immune serum does not deplete MeCP1 or MBD2. MBD2 was detected by western blot using S923 antiserum. **d**, Anti-MBD2 antibodies co-immunoprecipitate HDAC1, HDAC2 and RbAp48 from HeLa nuclear extracts. Co-migration with immunoglobulin heavy chain leads to aberrant migration of RbAp48 in the immunoprecipitated lane. **e**, Polypeptides retained on HDAC1 ( $\alpha$ HDAC1) and GST ( $\alpha$ GST) antibody columns after separation by SDS-PAGE followed by silver staining. Known components of the complex<sup>11</sup> are labelled (right). Mass spectrometric and sequence analysis of the 43-kD band from the fraction isolated using anti-HDAC1 antibodies identified a peptide (VLYIDIIHHGDGVVEAFYTTDR) contained in HDAC1, as well as a peptide (GLQGVGPGSNDLTLSSAVASALHTSS-APITGQVSAAVEK) that matches the MBD2 sequence. It is not known whether HDAC1\* represents an alternative form or degradation product of HDAC1. The data suggest roughly equimolar amounts of HDAC1\* and MBD2 in this band (data not shown). **f**, MBD2a, RbAp46, RbAp48 and HDAC1 were detected by western blot in nuclear extract (input) and in complexes that were affinity purified with anti-HDAC1 antibodies.



mately 49-kD band corresponding to full-length MBD2a (*in vitro*-translated MBD2a has an identical mobility; data not shown), plus an approximately 30-kD band that may be MBD2b or a degradation product. S923 does not cross-react with MeCP2, MBD1 or MBD4 (data not shown). We further tested the specificity of the S923 antiserum against soluble nuclear proteins by probing western blots of S923 immunoprecipitates with the less-specific R593 antiserum. R593 recognizes MBD3 (Fig. 1c), but only MBD2a was detected in the precipitated material (Fig. 1d).

We used the highly specific S923 antiserum to determine whether MBD2 is a component of the methyl-CpG binding activity (MeCP1; ref. 1) previously implicated as a transcriptional repressor<sup>9,13</sup>. S923 antibodies were able to quantitatively supershift the MeCP1 complex with a probe methylated at multiple sites, whereas pre-immune serum was inert (Fig. 1e, MeCG11 lanes). A non-specific complex with the non-methylated probe<sup>9</sup> was unaffected by either immune or pre-immune sera (Fig. 1e, CG11 lanes). When we used anti-MBD2 antibodies to deplete MBD2 from the HeLa extract, the MeCP1 complex was correspondingly depleted (Fig. 1f). Pre-immune serum did not deplete the MeCP1 complex, and neither antiserum affected the non-specific complex with non-methylated DNA. We conclude that MBD2 is a component of the MeCP1 complex.

Several studies have indicated that methylation-dependent transcriptional repression involves deacetylation<sup>6,7,14-17</sup>. As MeCP1 is a likely repressor of methylated genes<sup>9</sup>, we asked whether its constituent MBD2 is associated with deacetylases. We found that S923 immunoprecipitates contained up to 27% of the total histone deacetylase activity of HeLa extracts (Fig. 2a). Moreover, anti-HDAC1 antibodies efficiently immunodepleted both MeCP1 (Fig. 2b) and MBD2 (Fig. 2c) from the extract, whereas control non-immune serum had no effect. In addition, several known protein components of corepressor complexes were co-immunoprecipitated by MBD2 antibodies from a HeLa extract (Fig. 2d), indicating that MeCP1 activity is due to a complex between MBD2, deacetylases HDAC1 and HDAC2, and RbAp48. As a negative control, we found that retinoblastoma protein, which associates with HDACs (refs 18-20), was not immunoprecipitated by anti-MBD2 (data not shown). Mass fingerprinting analysis of proteins affinity-purified using an anti-

HDAC1 antibody<sup>11</sup> (Fig. 2e) indicated that MBD2 is a component of a corepressor complex. Western-blot analysis confirmed that MBD2 is present in complexes affinity purified with HDAC1 antibodies, together with HDAC1/2-associated proteins RbAp46 and RbAp48 (Fig. 2f).

A previous report claimed that a related methyl-CpG binding protein, MBD1, is a component of MeCP1 (ref. 3). Subsequent experiments with antibodies highly specific for MBD1 failed to supershift or immunodeplete MeCP1 (H.-H.N. and A.B., unpublished data) or to immunoprecipitate deacetylase activity (Fig. 2a). These and other results lead us to suspect that the anti-MBD1 antibody used in the previous study may have cross-reacted with MBD proteins that were unknown at the time. Recent work confirms the proposal<sup>3</sup> that MBD1 is a methylation-dependent transcriptional repressor, but argues against its involvement in MeCP1 of HeLa cells (H.-H.N. and A.B., unpublished data).

HeLa cells lack detectable amounts of the known methylation-dependent repressor MeCP2 (Fig. 3a), but are nevertheless able to repress transfected methylated genes<sup>9</sup>. We initially determined whether repression in the effective absence of MeCP2 was HDAC dependent by transiently transfecting a luciferase reporter gene that contained a cluster of 20 methylatable *HhaI* sites (CG11) immediately upstream of the SV40 promoter (Fig. 3b,c). The M.HhaI-methylated reporter construct was transcribed at approximately 30% of the level seen in non-methylated controls (Fig. 3b). The same construct lacking the CG11 cluster (originally used to







tion of binding sites during the prolonged assay (Figs 1b and 4a). Attempts to demethylate other oligonucleotide and plasmid substrates with bacterially expressed MBD2, *in vitro*-translated MBD2b or anti-MBD2 immunoprecipitates from HeLa extracts were similarly unsuccessful (data not shown). It may be that MBD2 has a dual role as both a DNA methylation-dependent repressor and an activator of genes that are silenced by methylation. Diametrically opposed activities might theoretically be combined in a protein that serves as a switch between active and inactive chromatin states. This speculative hypothesis depends, however, on verification of the demethylase activity of MBD2.

Our data uncover a novel mediator of methylated gene silencing. Although MeCP2 and MBD2 both repress transcription of methylated DNA via a deacetylase-containing corepressor complex, there are differences between them. Tightly bound MeCP2 (refs 2,21) has the character of a structural component of the chromosome that ensures long-term silencing of methylated sequences. MBD2/MeCP1, on the other hand, is released from nuclei by low salt<sup>1</sup>, suggesting that it is not stably complexed with DNA. The MeCP1 subset of the HDAC corepressor complex may be a catalytic complex that alters methylated DNA-containing chromatin during a transient visit. There is also a difference in binding specificity between MeCP1 and MeCP2. MeCP2 binds to single methylated sites<sup>22</sup>, whereas MeCP1 requires regions of dense CpG methylation<sup>1</sup>. Conceivably, each MeCP1 complex contains multiple molecules of MBD2, thereby enhancing binding to local clusters of methyl-CpG. Assessment of this possibility awaits future purification and compositional analysis of MeCP1. Already, the differences between MBD2 and MeCP2 suggest that they may have somewhat different roles in transcriptional silencing.

## Methods

**Plasmids.** We constructed plasmid pCMV-Gal4-MBD2a by inserting a *NaeI/SspI* fragment of *MBD2a* cDNA (ref. 4) into *Bam*HI-blunted pCMV-Gal4 vector<sup>5</sup> and pGEX5-MBD2a by inserting the same *NaeI/SspI* fragment of *MBD2a* cDNA into *Bam*HI-blunted pGEX5x1 vector (gift from S. MacNeill). We sequenced all constructs to confirm identity. CG11-pGL2 luciferase reporter was a gift from X. Nan (unpublished data). Plasmid pHis-dMTase was a gift from M. Szyf<sup>8</sup>. The Gal4-DNA polymerase- $\beta$  luciferase reporter<sup>23</sup> was a gift from J. Millbrandt.

**Transfection and reporter assays.** We transfected mouse fibroblast L cells and  $\beta$ -galactosidase assays essentially as described<sup>5,6</sup> and detected luciferase activity using a luciferase assay kit (Promega). Transfection mixtures included reporter (2  $\mu$ g; with Gal4-binding sites), effector (2  $\mu$ g; Gal4-MBD2a) and internal control reporter (2  $\mu$ g; either human *ACTB* promot-

er driving  $\beta$ -galactosidase or pGL2 SV40 promoter/enhancer driving luciferase). We transfected HeLa cells with CG11-pGL2 luciferase reporter (1  $\mu$ g; either mock methylated or methylated with *HhaI* methyltransferase) and human  $\beta$ -actin/ $\beta$ -galactosidase reporter (1  $\mu$ g) using Lipofectamine (Gibco BRL) according to the manufacturer's instructions. We treated cells with TSA (100 ng/ml) for 24 h before collection.

**Bandshifts, immunodepletion and immunoprecipitation assays.** We performed bandshift assays as described<sup>3</sup>. For immunodepletion of MBD2 or MeCP1 activity, we immobilized the antibodies (5–30  $\mu$ l serum) on protein-G sepharose (Pharmacia). After washing the beads with PBS (plus 1 M NaCl and 0.1% Triton X-100) and binding buffer (50 mM Hepes, pH 7.9, 150 mM NaCl, 0.5 mM EDTA, 10% glycerol, 0.1% Triton X-100), we incubated the beads with HeLa nuclear extract (40  $\mu$ g) in binding buffer for 2 h at 4 °C. Aliquots of the supernatant were then used for bandshift and western-blot analysis. We performed immunoprecipitations as described<sup>6</sup>. For immunoprecipitation of histone deacetylase activity<sup>6</sup>, we used HeLa whole cell extract, lysed in Tris (50 mM, pH 8.0), NaCl (150 mM), EDTA (0.5 mM), EGTA (0.5 mM), 1% NP40,  $\beta$ -mercaptoethanol (5 mM) and protease inhibitors, as input.

**Antibodies, western blots and affinity purification.** Antibody S923 was raised in sheep against GST-MBD2a fusion protein. In this preparation of affinity-purified recombinant protein, only 20% was full length, so the sample was biased toward the tagged amino-terminal portion of MBD2a. Antibody R593 was raised in rabbit against GST-MBD2b fusion protein. Anti-mSin3A (AK11), anti-HDAC2 (H-52) and anti-Rb (C-15) were purchased (Santa Cruz). Anti-RbAp48 was raised against the synthetic peptide CENYNDPEGSVDPEGQGS as described<sup>24</sup>. Antibodies against MeCP2 (ref. 5) have been described. We performed western blots as described<sup>3</sup>, except the buffer used was Tris HCl (50 mM, pH 8.0), NaCl (150 mM), 0.05% Tween-20 and 1% dried milk powder. Complexes were affinity purified using anti-HDAC1 antibodies as described<sup>11</sup>.

**Demethylase assays.** We carried out *in vitro* transcription/translation of MBD2 expression plasmids and demethylase assays as described<sup>8</sup>. We analysed restriction digests of MeCG11 probe by electrophoresis on 8% polyacrylamide gels.

## Acknowledgements

We thank J. Millbrandt, M. Szyf, S. MacNeill and X. Nan for plasmids and J. Davidson and A. Greig for technical assistance. This work was supported by grants from the Wellcome Trust to A.B. and B.M.T.H.-H.N. holds a Darwin Trust Scholarship. D.R. is supported by the Howard Hughes Medical Institute and the National Institutes of Health. Y.Z. holds an NIH Post-doctoral Fellowship.

Received 3 May; accepted 1 July 1999.

- Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L. & Bird, A.P. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**, 499–507 (1989).
- Lewis, J.D. *et al.* Purification, sequence and cellular localisation of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905–914 (1992).
- Cross, S.H., Meehan, R.R., Nan, X. & Bird, A. A component of the transcriptional repressor MeCP1 is related to mammalian DNA methyltransferase and thirathorax-like protein. *Nature Genet.* **16**, 256–259 (1997).
- Hendrich, B. & Bird, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18**, 6538–6547 (1998).
- Nan, X., Campoy, J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471–481 (1997).
- Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
- Jones, P.L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet.* **19**, 187–191 (1998).
- Bhattacharya, S.K., Ramchandani, S., Cervoni, N. & Szyf, M. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397**, 579–583 (1999).
- Boyes, J. & Bird, A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**, 1123–1134 (1991).
- Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell* **89**, 357–364 (1997).
- Zhang, Y. *et al.* SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell* **1**, 1021–1031 (1998).
- Yoshida, M., Horinouchi, S. & Beppu, T. Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and

- function. *Bioessays* **17**, 423–430 (1995).
- Boyes, J. & Bird, A. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J.* **11**, 327–333 (1992).
- Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G. & Baylin, S.B. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nature Genet.* **21**, 103–107 (1999).
- Eden, S., Hashimshony, T., Keshet, I. & Cedar, H. DNA methylation models histone acetylation. *Nature* **394**, 842 (1998).
- Ng, H.-H. & Bird, A. DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev.* **9**, 158–163 (1999).
- Selker, E.U. Trichostatin A causes selective loss of DNA methylation in *Neurospora*. *Proc. Natl. Acad. Sci. USA* **95**, 9430–9435 (1998).
- Brehm, A. *et al.* Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597–600 (1998).
- Magnaghi-Jaulin, L. *et al.* Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391**, 601–604 (1998).
- Luo, R.X., Postigo, A.A. & Dean, D.C. Rb interacts with histone deacetylase to repress transcription. *Cell* **92**, 463–473 (1998).
- Nan, X., Tate, P., Li, E. & Bird, A.P. DNA Methylation specifies chromosomal localization of MeCP2. *Mol. Cell. Biol.* **16**, 414–421 (1996).
- Nan, X., Meehan, R.R. & Bird, A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**, 4886–4892 (1993).
- Yeung, K.C., Inostroza, J.A., Mermelstein, F.H., Kannabiran, C. & Reinberg, D. Structure-function analysis of the TBP-binding protein Dr1 reveals a mechanism for repression of class II gene transcription. *Genes Dev.* **8**, 2097–2109 (1994).
- White, D.A., Belyaev, N.D. & Turner, B.M. Preparation of site-specific antibodies to acetylated histones. *Methods* (in press).



## DNA methylation and chromatin modification

Huck-Hui Ng\* and Adrian Bird†

DNA methylation and chromatin modification are two global mechanisms that regulate gene expression. Recent studies provide insight into the mechanism of transcriptional silencing by a methyl-CpG binding protein, MeCP2. MeCP2 is shown to interact with the Sin3/histone deacetylase co-repressor complex. Thus, this interaction can provide a mechanistic explanation for the long-known relationship between DNA methylation and chromatin structure. Moreover, several studies have shown that inhibition of histone deacetylases by specific inhibitors can reactivate endogenous genes or reporter constructs previously silenced by DNA methylation. Taken together, the data strongly suggest that DNA methylation can pattern chromatin modification.

### Addresses

Institute of Cell and Molecular Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3JR, UK  
\*e-mail: Huck.Hui.Ng@ed.ac.uk  
†e-mail: A.Bird@ed.ac.uk

Current Opinion in Genetics & Development 1999, 9:158–163

<http://biomednet.com/elecref/0959437X00900158>

© Elsevier Science Ltd ISSN 0959-437X

### Abbreviations

5-aza-dC	5-aza-2'-deoxycytidine
5-aza-C	5-azacytidine
ER	estrogen receptor
HAT	histone acetyltransferase
HDAC	histone deacetylase
MeCP	methyl-CpG binding protein
NF	nuclear factor
PR	progesterone receptor
SRC-1	steroid receptor coactivator-1
TF	transcription factor
TK	thymidine kinase
TRD	transcriptional repression domain
TSA	Trichostatin A

### Introduction

Methylation of cytosines at the carbon 5 position of CpG dinucleotides is a characteristic feature of many eukaryotic genomes. In vertebrates, 60–90% of all CpGs are methylated, leaving a minor part of the genome methylation free. Many of the remaining non-methylated CpGs (~15% of all CpGs in human DNA) are found in CpG islands [1], which usually include functional promoters.

A large body of evidence accumulated over the years suggests that DNA methylation is incompatible with transcriptional activity. Initially, many studies documented a strong correlation between DNA methylation and gene inactivity [2]. The loss of expression from artificially methylated gene constructs upon transfection into tissue culture cells strongly supported this hypothesis [3]. Furthermore, the application of a DNA-demethylating agent, 5-azacytidine (5-aza-C), was shown to bring about

the reactivation of previously methylated, silenced endogenous genes [4] and retroviruses [5].

### DNA methylation and chromatin

Early experiments demonstrated that artificially methylated DNA can adopt a distinctive chromatin structure upon integration into the genome [6]. This conformation renders the DNA refractory to nuclease or restriction endonuclease digestion and also leads to the loss of DNase-I-hypersensitive sites. On the other hand, unmethylated CpG islands, which are associated with promoters of many genes, possess a nuclease-sensitive chromatin structure that differs from the bulk of the methylated genome [7]. CpG island chromatin was found to contain highly acetylated histone H3 and H4.

Is the inactive chromatin structure associated with DNA methylation important for the loss of transcriptional activity? This question was critically addressed by experiments analyzing the kinetics of transcriptional arrest from methylated templates [8]. Both methylated and non-methylated versions of the herpes simplex virus thymidine kinase (HSV TK) gene were equally active during the first 8 hours after microinjection into rodent cells. It was only after longer periods (>8 hours) that the inhibitory effect of DNA methylation was felt on the methylated reporter construct, as measured by the accumulation of HSV TK RNA. It was hypothesized that the injected DNA was not fully assembled into chromatin at early times, and that the DNA methylation status had no effect on transcriptional activity in the absence of chromatin. Indeed, when an *in vitro* reconstituted chromatinized methylated reporter was introduced into the cells, it became transcriptionally inert. Subsequently Kass *et al.* [9] showed clear evidence for a time-dependent repression of methylated template following injection into *Xenopus* oocytes. Loss of transcriptional activity from the methylated template coincided with the appearance of a nucleosomal array in the vicinity of the promoter and the disappearance of engaged RNA polymerases. The above studies demonstrate that DNA methylation can lead to alteration of chromatin structure.

Until recently, the molecular mechanism by which DNA methylation represses transcription remained largely unknown but a few reports hinted at a connection with histone deacetylation. In one instance, an inhibitor of histone deacetylase (HDAC), sodium butyrate, was shown to elevate the expression of a methylated episomal reporter gene [10]. It was also reported that both sodium butyrate and Trichostatin A (TSA; a specific inhibitor of HDACs [11]) can substitute for the DNA-demethylating agent, 5-aza-2'-deoxycytidine (5-aza-dC) in restoring transcription from previously methylated and silenced plant ribosomal RNA genes [12\*]. The interchangeable effect of



TSA and a DNA-demethylating agent prompted the notion that DNA methylation and histone deacetylation may operate along a common mechanistic pathway to silence transcription.

### Histone hypoacetylation and transcriptional repression

It is evident from numerous studies (for review, see [13]) that transcriptional repressors often function in the context of chromatin by recruiting a common chromatin-modifying complex. For example, Mad-Max heterodimers and unliganded nuclear hormone receptors were found to associate with the mSin3 co-repressor complex [14–19]. This multi-protein complex contains catalytically active HDACs (HDAC1 and HDAC2). An underlying theme that emerged is the targeting of HDACs to the chromatin by sequence-specific DNA-binding proteins [20]. Deacetylation of histone H3 and H4 by the HDACs presumably leads to the formation of a chromatin environment that inhibits transcription. Indeed, HDAC inhibitors TSA or trapoxin were shown to relieve the repressive activity associated with the mSin3 co-repressor complex.

### DNA methylation and histone deacetylation

MeCP2 is a chromosomal protein that binds specifically to methylated DNA both *in vitro* and *in vivo*, via its methyl-CpG binding domain [21–23]. In addition, MeCP2 can 'invade' chromatin in a methylation-dependent manner and, *in vitro* at least, can displace histone H1 [24]. A highly basic central region of MeCP2 was shown to repress transcription *in vivo*. In a recent study [25\*\*], a region of MeCP2 encompassing the transcriptional repression domain (TRD) has been shown to bind components of the mSin3A/HDAC complex. The interaction between MeCP2 and the co-repressor complex was further substantiated by reciprocal co-immunoprecipitation experiments using native rat brain nuclear extracts. As mSin3A is the preferred binding partner of MeCP2 *in vitro*, it is proposed that MeCP2 recruits the co-repressor complex via interaction with mSin3A *in vivo*. TSA was able to relieve the TRD-mediated transcriptional repression, indicating that histone deacetylation is indeed a critical component of the repression mechanism.

Using biochemical fractionation methodology, Jones *et al.* [26\*\*] have found that *Xenopus* MeCP2 co-purifies with both the Sin3A protein and the HDAC activity through several chromatography steps [26\*\*], strongly suggesting that MeCP2 can also form a stable complex with the Sin3A/HDAC proteins in the *Xenopus* system. The association between MeCP2 and the Sin3A co-repressor complex was also confirmed by co-immunoprecipitation experiments [26\*\*]. As in mammals, repression by the TRD of *Xenopus* MeCP2 can be alleviated by TSA. More importantly, the authors also reported that TSA can relieve the inhibition imposed on methylated reporters injected into the oocytes, as the drug was able to restore transcription and hypersensitive sites of the methylated *hsp70*

promoter. These results show that deacetylation causes silencing of the methylated reporters and they suggest that MeCP2 may be the mediator responsible.

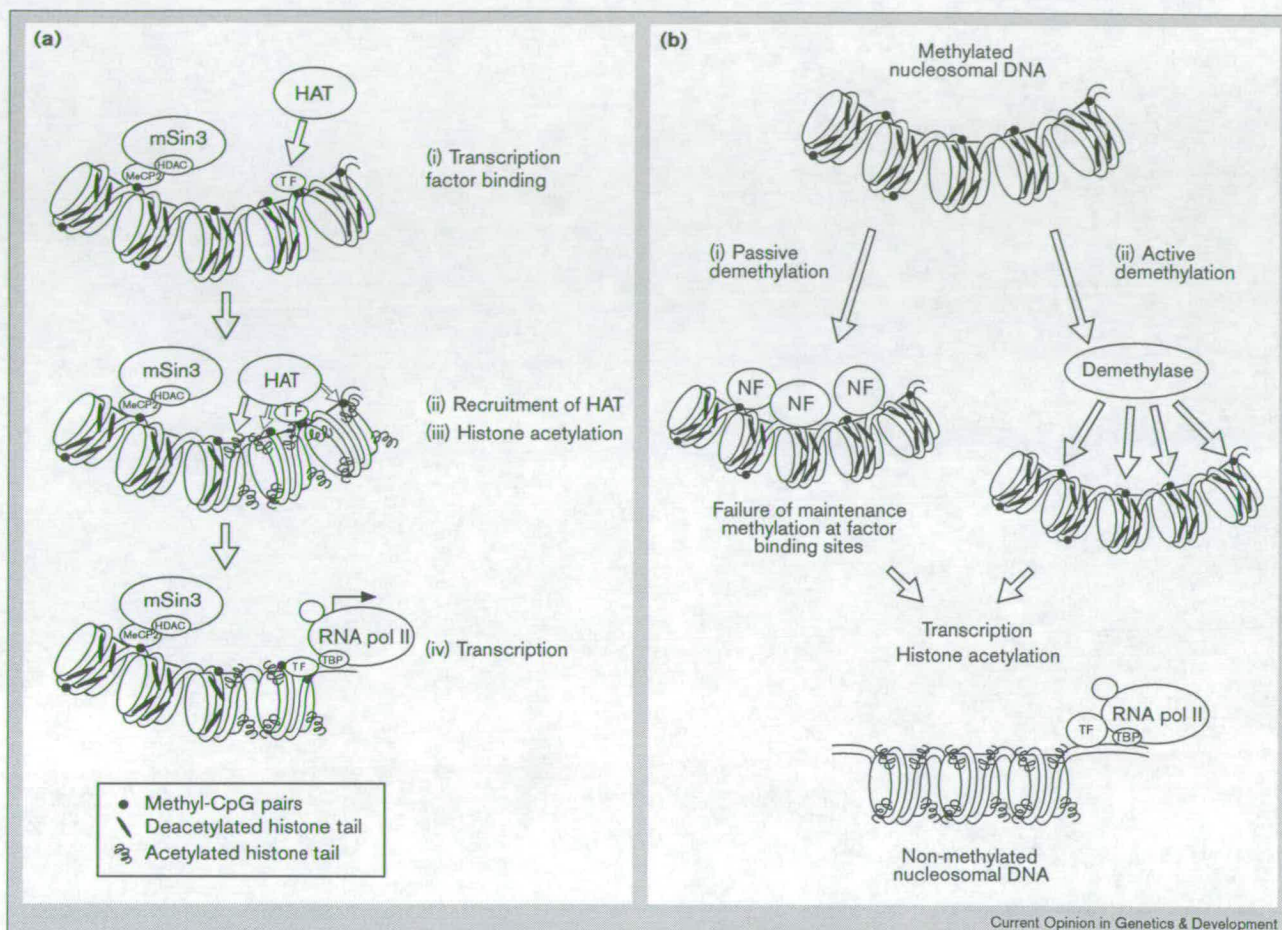
A causal link between DNA methylation and histone deacetylation is further strengthened by studies that show parallel effects of DNA-demethylating agents and TSA in reactivating transcription of sequences repressed by DNA methylation. In one study, mouse fibroblast L cells were stably transfected with either a methylated or a non-methylated *tk* gene [27\*\*]. Subsequent immunoprecipitation of chromatin with an antibody against acetylated histone H4 revealed that, in contrast to the non-methylated *tk* gene, the methylated *tk* gene was not enriched for acetylated histones. TSA treatment resulted in an increased expression of the methylated *tk* gene. In addition, TSA treatment increased the accessibility of the methylated *tk* gene locus to nucleases, suggesting that acetylation of histones can re-establish features of transcriptionally active chromatin despite the presence of DNA methylation. Similar results were obtained with integrated transgenes that were subjected to spontaneous epigenetic silencing with the concomitant appearance of DNA methylation and loss of histone hyperacetylation [28\*]. Again, both butyrate/TSA and 5-aza-C were able to restore transgene expression.

Are endogenous genes that are repressed by DNA methylation also affected by treatment with a histone deacetylase inhibitor? This question is addressed by two recent reports. The myeloid-specific mouse M-lysozyme gene is methylated and silenced in non-myeloid cells. When non-expressing cells — such as mouse fibroblasts, lymphocytic T-cells and myeloid precursor cells — are treated with TSA, the expression of M-lysozyme is induced even though the gene remains methylated [29\*]. The level of induction of M-lysozyme gene expression is comparable with that achieved by the demethylating agent 5-aza-dC. The effect of TSA on methylated endogenous genes has also been tested in the fungus *Neurospora crassa* [30\*\*]. Inhibition of histone deacetylation results in the re-expression of a methylated copy of the hygromycin resistance gene. The degree of derepression by TSA was found to be comparable to that caused by DNA demethylation using 5-aza-C. A similar result was also observed at the *Tad1am* locus.

Although the above-mentioned studies provide strong evidence that DNA methylation may influence the underlying histone modification pattern, it is important to note that not all data support this idea. For example, although TSA treatment can reactivate some methylated genes in *Neurospora crassa*, the expression of other methylated regions remains unaffected [30\*\*]. Similarly, 5-aza-C can reactivate dormant, methylated proviral genomes in chicken cells [5] but TSA in this system fails to restore transcription (RN Eisenman, personal communication). In general, caution needs to be exercised when interpreting the results with HDAC inhibitors such as TSA. It is important to be certain that the effect seen can be attributed



Figure 1



How might the inactive state of methylated nucleosomal DNA hypothetically be reversed? **(a)** Restoration of histone acetylation. (i) Binding of transcription factors (TFs) to methylated DNA can (ii) recruit protein(s) with HAT activity. Next (iii), the assembly of this protein complex can lead to the acetylation of core histones and (iv) subsequent transcription. **(b)** DNA demethylation. Methylated DNA may undergo demethylation either (i) through passive or (ii) active demethylation processes. The loss of DNA methylation through either

demethylation process may lead to the acetylation of histones associated with demethylated DNA. (i) Passive DNA demethylation involves the binding of nuclear factors (NFs) to the DNA. Persistent binding of nuclear factors in the vicinity of methylated DNA can lead to the failure of the maintenance methyltransferase to restore fully methylated CpG sites during DNA replication. (ii) Active DNA demethylation requires a demethylase that catalyzes a reaction leading to the removal of methyl groups from methylated cytosine residues.

directly to deacetylation of chromatin at the methylated genes under study.

### Important unanswered questions

#### Is transcriptional repression by MeCP2 entirely caused by histone deacetylation?

In both the mammalian and *Xenopus* systems, TSA treatment did not completely restore transcription from the reporter repressed by the TRD of MeCP2 [25\*\*,26\*\*]. It has been shown recently that transcriptional repression by the amino terminus of mSin3A is independent of the histone deacetylases [31]. In addition, a region in mSin3A was found to interact *in vitro* with a component of the basal transcription complex, TFIIB. As this TFIIB-

interacting region does not correspond to either of the two transcriptional repression domains in mSin3A, its functional significance remains to be tested *in vivo*. It is possible that there is a HDAC-independent mode of repression by the MeCP2/Sin3 complex, which needs to be understood.

It has been shown recently that MeCP2 can repress Sp1-activated transcription in *Drosophila* cells in a methylation-dependent manner [32]. It is unclear, however, whether MeCP2 is recruiting a co-repressor complex to repress the methylated reporter or if the repression is a consequence of steric exclusion of the Sp1 transcription factor from binding to the promoter by



MeCP2. It would be interesting to know if MeCP2 interacts with the *Drosophila* homologue of Sin3.

#### Is all indirect repression of methylated genes caused by MeCP2?

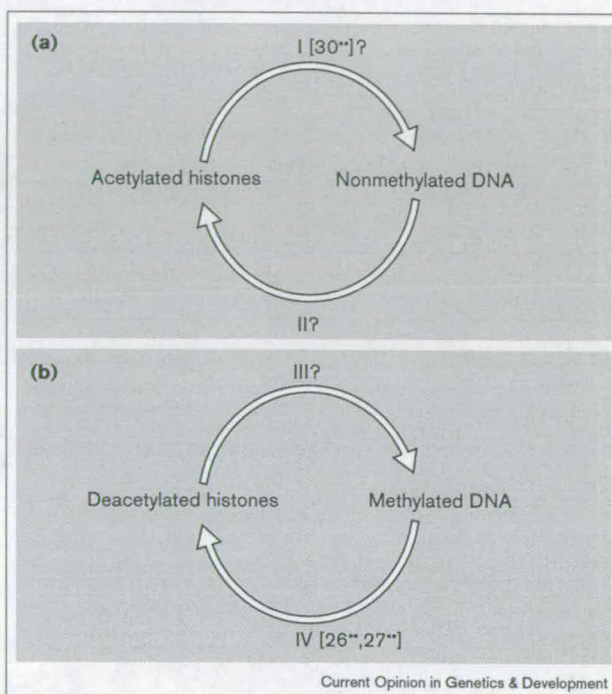
It has been shown previously that HeLa cells can efficiently repress methylated reporter constructs *in vivo* [33] but MeCP2 is below the limit of detection in HeLa cells and is therefore unlikely to repress methylated genes in this cell line (H-H Ng, A Bird, unpublished observation). Recently, a family of mammalian proteins that contain regions with homology to the methyl-CpG-binding domain of MeCP2 was discovered by searching the expressed sequence tag (EST) databases [34,35]. Interestingly, these other methyl-CpG-binding proteins can be detected in the nuclear extract of HeLa cells (H-H Ng, A Bird, unpublished observation). It is important to test if they are involved in the methylation-dependent repression seen in this cell line.

#### How can the inactive state of methylated nucleosomal DNA be reversed?

It is of great interest to understand how the transcriptional machinery gains access to specific loci (e.g. tissue-specific genes) amidst the bulk of the methylated genome. A recent report [36] suggests that even a densely methylated CpG island can be activated. The progesterone receptor (PR) gene is estrogen-responsive and, in certain human breast cancers, DNA methylation at its CpG island can lead to transcriptional inactivity. Treatment with 5-azadC can reactivate PR through DNA demethylation and activation of the estrogen receptor (ER) gene. Intriguingly, it was shown that the ectopic expression of ligand-bound ER can reactivate PR even in the presence of DNA methylation. Most importantly, induction of the silenced methylated PR by ER was blocked by co-expression of a dominant negative form of SRC-1A. SRC-1 is a coactivator with an intrinsic histone acetyltransferase (HAT) activity and can potentiate the activity of ligand-bound ER [37,38]. This result demonstrates that the association between ER and coactivator is crucial in mediating the induction of the methylated PR. Therefore, a possible mechanism to overcome the effect of DNA methylation is to recruit proteins with HAT activity that can reverse the state of histone acetylation (Figure 1a).

An alternative way of reversing the inactive state of methylated DNA is through DNA demethylation, a process that leads to the removal of the methyl group on methylated cytosine residues (Figure 1b). Active DNA demethylation activities have been reported in extracts from chick embryo [39] and certain tissue culture cells [40]. DNA demethylation can also occur via a passive route (i.e. not catalysed by enzymatic factors). Recent studies [41,42] lend support to this passive mode of DNA demethylation, which is dependent on DNA replication. Binding of nuclear factors (either transcription factors or EBNA-1) to the DNA was shown to be necessary for the loss of DNA methylation after DNA

Figure 2



Two self-propagating epigenetic cycles that could stabilize (a) transcriptionally competent or (b) transcriptionally incompetent chromatin states. The cycles are based upon hypothetical interactions between methylated/non-methylated DNA and deacetylated/acetylated histones. Preliminary evidence for the existence of arrow I comes from a recent study in the fungal system showing that the histone deacetylase inhibitor TSA can induce DNA demethylation at specific genomic loci [30\*], but this was not seen in mammalian systems [27\*\*,36\*,46\*]. It is not known whether absence of DNA methylation can induce histone hyperacetylation (arrow II). Arrow III is entirely hypothetical at this time but arrow IV, connecting DNA methylation with histone deacetylation, has firm biochemical support [26\*\*,27\*\*].

replication. It was hypothesized that the nuclear factors may block the access of maintenance DNA methyltransferase to the hemi-methylated sites that are formed during DNA replication, with the result that fully-methylated sites cannot be restored. The remaining hemi-methylated sites would subsequently be lost by dilution through further rounds of DNA replication, or these sites may be the substrate for active DNA demethylation.

#### DNA methylation and chromatin modification: a self-propagating epigenetic cycle?

An intriguing observation has been made by Selker [30\*] when using TSA to reactivate the expression of the hygromycin-resistance gene and *Tadlam* loci. The genomic methylation at these two regions was found to be reduced, raising the possibility that histone acetylation status can either directly or indirectly influence DNA methylation patterns. One hypothetical scenario is that TSA brings about histone acetylation, so that the methylated nucleosomal DNA becomes transcriptionally



active and consequently repels DNA methylation. If so, the status of chromatin modification can potentially feedback onto DNA methylation and reinforce the two modes of epigenetic silencing [30\*\*] (Figure 2a). Interestingly, a similar phenomenon of HDAC inhibitor-induced demethylation was observed in zebrafish embryos [43]. Butyrate treatment was found to inhibit *de novo* methylation of plasmids microinjected into zebrafish embryos.

Conversely, can the inactive state of chromatin attract DNA methylation (Figure 2b)? It is important to note that DNA methylation is not necessarily the primary cause of gene inactivity, but could be a secondary event that may follow transcriptional shutdown [44,45]. Perhaps the DNA methyltransferase — either independently or assisted by accessory proteins — is capable of reading the histone acetylation patterns on the chromatin and its *de novo* methyltransferase activity can respond differentially to different states of chromatin modification. In this case, deacetylated chromatin would provoke *de novo* methylation. This self-reinforcing mechanism, supported by both DNA methylation and histone deacetylation, could provide a stable state of inactive chromatin, unless overcome by other mechanisms (see Figure 1).

## Conclusion

There are now several studies indicating that gene silencing by DNA methylation involves histone deacetylation. The association between MeCP2 and a multiprotein complex capable of deacetylating histones provides a probable molecular mechanism of transcriptional repression by this protein. Now that other methyl-CpG-binding proteins are known, it will be important to establish their role in interpreting the methyl-CpG signal. It seems likely that the study of the MeCP protein family will greatly increase our understanding of the biological consequences of DNA methylation.

## Note added in proof

A recent study reports that genes silenced by CpG island methylation cannot be reactivated by TSA without prior partial demethylation of the DNA with 5-aza-dC [46\*]. It is suggested that CpG island methylation has a dominant effect over histone deacetylation.

## Acknowledgements

Huck-Hui Ng is a Darwin Trust Scholar. Research by Adrian Bird is supported by the Wellcome Trust. We thank Andrew Free and Xinsheng Nan for helpful comments on the manuscript; RN Eisenman for communicating unpublished observation; and Rainer Renkawitz for a pre-print.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
  - \*\* of outstanding interest
1. Antequera F, Bird A: **CpG islands**. In *DNA Methylation: Molecular Biology and Biological Significance*. Edited by Jost JP, Saluz HP. Basel, Switzerland: Birkhauser Verlag; 1993:169-185.
  2. Razin A, Riggs AD: **DNA methylation and gene function**. *Science* 1980, 210:604-610.
  3. Stein R, Razin A, Cedar H: ***In vitro* methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells**. *Proc Natl Acad Sci USA* 1982, 79:3418-3422.
  4. Jones PA, Taylor SM: **Cellular differentiation, cytidine analogs and DNA methylation**. *Cell* 1980, 20:85-93.
  5. Groudine M, Eisenman R, Weintraub H: **Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation**. *Nature* 1981, 292:311-317.
  6. Keshet I, Lieman-Hurwitz J, Cedar H: **DNA methylation affects the formation of active chromatin**. *Cell* 1986, 44:535-543.
  7. Tazi J, Bird A: **Alternative chromatin structure at CpG islands**. *Cell* 1990, 60:909-920.
  8. Buschhausen G, Wittig B, Graessmann M, Graessmann A: **Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene**. *Proc Natl Acad Sci USA* 1987, 84:1177-1181.
  9. Kass SU, Landsberger N, Wolffe AP: **DNA methylation directs a time dependent repression of transcription initiation**. *Curr Biol* 1997, 7:157-165.
  10. Hsieh CL: **Dependence of transcriptional repression on CpG methylation density**. *Mol Cell Biol* 1994, 14:5487-5494.
  11. Yoshida M, Kijima M, Akita M, Beppu T: **Potent and specific inhibition of mammalian histone deacetylase both *in vivo* and *in vitro* by trichostatin A**. *J Biol Chem* 1990, 265:17174-17179.
  12. Chen ZJ, Pikaard CS: **Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance**. *Genes Dev* 1997, 11:2124-2136.
  - In the plant system, TSA can substitute for an inhibitor of DNA methylation, 5-aza-dC, to derepress silent rRNA genes. This study provides an early indication for a link between DNA methylation and histone deacetylation.
  13. Pazin MJ, Kadonaga JT: **What's up and down with histone deacetylation and transcription?** *Cell* 1997, 89:315-328.
  14. Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, Eisenman RN: **Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression**. *Cell* 1997, 89:349-356.
  15. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE: **Histone deacetylase activity is required for full transcriptional repression by mSin3A**. *Cell* 1997, 89:341-347.
  16. Zhang Y, Iratni R, Erdjument-Bromage H, Tempst P, Reinberg D: **Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex**. *Cell* 1997, 89:357-364.
  17. Alland L, Muhle R, Hou H Jr, Potes J, Chin L, Schreiber-Agus N, DePinho RA: **Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression**. *Nature* 1997, 387:49-55.
  18. Heinzel T, Lavinsky RM, Mullen TM, Soderstrom M, Laherty CD, Torchia J, Yang WM, Brard G, Ngo SD, Davie JR *et al.*: **A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression**. *Nature* 1997, 387:43-48.
  19. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL, Evans RM: **Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase**. *Cell* 1997, 89:373-380.
  20. Grunstein M: **Histone acetylation in chromatin structure and transcription**. *Nature* 1997, 389:349-352.
  21. Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppensen P, Klein G, Bird A: **Purification, sequence, and cellular localisation of a novel chromosomal protein that binds to methylated DNA**. *Cell* 1992, 69:905-914.
  22. Nan X, Meehan RR, Bird A: **Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2**. *Nucleic Acids Res* 1993, 21:4886-4892.
  23. Nan X, Tate P, Li E, Bird AP: **DNA methylation specifies chromosomal localization of MeCP2**. *Mol Cell Biol* 1996, 16:414-421.
  24. Nan X, Campoy J, Bird A: **MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin**. *Cell* 1997, 88:471-481.



25. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A: **Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex.** *Nature* 1998, **393**:386-389.
- In this study, a region of MeCP2 that substantially overlaps with the transcriptional repression domain was shown to interact specifically with the mSin3A/HDAC complex. More importantly, the repressive effect of the transcriptional repression domain of MeCP2 can be alleviated by treatment with TSA – a specific HDAC inhibitor. The paper shows that a protein known to bind methyl-CpG represses via recruitment of a corepressor complex that can deacetylate histones.
26. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP: **Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription.** *Nat Genet* 1998, **19**:187-189.
- Xenopus* MeCP2 was found to co-fractionate biochemically with the Sin3/histone deacetylase complex. In addition, an inhibitor of histone deacetylases, Trichostatin A, was shown to relieve the repression mediated by the transcriptional repression domain of *Xenopus* MeCP2. Treatment with Trichostatin A was also able to restore the hypersensitive sites of a microinjected methylated reporter. The authors strongly suggest that methylation affects chromatin structure by inducing histone deacetylation and implicates MeCP2 in this phenomenon.
27. Eden S, Hashimshony T, Keshet I, Cedar H, Thorne AW: **DNA methylation models histone acetylation.** *Nature* 1998, **394**:842.
- Integrated copies of methylated HSV thymidine kinase transgenes were silenced and not enriched in acetylated histones. TSA treatment brought about the re-expression of the methylated HSV tk transgenes and rendered the methylated HSV tk transgenes DNase-I-sensitive. The results strongly suggest that the methylated transgene is silenced by deacetylation.
28. Pikaart MJ, Recillas-Targa F, Felsenfeld G: **Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators.** *Genes Dev* 1998, **12**:2852-2862.
- Insulator sequences can protect integrated transgenes from variability of expression. In the absence of such protective sequences, transgenes are susceptible to transcriptional silencing upon integration into the genome. DNA methylation and loss of histone hyperacetylation are two features of the silenced reporters. Interestingly, both 5-azac-C and histone deacetylase inhibitors were shown to restore the expression of the silenced reporters.
29. Ammerpohl A, Schmitz A, Steinmuller L, Renkawitz R: **Repression of the mouse M-lysozyme gene involves both hindrance of enhancer factor binding to the methylated enhancer and histone deacetylation.** *Nucleic Acids Res* 1998, **26**:5256-5260.
- In this study, the effect of TSA on a methylated endogenous gene was tested. The M-lysozyme gene is only active in cell types of myeloid lineage and is repressed by DNA methylation in non-myeloid cells. In non-expressing cells, TSA was shown to reactivate this myeloid-specific gene in the presence of DNA methylation without recruitment of *trans*-acting factors to the enhancer. When 5-aza-dC was used to reactivate the methylated gene through DNA demethylation, however, a footprint at the enhancer was restored. This suggests that DNA methylation can repress the M-lysozyme gene by two distinct mechanisms: histone deacetylation and direct interference with the binding of *trans*-acting factors to the enhancer.
30. Selker EU: **Trichostatin A causes selective loss of DNA methylation in *Neurospora*.** *Proc Natl Acad Sci USA* 1998, **95**:9430-9435.
- In the fungus *Neurospora*, TSA, an inhibitor of histone deacetylases, was shown to derepress two chromosomal genes which were previously silenced by DNA methylation. It is notable, however, that the drug did not affect the expression of other methylated sequences of the fungal genome (e.g. nucleolar rDNA) suggesting that DNA methylation may repress by mechanisms that do not involve histone deacetylation. Interestingly, the two derepressed genes after Trichostatin A treatment were found to undergo selective demethylation, suggesting that histone acetylation may either directly or indirectly cause loss of DNA methylation.
31. Wong CW, Privalsky ML: **Transcriptional repression by the SMRT-mSin3 corepressor: multiple interactions, multiple mechanisms, and a potential role for TFIIB.** *Mol Cell Biol* 1998, **18**:5500-5510.
32. Kudo S: **Methyl-CpG binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated.** *Mol Cell Biol* 1998, **18**:5492-5499.
33. Boyes J, Bird A: **DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein.** *Cell* 1991, **64**:1123-1134.
34. Cross SH, Meehan RR, Nan X, Bird A: **A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins.** *Nat Genet* 1997, **16**:256-259.
35. Hendrich B, Bird A: **Identification and characterization of a family of mammalian methyl-CpG binding proteins.** *Mol Cell Biol* 1998, **18**:6525-6537.
- A family of proteins with a region homologous to the DNA binding domain of MeCP2 was identified. The study of these other methyl-CpG binding proteins will provide insight into the biological consequences of DNA methylation.
36. Ferguson AT, Lapidus RG, Davidson NE: **Demethylation of the progesterone receptor CpG island is not required for progesterone receptor gene expression.** *Oncogene* 1998, **17**:577-583.
- 5-aza-dC can cause demethylation of the heavily methylated CpG island at the progesterone receptor gene and restores progesterone receptor expression through the activation of the estrogen receptor gene. However, ectopic expression of estrogen receptor can also reactivate a methylated progesterone receptor gene apparently without demethylation. As estrogen receptor can associate with proteins that possess HAT activity (SRC-1), it is likely that the HAT can reverse the de-acetylation state of methylated nucleosomal DNA.
37. McInerney EM, Tsai MJ, O'Malley BW, Katzenellenbogen BS: **Analysis of estrogen receptor transcriptional enhancement by a nuclear hormone receptor coactivator.** *Proc Natl Acad Sci USA* 1996, **93**:10069-10073.
38. Spencer TE, Jenster G, Burcin MM, Allis CD, Zhou J, Mizzen CA, McKenna NJ, Onate SA, Tsai SY, Tsai MJ, O'Malley BW: **Steroid receptor coactivator-1 is a histone acetyltransferase.** *Nature* 1997, **389**:194-198.
39. Fremont M, Siegmund M, Gaulis S, Matthies R, Hess D, Jost JP: **Demethylation of DNA by purified chick embryo 5-methylcytosine-DNA glycosylase requires both protein and RNA.** *Nucleic Acids Res* 1997, **25**:2375-2380.
40. Weiss A, Keshet I, Razin A, Cedar H: **DNA demethylation *in vitro*: involvement of RNA.** *Cell* 1996, **86**:709-718.
41. Matsuo K, Silke J, Georgiev O, Marti P, Giovannini N, Rungger D: **An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA.** *EMBO J* 1998, **17**:1446-1453.
- This study provides a possible molecular mechanism for DNA demethylation (see also [42]). The binding of transcription factors at the promoter region was shown to induce DNA demethylation. This process of demethylation is dependent on DNA replication but not transcription.
42. Hsieh CL: **Evidence that protein binding specifies sites of DNA demethylation.** *Mol Cell Biol* 1999, **19**:46-56.
- This study presents evidence that EBNA-1 binding at the Epstein-Barr virus latent replication origin contributes to the subsequent loss of methylation and that this passive mode of demethylation is dependent on DNA replication.
43. Collas P: **Modulation of plasmid DNA methylation and expression in zebrafish embryos.** *Nucleic Acids Res* 1998, **26**:4454-4461.
44. Enver T, Zhang J, Papayannopoulou T, Stamatoyannopoulos G: **DNA methylation: a secondary event in globin gene switching?** *Genes Dev* 1988, **2**:698-706.
45. Bird A: **The essentials of DNA methylation.** *Cell* 1992, **70**:5-8.
46. Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB: **Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer.** *Nat Genet* 1999, **21**:103-107.
- In this report, treatment of cells with TSA failed to restore the expression of four genes which are silenced by CpG island methylation. Intriguingly, these genes can be reactivated by TSA only after partial demethylation with 5-aza-dC. This study suggests that gene silencing by DNA methylation employs mechanisms in addition to histone deacetylation.

# Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation

Yi Zhang, Huck-Hui Ng,<sup>2</sup> Hediye Erdjument-Bromage,<sup>3</sup> Paul Tempst,<sup>3</sup> Adrian Bird,<sup>2</sup> and Danny Reinberg

Howard Hughes Medical Institute (HHMI), Division of Nucleic Acids Enzymology, Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854 USA;

<sup>2</sup>Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, UK; <sup>3</sup>Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10021 USA



# Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation

Yi Zhang,<sup>1</sup> Huck-Hui Ng,<sup>2</sup> Hediye Erdjument-Bromage,<sup>3</sup> Paul Tempst,<sup>3</sup> Adrian Bird,<sup>2</sup> and Danny Reinberg<sup>4</sup>

Howard Hughes Medical Institute (HHMI), Division of Nucleic Acids Enzymology, Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854 USA; <sup>2</sup>Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, UK; <sup>3</sup>Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10021 USA

ATP-dependent nucleosome remodeling and core histone acetylation and deacetylation represent mechanisms to alter nucleosome structure. NuRD is a multisubunit complex containing nucleosome remodeling and histone deacetylase activities. The histone deacetylases HDAC1 and HDAC2 and the histone binding proteins RbAp48 and RbAp46 form a core complex shared between NuRD and Sin3-histone deacetylase complexes. The histone deacetylase activity of the core complex is severely compromised. A novel polypeptide highly related to the metastasis-associated protein 1, MTA2, and the methyl-CpG-binding domain-containing protein, MBD3, were found to be subunits of the NuRD complex. MTA2 modulates the enzymatic activity of the histone deacetylase core complex. MBD3 mediates the association of MTA2 with the core histone deacetylase complex. MBD3 does not directly bind methylated DNA but is highly related to MBD2, a polypeptide that binds to methylated DNA and has been reported to possess demethylase activity. MBD2 interacts with the NuRD complex and directs the complex to methylated DNA. NuRD may provide a means of gene silencing by DNA methylation.

[Key Words: DNA methylation; histone deacetylase complex; nucleosome remodeling; gene silencing]

Received May 26, 1999; revised version accepted June 23, 1999.

Packaging of DNA into chromatin allows the cell to store its genetic information efficiently and has an important role in regulating gene expression (Workman and Kingston 1998). Dynamic changes in chromatin structure can facilitate or prevent the access of the transcription machinery to nucleosomal DNA, leading to transcription regulation. Recent studies have revealed two mechanisms by which chromatin structure can be altered. One mechanism involves multisubunit protein complexes that use the energy derived from ATP hydrolysis to alter the structure of, or 'remodel', nucleosomes (for review, see Tsukiyama and Wu 1997; Kadonaga 1998; Varga-Weisz and Becker 1998; Travers 1999). The other mechanism involves covalent modification of nucleosomes, in particular acetylation of the core histone tails and methylation of DNA (for review, see Grunstein 1997; Kuo and Allis 1998; Struhl 1998; Ng and Bird 1999).

Since the discovery of histone acetylation by Allfrey et al. (1964), a general correlation between histone acetylation and gene activity has been established (Hebbes et al. 1988). The enzymes that catalyze histone acetylation and deacetylation have been identified (Brownell et al. 1996; Taunton et al. 1996). Several transcriptional coactivators have histone acetyltransferase (HAT) activity, whereas several transcriptional corepressors have histone deacetylase activity (for review, see Grunstein 1997; Kuo and Allis 1998; Struhl 1998). In addition, mutagenesis studies with Gen5 and Rpd3, the prototypical histone acetyltransferase and deacetylase, respectively, confirmed the long-suspected role of histone acetylation and deacetylation in transcription regulation (Kadosh and Struhl 1998a; Kuo et al. 1998; Wang et al. 1998). Moreover, Rpd3/Sin3-dependent repression has been shown to be directly associated with the deacetylation of lysine 5 of histone H4 in the promoters of UME6-regulated genes (Kadosh and Struhl 1998b; Rundlett et al. 1998). However, how core histone acetylation/deacetylation leads to transcriptional activation/repression remains to be elucidated.

Methylation of cytosine at CpG dinucleotides is a common feature of many higher eukaryotic genomes.

<sup>1</sup>Present address: Lineberger Comprehensive Cancer Center, Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7295 USA.

<sup>4</sup>Corresponding author.

E-MAIL reinbedf@umdnj.edu; FAX (732) 235-5294.



Many studies have established a general correlation between DNA methylation and gene inactivation (Razin and Riggs 1980). However, the underlying molecular mechanism remained unknown until recently. It was found that MeCP2, a protein that specifically binds to methylated DNA, copurifies with the Sin3A/HDAC co-repressor complex and that the histone deacetylase inhibitor TSA relieves MeCP2-mediated transcriptional repression (Jones et al. 1998; Nan et al. 1998). Recently, four mammalian proteins containing regions homologous to the MeCP2 methyl-CpG-binding domain, MBD1-4, were identified by searching the expressed sequence tag (EST) databases (Hendrich and Bird 1998). Interestingly, MBD2 was claimed to be a DNA demethylase, and MBD4 was shown to be an endonuclease potentially involved in DNA mismatch repair (Bellacosa et al. 1999; Bhattacharya et al. 1999). The functions of MBD1 and MBD3 are unknown.

To develop a mechanistic understanding of how core histone acetylation regulates transcription, we have studied the histone deacetylases HDAC1 and HDAC2 (Taunton et al. 1996; Yang et al. 1996). Using a combination of conventional and affinity chromatography, we previously purified and characterized two HDAC1/HDAC2-containing histone deacetylase complexes, the Sin3A/HDAC complex, and the NuRD complex (Zhang et al. 1997, 1998a,b). The two protein complexes share four polypeptides: HDAC1, HDAC2, RbAp46, and RbAp48. In addition, each complex contains three unique polypeptides (Sin3A, SAP30, and SAP18 in the Sin3 complex, and Mi2, p70, and p32 in the NuRD complex). Interestingly, the NuRD complex also possesses nucleosome remodeling activity, most likely because of the presence of Mi2, a member of the SWI2/SNF2 helicase/ATPase family (Tong et al. 1998; Xue et al. 1998; Zhang et al. 1998a).

To gain insight into the function of the NuRD complex, we have identified its p70 and p32 subunits. We demonstrate that these polypeptides have an important role in modulating the histone deacetylase activity of NuRD. Furthermore, we provide evidence linking NuRD function to methylated DNA.

## Results

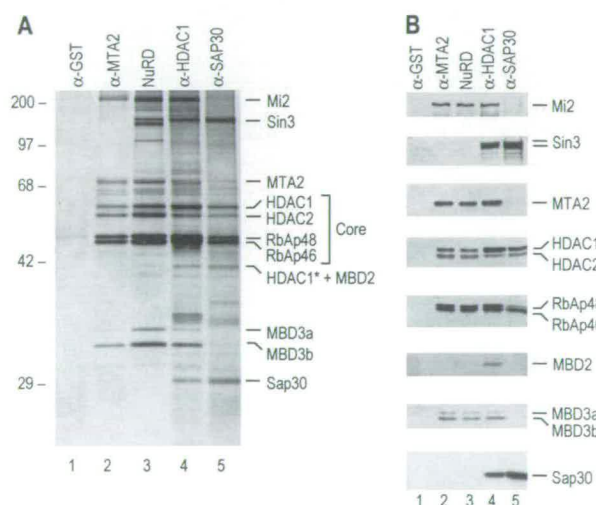
### *MTA2 and MBD3 are components of the NuRD complex*

Previously, we reported the purification of NuRD, a multisubunit complex containing both histone deacetylase and nucleosome remodeling activities (Zhang et al. 1998a). Through extensive purification using conventional methods, combined with affinity purification using antibodies against Mi2, the largest subunit of the complex, we determined that this complex is composed of seven subunits, including the SWI2/SNF2 helicase/ATPase domain-containing Mi2 protein, the two histone deacetylases HDAC1 and HDAC2, the two histone-binding proteins RbAp46 and RbAp48, and polypeptides of 70 and 32 kD (Zhang et al. 1998a). The conventionally pu-

rified complex also contains minor contaminating polypeptides that did not copurify with NuRD activity (Fig. 1A, lane 3; see also Fig. 3B of Zhang et al. 1998a).

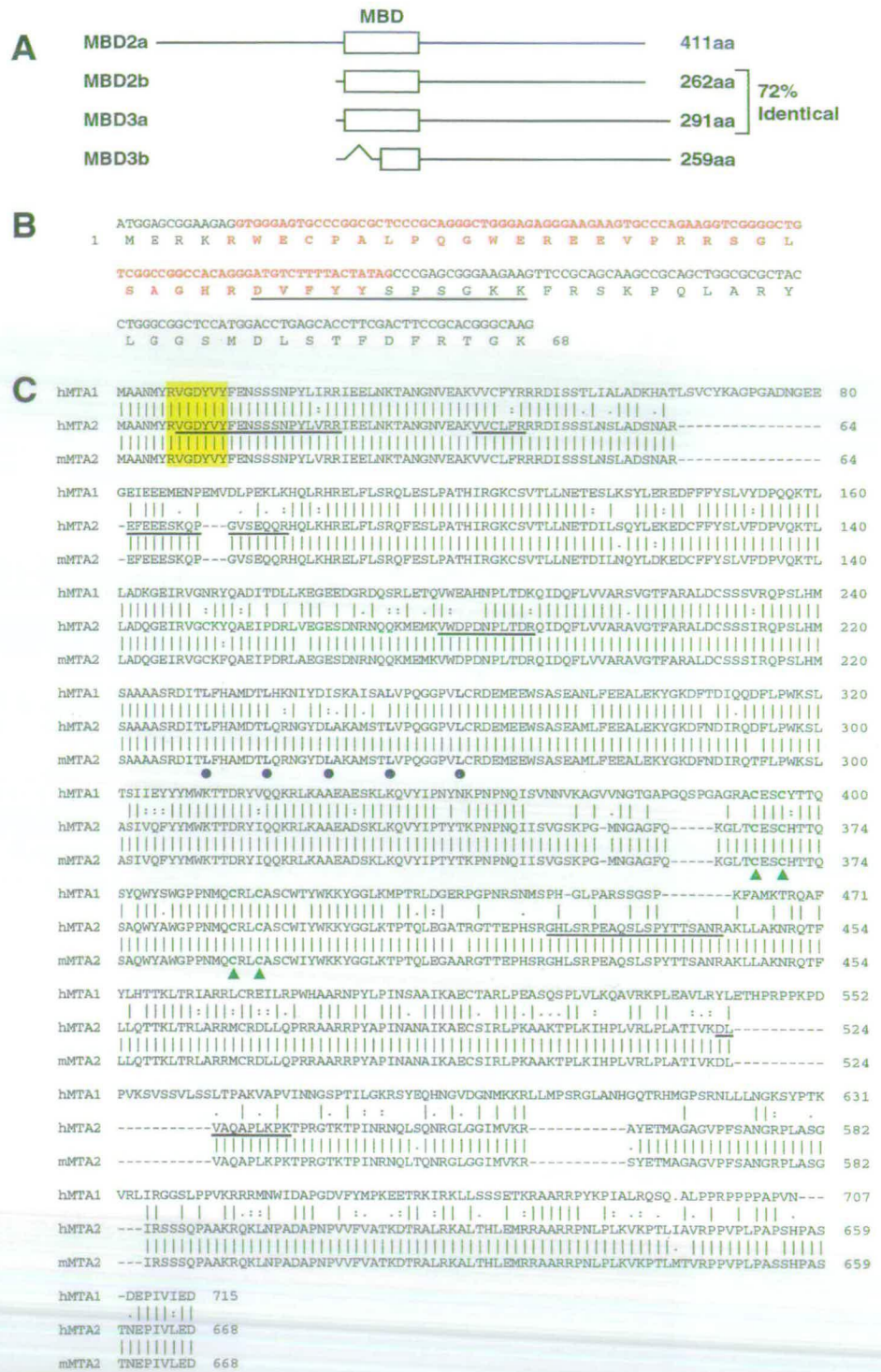
We expanded our previous studies by identifying each of the polypeptides present in the conventionally purified fraction of NuRD. Protein sequencing of the doublet migrating around 32 kD revealed that it corresponds to two in-frame spliced forms of MBD3, a member of a protein family containing the methyl-CpG binding domain (Hendrich and Bird 1998). We named the two spliced forms MBD3a and MBD3b. The major form in the NuRD complex is MBD3b, which only contains a portion of the methyl-CpG binding domain (Fig. 2A,B). Interestingly, MBD3 is highly related to MBD2 (80% similar, 72% identical; see Fig. 2A), a protein recently claimed to have DNA demethylase activity (Bhattacharya et al. 1999).

Sequencing of the 70-kD polypeptide identified it as a novel protein highly related (65% identical) to a candidate metastasis-associated protein, MTA1. Therefore, we refer to this polypeptide as MTA2 (Fig. 2C). Interestingly, MTA1 was reported to be a component of the



**Figure 1.** MTA2 and MBD3 are components of the NuRD complex. (A) Silver staining of an SDS-polyacrylamide gel containing purified NuRD (lane 3) and samples derived from different antibody columns, as indicated at top. The identity of each polypeptide was defined by microsequencing and/or Western blot analysis and is indicated at right. The histone deacetylase core complex present in both the Sin3/SAP30 complex and the NuRD complex is indicated with a bracket and denoted as Core. Peptide sequences derived from the bands labeled as MTA2, MBD3a, and MBD3b are presented in Fig. 2. Sequencing of the band labeled as HDAC1\* + MBD2 derived from HDAC1 antibody column suggest the presence of both HDAC1 and MBD2. Peptides specific for HDAC1 and MBD2 are VMTVSFHK and GLQGVGPGSNDETLLSAVASALHTSSAPITGQVSAAVEK, respectively. Mass-spectrometric analysis of the three polypeptides between Mi2 and MTA2 found in the conventionally purified NuRD fraction (lane 3) identified these contaminants as HCAP (GenBank accession no. AF020043), SA-1 (GenBank accession no. Z75330), and SB1/DXS423E (GenBank accession no. S78271). (B) Western blot analysis of the samples used in A. The proteins detected are indicated at right.





**Figure 2.** Sequence analysis of MBD3 and MTA2. (A) Schematic representation of the two forms of MBD2 and MBD3. The open box represent the methyl-CpG binding domain (MBD) initially identified in MeCP2 (also see B). The GenBank accession nos. for MBD2 and MBD3 are AF072242 and AF072247, respectively. The in-frame spliced form MBD3b only contains part of the MBD. (B) Sequences and splicing junctions of the MBD3a and MBD3b. The orange nucleotides that encode 5' portion of MBD are spliced out in MBD3b. The peptide sequence derived from the band labeled as MBD3a in Fig. 1A is underlined. Other peptides derived from the MBD3 bands of Fig. 1A have perfect match to MBD3: KQEELVQQRV, TMDLPK, GKPDNLALPVR, and NP-GVWLNTTQPLCK. (C) MTA2 is related to the metastasis-associated protein MTA1. Amino acid sequence alignment of human MTA2 (GenBank accession no. AB016591), mouse MTA2 (GenBank accession no. AF159259) with human MTA1 (GenBank accession no. U35113). Sequence alignment was performed using Gap of the GCG program (University of Wisconsin, Madison). Peptide sequences obtained from microsequencing of the band labeled MTA2 in Fig. 1A are underlined. The putative zinc-finger is indicated by green triangles. The putative leucine zipper is indicated by purple dots. The yellow box indicates a potential tyrosine kinase phosphorylation site. Human and mouse MTA2 are 98% identical and are 65% and 63% identical to human MTA1, respectively.

NURD complex based on four peptide sequences that happen to be common between MTA1 and MTA2 (Xue et al. 1998). Analysis of the amino acid sequences of MTA proteins identified one putative zinc-finger domain, one leucine zipper domain, and a potential tyrosine kinase phosphorylation site (Fig. 2C). The expression level of MTA1 was found to be elevated in metastatic breast cancer cell lines and metastatic cancer tissues, such as colorectal, gastric, and esophageal carci-

nomas (Toh et al. 1994, 1997, 1999). Similarly, we also found that the MTA2 expression level is elevated in cervical cancer tissue (data not shown). However, the cause-and-effect relationship between cancer metastasis and overexpression of MTA proteins is not known.

To further establish that MTA2 and MBD3 are integral components of the NuRD complex, antibodies against these polypeptides were produced. Western blot analysis using MTA2 antibodies demonstrated that MTA2 copu-



rified with the NuRD complex and its associated nucleosome remodeling and histone deacetylase activities (Zhang et al. 1998a). Affinity purification using anti-MTA2 antibodies resulted in the isolation of a protein complex composed of seven polypeptides, all present as major polypeptides in the conventionally purified NuRD complex (Fig. 1A, cf. lanes 2 and 3). The identity of the polypeptides present in the anti-MTA2 affinity purified complex was verified by Western blot analyses as shown in Fig. 1B (lane 2). The other polypeptides present in the conventionally purified NuRD complex (lane 3) were absent in the complex isolated through affinity chromatography using anti-Mi2 (Zhang et al. 1998a) or anti-MTA2 (Fig. 1A, lane 2) antibodies. Moreover, these polypeptides did not coelute with subunits of the NuRD complex through conventional chromatography (see Fig. 3B of Zhang et al. 1998a). Thus, we conclude that these polypeptides are contaminants.

Antibodies against MBD3 were capable of immunoprecipitating recombinant MBD3 protein, however, these antibodies failed to immunoprecipitate the NuRD complex (data not shown), suggesting that MBD3 is not accessible in the NuRD complex (see below).

Immunopurification experiments using antibodies against HDAC1, which should result in the isolation of most of the polypeptides associated with HDAC1, resulted in a complex pattern of polypeptides (Fig. 1A, lane 4). Comparison of the polypeptides present in this fraction with those present in the NuRD and Sin3 complexes established that most polypeptides in the anti-HDAC1 purified sample were present in the NuRD or Sin3 complex (Fig. 1A,B). Interestingly, the reported demethylase MBD2 was also present in the anti-HDAC1 purified sample (Fig. 1A,B, lanes 4; see below). In addition, we found that four polypeptides (HDAC1, HDAC2, RbAp48, RbAp46) were common to the Sin3 and NuRD complexes (Fig. 1A), whereas the other polypeptides (Sin3, SAP30, Mi2, MTA2, MBD3) were specific to one of the two complexes (Fig. 1A,B, lanes 2,5).

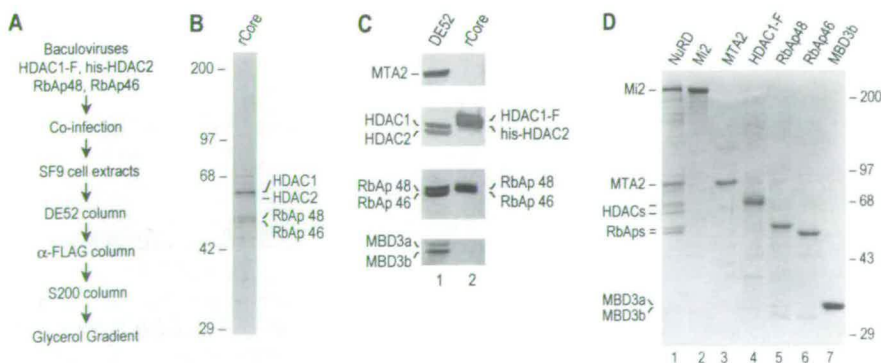
The studies described above establish that both MTA2

and MBD3 are integral components of the NuRD complex and suggest the existence of a shared core histone deacetylase complex composed of HDAC1, HDAC2, RbAp48, and RbAp46. These results also suggest a possible connection between the NuRD complex and methyl-CpG binding proteins.

#### *MTA2 promotes the assembly of a catalytically active histone deacetylase complex*

Having established that HDAC1, HDAC2, RbAp48, and RbAp46 are shared components of the Sin3 and NuRD complexes, we asked whether these four polypeptides can form a core protein complex. We began the studies by investigating whether a histone deacetylase core complex could be isolated from HeLa cells. These studies uncovered different complexes containing these four polypeptides (data not shown). However, we were unable to isolate a complex containing only these four polypeptides. A possible explanation is that the putative HDAC/RbAp core complex is limiting with respect to polypeptides that may associate with it.

We next asked whether an HDAC/RbAp core complex could be reconstituted from recombinant polypeptides. In agreement with our previous studies, which demonstrated that the HDAC and RbAp association requires cotranslation and/or the presence of a molecular chaperone (Zhang et al. 1998b), we failed to reconstitute the core complex using individually purified HDAC1 and HDAC2 from baculovirus-infected SF9 cells and RbAp46 and RbAp48 purified from *Escherichia coli* (data not shown). Thus, we performed coinfection experiments using recombinant baculoviruses expressing each of the four putative core subunits with Flag-tagged HDAC1. HDAC1 and its associated polypeptides were purified through multiple chromatographic steps, including ion exchange, immunoaffinity, gel filtration, and glycerol gradient sedimentation (Fig. 3A). Western blot analysis demonstrated cofractionation of the four polypeptides during each of the purification steps (data not shown).



**Figure 3.** Purification of a recombinant histone deacetylase core complex. (A) Schematic representation of the steps used to purify the histone deacetylase core complex. (B) Silver staining of an SDS-polyacrylamide gel containing the purified core complex. The identities of the major polypeptides and protein size markers are indicated. (C) Western blot analysis of a partially purified HeLa nuclear extract (lane 1) and the purified core complex (lane 2). The identities of the polypeptides are indicated. (D) Coomassie blue staining of an SDS-polyacrylamide gel containing the purified NuRD complex (lane 1) and individual recombinant NuRD components. Mi2, HDAC1, and MBD3b are purified from baculovirus-infected SF9 cells. MTA2 and RbAps were produced in *Escherichia coli*. Protein size markers are indicated.



Silver staining of the sample derived from the last purification step revealed the presence of four major polypeptides (Fig. 3B). The identity of these polypeptides was confirmed by Western blot analysis (Fig. 3C). The contaminating protein of about 70 kD (Fig. 3B) does not react with MTA2 antibodies (Fig. 3C). These results suggest that a complex containing HDAC and RbAp polypeptides can be formed. However, the histone deacetylase activity of this core complex was severely compromised compared to that of native NuRD complex when equal Western blot units of HDACs or RbAps were used (Table 1; data not shown).

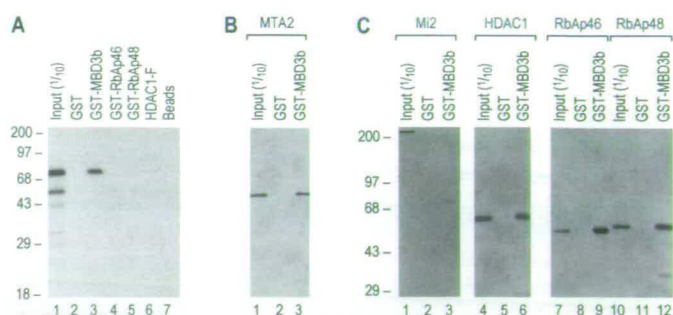
We next analyzed whether the addition of the other NuRD subunits (Mi2, MTA2, and MBD3) to the core complex could restore the activity to the level of the native NuRD complex. The addition of highly purified (Fig. 3D) recombinant Mi2 and MBD3b (or MBD3a, data not shown) purified from baculovirus infected-SF9 cells, or MTA2 produced in *Escherichia coli*, independently or in combination, did not affect the histone deacetylase activity (Table 1, and data not shown). In light of these negative results and to determine the identity of the subunits affecting enzymatic activity, we investigated the polypeptides of the NuRD complex that interact with subunits of the core HDAC/RbAp complex.

We began the study by asking whether subunits of the putative core complex, as GST-fusion or Flag-tagged proteins, could interact with MTA2 that was translated in vitro using the rabbit reticulolysate system. We found that none of the core subunits interacted with MTA2 in this assay (Fig. 4A). However, we observed an interaction between GST-MBD3b and MTA2 under the same conditions (Fig. 4A). We extended this finding by demonstrating a direct interaction using GST-MBD3b and MTA2 produced in *Escherichia coli* (Fig. 4B). We next analyzed whether MBD3b interacts with components of the core complex. This analysis demonstrated that MBD3b is able to interact with highly purified (Fig. 3D) HDAC1, RbAp48, and RbAp46 in a GST pull-down assay

**Table 1.** Effect of NuRD subunits on histone deacetylase activity

Complex/protein	<sup>3</sup> H-labeled acetate released (cpm)
—	32 ± 6
NuRD	6176 ± 213
Core (SF9)	301 ± 32
Core (SF9) + MTA2 ( <i>E. coli</i> )	357 ± 17
Core (SF9) + MBD3 (SF9)	308 ± 19
Core (SF9) + Mi2 (SF9)	324 ± 7
Core (SF9) + MTA2 (SF9)	2466 ± 250
MTA2 (SF9)	1604 ± 106

Different subunits of the NuRD complex were analyzed for histone deacetylase activity and for their ability to regulate the histone deacetylase activity of the HDAC/RbAp core complex. The activity was compared to that of the NuRD complex using equal Western blot units of HDACs. Assays were performed as described in Materials and Methods. The numbers shown represent an average of at least three independent assays.

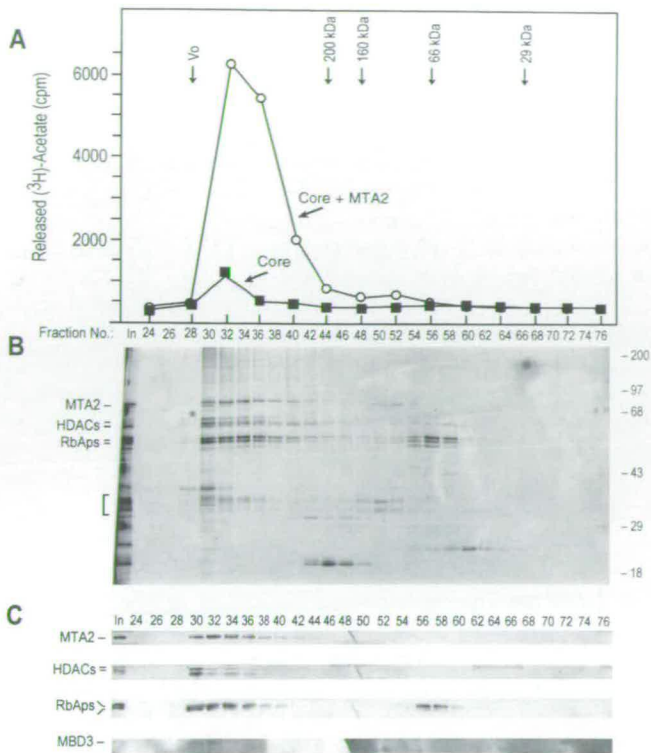


**Figure 4.** MBD3 interacts with MTA2 and components of the histone deacetylase core complex. (A) GST pull-down assays show that in vitro-translated MTA2 only interacts with MBD3. Equal amounts (5  $\mu$ l) of in vitro-translated and labeled MTA2 was incubated with 10  $\mu$ l of glutathione-agarose beads or anti-Flag beads coated with 2  $\mu$ g of proteins as indicated at the top. After extensive wash, bound proteins were eluted, resolved by SDS-PAGE, and visualized by autoradiography. Ten percent of input was loaded on lane 1. (B) MTA2 directly interacts with MBD3. An experimental procedure similar to A was used except that MTA2 was purified from *Escherichia coli* (Fig. 3D, lane 3). Bound proteins were eluted, resolved by SDS-PAGE, and visualized by Western blot. (C) MBD3 interacts with components of the core complex in a GST pull-down assay. Assays were performed as in B. Input proteins are indicated at top and are the same as those used in Fig. 3D.

(Fig. 4C, lanes 6,9,12). These interactions appear to be specific because MBD3b failed to interact with Mi2 under the same conditions (Fig. 4C, lane 3). The finding that MBD3 engages in multiple interactions with subunits of the core complex is consistent with the result demonstrating that antibodies against MBD3 could immunoprecipitate recombinant MBD3, but failed to immunoprecipitate the MBD3-containing NuRD complex, as discussed above. Collectively, these findings suggest that MBD3 is embedded within the NuRD complex.

To verify these in vitro protein-protein interaction studies, we coinfecting SF9 cells with five baculoviruses each expressing one of the four core subunits and MTA2. The complex was purified as described above (Fig. 3A). In light of the protein-protein interaction results described above, we expected that during affinity purification and gel-filtration chromatography the four subunits of the core complex would copurify, but would be separated from MTA2. However, to our surprise we observed copurification of MTA2 with the core HDAC/RbAp complex during affinity purification and gel-filtration chromatography (Fig. 5; data not shown). More importantly, we found that this complex is active in deacetylating core histones (Fig. 5A; Table 1). The activity was similar (within 2.5-fold) to that of the native NuRD complex (Table 1). Our interpretation of these results is that either the association of MTA2 with the core complex requires cotranslation or, alternatively, that mammalian MTA2 can associate with the endogenous insect MBD3 that mediates the association between MTA2 and core. Silver staining (Fig. 5B) and Western blot (Fig. 5C) analyses of the fractions derived from the gel-filtration column indicate the presence of MTA2 and the HDAC/





**Figure 5.** MTA2 is required for the formation of a functional histone deacetylase complex. (A) Histone deacetylase activities of the fractions derived from gel filtration S200 columns. Histone deacetylase core complex as well as core plus MTA2 were purified using the procedure described in Fig. 3A (see Materials and Methods for details). In the histone deacetylation assays using core- and core + MTA2-derived fractions approximately equal Western blot units of HDACs were used. The elution profiles of different size markers are indicated. (B) Silver staining of an SDS-polyacrylamide gel containing the core + MTA2 fractions derived from the S200 gel filtration column analyzed in A. The identities of the polypeptides are indicated and were confirmed by Western blot analysis which is shown in C. (C) Western blot analysis of the fractions shown in B.

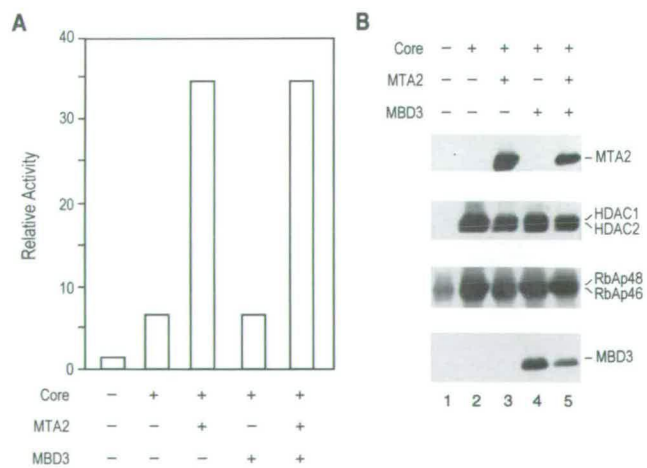
RbAp core polypeptides coeluting with HDAC activity. Moreover, the silver-staining analysis revealed polypeptides in the 30-kD range coeluting with histone deacetylase activity (Fig. 5B). Western blot analysis revealed that antibodies directed against the human MBD3 protein reacted, although weakly, with a polypeptide of ~30 kD that coeluted with the histone deacetylase activity. We concluded that the MBD3-immunoreactive 30-kD polypeptide is likely the SF9 cell-derived MBD3 (Fig. 5C). The silver-staining analysis demonstrates that the complex is highly pure and suggests that the coelution of HDACs, RbAps, MTA2, and insect-derived MBD3 is a functional association and not the result of protein aggregation.

To further analyze the specificity of MTA2 in directing the formation of an enzymatically active histone deacetylase complex, we performed coinfections as above, but coinfecting a baculovirus-expressing MBD3 with components of the HDAC core complex in the pres-

ence and absence of MTA2. Following affinity purification, the purified complexes were divided and analyzed for histone deacetylase activity (Fig. 6A) and protein composition by western blots (Fig. 6B). As above, the core complex was severely compromised in its enzymatic activity (Fig. 6A). The presence of MBD3 was without effect. However, coinfection with MTA2 in the presence or absence of MBD3 resulted in the recovery of a complex that was enzymatically active (Fig. 6A). Consistent with our finding that MTA2 can direct the formation of an enzymatically active histone deacetylase complex, we found that infection of SF9 cells with a recombinant baculovirus-expressing Flag-tagged MTA2 followed by purification on a column containing antibodies against the Flag tag resulted in the isolation of an MTA2-containing complex that was enzymatically active (Table 1). Silver staining and Western blot analysis revealed that in addition to MTA2, polypeptides corresponding to the insect HDAC and RbAp were present in the affinity-purified sample (data not shown). Collectively, these studies demonstrate that MTA2 directs the assembly of an active histone deacetylase complex and that the association of MTA2 with the core HDAC/RbAp complex requires MBD3.

#### *The NuRD complex interacts with MBD2 and is tethered to methylated DNA*

The finding that MBD3 is a component of the NuRD complex suggests a possible connection between the NuRD complex and methylated DNA. Therefore, we in-



**Figure 6.** MTA2 directs the formation of an enzymatically active histone deacetylase complex. (A) Relative histone deacetylase activities of different coinfections. SF9 cells were infected with different combinations of viruses. The complexes were affinity purified through Flag-tagged HDAC1 using antibodies against Flag. The purified complexes were divided for deacetylase assay shown in A and Western blot analysis shown in B. (+) Presence of the viruses in the coinfection. (B) Western blot analysis of the samples used in histone deacetylation assays shown in A. Antibodies against RbAps cross-reacted with bleached heavy chain in lane 1.

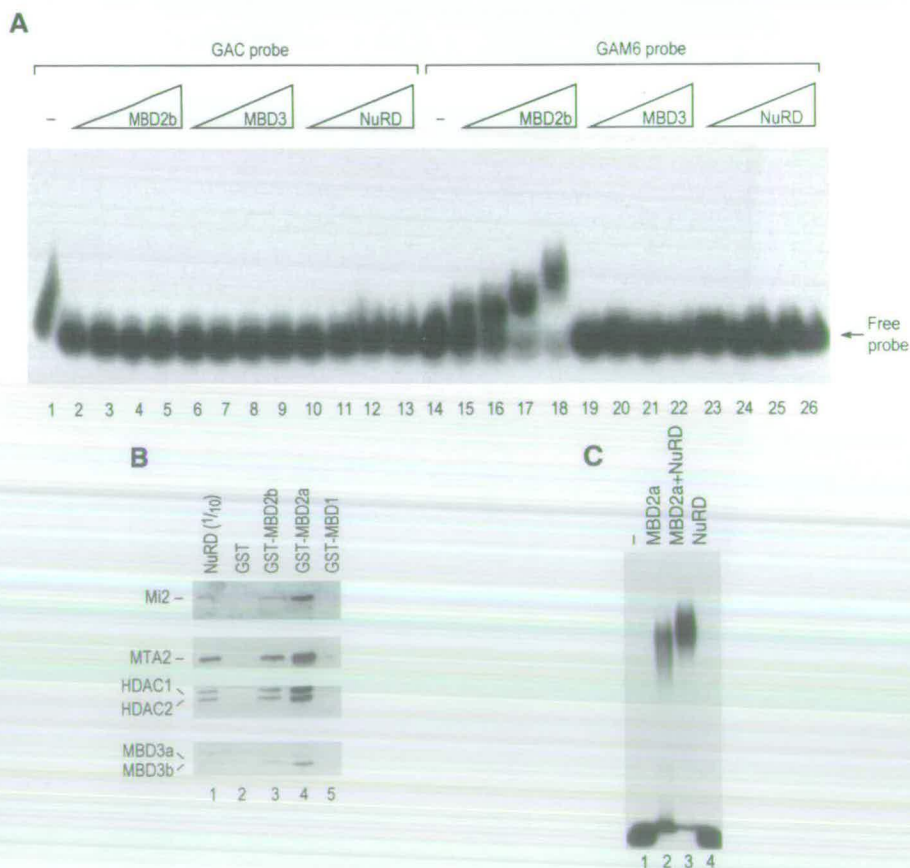


investigated whether the NuRD complex could specifically bind to methyl-CpG-containing DNA using a gel mobility-shift assay. We found that NuRD and MBD3 failed to bind to DNA. However, under the same conditions, MBD2 bound specifically to methylated DNA (Fig. 7A, cf. lanes 2–6 with 15–18). This is in agreement with previous studies (Hendrich and Bird 1998). Moreover, because MBD3 is likely embedded in the NuRD complex, and the major form of MBD3 present in the complex is MBD3b, which only contains part of the methyl-CpG-binding domain (Fig. 2A), it is not surprising that NuRD failed to directly bind to methylated DNA.

Because MBD3 is highly related to MBD2 and we found that MBD3 is able to interact with each component of the histone deacetylase core complex (Fig. 4C), we investigated whether MBD2 associates with histone deacetylase complexes. Immunoprecipitation experiments using antibodies against HDAC1 resulted in the immunoprecipitation of MBD2 (Fig. 1B, lane 4). However, as described above, MBD2 is absent from the NuRD complex isolated through conventional (Fig. 1B, lane 3) or affinity (Fig. 1B, lane 2) purification procedures. Nonetheless, the similarity between MBD2 and MBD3 proteins prompted us to analyze whether MBD2 could physically interact with the NuRD complex. In this experiment, full-length and a truncated form of MBD2 (Hendrich and Bird 1998) were analyzed. The repression domain of MBD1 (H-H. Ng and A. Bird, unpubl.), a protein related to MBD2 and MBD3 through the methyl-

CpG DNA-binding domain (Hendrich and Bird 1998), was used as a control. Different GST-MBD fusion proteins were independently incubated with purified NuRD complex and possible interactions were analyzed by GST pull-down assays followed by Western blot analyses using antibodies against different components of the NuRD complex. Whereas GST-MBD1 failed to interact with the NuRD complex, both forms of MBD2 interacted with the NuRD complex, and the full-length form of MBD2 interacted more efficiently (Fig. 7B). Thus, we conclude that MBD2 is able to interact directly with components of the NuRD complex.

The interaction between MBD2 and the NuRD complex prompted us to analyze whether MBD2 could tether the NuRD complex to methylated DNA. We analyzed this possibility using the gel mobility-shift assay with methylated DNA as the probe. Under these conditions, and in agreement with the studies presented above, NuRD failed to bind to DNA, but MBD2 bound to the methylated DNA specifically (Fig. 7C; data not shown). Importantly, the addition of NuRD to a DNA-binding assay containing MBD2 resulted in the production of a new DNA-protein complex that migrates slower than the MBD2-DNA protein complex. Additionally the amount of DNA in the NuRD-MBD2 ternary complex was greater than the amount of DNA in the binary MBD2-containing complex. The observed supershift is specific because addition of 3  $\mu$ g of BSA (10  $\times$  the amount of NuRD used) or other proteins was without affect (data



**Figure 7.** The NuRD complex can be targeted to methylated DNA by MBD2. (A) Gel mobility shift assay shown that the NuRD complex does not bind to methylated DNA directly. The GAM6 probe contains six methyl-CpG sites and was previously described (Nan et al. 1993). Binding reactions were resolved on a 2% agarose gel as described in Materials and Methods. (B) GST pull-down assays demonstrates an interaction between MBD2 and the NuRD complex. The assays were performed as described in Fig. 4B using different GST fusion proteins as indicated at the top. (C) Gel mobility-shift assay shown that the NuRD complex is able to super-shift the DNA-MBD2 binary complex (cf. lanes 2 and 3). The probe used is MeCG11, which contains 27 methyl-CpG sites and was described previously (Ng et al. 1999). Binding reactions were resolved on a 1% agarose gel as described in Materials and Methods.



not shown). Therefore, we conclude that NuRD is tethered to methylated DNA via MBD2 and that the interaction of NuRD with MBD2 stabilizes the MBD2–DNA protein complex.

## Discussion

In this study we have identified MTA2 and MBD3 as two subunits of the NuRD complex. MTA2 is related to the metastasis-associated protein MTA1, whereas MBD3 is related to MBD2, a methyl-CpG-binding domain-containing protein recently claimed to possess demethylase activity. In addition, we provide evidence that MTA2, through an interaction with MBD3, has an important role in modulating the enzymatic activity of the histone deacetylase core complex composed of HDAC1 and HDAC2 and RbAp48 and RbAp46. Furthermore, we demonstrate that the methyl-CpG binding protein MBD2 is able to tether the NuRD complex to CpG-methylated DNA.

### *The connection between NuRD and cellular proliferation*

The nucleosome remodeling and histone deacetylase NuRD complex contains seven subunits. In addition to the histone deacetylase core HDAC1, HDAC2, RbAp46, and RbAp48, it contains Mi2, MTA2, and MBD3. It is interesting to note that MTA2 is 65% identical to the metastasis-associated protein MTA1, whose expression level was found to be elevated in different types of cancer tissue as well as metastatic breast cancer cell lines (Toh et al. 1994, 1997, 1999). Similarly, we found that MTA2 is highly expressed in rapidly dividing cells at both the RNA and protein level (data not shown) indicating a correlation between MTA2 expression and cellular proliferation. Another component that links the NuRD complex to cellular transformation is MBD3. This protein was initially identified because it contains a domain related to the methyl-CpG-binding domain of MeCP2 (Hendrich and Bird 1998). However, in agreement with a previous report (Hendrich and Bird 1998), MBD3 is not able to bind to methyl-CpG specifically (Fig. 7A). Moreover, two alternately spliced forms of MBD3 were identified. The predominant form present in the NuRD complex is a spliced variant deleted of the methylated CpG-binding domain. Interestingly, MBD3 is highly related to MBD2, which was also identified as a colon cancer antigen (Scanlan et al. 1998). The fact that two components of the NuRD complex are linked to malignancy, in conjunction with the fact that patients with dermatomyositis, who produce antibodies against Mi2 (Seelig et al. 1996), have an increased risk to malignancy (Airio et al. 1995), suggests a link between the NuRD complex and malignancy. Indeed the accumulative evidence strongly suggest that histone deacetylases, in particular HDAC1 and HDAC2, are associated with polypeptides that regulate cellular proliferation. The first mammalian HDAC

complex to be described was Sin3, which associates with the Max–Mad heterodimer required for cellular differentiation (for review, see Amati and Land 1994; Pazin and Kadonaga 1997). Additionally, Sin3 associates with the corepressor NcoR/SMRT, which negatively regulates transcription of genes targeted by nuclear hormone receptors and has been linked with acute promyelocytic leukemia (Grignani et al. 1998; Lin et al. 1998). Furthermore, HDAC1 was found to be associated with the tumor suppressor Rb protein (Brehm et al. 1998; Luo et al. 1998; Magnaghi-Jaulin et al. 1998). Moreover, our recent studies have uncovered that SAP30, a component of the Sin3 complex (Laherty et al. 1998; Zhang et al. 1998b), associates with the Rb-binding protein RBP1 as well as with the p53-binding protein p33<sup>ING1</sup> (Y. Zhang and D. Reinberg unpubl.).

### *NuRD is targeted to methylated DNA*

Since the discovery of the first nucleosome remodeling factor, the SWI/SNF complex, about a dozen different chromatin remodeling factors, from different organisms, have been described (for review, see Bjorklund et al. 1999; Travers 1999). However, it is not known whether these remodeling factors function on a genome-wide basis or whether they are targeted to specific genes. Genetic studies indicate that transcription of only a fraction of genes is affected by the SWI/SNF complex (Holstege et al. 1998), but until recently it was not clear how the SWI/SNF complex is recruited to promoters of SWI/SNF-regulated genes. The first evidence suggesting that the SWI/SNF complex can be targeted to specific genes came from studies on the human  $\beta$ -globin promoter. It was demonstrated that a SWI/SNF-related chromatin remodeling factor E-RC1 is required for the erythroid Krüppel-like factor, ELKF, to activate transcription in vitro using a chromatin-assembled template (Armstrong et al. 1998). Recently, Nasmyth and coworkers demonstrated that the SWI/SNF complex is recruited to the HO promoter by SWI5 in vivo in a cell cycle-dependent manner (Cosma et al. 1999). Therefore, targeted nucleosome remodeling is at least one mechanism by which nucleosome remodeling complexes can function to activate transcription. Similarly, histone deacetylase complexes may also be targeted to specific promoters. For example, the histone deacetylase HDAC2 interacts with the transcription factor YY1, and thus could be recruited to promoters containing YY1-binding sites (Yang et al. 1996). Moreover, the Sin3/HDAC complex can be recruited by Mad–Max, Ume6, and unliganded nuclear hormone receptors to specific promoters (for review, see Pazin and Kadonaga 1997). Additionally, the Sin3–histone deacetylase complex can also be recruited via the methyl-CpG binding protein MeCP2 to methylated regions of the genome (Jones et al. 1998; Nan et al. 1998). Therefore, targeted deacetylation represents at least one mechanism by which histone deacetylase complexes can be recruited.

Previously, we and others have shown that the NuRD complex has both nucleosome-remodeling and histone



deacetylase activities (Tong et al. 1998; Xue et al. 1998; Zhang et al. 1998a). It was also shown that recruitment of the NuRD complex to a promoter results in transcription repression (Kehle et al. 1998; Xue et al. 1998). We proposed that the NuRD complex is targeted to specific genes by transcription factors (Zhang et al. 1998a). The NuRD-subunit MTA2 contains a zinc-finger (CX<sub>2</sub>CX<sub>17</sub>CX<sub>2</sub>C) belonging to the type found in transcription factors that bind to the GATA sequence involved in hematopoiesis and heart development (Orkin 1992; Lyons 1996). These observations prompted us to analyze whether MTA2 or the NuRD complex could bind to the human  $\beta$ -globin gene promoter using the gel mobility-shift assay. This assay failed to detect any DNA binding activity (data not shown). This negative result can be explained because several amino acids involved in GATA sequence recognition are not conserved in the MTA2 zinc-finger domain (Omichinski et al. 1993).

Similar to previous studies demonstrating that Mad, a protein that interacts with the Sin3-histone deacetylase corepressor complex, but is not an integral component of the complex, can recruit the corepressor complex to specific promoters (for review, see Pazin and Kadonaga 1997), we found that the methyl-CpG binding protein MBD2, although not an integral component of the NuRD complex (Fig. 1B), has the ability to directly interact with the NuRD complex and recruit it to methylated DNA (Fig. 7). The implication of this result is that the NuRD complex may play a role in methylation-associated gene silencing. In support of this hypothesis are recent findings demonstrating that artificial recruitment of MBD2 to promoters results in repression of transcription, which was reversed by the histone deacetylase inhibitor Trichostatin A (Ng et al. 1999). All of the evidence discussed above suggests that MBD2 functions as a transcriptional repressor. However, it is worth noting that MBD2 was recently reported to have demethylase activity (Bhattacharya et al. 1999). Because DNA methylation causes gene silencing, a demethylase is likely to function as a transcriptional activator. It remains to be determined how MBD2 activity can be converted from a repressor to an activator. Given the similarity between MBD2 and MBD3 (72% identical), we tested the recombinant MBD3 protein (produced in *Escherichia coli* or SF9 cells) and the native NuRD complex for demethylase activity and failed to detect such an activity (data not shown).

Can NuRD also be recruited to promoters by gene-specific DNA-binding protein? It has been shown that the *Drosophila* gap gene *hunchback* can interact with *Drosophila* Mi2 and potentially recruit the NuRD complex to repress *hunchback*-regulated HOX genes (Kehle et al. 1998). Recently, it was also reported that Ikaros, a zinc finger DNA-binding protein essential for lymphocyte lineage determination, through interaction with Mi2, recruits the NuRD complex to heterochromatin regions (Kim et al. 1999). It is therefore likely that DNA-binding proteins, which in principle have a restricted specificity as compared to MBD2, can also target the NuRD complex to specific genes. The interplay between

sequence-specific DNA-binding proteins and MBD2 in recruiting the NuRD complex to specific genes may represent another pathway for regulation of the NuRD complex.

#### *The function of the NuRD complex*

The NuRD complex was purified based on its histone deacetylase and nucleosome remodeling activities (Tong et al. 1998; Xue et al. 1998; Zhang et al. 1998a). The biochemical association of the histone deacetylase and nucleosome remodeling activities suggest that these two enzymatic activities are functionally related. It is conceivable that nucleosome remodeling may be required for nucleosomal histone deacetylation in vivo. Therefore, the presence of these two activities in one protein complex may represent an efficient way to facilitate dynamic changes in nucleosome structure. Based on the finding that a similar protein complex is also present in *Xenopus* eggs (Wade et al. 1998), we believe that the function of the NuRD complex is rather general. Several reports have already shed light on its potential biological function. Studies in *Drosophila* implicated NuRD in repression of homeotic (HOX) genes and Polycomb-group genes (Kehle et al. 1998). In addition, NuRD has been shown to be associated with the zinc-finger DNA-binding protein Ikaros in toroidal structures presumed to be associated with centromeric heterochromatin in the G<sub>1</sub> and S phase of the T-lymphocyte cell cycle (Brown et al. 1997; Kim et al. 1999). This suggests that NuRD is involved in centromeric silencing, consistent with the requirement of histone deacetylase activity for the maintenance of the underacetylated state of centromeric histones (Ekwall et al. 1997). It is interesting to note that Ikaros is required for the differentiation of all three lymphoid lineage (Georgopoulos et al. 1994). Because Ikaros can recruit NuRD, it is possible that NuRD may play an important role in lymphocyte development. We provide evidence indicating that NuRD is potentially involved in transcriptional repression of methyl-CpG through an interaction with the methyl-CpG-binding protein MBD2. Therefore, the two best-characterized histone deacetylase complexes, Sin3 and NuRD, are both involved in transcriptional repression of methylated DNA. MBD2 was recently shown to be a component of the MeCP1 complex (Meehan et al. 1989; Ng et al. 1999). Because MeCP2 binds methylated DNA much tighter than MeCP1 (Meehan et al. 1989; Lewis et al. 1992), it is possible that the Sin3/HDAC complex may be involved in long-term silencing of methylated DNA sequences, whereas the NuRD complex may be involved in transient silencing of some methylated genes.

To maintain epigenetic silencing of genes through multiple cell divisions, newly deposited histones that are acetylated in the cytoplasm must be deacetylated (Jeppesen 1997). The identity of the histone deacetylase complex responsible for this function is unknown. It is intriguing that Ikaros forms a higher-order toroidal structure with NuRD that colocalizes with components of the DNA replication machinery upon T-cell activation



(Avitahl et al. 1999; Kim et al. 1999). Therefore, the nucleosome-remodeling and histone deacetylase activities of the NuRD complex are ideal candidates for deacetylating the deposition-related acetylated lysine residues during chromatin maturation.

## Materials and methods

### *Purification of the NuRD complex and cloning of MTA2 and MBD3*

The procedure for conventional purification of the NuRD complex has been described (Zhang et al. 1998a). The purified NuRD complex derived from Mono S column was concentrated using a centricon concentrator and was resolved on a 10% SDS-polyacrylamide gel. After Coomassie blue staining, protein bands were excised and subjected to in-gel tryptic digestion. The previously identified polypeptides were verified by mass-spectrometric analysis (Erdjument-Bromage et al. 1998; Geromanos et al. 1998). Protein bands corresponding to p70 and p32 were sequenced as described (Zhang et al. 1997). The peptide sequences obtained were used to search the EST database. Multiple EST clones were identified and were used to construct the full-length cDNA clones encoding p70 and p32. Clones encoding alternative spliced forms of p32 (MBD3a and MBD3b) were identified in human and mouse by sequencing EST clones (see text). Affinity purification using anti-MTA2, anti-HDAC1, and anti-SAP30 was based on a previously published procedure (Zhang et al. 1998b) using HeLa nuclear extracts fractionated on phosphocellulose and DEAE-52 columns.

### *Baculoviruses, recombinant core complex, and antibodies*

Baculoviruses expressing RbAp46, RbAp48, and HDAC1 have been described (Hassig et al. 1997; Verreault et al. 1998). Baculovirus-expressing HDAC2 was constructed using the pFastBac HTb vector (GIBCO BRL) and was a gift from Dr. Ed Seto (Moffitt Cancer Center, Tampa, FL). cDNAs encoding His-tagged Mi2 and MBD3 and Flag-tagged MTA2 were constructed using the pVL1392 vector and the recombinant viruses were generated using BaculoGold DNA (Pharmigen). Histone deacetylase core complex and core plus MTA2 were purified using the procedure shown in Figure 3A. Extracts derived from SF9-infected cells (11 grams) were loaded onto a 35-ml DEAE-cellulose column and were eluted with six column volumes using a linear gradient of buffer C (20 mM Tris-HCl at pH 7.9, 0.2 mM EDTA, 10 mM  $\beta$ -ME, 0.2 mM PMSF, 10% glycerol) from 50 mM to 400 mM of KCl (BC50-BC400). The fractions containing the peak of the recombinant core complex (160–269 mM KCl) were pooled and incubated with 0.5 ml of anti-Flag M2 affinity gel (Eastman Kodak) at 4°C for 2 hr. Proteins were eluted with 0.5 ml of BC400 containing 0.1 mg of Flag peptide. The affinity purified core complex was further purified through a gel filtration Sephadex-200 (10/30) column. The complex eluted with an apparent mass of ~450 kD. Finally, the core complex was purified by sedimentation through a 15%-50% glycerol gradient. Recombinant HDAC1-Flag and MBD3-His proteins were purified using anti-Flag M2 affinity gel (Eastman Kodak) and Ni<sup>2+</sup>-NTA agarose (Qiagen), respectively. Recombinant RbAp48 and RbAp46 were generated by cleaving the GST fusion proteins produced in *Escherichia coli* with thrombin. Antibodies against Mi2, Sin3, HDAC1, HDAC2, RbAp48, RbAp46, SAP30, and MBD2 were described previously (Zhang et al. 1998a; Ng et al. 1999). Antibodies against MTA2 and MBD3 were generated by injecting recombinant proteins into rabbits.

### *Histone deacetylase assays*

Core histone octamers were purified from HeLa cells as described (Zhang et al. 1998b) and acetylated with yeast Hat1p in buffer containing 50 mM HEPES (pH 7.8), 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF, 10 mM sodium butyrate, 10% glycerol, 5  $\mu$ M [<sup>3</sup>H]-acetyl coenzyme A, and 0.5 mg/ml core histones. Acetylated core histones were purified on a phosphocellulose column. Histone deacetylase assays were performed as described (Zhang et al. 1998b) with 1  $\mu$ g of labeled core histones.

### *Gel mobility shift assays*

Gel mobility-shift assays were performed with modifications from a published procedure (Meehan et al. 1989; Nan et al. 1993). Data presented in Figure 7A used GAM6 probe and the binding reactions were carried out in 20 mM HEPES buffer at pH 7.9, 1 mM EDTA, 3 mM MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, 4% glycerol, 0.1% Triton X-100, and 500 ng of poly [d(G-C)] at 25°C for 30 min. Each reaction contains 0.1 ng of probe and 10–80 ng of recombinant MBDs or 50–300 ng of the NuRD complex. The reactions were loaded onto a 2% agarose gel and resolved with 0.5 $\times$  TBE containing 5 mM magnesium acetate and 3% glycerol. The supershift assay shown in Figure 7C used the MeCG11 probe in a reaction similar to the above except it also contained 100 mM NaCl, 0.1 mg/ml BSA, and 300 ng of *Escherichia coli* DNA as competitor. Approximately 200 ng of GST-MBD2a and 300 ng of NuRD were used. Reactions were incubated on ice for 2 hr before electrophoresis on a 1% agarose gel.

### *GST pull-down assays*

Approximately 2–3  $\mu$ g of GST fusion proteins were bound to glutathione-Sepharose beads and incubated with in vitro translated, or purified recombinant proteins, or purified NuRD complex at 4°C for at least 4 hr. Beads were then washed three times with buffer containing 300 mM KCl and 0.05% NP40 for three times before loading onto SDS-PAGE. The bound proteins were revealed by fluorography or Western blot analysis.

## Acknowledgments

We are grateful to Drs. E. Seto and S.C. Tsai for baculovirus expressing HDAC2, to S. Schreiber for baculovirus expressing HDAC1, to A. Verreault and B. Stillman for baculoviruses expressing RbAp46 and RbAp48, and to D. Gottschling for the plasmid encoding yeast Hat 1. We thank Dr. George Orphanides for critical reading of the manuscript. We also thank members of the Reinberg laboratory for stimulating discussions during the course of this work. Y.Z. is a recipient of a National Institutes of Health (NIH) fellowship (1F32GM19515-01). H.H.N. holds a Darwin Trust Scholarship. D.R. is supported by a grant from NIH (GM-48518) and from the HHMI. A.B. is supported by grants from the Wellcome Trust. P.T. is supported by grants from the National Science Foundation and the National Cancer Institute.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked 'advertisement' in accordance with 18 USC section 1734 solely to indicate this fact.

## References

- Airio, A., E. Pukkala, and H. Isomaki. 1995. Elevated cancer incidence in patients with dermatomyositis: A population based study. *J. Rheumatol.* 22: 1300–1303.



- Allfrey, V., R.M. Faulkner, and A.E. Mirsky. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci.* **51**: 786-794.
- Amati, B. and H. Land. 1994. Myc-Max-Mad: A transcription factor network controlling cell cycle progression, differentiation and death. *Curr. Opin. Genet. Dev.* **4**: 102-108.
- Armstrong, J.A., J.J. Bieker, and B.M. Emerson. 1998. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLf in vitro. *Cell* **95**: 93-104.
- Avitahl, N., S. Winandy, C. Friedrich, B. Jones, Y. Ge, and K. Georgopoulos. 1999. Ikaros sets thresholds for T cell activation and regulates chromosome propagation. *Immunity* **10**: 333-343.
- Bellacosa, A., L. Cicchillitti, F. Schepis, A. Riccio, A.T. Yeung, Y. Matsumoto, E.A. Golemis, M. Genuardi, and G. Neri. 1999. MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proc. Natl. Acad. Sci.* **96**: 3969-3974.
- Bhattacharya, S.K., S. Ramchandani, N. Cervoni, and M. Szyf. 1999. A mammalian protein with specific demethylase activity for mCpG DNA [see comments]. *Nature* **397**: 579-583.
- Bjorklund, S., G. Almouzni, I. Davidson, K.P. Nightingale, and K. Weiss. 1999. Global transcription regulators of eukaryotes. *Cell* **96**: 759-767.
- Brehm, A., E.A. Miska, D.J. McCance, J.L. Reid, A.J. Bannister, and T. Kouzarides. 1998. Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**: 597-601.
- Brown, K.E., S.S. Guest, S.T. Smale, K. Hahm, M. Merken-schlager, and A.G. Fisher. 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**: 845-854.
- Brownell, J.E., J. Zhou, T. Ranalli, R. Kobayashi, D.G. Edmondson, S.Y. Roth, and C.D. Allis. 1996. Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**: 843-851.
- Cosma, M.P., T. Tanaka, and K. Nasmyth. 1999. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* **97**: 299-311.
- Ekwall, K., T. Olsson, B.M. Turner, G. Cranston, and R.C. Allshire. 1997. Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* **91**: 1021-1032.
- Erdjument-Bromage, H., M. Lui, L. Lacomis, A. Grewal, R.S. Annan, D.E. McNulty, S.A. Carr, and P. Tempst. 1998. Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis. *J. Chromatogr. A* **826**: 167-181.
- Georgopoulos, K., M. Bigby, J.H. Wang, A. Molnar, P. Wu, S. Winandy, and A. Sharpe. 1994. The Ikaros gene is required for the development of all lymphoid lineages. *Cell* **79**: 143-156.
- Geromanos, S., J. Philip, G. Freckleton, and P. Tempst. 1998. Injection adaptable fine ionization source ('JaFIS') for continuous flow nano-electrospray. *Rapid Commun. Mass. Spectrom.* **12**: 551-556.
- Grignani, F., S. De Matteis, C. Nervi, L. Tomassoni, V. Gelmetti, M. Cioce, M. Fanelli, M. Ruthardt, F.F. Ferrara, I. Zamir, C. Seiser, M.A. Lazar, S. Minucci, and P.G. Pelicci. 1998. Fusion proteins of the retinoic acid receptor- $\alpha$  recruit histone deacetylase in promyelocytic leukaemia. *Nature* **391**: 815-818.
- Grunstein, M. 1997. Histone acetylation in chromatin structure and transcription. *Nature* **389**: 349-352.
- Hassig, C.A., T.C. Fleischer, A.N. Billin, S.L. Schreiber, and D.E. Ayer. 1997. Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell* **89**: 341-347.
- Hebbes, T.R., A.W. Thorne, and C. Crane-Robinson. 1988. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* **7**: 1395-1402.
- Hendrich, B. and A. Bird. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18**: 6538-6547.
- Holstege, F.C., E.G. Jennings, J.J. Wyrick, T.I. Lee, C.J. Hengartner, M.R. Green, T.R. Golub, E.S. Lander, and R.A. Young. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717-728.
- Jeppesen, P. 1997. Histone acetylation: A possible mechanism for the inheritance of cell memory at mitosis. *BioEssays* **19**: 67-74.
- Jones, P.L., G.J. Veenstra, P.A. Wade, D. Vermaak, S.U. Kass, N. Landsberger, J. Strouboulis, and A.P. Wolffe. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**: 187-191.
- Kadonaga, J.T. 1998. Eukaryotic transcription: An interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92**: 307-313.
- Kadosh, D. and K. Struhl. 1998a. Histone deacetylase activity of Rpd3 is important for transcriptional repression in vivo. *Genes & Dev* **12**: 797-805.
- . 1998b. Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. *Mol. Cell. Biol.* **18**: 5121-5127.
- Kehle, J., D. Beuchle, S. Treuheit, B. Christen, J.A. Kennison, M. Bienz, and J. Muller. 1998. dMi-2, a hunchback-interacting protein that functions in polycomb repression. *Science* **282**: 1897-1900.
- Kim, J., S. Sif, B. Jones, A. Jackson, J. Koipally, E. Heller, S. Winandy, A. Viel, A. Sawyer, T. Ikeda, R. Kingston, and K. Georgopoulos. 1999. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity* **10**: 345-355.
- Kuo, M.H. and C.D. Allis. 1998. Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays* **20**: 615-626.
- Kuo, M.H., J. Zhou, P. Jambeck, M.E. Churchill, and C.D. Allis. 1998. Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. *Genes & Dev* **12**: 627-639.
- Laherty, C.D., A.N. Billin, R.M. Lavinsky, G.S. Yochum, A.C. Bush, J.M. Sun, T.M. Mullen, J.R. Davie, D.W. Rose, C.K. Glass, M.G. Rosenfeld, D.E. Ayer, and R.N. Eisenman. 1998. SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. *Mol. Cell* **2**: 33-42.
- Lewis, J.D., R.R. Meehan, W.J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein, and A. Bird. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**: 905-914.
- Lin, R.J., L. Nagy, S. Inoue, W. Shao, W.H. Miller, Jr., and R.M. Evans. 1998. Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature* **391**: 811-814.
- Luo, R.X., A.A. Postigo, and D.C. Dean. 1998. Rb interacts with histone deacetylase to repress transcription. *Cell* **92**: 463-473.
- Lyons, G.E. 1996. Vertebrate heart development. *Curr. Opin. Genet. Dev.* **6**: 454-460.
- Magnaghi-Jaulin, L., R. Groisman, I. Naguibneva, P. Robin, S.



- Lorain, J.P., Le Villain, F., Troalen, D., Trouche, and A. Harel-Bellan. 1998. Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391**: 601–605.
- Meehan, R.R., J.D. Lewis, S. McKay, E.L. Kleiner, and A.P. Bird. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**: 499–507.
- Nan, X., R.R. Meehan, and A. Bird. 1993. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**: 4886–4892.
- Nan, X., H.H. Ng, C.A. Johnson, C.D. Laherty, B.M. Turner, R.N. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**: 386–389.
- Ng, H.H. and A. Bird. 1999. DNA methylation and chromatin modification. *Curr. Opin. Genet. Dev.* **9**: 158–163.
- Ng, H.H., Y. Zhang, B. Hendrich, C.A. Johnson, B.M. Turner, H. Erdjument-Bromage, P. Tempst, D. Reinberg, and A. Bird. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* (in press).
- Omichinski, J.G., G.M. Clore, O. Schaad, G. Felsenfeld, C. Trainor, E. Appella, S.J. Stahl, and A.M. Gronenborn. 1993. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* **261**: 438–446.
- Orkin, S.H. 1992. GATA-binding transcription factors in hematopoietic cells. *Blood* **80**: 575–581.
- Pazin, M.J. and J.T. Kadonaga. 1997. What's up and down with histone deacetylation and transcription? *Cell* **89**: 325–328.
- Razin, A. and A.D. Riggs. 1980. DNA methylation and gene function. *Science* **210**: 604–610.
- Rundlett, S.E., A.A. Carmen, N. Suka, B.M. Turner, and M. Grunstein. 1998. Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3. *Nature* **392**: 831–835.
- Scanlan, M.J., Y.T. Chen, B. Williamson, A.O. Gure, E. Stockert, J.D. Gordan, O. Tureci, U. Sahin, M. Pfreundschuh, and L.J. Old. 1998. Characterization of human colon cancer antigens recognized by autologous antibodies. *Int. J. Cancer* **76**: 652–658.
- Seelig, H.P., M. Renz, I.N. Targoff, Q. Ge, and M.B. Frank. 1996. Two forms of the major antigenic protein of the dermatomyositis-specific Mi-2 autoantigen. *Arthritis Rheum.* **39**: 1769–1771.
- Struhl, K. 1998. Histone acetylation and transcriptional regulatory mechanisms. *Genes & Dev* **12**: 599–606.
- Taunton, J., C.A. Hassig, and S.L. Schreiber. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**: 408–411.
- Toh, Y., S.D. Pencil, and G.L. Nicolson. 1994. A novel candidate metastasis-associated gene, *mta1*, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. *J. Biol. Chem.* **269**: 22958–22963.
- Toh, Y., E. Oki, S. Oda, E. Tokunaga, S. Ohno, Y. Maehara, G.L. Nicolson, and K. Sugimachi. 1997. Overexpression of the MTA1 gene in gastrointestinal carcinomas: correlation with invasion and metastasis. *Int. J. Cancer* **74**: 459–463.
- Toh, Y., H. Kuwano, M. Mori, G.L. Nicolson, and K. Sugimachi. 1999. Overexpression of metastasis-associated MTA1 mRNA in invasive oesophageal carcinomas. *Br. J. Cancer* **79**: 1723–1726.
- Tong, J.K., C.A. Hassig, G.R. Schnitzler, R.E. Kingston, and S.L. Schreiber. 1998. Chromatin deacetylation by an ATP-dependent nucleosome remodeling complex. *Nature* **395**: 917–921.
- Travers, A. 1999. An engine for nucleosome remodeling. *Cell* **96**: 311–314.
- Tsukiyama, T. and C. Wu. 1997. Chromatin remodeling and transcription. *Curr. Opin. Genet. Dev.* **7**: 182–191.
- Varga-Weisz, P.D. and P.B. Becker. 1998. Chromatin-remodeling factors: machines that regulate? *Curr. Opin. Cell. Biol.* **10**: 346–353.
- Verreault, A., P.D. Kaufman, R. Kobayashi, and B. Stillman. 1998. Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr. Biol.* **8**: 96–108.
- Wade, P.A., P.L. Jones, D. Vermaak, and A.P. Wolffe. 1998. A multiple subunit Mi-2 histone deacetylase from *Xenopus laevis* cofractionates with an associated Snf2 superfamily ATPase. *Curr. Biol.* **8**: 843–846.
- Wang, L., L. Liu, and S.L. Berger. 1998. Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo. *Genes & Dev* **12**: 640–653.
- Workman, J.L. and R.E. Kingston. 1998. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* **67**: 545–579.
- Xue, Y., J. Wong, G.T. Moreno, M.K. Young, J. Cote, and W. Wang. 1998. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol. Cell* **2**: 851–861.
- Yang, W.M., C. Inouye, Y. Zeng, D. Bearss, and E. Seto. 1996. Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc. Natl. Acad. Sci.* **93**: 12845–12850.
- Zhang, Y., R. Iratni, H. Erdjument-Bromage, P. Tempst, and D. Reinberg. 1997. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell* **89**: 357–364.
- Zhang, Y., G. LeRoy, H.P. Seelig, W.S. Lane, and D. Reinberg. 1998a. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95**: 279–289.
- Zhang, Y., Z.W. Sun, R. Iratni, H. Erdjument-Bromage, P. Tempst, M. Hampsey, and D. Reinberg. 1998b. SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell* **1**: 1021–1031.



# The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites

Brian Hendrich<sup>†</sup>, Ulrike Hardeland<sup>†‡</sup>, Huck-Hui Ng<sup>\*</sup>, Josef Jiricny<sup>‡</sup> & Adrian Bird<sup>\*</sup>

<sup>\*</sup> Institute of Cell and Molecular Biology, University of Edinburgh, The King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

<sup>‡</sup> Institute of Medical Radiobiology, August Forel-Strasse 7, CH-8008 Zürich, Switzerland

<sup>†</sup> These authors contributed equally to this work.

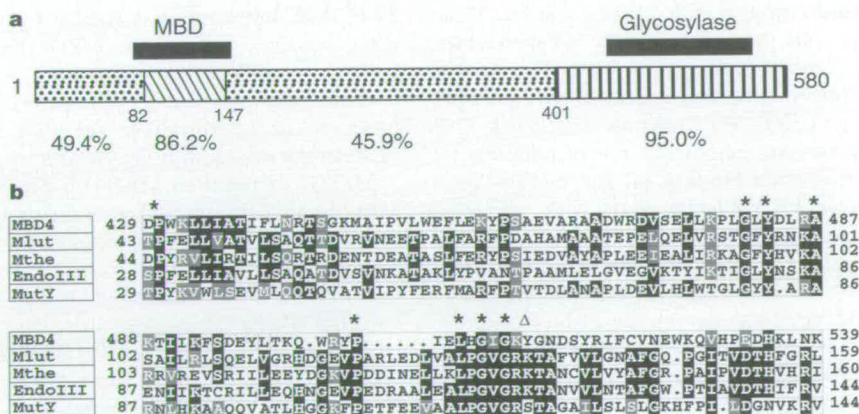
In addition to its well-documented effects on gene silencing, cytosine methylation is a prominent cause of mutations. In humans, the mutation rate from 5-methylcytosine (m<sup>5</sup>C) to thymine (T) is 10–50-fold higher<sup>1–4</sup> than other transitions and the methylated sequence CpG is consequently under-represented<sup>5</sup>. Over one-third of germline point mutations associated with human genetic disease<sup>6</sup> and many somatic mutations leading to cancer<sup>7,8</sup> involve loss of CpG. The primary cause of mutability appears to be hydrolytic deamination. Cytosine deamination produces mismatched uracil (U), which can be removed by uracil glycosylase<sup>9,10</sup>, whereas m<sup>5</sup>C deamination generates a G-T mismatch that cannot be processed by this enzyme. Correction of m<sup>5</sup>CpG-TpG mismatches may instead be initiated by the thymine DNA glycosylase, TDG<sup>11,12</sup>. Here we show that MBD4, an unrelated mammalian protein that contains a methyl-CpG binding domain<sup>13,14</sup>, can also efficiently remove thymine or uracil from a mismatched CpG site *in vitro*. Furthermore, the methyl-CpG binding domain of MBD4 binds preferentially to m<sup>5</sup>CpG-TpG mismatches—the primary product of deamination at methyl-CpG. The combined specificities of binding and catalysis indicate that this enzyme may function to minimize mutation at methyl-CpG.

MBD4 (ref. 14) was identified in a database search for proteins with a methyl-CpG-binding domain (MBD) resembling that of the transcriptional repressor MeCP2 (refs 13, 15). Hypothetical translation of the full-length human and mouse MBD4 complementary DNAs revealed an amino-terminal methyl-CpG-binding domain

and a carboxy-terminal region that was closely related to the glycosylase/endonuclease domains of bacterial repair proteins (Fig. 1a). Related proteins include the 8-oxoG:A mispair-specific adenine glycosylase MutY of *Escherichia coli*<sup>16</sup>, the G-T mismatch-specific thymine glycosylase Mig of *Methanobacterium thermoautotrophicum*<sup>17</sup>, the thymine glycol glycosylase EndoIII of *E. coli*<sup>18</sup> and the photodimer-specific UV-endonuclease of *Micrococcus luteus*<sup>19</sup> (Fig. 1b). Mouse and human MBD4 proteins are 86% and 95% identical in the methyl-CpG-binding domain and the C-terminal glycosylase-like domain, respectively, but are less conserved elsewhere (Fig. 1a).

On the basis of these relationships, we asked whether human MBD4 has DNA glycosylase or endonuclease activity. The presence of a methyl-CpG-binding domain led us to postulate that MBD4 might be involved in the processing of DNA damage associated with methylated CpG sites. Upon incubation of recombinant human MBD4 with a fluorescently labelled oligonucleotide duplex containing a single G-T mismatch in a CpG context, an abasic site was generated by the removal of the mispaired pyrimidine by glycosidic bond cleavage (Fig. 2a). In addition to the G-T mispair, G-U substrates were also cleaved (Fig. 2b), but the processing of other base/base mismatches, including C-A, C-C and G-G, was not observed (data not shown). That MBD4 is a mismatch specific T/U DNA glycosylase was confirmed by high-performance liquid chromatography (HPLC) analysis of the product mixture. Free T or U could be detected in reaction mixtures which contained MBD4 and G-T or G-U mismatches, respectively (data not shown). The sugar-phosphate backbone of DNA was not cleaved unless treated with hot alkali or the human AP-endonuclease HAP1 (ref. 20; and Fig. 2a). HAP1 cleaves oligonucleotide substrates at the 5' end of abasic sites. In the assay shown in Fig. 2a, HAP1 generates two fragments: a 5' fluorescein-labelled 23-mer terminated at its 3' end with a hydroxy group; and a 37-mer that has at its 5' end the baseless sugar phosphate from which the mispaired T has been removed by the action of MBD4. Neither glycosylase nor endonuclease activity could be detected on perfectly matched substrates containing symmetrically methylated, hemimethylated or non-methylated CpG sites (Fig. 2b). From these experiments, we conclude that MBD4 is a mismatch-specific T/U glycosylase that lacks endonuclease activity.

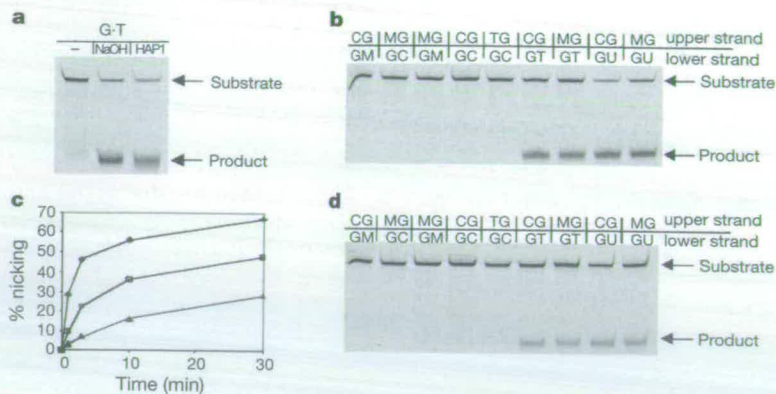
A CpG sequence context for the mismatch was preferred, but was not absolutely required, as G-T and G-U mismatches in other sequences were also processed, albeit at a reduced rate (Fig. 2c;



**Figure 1** MBD4 contains a conserved glycosylase-like domain and a methyl-CpG-binding domain. **a**, Diagram of MBD4 showing the N-terminal MBD and the C-terminal glycosylase domain. Percentage conservation between human and mouse proteins is shown below each shaded domain. **b**, Amino-acid sequence alignment of human MBD4 (amino acids 429–539) with the *E. coli* thymine glycol glycosylase EndoIII, the 8-oxoG:A-

specific adenine glycosylase MutY, the Mig G-T glycosylase of *M. thermoautotrophicum* (Mthe) and the *M. luteus* UV endonuclease (Mlut). Asterisks mark residues found in all the proteins; triangle indicates the residue that, when basic (K), is diagnostic of an endonuclease (glycosylase/lyase).





**Figure 2** MBD4 is a mismatch-specific T/U DNA glycosylase. **a**, MBD4 does not cleave the sugar-phosphate backbone of a mismatched DNA substrate (lane -). Cleavage occurs only upon the addition of hot alkali (lane NaOH) or human HAP1 endonuclease (lane HAP1). **b**, G-T and G-U mismatches are substrates for T glycosylase processing by full-length MBD4. Oligonucleotide substrate JJ, containing the appropriate mismatches or the corresponding perfectly matched, methylated or unmethylated controls, was incubated with MBD4 (see Methods). Apyrimidinic sites were then cleaved with hot alkali. Substrate

and product bands are indicated. **c**, Kinetics of MBD4 glycosylase action in CG and non-CG sequence contexts. G-U (diamonds) or G-T (squares) in a CG context is compared with G-T in a GG context (triangles). **d**, A C-terminal fragment of MBD4 that includes the glycosylase domain (amino acids 379–580; see Fig. 4a) can process G-U or G-T mismatches. **a**, **b** and **d** are fluorescence scans of denaturing 15% polyacrylamide gels. M denotes  $m^5C$ .

and data not shown). The methylation status of cytosine in CpG-TpG or CpG-UpG mismatches did not affect the glycosylase reaction in this assay (Fig. 2b). Substrate specificity *in vitro* is therefore similar to that of the other known mammalian G-T mismatch-specific DNA glycosylase, TDG<sup>11,12</sup>. As anticipated from the amino-acid sequence alignment (Fig. 1b), the glycosylase activity resides in the C-terminal region of MBD4, as a polypeptide corresponding to amino acids 379–580 could process G-T and G-U mismatches (Fig. 2d), whereas polypeptides in which the C terminus was deleted were inactive (data from Figs 2 and 3 are summarized in Fig. 4a; and data not shown). The substrate specificity and flanking sequence preference of full-length MBD4 and its C-terminal fragments were similar (compare Fig. 2b, d), indicating that the methyl-CpG-binding domain (see below) did not influence site preference in this assay.

the methyl-CpG binding domain (379–580; Fig. 3c) formed a complex with each of the mismatched probes (Fig. 3b), but a truncation lacking the C-terminal glycosylase-like domain did not form the complex (data not shown). As other DNA glycosylases bind tightly to the abasic site resulting from their catalytic action<sup>21</sup>, we suspected that the observed complexes also represented product binding by MBD4. Accordingly, we found that the glycosylase domain (379–580) and the intact form of MBD4 bound efficiently to abasic sites under the same conditions (Fig. 3c; and data not shown). Thus the N terminus of MBD4 binds to the MG-TG mismatch that is the substrate for base removal, whereas the C terminus, following processing, binds to the abasic product of the reaction.

Next, we investigated the DNA-binding specificity of MBD4. Mouse MBD4 binds to densely methylated DNA molecules *in vitro* and localizes to heavily methylated foci in mouse cell nuclei when overexpressed<sup>14</sup>. Bandshift assays with probes that contained non-base-paired regions, however, revealed additional specificity. Human MBD4 gave a strong complex with a probe that had three  $m^5CpG$ -TpG mismatches but bound relatively weakly when  $m^5CpG$ - $m^5CpG$  pairs were substituted at these sites (Fig. 3a, left panel). The probe with non-methylated, mismatched CpG-TpG sites also gave a weak complex. We conclude that both G-T mismatch and  $m^5C$  moieties are important for recognition by MBD4. Dependence of mismatch binding on the methyl-CpG-binding domain was shown by bandshift assays with truncated versions of MBD4. The N-terminal 165 residues, which include the DNA-binding domain, retained specificity for the methylated mismatch (Fig. 3a, right panel), whereas fragments that excluded all or part of the domain could not bind any of the probes (Fig. 4a; polypeptides 1–580 and 379–580). The results show that MBD4 preferentially binds to the mismatch that results from hydrolytic deamination of  $m^5C$  in mammalian genomic DNA.

MBD4 shows no apparent amino-acid sequence similarity to TDG, a protein that also excises mismatched T and U from G-T and G-U mismatches<sup>11,12</sup>. The two proteins probably evolved from different ancestors, as TDG is related to a bacterial G-U-processing enzyme, Mug<sup>22</sup>, whereas MBD4 seems to have arisen through the fusion of a glycosylase domain, such as that of bacterial MutY, to an N terminus that may target the enzyme to methylated mismatches. TDG does not contain a methyl-CpG-binding domain. Despite these sequence differences, MBD4 and TDG show very similar catalytic specificities *in vitro*, both enzymes preferring G-T and G-U mismatches within a CpG context<sup>23</sup>. In addition, both proteins bind to the apyrimidinic site that arises following mismatch processing (ref. 21; and Fig. 3b, c).

The complexes in Fig. 3a were formed under conditions in which processing of the mismatched site is undetectable (data not shown). After incubation for 20 min at 37 °C, MBD4–DNA complexes reproduced the specificity of the glycosylase reaction, which prefers mismatches in a CpG context, but is indifferent to methylation (Fig. 2b). The full-length protein and a C-terminal fragment lacking

MED1 (identical to MBD4) is reportedly a binding partner of MLH1, a protein implicated in mismatch repair<sup>24</sup>. It was proposed that MBD4/MED1 is an endonuclease that binds to hemi-methylated CpG sites and directs mismatch repair via a single strand nick towards the newly replicated strand, in a manner analogous to bacterial MutH<sup>25,26</sup>. Against this hypothesis, we find that MBD4 lacks endonuclease activity. The weak binding of MBD4 to hemi-methylated CpG sites compared with symmetrically methylated CpG<sup>24</sup> or  $m^5CpG$ -TpG mismatches also argues against the MutH homologue model.

We propose instead that the function of MBD4 is to counteract the mutability of  $m^5C$  by initiating conversion of  $m^5CpG$ -TpG mismatches back to  $m^5CpG$ -CpG. Among the sequences that we have tested, only  $m^5CpG$ -TpG shows both preferential binding by the N terminus of MBD4 and processing by its catalytic domain





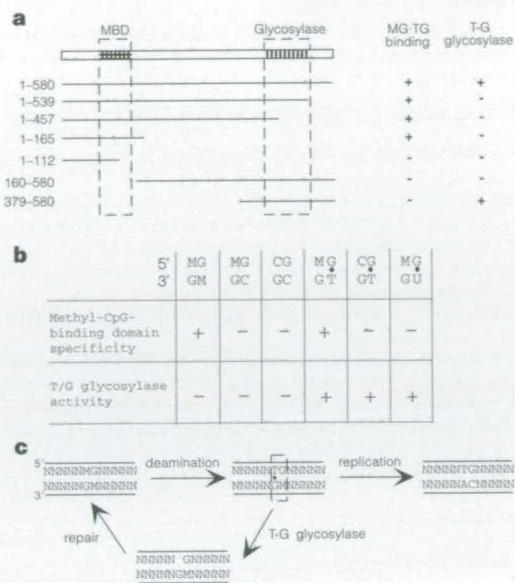
**Figure 3** DNA-binding specificities of MBD4 and its truncated derivatives to DNA containing methylated and non-methylated duplexes with or without G-U or G-T mismatches. **a**, Binding of full-length MBD4 (left) and the N-terminal 165 amino acids of MBD4 (right) to probe 'BH' (see Methods), which contains three potential CpG mismatch sites. Within each probe, all three sites were CpG-CpG duplex (CG-CG),  $m^5$ CpG- $m^5$ CpG duplex (MG-MG),  $m^5$ CpG-TpG mismatch (MG-TG) or CpG-TpG mismatch (CG-TG). Major complexes (C) are indicated. The slower migrating complexes in the MG-MG and MG-TG lanes of the 1-165 panel may be due to binding of more than one protein molecule to the probe. **b**, Full-length MBD4 incubated under conditions that favour the glycosylase reaction with fluorescently labelled probe JJ, in which a single CpG site was methylated or mismatched as shown (see Methods). Complexes were formed with G-U or T-G mismatches regardless of methylation status. For comparison, the complex between CG-TG mismatch and the T glycosylase TDG is shown (lane TDG). **c**, Under the same conditions, a C-terminal fragment (amino acids 379-580; see Fig. 4a) that can carry out the glycosylase reaction complexed with all mismatches in probe JJ, which was mismatched or methylated as shown. Asterisk denotes an abasic site.

(Fig. 4b). Although processing of mismatches *in vitro* was not accelerated by methylation, it is possible that *in vivo* the N-terminal domain targets MBD4 to these sites as they arise by hydrolytic deamination of  $m^5$ C (Fig. 4c). Repair of the resulting abasic site may involve interaction with MLH1 (ref. 24). Given the known importance of CpG to TpG transitions in genetic disease and cancer, it will be of interest to test the effect of MBD4 loss on mutation frequency and cancer predisposition. □

## Methods

### Protein expression.

MBD4 deletion constructs were made by PCR amplification from an MBD4 cDNA and cloning the resulting products into the pET6H vector. Recombinant proteins were expressed and purified as described<sup>14</sup>. N-terminal constructs were made by amplifying with the 5' primer 5'-CCTGCTCCATGGGCACGACTGGGCTG-3' and one of the



**Figure 4** The properties of MBD4 suggest a role in initiating repair of  $m^5$ CpG-TpG mismatches. **a**, Summary of DNA-binding and T glycosylase activity of truncated MBD4 polypeptides against  $m^5$ CpG-TpG (data is from Figs 2 and 3; and data not shown). The failure of the 160-580 protein to process mismatches is unexplained (aberrant protein folding?). **b**, Of the DNA sequences tested, only the methylated mismatch  $m^5$ CpG-TpG is both bound by the N terminus of MBD4 and processed by its glycosylase domain. **c**, A scheme for the role of T-G glycosylases in initiation of repair of deaminated methyl-CpG pairs. Solid circles in **b** and **c** indicate interstrand base mismatches.

following 3' primers: (580) 5'-GCGGGATCCTGAGCTTGAAGCTGCAG-3'; (539) 5'-GCGGGTACCAATCCATGAAGCTC-3'; (165) 5'-GCGGATCCGATGTC-AGGGCTGCCATG-3'; (112) 5'-GCGCAGCTGATAAAGTACACACTC-3'. The 1-457 construct was made by cloning the N-terminal *Nco*I fragment of the 1-580 construct into pET6H. C-terminal constructs were made by amplifying with the 580 primer (above) and one of the following 5' primers: (160) 5'-CGCCATGGCAGCCCTGACATC-3'; (379) 5'-GCGCCATGGGCTCTGAAATGGACAACAAC-3'.

### Enzymatic activity assay.

The enzymatic activity of the wild-type and mutant recombinant proteins was monitored using a 'nicking assay'. A 60-mer double-stranded oligonucleotide substrate containing different mismatches was prepared by annealing oligonucleotides to give the 'JJ' duplex: 5'-TAGACATTGCCCTCGAGGTACCATTGGATCCGATGTZRACCTCAAACCTAGACGAATCCG-3' 3'-ATCTGTAACGGGAGCTCCATGGTACCTAGGCTACAGYTGAGTTTGGATCTGCTTAAGGC-F-5' where F is fluorescein and R = G or A, Y = C,  $m^5$ C, T or U, and Z = C or  $m^5$ C (see underlined dinucleotide). A 0.5  $\mu$ M solution of the labelled upper strand and 1  $\mu$ M solution of the unlabelled upper strand were heated for 5 min at 95 °C in 10 mM Tris-HCl pH 8.0, 50 mM NaCl, and the solution was allowed to cool slowly to room temperature. In standard 'nicking reactions', 50 nM substrate and 50 nM protein were incubated for 1 h at 37 °C in a total volume of 20  $\mu$ l in 1 $\times$  nicking buffer (50 mM Tris-HCl pH 8.0, 1 mM DTT, 0.1 mg ml<sup>-1</sup> BSA, 1 mM EDTA). The reaction was stopped by adding NaOH to a final concentration of 90 mM and heating for 10 min at 99 °C. The NaOH treatment cleaves the AP sites resulting from the removal of the mispaired bases by the glycosylase. After adding 0.5  $\mu$ l 10 mg ml<sup>-1</sup> transfer RNA, 1/10 vol. 3 M NaOAc pH 5.2 and 3 vol. cold ethanol (-20 °C) to the reaction mixture, the DNA was precipitated at -20 °C for 1 h, pelleted by centrifugation and washed with 80% ethanol (-20 °C). The dried DNA was resuspended in 10  $\mu$ l formamide buffer (90% formamide, 1 $\times$  TBE) and heated for 5 min at 99 °C before loading on a 15% denaturing polyacrylamide gel. The fluorescein-labelled DNA fragments were detected using the blue fluorescence mode of the PhosphorImager (Storm 860, Applied Biosystems). The HAP1 endonuclease was expressed and purified from a clone donated by I. Hickson.

In some experiments, glycosylase activity was also assayed by incubating 1-10 ng recombinant protein with a radiolabelled 27-mer oligonucleotide,

5'-TCAGATTCCGCGCMGGCTGCGATAAGCT

3'-AGTCTAAGCGGGTTCGACGCTATTCTGA

(where M signifies  $m^5$ C), in 25 mM sodium phosphate pH 7.2, 10 mM EDTA, 50 mM NaCl and 1 mg ml<sup>-1</sup> BSA for 30 min at 37 °C. Piperidine was added to 1 M and the reaction was incubated at 90 °C for 30 min, after which the reactions were completely dried under vacuum. Samples were resuspended in 90% formamide gel loading buffer<sup>27</sup>, denatured at 95 °C for 5 min, and separated on a 15% denaturing polyacrylamide gel.



## Bandshift assays.

Standard bandshift reactions (Fig. 3a) utilized <sup>32</sup>P-labelled 'BH' probes that contained three CpG sites (underlined below) whose methylation/mismatch status was varied. The BH probes used in Fig. 3c were

5'TCAGATTCGCGCZGGCTGCGATAAGCTGZGCGGATCCZGGGAATTCAGCT3'

3'AGTCTAAGCGCGGYCGACGCTATTCGACGYGCCTAGGGYCCCTAAGTCGA5'

where Y = C, m<sup>5</sup>C, T or U, and Z = C or m<sup>5</sup>C (see underlined dinucleotides). The three CpG sites were identically modified/mismatched within each probe. Binding reactions were carried out at room temperature for 30 min in 20 mM HEPES pH 7.9, 25 mM NaCl, 10 mM β-mercaptoethanol, 1 mM EDTA, 4% glycerol, 1% digitonin and 50 ng sonicated *E. coli* DNA. The glycosylase reaction does not occur under these conditions (data not shown). Complexes were electrophoresed through 6% polyacrylamide gels in 0.5 × TBE at 4 °C.

Complexes formed under conditions favourable to the glycosylase reaction (Fig. 3b, c) utilized the fluorescent or <sup>32</sup>P-labelled JJ oligonucleotide (see above) as a probe. In standard gel-retardation reactions, 200 nM protein was incubated with 66 nM labelled oligonucleotide substrate and 333 nM unlabelled homoduplex oligonucleotide in 50 mM Tris-HCl pH 8.0, 1 mM DTT, 5% glycerol, 1 mM EDTA at 37 °C for 20 min. The samples were electrophoresed immediately through a 6% native 0.5 × TBE polyacrylamide gel for 45 min at 100 V. A probe with an abasic site was generated by treatment of oligonucleotide JJ containing a MG-UG mismatch with the enzyme uracil DNA glycosylase.

Received 25 June; accepted 21 July 1999.

- Duncan, B. K. & Miller, J. H. Mutagenic deamination of cytosine residues in DNA. *Nature* **287**, 560–561 (1980).
- Britten, R. J., Baron, W. F., Stout, D. B. & Davidson, E. H. Sources and evolution of human Alu repeated sequences. *Proc. Natl Acad. Sci. USA* **85**, 4770–4774 (1988).
- Sved, J. & Bird, A. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc. Natl Acad. Sci. USA* **87**, 4692–4696 (1990).
- Bulmer, M. Neighboring base effects on substitution rates in pseudogenes. *Mol. Biol. Evol.* **3**, 322–329 (1986).
- Bird, A. P. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8**, 1499–1594 (1980).
- Cooper, D. N. & Youssoufian, H. The CpG dinucleotide and human genetic disease. *Hum. Genet.* **78**, 151–15 (1988).
- Hollstein, M. *et al.* Database of p53 gene somatic mutations in human tumors and cell lines. *Nucleic Acids Res.* **22**, 3551–3555 (1994).
- Jones, P. A., Rideout, W. M., Shen, J.-C., Spruck, C. H. & Tsai, Y. C. Methylation, mutation and cancer. *BioEssays* **14**, 33–36 (1992).
- Lindahl, T. An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl Acad. Sci. USA* **71**, 3649–3653 (1974).
- Lindahl, T., Karran, P. & Wood, R. D. DNA excision repair pathways. *Curr. Opin. Genet. Dev.* **7**, 158–169 (1997).
- Wiebauer, K. & Jiricny, J. *In vitro* correction of G-T mispairs to G-C pairs in nuclear extracts from human cells. *Nature* **339**, 234–236 (1989).
- Neddermann, P. *et al.* Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* **271**, 12767–12774 (1996).
- Nan, X., Meehan, R. R. & Bird, A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**, 4886–4892 (1993).
- Hendrich, B. & Bird, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18**, 6538–6547 (1998).
- Nan, X., Campoy, J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471–481 (1997).
- Michaels, M. L., Pham, L., Nghiem, M., Cruz, C. & Miller, J. H. MutY, an adenine glycosylase active on G-A mispairs, has homology to endonuclease III. *Nucleic Acids Res.* **18**, 3841–3845 (1990).
- Horst, J. P. & Fritz, H. J. Counteracting the mutagenic effect of hydrolytic deamination of DNA 5-methylcytosine residues at high temperature: DNA mismatch N-glycosylase Mig.Mth of the

thermophilic archaeon *Methanobacterium thermoautotrophicum* THE. *EMBO J.* **15**, 5459–5469 (1996).

- Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H. & Cunningham, R. P. Purification and characterization of *Escherichia coli* endonuclease III from the cloned nth gene. *Biochemistry* **28**, 444–449 (1989).
- Shiota, S. & Nakayama, H. UV endonuclease of *Micrococcus luteus*, a cyclobutane pyrimidine dimer-DNA glycosylase/abasic lyase: cloning and characterization of the gene. *Proc. Natl Acad. Sci. USA* **94**, 593–598 (1997).
- Robson, C. N. & Hickson, I. D. Isolation of cDNA clones encoding a human apurinic/apyrimidinic endonuclease that corrects DNA repair and mutagenesis defects in *E. coli* xth (exonuclease III) mutants. *Nucleic Acids Res.* **19**, 5519–5523 (1991).
- Scharer, O. D., Nash, H. M., Jiricny, J., Laval, J. & Verdine, G. L. Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *J. Biol. Chem.* **273**, 8592–8597 (1998).
- Gallinari, P. & Jiricny, J. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature* **383**, 735–738 (1996).
- Sibghat-Ullah *et al.* Base analog and neighboring base effects on substrate specificity of recombinant human G:T mismatch-specific thymine DNA glycosylase. *Biochemistry* **35**, 12926–12932 (1996).
- Bellacosa, A. *et al.* MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proc. Natl Acad. Sci. USA* **96**, 3969–3974 (1999).
- Modrich, P. & Lahue, R. Mismatch repair in replication fidelity, genetic recombination and cancer biology. *Annu. Rev. Biochem.* **65**, 101–133 (1996).
- Jiricny, J. Replication errors: cha(lle)nging the genome. *EMBO J.* **17**, 6427–6436 (1998).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* 196–198 (Cold Spring Harbor, New York, 1982).

## Acknowledgements

We thank P. Hunziger for HPLC analysis; P. Schär, R. Wood and T. Lindahl for fruitful discussions; I. Hickson for the HAP1 endonuclease; and S. Tweedie for comments on the manuscript. This work was supported by grants from the Wellcome Trust (B.H. and A.B.), the Schweizerische Krebsliga (U.H. and J.J.) and by a Darwin Trust Scholarship to H.-H.N.

Correspondence and requests for materials should be addressed to A.B. (e-mail: A.Bird@ed.ac.uk).