Epigenetic regulation of gene silencing during stem cell commitment and differentiation.

Claudia Canzonetta

A thesis submitted for the degree of
Doctor of Philosophy at Imperial College School of Medicine

May 2007

Abstract

The B lineage specific mouse *λ5-VpreB1* locus was analysed to determine the patterns of epigenetic modification during lineage commitment and differentiation. The silenced state of the genes was investigated in precursor cells with different levels of plasticity to determine which epigenetic features of the domain are influenced by early fate decisions. Investigation of histone modifications linked with gene silencing and chromosome condensation such as di- and tri- methyl histone H3 lysine 9 and the double histone H3 tri-methyl lysine 9 phosphorylated serine 10, showed similar levels of enrichment in embryonic stem (ES) cells, mesenchymal stem cells and mature B cells. However, in ES cells, a region that was previously found to be marked by histone H3 lysine 4 di-methylation (Szutorisz et al., 2005), displayed a lower level of lysine 9 di-methylation compared to the rest of the locus. This feature was not observed in mature B cells. These results imply that silencing within the B cell lineage is not accompanied by major increases in inactivating modifications. In contrast, very high enrichment in histone H3 lysine 9 di-methylation was observed in non-lymphoid osteocyte cells. Analysis of DNA methylation across the *λ5-VpreB1* locus revealed significant epigenetic differences between the different stages of B cell development. In particular ES cells showed a complete absence of CpG methylation in the locus. Methylation was observed in repetitive regions at the pre-B cell stage, when the genes are expressed. DNA methylation increased in mature B cells, where it was detectable at a number of sites in the locus.

Acknowledgements

I would like to thank all the people in the Gene Regulation and Chromatin group at the Clinical Sciences Centre, Hammersmith Hospital for their help and for rendering the lab a very nice place were to work. I would like to thank especially my supervisor Dr. Niall Dillon for constant advice and support throughout the project. I would also like to give a very special mention to Roberto, without him, his love and his patient I would have never make it so far. Also thanks to my Mum and Dad, my brother Andrea and Pamela and their fantastic baby girls, Chiara and Luna because seeing them growing is just the most wonderful thing never happened to me. Thanks to all of my friends, both near and far, whose love, friendship and support has been, and continues to be, invaluable to me.

Table of contents

ABSTRACT 2

ACKNOWLEDGEMENTS 3

TABLE OF CONTENTS 4

LIST OF FIGURES 8

LIST OF TABLES 11

LIST OF ABBREVIATIONS 12

1. INTRODUCTION 13

1.1 OVERVIEW 13
1.2 EUKARYOTIC CHROMATIN ORGANIZATION 14
1.3 EUCHROMATIN AND HETEROCHROMATIN 18
1.3.1 EUCHROMATIN 19
1.3.2 HETEROCHROMATIN 19
1.4 OVERVIEW OF TRANSCRIPTION 21
1.5 CIS-ACTING REGULATORY SEQUENCES: PROMOTERS AND ENHANCERS 25
1.6 LOCUS CONTROL REGIONS (LCRS) 27
1.6.1 THE β-GLOBIN LOCUS CONTROL REGION 28
1.6.2 PROPERTIES OF LCRS: TRANSCRIPTIONAL ENHANCER ACTIVITY 29
1.7 DOMAINS AND GENETIC FUNCTIONAL UNITS 30
1.7.1 STRONG AND WEAK DOMAIN MODELS 30
1.7.2 LONG-RANGE CHROMOSOMAL INTERACTIONS 34
1.7.3 THE ACTIVE CHROMATIN HUB (ACH) 35
1.8 CHROMATIN REMODELLING 37
1.8.1 THE YEAST HO PROMOTER 39
1.8.2 THE MOUSE MMTV PROMOTER 39
1.9 EPIGENETIC REGULATION 40
1.10 HISTONE METHYLATION 42
1.10.1 HETEROCHROMATIN FORMATION: LYSINE 9 TRI-METHYLATION 43
1.10.2 HETEROCHROMATIN FORMATION: HISTONE H4 LYSINE 20 44
1.10.3 TRANSCRIPTIONAL SILENCING: HISTONE H3 LYSINE 9 MONO- AND DIMETHYLATION AND HISTONE H3 LYSINE 27 METHYLATION 45
1.10.4 HISTONE METYLATION AND TRANSCRIPTIONAL ACTIVATION 46
1.10.5 HISTONE H3 LYSINE 36 AND 79 48
1.11 DEMETHYLATION OF HISTONES 49
1.12 HISTONE ACETYLATION 51
1.13 HISTONE DEACETYLASES (HDACS) 54
1.13.1 REGULATION OF TANSCRIPTIONAL SILENCING BY HDACS 56
1.13.2 SIRTUIN DEACETYLASES 57
1.14 HISTONE PHOSPHORYLATION 58
1.15 UBIQUITINATION OF HISTONES 60
1.16 DNA METHYLATION 61

1.16.1 METHYL BINDING PROTEINS AND TRANSCRIPTIONAL REPRESSION 62
1.16.2 DNA METHYLATION, HISTONE DEACETYLATION AND H3K9 METHYLATION 64
1.17 EMBRYONIC STEM CELLS 65
1.17.1 CHROMATIN STRUCTURE IN ES CELLS 65
1.17.2 EMBRYONIC STEM CELLS REGULATORS 66
1.17.3 MODELS FOR EMBRYONIC STEM CELL PLURIPOTENCY 69
1.18 B CELL BIOLOGY 70
1.19 B CELL DEVELOPMENT 72
1.20 TRANSCRIPTIONAL REGULATION DURING B CELL DEVELOPMENT 74
1.20.1 ROLE OF EBF, PAX5 AND E2A DURING B CELL DEVELOPMENT 75
1.20.2 ROLE OF IKAROS, AIOLOS AND DEDALOS IN B CELL DEVELOPMENT 78
1.21 THE MOUSE λ5-VPREB1 LOCUS 80
1.21.1 TRANSCRIPTIONAL REGULATION OF THE Λ5-VPREB1 LOCUS 80
1.21.2 TRANSGENIC ANALYSIS OF THE Λ5-VPREB1 LOCUS 82

2. MATERIALS AND METHODS 84

2.1 STANDARD STOCK SOLUTIONS 84
2.2 SOLUTIONS FOR BACTERIAL CULTURE AND TRANSFORMATIONS 84
2.3 ISOLATION OF NUCLEIC ACIDS 84
2.3.1 MINI-PREP OF PLASMID DNA 84
2.3.2 MAXI-PREP OF PLASMID DNA 85
2.3.3 ISOLATION OF MOUSE GENOMIC DNA FROM BRAIN AND LIVER TISSUE 85
2.3.4 ISOLATION OF MOUSE GENOMIC DNA FROM ES, PRIMARY PRE-B AND MATURE B CELLS 85
2.3.5 OLIGONUCLEOTIDES 86
2.3.6 QUANTIFICATION OF NUCLEIC ACIDS 86
2.4 STANDARD MANIPULATIONS OF DNA 86
2.4.1 RESTRICTION ENZYME DIGESTION OF PLASMID DNA 86
2.4.2 RESTRICTION ENZYME DIGESTION OF GENOMIC DNA. 86
2.5 DNA ELECTROPHORESIS 86
2.5.1 AGAROSE GEL ELECTROPHORESIS 86
2.5.2 NON-DENATURING POLYACRYLAMIDE GEL ELECTROPHORESIS (PAGE) 87
2.5.3 ISOLATION OF DNA FROM AGAROSE GELS 87
2.6 RADIO-LABELLING OF DNA 87
2.7 GENERATION OF DNA PROBE TEMPLATES BY PCR 88
2.8 SOUTHERN BLOT ANALYSIS 89
2.9 HISTONE PROTEINS SDS PAGE 90
2.10 CELLS CULTURE 91
2.10.1 TRANSFORMED CELL LINES 91
2.10.2 ACTIVATION OF PRIMARY FOETAL PRE-B CELL CULTURES FROM FOETAL LIVER 93
2.10.3 ACTIVATION OF PRIMARY MATURE B CELL CULTURES 94
2.10.4 ES CELLS 94
2.10.5 C3H10T1/2 CELLS 95
2.11 CHROMATIN IMMUNOPRECIPITATION (CHIP) 96
2.11.1 UNFIXED CHROMATIN PREPARATION 96
2.11.2 IMMUNOPRECIPITATION OF UNFIXED CHROMATIN (N-CHIP) 98
2.11.3 FIXED CHROMATIN PREPARATION 100
2.11.4 IMMUNOPRECIPITATION OF FIXED CHROMATIN (X-CHIP) 102
2.12 REAL TIME PCR 103
2.12.1 PRIMERS 103
2.12.2 DATA ANALYSIS 106

3. PATTERNS OF HISTONE H3 AND H4 ACETYLATION AND H3 K4 DIMETHYLATION ACROSS THE λ5-VPREB1 LOCUS IN PRE-B AND MATURE B CELLS 107

3.1 INTRODUCTION 107
3.2 OVERVIEW OF CHIP 108
3.3 DATA ANALYSIS 110
3.4 PRIMERS 112
3.5 ANALYSIS OF HISTONE H3 AND H4 ACETYLATION ACROSS THE λ5-VPREB1 LOCUS IN ABELSON TRANSFORMED CELL LINES 114
3.6 ANALYSIS OF HISTONE H3 AND H4 ACETYLATION ACROSS THE *λ5-VPREB1* LOCUS IN PRIMARY PRE-B AND MATURE B CELLS 116
3.7 PATTERN OF HISTONE H3 LYSINE 4 DIMETHYLATION IN PRIMARY PRE-B AND MATURE B CELLS. 118
3.8 SUMMARY 120

4. INACTIVE VERSUS ACTIVE CHROMATIN MODIFICATIONS DURING B CELL DEVELOPMENT 121

4.1 INTRODUCTION 121
4.2 PATTERNS OF HISTONE H3K9ME2 ACROSS THE λ5-VPREB1 LOCUS DURING B CELL DEVELOPMENT 122
4.3 PATTERNS OF HISTONE H3K9ME3 AND H3K9ME3S10PH ACROSS THE λ5-VPREB1 LOCUS IN ES AND MATURE B CELLS 126
4.4 PATTERNS OF H3K9ME2 AND H3K4ME2 IN TISSUE-SPECIFIC AND UBIQUITOUSLY EXPRESSED GENES AT DIFFERENT STAGES OF B CELL DEVELOPMENT 129
4.5 PATTERNS OF HISTONE H3K9ME3 AND H3K9ME3S10PH AT TISSUE-SPECIFIC AND UBIQUITOUSLY EXPRESSED GENES IN ES AND MATURE B CELLS 132
4.6 SUMMARY 135

5. EPIGENETIC MODIFICATIONS OF TISSUE-SPECIFIC GENES IN THE C3H10T1/2 MESENCHYMAL STEM CELL LINE 136

5.1 INTRODUCTION 136
5.2 C3H10T1/2 DIFFERENTIATION 137
5.3 PATTERNS OF HISTONE H3K4ME2 AND HISTONE H3 DI-ACETYLATION ACROSS THE *λ5-VPREB1* LOCUS DURING C3H10T1/2 CELL DIFFERENTIATION 139
5.4. PATTERNS OF HISTONE H3 LYSINE 9 METHYLATION IN THE λ5-VPREB1 LOCUS IN UNDIFFERENTIATED AND DIFFERENTIATED C3H10T1/2 CELLS 140
5.4.1 HISTONE H3 LYSINE 9 DIMETHYLATION 141
5.4.2 HISTONE H3 K9 TRIMETHYLATION AND HISTONE H3 K9 TRIMETHYLATION / S10 PHOSPHORYLATION 142
5.5 HISTONE MODIFICATIONS IN TISSUE SPECIFIC AND UBIQUITOUSLY EXPRESSED GENES DURING IN VITRO OSTEOCYTE DIFFERENTIATION 145
5.5.1 HISTONE H3 ACETYLATION AND H3 LYSINE 4 DIMETHYLATION 145
5.5.2 HISTONE H3 LYSINE 9 DIMETHYLATION 147
5.5.3 HISTONE H3 LYSINE 9 TRIMETHYLATION 149
5.5.4 HISTONE H3 K9 TRIMETHYLATION / S10 PHOSPHORYLATION 150
5.7 SUMMARY 152

6. DNA METHYLATION ANALYSIS OF THE λ5-VPREB1 LOCUS 153

6.1 INTRODUCTION 153
6.2 STRATEGY FOR SOUTHERN BLOT ANALYSIS OF DNA METHYLATION 153
6.3 DNA METHYLATION ANALYSIS OF THE λ5-VPREB1 LOCUS 155
6.4 SUMMARY 158

7 DISCUSSION 159

7.1 OVERVIEW 159
7.2 THE ACTIVE *λ5-VPREB1* LOCUS 159
7.3 EPIGENETIC FEATURES OF THE *λ5-VPREB1* DOMAIN IN THE B CELL LINEAGE AND IN NON LYMPHOID LINEAGES 162
7.4 DEVELOPMENTAL VERSUS LONG TERM SILENCING OF THE B CELL-SPECIFIC *λ5-VPREB1* DOMAIN 163
7.5 HISTONE H3 K9 METHYLATION IS A GENERAL MARKER FOR GENE SILENCING DURING CELL DIFFERENTIATION 164
7.6 DNA METHYLATION OF THE *λ5-VPREB1* LOCUS DURING B CELL DEVELOPMENT 165

REFERENCE LIST 167

List of Figures

Fig.1.2.1	X-ray structure of the nucleosome core	15
Fig. 1.2.2	Organization of DNA within chromatin structure	17
Fig. 1.6.1	The Locus Control Region of the human β-globin gene cluster	29
Fig. 1.7.1	Insulator and enhancer	33
Fig. 1.7.2	Long range chromatin interactions	35
Fig.1.7.3	Long-range interactions in the mouse b-globin	36
Fig. 1.8.1	Models of actions of ATP-dependent chromatin remodellers	38
Fig. 1.9.1	Histone modifications identified in vivo	41
Fig.1.13.1	Schematic representation of complexes containing HDACs.	55
Fig.1.16.1	Methyl-CpG-binding proteins (MBPs)	62
Fig.1.18.1	B cell development	71
Fig.1.20.1	Transcriptional control of B-cell development	74
Fig.1.21.1	The λ5-VpreB1 locus	80
Fig.1.21.2	Location of sequences that activate λ5 and VpreB1 expression	82
Fig.2.10.1	FACS analysis of Abelson transformed pre-B cells	92
Fig.2.10.2	FACS analysis of primary pre-B cells	93
Fig.2.10.3	FACS analysis of LPS activated mature B cells	94
Fig.2.11.1	Native ChIP	97
Fig.2.11.3	Fixed ChIP	101
Fig. 3.2.1	Formaldehyde cross linking: generation of non specific signal	110
Fig.3.3.1	Typical output from a real time PCR experiments.	115
Fig. 3.4.1	Real time PCR primers and the λ5-VpreB1 locus	113
Fig. 3.5.1	Pattern of histone H3 and histone H4 acetylation across the λ5-VpreB1 locus in pre-B cell and mature B cell lines	115
Fig. 3.5.2	Pattern of histone H3 and histone H4 acetylation across the λ5-VpreB1 locus in primary cells	117
Fig. 3.5.3	Pattern of histone H3K4me2 λ5-VpreB1 locus in primary cells	119
Fig. 4.1.1	Model of B cell development	122
Fig. 4.2.1	Pattern of H3K9me2 across the λ5-VpreB1 locus at different stages of B-cell development.	124

Fig. 4.3.1 Pattern of H3K9me3 across the λ5-VpreB1 locus in ES and mature B cells 127

Fig. 4.3.2 Pattern of H3K9me3S10ph across the λ5-VpreB1 locus in ES and mature B cells. 128

Fig. 4.4.1 Patterns of H3K9me2 and H3K4me2in tissue specific and housekeeping genes. 131

Fig. 4.5.1 Patterns of H3K9me3 in ES and mature B cells in tissue-specific and housekeeping genes. 133

Fig. 4.5.2 Patterns of H3K4me3S10ph in ES cells and mature B cells in tissue-specific and housekeeping genes. 134

Fig. 5.1.1 Model system used to analyse epigenetic modifications during cell differentiation. 136

Fig.5.2.1 Differentiation of the mesenchymal stem cell line C3H10T1/2 into osteocytes. 138

Fig.5.2.2 Differentiation of the mesenchymal stem cell line C3H10T1/2 into osteocytes. 138

Fig. 5.3.1 Patterns of H3 acetylation and H3K4me2 across the λ5-VpreB1 locus in CHT and CHTD cells 140

Fig. 5.4.1 Patterns of histone H3K9me2 across the λ5-VpreB1 locus in ES cells, in CHT cells and in CHTD cells 141

Fig. 5.4.2 Patterns of histone H3K9me3 across the λ5-VpreB1 locus in CHT cells 143

Fig. 5.4.3 Patterns of histone H3K4me3S10ph across the λ5-VpreB1 locus in CHT and CHTD 144

Fig. 5.5.1 Patterns of H3K4me2 and H3 acetylation CHT and CHTD cells in tissue-specific and housekeeping genes. 146

Fig. 5.5.2 Patterns of H3K9me2 in ES cells, in CHT cells and CHTD cells in tissue-specific and housekeeping genes. 147

Fig. 5.5.3 Patterns of H3K9me3 in ES cells CHT and CHTD cells in tissue-specific and housekeeping genes. 149

Fig. 5.5.4 Patterns of H3K4me3S10ph in ES cells, CHT and CHTD cells in tissue-specific and housekeeping genes. 151

Fig.6.2.1 McrBC Southern blot strategy 154

Fig. 6.2.2 Etidium bromide staining of southern blots 154
Fig. 6.3.1 Southern blot of the 5'end of the λ5-VpreB1 domain 155
Fig. 6.3.2 Southern blot of the central region of the λ5-VpreB1 domain 156
Fig. 6.3.3 Southern blot of the 3'end of the λ5-VpreB1 domain 156
Fig. 6.3.4 Summary of the methylation sites detected by McrBC digestion. 157
Red arrows indicate methylated cytosine.

List of Tables

Table 1.3.1	General characteristics chromatin	18
Table 1.5.1	Glossary of terms	26
Table 1.6.1	Examples of gene loci with LCR activity	28
Table 1.12.1	Characteristics of HAT families	51
Table 1.18.1	Stages of B-lymphocyte development	71
Table 2.10.2	List of antibodies used in N-ChIP	99
Table 2.10.4	List of antibodies used in X-ChIP	102
Table 2.12.1a	Real Time PCR primers in the λ5-VpreB1 locus	104
Table 2.12.1b	House keeping and tissue specific genes real time PCR primers	105
Tab. 3.4.1.	Possible combination of concentration for Real Time PCR primers	113

List of abbreviations

ACH	active chromatin hub
bp	base pair
ChIP	chromatin immunoprecipitation
CHT	C3H10T1/2 cells
CHTD	Differentiated C3H10T1/2 cells
CPE	core promoter element
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
(E)	Embryonic
EBF	early B-cell factor
ES	embryonic stem cells
EtBr	Etidium bromide
HP1	heterochromatin protein 1
HS	hypersensitive site
kb	kilobase
LCR	locus control region
MMTV	mouse mammary tumour virus
MNase	micrococcal nuclease
N-ChIP	Native chromatin immunoprecipitation
PCR	polymerase chain reaction
PEV	position effect variegation
RAG	recombination activating gene
RNA	ribonucleic acid
RNase	Ribonuclease
TBP	TATA binding protein
UV	Ultraviolet
λ5	Lambda 5
X-ChIP	Cross linked chromatin immunoprecipitation

1. Introduction

1.1 Overview

Chromatin represents the condensed state of the DNA in the eukaryotic cell. The ability to package large genomes is associated with a fundamental shift in the logic of gene regulation between prokaryotes and eukaryotes. Whereas in prokaryotes the ground state is not restrictive, in eukaryotes transcriptional activity is generally impeded by nucleosomal packaging. In the past chromatin was regarded as a static, non-functional structure that evolved merely as a requirement for DNA packaging. Now there is experimental evidence that chromatin can provide control over gene expression (Felsenfeld, 1992) and a variety of proteins have been identified that modify or stabilise chromatin structure. Changes in chromatin structure within the compacted nuclear environment are thought to have an important regulatory role by modifying the accessibility of the transcription machinery to the DNA. Further, it is now recognised that heritable changes in gene expression occur without accompanying changes in linear DNA sequence. Research into 'epigenetic' control mechanisms of gene expression has intensified in recent years and there is now evidence that epigenetic memory acts to maintain differentiated states during cell division (Turner, 2002).

All the differentiated cell types that constitute a metazoan organism share the same genome in its highly compacted form, but they use only part of it to accomplish their functions because they are structurally and functionally heterogeneous. The establishment of these unique cell types is owned to correct temporal and spatial patterns of gene expression, which also need to be stably inherited by daughter cells throughout a number of cell divisions. Different patterns of gene expression start to be established at the stage of the embryonic stem cells (ES). ES cells have the ability to self-renew and to differentiate into all tissues of the embryo, a state that is known as multipotency. Differentiation and commitment of multipotent precursor cells to specific lineages implies the activation and/or silencing of many tissue specific genes. Most of the studies on gene regulation have focused on mechanisms of transcriptional activation. However, cell identity can also be described in terms of genes that are actively silenced and gene silencing is known to be equally prominent in cell fate decision, developmental progression and cell

activation (Fisher, 2002; Fisher and Merkenschlager, 2002; Grewal and Elgin, 2002).

The aim of this project was to investigate the role of epigenetic modifications in developmental and long term silencing of the *λ5-VpreB1* B cell-specific domain. The silenced state of the *λ5-VpreB1* locus was investigated in precursor cells with different levels of plasticity and during B cell development.

1.2 Eukaryotic chromatin organization

The genetic information of a eukaryotic cell is stored in DNA molecules that are over two meters in length. This material is compacted nearly one hundred thousand times in order to fit into the limited volume of the nucleus (Chakravarthy et al., 2005b). Folding of the DNA together with an equal mass of protein forms a compact nucleoprotein complex called chromatin (Widom, 1998). The first, and most well-defined, level of chromatin organization is the nucleosome, which consists of 147 base pairs (bp) of DNA wrapped in a left-handed spiral around the an octamer of basic proteins, the histones (Luger, 2003; Luger et al., 1997). About 80 per cent of genomic DNA is packaged into nucleosomes. The four core histone molecules H2A, H2B, H3 and H4 have been highly conserved through evolution (Wolffe, 1995). Histones are organised into the central $(H3/H4)_2$ tetramer and the two peripheral H2A/ H2B dimers. The assembly of a stable nucleosome core depends on the initial heterodimerization of H3 with H4 and the subsequent dimerization to form the $(H3/H4)_2$ tetramer (Eickbush and Moudrianakis, 1978). The $(H3/H4)_2$ tetramer can form a stable complex with 121 bp of DNA (Hayes et al., 1991). Histones H2A and H2B form a stable heterodimer in a manner that is structurally homologous to H3/H4, but do not self-assemble into stable tetramer complexes (Arents et al., 1991). Binding of the H2A/H2B dimers to either side of the $(H3/H4)_2$ tetramer extends binding of DNA up to 147 bp. DNA is wrapped around the histones about 1.65 times, severely distorting its path and reaching a level of compaction of about 7-fold (Luger et al., 1997). There are more than 116 direct and 358 water-bridged interactions between the histone octamer and the DNA, which make the nucleosome a quite stable particle (Davey et al., 2002; Luger and Richmond, 1998) (Fig.1.2.1).

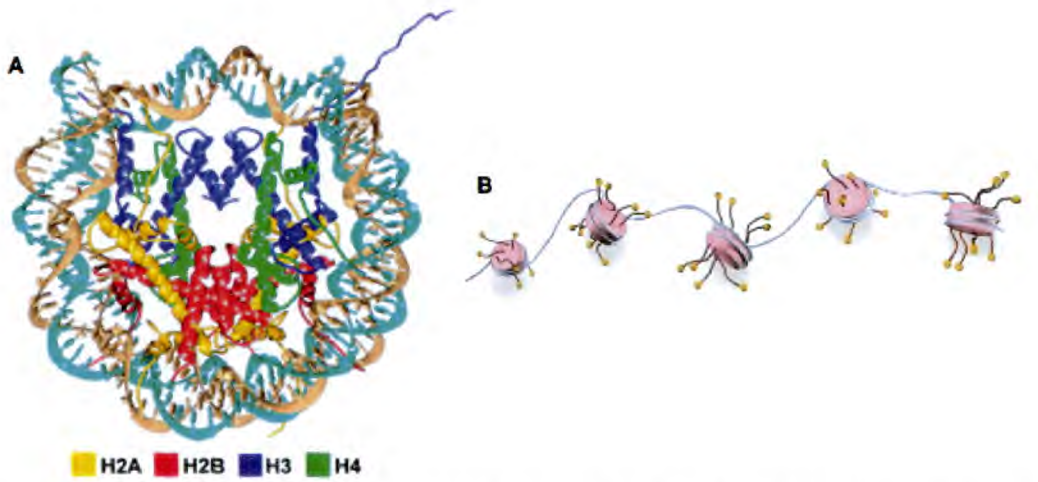


Fig.1.2.1 **A**) X-ray structure of the nucleosome core particle taken from (Luger et al., 1997). Half the core particle with four histone molecules and 73 bp of DNA is shown. The N-terminal tail of histone H3 is shown. **B**) The nucleosome at the lowest level of organisation in which approximately 146 base pairs of DNA is wound in two super helical turns around the outside of a histone octamer. The nucleosomes are connected to one another by short stretches of linker DNA. "Beads on a string" structure.

Nucleosomes are arranged like 'beads on a string', separated by short segments of linker DNA, which account for the remaining 20 per cent of genomic DNA. The nucleosomal array is the basic functional unit of chromatin (Van Holde, KE. 1989, 'Chromatin', Springer Verlag, New York) (Fig.1.2.1).

The nucleosome core particle (NCP) can be assembled with histone variants (Suto et al., 2000) or with histones carrying functionally important mutations (Flaus et al., 2004; Muthurajan et al., 2004). These modified NCPs present subtle, yet important, modulation of the histone-histone or histone-DNA interaction but maintain the overall structure unchanged (Chakravarthy et al., 2005b). For example H2A.Z, which shares only 60% sequence identity with H2A, forms an NCP with an extended acidic patch on the surface (Suto et al., 2000). This change in the NCP surface has been shown to be important for the interaction between nucleosomes and non-histone proteins such us HP1-α (Fan et al., 2004). MacroH2A, another H2A variants, contains a large C terminal region, which has been shown to contain a binding site for histone deacetylases (HDAC). Evidence for such an interaction has been obtained consistent with the presence of macro H2A in the transcriptionally repressed X chromosome and pericentric heterochromatin (Chakravarthy et al., 2005a).

Each histone can be divided into two functional domains: the histone fold, which is required for the organisation of the histone octamer and DNA binding, and the N- and C-terminal domains ('tails'), which are the main targets for post-translational modifications. The N-terminal domains (NTDs) of the histones form interactions with linker DNA and other nucleosomes, thereby modulating the formation of chromatin higher order structure (Hansen, 2002; Hayes and Hansen, 2001). Unfortunately, because of their unstructured nature, most of the N terminal domains of the core histones are missing from X-ray data. These domains are important in promoting chromatin folding and compaction (Luger and Hansen, 2005). Dorigo and colleagues (Dorigo et al., 2003) showed that the histone H4 NTD is needed to assemble nucleosomal arrays in to an extensively condensed 30nm diameter secondary chromatin structure. Other studies have confirmed the importance of the specific H4 NTD-H2A interaction (Dorigo et al., 2004; Fan et al., 2004).

The 30 nm fibre was originally suggested to have a solenoid structure with coiling of the linker DNA and formation of super helical turns (Fig.1.2.2). An alternative model proposed an irregular zigzag pattern with the linker DNA inside the fibre (Woodcock et al., 1993). In accordance with this model, Dorigo et al., have shown *in vitro* that nucleosome arrays are organised as a zigzag two-start helix (Dorigo et al., 2004). The two-start model comprises two crossed-linker adjacent stacks of helically arranged nucleosomes (Fig.1.2.2). Native chromatin can have more extreme compaction levels than the 30 nm fibre (Daban, 2000). However, less is known about how these fibres are further compacted (about 400-fold) into higher chromatin structures in the interphase nucleus. The 30 nm fibres could be further compacted or coiled, or just be arranged side to side, forming a hierarchy of folding levels, resulting in fibres of 300 nm (Hendzel et al., 1998).

During the past several years it has been established, by *in vitro* studies, that many nucleosome-binding proteins possess the ability to assemble specific secondary and tertiary chromatin structures. Among these proteins, the linker histone H1 represents a special case, since it occurs in high abundance (up to 1 molecule per nucleosome) and is usually considered to be a basic chromatin component. There is ample evidence that the globular domain of H1 binds close to the linker DNA entry–exit sites inducing a compact zigzag secondary structure and inhibiting nucleosome sliding and transcription in vitro (Woodcock et al., 2006).

Other chromatin architectural proteins are much less ubiquitous or abundant than H1.

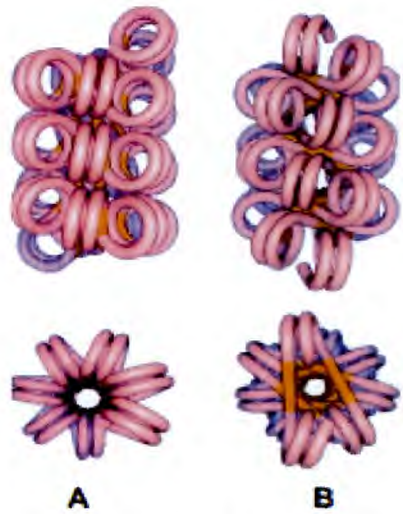


Fig. 1.2.2 Organisation of DNA within chromatin structure (adapted from Dorigo et al., 2004). (A) In the one-start solenoidal model nucleosomes are connected by short stretches of linker DNA and the 30 nm fibre is formed with linker histones (H1). The upper view shows the fibre axis running vertically. The lower views are down the fibre axis. The linker DNA running between nucleosome cores is orange. (B) The two-start twisted model of chromatin structure.

MeCP2, originally identified as a methyl CpG binding protein, is of particular interest because specific mutations in this protein have been associated with most cases of the human neurodevelopmental condition Rett syndrome (Caballero and Hendrich, 2005). It has been identified as a chromatin architectural protein on the basis of the profound compaction it induces in nucleosomal arrays (Georgel et al., 2003). The precise functions of MeCP2 are still uncertain, but there is evidence that, when bound to methylated DNA, it recruits HDACs, leading to local transcriptional repression (Suzuki et al., 2003). Furthermore, MeCP2 is required to stabilize large chromatin loops (Horike et al., 2005).

Polycomb is a large multicomponent complex that is required to silence some early development genes (Ringrose and Paro, 2004). It has also shown to be able to compact nucleosomal arrays by binding to nucleosomes. A single polycomb complex has been shown to be able to bind three to four nucleosomes in a sequence-independent way (Francis et al., 2004). Studies of the interactions of heterochromatin protein 1α (HP1α) with nucleosomal arrays identified a fundamentally different type of chromatin architectural protein (Fan et al., 2004).

These studies showed that HP1α prefers to bind to folded nucleosomal arrays and that it induces local rearrangement of secondary chromatin structures.

1.3 Euchromatin and heterochromatin

The first to observe different types of chromatin in the nucleus was Emil Heitz in 1928. Cytogenetic studies on plant cell nuclei stained with carmine acetic acid (Heitz, 1928) revealed a type of chromatin that remained condensed throughout the cell cycle. This darkly staining chromatin was called heterochromatin. The other type, which was undergoing normal cycles of condensation and de-condensation, was named euchromatin or 'true' chromatin.

General characteristics of chromatin

	Euchromatin (active genes)	**Constitutive heterochromatin** (repetitive DNA)	**Facultative heterochromatin** (silenced genes)
CpG	Hypomethylation	Methylation	Methylation
Histone modifications	H3/H4 hyperacetylation H3 K4 dimethylation H3 Arg 2, Arg 7, Arg 26 methylation. H3 posphorylation H3 K79 methylation	H3/H4 hypoacetylation H3 K9 trimethylation H4 K20 trimethylation	H3/H4 hypoacetylation H3 K27 trimethylation H3 K9 dimethylation H4 K20 monomethylation
Tipical protein factors	Activator proteins Chromatin remodelling complexes	HP(1) proteins Su(Var)3-9	PcG complex H(P)1 ? Activator proteins Chromatin-remodelling complexes
Replication timing	Predominantly early	Late	Variable
General chromatin features	"Open" chromatin	Compacted chromatin Clustered centromeres	Compacted chromatin
Transcription	Active gene transcription Non-genic transcription	Small RNA transcripts	No gene transcription Non-genic trascription unk

Table 1.3.1 General characteristics of euchromatin, constitutive heterochromatin and facultative heterochromatin (adapted from Arney and Fisher, 2004).

1.3.1 Euchromatin

Euchromatic regions of the genome are cytologically observed as more decondensed, and located near the centre of the interphase nucleus. Euchromatin normally contains a high density of functional genes and often replicates earlier in S-phase than heterochromatic DNA (Dimitri et al., 2005). Histones in euchromatin may be hyperacetylated (Table 1.3.1) and CpG dinucleotides in the associated DNA are often unmethylated, particularly in expressing genes (Eden et al., 1998).

Methylation-free islands containing clustered CpG sequences occur near the transcription start site of many ubiquitously expressed genes and approximately 40% of mammalian genes showing a tissue-specific expression pattern (Cross and Bird, 1995). A number of studies on accessibility of DNA within the nucleus to nucleases such as DNase I have shown that transcriptionally active or potentially active genes have nuclease hypersensitive sites (HS) located upstream and/or downstream of the gene. The increased sensitivity to DNase I observed at hypersensitive sites is likely to be due to displacement or modification of a nucleosome caused by transcription factor binding (Kornberg and Lorch, 1999).

1.3.2 Heterochromatin

Heterochromatin, or darkly stained chromatin as described by Heitz (1928), remains as a visible highly condensed region of the genome throughout the cell cycle. Constitutive heterochromatin is associated with repetitive DNA sequences and contains a low density of genes that replicate late in S phase. Heterochromatic DNA also tends to be highly methylated. It is predominantly located at pericentromeric chromosomal domains, which contribute to centromeres in mammals (Grewal and Moazed, 2003). Pericentromeric heterochromatin is associated with the presence of a regular arrangement of nucleosomes (Wallrath and Elgin, 1995). The heterochromatin surrounding centromeres (pericentromeric heterochromatin) is usually composed of repeat sequences that vary widely in size between species from the 5-7 bp satellite repeats in Drosophila to the 234 bp pericentromeric heterochromatin in mouse. The effect of such areas is to create a relatively uniform sequence composition over a large region and to reiterate sequences with specific properties i.e. factor binding sites and nucleosome

positioning elements. These features contribute to stable formation of condensed chromatin structures. Furthermore, the association of heterochromatin proteins such as heterochromatin protein 1 (HP1) with repetitive centromeric sequence motifs has shown to be important for efficient cohesion between sister-chromatids and chromosome segregation in fission yeast (Grewal and Moazed, 2003 and reference herein). Other experiments with fission yeast have indicated that small RNAs associated with the RNA interference (RNAi) pathway are transcribed from centromeric repeats (Volpe et al., 2002).

Since its identification, heterochromatin has been a useful model for studying chromatin condensation and heritable gene silencing. More recently, studies that focus on the mechanisms that underlie its structure have been carried out to help to determine the link between visible heterochromatin structures and gene silencing. The observation that heterochromatin is epigenetically marked by specific histone modifications has enhanced our understanding of how it is formed. Epigenetic marks that have been associated with heterochromatin include generalised hypoacetylation of histones (Jeppesen and Turner, 1993) and trimethylation of lysine 9 of histone H3 (H3K9me3) (see Table 1.3.1) (Dillon, 2004). The observation by Jenuwein and colleagues that the Su(Var)39H gene encodes a histone methyltransferase protein specific for H3K9 provided the first evidence that methylation of this residue is directly involved in heterochromatin formation (Rea et al., 2000). Highly specific methylation of H3K9 by Su(Var)39H histone methyltransferase creates a high-affinity HP1 binding site. In turn, HP1 interactions are thought to promote the formation of condensed chromatin structures (Nielsen et al., 2001).

Facultative heterochromatin is defined as a region of chromatin that becomes subject to large-scale condensation during development. The best-known example of facultative heterochromatin is formed during X chromosome inactivation in female mammals. During X inactivation, transition from a permissive chromatin state to transcriptionally inert heterochromatin is visible as the condensed Barr body located close to the nuclear periphery in female somatic cells (Barr, 1949). Facultative heterochromatin contains modifications that are associated with transcriptional repression: di-methylation of histone H3K9, hypoacetylation of lysine residues within histones H3 and H4, and DNA methylation (Table 1.3.1). Methylation of H3K27 is thought to be a feature of facultative heterochromatin and has been

associated with transcriptional silencing by Polycomb group proteins as well as being found on the inactive X chromosome.

1.4 Overview of transcription

In eukaryotes, transcription is a tightly regulated process by which three nuclear RNA polymerases (RNA pol I, pol II and pol III) and several small nuclear RNAs synthesise all of he RNAs encoded by the genome (Roeder and Rutter, 1969, reviewed by Sims et al., 2004). RNA pol I and III transcribe the genes for ribosomal RNA (rRNA) and transfer RNA (tRNA) respectively. RNA pol II is responsible for transcribing all the protein-encoding genes. RNA pol II consists of 12 subunits, which are remarkably conserved throughout eukaryotes (Young, 1991). The generation of a mature mRNA molecule involves multiple processes, some of which occur sequentially and others in parallel. The primary phases of this process form a cycle (transcription cycle) that includes pre-initiation complex (PIC) assembly, PIC activation, initiation, promoter clearance, elongation and termination (Kornberg, 1999).

Unlike the prokaryotic enzyme, eukaryotic RNA pol II is not able to recognise the promoters of its target genes. A series of general transcription factors (GTFs) are necessary for pol II to bind to the transcription start site (Orphanides et al., 1996). GTFs recognise conserved core promoter elements. Two important core promoter elements are the TATA box, located near position -30 to -25 and the initiator sequence (Inr) located at the start site of transcription. The TATA box is recognised by the TATA box binding protein (TBP), which is part of the TFIID complex, which also recognise Inr. TFIID contains many TBP-associated factors (TAFs), which are implicated in sequence specific binding (reviewed by (Butler and Kadonaga, 2002)).

Transcription initiation is thought to begin with the assembly of the RNA polymerase and five general initiation factors (TFIIB, TFIID, TFIIE, TFIIF and TFIIH) into a pre-initiation complex (PIC) at the promoter (Orphanides and Reinberg, 2002; Reese, 2003). In the first step, TBP binds sequence-specifically to the promoter to establish a nucleoprotein recognition site for pol II on the DNA (Burley and Roeder, 1996). In the second step, TFIIB binds to the polymerase and to the TBP subunit of TFIID to form the Dbpol II complex. TFIIF has the double role of stabilizing the Dbpol II intermediate and recruiting TFIIE and TFIIH into the complex (Dvir et al., 2001). Formation of this promoter-bound complex is sufficient

for a basal level of transcription. However, regulation of this process in response to environmental signals involves a number of additional co-activators. These are recruited to enhancers or upstream activating sequences (UAS) by gene-specific activators and operate by several distinct mechanisms involving alteration of the structure of chromatin and direct interaction with RNA pol II and GTFs. Three classes of general cofactors, including TBP-associated factors (TAFs), Mediator, and upstream activating sequence (USA)-derived positive and negative cofactors (PC1/PARP-1, PC2, PC3/DNA topoisomerase I, PC4, NC1/HMGB1) normally fine-tune the promoter activity in a gene-specific or cell-type-specific manner (reviewed by (Martinez, 2002)).

Once the PIC is fully assembled, an ATP-dependent process, which requires the action of the helicase TFIIH, cause the opening of the double stranded DNA at the transcription start site (promoter melting) and the formation of the open complex. Transcriptional initiation then occurs upon addition of the two initiating nucleoside triphosphates (NTPs) depending on the DNA sequence and formation of the first phosphodiester bond (Goodrich and Tjian, 1994; Holstege et al., 1996; Kim et al., 2000).

Before RNA pol II engages in productive transcript elongation it must go through a stage known as promoter clearance. During this stage the PIC is partially disassembled. A subset of GTFs remain at the promoter where they constitute a scaffold for the formation of the next transcription initiation complex (Yudkovsky et al., 2000; Zawel et al., 1995). This stage is accompanied by massive phosphorylation at the C-terminal repeat domain of the largest subunit of RNA pol II (CTD). The CTD contains multiple repeats of the hepta-peptide sequence YSPTSPS. The number of these repeats increases with genomic complexity: 26 in yeast, 45 in Drosophila, and 52 in mammals (Sims et al., 2004). It has been demonstrated that RNA pol II in the PIC is un-phosphorylated (Laybourn and Dahmus, 1989; Lu et al., 1991), whereas the transcription-competent form is heavily phosphorylated on its CTD (Christmann and Dahmus, 1981). Serine 2 and 5 are the major sites of posphorylation in the YSPTSPS repeat, but it is not well understood which of these phosphorylation event is most important for the disruption of the connections with the promoter. It is also possible that while leading to disruption of one set of interactions, CTD-phosphorylation might be relevant for establishing another set (Svejstrup, 2004).

To ensure efficient elongation, the phosphorylation state of the CTD acquired at the initiation-elongation transition needs to be maintained (Lee and Greenleaf, 1997; Marshall and Price, 1995). ChIP data suggest that phosphorylation at Ser 5 is associated with promoter-proximal regions of transcribed genes and is apparently not detected in the 3-end region. At the same time, the amount of the enzyme phosphorylated at Ser 2 increases toward the 3-end of genes (Cho et al., 2001). These results, which are controversial, suggest that CTD phosphorylation at Ser 5 correlates with transcription initiation and early elongation (promoter clearance), whereas Ser 2 phosphorylation is associated with RNA pol II farther away from the promoter (Cho et al., 2001). These data have been partly questioned by the finding that the Ser 2-specific antibody recognizes both the Ser 5 and Ser 2 phosphorylated forms (Jones et al., 2004) and the disappearance of Ser 5 phosphorylation at the 3' end will need to be confirmed.

A variety of evidence suggests that purified RNA polymerase II lacks the ability to catalyze mRNA chain elongation in vitro at rates sufficient to account for the rates of mRNA synthesis observed in vivo (reviewed by Reines et al., 1999; Shilatifard et al., 2003). Several factors have been isolated that can alter the kinetics of transcription in vitro. This class of RNA polymerase II elongation factors includes TFIIF, the Elongins, DSIF, CSB and the ELL family of proteins. All these factors have similar kinetic properties, they interact directly with elongating polymerase and boost its elongation rate in vitro but they do not have overlapping or redundant functions in cells. Mechanistic studies have demonstrated that these factors function by suppressing transient pausing by the polymerase at all or most steps of nucleotide addition (reviewed by Sims et al., 2004).

Other factors that do not have consequences on transcription kinetics have also been reported to be associated with the elongating RNA pol II. Among these is the Paf1 complex, which function at both the initiation and elongation stages of transcription (Mueller and Jaehning, 2002; Pokholok et al., 2002). It has been demonstrated that the Paf1 complex is physically associated with elongating RNA polymerase II and can function as a platform for the recruitment of the histone modifying machinery to the transcribing polymerase (Krogan et al., 2003a; Krogan et al., 2003b). The Elongator complex, which is associated with elongating, hyper phosphorylated, RNA pol II in yeast (Otero et al., 1999), has a HAT activity predominantly addressed to lysine 14 of histone H3 and lysine 8 of histone H4

(Winkler et al., 2002). In yeast it has been shown that transcription is reduced in genes that have low acetylation levels in the middle of the coding region, consistent with the idea that hypoacetylation can inhibit transcriptional elongation (Kristjuhan et al., 2002). The observation that mutation of the gene encoding the histone deacetylase Hos2 results in slow gene activation in yeast (Wang et al., 2002) coupled with the idea that hypoacetylation inhibits transcription suggest that not only hyperacelation of histones enables the passage of RNA pol II but also that immediate deacetylation after this passage could stabilize chromatin structure to ensure continued ability to be transcribed (Svejstrup, 2003; Svejstrup, 2004).

The last step before starting the next round of transcription, termination, is critical for successful gene expression. It allows the release of transcripts from the site of transcription and also releases pol II from the template, facilitating its recycling back to the promoter for further rounds of transcription. Transcript termination is dependent on the presence of a functional poly(A) signal (reviewed by Proudfoot, 1989). Cleavage and polyadenylation of the nascent transcript are distinct from termination, which in principle releases the RNA. It has been shown that RNA pol II can transcribe up to 2 kb after RNA cleavage, although cleavage is necessary for termination. The only termination factor associated with early transcript elongation complexes (TECs) known until now is the RNA pol I and II specific termination factor (TTF2). TTF2 can terminate transcription at most positions along a DNA template and was shown to act independently of the state of CTD phosphorylation (Jiang et al., 2004).

There is evidence that chromatin structure in the terminator region might play an important role and that ATP-dependent chromatin remodelling factors, such us yeast Chd1, Isw1 or Isw2 play a role in the termination process (Alen et al., 2002; Morillon et al., 2003). DNA dependent ATPases capable of disrupting ternary complexes (RNA polII/DNA/RNA) and terminating transcription have been identified in both prokaryotes (Richardson, 2002; Selby and Sancar, 1993) and eukaryotes (Xie and Price, 1997). Transcription is a tightly regulated process, which involves many steps and the concerted action of many different factors. The presence of cis-regulatory elements other than promoter element is without any doubt imposing a further level of control over the transcription process.

1.5 Cis-acting regulatory sequences: promoters and enhancers

DNA is packaged as chromatin within the nuclei of eukaryotes, a structure that must be disrupted when the genes are transcribed by RNA pol II. Many sequences located upstream of the transcriptional start site of the gene (promoter) and around or within the gene itself are important in the control of gene expression (Table1.5.1). DNase I hypersensitive sites are stretches of DNA, usually 200 bp, that are rich in binding sites for transcription factors. When a sufficient level of associated binding proteins are present, these sites are often exposed as sequences hypersensitive to digestion with DNAse I. Hypersensitive sites (HS) have a prominent role during cell differentiation since the formation of specific hypersensitive sites in progenitor cells often precedes transcriptional activation at later stages (Blom van Assendelft et al., 1989; Hatzis and Talianidis, 2002; Szutorisz et al., 2005).

Gene promoters were the first of a number of functional DNA sequences to be described that are required for gene expression and that are targets for regulatory proteins. Promoters are defined as spatially constrained regions that specify the site of transcriptional initiation. Protein-coding genes have them typically located immediately upstream of the site of transcriptional initiation.

Enhancers are DNA sequences that can activate a promoter and control its efficiency and rate of transcription independently of orientation and with some flexibility with respect to distance. Promoters and enhancers contain groups of transcription factor binding sites, which are often repeated. The spacing and combination of sites is believed to be significant for enhancer function (Zenke et al., 1986).

Functional definitions used to classify *cis* -acting sequences that affect transcription	
Promoter	Spacially constrained region that specifies site of transcriptional initiation in *in vitro* and *in vivo* assays. Promoters bind the basal transcriptional machinery and are generally located upstream of the site of transcriptional initiation for protein encoding genes
Enhancer	DNA element, linked in *cis* to the promoter (upstream or downstream), which acts to increase transcription in an orientation independent manner and with some flexibility with respect to distance. Originally defined in transient expression assays in cultured cells. However, the level of transcriptional enhancement or the assay used to measure it is not specified
Locus Control Region (LCR)	Dominant activating sequences that allow a gene to be expressed independently of its site of integration into the genome. LCRs confer position-independent and copy-dependent expression on a linked gene in transgenic mice and are capable of activating expression at single copy
DNase I hypersensitive site (HS)	Consist of a region of 150-300 bp with an increased sensitivity to DNase I digestion compared to surrounding chromatin. HS have a high density of binding sites for *trans* -acting factors, and are detected by digesting isolated nuclei with increasing amounts of DNase I

Table 1.5.1 Glossary of terms

In eukaryotes, activator proteins bind at distinct sites upstream or downstream of the promoter and interact in a synergistic manner with co-activators and core promoter elements (CPE) (Ptashne and Gann, 1997). The core promoter is defined as the minimal DNA region that is sufficient to direct low levels of activator-independent (basal) transcription by RNAP II in vitro. The core promoter typically extends 40 bp upstream and downstream of the start site of transcription and can contain several distinct sequence elements. In higher eukaryotes, CPEs are highly diverse in structure, and each of them is only found in a subset of genes. So far, only the TATA box and INR (initiator) element have been shown to be capable of directing accurate RNA pol II transcription initiation independent of other core promoter elements. Transcription is controlled by a combination of CPEs and gene-specific enhancer elements that co-operatively define expression patterns (Dynan, 1989). Enhancers are often found proximal to the core promoter and may include canonical sequence elements such as the GC box and CCAAT box (Nussinov, 1990).

A further level of complexity is added by distal enhancer elements such as hormone response element (HREs) (Archer et al., 2005; Kardassis et al., 1997).

Enhancers act in conjunction with linked promoters to increase transcription. They function as autonomous regulatory units with respect to orientation and distance from sites of initiation and often contain clusters of activator binding sites. The 'classical' enhancer effect was first characterised in studies of the *SV40* enhancer (Treisman et al., 1983). Weintraub and colleagues (1988) showed that the SV40 enhancer increases the probability of a transfected template to become active but its presence or absence does not change the activity (the amount of transcript produced by the gene in a given period of time) of an active template (Weintraub, 1988). Enhancers in transgenic mouse studies have been shown to increase the proportion of cells expressing the transgene (Sutherland et al., 1997).

1.6 Locus Control Regions (LCRs)

Early transgenic experiments indicated that inclusion of enhancer and promoter sequences within a construct was not sufficient to ensure correct levels of tissue specific expression. Also, mouse lines with multicopy transgene exhibited expression levels that had no relationship with the number of integrated transgene copies. These findings suggested that both chromosomal location and lack of transcriptional activation from cis-acting elements can exercitate a negative effect on tissue specific expression.

Repression resulting from translocation of genes to abnormal locations had been observed in very early studies (Dobzhansky, 1936; Lewis et al., 1992; Muller, 1930). The phenomenon was first observed in 1930 when Herman Muller described mutation in the drosophila *white mottled-4* gene that gave phenotypes with mottled/variegated, red or white eyes (Muller, 1930). This phenotype was traced to the fact that when the white-eye colour gene was juxtaposed next to pericentromeric heterochromatin it was often subject to unstable repression. Therefore, the variegation observed was the result of placement next to the heterochromatin and not due to mutations in the gene itself. This phenomenon was termed Position Effect Variegation (PEV).

Locus control regions (LCRs) are regulatory elements able to overcome such position effects. LCRs can direct tissue-specific expression of a transgene at high levels irrespective of the site of integration in the host genome (Grosveld et al., 1987). This implies that this element is capable of inducing the changes in chromatin structure that are required for expression of the genes. In their native

chromosomal locations, LCRs are required for the initial chromatin activation of a locus and subsequent gene transcription.

1.6.1 The β-globin locus control region

The first LCR was identified in the human β-globin locus (Grosveld et al., 1987), since then LCRs have been identified for many genes (Table 1.6.1). The human *β-globin* cluster consists of five erythroid specific genes arranged along the chromosome in the order in which they are expressed during development (Fig 1.6.1). Upstream of the genes there are five DNAse I hypersensitive sites (HSs) within a 20kb region. Each HS contains numerous binding sites for ubiquitous and erythroid restricted trans-acting factors. Early studies showed that a 5kb *β-globin* gene segment, including a 1.5kb promoter region give uniform gene expression in transgenic mice (Kollias et al., 1986; Magram et al., 1985; Townes et al., 1985). The gene was expressed only in a small proportion of transgenics, with expression levels far below the normal and subject to position effects. These studies implied that a regulatory element present in the native locus was missing in these constructs.

Locus or gene	Tissue specificity	Reference	Organism
b-globin gene	erythroid cells	Grosveld 1987	Human
CD2 gene	T cell specific	Greaves 1989	Human
CD4 gene	T cell specific	Boyes 1997	Human
TCR a/d	T cell specific	Diaz 1994	Human
MHC class II Ea gene	B cell specific	Carson 1993	Mouse
S100b gene	neuron	Friend 1992	Human
Lysozyme gene	macrophage	Bonifer 1990	Chicken
LAP	liver	Talbot 1994	Rat
l5-VpreB1 locus	B cell specific	Sabbattini 1999	Mouse

Table 1.6.1 Examples of gene loci with LCR activity modified from Li et al 1999

It was observed that in some forms of β-thalassemia the genes of the *β-globin* locus are intact but not expressed (Driscoll et al., 1989; Kioussis et al., 1983). A common defect underlying these conditions is the presence of large deletions upstream of the β-like globin genes (Duch and Hispanic deletions) (Fig. 1.6.1). The

hispanic deletion results in a closed chromatin conformation spanning the whole locus that leads to suppression of gene expression (Forrester et al., 1990). The existence of a regulatory element was also suggested by the presence of developmentally stable, erythroid-specific, DNAse I HSs located 5' of the ε-globin gene (Forrester et al., 1986; Tuan et al., 1985). Definitive evidence for the presence of the LCR came from transgenic mouse experiments. Linkage of this region to a *β-globin* gene resulted in expression of the gene at level comparable to the endogenous mouse *β-globin* gene in a position-independent, copy number-dependent manner (Grosveld et al., 1987).

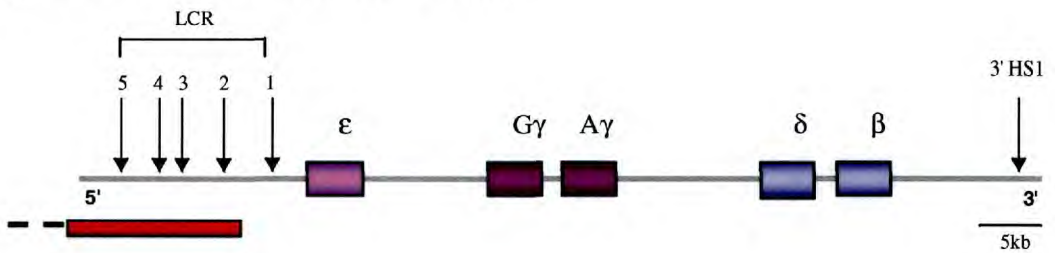


Figure 1.6.1 The Locus Control Region of the human β-globin gene cluster. HS are indicated by vertical arrows (Grosveld et al., 1987). The Hispanic β thalassemia deletion is shown as a red rectangle. The genes are arranged along the locus in the order in which they are expressed. The ε is expressed during embryogenesis, Gγ and Aγ during foetal development and δ and β in the adult.

1.6.2 Properties of LCRs: transcriptional enhancer activity

The most prominent property of the LCRs is their strong, transcription-enhancing activity. When the human *β-globin* LCR is deleted transcription of the human β-globin gene is usually less then 1% of the endogenous murine *β-globin* mRNA in transgenic mice (Kollias et al., 1986; Magram et al., 1985; Townes et al., 1985). Inclusion of the LCR increases the *β-globin* gene expression to a level comparable with that of the endogenous gene in transgenic mice indicating a strong enhancer activity of the LCR (Grosveld et al., 1987).

The *β-globin* LCR enhancer activity is also significant at its endogenous location as demonstrated by deletion experiments (Bender et al., 1998; Epner et al., 1998; Reik et al., 1998). The enhancer activity of the LCR resides in HS2, 3 and 4 but not in HS 1 and 5. HS 2 enhancer activity can be detected by transient transfection assays, while enhancer activity from HS 3 and 4 could only be detected when they

are integrated into chromatin. Importantly, the enhancer activity of the *β-globin* LCR is tissue specific. Expression of the β-globin genes is confined to erythroid cells when linked to the *β-globin* LCR. In addition, the LCR is able to enhance the expression of linked non-globin genes promoters in erythrocytes. However, ectopic expression was observed for these genes in some transgenic mice. The natural promoter in this case allowed expression outside the erythroid compartment. Tissue specific control of basal transcription may reside in the promoter while tissue specific enhancement of gene expression may be a property of the LCR. Very important for the understanding of enhancer activity of the *β-globin* LCR was the identification of the transcription factors mediating enhancer activity. Deletion analysis of factor binding motifs within the HS was carried out to determine their role (Philipsen et al., 1990; Pruzina et al., 1991; Talbot and Grosveld, 1991). This analysis defined the essential sequences of 5' HS sites 2, 3 and 4 in the human *β-globin* LCR. The HS fragments were shown to contain binding sites for erythroid-specific proteins GATA1 and NF-E2 and a number of general transcription factors that included SP1 and AP1. A notable feature was the recurring presence of a G-rich sequence flanked by GATA1 binding sites, particularly at HS3. At the core of HS3 are closely spaced triple repeats of GATA1 binding sites and G-rich sequences (Philipsen et al., 1990). When introduced into cultured cells or transgenic mice, HS3 was shown to be sufficient for copy number dependent expression of a β-globin gene (Fraser et al., 1990). Deletions of HS3 factor-binding sites revealed that a specific arrangement of binding sites and the cooperation of at least three proteins was required to achieve position independent expression (Philipsen et al., 1990).

1.7 Domains and genetic functional units

1.7.1 Strong and weak domain models

The characterisation of specific cis-acting regulatory sequences and the factors that bind to them is not enough to unravel the mechanisms of gene regulation. Correctly regulated expression of a eukaryotic gene *in vivo* is not simply a function of a single sequence element but more often of multiple sequences acting co-operatively and spread over large distances. These sequence elements define the complete functional domain that can correctly regulate a gene locus. In higher organisms, genomes comprise multiple unique gene loci with quite different patterns

of expression located in close proximity to one another on the chromosome. Autonomy of the adjacent loci must be maintained with respect to gene expression patterns. Understanding how adjacent loci are independently regulated has formed the basis for a number of models aimed at defining the functional gene expression domain (Dillon and Sabbattini, 2000).

The classical picture of domain arose from early light and electron microscopy studies of the interior of the nucleus where genomic DNA appeared to form loops anchored to a nuclear matrix (Manuelidis, 1990). The anchoring DNA sequences were postulated to participate in demarcating selected chromosome domains isolated by these 'boundaries'. The concept of a structural domain physically separated from neighbouring domains by boundaries was considered to mirror a functional compartmentalisation of gene loci. Recent functional assays have shown that these matrix attachment regions (MARs) do not generally constitute boundaries between active or inactive chromatin (Wang et al., 1996). The next generation of domain models proposed that transcriptionally active genes are located within region of "open" extended chromatin. Neighbouring inactive domains would instead have a closed compacted chromatin structure and the borders between the active and inactive domains would be defined by discrete boundary or insulator elements which would have the specific function of blocking against incursions from surrounding chromatin domains. This type of model has been termed the 'strong' domain model (Dillon and Sabbattini, 2000).

The strong domain model has been based mainly on studies of two loci, the drosophila *87A7 heat-shock* locus and the chicken *β-globin* locus. The *87A7 heat-shock* locus is the site of a heat-inducible chromatin puff on polytene chromosomes (Udvardy et al., 1985), which suggests the formation of a decondensed chromatin region upon activation of the genes. In 1991, Kellum and Schedl were the first to test by functional assay whether two DNAseI HS located at each side of the puff (scs and scs') could establish a domain of independent gene activity in vivo and insulate against chromosomal position effects (Fig.1.7.1). The scs and scs' sequences have been shown act as insulators against position effects on a white gene. The insulator activity was dependend on the the gene been positioned in between the two elements (Kellum and Schedl, 1991). Recently, experiments from the Geyer laboratory have cast doubt on the ability of the scs and scs' elements to form boundaries to chromatin domains (Kuhn et al., 2004). In the chicken *β-globin* locus

a DNAseI HS, HS4, which is located close to the edge of a DNase I sensitive and hyperacetylated region, also has the properties of an insulator (Chung et al., 1993). Upstream of this element is located the independently regulated folate receptor gene. The region between the *β-globin* and folate receptor loci is rich in repetitive sequences and in erythroid cells is largely DNase I insensitive, hypoacetylated and hypermethylated (Hebbes et al., 1994; Prioleau et al., 1999). Insulators are operationally defined as sequences that can interfere with enhancer-promoter interactions when placed between them, and that can buffer transgenes from position effects (Bell et al., 2001).

Insulators have now been identified associated with a number of unrelated gene loci and in a variety of organisms from yeast to man (Bell et al., 2001). They have also been frequently associated with sequences that have other functions (Dillon and Sabbattini, 2000). The scs' element in the *Drosophila heat shock* locus for example has shown to be also the promoter of the *Aurora* gene (Glover et al., 1995), which is involved in spindle formation. Among the best-characterised examples of a sequence that has insulator effects associated with other functions is the *Drosophila eve* promoter. Analysis of the insulator effect showed that it is partly dependent on the presence of a GAGA sequence located between the TATA box and the initiation site (Ohtsuki and Levine, 1998). Removal of the GAGA sequence did not abolish the insulator effect of the eve promoter. This result suggested that the insulator effect is dependent on synergism between several different promoter components rather than being solely generated by the GAGA element (Dillon and Sabbattini, 2000). The vertebrate insulator element cHS4 at the 5' end of the chicken globin locus prevents the spread of methylated condensed heterochromatin from the adjacent folate receptor gene and it also acts as an enhancer-blocker separating the regulatory elements of the two nearby loci (Fig.1.7.1) (Prioleau et al., 1999).

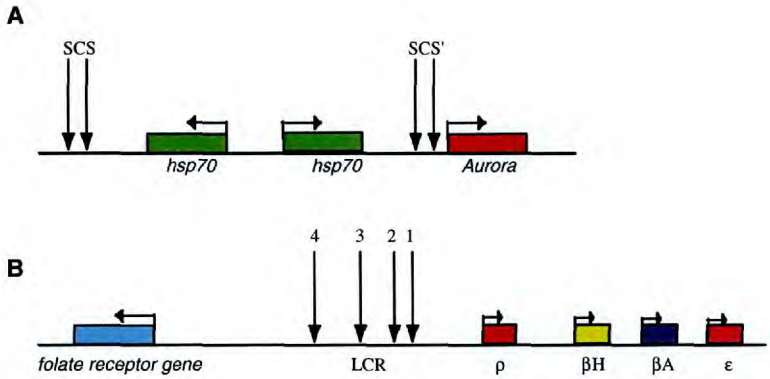


Figure 1.7.1 Insulator and enhancer blocking effects identified in different eukaryotic systems (adapted from Sabbattini and Dillon, 2000). (A) Insulator effects of 'specialised chromatin sequences' (SCS) at the Drosophila *87A7 heat-shock* locus. The SCS and SCS' insulators (indicated by vertical arrows) exist as a pair of strong DNase I hypersensitive sites flanking the Drosophila *87A7 heat-shock* locus. SCS' is also the *Aurora* gene promoter. These elements act as insulators and were identified on the basis of their ability to give position-insensitive expression when placed either side of a transgene (Kellum and Schedl, 1991). (B) The 5' end of the chicken globin LCR contains an insulator (HS4) which can block the spread of methylated condensed heterochromatin from the adjacent folate receptor gene.

The observations made above imply that insulator activity may not be a specific function. The insulator can be bound by factors that can disrupt productive interactions between promoter-enhancer, and/or it can have a blocking effect that prevents active or inactive chromatin from spreading into adjoining domains. Therefore communication between a gene and positive cis-regulatory elements, rather than the presence of specialized boundary elements, would define an active domain organisation. The 'weak' domain model proposed by Dillon and Sabbattini (2000) explain insulator distribution in terms of differential selection. It suggests that specificity of enhancer-promoter interactions is key to maintaining the functional autonomy of adjacent genes (Dillon and Sabbattini, 2000). Sequences that interfere with enhancer-promoter interactions i.e. insulators would be evolutionary selected against if they occurred within the domain but not at the edges. As a consequence, insulators would be found near the borders of domains without necessarily being selected to act as boundaries.

1.7.2 Long-range chromosomal interactions

In higher eukaryotes genes can overlap and cis-regulatory sequences can be found hundreds of kilobases away from the target gene they control. Often there are unrelated genes in between. In particular developmental and tissue-specific genes can be found in gene-dense regions and, unlike housekeeping genes, do not cluster on the chromosomes. Transcriptional activation of genes by the long-range action of regulatory sequences has long been recognised. Several models have been proposed to explain this phenomenon (Fig.1.7.2). In the looping model enhancers communicate with promoters through direct interactions between proteins bound to the DNA elements with the intervening DNA looping out (Ptashne, 1986).

An alternative model, the linking model, implies a role for the DNA in between the enhancer and promoter in supporting the transmission of a signal to create an environment favourable for transcription (Fig.1.7.2). According to this model transcription factors would bind to DNase I HS and the protein complexes assembled at the HS would extend along the chromatin fibre between the LCR and the target gene (Bulger and Groudine, 1999). Basic elements in this model are insulators, which have a position dependent ability to block the spreading of inhibitory or closed chromatin structures from neighbouring loci (Bell and Felsenfeld, 1999). Two other models have been proposed, the tracking or scanning model, which hypothesized that the transcription-activating complex recruited by the enhancer linearly tracks along chromatin until it encounters a competent promoter (Tuan et al., 1992). The facilitated tracking model instead incorporates aspects of both the looping and the tracking models (Blackwood and Kadonaga, 1998). An enhancer and its activation complex migrate along the chromatin fibre until they encounter the promoter. The chromatin between the enhancer and the promoter 'reels out' through the enhancer complex and form a loop, which is progressively enlarged during tracking.

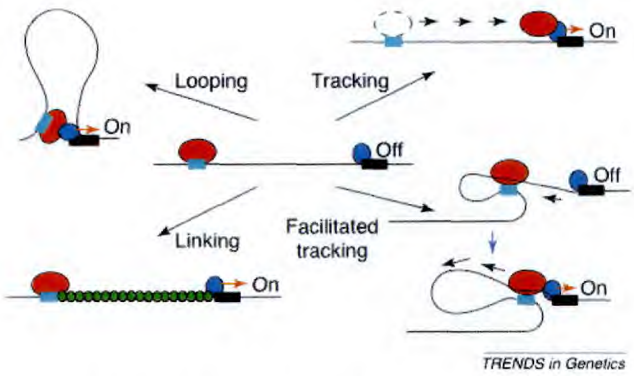


Fig. 1.7.2 Long-range chromatin interactions. From (Li, Q et al, 2006) Schematic representation of the looping, tracking, facilitated tracking and linking model. The first three models imply that the enhancer communicates with the gene promoter through direct interactions between proteins bound to the DNA elements with the intervening DNA looping out. The linking model instead proposes that "linking" proteins mediate the communication between the enhancer the gene.

1.7.3 The Active Chromatin Hub (ACH)

Recent studies would seem to support the looping model. Chromosome conformation capture (3C) and tagging and recovery of associated protein (TRAP) assay have been used to investigate the spatial organization of the βglobin locus in expressing and non-expressing tissues (Carter et al., 2002; Tolhuis et al., 2002). Tolhuis and colleagues used 3C to show that expressing adult mouse *β-globin* genes (*βmajor* and *β minor*) spatially interact with the HS of the LCR in erythroid cells. The inactive embryonic adult genes did not participate in this interaction and instead 'looped out' (Tolhuis et al., 2002). Carter et al., (2002) reported similar results using RNA TRAP. In particular they observed that the 5'HS2 is in close spatial proximity to the active globin genes (Carter et al., 2002). The finding that these HS interact with each other to form a spatial cluster with the active genes was not confined to the cis-regulatory elements of the LCR. Data collected by de Laat and Grosveld showed that two sets of HS at either side of the locus, 3' HS1 and two additional 5'HS, were also present in this spatial cluster and may be involved in globin gene regulation in erythroid cells. Importantly, the transcriptionally silent olfactory receptor (OR) genes located in between the 5' HS and the LCR, like the inactive embryonic globin genes, did not participate in clustering, but looped out (Fig.1.7.3). Analysis of non-expressing brain cells revealed that the whole region adopts a linear conformation in nuclear space in these cells (Tolhuis et al., 2002). Formation and/or stabilization of the interaction require trans-acting factors such as erythroid Kruppel

like factor (EKLF), GATA-1 and Fog-1 (Drissen et al., 2004; Vakoc et al., 2005). This complex structure containing multiple loops has been termed an "Active Chromatin Hub" (Fig.1.8.3). Using the same approach it was shown that a chromatin loop is formed between the T_H2 LCR and its cognate gene encoding interleukin 4 (*IL-4*), *IL-5* and *IL-13* in T-helper type 2 cells (Spilianakis and Flavell, 2004). In addition to enhancer-promoter loops, parent-specific loop formation was observed between the differentially methylated regions in imprinted insulin like factor 2 (*IGF2*) and *H19* a maternally expressed non-coding gene (Murrell et al., 2004). Taken together these studies show that chromatin looping brings cis-regulatory elements close together.

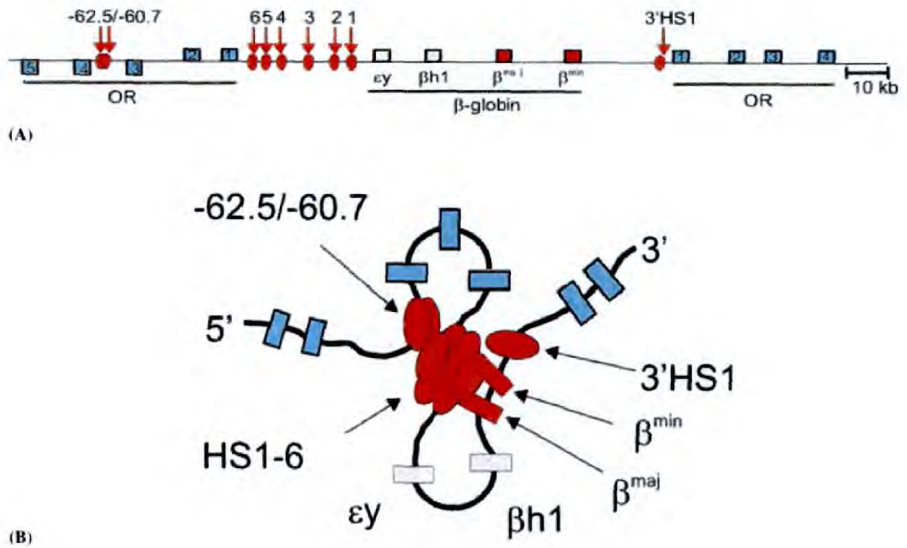


Figure 1.7.3 Long-range interactions in the mouse *β-globin* (from Tolhuis et al., 2002). (A) Schematic of the 200 kb mouse *β-globin* locus LCR containing several DNase I HS spread over a 15 kb region and others located 40-60 kb away from the genes in linear sequence. The grey boxes indicate the inactive genes (*εy* and *βh1*). In the 14.5dpc foetal liver the active globin genes (*βmaj* and *βmin*) are shown as red boxes. (B) Clustering of active genes and hypersensitive sites in the mouse hub, the active chromatin hub (ACH). Red loops indicate the active regions (hypersensitive sites and active genes) of the locus the inactive regions of the locus loop out, with the inactive *βh1* and *εy* globin genes on one loop, and the 5' OR 1-3 genes on another loop.

1.8 Chromatin remodelling

Nuclear processes such us transcription, replication, recombination and repair are regulated and initiated by DNA binding factors. Interaction of these factors with their target elements requires that that the tight and repressive chromatin organization is partially unravelled. Chromatin needs to accommodate transitions between active and inactive states. This dynamic, interconvertible structure has been achieved through two major general mechanisms. These are modification of the N terminal histone-tails through post translational modifications (discussed in detail in section 1.9) and the action of nucleosome remodelling factors which are able to alter histone-DNA interaction so that nucleosomal DNA becomes more accessible to regulatory proteins. These perturbations of nucleosomes may lead to the relocation of histone octamers from a particular DNA fragment to available acceptor DNA in cis- or trans- establishing a "fluid" state of chromatin in which the overall packaging of DNA is maintained, but individual sequences are transiently exposed to interacting factors (Kingston and Narlikar, 1999). Histone and DNA in nucleosomes are held together by a large number of weak interactions and therefore nucleosome remodelling requires an energy input, which is given by ATP hydrolysis. ATP dependent chromatin remodelling enzymes belong to the SNF2 family of DNA dependent ATPase, which all have a helicase-like ATP domain (Bork and Koonin, 1993; Henikoff, 1993). They can be divided into three classes based on the identity of their catalytic ATPase subunit. These are SWI/SNF, ISWI and Mi-2/CHD (Eisen et al., 1995). All of these complexes contain a highly conserved ATPase core domain flanked by N- and C- terminal domains that differ considerably between the three families.

The SWI/SNF family complexes are generally, but not exclusively, implicated in the activation of transcription. They comprise yeast Snf2 and Sth2, Drosophila melanogaster brahma (BRM) and mammalian BRM and brahama-like1 (BRG1). These proteins are characterised by the presence of a bromo-domain, which binds acetylated histones (Hassan et al., 2002). By contrast, ISWI family enzymes appear to have a role in transcriptional repression pathways and nucleosome assembly. ISWI proteins contain a SANT domain that is thought to act as a histone-binding domain (Boyer et al., 2004) and may recognise specifically modified histones. There

are two ISWI homologues in yeast (Isw1 and Isw2) and mammals (SNF2H and SNF2L). The third family, Mi-2/CHD complexes are involved in transcriptional repression and form complexes with HDACs. The Mi-2 family is characterised by the presence of two amino-terminal chromo-domain, which can interact with methylated histone tails (Bannister and Kouzarides, 2004; Lachner et al., 2001).

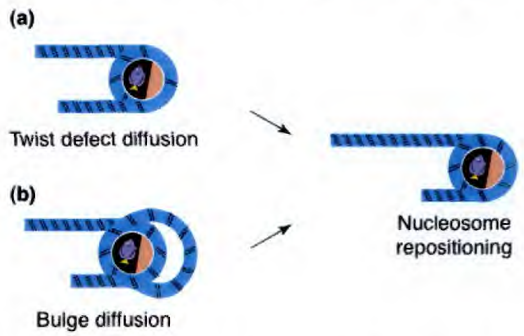


Figure 1.8.1 Models of actions of ATP-dependent chromatin remodellers (adapted from Flaus and Owen-Hughes, 2004). (A) Small local alterations from mean DNA twist propagate around the nucleosome. (B) The DNA is unpeeled from the histone octamer at the entry and exit point of the nucleosome. Rebinding of more distal sequences to the same contact points would cause the looping out of DNA to form a 'bulge'. As the bulge moves around the nucleosome, the nucleosome would appear to move in the direction of the bulge. In effect the nucleosome appears to make a sliding movement. Both the twist and bulge diffusion would result in nucleosome repositioning and exposure of previously occluded DNA sequence.

Two potential mechanisms by which the remodelling complexes alter the accessibility of DNA through nucleosome mobilisation are 'twist defect diffusion' and 'bulge diffusion' (Fig. 1.8.1). The first mechanism involves the repositioning of intact nucleosomes on DNA by small, local alterations that propagate around the nucleosome with the effect of 'sliding' the nucleosome along the DNA. In the second mechanism, DNA is 'unpeeled' from the histone octamer by changes in histone-DNA or histone-histone contacts, to form loops 'bulging out'. Movement of the bulge around the nucleosome would result in nucleosome repositioning. Bulge diffusion could also expose previously occluded DNA sequences (Flaus and Owen-Hughes, 2003; Flaus and Owen-Hughes, 2004).

Nucleosomes have been shown to prevent binding of TBP to the TATA element in vitro in the absence of an activator (Kuras and Struhl, 1999) but they only modestly inhibit binding of a variety of activator proteins to their target sites. Activators and repressors regulated by cell cycle and developmental signals can also recruit chromatin-modifying activities to promoters. The relationship between transcriptional regulation and nucleosomal organisation has been studied in depth

both in the mouse mammary tumour virus MMTV (Richard-Foy et al., 1987) and in the yeast HO promoters

1.8.1 The yeast HO promoter

The yeast *HO* gene is expressed in late G1 and encodes an endonuclease involved in mating-type switching. Correct regulation of *HO* transcription requires the transcription factor Swi5p (Stern et al., 1984), which is expressed early in cell cycle and has largely or completely disappeared in late G1. ChIP experiments by Cosma et al (1999) showed that at the end of anaphase, Swi5 binds to the *HO* promoter. Swi5 recruits the remodelling complex SWI/SNF, which facilitates the binding and action of the HAT complex SAGA. It is possible to speculate that SWI/SNF might either directly recruit SAGA or remodel the surrounding nucleosomes to make them better substrates for acetylation. Swi5 remains bound for only a short time while SWI/SNF and SAGA persist and facilitate the binding of the heterodimeric activator Swi4/Swi6 that finally triggers transcription (Cosma et al., 1999). It has also been shown that prior to activation, a domain of histone H3 and H4 acetylation of approximately 1 kilobase is established in mid-G1, prior to and independent of HO transcription. The presence of acetylated histones over the promoter implies that nucleosomes are not removed by the remodelling action of SWI/SNF. Nevertheless the establishment of this acetylation domain has been shown to require the presence of SWI/SNF, consistent with its prior recruitment to the promoter (Krebs et al., 1999).

1.8.2 The mouse MMTV promoter

At the mouse mammary tumor virus (*MMTV*) promoter, six precisely positioned nucleosomes play a key role in the regulation of transcription. Regulatory sequences for the *MMTV* promoter extend to 1.3 kb upstream of the gene and include a long terminal repeat (LTR) in which the six nucleosomes are positioned (designated nucleosomes A-F). Within the LTR, several cis-acting regulatory sites have been characterised. These include a negative regulatory element, a positive enhancer containing transcription factor binding sites that are required for tissue-specific expression, and hormone response element (HRE) binding sites for steroid receptors. The gene is silent in the absence of hormones but is rapidly activated in the presence of glucocorticoids or progestins (Smith et al., 1993). Upon activation the region

containing nucleosome B becomes hypersensitive to DNase I at the site where the HRE is located, implying a nucleosome core- remodelling event (Archer et al., 1992). The remodelled chromatin allows for the occupancy of all factor binding sites. This series of events results in to the formation of the pre-initiation complex and activation of the *MMTV* promoter leading to activation of the gene (Lee and Archer, 1994). There is evidence that the human SWI/SNF complex acts in concert with the estrogen receptor to remodel the nucleosomes over the promoter and activate transcription (Fryer and Archer, 1998; Trotter and Archer, 2004).

1.9 Epigenetic regulation

In 1942 Conrad Waddington defined epigenetics as 'the branch of biology that studies the causal interactions between genes and their products which bring the phenotype into being'. The term epigenetics is now used to describe heritable changes in genome function that occur without a change in nucleotide sequence within the DNA. Epigenetics is linked to cell identity and to how this identity is transmitted to daughter cells through the processes of DNA replication, chromatin assembly, and DNA re-packaging that accompany cell division (Turner, 2002). The amino terminal tails of the four histones have long been considered to be the main repository of this information. Histone tails are exposed on the nucleosome surface and subject to a variety of enzyme-catalyzed post-translational modifications such as lysine acetylation, lysine and arginine methylation, serine phosphorylation, ubiquitination and ADP ribosylation (Turner, 2000) (Fig. 1.9.1). Those modifications, together with DNA methylation, are considered epigenetic marks of gene expression that can be stably transmitted to progeny cells through many cell divisions. Recently it has been shown that the globular core domain of the histone proteins can be subject to covalent modifications as well (Mersfelder and Parthun, 2006). While modifications on the histone tails are thought to act primarily by altering the ability of non-histone proteins to interact with chromatin, it is likely that some core histone modifications will also affect chromatin structure. The locations and the evolutionary conservation of these modifications support this idea (Mersfelder and Parthun, 2006). Mapping of their positions onto the nucleosome crystal structure, has shown that they fall into clusters that can be organised into three different classes: those at the histone-histone interfaces, those at the histone

lateral surface and those at the solute accessible face (Cosgrove et al., 2004; Freitas et al., 2004).

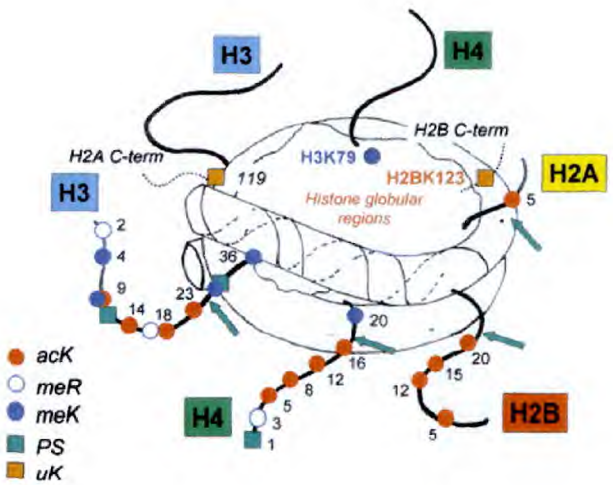


Fig. 1.9.1 Histone modifications identified in vivo (from Turner BM, 2002). Legend key. acK : acetylable lysine ; meR : methylable arginine ; meK : methylable lysine ; PS : phosphorylable serine ; uK : ubiquitinable lysine.

The idea that modifications on histone tails could be recognised by specific binding proteins, which in turn can influence gene expression was put forward a number of years ago (Turner, 1993; Turner et al., 1992) and this is now being substantiated by experimental data. Bromodomains, chromodomains and SANT domains are able to interact with chromatin and/or with its modified components and play a crucial role in the process of targeting specific histone modifying enzymes, to their substrates. What still remains controversial is the existence of a histone code constituted by specific histone tail modifications, or their combination, which could play a role not only in putting in place transcription patterns but also in determining their heritability (Jenuwein and Allis, 2001; Spotswood and Turner, 2002). Cells are obviously able to replicate while maintaining their identity and one of the question still to be answered is if cellular memory could actually operate through patterns of histone modifications. The implications of an epigenetic code determining cellular memory are tremendous. Changes to histone modifications, due to metabolic or environmental influences, would change not only the cells subjected to them but their progeny as well. Obviously a heritable epigenetic code should have a predictive value, which means it should be able to provide an indication that a certain function will occur when the specific set of modifications is in place. It has

been proposed a distinction to be made between a histone code, which refers to the combination of modifications known to be involved in ongoing cellular processes, and an epigenetic code, which instead refers to the heritable code that could be involved in cellular memory (reviewed by Nightingale, 2006).

1.10 Histone methylation

Histone methylation occurs on both arginine and lysine residues. While histone arginine methylation generally correlates with transcriptional activation, histone lysine methylation can signal either activation or repression, depending on the particular lysine residue that is methylated (Kouzarides, 2002; Zhang and Reinberg, 2001). Even within the same lysine residue, the biological consequence of methylation can differ depending on the methylation state (Santos-Rosa et al., 2002; Wang et al., 2003).

At least five methylatable lysine positions exist in the N termini of histone H3 (K4,K9,K27,K36) and H4 (K20); another occurs in the histone fold domain of histone H3 (K79). The position of the methylated lysine within histone H3 and H4 marks a gene for either activation or repression. Moreover, a lysine residue has the ability to be modified by mono-, di- or tri- methylation and histone methyltransferase (HMTases) display remarkable specificity for the level of methylation they catalyse (Zhang and Reinberg, 2001). In mice it has been shown that that pericentric heterochromatin is specifically enriched in histone H3 lysine 9 trimethylation (H3K9me3) and histone H4 lysine 20 trimethylation (H4K20me3). By contrast mono and dimethyl histone H3 lysine 9 and histone H4 lysine 20 are found in euchromatin (Peters et al., 2003; Rice et al., 2003; Schotta et al., 2004). It has been suggested that the differential methylation status at specific lysines recruits specific methyl binding proteins that are dedicated to unique functions. For example, three different isoforms of the heterochromatic protein 1 (HP1) are linked to different degrees of H3K9 methylation. HP1 α and β are localised to pericentric heterochromatin while the HP1γ isoform is localised to euchromatin (reviewed by Eissenberg and Elgin, 2000).

1.10.1 Heterochromatin formation: lysine 9 tri-methylation

Heterochromatin formation has been connected to methylation of lysine 9 thanks to studies of PEV in Drosophila. The identification of Su(var)39 as a suppressor of variegation established the importance of this gene, which encodes a HMTase, in chromatin modification (Tschiersch et al., 1994). The mammalian homologs SUV39H1 and Suv39h1 were later discovered to have specificity for histone H3 lysine 9 (Rea et al., 2000). It has been shown in mice that Suv39h1/Suv39h2 specifically catalyse trimetylation of lysine 9 (Peters et al., 2003; Rice et al., 2003). *Suv39h1/ Suv39h2* null cells showed a significant decrease in H3K9me3 within major satellite repeats compare to wild type cells (Peters et al., 2003). This decrease was concomitant with an increase in H3K9 mono-methylation. These finding suggest that H3 lysine 9 must be mono-methylated prior to conversion to the tri-methylated form by Suv39h1/h2. *In vitro* analysis demonstrated that mono-methyl H3 lysine 9 is the preferred substrate for both Suv39h1 and Suv39h2 (Rice et al., 2003).

The catalytic activity of SUV39H1 was shown to be dependent on the evolutionarily conserved SET domain, a protein domain that is present in many known chromatin modifier. The SET domain is found in more than 300 proteins in a wide range of species and was initially defined by homology between the suppressor of PEV, Su(var)3-9 (Tschiersch et al., 1994), the polycomb group (Pc-G) protein enhancer of zeste (E[z]) (Jones and Gelbart, 1993) and the activating thritorax group (trx-G) protein Thritorax (Jenuwein et al., 1998). Other prominent members include the *S. cerevisiae* Set1, which affects mating-type switching and telomeric silencing (Nislow et al., 1997), the *S. pombe* Clr4 which is involved in centromere function (Ekwall et al., 1996), and the human trx homologue HRX (Djabali et al., 1992; Gu et al., 1992; Tkachuk et al., 1992) which is implicated in the onset of translocation-induced leukaemia.

Metylation of histone H3 lysine 9 by SUV39H1 provides a binding site for the chromodomain of heterochromatin protein 1 (HP1). These mechanisms are highly conserved and in yeast Clr4 recruits Swi6 (Swi6 is the fission yeast homologues of HP1) (Bannister et al., 2001; Lachner et al., 2001). HP1 is a small protein with two conserved domains, the N-terminal chromodomain and the C-terminal

chromoshadow domain, which are critical for its function in heterochromatin. HP1 interacts stably with SUV39H1 via its chromoshadow domain and with the H3K9 methyl mark via its chromodomain in flies and mammals (Aagaard et al., 1999; Nakayama et al., 2001).

Studies in fission yeast have revealed a role for the RNA interference pathway in heterochromatin formation. The current model proposes that small interfering RNA (siRNA) molecules (produced by the RNA-directed RNA polymerase complex (RDRC) target the RNA-induced transcriptional silencing (RITS) complex to RNA transcripts that are associated with heterochromatic loci. This would lead to the recruitment of Clr4 and the methylation of H3-K9 (Hall et al., 2002; Motamedi et al., 2004; Verdel et al., 2004; Volpe et al., 2002). This model requires further clarification since one of the components of RITS, Chp1, contains a chromodomain that interacts specifically with methylated lysine 9 on histone H3. Also Clr4, and presumably histone H3K9 methylation, are required for the association of Chp1 and RITS with heterochromatin. If the RNAi machinery directs heterochromatin formation to repeat–rich regions, one would not expect it to be dependent upon the histone methyl transferase. The self-reinforcing loop model, were Clr4 is recruited by RITS/RDRC and methylation of H3K9 by Clr4 helps recruit RITS and RDRC might explain this contradiction (Noma et al., 2004; Yamada et al., 2005). In mice, two Clr4 homologues, SUV39h1 and Suv39h2 are responsible for H3K9 trimethylation at heterochromatic regions such us pericentric heterochromatin (Lachner et al., 2001; Peters et al., 2003; Rice et al., 2003). This modification recruits HP1α and HP1β, homologues of Swi6 (Bannister et al., 2001; Lachner et al., 2001). The contribution of the RNAi pathway to mammalian heterochromatin formation has not yet been established.

1.10.2 Heterochromatin formation: histone H4 lysine 20

Trimethylation of lysine 20 on histone H4 is also associated with heterochromatin formation. SUV4-20h1 and SUV4-20h2, two set domain-containing proteins, are responsible for this modification. It has been shown that H4-K20 methylation is abolished at pericentric regions in the absence of SUV39h1/SUV39h2 function (Schotta et al., 2004). Schotta et al proposed the following mechanism for heterochromatin formation in mammals. In the first step, SUV39h enzymes would be targeted to repeat-rich sequences at pericentric heterochromatin. Because

pericentric heterochromatin is also enriched for H3-K27 monomethylation, and the Suv39h enzymes prefer an H3-K9 monomethylated substrate (Peters et al., 2003), they suggested that the first step could require the activity of currently unknown H3-K27 and H3-K9 monomethylases. HP1 would then bind to H3-K9 trimethylated nucleosomes, and recruit the Suv4-20h enzymes, which in turn could trimethylate H4-K20. This model would involve collaboration of at least four distinct HMTases to induce the observed combination of H3-K9 trimethyl, H4-K20 trimethyl, and H3-K27 monomethyl marks at pericentric heterochromatin. This pathway is not conserved in fission yeast and might therefore be restricted to metazoans.

1.10.3 Transcriptional silencing: histone H3 lysine 9 mono- and dimethylation and histone H3 lysine 27 methylation

Histone H3 lysine 9 mono and di-methylation have been associated with silencing of euchromatic loci in mammals (Peters et al., 2003; Rice et al., 2003) but, histone H3 lysine 9 methylation, has been found completely absent from the "activated genome" of *S.cerevisiae* (Briggs et al., 2001). Mono and di-methylation at H3 lysine 9 in euchromatin are catalysed by G9a and G9a-related protein (GLP). Both form complexes with HP1γ and a subset of E2F transcription factors (Ogawa et al., 2002). G9a and GLP have also been found to be responsible for the vast majority of H3K9me2 and most H3K9me in mouse embryonic stem cells. Lack of *G9a* resulted in a global increase in histone modifications associated with transcription as well as in a dramatic phenotype of early embryonic lethality due to severe differentiation defects (Tachibana et al., 2002). These results imply that H3K9me and H3K9me2 play an important role in the establishment and maintenance of developmentally regulated gene silencing. It has been shown that H3K9me induced by the SET domain bifurcated 1 (SETDB1) HKMT can induce gene silencing in euchromatic regions, which involves recruitment of HP1 at a reporter gene promoter used (Ayyanathan et al., 2003). Spreading of HP1, which could contribute to heterochromatin formation, from the promoter region was not observed. These findings suggest that different processes mediate the downstream events that silence genes in heterochromatin and euchromatin even though both regions are marked by H3K9me and HP1 recruitment.

Histone H3 lysine 27 has been linked to several silencing phenomena including homeotic-gene silencing, X inactivation and genomic imprinting (Cao and Zhang,

2004). The Pc-G protein enhancer of zeste [E(z)] contains a SET domain and becomes a HMTase when complexed with another early-acting Pc-G protein, extra sex combs (Esc). The Drosophila E(z)-Esc complex (Czermin et al., 2002; Muller et al., 2002) and its mammalian Ezh-Eed counterpart (Cao et al., 2002; Kuzmichev et al., 2002) have an apparent preference for H3-K27 but might also target H3-K9. Ezh/Eed-mediated nucleosome methylation increases *in vitro* binding to histone H3 of the chromodomain protein polycomb (PC), which is a component of the polycomb repressive complex-1 (PRC1)(Czermin et al., 2002; Kuzmichev et al., 2002). In E(z) mutants, methylation of H3-K27, and probably also H3-K9, is impaired in a manner suggesting that extended H3-K27 di- and tri-methylation across several nucleosomes (Cao et al., 2002) or dual tri-methylation of H3-K27 and H3-K9 (Czermin et al., 2002) might induce stable recruitment of Pc-G complexes.

PcG proteins have shown to be involved in female X chromosome inactivation [reviewed by (Avner and Heard, 2001; Heard, 2004)]. H3-K9 methylation is associated with the inactive X chromosome (Xi) but H3-K27 tri-methylation might also be a prominent, if not the major, mark (Boggs et al., 2002; Heard et al., 2001; Peters et al., 2002). It has been observed that components of the mouse EZH2 and PRC1 complexes are transiently recruited to the inactive X chromosome during the initiation stage of both imprinted and random X inactivation. Recruitment of these proteins coincides with the appearance of methyl H3K27 (Plath et al., 2003; Silva et al., 2003). These findings are consistent with X-inactivation being independent of the Suv39h HMTases and of the HP1 proteins (Peters et al., 2002). However, there is no evidence for stable association of PC or other Pc-G complexes at the Xi (Silva et al., 2003).

1.10.4 Histone metylation and transcriptional activation

In addition to transcriptional repression, histone lysine methylation has been also associated with transcriptionally competent euchromatin. Sites of methylation in transcriptionally active regions are lysine 4 (H3K4), lysine 36 (H3K36) and lysine 79 (H3K79) on histone H3. In the case of H3K4 and H3K36, studies in yeast have identified two separate histone lysine methyl-transferase (HKMT) containing complexes that functionally associate with RNA pol II and its machinery during elongation resulting in histone methylation in the coding regions (Krogan et al., 2003; Li et al., 2002; Ng et al., 2003).

It has been shown that methylation of H3K4 impairs Suv39h1-mediated methylation of H3K9 and preclude binding of the histone deacetylase NuRD repression complex (Nishioka et al., 2002; Zegerman et al., 2002), which suggests a function for H3K4 in preventing heterochromatin spreading. In *S.cerevisiae* the HKMT Set1 is responsible for all the three methylated states of H3K4. Analysis of H3K4 methylation in different organisms indicate that both H3K4me2 and H3K4me3 are enriched at actively transcribed genes. While H3K4me2 appear to be generally distributed across the body of the active genes, chromatin immunoprecipitation experiments have shown that H3K4me3 is localised specifically at their 5' ends (Ng et al., 2003; Santos-Rosa et al., 2002). In fact H3K4me3 is linked to the association of Set1 with the phoshorylated form of elongating RNA pol II that is confined to the 5' region of genes. It has been proposed that association of Set1 with the phoshorylated RNApol II causes a conformational change into the catalytic site of the protein, which is then converted into a trimethylase (Ng et al., 2003). Another possibility is that Set1 becomes a trimethylase as a result of post translational modifications when it is associated with RNA pol II (Laribee et al., 2005; Schlichter and Cairns, 2005).

Lysine 4 methylation has also been linked with chromatin remodelling activity. Genome wide analysis by Chip-on-chip showed that the yeast Isw1, part of the Isw1 ATP dependent chromatin remodelling complex, is recruited to a specific set of genes in a manner that is dependent of H3K4 methylation. In particular, Isw1p is associated with chromatin only in presence of H3K4me2 or H3K4me3. The binding is not direct but probably mediated by an intermediary factor(s) (Santos-Rosa et al., 2003).

Recently two proteins have been shown to interact specifically with methylated H3K4, CHD1 and WDR5 (Flanagan et al., 2005; Sims et al., 2005; Wysocka et al., 2005). CHD1 (chromo-ATPase/helicase-DNA-binding protein1) is a member of the SNF2-like family of proteins, which has two chromodomains (Woodage et al., 1997), thus linking covalent histone modifications with chromatin remodeling. The WD40-repeat protein WDR5A, which is a common component of MLL1, MLL2, and hSet1 H3 K4 methyltransferase complexes (Dou et al., 2005), directly associates with histone H3 di- and trimethylated at K4. WDR5 is required for binding of the methyltransferase complex to the K4 dimethylated H3 tail as well as for global H3 K4 trimethylation (Wysocka et al., 2005).

It has also been shown that the largest subunit of NURF, BPTF (bromodomain and PHD finger transcription factor), mediates a direct preferential association with H3K4me3 tails via its PHD2 domain. Depletion of H3K4me3 has been shown to cause partial release of the NURF subunit from chromatin and defective recruitment of the associated ATPase, SNF2L (also known as ISWI), to the HOXC8 promoter (Wysocka et al., 2006). These findings confirmed methylation of H3 K4 as a binding platform for proteins involved in gene transcription.

1.10.5 Histone H3 lysine 36 and 79

In *S.cervisiae* Set2 is responsible for K36 methylation. Set2 is physically associated with the elongating form of RNApol II (Li et al., 2003; Xiao et al., 2003). The association in maintained throughout the body of transcribed genes. In yeast, the Dot1 enzyme is responsible for K79 methylation. Dot1 and its homologues are the only methyltransferases that lack a SET domain (Feng et al., 2002b). Dot1 has been shown to prevent the spreading of histone deacetylases (Lacoste et al., 2002; van Leeuwen et al., 2002). Recent studies provide evidence that H3K79 methylation might represent a binding site for the mammalian protein 53BP1, a conserved protein with properties of a DNA double-strand breaks (DSBs) sensor (Huyen et al., 2004). Huyen and colleagues showed that the 53BP1 domain that recruits the protein to sites of DSBs consists of two tandem tudor folds with a deep pocket at their interface. In vitro, the 53BP1 tandem tudor domain binds histone H3 methylated on Lys 79 using residues that form the walls of the pocket; these residues are also required for recruitment of 53BP1 to DSBs. Suppression of Dot1, the enzyme that methylates Lys 79 of histone H3, also inhibited recruitment of 53BP1 to DSBs. These findings indicate a role for H3K79 in DNA repair but there is no evidence of a link between 53BP1 and gene transcription. Another as yet unidentified protein, which would bind to methylated H3K79 might be involved in linking Dot1, H3K79 methylation and transcriptional regulation (Huyen et al., 2004).

1.11 Demethylation of histones

Unlike other histone modifications such as acetylation, methylation has long been considered a "permanent" modification. This view was based on the early observation that the half-lives of histones and total histone methyl groups were comparable, which led to the conclusion that histone methylation is stable and irreversible (Byvoet et al., 1972). Turnover of methylation marks has long been considered to be by dilution during replication, by exchange on chromatin, or by proteolytic cleavage of the protein portion carrying the methyl group (Ehrenhofer-Murray, 2004). Additionally, identification of the histone H3.3 variant as participating in replication-independent nucleosome assembley targeted to transcriptionally active loci, provided evidence of a mechanism for turnover of methylated histone H3 and gene activation (Henikoff and Ahmad, 2005). The view of histone methylation as enzymatically irreversible has recently been abandoned after the discovery of enzymes that can demethylate arginine as well as lysine residues. These include LSD1, which acts on H3K4me and H3K4me2 (Shi et al., 2004); JHDM1, which specifically demethylates H3K36me2 (Tsukada et al., 2006) and lately JMJD2, which demethylates H3K9me3 and H3K936me3 (Whetstine et al., 2006) .

The first example of an enzyme that was capable of turning over a methylated residue was the human peptidylarginine deiminase 4 (PAD4/PADI4). PAD4/PADI4 converts monomethyl-arginine in histone H3 and H4 to citrulline by demethylimination (Cuthbert et al., 2004; Wang et al., 2004). However, since PAD4/PADI4 catalyzed demethylimination does not generate arginine and given that PAD4/PADI4 also acts on unmethylated arginine to generate citrulline, PAD4/PADI4 is not a strict demethylase. A recent study has uncovered a true histone lysine demethylase. Shi and colleagues demonstrated that LSD1/ BHC110 is able to specifically demethylate mono- and dimethyl H3K4 in a FAD (flavin adenine dinucleotide)-dependent oxidative reaction (Shi et al., 2004). LSD1 was previously identified in several histone deacetylase complexes (Hakimi et al., 2002; Shi et al., 2003; You et al., 2001). When part of a complex with the androgen receptor, LSD1 has also been shown to demethylate H3-K9Me2, thereby expanding its substrate repertoire (Metzger et al., 2005).

A second histone demethylase, JHDM1, has also been identified (Tsukada et al., 2006). JHDM1 specifically demethylates H3-K36Me2 via an oxidative reaction catalyzed by the evolutionarily conserved JmjC domain. Both LSD1 and JHDM1 demethylate only di- or monomethylated histones. The most recent discovery is the one of the JMJD2 subfamily of histone lysine trimethyl demethylases. The JMJD2 family is composed of four members (JMJD2A, JMJD2B, JMJD2C, and JMJD2D) (Katoh and Katoh, 2004). JMJD2A interacts with the tumour suppressor Rb, with histone deacetylases (HDACs), and with the corepressor N-CoR and functions as a repressor in part through the recruitment of HDACs and N-CoR, respectively (Gray et al., 2005; Yoon et al., 2003a; Zhang et al., 2005). However, molecular mechanisms by which JMJD2A represses transcription are incompletely understood. JMJD2C is also known as GASC-1 (gene amplified in squamous cell carcinoma) and has been implicated in tumorigenesis (Yang et al., 2000). Cloos et al have shown that three members of the JMJD2C subfamily of proteins demethylate H3K9me3/me2 in vitro through a hydroxylation reaction requiring iron and α-ketoglutarate as co factors. Furthermore, they have demonstrated that ectopic expression of GASC-1 or other JMJD2 members markedly decreases H3K9me3/me2 levels, increases H3K9me1 levels, delocalizes HP1 and reduces heterochromatin in vivo (Cloos et al., 2006).

JMJD2A has been shown to act as a lysine trimethyl-specific histone demethylase, which catalyzes demethylation of H3K9Me3 and H3K36Me3 (Whetstine et al., 2006, Klose et al, 2006). In particular Klose et al showed that JHDM3A (JMJD2A) has a further demethylase activity directed towards lower modification states such as H3K9/H3K36me2/me1 and that overexpression of JHDM3A abrogates recruitment of HP1 (heterochromatin protein 1) to heterochromatin. This indicates a role for JHDM3A in antagonizing methylated H3K9 nucleated events (Klose et al., 2006). SiRNA-mediated knockdown of JHDM3A also showed to lead to increased levels of H3K9 methylation and upregulation of a JHDM3A target gene, ASCL2, indicating that JHDM3A may function in euchromatin to remove histone methylation marks that are associated with active transcription (Klose et al., 2006).

1.12 Histone acetylation

Acetylation and deacetylation of histone lysine residues in the N termini of histones play a fundamental role in many chromatin-based processes. Histone acetyltransferases (HATs) and deacetylases (HDACs) have shown to act as coregulators in transcription, DNA replication and DNA repair. Sites of acetylation include at least four highly conserved lysines in histone H4 (K5,K8, K12 and K16) and five in histone H3 (K9, K14, K18, K23 and K27) as well as less conserved sites in histone H2A and H2B. HATs can be organized into families, which display different mechanisms of histone substrate binding and catalysis (Table 1.12.1). The Gcn5/PCAF family of HAT proteins (GNATs) function as co-activators for a subset of transcriptional activators. They contain both a HAT domain and a bromodomain (de la Cruz et al., 2005).

Characteristics of HAT families

HAT group	HAT (and its associated complexes)	Histones acetylated by HAT complex	Interactions with other HATs
1. GNAT	Gcn5 (SAGA, ADA, A2)	H3, H2B	p300; CBP
	PCAF (PCAF)	H3, H4	p300; CBP
	Hat1 (HatB)	H4, H2A	
	Elp3 (elongator)	H2A, H2B, H3, H4	
	Hpa2	H3>H4	
2. MYST	esa1 (nuA4)	H2A, H4	
	MOF (MSL)	H4	
	Sas2	unknown	
	Sas3 (nuA3)	H3	
	MORF	H4>H3	
	Tip60	H4>H3, H2A	
	Hbo1 (ORC)	H3, H4	
3. p300/CBP	p300	H2A, H2B, H3, H4	PCAF; Gcn5
	CBP	H2A, H2B, H3, H4	PCAF; Gcn5
4. Basal transcriptior	TAFII250 (TFIID)	H3>>H2A	
factors	TFIIIC	H3, H4>H2A	
	Nut1 (mediator)	H3>>H4	
5. Nuclear receptor	ACTR*	H3>H4	p300: CBP; PCAF
cofactors	SRC1	H3>H4	p300: CBP; PCAF

* also known as RAC3, AIB1, PCIP, TRAM
histones acetylated by recombinant HAT

Table 1.12.1 Characteristics of HAT families (adapted from Roth et al 2001)

The p300/CBP family is another major group of nuclear HATs. The members of this family are more global regulators of transcription and contain a very large HAT

domain, a bromodomain and three cysteine-histidine rich domains (TAZ, PHD, ZZ) that are believed to mediate protein-protein interaction. The MYST family of HAT proteins contain a highly conserved 370 residue MYST domain, which uses a different catalytic mechanism from that shared by the other HATs (Yang, 2004). There are two main mechanisms by which histone acetylation and deacetylation are thought to play a role in transcriptional activation. Firstly acetylation of lysine residues could directly inhibit chromatin condensation by altering the normal core histone tail domain function. In vitro studies in fact have shown that above a certain threshold of acetyl groups per nucleosome condensation of chromatin is blocked (Tse et al., 1998). Secondly, covalently modified termini have shown to serve as a signal for the binding of trans-acting factors. For example, some activating factors, including several HATs, possess a region termed the bromodomain that interacts specifically with acetylated core histones (Dhalluin et al., 1999; Hassan et al., 2002).

Bromodomains are found in many eukaryotic transcription factors, such as the HAT Gcn5 and Taf1 (TFII250), which is the largest subunit of the TFIID transcription initiation complex, as well as in proteins in chromatin-remodeling complexes, such as Snf2 of the yeast SWI/SNF complex.

The presence of sites of localized histone acetylation (Litt et al., 2001) supports the idea that histone acetylation facilitates protein–DNA interactions within chromatin in general. Many cases in which activating transcription factors recruit HAT-containing co-activators to specific promoters have now been documented (Brown et al., 2000).

The mammalian HAT GCN5/PCAF and the homologous yeast SAGA complex, together with the CBP/p300 proteins, function as co-activators of transcription after their recruitment to promoters via their interaction with numerous DNA binding proteins. Histone acetyltransferase activity of these complexes is critical for activation of transcription (Berger, 1999; Struhl, 1999). In the case of the yeast HO promoter, activator-dependent recruitment of SWI/SNF is required for subsequent recruitment of SAGA followed by histone acetylation and DNA binding of an additional activator, SBF (Cosma et al., 1999; Krebs et al., 1999). Studies on PHO5 promoter instead showed that is SAGA-dependent histone acetylation that facilitates SWI/SNF induced chromatin remodelling and transcriptional activation (Syntichaki et al., 2000).

More recently Hassan et al. investigated the binding of the SWI/SNF chromatin remodeling complex and the SAGA and NuA4 HAT complexes to chromatin in vitro. The SWI/SNF complex was recruited to chromatin by the activator Gal4-VP16. The authors have shown that prior acetylation by SAGA or NuA4 complexes does not enhance SWI/SNF binding but does allow retention of the complex when the activator Gal4-VP16 is removed (Hassan et al., 2001). Retention of SWI/SNF requires the bromodomain of the Swi2/Snf2 subunit of the complex. They also showed that the SAGA complex itself is anchored to acetylated chromatin following removal of VP16, but only if the bromodomain of the Gcn5 subunit is intact. In contrast, the NuA4 complex, which lacks a bromodomain, is not retained following removal of Gal4-VP16 (Hassan et al., 2002).

Studies on the *IFN-β* promoter have also demonstrated the existence of a link between histone acetylation and chromatin remodeling (Agalioti et al., 2000). Agalioti and co-workers have shown that activation of the *IFN-β* gene in response to viral infection is accompanied by the progressive acetylation of H4K8, H3K9 and K14. *In vitro* experiments have shown that acetylation at K8 is necessary for recruitment of the SWI/SNF complex and that H3K9 acetylation and H3K14 acetylation are necessary for recruitment of TFIID, probably through binding of the double bromodomain of the TAFII250 subunit. Acetylation and transcription complex assembly required ordered recruitment of the GCN5 and CBP/PolII holoenzyme complexes followed by recruitment of SWI/SNF. SWI/SNF recruitment and subsequent chromatin remodelling was found to be greatly enhanced by histone acetylation, indicating that this modification is a prerequisite for the remodeling function of the SWI/SNF complex. These observations suggested a model in which transcription initiation involves sequential modifications at defined sites on the histone tails, each of which facilitates recruitment/binding of the next element.

Broad domains of acetylation have also been observed. In yeast the ground state of chromatin is characterized by intermediate levels of H3 and H4 acetylation, a state caused by a mix of untargeted HAT and HDAC activities (Vogelauer et al., 2000). Site-specific acetylation or deacetylation is believed to lead to locally restricted activation or repression of transcription, respectively. However, this overall level of flexibility seems to be a specific feature of the highly active yeast genome. In differentiated, higher eukaryotic cells, most of the genome consists of hypoacetylated, inactive chromatin, which may be considered the 'ground state'.

Activation of house-keeping and cell-type-specific genes may involve the acetylation of histones across broad chromatin domains. An example is provided by the *β-globin* loci, which exhibit a pattern of broad acetylation that correlates with transcriptional competence (Litt et al., 2001; Schubeler et al., 2001). Broadly distributed acetylation is also observed in the dosage compensated male X chromosome in Drosophila (Gu et al., 1998; Lucchesi, 1998), where the acetylation of histone H4 at lysine 16 (H4K16) by MOF (Males absent on the First) correlates with increased transcription of many genes throughout most of the male X chromosome (Akhtar and Becker, 2000; Smith et al., 2001). Broad acetylation of histone H3/H4 has shown to lead to partial decondensation and is not tightly correlated with active transcription per se, but rather marks regions of transcriptional competence (Verdone et al., 2005).

1.13 Histone deacetylases (HDACs)

Mammalian deacetylases can be grouped into four classes. Class I deacetylases (HDACs 1, 2, 3 and 8) share homology in their catalytic sites and are homologous to yeast Rpd3. Class II deacetylases include HDACs 4, 5, 6, 7, 9 and 10. Of these HDACs 4, 5, 7 and 9 share homology in two regions: the C-terminal catalytic domain and the N-terminal regulatory domain. These HDACs are homologous to yeast Hda1. HDAC6 and 10 have two regions of homology with the class II catalytic site (de Ruijter et al., 2003; Grozinger et al., 2001). The third class of deacetylases is the conserved nicotinamide adenine dinucleotide-dependent Sir2, or sirtuins, family of deacetylases (Grozinger et al., 2001). Class I and II HDACs are inhibited by trichostatin A (TSA), suberoylanilide hydroxamic acid (SAHA) and related compounds; class III HDACs are not inhibited by these compounds. The fourth class are proteins related to the recently identified human HDAC11, which contains conserved residues in the catalytic core regions shared by both class I and II mammalian enzymes (Gregoretti et al., 2004). The HDAC enzymes possess a highly conserved catalytic domain of approximately 390 amino acids and appear to deacetylate their substrates by activating a water molecule with a divalent zinc cation coupled to a histidine-aspartate charge-relay system (Finnin et al., 1999). The sirtuin deacetylases contain a conserved 275 amino acid catalytic domain, which is

unrelated to that of the other HDACs, and operates by a very different mechanism that requires NAD as a substrate (Frye, 2000).

HDACs are not redundant in function. Levels of HDACs 1, 2, 3, 5, 6, 7 and 10 are approximately the same in normal tissue, whereas HDAC4 is expressed in embryonic but not adult muscle tissue (de Ruijter et al., 2003; Khochbin et al., 2001). Class I HDACs are found almost exclusively in the nucleus; however, class II HDACs shuttle between the nucleus and cytoplasm upon certain cellular signals (de Ruijter et al., 2003; Grozinger et al., 2001). HDACs 1 and 3 can deacetylate all four core histones at every lysine tested, although with different degrees of efficacy (Turner, 2002). HDAC6 has greater deacetylation activity on histone 4, Lys 5 and Lys 8.

HDACs do not bind directly to DNA but are recruited by protein complexes that differ in their subunit composition (Khochbin et al., 2001) (Fig. 1.13.1). For example, HDACs 1 and 2 have been found in complexes with Sir3, NuRD (nucleosome remodeling and deacetylation) and CoREST (Andres et al., 1999). A "core complex," consisting of HDAC1 and 2 and the histone binding proteins RbAp46 and 48, has been described (Zhang et al., 1999) in all of them.

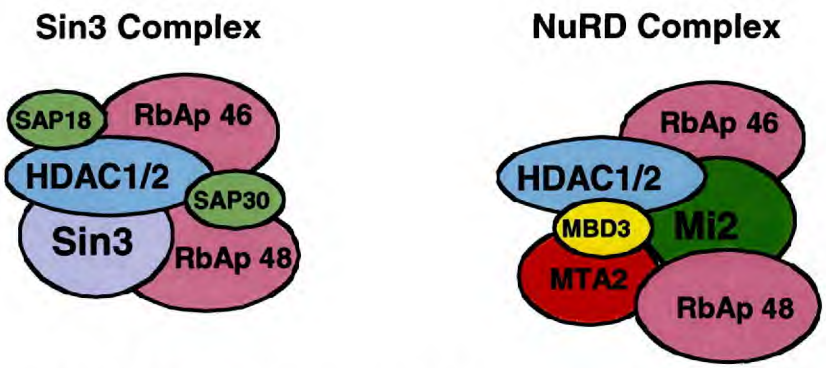


Fig.1.13.1 Schematic rapresentation of complexes containing HDACs.

The Sin3 complex consists of the core complex, along with SAP30 and SAP18. SAP30 is necessary to target the mSin3A-HDAC complex to a subset of repressors that utilize the nuclear hormone co-repressor N-CoR (Ayer, 1999), and SAP18 is required for maximal repression at some yeast promoters (Laherty et al., 1998; Lusser et al., 1997). The mSin3A protein instead serves as a scaffold for the assembly of the entire complex and for its interaction with various DNA binding proteins (Eilers et al., 1999). The NuRD complex (nucleosomal remodeling and

deacetylation) contains the HDAC1 and 2 core complex and the CHD3 and CHD4 proteins (also called Mi-2α and Mi-2β), which posses the DNA helicase/ATPase domain found in the SWI/SNF family of chromatin remodeling proteins (reviewed in Knoepfler and Eisenman, 1999).

Another distinguishing feature of NuRD is the inclusion of the protein MTA2. MTA2 is related to protein MTA1, which is overexpressed in metastatic tumor cells (Toh et al., 1994) and contains a zinc finger and a SANT domain (Aasland et al., 1996; Solari et al., 1999). Unlike Sin3 and NuRD, the CoREST complex contains HDAC1 and HDAC2 but does not appear to include RbAp46 or RbAp48. Two SANT domains have been identified in CoREST, the N-terminal one has shown to be required for the recruitment of HDAC1 to promoters (You et al., 2001).

1.13.1 Regulation of tanscriptional silencing by HDACs

In order for transcription to be silenced, HDACs must be targeted to the promoter regions of genes because they cannot directly interact with DNA or histone proteins. Four mechanisms for regulating HDAC activity have been described. The first three mechanisms involve recruiting HDACs to genes via interactions with specific DNA binding proteins, whereas the fourth mechanism involves a higher order regulatory system in which HDACs are inactivated by sequestration (reviewed in (Burke and Baniahmad, 2000; Wade, 2001).

In the simplest case, HDACs can directly interact with a DNA binding protein that specifically associates with a set of promoters. For example, HDAC1 appears to interact directly with the factor YY1 (Yang et al., 1996), whereas HDAC4, 5, 7, and 9 bind to the MEF (myeloid Elf-1 like factor) family of transcription factors. The second mechanism sees the HDAC proteins recruited by the interaction of a DNA-binding protein with other members of the complex (reviewed in (Cress and Seto, 2000)). Alternatively, the NuRD complex is recruited by the polycomb and hunchback repressors, which are involved in silencing gene expression from clusters of homeotic genes (Knoepfler and Eisenman, 1999). A third mechanism for regulating HDAC activity involves blocking the binding of a HDAC to a transcription factor. Such a mechanism regulates the interaction between HDAC4 and MEF2 (Myocyte enhancer factor 2) (Youn et al., 2000). The fourth mechanism for regulating HDAC activity involves sequestration of HDACs in an extranuclear compartment. This method of regulation thus operates at a higher level than the

previously described mechanisms. This process has been characterized for the class II HDACs 4, 5, and 7. These HDACs can be phosphorylated in their N-terminal domains. This phosphorylation both disrupts the HDAC-MEF interaction (McKinsey et al., 2001) and allows binding at the phosphorylated sites and results in the shuttling of the HDACs out of the nucleus into the cytoplasm (Grozinger and Schreiber, 2000).

1.13.2 Sirtuin deacetylases

Several sirtuins have been identified in prokaryotes and eukaryotes, and these can be divided into five classes based on sequence homology. S. cerevisiae has five sirtuins, including the founding member, Sir2, and four homologs, Hst1–4 (homolog of Sir two). Four sirtuins have been identified thus far in C. elegans, and five in Drosophila. Seven sirtuins (SIRT1-7) have been identified in humans. Several bacterial species, including E. coli, Streptomyces, and Bacillus subtilis, also have sirtuins (reviewed by (Frye, 2000). Yeast Sir2 is the best-characterized sirtuin protein. Sir2 is found in two major complexes, one of which is responsible for silencing expression from the telomeric and HM loci (Sir2, Sir3, and Sir4), and a second that suppresses genetic recombination by forming a closed chromatin structure at the rDNA locus (Sir2, Net1, Cdc14) (reviewed in Gartenberg, 2000).

When tethered to telomeric loci, Sir2 performs several functions aside from silencing gene expression. Telomeric Sir2 is required for the maintenance of genomic integrity and may play a role in the DNA damage response (Palladino et al., 1993). At the rDNA locus, yeast Sir2 inhibits genetic recombination (Straight et al., 1999). In yeast, Sir2 has been linked to aging, an effect that is thought to be mediated by its NAD dependence and repression of rDNA recombination [reviewed (Guarente, 2000)]. Regulation of transcriptional silencing by Sir2 is similar to that observed in the HDAC family of histone deacetylases in that Sir2 must be initially recruited to a specific promoter or to chromosomal domains via DNA binding proteins (Braunstein et al., 1996; Muth et al., 2001). Finally, Sir2 is generally regulated by a higher level of organization similar to the class II HDACs in that it is inactivated by sequestration within the nucleus (Cockell and Gasser, 1999).

1.14 Histone phosphorylation

Phosphorylation occurs in all four core histones, histone variants and histone H1 at specific serine and threonine residues. The overall charge of the histone is altered by phosphorylation and this can disrupt the electrostatic interactions between histones and DNA. Similar to acetylation, this modification is reversible, and we now know that acetylation or phosphorylation of histones can activate transcription whereas deacetylation or dephosphorylation of histones leads to transcriptional repression [reviewed by (Cheung et al., 2000a; Iizuka and Smith, 2003)].

In addition Ser 10 phosphorylation of histone H3 has a major role in chromosome condensation and segregation during mitosis and meiosis (Gurley et al., 1978). Mitosis-specific phosphorylation of histone H3 also occurs at Ser 28 (Goto et al., 1999) and at threonine 11 (Thr 11) (Preuss et al., 2003). Members of the aurora AIR2 – Ipl1 kinase family have been found to govern histone H3 phosphorylation at Ser 10 during mitosis in several organisms (de la Barre et al., 2000; De Souza et al., 2000; Giet and Glover, 2001). The action of these particular kinases is thought to be required for the proper recruitment of the condensin complex and assembly of the mitotic spindle in a phosphorylated histone H3-dependent manner (Giet and Glover, 2001). These kinases are counter-balanced by the activity of type 1 phosphatases (PP1) (Hsu et al., 2000; Murnion et al., 2001).

The link between phosphorylation of histone H3 and transcriptional activation was first postulated in 1991, when Mahadevan and co-workers observed a rapid phosphorylation of histone H3 molecules concomitant with the activation of the c-fos and c-jun immediate-early (IE) response genes. This effect was dependent upon the stimulation of the cells with growth factors, phorbol esters, protein synthesis and protein phosphatases inhibitors (Mahadevan et al., 1991). *In vitro*, experiments revealed that the yeast GCN5 histone acetyltransferase displayed a preference for binding to a portion of the histone H3 tail that is pre-phosphorylated at the Ser 10 position (Cheung et al., 2000b).

Histone H3 phosphorylation at serine 10 has been coupled to both acetylation and methylation of other residues on the same histone tail. Phosphorylation of histone H3 at serine 10 and acetylation of histone H3 lysine 9 and 14 are linked to transcriptional activation. Antibodies generated specifically to histone H3 tails that

were both phosphorylated and acetylated provided the first direct proof that the double-modified H3 isoform existed in vivo (Cheung et al., 2000b; Clayton et al., 2000).

Nucleosome particles containing acetylated and phosphorylated histone H3 tails have been detected at the promoters of several MAP-kinase-activated IE genes and at the promoters of c-fos pathway genes by chromatin immunoprecipitation (ChIP) assays (Cheung et al., 2000b). Phosphorylation and acetylation however are each deposited via independent pathways, suggesting that, at least in some cases, histone H3 phosphorylation at Ser 10 is not just a signal for subsequent acetylation at Lys 14. Analysis of c-jun induction suggests that the presence of a phosphorylated histone H3 tail does not necessarily predispose that particular tail to acetylation at Lys 14. In another study, the mouse mammary tumour virus (MMTV) promoter was shown to be inactivated via a mechanism whereby histone H1 is dephosphorylated in response to glucocorticoids. Removal of glucocorticoids resulted in the rephosphorylation of histone H1 and reactivation of transcription by the promoter. This study indicated that phosphorylated H1 bound at the MMTV promoter is linked with the activation of transcription in vivo (Lee and Archer, 1998).

Coupling of H3 phosphorylation at serine 10 and methylation of Lys 9 has also been studied. It has been shown that phosphorylation of a synthetic peptide by recombinant Ipl1 aurora kinase decreased when the H3K9 residue was methylated, and that phosphorylation of Ser 10 prevented methylation of Lys 9 by the SUV39H1 methyltransferase. Mouse embryonic fibroblasts with mutations in suv39h1 have increased levels of Ser 10-phosphorylated histone H3 (Rea et al., 2000). More recently two studies in particular have investigated the relationship between H3K9me3 and H3S10ph. Fischle et al have shown that H3S10ph during M phase cause release of HP1α, -β and -γ from chromatin even though the level of H3K9me3 remained unchanged. Inhibition of the mitotic kinase Aurora B, which is necessary for Ser 10 phosphorylation, caused retention of HP1 on mitotic chromosomes. These findings lead to the conclusion that H3S10ph is necessary for dissociation of HP1 from chromatin in M phase (Fischle et al., 2005). In contrast to these results Matescuu et al showed that HP1α binding to chromatin was actually favoured, rather than prevented HP1 by histone H3 serine 10 phosphorylation. In this case, dissociation of HP1 from the methylated histone H3 tails was observed only after the appearance of a third modification, acetylation of lysine 14, which occurs only in

prophase (Mateescu et al., 2004). However, it should be noted that most of the experiments showed by Fischle et al were performed with the isoform HP1β and not with HP1α. It has also been shown that decrease in the levels of JIL-1, a H3S10 kinase, lead to spreading of H3K9methylation and the presence of HP1 at ectopic sites. This finding suggest a mechanism for antagonizing the recruitment of HP1 by Su(var)3-9, making repression readily reversible (Zhang et al., 2006).

1.15 Ubiquitination of histones

Ubiquitination is a reversible modification, primarily seen on the histone tails of H2A and H2B. Ubiquitin (Ub) is a 76 amino acid protein that covalently attaches to a target protein through an isopeptide bond between its C-terminal glycine and the of a lysine residue on the acceptor protein (Pickart, 2001b). Ub modification is a complex process involving an Ub-activating enzyme (E1), an Ub-conjugating enzyme (E2), and E3 ligase enzyme; with E3 mediating specific recognition and ligation on the target. Poly-Ub chains can form, or alternatively, some proteins are mono-ubiquitinated on one or multiple lysine residues. The functional significance of these various forms of ubiqitination is still unclear. Generally, Ub induces proteasome-dependent degradation (Pickart, 2001a) and also regulates nuclear-cytoplasmic trafficking of transcriptional regulators (Shcherbik and Haines, 2004). In Drosophila, Sauer and colleagues identified a histone-ubiquitinating enzyme, the coactivator TAF II250, the central subunit of the TFIID complex, which is part of the basal transcription machinery. TAFIIs are known to interact with DNA, general transcription factors and transcription factors to facilitate activation of transcription and, binding of TFIID to promoter regions is thought to represent an initial step in transcription. They went on to show that TAFII250 mono-ubiquitinates histone H1 in vitro (Pham and Sauer, 2000). Other studies have also provided evidence that H2A ubiquitination is a novel epigenetic marker for the inactive X chromosome. A link was demonstrated between H2A ubiquitination and H2A lysine 119 ubiquitin E3 ligase Ring1b, a Polycomb group protein, and the initiation of X inactivation (Fang et al., 2004).

1.16 DNA methylation

DNA methylation in eukaryotes occurs at cytosine bases within CpG dinucleotides (CpGs) and it is associated with transcriptional repression. Early studies revealed that high levels of CpG methylation are observed in heterochromatic regions. Also, it has been shown that DNA that has been methylated *in vitro* when transfected into cells and integrated into the genome, forms a repressed chromatin structure. Other work revealed that unmethylated CpG-island chromatin is enriched in hyperacetylated histones (reviewed by Bird and Wolffe, 1999). In mice, DNA methylation is essential for viability, with targeted disruption of the DNA methyltransferase (Dnmt) enzymes resulting in lethality (Li et al., 1992; Okano et al., 1999).

DNA methyltransferases can be grouped into two general classes, maintenance and de-novo. Maintenance methyltransferase Dnmt1 (Bestor et al., 1988) copies pre-existing methylation patterns onto the new DNA strand. Dnmt1 recognises and modifies hemimethylated sites (Yoder and Bestor, 1998; Yoder et al., 1997) generated during DNA replication, preserving in this way the tissue-specific methylation pattern. The de novo methyltransferases Dnmt3 and Dnmt3b (Okano et al., 1998a; Xie et al., 1999) are mainly responsible for introducing cytosine methylation at previously unmethylated CpG sites. By this process methylation patterns can be changed by the introduction of new methyl groups at sites in which neither strand was previously methylated. A fourth DNA methyltransferase, DNMT2, shows weak DNA methyltransferase activity in vitro (Hermann et al., 2003), but targeted deletion of the *DNMT2* gene in embryonic stem cells causes no detectable effect on global DNA methylation, suggesting that this enzyme has little involvement in setting DNA methylation patterns (Okano et al., 1998b). DNMT3L is a DNMT-related protein that does not contain intrinsic DNA methyltransferase activity, but physically associates with DNMT3a and DNMT3b and modulates their catalytic activity (Suetake et al., 2004).

DNA methylation patterns undergo two periods of genome wide DNA methylation reprogramming, one during primordial germ cell differentiation and one during pre-implantation development (Dean et al., 2005). In mice it has been shown that after fertilization the DNA of the male pronucleus is rapidly demethylated. In contrast the

maternal genome initially maintains its methylation status, which then progressively drops until the 8 cell stage when the methylation content appear to be equivalent between maternal and paternal genome. De novo methylation at the late-morula early-blastocyst transition begins to restore methylation to both of the newly defined lineages of the blastocyst. De novo methylation results in an asymmetric amount of methylation in the differentiated lineages, with the inner cell mass (ICM) being hypermethylated in comparison to the trophectoderm (TE) (reviewed by (Dean et al., 2005).

There are two general mechanisms by which DNA methylation can inhibit gene expression: first, modification of cytosine bases can prevent the association of DNA-binding factors with their DNA recognition sequences (Watt and Molloy, 1988), second, DNA methylation can act via methyl binding proteins (MBPs), that recognize methylated CpG, and recruit co-repressors to silence gene expression directly (Boyes and Bird, 1991; Hendrich and Bird, 1998; Nan et al., 1998).

1.16.1 Methyl binding proteins and transcriptional repression

Methyl-CpG-binding proteins (MBPs) use transcriptional co-repressor molecules to silence transcription and to modify surrounding chromatin, providing a link between DNA methylation, chromatin remodelling and histone modifications (Meehan, 2003). Proteins that recognize methyl-CpG were identified more than a decade ago (Lewis et al., 1992; Meehan et al., 1989). They contain a methyl-binding domain (MBDs), which is responsible for binding to methylated CpG dinucleotides. So far six methyl binding protein have been identified, MeCP2, MBD1-4, and Kaiso (Fig.16.1).

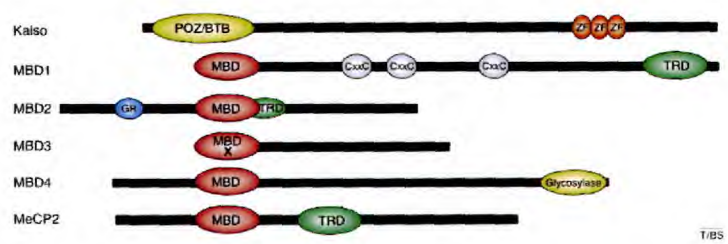


Fig.1.16.1 Methyl-CpG-binding proteins (MBPs) (from Klose, RJ 2006) Six mammalian MBPs have been characterized so far. Kaiso is an atypical MBP, because it depends on a zinc-finger domain (ZF) to recognize methylated DNA and a POZ/BTB domain to repress transcription. All the others MBPs possess a methyl-binding domain, even though the one in MBD3 does not recognize methylated DNA owing to crucial amino acid changes

MBD1 has shown to be associated with a H3 K9 methyltransferase activity and the ESET/SETBD1 H3 K9 histone methyltransferase has been shown to be a MBD1-interacting partner. ChIP studies suggested that H3 K9 methylation could be dependent on MBD1 and on DNA methylation at specific loci. MBD1 also forms a complex with CAF-1 (chromatin assembly factor 1) (Sarraf and Stancheva, 2004). This complex is thought to be involved in the maintenance of histone methylation patterns during DNA replication. MBD1 contains, in addition to the MBD and the transcription repressor domain (TRD), up to three zinc-co-ordinating CXXC domains, depending on the splice variant. The third CXXC domain has been shown to bind specifically to non-methylated CpGs.

MBD2 possesses an MBD domain that overlaps with a TRD domain and a GR repeat at the N terminus. MBD2 is part of the MeCP1/NURD complex that remodels and deacetylases methylated chromatin templates in vitro (Zhang et al., 1999). MBD3 is the only member of the family to have a MBD that do not recognize methylated CpGs owing to amino acid changes (Hendrich and Bird, 1998). The protein has also been identified as a core component of the Mi-2/NURD transcriptional corepressor complex containing HDACs (Feng et al., 2002a; Feng and Zhang, 2001). MBD4 has a C-terminal glycosilase domain that is important for its function in DNA repair (Hendrich et al., 1999; Millar et al., 2002).

MeCP2 is the founding member of the MBD family and contains the conserved MBD domain and an adjacent TRD domain. This protein is involved in a co-repressor complex containing Sin3 and HDAC, and with a histone methyltransferase activity directed against histone H3 lysine 9 *in vivo* (Fuks et al., 2003). Recently it has also been shown that the SWI/SNF ATPase remodelling complex is associated with transcriptional repression and it is part of the MeCP2 repressor complex (Harikrishnan et al., 2005). MeCP2 seems to have a reinforcing activity on repressive chromatin states by acting as a bridge between two global epigenetic modifications, DNA methylation and histone methylation. These findings extend the link between histone modifications, DNA methylation, chromatin remodelling activities and repression.

Kaiso instead is an atypical MBP that does not posses an MBD domain. It can recognize DNA sequences containing two methylated CpG dinucleotides or a related sequence that does not contain methyl-CpG DNA (Daniel et al., 2002; Prokhortchouk et al., 2001). Kaiso recognizes methyl-CpG using zinc-finger

domains present in the C terminus and mediates repression of methylated genes by associating with the histone-deacetylase-containing N-CoR co-repressor complex (Yoon et al., 2003b).

1.16.2 DNA methylation, histone deacetylation and H3K9 methylation

Histone deacetylation, and histone H3 lysine 9 methylation work hand-in-hand with DNA methylation to repress transcription (Freitag and Selker, 2005; Tariq and Paszkowski, 2004). The existence of an epigenetic 'conversation' between histones and DNA, involving cytosine methylation, histone deacetylation, and H3K9 methylation, and leading to transcriptional silencing, is now well established. What remains unclear is the precise sequence of events. As already mentioned, there is evidence that DNA methylation influences the histone modification pattern through the action of MeCP2 and MBD1. Nevertheless some studies have suggested that histone modifications are a prerequisite for DNA methylation. Observations made in fungi, plants and mammals highlight methylation at lysine 9 of H3 as a potential 'sign' for DNA methylation (Jackson et al., 2002; Lehnertz et al., 2003; Tamaru and Selker, 2001). These observations led to the hypothesis that methyl binding proteins could be responsible for setting up a self-reinforcing cycle of repression, which may be necessary for the maintenance and heritability of repressed states. A DNA methyltransferase (DNMT) bound to an adaptor molecule (e.g. HP1) would add a methyl group to DNA when the chromatin is marked by H3K9me. Generation of methylated DNA would then allow binding of MBD proteins to DNA. Some of these proteins are able to recruit histone deacetylase (HDAC) complexes in addition to H3 K9 methyltransferases (H3K9 HMTs). Deacetylation of histone H3 at lysine 9 is necessary for methylation to take place on this residue, therefore H3K9 deacetylation would be followed by histone methylation, which, in turn, might result in recruitment of the HP1 adaptor. Thus, epigenetic information in residue methylation states would flow from histone to DNA and back (Fuks, 2005). This self-reinforcing cycle might be relevant to situations in which a gene or locus are heritably 'locked' into a repressed chromatin state; for example, at hypermethylated genes in cancer or at imprinted genes.

1.17 Embryonic stem cells

The pre-implantation mouse embryo at the blastocyst stage is composed of three distinct cell types: the trophectoderm, the primitive endoderm, and the inner cell mass (ICM). The trophectoderm gives rise to the placenta and the primitive endoderm forms the visceral and parietal endoderm that lines the yolk sac cavity. The ICM is organized into a pluripotent epithelial layer (the epiblast) from which embryonic and other extraembryonic tissues are derived (Gardner and Beddington, 1988). Embryonic stem (ES) cells are derived from the inner cell mass (ICM). They can be propagated in culture in an undifferentiated state (self-renewal) while maintaining the capacity to generate any cell type in the body (pluripotency). Self-renewal requires that the ES cell genome maintain a cellular memory that specifies its pluripotent capacity. At the same time, the genome in pluripotent ES cells must be highly plastic to be able to enter any distinct differentiation pathway. Once differentiation is initiated, lineage specification is accomplished by the execution of genome-expression programmes, which are unique for each cell type. The molecular mechanism behind the self-renewal capacity and pluripotency of ES cells will be described in section 1.17.2. Recent findings point to a role for epigenetic mechanisms, including chromatin structure, chromatin dynamics and histone modifications, in these fundamental processes.

1.17.1 Chromatin structure in ES cells

Chromatin structure can influence gene function by affecting the accessibility of regulatory proteins to their target site and by modulating the affinity of transcriptional regulators with their targets (Meshorer et al., 2006). Studies in several systems indicate that ES cells are rich in euchromatin and, as differentiation progresses, accumulate highly condensed, transcriptionally inactive heterochromatin regions. For example, spatial heterochromatin rearrangement and an increase in the number of heterochromatin foci have been observed after differentiation of embryonal teratocarcinoma F9 stem cells (Cammas et al., 2002) and in murine ES cells (Meshorer et al., 2006). Direct visualization of centromeric heterochromatin with a probe against the major satellite-repeat sequence revealed a more diffuse heterochromatin structure in undifferentiated ES cells versus more compact heterochromatin with well-defined foci in ES-derived neuronal progenitor cells (NPCs) (Meshorer et al., 2006). Consistent with these observations, the genomic

disruption of chromatin remodelling proteins, including BRG1 (Bultman et al., 2000), SNF5 (Klochendler-Yeivin et al., 2000) and SNF2H (Stopka and Skoultchi, 2003), results in premature embryonic death prior to implantation. Lethality occurred at the blastocyst stage, when the inner cell mass (ICM), the source of all ES cells, is formed. In mouse ES cells, the chromatin remodelling NuRD (nucleosome remodelling and histone deacetylation) complex is essential for differentiation (Kaji, K. et al. 2006). Although the detailed mechanism of their functions is unknown, these observations strongly point towards an active role for chromatin-remodelling factors in the maintenance of stem-cell identity and the initial steps of differentiation.

Recent observations in pluripotent ES cells indicate that some DNA-binding proteins undergo remarkably transient interactions with native chromatin in vivo. Dynamic measurements and biochemical analysis indicate that residence times range from a few minutes for linker histones to a few seconds for most transcription factors, remodellers and structural proteins, including heterochromatin protein HP1 (Meshorer et al., 2006). The loosely bound proteins have been suggested to have two opposite effects. They could provide essential building blocks for the formation of chromatin domains as the cells engage their specific expression programmes by the targeted sequestration of genes into active and repressive chromatin domains. Chromatin-assembly factors and remodellers would facilitate their incorporation at these sites and would attract histone- and DNA modifying enzymes to mark these regions epigenetically. On the other side the dynamic nature of chromatin-associated proteins might also make it easier for regulatory factors to gain access to regulatory sequences, and to rapidly activate complex lineage-specific gene expression programmes (Meshorer et al., 2006)

1.17.2 Embryonic stem cells regulators

Self-renewal of mouse embryonic stem (ES) cell in culture depends upon extrinsic signals from leukaemia inhibitory factor (LIF) and bone morphogenetic protein (BMP). The key positive event occurring upon LIF stimulation of ES cells is the tyrosine phosphorylation, dimerization and translocation to the cell nucleus of the signal transducer and activator of transcription, STAT3 (Matsuda et al., 1999; Niwa et al., 1998). In the absence of LIF signalling, induced either by LIF withdrawal or by the expression of a dominant interfering form of STAT3, ES cells

differentiate into a morphologically mixed cell population expressing genes characteristic of endoderm and mesoderm (Niwa et al., 1998; Niwa et al., 2000). The BMP requirement for ES cell self-renewal is strictly dependent upon continued LIF stimulation, since in the absence of LIF, BMP directs differentiation into cells morphologically similar to those formed upon LIF withdrawal in the presence of serum. In contrast, without BMP, neural differentiation occurs (Harland, 2000). BMPs acutely stimulate the transcription of Id genes (Ruzinova and Benezra, 2003). That Id induction is the critical event occurring downstream of BMP was demonstrated by the circumvention of the BMP requirement in ES cells constitutively expressing a loxP flanked *Id1* transgene (Ying et al., 2003). Subsequent excision of the loxP flanked *Id1* expression cassette restored BMP dependence.

In the past few years some intrinsic regulators of ES cell pluripotency have been identified. The best characterized is the homeodomain protein Oct4, a member of the POU class of transcription factors. When the *Oct4* gene is deleted in vivo, blastocysts form in which cells become allocated to the interior but fail to adopt a pluripotent identity, differentiating instead into the trophectodermal lineage (Nichols et al., 1998). The use of a cell line in which Oct4 expression was solely under exogenous control allowed the examination of the effect of maintaining Oct4 expression in cells from which LIF is withdrawn. In this case an apparently similar differentiation occurred to that which occurs when LIF is withdrawn from genetically unmanipulated cells. This indicates that Oct4 is not simply a downstream mediator of LIF signalling and that Oct4 and LIF are both required for ES cell maintenance. Unexpectedly, when Oct4 is over-expressed, ES cells differentiate into cells expressing markers of endoderm and mesoderm (Niwa et al., 2000).

Another homeodomain protein, Nanog, has been shown to be important for ES cell function. When over expressed in ES cells this protein maintains self-renewal in the absence of STAT3 activation, conditions under which ES cells would normally differentiate (Chambers et al., 2003; Mitsui et al., 2003). Experiments in which a loxP flanked *Nanog* transgene was integrated into ES cells showed that Nanog is able to sustain ES cell pluripotency (Chambers et al., 2003). Genetic deletion of Nanog showed that the protein is required for the maintenance of pluripotency, since cells from the interior of the blastocyst do not maintain their pluripotency but rather differentiate into primitive endoderm cells (Mitsui et al., 2003). Compared to *Oct4*

mutants, the ability of *Nanog* mutant embryos to produce the temporally more mature primitive endoderm indicates that Nanog function is critical at a later time during development than Oct4. Deletion of *Nanog* from ES cells also results in primitive endoderm differentiation. While Nanog is capable of maintaining ES cells in the absence of STAT3 activation, self-renewal occurs with maximal efficiency when ES cells are both responding to STAT3 and overexpressing Nanog (Chambers et al., 2003). This cooperative interaction between Nanog and activated STAT3 suggests that each is dependent on the other for maximal function. The fact that Nanog overexpression blocks differentiation induced by LIF withdrawal could suggest that Nanog is downstream of STAT3 but is incapable of being maximally activated by STAT3. However, Nanog does not appear to be a transcriptional target of STAT3 nor is STAT3 activated by Nanog over expression (Chambers et al., 2003). In addition, genetic deletion of *Nanog* produces primitive endodermal differentiation whereas withdrawal of LIF causes differentiation to a mixed endoderm/mesoderm population. This suggests that either STAT3 may interact with molecules other than Nanog or that Nanog may be necessary in the ectoderm to allow development to a stage capable of mesodermal differentiation. Nanog function has been shown to be dependent upon continued *Oct4* expression (Chambers et al., 2003).

Very recently ChIP-Chip analysis of hES cells showed the presence of a complex network of interactions between the two variant homeodomain proteins, OCT4 and NANOG, and SOX2, a HMG factor, at specific promoter elements (Boyer et al., 2005). NANOG occupied 9% of the promoter regions, SOX2 7%, and OCT4 3%. These data point to a high contribution of these transcription factors to the cellular phenotype, rather than a model in which each regulates a very limited number of target genes in a cascade. The frequent co-occupancy of the gene promoters by OCT4, NANOG, and SOX2, led to the conclusion that these three factors work together to control a whole set of target genes in ES cells.

In a second level of analysis, Boyer et al. (2005) showed that roughly half of the genes bound by OCT4, NANOG, and SOX2 are expressed in ES cells. Among these are the promoters of the genes themselves, as well as components of signalling pathways such as the *TGF-β* and *Wnt*. Those pathways that have been implicated in ES cell self-renewal. Genes whose promoters are co-occupied by OCT4, NANOG,

and SOX2 and are not expressed include many transcription factors believed to be critical for germ-layer-specific differentiation.

Two regulatory network models have been proposed to account for these new findings. One model (feed forward loop) suggests that OCT4 and SOX2 converge on the regulation of NANOG, which, in turn, acts with these proteins to control a vast array of downstream targets. This model allows for both stability and developmental switching, depending on the activities and concentrations of the individual factors, and also suggests that NANOG maintains a central position in the regulatory hierarchy. The other model, consistent with the co-occupancy of the promoters of the genes that produce the factors, propose that OCT4, SOX2, and NANOG are each forming an auto-regulatory loop. It seems that OCT4, SOX2, and NANOG are likely to promote their own expression and the expression of other pluripotency/self-renewal genes, while simultaneously preventing expression of differentiation-promoting genes.

1.17.3 Models for embryonic stem cell pluripotency

Altered chromatin structure can contribute to the unique properties of stem cells, including pluripotency and their potential to self-renew. A number of distinct models have been proposed. The gene-specific early marking model envisions that tissue specific genes, not active in ES cells, are epigenetically marked for expression at a later stage. According to this model transcription competence of lineage-specific genes is determined by localized marks of histone modifications established by binding of sequence-specific transcription factors to discrete regions at the ES cell stage. These early transcription competence marks would help to maintain pluripotency by preventing the spread of repressive chromatin modifications and would act as recruitment centres for transcriptional activators as ES cells commit to different lineages (reviewed by Szutorisz and Dillon, 2005). After lineage commitment genes belonging to different fate choices would loose their mark and become permanently silenced. This early transcription competence marks model (ETCM) is based on the observation that several lineage-specific genes, which are repressed in ES cells, contain active chromatin marks (Chambeyron and Bickmore, 2004; Forsberg et al., 2000; Szutorisz et al., 2005; Vieira et al., 2004). Two other models have been proposed to explain ES cells pluripotency, the hierarchical activation (HA) and the promiscuous transcription (PT) model. The first proposes

that ES cells express, in addition to their housekeeping genes, a minimal set of genes that are responsible for self-renewal and maintenance of pluripotency. Following differentiation, these genes, specific for stem cells, would be silenced and expression of lineage-specific transcription factors would trigger sets of lineage-specific genes (Fisher, 2002; Szutorisz and Dillon, 2005). The promiscuous transcription (PT) model, quite similar to the HA model, suggest that ES cells express a set of stem cell-specific genes controlled by stem cell-specific transcription factors but, in contrast to a strictly hierarchical activation, proposes that the rest of the genome is expressed at low levels. Upon differentiation, a global reduction of gene activity would occur, causing silencing of the stem cell specific genes and activation of sets of lineage-specific genes. Activity of tissue specific genes would be reinforced by increased expression of lineage-specific transcription factors (Munoz-Sanjuan and Brivanlou, 2002).

1.18 B cell biology

B cells are generated from multipotent haematopoietic stem cells first in the foetal liver and, after birth, in the bone marrow. They are responsible for humoral immunity, which allows the immune system to respond to almost any antigen to which it is exposed, because of the diversity of the antibody repertoire. The B cell pathway can be divided in a series of discrete developmental states (Fig.1.18.1) (Singh et al., 2005). Each state can be described on the basis of a unique constellation of cell surface markers and/or gene expression patterns, as well as from the rearrangement status of the immunoglobulin (Ig) genes. The first functional B cells are known as immature B cells and are defined by expression of IgM at the cell surface. Immature B cells exit the bone marrow and enter the spleen, where they further differentiate through several transitional stages and eventually become mature follicular or marginal-zone B cells. These cells can also undergo terminal differentiation to become antibody producing plasma cells (Sabbattini and Dillon, 2005).

The expression of IgM at the cell surface depends on the correct rearrangements of the cell's heavy (IgH) and light (IgL) chains which pair to form the antigen-binding site. Specificity and variation are generated by the diverse assembly of the germiline V (variable), D (diversity) and J (joining) segments through a series of highly regulated site-specific recombination events. The newly created V(D)J

segments, are located upstream of another element of the receptor, the constant (C) region. As an example of the potential for recombination, the IgH locus contains approximately 200 variable genes, 12 diversity genes and 4 joining genes. Variations generated by V(D)J recombination are myriad because they do not simply rely on the number of gene segments. Further diversity is generated by the introduction of small insertions and deletions at the junctions between rearranged gene segments. However, not all gene rearrangement events are successful. A high proportion will produce a non in-frame DNA sequence that cannot be translated into protein. A successful gene rearrangement, termed as a productive rearrangement, leads to the synthesis of the protein product that represent the signal for the cell to progress to the next stage of development.

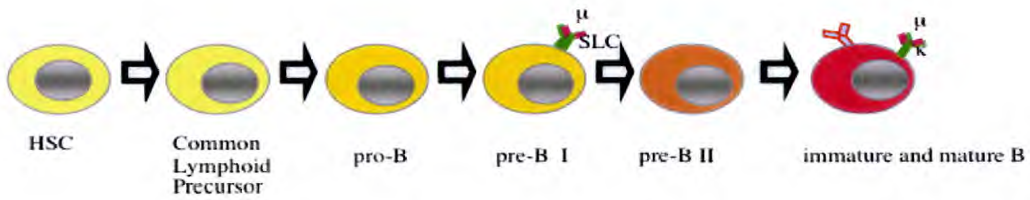


Fig.1.18.1 Stages of B cell development. HSC = haematopoietic stem cell. SLC = surrogate light chain.

Developmental stage	HSC	Lymphoid precursor	Pro-B	Pre-B I	Pre-B II	Immature and mature B
Transcription factors	E2A Ikaros	E2A Ikaros	E2A Ikaros EBF Pax5	E2A Ikaros EBF Pax5	E2A Ikaros EBF Pax5	E2A Ikaros EBF Pax5
Gene expression			*RAG* *λ5 and VpreB1*	- *λ5 and VpreB1* μ	*RAG* *λ5 and VpreB1* (c) μ	K?.
Rearrangement IgH IgL			DJH VKJK	VH to DJH	VDJ	surface IgM VDJ VκJκ;VλJλ
Surface receptor				SLC gp130	Cμ+	IgH IgL
Marker expression			B220+,CD19- c-kit+, CD43+	B220+,CD19+ c-kit+, CD43+	B220+,CD25+ CD19+	B220+,CD19+ IgM+/IgD+

Table 1.18.1 Stages of B-lymphocyte development (adapted from Martensson and Ceredig 2000)

1.19 B cell development

In stem cells, as well as in any other non-lymphoid cell, the immunoglobulin genes are in the germline configuration. The first rearrangement event occurs in early pro-B cells at the heavy chain locus, promoted by the recombination-activating proteins RAG1 and RAG2 which introduce double strand breaks at conserved recombination signal sequences (RSSs) that flanks all Ig genes. Artificial expression of the RAG proteins in mammalian cells other than B or T lymphocytes suggests that they are the only lymphocyte-specific factors required for this recombination event to occur (Schatz et al., 1989). In *RAG*$^{-/-}$ mice V(D)J recombination is completely abolished. These mice have neither mature B nor T cells and this causes a severe combined immunodeficiency syndrome (SCID) (Mombaerts et al., 1992; Shinkai et al., 1992). Severe immunodeficiency, called Omenn syndrome, is seen in humans with miss sense mutations in either of their *RAG* genes (Villa et al., 1998). At the early pro-B cell stage DH and JH segment are rearranged and the cell proceeds to the late pro-B cell (pre-B I) stage where VH to DJH rearrangement occurs. V(D)J rearrangements are subject to allelic exclusion, which means that productive recombination at one allele and expression of a functional IgH chain inhibit further rearrangements on the other allele. The phenotype of early pre-B cells is also characterized by the expression on the membrane proteins B220 and CD19, which are expressed throughout B-cell development. At this stage CD43 is also expressed, but it is downregulated at the immature B cell transition. C-kit, is expressed at the late pro-B cell stage and is lost when the cells move towards the large pre-B cell stage.

A productive VDJH joining leads to the expression of an intact immunoglobulin heavy chain (μ) at the cell surface as part of the pre-B cell receptor (pre-BCR). Association of the μ heavy chain with the SLC proteins *λ5* and *VpreB1*, and the signal transducing heterodimer Igα and β forms the pre-BCR (Karasuyama et al., 1990), which is mainly expressed within the cell and in small amounts at the cell surface. These events precede rearrangement and expression of the light chain genes and marks a critical step in B-cell differentiation as the pre-BCR mediates signalling for cell survival and proliferation. Early B cells that fail to express a functional pre-BCR are compromised in further development. At this stage, cell viability and

proliferation are dependent on additional factors such as interleukin-7 (IL-7) and its receptor (IL-7R). Evidence from studies in IL-7-deficient mice (von Freeden-Jeffry et al., 1995) and in mice treated with anti-IL-7 monoclonal antibodies in vivo (Grabstein et al., 1993), indicate that only small numbers of B lineage cells develop beyond the pro-B cell stage in the absence of IL-7. Stimulation by IL-7 also appears to play a role in the differentiation of B cells by promoting rearrangement of μ H chain genes in pro-B cells (Corcoran et al., 1998). It has been shown that certain precursor B cell lines rapidly undergo apoptotic cell death following IL-7 withdrawal (Griffiths et al., 1994).

After a few divisions large pre-B cells give rise to small pre-B cells, in which light chain rearrangements begins and the μ heavy chain is expressed in the cytoplasm only. At this stage the *RAG* genes are up-regulated and VL to JL rearrangements are initiated at the light-chain loci. Despite the presence of an active recombinase, IgH loci containing a DJH rearrangement are not targeted for further V-DJH recombination in pre-B cells. Upon complete rearrangement of the light-chain genes, the cells express a complete IgM molecule at their surface and become immature B cells.

The next stage of development sees the production of a δ heavy chain as well as of a μ heavy chain by a mechanism of alternative mRNA splicing, and it is marked by the additional appearance of IgD on the cell surface. Immature B cells are also called naïve B cells until they encounter their specific antigen.

1.20 Transcriptional regulation during B cell development

The control of B-lymphocyte development from haematopoietic stem cells (HSC) is known to involve the action of two cytokine receptors (Flk2/Flt3 and IL-7R) and at least six transcription factors (PU.1, Ikaros, E2A, Bcl11a, EBF, and Pax-5) which regulates distinct developmental aspects (Fig 1.20.1). The earliest regulatory events that seem to trigger B cell development are the loss of long-term self-renewal capacity of the HSC accompanied by the expression of the tyrosine kinase receptor foetal liver kinase 2/3 (Flk2/Flt3) within a subset of multipotential progenitor (MPP).

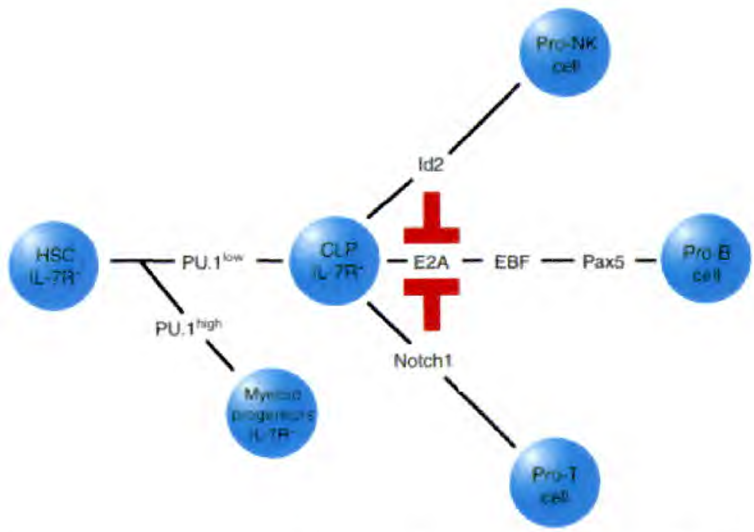


Fig.1.20.1 Transcriptional control of B-cell development (adapted from Busslinger, 2002). HSC = haematopoietic stem cell. At an early stage, low levels of PU.1 activate expression of the IL-7R gene and the cell continues down the lymphoid pathway (CLP = common lymphoid progenitor). Expression of the transcription factors E2A, EBF, and Pax5 further commit the cell to B-cell development. Notch 1 signalling specifies a T-cell fate and the activity of Id2 interferes with E2A and promotes commitment to the natural killer (NK) cell lineage.

These events also restrict the engagement to the myeloid pathway (Adolfsson et al., 2001). Targeted inactivation of *Flk2/Flt3* genes results in a severe deficiency in B lineage precursors (Mackarehtschian et al., 1995). The signalling pathway through which Flk2/Flk3 selectively favours the generation of B lineage progenitors is unknown but *in vitro* data suggest that activation of this receptor promotes expression of the interleukin-7 receptor (IL-7R) (Borge et al., 1999). The IL-7 receptor is expressed primarily on lymphoid progenitors and is important for the

development of both B and T lymphocytes (Kondo et al., 1997; Sudo et al., 1993). It has been shown that combined loss of Flk2/Flt3 and IL-7R results in a complete failure to develop B lineage cells during both foetal and adult haematopoiesis (Sitnicka et al., 2003; Vosshenrich et al., 2003).

The transcription factor PU.1 and Ikaros regulate expression of these two cytokine receptor genes. *Ikaros* null and *ikaros* $DN^{-/-}$ haematopoietic progenitors show impaired expression of Flk2/Flt3 (DeKoter et al., 1998; Nichogiannopoulou et al., 1999) and a reduction of 5-10 folds in c-Kit (tyrosine kinase receptor) levels (Wang et al., 1996). $PU.1^{-/-}$ progenitors, also present defects in the expression of IL-7R and complete absence of B lineage precursors. This is due to the fact that PU.1 is involved in directly regulating the transcription of the *IL-7Rα-chain* gene. Retroviral transduction of the *IL-7Rα-chain* in $PU.1^{-/-}$ cells partly restored B lymphopoiesis, suggesting additional critical functions for PU.1 in B lymphocyte development (DeKoter et al., 2002). It has been shown that low expression of PU.1 induces B cell development whereas high levels suppress the B cell fate and promote machrophage differentiation under in vitro culture conditions (DeKoter and Singh, 2000). This suggests that graded expression of PU.1 specifies either the myeloid or lymphoid fate of early progenitors by controlling their responsiveness to IL-7.

1.20.1 Role of EBF, Pax5 and E2A during B cell development

PU.1 affects multiple hematopoietic lineages, whereas the loss of other transcriptional regulators specifically affects early B cell development and the entry to the pro-B cell stage (Scott et al.,1994). These transcription factors include early B cell factor (EBF), E2A and B cell-specific activator protein (BSAP or Pax5) (Bain et al., 1994; Urbanek et al., 1994; Zhuang et al., 1994; Lin and Grosschedl, 1995; Singh, 1996; Reya and Grosschedl, 1998). EBF and E2A help to specify B-cell fate by inducing expression of the surrogate light genes *λ5* and *VpreB*, of the signalling molecules *Igα* (*mb-1*) and *Igβ* (*B29*), which form the pre-BCR, and of the two components of the rearrangement machinery, the RAG1 and RAG2 proteins. The *E2A* gene encodes, by differential splicing (Murre et al., 1989), the two basic helix-loop-helix proteins (bHLH) E12 and E47, whose expression and activities are induced during early B cell development (Murre et al., 1994). Although these proteins are widely expressed and normally form heterodimers with tissue-restricted

bHLH proteins, they specifically function as homodimers (E47 homodimer, also called B cell factor-1) in the B-lymphoid lineage (Murre et al., 1991; Shen and Kadesch, 1995). Targeted disruption of E2A in mice causes a severe and specific defect in early B cell development, which is arrested at its earliest stages prior to DH-JH rearrangements at the IgH locus (Bain et al., 1994; Zhuang et al., 1994).

EBF also forms homodimers and recognise DNA via an N-terminal domain containing a zinc finger motif (Hagman et al., 1993; Hagman et al., 1995). EBF is expressed in pro-, pre- and mature- B cell but not in terminally differentiated plasma cells. In addition EBF is also expressed in adipose tissue, in brain and in olfactory epithelium (Hagman et al., 1993; Hagman et al., 1991; Wang and Reed, 1993). Loss of EBF results in a similar arrest of early B cell development to that observed in E2A mutants (Lin and Grosschedl, 1995). The few early B cell precursors expressing B220 and IL-7R found in the bone marrow of $EBF^{-/-}$ mice fail to transcribe the *Rag1* and 2, *Igα* (*mb-1*), *Igβ* (*B29*), *λ5*, *VpreB* , *CD19* and *Pax5* genes. It has been shown that E2A and EBF act in synergy to induce transcription of the endogenous *λ5* and *VpreB1* genes when they are expressed ectopically in the early-proB cell line BaF3 (Sigvardsson et al., 1997). It has also been demonstrated that EBF and E2A bind to and activate the promoters of *λ5* (Sigvardsson, 2000; Sigvardsson et al., 1997), *VpreB* (Gisler and Sigvardsson, 2002), and *Igα* genes (Sigvardsson et al., 2002). *Igβ* seems to be regulated by EBF alone (Akerblad et al., 1999).

Chromatin immunoprecipitation experiments have confirmed the presence of E2A at all four target genes (Greenbaum and Zhuang, 2002). Some findings suggest that E2A may act upstream of EBF in the genetic hierarchy of B cell development. It has been shown that a fraction of $EBF^{-/-}$ early pro B cells express normal levels of E2A mRNA, whereas EBF transcripts are reduced in $E2A^{-/-}$ bone marrow cells (Bain et al., 1997; Lin and Grosschedl, 1995). Also, ectopic expression of E2A activates the endogenous EBF gene in a machrophage cell line while EBF is unable to induce E2A transcription in the same cells (Kee and Murre, 1998). Finally, the activity of the EBF promoter seems to depend on a functional E2A-binding site (Smith et al., 2002). It has been shown that transduction of EBF facilitates the rapid generation of B cell precursor from $PU.1^{-/-}$ progenitors. This suggests that the block to B cell

differentiation in the *PU.1*$^{-/-}$ embryos is due to impaired expression of the EBF gene (Medina et al., 2004).

The activation of B-cell specific genes and V(D)J recombination by E2A and EBF is not sufficient to commit B cell progenitor to the lymphoid lineage in the absence of Pax5, also known as B-cell-specific activator protein (BSAP) . B cell development is arrested at an early pro-B cells stage in the bone marrow of *Pax5*$^{-/-}$ mice (Urbanek et al., 1994). *Pax5*$^{-/-}$ pro-B cells can be cultured ex vivo in the presence of IL-7 and stromal cells. The characterization of these cells revealed that they are not yet fully committed to the B lymphoid lineage but instead retain a broad lympho-myeloid differentiation potential. Upon IL-7 withdrawal and appropriate cytokines stimulation the *Pax5* null pro-B cells can be induced to differentiate into a variety of lineages, including granulocytes, dendritic cells, osteoclast, NK cells and T cells. However this multilineage potential is suppressed by retroviral restoration of Pax5 expression (Nutt et al., 1999). Moreover, *Pax5*$^{-/-}$ pro-B cells can reconstitute T cell development in vivo when injected into *RAG2* deficient mice (Rolink et al., 1999).

Pax5 is expressed from the pro-B to the mature B cell stage (Adams et al., 1992) and it seems to play a role in maintaining B cell identity. Pax5 fulfils a dual role by repressing lineage-inappropriate genes and simultaneously activating B-cell-specific genes consolidating the B lymphoid gene expression program (Nutt et al., 1999). It has been shown to activate expression of essential components of the pre-BCR signalling pathway, such us *Ig*α (*mb-1*) (Fitzsimmons et al., 1996; Nutt et al., 1997), *CD19* (Kozmik et al., 1992; Nutt et al., 1998) and *BLNK* (Schebesta et al., 2002) and to repress the expression of several lineage inappropriate genes such as *M-CSF-R* and *Notch1* (Souabni et al., 2002). The mechanisms by which Pax5 regulates the transcriptional activity of its targets are likely to be diverse. It activates the transcription of *lef-1* and *mb-1* by recruitment of the Ets family of transcription factors (Nutt et al., 1998), which are known to activate acetyltransferase complexes (Li et al., 2000). It was also found that Pax5 transcriptional repressor activity involve interaction with Groucho-like co-repressor proteins, which act by reducing chromatin accessibility (Eberhard et al., 2000).

Pax5 seems to take part to the control of B lineage commitment by shutting down inappropriate signalling system and by simultaneously facilitating signal transduction from the pre-BCR and BCR, which both execute important checkpoints

functions in B cell development. E2A, EBF, and their target genes are normally expressed in *Pax*$^{-/-}$ pro-B cells (Nutt et al., 1999; Nutt et al., 1998), while Pax5 expression is reduced in pro-B cells of heterozygous *E2A*$^{+/-}$*EBF*$^{+/-}$ mice (O'Riordan and Grosschedl, 1999). It has also been shown that Pax5 is unable to rescue B cell development from either *PU.1*$^{-/-}$ or *EBF*$^{-/-}$ progenitors (Medina et al., 2004). All of these results indicate that Pax5 functions downstream of E2A and EBF in the control of B cell development.

1.20.2 Role of Ikaros, Aiolos and Dedalos in B cell development

Ikaros, together with Helios, Aiolos and Dedalos constitute a family of zinc finger transcription factors involved in haematopoietic cell fate decision (Georgopoulos, 2002). The Ikaros gene can give rise, by alternative splicing, to eight different Ikaros isoforms. These proteins can form homo- or heterodimers via two C-terminal zinc finger domains, while their DNA-binding specificity is determined by the N-terminal zinc finger (Georgopoulos et al., 1994). Homo- and heterodimerization of isoforms with intact DNA-binding domains greatly increases both their affinity for DNA and their transcriptional activity. Some of the Ikaros isoforms lack the DNA binding domain and function as dominant-negative proteins by sequestering other members into inactive heterodimers (Sun et al., 1996). It has been shown that targeted deletion of the N-terminal zinc finger exons results in a dominant negative (DN) Ikaros allele, which interferes with Aiolos and Helios function (Georgopoulos et al., 1994), whereas inactivation of the C-terminal zinc fingers eliminates the dimerization function and thus creates an Ikaros null allele. Ikaros DN mice fail to generate any B, T NK and DC cells. In the null mutation (Ikaros$^{-/-}$) all the component of the B cell pathway are absent, from the earliest described precursor in the foetal liver and bone marrow to the mature populations in peripheral lymphatic centres, NK development is prevented and a strong reduction of the DC differentiation is observed. All stages of foetal T cell differentiation are also absent, while an abnormal T-lymphopoiesis is present in post-natal mice (Wang et al., 1996).

In sharp contrast to the severe impairment in the production of B and T cells precursors in Ikaros knockout mice, common myeloid precursor and the more restricted granulocyte precursors are present at the near normal numbers (Nichogiannopoulou et al., 1999). These results suggest that Ikaros regulates HSC

activity and enables the differentiation of its multipotent progeny along the lymphoid pathway, whereas it might prohibit the myeloid fate. An Ikaros binding site within the *TdT* promoter has been shown to be important for promoter activity in immature lymphocytes, which suggested that Ikaros might function as a transcriptional activator (Lo et al., 1991). However, other findings suggest the functional activator of *TdT* transcription is the Ets protein Elf-1 (Ernst et al., 1996) and that Ikaros acts instead as a repressor or a competitive inhibitor. FISH and confocal microscopy studies have provided evidence of Ikaros co-localization with centromeric heterochromatin and inactive genes in B cells (Brown et al., 1997).

Mutation of an Ikaros binding site in the $\lambda 5$ gene promoter provided evidence that the protein can mediate transcriptional repression in vivo. Analysis of transgenic mice in fact showed that the mutation causes high levels of expression of the $\lambda 5$ transgene, in activated mature B-cells (Sabbattini et al., 2001). Integration of the same mutant transgene into pericentromeric heterochromatin provided a direct test for the role of centromeric localization in Ikaros mediated silencing. The pericentromeric integration did not rescue silencing, and it was concluded that Ikaros does not repress transcription solely through a re-localisation mechanism.

1.21 The mouse λ5-VpreB1 locus

The λ5 and VpreB genes are located on chromosome 16 in mice and on chromosome 22 in humans. In the mouse there is one *λ5* and three *VpreB* genes. *λ5* and *VpreB1* are closely linked and are separated by only 4.6Kb. They encode for the component of the surrogate light chain proteins and are expressed only at the pro- and pre-B cell stages (Li et al., 1993; Okabe et al., 1992). Analysis of the locus in transgenic mice has shown that all of the sequences required for efficient correctly regulated expression of the genes are located within a 18.3 Kb region containing both genes (Sabbattini et al., 1999). This region extends from 2.8 Kb upstream of the *VpreB1* transcription start site to 7.4 Kb downstream from the 15 poly A site (Fig.1.21.1).

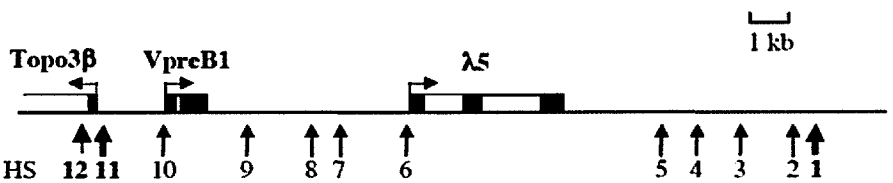


Figure 1.21.1 The *λ5* and *VpreB1* locus. Vertical arrows indicate the positions of DNase I hypersensitive sites in the domain. Bold arrows indicate constitutive HS and pro/pre-B cell specific HS are indicated by non-bold arrows.

1.21.1 Transcriptional regulation of the λ5-VpreB1 locus

The *λ5* and *VpreB1* are activated at the pro-B cell stage and are transcribed in pro- and pre-B cells before being silenced during the transition to the immature B cell stage (Dul et al., 1996; Karasuyama et al., 1994; Kudo and Melchers, 1987; Sakaguchi and Melchers, 1986). Transcription of the *λ5-VpreB1* locus is co-ordinately activated by an LCR which comprises a cluster of 5 DNase I HS sites located in a 6Kb region 3' of the λ5 gene (Minaee et al., 2005; Sabbattini et al., 1999). The promoters of the two genes belong to the initiator family of promoters in that they lack TATA boxes and initiate transcription at multiple sites (Mai and Martensson, 1995).

The gene promoters for both λ5 and *VpreB1* contain multiple binding sites for the transcription factors EBF and E2A (Martensson and Martensson, 1997; Sigvardsson et al., 1997). Evidence of the role of the two transcription factors in *λ5* and *VpreB1*

regulation has been provided by a study in which EBF and CBF-1 (E47 homodimer) were transfected into the Ba/F3 early pro-B cell line. Ba/F3 cells express a number of B cells markers but they represent a stage where *λ5* and *VpreB1* expression is not yet activated. Transfection of *λ5* was activated by expression of EBF alone but was enhanced by co-expression of CBF-1. The E47 homodimer acted in synergy with EBF to activate the *λ5* promoter but was not effective on his own. Mice lacking E2A and EBF gene products do not express *λ5* and *VpreB1* (Bain et al., 1994).

Another transcription factor that may play a role activating *λ5* and *VpreB1* is Pax5. Pax5 binds to the *λ5* and *VpreB1* promoters (Tian et al., 1997), but the significance of the binding is still unclear since Pax5 deficient mice showed normal expression levels for *λ5* and *VpreB1*, implying that Pax5 is not required for the activation of the genes (Rolink et al., 1999). It has also been shown that Pax5 is not essential for in vitro propagation of pro-B cells in the presence of interleukin-7 and stromal cells (Nutt et al., 1997). Pax5 could cooperate with EBF and E2A to activate *λ5* and *VpreB1* expression at the pre-B cell stage, as has been shown to be the case for *mb-1* gene, which encodes for the Igα subunit of the pre-BCR (Sigvardsson et al., 2002).

The *λ5* promoter also contains two adjacent binding sites for the transcription factor Ikaros (Ik1 and Ik2). Ikaros has been implicated as a transcriptional regulator of *λ5* and *VpreB1*, which would promote silencing in the transition to the immature B cell stage. The binding site Ik1 overlaps with a functional EBF binding site that has been shown to be important for promoter activity (Lo et al., 1991; Martensson and Martensson, 1997). The fact that Ikaros binds DNA more efficiently as a dimer (Sun et al., 1996) suggested that co-operative binding of Ikaros to the two sites could repress λ5 transcription by preventing binding of EBF to the overlapping site. Analysis of the effects of a mutation in the Ik-2 binding site in the *λ5* promoter has shown a substantial increase in expression of the gene in mature B-cells, which do not normally express *λ5* (Sabbattini et al., 2001). However, Ikaros is present in developing lymphocytes even when *λ5* and *VpreB1* are active. The implication is that there must be an additional determinant of Ikaros stage specificity. This may involve co-operation with other regulatory proteins such as IRF-4 and IRF-8 that have been shown to be involved in activating Ig light chain transcription (Lu et al., 2003). Another possibility is that post-translational modifications could affect the

ability of Ikaros to bind its DNA target site (Sabbattini and Dillon, 2005). Phosphorylation of Ikaros has been shown to block DNA binding during mitosis leading to dissociation from satellite sequences (Dovat et al., 2002).

1.21.2 Transgenic analysis of the λ5-VpreB1 locus

All of the regulatory sequences required for the full expression of the locus are contained in an 18.3 kb region that extends from an EcoRI site located 2.9kb upstream from the *VpreB1* transcription start site to a BamHI site 6.92 kb downstream from the *λ5* polyadenylation site (Minaee et al., 2005; Sabbattini et al., 1999). Transgenic analysis of the locus has shown that, in addition to the gene promoters, expression of *λ5* is regulated by an LCR-like element which includes a cluster of five DNase I HS located in a 6 kb region 3' of the *λ5* gene (Fig. 1.21.2). This region has shown to be required for position independent expression of the genes at single copy (Sabbattini et al., 1999) and to have strong enhancer activity (S. Minaee PhD Thesis 2002).

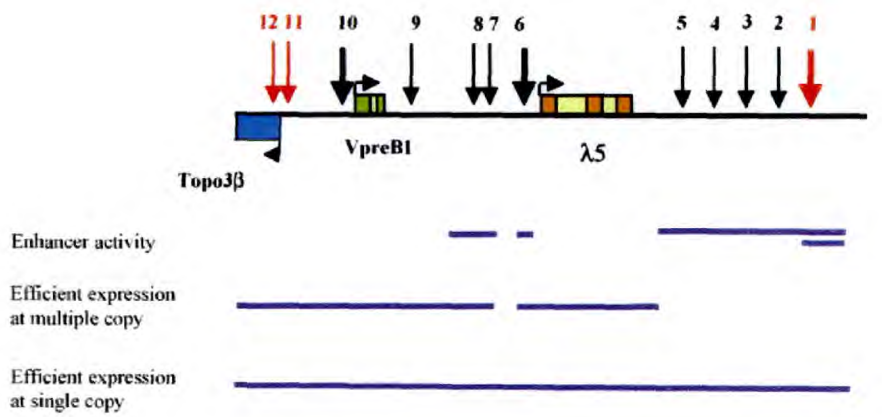


Fig. 1.21.2. Location of sequences that activate *λ5* and *VpreB1* expression. (from Sabbattini and Dillon, 2005). The blue lines represent fragments of the locus that have been shown to display enhancer activity or to give efficient expression of *λ5* and *VpreB1* in transgenic mice. Vertical arrows indicate the position of DNase I hypersensitive sites (black, pro/preB cell specific; red, constitutive).

Four of the hypersensitive sites (HS 2 to 5) are pro- and pre-B cell specific while HS1 has also been detected in non-lymphoid tissues. It has been shown that small *λ5* transgenic constructs lacking the 3' HSs are able to promote efficient copy-dependent expression when integrated into the genome as tandem repeats of three or more copies. This indicated that other elements located elsewhere within the locus

could contribute to the LCR effect. Sequences that can enhance transcription have been identified in the *λ5* promoter and in an intergenic region containing two hypersensitive sites (HS7 and HS8). Enhancer activity has also been detected in constitutive HS1 (Minaee et al., 2005; Sabbattini et al., 1999). Three-dimensional FISH studies of the mechanisms of transcriptional activation in heterochromatin showed that integration into pericentromeric heterochromatin of a *λ5* transgene and of the 3' HS resulted in PEV (Lundgren et al., 2000). Moreover, in the same study, was showed that the transgene was mostly located on the outside of the heterochromatin complex both in pre-B cells and fibroblasts, irrespective of whether it was transcribed or not. The external location in fibroblasts was dependent on a constitutive DNase I hypersensitive site (HS1). When the region containing HS1 was deleted, the transgene was embedded within the complex in most fibroblasts but was still located mainly on the outside of the complex in pre-B cells. The fact that these two different locations were observed in different cell types is particularly significant because it excludes the possibility that the external localization is an effect of different integration sites in the heterochromatin. EBF dosage showed to be important for localization of the *λ5* transgene as well. Reduction of EBF in pre-B cells, resulted in a reduced frequency of localisation of the transgene to the outside of the pericentromeric heterochromatin complex and in lower levels of transcription. The nature of the EBF sensitive element is not yet been established, but the *λ5* promoter region represent a strong candidate (Lundgren et al., 2000).

Sequence analysis identified a third gene, Topoisomerase 3β (*Topo3β*), which transcription start site is only 1.5 kb upstream from the one of the *VpreB1* gene. *Topo3β* is transcribed on the opposite strand respect to *VpreB1 and λ5* and has been shown to be expressed in all the tissues analysed (Szutorisz et al., 2005). The Topo3β gene promoter was also found to overlap with a CpG island and with an HS (HS11). The location of the *λ5-VpreB1* functional domain in close proximity to *Topo3β*, a housekeeping gene has implications for how the two genes are differentially regulated (Szutorisz et al., 2005).

2. Materials and methods

2.1 Standard stock solutions

10 x TBE	890 mM Tris, 890 mM Boric Acid, 20 mM EDTA, pH 8.0
20 x SSC	3 M NaCl, 0.3 M Tri-sodium Citrate, pH 7.0
TE buffer	10 mM Tris-HCl, pH 7.4, 1 mM EDTA
PBS-A	137 mM NaCl, 2.68 mM KCl, 4 mM Na_2HPO_4, 1.76 mM KH_2PO_4
10 % SDS	10 % weight/volume (w/v) Sodium Dodecyl Sulphate
50 x Denhardt's	10 g/l Ficoll (Type 400, Pharmacia), 1g/lPolyvinylpyrolidone, 10 g/l Bovine Serum Albumin (Fraction V)

2.2 Solutions for bacterial culture and transformations

LB-broth with ampicilin	10 g/l Bacto-tryptone (Merck), 5g/l Bacto-yeast extract (GIBCO BRL), 1 g/l NaCl, 100 mg/l Ampicilin
LB-agar with ampicilin	1.5 % Bacto-agar (Merck) in LB-broth, 10 µg/ml Ampicilin
2 x YT	8 g/l Tryptone, 5 g/l NaCl, 5 g/l yeast extract, 100 mg/l Ampicilin
Resuspension buffer TE RNaseA	50 mM Tris, pH 8.0, 10 mM EDTA, 100 µg/ml

Unless otherwise stated all chemicals were obtained from Sigma (Sigma-Aldrich.com).

2.3 Isolation of nucleic acids

2.3.1 Mini-prep of plasmid DNA

QIAGEN Ltd (UK) QIAprep Miniprep® (Crawley, West Sussex, UK)

A single colony of transformed bacteria was used to inoculate 4 ml of LB-broth (containing ampicillin 10 µg/ml) and incubated overnight in 37 °C shaker. QIAprep® Spin Miniprep Kit was used to prepare the plasmid DNA according to the manufacturers instructions.

2.3.2 Maxi-prep of plasmid DNA

QIAGEN Ltd (UK) EndoFree ™ Plasmid Maxi Kit

A 25 µl aliquot of starter culture described above was used to inoculate either 100 ml or 250 ml (high-copy plasmids, low-copy plasmids respectively) of 2x YT broth (containing ampicillin) and incubated overnight at 37 °C with shaking. The EndoFreeTM Maxi Kit was used to prepare plasmid DNA according to manufacturers instructions.

2.3.3 Isolation of mouse genomic DNA from brain and liver tissue

Tail buffer 0.05 M Tris-HCl, pH 8.0, 0.1 M EDTA, 0.1 M NaCl,1 % (w/v) SDS

Adult mice were sacrificed by cervical dislocation and brain and liver tissues were harvested and each placed into individual sterile 15ml falcon tube. To each tube, 1 ml Tail buffer and 20 µl proteinase K (10 mg/ml) were added. Samples were incubated at 55 °C overnight with rotation. The day after a volume of 1 ml 1:1 phenol-chloroform was added to each sample. The samples were then extracted by vigorous shaking for 2 minutes at RT on IKA vibrax VXR shaker moved to a 10 ml corex tube and recovered after spinning at 13000 g for 10 minutes in a centrifuge (the phenol-chloroform extraction step was repeated three times to ensure a clean preparation of DNA). The upper aqueous phase was moved to a new sterile falcon tube and 2.5 ml of 99.7-100% ethanol were added to precipitate the DNA. After addition of ethanol the DNA became visible as a white cotton-like precipitate, which could be spooled with a plastic rod. Each sample was placed into 500 µl of TE buffer. The DNA was allowed to fully dissolve in the TE solution overnight at 4 °C.

2.3.4 Isolation of mouse genomic DNA from ES, primary pre-B and mature B cells

5 x 10^7 cells for each cell types were collected and washed twice with ice cold PBS. Cells were then pelletted in 15ml falcon tubes and 1 ml Tail buffer and 20 µl proteinase K (10 mg/ml) were added. Samples were incubated at 55 °C overnight and the procedure described in section 2.4 was followed.

2.3.5 Oligonucleotides

Oligonucleotides were purchased from MWG BIOTECH (UK) LTD. (Milton Keynes, UK) and Sigma-Genosys Biotechnologies (Europe) Ltd., (Cambridge, UK). The lyophilised oligonucleotides were resuspended in sterile distilled water to a final concentration of 1 ng/µl (standard PCR) and stored at –20 °C.

2.3.6 Quantification of nucleic acids

To determine the concentration of genomic and plasmid DNA, a spectrophotometer was used to measure the OD_{260}. Concentration was estimated from optical density as solution containing 50 µg/ml of double stranded DNA has an absorbance of 1 at 260 nm. Concentration of DNA obtained from ChIP experiment was determined by PicoGreen® ds DNA Quantification Kit (Molecular probes, Cambridge Biosciences, Cambridge, UK) with a standard spectrofluorometer and fluorescein excitation (480 nm) and emission (520 nm) wavelengths.

2.4 Standard manipulations of DNA

2.4.1 Restriction enzyme digestion of plasmid DNA

All restriction enzymes and buffers used were obtained from NEW ENGLAND Biolabs® Inc. (NEB (UK) Ltd., Hertfordshire, UK). Unless otherwise stated reaction conditions used were as directed in the published protocols, in a total volume of 20 µl with incubation periods of 1-16 hours at 37 °C.

2.4.2 Restriction enzyme digestion of genomic DNA.

10 to 50 µg of genomic DNA were digested with up to 150-200 U of the appropriate restriction enzymes (range of 5-20 U/µl) in a final digest volume which varied from 25 up to 200 µl and incubated for a minimum of 2 to a maximum of 16 hours at 37 °C.

2.5 DNA electrophoresis

2.5.1 Agarose gel electrophoresis

Unless otherwise stated, gels were made at 0.8 % (w/v) agarose supplemented with EtBr to a final concentration 0.1 µg/µl. Bacteriophage λ DNA digested with

B*st*E II (NEB) was run alongside the samples as a molecular weight standard. In some cases a 100 bp DNA ladder (NEB) was used as a size marker. Electrophoresis was carried out at a potential of between 40-95 V. DNA was visualised on the EtBr stained gels by long-wave ultraviolet light (UV).

2.5.2 Non-denaturing polyacrylamide gel electrophoresis (PAGE)

Non-denaturing polyacrylamide gels were used to assess chromatin preparations to be used in ChIP experiment. The gels were prepared at 6.0 % (29:1) polyacrylamide and run in 0.5 % x TBE at 200 V for 3 hours then stained in an EtBr bath (EtBr final concentration of 0.1 μg/μl). Bands were visualised by long-wave ultraviolet light (UV).

2.5.3 Isolation of DNA from agarose gels

DNA samples were electrophoresed on 0.7 % (w/v) agarose gels with the appropriate molecular weight marker and assessed by long-wave UV transillumination as described. Gel slices of the required DNA fragments were carefully cut out using a scalpel. Fragments were purified using GENECLEAN *II*® Kit *BIO 101 Inc.*, (Anachem, La Jolla, CA, USA) as described in the manufacturers protocol. This method uses glass milk (silica matrix) to bind DNA. Agarose gel bands are first solubilised with sodium iodide at 55 °C. Proteins and most RNA do not bind to the silica matrix and are eliminated during the washes. The bound DNA is eluted using low-salt buffer. Modifications were made to increase incubation time to overnight on a rotator. Samples were eluted from the glass milk in a 30 μl volume of TE at 55 °C and quantified by spectrophotometry.

2.6 Radio-labelling of DNA

DNA probes between 0.35-2.0 kb were used at a concentration of 50 ng/μl per reaction. The probes were labelled by random-priming using STRATAGENE Prime It® random priming kit according to the manufacturers protocol. 50 μCi of ^{32}P α dATP was used in each reaction (3000 Ci/mmol, Amersham Pharmacia Biotech Inc, Piscataway, NJ, USA). Unincorporated label was removed using Microspin G-25 Columns (Amersham Pharmacia Biotech Inc.,) according to manufacturers instructions.

The reaction was assembled as follows:

50 ng DNA

2 µl Random 9mer Primer

distilled water to a final volume of 10 µl

The mixture was heated at 95 °C for 3-5 minutes and then chilled on ice for 1 minute. The radio-labelling reaction was then assembled by adding:

10 µl DNA/oligonucleotide reagent mixture (as above)

4 µl 5 x dATP buffer (STRATAGENE)

1 µl exo (-) KLENOW (5 U/µl) (STRATAGENE)

5 µl ^{32}P α dATP (Amersham)

Labelling was carried out for 1 hour at 37 °C, and stopped by adding 80 µl of stop solution (see section 2.2.11.4). The labelled probe was then denatured by heating at 95 °C for 5 minutes and immediately chilled on ice before adding it to hybridisation buffer (see section) preheated to 65 °C. Blots were hybridised in a hybridisation oven overnight at 65 °C.

2.7 Generation of DNA probe templates by PCR

DNA probes were generated by PCR with the following reaction mix:

5 ng plasmid DNA

1µl forward primer (1 ng/µl stock)

1 µl reverse primer (1 ng/µl stock)

4 µl $MgCl_2$ (50 mM stock, Bioline)

8 µl dNTPs (2.5 mM stock, Pharmacia Biotech)

1 µl Imolase DNA polymerase (Bioline)

741 µl distilled water

Primers were as described in (Minaee et al., 2005). PCR was carried out using a DNA thermal cycler. The samples were denaturated at 94 °C for 7 minutes, followed by 32 cycles of denaturation at 94 °C for 30 seconds, annealing at 55 °C for 1 minute and extension at 72 °C for 1 minute. A final extension cycle was performed at 72 °C for 10 minutes. The samples were stored at -20 °C until use.

2.8 Southern blot analysis

Southern blotting hybridisation solutions:

Pre-hybridisation buffer

15 ml 20 x SSC

20 ml 50 x Denhardt's

1 ml 10% SDS

500 µl 50 µg/ml denatured salmon sperm DNA

distilled water to 100 ml total volume

Hybridisation buffer

15 ml 20 x SSC

20 ml 50 x Denhardts

1 ml 10% SDS

500 µl 50 µg/ml denatured salmon sperm DNA

10 g Dextran Sulphate (Amersham Pharmacia Biotech AB, Uppsala Sweden)

distilled water to 100 ml total volume

Stop Solution 0.5% (w/v) SDS, 10 mM EDTA

Genomic DNA was extracted from liver, brain tissues and cultured cells as described in sections 2.3.3 and 2.34. DNA was digested for 2-3 hours with increasing amounts of McrBC. Digested DNA was extracted with an equal volume of 1:1 phenol-chloroform, ethanol precipitated and dissolved in 10 mM tris. The recovered DNA was then digested again overnight with the appropriate restriction enzymes at 37 °C and was run on a 0.9 % agarose gel for 16-20 hours at 35 V. 200 ng of DNA B*st*E II and 200 ng of 2Log ladder molecular weight markers (NEB) were used as size markers. The DNA was visualised by staining with ethidium bromide and exposure to UV. The gel was then first soaked for 20 minutes in 500 ml 0.25 M HCl, which resulted in partial depurination and fragmentation of the DNA. This procedure is necessary to give efficient transfer of larger DNA fragments onto the membrane. Afterward the gel was soaked in a 0.5 M NaOH/1.5 M NaCl solution for a minimum 40 minutes. Capillary transfer to Zeta-Probe® Genomic Tested Blotting Membrane (Bio-Rad Laboratories, Hercules, California, USA) was carried

out as described by (Botchan et al., 1976). The DNA was fixed to the membrane by baking for two hours at 80 °C. The membrane was transferred to a bottle and incubated with pre-hybridisation buffer at 65 °C for a minimum of 4 up to 16 hours in a rotating hybridisation oven (Hybaid Mini10). The membrane was hybridised in hybridisation buffer containing radio-labelled probe (prepared as described in sec 2.7) overnight at 65 °C. After incubation the membrane was washed twice with 500 ml 3 x SSC for 1 minute at room temperature in the bottle, transferred to a receptacle and washed twice with 0.1 x SSC, 0.1 % SDS for 20 minutes in a shaking water bath at 65 °C. The membrane was wrapped in cling-film and exposed to a PhosphorScreen (Molecular Dynamics) and left at room temperature for 3-16 hr before being exposed using a 445 SI phosphorimaging machine (Molecular Dynamics) and analysed using ImageQuant Version 5.2 Build 039 Copyright © 1997 Molecular Dynamics software. Blots were also visualised by exposure to Kodak Xomat film.

2.9 Histone proteins SDS PAGE

Histone proteins from native chromatin preparations were checked on sodium dodecyl sulfate (SDS) denaturing acrylamide gels.

Solutions:

Gel stacking:

5% (29:1) polyacrylamide

1% SDS

pH 6.8

Gel separating:

12% (29:1) polyacrylamide

1% SDS

pH 8.8

10x Glycine running buffer:

0.25M Tris base,

1.92M Glylcine

1% (w/v) SDS.

2x Loading buffer:

Tris-HCl 12.5 mM pH 6.8

4% (w/v) SDS

20% (v/v) glycerol

2% β mercaptoethanol

bromophenol blue in deionized water

Fixation solution:

40% methanol

10% acetic acid

Blu Coomassie solution:

40%Trichloric acetic acid

10% methanol

0.2% Blue coomassie

An equal volume of 2x loading buffer was added to 10-20 μg of chromatin. The sample was boiled at 95 °C for 5 minutes and chilled on ice for 2 minutes before being loaded on SDS polyacrylamide gel. Elecrophoresis was performed for one hour at 40 mA. The gel was fixed for 30 minutes / 1 hours in fixation solution and stained for 30 minutes at room temperature with light rocking in Blue Coomassie solution. Decoloration was performed with fixation solution until the bands of interest were not clearly visible. Gels were dried for 45 minutes at 80 °C in a gel drier.

2.10 Cells culture

2.10.1 Transformed cell lines

Abelson transformed pre-B cell (AB1028) and Mature B cell (WeHi) lines were grown in RPMI 1640 media (Sigma) with 10% foetal calf serum (Sigma), 2 mM glutamine (Gibco), 50 μg/ml gentamicyn (Sigma), 50 μM β-mercaptoethanol (Sigma). All cultures were grown at 37 °C.

The AB-1028 non-transgenic pre-B cell line was generated in the laboratory by Andrew Georgiou. Primary pre-B cell cultures from non-transgenic foetal livers were grown in the presence of interleukin-7 (IL-7) for 5-7 days. The cells were then split and transferred onto irradiated ST-2 stromal cell feeder layers. After a total of

12-15 days in culture, transformation of the pre-B cells with Abelson leukaemia virus (Waneck and Rosenburg, 1981) was carried out as described in Sabbattini et al., (1999). FACS analysis of the clone showed that the cells are ~73% positive for both the surrogate light chain (SLC) component λ5 and the B lineage specific marker CD45R (B220) (Fig.2.10.1).

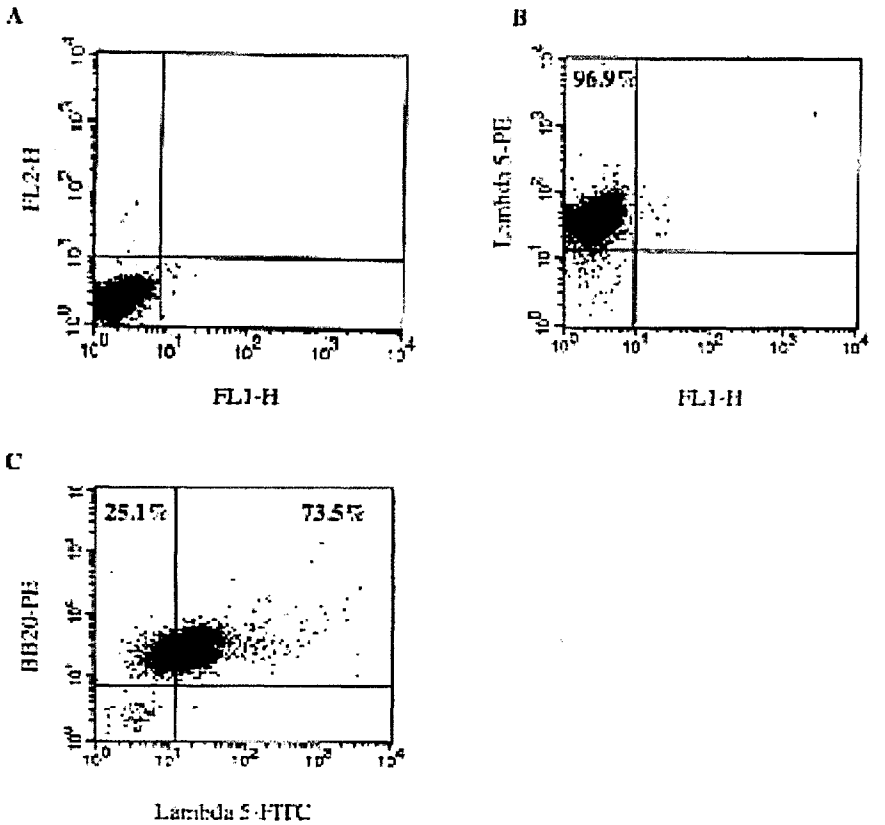


Fig. 2.10.1 FACS analysis of Abelson-transformed non-transgenic pre-B cell (A) Unstained negative control. (B) Percentage of cells positive for the Lambda5 intracellular marker. (C) Double staining for Lambda5 and cell surface B220 markers.

2.10.2 Activation of primary foetal pre-B cell cultures from foetal liver

In order to generate primary pre-B cell cultures from foetal samples, individual livers were dissected and homogenised by passing them through fine meshes. The cells were re-suspended in RPMI media with 15% foetal calf serum supplemented with IL-7 (2 ng/ml; R&D System) and cultured on γ-irradiated ST2 stromal feeder layers. After a total of 12-15 days in culture, FACS analyses of bulk IL-7 dependant foetal liver cultures stained with μ heavy chain was carried out by A. Georgiou to determine their Ig expression status and also to asses the heterogeneity of the cell populations. There was some variation between different bulk foetal liver cultures, but in general the cells were found to be between 10-30% μ+ and 95% B220+ and λ5+. A representative FACS analysis is showed in fig. 2.10.2.

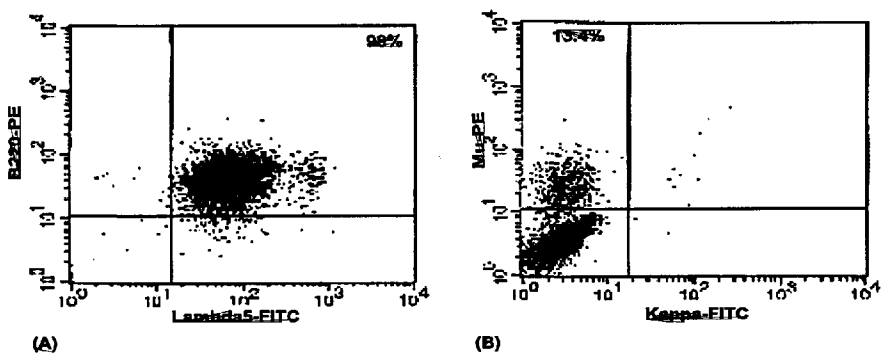


Fig. 2.10.2 FACS analysis of bulk foetal liver primary pre-B cell cultures (A) Percentage of bulk IL-7 dependant foetal liver cells positive for the Lambda5 intracellular marker and B220 cell surface marker. (B) Double staining for Kappa (κ) and Mu (μ) intracellular markers.

2.10.3 Activation of primary mature B cell cultures

Activated B cells were generated from non-transgenic spleens dissected from wild type mice. The spleens (3-5) were homogenised by passing through a fine sieve and re-suspended in 20ml of RPMI media. Cells were purified on Ficoll and then re-suspended in RPMI 15% foetal calf serum supplemented with LPS (2 ng/ml) at a final concentration of 0.5 x 10^6 cells/ml. After Ficoll the cell population is constituted mainly from B lymphocite (60%) and T cells (35%). After activation for three days with LPS, the cell population was constituted 95% from B lymphocites. FACS analysis was carried out by A. Georgiou on several samples during ottimization of the activation protocol. Cells were found 98% positive for the B cell specific surface marker B220 and negative for the T cell specific Thy1 marker. They also showed to be 93% positive for Kappa (k) that, in mature B cells, is expressed on the surface (Fig. 2.10.3).

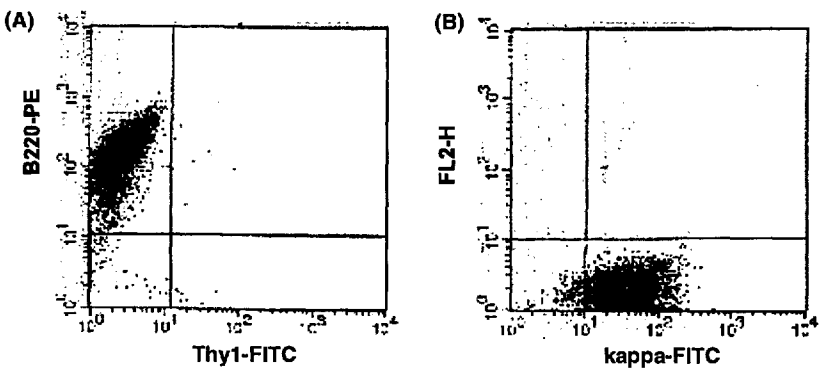


Fig. 2.10.3 FACS analysis of bulk LPS activated mature B cell cultures (A) Percentage of cells positive for the B220 cell surface marker and negative for the T cell specific Thy1 marker. (B) Staining for Kappa (κ) surface marker.

2.10.4 ES cells

E14 ES cells were cultured in 0.1% gelatin coated flasks on a γ-irradiated ST2 stromal feeder layer. Culture medium was Dulbecco minimal essential medium (DMEM) supplemented with 15% foetal calf serum, 2 mM glutamine (Gibco), 50 μgM β-mercaptoethanol (Sigma), 2 mM non-essential amminoacids (Sigma), 2 mM sodium pyruvate (Sigma), 1.8 Unit/μl LIF (Leukemia Inhibitory Factor from Chemicon).

2.10.5 C3H10T1/2 cells

C3H10T1/2 cells were expanded in complete expansion medium (CEM), constituted by DMEM (Dulbecco's Modified Eagle Medium) supplemented with 10% FCS, 100 U/ml penicillin, 100 μgr/ml streptomycin, 2 mM glutamax, 2 mM non-essential amino acids (sigma) and 2 mM sodium pyruvate (sigma). Cells were kept below 80-90% confluency to preserve their multipotent status as from manufacture directions.

Osteogenic differentiation was performed as in Peister et al., 2004 with minor modifications. 10^7 C3H10T1/2 cells were plated in 175 cm^2 flasks and left overnight in CEM. The day after the medium was changed to IMDM (Iscove's Modified Dulbecco Medium) supplemented with 10% FCS, 10% HS, 100 U/ml penicillin, 100 μg/ml streptomycin, 2 mM glutamax, 20 mM βglycerolphosphate, 50 ngr/ml thyroxine, 1 nM dexamethasone and 0.5 μM ascorbate (all from Sigma). The media was changed two times per week for three weeks. The cells were then fixed for maximum 2minutes with 1% formalin and then stained for phosphatase activity.

Adypocites differentiation was as well as in Peister et al., 2004, differentiation media was constituted from IMDM supplemented with 10% FCS, 10% HS, 100 U/ml penicillin, 100 μgr/ml streptomycin, 2 mM glutamax, 20 mM βglycerolphosphate, 5 μgr/ml insulin, 50 μM indomethacin, 10^{-6} M dexamethasone, 0.5 μM isobutyl methylxantine (IBMX). The medium was changed two times a week for three weeks. Cells were fixed with 10% formalin for 20 minutes and stained with 0.5% Oil Red O in methanol for 20 minutes.

2.11 Chromatin Immunoprecipitation (ChIP)

2.11.1 Unfixed Chromatin preparation

Preparation of unfixed chromatin and ChIP were carried out as previously described (Crane-Robinson et al., 1997; Hebbes et al., 1994; O'Neill and Turner, 1995) with some modifications.

Solutions (all the solution were supplemented with inhibitor cocktail tablets from Roche):

1 X TBS:

0.01 M Tris-HCl PH 7.4

3 mM $CaCl_2$

2 mM $MgCl_2$

5 mM sodium butyrate

Digestion Buffer:

50 mM Tris-HCl PH 7.4

1 mM $CaCl_2$

4 mM $MgCl_2$

5 mM sodium butyrate

0.32 M sucrose

Lysis Buffer:

1 mM Tris-HCl PH 7.4

0.2 mM EDTA

5 mM sodium butyrate

Other solutions:

PBS + 5 mM sodium butyrate

25% sucrose in TBS

50% sucrose in TBS

1% NP40 in TBS

10^8 cells for each sample were harvested and washed 3 times with ice cold PBS plus 5 mM Sodium Butyrate. Cell pellets were resuspended in 1 X TBS to a final concentration of 2 x 10^7 cell/ml. An equal volume of 1% NP40 solution and 1/2000 of the volume of 100 mM PMSF were then added and the cells suspension was left

for 40-60 minutes to equilibrate on ice while stirring. After stirring on ice for 40-60 minutes cells suspensions were homogenizer in a glass homogenizer (tight pestel) with nuclei extrusion. Homogenates were centrifuged at 600 rcf for 20 minutes at 4°C. The pellet was resuspended in 25% sucrose in TBS (20 ml) and 2 ml of 50% sucrose in TBS were underlined below. After centrifugation at 1500 rcf for 20 minutes at 4°C, the nuclear pellet was washed once in 25% sucrose in TBS and resuspended in digestion buffer to a concentration of 0.5 mg of DNA/ml. Chromatin was released from the nuclei by digestion with micrococcal nuclease (15 U/ml) for 5 minutes at 37 °C.

Digestion was stopped by addition of EDTA to a final concentration of 5 mM and by cooling on ice. Samples were centrifuged at 11600 rcf for 10 minutes, the supernatant (S1) was recovered and the pellet was resuspended in lysis buffer (1-2ml), left for 30 minutes at 0 °C and then frozen overnight at -20 °C to complete nuclei lysis. S1 was kept overnight at 4 °C after addition protease inhibitor. The day after the samples in the lysis buffer were centrifuged at 4000 rpm in eppendorf tubes and the second supernatant (S2) was collected. Chromatin concentration was determined by UV reading at 260 nm. MNase digestion was assessed on a 6% acrylamide gel or on a 1.2% agarose gel (Fig. 2.11.1). Evetually S1 and S2 were mixed together and used for the immuno precipitation. Proteins were also routinely checked by SDS PAGE as described in section

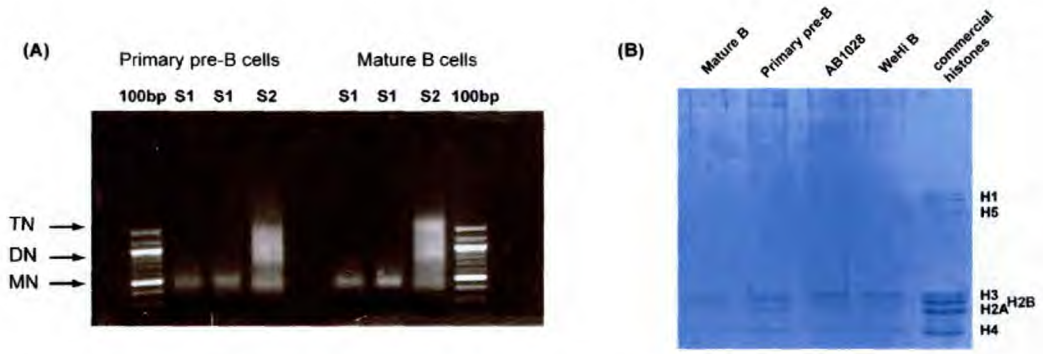


Fig.2.11.1 Native ChIP. (A) Agarose gel (1.2%).Chromatin extracted by micrococcal nuclease digestion (DNA-histone complexes) of nuclei from primary pre-B and mature B cells. S1 fraction showed to be constituted mainly by mononucleosome (MN). S2 fraction composition ranged between mono nucleosome and tri nucleosome arrays. (B) Denaturing protein gel. Proteins were extracted from micrococcal digested chromatin as described in sec 2.9 and run on a 12% separating gel to control the presence and the equimolarity of all the four core nucleosome.

2.11.2 Immunoprecipitation of unfixed chromatin (N-ChIP)

Solutions (all the solution were supplemented with inhibitor cocktail tablets from Roche):

ChIP buffer 1:

50 mM NaCl

1 0mM Tris-HCl PH 7.4

10 mM sodium butyrate

1 mM EDTA

ChIP buffer 2:

100 mM NaCl

10 mM Tris-HCl PH 7.4

10 mM sodium butyrate

1 mM EDTA

ChIP buffer 3:

150 mM NaCl

10 mM Tris-HCl PH 7.4

10 mM sodium butyrate

1 mM EDTA

Protein A agarose beads: before use protein A agarose beads were precleared as follows. 100 µl beads stock per IP sample were washed 2 times with ChIP buffer 1 and then incubated with 1ml ChIP buffer 1 containing 1 mg/ml BSA and 1 mg/ml salmon sperm DNA for 1hour with rotation. After incubation beads were spin down at 1200 rpm for 5 minutes and resuspended in 200 µl of ChIp buffer 1 per IP.

150 µg of unfixed chromatin were used for each immunoprecipitation. Samples were adjusted to 1 ml with ChiP buffer 1 and were pre-cleared with 200 µl of Protein A agarose beads each. The samples were left for 1 hour on a rotating platform at 4 °C. After 1 hour they were centrifuged at 1400 rpm for 5 minutes.
50-100 µl of pre-cleared chromatin were taken as input material. Immunoprecipitations were carried out overnight with rotation at 4 °C.

Immunoprecipitation:

150 µg of chromatin were immunoprecipitated with:

50 µl of anti acetyl histone H4 serum (Upstate cat. numb. 06-866)

50 µl of anti dimethyl histone H3 lysine 4 serum (Upstate cat. numb. 07-030)

20 µl of anti diacetylated histone H3 (200mgr pack Upstate cat. numb. 06-599)

15 µg of normal IgG as a control for aspecific immuno-precipitation (Santa Cruz Biothecnology).

A no antibody sample was also included.

Antibody	Host	Type	Specificity
Anti acetyl histone H4	rabbit	polyclonal	Histone H4 acetylated at lysines 5,8,12 and 16.
Anti acetyl histone H3	rabbit	polyclonal	Histone H3 acetylated at lysines 9 and 14
Anti dimethyl histone H4 lysine 4	rabbit	polyclonal	Histone H3 dimethylated at lysine 4

Tab. 2.10.2 List of antibodies used in unfixed chromatin immunoprecipitation. Data from Upstate (http://www.upstate.com).

After overnight incubation, 200 µl of protein A agarose beads were added to the samples. Samples were left on the rotating platform for 3 hours, and then centrifuged for 5 minutes at 1400 rpm. The supernatant was discarded and the pellet was washed 1 time with 1.5 ml ChIP buffer 1, 1 time with 1.5 ml of ChIP buffer 2 and two times with 1.5 ml of ChIP buffer 3.

The bound fraction was eluted twice with 250 µl of ChIP Buffer 1 with 1% SDS for 15 minutes on rotating wheel at room temperature. DNA was extracted two times with phenol/chloroform, ETOH precipitated and resuspended in 100 µl of tris 10 mM. Glycogen (0.1 µg) was added to each sample before ETOH precipitation to facilitate visualization of the pellet. DNA concentration was determined by Pico green as described in section 2.3.6.

2.11.3 Fixed Chromatin preparation

ChIP of formaldehyde-cross-linked chromatin was carried out essentially as described by (Soutoglou and Talianidis, 2002).

Solutions (all the solution were supplemented with inhibitor cocktail tablets from Roche):

Swelling buffer

25 mM Hepes pH 7.9

1.5 mM $MgCl_2$

10 mM KCl

0.1% NP-40

Sonication buffer

50 mM Hepes pH 7.9

140 mM NaCl

1 mM EDTA

1% Triton X-100

0.1% Na-deoxycholate

0.1% SDS

Wash buffer A

50 mM Hepes pH 7.9

140 mM NaCl

1 mM EDTA

1% Triton X-100

0.1% Na-deoxycholate

0.1% SDS

500 mM NaCl

Wash buffer B

20 mM Tris pH 8.0

1 mM EDTA

250 mM LiCl

0.5% NP-40

0.5% Na-deoxycholate

Elution buffer

50 mM Tris pH 8.0

1 mM EDTA

1% SDS

Protein G sepharose beads: before use protein protein G sepharose beads were precleared as follows. 100 µl beads stock/IP sample were washed 2 times with sonication buffer and then incubated with 1ml sonication buffer containing 1 mg/ml BSA and 1 mg/ml salmon sperm DNA for 1 hour with rotation. After incubation beads were spin down at 1200 rpm for 5 minutes and resuspended in 200 µl of sonication buffer.

1 x 10^8-3 x 10^8 cells for each sample were harvested and crosslinked by addition of formaldehyde to the medium to a final concentration of 1%. After 10 minutes incubation at 37 °C, the reaction was stopped by adding glycine to a final concentration of 0.125 M. Cells were incubated at RT for 5 minutes, and then washed 3 times with ice-cold PBS plus protease inhibitors. Cells were resuspended in 10ml of swelling buffer, chilled on ice for 10 minutes and homogenized in a glass homogenizer (tight pestel) with nuclei extrusion. The nuclei samples were centrifugated at 5000 rpm for 5 min at 4 °C and resuspended in sonication buffer. Chromatin was sonicated by using Soniprep 150 equipment to obtain DNA fragments of approximately 400-500 bp (Fig 2.11.3). After sonication the chromatin sample was spin at 14000 rpm for 15 minutes at 4°C, the supernatant was recovered and spin again. Chromatin was then used for immunoprecipitations.

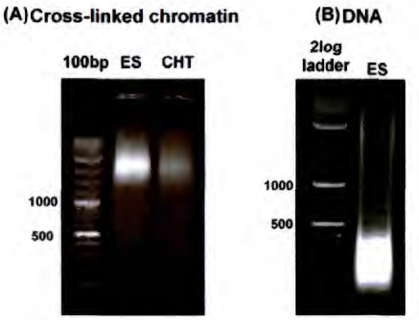


Fig.2.11.3 Fixed ChIP. Agarose gel (1.2%). (A) Chromatin before revarsal of cross linking from ES cells and C3H10T1/2 cells. (B) ES cells DNA after reversal of cross linking and phenol/chloroform extraction. Fragment size ranges between 100 and 500bp.

2.11.4 Immunoprecipitation of fixed chromatin (X-ChIP)

The total amount of chromatin needed for several immunoprecipitations (variable between 125 to 500 µg for each sample depending on the antibody used) was pre-cleared with protein-G-Sepharose beads. Chromatin was spin at 1300 rpm for 5 minutes, the supernatant was collected and the concentration adjusted to 500 µg/ml. Immunoprecipitations were carried out in 1 ml, 500 µl or 250 µl depending on the efficiency of the antibody used. 1/10 of the material used for the immunoprecipitation was kept aside as input. Immunoprecipitations were carried out overnight with rotation at 4°C with the following antibodyes:

Immunoprecipitations carried out with 125µgr of chromatin (in 250µl volume):

12.5 µl anti dimethylated histone H3 lysine 4 serum (Upstate, cat. numb. 07-030)

12.5 µg antibody diacetylated histone H3 (200gr pack, Upstate, cat. numb. 06-599)

Immunoprecipitations carried out with 250µg of chromatin (in 500µl volume):

25 µg anti trimethylated histone H3 lysine 9 (Upstate, cat. numb. 07-442)

125 µl anti H3K9me3/phSer10 (200µl pack, Abcam, cat. numb. ab5819)

25 µg anti-rabbit immunoglobulin G (IgG) (Santa Cruz Biothecnology)

Immunoprecipitations carried out with 500µg of chromatin (in 1ml volume):

50 µg anti dimethylated histone H3 lysine 9 (Upstate, cat.numb 07-441)

50 µg anti Flag (Sigma)

A no antibody, sample was also included.

Antibody	Host	Type	Specificity
Anti acetyl histone H3	rabbit	polyclonal	Histone H3 acetylated at lysines 9 and 14.
Anti dimethyl histone H4 lysine 4	rabbit	polyclonal	Histone H3 dimethylated at lysine 4
Anti dimethyl histone H3 lysine 9	rabbit	polyclonal	Histone H3 dimethylated at lysine 9
Anti trimethyl histone H3 lysine 9	rabbit	polyclonal	Histone H3 trimethylated at lysine 9
Anti H3K9me3/phSer10	rabbit	polyclonal	Histone H3 phosphorylated at residue serine 10 and tri methylated at residue lysine 9

Tab. 2.10.4 List of antibodies used in fixed chromatin immunoprecipitation. Data from Upstate and Abcam (http://www.upstate.com, http://www.abcam.com).

Immunoprecipitated samples were recovered as described in section 2.9.2. Washes were performed twice with sonication buffer, twice with wash buffer A, twice with wash buffer B and twice with TE before eluting two times with 250 µl of elution buffer.

After elution cross link was reversed adding 16 µl 5M NaCl and 1 µl 10 mg/ml DNase-free RNase A to the combined eluates and the input sample and incubating overnight at 65 °C. The day after 5 µl of 10 mg/ml proteinase K were added and samples were incubated for 2 hours at 45 °C before phenol/chloroform extraction and ETOH precipitation. Glycogen (0.1 µg) was added to each sample before ETOH precipitation to facilitate visualization of the pellet. DNA concentration was determined by Pico green as described in section 2.3.6.

2.12 Real time PCR

Real time PCR amplifications were performed on DNA Engine Opticon system (MJ research, Inc.). Reactions were set up in 20 µl using SYBR Green PCR Master Mix (Applied Biosystem). Reactions were carried out in duplicates on 2 ng of DNA for N-ChIP and 2 µl of DNA for X-ChIP.

2.12.1 Primers

Primers across the *λ5-VpreB1* locus and for housekeeping and tissue specific genes were designed by using Primer Select or Primer Express software. The sequence of the entire 19-kb *λ5-VpreB1* locus has been determined elsewhere (accession number AJ852426; Minaee et al., 2005). The positions of the primers in the locus were calculated relative to that of an EcoRI site 2.8 kb upstream from the initiation site for *VpreB1*.

Sequences, concentrations, and annealing temperatures are shown in table 2.10.1a and table 2.10.1b. The thermal cycling conditions were as follows: initial denaturation at 95 °C for 10 minutes, followed by 40 cycles at 95 °C for 15 seconds, annealing at a primer-dependent temperature for 15 seconds, and extension at 72 °C for 15seconds. For each experiment, the threshold was set to cross a point at which real-time PCR amplification was linear (0.02 to 0.05 for the majority of the experiments). See also section 3.4.

Primer	Primer sequence (5' - 3')	Position (kb)	Conc. (nM)	Annealing temp. (C)	Primer	Primer sequence (5' - 3')	Position (kb)	Conc. (nM)	Annealing temp. (C)
1	TTC TTT GCC ACA CCT CAC TAA T	0.204	900	52	29	GGA TAG ACT TTG CAT GTT TTT GAT	7.285	400	52
2	CCT TTT GGG GCT CCT CA	0.31	900	52	30	GGG GGA TAT TAC CTG CTC TTT	7.394	400	52
3	TGC CGG TCA CAG GTC AG	0.522	900	57	31	TGG ATA TCA GTC AGG CAG AG	7.869	900	60
4	GAT GGC TCA GCG CTC TTA CAA	0.623	900	57	32	CCC GGT TGT GGT TGG	7.968	900	60
5	TAC CCA AGT TTT CTC CTA ATC TGC	0.856	900	57	33	CTA AGA GGA GGG GGC TGT GC	8.487	900	70
6	GGC CTG CGG TTT GCG TTC T	0.974	900	57	34	GGA GGT GGG CTA TGG ATG TGG	8.586	900	70
83	AGT CCG AGA ACA GCC TGG GT	0.888	400	68	35	TCA GCA ACC ATT CAT CTA CTA TCA	9.001	900	65
84	AGT TGT GCT GCC CAC AGA GG	0.998	400	68	36	CCT GCC TCT TCC TCC CAA CT	9.119	900	65
708	TTTATTTTTATTTGATATGCATTGGTGTT	1.587	900	57	37	AGG CCC TAA CAG CTT CAT CTA CTC	9.777	400	55
808	ATA TGG CAG CTC ACA ACT ATC CC	1.665	400	57	38	GCA TCT GGG CCT CGG TTT A	9.894	900	55
11	TAG AAG GTT GAG ACA GCG AGT TAG	1.831	900	58	39	GGC TGC ACC TGG AAA ACC TTA	11.701	900	55
12	AGG TGC TTT GTT TTG TTT TGT CTT	1.934	900	58	40	ATC AAA ATC TTC CCC TCA ATC TGT	11.801	400	55
13	CAA ACC CAG AGC CAC AAA G	2.759	900	62	43	AGG AGA CCA GAA GGG GCA GTT	13.382	900	53
14	GGG GAG AGG GTC ACT GC	2.875	900	62	44	GGT TCA ATG TTT AAG AGC AAG TTT	13.491	900	53
15	GGG CTT CTT TGC TCC TCC TAT GG	3.027	900	60	45	ACC TTT GCT TTC TTG TTG TT	14.357	900	50
16	GAA GAT GCT AAT GGT GGC TGA TGC	3.137	900	60	46	TTG AAA TTT TTA GGG ACT GTG AGA	14.456	900	50
17	CAA GTG ACA GGT GTG GAG CAA GTT	3.939	900	60	47	CAC AGA AGA GCA GAG AAC AGA	15.322	900	60
18	GGA GAG CAC ACC CCA GTA GGA TTA	4.046	900	60	48	CAC CCC ATG AGA CAA CCA G	15.432	900	60
19	GTG ACA GAC CCG TTA CCA A	4.445	900	52	49	ACC CAG TAA GCA AGT TTT CA	16.262	900	55
20	AGA AAG AGA AGG GGA AAA AGA G	4.589	400	52	50	ATA AGC TCT CCT CCC TCA AG	16.372	400	55
25	TCC CCA TTG CCA GAT AGA GAC ACA	5.434	900	61	51	TGG AGT ATG AGG CAG TGA TTG TTA	16.731	400	53
26	TGG GCC CAA CAG ATT AAC ACA GAG	5.538	900	61	52	AGA GTT TTG ACG GCT TCC AGA	16.861	900	53
27	TGA CTT GCT TGT GCT TGC CTG GAC	5.732	400	58	53	TGT AAG CTC CTG TAT TTG GTT TTT	16.977	900	54
28	ATA ATA ACA AAG TAC TGA GAA AAC	5.841	900	58	54	GCC CTA CAG AGC TTT TGG AGA	17.103	900	54
57	TGG ATA ACT GCA GGA AGC TGT	6.608	600	68	55	ACC CCT CCC TCC CCG TAT G	17.927	900	62
58	GCA GTG CCA GAT CTC CAG AC	6.728	600	68	56	TCC TTC CCT GAC CAC TCC	18.031	900	62

Table 2.12.1a Sequence and position of PCR primers across the *λ5-VpreB1* locus. Positions in kilobases correspond to the 5' end of each primer pair. Primer pairs 708/808, 13/14 and 31/32 were respectively in the *Topo3β*, *VpreB1* and *λ5* promoters.

	gene	Primer	Primer sequence (5' - 3')	Position (kb)	Conc. (nM)	Annealing temp. (C)
housekeeping genes	β actin-p	forward	GCA GGC CTA GTA ACC GAG ACA	-336	400	61
		reverse	AGT TTT GGC GAT GGG TGC T		400	
	β actin-t	forward	TCC TGG CCT CAC TGT CCA C	+2832	400	62
		reverse	GTC CGC CTA GAA GCA CTT GC		400	
	β2-m-p	forward	AGT CCG GAT TGG CTG TGA GT	-54	400	61
		reverse	GAA GCG ACC GCG ACT GAA		400	
	β2-m-t	forward	CTG ATA CAT ACG CCT GCA GAG TTA A	+3408	400	61
		reverse	TGG AGC GTC CAG AAA GTA AGG		400	
	Topo3β-p	forward	TTT ATT TTT ATT TGA TAT GCA TTG GTG TT	-477	900	57
		reverse	ATA TGG CAG CTC ACA ACT ATC CC		400	
	Topo3β-t	forward	AGT CCG AGA ACA GCC TGG GT	+276	400	65
		reverse	AGT TGT GCT GCC CAC AGA GG		400	
Tissue-specific genes (non-B lineage)	MyoD-p	forward	GAG GTC AGT ACA GGC TGG AGG A	-308	400	61
		reverse	GAG GGC AGA GAG ACT GGC AG		400	
	MyoD-t	forward	GGG CAT CCC TAA GCG ACA	+1806	900	62
		reverse	CAC CTG ATA AAT CGC ATT GGG		900	
	NFL-p	forward	TTT GCT TTT GCG CAG AAT CC	-152	400	61
		reverse	TAT ACG CCG GGA CTC TGA CG		400	
	NFL-t	forward	CTC CTA CTT GAT GTC TGC TCG CT	+2723	400	61
		reverse	CAG CTT TCG TAG CCT CAA TGG		400	
	Amy-p	forward	TTT CTC TTT GGA ATG GTG CAA TAC A	-337	900	61
		reverse	GTG ACC TCC ACC AAC CCG TA		900	
	Amy-t	forward	ATG TTC GTA CCA AGG TGG CTG	+2701	400	61
		reverse	TAT GTC TCC AGG CCA CAT GTG		400	
B-cell specific genes	VpreB1-p	forward	CAA ACC CAG AGC CAC AAA G	-61	900	62
		reverse	GGG GAG AGG GTC ACT GC		900	
	VpreB1-t	forward	GGG CTT CTT TGC TCC TCC TAT GG	+207	900	60
		reverse	GAA GAT GCT AAT GGT GGC TGA TGC		900	
	λ5-p	forward	TGG ATA TCA GTC AGG CAG AG	-473	900	60
		reverse	CCC GGT TGT GGT TGG		900	
	λ5-t	forward	AGG CCC TAA CAG CTT CAT CTA CTC	+1731	400	55
		reverse	GCA TCT GGG CCT CGG TTT A		900	
	Blimp1-p	forward	GAT CGC TAG CTT CCT GTC GC	-312	300	61
		reverse	CGA AGT CGT CGG TCC TGT AA		300	
	Pax5-p	forward	AAA GTG GAA ACT TTT CCT CGC TGT C	-255	400	57
		reverse	AGT AAG CAT GGA TGA ATC CCG TTT C		400	

Table 2.12.1b Sequence and working conditions for housekeeping and tissue-specific genes. Position of the amplicons is given respect to the transcription start site as obtained from Sanger genome data base www.ensembl.org. For the β2 microglobulin gene (*β2-m*) the position of the primers are as in Lonergan M. et al. (1993) Mol Cell Biol 13: 6629-6639. "P" indicates promoter and "t" transcribed regions. Primer pairs Topo3β-p, Topo3β-t, VpreB1-p, VpreB1-t, λ5-p and λ5-t correspond respectively to primer pairs 708/808, 83/84, 13/14, 15/16, 31/32 and 37/38 in table 2.10.1a. *MyoD*=myoblast determination protein 1; *Nfl*=neurofilament light chain; *Amy2*=pancreatic α-amylase; *Pax5*=paired box gene 5, *Blimp1*=B-lymphocyte-induced maturation protein 1.

2.12.2 Data Analysis

Real-time PCR reactions were performed with the same amount of material for the immunoprecipitated and for the input sample used as a reference standard. To determine the enrichment of the immunoprecipitation for a specific antibody in a specific region, the relative abundance for each primer pair amplicon was determined from the difference in the number of PCR cycles necessary to reach a fixed threshold. Values were determined by dividing the amount of target sequence in the immunoprecipitated fraction by the amount of target sequence in the input DNA.

The amount of DNA produced in any PCR reaction can be described by the formula $X=X_0\ 2^{Ct}$ where X and X_0 are respectively the final and the initial DNA concentration of a target sequence and C_t is the number of cycles. The number of cycles required to reach the value X is inversely related to the initial target sequence concentration X_0.

$X(Input)= X_0(Input)\ 2^{CtInput}$

$X(IP)= X_0(IP)\ 2^{CtIP}$

If data are collected at a fixed threshold value the final concentration X became a constant and X(Input)=X(IP).

Therefore, for the same target sequence

$X_0(IP)\ 2^{CtIP} = X_0(Input)\ 2^{CtInput}$ from which:

$\text{Fold Difference} = X_0(IP) / X_0(Input) = 2^{(CtInput-CtIP)}$

The threshold value is set as the point in which the PCR reaction enters the linear range. Errors bars have been calculated as standard deviation on the whole set of value obtained for every single target sequence.

3. Patterns of histone H3 and H4 acetylation and H3 K4 dimethylation across the λ5-VpreB1 locus in pre-B and mature B cells

3.1 Introduction

Chromatin is a highly flexible structure that regulates gene expression through spatially and temporally coordinated changes between transcriptionally repressive/condensed states, and active/accessible states. Histone proteins, which are important structural components of this architecture, are now recognised for their central role in maintaining this dynamic equilibrium. Many studies have pointed to the flexible, N-terminal histone tails as targets of numerous post-translational modifications that can either activate or repress transcription, depending on the type of chemical modification and its location. This has led to the "histone code" hypothesis, which envisions that patterns of histone modifications could alter chromatin structure and promote recruitment of effector molecules recruitment. Recently, there are many indications pointing to histone modifications as important players in the maintenance of ES cell pluripotency and in the specification of their differentiation potential (Szutorisz et al., 2005; Vieira et al., 2004).

Some of the major questions in developmental biology concern how fate choices are made, what determines patterns of gene expression and how, in early precursor cells, information about activation, or silencing, at later stages are conserved through cycles of self renewal. The aim of this project was to study the chromatin structure of the *λ5-VpreB1* locus in different silenced states. The *λ5* and *VpreB1* genes are activated at the pro-B cell stage and are transcribed in pro- and pre-B cells before being silenced during the transition to the immature B cell stage (Dul et al., 1996; Karasuyama et al., 1994; Kudo and Melchers, 1987; Sakaguchi and Melchers, 1986). The tight stage-specific control that the genes are subjected to render this domain particularly suitable for the investigation of developmental silencing. Our goal was to characterise the silenced state of the locus in early precursor cells with different levels of plasticity and to assess whether the features of silenced states of the domain could be influenced by early fate decisions. For this purpose a chromatin

immunoprecipitation assay followed by real-time PCR analysis was used to analyse histone modifications linked with active and inactive chromatin across the domain.

3.2 Overview of ChIP

The most powerful tool that is available to investigate the in-vivo distribution of histone modifications across defined genomic regions is chromatin immunoprecipitation (ChIP). Technically there are two major variations in the ChIP protocol that differ in how the starting material (input chromatin) is prepared. Native chromatin immunoprecipitation (N-ChIP) uses chromatin prepared by micrococcal nuclease (MNase) digestion of cell nuclei. Cross linked chromatin immunoprecipitation (X-ChIP) uses chromatin cross-linked by the addition of formaldehyde to growing cells and fragmentation of the DNA by sonication. Histone proteins, which are tightly bound to the DNA, can be investigated both by X-ChIP and by N-ChIP. Non-histone proteins can only be analysed by X-ChIP since they are not retained on DNA during nuclease digestion. Both methods have advantages and disadvantages. N-ChIP allows a better control of the average size of the fragments generated. The length of the nucleosome arrays can be modulated by modifying the amount of MNase and the time of exposure. Resolution can be further increased purifying mono-nucleosome particles on a sucrose gradient. Control over the size of the generated fragments is very important since this represents a fundamental step to assure reproducibility of the experiments. However, it should also be noted that there is a risk of bias caused by selective nuclease digestion and over-digestion that can deplete specific regions from the input chromatin. Another disadvantage is that some of the nuclease-digested chromatin might be difficult to recover. Part of the material (in particular the bigger fragments) is trapped into the insoluble residual fraction left after the first extraction step. The amount of chromatin trapped in this fraction can be reduced incubating the nuclear pellet with hypotonic lysis buffer at 0 °C for half an hour followed by overnight freezing at -20°C. This procedure causes first the swelling of the nuclei and then their lysis, releasing most of the chromatin.

It has been shown that, *in vitro*, the histone octamer is capable of sliding along the DNA over significant distances in a temperature and salt dependent manner (Flaus et al., 1996; Hamiche et al., 2001; Meersseman et al., 1992). Therefore it is important to have tight control both of the temperature and of the salt content of the

solutions used to prevent chromatin rearrangement. One of the main advantages of N-ChIP is that most antisera are raised against un-fixed, synthetic peptide antigens and the epitopes they recognize may be masked or destroyed by formaldehyde treatment. However, proteins that bind weakly or indirectly to DNA can only be analysed by X-ChIP and formaldehyde treatment has been widely used in recent years.

There are obviously some issues that need to be addressed when using X-ChIP. One of the most important parameters is the time of exposure to the cross-linking agent. The most widely used time of cross-linking for the analysis of nucleosomal proteins is 10 min at a concentration of formaldehyde of 1% of the total volume (Kuo et al., 1996; Solomon et al., 1988). Longer exposure and higher concentrations of formaldehyde may leads to poor yield of immunoprecipitated (IP) material. Two potential problems should be considered; prolonged crosslinking might favour the binding of nucleosome associated proteins, thereby masking histone epitopes. It is also important to consider the chemistry of the reaction. Formaldehyde is a dipolar compound with a nucleophilic carbon atom that can react very quickly to form a Schiff base with the amino and imino groups of lysines, arginines and histidines, on amino acids, and adenines and cytosines on DNA. This means that some epitopes might be lost because of the engagement with formaldehyde reactive sites. Formaldehyde is also a moderate denaturing agent for proteins and it is known to interfere with secondary and, particularly, tertiary structures that could result in unfolding of the protein of interest.

Another parameter to be considered when adjusting fixation conditions is the susceptibility of the fixed material to mechanical shearing. Over-fixation leads to poor chromatin preparations since it becomes very difficult to sonicate the DNA to the desired length. Formaldehyde treatment of chromatin can also fix temporary superstructures such as loops (Fig. 3.2.1) and produce misleading positive results. Sequences not marked by the investigated modification can be immunoprecipitated because they become linked to the target sequence by crosslinked proteins.

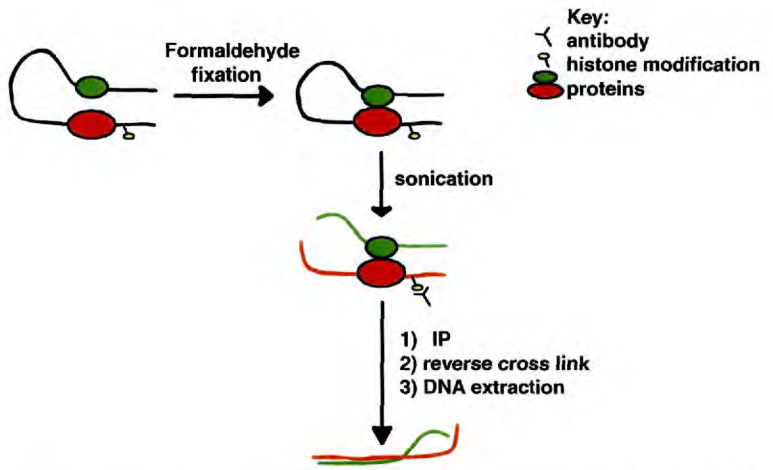


Fig. 3.2.1 Sequences not carrying any modification (green) can be crosslinked to the modified ones (red). This will cause them to be immunoprecipitated, creating a false positive result.

Another parameter that has to be considered is antibody specificity. ChIP grade antibodies purchased from Upstate LTD and Abcam LTD were tested for specificity by western blot and/or ELISA. To reduce the level of background in both fixed and native ChIP, respectively protein G and A beads were treated, prior to use, with BSA and salmon sperm DNA. Also, before elution the beads-immunoprecipitated sample complex was washed several times with buffers at increasing ionic strength. High salt concentration disrupts weak non-specific binding.

3.3 Data analysis

Immunoprecipitated material from X-ChIP and N-ChIP was analysed by real-time PCR and enrichment values were obtained from C(t) readings. Two different approaches were chosen depending on which experimental procedure was followed. When performing N-ChIP experiments levels of immunoprecipitated and input DNA were quantified by Pico Green (Molecular Probes) fluorescent dye and duplicate PCR reactions were carried out using 2ng of DNA per reaction. A typical output is shown in Fig. 3.3.1. The C(t) value was calculated from the intersection of the curve of interest with a manually set threshold (intensity of fluorescence corresponding to the beginning of the exponential phase of the PCR reaction).

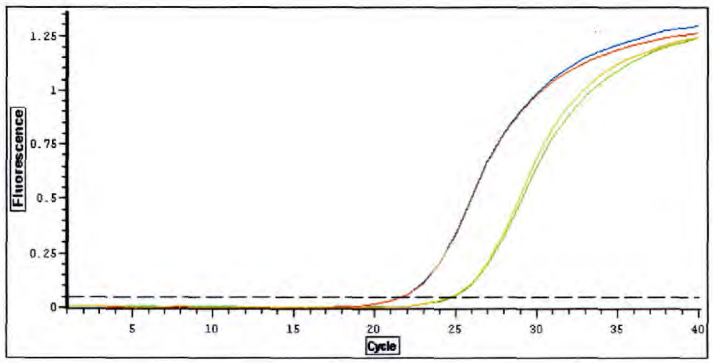


Fig.3.3.1 Typical output from a real time PCR experiments. Samples were run in duplicates, input is in green and yellow while the immunoprecipitated sample is in blue and red. The C(t) values relative to primer pair 83/84 (Topo3β transcribed region) are indicated below. PPB is abbreviation for primary pre-B cells. Immunoprecipitation was carried out with anti dimethyl lisine 4 histone H3 antibody (H3K4me2).

C(t)		C(t)	
(green) 83/84 PPB input	24.71757	(red) 83/84 PPB H3K4me2	21.67311
(yellow) 83/84 PPB input	24.78094	(blue) 83/84 PPB H3K4me2	21.59698

As described in section 2.12.2 of material and methods, the ratio between the fraction of target sequence in the immunoprecipitated sample and the fraction of target sequence in the input DNA (fold enrichment) can be calculated by the formula:

$$\mathbf{X_{0IP}/\ X_{0input} = 2^{[C(t)input - C(t)\ IP]}}$$

A fold enrichment equal to 20 would imply that in 2ng of immunoprecipitated samples there is 20 times more target sequence than in an equal amount of starting material.

When performing X-ChIP a different approach was chosen to analyse the data. After ethanol precipitation input and immunoprecipitated samples were dissolved in the appropriate volumes, DNA was not quantitated and PCRs were performed with 2µl of each sample. The percentage of the input immunoprecipitated with the specific antibody was calculated with an equation similar to the one described in sec 2.10.2.

$$\mathbf{X_{0IP}/\ X_{0input} = 2^{[C(t)input - C(t)\ IP]}}$$

A correction term was added to take into account the volumes in which input ($\mathbf{V_{Input}}$) and immunoprecipitated ($\mathbf{V_{IP}}$) samples were dissolved and the fact that only 1/10 of the total volume used for the specific immunoprecipitation was taken as input material (see section 2.11.4 matherial and methods).

$$\mathbf{X_{0IP}/\ X_{0Input}\,(\%) = \{[2^{(C(t)input - C(t)\ IP)}] \times [V_{IP}/10 \times V_{Input}]\} \times 100}$$

The values obtained represent, for each target sequence, the percentage of the input that has been immunoprecipitated.

The volume method allows detection of modifications that are widespread across the genome more efficiently than when equal amounts of DNA are used. When a modification is widely present in the genome the sequence profile of the IP sample will be similar to that of the input material and, therefore if equal quantities of DNA are analysed by PCR, the enrichment will not be not detectable. The volume method also allows the analysis of modifications that are present at very low level; very small percentages of immunoprecipitated input can be detected. Obviously when the percentage of immunoprecipitated sequence is very low, it is important to have good negative controls to validate the result. Appropriate controls include immunoprecipitations with non-specific antibodies or with no-antibody.

3.4 Primers

The first step that was necessary to investigate histone modifications in the *λ5-VpreB1* locus was to generate a set of primer pairs evenly spaced across the domain so that the distance between two adjacent primer sets would be less than 1kb. Primers were positioned in crucial regions of the domain such us hypersensitive sites and promoters (Fig. 3.4.1). A total of 60 primer pairs were tested from which a final set of 25 was chosen. Optimal design of the PCR primers was essential for accurate and specific quantification by real-time PCR. Detection is based on the binding of SYBR-Green dye into double stranded PCR products, a sequence-independent process. The sensitivity of detection using SYBR-Green may be compromised by the formation of primer-dimers, by lack of specificity of the primers and by the formation of secondary structures in the PCR product. These factors could lead to the creation of non-specific double-stranded DNA products, which would incorporate SYBR-Green and register a fluorescent signal. The primers were selected by Primer Select software according to the following criteria: 1) minimal propensity to form 3' annealed primer dimers, 2) ΔG value of the 3' end of the primers to be as low as possible to minimise the probability of mis-priming. The average length of the amplicons was maintained below 120 bp since longer PCR products could lead to non-linear incorporation of the fluorescent dye therefore to a decreased accuracy of the quantification.

Optimization of primer concentration was essential to avoid primer-dimer formation; each set described here worked best under different conditions, which were identified by testing combination of different concentrations for each forward and the reverse primer following the scheme showed in table 3.4.1.

Reverse Primer conc. (nM)	**Forward Primer conc. (nM)**			
	50 (nM)	300 (nM)	600 (nM)	900 (nM)
50 (nM)	50/50	300/50	600/50	900/50
300 (nM)	50/300	300/300	600/300	900/300
600 (nM)	50/600	300/600	600/600	900/600
900 (nM)	50/900	300/900	600/900	900/900

Tab. 3.4.1. Optimal operative concentration for the forward and reverse primer was investigated by testing all the possible combinations showed in this table.

Each primer pair was tested for all of the possible concentration combinations and the products were visualised on 1% agarose gels. The combinations giving the best yield were used in real time PCR experiment and the primer pairs showing the better signal/amplification efficiency ratio were chosen. Melting curves were also analysed to exclude non-specific amplification and primer-dimer formation in the presence of SYBR-Green. The position of the primers across the locus is shown in Fig. 3.4.1. The sequences, annealing temperatures and working concentrations are listed in table 2.12.1a section 2.12.1 matherials and methods.

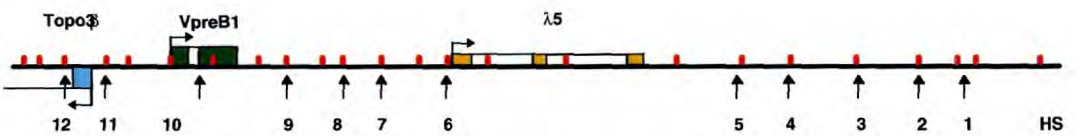


Fig. 3.4.1 Position of primers in of the λ5-VpreB1 locus. Primer positions are shown as red vertical lines. Black arrows represent HS sites. The λ5, VpreB1 and Topo3β genes are indicated by coloured boxes.

3.5 Analysis of histone H3 and H4 acetylation across the λ5-VpreB1 locus in Abelson transformed cell lines

Acetylation is thought to be able to activate transcription in two different ways, by direct inhibition of chromatin condensation and by generating a signal for the binding of general and/or specific transcription factors. The *λ5-VpreB1* locus is active in pro and pre-B cells and silenced at the immature B cell stage. Changes in the pattern of histone modifications during this transition could play an important role in the process of silencing the two genes, especially since positive regulators such as EBF and E2A are still present in immature and mature B cells. Histone H3 and H4 acetylation patterns across the locus were initially investigated in Abelson transformed pre-B (AB-1028) and in the immature B cell line (We-Hi 231). These cells have the advantage that they can be easily grown in large numbers providing sufficient material for chromatin preparation. Chromatin was extracted by MNase digestion and immunoprecipitated with antibodies against di-acetylated histone H3 and tetra-acetylated histone H4. In the pre-B cell line, histone H3 and H4 acetylation showed a pattern of peaks and troughs across the locus (Fig. 3.5.1 top panel). There were three main regions of H3 acetylation (red line), one of these was located between the *Topo3β* and *VpreB1* genes. A second region, which extended between *VpreB1* and *λ5* consisted of two peaks, the first located on HS 8 and the second upstream of the *λ5* promoter. The third region was located in the 3'end of the *λ5* gene. Histone H4 (blue line) presented a broadly similar pattern, with peaks between *Topo3β* and *VpreB1*, between *VpreB1* and *λ5*, and upstream of the *λ5* promoter. The 3' region of the locus, containinging HS 1 to HS5 has been shown to be required for full expression of both genes in transgenic experiments (Sabbattini et al., 1999). However, despites its functional importance, the region showed only low levels of H3 and H4 acetylation. This result suggests that functional elements of this type can exert their effects without necessarily being subject to active histone modifications.

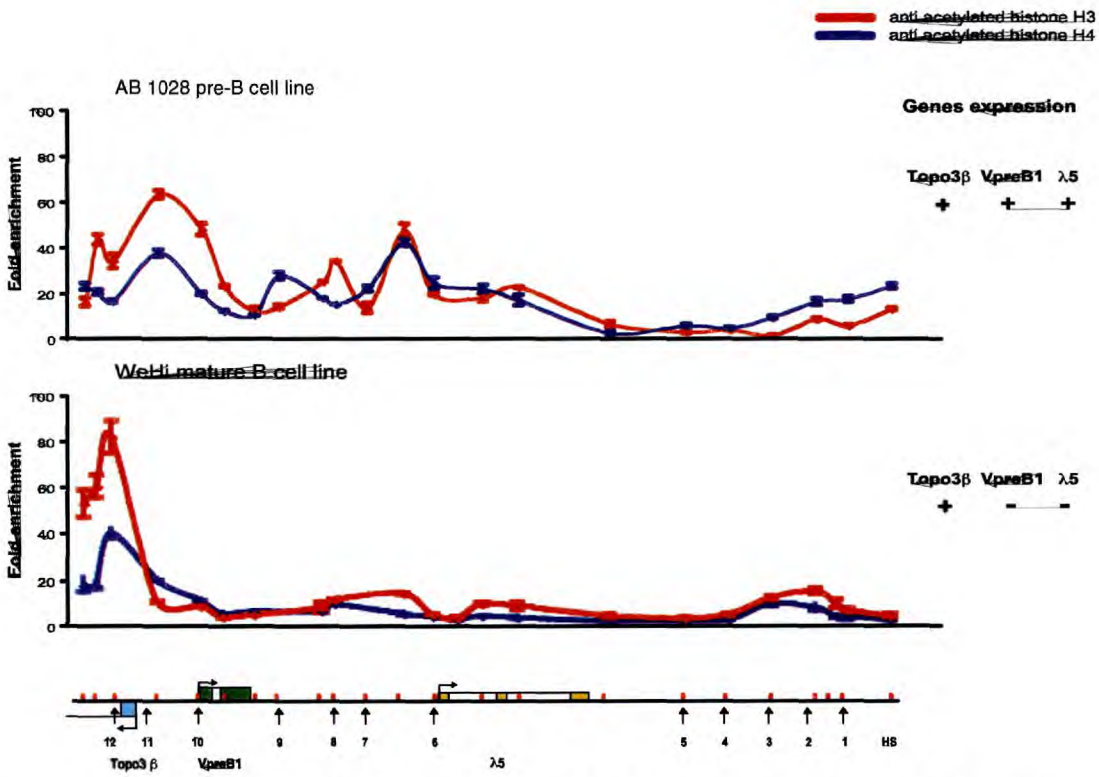


Fig. 3.5.1 Pattern of histone H3 (red) and histone H4 (blue) acetylation across the *λ5-VpreB1* locus in the pre-B cell line AB1028 and in the mature B cell line WeHi. Antibodies against anti acetylated histone H3 and anti acetylated histone H4 (both from Upstate) were used for the immunoprecipitations. On the *y* axis is shown the fold enrichment. The *x* axis represents the position of the amplicons across the *λ5-VpreB1* domain. Error bars indicate standard deviations relative to three different ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

The *λ5* and *VpreB1* genes and the neighbouring *Topo3β* gene are all expressed in pre-B cells. The promoter regions of these three expressed genes show a moderate level of acetylation but do not constitute peaks. Instead, peaks were located in the intergenic regions and at some of the HS. One explanation for the reduced degree of acetylation at promoters might be nucleosome depletion. Several independent studies (Bernstein et al., 2004; Liu et al., 2005; Sekinger et al., 2005) have observed nucleosome depletion at promoters in yeast suggesting that it might represent a fundamental feature of eukaryotic transcriptional regulation. However ChIP analysis of chromatin isolated from human cells with anti histone H3 antibody suggested that marked nucleosome depletion is not a feature of active mammalian chromatin (Bernstein et al., 2005). A recent study reported a correlation between nucleosome density in the vicinity of transcription start sites (TSS), histone H3 lysine 9 acetylation and gene expression level. The authors suggested that nucleosome

density is low in proximity to TSS at promoters that are significantly enriched in histone H3 K9 acetylation, regardless of their transcriptional activation status (Nishida et al., 2006).

Analysis of histone H3 and H4 acetylation was also performed on chromatin from the immature B cell line WeHi-231 (Fig. 3.5.1 bottom panel). H3 and H4 acetylation were enriched only on the *Topo3β* gene, which is expressed in mature B cells, and were largely absent from the rest of the locus.

The *λ5* and *VpreB1* genes are inactive in mature B cells, and the disappearance of histone modification marks suggest the formation of a repressive chromatin structure. Nevertheless the *λ5* promoter appeared to be undermethylated in mature B cells and still accessible to restriction enzymes (Sabbattini et al., 2001).

3.6 Analysis of histone H3 and H4 acetylation across the *λ5-VpreB1* locus in primary pre-B and mature B cells

A similar analysis of histone H3 and H4 acetylation and H3K4 methylation was conducted on primary pre-B cells and LPS activated primary B cells (Fig. 3.6.1). Primary pre-B and mature B cells cultures were established and characterised by FACS analysis as described in material and methods (section 2.10.2 and 2.10.3).

One of the main differences between the primary and transformed cells was that the levels of enrichment were much lower for both H3 and H4 acetylation for primary pre-B cells compared with the corresponding cell line. This was especially true for H4 acetylation. Expression studies have shown that *λ5* and *VpreB1* genes are slightly over expressed in the AB1028 cell line relative to primary pre-B cell cultures (Szutorisz, H., personal communication). Correlations between transcriptional activation and hyperacetylation have long been established (Braunstein et al., 1993; Hebbes et al., 1988; Jeppesen and Turner, 1993; Kuo et al., 1998), and particularly between hyperacetylation and gene reactivation (Mariani et al., 2003) and hyperacetylation and induction of transcription (Parekh and Maniatis, 1999). It is possible to speculate that over expression of the genes is correlated with their hyperacetylation in AB1028.

In primary pre-B cells H4 tetra-acetylation levels were generally lower (Fig. 3.6.1) and more uniform, with fewer peaks, compared to H3 di-acetylation levels observed both in primary pre-B cells and in the Abelson transformed cell line. The pattern observed across the locus suggests that H4 acetylation might not be involved in the establishment of an active λ5-VpreB1 domain in primary pre-B cells.

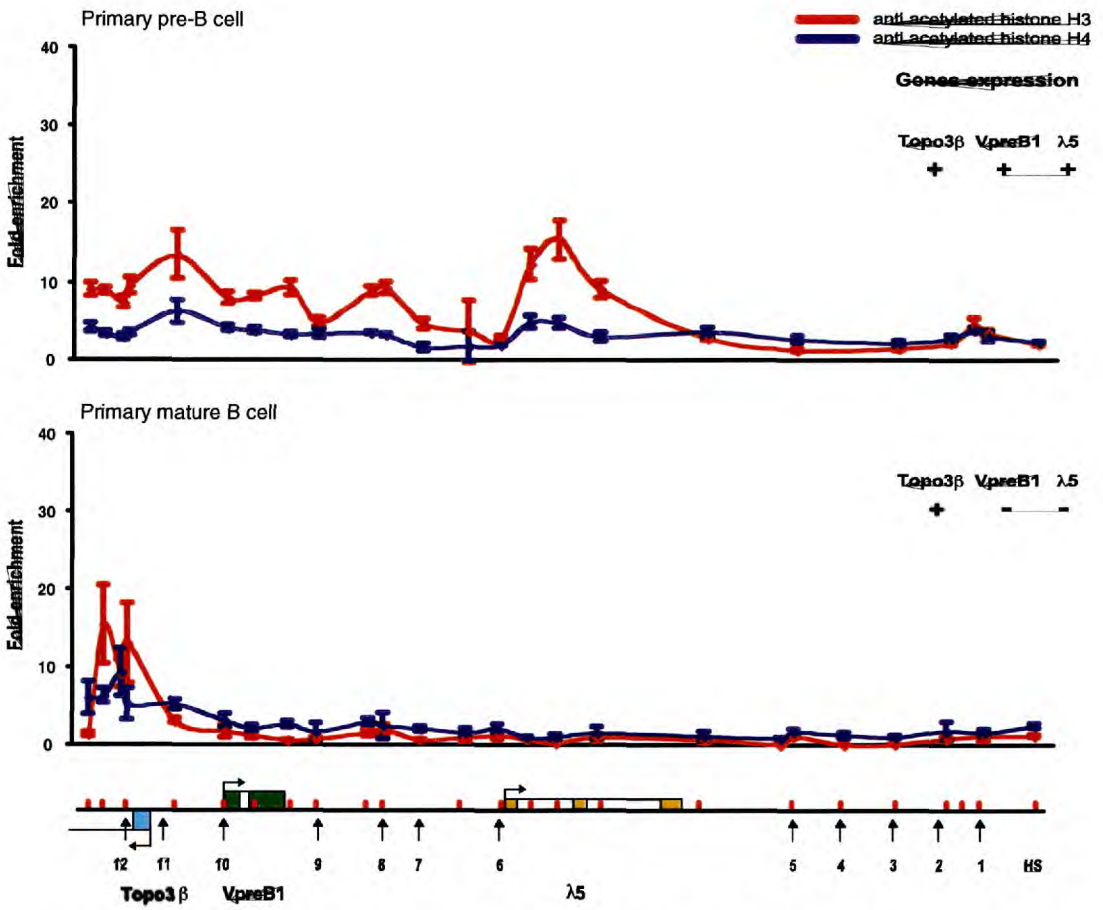


Fig. 3.6.1 Pattern of histone H3 (red) and histone H4 (blue) acetylation across the *λ5-VpreB1* locus in primary pre-B cells and in LPS activated mature B cell. Antibodies against anti acetylated histone H3 and anti acetylated histone H4 (both from Upstate) were used for the immunoprecipitations. On the *y* axis is shown the fold enrichment. The *x* axis represents the position of the amplicons across the *λ5-VpreB1* domain. Error bars indicate standard deviations relative to two different ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

In LPS-activated B cells, peaks of H3 and H4 acetylation were only observed on the Topo3β gene while the rest of the locus appeared depleted of both types of modifications. *Topo3β* is ubiquitously expressed (Szutorisz et al., 2005) and is transcribed on the opposite strand with respect to λ5 and VpreB1. Its TSS is only 1.5 Kb away from the VpreB1 start site. It is noteworthy that both the acetylation peaks

are shifted towards the 5'end of the *Topo3β* gene. Instead, in pre-B cells, where *VpreB1* is active, they span the region across the two adjacent genes. A very interesting feature is the sharp transition observed between the active *Topo3β* gene and the inactive *VpreB1*, which is underlined by a drop in the levels of histone modifications occurring approximately 600bp upstream of the *VpreB1* gene promoter. It is noteworthy that the transcription start sites of *Topo3β* and *VpreB1* are only 1.5kb apart.

These results suggest a tight controlled epigenetic determination of differentially expressed domains over a very short genomic distance. The short region comprised between the two genes does not contain any DNase I HS or intergenic promoter (Minaee et al., 2005) that can have an insulator function, which suggest that the separate histone modification profiles could be generated by the differential assembly of transcription factors complexes on their cis-regulatory elements.

The levels of enrichment on the Topo3β gene were lower than in the cell line suggesting an overall increment in gene expression in the transformed cells. Increased levels of expression have been observed for the Rag-1 and Rag-2 genes in Ableson transformed lymphoid cells (Rosenberg, 1994).

3.7 Pattern of histone H3 Lysine 4 dimethylation in primary pre-B and mature B cells.

Lysine residues within histones can be mono, di and tri-methylated. Different numbers of methyl groups on a given residue confer different biological readouts, which give methylation a greater combinatorial potential compared with other types of modifications. In contrast to acetylation, which correlates almost without exception with transcriptional activation, histone methylation can result in either transcription activation or repression, depending on which residues are modified. H3K4me2 correlates with a permissive state of chromatin, in which genes are either active or potentially active (Santos-Rosa et al., 2002). Other experiments in yeast have shown that the presence of H3K4me2 in the coding regions correlates mainly with transcriptional activity (Santos-Rosa et al., 2002).

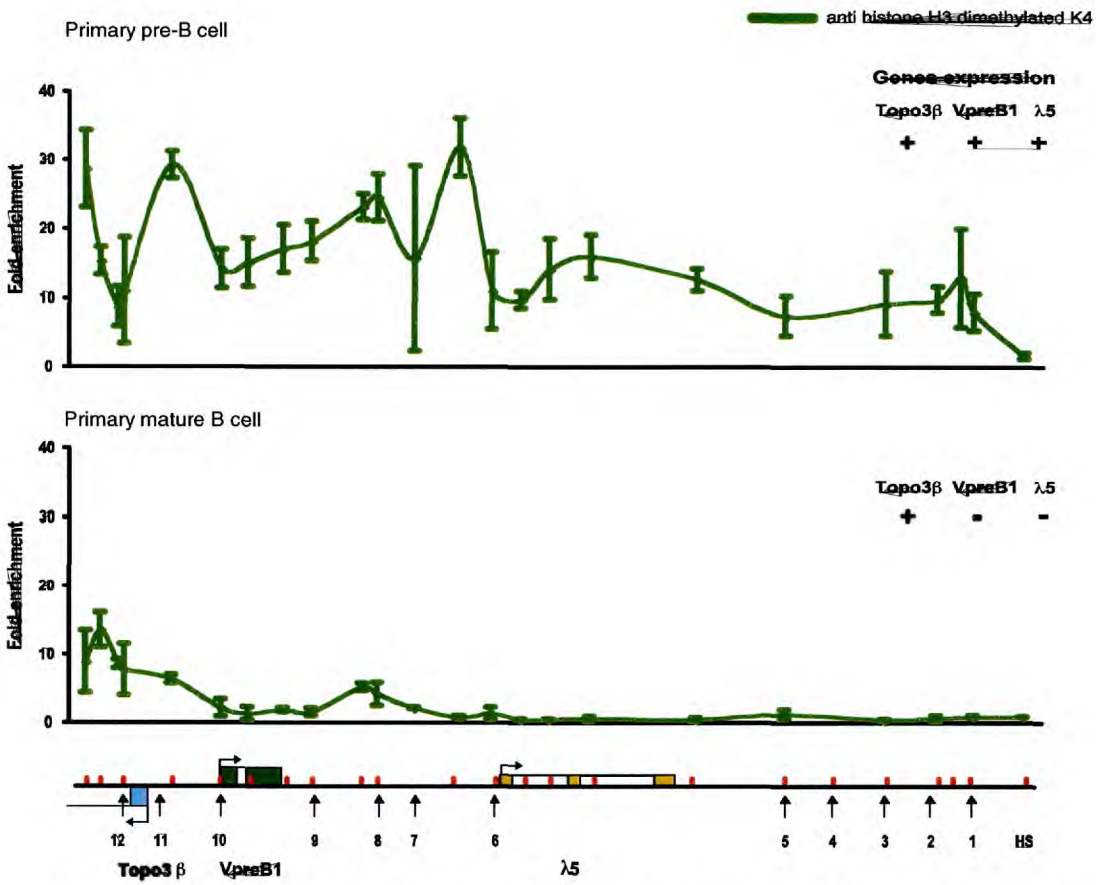


Fig. 3.7.1 Pattern of histone H3K4me2 (green) across the *λ5-VpreB1* locus in primary pre-B cells (top panel) and LPS activated mature B cell (bottom panel). An antibody against histone H3 dimethylated lysine 4 (Upstate) was used for the immunoprecipitations. On the *y* axis is shown the fold enrichment. The *x* axis represents the position of the amplicons across the *λ5-VpreB1* domain. Error bars indicate standard deviations relative to two different ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

In the *λ5-VpreB1* locus enrichment for H3K4me2 was spread across the entire domain (Fig. 3.7.1) in pre-B cells where the genes are active. Three highly enriched regions were found, one corresponding to a peak of H3 acetylation, between *Topo3β* and *VpreB1* and another located between the two pre-B cell specific genes. A broader region spanned the coding region of the *λ5* gene. In activated B cells the two broad domains on the *VpreB1* and *λ5* genes are lost and only the peak on the *Topo3β* is still present, even though the level of enrichment is quite reduced. It has been shown that H3K4me2 can be present in inactive genes (Cuthbert et al., 2004) but this does not seem to apply to the developmentally regulated *λ5* and *VpreB1*.

3.8 Summary

The active *λ5-VpreB1* locus in primary pre-B cells was found to be marked by the presence of three broad domains of H3K4me2, which spanned the *Topo3β* gene, the region between *VpreB1* and *λ5* and the *λ5* gene itself. Histone H3 acetylation was present as two major regions, one of these was located at HS8, and the other covered the length of the *λ5* gene. Histone H4 acetylation showed a low background level of enrichment across the entire domain and did not seem to participate in triggering the activation of the domain. Primary mature B cells showed peaks of the H3 and H4 acetylation and H3 lysine dimethylation only on the *Topo3β* gene while they were largely absent from the rest of the locus.

4. Inactive versus active chromatin modifications during B cell development

4.1 Introduction

A fertilised egg contains all of the genetic information required for the development of a new organism. During development the totipotent cells of the blastocyst will make fate choices that lead to the expression of some genes and to the silencing of others. Cellular identity is usually described in terms of expressed genes and most of the studies on gene regulation have focused on mechanisms of transcriptional activation. However, cell identity can also be described in terms of genes that are actively silenced. Gene silencing is known to be important for cell fate decisions, developmental progression and cell activation (Fisher, 2002; Fisher and Merkenschlager, 2002; Grewal and Elgin, 2002). One of the aims of this project was to investigate developmentally regulated silencing during the transition from the inner cell mass (ICM)/ES cell stage to multipotent and fully differentiated cells. There are least two types of situations where the *λ5-VpreB1* locus is silenced, one of these is established in mature B cells, as the result of stage specific silencing, during normal B cell development, while the other is observed in non lymphoid cells in which the locus has never been expressed. Investigation of the chromatin structure of the *λ5-VpreB1* locus led to the discovery of a small intergenic region marked by histone H3 acetylation and histone H3 lysine 4 dimethylation in ES cells (Szutorisz et al., 2005). This area, which was named early transcription competence mark (ETCM), is also a centre for recruitment of general transcription factors (GTFs) and Pol II in ES cells and contains an enhancer element that is active in pre-B cells. It has been proposed that the presence of a localized epigenetic mark in ES cells is important for establishing the transcriptional competence of the genes and that it could offer a solution to the problem of how ES cell plasticity is propagated during cell division (Szutorisz and Dillon, 2005). These results raise the question of whether active marks co-exist with repressive modifications and whether the silent, pluripotent state observed in ES cells is somehow different from the one observed in the fully committed B cell, after the genes have gone through both activation and silencing. To answer these questions analysis of histone H3 K9 dimethylation

(H3K9me2), histone H3 K9 trimethylation (H3K9me3) and of the double modification histone H3 K9 tri-methylation/S10 phosphorylation (H3K9me3S10ph) was performed at different stages of B cell development (Fig. 4.1.1). A number of housekeeping and tissue-specific gene promoters and transcribed regions were also analysed for comparison.

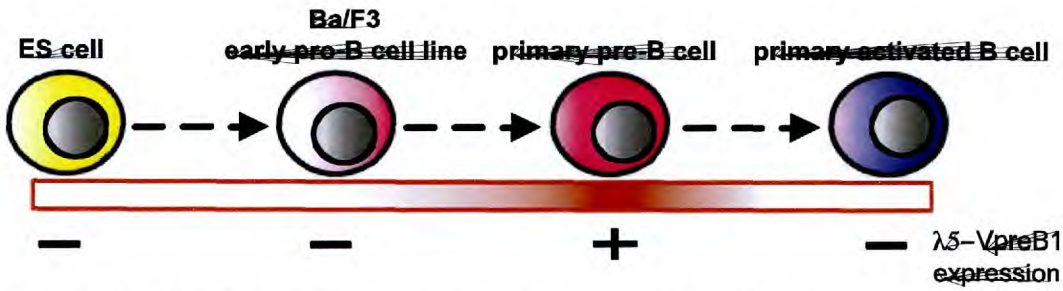


Fig. 4.1.1 Cells used to analyse epigenetic modifications during B cell development.

4.2 Patterns of histone H3K9me2 across the λ5-VpreB1 locus during B cell development

H3K9me2 has been associated with the silencing of euchromatic regions. There is evidence that the two HMTases, G9a and G9a related protein (GLP) responsible for the mono and di-methylation of lysine 9 (Tachibana et al., 2001; Tachibana et al., 2002), form complexes with HP1γ and a subset of E2F transcription factors, causing silencing of individual genes (Ogawa et al., 2002). These findings suggested that H3K9me2 could play a role in the silencing of the *λ5* and *VpreB1* genes. The pattern of the modification was investigated across the entire locus at different stages of B cell development. The results presented in this and the following chapters have been obtained by X-ChIP (see section 3.3). Attempts were also made to analyse H3K9me2 by N-ChIP but the outcome of the experiments was not satisfactory in terms of the yield of immunoprecipitated chromatin and of the fold enrichment at specific positions in the genome. The difficulties encountered in immunoprecipitating histone H3K9me2 with native chromatin preparations could be explained by the presence of an active demethylase. Histone methylation has long been considered to be a permanent modification. However, a number of histone demethylases have been identified recently. These include LSD1, which acts on H3K4me and H3K4me2 (Shi et al., 2004); JHDM1, which specifically demethylates

H3K36me2 (Tsukada et al., 2006) and lately JMJD2 which demethylates H3K9me3 and H3K936me3 (Whetstine et al., 2006).

In ES cells H3K9me2 was observed in the promoters and coding regions of the *λ5* and *VpreB1* genes, but not at the 5' end of the domain, where the active *Topo3β* gene is located (Fig. 4.2.1). It is also notable that the ETCM region (located in the intergenic region at HS7 and HS8) showed a decrease in H3K9me2 with the lowest level observed at HS8. A peak of H3K4me2 and H3 acetylation has been observed previously at this region in ES cells (Szutorisz et al., 2005).

The IL-3 dependent cell line Ba/F3 was used to analyse the pattern of histone H3 K9 dimethylation at the early pro-B cell stage. Ba/F3 cells are already committed to the B cell lineage and express B cell specific genes such us *mb-1*, *B29* and the germline Iμ. They do not express *λ5*, *VpreB1*, *Pax5*, *EBF*, *TdT*, *Rag1*, *CD19*, and *IL7R* (Sigvardsson et al., 1997). The pattern of H3K9me2 in Ba/F3 did not present any dramatic variation compared to what was observed in ES cells (Fig. 4.2.1).

In pre-B cells, where the two genes are fully active, H3K9me2 was observed at only one site, corresponding to primer pair 47/48 located on HS3. This has been identified as a repetitive LTR sequence that is also a site of DNA methylation in pre-B cells (see section 6.3). No specific immunoprecipitation was observed at any other site. Interestingly the background level in pre-B cells is higher than in the other cell types analysed. This feature, specific of the active *λ5-VpreB1* domain, could be due to a more open and accessible chromatin structure that is more accessible to antibody epitopes.

λ5 and *VpreB1* are repressed at the small resting pre-B cell stage, and are completely silenced in activated mature B cells. Surprisingly the pattern of H3K9me2 in mature B cells is quite similar to the one observed in ES cells. The modification is present in the genes promoters and coding regions, but is reduced the 3' end of the locus, where DNase I HS 1 to 5 are located. The level of H3K9me2 on the *λ5* gene is also slightly lower in mature B cells compared to ES cells.

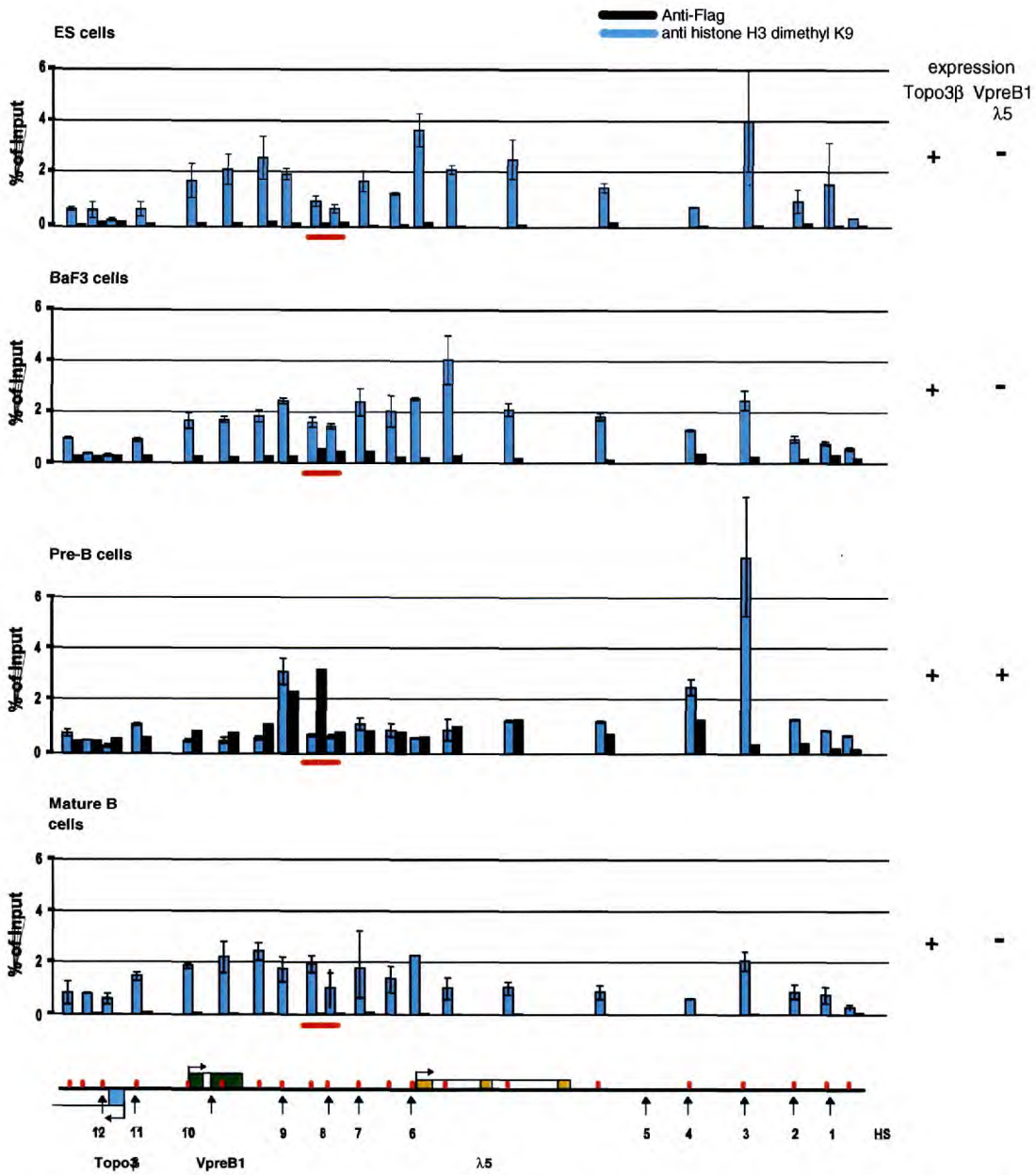


Fig. 4.2.1 Pattern of H3K9me2 (blue bars) across the λ*5-VpreB1* locus at different stages of B-cell development. An antibody against histone H3 dimethylated lysine 9 (Upstate) was used for the immunoprecipitations. A non-specific anti-Flag antibody was used to confirm that histone modification signals were detected above background (black bars). A horizontal red bar highlights the ETCM region. The percentage of immunoprecipitated input is shown on the *y* axis of each plot. The *x* axis represents positions across the λ*5-VpreB1* domain. The expression status of the genes at different stages of B-cell development is showed on the right side of each graph. Error bars indicate standard deviations calculated on minimum two ChIP experiment. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

These results show that, when silenced, the *λ5* and *VpreB1* genes are marked by histone H3 lysine 9 dimethylation on their promoters and coding regions. The ETCM region showed a reduced level of H3K9me2 that was especially pronounced in ES cells. It is possible to speculate that, in ES cells, recruitment of generic transcription factors and Pol II at the ETCM are necessary to prepare the locus for later activation while the presence of H3K9me2 at the gene promoters helps to keep them in a silenced state. In activated B cells silencing does not seem to require an increase in the level of H3K9me2 compared to ES cells. This suggest that the mechanism that triggers *λ5* and *VpreB1* silencing after activation in pre-B cells might involve the action of repressor molecules such as Ikaros (Sabbattini et al., 2001) and/or the presence of DNA methylation (chapter 6) rather than epigenetic mechanisms involving histone methylation.

4.3 Patterns of histone H3K9me3 and H3K9me3S10ph across the λ5-VpreB1 locus in ES and mature B cells

The pattern of H3K9me2 did not differ significantly in ES and mature B cells. In order to further investigate the silent states of the domain in ES and mature B cells other modifications associated with inactive chromatin were analysed. H3K9me3 has been mainly linked to heterochromatic regions and it is thought to provide a docking site for HP1 (Aagaard et al., 1999; Rea et al., 2000). Recently it has also been observed in the transcribed region of active genes (Vakoc et al., 2005). ChIP analysis of H3K9me3 showed extremely low levels across the entire *λ5-VpreB1* locus both in ES and activated B cells (Fig. 4.3.1 next page). This finding suggests that the *λ5-vpreB1* domain is in a quite open and accessible chromatin structure in both the cell types. This is in agreement with the fact that the *λ5* promoter was found to be undermethylated and accessible to restriction enzymes in activated B cells (Sabbattini et al., 2001).

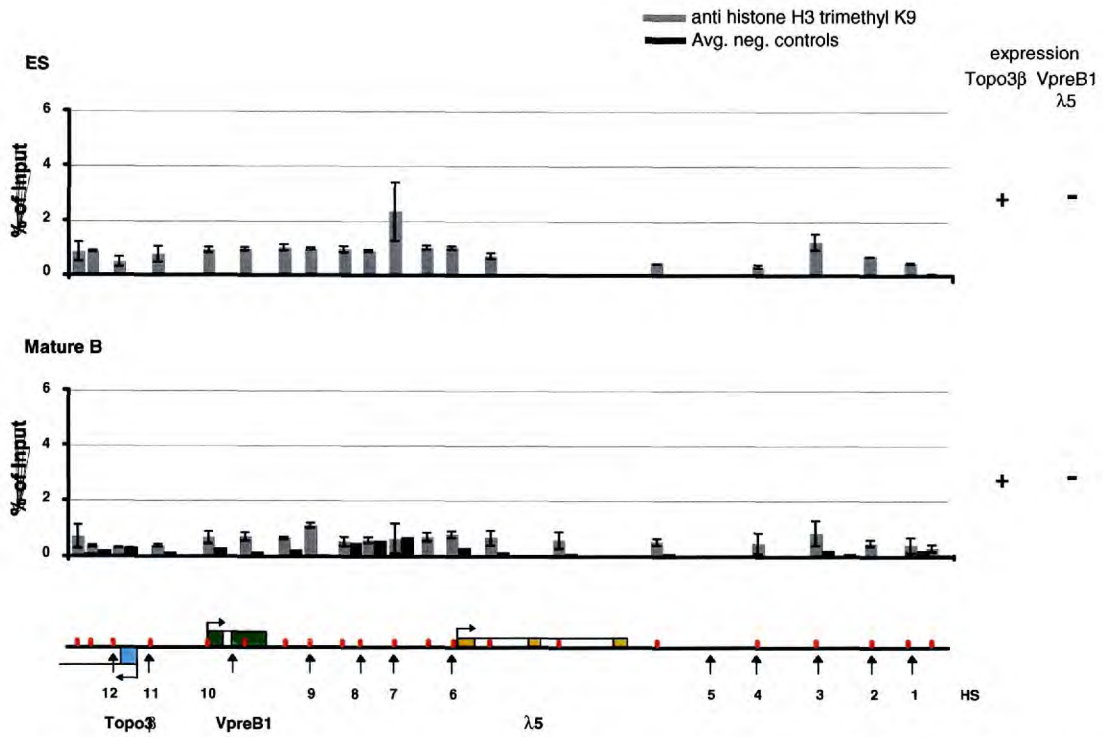


Fig. 4.3.1 Pattern of H3K9me3 across the λ*5-VpreB1* locus in ES (top panel) and mature B cells (bottom panel). An antibody against histone H3 trimethylated lysine 9 (Upstate) was used for the immunoprecipitations. Non-specific anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background (black bars). The percentage of immunoprecipitated input is shown on the *y* axis of each plot. The *x* axis represents positions across the λ*5-VpreB1* domain. The expression status of the genes is showed on the right side of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

The double trimethyl H3 K9 phospho S10 modification (H3K9me3S10ph) is known to be associated with condensation of mitotic chromosomes (Hirota et al., 2005) but its function is still not completely understood. Fischle et al, suggested a role for S10 phosphorylation in reducing the affinity of HP1 for the K9 methylation mark, thus allowing a dynamic control of the HP1-H3K9me3 interaction. Mateescu, B., et al (2004) instead proposed that association of HP1 with H3K9me3S10ph is more stable than its association with histone H3 tails that have only the methylation mark and that therefore the effect of the mitotic S10 phosphorylation is to tightly bind HP1 outside the heterochromatic foci promoting silencing. However, it has to be pointed out that different methyl groups and different HP1 isoforms were considered in the two studies. Recent unpublished results from Sabbattini, P. et al provide evidence that the double modification is associated with the formation of facultative heterochromatin in terminally differentiated cells.

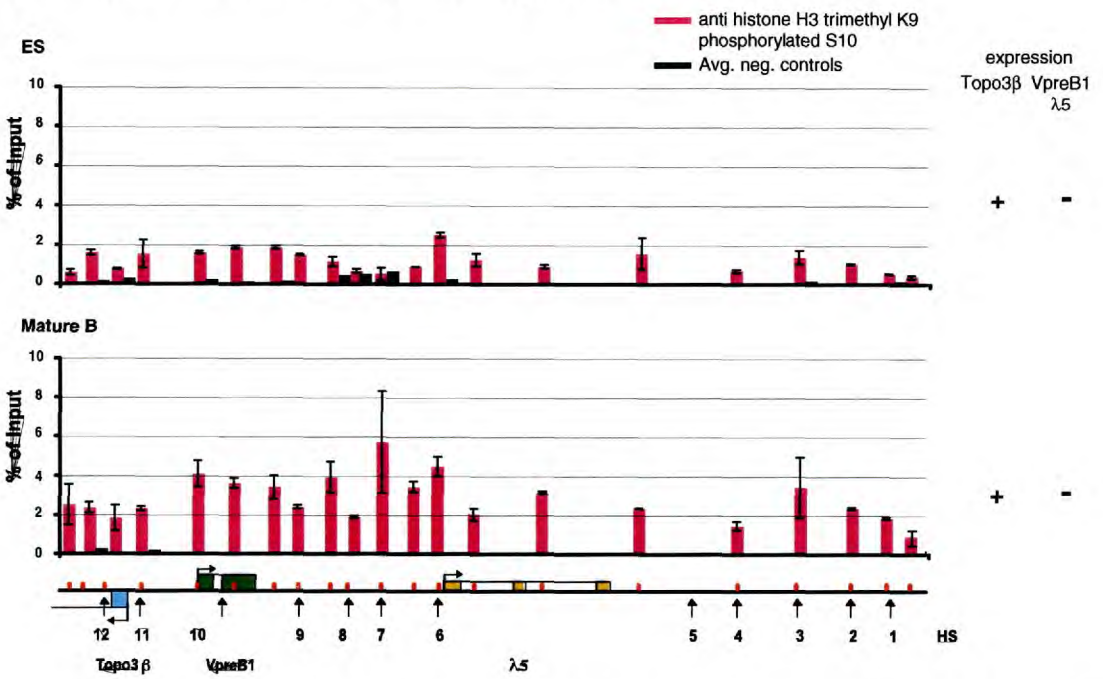


Fig. 4.3.2 Pattern of H3K9me3S10ph across the λ5-*VpreB1* locus in ES (top panel) and mature B cells (bottom panel). An antibody against histone H3 trimethylated lysine 9 phosporylated serine 10 (Abcam) was used for the immunoprecipitations. Non-specific anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background (black bars). The percentage of immunoprecipitated input is shown on the *y* axis of each plot. The *x* axis represents positions across the λ5-*VpreB1* domain. The expression status of the genes is showed on the right side of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

ChIP analysis of H3K9me3S10ph in ES cells showed a small enrichment compared to the negative controls across the all *λ5-VpreB1* locus (Fig. 4.3.2 top panel). In activated B cells (Fig. 4.3.2 bottom panel) the amount of immunoprecipitated input was higher and mainly homogeneous across the domain showing an increase (from 2 to 4%) in the *VpreB1* and *λ5* genes promoters and coding regions. Levels of H3K9me3S10ph in mature B cells showed to be higher than in ES cells; on average (3.0 ± 1.1)% versus (1.2 ± 0.6)%. This result raises the possibility that this modification is involved in determining the silenced state of the *λ5* and *VpreB1* genes in activated B cells.

4.4 Patterns of H3K9me2 and H3K4me2 in tissue-specific and ubiquitously expressed genes at different stages of B cell development

The analysis of histone modifications linked to inactive chromatin in the *λ5-VpreB1* locus during B cell development revealed few differences between the two silent states represented by ES and mature B cells. Depletion of H3K9me2 was observed at the ETCM region in ES cells, which was less pronounced in mature B cells. The *λ5* gene appeared more enriched in H3K9me2 in ES than in mature B cells. Levels of H3K9me3S10ph across the domain were higher in mature B cells than in ES cells, although this enrichment was generalised and not specific to regulatory regions of the locus. To further investigate the role of histone methylation in gene silencing and to generalise the results obtained from the analysis of the *λ5-VpreB1* locus, a set of tissue specific and housekeeping genes promoters and transcribed regions was analysed. Primers used for this analysis were described in materials and methods tab. 2.12.1b.

Current evidence suggests that active genes are enriched in H3K4me2 and depleted of H3K9me2, and that the contrary is true for inactive genes. In ES cells ubiquitously expressed genes followed this pattern while a number of tissue specific genes, such us *MyoD*, *VpreB1* and *λ5*, showed the simultaneous presence of the two modifications in the promoter and/or in the coding region (Fig. 4.4.1 ES cells, first panel). Other genes, such as *Neurofilament* (*Nfl*) and *Blimp1*, showed relatively high level of enrichment for H3K4me2 in their promoters. Hypermethylation of histone

H3 lysine 4 at *Nfl* promoter has been previously described in ES cells (Kimura et al., 2004) despite the gene not being expressed (Rohwedel et al., 1998). *Blimp1* is quite widely expressed in mouse and it has shown to be important for primordial germ cells (PGCs) specification (Ohinata et al., 2005) and, althought not expressed in ES cells (Ancelin et al., 2006) it has shown to be present at day 6.5 of embryonic development (Ohinata et al., 2005). Therefore it is possible to speculate that, in ES cells, H3K4me2 might mark the gene promoter in preparation for its activation early in embryonic development. Among the remaining tissue specific genes, *Amylase* showed the expected inactive/active modification pattern with high levels of H3K9me2 at the promoter and transcribed regions.

These results suggest that in ES cells the histone modifications associated with tissue specific genes are variable and do not conform to a single rigid pattern.

Commitment to the B cell pathway result in a much more defined pattern of histone modifications. In Ba/F3 early pro-B cells tissue-specific genes that are expressed outside the B lineage were all depleted for H3K4me2, both in the promoters and in the transcribed regions and were enriched in H3K9me2. Active genes presented an opposite pattern and were found to be enriched in H3K4me2 and devoid of HK9me2 (Fig. 4.4.1), Ba/F3, second panel). The *λ5* and *VpreB1* genes, which are not yet active at this developmental stage, showed an intermediate percentage of immunoprecipitation of both of these modifications. This result indicates that the two modifications are not mutually exclusive and that in the transition from an inactive to an active state the increase in the activating mark H3K4me2 happens before the inactivating one is completely erased.

In primary pre-B cells, the *λ5* and *VpreB1* genes are fully active, a status that is underlined by the high levels of H3K4me2 and the absence of H3K9me2. High levels of H3K4me2 marked all of the housekeeping genes examined, while silenced genes showed mainly the presence of H3K9me2. Silenced genes appeared to be more enriched in H3K9me2 in their transcribed regions than in their promoters (Fig.4.41 pre-B cells, third panel). Interestingly, depletion of lysine 9 dimethylation was observed at the silent *MyoD* and *Nfl* genes in pre-B cells and corresponded to enrichment in lysine 4 dimethylation.

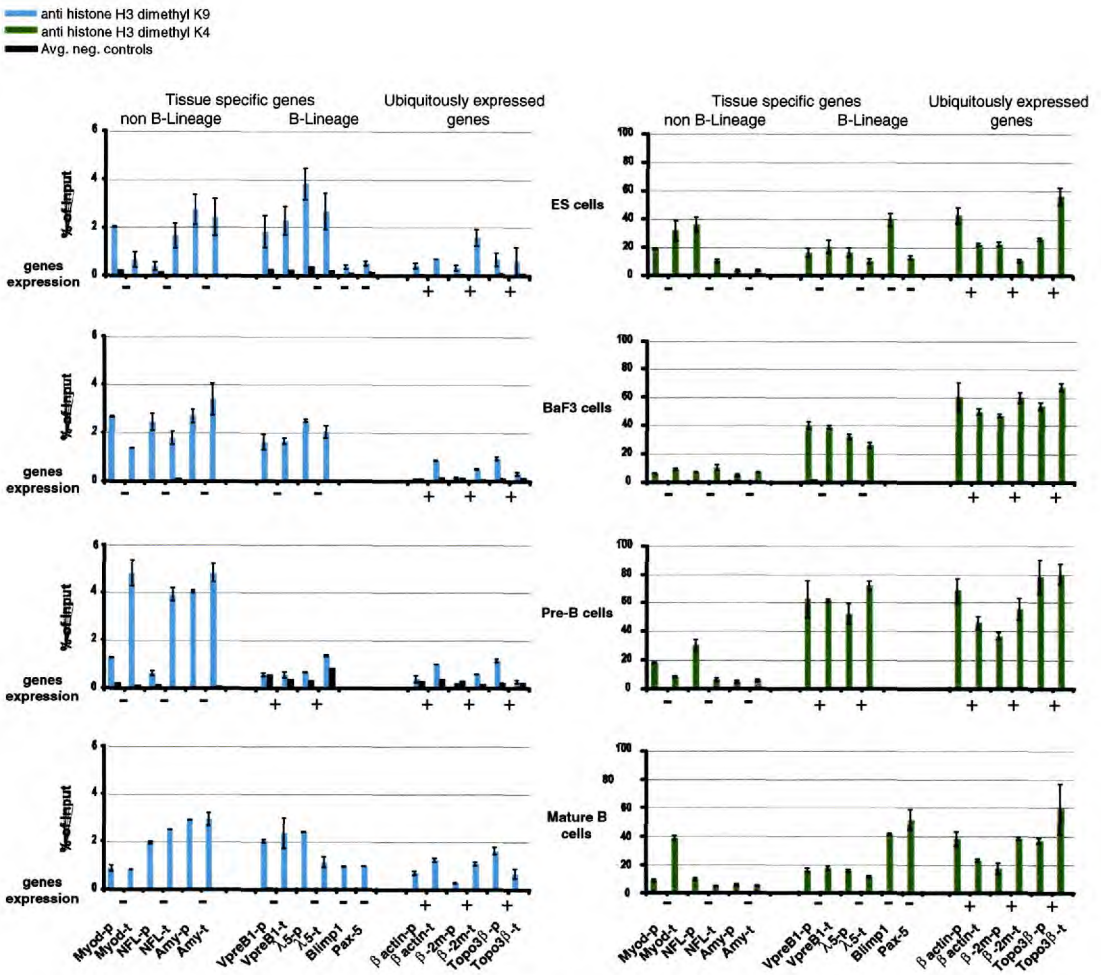


Fig. 4.4.1 Patterns of H3K9me2 (blue bars) and H3K4me2 (green bars) in tissue specific and housekeeping genes. Antibodies against histone H3 dimethylated lysine 9 and histone H3 dimethylated lysine 4 were used for the immunoprecipitations (both Upstate). Anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background. Values on the Y axis represents the percentage of immunoprecipitated input, on the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. P and t indicate respectively promoter and transcribed regions.

In mature B cells, among the tissue specific genes that are not expessed in the B lineage, the two modifications appear to mirror each other; any enrichment in H3K4me2 corresponded to a decrease in H3K9me2 and vice versa (Fig. 4.4.1 mature-B cells, fourth panel). The only exception was the *MyoD* gene, which showed a peak of H3K4me2 in the transcribed region. In contrast, genes active at this stage of B cell development such as *Blimp1* and *Pax5* presented high levels of H3K4me2 in their promoters (Fig 4.4.1, mature-B cells, fourth panel).

These results suggest that in ES cells lineage specific genes can have a variety of histone modification profiles. H3K9me2 and H3K4me2 were either both present at the same sites (*MyoD, λ5, VpreB1*) or alternating between the promoter and transcribed region of the gene (*NFL, β2m*). ES cells are able to differentiate into all lineages, a property that can be explained by the ability of these cells to keep lineage specific genes in a "poised" state from which it would be easier to progress into an active or inactive state. The results described in this section suggest that tissue-specific genes have a mixture of modifications associated with their promoters or coding regions in ES cells and are kept in a hybrid condition until a fate decision in taken and there is a need to silence or express them. This is confirmed by the observations made in Ba/F3 cells, since once the B cell pathway is engaged, the genes present a more defined pattern of modifications.

4.5 Patterns of histone H3K9me3 and H3K9me3S10ph at tissue-specific and ubiquitously expressed genes in ES and mature B cells

The results described in the previous section show that non-lymphoid tissue specific genes have a more defined pattern of H3K9me2 in mature B cells compared with ES cells (Fig. 4.4.1). These observations were extended by ChIP analysis of H3K9me3. Tissue specific genes that are not expressed in the B lineage were enriched in H3K9me3 in mature B cells but not in ES cells (Fig.4.5.1). *Myod* represented an exception and was found enriched in K4me2 (Fig. 4.4.1) and depleted in H3K9me3 similar to the pattern that was observed for H3K9me2 (Fig. 4.5.1) in its transcribed region. In contrast, B lineage-specific genes, even if inactive, were not marked by H3K9me3 in mature B cells. These results suggest that genes that undergo long term silencing during cell differentiation are marked by H3K9me3, and are probably heterochromatinised in fully differentiated cells. In ES cells, the chromatin of tissue specific genes seems to adopt a permissive conformation with reduced levels of heterochromatic modifications.

Housekeeping genes such as *β-actin* and *β2-m* showed high values of H3K9me3 in their transcribed regions both in ES and mature B cells. The presence of H3K9me3 in transcribed regions of active genes has been described previously (Vakoc et al., 2005) and it has been proposed, that this modification acts in

conjunction with the presence of HP1γ and RNA pol II to create a pattern that identifies transcribed regions in the genome.

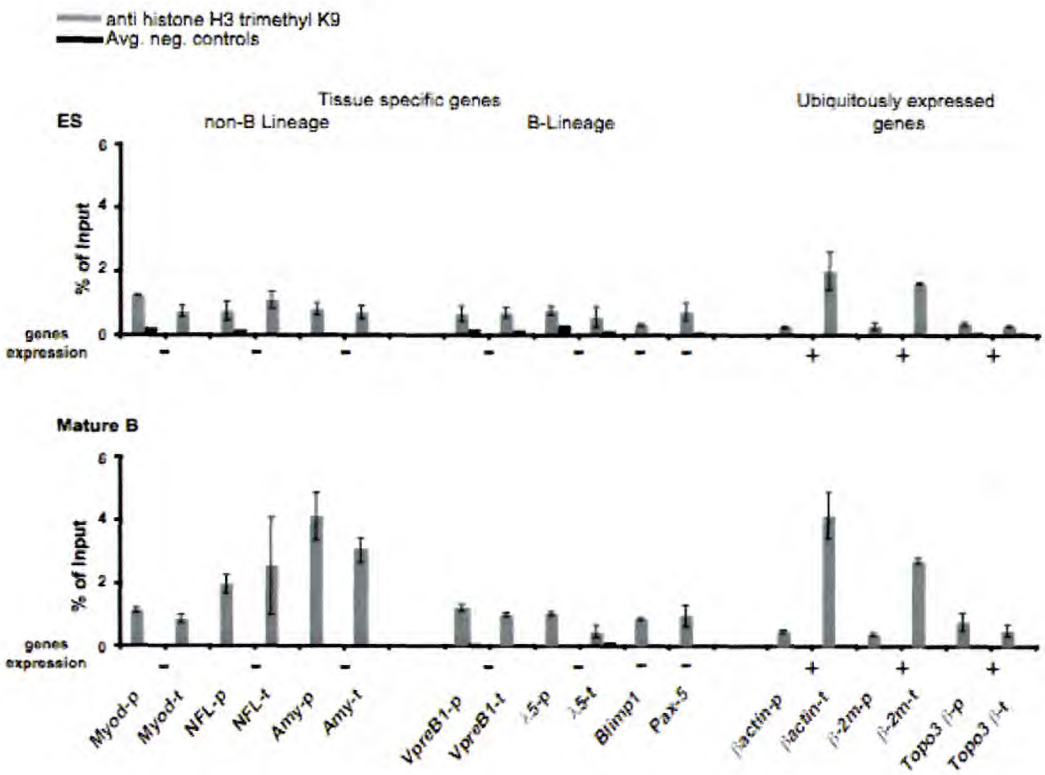


Fig. 4.5.1 Patterns of H3K9me3 (grey bars) in ES and mature B cells in tissue-specific and housekeeping genes. An antibodys against histone H3 trimethylated lysine 9 was used for the immunoprecipitations (Upstate). Anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background. Values on the y axis represents the percentage of immunoprecipitated input, on the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. P and t indicate respectively promoter and transcribed regions.

In mature B cells H3K9me3S10ph appeared enriched in all the tissue specific genes, both in their promoters and transcribed regions. Housekeeping genes, such as *βactin* and *β2-m*, showed instead a pattern similar to H3K9me3, with a higher level of enrichment in the transcribed region (Fig. 4.5.2). The amount of H3K9me3S10ph increased progressively from the housekeeping to the non-lymphoid lineage-specific genes suggesting that the presence of the double modification could be correlated to both long term and stage specific silencing. Nevertheless some enrichment was observed on *Blimp1* and *Pax5* promoters, which are expressed in mature B cells. In these two genes the level of H3K9me3S10ph is actually comparable to the level observed at *λ5* and *VpreB1* promoters, which are silent at this stage of development. In ES cells a background level of enrichment (1.5 ± 0.5)% was observed at most gene promoters and transcribed regions.

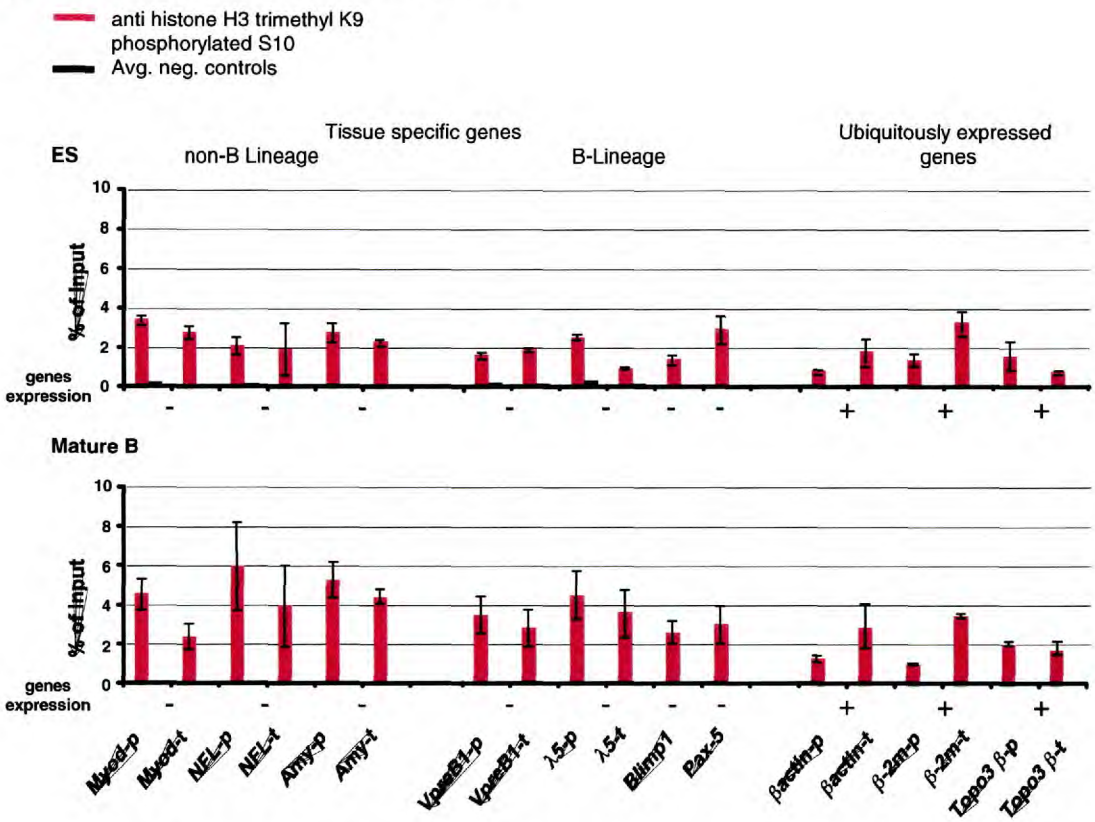


Fig. 4.5.2 Patterns of H3K4me3S10ph (pink bars) in ES cells (top panel) and mature B cells (bottom panel) in tissue-specific and housekeeping genes. An antibodys against histone H3 trimethylated lysine 9 phosphorylated S10 was used for the immunoprecipitations (Abcam). Anti-Flag, anti IgG and no antibody control were as negative controls. Values on the y axis represents the percentage of immunoprecipitated input, on the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. P and t indicate respectively promoter and transcribed regions.

4.6 Summary

The silent state of the *λ5* and *VpreB1* genes in ES and mature B cells was found to be marked by H3K9me2. Enrichment in the dimetylation mark was detected at the genes promoters and transcribed regions. In ES cells only a decrease in H3K9me2 was observed at HS8, a site that had previously shown to coincide with a peak of H3K4me2 (Szutorisz et al., 2005). Analysis of a set of tissue-specific genes in ES cells showed a varied pattern of histone H3 dimethylation at K9 and K4. The two modifications were often present at the same sites (i.e. *MyoD*, *λ5*, *VpreB1*) or alternating between the gene promoter and transcribed region (i.e. *NFL*, *β2m*). Histone H3 K9 trimethylation was instead largely absent both from the locus and from the lineage-specific genes analysed. This result suggested an open, permissive, chromatin conformation. In mature B cells, tissue-specific non-lymphoid genes were found enriched in both H3K9me2 and H3K9me3. The double modification H3K9me3S10ph presented a low, background, level in ES cells and was slightly enriched in mature B cells.

5. Epigenetic modifications of tissue-specific genes in the C3H10T1/2 mesenchymal stem cell line

5.1 Introduction

Recent findings suggest the possibility that genes might be marked in ES cells by chromatin modifications that are associated with gene activation. These localised epigenetic marks would keep the genes in a potentiated state that would render them ready for expression at later stage of development (Szutorisz et al., 2005; Vieira et al., 2004). According to this hypothesis when the cells commit to a specific lineage tissue-specific genes that are not expressed in that lineage would lose their marks and be silenced. Genes that need to be activated will acquire transcriptional competence through recruitment of specific transcription factors and expansion of the active epigenetic marks (Szutorisz and Dillon, 2005).

One of the aims of this project was to investigate how early cell fate decisions and differentiation can affect the chromatin structure of the *λ5-VpreB1* locus. The C3H10T1/2 mesenchymal stem cell line was chosen as a model system to analyse the features of the silent *λ5-VpreB1* domain in a multipotent progenitor with a restricted differentiation potential (which does not include the B cell pathway) and in its product of differentiation (Fig. 5.1.1). C3H10T1/2 cells have been used extensively as an in vitro model to examine mesenchymal differentiation into various lineages (Ahrens et al., 1993; Denker et al., 1999; Katagiri et al., 1990). The *λ5-VpreB1* locus has never been active and will never be activated in these cells even upon differentiation.

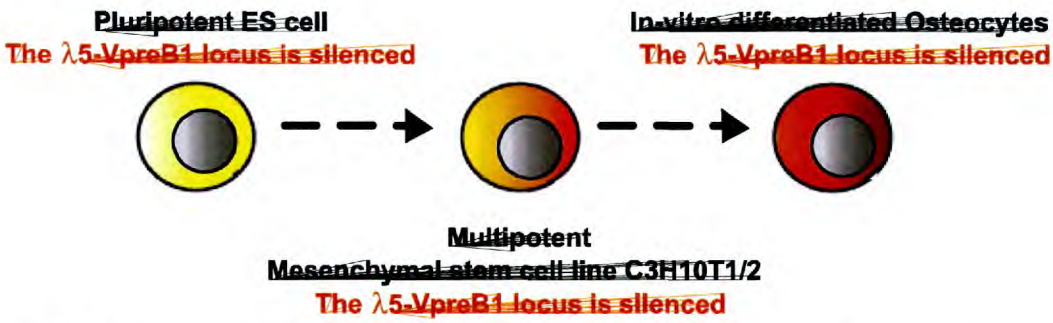


Fig. 5.1.1 model system used to analyse epigenetic modifications during cell differentiation.

5.2 C3H10T1/2 differentiation

Differentiation of the mesenchymal stem cell line C3H10T1/2 (CHT) towards osteocytes and adypocites was performed as described in Peister A., et al., (Blood, 2004) with minor modifications (see materials and methods and Fig. 5.2.1 and 5.2.2). Adipogenesis was partially effective since only 30-40% of the cells gave bright red staining with Oil Red and assumed the characteristic granular phenotype. Cells incubated with the osteogenesis medium showed instead a complete change in morphology during the first week of treatment and after three weeks 70%-80% gave a positive stain with alkaline phosphatase. Counting of the number of cells stained was difficult since the alkaline phosphatase coloured reaction could not be detected after staining of the nuclei with GIEMSA. Therefore the number of cells positive for alkaline phosphatase was determined as follows. Two 175 cm^2 flask containing osteocytes originating from the same differentiation experiment were analysed. One of the flasks was used for GIEMSA staining and counting of the nuclei. Cells in the other flask were treated with the alkaline phosphates reagent and those giving positive staining counted. Comparison of the number of cells positive for GIEMSA and alkaline phosphatase staining showed that an average of 70-80% of cells that could be considered to have successfully differentiated.

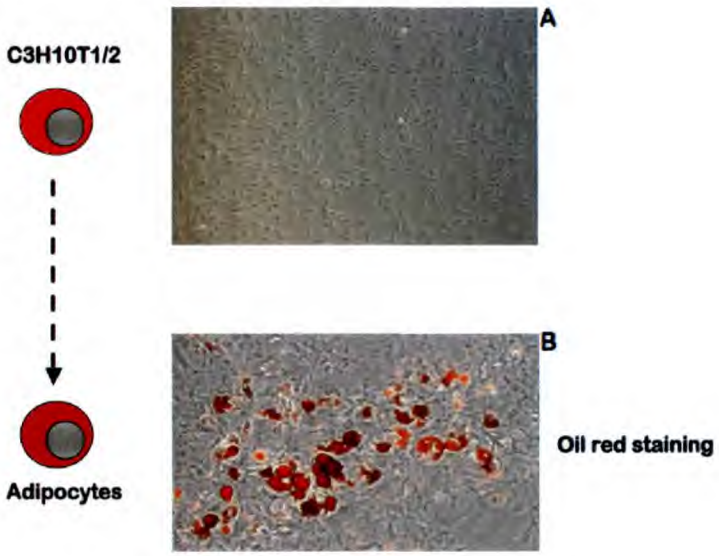


Fig.5.2.1 Differentiation of the mesenchymal stem cell line C3H10T1/2 into osteocytes. Cells were grown to 70% confluency (A) and incubated for three weeks with adipocytes differentiation media (B). Differentiation was assessed by Oil red staining (C).

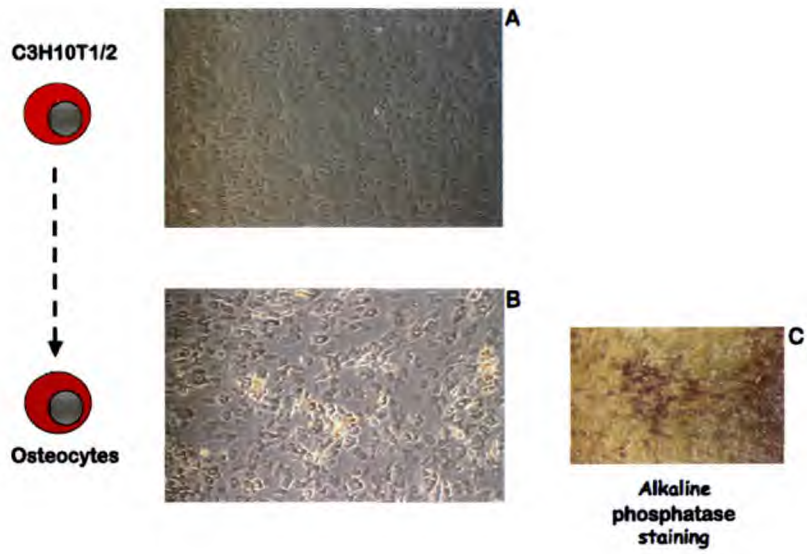


Fig.5.2.2 Differentiation of the mesenchymal stem cell line C3H10T1/2 into osteocytes. Cells were grown to 70% confluency (A) and incubated for three weeks with osteocytes differentiation media (B). Differentiation was assessed by alkaline phosphatase staining (C).

5.3 Patterns of histone H3K4me2 and histone H3 di-acetylation across the *λ5-VpreB1* locus during C3H10T1/2 cell differentiation

Histone H3 acetylation and H3K4me2 were analysed across the *λ5-VpreB1* locus in multipotent C3H10T1/2 cells (CHT) and in osteocytes (CHTD) obtained by differentiation of these cells. The aim of this analysis was to investigate the effects of long term silencing on the *λ5-VpreB1* locus. The pattern of H3 acetylation (red bars in Fig. 5.3.1, top panel) in the mesenchymal stem cell line showed a strong peak on the *Topo3β* gene, and a smaller one on HS8. A peak was also observed on HS3 (Fig. 5.3.1 right side, top panel), but the error bars indicated a high variability in the ChIP analysis of this specific site that raises questions about the significance of this particular result. The region containing the pre-B cell-specific hypersensitive sites 7 and 8 showed a rise in H3K4me2 (green bars in Fig. 5.3.1, top panel). Upon differentiation along the osteogenic pathway the pattern of histone modifications did not change dramatically. The small peak of H3 acetylation on HS8 disappeared, while the overall level remained constant (Fig.5.3.1 bottom panel). The locus showed a generalised low level of enrichment of H3K4me2 although this was significantly lower then the level observed at the active *Topo3β* promoter. From this analysis it can be concluded that the region marked by HS7 and HS8 presents a slight enrichment in H3 acetylation and H3K4me2, but it is not possible to discuss this result in relation to the already published data on ES cells since different chromatin preparation methods have been used. The experiment showed in this thesis have been performed with X-ChIP, while results obtained by Szutorisz et al were obtained with N-ChIP (Szutorisz and Dillon, 2005).

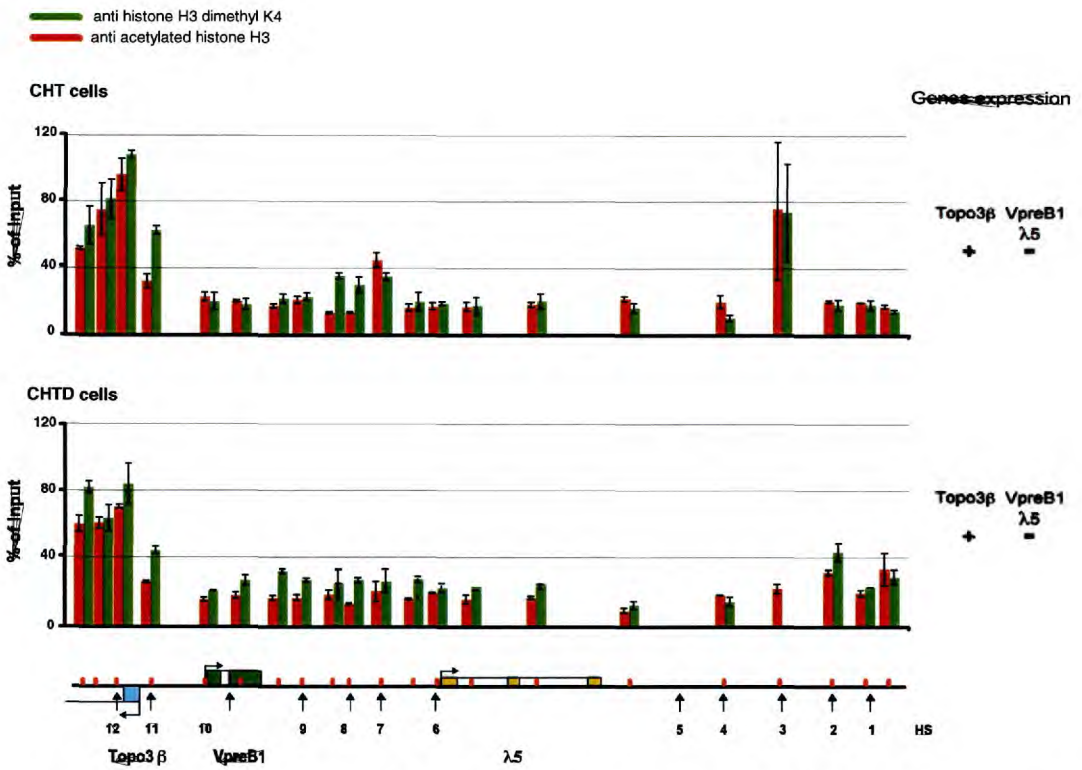


Fig. 5.3.1 Patterns of histone H3 acetylation (red) and histone H3K4me2 (green) across the λ*5-VpreB1* locus in the multipotent mesenchymal stem cell line CHT (top panel) and in its product of differentiation along the osteogenic pathway CHTD (bottom panel). Antibodies against acetylated histone H3 and histone H3 dimethylated K4 (both from Upstate) were used for the immunoprecipitations. The percentage of immunoprecipitated input is shown on the *y* axis. The *x* axis represents the position of the amplicons across the λ*5-VpreB1* domain Error bars indicate standard deviations relative to two different ChIP experiments. The locus map is in scale; red vertical lines indicate the positions of the amplicons, arrows indicate DNAse I HS. Negative controls (not shown) were calculated as the average of the values obtained with anti-flag, anti-IgG and no antibody and were always below 1% of IP input.

5.4. Patterns of histone H3 lysine 9 methylation in the λ5-VpreB1 locus in undifferentiated and differentiated C3H10T1/2 cells

Analysis of the *λ5-VpreB1* locus in undifferentiated and differentiated C3H10T1/2 cells showed that modifications characteristic of active chromatin are localised to the *Topo3β* gene and are largely absent from the rest of the locus. To investigate the epigenetic signature of this developmentally regulated domain when it is undergoing long term silencing, methylation of histone H3 at lysine 9 was investigated. Immunoprecipitations were performed with antibodies against di- and

trimethylated lysine 9 and against the double trimethylated lysine 9 phosphorylated serine 10 modification.

5.4.1 Histone H3 lysine 9 dimethylation

The pattern of H3K9me2 observed in the undifferentiated cells (Fig. 5.4.1) resembled that observed in ES cells, with the highest levels of enrichment located on the genes promoters and coding regions. In the differentiated osteocytes a dramatic increase in H3K9me2 was observed across the entire domain (Fig. 5.4.1, bottom panel). The enrichment was particularly high on the gene promoters and coding regions.

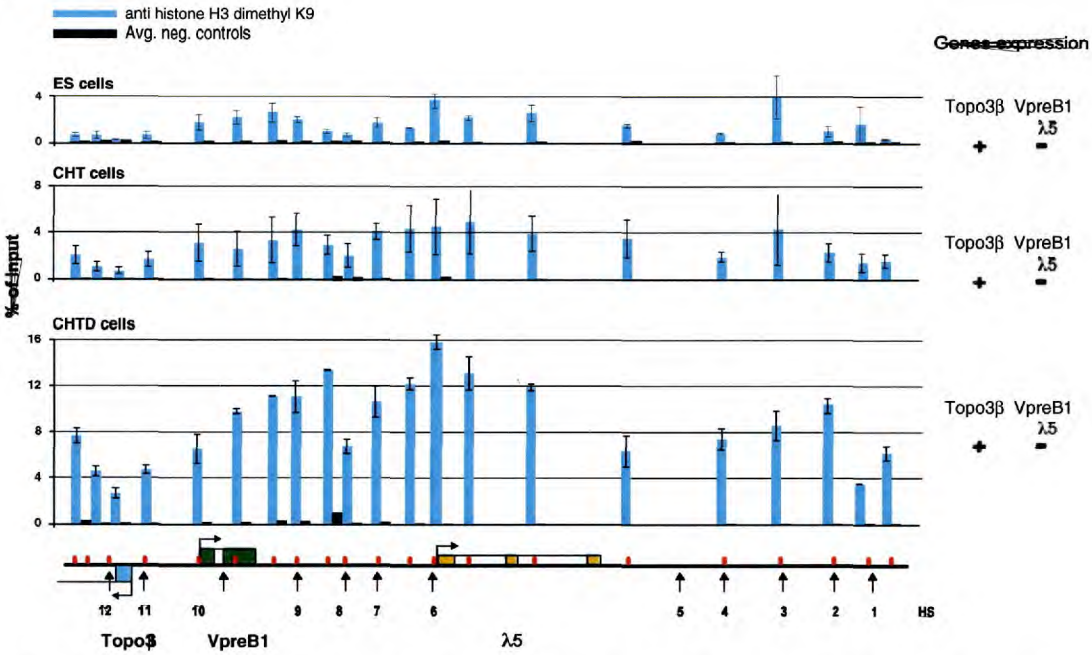


Fig. 5.4.1 Patterns of histone H3K9me2 (blue) across the *λ5-VpreB1* locus in ES cells, in the multipotent mesenchymal stem cell line C3H10T1/2 (CHT) and in its product of differentiation along the osteogenic pathway (CHTD bottom panel). An antibody against histone H3 dimethylated K9 (Upstate) was used for the immunoprecipitations. Negative controls were calculated as the average of the values obtained with anti-flag, anti-IgG and no antibody. The percentage of immunoprecipitated input is shown on the *y* axis. The *x* axis represents the position of the amplicons across the *λ5-VpreB1* domain Error bars indicate standard deviations relative to at least two different ChIP experiments. The locus map is in scale; red vertical lines indicate the positions of the amplicons, arrows indicate DNAse I HS.

Calculation of the average value of enrichment across the *λ5-VpreB1* locus confirmed that the level of H3K9me2 increased with the restriction of the differentiation potential. ES cells showed an average value of (1.6 ± 0.5) % while the value for the mesenchymal stem cell line was (3.3 ± 1.4) %. Mature B cells gave

a value of (1.4 ± 0.7) % of immunoprecipitation, which is very similar to the level in ES cells. This result when compared with the (8.5 ± 0.6) % observed in the differentiated osteocytes suggests a more prominent role for H3K9me2 in long term silencing than in stage-specific silencing of the *λ5-VpreB1* locus.

5.4.2 Histone H3 K9 trimethylation and Histone H3 K9 trimethylation / S10 phosphorylation

The presence of H3K9me3 is considered a hallmark of heterochromatin formation. The *λ5-VpreB1* locus was depleted of the modification in ES cells, showing that prior to activation, the domain is not in a heterochromatic configuration. Analysis of H3K9me3 in undifferentiated C3H10T1/2 cells revealed again a very similar pattern to the one observed in ES cells (Fig.5.4.2, top panel), suggesting that although λ5 and VpreB1 are on their way being permanent silenced, they are most probably not packaged into heterochromatin.

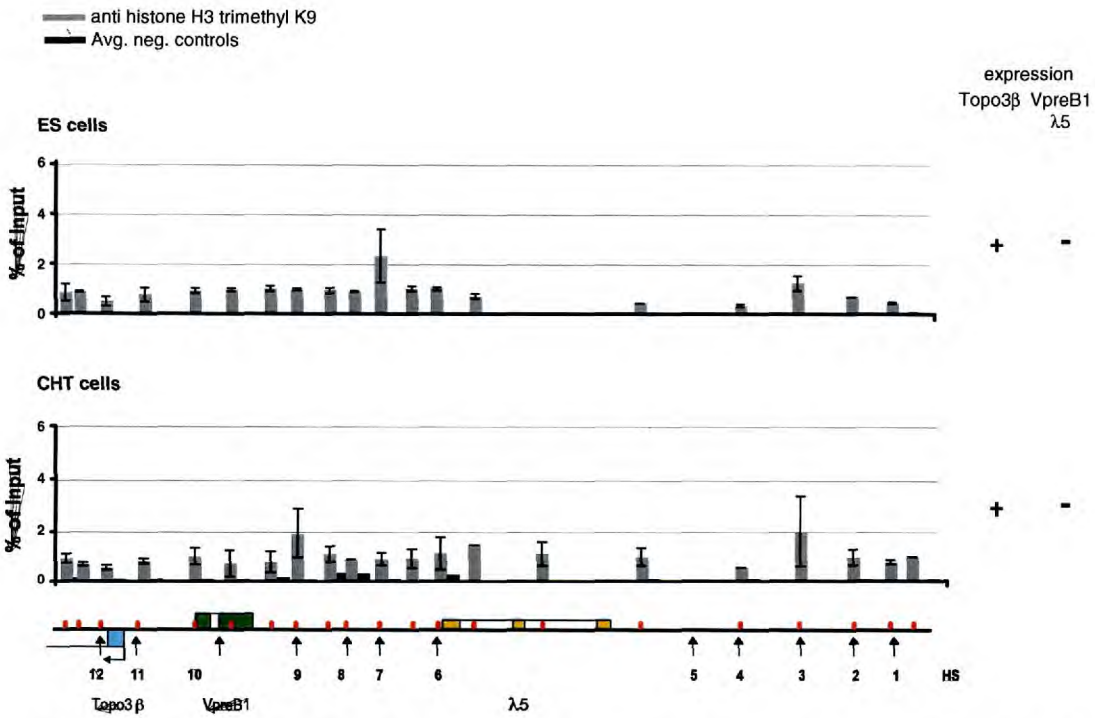


Fig. 5.4.2 Pattern of H3K9me3 across the λ5-*VpreB1* locus in ES (top panel) and in the multipotent mesenchymal stem cell line C3H10T1/2 (CHT, bottom panel). An antibody against histone H3 trimethylated lysine 9 (Upstate) was used for the immunoprecipitations. Non-specific anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background (black bars). The percentage of immunoprecipitated input is shown on the *y* axis of each plot. The *x* axis represents positions across the λ5-*VpreB1* domain. The expression status of the genes is showed on the right side of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

The result obtained from the analysis of the H3K9me3 ChIP experiments after differentiation showed large error bars, which made it difficult to drawn definitive conclusions and therefore have not been presented in this thesis.

The double H3K9me3S10ph modification was largely undetectable across the *λ5-VpreB1* locus in undifferentiated C3H10T1/2 (Fig. 5.4.3, CHT top panel), while the level increased upon differentiation along the osteogenic pathway (Fig. 5.4.3, CHTD bottom panel). The same behaviour was observed during B cell development, levels of H3K9me3S10ph were very low in ES cells and increased in mature B cells (section 4.3 Fig. 4.3.2).

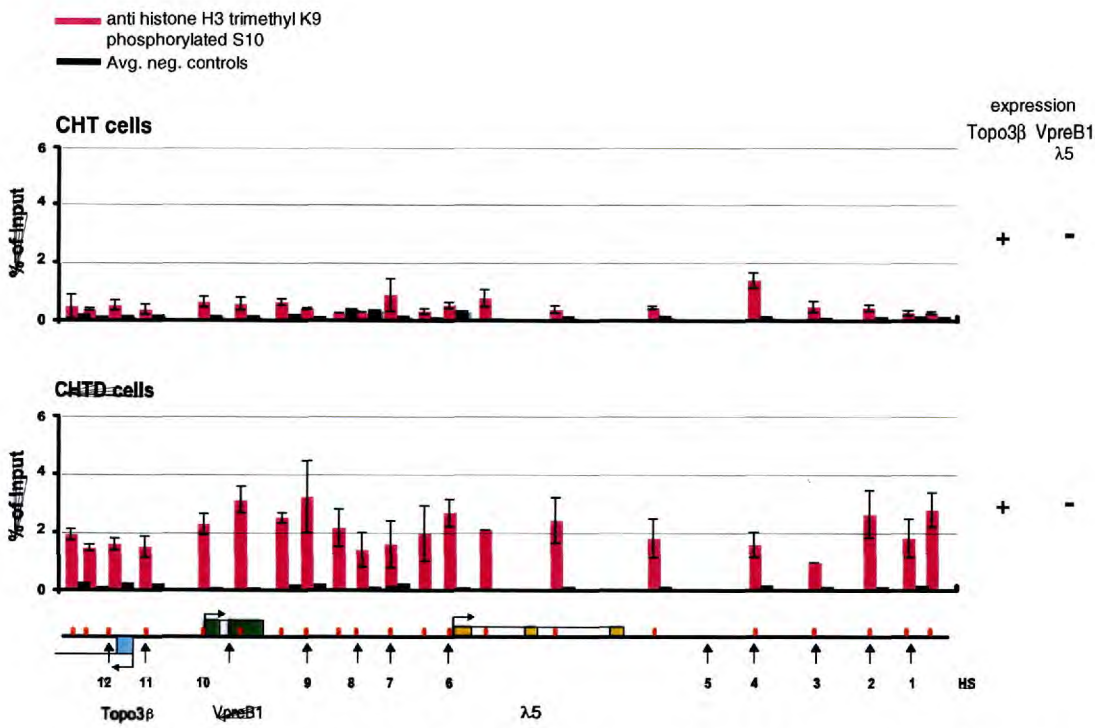


Fig. 5.4.3 Patterns of histone H3K4me3S10ph (pink bars), across the *λ5-VpreB1* locus in the multipotent mesenchymal stem cell line C3H10T1/2 (top panel) and in its product of differentiation along the osteogenic pathway (bottom panel). An antibody against histone H3 trimethylated lysine 9 phosporylated serine 10 (Abcam) was used for the immunoprecipitations. Non-specific anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background (black bars). The percentage of immunoprecipitated input is shown on the *y* axis of each plot. The *x* axis represents positions across the *λ5-VpreB1* domain. The expression status of the genes is showed on the right side of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. A locus map is shown below the plots; red vertical lines indicate amplicons positions, black arrows indicate the positions of the DNAseI hypersensitive sites.

5.5 Histone modifications in tissue specific and ubiquitously expressed genes during in vitro osteocyte differentiation

5.5.1 Histone H3 acetylation and H3 lysine 4 dimethylation

The analysis of histone H3 acetylation and histone H3K4me2 in the *λ5-VpreB1* locus showed that the level of enrichment of the two modifications at the genes promoters and transcribed regions was identical for the undifferentiated and differentiated C3H10T1/2. The same analysis was performed on a number of lineage-specific genes (Fig. 5.5.1) confirming this result. Tissue-specific genes were not enriched in H3 acetylation or H3K4me2 either in their promoters or in their transcribed regions, with the exception of *Nfl* and *Blimp1*. Housekeeping genes, instead, showed high percentages of immuno precipitation both in their promoters and transcribed regions. These results suggest that in multipotent precursor cells most of the genes belonging to alternative lineage choices are not marked by modifications characteristic of an active chromatin conformation.

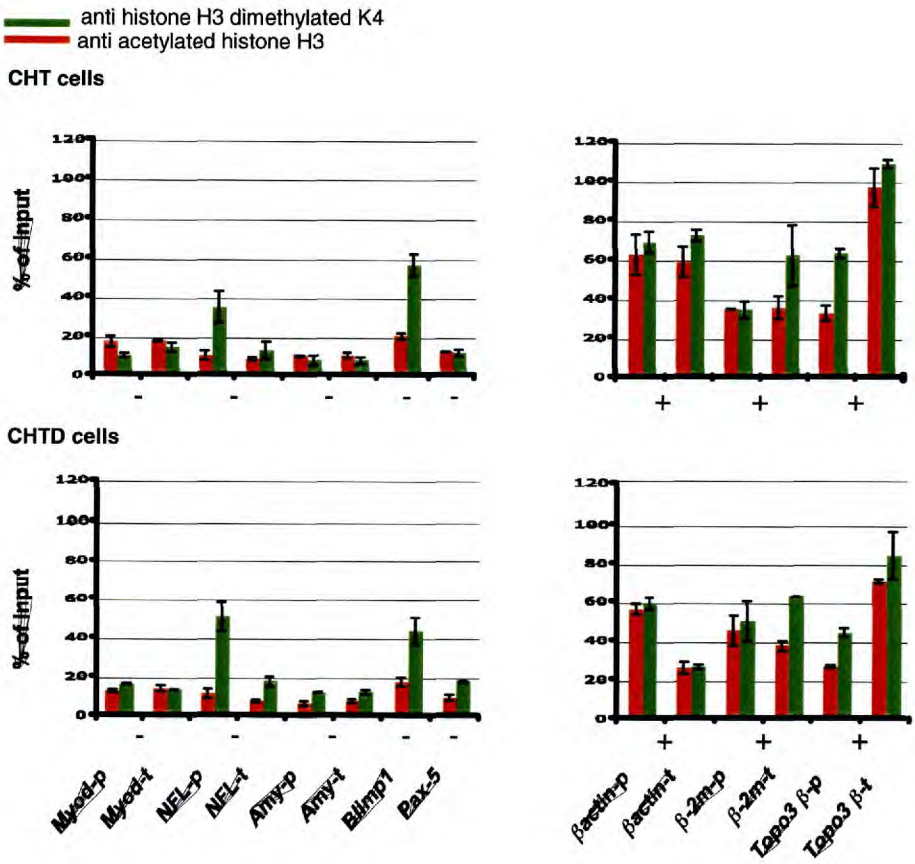


Fig. 5.5.1 Patterns of H3K4me2 (green bars) and H3 acetylation (red bars) in the multipotent undifferentiated C3H10T1/2 (CHT, top panel) and following differentiation along the osteogenic pathway (CHTD, bottom panel). Antibodies against acetylated histone H3 and histone H3 dimethylated K4 (both from Upstate) were used for the immunoprecipitations. Values on the y axis represents the percentage of immunoprecipitated input, on the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. Error bars indicate standard deviations relative to two different ChIP experiments. Negative controls (not shown) were calculated as the average of the values obtained with anti-flag, anti-IgG and no antibody and were always below 1% of IP input.

5.5.2 Histone H3 lysine 9 dimethylation

Analysis of H3K9me2 across the *λ5-VpreB1* locus showed a dramatic increase in the levels of enrichment of the modification in differentiated osteocytes compared to the multipotent progenitor cells. The study of the panel of tissue specific genes confirmed this result. In the undifferentiated cells (Fig. 5.5.2 top panel) the pattern was similar to that observed in ES cells (section 4.4.1).

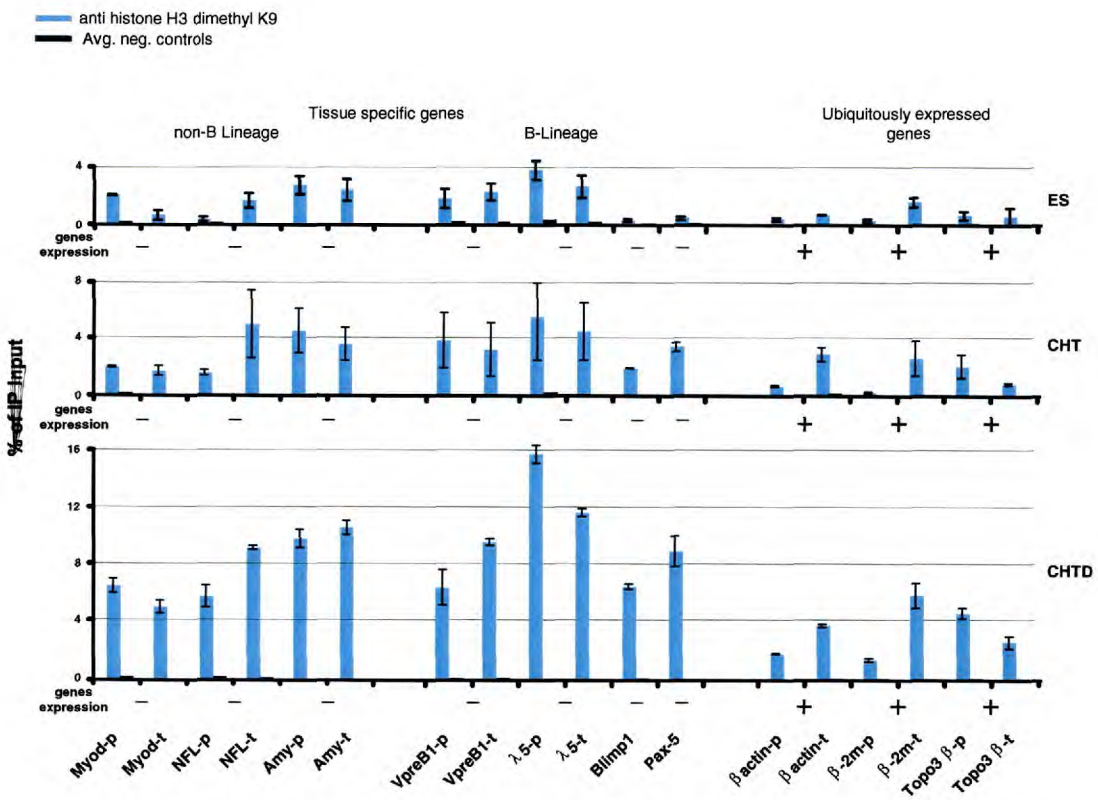


Fig. 5.5.2 Patterns of H3K9me2 (blue bars), in ES cells (top panel), in the multipotent mesenchymal stem cell line C3H10T1/2 (CHT middle panel) and in its product of differentiation along the osteogenic pathway (CHTD bottom panel). An antibody against histone H3 dimethylated lysine was used for the immunoprecipitations (Upstate). Anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background. Values on the y axis represents the percentage of immunoprecipitated input. On the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. P and t indicate respectively promoter and transcribed regions.

The enrichment was particularly low at the *MyoD*, *Nfl* and *Blimp1* promoters. It should be noted that *MyoD* and *Nfl* showed enrichment in H3K4me2 at the same sites in both the undifferentiated and the differentiated cells (section 5.5.1, Fig. 5.5.1). Upon differentiation (Fig. 5.6.1 bottom panel) the overall level of H3K9me2 rose dramatically in all of the tissue-specific genes. A small level of increase was also observed in the ubiquitously expressed gene *β-2m*, in particular in the transcribed regions. These results confirm the conclusion withdrawn in the specific case of the *λ5-VpreB1* locus that H3K9me2 plays an important role not only in developmental silencing and formation of facultative heterochromatin but also in long term silencing.

5.5.3 Histone H3 lysine 9 trimethylation

The analysis of the H3K9me3 at lineage specific gene promoters and transcribed regions in the mesenchymal stem cell line C3H10T1/2 showed very low levels of enrichment, comparable to ES cells. This result suggests that in a multipotent progenitor chromatin still retains an open conformation although the differentiation potential is severely restricted compared to ES cells.

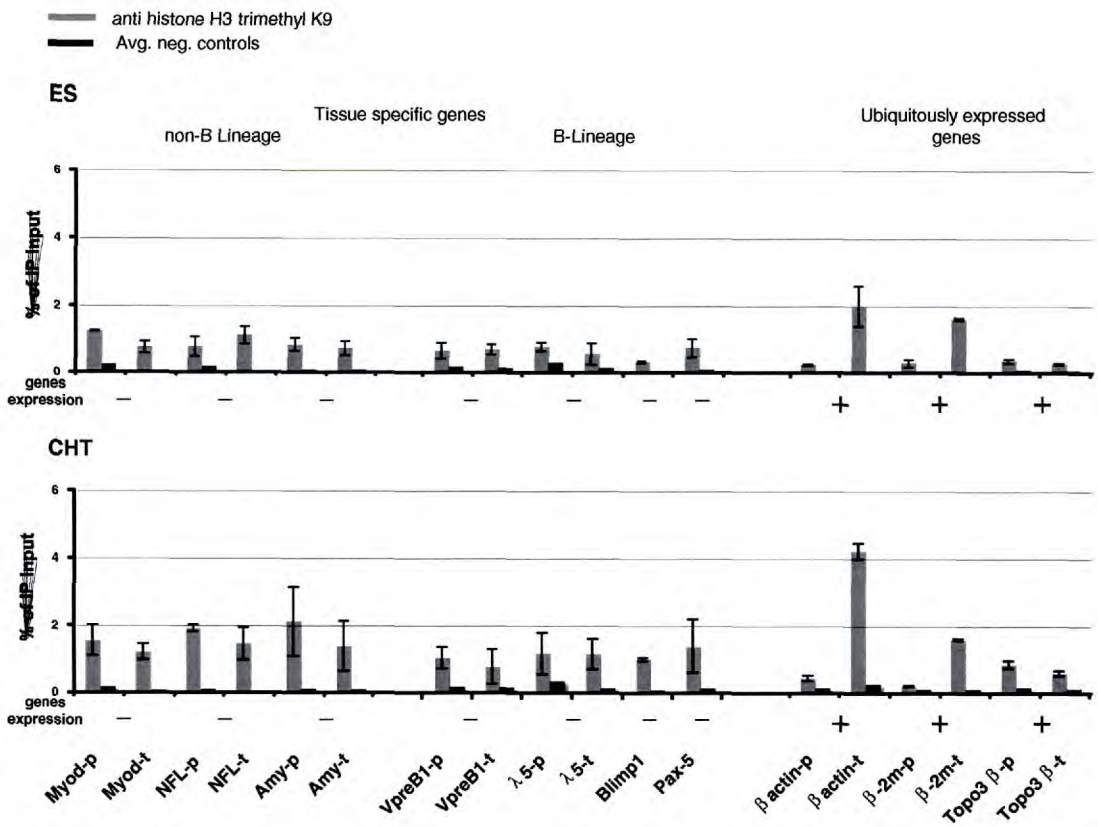


Fig. 5.5.3 Patterns of H3K9me3 (grey bars), in ES cells (top panel) and in the multipotent mesenchymal stem cell line C3H10T1/2 (bottom panel). An antibodys against histone H3 trimethylated lysine 9 was used for the immunoprecipitations (Upstate). Anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background. Values on the y axis represents the percentage of immunoprecipitated input, on the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. P and t indicate respectively promoter and transcribed regions.

5.5.4 Histone H3 K9 trimethylation / S10 phosphorylation

The double modification seems to play a role in silencing of the *λ5-VpreB1* locus in differentiated cells. Analysis of the panel of lineage-specific and housekeeping genes in ES cells is showing in Fig. 4.5.2 that H3K9me3S10ph presents a low level of enrichment independently of the gene being active or inactive. This low signal was not observed in the undifferentiated C3H10T1/2 cells (Fig. 5.4.3, CHT top panel). Analysis of selected gene promoters and coding regions showed levels of enrichment comparable to negative controls. H3K9me3S10ph has been associated with cells entering mitosis, the multipotent mesenchymal precursor cells are not dividing as rapidly as ES cells are. This data suggest that, in ES cells, the presence of the double modification could be actually due exclusively to the very rapid cell cycle. After differentiation the levels of enrichment increased in the lineage specific but not in the housekeeping genes (Fig. 5.5.4, bottom panel). Amylase, VpreB1 and λ5 seemed to be the less enriched among the tissue specific genes. Although H3K9me3S10ph has never been associated with the regulation of gene expression, the data herein shown provide some evidences that the double modification might be associated with both developmental and long term silencing.

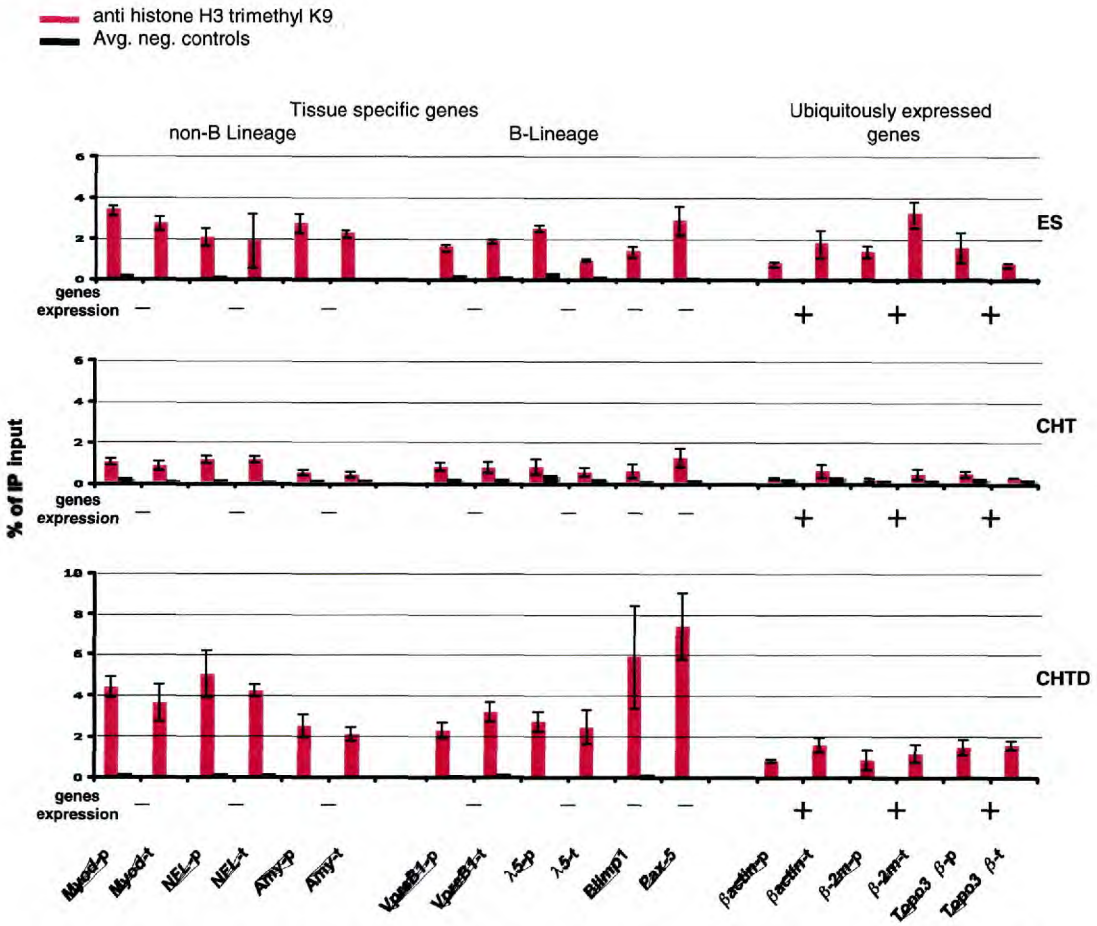


Fig. 5.5.4 Patterns of H3K4me3S10ph (pink bars), in ES cells, in the multipotent mesenchymal stem cell line C3H10T1/2 (CHT) and in its product of differentiation along the osteogenic pathway (CHTD, bottom panel). An antibodys against histone H3 trimethylated lysine 9 phosphorylated S10 was used for the immunoprecipitations (Abcam). Anti-Flag antibody, anti IgG and no antibody control were used to confirm that histone modification signals were detected above background. Values on the y axis represents the percentage of immunoprecipitated input, on the x axis are indicated the genes of interest. The expression status of the genes is showed on the bottom of each panel. Error bars indicate standard deviations calculated on minimum two ChIP experiments. P and t indicate respectively promoter and transcribed regions.

5.7 Summary

The results showed in this chapter demonstrate that the mesenchymal stem cell line C3H10T1/2 possesses a quite defined pattern of activating histone modifications. Restriction of differentiation potential caused depletion of H3 acetylation and H3K4me2 from gene promoters and transcribed regions of tissue specific genes. Nevertheless histone modifications characteristic of inactive chromatin were not enriched in the multipotent progenitors cells, suggesting the retaining of an open accessible chromatin conformation at this stage of development. Upon differentiation both the λ5-VpreB1 locus and non lymphoid tissue specific genes showed a dramatic increase in the levels of H3K9me2 and H3K9me3S10ph, suggesting that those modifications play a role in long term silencing of tissue specific genes.

6. DNA methylation analysis of the λ5-VpreB1 locus

6.1 Introduction

DNA methylation, histone deacetylation, and methylation of histone H3 at lysine 9 are the three best characterised covalent modifications associated with a repressed chromatin state; the existence of an epigenetic "interaction" between them leading to transcriptional silencing is well established although the precise sequence of events is still unclear. There is evidence that DNA methylation influences the histone modification pattern. In fact DNA methyltransferases (DNAMTs) and some methyl-CpG-binding domain (MBD) proteins have been shown to recruit repressor complexes containing histone deacetylases (HDACs) (Fuks et al., 2001; Robertson et al., 2000). There is also evidence that histone modification can influence DNA methylation. Observation made in fungi, plants and mammals identified methylation of histone H3 K9 as a signal for DNA methylation (Jackson et al., 2002; Lehnertz et al., 2003; Tamaru and Selker, 2001). In the previous chapters it has been shown that the silent state of the *λ5-VpreB1* locus in ES is characterised by the presence of H3K9me2 within the transcription units of the two genes. Histone H3K9me3 was absent from the domain in ES cells. After activation at the pro-B cell stage λ5 and VpreB1 are downregulated in immature B cells and completely silenced in mature B cells. This latest silent state was shown to be characterised by enrichment in H3K9me2 especially on the *VpreB1* gene and in the intergenic region. In order to further characterise the epigenetic profile of the silent *λ5-VpreB1* locus DNA methylation was investigated across the entire domain at different stages of B cell development.

6.2 Strategy for southern blot analysis of DNA methylation

The enzyme McrBC was used to digest genomic DNA extracted from ES, primary pre-B cells, activated mature B cells and brain tissue. McrBC is an endonuclease that cleaves DNA containing methyl-cytosine (Sutherland et al., 1992) on one or both strands. Sites on the DNA that are recognized by McrBC consist of two half sites of the form $(G/A)^{m}C$. These half sites can be separated by up to 3kb but the optimal separation is 50-100 base pairs (Panne et al., 1999; Stewart and Raleigh, 1998). Analysis of McrBC cutting sites across the locus compared to the

number of possible methylated CpGs showed that 75% of them could be recognized by McrBC and, if methylated, cut.

The Southern blot strategy used is shown in Fig. 6.2.1. The locus was divided into three broad regions of approximately 5kb each; three non-repetitive probes were used for end labeling. Probes 1 and 4 were previously described by (Minaee et al., 2005) and were used to detect methylated regions in the 3' and 5' end of the domain. The λ5 probe used for the central region of the locus was a restriction fragment generated by a SacI and SphI digestion and purified by agarose gel electrophoresis.

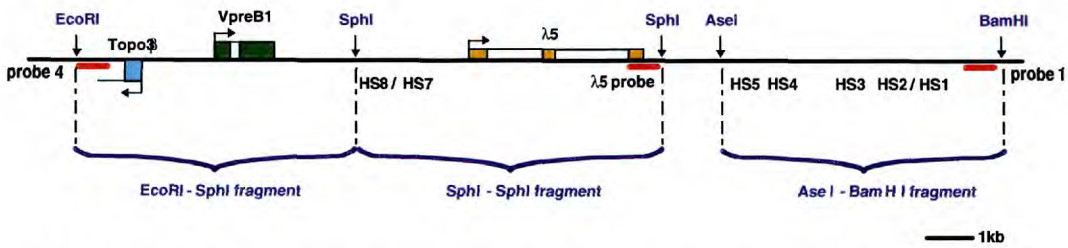


Fig.6.2.1 southern blot strategy. Horizontal red lines indicate the position of non-repetitive probes.

Genomic DNA was digested with increasing amounts of McrBC and purified. DNA was then run on agarose gel, etidium bromide stained (Fig.6.2.2), southern blotted and detected by indirect end labeling. Fragments cut by McrBC appeared as bands migrating below the parental restriction enzyme band.

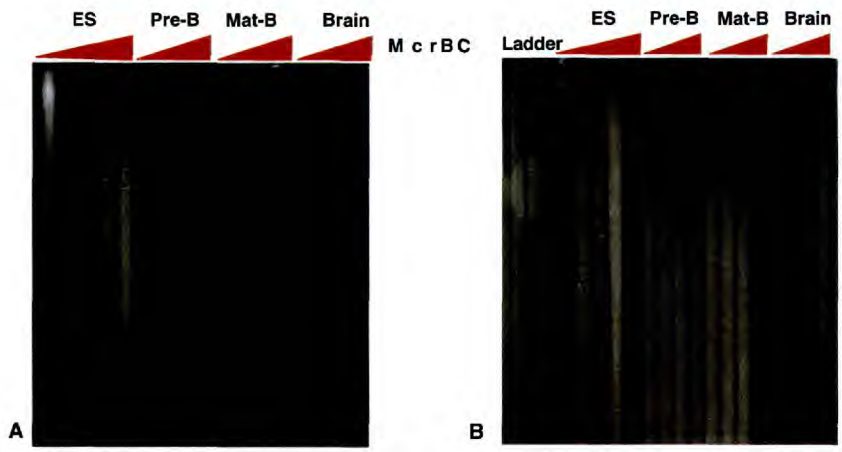


Fig. 6.2.2 Etidium bromide staining of southern blots showed respectively in Fig. 6.3.1 (A) and in Fig. 6.3.3 (B). Etidium bromide staining was used to ensure that genomic DNA was properly digested by McrBC and that the electrophoresis run was satisfactory.

6.3 DNA methylation analysis of the λ5-VpreB1 locus

To map the region containing the *Topo3β* and *VpreB1* genes a probe located at the 5' end of the locus was used (Minaee et al., 2005). Genomic DNA treated with McrBC was digested with EcoRI and SphI to yield a 5.4 kb parent fragment. Sites of methylation were detected in the second exon of the VpreB1 genes in mature B cells and in brain (Fig. 6.3.1). Three sites of methylation were observed in brain and one in mature B cells. No sites of methylation were observed in primary pre-B and in ES cells.

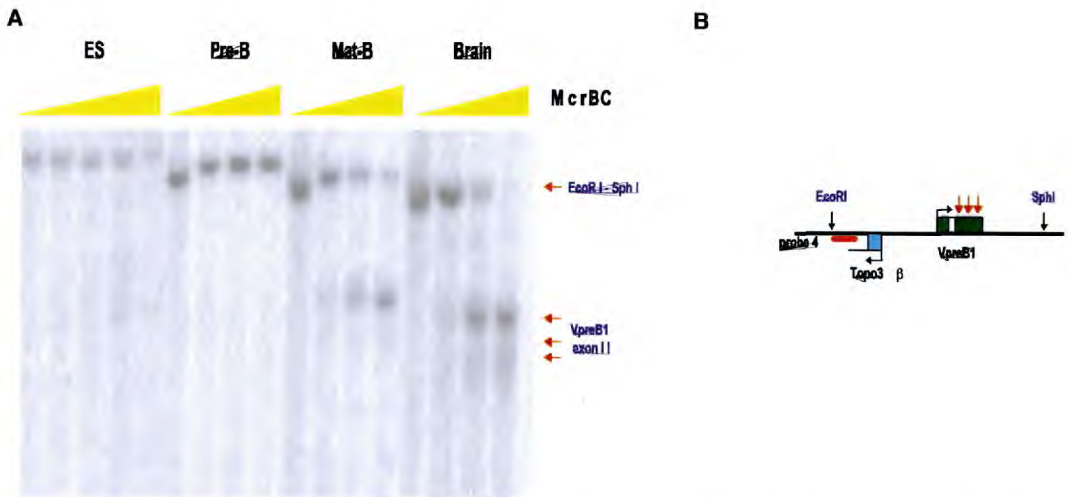


Fig. 6.3.1 A) DNA from McrBC treated ES cells, primary pre-B cells, primary mature B cells and brain tissue was digested EcoRI and SphI and hybridized with probe 4. Red arrows in the right margin indicate bands generated by McrBC cutting. B) Schematic illustration of the strategy used. Red vertical arrows indicate the position of the metylated residues on the VpreB1 gene. The red horizontal line indicates the position of the probe.

The central region of the locus, containing HS 7 and 8 and the *λ5* gene, was analysed using a probe spanning the third exon of *λ5* (Fig. 6.3.2). Genomic DNA treated with McrBC was digested with Sph I and produced a 6kb parental fragment. The analysis detected a methylated cytosine in *λ5* exon III in mature B cells, brain and surprisingly, in pre-B cells. Methylation at exon II instead was detected only in mature B cells and brain. ES cells did not show any methylation sites in this region. The presence of methylation in the third exon of the *λ5* gene in pre-B cells was unexpected since the gene is expressed at this stage of development.

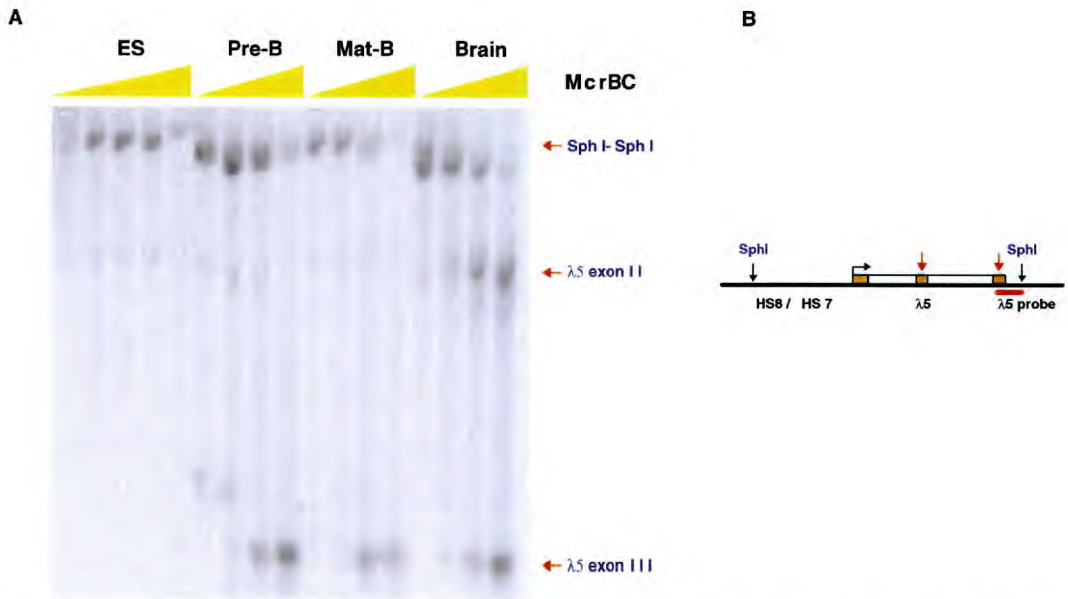


Fig. 6.3.2 A) DNA from McrBC treated ES cells, primary pre-B cells, primary mature B cells and brain tissue was digested Sph I and hybridized with the *λ5* probe. Red arrows in the right margin indicate bands generated by McrBC cutting. B) Schematic illustration of the strategy used. Red vertical arrows indicate the position of the metylated residues on the *λ5* gene. The red horizontal line indicates the position of the probe.

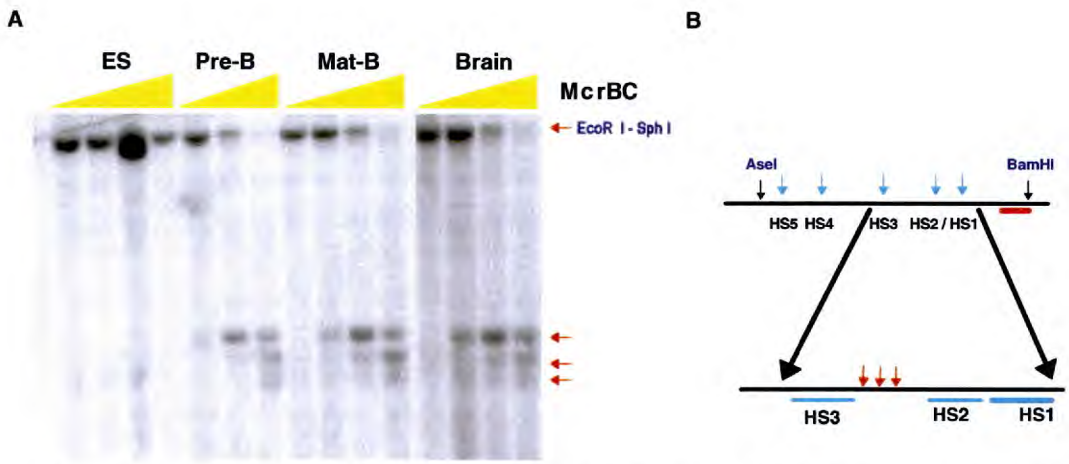


Fig. 6.3.3 A) DNA from McrBC treated ES cells, primary pre-B cells, primary mature B cells and brain tissue was digested Ase I and BamH I and hybridized with probe 3. Red arrows in the right margin indicate bands generated by McrBC cutting. B) Schematic illustration of the strategy used. Red vertical arrows indicate the position of the metylated residues on the *λ5* gene. The red horizontal line indicates the position of the probe. The region between HS 1 and 3 has been enlarged to show the position of the methylated cytosine

The region containing HSs 1 to 5 was analysed with a probe located at the 3' edge of the *λ5-VpreB1* domain (Minaee et al., 2005). Hypersensitive sites 1 to 5 are

required for efficient expression of both *λ5* and *VpreB1* in single and two copy integrations in transgenic mice (Sabbattini et al., 1999). This region also has a strong enhancer activity (Minaee S. characterisation of the *λ5 VpreB1* functional domain, London University 2002). McrBC digestion revealed a cluster of three methylated cytosines downstream from HS3 (LTR region) thet were present at the pre-B cell stage and conserved in mature B cells and brain. DNA methylation was not detected in ES cells in this region (Fig. 6.3.3).

The analysis of DNA methylation across the *λ5-VpreB1* locus with the McrBC enzyme revealed the presence of methylated cytosines in the λ5 exon III and downstream HS3 already in pre-B cells, when the genes are fully active. Methylation of the locus increased in mature B cells and reached its maximum in Brain, where additional sites were detected at VpreB1 exon II and λ5 exon III. Interestingly, ES cells did not show DNA methylation at any of these sites (Fig. 6.3.4).

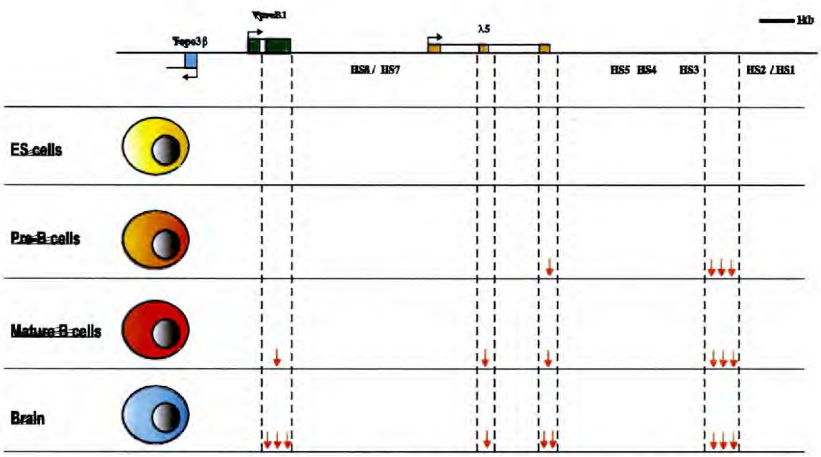


Fig. 6.3.4 Summary of the methylation sites detected by McrBC digestion. Red arrows indicate methylated cytosine.

6.4 Summary

McrBC analysis revealed a substantial difference between the methylation status of the λ5-VpreB1 locus in ES and Mature B cells. The domain appeared completely non-methylated in ES cells and showed instead methylation in the two genes and on HS3 in mature B cells. The amount of methylated sites increased in brain compared to mature B cells. These results therefore suggest that the λ5-VpreB1 locus is less methylated in ES cells than in Mature B cells. The results presented in this chapter can not be considered conclusive and bisulphite analysis of the genes will be necessary to determine the number of methylated cytosine in each cell types and asses if the methylation status of the locus can actually mark the difference between the two silent states observed in ES and mature B cells.

7 Discussion

7.1 Overview

The problem of how a single fertilised egg can develop into a complex multi-cellular organism is one of the major questions in biology. The fertilised egg and the cells of the 2- and 4-cell embryo are totipotent and can give rise to every cell type in the body. By the time that the blastocyst stage is reached a certain degree of differentiation as already occurred. The inner cell mass (ICM) can still give rise to all somatic lineages but lack the ability to differentiate into the placenta and the extra-embryonic membranes. Thus, the essential component of development is cell differentiation and the restriction of differentiation potential as cell commit to various fates. During development, tissue-specific patterns of gene expression need to be established, maintained throughout the life of the organism, and changed when and where appropriate. Many studies on gene regulation have focused on mechanisms of transcriptional activation. However, gene silencing is known to be equally prominent than gene activation in cell fate decisions and developmental progression. In the study described in this thesis the *λ5-VpreB1* locus was used as a model system to investigate the epigenetic regulation of gene silencing during stem cell commitment and differentiation. The locus has a number of advantages for this type of study. Its small size allows ChIP analysis to be carried at very high resolution and the tight tissue and stage-specific regulation of the genes makes it an excellent system for studying mechanisms of transcriptional control during cell commitment. The epigenetic features of developmental and long term silencing of the locus were investigated during B cell development and during *in vitro* differentiation of osteocytes from the mesenchymal stem cell line C3H10T1/2. The use of these two cells system allowed a comparison between stage-specific and long-term lineage-specific silencing of a developmentally regulated locus.

7.2 The active *λ5-VpreB1* locus

In order to understand the role of histone modifications in defining the active *λ5-VpreB1* locus, the pattern of histone H3 and H4 acetylation was analysed across the entire domain in pre-B cells. The analysis revealed the presence of distinct peaks of H3 acetylation that were localised in intergenic regions (between the Topo3β and

VpreB1 and on HS 7 and 8) and on the *λ5* gene. The 3' region of the locus did not show enrichment for H3 acetylation despite being able to promote full expression at single copy in euchromatin and variegated expression in heterochromatin (Lundgren et al., 2000).

Interestingly, the promoter regions of the *Topo3β*, *VpreB1* and *λ5* genes also appeared to be less enriched for histone acetylation compared to the body of the genes. Nucleosome remodelling or nucleosome loss at active promoters has been observed previously. Studies on the PHO5 promoter in yeast showed that a dramatic loss of acetylation signal occurs upon induction of the promoter and Chip analysis suggested that this could be caused by a progressive loss of histones (Reinke and Horz, 2003). Severe histone H3 hypoacetylation in the promoter regions of yeast genes does not affect the ability of TBP to bind TATA elements, or the ability to transcribe the gene (Kristjuhan et al., 2002). By contrast, similar hypoacetylation of the coding region correlates with inhibition of transcription (Eberharter and Becker, 2002). Our results indicate that transcriptional competence of the *λ5-VpreB1* domain is characterised by discrete peaks of H3 acetylation instead of the formation of a broad hyperacetylaed domain such as the one observed in the β globin locus (Litt et al., 2001; Schubeler et al., 2001). The fact that acetylation of histone H4 showed much lower levels of enrichment compared with H3 acetylation, suggests that if an "acetylation code" is established in the active *λ5-VpreB1* domain, it is likely to be mainly based on modifications to histone H3. The low uniform level of histone H4 acetylation that was observed is in agreement with previously published results describing H4 hypoacetylation across the human growth hormone and human β-globin loci (Kimura et al., 2004; Schubeler et al., 2001) but is in contrast with the observation of peaks of H4 acetylation in the active murine β-globin locus (Forsberg et al., 2000; Kiekhaefer et al., 2002).

Histone acetylation has been linked with transcriptional activation (Turner, 2000) and with a generalised "open" and accessible chromatin structure. It has been shown that histone acetylation may decrease the stability of the compacted 30nm chromatin fibre (Bird and Wolffe, 1999) and destabilise higher order structures (Tse et al., 1998). This unfolding of active chromatin regions may facilitate transcription and correlates with a variety of indirect observations such as accessibility to

endonuclease digestion, timing of replication, nuclear position and patterns of DNA methylation (Anguita et al., 2001).

In contrast to the localised peaks of histone H3 acetylation, levels of di-methyl H3K4 were found to be increased across the entire domain in pre-B cells. This result suggests that this epigenetic mark plays a major role in defining the extent of functional gene expression domains. Work carried out in parallel with this study showed that the first evidence of active histone modifications in the domain is the presence of a localised peak of histone H3 acetylation and H3K4 dimethylation at HS8 in ES cells. This mark expands to include HS7 in lymphoid precursors (Szutorisz et al., 2005) before expanding to cover the entire domain (this study). The spreading coincides with a dramatic increase in the binding of general transcription factors, TBP, Brg1 and pol II (Szutorisz et al., 2005).

An important question concerns the mechanisms by which the patterns of histone H3 acetylation and H3 K4 dimethylation are established in the *λ5-VpreB1* domain. Previous analysis of the domain has revealed the presence of 12 DNase I hypersensitive sites at the pre-B cell stage (Minaee et al., 2005). Although peaks of H3 acetylation are found at HS8 and HS11, several of the other acetylation peaks do not correlate with the presence of HS. In particular, a strong peak was observed within the first intron of the *λ5* gene. High levels of H3 K4 dimethylation were observed across the entire domain but peaks of this modification were also observed and several of these coincided with peaks of H3 acetylation. However, a particularly strong peak of K4 dimethylation upstream of the *λ5* promoter coincided with a region of H3 hypoacetylation which does not contain a DNase I HS. Theoretical considerations suggest that formation of sequence-specific histone modifications should be driven by binding of sequence specific transcription factors, which would then recruit histone methyltransferases and acetyltransferases. Histone modifications could also spread from centers of factor driven HAT and HMT recruitment. The results obtained for the *λ5-VpreB1* locus suggest a complex relationship between factor binding, HS formation and recruitment of histone modification enzymes. Ultimately, understanding this relationship will require functional studies involving deletion of specific sequences and analysis of the effect on histone modifications.

7.3 Epigenetic features of the *λ5-VpreB1* domain in the B cell lineage and in non lymphoid lineages

The two best-known traits of silent chromatin in mammalian cells are the absence of acetyl groups at the amino termini of histone H3 and H4 and DNA methylation. In recent years methylation of lysine 9 on histone H3 has also been shown to be a hallmark of silent genes. Di-methylation on histone H3K9 has been linked with formation of facultative heterochromatin while tri-methylation is correlated with constitutive heterochromatin and binding of HP(1) (reviewed by Smale, 2003).

The epigenetic profile of the *λ5-VpreB1* locus was analysed in ES cells and in the multipotent mesenchymal stem cell line C3H10T1/2 and was found to be remarkably similar in both cell types, which suggests that the first step toward developmental restriction might not involve dramatic epigenetic changes. H3K9me2 was found to be enriched in the *λ5* and *VpreB1* genes promoters and within transcribed regions in both cell types, although the levels were slightly higher in the multipotent precursor cells than in ES cells. In addition, in ES cells the region marked by HS7 and 8 showed a pronounced reduction in the level of H3K9me2 compared to the rest of the locus. HS7 and 8 have previously been shown to constitute a peak of H3K4me2 and a centre of recruitment of Pol II and general transcription factors in ES cells (Szutorisz et al., 2005). The presence of a trough of H3K9me2 at this specific site is in agreement with these findings.

The C3H10T1/2 multipotent stem cells showed a slightly higher level of H3K9me2 compared with ES cells and K9 trimethylation was largely absent from the locus at both of these stages. Differentiation of C3H10T1/2 cells into osteocytes resulted in a striking increase in the level of H3K9me2 across the entire domain. The sharp increase in H3K9me2 observed in the differentiated C3H10T1/2 cells is in agreement with recent published observations. ChIP on chip analysis of a total of 56 large regions containing highly conserved noncoding elements (HCNEs) in ES cells has revealed extended regions of H3 lysine 27 methylation that harbour smaller regions of H3 lysine 4 methylation. This pattern has been named "bivalent domains" and has been suggested as a mechanism that silences developmental genes in ES cells while keeping them poised for activation. Upon differentiation of the ES cells

the H3K27me marks expanded in regions containing genes that were not expressed in the differentiated cells (Bernstein et al., 2006).

Taken together these results lead to two important conclusions. Firstly a gene locus that is not expressed in mesenchymal lineages appears to retain a relatively plastic epigenetic profile in multipotent mesenchymal stem cells. Secondly differentiation into a specific lineage represents a major commitment event, which involves the acquisition of substantial levels of H3 K9 methylation as genes become permanently silenced. These results suggest that multipotent stem cells may be more epigenetically plastic than was originally thought and could be susceptible to trans-differentiation into other lineages. They are also in line with recent functional studies showing that neural stem cells, when used as donors for nuclear transplantation, produce embryonic stem cells at a higher efficiency than blastocysts derived from terminally differentiated neuronal donor cells (Blelloch et al., 2006).

7.4 Developmental versus long term silencing of the B cell-specific *λ5-VpreB1* domain

One of the aims of this study was to investigate differences between developmental and long term silencing of a tissue specific gene expression domain. The stage specific silencing of the *λ5-VpreB1* locus during B cell development made it possible to address this question by comparing the epigenetic profile of the silent domain in mature B cells and *in vitro* differentiated osteocytes.

Analysis of H3 K9 dimethylation in mature B cells gave the unexpected result that the modification was present at similar levels across the domain to those observed in ES cells with the exception of the ES cell-specific trough at HS8. In contrast, histone H3 K9 dimethylation was highly enriched across the domain in the differentiated C3H10T1/2 cells. H3 K9 trimethylation was shown to be largely absent from the domain in mature B cells. Taken together, these results indicate that there are significant differences between the epigenetic mechanisms that govern stage-specific repression and long-term lineage-specific silencing of transcription.

In general, stage specific silencing of gene expression is initiated by binding of repressors to promoters and enhancers with the activators often still present in the cells. Silencing of the *λ5-VpreB1* locus in mature B cells has been shown to depend on the presence of a binding site for the transcription factor Ikaros (Sabbattini et al.,

2001). EBF, which is involved in activating *λ5*, has been shown to be still present in the cells at this stage. The results described in this thesis show that repression at the mature B cell stage does not require the high levels of H3K9 dimethylation that are associated with long-term silencing of the locus in non lymphoid lineages. The findings from this study differ from the results of an analysis of Ikaros dependent silencing of the mouse *Dntt* gene, which is silenced and repositioned to pericentric heterochromatin during thymocyte maturation. *Dntt* silencing was accompanied by ordered deacetylation of histone H3 K9, loss of methylation of H3 K4 followed by bidirectional spreading of histone H3 lysine 9 from the promoter region to a distance of several kb upstream and downstream (Su et al., 2004). The difference in the epigenetic profiles of the two loci suggests that there is significant heterogeneity in the mechanisms that mediate stage specific silencing of transcription at different genes.

Whereas tissue-specific silencing depends on repressors, long-term lineage-specific silencing is more likely to result from absence of appropriate activators. It seems that in this situation genes are more susceptible to H3 K9 methylation and this could be important for fixing the silent state in fully differentiated cells.

7.5 Histone H3 K9 methylation is a general marker for gene silencing during cell differentiation

To generalise the results obtained for the *λ5-VpreB1* locus, a set of tissue specific and ubiquitously expressed genes was analysed both during B cell development and during *in vitro* differentiation of osteocytes. Histone H3K9 and H3K4 dimethylation presented a varied pattern at tissue specific gene promoters and transcribed regions in ES cells and in C3H10T1/2 mesenchymal stem cells. In general tissue specific genes showed lower levels of histone H3 K4 dimethylation compared to ubiquitously expressed genes while the contrary was observed for H3 K9 dimethylation in both cell types. Tri methylation of H3 K9 was not observed in tissue specific genes both in ES cells and in the mesenchymal stem cell line. Previous studies have shown that global levels of histone H3K9me3 are low in ES cells and increase upon differentiation (Lee et al., 2004). Interestingly tissue-specific genes that are not expressed in the B cell lineage showed enrichment of H3K9me3

in mature B cells. Ubiquitously expressed genes showed enrichment of H3K9me3 in their transcribed regions in all of the cell types that were analysed, this is in agreement with previous observations (Vakoc et al., 2005) and suggests a dual role for this modification in gene regulation and maintenance of a transcriptionally active or inactive status.

The double histone H3 K9 trimethylated/ S10 phosphorylated (H3K9me3S10ph) modification was also analysed in this study. This modification has been implicated in chromosome condensation and effects on binding of chromatin proteins during mitosis (Fischle et al., 2005). The results described in this thesis suggest a novel role for H3K9me3S10ph in gene silencing. Increased levels of the double modification were observed in tissue specific genes in mature B cell compared to ES cells and in osteocytes compared to C3H10T1/2 cells. Experiments conducted by Sabbattini et al suggest that phosphorylation at H3 S10 blocks the binding of HP1β to the neighbouring di-and tri-methyl H3 K9 allowing the binding of other factors involved in chromatin condensation and the formation of an alternative form of facultative heterochromatin (P. Sabbattini, personal communication).

In summary, the analysis of epigenetic modifications during development and differentiation described in this thesis suggests that non-expressed tissue-specific genes, in pluripotent and multipotent precursor, are marked by H3K9me2. This level of H3K9me2 is maintained at silent tissue-specific genes in fully committed somatic cells and reach its maximum in terminally differentiated cells. The results obtained in mature B cells, suggest that genes that are subject to long-term lineage-specific silencing will acquire the lysine 9 trimethylation mark in their promoters as well as in the transcribed regions in order to favour the formation of a long-term silent state.

7.6 DNA methylation of the *λ5-VpreB1* locus during B cell development

The patterns of H3K9me2 and H3K9me3 across the *λ5-VpreB1* locus in ES and mature B cells did not reveal major differences in the silent state of the domain between these two cell types. In order two further investigate epigenetic changes in the locus during B cell development, a Southern blot approach using the methylation sensitive enzyme McrBC was used to analyse DNA methylation across the entire domain. Interestingly, methylation was not observed in ES cells despite the fact that

the locus is silent at this stage. At the pre-B cell stage, when the genes are fully active two sites of CpG methylation were detected. One of these was at an LTR element located at HS3, while the other was in λ5 exon 3 and did not correspond to any known elements. The number of methylated cytosines increased progressively in mature B cells and reached its maximum in brain tissue. Analysis of ES cell genomic DNA with the methylation sensitive AvaI enzyme further supported the conclusion that the locus is undermethylated in ES cells.

It has been shown that there is a progressive loss of methylation in the early embryo, which reaches its lowest point in the morula and shows some increase by the blastocyst stage. Jackson et al showed that *Dnmt3a* and *3b* knock down in ES cells caused a global DNA hypometylation, which resulted in a failure to differentiate and inappropriate expression of trophectoderm-specific genes. Although these observations indicate that DNA methylation plays an important role in gene regulation in ES cells, the results described here suggest that at least some tissue-specific genes may be silenced in ES cells in the absence of methylation. This finding is further evidence that developmentally regulated genes are in a potentiated state in ES cells, ready for expression at later stages.

The presence of methylated residues in the locus in pre-B cells was unexpected and highlights the difficulty of making a direct correlation between DNA methylation and transcription. Analysis of the mouse *Dntt* locus during thymocyte maturation showed relatively low levels of methylation in unstimulated thymocytes where the gene is active. Inactivation following stimulation was accompanied by only a modest increase in methylation. The results obtained for the *λ5-VpreB1* locus show some similarities to these observations and suggest that DNA methylation may not be a primary silencer of developmentally regulated loci.

Reference List

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P. B., *et al.* (1999). Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. Embo J *18*, 1923-1938.

Aasland, R., Stewart, A. F., and Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. Trends Biochem Sci *21*, 87-88.

Adams, B., Dorfler, P., Aguzzi, A., Kozmik, Z., Urbanek, P., Maurer-Fogy, I., and Busslinger, M. (1992). Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. Genes Dev *6*, 1589-1607.

Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Monch, K., Astrand-Grundstrom, I., Sitnicka, E., Sasaki, Y., and Jacobsen, S. E. (2001). Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. Immunity *15*, 659-669.

Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. Cell *103*, 667-678.

Ahrens, M., Ankenbauer, T., Schroder, D., Hollnagel, A., Mayer, H., and Gross, G. (1993). Expression of human bone morphogenetic proteins-2 or -4 in murine mesenchymal progenitor C3H10T1/2 cells induces differentiation into distinct mesenchymal cell lineages. DNA Cell Biol *12*, 871-880.

Akerblad, P., Rosberg, M., Leanderson, T., and Sigvardsson, M. (1999). The B29 (immunoglobulin beta-chain) gene is a genetic target for early B-cell factor. Mol Cell Biol *19*, 392-401.

Akhtar, A., and Becker, P. B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Mol Cell *5*, 367-375.

Alen, C., Kent, N. A., Jones, H. S., O'Sullivan, J., Aranda, A., and Proudfoot, N. J. (2002). A role for chromatin remodeling in transcriptional termination by RNA polymerase II. Mol Cell *10*, 1441-1452.

Ancelin, K., Lange, U. C., Hajkova, P., Schneider, R., Bannister, A. J., Kouzarides, T., and Surani, M. A. (2006). Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol *8*, 623-630.

Andres, M. E., Burger, C., Peral-Rubio, M. J., Battaglioli, E., Anderson, M. E., Grimes, J., Dallman, J., Ballas, N., and Mandel, G. (1999). CoREST: a functional corepressor required for regulation of neural-specific gene expression. Proc Natl Acad Sci U S A *96*, 9873-9878.

Anguita, E., Johnson, C. A., Wood, W. G., Turner, B. M., and Higgs, D. R. (2001). Identification of a conserved erythroid specific domain of histone acetylation across the alpha-globin gene cluster. Proc Natl Acad Sci U S A *98*, 12114-12119.

Archer, A., Sauvaget, D., Chauffeton, V., Bouchet, P. E., Chambaz, J., Pincon-Raymond, M., Cardot, P., Ribeiro, A., and Lacasa, M. (2005). Intestinal apolipoprotein A-IV gene transcription is controlled by two hormone-responsive elements: a role for hepatic nuclear factor-4 isoforms. Mol Endocrinol *19*, 2320-2334.

Archer, T. K., Lefebvre, P., Wolford, R. G., and Hager, G. L. (1992). Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. Science *255*, 1573-1576.

Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A *88*, 10148-10152.

Avner, P., and Heard, E. (2001). X-chromosome inactivation: counting, choice and initiation. Nat Rev Genet *2*, 59-67.

Ayer, D. E. (1999). Histone deacetylases: transcriptional repression with SINers and NuRDs. Trends Cell Biol *9*, 193-198.

Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K., and Rauscher, F. J., 3rd (2003). Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. Genes Dev *17*, 1855-1869.

Bain, G., Maandag, E. C., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., and et al. (1994). E2A

proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell *79*, 885-892.

Bain, G., Robanus Maandag, E. C., te Riele, H. P., Feeney, A. J., Sheehy, A., Schlissel, M., Shinton, S. A., Hardy, R. R., and Murre, C. (1997). Both E12 and E47 allow commitment to the B cell lineage. Immunity *6*, 145-154.

Bannister, A. J., and Kouzarides, T. (2004). Histone methylation: recognizing the methyl mark. Methods Enzymol *376*, 269-288.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Barr, M. L. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleprotein synthesis. . Nature *163*, 676-677.

Bell, A. C., and Felsenfeld, G. (1999). Stopped at the border: boundaries and insulators. Curr Opin Genet Dev *9*, 191-198.

Bell, A. C., West, A. G., and Felsenfeld, G. (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic. Science *291*, 447-450.

Bender, M. A., Reik, A., Close, J., Telling, A., Epner, E., Fiering, S., Hardison, R., and Groudine, M. (1998). Description and targeted deletion of 5' hypersensitive site 5 and 6 of the mouse beta-globin locus control region. Blood *92*, 4394-4403.

Berger, S. L. (1999). Gene activation by histone and factor acetyltransferases. Curr Opin Cell Biol *11*, 336-341.

Bernstein, B. E., Liu, C. L., Humphrey, E. L., Perlstein, E. O., and Schreiber, S. L. (2004). Global nucleosome occupancy in yeast. Genome Biol *5*, R62.

Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J., 3rd, Gingeras, T. R., *et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

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.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The

carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol *203*, 971-983.

Blackwood, E. M., and Kadonaga, J. T. (1998). Going the distance: a current view of enhancer action. Science *281*, 60-63.

Blom van Assendelft, G., Hanscombe, O., Grosveld, F., and Greaves, D. R. (1989). The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. Cell *56*, 969-977.

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

Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006). Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells *24*, 2007-2013.

Boggs, B. A., Cheung, P., Heard, E., Spector, D. L., Chinault, A. C., and Allis, C. D. (2002). Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. Nat Genet *30*, 73-76.

Borge, O. J., Adolfsson, J., and Jacobsen, A. M. (1999). Lymphoid-restricted development from multipotent candidate murine stem cells: distinct and complimentary functions of the c-kit and flt3-ligands. Blood *94*, 3781-3790.

Bork, P., and Koonin, E. V. (1993). An expanding family of helicases within the 'DEAD/H' superfamily. Nucleic Acids Res *21*, 751-752.

Botchan, M., Topp, W., and Sambrook, J. (1976). The arrangement of simian virus 40 sequences in the DNA of transformed cells. Cell *9*, 269-287.

Boyer, L. A., Latek, R. R., and Peterson, C. L. (2004). The SANT domain: a unique histone-tail-binding module? Nat Rev Mol Cell Biol *5*, 158-163.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

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

Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M., and Broach, J. R. (1996). Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. Mol Cell Biol *16*, 4349-4356.

Braunstein, M., Rose, A. B., Holmes, S. G., Allis, C. D., and Broach, J. R. (1993). Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. Genes Dev *7*, 592-604.

Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., Winston, F., and Allis, C. D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev *15*, 3286-3295.

Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M., and Fisher, A. G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845-854.

Brown, C. E., Lechner, T., Howe, L., and Workman, J. L. (2000). The many HATs of transcription coactivators. Trends Biochem Sci *25*, 15-19.

Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes Dev *13*, 2465-2477.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., and Magnuson, T. (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Mol Cell *6*, 1287-1295.

Burley, S. K., and Roeder, R. G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annu Rev Biochem *65*, 769-799.

Burke, L. J., and Baniahmad, A. (2000). Co-repressors 2000. Faseb J *14*, 1876-1888.

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.

Byvoet, P., Shepherd, G. R., Hardin, J. M., and Noland, B. J. (1972). The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. Arch Biochem Biophys *148*, 558-567.

Caballero, I. M., and Hendrich, B. (2005). MeCP2 in neurons: closing in on the causes of Rett syndrome. Hum Mol Genet *14 Spec No 1*, R19-26.

Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.

Cammas, F., Oulad-Abdelghani, M., Vonesch, J. L., Huss-Garcia, Y., Chambon, P., and Losson, R. (2002). Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. J Cell Sci *115*, 3439-3448.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev *14*, 155-164.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chambeyron, S., and Bickmore, W. A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes Dev *18*, 1119-1130.

Chakravarthy, S., Gundimella, S. K., Caron, C., Perche, P. Y., Pehrson, J. R., Khochbin, S., and Luger, K. (2005a). Structural characterization of the histone variant macroH2A. Mol Cell Biol *25*, 7616-7624.

Chakravarthy, S., Park, Y. J., Chodaparambil, J., Edayathumangalam, R. S., and Luger, K. (2005b). Structure and dynamic properties of nucleosome core particles. FEBS Lett *579*, 895-898.

Cheung, P., Allis, C. D., and Sassone-Corsi, P. (2000a). Signaling to chromatin through histone modifications. Cell *103*, 263-271.

Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M., and Allis, C. D. (2000b). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. Mol Cell *5*, 905-915.

Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev *15*, 3319-3329.

Chow, C. M., Georgiou, A., Szutorisz, H., Maia e Silva, A., Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. EMBO Rep *6*, 354-360.

Christmann, J. L., and Dahmus, M. E. (1981). Monoclonal antibody specific for calf thymus RNA polymerases IIO and IIA. J Biol Chem *256*, 11798-11803.

Chung, J. H., Whiteley, M., and Felsenfeld, G. (1993). A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell *74*, 505-514.

Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C. (2000). Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. Embo J *19*, 3714-3726.

Cloos, P. A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K. H., and Helin, K. (2006). The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. Nature *442*, 307-311.

Cockell, M., and Gasser, S. M. (1999). Nuclear compartments and gene regulation. Curr Opin Genet Dev *9*, 199-205.

Corcoran, A. E., Riddell, A., Krooshoop, D., and Venkitaraman, A. R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature *391*, 904-907.

Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat Struct Mol Biol *11*, 1037-1043.

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

Crane-Robinson, C., Hebbes, T. R., Clayton, A. L., and Thorne, A. W. (1997). Chromosomal mapping of core histone acetylation by immunoselection. Methods *12*, 48-56.

Cress, W. D., and Seto, E. (2000). Histone deacetylases, transcriptional control, and cancer. J Cell Physiol *184*, 1-16.

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

Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004). Histone deimination antagonizes arginine methylation. Cell *118*, 545-553.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell *111*, 185-196.

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.

Daban, J. R. (2000). Physical constraints in the condensation of eukaryotic chromosomes. Local concentration of DNA versus linear packing ratio in higher order chromatin structures. Biochemistry *39*, 3861-3866.

Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol *319*, 1097-1113.

de la Barre, A. E., Gerson, V., Gout, S., Creaven, M., Allis, C. D., and Dimitrov, S. (2000). Core histone N-termini play an essential role in mitotic chromosome condensation. Embo J *19*, 379-391.

de la Cruz, X., Lois, S., Sanchez-Molina, S., and Martinez-Balbas, M. A. (2005). Do protein motifs read the histone code? Bioessays *27*, 164-175.

DeKoter, R. P., Lee, H. J., and Singh, H. (2002). PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. Immunity *16*, 297-309.

DeKoter, R. P., and Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science *288*, 1439-1441.

DeKoter, R. P., Walsh, J. C., and Singh, H. (1998). PU.1 regulates both cytokine-dependent proliferation and differentiation of granulocyte/macrophage progenitors. Embo J *17*, 4456-4468.

Denker, A. E., Haas, A. R., Nicoll, S. B., and Tuan, R. S. (1999). Chondrogenic differentiation of murine C3H10T1/2 multipotential mesenchymal cells: I. Stimulation by bone morphogenetic protein-2 in high-density micromass cultures. Differentiation *64*, 67-76.

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.

De Souza, C. P., Osmani, A. H., Wu, L. P., Spotts, J. L., and Osmani, S. A. (2000). Mitotic histone H3 phosphorylation by the NIMA kinase in Aspergillus nidulans. Cell *102*, 293-302.

Dean, W., Lucifero, D., and Santos, F. (2005). DNA methylation in mammalian development and disease. Birth Defects Res C Embryo Today *75*, 98-111.

Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature *399*, 491-496.

Dillon, N., and Sabbattini, P. (2000). Functional gene expression domains: defining the functional unit of eukaryotic gene regulation. Bioessays *22*, 657-665.

Dillon, N. (2004). Heterochromatin structure and function. Biol Cell *96*, 631-637.

Dimitri, P., Corradini, N., Rossi, F., and Verni, F. (2005). The paradox of functional heterochromatin. Bioessays *27*, 29-41.

Djabali, M., Selleri, L., Parry, P., Bower, M., Young, B. D., and Evans, G. A. (1992). A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. Nat Genet *2*, 113-118.

Driscoll, M. C., Dobkin, C. S., and Alter, B. P. (1989). Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. Proc Natl Acad Sci U S A *86*, 7470-7474.

Drissen, R., Palstra, R. J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev *18*, 2485-2490.

Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in Drosophila pseudoobscura hybrids. Genetics *21*, 113–135.

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.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003). Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol *327*, 85-96.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science *306*, 1571-1573.

Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B. S., and Smale, S. T. (2002). A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. Genes Dev *16*, 2985-2990.

Dul, J. L., Argon, Y., Winkler, T., ten Boekel, E., Melchers, F., and Martensson, I. L. (1996). The murine VpreB1 and VpreB2 genes both encode a protein of the surrogate light chain and are co-expressed during B cell development. Eur J Immunol *26*, 906-913.

Dvir, A., Conaway, J. W., and Conaway, R. C. (2001). Mechanism of transcription initiation and promoter escape by RNA polymerase II. Curr Opin Genet Dev *11*, 209-214.

Dynan, W. S. (1989). Modularity in promoters and enhancers. Cell *58*, 1-4.

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

Ehrenhofer-Murray, A. E. (2004). Chromatin dynamics at DNA replication, transcription and repair. Eur J Biochem *271*, 2335-2349.

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.

Eilers, A. L., Billin, A. N., Liu, J., and Ayer, D. E. (1999). A 13-amino acid amphipathic alpha-helix is required for the functional interaction between the transcriptional repressor Mad1 and mSin3A. J Biol Chem *274*, 32750-32756.

Eisen, J. A., Sweder, K. S., and Hanawalt, P. C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Res *23*, 2715-2723.

Eberhard, D., Jimenez, G., Heavey, B., and Busslinger, M. (2000). Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. Embo J *19*, 2292-2303.

Eberharter, A., and Becker, P. B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO Rep *3*, 224-229.

Eissenberg, J. C., and Elgin, S. C. (2000). The HP1 protein family: getting a grip on chromatin. Curr Opin Genet Dev *10*, 204-210.

Ekwall, K., Nimmo, E. R., Javerzat, J. P., Borgstrom, B., Egel, R., Cranston, G., and Allshire, R. (1996). Mutations in the fission yeast silencing factors clr4+ and rik1+ disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. J Cell Sci *109 (Pt 11)*, 2637-2648.

Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M. A., Fiering, S., Enver, T., Martin, D. I., Kennedy, M., Keller, G., and Groudine, M. (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell *2*, 447-455.

Ernst, P., Hahm, K., Trinh, L., Davis, J. N., Roussel, M. F., Turck, C. W., and Smale, S. T. (1996). A potential role for Elf-1 in terminal transferase gene regulation. Mol Cell Biol *16*, 6121-6131.

Fan, J. Y., Rangasamy, D., Luger, K., and Tremethick, D. J. (2004). H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. Mol Cell *16*, 655-661.

Fang, J., Chen, T., Chadwick, B., Li, E., and Zhang, Y. (2004). Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. J Biol Chem *279*, 52812-52815.

Feng, Q., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2002a). Identification and functional characterization of the p66/p68 components of the MeCP1 complex. Mol Cell Biol *22*, 536-546.

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

Feng, Q., and Zhang, Y. (2001). The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. Genes Dev *15*, 827-832.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. Nature *355*, 219-224.

Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. A., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999). Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature *401*, 188-193.

Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H., and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116-1122.

Fisher, A. G. (2002). Cellular identity and lineage choice. Nat Rev Immunol *2*, 977-982.

Fisher, A. G., and Merkenschlager, M. (2002). Gene silencing, cell fate and nuclear organisation. Curr Opin Genet Dev *12*, 193-197.

Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature *438*, 1181-1185.

Flaus, A., Luger, K., Tan, S., and Richmond, T. J. (1996). Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. Proc Natl Acad Sci U S A *93*, 1370-1375.

Flaus, A., and Owen-Hughes, T. (2003). Dynamic properties of nucleosomes during thermal and ATP-driven mobilization. Mol Cell Biol *23*, 7767-7779.

Flaus, A., and Owen-Hughes, T. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? Curr Opin Genet Dev *14*, 165-173.

Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hughes, T. (2004). Sin mutations alter inherent nucleosome mobility. Embo J *23*, 343-353.

Fitzsimmons, D., Hodsdon, W., Wheat, W., Maira, S. M., Wasylyk, B., and Hagman, J. (1996). Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. Genes Dev *10*, 2198-2211.

Forrester, W. C., Epner, E., Driscoll, M. C., Enver, T., Brice, M., Papayannopoulou, T., and Groudine, M. (1990). A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. Genes Dev *4*, 1637-1649.

Forrester, W. C., Thompson, C., Elder, J. T., and Groudine, M. (1986). A developmentally stable chromatin structure in the human beta-globin gene cluster. Proc Natl Acad Sci U S A *83*, 1359-1363.

Forsberg, E. C., Downs, K. M., Christensen, H. M., Im, H., Nuzzi, P. A., and Bresnick, E. H. (2000). Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. Proc Natl Acad Sci U S A *97*, 14494-14499.

Francis, N. J., Kingston, R. E., and Woodcock, C. L. (2004). Chromatin compaction by a polycomb group protein complex. Science *306*, 1574-1577.

Freitas, M. A., Sklenar, A. R., and Parthun, M. R. (2004). Application of mass spectrometry to the identification and quantification of histone post-translational modifications. J Cell Biochem *92*, 691-700.

Freitag, M., and Selker, E. U. (2005). Controlling DNA methylation: many roads to one modification. Curr Opin Genet Dev *15*, 191-199.

Frye, R. A. (2000). Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. Biochem Biophys Res Commun *273*, 793-798.

Fryer, C. J., and Archer, T. K. (1998). Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. Nature *393*, 88-91.

Fuks, F. (2005). DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev *15*, 490-495.

Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J Biol Chem *278*, 4035-4040.

Fuks, F., Burgers, W. A., Godin, N., Kasai, M., and Kouzarides, T. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. Embo J *20*, 2536-2544.

Gardner, R. L., and Beddington, R. S. (1988). Multi-lineage 'stem' cells in the mammalian embryo. J Cell Sci Suppl *10*, 11-27.

Gartenberg, M. R. (2000). The Sir proteins of Saccharomyces cerevisiae: mediators of transcriptional silencing and much more. Curr Opin Microbiol *3*, 132-137.

Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N., Woodcock, C. L., Wade, P. A., and Hansen, J. C. (2003). Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. J Biol Chem *278*, 32181-32188.

Georgopoulos, K. (2002). Haematopoietic cell-fate decisions, chromatin regulation and ikaros. Nat Rev Immunol *2*, 162-174.

Georgopoulos, K., Bigby, M., Wang, J. H., Molnar, A., Wu, P., Winandy, S., and Sharpe, A. (1994). The Ikaros gene is required for the development of all lymphoid lineages. Cell *79*, 143-156.

Giet, R., and Glover, D. M. (2001). Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. J Cell Biol *152*, 669-682.

Gisler, R., and Sigvardsson, M. (2002). The human V-preB promoter is a target for coordinated activation by early B cell factor and E47. J Immunol *168*, 5130-5138.

Glover, D. M., Leibowitz, M. H., McLean, D. A., and Parry, H. (1995). Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. Cell *81*, 95-105.

Goodrich, J. A., and Tjian, R. (1994). TBP-TAF complexes: selectivity factors for eukaryotic transcription. Curr Opin Cell Biol *6*, 403-409.

Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T., and Inagaki, M. (1999). Identification of a

novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. J Biol Chem *274*, 25543-25549.

Grabstein, K. H., Waldschmidt, T. J., Finkelman, F. D., Hess, B. W., Alpert, A. R., Boiani, N. E., Namen, A. E., and Morrissey, P. J. (1993). Inhibition of murine B and T lymphopoiesis in vivo by an anti-interleukin 7 monoclonal antibody. J Exp Med *178*, 257-264.

Gray, S. G., Iglesias, A. H., Lizcano, F., Villanueva, R., Camelo, S., Jingu, H., Teh, B. T., Koibuchi, N., Chin, W. W., Kokkotou, E., and Dangond, F. (2005). Functional characterization of JMJD2A, a histone deacetylase- and retinoblastoma-binding protein. J Biol Chem *280*, 28507-28518.

Greenbaum, S., and Zhuang, Y. (2002). Identification of E2A target genes in B lymphocyte development by using a gene tagging-based chromatin immunoprecipitation system. Proc Natl Acad Sci U S A *99*, 15030-15035.

Gregoretti, I. V., Lee, Y. M., and Goodson, H. V. (2004). Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. J Mol Biol *338*, 17-31.

Grewal, S. I., and Elgin, S. C. (2002). Heterochromatin: new possibilities for the inheritance of structure. Curr Opin Genet Dev *12*, 178-187.

Grewal, S. I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science *301*, 798-802.

Griffiths, S. D., Goodhead, D. T., Marsden, S. J., Wright, E. G., Krajewski, S., Reed, J. C., Korsmeyer, S. J., and Greaves, M. (1994). Interleukin 7-dependent B lymphocyte precursor cells are ultrasensitive to apoptosis. J Exp Med *179*, 1789-1797.

Grosveld, F., van Assendelft, G. B., Greaves, D. R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell *51*, 975-985.

Grozinger, C. M., Chao, E. D., Blackwell, H. E., Moazed, D., and Schreiber, S. L. (2001). Identification of a class of small molecule inhibitors of the sirtuin family of NAD-dependent deacetylases by phenotypic screening. J Biol Chem *276*, 38837-38843.

Grozinger, C. M., and Schreiber, S. L. (2000). Regulation of histone deacetylase 4 and 5 and transcriptional activity by 14-3-3-dependent cellular localization. Proc Natl Acad Sci U S A *97*, 7835-7840.

Gu, W., Szauter, P., and Lucchesi, J. C. (1998). Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of Drosophila melanogaster. Dev Genet *22*, 56-64.

Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. (1992). The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell *71*, 701-708.

Guarente, L. (2000). Sir2 links chromatin silencing, metabolism, and aging. Genes Dev *14*, 1021-1026.

Gurley, L. R., D'Anna, J. A., Barham, S. S., Deaven, L. L., and Tobey, R. A. (1978). Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. Eur J Biochem *84*, 1-15.

Hagman, J., Belanger, C., Travis, A., Turck, C. W., and Grosschedl, R. (1993). Cloning and functional characterization of early B-cell factor, a regulator of lymphocyte-specific gene expression. Genes Dev *7*, 760-773.

Hagman, J., Gutch, M. J., Lin, H., and Grosschedl, R. (1995). EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains. Embo J *14*, 2907-2916.

Hagman, J., Travis, A., and Grosschedl, R. (1991). A novel lineage-specific nuclear factor regulates mb-1 gene transcription at the early stages of B cell differentiation. Embo J *10*, 3409-3417.

Hall, I. M., Shankaranarayana, G. D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S. I. (2002). Establishment and maintenance of a heterochromatin domain. Science *297*, 2232-2237.

Hakimi, M. A., Bochar, D. A., Chenoweth, J., Lane, W. S., Mandel, G., and Shiekhattar, R. (2002). A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. Proc Natl Acad Sci U S A *99*, 7420-7425.

Hamiche, A., Kang, J. G., Dennis, C., Xiao, H., and Wu, C. (2001). Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. Proc Natl Acad Sci U S A *98*, 14316-14321.

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

Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J. M., Jones, P. L., Sif, S., and El-Osta, A. (2005). Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nat Genet *37*, 254-264.

Harland, R. (2000). Neural induction. Curr Opin Genet Dev *10*, 357-362.

Hassan, A. H., Neely, K. E., and Workman, J. L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. Cell *104*, 817-827.

Hassan, A. H., Prochasson, P., Neely, K. E., Galasinski, S. C., Chandy, M., Carrozza, M. J., and Workman, J. L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell *111*, 369-379.

Hatzis, P., and Talianidis, I. (2002). Dynamics of enhancer-promoter communication during differentiation-induced gene activation. Mol Cell *10*, 1467-1477.

Hayes, J. J., Clark, D. J., and Wolffe, A. P. (1991). Histone contributions to the structure of DNA in the nucleosome. Proc Natl Acad Sci U S A *88*, 6829-6833.

Hayes, J. J., and Hansen, J. C. (2001). Nucleosomes and the chromatin fiber. Curr Opin Genet Dev *11*, 124-129.

Heard, E. (2004). Recent advances in X-chromosome inactivation. Curr Opin Cell Biol *16*, 247-255.

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C. D., and Spector, D. L. (2001). Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell *107*, 727-738.

Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. Embo J *7*, 1395-1402.

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

Heitz, E. (1928). Das heterochromatin der moose. IJahrb Wiss Botanik *69*, 762-818.

Hendzel, M. J., Nishioka, W. K., Raymond, Y., Allis, C. D., Bazett-Jones, D. P., and Th'ng, J. P. (1998). Chromatin condensation is not associated with apoptosis. J Biol Chem *273*, 24470-24478.

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

Henikoff, S. (1993). Transcriptional activator components and poxvirus DNA-dependent ATPases comprise a single family. Trends Biochem Sci *18*, 291-292.

Henikoff, S., and Ahmad, K. (2005). Assembly of variant histones into chromatin. Annu Rev Cell Dev Biol *21*, 133-153.

Hermann, A., Schmitt, S., and Jeltsch, A. (2003). The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. J Biol Chem *278*, 31717-31721.

Hirota, T., Lipp, J. J., Toh, B. H., and Peters, J. M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. Nature *438*, 1176-1180.

Holstege, F. C., van der Vliet, P. C., and Timmers, H. T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. Embo J *15*, 1666-1677.

Horike, S., Cai, S., Miyano, M., Cheng, J. F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet *37*, 31-40.

Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F., *et al.* (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. Cell *102*, 279-291.

Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S., and Halazonetis, T. D. (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. Nature *432*, 406-411.

Iizuka, M., and Smith, M. M. (2003). Functional consequences of histone modifications. Curr Opin Genet Dev *13*, 154-160.

Jackson, J. P., Lindroth, A. M., Cao, X., and Jacobsen, S. E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556-560.

Jiang, Y., Liu, M., Spencer, C. A., and Price, D. H. (2004). Involvement of transcription termination factor 2 in mitotic repression of transcription elongation. Mol Cell *14*, 375-385.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol Life Sci *54*, 80-93.

Jones, R. S., and Gelbart, W. M. (1993). The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. Mol Cell Biol *13*, 6357-6366.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-1080.

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

Jones, J. C., Phatnani, H. P., Haystead, T. A., MacDonald, J. A., Alam, S. M., and Greenleaf, A. L. (2004). C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. J Biol Chem *279*, 24957-24964.

Karasuyama, H., Kudo, A., and Melchers, F. (1990). The proteins encoded by the VpreB and lambda 5 pre-B cell-specific genes can associate with each other and with mu heavy chain. J Exp Med *172*, 969-972.

Karasuyama, H., Rolink, A., Shinkai, Y., Young, F., Alt, F. W., and Melchers, F. (1994). The expression of Vpre-B/lambda 5 surrogate light chain in early bone marrow precursor B cells of normal and B cell-deficient mutant mice. Cell *77*, 133-143.

Kardassis, D., Tzameli, I., Hadzopoulou-Cladaras, M., Talianidis, I., and Zannis, V. (1997). Distal apolipoprotein C-III regulatory elements F to J act as a general modular enhancer for proximal promoters that contain hormone response elements. Synergism between hepatic nuclear factor-4 molecules bound to the proximal promoter and distal enhancer sites. Arterioscler Thromb Vasc Biol *17*, 222-232.

Katagiri, T., Yamaguchi, A., Ikeda, T., Yoshiki, S., Wozney, J. M., Rosen, V., Wang, E. A., Tanaka, H., Omura, S., and Suda, T. (1990). The non-osteogenic mouse pluripotent cell line, C3H10T1/2, is induced to differentiate into osteoblastic

cells by recombinant human bone morphogenetic protein-2. Biochem Biophys Res Commun *172*, 295-299.

Katoh, M., and Katoh, M. (2004). Identification and characterization of JMJD2 family genes in silico. Int J Oncol *24*, 1623-1628.

Kee, B. L., and Murre, C. (1998). Induction of early B cell factor (EBF) and multiple B lineage genes by the basic helix-loop-helix transcription factor E12. J Exp Med *188*, 699-713.

Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. Cell *64*, 941-950.

Khochbin, S., Verdel, A., Lemercier, C., and Seigneurin-Berny, D. (2001). Functional significance of histone deacetylase diversity. Curr Opin Genet Dev *11*, 162-166.

Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature *442*, 312-316.

Kiekhaefer, C. M., Grass, J. A., Johnson, K. D., Boyer, M. E., and Bresnick, E. H. (2002). Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain. Proc Natl Acad Sci U S A *99*, 14309-14314.

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

Kornberg, R. D. (1999). Eukaryotic transcriptional control. Trends Cell Biol *9*, M46-49.

Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2004). Histone code modifications on pluripotential nuclei of reprogrammed somatic cells. Mol Cell Biol *24*, 5710-5720.

Kingston, R. E., and Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev *13*, 2339-2352.

Kioussis, D., Vanin, E., deLange, T., Flavell, R. A., and Grosveld, F. G. (1983). Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. Nature *306*, 662-666.

Klochendler-Yeivin, A., Fiette, L., Barra, J., Muchardt, C., Babinet, C., and Yaniv, M. (2000). The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression. EMBO Rep *1*, 500-506.

Knoepfler, P. S., and Eisenman, R. N. (1999). Sin meets NuRD and other tails of repression. Cell *99*, 447-450.

Kollias, G., Wrighton, N., Hurst, J., and Grosveld, F. (1986). Regulated expression of human A gamma-, beta-, and hybrid gamma beta-globin genes in transgenic mice: manipulation of the developmental expression patterns. Cell *46*, 89-94.

Kouzarides, T. (2002). Histone methylation in transcriptional control. Curr Opin Genet Dev *12*, 198-209.

Krebs, J. E., Kuo, M. H., Allis, C. D., and Peterson, C. L. (1999). Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. Genes Dev *13*, 1412-1421.

Kristjuhan, A., Walker, J., Suka, N., Grunstein, M., Roberts, D., Cairns, B. R., and Svejstrup, J. Q. (2002). Transcriptional inhibition of genes with severe histone h3 hypoacetylation in the coding region. Mol Cell *10*, 925-933.

Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., *et al.* (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell *11*, 721-729.

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., *et al.* (2003b). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Kondo, M., Weissman, I. L., and Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell *91*, 661-672.

Kozmik, Z., Wang, S., Dorfler, P., Adams, B., and Busslinger, M. (1992). The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. Mol Cell Biol *12*, 2662-2672.

Kudo, A., and Melchers, F. (1987). A second gene, VpreB in the lambda 5 locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. Embo J *6*, 2267-2272.

Kuhn, E. J., Hart, C. M., and Geyer, P. K. (2004). Studies of the role of the Drosophila scs and scs' insulators in defining boundaries of a chromosome puff. Mol Cell Biol *24*, 1470-1480.

Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature *383*, 269-272.

Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E., and Allis, C. D. (1998). Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. Genes Dev *12*, 627-639.

Kuras, L., and Struhl, K. (1999). Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. Nature *399*, 609-613.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev *16*, 2893-2905.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lacoste, N., Utley, R. T., Hunter, J. M., Poirier, G. G., and Cote, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J Biol Chem *277*, 30421-30424.

Laherty, C. D., Billin, A. N., Lavinsky, R. M., Yochum, G. S., Bush, A. C., Sun, J. M., Mullen, T. M., Davie, J. R., Rose, D. W., Glass, C. K., *et al.* (1998). SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. Mol Cell *2*, 33-42.

Laribee, R. N., Krogan, N. J., Xiao, T., Shibata, Y., Hughes, T. R., Greenblatt, J. F., and Strahl, B. D. (2005). BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. Curr Biol *15*, 1487-1493.

Laybourn, P. J., and Dahmus, M. E. (1989). Transcription-dependent structural changes in the C-terminal domain of mammalian RNA polymerase subunit IIa/o. J Biol Chem *264*, 6693-6698.

Lee, H. L., and Archer, T. K. (1998). Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. Embo J *17*, 1454-1466.

Lee, J. H., Hart, S. R., and Skalnik, D. G. (2004). Histone deacetylase activity is required for embryonic stem cell differentiation. Genesis *38*, 32-38.

Lee, H. L., and Archer, T. K. (1994). Nucleosome-mediated disruption of transcription factor-chromatin initiation complexes at the mouse mammary tumor virus long terminal repeat in vivo. Mol Cell Biol *14*, 32-41.

Lee, J. M., and Greenleaf, A. L. (1997). Modulation of RNA polymerase II elongation efficiency by C-terminal heptapeptide repeat domain kinase I. J Biol Chem *272*, 10990-10993.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

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, B., Howe, L., Anderson, S., Yates, J. R., 3rd, and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *278*, 8897-8903.

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, J., Moazed, D., and Gygi, S. P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. J Biol Chem *277*, 49383-49388

Li, R., Pei, H., and Watson, D. K. (2000). Regulation of Ets function by protein - protein interactions. Oncogene *19*, 6514-6523.

Li, Y. S., Hayakawa, K., and Hardy, R. R. (1993). The regulated expression of B lineage associated genes during B cell differentiation in bone marrow and fetal liver. J Exp Med *178*, 951-960.

Lin, H., and Grosschedl, R. (1995). Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature *376*, 263-267.

Litt, M. D., Simpson, M., Recillas-Targa, F., Prioleau, M. N., and Felsenfeld, G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. Embo J *20*, 2224-2235.

Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N., and Rando, O. J. (2005). Single-nucleosome mapping of histone modifications in S. cerevisiae. PLoS Biol *3*, e328.

Lo, K., Landau, N. R., and Smale, S. T. (1991). LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. Mol Cell Biol *11*, 5229-5243.

Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc Natl Acad Sci U S A *88*, 10004-10008.

Lu, R., Medina, K. L., Lancki, D. W., and Singh, H. (2003). IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. Genes Dev *17*, 1703-1708.

Lucchesi, J. C. (1998). Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation. Curr Opin Genet Dev *8*, 179-184.

Luger, K. (2003). Structure and dynamic behavior of nucleosomes. Curr Opin Genet Dev *13*, 127-135.

Luger, K., and Hansen, J. C. (2005). Nucleosome and chromatin fiber dynamics. Curr Opin Struct Biol *15*, 188-196.

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.

Luger, K., and Richmond, T. J. (1998). DNA binding within the nucleosome core. Curr Opin Struct Biol *8*, 33-40.

Lundgren, M., Chow, C., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. (2000). Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell *103*, 733-743.

Lusser, A., Brosch, G., Loidl, A., Haas, H., and Loidl, P. (1997). Identification of maize histone deacetylase HD2 as an acidic nucleolar phosphoprotein. Science *277*, 88-91.

Mackarehtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P., and Lemischka, I. R. (1995). Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. Immunity *3*, 147-161.

McKinsey, T. A., Zhang, C. L., and Olson, E. N. (2001). Control of muscle development by dueling HATs and HDACs. Curr Opin Genet Dev *11*, 497-504.

Magram, J., Chada, K., and Costantini, F. (1985). Developmental regulation of a cloned adult beta-globin gene in transgenic mice. Nature *315*, 338-340.

Mahadevan, L. C., Willis, A. C., and Barratt, M. J. (1991). Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. Cell *65*, 775-783.

Mai, S., and Martensson, I. L. (1995). The c-myc protein represses the lambda 5 and TdT initiators. Nucleic Acids Res *23*, 1-9.

Manuelidis, L. (1990). A view of interphase chromosomes. Science *250*, 1533-1540.

Mariani, M. R., Carpaneto, E. M., Ulivi, M., Allfrey, V. G., and Boffa, L. C. (2003). Correlation between butyrate-induced histone hyperacetylation turn-over and c-myc expression. J Steroid Biochem Mol Biol *86*, 167-171.

Marshall, N. F., and Price, D. H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J Biol Chem *270*, 12335-12338.

Martensson, A., and Martensson, I. L. (1997). Early B cell factor binds to a site critical for lambda5 core enhancer activity. Eur J Immunol *27*, 315-320.

Martinez, E. (2002). Multi-protein complexes in eukaryotic gene transcription. Plant Mol Biol *50*, 925-947.

Mateescu, B., England, P., Halgand, F., Yaniv, M., and Muchardt, C. (2004). Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. EMBO Rep *5*, 490-496.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. Embo J *18*, 4261-4269.

Mersfelder, E. L., and Parthun, M. R. (2006). The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. Nucleic Acids Res *34*, 2653-2662.

Medina, K. L., Pongubala, J. M., Reddy, K. L., Lancki, D. W., Dekoter, R., Kieslinger, M., Grosschedl, R., and Singh, H. (2004). Assembling a gene regulatory network for specification of the B cell fate. Dev Cell *7*, 607-617.

Meehan, R. R. (2003). DNA methylation in animal development. Semin Cell Dev Biol *14*, 53-65.

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

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

Meshorer, E., Yellajoshula, D., George, E., Scambler, P. J., Brown, D. T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell *10*, 105-116.

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.

Minaee, S., Farmer, D., Georgiou, A., Sabbattini, P., Webster, Z., Chow, C. M., and Dillon, N. (2005). Mapping and functional analysis of regulatory sequences in the mouse lambda5-VpreB1 domain. Mol Immunol *42*, 1283-1292.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Mombaerts, P., Iacomini, J., Johnson, R. S., Herrup, K., Tonegawa, S., and Papaioannou, V. E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. Cell *68*, 869-877.

Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Morillon, A., Karabetsou, N., O'Sullivan, J., Kent, N., Proudfoot, N., and Mellor, J. (2003). Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. Cell *115*, 425-435.

Mueller, C. L., and Jaehning, J. A. (2002). Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. Mol Cell Biol *22*, 1971-1980.

Muller, H. J. (1930). Types of visible variations produced by X-rays in Drosophila melanogaster. J Genet *22*, 299–234.

Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., and Simon, J. A. (2002). Histone

methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell *111*, 197-208.

Murnion, M. E., Adams, R. R., Callister, D. M., Allis, C. D., Earnshaw, W. C., and Swedlow, J. R. (2001). Chromatin-associated protein phosphatase 1 regulates aurora-B and histone H3 phosphorylation. J Biol Chem *276*, 26656-26665.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat Genet *36*, 889-893.

Muth, V., Nadaud, S., Grummt, I., and Voit, R. (2001). Acetylation of TAF(I)68, a subunit of TIF-IB/SL1, activates RNA polymerase I transcription. Embo J *20*, 1353-1362.

Munoz-Sanjuan, I., and Brivanlou, A. H. (2002). Neural induction, the default model and embryonic stem cells. Nat Rev Neurosci *3*, 271-280.

Murre, C., Bain, G., van Dijk, M. A., Engel, I., Furnari, B. A., Massari, M. E., Matthews, J. R., Quong, M. W., Rivera, R. R., and Stuiver, M. H. (1994). Structure and function of helix-loop-helix proteins. Biochim Biophys Acta *1218*, 129-135.

Muthurajan, U. M., Bao, Y., Forsberg, L. J., Edayathumangalam, R. S., Dyer, P. N., White, C. L., and Luger, K. (2004). Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions. Embo J *23*, 260-271.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.

Murre, C., Voronova, A., and Baltimore, D. (1991). B-cell- and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. Mol Cell Biol *11*, 1156-1160.

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.

Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

Ng, H. H., Robert, F., Young, R. A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell *11*, 709-719.

Nichogiannopoulou, A., Trevisan, M., Neben, S., Friedrich, C., and Georgopoulos, K. (1999). Defects in hemopoietic stem cell activity in Ikaros mutant mice. J Exp Med *190*, 1201-1214.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Nielsen, S. J., Schneider, R., Bauer, U. M., Bannister, A. J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R. E., and Kouzarides, T. (2001). Rb targets histone H3 methylation and HP1 to promoters. Nature *412*, 561-565.

Nightingale, K.P., O'Neill, L.P., Turner, B.M. (2006). Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. Curr Opin Genet Dev. 16(2):125-36.

Nishida, H., Suzuki, T., Kondo, S., Miura, H., Fujimura, Y., and Hayashizaki, Y. (2006). Histone H3 acetylated at lysine 9 in promoter is associated with low nucleosome density in the vicinity of transcription start site in human cell. Chromosome Res *14*, 203-211.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C. D., Tempst, P., and Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes Dev *16*, 479-489.

Nislow, C., Ray, E., and Pillus, L. (1997). SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. Mol Biol Cell *8*, 2421-2436.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev *12*, 2048-2060.

Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S. I. (2004). RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. Nat Genet *36*, 1174-1180.

Nussinov, R. (1990). Sequence signals in eukaryotic upstream regions. Crit Rev Biochem Mol Biol *25*, 185-224.

Nutt, S. L., Heavey, B., Rolink, A. G., and Busslinger, M. (1999). Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature *401*, 556-562.

Nutt, S. L., Morrison, A. M., Dorfler, P., Rolink, A., and Busslinger, M. (1998). Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. Embo J *17*, 2319-2333.

Nutt, S. L., Urbanek, P., Rolink, A., and Busslinger, M. (1997). Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus. Genes Dev *11*, 476-491.

Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D. M., and Nakatani, Y. (2002). A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science *296*, 1132-1136.

Ohtsuki, S., and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. Genes Dev *12*, 3325-3330.

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.

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

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

O'Neill, L. P., and Turner, B. M. (1995). Histone H4 acetylation distinguishes coding regions of the human genome from heterochromatin in a differentiation-dependent but transcription- independent manner. Embo J *14*, 3946-3957.

O'Riordan, M., and Grosschedl, R. (1999). Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A. Immunity *11*, 21-31.

Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D. M., and Nakatani, Y. (2002). A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science *296*, 1132-1136.

Okabe, T., Bauer, S. R., and Kudo, A. (1992). Pre-B lymphocyte-specific transcriptional control of the mouse VpreB gene. Eur J Immunol *22*, 31-36.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S. C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., *et al.* (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature *436*, 207-213.

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

Orphanides, G., and Reinberg, D. (2002). A unified theory of gene expression. Cell *108*, 439-451.

Otero, G., Fellows, J., Li, Y., de Bizemont, T., Dirac, A. M., Gustafsson, C. M., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. (1999). Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. Mol Cell *3*, 109-118.

Palladino, F., Laroche, T., Gilson, E., Axelrod, A., Pillus, L., and Gasser, S. M. (1993). SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. Cell *75*, 543-555.

Panne, D., Raleigh, E. A., and Bickle, T. A. (1999). The McrBC endonuclease translocates DNA in a reaction dependent on GTP hydrolysis. J Mol Biol *290*, 49-60.

Parekh, B. S., and Maniatis, T. (1999). Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-beta promoter. Mol Cell *3*, 125-129.

Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., *et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell *12*, 1577-1589.

Peters, A. H., Mermoud, J. E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., and Jenuwein, T. (2002). Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. Nat Genet *30*, 77-80.

Pham, A. D., and Sauer, F. (2000). Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila. Science *289*, 2357-2360.

Philipsen, S., Talbot, D., Fraser, P., and Grosveld, F. (1990). The beta-globin dominant control region: hypersensitive site 2. Embo J *9*, 2159-2167.

Pickart, C. M. (2001a). Mechanisms underlying ubiquitination. Annu Rev Biochem *70*, 503-533.

Pickart, C. M. (2001b). Ubiquitin enters the new millennium. Mol Cell *8*, 499-504.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science *300*, 131-135.

Pokholok, D. K., Hannett, N. M., and Young, R. A. (2002). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. Mol Cell *9*, 799-809.

Preuss, U., Landsberg, G., and Scheidtmann, K. H. (2003). Novel mitosis-specific phosphorylation of histone H3 at Thr11 mediated by Dlk/ZIP kinase. Nucleic Acids Res *31*, 878-885.

Prioleau, M. N., Nony, P., Simpson, M., and Felsenfeld, G. (1999). An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. Embo J *18*, 4035-4048.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev *15*, 1613-1618.

Proudfoot, N. J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem Sci *14*, 105-110.

Pruzina, S., Hanscombe, O., Whyatt, D., Grosveld, F., and Philipsen, S. (1991). Hypersensitive site 4 of the human beta globin locus control region. Nucleic Acids Res *19*, 1413-1419.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. Nature *322*, 697-701.

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

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., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

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., and Jenuwein, T. (2000).

Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Reese, J. C. (2003). Basal transcription factors. Curr Opin Genet Dev *13*, 114-118.

Reik, A., Telling, A., Zitnik, G., Cimbora, D., Epner, E., and Groudine, M. (1998). The locus control region is necessary for gene expression in the human beta-globin locus but not the maintenance of an open chromatin structure in erythroid cells. Mol Cell Biol *18*, 5992-6000.

Reines, D., Conaway, R. C., and Conaway, J. W. (1999). Mechanism and regulation of transcriptional elongation by RNA polymerase II. Curr Opin Cell Biol *11*, 342-346.

Reinke, H., and Horz, W. (2003). Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. Mol Cell *11*, 1599-1607.

Rice, J. C., Briggs, S. D., Ueberheide, B., Barber, C. M., Shabanowitz, J., Hunt, D. F., Shinkai, Y., and Allis, C. D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell *12*, 1591-1598.

Richard-Foy, H., Sistare, F. D., Riegel, A. T., Simons, S. S., Jr., and Hager, G. L. (1987). Mechanism of dexamethasone 21-mesylate antiglucocorticoid action: II. Receptor-antiglucocorticoid complexes do not interact productively with mouse mammary tumor virus long terminal repeat chromatin. Mol Endocrinol *1*, 659-665.

Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L., and Wolffe, A. P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat Genet *25*, 338-342.

Richardson, J. P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim Biophys Acta *1577*, 251-260.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet *38*, 413-443.

Roeder, R. G., and Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature *224*, 234-237.

Rohwedel, J., Kleppisch, T., Pich, U., Guan, K., Jin, S., Zuschratter, W., Hopf, C., Hoch, W., Hescheler, J., Witzemann, V., and Wobus, A. M. (1998). Formation of postsynaptic-like membranes during differentiation of embryonic stem cells in vitro. Exp Cell Res *239*, 214-225.

Rolink, A. G., Nutt, S. L., Melchers, F., and Busslinger, M. (1999). Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. Nature *401*, 603-606.

Rosenberg, N. (1994). Abl-mediated transformation, immunoglobulin gene rearrangements and arrest of B lymphocyte differentiation. Semin Cancer Biol *5*, 95-102.

Sabbattini, P., and Dillon, N. (2005). The lambda5-VpreB1 locus--a model system for studying gene regulation during early B cell development. Semin Immunol *17*, 121-127.

Sabbattini, P., Lundgren, M., Georgiou, A., Chow, C.-M., Warnes, G., and Dillon, N. (2001). Binding of Ikaros to the λ5 promoter silences transcription through a mechanism that does not require heterochromatin formation. EMBO J *20*, 2812-2822.

Sabbattini, P., Georgiou, A., Sinclair, C., and Dillon, N. (1999). Analysis of mice with single copies and multiple copies of transgenes reveals a novel arrangement for the l5-VpreB1 locus control region. Mol Cell Biol *19*, 671-679.

Sakaguchi, N., and Melchers, F. (1986). Lambda 5, a new light-chain-related locus selectively expressed in pre-B lymphocytes. Nature *324*, 579-582.

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

Santos-Rosa, H., Schneider, R., Bernstein, B. E., Karabetsou, N., Morillon, A., Weise, C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2003). Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. Mol Cell *12*, 1325-1332.

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.

Schatz, D. G., Oettinger, M. A., and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. Cell *59*, 1035-1048.

Schebesta, M., Pfeffer, P. L., and Busslinger, M. (2002). Control of pre-BCR signaling by Pax5-dependent activation of the BLNK gene. Immunity *17*, 473-485.

Schubeler, D., Groudine, M., and Bender, M. A. (2001). The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. Proc Natl Acad Sci U S A *98*, 11432-11437.

Schlichter, A., and Cairns, B. R. (2005). Histone trimethylation by Set1 is coordinated by the RRM, autoinhibitory, and catalytic domains. Embo J *24*, 1222-1231.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev *18*, 1251-1262.

Schubeler, D., Groudine, M., and Bender, M. A. (2001). The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. Proc Natl Acad Sci U S A *98*, 11432-11437.

Selby, C. P., and Sancar, A. (1993). Molecular mechanism of transcription-repair coupling. Science *260*, 53-58.

Sekinger, E. A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol Cell *18*, 735-748.

Shen, C. P., and Kadesch, T. (1995). B-cell-specific DNA binding by an E47 homodimer. Mol Cell Biol *15*, 4518-4524.

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

Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003). The RNA polymerase II elongation complex. Annu Rev Biochem *72*, 693-715.

Shinkai, Y., Rathbun, G., Lam, K. P., Oltz, E. M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A. M., and et al. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell *68*, 855-867.

Shcherbik, N., and Haines, D. S. (2004). Ub on the move. J Cell Biochem *93*, 11-19.

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

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

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell *4*, 481-495.

Sigvardsson, M. (2000). Overlapping expression of early B-cell factor and basic helix-loop-helix proteins as a mechanism to dictate B-lineage-specific activity of the lambda5 promoter. Mol Cell Biol *20*, 3640-3654.

Sigvardsson, M., Clark, D. R., Fitzsimmons, D., Doyle, M., Akerblad, P., Breslin, T., Bilke, S., Li, R., Yeamans, C., Zhang, G., and Hagman, J. (2002). Early B-cell factor, E2A, and Pax-5 cooperate to activate the early B cell-specific mb-1 promoter. Mol Cell Biol *22*, 8539-8551.

Sigvardsson, M., O'Riordan, M., and Grosschedl, R. (1997). EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes. Immunity *7*, 25-36.

Sims, R. J., 3rd, Chen, C. F., Santos-Rosa, H., Kouzarides, T., Patel, S. S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem *280*, 41789-41792.

Sims, R. J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev *18*, 2437-2468.

Singh, H., Medina, K. L., and Pongubala, J. M. (2005). Contingent gene regulatory networks and B cell fate specification. Proc Natl Acad Sci U S A *102*, 4949-4953.

Sitnicka, E., Brakebusch, C., Martensson, I. L., Svensson, M., Agace, W. W., Sigvardsson, M., Buza-Vidas, N., Bryder, D., Cilio, C. M., Ahlenius, H., *et al.* (2003). Complementary signaling through flt3 and interleukin-7 receptor alpha is indispensable for fetal and adult B cell genesis. J Exp Med *198*, 1495-1506.

Smale, S. T. (2003). The establishment and maintenance of lymphocyte identity through gene silencing. Nat Immunol *4*, 607-615.

Smith, E. R., Allis, C. D., and Lucchesi, J. C. (2001). Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males. J Biol Chem *276*, 31483-31486.

Smith, E. M., Gisler, R., and Sigvardsson, M. (2002). Cloning and characterization of a promoter flanking the early B cell factor (EBF) gene indicates roles for E-proteins and autoregulation in the control of EBF expression. J Immunol *169*, 261-270.

Smith, C. L., Archer, T. K., Hamlin-Green, G., and Hager, G. L. (1993). Newly expressed progesterone receptor cannot activate stable, replicated mouse mammary tumor virus templates but acquires transactivation potential upon continuous expression. Proc Natl Acad Sci U S A *90*, 11202-11206.

Solari, F., Bateman, A., and Ahringer, J. (1999). The Caenorhabditis elegans genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. Development *126*, 2483-2494.

Souabni, A., Cobaleda, C., Schebesta, M., and Busslinger, M. (2002). Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. Immunity *17*, 781-793.

Solomon, M. J., Larsen, P. L., and Varshavsky, A. (1988). Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell *53*, 937-947.

Soutoglou, E., and Talianidis, I. (2002). Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. Science *295*, 1901-1904.

Sudo, T., Nishikawa, S., Ohno, N., Akiyama, N., Tamakoshi, M., Yoshida, H., and Nishikawa, S. (1993). Expression and function of the interleukin 7 receptor in murine lymphocytes. Proc Natl Acad Sci U S A *90*, 9125-9129.

Spilianakis, C. G., and Flavell, R. A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat Immunol *5*, 1017-1027.

Spotswood, H. T., and Turner, B. M. (2002). An increasingly complex code. J Clin Invest *110*, 577-582.

Stern, M., Jensen, R., and Herskowitz, I. (1984). Five SWI genes are required for expression of the HO gene in yeast. J Mol Biol *178*, 853-868.

Stewart, F. J., and Raleigh, E. A. (1998). Dependence of McrBC cleavage on distance between recognition elements. Biol Chem *379*, 611-616.

Stopka, T., and Skoultchi, A. I. (2003). The ISWI ATPase Snf2h is required for early mouse development. Proc Natl Acad Sci U S A *100*, 14097-14102.

Straight, A. F., Shou, W., Dowd, G. J., Turck, C. W., Deshaies, R. J., Johnson, A. D., and Moazed, D. (1999). Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. Cell *97*, 245-256.

Struhl, K. (1999). Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell *98*, 1-4.

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.

Sutherland, E., Coe, L., and Raleigh, E. A. (1992). McrBC: a multisubunit GTP-dependent restriction endonuclease. J Mol Biol *225*, 327-348.

Sutherland, H. G., Martin, D. I., and Whitelaw, E. (1997). A globin enhancer acts by increasing the proportion of erythrocytes expressing a linked transgene. Mol Cell Biol *17*, 1607-1614.

Suto, R. K., Clarkson, M. J., Tremethick, D. J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Biol *7*, 1121-1124.

Suzuki, M., Yamada, T., Kihara-Negishi, F., Sakurai, T., and Oikawa, T. (2003). Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression. Oncogene *22*, 8688-8698.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem *279*, 27816-27823.

Sun, L., Liu, A., and Georgopoulos, K. (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. Embo J *15*, 5358-5369.

Svejstrup, J. Q. (2003). Rescue of arrested RNA polymerase II complexes. J Cell Sci *116*, 447-451.

Svejstrup, J. Q. (2004). The RNA polymerase II transcription cycle: cycling through chromatin. Biochim Biophys Acta *1677*, 64-73.

Szutorisz, H., Canzonetta, C., Georgiou, A., Chow, C. M., Tora, L., and Dillon, N. (2005). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. Mol Cell Biol *25*, 1804-1820.

Szutorisz, H., and Dillon, N. (2005). The epigenetic basis for embryonic stem cell pluripotency. Bioessays *27*, 1286-1293.

Syntichaki, P., Topalidou, I., and Thireos, G. (2000). The Gcn5 bromodomain co-ordinates nucleosome remodelling. Nature *404*, 414-417.

Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. (2001). Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem *276*, 25309-25317.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev *16*, 1779-1791.

Talbot, D., and Grosveld, F. (1991). The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. Embo J *10*, 1391-1398.

Tamaru, H., and Selker, E. U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277-283

Tariq, M., and Paszkowski, J. (2004). DNA and histone methylation in plants. Trends Genet *20*, 244-251.

Tian, J., Okabe, T., Miyazaki, T., Takeshita, S., and Kudo, A. (1997). Pax-5 is identical to EBB-1/KLP and binds to the VpreB and lambda5 promoters as well as the KI and KII sites upstream of the Jkappa genes. Eur J Immunol *27*, 750-755.

Tkachuk, D. C., Kohler, S., and Cleary, M. L. (1992). Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell *71*, 691-700.

Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.

Toh, Y., Pencil, S. D., and Nicolson, G. L. (1994). A novel candidate metastasis-associated gene, mta1, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. J Biol Chem *269*, 22958-22963.

Townes, T. M., Lingrel, J. B., Chen, H. Y., Brinster, R. L., and Palmiter, R. D. (1985). Erythroid-specific expression of human beta-globin genes in transgenic mice. Embo J *4*, 1715-1723.

Treisman, R., Green, M. R., and Maniatis, T. (1983). cis and trans activation of globin gene transcription in transient assays. Proc Natl Acad Sci U S A *80*, 7428-7432.

Trotter, K. W., and Archer, T. K. (2004). Reconstitution of glucocorticoid receptor-dependent transcription in vivo. Mol Cell Biol *24*, 3347-3358.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. Embo J *13*, 3822-3831.

Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol *18*, 4629-4638.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol *18*, 4629-4638.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Tuan, D., Kong, S., and Hu, K. (1992). Transcription of the hypersensitive site HS2 senhancer in erythroid cells. Proc Natl Acad Sci U S A *89*, 11219-11223.

Tuan, D., Solomon, W., Li, Q., and London, I. M. (1985). The "beta-like-globin" gene domain in human erythroid cells. Proc Natl Acad Sci U S A *82*, 6384-6388.

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

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

Turner, B. M. (2002). Cellular memory and the histone code. Cell *111*, 285-291.

Turner, B. M., Birley, A. J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. Cell *69*, 375-384.

Udvardy, A., Maine, E., and Schedl, P. (1985). The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. J Mol Biol *185*, 341-358.

Urbanek, P., Wang, Z. Q., Fetka, I., Wagner, E. F., and Busslinger, M. (1994). Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. Cell *79*, 901-912.

Vakoc, C. R., Mandat, S. A., Olenchock, B. A., and Blobel, G. A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell *19*, 381-391.

Vakoc, C. R., Letting, D. L., Gheldof, N., Sawado, T., Bender, M. A., Groudine, M., van Leeuwen, F., Gafken, P. R., and Gottschling, D. E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell *109*, 745-756.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.

Verdone, L., Caserta, M., and Di Mauro, E. (2005). Role of histone acetylation in the control of gene expression. Biochem Cell Biol *83*, 344-353.

Vieira, K. F., Levings, P. P., Hill, M. A., Crusselle, V. J., Kang, S. H., Engel, J. D., and Bungert, J. (2004). Recruitment of transcription complexes to the beta-globin gene locus in vivo and in vitro. J Biol Chem *279*, 50350-50357.

Villa, A., Santagata, S., Bozzi, F., Giliani, S., Frattini, A., Imberti, L., Gatta, L. B., Ochs, H. D., Schwarz, K., Notarangelo, L. D., *et al.* (1998). Partial V(D)J recombination activity leads to Omenn syndrome. Cell *93*, 885-896.

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.

von Freeden-Jeffry, U., Vieira, P., Lucian, L. A., McNeil, T., Burdach, S. E., and Murray, R. (1995). Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. J Exp Med *181*, 1519-1526.

Vosshenrich, C. A., Cumano, A., Muller, W., Di Santo, J. P., and Vieira, P. (2003). Thymic stromal-derived lymphopoietin distinguishes fetal from adult B cell development. Nat Immunol *4*, 773-779.

Wade, P. A. (2001). Transcriptional control at regulatory checkpoints by histone deacetylases: molecular connections between cancer and chromatin. Hum Mol Genet *10*, 693-698.

Wallrath, L. L., and Elgin, S. C. (1995). Position effect variegation in Drosophila is associated with an altered chromatin structure. Genes Dev *9*, 1263-1277.

Wang, A., Kurdistani, S. K., and Grunstein, M. (2002). Requirement of Hos2 histone deacetylase for gene activity in yeast. Science *298*, 1412-1414.

Wang, D. M., Taylor, S., and Levy-Wilson, B. (1996). Evaluation of the function of the human apolipoprotein B gene nuclear matrix association regions in transgenic mice. J Lipid Res *37*, 2117-2124.

Wang, H., An, W., Cao, R., Xia, L., Erdjument-Bromage, H., Chatton, B., Tempst, P., Roeder, R. G., and Zhang, Y. (2003). mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. Mol Cell *12*, 475-487.

Wang, J. H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A. H., Bigby, M., and Georgopoulos, K. (1996). Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity *5*, 537-549.

Wang, M. M., and Reed, R. R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature *364*, 121-126.

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., *et al.* (2004). Human PAD4 regulates histone arginine methylation levels via demethylimination. Science *306*, 279-283.

Watt, F., and Molloy, P. L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev *2*, 1136-1143.

Weintraub, H. (1988). Formation of stable transcription complexes as assayed by analysis of individual templates. Proc Natl Acad Sci U S A *85*, 5819-5823.

Weiss, M. J., Dekker, J., and Blobel, G. A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol Cell *17*, 453-462.

Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases. Cell *125*, 467-481.

Widom, J. (1998). Structure, dynamics, and function of chromatin in vitro. Annu Rev Biophys Biomol Struct *27*, 285-327.

Winkler, G. S., Kristjuhan, A., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. (2002). Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels in vivo. Proc Natl Acad Sci U S A *99*, 3517-3522.

Wolffe, A. P. (1995). Centromeric chromatin. Histone deviants. Curr Biol *5*, 452-454.

Woodage, T., Basrai, M. A., Baxevanis, A. D., Hieter, P., and Collins, F. S. (1997). Characterization of the CHD family of proteins. Proc Natl Acad Sci U S A *94*, 11472-11477.

Woodcock, C. L., Skoultchi, A. I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res *14*, 17-25.

Woodcock, C. L., Grigoryev, S. A., Horowitz, R. A., and Whitaker, N. (1993). A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. Proc Natl Acad Sci U S A *90*, 9021-9025.

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

Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. Cell *79*, 875-884.

Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., Hall, M. C., Borchers, C. H., and Strahl, B. D. (2003). Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev *17*, 654-663.

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.

Xie, Z., and Price, D. (1997). Drosophila factor 2, an RNA polymerase II transcript release factor, has DNA-dependent ATPase activity. J Biol Chem *272*, 31902-31907.

Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D., and Grewal, S. I. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. Mol Cell *20*, 173-185.

Yang, W. M., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996). Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. Proc Natl Acad Sci U S A *93*, 12845-12850.

Yang, X. J. (2004). The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. Nucleic Acids Res *32*, 959-976.

Yang, Z. Q., Imoto, I., Fukuda, Y., Pimkhaokham, A., Shimada, Y., Imamura, M., Sugano, S., Nakamura, Y., and Inazawa, J. (2000). Identification of a novel gene,

Ying, Q. L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281-292.

Yoder, J. A., and Bestor, T. H. (1998). A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. Hum Mol Genet *7*, 279-284.

Yoder, J. A., Soman, N. S., Verdine, G. L., and Bestor, T. H. (1997). DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. J Mol Biol *270*, 385-395.

Yoon, H. G., Chan, D. W., Huang, Z. Q., Li, J., Fondell, J. D., Qin, J., and Wong, J. (2003a). Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. Embo J *22*, 1336-1346.

Yoon, H. G., Chan, D. W., Reynolds, A. B., Qin, J., and Wong, J. (2003b). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. Mol Cell *12*, 723-734.

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.

Youn, H. D., Grozinger, C. M., and Liu, J. O. (2000). Calcium regulates transcriptional repression of myocyte enhancer factor 2 by histone deacetylase 4. J Biol Chem *275*, 22563-22567.

Young, R. A. (1991). RNA polymerase II. Annu Rev Biochem *60*, 689-715.

Yudkovsky, N., Ranish, J. A., and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. Nature *408*, 225-229.

Zawel, L., Kumar, K. P., and Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. Genes Dev *9*, 1479-1490.

Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A., and Chambon, P. (1986). Multiple sequence motifs are involved in SV40 enhancer function. Embo J *5*, 387-397.

GASC1, within an amplicon at 9p23-24 frequently detected in esophageal cancer cell lines. Cancer Res *60*, 4735-4739.

Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. J Biol Chem *277*, 11621-11624.

Zhang, D., Yoon, H. G., and Wong, J. (2005). JMJD2A is a novel N-CoR-interacting protein and is involved in repression of the human transcription factor achaete scute-like homologue 2 (ASCL2/Hash2). Mol Cell Biol *25*, 6404-6414.

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J., and Johansen, K. M. (2006). The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in Drosophila. Development *133*, 229-235.

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

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.