4-17-2013 12:00 AM

Transcriptional Regulation by the Oncogenic ZNF217/CoREST Complex

Gobi Thillainadesan
The University of Western Ontario

Supervisor
Dr. Joseph Torchia
The University of Western Ontario

Graduate Program in Biochemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

TRANSCRIPTIONAL REGULATION BY THE ONCOGENIC ZNF217/COREST COMPLEX

(Thesis format: Integrated Article)

by

Gobi Thillainadesan

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

The ZNF217 transcription factor is an oncogene found within the 20q13 amplicon and is amplified and overexpressed in many cancers including breast and ovarian. Overexpression of ZNF217 leads to increased cell proliferation, survival, and causes resistance to TGFβ's anti-proliferative effects.

ZNF217 is a core constituent of a transcriptional complex that includes CoREST, HDAC1/2, LSD1, and the CtBP1/2. In this study, I have combined genome-wide biochemical approaches to identify genes directly regulated by ZNF217. I have identified the tumor suppressor and cell cycle inhibitor, $p15^{ink4b}$, as a direct target of the ZNF217 complex and demonstrated that ZNF217 represses the *p15*^{ink4b} gene by promoting a repressive chromatin environment and facilitating promoter DNA hypermethylation that involves a novel interaction with DNMT3A.

Furthermore, treatment of cells with TGFβ triggers DNA demethylation of the *p15*^{ink4b} promoter and the release of ZNF217/CoREST/DNMT3A complex. Subsequently, a novel activation complex is recruited that consists of SMAD2/3, CBP, and the DNA glycosylase TDG which precedes increases in $p15^{ink4b}$ protein expression. Knockdown of TDG, or its functional homolog MBD4, prevents TGF-β-dependent demethylation of the *p15*^{ink4b} promoter suggesting that the demethylation occurs through an active mechanism and is required for TGFβ dependant activation of gene expression. DNA immunoprecipitation experiments indicate that 5mC undergoes conversion to 5hmC in response to TGFβ treatment. AID/APOBEC2 deaminases are also required for the DNA demethylation by TGFβ supporting a mechanism whereby 5mC is hydroxylated to 5hmC

and then deaminated to 5hmU which is reverted to the unmethylated cytosine by the BER enzymes.

Overexpression of ZNF217 inhibits promoter demethylation and expression of the *p15*^{ink4b} gene in response to TGFβ by preventing recruitment of SMAD2/3/TDG complex. These findings suggest that the coregulator balance at promoters of genes is an important determinant of gene regulation and oncogenic amplifications such as ZNF217 can upset this balance causing deregulation of many genes. Taken together, these results establish the ZNF217 complex as a negative regulator of the *p15*^{ink4b} gene and may constitute an important link between amplification of ZNF217, increased cell proliferation and loss of TGFβ responsiveness in cancer.

Keywords: ZNF217, p15ink4b, Active DNA Demethylation, TGFβ, Cancer, TDG, MBD4, LSD1, CoREST, HDAC.

Co-Authorship Statement

Chapter 2 has been published as "Genome analysis identifies the p15ink4b tumor suppressor as a direct target of the ZNF217/CoREST complex" by Thillainadesan G, Isovic M, Loney E, Andrews J, Tini M, Torchia J in the journal Molecular and Cellular Biology. I participated in the design of the study, co-wrote the manuscript with Dr. Joe Torchia and conducted all of the experiments with the exception of the results for the following figures; Figure 2.1 was generated by Dr. Joe Torchia. Figure 2.2C was generated by Majdina Isovic. Figure 2.7B generated by Majdina Isovic. The analysis for the ChIP-DSL assay was conducted with the help of Joe Andrews who did the statistical analysis and generated lists of genes.

Chapter 3 has been published as "TGF-β-dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex" by Thillainadesan G, Chitilian JM, Isovic M, Ablack JN, Mymryk JS, Tini M, Torchia J in the journal Molecular Cell. I participated in the design of the study, co-wrote the manuscript with Dr. Joe Torchia and conducted all of the experiments with the exception of the results for the following figures; Figure 3.3D/3.4 was generated with the help of Jai Ablack who did the FACs experiments and analysis. Figures 3.5B, 3.6A, 3.6B, 3.10B, 3.10C, 3.13B, 3.13C and 3.15C were generated with the help of Jenny Chitilian who conducted the bisulphite conversion and sequencing of the samples. Figure 1E was conducted by Majdina Isovic. Figure 3.10D was generated with the help of Majdina Isovic who conducted the real-time analysis. Western blotting for Figures 3.13A and 3.5A was done by Majdina Isovic. Figure 3.11 was done by Dr. Joe Torchia.

Acknowledgments

The composure of this dissertation was made possible through the guidance and the assistance of many individuals who in their own way have contributed to the completion of this study.

I would like to begin by thanking the department of Biochemistry that consists of an intelligent collection of scientists and staff for their part(s). In particular, Barb Green for her much appreciated and unmatchable administrative service. Dr. David Litchfield and Dr. David Edgel for allowing me to serve as an assistant mentor, which I feel has enhanced my graduate experience, and also for their valuable words of wisdom. I would also like to thank Dr. Fred Dick and Dr. David Rodenhiser for serving as my graduate committee members and providing me with insightful suggestions that have helped to improve this study.

I would like to acknowledge the many scientists that collectively provide a cohesive scientific environment and the administrative staff at LRCP including Gail Howard and Denise Power.

My utmost gratitude to Dr. Joe Torchia for welcoming me into his lab and providing me with the many resources needed for this study. Also, his guidance and understanding through some critical stages of my graduate school is much appreciated and will not be forgotten. In addition, his immense scientific knowledge and aptitude has provided the inspiration I needed to complete this study.

My deepest thanks and appreciation to Dr. Joe Mymryk, who had kind concern and consideration regarding my academic progression. Also for his sincerity, kindness, moral support and encouragement. Futhermore, his capabilities as a scientist and his scientific outlook are inspirational. In addition, I would like to thank Dr. Joe Mymryk as

well as many of his graduate students for their collaborations including Greg Fonseca and Jai N. Ablack.

I would like to thank Dr. Susan Kohalmi who has provided me with a strong scientific foundation which has helped me throughout my study and I would also like to thank her for the encouragment and continued invaluable friendship.

I would like to thank the past and present members of Dr. J. Torchia's lab including Dr. Ester Loney and Bart Kolendowski. Niamh Coughlan, for being a good friend, providing support when needed and contributing to the pleasant experience. Majdina Isovic, for being my oldest companion in the lab and also as “the favorite” technician. In addition, her motivational support and the shared laughs/conversations will not be forgotten. Jenny Chitillian, who has ultimately been a fantastic friend and has filled a significant portion of my graduate years with laughter, insightful conversations and happiness. I would like to thank her for her indispensible support and kindness.

I would like to thank Oliva Palender for her emotional and academic support and her continued unconditional friendship. I would also like to thank Srikanth Talluri for his scientific insights and friendly support.

Finally and, most importantly, this entire experience and this dissertation would not have been at all possible without the unconditional, all inclusive support from Randeep Singh. I would like to thank her for her support in all aspects of my life and her friendship which is long lasting and significant. Her intelligent insights throughout this study and kindness are nonpareil. There is a lack of words to describe the tremendous help and happiness she has provided me over many years, including my years as a graduate student.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

5cC	5 carboxylcytosine
5fC	5 formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxmethyluracil
5hU	5 hydroxyuracil
5mC	5-methylcytosine
ADA2	Adaptor 2
AID	Activation-Induced Deaminase
AML1-ETO	Acute myeloid leukemia-1 and the eight-twenty-one corepressor
AP	Apurinic/Apyrimidinic
APC	Adenomatous Polyposis Coli
APE	AP Endonuclease
APOBEC	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
ATF3	Activating Transcription Factor 3
ATP	Adenine Triphosphate
AURKA	Aurora Kinase A
BAF	BRG1/BRM Associated Factors
Bdnf	brain-derived neurotophic factor
BER	Base Excision Repair
bHLH	basic helix-loop-helix
BMP	Bone Morphogenic Proteins
BRD	Bromodomains
Bre2	Brefeldin-A sensitivity protein 2
BRG1	Brahma-related gene 1
BRM	Brahma
cAMP	Cyclic Adenosine Monophosphate
CASTing	Cyclic Amplification and Selection of Targets
CBP/p300	CREB binding protein/p300
CDC25A	Cell Division Cycle 25 A
CDK	Cyclin Dependent Kinase
CHD	Chromo-ATPase/Helicase-DNA binding domain
ChIP	Chromatin Immunoprecipitation
ChIP-DSL	Chromatin Immunoprecipitation with Directed Selection and Ligation
ChIP-reChIP	Sequential ChIP analysis
cip1	CDK-interacting protein 1
COMPASS	Complex Proteins Associated with Set1
CoREST	Corepressor of REST
CREB	cAMP response element-binding protein

Ct	Cycle Threshold
CtBP	C-terminal binding protein
CTD	C-terminal domain
C-terminal	Carboxyl Terminal End
D2HG	D-2-hydroxyglutarate
dimetK4-H3	dimethyl K4 on Histone H3
DME	Demeter
DNA	Deoxyribonucleic Acid
DNMT	DNA methyltransferases
DNMT3L	DNMT3-like
dNTPs	Deoxyribonucleoside Triphosphates
DOT1L	Dot1-like Protein
E1A	Early Region 1A
EMT	Epithelial to Messenchymal Transition
ERBB3	V-ERBB2 avian erythroblastic leukemia viral oncogene homolog 3
ERK	Extracellular Signal Regulated Kinases
ERα	Estrogen Receptor Alpha
Ez	Enhancer of Zeste protein
Ezh1/Ezh2	Homologous of Enhancer of Zeste protein 1/2
FAD	Flavin Adenine Dinucleotide
FoxO	Forkhead Box O
GADD45α	Growth Arrest and DNA Damage Inducible Alpha
GBM	Glioblastoma Multiforme
GCN5p	General Control Nonrepressed
GRB2	Growth Factor Receptor Bound Protein 2
GSCs	Glioma Stem Cells
GTF	General Transcription Factor
HAT	Histone Acetyl Transferase
HDAC	Histone Deacetylase
HDACi	HDAC inhibitors
HMEC	Human Mammary Epithelial Cells
HP1	Heterochromatin protein 1
HRP	Horseradish Peroxidase Enzyme
H-SAM	S-adenosyl-methionine
HSP90	Heat Shock Protein 90
ID1	Inhibitor of DNA Binding
IDH	Isocitrate dehydrogenase
ink	Inhibitors of CDKs
INO80	Inositol requiring 80
Inr	Initiator Element

IP	Immunoprecipitated
ISW	Imitation SWI
JARID	Jumonji at-rich interactive domain
JHD2P	Jmjc domain containing histone demethylase-2
JHDM	JmjC domain containing demethylase
JmjC	Jumonji C
JMJD	JHDM2 jumonji domain containing protein
K9/K14-H3	Acetylated K9/14 on Histone H3
KAT	Lysine Acetyl Transferase
KDM	Lysine demethylase
kip1	Kinase interacting protein 1
LAP1/2	Liver-enriched transcriptional Activator Protein 1 and 2
LAZ3/BCL6	Lymphoma-associated zinc finger-3/B cell lymphoma 6
LIF	Leukemia Inhibitory Factor
LIP	Transcriptional Inhibitor Protein
LM-PCR	Ligation Mediated PCR
LSD1	Demethylase 1
LSD1	Lysine Demethylase I
MAP	Mitogen Activated Protein
MBD	Methyl-CpG Binding Domain
MBT	Malignant Brain Tumor
MeDIP	Methylated DNA Immunoprecipitation
MEFs	Mouse Embryonic Fibroblasts
mESCs	Mouse Embryonic Stem Cells
MH1 or 2	Mad homology 1 or 2
miRNA	Micro RNA
MIZ1	Msx Interacting Zinc Finger 1
MLH1	mut L homolog 1
MLL	Mixed Lineage Leukemia
MMR	mismatch repair
mRNA	messenger RNA
MSI	Microsatellite Instability
MUG	Mismatch-specific Uracil-DNA Glycosylase
MYBL2	V-Mybavian Myeloblastosis Viral Oncogene Homolog-Like 2
MYC	Myelocytomatosis Viral Oncogene Homolog
NABC1	Sodium-Coupled Borate Cotransporter 1
N-CoR	Nuclear Receptor Co-repressor
NTD	NH_2-Terminal Domain
N-terminal	Amino-Terminal End
NuRD	Nucleosome Remodelling and Histone Deacetylase

P/CAF	p300/CBP-Associated Factor
PcG	Polycomb Group
PGC	Primordial Germ Cells
PHD	Plant Homeodomain
PI	Propidium Iodide
PIC	Pre-Initiation Complex
PKC	Protein kinase C
PLDLS	Pro-X-Asp-Leu-Ser
Pol B	Polymerase B
pRb	phosphorylated Rb
PRC	Polycomb Repressor Complex
PRMT	Protein Arginine Methyl Transferase
PTM	Post Translational Modification
RAR	Retinoic Acid Receptor
Rb	Retinoblastoma Protein
RBPB5	Retinoblastoma binding protein 5
REST	Repressor Element Silencing Transcription Factor
RNA	Ribonucleic Acid
RNA-pol	RNA polymerase
RNA-pol-I	RNA polymerase I
RNA-pol-II	RNA polymerase II
ROS1	Repressor of Silencing 1
RPD3p	Reduced potassium dependency 3
rRNA	Ribosomal RNA
RRT	Arg-Arg-Thr
RSF	Remodeling and Spacing Factor
R-SMADs	Receptor Regulated SMADs
SAGA	Spt–Ada–Gcn5 acetyltransferase
SAM	S-adenyl methionine
SANT	SWI3, ADA2, N-CoR, TFIIB
SBE	SMAD binding element
SF2	Superfamily 2
SLIK	SAGA-like
SMAD	SMA- and MAD-related
Snf2h	Sucrose Nonfermenting 2 Homolog
SP1	Specificity Protein 1
SRC1	Steroid Receptor Coactivator 1
STAT3	Signal Transducers and Activators of Transcription 3
SUV39	Suppressor of Variegation 3-9
SWI/SNF	Switching Defective/sucrose Non-Fermenting

TAF	TBP Associated Factor
TBP	TATA Box Binding Protein
TDG	Thymine DNA Glycosylase
TET 1-3	Ten Eleven Translocation 1-3
TF	Transcription factor
TFIIB	Transcription Factor IIB
TFIID	Transcription Factor IID
TFIIE	Transcription Factor IIE
TFIIF	Transcription Factor IIF
TFIIH	Transcription Factor IIH
TGFβ	Transforming Growth Factor β
TGFβRI or II	TGFβ Receptor I or II
TIP60	TAT interacting protein
TRD	Transcriptional Repression Domain
TRIM33	Tripartite Motif Containing 33
trimet K27-H3	trimethyl K27 on Histone H3
tRNA	transfer RNAs
TSA	Trichostatin A
TSS	Transcription Start Site
UHRF	Ubiquitin-like containing PHD and RING finger domain
WCE	Whole Cell Extracts
WD40	Tryptophan-Aspartic acid dipeptide with 40 amino acid residues
WDR5	WD repeat domain 5
YY1	Yin Yang 1
ZNF	Zinc Finger
αKG	Alpha-ketogluterate

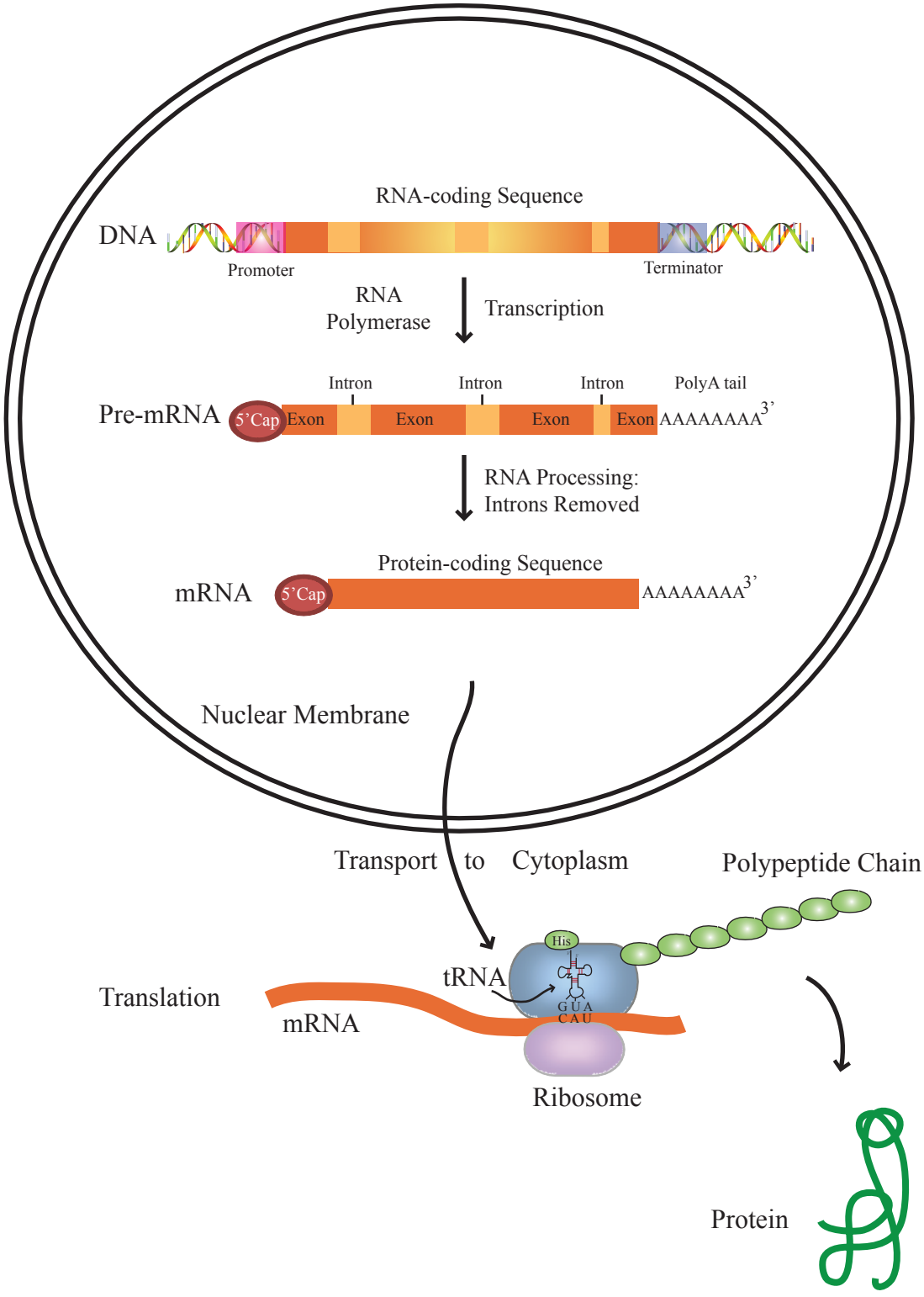
Chapter 1: General Introduction

1.1 Overview

Cell-specific phenotypes resulting from embryonic development are dictated through a succession of signals that trigger elaborate and accurate patterns of gene expression. Gene expression can be regulated in many ways that can be understood through consideration of the central dogma (Figure 1.1). The central dogma describes how the message encrypted in the deoxyribonucleic acid (DNA) is decoded and begins through the production of ribonucleic acid (RNA) molecules complementary to the DNA molecule, referred to as transcription, and then the production of linked amino acids (peptides) translated from the RNA to make complex proteins, referred to as translation (Bustamante et al., 2011; Franklin and Vondriska, 2011). In eukaryotes, RNA is synthesized in the nucleus through an enzyme, RNA polymerase (RNA-pol), followed by processing of the RNA such as addition of poly Adenosine tail and splicing. The RNA is then exported out to the cytoplasm to be translated. Translation of the RNA molecule to generate individual proteins is performed by a large protein complex, the ribosome (Rodrigo-Brenni and Hegde, 2012). The ribosome is composed of both protein and RNA subunits that can recruit structures called transfer RNAs (tRNA), structural RNA molecules that carry a specific amino acid. Within the ribosome the tRNA molecules recognize triple nucleotide base units of the messenger RNA (mRNA), referred to as codons. A catalytic unit of the ribosomal RNA catalyzes a peptide linkage to the carboxyl-terminal end of a growing peptide chain which will, following completion and proper folding, become a functional protein (Jackman and Alfonzo, 2012). The translated proteins then become the building blocks for structural integrity and enzymes that carry out many complex reactions in the cell.

Figure 1.1: The central dogma.

In the nucleus, DNA is transcribed by RNA polymerase to pre-RNA which is then processed by addition of a poly-A tail, 5'cap and the removal of introns to yield the mature RNA. The RNA is then transported out of the nucleus into the cytoplasm where it is translated by the ribosome into a functional protein.


RNA-coding Sequence
DNA
Promoter
Terminator
RNA
Polymerase
Transcription
Intron
Intron
Intron
PolyA tail
Pre-mRNA
5'Cap
Exon
Exon
Exon
Exon
AAAAAAAA3'
RNA Processing:
Introns Removed
Protein-coding Sequence
mRNA
5'Cap
AAAAAAAA3'
Nuclear Membrane
Transport to Cytoplasm
Polypeptide Chain
His
tRNA
G U A
C A U
Translation
mRNA
Ribosome
Protein

1.2 Eukaryotic Transcription

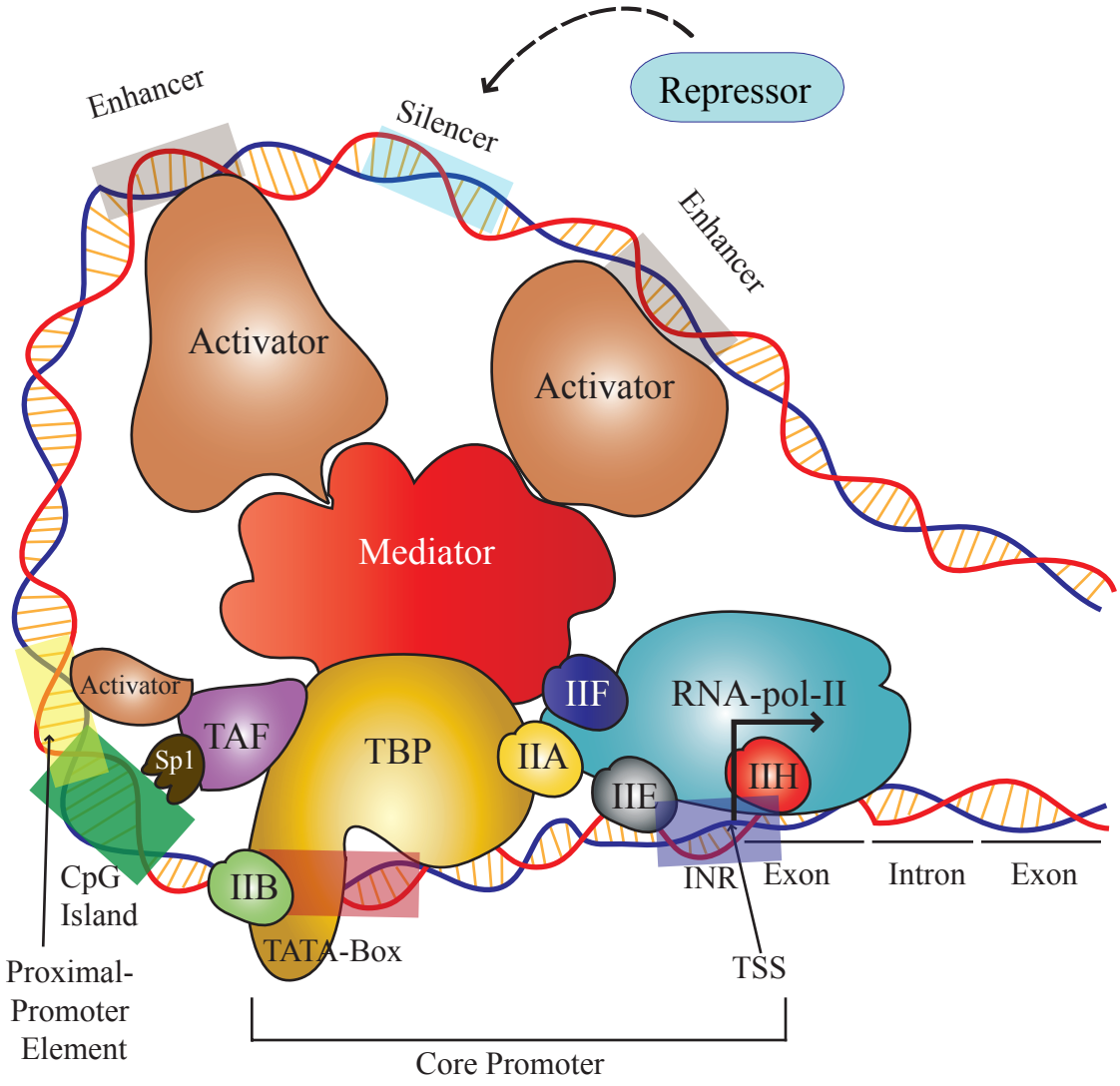
Transcription can be divided into three major phases, initiation, elongation and termination. The control of transcription initiation is considered a critical and rate limiting step in determining cell protein composition. In eukaryotes, there are three major classes of proteins that dictate transcription initiation: (1) Sequence specific transcription factors (TFs), many of which are expressed in a tissue specific fashion and can serve as activators or repressors (Garcia-Huerta et al., 2012; Sharon et al., 2012), (2) the basal transcriptional machinery consisting of core proteins that include the ubiquitously expressed RNA polymerase II (RNA-pol-II) and general transcription factors (GTFs) that facilitate the loading of the polymerase onto the DNA, and (3) transcriptional coregulators, that can either directly interact with the basal transcriptional machinery or that can modify the promoter region to indirectly facilitate or inhibit the loading of the transcriptional machinery. These three classes of proteins work in concert to determine the amount and types of crucial cellular contituents (proteins, RNA etc.) produced that ultimately dictate the phenotype of the cell (Figure 1.2).

1.2.1 *DNA*

Proteins within a cell are strategically synthesized through specific instructions contained in the genetic “blue print” of the cell, the DNA. DNA consists of two long chains of linear polymers, which make up the individual chromosomes, and each chain is composed of nucleotides. A nucleotide contains three components; a negatively charged phosphate group, a pentose sugar molecule and a nucleotide base. The pentose sugar molecule is linked to the phosphate group by a phosphoester bond to the 5’ carbon and

Figure 1.2: Transcription initiation.

This diagram depicts the regulation of gene expression through the concerted actions of many cis-regulatory elements which consists of core promoter elements, proximal promoter elements and elements located at further distances from the TSS such as enhancers, and silencers. Transcription is initiated by binding of GTFs to the core promoter in the following sequence; TBP binds to the TATA box followed by the binding of TFIIB, TFIIF and RNA-pol-II complex, TFIIE, and TFIIH. Two additional protein complexes can also regulate transcription initiation in eukaryotes; TAFs interact directly with activators that bind proximal promoter elements and Mediator complex is responsible for facilitating the interaction between long range activators, which bind to enhancer elements, and the GTFs. CpG islands are also located near the TSS that can recruit SP1 to facilitate transcription.


Repressor
Enhancer
Silencer
Enhancer
Activator
Activator
Mediator
Activator
TAF
Sp1
TBP
IIF
IIA
IIE
RNA-pol-II
IIH
IIB
TATA-Box
CpG
Island
Proximal-
Promoter
Element
INR
Exon
Intron
Exon
TSS
Core Promoter

the 1’ carbon is linked to the base through a glycosidic bond. Adjacent nucleotides are linked through a phosphodiester bond created between the phosphate group and the 3’ hydroxyl of the pentose sugar. Four different bases make up the DNA; the purines, adenine (A) and guanine (G), pairs with the pyrimidines, thymine (T) and cytosine(C), respectively. Unique combinations of the four bases can form “factors”, as coined by Gregor Mendel, which are units of information, referred to as genes (Reid and Ross, 2011). A gene is a hereditary unit resulting in identifiable traits and is the simplest unit that can produce a functional protein. Most eukaryotic genes contain introns, which consist of DNA sequence that is excluded from the resulting protein, and exons, DNA sequence containing the necessary information to make a complete protein (Figure 1.1). The DNA chains are arranged in an anti-parallel orientation, held together by complementary base pairing to form a double helix molecule of approximately 22-26 angstroms (Watson and Crick, 1953, 2003). The resulting double helix also contains two unequally sized grooves, the major groove and the minor groove. The mammalian genome consists of 23 homologous chromosomes pairs and 2 sex chromosomes and contains more than 25 000 protein coding genes.

1.2.2 *RNA and Amino Acids*

During transcription, DNA is read from 3’ to 5’ direction by the RNA-pol and generates the complementary RNA molecule in the 5’ to 3’ direction. The resulting RNA molecules are single strand linear polymers and, unlike DNA, contain a ribose instead of a deoxyribose sugar (Cheatham and Kollman, 1997). RNA molecules also contain the base uracil (U) in place of T. Many types of RNAs are transcribed from the eukaryotic genome. Ribosomal RNA molecules (rRNA) are part of the ribosomal protein-RNA

translational machinery and are transcribed by RNA polymerase I (RNA-pol-I) (Reeder, 1990). tRNAs are structural RNAs that carry the necessary amino acid to the ribosome during translation (Lambowitz and Perlman, 1990). Recently, many non-coding RNA (ncRNA) molecules involved in structural and functional roles within the cell have been identified. One such RNA molecule is the micro RNA (miRNA) which regulate transcription of genes through downregulation of mRNAs (Leonardo et al., 2012).

The mRNAs are transcribed from protein coding genes by RNA-pol-II (Figure 1.1). In prokaryotes, transcription and translation occur simultaneously and the resulting RNA is translated without further processing (Mitchell et al., 1997). In eukaryotes, the transcribed RNA molecule (pre-mRNA) is processed by an RNA-protein complex called the spliceosome which removes introns and splice together exons corresponding to an individual gene (Bonnal et al., 2012; Nielsen and Staley, 2012). In many cases, unique combinations of exons can be formed, called splice variants, that results in multiple protein isoforms being derived from a single gene (Herbert and Rich, 1999). Further processing of the pre-mRNA involves the addition of a 5' cap and a poly-adenylated 3' tail, which are important determinants of mRNA stability. The resulting mRNA is loaded onto the ribosome and is read in the 5' to 3' direction one codon at a time and each codon dictates the amino acid added to the growing polypeptide chain (Figure 1.1)(Nakamoto, 2009).

There are 20 amino acids that make up the individual eukaryotic proteins and amino acids can be represented by several codons (Nakamoto, 2009). The backbone of every amino acid has a simple configuration, consisting of an alpha-carbon joined to an amino group and a carboxyl group (NH2-CH-COOH). During translation, a peptide bond

is formed between the carboxyl group and the amino group of an adjacent amino acid. The resulting peptide contains an amino-terminal end (N-terminal) and a carboxyl terminal end (C-terminal). The side chains that are attached to the alpha-carbon defines the unique properties of the amino acids. Importantly, the side chains can participate in protein-protein interactions, can be covalently modified (phosphorylation, methylation, ubiquitination, acetylation and sumoylation), function as proton donors/acceptors during enzymatic reactions, and/or influence the secondary structure and function of the protein (Gray, 2003; Moreira et al., 2007).

1.2.3 *RNA Polymerase*

RNA-pol-II is the fundamental complex that catalyzes that the transcription of mRNA precursors as well as microRNAs. Eukaryotic RNA-pol-II contains 10-12 subunits, Rpb1 to Rpb12. The subunits are structurally and functionally conserved between yeast and humans and the core eukaryotic RNA-pol-II also shares similarity with prokaryotic RNA polymerase (Cramer et al., 2000). Initial binding of RNA-pol-II to DNA is performed by the Rpb1 subunit that is also capable of identifying a transcriptional start site (TSS). Rpb1 and Rpb2, together, form the active site of the enzyme, where ribonucleoside triphosphates (NTPs) are polymerized. Structural studies indicate that Rpb5 on one side and Rpb1 and Rpb9 on the other act as a mechanical jaw to help position DNA for transcription and Rpb1, Rpb2 and Rpb6 act as a sliding clamp allowing DNA to pass through the RNA-pol while also stabilizing the DNA-RNA-polymerase complex (Cramer et al., 2008). The remaining subunits contribute to various transcriptional processes such as TSS identification (or binding) transcriptional elongation and interaction with additional activator proteins. For example, Rpb4 and

Rpb7 appear to function as stress responsive dissociable subunits and display tissue specificity (Khazak et al., 1998). DNA assumes a non-linear structure when bound by RNA-pol-II forming a sharp bend at the site of interaction, which is speculated to provide torsional stress to the double helix hence lowering the energy required to break the hydrogen bonds between the complementary strands (Coulombe and Burton, 1999). The growing nascent RNA strand is fed through two pores that are formed by the Rpb1 and Rpb2 subunits.

1.2.4 *Eukaryotic RNA-pol-II Regulator DNA Regions*

Regulation of gene expression occurs through the concerted actions of short DNA sequences known as *cis*-regulatory elements that are found within promoters, enhancers and introns. *Cis*-regulatory elements bind TFs that function as activators, repressors or insulators (Figure 1.2) (Gaszner and Felsenfeld, 2006; Petrykowska et al., 2008).

Core promoters consist of DNA sequence immediately upstream (5') of TSSs and contain diverse sequence elements that direct transcription initiation (Butler and Kadonaga, 2002). Variability in core promoter sequence is crucial for transcription regulation since it provides a gene specific function for activators and repressors. Eukaryotic core promoters of protein coding genes can contain three different promoter elements that dictate RNA-pol-II loading; the TATA box, the initiator element (Inr), and the GC-rich elements (Figure 1.2). The TATA box was identified as early as 1979 when promoter regions of *Drosophila melanogaster*, mammals and viral RNA-pol-II protein coding sequences were compared and analyzed (Breathnach and Chambon, 1981). The TATA box has a consensus sequence that is TATAXAAX (X represents A, T, C or G)

and is found 25-35 bp upstream of the TSS (Carninci et al., 2006; Hahn et al., 1989). Recent comparative sequence analysis of many *D. melanogaster* core promoters has found that as many as 32-43% of promoters contain a TATA box (Kutach and Kadonaga, 2000; Ohler et al., 2002). In humans, a similar analysis of 1031 core promoters revealed that 32% contain the TATA box (Suzuki et al., 2001). Depending on the tissue type, mutations within the TATA box can result in a drastic reduction in transcription (Duan et al., 2002). TATA box acts to position RNA-pol-II at the TSS and because the TATA box consensus sequence is not symmetrical, it also confers directionality of transcription (Singh et al., 1997).

Some eukaryotic genes contain an Inr element which is found at the TSS and consists of a degenerative consensus sequence Y-Y-A (+1)-N-T/A-Y-Y-Y-3', where Y represents any pyrimidine and N represents (A,G,C or T) (Figure 1.2) (Javahery et al., 1994). Pyrimidines at position -2, +4, and +5 are essential for Inr activity and increasing the number of pyrimidines around this position correlates with increased transcription (Kutach and Kadonaga, 2000). The Inr sequence functions independently of other regulatory elements and transcription from Inr-containing promoters is comparable to TATA box regulated promoters (Corden et al., 1980; Grosschedl and Birnstiel, 1980). Some eukaryotic promoters contain both the TATA box and the Inr element where both elements function synergistically (Malecova et al., 2007). An Inr element is found in approximately 69 % of core *D. melanogaster* promoters but the prevalence of these elements in mammals is yet to be determined (Kutach and Kadonaga, 2000).

Promoters of some eukaryotic genes lack both the TATA box and the Inr element and do not contain a well-defined TSS. These genes may have multiple TSSs that span

between 20-200 bp, giving rise to mRNAs that vary in size. These genes often contain CG dinucleotide repeats, also known as CpG islands. CG dinucleotide repeats can serve as specificity protein 1 (SP1) binding sites and studies have demonstrated that transcription often begins 40-80 bp downstream of SP1 binding. This suggests that SP1 may direct the transcriptional machinery to these sites (Figure 1.2) (Kwon et al., 1999). The human genome is estimated to have approximately 29,000 CpG islands and half of the protein coding genes contain at least one CpG island in the vicinity of the TSS (Antequera and Bird, 1993; Suzuki et al., 2001). Interestingly, it has been found that when an Inr element is inserted downstream of an SP1 binding site, higher levels of transcription can be attained (Butler and Kadonaga, 2002; Smale, 2001). Combinatorial effects of the proximal promoter elements can thus confer multiple transcriptional states of genes, adding complexity and control of gene transcription (Petrykowska et al., 2008).

The promoter may also contain other *cis* proximal promoter elements that can activate or repress transcription. In addition, regulatory elements may also be found many kilobases away, within enhancer regions (Figure 1.2) (Lenhard et al., 2012). Enhancers are DNA sequences consisting of one or more binding sites for a variety of TFs that can regulate transcription independent of their position within the genome, distance or direction (Banerji et al., 1981). In fact, enhancers have been known to activate transcription of genes located on a separate chromosome (Geyer et al., 1990; Lomvardas et al., 2006). It has been shown that TFs bound at enhancers can directly interact with the core transcriptional machinery by a mechanism involving looping out of the intervening sequences (Schoenfelder et al., 2010).

1.2.5 *Transcription Factors*

TFs are proteins that bind to *cis*-regulatory elements and regulate gene expression by facilitating the loading of RNA-pol-II and activating transcription, or hindering its binding and repressing transcription (Petrykowska et al., 2008). The activity of a single TF bound to DNA may regulate transcription alone or may be affected by the presence of other TFs that bind to adjacent sites. For example, Yin Yang 1 (YY1) is a transcriptional repressor or activator depending on the presence of specific factors. This suggests that the context in which the TFs are found is an important determinant of gene regulation (McKenna and O'Malley, 2002; Shi et al., 1997). The human genome encodes approximately 2000 distinct TFs representing approximately 5 % of the human genome (Tupler et al., 2001). TFs can be classified into many different families according to their DNA binding domains or the class of genes they regulate. For example, the homeodomain protein family contains domains that can specifically bind and regulate homeotic genes, which determine the development of body plan (Blyth, 2012; Verzi et al., 2012). Helix-loop-helix (HLH) proteins and leucine-zipper proteins function as dimers and contain the hydrophobic amino acid leucine at every seventh position that is required for the dimerization (Baxevanis and Vinson, 1993; Oshaben et al., 2012; Zhao et al., 2012).

The largest class of TFs are the zinc finger (ZNF) family of proteins (Laity et al., 2001). ZNF proteins use zinc ions to contour a relatively short, 23-50 amino acids, polypeptide to form compact projected structures. Two major classes of ZNF proteins are present in eukaryotes, referred to as the C4 ZNF and C2H2 ZNF proteins (Iuchi, 2001; Seo et al., 2012). The C4 ZNF proteins contain four conserved cysteines within a short

stretch of polypeptide that can coordinate a zinc ion and bend the polypeptide to form the fingers. The C4 ZNF proteins contain only two such fingers and generally bind to DNA as a dimer. This class of TF encompasses the nuclear hormone receptors that are regulated by ligand binding (Knegtel et al., 1995). The C2H2 class of ZNF proteins contains two conserved cysteines adjacent to two conserved histidines that bind and coordinate a zinc ion to form the individual fingers. This class of proteins can contain several adjacent fingers which wrap around the double helix DNA and insert into the major groove (Iuchi, 2001; Quinlan et al., 2007).

Several mechanisms exist for the regulation of TFs that provides an additional level of transcriptional control. Many TFs, such as the steroid hormone receptors are sequestered in the cytoplasm thus preventing access to a DNA binding site. Additionally, TFs are subject to post-translational modifications such as phosphorylation, acetylation and methylation, that may also result in alterations to subcellular localization, activity and ability to recognize binding sites on DNA (Zhang and Reinberg, 2001).

1.2.6 *Mechanism of Eukaryotic Transcription Initiation*

Transcription initiation is considered to be an integrative process and is the result of the combinatorial binding of TFs to the core promoter, proximal promoter elements and enhancers (Figure 1.2). *In vitro* experiments have demonstrated that RNA-pol alone is sufficient for transcription but the eukaryotic RNA-pol-II requires additional proteins, such as TFs, to be targeted to the promoter region of genes *in vivo*. These proteins can form a complex with RNA-pol-II and is referred to as the pre-initiation complex (PIC) (Walter, 1967; Sekine et al., 2012). In addition, the PIC also assists in strand separation,

dictates start site selection and provide directionality to transcription. DNA foot-printing and electro-mobility shift assays have been used to characterize the proteins of the PIC that bind to DNA, referred to as the general transcription factors (GTFs) (Orphanides et al., 1996). Transcription is initiated by binding of the TATA box binding protein (TBP) through its C-terminal region. TBP is highly conserved between yeast and humans and its structure resembles a saddle with both non-identical halves of the protein exhibiting a dyad symmetry (Kays and Schepartz, 2000). Interestingly, much like a saddle, TBP interacts with the minor groove of the DNA double helix with each half of the protein on both sides of the DNA and this interaction causes a significant bend in the DNA. Transcription factor IIB (TFIIB) then binds to TBP, through its C-terminal end, and makes direct contact with the TATA box (Sainsbury et al., 2012). TFIIB demonstrates directionality with its N-terminal domain extending towards the TSS. This is followed by recruitment of a preformed complex consisting of transcription factor IIF (TFIIF) and RNA-pol-II and the catalytic site of RNA-pol-II is positioned over the TSS. Before the DNA double strand can be separated to expose the template strand, two other proteins must bind, transcription factor IIE (TFIIE) and transcription factor IIH (TFIIH) (Kim et al., 2000; Kim et al., 1997). TFIIH has helicase activity and uses the energy derived from adenine triphosphate (ATP) breakdown separate the double stranded DNA providing an open template for RNA-pol-II to initiate transcription (Kim et al., 2000).

Two additional protein complexes can also regulate transcription initiation in eukaryotes; TBP associated factors (TAFs) and Mediators (Figure 1.2). TAFs serve as scaffolds for the assembly of the PIC and play a critical role in promoter recognition (D'Alessio et al., 2009). A multi-subunit complex consisting of TBP and approximately

12 TAFs, generally referred to as transcription factor IID (TFIID), is often found as part of the eukaryotic PIC. Eukaryotic genes can be characterized as TAF dependent or independent and genes requiring TAFs achieve specificity through differential tissue specific expression of TAF proteins (D'Alessio et al., 2009; Kuras et al., 2000; Raha et al., 2005). TAFs can also interact directly with activators that bind proximal promoter elements thus increasing gene expression (Shen and Green, 1997).

The Mediator is a large multi-protein complex containing as many as 20 subunits. Seven Mediator complexes have been identified in humans (Kim et al., 1994; Myers and Kornberg, 2000). Mediator complex is responsible for facilitating the interaction between long range activators, which bind to enhancer elements, and the GTFs. This increases the formation of the PIC at TSS and enhances gene transcription (Ptashne and Gann, 1997; Scafe et al., 1990). The Mediator has also been shown to increase basal transcription by approximately 10-fold and promote phosphorylation of the C-terminal domain (CTD) of RNA-pol-II. The CTD consists of up to 52 repeats with a consensus Tyr-Ser-Pro-Thr-Ser-Pro-Ser and is an essential component of the RNA pol II that plays an important role in transcription initiation, RNA capping and splicing (Kim et al., 1994; Lee et al., 1997).

Transition from the initiation phase to the elongation phase involves the dissociation of RNA-pol-II from the majority of the GTFs associated with the PIC. However, certain factors that facilitate the elongation by RNA-pol-II, remain associated. In addition, the transition into the elongation phase requires several covalent modifications to the CTD of RNA-pol-II such as phosphorylation (Wade and Struhl, 2008). RNA-pol-II then transcribes towards the 3’-end of the gene, effectively elongating the mRNA as it proceeds. Termination of transcription can occur in several ways but one

accepted mechanism involves the addition of the poly A tail to the nascent mRNA. This allows for the recruitment of RNA binding proteins that can promote the dissociation of the RNA-pol-II/RNA complex from the DNA(Kuehner et al., 2011).

In summary, the PIC is required for the loading of RNA-pol-II and in many cases, TAFs are required as part of the PIC for promoter recognition and, along with Mediators, facilitate the activation by long range enhancers. Although GTFs, TAFs and Mediators are essential for transcription initiation, the regulation of their binding and activity is controlled by epigenetic mechanisms.

1.3 Epigenetic Regulation of Gene Expression: Overview

Epigenetic regulation is defined as "the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states" of genes (Bird, 2007). One component of epigenetic regulation occurs as a result of the packaging of DNA into chromatin, a nucleoprotein complex that allows for gene regulation through nucleosome remodelling and post-translational covalent modifications (Gu and Roeder, 1997). Another component of epigenetic control involves DNA methylation at cytosine within CpG dinucleotides and CNG trinucleotides resulting in the formation of 5-methylcytosine (5mC). DNA methylation is generally considered to be repressive to transcription and is important for the maintenance of chromosome stability. (Bird, 1986; Deaton and Bird, 2011). These two forms of epigenetic control are major determinants of the transcriptional state of the cell and are largely defined by a repertoire of proteins with enzymatic activity that are targeted to specific sites throughout the genome.

1.3.1 *Chromatin: General Features*

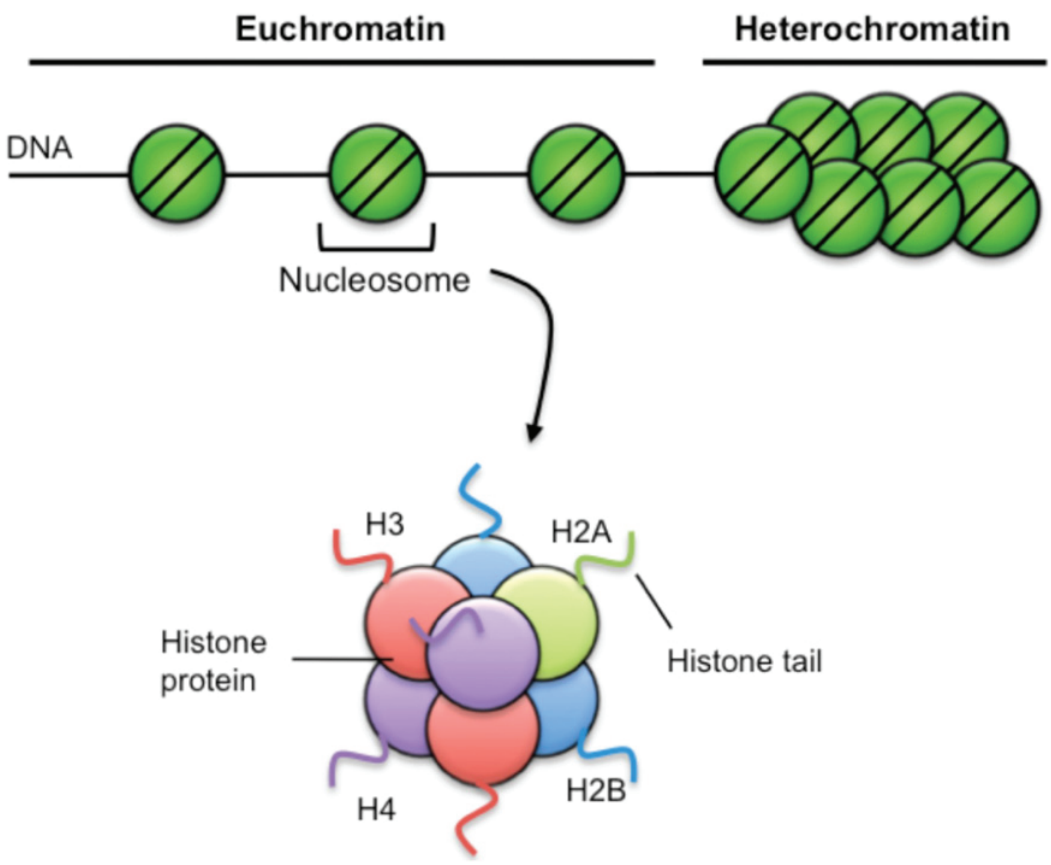
The nucleosome is the fundamental repeating unit of chromatin and consists of an octamer of core histones containing two copies each of histone H3, H4, H2A and H2B, interacting with each other through globular domains (Figure 1.3) (Eickbush and Moudrianakis, 1978; Luger et al., 1997). Approximately 147 bps of DNA is wrapped around each nucleosome in 1.7 helical turns and the nucleosomes are connected by 10-16 bps of linker DNA (Kornberg and Thomas, 1974).

Each histone has an amino-terminal tail consisting of 20 to 37 amino acids, and although the molecular structure of the tails have not been determined, it is believed that the histone tails protrude out from the surface of the nucleosome (Szerlong and Hansen, 2011). A fifth histone, the histone H1, interacts with nucleosomes and serves as a linker between successive nucleosomes to stabilize and promote a higher order structure of chromatin. Histone H1 has been implicated in promoting transcriptional repression by blocking nucleosome repositioning, obstructing activator binding and promoting heterochromatization (Caterino and Hayes, 2011; Shen and Gorovsky, 1996).

Chromatin is packaged into two general states that are highly dynamic and contribute to the maintenance of cellular identity, heterochromatin and euchromatin. Heterochromatin consists of highly dense regions and contains genes which are not being transcribed. Euchromatin consists of less dense regions of the genome which contains many genes that are either being actively transcribed or are poised for transcription (Figure 1.3). Euchromatin represents the first order of compaction and when examined by electron microscopy, appears as a 11 nm fiber composed of DNA wrapped around many

Figure 1.3: The chromatin structure.

The top represents chromatin packaging which is dynamic and generates two states, heterochromatin which is a highly dense region, and euchromatin, a less dense region of the genome. Euchromatin consists of approximately 147 bp of DNA wrapped around each nucleosome and the nucleosomes are connected by 10-16 bp of linker DNA. The bottom depicts a nucleosome which consists of an octamer of core histones containing two copies each of histone H3, H4, H2A and H2B. Each histone has an amino-terminal tail that protrudes out from the surface of the nucleosome. Adapted from (Russ et al., 2012).


Euchromatin
Heterochromatin
DNA
Nucleosome
H3
H2A
Histone protein
Histone tail
H4
H2B

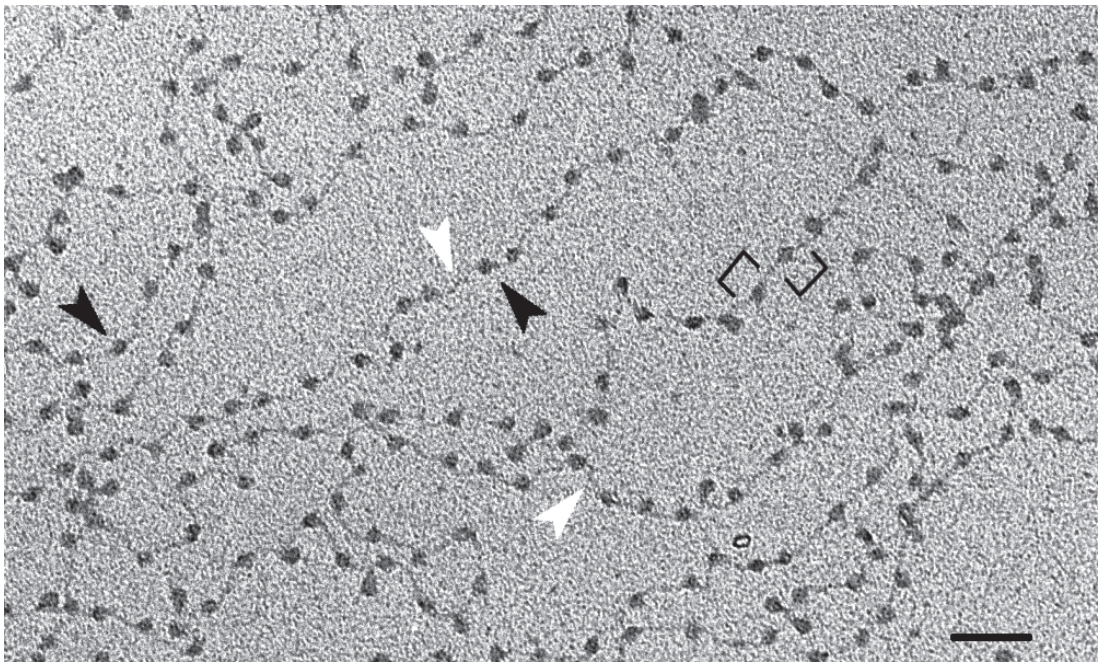
repeating nucleosomes which resemble “beads on a string” (Figure 1.4) (Szerlong and Hansen, 2011). The second level of compaction consists of the 30 nm fiber which is found in heterochromatin (Gu and Roeder, 1997). Though the crystal structure of the 11nm fiber has been solved the exact structure of the 30nm fiber remains unclear (Kornberg and Thomas, 1974). Chromatin is also able to undergo several levels of higher order compaction and it has been hypothesized that long-range interactions between nucleosomal arrays is thought to play a role in additional compaction although the dynamics remains unclear (Li and Reinberg, 2011). Chromatin itself plays an extremely important role in the regulation of gene transcription because the winding of DNA around nucleosomes, as well as the higher order packaging of chromatin renders the DNA largely inaccessible. Consequently, transcription requires that genes become accessible to both sequence specific TFs and the transcriptional machinery (Narlikar et al., 2002; Urnov and Wolffe, 2001).

In order to accomplish this, eukaryotic cells utilize coregulator proteins which alter chromatin structure and make the DNA more accessible. Coregulator proteins can be divided into two categories. The first are the ATP-dependent nucleosome remodeling proteins, which often exist as part of multi-protein complexes and utilize the energy derived from the hydrolysis of ATP to reposition nucleosomes. The second category consists of coregulators that modify the chemistry of histone proteins by catalyzing post-translational modifications to the amino terminal tails of the histones (Turner, 1993). Both types of coregulator proteins can induce transcriptional changes important for gene regulation.

Figure 1.4: Transmission electron image of euchromatin.

An image generated using Transmission Electron Microscope of Chicken erythrocyte nuclei that were allowed to decondense. Decondensed chromatin appears as a 10nm fiber composed of DNA wrapped around many repeating nucleosomes which resemble a "beads on a string" model. The black arrows point to nucleosomes and the white arrows point to the linker region between nucleosomes.

Figure adapted from The Cell (An image library; http://www.cellimagelibrary.org/images /709).



The ATP-dependent chromatin remodeling complexes can restructure chromatin to make DNA more accessible for transcription, or can generate highly compact chromatin structures that repress transcription. All of the chromatin remodeling complexes contain a subunit that belongs to the superfamily 2 (SF2) of helicases and contains an intrinsic ATPase domain which is required for separating DNA from nucleosomes (Narlikar et al., 2002). The ATPase subunit associates with additional subunits and each of the individual subunits possess specific activities that contribute to the function of the complex, such as regulating the ATPase activity, mediating interactions with chromatin, and/or protein-protein interactions (Clapier and Cairns, 2009).

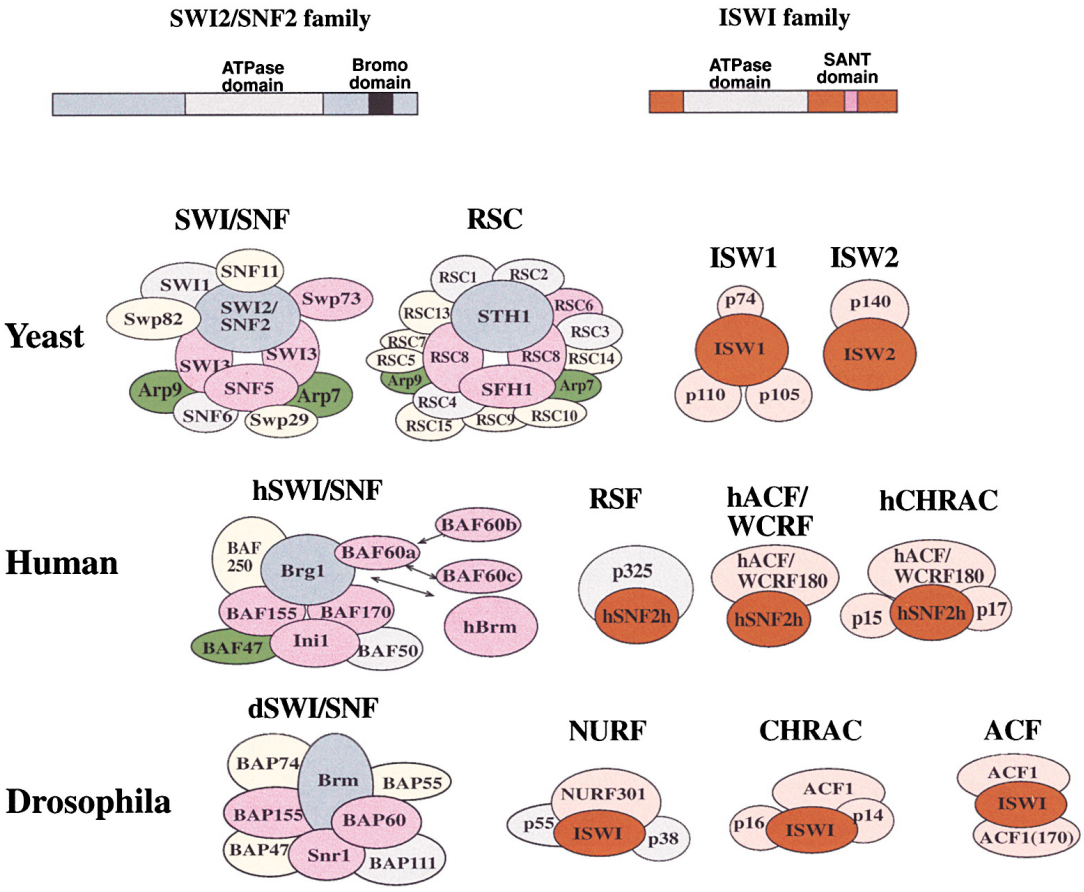
The two major classes of ATP-dependent chromatin remodelers are the, switching defective/sucrose non-fermenting (SWI/SNF) and imitation SWI (ISWI), based on the homology of their ATPase domain to yeast and *D. melanogaster* proteins, respectively (Figure 1.5) (Gangaraju and Bartholomew, 2007). Complexes within the SWI/SNF class of remodelers consist of eight to fifteen subunits, while the ISWI class of remodelers are smaller and consist of two to five subunits (Lall, 2007). More recently, two additional families of chromatin remodelers have been identified, the chromo-ATPase/Helicase-DNA binding domain (CHD) and inositol requiring 80 (INO80) (Clapier and Cairns, 2009).

Despite similar biochemical properties it has been shown that the chromatin remodeling complexes can generate distinct rearranged products of the nucleosomes. For example, the SWI/SNF class of remodelers can generate stable protruding loops of DNA around nucleosomes and this is postulated as a mechanism to expose DNA at promoters

Figure 1.5: ATP-dependent chromatin remodelers.

The diagram represents two major classes of ATP-dependent chromatin remodelers; SWI/SNF and ISWI. Complexes within the SWI/SNF class of remodelers consist of eight to fifteen subunits, while the ISWI class of remodelers is smaller and consist of two to five subunits. The ATPase subunits also contain domains such as the bromo-domain and SANT domain which are chromatin binding domains. Adapted from (Narlikar et al., 2002).

SWI/SNF: SWItch/Sucrose NonFermentable. RSC: Remodels structure of chromatin. ISW1/2: Imitation switch 1/2. RSF: Remodeling and spacing factor. ACF: Assembly and chromatin remodeling factor. CHRAC: Chromatin accessibility complex. NURF: Nucleosome remodeling factor.


SWI2/SNF2 family
ATPase domain
Bromo domain
ISWI family
ATPase domain
SANT domain
SWI/SNF
RSC
ISW1
ISW2
Yeast
SNF11
SWI1
Swp73
Swp82
SWI2/ SNF2
SWI3
SWI3
Arp9
SNF5
Arp7
SNF6
Swp29
RSC1
RSC2
RSC13
STH1
RSC6
RSC3
RSC7
RSC5
RSC8
RSC8
RSC14
Arp9
SFH1
Arp7
RSC4
RSC15
RSC9
RSC10
p74
ISW1
p110
p105
p140
ISW2
hSWI/SNF
RSF
hACF/ WCRF
hCHRAC
Human
BAF 250
Brg1
BAF60a
BAF60b
BAF60c
hBrm
BAF155
BAF170
BAF47
Ini1
BAF50
p325
hSNF2h
hACF/ WCRF180
hSNF2h
hACF/ WCRF180
p15
hSNF2h
p17
dSWI/SNF
NURF
CHRAC
ACF
Drosophila
BAP74
Brm
BAP55
BAP155
BAP60
BAP47
Snr1
BAP111
NURF301
p55
ISWI
p38
ACF1
p16
ISWI
p14
ACF1
ISWI
ACF1(170)

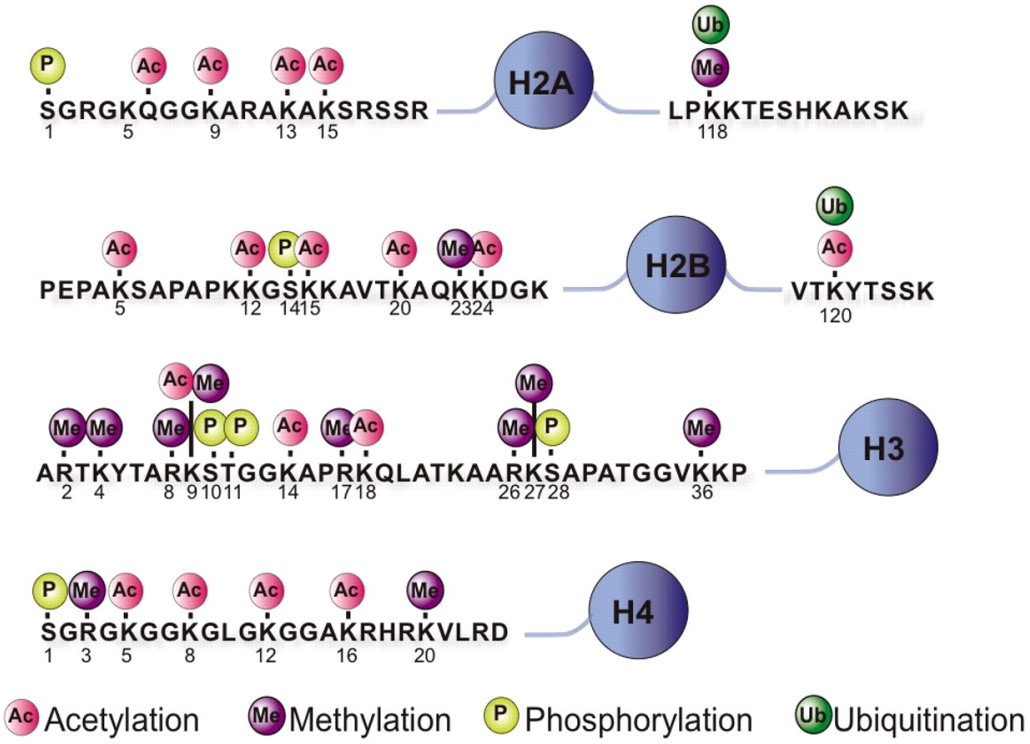
within chromatin dense regions to facilitate transcription. SWI/SNF remodelers can also transfer nucleosomes to different regions of the DNA as well as remove H2A/H2B dimers or entire nucleosomes (Fan et al., 2004; Narlikar et al., 2002). Both of these mechanisms have been implicated in generating nucleosome free regions of DNA. Alternatively, the ISWI class of remodelers can generate evenly spaced nucleosomes over long stretches of DNA which generally promotes higher order chromatin structures and represses transcription (Cairns, 2005; Cryderman et al., 1999; Gangaraju and Bartholomew, 2007).

Human SWI/SNF complexes include the Brahma-related gene 1 (BRG1)/Brahma (BRM) associated factors (BAF) chromatin remodeler that contain the BRG1/BRM proteins as the ATPase subunit along with eight other subunits (Figure 1.5) (Steinberg et al., 2012; Wu, 2012). An example of the human ISWI is the remodeling and spacing factor (RSF) which consists of sucrose non-fermenting 2 homolog (Snf2h) as the ATPase domain and one other subunit (Figure 1.5) (LeRoy et al., 1998). Chromatin remodelers can be targeted to gene promoters by interacting with TFs, such as the the glucocorticoid receptor which binds to the human BAF chromatin remodeling complex to promote transcriptional activation in response to glucocorticoids (Hsiao et al., 2003) Alternatively, they can be targeted to the genome by interacting with the chromatin through intrinsic chromatin binding domains contained within specific subunits (Clapier and Cairns, 2009).

The second class of coregulators catalyzes the post translational modifications of the histones. Due to their accessibility outside of the core nucleosome the amino acids on the histone tails are highly amenable to covalent modifications (Figure 1.6). These

Figure 1.6: Chromatin tail modifications.

The diagram represents histone modifications that occur on histones H3, H4, H2A and H2B. Each histone has an amino-terminal tail consisting of 20 to 37 amino acids. The modifications include methylation, phosphorylation, acetylation, and ubiquitination. Adapted from (Kato et al., 2010).


P
Ac
Ac
Ac
Ac
SGRGKQGGKARAKAKSRSSR
1
5
9
13
15
H2A
Ub
Me
LPKKTESHKAKSK
118
Ac
Ac
P
Ac
Ac
Me
Ac
PEPAKSAPAPKKGSKKAVTKAQKKDGK
5
12
1415
20
2324
H2B
Ub
Ac
VTKYTSSK
120
Ac
Me
Me
Me
Me
P
P
Ac
Me
Ac
Me
Me
P
Me
ARTKYTARKSTGGKAPRKQLATKAARKSAPATGGVKKP
2
4
8
9
10
11
14
17
18
26
27
28
36
H3
P
Me
Ac
Ac
Ac
Ac
Me
SGRGKGGKGLGKGGAKRHRKVLRD
1
3
5
8
12
16
20
H4
Ac Acetylation
Me Methylation
P Phosphorylation
Ub Ubiquitination

modifications include the acetylation of lysines, methylation of lysines and arginines, as well as phosphorylation of serines (Strahl and Allis, 2000). The consequences of histone covalent modifications varies depending on which amino acid is modified, as well as the type of modification but tend to result in two possible consequences. First, the modification can alter the charge on the specific amino acid which can influence interactions between histones and DNA leading to changes in chromatin structure. Second, histone modifications can affect the recruitment or repulsion of non-histone effector molecules resulting in downstream transcriptional effects. With respect to transcription, histone modifications are grouped into two categories: those resulting in transcriptional activation and those resulting in transcriptional repression (Strahl and Allis, 2000). It has been postulated that the type, number and combination of specific histone modifications, generate a type of code known as "the histone code" that dictate a downstream transcriptional event and consequently a biological outcome (Strahl and Allis, 2000; Turner, 2000, 2007).

Chromatin modifications are mediated by enzymes that add or remove specific moieties at the histone tails. To gain access to their histone substrates these enzymes must be recruited to the DNA and one mechanism involves recruitment by specific TFs (Kang et al., 2004). This mode of recruitment provides a degree of selectivity of gene expression because only the genes that are targeted by these enzymes will be regulated by the corresponding histone modification. Enzymes that catalyze the addition or removal of certain chromatin marks on histone tails with a transcriptional outcome are referred to as co-activators or co-repressors. (Turner, 1993).

1.3.2 *Lysine Acetylation*

The presence of acetylated histones in locations of active transcription was noted as early as 1978 when chromatin was found to be less resistant to DNAase I digestion after treatment with sodium butyrate, a potent histone deacetylase (HDAC) inhibitor (Simpson, 1978). Since then, it has been shown that lysine acetylation plays a fundamental role in many cellular processes such as DNA replication where histones are acetylated before the assembly into the newly synthesized strand, and importantly in transcriptional regulation (Sawan and Herceg, 2010; Strahl and Allis, 2000). Lysine acetylation has been detected *in vitro* and *in vivo.* In addition, *in vitro* reconstitution of DNA in the context of chromatin have confirmed the requirement for lysine acetylation in transcriptional activation (Loyola et al., 2001). Lysine acetylation has been characterized most frequently at residue 9, 14, 18, 23 and 56 on histone H3 (H3K9, 14, 18, 23, 56) as well as residues 5, 8, 12, and 16 of histone H4. Acetylation at these residues has been associated with transcriptional activation of genes. Acetylation may facilitate transcription by neutralizing the positive charge on the lysine residues, which normally contributes to interactions between the chromatin and DNA (Hansen, 2002). This mechanism is controversial since acetylation can only neutralize approximately 30% of the charge which may not be sufficient to release the DNA from histones (Struhl, 1998). Alternatively, acetylated lysines can serve as binding sites for the bromodomain (BRD), a consensus motif found in many transcriptional regulatory proteins (Struhl, 1998). The BRD is a well conserved left handed helix bundle consisting of four helices, more commonly referred to as the BRD fold. Within the human genome, 46 BRD-

containing proteins have been identified and many of these proteins contain multiple BRD domains (Schultz et al., 2000).

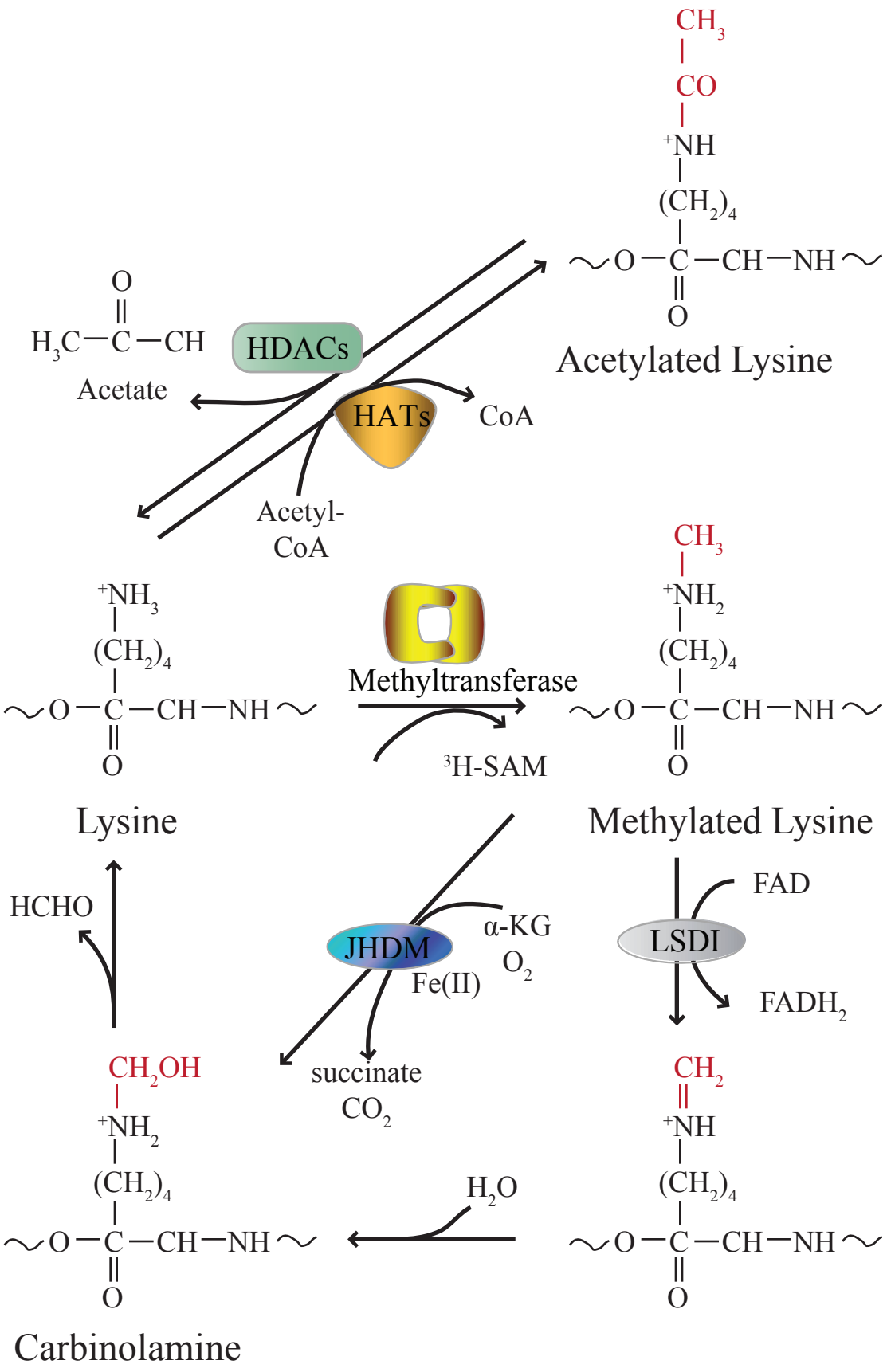
BRD-containing proteins can facilitate transcription in many ways. For example, BRD1 and BRD2 proteins can enhance transcription on nucleosomal templates *in vitro* that is dependent on histone acetylation. The ability of BRD1 and BRD2 to facilitate transcription was attributed to their capability to bind acetylated histones and their intrinsic chaperone activity which was found to temporarily remove histones to allow the progression by RNA-pol-II (LeRoy et al., 2008).

Acetylation is catalyzed by the histone acetyl transferases (HATs), also known as lysine acetyltransferases (KATs). HATs transfer an acetyl group from the co-factor acetyl-CoA to the terminal amino group of the side chain (ε-amino group) on lysines (Figure 1.7) (Wiegand and Brutlag, 1981). There are two types of HATs in eukaryotic cells, type A and B. Type B-HATs are exclusively cytoplasmic and are mainly involved in the acetylation of histone H4 prior to incorporation into the newly synthesized DNA(Brownell and Allis, 1996). Type A-HATs are directly involved in transcriptional regulation and chromatin assembly (Ruiz-Garcia et al., 1997).

The first HAT was cloned from *Tetrahymena thermophila* and is homologue to the yeast transcriptional adaptor/coactivator protein, general control nonrepressed (GCN5p). Using free nucleosomes as substrates, it was subsequently shown that recombinant GCN5p possesses intrinsic HAT activity (Brownell et al., 1996). These experiments established an important link between acetylation and transcriptional activation by adaptor proteins (Georgakopoulos and Thireos, 1992). Additionally, it was

Figure 1.7: Acetylation/methylation of lysines.

HATs transfer an acetyl group from the co-factor acetyl-CoA to the terminal amino group of the side chain on lysines. HDACs can catalyze the removal of acetyl groups from proteins by an ion exchange mechanism utilizing the co-factors zinc or NAD^+. Methylation is also added to the amino terminal of the lysine side chain and this is catalysed by the MLL proteins using S-adenosyl-methionine (H-SAM) as the methyl donor. Methyl-lysines can be demethylated by LSD1 which utilize FAD to oxidize methyl-lysine to an imine intermediate which is then hydrolyzed to yield unmethylated lysine and formaldehyde. The JHDM enzymes can also demethylate lysines and this class of enzymes uses iron and α-ketoglutarate as co-substrates and demethylation results in the generation of formaldehyde and succinate. Molecules indicated in red represent modifications to lysine.


HDACs
Acetate
HATs
CoA
Acetyl-
CoA
Acetylated Lysine
Methyltransferase
^{3}H-SAM
Lysine
Methylated Lysine
HCHO
JHDM
α-KG
O_2
Fe(II)
succinate
CO_2
LSDI
FAD
$FADH_2$
H_2O
Carbinolamine

found that GCN5p contains a BRD, implicating this domain in targeting of HATs to genes (Sanchez and Zhou, 2009). Many additional HATs have been identified in eukaryotes, such as cAMP response element-binding protein and its close homologue, binding protein 300 (CBP/p300), TAT interacting protein (TIP60), and p300/CBP-Associated factor (P/CAF) (Bararia et al., 2008; Sartorelli et al., 1999; Verreault et al., 1998). Many of these HATs exhibit specificity towards their substrates through recognition of consensus motifs (Peserico and Simone, 2011). For example, GCN5p contains a consensus motif for H3K14 consisting of GKXP, where X represents any amino acid (Rojas et al., 1999).

The CBP/p300 proteins, are ubiquitous HATs that are utilized by many classes of TFs. For example, p300 was found to be an important component of the estrogen receptor (ER) nuclear receptor activation complex that is targeted to specific genes following β-estradiol treatment (Hanstein et al., 1996). CBP/p300 is also required for the transcriptional response induced by the transforming growth factor β (TGFβ) (Feng et al., 1998; Janknecht et al., 1998; Nishihara et al., 1998; Pouponnot et al., 1998). Homozygous deletion of p300 or CBP in mice is embryonic lethal and exhibit numerous developmental defects such as defects in neural tube closure, reduced cell proliferation and disrupted cardiac development (Oike et al., 1999; Yao et al., 1998).

Acetyl groups are removed from lysine residues by histone deacetylase (HDACs) class of enzymes. HDACs catalyze the removal of acetyl groups by an ion exchange mechanism utilizing the co-factors zinc or NAD^+ (Figure 1.7) (Dokmanovic et al., 2007). In mammals, there are four major classes of HDAC enzymes based on sequence

homology. Class I (HDAC 1-3, 8), class IIa (HDAC 4, 5, 7, 9), class IIb (6, 10), class III (SIRT 1-7) and class IV (HDAC11). Class I, II and IV are zinc dependent while Class III is NAD^+ dependent (Bali et al., 2005). Deacetylation by HDACs plays a critical role in regulating histone as well as non-histone proteins. For example, HDAC6 is primarily cytosolic and deacetylates heat shock protein 90 (HSP90) which is required for HSP90 to bind ATP and associate with its client proteins (Bali et al., 2005).

In mammals, the class I enzymes (HDAC1 and HDAC2) share the most sequence identity (65%) with the yeast reduced potassium dependency 3 (RPD3p) being the major deacetylase that regulates histone acetylation in yeast (de Ruijter et al., 2003; Kurdistani et al., 2002). HDAC1 and 2 are ubiquitously expressed and share 87% amino acid similarity in mammals (Yang and Seto, 2008). Deletion of HDAC1 in mice causes defects in cell proliferation and growth, resulting in embryonic lethality, whereas deletion of HDAC2 causes prenatal lethality resulting from cardiac defects (Montgomery et al., 2007). With respect to deacetylation of histones, HDAC1 and HDAC2 are considered to be functionally redundant. However they also possess different non-histone substrates which could account for the differences in the phenotypes of the individual knockouts (Jurkin et al., 2011). For example, during differentiation of neuronal precursors knockdown of both HDACs, but not individual knockdowns, result in severe brain abnormalities (Sun et al., 2011). HDAC 1 and 2 are found in the nucleus as components of multiprotein repressor complexes such as Sin3A, nucleosome remodelling and histone deacetylase complex (NuRD) and the corepressor of REST (CoREST) complex (Peserico and Simone, 2011).

Alterations in HATs and HDACs have been identified in numerous cancers. For example, disruption of CBP or p300 resulting from missense mutations, or mutations leading to p300 truncations are associated with colorectal, gastric and other epithelial cancers (Giles et al., 1998; Goodman and Smolik, 2000). Overexpression of the transcriptional repressor lymphoma-associated zinc finger-3/B-cell lymphoma 6 (LAZ3/BCL6) in non-Hodgkin's lymphom causes aberrant recruitment of HDACs resulting in deregulated transcriptional repression (Dhordain et al., 1998). The fusion gene acute myeloid leukemia-1 and the eight-twenty-one corepressor (AML1-ETO) recruit HDACs and acts as a potent repressor of transcription that associates with acute myloid leukaemia (Wang et al., 1998).

Interestingly, many structurally diverse HDAC inhibitors have been developed that inhibit cell proliferation *in vitro* and *in vivo* by inducing cell cycle arrest, differentiation and/or apoptosis (Marks et al., 2001; Munster et al., 2001). HDAC inhibitors have provided a valuable tool to study the physiological role of HDACs and are being tested therapeutically to treat specific cancers such as those mediated by the AML1-ETO fusion proteins (Wang et al., 1998).

1.3.3 *Histone Lysine Methylation*

Histone methylation was first identified and characterized as early as 1968 (Murray, 1964). Much like acetylation, methylation of histone tails is an important regulator of transcription (Figure 1.6). However, unlike acetylation, methylation does not alter the charge of the histone tail, suggesting that this modification does not play a direct role in DNA histone interactions (Hansen, 2002; Martin and Zhang, 2005). Furthermore,

methylation of histone tails can be mono-, di-, or tri-methylated and, depending on the residue that is methylated, can be associated with either transcriptional activation or repression (Greer and Shi, 2012). Notably, methylation at H3K4 is generally linked to transcriptional activation and the degree of methylation at H3K4 confers varying states of transcription. In vertebrates H3K4me3 and H3K4me2 are often colocalized and associate with the 5' regions of actively transcribed genes (Bernstein et al., 2002; Sims and Reinberg, 2006; Wysocka et al., 2005b). H3K4me2 is also found in the coding regions of active genes and H3K4me1 appears to be mostly localized to the 3' region of transcribed genes (Heintzman et al., 2007; Santos-Rosa et al., 2002). In contrast, di and tri-methylation of H3K9 and H3K27 are considered hallmarks of transcriptionally repressed genes (Young et al., 2010; Zee et al., 2010). Interestingly, in some contexts the presence of H3K4me3 can co-occupy the same promoter regions of genes as the repressive methylation mark, H3K27me3 (Heintzman et al., 2007; Orford et al., 2008; Santos-Rosa et al., 2002). This phenomenon is known as bivalency and is generally found at lineage specific genes within pluripotent cells, such as embryonic stem cells. The presence of transcriptionally repressive marks are thought to maintain lineage commitment genes in the off position while the activation marks hold the promoters in a "poised" state, prepared for rapid activation upon appropriate differentiation signals (Bernstein et al., 2006).

Methylation can also occur at arginine residues (Figure 1.6). Arginines can be mono- or di-methylated, and di-methylation of arginines can occur in two conformations, symmetric or asymmetric which can result in distinct phenotypic outcomes (Chen et al.,

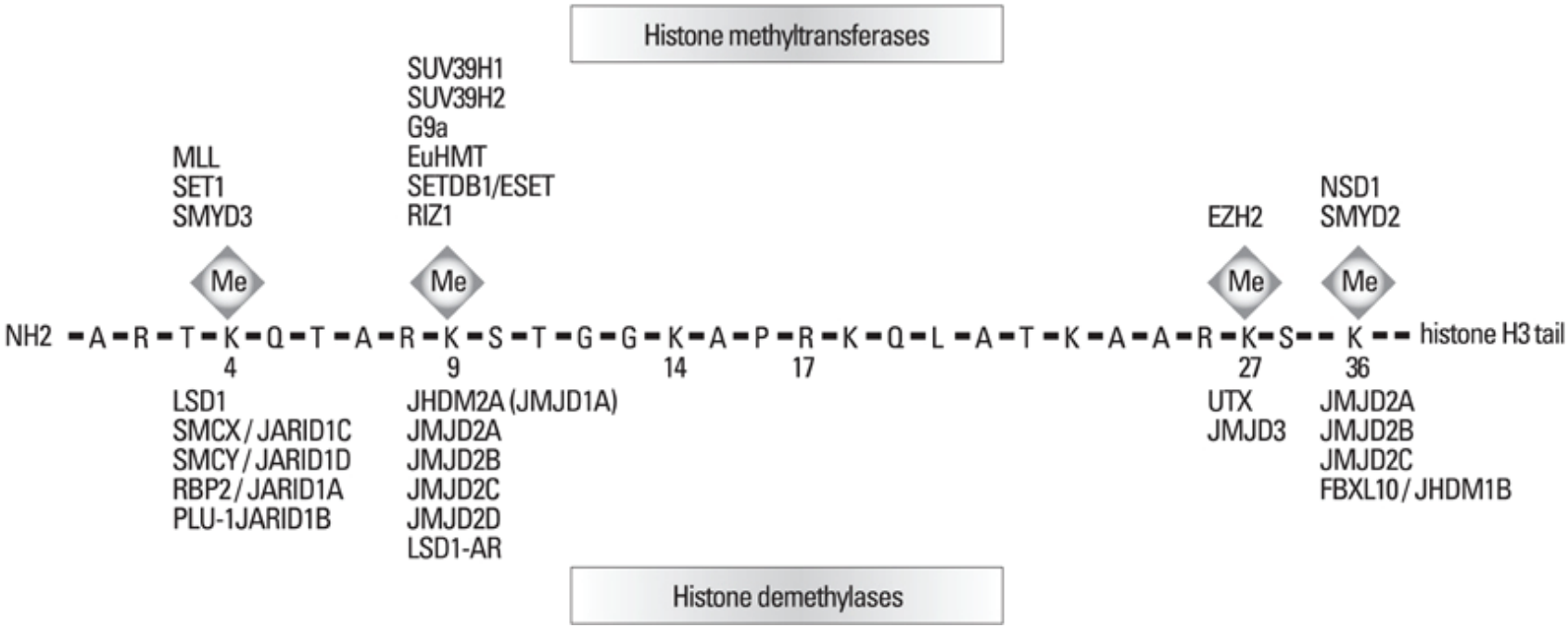
2011). Methylation at histidines is rare in eukaryotes and only the mono-methylation at these residues has been reported (Paik et al., 2007).

Histone methylation is mediated by three classes of enzymes. The DOT1-like proteins (DOT1L) methylate H3K79, and have been implicated in telomeric silencing and transcriptional regulation (Feng et al., 2002). The protein arginine methyl transferase (PRMT) family methylate an array of histone tail arginine residues and are associated with many cellular outcomes and transcriptional states (Wolf, 2009).

The SET domain-containing proteins are a large family of lysine methyltransferases that regulate transcription of many genes (Figure 1.7 and Figure 1.8) (Rea et al., 2000). The SET domain is the catalytically active component of the enzyme and amino acids surrounding the SET domain dictates the amino acid motif recognized by the enzyme (Bannister and Kouzarides, 2011). One of the first SET domain methyl transferases identified is the suppressor of variegation 3-9 (SUV39) which specifically methylates H3K9. Many additional proteins within this family have been identified. This includes G9a/GLP1 which can di- and tri-methylate H3K9. A unique feature of this enzyme is that it contains an ankyrin domain which can specifically bind to the H3K9 modification and as a result, this enzyme is capable of propagating the H3K9me mark from a narrow region to extended regions of the genome (Chen et al., 2012). In addition, enhancer of zeste (Ez) proteins (Ezh1 and Ezh2), are part of polycomb repressor complex 2 (PRC2) that can trimethylate H3K27 (Ezhkova et al., 2011). PRC2 establishes long term repression at homeobox genes to maintain stem cell identity and plasticity, and tri-methylation of H3K27 mediated by Ezh1/2 is one of the critical marks required for this repression (Richly et al., 2011).

Figure 1.8: Histone methyltransferases/demethylases.

The diagram lists the enzymes that catalyse the addition or removal of methyl groups from histone H3 lysines. The methyltransferases are listed on the top and demethylases are listed at the bottom of the histone mark they are associated with. Adapted from (Kondo, 2009).


Histone methyltransferases
SUV39H1
SUV39H2
G9a
EuHMT
SETDB1/ESET
RIZ1
MLL
SET1
SMYD3
EZH2
NSD1
SMYD2
Me
Me
Me
Me
NH2 – A – R – T – K – Q – T – A – R – K – S – T – G – G – K – A – P – R – K – Q – L – A – T – K – A – A – R – K – S – K – histone H3 tail
4
9
14
17
27
36
LSD1
SMCX / JARID1C
SMCY / JARID1D
RBP2 / JARID1A
PLU-1JARID1B
JHDM2A (JMJD1A)
JMJD2A
JMJD2B
JMJD2C
JMJD2D
LSD1-AR
UTX
JMJD3
JMJD2A
JMJD2B
JMJD2C
FBXL10 / JHDM1B
Histone demethylases

The majority of SET domain containing proteins recognize and methylates H3K4, a covalent modification that is associated with transcriptional activation. The first H3K4 methyltransferase to be discovered was SET1 in *Saccharomyces cerevisiae* (Miller et al., 2001). SET1 is assembled into a multi-subunit complex called complex proteins associated with Set1 (COMPASS), which consists of approximately seven subunits that are essential for regulating SET1 activity. SET1 is capable of mono-, di-, or tri-methylating H3K4 and its specificity is dependent on the subunit composition of the COMPASS complex. For example, the absence of brefeldin-A sensitivity protein 2 (Bre2/Csp60) subunit from the COMPASS complex prevents H3K4-trimethylation and the presence of Cps25 is required for di- and tri-methyaltion of H3K4 (Takahashi et al., 2011).

There are six human orthologues (MLL1 – MLL4, SET1A and SET1B) to the yeast SET1, and all six belong to the mixed lineage leukemia (MLL) family of proteins (Milne et al., 2002). Similar to the yeast COMPASS complex, all of the human MLL proteins are constituents of multi-subunit complexes and the methyltransferase activity of MLL proteins is also regulated by the associated subunits (Patel et al., 2009). In addition, the MLL complexes associate with many other activating complexes such as the MOF (a MYST family histone acetyltransferase) complex which also contains a HAT capable of acetylating H4K16, a mark that is also associated with transcriptional activation (Dou et al., 2005; Glaser et al., 2006; Milne et al., 2002). Collectively, the existence of many histone methyltransferases that uniquely modify a specific histone tail amino acid, represents the variety of ways eukaryotic cells can fine tune transcriptional regulation.

1.3.4 *Readers of Histone Methylation*

Lysine methylation serves as a marker for the recruitment of effecter proteins containing methyl-lysine binding domains that convert this modification to an appropriate biological outcome. Several classes of methyl-lysine binding domains have been identified in mammals and include WD40 repeats, plant homeodomain (PHD) fingers, CW domains, PWWP domains and ankryin repeats. Proteins belonging to the Royal super family also bind methylated histones. All of these proteins contain domains such as chromobarrels, chromodomains, double chromodomains, Tudor domains, and malignant brain tumor (MBT) repeats which have varying affinities for different methylated lysines (Taverna et al., 2007).

Heterochromatin protein 1 (HP1) contains a chromodomain that can bind H3K9me3 and repress transcription by stabilizing macromolecular complexes between nucleosomes, as well as preventing transcriptionally activating histone variants, such as histone H3.3, from being deposited (Cheutin et al., 2003; Jacobs and Khorasanizadeh, 2002; Janicki et al., 2004). In contrast, CHD1 is an ATP-dependent chromatin remodeler and contains two chromodomains which specifically binds H3K4me3 (Pray-Grant et al., 2005). CHD1 promotes transcriptional activation by altering chromatin structure through the SWI/SNF ATPase domain and interacting with coactivator complexes such as the histone Spt-Ada-Gcn5 acetyltransferase (SAGA) and SAGA-like (SLIK) complex (Pray-Grant et al., 2005; Simic et al., 2003). Collectively, these studies suggest that there are many lysine methylation readers capable of propogating this modification and determining a specific transcriptional outcome.

1.3.5 *Histone Lysine Demethylation*

Lysine methylation is also a reversible modification and the first demethylase discovered was lysine specific demethylase 1 (LSD1), also referred to as KDM1A (Shi et al., 2004). LSD1 is a flavin adenine dinucleotide (FAD)-dependant amine oxidase which catalyzes the demethylation of mono- and di-methylation of H3K4 and H3K9. This is accomplished by utilizing FAD to oxidize methyl-lysine to an imine intermediate which is then hydrolyzed to yield unmethylated lysine and formaldehyde (Figure 1.7) (Shi et al., 2004). LSD1 was initially identified as a component of the repressor element silencing transcription factor (REST) complex that is recruited to RE1 elements found at many neuron-specific genes (Lee et al., 2005; You et al., 2001). Purified recombinant LSD1 is capable of demethylating H3K4me1 and H3K4me2 in peptides and bulk histones, but not from nucleosomes. However, LSD1 is capable of demethylating H3K4me2 from nucleosomes as a component of the REST complex (Shi et al., 2005). Downregulation of LSD1 using small-interfering RNAs (siRNA) demonstrates an over-all increase in H3K4me2, mainly at neuronal specific genes (Amente et al., 2010). For this reason, LSD1 was initially implicated in transcriptional repression.

However, LSD1 is also required in many hormone dependent processes, such as estrogen and androgen receptor signaling where it acts as a H3K9 demethylase (mono or di) and is associated with transcriptional activation of many genes (Krig et al., 2010; Metzger et al., 2005). This dual mode of action by LSD1 is speculated to be context dependant such that specific protein associations alter the chemistry of the active site, to yield distinct specificities (Wysocka et al., 2005a). LSD1 has been identified as an integral component of several transcription complexes such as C-terminal binding protein

1 (CtBP1), NuRD and ZNF217 complexes (Cowger et al., 2007; Wang et al., 2007; Wang et al., 2009). LSD1 is also essential for mammalian development and homozygous knockout of LSD1 causes embryonic lethality by E7.5 (Wang et al., 2007).

Demethylation of mono, di or tri-methylated lysines is accomplished by a second class of lysine demethylases, the jumonji C (JmjC) domain-containing iron-dependent dioxygenases (Figure 1.8) (Takeuchi et al., 2006). This class of enzymes use iron and α-ketoglutarate as co-substrates and demethylation results in the generation of formaldehyde and succinate (Figure 1.7) (Klose et al., 2006). There are six JmjC domain containing families of demethylases within this class; JmjC domain containing demethylase I (JHDM1), jumonji domain containing protein 2 (JMJD2), jumonji at-rich interactive domain (JARID1), Lysine demethylase 6 (KDM6), KDM7 and KDM8 (Zhang et al., 2012). The first JmjC domain containg demethylase was identified in *S. cerevisiae* and was named the Jmjc domain containing histone demethylase-2 (JHD2P) which is capable of demethylating H3K4me2 or H3K4me3 (Liang et al., 2007). The mammalian homologues of JHD2P includes the JARID1 family members, JARID1a,b,c and d (Christensen et al., 2007).

JmjC domain containing demethylases are targeted to specific genes for transcriptional regulation. Interestingly, JMJD2 is found in a complex also containing LSD1 and mediates androgen receptor signaling. Within this complex, JMJD2 and LSD1 cooperate to demethylate H3K9 and activate androgen responsive genes (Wissmann et al., 2007). JARID1B represses many developmentally important genes by demethylating H3K4 and is crucial for the differentiation of embryonic stem (ES) cells towards a neural

lineage (Schmitz et al., 2011). JMJD5, part of the KDM8 family of demethylases, demethylates H3K36 and knockout in mice demonstrated the upregulation of the tumor suppressor gene p53 resulting in embryonic lethality (Oh and Janknecht, 2012).

1.4 Cytosine Methylation: Overview

DNA methylation is an epigenetic modification that plays a pivotal role in many biological processes such as silencing of transposable elements, defense against viral sequences, gene imprinting during development and X-inactivation. Importantly, the loss of normal DNA methylation patterns in mice often results in embryonic lethality and is a hallmark of all cancers (Li et al., 1992; Okano et al., 1999; Severson et al., 2012).

DNA methylation was discovered nearly 66 years ago, and it was immediately proposed to be involved in the regulation of gene expression (McCarty, 1946). Shortly after, it was demonstrated that DNA methylation occurs exclusively at the cytosine nucleotide (Hotchkiss, 1948). However, it was not until 40 years later that a correlation between DNA methylation and gene transcription was established (Compere and Palmiter, 1981). Although methylated cytosine only accounts for 1-4 % of the nucleic acids, it was shown to markedly affect transcription of many genes and numerous studies have found a high degree of correlation between DNA methylation and transcriptional silencing (Bird, 2002; Ehrlich et al., 1982).

Cytosine methylation occurs within the sequence context CG (CpG) dinucleotide or CNG trinucleotide. There are approximately 28 million CpG dinucleotides (http://genome.ucsc.edu/) in the human genome and 60-90 % of the CpGs are methylated (Siegfried and Cedar, 1997; Xie et al., 2009). The majority of CpGs are found within

repetitive DNA sequences and intergenic regions that contain transposable and/or retroviral elements. Approximately 45 % of the human genome contains transposable elements that are kept silenced by DNA methylation. This methylation dependent suppression is critical because active transposable elements can cause mutations and gene disruptions that are deleterious to the cell (Giordano et al., 2007; Lee et al., 2012).

CpG dinucleotides are also found at a much higher frequency within dense clusters referred to as CpG islands. Approximately 70 % of human genes have a CpG island associated with their promoter region and many are highly conserved between mice and humans (Figure 1.2) (Deaton and Bird, 2011). The majority of CpG islands at promoters are not methylated in somatic cells although some CpG islands become methylated during development (Feltus et al., 2003). An interesting feature of CpG islands is that they are commonly found in areas where there are fewer nucleosomes, and the nucleosomes which are found in these areas, usually contain histone tail modifications associated with transcriptional activation (Choi, 2010). Collectively, these observations suggest that CpG islands are involved in transcriptional regulation of genes.

Cancer cells exhibit drastically altered patterns of DNA methylation that includes a global hypomethylation along with hypermethylation of many promoter CpG islands. Promoter regions that are normally unmethylated in somatic cells are found hypermethylated in tumor cells and this is associated with silencing of many tumor suppressor genes (Das and Singal, 2004; Esteller, 2002; Teng et al., 2011). However, CpG methylation is also found within gene bodies and is less predictive of transcription. Infact, in some cases, it has been demonstrated, to be associated with gene expression (Aran et al., 2011; Mohn et al., 2008).

1.4.1 *Molecular Determinants of DNA Methylation*

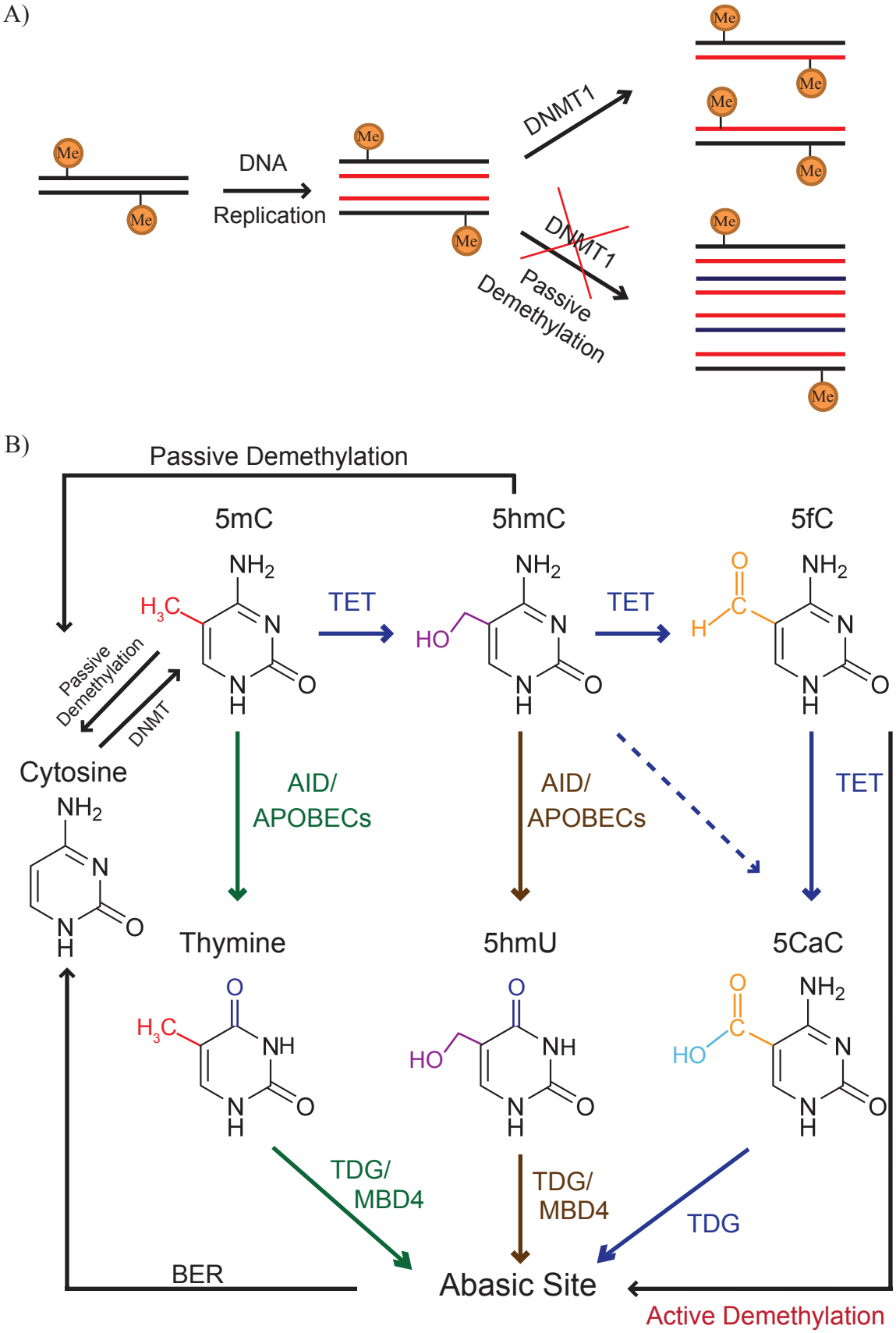
Cytosine methylation is catalyzed by the DNA methyltransferases (DNMT1, DNMT3A, and DNMT3B) which transfer a methyl group from the cosubstrate S-adenyl methionine (SAM) to the 5' position of cytosine (5mC). All of the DNMTs contain a carboxy-terminal methyltransferase domain and a cysteine rich zinc binding domain (CXXC), both of which are conserved between mouse and human (Xie et al., 1999). Mouse knockouts of DNMTs demonstrate the importance of these enzymes during development. Deletion of DNMT3A, 3B or DNMT1 is embryonic lethal and in all cases, hypomethylation was observed (Li et al., 1992; Okano et al., 1999).

DNMT1 is the maintenance methyltransferase that recognizes hemi-methylated DNA and during DNA replication, methylates the unmethylated daughter strand. This allows the 5mC mark to be stably conveyed following subsequent cell divisions (Figure 9A). DNMT1 contains six additional domains within the N-terminal portion of the protein which facilitate protein-protein interactions and targeting of DNMT1 to replication foci which are locations within the genome where the DNA is actively being replicated (Jurkowska et al., 2011). Interestingly, embryonic stem cells remain viable following knockout of DNMT1, however, upon differentiation massive cell death was observed suggesting that methylation is required for proper differentiation (Li et al., 1992).

DNMT3A and DNMT3B are highly homologous proteins, based on amino acid sequence, and are responsible for establishing *de novo* patterns of methylation, the majority of which occurs during early development (Reik et al., 2001). The N-terminus of

Figure 1.9: DNA demethylation mechanisms.

A. Passive DNA demethylation. Top branch represents maintenance of DNA methylation on the newly synthesized strand by DNMT1. Bottom branch represents inhibition of methylation by DNMT1 which leads to passive DNA demethylation following subsequent DNA replications. Black lines represent parental DNA strands, red lines represent newly synthesized strands after first division and blue lines represents newly synthesized strands after second division. Me represents methylation. **B.** Active DNA demethylation pathways. The pathway indicated in green involves deamination of 5mC by the AID/APOBEC deaminase enzymes to T which is excised by TDG or MBD4. The pathway indicated in blue involves oxidation of 5mC by the TET proteins to 5hmC, 5fC and 5CaC which is then recognized and cleaved by TDG. The pathway indicated in brown involves hydroxylation by TETs to 5hmC followed by deamination by AID/APOBECs to 5hmU which is cleaved by TDG or MBD4. The cleavage by TDG/MBD4 glycosylases results in an aprunic/aprymidinic site which is repaired by the BER machinery and replaced with the unmethylated cytosine. The generation of the 5hmC has also been observed to facilitate replication dependent passive DNA demethylation.


A)
Me
DNA
Replication
DNMT1
DNMT1
Passive
Demethylation
B)
Passive Demethylation
5mC
5hmC
5fC
TET
TET
Passive
Demethylation
DNMT
Cytosine
AID/
APOBECs
AID/
APOBECs
TET
Thymine
5hmU
5CaC
TDG/
MBD4
TDG/
MBD4
TDG
BER
Abasic Site
Active Demethylation

DNMT3A/B contains a PHD-like ADD domain that binds histone H3, and a PWWP domain which is involved in direct DNA binding and targeting of DNMT3A/B (Ge et al., 2004; Otani et al., 2009; Qiu et al., 2002; Shirohzu et al., 2002; Zhang et al., 2010). DNMT3A/B have also been found as part of multi-protein repressor complexes that are targeted to specific genes. For example, DNMT3A associates with the PRC complex to establish *de novo* methylation at many gene promoters (Mohammad et al., 2009; Vire et al., 2006). Alternatively, the underlying chromatin structure may also contribute to DNMT3A/B binding. DNMT3A/B possesses higher affinity towards unmethylated rather than methylated H3K4 and binds to this moiety through the ADD domain (Otani et al., 2009; Zhang et al., 2010). It is likely that both of these mechanisms are functional in eukaryotic cells to establish selective *de novo* DNA methylation patterns. Recently, it was demonstrated that the deletion of DNMT3A in hematopoietic stem cells resulted in the loss of differentiation emphasizing the important role of methylation during differentiation (Challen et al., 2012).

An additional member of this family, the DNMT3-like (DNMT3L) protein, is also involved in DNA methylation. DNMT3L is homologous to DNMT3A/B but lacks the catalytic domain. DNMT3L has been shown to dimerize and enhance the activities of DNMT3A or DNMT3B and is required for genomic imprinting, compaction of the X chromosome, and silencing of retroviral transposons (Hata et al., 2002).

1.4.2 *Consequences of DNA Methylation*

Although it is well accepted that DNA methylation is associated with transcriptional silencing, it is not clear whether methylation is an instructive modification

that dictates transcriptional repression, and if so, how the methyl marks confer this repression (Jones, 2012). One theory is that 5mC could directly prevent the binding of TFs that are required to initiate and/or facilitate transcription (Tate and Bird, 1993). Alternatively, many proteins selectively bind 5mC and can inhibit transcription by physically interfering with activators, or by associating with histone modifying proteins such as HDACs that promote localized chromatin condensation (Kavalali et al., 2011; Lewis et al., 1992).

There are three classes of proteins that bind methylated CpGs in mammals. First, ubiquitin-like containing PHD and RING finger domain (UHRF) proteins direct DNMT1 to hemi-methylated DNA (Achour et al., 2008). Second, a subclass of zinc finger proteins selectively bind methylated DNA. For example, the TF Kaiso binds to two consecutively methylated CpG dinucleotides (Daniel et al., 2002). Finally, methyl-CpG binding domain (MBD proteins) containing proteins recognize and bind methylated CpGs indiscriminately. Seven such proteins have been identified in mammals, MeCP1, MeCP2, MBD1, MBD2A/B, MBD3 and MBD4 (Hendrich and Bird, 1998; Lewis et al., 1992). MeCP1/2, MBD1 and MBD2 can directly bind methylated CpGs and recruit additional repressor proteins through an intrinsic transcriptional repression domain (TRD). For example, MBD proteins directly interact with HDACs through their TRD domain and recruit HDACs to methylated promoters, resulting in a hypoacetylated and more stably silenced state of genes (Kavalali et al., 2011; Nan et al., 1998; Sarraf and Stancheva, 2004). MeCP1/2 have been biochemically purified and are components of transcriptional corepressor complexes which also contain HDACs1/2 and Sin3A (Kavalali et al., 2011). It has also been demonstrated that MeCP2 can promote heterochromatin formation (Nan

et al., 1997). *Mecp2* null mice are viable but have neurological disorders that mimic Rett syndrome resulting in death at 6 and 12 weeks postnatal (Chen et al., 2001; Guy et al., 2001). Knockout of the various *Mbd* genes did not have any obvious developmental defects with the exception of *Mbd3* which is embryonic lethal (Hendrich et al., 2001; Martin Caballero et al., 2009; Zhao et al., 2003).

1.4.3 *DNA Demethylation*

In mammals, 5mC has long been regarded as a stable, heritable, epigenetic mark. In dividing cells, DNA can be demethylated through a passive mechanism whereby the newly synthesized sister strand is not remethylated by DNMT1, following DNA replication (Figure 1.9A) (Bird and Wolffe, 1999). This would result in the dilution of the methylation marks after subsequent divisions. In contrast, a much more rapid DNA demethylation mechanism has been described that occurs in response to specific signals and is DNA replication independent. This is known as active DNA demethylation and was first described in plants involving DNA glycosylases (Zhu, 2009).

Although orthologues of plant DNA glycosylases have not been identified, numerous studies have provided evidence for the existence of active DNA demethylation in mammals. Following fertilization the paternal genome, but not the maternal, is rapidly demethylated before the onset of DNA replication, indicative of an active process (Mayer et al., 2000; Oswald et al., 2000). In addition, there have been several examples of active demethylation at specific loci in somatic cells. The interleukin-2 promoter region is rapidly demethylated within 20 minutes following T-cell (T-lymphocytes) activation by anti-CD3 plus anti-CD28 and the observed demethylation was independent of DNA

replication (Bruniquel and Schwartz, 2003). Depolarization of post-mitotic neurons with KCl resulted in DNA demethylation of the brain-derived neurotophic factor (*bdnf*) gene promoter with subsequent increases in gene expression (Martinowich et al., 2003). Binding of the estrogen receptor to the *pS2* promoter is rapid and cyclical and is accompanied by rapid cycles of methylation and demethylation (Kangaspeska et al., 2008).

Although the mechanism of active DNA demethylation in mammals is unclear, current evidence points to the involvement of the BER pathway. In mammals, the deamination of C and 5mC can occur spontaneously to generate U and T, respectively. If not corrected, the resulting G:U and G:T mispairs can lead to mutations following subsequent rounds of DNA replication. In fact, it has been estimated that over one third of mutations arising in cancer are a result of spontaneous deamination of cytosine and 5 mC (Jones et al., 1992). To correct this spontaneous decay, mammalian cells have developed repair mechanisms that recognize and correct this lesion. During BER the G:T and G:U mispairs are initially recognized by DNA glycosylases that cleave the thymine/uracil base from its sugar back bone at the glycosidic bond. Two DNA glycosylases are known for this function, thymine DNA glycosylase (TDG) and MBD4 also known as MED1 (Cortellino et al., 2003; Robertson et al., 2009). Both of these enzymes produce an apurinic/apyrimidinic (AP) site which is cleaved by AP endonuclease (APE) and this allows DNA polymerase β to incorporate the correct nucleotide (Robertson et al., 2009).

1.4.4 *The DNA Glycosylases TDG and MBD4*

TDG belongs to the *Escherichia coli* mismatch-specific uracil-DNA glycosylase (MUG) superfamily of enzymes that are capable of excising uracil from single strand or double strand DNA and are implicated in the processing of G:T/G:U mispairs arising from spontaneous mutations (Lee et al., 2011). In addition, numerous studies have implicated TDG in gene transcription involving two mechanisms which may be interrelated; as a coactivator for various TFs and in active demethylation of 5mC.

A direct role for TDG in transcription was initially inferred when it was identified, by yeast two-hybrid screening, as an interacting partner for the ligand bound retinoic acid receptor (RAR) (Um et al., 1998). Since then it has been shown to directly interact with other nuclear receptors such as the ligand bound estrogen receptor alpha (ERα) and is recruited to ERα responsive gene promoters (Chen et al., 2003). In addition, TDG interacts with coactivators such as steroid receptor coactivator 1 (SRC1) and CBP (Lucey et al., 2005; Tini et al., 2002). TDG's interaction with CBP/p300 greatly enhanced the transcriptional activation by CBP. This study further demonstrated that CBP acetylated TDG at the amino terminus and that the acetylation promoted the dissociation of CBP from TDG. Interestingly, the CBP-TDG complex still retained the ability to cleave the mispaired G:T or G:U (Tini et al., 2002). Collectively these studies implicate TDG in transcriptional activation although the exact mechanism remained unclear.

TDG knockout mice have recently been generated in order to evaluate the contribution of TDG to DNA demethylation *in vivo*. Knockout of TDG in mice is

embryonic lethal caused by cardiovascular defects. Importantly, TDG knockout resulted in increased promoter DNA methylation which was attributed to the loss of TDG's glycosylase activity, coinciding with aberrant transcriptional repression of genes that are developmentally regulated and genes involved in retinoic acid receptor signaling (Cortazar et al., 2011; Cortellino et al., 2011). TDG was also required for recruiting the histone acetyltransferase CBP/p300 and the MLL proteins to target genes (Cortazar et al., 2011; Cortellino et al., 2011). This was consistent with the loss of active chromatin marks such as H3K4me3, and increases in repressive marks, H3K27me3 and H3K9me3 in the TDG null mice (Cortazar et al., 2011). Surprisingly, loss of TDG in ES cells or mouse embryonic fibroblasts (MEFs) did not show any significant differences in cell survival and mutation frequency in response to ionizing radiation or treatment with H2O2, suggesting that TDG's contribution to BER is relatively minor (Cortellino et al., 2011). These findings suggest that TDG plays a central role in gene-specific transcription and in epigenetic stability.

The human MBD4 protein was initially isolated through its interaction with the mut L homolog 1 (MLH1), a mismatch repair (MMR) protein, in a yeast two-hybrid screen (Bellacosa et al., 1999). MBD4 belongs to the MBD-containing proteins and consists of an N-terminal MBD and a C-terminal glycosylase domain (Sjolund et al., 2012). Although MBD4 is not homologous to TDG, it does possess similar activities towards U:G and T:G mispairs resulting from deamination of cytosine and 5mC, respectively (Hendrich and Bird, 1998). Surprisingly, *Mbd4* knockout (*Mbd4*$^{-/-}$) mice are viable and develop normally, however, mutation analysis demonstrated that *Mbd4*$^{-/-}$ showed a threefold increase in mutation rate compared to wild type (Millar et al., 2002).

In addition, when *Mbd4*$^{-/-}$ mice were crossed with mice containing heterozygous mutated Adenomatous polyposis coli (*Apc*) allele (*Apc*$^{min/+}$), that is prone to intestinal tumourgenesis, to generate a double mutant mice (*Mbd4*$^{-/-}$/*Apc*$^{min/+}$), an increase in intestinal tumor formation was observed compared to the *Apc*$^{min/+}$ background alone. Further examination of the tumors showed an increase in C:G to T:A transitions within the *Apc* allele suggesting that MBD4 functions as a tumor suppressor by preventing C to T transitions and its deletion/mutation can promote tumors in a cancer prone background (Wong et al., 2002).

One study demonstrates that MBD4 mutations were observed in approximately 20% of colorectal carcinomas, endometrial carcinomas and pancreatic primary tumours that also demonstrate microsatellite instability (MSI) (Riccio et al., 1999; Sjolund et al., 2012). MSI arises often through mutations in a stretch of poly A DNA sequence that either shortens or lengthens the sequence contributing to genomic instability (Giannini et al., 2002). Collectively these studies demonstrate that inactivation of MBD4 in human cancers can increase mutation frequency and can accelerate tumor progression in certain genetic background (Sjolund et al., 2012).

It has been reported that, MBD4 localizes to heterochromatinized regions of the genome and has been proposed to function in transcriptional repression (Hendrich and Bird, 1998). MBD4 represses transcription of *p16*^{ink4a} and *Mlh* genes in methylation dependent manner through direct association with Sin3a and HDAC1 (Kondo et al., 2005).

In contrast, recent studies using zebrafish as a model system have implicated MBD4 in active DNA demethylation. It was found that the glycosylase activity of MBD4 was essential for demethylating and activating an ectopically introduced methylated cDNA containing plasmid that involved recognizing and cleaving G:T mispairs. In addition, the zebrafish MBD4 could be functionally compensated by the human version suggesting the conserved nature of MBD4 function (Rai, Huggins et al. 2008). Collectively these studies demonstrate the versatility of MBD4 and it is possible that MBD4 functions redundantly in a variety of roles described in a context dependent manner.

The mechanism of active demethylation involving BER has become an area of intense research. However, it is becoming apparent that active DNA demethylation cannot be explained by a single unified mechanism but involves a cascade of enzymes that utilize different pathways (Figure 1.9B). One pathway involves the deaminase enzymes that convert 5mC to T to generate a T:G mispair that is excised by TDG or MBD4 and the BER machinery. The second pathway involves the metabolism of 5mC into various oxidation products by the ten-eleven translocation (TET) family of enzymes. Many of these oxidation products are recognized and excised by TDG and the BER enzymes.

1.4.5 *Active Demethylation involving Deamination of 5mC*

The deamination of 5mC to T, generates a G:T mismatch that is excised and replaced with cytosine by the concerted actions of DNA glycosylases (TDG or MBD4) and the BER pathway (Figure 1.9B) (Hashimoto et al., 2012a; Hashimoto et al., 2012b).

In mammals, deamination is mediated by the activation induced deaminase (AID)/apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC), a large family of zinc-coordinating enzymes involved in purines and pyrimidine metabolism. AID, a 24kDa protein, was initially shown to play a role in inducing variability in immunoglobulins, a process by which the mRNA generated from the immunoglobulin gene is deaminated to cause C to U transitions (Muramatsu et al., 2000). This provides functionally different products to generate high affinity antibodies during a process referred to as affinity maturation (Muramatsu et al., 2000; Muramatsu et al., 1999). The APOBEC family of proteins consists of four members, APOBEC 1-4 (Smith et al., 2012). The structure of both AID and APOBEC proteins are highly similar, consisting of alternating α helices and β sheets. The second and the third α helices are positioned to form a pocket consisting of histidine and two cysteines that hold a zinc ion crucial for the deamination of cytidine bases (Conticello et al., 2005).

Genome wide analysis demonstrated that knockdown of AID in primordial germ cells (PGCs) resulted in three times the methylation found in wild-type cells (Popp et al., 2010). The AID protein is required for the reprogramming of somatic cells towards pluripotency, which occurs in the absence of DNA replication, and is mediated by DNA demethylation and activation of key pluripotency genes such as *Oct4* and *Nanog* (Bhutani et al., 2010). In zebra fish AID/APOBEC2 function in a complex containing MBD4 and the growth arrest and DNA damage inducible 45 alpha (GADD45α) protein to demethylate an ectopically introduced methylated cDNA that involves deamination of 5mC to T (Rai et al., 2008). More recently, coimmunoprecipitation studies have demonstrated that AID, GADD45α and TDG interact when overexpressed in HEK293

and p19 cells suggesting that all three proteins function together as a DNA demethylase (Cortellino et al., 2011).

1.4.6 *Oxidation of 5mC by the TET proteins*

A second mechanism of active DNA demethylation involve the metabolism of 5mC to 5'-hydroxy-methylcytosine (5hmC) that is mediated by the TET family of enzymes (Figure 1.9B) (Ito et al., 2010). 5hmC is a reaction intermediate that is not recognized by either TDG or MBD4. However 5hmC is oxidized further by an interative process involving the TET proteins to 5'formyl-cytosine (5fC) and then to 5'carboxyl-cytosine (5CaC) (Figure 1.9B) (Ito et al., 2011). Alternatively 5hmC can be deaminated to 5-hydroxymethyluracil (5hmU) (Figure 1.9B) (Cortellino et al., 2003). 5fC, 5CaC and 5hmU are all efficient substrates for TDG *in vitro* (Cortellino et al., 2003; Maiti and Drohat, 2011).

It has been shown in mESCs that TDG knockdown results in an accumulation of 5caC (He et al., 2011). However the amount of 5caC, under TDG-depleted conditions, is still well below the level of 5fC found under normal conditions. This finding is somewhat paradoxical given the high efficiency of TET proteins to convert 5fC to 5caC (He et al., 2011). If TDG is the only enzyme capable of processing 5caC, one would predict that 5caC to accumulate at a much higher level under TDG depleted conditions. This suggests that additional enzymes that possess 5caC processing activity may exist in mammalian cells.

Although derivatives of 5mC are generally regarded as metabolic intermediates, accumulation of these derivatives in some tissues have been observed and a very recent

study using quantitative mass spectrometry in combination with affinity purification has identified proteins which bind selectively to 5mC, 5hmC, 5fC and 5caC, suggesting that these products may represent novel epigenetic marks (Spruijt et al., 2013). In the adult, levels of 5hmC vary between tissues with high levels found in the central nervous system, accounting for 0.3 to 0.7 % of all cytosines (Kriaucionis and Heintz, 2009; Penn et al., 1972). 5hmC is also believed to play a role during development. High levels of 5hmC are observed immediately after fertilization but drops sharply as the embryo develops (Lee et al., 2002; Oswald et al., 2000; Wossidlo et al., 2011). Levels of 5hmC are elevated once again as the embryo reaches the blastocyst stage (~E11.5 to E12.5) and in particular, the inner cell mass contains high levels of 5hmC. The changes in the levels of 5hmC observed during development, inversely correlates with the expression of the TET enzymes. Surprisingly, TET1 or TET2 knockout mice are viable and grow normally whereas knockout of TET3 is embryonic lethal (Tan and Shi, 2012).

All three TET enzymes contain a C-terminal dioxygenase domain and are capable of converting 5mC to 5hmC, in the presence of ATP and alpha-ketogluterate (αKG) as co-substrates (Ficz et al., 2011; Ito et al., 2010; Ko et al., 2010; Szwagierczak et al., 2010). However, only TET1 and TET3 contain an amino terminal CXXC domain that possesses DNA binding activity and can target TET1/3 to gene promoters (Xu et al., 2012). In mESCs TET1 and 5hmC colocalize to many promoter CpG islands at promoters and maintain pluripotency through promoter DNA demethylation and expression of key genes such as *Oct4* and *Nanog* (Ito et al., 2010; Williams et al., 2011; Xu et al., 2011). Consistently, downregulation of TET1 resulted in reduced 5hmC levels coinciding with the loss of stem cell identity (Freudenberg et al., 2012).

The metabolism of 5mC by TET proteins also plays an important role in cancer. Mutations in TET2 have been identified in a subset of myeloid leukemias and correlate with DNA hypermethylation and low levels of 5hmC in patients (Ko et al., 2010). Furthermore, mice that have a conditional knockout of *Tet2* demonstrate an increase in hematopoietic stem cell self renewal and myeloid transformation (Moran-Crusio et al., 2011; Tefferi et al., 2009). Isocitrate dehydrogenase 1 and 2 (IDH1/2) convert isocitrate to αKG which is the major co-substrate for TET proteins. Studies have shown that IDH1/2 are mutated in approximately 80% of adult glioblastomas and in acute myeloid leukemias (Figueroa et al., 2010; Gross et al., 2010; Yan et al., 2009). These mutations are restricted to specific residues that cause the IDH proteins to acquire novel enzymatic activity and converts αKG to D-2-hydroxyglutarate (D2HG) in a reaction that reduces NADP to NADPH. D2HG acts as a competitive inhibitor of αKG dependent enzymes, such as the TETs, and results in elevated 5mC levels in DNA, generating a genome-wide hypermethylation signature characteristic of transformed cells (Gross et al., 2010; Ward et al., 2010). Furthermore, TET1 was shown to be downregulated in prostate and breast cancer tissues and downregulation of TET1 causes invasion, tumor growth, and cancer metastasis in prostate xenograft models (Hsu et al., 2012).

1.5 The Cell Cycle and the Restriction Point

Cell division, the generation of two identical daughter cells from a parent cell, is divided into four stages that are sequentially orchestrated by cyclins and cyclin-dependent-kinases (CDKs). The cell cycle consists of two gap phases, G1 and G2, interspersed between two main phases, synthesis (S) and mitosis (M), during which the chromosomes are replicated and segregated, respectively (Vermeulen et al., 2003). Proper

regulation of cell division is essential for development and de-regulated cell division is an underlying cause of cancer (Mitra et al., 2012). A major point of cell cycle regulation exists within the G1 phase of the cell cycle referred to as the restriction (R) point. Following mitosis, cells require constant stimulation by mitogens, such as growth factors, to continue past the R point and once the cell cycle passes this point, the cells are fully committed to cell division. Alternatively, the lack of mitogen stimulation or growth inhibitory signals can cause the cells to exit the cell cycle prior to the R point and enter a quiescent (G0) stage (Zetterberg et al., 1995).

The retinoblastoma protein (Rb) is critical for maintaining cells in G0. Rb belongs to the pocket protein family which also contains p107 and p130. The Rb family of proteins binds to E2F TFs and prevents the transcription of genes necessary for S-phase entry. Upon mitogen stimulation, Cyclin-cyclin dependent kinase (CDK) complexes are activated which result in the phosphorylation of Rb. This releases E2Fs and activates E2F-dependent transcription. Conversely, growth inhibitory signals can prevent cell division by upregulating CDK inhibitors and downregulating CDK activators (Weinberg, 1996). Common inhibitors of CDKs (ink) are $p15^{ink4b}$, $p16^{ink4a}$, $p27^{kip1}$ (Kinase interacting protein 1), $p57^{kip2}$ and $p21^{cip1}$ (CDK-interacting protein 1), Activators of of CDKs include the cell division cycle 25 A (CDC25A). The $p15^{ink4b}$ and $p16^{ink4a}$ proteins are cyclin D-CDK4/6 inhibitors that bind the catalytic site of the kinase and prevent phosphorylation of Rb. $p27^{kip1}$, $p57^{kip2}$ and $p21^{cip1}$ can inhibit a wide range of cyclin-CDK complexes by a similar mechanism (Sherr and Roberts, 1995). CDC25A is a protein phosphatase that dephosphorylates CDK2 and CDK1 promoting progression through the cell cycle (Draetta and Eckstein, 1997).

Cell division can be regulated by many external signals. For example, many growth factors bind to their respective receptors on the cell surface to activate Ras which initiates a kinase cascade that results in the transcriptional upregulation of cyclin D1 (Jiang et al., 2011; Liu et al., 1995). Ras, or components of the Ras-pathway, are often mutated and/or constitutively activated in many cancers (Downward, 2003). Conversely, the TGFβ cytokine can activate a critical tumor suppressor pathway, leading to the expression of many cell cycle inhibitors that can cause cell growth arrest at G1 phase (Planas-Silva and Weinberg, 1997; Su et al., 2007).

1.6 The TGFβ Signalling Pathway

TGFβ is a family of secreted growth factors that consists of over 30 types of ligands in humans that are subdivided into two major subfamilies, TGFβ-activin-nodal and bone morphogenic proteins (BMP) (Greenwald et al., 2003).

TGFβ ligands bind to TGFβ Receptor I (TGFβRI) and TGFβRII which belong to the serine/threonine kinase receptor family of proteins. In mammals, there are seven type I receptors and five type II receptors that can dimerize in various combinations to activate two pathways, a SMAD-dependent and SMAD-independent pathway (Manning et al., 2002; Massague, 2012). The SMAD-dependent pathway consists of eight SMAD proteins (SMAD1-8) which are structurally similar consisting an N-terminal mad homology 1 (MH1) and a C-terminal MH2 domains linked by an unstructured region. The MH1 domain interacts with DNA, and the MH2 domain can interact with many proteins such as the activated TGFβ receptor or nuclear pore proteins which mediate the translocation of SMADs into the nucleus. SMADS 1-3, 5 and 8 also contain an SXS motif at the C-

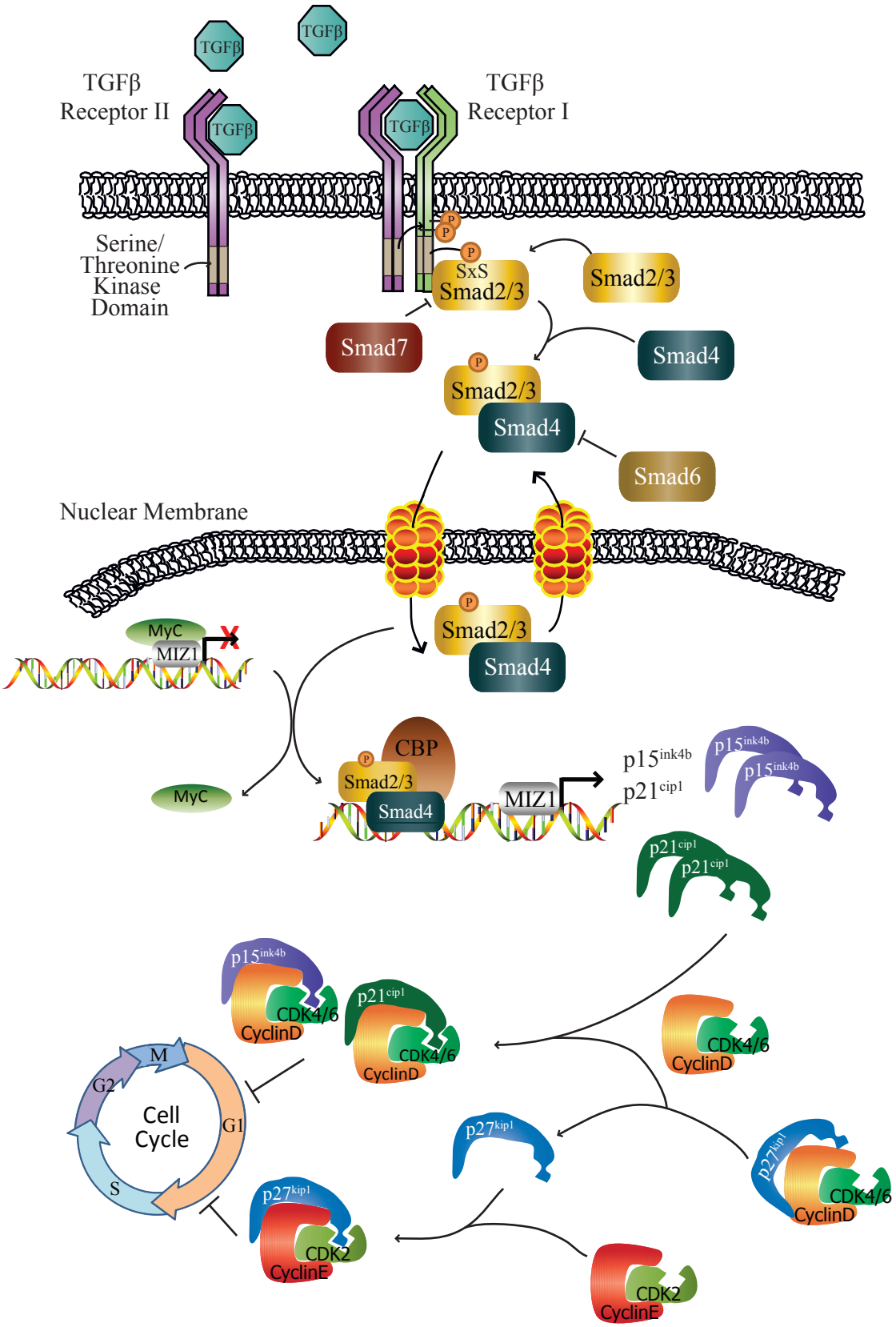
terminal end and phosphorylation of this motif promotes complex formation with SMAD4 (Abdollah et al., 1997; Shi and Massague, 2003).

Binding of TGFβ ligand promotes the heterodimerization of TGFβRI with TGFβRII (Figure 1.10). The TGFβRII then phosphorylates and activates the kinase domain of TGFβRI which, in turn, phosphorylates receptor regulated SMADs (R-SMADs) promoting heterodimerization with SMAD4. The Nodal/Activin class of TGFβ ligands activates SMAD2 and SMAD3 and the BMP subfamily of TGFβ ligands activate SMAD1, SMAD5 and SMAD8 (Massague, 2012). Two additional proteins complete the SMAD family, SMAD 6 and 7, which are inhibitory SMADs that negatively regulate TGFβ signaling. SMAD7 prevents the phosphorylation of R-SMADs by competitive TGFβRI binding. SMAD6 binds to SMAD4 and inhibits its interaction to the R-SMADs (Itoh et al., 2001).

The R-SMADs/SMAD4 complex translocate into the nucleus where it binds to specific regions of the genome to regulate gene transcription (Figure 1.10) (Shi and Massague, 2003). SMADs can be targeted to genes indirectly by associating with tissue-specific proteins, such as tripartite motif containing 33 (TRIM33) and inhibitor of DNA binding 1 (ID1) (Massague et al., 2005; Young, 2011). Alternatively, the SMADs can directly bind to DNA through a SMAD binding element (SBE) consisting of a consensus CAGAC sequence. The DNA bound SMAD complex also associates with various coregulators such as CBP/p300 to activate transcription (Pouponnot et al., 1998). Alternatively SMAD complexes can repress some genes by associating with HDACs and CtBP1/2 (Massague et al., 2005).

Figure 1.10: TGFβ signaling leading to cell growth arrest.

The model depicts the signaling events that result from activation of the cell surface TGFβ receptors to transcriptional activation by the SMAD complexes. TGFβ1 ligand binding promotes the formation of TGFβRII/TGFβRI heterodimerization. The kinase domain of TGFβRII phosphorylates TGFβRI resulting in TGFβRI kinase activation. TGFβRI then specifically phosphorylates SMAD2/3 which heterodimerizes with SMAD4 and localizes to the nucleus. SMAD7 can inhibit the interaction of SMAD2/3 with TGFβRI and SMAD6 can inhibit heterodimerization between SMAD4/SMAD2/3. In the nucleus, SMAD4/SMAD2/3 complex interacts with co-regulators such as CBP/p300 and regulate transcription of genes. The TGFβ cytostatic program includes the increased expression of the CDK inhibitors $p15^{ink4b}$ and $p21^{cip1}$ through the binding of SMAD activation complex and release of MYC repression from the promoters. Upregulation of $p15^{ink4b}$ and $p21^{cip1}$ leads to the binding and inhibition of cyclin D-CDK4/6. This promotes the release of $p27^{kip1}$ from the cyclin D-CDK4/6 complex and $p27^{kip1}$ can inactivate cylin E-CDK2. These events ultimately lead to cell growth arrest at G1 phase.


TGFβ
TGFβ
TGFβ Receptor II
TGFβ
TGFβ
TGFβ Receptor I
Serine/ Threonine Kinase Domain
P
P
P
SxS
Smad2/3
Smad2/3
Smad7
Smad4
P
Smad2/3
Smad4
Smad6
Nuclear Membrane
MyC
MIZ1
X
P
Smad2/3
Smad4
MyC
CBP
P
Smad2/3
Smad4
MIZ1
p15ink4b
p21cip1
p15ink4b
p15ink4b
p21cip1
p21cip1
p15ink4b
CDK4/6
CyclinD
p21cip1
CDK4/6
CyclinD
CDK4/6
CyclinD
M
G2
Cell Cycle
G1
S
p27kip1
p27kip1
CDK2
CyclinE
p27kip1
CDK4/6
CyclinD
CDK2
CyclinE

TGFβ can also regulate gene expression independent of SMADs. For example, in some cells, the activated TGFβRs initiate the mitogen activated protein (MAP) kinase signaling cascade involving Ras and Raf that leads to the activation of the extracellular signal regulated kinase (ERK) protein. ERK then translocates into the nucleus and mediates transcription of many genes (Lafontaine et al., 2011; Zhang, 2009).

The TGFβ signaling pathway is essential for mammalian development and is required for an array of cellular responses, such as axis formation and the subsequent patterning to generate tissues during embryogenesis (Little and Mullins, 2006; O'Connor et al., 2006). Genetic ablation of the TGFβ signaling proteins, such as *TfgβrII*, *Smad2* or *Smad4* leads to embryonic lethality in mice (Goumans and Mummery, 2000; Larsson et al., 2001; Oshima et al., 1996; Weinstein et al., 1998). In the adult, TGFβ maintains homeostasis by playing a role in a wide variety of cellular processes which includes cell differentiation, growth, apoptosis and migration (Derynck and Akhurst, 2007; Massague et al., 2000).

In addition, TGFβ is a potent anti-proliferative cytokine and coordinates this function through transcriptional regulation of genes that primarily target cell growth arrest at the G1 phase of the cell cycle (Figure 1.10) (Hanahan and Weinberg, 2000; Laiho et al., 1990). In many cell types, TGFβ upregulates the expression of cyclin dependent kinase inhibitors $p15^{ink4b}$ and $p21^{Cip1}$ and downregulates the expression of proteins that promote cell division, including CDC25A, myelocytomatosis viral oncogene homolog (MYC), ID1, ID2 and ID3 (Brown and Bhowmick, 2004; Feng et al., 2000; Kang et al., 2003).

G1/S phase transition requires the actions of cyclin D-CDK4/6 and cyclin E-CDK2 complexes to phosphorylate Rb. The TGFβ induced CDK inhibitors, $p15^{ink4b}$ and $p21^{Cip1}$, bind and inhibit these CDKs (Besson et al., 2008). In dividing cells, CDK inhibitor $p27^{kip1}$ is bound to cyclin D-CDK4/6 and the $p27^{kip1}$-cyclin D-CDK4/6 complex still retains its kinase activity. The binding of $p15^{ink4b}$ to cyclin D-CDK4/6 complex dislodges $p27^{kip1}$ allowing it to bind and repress cyclin E-CDK2 complex promoting complete repression of the CDKs (Figure 1.10) (Blain et al., 1997).

The TGFβ anti-proliferative response is also dependant on MYC levels. MYC acts as a repressor of the msx interacting zinc finger 1 (MIZ1) which is bound to the $p15^{ink4b}$ and $p21^{cip1}$ promoter and is required for the activation of $p15^{ink4b}$ and $p21^{cip1}$ genes. TGFβ causes downregulation of MYC and this relieves its repression of MIZ1 and allows activation of $p15^{ink4b}$ and $p21^{cip1}$ gene.

MYC is an activator of *ID2* transcription and downregulation of *MYC* by TGFβ also leads to *ID2* silencing. The ID proteins promote growth by preventing differentiation through inhibition of the basic helix-loop-helix (bHLH) TFs and ID2 promotes cell division by physically interacting and sequestering Rb. TGFβ also stimulates the expression of activating transcription factor 3 (*ATF3*) and ATF3 interacts with the SMAD complex to repress *ID1* transcription (Kang et al., 2003).

The anti-proliferative effects of TGFβ are often lost during cellular transformation and this feature is considered a hallmark of cancer (Hanahan and Weinberg, 2000; Reiss, 1997). The link between TGFβ signaling and cancer was demonstrated using mouse genetic studies. Overexpression of a dominant negative *TfgβrII* or deletion of *Smad4* in mice resulted in increased tumor frequency compared to wild type. Treatment of these

mice with carcinogens lead to increased accumulation of tumors compared to the wild type mice (Barcellos-Hoff and Ewan, 2000; Go et al., 2000; Tang et al., 1998; Xu et al., 2000). Frequent mutations of *TGFβR*s and *SMAD*s have been identified in specific types of cancers (Goggins et al., 1998; Grady et al., 1999). Bi-allelic loss of *SMAD4* has been observed in approximately 50 % of pancreatic cancers and one third of metastatic colon tumors (Hahn et al., 1996; Miyaki et al., 1999). *TGFβRII* mutations have been identified in human colorectal and gastric carcinomas (Markowitz et al., 1995). Deletion/mutations of *TGFβRI* have been identified with ovarian, breast, pancreatic, biliary, and colon cancers (Chen et al., 1998; Goggins et al., 1998; Pasche et al., 1999; Wang et al., 2000).

In summary, during the early stages of transformation the cytostatic and apoptotic functions of TGFβ help restrain growth of mammalian cells and TGFβ acts as a tumour suppressor. However in later stages of cancer the TGFβ pathway is disrupted resulting in hyperproliferation and tumor promotion (Kang et al., 2005). How specific oncogenic signals attenuate the tumor-suppressor functions of TGFβ and turn this cytokine into a tumor promoter is currently unclear.

1.7 Zinc Finger 217 Protein

ZNF217 is a Kruppel-like TF of approximately 1100 amino acids and contains 8 Cys2 – His2 zinc fingers. ZNF217 mainly localizes to the nucleus and binds directly to DNA (Nunez et al., 2011; Vandevenne et al., 2013). Two consensus sites for ZNF217 have been identified. Using an *in vitro* approach the consensus CAGAAY (Y represents C or T) was isolated, and using a sequence comparison analysis of global ZNF217 targets, the consensus ATTCCNAC (N represents any nucleotide) was identified (Cowger et al., 2007; Krig et al., 2007). Recent crystal structures demonstrate that

ZNF217 interacts with DNA through the 6th and 7th zinc fingers to the consensus sequence, CAGAAY (Nunez et al., 2011). Subsequently, an extended version of this consensus sequence was identified which conferred stronger binding, (T/A)(G/A)CAGAA(T/G/C) (Vandevenne et al., 2013). Mutation of this site only resulted in a modest decrease in ZNF217 binding and it was demonstrated that a lower affinity non-specific interaction between DNA and ZNF217 in solution also exists. *In vitro* assays demonstrate that ZNF217 is a transcriptional repressor and is a component of a 1.5 MDa repressor complex (Cowger et al., 2007; Vandevenne et al., 2013). This complex also consists of CoREST and several histone modifying enzymes, including HDACs (1 or 2), LSD1 and the CtBP1 (Cowger et al., 2007; Kuppuswamy et al., 2008; You et al., 2001; Zhang et al., 2001). ZNF217 directly interacts with both CoREST and CtBP1 through its N-terminal and C-terminal portions, respectively. Both motifs are essential for conferring the transcriptional repressor properties of ZNF217 (Cowger et al., 2007; Quinlan et al., 2006a).

CoREST contains two SANT domains (SWI3, adaptor 2 (ADA2), nuclear receptor co-repressor (N-CoR), TFIIB) which are domains capable of binding unmodified histone tails. However in the context of CoREST, the SANT domain is involved in protein-protein interaction. CoREST interacts with HDAC1/2 through the amino terminal SANT domain and LSD1 through the linker region between the two SANT domains (Shi et al., 2005; You et al., 2001). *In vitro* experiments have shown that CoREST acts as an auxiliary protein and enhances the activity of LSD1 on methylated nucleosomes (Shi et al., 2005).

CtBP1 and CtBP2 are proteins that share a high degree of amino acid sequence similarity and function primarily as adaptor proteins by facilitating the formation of multi-protein complexes (Chinnadurai, 2002; Sewalt et al., 1999; Shi et al., 2003; Siles et al., 2013). CtBP proteins contain two hydrophobic binding pockets at each end of the protein which bind many PXDLS (where X represents any amino acid) motif containing proteins such as early region 1A (E1A) (Quinlan et al., 2006b). These interactions are further enhanced by the RRT binding surface housed between the PXDLS pockets (Kuppuswamy et al., 2008). It has been reported that CtBP proteins possess NADH-dependant dehydrogenase activity and undergo NAD(H) dependent dimerization although its dehydrogenase function is controversial (Kumar et al., 2002; Schaeper et al., 1995). ZNF217 also contains the PXDLS and RRT motifs within the C-terminal end, where CtBP1 was observed to bind (Quinlan et al., 2006a). CtBP2, and not CtBP1, is crucial during development and overexpression of CtBPs in cancers correlates with increased tumor growth and epithelial to messenchymal transition (Chinnadurai, 2009; Hildebrand and Soriano, 2002).

Interestingly ZNF217 has also been shown to activate transcription in some contexts. ZNF217 directly binds and activates the promoter of the *ERBB3* receptor tyrosine kinase (V-ERB-B2 avian erythroblastic leukemia viral oncogene homolog 3) gene that can as an oncogene. In contrast, CtBP2 represses *ERBB3,* but ZNF217 and CtBP2 are both found at the *ERBB3* promoter suggesting a complex mechanism of transcriptional regulation. However, the mechanism by which ZNF217 functions as an activator still remains unclear (Krig et al., 2010).

ZNF217 is found at the 20q13 chromosomal region which is found amplified in a number of different cancer types including ovarian and breast (Collins et al., 1998). Cancers possessing the 20q13 amplification are associated with reduced patient survival, increased cell proliferation and increased tumor aggressiveness and grade (Tanner et al., 1995). Positional cloning of the 1 MB region at 20q13 allowed for the isolation of a highly amplified central 260 Kb region and identified ZNF217 as the strongest candidate within the maximally amplified region (Collins et al., 1998).

ZNF217 amplification and overexpression has been shown to correlate with lower patient survival in many cancer types such as ovarian, colon and breast (Collins et al., 1998; Ginestier et al., 2006). ZNF217 amplification in breast tumors is associated with increased tumor grade and aggressiveness with poor prognosis (Tanner et al., 1995). Recently it was demonstrated that ZNF217 is overexpressed in 71.2% of Glioblastoma multiforme (GBM) an aggressive type of brain tumor that arises from a subpopulation of Glioma stem cells (GSCs). It is postulated that ZNF217 contributes to this cancer type as a consequence of its ability to prevent differentiation of GSCs and its sustained expression in hypoxic conditions (Mao et al., 2011). ZNF217 overexpression was also found to have increased self-renewal capacity, expression of mesenchymal markers, motility, and metastasis (Rahman et al., 2012). Consequently ZNF217 overexpression represents a strong prognostic marker associated with poor clinical outcome (Littlepage et al., 2012).

Numerous studies have provided strong evidence that ZNF217 overexpression contributes to the transformed phenotype. Transduction of finite life-span human mammary epithelial cells (HMEC) with a retrovirus consisting of a construct that

overexpresses ZNF217, gives rise to immortalized cells with increased telomerase activity coinciding with stable telomere length, increased cell proliferation and resistance to TGFβ growth inhibition which are all key characteristics of transformed cells (Hanahan and Weinberg, 2000; Nonet et al., 2001). Transduction of ovarian epithelial cells with a ZNF217 overexpressing retrovirus showed similar characteristics, as well as increased survival in reduced serum conditions and anchorage independent cell growth. Furthermore, silencing of ZNF217 in the ovarian epithelial cells using siRNA caused growth arrest demonstrating that the sustained proliferation was due to ZNF217 (Li et al., 2007). In addition to a proliferative advantage, many cancer cell types which have ZNF217 overexpression also develop resistance to apoptotic signals and chemotherapeutic drugs such as doxorubicin (Song et al., 2005). Furthermore, ZNF217 was demonstrated to function together with the TGFβ-SMAD signaling pathway to promote epithelial to messenchymal transition (EMT). Coincidently, overexpression of ZNF217 promoted migration and invasion in breast cancer cells, increased lung/node metastasis in mice and correlates with poor prognosis (Vendrell et al., 2012).

In summary, ZNF217's oncogenic potential is compelling but the molecular mechanisms employed by ZNF217 to drive and sustain oncogenic properties, such as increased proliferation survival, and resistance to TGFβ, remain elusive.

1.8 Hypothesis and Summary of Objectives

Strong evidence suggests that overexpression of ZNF217 causes patterns of cellular deregulation that are consistent with oncogenesis. However, its molecular mechanism of action remains poorly understood. I have conceptualized a global hypothesis which encompasses the entire scope of my study as follows; Overexpression

of ZNF217, as has been found in cancers, causes disruption of transcriptional regulation resulting in aberrant expression of specific target genes which provides a growth advantage to the host cell. Ultimately, my goal was to identify gene targets for ZNF217 that would provide insight into its molecular mechanisms of action as well as to better understand its oncogenic potential.

In chapter 2 of my thesis, I utilized a modified chromatin immunoprecipitation (ChIP)-on-chip approach known as chromatin immunoprecipitation with directed selection and ligation (ChIP-DSL) to identify targets on a 20,000 gene promoter array. This was done in conjunction with microarray gene expression analysis following ZNF217 knockdown. By comparing the results of both approaches, I identified a group of genes directly regulated by ZNF217. Among the genes identified, I focused on the $p15^{ink4b}$ CDK inhibitor because it was known that ZNF217 overexpression was associated with increased cell proliferation and TGFβ resistance; functions that are tightly linked to $p15^{ink4b}$ activity. Furthermore, the transcriptional activation of 15^{ink4b} has been well characterized, and so it provided me with an ideal model system to study the molecular mechanisms of transcriptional regulation employed by ZNF217. I showed that there is a good correlation between recruitment of the ZNF217 complex and repression of $p15^{ink4b}$. Using ChIP, I mapped the ZNF217 binding site to the proximal promoter region of $p15^{ink4b}$ and I established a good correlation between the presence of ZNF217, specific histone marks and TGFβ signaling. I demonstrated a coregulator exchange following TGFβ treatment that involves the release of ZNF217 and the binding of SMADs at the $p15^{ink4b}$ promoter.

In chapter 3 I extended the findings from chapter 2 in several ways. First, I have used ChIPseq which has much greater resolving power than ChIP-DSL and this allowed me to identify binding sites genome-wide. Additionally, I have performed the ChIPseq analysis for two components of the ZNF217/CoREST complex, ZNF217 and CTBP1 which was not done in chapter 2. I have performed some analysis on this data to define binding sites of these proteins within the genome. This demonstrated a significant overlap between ZNF217 and CtBP1, which confirms previous biochemical observations regarding the association of ZNF217 and CtBP1 in the same complex (Cowger et al., 2007; Quinlan et al., 2006a). I have also performed knockdown experiments in order to examine a requirement for specific components of the ZNF217 complex. I also conducted additional experiments regarding the downstream events resulting from the ZNF217/p15^{ink4b} association, such as effects on the cell cycle and Rb status.

Chapter 3 also broadened the mechanistic aspects which began in chapter 2 by examining the dynamics of DNA methylation at the *p15*^{ink4b} promoter. I observed a strong correlation between DNA hypermethylation of *p15*^{ink4b} promoter and silencing of the *p15*^{ink4b} gene in MCF7 cells, and based on my preliminary observations in chapter 2, I formulated a secondary hypothesis as follows: In addition to coregulator exchange, TGFβ stimulation causes active DNA demethylation and this mechanism is inhibited by ZNF217 overexpression. I was able to verify this hypothesis experimentally at the *p15*^{ink4b} promoter in HaCAT cells. I identified a novel activator complex consisting of SMAD2/3, TDG and CBP which bound the *p15*^{ink4b} promoter in close proximity to the region occupied by the ZNF217 complex and mediated the TGFβ induced active DNA demethylation. Finally, overexpression of ZNF217 using adenovirus, abrogated the active

DNA demethylation imposed by TGFβ, by preventing the binding of the demethylase protein complex, coinciding with repression of the *p15*^{ink4b} gene.

1.9 References

Abdollah, S., Macias-Silva, M., Tsukazaki, T., Hayashi, H., Attisano, L., and Wrana, J.L. (1997). TbetaRI phosphorylation of Smad2 on Ser465 and Ser467 is required for Smad2-Smad4 complex formation and signaling. J Biol Chem *272*, 27678-27685.

Achour, M., Jacq, X., Ronde, P., Alhosin, M., Charlot, C., Chataigneau, T., Jeanblanc, M., Macaluso, M., Giordano, A., Hughes, A.D.*, et al.* (2008). The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. Oncogene *27*, 2187-2197.

Amente, S., Bertoni, A., Morano, A., Lania, L., Avvedimento, E.V., and Majello, B. (2010). LSD1-mediated demethylation of histone H3 lysine 4 triggers Myc-induced transcription. Oncogene *29*, 3691-3702.

Antequera, F., and Bird, A. (1993). Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci U S A *90*, 11995-11999.

Aran, D., Toperoff, G., Rosenberg, M., and Hellman, A. (2011). Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet *20*, 670-680.

Bali, P., Pranpat, M., Bradner, J., Balasis, M., Fiskus, W., Guo, F., Rocha, K., Kumaraswamy, S., Boyapalle, S., Atadja, P., *et al.* (2005). Inhibition of histone deacetylase 6 acetylates and disrupts the chaperone function of heat shock protein 90: a novel basis for antileukemia activity of histone deacetylase inhibitors. J Biol Chem *280*, 26729-26734.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Bararia, D., Trivedi, A.K., Zada, A.A., Greif, P.A., Mulaw, M.A., Christopeit, M., Hiddemann, W., Bohlander, S.K., and Behre, G. (2008). Proteomic identification of the MYST domain histone acetyltransferase TIP60 (HTATIP) as a co-activator of the myeloid transcription factor C/EBPalpha. Leukemia *22*, 800-807.

Barcellos-Hoff, M.H., and Ewan, K.B. (2000). Transforming growth factor-beta and breast cancer: Mammary gland development. Breast Cancer Res *2*, 92-99.

Baxevanis, A.D., and Vinson, C.R. (1993). Interactions of coiled coils in transcription factors: where is the specificity? Curr Opin Genet Dev *3*, 278-285.

Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A.T., Matsumoto, Y., Golemis, E.A., Genuardi, M., and Neri, G. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc Natl Acad Sci U S A *96*, 3969-3974.

Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T., and Schreiber, S.L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci U S A *99*, 8695-8700.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Besson, A., Dowdy, S.F., and Roberts, J.M. (2008). CDK inhibitors: cell cycle regulators and beyond. Dev Cell *14*, 159-169.

Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature *463*, 1042-1047.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

Bird, A. (2007). Perceptions of epigenetics. Nature *447*, 396-398.

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

Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451-454.

Blain, S.W., Montalvo, E., and Massague, J. (1997). Differential interaction of the cyclin-dependent kinase (Cdk) inhibitor p27Kip1 with cyclin A-Cdk2 and cyclin D2-Cdk4. J Biol Chem *272*, 25863-25872.

Blyth, N.J. (2012). Mechanisms and techniques of reprogramming - Using PDX-1 homeobox protein as a novel treatment of insulin dependent diabetes mellitus. Diabetes Metab Syndr *6*, 113-119.

Bonnal, S., Vigevani, L., and Valcarcel, J. (2012). The spliceosome as a target of novel antitumour drugs. Nat Rev Drug Discov *11*, 847-859.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. Annu Rev Biochem *50*, 349-383.

Brown, K., and Bhowmick, N.A. (2004). Linking TGF-beta-mediated Cdc25A inhibition and cytoskeletal regulation through RhoA/p160(ROCK) signaling. Cell Cycle *3*, 408-410.

Brownell, J.E., and Allis, C.D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Curr Opin Genet Dev *6*, 176-184.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

Bruniquel, D., and Schwartz, R.H. (2003). Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. Nat Immunol *4*, 235-240.

Bustamante, C., Cheng, W., and Mejia, Y.X. (2011). Revisiting the central dogma one molecule at a time. Cell *144*, 480-497.

Butler, J.E., and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev *16*, 2583-2592.

Cairns, B.R. (2005). Chromatin remodeling complexes: strength in diversity, precision through specialization. Curr Opin Genet Dev *15*, 185-190.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A., Taylor, M.S., Engstrom, P.G., Frith, M.C., *et al.* (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet *38*, 626-635.

Caterino, T.L., and Hayes, J.J. (2011). Structure of the H1 C-terminal domain and function in chromatin condensation. Biochem Cell Biol *89*, 35-44.

Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., *et al.* (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet *44*, 23-31.

Cheatham, T.E., and Kollman, P.A. (1997). Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA, and DNA:RNA hybrid duplexes. Journal of the American Chemical Society *119*, 4805-4825.

Chen, C., Nott, T.J., Jin, J., and Pawson, T. (2011). Deciphering arginine methylation: Tudor tells the tale. Nat Rev Mol Cell Biol *12*, 629-642.

Chen, D., Lucey, M.J., Phoenix, F., Lopez-Garcia, J., Hart, S.M., Losson, R., Buluwela, L., Coombes, R.C., Chambon, P., Schar, P., *et al.* (2003). T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. J Biol Chem *278*, 38586-38592.

Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet *27*, 327-331.

Chen, T., Carter, D., Garrigue-Antar, L., and Reiss, M. (1998). Transforming growth factor beta type I receptor kinase mutant associated with metastatic breast cancer. Cancer Res *58*, 4805-4810.

Chen, X., Skutt-Kakaria, K., Davison, J., Ou, Y.L., Choi, E., Malik, P., Loeb, K., Wood, B., Georges, G., Torok-Storb, B., *et al.* (2012). G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. Genes Dev *26*, 2499-2511.

Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B., and Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science *299*, 721-725.

Chinnadurai, G. (2002). CtBP, an unconventional transcriptional corepressor in development and oncogenesis. Mol Cell *9*, 213-224.

Chinnadurai, G. (2009). The transcriptional corepressor CtBP: a foe of multiple tumor suppressors. Cancer Res *69*, 731-734.

Choi, J.K. (2010). Contrasting chromatin organization of CpG islands and exons in the human genome. Genome Biol *11*, R70.

Christensen, J., Agger, K., Cloos, P.A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell *128*, 1063-1076.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

Collins, C., Rommens, J.M., Kowbel, D., Godfrey, T., Tanner, M., Hwang, S.I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K., *et al.* (1998). Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. Proc Natl Acad Sci U S A *95*, 8703-8708.

Compere, S.J., and Palmiter, R.D. (1981). DNA methylation controls the inducibility of the mouse metallothionein-I gene lymphoid cells. Cell *25*, 233-240.

Conticello, S.G., Thomas, C.J., Petersen-Mahrt, S.K., and Neuberger, M.S. (2005). Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Mol Biol Evol *22*, 367-377.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C., and Chambon, P. (1980). Promoter sequences of eukaryotic protein-coding genes. Science *209*, 1406-1414.

Cortazar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F.*, et al.* (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature *470*, 419-423.

Cortellino, S., Turner, D., Masciullo, V., Schepis, F., Albino, D., Daniel, R., Skalka, A.M., Meropol, N.J., Alberti, C., Larue, L.*, et al.* (2003). The base excision repair enzyme MED1 mediates DNA damage response to antitumor drugs and is associated with mismatch repair system integrity. Proc Natl Acad Sci U S A *100*, 15071-15076.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D.*, et al.* (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67-79.

Coulombe, B., and Burton, Z.F. (1999). DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. Microbiol Mol Biol Rev *63*, 457-478.

Cowger, J.J., Zhao, Q., Isovic, M., and Torchia, J. (2007). Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. Oncogene *26*, 3378-3386.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J.*, et al.* (2008). Structure of eukaryotic RNA polymerases. Annu Rev Biophys *37*, 337-352.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science *288*, 640-649.

Cryderman, D.E., Tang, H., Bell, C., Gilmour, D.S., and Wallrath, L.L. (1999). Heterochromatic silencing of Drosophila heat shock genes acts at the level of promoter potentiation. Nucleic Acids Res *27*, 3364-3370.

D'Alessio, J.A., Wright, K.J., and Tjian, R. (2009). Shifting players and paradigms in cell-specific transcription. Mol Cell *36*, 924-931.

Daniel, J.M., Spring, C.M., Crawford, H.C., Reynolds, A.B., and Baig, A. (2002). The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. Nucleic Acids Res *30*, 2911-2919.

Das, P.M., and Singal, R. (2004). DNA methylation and cancer. J Clin Oncol *22*, 4632-4642.

de Ruijter, A.J., van Gennip, A.H., Caron, H.N., Kemp, S., and van Kuilenburg, A.B. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J *370*, 737-749.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev *25*, 1010-1022.

Derynck, R., and Akhurst, R.J. (2007). Differentiation plasticity regulated by TGF-beta family proteins in development and disease. Nat Cell Biol *9*, 1000-1004.

Dhordain, P., Lin, R.J., Quief, S., Lantoine, D., Kerckaert, J.P., Evans, R.M., and Albagli, O. (1998). The LAZ3(BCL-6) oncoprotein recruits a SMRT/mSIN3A/histone deacetylase containing complex to mediate transcriptional repression. Nucleic Acids Res *26*, 4645-4651.

Dokmanovic, M., Clarke, C., and Marks, P.A. (2007). Histone deacetylase inhibitors: overview and perspectives. Mol Cancer Res *5*, 981-989.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873-885.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer *3*, 11-22.

Draetta, G., and Eckstein, J. (1997). Cdc25 protein phosphatases in cell proliferation. Biochim Biophys Acta *1332*, M53-63.

Duan, Z.J., Fang, X., Rohde, A., Han, H., Stamatoyannopoulos, G., and Li, Q. (2002). Developmental specificity of recruitment of TBP to the TATA box of the human gamma-globin gene. Proc Natl Acad Sci U S A *99*, 5509-5514.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res *10*, 2709-2721.

Eickbush, T.H., and Moudrianakis, E.N. (1978). The histone core complex: an octamer assembled by two sets of protein-protein interactions. Biochemistry *17*, 4955-4964.

Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene *21*, 5427-5440.

Ezhkova, E., Lien, W.H., Stokes, N., Pasolli, H.A., Silva, J.M., and Fuchs, E. (2011). EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. Genes Dev *25*, 485-498.

Fan, H.Y., Narlikar, G.J., and Kingston, R.E. (2004). Noncovalent modification of chromatin: different remodeled products with different ATPase domains. Cold Spring Harb Symp Quant Biol *69*, 183-192.

Feltus, F.A., Lee, E.K., Costello, J.F., Plass, C., and Vertino, P.M. (2003). Predicting aberrant CpG island methylation. Proc Natl Acad Sci U S A *100*, 12253-12258.

Feng, Q., Wang, H., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol *12*, 1052-1058.

Feng, X.H., Lin, X., and Derynck, R. (2000). Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-beta. EMBO J *19*, 5178-5193.

Feng, X.H., Zhang, Y., Wu, R.Y., and Derynck, R. (1998). The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for smad3 in TGF-beta-induced transcriptional activation. Genes Dev *12*, 2153-2163.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *473*, 398-402.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., *et al.* (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell *18*, 553-567.

Franklin, S., and Vondriska, T.M. (2011). Genomes, proteomes, and the central dogma. Circ Cardiovasc Genet *4*, 576.

Freudenberg, J.M., Ghosh, S., Lackford, B.L., Yellaboina, S., Zheng, X., Li, R., Cuddapah, S., Wade, P.A., Hu, G., and Jothi, R. (2012). Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic Acids Res *40*, 3364-3377.

G. WALTER, W.Z., P. PALM, E. FUCHS (1967). Initiation of DNA-Dependent RNA synthesis and the Effect of Heparin on RNA Polymerase. European J Biochem *3*, 194-201.

Gangaraju, V.K., and Bartholomew, B. (2007). Mechanisms of ATP dependent chromatin remodeling. Mutat Res *618*, 3-17.

Garcia-Huerta, P., Diaz-Hernandez, M., Delicado, E.G., Pimentel-Santillana, M., Miras-Portugal, M.T., and Gomez-Villafuertes, R. (2012). The Specificity Protein Factor Sp1 Mediates Transcriptional Regulation of P2X7 Receptors in the Nervous System. J Biol Chem.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet *7*, 703-713.

Ge, Y.Z., Pu, M.T., Gowher, H., Wu, H.P., Ding, J.P., Jeltsch, A., and Xu, G.L. (2004). Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. J Biol Chem *279*, 25447-25454.

Georgakopoulos, T., and Thireos, G. (1992). Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. EMBO J *11*, 4145-4152.

Geyer, P.K., Green, M.M., and Corces, V.G. (1990). Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. EMBO J *9*, 2247-2256.

Giannini, G., Ristori, E., Cerignoli, F., Rinaldi, C., Zani, M., Viel, A., Ottini, L., Crescenzi, M., Martinotti, S., Bignami, M.*, et al.* (2002). Human MRE11 is inactivated in mismatch repair-deficient cancers. EMBO Rep *3*, 248-254.

Giles, R.H., Peters, D.J., and Breuning, M.H. (1998). Conjunction dysfunction: CBP/p300 in human disease. Trends Genet *14*, 178-183.

Ginestier, C., Cervera, N., Finetti, P., Esteyries, S., Esterni, B., Adelaide, J., Xerri, L., Viens, P., Jacquemier, J., Charafe-Jauffret, E.*, et al.* (2006). Prognosis and gene expression profiling of 20q13-amplified breast cancers. Clin Cancer Res *12*, 4533-4544.

Giordano, J., Ge, Y., Gelfand, Y., Abrusan, G., Benson, G., and Warburton, P.E. (2007). Evolutionary history of mammalian transposons determined by genome-wide defragmentation. PLoS Comput Biol *3*, e137.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K.R., Aasland, R., Anastassiadis, K., Ang, S.L., and Stewart, A.F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development *133*, 1423-1432.

Go, C., He, W., Zhong, L., Li, P., Huang, J., Brinkley, B.R., and Wang, X.J. (2000). Aberrant cell cycle progression contributes to the early-stage accelerated carcinogenesis in transgenic epidermis expressing the dominant negative TGFbetaRII. Oncogene *19*, 3623-3631.

Goggins, M., Shekher, M., Turnacioglu, K., Yeo, C.J., Hruban, R.H., and Kern, S.E. (1998). Genetic alterations of the transforming growth factor beta receptor genes in pancreatic and biliary adenocarcinomas. Cancer Res *58*, 5329-5332.

Goodman, R.H., and Smolik, S. (2000). CBP/p300 in cell growth, transformation, and development. Genes Dev *14*, 1553-1577.

Goumans, M.J., and Mummery, C. (2000). Functional analysis of the TGFbeta receptor/Smad pathway through gene ablation in mice. Int J Dev Biol *44*, 253-265.

Grady, W.M., Myeroff, L.L., Swinler, S.E., Rajput, A., Thiagalingam, S., Lutterbaugh, J.D., Neumann, A., Brattain, M.G., Chang, J., Kim, S.J., *et al.* (1999). Mutational inactivation of transforming growth factor beta receptor type II in microsatellite stable colon cancers. Cancer Res *59*, 320-324.

Gray, M.R.B.a.I.C. (2003). Bioinformatics for Geneticist (John Wiley & Sons Canada, Ltd), pp. 293-314.

Greenwald, J., Groppe, J., Gray, P., Wiater, E., Kwiatkowski, W., Vale, W., and Choe, S. (2003). The BMP7/ActRII extracellular domain complex provides new insights into the cooperative nature of receptor assembly. Mol Cell *11*, 605-617.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet *13*, 343-357.

Gross, S., Cairns, R.A., Minden, M.D., Driggers, E.M., Bittinger, M.A., Jang, H.G., Sasaki, M., Jin, S., Schenkein, D.P., Su, S.M., *et al.* (2010). Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. J Exp Med *207*, 339-344.

Grosschedl, R., and Birnstiel, M.L. (1980). Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H2A histone gene transcription in vivo. Proc Natl Acad Sci U S A *77*, 7102-7106.

Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595-606.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet *27*, 322-326.

Hahn, S., Buratowski, S., Sharp, P.A., and Guarente, L. (1989). Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. Proc Natl Acad Sci U S A *86*, 5718-5722.

Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., *et al.* (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science *271*, 350-353.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Hansen, J.C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu Rev Biophys Biomol Struct *31*, 361-392.

Hanstein, B., Eckner, R., DiRenzo, J., Halachmi, S., Liu, H., Searcy, B., Kurokawa, R., and Brown, M. (1996). p300 is a component of an estrogen receptor coactivator complex. Proc Natl Acad Sci U S A *93*, 11540-11545.

Hashimoto, H., Hong, S., Bhagwat, A.S., Zhang, X., and Cheng, X. (2012a). Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. Nucleic Acids Res.

Hashimoto, H., Zhang, X., and Cheng, X. (2012b). Excision of thymine and 5-hydroxymethyluracil by the MBD4 DNA glycosylase domain: structural basis and implications for active DNA demethylation. Nucleic Acids Res.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983-1993.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., *et al.* (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303-1307.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.

Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol *18*, 6538-6547.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A., and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev *15*, 710-723.

Herbert, A., and Rich, A. (1999). RNA processing and the evolution of eukaryotes. Nat Genet *21*, 265-269.

Hildebrand, J.D., and Soriano, P. (2002). Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. Mol Cell Biol *22*, 5296-5307.

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. J Biol Chem *175*, 315-332.

Hsiao, P.W., Fryer, C.J., Trotter, K.W., Wang, W., and Archer, T.K. (2003). BAF60a mediates critical interactions between nuclear receptors and the BRG1 chromatin-remodeling complex for transactivation. Mol Cell Biol *23*, 6210-6220.

Hsu, C.H., Peng, K.L., Kang, M.L., Chen, Y.R., Yang, Y.C., Tsai, C.H., Chu, C.S., Jeng, Y.M., Chen, Y.T., Lin, F.M., *et al.* (2012). TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. Cell Rep *2*, 568-579.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129-1133.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300-1303.

Itoh, F., Asao, H., Sugamura, K., Heldin, C.H., ten Dijke, P., and Itoh, S. (2001). Promoting bone morphogenetic protein signaling through negative regulation of inhibitory Smads. EMBO J *20*, 4132-4142.

Iuchi, S. (2001). Three classes of C2H2 zinc finger proteins. Cell Mol Life Sci *58*, 625-635.

Jackman, J.E., and Alfonzo, J.D. (2012). Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley Interdiscip Rev RNA.

Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science *295*, 2080-2083.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., *et al.* (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Janknecht, R., Wells, N.J., and Hunter, T. (1998). TGF-beta-stimulated cooperation of smad proteins with the coactivators CBP/p300. Genes Dev *12*, 2114-2119.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S.T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol *14*, 116-127.

Jiang, H., Grenley, M.O., Bravo, M.J., Blumhagen, R.Z., and Edgar, B.A. (2011). EGFR/Ras/MAPK signaling mediates adult midgut epithelial homeostasis and regeneration in Drosophila. Cell Stem Cell *8*, 84-95.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet *13*, 484-492.

Jones, P.A., Rideout, W.M., 3rd, Shen, J.C., Spruck, C.H., and Tsai, Y.C. (1992). Methylation, mutation and cancer. Bioessays *14*, 33-36.

Jurkin, J., Zupkovitz, G., Lagger, S., Grausenburger, R., Hagelkruys, A., Kenner, L., and Seiser, C. (2011). Distinct and redundant functions of histone deacetylases HDAC1 and HDAC2 in proliferation and tumorigenesis. Cell Cycle *10*, 406-412.

Jurkowska, R.Z., Jurkowski, T.P., and Jeltsch, A. (2011). Structure and function of mammalian DNA methyltransferases. Chembiochem *12*, 206-222.

Kang, Y., Chen, C.R., and Massague, J. (2003). A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. Mol Cell *11*, 915-926.

Kang, Y., He, W., Tulley, S., Gupta, G.P., Serganova, I., Chen, C.R., Manova-Todorova, K., Blasberg, R., Gerald, W.L., and Massague, J. (2005). Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. Proc Natl Acad Sci U S A *102*, 13909-13914.

Kang, Z., Janne, O.A., and Palvimo, J.J. (2004). Coregulator recruitment and histone modifications in transcriptional regulation by the androgen receptor. Mol Endocrinol *18*, 2633-2648.

Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., and Reid, G. (2008). Transient cyclical methylation of promoter DNA. Nature *452*, 112-115.

Kato, S., Inoue, K., and Youn, M.-Y. (2010). Emergence of the osteo-epigenome in bone biology. IBMS BoneKEy *7*, 314-324.

Kavalali, E.T., Nelson, E.D., and Monteggia, L.M. (2011). Role of MeCP2, DNA methylation, and HDACs in regulating synapse function. J Neurodev Disord *3*, 250-256.

Kays, A.R., and Schepartz, A. (2000). Virtually unidirectional binding of TBP to the AdMLP TATA box within the quaternary complex with TFIIA and TFIIB. Chem Biol *7*, 601-610.

Khazak, V., Estojak, J., Cho, H., Majors, J., Sonoda, G., Testa, J.R., and Golemis, E.A. (1998). Analysis of the interaction of the novel RNA polymerase II (pol II) subunit hsRPB4 with its partner hsRPB7 and with pol II. Mol Cell Biol *18*, 1935-1945.

Kim, T.K., Ebright, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science *288*, 1418-1422.

Kim, T.K., Lagrange, T., Wang, Y.H., Griffith, J.D., Reinberg, D., and Ebright, R.H. (1997). Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. Proc Natl Acad Sci U S A *94*, 12268-12273.

Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell *77*, 599-608.

Klose, R.J., Kallin, E.M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet *7*, 715-727.

Knegtel, R.M., van Tilborg, M.A., Boelens, R., and Kaptein, R. (1995). NMR structural studies on the zinc finger domains of nuclear hormone receptors. EXS *73*, 279-295.

Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R., *et al.* (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature *468*, 839-843.

Kondo, E., Gu, Z., Horii, A., and Fukushige, S. (2005). The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes. Mol Cell Biol *25*, 4388-4396.

Kondo, Y. (2009). Epigenetic cross-talk between DNA methylation and histone modifications in human cancers. Yonsei Med J *50*, 455-463.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Krig, S.R., Jin, V.X., Bieda, M.C., O'Geen, H., Yaswen, P., Green, R., and Farnham, P.J. (2007). Identification of genes directly regulated by the oncogene ZNF217 using chromatin immunoprecipitation (ChIP)-chip assays. J Biol Chem *282*, 9703-9712.

Krig, S.R., Miller, J.K., Frietze, S., Beckett, L.A., Neve, R.M., Farnham, P.J., Yaswen, P.I., and Sweeney, C.A. (2010). ZNF217, a candidate breast cancer oncogene amplified at 20q13, regulates expression of the ErbB3 receptor tyrosine kinase in breast cancer cells. Oncogene *29*, 5500-5510.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol *12*, 283-294.

Kumar, V., Carlson, J.E., Ohgi, K.A., Edwards, T.A., Rose, D.W., Escalante, C.R., Rosenfeld, M.G., and Aggarwal, A.K. (2002). Transcription corepressor CtBP is an NAD(+)-regulated dehydrogenase. Mol Cell *10*, 857-869.

Kuppuswamy, M., Vijayalingam, S., Zhao, L.J., Zhou, Y., Subramanian, T., Ryerse, J., and Chinnadurai, G. (2008). Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex. Mol Cell Biol *28*, 269-281.

Kuras, L., Kosa, P., Mencia, M., and Struhl, K. (2000). TAF-Containing and TAF-independent forms of transcriptionally active TBP in vivo. Science *288*, 1244-1248.

Kurdistani, S.K., Robyr, D., Tavazoie, S., and Grunstein, M. (2002). Genome-wide binding map of the histone deacetylase Rpd3 in yeast. Nat Genet *31*, 248-254.

Kutach, A.K., and Kadonaga, J.T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. Mol Cell Biol *20*, 4754-4764.

Kwon, H.S., Kim, M.S., Edenberg, H.J., and Hur, M.W. (1999). Sp3 and Sp4 can repress transcription by competing with Sp1 for the core cis-elements on the human ADH5/FDH minimal promoter. J Biol Chem *274*, 20-28.

Lafontaine, L., Chaudhry, P., Lafleur, M.J., Van Themsche, C., Soares, M.J., and Asselin, E. (2011). Transforming growth factor Beta regulates proliferation and invasion of rat placental cell lines. Biol Reprod *84*, 553-559.

Laiho, M., DeCaprio, J.A., Ludlow, J.W., Livingston, D.M., and Massague, J. (1990). Growth inhibition by TGF-beta linked to suppression of retinoblastoma protein phosphorylation. Cell *62*, 175-185.

Laity, J.H., Lee, B.M., and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. Curr Opin Struct Biol *11*, 39-46.

Lall, S. (2007). Primers on chromatin. Nat Struct Mol Biol *14*, 1110-1115.

Lambowitz, A.M., and Perlman, P.S. (1990). Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. Trends Biochem Sci *15*, 440-444.

Larsson, J., Goumans, M.J., Sjostrand, L.J., van Rooijen, M.A., Ward, D., Leveen, P., Xu, X., ten Dijke, P., Mummery, C.L., and Karlsson, S. (2001). Abnormal angiogenesis but intact hematopoietic potential in TGF-beta type I receptor-deficient mice. EMBO J *20*, 1663-1673.

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., 3rd, Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K., *et al.* (2012). Landscape of somatic retrotransposition in human cancers. Science *337*, 967-971.

Lee, H.W., Dominy, B.N., and Cao, W. (2011). New family of deamination repair enzymes in uracil-DNA glycosylase superfamily. J Biol Chem *286*, 31282-31287.

Lee, J., Inoue, K., Ono, R., Ogonuki, N., Kohda, T., Kaneko-Ishino, T., Ogura, A., and Ishino, F. (2002). Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. Development *129*, 1807-1817.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432-435.

Lee, Y.C., Min, S., Gim, B.S., and Kim, Y.J. (1997). A transcriptional mediator protein that is required for activation of many RNA polymerase II promoters and is conserved from yeast to humans. Mol Cell Biol *17*, 4622-4632.

Lenhard, B., Sandelin, A., and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nat Rev Genet *13*, 233-245.

Leonardo, T.R., Schultheisz, H.L., Loring, J.F., and Laurent, L.C. (2012). The functions of microRNAs in pluripotency and reprogramming. Nat Cell Biol *14*, 1114-1121.

LeRoy, G., Orphanides, G., Lane, W.S., and Reinberg, D. (1998). Requirement of RSF and FACT for transcription of chromatin templates in vitro. Science *282*, 1900-1904.

LeRoy, G., Rickards, B., and Flint, S.J. (2008). The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. Mol Cell *30*, 51-60.

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

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

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Curr Opin Genet Dev *21*, 175-186.

Li, P., Maines-Bandiera, S., Kuo, W.L., Guan, Y., Sun, Y., Hills, M., Huang, G., Collins, C.C., Leung, P.C., Gray, J.W.*, et al.* (2007). Multiple roles of the candidate oncogene ZNF217 in ovarian epithelial neoplastic progression. Int J Cancer *120*, 1863-1873.

Liang, G., Klose, R.J., Gardner, K.E., and Zhang, Y. (2007). Yeast Jhd2p is a histone H3 Lys4 trimethyl demethylase. Nat Struct Mol Biol *14*, 243-245.

Little, S.C., and Mullins, M.C. (2006). Extracellular modulation of BMP activity in patterning the dorsoventral axis. Birth Defects Res C Embryo Today *78*, 224-242.

Littlepage, L.E., Adler, A.S., Kouros-Mehr, H., Huang, G., Chou, J., Krig, S.R., Griffith, O.L., Korkola, J.E., Qu, K., Lawson, D.A.*, et al.* (2012). The transcription factor ZNF217 is a prognostic biomarker and therapeutic target during breast cancer progression. Cancer Discov *2*, 638-651.

Liu, J.J., Chao, J.R., Jiang, M.C., Ng, S.Y., Yen, J.J., and Yang-Yen, H.F. (1995). Ras transformation results in an elevated level of cyclin D1 and acceleration of G1 progression in NIH 3T3 cells. Mol Cell Biol *15*, 3654-3663.

Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Loyola, A., LeRoy, G., Wang, Y.H., and Reinberg, D. (2001). Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription. Genes Dev *15*, 2837-2851.

Lucey, M.J., Chen, D., Lopez-Garcia, J., Hart, S.M., Phoenix, F., Al-Jehani, R., Alao, J.P., White, R., Kindle, K.B., Losson, R., *et al.* (2005). T:G mismatch-specific thymine-DNA glycosylase (TDG) as a coregulator of transcription interacts with SRC1 family members through a novel tyrosine repeat motif. Nucleic Acids Res *33*, 6393-6404.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J Biol Chem *286*, 35334-35338.

Malecova, B., Gross, P., Boyer-Guittaut, M., Yavuz, S., and Oelgeschlager, T. (2007). The initiator core promoter element antagonizes repression of TATA-directed transcription by negative cofactor NC2. J Biol Chem *282*, 24767-24776.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science *298*, 1912-1934.

Mao, X.G., Yan, M., Xue, X.Y., Zhang, X., Ren, H.G., Guo, G., Wang, P., Zhang, W., and Huo, J.L. (2011). Overexpression of ZNF217 in glioblastoma contributes to the maintenance of glioma stem cells regulated by hypoxia-inducible factors. Lab Invest *91*, 1068-1078.

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R.S., Zborowska, E., Kinzler, K.W., Vogelstein, B., *et al.* (1995). Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science *268*, 1336-1338.

Marks, P., Rifkind, R.A., Richon, V.M., Breslow, R., Miller, T., and Kelly, W.K. (2001). Histone deacetylases and cancer: causes and therapies. Nat Rev Cancer *1*, 194-202.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol *6*, 838-849.

Martin Caballero, I., Hansen, J., Leaford, D., Pollard, S., and Hendrich, B.D. (2009). The methyl-CpG binding proteins Mecp2, Mbd2 and Kaiso are dispensable for mouse embryogenesis, but play a redundant function in neural differentiation. PLoS One *4*, e4315.

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y.E. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. Science *302*, 890-893.

Massague, J. (2012). TGFbeta signalling in context. Nat Rev Mol Cell Biol *13*, 616-630.

Massague, J., Blain, S.W., and Lo, R.S. (2000). TGFbeta signaling in growth control, cancer, and heritable disorders. Cell *103*, 295-309.

Massague, J., Seoane, J., and Wotton, D. (2005). Smad transcription factors. Genes Dev *19*, 2783-2810.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501-502.

McCarty, M. (1946). Purification and Properties of Desoxyribonuclease Isolated from Beef Pancreas. J Gen Physiol *29*, 123-139.

McKenna, N.J., and O'Malley, B.W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. Cell *108*, 465-474.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436-439.

Millar, C.B., Guy, J., Sansom, O.J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P.D., Bishop, S.M., Clarke, A.R., and Bird, A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science *297*, 403-405.

Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci U S A *98*, 12902-12907.

Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell *10*, 1107-1117.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. Cell *91*, 457-466.

Mitra, A.P., Hansel, D.E., and Cote, R.J. (2012). Prognostic value of cell-cycle regulation biomarkers in bladder cancer. Semin Oncol *39*, 524-533.

Miyaki, M., Iijima, T., Konishi, M., Sakai, K., Ishii, A., Yasuno, M., Hishima, T., Koike, M., Shitara, N., Iwama, T., *et al.* (1999). Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. Oncogene *18*, 3098-3103.

Mohammad, H.P., Cai, Y., McGarvey, K.M., Easwaran, H., Van Neste, L., Ohm, J.E., O'Hagan, H.M., and Baylin, S.B. (2009). Polycomb CBX7 promotes initiation of heritable repression of genes frequently silenced with cancer-specific DNA hypermethylation. Cancer Res *69*, 6322-6330.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell *30*, 755-766.

Montgomery, R.L., Davis, C.A., Potthoff, M.J., Haberland, M., Fielitz, J., Qi, X., Hill, J.A., Richardson, J.A., and Olson, E.N. (2007). Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes Dev *21*, 1790-1802.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X.*, et al.* (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell *20*, 11-24.

Moreira, I.S., Fernandes, P.A., and Ramos, M.J. (2007). Hot spots--a review of the protein-protein interface determinant amino-acid residues. Proteins *68*, 803-812.

Munster, P.N., Troso-Sandoval, T., Rosen, N., Rifkind, R., Marks, P.A., and Richon, V.M. (2001). The histone deacetylase inhibitor suberoylanilide hydroxamic acid induces differentiation of human breast cancer cells. Cancer Res *61*, 8492-8497.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell *102*, 553-563.

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., and Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem *274*, 18470-18476.

Murray, K. (1964). The Occurrence of Epsilon-N-Methyl Lysine in Histones. Biochemistry *3*, 10-15.

Myers, L.C., and Kornberg, R.D. (2000). Mediator of transcriptional regulation. Annu Rev Biochem *69*, 729-749.

Nakamoto, T. (2009). Evolution and the universality of the mechanism of initiation of protein synthesis. Gene *432*, 1-6.

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

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Narlikar, G.J., Fan, H.Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475-487.

Nielsen, K.H., and Staley, J.P. (2012). Spliceosome activation: U4 is the path, stem I is the goal, and Prp8 is the keeper. Let's cheer for the ATPase Brr2! Genes Dev *26*, 2461-2467.

Nishihara, A., Hanai, J.I., Okamoto, N., Yanagisawa, J., Kato, S., Miyazono, K., and Kawabata, M. (1998). Role of p300, a transcriptional coactivator, in signalling of TGF-beta. Genes Cells *3*, 613-623.

Nonet, G.H., Stampfer, M.R., Chin, K., Gray, J.W., Collins, C.C., and Yaswen, P. (2001). The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. Cancer Res *61*, 1250-1254.

Nunez, N., Clifton, M.M., Funnell, A.P., Artuz, C., Hallal, S., Quinlan, K.G., Font, J., Vandevenne, M., Setiyaputra, S., Pearson, R.C., *et al.* (2011). The multi-zinc finger protein ZNF217 contacts DNA through a two-finger domain. J Biol Chem *286*, 38190-38201.

O'Connor, M.B., Umulis, D., Othmer, H.G., and Blair, S.S. (2006). Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. Development *133*, 183-193.

Oh, S., and Janknecht, R. (2012). Histone demethylase JMJD5 is essential for embryonic development. Biochem Biophys Res Commun *420*, 61-65.

Ohler, U., Liao, G.C., Niemann, H., and Rubin, G.M. (2002). Computational analysis of core promoters in the Drosophila genome. Genome Biol *3*, RESEARCH0087.

Oike, Y., Takakura, N., Hata, A., Kaname, T., Akizuki, M., Yamaguchi, Y., Yasue, H., Araki, K., Yamamura, K., and Suda, T. (1999). Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. Blood *93*, 2771-2779.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Orford, K., Kharchenko, P., Lai, W., Dao, M.C., Worhunsky, D.J., Ferro, A., Janzen, V., Park, P.J., and Scadden, D.T. (2008). Differential H3K4 methylation identifies developmentally poised hematopoietic genes. Dev Cell *14*, 798-809.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev *10*, 2657-2683.

Oshaben, K.M., Salari, R., McCaslin, D.R., Chong, L.T., and Horne, W.S. (2012). The Native GCN4 Leucine-Zipper Domain Does Not Uniquely Specify a Dimeric Oligomerization State. Biochemistry *51*, 9581-9591.

Oshima, M., Oshima, H., and Taketo, M.M. (1996). TGF-beta receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis. Dev Biol *179*, 297-302.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr Biol *10*, 475-478.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. EMBO Rep *10*, 1235-1241.

Paik, W.K., Paik, D.C., and Kim, S. (2007). Historical review: the field of protein methylation. Trends Biochem Sci *32*, 146-152.

Pasche, B., Kolachana, P., Nafa, K., Satagopan, J., Chen, Y.G., Lo, R.S., Brener, D., Yang, D., Kirstein, L., Oddoux, C., *et al.* (1999). TbetaR-I(6A) is a candidate tumor susceptibility allele. Cancer Res *59*, 5678-5682.

Patel, A., Dharmarajan, V., Vought, V.E., and Cosgrove, M.S. (2009). On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. J Biol Chem *284*, 24242-24256.

Penn, N.W., Suwalski, R., O'Riley, C., Bojanowski, K., and Yura, R. (1972). The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. Biochem J *126*, 781-790.

Peserico, A., and Simone, C. (2011). Physical and functional HAT/HDAC interplay regulates protein acetylation balance. J Biomed Biotechnol *2011*, 371832.

Petrykowska, H.M., Vockley, C.M., and Elnitski, L. (2008). Detection and characterization of silencers and enhancer-blockers in the greater CFTR locus. Genome Res *18*, 1238-1246.

Planas-Silva, M.D., and Weinberg, R.A. (1997). The restriction point and control of cell proliferation. Curr Opin Cell Biol *9*, 768-772.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101-1105.

Pouponnot, C., Jayaraman, L., and Massague, J. (1998). Physical and functional interaction of SMADs and p300/CBP. J Biol Chem *273*, 22865-22868.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434-438.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. Nature *386*, 569-577.

Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat Struct Biol *9*, 217-224.

Quinlan, K.G., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., Bolognesi, M., and Crossley, M. (2006a). Specific recognition of ZNF217 and other zinc finger proteins at a surface groove of C-terminal binding proteins. Mol Cell Biol *26*, 8159-8172.

Quinlan, K.G., Verger, A., Kwok, A., Lee, S.H., Perdomo, J., Nardini, M., Bolognesi, M., and Crossley, M. (2006b). Role of the C-terminal binding protein PXDLS motif binding cleft in protein interactions and transcriptional repression. Mol Cell Biol *26*, 8202-8213.

Quinlan, K.G., Verger, A., Yaswen, P., and Crossley, M. (2007). Amplification of zinc finger gene 217 (ZNF217) and cancer: when good fingers go bad. Biochim Biophys Acta *1775*, 333-340.

Raha, T., Cheng, S.W., and Green, M.R. (2005). HIV-1 Tat stimulates transcription complex assembly through recruitment of TBP in the absence of TAFs. PLoS Biol *3*, e44.

Rahman, M.T., Nakayama, K., Rahman, M., Katagiri, H., Katagiri, A., Ishibashi, T., Ishikawa, M., Iida, K., Nakayama, N., Otsuki, Y., *et al.* (2012). Gene amplification of ZNF217 located at chr20q13.2 is associated with lymph node metastasis in ovarian clear cell carcinoma. Anticancer Res *32*, 3091-3095.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201-1212.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Reeder, R.H. (1990). rRNA synthesis in the nucleolus. Trends Genet *6*, 390-395.

Reid, J.B., and Ross, J.J. (2011). Mendel's genes: toward a full molecular characterization. Genetics *189*, 3-10.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089-1093.

Reiss, M. (1997). Transforming growth factor-beta and cancer: a love-hate relationship? Oncol Res *9*, 447-457.

Riccio, A., Aaltonen, L.A., Godwin, A.K., Loukola, A., Percesepe, A., Salovaara, R., Masciullo, V., Genuardi, M., Paravatou-Petsotas, M., Bassi, D.E., *et al.* (1999). The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. Nat Genet *23*, 266-268.

Richly, H., Aloia, L., and Di Croce, L. (2011). Roles of the Polycomb group proteins in stem cells and cancer. Cell Death Dis *2*, e204.

Robertson, A.B., Klungland, A., Rognes, T., and Leiros, I. (2009). DNA repair in mammalian cells: Base excision repair: the long and short of it. Cell Mol Life Sci *66*, 981-993.

Rodrigo-Brenni, M.C., and Hegde, R.S. (2012). Design principles of protein biosynthesis-coupled quality control. Dev Cell *23*, 896-907.

Rojas, J.R., Trievel, R.C., Zhou, J., Mo, Y., Li, X., Berger, S.L., Allis, C.D., and Marmorstein, R. (1999). Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. Nature *401*, 93-98.

Ruiz-Garcia, A.B., Sendra, R., Pamblanco, M., and Tordera, V. (1997). Gcn5p is involved in the acetylation of histone H3 in nucleosomes. FEBS Lett *403*, 186-190.

Russ, B.E., Denton, A.E., Hatton, L., Croom, H., Olson, M.R., and Turner, S.J. (2012). Defining the molecular blueprint that drives CD8(+) T cell differentiation in response to infection. Front Immunol *3*, 371.

Sainsbury, S., Niesser, J., and Cramer, P. (2012). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. Nature.

Sanchez, R., and Zhou, M.M. (2009). The role of human bromodomains in chromatin biology and gene transcription. Curr Opin Drug Discov Devel *12*, 659-665.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407-411.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol Cell *15*, 595-605.

Sartorelli, V., Puri, P.L., Hamamori, Y., Ogryzko, V., Chung, G., Nakatani, Y., Wang, J.Y., and Kedes, L. (1999). Acetylation of MyoD directed by PCAF is necessary for the execution of the muscle program. Mol Cell *4*, 725-734.

Sawan, C., and Herceg, Z. (2010). Histone modifications and cancer. Adv Genet *70*, 57-85.

Scafe, C., Chao, D., Lopes, J., Hirsch, J.P., Henry, S., and Young, R.A. (1990). RNA polymerase II C-terminal repeat influences response to transcriptional enhancer signals. Nature *347*, 491-494.

Schaeper, U., Boyd, J.M., Verma, S., Uhlmann, E., Subramanian, T., and Chinnadurai, G. (1995). Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation. Proc Natl Acad Sci U S A *92*, 10467-10471.

Schmitz, S.U., Albert, M., Malatesta, M., Morey, L., Johansen, J.V., Bak, M., Tommerup, N., Abarrategui, I., and Helin, K. (2011). Jarid1b targets genes regulating development and is involved in neural differentiation. EMBO J *30*, 4586-4600.

Schoenfelder, S., Clay, I., and Fraser, P. (2010). The transcriptional interactome: gene expression in 3D. Curr Opin Genet Dev *20*, 127-133.

Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., and Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res *28*, 231-234.

Sekine, S., Tagami, S., and Yokoyama, S. (2012). Structural basis of transcription by bacterial and eukaryotic RNA polymerases. Curr Opin Struct Biol *22*, 110-118.

Seo, K.W., Lee, S.K., Bhandari, D.R., Park, S.B., Roh, K.H., Yang, S.R., Kang, S.K., and Kang, K.S. (2012). ZNF 281 knockdown induced osteogenic differentiation of human multipotent stem cells in vivo and in vitro. Cell Transplant.

Severson, P.L., Tokar, E.J., Vrba, L., Waalkes, M.P., and Futscher, B. (2012). Agglomerates of aberrant DNA methylation are associated with toxicant-induced malignant transformation. Epigenetics *7*.

Sewalt, R.G., Gunster, M.J., van der Vlag, J., Satijn, D.P., and Otte, A.P. (1999). C-Terminal binding protein is a transcriptional repressor that interacts with a specific class of vertebrate Polycomb proteins. Mol Cell Biol *19*, 777-787.

Sharon, N., Mor, I., Zahavi, E., and Benvenisty, N. (2012). DUXO, a novel double homeobox transcription factor, is a regulator of the gastrula organizer in human embryonic stem cells. Stem Cell Res *9*, 261-269.

Shen, W.C., and Green, M.R. (1997). Yeast TAF(II)145 functions as a core promoter selectivity factor, not a general coactivator. Cell *90*, 615-624.

Shen, X., and Gorovsky, M.A. (1996). Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell *86*, 475-483.

Sherr, C.J., and Roberts, J.M. (1995). Inhibitors of mammalian G1 cyclin-dependent kinases. Genes Dev *9*, 1149-1163.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., and Casero, R.A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shi, Y., Lee, J.S., and Galvin, K.M. (1997). Everything you have ever wanted to know about Yin Yang 1. Biochim Biophys Acta *1332*, F49-66.

Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell *113*, 685-700.

Shi, Y., Sawada, J., Sui, G., Affar el, B., Whetstine, J.R., Lan, F., Ogawa, H., Luke, M.P., and Nakatani, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature *422*, 735-738.

Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell *19*, 857-864.

Shirohzu, H., Kubota, T., Kumazawa, A., Sado, T., Chijiwa, T., Inagaki, K., Suetake, I., Tajima, S., Wakui, K., Miki, Y., *et al.* (2002). Three novel DNMT3B mutations in Japanese patients with ICF syndrome. Am J Med Genet *112*, 31-37.

Siegfried, Z., and Cedar, H. (1997). DNA methylation: a molecular lock. Curr Biol *7*, R305-307.

Siles, L., Sanchez-Tillo, E., Lim, J.W., Darling, D.S., Kroll, K.L., and Postigo, A. (2013). ZEB1 imposes a temporary stage-dependent inhibition of muscle gene expression and differentiation via CtBP-mediated transcriptional repression. Mol Cell Biol.

Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A., and Arndt, K.M. (2003). Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. EMBO J *22*, 1846-1856.

Simpson, R.T. (1978). Structure of chromatin containing extensively acetylated H3 and H4. Cell *13*, 691-699.

Sims, R.J., 3rd, and Reinberg, D. (2006). Histone H3 Lys 4 methylation: caught in a bind? Genes Dev *20*, 2779-2786.

Singh, U., Rogers, J.B., Mann, B.J., and Petri, W.A., Jr. (1997). Transcription initiation is controlled by three core promoter elements in the hgl5 gene of the protozoan parasite Entamoeba histolytica. Proc Natl Acad Sci U S A *94*, 8812-8817.

Sjolund, A.B., Senejani, A.G., and Sweasy, J.B. (2012). MBD4 and TDG: Multifaceted DNA glycosylases with ever expanding biological roles. Mutat Res.

Smale, S.T. (2001). Core promoters: active contributors to combinatorial gene regulation. Genes Dev *15*, 2503-2508.

Smith, H.C., Bennett, R.P., Kizilyer, A., McDougall, W.M., and Prohaska, K.M. (2012). Functions and regulation of the APOBEC family of proteins. Semin Cell Dev Biol *23*, 258-268.

Song, G., Ouyang, G., and Bao, S. (2005). The activation of Akt/PKB signaling pathway and cell survival. J Cell Mol Med *9*, 59-71.

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W., Bauer, C., Munzel, M., Wagner, M., Muller, M., Khan, F.*, et al.* (2013). Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell *152*, 1146-1159.

Steinberg, X.P., Hepp, M.I., Fernandez Garcia, Y., Suganuma, T., Swanson, S.K., Washburn, M., Workman, J.L., and Gutierrez, J.L. (2012). Human CCAAT/enhancer-binding protein beta interacts with chromatin remodeling complexes of the imitation switch subfamily. Biochemistry *51*, 952-962.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev *12*, 599-606.

Su, C., Deaton, R.A., Iglewsky, M.A., Valencia, T.G., and Grant, S.R. (2007). PKN activation via transforming growth factor-beta 1 (TGF-beta 1) receptor signaling delays G2/M phase transition in vascular smooth muscle cells. Cell Cycle *6*, 739-749.

Sun, G., Fu, C., Shen, C., and Shi, Y. (2011). Histone deacetylases in neural stem cells and induced pluripotent stem cells. J Biomed Biotechnol *2011*, 835968.

Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y.*, et al.* (2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Res *11*, 677-684.

Szerlong, H.J., and Hansen, J.C. (2011). Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. Biochem Cell Biol *89*, 24-34.

Szwagierczak, A., Bultmann, S., Schmidt, C.S., Spada, F., and Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res *38*, e181.

Takahashi, Y.H., Westfield, G.H., Oleskie, A.N., Trievel, R.C., Shilatifard, A., and Skiniotis, G. (2011). Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. Proc Natl Acad Sci U S A *108*, 20526-20531.

Takeuchi, T., Watanabe, Y., Takano-Shimizu, T., and Kondo, S. (2006). Roles of jumonji and jumonji family genes in chromatin regulation and development. Dev Dyn *235*, 2449-2459.

Tan, L., and Shi, Y.G. (2012). Tet family proteins and 5-hydroxymethylcytosine in development and disease. Development *139*, 1895-1902.

Tang, B., Bottinger, E.P., Jakowlew, S.B., Bagnall, K.M., Mariano, J., Anver, M.R., Letterio, J.J., and Wakefield, L.M. (1998). Transforming growth factor-beta1 is a new form of tumor suppressor with true haploid insufficiency. Nat Med *4*, 802-807.

Tanner, M.M., Tirkkonen, M., Kallioniemi, A., Holli, K., Collins, C., Kowbel, D., Gray, J.W., Kallioniemi, O.P., and Isola, J. (1995). Amplification of chromosomal region 20q13 in invasive breast cancer: prognostic implications. Clin Cancer Res *1*, 1455-1461.

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

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Tefferi, A., Lim, K.H., and Levine, R. (2009). Mutation in TET2 in myeloid cancers. N Engl J Med *361*, 1117; author reply 1117-1118.

Teng, I.W., Hou, P.C., Lee, K.D., Chu, P.Y., Yeh, K.T., Jin, V.X., Tseng, M.J., Tsai, S.J., Chang, Y.S., Wu, C.S., *et al.* (2011). Targeted methylation of two tumor suppressor genes is sufficient to transform mesenchymal stem cells into cancer stem/initiating cells. Cancer Res *71*, 4653-4663.

Tini, M., Benecke, A., Um, S.J., Torchia, J., Evans, R.M., and Chambon, P. (2002). Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell *9*, 265-277.

Tupler, R., Perini, G., and Green, M.R. (2001). Expressing the human genome. Nature *409*, 832-833.

Turner, B.M. (1993). Decoding the nucleosome. Cell *75*, 5-8.

Turner, B.M. (2000). Histone acetylation and an epigenetic code. Bioessays *22*, 836-845.

Turner, B.M. (2007). Defining an epigenetic code. Nat Cell Biol *9*, 2-6.

Um, S., Harbers, M., Benecke, A., Pierrat, B., Losson, R., and Chambon, P. (1998). Retinoic acid receptors interact physically and functionally with the T:G mismatch-specific thymine-DNA glycosylase. J Biol Chem *273*, 20728-20736.

Urnov, F.D., and Wolffe, A.P. (2001). Chromatin remodeling and transcriptional activation: the cast (in order of appearance). Oncogene *20*, 2991-3006.

Vandevenne, M., Jacques, D.A., Artuz, C., Nguyen, C.D., Kwan, A.H., Segal, D.J., Matthews, J.M., Crossley, M., Guss, J.M., and Mackay, J.P. (2013). New insights into

DNA recognition by zinc fingers revealed by structural analysis of ZNF217. J Biol Chem.

Vendrell, J.A., Thollet, A., Nguyen, N.T., Ghayad, S.E., Vinot, S., Bieche, I., Grisard, E., Josserand, V., Coll, J.L., Roux, P., *et al.* (2012). ZNF217 is a marker of poor prognosis in breast cancer that drives epithelial-mesenchymal transition and invasion. Cancer Res *72*, 3593-3606.

Vermeulen, K., Van Bockstaele, D.R., and Berneman, Z.N. (2003). The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Prolif *36*, 131-149.

Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. Curr Biol *8*, 96-108.

Verzi, M.P., Shin, H., San Roman, A.K., Liu, X.S., and Shivdasani, R.A. (2012). Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. Mol Cell Biol.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., *et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Wade, J.T., and Struhl, K. (2008). The transition from transcriptional initiation to elongation. Curr Opin Genet Dev *18*, 130-136.

Wang, D., Kanuma, T., Mizunuma, H., Takama, F., Ibuki, Y., Wake, N., Mogi, A., Shitara, Y., and Takenoshita, S. (2000). Analysis of specific gene mutations in the transforming growth factor-beta signal transduction pathway in human ovarian cancer. Cancer Res *60*, 4507-4512.

Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S., and Liu, J.M. (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 U S A *95*, 10860-10865.

Wang, J., Scully, K., Zhu, X., Cai, L., Zhang, J., Prefontaine, G.G., Krones, A., Ohgi, K.A., Zhu, P., Garcia-Bassets, I., *et al.* (2007). Opposing LSD1 complexes function in developmental gene activation and repression programmes. Nature *446*, 882-887.

Wang, Y., Zhang, H., Chen, Y., Sun, Y., Yang, F., Yu, W., Liang, J., Sun, L., Yang, X., Shi, L., *et al.* (2009). LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. Cell *138*, 660-672.

Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Coller, H.A., Cross, J.R., Fantin, V.R., Hedvat, C.V., Perl, A.E., *et al.* (2010). The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell *17*, 225-234.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738.

Watson, J.D., and Crick, F.H. (2003). A structure for deoxyribose nucleic acid. 1953. Nature *421*, 397-398; discussion 396.

Weinberg, R.A. (1996). E2F and cell proliferation: a world turned upside down. Cell *85*, 457-459.

Weinstein, M., Yang, X., Li, C., Xu, X., Gotay, J., and Deng, C.X. (1998). Failure of egg cylinder elongation and mesoderm induction in mouse embryos lacking the tumor suppressor smad2. Proc Natl Acad Sci U S A *95*, 9378-9383.

Wiegand, R.C., and Brutlag, D.L. (1981). Histone acetylase from Drosophila melanogaster specific for H4. J Biol Chem *256*, 4578-4583.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343-348.

Wissmann, M., Yin, N., Muller, J.M., Greschik, H., Fodor, B.D., Jenuwein, T., Vogler, C., Schneider, R., Gunther, T., Buettner, R.*, et al.* (2007). Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. Nat Cell Biol *9*, 347-353.

Wolf, S.S. (2009). The protein arginine methyltransferase family: an update about function, new perspectives and the physiological role in humans. Cell Mol Life Sci *66*, 2109-2121.

Wong, E., Yang, K., Kuraguchi, M., Werling, U., Avdievich, E., Fan, K., Fazzari, M., Jin, B., Brown, A.M., Lipkin, M.*, et al.* (2002). Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. Proc Natl Acad Sci U S A *99*, 14937-14942.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun *2*, 241.

Wu, J.I. (2012). Diverse functions of ATP-dependent chromatin remodeling complexes in development and cancer. Acta Biochim Biophys Sin (Shanghai) *44*, 54-69.

Wysocka, J., Milne, T.A., and Allis, C.D. (2005a). Taking LSD 1 to a new high. Cell *122*, 654-658.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005b). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859-872.

Xie, H., Wang, M., Bonaldo Mde, F., Smith, C., Rajaram, V., Goldman, S., Tomita, T., and Soares, M.B. (2009). High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. Nucleic Acids Res *37*, 4331-4340.

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

Xu, X., Brodie, S.G., Yang, X., Im, Y.H., Parks, W.T., Chen, L., Zhou, Y.X., Weinstein, M., Kim, S.J., and Deng, C.X. (2000). Haploid loss of the tumor suppressor Smad4/Dpc4 initiates gastric polyposis and cancer in mice. Oncogene *19*, 1868-1874.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., *et al.* (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell *42*, 451-464.

Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J.G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T., *et al.* (2012). Tet3 CXXC Domain and Dioxygenase Activity Cooperatively Regulate Key Genes for Xenopus Eye and Neural Development. Cell *151*, 1200-1213.

Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G.J., *et al.* (2009). IDH1 and IDH2 mutations in gliomas. N Engl J Med *360*, 765-773.

Yang, X.J., and Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. Nat Rev Mol Cell Biol *9*, 206-218.

Yao, T.P., Oh, S.P., Fuchs, M., Zhou, N.D., Ch'ng, L.E., Newsome, D., Bronson, R.T., Li, E., Livingston, D.M., and Eckner, R. (1998). Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. Cell *93*, 361-372.

You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc Natl Acad Sci U S A *98*, 1454-1458.

Young, N.L., Dimaggio, P.A., and Garcia, B.A. (2010). The significance, development and progress of high-throughput combinatorial histone code analysis. Cell Mol Life Sci *67*, 3983-4000.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.

Zee, B.M., Levin, R.S., Xu, B., LeRoy, G., Wingreen, N.S., and Garcia, B.A. (2010). In vivo residue-specific histone methylation dynamics. J Biol Chem *285*, 3341-3350.

Zetterberg, A., Larsson, O., and Wiman, K.G. (1995). What is the restriction point? Curr Opin Cell Biol *7*, 835-842.

Zhang, C.L., McKinsey, T.A., Lu, J.R., and Olson, E.N. (2001). Association of COOH-terminal-binding protein (CtBP) and MEF2-interacting transcription repressor (MITR) contributes to transcriptional repression of the MEF2 transcription factor. J Biol Chem *276*, 35-39.

Zhang, X., Wen, H., and Shi, X. (2012). Lysine methylation: beyond histones. Acta Biochim Biophys Sin (Shanghai) *44*, 14-27.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., *et al.* (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res *38*, 4246-4253.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev *15*, 2343-2360.

Zhang, Y.E. (2009). Non-Smad pathways in TGF-beta signaling. Cell Res *19*, 128-139.

Zhao, L., Gao, L., Wang, H., Chen, X., Wang, Y., Yang, H., Wei, C., Wan, X., and Xia, T. (2012). The R2R3-MYB, bHLH, WD40, and related transcription factors in flavonoid biosynthesis. Funct Integr Genomics.

Zhao, X., Ueba, T., Christie, B.R., Barkho, B., McConnell, M.J., Nakashima, K., Lein, E.S., Eadie, B.D., Willhoite, A.R., Muotri, A.R., *et al.* (2003). Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. Proc Natl Acad Sci U S A *100*, 6777-6782.

Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet *43*, 143-166.

Chapter 2: Genome Analysis Identifies the $p15^{ink4b}$ Tumor Suppressor as a Direct Target of the ZNF217/CoREST Complex

2.1 Introduction

ZNF217 is a candidate oncogene found at the core of the 20q13.2 amplicon (Collins et al., 1998). Amplification and overexpression of ZNF217 has been found in a significant proportion of tumours and transformed cell types of epithelial origin including those of the breast, colorectal, ovarian and prostate, ranging in amplification frequency from 10 to 40% (Suzuki et al., 2004; Tanner et al., 2000; Watanabe et al., 2002; Weiss et al., 2003). Several studies have also established correlations between amplification and overexpression of ZNF217 and clinical outcome, with increased ZNF217 expression correlating with a poor prognosis (Bar-Shira et al., 2002; Chin et al., 2006; Hidaka et al., 2000; Tanner et al., 2000; Weiss et al., 2003).

Insights into the oncogenic role of ZNF217 have come primarily from studies using finite lifespan human epithelial cells. Forced expression of ZNF217 by retroviral gene transfer in human mammary epithelial cells promotes loss of senescence, immortalization, and resistance to growth inhibition by TGFβ (Nonet et al., 2001). In addition, prolonged growth of ZNF217-immortalized cells display chromosomal instability, as well as telomere crisis and telomerase reactivation (Chin et al., 2004). More recently, transduction of finite lifespan ovarian cells with a ZNF217 retrovirus has also been shown to promote cellular immortalization, increased cellular proliferation and telomerase activity, as well as anchorage independent cell growth (Kwon et al., 2007). ZNF217 overexpression was also shown to suppress spontaneous and doxorubicin-induced apoptosis, suggesting that ZNF217 may promote oncogenic transformation by increasing cell survival (Huang et al., 2005).

The ZNF217 protein contains 8 Kruppel-like zinc fingers suggesting that it most likely functions as a transcription factor. Molecular mapping studies using a transcription-based reporter assay and various regions of ZNF217 fused to the GAL4 DNA binding domain have identified two repression domains located within the carboxy terminus (Cowger et al., 2006; Quinlan et al., 2006). Biochemical purification studies in combination with mass spectrometry have identified ZNF217 as a constituent of several-related transcriptional repressor complexes (Cowger et al., 2006; Hakimi et al., 2003; Lee et al., 2005; Shi et al., 2004; Shi et al., 2005; You et al., 2001). Comparative analysis of each of the purified complexes suggests that the ZNF217 complex is very similar to the CoREST complex previously implicated in neuronal differentiation (You et al., 2001), and consists of 3 core proteins: the histone deacetylase 2, the lysine demethylase I (LSD1) and the Corepressor of Rest (CoREST). In addition, the carboxy terminus of ZNF217 interacts directly with the CtBP 1/2 corepressor and this interaction is, in part, essential for the repressor function of ZNF217 (Cowger et al., 2006; Quinlan et al., 2006).

CASTing (Cyclic amplification and selection of targets) analysis, using degenerate oligonucleotides and the region of ZNF217 encompassing the sixth and seventh zinc fingers has identified a core recognition sequence consisting of CAGAAY (where Y is A, G orT) (Cowger et al., 2006). This sequence has been identified within the E-cadherin promoter and chromatin immunoprecipitation (ChIP) assays have shown that the ZNF217 complex is present on the E-cadherin promoter in breast cancer cells (Cowger et al., 2006). More recently, a bioinformatics approach in conjunction with ChIP analysis was used to identify a consensus ZNF217 binding site (ATTCNAC), in ZNF217

target genes. Interestingly, 65% of the genes identified in the ChIP screen also contained the CAGAAY motif suggesting that ZNF217 may use multiple zinc fingers to bind specific target genes (Krig et al., 2007).

In the present study, we have used a two-step approach to identify ZNF217 targets. First, we employed siRNA-mediated gene silencing of ZNF217 coupled with microarray screening to identify genes with altered expression. Secondly, we used chromatin immunoprecipitation with directed selection and ligation (ChIP-DSL) to identify promoters directly bound by ZNF217. By comparative analysis of genes identified using both approaches, we have identified a subset of genes directly regulated by the ZNF217 complex. Our analysis has focused on the $p15^{ink4b}$ tumour suppressor gene as a critical ZNF217 target. ChIP analysis in both MCF7 and HaCAT cells confirmed that the ZNF217 complex occupies a region of p15ink4B promoter that is critically important for transcriptional activation. Stimulation of HaCAT cells with TGFβ resulted in a rapid release of ZNF217 and a concomitant recruitment of SMAD2 protein to the $p15^{ink4b}$ gene promoter, which preceded increases in protein expression. Importantly, ZNF217 downregulation and TGFβ stimulation have similar affects on the chromatin modifications surrounding the $p15^{ink4b}$ promoter suggesting that ZNF217 and TGFβ are functioning through convergent mechanisms. Our results suggest that a coactivator/corepressor balance may constitute an important parameter regulating $p15^{ink4b}$ expression and establishes a possible link between overexpression of ZNF217 and the loss of TGFβ stimulation at selected targets.

2.2 Materials and Methods

2.2.1 *Plasmids, antibodies, reagents and culture conditions.*

The affinity purified anti-ZNF217 antibody was generated as previously described (Cowger et al., 2006). A complete list of primers used can be found in Table 2.1. Antibodies used in this study are listed in Table 2.2. TGFβ was purchased from R&D systems. The siRNA was purchased from Dharmacon. ZNF217 siRNA1: GCAAAUAACCUCAUCUGUAUU, ZNF217 siRNA2: GAACAGAACCUCCCAAGG AUU. Control siRNA: scrambled pool siRNA (sequence not revealed by manufacturer). MCF7 (breast cancer cells) and HaCAT cells (Immortilized skin keratinocytes) were grown in humid 37°C incubators containing 5% CO_2. HaCAT cells were grown to approximately 75% percent confluence and were passaged once every 3 days. MCF7 cells were grown to full confluence and were passaged once every 3 days.

2.2.2 *RNA isolation and real-time PCR.*

Total cellular RNA was isolated using RNA EZ kit (Qiagen). The quality and quantity of RNA were evaluated by measuring OD 260/280. In addition, RNA quality was evaluated by agarose gel electrophoresis. RNA knockdown was conducted using Oligofectamine (Invitrogen Cat # 12252-011) according to manufacturer's recommendations. For real-time PCR analysis, 0.2 μg of RNA was reverse-transcribed with TaqMan reverse transcriptase (Applied Biosystems) using random hexamers to generate cDNA. All amplicons were detected using the 5' nuclease (Taqman) assay with 5' TAMRA- labelled probes. Probes were already predesigned and quality tested (Applied Biosystems). Reactions were performed according to the manufacturer's recommendations (Applied biosystems) and were run in replicates of two, in a 96-well

Table 2.1: Primers used in ChIP experiments. Primer sequences are derived from 1Kb upstream region of genes.

Gene	Forward Primer 5' - 3'	Reverse Primer 5' - 3'	Region amplified
ABCA12	GAGCCTGAAACAATGTGG	CTATGTTGGCCAGGTCC	-769/-588
EFNB2	GTCGCTGTTTCCACGTC	CGGGCTGACAGGTGAGC	-848/-717
GJA1	CTCTCTAGTGGGCTTGAG	CCATGTCTCCAGAAAACTAAG	-934/-776
NOTCH1	GCGCCAGCGGCAGATC	TCGCGGACGGATTGTGC	-738/-509
VAV3	GCGCAAAAGTTCTGGGG	GCCGTTGCTGTTCTGG	-632/-536
RAC3	GGCGACTGTTGGTGGTGT	AACGCGCTGTATTTCCAAAC	-632/-484
SF3A1	TGGTGAAACCCCGTCTCTAC	GCGATCTTGACTCTCTGCAA	-583/-451
MAN1A1	GGCGGGAGAGACATACAAGT	GAAGCACGGCTTTACTCCAG	-693/-505
MCM8	TGGACGGCCAGATATGAAAT	GTGCCATTCTTGGCTCTCTC	-996/-870
CDC25C	TCCCAAAGTGCTGGGATTAC	AATTCCGTTGCAGGGAAAG	-791/-638
$p15^{ink4b}$ Standard PCR			
Primer Set 4	CTCGGTCACAAGGGAGC	AATGCTGGCTGCACTGC	-359/-203
Primer Set 3	GCACACGCAAAACATGATTC	GCGACAGCTCTGCACC	-578/-420
Primer Set 1	CCTAGGAAGATTAGGAAGG	CCCACTTTGTCAGGTATC	-950/-779
Realtime PCR			
Primer Set 4	CCAACGTCTCCACAGTGAAA	AATGCTGGCTGCACTGCT	-340/-204
Primer Set 3	CATGATTCTCGGGATTTTTCTC	GACAGCTCTGCACCTGTCAT	-566/-426
Primer Set 2	CCTGACAAAGTGGGTTTAAATAGG	GAATCATGTTTTGCGTGTGC	-775/-578
Primer Set 1	AGGAAGATTAGGAAGGGGAAA	CCCACTTTGTCAGGTATCTTATTTT	-947/-786

Table 2.2: Antibodies used for immunoblotting and chromatin immunoprecipitation.

Antibodies	Company	Catalogue Number
EHF	Lab Vision	RT-1912
MAP2	Santa Cruz	sc-20172
$p15^{INK4B}$	Santa Cruz	sc-612
EFNB2	Santa Cruz	sc-15397
GJA1	Sigma	c6219
LSD1	Bethyl Laboratories	A300-216A
CoREST	Bethyl Laboratories	A300-130A
SMAD2/3	Santa Cruz	sc-6032
Anti-Acetyl-Lysine 9/14 H3	Upstate	06-599
Anti-dimethyl-Lysine 4 H3	Upstate	05-790
Anti-trimethyl-Lysine 27 H3	Upstate	07-449

format. Each reaction included 18S RNA as a control for normalization, and reactions lacking cDNA served as negative controls. Two independent experiments were performed for each gene, and a mean value was obtained and compared to the mean expression level of each gene from cells transfected with control siRNA. Applied Biosystems 7500 Real Time PCR System software was used to identify cycle threshold (Ct) for each reaction.

2.2.3 *RNA microarray analysis.*

Total RNA was extracted from MCF7 cells transfected with ZNF217 siRNA or mock. Independent biological triplicates were performed for each siRNA, and including control transfections, brought the number of independent transfection experiments to nine. cDNA was prepared from control and each siRNA-transfected sample, labelled and hybridized to HgU133A + 2 human affymetrix DNA microarray and a list of genes was then created for all 9 experiments. The hybridization, washing, scanning and analysis of genechips were performed at the University of Western Ontario, Robarts Genomic Centre (London, Ontario, Canada).

An average intensity of siRNA knockdowns (RNAi (1 and 2) was compared to control non treated sample. Three biological replicates were done for each array and the data was transformed using Robust Multi-Array normalization (Bolstad et al., 2003) and values below 0.01 were set to 0.01. Each measurement was normalized by dividing all measurements in that sample by the 50^{th} percentile. Ratios were then calculated for all 9 samples against the median of the control samples (1, 4, and 9). A student t-test statistical analysis was conducted and false positives were reduced using Benjamini and Hochberg false discovery rate.

2.2.4 *Western blot analysis.*

Cells were washed twice in phosphate buffered saline (PBS) harvested and lysed in lysis buffer (~300 ul/10 cm plate) consisting of 20 mM Tris (pH7.9), 300 mM KCl, 0.1% NP40, 10% glycerol, 0.1 mM DTT, 0.5 mM EDTA, 0.5 mM EGTA and protease inhibitor cocktail (Thermo Scientific Cat# 1862209). For experiments involving detection of $p15^{ink4b}$, RIPA buffer was used to prepare the cell extracts consisting of 50 mM Tris (pH 8.0), 150 mM NaCl, 1% NP-40 and 0.1% SDS. Extracts were centrifuged for 10 min at 16,000xg at 4°C and the soluble extracts were retained. Samples were normalized for protein content and were separated by SDS-PAGE, transferred to nitrocellulose or PVDF membrane and blocked overnight in PBS containing 0.1% TWEEN20 and 5% nonfat dried milk. The appropriate antibodies were then diluted (according to manufacturer's recommendations) in blocking buffer and the membrane was probed for 2 hr at room temperature with rocking followed by the appropriate secondary antibody (1/10 000 dilution) for 1 hr. Proteins were detected using ECL according to the manufacturers recommendations (Amersham). ZNF217 antibody was diluted at 1/3000 dilution.

2.2.5 *Purification of the ZNF217 complex.*

ZNF217 was purified from MCF-7 cells essentially as previously described (13). Approximately 20 mg of nuclear extract was loaded onto a 10 ml phosphocellulose P11 column. The column was then washed using buffer A (20 mM Tris [pH 7.9] 0.5 mM EDTA, 10% glycerol, and 0.5 mM DTT and 100 mM KCl) and ZNF217 was eluted with buffer A containing 0.3M KCl. The ZNF217-containing fraction was assayed by western blotting and then loaded, and reloaded five times, onto an anti-ZNF217 immunoaffinity column that was generated by crosslinking affinity purified ZNF217 antibody to protein

A Sepharose according to standard procedures (Harlow and Lane, 1999). The column was then washed with buffer A containing 0.3 M KCl and 0.1% NP40. Bound proteins were eluted with 100 mM glycine (pH 2.8) containing 100 mM KCl and analyzed for various proteins by western blotting.

2.2.6 ***Chromatin immunoprecipitation assay.***

MCF-7 cells were cross-linked with 1% formaldehyde at room temperature for 10 min. Cross-linking was quenched by immediately washing cells twice with ice-cold PBS. Cells were washed twice with ice-cold PBS containing 0.5 mM EDTA and harvested. Cells pellets were lysed in 0.3 ml of cell lysis buffer (50 mM Tris-HCl [pH 8.1], 10 mM EDTA, 1% SDS, and protease inhibitors) and incubated on ice for 10 min. Cell lysates were sonicated to yield DNA fragments ranging in size from 300- to 1,000-bp. Approximately 450μg of the cross-linked, sheared chromatin solution was used for immunoprecipitation (IP) with. A small portion of each IP was saved as input DNA (5%). Supernatants were diluted 10-fold in dilution buffer (20 mM Tris-HCl [pH 8.1], 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, and protease inhibitors) and precleared with 60 μl of 50% slurry protein A-Sepharose containing 2.5 μg of sheared salmon sperm DNA for 2 h at 4°C. IP was performed overnight at 4°C with 1.5-4 μg of the antibodies. 60 μl of protein A-Sepharose containing 2.5 μg of salmon sperm DNA per ml was added to the solution and incubated for 1 h at 4 °C. The beads were washed one time with wash buffer I (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, 150mM NaCl), wash buffer II (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, 500mM NaCl), wash buffer III (0.25 M LiCl; 1% NP-40; 1% Na-Deoxycholate; 1 mM EDTA; 10 mM TrisHCl) and twice with TE buffer. Immunocomplexes were extracted twice with

200 µl elution buffer (1% SDS-0.1 M $NaHCO_3$). NaCl was added to a final concentration of 200 mM and the cross-linking was reversed by heating at 65 °C overnight. The DNA was purified using Qiagen PCR purification spin columns. For analysis by conventional PCR, conditions were as follows: initial denaturing cycle of at 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 sec, 52 °C for 30 sec and 72 °C for 1 min, and a final elongation step of 72 °C for 10 min. For experiments involving TGFβ treatment, HaCAT cells were plated to approximately 90% confluence and treated with 150 pM TGFβ for 90 min prior to ChIP analysis.

For some experiments, DNA isolated from ChIP experiments was subjected to quantitation by real time PCR using Brilliant SYBR green master mix (Stratagene; 600548). Primers were identified using the Primer Express program (Stratagene) and tested to establish optimum reaction conditions. Reactions were performed in a 25ul volume according to manufacturer's recommendations. The reaction was carried out and measured using Mx3000P realtime instrument. Standard curves were generated using total input DNA (copy number range: 8X105 to 8X101). The IP and IgG DNA copy number was calculated by extrapolating their respective Ct value from the standard curve. The nonimmune IgG copy number was subtracted from IP DNA copy number. The resulting IP copy number was normalized against the total input DNA by dividing the IP by input and expressing the IP as a percentage of the input DNA. All measurements were done in duplicates and an average Ct value was used to calculate copy number. Two independent realtime reactions were done for each experiment.

2.2.7 *ChIP-DSL assay.*

Chromatin immunoprecipitation coupled to DNA selection and Ligation (ChIP-

DSL) was used to assess global promoter occupancy by ZNF217. MCF7 cells were cross-linked with formaldehyde and subjected to standard ChIP assay using affinity purified anti-ZNF217 antibody. The procedure for oligonucleotide annealing, solid phase selection ligation and PCR amplification were performed exactly as described (Aviva Systems Biology; H20K, Cat# AK-0504). The antibody-enriched DNA and the total input were biotinylated followed by annealing to the 40mer oligonucleotide pool. The DNA-oligonucleotide complexes are then selected by binding to streptavidin-conjugated magnetic beads, while the non-annealed oligonucleotides are washed away. Correctly paired 40mers are then ligated to form the corresponding 80mer which is flanked by both universal primer annealing sites (T3 and T7) giving rise to a complete amplicon. A PCR reaction was then conducted on the amplicons using fluorescently labeled T7 and regular T3 primers. Total input DNA was PCR amplified using Cy5 (green) labelled T7 primer and the IP sample was amplified using Cy3 (red) labeled T7 primer. The PCR products are co-hybridized to the 40mer array (Hu20K) to derive an enrichment ratio for each target. After hybridization and washing, array slides were scanned on a One Virtek (Bio-Rad) Chip Reader, and the ArrayVision (v6.0) software package (Genomic Centre, London, Ontario, Canada) was used to quantify fluorescence intensity. The Chip on chip intensity values were normalized using a Lowess curve, which was fit to the log intensity versus log-ratio plot and 20% of the data was used to calculate the Lowess fit at each point. Following normalization, a two-sided student's T-test was conducted where standard deviation of the replicates was used to calculate a p-value. Fold change was calculated for each gene using a mean value that was calculated from all three biological replicates.

2.2.8 *Ingenuity Pathways Systems analysis.*

Ingenuity Pathways Systems (http://www.ingenuity.com) analysis was employed to group statistically significant genes. Genes that had at least a single enrichment were imported into the Ingenuity systems. Only 1215 genes were found in the system database and only those genes were used for further analysis. The significance value associated with a function is expressed as a p-value, which is calculated using the right-tailed Fisher Exact Test. This is done by comparing the number of genes from the gene expression profile that participate in a given function, relative to the total number of occurrences of those genes in all functional annotations stored in the Ingenuity Pathways database.

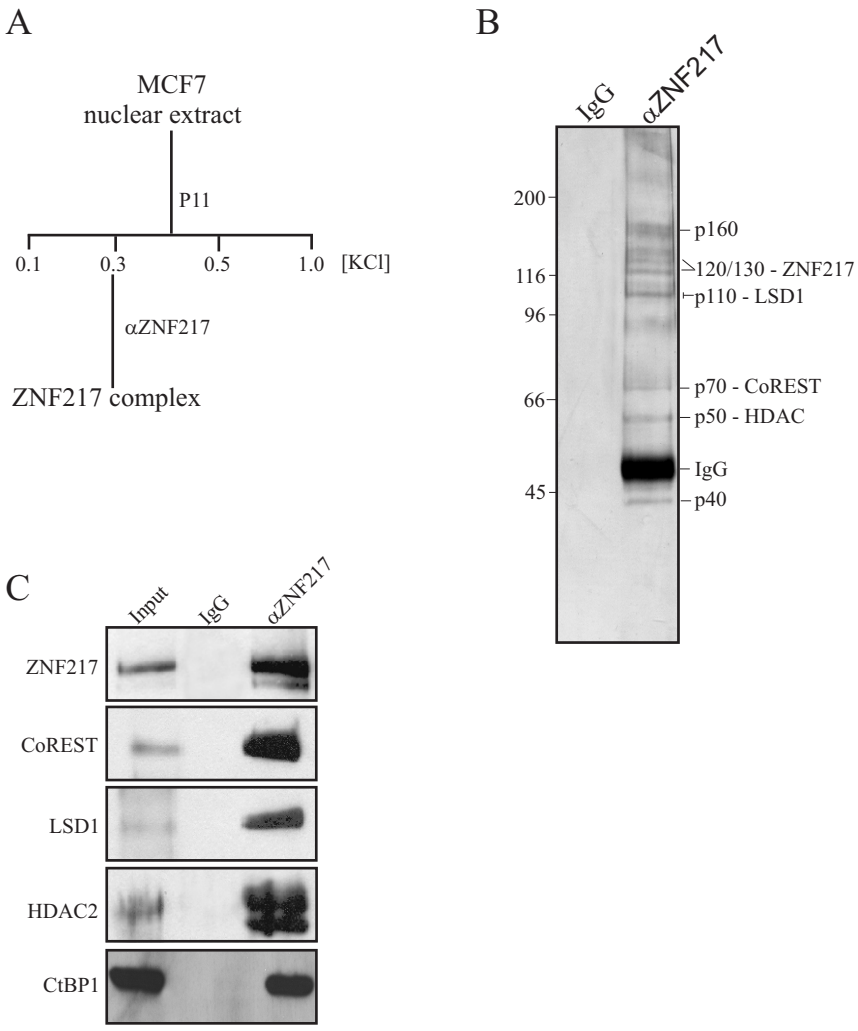
2.3 Results

2.3.1 *Purification of ZNF217 from MCF-7 breast cancer cells.*

Using protein purification in combination with mass spectrometry, we, and others, have recently shown that ZNF217 is a component of a core complex consisting of CoREST, LSD1 CtBP1 and HDAC2 (Cowger et al., 2006; Hakimi et al., 2003; Lee et al., 2005; Shi et al., 2004; Shi et al., 2005; You et al., 2001). To assess the integrity of the ZNF217 complex in MCF7 breast cancer cells we purified ZNF217 by immunoaffinity chromatography using a similar approach (Figure 2.1A). SDS-PAGE analysis of the immunopurified ZNF217 followed by silver staining indicated that the profile of proteins was similar to the ZNF217 complex we recently purified (Cowger et al., 2006) and contains at least three additional polypeptides that were not present in the control immunopurification using rabbit IgG (Figure 2.1B).Western blot analysis using selected antibodies indicated that CoREST, LSD1, HDAC2 and CtBP1 all copurified with

Figure 2.1: Purification of the ZNF217 complex from MCF-7 nuclear extracts.

(A) Purification scheme used to purify the ZNF217 complex from MCF7 cell nuclear extracts. ZNF217 was partially purified by passing MCF7 cell nuclear extracts through a P11 phosphocellulose column prior to immunoaffinity chromatography using ZNF217 antibody. (B) Silver stain SDS-PAGE gel of the purified proteins. A 15 µl aliquot of the purified ZNF217 complex was analyzed by SDS-PAGE followed by silver staining. IgG; represents affinity purification using a rabbit IgG nonimmune affinity column. (C) Western blotting of various proteins found in the ZNF217 complex. A 15 µl aliquot of the purified ZNF217 complex was analyzed by SDS-PAGE followed by western blotting using various antibodies indicated on the left of the figure.


A
MCF7
nuclear extract
P11
0.1
0.3
0.5
1.0
[KCl]
αZNF217
ZNF217 complex
B
IgG
αZNF217
200
116
96
66
45
p160
120/130 - ZNF217
p110 - LSD1
p70 - CoREST
p50 - HDAC
IgG
p40
C
Input
IgG
αZNF217
ZNF217
CoREST
LSD1
HDAC2
CtBP1

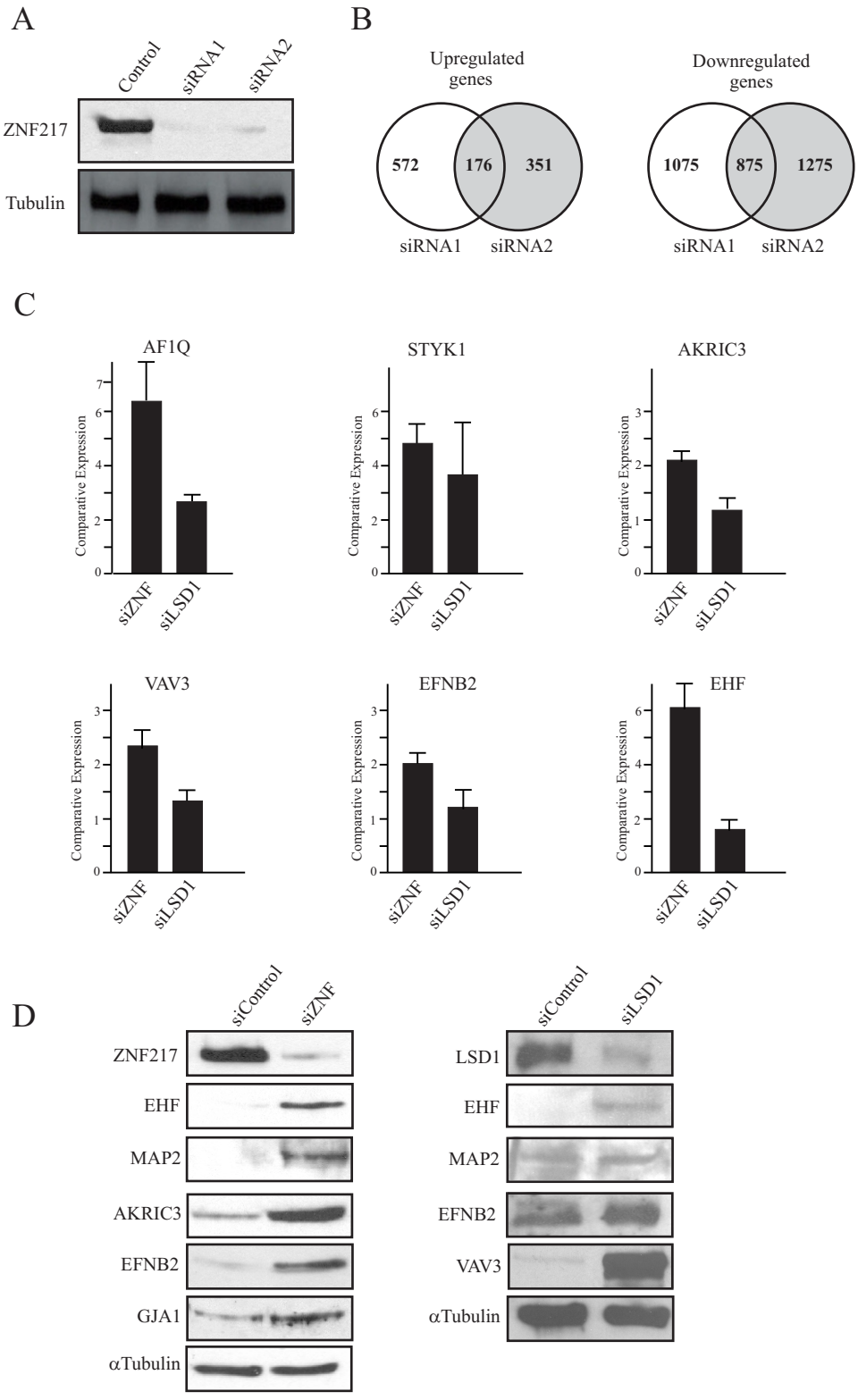
ZNF217 (Figure 2.1C). These results suggest that the ZNF217 complex found in MCF-7 cells is similar to the complex previously identified in HeLa cells (Cowger et al., 2006).

2.3.2 *Identification of gene expression changes in ZNF217-depleted MCF-7 cells.*

To identify genes that are regulated by ZNF217 we initially performed expression array analysis on ZNF217-depleted cells. Two small interfering RNAs (siRNA1 and siRNA2), which recognize distinct regions of the ZNF217 transcript, were transfected into MCF7 cells. The use of multiple siRNAs minimizes potential off target effects associated with individual siRNAs. Western blot analysis of cell lysates prepared 72 hr after transfection confirmed that ZNF217 protein levels were significantly reduced (Figure 2.2A). cDNA from three independent cultures of control and siRNA-treated MCF7 cells were prepared, fluorescently labelled and hybridized to Affymetrix arrays. In our initial analysis, we identified genes displaying statistically significant changes in expression in 3 out of 3 replicates for each siRNA. From this preliminary list of genes we identified those genes which were common to both siRNAs. Using this approach, we identified 176 genes which were significantly upregulated and 875 genes which were downregulated following ZNF217 knockdown (Figure 2.2B, Thillainadesan et al. 2008; Supplementary Table 3). In our initial analysis, we focused on those genes which are upregulated by ZNF knockdown. This is based on the observation that ZNF217 is generally believed to function as a transcriptional repressor and depletion of ZNF217 should result in derepression of target genes. To confirm the results of the expression screen, quantitative real-time PCR analysis and western blotting was performed on randomly selected genes (Figure 2.2C and 2.2D) which indicated significant upregulation following ZNF217 depletion. Realtime PCR analysis and western blotting was also

Figure 2.2: Genome-wide expression screen to identify changes in gene expression associated with ZNF217 depletion.

(A) Western blot of ZNF217-depleted MCF7 cells. MCF7 cells were transfected with siRNA recognizing two different regions of ZNF217 (siRNA1 or siRNA2). Cells were incubated for 72 hr prior to analysis of whole cell extracts by western blotting using ZNF217 antibody. (B) Venn diagrams depicting the overlap in genes upregulated or downregulated from cells transfected with either siRNA1 or siRNA2. (C) Real-time PCR analysis of selected genes identified as significantly upregulated following ZNF217 or LSD1 knockdown using siRNA. Each bar represents the mean relative expression as compared to the expression in cells transfected with control siRNA. (D) Western blot of selected genes significantly upregulated following ZNF217 or LSD1 gene knockdown using siRNA. MCF7 cells were transfected with siRNA targeting either ZNF217 or LSD1. After 72 hr, whole cell extracts were prepared and proteins were analyzed by SDS-PAGE followed by western blotting using the antibodies indicated on the left of each panel.


A
Control
siRNA1
siRNA2
ZNF217
Tubulin
B
Upregulated genes
572
176
351
siRNA1
siRNA2
Downregulated genes
1075
875
1275
siRNA1
siRNA2
C
AF1Q
STYK1
AKRIC3
VAV3
EFNB2
EHF
Comparative Expression
siZNF
siLSD1
D
siControl
siZNF
siLSD1
ZNF217
EHF
MAP2
AKRIC3
EFNB2
GJA1
αTubulin
LSD1
VAV3

performed on LSD1-depleted MCF7 cells to examine whether gene expression was dependent on other components of the ZNF217 complex. For the majority of genes examined, upregulation was observed in response to LSD1 knockdown although the levels of upregulation were generally lower than those observed following ZNF217 downregulation most likely resulting from our inability to fully downregulate LSD1.

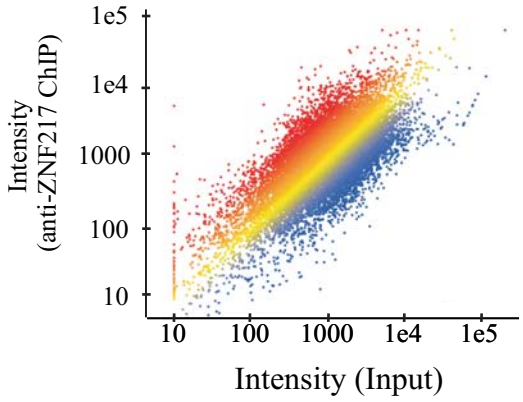
2.3.3 *Genome-wide identification of human promoters bound by ZNF217.*

To determine which of the genes identified in the expression analysis are directly bound by ZNF217, we used a genome-wide chromatin immunoprecipitation assay based on DNA selection and ligation (ChIP-DSL) (Garcia-Bassets et al., 2007; Kwon et al., 2007). A standard ChIP assay was performed on MCF-7 cells using affinity purified ZNF217 antibody. The resulting ZNF217-enriched and input DNA were then biotinylated and combined with 20,000 predesigned oligonucleotides pairs, each representing one-half of an 80mer sequence, corresponding to promoter regions of target genes and flanked by a T3 and T7 primer. After annealing, the biotinylated DNA was purified using streptavidin-conjugated magnetic beads. Adjacent oligonucleotides, that have annealed to the immunoprecipitated DNA, are ligated creating complete amplicons which are then amplified with fluorescently labelled primers and hybridized to a 20,000 gene promoter array. We first validated the functionality of the platform by demonstrating a normalized distribution of the intensity values (Figure 2.3A). From three independent analysis we identified 1431 promoters which are directly bound by ZNF217 (Figure 2.3A and 2.3B and Supplementary Table 4 from Thillainadesan et al. 2008). To identify the nature of the target genes identified, the targets were analyzed using the Ingenuity Analysis program (http://www.ingenuity.com/). Based on this analysis we found that approximately 25% of

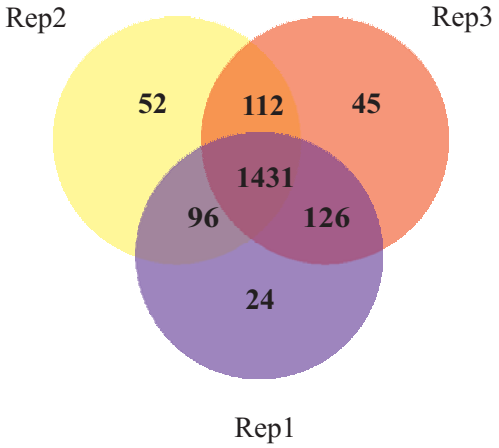
Figure 2.3: ChIP-DSL analysis of ZNF217 target genes in MCF7 cells.

(A) Scatterplot of specific ChIP (y axis) versus total X axis of averaged values from three independent biological replicates demonstrating a normal cluster distribution. Data in red indicates genes significantly enriched. (B) Venn diagram depicting the overlap in genes enriched from three independent ChIP-DSL experiments (Rep 1-3). 1-fold change and a P value < 0.05(calculated using student's T-test) yielded 1431 genes common to all three replicates. (C) Ingenuity functional analysis revealed majority of genes whose promoter regions are bound by ZNF217 to be involved in various diseases. Genes were scored with a significance value which is a log p-value calculated using the Fishers exact test measuring the uniqueness of a gene within a function. (D) Targets corresponding to molecular function were also categorized revealing approximately 25% of ZNF217 target genes are transcription factors.

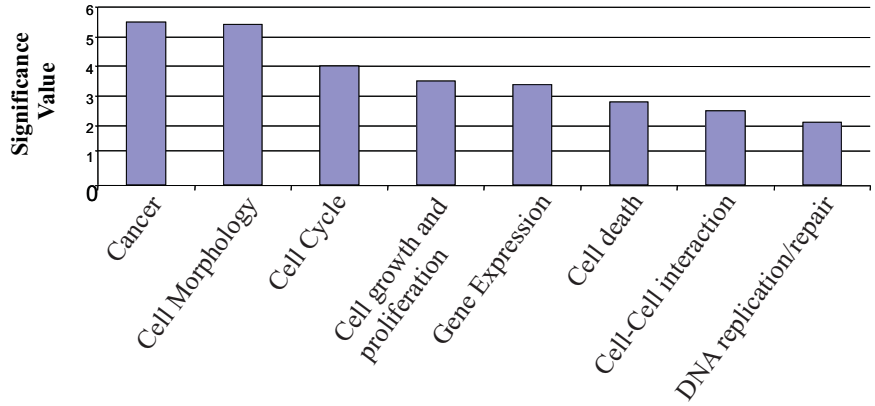
A


Intensity (anti-ZNF217 ChIP)
1e5
1e4
1000
100
10
10
100
1000
1e4
1e5
Intensity (Input)

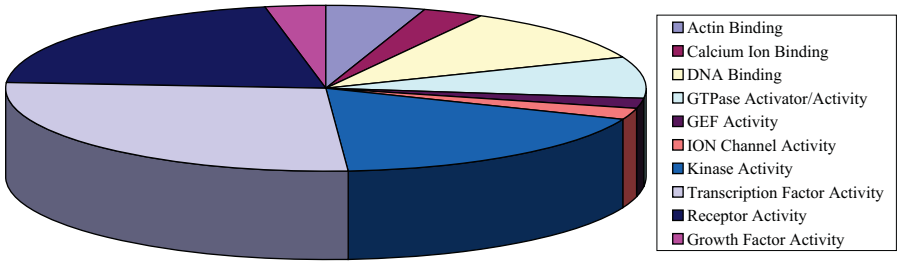
B


Rep2
Rep3
52
112
45
1431
96
126
24
Rep1

C


Significance Value
6
5
4
3
2
1
0
Cancer
Cell Morphology
Cell Cycle
Cell growth and proliferation
Gene Expression
Cell death
Cell-Cell interaction
DNA replication/repair

D

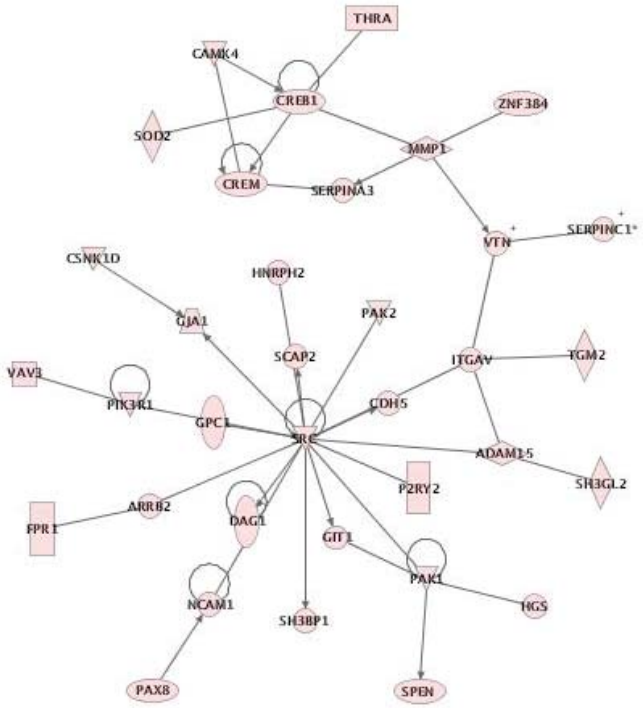

Actin Binding
Calcium Ion Binding
DNA Binding
GTPase Activator/Activity
GEF Activity
ION Channel Activity
Kinase Activity
Transcription Factor Activity
Receptor Activity
Growth Factor Activity

the targets identified are transcription factors suggesting that ZNF217 may play a prominent role in various differentiation processes (Figure 2.3D). Interestingly, in terms of biological function, genes uniquely related to various aspects of cancer were ranked as the most significant, and approximately 73 genes were found to be consistently present in more than one cancer category (Supplementary Table 5 from Thillainadesan et al. 2008;). Additionally, a significant number of genes related to cell morphology and various aspects of the cell cycle were also identified. Network analysis revealed two highly significantly networks of interacting genes (Figure 2.4) with many of the genes within each network linked to tumour suppressor activity.

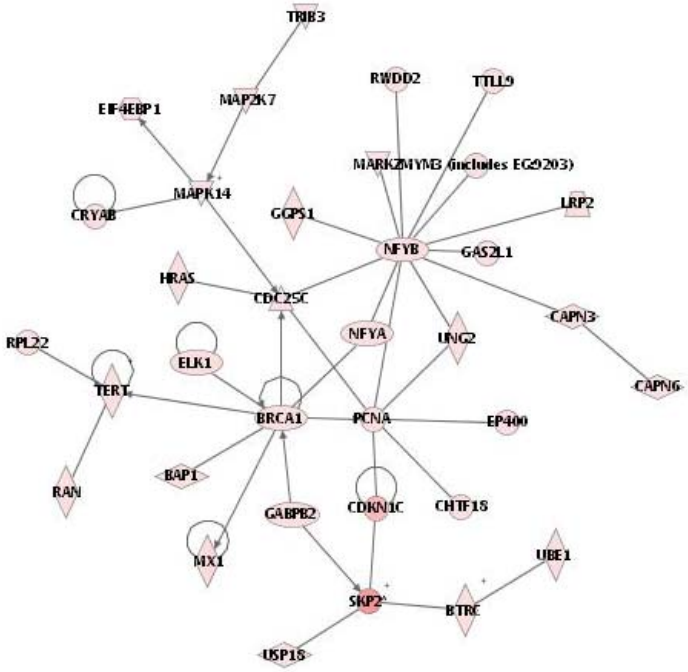
By comparing the ZNF217 target genes, with genes undergoing a significant change in expression following ZNF217 knockdown, we identified 9 genes directly bound by ZNF217 which were found to be significantly upregulated (are repressed by ZNF217), and 45 genes which were significantly downregulated (are activated by ZNF217) (Table 2.3). The results of the ChIP-DSL screen were confirmed by standard ChIP using ZNF217 antibody and oligonucleotide primers corresponding to specific regions within selected promoters (Figure 2.5A and 2.5B, Figure 2.6). Surprisingly, for the vast majority of the target genes identified, no change in expression levels was observed in ZNF217-depleted MCF7 cells. The reason for this are not entirely clear, but our results are consistent with several recent ChIP-chip studies published for other human transcription factors and coregulators suggesting that loss of a single factor may not be sufficient to alter the expression of a gene, perhaps as a result of secondary repressive modifications which are not alleviated following ZNF217 knockdown (Bracken et al., 2006; Carroll et al., 2006; Cawley et al., 2004; Scacheri et al., 2006).

Figure 2.4: ChIP-DSL data analysis.

Network (1 and 2) diagram highlighting interrelationships between specific genes identified in the ZNF217 ChIP-DSL analysis.



Network 1



Network 2

Table 2.3: List of genes directly regulated by ZNF217.

Comparative analysis of ChIP-chip data and expression analysis following ZNF217 depletion in MCF7 cells allowed for the identification of ZNF217 target genes.

[a]KSHV, Kaposi's sarcoma-associated herpesvirus.

Direction of regulation and protein name	Accession	Description
Genes repressed by ZNF217		
MGC45400	AI743979	transcription elongation factor 8
VAV3	AF118886	cell motility
ABCA12	AL080207	membrane transporter
p15INK4b	AW444761	cdk inhibitor
FLJ39370	AI110850	paired like homedomain 2
MAN1A1	BG287153	Glycosylase
EFNB2	BF001670	transmembrane ligand
THY28	NM_014174	thymocyte nuclear protein
FLJ11280	AL561943	Fam 63A
Genes activated by ZNF217		
FAM35A	NM_019054	hypothetical protein MGC5560
DAZAP1	BF512907	DAZ associated protein 1
LOC55971	AA496034	insulin receptor tyrosine kinase substrate
C18orf25	AI823360	mRNA sequence of wh53b02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA
GAJ	AY028916	meiotic nuclear divisions 1 homolog (S. cerevisiae)
C13orf3	AI829603	Chromosome 13 open reading frame 3
SF3A1	BF129339	splicing factor 3a, subunit 1, 120kDa
KLIP1	AA460299	KSHV[a] latent nuclear antigen interacting protein 1
BRCA1	NM_007295	breast cancer 1, early onset
ETF1	NM_004730	eukaryotic translation termination factor 1
TCF3	M31523	transcription factor 3
MCM8	BC005170	minichromosome maintenance deficient 8
LBR	NM_002296	lamin B receptor
CaMKIINα	NM_018584	calcium/calmodulin-dependent protein kinase II
PFKFB3	NM_004566	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
ZNF395	AK021850	papillomavirus regulatory factor PRF-1
SUSD2	Z92546	sushi domain containing 2
XPO4	BF968638	exportin 4
SF3A1	NM_005877	splicing factor 3a, subunit 1, 120kDa
AD-017	NM_018446	glycosyltransferase AD-017
FLJ90022	AW264102	hypothetical protein FLJ90022
PIR51	BE966146	mRNA sequence of wh53b02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA
RAC3	NM_005052	rho family, small GTP binding protein
IRX2	AI928035	iroquois homeobox protein 2
POP7	BC001430	processing of precursor, S. Cerevisiae
FLJ11029	BG165011	Hypothetical Protein
CDC25C	NM_001790	cell division cycle 25C
TUBB2	W72331	tubulin, beta, 2
CECR5	NM_017829	cat eye syndrome chromosome region, candidate 5
ZBTB8	AW006067	wz92a02.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2566250 3', mRNA sequence.
DCXR	NM_016286	dicarbonyl/L-xylulose reductase
FLJ11127	NM_019018	Hypothetica protein
CDCA8	BC001651	cell division cycle associated 8
FGFRL1	AF312678	fibroblast growth factor receptor-like 1

Figure 2.5: ChIP analysis of selected ZNF217 targets.

(A) Direct ZNF217 targets which were found significantly upregulated following ZNF217 depletion. Proliferating MCF7 cells were crosslinked with 1% formaldehyde and ChIP was performed with either control IgG or αZNF217 antibody. The recovered DNA was then assayed by PCR using oligonucleotide primers corresponding to the promoter region of specific genes identified in the ChIP-DSL analysis. (B) Direct ZNF217 targets which were found significantly downregulated in ZNF217-depleted MCF7 cells.

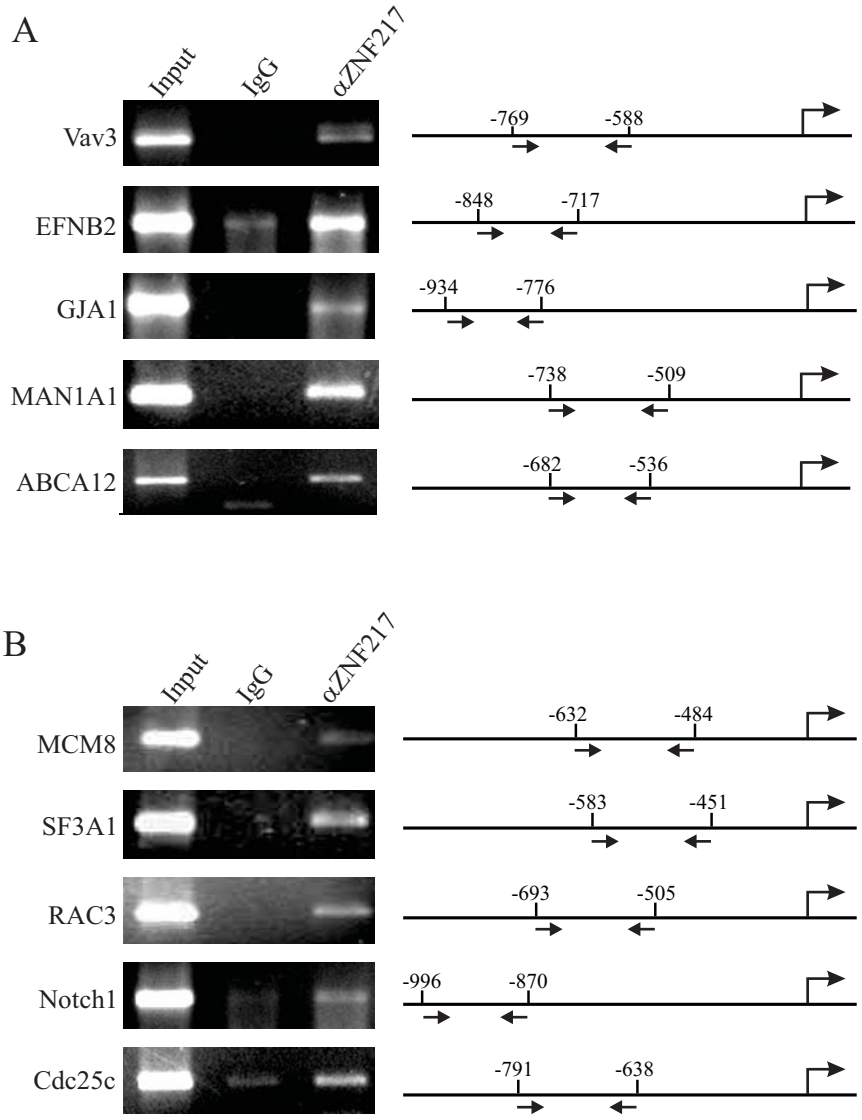
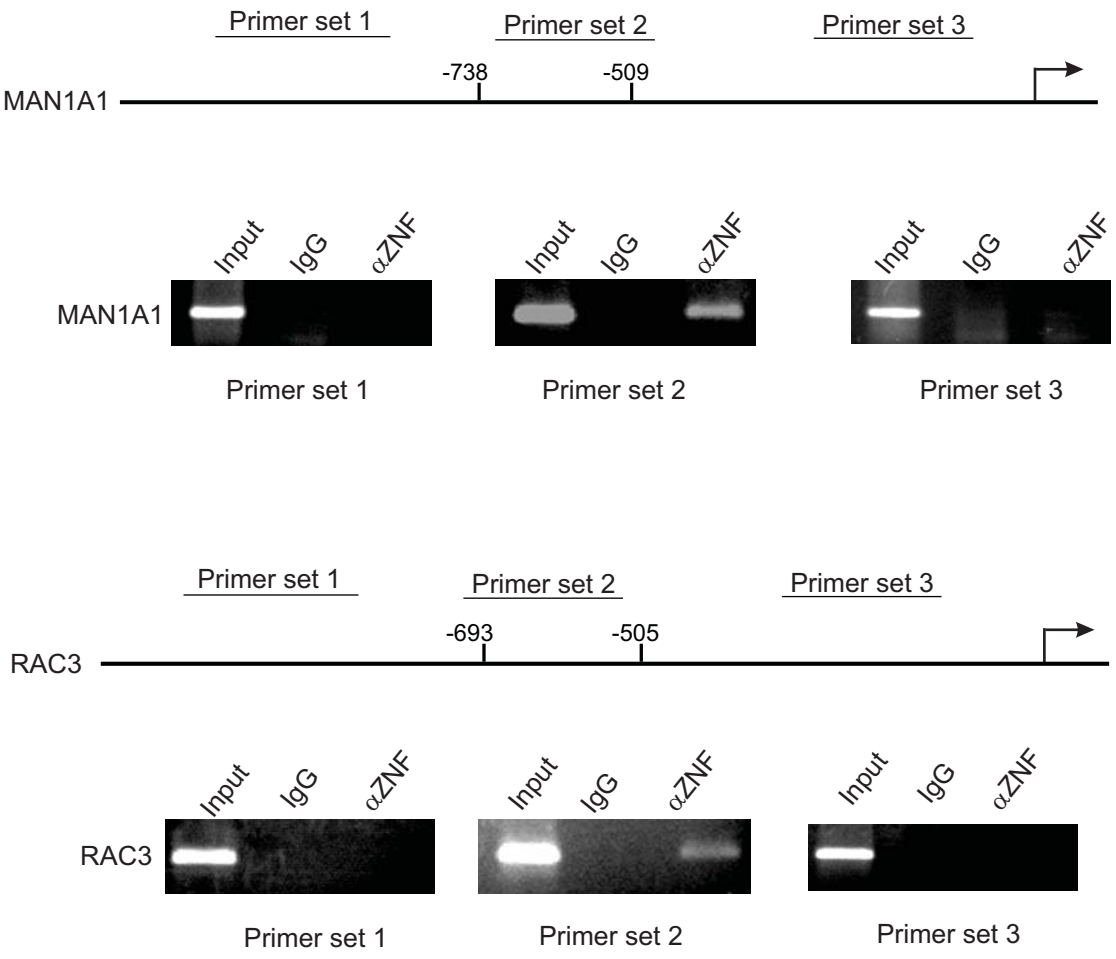

A
Input
IgG
αZNF217
Vav3
EFNB2
GJA1
MAN1A1
ABCA12
-769
-588
-848
-717
-934
-776
-738
-509
-682
-536
B
Input
IgG
αZNF217
MCM8
SF3A1
RAC3
Notch1
Cdc25c
-632
-484
-583
-451
-693
-505
-996
-870
-791
-638

Figure 2.6: ChIP PCR analysis of the MAN1A1 and RAC3 target genes.

ChIP experiments were performed in MCF7 cells using ZNF217 antibody MAN1A1 and RAC3 genes were analyzed using specific oligonucleotide primers (primer set 1, 2 or 3).


Primer set 1
Primer set 2
Primer set 3
-738
-509
MAN1A1
Input
IgG
αZNF
MAN1A1
Primer set 1
Primer set 2
Primer set 3
Primer set 1
Primer set 2
Primer set 3
-693
-505
RAC3
RAC3
Primer set 1
Primer set 2
Primer set 3

2.3.4 *$p15^{ink4b}$ is a direct target of the ZNF217 complex.*

Among the ZNF217 gene targets identified, we focused our initial investigation on the $p15^{ink4b}$ tumour suppressor gene. $p15^{ink4b}$ is a member of the INK4 family of CDK inhibitors which causes cell cycle arrest by directly inhibiting cdk4 and 6 (Kim and Sharpless, 2006). Inactivation of $p15^{ink4b}$ has been found in a wide spectrum of cancers, suggesting that direct silencing of $p15^{ink4b}$ by ZNF217 may contribute to its oncogenic properties (Latres et al., 2000; Melendez et al., 2000).

As shown in figure 2.7, downregulation of ZNF217, or LSD1, using siRNA resulted in significant increases in $p15^{ink4b}$ levels. To confirm binding of ZNF217 to the INK4b promoter and more accurately define its binding site, we performed ChIP analysis in MCF7 cells using pairs of oligonucleotides encompassing approximately 150 base pair intervals (Figure 2.8A). The promoter arrays used in the ChIP-DSL screen consist of unique 80 mer sequences located within 1 kb upstream from the transcription start site for each gene, therefore we restricted our analysis to this region of the $p15^{ink4b}$ promoter. Chip analysis indicated that ZNF217 is highly enriched within a region of the promoter encompassing nucleotides (nt) -566 to -426 (Figure 2.8B). In addition, both LSD1 and CoREST were also found predominantly within this region confirming that the p15ink4B gene is a target for the ZNF217 complex.

To determine whether the presence, or absence, of ZNF217 corresponded to specific chromatin marks at the $p15^{ink4b}$ promoter, ChIP analysis was performed using antibodies corresponding to acetylated K9/14 on histone H3 (K9/K14-H3) and dimethyl K4 on histone H3 (dimetK4-H3), which are generally associated with transcriptionally active genes, as well as trimethyl K27 on histone H3 (trimet K27-H3), a marker for

Figure 2.7: The $p15^{ink4b}$ gene is regulated by the ZNF217 complex.

(A) Western blot of $p15^{ink4b}$ protein following ZNF217 knockdown in MCF7 cells. MCF7 cells were transfected with siRNA recognizing ZNF217 (siRNA1). Cells were incubated for 72 hr prior to analysis of cell extracts by western blotting using ZNF217 or $p15^{ink4b}$ antibody. (B) Real-time PCR analysis of the $p15^{ink4b}$ gene following knockdown of ZNF217 or LSD1 using siRNA. MCF7 cells were transfected with siRNA recognizing ZNF217 (siRNA1) or LSD1. RNA was then prepared, reverse-transcribed and real-time analysis performed. Each bar represents the mean comparative expression relative to cells transfected with control siRNA of two independent experiments.

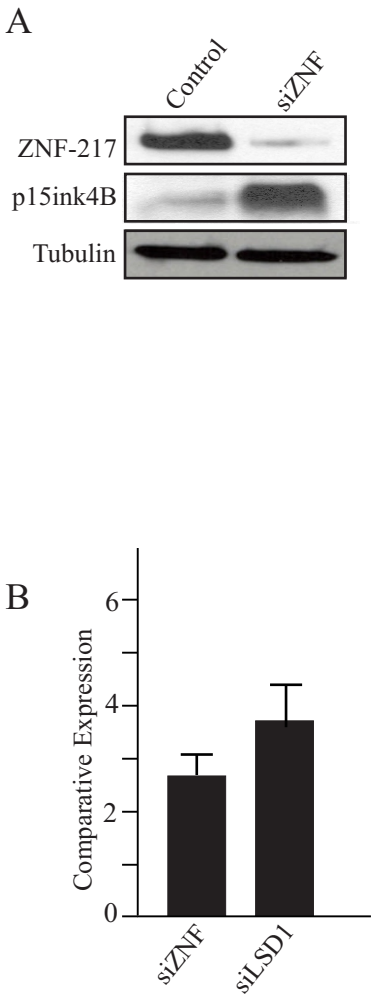
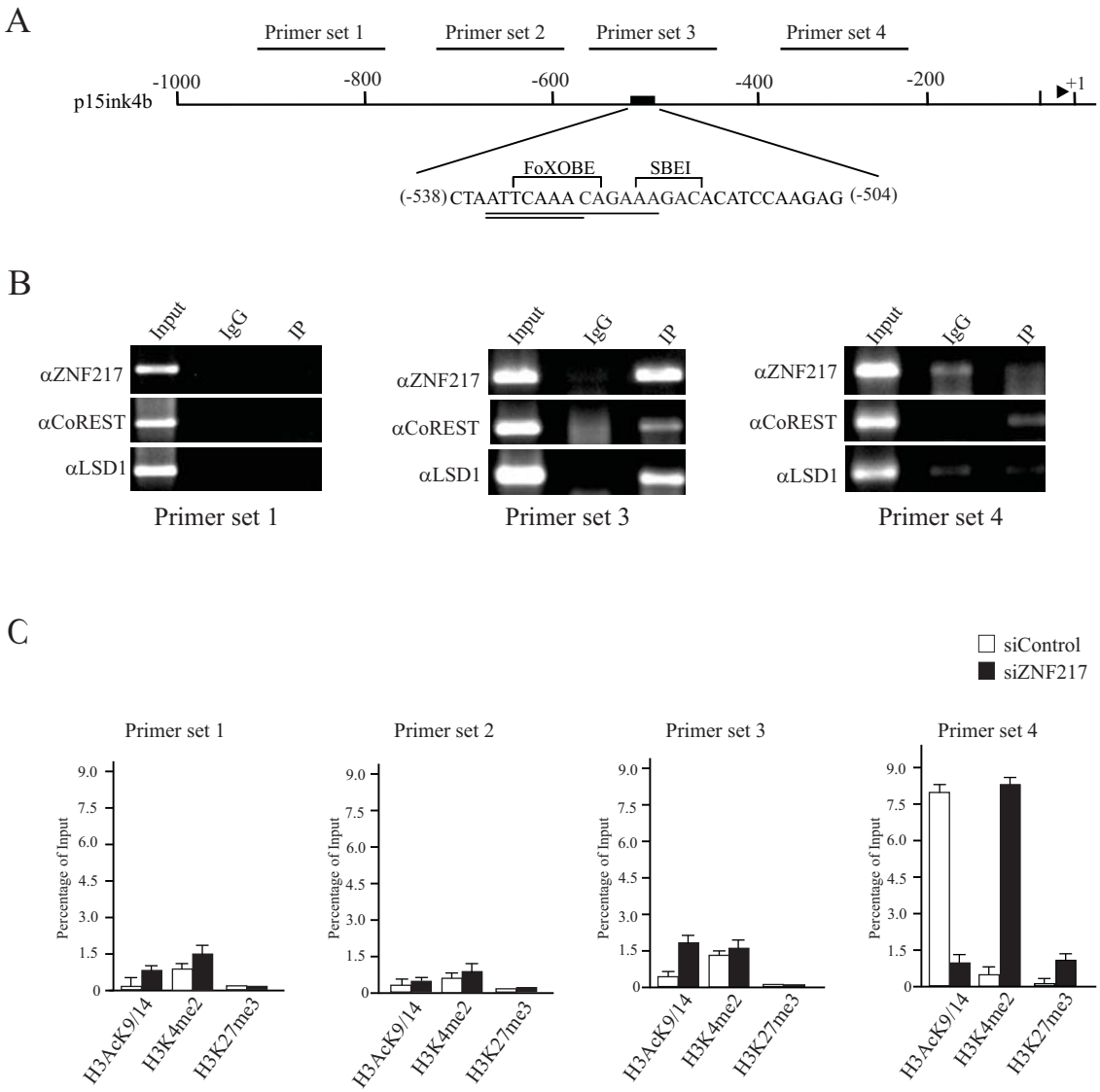

A
Control
siZNF
ZNF-217
p15ink4B
Tubulin
B
Comparative Expression
6
4
2
0
siZNF
siLSD1

Figure 2.8: The $p15^{ink4b}$ promoter is a direct target of the ZNF217 complex.

(A) Schematic representation of the human $p15^{ink4b}$ promoter showing the nucleotide sequence encompassing the Smad binding element (SBE1) and the FoxO3 binding element (FoxO). The underlined sequences represent consensus ZNF217 binding sites identified in (29), double line; and (12) single line. (B) ChIP analysis of the $p15^{ink4b}$ promoter. MCF7 cells were crosslinked with 1% formaldehyde and ChIP was performed with either control antibody (IgG) or the specific antibody (IP) indicated on the left. The recovered DNA was then assayed by PCR using pairs of oligonucleotides encompassing specific regions of the $p15^{ink4b}$ promoter (primer set 1, 3 or 4). (C) ZNF217-dependent changes in histone marks across the $p15^{ink4b}$ promoter. MCF7 cells were transfected with siRNA recognizing ZNF217 (siRNA1). After 72 hr, cells were crosslinked with 1% formaldehyde and ChIP was performed with either control antibody (IgG) or the histone modification-specific antibodies indicated at the bottom of the figure. The recovered DNA was then assayed by realtime PCR using pairs of oligonucleotides encompassing specific regions of the $p15^{ink4b}$ promoter as indicated in Figure 2.6A (primer set 1-4).


A
Primer set 1
Primer set 2
Primer set 3
Primer set 4
-1000
-800
-600
-400
-200
+1
p15ink4b
FoXOBE
SBEI
(-538) CTAATTCAAA CAGAAAGACACATCCAAGAG (-504)
B
Input
IgG
IP
αZNF217
αCoREST
αLSD1
Primer set 1
Primer set 3
Primer set 4
C
siControl
siZNF217
Primer set 1
Primer set 2
Primer set 3
Primer set 4
Percentage of Input
9.0
7.5
6.0
4.5
3.0
1.5
0
H3AcK9/14
H3K4me2
H3K27me3

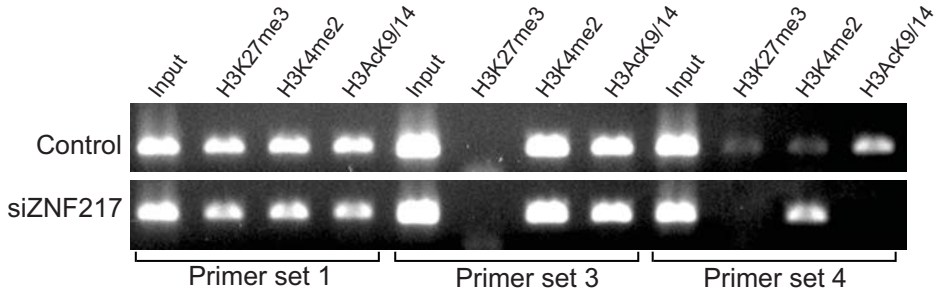
transcriptional repression (Baylin and Ohm, 2006; Berger, 2007; Ting et al., 2006). Realtime PCR analysis identified several significant changes following ZNF217 knockdown. First, a statistically significant increase in acetylation was observed at nt -566 to -426 (primer set 3), containing the ZNF217 binding region. However, the major changes were found downstream of the ZNF217 complex binding site, at nt -340 to -204 (primer set 4). These included a dramatic increase in dimetK4-H3, which would be consistent with the loss of LSD1. Surprisingly, in the presence of ZNF217, this region was found to be highly acetylated at K9/14 on histone H3 and following ZNF217 knockdown, a complete loss of acetylation at K9/14-H3 was observed, despite the observation that p15ink4B is highly expressed (Figure 2.8C, Figure 2.9).

Previous studies have shown that TGFβ stimulates rapid binding of SMAD, FoxO and CEBPβ to the $p15^{ink4b}$ promoter resulting in transcriptional activation of the $p15^{ink4b}$ gene (Gomis et al., 2006a; Gomis et al., 2006b). Interestingly, the ZNF217 binding region of the $p15^{ink4b}$ promoter encompasses both, a SMAD binding site flanked by a forkhead binding element at nt -504 to -538. Based on the proximity of the ZNF217 binding region to the SMAD binding site, we speculated that binding of ZNF217 and SMADs may be mutually exclusive and that coregulator exchange, in response to TGFβ, is a prerequisite for transcriptional activation of the $p15^{ink4b}$ gene. However, in preliminary experiments upregulation of $p15^{ink4b}$ gene in response to TGFβ was not observed in MCF-7 cells (data not shown). Therefore, to examine dynamic changes in ZNF217 complex assembly at the ink4b gene promoter, we used the HaCAT keratinocyte cell line, a well established model for TGFβ-responsive events. For these experiments, HaCAT cells were stimulated with TGFβ and promoter occupancy was assessed by ChIP

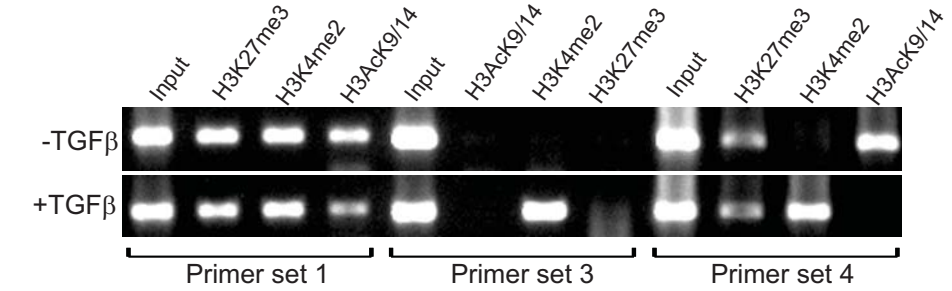
Figure 2.9: ChIP-PCR analysis of the p15^{ink4b} promoter.

Antibodies against trimet H3K27, dimetH3K4 and AcH3K9/14 were used for ChIP experiments and the p15^{ink4b} promoter was analyzed by conventional PCR using the indicated primers A) ChIP-PCR analysis of the p15^{ink4b} promoter following ZNF217 knockdown in MCF7 cells using siRNA B) ChIP-PCR analysis of the p15^{ink4b} promoter following stimulation of HaCAT cells with 150 pM of TGFβ for 90 min.

A


Input
H3K27me3
H3K4me2
H3AcK9/14
Input
H3K27me3
H3K4me2
H3AcK9/14
Input
H3K27me3
H3K4me2
H3AcK9/14
Control
siZNF217
Primer set 1
Primer set 3
Primer set 4

B


Input
H3K27me3
H3K4me2
H3AcK9/14
Input
H3AcK9/14
H3K4me2
H3K27me3
Input
H3K27me3
H3K4me2
H3AcK9/14
-TGFβ
+TGFβ
Primer set 1
Primer set 3
Primer set 4

assay using specific antibodies recognizing either ZNF217 or SMAD2. As shown in figure 2.10A, p15ink4B protein levels are strongly upregulated upon TGFβ stimulation. Importantly, ChIP analysis indicated that stimulation with TGFβ for 90 minutes resulted in a rapid loss of ZNF217 and a concomitant increase in SMAD2 binding from the same region of the $p15^{ink4b}$ promoter (Figure 2.10B).

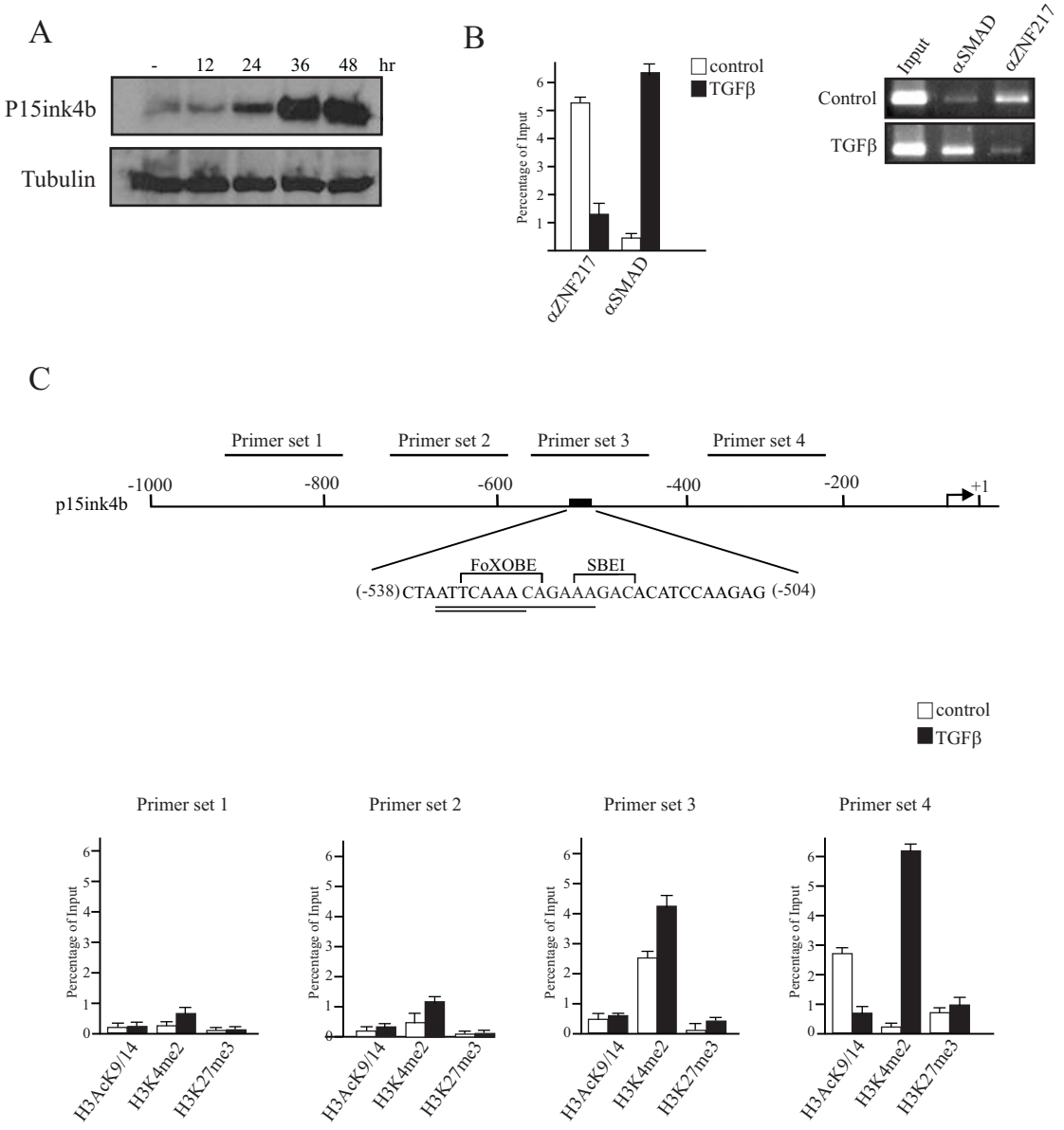
To examine the effect of TGFβ on specific chromatin marks, HaCAT cells were stimulated with TGFβ for 90 minutes and ChIP assays were performed using antibodies recognizing acetylated K9/K14-H3, dimethyl K4-H3 and trimethyl K27-H3 (Figure 2.9, Figure 2.10C,). The major changes were again confined to the nt -340 to -204 region of the $p15^{ink4b}$ promoter, and included increases in dimethylation at H3-K4 and deacetylation at K9/K14-H3. Collectively, these results suggest that loss of ZNF217, and TGFβ stimulation result in similar changes chromatin modifications at the $p15^{ink4b}$ proximal promoter.

2.4 Discussion

The ZNF217 gene codes for a Kruppel-like transcription factor and represents a strong candidate oncogene associated with the 20q13.2 amplification. Insights into the mechanism of ZNF217 have come from studies which have identified ZNF217 as a major constituent of several related transcriptional repressor complexes containing HDAC2 and LSD1 (Cowger et al., 2006; Hakimi et al., 2003; Lee et al., 2005; Shi et al., 2004; Shi et al., 2005; You et al., 2001). Thus, ZNF217 may promote the development of cancer by inappropriately targeting histone modifying enzymes to genes which are essential for

Figure 2.10: TGFβ-inducible release of ZNF217 from the $p15^{ink4b}$ promoter.

(A) Western blot of cells following stimulation with TGFβ. HaCAT cells were stimulated with TGFβ. After various intervals following stimulation, cell extracts were prepared and western blotting was performed using antibodies against $p15^{ink4b}$ or tubulin. Lane 1(-) cells receiving no TGFβ were grown for 48 hr prior to lysis. (B) ChIP analysis of TGFβ-stimulated HaCAT cells. HaCAT were stimulated with TGFβ for 90 minutes. Cells were then crosslinked with 1% formaldehyde and ChIP was performed with either control antibody (IgG), SMAD2 or ZNF217 antibodies. The recovered DNA was then assayed by standard PCR (right panel), or by realtime PCR using pairs of oligonucleotides encompassing the ZNF217 binding region of the $p15^{ink4b}$ promoter (primer set 3). (C) TGFβ-dependent changes in histone marks. HaCAT cells were stimulated with TGFβ for 90 minutes. Cells were then crosslinked with 1% formaldehyde and ChIP was performed with either control antibody (IgG) or the modification specific antibodies indicated at the bottom of the figure. The recovered DNA was then assayed by realtime PCR using pairs of oligonucleotides encompassing specific regions of the $p15^{ink4b}$ promoter as indicated (primer set 1-4).


A
-
12
24
36
48
hr
P15ink4b
Tubulin
B
control
TGFβ
Percentage of Input
αZNF217
αSMAD
Input
αSMAD
αZNF217
Control
TGFβ
C
Primer set 1
Primer set 2
Primer set 3
Primer set 4
-1000
-800
-600
-400
-200
+1
p15ink4b
FoXOBE
SBEI
(-538) CTAATTCAAA CAGAAAGACACATCCAAGAG (-504)
control
TGFβ
Primer set 1
Primer set 2
Primer set 3
Primer set 4
Percentage of Input
H3AcK9/14
H3K4me2
H3K27me3

normal cell growth and differentiation. In the present study, we have used global genomic approaches to identify genes which are directly regulated by ZNF217.

Using ChIP-DSL we were able to identify over 1400 gene promoters bound by the ZNF217 transcription factor in a native chromatin context. Furthermore, in conjunction with siRNA knockdown and microarray analysis, we established that 54 genes are directly regulated by ZNF217. Thus, for the vast majority of the target genes identified, changes in gene expression were not correlated with ZNF217 promoter occupancy in MCF7 cells. This is not entirely unexpected as gene expression is dictated by the repertoire of transcription factors and coregulators that transiently occupy target promoters, and not by a single factor.

While this manuscript was in preparation, a genome-wide ChIP-chip analysis identified approximately 1045 genes, bound by ZNF217, in MCF7 and Ntera2 cell lines, with approximately 745 genes common to both cell types (30). However, only 54 of the target genes identified were found to be directly regulated by ZNF217. A comparison of target genes identified in our study, and the study of Krig et al., indicated 7% overlap (73 genes), and of the 54 genes directly regulated by ZNF217, only 4 were found in our ChIP-DSL analysis. These differences are most likely attributed to variations in experimental design, the specific antibodies used which may recognize different epitopes, and more importantly, differences in the microarray platforms used. Unlike the more conventional ChIP-chip approach which uses ligation mediated PCR (LM-PCR) to amplify the immunoprecipitated DNA, in ChIP-DSL, the DNA is not directly amplified but is used only as a template for annealing of complimentary oligonucleotides corresponding to specific promoter regions. Consequently, this greatly reduces the

complexity of the hybridization mixture and may avoid certain biases inherent in the LM-PCR strategy. However, ChIP-DSL can only identify targets found within 1 kilobase of the transcription start site. Thus, ZNF217-dependent regulatory sites located many kilobases upstream would not be detectable by ChIP-DSL.

Surprisingly, 45 of the 54 target genes we identified are downregulated upon ZNF217 knockdown suggesting that ZNF217 may also function in transcriptional activation. This finding is consistent with recent ChIP on Chip studies of other transcriptional repressor proteins. For example, using ChIP-DSL, recent studies have demonstrated that LSD1 is recruited to target genes which are transcriptionally active as well as repressed (Garcia-Bassets et al., 2007; Kwon et al., 2007). LSD1 has been shown to function in both, transcriptional repression by demethylating H3-K4, and in activation of specific genes by removing the dimethyl mark from H3-K9 rather than H3-K4. Although the underlying mechanism for defining LSD1 specificity is unclear, *in vitro* studies have suggested that LSD1 activity is allosterically regulated through interaction with other proteins such as CoREST, BHC80 and the ligand-bound androgen receptor (Lee et al., 2005; Metzger et al., 2005). In addition, surrounding histone marks found at specific promoters may also play a role in substrate recognition by LSD1 (Forneris et al., 2006; Forneris et al., 2005). Thus, the substrate specificity of LSD1 may be an important determinant for defining the transcriptional activity of ZNF217 at selected targets.

To obtain insight into the mechanism of repression of ZNF217, we focused our analysis on the $p15^{ink4b}$ gene. The $p15^{ink4b}$ gene is found within a 35 kb stretch of DNA, the INK4 locus, which also contains p16ink4A and p14ARF a splice variant of p16ink4A. The entire locus has been found frequently deleted or mutated in many types

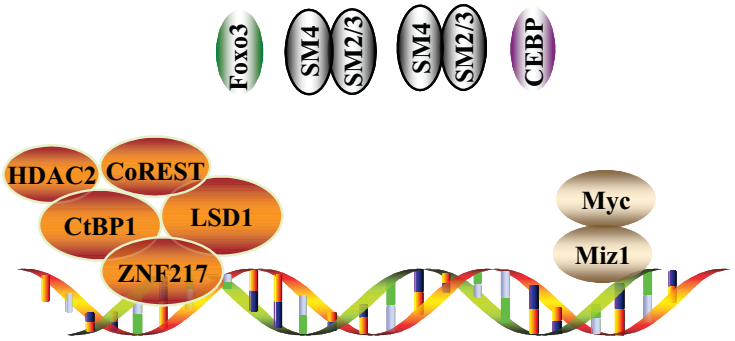
of cancer (Ghiorzo et al., 2004; Sill et al., 1995). Interestingly, we did not detect changes in expression levels of either $p16^{ink4A}$ or p14ARF, based on the microarray expression analysis, indicating that the repressive effects of the ZNF217 complex are specific for $p15^{ink4b}$.

In many epithelial cell lines, $p15^{ink4b}$ is rapidly upregulated in response to TGFβ and under normal growth conditions, contributes to the TGFβ–dependent cytostatic program. Previous studies using HaCAT cells have shown that the induction of $p15^{ink4b}$ occurs predominantly at the level of transcription through a dual mechanism involving downregulation of c-myc and the recruitment of activating transcription factors to the promoter region (Gil and Peters, 2006; Seoane et al., 2001; Staller et al., 2001). Myc acts as a negative regulator of the $p15^{ink4b}$ gene by preventing the transcription factor Miz1 from activating $p15^{ink4b}$ transcription (Seoane et al., 2001; Staller et al., 2001). The addition of TGFβ suppresses myc expression, depleting the cellular pools of myc available for binding to Miz1 which, in turn, relieves active repression of the p15INK4b gene. Concurrently, SMAD proteins, as well as other transcription factors, bind to specific DNA elements in the promoter region and, in association with Miz1, elicit full activation of the $p15^{ink4b}$ gene (Feng et al., 2002; Feng et al., 2000).

The identification of the ZNF217 complex as a negative regulator of the $p15^{ink4b}$ gene in breast cancer cells adds an additional layer of complexity in our understanding of the molecular events regulating $p15^{ink4b}$ gene transcription (Figure 2.11). In MCF7 cells the levels of p15ink4B are virtually undetectable and downregulation of ZNF217 using

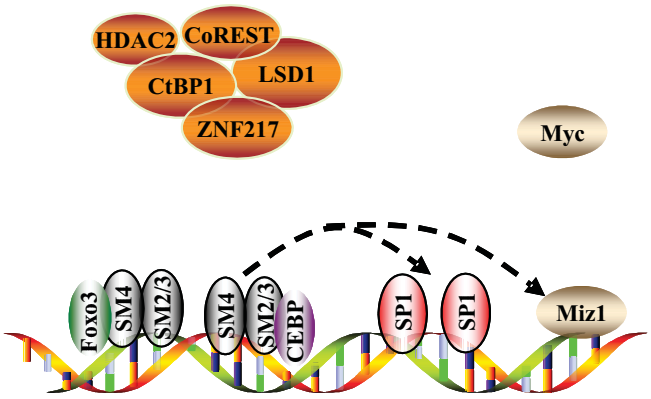
Figure 2.11: Model highlighting the role of ZNF217 complex in $p15^{ink4b}$ expression.

In normal proliferating epithelial cells, the ZNF217 complex is bound to the $p15^{ink4b}$ promoter and expression of $p15^{ink4b}$ is repressed. Stimulation with TGFβ causes a release of the ZNF217 complex and a concomitant binding of activating transcription factors which include SMADs (SM2/3 and SM4), CEBPβ, FoxO3 and SP1. Additionally, associations between adjacent transcription factors, as well as downregulation of Myc following TGFβ stimulation, may result in full activation of the $p15^{ink4b}$ gene.



p15^{ink4B} Off

TGFβ



p15^{ink4B} On

Cell cycle arrest

siRNA resulted in a dramatic increase in $p15^{ink4b}$ protein levels, indicating that loss of ZNF217 alone was sufficient to relieve repression of the $p15^{ink4b}$ gene in this cell type. ChIP analysis indicated that the ZNF217 binds a region of the promoter that has previously been shown to encompass a SMAD binding element flanked by a FoxO site (Gomis et al., 2006a; Gomis et al., 2006b). Recent studies have shown that upregulation of $p15^{ink4b}$, in response to TGFβ, is dependent on rapid binding of both SMAD proteins and FoxO3 to their respective sites. Interestingly, a CAGAAA and an ATTCAA motifs directly overlap the SMAD and FoxO3 binding sites, respectively. Both of these elements have been identified as putative consensus ZNF217 binding sites in independent studies (Cowger et al., 2006; Krig et al., 2007). Based on these observations, we speculated that TGFβ-dependent activation of the $p15^{ink4b}$ gene is dependent on release of ZNF217. This was indeed confirmed in HaCAT cells where treatment with TGFβ resulted in a rapid loss in ZNF217 and a concomitant increase in SMAD2 binding, at nt -566 to -426, which preceded increases in $p15^{ink4b}$ protein expression. Collectively, these findings establish the ZNF217 complex as a novel suppressor of the $p15^{ink4b}$ gene (Figure 2.11).

We have also examined the relationship between promoter occupancy by ZNF217 and changes in chromatin modifications at the $p15^{ink4b}$ promoter. Following ZNF217 knockdown, a dramatic increase in dimethylation of K4-H3 was observed at nt -340 to -204. Dimethyl K4-H3 serves as a major substrate for LSD1, and suggests that targeting of LSD1 may depend directly on ZNF217 recruitment. Unexpectedly, we have found that -340 to -204 is also highly acetylated at K9/K14 of histone H3, when the $p15^{ink4b}$ gene is not expressed, and is deacetylated upon ZNF217 knockdown. Although deacetylation of histones is generally associated with transcriptional silencing of genes, several reports

have demonstrated a requirement for HDAC activity in transcriptional activation. For example, analysis of the IFN-β promoter has shown that deacetylation of specific lysine residues serves as a prerequisite for transcriptional activation (Agalioti et al., 2002). Similarly, removal of yeast repressor proteins RPD3 and SIN3 resulted in decreased transcription in a number of genes and increased silencing (De Rubertis et al., 1996). One possibility is that hyperacetylation at K9/K14 is required for binding of repressor proteins containing specific recognition motifs, such as bromodomains, which bind to acetylated lysines with high affinity. For example BRD4, a bromodomain-containing transcriptional adaptor protein, inhibits HPV transcription by binding to acetylated lysines on histone H3 and blocks recruitment of the core transcriptional machinery (Wu et al., 2006). Consequently, the loss of acetylation at K9/K14, and increases in dimethyl K4-H3, following ZNF217 release, may function cooperatively to target additional proteins to the $p15^{ink4b}$ promoter which, in turn, may increase the accessibility of the chromatin to the transcriptional machinery and facilitate transcriptional activation of the $p15^{ink4b}$ gene.

Similar changes in chromatin marks were also observed in HaCAT cells following stimulation with TGFβ. Importantly, the changes we have observed coincide with the release of ZNF217, indicating that loss of ZNF217 and stimulation with TGFβ are most likely part of a concerted signaling mechanism regulating the $p15^{ink4b}$ gene. The differences in the magnitude of the changes in covalent modifications between MCF7 and HaCAT cells may be attributed to differences in cell types, as well as experimental conditions. Covalent modifications related to a specific transcriptional outcome often involve repetitive cycles of association and dissociation of transcription factors and a large number of coregulator proteins in a sequential manner (Metivier et al., 2003).

Consequently, 90 minute following TGFβ stimulation may not be optimal for the establishment of specific changes in chromatin structure, even though increases in SMAD2 binding and loss of ZNF217 are clearly evident by this time point. Nevertheless, the changes in chromatin marks, following loss of ZNF217, or stimulation of cells with TGFβ are quite striking and suggest that the -340 to -204 region is dynamically regulated and most likely plays a critical role in determining expression of the $p15^{ink4b}$ gene.

Many tumours and cancer cell lines develop resistance to the growth inhibitory effects of TGFβ which in some cases results from loss-of-function mutations in TGFβ receptors or SMAD proteins (Pardali and Moustakas, 2007). However, the majority of tumours which have lost TGFβ–dependent functional responses retain an intact TGFβ signaling mechanism suggesting that specific downstream defects may be involved. This may include overexpression of specific proteins resulting in a deregulated transcriptional response at TGFβ-dependent target genes. The c-ski oncoprotein which is overexpressed in a subset of leukemic patients, negatively regulates TGFβ signaling by interfering with the formation of SMAD complexes at target genes resulting in abnormal silencing of transcription (He et al., 2003; Luo, 2004; Luo et al., 1999). More recently, it has been shown that excess levels of the C/EBPβ isoform liver enriched transcriptional inhibitor protein (LIP) may, in part, account for the loss of TGFβ responsiveness (Gomis et al., 2006b). C/EBPβ consists of multiple isoforms, a liver-enriched transcriptional activator protein 1 and 2 LAP1/2, and LIP which lacks regulatory domains found in LAP and functions as a dominant negative for C/EBPβ-dependent transcription. LAP binds directly to the $p15^{ink4b}$ promoter and is required for induction of $p15^{ink4b}$ gene in response

to TGFβ. A high LIP/LAP ratio has been reported in some metastatic breast cancers and lowering of the LIP/LAP ratio by overexpression of LAP restores both TGFβ-dependent induction of $p15^{ink4b}$ and the growth inhibitory response to TGFβ. Finally, it has been shown that human mammary epithelial cells, aberrantly expressing ZNF217, become immortalized and develop resistance to the growth inhibitory properties of TGFβ (Nonet et al., 2001).

The identification of the $p15^{ink4b}$ gene as a direct target for the ZNF217 corepressor complex represents a potentially novel link between amplification of ZNF217 and the loss of TGFβ responsiveness in breast cancer. The proximity of the ZNF217 binding region to the SMAD binding is consistent with this hypothesis, and suggests that the balance between coactivators and corepressor proteins at the level of gene transcription represents a critical regulatory mechanism and an important determinant of cell growth.

2.5 References

Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. Cell *111*, 381-392.

Bar-Shira, A., Pinthus, J.H., Rozovsky, U., Goldstein, M., Sellers, W.R., Yaron, Y., Eshhar, Z., and Orr-Urtreger, A. (2002). Multiple genes in human 20q13 chromosomal region are involved in an advanced prostate cancer xenograft. Cancer Res *62*, 6803-6807.

Baylin, S.B., and Ohm, J.E. (2006). Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? Nat Rev Cancer *6*, 107-116.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407-412.

Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics *19*, 185-193.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev *20*, 1123-1136.

Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., *et al.* (2006). Genome-wide analysis of estrogen receptor binding sites. Nat Genet *38*, 1289-1297.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., *et al.* (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell *116*, 499-509.

Chin, K., de Solorzano, C.O., Knowles, D., Jones, A., Chou, W., Rodriguez, E.G., Kuo, W.L., Ljung, B.M., Chew, K., Myambo, K., *et al.* (2004). In situ analyses of genome instability in breast cancer. Nat Genet *36*, 984-988.

Chin, K., DeVries, S., Fridlyand, J., Spellman, P.T., Roydasgupta, R., Kuo, W.L., Lapuk, A., Neve, R.M., Qian, Z., Ryder, T., *et al.* (2006). Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell *10*, 529-541.

Collins, C., Rommens, J.M., Kowbel, D., Godfrey, T., Tanner, M., Hwang, S.I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K., *et al.* (1998). Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. Proc Natl Acad Sci U S A *95*, 8703-8708.

Cowger, J.J., Zhao, Q., Isovic, M., and Torchia, J. (2006). Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. Oncogene.

De Rubertis, F., Kadosh, D., Henchoz, S., Pauli, D., Reuter, G., Struhl, K., and Spierer, P. (1996). The histone deacetylase RPD3 counteracts genomic silencing in Drosophila and yeast. Nature *384*, 589-591.

Feng, X.H., Liang, Y.Y., Liang, M., Zhai, W., and Lin, X. (2002). Direct interaction of c-Myc with Smad2 and Smad3 to inhibit TGF-beta-mediated induction of the CDK inhibitor p15(Ink4B). Mol Cell *9*, 133-143.

Feng, X.H., Lin, X., and Derynck, R. (2000). Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-beta. Embo J *19*, 5178-5193.

Forneris, F., Binda, C., Dall'Aglio, A., Fraaije, M.W., Battaglioli, E., and Mattevi, A. (2006). A highly specific mechanism of histone H3-K4 recognition by histone demethylase LSD1. J Biol Chem *281*, 35289-35295.

Forneris, F., Binda, C., Vanoni, M.A., Battaglioli, E., and Mattevi, A. (2005). Human histone demethylase LSD1 reads the histone code. J Biol Chem *280*, 41360-41365.

Garcia-Bassets, I., Kwon, Y.S., Telese, F., Prefontaine, G.G., Hutt, K.R., Cheng, C.S., Ju, B.G., Ohgi, K.A., Wang, J., Escoubet-Lozach, L., *et al.* (2007). Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. Cell *128*, 505-518.

Ghiorzo, P., Pastorino, L., Bonelli, L., Cusano, R., Nicora, A., Zupo, S., Queirolo, P., Sertoli, M., Pugliese, V., and Bianchi-Scarra, G. (2004). INK4/ARF germline alterations in pancreatic cancer patients. Ann Oncol *15*, 70-78.

Gil, J., and Peters, G. (2006). Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. Nat Rev Mol Cell Biol *7*, 667-677.

Gomis, R.R., Alarcon, C., He, W., Wang, Q., Seoane, J., Lash, A., and Massague, J. (2006a). A FoxO-Smad synexpression group in human keratinocytes. Proc Natl Acad Sci U S A *103*, 12747-12752.

Gomis, R.R., Alarcon, C., Nadal, C., Van Poznak, C., and Massague, J. (2006b). C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. Cancer Cell *10*, 203-214.

Hakimi, M.A., Dong, Y., Lane, W.S., Speicher, D.W., and Shiekhattar, R. (2003). A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. J Biol Chem *278*, 7234-7239.

Harlow, E., and Lane, D. (1999). Using Antibodies: A laboratory manual (Cold Spring Harbour, Cold Spring Harbour Laboraory Press).

He, J., Tegen, S.B., Krawitz, A.R., Martin, G.S., and Luo, K. (2003). The transforming activity of Ski and SnoN is dependent on their ability to repress the activity of Smad proteins. J Biol Chem *278*, 30540-30547.

Hidaka, S., Yasutake, T., Takeshita, H., Kondo, M., Tsuji, T., Nanashima, A., Sawai, T., Yamaguchi, H., Nakagoe, T., Ayabe, H., *et al.* (2000). Differences in 20q13.2 copy number between colorectal cancers with and without liver metastasis. Clin Cancer Res *6*, 2712-2717.

Huang, G., Krig, S., Kowbel, D., Xu, H., Hyun, B., Volik, S., Feuerstein, B., Mills, G.B., Stokoe, D., Yaswen, P., *et al.* (2005). ZNF217 suppresses cell death associated with chemotherapy and telomere dysfunction. Hum Mol Genet *14*, 3219-3225.

Kim, W.Y., and Sharpless, N.E. (2006). The regulation of INK4/ARF in cancer and aging. Cell *127*, 265-275.

Krig, S.R., Jin, V.X., Bieda, M.C., O'Geen, H., Yaswen, P., Green, R., and Farnham, P.J. (2007). Identification of genes directly regulated by the oncogene ZNF217 using ChIP-chip assays. J Biol Chem.

Kwon, Y.S., Garcia-Bassets, I., Hutt, K.R., Cheng, C.S., Jin, M., Liu, D., Benner, C., Wang, D., Ye, Z., Bibikova, M., *et al.* (2007). Sensitive ChIP-DSL technology reveals an extensive estrogen receptor alpha-binding program on human gene promoters. Proc Natl Acad Sci U S A *104*, 4852-4857.

Latres, E., Malumbres, M., Sotillo, R., Martin, J., Ortega, S., Martin-Caballero, J., Flores, J.M., Cordon-Cardo, C., and Barbacid, M. (2000). Limited overlapping roles of P15(INK4b) and P18(INK4c) cell cycle inhibitors in proliferation and tumorigenesis. Embo J *19*, 3496-3506.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432-435.

Luo, K. (2004). Ski and SnoN: negative regulators of TGF-beta signaling. Curr Opin Genet Dev *14*, 65-70.

Luo, K., Stroschein, S.L., Wang, W., Chen, D., Martens, E., Zhou, S., and Zhou, Q. (1999). The Ski oncoprotein interacts with the Smad proteins to repress TGFbeta signaling. Genes Dev *13*, 2196-2206.

Melendez, B., Malumbres, M., de Castro, I.P., Santos, J., Pellicer, A., and Fernandez-Piqueras, J. (2000). Characterization of the murine p19(ARF) promoter CpG island and its methylation pattern in primary lymphomas. Carcinogenesis *21*, 817-821.

Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003). Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. Cell *115*, 751-763.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436-439.

Nonet, G.H., Stampfer, M.R., Chin, K., Gray, J.W., Collins, C.C., and Yaswen, P. (2001). The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. Cancer Res *61*, 1250-1254.

Pardali, K., and Moustakas, A. (2007). Actions of TGF-beta as tumor suppressor and pro-metastatic factor in human cancer. Biochim Biophys Acta *1775*, 21-62.

Quinlan, K.G., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., Bolognesi, M., and Crossley, M. (2006). Specific recognition of ZNF217 and other zinc-finger proteins at a surface groove of CtBPs. Mol Cell Biol.

Scacheri, P.C., Davis, S., Odom, D.T., Crawford, G.E., Perkins, S., Halawi, M.J., Agarwal, S.K., Marx, S.J., Spiegel, A.M., Meltzer, P.S., *et al.* (2006). Genome-wide analysis of menin binding provides insights into MEN1 tumorigenesis. PLoS Genet *2*, e51.

Seoane, J., Pouponnot, C., Staller, P., Schader, M., Eilers, M., and Massague, J. (2001). TGFbeta influences Myc, Miz-1 and Smad to control the CDK inhibitor p15INK4b. Nat Cell Biol *3*, 400-408.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., and Casero, R.A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell *19*, 857-864.

Sill, H., Goldman, J.M., and Cross, N.C. (1995). Homozygous deletions of the p16 tumor-suppressor gene are associated with lymphoid transformation of chronic myeloid leukemia. Blood *85*, 2013-2016.

Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Moroy, T., Bartek, J., Massague, J., Hanel, F.*, et al.* (2001). Repression of p15INK4b expression by Myc through association with Miz-1. Nat Cell Biol *3*, 392-399.

Suzuki, S., Egami, K., Sasajima, K., Ghazizadeh, M., Shimizu, H., Watanabe, H., Hasegawa, H., Iida, S., Matsuda, T., Okihama, Y.*, et al.* (2004). Comparative study between DNA copy number aberrations determined by quantitative microsatellite analysis and clinical outcome in patients with stomach cancer. Clin Cancer Res *10*, 3013-3019.

Tanner, M.M., Grenman, S., Koul, A., Johannsson, O., Meltzer, P., Pejovic, T., Borg, A., and Isola, J.J. (2000). Frequent amplification of chromosomal region 20q12-q13 in ovarian cancer. Clin Cancer Res *6*, 1833-1839.

Thillainadesan, G., Isovic, M., Loney, E., Andrews, J., Tini, M., and Torchia, J. (2008). Genome analysis identifies the p15ink4b tumor suppressor as a direct target of the ZNF217/CoREST complex. Mol Cell Biol *28*, 6066-6077.

Ting, A.H., McGarvey, K.M., and Baylin, S.B. (2006). The cancer epigenome--components and functional correlates. Genes Dev *20*, 3215-3231.

Watanabe, T., Imoto, I., Katahira, T., Hirasawa, A., Ishiwata, I., Emi, M., Takayama, M., Sato, A., and Inazawa, J. (2002). Differentially regulated genes as putative targets of amplifications at 20q in ovarian cancers. Jpn J Cancer Res *93*, 1114-1122.

Weiss, M.M., Snijders, A.M., Kuipers, E.J., Ylstra, B., Pinkel, D., Meuwissen, S.G., van Diest, P.J., Albertson, D.G., and Meijer, G.A. (2003). Determination of amplicon boundaries at 20q13.2 in tissue samples of human gastric adenocarcinomas by high-resolution microarray comparative genomic hybridization. J Pathol *200*, 320-326.

Wu, S.Y., Lee, A.Y., Hou, S.Y., Kemper, J.K., Erdjument-Bromage, H., Tempst, P., and Chiang, C.M. (2006). Brd4 links chromatin targeting to HPV transcriptional silencing. Genes Dev *20*, 2383-2396.

You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc Natl Acad Sci U S A *98*, 1454-1458.

Chapter 3: TGF-β-Dependent Active Demethylation and Expression of the $p15^{ink4b}$ Tumor Suppressor Are Impaired by the ZNF217/CoREST Complex

3.1 Introduction

Transforming growth factor β (TGFβ) is an essential cytokine that coordinates a complex antiproliferative program that includes induction of the cyclin dependent kinase (cdk) inhibitors $p15^{ink4b}$ and $p21^{cip1}$. The *$p15^{ink4b}$* is a tumour suppressor and a component of the ink4 locus, which also contains the *$p16^{ink4a}$* and *ARF* genes. The regulation of $p15^{ink4b}$ involves several interrelated mechanisms which operate primarily at the transcriptional level. One mechanism involves silencing by the c-myc oncoprotein which is recruited to the $p15^{ink4b}$ promoter through interactions with Miz-1 and represses its transactivation properties. TGFβ suppresses myc expression, which relieves repression of Miz1 (Seoane et al., 2001; Staller et al., 2001). At the same time, SMAD proteins are rapidly recruited to the promoter, along with other sequence-specific transcription factors, resulting in activation of $p15^{ink4b}$ gene. The Corepressor of Rest (CoREST) complex also plays a role in silencing of the $p15^{ink4b}$ gene in some cell types (Thillainadesan et al., 2008). This complex consists of the ZNF217 oncoprotein, and at least 4 additional proteins: the histone deacetylase 1/2, the lysine demethylase 1 (LSD1), the Corepressor of Rest (CoREST) and the C terminal binding protein 1 or 2 (CtBP 1/2) (Cowger et al., 2007; Lee et al., 2005; You et al., 2001). The ZNF217 gene is found at the core of the 20q13.2 amplicon (Collins et al., 1998) and is amplified and overexpressed in many cancers ranging in amplification frequency from 10 to 40% (Bar-Shira et al., 2002; Li et al., 2007a; Rooney et al., 2004). Several studies have shown that overexpression of ZNF217 is associated with cellular immortalization, increased cellular proliferation and resistance to TGFβ (Krude, 1999; Li et al., 2007a; Nonet et al., 2001; Sun et al., 2008; Thollet et al.).

Aberrant transcriptional silencing of $p15^{ink4b}$ involving promoter DNA hypermethylation is frequently associated with hematologic and other cancers (Aggerholm et al., 2006; Papageorgiou et al., 2007; Tsellou et al., 2005). Nevertheless, in many cell types TGFβ as well as other stimuli rapidly induce $p15^{ink4b}$ expression suggesting that mechanisms are in place for reversing DNA methylation (Gomis et al., 2006a; Gomis et al., 2006b; Warner et al., 1999). DNA demethylation can occur passively when newly synthesized DNA remains unmethylated following DNA replication, or by an active mechanism which is replication independent and involves enzymes that act directly on 5-methylcytosine (5mC). For example, the interleukin-2 promoter undergoes rapid DNA demethylation following T lymphocyte activation in a replication independent manner (Bruniquel and Schwartz, 2003). In MCF7 cells, the pS2/TFF1 gene undergoes cyclical demethylation and remethylation shortly after β-estradiol treatment which coincides with estrogen receptor binding and activation (Metivier et al., 2008). While a unified mechanism responsible for active demethylation has not been definitively established, one mechanism which has recently gained momentum involves base excision repair (BER) initiated by the DNA glycosylases, thymine DNA glycosylase (TDG) and methyl-CpG binding domain 4 (MBD4). It has been postulated that these enzymes may directly cleave the glycosidic bond between the 5mC and deoxyribose creating an abasic site which is then removed and repaired by the BER machinery. However, while MBD4 has been shown to excise 5mC under some conditions (Kim et al., 2009), in the majority of studies demethylation involving TDG and MBD4 requires enzymatic conversion of 5mC to a more favourable substrate. For example, the activation-induced deaminase (AID) or Apolipoprotein B mrRNA editing

enzymes (APOBEC 1-4) can deaminate 5mC to thymine to generate a G:T mispair (Wu and Zhang, 2010). TDG and MBD4 have a high affinity for G:T mispairs and excision of the mispair initiates BER which effectively replaces the 5mC. This mechanism is supported by recent findings in zebrafish where active DNA demethylation is mediated by the cooperative activities of AID, MBD4 and an auxiliary factor GADD45α (Rai et al., 2008).

Another mechanism may involve the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) catalyzed by the Ten Eleven Translocation (TET 1-3) proteins (Wu and Zhang, 2011). It has been postulated that 5hmC is an intermediary metabolite which can be deaminated to 5-hydroxmethyluracil (5hmU) creating a 5hmU:G mismatch that is efficiently recognized and excised by TDG (Cortellino et al., 2011; Hashimoto et al., 2012). Alternatively, a more recently implicated mode of TET-mediated demethylation involves the conversion of 5hmC to 5 formylcytosine (5fC) and 5 carboxylcytosine (5cC) (He et al., 2011; Ito et al., 2011; Krude, 1999; Maiti and Drohat, 2011). TDG possesses robust activity against both 5fC and 5cC which could account for the relatively low levels of both derivatives in the mammalian genome. Collectively, these findings suggest that active DNA demethylation is a multi-step process that links TET-mediated oxidation of 5mC to the BER pathway initiated by the TDG and MBD4 DNA glycosylases.

In the present study we have examined the mechanism of dynamic DNA methylation and demethylation of the $p15^{ink4b}$ promoter in response to TGFβ. Using ChIP-seq we demonstrate that the $p15^{ink4b}$ tumour suppressor gene is a target for the

ZNF217/CoREST complex and that promoter hypermethylation and silencing is, in part, dependent on an association between DNMT3A and ZNF217/CoREST. TGFβ stimulates active demethylation of the $p15^{ink4b}$ promoter by a mechanism involving loss of the DNMT3A/ZNF217/CoREST complex and recruitment of SMAD2/3, the CBP acetyltransferase and TDG or MBD4 to the same promoter region. Remarkably, overexpression of ZNF217 largely inhibits active DNA demethylation and expression of $p15^{ink4b}$ by preventing recruitment of the DNA demethylation machinery. Collectively, the present study suggests that active DNA demethylation involving BER is an essential step in $p15^{ink4b}$ induction and that deregulation of this mechanism may contribute to aberrant silencing of this tumour suppressor gene in cancer.

3.2 Materials and Methods

3.2.1 *Plasmids, Antibodies, and Reagents.*

The ZNF217 antibody was generated as described (Cowger et al., 2007). The antibody used for ChIP experiments involving TDG was prepared as described (Mohan et al., 2007). For immunoprecipitation and western blotting experiments (Figure 3.6E) the antibodies used are described in (Mohan et al., 2007; Hardeland et al., 2002). $p15^{ink4b}$and $p21^{cip1}$ RNA levels were quantitated using Taqman (Applied Biosystems). The mouse TDG cDNA was prepared as described (Mohan et al., 2007). Recombinant Flag-tagged TDG was generated using the Bac-to-Bac baculovirus expression system (Invitrogen) exactly as described. Human TGFβ1 was purchased from R&D Systems (Cat# 100-B-001). L-mimosine was obtained from Sigma. A list of commercial antibodies, oligonucletide primers and siRNAs used in this study can be found in Table 3.1, Table 3.2 and Table 3.3.

Table 3.1: Antibodies used in this study.

Antibody	Cat#	Supplier
CoREST	BL1039	Bethyl Laboratories
CtBP1	612042	BD-Western Blotting
CtBP1	sc-11390	Santa Cruz Biotechnology, Inc. ChIP
LSD1	A300-216A	Bethyl Laboratories
DNMT3A	sc-20703	Santa Cruz Biotechnology, Inc.
SMAD2/3	sc-6202	Santa Cruz Biotechnology, Inc.
CBP	sc-369	Santa Cruz Biotechnology, Inc.
AID	sc-14680	Santa Cruz Biotechnology, Inc.
TDG	Tini and Scharr	Tini and Scharr Lab
APE/REF1	Tini(sc-17774)	Santa Cruz Biotechnology, Inc.
DNA Ligase I	sc-47703	Santa Cruz Biotechnology, Inc.
DNA Polβ	sc-48810	Santa Cruz Biotechnology, Inc.
MBD4	A-1009	Epigentek
ZNF217	Torchia	Torchia Lab
Phospho-Rb	# 9308	Cell Signalling
5mC	ChIPgrade	Millipore
5hmC	39770	Active motif

Table 3.2: qPCR primers used for detection of methylation and chromatin immunoprecipitation of the $p15^{ink4b}$ and $p21^{cip1}$ promoters.

Name	Forward 5' → 3'	Reverse 5' → 3'	Coordinates[a]
Methyl-specific p15	TTGG**CG**TATG**CG**T TTTAGTATT	AACTAAACCAAA AAAC**CG**A**CG**T	-194/-76
Unmethyl-specific p15	TTGG**TG**TATG**TG**T TTTAGTATT	AACTAAACCAAA AAAC**CA**A**CA**T	-194/-76
p15 Bisulphite sequencing	GAGTTGAGGGTAG TGGTGAATATTT	TTTTCAACTAAA CCAAAAAACC	-298/-71
p15 MeDIP/ChIP[b]	CATGATTCTCGGG ATTTTTCTC	GACAGCTCTGCA CCTGTCAT	-598/-458
p15 upstream MeDIP/ChIP[c]	CTAGGAAATGGGG GTTGGAT	CCCCCAAATCCC TGTAGAAT	-1463/-1285
Unmethyl-specific p21	T**TGTG**ATAAGGA GATTTTAGGGAAC	AT**CA**CTATACCG ATAAAAAAACGA A	+784/+890
p21 MeDIP[d]	GAGGAAAAGCAT CTTGGAG	AAATAGACGGGA GCAACG	+1421/+1534

[a]Measured in base pairs from transcriptional start site.
[b]Primers to detect mid region of the p15 promoter following ChIP/MeDIP.
[c]Control primers used to detect upstream region of the p15 promoter following ChIP/MeDIP.
Underlined and bolded nucleotides represent CpG sites that were targeted.
[d](Koinuma et al., 2009)

Table 3.3: Silencing RNAs (siRNA) used in this chapter.

siRNA	Cat#	Supplier
CtBP1	J-0008609-11	Dharmacon
DNMT3a	sc-37757	Santa Cruz Biotechnology, Inc.
MBD4	sc-37763	Santa Cruz Biotechnology, Inc.
ZNF217 #1	CUSTOM	GCAAAUAACCUCAUCUGUAUU
mTDG	TINI	Dharmacon
Control Pool	D-00180-10-20	Dharmacon
APOBEC2	sc-95404	Santa Cruz Biotechnology, Inc.
AID	CUSTOM	UUCAAAAAUGUCCGCUGGGCUU

3.2.2 *Cell Culture, Adenovirus Infections, and Transfections.*

MCF7 and HaCAT cells were maintained in DMEM containing 10% FBS. For experiments involving TGFβ treatments, HaCAT cells were grown to approximately 60% confluency and serum starved for 24 hrs. Cells were then washed with PBS and resuspended in DMEM containing 10% FBS in the presence of vehicle (4 mM HCl containing 0.1% human BSA), or 150 pM TGFβ1 for the indicated time period prior to analysis. For adenovirus infections, HaCAT cells were infected with a recombinant adenovirus expressing ZNF217 ($1x10^6$ PFU) or GFP, in media containing 8ug/ml Polybrene, for 24 hrs prior to analysis. All siRNAs were delivered into cells using Oligofectamine (Invitrogen) according to the manufacturer's recommendations. Unless otherwise indicated, cells were harvested 72 hrs after siRNA transfection. Protein extracts were made and western blotting was performed as described (Thillainadesan et al., 2008). Western blots were quantitated by densitometry using the software ImageQuant 5.1. Unless otherwise indicated, the data was normalized using alpha-tubulin.

3.2.3 *Flow Cytometry.*

MCF7 cells were transfected with either control or ZNF217 siRNA for 72 hours and then treated with 1μg/ml of Nocodazole (Sigma). At the indicated time points post nocodazole treatment cells were fixed in ethanol, washed twice in PBS containing 1% FBS and stained with propidium iodide. Single cells were gated out and %age of cells in G1, S or G2/M phase was determined by DNA content using the cell cycle analysis module (Watson-Pragmatic model) of FlowJo (Tree Star Inc.).

3.2.4 *Chromatin Immunoprecipitation Assay.*

ChIP, ChIP-ReCHIP and DNA immunoprecipitation assays were performed as described (Thillainadesan et al., 2008). For ChIP-ReChIP assays, protein-DNA complexes were eluted in 25 µl 10 mM DTT incubated at 37° C for 30 min. The eluted DNA was diluted 10 fold in ReChIP buffer (20 mM Tris-HCl [pH 8.1], 0.1% Triton X-100, 2 mM EDTA, 150 mM NaCl) and incubated with a second antibody (1.5 to 4 µg) overnight. Antibody-protein complexes were immunprecipitated using protein A Sepharose beads (Upstate) for 2 hrs at 4° C. The beads were washed once with wash buffer I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, 150 mM NaCl), wash buffer II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, 500 mM NaCl), and Tris-EDTA buffer. DNA extraction, de-crosslinking and purification were done as previously described (Thillainadesan et al., 2008). Sequential ChIP with two specific antibodies was always matched with a control sequential ChIP performed with the first specific antibody and a second non-specific IgG antibody. The chromatin immunoprecipitation assay used prior to deep sequencing analysis was modified slightly. Sonication conditions were altered such that the majority of the genomic fragments were in the 250-350 bp range. In addition, Protein A/G magnetic Dynabeads (Invitrogen) were utilized for immunoprecipitation of the protein DNA complexes.

3.2.5 *DNA immunoprecipitation Assay*

MeDIP was conducted as previously described (Cortazar et al., 2011). Genomic DNA was isolated using the GeneElute Mammalian Genomic DNA miniprep kit (Sigma G1N70) and sonicated using a Bioruptor (15s ON/30s OFF) for 15 min to yield DNA fragments in the range of 200bp – 600bp. 4µg of the fragmented DNA was diluted in

450µl of TE and heat denatured at 95°C for 10 min followed by addition of 50 µl of cold 10X immunoprecipitation buffer (100mM sodium phosphate pH 7.0, 1.4 M NaCl, 0.5% Triton X-100). The DNA was incubated with 10 µg of ChIP grade anti-mC antibody (Millipore cat# MABE146) and 1/500 dilution of the anti 5hmC antibody (Active motif cat# 39770) for 2 hr at 4 C. The antibody/DNA complex was captured using magnetic protein G beads for 2hr at 4°C and washed three times with 700 µl of 1X IP buffer. DNA was eluted in elution buffer (50mMTris -HCl pH 8.0, 10mM EDTA, 0.5% SDS and 0.25 mg ml^{-1} proteinase K) for 3hr at 50 C. Input DNA as well as the IP samples were purified using the PCR purification kit (Qiagen) and quantitated by realtime PCR.

3.2.6 *In vitro glycosylation assay*

The *in vitro* DNA glycosylation assay was performed essentially as described with slight modifications (Cortellino et al.; Hardeland et al., 2000). The substrate DNAs consisted of T/G mismatch oligonucleotides CAATCCTAGCTGACACGATGTGGCCA ATGGC ATGACT (top) and GAGTCATGCCATTGGCCACATTGTGTCAGCTAG GATTG (bottom), hemi-methylated oligonucleotides CAATCCTAGCTGACACGATG TGGCC AATGGCATGACT (top) and GAGTCATGCCATTGGCCACAT^{m}CGTGTCA GCTAG GATTG (bottom), or symmetrically methylated oligonucleotides CAATCC TAGCTGACA^{m}CGATGTGGCCAATGGCATGACT (top) and GAGTCATGCCATTG GCCACAT^{m}CGTGTCAGCTAGGATTG (bottom). Prior to annealing, the bottom strand oligonucleotide was radiolabelled at the 5' end with T4 polynucleotide kinase (Fermentas) and γ-^{32}P-ATP (3000 Ci/mmol). The annealed substrates (1 pmol) were incubated for 30 min at 37°C in a reaction buffer consisting of 25 mM HEPES pH 8.0, 1 mM DTT, 0.1 mg/ml BSA, 1 mM EDTA] and various concentrations of purified

recombinant FLAG-TDG (100-500 ng) in a total volume of 20 µl. For some experiments, the reactions were combined with ~10 µl of TDG immunopurified from transfected HaCAT cells. The reactions were terminated by the addition of 2 µl of 1 M NaOH followed by incubation at 95 C for 5 min. Gel loading buffer (90% formamide , (1:1/ v:v) was added to each sample and subjected to 8.0 M urea/20% PAGE. The gel was then dried and analyzed by autoradiography.

3.2.7 *Dot Blot Analysis*

Genomic DNA was quantified using a nanodrop spectrophotometer (ND-1000) and 300ng of DNA along with two subsequent serial dilutions (1/10) were used for the experiment. DNA was then denatured at 99°C for 5min in a 0.1M sodium hydroxide solution and neutralized on ice using 0.1vol of 6.6M ammonium acetate. 5µl of the denatured DNA was manually spotted onto an AmershamHybond-N+ membrane and dried in a hybridization oven at 80°C for 10min. The DNA was then UV-crosslinked using a transluminator for 90 sec and blocked overnight at 4°C in blocking buffer (10% milk, PBS-T (PBS+ 0.1% Tween-20). The membrane was incubated with 10 µg of anti 5mC antibody (Millipore 33D3) diluted in 10ml of blocking buffer for 2hr followed by two washes with PBS-T. A mouse secondary antibody conjugated to a horseradish peroxidase enzyme (HRP) was incubated with the membrane for 45min followed by three washes with PBS-T. The chemiluminescent signal was detected and by ECL (Amersham). To confirm loading the same blots were also staining with 0.2% methylene blue for 30 min.

3.2.8 *Immunoprecipitations*

HaCAT cells were harvested at approximately 60% confluency and whole cell extracts (WCE) were prepared as described (Thillainadesan et al., 2008). Approximately 750 μg of WCE was incubated in the presence of 5μg of TDG antibody overnight at 4°C. The following day immunocomplexes were captured using Protein A/G agarose beads, washed four times with lysis buffer (50mM Tris pH 8, 0.3M KCl, 0.5mM EDTA/EGTA, 0.25% NP-40) and complexes were analyzed by SDS-PAGE and western blotting with the appropriate antibodies. For experiments involving overexpressed TDG, 293 embryonic cells were transfected with 10μg of Flag tagged TDG with SuperFect (Qiagen). The following day, cells were treated with 150pM TGFβ for 90min and WCE were prepared. TDG was immunopurified using Flag Sepharose and the immune complexes were then tested for glycosylase activity.

3.2.9 *Genomic Sodium Bisulphite Analysis.*

Genomic DNA was isolated from cells using DNeasy Tissue (Qiagen cat# 59104) and bisulphite converted using EpiTect Bisulfite Kit (Qiagen cat#69504). For sequencing, bisulphite converted DNA was PCR amplified using methyl sequencing primers and cloned directly into TA cloning vector (Invitrogen Cat# 45-0046) and sequenced using the T7 primer. For methylation analysis using PCR, bisulphite converted DNA was subjected to methyl and/or unmethylated specific quantitative PCR as indicated. The quantitative reactions were carried out as previously described (Thillainadesan et al., 2008) using Brilliant SYBR® Green mastermix (Agilent Technologies cat# 600548) and quantitated with a Mx3000P (Stratagene) real-time PCR instrument.

3.2.10 *ChIPseq Analysis.*

ChIP assays were performed using either ZNF217 or CtBP1 antibodies. The ChIP-DNA was sequenced at the University of British Columbia (http://www.cmmt.ubc.ca/ facilities/services/sequencing) using the Illumina sequencing platform. Approximately, 2.1 million raw reads for ZNF217 and 1.7 million raw reads for CtBP1 were obtained. Each read corresponded to approximately 75bp of DNA. The obtained reads were screened using PERL scripts that eliminated reads based on the following criteria; reads that were shorter than 75bp, reads that had gaps and reads that had repeat nucleotides longer than 10bp. The later criteria for eliminating repeat sequence is based on the fact that larger than 10bp repeats were not found within the genome according to blast results (http://www.ncbi.nlm.nih.gov/blast) and the repeats may have been a result of sequence duplication events. The obtained reads were then aligned using Bowtie 0.12.5 with Hg19 human genome. Alignment using Bowtie, resulted in 85.8% and 87.3% of aligned reads to the Hg19 human genome for ZNF217 and CtBP1, respectively. Peaks were identified using CisGenome (Ji et al., 2008) and MACS (Zhang et al., 2008).

Negative binomial model was used to calculate the expected percentage of windows that contained a certain number of reads. Percentage of false positive rate was calculated by dividing the observed number of windows by the expected number of windows containing the same number of reads. A false positive percentage of 6.3% and 5.1% corresponded to a read number of 8 for ZNF217 and CtBP1, respectively. A cut-off value of 8 reads was used to control for false positives peaks.

3.3 Results

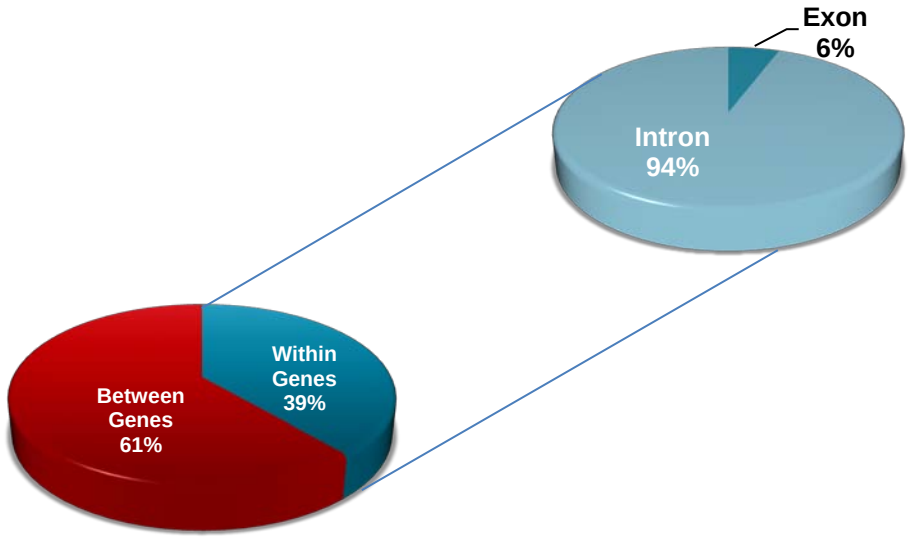
3.3.1 *ZNF217 Is Essential for Recruitment of the CoREST Complex to the p15^{ink4b} Promoter.*

To define the genome occupancy of ZNF217/CoREST complex, we performed ChIPseq in MCF7 cells using antibodies directed against ZNF217 and CtBP1 (Table S2 from Thillainadesan et al., 2012). Examination of the normalized genome-wide distribution of ZNF217 indicated that only 2% of ZNF217 binding could be annotated to promoters of specific genes within 1 kb of a transcription start site (TSS), suggesting that the ZNF217/CoREST complex plays a role in repression at a distance (Figure 3.1 and Table S1 from Thillainadesan et al. 2012). Approximately, 65.7% (FPR of 3.0%) of CtBP1 binding sites were found to overlap with ZNF217 binding which is consistent with earlier observations using ChIP on chip (Krig et al., 2007) and indicate that both constituents are recruited as a complex to the majority of genomic sites. Of the 1202 CtBP1 binding sites annotated within 50 kb of a TSS, 979 (81.4%) binding sites are also bound by ZNF217 (Figure 3.2). Comparison of the ZNF217 and CtBP1 binding sites to a microarray expression analysis of ZNF217–depleted MCF7 cells (Thillainadesan et al., 2008) identified several common genes involved in cell cycle control. Importantly, ZNF217/CtBP1 binding was detected in a region of the p15^{ink4b} promoter that overlaps directly with a previously identified regulatory region containing a SMAD/FoxO binding site (Figure 3.3A and 3.3B) (Gomis et al., 2006b; Seoane et al., 2001; Staller et al., 2001; Thillainadesan et al., 2008). Knockdown of ZNF217 did not affect expression of p16^{ink4a} or ARF based on microarray analysis suggesting that repression by ZNF217 is selective for p15^{ink4b} in the ink 4 locus (Krig et al., 2007; Thillainadesan et al., 2008). To confirm the ChIPseq data, we examined occupancy by the CoREST complex following

Figure 3.1: Genomic analysis of ZNF217 binding in MCF7 cells.

(A) Pie chart indicating the genomic distribution of ZNF217 binding in MCF7 cells. (B) Pie chart indicating the distribution of ZNF217 binding based on the distance from a transcription start site.

A


Exon
6%
Intron
94%
Within
Genes
39%
Between
Genes
61%

B

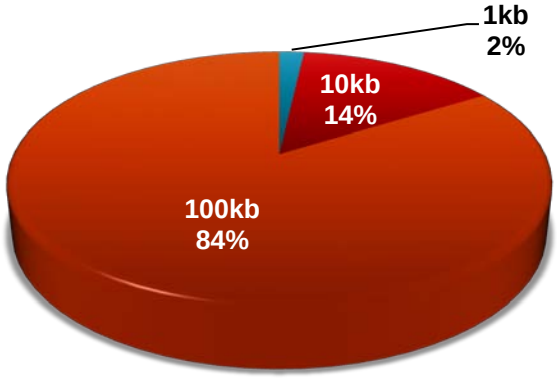

1kb
2%
10kb
14%
100kb
84%

Figure 3.2: Overlap between ZNF217 and CtBP1 binding.

(A**)** The pie chart indicates the overlap in binding sites between CtBP and ZNF217. Peaks were considered overlapping if individual peaks were located within 500 bps of each other. (B) Pie chart indicating overlap in neighboring genes. The neighboring genes were annotated within 50 kb from the TSS and the percentage of genes common to ZNF217 and CtBP1 are indicated. (C) Venn diagram indicating the number of genes annotated for ZNF217 and CtBP1. Shown in the square are selected genes.

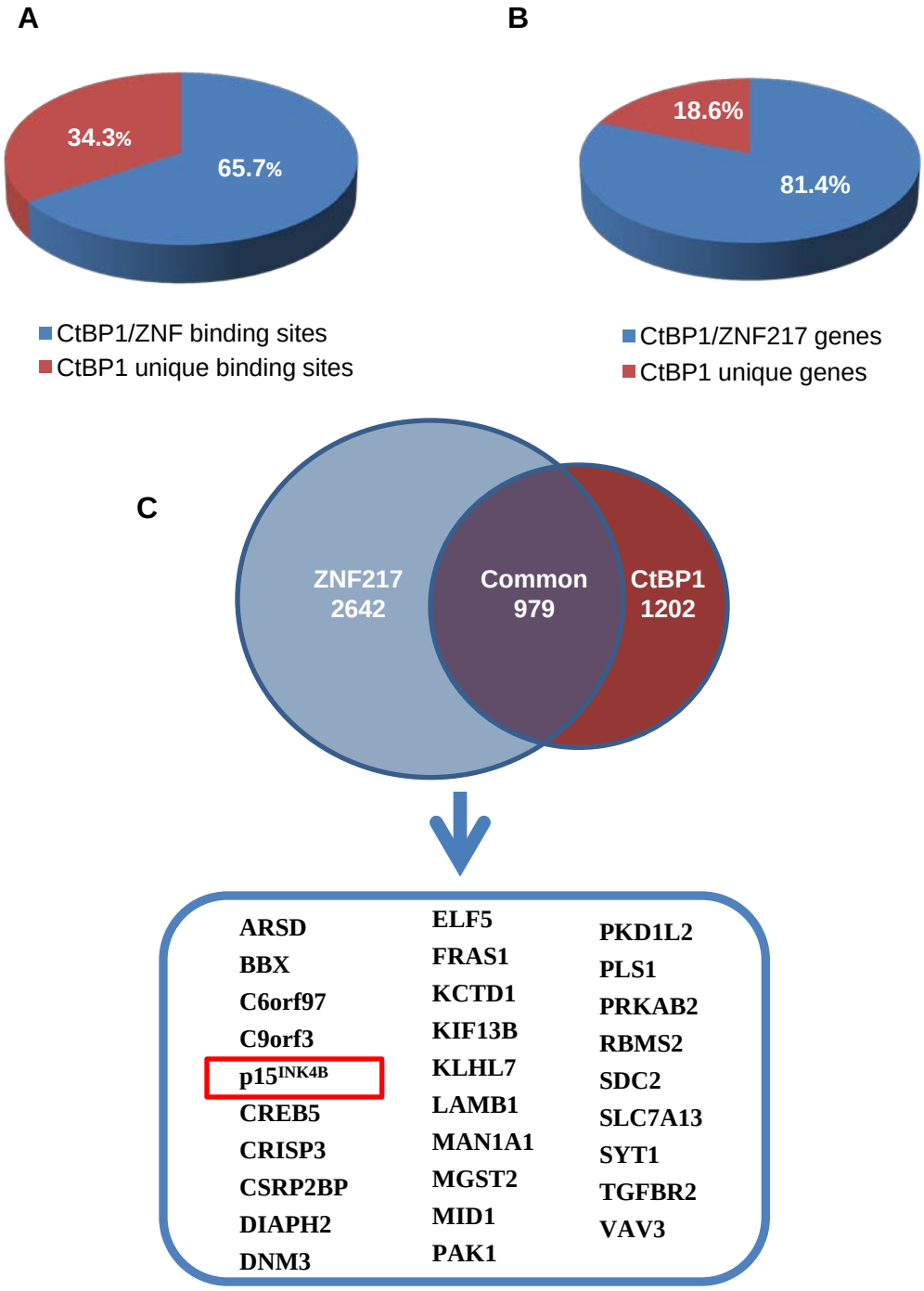
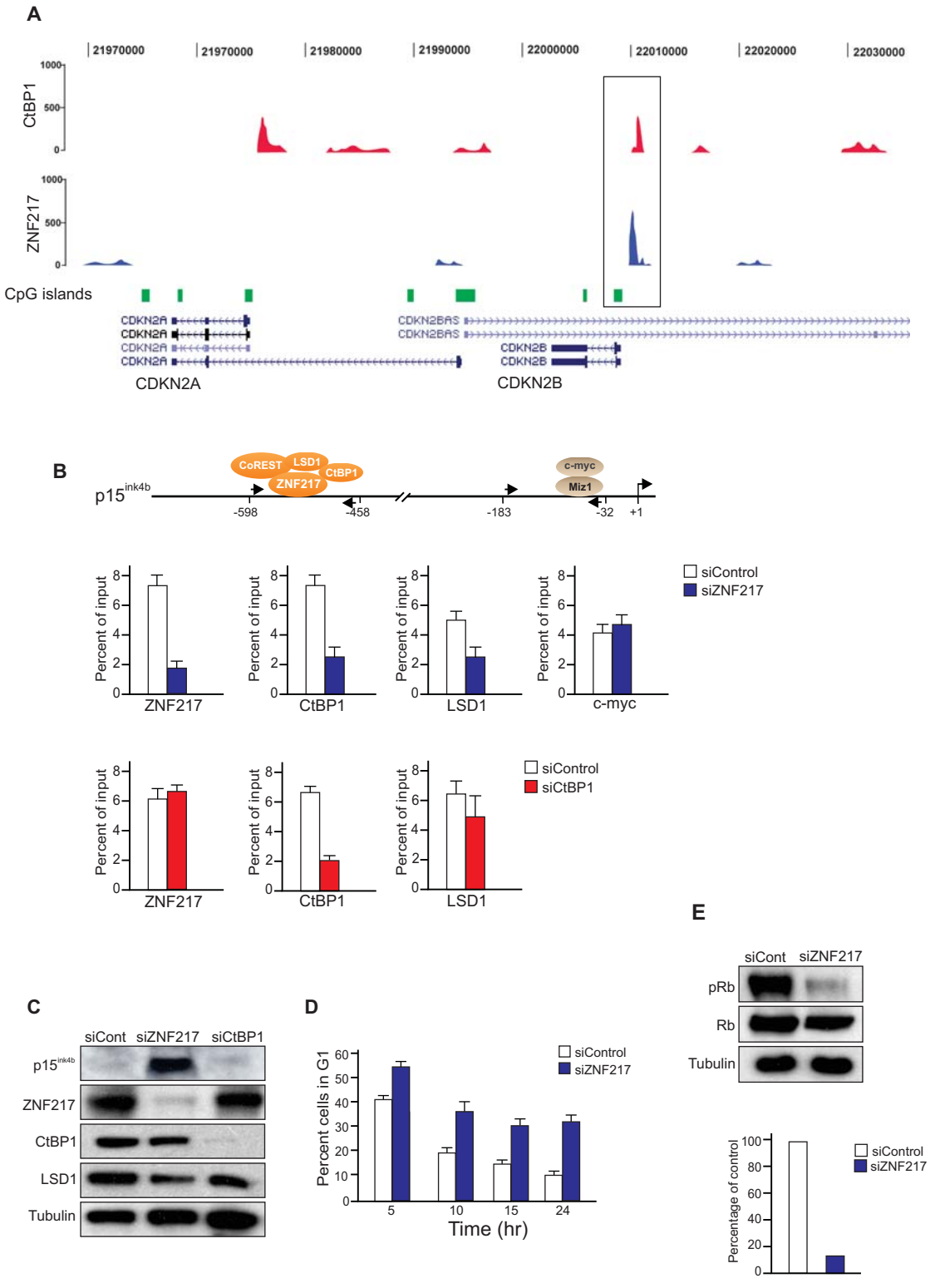

A
34.3%
65.7%
CtBP1/ZNF binding sites
CtBP1 unique binding sites
B
18.6%
81.4%
CtBP1/ZNF217 genes
CtBP1 unique genes
C
ZNF217
2642
Common
979
CtBP1
1202
ARSD
BBX
C6orf97
C9orf3
p15INK4B
CREB5
CRISP3
CSRP2BP
DIAPH2
DNM3
ELF5
FRAS1
KCTD1
KIF13B
KLHL7
LAMB1
MAN1A1
MGST2
MID1
PAK1
PKD1L2
PLS1
PRKAB2
RBMS2
SDC2
SLC7A13
SYT1
TGFBR2
VAV3

Figure 3.3: Transcriptional repression of the $p15^{ink4b}$ gene by the ZNF217/CoREST complex.

(A) ChIP-seq tracings for CtBP1 and ZNF217 at the ink 4 locus. The square indicates enrichment of ZNF217 and CtBP1 to the same region of the $p15^{ink4b}$ (CDKN2B) promoter. Genes have been correlated to fit the UCSC annotations. (B) ZNF217 is required for targeting the CoREST complex. MCF7 cells were transfected with the indicated siRNA and ChIP-qPCR analysis of the $p15^{ink4b}$ promoter was performed using the indicated antibodies. Error bars indicate standard error of the mean. (C) Western blotting analysis of specific proteins (as indicated on the left) from MCF7 cells following transfection with siRNAs for ZNF217 or CtBP1. (D) Knockdown of ZNF217 causes an increase in the percentage of cells found in G1. Cells were transfected with the indicated siRNAs and treated with nocodazole for 24 hrs. The percentage of cells in G1 was then determined by flow cytometry. Shown is a representative experiment from 3 independent experiments as depicted in Figure 3.4. (E) Knockdown of ZNF217 causes a decrease in phosphorylated Rb. MCF7 cells were transfected with the indicated siRNAs. Endogenous Rb status was analyzed by western blotting using the indicated antibodies as shown on the left. The amount of phosphorylated Rb (pRb) was quantified using densitometry and normalized to total Rb and Tubulin.


A
21970000
21970000
21980000
21990000
22000000
22010000
22020000
22030000
CtBP1
ZNF217
CpG islands
CDKN2A
CDKN2B
B
p15ink4b
CoREST
LSD1
CtBP1
ZNF217
c-myc
Miz1
-598
-458
-183
-32
+1
Percent of input
siControl
siZNF217
siCtBP1
ZNF217
CtBP1
LSD1
c-myc
C
siCont
siZNF217
siCtBP1
p15ink4b
ZNF217
CtBP1
LSD1
Tubulin
D
Percent cells in G1
Time (hr)
E
pRb
Rb
Tubulin
Percentage of control

knockdown of either ZNF217 or CtBP1 (Figure 3.3B). The recruitment of CtBP1 and LSD1 was reduced by approximately 60% following ZNF217 knockdown, whereas knockdown of CtBP1 did not affect recruitment of either ZNF217 or LSD1 suggesting that ZNF217 is required for targeting of the ZNF217/CoREST complex.

We also examined the occupancy by c-myc which is a well-known negative regulator of the $p15^{ink4b}$ gene. c-myc represses $p15^{ink4b}$ expression by inhibiting transactivation of Miz1 at the core promoter region further downstream of the ZNF217 binding site (Seoane et al., 2001; Staller et al., 2001;). However, in MCF7 cells, c-myc was found on the promoter regardless of the presence of ZNF217 suggesting that its role in regulating $p15^{ink4b}$ is independent of the ZNF217/CoREST complex (Figure 3.3B). Expression analysis indicated that knockdown of ZNF217 resulted in robust activation of $p15^{ink4b}$ expression in MCF7 cells (Figure 3.3C). In contrast, knockdown of CtBP1 did not stimulate expression of the $p15^{ink4b}$ gene suggesting that CtBP1 may not be essential for repression, or a compensatory role may be provided by the homologous protein CtBP2.

3.3.2 *ZNF217-Dependent Repression of the $p15^{ink4b}$ Gene Is Essential for Cell-Cycle Progression.*

The $p15^{ink4b}$ gene is a cyclin dependent kinase inhibitor that forms a complex with cdk4 or 6 and prevents their activation resulting in inhibition of cell cycle progression at the G1/S transition. To assess the role of ZNF217 in cell cycle progression we performed flow cytometry and as expected, knockdown of ZNF217 increased the percentage of cells remaining in G1 relative to control (Figure 3.3D and Figure 3.4), and caused an eight-fold reduction in phosphorylated Rb levels (Figure 3.3E). These results suggest that

Figure 3.4: Cytometric analysis of MCF7 cells.

MCF7 cells were transfected with scrambled siRNA or siRNA targeting ZNF217. After 72 hrs cells were treated with nocodazole. At the indicated time points post nocodazole treatment cells were fixed in ethanol, washed twice in PBS containing 1% FBS and stained with propidium iodide (PI).

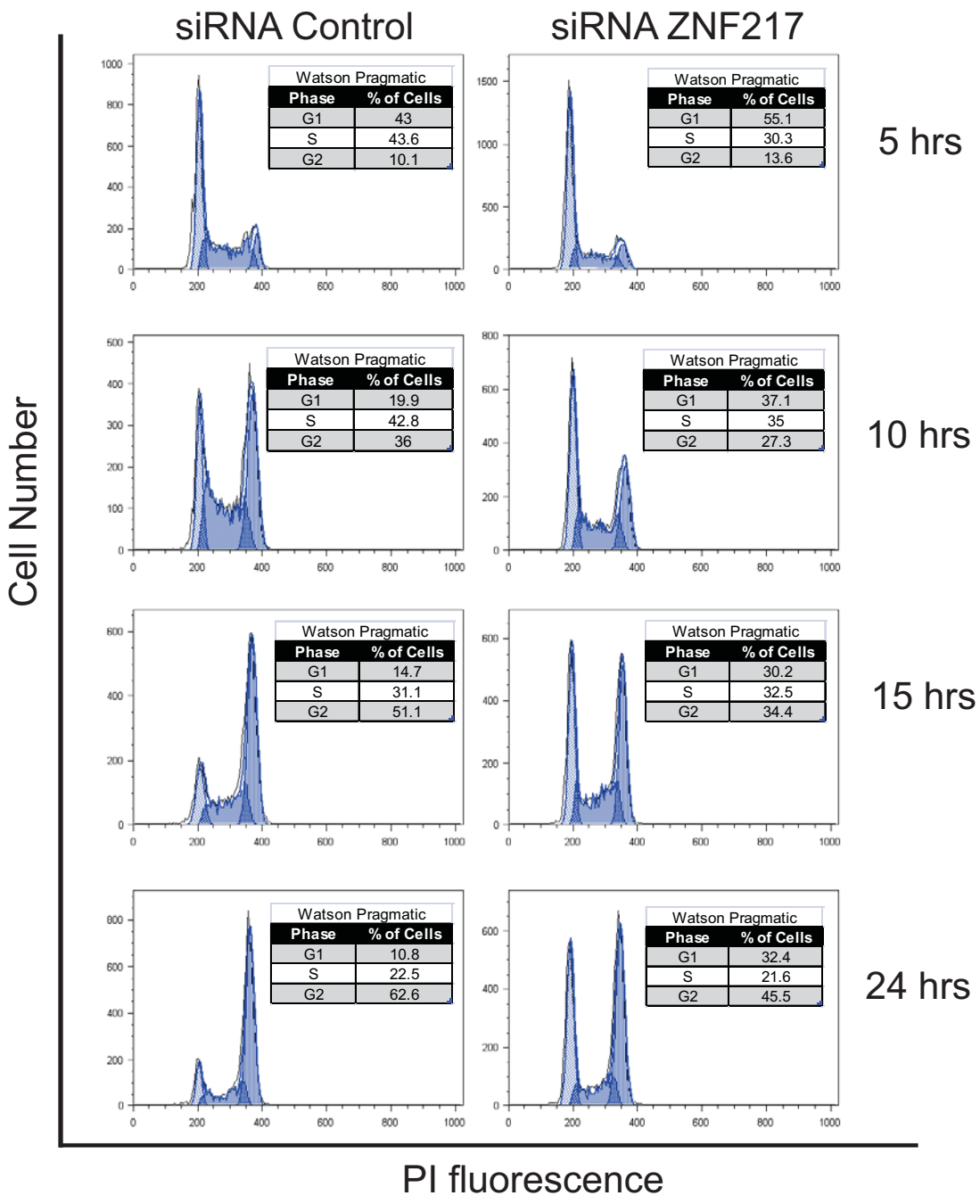
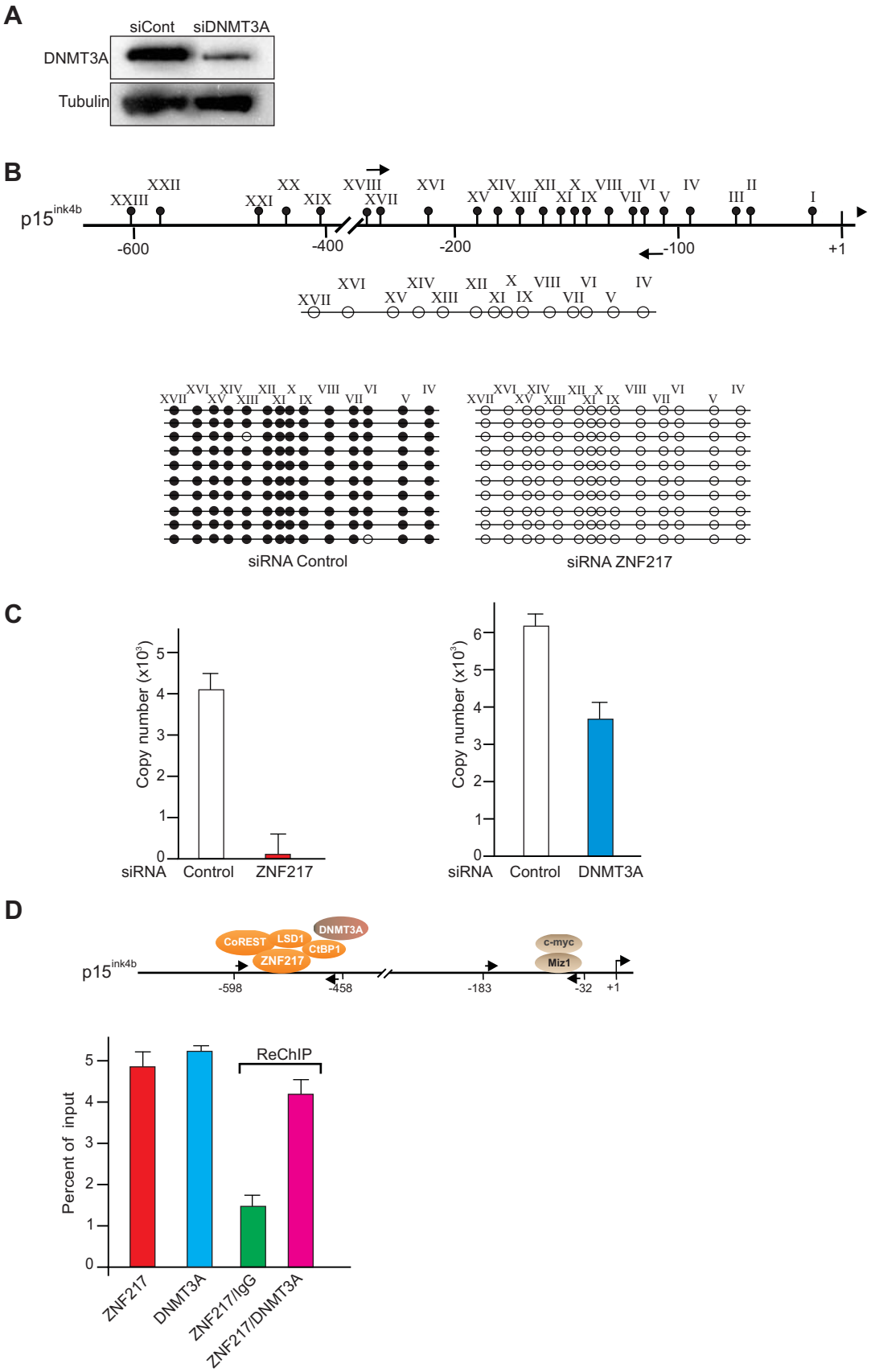

siRNA Control
siRNA ZNF217
5 hrs
10 hrs
15 hrs
24 hrs
Watson Pragmatic
Phase
% of Cells
G1
S
G2
43
43.6
10.1
55.1
30.3
13.6
19.9
42.8
36
37.1
35
27.3
14.7
31.1
51.1
30.2
32.5
34.4
10.8
22.5
62.6
32.4
21.6
45.5
Cell Number
PI fluorescence

Figure 3.5: Knockdown of ZNF217 or DNMT3A causes DNA demethylation of the $p15^{ink4b}$ promoter.

(A) Knockdown of DNMT3A using siRNA. MCF7 cells were transfected with the indicated siRNA and after 72 hrs, cell extracts were prepared and western blot analysis was performed with the indicated antibodies. (B) DNA methylation analysis of the $p15^{ink4b}$ promoter. MCF7 cells were transfected with the indicated siRNAs, DNA was extracted, and the core CpG region of the $p15^{ink4b}$ promoter was analyzed by sodium bisulphite sequencing. The human $p15^{ink4b}$ promoter with the CpGs indicated in Roman numerals is shown. White and black circles indicate unmethylated and methylated cytosines, respectively. (C) Methylation-specific PCR of the $p15^{ink4b}$ promoter. Samples were treated as described in (A) and then analyzed by PCR using specific oligonucleotides recognizing the methylated sequence contained within the core CpG island. Error bars indicate standard error of the mean. (D) ZNF217 and DNMT3A co-occupy the $p15^{ink4b}$ promoter in MCF7 cells. ChIP or sequential ChIP (ChIP-ReChIP) was performed with the indicated antibodies. The IgG control values for ChIP were subtracted from values obtained using specific antibodies. Error bars indicate standard error of the mean.


A
siCont
siDNMT3A
DNMT3A
Tubulin
B
p15ink4b
-600
-400
-200
-100
+1
siRNA Control
siRNA ZNF217
C
Copy number (x10³)
siRNA Control ZNF217
siRNA Control DNMT3A
D
CoREST
LSD1
DNMT3A
CtBP1
ZNF217
c-myc
Miz1
-598
-458
-183
-32
+1
ReChIP
Percent of input
ZNF217
DNMT3A
ZNF217/IgG
ZNF217/DNMT3A

repression of $p15^{ink4b}$ by ZNF217/CoREST prevents inhibition of cyclinD-cdk4/6 complexes thus allowing for increased phosphorylation of the retinoblastoma protein (Rb) and G1/S phase progression.

3.3.3 *Methylation of the $p15^{ink4b}$ Promoter Is Dependent on the ZNF217/CoREST Complex.*

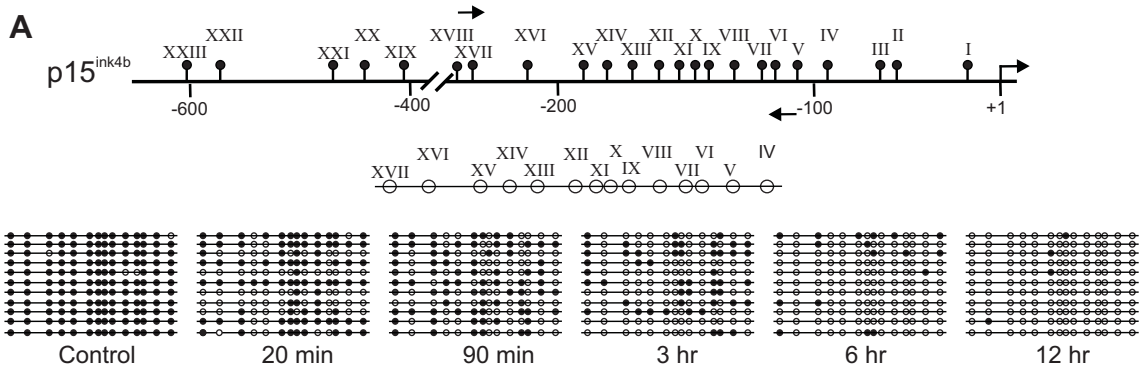
The human $p15^{ink4b}$ promoter contains a CpG island within 1 kb of the transcription start site which is highly methylated in many cancer cell lines (Aggerholm et al., 2006; Papageorgiou et al., 2007) and also contains the ZNF217/CoREST binding region (Figure 3.3A). To assess whether the ZNF217/CoREST complex influences the DNA methylation of the $p15^{ink4b}$ gene, we carried out knockdown experiments and assessed the methylation status of the $p15^{ink4b}$ promoter by sodium bisulphite sequencing and methylation-specific PCR. Knockdown of ZNF217 resulted in dramatic promoter demethylation (greater than 95%) which coincided with increased $p15^{ink4b}$ expression (Figures 3.5A-3.5C). To determine the mechanism involved in maintaining the $p15^{ink4b}$ in a hypermethylated state, we initially examined the role of the *de novo* methyltransferase DNMT3A which is often associated with repression at specific genomic loci (Hervouet et al., 2009). Knockdown of DNMT3A using siRNA reversed hypemethylation of the $p15^{ink4b}$ promoter although the reversal was not complete suggesting that other DNMT family members may also be involved. Sequential ChIP analysis (ChIP-reChIP) using ZNF217 and DNMT3A-specific antibodies indicated that both proteins co-occupy the same region of the promoter in MCF7 cells (Figure 3.5D). These results suggest that the ZNF217/CoREST complex and specific DNMTs combine to generate a hypermethylated state resulting in repression of the $p15^{ink4b}$ gene.

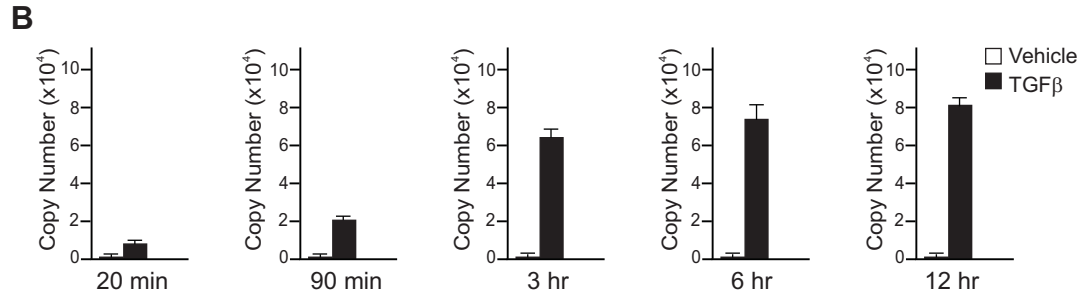
3.3.4 *TGF-β Stimulates Active Demethylation of p15^{ink4b} via Recruitment of TDG.*

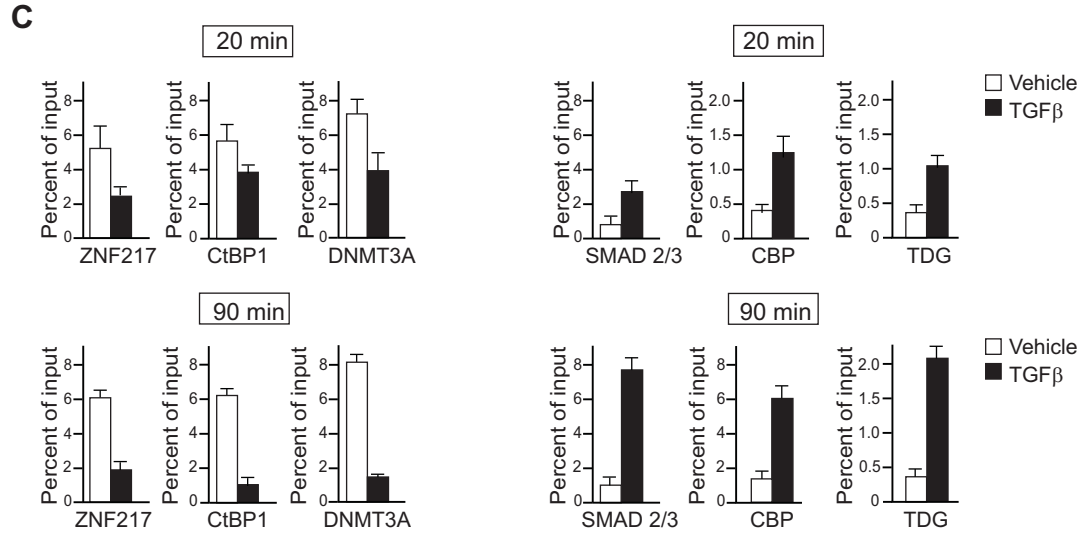
To examine whether the p15^{ink4b} promoter undergoes signal-dependent, temporal changes in DNA methylation, serum-starved HaCAT cells were treated with TGFβ for various time periods followed by sodium bisulphite sequencing and methylation-specific PCR. In untreated cells, the core CpG island is highly methylated and demethylation was detected within 20 minutes after TGFβ treatment while nearly complete demethylation was evident by 3 hours (Figure 3.6A and 3.6B). Similar results were obtained by methylated DNA immunoprecipitation (MeDIP) using a mC-specific antibody (Figure 3.7). Furthermore, DNA demethylation in response to TGFβ was not inhibited by pretreatment of HaCAT cells with the DNA replication inhibitor L-mimosine (Figure 3.8). The rapidity of this response in quiescent cells, and the observation that both TGFβ and L-mimosine arrest HaCAT cells at G1 (Krude, 1999; Warner et al., 1999) strongly indicates that DNA demethylation is active, as opposed to passive demethylation that requires DNA replication. ChIP analysis of the p15^{ink4b} promoter in HaCAT cells treated with TGFβ demonstrated loss of ZNF217, CtBP1 and DNMT3A and a concomitant increase in SMAD, CBP and TDG binding to the same promoter region (Figure 3.6C). These results are consistent with earlier findings demonstrating that TGFβ-dependent activation of p15^{ink4b} promotes recruitment of SMAD2/3 and p300 (Seoane et al., 2001; Staller et al., 2001). The coregulator exchange was detected as early as 20 minutes following TGFβ treatment and preceded the large changes in TGFβ-dependent DNA demethylation. ChIP-ReChIP indicated that on the p15^{ink4b} promoter, TDG associates with SMAD proteins as well as CBP (Figure 3.6D and Figure 3.14). Consistent with

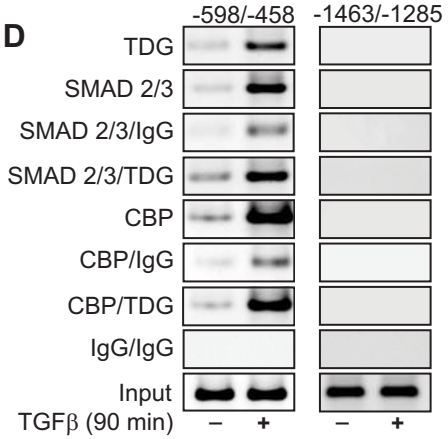
Figure 3.6: TGFβ-dependent demethylation of the p15^{ink4b} promoter.

(A) TGFβ stimulates DNA demethylation of the p15^{ink4b} promoter. HaCAT cells were treated with TGFβ for the indicated time periods and the p15^{ink4b} promoter was analyzed by sodium bisulphite sequencing. White and black circles indicate unmethylated and methylated CpGs, respectively. The human p15^{ink4b} promoter with the CpGs indicated in Roman numerals is shown. (B) PCR analysis of the p15^{ink4b} promoter. Samples were treated as described in (A) and then analyzed by PCR using specific oligonucleotides recognizing the unmethylated sequence within the core CpG island as indicated by the arrows in (A). (C) TGFβ stimulates a coregulator switch at the p15^{ink4b} promoter. HaCAT cells were treated with 150 pM TGFβ for 20 minutes or 90 minutes and ChIP-qPCR was performed with the indicated antibodies. IgG control values for ChIP were subtracted from values obtained using specific antibodies. Error bars indicate standard error of the mean. (D) TDG associates with both CBP and SMAD 2/3 in response to TGFβ. HaCAT cells were treated as in (C) and ChIP was performed using a single antibody, or the combination of antibodies (ChiP-ReChIP) indicated on the left. (E) Co-immunoprecipitation of endogenous SMAD 2/3 and TDG. HaCAT cells were treated with TGFβ for 90 minutes and whole cell extracts were prepared. SMAD 2/3 was then immunoprecipitated and western blotting performed using the TDG antibody.


A
p15ink4b
XXII
XXIII
XX
XXI
XIX
XVIII
XVII
XVI
XIV
XV
XIII
XII
XI
X
IX
VIII
VII
VI
V
IV
III
II
I
-600
-400
-200
-100
+1
Control
20 min
90 min
3 hr
6 hr
12 hr


B
Copy Number (x10⁴)
20 min
90 min
3 hr
6 hr
12 hr
Vehicle
TGFβ


C
20 min
90 min
Percent of input
ZNF217
CtBP1
DNMT3A
SMAD 2/3
CBP
TDG
Vehicle
TGFβ


D
-598/-458
-1463/-1285
TDG
SMAD 2/3
SMAD 2/3/IgG
SMAD 2/3/TDG
CBP
CBP/IgG
CBP/TDG
IgG/IgG
Input
TGFβ (90 min)

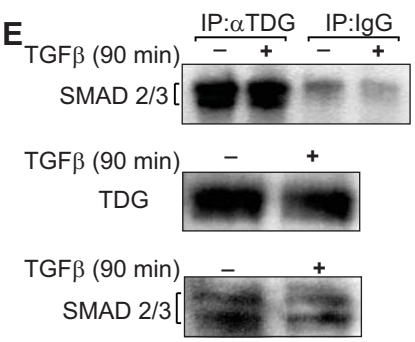

E
IP:αTDG
IP:IgG
TGFβ (90 min)
SMAD 2/3
TDG

Figure 3.7: Time dependent demethylation of the $p15^{ink4b}$ promoter in response to TGFβ.

Serum starved HaCAT cells were treated with TGFβ for the indicated times, cells were then harvested and methyl DNA immunoprecipitation (MeDIP) was performed using a 5mC-specific antibody. The $p15^{ink4b}$ promoter was quantified by qPCR. The control (C) was assayed 360 min after vehicle treatment.

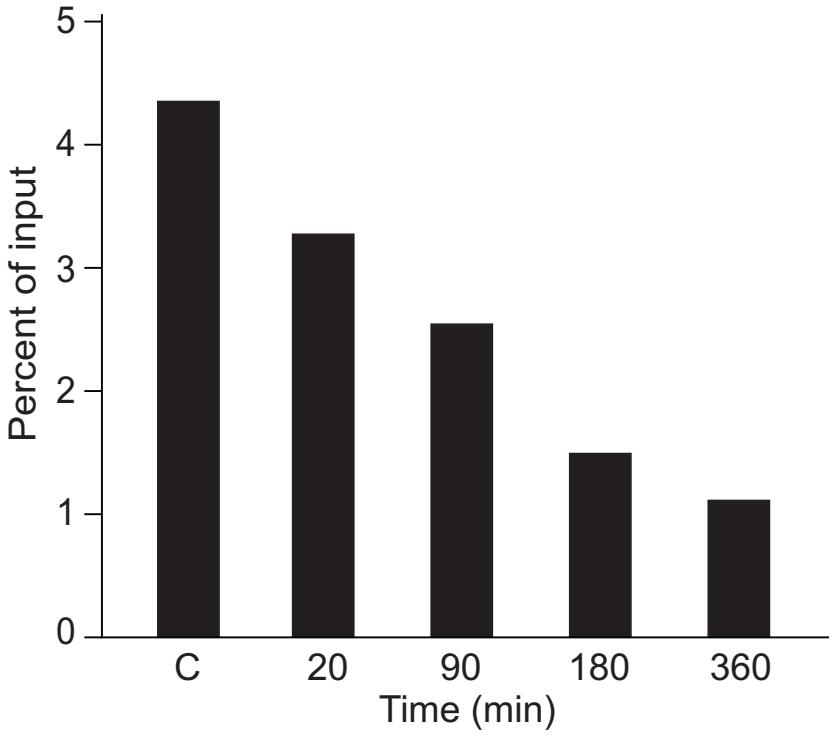
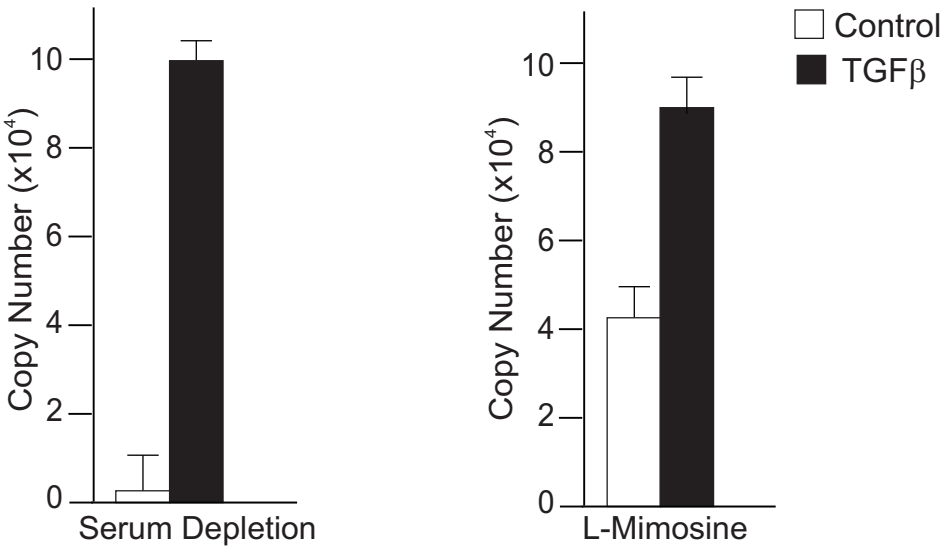

5
4
3
2
1
0
Percent of input
C
20
90
180
360
Time (min)

Figure 3.8: PCR analysis of the $p15^{ink4b}$ promoter.

Cells were serum starved or grown in the presence of L-Mimosine (0.5 mM) for 24 hr prior to TGFβ treatment for 90 minutes. DNA was then isolated sodium bisulphite treated and the $p15^{ink4b}$ promoter was analyzed by qPCR using oligonucleotides recognizing the unmethylated sequence within the core CpG island.


Control
TGFβ
Copy Number (x10^4)
10
8
6
4
2
0
Serum Depletion
Copy Number (x10^4)
10
8
6
4
2
0
L-Mimosine

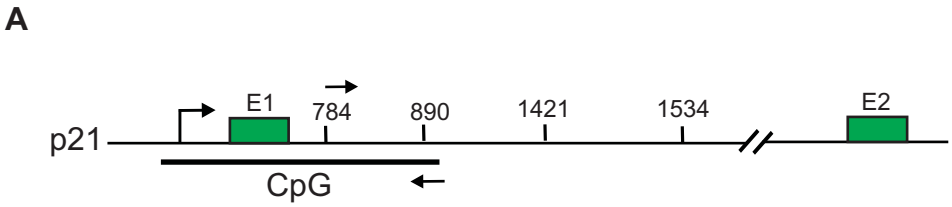
these findings, we were also able to coimmunoprecipitate endogenous TDG and SMAD2/3 from whole cell extracts using TDG-specific antibodies from both control as well as TGFβ-treated cells (Figure 3.6E). These results strongly suggest that that TDG, SMAD2/3 and CBP form multiprotein complex(es) and that targeting of this complex represents a critical step in the induction of $p15^{ink4b}$ expression. We also examined the methylation status of the $p21^{cip1}$ gene. The $p21^{cip1}$ is also a critical regulator of cell cycle progression and is also activated at the transcriptional level by the canonical TGFβ-SMAD signaling pathway (Gomis et al., 2006a). Treatment of either serum starved or L-mimosine-treated HaCAT cells with TGFβ for 90 minutes resulted in rapid DNA demethylation of the $p21^{cip1}$ promoter and a concomitant increase in expression based on qPCR analysis (Figure 3.9).

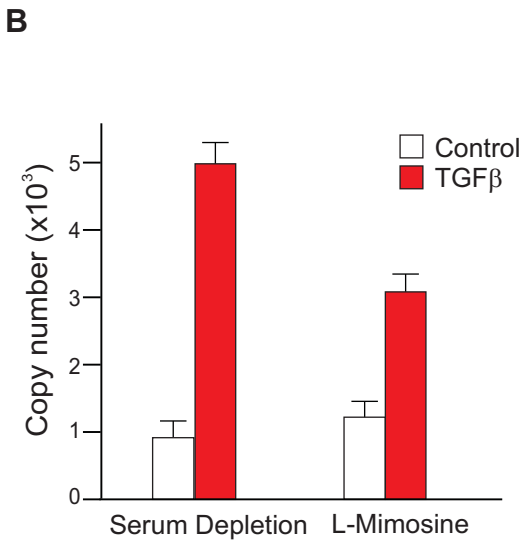
To assess if TDG is required for TGFβ-dependent demethylation of the $p15^{ink4b}$ promoter, HaCAT cells were transfected with TDG-specific siRNA prior to treatment with TGFβ for 90 minutes. TDG knockdown in HaCAT cells abrogated promoter demethylation and decreased $p15^{ink4b}$ mRNA levels by approximately 50% in response to TGFβ treatment (Figures 3.10A-3.10D). These findings demonstrate that TDG regulates TGFβ–dependent promoter demethylation and expression of the $p15^{ink4b}$ tumour suppressor gene.

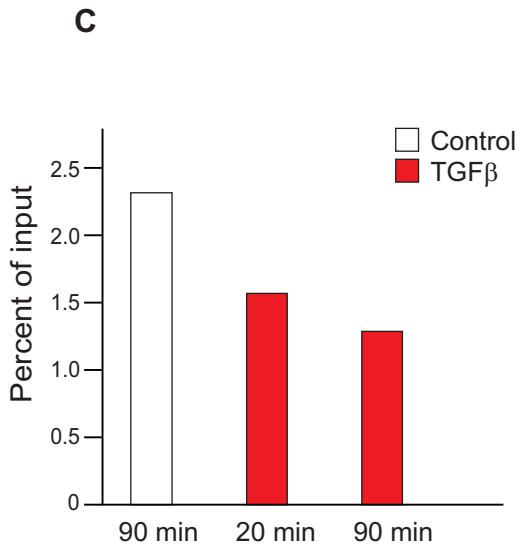
We also performed dot blot analysis of bulk genomic DNA using an anti-5mC antibody to examine whether TGFβ causes global changes in DNA demethylation. Although only semi quantitative, we consistently observed a 2-3 fold decrease in DNA

Figure 3.9: Active demethylation of the p21 gene.

(A) Schematic diagram of the p21 gene. p21 contains one CpG island which extends from the promoter, through exon 1 (E1) and into the first intron. (B) Methylation analysis of the p21 gene. HaCAT cells were serum starved or pretreated with L-Mimosine for 24 hrs prior to treatment with TGFβ for 90 minutes. DNA was isolated, sodium bisulfite treated and the p21gene was analyzed by qPCR with primers recognizing the unmethylated CpG island (784/890). (C) MeDIP analysis of the p21 gene. Cells were serum starved and then treated with TGFβ for the indicated times. (D) RNA analysis of the p21 gene. HaCAT cells were treated with TGFβ for 3hrs, RNA was extracted and the p21 gene was analyzed by qPCR. Error bars indicated standard error of the mean.


A
E1
784
890
1421
1534
E2
p21
CpG


B
Control
TGFβ
Copy number (x10^3)
0
1
2
3
4
5
Serum Depletion
L-Mimosine


C
Control
TGFβ
Percent of input
0
0.5
1.0
1.5
2.0
2.5
90 min
20 min
90 min

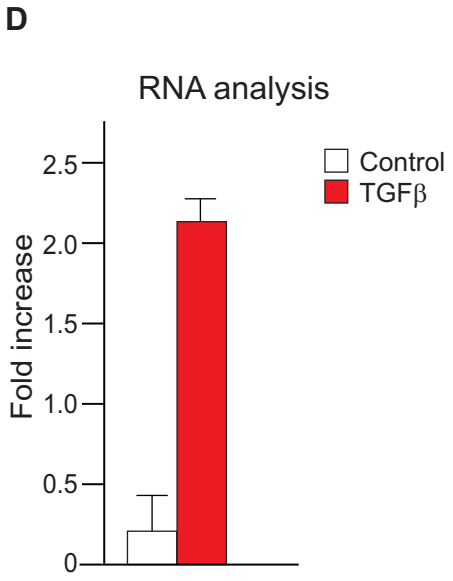
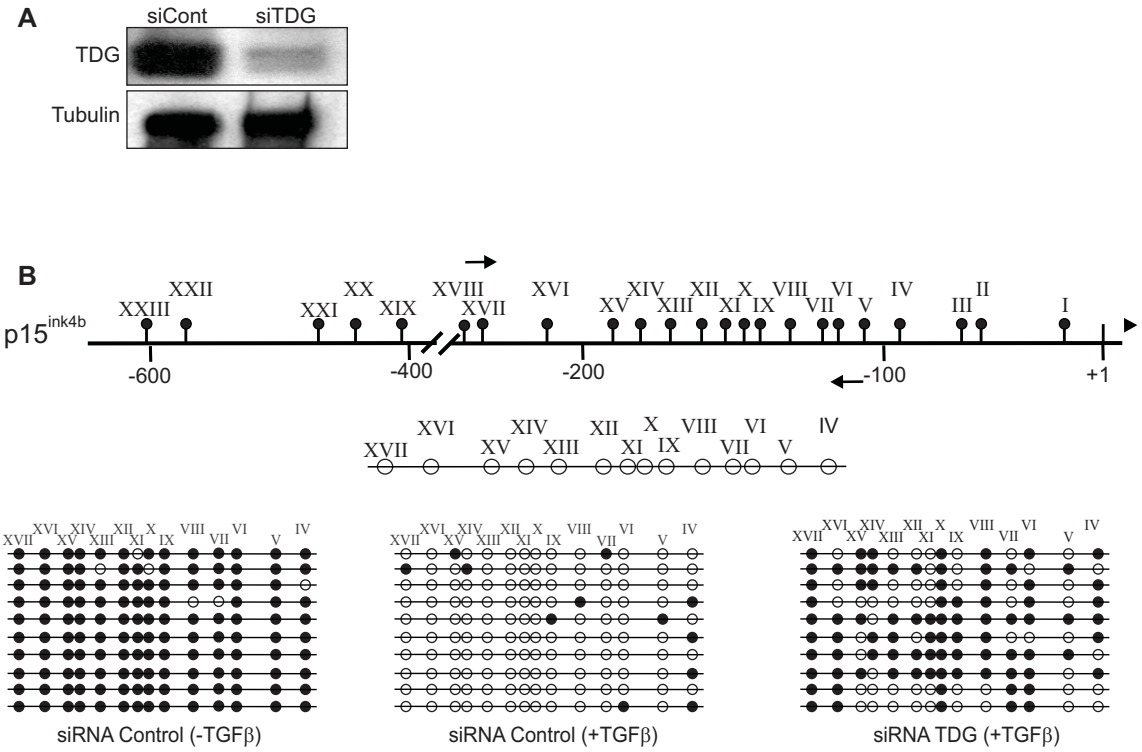

D
RNA analysis
Control
TGFβ
Fold increase
0
0.5
1.0
1.5
2.0
2.5

Figure 3.10: TDG is required for TGFβ-dependent DNA demethylation.

(A) Knockdown of TDG using siRNA. HaCAT cells were transfected with the indicated siRNA and after 72 hr cell extracts were prepared and western blot analysis was performed. (B) TDG knockdown inhibits TGFβ–dependent DNA demethylation**.** HaCAT cells were transfected with the indicated siRNAs and then treated with TGFβ for 90 minutes. The $p15^{ink4b}$ promoter was then analyzed by sodium bisulphite sequencing. White and black circles indicate unmethylated and methylated cytosines, respectively. (C) Methylation-specific PCR of the $p15^{ink4b}$ promoter. HaCAT cells were treated as described in (A) and analyzed by methylation-specific PCR. Error bars indicate standard error of the mean. (D) Knockdown of TDG inhibits $p15^{ink4b}$ expression. HaCAT cells were transfected with the indicated siRNA and then stimulated with TGFβ for 6 hrs. RNA was then extracted and analyzed by qPCR. (E) Dot blot analysis of genomic 5mC. HaCAT cells were treated with TGFβ for 90 minutes. Bulk genomic DNA was isolated, sonicated, crosslinked to nitrocellulose membrane and probed with a monoclonal antibody specific for 5mC. Recognition of DNA by the anti-5mC antibody is shown in the top panel, loading control is shown by the methylene blue stain in the bottom panel. The blot was quantitated by densitometry and a representative experiment is shown from two independent experiments.


A
siCont
siTDG
TDG
Tubulin
B
p15ink4b
-600
-400
-200
-100
+1
siRNA Control (-TGFβ)
siRNA Control (+TGFβ)
siRNA TDG (+TGFβ)
C
Methyl-specific PCR
Vehicle
TGFβ
Copy Number (x10⁴)
siRNA Control TDG
D
RNA analysis
Vehicle
TGFβ
Fold increase
siRNA Control TDG
E
Dot blot
TGFβ + - +
α5mC
Methylene blue
siRNA TDG Control Control
Vehicle
TGFβ
Percentage of control
siRNA Control TDG

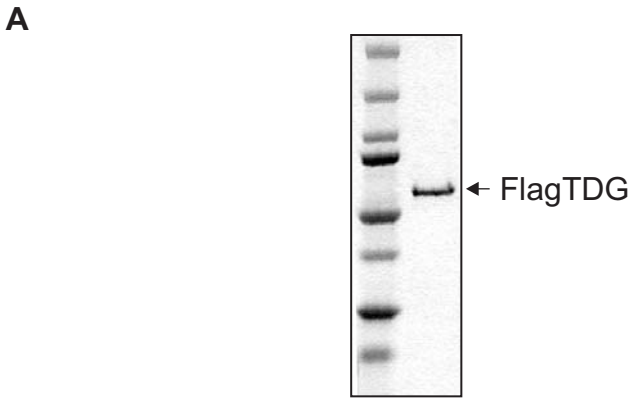
methylation 90 minutes after TGFβ treatment (Figure 3.10E). Remarkably, demethylation was reversed following knockdown of TDG suggesting that the effects of TGFβ on DNA demethylation are genome wide and extend beyond selective promoters.

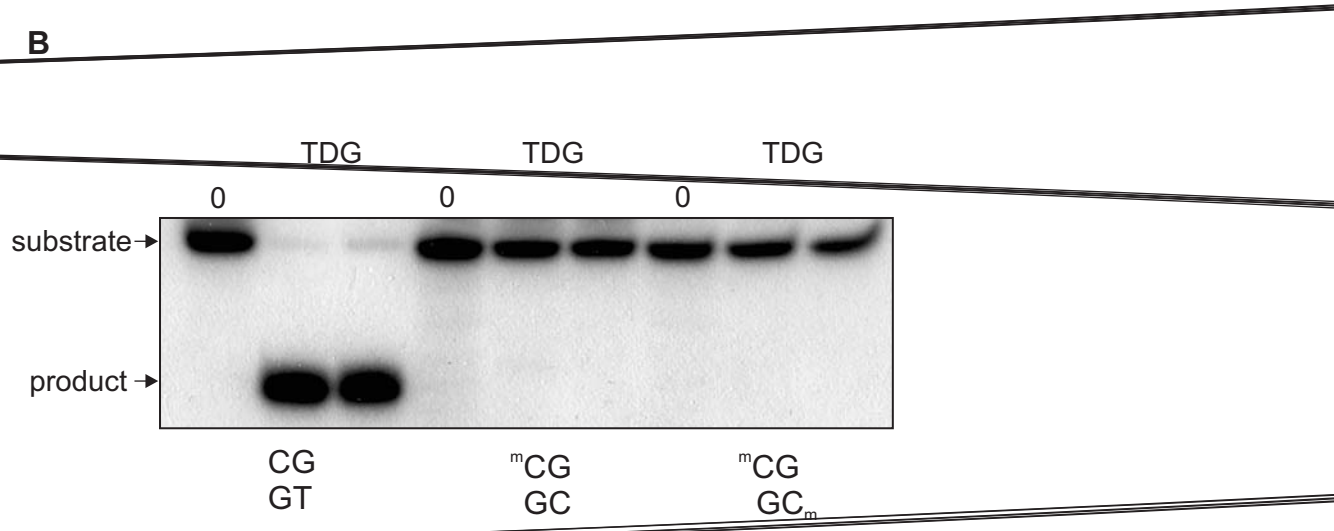
3.3.5 *TGF-β Treatment Stimulates Conversion of 5mC to 5hmC at the $p15^{ink4b}$ Promoter.*

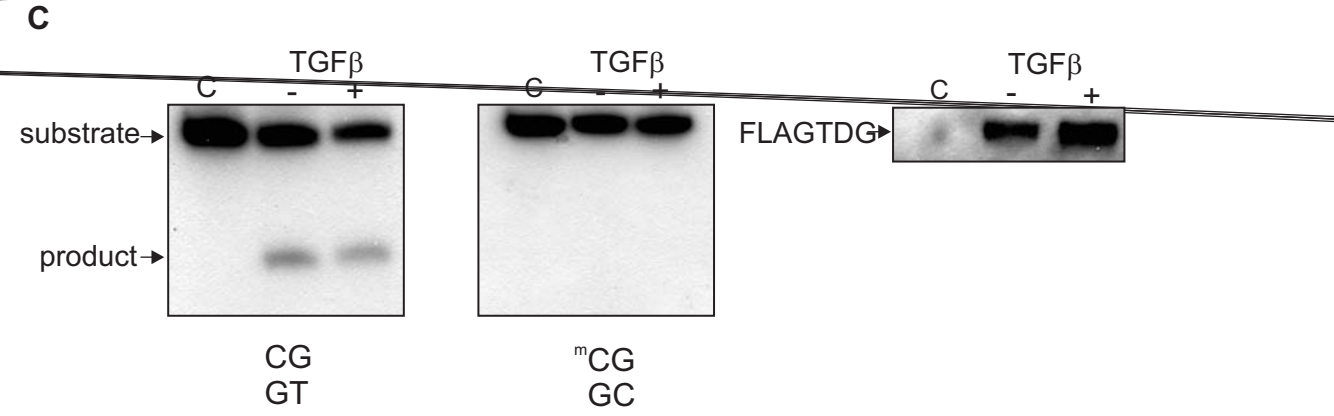
Recent studies have shown that knockout of TDG in mice causes promoter-specific DNA hypermethylation and is embryonic lethal (Cortazar et al., 2011; Cortellino et al., 2011). However, we and others have been unable to demonstrate that purified recombinant TDG possesses intrinsic 5mC demethylase activity *in vitro* (Figure 3.11 and (Cortazar et al., 2011; Cortellino et al., 2011). One possibility is that 5mC is deaminated to thymine by the AID/ABOBEC class of proteins. This would generate a G:T mispair that can be recognized and excised by TDG and subsequently repaired by the BER machinery. Alternatively, mounting evidence suggests that 5mC can be oxidized to 5hmC by TET proteins. Although 5hmC is not a substrate for TDG, it may be deaminated by AID/APOBECs to 5hmU which is recognized and excised by TDG (Cortellino et al., 2011). In this model, we would predict that knockdown of TDG or AID should lead to the accumulation of 5hmC. To assess whether hmC is generated at the $p15^{ink4b}$ promoter, we performed DNA immunoprecipitation assays using antibodies specific for 5mC (MeDIP) or hmC (hMeDIP). Surprisingly, TGF□ treatment following TDG knockdown decreased 5mC levels further (greater demethylation was observed) whereas the 5hmC levels accumulated in the same experiment (Figure 3.12A and 3.12B). A similar trend was observed following the combined knockdown of AID and APOBEC2. These results are consistent with a mechanism suggesting that 5hmC represents an intermediary

Figure 3.11: TDG lacks 5mC DNA glycosylase activity.

Base release assays were conducted with synthetic DNA duplexes containing either a CG/GT mispair, hemimethylated (CG/GCm) or fully methylated (mCG/GCm) CpGs, in the presence of increasing concentrations of purified recombinant TDG generated using baculovirus (A). Shown in (B) is a denaturing polyacrylamide gel with the intact substrate and cleaved protein indicated at the top and bottom of the gel, respectively. (C) Cells were transfected with Flag-tagged TDG and after 24hrs were treated with TGFβ for 90 minutes. Cell extracts were prepared and TDG was immunopurified and tested for base cleavage activity. Western blot of the overexpressed and immunopurified Flag-tagged TDG recombinant protein is shown on the right.


A
FlagTDG


B
TDG
TDG
TDG
0
0
0
substrate
product
CG
GT
mCG
GC
mCG
GCm


C
TGFβ
C
-
+
substrate
product
CG
GT
TGFβ
C
-
+
mCG
GC
TGFβ
C
-
+
FLAGTDG

product during active demethylation at the $p15^{ink4b}$ promoter. ChIP analysis demonstrated that AID is recruited to the $p15^{ink4b}$ promoter in response to TGFβ and sequential ChIP assays using AID and TDG-specific antibodies confirm co-occupancy suggesting that the activities of TDG and AID may be functionally coupled (Figure 3.12C).

Finally, we have also found that the BER enzymes, apurinic/apyrimidinic endonuclease-I (APE-1), DNA ligase I and polymerase (Pol) B are recruited to the $p15^{ink4b}$ promoter following treatment with TGFβ which is consistent with the requirement of the BER pathway for active promoter demethylation (Figure 3.12D).

Like TDG, MBD4 possesses a similar DNA glycosylase activity at G:T mispairs and also plays a role in active demethylation at selective targets (Kim et al., 2009). Remarkably, knockdown of MBD4 also reduced TGFβ-dependent demethylation consistent with the recruitment of this factor to the $p15^{ink4b}$ promoter (Figures 3.13A-3.13C). However, ChIP-ReChIP assays indicated that MBD4 does not co-occupy the $p15^{ink4b}$ promoter with SMAD2/3 suggesting that the roles of MBD4 and TDG may be functionally redundant but that their mechanism of recruitment differs (Figures 3.13C and Figure3.14).

3.3.6 *ZNF217 Overexpression Abrogates $p15^{ink4b}$ Promoter Demethylation and Expression.*

Based on the overlap between ZNF217/CoREST and SMAD/CBP/TDG binding, we speculated that overexpression of ZNF217 may prevent DNA demethylation by interfering with coregulator exchange in response to TGFβ. To test this hypothesis, we infected HaCAT cells with a ZNF217 expressing adenovirus and examined the status of

Figure 3.12: TGFβ-dependent demethylation of the $p15^{ink4b}$ promoter involves conversion of 5mC to 5hmC.

(A) Schematic diagram of the p15 promoter with the CpG islands indicated as a closed ball and stick. (B) DNA immunoprecipitation analysis of the $p15^{ink4b}$ promoter. HaCAT cells were transfected with the indicated siRNAs and then treated with TGFβ or vehicle for 90 minutes. DNA was isolated and DNA immunoprecipitation assays were performed using 5mC or 5hmC-specific antibodies. Aliquots from the same DNA samples were used to perform MeDIP and hMeDIP. The data is expressed as the ratio of the percentage of input DNA from the TGFβ treated and vehicle treated cells from three independent experiments. (C) AID is recruited to the $p15^{ink4b}$ promoter. HaCAT cells were treated with 150 pM TGFβ for 20 mins or 90 mins and ChIP-ReChIP assays followed by qPCR were performed with the indicated antibodies. (D) Components of the BER machinery are recruited to the $p15^{ink4b}$ promoter. Experiments were performed as described in (B) and ChIP assays were performed with the indicated antibodies. Shown are the ChIP assays from samples treated with TGFβ for 90 mins. No significant change in recruitment was observed 20 mins following TGFβ treatment (data not shown). Error bars indicate standard error of the mean.

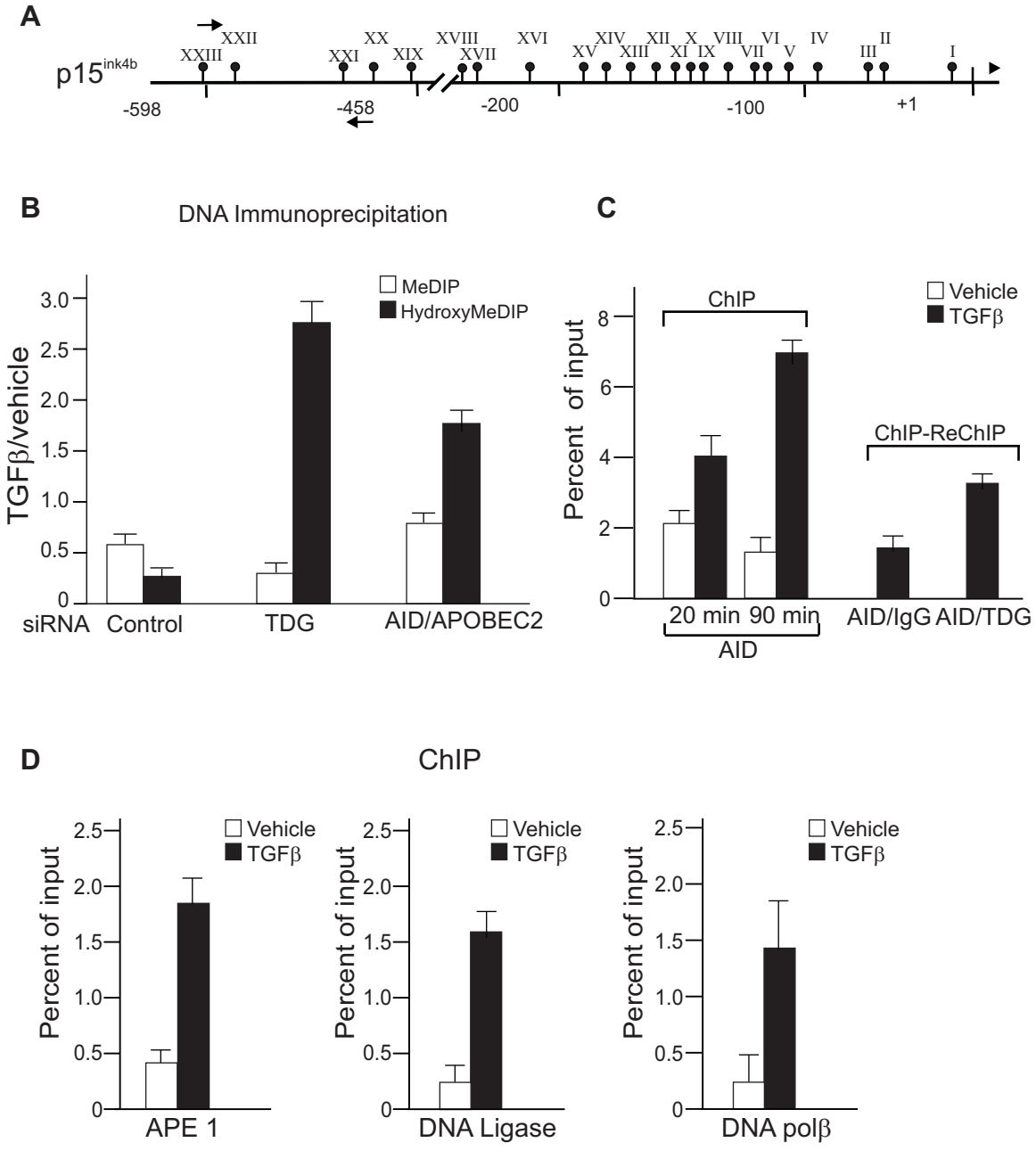

A
p15ink4b
XXII
XXIII
XX
XXI
XIX
XVIII
XVII
XVI
XV
XIV
XIII
XII
XI
X
IX
VIII
VII
VI
V
IV
III
II
I
-598
-458
-200
-100
+1
B
DNA Immunoprecipitation
MeDIP
HydroxyMeDIP
TGFβ/vehicle
siRNA Control
TDG
AID/APOBEC2
C
ChIP
ChIP-ReChIP
Vehicle
TGFβ
Percent of input
20 min
90 min
AID
AID/IgG
AID/TDG
D
ChIP
APE 1
DNA Ligase
DNA polβ

Figure 3.13: MBD4 is required for TGFβ-dependent DNA demethylation.

(A) Knockdown of MBD4 using siRNA. HaCAT cells were transfected with the indicated siRNA and after 72 hrs. cell extracts were prepared and western blot analysis was performed. (B) Knockdown of MBD4 inhibits TGFβ–dependent DNA demethylation**.** HaCAT cells were transfected with the indicated siRNAs and then stimulated with TGFβ for 90 mins. The $p15^{ink4b}$ promoter was analyzed by sodium bisulphite sequencing. White and black circles indicate unmethylated and methylated cytosines, respectively. (C) Methylation-specific PCR of the $p15^{ink4b}$ promoter. HaCAT cells were treated as described in (A) and analyzed by methylation-specific PCR. Error bars indicate standard error of the mean. (D) MBD4 and TDG do not co-occupy the $p15^{ink4b}$ promoter. HaCAT cells were treated with 150 pM TGFβ for 90 minutes and then ChIP or ChIP-ReChIP assays were performed with the indicated antibodies. Error bars indicate standard error of the mean.

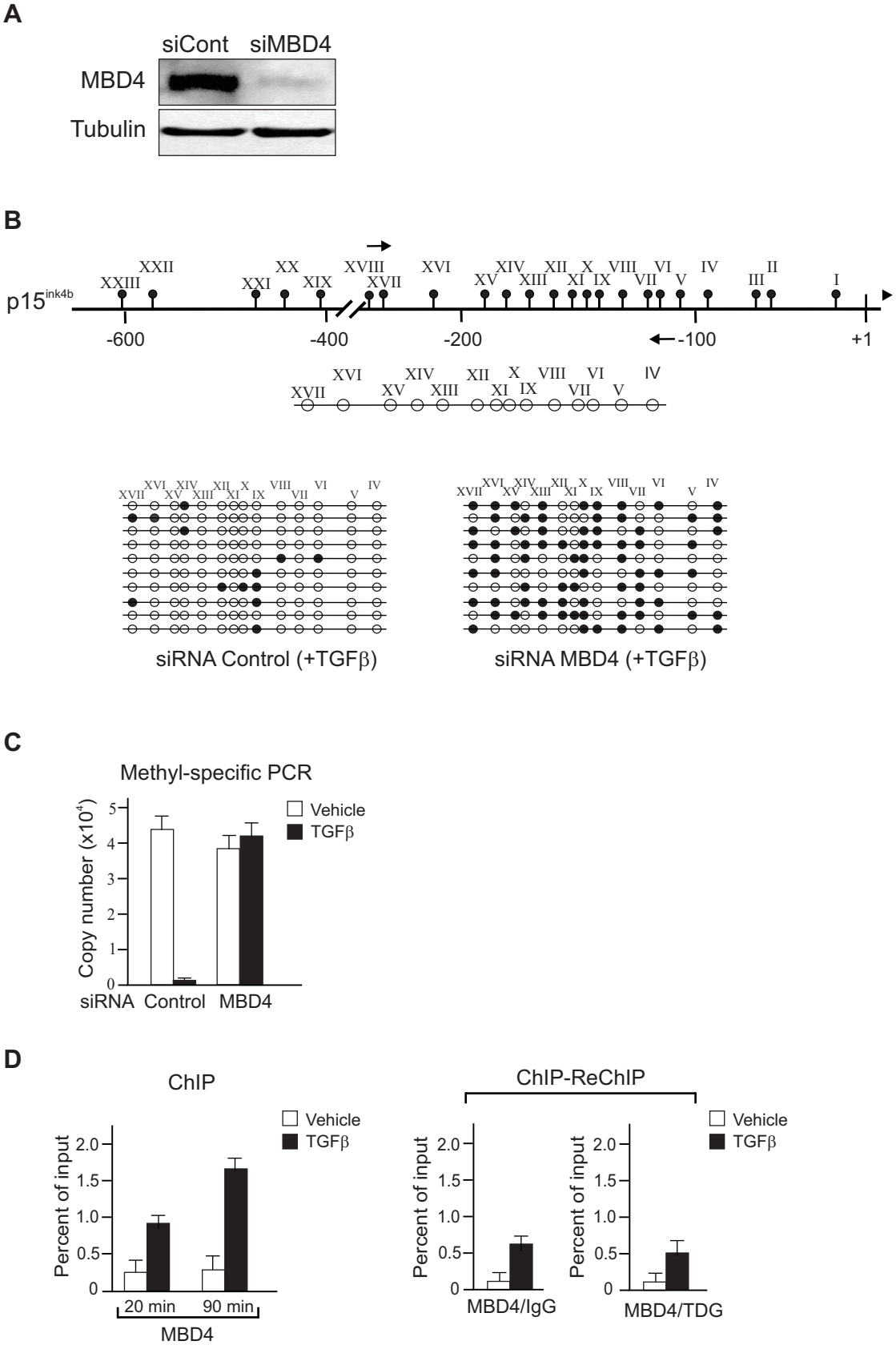
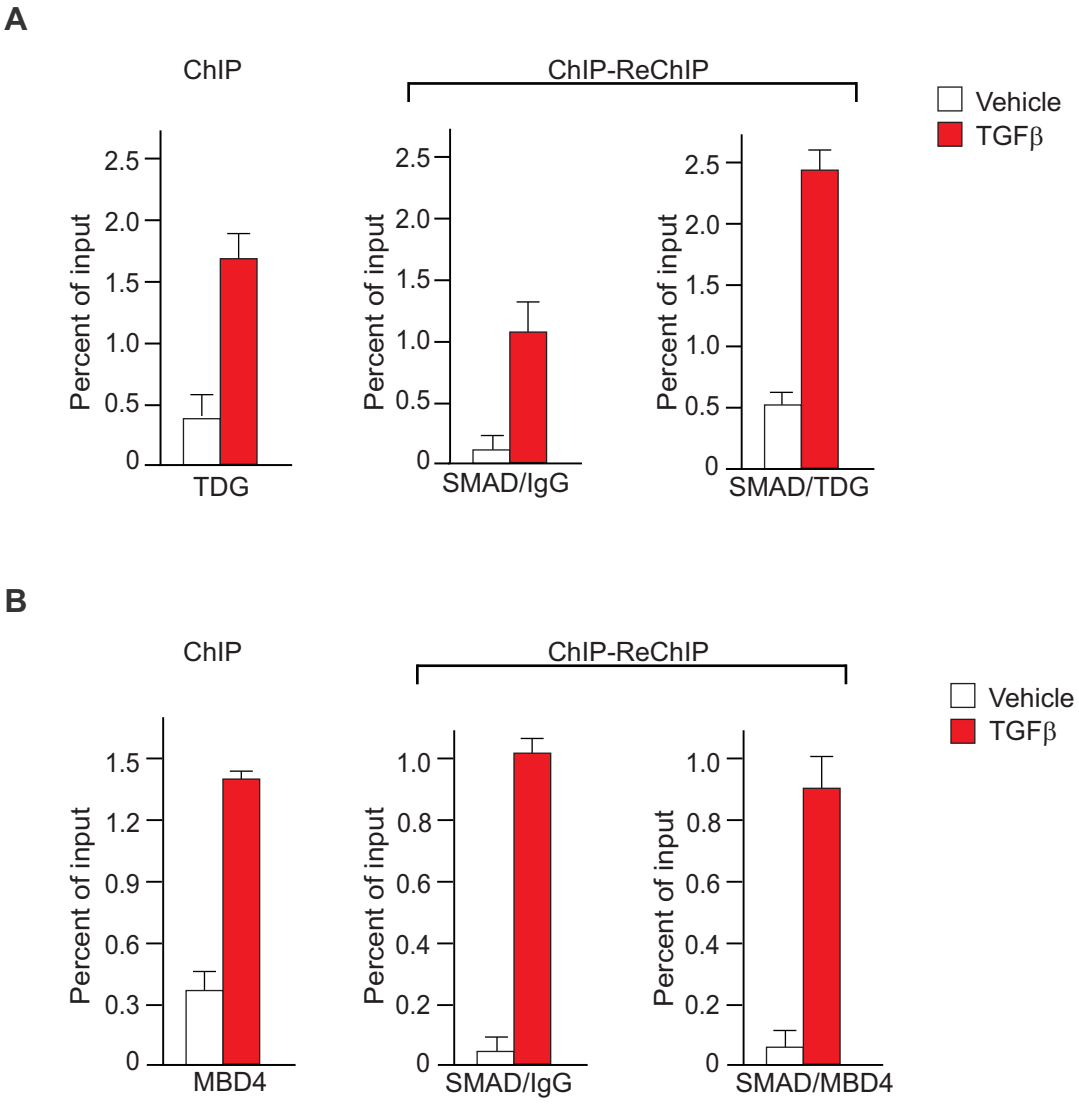

A
siCont
siMBD4
MBD4
Tubulin
B
p15ink4b
-600
-400
-200
-100
+1
siRNA Control (+TGFβ)
siRNA MBD4 (+TGFβ)
C
Methyl-specific PCR
Copy number (x10⁴)
Vehicle
TGFβ
siRNA Control MBD4
D
ChIP
ChIP-ReChIP
Percent of input
Vehicle
TGFβ
20 min
90 min
MBD4
MBD4/IgG
MBD4/TDG

Figure 3.14: Quantitative ChIP-ReCHIP analysis of the p15^{ink4b} promoter.

HaCAT cells were treated with TGFb for 90 minutes and ChIP or ChIP-ReChIP assays were performed with the antibodies indicated and quantified by qPCR. In (A) ChIP-ReChIPs were performed using antibodies against SMAD and TDG (also shown by conventional PCR in Figure 3D) (B) ChIP-ReChIPs were performed using antibodies against SMAD and MBD4. IgG control values were subtracted from values obtained using specific antibodies. Error bars indicate standard error of the mean.


A
ChIP
ChIP-ReChIP
Vehicle
TGFβ
Percent of input
2.5
2.0
1.5
1.0
0.5
0
TDG
SMAD/IgG
SMAD/TDG
B
ChIP
ChIP-ReChIP
Vehicle
TGFβ
Percent of input
1.5
1.2
0.9
0.6
0.3
0
1.0
0.8
0.6
0.4
0.2
MBD4
SMAD/IgG
SMAD/MBD4

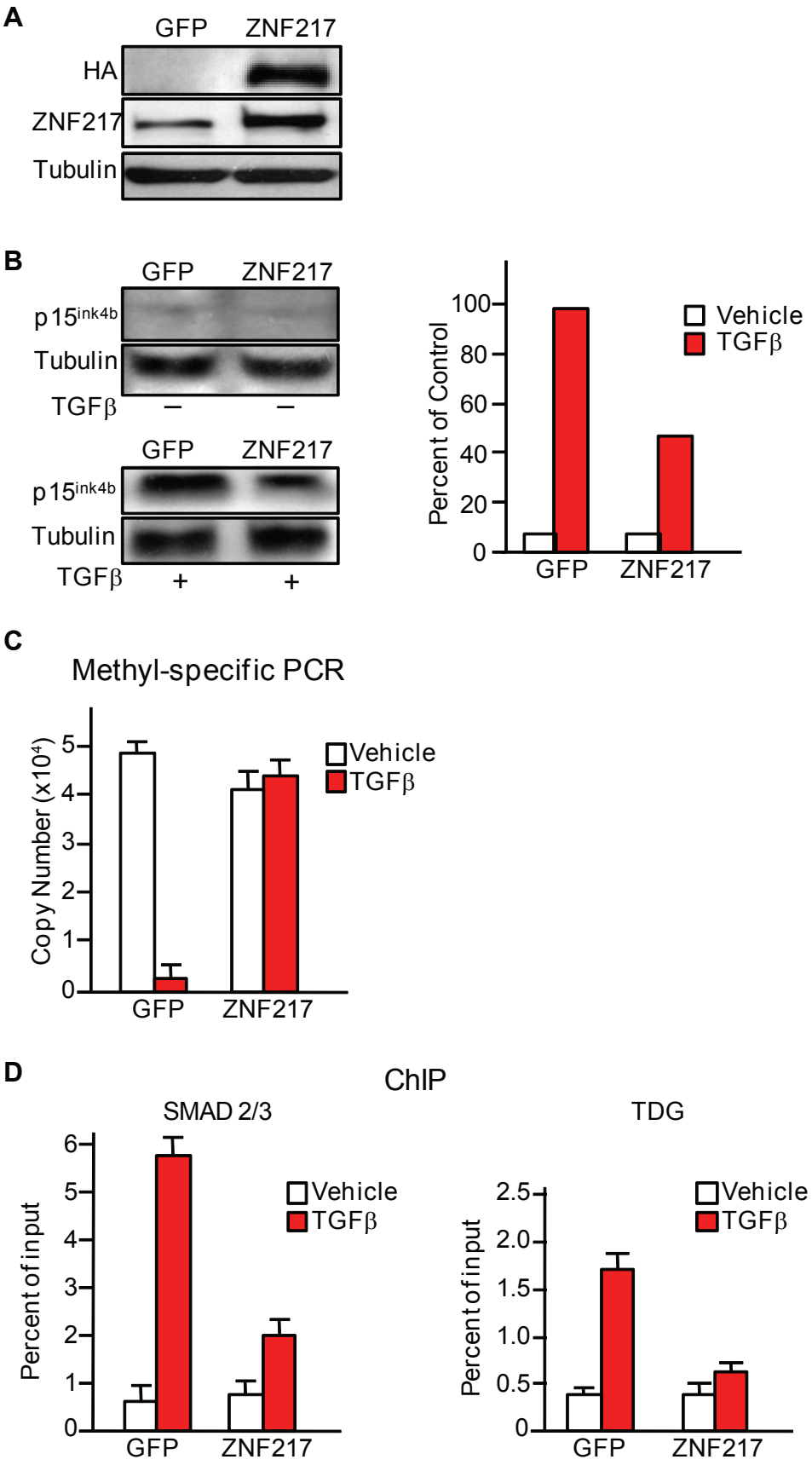
the $p15^{ink4b}$ gene (Figure 3.15A and 3.15B). Overexpression of ZNF217 by approximately 3-fold, resulted in a 60% reduction in TGFβ-dependent induction of the $p15^{ink4b}$ gene and inhibited promoter demethylation based on methylation-specific PCR (Figure 3.15C). To examine the consequences of ZNF217 overexpression on promoter occupancy, we performed ChIP analysis on transduced HaCAT cells which indicated that ZNF217 reduced the recruitment of SMAD2/3 and TDG to the $p15^{ink4b}$ promoter (Figure 3.15D). Collectively, these results indicate that overexpression of ZNF217 is sufficient to impair TDG-dependent DNA demethylase activity by preventing recruitment of critical components required for active DNA demethylation and transcriptional activation of the $p15^{ink4b}$ gene.

3.4 Discussion

The $p15^{ink4b}$ gene is a key regulator of cell proliferation that inhibits cell cycle progression by blocking the activity of cyclin dependent kinases 4 and 6 (Kim and Sharpless, 2006). Herein we demonstrate that transcriptional silencing of the $p15^{ink4b}$ gene by the ZNF217/CoREST complex involves promoter hypermethylation mediated in part by recruitment of DNMT3A. Furthermore, we show for the first time that the TGFβ/SMAD pathway triggers active demethylation and gene expression by targeting the BER enzymes to the $p15^{ink4b}$ promoter. Importantly, overexpression of the ZNF217 oncogene, which is associated with a loss of proliferative control in breast and ovarian cancer, inhibits $p15^{ink4b}$ expression by impairing the TGFβ-dependent recruitment of cofactors involved in active demethylation. These results highlight the dynamic nature of

Figure 3.15: Overexpression of the ZNF217 oncogene inhibits TGFβ–dependent DNA demethylation of the $p15^{ink4b}$ promoter.

(A) Overexpression of HA-tagged ZNF217 in HaCAT cells. HaCAT cells were infected with the indicated adenoviruses (GFP or ZNF217). After 24 hrs, extracts were prepared and western blotting was performed with the indicated antbodies. (B) Overexpression of ZNF217 inhibits $p15^{ink4b}$ protein expression. (Left) HaCAT cells were infected with the indicated adenoviruses and after 24 hrs cells were stimulated with TGFβ and western blotting was performed using the indicated antibodie shown on the left. (Right) The amount of $p15^{ink4b}$ protein expression was quantified using densitometry and normalized to Tubulin. (C) Overexpression of ZNF217 blocks TGFβ-dependent demethylation. Cells were infected with the GFP or ZNF217adenoviruses then stimulated with TGFβ as described for 90 minutes and the $p15^{ink4b}$ promoter was analyzed by methylation-specific PCR. Error bars indicate standard error of the mean. (D) Overexpression of ZNF217 inhibits recruitment of SMAD2/3 and TDG to the the $p15^{inFk4b}$ promoter. HaCAT cells were infected with the GFP or ZNF217 adenoviruses and then stimulated with TGFβ for 90 minutes. ChIP-qPCR was performed using antibodies recognizing SMAD 2/3 or TDG. Error bars indicate standard error of the mean.


A
GFP
ZNF217
HA
ZNF217
Tubulin
B
GFP
ZNF217
p15ink4b
Tubulin
TGFβ
–
–
GFP
ZNF217
p15ink4b
Tubulin
TGFβ
+
+
Percent of Control
100
80
60
40
20
0
Vehicle
TGFβ
GFP
ZNF217
C
Methyl-specific PCR
Copy Number (x10⁴)
5
4
3
2
1
0
Vehicle
TGFβ
GFP
ZNF217
D
ChIP
SMAD 2/3
Percent of input
6
5
4
3
2
1
0
Vehicle
TGFβ
GFP
ZNF217
TDG
Percent of input
2.5
2.0
1.5
1.0
0.5
0
Vehicle
TGFβ
GFP
ZNF217

DNA methylation/demethylation and suggest that coregulator balance is a critical determinant of the methylation status of the $p15^{ink4b}$ promoter, and possibly other tumour suppressor genes (Figure 3.16).

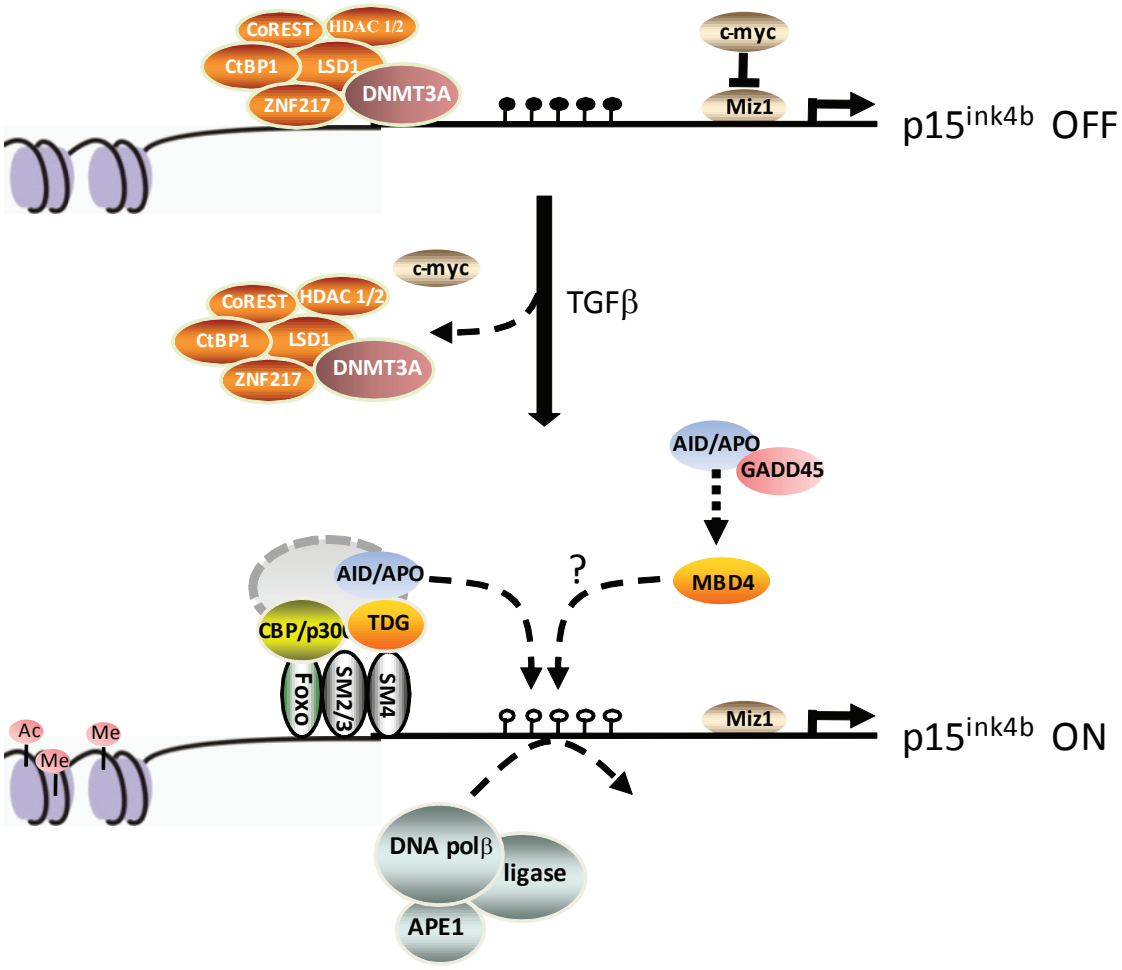
Recent *in vivo* studies in mice have established the importance of DNA glycosylases and BER in active gene-specific demethylation. *Tdg* knockout mice display embryonic lethality and gene expression defects involving hypermethylation of numerous developmentally regulated genes, as well as the loss of active chromatin marks and the accumulation of repressive marks at selected genes (Cortazar et al., 2011). These findings are consistent with a central role of TDG in gene-specific demethylation as well as in other epigenetic events such as the recruitment of transcriptional cofactors CBP/p300. Interestingly, differential gene expression analysis of *Tdg* null MEFs revealed that $p15^{ink4b}$ is significantly downregulated suggesting that cell cycle defects may also contribute to the phenotype observed in the knockout mice (Cortellino et al., 2011).

3.4.1 *Involvement of 5hmC in TGF-β-Dependent Active Demethylation.*

A controversial issue in recent studies has centered specifically on whether DNA glycosylases can directly excise 5mC from DNA, or if enzymatic conversion of 5mC is required prior to its removal. Our data suggests that 5hmC is generated in response to TGFβ and that both 5mC and 5hmC levels decreased at the $p15^{ink4b}$ promoter following TGFβ treatment. However, under TDG-depleted conditions, 5hmC levels accumulated suggesting that TDG most likely functions at a step downstream of 5hmC formation. These effects were not apparent when using bisulphite genomic sequencing because

Figure 3.16: Model depicting the mechanism of TGFβ-dependent demethylation of the $p15^{ink4b}$ promoter.

In normal proliferating epithelial cells, the ZNF217/CoREST/DNMT3A complex is bound to the $p15^{ink4b}$ promoter along with c-myc and the promoter is hypermethylated (filled circles). Stimulation with TGFβ causes release of the ZNF217/CoREST/DNMT3A complex and c-myc and a concomitant binding of activating transcription factors in association with CBP/p300, TDG and AID. 5mC is then oxidized to 5hmC and may undergo deamination by AID/APOBEC proteins to generate 5hydroxyuracil (5hU) which is then processed by TDG and repaired by the BER enzymes (APE1, DNA polymeraseβ and DNA ligase) which reintroduce an unmethylated cytosine (open circles). Alternatively, TGFβ may stimulate recruitment of MBD4 which triggers active demethylation.


CoREST
HDAC 1/2
CtBP1
LSD1
ZNF217
DNMT3A
c-myc
Miz1
p15ink4b OFF
TGFβ
AID/APO
GADD45
MBD4
?
CBP/p30
TDG
FOXO
SM2/3
SM4
Ac
Me
Me
DNA polβ
ligase
APE1
p15ink4b ON

5hmC does not undergo a C-to-T transition after bisulfite treatment, and thus cannot be distinguished from 5mC by that technique (Huang et al., 2010). The accumulation of 5hmC was also observed following the combined knockdown of AID and APOBEC2, although the effects were not as great as those observed following TDG knockdown. Additionally, ChIP analysis suggests that AID and TDG co-occupy the same region of the $p15^{ink4b}$ promoter in response to TGFβ suggesting that deamination may also be required at selected TGFβ-dependent targets.

The involvement of AID/APOBEC2 suggests two possible scenerios for deamination-dependant demethylation at the $p15^{ink4b}$ promoter. First, deamination of 5mC to thymine could generate a G:T mispair which is then recognized and processed by TDG. Such a mechanism infers that deamination and TDG-dependant glycosylation would have to be tightly coupled in order to explain the lack of accumulation of G:T mispairs following knockdown of TDG. A deamination-coupled mechanism is supported by transgenic studies in zebrafish embryos where the presence of a G:T base pair intermediate could only be detected when a catalytically inactive MBD4 mutant was coexpressed with AID (Rai et al., 2008).

Secondly, it has been postulated that 5hmC can be deaminated to produce 5hmU and the resulting 5hmU:G mispair is subsequently repaired by DNA glycosylases and the BER pathway (Cortellino et al., 2011). Although biochemical evidence for the deamination of 5hmC is lacking, recent studies have demonstrated that combined overexpression of TET1 and AID cooperate in demethylating a 5hmC containing DNA duplex (Guo et al., 2011). Furthermore, TDG and MBD4 exhibit robust activity against 5

hmU (Cortellino et al., 2011; Hashimoto et al., 2012). Thus, the observation that 5hmC is generated at the $p15^{ink4b}$ promoter and accumulates under TDG depleted conditions makes this mechanism plausible.

Our results do not discount the possibility that a deaminase-independent demethylation pathway may also play a role in active demethylation of the $p15^{ink4b}$ promoter. 5hmC can undergo iterative oxidation to 5fC and then to 5cC and TDG possesses robust excision activity towards both metabolites (He et al., 2011; Ito et al., 2011; Maiti and Drohat, 2011).

3.4.2 *Requirement for TDG or MBD4 in Active Demethylation.*

A surprising result of our study was that knockdown of TDG or MBD4 prevented TGFβ-dependent active demethylation suggesting a requirement for both DNA glycosylases at the $p15^{ink4b}$ promoter. Furthermore, ChIP assays indicated that TDG and MBD4 were targeted to a similar region of the ink4 locus and with similar kinetics based on two different points. However, the fold enrichment of the ChIP-reChIP analysis was not greater than that observed for each individual ChIP suggesting that MBD4 and TDG do not co-occupy the same promoter region. This suggests that the activities of TDG and MBD4 with respect to DNA demethylation are redundant based on our assay conditions and consequently different population of cells may utilize either TDG or MBD4. A partially redundant role for TDG has also been reported in zebrafish embryos where active DNA demethylation is mediated by the cooperative activities of AID, MBD4/TDG and GADD45α (Rai et al., 2008). However, this conclusion must be interpreted with caution since we have not performed an extensive analysis of MBD4 occupancy at the ink4 locus and the mechanism of MBD4 recruitment is not entirely clear. Unlike TDG,

the recruitment of MBD4 does not appear to involve direct binding with SMAD2/3 upon treatment with TGFβ. Therefore, it is possible that TDG and MBD4 activities are cooperative at some level. This is also supported by the significant differences in their substrate specificities (Cortazar et al., 2007). For example while both TDG and MBD4 recognize AID-generated 5hmU (Hashimoto et al., 2012), MBD4 does not recognize 5fC or 5cC which have been shown to be efficiently processed by TDG (He et al., 2011; Ito et al., 2011). MBD4 also contains a methyl-CpG binding domain (MBD), that is not found in TDG, and which could play a major role in targeting MBD4. Thus, in response to TGFβ, the recruitment of chromatin remodeling factors along with the accumulation of activating histone marks may create a nucleosome free area at the promoter to facilitate binding of MBD4 and, in association with AID and GADD45, promote active demethylation. This mechanism may be facilitated by TDG which can be targeted indirectly to promoter regions by interacting with SMAD2 or 3, CBP or its close homologue p300 (Chiang et al.; Kim and Um, 2008; Li et al., 2007b; Neddermann et al., 1996; Tini et al., 2002). The association between CBP/p300 and TDG has been well documented and the TDG knockout studies in mice supports a mechanism whereby recruitment of CBP/p300 to retinoic acid receptor (RAR) target genes is dependent on TDG (Cortazar et al., 2011; Cortellino et al., 2011; Tini et al., 2002). The present work is consistent with this hypothesis as we have found an association between TDG, SMAD2/3 and CBP on the $p15^{ink4b}$ promoter.

Post translational modifications (PTMs) may also play a critical role in the activities of MBD4 and TDG. A recent study has shown that MBD4 is found in a complex with the vitamin D receptor and the protein kinase C (PKC)-dependent

phosphorylation of MBD4 potentiates its DNA glycosylase activity for 5mC at the CYP27B1 gene promoter (Kim et al., 2009). This was not observed for TDG despite the fact that TDG is also a substrate for PKC (Mohan et al., 2010). In our experiments, overexpression and immunopurification of TDG from TGFβ-treated cells did not stimulate glycosylase activity towards 5mC. Nevertheless, the NH_2-terminal domain (NTD) of mammalian TDG is essential for tight binding to abasic sites and processing of G:T mispairs and undergoes a variety of post translational modifications that dramatically alter interactions with DNA and accessory factors (Mohan et al., 2010; Mohan et al., 2007; Tini et al., 2002). Consequently, PTMs may also play a central role in the substrate recognition and targeting by TDG as well as MBD4 and warrants further investigation.

3.4.3 *Implications of Active Demethylation in TGF-β Signaling Event.*

Active demethylation may have important ramifications with respect to global TGFβ signalling. A CpG rich region of the $p21^{cip1}$ promoter also undergoes active demethylation in response to TGFβ signalling. Furthermore, we observed significant demethylation of bulk genomic DNA from TGFβ-treated cells which surprisingly was TDG-dependant, based on dot blot analysis. Although this may appear somewhat paradoxical when considering the effects observed at the $p15^{ink4b}$ promoter using MeDIP, bulk genomic DNA consists mainly of repetitive elements and transposons and consequently the effects of TGFβ on gene-specific promoter methylation may be different.

Many cancers lose TGFβ dependent functional responses, which in some cases results from a loss of function mutations in various components of this pathway. More

frequently however specific downstream defects are involved. For example, similar to ZNF217, the c-ski oncoprotein is overexpressed in a subset of leukemic patients and negatively regulates TGFβ signaling by interfering with the formation of SMAD complexes at target genes resulting in abnormal silencing of transcription (Lu and Chen, 2003). Overproduction of TGFβ has also been found in many tumours and is associated with induction of epithelial-mesenchymal transition and increased tumour invasiveness (Derynck et al., 1987). Thus, deregulated TGFβ production may contribute to the abnormal DNA methylation patterns associated with malignant transformation.

It has recently been shown that the loss of the Adenomatous polyposis coli (APC) tumour suppressor gene causes upregulation of components of the DNA demethylase machinery resulting in hypomethylation of numerous developmental genes and impaired differentiation (Rai et al., 2010). We have identified signal-dependent mechanisms responsible for epigenetic regulation of the $p15^{ink4b}$ gene which rely on the exchange of a promoter-bound silencing complex with activating transcription factors and components of an active DNA demethylation machinery consisting of DNA glycosylases, other BER enzymes and accessory factors. These findings highlight the expanding role of DNA glycosylases and BER in epigenetic regulation and identify important molecular mechanisms that may be targeted by oncogenic processes.

3.5 References

Aggerholm, A., Holm, M.S., Guldberg, P., Olesen, L.H., and Hokland, P. (2006). Promoter hypermethylation of p15INK4B, HIC1, CDH1, and ER is frequent in myelodysplastic syndrome and predicts poor prognosis in early-stage patients. Eur J Haematol *76*, 23-32.

Bar-Shira, A., Pinthus, J.H., Rozovsky, U., Goldstein, M., Sellers, W.R., Yaron, Y., Eshhar, Z., and Orr-Urtreger, A. (2002). Multiple genes in human 20q13 chromosomal region are involved in an advanced prostate cancer xenograft. Cancer Res *62*, 6803-6807.

Bruniquel, D., and Schwartz, R.H. (2003). Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. Nat Immunol *4*, 235-240.

Chiang, S., Burch, T., Van Domselaar, G., Dick, K., Radziwon, A., Brusnyk, C., Edwards, M.R., Piper, J., Cutts, T., Cao, J., *et al.* (2010). The interaction between thymine DNA glycosylase and nuclear receptor coactivator 3 is required for the transcriptional activation of nuclear hormone receptors. Mol Cell Biochem *333*, 221-232.

Collins, C., Rommens, J.M., Kowbel, D., Godfrey, T., Tanner, M., Hwang, S.I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K., *et al.* (1998). Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. Proc Natl Acad Sci U S A *95*, 8703-8708.

Cortazar, D., Kunz, C., Saito, Y., Steinacher, R., and Schar, P. (2007). The enigmatic thymine DNA glycosylase. DNA Repair (Amst) *6*, 489-504.

Cortazar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., Macdougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., *et al.* (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., *et al.* (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67-79.

Cowger, J.J., Zhao, Q., Isovic, M., and Torchia, J. (2007). Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. Oncogene *26*, 3378-3386.

Derynck, R., Goeddel, D.V., Ullrich, A., Gutterman, J.U., Williams, R.D., Bringman, T.S., and Berger, W.H. (1987). Synthesis of messenger RNAs for transforming growth factors alpha and beta and the epidermal growth factor receptor by human tumors. Cancer Res *47*, 707-712.

Gomis, R.R., Alarcon, C., He, W., Wang, Q., Seoane, J., Lash, A., and Massague, J. (2006a). A FoxO-Smad synexpression group in human keratinocytes. Proc Natl Acad Sci U S A *103*, 12747-12752.

Gomis, R.R., Alarcon, C., Nadal, C., Van Poznak, C., and Massague, J. (2006b). C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. Cancer Cell *10*, 203-214.

Guo, J.U., Su, Y., Zhong, C., Ming, G.L., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423-434.

Hardeland, U., Steinacher, R., Jiricny, J., and Schar, P. (2002). Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J *21*, 1456-1464.

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res, 1-9.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., *et al.* (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303-1307.

Hervouet, E., Vallette, F.M., and Cartron, P.F. (2009). Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. Epigenetics *4*, 487-499.

Huang, Y., Pastor, W.A., Shen, Y., Tahiliani, M., Liu, D.R., and Rao, A. (2010). The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One *5*, e8888.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300-1303.

Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nat Biotechnol *26*, 1293-1300.

Kim, E.J., and Um, S.J. (2008). Thymine-DNA glycosylase interacts with and functions as a coactivator of p53 family proteins. Biochem Biophys Res Commun *377*, 838-842.

Kim, M.S., Kondo, T., Takada, I., Youn, M.Y., Yamamoto, Y., Takahashi, S., Matsumoto, T., Fujiyama, S., Shirode, Y., Yamaoka, I., *et al.* (2009). DNA demethylation in hormone-induced transcriptional derepression. Nature *461*, 1007-1012.

Kim, W.Y., and Sharpless, N.E. (2006). The regulation of INK4/ARF in cancer and aging. Cell *127*, 265-275.

Koinuma, D., Tsutsumi, S., Kamimura, N., Taniguchi, H., Miyazawa, K., Sunamura, M., Imamura, T., Miyazono, K., and Aburatani, H. (2009). Chromatin immunoprecipitation on microarray analysis of Smad2/3 binding sites reveals roles of ETS1 and TFAP2A in transforming growth factor beta signaling. Mol Cell Biol *29*, 172-186.

Krig, S.R., Jin, V.X., Bieda, M.C., O'Geen, H., Yaswen, P., Green, R., and Farnham, P.J. (2007). Identification of genes directly regulated by the oncogene ZNF217 using ChIP-chip assays. J Biol Chem. *282(13*), 9703-9712.

Krude, T. (1999). Mimosine arrests proliferating human cells before onset of DNA replication in a dose-dependent manner. Exp Cell Res *247*, 148-159.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432-435.

Li, P., Maines-Bandiera, S., Kuo, W.L., Guan, Y., Sun, Y., Hills, M., Huang, G., Collins, C.C., Leung, P.C., Gray, J.W., *et al.* (2007a). Multiple roles of the candidate oncogene ZNF217 in ovarian epithelial neoplastic progression. Int J Cancer *120*, 1863-1873.

Li, Y.Q., Zhou, P.Z., Zheng, X.D., Walsh, C.P., and Xu, G.L. (2007b). Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. Nucleic Acids Res *35*, 390-400.

Lu, Z.H., and Chen, J. (2003). [Role of Ski/SnoN protein in the regulation of TGF-beta signal pathway]. Zhongguo Yi Xue Ke Xue Yuan Xue Bao *25*, 233-236.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J Biol Chem *286*, 35334-35338.

Metivier, R., Gallais, R., Tiffoche, C., Le Peron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G., *et al.* (2008). Cyclical DNA methylation of a transcriptionally active promoter. Nature *452*, 45-50.

Mohan, R.D., Litchfield, D.W., Torchia, J., and Tini, M. (2010). Opposing regulatory roles of phosphorylation and acetylation in DNA mispair processing by thymine DNA glycosylase. Nucleic Acids Res *38*, 1135-1148.

Mohan, R.D., Rao, A., Gagliardi, J., and Tini, M. (2007). SUMO-1-dependent allosteric regulation of thymine DNA glycosylase alters subnuclear localization and CBP/p300 recruitment. Mol Cell Biol *27*, 229-243.

Neddermann, P., Gallinari, P., Lettieri, T., Schmid, D., Truong, O., Hsuan, J.J., Wiebauer, K., and Jiricny, J. (1996). Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. J Biol Chem *271*, 12767-12774.

Nonet, G.H., Stampfer, M.R., Chin, K., Gray, J.W., Collins, C.C., and Yaswen, P. (2001). The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. Cancer Res *61*, 1250-1254.

Papageorgiou, S.G., Lambropoulos, S., Pappa, V., Economopoulou, C., Kontsioti, F., Papageorgiou, E., Tsirigotis, P., Dervenoulas, J., and Economopoulos, T. (2007). Hypermethylation of the p15INK4B gene promoter in B-chronic lymphocytic leukemia. Am J Hematol *82*, 824-825.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201-1212.

Rai, K., Sarkar, S., Broadbent, T.J., Voas, M., Grossmann, K.F., Nadauld, L.D., Dehghanizadeh, S., Hagos, F.T., Li, Y., Toth, R.K., *et al.* (2010). DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. Cell *142*, 930-942.

Rooney, P.H., Boonsong, A., McFadyen, M.C., McLeod, H.L., Cassidy, J., Curran, S., and Murray, G.I. (2004). The candidate oncogene ZNF217 is frequently amplified in colon cancer. J Pathol *204*, 282-288.

Seoane, J., Pouponnot, C., Staller, P., Schader, M., Eilers, M., and Massague, J. (2001). TGFbeta influences Myc, Miz-1 and Smad to control the CDK inhibitor p15INK4b. Nat Cell Biol *3*, 400-408.

Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Moroy, T., Bartek, J., Massague, J., Hanel, F., *et al.* (2001). Repression of p15INK4b expression by Myc through association with Miz-1. Nat Cell Biol *3*, 392-399.

Sun, G., Zhou, J., Yin, A., Ding, Y., and Zhong, M. (2008). Silencing of ZNF217 gene influences the biological behavior of a human ovarian cancer cell line. Int J Oncol *32*, 1065-1071.

Thillainadesan, G., Isovic, M., Loney, E., Andrews, J., Tini, M., and Torchia, J. (2008). Genome analysis identifies the p15ink4b tumor suppressor as a direct target of the ZNF217/CoREST complex. Mol Cell Biol *28*, 6066-6077.

Thillainadesan, G., Chitilian, J.M., Isovic, M., Ablack, J.N., Mymryk, J.S., Tini, M., and Torchia, J. (2012). TGF-beta-dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex. Mol Cell *46*, 636-649.

Thollet, A., Vendrell, J.A., Payen, L., Ghayad, S.E., Ben Larbi, S., Grisard, E., Collins, C., Villedieu, M., and Cohen, P.A. (2010). ZNF217 confers resistance to the pro-apoptotic signals of paclitaxel and aberrant expression of Aurora-A in breast cancer cells. Mol Cancer *9*, 291.

Tini, M., Benecke, A., Um, S.J., Torchia, J., Evans, R.M., and Chambon, P. (2002). Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell *9*, 265-277.

Tsellou, E., Troungos, C., Moschovi, M., Athanasiadou-Piperopoulou, F., Polychronopoulou, S., Kosmidis, H., Kalmanti, M., Hatzakis, A., Dessypris, N., Kalofoutis, A., *et al.* (2005). Hypermethylation of CpG islands in the promoter region of the p15INK4B gene in childhood acute leukaemia. Eur J Cancer *41*, 584-589.

van Riggelen, J., Muller, J., Otto, T., Beuger, V., Yetil, A., Choi, P.S., Kosan, C., Moroy, T., Felsher, D.W., and Eilers, M. (2010). The interaction between Myc and Miz1 is required to antagonize TGFbeta-dependent autocrine signaling during lymphoma formation and maintenance. Genes Dev *24*, 1281-1294.

Warner, B.J., Blain, S.W., Seoane, J., and Massague, J. (1999). Myc downregulation by transforming growth factor beta required for activation of the p15(Ink4b) G(1) arrest pathway. Mol Cell Biol *19*, 5913-5922.

Wu, H., and Zhang, Y. (2011). Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes Dev *25*, 2436-2452.

Wu, S.C., and Zhang, Y. (2010). Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol *11*, 607-620.

You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc Natl Acad Sci U S A *98*, 1454-1458.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Chapter 4: General Discussion

4.1 Discussion: Overview

ZNF217 is a Kruppel-like transcription factor and is a constituent of a core transcriptional repressor complex consisting of LSD1, HDAC 1/2, CoREST and CtBP1(Vandevenne et al., 2013). ZNF217 is found within the 20q13.2 region that is frequently amplified in cancers and amplification and overexpression of ZNF217 represents a marker of poor prognosis (Littlepage et al., 2012). Overexpresssion of ZNF217 in normal HMECs leads to cellular immortalization, loss of senescence, increased proliferative capacity and TGFβ resistance which are all hallmarks of cellular transformation (Littlepage et al., 2012; Nonet et al., 2001). In this thesis, I have used global genome-wide approaches to identify novel transcriptional targets for ZNF217 in MCF7 cells. I identified the *p15*^{ink4b}gene as a direct target for ZNF217. I showed that *p15*^{ink4b} is regulated at the transcriptional level by a dyamic interplay between the ZNF217 repressor complex and components of the TGFβ signaling pathway. During the course of my analysis I uncovered a novel mechanism of TGFβ signaling involving active DNA demethylation. I further demonstrated that ZNF217 overexpression impairs the TGFβ dependent transcriptional activation and expression of p15^{ink4b} by blocking DNA demethylation.

4.2 Global Approaches to Identify ZNF217 Targets

Chromatin immunoprecipitation (ChIP) in conjunction with technologies to assess genome-wide occupancy, represent powerful approaches to identify transcription factor binding sites. In chapter 2, I employed a variation of the ChIP-on-chip approach, ChIP-DSL, to identify gene promoters that are bound by ZNF217. ChIP-DSL provides several advantages over the traditional ChIP-on-chip approach. First, traditional ChIP-on-chip

utilizes random oligonucleotides to amplify the ChIP-DNA and this could introduce amplification biases in the technique. ChIP-DSL reduces this bias by using the ChIPed DNA as templates for the selection of predesigned oligonucleotides which are then PCR amplified using the same pair of specific primers. Second, by using unique predesigned oligonucleotides, the approach avoided repetitive sequences that are abundant within the human genome and can interfere with hybridization to the microarray. Third, ChIP-DNA that was not completely de-crosslinked would be more tolerable since the ChIP-DNA was not directly used for PCR amplification which is traditionally used for the ChIP-on-chip approach (Kwon et al., 2007).

Using ChIP-DSL, I identified 1431 gene promoters as ZNF217 targets but only 9 were found to be directly repressed by ZNF217 and 45 targets that were activated. The relatively low number of promoters was surprising because we expected ZNF217 a large set of genes to collectively confer oncogenesis. However, a limitation of ChIP-DSL is that it identifies binding sites that are only found 1 kb upstream of the TSS and studies examining genome-wide binding of other TFs such as ERα have shown that only 3-8% of the ER binding sites are found within 1 kb and approximately only 4% of the targets were transcriptionally regulated by ERα (Kwon et al., 2007). Furthermore, a recent genome analysis using 125 different cell and tissue types have annotated DNase I hypersensitive sites (DHSs) within the human genome that indicates euchromatic regions. These regions were found to encompass many regulatory genomic elements including enhancers, promoters, insulators, and silencers. This analysis indicates that with respect to the TSS, approximately 3% of DHSs are localized to the TSS, 5% within 2.5Kbs and 95% are positioned further away (Thurman et al., 2012). The ChIP-DSL approach would

have failed to identify many targets where ZNF217 binds to promoter regions that are greater than 1 kb from the TSS.

To circumvent this problem, I used ChIP sequencing (ChIPseq) to identify ZNF217 binding sites genome-wide at single nucleotide resolution. The single nucleotide resolution consequently allows for easier validation and characterization of putative DNA binding motifs following mapping of the DNA to the genome. ChIP-DNA with a high degree of repetitiveness can also be better represented using ChIPseq compared ChIP-on-chip method since the ChIP-on-chip method relies on hybridization to complementary strands. Futhermore, ChIPseq has a very large dynamic range with the capability to detect regions of DNA bound lightly and regions with highly enriched binding. For ChIP on chip approaches, abundantly bound ChIPed DNA can often confer intense signal that generally has a bleaching effect and sparsely bound regions are too low for detection (Park, 2009). Due to the dynamic nature and high resolution mapping of the binding site, bound regions can be quantitated to confer strength of binding. Using this approach I was able to expand the number of genes whose promoters were occupied by ZNF217 (within 5Kb from TSS). I also mapped the binding sites of both ZNF217 and CtBP1 throughout the genome which indicated a significant overlap and this observation is consistent with earlier biochemical assays showing that they are components of the same complexes (Cowger et al., 2007; Quinlan et al., 2006). Surprisingly, only a small overlap between ChIP-DSL and ChIPseq was observed and the reasons for this are unclear but may be a result of the differences between the two assays, as discussed above. Nevertheless, both assays confirmed the binding of ZNF217 to a similar region of the $p15^{ink4b}$ promoter.

Importantly, the ChIPseq analysis excluded binding to the other regions at the *ink4* locus, highlighting that the binding is unique to the p15ink4b gene.

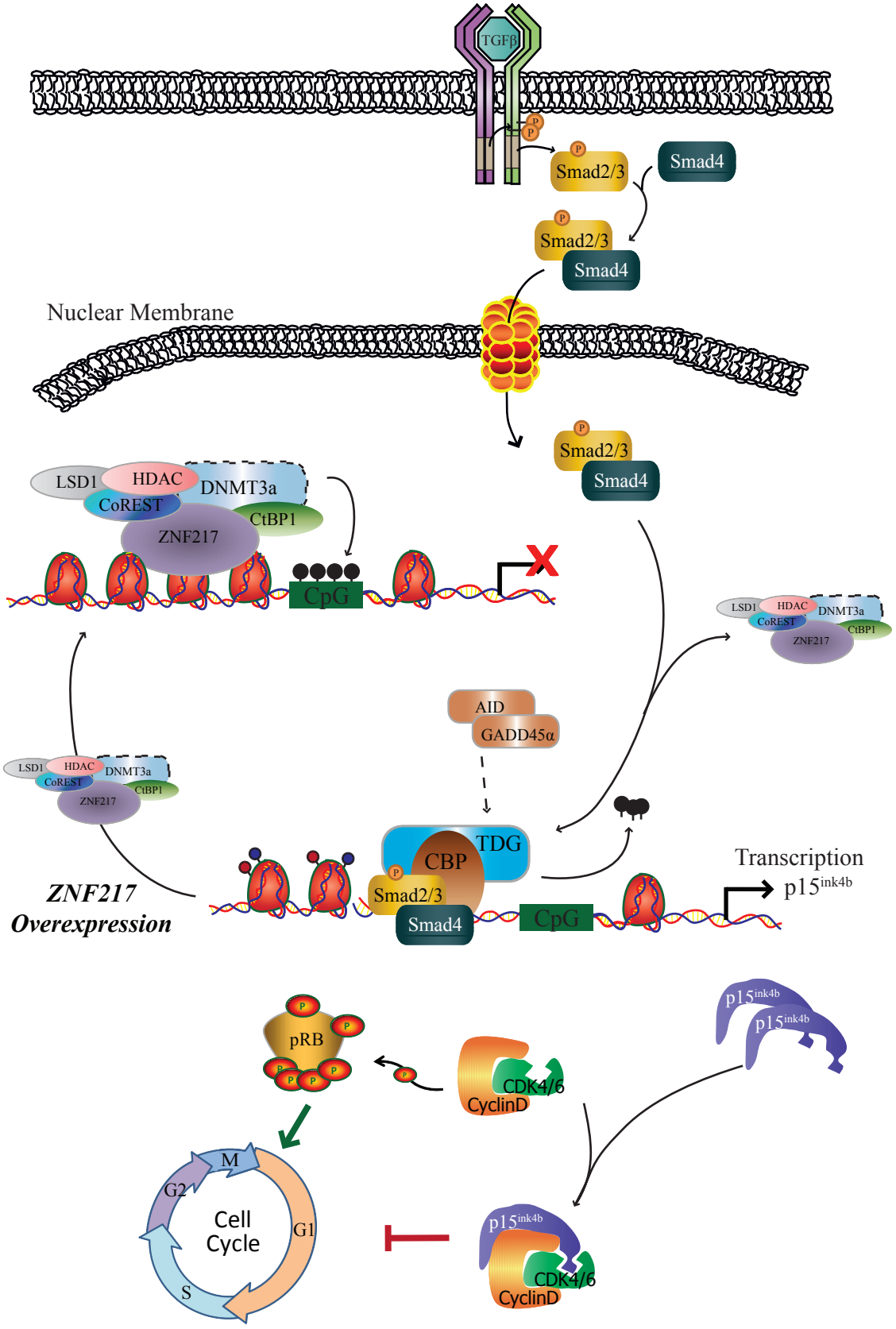
The identification of $p15^{ink4b}$ is consistent with the FACS analysis which demonstrates that ZNF217 overexpression facilitates the transition of cells from G1 to S phase and is in agreement with published reports in other cell lines indicating that overexpression of ZNF217 causes an increase in cell proliferation (Thollet et al., 2010; Vendrell et al., 2012). Silencing of ZNF217 in MCF7 cells also dramatically decreased the phosphorylation of the Rb protein suggesting that ZNF217 promotes the transition through G1/S by maintaining Rb in a hyperphosphorylated state. It is possible that the hyposphosphorylated state of Rb and the G1 arrest observed following ZNF217 downregulation could be as a result of $p15^{ink4b}$ expression (Figure 4.1).

4.3 Transcriptional Regulation of $p15^{ink4b}$

The $p15^{ink4b}$ gene is part of the *ink4* locus, which is encoded on chromosome 9p21, and also contains $p16^{ink4a}$ and $p14^{ARF}$ genes. $p15^{ink4b}$ and $p16^{ink4a}$ are often silenced through DNA hypermethylation and co-deleted in many cancers (Simboeck et al., 2011). The expression of genes in the *ink4* locus provides an effective barrier against cellular hyperproliferation that is associated with many cancers and as a result, the transcription of the locus is tightly regulated. For example, the polycomb group complex consisting of PRC1 and PRC2 are important repressors of the *ink4* locus in many cell types (Bracken et al., 2007). Surprisingly, knockout of $p15^{ink4b}$ in mice did not result in tumor formation and MEFs were still responsive to TGFβ signaling. This suggests that there may be compensatory mechanisms. However, deletion of both $p16^{ink4a}$ and $p15^{ink4b}$ significantly

Figure 4.1: Model of $p15^{ink4b}$ regulation by ZNF217 and TGFβ.

Normally, the $p15^{ink4b}$ promoter is bound by the ZNF217 repressor complex and its expression is repressed. The repression is mediated by the removal of H3K4me by LSD1 and H3K9/14Ac by HDAC1/2 and promoter hypermethylation at the CpG island by DNMT3A. TGFβ stimulation releases the ZNF217 repressor complex and promotes the binding of SMAD/TDG/CBP activation complex to the same region of the promoter. TDG along with AID and TET facilitates active DNA demethylation of the promoter and CBP acetylates the histones leading to the activation of $p15^{ink4b}$. $p15^{ink4b}$ can inhibit cyclin D-CDK4/6 complex. This results in the hypophosphorylation of Rb and inhibition of the cell cycle at G1. However, overexpression of ZNF217 can prevent the binding of SMAD/TDG/CBP complex to the promoter leading to the repression of p15ink4b. This can lead to the activation of cyclin D-CDK4/6 resulting in the phosphorylation of Rb and ultimately G1/S transition. Black circles represent DNA methylation. Blue and Red circles indicate H3K4me and H3K9/14Ac. Green arrow represents activation of the cell cycle. Red arrow indicates inhibition of the cell cycle.


TGFβ
Smad2/3
Smad4
Smad2/3
Smad4
Nuclear Membrane
Smad2/3
Smad4
LSD1
HDAC
DNMT3a
CoREST
CtBP1
ZNF217
CpG
LSD1
HDAC
DNMT3a
CoREST
CtBP1
ZNF217
AID
GADD45α
LSD1
HDAC
DNMT3a
CoREST
CtBP1
ZNF217
TDG
CBP
Smad2/3
Smad4
CpG
Transcription
p15ink4b
ZNF217
Overexpression
pRB
CDK4/6
CyclinD
p15ink4b
p15ink4b
M
G2
Cell
Cycle
G1
S
p15ink4b
CDK4/6
CyclinD

increased the tumor incidence suggesting functional redundancy (Krimpenfort et al., 2007; Ortega et al., 2002).

Expression of $p15^{ink4b}$ is regulated primarily at the transcriptional level and is cell type specific (Figure 4.1) (Siegel and Massague, 2003). The $p15^{ink4b}$ proximal promoter contains two SMAD binding regions (SBR), SBR1 (-538/506) and SBR2 (-443/-385). SBR1 is flanked by the fork-head binding element. TGFβ treatment promotes binding of the SMAD2/3 and SMAD4 heterodimers to the promoter and, along with FOXO factors, facilitate the transcription of $p15^{ink4b}$. ZNF217 binds within −566 to −426 of the $p15^{ink4b}$ promoter which contains the ZNF217 consensus binding sites that overlap with the SMAD2/3 and FOXO binding elements suggesting that binding between the SMAD2/3 and ZNF217 is mutually exclusive.

A complementary mechanism has been postulated for $p15^{ink4b}$ regulation in some cell types and involves the MIZ1 transcription factor which binds to the Inr element found at the proximal promoter region (-150/+75) and activates $p15^{ink4b}$(Figure 4.1). MIZ1 is repressed by the interaction with MYC, which prevents recruitment of coactivators, such as p300 (Staller et al., 2001). TGFβ treatment promotes the downregulation and dissociation of MYC from MIZ1, relieving the repression, and along with binding of SMADS, FOXO and SP1 transcription factors facilitates the activation of the $p15^{ink4b}$ gene.

In this study, knockdown of ZNF217 caused a robust increase in $p15^{ink4b}$ expression but did not affect the occupancy by MYC at the promoter, demonstrating that ZNF217-mediated repression of $p15^{ink4b}$ is not dependent on MYC binding. Although the

reason for this discrepancy is unclear it is possible that ZNF217 and MYC occupy the promoter of *p15*^{ink4b} and repress its expression in distinct cell populations. Alternatively, ZNF217 and MYC may both be required for the repression of *p15*^{ink4b} and the loss of either protein can lead to the relief of repression. Based on the established role that MYC plays in regulating p15^{ink4b} in some cell types, this observation will require further investigation.

4.4 ZNF217 Represses *p15*^{ink4b} through Chromatin Modifications

Treating cells with formaldehyde crosslinks protein-DNA and protein-protein interactions, so it is formally possible that ZNF217 could bind to the *p15*^{ink4b} promoter through association with other transcription factors and that CtBP1may bridge this association through its PXDLS binding cleft (Quinlan et al., 2007). However, the ZNF217 binding region at the p15^{ink4b} promoter contains both of the consensus binding sites identified for ZNF217. Furthermore, I have demonstrated that knockdown of ZNF217 causes the loss of promoter occupancy by the other constituents of the ZNF217 complex such as LSD1 and CtBP1 whereas CtBP1 knockdown did not affect the recruitment of either ZNF217 or LSD1. This suggests that ZNF217 is required for the recruitment of the chromatin modifying enzymes to the *p15*^{ink4b} promoter and ZNF217's recruitment is independent of CtBP1. Consistent with this, recent crystallographic studies demonstrate a direct interaction between ZNF217 and DNA (Nunez et al., 2011; Vandevenne et al., 2013).

I have also demonstrated that the binding of the ZNF217 complex is associated with the lack of activating chromatin marks that coincide with the inactive state of p15^{ink4b}. In proliferating MCF7 cells H3K4 is unmethylated and knockdown of ZNF217

caused a dramatic increase in H3K4me2 and a modest increase in H3K9/14 acetylation at the proximal promoter region of *p15*^{ink4b}. Surprisingly, there was a robust decrease in acetylation following ZNF217 silencing near the TSS that coincided with transcriptional activation of *p15*^{ink4b}. H3K9/14 acetylation is normally associated with transcriptional activation and the reasons for this paradoxical result are unclear. It is possible however, that the H3K9/14 acetylation at the TSS serves as a docking site for certain BRD-containing proteins that facilitate repression. Alternatively, activated gene promoters recruit many proteins such as the GTFs to the TSS and it is possible that downregulation of ZNF217 does infact increase histone acetylation but our ChIP assay was unable to detect this modification due to masking by the occupying proteins (Agalioti et al., 2002).

Our analysis was limited to the region encompassing the ZNF217 binding site. However, it has been documented that these marks can rapidly spread over extended regions in either direction. During thymocyte maturation, the silencing of the mouse terminal transferase (*dntt*) gene was initiated by the removal of H3K9Ac and H3K4me and a gain of H3K9me at a specific region of the promoter (~200bp) at approximately 50bp upstream of the TSS. The pattern of these repressive marks then spread rapidly over large regions of DNA upstream and downstream from the TSS to a maximum distance of approximately 12Kbs (Su et al., 2004).

The spreading of these marks may be mediated by chromatin modifying enzymes that can recognize and bind a specific histone tail mark. For example, G9a/GLP proteins bind H3K9me1/2 through a hydrophobic ANKYRIN (ANK) domain and can place mono- or di-methylation marks on adjacent nucleosomes through the catalytic SET domain. G9a/GLP-mediated initiation and spreading of these marks has also been

observed on the *Oct3/4* genes (Collins and Cheng, 2010). Kinetic ChIP experiments following TGFβ treatment could be useful in assessing whether activating chromatin marks that nucleate at the $p15^{ink4b}$ promoter is spread over the entire *ink4* locus.

4.5 Role of Methylation in Transcription

In chapter 3, I demonstrated that expression of the $p15^{ink4b}$ promoter is dynamically regulated and involves active DNA demethylation in response to TGFβ signaling. In dividing HaCAT cells, the expression of $p15^{ink4b}$ is minimal and its proximal promoter is methylated and occupied by the ZNF217/DNMT3A repressor complex.

A point of controversy in recent years has been whether transcriptional repression is a direct consequence of DNA methylation, or whether DNA methylation is a secondary mark that serves to reinforce the state imposed by repressive chromatin modifications. However, multiple studies have demonstrated the dominancy of DNA methylation in transcriptional repression. Treatment of cells with 5-azacytidine results in activation of several repressed genes (Jones, 1985). Transfection of a methylated DNA template, but not the unmethylated template, results in the assembly of highly compacted chromatin structures that are nuclease-resistant and coincides with transcriptional repression (Buschhausen et al., 1987; Keshet et al., 1986). Equivalent transcription was observed immediately after injection of methylated and unmethylated DNA in Xenopus oocyte nuclei. However, following assembly of chromatin, the methylated DNA conferred transcriptional repression coinciding with nuclease resistance (Kass et al., 1997). These studies suggest a direct causal relationship between DNA methylation and transcriptional repression.

However there are also studies which suggest that DNA methylation is a secondary effect that occurs as a consequence of changes in chromatin marks and serves to reinforce this repression. The hypoxanthine guanine phosphoribosyltransferase (*hprt*) gene on the X chromosome is methylated after X chromosome inactivation (Lock et al., 1987). Several studies in cancer cells have demonstrated that repression by the PcG complex, through methylation of H3K27, precedes DNA methylation (Gal-Yam et al., 2008; Ohm et al., 2007; Schlesinger et al., 2007). Embryonic carcinoma cells that differentiate following retinoic acid treatment show that the repression of the pleuripotent genes begins with the appearance of a nucleosome at the *oct4* enhancer and *nanog* promoter followed by the occupancy by DNMT3A and DNA methylation (You et al., 2011). It is clear from all of these studies that although DNA methylation and histone modifications are carried out by different sets of enzymes, there is a tight relationship between these two epigenetic marks that is important for defining transcriptional outcome. My analysis of the $p15^{ink4b}$ promoter demonstrated that both chromatin tail modifications and DNA methylation associated with the transcriptional state of the gene and based on the observations presented above, both of these marks may mutually confer a transcriptional state at $p15^{ink4b}$.

Methylation of H3K4 inhibits DNMT binding and prevents *de novo* methylation of the underlying DNA. Consequently, H3K4me is permissive for transcriptional activation and is often enriched in the promoter regions of actively transcribed genes (Chotalia et al., 2009; Meissner et al., 2008; Ooi et al., 2007). In our study TGFβ treatment of HaCAT cells resulted in the release of DNMT3A from the $p15^{ink4b}$ promoter which coincided with the increase of H3K4me2 mark suggesting the possibility that the

increase in H3K4me2 mark inhibited the binding of DNMT3A. In contrast, knockdown of TDG prevented TGFβ mediated DNA demethylation at the promoter and conseqently inhibited the expression of $p15^{ink4b}$. These observations imply that DNA methylation can act as both an instructive mark and an epigenetic mark to reinforce repression of an already inactive chromatin however it is most likely the cross-talk between DNA methylation and chromatin that collectively confer repression.

4.6 ZNF217/DNMT3A Complex Methylates the *p15^{ink4b}* Promoter

In MCF7 cells the CpG island at the *p15^{ink4b}* promoter is hypermethylated and $p15^{ink4b}$ is silenced. However, downregulation of ZNF217 caused a dramatic DNA demethylation of the promoter which coincided with robust expression of *p15^{ink4b}*. Downregulation of DNMT3A resulted in only a partial decrease in promoter methylation. The reasons for this are not entirely clear but can be attributed to the functional redundancy between DNMT3A and DNMT3B (Kaneda et al., 2004). Using ChIP-reChIP assay, I demonstrated that DNMT3A and ZNF217 colocalize to the same region on the *p15^{ink4b}* promoter. These results suggest a mechanism whereby ZNF217 may recruit DNMT3A and methylate the promoter (Figure 4.1). The formation of chromatin modifying complexes containing DNMTs has been previously observed. For example DMNT3A/B interact with EZH2 of the PRC2 complex and this interaction is responsible for targeting DNMTs to PRC2 target genes (Vire et al., 2006). Interestingly, about 49% of genes that are hypermethylated in colon cancer are targets for PcG complex and the PcG targets are 12 times more likely to acquire hypermethylation in colon cancers (McGarvey et al., 2008; Widschwendter et al., 2007).

It has been demonstrated that deregulated targeting of methyltransferases correlates with hypermethylation at promoters in cancers. For example, upregulated expression of DNMT3B1 within the *Apc*$^{min/+}$ background exacerbated colorectal cancer and correlated with hypermethylation of tumor suppressor genes (Linhart et al., 2007). Hypermethylation of CpG islands at promoter regions of genes and hypomethylation at repetitive elements are hallmarks of cancer (Jones and Baylin, 2002). These observations suggest a potential link between DNA methylation and ZNF217 that may play a role in promoter hypermethylation in cancers where ZNF217 is overexpressed.

4.7 The Mechanism of DNA demethylation at the p15ink4b promoter

Serum starvation causes cells to enter a quiescent state (G1/G0) resulting in growth arrest (Khammanit et al., 2008). Under these conditions, my result demonstrates that the *p15*^{ink4b} promoter is hypermethylated. My results also indicate that the treatment of serum starved HaCAT cells with TGFβ triggered DNA demethylation at the *p15*^{ink4b} promoter as early as 20 mins after treatment and demethylation was nearly complete by 3 hrs following treatment. The rapidity of this response suggests that DNA demethylation at the *p15*^{ink4b} promoter cannot occur through the passive mechanism which requires DNA replication. However, it has been demonstrated that a small percentage of cells may still undergo cell division following serum starvation (Khammanit et al., 2008). Therefore to confirm that the passive mechanism is not responsible for demethylation of the *p15*^{ink4b} promoter, I treated HaCAT cells with L-mimosine which has been shown induce growth arrest at late G1 phase of the cell cycle before initiation of DNA replication (Krude, 1999). Treatment of cells with TGFβ following treatment with L-mimosine still resulted in DNA demethylation. However, I also observed DNA demethylation of the promoter in

cells treated with L-mimosine alone. L-mimosine is toxic to cells suggesting that the observed demethylation at the *p15*^{ink4b} promoter is as a result of treatment with this amino acid since toxicity may upregulate cell cycle inhibitors including p15^{ink4b} in an attempt to halt cell division (Jenkins et al., 2011; Tsai and Ling, 1971). These observations collectively support an active DNA demethylation mechanism at the *p15*^{ink4b} promoter following TGFβ stimulation.

Two interrelated mechanisms have been postulated for active DNA demethylation. One mechanism is mediated by the AID/APOBEC enzymes that catalyze deamination of 5mC to T. This generates a G:T mispair which is then cleaved by TDG or MBD4 and then reverted to G:C pairs by components of the BER machinary (Rai et al., 2008). The other mechanism involves hydroxylation of 5mC to 5hmC by the TET enzymes. The 5hmC intermediate is oxidized further by the TET enzymes to generate 5fC and 5caC which are recognized and removed by TDG and the BER machinery (Guo et al., 2011a).

ChIP analysis demonstrated that treatment with TGFβ increased recruitment of AID to the the *p15*^{ink4b} promoter. However, knockdown of AID did not abrogate DNA demethylation in response to TGFβ but knockdown of both AID and APOBEC2 demonstrated a modest decrease in TGFβ dependant DNA demethylation. This suggests a deamination-coupled mechanism which is supported by transgenic studies in zebrafish embryos where the co-expression of AID/APOBEC2 and MBD4 is required for demethylation of an ectopically introduced methylated DNA template (Rai et al., 2008). However, DNA immunoprecipitation analysis of the *p15*^{ink4b} promoter indicated enrichment in 5hmC from TGFβ treated HaCAT cells when either TDG or the

deaminases (AID and APOBEC2) were silenced using siRNA. This result suggests a mechanism whereby 5mC is converted to 5hmC prior to removal. However, given the involvement of AID/ABOBEC2 we favor a mechanism whereby 5hmC is deaminated to 5hmU which is then excised by TDG or MBD4. Recent studies have demonstrated that combined overexpression of TET1 and AID cooperate in demethylating a DNA duplex and that 5hmC proved to be a more effective substrate for the deaminases than 5mC (Guo et al., 2011b). Consistent with this, it was demonstrated that in the parietal cortex of psychotic patients upregulation of TET1 and downregulation of APOBEC3A/3B correlates with increased 5hmC levels (Dong et al., 2012). In addition, TDG and MBD4 have robust glycosylase activity against 5hmU:G mispair (Cortellino et al., 2011; Hashimoto et al., 2012a; Morera et al., 2012). However, a recent study has called into question the ability of AID to process 5hmC (Nabel et al., 2012). They demonstrated that AID/APOBECs have decreased deamination activity against 5mC compared to C and no detectable activity against 5hmC *in vitro*. This was attributed to the bulkiness of the hydroxymethyl group which disfavored deamination activity.

5hmC represents 0.3-0.7% and 4-6% of all cytosines in neurons and ESC, respectively, suggesting that 5hmC may be a bona fide epigenetic mark, in addition to being a metabolic intermediate (Kriaucionis and Heintz, 2009; Szulwach et al., 2011; Tahiliani et al., 2009). Furthermore, genomic analysis reveals that 5hmC is found at regulatory regions such as promoters near TSSs, enhancers and exons (Ficz et al., 2011; Pastor et al., 2011; Stroud et al., 2011; Wu et al., 2011). 5hmC levels also correlate well with DNAase I hypersensitive regions and it is often found adjacent to transcription factor binding sites. Ultimately, 5hmC is associated with transcriptional activation (Yu et

al., 2012). In contrast, the derivatives of 5hmC (5fC and 5CaC) are much less abundant consistent with the theory that the 5hmC mark may play a more direct role in transcription (Ito et al., 2011). 5hmC could activate transcription by preventing the binding of the repressive methyl binding proteins such as MeCP2 (Valinluck et al., 2004). However, a recent analysis has shown that MeCP2 binds to 5hmC and facilitates transcription whereas binding to 5mC by MeCP2 promotes repression. In accordance with these observations the authors propose that 5hmC is a stable epigenetic mark that confer transcriptional activation (Mellen et al., 2012). At the *p15*^{ink4b} promoter, it is possible that conversion of 5mC to 5hmC in response to TGFβ treatment may provide an initial rapid transient mechanism to activate transcription, perhaps through switching the activity of MeCP2 that is already bound to the promoter from a repressor to an activator. The sustained treatment with TGFβ could result in full demethylation leading to the expression of p15^{ink4b} for longer time points which is required for the maintenance of the cells in a quiescent state. Further investigation is required to decipher such mechanism.

Our results do not discount the possibility that multiple demethylation pathways play a role in active demethylation at the *p15*^{ink4b} promoter. Structural analysis has demonstrated that TDG is a highly versatile enzyme, capable of cleaving the glycosidic bond of several intermediate derivatives of 5mC that is mispaired with G including T, U, 5hmU, 5fC and 5CaC (Hashimoto et al., 2012b). The generation of these metabolic intermediates is believed to be very rapid and transient. Unfortunately, the availability of techniques and products, such as modification specific antibodies, to quantitatively detect many of these metabolic intermediates with high sensitivity are somewhat limited.

Consequently the relevancy of many of these intermediates must await further clarification.

4.8 The Activation Complex; SMAD/TDG/CBP

We have identified a novel activation complex consisting of SMAD2/3, CBP and TDG, that assembles at the *p15*^{ink4b} promoter in response to TGFβ (Figure 4.1). This complex is required for the active DNA demethylation and coincides with occupancy by several BER enzymes. Our immunoprecipitation assay suggests that SMAD2/3 and TDG physically interact and our ChIP-reChIP assay demonstrates that TDG interacts with both SMAD2/3 and CBP at the *p15*^{ink4b} promoter. Studies using breast cancer cells have shown that TDG associates with CBP and the coactivator SRC3 in response to β-estradiol treatment and these associations are required for DNA demethylation of the PS2 promoter (Chen et al., 2003; Lucey et al., 2005; Tini et al., 2002). More recently, it has been demonstrated in MEFs that TDG associates with both the liganded RAR and p300 at some RAR target genes. Furthermore, the authors demonstrated that the recruitment of p300 to the RAR is dependent on TDG (Cortellino et al., 2011). Although we have not examined the assembly of the SMAD/TDG/CBP complex, it is possible that TDG may act in a similar manner to recruit CBP to the SMAD heterodimers.

A surprising result of this study is that TDG and MBD4 both appear to mediate the active DNA demethylation and activation of *p15*^{ink4b} gene and are targeted to the same region of the promoter in response to TGFβ. However, based on ChIP-reChIP analysis, we did not observe an interaction between MBD4 and SMADs or MBD4 and TDG at the *p15*^{ink4b} promoter suggesting that MBD4 may work independently of TDG and SMADs. It is possible that TDG and MBD4 are functionally redundant and mediate

the actions of TGFβ in different cell populations. Although they are non-homologous proteins, the catalytic activities of TDG and MBD4 are very similar. Another possibility is that the promoter may be initially targeted by the SMAD-TDG-CBP complex promoting an open chromatin conformation which then facilitates the binding of MBD4 through its methyl binding domain. Binding of MBD4 may promote a more complete demethylated state of the promoter and since coregulator binding to the promoter is dynamic, MBD4 may also ensure that the CpG sites remain demethylated by counteracting the activities of DNMT containing complexes, such as the ZNF217 complex.

4.9 ZNF217 Overexpression Impairs TGFβ Induced Coregulator Exchange

HaCAT cells are immortalized keratinocytes which contain normal physiological levels of ZNF217. To examine the consequences of ZNF217 overexpression on TGFβ responsiveness of $p15^{ink4b}$, I transduced ZNF217 in HaCAT cells using adenovirus. ZNF217 overexpression inhibited the promoter demethylation and expression of $p15^{ink4b}$, compared to the mock transduced cells, following TGFβ treatment. ChIP analyses demonstrate an increase in ZNF217 occupancy at the $p15^{ink4b}$ promoter and a reduction in occupancy by SMAD2/3 and TDG. However, no change in the cellular levels of TDG or SMAD2/3 was observed following overexpression of ZNF217. Taken together, with the observation that ZNF217 and SMAD2/3 have overlapping binding sites within the $p15^{ink4b}$ promoter, these results suggests that ZNF217 interferes with coregulator exchange and this imbalance inhibits $p15^{ink4b}$ expression (Figure 4.1).

Deregulation in DNA methylation/demethylation has recently been documented *in vivo* using zebrafish where, the loss of the *APC* tumour suppressor gene causes upregulation of the DNA demethylase machinery (MBD4/TDG, AID and GADD45) and prevents differentiation of intestinal stem cells resulting from a concomitant hypomethylation of key intestinal genes (Rai et al., 2010). Similar findings have been observed in human colon adenomas obtained from patients that harbor mutations at *APC*. This study complements our findings and suggests that oncogenic events leading to the dysregulation of active DNA demethylation could have broad implications in cancer progression.

4.10 Summary and Model of Transcriptional Regulation at p15ink4b

The $p15^{ink4b}$ gene is a CDK inhibitor and an important component of the TGFβ cytostatic response. In this study, I demonstrate for the first time that promoter hypermethylation and transcriptional silencing of the $p15^{ink4b}$ gene involves an association between the ZNF217 complex and DNMT3A/B. Occupancy by the ZNF217 complex also leads to a repressive chromatin landscape at the promoter. Treatment with TGFβ triggers the release of the ZNF217/CoREST/DNMT3A complex and the concomitant binding of an activation complex that includes SMAD 2/3, TDG and CBP. The association of this complex with the promoter is required for the TGFβ mediated active DNA demethylation and activation of $p15^{ink4b}$. Additional constiuents may be part of this complex and could be required to mediate the active DNA demethylation including AID/APOBEC and possibly GADD45α. Collectively, this complex may regulate transcription of genes by simultaneously acetylating histone tails and demethylating the CpG dinucleotides at the promoters of genes (Figure 4.1).

Our data also shows that active DNA demethylation at the *p15*^{ink4b} promoter involves conversion of 5mC to 5hmC which suggests the involvement of TET proteins in this mechanism although this has not been formally demonstrated. Indeed, our dot blot analysis demonstrated a global DNA demethylation in response to TGFβ stimulation and this demethylation could be at promoters of genes, such as *p15*^{ink4b} and *p21*^{Cip1}, affecting transcription or at repetitive sequences/transposons. Finally, I have shown that overexpression of the ZNF217 oncogene, which is associated with a loss of proliferative control and TGFβ resistance in breast and ovarian cancer, inhibits p15^{ink4b} expression by impairing the TGFβ-dependent recruitment of the activation complex. These results highlight the dynamic nature of DNA methylation/demethylation and suggest that coregulator balance is a critical determinant of the status of the *p15*^{ink4b} gene (Figure 4.1).

In many cancer cells, it is possible that ZNF217 overexpression sequesters a limited pool of the other constitutents of the ZNF217 complex recruiting them to promoters of genes where they would not be bound under normal physiological conditions. This can severely upset the normal delicate coregulator balance leading to the aberrant transcriptional regulation of target genes. FACs analysis demonstrates that this deregulation of genes can result in the bypassing of the major check-point at G1 phase of the cell cycle and lead to the promotion of G1/S transition ultimately causing increased cell division when ZNF217 is overexpressed. ZNF217 transduced cell lines also demonstrate resistance to the anti-proliferative effect mediated by TGFβ. *p15*^{ink4b} provides a potential link to the increased G1/S phase transition and TGFβ resistance seen following the overexpression of ZNF217 (Figure 4.1). ZNF217 may target additional gene promoters in a similar fashion and deregulate their expression to collectively

provide tumors having the ZNF217 overexpression with TGFβ resistance, increased proliferation, resistance to apoptosis, EMT and metastasis, ultimately leading to the poor prognosis that is associated with its overexpression.

4.11 Significance and Future Directions

In cancer, aberrant DNA methylation is thought to be an early indicator of the disease and global methylation changes are often observed as a result of epigenetic re-programming (Novik et al., 2002). The recent identification of active DNA demethylation mechanims has provided a new level of complexity in gene regulation and establishes a new role for this mark as a dynamic regulator of transcription. However, the exact mechanism of DNA demethylation remains unclear and will require further studies.

To complete our understanding of the TGFβ pathway, additional targets for the activation complex SMAD/TDG/CBP needs to be elucidated. Since advanced techniques are becoming more available, global approaches such as ChIPseq in conjunction with RNAseq should be utilized to isolate a subset of targets that are bound by the components of this complex and transcriptionally activated in response to TGFβ. In conjunction with this, methylation status at the gene promoters can also be assessed using genome-wide MeDIP assays.

Our knowledge of the mechanism employed by ZNF217 to transcriptionlly repress genes may allow for better treatment strategies for ZNF217 amplified cancers. For example, HDAC inhibitors may not be effective in ZNF217 overexpressed cancers since the mode of transcriptional repression by this complex also involves histone tail demethylation and DNA methylation (Gryder et al., 2012). Drugs that target all three

enzymatic activities associated with the ZNF217 complex, histone deacetylase, histone demethylase and DNA methyltransferase may be required to effectively treat tumors that have the ZNF217 amplification. In the future, ChIPseq in conjunction with RNAseq for ZNF217 needs to be conducted to expand the number of genes that are regulated directly by ZNF217 to completely understand the oncogenesis associated with ZNF217 overexpression. Furthermore, the identification of additional targets for the SMAD activation complex and ZNF217 can be compared to isolate a complete set of genes that confer resistance to the anti-proliferative effects of TGFβ when ZNF217 is amplified.

Collectively, this study provides a comprehensive understanding of the mechanisms that underlies the transcriptional regulation of genes by TGFβ and how this regulation is affected by oncogenes.

4.12 Reference

Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. Cell *111*, 381-392.

Bracken, A.P., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C., Theilgaard-Monch, K., Minucci, S., Porse, B.T., Marine, J.C., *et al.* (2007). The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. Genes Dev *21*, 525-530.

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

Chen, D., Lucey, M.J., Phoenix, F., Lopez-Garcia, J., Hart, S.M., Losson, R., Buluwela, L., Coombes, R.C., Chambon, P., Schar, P., *et al.* (2003). T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. J Biol Chem *278*, 38586-38592.

Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., and Kelsey, G. (2009). Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev *23*, 105-117.

Collins, R., and Cheng, X. (2010). A case study in cross-talk: the histone lysine methyltransferases G9a and GLP. Nucleic Acids Res *38*, 3503-3511.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., *et al.* (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67-79.

Cowger, J.J., Zhao, Q., Isovic, M., and Torchia, J. (2007). Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. Oncogene *26*, 3378-3386.

Dong, E., Gavin, D.P., Chen, Y., and Davis, J. (2012). Upregulation of TET1 and downregulation of APOBEC3A and APOBEC3C in the parietal cortex of psychotic patients. Transl Psychiatry *2*, e159.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *473*, 398-402.

Gal-Yam, E.N., Egger, G., Iniguez, L., Holster, H., Einarsson, S., Zhang, X., Lin, J.C., Liang, G., Jones, P.A., and Tanay, A. (2008). Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. Proc Natl Acad Sci U S A *105*, 12979-12984.

Gryder, B.E., Sodji, Q.H., and Oyelere, A.K. (2012). Targeted cancer therapy: giving histone deacetylase inhibitors all they need to succeed. Future Med Chem *4*, 505-524.

Guo, J.U., Su, Y., Zhong, C., Ming, G.L., and Song, H. (2011a). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. Cell Cycle *10*, 2662-2668.

Guo, J.U., Su, Y., Zhong, C., Ming, G.L., and Song, H. (2011b). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423-434.

Hashimoto, H., Hong, S., Bhagwat, A.S., Zhang, X., and Cheng, X. (2012a). Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. Nucleic Acids Res *40*, 10203-10214.

Hashimoto, H., Hong, S., Bhagwat, A.S., Zhang, X., and Cheng, X. (2012b). Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. Nucleic Acids Res.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300-1303.

Jenkins, N.C., Liu, T., Cassidy, P., Leachman, S.A., Boucher, K.M., Goodson, A.G., Samadashwily, G., and Grossman, D. (2011). The p16(INK4A) tumor suppressor regulates cellular oxidative stress. Oncogene *30*, 265-274.

Jones, P.A. (1985). Altering gene expression with 5-azacytidine. Cell *40*, 485-486.

Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. Nat Rev Genet *3*, 415-428.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900-903.

Kass, S.U., Landsberger, N., and Wolffe, A.P. (1997). DNA methylation directs a time-dependent repression of transcription initiation. Curr Biol *7*, 157-165.

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

Khammanit, R., Chantakru, S., Kitiyanant, Y., and Saikhun, J. (2008). Effect of serum starvation and chemical inhibitors on cell cycle synchronization of canine dermal fibroblasts. Theriogenology *70*, 27-34.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Krimpenfort, P., Ijpenberg, A., Song, J.Y., van der Valk, M., Nawijn, M., Zevenhoven, J., and Berns, A. (2007). p15Ink4b is a critical tumour suppressor in the absence of p16Ink4a. Nature *448*, 943-946.

Krude, T. (1999). Mimosine arrests proliferating human cells before onset of DNA replication in a dose-dependent manner. Exp Cell Res *247*, 148-159.

Kwon, Y.S., Garcia-Bassets, I., Hutt, K.R., Cheng, C.S., Jin, M., Liu, D., Benner, C., Wang, D., Ye, Z., Bibikova, M., *et al.* (2007). Sensitive ChIP-DSL technology reveals an extensive estrogen receptor alpha-binding program on human gene promoters. Proc Natl Acad Sci U S A *104*, 4852-4857.

Linhart, H.G., Lin, H., Yamada, Y., Moran, E., Steine, E.J., Gokhale, S., Lo, G., Cantu, E., Ehrich, M., He, T., *et al.* (2007). Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. Genes Dev *21*, 3110-3122.

Littlepage, L.E., Adler, A.S., Kouros-Mehr, H., Huang, G., Chou, J., Krig, S.R., Griffith, O.L., Korkola, J.E., Qu, K., Lawson, D.A., *et al.* (2012). The transcription factor ZNF217 is a prognostic biomarker and therapeutic target during breast cancer progression. Cancer Discov *2*, 638-651.

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

Lucey, M.J., Chen, D., Lopez-Garcia, J., Hart, S.M., Phoenix, F., Al-Jehani, R., Alao, J.P., White, R., Kindle, K.B., Losson, R., *et al.* (2005). T:G mismatch-specific thymine-DNA glycosylase (TDG) as a coregulator of transcription interacts with SRC1 family members through a novel tyrosine repeat motif. Nucleic Acids Res *33*, 6393-6404.

McGarvey, K.M., Van Neste, L., Cope, L., Ohm, J.E., Herman, J.G., Van Criekinge, W., Schuebel, K.E., and Baylin, S.B. (2008). Defining a chromatin pattern that characterizes DNA-hypermethylated genes in colon cancer cells. Cancer Res *68*, 5753-5759.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., *et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766-770.

Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. Cell *151*, 1417-1430.

Morera, S., Grin, I., Vigouroux, A., Couve, S., Henriot, V., Saparbaev, M., and Ishchenko, A.A. (2012). Biochemical and structural characterization of the glycosylase domain of MBD4 bound to thymine and 5-hydroxymethyuracil-containing DNA. Nucleic Acids Res.

Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat Chem Biol *8*, 751-758.

Nonet, G.H., Stampfer, M.R., Chin, K., Gray, J.W., Collins, C.C., and Yaswen, P. (2001). The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. Cancer Res *61*, 1250-1254.

Novik, K.L., Nimmrich, I., Genc, B., Maier, S., Piepenbrock, C., Olek, A., and Beck, S. (2002). Epigenomics: genome-wide study of methylation phenomena. Curr Issues Mol Biol *4*, 111-128.

Nunez, N., Clifton, M.M., Funnell, A.P., Artuz, C., Hallal, S., Quinlan, K.G., Font, J., Vandevenne, M., Setiyaputra, S., Pearson, R.C., *et al.* (2011). The multi-zinc finger protein ZNF217 contacts DNA through a two-finger domain. J Biol Chem *286*, 38190-38201.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W., *et al.* (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet *39*, 237-242.

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Ortega, S., Malumbres, M., and Barbacid, M. (2002). Cyclin D-dependent kinases, INK4 inhibitors and cancer. Biochim Biophys Acta *1602*, 73-87.

Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet *10*, 669-680.

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., *et al.* (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature *473*, 394-397.

Quinlan, K.G., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., Bolognesi, M., and Crossley, M. (2006). Specific recognition of ZNF217 and other zinc finger proteins at a surface groove of C-terminal binding proteins. Mol Cell Biol *26*, 8159-8172.

Quinlan, K.G., Verger, A., Yaswen, P., and Crossley, M. (2007). Amplification of zinc finger gene 217 (ZNF217) and cancer: when good fingers go bad. Biochim Biophys Acta *1775*, 333-340.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201-1212.

Rai, K., Sarkar, S., Broadbent, T.J., Voas, M., Grossmann, K.F., Nadauld, L.D., Dehghanizadeh, S., Hagos, F.T., Li, Y., Toth, R.K., *et al.* (2010). DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. Cell *142*, 930-942.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E., *et al.* (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet *39*, 232-236.

Siegel, P.M., and Massague, J. (2003). Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. Nat Rev Cancer *3*, 807-821.

Simboeck, E., Ribeiro, J.D., Teichmann, S., and Di Croce, L. (2011). Epigenetics and senescence: learning from the INK4-ARF locus. Biochem Pharmacol *82*, 1361-1370.

Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Moroy, T., Bartek, J., Massague, J., Hanel, F., *et al.* (2001). Repression of p15INK4b expression by Myc through association with Miz-1. Nat Cell Biol *3*, 392-399.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol *12*, R54.

Su, R.C., Brown, K.E., Saaber, S., Fisher, A.G., Merkenschlager, M., and Smale, S.T. (2004). Dynamic assembly of silent chromatin during thymocyte maturation. Nat Genet *36*, 502-506.

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., *et al.* (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat Neurosci *14*, 1607-1616.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Thollet, A., Vendrell, J.A., Payen, L., Ghayad, S.E., Ben Larbi, S., Grisard, E., Collins, C., Villedieu, M., and Cohen, P.A. (2010). ZNF217 confers resistance to the pro-apoptotic signals of paclitaxel and aberrant expression of Aurora-A in breast cancer cells. Mol Cancer *9*, 291.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., *et al.* (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75-82.

Tini, M., Benecke, A., Um, S.J., Torchia, J., Evans, R.M., and Chambon, P. (2002). Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell *9*, 265-277.

Tsai, W.C., and Ling, K.H. (1971). Toxic action of mimosine. I. Inhibition of mitosis and DNA synthesis of H.Ep-2 cell by mimosine and 3,4-dihydroxypyridine. Toxicon *9*, 241-247.

Valinluck, V., Tsai, H.H., Rogstad, D.K., Burdzy, A., Bird, A., and Sowers, L.C. (2004). Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). Nucleic Acids Res *32*, 4100-4108.

Vandevenne, M., Jacques, D.A., Artuz, C., Nguyen, C.D., Kwan, A.H., Segal, D.J., Matthews, J.M., Crossley, M., Guss, J.M., and Mackay, J.P. (2013). New insights into DNA recognition by zinc fingers revealed by structural analysis of ZNF217. J Biol Chem.

Vendrell, J.A., Thollet, A., Nguyen, N.T., Ghayad, S.E., Vinot, S., Bieche, I., Grisard, E., Josserand, V., Coll, J.L., Roux, P., *et al.* (2012). ZNF217 is a marker of poor prognosis in breast cancer that drives epithelial-mesenchymal transition and invasion. Cancer Res *72*, 3593-3606.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., *et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I., *et al.* (2007). Epigenetic stem cell signature in cancer. Nat Genet *39*, 157-158.

Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev *25*, 679-684.

You, J.S., Kelly, T.K., De Carvalho, D.D., Taberlay, P.C., Liang, G., and Jones, P.A. (2011). OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes. Proc Natl Acad Sci U S A *108*, 14497-14502.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., *et al.* (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell *149*, 1368-1380.

Appendices

Appendix A: License Agreement for Chapter 2

Copyright

To maintain and protect the Society's ownership and rights and to continue to afford scientists the opportunity to publish in high-quality journals, ASM requires the corresponding author to sign a copyright transfer agreement on behalf of all the authors. Unless this agreement is executed (without changes and/or addenda), ASM will not publish the article.

In the copyright transfer agreement signed by an author, ASM grants to that author (and coauthors) the right to republish discrete portions of his/her (their) article in any other publication (print, CD-ROM, and other electronic forms) of which he/she is (they are) the author(s) or editor(s), on the condition that appropriate credit is given to the original ASM publication. This republication right also extends to posting on a host computer to which there is access via the Internet. Except as indicated below, significant portions of the article may not be reprinted/posted without ASM's prior written permission, however, as this would constitute duplicate publication.

Authors may post their own published articles on their personal or university-hosted (but not corporate, government, or similar) websites without ASM's prior written permission provided that appropriate credit is given (i.e., the copyright lines shown at the bottom of the first page of the PDF version).

Works authored solely by U.S. government employees are not subject to copyright protection, so there is no copyright to be transferred. However, the other provisions of the copyright transfer agreement, such as author representations of originality and authority to enter into the agreement, apply to U.S. government employee-authors as well as to other authors.

Copyright for supplemental material (see “Supplemental Material”) remains with the author, but a license permitting the posting by ASM is included in the article copyright transfer agreement. If the author of the article is not also the copyright owner of the supplemental material, the corresponding author must send to ASM signed permission from the owner that allows posting of the material, as a supplement to the article, by ASM. The corresponding author is also responsible for incorporating into the supplemental material any copyright notices required by the owner.

Appendix B: License Agreement for Chapter 3

ELSEVIER LICENSE
TERMS AND CONDITIONS

Mar 05, 2013

This is a License Agreement between Gobi Thillainadesan ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Gobi Thillainadesan
License number	3102690415586
License date	Mar 05, 2013
Licensed content publisher	Elsevier
Licensed content publication	Molecular Cell
Licensed content title	TGF-β-Dependent Active Demethylation and Expression of the p15Tumor Suppressor Are Impaired by the ZNF217/CoREST Complex
Licensed content author	Gobi Thillainadesan,Jennifer Mary Chitilian,Majdina Isovic,Jailal Nicholas George Ablack,Joe Stephen Mymryk,Marc Tini,Joseph

	Torchia
Licensed content date	8 June 2012
Licensed content volume number	46
Licensed content issue number	5
Number of pages	14
Start Page	636
End Page	649
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Transcriptional Regulation by the Oncogenic ZNF217/CoREST Complex
Expected completion date	Apr 2013
Estimated size (number of pages)	300
Elsevier VAT number	GB 494 6272 12

Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be

altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Website**: The following terms and conditions apply to electronic reserve and author websites:
Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:
This license was made in connection with a course,
This permission is granted for 1 year only. You may obtain a license for future website posting,
All content posted to the web site must maintain the copyright information line on the bottom of each image,

A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com , and
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx . As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. **Author website** for books with the following additional clauses:
Authors are permitted to place a brief summary of their work online only.
A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. or for books to the Elsevier homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. **Other Conditions**:

v1.6

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500971028. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:

Copyright Clearance Center

Dept 001

P.O. Box 843006

Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

Curriculum Vitae

Gobi Thillainadesan

Education:

University of Western Ontario
London regional cancer centre — London, Ontario
Ph.D — 2006-2013

University of Western Ontario — London, Ontario
Bachelor of Science,Honors Genetics (Hon.B.Sc) — 2001-2006

Professional Experience:

Dr. Joe Torchia — London, Ontario
Research assistant — May 2006 - Sep 2006

Dr. Susanne Kohalmi — London, Ontario
Research assistant/ — Sep 2004 - May 2006
Honor's 4th year project

Dr. Vava Grbic — London, Ontario
Research assistant — Sep 2003- May 2003

Scholarships:

External:
Frederick Banting and Charles Best Canada Graduate Scholarships - Doctoral
Issuing party: Canadian Institute of Health Research (CIHR;Canada-National).
Period of Tenure: Sep 2009 - Sep 2012

Ontario Graduate Scholarship (OGS)
Issuing party: Provincial (Ontario)
Location of Tenure of Award: Received but did not accept.
Received: May 2009

Ontario Graduate Scholarship in Science and Technology (OGSST)
Issuing party: Provincial (Ontario)
May 2007 - May 2008

Internal:
CIHR: The London Strategic Training Initiative in Cancer Research and Technology Transfer (CaRTT)
Issuing party: Regional (London)
May 2007 - Sep 2009

Schulich Graduate Scholarship and Western Graduate Research Scholarship

Issuing party: University of Western Ontario (London)
Sep 2006 - Sep 2011

Awards:

Institute of Cancer Research Publication Prize (CIHR)
For the best publication 2013

Schulich School of Medicine and Dentistry Publication Award
For the best publication 2012

Morris Kroll Memorial Scholarship in Cancer Research
For excellence in the field of cancer research. 2009

Oncology Research and Education Conference Presentation Award
For the best presenter at the conference 2007

Publications:

Coughlan N, **Thillainadesan G**, Andrews J, Isovic M, Torchia J
ß-Estradiol-dependent activation of the Jak/Stat pathway requires p/CIP and CARM1.
.Biochim Biophys Acta. 2013 Feb 20.

DeKoter RP, Geadah M, Khoosal S, Xu LS, **Thillainadesan G**, Torchia J, Chin SS, Garrett-Sinha LA.
Regulation of follicular B cell differentiation by the related E26 transformation-specific transcription factors PU.1, Spi-B, and Spi-C. J Immunol. 2010 Dec 15;185(12):7374-84.

Fonseca GJ, **Thillainadesan G**, Yousef AF, Ablack JN, Mossman KL, Torchia J, Mymryk JS.
Adenovirus evasion of interferon-mediated innate immunity by direct antagonism of a cellular histone posttranslational modification. Cell Host Microbe. 2012 Jun 14;11(6):597-606.

Ablack JN, Cohen M, **Thillainadesan G**, Fonseca GJ, Pelka P, Torchia J, Mymryk JS. J Cellular GCN5 is a Novel Regulator of Human Adenovirus E1A-Conserved Region 3 Transactivation. Virol. 2012 May 23.

Thillainadesan G, Chitilian JM, Isovic M, Ablack JN, Mymryk JS, Tini M, Torchia J.
TGF-β-Dependent Active Demethylation and Expression of the p15(ink4b) Tumor Suppressor Are Impaired by the ZNF217/CoREST Complex.Mol Cell. 2012 Jun8;46(5):636-49.

Corkery D, **Thillainadesan G**, Coughlan N, Mohan RD, Isovic M, Tini M, Torchia J.
Regulation of the BRCA1 gene by an SRC3/53BP1 complex. BMC Biochem. 2011 Sep13;12:50.

Xu S, Sokalski K, Hotke K, Piskorz J, Zarnett O, **Thillainadesan G**, Torchia J, DeKoter

RP.
Regulation of B cell linker protein transcription by PU.1 and Spi-B in murine B cell acute lymphoblastic leukemia. J Immunol. 2012 Oct 1;189(7):3347-54

Thillainadesan G, Isovic M, Loney E, Andrews J, Tini M, Torchia J.
Genome analysis identifies the p15ink4b tumor suppressor as a directtarget of the ZNF217/CoREST complex. Mol Cell Biol. 2008 Oct;28(19):6066-77.

Publications in preparation:

Chitilian JM, Chang WY, Walker E, **Thillainadesan G**, Stanford WL, and Torchia J
The role of the transcriptional coregulator p/CIP in stem cell pleuripotency. In preparation.

Conferences/Oral presentations:

Thillainadesan G.
Schulich School of Medicine and Dentistry (Biochemistry) Talks
(2006 - 2011)

Thillainadesan G.
Cancer Research Laboratories Talks
(2006 - 2010)

Thillainadesan G,Torchia J.
Global Identification of Target Genes for the Oncoprotein ZNF217 in breast cancer.
2007-Invited speaker at the Oncology Research and Education Day (London, Ontario).

Thillainadesan G, Kohalmi S
Determination of Subcellular localization of ArogenateDehydratase (ADT) in Arabidopsis thaliana.
2004-Biology Day Conference, Southwestern Ontario.

Conferences/Poster presentations:

Thillainadesan G, Chitilian JM, Isovic M, Ablack JN, Mymryk JS, Tini M, Torchia J.
TGFβ dependent active demethylation and expression of the p15ink4b tumour suppressor gene is impaired by the ZNF217 oncogene.
2012-Oncology Research and Education Day (London, Ontario).

Thillainadesan G, Chitilian JM, Isovic M, Ablack JN, Mymryk JS, Tini M, Torchia J.
TGFβ dependent active demethylation and expression of the p15ink4b tumour suppressor gene is impaired by the ZNF217 oncogene.

2012-Epigenetics of Cancer and Stem Cells (Toronto, Ontario).

Thillainadesan G, Chitilian J, Torchia J.
ChIPseq analysis of ZNF217 and CtBP1 in breast cancer.
2011-Gordon Conference: Cancer Genetics & Epigenetics (Ventura, California).

Thillainadesan G, Torchia J.
ZNF217 oncogene maintains DNA methylation at the p15ink4b promoter.
2011-Oncology Research and Education Day (London, Ontario).

Thillainadesan G, Isovic M, Loney E, Andrews J, Tini M,Torchia J.
Genome analysis identifies the p15ink4b tumor suppressor as a direct target of the ZNF217/CoREST complex.
2010-Oncology Research and Education Day (London, Ontario).

Thillainadesan G, Isovic M, Loney E, Andrews J, Tini M,Torchia J.
Global Identification of Targets and Characterization of the oncoprotein ZNF217's role on the p15ink4b promoter in breast cancer.
2009-Oncology Research and Education Day (London, Ontario).

Thillainadesan G,Torchia J.
Identification of Tumor Suppressor Target Genes for the Oncoprotein ZNF217 in breast cancer.
2008-Oncology Research and Education Day (London, Ontario).

Thillainadesan G, Isovic M, Loney E, Andrews J, Tini M, Torchia J.
Global Identification of ZNF217 Direct Target Genes In Breast Cancer.
2008-American Association of Cancer Research (SanDiego, California).

Kohalmi S, **Thillainadesan G**.
Determination of Subcellular localization of ArogenateDehydratase(ADT) in Arabidopsis thaliana.
2006-Genetic Society Meeting Conference(McMaster, Hamilton, Ontario).

Thillainadesan G, Torchia J.
Global Identification of Transcription Factor Target Genes in Breast Cancer Cells.
2006-Biochemistry Day Conference, (UWO,London, Ontario).