CHARACTERISING THE MLL COMPLEX: EPIGENETIC REGULATION OF *Hoxa* GENES

By

MARIANNE BUSSIÈRE

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Chromatin and Gene Expression Group
Institute of Biomedical Research
College of Medical and Dental Sciences
The University of Birmingham
November 2009

Abstract

Characterising the MLL complex: epigenetic regulation of *Hoxa* genes

The *Mixed-Lineage Leukaemia* (MLL1) protein is a key developmental factor that acts to regulate genes via its histone methyl-transferase activity. This study aimed to examine how MLL1 and its associated complex contribute to gene regulation in a functionally relevant cell background. To assess dynamic processes involved, we developed a haematopoietic Stem Cell (hSC) -based system, in which *Hoxa* genes are down-regulated in a differentiation-induced manner.

I characterised the histone modification distribution on two MLL-target genes showing the largest changes upon differentiation: *Hoxa4* and *Hoxa5*. When active, these genes are associated with a peak of "activating" histone marks (H3K4me3, H3K9ac, H4K16ac) over the transcriptional start site, which are lost when the gene is repressed. This correlated with the recruitment of enzymes that deposit these marks, including components of the MLL complex (MLL^C, MLL^N and menin) as well as HATs that may be associated with the complex (CBP and MOF). Interestingly, the location of these proteins does not always correlate with the marks they deposit. We show that the dual mark H3K9acS10p is present on active *Hoxa4* and *Hoxa5* genes, and correlates with the presence of the histone kinase Msk1. We speculate that Msk1 contributes to regulating MLL1 HMT'ase activity on these genes.

À mes parents,
Paul et Véra

Acknowledgments

I would like to thank my supervisor Dr. Karl Nightingale, for his support and encouragement during this project. I benefited from his precious advice, which gave me confidence. I am also deeply grateful for the patience and calmness he showed throughout my Ph.D, and I am proud to have been his first PhD student.

I also would like to thanks Prof. Bryan Turner and Dr. Laura O'Neill for their support.

Also special thanks to John who accepted some extra work during these past couples of weeks.

Many thanks to the "Chromatin and Gene Regulation Group" and the Bioinformatic Dreamteam for the great ambiance they managed to establish between the groups, the fun lunchtimes, and all the moments that made the lab feel like a second home.

Huge thanks to all the friends who were there for me, in particular Amandine, Elsa, Nil, Anna, Tammy, Natascha, and also Karine, Cécile and Céline despite the distance.

My sincere gratitude goes to my parents and my brothers, Pierre and Eric, whose love and encouragement provided me with the energy to carry through this fascinating project.

Je remercie également mes grand-parents pour leur précieux soutien, ainsi que toute ma famille, en particulier ma grande-tante Jeanine, dont la vie et la carrière ont été, et restent, une source d'inspiration.

Abbreviations

abd-A: *abdominal A*

Abd-B: *Abdominal-B*

Acetyl CoA: acetyl Coenzyme A

Antp: *Antennapedia*

AIRE: AutoImmune Regulator

ALL: Acute Lymphoid Leukaemia

AML: Acute Myeloid Leukaemia

APS: Ammonium Persulfate

ASH2L: Absent, Small or Homeotic-Like 2

ATP: Adenosine TriPhosphate

AUT: Acid Urea Triton

bp: base pair

BPTF: Bromodomain and PHD finger Transcription Factor

BRM: Brahma

BSA: Bovine Serum Albumin

CBP: CREB-Binding Protein

C-ChIP: Carrier-Chromatin Immunoprecipitation

cDNA: complementary DNA

Chd1p: Chomo-ATPase/helicase-DNA binding domain 1

CHO: Chinese Hamster Ovary

Chromodomain: Chomatin Organisation Modifier

CLP: Common Lymphoid Progenitor

CMP: Common Myeloid Progenitor

COMPASS: Complex of Protein Associated with Set1

Co-REST: Co-repressor to RE1 Silencing Transcription factor/neural restrictive silencing factor

Cps35: Compass 35

DCC: Dosage Compensation Complex

Dfd: *Deformed*

DMEM: Dulbecco's Modified Eagle Medium

DNA: DeoxyriboNucleic Acid

DNMT: DNA Methyl-Transferase

DPY-30:DumPY-30

EDTA: Ethylene Diamine Tetraacetic Acid

EGTA: Ethylene Glycol Tetraacetic Acid

ENL: Eleven-Nineteen Leukaemia

ESC: Embryonic Stem Cell

ESC: Extra Sex Combs

E(Z): Enhancer of Zeste

FACS: Fluorescence-Activated Cell Sorting

FAD: Flavin Adenine Dinucleotide

FBS: Foetal Bovine Serum

GAPDH: GlycerAldehyde 3-Phosphate DeHydrogenase

GASC-1: Gene Amplified in Squamous protein 2C

Gcn5p: General control of amino acid synthesis protein 5

GMP: Granulocyte-Macrophage Progenitor

GNAT: Gcn5p-related Acetul-Transferase

GP6: Glycoprotein VI

HAT: Histone Acetyl-Transferase

HDAC: Histone Deacetylase

HDM: Histone Demethylase

HMG: High Mobility Group

HMT: Histone Methyl-Transferase

Hos2: High osmolarity sensitivity protein 2

Hox: *Homeobox*

HP1: Heterochromatin-associated Protein 1

HPC: Haematopoietic Progenitor Cell

HRX: Homologue of trithorax

hSC: haematopoietic Stem Cell

ING2: Inhibitor of Growth 2

JHDM1: JmjC domain-containing Histone Demethylase 1

Jmjc: Jumonji c

JMJD2A: Jumonji Domain-containing protein-2-A

Jnk: Jun N-terminal Kinase

lab: *labial*

LEGDF: Lens Epithelium-Derived Growth Factor

LSD1: Lysine Specific Demethylase 1

MeCP2: Methyl-CpG-binding protein

MEP: Megakaryocyte-Erythrocyte Progenitor

MLL: Myeloid/Lymphoid or Mixed Lineage Leukaemia

MOF: Males absent On the First

Msk1: Mitogen- and Stress- activated Protein Kinase 1

N-ChIP: Native-Chomatin Immunoprecipitation

NFκB: Nuclear Factor kappa-light-chain-enhancer of activated B

NuRD: Nucleosome Remodelling and histone Deacetylation

NURF: Nucleosome Remodelling Factor

PAGE: PolyAcrylamide Gel Electrophoresis

Pb: *Proboscipedia*

PBS: Phosphate Buffered Saline

Pc: Polycomb

PCA: Perchloric Acid

PCAF: P300/CBP Associated Factor

PcG: Polycomb Group

PCR: Polymerisation in Chain Reaction

PH: Polyhomeotic

PHD: Plant Bromodomain

PRMT: Protein Arginine Methyl-Transferase

PSC: Posterior Sex Combs

PTD: Partial genomic Tandem Duplication

RBBP5: Retinoblastoma Binding Protein 5

RNA: RiboNucleic Acid

RTQ-PCR: Real-Time Quantitative- PCR

SAGA: Spt-Ada-Gcn5-Acetyl-transferase

SAM: S-adenosyl Methionine

Sbf1: SET binding factor 1

Scr: *Sex comb reduced*

SDS: Sodium Dodecyl Sulfate

SET: Suppressor of variegation, Enhancer of zeste andTrithorax

SL2: Schneider line-2

SUV39H1: Suppressor of Variegation 3-9 Homolog 1)

SU(Z): Suppressor of Zeste

SWI/SNF: SWItch/Sucrose Non-Fermentable

TAC: Trithorax Activating Complex

TAD: Transcriptional Activation Domain

TAF: TBP Associated Factor

TASPASE1: Threonine Aspartase 1

TBP: TATA box Binding Protein

TCA: Trichloroacetic Acid

TEB: Tris-EDTA Boric acid buffer

TPO: Thrombopoietin

Trx: Trithorax

trxG: trithorax Group

TSS: Transcription Start Site

Ubx: *Ultrabithorax*

UTX: Ubiquitously Transcribed TPR on X

WBP7: WW Binding Protein 7

WDR5: WD-Repeat protein-5

X-ChIP: Cross-linking Chomatin Immunoprecipitation

YY1: Yin Yang-1

Table of Contents

Table of figures

1. INTRODUCTION

Almost every cells of a multi-cellular organism possesses the same genetic information asthe zygote from which it has developed, based on identical copies of the same DNA sequence. Upon differentiation cells must adopt a specific pattern of gene expression, which defines their function and identity. A pattern of gene expression is specific to individual cells and can vary in a dynamic manner through time, or in response to environmental influences (Jaenisch and Bird, 2003): genes can be selectively activated, repressed or maintained in their activity (Figure 1.1). These patterns of gene transcription are regulated via networks of activating or repressive transcription factors, but also by a range of "*epigenetic mechanisms*". Several definitions of "epigenetic" have been proposed (Berger *et al.*, 2009; Bird, 2007; Probst *et al.* 2009). Current definitions propose that epigenetic mechanisms correspond to changes in phenotype, or gene expression that occur without altering the DNA sequence (Berger *et al.*, 2009). In this study, the term "epigenetic mechanisms" is used to describe changes in the pattern of gene expression influenced by processes including nucleosome positioning, the distribution of histone post-translational modifications, and binding of chromatin-associated proteins and DNA methylation. Many of these mechanisms interact at specific genes, yet regulatory processes using histone modification play a central role as these are sensitive to intrinsic and environmental factors as these can impact on the activity of specialised histone modifying enzymes (Bird, 2007; Jaenisch and Bird, 2003). This has functional effects as histone modifications can either have direct effects on chromatin structure and/or be selectively recognised by "effector" proteins through functional domains (Jenuwein and Allis, 2007; Taverna, 2007), which subsequently act to govern gene activity (Campos and Reinberg, 2009) (Figure 1.2). The biological processes in which "*epigenetic mechanisms*" act are varied in both extent and duration, and range from the permanent chromosome-wide

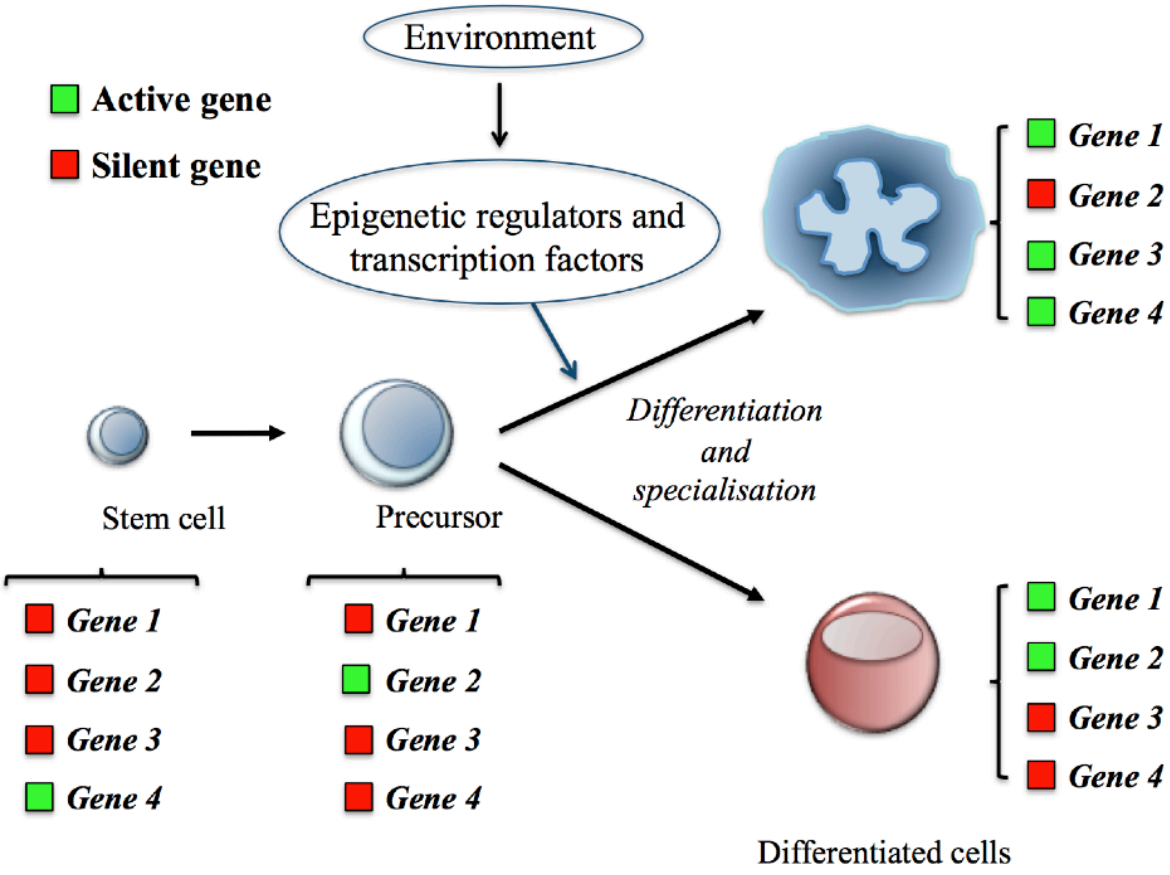


Figure 1.1: Model of cell fate and gene activity (based on Jaenisch and Bird, Nature, 2003; Ng J.H. *et al.,* Mutation research, 2008; and Lennartsson A. and Ekwall K., *Biochim. Biophys Acta*, 2009)

Almost every cell possesses the same genetic information. However, their cell fate or phenotype is determined by the selective activation or repression of genes throughout development, generating a pattern of gene expression which is transmitted through cell generations. The specialisation associated with the establishment of a pattern of gene expression is specific to individual cells, and can be governed by epigenetic mechanisms.

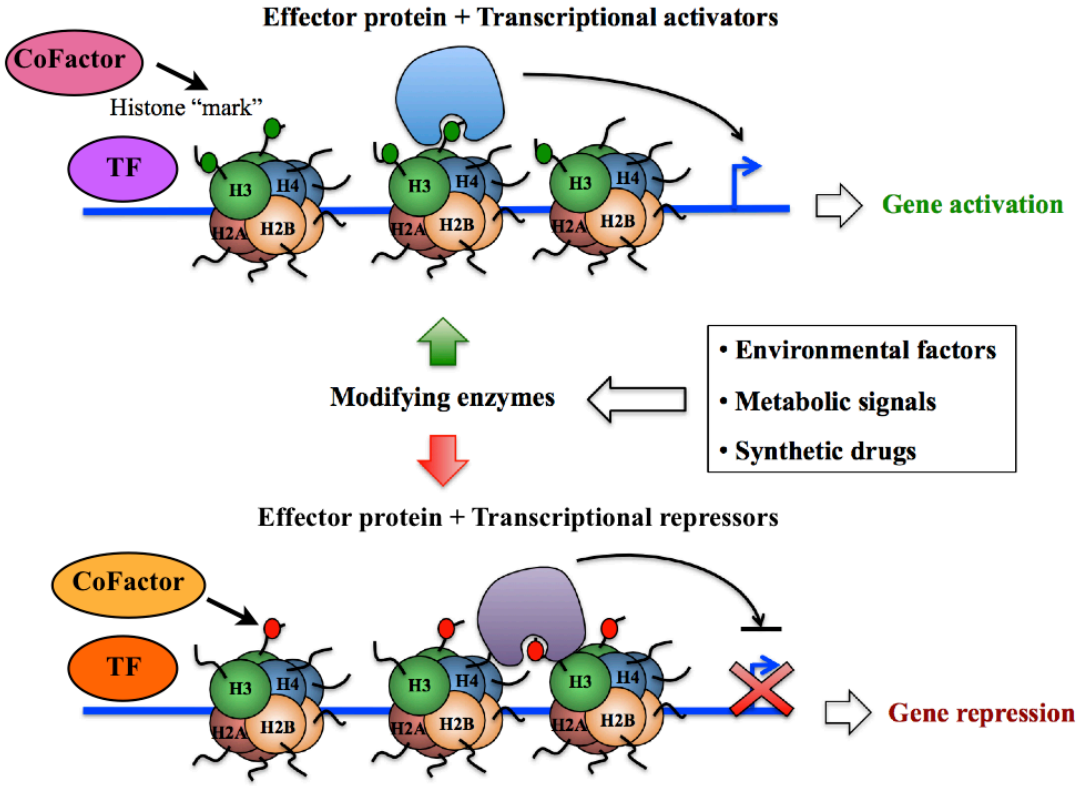


Figure 1.2 : Theoretical model of epigenetic regulation (based on Jaenisch R. and Bird A., Nature, 2003; and Imhof A., Briefings in functional genomics & proteomics, 2006)

Epigenetic mechanisms are influenced by intrinsic and environmental factors, metabolic signals, or synthetic or naturally occuring small molecules. Modifying enzymes (Cofactors) are recruited by transcriptional activators or repressors (TF) and add or remove histone modifications (or histone "marks"), which are then recognised by specific effector proteins through recognition domains, leading to subsequent functional effects and the possible modulation of gene activity.

silencing associated with X-inactivation in mammals (Heard, 2004; Chow & Heard, 2009), to more gene- or cluster-specific regulation like the maintenance of homeotic gene expression (Lewis, 1978; Kundu & Peterson, 2009).

1.1. Epigenetic modifications

1.1.1. Chromatin organisation

The human genetic information in every cell is contained in two metres of genomic DNA per cell. In order to be packed into the eukaryotic nucleus, DNA associates with small nuclear basic proteins termed histones, and folds into a series of higher order structures associated with increased levels of compaction (Figure 1.3). This highly organised structure is called chromatin. Histones are highly conserved proteins, being essentially identical in yeast and human, and assemble into nucleosomes, the fundamental unit of chromatin (Oudet *et al.*, 1975). The nucleosome is composed of the nucleosome core, the "linker DNA" and the linker histone H1 (Simpson, 1978).

The nucleosome core particle is a structure composed of an octamer of core histones, around which wraps 147 base pairs of DNA in 1.75 turns. One nucleosome is composed of a H3-H4 tetramer and two H2A-H2B dimers (Thomas and Kornberg, 1975). We understand the structure of the nucleosome due to high-resolution crystallography studies (e.g. Figure 1.4, Luger *et al*., 1997). Core histone proteins share a similar structural motif. Each histone displays a "U-shape" structure: the association of two of these structures in a "handshake" motif leads to the formation of a dimer, which are subsequently arranged to form "crescent-shape" heterodimers (Luger *et al*., 1997). Each individual histone comprises a globular or "histone fold" domain, responsible for histone-histone interactions and histone-DNA interactions, and a "tail" domain, which are structurally undefined but highly evolutionarily

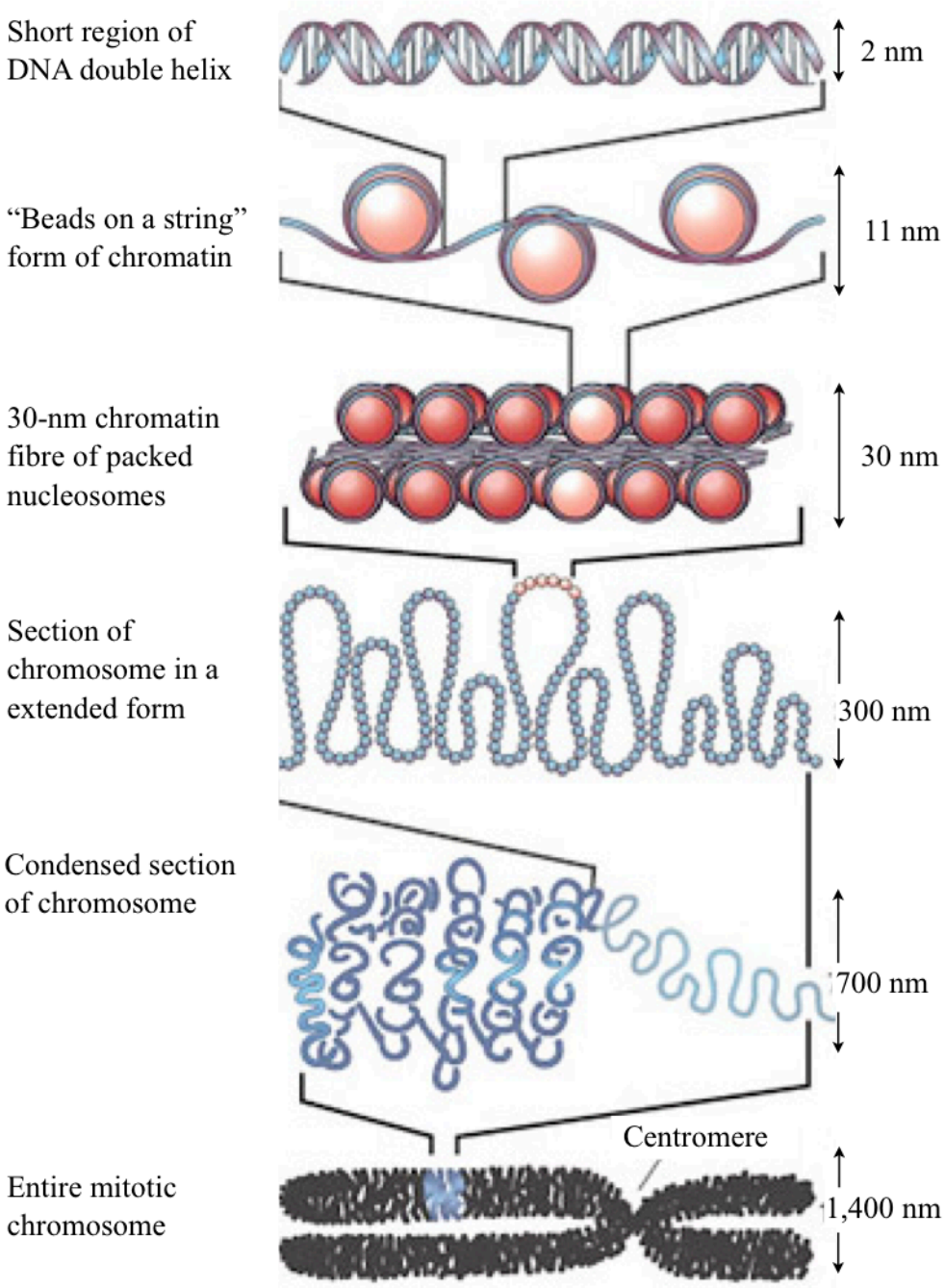


Figure 1.3: Different levels of condensation of chromatin (from Felsenfeld,G. and Groudine, M. , Nature 2003)

"Naked" DNA is first wrapped twice around nucleosomes, leading to the formation of a "beads on a string" structure. Then, nucleosomes are packed and the chromatin is condensed into higher order structures, which eventually fold into the structures associated with the mitotic chromosome, the highest condensed chromatin conformation.

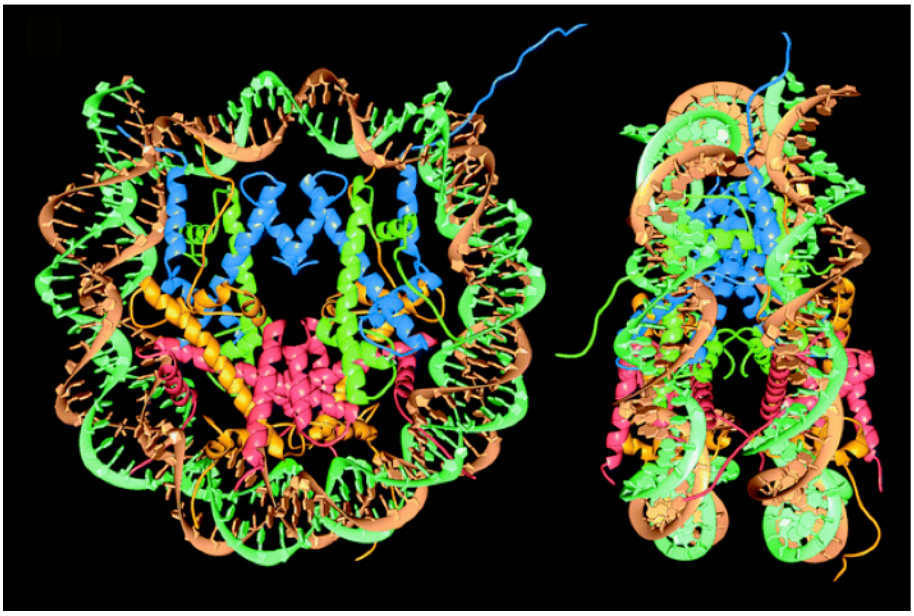


Figure 1.4: Crystal structure of the nucleosome core particle (from Luger K. *et al.*, Nature, 1997)

Nucleosome core particles are composed of two H2A-H2B dimers and a H3-H4 tetramer, and is associated with 1.75 turns of DNA. Each histone displays a "U-shape" structure and associates with a partner to form a "crescent-shape" heterodimer. The figure shows the particle "face on" (left) or a side view (right) of a single nucleosome particle, with the two strands of 147 base pairs DNA indicated (green and orange), and the four core histones; H3 (Blue), H4 (Green), H2A (Yellow) and H2B (Red). Note the histone tails are not fully shown as these were not visible in this crystal structure.

conserved. The N-terminal tails of all the four core histones and the C-terminal tail of histone H2A are unstructured, highly basic domains (though they are thought to adopt a defined α-helical conformation when binding within the chromatin fibre, (Baneres *et al*., 1997), characterised by their high level of positively-charged residues, and protruding from the nucleosome, thus being exposed to chromatin-associated proteins (Vettese-Dadey *et al*., 1996; Luger *et al*., 1997; Luger and Richmond, 1998; Zheng and Hayes, 2003). While the core histone is responsible for DNA binding, the histone tails have a role in the recruitment of specific protein factors, intervening either directly or indirectly in the modulation of higher order chromatin structures (Hansen, 2002; Hansen *et al.*, 1989; Fletcher and Hansen, 1995; Luger, 2003). As such the histone tails are involved in the formation and modulation of higher order chromatin structures (Hansen, 2002; Hansen *et al.,* 1998). It has been suggested that the assembly of chromatin can be transcriptionally repressive. In this model, the packaging of DNA in nucleosomes prevents the binding of DNA-binding factors (i.e. Lee et al, 1993), either due to the distortion of the highly constrained DNA, or because of steric occlusion generated by the core histones or histone H1 (Juan *et al*., 1997; Luger *et al.*, 1997). Chromatin is typically packaged with nucleosomes at ~200bp intervals, generating the intervening "linker DNA" (Oudet *et al*., 1975). Thus, each nucleosome is linked to another by "linker DNA" and bound to a linker histone, such as H1, which influences and stabilises chromatin higher order structure (Hansen *et al*., 1989; Hansen, 2002). The association of about 160bp of DNA with the nucleosome, and associated with the linker histone H1, is termed the "chromatosome" (Simpson, 1978). The linker histone H1 is positioned outside the core octamer, adjacent to the site of entrance and exit of the DNA from the nucleosome (Figure 1.5) (Travers, 1999). Chromatin depleted of histone H1 generates an 11-nm diameter fibre termed "beads-on-a-string" (Hansen, 2002). This structure can be subsequently further

compacted in a higher order 30-nm diameter fibre though the molecular details of this structure remain unresolved and controversial (Robinson and Rhodes, 2006). Two models of 30-nm fibre structure have been proposed (Wu *et al.*, 2007): a "one-start" stack of nucleosomes in which the 30-nm fibre is packaged as a solenoid, and constitutes a continuous helix of nucleosomes linked by bent "linker DNA" (Finch and Klug, 1976), and a "two-start" stack of nucleosomes, composed of a helix arranged in a conformation where nucleosomes adopt a "zig-zag" conformation. In this model, the two stacks of nucleosomes are linked by linker DNA which can adopt two different orientations, corresponding to either a helical ribbon model (Worcel *et al*., 1981) or a crossed-linker model, depending on the linker DNA length (Williams *et al*., 1986) (Figure 1.6, Dorigo *et al*., 2004). In addition, Richmond and colleagues recently proposed a "two-start" crossed-linker model comprising a four-nucleosome "core", but missing the linker histone (Schalch, *et al.*, 2005). This structure was used as a model for 30-nm higher order chromatin compaction (Robinson and Rhodes, 2006). Though the "one-start" model was initially preferred because of its simplicity, some experimental data suggests that chromatin adopts a "two-start" stack conformation, as this correlates with the high degree of packaging observed with some fibres (Dorigo *et al*., 2004). In addition, the transition from the supercoiled helical structure to the relaxed crossed-ribbon structure is consistent with variable and reversible accessibility of chromatin (Turner, 2002; Margueron, Trojer *et al*., 2005; Wu *et al*., 2007). It was also proposed that higher order chromatin structure displays at least two levels of compaction, one "loose", lacking histone H1, and a more compact structure containing the linker histone and corresponding to the 30-nm fibre (Robinson and Rhodes, 2006). A more recent study by Routh *et al.* proposed that the degree of compaction of the 30-nm structure is also dependent on the nucleosome repeat length (Routh *et al.*, 2008). Further compaction into higher order chromatin structures is

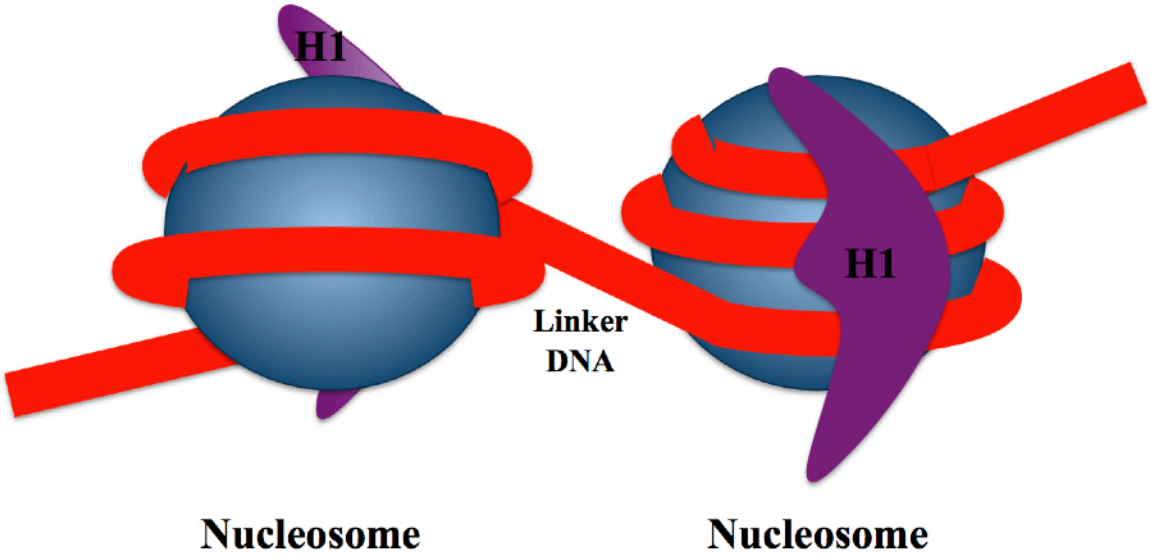


Figure 1.5: Histone H1

Schematic picture of a nucleosome linked to another by the linker DNA and is bound by linker histone H1, which maintains chromatin higher order structure in a stable conformation (based on Travers A., Trends Biochem Sci, 1999).

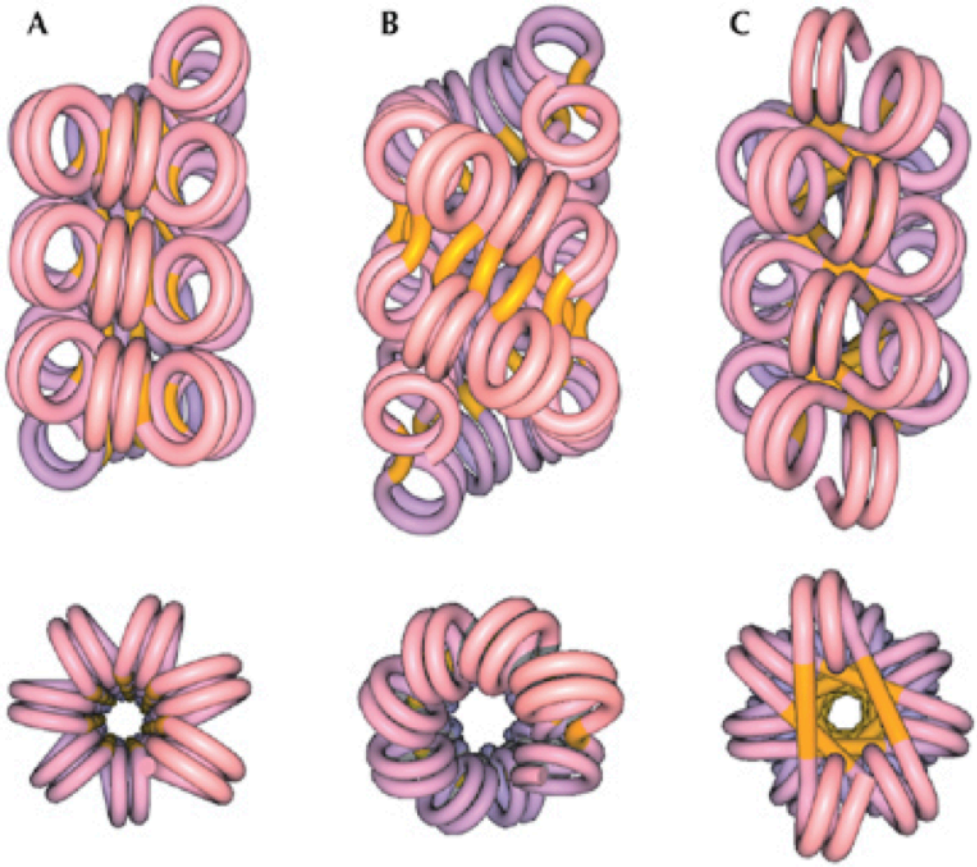


Figure 1.6: Model proposed by Dorigo *et al.* for the 30-nm chromatin fibre (from Dorigo *et al.*, Science, 2004)

Nucleosome core particles can associate into a higher order structure, forming a 30-nm chromatin fibre. Two models of the fibre were proposed: the solenoid helix conformation, comprising a "one-start" stack of nucleosomes, and in which the linker- DNA is bent (A), and the zig-zag helix conformation, comprising a "two-start" stack of nucleosomes, in which adjacent nucleosomes are linked by a straight linker DNA (B and C). This second model is consistent with two possibilities for the orientation of the linker DNA, which are associated with either an helical ribbon model (B), or a crossed- linker model (C).

necessary to enable the condensation of the chromatin into the structures seen as metaphase chromosomes, but this is not yet well understood.

1.1.2. Chromatin dynamics

Linker histones are generally found in a 1:1 ratio with the core histones, and are thought to maintain the chromatin fibre in higher order structures (Thoma *et al*., 1979; Hansen *et al*., 1989). However, it has been reported that the linker histone is only required for the stabilisation of this highly organised structure, while the core histones are involved in both the formation and the stabilisation of higher order chromatin structure (Schwarz and Hansen, 1994). Indeed, H1-depleted chromatin can be folded into higher order structures at appropriate ionic conditions (Hansen *et al*., 1989), but the loss of the core histone tail domains (Garcia-Ramirez *et al*., 1992), notably the H4 tail prevent compaction (Dorigo *et al*., 2003), even in the presence of the linker histone (Allan *et al*., 1982). This raises the question of if (or how) the histones contribute to the dynamic process(es) of chromatin folding. It is essential that the successive levels of chromatin structure are reversible in order to enable access for functional processes. This requirement for dynamic folding appears to be controlled by a combination of linker-histones, other chromatin-associated proteins (like the heterochromatin-associated protein HP1, Hiragami and Festenstein, 2005) and the post-translational modifications of the N-terminal tails of the histones. For example, the degree of histone acetylation (Garcia-Ramirez *et al.*, 1992), specifically at H4K16 (Robinson et al., 2008) has been shown to modulate the folding of the chromatin fibre. Histone acetylation is a lysine amidation, in which the positive charge associated with this residue is neutralised, weakening its affinity for negatively charged DNA. This is thought to lead to the decrease of the interactions between histones and DNA involved in chromatin folding and stabilisation. A major effect of histone acetylation is on the linker DNA, leading to an extended

conformation of the chromatin between nucleosomes (Garcia-Ramirez *et al*., 1995). In addition to its role in the dynamics of chromatin organisation, the acetylation of the histones facilitates the accessibility of nucleosomal DNA to transcription factors (Lee *et al*., 1993; Utley *et al*., 1998) (Figure 1.7). This is thought to play a key role in transcriptional regulation, as the increased binding of transcription factors, correlates with an active state of gene expression.

The functional links between histone modification and transcription were shown by Utley *et al.* (1998), who demonstrated that the yeast histone acetyl transferase complex SAGA was specifically recruited by a promoter-binding transcriptional activator, leading to nucleosomal acetylation, which stimulates active transcription (Utley *et al.,* 1998). Inversely, the loss of histone acetylation leads to a stronger interaction of the histone tails with the DNA, which is thought to contribute to a compacted "inaccessible" chromatin structure.

The structures (and regulators) of chromatin condensation remain unclear, but the suggestion that the initial stage(s) of chromatin compaction are determined by modification of the histone tails, represents a mechanism by which the cellular environment can impact on both chromatin structure and transcription (i.e. Murayama *et al*., 2008). Though histones and nucleosomes were initially acknowledged for their structural role in chromatin compaction, they are now recognised for their essential roles in regulating nuclear functions. The idea that histone post-translational acetylation may affect transcriptional activity (i.e. an increase of acetylation levels is associated with transcription activation) was suggested in 1964 by Allfrey *et al*. (Allfrey *et al.*, 1964). Subsequent studies showed that a high degree of histone acetylation (*hyperacetylation*) correlates with the active transcriptional state at the level of chromosomal domains (Turner, 1998) and individual genes (Hebbes *et al*., 1988) and has a causal role in facilitating transcription (Lee *et al*., 1993; Davie, 1997; Ura *et al.,* 1997; Utley

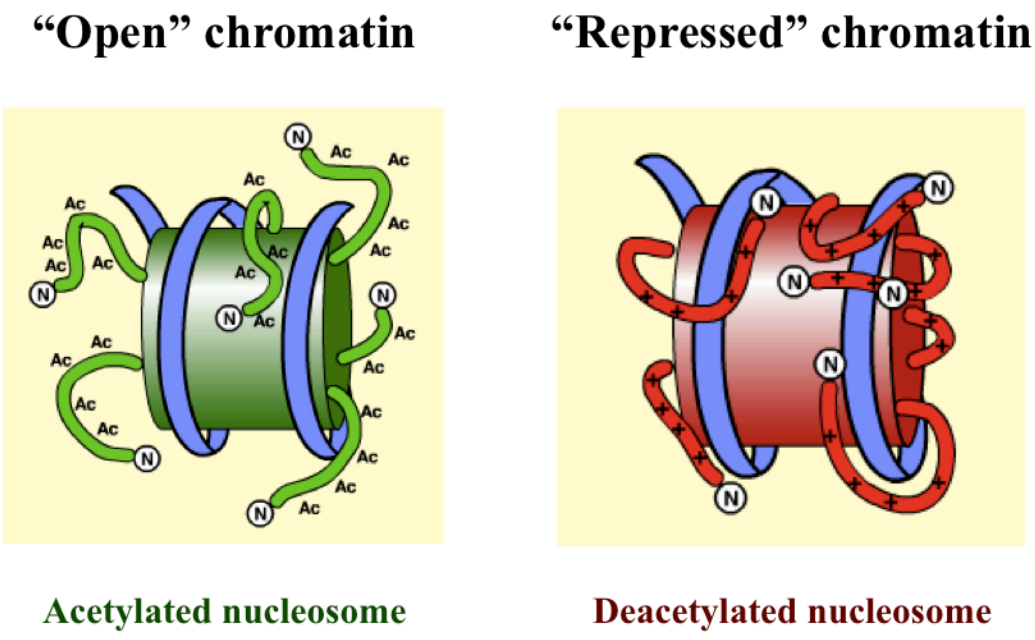


Figure 1.7: Histone acetylation and accessibility (from Schreiber, S., Cambridge, MA. http://www.broadinstitute.org/chembio/lab_schreiber/anims/animations/smdbHDAC.php)

In the nucleosome, the N-terminal tails of histones are subject to acetylation. Histone acetylation (Left) results in the loss of histone-DNA contacts, resulting in the acquisition of an "open" chromatin conformation. This is more accessible to transcription factors, correlating with the active transcriptional state. In contrast, histone deacetylation (Right) results in the formation of a more compact chromatin conformation, which prevents access of transcription factors to the DNA and is associated with gene silencing.

et al., 1998). In contrast, incorporation of linker histone H1 represses transcription (Kamakaka and Thomas, 1990) and correlates with a restriction of the mobility of histones (Ura *et al*., 1997).

Nucleosome assembly is essential in the structural organisation of chromatin. However, the recognition that the core histone tails can modulate the accessibility of the DNA to transcription factors and thereby regulate transcription factor access, and subsequent gene transcription, was important in the understanding of the functional role of chromatin (Bassett *et al.*, 2009; Zheng and Hayes, 2003). These are also essential for maintaining both chromatin conformation and a dynamic chromatin structure through the action of ATP-dependent chromatin remodelling complexes and associated factors (Clapier and Cairns, 2009). Post-translational modifications may cause the tail to adopt new conformations and/or lead to new nucleosomal interactions, altering the properties of the chromatin. These remodelers have the capacity to modulate the composition of nucleosomes, therefore favoring DNA exposure to additional factors (Clapier and Cairns, 2009. Histone modifications on histone residues have also been shown to have a role as a binding platform for other non-histone proteins, leading to downstream functional events (Strahl and Allis, 2000; Kurdistani and Grunstein, 2003). A wide variety of post-translational modifications are found on the histones and range from large and complex groups (i.e. sumoylation, ubiquitylation, ADP-ribosylation), which have been suggested to impact on chromatin structure, to small modifications, which have minimal charge effects (like lysine mono-methylation) (Hansen, 2002; Ruthenburg *et al.*, 2007; Turner, 2000). This suggests that for the majority of histone modifications, the role of modification-binding "effector" proteins is a key concept in their function and in the correlation between histone modifications and the transcriptional activity of genes (Jenuwein and Allis, 2001).

Specific post-translational histone modifications are also associated with different functional states of chromatin, depending on their level of enrichment and their distribution. As previously discussed, histone modifications are involved in modulating chromatin structure. Specific post-translational histone modifications are also associated with different functional states of chromatin, depending on their level of enrichment and their distribution. This has led to the concept that these modifications correspond to a "mark" that can be recognised or "read" by specific proteins, with subsequent functional consequences for the associated chromatin (Strahl and Allis, 2000; Jenuwein and Allis, 2001; Turner, 2000; Turner 2002).

1.1.3. Histone modifications and the histone code

Following the hypothesis that that acetylation might be associated with functional effects on chromatin in *Drosophila* (Turner *et al.*, 1992), the recognition that a range of histone modifications are found in chromatin, and play a role in regulating functional processes has recently led to the proposal of the "epigenetic" (Turner, 2000), or "histone code" hypothesis (Strahl and Allis, 2000). This suggests that one, or a combination of histone modifications generate an "epigenetic mark", which is specifically recognised by proteins, with subsequent functional consequences for the underlying DNA (Strahl and Allis, 2000; Turner, 2000; Jenuwein and Allis, 2001; Turner, 2002). This concept is associated with a number of ideas that will be subsequently explored.

1.1.4. Histone post-translational modifications

The N-terminal histone "tails" are subject to a broad range of post-translational modifications (including methylation, acetylation, phosphorylation, ubiquitination, biotinylation, glycosylation, carbonylation, ADP-ribosylation and sumoylation) (Figure 1.8). The functional consequences of many of these marks remain unclear, but lysine phosphorylation,

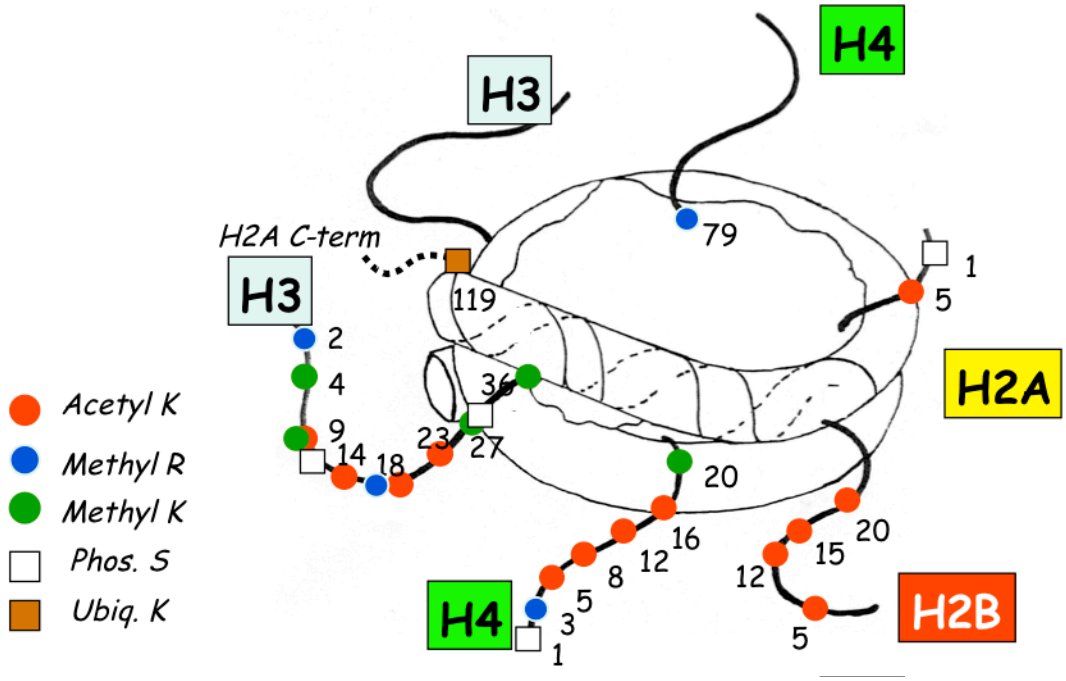


Figure 1.8: Sites of post-translational modifications on histone tails (from Turner BM, Nature Structural and Molecular Biology, 2005)

Histone tails are subject to post-translational modifications primarly on their N-terminal tails. These are epigenetic marks. A wide range of marks on histone residues are shown, notably acetylation (Acetyl K), lysine and arginine methylation (Methyl K, Methyl R), serine phosphorylation (Phos. S) and lysine ubiquitylation (Ubiq. K).

lysine and arginine methylation and lysine acetylation (at specific residues) have been well studied. This has led to the concept that there are two classes of marks; particular histone modifications, termed "activating" marks, are enriched at sites of actively transcribed chromatin (i.e. H3 hyperacetylation, H3K4 methylation, etc.) whereas others, termed "repressive" marks, are predominantly found in silent chromatin (i.e. H3K9me, H3K27me), (Margueron *et al.*, 2005; Turner, 2005). It should be noted that the functional role of these "marks" are highly modification and residue specific. The same modification at two different residues, or the deposition of two different modifications on the same residue typically has different functional effects, for example methylation of H3K4 (Santos-Rosa *et al*., 2002; Wysocka *et al*., 2006) typically has the opposite transcriptional consequences to H3K9 methylation (Bannister *et al*., 2001) but this seems to be determined by the chromatin context (i.e. Shi *et al*., 2006). Furthermore, the number of methyl groups incorporated at a specific residue also impacts on the functional outcome of a "mark"; lysines can be mono, di or tri-methylated (Santos-Rosa *et al*., 2002), whereas arginines can be mono or di-methylated (asymmetric or symmetric) (Margueron *et al*., 2005; Kirmizis et al., 2009).

1.1.5. Spatial distribution of histone modifications

Biochemical analysis of histone modifications, typically using chromatin immunoprecipitation (ChIP) approaches, has shown that different histone marks display distinct patterns in chromatin (Pokholok *et al*., 2005). Numerous studies have shown that promoter regions of active genes are associated with histone acetylation and H3K4 methylation (Hebbes *et al*., 1988), whereas inactive genes are enriched with H3K27 methylation (O'Neill *et al*., 2003).

More recent studies, using ChIP combined with gene microarray analysis ("ChIP-chip"), reveal consistent genome-wide patterns in both yeast and mammalian cell lines (Pokholok *et al*., 2005; Vakoc *et al.*, 2006; Wang *et al.,* 2008) (Figure 1.9). In active genes, modifications such as H3K4 methylation (H3K4me3) and H3K9 acetylation (H3K9ac) are enriched within the promoter regions, while some others are more abundant in the coding region. For example, in the yeast *S. cerevisiae*, the promoter and the 5' end of genes are enriched with acetyl marks (H3K9ac), whereas the coding regions of actively transcribed genes are enriched with tri-methylation of H3K4, H3K36 and H3K79 (Pokholok *et al*., 2005; Kouzarides, 2007). Similarly, in mammals, mono, di and tri-methylation of H3K4 were found to be elevated at the transcription start sites (TSS) of active genes (Bernstein *et al.,* 2005; Barski *et al*., 2007). More specifically, where all H3K4me marks are enriched around the transcription start site, H3K4me1 and H3K4me2 typically display two major peaks of enrichment, on each side of the TSS, whereas the tri-methyl mark is closer to the transcription start site (Barski *et al.,* 2007). In mammalian chromatin, there is also a high level of H3K79me3 near the transcription start site (TSS) of active genes (Vakoc *et al*., 2006). In contrast, other marks present a clearly different distribution; H3K36me3, H3K9me3 and H3K79me3, and monomethylated H4K20 are enriched at the 3' end of genes, and spread across the coding region. These marks also contribute to transcriptional repression (Vakoc *et al.,* 2006).

A similar pattern is seen with repressive modifications, with both TSS specific, and coding-region-specific distributions. H3K27me3, a mark that correlates with gene repression, is slightly elevated in the promoter regions of transcriptionally inactive genes, and is associated with a broad distribution over the coding region, whereas this mark is decreased in active genes (Barski *et al.*, 2007). Similarly, H3K9me2/3, a mark associated with gene repression

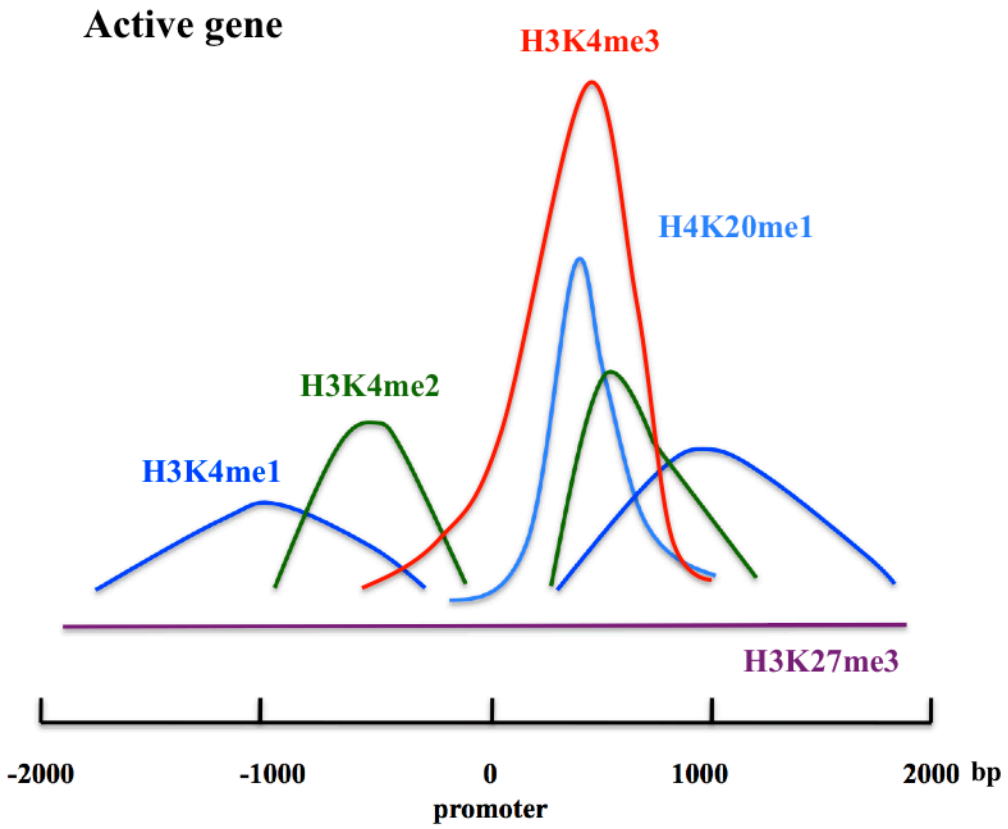


Figure 1.9: Distribution of histone marks on an "average" human active gene (based on Barski, A., *et al.*, Cell, 2007)

In active genes, histone marks show a characteristic pattern of distribution. For example, an histone mark associated with transcriptionally active gene, H3K4me3 is enriched at the transcription start site, whereas H3K4me2 and H3K4me1 are enriched in the proximity of the TSS, on each side. H4K20me1 is enriched at promoter of active genes. H3K27me3, a mark associated with gene repression, shows no enrichment on active genes. Note that the relative abundance of different marks cannot be compared.

and the formation of heterochromatin (Bannister *et al.*, 2001), is increased around the transcription start site in silent genes (H3K9me2), and reduced in active genes (Barski *et al.*, 2007). Furthermore, the degree of modification adds another level of complexity: while H3K27me3 and H3K9me2 are associated with gene silencing, H3K27me1 and H3K9me1 are associated with actively transcribed genes, particularly surrounding the transcription start site (Barski *et al.,* 2007). Similarly, where H3K4me3 is associated with gene activation only, high levels of mono and di H3K4me correlate with both active and inactive genes (Santos-Rosa *et al*., 2002). The observation that different marks are associated with different functional states of genes and adopt distinct genomic locations suggests they have divergent functional effects, and achieve them in different ways. For example, it was proposed that the enrichment of repressive marks (such as H4K20me) within the transcribed region of genes, may be involved in the prevention of aberrant "adventitious activation" by RNA Polymerase II, as seen with H3K36me (Joshi and Struhl, 2005; Vakoc *et al*., 2006).

Finally, interestingly, two functionally "opposite" marks, H3K4me3 and H3K27me3, are observed in the same region of key genes in the context of stem cell differentiation. These were termed "bivalent domains" and are proposed to enable the gene to adopt a different functional state (i.e. active to silent) as differentiation progresses, having the potential to be subsequently associated with either of these two marks (Bernstein *et al*., 2006; Azuara *et al*., 2006). However, recent data suggests that bivalency at least partially reflects the pooling of functionally distinct populations of cells (Akkers *et al.,* 2009).

1.1.6. "Cross-talk", or the functional interaction of modifications

The divergent distributions of histone modifications suggest their functional effects are via a number of mechanisms, but often these distributions overlap suggesting different marks act

in a co-ordinated manner. At a mechanistic level, this "cross-talk" implies that several marks act together to regulate processes, but it also suggests that some marks influence the deposition of other modifications, either *in cis* (on the same histone tail) or *in trans* (on different histone tails) (Nightingale *et al.,* 2006; Latham and Dent, 2007). The simplest examples are where one modification can affect the deposition of other modifications on the same residue, and thus are mutually exclusive; for example, H3K9 acetylation and H3K9 methylation are marks with opposing transcriptional outcomes.

Other interrelated modifications interact *in trans*, or between histones. For example, in the yeast *S. cerevisiae*, lysine 123 mono-ubiquitination of histone H2B precedes H3K4 and H3K79 trimethylation by the COMPASS and Dot1 methyl-transferases respectively (Briggs *et al*., 2002; Sun and Allis, 2002). In this case, deposition of the ubiquitin mark is catalysed by the ubiquitin-conjugating-enzyme Rad6/Bre1 (Sung *et al.,* 1988; Robzyk *et al.*, 2000), and leads to the direct or indirect recruitment of the histone methyl-transferase COMPASS (Complex of Proteins Associated with Set1) (Dover *et al.,* 2002; Nakanishi et al., 2009). This pathway is also conserved from yeast (through H2BK123Ub1) to human (through H2B120Ub1) (Kim *et al*., 2009). As a crucial element in this cross-talk process, a study from Lee *et al.* showed that Swd2 is directly or indirectly recruited to the chromatin in a mono-ubiquitination dependent manner, and that the subsequent association of Swd2 with COMPASS is essential for H3K4 trimethylation (Lee *et al*., 2007). However, this study contradicts other studies proposing that Swd2 is the key element in Set1 stability (Dichtl *et al.*, 2004), but that H2B is not essential for Swd2 recruitment to the chromatin (Vitaliano-Prunier *et al.*, 2008). In this last study, it is proposed that Swd2 contributes to the unmodified H3-dependent recruitment of COMPASS to the chromatin. Ubiquitylation of both H2B and Swd2 would then favor the recruitment of the COMPASS subunit Spp1 contributing to H3K4

methylation (Dehe *et al.*, 2006; Morillon *et al.*, 2005; Vitaliano-Prunier *et al.*, 2008). Swd2 also binds to Dot1, and is required for H3K79 trimethylation (Lee *et al.,* 2007). This H2B modification is transient though, as the mono-ubiquitination is then removed by a component of the Gcn5-containing SAGA complex, the ubiquitin protease Ubp8 (Dover *et al*., 2002; Henry *et al*., 2003; Margueron *et al*., 2005; Latham and Dent, 2007). Cross-talk can also occur between modifications on different residues on the same histone, *in cis*. A particularly well-studied example of cross-regulation involves H3K9ac, H3K14ac and H3S10p, where the phosphorylation of H3S10 is followed by acetylation of H3K14 (Cheung *et al*., 2000). This interrelation was also observed in yeast, where H3K14 acetylation is deposited by the Gcn5 acetyl-transferase (Lo *et al*., 2000). In addition, H3S10 phosphorylation prevents H3K9 methylation (Rea *et al*., 2000), blocking the recruitment of HP1 to chromatin (Fischle *et al*., 2005; Latham and Dent, 2007). Similarly, H3K4me3, H3K9ac and H3S10p are present on the same nucleosomes, on transcriptionally active genes in mouse (Clayton *et al*., 2000; Hazzalin and Mahadevan, 2005). These studies found histone H3 methylated at lysine 4 is subject to dynamic acetylation, whereas this was not observed with H3K9me. This was consistent with their findings that both H4K9ac and H3S10p were transient marks that correlated with gene activation, whereas H3K4me3 is present before gene activation and maintained for a few hours afterwards (Hazzalin and Mahadevan, 2005).

The specific degree of histone modification seems to be determined by other adjacent modifications. Mass-spectroscopy (Churikov *et al*., 2004) and subsequent Western analysis showed that increased acetylation of histone H3 is associated with a high level of H3K4 methylation (di and tri particularly) (Nightingale *et al.*, 2007), whereas methylation at the residue H3R2 appears to inhibit H3K4 methylation (Kimizis *et al*., 2007).

H3K4 methylation can also influence other modifications *in trans*, and leads to the acetylation of histone H4 at lysine 16. This cross talk is the result of the association of two enzymes, the histone methyl-transferase MLL1 and the H4K16-specific histone acetyl-transferase MOF (Dou *et al.,* 2005). This is consistent with their "activating" roles; H4K16 acetylation prevents higher order chromatin compaction (Robinson *et al*., 2008) and modulates the interaction of chromatin-associated proteins such as chromatin-remodelling enzymes (Shogren-Knaak *et al.*, 2006; Shogren-Knaak and Peterson, 2006). In addition, H4K16 acetylation also contributes to the prevention of heterochromatin propagation associated with silencing in yeast (Shia *et al.*, 2006). Similarly, in budding yeast, H3K4 and H3K36 methylation are required prior to H3K14 acetylation by the histone acetyl-transferase complex NuA3 (Martin *et al.*, 2006). NuA3 is recruited to these marks through the PHD finger of its subunit Yng1p, which binds H3K4me (Martin *et al.,* 2006). Modifications of these residues may also be required for H4K8 acetylation (Morillon *et al.,* 2005; Latham and Dent, 2007).

In many cases, the mechanistic basis of these interactions is unknown, but it is assumed that a combination of multiple marks generates a "platform" for the recognition and the recruitment of modification-recognising "effector" proteins, thus leading to the recruitment of other factors, or to chromatin rearrangement.

1.2. Effector proteins

Histone modification is a dynamic process which is influenced by the environment and is associated with a biological response. A modification can be deposited on a residue depending on the functional status of the chromatin, and then be removed or lost when this condition is changed (Nightingale *et al*., 2006; Vermilyea *et al*., 2009). Specific histone-

modifying enzymes mediate the regulation of these modifications. Thus, histone modification acts as a platform for the recruitment of modification-specific "effector proteins", which then influence the adjacent chromatin environment and its functional role (Jenuwein & Allis, 2001). These subsequent effects can be both "activating" or "silencing". Effector proteins are involved in the recruitment of specific enzyme activities to chromatin (e.g. HATs or chromatin remodelling complexes), but they can also maintain repressed genes or stabilize heterochromatin (e.g. HP1, or Polycomb-group proteins). This can favour the establishment of processes such as transcription, repair (Tsukada *et al*., 2005), replication (Goren *et al.,* 2008), and gene silencing. Complexes containing histone modification enzymes have been shown to associate with transcription factors as well, linking functional regulators and histone modification (Kuo *et al.,* 1996; Spencer *et al*., 1997; Kouzarides, 2007).

Effector proteins are recruited to specific marks via specialised domains that recognise both the residue, and the precise post-translational modification. These "domains" were historically recognised in many chromatin-binding proteins, via their high degree of conservation, and act to recruit both the protein, and its associated complex to specific loci.

1.2.1. Bromodomain-containing proteins

Acetylation was the first histone modification identified, and was proposed to play a role in transcription (Allfrey *et al*., 1964). Lysine histone marks can be identified and "read" by specialised protein domains called bromodomains, found in a wide range of chromatin-associated proteins, in particular in histone acetyl-transferases (Dhalluin *et al.,* 1999). The bromodomain was the first domain recognised to interact with histone modifications, being specific for lysine acetylation marks (Dhalluin *et al*., 1999). This conserved protein motif consists of around 110 amino acid residues and is found in a wide variety of chromatin-

associated proteins, such as histone acetyl-transferases, chromatin remodelling complexes and transcription factors (Mizzen *et al*., 1996; Jacobson *et al*., 2000).

The first histone-binding motif to be structurally characterised was the bromodomain of the p300/CBP-associated factor, PCAF (homologue of the *S. cerevisiae* histone acetyl-transferase Gcn5p) (Dhalluin *et al.*, 1999). Bromodomains display a conserved structure, constituted of a cluster of four helices, with a left-handed twist, and two interhelical loops at the end of the structure, forming a hydrophobic binding pocket, into which acetyl-lysine marks are inserted (Dhalluin *et al.*, 1999; Zeng *et al.*, 2008). This is kept in place in the binding pocket through a hydrogen bond, constituting an anchoring point to stabilise the interaction between the residue and the protein (Taverna *et al*., 2007; Zeng *et al*., 2008). However, despite the conservation of this structure, the amino acid composition of the loops is able to discriminate for the binding of specific residues (Mujtaba *et al*., 2007). Distinct bromodomain-containing proteins selectively bind acetyl lysine marks depending on the other residues surrounding the modified lysine (Mujtaba *et al*., 2007; Zeng *et al*., 2008; Zhang *et al*., 2010). Gcn5p recognises H4K16 acetylation (Owen *et al*., 2000), whereas PCAF (p300/CBP-associated factor) binds the acetylated lysine 50 residue of the HIV1 (human immunodeficiency virus type 1) transactivator protein Tat (Mujtaba *et al*., 2002). Similarly, bromodomain-containing transcription complexes also bind selectively to specific acetylated residues. For example, the BRG1 subunit of the SWI/SNF complex binds H4K8ac, whereas TAF1 ($TAF_{II}250$; Jacobson *et al.,* 2000; Agalioti *et al*., 2002; Taverna *et al*., 2007), or the TAF1 homologue, Brdt (Moriniere *et al*., 2009), both of which contain two bromodomains, bind two adjacent acetylated marks.

1.2.2. Chromodomain-containing proteins

Chromodomains (for Chromatin Organisation Modifier) are highly conserved 50 amino-acid protein motifs, originally defined by regions of high sequence homology found within two *Drosophila* proteins, Pc and HP1 (Pearce *et al*., 1992). Chromodomains were then described in a variety of other proteins. In the mouse HP1-like protein, the chromodomain adopts a three-stranded anti-parallel ß-pleated sheet structure backing onto an α-helix. The domain is composed of a highly conserved hydrophobic core. Differences between chromodomains are mostly found in the bends connecting the sheets (Jones *et al*., 2000).

H3K9 is a substrate for the SUV39H1 histone methyl-transferase (Suppressor of Variegation 3-9 Homolog 1, or SuVar 3-9), which generates a mark (H3K9me3) recognised by a highly conserved chromodomain on heterochromatin-associated protein 1 (HP1) (Nielsen *et al*., 2002; Lu *et al*., 2009). Recruitment of this factor favours heterochromatin formation and gene repression (Lachner *et al.,* 2001). However, chromodomains are also specific for a number of methylated marks, for example the Polycomb (Pc) chromodomain recognises histone H3K27me3, a mark deposited by Enhancer of Zeste [E(Z)], and both proteins show some flexibility of binding as HP1 and Pc chromodomains target di- and tri-methyllysines (Lachner *et al*., 2001; Margueron *et al*., 2005; Taverna *et al.,* 2007). Chromodomains are also found in histone acetyl-transferases. In the yeast SAGA (Spt-Ada-Gcn5-acetyltransferase) complex, the Chd1p (Chromo-ATPase/helicase-DNA binding domain 1) subunit contains two N-terminal chromodomains (Pray-Grant *et al.*, 2005; Tran *et al.*, 2000; Woodage *et al.*, 1997). In human CHD1, these two chromodomains are bridged by a two-helix linker element, forming a helix-turn-helix structure. This linker segment juxtaposes the two chromodomains, allowing them to cooperate by constituting a new recognition pocket to interact with H3K4me (Flanagan *et al*., 2005; Taverna *et al*., 2007). It should be noted that

the phosphorylation of the adjacent H3T3 residue weakens the binding of CHD1 for H3K4me (Flanagan *et al*., 2005). This is also seen in the *Drosophila* and human CHD1 orthologues. Chd1p binds H3K4, with a preference for di and trimethylated residues, reflecting the specificity of one of its chromodomains (Pray-Grant *et al.,* 2005). Interestingly, more divergent chromodomains are thought to bind RNA. In *Drosophila*, the histone acetyl-transferase MOF interacts *in vivo* with RNA through its chromodomain, an interaction, which may contribute to the recruitment of MOF to the Dosage Compensation Complex (DCC) at the *Drosophila* male X chromosome (Akhtar *et al*., 2000).

1.2.3. Tudor domain-containing proteins

Methyl-lysine marks are also recognised by proteins involved in their removal. In the JMJD2 family, this recognition is mediated by the catalytic-core region, which comprises substrate binding sites and the JmjN and the JmjC catalytic domain (Chen *et al.*, 2006). For example, H3K4me3 and H4K20me3 are individually targeted by the hybrid Tudor domains of the histone lysine demethylase JMJD2A (Taverna *et al*., 2007; Adams-Cioaba and Min, 2009). This 50 amino-acid conserved motif was first identified in the *Drosophila* Tud protein (Boswell and Mahowald, 1985; Adams-Cioaba and Min, 2009). JMJD2A, which is a protein of the Jmjc histone demethylase superfamily, contains two Tudor domains. This double tudor domain binds methylated H3K4 and H4K20 (Huang *et al*., 2006; Kim *et al.*, 2006) by adopting a two "lobed" structure, where only the second "lobe" can interact. This is a complex structure, requiring residues from both Tudor motifs to generate an "aromatic cage" binding pocket, used for the insertion of a trimethyl-lysine residue (Huang *et al*., 2006; Taverna *et al*., 2007).

1.2.4. PHD finger-containing proteins

Methyl-lysine marks can also be read by PHD-finger-containing proteins. Plant homeodomains (PHD) are constituted of around 60 amino acid residues, characterised by a motif consisting of four cysteines, one histidine, three cysteines ($Cys_4HisCys_3$). This motif binds two Zn^{2+} ions, coordinating the structure of the PHD finger. The PHD finger displays two flexible loops, the second loop being used to interact with specific nuclear proteins (Bienz, 2006). PHD fingers promote the recruitment of enzyme complexes by binding methylated marks, for example, the PHD fingers of BPTF and ING2 are composed of a H3K4me-binding pocket, containing two aromatic residues forming an aromatic cage (Taverna *et al*., 2007).

PhD finger binding is now known to be highly specific; the human protein BPTF and its *Drosophila* orthologue NURF301, a subunit of the NURF chromatin-remodelling complex, contain a C-terminal PHD finger, which binds preferentially to H3K4me3, rather than the H3K4me2 mark. This is consistent with its ability to stimulate transcription *in vitro* (Li *et al*., 2006), and the contribution of H3K4 methylation to transcription in *S.cerevisiae* (Santos-Rosa *et al*., 2002; Pinskaya and Morillon, 2009). Moreover, the PHD finger domain is situated close to the C-terminal bromodomain, which suggests that these two domains may function in a cooperative manner in the recognition of histone marks (Wysocka *et al.,* 2006; Taverna *et al*., 2007). In addition, the alternative splicing of the NURF301 subunit leads to the formation of two functionally distinct NURF complexes, one containing the full-length NURF301 protein, which recognises H3K4me3 as well as H4K16ac, and a complex containing a truncated form of NURF301, which cannot target these residues (Kwon *et al*., 2009). In contrast, the human ING2 protein, a subunit of the histone acetyl-transferase complex Sin3 also interacts with H3K4me3, but is linked to gene repression (Pena *et al*.,

2006; Shi *et al*., 2006; Taverna *et al*., 2007). More recent data confirms that PHD fingers can recognise the degree of methylation in a specific manner. For example, in contrast with BPTF and ING2, which bind preferentially to di- and tri-methylated H3K4, another PHD domain-containing protein binds specifically to non-methylated H3K4 (H3K4me0). The autoimmune regulator (AIRE) protein possesses two PHD fingers, through which the histone H3 is bound. The recruitment of AIRE to non-methylated H3K4 on the tissue-restricted antigen promoter then favours the recruitment of other factors, such as the histone acetyl-transferase CBP and is associated with transcription activation (Musco and Peterson, 2008; Org *et al.*, 2008).

In addition to chromatin-remodelling complexes, PHD fingers are also used by transcription factors to recognise histone modifications. TAF subunits of the basal transcription factor TFIID bind post-translational histone marks through different motifs. TAF3 binds H3K4me3 and H3K4me2 through its C-terminal PHD finger domain (Vermeulen *et al*., 2007). Interestingly, H3R2 dimethylation prevents TAF3 binding, by impairing H3K4 recognition, but does not impact on the binding of two other PHD-finger-containing proteins, ING2 and BPTF (Vermeulen *et al*., 2007). As discussed previously, TAF1 and its homologue Brdt, (Moriniere *et al*., 2009) contain a double bromodomain, suggesting that the cooperative recognition of multiple marks (H3K4me3 and H3 or H4 acetylation) by subunits of TFIID favours the interaction of the complex with transcriptionally active promoters (Vermeulen *et al*., 2007).

1.2.5. Phosphorylation-targeting proteins

Histones can also be modified by the phosphorylation of serine and threonine residues, marks with a range of functional effects. For example, H3S10p is linked with early gene expression (Clayton *et al*., 2000; Soloaga *et al*., 2003), early transcription elongation (Ivaldi *et al*., 2007;

Zippo *et al*., 2009), and the onset of mitosis (Di Croce, 2005). Transient phosphorylation of H3S10 during early gene induction is mediated by mitogen-and stress-activated kinases 1 and 2 (Msk1/2) (Soloaga *et al*., 2003; Taverna *et al*., 2007). The H3S10p mark is recognised by conserved phosphoserine-binding modules, present in the mammalian 14-3-3 proteins (Bmh1p and Bmh2p in yeast) (Macdonald *et al*., 2005; Winter *et al*., 2008). Furthermore, the affinity of the binding is increased by adjacent acetylated H3K9 and H3K14 residues, suggesting that these acetylated marks stabilise the interaction (Winter *et al.*, 2008). These proteins have the capacity to homo or heterodimerise, binding to at least two phosphoserine sites. 14-3-3 proteins also bind to H3S28p (Taverna *et al*., 2007).

Thus to summarise, the interactions between histone modifications and specific effector proteins are complex. Specific histone residues are recognised by distinct domains. In addition, the degree of modification and the residues flanking a target histone mark influence this selective recognition, allowing specific functional proteins to be recruited. This specificity can be crucial when two proteins are competing to bind an adjacent modified residue. Modified residues can prevent the recruitment of some effectors and/or stimulate the recruitment of other proteins, potentially with diametrically opposite functional effects.

1.3. Modifying enzymes

The abundance of a specific histone modification at a locus is determined by the balance of activities of two classes of opposing enzymes - depositing and modification-removing enzymes. These are specific to particular classes of modification; for example, histone acetyl-transferases catalyse the addition of acetyl marks, while histone methyl-transferases catalyse the addition of methyl marks. Other enzymes exist to catalyse the modification of other type

of histone marks. Furthermore, these marks can also be removed, in processes that involve other types of enzymes, histone deacetylases or histone demethylases (Figure 1.10).

1.3.1. Regulation of histone acetylation

Histone acetyl-transferases (HATs)

Histone acetylation is primarily at the histone tail (except the core-located H3K56 in yeast and *Drosophila* (Xu *et al.*, 2005)) and is catalysed by Histone Acetyl Transferases (HATs) while its removal is catalysed by Histone Deacetylases (HDACs) (Turner, 2002). Histone acetyl transferases catalyse the transfer of acetate groups from the cofactor acetyl Coenzyme A (acetyl CoA) to histones (Figure 1.11). Functionally it is involved in modulating chromatin structure. Early observations indicated that HATs were involved in transcriptional regulation, as actively transcribed regions of chromatin were found to be associated with hyperacetylation, while transcriptional silent regions were associated with hypoacetylation (Hebbes *et al*., 1988; Turner, 1991; Davie, 1997; Turner, 1998; Utley *et al*., 1998). This was more recently followed by the finding that transcription activators, like yeast Gcn5, acetylate histones (Kuo *et al*., 1996). Two processes were suggested to explain the link between histone acetylation and transcription regulation. An earlier hypothesis suggested that acetylation neutralises lysine's positive charge, resulting in the reduction in histone - DNA interactions, thus leading to the formation of an accessible chromatin conformation. The second proposition proposes that histone modifications function as a platform for recognition by effector proteins. Experimental data suggests that both of these hypotheses are correct for histone acetylation.

Histone acetyl-transferases constitute a diverse series of proteins that form five families: Gcn5-related acetyltransferases (GNATs), MYST-related HATs, p300/CBP HATs, general

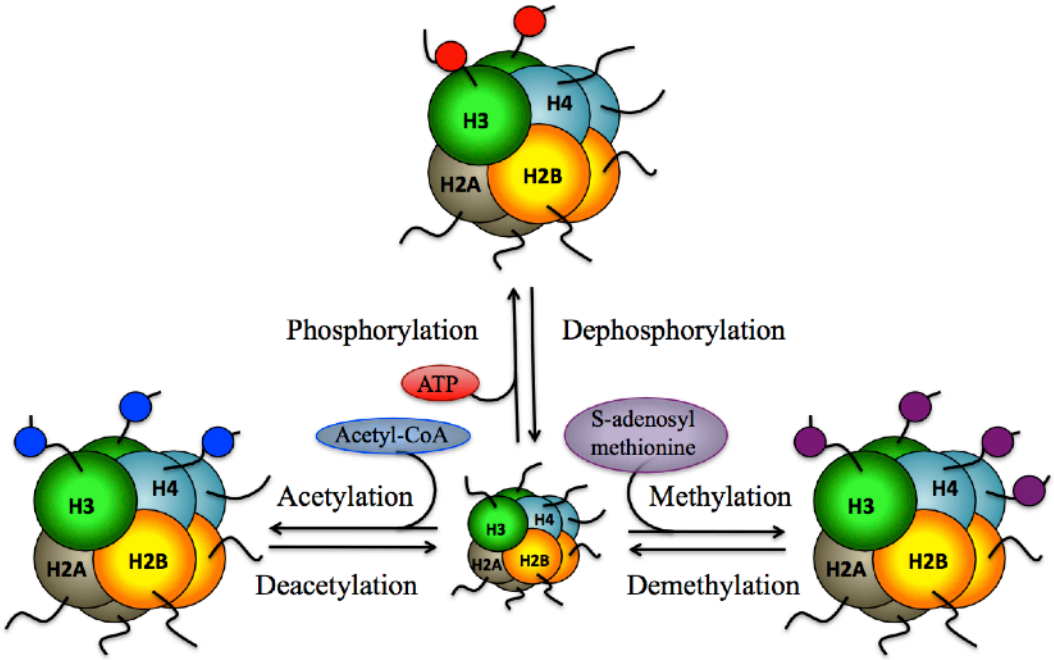


Figure 1.10: A dynamic mechanism of post-translational modifications

Histone and non-histone proteins can be post-translationally and reversibly modified. These modifications are highly dynamic: modifications can be added and then removed by specific enzymes. Three major mechanisms have been especially well-studied. Histone acetylation corresponds to the transfer of an acetyl group to a histone residue and is opposed to deacetylation. Histone lysine and arginine methylation correspond to the transfer of a methyl group from an S-adenosyl cofactor to a lysine or arginine residue. This process is reversed by demethylation. Histone phosphorylation, in which the transfer of a phosphate requires ATP, is reversed by dephosphorylating phosphatase enzymes.

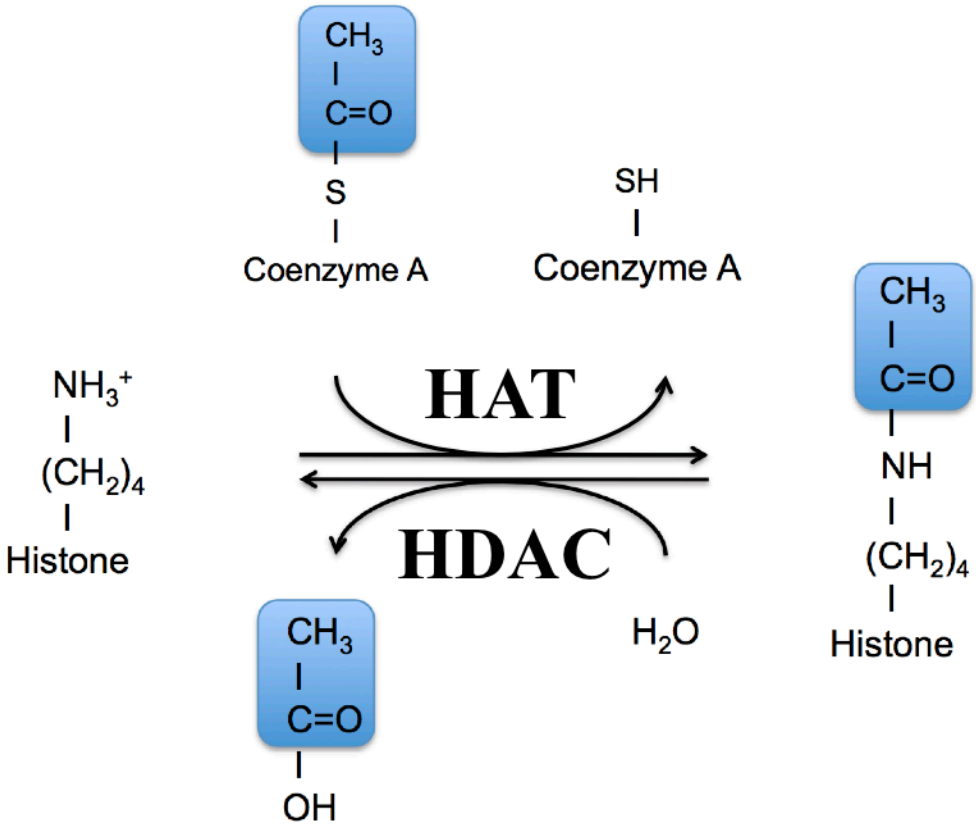


Figure 1.11: Acetylation and deacetylation mechanism

Histone and non-histone proteins can be post-translationally and reversibly modified. Acetylation is catalysed by one of a broad family of structurally diverse histone acetyl-transferases (HATs), while histone deacetylation is catalysed by histone deacetylases (HDACs). Acetylation corresponds to the transfer of an acetate group from the cofactor acetyl-Coenzyme A to the protein.

transcription factors HATs and nuclear hormone-related HATs (Carrozza *et al*., 2003). The finding that a wide range of enzymes is involved in regulating essentially the same chemistry suggests that they act in different processes or at specific loci (Lee and Workman, 2007). Experimental data suggests that there is functional redundancy between some histone acetyl-transferases.

Most acetyl-transferases act as members of multiprotein complexes, whose members are essential in the selective targeting of genes (Berndsen and Denu, 2008). Thus, enzymes function in a cooperative manner to recognise specific histone modifications and generate an appropriately targeted, regulated effect. Acetyl-transferases can interact with histone methyl-transferases, in order to cooperate in the regulation of gene transcription. For example MOF, a histone acetyl-transferase from the MYST family, catalyses the acetylation of a specific histone mark at H4K16 (Carrozza *et al*., 2003). MOF is also associated with the histone methyl-transferase MLL1 (Dou *et al*., 2005). Similarly, the broad specificity histone acetyl-transferase, CBP, which catalyses the acetylation of multiple histone (Das *et al*., 2009; Tie *et al*., 2009) and non-histone proteins (Gu and Roeder, 1997), can also associate with MLL1 (Ernst *et al*., 2001).

Histone deacetylases (HDACs)

Histone deacetylases (HDACs) (i.e. Figure 1.11) can be grouped into two families, the Silent Information Regulator Two (Sir2) family of NAD^- dependent HDACs, or "sirtuins", and the initial HDAC family. This second family is divided into class I (HDACs 1, 2, 3 and 8, Cunliffe, 2008) and class II (HDACs 4, 5, 6, 7 9, 10). In yeast, the principal deacetylases are Rpd3 and Hda1 (Kurdistani and Grunstein, 2003), where Rpd3 belongs to class I. Class I HDACs are principally found in the nucleus, while class II can transit in and out of the

nucleus (De Ruijter *et al*., 2003). As with histone acetyl-transferases; histone deacetylases are organised in complexes to mediate transcriptional silencing: e.g. the Sin3 complex (Zhang *et al*., 1997), the chromatin remodeller NuRD (Xue *et al*., 1998) and Co-REST (You *et al*., 2001). Hos2, a HDAC that deacetylates histones H3 and H4, was shown to be associated with coding region of genes only during active transcription in yeast (Kurdistani and Grunstein, 2003). Acetyl-transferases and deacetylases also have other substrates. The first non-histone target for acetyl-transferase identified was p53 (Gu and Roeder, 1997), and is now known to be essential for its activation (Tang *et al*., 2008). Thus post-translational acetylation and deacetylation can affect diverse factors, such as transcription factors like YY1, cellular proteins like Hsp90 and tubulin, and viral proteins (Glozak *et al.,* 2005). For example, some HATs and HDACs are associated with DNA binding proteins such as the transcription factor YY1 (Mokrani *et al*., 2006), whose regulation by acetylation and deacetylation influences its ability to bind DNA (Yao *et al*., 2001). The level of histone acetylation at specific loci in chromatin is generated by a balance of the activity between histone acetyl-transferases and histone deacetylases, modulating chromatin conformation and gene regulation. Other histone modifications are also involved in transcriptional regulation of genes including methylation and demethylation.

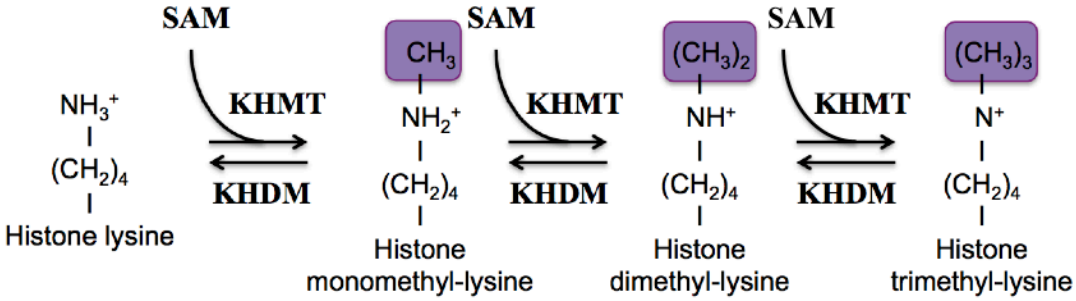
1.3.2. Regulation of methylation

Histone methylation is catalysed by Histone Methyl Transferases (HMTs) on arginine and lysine residues. This reflects the activity of two families of Histone Methyl Transferases: the histone lysine methyl transferases (HKMTs), and a family of protein arginine methyl-transferases (PRMTs, Krause *et al*., 2007). We will essentially focus on lysine modification.

Lysine histone methyl-transferases (KHMTs)

Lysine methylation is the most studied process of methylation on the histone N-terminal tails. This process corresponds to the transfer of one to three methyl groups from the cofactor S-adenosyl methionine (SAM) to lysine by a histone lysine N-methyl-transferase (Figure 1.12). This is catalysed by the SET domain (Rea *et al.*, 2000). The evolutionarily conserved SET domain was initially identified in modifier of PEV (Position Effect of Variegation) as a region of homology to the C-terminal regions of three proteins; *Drosophila* histone Su(var)3-9, Enhancer of zeste [E(z)] and Trithorax (Trx) (Tschiersch *et al*., 1994). Human SUV39H1 and *S. pombe* Clr4, were the first histone methyl-transferases identified containing a SET domain, suggesting that other proteins containing this domain might have histone methyl-transferase activity (Rea *et al*., 2000). These HMTs catalyse methylation of H3K9. More recently, other HMTs have been identified, like *Drosophila* Ash1, which catalyse methylation of H3K4,K9 and H4K20 (Beisel *et al*., 2002), and the human histone H3K9-specific methyl-transferases G9a (Tachibana *et al*., 2001) GLP, Suv39H2 and PRDM2 (Wu *et al*., 2010), though there are indications that these protieins are not all specific to core histone substrates (Huang *et al*., 2010; Trojer *et al*., 2009). This domain of homology was also identified in the human homologue of TRX, ALL1/MLL (Tschiersch *et al*., 1994).

Studies in the yeast *S. cerevisiae* indicate broad similarities between mammalian and yeast histone methyl-transferases. Sequence alignment identified six genes containing SET domains, termed *set1-6* in *S.cerevisiae* (Pijnappel *et al*., 2001). In particular, the protein Set1 was found to display similarities with the SET domain of Trx. Set1 is part of a multi-protein complex, the Set1 complex (Set1C) or COMPASS (Complex Proteins Associated with Set1), which has histone methyl-transferase activity, specifically catalysing the deposition of methyl groups on H3K4 (Miller *et al*., 2001; Roguev *et al.*, 2001). Subunits of COMPASS are



Methylation

Figure 1.12: Methylation by HMTs

Histones can be post-translationally and reversely methylated on their lysine residues. Lysine mono, di and tri-methylation is catalysed by histone lysine methyl-transferases (KHMTs). This process corresponds to the transfer of one to three methyl group(s) from the cofactor S-adenosyl Methionine (SAM) to the lysine of a histone. Methyl groups can be removed by one, or more lysine demethylases (KHDMs), through mechanisms described in the following figures.

similar to Trithorax associated proteins (Miller *et al.*, 2001) and include a H3 binding subunit Swd2 (Vitaliano-Prunier *et al*., 2008), and Wdr82, which is required for the recognition of ubiquitinated H2B in chromatin (Wu *et al*., 2008). SET domain-containing proteins present some structural differences though: the SET domain of *Drosophila* SUVH9H1 possesses a cysteine-rich motif in its N-terminal region (Rea *et al*., 2000), which is required for histone methyl-transferase activity, whereas this "pre-SET" domain is not essential in Set1 (Roguev *et al*., 2001). Similar to COMPASS, these characteristics and specific enzymatic activity can be found in the MLL complex, which will be discussed later.

Histone demethylases (HKDMs)

Until recently, histone methylation was thought to be irreversible, although histone demethylase activities were detected in whole cell extracts in the early 1970's (Paik and Kim, 1973). Two families of proteins displaying histone lysine demethylase activity (HKDM) have been identified: LSD1 (Shi *et al*., 2004), and the JMJC class, JHDM1 (Tsukada *et al*., 2006). These proteins are structurally and functionally distinct, using different chemistries to act on distinct histone modifications (Marmorstein and Trievel, 2009). LSD1 induces lysine 4 demethylation on histone H3 via an amine oxidation that removes methyl groups (Shi *et al*., 2004), while JHDM1 is implicated in H3K36 demethylation via a hydroxylation mechanism (Tsukada *et al*., 2006).

LSD1 specifically demethylates histone marks associated with active transcription, such as H3K4me but not H3K9me and it was first proposed that flanking residues can modulate substrate specificity (Shi *et al*., 2004). However, subsequent experiments showed that residue-specificity can also be modified by proteins associated with LSD1, for example, the association of LSD1 with the androgen receptor leads to the removal of methyl marks on

H3K9, and to the derepression of target genes (Metzger *et al*., 2005). Furthermore, this histone demethylase also discrimates between di- and tri-methylation. LSD1 can demethylate mono- or di-methylated histones, but not tri-methylated histone residues. This reflects its amine oxidase activity; LSD1 would need a protonated nitrogen to demethylate tri-methyl residues (Shi *et al*., 2004). LSD1 is a flavin-containing protein, which can bind the cofactor FAD in order to catalyse the amine oxidation reaction (Shi *et al*., 2004) (Figure 1.13). Recent biochemical characterisation indicates that LSD1 is a component of the NuRD chromatin remodelling and deacetylase complex (Wang *et al*., 2009).

The JmjC domain is evolutionarily conserved and more than 100 proteins containing this domain have been identified (Tian and Fang, 2007). Demethylation involves binding to Fe(II) and α-ketoglutarate cofactors through its JmjC domain, and demethylation via hydroxylation leads to the formation of succinate and formaldehyde (Figure 1.14) (Tsukada *et al*., 2006). Ongoing biochemical characterisation indicates that these proteins are specific for a wide variety of residues - JHDM1A preferentially demethylates H3K36me2 (Tsukada *et al*., 2006), whereas JHDM2A, can specifically demethylate mono- and di-methylated H3K9 (Yamane *et al*., 2006), JMJD3 is specific for H3K27me (Burgold *et al*., 2007) and GASC-1 was shown to demethylate di- and tri-methylated H3K9, to form mono-methylated H3K9 (Cloos *et al*., 2006). Similarly, JHDM3A (or JMJD2A) is necessary for the demethylation of trimethylated H3K9 and H3K36 (Klose *et al*., 2006). Perhaps not surprisingly, JMJD2 proteins are also found to demethylate residues on linker histones (Trojer *et al*., 2009).

In summary, transcriptional regulation is modulated by distinct processes, involving a diverse range of proteins depositing or removing histone methyl or acetyl groups. These specific enzymes modify residues selectively, depending on their degree of modification, and with a specificity determined by enzyme-associated subunits. Thus, the association of catalytic

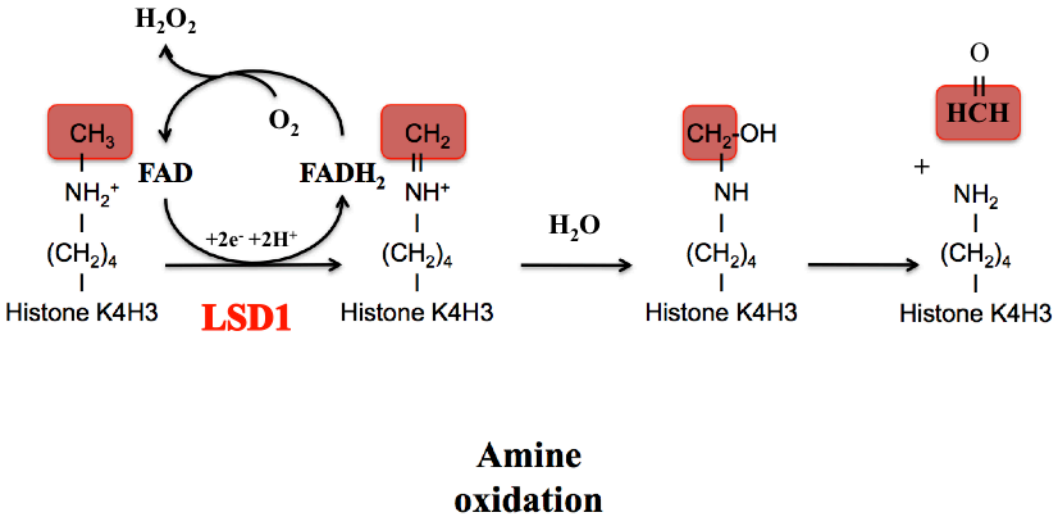


Figure 1.13: Demethylation by LSD1 (based on Shi Y. *et al.*, Cell, 2004)

Mechanism of LSD1-catalysed demethylation of a monomethylated lysine. Amine oxidation is catalysed by the flavin-containing amine oxidase LSD1 and the cofactor FAD, generating an imine intermediate. This substrate is then hydrolysed to form an amine and an aldehyde (HCHO). For simplicity, the mechanism is represented for demethylation of monomethylated lysine, but the same mechanism is used for dimethylated residues.

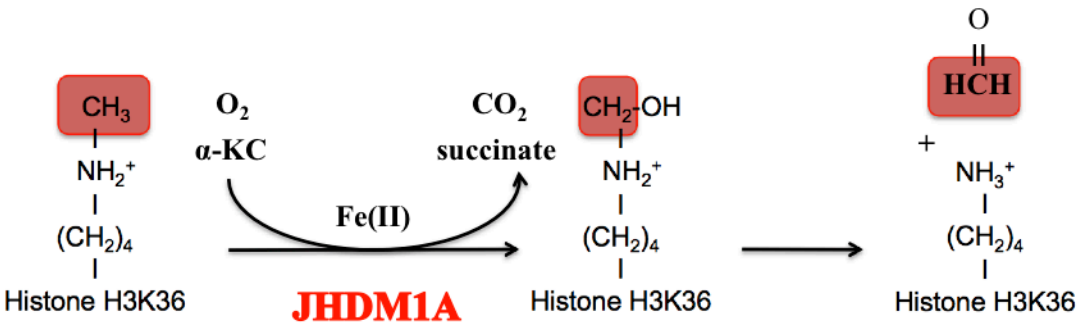


Figure 1.14: Demethylation by JHDM1A (based on Tsukada Y. *et al.*, Nature, 2006)

Mechanism of JHDM1A-catalysed demethylation of a monomethylated lysine. The amine oxidation is catalysed by JHDM1A and the cofactors Fe(II) and alpha-ketoglutarate (**α-KC**), generating an imine intermediate. This substrate is then hydrolysed to form an amine and an aldehyde (HCHO). For simplicity, the mechanism is represented for demethylation of monomethylated lysine, but the same mechanism is used for dimethylated residues.

enzymes with other factors into a multiprotein complex constitutes an essential and complex means of regulating chromatin remodelling and gene regulation.

1.4. The MLL family of histone methyl transferases

1.4.1. MLL1 protein (KMT2A)

Mixed Lineage Leukaemia (MLL1 or ALL1, HRX) is the archetypal protein of the family of MLL histone methyl-transferases. *MLL* was originally identified at the site of reciprocal chromosomal translocation frequently observed in 10% of human acute myeloid (AML) and acute lymphoid (ALL) leukaemias, as well as in therapy-induced secondary leukaemia treated with DNA topoisomerase II inhibitors. This gene was found to span the 11q23 region, a chromosomal breakpoint involved in leukaemia (Ziemin-Van Der Poel *et al*., 1991). These rearrangements involving *MLL1* result in the generation of oncogenic fusion proteins, corresponding to the expression of the N-terminal region of MLL fused to one of over 50 partners. As such the catalytic C-terminal region is lost. Chromosomal translocations can also lead to the generation of partial genomic tandem duplication of *MLL1* (PTD), corresponding to the association of an additional MLL1 N-terminal region to MLL1 (Slany, 2009).

At a protein level, MLL1 is a histone methyl-transferase (HMT) belonging to the MLL family. MLL is the mammalian homologue of *Drosophila* Trithorax (TRX), a positive regulator of gene expression, especially the homeotic *Hox* genes (Gu *et al*., 1992; Tkachuk *et al*., 1992). TRX is one of a group of proteins involved in "cellular memory", acting to maintain pre-established transcription patterns during development (Kundu and Peterson, 2009). *MLL* encodes for a large protein of 3,969 residues with an expected mass of about 430kDa. This large protein is post-translationally cleaved in the cytoplasm into a large 320kDa N-terminal protein (MLL^N) and a small 180kDa C-terminal protein (MLL^C), which

is achieved by a specialized threonine-aspartase endopeptidase TASPASE1 (Hsieh *et al*., 2003) (Figure 1.15a). The two polypeptides then non-covalently associate in the nucleus, forming an active heterodimeric complex (Nakamura *et al*., 2002; Hsieh *et al*., 2003; Slany, 2005a), and stabilising each other (Hsieh *et al*., 2003). Interestingly MLL^C and MLL^N have antagonistic transcriptional properties (Yokoyama *et al*., 2002).

MLL bears multiple conserved domains with enzymatic activities (Figure 1.15b), with several domains identified in the MLL^C fragment involved in epigenetic modification of chromatin. MLL^N contains three AT hook motifs involved in DNA binding (Tkachuk *et al.*, 1992), a bromodomain, three PHD zinc finger domains (Fair *et al*., 2001) and a region with homology to DNA methyltransferase (CxxC), while MLL^C contains the SET domain involved in the methylation of H3K4 (Milne *et al*., 2002; Nakamura *et al*., 2002), a binding site for the histone acetyltransferase CBP and a domain with transcriptional activation capacity (TAD) (Slany, 2005b). The structure of the SET domain of MLL1 has evolved to optimise association with both the target lysine and the SAM cofactor, and thereby enables the optimal methyl-transferase activity of MLL1. The domain comprises three distinct regions, SET-N, SET-I and SET-C, whose residues form a channel holding the target lysine residue in position for the efficient transfer of the methyl group (Southall *et al*., 2009). Among these regions, the orientation of SET-I is essential for the proper association with the substrate, modulating the accessibility of the channel, and optimising MLL1 activity, though it is thought that this modulation requires the intervention of other factors (Southall *et al*., 2009).

In mice, double knockout (*Mll* $^{-/-}$) is embryonic lethal at 10.5 day post coitum, this being associated with defects in yolk sac (where blood cells are first detected during embryogenesis), foetal liver haematopoiesis, and segmental identity (Hess *et al*., 1997),

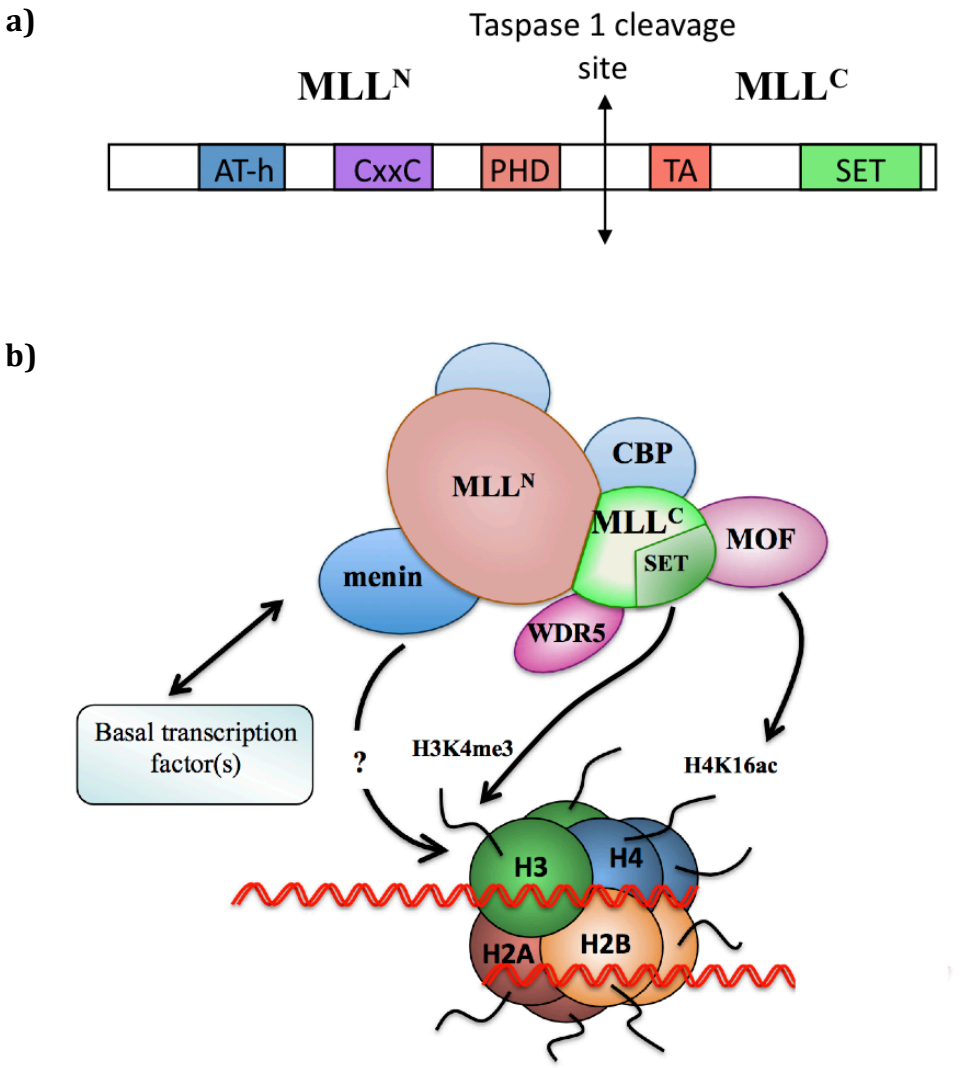


Figure 1.15: MLL1 and MLL complex (adapted from Slany RK, 2005)

a) MLL1 is post-translationally cleaved by a specific threonine aspartase enzyme, Taspase 1, into a large N-terminal fragment (MLL^N) and a small C-terminal fragment (MLL^c). MLL contains several functional domains, the most important being the SET domain involved in the methylation of lysine 4 on histone 3. MLL1 also contains a TA domain that binds CBP, an AT hook domain that binds DNA, a PHD domain that may be involved in protein-protein interactions, and a CxxC domain that selectively binds non-methyl-CpG DNA. The two polypeptides MLL^c and $MLL\text{-}^N$ are then non-covalently bound in an active heterodimeric complex.

b) MLL is associated with 29 other proteins in a large complex. Among them are found Histone Acetyl-Transferase (HAT) proteins such as MOF and CBP, proteins involved in DNA binding such as menin and proteins involved in histone modification or the recognition of histone marks. WDR5 is implicated in complex stabilisation, as well as the direct contact with the H3 substrate.

whereas heterozygous *Mll*$^{-/+}$ mice are characterised by retarded growth, homeotic transformations of the body axis and defects in haematopoiesis (Yu *et al*., 1995). This highlights the central role of MLL1 in development, and as with TRX, MLL is now known to act at the *Hox* genes (Yu *et al*., 1998).

The histone methyl-transferase MLL associates with transcriptionally active genes, such as the active promoters of *Hox* genes (such as *Hoxa9*), in a cell-type and differentiation-stage dependent manner (Milne *et al*., 2002; Milne *et al*., 2005), and contributes to their regulation (Milne *et al*., 2002). Enzymatic H3K4 methyltransferase activity was first described in *S. Cerevisiae,* where a protein with MLL homology, Set1 was shown to deposit H3K4 methylation in the promoter and the coding region of genes (Briggs *et al*., 2001). More recent studies have shown that Set1 is a component of the COMPASS complex, which is capable of generating mono-, di-, and tri-methyl H3K4 marks (Schneider *et al*., 2005), and this has functional effects by recruiting HDAC activities to the 5' transcribed regions of genes (Kim and Buratowski, 2009). The nature of the mark deposited is key as H3K4 di-methylation is associated with both active and inactive genes, while the tri-methyl mark is associated with active genes (Santos-Rosa *et al*., 2002). A similar study showed that MLL1 binds at the transcription start site and on 5' regions of actively transcribed genes, correlating with an enrichment of H3K4 trimethylation (Guenther *et al*., 2005). H3K4 specific methyl-transferase activity was then further confirmed by Milne *et al.* (Milne *et al.*, 2002). Despite its diverse functional domains, MLL1 acts in a large complex, requiring the association of other factors to function optimally.

1.4.2. The MLL1 histone methyl-transferase complex

As with yeast COMPASS, MLL is part of a multi-protein complex including many factors involved in chromatin modification and gene transcription (Nakamura *et al*., 2002; Dou *et al*., 2005). A MLL-containing complex purified by Yokoyama *et al.*, called MLL/HCP complex, was composed of eight polypeptides displaying similarities with the proteins of the Set1 complex, including proteins such as WDR5, ASH2L, and RBBP5 (Yokoyama *et al.,* 2004) (Figure 1.16). These MLL1-associated polypeptides appear to assemble in a stable subcomplex interacting with MLL^C through its SET domain (Yokoyama *et al*., 2004). This MLL "core complex" is similar to the SET1 family members of histone methyl-transferase complexes.

Of the core complex, the WDR5 subunit is essential for the maintenance of MLL1 complex stability at promoters and facilitates the interaction of SET domain with the H3K4me2 substrate (Wysocka *et al.*, 2005). This activity reflects WDR5's ability to associate with methylated histone H3, and increase the conversion from di to trimethylation of H3K4 by MLL1 (Wysocka *et al.*, 2005). WDR5 possesses a specific WD40-repeat motif, constituted of seven ß-propeller "blades", each composed of four anti-parallel ß-sheets, which interact with the N-terminal of histone H3, at the motif Ala1-Arg2-Thr3-Lys4 of histone H3 through an association of hydrogen bond and van der Waals contacts (Couture *et al*., 2006; Ruthenburg *et al*., 2006). Post-translational modification of H3R2 and H3T3 prevent binding of histone H3K4me2 by WDR5 (Couture *et al.*, 2006). Both methylation of H3R2 and phosphorylation of H3T3 disrupt WDR5 binding histone H3, by either disrupting hydrogen

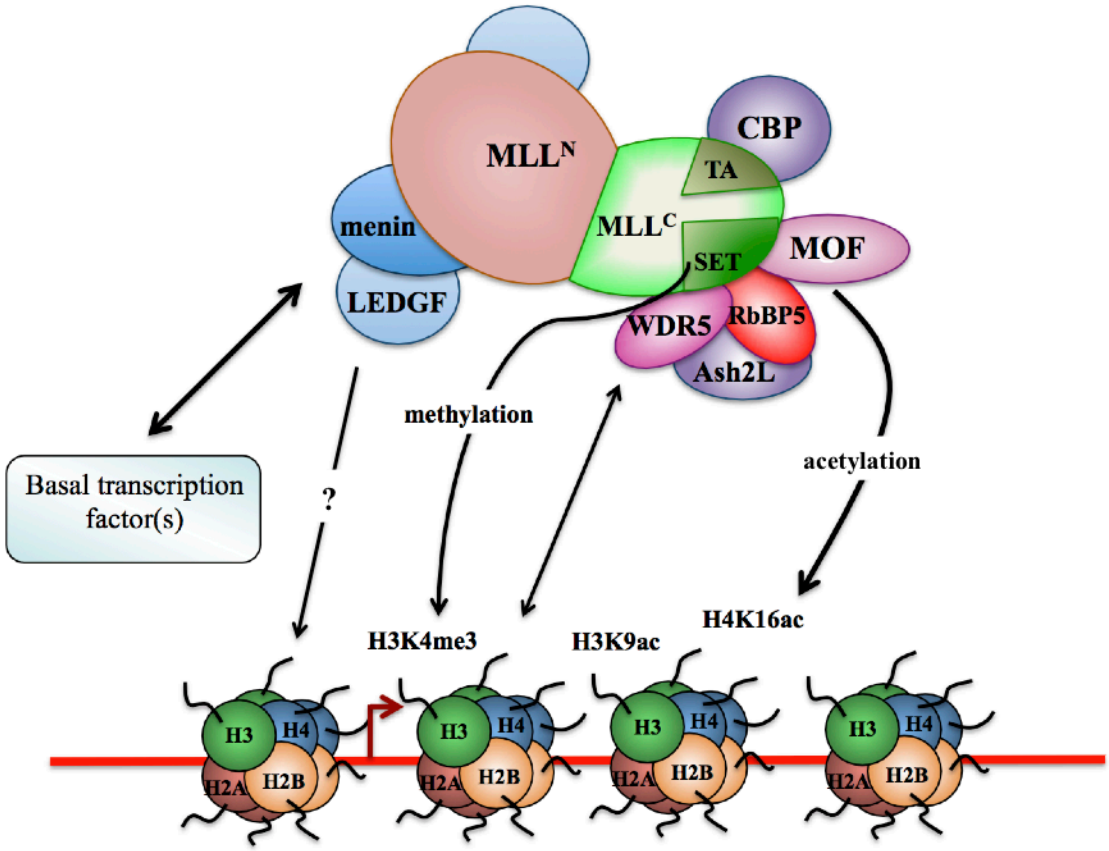


Figure 1.16: MLL1 complex and associated proteins

MLL1 is associated with diverse factors in a multi-protein complex. Some are associated either with the MLL N-terminal region (like menin, or LEDGF through menin), or with the C-terminal region. Specifically, the acetyl-transferase CBP is associated with the transcriptional activation domain, or TA of MLL, whereas the histone acetyl-transferase MOF, WDR5, RbBP5 and Ash2L interact with the SET domain.

bond or ionic interactions or sterically blocking the WDR5 T3 binding pocket (Couture *et al*., 2006). Furthermore, two independent studies have shown that the N-SET region of MLL1 binds the WDR5 binding pocket through a histone H3-like motif (ARAEVHLRKSAFD) (Song and Kingston, 2008), also called WDR5 interaction or "*Win*" motif (GSARAEVHLRKS) (Patel *et al*., 2008; Patel *et al*., 2008), which is conserved among SET1 family members. The WDR5 binding pocket that binds this MLL1 specific motif is also required to bind histone H3, so both compete for WDR5 binding (Patel *et al*., 2008; Song and Kingston, 2008).

In association with WDR5, RbBP5, another WD40–repeat protein, is required to stabilise MLL^{C}'s interaction with the core complex. WDR5 and RbBP5's interaction with MLL depends on the SET domain (Yokoyama *et al.*, 2004). RbBP5 is also required for the association of AshL2 with the core complex, and the activity of both proteins is required for the di and trimethylation of H3K4 (Dou *et al.*, 2006). In particular, Ash2L regulates H3K4 trimethylation activity at the 5' end of genes in human cells (Steward *et al.*, 2006). However, in contrast with MLL1 methylation processes already described in the literature, another process of methylation involving MLL1 and the core complex has recently been described. Patel *et al.* (2009) proposed that the isolated MLL1 SET domain is a H3K4 monomethyl-transferase, which can then acquire H3K4 dimethyl-transferase activity in the presence of the core complex; this is consistent with observation that MLL1 efficiency is impaired without additional factors, like Ash2L, RbBP5 or WDR5 (Southall *et al*., 2009).

Biochemical analysis showed that the presence of a phenylalanine at position 3,942 of MLL1 modulates its histone methyl-transferase activity, allowing it to act as a trimethyl-transferase. It was proposed that MLL1 trimethyl-transferase activity could be generated by the interaction of MLL1 with other proteins, which act to influence the position of the tyrosine

residue in the SET domain (Patel *et al*., 2009). This was reinforced by Dou *et al.* (2005), who proposed that MLL1 needs association with another component to be fully active (Dou *et al*., 2005).

Altogether, WDR5, ASHL2 and RbBP5 form a structurally independent complex, which constitutes a common platform for catalytic HMTs (Dou *et al*., 2006; Steward *et al.*, 2006). In addition to this initial subcomplex, another protein, DPY-30, was proposed to be part of the core complex (Miller *et al*., 2001; Patel *et al*., 2009). This protein is involved in the recruitment of factors to the dosage compensation machinery in *C. elegans* (Hsu *et al*., 1995).

MLL1 is also associated with histone acetyl-transferases such as MOF and CBP, and these are also known to contribute to MLL function. MOF is an acetyl-transferase catalysing H4K16 acetylation. This enzyme is associated with the MLL1 SET domain through its zinc finger (Dou *et al.*, 2005), suggesting that both proteins function in a coordinated manner for the activation of transcription. A model was proposed, in which the MLL1 complex is recruited to target genes either following the binding of WDR5 to H3K4me-marked nucleosomes, or through the binding of the MLL1 bromodomain to H4K16ac-marked nucleosomes (Dou *et al*., 2005). This does not make clear how these marks are originally deposited though. CBP binds to the MLL1 activation domain, corresponding to MLL amino acids 2,829 to 2,883, through a "KIX", or CREB-binding domain (Ernst *et al*., 2001). This interaction is essential for transcriptional activation by MLL1 (Ernst *et al*., 2001), presumably reflecting CBP's ability to acetylate a wide range of non-histone proteins. In addition, it has recently been shown that CBP mediates the acetylation of residues such as H3K56 (Das *et al*., 2009). CBP also mediates acetylation of H3K27 in *Drosophila*, thereby preventing the methylation of the same residue by TRX, the MLL1 orthologue (Tie *et al*., 2009).

The N-terminal fragment of MLL1 is associated with menin, a tumor suppressor protein, which can bind DNA (La *et al*., 2004), interacts with several transcription factors such as NF-κB (Heppner *et al*., 2001) and is proposed to contribute to DNA recruitment of the MLL complex (Yokoyama *et al.,* 2004). Menin is essential for maintaining *Hoxa9* expression suggesting that it could influence MLL1 function (Yokoyama *et al*., 2004; Popovic and Zeleznik-Le, 2005; Slany, 2005b). Interestingly, menin interacts with both wild type MLL1 and MLL-fusion proteins (Yokoyama *et al*., 2004) (see later), suggesting it interacts with the N-terminal region (MLL^N).

Purification of the MLL-ENL fusion/menin complex identified a 75kDa subunit called LEDGF (Yokoyama and Cleary, 2008). LEDGF is implicated in transcriptional coactivation (Ge *et al.*, 1998), and colocalises with MLL target genes, including *Hoxa9* (Yokoyama and Cleary, 2008), associating with transcriptionally active chromatin through its conserved PWWP domain (Turlure *et al.*, 2006). This motif is related to "royal family" domains, having similarities with Tudor domains. It is thought that menin is only required to recruit or stabilise LEDGF in the MLL1 complex, which thereby allows MLL1 to target transcriptionally active chromatin through (Yokoyama and Cleary, 2008).

To summarise, it is clear that the MLL1 complex requires cofactors and associated proteins to be recruited to chromatin on target genes and be fully active. Furthermore, the proper association of both MLL^C and MLL^N, which have transcriptionally antagonistic properties, is essential in the balanced function of MLL1. It is perhaps not surprising that the disruption of this balance would thus have consequences for downstream regulation. Indeed, the generation of MLL-fusion proteins leads to the disregulation of gene expression.

1.4.3. MLL-Fusion proteins

A leading reason why MLL1 is a focus of widespread research is the finding that chromosomal translocations are able to generate MLL fusion proteins that disrupt the proper regulation of genes, notably the *Hox* genes. These translocations typically result in the loss of the C-terminal region of the MLL protein, including the PHD domains (Fair *et al*., 2001), the transactivation domain and the SET domain, which are replaced by one of 54 translocation partners fused to MLL^N (Slany, 2009). Early papers suggest that the loss of PHD domains prevents the recruitment of corepressors (Fair *et al*., 2001), however, this may also impact on MLL-chromatin interactions. Importantly, it is thought that the novel properties characteristic of MLL-fusion proteins explain the resulting gene disregulation: cell transformation by MLL fusions necessitates the loss of the PHD domains (Chen *et al*., 2008), whose presence would reduce association with *Hoxa9* and impair *Hoxa9* up-regulation (Muntean *et al.*, 2008), but requires the CxxC domain to associate to non-methylated CpG DNA sites (Ayton *et al.*, 2004). Furthermore, menin remains associated with the N-terminal region of MLL-fusion proteins (Yokoyama *et al*., 2004), and still acts to recruit LEDGF to MLL-fusion proteins (Yokoyama and Cleary, 2008). Thus, the composition of the "remaining MLL" after translocation is critical for leukaemic transformation. Furthermore, MLL fusion partner proteins are essential for transformation; these are mostly transcription factors, like AF4 (Gu *et al.,* 1992), AF9 (Dobson *et al.,* 1999) and ENL (Tkachuk *et al.,* 1992), where the generation of "MLL-transcription factors fusions" is associated with a gain-of-function on gene expression (Lavau *et al*., 1997). MLL can also, but rarely, fuse with proteins normally found in the MLL complex, such as CBP, which is also sufficient for transformation (Yokoyama and Cleary, 2008). The fusion of the chimeric MLL1 with an acetyl-transferase,

and the permanent recruitment of this protein lead to the hyperacetylation of MLL target genes and thus aberrant up-regulation (Slany, 2009).

The process of MLL fusion generates a truncated MLL molecule, which loses a number of regulatory domains, but gains a range of activating factors, resulting in novel functions due to the constitutive activation of MLL. Thus it is proposed that MLL translocation partners act to change MLL into a permanent transcriptional activator (Zeisig *et al*., 2003), though recent analysis of MLL target genes suggest that the pattern of histone modification brought to MLL target genes by fusion proteins may lead to deregulation (Guenther *et al*., 2008; Krivtsov *et al*., 2008)

Other MLL proteins are present in metazoans: four other MLL's have been described in human, and three in mouse, but they remain poorly characterised. Members of the MLL family share a high degree of homology within the same species, but MLL proteins from human and mouse also share a high degree of homology (Figure 1.17). This is typically focused on the conserved domains (Figure 1.18).

1.4.4. MLL4 (KMT2D)

Human MLL4 (MLL2, WBP7, HRX2, TRX2) was initially named MLL2, and is the second human homologue to the *Drosophila* trithorax gene to be described. The gene is located on the region 19q13.1 in human (Fitzgerald and Diaz, 1999) and displays a high degree of homology with MLL1 and contains the same conserved domains observed in MLL1 (Huntsman, Chin *et al*., 1999). The observation that both proteins interact with similar factors, suggests there are overlapping functions between the two proteins (Huntsman, Chin *et al*., 1999). Human MLL4 is a 2,715 residue protein and like MLL1 contains three AT-hook

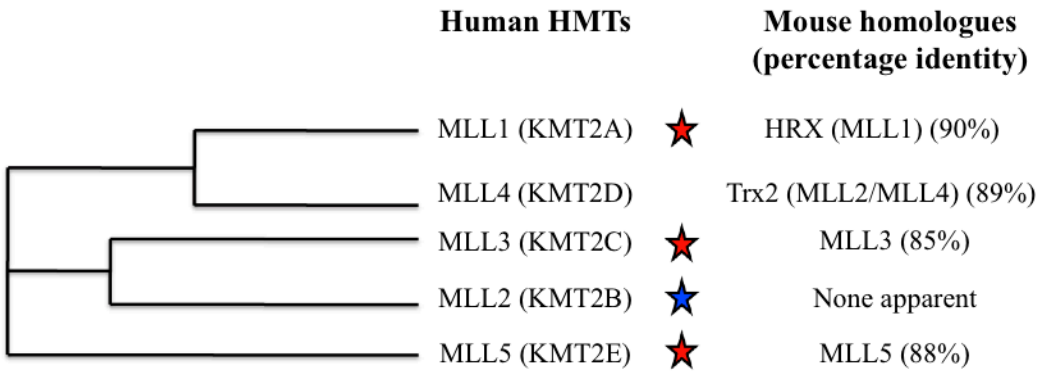


Figure 1.17: Human MLL HMTs

Cladogram showing sequence alignment of human MLL proteins (alignment software: *ClustalW2*). Proteins with links to leukaemia (★) and solid tumours (★) are indicated. The equivalent mouse MLL homologues and their overall sequence conservation are indicated (http://services.uniprot.org).

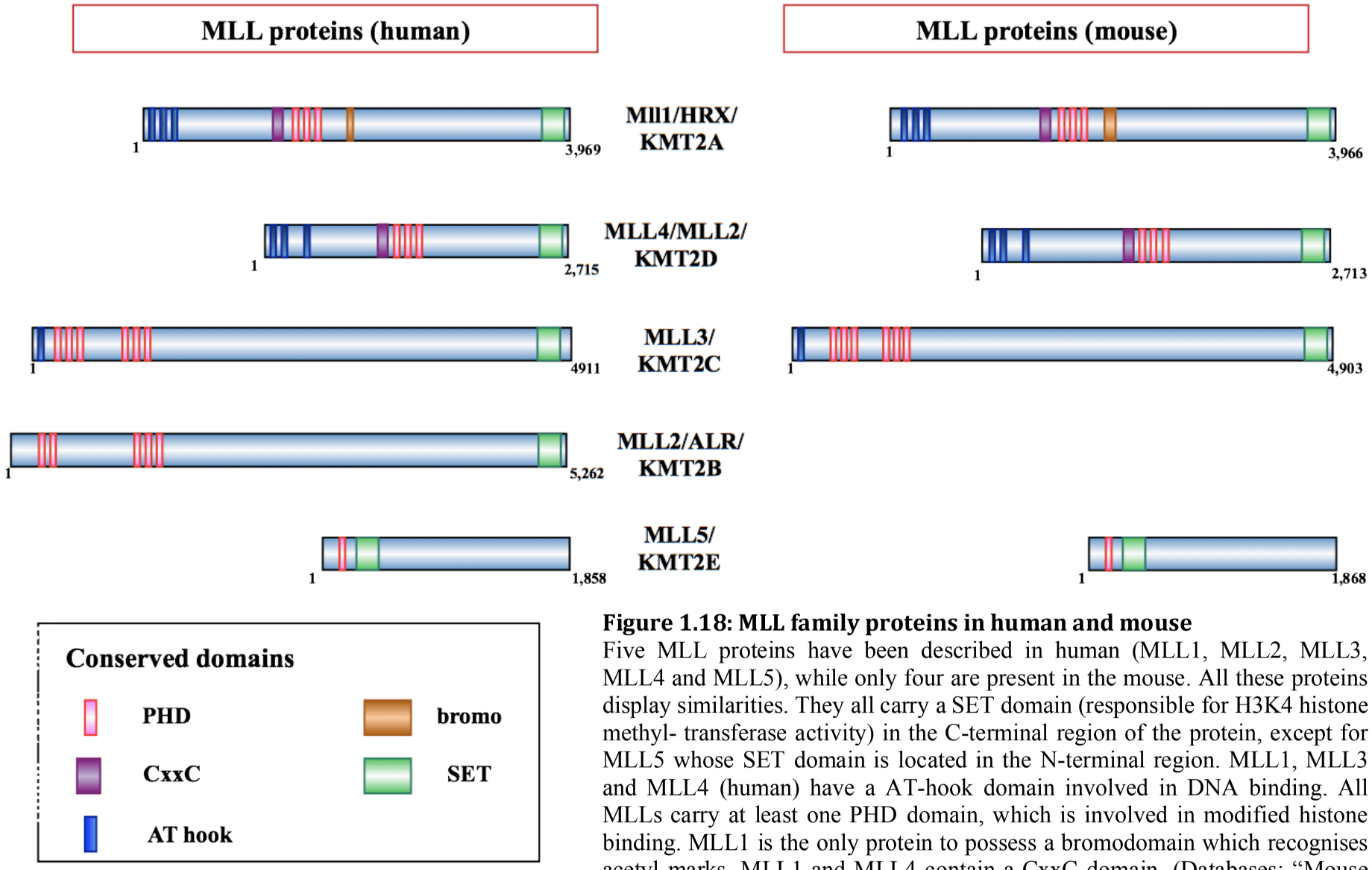


Figure 1.18: MLL family proteins in human and mouse
Five MLL proteins have been described in human (MLL1, MLL2, MLL3, MLL4 and MLL5), while only four are present in the mouse. All these proteins display similarities. They all carry a SET domain (responsible for H3K4 histone methyl- transferase activity) in the C-terminal region of the protein, except for MLL5 whose SET domain is located in the N-terminal region. MLL1, MLL3 and MLL4 (human) have a AT-hook domain involved in DNA binding. All MLLs carry at least one PHD domain, which is involved in modified histone binding. MLL1 is the only protein to possess a bromodomain which recognises acetyl marks. MLL1 and MLL4 contain a CxxC domain. (Databases: "Mouse Genome Informatics" and "Uniprot")

domains, one CxxC motif, three PHD domains and a SET domain (Huntsman *et al*., 1999). In addition, the mouse protein contains a HMG binding motif, a domain found in a number of DNA and chromatin binding proteins. MLL4 is part of a similar protein complex to MLL1 (including ASH2, RbBP5, and WDR5) (Hughes *et al*., 2004; Yokoyama, Wang *et al*., 2004), and also associates with menin (Hughes *et al*., 2004).

In mouse, *Mll4* knockdown leads to a reduction of cell growth and an increase of apoptosis in *Mll4*$^{-/-}$ ES cells, possibly because MLL4 contributes to *Bcl2* expression, an anti-apoptotic gene (Lubitz *et al*., 2007). Furthermore, differentiation is delayed in *Mll4*$^{-/-}$ cells (Lubitz *et al*, 2007). This suggests that *Mll4* and *Mll1* are involved in distinct mechanisms of development, and act to regulate different genes (Glaser *et al*., 2006).

1.4.5. Other MLL family members

In total, five MLL family members have been identified in human (four in mouse) such that, in addition to MLL1 and MLL4, the MLL family is composed of MLL2, MLL3 and MLL5 (MLL3 and MLL5 in mouse). All of these proteins share high sequence homologies, notably in shared specific motifs. In particular, all MLL family members display a SET domain located in their C-terminal region, except for MLL5, which has a N-terminal SET domain.

MLL2 (or ALR) is a 5,262 residue protein containing a SET domain, PHD fingers and a HMG domain. The *MLL2* gene maps to chromosome 12q12-13 (Prasad *et al.*, 1997). *MLL2* knockdown experiments in Hela cells have been used to identify its target genes, many of which are implicated in development, cell adhesion, cytoskeleton organisation and transcriptional regulation. This is consistent with findings that *MLL2* knockdown leads to a reduction of the migration capacity of HeLa cells (Issaeva *et al*., 2007).

The *MLL3* gene maps to chromosome band 7q36, a region whose deletion is commonly found in malignant myeloid disorders (Ruault *et al.*, 2002), and encodes for a predicted protein of 4,911 residues. As with MLL1, MLL3 is characterised by conserved protein domains, including; (1) the N-terminal domain carrying six PHD fingers, divided into two groups, (2) an HMG box, and (3) a C-terminal SET domain (Ruault *et al.*, 2002; Brun, Gasca *et al.*, 2006).

MLL5 remains the least well-characterized MLL family member. The *Mll5* gene was located to chromosome band 7q22 in a screen for candidate myeloid leukaemia tumor suppressor genes within this region in a mouse model (Emerling *et al.*, 2002). *Mll5* is highly conserved in mouse and human, and a small 1,858 residue protein, containing (1) a single PHD domain, and (2) a SET domain located in the N-terminal region of the protein, rather than at the C-terminal as in MLL1. In contrast, it does not contain AT hooks (which suggests that MLL5 might not bind DNA) nor a methyl-transferase homology domain, unlike other MLL family proteins. These domains are involved in the transforming ability of MLL fusion proteins (Slany *et al*., 1998; Deng *et al*., 2004), suggesting that MLL5 does not regulate transcription directly. It has been suggested that the *MLL5* gene could contribute to leukemogenesis by haplo-insufficiency, as a single functional copy of *MLL5* is not sufficient to maintain a wild-type condition (Emerling *et al.,* 2002). MLL5 has a role in haematopoiesis in mice, as mice deficient in *Mll5* have more megakaryocyte-erythrocyte progenitors than granulocyte-macrophage progenitors (Zhang *et al*., 2009). It has also been shown that a *MLL5* over-expression induces growth arrest in G1 phase (Deng *et al*., 2004).

Thus, MLL histone methyl-transferases regulate gene transcription through multiple and cooperative mechanisms. To date, research has focused on MLL1 because of its role in the

disregulation of genes in leukaemia, but also for its involvement in *Hox* gene regulation during development.

1.5. The contribution of *Hox* genes to differentiation

1.5.1. *Hox* genes

Cell differentiation and organ development need to be tightly regulated and coordinated, and are controlled by a group of master regulatory genes. These genes, described as homeobox (*Hox*) genes, are highly conserved through evolution and encode DNA-binding transcription factors. *Hox* genes were first identified following observations made in mutants of the fruit fly *Drosophila melanogaster* (Lewis, 1978). For example, mutation of one of the ANT-C genes, *Antennapedia* (*Antp*), which is responsible for the identity of the second thoracic segment, leads to gross changes in body morphology. Failure to express this *Hox* gene properly results in the development of antennae instead of legs in the second thoracic segment, whereas its inappropriate expression in the eye-antennal imaginal disc leads to the production of legs instead of antennae (Struhl, 1981; Gehring, 1987). Similarly, mutants of the *Hox* gene *Ultrabithorax* (*Ubx*) are associated with the development of wings instead of halteres in the third thoracic segment. These changes were described as "*homeotic tranformations*", and are associated with the conversion of a body segment to another one. In *Drosophila*, *Hox* genes are organised in two complexes of genes, associated with segments of the body axis: the first complex, called the Antennapedia complex (ANT-C), is constituted of five homeobox genes, *labial* (*lab*), *Proboscipedia* (*Pb*), *Deformed* (*Dfd*), *Sex comb reduced* (*Scr*) and *Antennapedia* (*Antp*). These genes are responsible for the establishment of the head and the first thoracic segment. The second complex, the Bithorax complex (BX-C) (Lewis, 1978), is composed of the genes *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-*

B (*Abd-B*), and are involved in the development of the second and third thoracic segments, and the abdominal segments.

In mammals, *Hox* genes are composed of 39 genes, which are organised into four clusters (A-D) (Figure 1.19) located on distinct chromosomes: on chromosomes 7, 17, 12 and 2 in human and on chromosomes 6, 11, 15 and 2 in the mouse, respectively. Each cluster is constituted of nine to eleven individual genes, corresponding to 13 paralogous groups: each cluster is composed of a specific combination of these genes, with none of them containing a complete set of paralogues. These clusters can be aligned with each other on the basis of sequence homology, as with the *Drosophila HOM-C* cluster.

One of the more interesting observations on these gene clusters is the relationship between the relative position of *Hox* genes on chromosomes and the spatial gradient of *Hox* gene expression along the body axis; *Hox* genes positioned at the 3' end of the locus are expressed earlier than *Hox* genes situated at the 5' end (Krumlauf, 1994). In addition, in limb development, expression of *Hox* genes at the 3' end of the locus was observed the proximal regions, while expression of *Hox* genes from the 5' end was observed in distal regions (Lappin, Grier *et al*., 2006). Therefore, *Hox* gene arrangement follows a co-linear pattern, where expression of these key regulator genes shows a spatio-temporal pattern along the anterior-posterior body axis in mammals and insects during development (Lewis, 1978; Krumlauf, 1994; Lappin *et al*., 2006). These genes express transcription factors, and contain a highly conserved domain, the homeobox, which determines their DNA binding activity.

1.5.2. Long-term regulation of *Hox* genes

Gene regulation is thought to be determined by a range of genetic and epigenetic mechanisms (Jaenisch and Bird, 2003; Jenuwein and Allis, 2001; Li, 2002). Histone modifications are

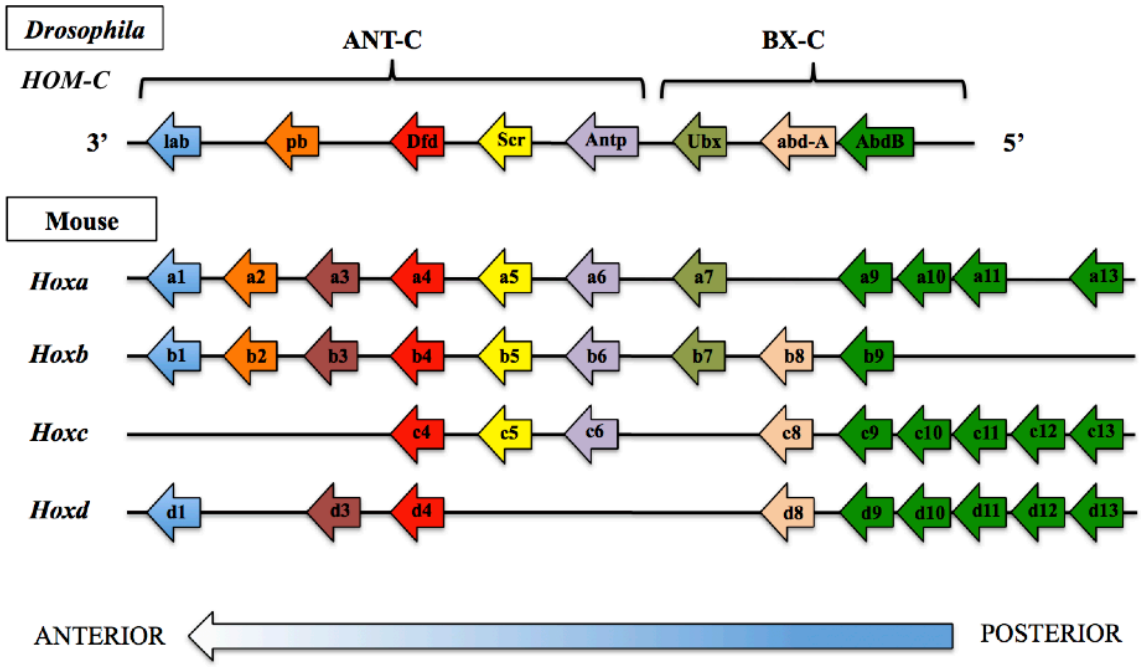


Figure 1.19: *Hox* gene clusters in *Drosophila melanogaster* and mouse (adapted from Carroll S. B., 1995)

In *Drosophila*, Hox genes are gouped in two clusters: the Antennapedia (ANT-C) and the Bithorax (BX-C) complexes, Corresponding *Hox* genes are organised in four clusters in mammals (*Hoxa*, *b*, *c* and *d*), constituted of 13 paralogous groups. *Hox* genes are expressed following a spatio-temporal gradient along the anterior-posterior body axis, which is mirrored by their position on the chromosome.

involved in the ongoing regulation of gene transcription (Jenuwein and Allis, 2001) as part of the mechanism of short term regulation. However, another level of regulation ensures the long-term maintenance of gene expression, and ultimately its transmission through to the next generation of cells (Probst *et al.* 2009). These mechanisms include DNA methylation (Bird, 2002; Jaenisch and Bird, 2003) and the action of Trithorax and Polycomb group proteins, which appear to be key to *Hox* cluster regulation (Orlando and Paro, 1995; Soshinikova and Duboule, 2008).

DNA methylation

The mammalian mechanism of long-term cellular memory of gene expression pattern can be determined by DNA methylation (Sen *et al*., 2010), which corresponds to the addition of a methyl group to CpG dinucleotides in DNA. Vertebrate genomes are characterised by the presence of GC-rich regions with a high density of CpG dinucleotides at the promoter sequences of many genes (about 60% in human genes), termed CpG islands. During development, some CpG islands showed increased levels of methylation ("*hypermethylation*"), which is associated with promoter silencing. This methylation is, for example, involved in the process of X chromosome inactivation that occurs in the early embryo (Courtier *et al*., 1995). X-inactivation in female mammals can be mediated partly through DNA methylation of X-chromosome (Li, 2002). DNA methylation is also required for spematogenesis and has been suggested to regulate chromosome architecture (Li, 2002; Trasler, 2009). CpG islands can also in certain cases become hypermethylated in specific tissues, in permanent cell lines and in abnormal cells such as transformed cells (Bird, 2002).

DNA methylation is maintained during replication by copying the parental-strand methylation in a semi-conservative mechanism (Li, 2002). This is performed by DNA

methyltransferase 1 (DNMT1) (Li, 2002). However, DNA methylation can be established *de novo* by two enzymes: DNMT3A and DNMT3B, in a mechanism which is linked to the deposition of H4R3 methylation, a mark involved in gene silencing (Zhao *et al*., 2009). This occurs mostly in early embryonic cells (Okano *et al*., 1999).

It has been suggested that DNA methylation is a secondary event, acting to stabilise transcriptional repression on genes that are already silenced by other means, rather than acting to silence active promoters. Nevertheless, transcription repression by DNA methylation depends on the location and density of methyl-CpGs relative to the promoter. Typically it is associated with long-term repression, but methylation can also be reversed (Barreto *et al*., 2007). It has been proposed that histone modification, as determined by gene transcriptional activity, influences the level of DNA methylation. Thus, methyl-CpG sites could either prevent the binding of factors required for gene activation, or favour the recruitment of transcriptional repressor complexes containing HDAC activities (Bird, 2002; Fuks, 2005). For example, it has been shown that the PcG (Polycomb Group) protein EZH2 leads to EZH2-target promoter methylation by recruiting DNA methyl-transferases (Vire *et al.*, 2006). This is driven by complexes containing DNA methyl-binding proteins, the archetypal protein being MeCP2, which binds methyl CpG sites (Nan *et al.*, 1993) and recruits histone deacetylases, such as Sin3/histone deacetylase co-repressor complex to targeted promoters, thereby mediating transcriptional silencing (Nan *et al.*, 1998; Ng and Bird, 1999).

Polycomb and Trithorax groups

The long-term maintenance of gene expression can also be regulated by a specific group of proteins, which act to maintain genes in a transcriptional state once the initiating signals are

lost. These regulators are highly evolutionarily conserved, and form the polycomb (PcG) and trithorax group (trxG) families of proteins. These regulators were first identified in *Drosophila*, and constitute multiprotein complexes, which bind chromatin to modulate its conformation (Ringrose and Paro, 2004). PcG proteins maintain transcriptional repression, whereas trxG proteins maintain transcriptional activation through generations of cells (Ringrose and Paro, 2004). In *Drosophila*, two PcG complexes have been identified, PRC1 and PRC2. PRC2 core complex is composed of the histone methyl-transferase E(Z) (Enhancer of zeste), SU(Z) (Suppressor of zeste) and ESC (Extra Sex Combs) (Muller *et al.*, 2002). PRC1 is primarly composed of four subunits; PC (Polycomb), PSC (Posterior Sex Combs), PH (Polyhomeotic), and dRing (an ubiquitin ligase) (Francis *et al.*, 2001). Methylation of H3K27 is catalysed by the enzymatic subunit E(Z) of PRC2 (Muller *et al*., 2002). This histone modification is then recognised by the PC chromodomain, which leads to the recruitment of PRC1 to chromatin (Cao *et al*., 2002). TrxG proteins have also been identified, like Brahma (BRM, corresponding to yeast SWI/SNF nucleosome-remodelling complex), and TAC1 (Trithorax Activating Complex 1). The BRM complex contains Brahma, Moira and OSA, whereas TAC1 contains TRX, CBP and an antiphosphatase Sbf1 (Ringrose and Paro, 2004). These proteins have human homologues, like MLL1, which play similar regulatory roles.

These two modes of transmission of gene expression patterns are essential in regulating *Hox* genes. They thus play a critical role in the correct regulation of genes and the generation of downstream mechanisms, such as cell differentiation.

1.5.3. The *Hox* genes: Developmental role and regulation

In addition to their role in the regulation of the anterior-posterior body axis of embryos, *Hox* genes are essential for the regulation of haematopoiesis. Haematopoietic stem cells (HSC) are multipotent cells, which have the capacity of self-renewal, but also to differentiate into different cell lineages (Iwasaki and Akashi, 2007). During haematopoiesis, stem cells can generate myeloid (granulocytes, monocytes, macrophages, erythrocytes, megakaryocytes and mast cells) or lymphoid (T, B and Natural killer cells) lineages, first giving rise to specific progenitor cells, which then develop into a variety of differentiated cell types (Iwasaki and Akashi, 2007) (Figure 1.20). Progenitor cells are grouped into distinct populations: common myeloid progenitors (CMPs), granulocyte-macrophage progenitors (GMPs), megakaryocyte-erythrocyte progenitors (MEPs), common lymphoid progenitors (CLPs), and pro-T and pro-B cells (Terskikh *et al*., 2003). The cell fate decisions of HSC's are dependent on a broad variety of factors, including the *Hox* genes. In mammals, *Hox* genes of the a, b and c clusters are expressed in subpopulations of haematopoietic cells (Sauvageau *et al*., 1994; Pineault *et al*., 2002). Loss of *Hox* genes (for example *Hoxa5* (Fuller *et al*., 1999) or *Hoxa9* (Izon *et al*., 1998)) leads to defects of haematopoiesis. Interestingly, *Hox* genes are co-linearly down-regulated upon differentiation, with *Hox* genes at the 3' end of clusters being preferentially down-regulated at early stages, whereas *Hox* genes located in the 5' region of clusters being down-regulated in later stages (Sauvageau *et al*., 1994).

In addition to playing a role in global differentiation, specific *Hox* genes are key regulators of cell type determination. Indeed, expression of distinct *Hoxa* genes is specifically associated with cell type specific expression and with the stage of cell maturation. For example, *HOXA5* expression is proposed to influence cell identity: reduction of *HOXA5* favours the generation or proliferation of erythroid progenitor cells (megakaryocytes and erythrocytes), whereas it is

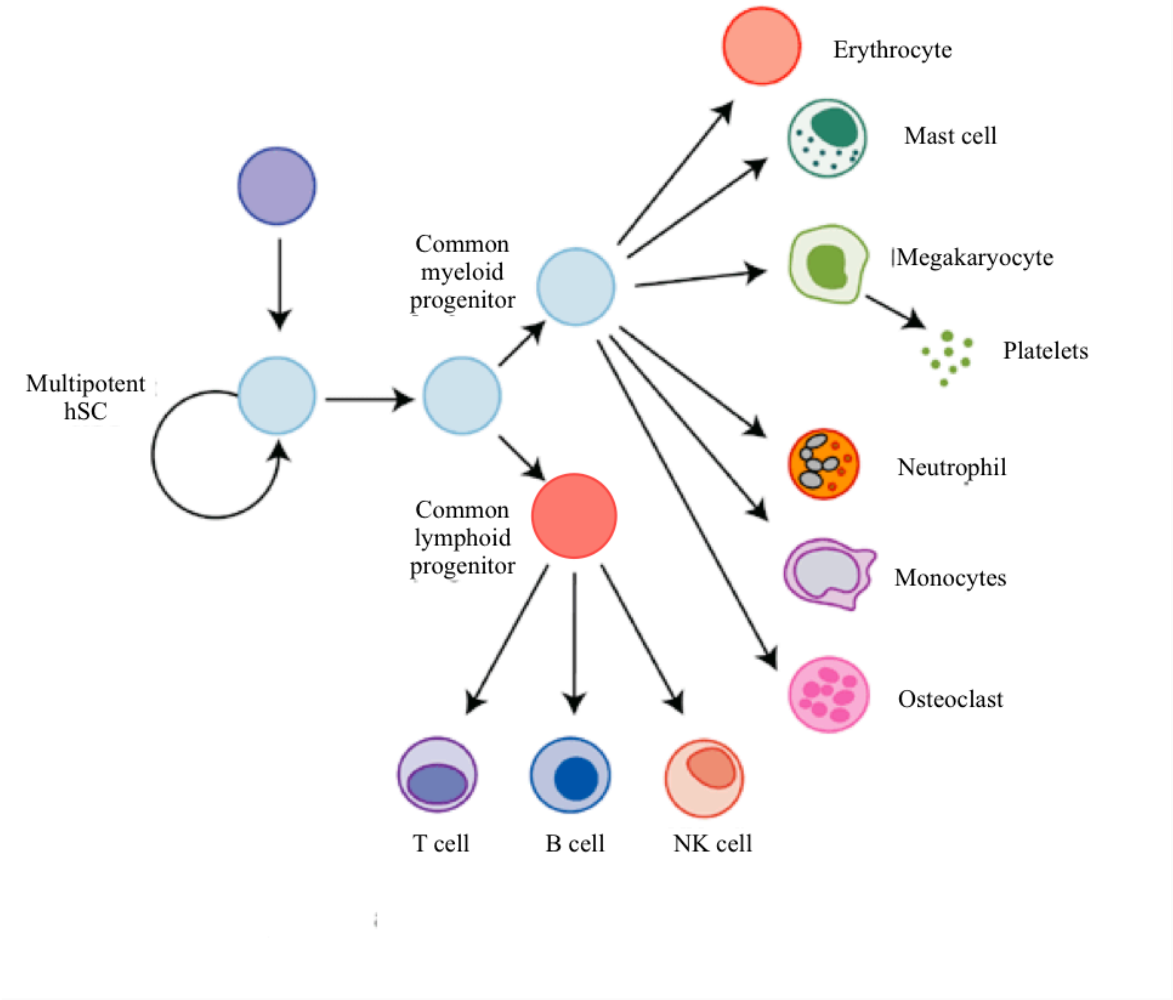


Figure 1.20: Haematopoiesis and the different cell lineages (from Qasim W., Expert Reviews in Molecular Medicine, 2004)

Bone marrow is composed of a range of cell types including haematopoietic stem cells (hSC) that can differentiate into different cell lineages. Multipotent hSCs can either self-renew or become common lymphoid or myeloid progenitor cells. The lymphoid lineage produces T cells, B cells, and NK cells (natural killer). The myeloid lineage produces erythrocytes, mast cells, megakaryocytes, neutrophils, monocytes and osteoclasts.

required for maturation and proliferation of myeloid progenitor cells (granulocytes and monocytes) (Crooks *et al*., 1999; Fuller *et al.,* 1999). Determination of cell fate is then maintained in the long-term by the DNA methylation of individual genes. In adult haematopoietic cells (but not in embryonic cells), *HOXA5* CpG islands are hypermethylated, correlating with a loss of gene expression (Strathdee *et al*., 2007b). This observation was also made with *Hoxa5* in mouse (Hershko *et al.,* 2003). *Hoxa4* has also been described as a gene required for cell type specificity. Hypermethylation of promoter CpG islands of *Hoxa4* and *Hoxa5* leads to repression of gene expression, and is associated with the progression of leukaemia (Strathdee *et al*., 2007a).

Hoxa cluster regulation: Role of transcription factors

Specific factors are involved in tissue- and developmental-specific gene regulation in order to determine cell-fate upon differentiation. Genes are thought to be mainly regulated at the transcriptional level, by the "basal" transcription machinery. The association of this complex with a specific gene requires the recruitment of a diversity of sequence-specific DNA-binding transcription factors, transcriptional co-activators such as ATP-dependant chromatin remodellers (Clapier and Cairns, 2009), and the basal transcription machinery, which includes RNA polymerases (Campos and Reinberg, 2009; Martinez, 2001). ATP-dependent chromatinremodelling complexes such as SWI/SNF also contribute to the accessibility of chromatin to other factors (Mellor, 2005).

This indicates that epigenetic factors, including Polycomb and Trithorax family members (as discussed), histone variants such as macroH2A (Buschbeck *et al*., 2009), and microRNAs contribute to the co-linear regulation observed on *Hox* gene clusters (Lemons and McGinnis, 2006). However, the transcription factors contributing to this regulation remain relatively

unclear. *Hox* regulation appears to be determined by both inter-genic and "global" enhancers, which respond to a range of developmental signals (Dubole, 1998). Retinoic acid response elements (RAREs) are found in many *Hox* genes, and exogenous retinoic acid addition is known to induce genes in the *Hoxb* cluster (Roelen *et al*., 2002). These are found to regulate *Hoxa*1 (Dupe *et al.,* 1997), as are estrogen response elements found to regulate *HOXA10* in some cell types (Akbas *et al*., 2004). Similarly YY1 (*Hoxa11*; Luke *et al*., 2006) and NF-κB (*HOXA9*; Trivedi *et al*., 2009) are known to mediate extracellular signals on multiple *Hox* genes, and are known to contribute to the regulation of individual *HOXA* genes. Interestingly, NF-kB has been found to interact with the H3S10p-specific histone kinase Msk1 (Reber *et al.*, 2009), suggesting a potential mechanism to recruit this modification to these MLL target genes.

Hoxa Regulation: Role of the MLL complex

Consistent with the central role of Trithorax family members on *Hox* gene regulation (Yu *et al*., 1995), MLL1 is also required for haematopoiesis (Hess *et al.*, 1997; Ernst *et al.*, 2002), for the proper regulation of body segment identity and the positive regulation of *Hox* gene expression (Guenther *et al*., 2005). MLL1 is required for the maintenance of gene activity on some *Hox* genes, like *Hoxa7, Hoxc8* (Yu *et al.*, 1998) and *Hoxa9, Hoxa10* (Ernst *et al.*, 2004b) but not for the initiation of these genes (Yu *et al.*, 1998). Furthermore, *Mll*$^{-/-}$ cells are associated with a reduction of expression of some *Hox* genes and defects in haematopoiesis, which can be "rescued" by the re-expression of *Hox* genes (Ernst *et al.*, 2004b). The MLL1-dependent regulation of *Hox* genes was also described for some *Hoxb* and *Hoxc* cluster genes (Ernst *et al.*, 2004b; Feinberg *et al.*, 2006). In addition, *Hox* gene expression is abolished in *Mll*$^{-/-}$ embryos and over-expressed in leukemogenic cells (Yu *et al.,* 1995; Milne *et al.*, 2002; Ayton and Cleary, 2003). This arises as in normal cells haematopoiesis is associated with the

down-regulation of *Hox* genes, whereas in leukaemia, MLL1 is transformed into a permanent transcriptional activator, associated with the up-regulation of MLL target gene expression (Figure 1.21). For example, MLL fusion proteins induce disregulation of the *HOXA9* gene by preventing its down regulation in differentiation.

Altogether, this shows the critical interrelation between MLL and *Hox* gene regulation during haematopoiesis. The disregulation of these processes can lead to the development of abnormal cells.

1.5.4. MLL fusion proteins and acute myeloid leukaemia (AML)

Metastatic cancers typically develop from normal cells via benign tumors. This process is typically associated with an accumulation of mutations in growth regulating genes leading to their aberrant expression. However, epigenetic mutations also play a central role in carcinogenesis: cancer cells are typically globally hypomethylated (Feinberg *et al*., 2006), but also show hypermethylation at tumor suppressor gene promoters associated with silencing and hypomethylation of oncogene promoters associated with inappropriate over-expression of gene (Feinberg *et al*., 2006; Kristensen *et al*., 2009). A number of type of leukaemia are typically associated with specific chromosomal rearrangements, e.g. Acute Lymphoblastic Leukaemia (ALL) and Acute Myeloid Leukaemia (AML) (Feinberg *et al*., 2006). Frequently these translocations involve the mixed lineage leukaemia (*MLL*) gene on chromosome band 11q23, and the corresponding *MLL* rearrangements are present in almost all types of haematologic malignancies (Slany, 2005a). It has been shown that the translocation event, involving the splitting of the *MLL* coding region, alters the nature of the *MLL* transcript, in which the DNA sequence is either partly lost or is ligated to new material (Ziemin-Van Der Poel *et al.*, 1991). Thus chromosomal rearrangements lead to the

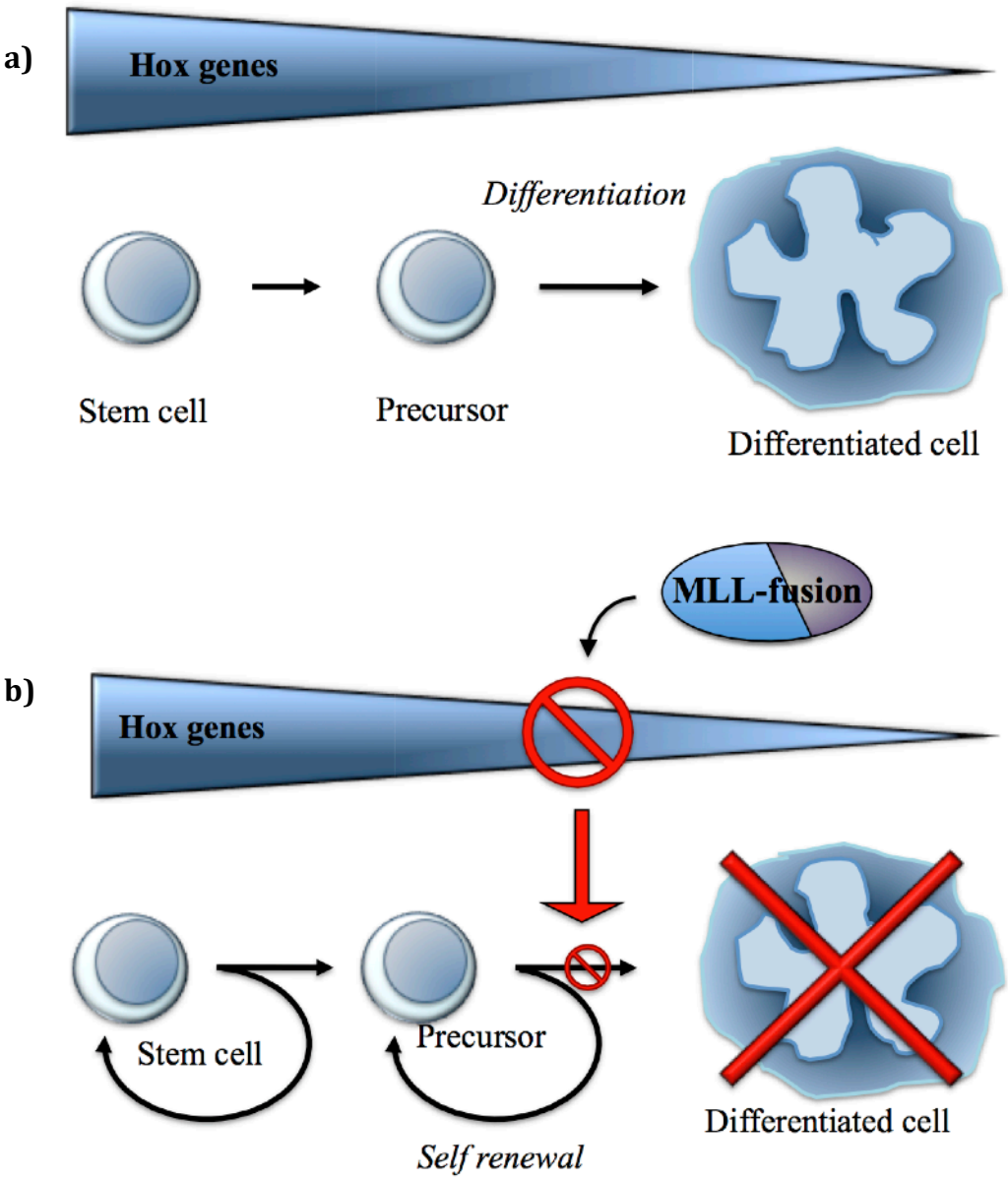


Figure 1.21: *Hox* gene expression in normal cells and leukaemic cells

In normal cells **(a)**, differentiation is associated with the down regulation of the *Hox* genes and *Meis1* expression, while in acute myeloid leukaemia (AML) **(b)** this down regulation is prevented by MLL fusion proteins. The cells do not differentiate and continue unregulated self-renewal.

generation of novel oncogenic proteins, called fusion proteins, containing the amino-terminus of MLL fused with one of the 54 partner proteins (Gu *et al.,* 1992; Tkachuk *et al*., 1992; Slany, 2009). Partners are diverse and include transcription factors like AF4 (Gu *et al*., 1992) and AF9 (Dobson et a*l*., 1999), or histone acetyl-transferases such as CBP (Yokoyama and Cleary, 2008). MLL can also associate with another N-terminal region of MLL, generating MLL dimers (Slany, 2009). MLL fusion proteins lack the SET domain, the CBP-binding domain and the PHD finger domains, and compromise the balance between transcriptional activation and repression. Several genes are upregulated by MLL fusion proteins, including *HOXA9* and the *Hox* co-factor *MEIS1*, found to be over-expressed in human leukaemias with MLL rearrangements (Zeisig *et al*., 2004; Feinberg, Ohlsson *et al*., 2006).

1.6. Aim of the study

The aim of this study was to examine and understand how the histone methyl-transferase MLL1 and its core complex regulate gene expression. In order to assess the dynamic processes involved, we have developed a differentiation model to examine MLL target gene regulation at genes that show differentiation-induced changes in transcriptional activity.

Our initial studies characterise "epigenetic" events on MLL target genes during differentiation. MLL1 targets a broad number of genes, but we chose to characterise histone modifications on the *Hoxa* cluster genes as these are well studied due to their key contribution to differentiation process. Then, we aimed to understand the underlying regulation of these events by characterising the binding of the proteins associated with depositing of these histone modifications, and thus contributing to the regulation of *Hoxa* genes activity upon differentiation.

Finally we aimed to assess the contribution of the MLL complex to these events by using siRNA technology to induce a MLL knockdown, to characterise the changes in histone modification in the absence of the MLL complex.

2. MATERIALS AND METHODS

2.1. Cell culture

2.1.1. HPC-7 cells

As a model, HPC-7, an immortalised Mouse Embryonic Stem Cell (ESC-like) line was used. This cell line was generated by expressing the LIM-homeobox gene Lhx2 in Haematopoietic Progenitor Cells (HPCs), derived from embryonic stem cells differentiated in vitro, and was established by the Carlsson Lab (*Karolinska Institute, Sweden*) (Pinto Do *et al*., 1998; Pinto Do *et al*., 2002; Richter *et al*., 2006). Lhx2 is a transcription factor involved in asymmetric cell division, tissue specification and differentiation of specific cell types. HPC-7 cells have the capacity to differentiate into several different cell types, notably into megakaryocytes, using thrombopoietin.

Cells were grown in "HPC-7 Growth Medium". Cells were maintained at $1x10^6$ per ml in growth medium with 1/10 CHO/SCF conditioned medium. Cells were split every two days. The CHO/SCF conditioned medium was obtained from cultivated CHO cells.

HPC-7 Growth Medium: *Invitrogen* StemPro-34 SFM medium, supplemented with StemPro-34 Nutrient Supplement (*Invitrogen*), 1/100 penicillin (10,000 U/ml)/ streptomycin (10,000 µl/ml), 1/100 L-glutamine (200 mM) (*Invitrogen*) and 1/10 CHO/SCF conditioned medium

2.1.2. CHO cells and CHO conditioned medium

CHO (Chinese Hamster Ovary) cells were kept in DMEM, 10% FBS (Foetal Bovine Serum), 1/100 penicillin/streptomycin (*Invitrogen*), and 1/100 L-glutamine (*Invitrogen*), until confluence. Cells were expanded three-fold as follows: CHO cells were washed with DMEM

and then resuspended with 5ml 0.05% Trypsin-EDTA (*Invitrogen*). After five minutes at 37°C, trypsination was stopped by adding 25ml of DMEM, 10% FBS. Cells were split into three flasks, and the same volume of DMEM, 10% FBS medium with supplements was added. CHO cells were maintained at 37°C until confluence, then left for two more days at 37°C. The cells were then pelleted and resuspended in StemPro-34, 0.5% FBS, 1/100 penicillin/streptomycin, 1/100 L-glutamine final concentrations, which was repeated the following day. The cells were then left for two days at 37°C before the medium was collected, filtered (0.45µm filter) and aliquoted. The medium collected is subsequently called "CHO/SCF conditioned medium". Aliquots were kept at -20°C.

2.1.3. HPC-7 cells differentiation to megakaryocytes

HPC-7 cells differentiate into megakaryocytes in the presence of thrombopoietin (TPO) over six days. The differentiation was achieved in 75 cm^3 flasks, with $20x10^6$ cells in 20ml HPC-7 Growth Medium seeded per flask. On the initial day, 1/50 CHO/SCF conditioned medium and 100ng/ml murine rTPO (recombinant ThromboPOietin) (*PrePro Tech*) were added. Cells were expanded two-fold on day two. On day three, the medium was changed to HPC-7 Growth Medium, 1/100 CHO/SCF conditioned medium and 100 ng/ml rTPO. On day five, the medium was changed to HPC-7 Growth Medium, 1/200 CHO/SCF conditioned medium and 100ng/ml rTPO. The cells were collected on day six, washed twice with cold Phosphate Buffer Saline (PBS), including the histone deacetylase inhibitor (HDACi) sodium butyrate (0.5mM), and the phosphatase inhibitor okadaic acid (100µM) where cells were used to examine histone marks. A first Bovine Serum Albumin (BSA) gradient (3% BSA overlaid with 1.5% BSA) was prepared to separate the megakaryocytes (bigger and heavier than undifferentiated HPC-7 cells) from undifferentiated cells. Again, the PBS/BSA gradient was supplemented with 0.5mM sodium butyrate and 100µM okadaic acid to prevent the loss of

histone marks. The separation in the BSA gradient was performed on ice for two hours. Then, the bottom 1ml was pelleted and washed with cold PBS, or cold PBS/butyrate/okadaic Acid before performing a second BSA gradient separation to ensure only megakaryocytes are retained. The cells were finally washed with cold PBS or cold PBS/butyrate/okadaic Acid and used or stored at -80°C.

2.1.4. SL2 cells

Carrier chromatin immunoprecipitation requires a cell type from a species different to HPC-7 as a carrier. *Drosophila* SL2 (Schneider line-2) cells were cultured in Schneider's medium supplemented with 8% foetal calf serum (FCS, *Gibco*) and antibiotics (1/100 penicillin/streptomycin, *Gibco*). *Drosophila* cells were kept incubated at 26°C. Cells were washed, pelleted and kept frozen at -80°C for subsequent C-ChIP experiments.

2.2. Protein analysis

2.2.1. Isolation of proteins from HPC-7 cells

Proteins were extracted from HPC-7 cells ($2x10^7$) using a whole cell extract lysis buffer. Cells were washed three times with ice-cold PBS and pelleted. 75µl Lysis buffer was added per ependorf tube for 30 minutes, with the tubes kept on ice. Cells were then spun down at 13,000 rpm for 10 minutes at 4°C. The protein concentration was quantified using Coomassie Plus ™ Protein Assay Reagent (*Thermo Scientific*) at 595 nm.

Whole Cell Lysis Buffer: 20mM Tris-HCl pH7.4, 100mM EDTA, 10mM NaCl, 1% Triton X-100, 1mM NaF, 1 mM ß-glycerophosphate, 1mM EGTA, 5mM sodium pyrophosphate, Protease cocktail inhibitors (*Roche*)

2.2.2. Isolation of histones from HPC-7 cells

Changes in histone modifications in undifferentiated HPC-7 cells and megakaryocytes were assessed using histone extraction. Cells from both cell types were pelleted and washed twice with cold PBS containing a final concentration of 5mM sodium α-butyrate and 100µM okadaic Acid.

Megakaryocytes were separated from undifferentiated HPC7 cells on two BSA gradients (1.5-3% BSA in cold PBS with 5mM butyrate and 100µM okadaic Acid). The usual method to isolate histones is to perform an acid extraction.

- Histone extraction with 0.4M HCl

HPC-7 cells were washed into cold PBS/5 mM butyrate/100 µM okadaic Acid and then pelleted. Then megakaryocytes pellets were resuspended in 1ml Triton Extraction Buffer (TEB) whilst pellets from undifferentiated HPC-7 cells were resuspended in 4ml of the same buffer and incubated for 10 minutes on ice. Samples were spun at 2,000 rpm for 10 minutes. The pellets were subsequently resuspended in half the initial volume of TEB and then spun again at 2,000 rpm for another 10 minutes. 500µl of 0.4M HCl was finally added per tubes, which were left for three hours on ice. After extraction, the tubes were centrifuged at 4°C and supernatants were retained.

TEB: 4mM Na-α-butyrate, 0.5% Triton X-100, 2mM PMSF, 0.02% Na-Azide in PBS

- Extraction with 5% perchloric acid (PCA) (Linker histone extraction)

Linker histones were extracted from cell pellets according to the protocol developed by Paulson *et al.* (Paulson, Ciesielski *et al.*, 1994). Cells were chilled on ice, pelleted, washed with cold PBS/5mM butyrate. 300µl of 5% Perchloric Acid was used to resuspend the

pellets, which were left 15 minutes on ice in order to extract linker histones. After an initial centrifugation for 10 minutes at 13,000 rpm (4°C), 100µl of 100% Trichloroacetic Acid (TCA) was added to the supernatant, to precipitate the proteins extracted. This was performed on ice for 35 minutes. The extracts were pelleted (13,000 rpm of centrifugation at 4°C for 10 minutes) and finally washed gently and briefly twice with acetone. The pellets were dried and resuspended into 50 µl of 1mM HCl.

Once the proteins have been quantified, half a volume of 2x Sodium Dodecyl Sulfate (SDS) loading buffer was added. The proteins were denatured by adding the tubes at 95°C for 5 minutes. The samples were loaded on a 15% SDS gel to visualize them.

2.2.3. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and gel transfer

Proteins were separated according to size following the method developed by Laemmli (Laemmli, 1970). The samples were loaded on an acrylamide gel whose concentration depends on the size of the proteins to be separated. A 15% SDS gel is prepared for separating histones (9-15 kDa) while a 7% SDS gel is prepared to examine MLL (The C-terminal domain of MLL1 is 180 kDa). The resolving gel was overlaid with a few drops of water-saturated isobutanol to prevent evaporation during polymerisation. Isobutanol was thoroughly washed away with distilled water before addition of the stacking gel. Protein samples were incubated at 95°C for 5 minutes and loaded onto the gel through SDS reservoir buffer. The gel was electrophoresed at 400V, 40mA and 20W, for three to five hours depending on the size of the proteins of interest, in a standard vertical electrophoresis unit (Hoefer "SE 600 Ruby", *Amersham*).

Proteins separated were then transferred (as described (Towbin *et al*., 1979) onto a Hybond C nitrocellulose membranes (*Amersham*) for three hours (180V/300mA/20W) in a *Bio-Rad* TransBlot Cell through transfer buffer. The efficiency of transfer was checked by soaking the membrane in Ponceau Red. The membrane was then incubated for one hour or overnight at 4˚C in a Blocking Buffer in order to prevent non-specific antibody binding. The membrane can also be stored in PBS/0.1% Tween.

Resolving gel for histones (15%): 15% Bis-acrylamide (30% Acrylamide-0.8% Bis), 250µM Tris-HCl pH8.8, 300µl 10%SDS, 100µl 10% APS (Ammonium Persulfate) and 30µl TEMED made up to 30ml with distilled water

Stacking gel: 1ml 30% Acrylamide-1.6% Bis, 250µl Tris-HCl pH6.8, 100 µl 10% SDS, 100µl 10% APS and 10µl TEMED made up to 10ml with distilled water

SDS reservoir buffer: 50mM Tris, 0.384M glycine, 0.1% SDS

Blocking buffer: 10% dried milk powder, PBS/0.1% Tween-20

Transfer buffer: 25mM Tris, 192mM glycine, 20% methanol

2.2.4. Western blotting

Antibodies were obtained from commercial and in-house sources. The membrane was then incubated with a first antibody (rabbit or mouse) diluted in PBS-0.1%Tween, with 5% milk: H3K4me1 (204) 1/400-fold dilution, H3K4me2 (R148) 1/400-fold dilution, H3K4me3 (612) 1/2,000-fold dilution, H3S10p (Ab12181, *AbCam*) 1/1,000-fold dilution, H4K8ac (R228) 1/1,000-fold dilution, H4K16ac (R252) 1/1,000-fold dilution, H3T3p (Ab17352, *AbCam*) 1/20-fold dilution, rabbit polyclonal ß-actin (Ab8227, *AbCam*), Msk1 (Ab32190, *AbCam*) 1/750-fold dilution, KAT3A/CBP (Ab3652, *AbCam*) 1/200 fold dilution, NFκB/p65 (Ab7970, *AbCam*) 1/1,000 fold dilution, menin (Ab2605, *AbCam*) 1/2,000-fold dilution for 1-2 hours or overnight at 4°C, and the mouse monoclonal MLL^{C}/HRX (clone 9-12, *Upstate*)

1/200-fold dilution, H3K4me3 (184) 1/400-fold dilution, H3K9me3 (Ab8898-100, *AbCam*) 1/500-fold dilution overnight at 4˚C. The membrane was washed twice with PBS/0.1% Tween and once with PBS/0.1%Tween/5% dried milk (15 minutes each). The secondary antibody (Anti-rabbit, *Li-cor* 1/5,000-fold dilution or goat anti-mouse HRP IgG peroxidase conjugate, A4416, *Sigma*, 1/2,000-fold dilution) was then added. The membrane was washed twice (10 minutes each) with PBS/0.1% Tween. Binding was either visualised by using the "Odyssey" (*Li-cor*) imaging system, or by ECL western Blotting detection reagents (*Amersham*) and then exposing the membrane to autoradiography film (*Amersham*) for 30 seconds to 30 minutes. Histone loading was typically normalised using an antibody raised against the C-terminal of histone H3 (*AbCam*), a domain with no known histone modifications.

2.2.5. Validation of antibody specificity

All histone-modification specific antibodies were assessed for specificity prior to use. In-house generated antibodies (i.e. H3K4me1, me2, and me3, H4K8ac) were assessed by competitive ELISA assays, in which the antibody of interest was bound to immobilised tissue-culture derived histones, and the interaction challenged by histone peptides containing or lacking the appropriate modifications. These assays were performed by first coating a microtitre plate with 50µl histones (10µg/mL) per well, which were left overnight to bind. These were removed, the wells washed with 2x PBS, and blocked with 1% BSA/PBS ("blocker", 4°C, 60 min.), prior to washing with 1x PBS. Peptide dilutions are prepared with PBS and 0.2mM 2mercaptoethanol (fresh). A serial dilution of the test antibody is generated with blocker buffer (in the range 10µg/ml to 0.31µg/ml), and 50µl/well incubated at 4°C (60 min.). In each well are deposited either the blocker, 100µl of antibody completed with 100µl of blocker or 100µl of appropriate peptide dilution completed with 100µl of antibody. The

well is subsequently washed twice with PBS/1M NaCl, PBS-Tween (0.1%) and PBS. The secondary antibody (1/1000) is then applied at 50µl/well, and incubated for 60 minutes at room temperature, prior to washing twice with PBS and four times with ddH2O, before applying the substrate (100µl/well) and left for 10 minutes. The reaction is terminated by addition of 5% H_2SO4 (100µl/well). The plate is then read.

In contrast commercial histone modification specific antibodies were screened using two approaches. We initially ascertained that the antibodies recognised the correct histone by Western blotting when probing tissue culture cell-derived histones (*K. Nightingale, personal communication*). A second screen used an in-house generated histone peptide microarray, in which 183 modified histone peptides are spotted on glass slides, including a range of peptides containing H3K27me, H3K9acS10p, and peptides containing multiple adjacent modifications (i.e H3K9ac, H3S10p etc.). This substate is probed by standard Western blotting conditions (with the exception that wash steps are higher stringency than required in Western blotting (1x PBS, 0.1% Tween, 1M NaCl). Primary antibody binding was detected by fluorescent-labelled anti-mouse IgG (800nm, Licor), using an "Oddyssey" scanner (*Licor*). Antibodies against non-histone proteins (*i.e. Menin, MLL-C, etc)* were assessed by Western blotting of HPC-7 - derived nuclear extracts, to ensure a single band of the correct molecular weight was detected.

2.2.6. Acid-Urea-Triton (AUT) polyacrylamide electrophoresis

In order to characterise the degree of modification of histones, proteins were separated on an AUT-PAGE according to their charge, as described (Bonner *et al*., 1980). Gel solutions were degassed before addition of 500µl Triton X-100, 150µl TEMED and 2ml 0.0004% riboflavin per 30ml gel solution and photopolymerised in front of a 2 X 15 W light source. The resolving gel was overlaid with a few drops of water-saturated isobutanol to prevent

evaporation during polymerisation. Isobutanol was washed off the resolving gel thoroughly with distilled water before addition of the stacking gel. It was degassed before addition of 100µl TEMED and 700µl 0.0004% riboflavin per 10ml gel solution. Photo-polymerisation was carried out as for the resolving gel.

Resolving gel: 12% acrylamide, 0.32% N,N-bis-acrylamide, 8M urea, 1M glacial acetic acid and 0.05M ammonia

Stacking gel: 4% acrylamide, 0.21% bis-acrylamide, 8M urea, 1M glacial acetic acid and 0.15M ammonia

Protein samples were prepared for AUT-PAGE as following:

1 volume of protein sample (20-50µg) was mixed with 2 volumes of AUT-loading buffer, vortexed and spun down at 13,000 rpm for 15 minutes, before loading through AUT reservoir buffer. Gels were run for 15 hours at 150V, 30mA, 10W at 10°C using reversed polarity.

AUT-loading buffer: 8M urea, 5% ß-mercaptoethanol, 1M glacial acetic acid, completed with a few drops of tracking dye pyroninY

AUT-reservoir buffer: 1M glycine, 0.1M glacial acetic acid

2.2.7. Coomassie blue and silver staining

Proteins were revealed by staining the gel with Coomassie brilliant blue, for half an hour. After destaining in several baths of 30% methanol/ 10% acetic acid, the gel was soaked in 50% methanol overnight to remove all of the remaining Coomassie blue stain. A more accurate and sensitive method to detect proteins is silver staining, achieved after overnight incubation with 50% methanol.

A first "Solution A" was prepared with 0.4g of silver nitrate in 2ml of water. 10.5ml of 0.36% sodium hydroxide (NaOH) was mixed with 800µl of 14.8M ammonium hydroxide (NH_4OH), which constitutes Solution B. Solution A was added to Solution A drop by drop with constant mixing. The silver nitrate must be totally dissolved. Then the solution was made up to 50ml with water. The gel was stained in this final solution for 15 minutes with gentle agitation, and then washed 5 minutes with water. The developing solution (2.5ml citric acid, 50µl 38% formaldehyde made up to 500 ml water) was added until bands were visible. After washing the gel with water, the reaction was stopped with 45% methanol/10% acetic acid solution. The gel was finally stored in 50% methanol.

2.2.8. Immuno-microscopy of histone distributions

The distribution of histone modifications in undifferentiated HPC-7 cells, HPC-7-derived megakaryocytes, and mouse bone-marrow-derived megakaryocytes (*a gift from Dr. S. Watson, University of Birmingham*) were examined by immunofluorescence microscopy using a standard protocol including a crosslinking/fixing step.

Briefly, cells are harvested at 1,000rpm (10 min, 4°C), and washed twice with PBS. $2x10^5$ cells/ml are resuspended in 0.1M KCl and incubated for 10 minutes at room temperature. Cells were mounted on glass slides (VWR 1.0-1.2mm thick, ethanol washed) with a Shandon Cytospin 4. $4x10^4$ cells are added per chamber and spun down at 1,500 rpm, 10 min. The location of cells weas marked and the slide immersed in KCM buffer (10 min, RT). Primary antibodies were made up in KCM/0.1% BSA and 50µl applied per slide, this was covered with Parafilm, and incubated (60 min) in a humid chamber at 4°C. Slides are washed twice with KCM for 10 minutes. 50µl FITC-conjugated secondary antibody (Goat anti-rabbit FITC x50 dilution, *Sigma*) is applied and incubated for one hour at 4°C in a humid chamber. Slides were washed twice with KCM (10 min.), fixed in 4% (v/v) formaldehyde/KCM (10 min.),

briefly washed with distilled water, and mounted in 7.25µl DAPI (1µg/ml, *Sigma*) in Vectorshield (Vector Labs) prior to analysis.

In-house generated antibodies have been examined for their specificity in the procedure by competition using histone tail peptides either containing or lacking the appropriate modification (*Dr. L.O'Neill, Uni. Birmingham, personal communication*).

2.3. Native and Carrier chromatin immunoprecipitation (Amended by Turner's group)

To examine the pattern of epigenetic marks on specific genes, in undifferentiated cells and megakaryocytes, two different chromatin immunoprecipitation protocols were used. Because of the limited amount of cells, the examination of histone marks in differentiated HPC-7 cells used the Carrier-Chromatin Immunoprecipitation (C-ChIP) protocol (Figure 2.1) (O'Neill *et al*., 2006), while the examination of histone marks in undifferentiated cells used the Native-Chromatin immunoprecipitation (N-ChIP) protocol.

2.3.1. Chromatin isolation from SL2/HPC-7 cells

For the carrier chromatin immunoprecipitation (C-ChIP) procedure, HPC-7 cells were mixed with *Drosophila* SL2 cells. Drosophila SL2 cells were pelleted and washed three times with ice-cold PBS, 5mM butyrate. Cells were resuspended to $5x10^7$ cells/ml and a 1ml aliquot was mixed with 10^5 megakaryocytes.

2.3.2. Chromatin preparation for C-ChIP

HPC-7 cells (8 flasks of 20ml, generating about 10^5 cells) were differentiated before use in the C-ChIP following the differentiation protocol (see paragraph 2.1.3 above). The cell pellet was resuspended into 400µl of NB buffer and 100µl was added to each aliquot (4) of frozen

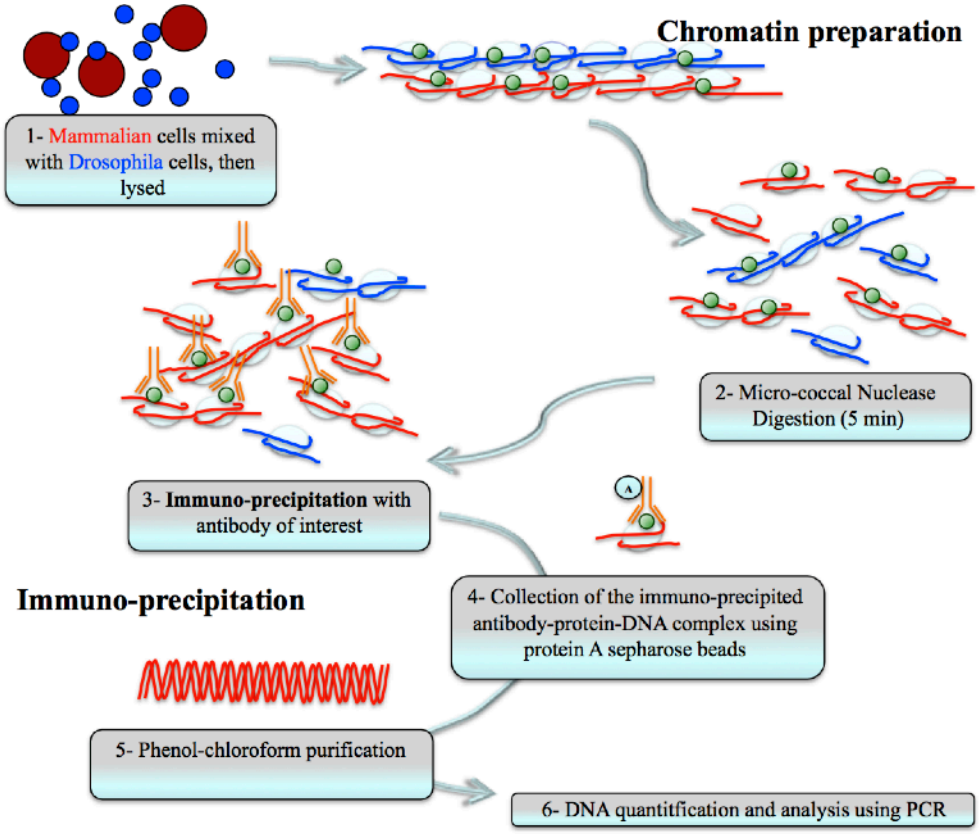


Figure 2.1: Carrier Chromatin Immunoprecipitation

Cartoon of step in C-ChIP protocol.

SL2 cells ($5x10^7$). Then each fraction (about $2.5x10^4$ cells) was made up to 1ml with NB buffer. The cells (HPC-7 / SL2 / NB buffer) were then transferred into 7ml bijou tubes containing 1mL of NB buffer/1%, tween 40 and kept on ice. 10µl of PMSF was added prior to stirring the samples on ice for 15 minutes. Each sample was kept separated and put into a different ice-cold Dounce homogeniser. Five strokes were given with a tight pestle ("A"), twice. In should be noted that chromatin release could not be checked by microscopy at this stage, due to the high amount of debris generated by frozen SL2 cells. However, the yield of chromatin obtained was subsequently checked and optimised by carefully following this protocol. Then, the samples were transferred to 50ml Falcon tubes and spun down for 15 minutes, at 800g, 4˚C. The supernatant was removed and the pellets were dislodged before being resuspended in 10mL of 5% sucrose/ NB buffer. At this step, the nuclei pellet must be white and a thin layer. The pellets were spun down at 3,000 rpm, for 10 minutes at 4˚C, dislodged and then resuspended into 2 ml of digestion buffer. The quantity of chromatin was checked by spectrophotometer (260nm, 280nm and 260/280 ratio) (Ultrospec 2100 pro UV/Visible) by adding 20µl of each sample to 300µl of 0.1% SDS. The chromatin pellets were spun down at 2,000 rpm for 10 minutes at 4˚C, and resuspended in digestion buffer in order to obtain a final concentration of 500µg chromatin per ml. 1ml of each sample was digested with 50U (10U/µl) of Micrococcal nuclease (Amersham Biosciences), at 28˚C for 5 minutes. The reaction was then stopped on ice by adding 0.1M EDTA. The samples were spun down at 13,000 rpm for 10 minutes at 4˚C. The supernatants were collected (S1 fraction) and kept in the fridge. The pellets were resuspended into 500µl of lysis buffer. Samples were put into dialysis membrane (Dialysis Tubing Cellulose Membrane, Sigma-Aldrich), and dialysed into Lysis buffer overnight. Each sample were transferred into individual ependorf tubes and spun down at 13,000 rpm, for 10 minutes at 4˚C. Supernatants

were kept (S2 fractions). The pellets were dislodged and resuspended into 200µl of lysis buffer. The quality of the chromatin was checked again with the spectrometer as before and by running a 1.2% agarose gel with 2µg of each sample. The S1 and S2 fractions were then pooled.

NB Buffer: 15mM Tris-HCl pH 7.4, 60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 0.5mM ß-mercaptoethanol, 0.1 mM PMSF, 5mM Na-Butyrate

Digestion Buffer: 0.32M Sucrose, 50mM Tris-HCl pH 7.4, 4mM $MgCl_2$, 1mM $CaCl_2$, 0.1mM PMSF, 5mM Na-Butyrate

Lysis Buffer: 1mM Tris-HCl pH7.4, 0.2mM EDTA, 5mM Na-Butyrate, 0.2mM PMSF, protease cocktail inhibitors (*Roche*)

2.3.3. Chromatin preparation for N-ChIP

The N-ChIP protocol was used on undifferentiated HPC-7 cells. This protocol is similar to the C-ChIP protocol.

$1x10^8$ cells were harvested and washed three times with ice-cold PBS, 5mM Na-Butyrate. The cell pellet was then resuspended into 5ml ice-cold TBS with 0.02mM PMSF. 5 ml of 1% ice-cold Tween 40 / TBS, protease inhibitors (*Roche*) and PMSF, was added. The samples were stirred on ice for 25 minutes. The lysate was transferred into an ice-cold Dounce homogeniser. Homogenisation was performed by giving 5 strokes with the tight "A" Pestle. Nuclei release was checked under microscope before carrying on with the preparation. Homogenates were pooled and centrifuged at 2,000 rpm for 20 minutes at 4˚C. A white pellet was obtained, corresponding to the nuclei. Pellets were first dislodged, and then washed in 10 ml 25% sucrose / TBS and centrifuged at 3,000 rpm for 10 minutes at 4˚C. Pellets were then resuspended in 5ml of Digestion Buffer (see before: C-ChIP), quantified by optical density

(260 nm) and then centrifuged at 2,000 rpm for 10 minutes at 4˚C. Pellets were dislodged and resuspended in Digestion Buffer in order to obtain a 0.5mg/ml final chromatin concentration. As with the C-ChIP protocol, the quality of the chromatin was checked with the spectrometer as before and by running a 1.2% agarose gel with 2 µg of each sample (Figure 2.2). From this point, the protocol follows the C-ChIP protocol. After DNA precipitation, the pellets were resuspended in 250µl of water.

TBS Buffer: 15mM NaCl, 10mM Tris-HCl pH7.5, 3mM $CaCl_2$, 2mM $MgCl_2$, 5mM Na-butyrate

2.3.4. Chromatin immunoprecipitation and DNA extraction

Immunoprecipitations were performed with about 75µg of chromatin in a volume of up to 1.3ml of antibody and incubation buffer or (15µl) Pre-Immune serum and incubation buffer. (Antibodies used: 100µl of H3K4me3 (Lab 612), H3K9me2 (616), H3K9ac (607), H4K16ac (252), or 15µl of H3K27me3 (07-449, *Upstate*), H3K9acS10p (Ab12181, *AbCam*)). The immunoprecipitation was performed at 4˚C overnight on a rotating wheel. 200µl of protein A Sepharose™ CL-4B (*GE Healthcare*) was added to each ependorf containing samples. Then the tubes were put on a high speed rotating wheel for three hours at room temperature. From this step, only siliconised tubes and pipettes were used. 15ml siliconised tubes were prepared with 9ml of buffer B. After incubation with the beads, the samples were spun down at 13,000 rpm for 10 minutes. The supernatants were kept (and labelled "Unbound Fraction") and transferred to siliconised tubes for protein extraction. The beads were resuspended into 1mL of buffer B and transferred into siliconised tubes with the 9ml of buffer B. The beads were spun down at 2,000 rpm for 10 minutes at 4˚C, then washed with 10ml of buffer C, then with 10mL of buffer C*. The beads were finally resuspended into 1ml of buffer C* and transferred

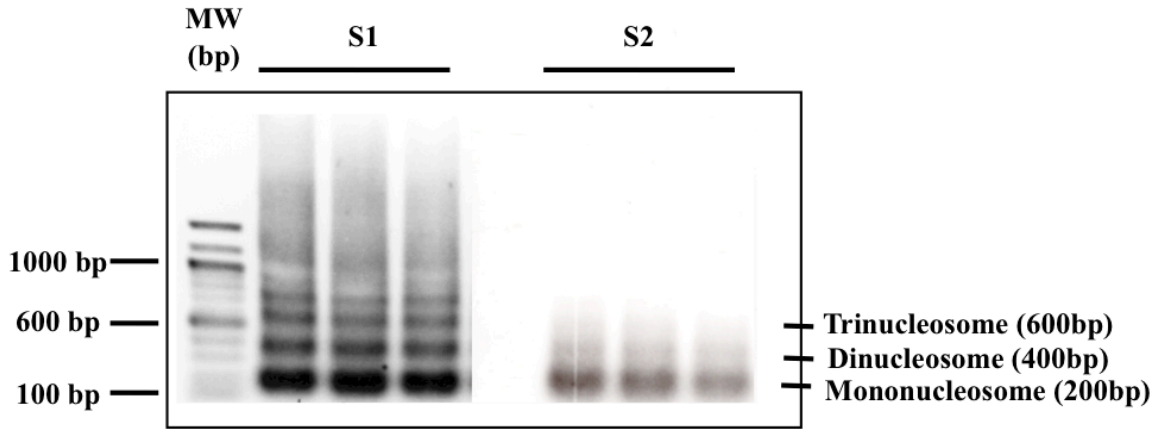


Figure 2.2: Chromatin analysis prior to N-ChIP

The chromatin extracted was treated with miccrococcal nuclease in order to obtain fragments of an appropriate size (mono, di, tri nucleosomes, etc) before performing the immunoprecipitation. S1 is the supernatant generated after centrifugation of digested chromatin (see Materials and methods), whereas S2 is the supernatant of pelleted chromatin after over-night dialysis. The quality of the chromatin was examined on a 1.2% agarose gel and visualised with ethidium bromide. S1 (lane 1-3) and S2 (lane 5-7)

MW= a molecular weigh marker

into a 1.5ml siliconised ependorf tubes and spun down at 13,000 rpm for 10 minutes. The beads were resuspended into 250µl of incubation buffer / SDS 1% left for 15 minutes on a high speed rotating wheel at room temperature, before being spun down at 13,000 rpm for 10 minutes. Supernatants were kept in siliconised tubes while the chromatin was again eluted from the beads in the same way. Supernatants from these two elutions were then pooled. 500µl of incubation buffer was added to each pool and transferred into 5ml siliconised tubes (labelled "Unbound fractions"). A Phenol-Chloroform extraction was then performed on the Bound and Unbound fractions to extract the DNA. DNA was precipitated with 5µl glycogen (*Roche*), 80µl of 5M LiCl and cold ethanol up to 4ml in a polypropylene tube. Proteins were precipitated with 4µl of 10 M H_2SO_4, 20µl of 1mg/mL BSA and cold acetone up to 4ml. Precipitations were performed overnight at -20˚C. DNA from Bound and Unbound fractions was spun down at 3,000 rpm 20 minutes at 4˚C, and the supernatants discarded. The pellets were dried and resuspended into 120µl of water. DNA was quantified using a Nanodrop® ND-1000 spectrometer.

Incubation Buffer: 75mM NaCl, 20mM Tris-HCl pH7.5, 20mM Na-Butyrate, 5mM Na_2EDTA, 0.1mM PMSF fresh

Buffer B: 50mM Tris-HCl pH7.5, 10mM EDTA, 100mM NaCl, 5mM Butyrate

Buffer C: 50mM Tris-HCl pH7.5, 10mM EDTA, 150mM NaCl, 5mM Butyrate

Buffer C*: 50mM Tris-HCl pH7.5, 10mM EDTA, 175mM NaCl, 5mM Butyrate

PCR analysis of immunoprecipitated material was performed with 10ng of DNA, on 384-well plates, on the ABI-7900 PCR machine. Only bound fractions were used for analysis. In order to normalise the data, Bound fractions obtained after immunoprecipitation of the chromatin with the pre-immune IgG, were used as an input. A new input was made for every ChIP.

2.4. Formaldehyde Cross-linking Chromatin Immunoprecipitation

In order to detect the distribution of MLL complex components on loci in the *Hoxa* gene cluster and other target genes, X-ChIP (Cross-linked Chromatin Immuno-Precipitation) was performed. This protocol has been adapted from an AbCam protocol, initially used in the Kouzaride lab (Cambridge University) (Figure 2.3).

2.4.1. Chromatin preparation

$5x10^6$ undifferentiated cells were pelleted and resuspended in a final concentration of 0.75% formaldehyde was immediately added for 10 minutes at room temperature and gently mixed. Cross-linking was stopped by the addition of glycine (1 M) to a 0.125 M final concentration. The cells were then washed twice with ice-cold PBS, and spun down at 1,200 rpm, 5 min at 4˚C. A final wash was performed with 1ml PBS per 10^7 cells. The cells were split into siliconised ependorf tubes, with 10^7 cells per tube. They were then resuspended in 1ml of lysis buffer, and divided into three siliconised ependorf tubes. The samples were sonicated at medium power for 8 min with 30 seconds On-Off (*Bioruptor, Diagenode*), and then quickly centrifuged at 13,000 rpm. The supernatant fractions were pooled, and 50µl samples were kept (labelled "Input"). 400µl of lysis buffer and 5 µl of proteinase K (20 mg/ml) were added to the input. The crosslinking was reversed by incubation at 65˚C overnight, while the remaining chromatin was kept at 4˚C. DNA from the input was then extracted by phenol-chloroform and quantified. The sonicated chromatin was checked by loading 10 µl of the input on a 1.2 % agarose gel (Figure 2.4).

Lysis Buffer: 1% SDS, 10mM EDTA, 50mM Tris-HCl pH8, protease cocktail inhibitors (*Roche*

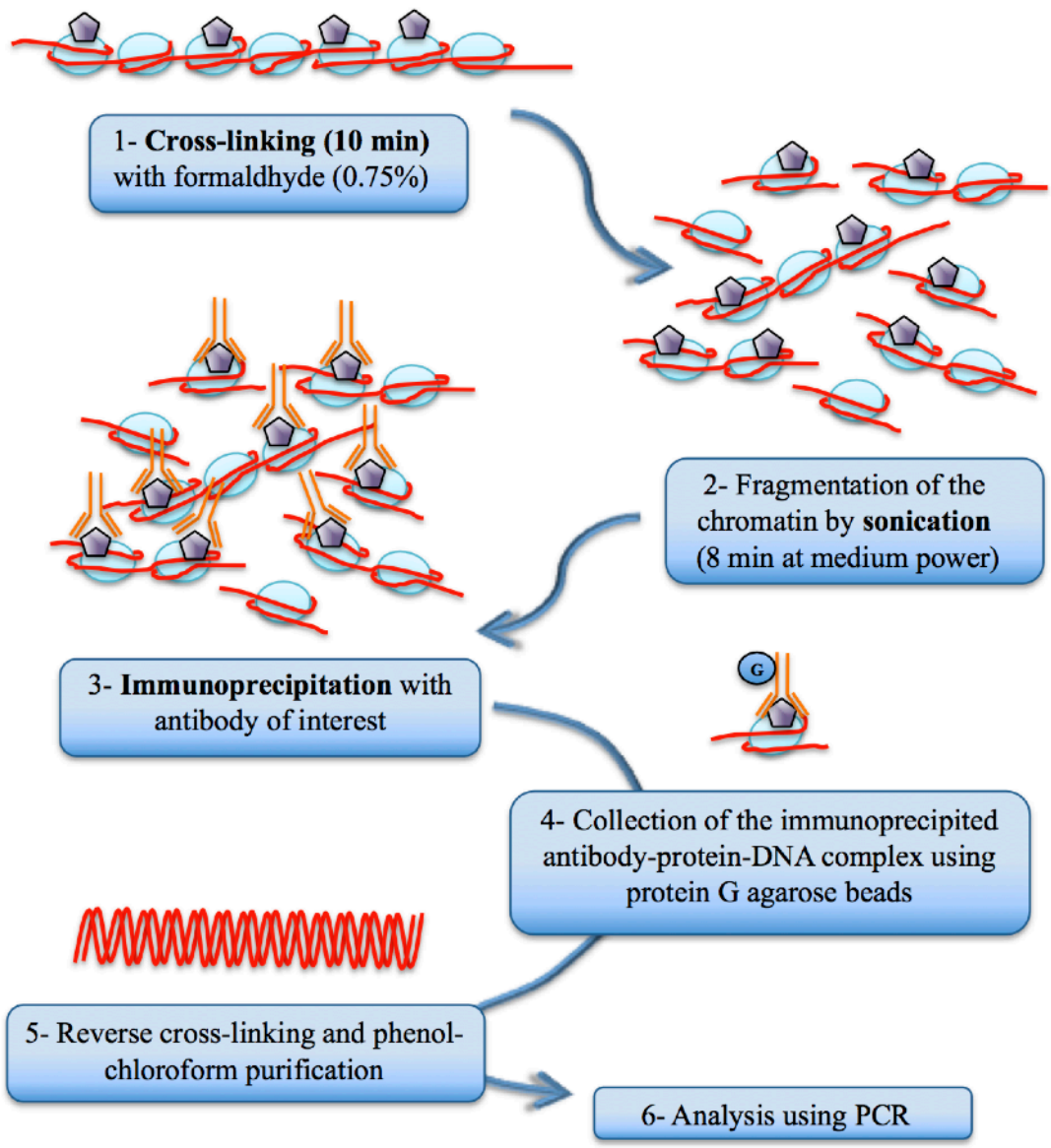


Figure 2.3: Cross-Linking Chromatin Immunoprecipitation

Cartoon of steps in X-ChIP protocol.

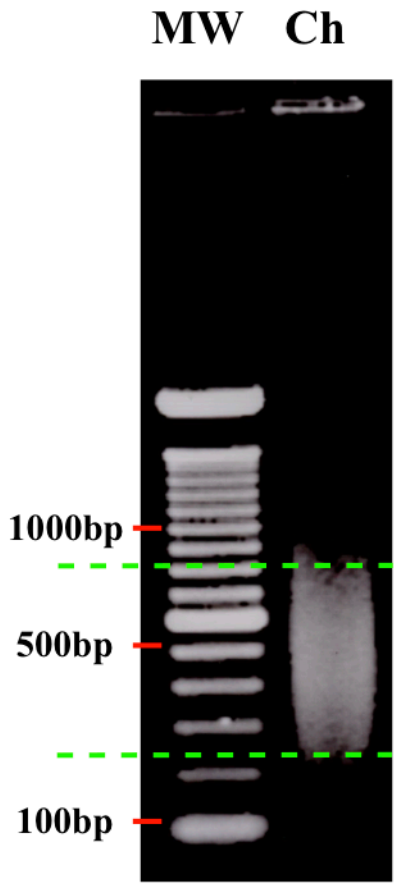


Figure 2.4: Assessment of cross-linked chromatin

The chromatin extracted was sonicated in order to obtain fragments of the appropriate size (between 200 and 700 bp) before performing the immunoprecipitation. The quality of the chromatin was examined on a 1.2% agarose gel, and visualised by ethidium bromide. MW= molecular weigh markers, Ch= chromatin.

2.4.2. Chromatin Immunoprecipitation

The chromatin was prepared with the dilution buffer up to 1 ml into a siliconised ependorf tube. Each fraction contains 25µg chromatin. Two controls were prepared: a fraction with no antibody and another one containing 10µl of Pre-Immune serum. Antibodies were added to the fractions: H3 C-terminal (Ab1791, AbCam) 1/100 as a positive control, the mouse monoclonal MLL^C/HRX (clone 9-12, Upstate) 1/50, menin (Ab2605, *AbCam*) 1/50, MLL-N (L. Ringrose Group, University of Vienna, Austria), RNA Pol2 (Ab24758, *AbCam*), Msk1 [PhosphoS376] (Ab32190, *AbCam*), KAT3A/CBP[AC238] (Ab3652, *AbCam*), MLL4 (Ab39267, *AbCam*). The fractions were immunoprecipitated at 4˚C on rotating wheel overnight. 40µl of protein A agarose beads supplemented with salmon sperm DNA (1 mg/5 ml) (Salmon Sperm DNA/Protein A Agarose, *Millipore/Upstate*) was added and the fractions were incubated two hours at room temperature on a wheel with fast rotation. After centrifugation at 8,000 rpm, for one minute at room temperature, the beads were washed three times with a low salt buffer. The final wash was done with a high salt buffer. The DNA was eluted by adding 450µl of elution buffer (1% SDS, 100mM $NaHCO_3$) incubated 15 minutes at room temperature with rotation. After 1 minute of centrifugation at 8,000 rpm, the supernatant was removed with siliconised tips and kept into a siliconised ependorf tube. 5µl of proteinase K (20 mg/ml) was added to deproteinise the sample whilst reversing the cross-links at 65˚C overnight. The DNA was extracted by phenol-chloroform extraction, and resuspended into 50µl of water.

Dilution buffer: 2% Triton X-100, 0.1% SDS, 2mM EDTA, 150mM NaCl, 20mM Tris-HCl pH8, protease cocktail inhibitors (*Roche*)

Low Salt buffer: 0.1% SDS, 1% Triton X-100, 2mM EDTA, 150mM NaCl, 20mM Tris-HCl pH8

High Salt Buffer: 0.1% SDS, 1% Triton X-100, 2mM EDTA, 500mM NaCl, 20mM Tris-HCl pH8

2.5. Gene expression analysis

2.5.1. RNA extraction (Stratagene Kit or Qiagen Kit)

Total RNA was isolated from HPC-7 cells using either QIAGEN (QIAshredder™, RNase Free DNase Set and RNeasy® Minikit) or Stratagene (Absolutely RNA® Miniprep) Kits according to the manufacturer's instructions. RNA was quantified using an UV/visible spectrophotometer (260 nm, 280 nm) (Ultraspec 2100 pro).

Subsequently, the RNA could either be used directly to perform a one step RT-PCR (for Reverse-Transcriptase Polymerase Chain Reaction), or be reverse transcribed in order to use the cDNA obtained in a semi-quantitative hot PCR. The aim of this reaction was to amplify the cDNA using primers specific to genes of interest. The first experiments were done with the latter method, before I switched to RTQ-PCR (Real-Time Quantitative PCR).

2.5.2. Reverse transcription and semi-quantitative PCR

SuperScript™ III RNase H - Reverse Transcriptase kit (*Invitrogen*) was used according to manufacturer's instructions. 1µg of RNA was denatured at 65˚C for 5 minutes along with 2pmol of each primer (either primer specific cDNA synthesis, or 1µl of a random Oligo$(dT)_{20}$ primer (50µM) Invitrogen), 1µl of dNTP mix (100mM) (*Invitrogen*) and made up to 13µl with RNase free water. The mix was then kept on ice for around 1 minute to stop the reaction

and avoid non-specific annealing. The mix was then completed with 4µl of 5X First Strand Buffer, 1µl of 0.1M dithiothreitol (DTT), 1µl of Protector RNase Inhibitor (40 U/µl) (*Roche*) and 1µl of SuperScript™ III Reverse Transcriptase (200 U/µl) (*Invitrogen*). The synthesis was performed by incubation at 55˚C for 1 hour, and the reaction stopped by heating the mix at 75˚C for 15 minutes. cDNA was stored at -20°C.

To quantify the expression of *Mll* genes, I first performed several semi-quantitative hot PCR by adding [α-^{32}P]-dATP in the reaction, which allowed the sensitive detection of the products. Each reaction contained 45 µl of ReddyMix PCR MasterMix (1,1x) (*ABgene*), 1µl of each primer (10µM), 2µl of DNase-free water and 1µl of cDNA template. 0.1µl of [α-^{32}P]-dATP (*Amersham*) was added to each reaction. The number of cycles required differs according to the primers. The programme was set as follows: a first step in which the samples were heated 2 minutes at 95° permitting to activate the Taq-polymerase. This first initialization step was also essential to ensure that most of the cDNA template and the primers were denatured. The amplification was then performed in three steps: (1) denaturation of the double-strand cDNA at 95°C for 30 seconds, (2) annealing at 54°C for 30 seconds (the temperature of annealing depends of the set of primers), (3) elongation and synthesis of the new strand at 72°C for 30 seconds (this corresponds to the optimum temperature for the polymerase activity). The time of extension depended on the length of the cDNA fragment. These three steps were performed for about 30 cycles depending on the primers. A final elongation step at 72°C for 10 minutes was performed to ensure that any remaining single-stranded cDNA was extended. The samples were stored at 4°C. The samples were then loaded on a 6% polyacrylamide gel. After drying the gel, it was exposed on a Phospho-Imager cassette from 1 hour to overnight, and subsequently scanned with the

Typhoon™ 9200 scanner (*Amersham*). The products obtained were quantified by using the software ImageQuant™ v2003.01 (*Amersham biosciences*).

2.5.3. Real-Time Quantitative PCR (or RTQ-PCR) with Reverse Transcription

In order to obtain more accurate results, the next cDNA amplification used RTQ-PCR. This allowed simultaneous RNA reverse transcription and cDNA amplification in one step, and was performed according to manufacturer's instructions. The reaction was achieved with commercial primers in the case of the genes *Hoxa9*, *Meis1* and *ß-actin*. The mix per tube contained the following: 12.5µl of QuantiTech™ Sybr® Green Mix (*Qiagen*), 0.25µl of Rtmix (*Qiagen*), 1.25µl of each primers for the *Mll* genes, or 2.5µl of the commercial primers, diluted to 24µl with DNase free water. Another kit, Brilliant®II SYBR® Green QRT-PCR Master Mix 1-Step kit from *Stratagene* was also used, according to the manufacturer's instructions. Three analyses were performed for each sample. A control was performed without Rtmix, also in triplicate, in order to ensure the non-contamination of the products. 1µl of cDNA (100 ng/µl) was added to each mix (2µl, i.e. 200 ng/µl in the case of *Hoxa9*). The samples were loaded in a 96-well optical reaction plate with barcode (*ABI PRISM™, Applied Biosystems*). The RTQ-PCR was performed on the 7500 Real Time PCR System (*Applied Biosystems)* or the MxPro-Mx3000P *(Stratagene*) machine. The programme was set up as following: reverse transcription was achieved by heating the samples at 50°C for 30 minutes. The samples were then heated at 95°C for 15 minutes (10 minutes for *Stratagene*), 94°C for 15 seconds for the denaturation (95°C for 30 minutes for *Stratagene*), 54°C for 30 seconds for the annealing (55-60°C for 1 minute for *Stratagene*), 72°C for 40 seconds for the elongation (30 seconds for *Stratagene*). These three last steps were performed 40-42 times. The real time PCR employed the characteristics of a fluorescent dye, SYBR® Green, which binds to double-strand DNA. PCR products are thus visualized and quantified after each

cycle of amplification, due to the emission of fluorescence by the SYBR® Green. The results were analysed and quantified either with the 7500 System SDS Software (*Applied Biosystem*) or the Stratagene Software.

2.5.4. Real Time PCR

N-ChIPs and the C-ChIP results were analysed by Real-Time PCR, using 384 well plates on the ABI-PRISM 7900HT Sequence Detection System PCR machine from Applied Biosystems. The mix per tube was constituted as following: 5µl of SYBR® Green 2x (*QuantiTect™ SYBR® Green kit (200) from Qiagen*), 0.5µl of each primer, 10ng of DNA and DNase free water up to 10µl. The programme was set up as following: (1) the samples were first heated at 95°C for 15 minutes, (2) then at 94°C for 15 seconds for the denaturation, (3) at 58°C for 30 seconds for the annealing, (4) and at 72°C for 30 seconds for the elongation. These three last steps were performed 45 times. The efficiency of the reaction was checked by examining the standard curve generated from a gradient of concentration (generated from successive dilutions of the product of interest), for a specific primer set (Figure 2.5). A dissociation curve step was added at the end, in order to check the specificity of the product amplified (Figure 2.6). A single product will generate a single peak, corresponding to the melting temperature for a specific amplified fragment. The results were analysed and quantified with the SDS 2.2.2 Software (*Applied Biosystem*). The ChIP results are displayed as the ratio of the amount of DNA in bound fractions and normalised to the input (amount of DNA in the bound fraction performed with the pre-immune IgG).

2.6. Co-Immnunoprecipitation

$10x10^8$ HPC-7 cells were harvested and washed twice with ice-cold PBS, 5mM butyrate, before to being resuspended in 300µl of NP40 150mM NaCl buffer supplemented with

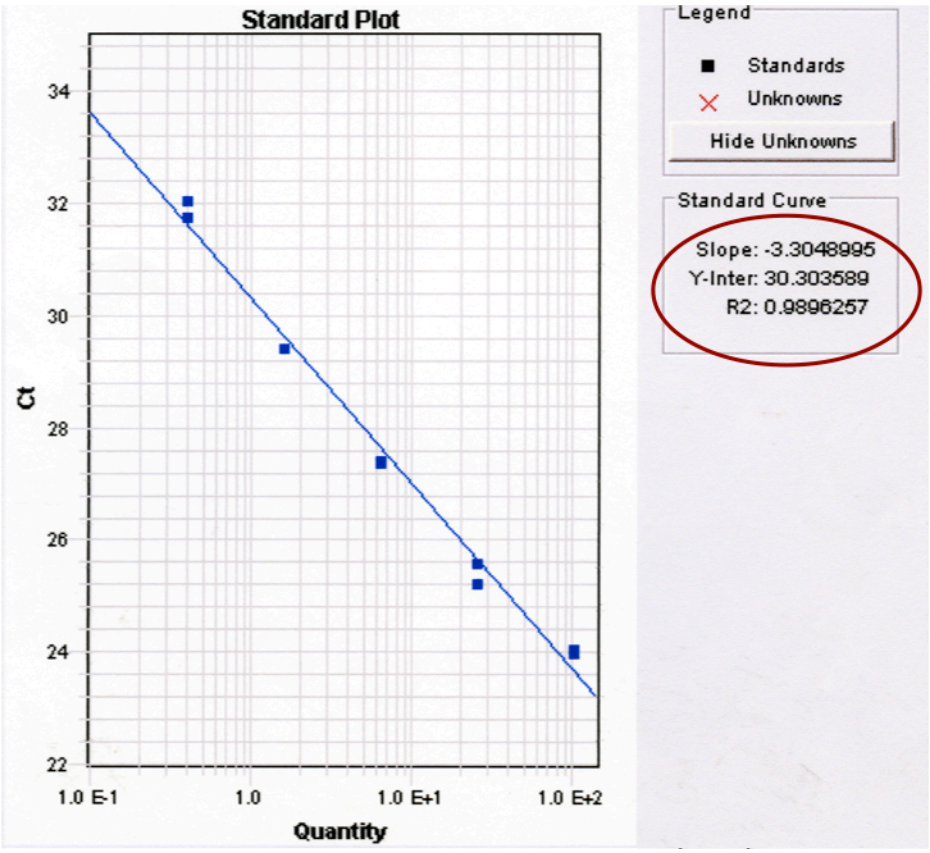


Figure 2.5: Assessing the efficiency of primer sets

Analysis of transcript abundance was performed using Real-Time PCR (and QRT-PCR). The quality of primer sets designed is checked by ensuring the standard curve generated from a gradient of different concentrations of a product displays a correct slope (-3.3) and a high percentage of efficiency.

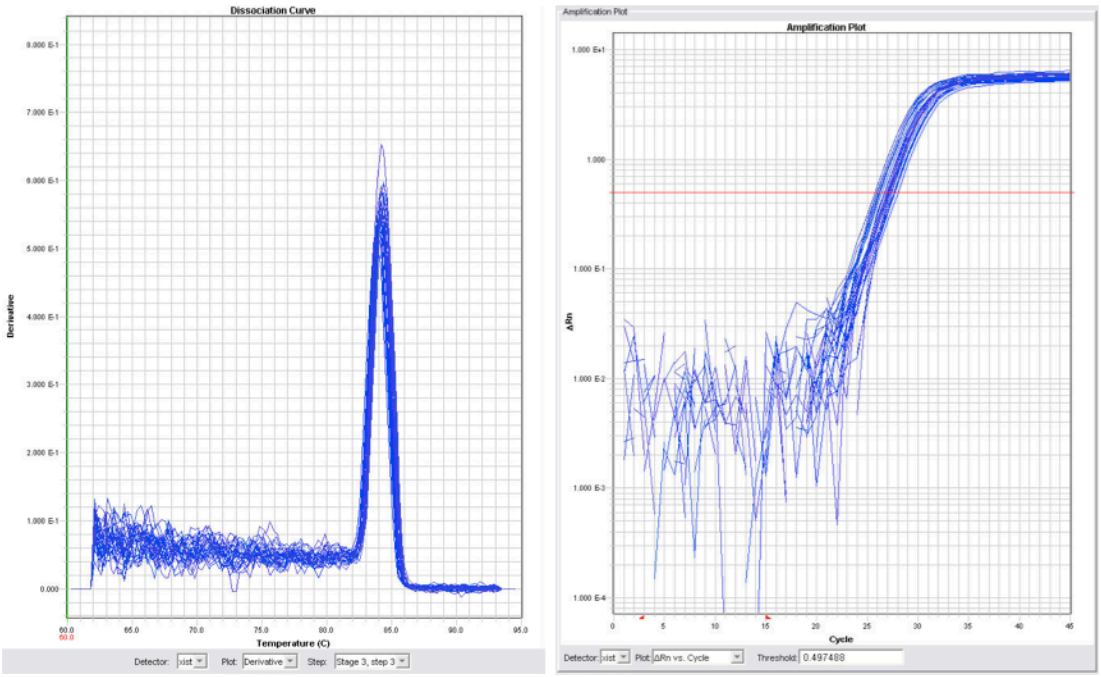


Figure 2.6: Assessing the quality of the QRT-PCR analysis

Analysis of transcript abundance was performed using Real-Time PCR (and QRT-PCR). The specificity of primer sets designed is checked by ensuring a single species is generated in the dissociation curve (*Left panel*) and generates smooth, consistent amplification curves, with an appropriate Ct threshold (*Right panel*).

PMSF and protease cocktail inhibitors (*Roche*). The cells were well vortexed and left for 20 minutes on ice to lyse. The lysate was then spun down at 13,000 rpm for 20 minutes. The supernatant was kept.

Protein A agarose beads were washed beforehand with NP40 150mM buffer. The lysate was then pre-cleared, by adjusting its volume up to 400µl with NP40 150mM buffer, supplemented with PMSF and protease cocktail inhibitors (*Roche*). Subsequently, 40µl of pre-washed agarose beads were added onto the lysate for 30 minutes at 4°C, on a rotating wheel. It is then spun down by centrifugation at 13,000 rpm for one minute, and the supernatant was kept. 5µg of the appropriate antibody was added to the pre-cleared lysate, and left overnight at 4°C. 40µl of protein A agarose was then added and left on the wheel for three hours before being spun down for three minutes at 3,000 rpm, 4°C. The supernatant was then kept (labelled "S1"). Beads were washed three times with NP40 150mM buffer (supplemented with PMSF and protease cocktail inhibitors), and once with NP40 200mM buffer (supplemented with PMSF and protease cocktail inhibitor). Samples were finally resuspended into 40µl Laemmli buffer.

NP40 150mM NaCl: 1% Nonidet P40, 10% glycerol, 50mM Tris pH7.5, 0.1% Sodium Azide, 150mM NaCl

NP40 200mM NaCl: 1% Nonidet P40, 10% glycerol, 50mM Tris pH7.5, 0.1% Sodium Azide, 200mM NaCl

Laemmli 4X: 200mM Tris-HCl pH6.8, 8% SDS, 40% glycerol, 2.4M ß-mercaptoethanol, Bromophenol blue

2.7. siRNA: cloning and DNA preparation

Primers were designed and to generate siRNA in order to perform *Mll* knockdowns for the mouse *Mll1*, *Mll2*, *Mll3* and *Mll5* genes (Figure 2.7).

The short hairpins RNA (*Mll3* and *Mll5*) used for the knockdown were designed (sequence identification using InvivoGen's siRNA Wizard™ program, "http://www.sirnawizard.com/") and cloned using a BLOCK-IT™ U6 RNAi Entry Vector Kit (*Invitrogen*). For each gene, six different single-stranded DNA oligos have been designed (three top strand oligos and three bottom strand oligos). Once the oligos are synthesized, double-stranded oligos are generated, by annealing equal amounts of each single-strand oligo. The reaction (5µl "Top strand" DNA oligo 200µM, 5µl "Bottom strand" DNA oligo 200µM, 2µl 10X Oligo Annealing Buffer (*Invitrogen*), 8µl DNase/RNase-free water) was incubated at 95˚C for 4 minutes. The reaction was cooled for 10 minutes at room-temperature, and then centrifuged briefly. A 100-fold dilution was made with RNase/DNase-free water, then another dilution (1/100) was made with 1/10 10X oligo Annealing buffer in RNase/DNase-free water to obtain a 5nM double-stranded products.

The double-stranded oligo was then cloned (ligation reaction) into the pENTR™/U6 vector. Ligation requires: 4µl of 5X ligation buffer, 2µl of pENTR™/U6 (0.5 ng/µl), 1µl of ds oligo (5nM), 12µl of DNase/RNase free water and 1µl of T4 DNA ligase (1 U/µl) were mixed and incubated 30 minutes at room-temperature, and then placed on ice. The ligation reaction was then transformed into One Shot® TOP10 chemically Competent *E.coli*. 20µl of TOP10 (*Invitrogen*) are added to 2µl of ligation reaction. The mix was left 5 minutes on ice, heat shocked for 30 seconds at 42˚C and then incubated for 1 minute again on ice. 200µl of LB (10 g Bacto-tryptone / 5g yeast extract / 10g NaCl for 1 litre) were added to the bacteria,

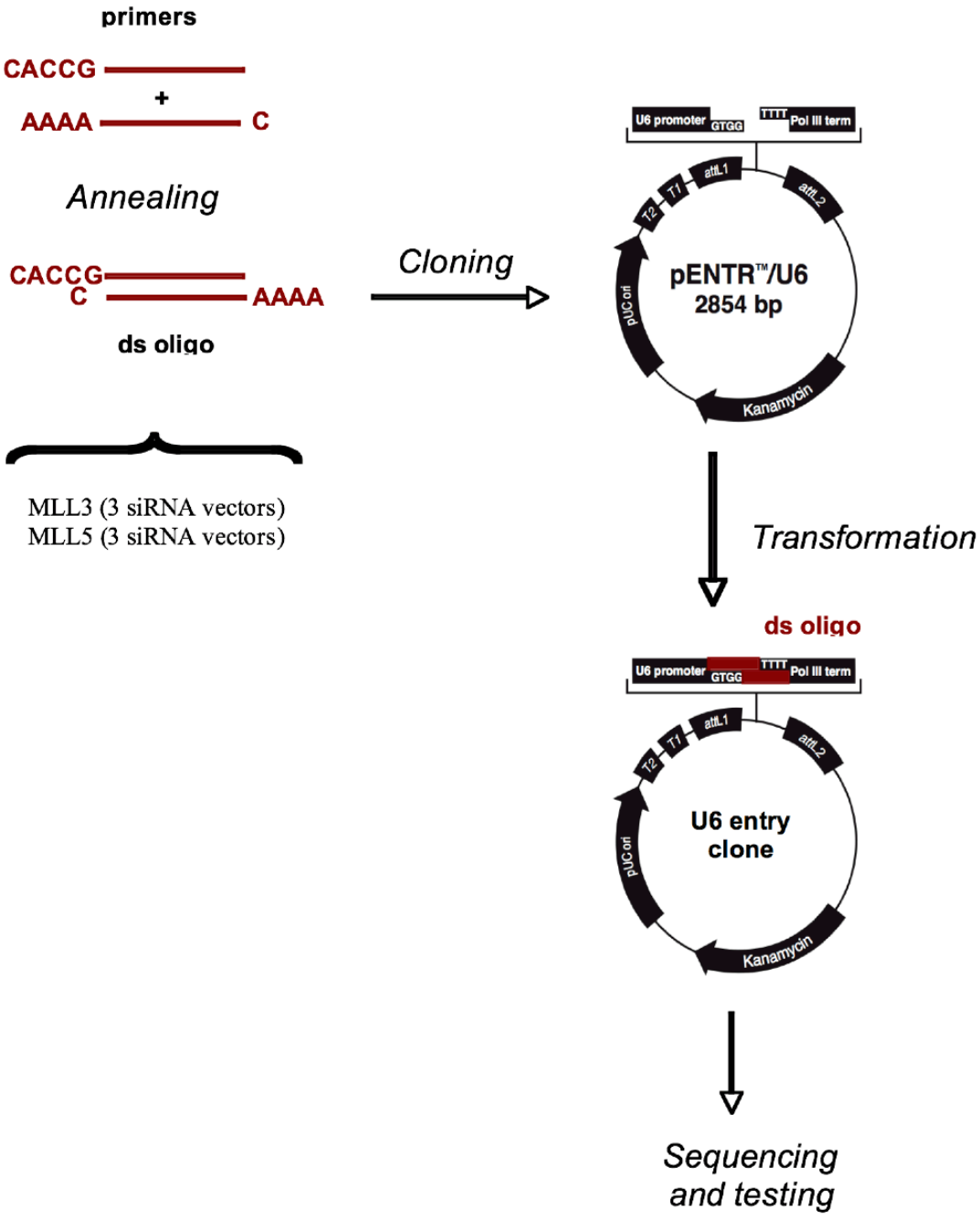


Figure 2.7: Establishing siRNA knockdown protocols

Three vectors were designed to knockdown *Mll3* or *Mll5* gene expression using Block-It pENTR/U6 vector kit (*Invitrogen*). Double strand fragments were generated then cloned in the pENTR/U6 vector. Clones were selected by kanamycin resistance.

The sequence of clones was checked and a higher amount of DNA was subsequently generated to be used for cell transfection. Three vectors have been cloned for each target. Each vector has then been tested for its efficiency.

which were shaken (300 rpm) for 1 hour at 37˚C. Finally, 50µl and 150µl of TOP10 were spread onto LB-agar-kanamycin (50 µl/ml) plates (*Mll3* and *5*) or LB-agar-ampicillin (100 µl/ml) plates (*Mll1* and 2*)* and incubated at 37˚C overnight. The day after, colonies were picked to prepare the plasmid according to a classical alkaline lysis technique. Once extracted from the bacteria, the DNA was sequenced (Functional Genomics and Proteomics Laboratory, University of Birmingham) to check if the double-stranded oligo has been correctly integrated. Correct vectors were then transformed again into TOP10 *E. Coli* to produce a high quantity of DNA. The DNA obtained was used to transfect HPC-7 cells and inactivate selective gene expression.

Other *Mll* (*Mll1*, *Mll2* and *Mll5*) double-stranded oligo pre-designed products were also used: Mouse GIPZ lentiviral shRNAmir clones [*Open Biosystems*]. (Clone Id: Mll2 V2LMM_193085 and V2LMM_151544; Mll1 V2LMM_96426 and V2LMM_96428; Mll5 V2LMM_107211). These clones were spread onto LB-agar-carbenicillin (100 µg/mL) plates and incubated at 37˚C overnight. DNA was isolated from the colonies picked as above. pGIPZ vectors express TurboGFP allowing measurement of transfection efficiency.

2.8. Transfection of HPC-7 cells

Knockdown vectors were transfected into HPC-7 cells using a commercial procedure (Cell Line Nucleofector® Kit L, Amaxa Biosystems). A 12-well plate was prepared with 1 ml of 2x HPC-7 Growth Medium (for 11ml final: 7.8 ml StemPro-34, 572µl StemPro-34 Nutrient Supplement, 1/50 penicillin/ streptomycin, 1/50 glutamine, 2.2ml CHO/SCF conditioned medium) and 400 µl unsupplemented StemPro medium per well. $2x10^6$ cells were required per nucleofection and per well. An Amaxa biosystems Cell line nucleofector® kit L was used according to manufacturer's instructions. Cells were centrifuged at 90 rcf for 10 minutes at

room temperature (RT), and resuspended in 100µl of Nucleofector Solution L. 2µg of DNA was added per cell suspension (DNA from siRNA clones or a plasmid expressing GFP was used as a control). The cell suspension was transferred into an appropriate cuvette for transfection with the Amaxa Biosystems Nucleofector™ II (programme X-001). 500µl of pre-warmed RPMI 1640 medium [*Gibco*] were immediately added after transfection. The cells were then transferred into an eppendorf tube and left 30 minutes to an hour at 37˚C for recover. The cells were finally transferred into pre-warmed 12-well plate and incubated for the required time period.

2.9. Flow Cytometry Analysis

After transformation, the cells were incubated overnight to allow them to revive before cell sorting, and preparation for Flow Cytometry analysis (FACS: **F**luorescence-**A**ctivated **C**ell **S**orting). The cells are harvested, spun at 850 rpm (90g) for five minutes. The pellet is resuspended in StemPro-34 media (*Invitrogen*), plus supplements (300 µl per $10x10^6$ cells) with no FCS. The cells are filtered through a cell strainer immediately before sorting. The filter is placed on top of a FACS tube. Cells are sorted in collection tubes (one per sample) containing 250µl Stem Pro media plus supplements.

3. RESULTS - CHARACTERISATION OF THE HAEMATOPOIETIC STEM CELL DIFFERENTIATION MODEL

A differentiation model was established using Haematopoietic Progenitor Cells (HPC-7), an immortalised mouse embryonic stem cell line derived from haematopoietic stem cells transformed with the transcription factor Lhx2 (Pinto d'O *et al*., 1998). HPC7 cells have the capacity to differentiate into a variety of haematopoietic cell lineages, including monocytes and megakaryocytes the lineage we focus on in this work.

3.1. Generating megakaryocytes

HPC-7 cells present several advantages for this study, notably that they can differentiate under controlled conditions to megakaryocytes, which are then easily purified from undifferentiated cells. Haematopoietic stem cells were maintained in the presence of steel factor (conditioned media), and self-renewed until differentiation was induced by adding a unique cytokine, thrombopoietin (TPO), involved in megakaryopoiesis and thrombopoiesis. Differentiated cell morphology differs from the undifferentiated stem cells: megakaryocytes become substantially larger, with an increase in ploidy (Figure 3.1a). Megakaryocytes were separated from the small undifferentiated cells using two consecutive BSA gradients. The purity of differentiated cell populations was checked by FACS analysis: a homologous population of megakaryocytes can be isolated as characterised by the number of cells containing greater than 8N (the majority are 16 and 32N) (Figure 3.1b). Moreover, the differentiation process was characterised by changes in gene expression, notably at *Hoxa9* and *Meis1*, genes dysregulated in leukaemia. Relative expression for these particular genes was assessed after three and six days of TPO-induced differentiation: *Hoxa9* gene expression

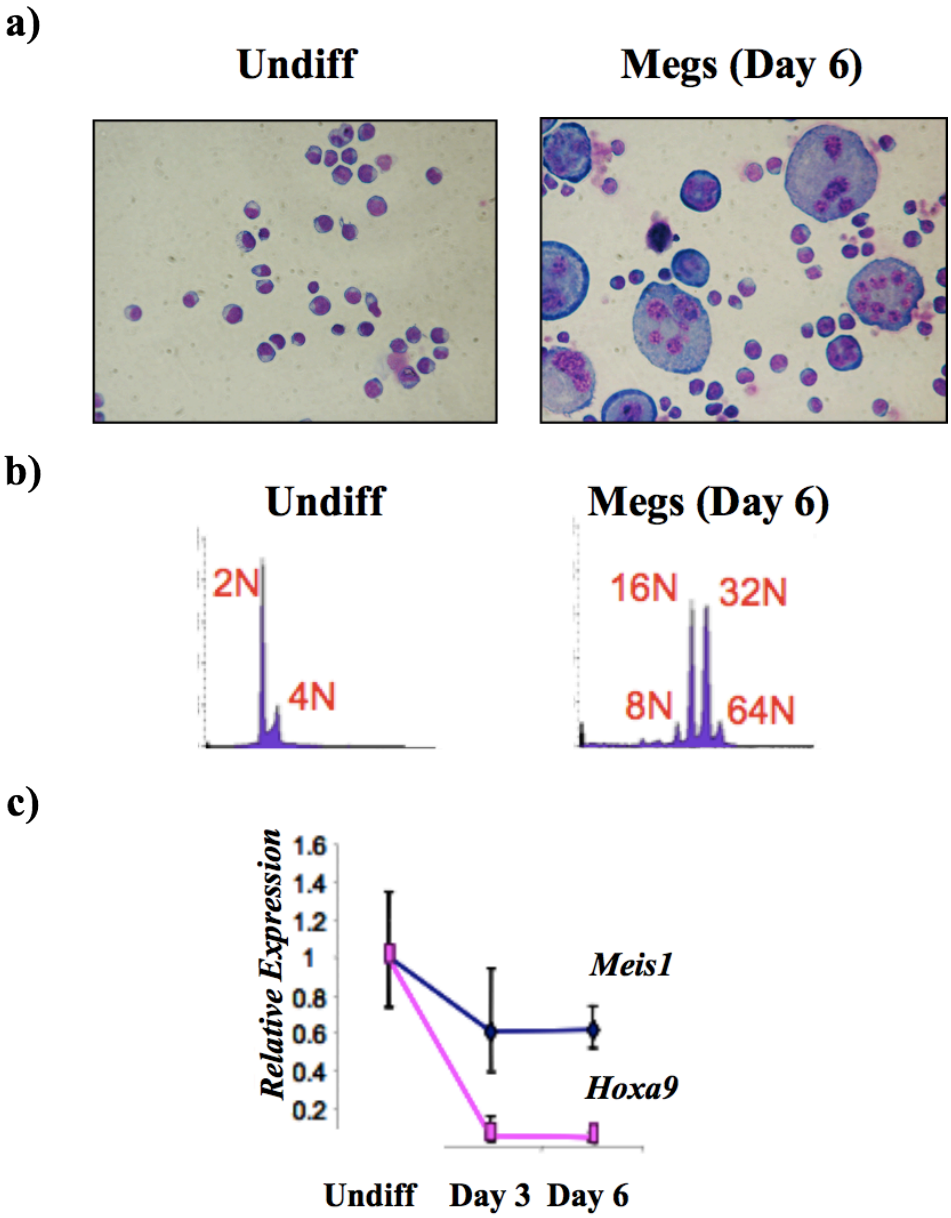


Figure 3.1: Our model for differentiation: HPC-7 cells

Cells were differentiated for six days in the presence of thrombopoietin. Megakaryocytes were isolated by fractionation on BSA gradients. **a)** Cell morphology of undifferentiated HPC-7 cells (*left panel*) and after 6 days exposure to TPO (*right panel*). Morphology was assessed by *Diff-Quick* staining of cytospin preparations. **b)** FACS analysis of undifferentiated cells (Undiff) and megakaryocytes (Megs). **c)** Real-time PCR analyses of relative gene expression in differentiating HPC-7 cells were performed with samples taken at different points in the differentiation process (Undifferentiated cells, Day 3 and Day 6). RNA samples were analysed in triplicate, and gene expression expressed relative to *ß-actin* expression (Dr. F. Tavner).

displays a substantial decrease (95%) whereas *Meis1* gene expression is slightly reduced (40%) (Figure 3.1c, data from Dr. F. Tavner).

3.2. Characterising megakaryocyte maturity

Previous data (Figure 3.1b) showed that a pure fraction of differentiated cells could be obtained using this differentiation protocol with HPC-7 cells. However, expression of particular control genes was assessed in order to confirm that the cells produced were fully mature megakaryocytes. Three genes involved at specific stages of development were used as controls. PU.1 is the product of the *Spi-1/Sfp-1* gene and a haematopoietic transcription factor required in early development. Only a few mature haematopoietic cells express PU.1, including monocytes, macrophages, lymphocytes B, mast cells and neutrophils (Valledor *et al*., 1998). PU.1 expression in mature megakaryocytes is therefore not expected, and as can be observed, *Sfpi-1* is four-fold down regulated in our megakaryocyte populations (Figure 3.2a). Another gene, whose expression was checked, is *Gata-1*, which encodes GATA binding Protein 1, present in progenitor cells. *Gata-1* expression is increased 18-fold in megakaryocytes (Figure 3.2b). GATA-1 drives the transcription of the *Gp6* gene. GP6 (Glycoprotein VI) is a platelet receptor exclusively expressed in the megakaryocyte lineage (Jandrot-Perrus *et al*., 2000; Holmes *et al*., 2002). *Gp6* gene expression considerably increased during HPC-7 differentiation (~3500-fold), indicating that the cells obtained are mature megakaryocytes (Figure 3.2c).

3.3. Changes in *Hoxa* gene activity during differentiation

Haematopoiesis is associated with the controlled regulation of particular genes by protein complexes, such as the MLL complex. Key genes in the differentiation of haematopoietic cells are the *Hoxa* genes. These genes encode transcription factors essential to determine the

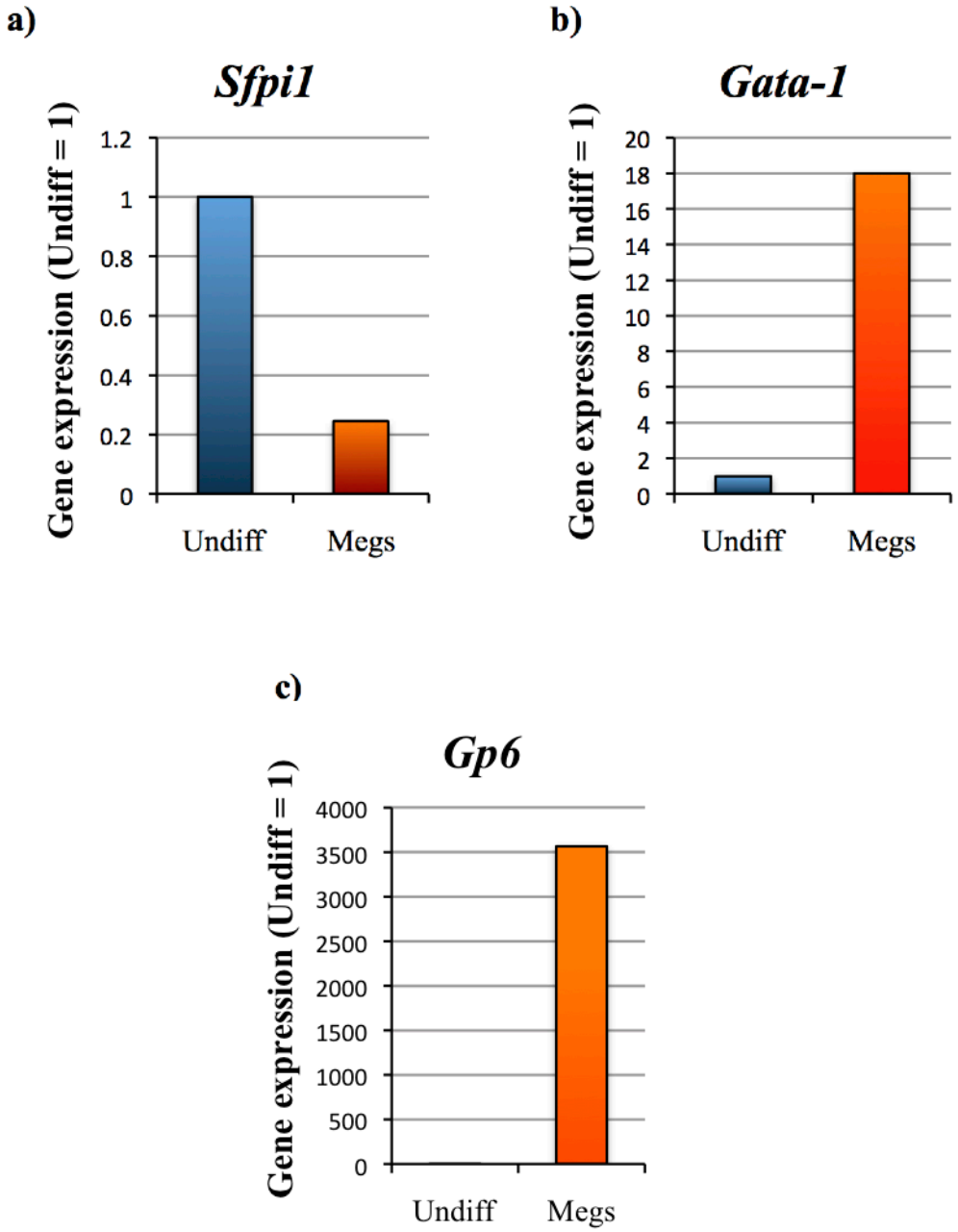


Figure 3.2: Expression of control genes in generated megakaryocyte cell population

To assess the quality of our model, control genes (*Sfpi1*, *Gp6* and *Gata-1*) were used to evaluate megakaryocytes maturity. PU.1, encoded by the gene *Sfpi1* (**a**), is down-regulated in megakaryocytes whereas *Gata-1* is expressed in megakaryocytes (**b**). *Gp6* is exclusively expressed in mature megakaryocytes (**c**). RNA was extracted from undifferentiated HPC-7 cells (Undiff) and megakaryocytes (Megs), and analysed in triplicate using *Gapdh* as a control. PCR was performed in triplicate from one experiment.

identity of regions of the embryonic anterior-posterior body axis during development (Krumlauf, 1994). These genes are distributed in a co-linear organisation, and are temporally expressed according to this spatial organisation (Figure 3.3a). We examined the changes of the *Hoxa* cluster (*Hoxa1-13*) gene expression during differentiation by assessing changes in transcript abundance in undifferentiated HPC-7 cells and megakaryocytes. In the first instance, Ct (Cycle threshold) values were examined along the *Hoxa* cluster gene (*Hoxa1-13*), reflecting transcript abundance (Figure 3.3b). It should be noted that *Gapdh* transcript abundance does not change during differentiation. Apart from genes at the edge of the cluster (*Hoxa1* and *Hoxa13*), there is a global increase of Ct values, indicating a reduction in transcript abundance at these genes during differentiation from undifferentiated HPC-7 cells to megakaryocytes. This is clearest for *Hoxa5*, which shows three cycles of difference between undifferentiated HPC-7 cells and megakaryocytes indicating a six-fold decrease in transcript abundance. Then, *Hoxa* gene expression normalised with *Gapdh* transcript abundance was examined. This alternative way of presenting this information (Figure 3.3c) shows a broad reduction of *Hoxa* gene expression on the *Hoxa* cluster in megakaryocytes. However, it must be noted that the transcript abundance of the genes at the edges of the cluster (*Hoxa1* and *Hoxa13*) display wide variations, while other genes, especially *Hoxa5* and *Hoxa6*, show a consistent reduction in expression. It should be noted that *Hoxa5* and *Hoxa6* Ct values are low (around 23 for *Hoxa5* in undifferentiated cells) in comparison with other *Hoxa* genes (around 34 for *Hoxa13* in undifferentiated cells), reflecting the relatively high abundance of these transcripts in the cells. Similarly, the change of Ct value is bigger for *Hoxa5* (going from 23 to 25) than for *Hoxa13* (going from 34.74 to 34.44, *Supplementary data*). Therefore, assuming equivalent primer efficiency, the biggest changes were observed at *Hoxa5* and *Hoxa6*.

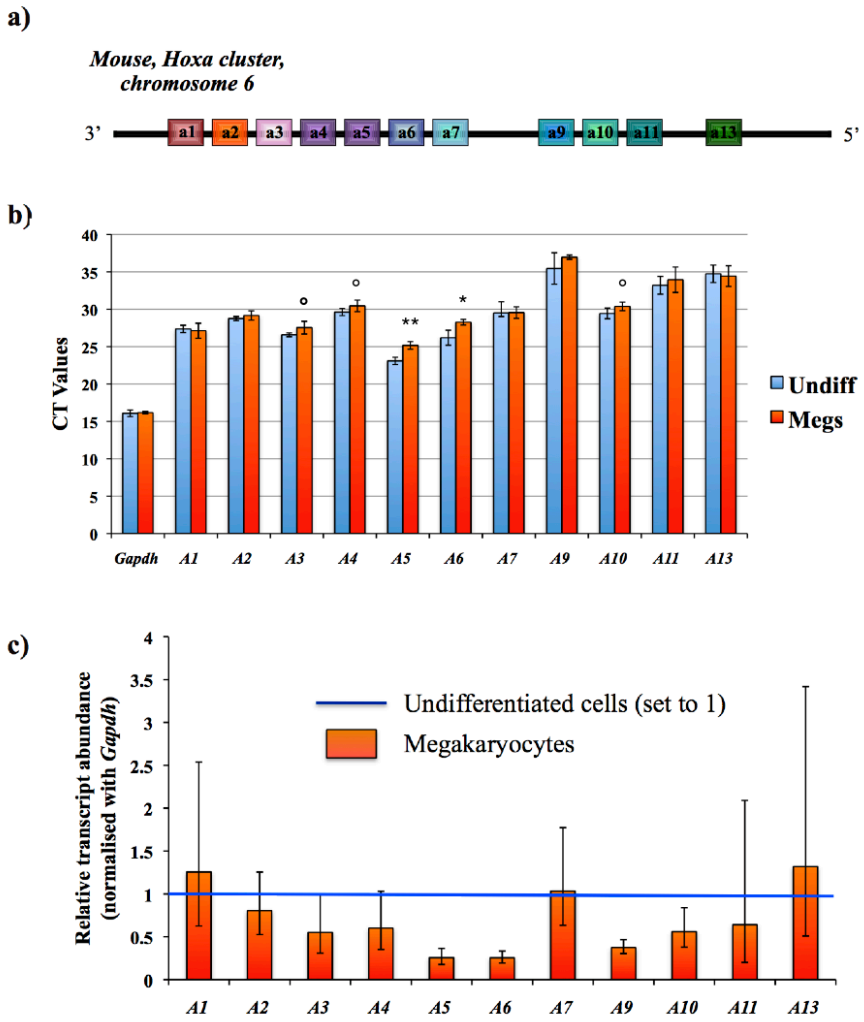


Figure 3.3: *Hoxa* gene cluster transcript levels in undifferentiated HPC-7 cells and megakaryocytes

Quantification by Real-Time PCR of levels of gene transcripts (normalised with *Gapdh*) along the *Hoxa* cluster in undifferentiated HPC-7 cells (blue) and megakaryocytes (orange).

a) Figure of *Hoxa* cluster, showing the positions of *Hoxa* genes.

b) Changes in transcript abundance (Ct values) in undifferentiated HPC-7 cells (blue) and megakaryocytes (orange). It should be noted that *Gapdh* abundance does not change. (**p<0.01, *p<0.05, °p<0.5, student test, comparing CT values in undiff. HPC-7 and megakaryocytes samples).

c) Changes in *Hoxa* gene expression upon differentiation, with the level of transcripts in undifferentiated cells set at 1.0. It should be noted that the transcript abundance of the genes at the edges of the cluster (*Hoxa1* and *Hoxa13*) display wide variations, reflected by wide error bars. qRT-PCR was performed in triplicate from more than two biological experiments.

It should also be noted that the decrease of the abundance of *Hoxa* gene transcripts is lower than expected. Transcript stability was next examined in order to assess if this could impact on RNA abundance.

3.4. Degree of stability of *Hoxa* gene transcripts

The stability of *Hoxa* gene transcripts was assessed by exposing undifferentiated HPC-7 cells to the toxin α-amanitin, which is an inhibitor of the RNA Polymerase II (Figure 3.4). Cells were treated for 8, 16, 24 and 48 hours before being washed and harvested. RNA was extracted, and RT-QPCR performed with the QIAGEN QuantiTect SYBR Green RT-PCR Kit©. Values correspond to the relative abundance of specific *Hoxa* genes (*Hoxa3*, *Hoxa4* and *Hoxa5*), normalised to the control *Gapdh* gene. Presentation of transcript abundance data indicate that the abundance of these transcripts is maintained until 24 hours of treatment, and is even slightly increased for *Hoxa3* and *Hoxa5* transcripts, before starting to decrease following 48 hours treatment. Therefore, transcripts of *Hoxa* genes are still detectable in cells 48 hours after the cessation of polymerase activity, suggesting that *Hoxa* gene transcripts are stable for extended periods. This stability could explain the apparently limited down-regulation of the *Hoxa* genes (i.e. *Hoxa1-4, Hoxa7, 11* and *13*) observed in megakaryocytes (i.e. Figure 3.3). Subsequent analysis of both MLL complex (i.e. Figure 5.2) histone modifications (i.e. Figure 4.3), and RNA polymerase II (i.e. Figure 5.6) give conflicting indications of the changes that are occurring on these genes upon differentiation (*Discussed later*).

3.5. General distribution of epigenetic marks on cells

These data established that there is a global down regulation of *Hoxa* gene expression during differentiation, and therefore a substantial difference in gene expression between

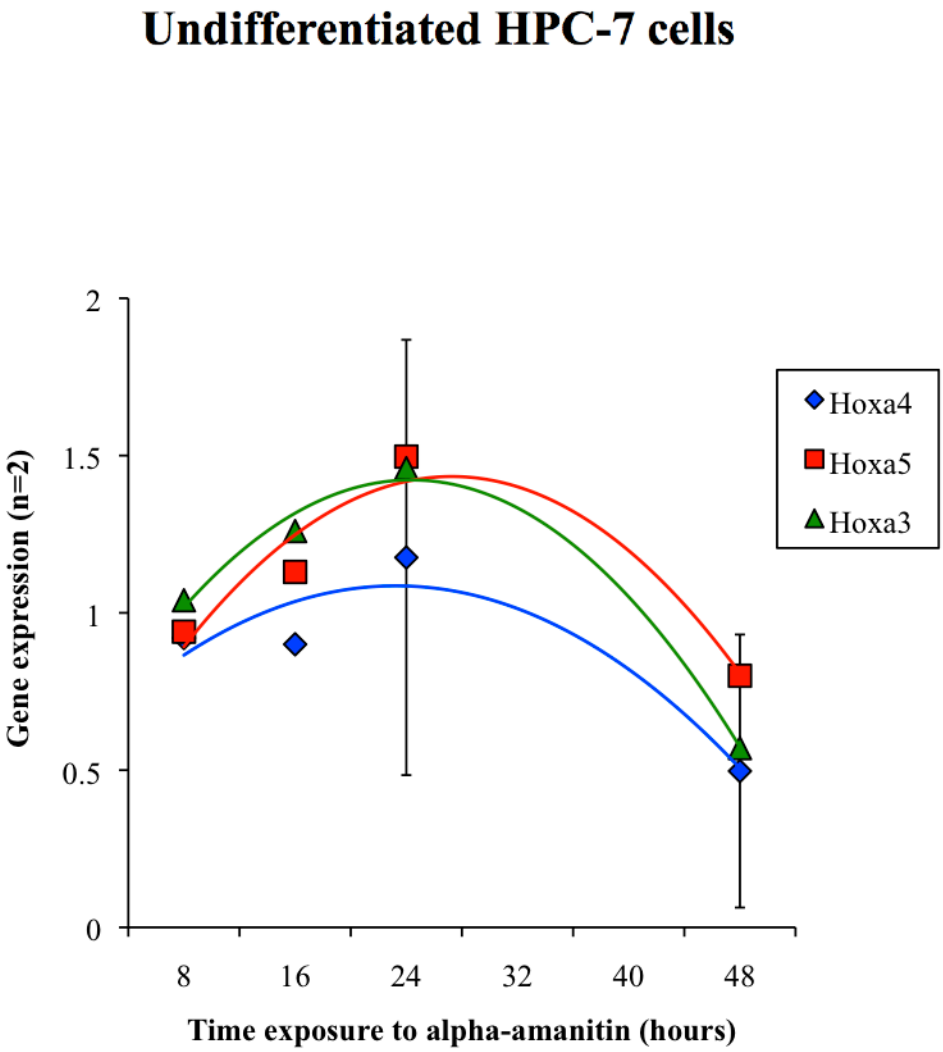


Figure 3.4: Time exposure to α-amanitin in undifferentiated HPC-7 cells

Undifferentiated HPC-7 cells were treated with alpha-amanitin for 8, 16, 24 and 48 hours. PCR was performed with primer sets corresponding to the coding region of *Hoxa3*, *Hoxa4* and *Hoxa5*, and was performed in triplicate from two experiments. Error bars represent the standard error of the mean.

undifferentiated HPC-7 cells and megakaryocytes. We next examined whether these changes in transcriptional activity reflected global changes in histone modifications, using immunofluorescence microscopy (Figure 3.5). Both undifferentiated HPC-7 cells (*Upper panels*) and megakaryocytes (*Lower panels*) were stained with histone modification-specific antibodies and compared to DAPI-stained nuclei. In megakaryocytes, cells are polyploid (as described previously), which are characterised by nuclear "lobes". One potential issue arises from our use of megakaryocytes; that is whether all copies of the *Hoxa* genes are equally regulated, in the whole genome, particularly as ChIP experiments assume that all copies of *Hoxa* genes are assembled into equivalent chromatin. Antibodies against activating marks (H3K4me3, H3K9ac) (Figure 3.5, **a** and **b**) and antibodies against repressive marks (H3K9me2, H3K27me3) (Figure 3.5, **c** and **d**) were used. As expected in undifferentiated cells, the histone marks are uniformly spread throughout the nucleus, though H3K9me2 staining shows a punctate distribution (**c**). It should be also noted that this mark is concentrated in the centromeres of metaphase chromosomes. In contrast, in megakaryocytes, histone marks are distributed overall, but highly concentrated regions of histone marks are associated with what could correspond to the "lobes", or structures "budding off" from the nucleus, as well as what appear to be "vacuoles". A similar distribution was observed for activating marks (H3K4me3, H3K9ac) (**a** and **b**) and repressive marks (H3K9me2, H3K27me3) (**c** and **d**) with the exception of H3K9me2, which shows punctate distribution as in undifferentiated cells.

These images were unexpected, but do not appear to reflect a technical artefact: we observed similar uneven distributions of histone modifications in cells, which were cross-linked either before or after incubation with antibodies. It was then essential to investigate whether this distribution reflected the histones in primary haematopoietic Stem Cells (hSC), as this could

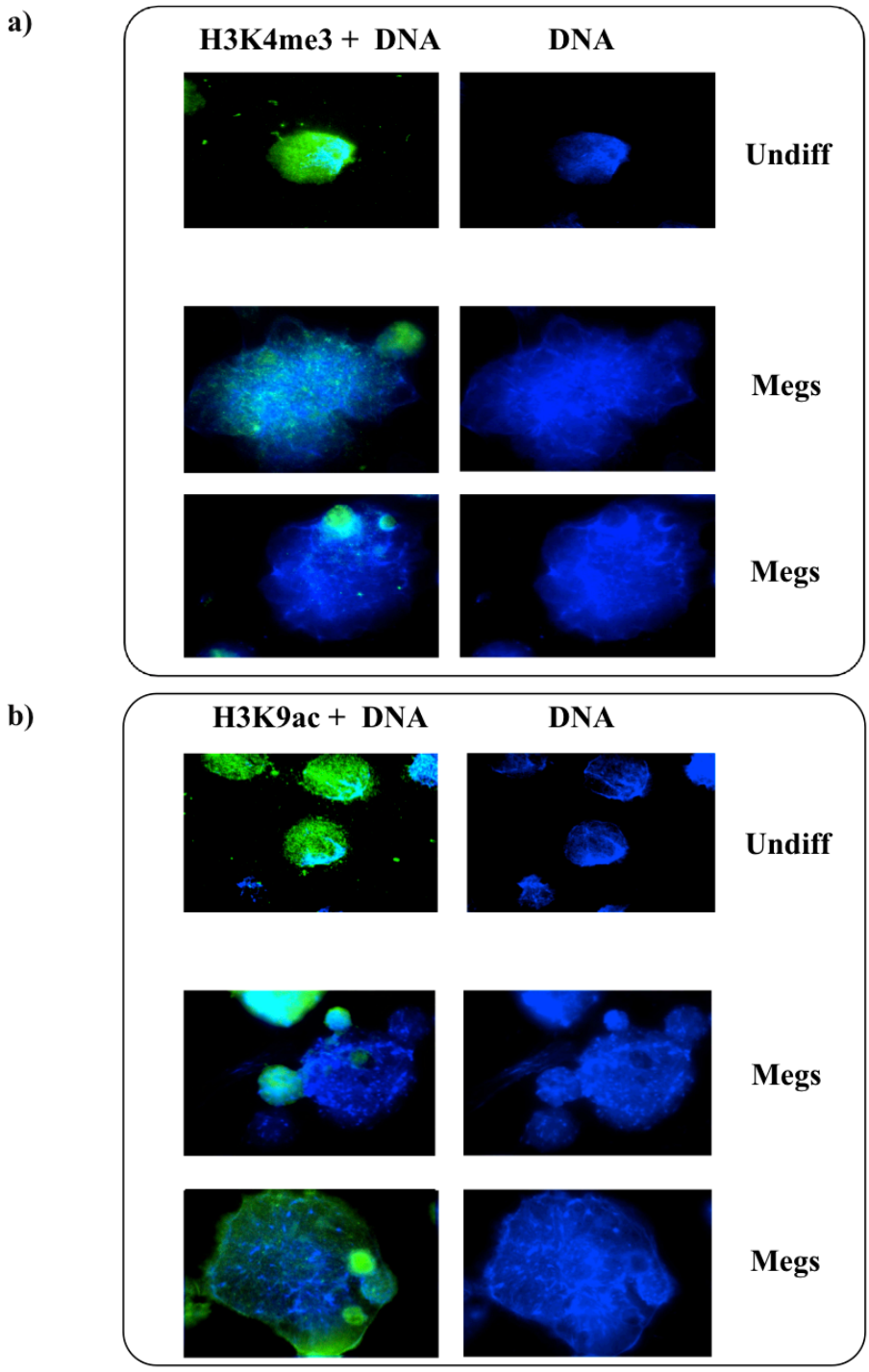


Figure 3.5 a & b: Staining with H3K4me3 (a) and H3K9ac (b) antibodies and revelation by immunofluorescence

Undifferentiated HPC-7 cells (*Upper panels*) and megakaryocytes (*Lower panels*) have been stained with DAPI alone to highlight the DNA (*Right panels*) or with the antibody of interest on top of it (*Left panels*). Magnification X400, Megs; X600, Undiff.

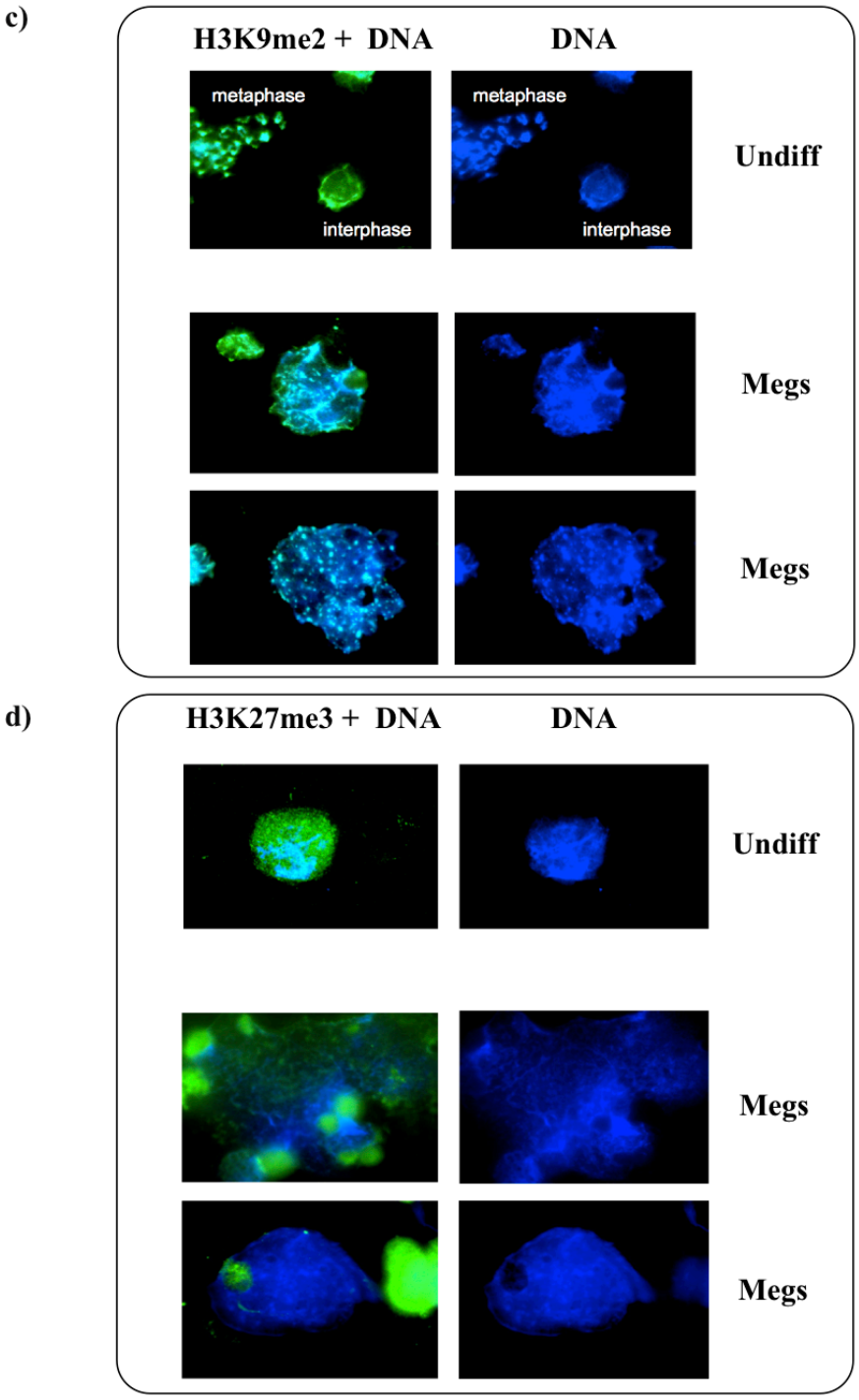


Figure 3.5 c & d: Staining with H3K9me2 (c) and H3K27me3 (d) antibodies and revelation by immunofluorescence

Undifferentiated HPC-7 cells (*Upper panels*) and megakaryocytes (*Lower panels*) have been stained with DAPI alone to highlight the DNA (*Right panels*) or with the antibody of interest on top of it (*Left panels*). Interphase and metaphase cells are indicated (**c**). Magnification X400, Megs; X600, Undiff.

reflect a HPC-7 specific anomaly. These cells were generated, and treated in a similar way to the HPC-7 cells and megakaryocytes, in order to confirm that they both are characterised by the same distribution of epigenetic marks. The morphology of undifferentiated hSC and differentiated hSC after five days of treatment with thrombopoietin was first examined (Figure 3.6). Undifferentiated primary hSC are small and round cells, while TPO-induced megakaryocytes appear similar to the cells produced from HPC-7 cells (Figure 3.6a). Histone marks display a similar distribution of H3K9me2 and H3K27me3, showing punctate (H3K9me2) and "all over" staining with concentrations in "vacuoles" (Figure 3.6b). This suggests that our previous observations in HPC-7 do not seem to reflect a cell-specific artefact, as undifferentiated cells stain as expected and most antibodies give the same distribution. These "lobe/vacuole" structures could thus either reflect areas where genome is replicating, or correspond to the processes that lead to the break down of cells to generate platelets. The latter seems unlikely though, as these "lobe/vacuole" structures were also observed in nuclei of smaller non-mature megakaryocytes.

3.6. Global chromatin changes during differentiation

The previous immunofluorescence studies indicate that the histone modifications examined are present in both undifferentiated HPC-7 cells and megakaryocytes, but these studies cannot give insight into the abundance of histone marks in these two cell types. Previous studies have detected changes in global histone modifications upon differentiation, notably in histone acetylation (Marin-Husstege *et al.*, 2002; Lee *et al.*, 2004; Wu and Sun, 2006). This is why global changes in histone modifications were then examined during differentiation.

We initially examined the histone proteins in both undifferentiated HPC-7 cells and megakaryocytes to assess if large changes occurred in histone isoform variants. Histones

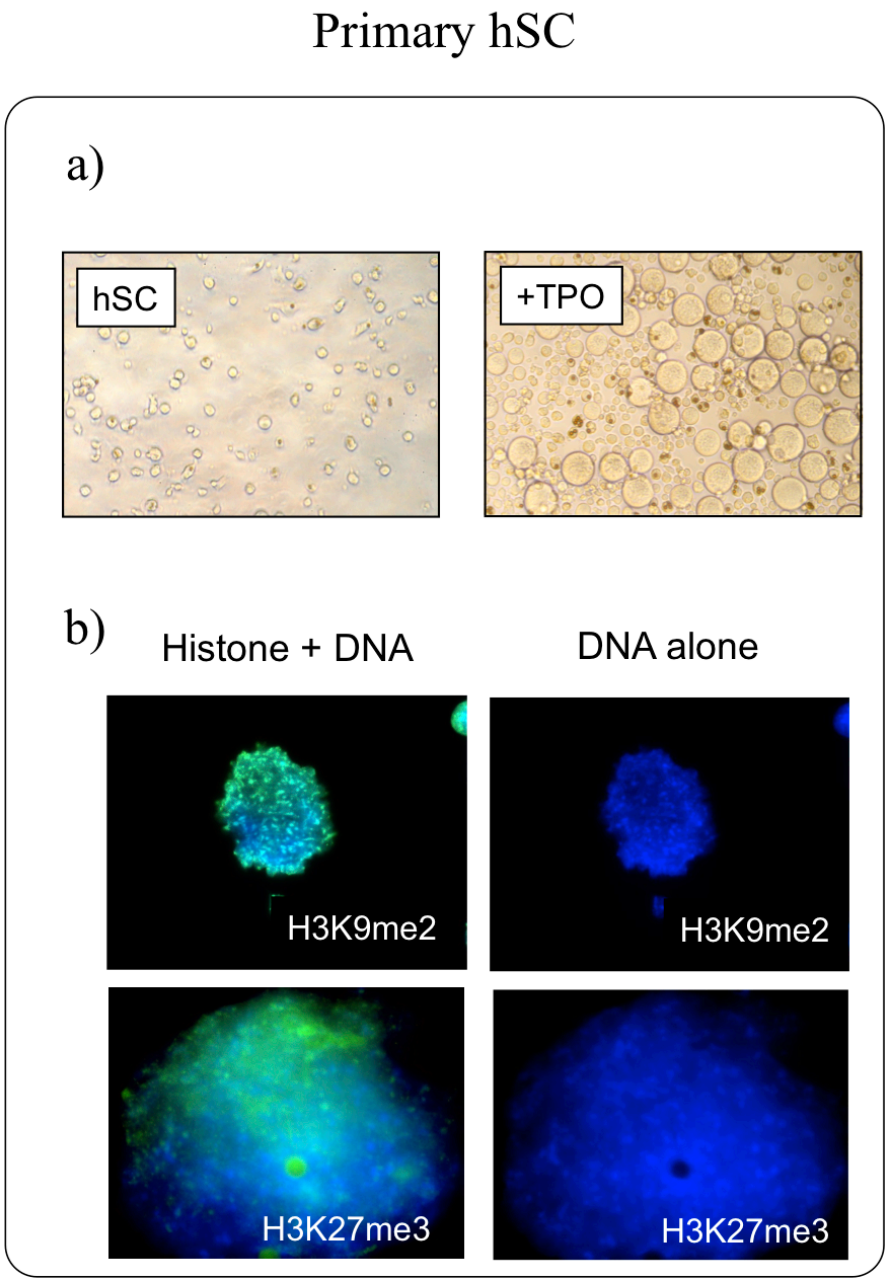


Figure 3.6: Characterising primary haematopoietic Stem Cells (hSCs)

a) Cell morphology of undifferentiated primary hSC and after five days exposure to TPO. Morphology was assessed by microscope observation, magnification (X20).

b) Undifferentiated haematopoietic Stem Cells (*Lower panels*) were stained with DAPI alone to highlight the DNA (*Right panels*) or with the antibody of interest (H3K9me2 or H3K27me3) (*Left panels*). Magnification (X400).

were isolated by acid (HCl) extraction, monitored for degradation by SDS gel (Figure 3.7) and the loading adjusted by Western blotting with an H3 C-terminal antibody (Data not shown). This is an ideal control, as the C-terminal region of the histone H3 contains no known modifications. No substantial changes are observed in histone abundance, though a novel high molecular weight band is observed only in differentiated HPC-7 cells (Figure 3.7, "?"). However, this novel protein is not extracted with perchloric acid (PCA), using the standard protocol for extracting linker histones, suggesting that it is not a linker histone.

A 2-dimentional SDS-AUT gel was performed in order to identify potential general changes in the histones occurring between undifferentiated HPC-7 cells and megakaryocytes (Figure 3.8). Histones were first separated according to their degree of acetylation, by migration on an AUT (Acid-Urea-Triton) resolving gel. This gel also allows the separation of modified histone variants and subtypes of histones on the basis of their charge and mass. Then, histones were separated according to their size, on a SDS denaturing gel. Comparing the two gels (a and b) shows no substantial changes in the distribution of the variant histones H3.1, H3.2 and H3.3, in spite of the fact that they appear to be smeared in megakaryocytes, which could be due to the acquisition of a phosphate. Furthermore, no noticeable changes in the degree of histone acetylation were shown, with the exception of a novel band revealed in undifferentiated HPC-7 cells (arrowed). The proximity of this new band with that of histone H2B, and its reduced mobility in the TAU dimension suggests that this probably reflects the acetylation of H2B. Subsequent mass spectroscopic analysis of this histone "spot" has shown that this contains a number of histone acetyl - methyl isoforms, including modifications at novel sites (*K. Nightingale, H. Cooper, Personal communication*).

Subsequently, different histone marks were more specifically examined by Western blot. These include marks associated with both the active (H3K4me1,2,3, H4K8ac and H4K16ac),

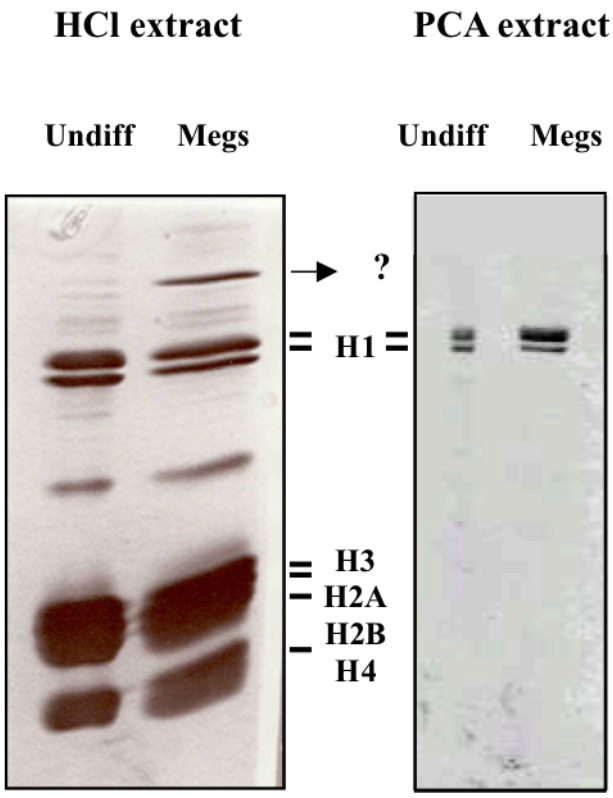


Figure 3.7: Changes in histone proteins between undifferentiated HPC-7 cells and megakaryocytes

Histone proteins were examined in undifferentiated HPC-7 cells (Undiff) and megakaryocytes (Megs). Histones were acid extracted (HCl) and then examined by silver staining. An unknown protein (?) is observed in megakaryocytes. This is not observed in Perchloric Acid (PCA) extracted samples (*Right panel*).

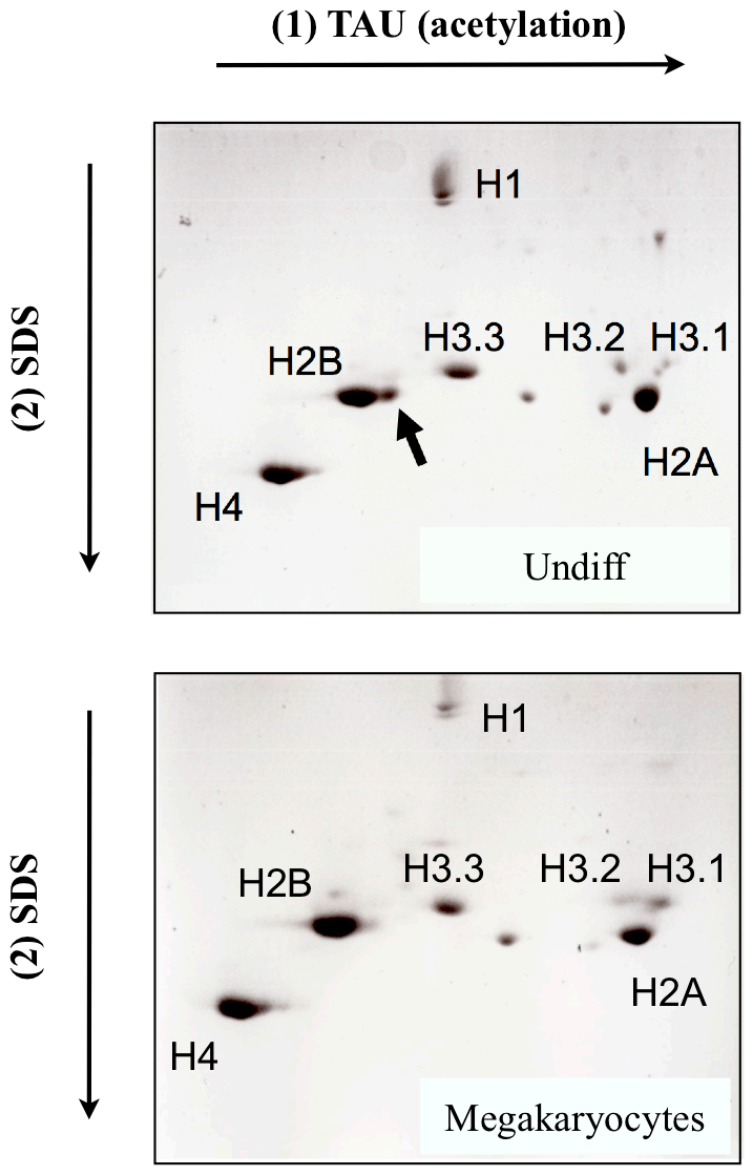


Figure 3.8: 2D gel identifying histone changes in undifferentiated HPC-7 cells and megakaryocytes

Histones from undifferentiated HPC-7 (**a**) and megakaryocytes (**b**) were first separated on a SDS denaturing gel, depending on their molecular weight. Then the proteins separated a second time on a TAU (Triton-Acid-Urea) gel, on the base of the charge and mass, according to their degree of acetylation. In undifferentiated HPC-7 cells (**a**), it should be noted the presence of a novel band (*arrow*), that is not observable in megakaryocytes (**b**).

and repressive (H3K9me2) states of gene expression and a mitotic chromatin (H3S10p) (Figure 3.9a). Quantification was normalised using an antibody against the H3 C-Terminal. Quantification of the abundance of these histone marks in megakaryocytes is shown relative to the abundance in undifferentiated HPC-7 cells, which was set at 1 (Figure 3.9b). Broadly most marks do not show substantial changes: H3K4me1 and H3K4me2 and H3K9me3 are slightly increased in megakaryocytes, whereas H3K4me3 and H4K16ac show small reductions, but these data need to be reproduced. No difference is observed with the acetylated H4K8 mark. In contrast, the most significant difference is observed with the phosphorylated H3S10 mark, which is increased by three-fold in megakaryocytes. This is consistent with the increase in mitotic activity in the polyploid megakaryocytes.

3.7. Changes in the MLL1 Histone Methyl-Transferase distribution during differentiation

Our previous studies assessing the abundance and distribution of histone modifications in undifferentiated HPC-7 cells and megakaryocytes showed that differentiation is not associated with large changes in histone modifications. This raised the question whether this could reflect a change in the abundance of specific histone methyl-transferase enzymes. The MLL family of histone methyl-transferases possesses a SET domain, which gives these enzymes the capacity to methylate H3K4. Moreover, the archetypal MLL protein, MLL1, plays an important role in haematopoiesis (Ernst *et al*., 2002; Ernst *et al*., 2004a). *Mll1* gene expression was therefore examined in undifferentiated HPC-7 cells and megakaryocytes (Figure 3.10a). The four MLL family members are expressed in HPC-7 cells. *Mll1*, *Mll4*, *Mll3* and *Mll5* transcript abundance was initially quantified by semi-quantitative PCR using [α-^{32}P]-dATP labelled nucleotide incorporation, and normalised with ß-actin gene expression

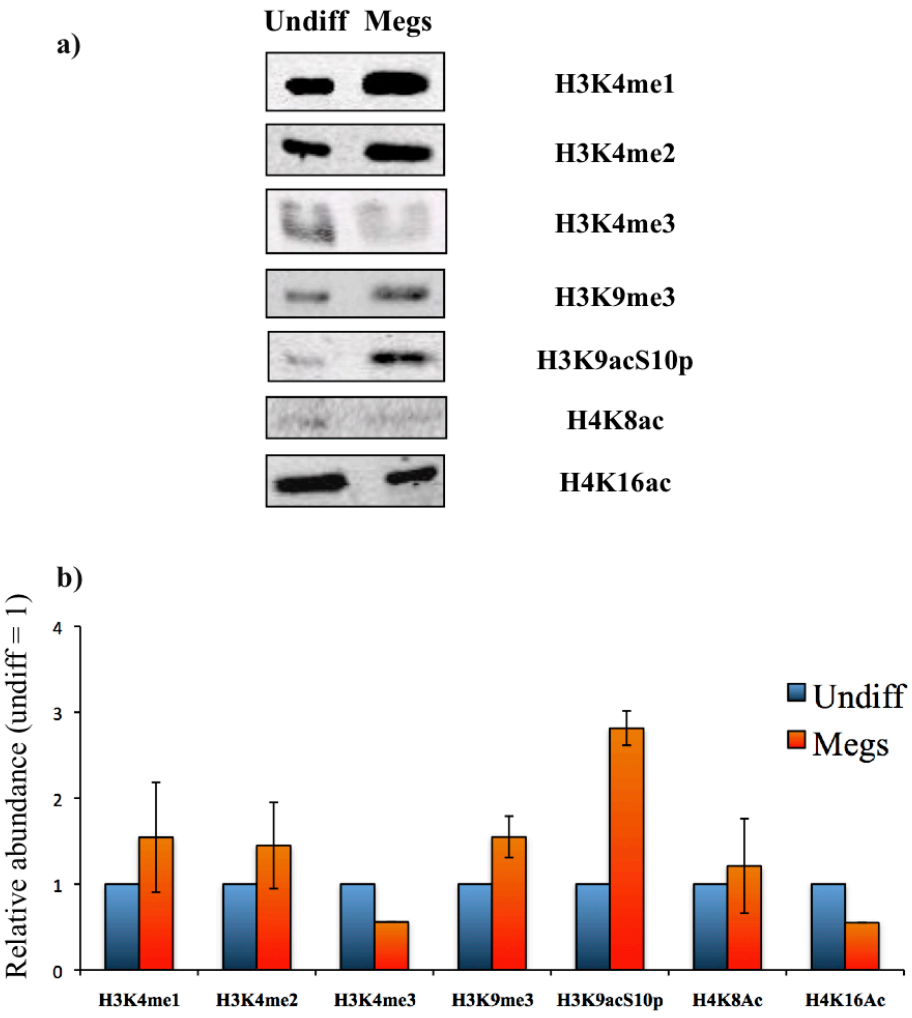


Figure 3.9: Changes in global histone modification between undifferentiated HPC-7 cells and megakaryocytes

Histone marks were examined in undifferentiated HPC-7 cells (Undiff) and megakaryocytes (Megs). H3K4 mono, di and trimethylation (H3K4me1,2,3 respectively), H3K9 trimethylation (H3K9me3), H4K8 acetylation (H4K8ac), H4K16 acetylation (H4K16ac) and H3K9acS10 phosphorylation (H3K9S10p) were examined.

a) Histone mark modification was examined at a protein level by Western blot in undifferentiated cells (Undiff) and megakaryocytes (Megs). The protein level was normalised with the H3 C-Terminal antibody.

b) Histones were normalised to the H3 C-Terminal. This chart displays histone mark abundance in megakaryocytes (Megs) relative to undifferentiated HPC-7 cells (Undiff), which are set at 1. Values are the mean of one or two experiments. Error bars represent the standard error of the mean.

(Figure 3.10a). According to this first experiment, only *Mll1* gene expression displayed a change between the cell types, showing a three-fold increase in megakaryocytes (Figure 3.9b). Otherwise, no significant differences were observed for the other *Mll* genes. These first results were confirmed by quantitative PCR (Real-time QRT-PCR) (Figure 3.10c). This result initially seems to be incompatible with the global changes in H3K4me3 during differentiation; this will be addressed in the following chapter.

3.8. Summary: Establishing a model of differentiation

In this model, TPO-induced differentiation of HPC-7 cells generates a clear morphologically distinct cell type: megakaryocytes. Their degree of maturity was validated by an increase in ploidy, and the detection of specific control genes, such as *Gp6*, which are exclusively expressed in mature megakaryocytes. Differentiation is associated with down regulation of MLL-target genes in the *Hoxa* gene cluster, as well as the key leukogenic gene *Meis1*. These changes are associated with only small changes in the global abundance of histone modifications, suggesting that targeted changes at individual loci, rather than genome-wide effects determine gene activity upon differentiation. Together, these observations suggest that this is an appropriate model to study regulation at MLL1 target genes.

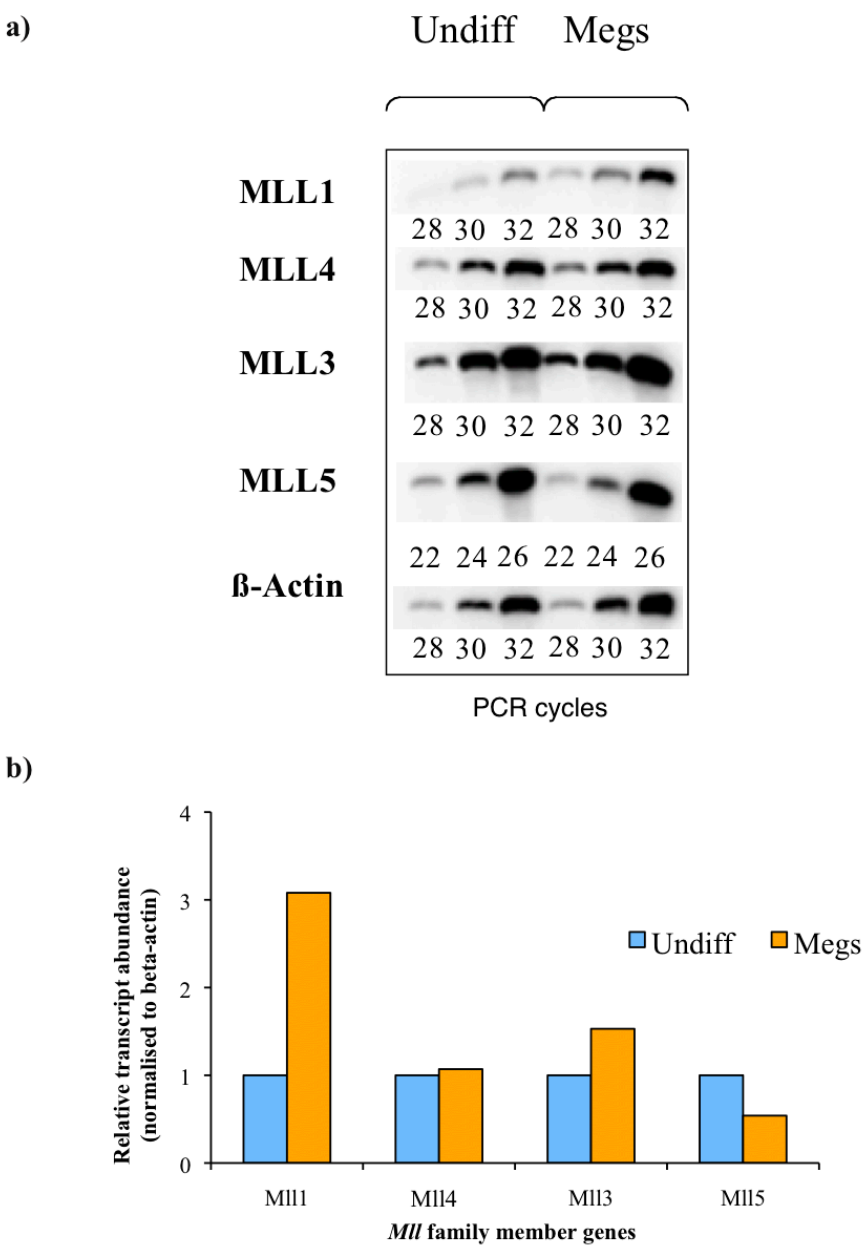


Figure 3.10: *Mll* family members transcript levels in undifferentiated HPC-7 cells and in megakaryocytes

a) Gene expression (*Mll1*, *Mll4* *Mll3*, *Mll5*) was quantified in undifferentiated cells (Undiff) and in megakaryocytes (Megs) by semi-quantitative PCR, by adding alpha-^{32}P in the reaction.

b) Abundance of *Mll* family members in undifferentiated cells (Undiff) and megakaryocytes (Megs). Mll transcript abundance is set at 1 in undifferentiated HPC-7 cells. The quantification was normalised using ß-actin gene expression as a control.

c)

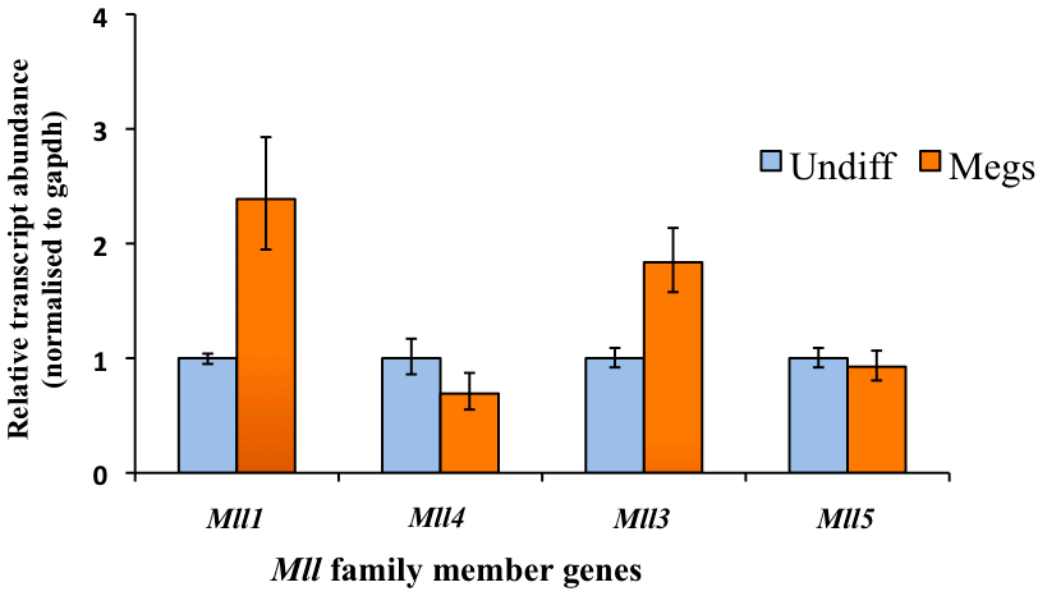


Figure 3.10 c: *Mll* family member gene transcript abundance in undifferentiated HPC-7 cells and in megakaryocytes

Quantification by real-time PCR of *Mll* family member transcript abundance (normalised with *Gapdh*). Transcript abundance in undifferentiated HPC-7 is set at 1.

4. CHARACTERISING HISTONE MODIFICATIONS ON MLL-TARGET GENES

In the previous chapter, a model system of differentiation was established to analyse MLL1 complex-driven changes on MLL target genes. It was established that genes of the *Hoxa* cluster, and the leukaemic gene *Meis1* are down regulated upon differentiation. Also, no global changes in histone modifications were observed, and previous results were unlikely to reflect changes in MLL1 abundance.

This chapter focussed on changes at the *Hoxa4* and *Hoxa5*. *Hoxa5* gene showed a substantial and reproducible fall in transcript abundance upon differentiation, suggesting an efficient silencing process. These genes are also implicated in haematopoiesis (Fuller *et al.,* 1999; Strathdee *et al*., 2006). Notably, *Hoxa4* was described to be the most 3' gene involved in haematopoiesis (Sauvageau *et al*., 1994). We examined epigenetic changes associated with specific *Hoxa* genes and control genes. This investigation focussed on number of activating and repressive marks, using Native chromatin immunoprecipitation (N-ChIP) in undifferentiated HPC-7 cells and Carrier chromatin immunoprecipitation (C-ChIP) in megakaryocytes, where material is more limiting. In the experiments performed, the housekeeping gene *Gapdh* was used as a control.

4.1. How is the *Hoxa* gene cluster globally regulated?

To assess the contribution of epigenetic regulations, histone modifications were examined on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells when the genes are active, and in megakaryocytes when they are repressed. Specific primer sets were designed in order to assess the distribution of epigenetic marks at specific sites along these genes (Figure 4.1).

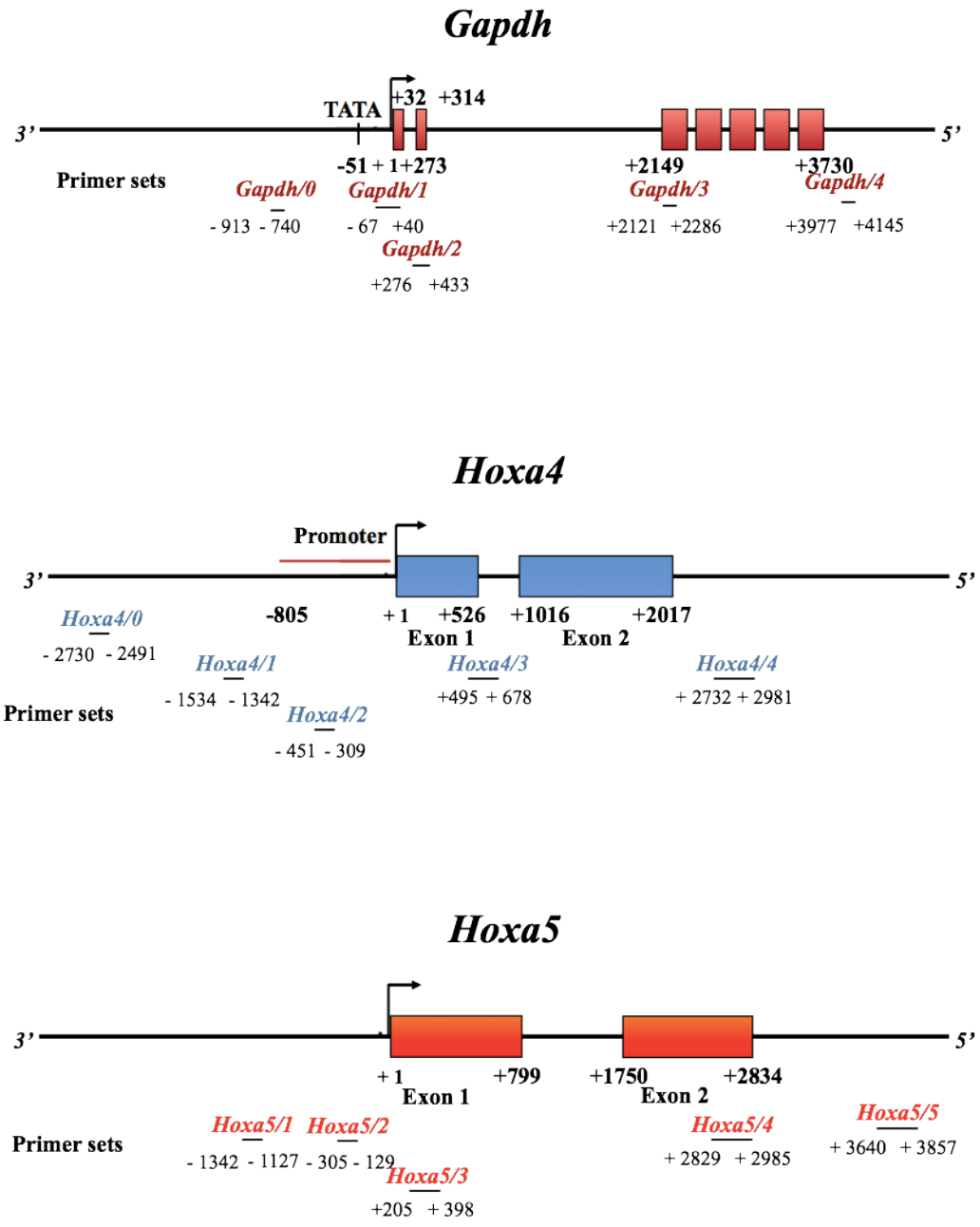


Figure 4.1: Representation of the distribution of primer sets used for PCR analysis over the *Gapdh*, *Hoxa4* and *Hoxa5* genes

The localisation of specific regions of the gene amplified by primer sets are represented below the name of each primer set. The arrow corresponds to the transcription start site (TSS), and boxes indicate exons. Promoter region is indicated with red bar and deduced from region of high homology in mammals. A TATA box could not be found on *Hoxa4* and *Hoxa5*.

Sites were located in the upstream "promoter" region of the genes, as well as at the transcription start sites (TSS), in some coding regions and downstream of these genes. The distribution of histone marks was assessed in undifferentiated HPC-7 cells using unfixed chromatin, performing Native Chromatin Immunoprecipitation (N-ChIP). However, the low number of megakaryocytes available (~10^5 cells) required a different procedure, in which the initial Native Chromatin Immunoprecipitation protocol was performed using *Drosophila* SL2 cells as a carrier. This method, developed by the Turner group (Birmingham University, UK) and called C-ChIP (O'Neill *et al*., 2006), was used on megakaryocytes. These two methods differ in the way of preparing the chromatin. Because of this first different step, it was important to assess whether the data obtained from these two distinct cell types, and by different procedures, could be compared quantitatively. In order to validate that these two methods can be compared, histone modifications were first examined on the control, housekeeping gene *Gapdh* and then compared in undifferentiated HPC-7 cells and megakaryocytes.

4.1.1. Examination of histone distributions on *Gapdh*

It is essential to notice that *Gapdh* transcript abundance does not change significantly (1% down/up regulation) between undifferentiated HPC-7 cells and megakaryocytes (Cf. Figure 3.3b and Supplementary data). This particular gene is therefore ideal to assess if the two methods can be compared.

Histone modifications associated with an active state of gene expression (H3K4me3), and one characterising a repressive aspect of gene expression (H3K27me3) were chosen to evaluate the results. The results are displayed as the amount of DNA bound, where a value corresponding to non-specific binding (derived from pre-immune binding) is subtracted. This

approach reflects the technical difficulties of adapting ChIP to real-time PCR. Whilst our group is typically using a Bound/Unbound ratio to quantify binding, we used this approach as we found the material in unbound C-ChIP fractions could not be quantified by real-time PCR. This may reflect the large excess of *Drosophila* DNA to mouse DNA, which generates a high background, but Real-Time PCR also appears to be more sensitive to this problem than conventional PCR. The locations of the primer sets used are presented above the graphs (Figure 4.2). Both undifferentiated HPC-7 cells (Figure 4.2, *Left upper panel*) and megakaryocytes (*Left lower panel*) display a peak of trimethylated H3K4 at the transcription start site (TSS) of *Gapdh,* consistent with numerous studies of this mark present at transcriptionally active genes (Margueron *et al*., 2005). These show similar levels of enhancement, but the peak in megakaryocytes is broader, with the mark also being detected upstream in megakaryocytes (*Gapdh/1*; Fig. 4.2) when it is absent in undifferentiated cells. This apparent broader peak of modification in megakaryocytes is likely to reflect the reduced extent of digestion and lower resolution of the chromatin obtained from these cells - we note that the *Gapdh/1* and *Gapdh/2* primers sets are separated by only ~200 base pairs. H3K27me3 was also monitored (Figure 4.2, *Right panels*). Its pattern along the *Gapdh* gene in megakaryocytes suggests that there are no significant changes during differentiation.

This experiment was important to establish that similar data could be obtained from what should be a similarly-regulated housekeeping gene in two cell types. This was first confirmed qualitatively, as similar patterns of distribution were observed in the two cell types. This is also consistent with what was previously described: peaks of active marks are typically associated with the "promoter" region (Liu *et al.*, 2005; Pan *et al.*, 2007). Furthermore, values from bound (specific – non specific) are quantitatively similar, suggesting that the two cell types can be compared using the two ChIP approaches. The difference of detection at

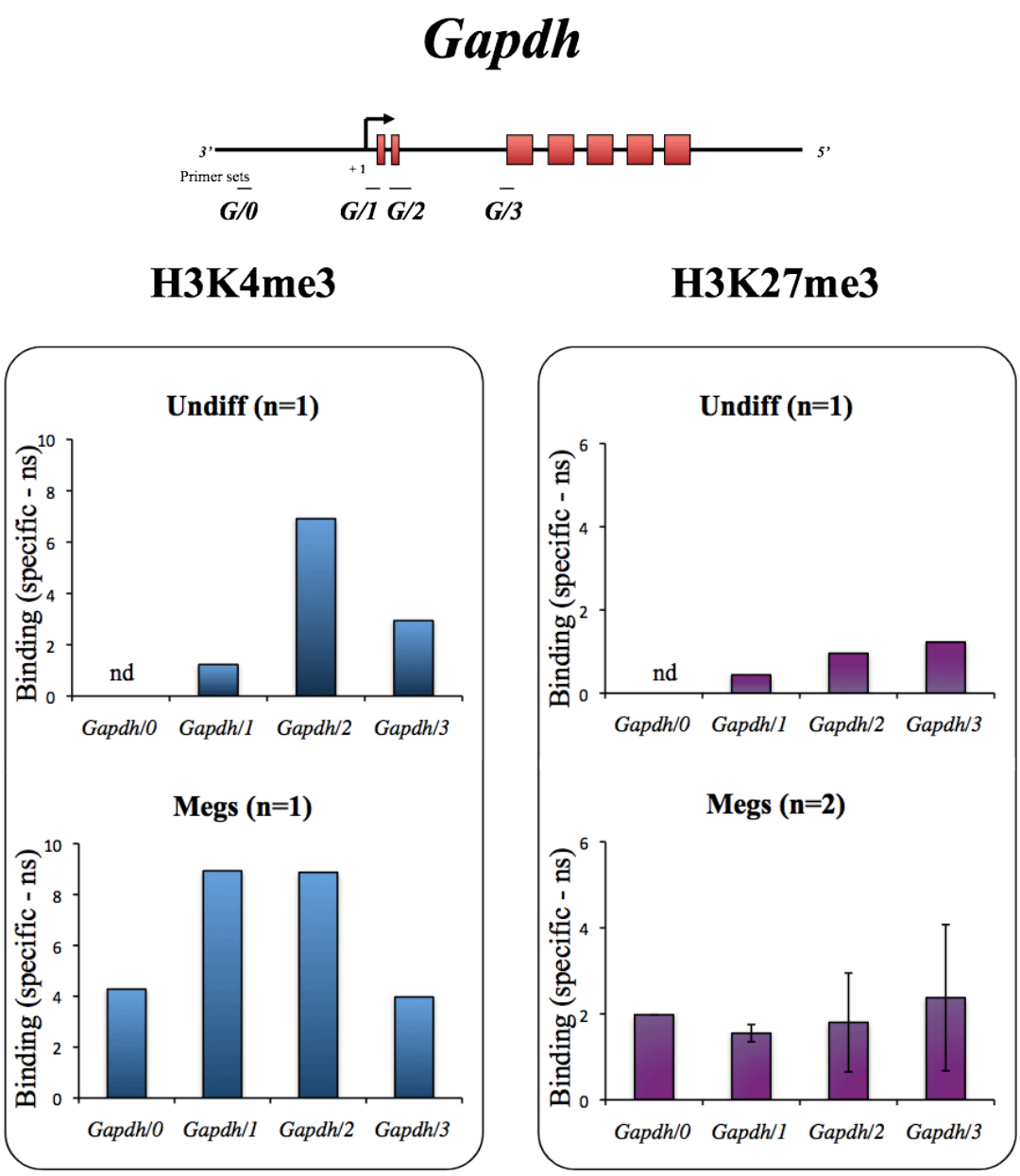


Figure 4.2: Distribution of H3K4me3 and H3K27me3 histone modifications on the *Gapdh* control gene in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibodies against **H3K4me3** and **H3K27me3**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Gapdh* gene and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

Gapdh1 with H3K4me3 suggests that chromatin from megakaryocytes may be lower resolution than from the undifferentiated cells.

4.1.2. What histone marks are up and down-regulated on *Hoxa4* and *Hoxa5* genes?

Once the legitimacy of comparing N-ChIP and C-ChIP data in these two cell types had been established, histone modifications on genes of interest (*Hoxa4* and *Hoxa5*) were then examined. In particular, two histone marks related to an active transcription state (H3K4me3, H3K9ac) and two marks corresponding to the repressive transcription state (H3K27me3, H3K9me2), were assessed for their distribution at defined sites along these genes.

H3K4me3 was first examined (Figure 4.3). In undifferentiated HPC-7 cells, a peak of this activating mark was observed close to the transcription start site of *Hoxa4* (Figure 4.3, *Left panels*), and shows a high abundance when compared to the regions surrounding this site. In contrast, there is a clear decrease (12-fold) of this histone enrichment in megakaryocytes. Similarly, an enrichment of H3K4me3 was observed over the transcription start site of *Hoxa5* (Figure 4.3, *Right panels*), which is then substantially decreased in megakaryocytes (20-fold decrease). *Hoxa5* displays a similar pattern of modification as *Hoxa4*: a substantial decrease of H3K4me3 at the transcription start of these genes in megakaryocytes. The reduction of H3K4me3 abundance in megakaryocytes suggests the requirement of a histone H3K4 demethylase, or histone turn-over.

Another "activating" mark, H3K9ac, was then examined (Figure 4.4). This particular histone mark is deposited by a number of HATs, including Gcn5 as part of the SAGA complex (Grant *et al.*, 1999), but it is also a mark of replication-coupled chromatin assembly (Kuo *et al*., 1996). Once again, a peak of acetylation is observed over the transcription start site of

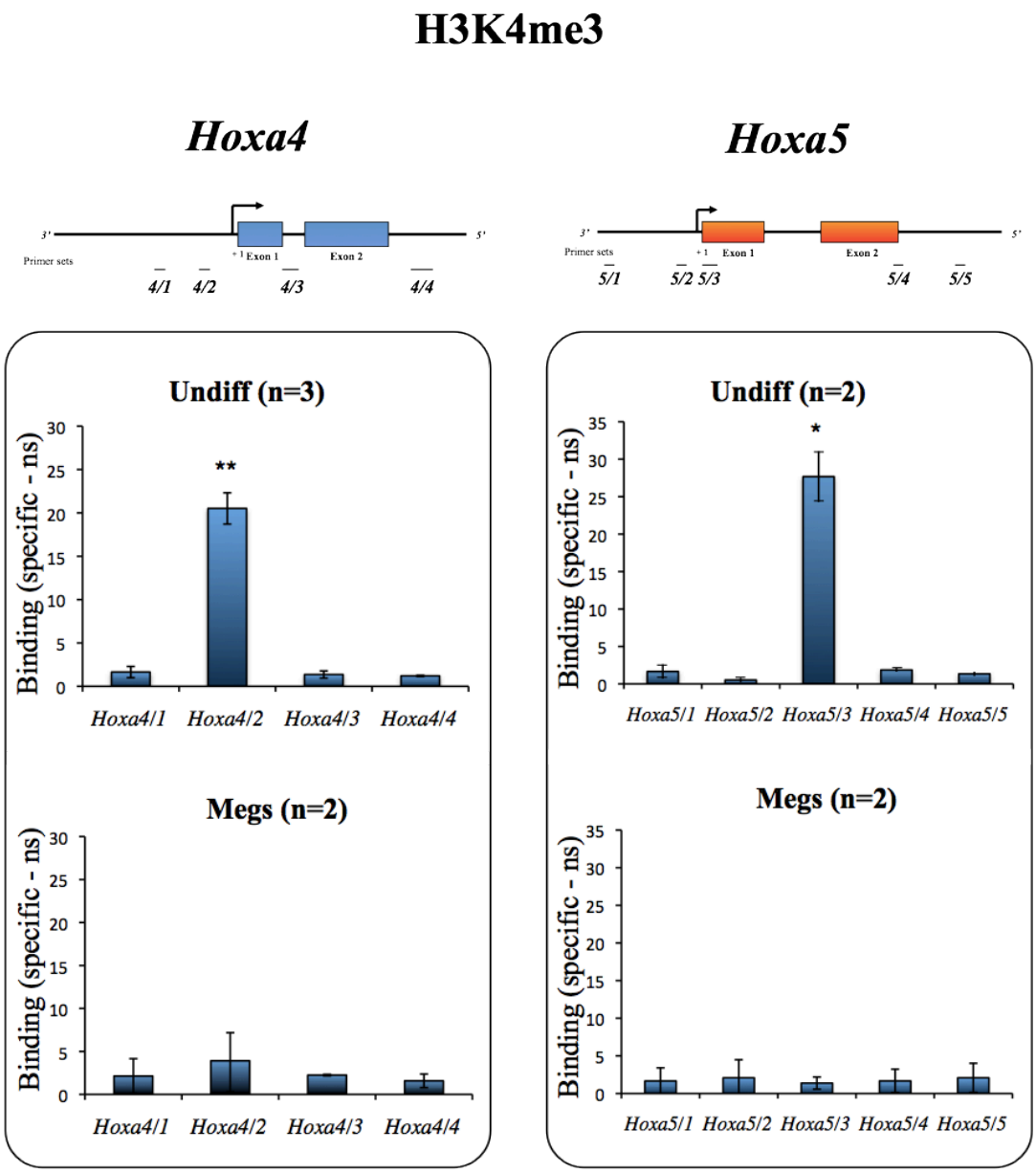


Figure 4.3: Distribution of H3K4me3 histone modifications on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against **H3K4me3**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from two or three ChIP experiments. Error bars represent the standard error of the mean. (**p<0.001, *p<0.02, Student test, comparing *Hoxa4/1* and *Hoxa4/2, Left panel.* Comparing *Hoxa5/2* and *Hoxa5/3, Right panel*).

Hoxa4 in undifferentiated cells (Figure 4.4, Left *Upper panel*). This enrichment is not as substantial as with H3K4me3 though, when compared with adjacent regions.

Interestingly, this mark spreads throughout the coding region of the gene after differentiation (Figure 4.4, *Lower panel*). Similarly, H3K9ac shows an enrichment at the TSS of *Hoxa5* in undifferentiated HPC-7 cells (Figure 4.4, *Right upper panel*). This is lost in megakaryocytes but we observed a broad enrichment of this mark throughout the locus (Figure 4.4 *Lower panel*). In contrast to H3K4me3, the distribution of H3K9ac is counter-intuitive, as this mark appears to increase in megakaryocytes with substantial increases in this mark in both the promoter and the coding regions. This is variable, as the increased errors bars indicate. An alternative interpretation of these data is that the distinct peak at the TSS seen in undifferentiated HPC-7 cells is lost, and replaced by a generalised distribution of the acetyl mark on silent genes.

The distribution of an additional "activating mark" was examined on *Hoxa4*: H4K16ac (Figure 4.5). H4K16ac is a mark deposited by another member of the MLL complex, the histone acetyl-transferase MOF. In undifferentiated cells, a slight enrichment of H4K16 acetylation is observed over the TSS of *Hoxa4* (*Hoxa4/2*) (Figure 4.5, *Upper panel*). This mark is reduced by about three-fold in megakaryocytes (Figure 4.5, *Lower panel*). This result reinforces the idea of the MLL complex involvement in the regulation of *Hoxa* genes.

Interestingly, all these "activating" marks examined display a similar pattern of distribution on *Hoxa4* in undifferentiated cells. So far, an enrichment of active histone marks, associated with the TSS of active genes, was observed in undifferentiated HPC-7 cells, which was substantially decreased after differentiation into megakaryocytes. This pattern of histone modification correlates with the global down-regulation of *Hoxa* gene expression during

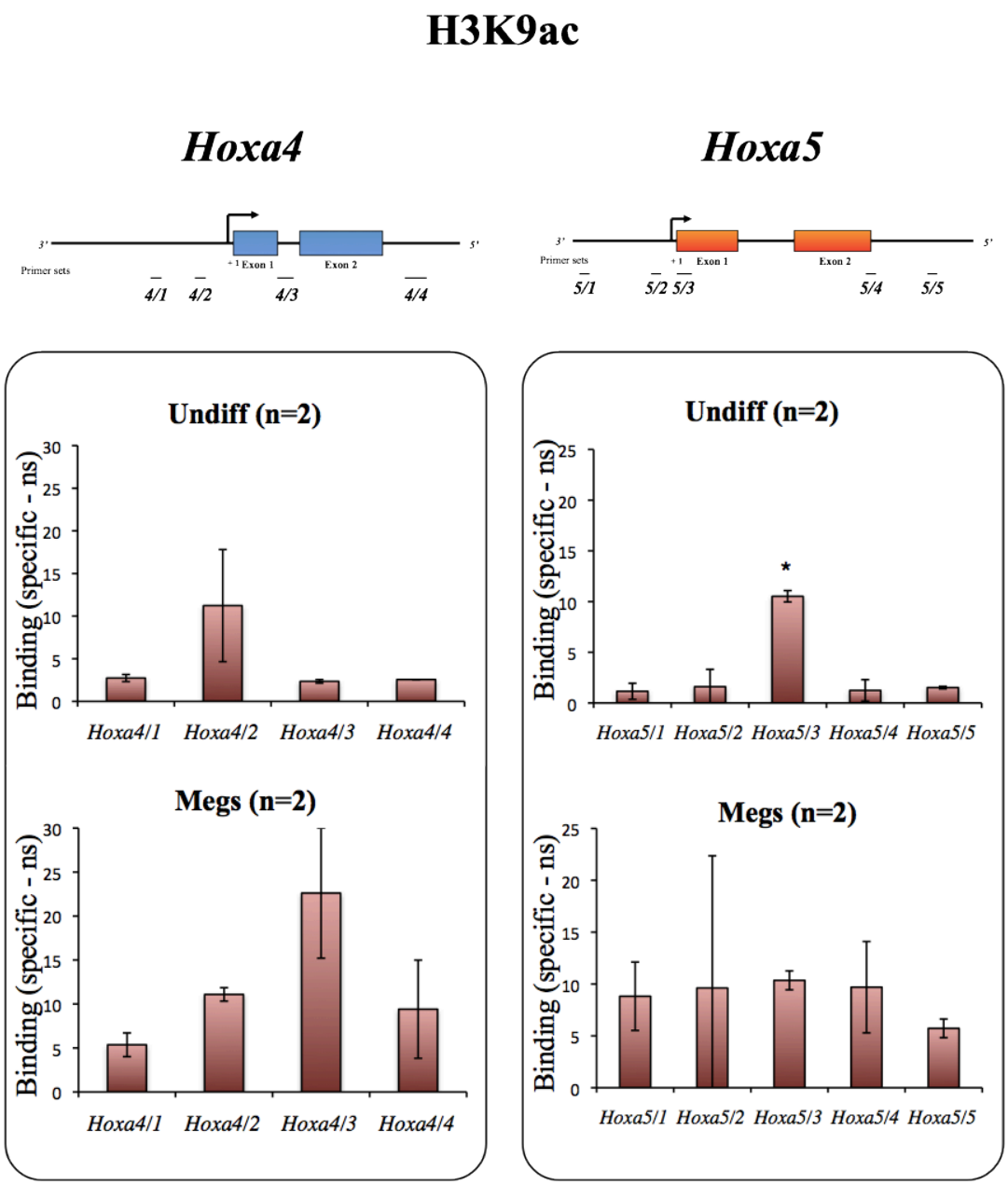


Figure 4.4: Distribution of H3K9ac histone modifications on the *Hoxa4* and *Hoxa5* in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against **H3K9ac**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from two ChIP experiments. Error bars represent the standard error of the mean. (*p<0.1, student test, comparing binding at *Hoxa5/2* and *Hoxa5/3*)

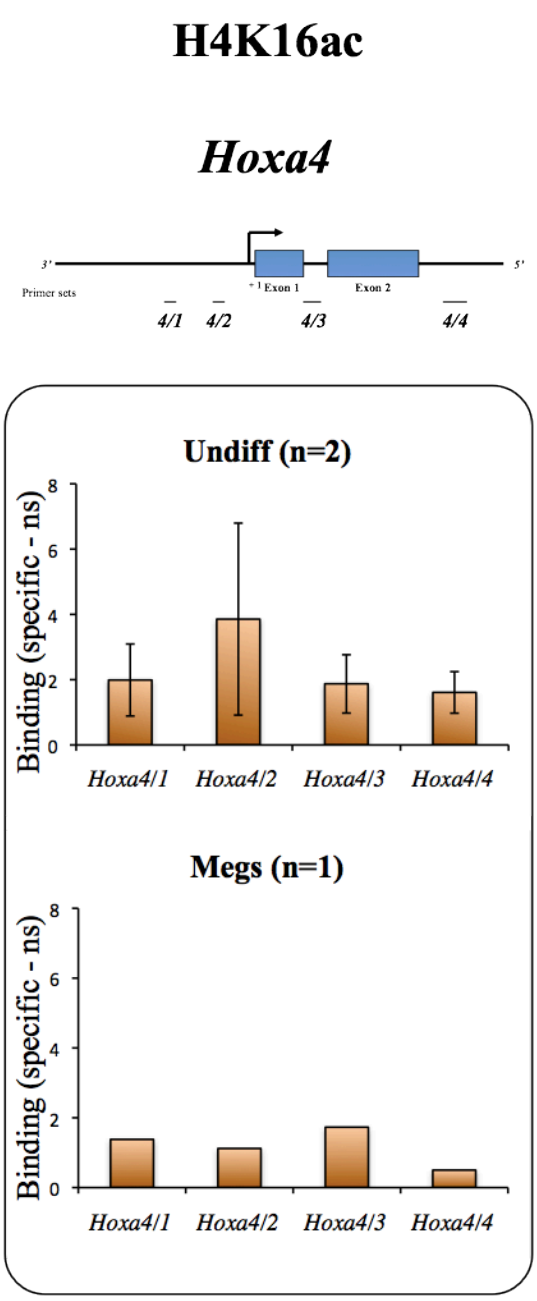


Figure 4.5: Distribution of H4K16ac histone modifications on the *Hoxa4* gene in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panel*) were differentiated into megakaryocytes (Megs, *Lower panel*). Chromatin was extracted and immunoprecipitated with antibody against **H4K16ac**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* gene and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

differentiation (i.e. Figure 3.3). We next asked whether "repressive" histone marks, typically associated with silent genes, show corresponding increases upon gene silencing.

In order to check repressive histone marks, the distribution of H3K27me3 was first assessed (Figure 4.6). In undifferentiated HPC-7 cells, this mark shows a low level, or "background" level, of abundance with no particular pattern on *Hoxa4* (Figure 4.6, *Left panels*). Moreover, no changes of H3K27 methylation were highlighted in megakaryocytes either. As with *Hoxa4*, *Hoxa5* displays similar "overall" pattern in undifferentiated cells and megakaryocytes (Figure 4.6, *Right panels*). Our trial to find a control gene in order to examine the presence of H3K27me3 on an inactive gene failed to detect enrichment for this mark.

The distribution of another histone mark associated with repression, H3K9me2, was also examined on *Hoxa4* (Figure 4.7). This shows a low level of enrichment (values ~2-3) in undifferentiated HPC-7 cells, where a slight but probably not significant enrichment is observed over both the promoter and the coding regions of the gene (Figure 4.7, *Upper panel*). As with H3K27me3, this distribution is not significantly altered in megakaryocytes (Figure 4.7, *Lower panel*). As a result, no change is observed with the repressive marks (H3K27me3 and H3K9me2) upon differentiation. This may reflect the low levels of these marks in both cell types.

The distribution of the association of two adjacent histone marks (H3K9acS10p) was a particular interest for this study. H3S10p is a marker of mitotic chromatin, while its neighbour H3K9ac characterises actively transcribed genes; however, the dual mark has also been shown to be associated with immediate early gene activation (Dyson *et al*., 2005). Importantly, in vitro studies indicate that phosphorylation of H3S10 prevents the methylation of H3K9. Moreover, it was also reported that H3S10 phosphorylation coupled

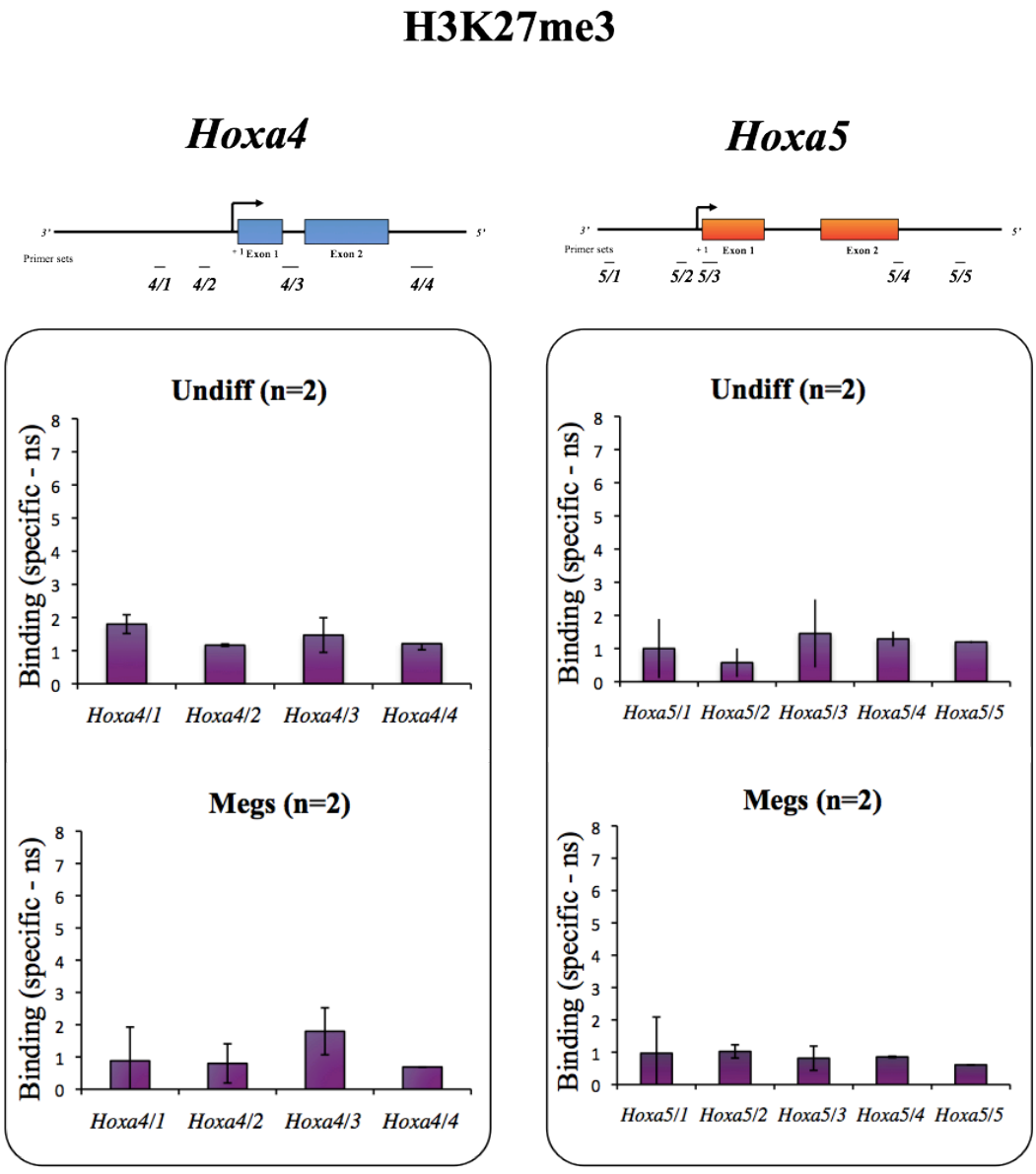


Figure 4.6: Distribution of H3K27me3 histone modifications on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against **H3K27me3**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from two ChIP experiments. Error bars represent the standard error of the mean.

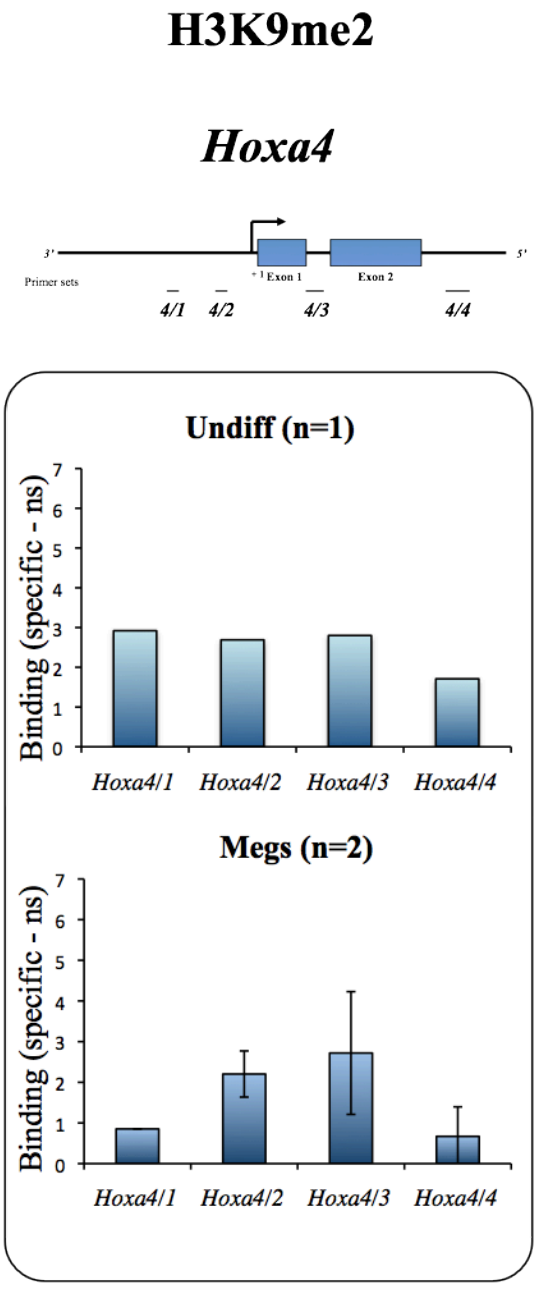


Figure 4.7: Distribution of H3K9me2 histone modifications on the *Hoxa4* gene in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panel*) were differentiated into megakaryocytes (Megs, *Lower panel*). Chromatin was extracted and immunoprecipitated with antibody against **H3K9me2**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* gene and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

with H3K9 acetylation leads to the activation of genes (Jenuwein and Allis, 2001; Li, Hall *et al.*, 2002). Previous in vitro studies with the MLL1 SET domain show that the combination of H3K9ac and H3S10p stimulates MLL1 SET domain histone methyl-transferase activity *in vitro* (Nightingale *et al*., 2007). It was of interest if this in vitro observation had in vivo consequences for MLL1 regulation. The phosphorylation of H3S10 seems to be involved in the regulation of genes through its association with other marks. Could therefore the dual modification, H3K9acS10p, contribute to the regulation of the *Hoxa* genes in this model?

4.2. Is H3S10p contributing to the regulation of the *Hoxa* genes?

To address this question, the distribution of phosphorylated H3S10 associated with the acetylated H3K9 was examined using an antibody recognising the dual mark, H3K9acS10p, in undifferentiated HPC-7 cells and in megakaryocytes (Figure 4.8). Examination of the distribution of this association of epigenetic marks on *Hoxa4* (Figure 4.8, *Left panels*) in undifferentiated HPC-7 cells, shows a clear enrichment over the transcription start site, correlating with the peaks of H3K4me3 and H3K9ac. This peak is removed in megakaryocytes. Similarly, an analogous pattern is seen on *Hoxa5* (Figure 4.8, *Right panels*): an enrichment of this dual mark over the transcription start site in undifferentiated HPC-7 cells, also followed by a four-fold reduction in megakaryocytes. It should be noted that a general increase of H3S10p was previously observed in megakaryocytes (i.e. Figure 3.8). Localised *Hoxa*-specific decrease in H3S10p suggests that proteins with phosphatase activity contribute to *Hoxa* regulation.

Moreover, the distribution of this phosphorylated mark follows the distribution of the H3K4 methylation, in undifferentiated HPC-7 cells and in megakaryocytes. This suggests that these

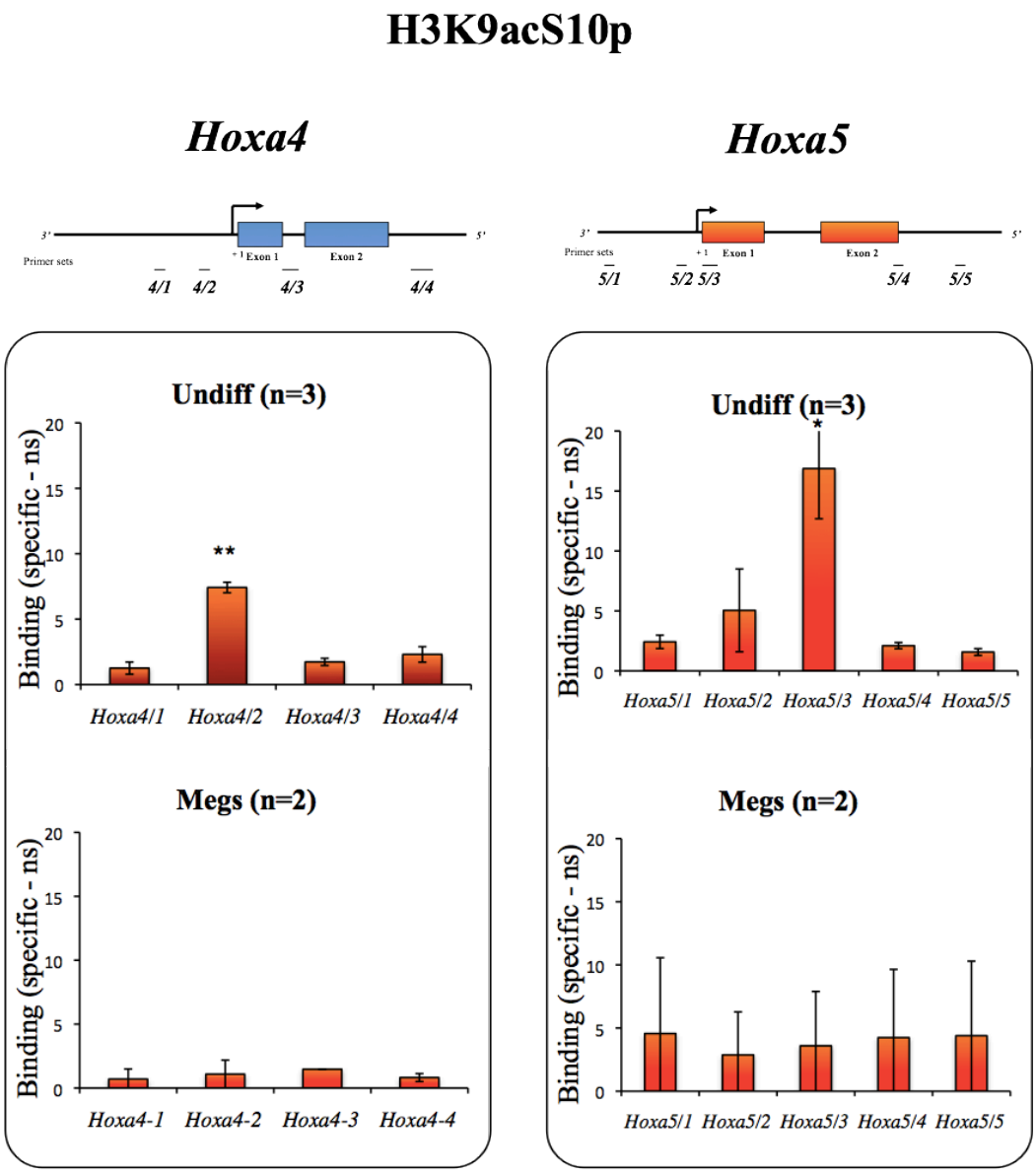


Figure 4.8: Distribution of H3K9acS10p histone modifications on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against the dual **H3K9acS10p**. The graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from two or three ChIP experiments. Error bars represent the standard error of the mean. (**p<0.001, *p<0.02, student test, comparing binding at *Hoxa4/1* and *Hoax4/2 Left panel*. Comparing *Hoxa5/2* and *Hoxa5/3*, *Right panel*).

marks may be related, though it is important to note that the apparent correlation between these two marks does not necessarily imply that the enzymes or processes that drive these changes are linked. This is consistent with our knowledge of the MLL complex, which has no known histone kinase subunit.

This suggests novel kinase and phosphatase proteins (or kinase and phosphatase complexes) take part in the regulation of *Hoxa* genes through their histone modification activity (Figure 4.9). The identity of the kinase protein will be discussed in the following chapter.

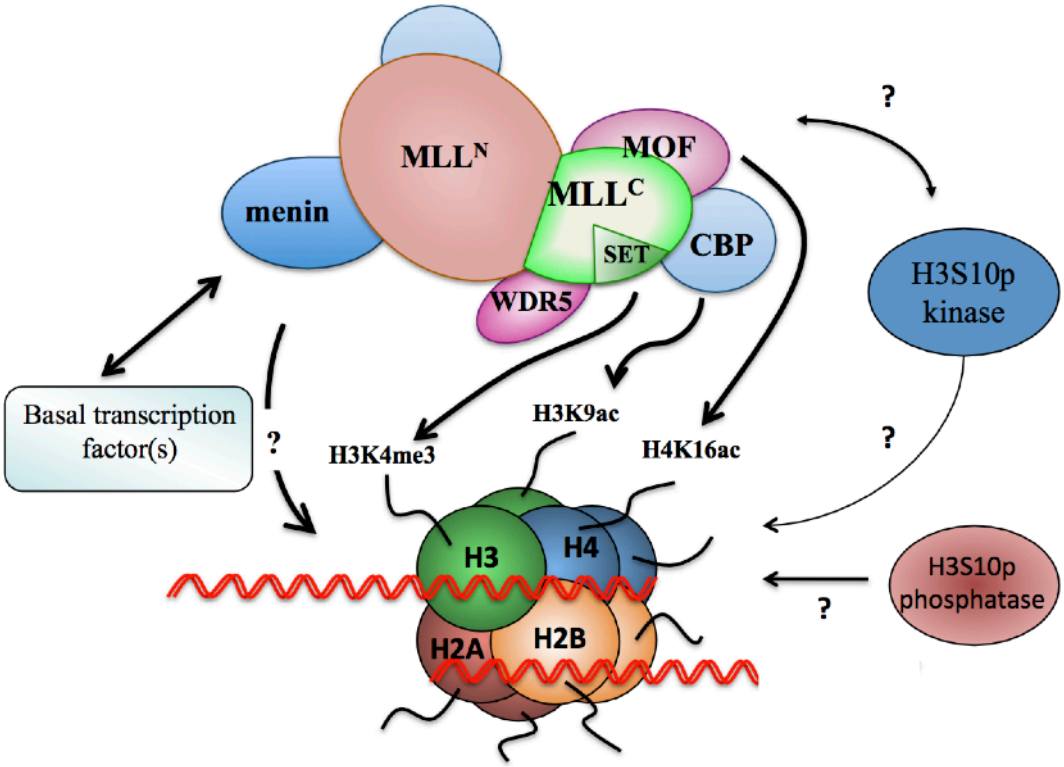


Figure 4.9: Suggested model of interaction

Other proteins were proposed to interact with the MLL complex in order to participate in the regulation of *Hoxa* genes during differentiation: a histone H3S10p-specific kinase, and a histone phosphatase

5. HOW ARE HISTONE MARKS REGULATED?

Specific enzymes deposit histone modifications. Histone methylation is deposited by histone methyl-transferases on lysines and arginines, whereas histone acetylation is deposited by histone acetyl-transferases on lysines. These modifications are typically recognised by "effector proteins", which induce functional changes in chromatin and are associated with regulation of gene expression. In this chapter we focus on the histone methyl-transferase MLL, which is involved in the regulation of *Hox* gene expression during haematopoiesis (Ernst *et al*., 2004b). As described earlier, *Mll1* is expressed in undifferentiated HPC-7 cells and in megakaryocytes.

5.1. Global abundance of MLL complex proteins upon differentiation

As differentiation is associated with the down-regulation of several MLL-target genes, (Chapter 4), we initially wondered whether this was driven by the down-regulation of proteins of the MLL complex. MLL1 is part of a multi-protein complex, including the DNA binding protein menin, and other proteins such as transcriptional co-activating factors and enzymes involved in the modification of histone marks. As the enzyme of interest, MLL abundance was first assessed in undifferentiated HPC-7 cells and in megakaryocytes (Figure 5.1a), using an antibody raised against the C-terminal part of MLL1 (*Upstate*). The protein loading was normalised with ß-actin. It was observed that MLL1 abundance is similar in both HPC-7 cells and megakaryocytes. This is perhaps a surprise, given the substantial changes between these two cell types, especially the increase in ploidy. Could then MLL1 still be present in the cells, but relocated, regulated by another protein in the MLL complex? As a component of the MLL complex, menin is a DNA binding protein, presumed to be involved in the recruitment of the MLL complex to the DNA. As with MLL1, menin abundance was

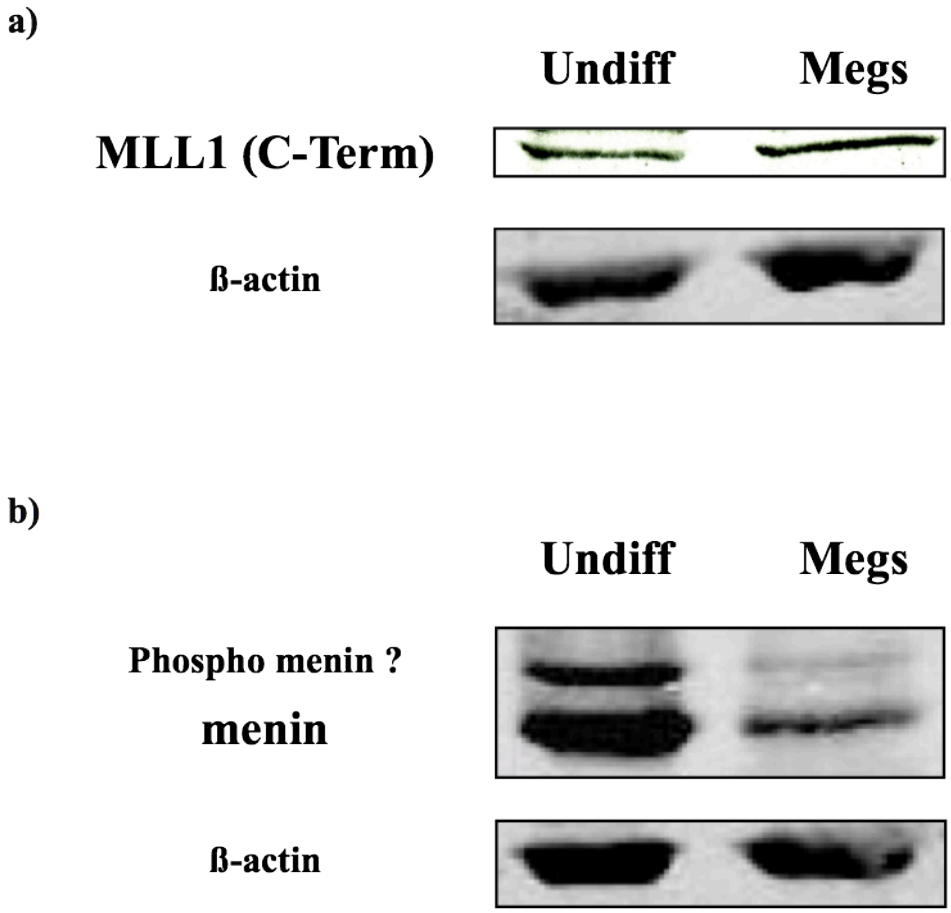


Figure 5.1: Protein quantification of (a) MLL1 and (b) menin in undifferentiated HPC-7 cells and megakaryocytes

MLL1 and menin abundance was assessed by Western blotting with an anti-MLL1-C terminal antibody (*UpState*), and an anti-menin antibody (*AbCam*), in undifferentiated HPC-7 (Undiff) and in megakaryocytes (Megs). Equal loading was assessed using ß-actin.

assessed by Western blotting using a menin-specific antibody (*AbCam*) (Figure 5.1b). Interestingly, there is a substantial drop in the abundance of menin in megakaryocytes. This reduction is also observed for what is assumed to be phosphorylated menin (*Upper band*). Since menin is involved in the recruitment of the MLL complex to the DNA, the reduction of the abundance of menin could suggest reduced MLL1 recruitment.

This suggests a possible mechanism for the relocation of the MLL complex during differentiation. However this is a global change in protein, we were interested in similar events occurring at individual MLL-target genes.

5.2. Is there a relationship between histone marks and proteins of the MLL complex?

In order to examine the relationship between the distribution of histone marks and the presence of proteins of the MLL complex, the distribution of some components of the MLL complex was assessed on *Hoxa4* and *Hoxa5* in both undifferentiated HPC-7 cells and megakaryocytes. This investigation was performed using chromatin immunoprecipitation on formaldehyde cross-linked chromatin (X-ChIP) in both cell types. The proteins analysed were MLL1, menin, the MLL complex component histone acetyl-transferase CBP as well as the complex RNA Polymerase II.

The results displayed correspond to the ratio of the amount of DNA in bound fractions and normalised to the input (amount of DNA in the bound fraction performed with the pre-immune IgG). The distribution of the primer sets along the gene are represented on a picture above the graphs.

5.2.1. Distribution of MLL1 binding on *Hoxa* genes

The distribution of specific proteins of the MLL complex was analysed over the *Hoxa4* and *Hoxa5* genes, and compared with the deposition of histone marks (Figure 5.2). MLL1 was then the first protein examined. The antibody used to assess MLL1 distribution over these genes targets the C-terminal part of MLL1 (*Upstate/Millipore*). On *Hoxa4* (Figure 5.2, *Left panels*), an enrichment of MLL1 was observed upstream of the transcription at start site of the gene (*Hoxa4/1* primer set) in undifferentiated HPC-7 cells (*Upper panel*). Surprisingly, this location is upstream of the site of enrichment of H3K4me3 (Cf. Figure 4.3). There is a 18-fold drop of this enrichment in megakaryocytes (*Lower panel*), consistent with the reduction in H3K4me3 at this gene after differentiation. Similarly, on the *Hoxa5* gene (Figure 5.2, *Right panels*), an enrichment of the C-terminal part of MLL1 is detected in undifferentiated HPC-7 cells, again just upstream of the transcription start site (*Upper panel*) (*Hoxa5/2*). Again, it is noticeable that this enrichment is situated upstream of the enrichment of H3K4me3 (*Hoxa5/3*), and is lost in megakaryocytes when the gene is silenced. Furthermore, it should be noticed that the enrichment detected over the *Hoxa4* gene is larger and more variable than that seen on *Hoxa5*: this may reflect that the primer sets used to characterise this enrichment on *Hoxa5* (*Hoxa5/2,* at -129 bp) are closer to the transcription start site than the primer sets used on *Hoxa4* (*Hoxa4/1*, at -1,342 bp). However, the peak of H3K4me3 is at the TSS on both genes.

In order to examine the apparent inconsistency between the site of MLL1 binding, and the mark that it deposits, the distribution of MLL1 on *Hoxa4* was assessed using an alternative antibody against the N-terminal of MLL1, in undifferentiated HPC-7 cells and megakaryocytes (Figure 5.3) (thanks to L. Ringrose's group, Vienna, Austria). This indicates that the N-terminal domain displays a similar pattern of enrichment to that observed with the

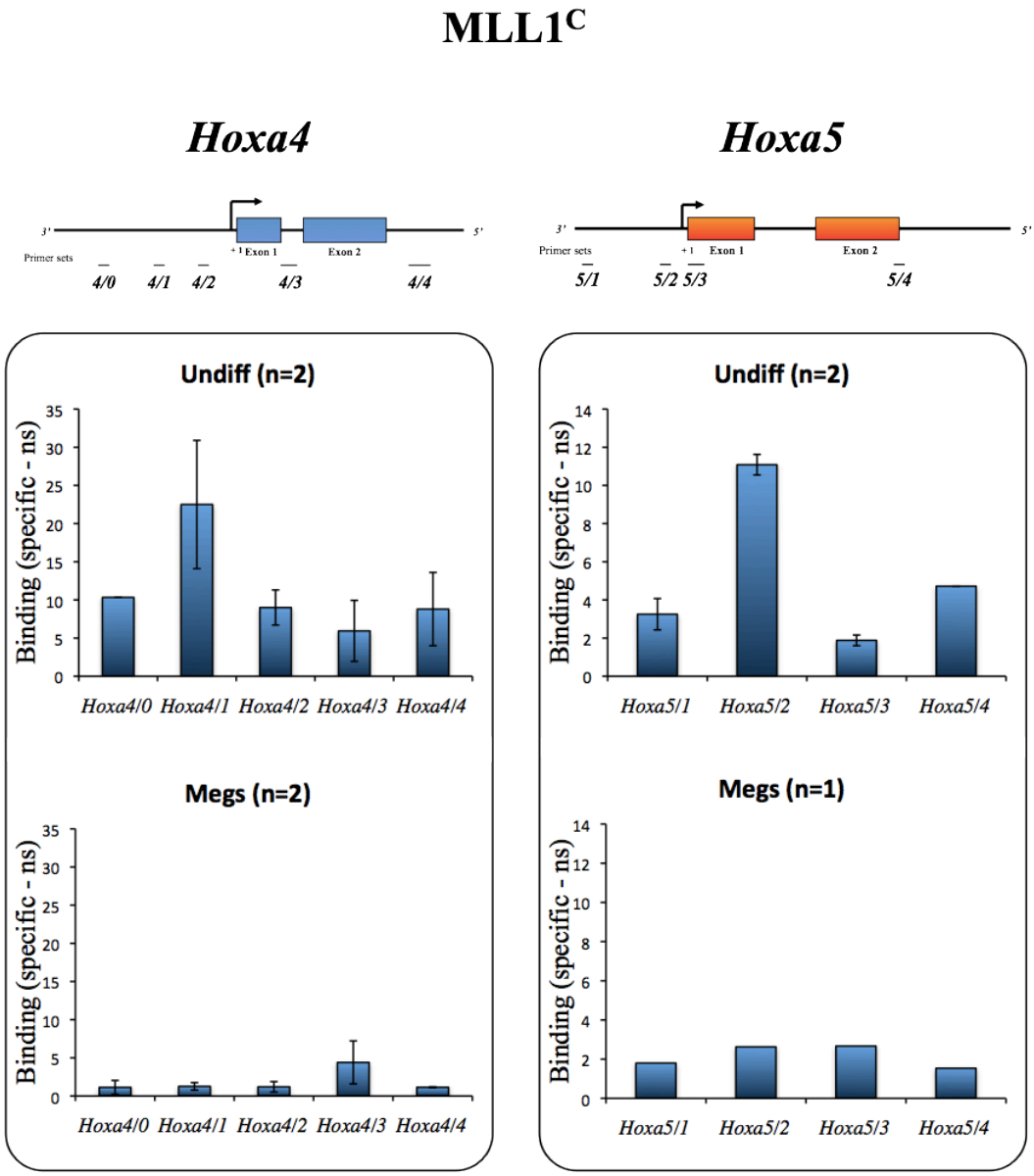


Figure 5.2: Distribution of Histone Methyl Transferase MLL1 (C-Term) on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against the C-terminal part of MLL1. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (obtained from pre-immune bound) is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes, and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

C-terminal domain on *Hoxa4* in undifferentiated HPC-7 cells (*Upper panel*). The peak over *Hoxa4/1* is then lost in megakaryocytes (*Lower panel*), as with the C-terminal domain of MLL1 (Cf. Figure 5.2). This confirms that the two domains of MLL1 protein are present and recruited to these genes, and that the MLL complex does occupy a promoter, rather than TSS-adjacent site.

5.2.2. Distribution of menin

As a key factor in the recruitment of the MLL complex, the distribution of menin was subsequently examined on *Hoxa4* and *Hoxa5* (Figure 5.4). An enrichment of menin binding was observed upstream of the TSS on *Hoxa4* in undifferentiated HPC-7 cells (Figure 5.4, *Left panels*), similar to the one observed with the C and N-terminal domains of MLL1 (Cf. Figures 5.2 and 5.3). This correlates with the idea that the MLL complex is recruited to the DNA along with menin. Moreover, there is a loss of menin binding in megakaryocytes (*Lower panel*). As shown previously, there is a global reduction in menin abundance upon differentiation, whereas the global MLL1 abundance displayed no changes. But these X-ChIP results show that the presence of MLL1 on the *Hoxa4* gene decreases, in parallel with a reduction in menin binding at the same site (Cf. Figures 3.4, 3.5 and 3.6). This confirms earlier reports that MLL1 and menin act as part of the same complex, both binding the same sites on the *Hoxa4* gene. This could suggest that in megakaryocytes, the down-regulation of menin impacts on MLL1 function, which could be either relocated to different loci in the cell after differentiation, or become functionally inactivated, despite being maintained at pre-differentiation levels. Looking at the distribution of menin on the *Hoxa5* gene (Figure 5.4, *Right panels*), similar observations were made: there is an enrichment of menin upstream to the gene (*Hoxa5/2*) (*Upper panel*), which is lost in megakaryocytes (*Lower panel*). Interestingly, the distribution of menin on both *Hoxa* genes follows that of the C-terminal

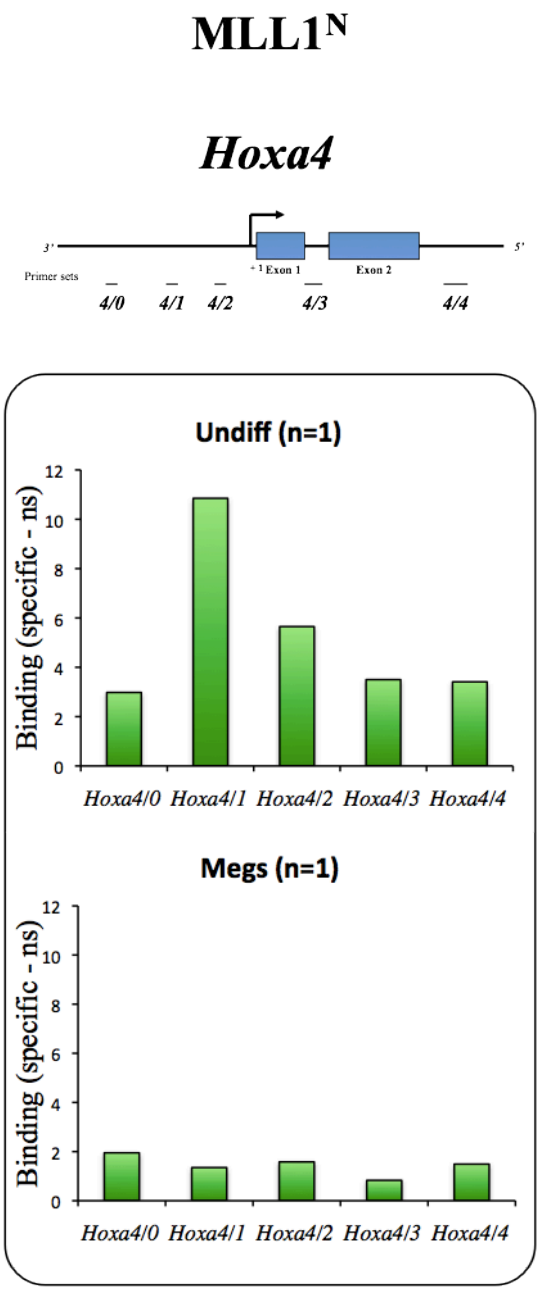


Figure 5.3: Distribution of Histone Methyl Transferase MLL1 (N-Term) on the *Hoxa4* gene in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panel*) were differentiated into megakaryocytes (Megs, *Lower panel*). Chromatin was extracted and immunoprecipitated with antibody against the N-terminal part of MLL1. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* gene, and was performed in triplicate from one ChIP experiment.

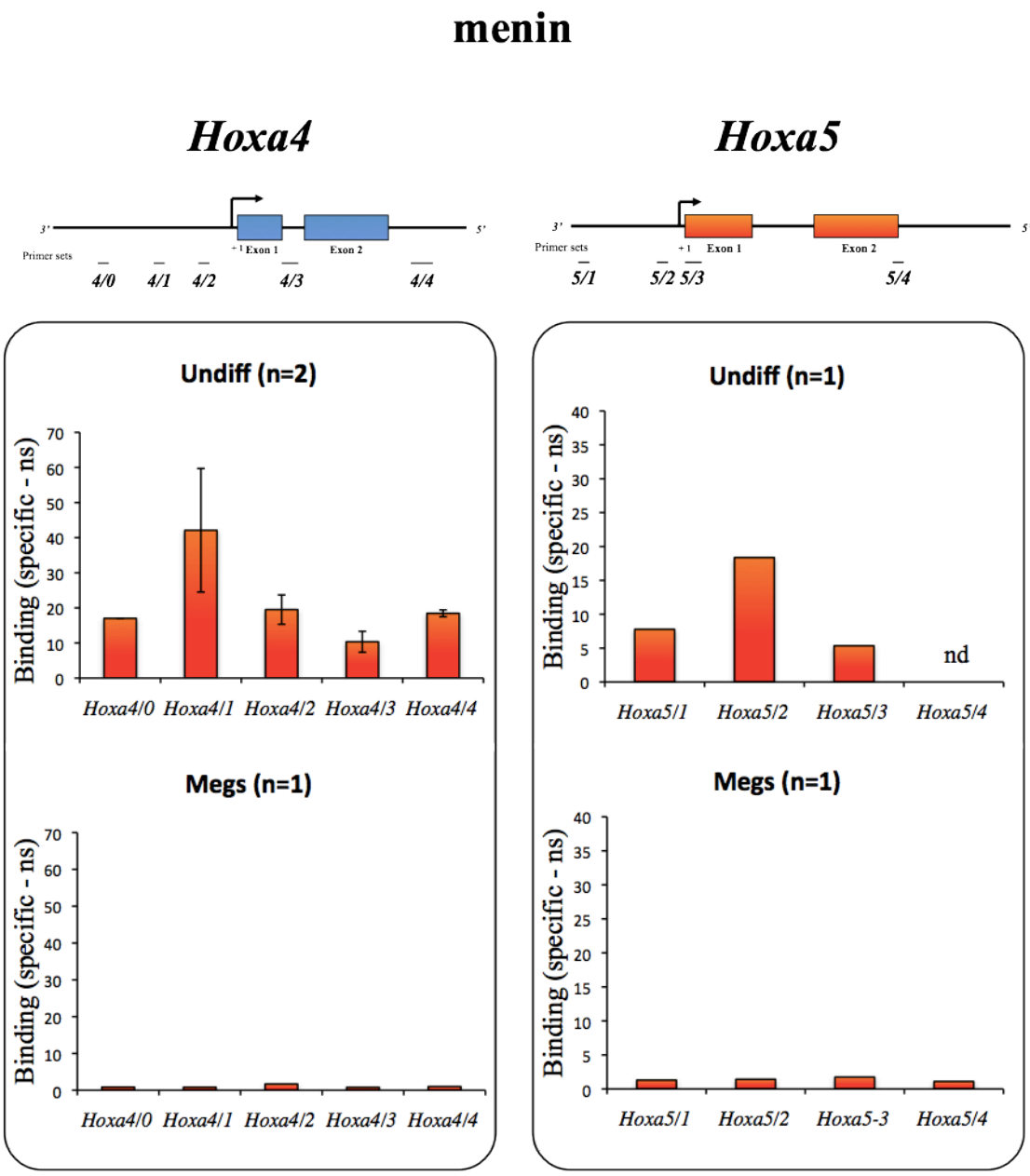


Figure 5.4: Distribution of menin on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against menin. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes, and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

part of MLL1. Furthermore, these sites are detected upstream of the TSS of the gene (*Hoxa4/1* and *Hoxa5/2*), whilst a peak of H3K4 methylation was observed over the TSS (Cf. Figure 4.3: *Hoxa4/2* and *Hoxa5/3*).

5.2.3. Distribution of a specific histone acetyl-transferase: CBP

Previous studies indicate that transcriptional co-activators, like CBP and MOF, may modify components of the basal machinery or Polymerase II, as well as targeting chromatin. However, recent studies also indicate that both histone acetyl-transferases are members of the MLL complex. Therefore, it was first investigated how another enzyme belonging to the MLL complex, CBP (CREB Binding Protein), was recruited along these genes, in undifferentiated HPC-7 cells (Figure 5.5). CBP is a histone acetyl transferase involved in the acetylation of a number of residues. Assessment was performed by X-ChIP, as with MLL1 and menin. It was first observed that the *Hoxa4* gene is enriched with CBP, in the coding region of the gene (*Hoxa4/3*) (Figure 5.5, *Left panel*). Although, it is striking that the CBP distribution does not reflect the distribution of H3K9ac in undifferentiated cells, it does reflect the distribution in megakaryocytes (Cf. Figure 4.4). Thus, unlike MLL1 and menin, CBP appears to be located in a more downstream region. On the other hand, CBP is enriched over the TSS on *Hoxa5* (Figure 5.5, *Right panel*) in undifferentiated HPC-7 cells. This enrichment is slightly weaker than on *Hoxa4* though. Moreover, CBP is also observed to bind on the downstream region of the coding part of the gene. It is clear that CBP distribution warrants further investigation.

5.2.4. Distribution of RNA Polymerase II

Finally, the distribution of the RNA Polymerase II was assessed on the *Hoxa4* and *Hoxa5* genes (Figure 5.6). On *Hoxa4* (Figure 5.6, *Left panels*), in undifferentiated HPC-7 cells

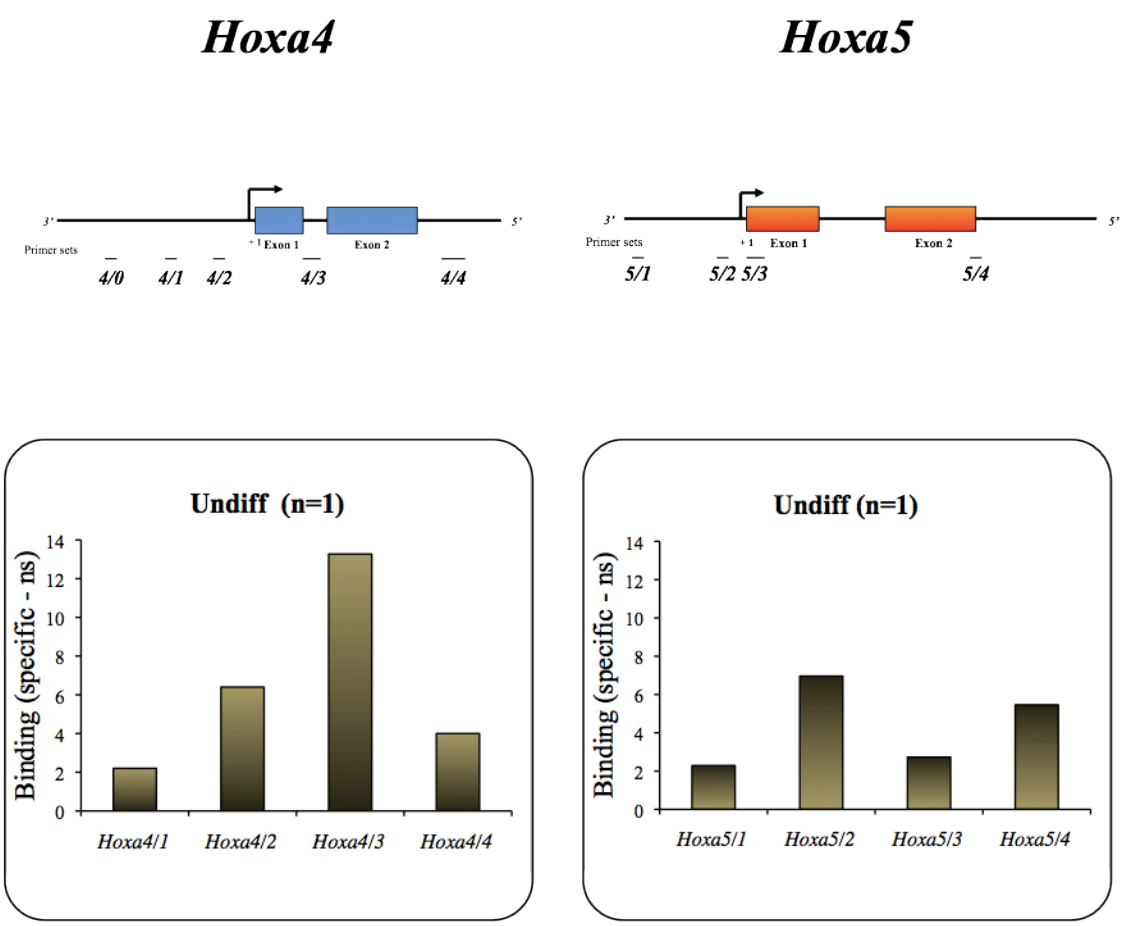


Figure 5.5: Distribution of Histone Acetyl Transferase CBP on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells

Chromatin of undifferentiated HPC-7 cells (Undiff) was extracted in undifferentiated cells, and then immunoprecipitated with antibody against CBP. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes, and was performed in triplicate from one ChIP experiment.

(*Upper panel*), an enrichment of RNA Polymerase II binding was observed over the TSS (*Hoxa4/2*). Thus, RNA Polymerase II appears to bind on, or near to the site carrying a high level of H3K4 methylation. Surprisingly, this protein remains bound on this same site in megakaryocytes (*Lower panels*), despite *Hoxa4* gene expression being substantially reduced in megakaryocytes. As with *Hoxa4*, RNA Polymerase II is present on *Hoxa5* (Figure 5.6, *Right panels*), with a peak of enrichment upstream to the TSS of the gene (*Hoxa5/2*) in undifferentiated HPC-7 cells (*Upper panel*), but in megakaryocytes (*Lower panel*), binding decreases, and the peak shifts some 100bp downstream (*Hoxa5/3*). This might reflect the down-regulation of this gene in megakaryocytes (characterised by a reduction of *Hoxa5* transcript abundance). However, the presence of RNA Polymerase II on both *Hoxa4* and *Hoxa5* genes in megakaryocytes, whilst the abundance of the transcripts of these genes is decreased, could also reflect RNA Polymerase promoter-proximal pausing, a mechanism seen on silenced genes (Strobl and Eick, 1992).

5.2.5. Could another histone methyl-transferase be involved in *Hoxa* regulation?

These results indicate that active *Hoxa* genes are associated with an enrichment of H3K4 methylation, which associates with a peak of MLL1 at an adjacent site. However, MLL1 is not the only enzyme with the binding capacity to generate this mark (H3K4me3). Indeed, mouse and human MLL1 possess a close homologue, MLL4, which can also trimethylate H3K4. Mouse *Mll1* and *Mll4* genes are orthologues (Glaser *et al*., 2006; Bach *et al*., 2009) and share a high degree of homology, contain very similar SET domains (Glaser *et al*., 2006), and associate with the same proteins including WDR5, ASH2, RbBP5 and menin (Eissenberg *et al.* 2009). The distribution of MLL4 was then assessed using an antibody (Ab39267) raised against the C-terminal domain of human MLL4 (also called MLL2), the homologue of the *Drosophila* trithorax gene (Huntsman *et al*., 1999), and that we will call "MLL4". The

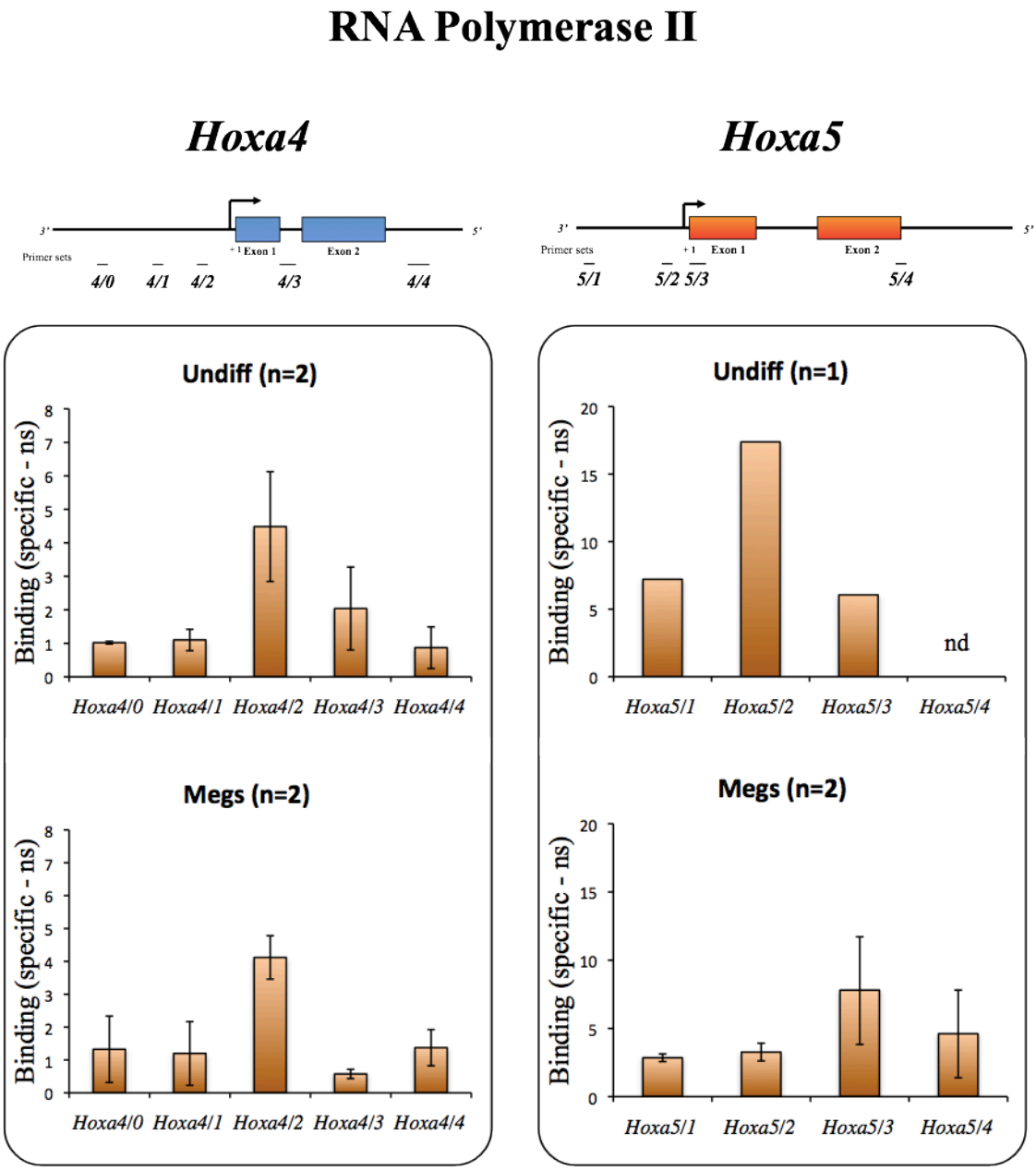


Figure 5.6: Distribution of the RNA Polymerase II on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against the C-terminal part of RNA Polymerase II. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes, and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

distribution of MLL4 was examined on *Hoxa4* and *Hoxa5* in undifferentiated HPC-7 cells (Figure 5.7), and showed an enrichment of this protein over the TSS of *Hoxa4* (Figure 5.7, *Left panel*), correlating with an enrichment of H3K4 methylation at the same site. Thus, MLL4 occupies a site just downstream of MLL1, though it seems unlikely that these two large complexes co-occupy the same promoter. On *Hoxa5* (Figure 5.7, *Right panel*), an enrichment of MLL4 is observed over the transcription start site of the gene, as with MLL1. However, MLL4 is also present at a same level within the body of the gene, raising the question of whether both complexes can bind simultaneously, how they are recruited to the promoter, and what the functional consequences of the two complexes are.

The presence of these two histone methyl-transferases at different sites within the *Hoxa4* gene suggests that the MLL1 and MLL4 antibodies recognise distinct proteins. However, MLL1 and MLL4 are highly similar, so we checked whether the corresponding antibodies were specific for their target proteins. This was assessed by Western blotting using both anti-MLL1 and anti-MLL4 antibodies to probe whole cell extract: in both case, bands of distinct size were detected, corresponding either to $MLL1^{C}$ (MW=180kDa) or MLL4 (MW=293kDa). The blot performed with the MLL4 antibody shows multiple proteins in addition to the one of expected size though (Data not shown), suggesting it is at least partially non-specific.

In the previous chapter, it was highlighted that another two histone modifying enzymes could also be involved in the regulation of *Hoxa* genes, through the phosphorylation and dephosphorylation of H3S10. Indeed, it has been shown that the *Hoxa4* and *Hoxa5* promoters are enriched with H3S10 phosphorylation in undifferentiated cells. This pattern of distribution mirrors that of H3K4 methylation and is also reduced in megakaryocytes. Only a

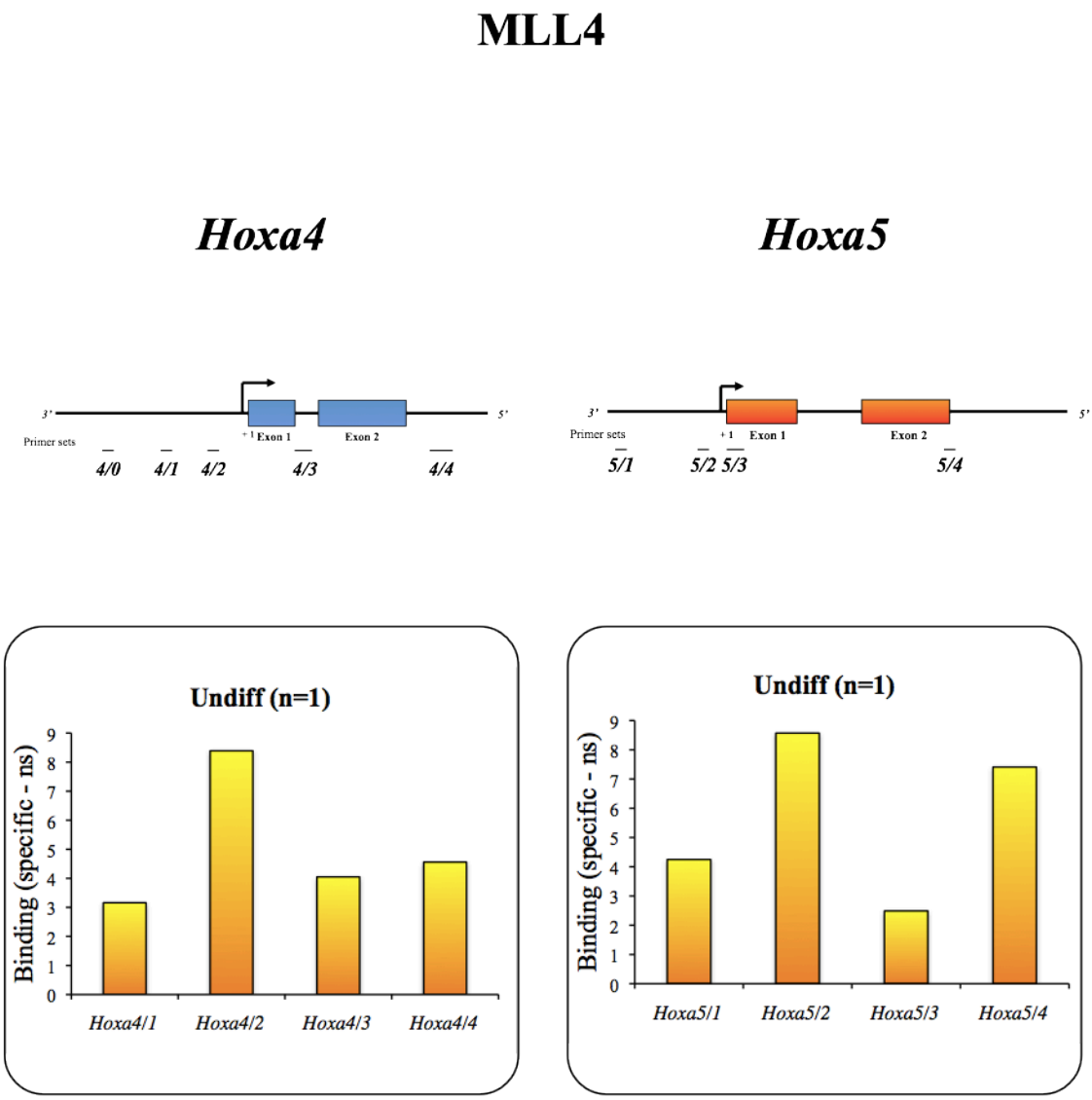


Figure 5.7: Distribution of Histone Methyl Transferase MLL4 on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells

Chromatin of undifferentiated HPC-7 cells (Undiff) was extracted in undifferentiated cells, and then immunoprecipitated with antibody against MLL4. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes, and was performed in triplicate from one ChIP experiment.

small number of kinases are known to show H3-specifity, or might be involved in the regulation of *Hoxa* genes. One appeared to be a good candidate: Msk1 (Figure 5.8). Indeed, Msk1 has been shown to phosphorylate H3 on serine 10 (Vicent *et al.*, 2006), and is recruited to promoters, leading to H3S10 phosphorylation over the promoter of these genes (Bruck *et al.*, 2009). Moreover, it was suggested that the phosphorylated H3S10 mark, in association with other histone modifications, could be a platform for the subsequent recruitment of chromatin-remodelling complexes (MacDonald *et al.*, 2005; Vicent *et al.*, 2006; Bruck *et al*., 2009). We therefore examined Msk1 distribution over the *Hoxa4* and *Hoxa5* genes to determine whether it contributes to MLL target gene regulation.

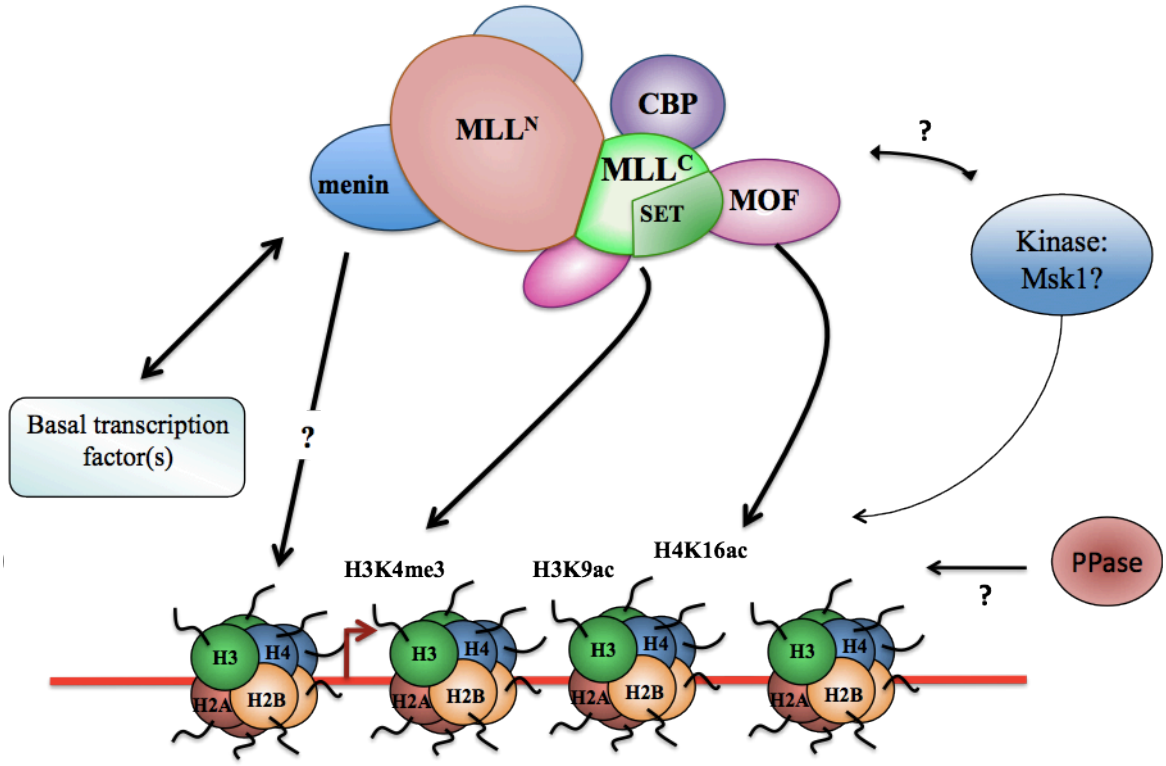


Figure 5.8: Suggested model of interaction of Msk1 with the MLL complex

Histone modification distributions suggest that a histone H3S10-specific kinase and phosphatase contribute to *Hoxa* gene regulation. It is unclear whether the kinase Msk1, which is detected on these genes, is part of the MLL complex, or is recruited independently.

6. DOES MSK1 CONTRIBUTE TO *HOXA* GENE REGULATION?

6.1. How is Msk1 distributed along *Hoxa* genes?

Msk1 (Mitogen- and Stress- activated Protein Kinase 1) is a protein kinase involved in the phosphorylation of H3S10 (Soloaga *et al.*, 2003), but is also downstream of the Jnk signalling pathway (Huang *et al.*, 2006). Our data suggests that H3S10p contributes to *Hoxa* gene activation. In order to examine the role of Msk1 as part of the mechanism of regulation of *Hoxa* genes, its distribution was examined on *Hoxa4* and *Hoxa5*, in undifferentiated HPC-7 cells and megakaryocytes (Figure 6.1). An enrichment of Msk1 was observed over the TSS of *Hoxa4* in undifferentiated HPC-7 cells (*Hoxa4/2*) (Figure 6.1, *Upper left panel*), whereas this peak is lost in megakaryocytes (*Lower panel*). Interestingly, this distribution mirrors the patterns of both H3K9acS10p and RNA polymerase II, in undifferentiated HPC-7 cells, with the peak of Msk1 binding just downstream of the MLL1 and menin binding sites. In contrast, on the *Hoxa5* gene (Figure 6.1, *Right panels*), there is only a slight enrichment of Msk1 over the TSS of the gene in undifferentiated HPC-7 cells (*Upper panel*), but a similar absence of the protein in megakaryocytes (*Lower panel*). This distribution also follows the distribution of proteins of the MLL complex and of RNA polymerase II (with a peak on *Hoxa5/2*), with the Msk1 binding site positioned just upstream of the peaks of H3K9acS10p and H3K4 trimethylation (*Hoxa5/3*). However our inability to detect an equivalent peak of Msk1 on *Hoxa5* is surprising, given the large peak of H3S10p found on the gene. Thus, on both genes, the sites of Msk1 binding correlates with the modifications deposited by proteins of the MLL complex, but the relative abundance of the enzymes and modifications do not correlate. These data suggest that Msk1 is either part of the MLL complex or acts together with the MLL complex to regulate *Hoxa* gene expression.

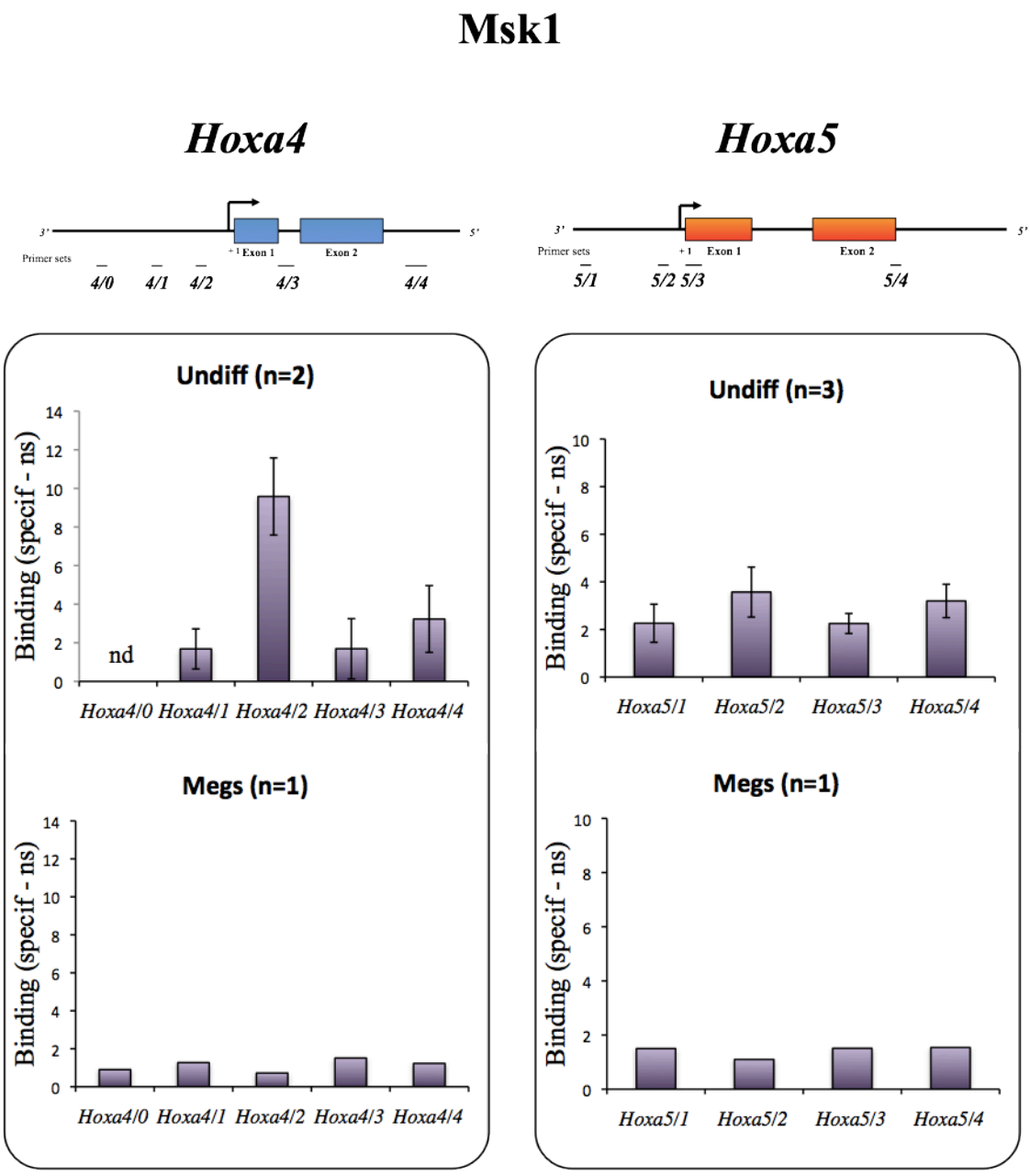


Figure 6.1: Distribution of Msk1 on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (Undiff, *Upper panels*) were differentiated into megakaryocytes (Megs, *Lower panels*). Chromatin was extracted and immunoprecipitated with antibody against Msk1. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes, and was performed in triplicate from one, two or three ChIP experiments. Error bars represent the standard error of the mean.

In order to assess whether Msk1 is a component of the MLL complex, we performed co-immunoprecipitation assays using several antibodies recognising components of the MLL complex (MLL^C, MLL^N, and CBP) and Msk1. We also used NFκB as an Msk1 - NFκB interaction has been previously described (Reber *et al*., 2009) (Figure 6.2). A whole-cell extract of undifferentiated HPC-7 cells was used as a source of material, and immunoprecipitated with either MLL1 (C- or N-terminal), NF-κB, Msk1 or CBP antibody. Each of these samples was separated on a 7% SDS gel, before transfer to a nitrocellulose membrane. The interaction between MLL1 and Msk1 was assessed by Western Blotting of a membrane containing proteins immunoprecipitated with either MLL^C, MLL^N or CBP antibodies. Western blotting with Msk1 was performed several times with reproducible results (Figure 6.2a). Msk1 (87kDa) is observed to associate with MLL^C and NFκB, but the interaction is consistently weaker for binding to MLL^C (lane 1) than to NFκB (lane 5). In contrast, no interaction of Msk1 was observed with MLL^N (lane 2) or CBP (lane 4).

As a control, the "reverse" IP was performed (Figure 6.2b). An anti-NFκB was shown to successfully immunoprecipitate NFκB (lane 2), but failed to co-immunoprecipitate Msk1 (lane 3). Similarly, probing with MLL^C or MLL^N antibodies failed to show the reverse interaction of MLL^C with Msk1. These experiments suggest that if MLL^C and Msk1 interact, these interactions are likely to be weak or transitory (Figure 6.2a). Further IP studies with other Msk antibodies (currently unavailable) will be necessary to confirm this interaction.

6.2. Importance of Msk1 as an actor of *Hoxa* gene regulation?

The previous data seems to suggest that Msk1 is an essential component of the mechanism of regulation of the *Hoxa* genes. We intended to assess the importance of Msk1 in the regulation of *Hoxa* genes, by artificially altering the distribution of Msk1 deposited histone

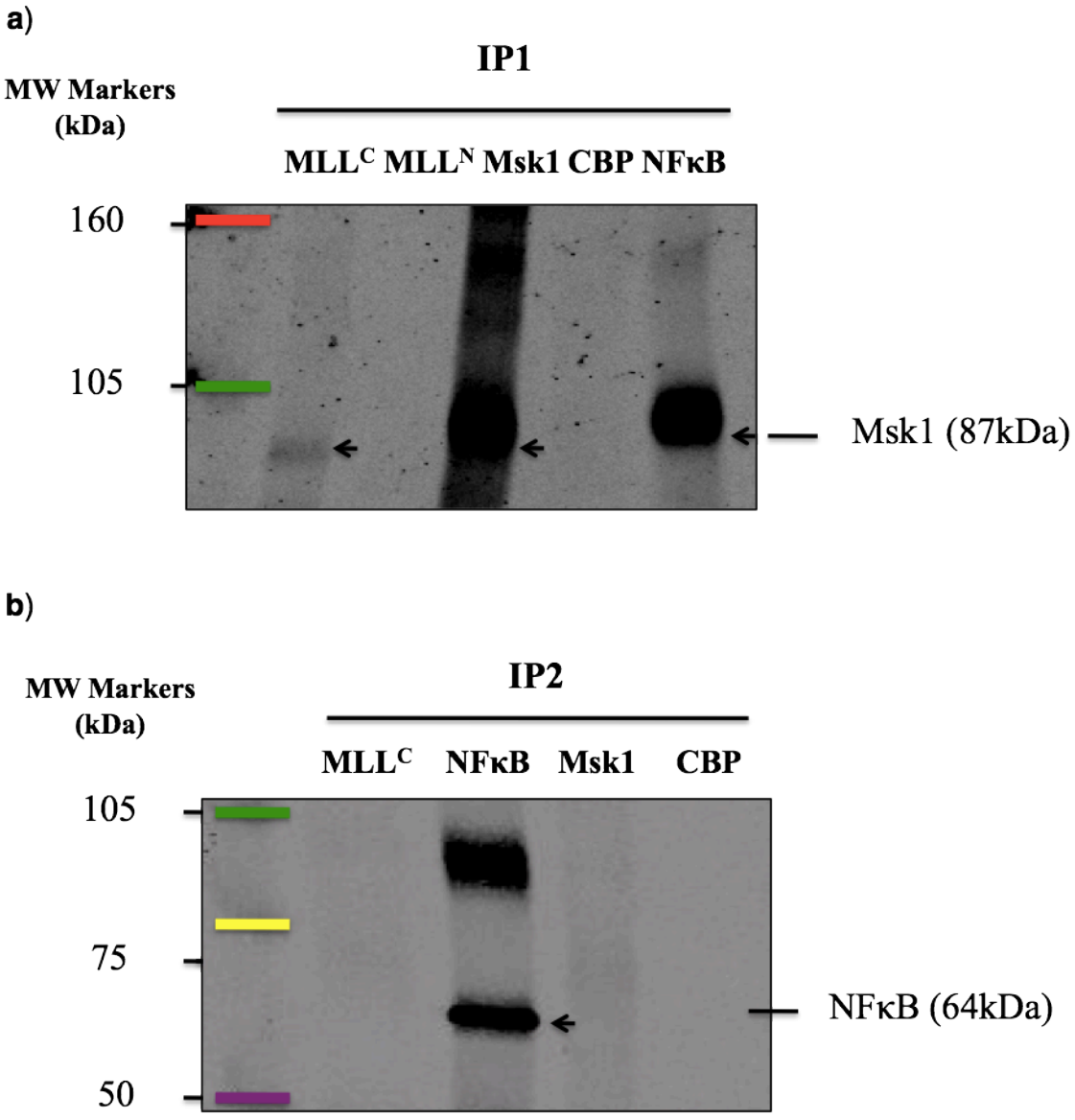


Figure 6.2: Co-immunoprecipitation of MLL complex interacting proteins

Interactions between MLL, NFκB, CBP and Msk1 were assessed by immunoprecipitation with the corresponding antibodies (anti-MLL1^{c} antibody (*Upstate*), anti-MLL1^{N} antibody (*Upstate*), anti-NFκB antibody (*AbCam*), anti-Msk1 antibody (*AbCam*) and anti-CBP antibody (*Upstate*), and then by Western Blotting with (**a**) Msk1 or (**b**) NFκB antibodies. Proteins of interest are indicated with an arrow.

modifications. Thus, two different drugs were used to alter the normal abundance of phosphorylation of H3S10 in HPC-7 cells, in order to examine whether this induced changes in *Hoxa* gene expression.

6.2.1. What are the global changes on histone modifications after treatment with okadaic acid?

In a first instance, the toxin okadaic acid, which strongly inhibits serine/threonine phosphatases, was used in order to increase the abundance of phosphorylated histone marks by preventing their removal. Undifferentiated HPC-7 cells were treated with a range of okadaic acid concentrations, and a time-course experiment was performed, in order to determine the effect of this agent on global histone modifications. Cells were treated with 1, 5 or 10nM okadaic acid, histones were extracted and histone modification abundance quantified by Western blotting using a range of appropriate antibodies including H3K9acS10p (Ab12181), H3S10p (Ab14955), H3T3p (Ab17352). We also analysed the non-phospho-isoform H4K16ac (Lab 252) as a control to examine whether okadaic acid only has impacts on histone phosphorylation (Figure 6.3a). The bottom panel shows after Coomassie staining that histones were equally loaded ("Total histone"), though Western analysis with a Histone H3 C-terminal antibody (*AbCam*) was used to normalise for protein loading. All marks examined show an increased abundance following the addition of a gradient of okadaic acid (Figure 6.3b). The highest increase is observed for the dual mark H3K9acS10p, which shows a ~2.5-fold rise, whilst H3S10p is only slightly increased (It appears to be reduced due to unequal histone loading). In addition, another phosphorylated mark, H3T3p, shows a ~2-fold increase after treatment with 10nM okadaic acid, confirming that the agent is not selective for serine-specific phosphatases. Perhaps not surprisingly, H4K16ac abundance, which would not be expected to increase as a direct effect of okadaic acid, also shows a small

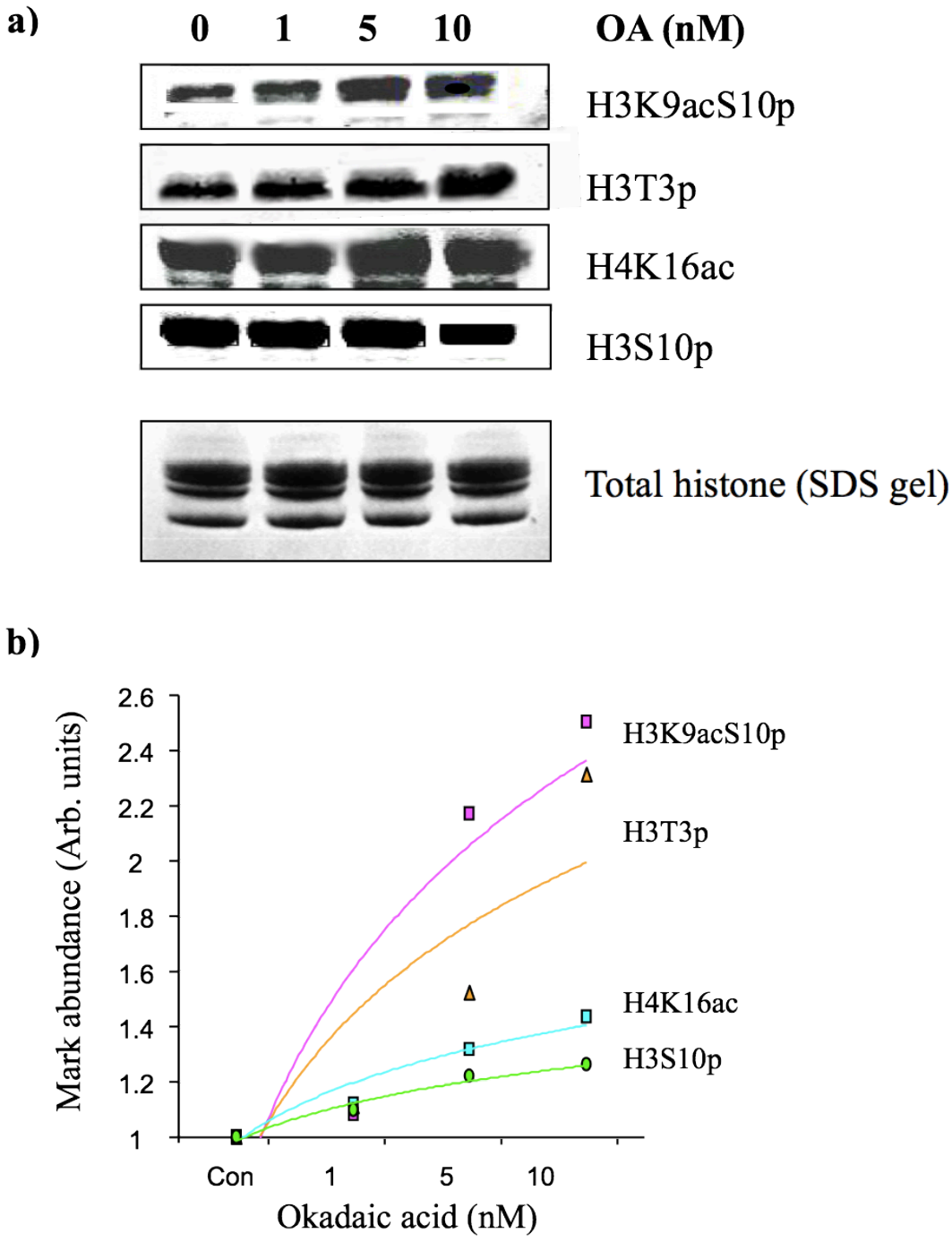


Figure 6.3: Quantification of histone modifications after exposure to gradient of concentration of okadaic acid in undifferentiated HPC-7 cells

Undifferentiated HPC-7 cells were treated with 1, 5 or 10nM of okadaic Acid for 24 hours. Histones were extracted and specific marks were examined by **(a)** Western blotting using appropriate antibodies (H3K9acS10p, H3T3p, H4K16ac and H3S10p). The gel is shown to be equally loaded (SDS gel). **(b)** Quantification is presented on a graph, normalised to the control, set at 1.

increase, suggesting that either "histone cross-talk" effects, or secondary, indirect gene effects take place after treatment.

Subsequently, a time-course experiment with 10nM okadaic Acid treatment of undifferentiated HPC-7 cells was performed in order to identify the ideal window of treatment with this toxin. Undifferentiated HPC-7 cells were treated with 10nM okadaic acid for 1, 2, 4, 8, 12, 18 and 24 hours. Histone were extracted as previously described and their abundance was quantified by Western blotting using H3K9acS10p or H4K16ac antibodies (Figure 6.4), using detection by the H3 C-terminal antibody (*AbCam*) to normalise for protein loading. As shown before, H4K16ac shows a small change (10% increase) during the time-course. By this assessment, global H3K9acS10p modification seems to be the highest between four and eight hours treatment.

In undifferentiated HPC-7 cells, active *Hoxa* genes were associated with H3S10p enrichment over the promoter of genes, whilst a reduction of *Hoxa* gene expression in megakaryocytes was associated with the loss of this mark. We hypothesised that the increase in H3S10P abundance due to okadaic acid treatment would lead to an increase of MLL target gene expression in both undifferentiated HPC-7 cells and megakaryocytes (Figure 6.5). We therefore examined the transcript abundance of specific *Hoxa* genes following treatment of undifferentiated HPC-7 cells and megakaryocytes with 1, 5 or 10nM okadaic acid for 8 (data not shown) and 24 hours (Figure 6.6). After eight hours of treatment, slight changes of the level of MLL target gene expression in undifferentiated HPC-7 cells were observed, with only small changes at *Hoxa4* and *Gapdh* (1.2 and 1.5-fold changes respectively). No changes were observed with *Hoxa2*, *Hoxa3* and *Hoxa5*.

Changes in the level of gene expression following 24 hours of 10nM okadaic acid treatment were then examined in undifferentiated HPC-7 cells and megakaryocytes (Figure 6.6). The

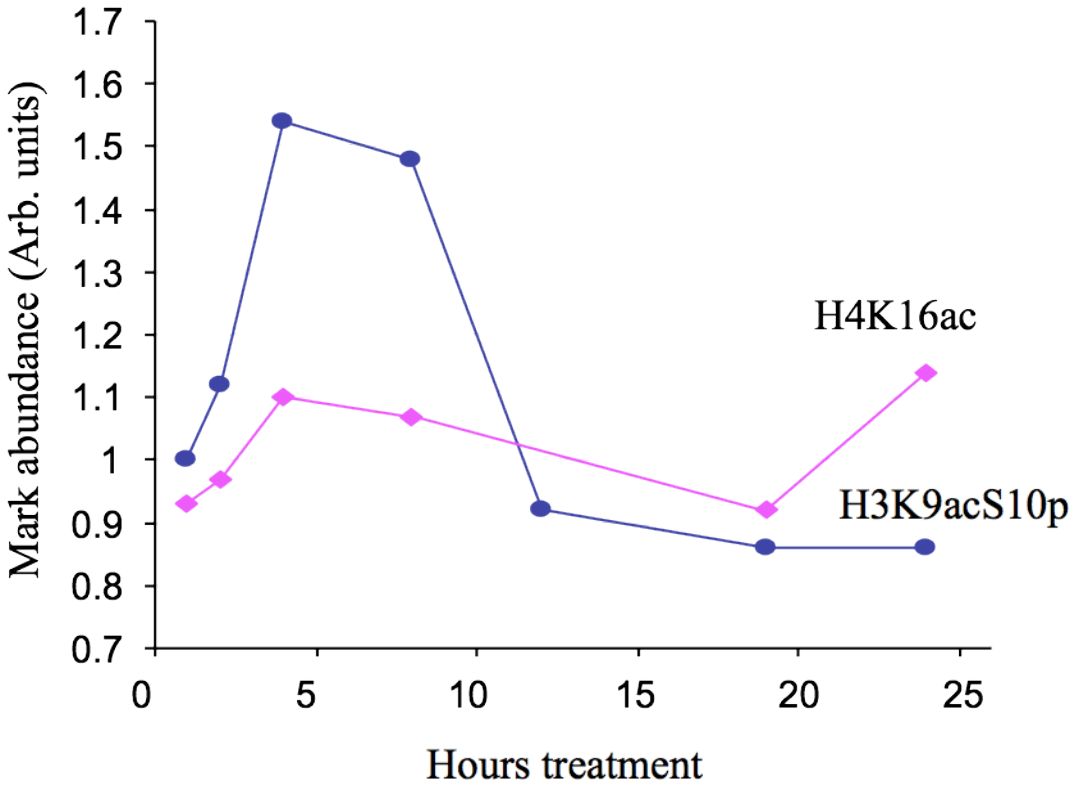


Figure 6.4: Effect of okadaic acid on histone modification abundance in undifferentiated HPC-7 cells

Undifferentiated HPC-7 cells were treated with 10nM okadaic acid for 1, 2, 4, 8, 12, 18 and 24 hours. Histones were extracted and specific marks were examined by Western blotting using appropriate antibodies (H3K9acS10p, and H4K16ac). Mark abundance is normalised to untreated control cells.

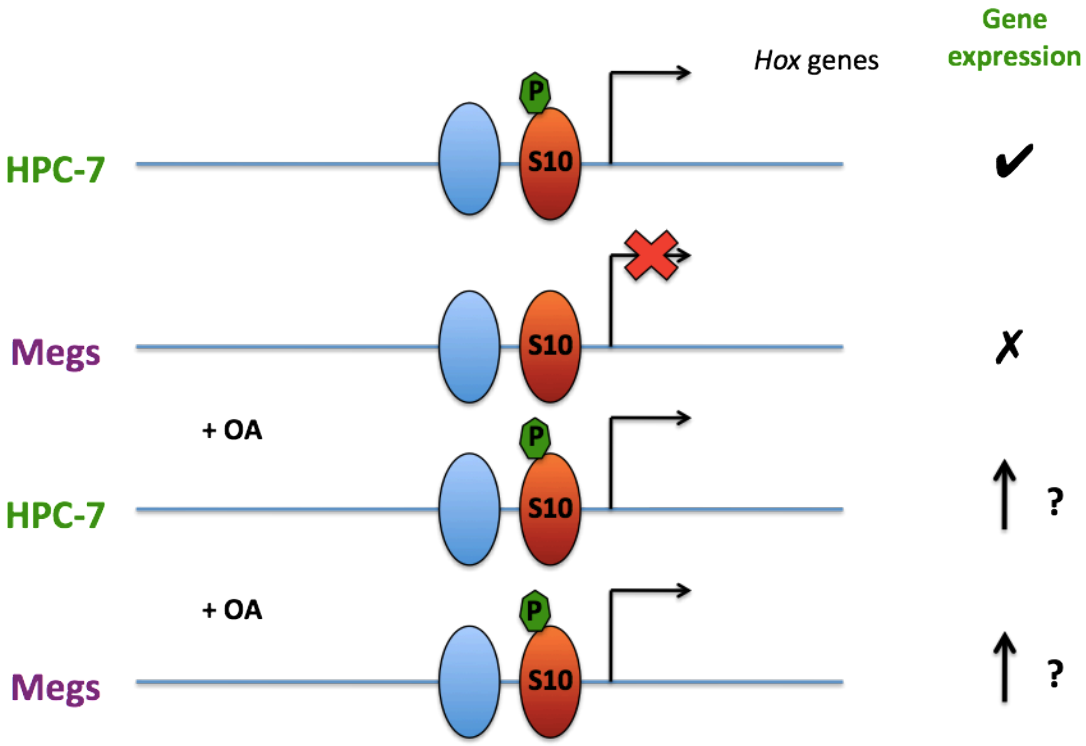


Figure 6.5: Model of changes possible after treatment with okadaic acid

In undifferentiated HPC-7 cells, active *Hoxa* genes are associated with a H3S10p enrichment over the promoter of genes, whilst a reduction of *Hoxa* gene expression in megakaryocytes is associated with the loss of this mark. In this model, it is suggested that the prevention of phosphate removal at H3S10 by the phosphatase inhibitor, okadaic acid, would lead to an increase of gene expression in both undifferentiated HPC-7 cells and megakaryocytes.

expression of several genes (*Hoxa2*, *Hoxa3*, *Hoxa4*, *Hoxa5* and *Gapdh*) was examined in response to this treatment, where gene expression is normalised with the level of expression in control cells (cells treated with DMSO). In undifferentiated cells (Figure 6.6, *Upper panel*), an increase of gene expression *Hoxa4* and *Gapdh* (1.4 and 3-fold respectively) was observed in correlation with an increased concentration of okadaic acid. No changes were observed for *Hoxa2, Hoxa3* and *Hoxa5*. In megakaryocytes (Figure 6.6, *Lower panel*), only *Hoxa3* showed a 8-fold increase of its level of expression following okadaic acid treatment, possibly reflecting the lower abundance of this transcript in megakaryocytes. No changes were observed for *Hoxa2*, *Hoxa4*, *Hoxa5* and *Gapdh*. It should be noted that the addition of okadaic acid was performed after five days of megakaryocyte differentiation, suggesting that the *Hoxa* genes would be expected to have already lost H3S10 phosphorylation, as the genes are down-regulated.

6.2.2. What are the local changes on *Hoxa* genes after treatment with okadaic acid?

In undifferentiated HPC-7 cells, okadaic acid induced an increase in gene expression at three MLL-target genes; *Hoxa4*, *Gapdh* and *Hoxa5*. To investigate if this increased level of gene expression is associated with an increase of histone phosphorylation, the distributions of H3S10 phosphorylation, H3K9acS10 phosphorylation and another phosphorylation mark, H3T3p, were examined on these genes after treatment with okadaic acid (10nM okadaic acid for 24 hours). As a control, cells were cultured with the addition of DMSO in the medium. The distribution of these marks along these genes was assessed by N-ChIP.

The distribution of H3K9acS10p was first examined along *Hoxa4* and *Hoxa5* genes in untreated (*Upper panels*) and treated HPC-7 cells (*Lower panels*) (Figure 6.7). Both *Hoxa4* and *Hoxa5* genes display the same distribution as previously described (Cf. Figure 4.8). On

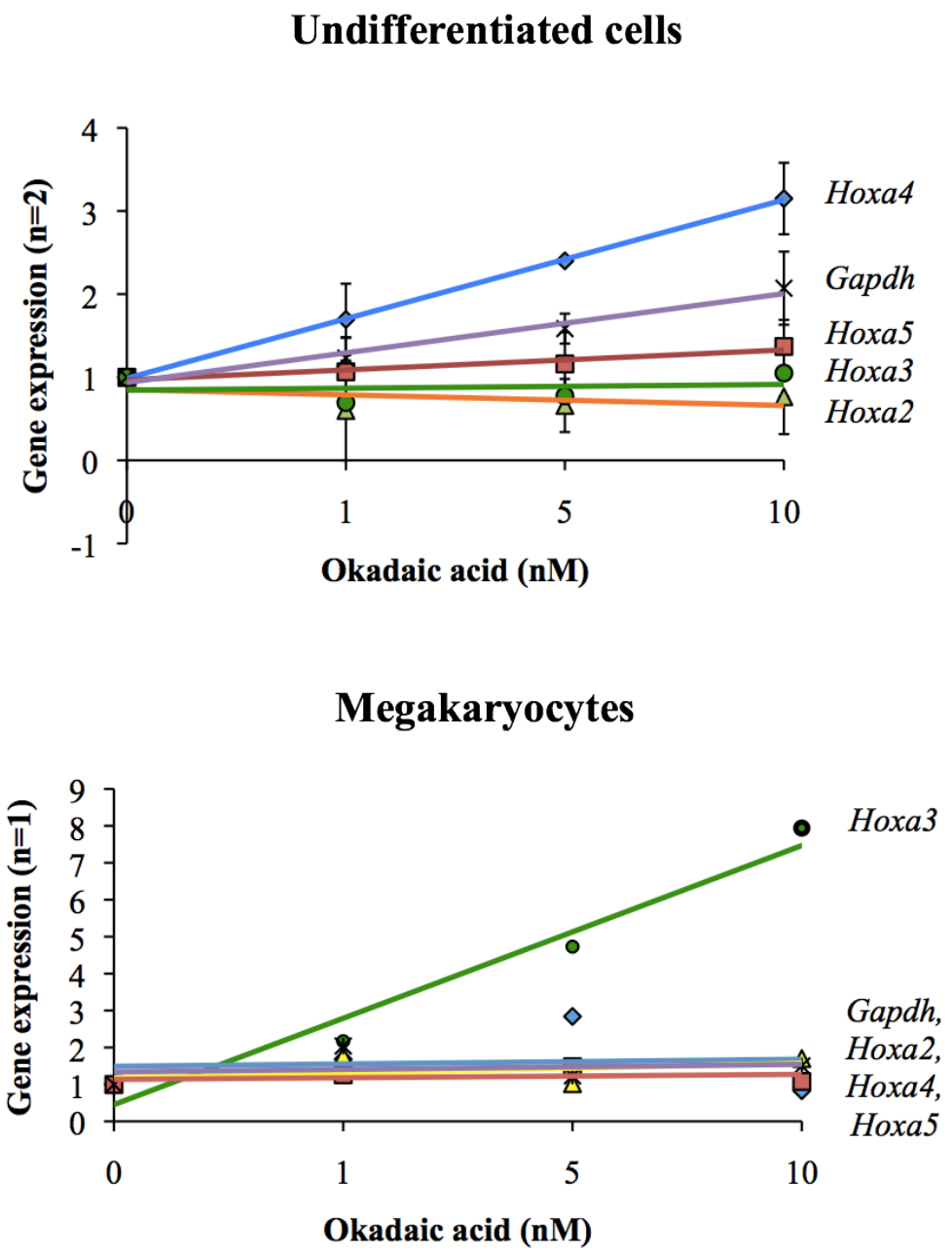


Figure 6.6: Effect of okadaic acid on gene expression in undifferentiated HPC-7 cells and megakaryocytes

Undifferentiated HPC-7 cells (*Upper panel*) or megakaryocytes (*Lower panel*) were treated with 1, 5 or 10nM okadaic acid for 24 hours. qRT-PCR was performed with primer sets corresponding to the coding region of *Hoxa2, Hoxa3*, *Hoxa4*, *Hoxa5* and *Gapdh*, and was performed in triplicate from one or two experiments. Error bars represent the standard error of the mean.

the *Hoxa4* gene (Figure 6.7, *Left panels*), an enrichment of H3K9acS10 phosphorylation was observed over the transcription start site of the gene (*Hoxa4/2*) (*Upper panel*), which showed a 1.5 fold increase after treatment with 10nM okadaic acid (*Lower panel*). Similarly, on *Hoxa5* (Figure 6.7, *Right panels*), an enrichment of the dual mark H3K9acS10p was observed over the TSS (*Hoxa5/3*) (*Upper panel*), which is then increased by 10-fold after treatment with 10nM okadaic acid (*Lower panel*). This suggests that the enhancement of *Hoxa4* and *Hoxa5* gene expression in undifferentiated HPC-7 cells, after treatment with okadaic acid, is associated with an increase in H3K9acS10p abundance at their promoters, but also implies that an unknown phosphatase is acting at these loci.

Then, the distribution of the particular mark H3S10p was assessed on the *Hoxa4* and *Hoxa5* genes in undifferentiated HPC-7 cells (Figure 6.8). Surprisingly, on *Hoxa4* (Figure 6.8, *Left panels*), this mark is essentially absent in untreated cells (*Upper panel*), and only shows a minor increase in treated cells (*Lower panel*). This is difficult to explain given the peak of H3K9acS10p at the TSS of *Hoxa4*, and that a similar peak of H3S10p is observed at the TSS of *Hoxa5* (*Hoxa5/2*) in untreated cells (Figure 6.8, *Right panels*). This enrichment is then increased by 30-fold in treated cells (*Lower panel*). Interestingly, the enrichment of H3S10p is situated on a site just upstream (*Hoxa5/2*) to the site enriched with H3K9acS10p (*Hoxa5/3*). Thus, in both cases, *Hoxa4* has lower levels of these marks (H3K9acS10p, H3S10p) despite showing similar levels of H3K4me3 (Cf. Figure 4.3). This is consistent with the output of the two genes: *Hoxa4* (Ct = 30) is much less transcriptionally active than *Hoxa5* (Ct = 23) (Cf. Figure 3.3), and may suggest that these phospho-marks, rather than H3K4me3, reflect the level of transcription.

Finally, the distribution of another phosphorylated mark, H3T3p, was assessed in untreated and okadaic acid treated HPC-7 cells, on *Hoxa4* and *Hoxa5* (Figure 6.9). This mark is a

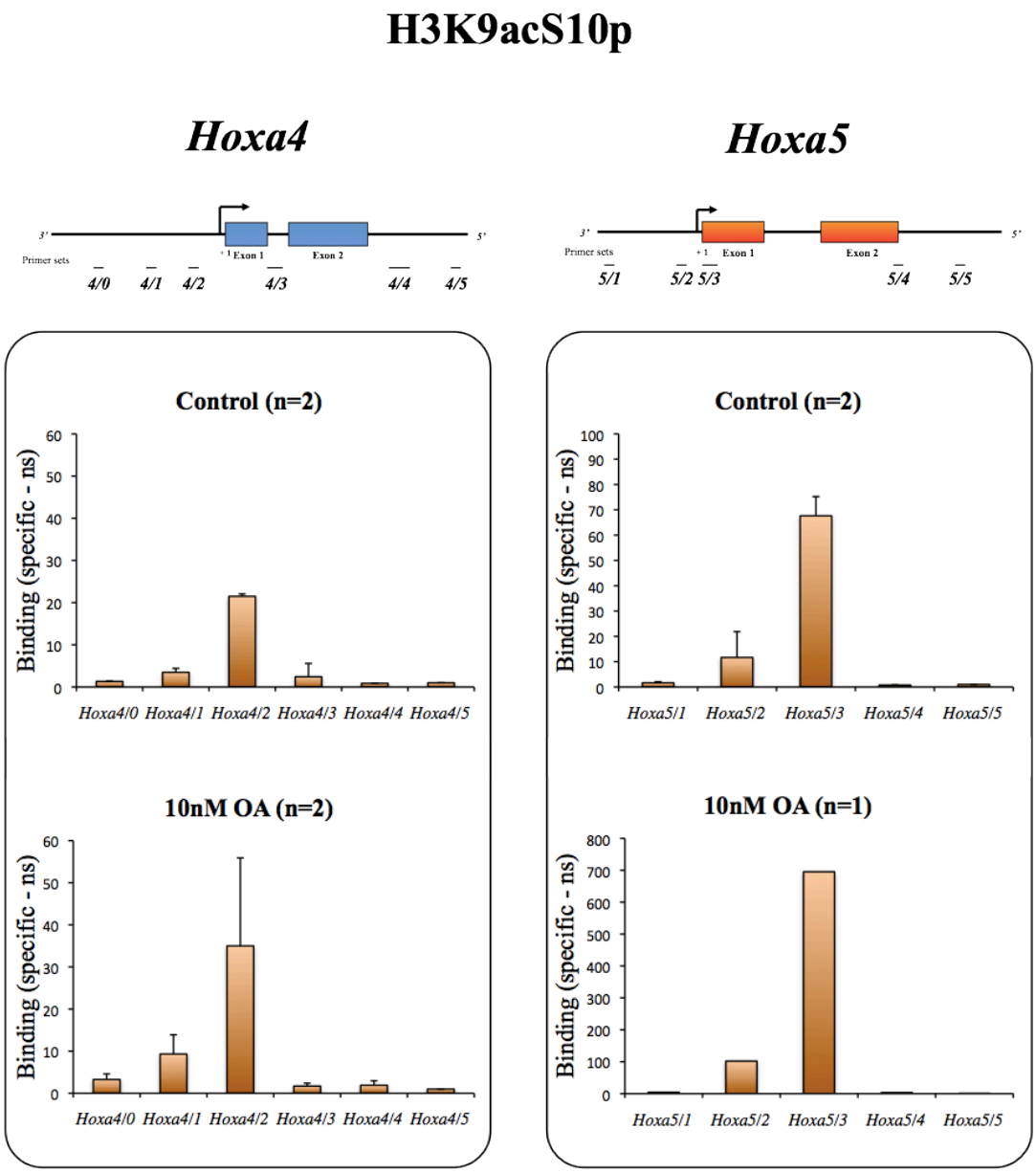


Figure 6.7: Distribution of H3K9acS10p histone modifications on the *Hoxa4* and *Hoxa5* genes in the control undifferentiated HPC7 cells and after treatment with 10nM of okadaic acid

Undifferentiated HPC7 cells were treated with either DMSO (as a control, *Upper panels*) or with 10nM of okadaic acid (*Lower panels*), for 24 hours. Chromatin was extracted and immunoprecipitated with antibody against **H3K9acS10p**. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

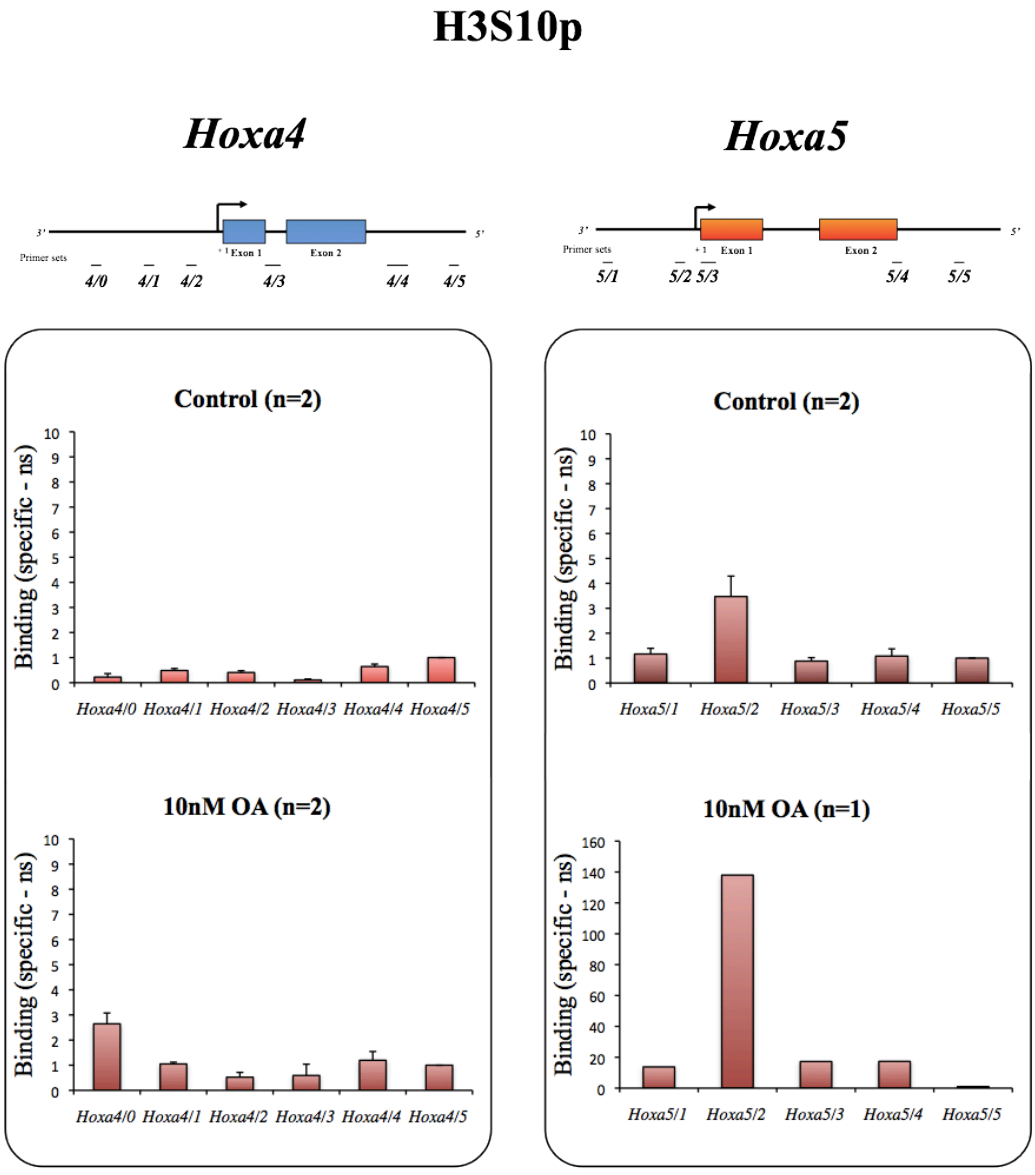


Figure 6.8: Distribution of H3S10p histone modifications on the *Hoxa4* and *Hoxa5* genes in the control undifferentiated HPC7 cells and after treatment with 10nM of okadaic acid

Undifferentiated HPC7 cells were treated with either DMSO (as a control, *Upper panels*) or with 10nM of okadaic acid (*Lower panels*), for 24 hours. Chromatin was extracted and immunoprecipitated with antibody against **H3S10p**. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from one or two ChIP experiments. Error bars represent the standard error of the mean.

control, as we have little insight into its regulatory role, but it is known to have an inhibitory effect on MLL activity (Southall *et al.*, 2009). Interestingly, H3T3p displays a similar pattern of distribution on *Hoxa4* to that of H3S10p (Figure 6.9, *Left panels*). Indeed, in untreated cells (*Upper panel*), the low levels of H3T3p, suggesting this mark is essentially absent, indicate that this does not contribute to *Hoxa4* regulation (*Lower panel*). However, on *Hoxa5* (Figure 6.9, *Right panels*), an enrichment of H3T3p was observed over the transcription start site of the gene (*Upper panel*), which is then reduced (three-fold decrease) after treatment with okadaic acid (*Lower panel*) (Southall *et al.*, 2009). This is in contrast with the global increase in this mark upon okadaic acid treatment, but is consistent with the increase in *Hoxa5* gene activity.

6.2.3. Changes after treatment with an inhibitor of H3S10 phosphorylation?

Having investigated gene expression and histone modification changes occurring after treatment with a phosphatase inhibitor, we examined the changes induced following treatment with an inhibitor of H3S10 phosphorylation, the Jun N-terminal Kinase inhibitor SP6000125 (Huang *et al.,* 2006). As with okadaic acid, a time-course was performed in order to determine the effect of this agent on global histone modifications (Figure 6.10). Undifferentiated HPC-7 cells were treated with 10nM Jnk inhibitor for 1, 2, 4, 8, 12 and 24 hours, and the global changes in H3K9acS10p, H3S10p and H4K16ac were examined. All three histone marks display a major decrease in abundance between four and eight hours (~30%). The parallel response H4K16ac is an indication that these cannot be "clean" experiments in which a single class of modifications are modulated without functional impacts elsewhere in the nucleus. This is perhaps not surprising given that the five-hour treatment allows a broad range of processes to occur. The impact on H3K16ac could reflect direct histone cross-talk, for example MOF, an H4K16 acetyl-transferase, is dynamically

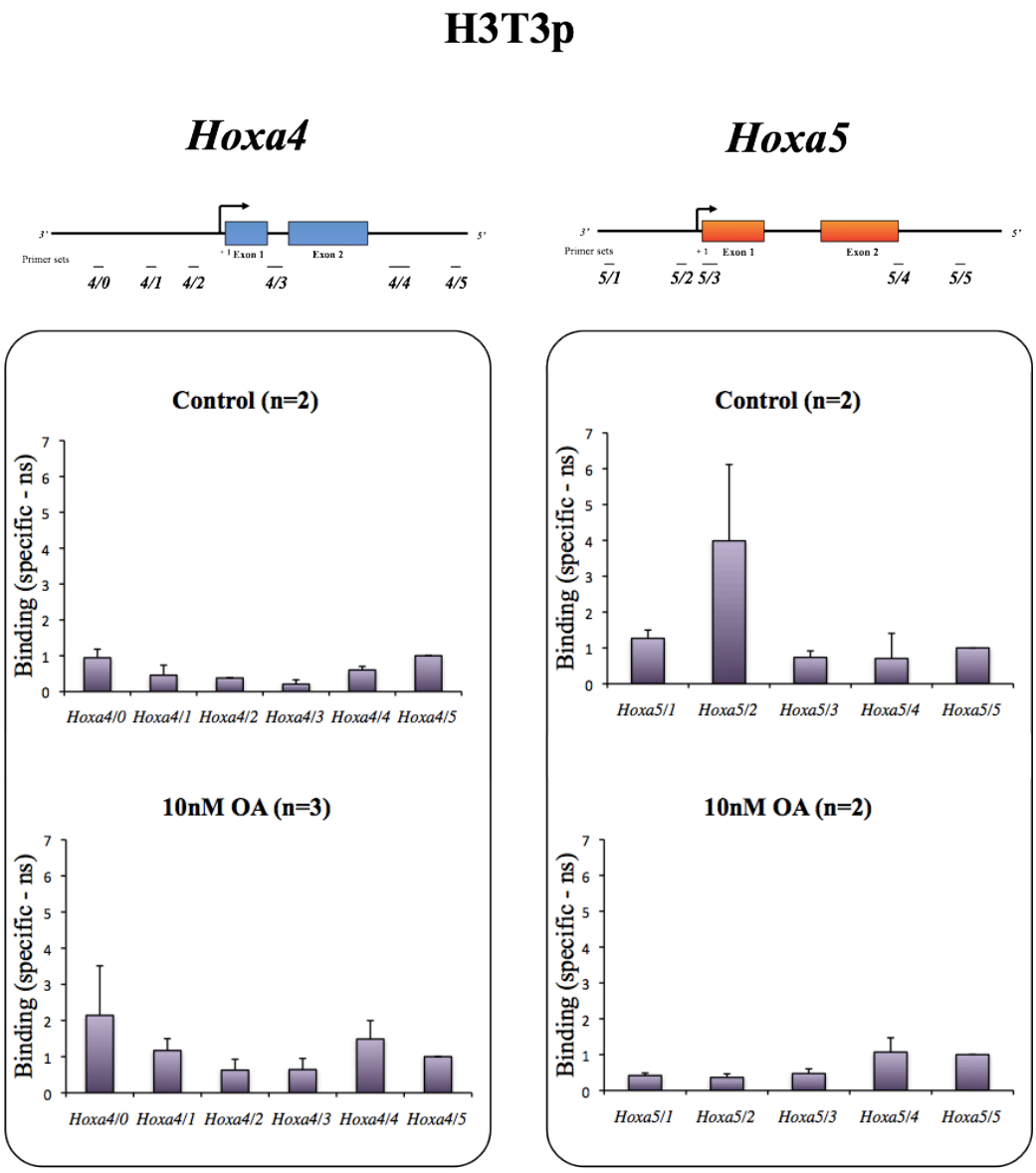


Figure 6.9: Distribution of H3T3p histone modifications on the *Hoxa4* and *Hoxa5* genes in the control undifferentiated HPC7 cells and after treatment with 10nM of okadaic acid

Undifferentiated HPC7 cells were treated with either DMSO (as a control, *Upper panels*) or with 10nM of okadaic acid (*Lower panels)*, for 24 hours. Chromatin was extracted and immunoprecipitated with antibody against **H3T3p**. Graphs display the relative enrichment [Bound (specific) - Bound (pre-immune)], where non-specific binding (ns) obtained from pre-immune bound is subtracted. PCR was performed with primer sets corresponding to specific sites on the *Hoxa4* and *Hoxa5* genes and was performed in triplicate from two or three ChIP experiments. Error bars represent the standard error of the mean.

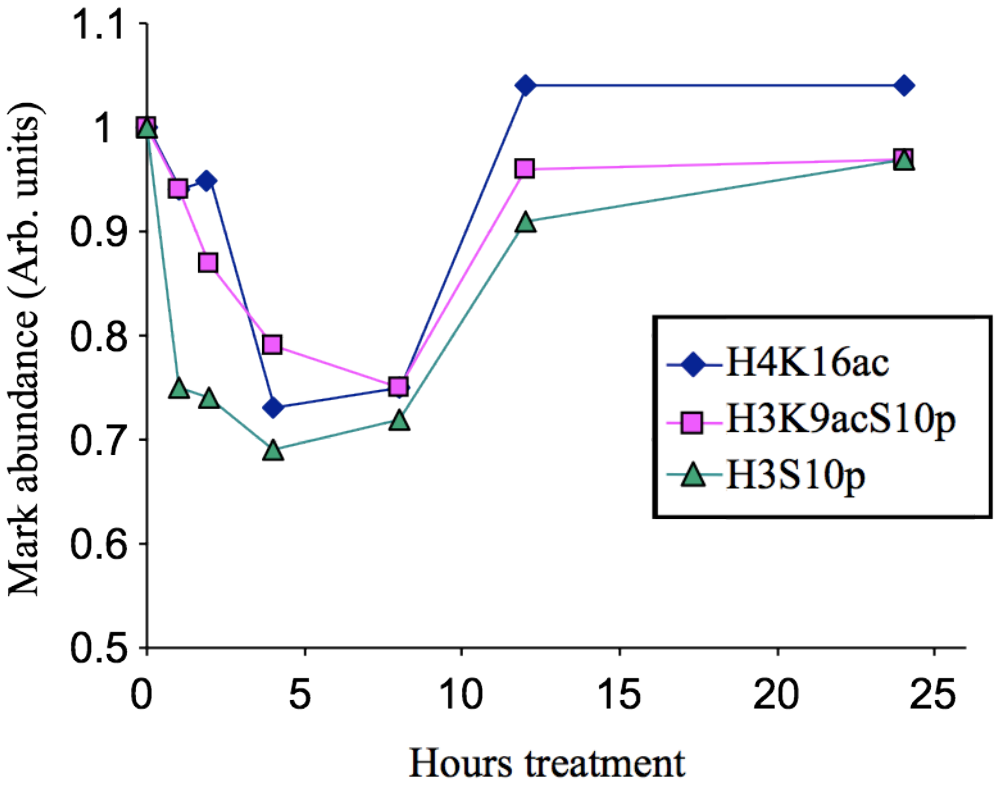


Figure 6.10: Changes in cellular histone modifications after exposure to 10nM Jnk inhibitor in undifferentiated HPC-7 cells

Undifferentiated HPC-7 cells were treated with 10nM Jnk inhibitor (SP6000125) for 1, 2, 4, 8, 12, 18 and 24 hours. Histones were extracted and specific marks were examined by Western blotting using appropriate antibodies (H3K9acS10p, H3S10p and H4K16ac). Mark abundance is normalised to untreated control cells. The protein level was normalised with the H3 C-terminal antibody.

recruited to chromatin by 14-3-3, a protein that recognises the H3S10p mark (Zippo *et al.*, 2009). However, this could also reflect indirect or secondary gene effects (i.e. increased H3S10p stimulates gene activity, and these gene products increase H4K16ac elsewhere in the genome.)

Transcript abundance was then assessed in undifferentiated HPC-7 cells treated with 1, 25 or 50nM Jnk inhibitor for eight hours in order to examine the changes in gene expression following the inhibition of H3S10 phosphorylation (Figure 6.11). Levels of gene expression were examined for *Hoxa2*, *Hoxa3*, *Hoxa4*, *Hoxa5* and *Gapdh*. The graph displays the transcript abundance of each of these genes, normalised to the transcript abundance of ß-actin. Generally, these show only small effects after eight hours of treatment, but a number of genes show increases (rather than the expected decreases) in transcript abundance, for example, *Hoxa2*, *Hoxa3* and *Hoxa5* show 1.4 - 2-fold increases at 50nM, correlating with the increase of the concentration of the inhibitor. No substantial changes were observed with *Hoxa4* and *Gapdh*. Despite this treatment giving an unexpected increase in transcriptional activity, it is interesting to note that the effects at specific genes are the reverse of those seen with okadaic acid (okadaic acid induces changes in the order *Hoxa4*, *Gapdh*, *Hoxa5*, *Hoxa3*, *Hoxa2* with *Hoxa4* being the most expressed, i.e. Figure 6.6, whereas the Jnk inhibitor induces changes in the order *Hoxa2*, *Hoxa3*, *Hoxa5, Hoxa4* and *Gapdh* with *Hoxa* 2 being the most expressed in cells treated with the highest concentration of Jnk inhibitor, ie. Figure 6.11).

The contribution of MLL in the regulation of *Hoxa* was examined in the previous chapter. It was shown that up-regulation of specific *Hoxa* genes is related to the recruitment of MLL1 to the promoter of these genes, and associated with the methylation of H3K4 over the transcription start site. Moreover, other proteins of the MLL complex, such as CBP and

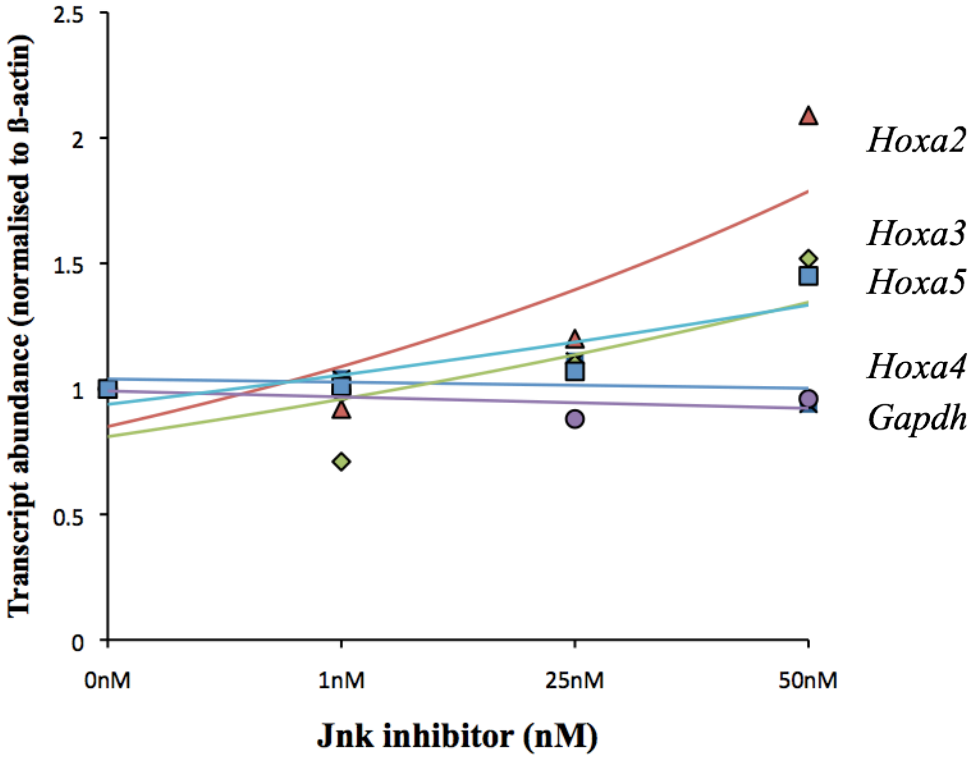


Figure 6.11: *Hoxa* gene expression in response to Jnk inhibitor treatment

Undifferentiated HPC-7 cells were treated with 1, 25 or 50nM Jnk inhibitor for 8 hours. PCR was performed with primer sets corresponding to the coding region of *Hoxa2, Hoxa3*, *Hoxa4*, *Hoxa5* and *Gapdh*, and was performed in triplicate from one experiment.

menin, were shown to interact with the same sites. The presence of CBP was shown to be associated with the acetylation of H3K9 over the transcription start site of these genes. In addition, it was shown that the TSS of MLL target genes was also enriched with H3K9acS10 phosphorylation on active genes, followed by a decrease of this mark upon transcriptional silencing. In this chapter we show a correlation between the distribution of this mark and the distribution of the kinase Msk1, on the *Hoxa4* and *Hoxa5* genes, implying that this mark is regulated by the kinase. In addition, co-immunoprecipitation experiments detect a weak interaction between MLL^C and Msk1, though this observation needs to be confirmed using other antibodies for the observation to be secure. Artificial (i.e. drug-induced) changes in histone phosphorylation impact on the activity of MLL-target genes, but the mechanism of their activity is clearly too complex for simple relationships to be identified.

7. ESTABLISHING SIRNA KNOCKDOWN PROTOCOLS FOR MLL1 FAMILY MEMBERS

In this final chapter, we discuss experiments that aimed to assess the contribution of MLL histone methyl-transferases to gene expression and to examine which histone modifications MLL complex enzymes bring to the promoter, and whether other enzymes are required. We proposed to use siRNA-mediated knockdown to address those questions. In order to develop this approach, vectors that generate siRNAi against *Mll1* were designed using an open access website "siRNA-Wizard". As members of the MLL family, vectors were also generated for *Mll3* and *Mll5* to investigate their contribution. Three sequences were designed to knock down *Mll3* and *Mll5* (Sequences in supplementary data) and were cloned into a commercial knock down vector system (pENTR/U6) (*Invitrogen*). In contrast commercial vectors were used to knock down *Mll1* (*Open BioSystems*). MLL1 contribution was first discussed in this chapter.

HPC-7 cells cannot be efficiently transfected using standard lipofection reagents (i.e. Lipofectamine 2000 or DMRYE C, *Invitrogen*). However, the transfection efficiency of Amaxa nucleofection in HPC-7 cells was previously optimised by Dr Fiona Tavner by transfecting HPC-7 cells with a GFP expressing vector or without DNA as a control (Figure 7.1). This resulted in significant cell death, showing only 50.5% cell viability after 24 hours transfection (as assessed by Flow cytometry), but the percentage of transfection in these remaining cells was shown to be 81% (Figure 7.1). Because of cell death, the analysis was limited to 24 hours post-transfection.

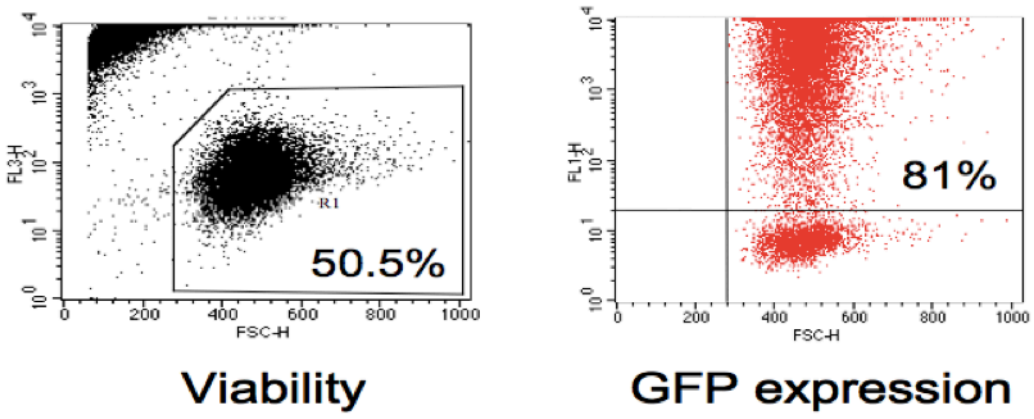


Figure 7.1: Optimisation of transfection in HPC-7 cells

The efficiency of Amaxa nucleofection-based transient transfection was previously assessed by GFP expression and subsequent flow cytometry by Dr Fiona Tavner. It was observed that 50.5% of cells are viable after transfection (Left panel). 81% of these cells have been efficiently transfected (Right panel). Cell viability was assessed by propidium iodide exclusion

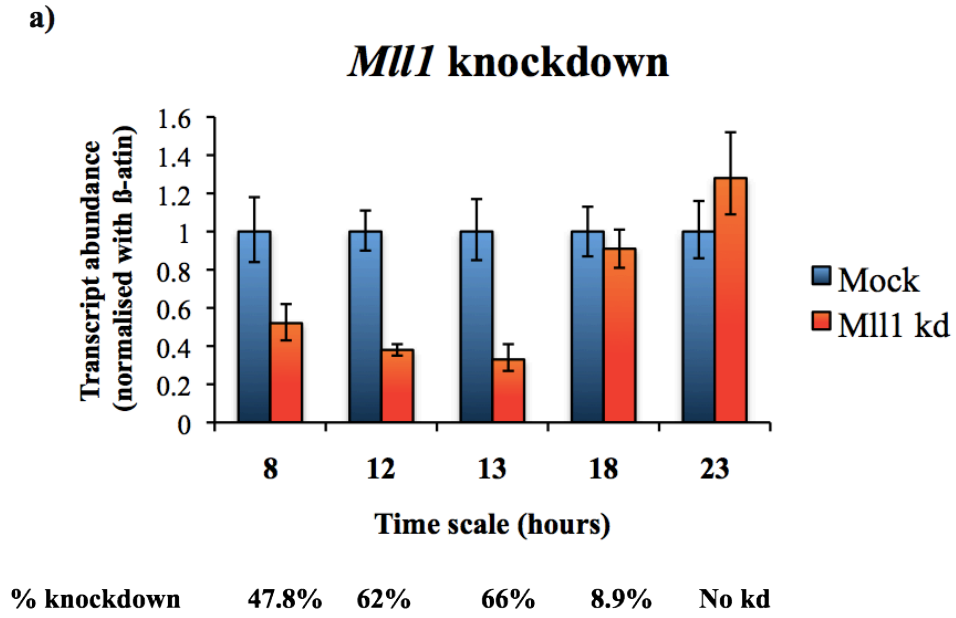
7.1. Establishing conditions for *Mll1* knockdown

7.1.1. Knockdown of *Mll* transcripts

In order to optimise the knockdown, a time-course after transient transfection was performed (Figure 7.2a), initially using *Mll1* transcript abundance to determine the optimal period for *Mll1* knockdown. Gene expression was quantified in relation to *Mll1* transcripts in "mock" transfected cells. For example, *Mll1* gene expression was knocked down by ~50% at eight hours post-transfection, by 66% at 13 hours, but by only ~10% at 18 hours. By 23 hours post-transfection the knockdown was poor, suggesting that the siRNA was degraded and *Mll1* transcript levels had recovered. These results allowed us to choose the 12-13 hours time window as the optimum point for transcript knockdown. Importantly, transfection with *Mll1* siRNA did not affect the abundance of the control gene, ß-actin, which was used to normalise *Mll1* gene expression, suggesting that MLL1 knockdown was specific rather than a global response to the transfection technology.

7.1.2. Knockdown of MLL1 protein

The reduction of *Mll1* gene expression was confirmed by assessing MLL1 protein abundance at 15 (Data not shown) and 18 hours post-transfection (Figure 7.2b). Knockdowns were assessed by western blot using a MLL1 C-terminal antibody (*Upstate/Millipore*). From these two time points, the optimal time for MLL1 protein knockdown was shown to be at 18 hours post-transfection (Figure 7.2b), with a 77% protein knockdown obtained, suggesting it takes several hours after transcript depletion (maximal at 13 hours post-transfection) to be reflected in protein abundance. Nonetheless, this represents an experimentally useful level of MLL1 knockdown, which we subsequently used to examine MLL1 function.



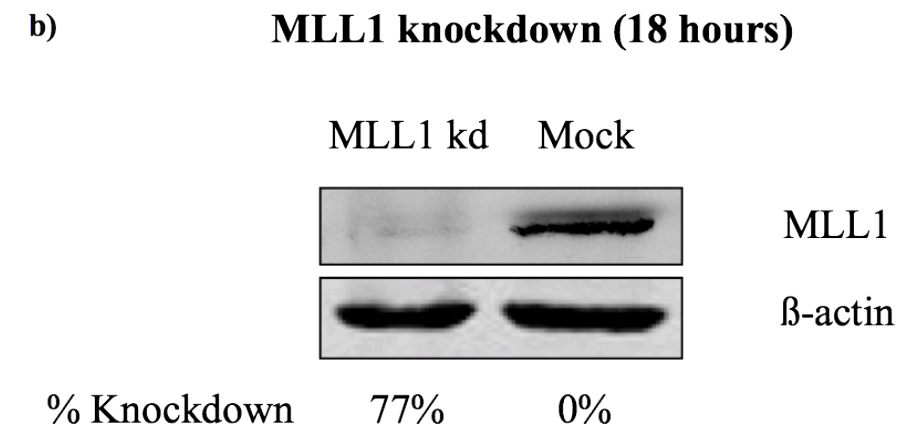


Figure 7.2: Assessing MLL1 knockdown: transcript and protein levels

a) *Mll1* expression was quantified over a time scale after transfection (8, 12, 13, 18 and 23 hours). HPC-7 undifferentiated cells were transfected with 2 *Mll1* siRNA vectors (kd1 and kd2). Gene expression was normalised with ß-actin gene expression (set at 1). Knockdowns are relative to the mock (transfection with a vector containing GFP). Optimal *Mll1* knockdown (66%) was observed after around 13 hours of transfection.

b) *Mll1* knockdown were examined at a protein level, 18 hours after transfecting HPC-7, with GFP (Mock) or with siRNA vectors (MLL1 knockdown). The percentage MLL1 knockdown was assessed by Western blotting with an anti-MLL1^{c} antibody (*UpState/Millipore*), using ß-actin as a loading control. A 77 % knockdown was obtained.

7.1.3. *Mll1* knockdown: Impact on MLL1 target genes

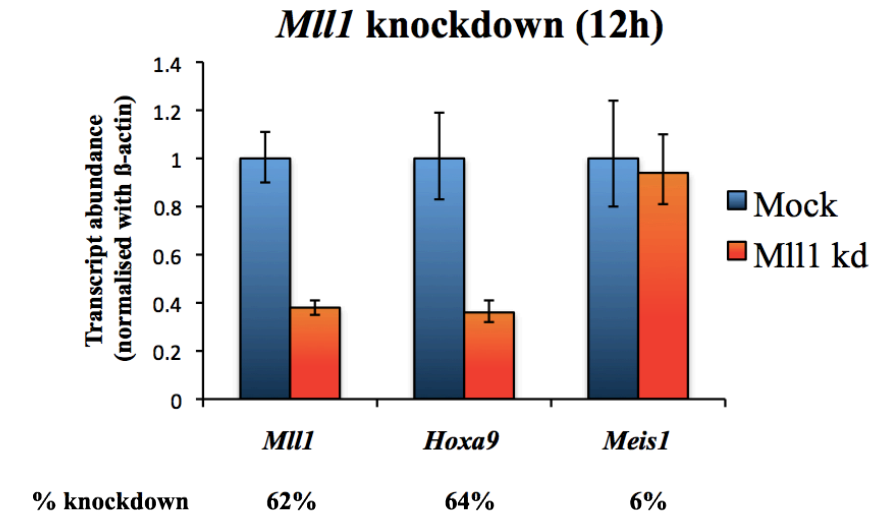
MLL1 is involved in *Hoxa* gene expression. A reduction of *Hoxa9* gene expression has already been shown after transfection with *MLL1* siRNA in HeLa cells, reflecting reduced H3K4 methylation on the *Hoxa9* promoter (Nakamura *et al*., 2002). This gene and *Meis1*, another MLL1 target gene (Milne *et al*., 2005), are up-regulated by MLL-fusion proteins. Thus, a reduction of *Hoxa9* and *Meis1* gene expression was expected after siRNA-mediated MLL1 knockdown.

Initial studies examined *Hoxa9* gene expression after 12 hours of transfection with *Mll1* siRNA vectors: a 64% reduction in *Hoxa9* expression was observed, which is consistent with the level of MLL1 depletion (62% knockdown) (Figure 7.3a). In contrast, *Meis1* gene expression shows no substantial difference at first glance. This difference between *Hoxa9* and *Meis1* sensitivity to MLL1 knockdown may reflect two underlying causes: *Meis1* may be more stable, but it also is a substantially more abundant transcript (*Meis1* Ct value is around 22 while *Hoxa9* Ct value is around 36). Assuming that *Hoxa9* and *Meis1* primers efficiency and sensitivity is the same, Ct values are therefore comparable; the results suggest that the impact of *Mll1* knockdown on *Meis1* transcripts is also substantial. Following this argument, Ct values indicate that there are 32,000 more transcripts of *Meis1* than *Hoxa9*. Therefore, a 60% depletion of *Hoxa9* transcript levels is substantially smaller at absolute transcript levels than a 6% reduction in *Meis1*. This small effect on *Meis1* levels suggests that MLL1 knockdown is insufficient for abundant MLL1-driven transcripts.

7.1.4. *Mll1* knockdown: impact on other MLL family member transcripts

The observation that MLL1 is a member of a family of related methyl-transferases, which all contain chromatin-binding and SET domains, suggests there may be some functional

a)



b)

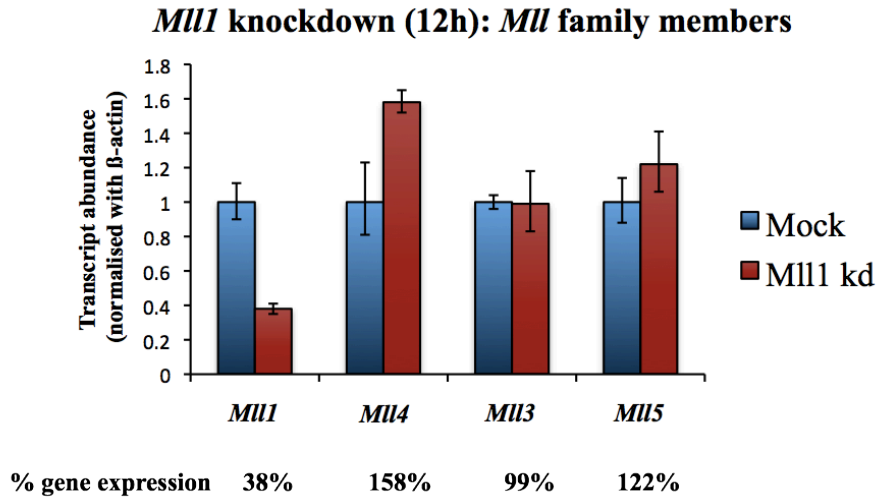


Figure 7.3: Gene expression after *Mll1* knockdown (12 hours post-transfection)

Quantification of *Hoxa9*, *Meis1* and *Mll* family gene expression was examined 12 hours after transfection with *Mll1* siRNA. Gene expression was normalised with ß-actin gene expression (with ß-actin expression set at 1). Knockdowns are relative to the mock (transfection with GFP). PCR was performed in triplicate from one experiment. Error bars represent the standard error of the mean.

a) MLL1 target gene transcript abundance: *Mll1* transcript, and *Hoxa9* transcripts are reduced at 12 hours post-transfection (36%), whereas *Meis1* shows minimal changes.

b) MLL familly members transcripts: MLL1 knockdown is associated with upregulation of *Mll4* (158%), whereas *Mll3* and *Mll5* transcripts remain unchanged.

redundancy between the different family members. To assess this, the impact of *Mll1* knockdown on the transcript abundance of other MLL family members (i.e. *Mll4*, *Mll3* and *Mll5*) was then examined (Figure 7.3b). We observed that conditions that lead to a 60% knockdown in *Mll1* transcripts are associated with minimal changes in the abundance of *Mll3* and *Mll5,* two genes coding for proteins with substantial differences with MLL1. In *contrast Mll4*, a gene coding for a protein homologous to MLL1 and which forms a similar complex, is upregulated by ~1.6-fold. This observed enhancement suggests that MLL1 and MLL4 functionally interact at the same level of gene regulation.

Therefore, using siRNA technology, *Mll1* knockdown was optimised, with an optimal time point at around 13 hours for transcript downregulation and at around 18 hours post-transfection to maximise protein depletion. In both cases, about 70% *Mll1* knockdown was obtained. This knockdown has functional consequences, as the expression of *Hoxa9*, a MLL1-regulated gene, is subsequently decreased (to 36%). Nevertheless, *Meis1* expression, another MLL1 target, is unchanged at 12 hours. This may reflect the time point used being too early to observe changes at *Meis1* expression. Other MLL family members were investigated for their expression in *Mll1* knockdown cells: *Mll3* and *Mll5* expression remained unchanged though a significant increase of *Mll4* expression was observed (158%). Therefore, *Mll4* may be down-regulated by MLL1. This result could explain *Meis1* expression maintenance: *Meis1* could be positively regulated by MLL4 in the absence of MLL1. Therefore, this shows that MLL1 contributes to the regulation of specific genes, and that its presence has an effect on other potential factors of gene regulation.

7.2. *Mll3* and *Mll5* knockdowns

A similar approach has also been used to establish knockdowns for other MLL family members: MLL3 and MLL5. Three vectors were designed for each *Mll* type (*Mll5*-1, *Mll5*-2, *Mll5*-3, *Mll3*-1, *Mll3*-2 and *Mll3*-3) and cloned into "Block-it" vectors (*Invitrogen*). A first step was to examine knockdown efficiency and to identify optimal conditions for the knockdowns (Figure 7.4). For each *Mll* gene, 2 µg of each vector was transfected, either individually or with the other two vectors (2µg in total). As a control, HPC-7 cells were transfected with 2 µg of a vector expressing GFP ("mock"). *Mll3* and *Mll5* knockdowns were quantified by semi-quantitative PCR by using [^{32}P]-dCTP incorporation, and normalised with the "mock" transfection. The knockdown of *Mll3* and *Mll5* transcripts were assessed at 8 and 24 hours post-transfection. Optimal knockdown was established with the *Mll5-2* vector (73.9% knockdown at 24 hours post-transfection) and the *Mll3-1* vector (83% knockdown at 8 hours post-transfection). These vectors were then used to perform subsequent knockdowns.

We found a time-window between six and ten hours post-transfection for *Mll3* knockdown (optimal at eight hours - 72%, Figure 7.5a), and between 12 and 24 hours post-transfection for *Mll5* knockdown (70%, Figure 7.5b). It was then intended to assess the impact of these knockdowns on other *Mll1* gene expression as well as *Hoxa9* and *Meis1* gene expression. At the time these experiments were performed, we were lacking MLL3 and MLL5 antibodies in order to confirm these results at the protein level.

Finally, the level of cell death was assessed for each *Mll* knockdown (Data not shown). Both Mock and *Mll* siRNA transfected cells, display a high percentage of cell death, which increases during the 12 first post-transfection hours (30-50% cell death) before reaching a

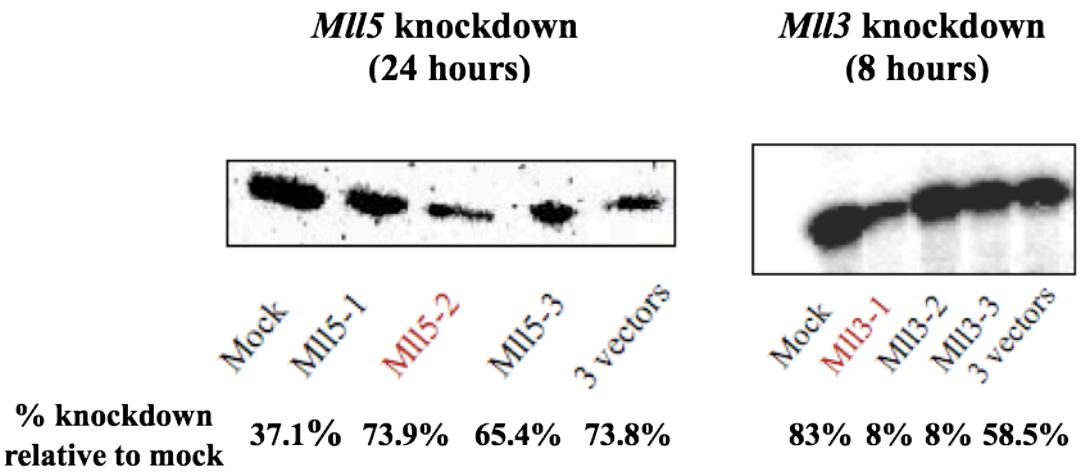


Figure 7.4: Optimisation of knockdown vectors in HPC-7 cells and transfection

Knockdowns for both Mll3 and Mll5 were examined for each vector individually or all together, normalised with the mock (transfection with GFP). Knockdowns were quantified by semi-quantitative PCR, using [^{32}P]-dATP incorporation. A 73.2 % knockdown was obtained with the Mll5-2 vector after 24 hours of transfection, while a 83% knock-down was obtained with the Mll3-1 vector after eight hours of transfection.

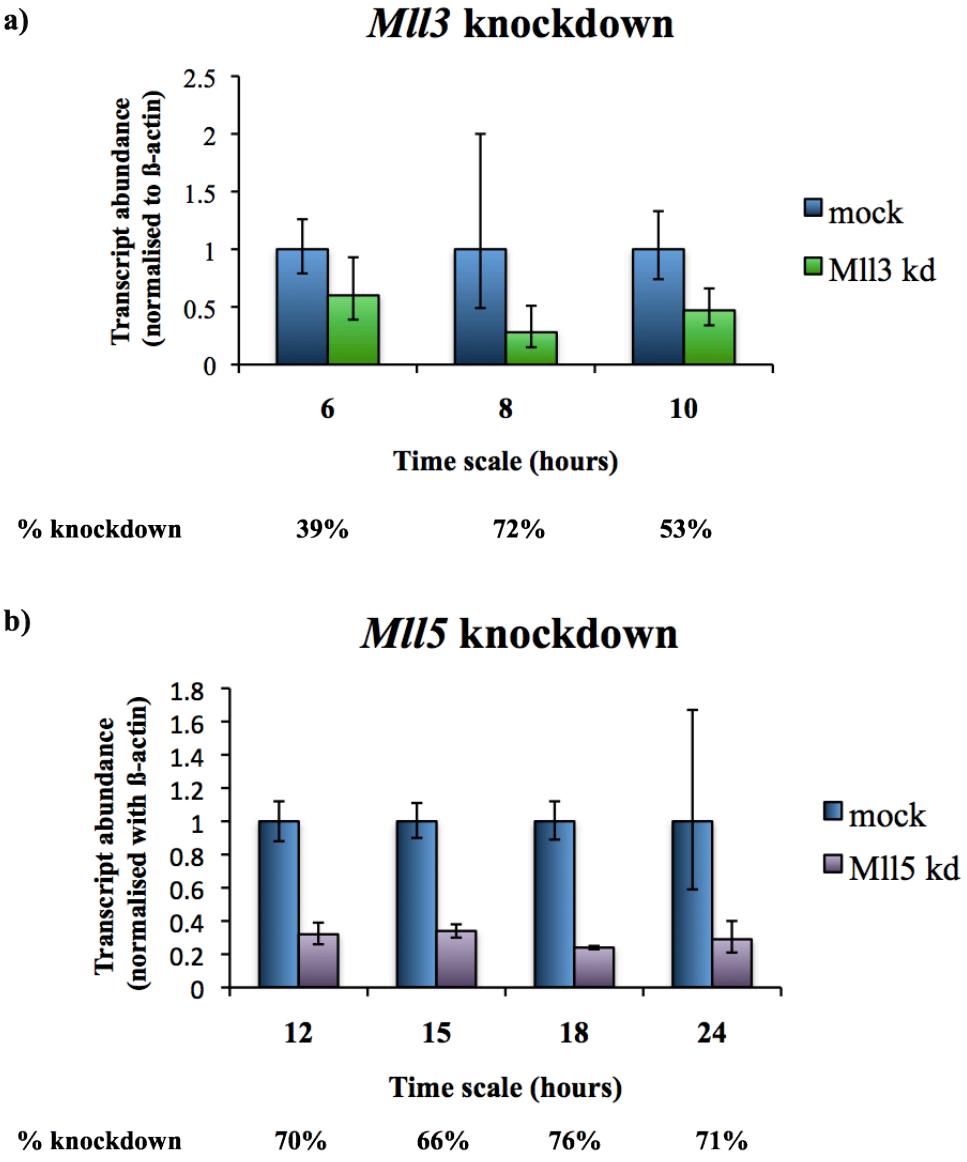


Figure 7.5: Establishing *Mll3* and *Mll5* knockdown conditions

A time-scale was established to find the optimum time point for *Mll3* and *Mll5* knockdown. Quantifications were normalised with ß-actin (set at 1). Knockdowns are relative to the mock (transfection with GFP).

a) *Mll3* transcript abundance was examined by real-time PCR after 6, 8 and 10 hourspost-transfection transfection. An optimal knockdown was obtained at 8 hours post-transfection (72%).

b) *Mll5* transcript abundace was examined by real-time PCR after 12, 15, 18 and 24 hours of transfection. Optimal knockdown is on extended period from 12 to 18 hours post-transfection;

plateau. However, this seems to be highly variable, as only about 25% cell death was observed after 18 hours transfection for the last experiment (see previously).

As with *Mll1* knockdown, efficient *Mll3* and *Mll5* knockdowns were achieved. The optimisation of another MLL family member knockdown was trialled, but failed. However, these experiments on MLL family members are just the premise to what could be done to assess their contribution to gene regulation. The assessment of level of protein after performing these knockdowns also needs to be examined.

Because the transfection rate is at 81% and we cannot get a higher percentage of knockdown than 70%, it might be that the decrease of transcript abundance we expect is masked by transcripts from non-transfected cells. In order to optimise these knock downs, it was suggested to transfect HPC-7 cells with a vector containing both the sequence generating siRNA against *Mll* and GFP. The transfected cells would then be collected by flow cytometry, GFP-transfected cells kept. This would therefore increase the percentage of knockdown observed, and increase the quality of the experiment. Nevertheless, this also implies that a smaller yield of cells would be collected after sorting. Therefore, more HPC-7 cells would need to be transfected.

8. DISCUSSION

The expression of genes is cell-specific, time and environment dependent, and requires the contribution of multiple genetic and epigenetic mechanisms. More specifically, differentiation requires the coordinated activity of multiple genes to determine their cell fate. In this study, we aimed to elucidate the role of a key developmental regulator, the histone methyl-transferase MLL1 and its core complex components on target gene regulation.

Establishing and characterising a differentiation model

Previous studies have shown that MLL1 is part of a multi-protein complex (Nakamura *et al*., 2002; Yokoyama *et al*., 2004), whose subunits catalyse the deposition of methyl and acetyl groups on specific histone residues, such as H3K4me3 catalysed by MLL1 and H4K16ac catalysed by MOF (Dou *et al.,* 2005). MLL1 targets several genes (Wang *et al*., 2009), in particular the *Hoxa* genes (Milne *et al*., 2002). However, little is known about the process(es) of gene regulation to which MLL1 contributes. The protein composition of the MLL complex is diverse and complicated, and still only poorly understood. Furthermore, studies characterising differences between actively transcribed and silent MLL target genes have generally focussed on comparing two separate cell lines (such as wild type cells versus *Mll*-deficient cells or leukaemic cells), without looking at the dynamic changes on MLL target genes that occur in a single cell type. We therefore developed a haematopoietic differentiation model to give insight into the differentiation-induced transcriptional changes at MLL target genes. This is an appropriate model as MLL1 has a critical role in haematopoiesis (Yu *et al*., 1995; Hess *et al*., 1997). We made use of HPC-7 cells, a cell line derived from haematopoietic stem cells, in which differentiation can be induced to generate a range of differentiated cell types, focussing on megakaryocytes as these cells are easy to

purify to a homogenous population in quantities that permit biochemical analysis. Upon megakaryocyte differentiation, morphological and epigenetic changes occur, simultaneously with the appearance of characteristic large cells. Similarly, megakaryocyte nuclei show an increase of ploidy, due to endoremitosis; DNA is replicated, followed by incomplete mitosis due to the absence of cytokinesis and karyokinesis (Vitrat *et al*., 1998). Thus from 2N undifferentiated cells we were able to obtain a "pure" population of differentiated cells with ploidy ranging from 16N to 32N (i.e. Figure 3.1). The degree of cell maturity was also validated by the detection of specific control genes (i.e. Figure 3.2), such as *Gata-1* and *Gp6*, which are exclusively expressed in mature megakaryocytes (Jandrot-Perrus *et al.*, 2000; Holmes *et al*., 2002). In contrast, down-regulation of the *Spi-1/Sfp-1* gene (PU.1, a haematopoietic transcription factor and marker of early development), was observed upon differentiation (Valledor *et al*., 1998). Together, these results indicate that morphologically and transcriptionally distinct megakaryocytes can be obtained through differentiation of HPC-7 cells.

Our use of megakaryocytes raised the question whether all copies of the genome in these cells were equally regulated. This is a key question as subsequent ChIP analysis will "pool" and average marks on the multiple copies of the gene examined. We thus used immunostaining to examine the histone modification distribution in these giant cells. Megakaryocyte nuclei are characterised by multiple lobes, and vacuoles. As expected, undifferentiated cells show "all-over" staining throughout the nuclei for most histone modifications (i.e. Figure 3.5, *Upper panels*), though H3K9me shows a "punctate" appearance (i.e. Figure 3.5c). In contrast the megakaryocytes show an uneven distribution of some histone modifications, with high concentrations in both "vacuoles", and areas of the nuclei where material is "budding off" from the bulk of the nucleus (i.e. Figure 3.5, *Lower*

panels). This was observed after fixing both prior to, or after immunostaining, and a similar observation was made in megakaryocytes generated from primary stem cells suggesting this is not a protocol-dependent, or HPC-7-associated artefact. Interestingly, these observations are only associated with some marks (H3K9ac, H3K4me3), the "punctate" distribution of H3K9me is similar in both undifferentiated cells and megakaryocytes. Of these marks, H3K9ac is known to be associated with newly deposited chromatin (Kuo *et al*., 1996), so we can speculate that these higher concentrations of histones are associated with stores of histones required for generating new copies of the genome. This suggests that not all copies of a gene are in equivalent environments, however the small numbers of cells with these "buds" (~10%) suggests this is a relatively minor issue (Note: this will be discussed later).

Previous studies have suggested that differentiation is associated with global changes in histone modifications (Rice *et al*., 2007). Therefore, a 2-dimentional SDS-AUT gel was performed to give a global overview of the changes upon differentiation and to identify potential general changes in the histones. This analysis showed there were no large changes in histone isoform though a novel band was observed in undifferentiated HPC-7 cells but not in megakaryocytes (i.e. Figure 3.8), which is likely to be associated with H2B acetylation. Mass spectrometric analysis of this histone "spot" suggests that this contains a combination of modifications, including acetylation at novel residues (*K. Nightingale & Helen Cooper, personal communication*). We then performed Western analysis to assess whether changes actually occurred in this system. Comparison of undifferentiated HPC-7 and megakaryocytes showed few significant changes in the global abundance of histone modifications, though a small increase of H3K4me2 and H3K4me1 was seen (i.e. Figure 3.9). It is difficult to speculate on the molecular basis of this observation, as these changes could reflect differentiation-induced changes in the abundance (or activity) of multiple enzyme complexes,

including H3K4-specific methyl-transferases (i.e. MLL family members), demethylases (i.e. LSD1, JMJD2), or the histone modifications that impact on these enzymes' activity. Nevertheless this could be consistent with a decrease in MLL1 activity as though MLL1 itself does not show major changes between undifferentiated HPC-7 and megakaryocytes (i.e. Figure 5.1), whereas the DNA recruiting subunit menin does (discussed later). These studies also showed a significant global increase of H3S10p in megakaryocytes. This is not surprising as the mark is involved in two distinct processes: in transcriptional activation (H3S10 phosphorylation is followed by H3K14 acetylation) (Clayton *et al.*, 2000; Lo *et al.*, 2000; Hans and Dimitrov, 2001), but also in chromosomal condensation during mitosis (Gurley *et al.*, 1978; Hendzel *et al.*, 1997). Thus, this mark reflects the high proportion of chromatin involved in the endomitosis process in megakaryocytes.

Looking at more specific changes, we examined the epigenetic modifications on defined genes upon differentiation, notably on MLL target genes in the *Hoxa* gene cluster. Initial characterisation of *Hoxa* cluster gene expression showed a broad decrease in *Hoxa* genes (1-13) transcript abundance upon differentiation (i.e. Figure 3.3), correlating with observations made in other models (Sauvageau *et al*., 1994). This decrease is more significant in the middle of the cluster (i.e. *Hoxa5)*, rather than at the edges (i.e. *Hoxa1* and *Hoxa13*). Interestingly, these central genes, especially within the region from *Hoxa4* to *Hoxa11,* are implicated during differentiation (Strathdee *et al*., 2007a; Strathdee *et al*., 2007b; Atkinson *et al*., 2008). Although we observed a substantial down-regulation at many *Hoxa* genes, we did not see a total loss of expression upon differentiation. This may reflect the high stability of their RNA products, as we find that the *Hoxa* gene transcripts are stable for up to 24 hours in α-amanitin treated cells. However, this down-regulation is sufficiently significant to be considered as a cause (or consequence) of differentiation, and to investigate the associated

mechanisms. Together, these observations suggest that this is an appropriate model to study the regulation of differentiation-induced changes at MLL1 target genes.

This is a powerful model, but there are some limitations to its use here: this study focussed on the changes occurring at two extreme stages of the differentiation process, but it would be interesting to investigate changes occurring throughout the differentiation process in order to obtain a more accurate overview of mechanisms involved in the regulation of gene expression. Similarly, the processes characterised with this model could be megakaryocyte-specific, and ideally would be validated by characterising the changes associated with a different cell lineage. To this purpose, monocytes were successfully generated in the laboratory. However, it has not been practical to examine these cells, as too few monocytes were obtained to enable biochemical studies.

Characterising histone modification changes on Hoxa4 and Hoxa5

Once the model was established, we examined the epigenetic changes occurring at specific genes in response to the differentiation process. The histone methyl-transferase MLL1 is an important regulator of haematopoiesis, especially through its role *Hoxa* gene maintenance (Ernst *et al.*, 2004a; Ernst *et al.*, 2004b; Rice *et al.*, 2007). For this reason, we chose to examine changes occurring on *Hoxa* genes. In numerous studies, the most examined *Hoxa* gene is *Hoxa9* as a target of MLL fusion proteins; *Hoxa9* is upregulated in leukaemia as a consequence of the generation of MLL-fusion protein (Zaisig *et al.*, 2004). However, in our model, *Hoxa9* transcript levels are too low in abundance to be accurately examined (i.e. Figure 3.3). In contrast we focussed on *Hoxa4* and *Hoxa5*, two genes with abundant transcripts, which are known to be relevant to the differentiation process. Our data initially suggested that these two genes act differently upon differentiation; *Hoxa5* showed a

significant down-regulation (i.e. a four-fold drop), wheras *Hoxa4* did not show a significant reduction in transcriptional output. *Hoxa4* is the most 3' gene in the *Hoxa* cluster to be involved during differentiation (Sauvageau *et al*., 1994), and *Hoxa5* down-regulation is required for the generation of megakaryocytes (Crooks *et al.*, 1999; Fuller *et al.,* 1999). Thus, although we know that MLL-driven *Hoxa* genes are essential for haematopoiesis, little is known about how MLL regulates these events, and how this contributes to differentiation. Therefore, the distribution of histone marks on the *Hoxa4* and *Hoxa5* genes during differentiation was examined. We examined marks associated with either transcriptional activation (H3K4me3, H4K16ac and H3K9ac) or transcriptional silencing (H3K27me3 and H3K9me2). In order to investigate the distribution of these marks, two forms of chromatin immunoprecipitation were used: carrier chromatin immunoprecipitation (CChIP) and native chromatin immunoprecipitation (NChIP). NChIP is principally used to assess histone mark modifications, as the histones are firmly attached to the DNA. However, this technique has been developed for a large number of cells (typically more than 10^6), whereas this stem cell system can only yield small numbers of megakaryocytes (~10^4). Therefore, CChIP, a technique that uses *Drosophila* SL2 cells during the preparation of chromatin, was used in order to improve the efficiency of immunoprecipitation. However, this raised the question of the quantitative comparability between those two ChIP procedures. Thus, an initial experiment compared the two approaches, using a *Gapdh* housekeeping gene that shows similar transcript abundance in the two cell types. We found that the binding values observed for marks over the housekeeping gene *Gapdh* are both qualitatively and quantitatively similar (i.e. Figure 4.2), suggesting that the two techniques can be used to compare undifferentiated HPC-7 cells and megakaryocytes. We thus examined the distribution of histone marks on *Hoxa4* and *Hoxa5* in these two cell types.

Analysis of these data (i.e. Figures 4.3; 4.4; 4.5) shows a consistent pattern of distribution of activating marks (H3K4me3, H3K9ac and H4K16ac), with a peak at the transcription start site of the genes, but with minimal binding both upstream of the gene, or within the body of the coding region. This is consistent with earlier studies, though we note that ChIP-chip studies often yield broader peaks, perhaps reflecting the size of immunoprecipitated material (Atkinson, Koch *et al*., 2008). In undifferentiated HPC-7, *Hoxa* genes were associated with an enrichment of H3K4me3, H3K9ac and H4K16ac close to the transcription start site. These "activating" marks were subsequently lost in megakaryocytes, which is associated with a 12-fold and a 20-fold decrease of H3K4me3 on *Hoxa4* and *Hoxa5* respectively, and a three-fold decrease of H4K16ac on *Hoxa4*. This loss of "activating" marks is consistent with the repression of *Hoxa5* upon differentiation (i.e. Figure 3.3), however it is suprising for *Hoxa4*, which does not show a significant decrease in transcriptional activity. This suggests that *Hoxa4* is in fact down-regulated, but that this is masked by *Hoxa4* transcript stability. Interestingly, the peak of H3K9ac is also lost on *Hoxa4* in megakaryocytes, but to be replaced by a broader distribution of acetylation in the body of the gene.

The high level of H3K9 acetylation retained on the *Hoxa* genes in megakaryocytes is surprising, given that this is a mark typically associated with gene activation (i.e. Mizzen *et al.*, 1996). However, this unexpected distribution toward the body of the genes is reminiscent of the H3K9me2 distribution (i.e. Figure 4.7) on *Hoxa4*, which is also in the body of the gene. Neither of these apparently antagonistic marks are at the transcription start site of the gene - possibly reflecting the unclear, transcriptional status of this gene in megakaryocytes (i.e. Figure 3.3), and/or that the cells are in a intermediate state, carrying both marks in the body of the gene when examined. The fact that H3K4me3 and H4K16ac (i.e. Figures 4.3; 4.5 respectively) are decreased on *Hoxa4* in megakaryocytesis is consistent with its down-

regulation, and also suggests that these changes are temporally regulated. However, in addition to this change of distribution, there is an increase of H3K9 acetylation. This mark is also associated with newly synthesised chromatin and this broad distribution may reflect the "pooling" of the marks on *Hoxa4* and *Hoxa5* from different regions of the nucleus in different contexts, in which *Hoxa4* and *Hoxa5* are in both silent chromatin regions and replicative chromatin (Figure 8.1).

This analysis also showed that active *Hoxa4* and *Hoxa5* promoters are associated with H3S10p, a mark associated with actively transcribed genes (Hazzalin and Mahadevan, 2005; Ivaldi *et al.*, 2007), and the dual mark H3K9acS10p which is also associated with early gene activation (Dyson *et al.*, 2005). This mark was examined on MLL-target genes as previous *in vitro* studies showed that a combination of H3K9 acetylation and H3S10 phosphorylation stimulated MLL SET domain activity (Nightingale *et al.,* 2007). In mammals, H3S10 phosphorylation favours the recruitment of Gcn5 (Cheung *et al*., 2000). In addition, it was shown in yeast that, as part of the SAGA complex, Gcn5 deposits H3K9ac (Grant *et al.*, 1999), a mark seen on the *Hoxa* genes studied. We show that the dual mark H3K9acSer10p has a similar distribution to H3K9ac or H3K4me3 on the active *Hoxa4* and *Hoxa5* genes (i.e. Figures 4.4; 4.3, *Upper panels*): a peak of H3K9acS10p is observed near the transcription start site in undifferentiated cells (i.e. Figure 4.8), which is subsequently lost in megakaryocytes. This similarity in the distribution of H3K4me3 and H3K9acS10p confirms that the mark is involved in *Hox* gene regulation, suggesting that is might be involved in regulating MLL target genes, consistent with its role on other genes. However whether it plays a role in regulating MLL activity as observed *in vitro* remains unclear. It is curious that only the dual mark H3K9acS10p had an influence on the SET domain activity, in contrast to individual marks that had no noticeable effect (Nightingale *et al.,* 2007). However, previous

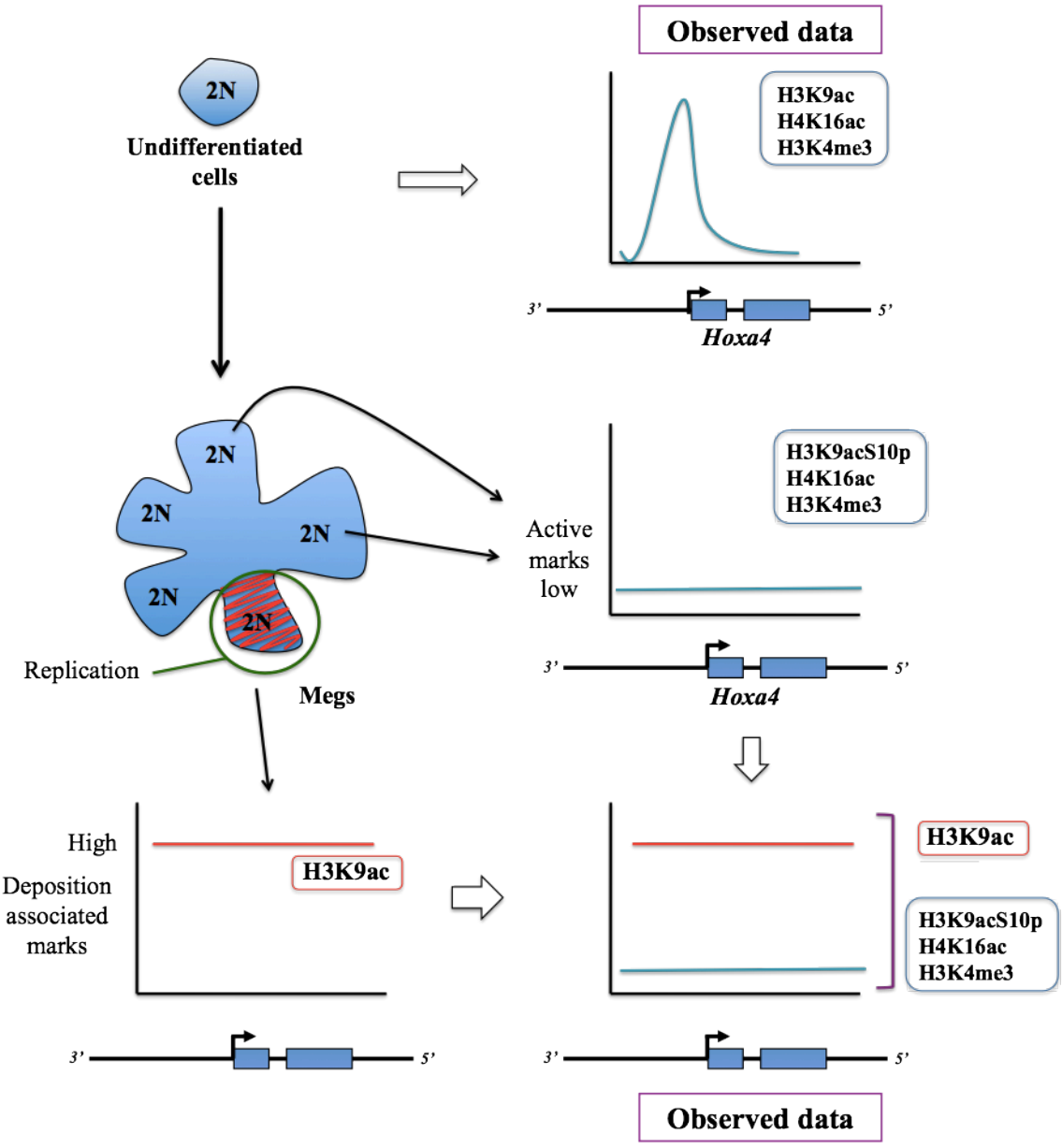


Figure 8.1: Hypothesis to explain our observation of a surprising H3K9ac distribution:

H3K9ac is a mark associated with transcriptional activation (as seen in undifferentiated cells). However, an increase of H3K9ac was observed in megakaryocytes on the silent *Hoxa4* and *Hoxa5* loci, while other activating marks were reduced. The mammalian H3K9ac mark is associated with newly synthesised chromatin, so the broad distribution observed may reflect the "pooling" of the marks on *Hoxa4* and *Hoxa5* from different regions of the nucleus in different contexts (i.e. in silent chromatin regions and replicative chromatin).

studies have shown the recognition of multiple marks by the same protein (for example, TAF1 binds H3K9ac and H3K14ac through its double bromodomain, (Vermeulen *et al.*, 2007)) suggesting that the cooperative recognition of H3K9ac and H3S10p may act to stabilise catalytic activity at H3K4me3. H3K9ac and H3S10p are adjacent residues, and this also suggests that these marks are simultaneously recognised to regulate MLL SET domain activity. This is consistent with previous data from structural studies (Southall *et al*., 2009).

The pattern of distribution of "activating marks" correlates with previous studies (Hebbes *et al.*, 1988; Vakoc *et al*., 2006; Wang *et al.,* 2008), and their decrease in megakaryocytes is consistent with the observed (*Hoxa*5), and assumed (*Hoxa4*) down-regulation upon differentiation (i.e. Figure 3.3). The expected removal of "activating" marks following gene repression raised the question whether "silencing" marks also contribute to the change in transcriptional activity. However, these marks are not abundant on *Hoxa* genes. H3K27me3 is broadly distributed on both *Hoxa* genes, but shows a very low level of enrichment (i.e. Figure 4.6), whereas H3K9me2 shows a slight enrichment over the promoter and the coding region (i.e. Figure 4.7). This pattern is replaced by a broad distribution in megakaryocytes, notably at the 5' end of the gene, similar to what was observed with H3K9ac. The absence of the H3K27me3 mark on repressed genes is surprising and contradicts what would have been expected; an enrichment associated with repressed transcription (Barski *et al*., 2007). Similarly, investigation of the presence H3K27me3 on a gene known to be silent in undifferentiated HPC-7 cells (i.e. *Gp3*) failed, despite the antibody being known to work well in ChIP assays. This may reflect megakaryocyte-specific mechanisms; there may be a low level of this mark in these cells. However, global staining with H3K27me3 was observed in HPC-7 in immunofluorescence assays (i.e. Figure 3.5d). Alternatively, this may reflect a regulatory role for CBP, a known subunit of the MLL complex, which is known to target

H3K27. Recent studies showed that H3K27me deposition was prevented by acetylation of the same residue (Tie *et al*., 2009), suggesting that CBP-deposited marks may persist into the megakaryocytes. Examination of H3K27ac and CBP distribution upon differentiation may help to understand the failure of H3K27me3 to increase with *Hoxa5* gene silencing. A final alternative hypothesis could be that this "repressive" mark is deposited only on long-term silenced genes, and that our examination of this mark might be too premature.

One aspect of these data is that they do not give indications to the relative abundance of the different gene transcripts. This is difficult to assess with any accuracy, but if we assume that qRT-PCR Ct values give a broad indication of transcript abundance then *Hoxa4* (Ct value ~ 28) and *Hoxa5* (Ct value ~ 23) are far more abundant than *Hoxa9* (Ct value ~ 36). This is consistent with the preliminary analysis of CpG island methylation in differentiating HPC-7 cells made by Gordon Strathdee (*Newcastle*; *Data not shown*): the *Hoxa5* and *Hoxa4* CpG islands are unmethylated whereas the *Hoxa9* CpG island was methylated in both undifferentiated cells and megakaryocytes. This is consistent with the low *Hoxa9* transcript abundance, as promoter methylation is associated with an inactive state of the gene. The absence of methyl marks on CpG islands on the down-regulated *Hoxa5* gene *(and Hoxa4,* which we assume is also down-regulated) could indicate that the megakaryocytes were examined too early, before the establishment of these marks. This is consistent with the absence of H3K27me3 on repressed *Hoxa5 (*and apparently repressed *Hoxa4)*. Studies of primary cells should establish whether this is a cell-line, or cell-type specific issue, as the CpG islands of the *Hoxa* genes are methylated when repressed in other cell lines (Strathdee *et al.*, 2006; Strathdee *et al*., 2007b).

In summary, differentiation correlates with both observed (*Hoxa5*), and assumed (*Hoxa4*) gene down-regulation, which is associated with the loss of "active" marks in the gene

promoters. However, gene repression did not correlate with either an increase of "silencing" marks, or an increase of DNA methylation on *Hoxa4* and *Hoxa5*. We therefore examined the enzyme complexes responsible for regulating these marks.

How are histone modifications regulated?

The distinct histone marks described above are associated with specific regulatory enzymes. We therefore initially examined the global abundance of the key enzyme regulators of these marks, and we then investigated these proteins at a gene-specific level using immunoprecipitation of formaldehyde cross-linked chromatin.

In mammals, six proteins are known or assumed to have H3K4-specific histone methyl-transferase activity: Set1A, Set1B, MLL1, MLL2, MLL3 and MLL4. During megakaryocyte differentiation, as with other active marks, H3K4me3 abundance is rapidly lost from the *Hoxa4* and *Hoxa5* promoters, correlating with their observed (*Hoxa5*), and presumed (*Hoxa4*), down regulation. This does not reflect changes in MLL family transcripts (i.e. Figure 3.9), or in global MLL1 protein abundance upon differentiation (i.e. Figure 5.1, *Upper panel*). This suggests that in megakaryocytes, MLL1 is still present, but unable to maintain *Hoxa* gene expression. In contrast, menin, a component of the MLL complex, shows a substantial drop in megakaryocytes (i.e. Figure 5.1, *Lower panel*). This subunit is believed to act with LEDGF in determining DNA binding of the MLL complex, so this is a potential mechanism to reduce MLL activity without the wholesale removal of the MLL enzyme.

Our analysis of specific genes showed that differentiation is also associated with a decrease of H4K16ac (i.e. Figure 4.5), H3K4me3 (i.e. Figure 4.3) and H3K9acS10p (i.e. Figure 4.8) at the *Hoxa* genes. The simultaneous loss of these marks at the promoter of *Hoxa4* and *Hoxa5* suggests the synchronised release of the associated deposition enzymes or a common

regulatory mechanism. The patterns of distribution of H3K4me3 and H4K16ac correlate spatially and temporally through differentiation, consistent with the association of the histone acetyl transferase MOF (a H4K16 specific HAT) with MLL1 (Dou *et al*., 2005) in the same complex. This raised the question of whether the marks observed are regulated by MLL-associated factors, in addition to MLL1.

After initially optimising X-ChIP, the binding of several MLL1 complex-associated factors that could contribute to mark regulation was examined. A MLL partner, menin was described to associate with MLL1 on *Hoxa* gene promoters in leukaemic cells, and to be required for the maintenance of the expression of these genes (Yokoyama *et al.*, 2005). In this study, both MLL (including both MLL^C and MLL^N subunits) and the MLL-associated factor menin bound upstream to the transcription start site on *Hoxa4* and *Hoxa5*, whereas H3K4me3 was found closer to the transcription start site. Interestingly, menin binding was stronger than MLL1 binding on these genes, which may reflect menin's tighter interaction with the DNA. This correlates with observations that menin recruits the MLL complex to transcriptionally active genes through its interaction with the protein LEDGF (Yokoyama and Cleary, 2008). The loss of either menin or LEDGF prevents MLL1 recruitment. It has also been shown that menin knockdown induces down-regulation of *Hoxa9*, comparable to that observed for MLL knockdown (Yokoyama *et al*., 2004). Interestingly, in contrast to MLL1 abundance, which remains unchanged upon differentiation, menin levels are globally decreased in megakaryocytes (see earlier). Menin could be an important factor in determining MLL1 regulatory function in haematopoiesis.

In contrast to MLL1 and menin, the histone acetyl-transferase CBP binds in the coding region of *Hoxa4*, whereas it is found at the transcription start site of *Hoxa5* in undifferentiated HPC-7 cells (i.e. Figure 5.5). This way, CBP is closer to the MLL-deposited

histone marks observed than menin and MLL1 (i.e. Figure 5.5). This is surprising given that this histone acetyl-transferase is thought to be recruited by the MLL complex (Ernst *et al*, 2001), and would be expected to correlate with MLL^C binding. However, previous studies showed that CBP can also be associated to the phosphorylated p65 subunit of NFκB at the H3S10p site (Reber *et al.*, 2009), so this may be an alternative route to recruiting this HAT to *Hoxa4* and *Hoxa5* (discussed later).

Similarly, RNA polymerase II was found over the transcription start site of both *Hoxa4* and *Hoxa5* in undifferentiated HPC-7 cells (i.e. Figure 5.6), which correlates with a previous study showing that H3K4me is located on transcriptionally active genes, associated with MLL1 and RNA Pol II binding (Guenther *et al*., 2005). In megakaryocytes, RNA polymerase II is maintained at the same location on *Hoxa4*. This is consistent with the unclear transcriptional status of *Hoxa4* in megakaryocytes; the loss of "activating" histone marks suggests that the gene is down-regulated, but this is not apparent from transcript levels. In contrast RNA Polymerase II binding is reduced and moved downstream on *Hoxa5*, which does show clear differentiation-induced repression. The presence of RNA polymerase II on both genes, in particular on *Hoxa5*, whilst their expression is down-regulated suggests that RNA polymerase II may be subject to a pausing mechanism (Strobl and Eick, 1992). This would be consistent with the retention of RNA polymerase II near the transcription start site of *Hoxa4* and *Hoxa5*. Interestingly, H3S10p appears to play a role in the release of promoter-proximal "paused polymerase" in *Drosophila* (Ivaldi *et al*., 2007; Zippo *et al*., 2009), consistent with a role for this mark in this system. In addition, the weak but distinct enrichment of H3K9me2 in the coding region of *Hoxa4* may be acting to regulate RNA polymerase II elongation (Vakoc *et al.*, 2006).

Our data also indicates there are functional interactions between some MLL family members. We show that *Mll1* knockdown is associated with an increase of *Mll4* transcript abundance (i.e. Figure 7.3b). Interestingly, menin is also associated with this MLL histone methyl-transferase complex (Hughes *et al.,* 2004), thus, the loss of MLL1 could be functionally redundant with a similar MLL4-menin (and other components) complex, though they cannot compensate for each other, as the loss of either molecule is embryonic lethal (Yu *et al.*, 1995; Glaser *et al.*, 2006). As with MLL1, MLL4 enzymatic activity is mediated by a SET domain, generated by proteolytic cleavage by Taspase-1 (Takeda *et al.,* 2006), and tri-methylates lysine K4 on histone H3 (Hugues *et al.*, 2004). We present data suggesting that MLL4 is present on *Hoxa4* and *Hoxa5* genes. However, validation of the specificity of this antibody was not convincing, and this would need to be re-examined using another antibody source to give confidence in these data.

To summarise, histone marks and enzymes display distinct patterns of distribution on active and repressed *Hoxa4* (Figure 8.2) and *Hoxa5* (Figure 8.3) genes, consistent with gene activity. On active genes, both genes show an enrichment of H3K4me3, H3K16ac, H3K9ac and H3K9S10p over the transcription start site, correlating with the binding of MLL1, MOF (not examined, but an MLL-associated protein) and Msk1 (see later). However, the binding sites of these enzymes did not always correlate with the sites where histone marks are deposited, especially on *Hoxa4* where the primers are more separated. For example on *Hoxa4*, whilst Msk1 binds over the transcription start site and correlates with the H3S10p mark, MLL-complex proteins (MLL^C, MLL^N and menin) bind upstream of the TSS. Similarly, on *Hoxa5* histone marks are localised at the TSS whereas the depositing enzymes are located just upstream. This suggests that MLL^C, MLL^N and menin interact as a complex, but this is less clear for CBP, as on *Hoxa4* this enzyme is found in the body of the gene.

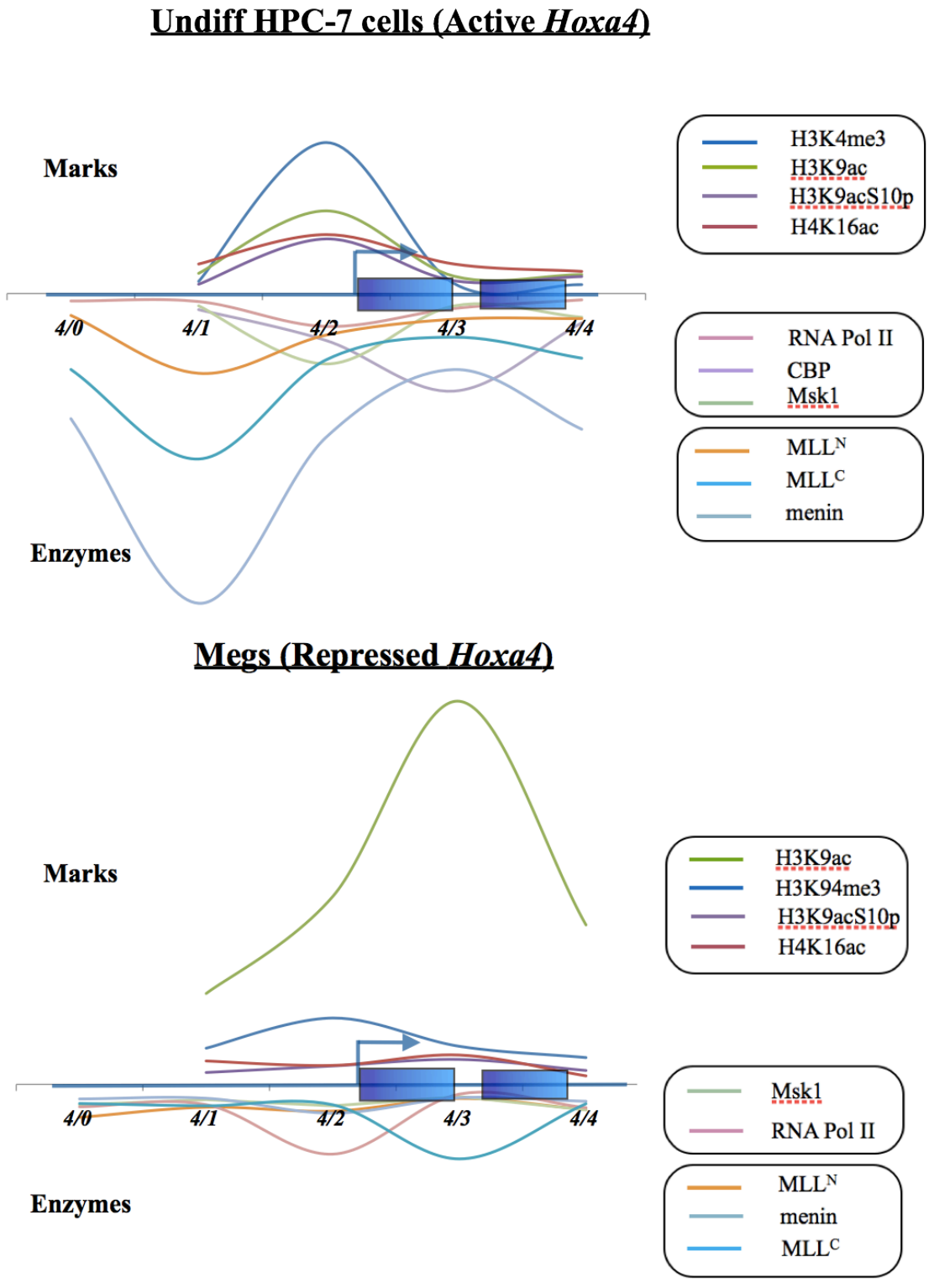


Figure 8.2: Alignment of mark and enzyme distributions on *Hoxa4*

Distribution of histone marks and modifying enzymes on *Hoxa4* in undifferentiated HPC-7 cells (*Top*) and in megakaryocytes (*Bottom*). Note: The relative sizes of peaks are arbitrary as different antibodies cannot be compared

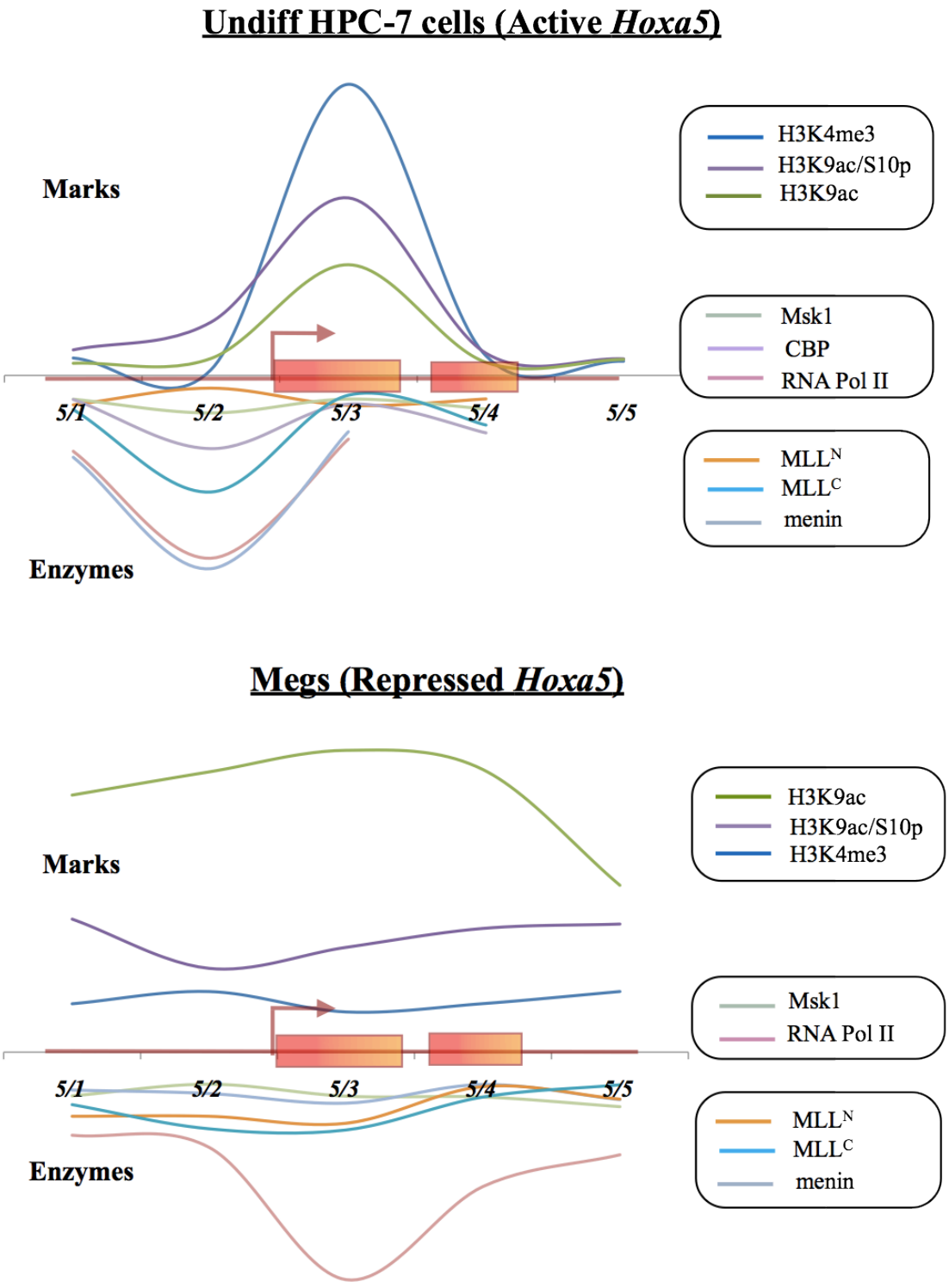


Figure 8.3: Alignment of mark and enzyme distributions on *Hoxa5*

Distribution of histone marks and modifying enzymes on *Hoxa5* in undifferentiated HPC-7 cells (*Top*) and in megakaryocytes (*Bottom*).

Binding at these primer sets (4/1 and 4/2) is separated by ~900bp - this could reflect the presence of a large complex with distinct domains interacting with DNA (and cross-linking) and catalytic sites. Another explanation would be that an element of the promoter looping occurs, allowing distant regions to be associated with different components of a same complex. Studies have shown long-range interactions, in which regulatory elements are brought close to the TSS of target genes causing the formation of DNA looping (Fraser, 2006; Fraser and Bickmore, 2007; Göndör and Ohlsson, 2008; Wallace and Felsenfeld, 2007).

Another alternative to explain the apparent dislocation between modifying enzyme binding sites and the location of the marks is that this reflects the protocol used to generate our chromatin. A recent publication suggests that nucleosomes that simultaneously contain two histone variants, H3.3 and H2A.Z, are less stable, and susceptible to displacement when challenged with the salt conditions we use to isolate chromatin (150mM NaCl, *Section 2.3.4,"'Buffer C"* Jin *et al.,* 2009). This paper finds that "double variant" (H3.3/H2A.Z) nucleosomes typically occupy a position at and 5' to the transcriptional start site (TSS) of active genes, this correlates with the sites where MLL-deposited marks would be expected. Given the lower resolution of our experiments the loss of these (potentially H3K4me / H4K16ac - marked) nucleosomes may explain the apparent absence of these modifications at specific sites in chromatin. However, this is unlikely to be a total explanation; other workers also see a disparity between the sites of MLL binding and H3K4 methylation (*G. Volpe PhD thesis, University of Birmingham*), but use a formaldehyde cross-linking procedure which would be expected to stabilise double variant nucleosomes in the chromatin.

Msk1 and H3S10p mark

As previous *in vitro* assays have shown that the dual mark H3K9acS10p acts to regulate MLL SET domain activity, the novel observation that this mark is found on MLL-target genes *in vivo* (i.e. Figure 4.8) suggested that we analyse Msk1, a protein kinase thought to be involved in H3S10 phosphorylation. This showed that Msk1 plays a role in regulation of *Hoxa4* and *Hoxa5* genes, as our experiments showed a correlation between the binding of Msk1 (i.e. Figure 6.1) and the enrichment of H3K4me3 (i.e. Figure 4.3) and H3K9acS10p (i.e. Figure 4.8) upon the TSS of these genes. Similarly, the kinase is absent when these genes are repressed in megakaryocytes (i.e. Figure 6.1, *Lower panels*). Previous data suggest Msk1 interacts with promoters via interactions with NFκB (Reber *et al*., 2009), but initial scanning for NFκB binding sites (GGG RNN YYC C, where R=purine Y=pyrimidine), did not reveal this sequence in *Hoxa4* or *Hoxa5*. The observation that Msk1 binds a similar location to the MLL complex on these genes prompted our co-immunoprecipitation experiments. Msk1 was found to associate with MLL^C (i.e. Figure 6.2a), though this association seems to be weak. These data suggest that Msk1 and MLL1 may interact in a transitory manner, though the co-immunoprecipitation could also reflect an indirect interaction via the cross-linking of the common histone substrate (Figure 8.4). In addition, Msk1 was shown to associate with NFκB but not with CBP (i.e. Figure 6.2a), whereas Reber *et al.* (2009) showed that p65 (NFκB), CBP and Msk1 co-immunoprecipitated at H3S10-marked loci. The failure to observe any interaction between CBP and Msk1, whilst these two proteins are closely bound on *Hoxa4* and *Hoxa5* genes could be explained by the high molecular weight of CBP (265kDa), which is difficult to immuno-precipitate. Considering the diverse interactions attributed to CBP, this failure could also be the consequence of the immunoprecipitation of a CBP-containing complex. These experiments suggest that Msk1

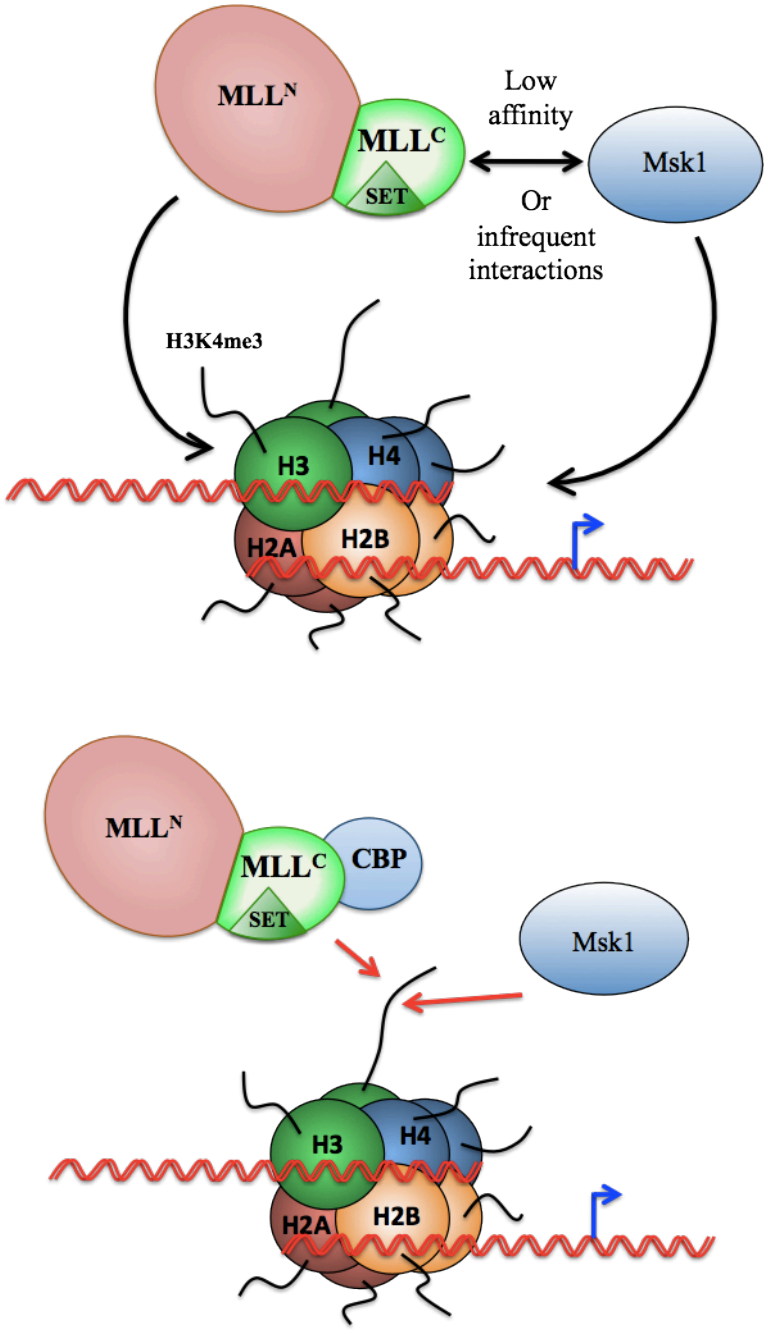


Figure 8.4: Msk1 contribution to MLL target genes

These two cartoons illustrate alternative mechanisms of how Msk1 contributes to MLL target gene regulation, in line with the ChIP and immunoprecipitation results. Msk1 either interacts with the MLL complex in a transitory manner (*Upper panel*), or is present at the same time as MLL complex components on a specific region of the gene (*Lower panel*), and co-immunoprecipitates indirectly, probably via the histone proteins.

might influence MLL1 enzymatic activity and target gene regulation. We note that Msk1 is downstream of the MAPK signalling pathway (Dyson *et al.*, 2005) and recent data suggest that Msk1 phosphorylation contributes to gene regulation (Bruck *et al.*, 2009; Reber *et al.*, 2009). These studies are consistent with the proposal that Msk1 represents a mechanism by which the environment may impact on MLL1 activity.

Our study shows that H3K9acS10p was found on active *Hoxa4* and *Hoxa5* genes, and correlates with sites bound by the histone kinase Msk1. These observations are consistent with previous findings on this mark and this enzyme (Soloaga *et al.*, 2003; Hazzalin and Mahadevan, 2005; Ivaldi *et al.*, 2007; Zippo *et al.*, 2009). However, it was still unclear whether this contributes to *Hoxa* gene regulation, and whether this influences MLL activity, as seen in vitro (Nightingale *et al.*, 2007). We therefore tested whether H3S10 phosphorylation had a direct effect on *Hoxa* gene expression by using small molecule inhibitors to modulate phosphorylated H3S10 abundance on *Hoxa* genes. This abundance was modulated either by increasing H3S10p mark using a phosphatase inhibitor (okadaic acid) or by reducing it using the Jnk kinase inhibitor (SP600125, Huang *et al.*, 2006). However, interpreting these experiments was more complicated than expected. For example, okadaic acid does induce a global increase in H3S10p in chromatin (i.e. Figure 6.3), and this is also observed on the *Hoxa5* gene (i.e. Figure 6.8). However, this effect is not directly translated into an increase in transcription in either undifferentiated cells or megakaryocytes (i.e. Figure 6.6).

This may reflect the role of H3T3 phosphorylation on MLL activity. An interesting result came from the observation of the H3T3p mark after okadaic acid treatment. Whilst no change was observed on *Hoxa4*, a loss of H3T3p was observed on *Hoxa5* (i.e. Figure 6.9), correlating with an increase in H3K9acS10p and H3S10p enrichment on the same gene.

Previous studies suggested an antagonism between H3T3 phosphorylation and H3K4 methylation; as H3T3 phosphorylation induces a decrease of MLL SET domain activity (Southall, Wong *et al.*, 2009). The increased abundance of this repressive mark (i.e. Figure 6.9) may antagonise the increased H3S10 phosphorylation induced on the *Hoxa* genes. This complexity is difficult to analyse, especially as there may be other impacts on other histones and non-histone proteins. Examination of H3K4me3 abundance, and its correlation with the balance of H3K9acS10p and H3T3p on a specific locus may help to dissect the role of these marks. Experiments with the Jnk inhibitor were affected by modifications at a range of residues and were difficult to interprete for the same reasons.

Pathways to gene activation at MLL target genes

These data suggest that Msk1 contributes to *Hoxa4* and *Hoxa5* regulation. However this is likely to be via a complex pathway of events, involving multiple histone marks and regulator proteins. A recent study consistent with a role for Msk1 shows that histone crosstalk between H3S10p and H4K16ac contributes to epigenetic regulation (Zippo *et al*., 2009). This study showed that, under serum stimulation, the 14-3-3 protein binds to phosphorylated H3S10 (consistent with Macdonald *et al.*, 2005) and leads to the co-recruitment of the H4K16 specific acetyl-transferase MOF. The recognition of this dual H3S10p/H4K16ac mark by the double bromodomain protein BRD4 promotes transcriptional elongation. These data, and findings from this project suggest that Msk1 deposition of H3S10p can potentially play an initiating role on the *Hoxa* genes.

In a speculative mechanism (Figure 8.5), we propose that initial Msk1 deposition of a phosphate group on H3S10 is recognised by a complex composed of the 14-3-3 protein and the histone acetyl-transferase MOF (Zippo *et al.*, 2009). This results in the deposition of an

acetyl group on H4K16 at the same location as H3S10p, which is consistent with our data (i.e. Figure 4.5 and 4.8) (An alternative mechanism, in which MOF is recruited by MLL (Dou *et al*., 2005) is also consistent with our observations). In addition, the acetyl-transferase Gcn5 is recruited to either H3S10p (Cheung *et al*., 2000) or H4K16ac (in yeast) (Owen *et al*., 2000), and acts to deposit an acetyl group on H3K9 residue (Grant *et al.*, 1999). The combination of these marks can be followed by two possible pathways. (1) The P-TEFb/BRD4 complex, comprising the bromodomain protein BRD4 and the transcriptional elongation factor P-TEFb, preferentially associates this combination of marks (i.e. H3K9ac, H3S10p and H4K16ac). This complex acts to recruit RNA Pol II (Zippo *et al.*, 2009). An alternative (2) is that the SET domain of MLL recognises and is activated by the dual mark H3K9acS10p (Nightingale *et al*., 2007), such that the level of H3K4me3 is increased. These two pathways are non-exclusive and both may act to regulate a single gene.

Developing MLL knockdown in HPC-7 cells

The studies described give insight into the distributions of histone modification, and correlate this with enzyme binding, according to the transcriptional state of genes. However, to assess the contribution of MLL1 to these mechanisms, we proposed to use MLL1 knockdown to examine the changes occurring as a consequence. Using siRNA technology, MLL1 knockdown was optimised, with a time point at ~18 hours post-transfection found to maximise protein depletion (~70% MLL1 knockdown obtained). This knockdown has functional consequences, as the expression of *Hoxa9*, a MLL-regulated gene, is repressed. However, *Meis1*, an abundant MLL target transcript, remained unchanged. This may reflect incomplete knockdown, such that the *Meis1* decrease we expect is masked by transcripts from non-transfected cells. According to qRT-PCR, *Meis1* transcripts are more abundant

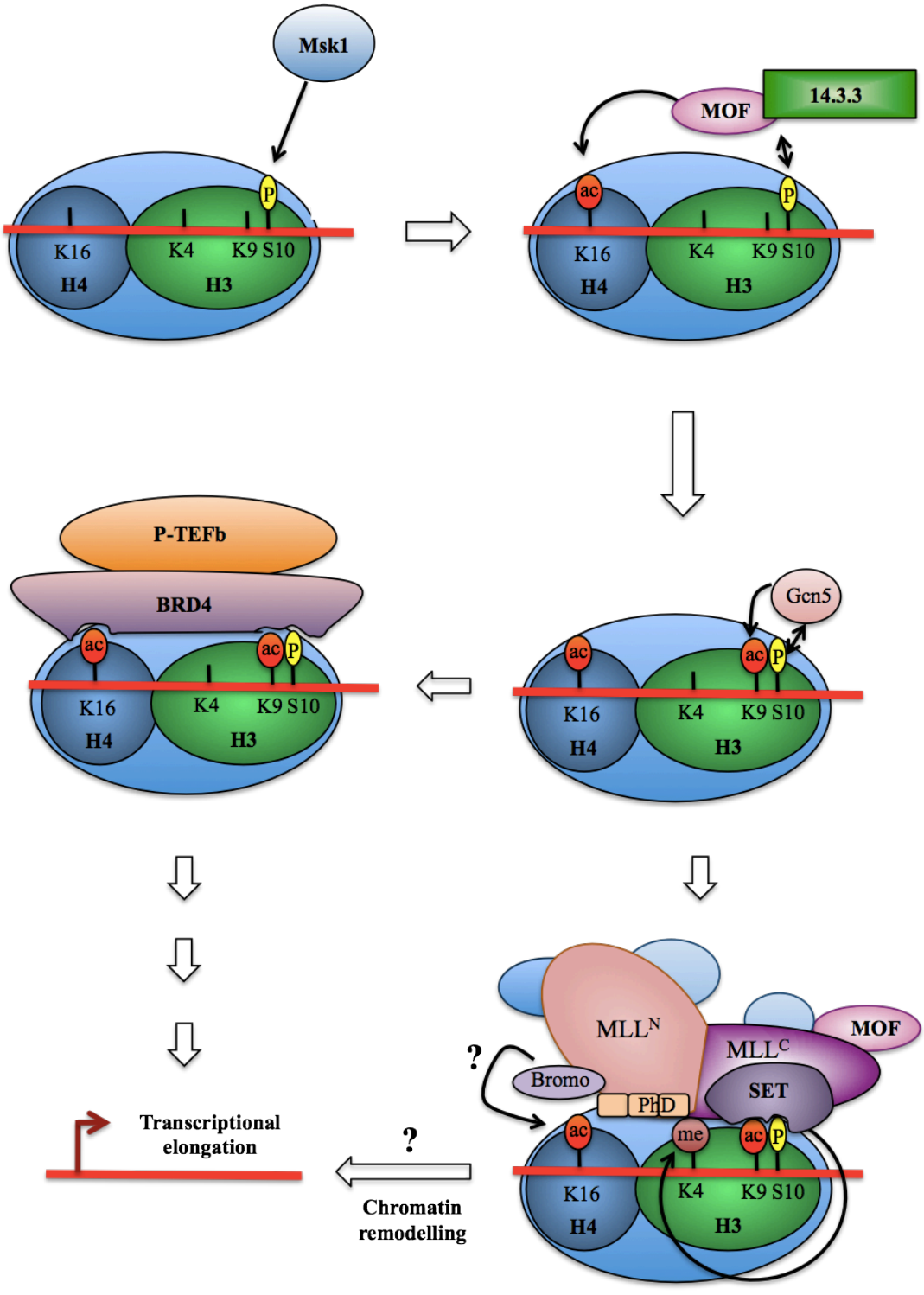


Figure 8.5: Pathways to gene activation at MLL target genes

A speculative cartoon of the factors involved in regulation of MLL target genes.

(>1000-fold) than *Hoxa9* suggesting that *Mll1* knockdown is more likely to influence the transcript levels of lower expressed genes. Alternatively the *Meis1* transcript may be more stable than *Hoxa9*. In addition, other MLL family members were investigated for their expression in *Mll1* knockdown cells; *Mll3* and *Mll5* transcript levels remained unchanged, though a significant increase of *Mll4* expression was observed (158%). This suggests these two genes functionally interact, for example *Mll4* might be down-regulated by MLL1. This result may also explain the maintenance of *Meis1* expression (i.e., *Meis1* could be positively regulated by MLL4 in the absence of MLL1), but *Meis1* has not been identified as a MLL4 target gene to date. These preliminary experiments were based on the examination of a heterogenous population of cells (transfected and non-transfected), This could be improved by using a GFP-tagged knockdown vector to transfect cells, which would allow the isolation of GFP-expressing transfected cells by FACS. However, this approach is likely to need a substantial increase in the number of cells transfected.

The establishment of knockdown approaches for HPC-7 cells would permit the analysis of both MLL and Msk1 contributions to *Hoxa4* and *Hoxa5* regulation. This study showed evidence of the distribution of histone marks, in association with the pattern of binding of MLL complex-associated factors. To investigate MLL target gene regulation, these distributions would be re-examined in MLL or Msk1 knockdown cells. In addition, knockdowns followed by microarray analysis of expression patterns would permit the identification of a broad range of MLL and Msk1 target genes.

9. REFERENCES

Adams-Cioaba, M. A., and Min, J. Structure and function of histone methylation binding proteins. Biochem Cell Biol, v.87, n.1, Feb, p.93-105. 2009.

Agalioti, T., Chen, G., and Thanos, D. Deciphering the transcriptional histone acetylation code for a human gene. Cell, v.111, n.3, Nov 1, p.381-92. 2002.

Agger, K., Cloos, P. A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E., and Helin, K. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature, v.449, n.7163, Oct 11, p.731-4. 2007.

Akbas, G. E., Song, J., and Taylor, H. S. HOXA10 estrogen response element (ERE) is differentially regulated by 17 beta-estradiol and diethylstilbestrol (DES). J Mol Biol, v.340, n.5, Jul 23, p.1013-23. 2004.

Akhtar, A., Zink, D., and Becker, P. B. Chromodomains are protein-RNA interaction modules. Nature, v.407, n.6802, Sep 21, p.405-9. 2000.

Akkers, R. C., van Heeringen, S. J., Jacobi, U. G., Janssen-Megens, E. M., Francoijs, K. J., Stunnenberg, H. G., and Veenstra, G. J. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos. Dev Cell, v.17, n.3, Sep, p.425-34. 2009.

Allan, J., Harborne, N., Rau, D. C., and Gould, H. Participation of core histone "tails" in the stabilization of the chromatin solenoid. J Cell Biol, v.93, n.2, May, p.285-97.1982.

Allfrey, V. G., Faulkner, R., and Mirsky, A. E. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A, v.51, May, p.786-94. 1964.

Atkinson, S. P., Koch, C. M., Clelland, G. K., Willcox, S., Fowler, J. C., Stewart, R., Lako, M., Dunham, I., and Armstrong, L. Epigenetic Marking Prepares the Human HOXA Cluster for Activation During Differentiation of Pluripotent Cells. Stem Cells, Feb 21. 2008.

Ayton, P. M., and Cleary, M. L. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. Genes Dev, v.17, n.18, Sep 15, p2298-307. 2003.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., and Fisher, A. G. Chromatin signatures of pluripotent cell lines. Nat Cell Biol, v.8, n.5, May, p.532-8. 2006.

Bach, C., Mueller, D., Buhl, S., Garcia-Cuellar, M. P., and Slany, R. K. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. Oncogene, v.28, n.6, Feb 12, p.815-23. 2009.

Baneres, J. L., Martin, A., and Parello, J. The N Tails of histones H3 and H4 adopt a highly structured conformation in the nucleosomes. J Mol Biol, v.273, n.3, Oct 31, p.503-8. 1997.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature, v.410, n.6824, Mar 1, p.120-4. 2001.

Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S. K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., and Niehrs, C. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature, v.445, N.7128, Feb 8, p.671-5. 2007.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. High-resolution profiling of histone methylations in the human genome. Cell, v.129, n.4, May 18, p.823-37. 2007.

Bassett, A., Cooper, S., Wu, C., and Travers, A. The folding and unfolding of eukaryotic chromatin. Curr Opin Genet Dev, v.19, n.2, Apr, p.159-65. 2009.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature, v.419, n.6909, Oct 24, p.857-62. 2002.

Berger, S. L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. An operational definition of epigenetics. Genes Dev, v.23, n.7, Apr 1, p.781-3. 2009.

Berndsen, C. E., and Denu, J. M. Catalysis and substrate selection by histone/protein lysine acetyltransferases. Curr Opin Struct Biol, v.18, n.6, Dec, p.682-9. 2008.

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., Schreiber, S. L., and Lander, E. S. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell, v.120, n.2, Jan 28, p.169-81. 2005.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., and Lander, E. S. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell, v.125, n.2, Apr 21, p.315-26. 2006.

Bienz, M. The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci, v.31, n.1, Jan, p.35-40. 2006

Bird, A. DNA methylation patterns and epigenetic memory. Genes Dev, v.16, n.1, Jan 1, p.6-21. 2002.

Bird, A. Perceptions of epigenetics. Nature, v.447, n.7143, May 24, p.396-8. 2007.

Bonner, W. M., West, M. H., and Stedman, J. D. Two-dimensional gel analysis of histones in acid extracst of nuclei, cells, and tiddues. Eur J Biochem, v.109, N.1, Aug, p.17-23. 1980.

Boswell, R. E., and Mahowald, A. P. Tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. Cell, v.43, n.1, Nov, p.97-104. 1985.

Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., Winston, F., and Allis, C. D. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev, v.15, n.24, Dec 15, p.3286-95. 2001.

Briggs, S. D., Xiao, T., Sun, Z. W., Caldwell, J. A., Shabanowitz, J., Hunt, D. F., Allis, C. D., and Strahl, B. D. Gene silencing: trans-histone regulatory pathway in chromatin. Nature, v.418, n.6897, Aug 1, p.498. 2002.

Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de The, H., and Rochette-Egly, C. A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. Embo J, v.28, n.1, Jan 7, p.34-47. 2009.

Brun, M. E., Gasca, S., Girard, C., Bouton, K., De Massy, B., and De Sario, A. Characterization and expression analysis during embryo development of the mouse ortholog of MLL3. Gene, v.371, n.1, Apr 12, p.25-33. 2006.

Burgold, T., Spreafico, F., De Santa, F., Totaro, M. G., Prosperini, E., Natoli, G., and Testa, G. The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment. PLoS One, v.3, n.8, p.e3034. 2008.

Buschbeck, M., Uribesalgo, I., Wibowo, I., Rue, P., Martin, D., Gutierrez, A., Morey, L., Guigo, R., Lopez-Schier, H., and Di Croce, L. The histone variant macroH2A is an epigenetic regulator of key developmental genes. Nat Struct Mol Biol, v.16, n.10, Oct, p.1074-9. 2009.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. Role of histone H3 lysine 27 methylation in Polycomb-goup silencing. Science, v.298, n.5595, Nov 1, p.1039-43. 2002.

Campos, E. I., and Reinberg, D. Histones: annotating chromatin. Annu Rev Genet, v.43, p.559-99. 2009.

Carrozza, M. J., Utley, R. T., Workman, J. L., and Cote, J. The diverse functions of histone acetyltransferase complexes. Trends Genet, v.19, n.6, Jun, p.321-9. 2003

Chen, J., Santillan, D. A., Koonce, M., Wei, W., Luo, R., Thirman, M. J., Zeleznik-Le, N. J., and Diaz, M. O. Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. Cancer Res, v.68, n.15, Aug 1, p.6199-207. 2008.

Chen, Z., Zang, J., Whetstine, J., Hong, X., Davrazou, F., Kutateladze, T. G., Simpson, M., Mao, Q., Pan, C. H., Dai, S., Hagman, J., Hansen, K., Shi, Y., and Zhang, G. Structural insights into histone demethylation by JMJD2 family members. Cell, v.125, n.4, May 19, p.691-702. 2006.

Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M., and Allis, C. D. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factors stimulation. Mol Cell, v.5, n.6, Jun, p.905-15. 2000.

Chow, J., and Heard, E. X inactivation and the complexities of silencing a sex chromosome. Curr Opin Cell Biol, v.21, n.3, Jun, p.359-66. 2009.

Churikov, D., Siino, J., Svetlova, M., Zhang, K., Gineitis, A., Morton Bradbury, E., and Zalensky, A. Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. Genomics, v.84, n.4, Oct, p.745-56. 2004.

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

Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C. Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. Embo J, v.19, n.14, Jul 17, p.3714-26. 2000.

Cloos, P. A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K. H., and Helin, K. The putative oncogene GASC1 demethylases tri- and dimethylated lysine 9 on histone H3. Nature, v.442, n.7100, Jul 20, p.307-11. 2006.

Courtier, B., Heard, E., and Avner, P. Xce haplotypes shows modified methylation in a region of the active X chromosome lying 3' to Xist. Proc Natl Acad Sci U S A, v.92, n.8, Apr 11, p.3531-5. 1995.

Couture, J. F., Collazo, E., and Trievel, R. C. Molecular recognition of histone H3 by the WD40 protein WDR5. Nat Struct Mol Biol, v.13, n.8, Aug, p.698-703. 2006.

Crooks, G. M., Fuller, J., Petersen, D., Izadi, P., Malik, P., Pattengale, P. K., Kohn, D. B., and Gasson, J. C. Constitutive HOXA5 expression inhibits erythropoiesis and increases myelopoiesis from human hematopoietic progenitors. Blood, v.94, n.2, Jul 15, p.519-28. 1999.

Cunliffe, V. T. Eloquent silence: developmental functions of Class I histone deacetylases. Curr Opin Genet Dev, v.18, n.5, Oct, p.404-10. 2008.

Das, C., Lucia, M. S., Hansen, K. C., and Tyler, J. K. CBP/p300-mediated acetylation of histone H3 on lysine 56. Nature, v.459, n.7243, May 7, p.113-7. 2009.

Davie, J. R. Nuclear matrix, dynamic histone acetylation and transcriptionally active chromatin. Mol Biol Rep, v.24, n.3, Aug, p.197-207. 1997.

De Ruijter, A. J., van Gennip, A. H., Caron, H. N., Kemp, S., and van Kuilenburg, A. B. Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J, v.370, n.Pt 3, Mar 15, p.737-49. 2003.

Dehe, P. M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodriguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A., Rosaleny, L. E., Tordera, V., Chavez, S., Stewart, A. F., and Geli, V. Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. J Biol Chem, v.281, n.46, Nov 17, p.35404-12. 2006.

Deng, L. W., Chiu, I., and Strominger, J. L. MLL 5 protein forms intranuclear foci, and overexpresion inhibits cell cycle progression. Proc Natl Acad Sci U S A, v.101, n.3, Jan 20, p.757-62. 2004.

Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. Structure and ligand of a histone acetyltransferase bromodomain. Nature, v.399, n.6735, Jun 3, p.491-6. 1999.

Di Groce, L. Chromatin modifying activity of leukaemia associated fusion proteins. Hum Mol Genet, v.14 Spec No 1, Apr 15, p.R77-84. 2005.

Dichtl, B., Aasland, R., and Keller, W. Functions for S. cerevisiae Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. Rna, v.10, n.6, Jun, p.965-77. 2004.

Dobson, C. L., Warren, A. J., Pannell, R., Forster, A., Lavenir, I., Corral, J., Smith, A. J., and Rabbitts, T. H. The mll-AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis. Embo J, v.18, n.13, Jul 1, p.3564-74. 1999.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J.Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol, v.327, n.1, Mar 14, p.85-96. 2003.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science, v.306, n.5701, Nov 26, p.1571-3. 2004.

Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., and Roeder, R. G. Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol, v.13, n.8, Aug, p.713-9. 2006.

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. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell, v.121, n.6, Jun 17, p.873-85. 2005.

Dover, J., Schneider, J., Tawiah-Boateng, M. A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem, v.277, n.32, Aug 9, p.28368-71. 2002.

Dupe, V., Davenne, M., Brocard, J., Dolle, P., Mark, M., Dierich, A., Chambon, P., and Rijli, F. M. In vivo functional analysis of the Hoxa-1 3' retinoic acid response element (3'RARE). Development, v.124, n.2, Jan, p.399-410. 1997.

Dyson, M. H., Thomson, S., Inagaki, M., Goto, H., Arthur, S. J., Nightingale, K., Iborra, F. J., and Mahadevan, L. C. Map kinase-mediated phosphorylation of distinct pools of histone H3 at S10 or S28 via mitogen- and stress-activated kinase 1/2. J Cell Sci, v.118, n.10, May 15, p.2247-59. 2005

Eissenberg, J. C. and A. Shilatifard. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Dev Biol, Aug 21. 2009.

Emerling, B. M., Bonifas, J., Kratz, C. P., Donovan, S., Taylor, B. R., Green, E. D., Le Beau, M. M., and Shannon, K. M. MLL5, a homolog of Drosophila trithorax located within a segment of chromosome band 7q22 implicated in myeloid leukemia. Oncogene, v.21, n.31, Jul 18, p.4849-54. 2002.

Ernst, P., Fisher, J. K., Avery, W., Wade, S., Foy, D., and Korsmeyer, S. J. Definitive hematopoiesis requires the mixed-lineage leukemia gene. Dev Cell, v.6, n.3, Mar, p.437-43. 2004a.

Ernst, P., Mabon, M., Davidson, A. J., Zon, L. I., and Korsmeyer, S. J. An MLL-dependent Hox program drives hematopoietic progenitor expansion. Curr Biol, v.14, n.22, Nov 23, p.2063-9. 2004b.

Ernst, P., Wang, J., Huang, M., Goodman, R. H., and Korsmeyer, S. J. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol Cell Biol, v.21, n.7, Apr, p.2249-58. 2001.

Ernst, P., Wang, J., and Korsmeyer, S. J. The role of MLL in hematopoiesis and leukemia. Curr Opin Hematol, v.9, n.4, Jul, p.282-7. 2002.

Fair, K., Anderson, M., Bulanova, E., Mi, H., Tropschug, M., and Diaz, M. O. Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. Mol Cell Biol, v.21, n.10, May, p.3589-97. 2001.

Feinberg, A. P., Ohlsson, R., and Henikoff, S. The epigenetic progenitor origin of human cancer. Nat Rev Genet, v.7, n.1, Jan, p.21-33. 2006.

Finch, J. T., and Klug, A. Solenoidal model for superstructure in chromatin. Proc Natl Acad Sci U S A, v.73, n.6, Jun, p.1897-901. 1976.

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. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature, v.438, n.7071, Dec 22, p.1116-22. 2005.

FitzGerald, K. T., and Diaz, M. O. MLL2: a new mammalian member of the trx/MLL family of genes. Genomics, v.59, n.2, Jul 15, p.187-92. 1999.

Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature, v.438, n.7071, Dec 22, p.1181-5. 2005.

Fletcher, T. M., and Hansen, J. C. Core histone tail domains mediate oligonucleosome folding and nucleosomal DNA organization through distinct molecular mechanisms. J Biol Chem, v.270, n.43, Oct 27, p.25359-62.1995.

Francis, N. J., Saurin, A. J., Shao, Z., and Kingston, R. E. Reconstitution of a functional core polycomb repressive complex. Mol Cell, v.8, n.3, Sep, p.545-56. 2001.

Fraser, P. Transcriptional control thrown for a loop. Curr Opin Genet Dev, v.16, n.5, Oct, p.490-5. 2006.

Fraser, P., and Bickmore, W. Nuclear organization of the genome and the potential for gene regulation. Nature, v.447, n.7143, May 24, p.413-7. 2007.

Fuks, F. DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev, v.15, n.5, Oct, p.490-5. 2005.

Fuller, J. F., McAdara, J., Yaron, Y., Sakaguchi, M., Fraser, J. K., and Gasson, J. C. Characterization of HOX gene expression during myelopoiesis: role of HOX A5 in lineage commitment and maturation. Blood, v.93, n.10, May 15, p.3391-400. 1999.

Garcia-Ramirez, M., Dong, F., and Ausio, J. Role of the histone "tails" in the folding of oligonucleosomes depleted of histone H1. J Biol Chem, v.267, n.27, Sep 25, p.19587-95. 1992.

Garcia-Ramirez, M., Rocchini, C., and Ausio, J. Modulation of chromatin folding by histone acetylation. J Biol Chem, v.270, n.30, Jul 28, p.17923-8. 1995.

Ge, H., Si, Y., and Roeder, R. G. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. Embo J, v.17, n.22, Nov 16, p.6723-9. 1998.

Gehring, W. J. Homeo boxes in the study of development. Science, v.236, n.4806, Jun 5, p.1245-52. 1987.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K. R., Aasland, R., Anastassiadis, K., Ang, S. L., and Stewart, A. F. Multiple epigenetic maintainance factors implicated by the loss of Mll2 mouse development. Development, v.133, n.8, Apr, p.1423-32. 2006.

Glozak, M. A., Sengupta, N., Zhang, X., and Seto, E. Acetylation and deacetylation of non-histone proteins. Gene, v.363, Dec 19, p.15-23. 2005

Gondor, A., and Ohlsson, R. Chromatin insulators and cohesins. EMBO Rep, v.9, n.4, Apr, p.327-9. 2008.

Goren, A., Tabib, A., Hecht, M., and Cedar, H. DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin. Genes Dev, v.22, n.10, May 15, p.1319-24. 2008.

Grant, P. A., Eberharter, A., John, S., Cook, R. G., Turner, B. M., and Workman, J. L. Expanded lysine acetylation specificity of Gcn5 in native complexes. J Biol Chem, v.274, n.9, Feb 26, p.5895-900. 1999.

Gu, W., and Roeder, R. G. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell, v.90, n.4, Aug, p.595-606. 1997.

Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. 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, v.71, n.4, Nov 13, p.701-8. 1992.

Guenther, M. G., Jenner, R. G., Chevalier, B., Nakamura, T., Croce, C. M., Canaani, E., and Young, R. A. Global and Hox-specific roles for the MLL1 methyltransferase. Proc Natl Acad Sci U S A, v. 102, n.24, Jun 14, p.8603-8. 2005.

Guenther, M. G., Lawton, L. N., Rozovskaia, T., Frampton, G. M., Levine, S. S., Volkert, T. L., Croce, C. M., Nakamura, T., Canaani, E., and Young, R. A. Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. Genes Dev, v.22, n.24, Dec 15, p.3403-8. 2008.

Gurley LR, D'Anna JA, Barham SS, Deaven LL, Tobey RA. Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. Eur J Biochem, v.84, n.1, Mar, p.1-15. 1978.

Hansen, J. C. Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu Rev Biophys Biomol Struct, v.31, p.361-92. 2002.

Hansen, J. C., Ausio, J., Stanik, V. H., and van Holde, K. E. Homogeneous reconstituted oligonucleosomes, evidence for salt-dependent folding in the absence of histone H1. Biochemistry, v.28, n.23, Nov 14, p.9129-36. 1989.

Hansen, J. C., Tse, C., and Wolffe, A. P. Structure and function of the core histone N-termini: more than meets the eye. Biochemistry, v.37, n.51, Dec 22, p.17637-41. 1998.

Hazzalin, C. A., and Mahadevan, L. C. Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun. PLos Biol, v.3, n.12, Dec, p.e393. 2005.

Heard, E. Recent advances in X-chromosome inactivation. Curr Opin Cell Biol, v.16, n.3, Jun, p.247-55. 2004.

Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. A direct link between core histone acetylation and transcriptionally active chromatin. Embo J, v.7, n.5, May, p.1395-402. 1988.

Hendzel MJ, Wei Y, Mancini MA, Van Hooser A, Ranalli T, Brinkley BR, Bazett-Jones DP, Allis CD. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. Chromosoma, v.106, n.6, Nov, p.348-60. 1997

Henry, K. W., Wyce, A., Lo, W. S., Duggan, L. J., Emre, N. C., Kao, C. F., Pillus, L., Shilatifard, A., Osley, M. A., and Berger, S. L. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes Dev, v.17, n.21, Nov 1, p.2648-63. 2003.

Heppner, C., Bilimoria, K. Y., Agarwal, S. K., Kester, M., Whitty, L. J., Guru, S. C., Chandrasekharappa, S. C., Collins, F. S., Spiegel, A. M., Marx, S. J., and Burns, A. L. The tumor suppressor protein menin interacts with NF-kappaB proteins and inhibits NF-kappaB-mediated transactivation. Oncogene, v.20, n.36, Aug 16, p.4917-25. 2001.

Hershko, A. Y., Kafri, T., Fainsod, A., and Razin, A. Methylation of HoxA5 and HoxaB5 and its relevance to expression during mouse development. Gene, v.302, n.1-2, Jan 2, p.65-72. 2003.

Hess, J. L., Yu, B. D., Li, B., Hanson, R., and Korsmeyer, S. J. Defects in yolk sac hematopoiesis in Mll-null mbryos. Blood, v.90, n.5, Sep 1, p.1799-806. 1977.

Hiragami, K., and Festenstein, R. Heterochromatin protein 1: a pervasive controlling influence. Cell Mol Life Sci, v.62, n.23, Dec, p.2711-26. 2005.

Hirose, Y., and Ohkuma, Y. Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. J Biochem, v.141, n.5, May, p.601-8. 2007.

Holmes, M. L., Bartle, N., Eisbacher, M., and Chong, B. H. Cloning and analysis of the thrombopoietin-induced megakaryocyte-specific glycoprotein VI promoter and its regulation by GATA-1, Fli-1, and Sp1. J Biol Chem, v.277, n.50, Dec 13, p.48333-41. 2002.

Hsieh, J. J., Cheng, E. H., and Korsmeyer, S. J. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. Cell, v.115, n.3, Oct 31, p.293-303. 2003.

Hsieh, J. J., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S. J. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol, v.23, n.1, Jan, p.186-94. 2003.

Hsu, D. R., Chuang, P. T., and Meyer, B. J. DPY-30, a nuclear protein essential early in embryogenesis for Caenorhabditis elegans dosage compensation. Development, v.121, n.10, Oct, p.3323-34. 1995.

Huang, J., Dorsey, J., Chuikov, S., Zhang, X., Jenuwein, T., Reinberg, D., and Berger, S. L. G9A and GLP methylate lysine 373 in the tumor suppressor p53. J Biol Chem, Jan 29.

Huang, W., Batra, S., Korrapati, S., Mishra, V., and Mehta, K. D. Selective repression of low-density lipoprotein receptor expression by SP600125: coupling of histone H3-Ser10 phosphorylation and Sp1 occupancy. Mol Cell Biol, v.26, n.4, Feb, p.1307-17. 2006.

Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R. M. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science, v.312, n.5774, May 5, p.748-51. 2006.

Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacharjee, A., Biondi, C. A., Kay, G. F., Hayward, N. K., Hess, J. L., and Meyerson, M. Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell, v.13, n.4, Feb 27, p.587-97. 2004.

Huntsman, D. G., Chin, S. F., Muleris, M., Batley, S. J., Collins, V. P., Wiedemann, L. M., Aparicio, S., and Caldas, C. MLL2, the second human homolog of the Drosophila trithorax gene, maps to 19q13.1 and is amplified in solid tumor cell lines. Oncogene, v.18, n.56, Dec 23, p.7975-84. 1999.

Imhof, A. Epigenetic regulators and histone modification. Brief Funct Genomic Proteomic, v.5, n.3, Sep, p.222-7. 2006.

Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C. M., Nakamura, T., Mazo, A., Eisenbach, L., and Canaani, E. Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Mol Cell Biol, v.27, n.5, Mar, p.1889-903. 2007.

Ivaldi, M. S., Karam, C. S., and Corces, V. G. Phosphorylation of histone H3 at Ser10 facilitates RNA polymerase II release from promoter-proximal pausing in Drosophila. Genes Dev, v.21, n.21, Nov 1, p.2818-31. 2007.

Iwasaki, H., and Akashi, K. Myeloid lineage commitment from the hematopoietic stem cell. Immunity, v.26, n.6, Jun, p.726-40. 2007.

Izon, D. J., Rozenfeld, S., Fong, S. T., Komuves, L., Largman, C., and Lawrence, H. J. Loss of function of the homeobox gene Hoxa-9 perturbs early T-cell development and induces apoptosis in primitive thymocytes. Blood, v.92, n.2, Jul 15, p.383-93. 1998.

Jacobson, R. H., Ladurner, A. G., King, D. S., and Tjian, R. Structure and function of a human TAFII250 double bromodomain module. Science, v.288, n.5470, May 26, p.1422-5. 2000.

Jaenisch, R., and Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet, v.33 Suppl, Mar, p.245-54. 2003.

Jandrot-Perrus, M., Busfield, S., Lagrue, A. H., Xiong, X., Debili, N., Chickering, T., Le Couedic, J. P., Goodearl, A., Dussault, B., Fraser, C., Vainchenker, W., and Villeval, J. L. Cloning, characterization, and functional studies of human and mouse glycoprotein VI: a platelet-specific collagen receptor from the immunoglobulin superfamily. Blood, v.96, n.5, Sep 1, p.1798-807. 2000.

Jenuwein, T., and Allis, C. D. Translating the histone code. Science, v.293, n.5532, Aug 10, p.1074-80. 2001.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet, v.41, n.8, Aug, p.941-5. 2009.

Jones, D. O., Cowell, I. G., and Singh, P. B. Mammalian chromodomain proteins: their role in genome organisation and expression. Bioessays, v.22, n.2, Feb, p.124-37. 2000.

Joshi, A. A., and Struhl, K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell, v.20, n.6, Dec 22, p.971-8. 2005.

Juan, L. J., Utley, R. T., Vignali, M., Bohm, L., and Workman, J. L. H1-mediated repression of transcription factor binding to a stably positioned nucleosome. J Biol Chem, v.272, n.6, Feb 7, p.3635-40. 1997.

Kamakaka, R. T., and Thomas, J. O. Chromatin structure of transcriptionally competent and repressed genes. Embo J, v.9, n.12, Dec, p.3997-4006. 1990.

Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y., and Bedford, M. T. Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep, v.7, n.4, Apr, p.397-403. 2006.

Kim, J., Guermah, M., McGinty, R. K., Lee, J. S., Tang, Z., Milne, T. A., Shilatifard, A., Muir, T. W., and Roeder, R. G. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell, v.137, n.3, May 1, p.459-71. 2009.

Kim, T., and Buratowski, S. Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. Cell, v.137, n.2, Apr 17, p.259-72. 2009.

Kirmizis, A., Santos-Rosa, H., Penkett, C. J., Singer, M. A., Green, R. D., and Kouzarides, T. Distinct transcriptional outputs associated with mono- and dimethylated histone H3 arginine 2. Nat Struct Mol Biol, v.16, n.4, Apr, p.449-51. 2009.

Kirmizis, A., Santos-Rosa, H., Penkett, C. J., Singer, M. A., Vermeulen, M., Mann, M., Bahler, J., Green, R. D., and Kouzarides, T. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. Nature, v.449, n.7164, Oct 18, p.928-32. 2007.

Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature, v.442, n.7100, Jul 20, p.312-6. 2006.

Kouzarides, T. Chromatin modifications and their function. Cell, v.128, n.4, Feb 23, p.693-705. 2007.

Krause, C. D., Yang, Z. H., Kim, Y. S., Lee, J. H., Cook, J. R., and Pestka, S. Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential. Pharmacol Ther, v.113, n.1, Jan, p.50-87. 2007.

Krebs, J. E., Fry, C. J., Samuels, M. L., and Peterson, C. L. Global role for chromatin remodeling enzymes in mitotic gene expression. Cell, v.102, n.5, Sep 1, p.587-98. 2000.

Kristensen, L. S., Nielsen, H. M., and Hansen, L. L. Epigenetics and cancer treatment. Eur J Pharmacol, v.625, n.1-3, Dec 25, p.131-42. 2009.

Krivtsov, A. V., Feng, Z., Lemieux, M. E., Faber, J., Vempati, S., Sinha, A. U., Xia, X., Jesneck, J., Bracken, A. P., Silverman, L. B., Kutok, J. L., Kung, A. L., and Armstrong, S. A. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell, v.14, n.5, Nov 4, p.355-68. 2008.

Krumlauf, R. Hox genes in vertebrate development. Cell, v.78, n.2, Jul 29, p.191-201. 1994.

Kundu, S., and Peterson, C. L. Role of chromatin states in transcriptional memory. Biochim Biophys Acta, v.1790, n.6, Jun, p.445-55. 2009.

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. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature, v.383, n.6597, Sep 19, p.269-72. 1996.

Kurdistani, S. K., and Grunstein, M. Histone acetylation and deacetylation in yeast. Nat Rev Mol Cell Biol, v.4, n.4, Apr, p.276-84. 2003.

Kwon, S. Y., Xiao, H., Wu, C., and Badenhorst, P. Alternative splicing of NURF301 generates distinct NURF chromatin remodelling complexes with altered modified histone binding specifities. PLoS Genet, v.5, n.7, Jul, p.e.1000574. 2009.

La, P., Silva, A. C., Hou, Z., Wang, H., Schnepp, R. W., Yan, N., Shi, Y., and Hua, X. Direct binding of DNA by tumor suppressor menin. J Biol Chem, v.279, n.47, Nov 19, p.49045-54. 2004.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature, v.410, n.6824, Mar 1, p.116-20. 2001.

Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, v.227, n.5259, Aug 15, p.680-5. 1970.

Lappin, T. R., Grier, D. G., Thompson, A., and Halliday, H. L. HOX genes: seductives science, mysterious mechanisms. Ulster Med J, v.75, n.1, Jan, p.23-31. 2006.

Latham, J. A., and Dent, S. Y. Cross-regulation of histone modifications. Nat Struct Mol Biol, v.14, n.11, Nov 5, p.1017-1024. 2007.

Lavau, C., Szilvassy, S. J., Slany, R., and Cleary, M. L. Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. Embo J, v.16, n.14, Jul 16, p.4226-37. 1997.

Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell, v.72, n.1, Jan 15, p.73-84. 1993.

Lee, J. H., Hart, S. R., and Skalnik, D. G. Histone deacetylase activity is required for embryonic stem cell differentiation. Genesis, v.38, n.1, Jan, p.32-8. 2004.

Lee, J. S., Shukla, A., Schneider, J., Swanson, S. K., Washburn, M. P., Florens, L., Bhaumik, S. R., and Shilatifard, A. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell, v.131, n.6, Dec 14, p.1084-96. 2007.

Lee, K. K., and Workman, J. L. Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol, v.8, n.4, Apr, p.284-95. 2007.

Lemons, D., and McGinnis, W. Genomic evolution of Hox gene clusters. Science, v.313, n.5795, Sep 29, p.1918-22. 2006.

Lennartsson, A., and Ekwall, K. Histone modification patterns and epigenetic codes. Biochim Biophys Acta, v.1790, n.9, Sep, p.863-8. 2009.

Lewis, E. B. A gene complex controlling segmentation in Drosophila. Nature, v.276, n.5688, Dec 7, p.565-70. 1978.

Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet, v.3, n.9, Sep, p.662-73. 2002

Li, H., Ilin, S., Wang, W., Duncan, E. M., Wysocka, J., Allis, C. D., and Patel, D. J. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature, v.442, n.7098, Jul 6, p91-5. 2006.

Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N., and Rando, O. J. Single-nucleosome mapping of histone modifications in S. cerevisiae. PLoS Biol, v.3, n.10, Oct, p.e328. 2005.

Lo, W. S., Trievel, R. C., Rojas, J. R., Duggan, L., Hsu, J. Y., Allis, C. D., Marmorstein, R., and Berger, S. L. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. Mol Cell, v.5, n.6, Jun, p.917-26. 2000.

Lu, Z., Lai, J., and Zhang, Y. Importance of charge independent effects in readout of the trimethyllysine mark by HP1 chromodomain. J Am Chem Soc, v.131, n.41, Oct 21, p.14928-31. 2009.

Lubitz, S., Glaser, S., Schaft, J., Stewart, A. F., and Anastassiadis, K. Increased apoptosis and skewed differentiation in mouse embryonic stem cells lacking the histone methyltransferase mll2. Mol Biol Cell, v.18, n.6, Jun, p.2356-66. 2007.

Luger, K. Structure and dynamic behavior of nucleosomes. Curr Opin Genet Dev, v.13, n.2, Apr, P.127-35. 2003.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature, v.389, n.6648, Sep 18, p.251-60. 1997.

Luger, K., and Richmond, T. J. The histone tails of the nucleosome. Curr Opin Genet Dev, v.8, n.2, Apr, p.140-6. 1998.

Luke, M. P., Sui, G., Liu, H., and Shi, Y. Yin Yang 1 physically interacts with Hoxa11 and represses Hoxa11-dependent transcription. J Biol Chem, v.281, n.44, Nov 3, p.33226-32. 2006.

Macdonald, N., Welburn, J. P., Noble, M. E., Nguyen, A., Yaffe, M. B., Clynes, D., Moggs, J. G., Orphanides, G., Thomson, S., Edmunds, J. W., Clayton, A. L., Endicott, J. A., and Mahadevan, L. C. Molecular basis for the recognition of phosphorylated and phosphoacetylated histone H3 by 14-3-3. Mol Cell, v.20, n.2, Oct 28, p.199-211. 2005.

Margueron, R., Trojer, P., and Reinberg, D. The key to development: interpreting the histone code? Curr Opin Genet Dev, v.15, n.2, Apr, p.163-76. 2005.

Marin-Husstege, M., Muggironi, M., Liu, A., and Casaccia-Bonnefil, P. Histone deacetylase activity is necessary for oligodendrocyte lineage progression. J Neurosci, v.22, n.23, Dec 1, p.10333-45. 2002.

Marmorstein, R., and Trievel, R. C. Trievel. Histone modifying enzymes: structures, mechanisms, and specificities. Biochim Biophys Acta, v.1789, n.1, Jan, p.58-68. 2009.

Martin, D. G., Grimes, D. E., Baetz, K., and Howe, L. Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. Mol Cell Biol, v.26, n.8, Apr, p.3018-28. 2006.

Martinez, E. Multi-protein complexes in eukaryotic gene transcription. Plant Mol Biol, v.50, n.6, Dec, p.925-47. 2002.

Mellor, J. The dynamics of chromatin remodeling at promoters. Mol Cell, v.19, n.2, Jul 22, p.147-57. 2005.

Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R., and Schule, R. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature, v.437, n.7057, Sep 15, p.436-9. 2005.

Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci U S A, v.98, n.23, Nov 6, p.12902-7. 2001.

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., and Hess, J. L. MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell, v.10, n.5, Nov, p.1107-17. 2002.

Milne, T. A., Dou, Y., Martin, M. E., Brock, H. W., Roeder, R. G., and Hess, J. L. MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A, v.102, n.41, Oct 11, p.14765-70. 2005.

Milne, T. A., Hughes, C. M., Lloyd, R., Yang, Z., Rozenblatt-Rosen, O., Dou, Y., Schnepp, R. W., Krankel, C., Livolsi, V. A., Gibbs, D., Hua, X., Roeder, R. G., Meyerson, M., and Hess, J. L. Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. Proc Natl Acad Sci U S A, v.102, n.3, Jan 18, p.749-54. 2005.

Mizzen, C. A., Yang, X. J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L., Kouzarides, T., Nakatani, Y., and Allis, C. D. The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell, v.87, n.7, Dec 27, p.1261-70. 1996.

Mokrani, H., Sharaf el Dein, O., Mansuroglu, Z., and Bonnefoy, E. Binding of YY1 to the proximal region of the murine beta interferon promoter is essential to allow CBP recruitment and K8H4/K14H3 acetylation on the promoter region after virus infection. Mol Cell Biol, v.26, n.22, Nov, p.8551-61. 2006.

Morillon, A., Karabetsou, N., Nair, A., and Mellor, J. Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. Mol Cell, v.18, n.6, Jun 10, p.723-34. 2005.

Moriniere, J., Rousseaux, S., Steuerwald, U., Soler-Lopez, M., Curtet, S., Vitte, A. L., Govin, J., Gaucher, J., Sadoul, K., Hart, D. J., Krijgsveld, J., Khochbin, S., Muller, C. W., and Petosa, C. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. Nature, v.461, n.7264, Oct 1, p.664-8. 2009.

Mujtaba, S., Zeng, L., and Zhou, M. M. Structure and acetyl-lysine recognition of the bromodomain. Oncogene, v.26, n.37, Aug 13, p.5521-7. 2007.

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. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell, v.111, n.2, Oct 18, p.197-208. 2002.

Muntean, A. G., Giannola, D., Udager, A. M., and Hess, J. L. The PHD fingers of MLL block MLL fusion protein-mediated transformation. Blood, v.112, n.12, Dec 1, p.4690-3. 2008.

Murayama, A., Ohmori, K., Fujimura, A., Minami, H., Yasuzawa-Tanaka, K., Kuroda, T., Oie, S., Daitoku, H., Okuwaki, M., Nagata, K., Fukamizu, A., Kimura, K., Shimizu, T., and Yanagisawa, J. Epigenetic control of rDNA loci in response to intracellular energy status. Cell, v.133, n.4, May 16, p.627-39. 2008.

Musco, G., and Peterson, P. PHD finger of autoimmune regulator: an epigenetic link between the histone modifications and tissue-specific antigen expression in thymus. Epigenetics, v.3, n.6, Nov, p.310-4. 2008.

Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C. M., and Canaani, E. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell, v.10, n.5, Nov, p.1119-28. 2002.

Nakanishi, S., Lee, J. S., Gardner, K. E., Gardner, J. M., Takahashi, Y. H., Chandrasekharan, M. B., Sun, Z. W., Osley, M. A., Strahl, B. D., Jaspersen, S. L., and Shilatifard, A. Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. J Cell Biol, v.186, n.3, Aug 10, p.371-7. 2009.

Nan, X., Meehan, R. R., and Bird, A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Res, v.21, n.21, Oct 25, p.4886-92. 1993.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature, v.293, n.6683, May28, p.386-9. 1998.

Ng, H. H., and Bird, A. DNA methylation and chromatin modification. Curr Opin Genet Dev, v.9, n.2, Apr, p.158-63. 1999.

Ng, J. H., Heng, J. C., Loh, Y. H., and Ng, H. H. Transcriptional and epigenetic regulations of embryonic stem cells. Mutat Res, v.647, n.1-2, Dec 1, p.52-8. 2008.

Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V., and Laue, E. D. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature, v.416, n.6876, Mar 7, p.103-7. 2002.

Nightingale, K. P., Gendreizig, S., White, D. A., Bradbury, C., Hollfelder, F., and Turner, B. M. Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation. J Biol Chem, v.282, n.7, Feb 16, p.4408-16. 2007.

Nightingale, K. P., O'Neill, L. P., and Turner, B. M. Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. Curr Opin Genet Dev, v.16, n.2, Apr, p.125-36. 2006.

O'Neill, L. P., Randall, T. E., Lavender, J., Spotswood, H. T., Lee, J. T., and Turner, B. M. X-linked genes in female embryonic stem cells carry an epigenetic mark prior to the onset of X inactivation. Hum Mol Genet, v.12, n.15, Aug 1, p.1783-90. 2003.

O'Neill, L. P., VerMilyea, M. D., and Turner, B. M. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. Nat Genet, v.38, n.7, Jul, p.835-41. 2006.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell, v.99, n.3, Oct 29, p.247-57. 1999.

Org, T., Chignola, F., Hetenyi, C., Gaetani, M., Rebane, A., Liiv, I., Maran, U., Mollica, L., Bottomley, M. J., Musco, G., and Peterson, P. The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. EMBO Rep, v.9, n.4, Apr, p.370-6. 2008.

Orlando, V., and Paro, R. Chromatin multiprotein complexes involved in the maintenance of transcription patterns. Curr Opin Genet Dev, v.5, n.2, Apr, p.174-9. 1995.

Oudet, P., Gross-Bellard, M., and Chambon, P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell, v.4, n.4, Apr, p.281-300. 1975.

Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. Embo J, v.19, n.22, Nov 15, p.6141-9. 2000.

Paik, W. K., and Kim, S. Enzymatic demethylation of calf thymus histones. Biochem Biophys Res Commun, v.51, n.3, Apr 2, p.781-8. 1973.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G. A., Stewart, R., and Thomson, J. A. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell, v.1, n.3, Sep 13, p.299-312. 2007.

Patel, A., Dharmarajan, V., and Cosgrove, M. S. Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. J Biol Chem, v.283, n.47, Nov 21, p.32158-61. 2008.

Patel, A., Dharmarajan, V., Vought, V. E., and Cosgrove, M. S. On the Mechanism of Multiple Lysine Methylation by the Human Mixed Lineage Leukemia Protein-1 (MLL1) Core Complex. J Biol Chem, v.284, n.36, Sep 4, p.24242-56. 2009.

Patel, A., Vought, V. E., Dharmarajan, V., and Cosgrove, M. S. A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. J Biol Chem, v.283, n.47, Nov 21, p.32162-75. 2008.

Paulson, J. R., Ciesielski, W. A., Schram, B. R., and Mesner, P. W. Okadaic acid induces dephosphorylation of histone H1 in metaphase-arrested HeLa cells. J Cell Sci, v.107 (Pt 1), Jan, p.267-73. 1994.

Pearce, J. J., Singh, P. B., and Gaunt, S. J. The mouse has a Polycomb-like chromobox gene. Development, v.114, n.4, Apr, p.921-9. 1992.

Pena, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., Zhao, R., and Kutateladze, T. G. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature, v.442, n.7098, Jul 6, p.100-3. 2006.

Pijnappel, W. W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Seraphin, B., Aasland, R., and Stewart, A. F. The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. Genes Dev, v.15, n.22, Nov 15, p.2991-3004. 2001.

Pineault, N., Helgason, C. D., Lawrence, H. J., and Humphries, R. K. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. Exp Hematol, v.30, n.1, Jan, p.49-57. 2002.

Pinskaya, M., and Morillon, A. Histone H3 lysine 4 di-methylation: a novel mark for transcriptional fidelity? Epigenetics, v.4, n.5, Jul, p.302-6. 2009.

Pinto do, O. P., Kolterud, A., and Carlsson, L. Expression of the LIM-homeobox gene LH2 generates immortalized steel factor-dependent multipotent hematopoietic precursors. Embo J, v.17, n.19, Oct 1, p.5744-56. 1998.

Pinto do, O. P., Richter, K., and Carlsson, L. Hematopoietic progenitor/stem cells immortalized by Lhx2 generate functional hematopoietic cells in vivo. Blood, v.99, n.11, Jun 1, p.3939-46. 2002.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K., and Young, R. A. Genome-wide map of nucleosome acetylation and methylation in yeast. Cell, v.122, n.4, Aug 26, p.517-27. 2005.

Popovic, R., and Zeleznik-Le, N. J. MLL: how complex does it get? J Cell Biochem, v.95, n.2, May 15, p.234-42. 2005.

Prasad, R., Zhadanov, A. B., Sedkov, Y., Bullrich, F., Druck, T., Rallapalli, R., Yano, T., Alder, H., Croce, C. M., Huebner, K., Mazo, A., and Canaani, E. Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to Drosophila trithorax. Oncogene, v.15, n.5, Jul 31, p.549-60. 1997.

Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., 3rd, and Grant, P. A. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature, v.433, n.7024, Jan 27, p.434-8. 2005.

Probst, A. V., Dunleavy, E., and Almouzni, G. Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol, v.10, n.3, Mar, p.192-206. 2009.

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. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature, v.406, n.6796, Aug 10, p.593-9. 2000.

Reber, L., Vermeulen, L., Haegeman, G., and Frossard, N. Ser276 phosphorylation of NF-kB p65 by MSK1 controls SCF expression in inflammation. PLoS One, v.4, n.2, p.e4393. 2009.

Rice, K. L., Hormaeche, I., and Licht, J. D. Epigenetic regulation of normal and malignant hematopoiesis. Oncogene, v.26, n.47, Oct 15, p.6697-714. 2007.

Richter, K., Wirta, V., Dahl, L., Bruce, S., Lundeberg, J., Carlsson, L., and Williams, C. Global gene expression analyses of hematopoietic stem cell-like cell lines with inducible Lhx2 expression. BMC Genomics, v.7, p.75. 2006.

Ringrose, L., and Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet, v.38, p.413-43. 2004.

Robinson, P. J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R. G., and Rhodes, D. 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. J Mol Biol, v.381, n.4, Sep 12, p.816-25. 2008.

Robinson, P. J., and Rhodes, D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Curr Opin Struct Biol, v.16, n.3, Jun, p.336-43. 2006.

Robzyk, K., Recht, J., and Osley, M. A. Rad6-dependent ubiquitination of histone H2B in yeast. Science, v.287, n.5452, Jan 21, p.501-4. 2000.

Roelen, B. A., de Graaff, W., Forlani, S., and Deschamps, J. Hox cluster polarity in early transcriptional availability: a high order regulatory level of clustered Hox genes in the mouse. Mech Dev, v.119, n.1, Nov, p.81-90. 2002.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. Embo J, v.20, n.24, Dec 17, p.7137-48. 2001.

Routh, A., Sandin, S., and Rhodes, D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc Natl Acad Sci U S A, v.105, n.26, Jul 1, p.8872-7. 2008.

Ruault, M., Brun, M. E., Ventura, M., Roizes, G., and De Sario, A. MLL3, a new human member of the TRX/MLL gene family, maps to 7q36, a chromosome region frequently deleted in myeloid leukaemia. Gene, v.284, n.1-2, Feb 6, p.73-81. 2002.

Ruthenburg, A. J., Li, H., Patel, D. J., and Allis, C. D. Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol, v.8, n.12, Dec, p.983-94. 2007.

Ruthenburg, A. J., Wang, W., Graybosch, D. M., Li, H., Allis, C. D., Patel, D. J., and Verdine, G. L. Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. Nat Struct Mol Biol, v.13, n.8, Aug, p.704-12. 2006

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T. Active genes are tri-methylated at K4 of histone H3. Nature, v.419, n.6905, Sep 26, p.407-11. 2002.

Sauer, F., and Tjian, R. Mechanisms of transcriptional activation: differences and similarities between yeast, Drosophila, and man. Curr Opin Genet Dev, v.7, n.2, Apr, p.176-81. 1997.

Sauvageau, G., Lansdorp, P. M., Eaves, C. J., Hogge, D. E., Dragowska, W. H., Reid, D. S., Largman, C., Lawrence, H. J., and Humphries, R. K. Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. Proc Natl Acad Sci U S A, v.91, n.25, Dec 6, p.12223-7. 1994.

Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature, v.436, n.7047, Jul 7, p.138-41. 2005.

Schneider, J., Wood, A., Lee, J. S., Schuster, R., Dueker, J., Maguire, C., Swanson, S. K., Florens, L., Washburn, M. P., and Shilatifard, A. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. Mol Cell, v.19, n.6, Sep 16, p.849-56. 2005.

Schwarz, P. M., and Hansen, J. C. Formation and stability of higher order chromatin structures. Contributions of the histone octamer. J Biol Chem, v.269, n.23, Jun 10, p.16284-9. 1994.

Sen, G. L., Reuter, J. A., Webster, D. E., Zhu, L., and Khavari, P. A. DNMT1 maintains progenitor function in self-renewing somatic tissue. Nature, v.463, n.7280, Jan 28, p.563-7. 2010.

Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M. R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B. R., Ayer, D. E., Kutateladze, T. G., Shi, Y., Cote, J., Chua, K. F., and Gozani, O. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature, v.442, n.7098, Jul 6, p.96-9. 2006.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell, v.119, n.7, Dec 29, p.941-53. 2004.

Shia, W. J., Li, B., and Workman, J. L. SAS-mediated acetylation of histone H4 Lys 16 is required for H2A.Z incorporation at subtelomeric regions in Saccharomyces cerevisiae. Genes Dev, v.20, n.18, Sep 15, p.2507-12. 2006.

Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science, v.311, n.5762, Feb 10, p.844-7. 2006.

Shogren-Knaak, M., and Peterson, C. L. Switching on chromatin: mechanistic role of histone H4-K16 acetylation. Cell Cycle, v.5, n.13, Jul, p.1361-5. 2006.

Simpson, R. T. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry, v.17, n.25, Dec 12, p.5524-31. 1978.

Slany, R. K. Chromatin control of gene expression: mixed-lineage leukemia methyltransferase SETs the stage for transcription. Proc Natl Acad Sci U S A, v.102, n.41, Oct 11, p.14481-2. 2005a.

Slany, R. K. When epigenetics kills: MLL fusion proteins in leukemia. Hematol Oncol, v.23, n.1, Mar, p.1-9. 2005b.

Slany, R. K. The molecular biology of mixed lineage leukemia. Haematologica, v.94, n.7, Jul, p.984-93. 2009.

Slany, R. K., Lavau, C., and Cleary, M. L. The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. Mol Cell Biol, v.18, n.1, Jan, p.122-9. 1998.

Soloaga, A., Thomson, S., Wiggin, G. R., Rampersaud, N., Dyson, M. H., Hazzalin, C. A., Mahadevan, L. C., and Arthur, J. S. MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14. Embo J, v.22, n.11, Jun 2 , p.2788-97. 2003

Song, J. J., and Kingston, R. E. WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. J Biol Chem, v.283, n.50, Dec 12, p.35258-64. 2008.

Soshnikova, N., and Duboule, D. Epigenetic regulation of Hox gene activation: the waltz of methyls. Bioessays, v.30, n.3, Mar, p.199-202. 2008.

Southall, S. M., Wong, P. S., Odho, Z., Roe, S. M., and Wilson, J. R. Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. Mol Cell, v.33, n.2, Jan 30, p.181-91. 2009.

Spencer, T. E., Jenster, G., Burcin, M. M., Allis, C. D., Zhou, J., Mizzen, C. A., McKenna, N. J., Onate, S. A., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. Steroid receptor coactivator-1 is a histone acetyltransferase. Nature, v.389, n.6647, Sep 11, p.194-8. 1997.

Steward, M. M., Lee, J. S., O'Donovan, A., Wyatt, M., Bernstein, B. E., and Shilatifard, A. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nat Struct Mol Biol, v.13, n.9, Sep, p.852-4. 2006.

Strahl, B. D., and Allis, C. D. The language of covalent histone modifications. Nature, v.403, n.6765, Jan 6, p.41-5. 2000.

Strathdee, G., Holyoake, T. L., Sim, A., Parker, A., Oscier, D. G., Melo, J. V., Meyer, S., Eden, T., Dickinson, A. M., Mountford, J. C., Jorgensen, H. G., Soutar, R., and Brown, R. Inactivation of HOXA genes by hypermethylation in myeloid and lymphoid malignancy is frequent and associated with poor prognosis. Clin Cancer Res, v.13, n.17, Sep 1, p.5048-55. 2007a.

Strathdee, G., Sim, A., Parker, A., Oscier, D., and Brown, R. Promoter hypermethylation silences expression of the HoxA4 gene and correlates with IgVh mutational status in CLL. Leukemia, v.20, n.7, Jul, p.1326-9. 2006.

Strathdee, G., Sim, A., Soutar, R., Holyoake, T. L., and Brown, R. HOXA5 is targeted by cell-type-specific CpG island methylation in normal cells and during the development of acute myeloid leukaemia. Carcinogenesis, v.28, n.2, Feb, p.299-309. 2007b.

Strobl, L. J., and Eick, D. Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo. Embo J, v.11, n.9, Sep, p.3307-14. 1992.

Struhl, G. A homoeotic mutation transforming leg to antenna in Drosophila. Nature, v.292, n.5824, Aug 13, p.635-8. 1981.

Sun, Z. W., and Allis, C. D. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature, v.418, n.6893, Jul 4, p.104-8. 2002.

Sung P, Prakash S, Prakash L. The RAD6 protein of Saccharomyces cerevisiae polyubiquitinates histones, and its acidic domain mediates this activity. Genes Dev, v.2, n.11, Nov, p.1476-85. 1988.

Swigut, T., and Wysocka, J. H3K27 demethylases, at long last. Cell, v.131, n.1, Oct 5, p.29-32. 2007.

Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. 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, v.276, n.27, Jul 6, p.25309-17. 2001.

Tang, Y., Zhao, W., Chen, Y., Zhao, Y., and Gu, W. Acetylation is indispensable for p53 activation. Cell, v.133, n.4, May 16, p.612-26. 2008.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol, v.14, n.11, Nov, p.1025-40. 2007.

Terskikh, A. V., Miyamoto, T., Chang, C., Diatchenko, L., and Weissman, I. L. Gene expression analysis of purified hematopoietic stem cells and committed progenitors. Blood, v.102, n.1, Jul 1, p.94-101. 2003.

Thoma, F., Koller, T., and Klug, A. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol, v.83, n.2 Pt 1, Nov, p.403-27. 1979.

Thomas, J. O., and Kornberg, R. D. An octamer of histones in chromatin and free in solution. Proc Natl Acad Sci U S A, v.72, n.7, Jul, p.2626-30. 1975.

Tian, X., and Fang, J. Current perspectives on histone demethylases. Acta Biochim Biophys Sin (Shanghai), v.39, n.2, Feb, p.81-8. 2007.

Tie, F., Banerjee, R., Stratton, C. A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M. O., Scacheri, P. C., and Harte, P. J. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development, v.136, n.18, Sep, p.3131-41. 2009.

Tkachuk, D. C., Kohler, S., and Cleary, M. L. Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell, v.71, n.4, Nov 13, p.691-700. 1992.

Towbin, H., Staehelin, T., and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A, v.76, n.9, Sep, p.4350-4. 1979.

Tran, H. G., Steger, D. J., Iyer, V. R., and Johnson, A. D. The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. Embo J, v.19, n.10, May 15, p.2323-31. 2000.

Trasler, J. M. Epigenetics in spermatogenesis. Mol Cell Endocrinol, v.306, n.1-2, Jul 10, p.33-6. 2009.

Travers, A. The location of the linker histone on the nucleosome. Trends Biochem Sci, v.24, n.1, Jan, p.4-7. 1999.

Trivedi, C. M., Patel, R. C., and Patel, C. V. regulation of HOXA9 expression by nuclear factor kappa B (NF-kappaB) and HOXA9. Gene, v.408, n.1-2, Jan 31, p.187-95. 2008.

Trojer, P., Zhang, J., Yonezawa, M., Schmidt, A., Zheng, H., Jenuwein, T., and Reinberg, D. Dynamic Histone H1 Isotype 4 Methylation and Demethylation by Histone Lysine Methyltransferase G9a/KMT1C and the Jumonji Domain-containing JMJD2/KDM4 Proteins. J Biol Chem, v.284, n.13, Mar 27, p.8395-405. 2009.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. 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, v.13, n.16, Aug 15, p.3822-31. 1994.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. Histone demethylation by a family of JmjC domain-containing proteins. Nature, v.439, n.7078, Feb 16, p.811-6. 2006.

Tsukuda, T., Fleming, A. B., Nickoloff, J. A., and Osley, M. A. Chromatin remodelling at a DNA double-strand break site in Saccharomyces cerevisiae. Nature, v.438, n.7066, Nov 17, p.379-83. 2005.

Turlure, F., Maertens, G., Rahman, S., Cherepanov, P., and Engelman, A. A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. Nucleic Acids Res, v.34, n.5, p.1653-65. 2006.

Turner, B. M. Histone acetylation and control of gene expression. J Cell Sci, v.99 (Pt 1), May, p.13-20. 1991.

Turner, B. M. Histone acetylation as an epigenetic determinant of long-term transcriptional competence. Cell Mol Life Sci, v.54, n.1, Jan, p.21-31. 1998.

Turner, B. M. Histone acetylation and an epigenetic code. Bioessays, v.22, n.9, Sep, p.836-45. 2000.

Turner, B. M. Cellular memory and the histone code. Cell, v.111, n.3, Nov 1, p.285-91. 2002.

Turner, B. M. Reading signals on the nucleosome with a new nomenclature for modified histones. Nat Struct Mol Biol, v.12, n.2, Feb, p.110-2. 2005.

Turner, B. M., Birley, A. J., and Lavender, J. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. Cell. v.69, n.2, Apr 17, p.375-84. 1992.

Ura, K., Kurumizaka, H., Dimitrov, S., Almouzni, G., and Wolffe, A. P. Histone acetylation: influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression. Embo J, v.16, n.8, Apr 15, p.2096-107. 1997.

Utley, R. T., Ikeda, K., Grant, P. A., Cote, J., Steger, D. J., Eberharter, A., John, S., and Workman, J. L. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. Nature, v.394, n.6692, Jul 30, p.498-502. 1998.

Vakoc, C. R., Sachdeva, M. M., Wang, H., and Blobel, G. A. Profile of histone lysine methylation across transcribed mammalian chromatin. Mol Cell Biol, v.26, n.24, Dec, p.9185-95. 2006.

Valledor, A. F., Borras, F. E., Cullell-Young, M., and Celada, A. Transcription factors that regulate monocyte/macrophage differentiation. J Leukoc Biol, v.63, n.4, Apr, p.405-17. 1998.

Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W., van Schaik, F. M., Varier, R. A., Baltissen, M. P., Stunnenberg, H. G., Mann, M., and Timmers, H. T. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell, v.131, n.1, Oct 5, p.58-69. 2007.

VerMilyea, M. D., O'Neill, L. P., and Turner, B. M. Transcription-independent heritability of induced histone modifications in the mouse preimplantation embryo. PLoS One, v.4, n.6, p.e6086. 2009.

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane- Robinson, C., Allis, C. D., and Workman, J. L. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. Embo J, v.15, n.10, May 15, p.2508-18. 1996.

Vicent, G. P., Ballare, C., Nacht, A. S., Clausell, J., Subtil-Rodriguez, A., Quiles, I., Jordan, A., and Beato, M. Induction of progesterone target genes requires activation of Erk and Msk kinases and phosphorylation of histone H3. Mol Cell, v.24, n.3, Nov 3, p.367-81. 2006.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y., and Fuks, F. The Polycomb group protein EZH2 directly controls DNA methylation. Nature, v.439, n.7078, Feb 16, p.871-4. 2006.

Vitaliano-Prunier, A., Menant, A., Hobeika, M., Geli, V., Gwizdek, C., and Dargemont, C. Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. Nat Cell Biol, v.10, n.11, Nov, p.1365-71. 2008.

Vitrat, N., Cohen-Solal, K., Pique, C., Le Couedic, J. P., Norol, F., Larsen, A. K., Katz, A., Vainchenker, W., and Debili, N. Endomitosis of human megakaryocytes are due to abortive mitosis. Blood, v.91, n.10, May 15, p.3711-23. 1998.

Wallace, J. A., and Felsenfeld, G. We gather together: insulators and genome organization. Curr Opin Genet Dev, v.17, n.5, Oct, p.400-7. 2007.

Wang, P., Lin, C., Smith, E. R., Guo, H., Sanderson, B. W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L. M., Ge, K., Krumlauf, R., and Shilatifard, A. Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Mol Cell Biol, Aug 24. 2009.

Wang, Y., Zhang, H., Chen, Y., Sun, Y., Yang, F., Yu, W., Liang, J., Sun, L., Yang, X., Shi, L., Li, R., Li, Y., Zhang, Y., Li, Q., Yi, X., and Shang, Y. LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. Cell, v.138, n.4, Aug 21, p.660-72. 2009.

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet, v.40, n.7, Jul, p.897-903. 2008.

Williams, S. P., Athey, B. D., Muglia, L. J., Schappe, R. S., Gough, A. H., and Langmore, J. P. Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. Biophys J, v.49, n.1, Jan, p.233-48. 1986.

Winter, S., Simboeck, E., Fischle, W., Zupkovitz, G., Dohnal, I., Mechtler, K., Ammerer, G., and Seiser, C. 14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation. Embo J, v.27, n.1, Jan 9, p.88-99. 2008.

Woodage, T., Basrai, M. A., Baxevanis, A. D., Hieter, P., and Collins, F. S. Characterization of the CHD family of proteins. Proc Natl Acad Sci U S A, v.94, n.21, Oct 14, p.11472-7. 1997.

Worcel, A., Strogatz, S., and Riley, D. Structure of chromatin and the linking number of DNA. Proc Natl Acad Sci U S A, v.78, n.3, Mar, p.1461-5. 1981.

Wu, C., Bassett, A., and Travers, A. A variable topology for the 30-nm chromatin fibre. EMBO Rep, v.8, n.12, Dec, p.1129-34. 2007.

Wu, H., Min, J., Lunin, V. V., Antoshenko, T., Dombrovski, L., Zeng, H., Allali-Hassani, A., Campagna-Slater, V., Vedadi, M., Arrowsmith, C. H., Plotnikov, A. N., and Schapira, M. Structural biology of human H3K9 methyltransferases. PLoS One, v.5, n.1, 2010.

Wu, H., and Sun, Y. E. Sun. Epigenetic regulation of stem cell differentiation. Pediatr Res, v.59, n.4 Pt 2, Apr, p.21R-5R. 2006.

Wu, M., Wang, P. F., Lee, J. S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Mol Cell Biol, v.28, n.24, Dec, p.7337-44. 2008.

Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell, v.121, n.6, Jun 17, p.859-72. 2005.

Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., Wu, C., and Allis, C. D. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature, v.442, n.7098, Jul 6, p.86-90. 2006.

Xu, F., Zhang, K., and Grunstein, M. Acetylation in histone H3 globular domain regulates gene expression in yeast. Cell, v.121, n.3, May 6, p.375-85. 2005.

Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J., and Wang, W. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell, v.2, n.6, Dec, p.851-61. 1998.

Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. Cell, v.125, n.3, May 5, p.483-95. 2006.

Yao, Y. L., Yang, W. M., and Seto, E. Regulation of transcription factor YY1 by acetylation and deacetylation. Mol Cell Biol, v.21, n.17, Sep, p.5979-91. 2001.

Yokoyama, A., and Cleary, M. L. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell, v.14, n.1, Jul 8, p.36-46. 2008.

Yokoyama, A., Kitabayashi, I., Ayton, P. M., Cleary, M. L., and Ohki, M. Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. Blood, v.100, n.10, Nov 15, p.3710-8. 2002.

Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D. J., Kitabayashi, I., Herr, W., and Cleary, M. L. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol Cell Biol, v.24, n.13, Jul, p.5639-49. 2004.

You, A., Tong, J. K., Grozinger, C. M., and Schreiber, S. L. CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc Natl Acad Sci U S A, v.98, n.4, Feb 13, p.1454-8. 2001.

Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E., and Korsmeyer, S. J. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. Proc Natl Acad Sci U S A, v.95, n.18, Sep 1, p.10632-6. 1998.

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A., and Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. Nature, v.378, n.6556, Nov 30, p.505-8. 1995.

Zeisig, B. B., Milne, T., Garcia-Cuellar, M. P., Schreiner, S., Martin, M. E., Fuchs, U., Borkhardt, A., Chanda, S. K., Walker, J., Soden, R., Hess, J. L., and Slany, R. K. Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. Mol Cell Biol, v.24, n.2, Jan, p.617-28. 2004.

Zeisig, B. B., Schreiner, S., Garcia-Cuellar, M. P., and Slany, R. K. Transcriptional activation is a key function encoded by MLL fusion partners. Leukemia, v.17, n.2, Feb, p.359-65. 2003.

Zeng, L., Zhang, Q., Gerona-Navarro, G., Moshkina, N., and Zhou, M. M. Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. Structure, v.16, n.4, Apr, p.643-52. 2008.

Zhang, Q., Chakravarty, S., Ghersi, D., Zeng, L., Plotnikov, A. N., Sanchez, R., and Zhou, M. M. Biochemical profiling of histone binding selectivity of the yeast bromodomain family. PLoS One, v.5, n.1, 2010.

Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. Cell, v.89, n.3, May 2, p.357-64. 1997.

Zhang, Y., Wong, J., Klinger, M., Tran, M. T., Shannon, K. M., and Killeen, N. MLL5 contributes to hematopoietic stem cell fitness and homeostasis. Blood, v.113, n.7, Feb 12, p.1455-63. 2009.

Zhao, Q., Rank, G., Tan, Y. T., Li, H., Moritz, R. L., Simpson, R. J., Cerruti, L., Curtis, D. J., Patel, D. J., Allis, C. D., Cunningham, J. M., and Jane, S. M. PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. Nat Struct Mol Biol, v.16, n.3, Mar, p.304-11. 2009.

Zheng, C., and Hayes, J. J. Structures and interactions of the core histone tail domains. Biopolymers, v.68, n.4, Apr, p.539-46. 2003.

Ziemin-van der Poel, S., McCabe, N. R., Gill, H. J., Espinosa, R., III, Patel, Y., Harden, A., Rubinelli, P., Smith, S. D., LeBeau, M. M., Rowley, J. D., and Manuel O. Diaz. Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. Proc Natl Acad Sci U S A, v.88, n.23, Dec 1, p.10735-9. 1991.

Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription élongation. Cell, v.138, n.6, Sep 18, p.1122-36. 2009.

TABLE OF SUPPLEMENTARY DATA

Supplementary data 1

Relative expression of control genes in undifferentiated HPC7 cells and megakaryocytes

	HPC-7	Megakaryocytes
Pu.1	1	0.24
Gata-1	1	3565.78
Gp6	1	18

Ct values of *Hoxa* genes in undifferentiated HPC7 cells and megakaryocytes

In megakaryocytes

Ct values	***Gapdh***	***A1***	***A2***	***A3***	***A4***	***A5***	***A6***	***A7***	***A9***	***A10***	***A11***	***A13***
Megs1	16.03		29.86	27.22	31.48	25.23	28.39			31.04	35.03	33.64
Megs2	16.35	28		28.78	30.51	25.78	28.53	30.34		30	31.98	35.37
Megs3	16.1	27.35	28.65	26.82	29.64	24.54	28.43	28.8	37.16	30.08		32
Megs4	16.23	26.03	28.98	27.32	30.17	25.1	27.71	29.52	36.75		34.82	35.83
Mean	16.18	27.13	29.16	27.54	30.45	25.16	28.27	29.55	36.96	30.37	33.94	34.44

In undifferentiated HPC-7 cells

Ct values	***Gapdh***	***A1***	***A2***	***A3***	***A4***	***A5***	***A6***	***A7***	***A9***	***A10***	***A11***	***A13***
Undiff1	15.62	27.01	28.52	26.29	30.04	22.74	25.52	28.41		28.93	32.36	34.62
Undiff2	16.5	27.71	28.69	26.7	29.1	23.65	27.35	28.91	33.95	29.94		35.97
Undiff3	16.12		29.05	26.74	29.72	22.89	25.7	31.19	36.93		34.05	33.62
Mean	16.08	27.36	28.76	26.58	29.62	23.1	26.19	29.5	35.44	29.44	33.21	34.74

Avg Ct values	***Gapdh***	***A1***	***A2***	***A3***	***A4***	***A5***	***A6***	***A7***	***A9***	***A10***	***A11***	***A13***
Undiff	16.08	27.36	28.76	26.58	29.62	23.1	26.19	29.5	35.44	29.44	33.21	34.74
Megs	16.18	27.13	29.16	27.54	30.45	25.16	28.27	29.55	36.96	30.37	33.94	34.44

Relative transcript abundance of *Hoxa* genes in megakaryocytes (relative to transcript abundance in undifferentiated HPC7 cells)

$$\text{Relative value} = 2^{\wedge -(\Delta\Delta Ct)}$$
$$\Delta\Delta Ct = \Delta Ct_{\text{megs}} - \Delta Ct_{\text{undiff}}$$
$$\Delta Ct_{\text{undiff}} = (Ct_{Hoxa} - Ct_{Gapdh}) \textit{ in undiff}$$
$$\Delta Ct_{\text{megs}} = (Ct_{Hoxa} - Ct_{Gapdh}) \text{ in megs}$$

Transcript abundance	**Undiff**	***A1***	***A2***	***A3***	***A4***	***A5***	***A6***	***A7***	***A9***	***A10***	***A11***	***A13***
Megs	1	1.25	0.8	0.55	0.6	0.25	0.25	1.03	0.37	0.55	0.64	1.31

Supplementary data 2

Distribution of histone marks in undifferentiated HPC-7 cells on *Hoxa4*, *Hoxa5* and *Gapdh*

H3K4me3

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	2	18.7	1.79	1.15
Value 2	0.89	22.31	0.96	1.18
Value 3	2.03	20.53	1.32	1.27
Average	1.64	20.51	1.35	1.2
Std error	0.65	1.80	0.42	0.06

H3K9ac

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	2.44	6.58	2.2	2.56
Value 2	3.03	15.88	2.49	2.56
Average	2.735	11.23	2.345	2.56
Std error	0.42	6.57	0.20	0

H4K16ac

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Values 1	1.21	1.77	1.24	1.16
Value 2	2.77	5.94	2.51	2.06
Average	1.99	3.855	1.875	1.61
Std error	1.10	2.95	0.89	0.64

H3K27me3

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	2	1.14	1.84	1.08
Value 2	1.6	1.19	1.1	1.34
Average	1.8	1.165	1.47	1.21
Std error	0.28	0.03	0.52	0.18

H3K9me2

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	2.92	2.69	2.8	1.71

H3K9acS10p

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	1.23	7.21	2.05	2.78
Value 2	0.8	7.16	1.57	1.64
Value 3	1.73	7.88	1.57	2.49
Average	1.25	7.41	1.73	2.30
Std error	0.46	0.40	0.27	0.59

H3K4me3

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	2.34	0.72	30.06	2.02	1.55
Value 2	0.96	0.32	25.28	1.7	1.14
Average	1.65	0.52	27.67	1.86	1.34
Std error	0.97	0.28	3.38	0.23	0.29

H3K9ac

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	1.72	2.82	10.92	1.99	1.43
Value 2	0.59	0.39	10.12	0.49	1.61
Average	1.15	1.60	10.52	1.24	1.52
Std error	0.79	1.72	0.56	1.06	0.12

H3K27me3

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	1.63	0.88	2.18	1.45	1.22
Value 2	0.37	0.269	0.73	1.13	1.18
Average	1	0.57	1.45	1.29	1.2
Std error	0.89	0.43	1.02	0.22	0.02

d

H3K9acS10p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	2.19	4.34	12.04	1.95	1.49
Value 2	2.02	2.01	19.29	1.97	1.34
Value 3	3.07	8.81	19.29	2.41	1.9
Average	2.42	5.05	16.87	2.11	1.57
Std error	0.56	3.45	4.18	0.26	0.28

H3K4me3

	Gapdh/0	*Gapdh/1*	*Gapdh/2*	*Gapdh/3*
Value	nd	1.23	6.91	2.94

H3K27me3

	Gapdh/0	*Gapdh/1*	*Gapdh/2*	*Gapdh/3*
Value	nd	0.44	0.96	1.23

e

Supplementary data 3

Distribution of histone marks in megakaryocytes on *Hoxa4*, *Hoxa5* and *Gapdh*

H3K4me3

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	3.57	6.48	2.31	2.14
Value 2	0.7	1.35	2.18	1.02
Average	2.13	3.91	2.24	1.58
Std error	2.03	3.62	0.09	0.79

H3K9ac

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	4.4	10.55	27.85	13.36
Value 2	6.32	11.63	17.38	5.46
Average	5.36	11.09	22.61	9.41
Std error	1.35	0.76	7.40	5.58

H4K16ac

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	1.38	1.12	1.73	0.5

H3K27me3

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	1.62	1.23	2.31	0.69
Value 2	0.14	0.37	1.28	nd
Average	0.88	0.8	1.795	0.69
Std error	1.04	0.61	0.73	nd

H3K9me2

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	0.96	0.98	1.89	0.85
Value 2	0.24	1.74	3.24	0.12
Value 3	1.36	3.89	3.03	1.04
Average	0.85	2.20	2.72	0.67
Std error	0.56	1.51	0.72	0.48

H3K9acS10p

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	0.14	0.31		0.6
Value 2	1.26	1.86	1.47	1.04
Average	0.7	1.08	1.47	0.82
Std error	0.79	1.09	nd	0.31

H3K4me3

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	2.9	3.78	1.95	2.77	3.45
Value 2	0.42	0.38	0.81	0.56	0.71
Average	1.66	2.08	1.38	1.66	2.08
Std error	1.75	2.40	0.80	1.56	1.93

H3K9ac

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	11.15	18.63	11	12.81	6.36
Value 2	6.49	0.6	9.71	6.58	5.09
Average	8.82	9.61	10.35	9.69	5.72
Std error	3.29	12.74	0.91	4.40	0.89

H3K27me3

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	1.76	1.17	1.08	0.84	0.61
Value 2	0.18	0.88	0.55	0.87	
Average	0.97	1.02	0.81	0.85	0.61
Std error	1.11	0.20	0.37	0.02	nd

H3K9acS10p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	0.26	0.44	0.51	0.38	0.16
Value 2	8.87	5.31	6.68	8.11	8.63
Average	4.56	2.875	3.59	4.24	4.39
Std error	6.08	3.44	4.36	5.46	5.98

H3K4me3

	Gapdh/0	*Gapdh/1*	*Gapdh/2*	*Gapdh/3*
Value	4.28	8.93	8.87	3.97

H3K27me3

	Gapdh/0	*Gapdh/1*	*Gapdh/2*	*Gapdh/3*
Value 1	nd	1.4	0.98	1.17
Value 2	1.98	1.7	2.62	3.58
Average	1.98	1.55	1.8	2.37
Std error	nd	0.21	1.16	1.70

Supplementary data 4

Distribution of proteins in undifferentiated HPC-7 cells on *Hoxa4*, *Hoxa5* and *Gapdh*

MLL^{C}

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	10.33	28.44	10.7	8.81	12.21
Value 2		16.56	7.31	3.07	5.38
Average	10.33	22.5	9.01	5.94	8.79
Std error	nd	8.40	2.39	4.05	4.82

MLL^{N}

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	2.98	10.85	5.65	3.5	3.41

menin

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	17.02	29.65	16.56	12.46	19.15
Value 2		54.56	22.47	8.22	17.75
Average	17.02	42.10	19.51	10.34	18.45
Std error	nd	17.61	4.18	2.99	0.98

Pol II

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	0.99	1.33	3.32	2.92	1.31
Value 2	1.04	0.87	5.65	1.15	0.43
Average	1.02	1.1	4.45	2.03	0.87
Std error	0.04	0.32	1.64	1.24	0.62

CBP

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	2.2	6.4	13.27	4

MLL4/2

	Hoxa4/1	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	3.16	8.39	4.05	4.56

Msk1

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	nd	2.42	8.16	0.58	4.46
Value 2	nd	0.94	11	2.8	2
Average	nd	1.68	9.58	1.69	3.23
Std error	nd	1.04	2.01	1.56	1.73

MLL^C

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value 1	2.67	11.47	2.08	4.72
Value 2	3.83	10.7	1.68	
Average	3.25	11.08	1.88	4.72
Std error	0.82	0.54	0.28	nd

menin

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	7.78	18.37	5.35	nd

Pol II

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	7.21	17.38	6.06	nd

CBP

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	2.29	6.96	2.73	5.46

MLL4/2

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	4.25	8.57	2.49	7.41

Msk1

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value 1	2.78	2.51	2.62	3.86
Value 2	1.33	4.62	1.79	3.27
Value 3	2.67	3.58	2.34	2.46
Average	2.26	3.57	2.25	3.19
Std error	0.80	1.05	0.42	0.70

MLL^{C}

	Gapdh4/0	*Gapdh4/1*	*Gapdh4/2*	*Gapdh4/3*	*Gapdh4/4*
Value	nd	6.58	1.97	2.78	2.2

menin

	Gapdh4/0	*Gapdh4/1*	*Gapdh4/2*	*Gapdh4/3*	*Gapdh4/4*
Value	nd	4.71	0.52	1.235	1.798

Supplementary data 5

Distribution of proteins in megakaryocytes on *Hoxa4*, *Hoxa5* and *Gapdh*

MLL^{C}

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	0.22	1.76	1.89	1.59	1.21
Value 2	2.04	0.77	0.52	7.21	1.08
Average	1.13	1.26	1.20	4.4	1.14
Std error	0.91	0.49	0.68	2.81	0.06

MLL^{N}

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	1.95	1.35	1.58	0.83	1.49

menin

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	0.86	0.82	1.7	0.79	1

Pol II

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value 1	0.61	0.51	4.59	0.47	0.98
Value 2	2.04	1.89	3.65	0.68	1.77
Average	1.32	1.2	4.12	0.57	1.37
Std error	1.01	0.97	0.66	0.14	0.55

Msk1

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*
Value	1.2	0.9	1.27	0.72	1.51

MLL^{C}

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	1.8	2.63	2.67	1.54

menin

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	1.3	1.42	1.75	1.11

Pol II

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value 1	3.05	3.73	4.99	2.32
Value 2	2.65	2.8	10.62	6.9
Average	2.85	3.26	7.80	4.61
Std error	0.28	0.65	3.98	3.23

Msk1

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*
Value	1.5	1.1	1.51	1.54

MLL^{C}

	Gapdh4/0	*Gapdh4/1*	*Gapdh4/2*	*Gapdh4/3*	*Gapdh4/4*
Value	0.97	1.17	2.63	1.54	0.17

menin

	Gapdh4/0	*Gapdh4/1*	*Gapdh4/2*	*Gapdh4/3*	*Gapdh4/4*
Value	5.31	7.83	19.83	13.26	12.04

Supplementary data 6

Distribution of histone marks on *Hoxa4* and *Hoxa5* in the control undifferentiated HPC-7 cells

H3K9acS10P

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*	*Hoxa4/5*
Value 1	1.36	2.81	21.89	4.66	0.83	1
Value 2	1.27	4.15	21.10	0.17	0.90	1
Average	1.32	3.48	21.49	2.41	0.86	1
Std error	0.06	0.94	0.56	3.17	0.04	0

H3S10p

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*	*Hoxa4/5*
Value 1	0.11	0.54	0.35	0.13	0.57	1
Value 2	0.32	0.42	0.45	0.07	0.71	1
Average	0.22	0.48	0.40	0.10	0.64	1
Std error	0.14	0.08	0.07	0.03	0.10	0

H3T3p

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*	*Hoxa4/5*
Value 1	0.77	0.25	0.38	0.29	0.67	1
Value 2	1.11	0.65	0.37	0.11	0.52	1
Average	0.94	0.45	0.37	0.20	0.60	1
Std error	0.24	0.28	0.01	0.12	0.10	0

NB: Values for the *Hoxa4/5* primer set has been set at 1, and data for the other primer sets are relative to this one.

H3K9acS10p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	1.37	4.43	62.25	0.69	1
Value 2	1.94	18.89	73.00	0.80	1
Average	1.66	11.66	67.62	0.75	1
Std error	0.40	10.22	7.59	0.07	0

H3S10p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	1.32	4.05	0.77	0.86	1
Value 2	0.99	2.88	0.97	1.29	1
Average	1.16	3.47	0.87	1.08	1
Std error	0.23	0.82	0.14	0.29	0

H3T3P

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	1.10	2.47	0.87	0.76	1
Value 2	1.43	5.49	0.60	0.65	1
Average	1.26	3.98	0.73	0.70	1
Std error	0.23	2.13	0.18	0.07	0

NB: Values for the *Hoxa5/5* primer set has been set at 1, and data for the other primer sets are relative to this one.

Supplementary data 7

Distribution of histone marks on *Hoxa4* and *Hoxa5* in the undifferentiated HPC-7 cells after treatment with 10nM of okadaic acid

H3K9acS10p

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*	*Hoxa4/5*
Value 1	2.33	6.08	20.19	1.21	1.16	1
Value 2	4.22	12.55	49.80	2.20	2.65	1
Average	3.27	9.31	34.99	1.71	1.91	1
Std error	1.33	4.58	20.94	0.69	1.05	0

H3S10p

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*	*Hoxa4/5*
Value 1	2.33	1.10	0.66	0.27	0.94	1
Value 2	2.95	1	0.38	0.91	1.44	1
Average	2.64	1.05	0.52	0.59	1.19	1
Std error	0.43	0.07	0.19	0.45	0.35	0

H3T3p

	Hoxa4/0	*Hoxa4/1*	*Hoxa4/2*	*Hoxa4/3*	*Hoxa4/4*	*Hoxa4/5*
Value 1	1.32	0.81	0.28	0.90	1.43	1
Value 2	3.73	1.47	0.87	0.73	2.02	1
Value 3	1.36	1.22	0.71	0.28	1	1
Average	2.14	1.16	0.62	0.64	1.48	1
Std error	1.37	0.33	0.30	0.31	0.51	0

H3K9acS10p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value	4.28	101.92	695.22	3.25	1

H3S10p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	0.93	1.65	0.53	1.05	1
Value 2	13.75	138	17.25	17.33	1
Average	7.34	69.82	8.89	9.19	1
Std error	9.05	96.40	11.81	11.51	0

H3T3p

	Hoxa5/1	*Hoxa5/2*	*Hoxa5/3*	*Hoxa5/4*	*Hoxa5/5*
Value 1	0.36	0.28	0.37	0.78	1
Value 2	0.46	0.44	0.56	1.35	1
Average	0.41	0.36	0.47	1.07	1
Std error	0.07	0.10	0.13	0.40	0

Supplementary data 8

PCR Primer Set used for ChIP

Gene Name	Forward 5'-3'	Reverse 5'-3'	Tm
Gapdh-0	tgagcctcctccaattcaac	ccaggaagacgcttaaaag	58°C
Gapdh-1	caccatccgggttcctataa	attttcacctggcactgcac	58°C
Gapdh-2	cgtcccgtagacaaaatggt	tgaccttgaggtctccttgg	58°C
Gapdh-3	tcaggtcccctgtttcttgt	gaatttgccgtgagt	58°C
Gapdh-4	tggacagcactgacttccag	agcccctcttcatggatacc	58°C
HoxA4/0	ccctctctgctccatttcag	cctttcctgaaccttccaca	58°C
HoxA4/1	gccgttgttctatcctgctc	tgatgcctcactcgtacctg	58°C
HoxA4/2	agaggcctaggacagacgtgtggatgctgctagccttcag		58°C
HoxA4/3	ccctggatgaagaagatcca	agggaccctgggtacaaaag	58°C
HoxA4/4	tgttggaaggaagccagact	aaaatcccccaaactgctct	58°C
HoxA4/5	attgtctgccctttgtcagg	caggggtcgttttggtaaga	58°C
HoxA5/1	atcggctctggctactgaaa	agtcgctcccaagctgtaaa	58°C
HoxA5/2	agccggggaaataaagttgt	ggggtcgaattgaggttaca	58°C
HoxA5/3	agccacaaatcaagcacaca	agatccatgccattgtagcc	58°C
HoxA5/4	gcaagctgctctttctgctt	cttctggcctgaggtttctg	58°C
HoxA5/5	tcttttgcaggctacgtgtg	caggcagagggaaaagtttg	58°C
MLL1	ggctccagcaagaacaaaagtcacacctgcaaatgagagc		55°C
MLL2	cgctgttccatcagtgagaa	cttcgcctttcaggtactcg	55°C
MLL3	gctccctcagcaaagacaac	gctgtgtctcagccttttcc	55°C
MLL5	agtcctccacacctcacacc	gttgcctgctgaggattagc	55°C
Actin	ggacttcgagcaagagatgg	agcactgtgttggcgtacag	55°C
mPU.1	atggaagggttttccctcaccgccc	gtccacgctctgcagctctgtgaa	54°C
GATA1	ggaattcgggccccttgtgaggccagagag	cggggtacctcacgctccagccagattcgaccc	
GPVI	atgtctccagcctcacccac	ctaggccagtgggagggg	54°C