

# EPIGENETIC REGULATION AT MLL1 TARGET GENES

By  
MAAIKE WIERSMA



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

Chromatin and Gene Expression Group  
College of Medical and Dental Sciences  
University of Birmingham  
September 2014

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

# Abstract

The mixed-lineage leukaemia 1 protein is a histone methyl-transferase that deposits the gene activating H3K4 trimethyl mark, and is often mutated in leukaemia. MLL1 is normally associated with a cohort of cofactors, but the mechanisms regulating the histone methyl-transferase activity remain unclear. Here I examine the role of Msk1, a downstream kinase of the MAP-kinase pathway, in regulating MLL1 activity. Msk1 is known to deposit the H3S10 phosphorylation mark, which was found to stimulate MLL1's methylation activity *in vitro*.

Here I demonstrate that MLL1 and Msk1 can be immunoprecipitated and their patterns of genomic binding show an overlap at ~30 of sites, suggesting a direct functional interaction.

In transient MLL1 and Msk1 knock-down cells, known MLL1 target genes were down-regulated and at a global level, 30% of all responding genes were regulated in the same manner. Furthermore, key histone modifications at MLL1 target genes change in Msk1 knock-down cells, suggesting that histone cross-talk within the MLL1 complex acts as a means of gene regulation. Finally, cell cycle studies suggest MLL1-Msk1 cross-talk may stimulate MLL1-driven gene expression after mitosis. These findings suggest that MLL1 is regulated by Msk1 and therefore by extracellular signals via the MAP-kinase pathway.

# Acknowledgments

I am using this opportunity to express my appreciation to all those who provided me the possibility to complete my PhD. I am grateful for their guidance, criticism, advice and support throughout the course.

I specially thank my supervisor Dr Karl Nightingale, whose suggestions and encouragement helped me to coordinate my project and to write this report. I consider myself fortunate to have had a supervisor who cared so much about my work, and who was always approachable.

Furthermore I would like to acknowledge Prof Bryan Turner for welcoming me to the Chromatin and Gene expression group, supporting my work with materials and stimulating ideas. My thanks also go to Dr Laura O'Neil for her well appreciated counsel.

I also would like to express my warmest gratitude to all members of the Chromatin and Gene Expression group, who not only helped me with techniques, analysis and proof-reading, but also created an enjoyable working atmosphere. Thank you, Hannah, Edith, John, Charlotte and Nil.

Beyond science I want to thank my friends and my fellow PhDs (Sarah in particular) for their moral support, fun times outside the lab and keeping things in perspective.

Finally, I must express my sincerest gratitude to my parents and my sisters, who are there and supported me throughout all ups and downs (not only) of my research. I am so thankful to have you.

# Abbreviations

The 'International System of Units' (SI) was used.

°C	Degree Celsius
3'	3-prime
5'	5-prime
ac	Acetylation
ADP	Adenosine diphosphate
AF10	Interacts with Dot1 Histone Methyl-transferase
AF4	Founder of AF4 Family, Member of EAP (ENL associated proteins)
AF9	ENL Homologue
ALL	Acute Lymphoblastic Leukaemia
AML	Acute Myeloid Leukaemia
Amp	Ampicillin
ASH	(Absent, Small, or homeotic)-Like
AT-hook	Adenine-Thymine-Hook
BAF	Brahma-Related Gene 1-Associated Factor
BHC80	Component of BRAF-HDAC complex (BHC)
BIR	Baculovirus IAP Repeat
BMI1	B-lymphoma Moloney murine leukemia virus insertion region-1
bp	Base-Pair
BPTF	Bromodomain PHD-Finger Transcription-Factor
BRD4	Bromodomain-Containing Protein 4
brm-gene	Brahma Gene
BSA	Bovine Serum Albumin
CBP	CREB-Binding Protein
CCHIP	Carrier Chromatin Immuno-Precipitation
cDNA	Complementary DNA
ChIP	Chromatin Immuno-Precipitation

CO <sub>2</sub>	Carbon Dioxide
CON	Control
CpG island	Cytosine-Phosphate-Guanine
CREB-ATF1	cAMP Response Element Binding Protein-Activation Transcription Factor-1
CTCF	CCCTC-Binding Factor
C-terminal	Carboxy-Terminal
CTK	C-Terminal Kinase
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
dH <sub>2</sub> O	Distilled Water
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
Dot1	Disruptor of Telomeric Silencing 1
DPF	Double PHD Finger
DPY-30	dpy-30 homologue (C. elegans)
DTT	Dithiothreitol
EAP	ENL Associated Proteins
EDTA	Ethylene diamine tetra-acetate
ELL	Polymerase II Elongation Factor
ENL	Binds histone H3, assembles EAP nuclear elongation/chromatin modifying complex
ENN	Interacts with Histone Methyl-transferase PRMT1
EP300	E1A-Binding Protein p300; also known as p300, Histone Acetyltransferase
ERK	Extracellular Signal-Regulated Kinases
FBS	Foetal Bovine Serum
FC	Fold-Change
FDR	False Discovery Rate
FITC	Fluorescein Isothiocyanate
FOSL1	FOS-like antigen 1
FPLC system	Fast Protein Liquid Chromatography System
FRET	Fluorescence Resonance Energy Transfer

G1/S/G2/M	Cell cycle stages: Gap-phase 1/ DNA-synthesis phase/ Gap-phase 2/ Mitosis
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GFP	Green Fluorescent Protein
GTP	Guanosine Triphosphate
h	hour
HDAC	Histone Deacetylase
HEK	Human Embryonic Kidney
HEPES	N-2-Hydroxyethylpiperazine-N'-2-Ethanesulfonic Acid
HMT	Histone Methyl-transferase
HOX	Homeobox
HP1	Heterochromatin Protein 1
HPC2	Human Polycomb protein 2
Ifnar2	Interferon alpha/beta receptor 23
IgG	Immunoglobulin G
ING2	Inhibitor of growth protein 2
IP	Immuno-Precipitation
JARID	Jumonji, AT-Rich Interactive Domain
JMJD	Jumonji domain-containing protein
kb	kilo-base
KD	Knock-Down
Kif20b	kinesin family member 20B
LB	Lysogeny Broth or Luria-Bertani medium
LCL	Lymphoblastoid Cell Line
LEDGF	Lens Epithelium-Derived Growth Factor
LSD1	Lysine-Specific Demethylase 1
MACS	Model-based Analysis of ChIP-Seq
MAP	Mitogen-Activated Protein
MAPK	Mitogen-Activated Protein Kinase
MAPKK	Mitogen-Activated Protein Kinase Kinase
MAPKKK	Mitogen-Activated Protein Kinase Kinase Kinase

MBLR	Mannma binding lectin receptor
MBT	Malignant Brain Tumour
me	Methylation
MEK	Mitogen Activated Protein/Extracellular Signal-Regulated Kinase
MEM	Minimum Essential Medium
min	Minute
MKK	MAP (Mitogen-Activated Protein) Kinase Kinase
MLL	Mixed Lineage Leukaemia
MLL <sup>C</sup>	Mixed Lineage Leukaemia, C-terminus
MLL <sup>N</sup>	Mixed Lineage Leukaemia, N-terminus
MOF	orthologue of Drosophila males absent on the first (MOF)
mRNA	Messenger Ribonucleic Acid
Msk	Mitogen- and Stress-activated protein kinase
MT	Methyl-transferase-Homology
MW	Molecular Weight
ncRNA	Non-Coding Ribonucleic Acid
NFκB	Nuclear Factor Kappa B
NGF	Nerve Growth Factor
N-terminal	Amino-terminal
NTK	N-terminal Kinase
NURF	Nucleosome Remodelling Factor
OD	Optical Density
p/ph	phosphorylation
PBS	Phosphate-Buffered Saline
PCG	Polycomb-Group Proteins
PCR	Polymerase Chain Reaction
PEI	Poly Ethylene Imine
pH	Potentia hydrogenii
PHD	Plant-Homeo-Domain
PI	Propidium Iodide



PMSF	Phenylmethylsulfonylflouride
Pol II	RNA Polymerase II
pTEFb	Positive Transcription Elongation Factor b
Puro	Puromycin
QC	Quality Control
qPCR	quantitative PCR
R	Arginine
RbBP5	Retinoblastoma-Binding Protein 5
RNA	Ribonucleic Acid
rpm	Rotation per minute
RPMI	Roswell Park Memorial Institute
RT	Room-Temperature
RT- qPCR	Reverse Transcription quantitative PCR
S	Serine
SDS	Sodium-Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
sec	second
SET	Su(var)3-9 and 'Enhancer of zeste'
shRNA	short Hairpin RNA
SL2	Schneider line-2
SNL	Subnuclear Localization Signals
SV-40	Simian virus 40
T	Threonine
TAD	Transactivation Domain
TBE	Tris Borate EDTA
TBS	Tris Buffered Saline
TE	Tris EDTA
TEB	Triton Extraction Buffer
TEMED	N, N, N', N'-Tetramethylethylendiamine
Tet3	Tetracycline Resistance Gene

TNF $\alpha$	Tumour Necrosis Factor $\alpha$
Tpl2	Tumour Progression Locus 2
Tris	Tris-(hydroxymethyl)-aminomethane
TRITC	Tetramethyl Rhodamine Iso-Thiocyanate
TRX	Trithorax
TSS	Transcription Start Site
Ttk	TTK protein kinase
Tween	Polyoxyethylene-sorbitan-monolaurate
U	Unit
ubi	Ubiquitination
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
UV	Ultraviolet
v/v	Volume/Volume
w/v	Weight/Volume
WD40	Domain with conserved Tryptophan (W) and Aspartate (D) repeats
WDR5	WD-repeat protein-5
XChIP	Formaldehyde Cross-linking Chromatin Immunoprecipitation

# Table of Content

.....	1
1 Introduction.....	1
1.1 Chromatin Organization.....	3
1.2 Histone Modifications.....	8
1.3 Histone Crosstalk.....	13
1.4 Readout of histone modifications.....	16
1.5 The histone methyltransferase MLL1.....	20
1.6 MLL1 in leukaemia.....	24
1.7 MLL1 targets the Hox genes.....	26
1.8 MLL1 in the context of histone modifications.....	29
1.9 Msk1.....	31
1.10 Aims.....	35
2 Materials and Methods.....	36
2.1 Tissue Culture.....	36
2.1.1 Cultivation of mouse embryonic fibroblasts (MEFs).....	36
2.1.2 Cultivation of LCL cells.....	37
2.1.3 Cultivation of SL2 cells.....	37
2.1.4 Cultivation of HEKs.....	37
2.2 Transfection methods.....	38
2.2.1 Amplification of transfection vectors.....	38
2.2.2 Transfection of MEFs.....	40
2.2.3 Transfection of HEKs.....	40
2.3 Protein Methods.....	41
2.3.1 Preparation of whole cell extract.....	41
2.3.2 Isolation of histones.....	41
2.3.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).....	42
2.2.4 Staining protein gels.....	42
2.3.5 Immuno-detection of proteins on nitrocellulose membranes.....	43
2.4 Functional analysis.....	45

2.4.1 RNA extraction .....	45
2.4.2 Quantitative PCR (qPCR) .....	45
2.4.3 Reverse Transcription quantitative PCR (RT-qPCR).....	46
2.4.4 Formaldehyde Cross-linking Chromatin Immunoprecipitation (XChIP).....	47
2.4.5 Carrier Chromatin Immunoprecipitation (CChIP).....	49
2.4.6 FLAG purification.....	51
2.4.7 Size exclusion chromatography.....	52
2.4.8 Elutriation of LCLs.....	52
2.4.9 Flow Cytometer Analysis (FACS).....	53
2.4.10 Immuno-fluorescence Microscopy.....	54
2.4.11 Co-immunoprecipitation .....	55
2.4.12 Gene Expression Arrays.....	56
3 Results .....	57
3.1 Characterizing the interaction of MLL1 and Msk1 .....	57
3.1.1 Co-Immunoprecipitation experiments with MLL1 and Msk1 .....	57
3.1.2 FLAG-tagged MLL1 complex.....	64
3.2 MLL1 and Msk1 binding in vivo .....	68
3.2.1 Establishing an XChIP protocol for MEFs .....	68
3.2.2 Analysis of the sequencing results of MLL1 and Msk1.....	71
3.3 Characterising the functional impact of MLL1 and Msk1 regulation .....	77
3.3.1 Establishment of a transient knock-down of MLL1 and Msk1 in MEFs .....	79
3.3.2 Changes of MLL1 target genes in knock-down cells .....	87
3.3.3 Changes of gene expression on a genome wide level in knock-down cells.....	89
3.3.5 Changes of histone modifications on MLL1 target genes in MLL1 and Msk1 knock-down cells.....	99
3.4 MLL1 and Msk1 during the cell cycle .....	105
3.4.1 Elutriation of LCLs.....	106
3.4.2 Analysis of MLL1 and Msk1 .....	110

3.4.3 Analysis of histone modifications.....	114
4 Discussion .....	122
4.1 Epigenetic regulation by MLL1 .....	122
4.2 Is the Histone kinase Msk1 involved in MLL1 regulation? .....	124
4.3 Do MLL1 and Msk1 physically interact? .....	125
4.4 Do MLL1 and Msk1 bind to the same genomic sites? .....	127
4.5 What is the functional impact of MLL1 and Msk1 interaction? .....	128
4.6 Do MLL1 and Msk1 change during the cell cycle? .....	131
4.7 How do the results support our hypothesis? .....	133
4.8 Perspectives .....	135
5 Bibliography .....	138

# Table of Figures and Tables

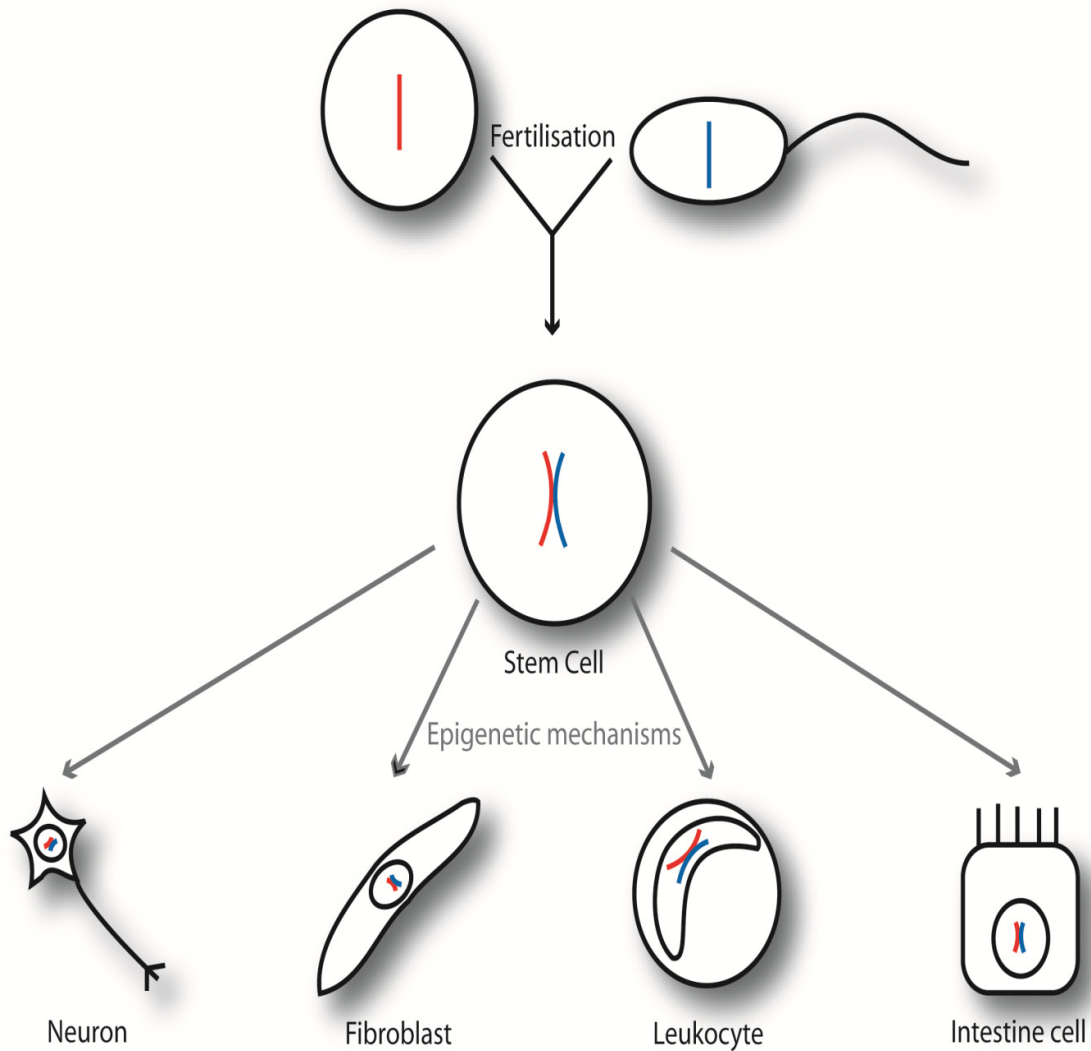
Figure 1.1: Development and Epigenetic mechanisms.....	2
Figure 1.2: Chromatin Organization. ....	7
Figure 1.3: The Nucleosome.....	12
Figure 1.4: Histone cross-talk. ....	15
Figure 1.5: Histone modification binding domains.....	19
Figure 1.6: MLL domain structure and the MLL complex. ....	23
Figure 1.7: Hox genes in fly, mouse and human. ....	28
Figure 1.8: Msk1 is a downstream target of the MAPK signalling pathway. ....	33
Figure 1.9: Msk1 domain structure and comparison to Msk2.....	34
Figure 2.1: Vector maps. ....	39
Table 2.1: Antibodies used for Western analysis, including source and used dilutions .....	44
Table 2.2: qPCR primer sets for ChIP analysis .....	46
Table 2.3: Commercial primers used for RT-qPCR.....	47
Table 2.5: Antibodies used for ChIP.....	49
Table 2.4: Antibodies used for Immunofluorescence microscopy.....	55
Table 2.5: Antibodies used for Co-IP .....	56
Figure 3.1: Representation of the Co-IP experiments. ....	59
Figure 3.2: Co-Immunoprecipitation experiments for MLL1 and Msk1. ....	61
Figure 3.3: Co-immunoprecipitation for MLL1 and Msk1.....	63
Figure 3.4: Overexpression of FLAG-MLL1 in HEK293s. (.....	65
Figure 3.5: MLL1-FLAG complex. ....	67
Figure 3.6: Chromatin preparation for Next Generation Sequencing. ....	70
Figure 3.7: Sequencing Analysis – peak identification.....	72
Figure 3.8: Sequencing Analysis – TSS probe analysis. ....	74
Figure 3.9: An analysis of MLL1 and Msk1 binding correlations. ....	75
Figure 3.10: Experimental design/workflow. ....	78
Figure 3.11: Establishing MEF transfection protocols. ....	80
Figure 3.12: Establishing Msk1 knock down protocol. ....	82
Figure 3.13: Establishing the Msk1 knock-down protocol.....	84

Figure 3.14: Establishing MLL1 knock-down protocol.....	86
Figure 3.15: Effect of MLL1 or Msk1 knock-down on HoxA and Meis1 gene expression. ....	88
Figure 3.16: Microarray analysis of MLL1 and Msk1 knock-down cells. ....	91
Figure 3.17: Significance analysis of gene expression changes for MLL1 and Msk1 KD. ....	92
Figure 3.18: Cluster and gene annotation analysis I.....	95
Figure 3.19: Cluster and gene annotation analysis II.....	96
Figure 3.20: Validation of microarray data by qPCR.....	97
Figure 3.21: Chromatin Isolation from knock-down cells.....	100
Figure 3.22: Effect of MLL1 and Msk1 knock-down on H3K4me3, H3K27me3 and H3K9ac/S10p distribution at HoxA4. ....	103
Figure 3.23: Effect of MLL1 and Msk1 knock-down on H3K4me3, H3K27me3 and H3K9ac/S10p distribution at HoxA5. ....	104
Figure 3.24: The elutriation system. ....	107
Figure 3.25: Representative FACS analysis of elutriated cells.....	109
Figure 3.26: MLL1 and Msk1 abundance during the cell cycle. ....	111
Figure 3.27: Nuclear distribution of MLL1 and Msk1 in G1 phase. ....	112
Figure 3.28: MLL1 and Msk1 distribution on mitotic chromosomes ....	113
Figure 3.29: Histone modification abundance during the cell cycle. ....	115
Figure 3.30: H3K4me3 and H3K9ac distribution in G1 phase cells.....	117
Figure 3.31: Histone modifications on mitotic chromosomes.....	118
Figure 3.32: H3K4me3 and H3K9ac distribution on mitotic chromosomes ....	119
Figure 3.33: H3K27me3 and H3S10p distribution on mitotic chromosomes. ....	120
Figure 4.1: Msk1 is a member of the MLL1 complex.....	134

# 1 Introduction

Development starts with fertilization, resulting in a stem cell with a unique and individual set of genes. This stem cell gives rise to every cell in the final organism, which all carry the same genetic information, but vary significantly in their phenotype defined by characteristic patterns of gene expression and silencing (Figure 1.1). These patterns are established by epigenetic mechanisms that do not alter the DNA sequence and need to be maintained through DNA replication, chromatin assembly and DNA condensation in mitosis. Bird defines epigenetic events as “the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states” (Bird 2007). Although epigenetic events can be short lived, under certain circumstances they need to be stable to be passed onto daughter cells, which led to the hypothesis of the “epigenetic memory” of the cell (Riggs & Porter 1996). The foundation of all epigenetic mechanisms are proteins that can establish, read, modify and erase epigenetic modifications and regulate the access of the transcription machinery to the DNA (Inbar-Feigenberg et al., 2013). Several mechanisms are known, including DNA methylation (Berger et al. 2009), post-translational histone modification (Turner 2005), histone variants (Khorasanizadeh 2004), nucleosome positioning (Schones et al. 2008) and ncRNA (ENCODE project Consortium 2012), which work together to establish the ‘epigenetic state’ of individual genes and generally the whole cell (Jirtle & Skinner 2007). However, epigenetic mechanisms do not only define constitutive gene expression, but also contribute to the response to extracellular signalling (Jaenisch & Bird 2003). The epigenome therefore has an important regulatory function and directly links the environment to gene activity (Franklin & Mansuy 2010).





**Figure 1.1: Development and Epigenetic mechanisms.** All the cells in an individual carry the same underlying genetic information, which was created at conception. Despite the common DNA-sequence, differentiated cells vary in their phenotype due to different patterns of gene expression. Epigenetic mechanisms play an important role in establishing these patterns during development.

## 1.1 Chromatin Organization

In prokaryotes, genes are in an unrestricted ground state and gene expression is only dependent on RNA Polymerase II binding to the promoter. The eukaryotic genome however is much more complex than the prokaryotic, with long stretches of non-protein-coding sequences. Hence gene expression not only depends on RNA Polymerase II access, but also on transcription factor binding and regulatory sequences, which can be several megabases away from the promoter (Struhl et al. 1999). In order to initiate transcription 'transcription factories', a large battery of proteins (e. g. transcription factors), associate at active genes and participate in transcription and co-transcriptional RNA processing (Buckley & Lis 2014). A key event is the recruitment of RNA Polymerase II and the general transcription factors at the core promoter, forming the Pre-Initiation-Complex, in which the polymerase is in a poised state. Only after its phosphorylation at the C-terminal domain (Phatnani & Greenleaf 2006) does Polymerase II leave the complex and move on to the elongating state (Müller & Tora 2014). The complexity of the eukaryotic genome and the gene expression require a high level of organization of the DNA in the nucleus.

Therefore in eukaryotes the primary role of the nucleus is the storage, retrieval and replication of genetic information, which requires the organization and compaction of ~2 m of DNA in mammals (Woodcock & Ghosh 2010). In Figure 1.2 the different levels of chromatin organisation are shown. The 2nm DNA fibre is wrapped twice around the nucleosome, forming an open zigzag string or 11nm fibre, often referred as the "beads on a string" structure (Olins & Olins 1974). Adjacent nucleosomes are linked with DNA, which varies in length in a cell- and species-specific manner. Nucleosomes subsequently self-assemble into the "30nm fibre", a regular array in which nucleosomes are packed on top of

one other (Robinson & Rhodes 2006). The 30nm is the first level of transcriptionally inactive chromatin. It has been assumed that chromosomes in interphase consist of either the 11nm (active) or 30nm (inactive) chromatin fibre (Gall 1966, Dorigo et al. 2004). Chromatin can be further compacted, aided by chaperone proteins, which prevent inappropriate interaction. The mechanisms however are unknown (Hansen et al. 2010). *In vitro* studies on chromosomes propose that condensin acts as a scaffolding and interlinking protein, which supports higher order structures (Maeshima & Eltsov 2008). During mitosis chromosomes are visible. They are the most condensed form of DNA and about 10,000 times shorter than the non-condensed DNA fibres (Felsenfeld & Groudine 2003).

However, recent reports claim, that this model of hierarchical DNA packaging (which is mostly based on *in vitro* studies) needs to be revised. It is now known that chromosomes are divided into hundreds of domains with different protein compositions (Bickmore & van Steensel 2013). New technologies (like Chromatin Conformation Capture techniques) have found that *in vivo* chromatin is packed into 11-nm fibres, which form 3-dimensional chromatin domains independent of higher order structures (Razin & Gavrilov 2014). It has been shown, that interphase chromosomes occupy distinct chromosome territories. These 'topological domains' have well defined boundaries, which are engaged with proteins such as the insulator protein CTCF (Bell et al. 1999) and are conserved across different species (Dixon et al. 2012).

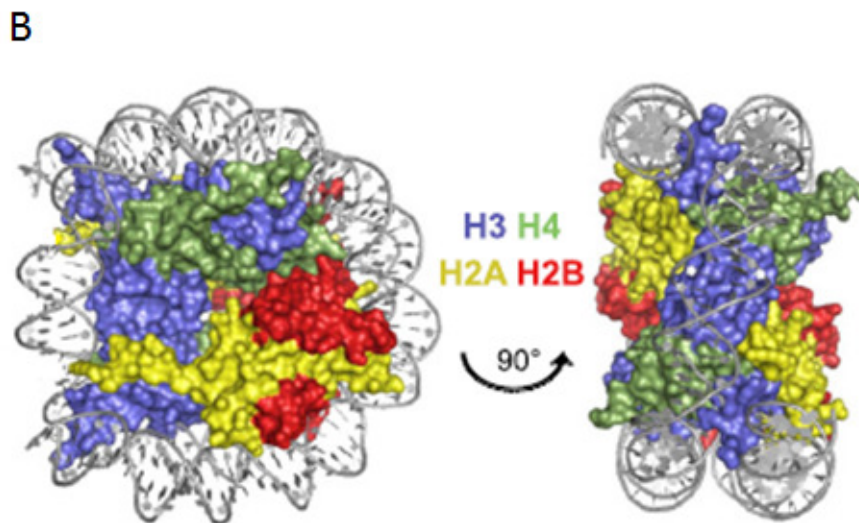
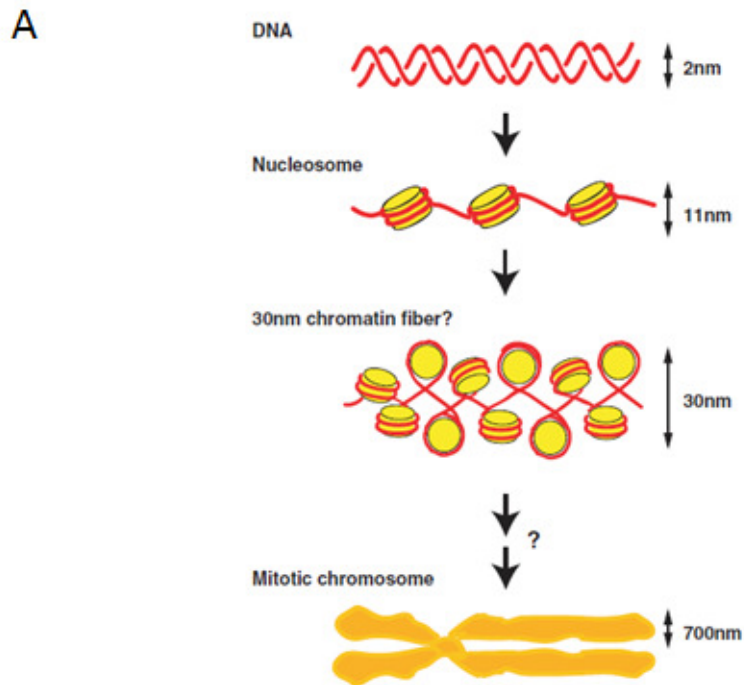
In chromatin formation the nucleosome plays a central role, as the basic unit of chromatin condensation. It is basically the same in all eukaryotes and is one of the most highly conserved structures (Luger et al. 1997). The nucleosome consists of a stretch of DNA, which is wound around eight histone proteins. Two copies of each histone protein (H2A, H2B, H3

and H4), as shown in Figure 3.2 B, assemble together and form a histone octamer. Around this 145-147 base-pairs of DNA are wrapped in a left handed super-helix (Luger et al. 1997). The shape and the charge on the nucleosome surface facilitate most of the histone-histone and histone-DNA interactions. This makes the nucleosome an essential component for higher chromatin structure (Garcia-Ramirez et al. 1992; Chodaparambil et al. 2007). From this nucleosome core the N-terminal domains of the histones (“tails”) protrude (Figure 3.3 A). They are flexible and highly mobile with a large number of charged residues (Banères et al. 1997). In low salt conditions they dissociate from their original binding positions on the nucleosome and therefore cannot be crystallized. However, *in vivo* studies have shown that they can make contact with other histones (du Preez & Patterson 2013). The N-terminal tails can interact within the same nucleosome but also with neighbouring nucleosomes. It is now thought that the tails support the formation of the higher order structures and may even be involved in the binding to the DNA or chromatin proteins (Zheng & Hayes 2003). For example tails of histone H4 bind to defined regions (‘acidic patches’) of neighbouring nucleosomes and help to compact the higher order structure of chromatin (Kalashnikova et al. 2013). In higher eukaryotes the nucleosome is stabilized by the linker histone H1, which sits on top of the core-nucleosome. H1 is positively charged and stabilizes the binding of the negatively charged DNA to the nucleosome (Harshman et al. 2013). Furthermore, recent findings suggest that histone H1 binding, in concert with the length of linker DNA between two nucleosomes, may have an influence on the shape and formation of the 30 nm fibre (Routh et al. 2008).

Although chromatin needs to be highly organized, it also needs to be kept flexible and accessible in order to be able to respond to environmental and developmental stimuli (Lange

et al. 2011). In general there are two different types of chromatin. In 1928, Heitz observed, that some chromatin regions stayed condensed during the cell cycle (Heitz 1928). He called those regions heterochromatin (Zacharias 1995). Today it is known, that heterochromatin is highly condensed and gene-poor. This is generally distinguished into constitutive heterochromatin, in which genes are poorly expressed and found at distinct genomic locations (eg. centromeres or telomeres, Tamaru 2010). In addition facultative heterochromatin exists, in genomic regions that can adopt an open or compact conformation and that are often associated with differentiation. The most prominent example for facultative heterochromatin formation is the inactivation of the second X-chromosomes in female cells (Trojer & Reinberg 2007). Euchromatin on the other hand is less condensed and transcriptionally active. It is thought that insulator elements in the DNA sequence prevent the interference of hetero- and euchromatin and which also block enhancer activity at inappropriate promoters. They are bound by proteins, like CTCF a highly conserved zinc-finger DNA binding protein (Bell et al. 1999; Ghirlando et al. 2012).

It is not yet understood how hetero- and euchromatin domains are formed. Yet nucleosome arrays are highly dynamic even in a very condensed state. In a study, which used FRET microscopy on fluorescence-labelled nucleosome arrays, it was shown that chromatin undergoes conformational fluctuations, exposing sites on the nucleosome (Poirier et al. 2010). These spontaneous chromatin fibre dynamics allow for the invasion and action of DNA processing protein complexes and may allow chromatin conformation changes (Poirier et al. 2010). Nevertheless, it is now accepted that the chromatin state influences gene transcription (Dillon 2004) and that compaction of chromatin leads to inhibition of transcription initiation and elongation by RNA polymerase II (Hansen & Wolffe, 1992)



**Figure 1.2: Chromatin Organization.** (A) The DNA double-helix is first wrapped twice around the nucleosome core, the fundamental packaging unit of chromatin, leading to an open “bead on a string” structure. This fibre is then further condensed into the 30-nm fibre and subsequently forms unknown higher order structures. Finally the chromosome in mitosis is the most condensed form of chromatin (Maeshima & Eltsov 2008). (B) X-ray crystal structure of the nucleosome core, which consists of histone proteins (H2A, H2B, H3 and H4 (Weber & Henikoff 2014, adapted from Luger et al. 1997).

## 1.2 Histone Modifications

Nucleosomes are an important structural component of chromatin. However, the histone tails, which project from the nucleosome core, are targeted by specialized enzymes that deposit modifications on specific amino acid residues after translation (Bhaumik et al. 2007). Most common are acetylation of lysines, methylation of lysines and arginines, and phosphorylation of serines, threonines or tyrosines. These modifications are small chemical groups and are deposited at distinct residues of the histone tails by specialized histone modifying enzymes, for example methyltransferases deposit and demethylases remove methyl-groups (Figure 1.3 A & B). These enzymes normally work in pairs, in which one class of enzymes oppose each others effect. The level of histone modification depends on a dynamic equilibrium of the enzymes (Turner 2012). While residues can only be mono-acetylated or mono-phosphorylated, methylation shows a greater extent of variation. Lysine methylation can differ between mono-, di- or tri-methylation, whereas di-methylated arginines can be symmetric or asymmetric (Kouzarides 2007). However the landscape of post-translational modifications is more complex with many more modifications known and modifications found on the fold of the N-terminal domains or even within the globular domain (Rothbart & Strahl 2014). By now over 100 modifications are known with a great variety of formats (Xu et al. 2014). For example, peptides can be added to the histone tails, like ubiquitin (West & Bonner 1980) or SUMO (Shiio & Eisenman 2003), and complex chemical groups like ADP-ribose (Martinez-Zamudio & Ha 2012) or biotin (Camporeale et al. 2004). However, post-translational modifications can also affect the histone tail itself, like deimination (conversion of arginine residues into citrulline, Cuthbert et al. 2004) or histone tail clipping (histone H3 is shortened by cleavage after alanine 21, Santos-rosa et al. 2012).

It is generally accepted, that histone modifications play an important role in gene regulation, though the precise mechanism for all modifications is still elusive (Henikoff & Shilatifard 2011). Nevertheless, key histone modifications have been studied intensively and are associated with specific functions. Histone H3 lysine 9 acetylation (H3K9ac) and histone H3 lysine 4 tri-methylation (H3K4me3) are generally associated with active genes, whereas histone H3 lysine 27 tri-methylation (H3K27me3) is generally associated with silent genes (Barski et al. 2007). CHIP-Seq experiments generated high-resolution maps of these histone modifications across the genome, revealing their position across active or silent genes (Figure 1.3 C). These experiments showed a spike of H3K4me3 and H3K9ac over the promoter region (TSS) of active genes and a lower background level of H3K27me3. In contrast, on silent genes H3K27me3 rises slightly, whereas H3K4me3 and H3K9ac abundance stays at a background level (Yin et al. 2011).

In general, cell type specific gene expression is controlled by the availability of transcription factors and their accessibility to chromatin, which is controlled by epigenetic mechanisms, like histone modifications (Schones & Zhao 2008). So far three theories have been proposed to explain their functions.

(I) Histone tails contain a high percentage of positively charged arginine and lysine residues, which facilitates interaction with the negatively charged DNA. It has been shown that acetylation alters the charge of histone tails by neutralizing the positive charge of lysine residues. Hyper-acetylation of the histone tails leads to an accumulated effect of charge neutralisation of the tails and therefore to an opening of chromatin structure allowing gene transcription (Dion et al. 2005). For example, artificially induced hyper-acetylation at H4K16 inhibited the formation of nucleosome arrays into more compact fibres (Shogren-Knaak et



al. 2006). A similar role has been proposed for phosphorylation, which brings a negative charge to the histone tail and therefore may create repulsion between the negatively charged DNA and the histone (Mahadevan et al. 1991; Banerjee & Chakravarti 2011).

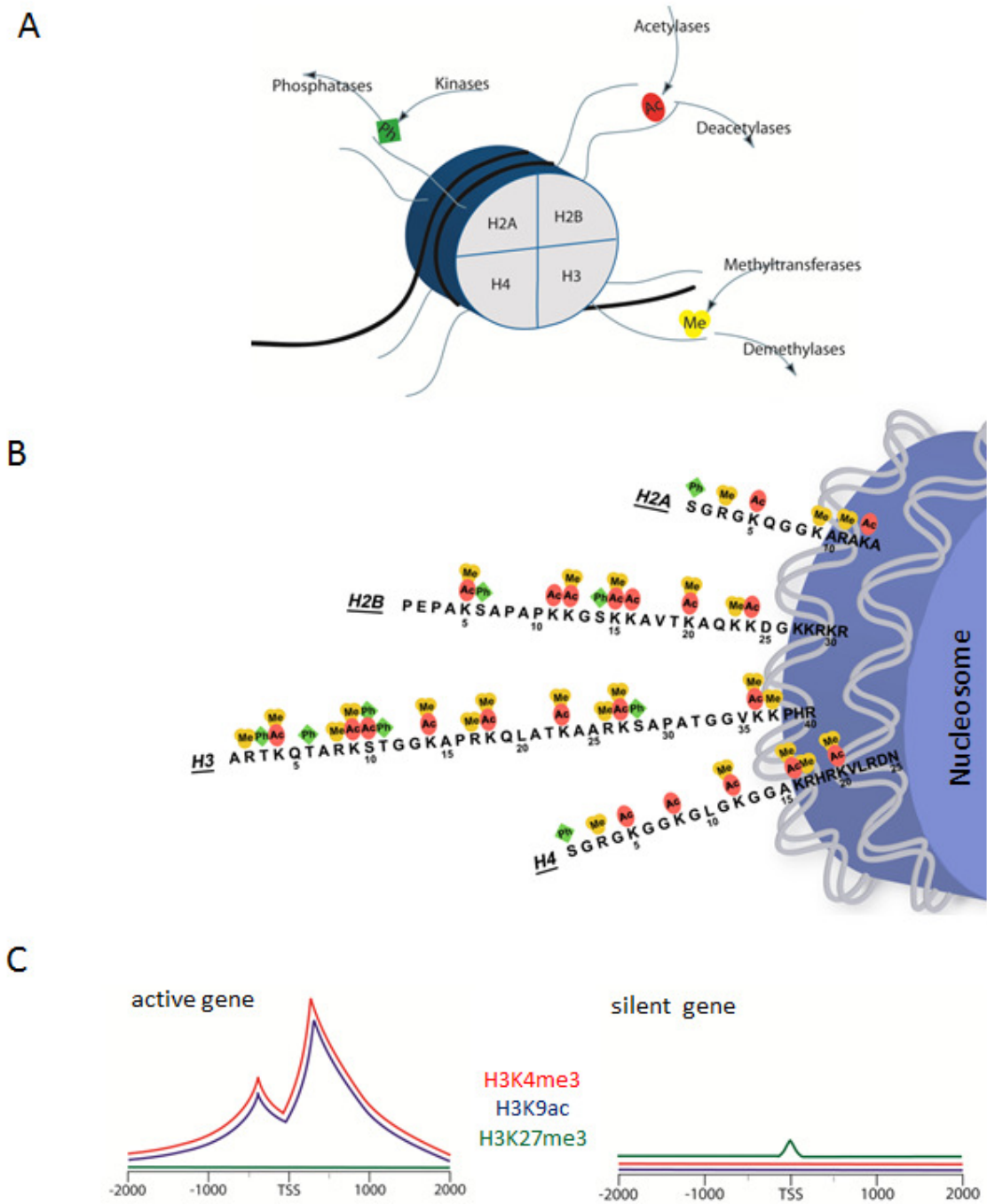
(II) Certain histone modifications are associated with specific transcriptional states. Acetylation in general is associated with actively transcribed genes, as well as H3K4me3. H3K27me3 on the other hand is associated with repressed genes (Lee & Mahadevan 2009). In the “histone code” model it is suggested that chromatin binding proteins can be recruited by histone modifications as they contain specialized domains, which recognize defined histone modifications. Thus different histone modifications can function together to signal distinct chromatin states (Strahl & Allis 2000).

(III) More recently it has been proposed, that histones work as signalling integration platforms. There are a variety of histone modifications and most do not seem to have a direct functional output. This leads to the suggestion that histone modifications summate the response to multiple external signals, which results in only a small directed output. Although signals on their own can either activate or inactivate gene expression they are interpreted in the context of the histone landscape. Thus, although every incoming signal changes the histone modification landscape, it is proposed that only the entirety of the modifications is interpreted, via recruiting proteins to the DNA (Badeaux & Shi 2013).

Histone modifications can be stable and unchanging throughout a cell’s life, or highly dynamic and rapidly changing, appearing or disappearing on a nucleosome within minutes after an external stimulus (Kouzarides 2007). For example the inactivation of one X-chromosome in female mammalian cells is associated with the stable establishment of facultative heterochromatin. This is hypoacetylated and hallmarked with H3K27me3,

H3K9me2, H4K20me1 and H2AK119ub1, whereas constitutive heterochromatin is hypoacetylated and marked with H3K27me1, H3K9me3 and H4K20me3. Nonetheless the composition of histone modification on the X chromosome can vary in different regions and probably the combination of different modifications signal the chromatin state (Brinkman et al. 2006). In contrast, a study on histone modifications in yeast during replication has shown that replication is accompanied by dynamically changing modifications, especially histone H3 and H4 acetylation, such as H3K56ac, which increases from low levels in G1 to high levels in S-phase and then drops back (Unnikrishnan et al. 2010).

Additionally different modifications have different turnovers. Histone methylation usually has a slower turnover than acetylation or phosphorylation, which only have a half-life of 1-5 minutes. Furthermore, silencing marks are more stable than activating histone marks (Clayton et al. 2006; Zee et al. 2010). This indicates that long-term changes, as well as immediate, short term changes to the chromatin landscape can be mediated by histone modifications (Lee & Mahadevan 2009).



**Figure 1.3: The Nucleosome.** (A) Nucleosomes are histone octamers with ~146 bp DNA wrapped around it. Histones have two domains: the globular domain and the N-terminal 'tail' domain. The globular domain forms the core of the nucleosome, around which the DNA is wrapped, while the tail domains protrude from it. (B) Modifications are highly specific for distinct residues on the tails (Ph=phosphorylation, Me=methylation, Ac=acetylation, figure by Charlotte Rutledge). (C) ChIP-Seq experiments generate high resolution maps of key histone modification distribution across genes. These indicated the average pattern of modifications across the transcriptional start site (TSS) (adapted from Barski et al. 2007 and Yin et al.

### 1.3 Histone Crosstalk

The vast number of histone modifications in a small area implicates that they influence each other, providing regulatory cross-talk. For example histone H3 phosphorylation at serine 10 (H3S10p) is involved in two opposite chromatin states. Mitotic chromosomes are highly phosphorylated and have very compacted chromatin (Hsu et al. 2000). On the other hand active genes are also marked with H3S10p, implicating that it's readout is influenced by other modifications (Mahadevan et al. 1991; Nightingale et al. 2007).

Cross-talk can take place in various forms (Figure 1.4). It can be limited to one histone tail (cis) or can occur between modifications on different histone tails (trans). The outcome of the affected modification can be enhanced (positive cross-talk) or can be blocked (negative cross-talk) (Nightingale et al. 2006). Finally it can be direct, affecting the histone modifications on the tails, or it can be indirect, affecting histone binding proteins (Latham & Dent 2007; Fischle 2008).

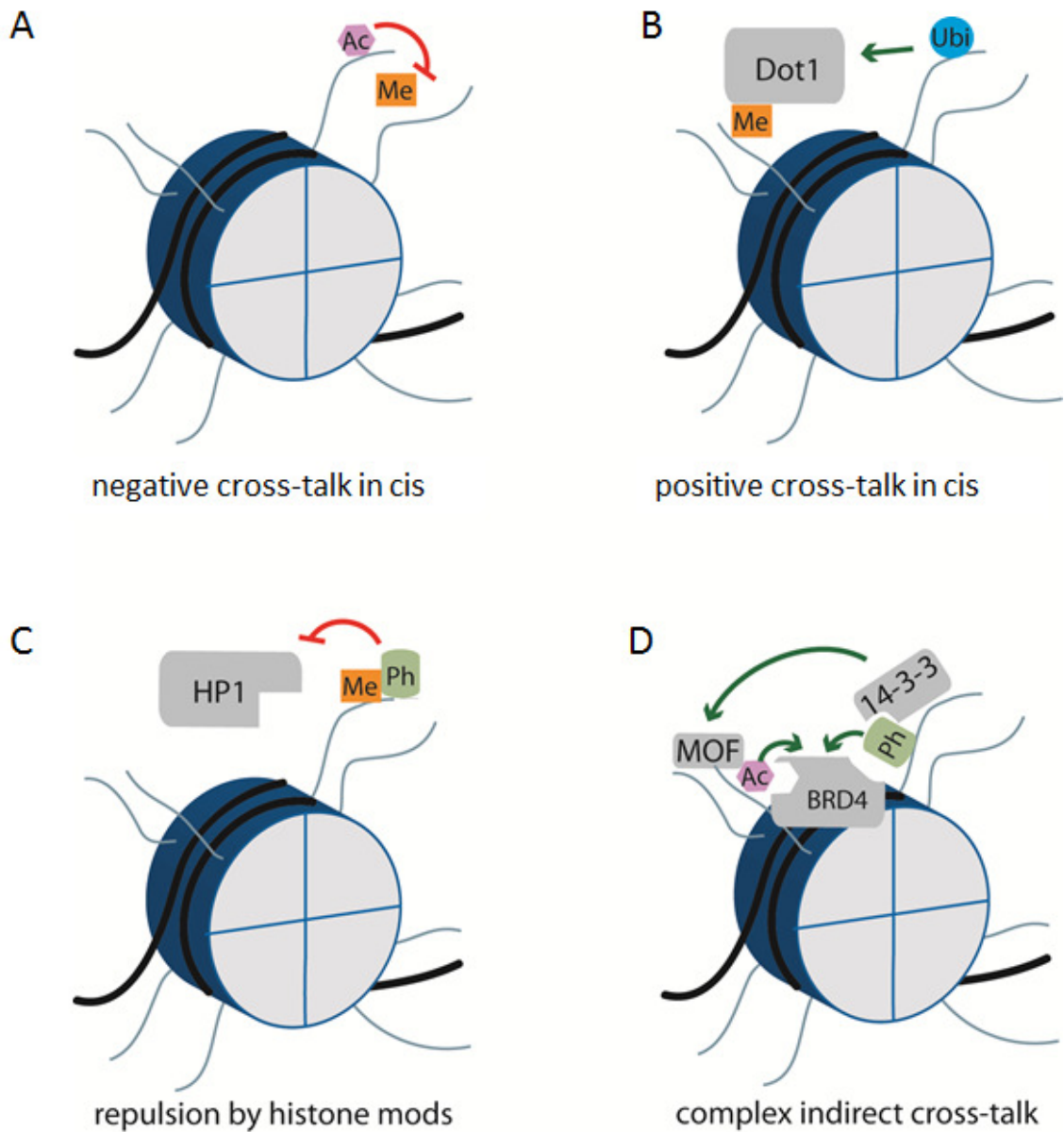
The most obvious form of direct cross-talk is shown in Figure 1.4 A. Some residues can be modified in different ways. H3K4 can be either acetylated or methylated, so blocking each other. Another example of negative direct cross-talk is in ES cells where key developmental genes are kept in a bivalent state by marking them with the activating H3K4me3 mark, as well as with the repressive H3K27me3 mark, neutralizing the outcome of the individual modifications (Bernstein et al. 2006).

However cross-talk can also involve the enzymes, which deposit modifications. Dot1 is a methyltransferase that targets H3K79, yet for efficient methylation, H2BK123 needs to be ubiquitinated (Figure 1.4 B). In contrast, for the ubiquitination of H2B, methylation of H3K79 is not required (Ng et al. 2002).

In Figure 1.4 C, an example of indirect cross-talk is given. The HP1 protein is an essential regulator of epigenetic gene expression and important for chromatin stability during mitosis. It can bind to methylated lysine 26 on histone H1.4, though when the same histone is phosphorylated at serine 27, HP1 cannot bind anymore (Daujat et al. 2005).

Another level of complexity in cross-talk was described for the regulation of the FOSL1 gene (Figure 1.4 D). After stimulation the FOSL1 enhancer is phosphorylated at H3S10. The 14-3-3 protein can bind then to the phosphorylated mark and activate the acetyltransferase MOF, which deposits H4K16 acetylation. At this point BRD4 can bind to the two marks and trigger the release of the paused polymerase II at the FOSL1 promoter (Zippo et al. 2009).

Crosstalk of different types of histone modifications acting in a combinatorial manner, allows fine-tuning of nuclear events and the integration of different cellular signalling pathways at the chromatin level (Zhang & Pradhan 2014). It also enables the histones to carry a vast amount of information. For example the first 10 amino acids of histone H3.1 (ARTKQTARKS) can be targeted by 19 different modifications, some of them exclusive. This results in 60 different modified peptides, expanding the number of possible single modified histone tails (Schwammle et al. 2014). More complexity is added by the fact that histone modifications also cross-talk to other epigenetic pathways, like DNA methylation and ncRNAs (Suganuma & Workman 2011).



**Figure 1.4: Histone cross-talk.** Histone modifications can influence each others functional output. Here are some examples given with more details in the text. (A) Negative cross-talk means that one modification blocks another one. Here H3K4ac prevents H3K4me3 in cis (on the same histone) (B) Positive cross-talk means one modification enhances another modification. Here H2B ubiquitination stimulates Dot1 methyltransferase activity in trans (on another histone). (C) Cross- talk can be indirect not affecting any histone modifications. For example HP1 cannot bind to its normal target H1.4K26 methyl-mark, when blocked by the H1.4S27 phospho-mark. (D) Cross-talk can be very complex. Here the 14-3-3 protein binds to H3S10p, which then activates MOF, which deposits the H4K16 acetyl-mark. This allows the binding of BRD4 to H3S10p and H4K16ac.

## 1.4 Readout of histone modifications

The histone modification landscape is a flexible, constantly changing environment. This requires specialized proteins, which can either deposit histone modifications ('writer'), remove them ('eraser') or recognize them ('reader') (Strahl & Allis 2000). The level of histone modifications at a locus is therefore determined by the balanced activity of the writer and the eraser proteins (Figure 1.3 A). For example, acetylation is deposited by acetyltransferases and removed by deacetylases (Wang et al. 2009). Phosphorylation is set by kinases in response to upstream signalling pathways and opposed by phosphatases (Shimada et al. 2010). Lysine methylation is catalysed by methyltransferases, which nearly all contain catalytic SET domains and removed by demethylases (Black & Whetstone 2013). Histone modifications provide highly selective binding sites for reader proteins, a class of proteins that contain different highly conserved domains, which recognize histone modifications at certain residues (Mellor 2006). In Figure 1.5 a selection of them are shown.

The bromo-domain for example bind specifically to acetylated lysines (Sanchez et al. 2014). This domain was first discovered on the *brm*-gene in *Drosophila* and it was recognized that this 77 amino-acid motif is found in other eukaryotic regulatory proteins (Tamkun et al. 1992). Recognition of acetyl-lysine was thought to be limited to bromo-domains, but more recently the DPF (double PHD finger) domain of human DPF3b, which functions in association with the BAF chromatin remodelling complex to initiate gene transcription has been associated with binding acetyl marks (Zeng et al. 2010).

A variety of methylation binding proteins have been described, which is probably due to the variety of methyl marks. The chromo-domain of Polycomb and HP1 proteins of *Drosophila* was described as a methyl-binding motif that binds to H3K27me3 (Paro & Hogness 1991).

This domain is a member of the 'royal family' of domains, a structurally related group of folds, which bind methyl groups (Mallette & Richard 2012) and also includes the tudor-domain and/or the MBT domain. The tudor-domain has been reported to bind to different methyl-marks, for example in the SMN protein, which is linked to spinal-muscular atrophy, the tudor-domain binds to symmetrically di-methylated arginines (Selenko et al. 2001), whereas in the demethylase JMJD2A the domain binds to methylated H3K4 and H4K20 (Huang et al. 2006). The MBT (malignant brain tumor) domain can recognize mono- and di-methylated lysines at various positions on the histone tails. PHD (plant homeo domain) fingers (not a tudor domain family member) are found in a variety of different proteins, where they interact specifically with H3K4me3 (e. g. the PHD finger of the chromatin remodeller NURF; Wysocka et al. 2006). However, there is some overlap between their binding sites as some domains can recognize the same methylation marks (Bannister & Kouzarides 2011).

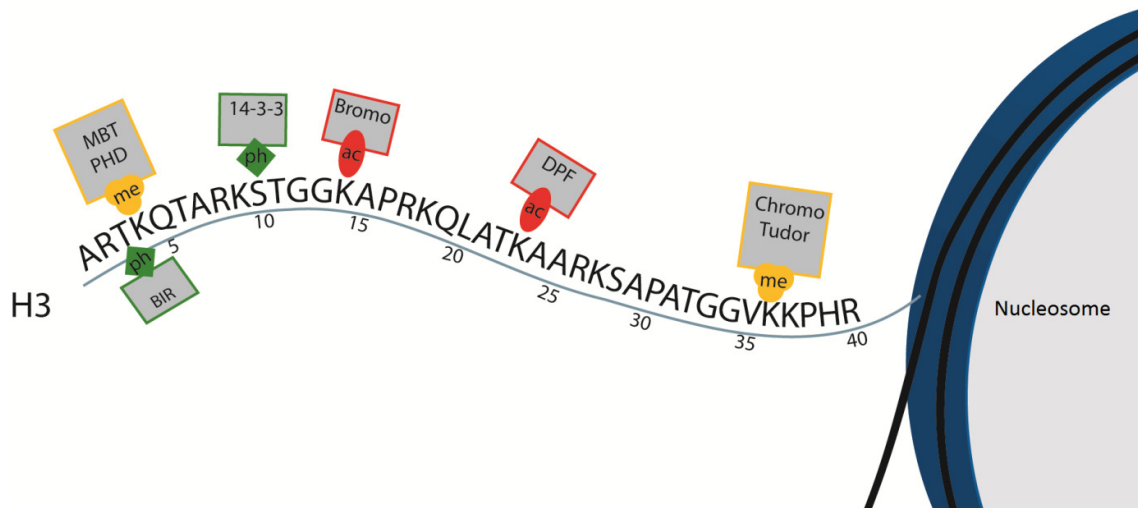
Phosphorylation is recognized by the well-conserved 14-3-3 family of phospho-specific binding proteins. These are independent dimeric proteins that offer two sites for protein binding and therefore can act as adaptors between phosphorylated histones (H3S10p and H3S28p) and the effector protein (Macdonald et al. 2005). Recently the BIR (Baculovirus IAP Repeat) domain of Survivin has been shown to bind to H3T3 phosphorylation (Niedzialkowska et al. 2012).

It has also been described that unmodified histone tails can be recognized by specialized PHD fingers and WD40 domains, which would make unmodified histone tails signals in their own right. For example the PHD finger of BHC80, a member of the LSD1 demethylase complex, binds specifically to H3K4me0 and this interaction is abolished when H3K4 is



methyated (Lan et al. 2007). Likewise, the human autoimmune regulator AIRE was also shown to interact with the unmodified H3K4 via its PHD fingers (Chakravarty et al. 2009).

The binding affinities of binding domains have been shown to be relatively weak, which allows rapid 'on-off' binding (Daniel et al. 2005). Also, binding domains are often found in close proximity to each other or contain several copies of a specific reader domain. For example, in ING2 a bromo-domain is found next to PHD fingers forming the BPTF (bromo-domain and PHD domain transcription factor) domain (Shi et al. 2011). This could allow co-binding, which would be expected to strengthen the binding of the effector protein to chromatin. Furthermore, the combinatorial readout of histone modification can modulate and specify the recruitment of reader proteins (Musselman et al. 2012).



**Figure 1.5: Histone modification binding domains.** Histone modifications are recognized by conserved binding protein modules, which are highly specific for distinct modifications at specific residues. Bromo-domains and DPF domains are known to bind to acetylated residues. Chromo-, Tudor- and MBT domains, as well as PHD fingers bind to methyl-marks. Phosphorylated residues are recognized by 14-3-3 proteins. It has also been reported that the BIR domain can bind to phospho-marks.

## 1.5 The histone methyltransferase MLL1

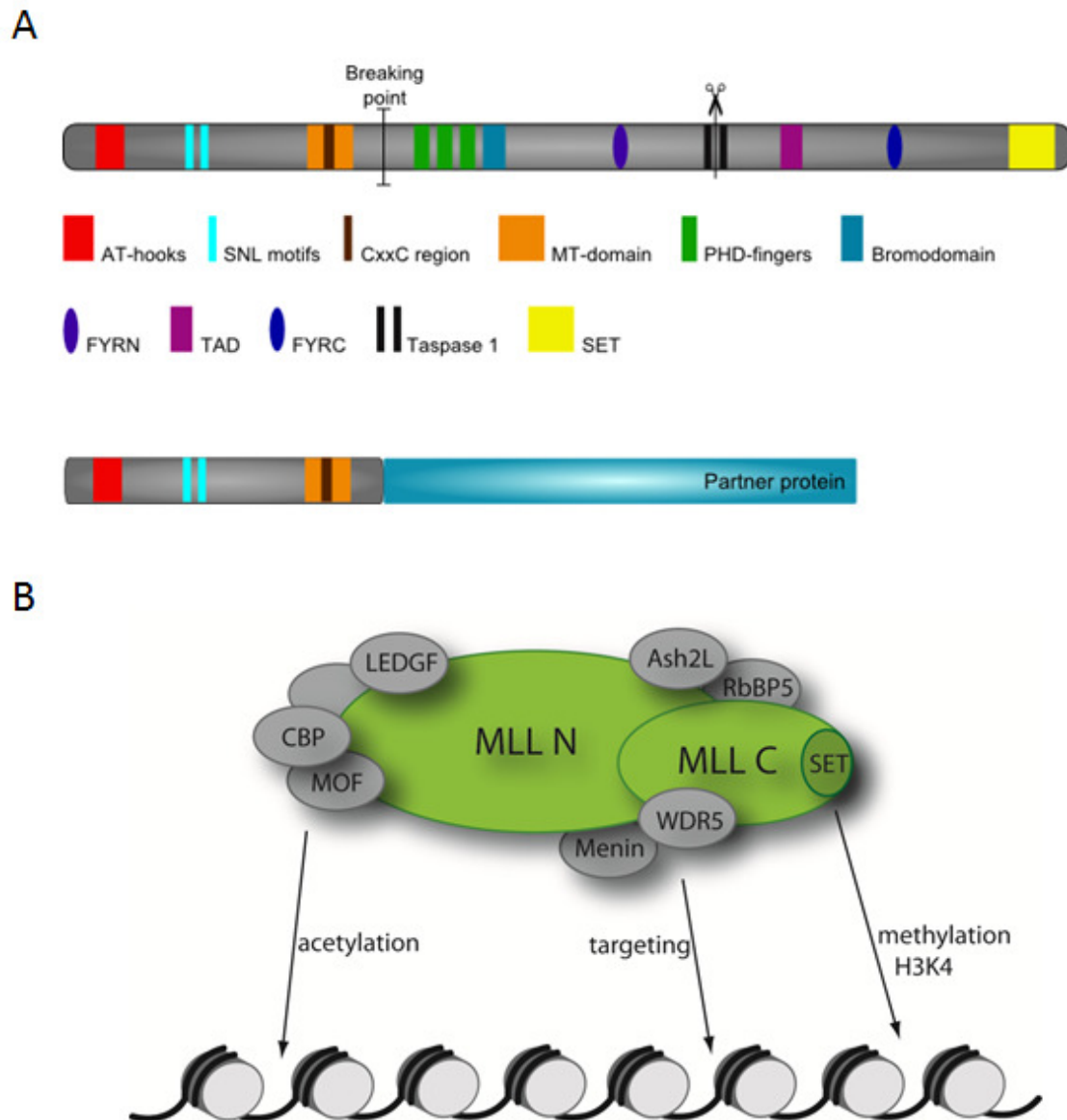
Eukaryotic gene expression is controlled by modification of chromatin structure through the enzymatic action of transcriptional regulators. MLL1 modifies chromatin structure by regulating the degree of H3K4 methylation (Liedtke & Cleary 2009). As it was first discovered about 20 years ago, chromosomal diagnostics of acute lymphoid leukaemia (ALL) and acute myeloid leukaemia (AML) patients, showed a recurring chromosomal translocation involving chromosome 11, band q23. The gene on this frequent chromosome break point was therefore termed mixed-lineage leukaemia (MLL) (Ziemin-van der Poel et al. 1991).

MLL1 is a large protein of 3900 amino-acids, which is cleaved by the Taspase 1 enzyme after translation into a 320kDa N-terminus (MLL<sup>N</sup>) and a 180kDa C-terminus (MLL<sup>C</sup>), which are bound together and form the MLL1 core complex (Yokoyama et al. 2002). The domain structure is complex (Figure 1.6 A). In control of the hetero-dimerization of the MLL1 core complex are the FYR-domains on the N- and the C-terminus (Ansari & Mandal 2010). MLL1 also contains multiple domains, which are involved in targeting MLL1 to its target sites. The AT-hooks on the N-terminus bind to the minor groove of AT-rich DNA without a specific recognition sequence (Zeleznik-Le et al. 1994). This domain is followed by two subnuclear localization signals (SNLs) and further downstream by the methyl-transferase-homology (MT) domain, which contains a specialized CxxC region and binds exclusively to non-methylated CpG islands, a characteristic of transcriptionally active genes (Birke et al. 2002). In addition, this region has also been shown to recruit repressing factors, like the polycomb repressor proteins HPC2 and BMI1, or the histone deacetylases HDAC1 and HDAC2 (Xia et al. 2003). MLL1 also contains reader domains, which bind to modified histones. The PHD-fingers bind to methylated lysines and the bromo-domain bind acetylated residues (Daser &

Rabbitts 2004). On the C-terminus the TAD and the SET domain can be found. The TAD domain is a conserved transcriptional activation domain, which binds to the coactivator CBP (Bannister & Kouzarides 1996; Arai et al. 2010). The SET (Suppressor of Variegation, Enhancer of Zeste, and Trithorax) domain is responsible for the transcriptional activation and holds the HMT activity (Milne et al. 2002; Ruthenburg et al. 2007). SET proteins are highly conserved. In *Saccharomyces cerevisiae*, Set1 is a member of the COMPASS (complex proteins associated with Set1) complex, which is essential for the mono, di- and tri-methylation of histone H3 at lysine4 and therefore the stimulation of transcription in yeast (Miller et al. 2001). In *Caenorhabditis elegans*, the SET-1 and SET-2 proteins have been identified as nematode counterparts to the yeast SET1 protein (Simonet et al. 2007). In *Drosophila melanogaster* Trithorax group proteins (TRX, ASH1 and ASH2) contain SET domains (Petruk et al. 2001), whereas in mammals at least six proteins contain SET domains: MLL1, MLL2, MLL3, MLL5, SET1A and SET1B (Wang et al. 2009). After the post-translational cleavage of MLL1, it is incorporated into a multi-protein complex (Figure 1.6 B). The core of the complex consists of MLL1 and the WARD proteins WDR5, Ash2L, RbBP5 and DPY-30. The association of these five proteins is sufficient to catalyse the methylation of histone 3 (Dharmarajan et al. 2012). It has been found that RbBP5 is essential to maintain mono- and di-methylation of histone H3, whereas Ash2L and DPY-30 are needed for the tri-methylation. These proteins are critical for the stability of the MLL1 complex, as in yeast the SET1-complex is shown to become unstable when any of the WARD proteins are depleted (Ernst & Vakoc 2012). On the other hand WDR5 can bind H3K4, but does not distinguish between its different methylation states. It was therefore thought that WDR5 presents the peptide for further methylation to the SET domain (Ruthenburg et al. 2006). The crystal

structure of MLL1 with the bound H3K4me3 mark however revealed that this is sterically not possible. Moreover WDR5 binding, in contrast to RbBP5 and Ash2L binding, did not stimulate methyl-transferase activity (Southall et al. 2009). It was shown that WDR5 binds to a conserved arginine containing motif (Win=WDR5 interaction motif) that is essential for the assembly of the MLL1 core complex and H3K4 dimethylation activity (Patel et al. 2008).

Furthermore, in the complex the acetyl-transferases CBP and MOF can be found. MOF specifically acetylates lysine 16 on histone 4, a modification known to relax chromatin structure via charge neutralization (Dou et al. 2005). CBP binds to MLL1's TAD domain and is necessary for its transcriptional activation (Arai et al. 2010). The complex also contains a number of proteins involved in targeting. It has been shown that MLL1 recruitment is highly dependent on menin (Milne et al. 2005). Later it was found that MLL1 and menin form a binding surface for LEDGF, which binds to chromatin via a PWWP domain, a motif structurally related to Tudor-/Chromo-domains (Yokoyama & Cleary 2008). The exact number of proteins, which are involved in this macromolecular complex it is still unknown with numbers ranging between 5 (WARD proteins) and up to 29 protein members (Nakamura et al. 2002). This is probably down to two main reasons. Firstly not all protein interactions are stable and it may be that some proteins are only transient members of the complex. For example MOF and MLL1 are recruited together (Dou et al. 2005), whereas CBP interacts only to activate transcription (Ernst et al. 2001). Secondly, it may also be that on different target sites, different proteins may be involved and/or that these proteins may be at the same target site, but not part of the MLL1 complex. In general it can be summarised that the MLL complex coordinates three major mechanisms of chromatin biology: methylation, acetylation and nucleosome remodelling (Slany 2009).



**Figure 1.6: MLL domain structure and the MLL complex.** (A) MLL1 has a complex domain structure. At the extreme end of the N-terminus AT-hooks can be found, next to Subnuclear-location domains, followed by a MT domain, with a characteristic CxxC-motif. These domains are involved in the non-specific binding of MLL1 to the DNA. A PHD-finger cassette for the recognition of methylated lysines and a bromo-domain for binding to acetylated histones are located at the N-terminus. On the C-terminus the SET domain is found, which holds the methyl-transferase activity. After translation the protein is cleaved by the Taspase1 enzyme at its cleavage site. The FYRN and FYRC domains are needed for the association of the two termini. In MLL fusion proteins the MLL1 protein breaks at a specific site and the truncated N-terminus is fused to a partner protein. Lower panel indicates a MLL-fusion protein, where the MLL1 protein is fused to one of ~70 partner proteins. (B) MLL1 is a large protein, which is cleaved into two fragments (MLL<sup>C</sup> and MLL<sup>N</sup>). They are bound tightly together, associate with a cohort of other proteins and form a multi-protein complex. These proteins can be involved in targeting (Menin, LEDGF), histone acetylation (CBP, MOF) or methylation (WDR5, RbBP5, Ash2L, DPY-30).

## 1.6 MLL1 in leukaemia

The MLL1 (mixed lineage leukaemia) gene is located on the long arm of chromosome 11, at locus 11q23, which is associated with aggressive types of leukaemia (Kaneko et al. 1982). Abnormalities of the MLL1 gene are found in ~70% of infant leukaemia and in ~10% of adult acute myeloid leukaemia (AML) (Felix 1998; Krivtsov & Armstrong 2007). Although the overall cure rate of children with leukaemia is good (78-85% at 5 years after the completion of therapy), the prognosis for leukaemia with MLL1 rearrangements is poor with an approximate cure rate of 20% (Slany 2005; Brassesco et al. 2009).

So far three mechanisms are known by which abnormalities at this location can transform MLL1 into an onco-protein. The most common mechanism is a balanced translocation that leads to the formation of fusion-proteins (see Figure 1.6 A). To date, nearly 100 different chromosome rearrangements have been described involving 11q23 and ~70 fusion-proteins have been characterized. These translocations seem to appear *in utero*, but it is unknown why chromosome 11 tends to break in this restricted 8.3 kb region (Ford 1993; Sim & Liu 2001). However, only five translocations occur in more than 66% of clinical cases with MLL rearrangements (MLL-AF4, MLL-AF9, MLL-AF10, MLL-ENL and MLL-ELL; Brassesco et al. 2009). The fusion-proteins share (with some exceptions) a common protein structure, in which only a small part of the N-terminus is conserved and fused together with the partner protein (Zhang et al. 2012). Less common transformations of the MLL1 gene are internal tandem duplications within the MLL coding region or amplification of the unaltered MLL1 gene itself (Basecke et al. 2006).

All fusion-proteins lose the C-terminus with the SET domain, which leads to a loss of H3K4 methylation and one would expect a decrease of target gene expression. However, the fusion partners induce transcriptional activity at MLL1 target genes, which leads to a gain of function and therefore they are crucial for leukaemogenesis (Dharmarajan & Cosgrove 2009). Despite the vast numbers of fusion partner proteins, according to Slany, only four molecular pathways lead to the oncogenic activation of MLL1 target genes (Slany 2009). (I) The most frequent fusion partners (eg MLL-AF4 and MLL-ENL) interact with the EAP complex. In this complex DOT1L, a histone H3K79 methyl-transferase, and pTEFb, which phosphorylates the C-terminal repeat domain of RNA polymerase II, can be found. (II) Fusion partners can be histone acetyltransferases (e.g. MLL-CBP or MLL-EP300) or (III) can bring additional methyl-transferase activity (eg MLL-EEN) to the target loci. (IV) It was also shown, that fusion partners supply dimerization domains, which leads to the dimerization of MLL1 and gene activation by unknown mechanisms (Slany 2009).

Although the mechanisms may vary, a hallmark of all MLL1 associated leukaemias is the irregular expression of MLL1 target genes. Especially HoxA7 and HoxA9 have been reported to be essential for the transformation of myeloid progenitors (Ayton & Cleary 2003). It was also shown that besides the up-regulation of these Hox genes, the up-regulation of the Hox cofactor Meis1 plays an important role in leukaemogenesis (Zeisig et al. 2004).



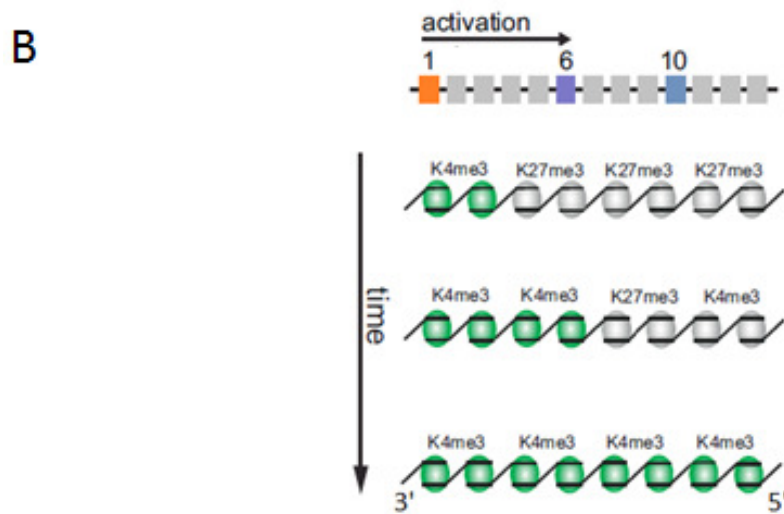
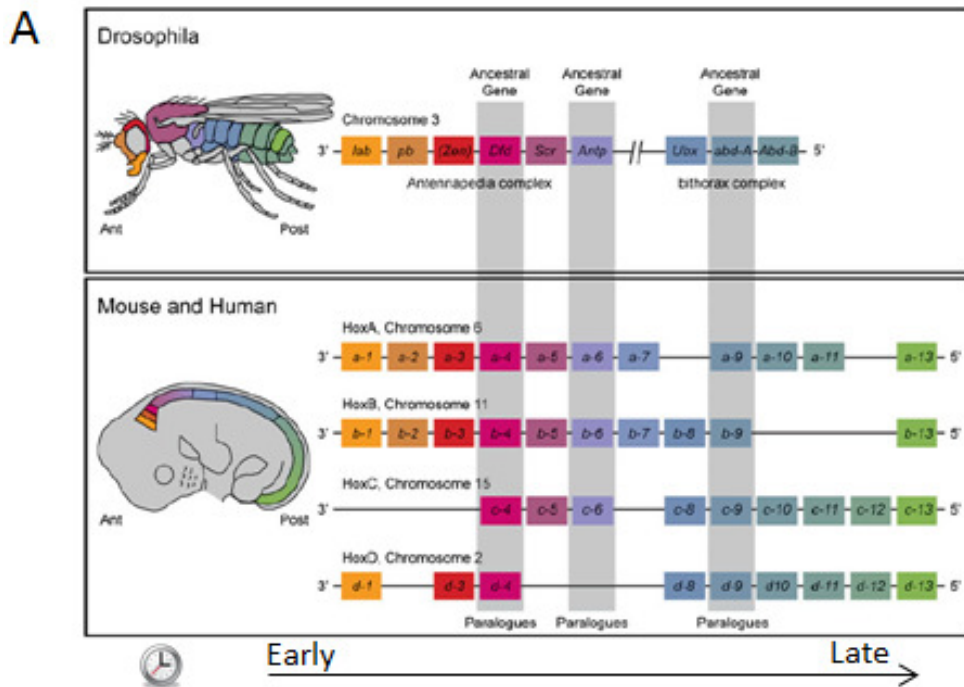
## 1.7 MLL1 targets the Hox genes

MLL1 is an essential protein. Knock-out studies in mice show, that MLL<sup>-/-</sup> embryos die at embryonic day 10.5. Embryonic lethality is accompanied by retarded growth, haematopoietic abnormalities and homeotic transformations of the axial skeleton (Yu et al. 1995; Yu et al. 1998). It was found that MLL1 is required for the maintenance of normal numbers of haematopoietic progenitors and their proper differentiation (Hess et al. 1997). These changes were accompanied by altered Hox gene expression and altered levels of histone methylation (Terranova et al. 2006). Later it was shown that in the adult, MLL1 also maintains haematopoietic stem cells and progenitors (Jude et al. 2007). Moreover, MLL regulates Hox gene expression through direct binding to promoter sequences (Milne et al. 2002) and by maintaining Hox gene expression, induces proliferation and differentiation of haematopoietic progenitors (Argiropoulos & Humphries 2007).

The Hox genes encode homeodomain transcription factors, which are critical for the proper establishment of regional identities in developing organisms. First discovered in *D. melanogaster*, mutations in the homeotic genes were found to lead to the developmental replacement of one body part by another (Frischer et al. 1986). They are highly conserved and in many species are grouped into genomic clusters with a similar organization. As shown in Figure 1.7 A, the single Hox gene cluster in fly was multiplied through subsequent gene duplication and in vertebrates four paralogous Hox clusters (Hox A, B, C and D) are found (Holland 2013).

As shown in Figure 1.7 B, the Hox genes are uniquely regulated as a result of their genomic organization. Their gene activity depends on their position in the cluster, what is known as 'collinearity' (Lewis 1978; Graham et al. 1989). This can take two forms. Firstly, the Hox genes, which are located at the 3' end of the cluster, tend to be expressed in more anterior regions of the organism, whereas genes at the 5' are expressed in the posterior region. This association between the position of the Hox gene in the cluster and its expression in the organism is referred to as spatial collinearity (Gaunt et al. 1988). Secondly, genes at the 3' are activated first and progressively the other Hox genes follow, according to their consecutive position in the cluster, which is known as temporal collinearity (Izpisua-Belmonte et al. 1991). Spatial and temporal regulation of the Hox genes occurs in parallel, which led to the hypothesis that Hox gene clustering is essential for coordinating the varied timing of gene activation. However, studies in mutant mice, in which the HoxD cluster was mutated, showed no systematic impact upon subsequent expression patterns, suggesting the spatial and temporal regulatory mechanisms could be uncoupled (Tschopp et al. 2009).

Chromatin plays an important role in the activation of the Hox genes (Figure 1.7 B). In the early mouse embryo the Hox genes are transcriptionally inactive and their chromatin is heavily marked with H3K27me3. Around day 8.5 the 3' end of the Hox cluster becomes active and the chromatin is marked with H3K4me3. As activation of the Hox genes progresses, more and more of the H3K27me3 marks are replaced with H3K4me3. By day 9.5 the whole cluster is transcriptionally active and covered with H3K4me3 (Montavon & Soshnikova 2014).



**Figure 1.7: Hox genes in fly, mouse and human.** (A) In mouse and human 39 Hox genes are found in four clusters of Hox A, B, C and D on different chromosomes. Marked in grey are the paralogue genes and their corresponding genes in *D. Melanogaster*. The order of the Hox genes on the chromosome corresponds to the order of the body parts that are controlled by them: genes at the 3' position of the chromosome are expressed in anterior body structures whereas the genes in 5' positions are expressed in posterior body structures. Furthermore, Hox genes are expressed in a progressive way, starting with the 3' genes expressed first advancing towards the 5', which are expressed last (Pang & Thompson 2011). (B) This expression is accompanied by changes in chromatin structure. In an inactive state, the whole cluster is labelled with H3K27me3, but upon Hox gene activation they are marked with H3K4me3 (Montavon & Soshnikova 2014).

## 1.8 MLL1 in the context of histone modifications

In general Hox genes are regulated by the Polycomb and Trithorax protein complexes, which are highly conserved and work as antagonistic partners in gene expression (Schuettengruber et al. 2007). Although their first and best characterized targets are the Hox genes, they also regulate many other genes. Polycomb and Trithorax proteins bind to many more loci making them epigenetic master regulators, which have an regulatory impact on cell-cycle control, tissue growth and cell proliferation (Zink & Paro 1989; Schumacher & Magnuson 1997; Mendenhall & Bernstein 2008; Christophersen & Helin 2010; Kolybaba & Classen 2014).

Polycomb and Trithorax proteins act as member of large protein complexes, which act primarily by depositing or reading, histone modifications like histone methylation, ubiquitylation and acetylation. Polycomb proteins repress, while Trithorax proteins activate transcription (Klymenko & Müller 2004). For example, tri-methylation of lysine 27 on histone 3 (H3K27me3) is deposited by PRC2 (polycomb repressive complex 2) and is associated with gene silencing (Schwartz et al. 2006), whereas tri-methylation of lysine 4 on histone 3 (H3K4me3) is deposited by SET proteins and activates gene expression, antagonizing Polycomb-induced silencing (Tie et al. 2014). In summary the activating and silencing of Trithorax and Polycomb proteins is tightly coupled and balanced (Steffen & Ringrose 2014).

Both Polycomb and Trithorax proteins deposit methyl-marks on chromatin, which requires close teamwork with demethylases in order to remove the opposing methyl-mark before deposition of a new mark. For example, whereas the UTX and the JMJD3 H3K27 demethylases are essential for proper development, UTX also associates with the H3K4me3 histone methyl-transferase MLL2 (Agger et al. 2007). Lid (Little imaginal discs), a jumonji C domain-containing protein with trimethyl H3K4 demethylating activity was shown to

reverse TRX-mediated methylation in *Drosophila* (Eissenberg et al. 2007). In humans JARID1d, another jumonji C-domain-containing protein was found to specifically demethylate trimethyl H3K4 and interacts with a Polycomb-like protein (Ring6a/MBLR) to bind to chromatin (Lee et al. 2007).

In mammals MLL1 is one member of the Trithorax protein family, which deposits the H3K4me3 mark. This makes H3K4me3 and H3K27me3 key histone modifications at MLL1 target sites (Canaani et al. 2004). *In vitro* studies with a recombinant MLL1 SET domain found that the methyl-transferase activity was stimulated by acetyl- and phospho-marks on the same histone (H3K9ac and H3S10p). The effect was biggest when both marks were on the histone at the same time, as a phosphoacetyl-mark (H3K9ac-S10p; Nightingale et al. 2007). In contrast, two different studies showed the phosphorylation of serine28 on histone 3 (H3S28p) antagonizes Polycomb protein 2 (PRC2)-mediated silencing and activates target genes. This suggests that histone phosphorylation may play a functional role in regulating TRX and PCG action. In both studies the phosphorylation mark was deposited by Msk1 (Gehani et al. 2010; Lau & Cheung 2011).

## 1.9 Msk1

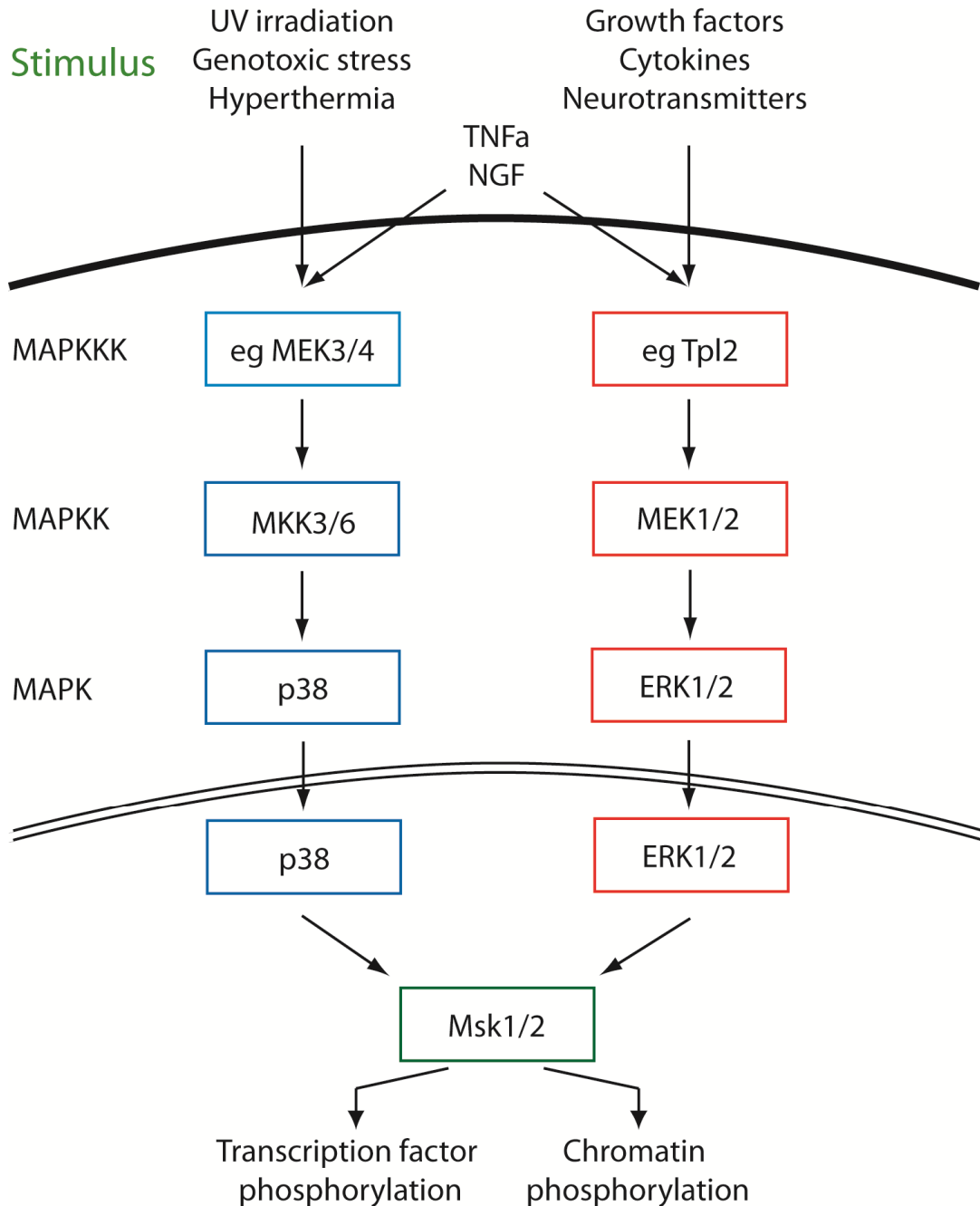
Msk1 (mitogen–and-stress-activated kinase 1) is a 90 kDA protein, which was described first in 1998 as a downstream target of the MAPK (mitogen-activated-protein-kinase) pathways (Deak et al. 1998). MAPK are serine-threonine kinases that can react to a range of different stimuli on the cell surface and act by consecutive cascade-phosphorylation of downstream targets in the cytoplasm and the nucleus (Edmunds & Mahadevan 2004). In mammals four MAPKs are known: ERK (extracellular regulated kinase) 1/2 and 5 and p38. These proteins are essential regulators for many regulatory processes, including those on the DNA (Klein et al. 2013). Msk1 is a downstream target of the ERK1/2 or the p38 pathway and is universally expressed in cells and found in mammals, nematode and fly, but not in yeast or plant (Hauge & Frödin 2006).

Thus far Msk1 and its homologue Msk2 have only one well-established function. As shown in Figure 1.8, they regulate gene expression by phosphorylation of either transcription factors or histones (Vermeulen et al. 2009). For example, Msk1 and Msk2 are required for the phosphorylation of the transcription factors CREB and ATF1. Without this phosphorylation CREB and ATF1 cannot properly activate their target genes (Wiggin et al. 2002). Another transcription factor targeted by Msk1 is NFκB, which is involved in the immune and inflammatory response. Msk1 phosphorylation of NFκB is essential for its transactivation capacity (Vermeulen et al. 2003). Msk1 can phosphorylate histone H3 at serine 28 or serine 10, however the two marks are not found on one histone tail together and Msk1 does not have an intrinsic specificity for one residue, suggesting that Msk1 kinase activity is context-dependent, potentially reflecting its genomic location (Arthur 2008). Thus far no consensus phosphorylation motif has been identified. It is likely that Msk1 cooperates with other

proteins, so that only a single site (i.e. S28 or S10) is available for phosphorylation. Alternatively the interaction with partner proteins may modulate Msk1's substrate specificity (Dyson et al. 2005).

Msk1 is transcribed at basal levels in the cell but kept in an inactivated, non-phosphorylated state. After the binding of either ERK1/2 or p38 at the MAPK docking site, the linker and the CTK (C-terminal kinase) domain is phosphorylated. The CTK domain auto-phosphorylates the rest of the linker domain and the NTK (N-terminal kinase), which leads to an allosteric change that activates the NTK domain (Msk1 domain structure and phosphorylation sites are shown in Figure 1.9 A). This conformational change leads to a rapid 10-50 fold increase of Msk1 activity in the cell (McCoy et al. 2005). However, how Msk1 is inactivated and whether this is coupled to the MAPK pathway is not clear. Thus far no phosphatase that might inactivate Msk1 has been identified (Duda, 2012).

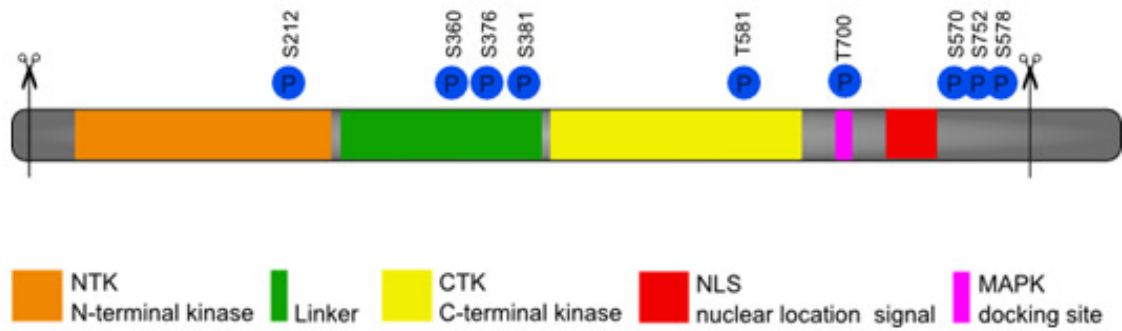
Besides Msk1 there is one homologue in humans, Msk2. Msk1 and Msk2 share about 76% protein identity. The core sequence of Msk1 and Msk2 are essentially the same. As shown in Figure 1.9 B Msk2 lacks the extreme N-terminus (amino acids 1-16) and the extreme C-terminus (amino acids 574-805) of Msk1 (Blast, Basic local alignment search tool). Functionally, only a few differences between Msk1 and Msk2 have been observed. It seems that in cells where both homologues are present, they show only minor differences in substrate choice, though Msk1 has a shorter activity period after activation (~30 min) than Msk2 (~3 hours) (Duda, 2012).



**Figure 1.8: Msk1 is a downstream target of the MAPK signalling pathway.** In humans four MAPK pathways are known, but only two activate Msk1 and its homologue Msk2. A variety of different stimuli at the cell surface activate the intra-cellular phosphorylation cascade, which leads to the activation of Msk1/2 in the nucleus.



A



B



**Figure 1.9: Msk1 domain structure and comparison to Msk2.** (A) Msk1 is constantly expressed in cells and kept in an un-phosphorylated, inactive state. Upstream kinases have to bind to the MAPK docking site and activate the C-terminal kinase, which then auto-phosphorylates Msk1 (Phosphorylation sites are indicated with blue circles and P). This activates the N-terminal kinase, which then auto-phosphorylates its tail domain for full activation and Msk1 targets (adapted from Duda, 2012). (B) Msk1 and Msk2 show 76% sequence identity (BLAST result, visualised with Jalview software: regions of high sequence identity are shown in red, those of low identity in orange). The core sequence of Msk1 and Msk2 are similar, but Msk2 is shorter than Msk1. The extreme N-terminus (AA 1-18) and the extreme C-terminus of Msk1 (AA 574-802) are not present in Msk2. The scissor-symbol in (A) indicates the missing termini, showing that the core-domain structure is conserved.

## 1.10 Aims

It is now well established that there is a close link between the chromatin state and the activity of genes. Histone modifications provide a layer of epigenetic information that can determine the functional state of chromatin. These are deposited by specific chromatin modifying enzymes, but also can modulate the activity or the binding of these enzymes. This results in a close interaction and inter-dependence between histone modifications and chromatin modifying enzymes, providing a network of regulatory information, which can integrate information from outside and inside the cell. MLL1 is a conserved histone methyltransferase, which deposits H3K4me3, a key regulatory mark in gene expression. This complex integrates histone methylation and acetylation, making it a powerful regulator of gene activity. Data from our laboratory has previously shown that MLL1's SET MTase activity is stimulated by histone 3 phospho-acetylation (H3K9ac/S10p; Nightingale et al. 2007). Acetyl-transferases are members in the MLL1 complex and acetylation is closely linked to active genes, however, to date no histone kinase has been associated with MLL1. In mammalian cells, two chromatin-associated histone kinases are known, Msk1 and Msk2, to play a role in PCG regulation.

We therefore hypothesized that Msk1 is involved in the regulation of the MLL1 complex. In this thesis I explore the nature of the MLL1 and Msk1 interaction and the functional implication of this. To investigate the nature of this interaction we examined the components of the MLL1 complex and characterize complex subunits. We also set out to explore the distribution of MLL1 and Msk1 in vivo and identify genes that were regulated by these proteins. Finally we examined the function of MLL1 and Msk1 through the cell cycle to gain a greater understanding of the regulation of MLL1 in a physiological context.

## 2 Materials and Methods

### 2.1 Tissue Culture

#### 2.1.1 Cultivation of mouse embryonic fibroblasts (MEFs)

Mouse embryonic fibroblasts (MEFs) are a widely used 'primary' cell-line (Smith 2001). Here MEFs were isolated from Balb/c wild-type mouse embryos between day 12 – 15 of gestation. Normally five E15 embryos were processed at a time. They were isolated, the heads and inner red organs removed and then the tissue dispersed in 10ml of 0.05% trypsin-EDTA (Gibco). The embryos were minced for ~2 min with a scalpel and then incubated for 20 min at 37°C. 10ml of culture medium [DMEM (Gibco), 1/100 penicillin (10000 U/ml, Gibco), 1/100 streptomycin (10000 U/ml, Gibco), 1/100 L-glutamine (200mM, Invitrogen), 1/100 MEM non-essential amino acids, Gibco), 1/1000 2-mercaptoethanol (Gibco), 10% v/v foetal calf serum (Invitrogen)] was then added and a single cell suspension was generated by pipetting vigorously for ~2 min. 8ml of the cell suspension was aliquoted into a 150cm<sup>2</sup> flask with 20ml MEF Growth Medium, which was passage 1. Cells were incubated at 37°C and 5% CO<sub>2</sub>, and depending on their density were divided 1:4 every 3<sup>rd</sup> day. Only cells between passage 2 and 4 were used.

### **2.1.2 Cultivation of LCL cells**

LCLs are a human lymphoblastoid cell line with a normal karyotype, which are routinely used for karyotyping (Terrenoire et al. 2010). The cells were grown in RPMI medium [RPMI 1640 (Gibco), 1/100 penicillin (10000 U/ml, Gibco), 1/100 streptomycin (10000 U/ml, Gibco), 1/100 L-glutamine (200mM, Invitrogen), 10% v/v foetal calf serum (Invitrogen)]. After three to four days, dependent on their density, the cells were divided 1:5 and incubated in fresh medium at 37° C and 5% CO<sub>2</sub>.

### **2.1.3 Cultivation of SL2 cells**

Drosophila SL2 (Schneider line-2) is a cell line, which was isolated from late embryonic stages of Drosophila (Schneider 1972). The cells were cultured in Schneider's medium [Insect Xpress (Lonza), 5% v/v foetal bovine serum (Invitrogen), 1/100 penicillin (10000 U/ml, Gibco), 1/100 streptomycin (10000 U/ml, Gibco)]. The cells were incubated at 26°C and divided every 5 days 1:10.

### **2.1.4 Cultivation of HEKs**

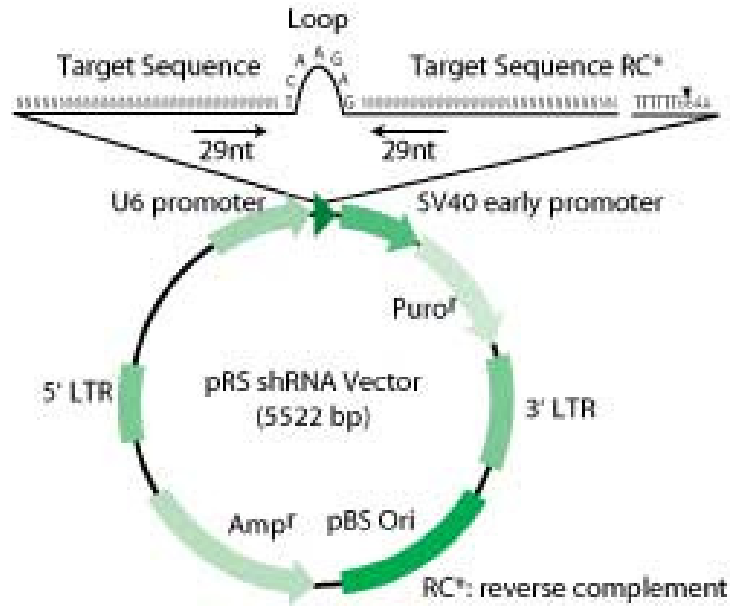
Human embryonic kidney cells (HEKs) were generated by the transformation of embryonic kidney tissue with the SV-40 virus (Sheint & Enders 1962) and later with the adenovirus type 12 (Zur Hausen 1967). The cells were cultured in DMEM [DMEM (Gibco), 1/100 penicillin (10000 U/ml, Gibco), 1/100 streptomycin (10000 U/ml, Gibco), 1/100 L-glutamine (200mM, Invitrogen), 10% v/v foetal calf serum (Invitrogen)]. Every 4 days the cells were divided 1:10 and incubated in fresh medium at 37°C at 5% CO<sub>2</sub>.

## 2.2 Transfection methods

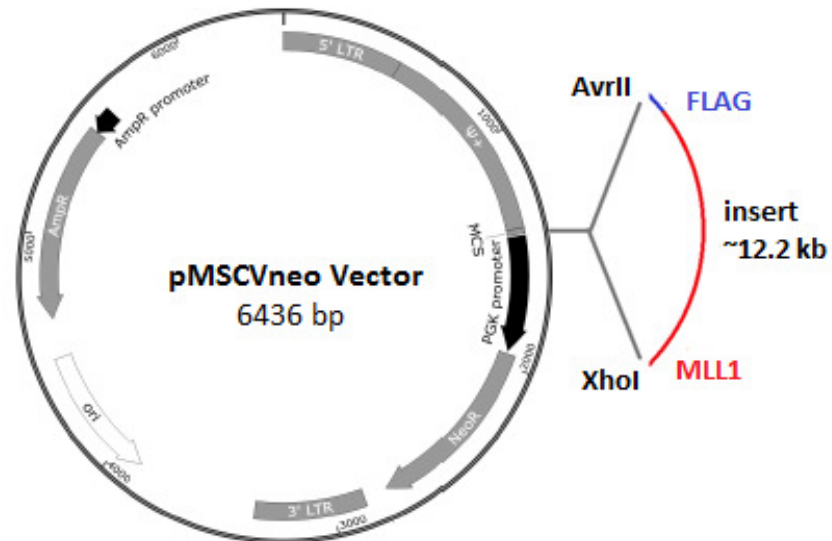
### 2.2.1 Amplification of transfection vectors

Three transfection vectors were used: Msk1 knock-down vector (Figure 2.1 A, Origene TR504795), MLL1 knock-down vector (Origene, TR517798) and a flag-tagged MLL1 (Figure 2.1 B, pMSCVneo, supplied from R. Slany, University of Erlangen). Vector DNA needed to be amplified and was therefore introduced into E.coli ( $\alpha$ -select gold efficiency, Bioline). Vector DNA (~20ng) was mixed with 50  $\mu$ l of E.coli and incubated on ice for 30 minutes prior to a 30 second heat-shock at 42°C. The bacteria were subsequently cultured in 1 ml LB [1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl] without antibiotics for 1 hour and then plated on a LB-plate [1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, 1.5% w/v Agar] with added ampicillin (final concentration 100  $\mu$ g/ml, Sigma) and incubated overnight at 37°C. The next day an isolated clone was picked and cultured in 5ml of LB with added ampicillin (final concentration 100 $\mu$ g/ml) and cultured at 37°C with shaking over the day. In the evening 500ml of LB medium with ampicillin (final concentration 100 $\mu$ g/ml) were inoculated with the day culture and incubated overnight at 37°C with shaking. The next morning the bacteria were collected by centrifugation (4000g, 10 minutes, 4°C) and the plasmid DNA purified using the GeneElute Plasmid MaxiPrep Kit (Sigma) according to the manufacturer's instructions.

A



B



**Figure 2.1: Vector maps.** The maps for the vectors used for transfection are shown. (A) The pRS vector with a shRNA target sequence for either MLL1 or Msk1 knock-down (Origene) is shown. (B) The pMSCVneo Vector, which carried the Flag-MLL cDNA is shown.

### **2.2.2 Transfection of MEFs**

MEFs were used to examine the effect of MLL1 and Msk1 knock-downs. For one transfection  $4 \times 10^6$  MEFs were harvested with trypsin and collected by centrifugation (300g, 5 minutes, RT), before washing them twice with PBS [phosphate-buffered saline: 1.37M NaCl, 27mM KCl, 100mM  $\text{Na}_2\text{HPO}_4$ , 18mM  $\text{KH}_2\text{PO}_4$ ]. The cells were resuspended in 400 $\mu$ l of transfection buffer [10mM HEPES, pH7.5, 135mM KCl, 2mM  $\text{MgCl}_2$ , 5mM EGTA, 25% FBS] and 20 $\mu$ g of knock-down vector DNA were added. The cell suspension was transferred into 0.4cm cuvettes (BioRAD, Cat# 1652088) and electroporated (300V, 950 $\mu$ F, 1000 $\Omega$ , BioRad GenePulser XCell) prior to being transferred into a 10cm dish with 10ml of MEF growth medium and incubated at 37°C and 5%  $\text{CO}_2$  for 5 hours. Puromycin (Gibco, final concentration 10  $\mu$ g/ml) was added to select for transfected cells and the cells were incubated again at 37°C and 5%  $\text{CO}_2$ . MEFs were harvested after a total of 20 hours post-transfection.

### **2.2.3 Transfection of HEKs**

HEKs were used for the expression of FLAG-MLL1 protein and therefore transfected with the pMSCV-FLAG-MLL1 vector (Figure2.1). The day before transfection the cells were divided and seeded at a density of  $4 \times 10^5$  cells per 10cm dish in complete medium [DMEM (Gibco), 1/100 penicillin (10000 U/ml, Gibco), 1/100 streptomycin (10000 U/ml, Gibco), 1/100 L-glutamine (200mM, Invitrogen), 10% v/v foetal calf serum (Invitrogen)]. The next day, a mix of 15 $\mu$ g vector DNA, 1000 $\mu$ l OptiMEM media (Gibco) and 60 $\mu$ l of Polyethyleneimine (1mg/ml stock solution, Polyscience Inc.) per plate was made up, vortexed thoroughly, incubated at room temperature for 30 minutes and added drop-wise to the plates. The cells were cultured at 37°C at 5%  $\text{CO}_2$  for 3 days and then harvested.

## **2.3 Protein Methods**

### **2.3.1 Preparation of whole cell extract**

Cells were collected and washed three times with ice-cold PBS before the cell pellet was resuspended in 500µl lysis buffer [10mM Tris, pH 7.9, 150mM NaCl, 5mM EDTA, 0.5mM EGTA, 1mM β-mercaptoethanol, 1% v/v NP40, 25% v/v glycerol]. The cell suspension was left on a rotating wheel at 4°C for 30 minutes, before sonicating at high amplitude for one minute (Bioruptor, Diagenode). The suspension was then centrifuged at 13000 rpm for 10 minutes at 4°C, and the supernatant stored at -20°C.

### **2.3.2 Isolation of histones**

Cells were collected and washed twice with cold PBS, to which 10mM sodium butyrate was added. Cells were resuspended in Triton extraction buffer [TEB: 1xPBS, 4mM sodium butyrate, 0.5% v/v Triton X-100, 2mM PMSF, 0.02% v/v sodium azide] to a final concentration of  $10^7$  cells per ml and left on ice for 10 minutes, before centrifuging for 10 minutes at 2000 rpm at 4°C. The pellet was resuspended in half the volume of TEB and centrifuged again at 2000 rpm for 10 minutes at 4°C. The pellet was then resuspended in 0.4 M HCl at a concentration of  $4 \times 10^7$  cells per ml and left on ice for 3 hours at 4°C. After centrifugation to pellet the nuclear debris, the supernatant with the solubilized histones was removed and stored at -20°C.



### **2.3.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)**

Proteins were separated according to size following the method developed by Laemmli (Laemmli, 1970). Samples were loaded on an acrylamide gel, with a concentration dependent on the size of the proteins to be separated. Generally 50µg of whole-cell extract or 10µg of purified protein was loaded. For the detection of Msk1 (~90kDa MW) and Actin (~50kDa MW) a 10% SDS gel [10ml Acrylamide (30% Acrylamide - 0.8% Bis), 7.5ml 1.5mM Tris, pH 8.8, 300µl 10% w/v SDS, 100µl 10% w/v ammonium persulfate, 30µl TEMED, made up to a total volume of 30ml with distilled water] was prepared. For MLL1 (~300 kDa MW) a 4% gel was prepared. The resolving gel was overlaid with 100% isopropanol to prevent evaporation during polymerisation. This was washed away before the addition of the stacking gel [1ml Bis-acrylamide (30% Acrylamide-1.6% Bis), 6.3ml 0.5mM Tris, pH 6.8, 100µl 10% w/v SDS, 100µl 10% w/v ammonium persulfate, 10µl TEMED, made up to a volume of 10ml with distilled water]. Protein samples were resuspended in 10xloading buffer [20% SDS, 500mM Tris, pH 7.6, 50% v/v glycerol, 1M DTT, 1mg/ml bromophenol blue], incubated at 95°C for 10 minutes and loaded onto the gel through SDS reservoir buffer [50mM Tris, 0.384M glycine, 0.1% v/v SDS]. The gel was electrophoresed at constant 60mA, for 3 hours in a standard vertical electrophoresis unit (Hoefer "SE 600 Ruby", GE Healthcare)

### **2.2.4 Staining protein gels**

To visualise proteins in SDS gels they were stained either with Coomassie blue or with silver-nitrate. For Coomassie blue staining the gel was soaked in Coomassie-solution [1.25% w/v Coomassie Brilliant Blue R250 (Sigma), 10% v/v isopropanol, 30% v/v methanol, 10% v/v

acetic acid] for 30 minutes. The background was removed by soaking the gel in Destain-Solution [30% v/v methanol, 10% v/v acetic acid] until only the protein bands were visible. For silver staining the gel was soaked overnight in 50% v/v methanol to fix the proteins. For the staining solution, Solution A [0.8g Silver nitrate, 2ml distilled water] and Solution B [10.5ml 0.36% w/v NaOH, 800µl 14.2M NH<sub>4</sub>OH] were prepared independently, slowly mixed together with constant stirring and finally made up to 50ml with distilled water. The gel was soaked for 15 minutes in the Staining solution and then washed with distilled water for 15 minutes. The gel was then soaked in the Developing solution [2.5ml 1% w/v citric acid, 50µl 38% formaldehyde, made up to 500ml with distilled water] until the protein bands were visible. The reaction was stopped by adding Stopping solution [45% v/v Methanol, 10% v/v acetic acid].

### **2.3.5 Immuno-detection of proteins on nitrocellulose membranes**

Western blotting was performed according to a standard protocol (Towbin et al., 1979). Briefly, the separated proteins were transferred onto a Hybond-C nitrocellulose membrane (Amersham) in a Trans-Blot Cell (BioRAD) for three hours (180V, 300mA, 20W) using transfer buffer [25mM Tris, 192mM glycine, 20% v/v methanol]. The efficiency of the transfer was checked by Ponceau Red (Sigma) staining. The membrane was then incubated for one hour at 4°C in blocking buffer [10% w/v dried milk powder, 1xTBS (50mM Tris, pH 7.5, 150mM NaCl)]. Primary antibodies were either generated in house or obtained from commercial sources (listed below). The membrane was then incubated with the primary antibody diluted in blocking buffer at room temperature at the recommended dilution. The membrane was typically washed three times, each for 10 minutes, with TBS-T [1xTBS (50mM Tris, pH 7.5,

150mM NaCl), 0.1% v/v Tween-20]. The secondary antibody was then added in blocking buffer and incubated for one hour at room temperature. The membrane was washed three times, each for 10 minutes, with TBS-T. Binding was visualised by using the “Odyssey” (LiCor) imaging system, which activates the fluorochrome coupled to the secondary antibody (LiCor).

<b>Antibody</b>	<b>Source</b>	<b>Dilution/Animal</b>
Msk1-p	Abcam ab81294	1:500 / rabbit
Msk1-p	Abcam ab32190	1:500 / rabbit
Msk1	Bethyl A302-747A	1:500 / rabbit
MLL <sup>C</sup>	Millipore 05-765	1:500 / mouse
MLL <sup>N</sup>	Millipore 05-764	1:500 / mouse
Msk2	R&D systems MAB2310	1:100 / rat
Actin	Abcam ab1801	1:1000 / rabbit
Pol II	Abcam ab5408	1:1000 / mouse
NFκB	Abcam ab7970	1:1000 / rabbit
Histone H3K4me3	In house R612	1:500 / rabbit
Histone H3K9ac	In house R607	1:500 / rabbit
Histone H3K27me3	Millipore 07-449	1:1000 / rabbit
Histone H3S10p	Abcam ab32107	1:2500 / rabbit
Histone H3K9acS10p	Abcam ab12181	1:1000 / rabbit
Histone H3, C-terminal	Abcam ab1791	1:2500 / rabbit
Licor 2 <sup>nd</sup> , α-rabbit IgG	Rockland 611-131-121	1:5000 / goat
Licor 2 <sup>nd</sup> , α-mouse IgG	Rockland 610-131-121	1:5000 / goat
AlexaFluor647 2 <sup>nd</sup> , α-rat IgG	Life technologies A-21247	1:500/ goat

**Table 2.1: Antibodies used for Western analysis, including commercial source and used dilutions**

## **2.4 Functional analysis**

### **2.4.1 RNA extraction**

Total RNA was isolated from cells using QIAGEN Kits (RNase free DNase Set and RNeasy Minikit) according to the manufacturer's instructions. RNA was quantified using a Implen Nanophotometer and stored at -80°C.

### **2.4.2 Quantitative PCR (qPCR)**

For CChIP analysis, qPCRs was performed with the primers in Table 2.2. "SensiMix SYBR Hi-Rox" (Bioline) was used. For one reaction 0.25µg of the forward and reverse primers were added to 15ng of DNA and 20µl of master mix. PCR samples were pipetted in a 384-well optical reaction plate (Applied Biosystems) and loaded into a 7900 HT machine (Applied Biosystems). The programme was set up as following: samples were heated at 95°C (15 minutes), then at 95°C (30 seconds), at 58 °C (30 seconds) and at 72°C (30 seconds). The last three steps were performed 40 times. In order to evaluate primer specificity, a dissociation curve was created after each run. qPCR results were analysed by calculating the relative expression ( $2^{-\Delta Ct}$ ) of the bound fraction compared to the unbound fraction of the same CChIP.

Gene Name	Forward	Reverse	TM
HoxA4-1	gccgttggttctatcctgctc	tgatgcctcactcgtacctg	58°C
HoxA4-2	agaggcctaggacagacgtg	tggatgctgctagccttcag	58°C
HoxA4-3	tgttggaaggaagccagact	aaaatcccccaaactgctct	58°C
HoxA5-1	atcggctctggctactgaaa	agtcgctccaagctgtaaa	58°C
HoxA5-2	agccggggaaataaagtgtg	ggggtcgaattgaggttaca	58°C
HoxA5-3	gcaagctgctctttctgctt	cttctggcctgaggtttctg	58°C

**Table 2.2: qPCR primer sets for CChIP analysis**

### **2.4.3 Reverse Transcription quantitative PCR (RT-qPCR)**

This method allowed simultaneous RNA reverse transcription and cDNA amplification in one step and was performed according to the manufacturer's instructions. It was used to analyse gene expression changes in knock-down cells. The reaction was achieved with commercial primers (see table 2.3) and the "Verso 1-step qPCR Syber green mix" (Thermo Scientific). For one reaction 0.1µl of enzyme-mix, 5µl of Syber buffer, 0.5µl of RT-enhancer, 1µl of oligos and 25ng RNA template were mixed and water added to make up to 10µl. Reactions were pipetted in a 384-well optical reaction plate (Applied Biosystems) and loaded into a 7900 HT machine (Applied Biosystems). The programme was set up as following: samples were heated at 50°C (15 minutes), then at 95°C (15 minutes), at 95°C (15 seconds), at 58°C (30 seconds) and at 72°C (30 seconds). The last three steps were performed 40 times. RT-qPCR results were analysed by normalizing the expression of genes in knock-down cells to the expression of the same genes in control cells ( $\Delta^{Ct}$ ). The relative expression of genes was calculated by comparing them to the reference gene expression of Actin ( $2^{-\Delta\Delta Ct}$ ).

Gene	Qiagen primer	Gene	Qiagen primer
$\beta$ -Actin	QT01136772	HoxA7	QT00168707
Msk1 (Rps6ka5)	QT00141554	HoxA9	QT00108885
Msk2 (Rps6ka4)	QT00116760	HoxA10	QT00240212
Meis1	QT00172557	HoxA11	QT00250404
HoxA1	QT00248322	HoxA13	QT0027642
HoxA2	QT0062812	MLL1	QT00240954
HoxA3	QT00138264	MLL2	QT01075620
HoxA4	QT00174986	MLL3	QT00310527
HoxA5	QT00098819	MLL5	QT00279426
HoxA6	QT00140742		

**Table 2.3: Commercial primers used for RT-qPCR**

#### **2.4.4 Formaldehyde Cross-linking Chromatin Immunoprecipitation (XChIP)**

XChIP was used to define the distribution of MLL1 and Msk1 binding in the genome. For one IP  $5 \times 10^6$  MEFs were trypsinized (0.05% Trypsin-EDTA, Gibco) and brought into a single-cell suspension. The cells were resuspended into MEF growth medium at a final concentration of  $10^6$  cells per ml. The proteins were cross-linked to the DNA by adding formaldehyde drop wise to a final concentration of 1% v/v, and incubated at room temperature with gentle agitation for 10 minutes. The reaction was stopped by adding 1/10 volume of 2M glycine in PBS [137mM NaCl, 2.7mM KCl, 10mM  $\text{Na}_2\text{HPO}_4$ , 1.8mM  $\text{KH}_2\text{PO}_4$ ] and the cells subsequently pelleted by centrifugation for 5 minutes at 300g at 4°C and washed twice with cold PBS. The cells were then resuspended in Lysis buffer [1% w/v SDS, 10mM EDTA, 50mM Tris pH 8, Protease Inhibitor cocktail (Roche)] to a final concentration of  $10^7$  cells per ml and sonicated to fragment the chromatin (high power, 30 seconds ON/OFF, 7 cycles, Bioruptor Diagenode).

A 100µl input sample was taken to assess the quality of the chromatin. 20µl of Proteinase K (20mg/ml) was added to the chromatin and crosslinking reversed by incubation at 65°C for 4 hours. The DNA was loaded on a 1% agarose gel to verify the size distribution (ideally ~200-300 bp). The remaining chromatin was diluted 1:2 in Dilution buffer [1% v/v Triton X-100, 2mM EDTA, 150mM NaCl, 20mM Tris pH 8, protein inhibitor cocktail (Roche)] and IgG coated magnetic beads (Thermo Scientific) prepared. For a single immuno-precipitation 15µl of beads were washed with citrate-phosphate buffer [CP buffer: 0.47% w/v citric acid, 0.92% w/v Na<sub>2</sub>HPO<sub>4</sub>] and finally resuspended in 20µl CP buffer with 0.5% w/v BSA. Then 4µg of antibody against the desired target protein was added to the beads and rotated for 2 hours at 4°C. The beads were then washed with CP buffer and resuspended in 15µl of CP buffer with 0.5% w/v BSA, before adding the beads to the quality controlled chromatin. The chromatin-beads suspension was rotated for 4 hours at 4°C. The beads were then washed twice with low salt buffer [1% v/v Triton X-100, 0.1% w/v SDS, 2mM EDTA, 150mM NaCl, 20mM Tris pH 8] and twice with high salt buffer [1% v/v Triton X-100, 0.1% w/v SDS, 2mM EDTA, 500mM NaCl, 20mM Tris pH 8]. The beads were then incubated for 15 minutes at room temperature in 50µl Elution buffer [100mM NaHCO<sub>3</sub>, 1% w/v SDS]. The supernatant was kept and the elution repeated. The second supernatant was pooled with the first and cross-linking reversed by adding 20µl proteinase K (20 mg/ml) and incubation at 65°C overnight. DNA was extracted using a QIAquick PCR purification Kit (QIAGEN) and tested for specificity by PCR [95°C (15 minutes), then at 95°C (30 seconds), at 58 °C (30 seconds) and at 72°C (30 seconds) for a total of 15 cycles]. DNA was sent to GATC Biotech (Konstanz, Germany) for library preparation and single-end sequencing. For analysis, SeqMonk software (Babraham Bioinformatics) was used.

Antibody	Source	Animal
Msk1	Abcam ab81294	rabbit
MLL <sup>N</sup>	Millipore 05-764	mouse
Histone H3	Abcam ab1791	rabbit

**Table 2.5: Antibodies used for XChIP**

#### **2.4.5 Carrier Chromatin Immunoprecipitation (CChIP)**

Carrier ChIP is a technology used to characterise DNA binding in small cell numbers (O'Neill et al. 2006). For one IP  $2 \times 10^7$  SL2 carrier cells were needed, as well as  $\sim 5 \times 10^5$  target cells. The cells were harvested and washed twice in cold NB buffer [15mM Tris pH 7.4, 5mM  $MgCl_2$ , 15mM NaCl, 60mM KCl, 0.1mM EGTA, 0.5mM  $\beta$ -mercaptoethanol, 5mM Na-butyrate, 0.1mM PMSF] and brought to a concentration of  $2 \times 10^7$  cells per ml in NB buffer. An equal volume of 1% Tween20 in NB buffer was added and  $1/200^{\text{th}}$  of 100mM PMSF. The cells were stirred on ice for 15 minutes and then homogenised in a glass homogeniser (tight fit) with 1 stroke per ml of cell suspension, before centrifuging at 2000 rpm for 20 minutes at 4°C. The nuclear pellet was then resuspended in 5% sucrose in NB buffer in half of the previous volume and centrifuged at 3000rpm for 10 minutes at 4°C. The nuclei were then resuspended in the same volume of Digestion buffer [0.32M sucrose, 50mM Tris pH 7.4, 4mM  $MgCl_2$ , 1mM  $CaCl_2$ , 5mM Na-butyrate, 0.1mM PMSF] and the chromatin content measured using an Implen Nanophotometer. The nuclei were centrifuged at 2000rpm for 10 minutes at 4°C and then brought to a final concentration of 0.5mg/ml with Digestion buffer, to which micrococcal nuclease (final concentration 50U/ml) was added. This was incubated at 37°C for five minutes, before 10 $\mu$ l 0.5M EDTA was added to terminate the digestion and



the chromatin placed on ice. It was subsequently centrifuged at 2000rpm for 5 minutes at 4°C and the supernatant retained (S1) and stored at 4°C. The pellet was resuspended in 1ml of Lysis buffer [1mM Tris pH 7.4, 0.2mM EDTA, 5mM Na-butyrate, 0.2mM PMSF] and dialysed (10kDa cut-off dialysis tubing, Sigma) overnight against 2 L Lysis buffer at 4°C overnight. The next day the chromatin was centrifuged at 1000g for 10 minutes at 4°C, the supernatant (S2) and pellet (resuspended in 200µl Lysis buffer, P) were retained. Chromatin integrity was checked on a 1% agarose-gel. Generally S1 and S2 were used and pooled, 10x Incubation buffer [50mM NaCl, 20mM Tris pH 7.4, 20mM Na-butyrate, 5mM EDTA, 0.1mM PMSF] was added and made up to 1ml with distilled water. This was subsequently pre-cleared with 100µl Protein A Sepharose (GE Healthcare) for 1 hour at 4°C. The chromatin was centrifuged for 10 minutes at 13000 rpm at 4°C and the pre-cleared supernatant taken and the antibody of interest added before incubation overnight at 4°C with slow rotation. The volumes of antibodies used were:

H3K4me3 (R612, in house)	50µl per 100µg of chromatin
H3K9acS10p (ab12181, Abcam)	10µl per 100µg of chromatin
H3K27me3 (07-449, Millipore)	10µl per 100µg of chromatin

The next day 100µl of pre-swollen protein A sepharose beads were added and incubated for 3 hours at room temperature with rotation. The beads were then centrifuged (13000rpm, 10 min, 4°C) and the supernatant removed and retained (labelled 'UNBOUND'). The beads were then washed with Buffer A [50mM Tris pH7.4, 10mM EDTA, 5mM Na-butyrate, 50mM NaCl, 0.1mM PMSF], Buffer B [50mM Tris pH7.4, 10mM EDTA, 5mM Na-butyrate, 100mM NaCl, 0.1mM PMSF] and Buffer C [50mM Tris pH7.4, 10mM EDTA, 5mM Na-butyrate, 150mM NaCl, 0.1mM PMSF], centrifuging the beads for 5 minutes at 2000rpm at 4°C. The pellet was

then resuspended in 1% w/v SDS in Incubation buffer and incubated at room temperature for 15 minutes with rotation, centrifuged (13000rpm, 5 minutes) and the supernatant kept (labelled 'BOUND1'). The elution was repeated and the supernatant kept (labelled 'BOUND2'). Samples BOUND1 and BOUND2 were pooled and the DNA extracted using a QIAquick PCR purification Kit (QIAGEN).

#### **2.4.6 FLAG purification**

We use HEK293 cells to overexpress FLAG tagged MLL1 protein and subsequently purify the MLL1 complex. HEK293 cells were transfected with an expression vector coding for FLAG-tagged MLL protein (Figure 2.1 B: pMSCVneo, R. Slany, University Erlangen). After 3 days of transfection the cells were harvested and washed twice with TBS [50mM Tris, 150mM NaCl, pH 7.4]. The cells were then resuspended in Lysis buffer [1xTBS, 1 mM EDTA, 1% v/v Triton X-100], for one 10cm dish 500µl Lysis buffer was used. Cells were stirred on ice for 30 minutes before they were centrifuged for 15 minutes at 13000 rpm and 4°C. 50µl of anti-FLAG M2 magnetic beads (Sigma) were added and incubated for 2 hours at room temperature with rotation. The beads were then washed with TBS. Washing was monitored by measuring the absorbance of the supernatant at 280nm and continued until the absorbance of the wash solution was  $OD_{280} < 0.05$ . FLAG-MLL protein was then eluted from the beads using 5µg of 3x FLAG peptide (Sigma), diluted in 1ml TBS. 500µl of Elution Solution was added to the beads and incubated at 4°C for 30 minutes with rotation. The supernatant was collected and another 500µl of Elution Solution added and incubated at 4°C for 30 minutes with rotation. The supernatants were pooled and the protein concentrated using a Centrifugal Filter Unit (Millipore, Cut-off 30kDa), at 6000rpm for 10 minutes at 4°C.

#### **2.4.7 Size exclusion chromatography**

Protein purification by size exclusion chromatography was performed using the Pharmacia FPLC system with a BioRad BioSil250 SEC column (7.8x300mm). Prior to loading protein samples, the column was equilibrated with TBS buffer [50mM Tris, 150mM NaCl, pH 7.4], and then ~200µg of protein in 50µl TBS loaded. During the gel filtration run defined fractions were collected and elution profiles were generated by measuring the OD<sub>280</sub> of the flow-through. The elution profile allowed the identification of peak fractions, which were pooled and analysed by SDS-PAGE.

#### **2.4.8 Elutriation of LCLs**

Elutriation is a centrifugal separation technique, which separates cell according to size. This allows the separation of cells into different stages of the cell cycle (Kauffman et al. 1990). For elutriation the JE-5.0 Elutriation System (Beckman Coulter) was used with an Avanti J-26S XP Centrifuge (Beckman Coulter). Elutriation was performed at room-temperature and all used buffers were freshly prepared and used at room-temperature. Prior to use the chamber was assembled, fitted into the centrifuge rotor, connected to the pump and washed with 70% ethanol. The system was sub-sequently washed with 0.75xPBS and finally with EB buffer [0.75xPBS, 1% w/v BSA, 5mM EDTA] and any remaining air-bubbles removed. Whilst equilibrating the system the LCLs were prepared. The cells were colcemid treated for three hours before the elutriation run to arrest cells in the cell cycle. Colcemid (Biochrom), a microtubule-depolymerizing agent (Yang et al. 2010), was added to the medium to a final concentration of 0.1 µg/ml and incubated for 3 hours. ~400x10<sup>6</sup> cells were used per run, washed with EB buffer and finally resuspended in 10ml EB buffer, before passing them

through a cell filter (50µm cut-off) to remove any clumps. The centrifuge was set to 1800 rpm and the pump to a flow rate of 110 for loading the cells into the system. Fractions of 200ml were collected before increasing the flow rate. Fractions were collected at a flow rate of 11.0, 12.5, 14.5, 16.5 and 18.5 ml/min. The collected cells were centrifuged (1200 rpm, 5 min at 4°C) and washed twice with cold 1xPBS.  $1 \times 10^6$  cells were taken for flow cytometer analysis and the remaining cells were either used immediately for further analysis or stored as a pellet at -20°C.

#### **2.4.9 Flow Cytometer Analysis (FACS)**

FACS analysis was used to define the ploidy and the cell cycle phase of elutriated cell populations. About  $1 \times 10^6$  cells were collected by centrifugation (1200 rpm, 5 minutes) per analysis. The supernatant was discarded, the cell pellet resuspended in 1ml of 80% ethanol/PBS and stored overnight at -20°C. After incubation the cells were centrifuged (800rpm, 10 minutes) and washed with cold PBS. The cells were centrifuged again at 800rpm for 10 minutes and the pellet resuspended in 500µl PBS with propidium iodide (PI, final concentration 0.1 mg/ml). The samples were incubated for 30-60 minutes at room temperature and then measured (Cyan ADP, Beckman Coulter). PI is excited at a wavelength of 488nm and emits light at 590nm, therefore the PE channel was used to measure PI incorporation into the DNA.

#### **2.4.10 Immuno-fluorescence Microscopy**

To characterise protein distribution on chromatin, microscopy of immuno-fluorescence labelled G1 or G2/M cells was used. For microscopy, the cells were washed twice with PBS and counted. The cells were diluted to a concentration of  $1 \times 10^5$  cells/ml in KCM [120mM KCl, 20mM NaCl, 10mM Tris/HCl pH 8, 0.5mM EDTA, 0.1% Triton X-100]. Ethanol-washed slides were fixed into chambers for the Cytospin (Thermo Scientific, Shandon Cytospin 4). 200 $\mu$ l of cells were put in the loading funnels and spun for 10 minutes at 1000rpm at room temperature. The slides with the cells were incubated in a bath of KCM for 10 minutes, before incubating with the primary antibody. Antibodies were diluted to the recommended concentration with KCM+0.1% w/v BSA. The slides were stored in a humidified chamber at 4°C for one hour. After incubation the slides were washed twice in a bath of KCM for 10 minutes, before adding the secondary antibody, which was labelled with a fluorochrome. The slides were incubated at 4°C for an hour again, prior to being washed twice with KCM for 10 minutes, and fixed by incubation in a bath of 4% formaldehyde/KCM for 10 minutes. Afterwards the slides were washed with dH<sub>2</sub>O, mounted with 10 $\mu$ l DAPI in Vectorshield solution (1 $\mu$ g/ml, Vector Laboratories) and then sealed with coverslips.

Antibody	Source	Dilution/Animal
Msk1	Abcam ab32190	1:1000 / rabbit
MLL <sup>N</sup>	Millipore 05-764	1:1000 / mouse
Histone H3K4me3	In house R612	1:500 / rabbit
Histone H3K9ac	In house R607	1:500 / rabbit
Histone H3S10p	Abcam ab32107	1:1000 / rabbit
FITC 2 <sup>nd</sup> , $\alpha$ -rabbit IgG	Abcam ab6717	1:1000 / goat
TRITC 2 <sup>nd</sup> , $\alpha$ -mouse IgG	Dako R0270	1:1000/ goat

**Table 2.4: Antibodies used for Immunofluorescence microscopy**

#### 2.4.11 Co-immunoprecipitation

Co-immunoprecipitation was used to verify the interaction of MLL1 complex subunits. It was performed in LCLs, HEKs and MEFs. For one IP  $1 \times 10^8$  cells were collected and washed twice in ice-cold PBS. The final cell pellet was resuspended in 500 $\mu$ l NP40 buffer [1% v/v NP-40, 10% v/v glycerol, 50mM Tris pH 7.5, 0.1% w/v sodium azide, 150mM NaCl]. The mixture was vortexed and incubated on ice for 20 minutes for cell lysis. Cells were then sonicated (high power, 10 seconds, Bioruptor Diagenode), the resultant lysates centrifuged for 20 minutes at 13000 rpm, and the supernatant kept. Cell lysates were pre-cleared by adding Protein A or G sepharose beads, which were soaked overnight in NP40 buffer. The mixture was incubated on a rotating wheel for 30 minutes at 4°C before centrifuging at 13000 rpm for one minute. The cleared supernatant was incubated overnight with 15 $\mu$ l of antibody at 4°C on a rotating wheel. Then 40 $\mu$ l of sepharose beads were added and incubated for 3 hours on a rotating wheel at room-temperature. After centrifugation (3000 rpm, 1 minute), the supernatant was kept and the beads washed three times with high salt NP40 buffer [1% v/v NP-40, 10% v/v glycerol, 50mM Tris pH 7.5, 0.1% w/v sodium azide, 200mM NaCl]. The

beads and the supernatant samples were mixed with SDS-loading buffer [20% w/v SDS, 500mM Tris, pH 7.6, 50% glycerol, 1M DTT, Bromophenol blue] and analysed by SDS gel and Western analysis.

<b>Antibody</b>	<b>Source</b>	<b>Animal/Sepharose</b>
Msk1	Abcam ab81294	rabbit/sepharose A
MLL <sup>N</sup>	Millipore 05-764	mouse/sepharose G
MLL <sup>C</sup>	Millipore 05-765	mouse/sepharose G
NFκB	Abcam ab7970	rabbit/sepharose A

**Table 2.5: Antibodies used for Co-IP**

#### **2.4.12 Gene Expression Arrays**

Expression analysis was performed on MEFs using the NimbleGen system according to the manufacturer's "NimbleGen Arrays User's Guide". The array-slide contained 12 individual arrays with 42,576 genes in triplicate on each. Therefore four control samples, four MLL1 knock-down and four Msk1 knock-down samples were loaded. Firstly, 10µg of RNA per array were isolated (RNeasy, Qiagen) and reverse transcribed into cDNA using the cDNA synthesis System (NimbleGen). cDNA samples were subsequently labelled with the NimbleGen One-Color DNA labelling kit. Then the samples were hybridized on the array (100718\_MM9\_EXP\_HX12) and washed (NimbleGen Hybridization, Sample Tracking Control Kit & NimbleGen Wash Buffer). Finally the array was scanned with the MS 200 Microarray Scanner (NimbleGen). Data were processed with R (background correction and normalisation), and analysed with MeV (clustering, statistical analysis and visualisation). They are freely available, specialized tools to analyse microarray data, which incorporate a number of different algorithms.

# 3 Results

## 3.1 Characterizing the interaction of MLL1 and Msk1

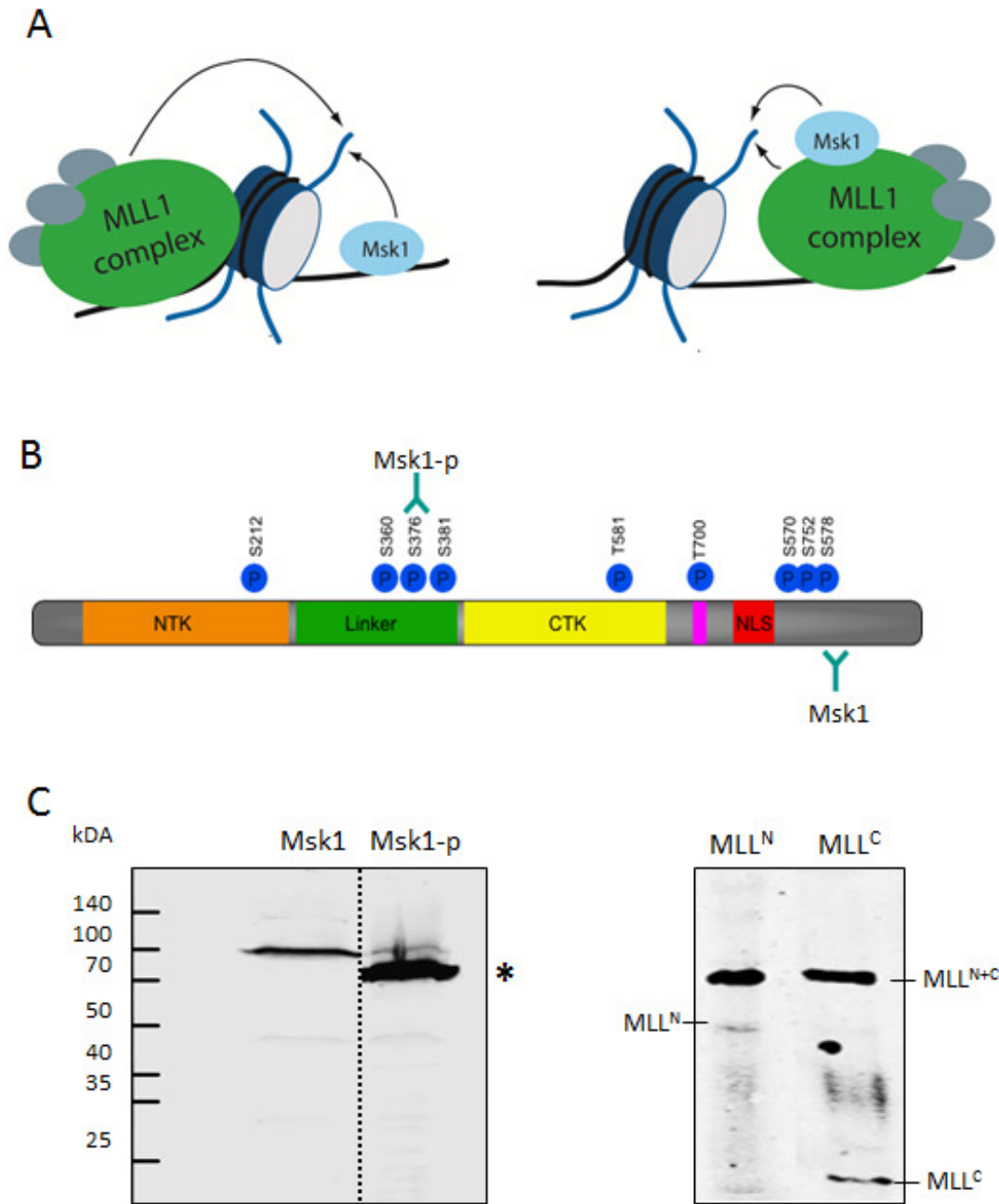
MLL1 is a histone methyltransferase, which deposits methyl marks (me1-me3) on lysine 4 of histone H3 (H3K4me3). The methylation of histone 3 is involved in gene activation and is functionally linked with the acetylation of histone 3 and 4. We have previously shown that the processivity of the HMTase activity of the MLL SET domain is stimulated *in vitro* by acetylation together with phosphorylation at H3S10 (Nightingale et al. 2007). H3 acetylation is an abundant mark, whereas H3 phosphorylation only appears on a sub-fraction of nucleosomes in interphase cells and is deposited by ERK and p38 MAP kinases, which act through their downstream kinases Msk1 and 2 (Dyson et al. 2005; Soloaga et al. 2003). We therefore hypothesized that Msk1 deposits the phosphorylation mark at H3S10 to stimulate MLL1 activity at specific loci. Here I wanted to explore if MLL1 and Msk1 interact *in vivo* and the nature of the interaction.

### 3.1.1 Co-Immunoprecipitation experiments with MLL1 and Msk1

If we assume Msk1 and MLL1 do functionally interact, two scenarios are possible. Firstly, MLL1 and Msk1 may act on the same sites on chromatin, but bind independently and do not directly interact (Figure 3.1 A, left). Alternatively, Msk1 may be a member (permanent or transient) of the MLL1 complex, similar to the histone acetyltransferases MOF or CBP (Figure 3.1 A, right). We therefore performed co-immunoprecipitation experiments to see if in a MLL1 pull-down, Msk1 can be detected and vice versa. Msk1 is constantly expressed in cells,



but in an inactive, un-phosphorylated state, and only after phosphorylation is it functionally active (Duda 2012). As Msk1 phosphorylation directly affects the activation status of Msk1, we used two different Msk1 antibodies (Figure 3.1 B): (1) A Msk1 antibody (Msk1-p), which was directed against the phosphorylated serine 376 residue (for LCLs and HEKs) or serine 360 (for MEFs) and would bind active Msk1, due to species reactivity. (2) An antibody (Msk1) raised against the non-phosphorylated region between residues 752 and 802. This region is phosphorylated during the last stages of Msk1 activation and therefore the un-phosphorylated region represents inactive Msk1 (Arthur 2008). All antibodies used for the co-immunoprecipitation experiments [un-phosphorylated Msk1 (Msk1), phosphorylated Msk1 (Msk1-p) and MLL1 (N- or C-terminus)] were tested for specificity by Western analysis, which also confirmed MLL1 and Msk1 expression in LCLs (Figure 3.1 C). To test the Msk1 antibodies the same cell lysate was loaded on two adjacent lanes, which were separated (dashed line) and one lane was probed with the anti-Msk1 antibody and the other with the anti-Msk1-p antibody. The two membrane pieces then were re-joined and scanned together in order to highlight binding differences. The Msk1 antibody only detected one band (~90 kDa), indicating it was specific for non-phosphorylated Msk1. The Msk1-p antibody recognizes both the un-phosphorylated Msk1 (weak band on top), but mainly a double band, which is shifted downwards. This is consistent with Msk1 having multiple phosphorylation sites, which impacts on the electrophoretic mobility of the protein. These findings indicated that both Msk1 antibodies were specific for their targets (Figure 3.1 C, left). In contrast, the MLL1 antibodies specificity for the individual MLL termini were low (MLL<sup>N</sup> is ~320kDa and MLL<sup>C</sup> is ~180 kDa) and only the intact complex (full length MLL is ~430 kDa) was recognized reliably with both antibodies (Figure 3.1 C, right).

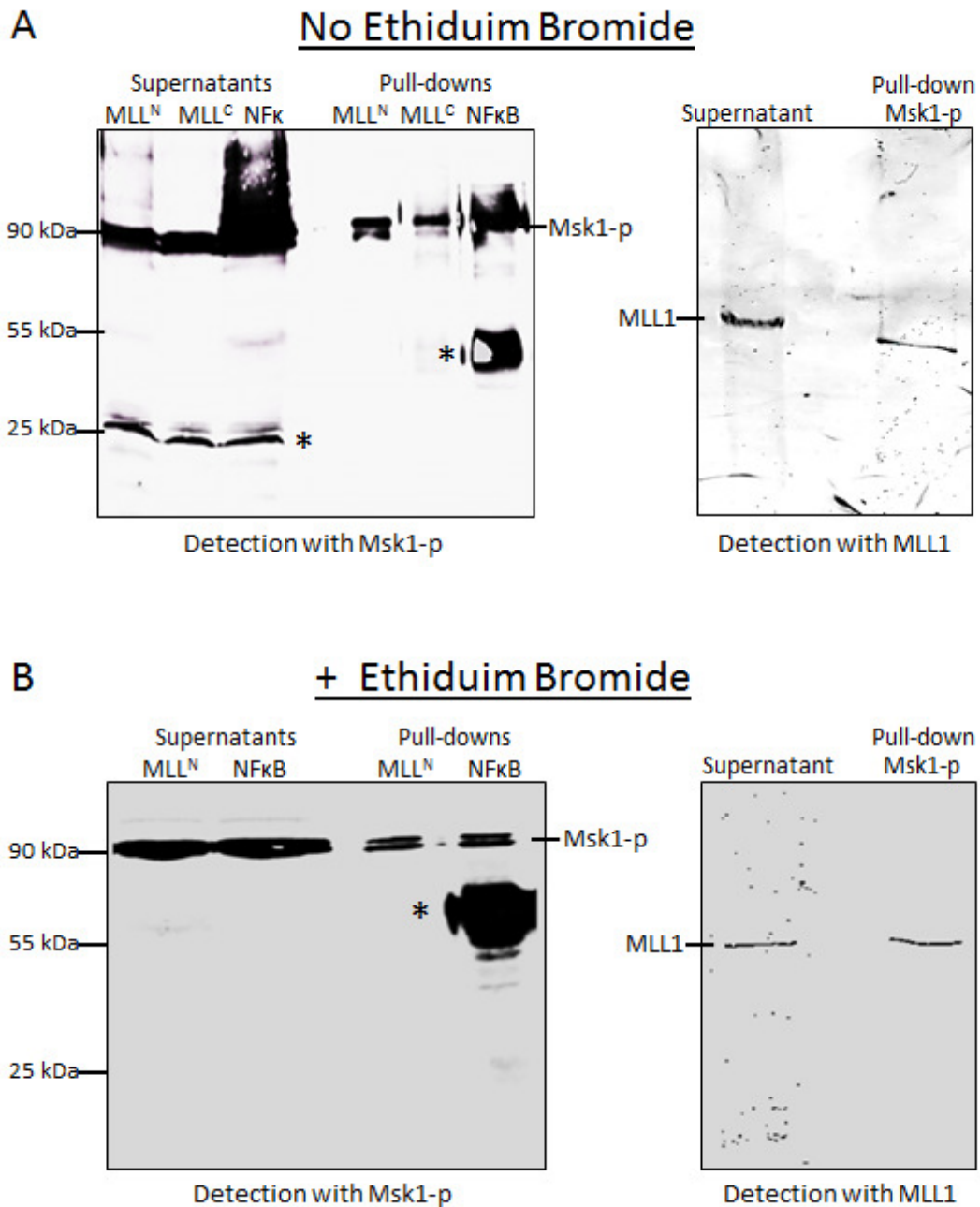


**Figure 3.1: Representation of the Co-IP experiments.** (A) MLL1 and Msk1 have two options to interact with each other: they might act independently at the same site on chromatin (left) or they might form a complex (right). The Co-IP experiments should address which interaction is more likely. (B) The domain structure of Msk1 is shown, as well as known phosphorylation sites. The epitopes for the antibodies against Msk1 used for the Co-IP are indicated. (C) Anti-Msk1 (left panel) and the anti-MLL1 antibodies (right panel) used for the experiments were tested by western blotting on whole cell extract for specificity and expression of MLL1 and Msk1 in LCLs. Note the altered mobility of phosphorylated Msk1 (\*).

First the interaction between the phosphorylated Msk1 and MLL1 was examined (Figure 3.2 A, left). A cell lysate was prepared and a pull-down with the MLL<sup>N</sup>, the MLL<sup>C</sup> and NFκB antibodies was performed. NFκB was used as a positive control as it is known that it binds strongly to Msk1 (Vermeulen et al. 2003). In the pull-down with the MLL antibodies a strong clear band of the right size (~90 kDa) could be detected, as well as in the control pull-down with NFκB. In the control pull-down a prominent band of smaller size (~55kDa, asterisk), which appears as a smear, could also be detected. Furthermore, in the input samples small bands of ~25 kDa were found (marked with an asterisk). Later experiments showed that these were the heavy or light chain of the used antibodies (data not shown). As the MLL<sup>N</sup> and the MLL<sup>C</sup> antibodies showed no differences, for further experiments only the MLL<sup>N</sup> antibody was used.

To validate the binding, the co-IP was reversed (Figure 3.2 A, right). A pull-down with Msk1-p was performed and the western probed for MLL1<sup>N</sup>. The western showed a clear MLL1 band. This finding supports the hypothesis that MLL1 and Msk1 interact.

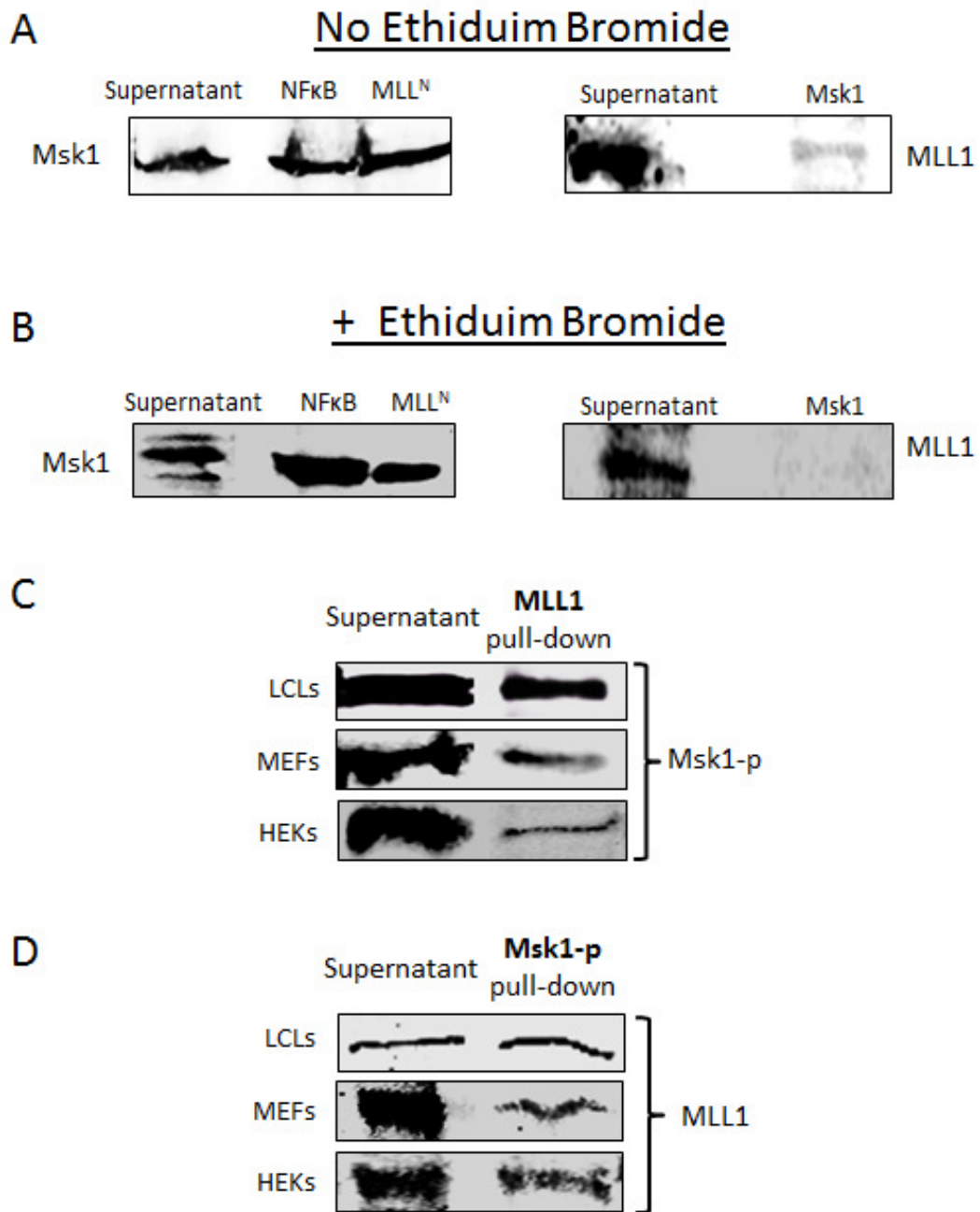
However, this experimental set up, did not rule out that the interaction between MLL1 and Msk1 was bridged by chromatin. In order to release the proteins from the DNA a high concentration of ethidium bromide (50µg/ml) was added (Lai & Herr, 1992; Schröter et al., 1985). Under these conditions (Figure 3.2 B), phosphorylated Msk1 was still detected in the MLL pull-down (A) and MLL1 could be still observed in the Msk-p pull-down (B).



**Figure 3.2: Co-Immunoprecipitation experiments for MLL1 and Msk1.** (A) Co-IP experiments were established in LCLs. Left: A cell extract was prepared and MLL<sup>N</sup>, MLL<sup>C</sup> and NFκB were pulled down with antibodies. The beads and their corresponding Input samples were loaded on a 12% gel. It is known that Msk1 and NFκB bind, so this IP was used as a positive control. Right: A Co-IP for Msk-p was performed and the beads and the supernatant sample were loaded on a 4% gel. (B) Co-IP including ethidium bromide. During the preparation of the cell extract 50 μg/ml ethidium bromide was added to release chromatin proteins from the DNA. Asterisks indicate the IgG heavy (55 kDa) or light (25 kDa) chain of the antibodies.

In the next experiment, the interaction between the non-phosphorylated Msk1 (inactive) and MLL1 was examined and therefore the immunoprecipitation repeated with the antibody that recognized non-phosphorylated Msk1 (Figure 3.1 B). In the pull-down with MLL<sup>N</sup> a strong band for Msk1 was detected (Figure 3.3 A, left), which also showed up in the positive NFκB pull-down. In the reverse experiment however, only a very weak Msk1 pull-down could be detected (Figure 3.3 A, right). The experiment was repeated with ethidium bromide added to the lysates (Figure 3.3 B). This confirmed that Msk1 could be detected in the MLL1 pull-down, but no MLL1 could be detected in the pull-down with non-phosphorylated Msk1. The interpretation of these findings are not clear, as it could suggest that MLL1 interacts with Msk1 in its active, phosphorylated form, but as not in the inactive non-phosphorylated form. However it is also possible that the antibody disrupts the MLL1-Msk1 interaction.

In order to validate and expand these Co-IP results, IPs with MLL<sup>N</sup> and Msk1-p were repeated in different human and murine cell lines, with ethidium bromide (50μg/ml) added to the cell lysates. As shown in Figure 3.3 C and D, experiments using human HEKs and mouse embryonic fibroblasts confirmed the binding of MLL1 and Msk1-p.

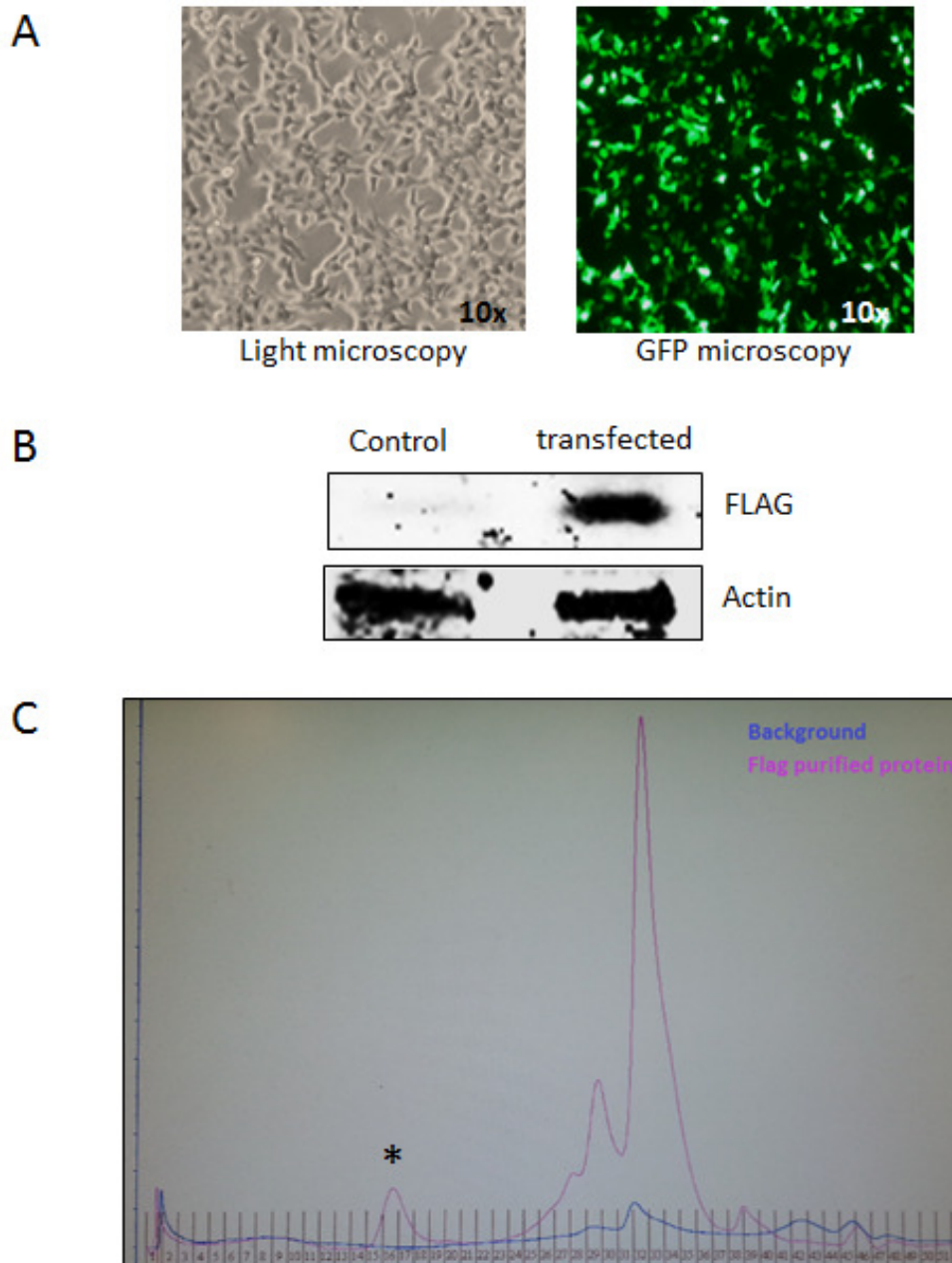


**Figure 3.3: Co-immunoprecipitation for MLL1 and Msk1.** (A) The Co-IP experiment was repeated using an antibody that recognizes un-phosphorylated Msk1. In the MLL1 pull-down Msk1 could be detected, but only a weak MLL1 band in the Msk1 pull down. (B) The experiment was repeated with 50µg/ml ethidium bromide in the cell extract. Msk1 could be detected in the MLL1 pull -down, but in the Msk1 pull-down no MLL1 could be detected. (C) The Co-IP experiments were repeated in MEFs and HEKs. Ethidium bromide was added to all cells extracts. In all tested cell lines Msk1-p could be detected in the MLL1 pull-down. (D) In the reverse experiments MLL1 could be detected in the Msk1-p pull-downs.

### 3.1.2 FLAG-tagged MLL1 complex

In a second approach to characterize the putative MLL1 and Msk1 interaction, we over-expressed FLAG-tagged MLL1 (pMSCVneo vector was a gift of Slany, University of Erlangen) in HEK cells and purified the associated protein complex. HEK cells were transfected with the FLAG-MLL1 encoding vector or with a GFP containing vector to control transfection efficiency (Figure 3.4 A) and transfected cells were shown to express FLAG MLL1 by western analysis (Figure 3.4 B).

In the next step the transfected cells were harvested, lysed and the MLL1-Flag protein purified using  $\alpha$ -FLAG IgG coated beads. A concentrated sample was run on a size exclusion gel filtration column to analyse the purity of the immuno-precipitated complex protein (Figure 3.4 C). As a control a background sample was run (supernatant of the concentration process). A single peak was detected in the elution profile of the FLAG purified protein sample (fractions 15-16, 3.75-4.5ml, marked with an asterisk), which couldn't be detected in the background profile. However, in the fractions 26-36 (6.5-9 ml) multiple peaks could be detected in both elution profiles, which are likely to reflect non-specific binding of proteins to the beads (MW ~30 kDa).

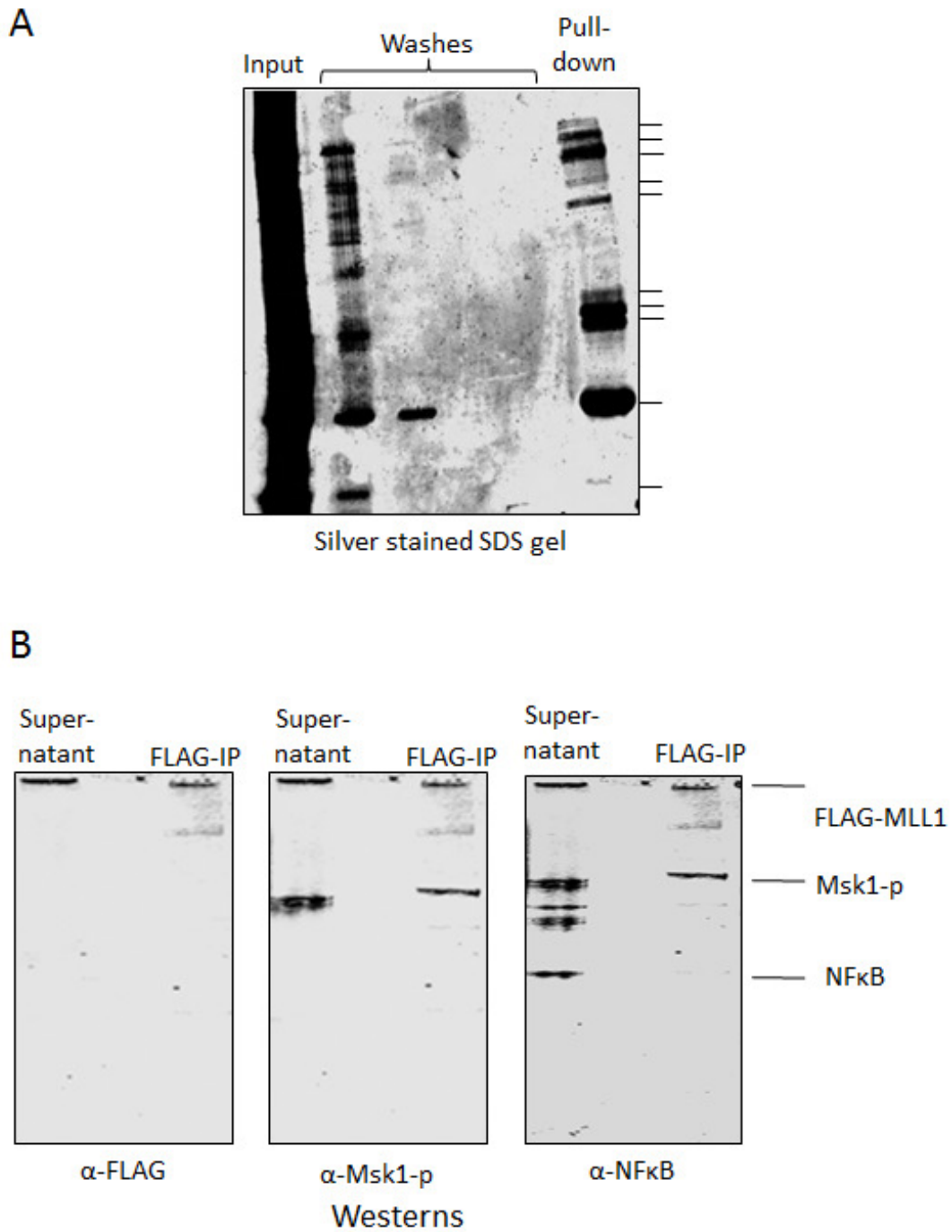


**Figure 3.4: Overexpression of FLAG-MLL1 in HEK293s.** (A) HEK cells were serum starved overnight, before incubation for 7 hours with transfection medium. After 3 days the cells were collected. Cells transfected with a GFP expressing vector were controlled for transfection efficiency with fluorescence-microscopy. (B) The transfected cells were checked for FLAG expression. A single band was detected, which was absent in non-transfected cells (Control). Actin was used as a loading control. (C) A concentrated sample and background sample (supernatant of the concentrated protein) was run on a size-exclusion gel-filtration column to control the quality of the purified protein. A unique peak in the FLAG-protein elution profile (fraction 15-18, asterisk), is likely to be a large molecular weight complex. In fraction 26-36 multiple peaks can be found in both elution profiles, which are likely to be non-specific protein binding to the beads



The purified protein complex was loaded on a gel and stained to visualize protein bands (Figure 3.5 A). A number of bands (~10x) could be detected, which we assumed were the MLL1 complex components. The protein was then analysed by western (Figure 3.5 B). The membrane was firstly probed with an antibody against the FLAG-tag and visualized a band at the top of the membrane, in both the Input and the purified protein sample (left panel). Secondly an antibody against Msk1 (phosphorylated) was used and visualized a band in the upper third of the membrane, in both (Input and purified protein) sample (middle panel). Lastly the membrane was probed with an NFκB antibody, as this is a known interaction partner of Msk1. This detected a band in the Input sample but not in the purified protein complex. The smear between the Msk1 band and the NFκB band (marked with \*) was cross-reaction of the NFκB antibody (right panel).

Taken together these results strongly support, that MLL1 does interact with Msk1 directly. The interaction is unlikely to be bridged by chromatin and may be limited to the active, phosphorylated form of Msk1. This suggested that active Msk1 is probably a member of the MLL1 complex. However, it is not yet clear if the MLL1-Msk1 interaction is a transient interaction or whether Msk1 is a permanent member of the MLL1 complex.



**Figure 3.5: MLL1-FLAG complex.** (A) The MLL-complex was isolated by FLAG-bead purification. On a SDS gel stained with silver-nitrate, a number of bands can be detected (indicated at the side with bars) (B) Western blotting, FLAG-tagged MLL (left), an Msk1 can be detected in the FLAG-bead pull-down (middle). In contrast, NFκB (a known interaction partner of Msk1) cannot be detected in the purified complex.

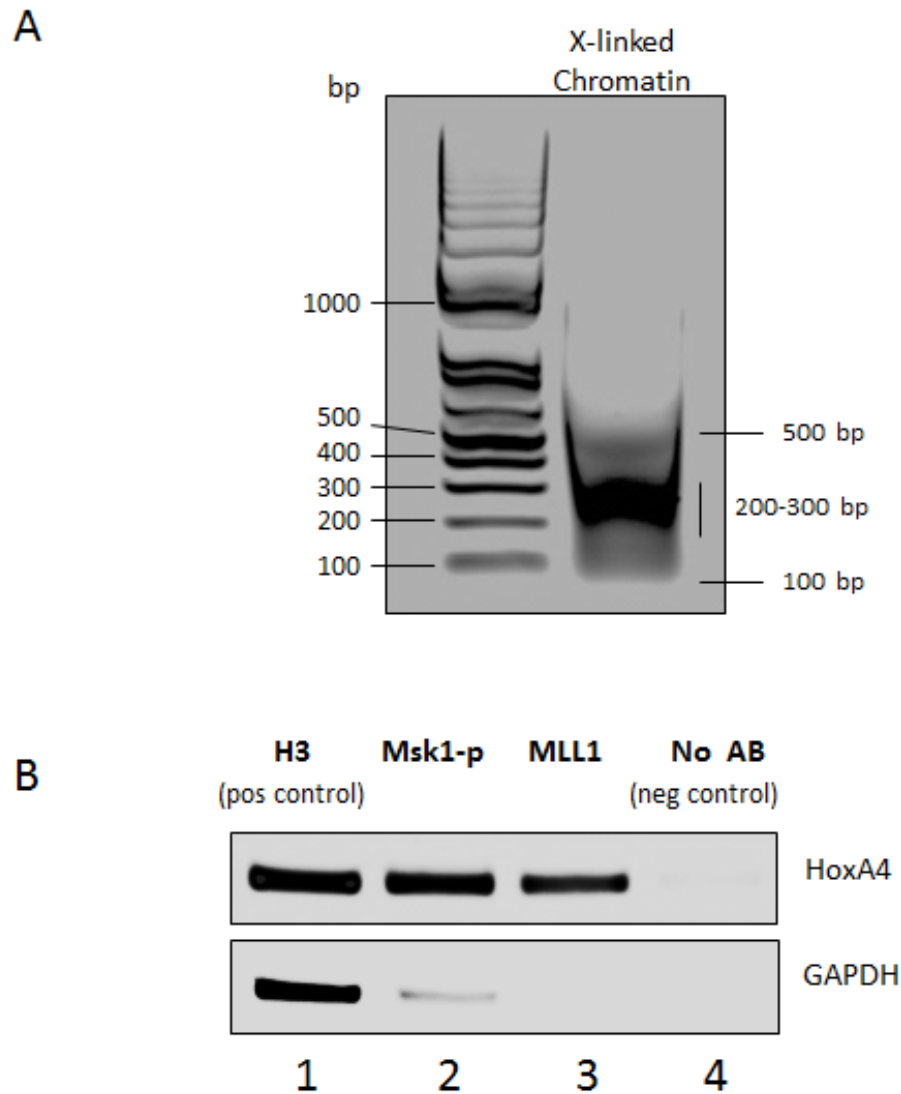
## **3.2 MLL1 and Msk1 binding in vivo**

In the previous experiments, we used immunoprecipitation and purification of FLAG-tagged MLL1 complex to investigate MLL1 and Msk1 interaction. These co-immunoprecipitation experiments suggested that MLL1 interacts strongly with Msk1. Moreover, analysis of a FLAG-purified MLL1 complex showed that Msk1-p can be found in the purified MLL1 complex. Following these experiments we wanted to examine the distribution of MLL1 and Msk1 binding in vivo and therefore used a Chromatin-immunoprecipitation followed by sequencing (ChIP-Seq) approach in mouse embryonic fibroblasts (MEFs).

### **3.2.1 Establishing an XChIP protocol for MEFs**

We chose MEFs because these are primary non-transformed cells with a normal karyotype, and are therefore a good model to study MLL1 and Msk1 binding under normal physiological conditions. Chromatin-immunoprecipitation for DNA-bound transcription factors requires cross-linking to the DNA. The cells were therefore treated with formaldehyde and the extracted chromatin was fragmented with sonication. This step was crucial, as sequencing requires chromatin fragments of ~250 bp. Sonication optimization (7 cycles, high intensity) resulted in a narrow spread of DNA fragments ranging from 100-500bp, with the majority in the 200-300bp range (Figure 3.6 A). Every batch of chromatin was validated in this way before use, as appropriate chromatin fragment sizes are essential for immunoprecipitation and subsequent library preparation.

For the immunoprecipitation the same antibodies were used as for the co-IPs (Chapter 3.1), one directed against the N-terminal domain of MLL1 and one directed against the phosphorylated serine 360 of Msk1 (activated Msk1, Figure 1.9 A). Additionally an antibody against the histone H3 C-terminal tail was used as a positive control. This domain is not known to contain modifications, so binding is assumed to be unaltered by histone modification. To control for non-specific binding, a 'mock' immunoprecipitation was performed, in which no antibodies were used. The precipitated chromatin was analysed in a short PCR (15 cycles) for amplification of two genes: HoxA4, a known MLL1 target gene and GAPDH as a reference gene (Figure 3.6 B). The pull-down with the H3-tail antibody (positive control, lane 1) showed amplification for both genes, indicating that the precipitation was efficient. Furthermore, the no-antibody pull-down (negative control, lane 4) showed no amplification of the two genes, indicating that non-specific binding was minimal. In contrast, chromatin pulled down with the MLL1 antibody amplified HoxA4, but not GAPDH. Likewise chromatin pulled down with the Msk1 antibody, showed strong amplification of HoxA4, but only a weak one for GAPDH. This suggested both precipitations were successful and pulled down known target genes. Immunoprecipitated DNA was subsequently sent to GATC Biotech for library preparation and 50bp single-end sequencing on the Illumina HiSeq2000. A coverage of 113,852,000 sequenced reads in the Msk1 library and 82,796,700 sequenced reads in the MLL1 library was achieved.

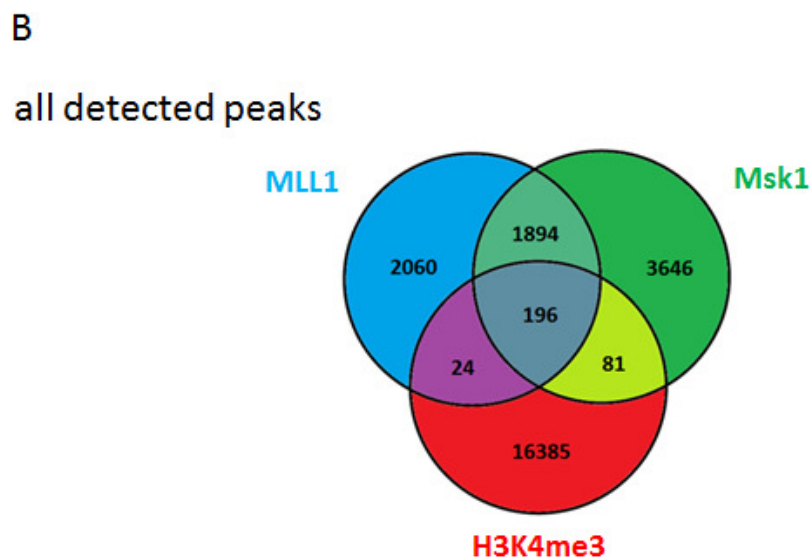
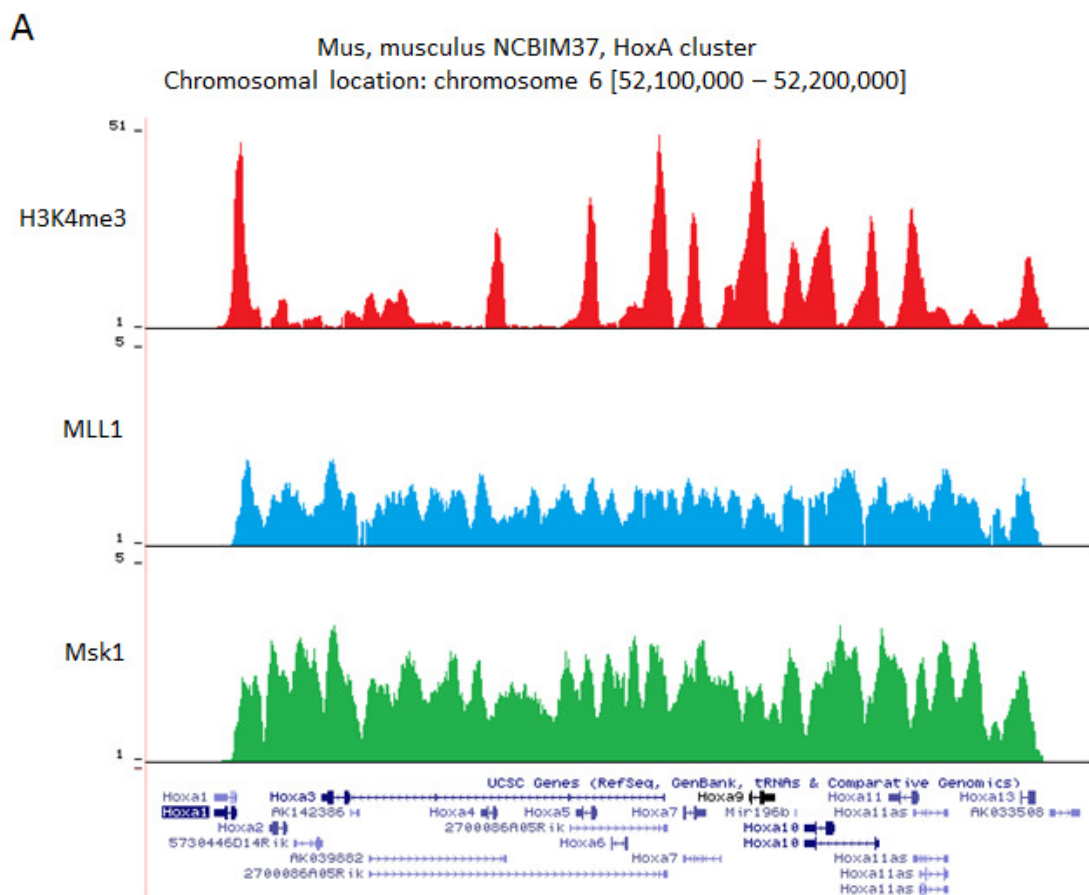


**Figure 3.6: Chromatin preparation for Next Generation Sequencing.** (A) Chromatin was fixed with formaldehyde and then sonicated. For sequencing small chromatin fragments ~250 bp were needed. 7 cycles (30 sec on/off) at high intensity resulted in chromatin between 100 and 500 bp, but with the majority of fragments between 200-300 bp. (B) In order to control the quality of the pull down, a positive control (antibody that recognizes the histone H3 C-terminal tail) and a negative control (no antibody control) were performed. A PCR was performed on the precipitated DNA to validate the antibody binding efficiency. HoxA4 was chosen, as it is a known MLL1 target gene. No binding was observed in the negative control (lane4), binding of Msk1 and MLL1 at HoxA4 was observed (lane2 & 3 upper panel), whilst no or little binding was detected in the GAPDH reference gene (lane2 & 3, lower panel). The primers were placed over the TSS of the two genes.

### 3.2.2 Analysis of the sequencing results of MLL1 and Msk1

For the analysis the Galaxy/Cistrom server was used (<https://usegalaxy.org>). In Figure 3.7 A, a screen-shot of the raw data is shown (wiggle plot). The ChIP-Seq data of MLL1, Msk1 and H3K4me3 for primary MEFs (Karolchik et al., 2014; Rosenbloom et al., 2013; Data generated by B. Ren, Ludwig Institute for Cancer Research) are shown over the *Mus musculus* HoxA cluster [Chromosome 6, 52,100,000 – 52,200,000], including known RNA products from these loci and major transcription start sites (TSS). The peak profile of H3K4me3 shows high enrichment and well defined peaks. The profiles of MLL1 and Msk1 show less enrichment and a more ‘spiky’ appearance, but appear to bind preferably to TSS where H3K4me3 can be detected. The first step was to identify regions, which were significantly enriched for reads. Therefore a peak detection algorithm (‘peak caller’) was run. In Galaxy it is based on the MACS peak caller for peak detection.

In order to compare regions bound by MLL1 or Msk1 and the H3K4me3 distribution, a Venn diagram was created (Galaxy tool, Figure 3.7 B). All binding sites of MLL1, Msk1 and H3K4me3 identified with the ‘peak caller’ were compared (Figure 3.7 B). The analysis showed overlap between MLL1 and Msk1 binding sites (2090 binding sites), however, the majority of sites were occupied by either MLL1 only (2060 sites) or Msk1 only (3646 sites). These data, showing a 30-40% overlap, between the sites, represent compelling data that the two proteins can act in tandem to regulate genes. The data also showed that the general MLL1/Msk1 binding does not correlate with H3K4me3, which was more abundant.

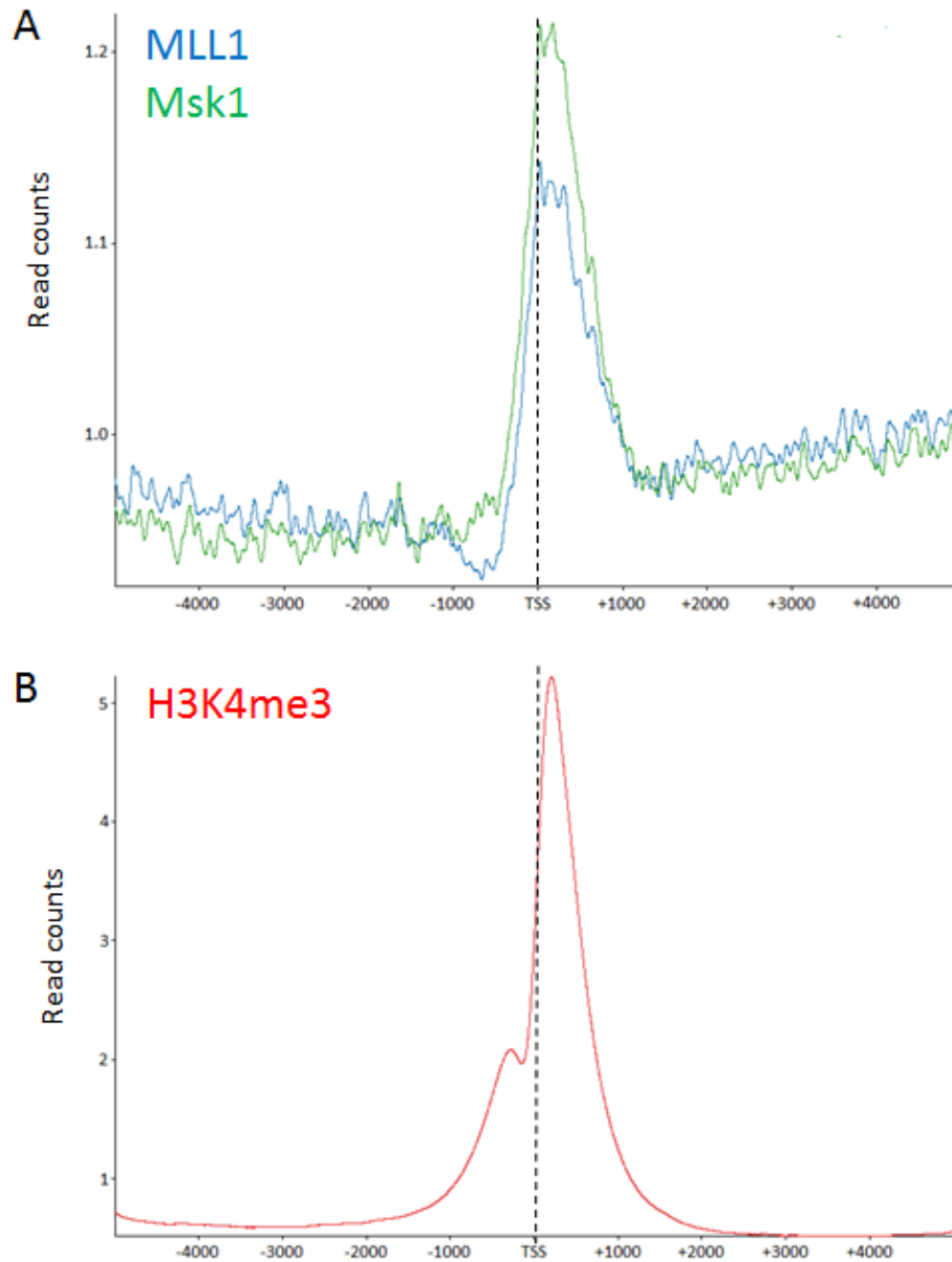


**Figure 3.7: Sequencing Analysis – peak identification.** (A) Screen Shot raw peak data over the HoxA cluster is shown. In order to identify high confidence binding of MLL1, Msk1 and H3K4me3 (Rosenbloom et al. 2013; Karolchik et al. 2014) to the genome, the MACS14 ‘peak caller’ algorithm was run. (B) Venn-diagram, comparing the binding sites of MLL1, Msk1 and H3K4me3, identified by the ‘peak caller’. MLL1 and Msk1 show an overlap of a subgroup of peaks (~20%). However, the majority of MLL1 and Msk1 do not correlate nor does it correlate with H3K4me3.

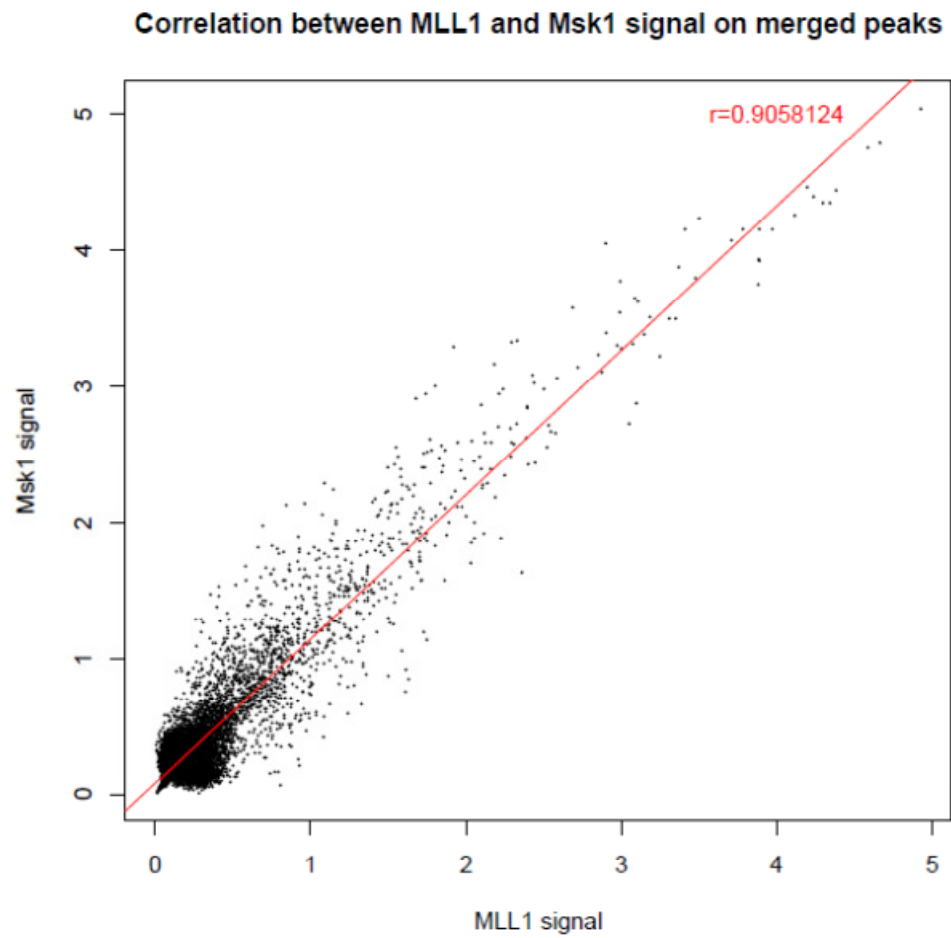
In the next step a probe trend plot was created (Figure 3.8). Probe trend plots show an average pattern or distribution of reads across known TSS sites (Galaxy/Cistrone CEAS tool). Here the average read pattern over each TSS (-5000 and +5000bp to each side) bound by MLL1 and Msk1 were analysed (Figure 3.8 A). The trend plot showed that MLL1 and Msk1 bind preferentially on the TSS, but that they also can be found over the adjacent 1000bp, with a broad peak of binding from  $\sim$ -500 to +1000. The two profiles are extremely similar and show profiles typical for TSS binding factors. In comparison the trend plot for H3K4me3 is shown (Figure 3.8 B), with a typical methylation peak over the TSS (compare Figure 1.3). However, while MLL1/Msk1 seem to bind directly over the TSS (indicated by black line), this position seems to be non-methylated, probably due to nucleosome redistribution by bound transcription factors at the TSS. H3K4me3 is mainly found  $\sim$ -500bp (small peak) and  $\sim$ +500bp from the TSS (large peak).

In the next step of the analysis the two peak data set (MLL1 binding vs. Msk1 binding) were compared, with the focus on the peaks that were bound by both MLL1 and Msk1. In Figure 3.9 the scatter plot is shown. The R-value of the correlation is 0.91, indicating that the binding patterns of MLL1 and Msk1 on the shared peaks were similarly strong.





**Figure 3.8: Sequencing Analysis – TSS probe analysis.** (A) The probe trend plot shows the average distribution of reads across TSS loci. It covers the region -5000 bp and +5000 bp around the TSS and shows an enrichment of MLL1 and Msk1 binding over the TSS. The black line represents the TSS. (B) Here the same analysis was done for H3K4me3. H3K4me3 is enriched ~-500 and +500 bp from the TSS, but not on the TSS itself. Data from Encode, generated by B. Ren, Ludwig Institute for Cancer Research (Rosenbloom et al. 2013; Karolchik et al. 2014)



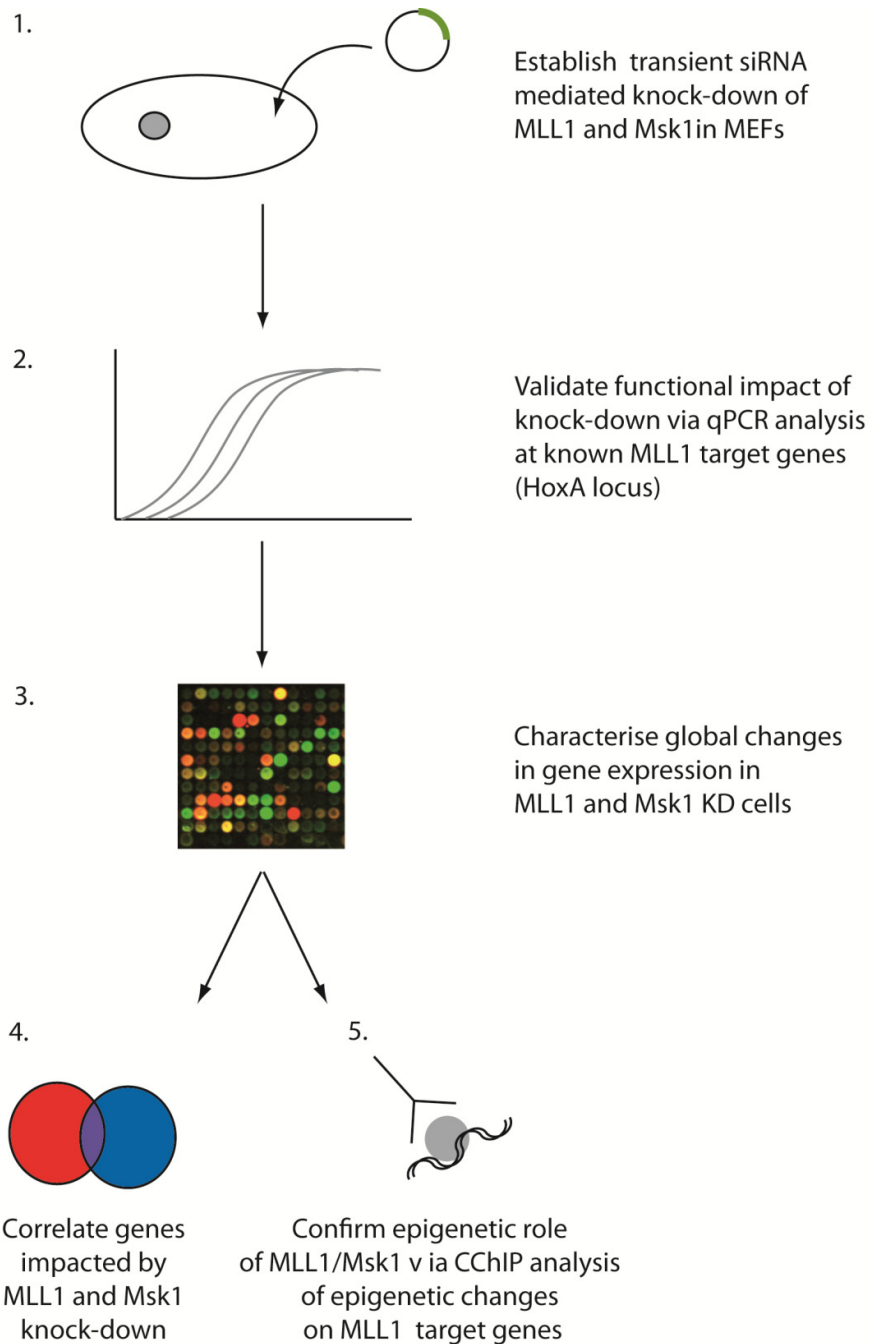
**Figure 3.9: An analysis of MLL1 and Msk1 binding correlations.** Significant binding of MLL1 and Msk1 were identified with the 'MACS peak caller'. It showed that MLL1 and Msk1 binding strength on the shared peaks correlates highly.

In this analysis MLL1 and Msk1 were cross-linked to the DNA in order to study their distribution of binding under physiological conditions and to identify MLL1 and Msk1 target genes. We have shown that MLL1 and Msk1 binding strongly correlate, but this also matches with the distribution of H3K4 tri-methylation in chromatin. These findings support our previous results that indicated there is a physical interaction between MLL1 and Msk1. The sequencing results confirmed the Hox genes are MLL1 target genes, although not the entire cluster was bound, consistent with the previous analysis of MLL1 target genes.

### **3.3 Characterising the functional impact of MLL1 and Msk1 regulation**

In the last section we established that MLL1 and Msk1 physically interact and bind to the same sites in the genome. In this section we examine if these proteins have a common regulatory effect at MLL1 target genes, using a MLL1 and Msk1 knock-down approach.

In Figure 3.10 an overview of the experiments is shown. Firstly a transient shRNA-mediated knock-down of MLL1 and Msk1 needed to be established. We chose to use MEFs to allow comparison with the binding experiments that have been described. Transient knock-down was used rather than a stable knock-out, as we wanted to characterize the immediate changes in gene expression in response to MLL1 and Msk1 down-regulation as these are likely to reflect direct effects. In the next step we assessed if these knock-downs led to a change in the expression of known MLL1 target genes (i.e. HoxA locus; Hess 2004). Previous data have shown that MLL1 and Msk1 bind a large number of sites in the genome (Chapter 3.2). Therefore changes in expression on a genome-wide scale were studied using a microarray approach. In the final experiment two representative MLL1 target genes, HoxA4 and HoxA5 were examined in knock-down cells to characterize the changes of key MLL1-associated histone modifications (H3K4me3, H3K27me3 and H3K9acS10p) in response to reduced levels of MLL1 and Msk1.



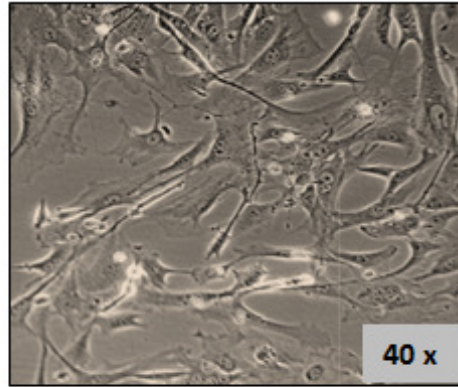
**Figure 3.10: Experimental design/workflow.** (1) A transient knock-down for MLL1 and Msk1 in MEFs needed to be established. (2) The functional impact was validated, as the HoxA cluster was screened for changes in expression in response to the knocked down proteins. (3) Microarray analysis was performed to examine the effect of MLL1 and Msk1 KD on gene expression on a global level. (4) Genes impacted by MLL1 and Msk1 knock-down were correlated to MLL1/Msk1 genomic binding and two HoxA genes were screened for changes in key histone modifications.

### **3.3.1 Establishment of a transient knock-down of MLL1 and Msk1 in MEFs**

The focus of this part of the study was to evaluate the functional impact of MLL1 and Msk1 on gene regulation in a physiological context. We decided to work with mouse embryonic fibroblasts (MEFs), as these primary cells have no genomic alterations (e.g. oncogenes or virus-infection) and as embryonic derived cells, they express HoxA genes, known MLL1 target genes (Hanson et al. 1999).

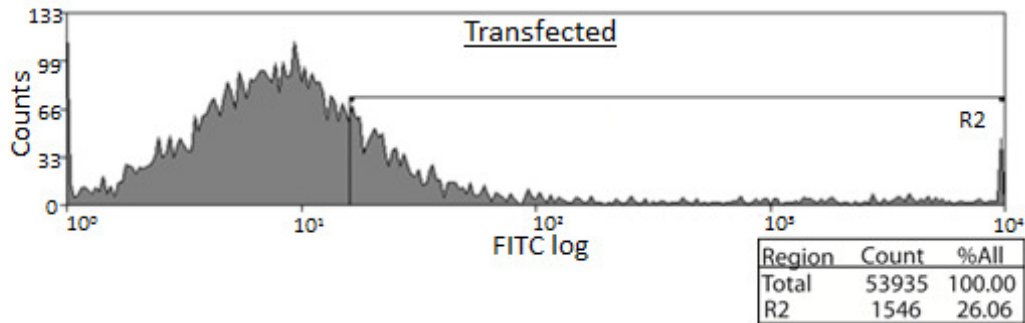
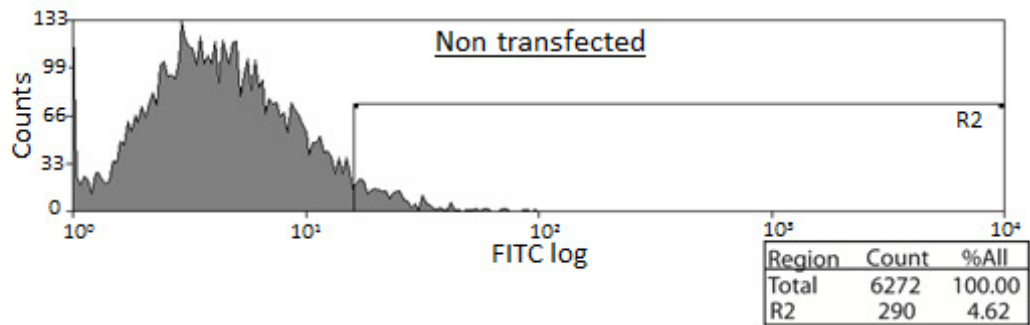
Initial experiments established conditions to transfect these cells with MLL1 and Msk1 siRNA knock-down vectors (Origene, Msk1: TR504795, MLL1: TR517798) which deliver 4-5 shRNA vectors per kit and a GFP- containing vector, which was used as a control. Once the vector is in the target cell it expresses short hairpin RNA, which is processed into siRNA that interacts with the RISC complex. This subsequently results in the cleavage of the target mRNA, resulting in reduced transcript and protein abundance (Macrae et al. 2006). Initially a range of different transfection methods were tested using the GFP vector, including lipofectamine (Sigma), X-fect (Clontech) and electroporation (Gene pulser XCell, BioRad). Evaluation of transfection efficiency using GFP fluorescence showed that all methods showed good cell viability (80% - 95%), however the transfection efficiency was low in all methods (for electroporated MEFs: Figure 3.11). This indicates a shift in the main peak of GFP positive cells, but only ~20% of the cells were highly transfected, whereas with X-fect and lipofectamine only ~10% were GFP-positive (data not shown). Therefore electroporation was used in subsequent experiments.

A



Transfected MEFs

B



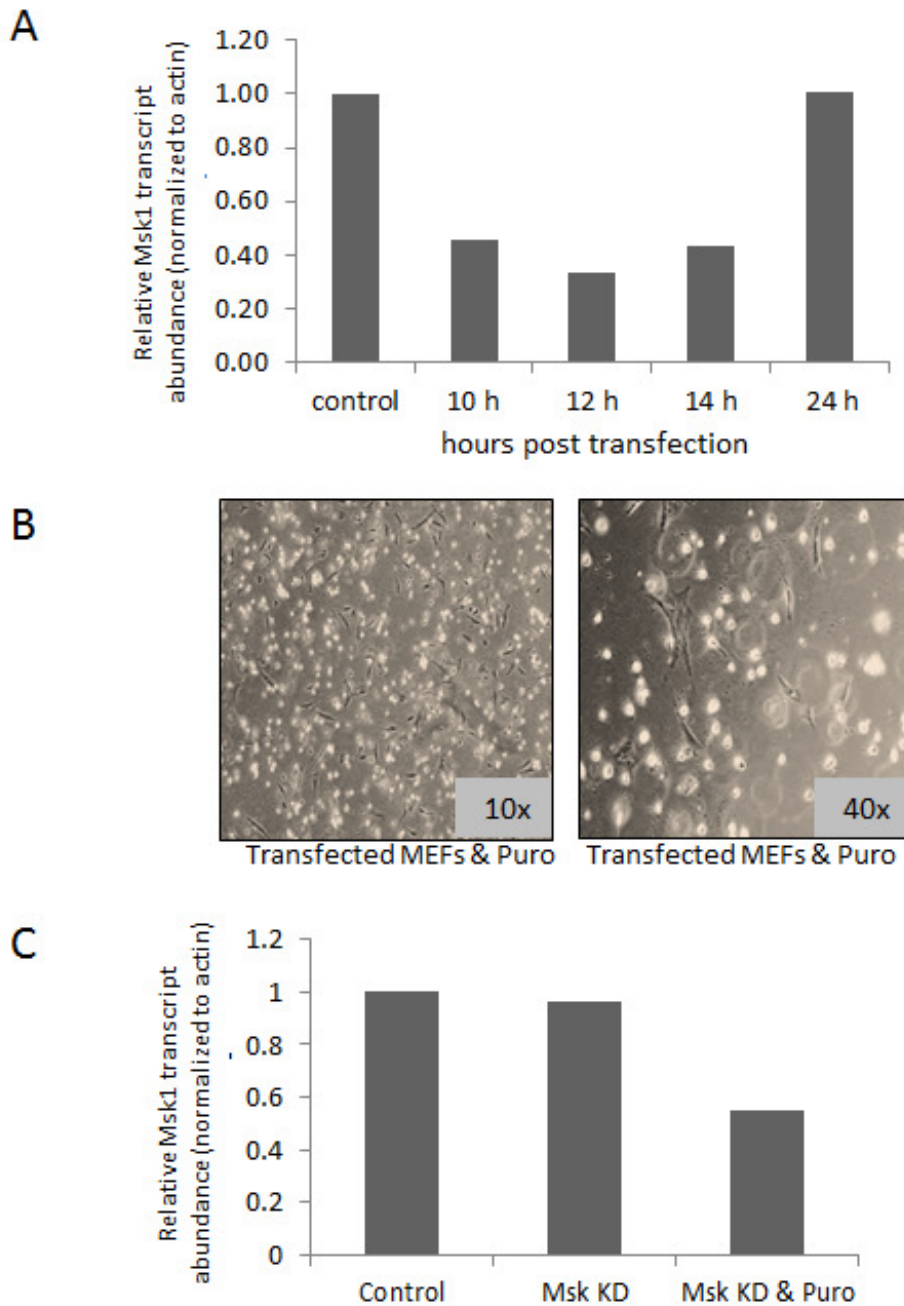
**Figure 3.11: Establishing MEF transfection protocols.** (A) Cell viability was controlled with light microscopy and trypan blue staining. Separation of live cells (attached to the plastic) and dead cells (floating) was possible. Here electroporated cells are shown. With optimised conditions, cell survival of 80% was achieved (B) The transfection efficiency of GFP was controlled by FACS analysis with cell count on y-axis and GFP fluorescence indicated on the x-axis. Lower panel indicates quantification of the cells in the R2 gate, which corresponds to cells. A shift in the measured peak could be detected, however the rate of highly transfected was around 22%.

We initially established the protocol for Msk1. In the next step the available shRNA vectors were tested for the best reduction of the target transcript. Out of the five available vectors only one showed a significant reduction of Msk1 RNA levels (data not shown) and was chosen for further experiments.

As it was a transient knock-down only a short-lived reduction in RNA levels was expected. We therefore examined the reduction of Msk1 RNA levels over an extended time scale and found that at ~12 hours post-transfection Msk1 RNA abundance was at the lowest and after ~24 hours this had recovered. However, even at the maximal knock-down, ~40% of Msk1 transcripts remained (Figure 3.12 A).

We suspected that low transfection efficiency was the reason why the knock-down levels were so low, as the majority of the cells were only poorly transfected. As these shRNA mir vectors contained a puromycin resistance cassette (Figure 2.1.A), we used this antibiotic to select for the transfected cells. MEFs were transfected and rested for six hours to allow the cells to recover and reattach to the culture flasks. Then they were treated with puromycin, and 20 hours after transfection the cells were collected for analysis. As shown in Figure 3.12 B, the living cells remained bound to the flask, whereas dead cells became detached and could be removed. qPCR analysis of the puromycin-selected cells showed that after 24 hours Msk1 RNA was still reduced by ~50% while the untreated cells had fully recovered (Figure 3.12 C).

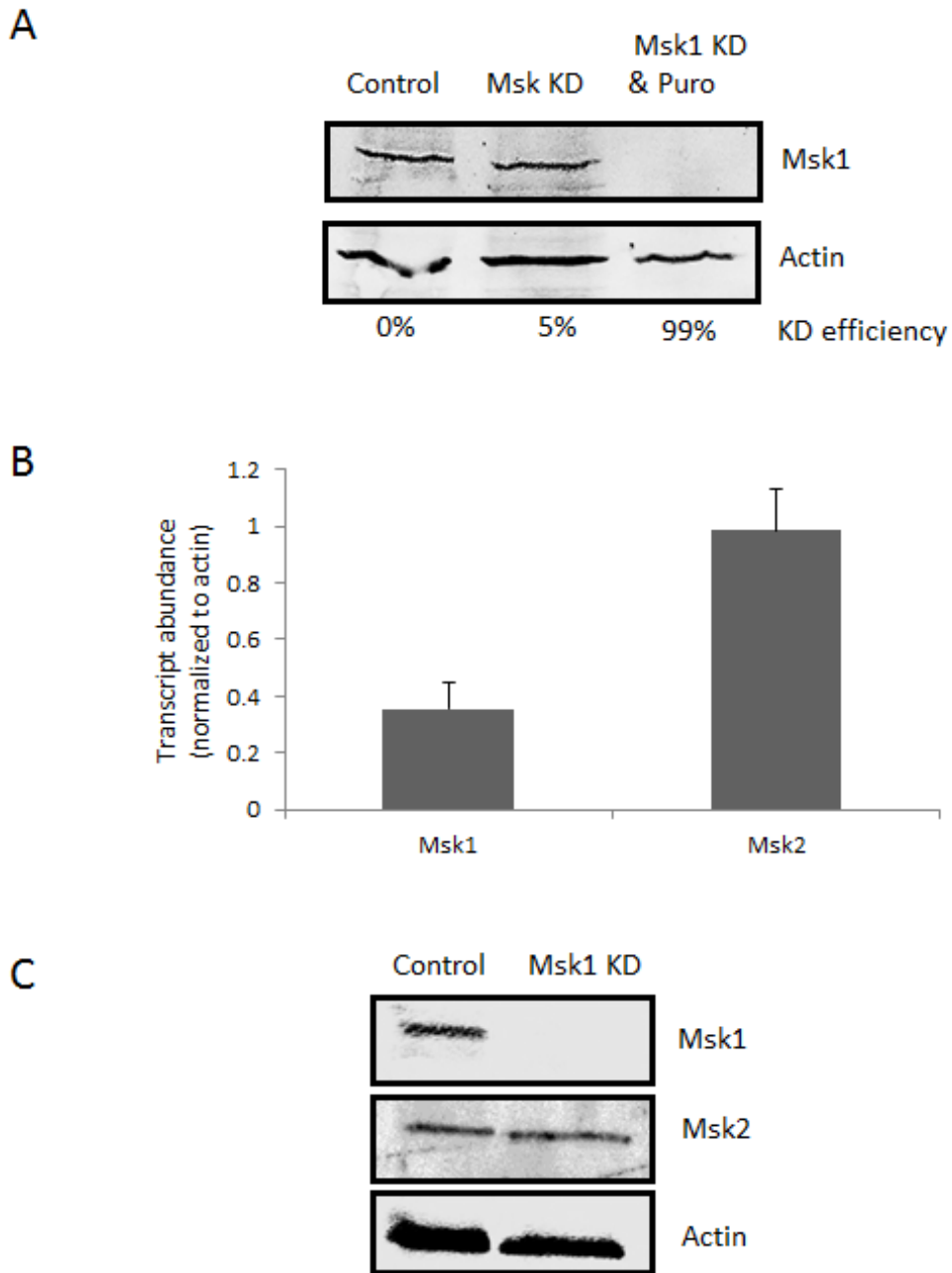




**Figure 3.12: Establishing Msk1 knock down protocol.** (A) The knock-down was transient and therefore the time-point with the maximal effect needed to be identified. For a control, cells were mock-transfected with a GFP vector. Cells were collected at the indicated time-points post transfection (n=2). Msk1 RNA levels were lowest after ~12 hours (about 35%) but recovered completely after 24 hours. (B) As transfection efficiency was low (~20%) transfected cells were treated with puromycin to the non-transfected cells. The surviving cells remain attached to the flask, whereas dead cells would float in the medium. Images were taken 24 hours after transfection. (C) Cells that were treated with puromycin showed significantly lower Msk1 transcript levels after 24 hours than those that were not treated, indicating that non-transfected cells are susceptible to puromycin-induced killing.

In order to see, if the changes Msk1 in transcript abundance would result in a change of protein level, western blots were performed. Previous studies suggest that protein levels reduce slower than RNA levels and that 20 hours post-transfection is likely to be the optimal time for reduced Msk1 protein knock-down, so the cells were harvested at this time point (both treated and untreated with puromycin) and analysed (Figure 3.13 A). The membrane was first probed with an antibody against Msk1 and then with an antibody against  $\beta$ -actin to control for loading. Cells, which were not selected with puromycin had only minimally lower Msk1 protein levels than the mock-transfected cells (Control). In contrast, puromycin-selected cells had extremely efficient Msk1 knock-down (~99% KD). This suggests that poorly transfected cells, which would buffer the observed reduction of RNA or protein levels, die during antibiotic selection and only highly transfected cells survived.

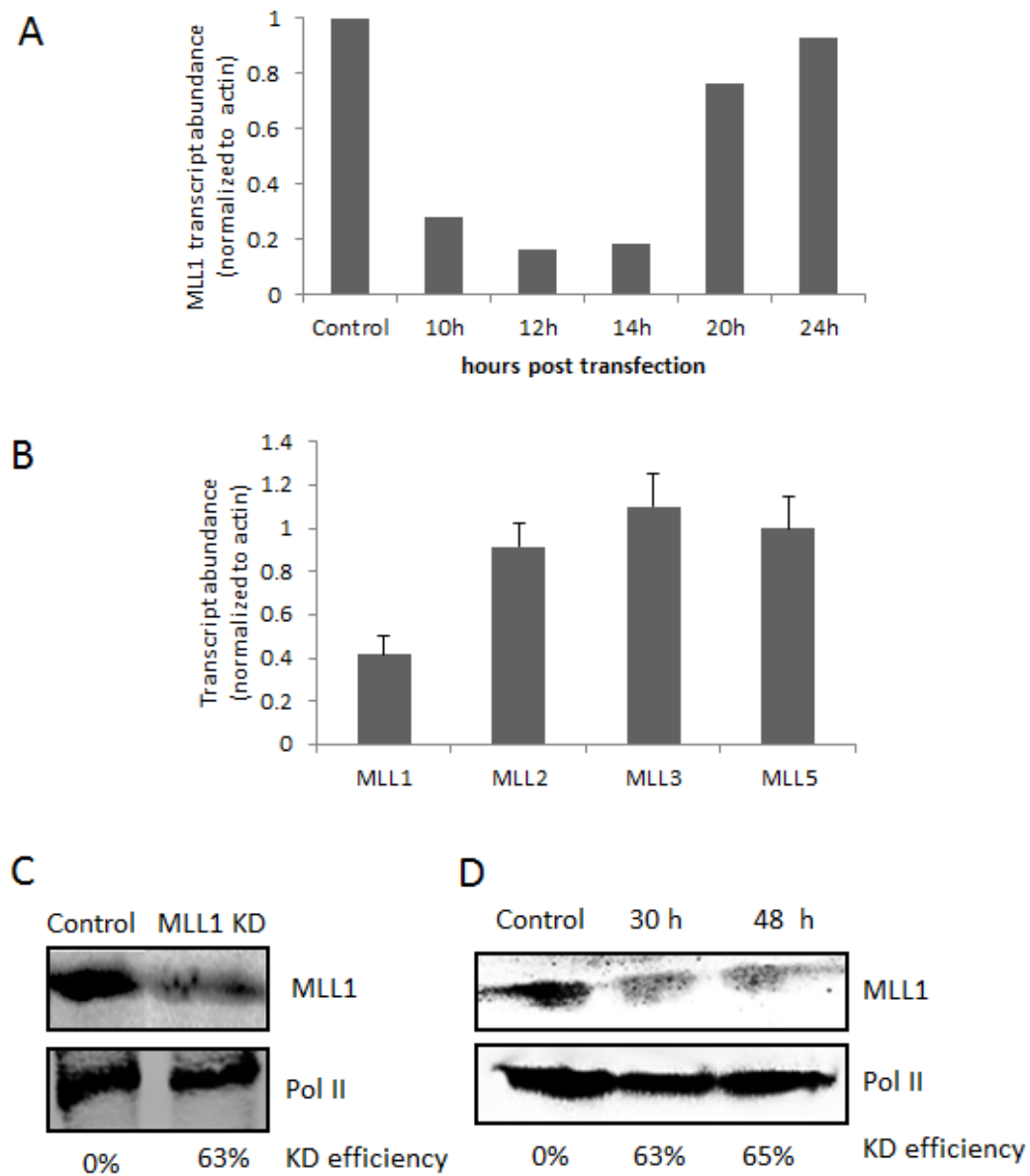
Next, the specificity of the Msk1 knock-down was confirmed. In order to exclude off-target effects, the impact of the knock-down on a similar gene/protein to Msk1 was examined. Msk2 is a mammalian homologue of Msk1, which shares ~76% sequence identity and was therefore chosen to verify the sequence specificity of the shRNA. In Figure 3.13 B Msk1 and Msk2 transcript abundance was analysed 20 hours after transfection. Msk1 RNA levels were, as seen before, reduced to ~40% whereas Msk2 transcript levels were comparable with mock transfected cells (Control). Subsequent western analysis confirmed that only Msk1 was affected, as Msk2 levels remained stable (Figure 3.13 C).



**Figure 3.13: Establishing the Msk1 knock-down protocol.** (A) Knock-down cells treated with puromycin showed a significant lower level of Msk1 protein after 20 hours. Msk1 abundance in GFP-control transfected cells (Control) and those in KD cells either untreated (Msk1) or treated with puromycin (Msk1 KD & Puro) are shown. Actin is used as a loading control. The extent of knock-down is shown below the blots. (B) Evaluating knock-down specificity. Cells were harvested after 20 hours correlating with the lowest protein levels. At this time point Msk1 transcript levels were still reduced by ~60% whereas Msk2 transcript levels were comparable to mock-transfected cells (control). Error bars indicate SD (n=3). (C) After 20 hours Msk1 protein is efficiently knocked-down, whereas Msk2 protein levels were unaltered.

Having established the conditions for the Msk1 knock-down, these were applied to the MLL1 knock-down. For MLL1 knock-down, four vectors were provided (Origene), which were tested and the most efficient chosen for succeeding experiments. In the next step, MLL1 transcript levels were monitored over time. Similar to the Msk1 knock-down, MLL1 RNA levels were lowest ~12 hours after transfection (~20% of original level) and fully recovered after 24 hours when not treated with puromycin (Figure 3.14 A).

Next the specificity of the MLL1 shRNA was controlled by analysing the transcript levels of the MLL1 homologues MLL2, MLL3 and MLL5 in MLL1 knock-down cells. As shown in Figure 3.14 B, MLL1 levels were reduced about 60%, whereas the MLL1 homologues had comparable transcript levels to the mock-transfected cells (Control). Finally, the level of MLL1 protein in the knock-down cells was determined. We treated transfected cells with puromycin and harvested the cells 20 hours post transfection (Figure 3.14 C). These were the same conditions as used for the Msk1 knock-down, however, the MLL1 protein levels were only reduced by ~60%. Analysis of transfected cells at later time-points (30 and 48 hours after transfection), found MLL1 protein at ~40% of control cell levels (Figure 3.14 D). Furthermore, the selected cells did not cope with the reduced levels of MLL1 and no viable cells were found 72 hours after transfection, even when cells were incubated in medium without puromycin. For the following experiments MLL1 knock-down cells were treated with puromycin after 6 hours and harvested after 20 hours as for the Msk1 knock-down protocol.

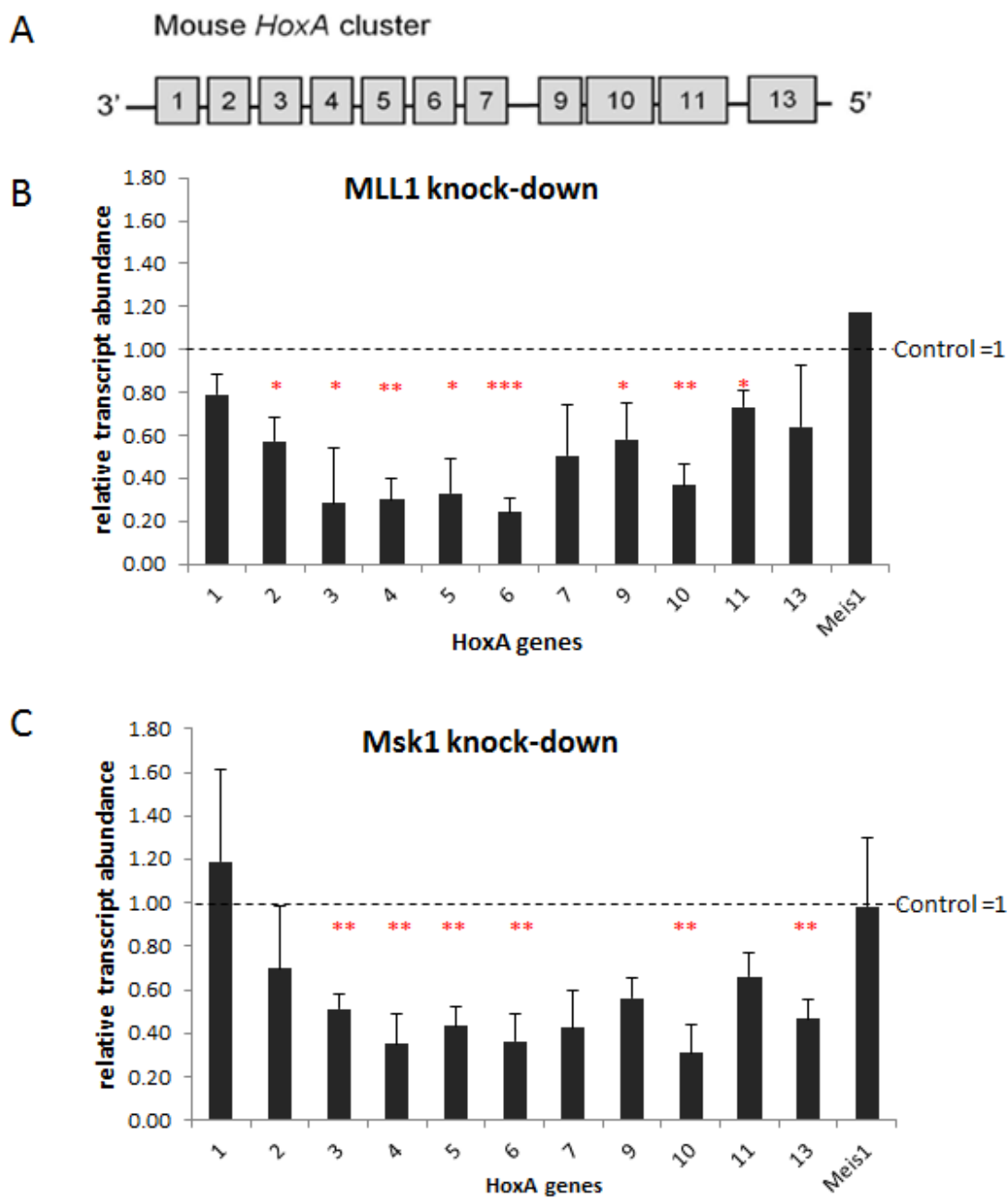


**Figure 3.14: Establishing MLL1 knock-down protocol.** (A) As the knock-down was transient, cells were harvested at indicated time-points. MLL1 transcripts were maximally knocked down at ~12 hours post transfection and recovered by 24 hours. (B) In order to test the specificity of the MLL shRNA vector, the transcript levels of MLL2, MLL3 and MLL5 were analysed in MLL1 knock-down cells. At 20 hours after transfection and treatment with puromycin, MLL1 transcript levels were ~40% of control levels, whereas MLL2, MLL3, MLL5 and control cell transcript levels were ~100%. Transcript abundance was compared with the MLL1 transcript levels in mock-transfected (Control) cells, which was normalized to 1.0 (C) Under the same conditions MLL1 protein was reduced by 63%. RNA Polymerase II (Pol II) was used as a loading control. (D) Cells harvested 30 or 48 hours post-transfection showed no further reduction in MLL1 protein

### 3.3.2 Changes of MLL1 target genes in knock-down cells

In the previous section conditions were established to create cells with significantly reduced levels of either MLL1 (~63% knock-down) or Msk1 (~99% knock-down). These cells were then studied for changes of expression at MLL1 target genes. We focussed on the HoxA genes and Meis1, as MLL1 is known to play an essential role in haematopoiesis through the maintenance of Hox gene expression (Argiropoulos & Humphries 2007b). Likewise, in MLL rearranged leukaemias, several HoxA genes and the oncogene Meis1 are consistently up-regulated (Horton et al., 2005; Orlovsky et al., 2011).

In order to capture changes in expression, RNA from knock-down and mock-transfected cells was isolated and analysed with qPCR. In Figure 3.15 A, a cartoon of the HoxA cluster is shown, which consists of 11 genes in mouse. In MLL1 knock-down cells, the expression of nearly all HoxA genes was down-regulated. Hox genes at the edge of the cluster (HoxA 1, 2 and HoxA 11, 13) showed only a moderate reduction of transcript levels (~30%). Similarly HoxA 7 and 9 were down-regulated by ~40%, whereas the greatest impact was seen on the expression of HoxA 3, 4, 5, 6 and 10, which were reduced by ~70% (Figure 3.15 B). A strikingly similar response was observed in Msk1 knock-down cells (Figure 3.15 C). HoxA 1, 2 and 11 showed no or moderate reduction in expression, HoxA 9 was down-regulated by ~40%, whereas all the other HoxA gene transcripts were reduced by 50-60%. These findings suggest Msk1 contributes to the regulation of these known MLL1 target genes. However, Meis1 expression was not changed in MLL1 or in Msk1 knock-down cells. We next examined whether this was consistent across the genome.



**Figure 3.15: Effect of MLL1 or Msk1 knock-down on HoxA and Meis1 gene expression.** The Hox genes are well known MLL1 target genes. Therefore the effect of MLL1 and Msk1 knock-down on HoxA genes was examined. (A) Cartoon of the murine HoxA cluster. (B) The changes of expression in MLL1 knock-down cells at 20 hours post-transfection. (C) The changes of expression in Msk1 knock-down cells at 20 hours post transfection. Transcript levels in mock-transfected cells (Control) are set at 1.0. Transcript abundance was normalized to  $\beta$ -actin and is presented +/- SD (n=3). Red asterisks indicate significantly down-regulated genes (paired t-test.  $p < 0.05$ \*,  $p < 0.01$ \*\*,  $p < 0.001$ \*\*\*).

### **3.3.3 Changes of gene expression on a genome wide level in knock-down cells**

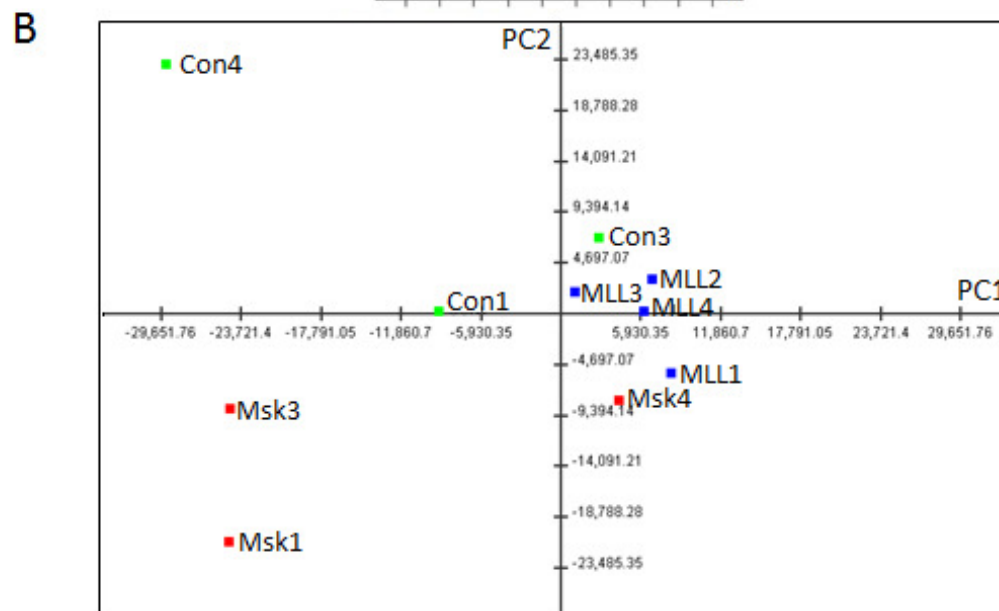
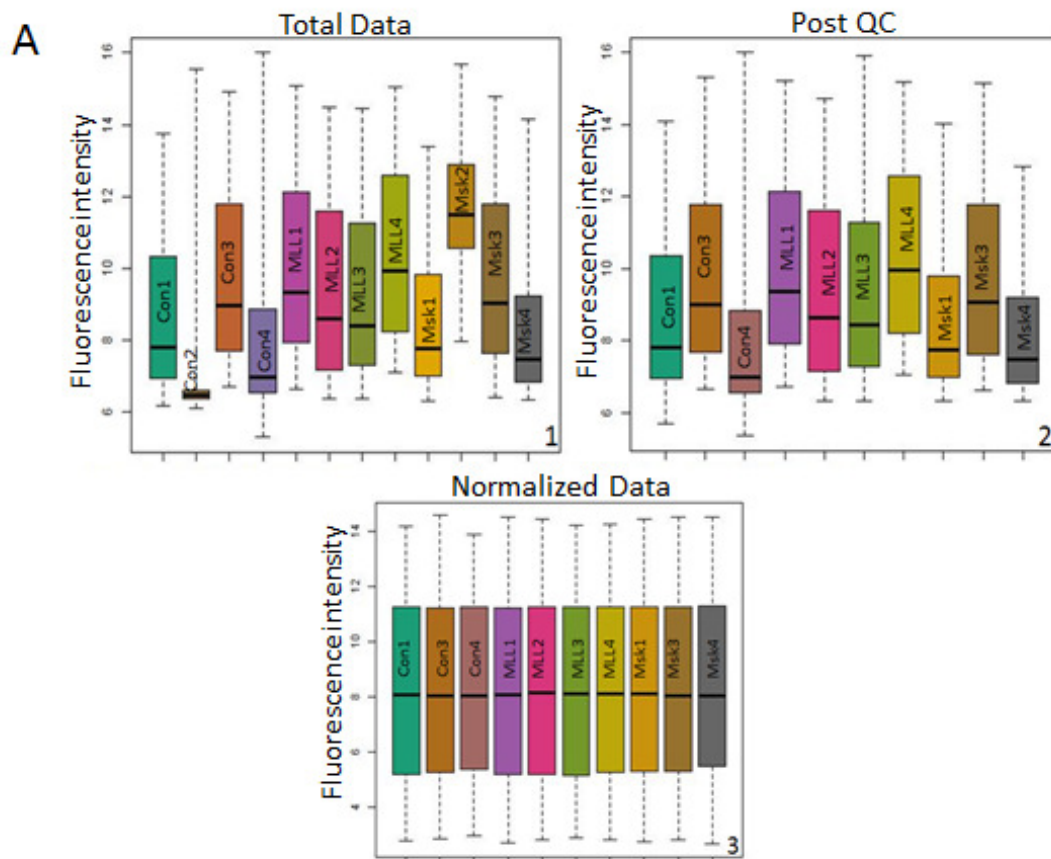
Our findings that the HoxA genes were down-regulated in MLL1 or Msk1 knock-down cells in a comparable manner suggested that MLL1 and Msk1 are functionally connected, at least at the HoxA loci. Our genome-wide analysis (Chapter 3.2) showed that MLL1 and Msk1 bind to a number of same sites in the genome, suggesting these proteins regulated a larger number of genes beyond the HoxA locus. Therefore changes in gene expression on a genome-wide scale were examined.

For this analysis the 100718\_MM9\_EXP\_HX12 microarray from NimbleGen was used. This slide contains 12 individual arrays, each representing 42,576 probes in triplicate. Four samples for each knock-down and mock-treated cells were prepared. RNA was extracted, reverse-transcribed into cDNA and labelled with Cy5. The labelled samples were hybridized onto the arrays and scanned. In the first step of the analysis, the quality of the scan was validated (Figure 3.16 A). Ideally the same amount of cDNA should hybridize to each array and give comparable fluorescence signals. However signal intensity can vary due to slightly different loading, inefficient hybridization or differential scanning quality on different parts of the slide. Hence, for each array the range of signal readings was analysed. These are presented as box and whisker plots, where the median of the box-plots represents the average density of the scan and the intensity of 75% of the features on the array lie within the box. In this format, the maximum and the minimum intensities lie at the end of the whiskers. Two arrays showed a significant difference in signal intensity from the other samples, suggesting that there was a problem with either the labelling or hybridization (e.g. Control2 & Msk2, Figure 3.16 A). For further analysis these two samples were excluded (Graph 1 and 2). In order to make the arrays comparable a quantile normalisation was

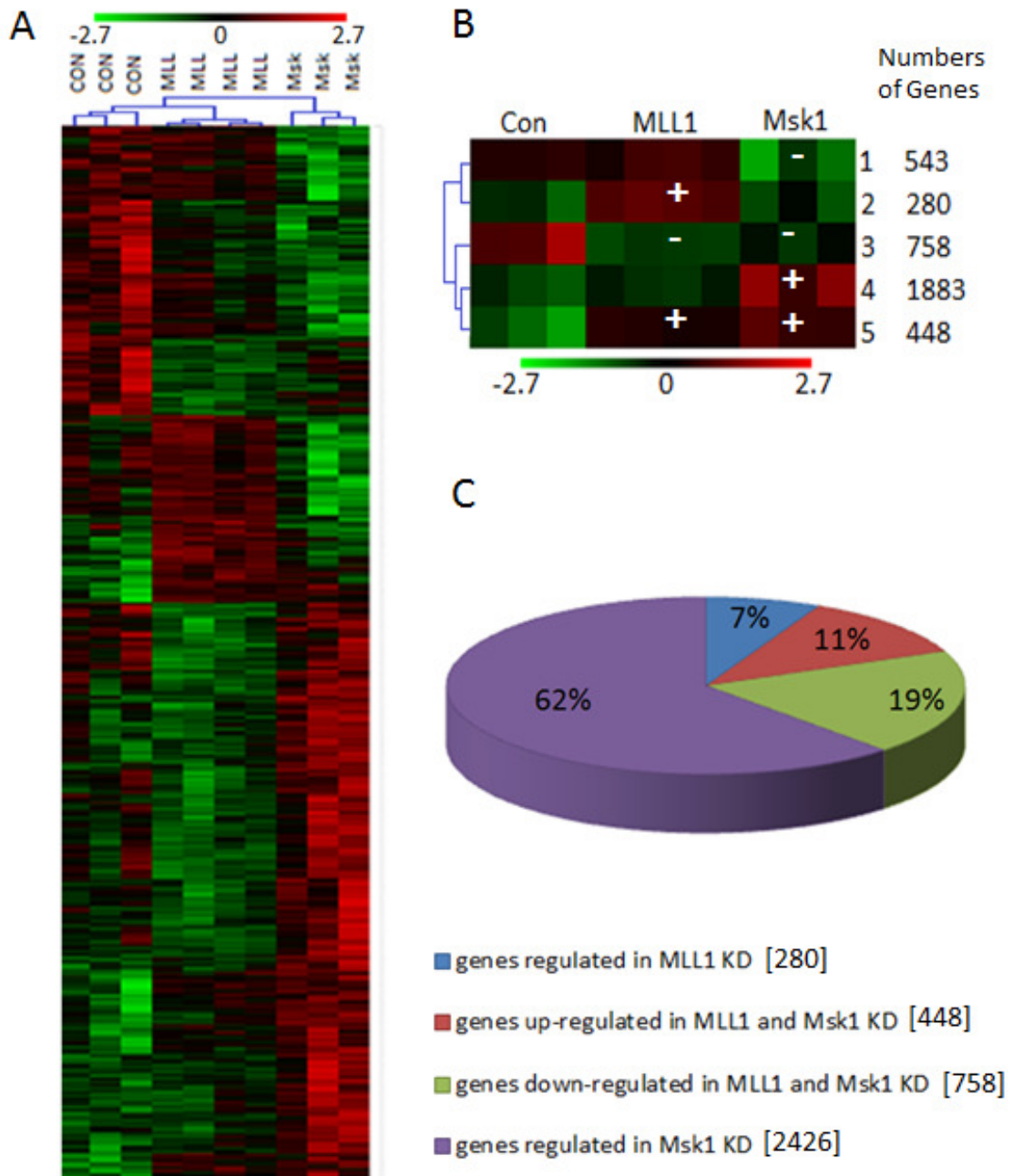


performed using the RMA (Robust-Multi-Array Average) pre-processing methodology implemented in R (Krause & Olson 2000). This normalization process adjusts for systematic errors arising from differences in the method or dye intensity to make further analysis meaningful (Graph 3). The normalised data were then analysed by principal component analysis (Figure 3.16 B), which is a mathematical algorithm that identifies patterns in the data sets, by identifying maximal trends, called principal components (Ringnér 2008). The analysis showed that the four samples for the MLL1 knock-down cells were closely clustered near the origin, indicating that they were quite similar (Figure 3.16 B). In contrast the three samples from Msk1 knock-down cells were mainly located to the lower left quadrant, whereas the three samples from control cells were located mainly in the upper left quadrant, and showed a wider variation between the data sets. However, this was acceptable, as each cell-condition formed a cluster with a defined territory.

In order to identify differentially expressed genes between the control and the knock-down cells, SAM (significance analysis of microarrays) analysis was performed in MeV (Tusher et al. 2001). The results of changes in expression were represented in a heat map (Figure 3.17 A). At the top of the heat map, a hierarchical tree diagram is shown, which visualizes the relationship between the data sets and shows that the three cell conditions are significantly different from each other. Each line represents a gene, which showed a change of fluorescence during scanning in the different cell types. These changes in fluorescence intensity correlate with changes in expression. In total 3912 genes were statistically different between the knock-down and the control cells (fold change cut-off:  $FC > 2$ ). The colouring scale is relative, in which higher gene expression is marked in red and lower gene expression in green.



**Figure 3.16: Microarray analysis of MLL1 and Msk1 knock-down cells.** (A) In the first step the microarray scans were quality controlled. Two extreme outliers were excluded from the analysis (Con2 & Msk2: compare panels 1 and 2) and all remaining samples were normalized (panel 3). (B) A Principal Component Analysis was performed to evaluate the differences between the three groups: Control (1, 3 & 4), MLL KD (1, 2, 3 & 4) and Msk KD (1, 3 & 4).



**Figure 3.17: Significance analysis of gene expression changes for MLL1 and Msk1 KD.** (A) To analyse a SAM (Significance analysis of microarrays) was performed to find significant changes between patterns of gene expression in Control, MLL1 knock-down and Msk1 knock-down cells. Gene expression changes are indicated by colour intensity of a range from 2.7 down (green) to 2.7 up (red). Samples are clustered by hierarchical tree at the top of the plot. (B) In total 3912 genes changed and could be clustered into five groups. The number of genes in each cluster is indicated on the right, the hierarchical tree on the left. (C) Out of these 3912 genes only 19% were down-regulated in both MLL1 and Msk1 knock-down cells (Cluster 3) and only 11% were up-regulated in both knock-downs (Cluster 5). This indicates the majority of genes respond differently to MLL1 or Msk1 knock-down.

Approximately ~9% of the genes on the microarray (3912 genes) showed changes in expression and could be clustered into five groups (using SOTA algorithm in MeV, Figure 3.17 B). Interestingly, of these genes, only ~11% were up-regulated and 19% down-regulated together in both knock-downs when compared with the control cells (Figure 3.17 C). This indicates that the majority of genes respond differently to MLL1 or Msk1 knock-down, suggesting they are regulated differently.

In order to understand which cellular processes are affected by the knock-down of MLL1 and Msk1, the genes were functionally annotated using DAVID. DAVID is a biological data-base and analytical tool, which compares a gene list in a list to the data base in order to correlate the genes to known biological functions (Huang et al. 2008). The gene list for each cluster was analysed individually and only genes with a FDR (false-discovery-rate) of lower than 15% were considered (Figure 3.18 and 3.19). This identified five clusters:

In *cluster 1* genes were grouped which were down-regulated in only Msk1 knock-down cells. These genes were functionally enriched with the cell cycle, chromosome condensation and DNA repair.

*Cluster 2* contains genes which were up-regulated in only MLL1 knock-down cells. Although 280 genes were in this cluster, only 66 genes could be connected to a specific function (eg. 21 genes to cellular stress-response). This suggests that these changes in gene expression were indirect responses to the changes in the cell, or are unlikely to be a coordinated response to the knock-down.

In *cluster 3* genes were down-regulated in both MLL1 and Msk1 knock-down cells. Most of these genes were involved in regulation of MAPK activity (81 genes), the immune response (57 genes) and related functions (e.g. cytokine activity, 16 genes). MAPKs are known to have

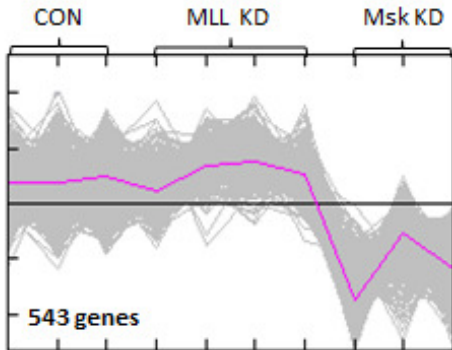
a regulatory influence on the immune response (Liu, Shepherd, & Nelin, 2007; Arthur & Ley, 2013), so the down-regulation of these genes in Msk1 knock-down cells is expected. This also supports a co-regulatory function of MLL1 and Msk1. However, a large number of these genes are also annotated with cellular maintenance (e.g. sugar/lectin binding, 42 genes).

*Cluster 4* genes were up-regulated in Msk1 knock-down cells and were found to be related to MLL1 and Msk1 function. For example histone protein genes (79 genes), transcription associated genes (72 + 55 + 33 genes) or protein kinases (15 genes) were up-regulated. This could be a cellular reaction to compensate for the loss of Msk1, which is an important kinase and involved in many cellular processes.

Finally, *cluster 5* genes were up-regulated in MLL1 and Msk1 knock-down cells. These genes were all associated with cellular maintenance (e.g. homeostatic processes, 35 genes).

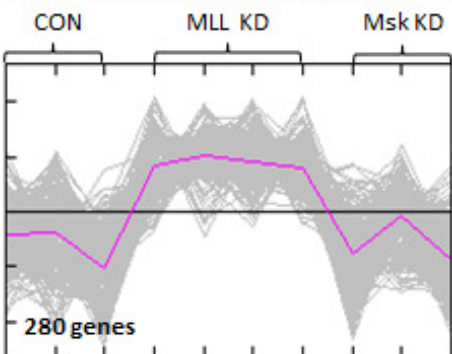
In order to verify the microarray findings, qPCR analysis of selected genes were performed to confirm their up- or down-regulation. Two candidate genes, which were up-regulated and two, which were down-regulated in each knock-down were chosen. For the MLL1 knock-down *Il1b* and *Ly9* (down-regulated) and *Hemt1* and *Hsph1* (up-regulated) and for the Msk1 knock-down *Ttk* and *Kif20b* (up-regulated) and *Ifnar2* and *Tet3* (down-regulated) were examined. The results are shown in Figure 3.20 and confirmed the microarray data.

Cluster 1: Genes only down in **Msk1** KD



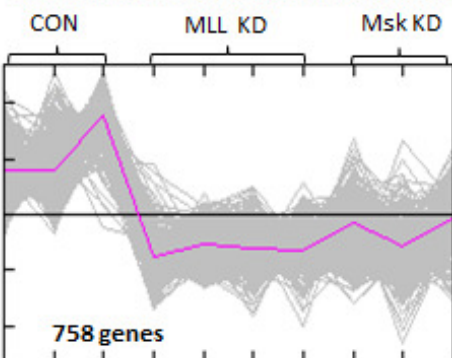
Cell cycle	404
Chromosome/chromosome condensation	160
Microtubule/Cytoskeleton	241
DNA repair	114

Cluster 2: Genes only up in **MLL1** KD



Interphase	22
Cell junction	12
Cellular response to stress	21
MAPK signaling/protein kinases	11

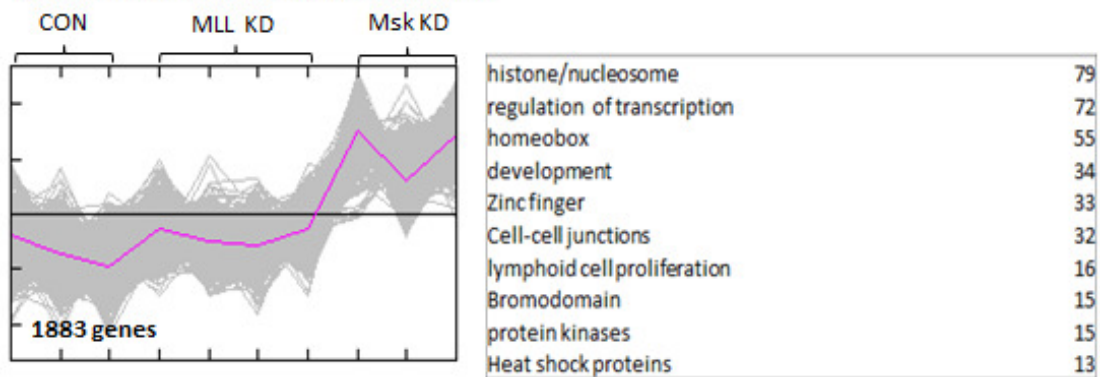
Cluster 3: Genes down in **MLL1** and **Msk1** KD



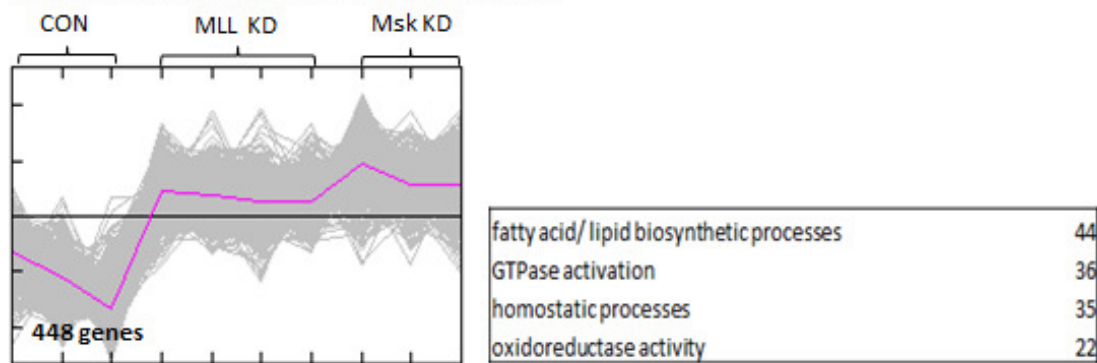
regulation of MAPK activity	81
cellular amino acid catabolic process	30
sugar/lectin binding	42
growth factor/regulation of cell division	18
immune-response signaling	18
immune response	57
lymphocyte mediated immunity	17
cytokine activity	16

**Figure 3.18: Cluster and gene annotation analysis I.** Each gene cluster was then analysed with DAVID, a tool for functional annotation of gene lists. The changes in gene expression in the cluster is shown in the left box and the associated annotations with their corresponding hits in the right box (false discovery rate cut-off: 15%, Enrichment>1.5). Significant biological associations and the number of genes involved are shown. Plots indicate the range of gene expression changes between control (1-3), MLL1 KD (1-4) and Msk1 KD (1-3) data set. The purple line indicates the average.

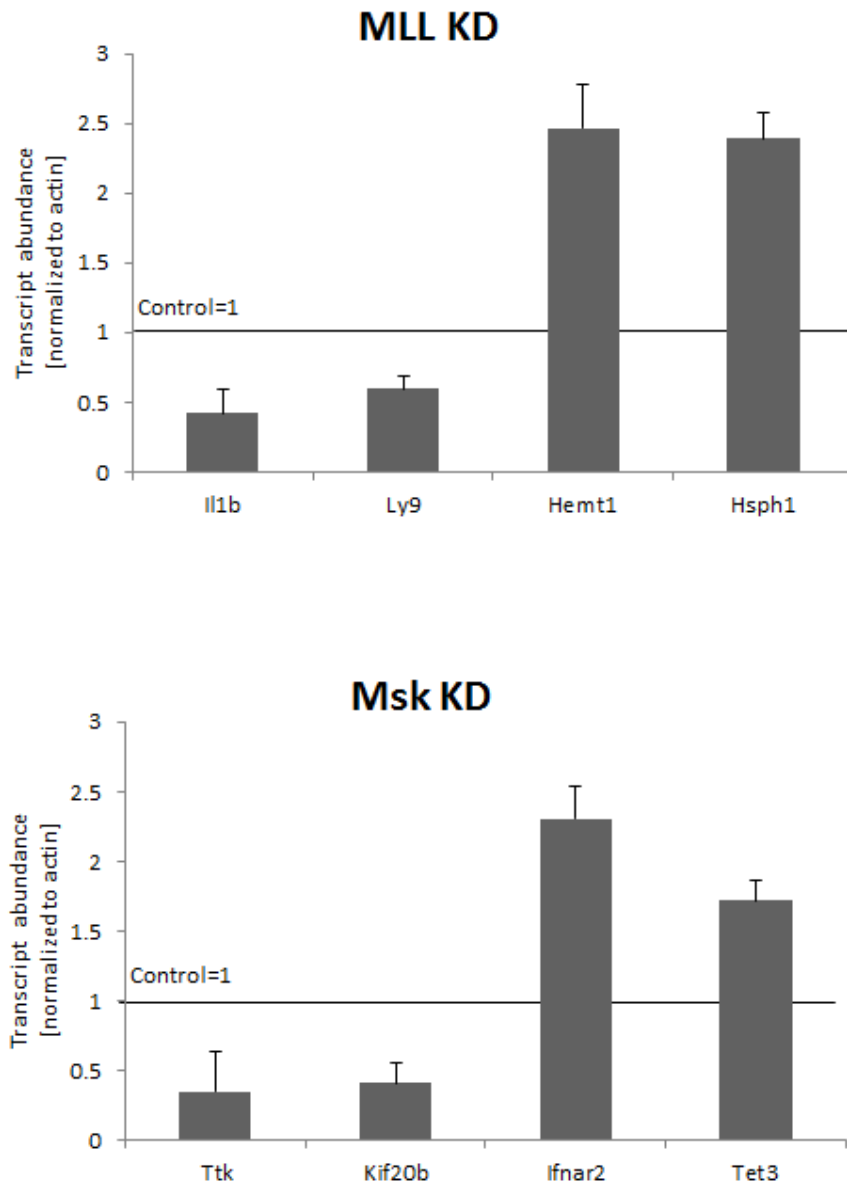
**Cluster 4: Genes only up in Msk1 KD**



**Cluster 5: Genes up in MLL1 and Msk1 KD**



**Figure 3.19: Cluster and gene annotation analysis II.** Each gene cluster was then analysed with DAVID, a tool for functional annotation of gene lists. The changes in gene expression in the cluster is shown in the left box and the associated annotations with their corresponding hits in the right box (false discovery rate cut-off: 15%, Enrichment>1.5). Significant biological associations and the number of genes involved are shown. Plots indicate the range of gene expression changes between control (1-3), MLL1 KD (1-4) and Msk1 KD (1-3) data set. The purple line indicates the average.



**Figure 3.20: Validation of microarray data by qPCR.** To validate the gene up-/down-regulation data from microarray analysis, gene expression changes in Msk1 and MLL1 knock-down cells were analysed by qPCR. Two genes that were up-regulated (Hemt1 and Hsph1) and two that were down-regulated (Il1b and Ly9) in MLL1 knock-down cells, furthermore, two genes that were up-regulated (Ifnar2 and Tet3) and two that were down-regulated (Ttk and Kif20b) in Msk1 knock-down cells, were tested. Graphs indicate transcript abundance normalized to  $\beta$ -actin and indicate  $\pm$  SD (n=3). All changes were statistically significant (paired t-test,  $p < 0.05$ ). Data were normalised to gene expression levels in control cells (Control=1).

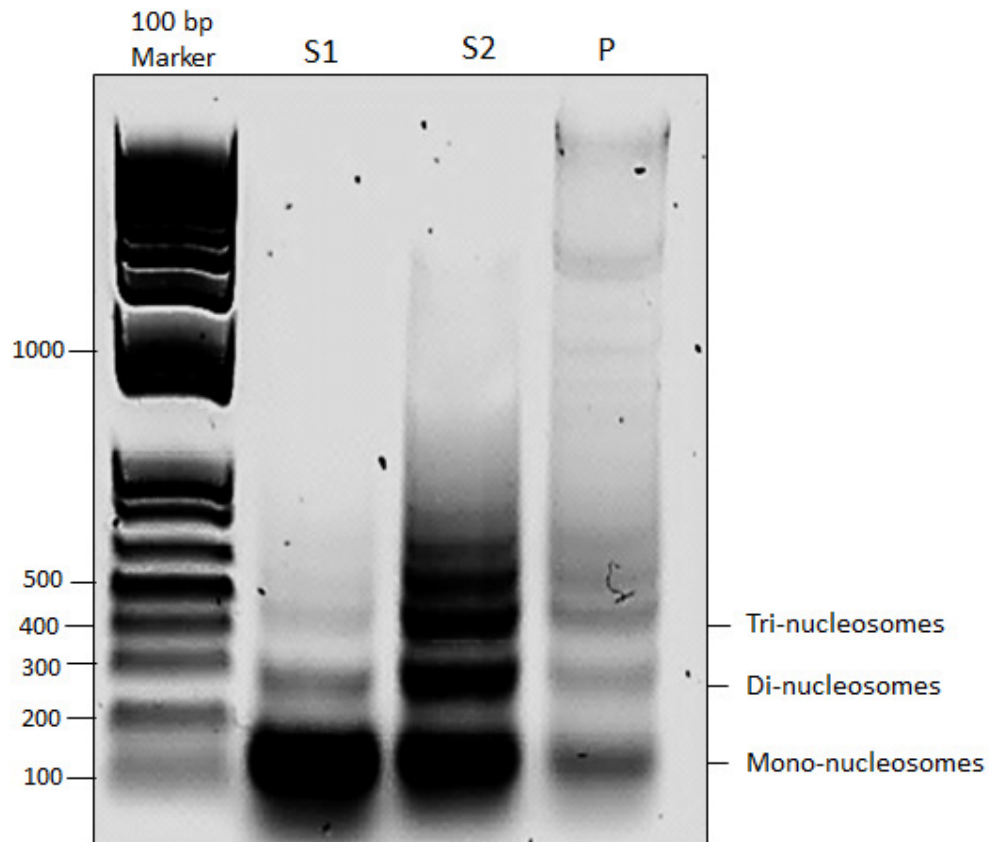


This analysis characterised the changes in gene expression induced in response to the knock-down of MLL1 and/or Msk1. In general, annotated genes reflected MLL1- and Msk1-related functions (eg. cellular response to stress, *cluster 2*). Notably, it seemed that the Msk1 knock-down had a more dramatic effect on gene expression than MLL1: 543 were down- and 1883 up in Msk1 knock-down cells, whereas only 280 genes changed expression upon MLL knockdown. This may reflect the extent of knock-down, as MLL knock-down was less efficient than Msk1, however, a large number of affected genes were not annotated with a specific function in DAVID, indicating a limitation of this type of analysis. Finally DAVID cannot distinguish between changes in gene expression in direct response to the knock-down and indirect effects which reflect secondary and tertiary changes resulting from the initial response.

### **3.3.5 Changes of histone modifications on MLL1 target genes in MLL1 and Msk1 knock-down cells**

We established that knock-down of MLL1 and Msk1 induces changes in gene expression at a large number of genes with a diversity of cellular functions. As MLL1 works as a transcriptional activator by modulating histone modification, we then examined if down-regulation of MLL1 or Msk1 would induce changes in the histone modification of target loci. Previous studies showed that MLL1 deposits H3K4me3 (Wang et al. 2009) and is stimulated by H3K9ac/S10p in vitro (Nightingale et al. 2007). MLL1 is a mammalian trithorax protein (TRX), regulators which are counter-acted by polycomb proteins (PCG), which deposit H3K27me3. For these reasons, these three marks were examined on the MLL1 target genes HoxA 4 and 5 in context of MLL1 or Msk1 knock-down cells. These genes were chosen because they showed large reductions in RNA levels in MLL1 and Msk1 knock-down cells (Figure 3.15) and we assumed that any changes in histone modifications would be the easiest to detect at these loci.

In order to study the distribution of modifications, a carrier Chip (CChIP) approach was used, as this technique requires a small number of target cells (i.e.  $1 \times 10^5$ ), in this case knock-down cells. Target cells were mixed with SL2 Drosophila cells to bulk up the cellular material and minimise the loss of target chromatin during the procedure. It was important to use a Drosophila cell line as this prevents non-specific amplification of the carrier at the end of the ChIP procedure. Cells were processed similar to native ChIP, allowing the digestion of chromatin with micrococcal nuclease, which resulted in chromatin fragments between  $\sim 100\text{bp}$  and  $\sim 500\text{bp}$  (Figure 3.21)



**Figure 3.21: Chromatin Isolation from knock-down cells.** For the ChIP analysis, knock-down cells were mixed with *Drosophila* SL2 cells (“carrier cells”) and the chromatin was digested with micrococcal nuclease, resulting in the typical ‘chromatin ladder’. Mono-di- and tri-nucleosomes are labelled, as is a 100bp ladder. S1, S2 and P are fractions of chromatin obtained during the chromatin precipitation procedure (see Chapter 2.4.5 Carrier Chromatin Immuno-precipitation). Only S1 and S2 were used for immuno-precipitation.

This allowed higher resolution analysis of the target genes than using cross-linked chromatin, which normally generates ~500bp fragments. In Figure 3.22 A and 3.23 A, the primers for HoxA4 and 5 are shown. These were designed to amplify a reference area at the promoter 5' to the TSS (primers 1), to amplify the TSS (primers 2), or an intra-genic region at the end of at the gene (primers 3). All primers were tested prior to the experiment for specificity to murine DNA.

In Figure 3.23 B the results for the ChIP on HoxA4 are shown. As expected, H3K4me3 levels, both 5' (HoxA4.1) or at the end of the gene (HoxA4.3), are low and are not altered in the different cell conditions. In contrast, a peak of this mark is seen over the TSS in control cells and is reduced in MLL1 (-40%) and Msk1 (-50%) knock-down cells. T-test analysis indicates this is a significant decline in both MLL1 ( $p=0.040$ ) and Msk1 knock-down cells ( $p=0.010$ )

A similar pattern of enrichment is seen for H3K9ac/S10p, with low levels of H3K9ac/S10p in the upstream region and at the end of the gene. Likewise, enrichment was observed over the TSS in the mock-transfected cells, which was diminished in MLL1 (-50%,  $p=0.009$ ) and Msk1 (-40%,  $p=0.010$ ) knock-down cells.

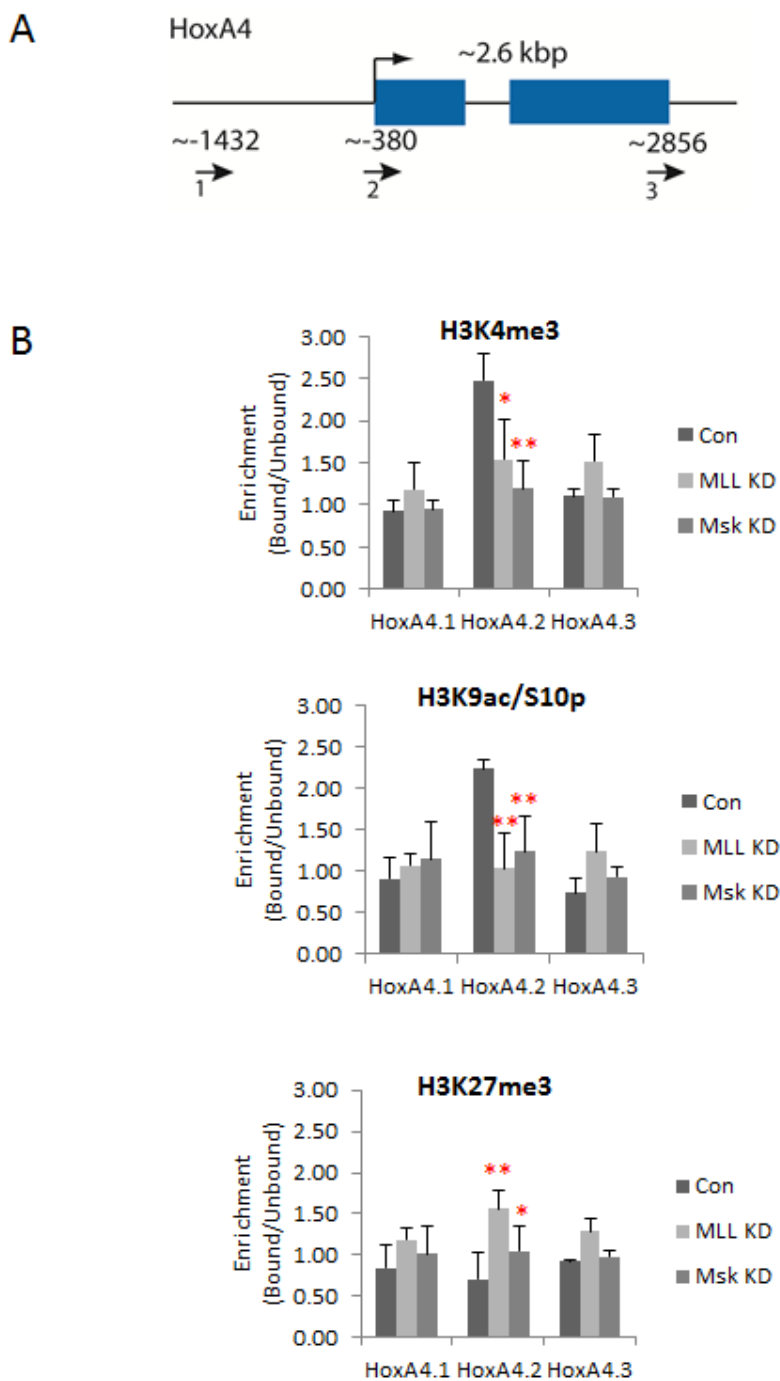
Finally, no enrichment of H3K27me3 was observed either 5' or 3' to the gene, or over the TSS. Interestingly a small increase in this mark (ca. 1.5 fold), which is significant ( $p=0.005$ ) is seen in MLL1 knock-down cells. In Msk1 knock-down cells the same trend was observed ( $p=0.046$ ).

Similar results were seen for HoxA5 (Figure 3.23 B). H3K4me3 abundance is low both upstream (HoxA5.1) and at the end of the gene (HoxA5.3) in the different cells. However, over the TSS (HoxA5.2) an enrichment of H3K4me3 was observed in control cells, which was reduced in MLL1 (-60%,  $p=0.003$ ) and Msk1 (-50%,  $p=0.002$ ) knock-down cells.

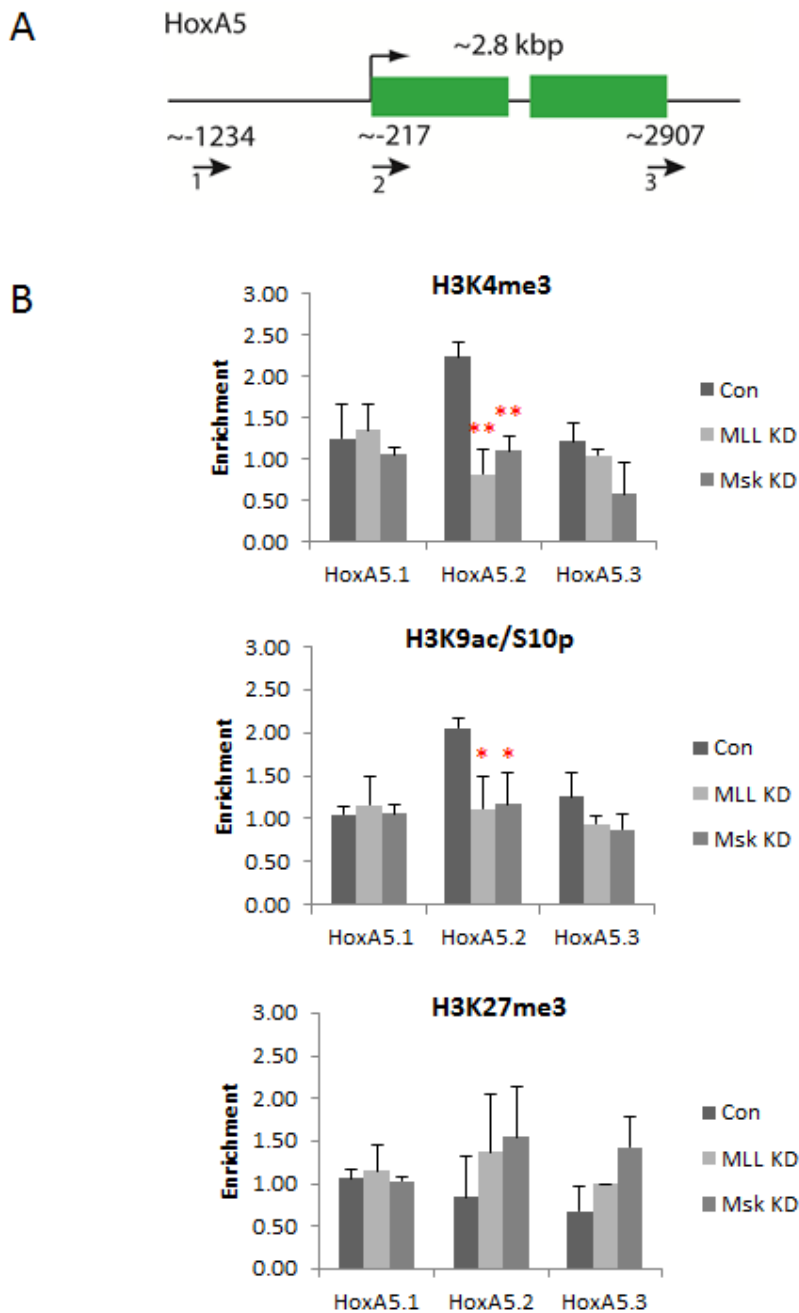
Similarly, for H3K9ac/S10p, an enrichment over the TSS observed in mock-transfected cells, was diminished in MLL1 (-50%,  $p=0.014$ ) and Msk1 (-40%,  $p=0.019$ ) knock-down cells.

Finally, no enrichment of H3K27me3 was observed upstream (5') or the end of the gene under the different cell conditions. Over the TSS no enrichment was observed in mock-transfected cells, but in knock-down cells a small increase in H3K27me3 levels were seen. However these changes were not statistically significant (MLL1  $p=0.413$  and Msk1  $p=0.183$ ).

In general, these experiments show that knock-down of MLL1 and Msk1 changes the histone modification landscape at MLL1 target genes. Furthermore, the changes in MLL1 knock-down cells were analogous to those in Msk1 knock-down cells. Moreover it confirmed the antagonistic action of the trithorax (i.e. H3K4me3 and H3K9ac/S10p levels down) and polycomb group proteins (ie. H3K27me3 up) at these MLL1 targets. These findings support the hypothesis that MLL1 and Msk1 work in concert to regulate gene expression using epigenetic mechanisms.



**Figure 3.22: Effect of MLL1 and Msk1 knock-down on H3K4me3, H3K27me3 and H3K9ac/S10p distribution at HoxA4.** (A) Gene map for HoxA4 (boxes = exons). Primers were designed to screen multiple sites on the HoxA4 gene (representative MLL1 target gene). (B) MEFs were transfected with a Msk1 or a MLL1 knock-down construct. After 20 hours the cells were harvested and ChIP analysis of H3K4me3, H3K27me3 and H3K9ac/S10p was performed. Graphs present these modifications at three sites over the HoxA4 gene, with the abundance in mock transfected cells normalized to 1.0. Red asterisks mark significant changes in modification levels (paired t-test,  $p < 0.05$ \*,  $p < 0.01$ \*\* ,  $p < 0.001$ \*\*\*).



**Figure 3.23: Effect of MLL1 and Msk1 knock-down on H3K4me3, H3K27me3 and H3K9ac/S10p distribution at HoxA5.** (A) Gene map for HoxA5 (boxes = exons). Primers were designed to screen multiple sites on the HoxA5 gene (representative MLL1 target gene). (B) MEFs were transfected with a Msk1 or a MLL1 knock-down construct. After 20 hours the cells were harvested and ChIP analysis of H3K4me3, H3K27me3 and H3K9ac/S10p was performed. Graphs present these modifications at three sites over the HoxA5 gene, with the abundance in mock transfected cells normalized to 1.0. Red asterisks mark significant changes in modification levels (paired t-test,  $p < 0.05^*$ ,  $p < 0.01^{**}$ ,  $p < 0.001^{***}$ ).

### **3.4 MLL1 and Msk1 during the cell cycle**

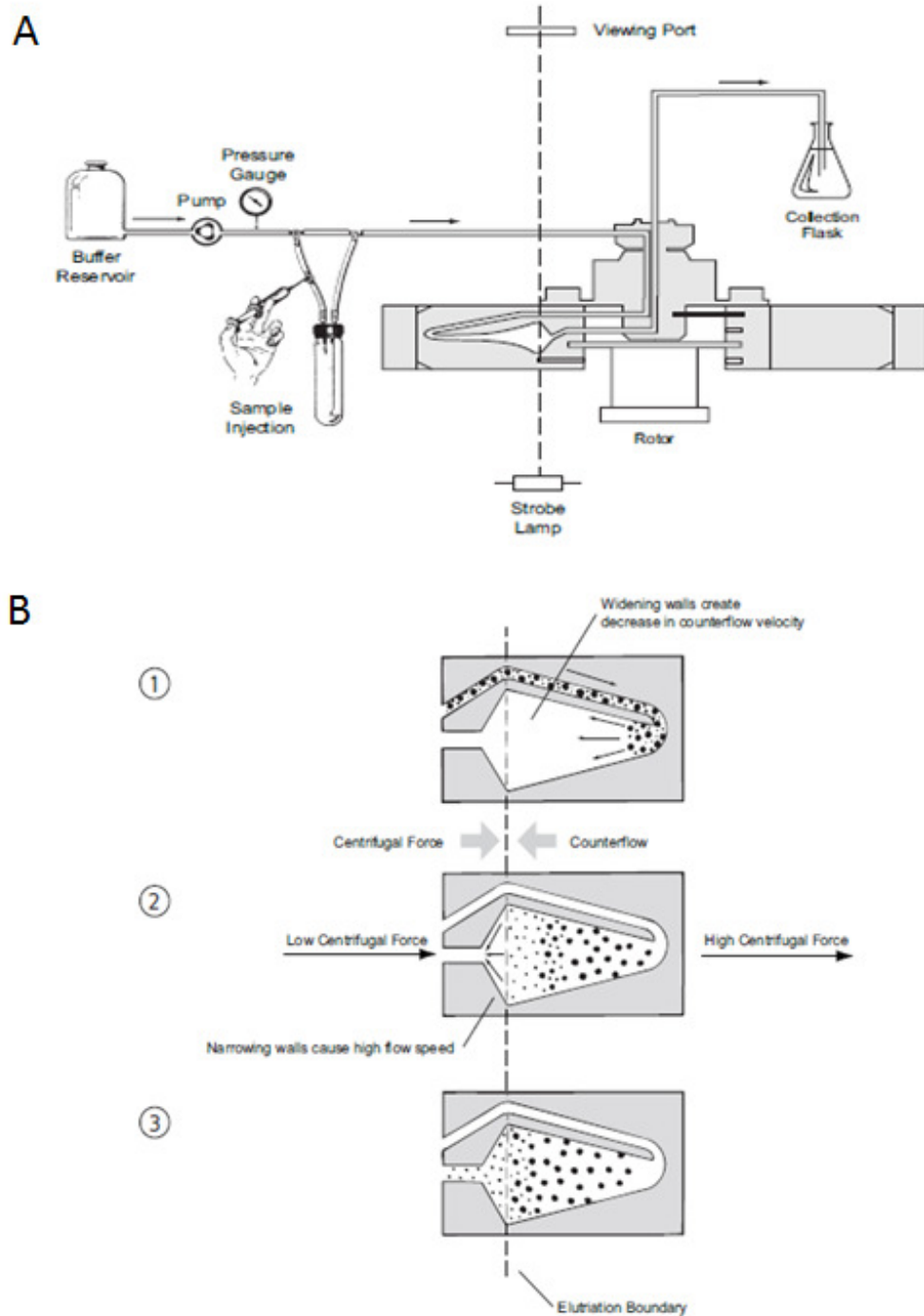
Previous studies have linked MLL1 to cell cycle regulation and our analysis of the distributions of MLL1 and Msk1 binding indicated that both proteins bind to genes, which are functionally linked to the cell cycle (see Chapter 3.3, Figure 3.18 and 3.19). This argues for a regulatory impact of MLL1 and Msk1 on the cell cycle, however, previous reports disagree about MLL1 function during the cell cycle. Several studies suggest that MLL1 may have an active role in cell cycle progression. For example MLL1, in concert with menin, has been shown to regulate cyclin-dependent inhibitors (Milne et al. 2005). This was consistent with observations that MLL1 abundance oscillates throughout the cell cycle, with peaks in G1/S and M-phase, but remains associated with chromatin even at its lowest levels during mitosis (Caslini et al. 2000; Liu et al. 2007). In contrast, another study showed that MLL1 is displaced during mitosis (Ennas et al. 1997) and that H3K4me3, although still found in condensed chromatin, changes dynamically during the cell cycle (Mishra et al. 2009). More recently, it has been suggested that MLL1 retention at gene promoters during mitosis accelerates transcription reactivation following mitotic exit (Blobel et al. 2009). Taken together, there is controversy whether MLL1 is displaced during mitosis and MLL1's impact on cell cycle progression. Therefore we examined the dynamics of MLL1 and Msk1 during the cell cycle, focussing on their association to chromatin as well as the dynamics of key histone modifications associated with the MLL1 complex.



### 3.4.1 Elutriation of LCLs

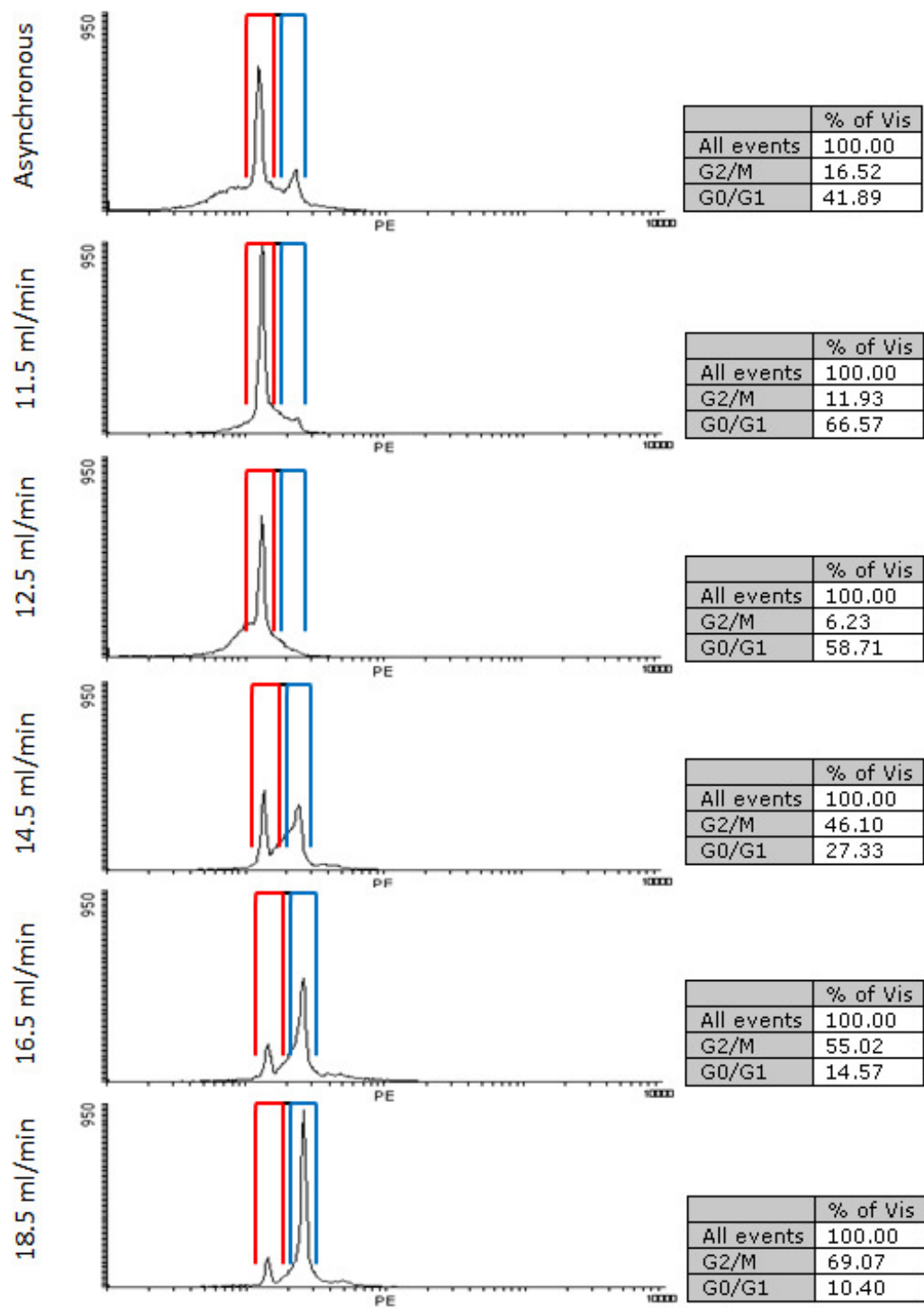
Elutriation is a technique to separate particles, which are in suspension, according to their size. As shown in Figure 3.24 A, a specialized elutriation rotor is needed. In this rotor an elutriation chamber is fitted. The rotor is spun at a constant speed, which is dependent on the average cell size, creating a centrifugal force. Cells are injected in a constant flow of buffer, which can be adjusted by altering the pump speed. The cells thus move with the fluid through the elutriation chamber, creating a counter flow to the centrifugal force (Figure 3.24 B). The cells initially enter the chamber with a lower flow rate, which loads them into the chamber (Figure 3.24 B). In this stress field of two counteracting forces the cells are aligned, with the bigger cells accumulating at the bottom of the chamber and the smaller ones towards the exit. By increasing the pump speed, the flow rate of the buffer can be increased and this drives the cells out of the chamber and different fractions can be collected (Grant & Morrison 1979; Banfalvi 2011).

In the different stages of the cell cycle the cells vary in size (e.g. G1 cells are smaller than G2 cells), therefore elutriation can be used to isolate fractions of cells at different stages of the cell cycle (Kauffman et al. 1990). However, the shape of the cells can influence their streaming behaviour. For this reason, the elutriation buffer contains a relatively low salt concentration (0.75 x PBS), which causes the cells to swell up and adopt a rounder morphology equalizing irregularities on the cell surface. The elutriation buffer also contained BSA and EDTA to prevent cells sticking to each other, or the walls of the elutriation system. In this study we used LCLs, a human lymphoblastoid cell line with a normal karyotype, that grows in suspension and can be cultured in large volumes.



**Figure 3.24: The elutriation system.** (A) During elutriation cells are separated according to their size, density and shape. The separating chamber is fixed in a centrifuge and the cells are pumped through it with a stream of buffer that runs in the opposite direction of sedimentation. (B) Two opposing forces (the centrifugal force and the counter flow) create a force field in the separating chamber, in which the cells align according to their characteristics. Generally smaller cells will sediment at the front and bigger ones towards the back of the chamber. By increasing the pump speed (=counter flow) cell fractions can be eluted from the chamber in a controlled manner (The JE-5.0 Elutriation System, Beckman-Coulter Handbook).

LCLs were loaded into the chamber at a pump speed of 11 ml/min. Fractions of 200 ml were collected and then the pump speed was increased. In total five fractions were collected at a pump speed of 11.0, 12.5, 14.5, 16.5 and 18.5 ml/min. These were kept on ice during collection, before subsequent analysis by FACS analysis to confirm that the cells were in the expected cell cycle stage. In Figure 3.25, a representative FACS analysis of one elutriation run is shown. In the top panel, a FACS profile of unsorted LCLs is shown (Asynchronous). In this sample ~42% of the cells were in G1 (red box) and only ~16.5% were in G2/M (blue box). In the first collected fraction (Fraction 11.5 ml/min, Panel 2) ~67% of the cells were in G1 and ~12% in G2/M. With increased pump speed the number of G1 cells decreases and the number of G2/M cells increases. In the last collected fraction (Fraction 18.5 ml/min, Panel 6) only ~11% of the cells were in G1 and ~70% of the cells were in G2/M. For further analysis the 11.5 ml/min (G1) and the 18.5 ml/min (G2/M) fractions were collected and compared to unsynchronized LCLs. In the initial experiments the LCLs were untreated. However, when examining the G2/M phase under the microscope, only ~10 % of the cells in the fraction were actually in M-phase. In order to increase the proportion of cells in mitosis, LCLs were treated with colcemid for 3 hours prior to sorting, which increased the proportion of cells in M-phase to ~30% (Data not shown).



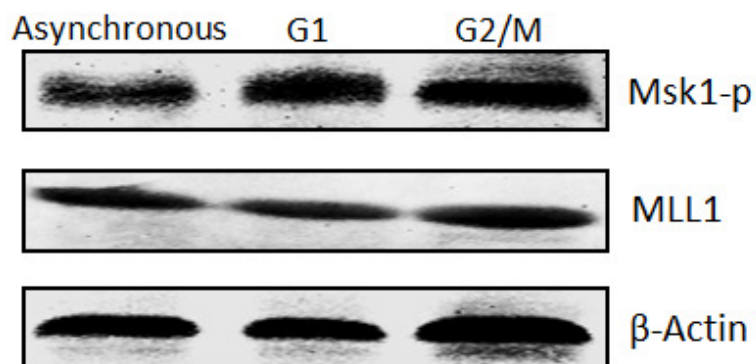
**Figure 3.25: Representative FACS analysis of elutriated cells.** The collected cell fractions were dependent on the pump speed (11-18.5 ml/min). Cells were treated with propidium iodide, which stains the DNA quantitatively, and FACS analysed. Graphs indicate cell number (y-axis) against PI fluorescence/cell ploidy (x-axis). The cell-cycle profiles are shown, with G0/G1 cells gated in the red box and G2/M cells in the blue box. In the table the proportion of cells in these cell cycle stages are shown. A sample of cells was taken before elutriation (Asynchronous) for reference. Generally the cell fraction harvested at 11 ml/min was used as the G1 sample and the cell fraction harvested at 18.5 ml/min as the G2/M sample.

### 3.4.2 Analysis of MLL1 and Msk1

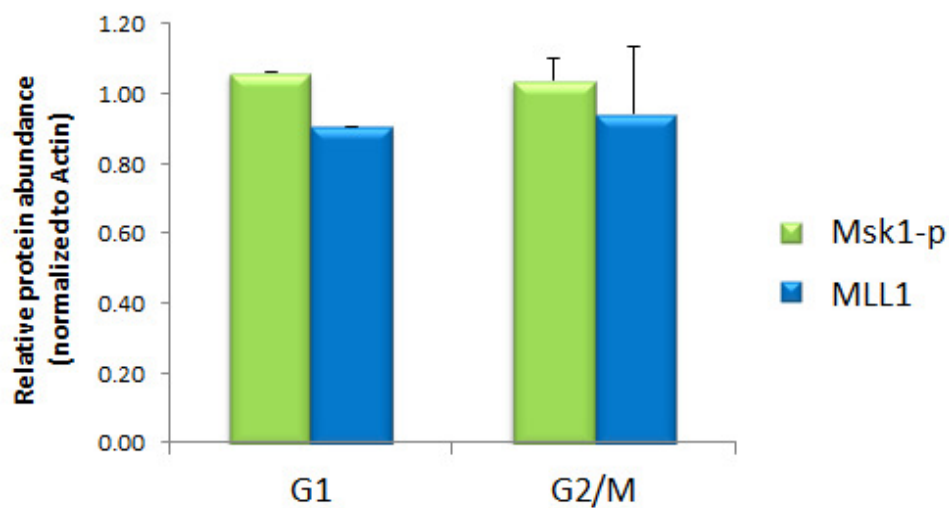
In the first part of the analysis, fractions were collected and cell lysates prepared for western analysis, to examine cell-cycle dependent changes in MLL1 or Msk1 abundance. In Figure 3.26 A, a representative western blot is shown. We loaded an asynchronized sample, a G1- and a G2/M-enriched sample and subsequently probed for MLL<sup>N</sup>, Msk1-p and  $\beta$ -Actin, which acted as a loading control. However, no significant changes were detected. In 3.26 B the quantitated and normalized results of three experiments are shown, confirming that both proteins had stable levels throughout the different stages of the cell cycle.

In a second part of this analysis, we examined the association of the two proteins with chromatin at different stages of the cell cycle. We used immuno-fluorescence microscopy to visualize the binding of these proteins in cells in G1 (Figure 3.27) and on mitotic chromosomes (Figure 3.28). In order to eliminate non-specific binding of the secondary, fluorescence-labelled antibody, we performed a no-primary antibody control (Figure 3.27 A and 3.28 A). No non-specific signal was detected either on cells in G1, or on mitotic chromosomes. In G1 cells, MLL1 and Msk1 show a strictly nuclear location, as their binding corresponded to the counter staining of the DNA with DAPI. Both proteins bind throughout the nucleus with no specific prevalence. A similar situation was observed on mitotic chromosomes. MLL1 and Msk1 bind with an 'overall' coverage of the chromosomes, indicating that the two proteins remain associated with the chromatin, even at its most condensed form during mitosis. The red and green stained patches were artefacts, as cell debris tends to absorb the secondary antibodies and those patches were not observed on all mitotic chromosomes.

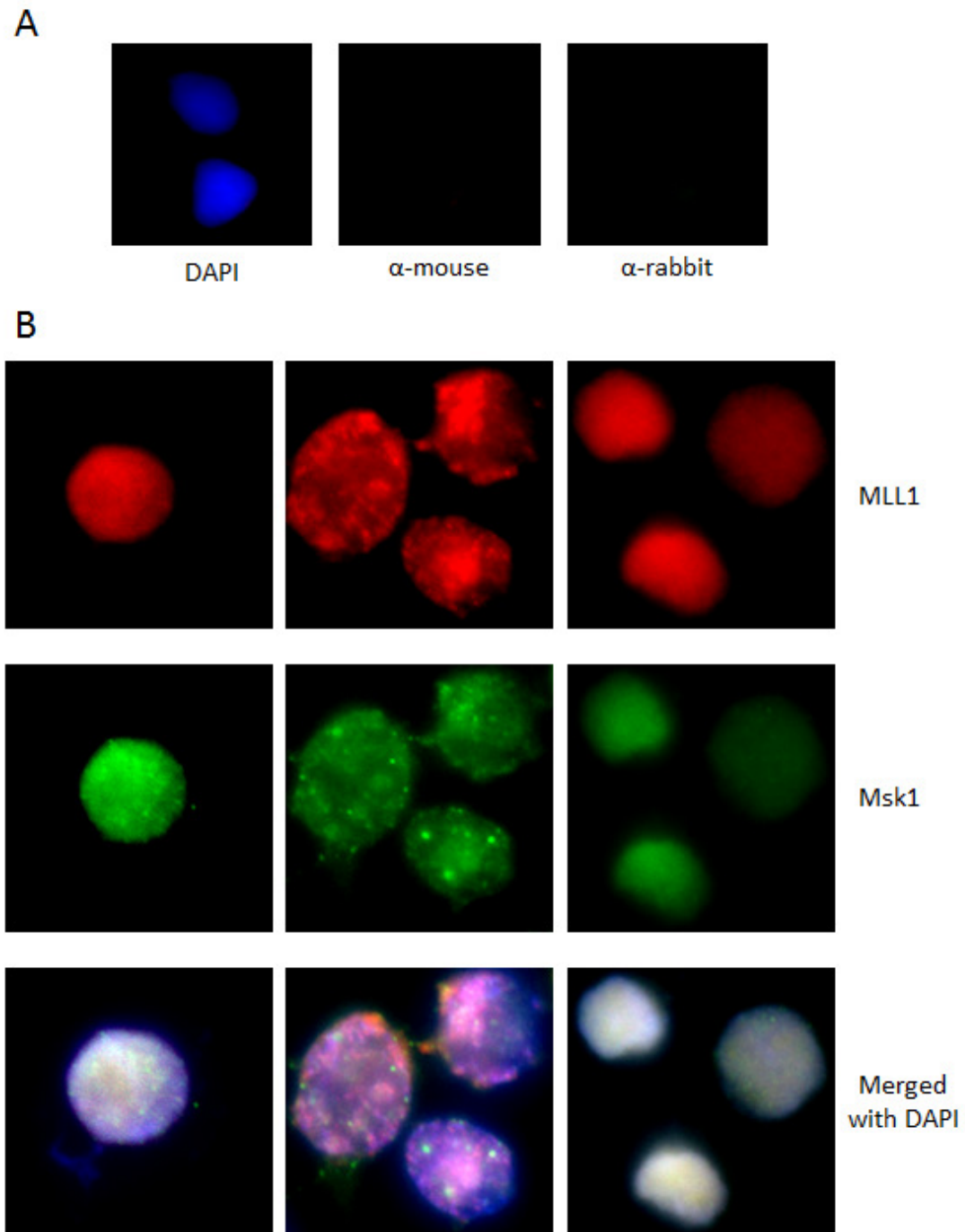
A



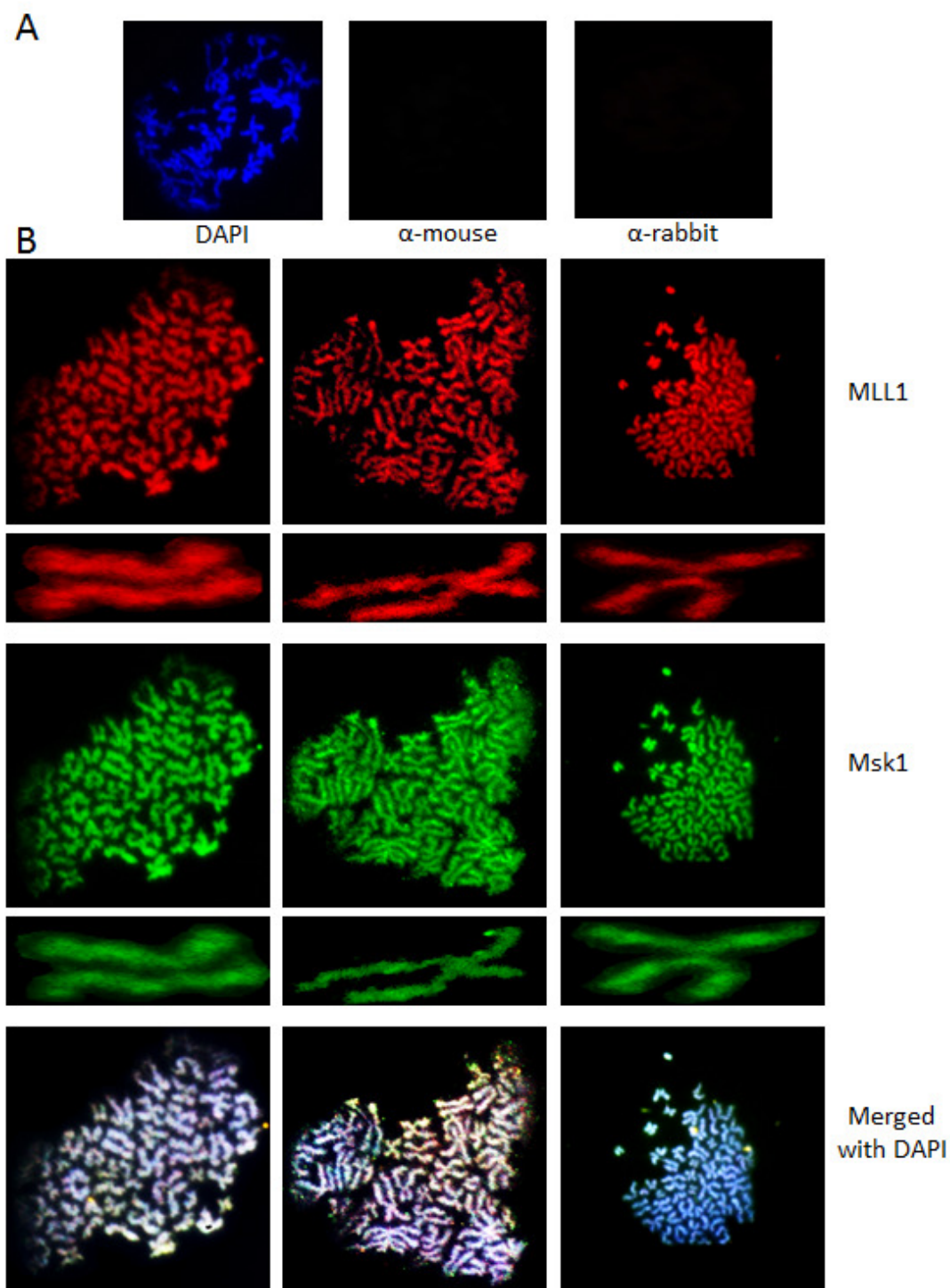
B



**Figure 3.26: MLL1 and Msk1 abundance during the cell cycle.** (A) Whole cell lysates of cell-cycle enriched fractions were analysed by western. The western was probed for MLL1, Msk1-p and for loading control with  $\beta$ -Actin. Representative blots are shown (B) The experiment was repeated three times and the westerns were quantified. No significant differences in the two cell cycle stages could be detected for MLL1 or Msk1. Significant changes are marked with asterisks (paired t-test,  $p < 0.05$  \*,  $p < 0.01$  \*\*,  $p < 0.001$  \*\*\*).



**Figure 3.27: Nuclear distribution of MLL1 and Msk1 in G1 phase.** (A) Secondary antibody specificity control. DNA was counter-stained with DAPI (left). Cells in G1 phase were stained with the secondary  $\alpha$ -mouse IgG (red, middle) or  $\alpha$ -rabbit IgG (green, right) antibody to check for non-specific cross reactions. (B) Cell cycle specific fractions were collected and prepared for immuno-fluorescence microscopy. Red=MLL1, Green=Msk1, Blue=DAPI (100 x magnification).

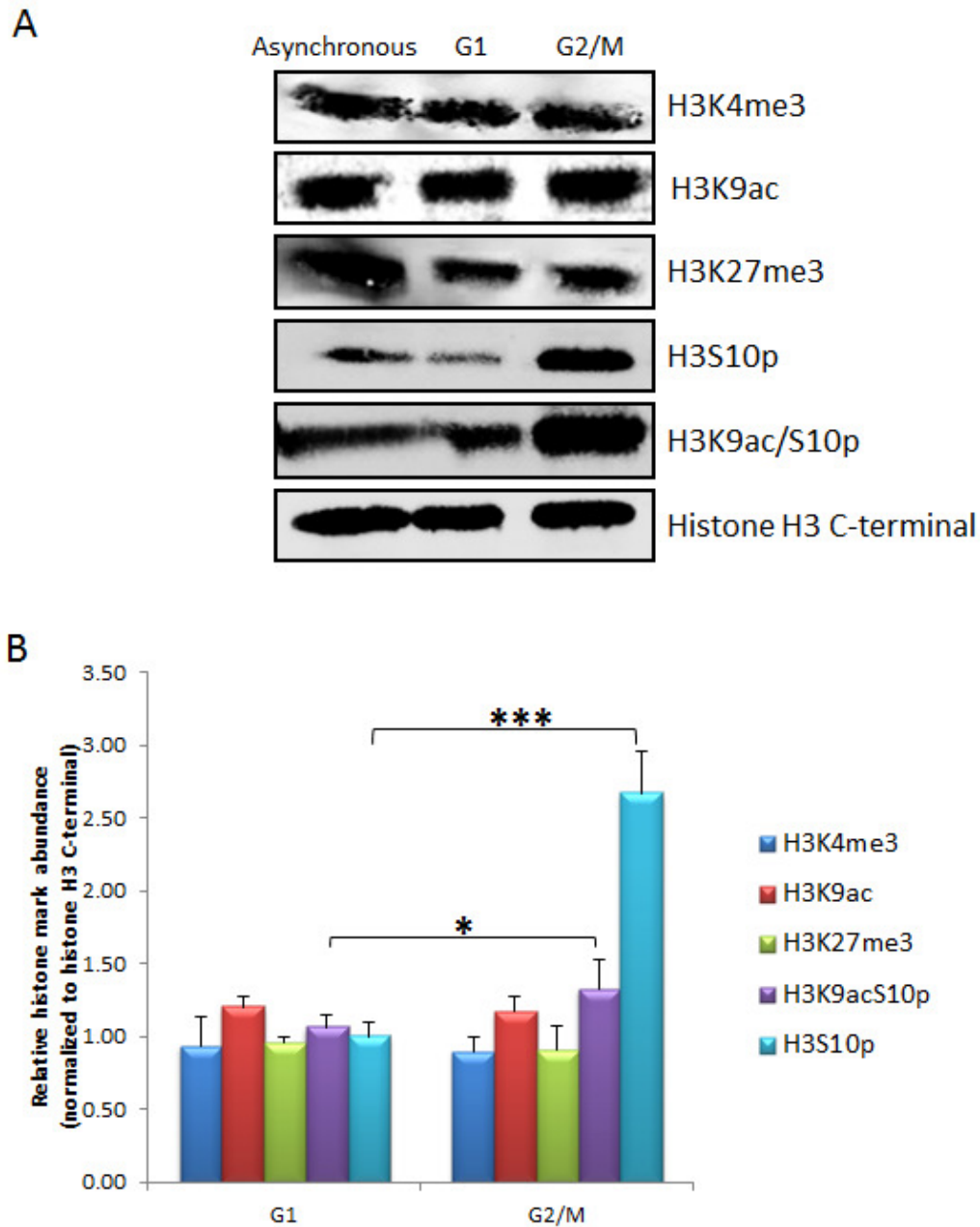


**Figure 3.28: MLL1 and Msk1 distribution on mitotic chromosomes.** (A) Secondary antibody specificity control. DNA was counter-stained with DAPI (left). Cells in M-phase were stained with the secondary  $\alpha$ -mouse IgG (red, middle) or  $\alpha$ -rabbit IgG (green, right) antibody to check for non-specific cross reactions. (B) Cells fractions were collected and immediately prepared for Immuno-fluorescence microscopy. Chromosome 1 is shown enlarged. Red=MLL1, Green=Msk1, Blue=DAPI (100 x magnification).



### 3.4.3 Analysis of histone modifications

In a subsequent analysis, we examined the global changes in histone modification abundance involved in MLL1 function in the different stages of the cell cycle. Histones were acid-extracted and analysed by western blotting. We examined the 'active marks' H3K4me3 and H3K9ac, which are deposited by components of the MLL1 complex (Slany 2009a) and analysed H3K27me3, a 'silencing mark', deposited by Polycomb proteins, as PRC2 (polycomb repressive complex 2) is proposed to be activated by dense chromatin (Yuan et al. 2012). Furthermore H3S10p was examined as mitotic chromatin is marked with this modification (Hsu et al. 2000; Hauf et al. 2003) and therefore functioned as an internal control. Finally H3K9acS10p was examined because this modification stimulates the SET activity of MLL1 (Nightingale et al. 2007). Western blots were labelled with an anti-histone H3 C-terminal antibody, which was used as a loading control. Representative westerns are presented in Figure 3.30 A and the results of three individual experiments, which were normalized to an unsorted, asynchronous cell population, are shown in Figure 3.30 B. We found H3K4me3 abundance is comparable in the G1 and the G2/M cells, as were H3K27me3 and for H3K9ac. In contrast, H3K9ac/S10p was significantly elevated in G2/M (~+ 35%, paired t-test:  $p=0.049$ ) and as expected, H3S10p levels were highly elevated in G2/M cells (~+170%, paired t-test:  $p=0.000$ ).

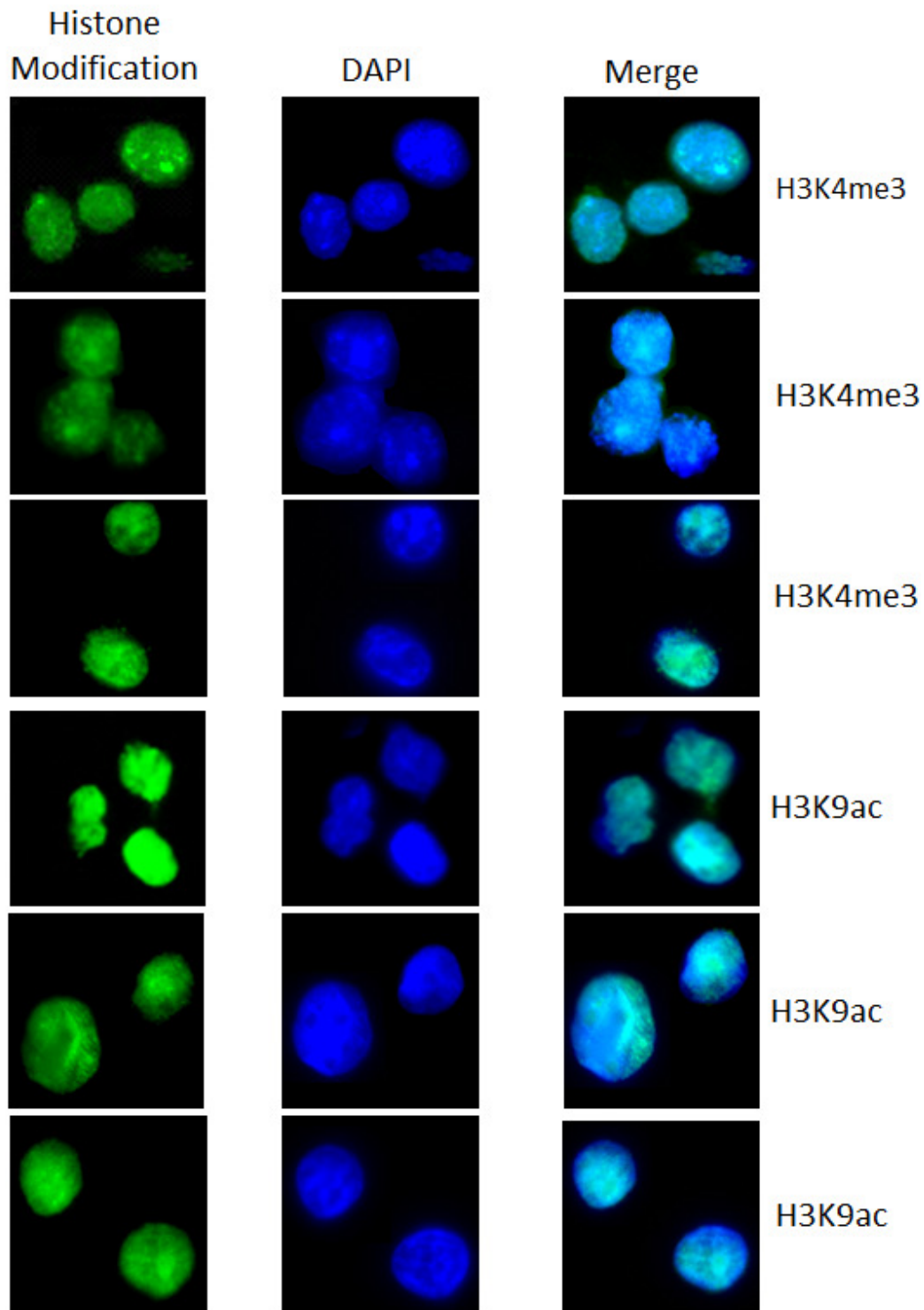


**Figure 3.29: Histone modification abundance during the cell cycle.** (A) Histones were extracted from cell-cycle enriched fractions and analysed by western. The western was probed for key MLL-related histone modifications. Histone H3 C-terminal was used as a loading control. Representative blots are shown. (B) The experiment was repeated three times and the westerns quantified. Most histone modifications showed no differences in the different cell cycle stages. Only H3S10p showed enrichment during G2/M. Significant changes are marked with asterisks (paired t-test,  $p < 0.05^*$ ,  $p < 0.01^{**}$ ,  $p < 0.001^{***}$ ).

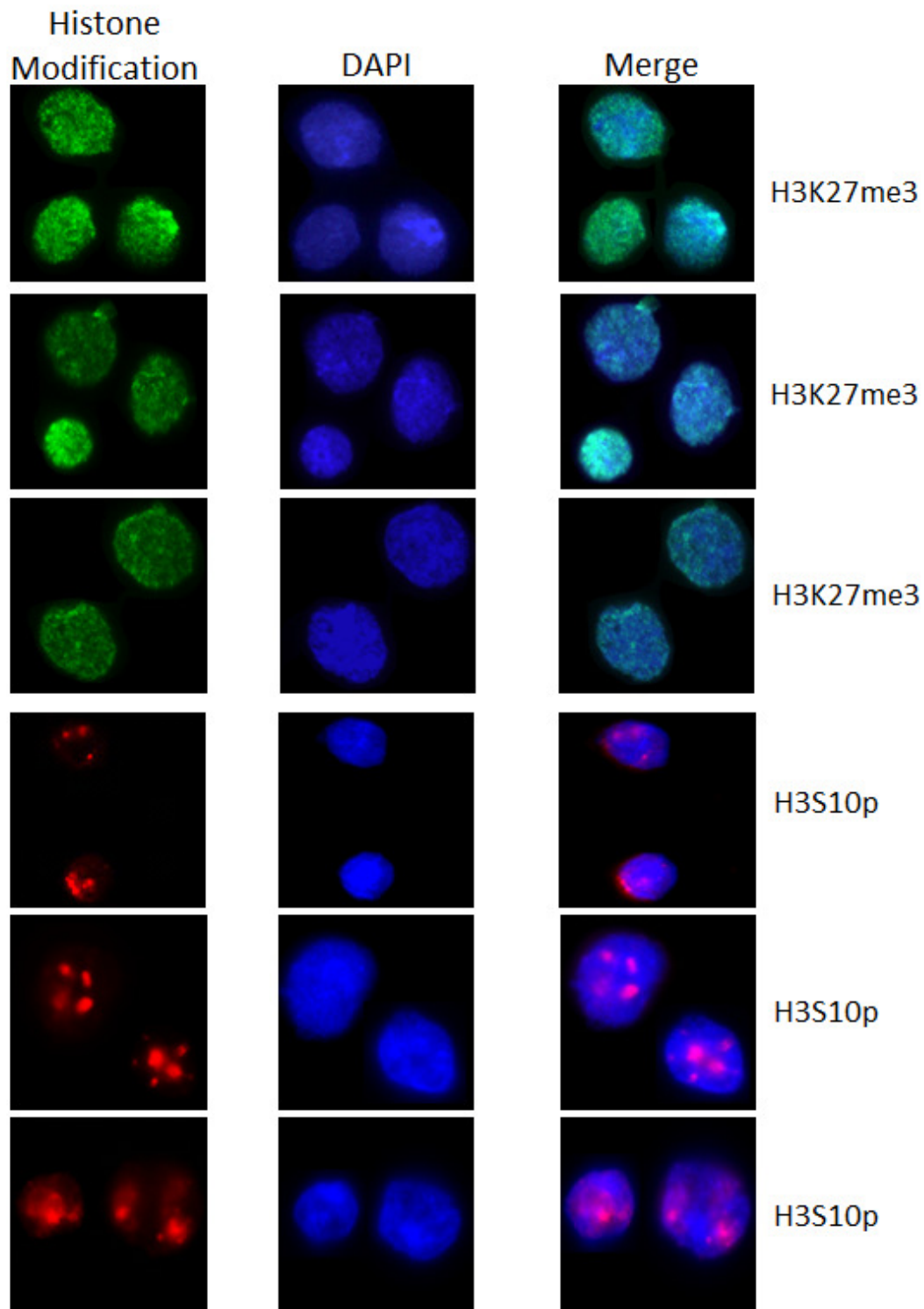
In a first analysis we explored the pattern of histone modification on the chromatin in G1 and mitotic cells using immuno-flourescence-microscopy [Figure 3.30 (H3K4me3 and H3K9ac in G1), Figure 3.31 (H3K27me3 and H3S10p in G1), Figure 3.32 (H3K4me3 and H3K9ac in M-phase) and in Figure 3.33 (H3K27me and H3 S100 in M-phase)].

In G1 phase H3K4me3, H3K27me3 and H3K9ac are associated with the chromatin throughout the nucleus. However, whilst H3K9ac showed an 'all-over' distribution, H3K4me3 and H3K27me3 showed a 'speckled' distribution with binding hotspots that do not correlate with DAPI dense regions. H3S10p was less abundant in the nucleus and showed concentration to a few areas.

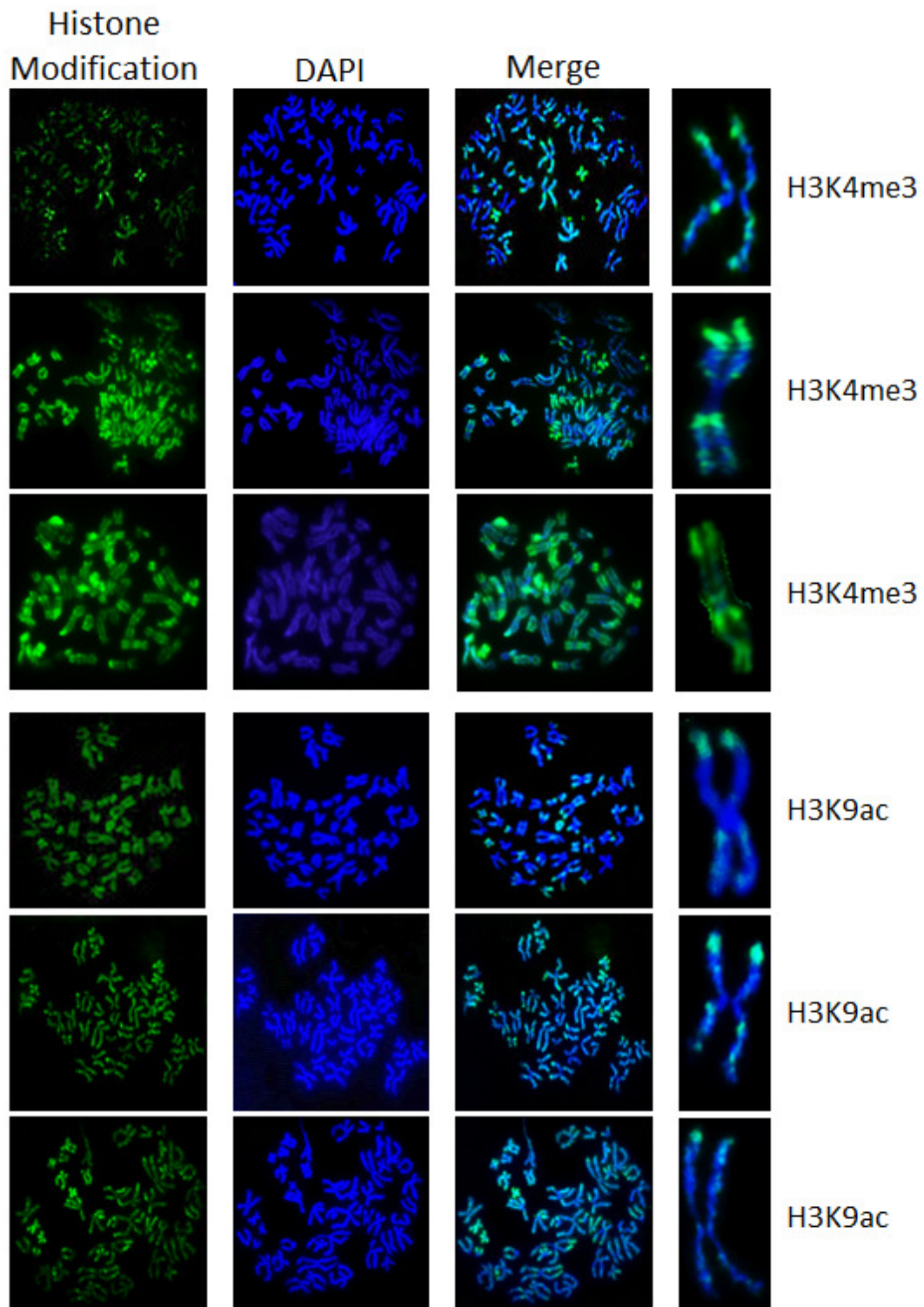
On the mitotic chromosomes H3K4me, H3K27me3 and H3K9ac bind in distinct patterns to the chromosomes, revealing highly methylated or acetylated areas, which have been described in detail previously (Terrenoire et al. 2010). The H3K4me3 and H3K9ac banding pattern seem to correlate. For example on chromosome 1, which is shown enlarged in Figure 3.31, a highly stained pole on the top of the p-arm, a distinctive band on the q-arm right beneath the centromere and diffuse staining on the q-arm for both modifications are observed. In contrast, H3K27me3 showed a distinctively different distribution. The pole of the p-arm is highly stained, several weaker bands on the p- and the q-arm can be observed and only the centromere is depleted completely of H3K27me3. Finally H3S10p is detected all over the chromosome, with highly stained caps on the p- and the q-arm and the centromere.



**Figure 3.30: H3K4me3 and H3K9ac distribution in G1 phase cells.** Cells fractions were collected and prepared for immuno-fluorescence microscopy. Green=indicated histone modification, Blue=DAPI (Magnification: 100x).

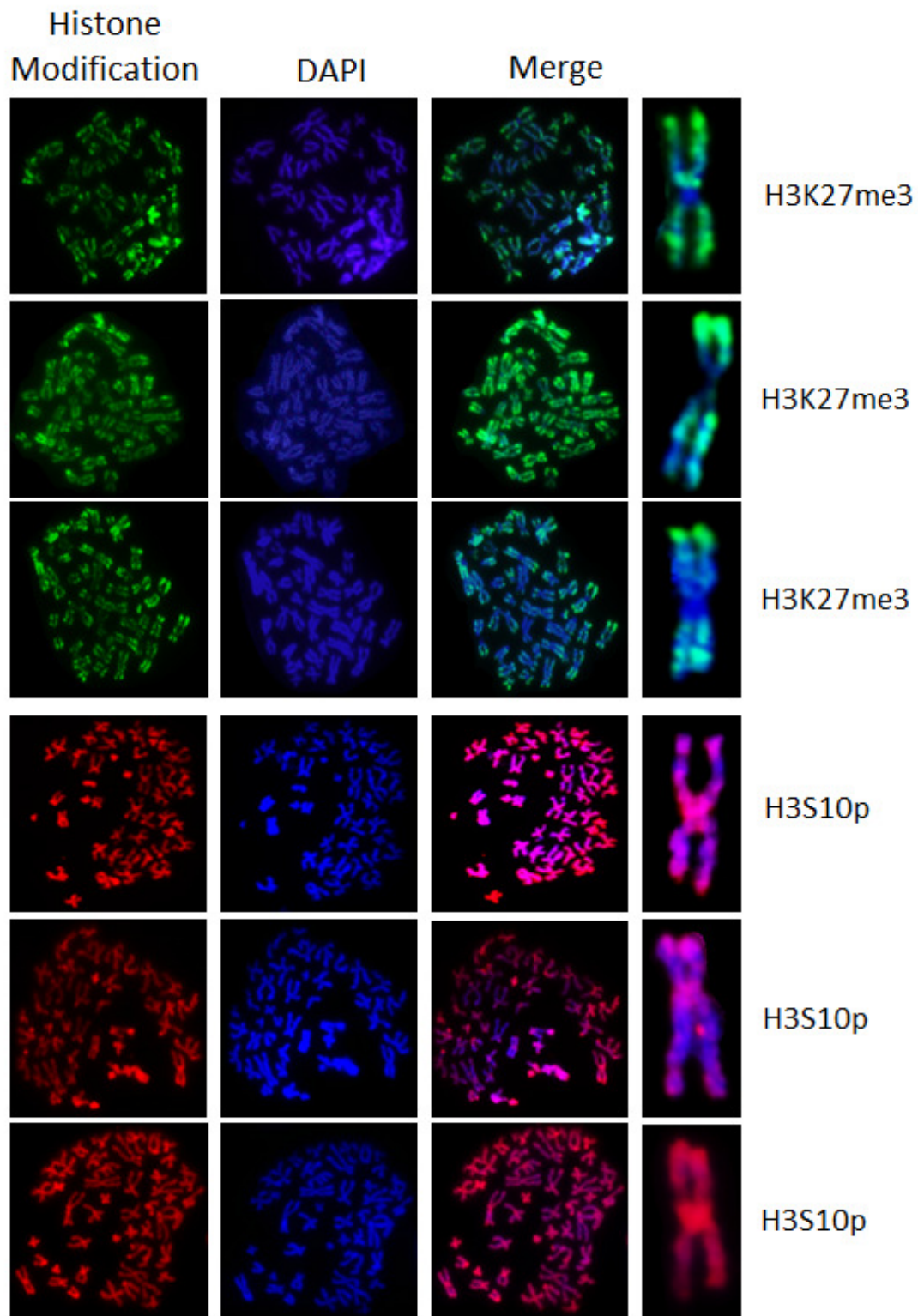


**Figure 3.31: Histone modifications on mitotic chromosomes.** Cells fractions were collected and immediately prepared for Immuno-fluorescence microscopy. Chromosome 1 is shown enlarged. Red/Green=indicated histone modification, Blue=DAPI (Magnification: 100x).



**Figure 3.32: H3K4me3 and H3K9ac distribution on mitotic chromosomes.** Cells fractions were collected and prepared for immuno-fluorescence microscopy. Chromosome 1 is shown enlarged. Green=indicated histone modification, Blue=DAPI (Magnification: 100x).





**Figure 3.33: H3K27me3 and H3S10p distribution on mitotic chromosomes.** Cells fractions were collected and prepared for immuno-fluorescence microscopy. Chromosome1 is shown enlarged. Red/Green=indicated histone modification, Blue=DAPI (Magnification: 100x).

Taken together, we have shown that the MLL1 and Msk1 abundance does not change in G1 or G2/M and is comparable to the protein levels found in unsynchronized cells. Furthermore, MLL1 and Msk1 were associated with chromatin in G1 as well as with mitotic chromosomes, the most condensed form of chromatin. Neither proteins show any local preferences on the chromatin, but bind throughout equally the chromatin in G1 and on the chromosomes. Likewise, H3K4me3, H3K9ac and H3K27me3 do not show any differences in abundance in any stage of the cell cycle, but the distribution pattern changes (from an overall binding to chromatin in G1, to specific banding patterns on chromosomes). H3S10p was noticeably different. It showed a vast upregulation during G2/M on both western blots and microscopy slides. In G1 cells the mark showed localization to discrete areas, while on mitotic chromosomes the mark was found all over them. Interestingly H3K9acS10p was elevated in G2/M cells, but not in G1. However, the antibody was not suitable for microscopy analysis and therefore no conclusion on its distribution in chromatin can be made.



## 4 Discussion

Chromatin and histone modifications play an important role in many physiological, cellular processes, like development (Pengelly et al. 2013), response to external stimuli (Tittel-Elmer et al. 2010) and disease (Burke & Bhatla 2014). It has been established, that certain histone modifications are associated to certain gene activation states. For example acetylation and tri-methylation of histone H3K4 are associated with the active gene state and tri-methylation of histone H3K27 with a repressive gene state (Kouzarides 2007). A large number of proteins are known to deposit ('writer'), remove ('eraser') or recognise ('reader') specific histone modifications, with crucial functions in physiological or pathological contexts (Jakovcevski & Akbarian 2012).

### 4.1 Epigenetic regulation by MLL1

The histone methyltransferase MLL1 (mixed lineage leukaemia 1) is an epigenetic writer protein, which is often mutated in leukaemia, especially in childhood acute myeloid leukaemia (AML) (Krivtsov & Armstrong 2007). Studies with MLL1<sup>-/-</sup> mice have shown that the homozygous knock-out of MLL1 is embryonically lethal, making MLL1 an essential epigenetic modifier (Hess et al. 1997). The MLL1 protein functions within a large multi-protein complex, which regulates the degree of histone H3K4 methylation deposited at target genes (Patel et al. 2009), however it is not yet known how MLL1 is targeted to its sites (Ansari & Mandal 2010). The SET domain of MLL1 is primarily a mono-methyltransferase and only in collaboration with the conserved group of WARD proteins (WDR5, Ash2L, RbBP5 and

DPY-30) H3K4me1 can be sequentially catalysed further (Dharmarajan & Cosgrove 2009). Although these proteins form the MLL1 core complex and are essential for proper histone H3K4 methylation in vitro, many more proteins (~30) have been associated with the MLL1 complex (Nakamura et al. 2002). The complete MLL1 complex acts on chromatin by methylation and acetylation (Slany 2009). In MLL1-associated leukaemias the MLL1 protein is disrupted by chromosomal translocation, which leads to the formation of fusion-proteins and a disruption of the MLL1 protein complex (Ford, 1993). The C-terminus of MLL1 is lost and with it the SET domain, resulting in the a mis-regulation of MLL1 target genes (Ayton & Cleary 2003). Well known MLL1 target genes are the Hox genes, which are key transcription factors during development and haematopoiesis (Hess et al. 1997; Jude et al. 2007).

MLL1 is a mammalian Trithorax protein, which in concert with Polycomb proteins regulate their target genes by methylation of chromatin. Trithorax proteins deposit the 'activating' H3K4me3 mark and Polycomb proteins deposit the 'silencing' H3K27me3 mark, antagonizing each other's function (Steffen & Ringrose 2014). Here we examine the role of Msk1, a histone kinase, as a potential member and co-regulator of the MLL1 complex.

## 4.2 Is the Histone kinase Msk1 involved in MLL1 regulation?

Msk1 (mitogen-and-stress-induced kinase1) is a downstream kinase of the MAPK pathway (Deak et al. 1998), which phosphorylates either transcription factors or histones (Duda 2012). It has been shown, that Msk1 is involved in the activation of immediate-early (IE) genes (Thomson et al. 2001) and it has been suggested, mainly based on work in yeast, that MAPKs might be functional components of transcriptional complexes, which might recruit additional factors to chromatin (Alepez et al. 2001; Thomson et al. 2001).

The findings, (I) that phosphorylation stimulates MLL1's SET activity (Nightingale et al. 2007), (II) that Msk1 induced phosphorylation lead to a loss of Polycomb-mediated gene silencing (Lau & Cheung 2011), (III) that Msk1 is involved in the activation of IE genes (Drobic et al. 2010) and (IV) that in yeast, MAPKs can be part of transcriptional complexes (Edmunds & Mahadevan 2004), lead to the hypothesis, that Msk1 is involved in the regulation of MLL1 activity.

Here we tested this hypothesis. In the first set of experiments the interaction of MLL1 and Msk1 was explored. These experiments were complemented with an analysis of the genomic distribution of MLL1 and Msk1, followed with a third set of experiments, which examined the functional impact of MLL1 and Msk1 on gene activity. A final experiment examined changes in MLL1 and Msk1 abundance and key histone modifications during the cell cycle.

### **4.3 Do MLL1 and Msk1 physically interact?**

MLL1 forms a multi-subunit complex and many proteins have been found to interact with it (Nakamura et al. 2002). One possibility how Msk1 might contribute to MLL1 regulation is by directly interacting with MLL1 and even be a component of the complex. To address this, co-immunoprecipitation experiments were carried out. However, both proteins are known to bind to chromatin. Therefore it needed to be considered that MLL1 and Msk1 might not interact directly, but might bind to adjacent sites in chromatin. Initial experiments were carried out in a human cell line. It showed that MLL1 and the phosphorylated, active form of Msk1 can be pulled down together (Figure 3.2). In order to release chromatin-bound proteins from the DNA, ethidium bromide was added to the lysis buffer but the proteins still efficiently co-immunoprecipitated. This indicated a direct interaction between MLL1 and Msk1, which was independent of chromatin. However, pull-downs with MLL1 and non-phosphorylated Msk1 were not as conclusive. Although Msk1 was detected in the MLL1 pull-down, MLL1 was not detected in the Msk1 pull-down. This can be interpreted in two ways: It might be that MLL1 interacts preferably with the activated Msk1, but not with the unphosphorylated, inactive form. It is also possible that this specific epitope is important for MLL1/Msk1 interaction, which the bound antibody interrupts. In order to extend these findings co-IP experiments were repeated in other cell types (human and mouse), which confirmed the interaction of MLL1 and phosphorylated Msk1. A potential problem with co-IP experiments is the non-specific protein binding to the IP antibody and/or the beads. Although the cell lysate was pre-cleared and the beads were washed with higher ionic strength buffer to disrupt non-specific binding, some back-ground has to be expected. We

observed antibody contamination, showing up as a ~25kDa (light chain) or a ~50kDa (heavy chain) on a SDS gel, as seen in Figure 3.2, that obscure the results.

To examine the question from another angle, FLAG-tagged MLL1 was over-expressed in HEK293 cells and then purified using beads, to which anti-FLAG-antibodies were cross-linked. The FLAG-tag is a small peptide (DYKDDDDK), which was engineered not to disturb the protein function of the tagged protein and to prevent non-specific binding. Here the FLAG-tag was fused to the N-terminus of MLL1 and introduced into the cells using a gentle transfection method (no cell death occurred). Cells were lysed and the protein precipitated, washed with low-ionic strength buffer and eluted with 3xFLAG peptide, to ensure the integrity of the MLL1 complex. Complex purity was characterized by gel-filtration. Although the sample contained a lot of small protein contamination (Figure 3.5 C), a single peak in the FLAG-prep (Figure 3.5 C, indicated with \*) was observed, corresponding to a large molecular weight complex. This peak corresponded to a complex of several large molecular weight proteins (Figure 3.5 A, indicated with arrow). For western analysis the complex was run on a 4-12% gradient gel to obtain better separation of the complex components. The membrane was probed with anti-FLAG, anti-Msk1p and anti-NFκB antibody. NFκB (a strong interaction partner of Msk1 (Vermeulen et al. 2003)) was only found in the input suggesting that the protein complex did not contain non-specific proteins. FLAG-MLL 1 protein and the phosphorylated Msk1 were also found in the complex, suggesting that Msk1 is a stable member of the MLL1 complex. However, to confirm this, a large amount of MLL1-complex needs to be purified and separated over the gel-filtration column, so that components can be characterized by mass-spectroscopy. Furthermore, these results do not reveal if Msk1 is a permanent member of the MLL1 complex or if this is a transient interaction.

#### **4.4 Do MLL1 and Msk1 bind to the same genomic sites?**

The first set of experiments indicated that MLL1 and phosphorylated Msk1 interact with each other and that Msk1 might be a member of the MLL1 complex. This suggested that the two proteins would be found at the same genomic sites. We addressed this directly by cross-linked chromatin immunoprecipitation. Proteins were cross-linked to chromatin and MLL1 and phosphorylated Msk1 were immunoprecipitated. The bound DNA was inserted into libraries for high-throughput sequencing on the Illumina Next-generation HiSeq2000 Sequencing platform. Direct comparison of the distribution of MLL1 and Msk1 binding showed an overlap of MLL1 and Msk1 binding sites at a third of binding sites (~35%). These binding sites were compared with H3K4 tri-methylation of histones, which showed a small correlation of H3K4me3 with MLL1 and Msk1 (~2% overlap). A trend-plot over transcription start sites (TSS) revealed a peak of MLL1 and Msk1 over the TSS, substantiating the proposition that Msk1 is a member of the MLL1 complex. The scatterplot analysis, which compared MLL1 and Msk1 binding strength at their shared binding sites, showed a strong correlation for their binding signals, which strengthens the hypothesis

Functional annotation analysis of these genes with DAVID did not reveal unique groups of genes. Only three groups (Translation, 3 counts, 0.37 fold enrichment; Regulation of transcription, 4 counts, 0.21 fold enrichment; Phospho-protein, 10 counts, 0.17 fold enrichment) were flagged, but they all had quite low enrichments and a high false-discovery-rate (FDR ~32) and were therefore not statistically significant.

Nevertheless, it was shown that MLL1 and Msk1 bind to a number of sites in the genome and can be found at a subgroup of binding sites. These findings support the hypothesis that Msk1 can be a member of the MLL1 complex.

## 4.5 What is the functional impact of MLL1 and Msk1 interaction?

Having established an interaction between MLL1 and Msk1, the functional impact of this interaction was examined by knocking these regulators down *in vivo*. A transient knock-down in primary cells was chosen, as they would best represent the early changes in gene transcription in a physiological context. This was also the only way to do the experiment as the transfected cells started to die at around day 2 post-transfection and stable knock-downs in CCEER (embryonic stem cell line) for MLL1 or Msk1 were not viable. However, while the Msk1 knock-down resulted in substantial reduction in protein levels (reduced by ~99%), MLL1 levels were not reduced by more than ~63%. Even after three days post knock-down, when ~70% of the transfected cells were dead, MLL1 levels remained stable, indicating that lower levels of MLL1 cannot be tolerated. Therefore cells with partial knock-down of MLL1 were used for further experiments.

Initially we examined the impact of knock down at known MLL1 target genes. The HoxA genes were chosen as they are important genes in development and haematopoiesis and are known to be de-regulated in leukaemia. This showed a significant reduction in gene expression for most HoxA genes, especially those in the middle of the cluster (HoxA 2-11), an effect seen in both knock-downs. These findings suggest that both MLL1 and Msk1 contribute to HoxA gene regulation. It is known that Hox genes are regulated in clusters and that genes in a cluster can cross-regulate each other, leading to the effect of collinearity (Montavon & Duboule 2013). Apparently MLL1 and Msk1 binding and H3K4 methylation at a few sites in the HoxA cluster is sufficient to regulate the majority of the cluster. The expression of the oncogene Meis1, which is often mis-regulated in leukaemias was also

examined, but no changes in expression were found, suggesting that MLL1/Msk1 do not regulate this gene.

As we have shown previously, MLL1 and Msk1 bind to many sites in the genome, therefore in the next experiment the global impact of MLL1 and Msk1 on gene regulation was examined using microarray analysis. Although 3912 genes changed their expression in MLL1 or Msk1 knock-down cells, only 11% of these genes were down-regulated in both MLL1 and Msk1 knock-down cells and only 19% were up-regulated together. The majority of genes showed different responses to MLL1 or Msk1 knock-down. This could have a number of reasons. (I) Msk1 modifies chromatin, but also transcription factors and is known to bind to other proteins, like NFκB or CREB (Arthur 2008). It is therefore likely that Msk1 knock-down would also functionally impact other genes beyond MLL1 targets. In fact 63% of the responding genes change their expression only in the Msk1 knock-down. (II) The MLL1 protein levels were only reduced by about 60-70%. As a certain level of MLL1 seems to be required for cell survival, it could be that the effect on MLL1 target genes is buffered. (III) Finally a number of MLL1 and Msk1 homologues exist in cells. It is possible that these proteins are redundant with MLL1 and Msk1 and permit cell functionality.

Nevertheless, these results showed that down-regulation of MLL1 and Msk1 lead to the same response in a significant number of genes, indicating a coordinated action of the two proteins. However, the fact that the majority of gene expression changes occurred only in one knock-down cell type (63% of genes only changed in Msk1 knock-down cells) suggest that MLL1 and Msk1 do not have an exclusive interaction. This finding is consistent with the results of the CHIP-Seq experiments which showed that MLL1 and Msk1 share ~35% of their genomic binding sites.



As MLL1 and Msk1 are both chromatin modifiers, the impact of their knock-down on key histone modification on MLL1 target genes was examined. HoxA4 and 5 were chosen because they showed a substantial reduction of transcription levels in knock-down cells (Figure 3.15) and we reasoned this would make the detection of changes in histone modifications easier. Primers were designed to cover three areas: a control area 5' to the gene, the TSS and a region 3' at the end of the gene. Levels of histone modification in knock-down cells were compared to mock-transfected cells. Although the changes were subtle, several trends could be detected, some of which were statistically significant. We found H3K4me3 and H3K9ac/S10p levels would drop at TSS in knock-down cells, whereas H3K27me3 levels would rise. These findings showed the impact of MLL1 and Msk1 gene regulation at an epigenetic level. The reduction of HoxA4 and HoxA5 was accompanied by the reduction of H3K4me3 (deposited by MLL1) and H3K9ac/S10p (deposited by Msk1 and CBP, a member of the MLL1 complex). As a consequence the levels of H3K27me3 (deposited by Polycomb proteins counter-balancing Trithorax proteins) showed a trend to increase. This transcriptional ON/OFF switch by Trithorax and Polycomb protein has been described for other target genes as well, for example at *Drosophila's Ubx* gene (Papp & Müller 2006). It was found that in the repressive state the *Ubx* gene was bound by Polycomb proteins and tri-methylated at H3K27, whereas in the active state this mark was absent and Trithorax protein binding at the promoter was observed in concert with H3K4 tri-methylation at this region (Papp & Müller 2006).

## 4.6 Do MLL1 and Msk1 change during the cell cycle?

The sequencing and microarray experiments indicated that a number of genes, which are bound by MLL1 and Msk1 and change their expression in knock-down cells, were associated with the cell cycle (Figure 3.21). Furthermore, previous reports associated MLL1 the cell cycle. However there has been controversy concerning how MLL1 is regulated. Data suggest that MLL1 is either evicted from chromatin during mitosis (Ennas et al. 1997, Mishra et al. 2009) or that it stays on chromatin throughout the cell cycle stages were it was (Caslini, Shilatifard, et al. 2000), and having an active role in cell-cycle regulation (Milne et al. 2005, Liu et al. 2007, Blobel et al. 2009). We therefore decided to examine MLL1 and Msk1 protein abundance in different cell cycle stages, as well as their association to chromatin. In parallel, key histone modifications associated with MLL1 function were examined. In order to separate cells in different stages of the cell cycle, elutriation was used, a method, which separates cells according to their different size and shape at different cell cycle stages. This method has the advantage that cells were treated the same way for the same time.

We found that on a global level MLL1 and (phosphorylated) Msk1 protein levels in G2/M and G1 were comparable to an unsorted cell population (Figure 3.5.3). We subsequently analysed cells at a single cell level for MLL1 and Msk1 nuclear occupancy. In G1 the two proteins were found throughout the nucleus, indicating even binding of MLL1 and Msk1 throughout chromatin, which was in accordance with our distribution data. The most condensed form of chromatin, metaphase chromosomes during mitosis, were examined and showed that MLL1 and Msk1 binding throughout the DNA as well. This indicated that MLL1 and Msk1 are not excluded from the chromatin during mitosis, however, the resolution of the microscope was not high enough to distinguish if MLL1 and Msk1 remain bound to the

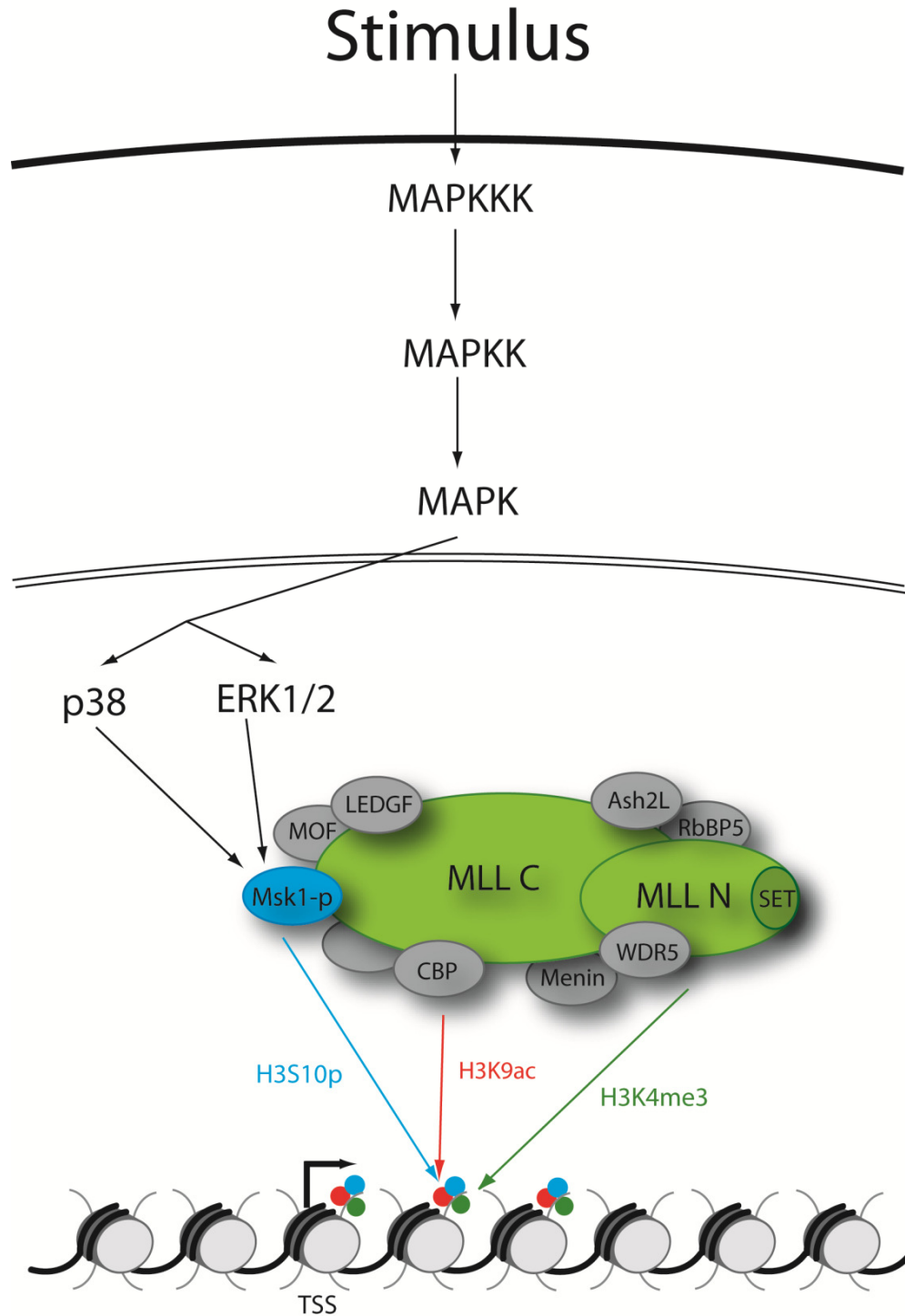
chromatin or if they align on the outside of the chromosome in the perichromosomal layer (chromosome periphery). It has been proposed that this layer may act like a skin protecting the chromosome surface and is enriched in nuclear proteins (Ohta et al. 2011).

We then examined the key histone modifications associated to MLL1 function. At a global level, H3K4me3 did not show any differences in G1 and G2/M compared to unsynchronized cells (Figure 3.5.6). Neither did H3K27me3. H3K9ac levels were slightly higher in both G1 and G2/M, presumably due to gene activity in these cell populations (~60% of G2/M were in G2 phase and H3K9 acetylation is a general mark of active transcription). In contrast, H3S10 phosphorylation levels were highly enriched in G2/M (~+170%). This is consistent with the known H3S10p function as a marker of mitotic chromatin (Hsu et al. 2000) and therefore functioned as an internal control, as significant changes in histone modifications could be detected although only ~30% of the cells in the G2/M population were in mitosis. The dual mark, H3K9acS10p, showed no enrichment in G1 cells, but in the G2/M population an enrichment of ~35% was observed. When the histone modifications were examined at a single cell level by fluorescence microscopy, these showed a shift from an overall distribution for H3K4me3, H3K27me3 and H3K9ac with some hotspots in G1 cells, to a banding pattern in mitotic chromosomes that has been described before (Terrenoire et al. 2010). In contrast, whereas H3S10p was only found in few areas in G1 cells, mitotic chromosomes showed an all-over distribution (Figure 3.32 and 3.32). No data were generated for H3K9acS10p as the antibody was not compatible with immunofluorescence. In summary we have shown that MLL1 and Msk1 binding is stable throughout cell cycle and that both stay closely associated to chromatin during mitosis. Although H3K4me3 levels were not changed in G1 or G2/M, H3K9acS10p levels were up-regulated in mitosis,

suggesting this may have a potential role for MLL1 to restart transcription after mitosis. This is consistent with proposals that MLL1 functions as a 'mitotic bookmarker' (Kadauke & Blobel 2013), as MLL1 is bound to chromatin during mitosis and shows a mitosis-specific shift in genomic occupancy of MLL1, in favour of genes highly expressed in interphase. Those changes should promote rapid transcriptional activation to facilitate M/G1 transition (Blobel et al. 2009) . Our results support this hypothesis and suggest that Msk1 might deposit the H3K9ac-associated H3S10p mark, which stimulates MLL1's SET activity and therefore promote transcriptional activation.

#### **4.7 How do the results support our hypothesis?**

The aim of this thesis was to examine whether Msk1 contributes to MLL1 function. As discussed, the data indicate that Msk1 is a potential member of the MLL1 complex. Not only were they precipitated together, they also correlate to sub-set of genomic sites on chromatin. It was further shown that MLL1/Msk1 occupy many sites independently and supported by expression studies, which showed that a subset of genes are regulated by both MLL1 and Msk1, but the majority of genes are not, suggesting that the MLL1/Msk1 interaction is not exclusive and that Msk1 regulates genes independently of MLL1. Moreover, we show that MLL1 and Msk1 are likely to regulate gene expression by depositing histone marks (H3K4me3 by MLL1 and H3K9ac/S10p by Msk1). Finally the elutriation experiments indicated that the MLL1 complex (with Msk1) remains associated with chromatin during the cell cycle and that, although H3K4me3 is stable throughout the cell cycle, the H3K9ac/S10p mark is elevated during mitosis, offering a possible mechanism of how MLL1 might promote gene reactivation at the mitosis-interphase change.



**Figure 4.1: Msk1 is a member of the MLL1 complex.** Our results indicated that Msk1 can be a member of the MLL1 complex. We suggest that Msk1 deposits the H3S10p mark, which together with the H3K9ac mark stimulates H3K4me3 depositing by MLL1. Our data showed that in a cell cycle context this combination of marks may help exit of mitosis, supporting the hypothesis that MLL1 is a bookmarker during mitosis. This suggests that MLL1 activity can be directly regulated in response to MAPK signalling.

## 4.8 Perspectives

So far it is not completely understood how MLL1 is regulated. However, given the connection of MLL1 with leukaemia, this is an essential question that needs to be addressed. In the progress of understanding MLL1 function we found that MLL1 binds to many genomic sites (~4000 sites, Chapter 3.2.2). This finding is consistent with previously reported ChIP-ChIP data, which also showed that MLL1 occupancy correlated with Pol II promoter occupancy (Guenther et al. 2005). We would like to investigate this interaction further using ChIP-Seq. As Pol II can be paused (CTD domain is non-phosphorylated), or actively elongating (CTD domain is phosphorylated) (Phatnani & Greenleaf 2006), different antibodies could be used to distinguish whether MLL1 binding correlates with paused or elongating RNA Polymerase II. This study is underway.

Our results also showed that the MLL1 distribution correlated with the Msk1 distribution on a third of sites and that MLL1 interacts physically with Msk1. These data suggest that they are found in one complex. The nature of this interaction remains elusive. It is known that in vitro a stable core complex of MLL1 and the WARD proteins (WDR, Ash2L, RbBP5 and Dpy-30) is needed to support the enzymatic activity for H3K4 methylation (Ernst & Vakoc 2012). However, such a core complex is not the same as a stable functional complex in vivo, where not only the enzymatic function is needed, but also targeting and regulating elements. It has been established, that MLL1 forms a multi-protein complex on chromatin (Nakamura et al. 2002, Steward et al. 2006, Patel et al. 2009), yet studies of MLL1 complex components have been controversial. For example, menin has been reported to be a stable member of this complex (Yokoyama et al. 2005), but a recent study on B-cell differentiation suggested that only a subgroup of genes are co-regulated by MLL1 and menin, whilst a larger proportion of

genes are regulated through distinct pathways (Li et al. 2013). Those discrepancies give rise to questions about the composition of the MLL1 complex itself. It is possible that the MLL1 complex consists of different proteins, depending on the cell type, or stage of differentiation, or even on the target site or stage of transcription. Current technologies only allow the investigation of cell populations, which are heterogeneous in respect to cell-cycle stage and state of transcription even in a single cell type, making it impossible to address this problem in the moment. However target site dependent MLL1 complex composition would explain why menin (Li et al. 2013), or here Msk1 seem to co-regulate only a small proportion of MLL1 target genes. Polycomb proteins, which also work as histone methyl-transferases, are known to form a small core complex with possible additional components (Aldiri & Vetter 2012). PRC2 (polycomb repressive complex 2) contains four core components: EZH 1/2, SUZ12, EED and RBBP7/4. However on most PRC2 target sites PCL (Wang et al. 2004), AEBP2 (Kim et al. 2009) and JARID2 (Li et al. 2010) can be found. Although the precise function of these additional components is not yet completely understood (ie they are not essential for methyltransferase activity) it has been suggested that they modulate PRC2 activity and promote DNA binding (Margueron & Reinberg 2011). A similar situation may be applicable for MLL1.

It also remains elusive if Msk1 is a stable member of the MLL1 complex. CBP, an acetyltransferase within the MLL1 complex only interacts transiently to activate transcription (Ernst et al. 2001). Likewise, it is possible, that Msk1 only interacts with MLL1 after stress-signals from the cell surface. A transient interaction with MLL1 to activate MLL1 target genes is consistent with the known Msk1 interaction of Msk1 and NFκB (Vermeulen et al. 2003) or CREB (Deak et al. 1998). Msk1 is activated upon cell-surface stimulation and phosphorylates

p65, which leads to the binding of NFκB p65 to its target sites and the interaction of CBP (Reber et al. 2009). Msk1 also phosphorylates and activates CREB dependent transcription. Whilst the significance of CBP to CREB activation by PKA is established, it is less clear if they are also involved in Msk1-dependent CREB activation (Naqvi et al. 2014). However, it emerges that CBP is a common co-activator for MLL1, NFκB and CREB, which leads to the speculation if CBP might mediate Msk1 dependent activation.

Our data also showed that MLL1 and Msk1 interaction leads to cross-talk of histone modifications at MLL1 target genes. H3K9ac is deposited by acetyltransferases found in the MLL1 complex and H3S10p is added by Msk1. In vitro it was shown that these two marks together stimulate MLL1's SET domain to deposit the H3K4me3 mark (Nightingale et al. 2007). The down-regulation of MLL1 led to a decrease in the H3K4me3 mark and the H3K9ac/S10p mark. The same was observed for a down-regulation of Msk1, supporting our hypothesis that there is a close functional interaction between MLL1 and Msk1 and suggesting histone cross-talk is a means to activate MLL1. We haven't investigated the possibility that Msk1 might phosphorylate MLL1 first to activate it and later after recruitment to chromatin, histone H3 to stimulate MLL1 activity further. It has been shown that MLL1 can be regulated by direct phosphorylation (Liu et al. 2010) and this course of events may explain why Msk1 only phosphorylates a small proportion of histone H3, although Msk1 itself doesn't have any preferences in substrate choice *in vitro* (Arthur 2008).



## 5 Bibliography

Agger, K., P. A. Cloos, J. Christensen, D. Pasini, S. Rose, J. Rappsilber, I. Issaeva, E. Canaani, A. E. Salcini and K. Helin (2007). "UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development." Nature **449**(7163): 731-734.

Aldiri, I. and M. L. Vetter (2012). "PRC2 during vertebrate organogenesis: a complex in transition." Dev Biol **367**(2): 91-99.

Alepuz, P. M., A. Jovanovic, V. Reiser and G. Ammerer (2001). "Stress-induced map kinase Hog1 is part of transcription activation complexes." Mol Cell **7**(4): 767-777.

Ansari, K. I. and S. S. Mandal (2010). "Mixed lineage leukemia: roles in gene expression, hormone signaling and mRNA processing." Febs j **277**(8): 1790-1804.

Arai, M., H. J. Dyson and P. E. Wright (2010). "Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain." FEBS Lett **584**(22): 4500-4504.

Argiropoulos, B. and R. K. Humphries (2007). "Hox genes in hematopoiesis and leukemogenesis." Oncogene **26**(47): 6766-6776.

Arthur, J. S. (2008). "MSK activation and physiological roles." Front Biosci **13**: 5866-5879.

Arthur, J. S. and S. C. Ley (2013). "Mitogen-activated protein kinases in innate immunity." Nat Rev Immunol **13**(9): 679-692.

Ayton, P. M. and M. L. Cleary (2003). "Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9." Genes Dev **17**(18): 2298-2307.

Badeaux, A. I. and Y. Shi (2013). "Emerging roles for chromatin as a signal integration and storage platform." Nat Rev Mol Cell Biol **14**(4): 211-224.

Baneres, J. L., A. Martin and J. Parello (1997). "The N tails of histones H3 and H4 adopt a highly structured conformation in the nucleosome." J Mol Biol **273**(3): 503-508.

Banerjee, T. and D. Chakravarti (2011). "A peek into the complex realm of histone phosphorylation." Mol Cell Biol **31**(24): 4858-4873.

Banfalvi, G. (2011). "Synchronization of mammalian cells and nuclei by centrifugal elutriation." Methods Mol Biol **761**: 25-45.

Bannister, A. J. and T. Kouzarides (1996). "The CBP co-activator is a histone acetyltransferase." Nature **384**(6610): 641-643.

Bannister, A. J. and T. Kouzarides (2011). "Regulation of chromatin by histone modifications." Cell Res **21**(3): 381-395.

Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-837.

Basecke, J., J. T. Whelan, F. Griesinger and F. E. Bertrand (2006). "The MLL partial tandem duplication in acute myeloid leukaemia." Br J Haematol **135**(4): 438-449.

Bell, A. C., A. G. West and G. Felsenfeld (1999). "The protein CTCF is required for the enhancer blocking activity of vertebrate insulators." Cell **98**(3): 387-396.

Berger, S. L., T. Kouzarides, R. Shiekhatar and A. Shilatifard (2009). "An operational definition of epigenetics." Genes Dev **23**(7): 781-783.

Bernstein, B. E., T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber and E. S. Lander (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell **125**(2): 315-326.

Bhaumik, S. R., E. Smith and A. Shilatifard (2007). "Covalent modifications of histones during development and disease pathogenesis." Nat Struct Mol Biol **14**(11): 1008-1016.

Bickmore, W. A. and B. van Steensel (2013). "Genome architecture: domain organization of interphase chromosomes." Cell **152**(6): 1270-1284.

Bird, A. (2007). "Perceptions of epigenetics." Nature **447**(7143): 396-398.

Birke, M., S. Schreiner, M. P. Garcia-Cuellar, K. Mahr, F. Titgemeyer and R. K. Slany (2002). "The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation." Nucleic Acids Res **30**(4): 958-965.

Black, J. C. and J. R. Whetstine (2013). "Tipping the lysine methylation balance in disease." Biopolymers **99**(2): 127-135.

Blobel, G. A., S. Kadauke, E. Wang, A. W. Lau, J. Zuber, M. M. Chou and C. R. Vakoc (2009). "A reconfigured pattern of MLL occupancy within mitotic chromatin promotes rapid transcriptional reactivation following mitotic exit." Mol Cell **36**(6): 970-983.

Brassasco, M. S., A. P. Montaldi, D. E. Gras, M. L. Camparoto, N. M. Martinez-Rossi, C. A. Scrideli, L. G. Tone and E. T. Sakamoto-Hojo (2009). "Cytogenetic and molecular analysis of MLL rearrangements in acute lymphoblastic leukaemia survivors." Mutagenesis **24**(2): 153-160.

- Brinkman, A. B., T. Roelofsen, S. W. Pennings, J. H. Martens, T. Jenuwein and H. G. Stunnenberg (2006). "Histone modification patterns associated with the human X chromosome." EMBO Rep **7**(6): 628-634.
- Buckley, M. S. and J. T. Lis (2014). "Imaging RNA Polymerase II transcription sites in living cells." Curr Opin Genet Dev **25**: 126-130.
- Burke, M. J. and T. Bhatla (2014). "Epigenetic modifications in pediatric acute lymphoblastic leukemia." Front Pediatr **2**: 42.
- Camporeale, G., E. E. Shubert, G. Sarath, R. Cerny and J. Zemleni (2004). "K8 and K12 are biotinylated in human histone H4." Eur J Biochem **271**(11): 2257-2263.
- Canaani, E., T. Nakamura, T. Rozovskaia, S. T. Smith, T. Mori, C. M. Croce and A. Mazo (2004). "ALL-1/MLL1, a homologue of Drosophila TRITHORAX, modifies chromatin and is directly involved in infant acute leukaemia." Br J Cancer **90**(4): 756-760.
- Caslini, C., A. S. Alarcon, J. L. Hess, R. Tanaka, K. G. Murti and A. Biondi (2000). "The amino terminus targets the mixed lineage leukemia (MLL) protein to the nucleolus, nuclear matrix and mitotic chromosomal scaffolds." Leukemia **14**(11): 1898-1908.
- Caslini, C., A. Shilatifard, L. Yang and J. L. Hess (2000). "The amino terminus of the mixed lineage leukemia protein (MLL) promotes cell cycle arrest and monocytic differentiation." Proc Natl Acad Sci U S A **97**(6): 2797-2802.
- Chakravarty, S., L. Zeng and M. M. Zhou (2009). "Structure and site-specific recognition of histone H3 by the PHD finger of human autoimmune regulator." Structure **17**(5): 670-679.
- Chodaparambil, J. V., A. J. Barbera, X. Lu, K. M. Kaye, J. C. Hansen and K. Luger (2007). "A charged and contoured surface on the nucleosome regulates chromatin compaction." Nat Struct Mol Biol **14**(11): 1105-1107.

Christophersen, N. S. and K. Helin (2010). "Epigenetic control of embryonic stem cell fate." J Exp Med **207**(11): 2287-2295.

Clayton, A. L., C. A. Hazzalin and L. C. Mahadevan (2006). "Enhanced Histone Acetylation and Transcription: A Dynamic Perspective." Molecular Cell **23**(3): 289-296.

ENCODE Project Consortium (2012). "An integrated encyclopedia of DNA elements in the human genome." Nature **489**(7414): 57-74.

Cuthbert, G. L., S. Daujat, A. W. Snowden, H. Erdjument-Bromage, T. Hagiwara, M. Yamada, R. Schneider, P. D. Gregory, P. Tempst, A. J. Bannister and T. Kouzarides (2004). "Histone deimination antagonizes arginine methylation." Cell **118**(5): 545-553.

Daniel, J. A., M. G. Pray-Grant and P. A. Grant (2005). "Effector proteins for methylated histones: an expanding family." Cell Cycle **4**(7): 919-926.

Daser, A. and T. H. Rabbitts (2004). "Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis." Genes Dev **18**(9): 965-974.

Daujat, S., U. Zeissler, T. Waldmann, N. Happel and R. Schneider (2005). "HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding." J Biol Chem **280**(45): 38090-38095.

Deak, M., A. D. Clifton, L. M. Lucocq and D. R. Alessi (1998). "Mitogen- and stress-activated protein kinase-1 (MSK1) is directly activated by MAPK and SAPK2/p38, and may mediate activation of CREB." Embo j **17**(15): 4426-4441.

Dharmarajan, V. and M. S. Cosgrove (2011). Biochemistry of the Mixed Lineage Leukemia 1 (MLL1) Protein and Targeted Therapies for Associated Leukemia. Acute Leukemia - The Scientist's Perspective and Challenge. P. M. Antica.

Dharmarajan, V., J. H. Lee, A. Patel, D. G. Skalnik and M. S. Cosgrove (2012). "Structural basis for WDR5 interaction (Win) motif recognition in human SET1 family histone methyltransferases." J Biol Chem **287**(33): 27275-27289.

Dillon, N. (2004). "Heterochromatin structure and function." Biol Cell **96**(8): 631-637.

Dion, M. F., S. J. Altschuler, L. F. Wu and O. J. Rando (2005). "Genomic characterization reveals a simple histone H4 acetylation code." Proc Natl Acad Sci U S A **102**(15): 5501-5506.

Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu and B. Ren (2012). "Topological domains in mammalian genomes identified by analysis of chromatin interactions." Nature **485**(7398): 376-380.

Dorigo, B., T. Schalch, A. Kulangara, S. Duda, R. R. Schroeder and T. J. Richmond (2004). "Nucleosome arrays reveal the two-start organization of the chromatin fiber." Science **306**(5701): 1571-1573.

Dou, Y., T. A. Milne, A. J. Tackett, E. R. Smith, A. Fukuda, J. Wysocka, C. D. Allis, B. T. Chait, J. L. Hess and R. G. Roeder (2005). "Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF." Cell **121**(6): 873-885.

Drobic, B., B. Perez-Cadahia, J. Yu, S. K. Kung and J. R. Davie (2010). "Promoter chromatin remodeling of immediate-early genes is mediated through H3 phosphorylation at either serine 28 or 10 by the MSK1 multi-protein complex." Nucleic Acids Res **38**(10): 3196-3208.

du Preez, L. L. and H. G. Patterson (2013). "Secondary structures of the core histone N-terminal tails: their role in regulating chromatin structure." Subcell Biochem **61**: 37-55.

Duda, K. and M. Frödin (2012). Stimuli That Activate MSK in Cells and the Molecular Mechanism of Activation. Austin (TX), Landes Bioscience.

Dyson, M. H., S. Thomson, M. Inagaki, H. Goto, S. J. Arthur, K. Nightingale, F. J. Iborra and L. C. Mahadevan (2005). "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 **118**(Pt 10): 2247-2259.

Edmunds, J. W. and L. C. Mahadevan (2004). "MAP kinases as structural adaptors and enzymatic activators in transcription complexes." J Cell Sci **117**(Pt 17): 3715-3723.

Eissenberg, J. C., M. G. Lee, J. Schneider, A. Ilvarsonn, R. Shiekhattar and A. Shilatifard (2007). "The trithorax-group gene in *Drosophila* little imaginal discs encodes a trimethylated histone H3 Lys4 demethylase." Nat Struct Mol Biol **14**(4): 344-346.

Ennas, M. G., C. Sorio, R. Greim, M. Nieddu, A. Scarpa, S. Orlandini, C. M. Croce, G. H. Fey and R. Marschalek (1997). "The human ALL-1/MLL/HRX antigen is predominantly localized in the nucleus of resting and proliferating peripheral blood mononuclear cells." Cancer Res **57**(10): 2035-2041.

Ernst, P. and C. R. Vakoc (2012). "WRAD: enabler of the SET1-family of H3K4 methyltransferases." Brief Funct Genomics **11**(3): 217-226.

Ernst, P., J. Wang, M. Huang, R. H. Goodman and S. J. Korsmeyer (2001). "MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein." Mol Cell Biol **21**(7): 2249-2258.

Felix, C. A. (1998). "Secondary leukemias induced by topoisomerase-targeted drugs." Biochim Biophys Acta **1400**(1-3): 233-255.

Felsenfeld, G. and M. Groudine (2003). "Controlling the double helix." Nature **421**(6921): 448-453.

Fischle, W. (2008). "Talk is cheap--cross-talk in establishment, maintenance, and readout of chromatin modifications." Genes Dev **22**(24): 3375-3382.

Ford, A. M., S. A. Ridge, M. E. Cabrera, H. Mahmoud, C. M. Steel, L. C. Chan and M. Greaves (1993). "In utero rearrangements in the trithorax-related oncogene in infant leukaemias." Nature **363**(6427): 358-360.

Franklin, T. B. and I. M. Mansuy (2010). "Epigenetic inheritance in mammals: evidence for the impact of adverse environmental effects." Neurobiol Dis **39**(1): 61-65.

Frischer, L. E., F. S. Hagen and R. L. Garber (1986). "An inversion that disrupts the Antennapedia gene causes abnormal structure and localization of RNAs." Cell **47**(6): 1017-1023.

Gall, J. (1963). "Chromosome fibers from an interphase nucleus." Science **139**(3550): 120-121.

Garcia-Ramirez, M., F. Dong and J. Ausio (1992). "Role of the histone "tails" in the folding of oligonucleosomes depleted of histone H1." J Biol Chem **267**(27): 19587-19595.

Gaunt, S. J., P. T. Sharpe and D. Buboule (1988). "Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan." Development **104**: 169-179.

Gehani, S. S., S. Agrawal-Singh, N. Dietrich, N. S. Christophersen, K. Helin and K. Hansen (2010). "Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation." Mol Cell **39**(6): 886-900.

Ghirlando, R., K. Giles, H. Gowher, T. Xiao, Z. Xu, H. Yao and G. Felsenfeld (2012). "Chromatin domains, insulators, and the regulation of gene expression." Biochim Biophys Acta **1819**(7): 644-651.

Graham, A., N. Papalopulu and R. Krumlauf (1989). "The murine and Drosophila homeobox gene complexes have common features of organization and expression." Cell **57**(3): 367-378.



Grant, W. D. and M. Morrison (1979). "Resolution of cells by centrifugal elutriation." Anal Biochem **98**(1): 112-115.

Guenther, M. G., R. G. Jenner, B. Chevalier, T. Nakamura, C. M. Croce, E. Canaani and R. A. Young (2005). "Global and Hox-specific roles for the MLL1 methyltransferase." Proc Natl Acad Sci U S A **102**(24): 8603-8608.

Hansen, J. C., J. K. Nyborg, K. Luger and L. A. Stargell (2010). "Histone chaperones, histone acetylation, and the fluidity of the chromogenome." J Cell Physiol **224**(2): 289-299.

Hansen, J. C. and A. P. Wolffe (1992). "Influence of chromatin folding on transcription initiation and elongation by RNA polymerase III." Biochemistry **31**(34): 7977-7988.

Hanson, R. D., J. L. Hess, B. D. Yu, P. Ernst, M. van Lohuizen, A. Berns, N. M. van der Lugt, C. S. Shashikant, F. H. Ruddle, M. Seto and S. J. Korsmeyer (1999). "Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development." Proc Natl Acad Sci U S A **96**(25): 14372-14377.

Harshman, S. W., N. L. Young, M. R. Parthun and M. A. Freitas (2013). "H1 histones: current perspectives and challenges." Nucleic Acids Res **41**(21): 9593-9609.

Hauf, S., R. W. Cole, S. LaTerra, C. Zimmer, G. Schnapp, R. Walter, A. Heckel, J. van Meel, C. L. Rieder and J. M. Peters (2003). "The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint." J Cell Biol **161**(2): 281-294.

Hauge, C. and M. Frodin (2006). "RSK and MSK in MAP kinase signalling." J Cell Sci **119**(Pt 15): 3021-3023.

Hazzalin, C. A. and L. C. Mahadevan (2005). "Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun." PLoS Biol **3**(12): e393.

Heitz, E. (1928). Das Heterochromatin der Moose. Jahrb Wiss Botanik. **69**: 762–818.

Henikoff, S. and A. Shilatifard (2011). "Histone modification: cause or cog?" Trends Genet **27**(10): 389-396.

Hess, J. L. (2004). MLL, Hox genes, and leukemia: the plot thickens.

Hess, J. L., B. D. Yu, B. Li, R. Hanson and S. J. Korsmeyer (1997). "Defects in yolk sac hematopoiesis in Mll-null embryos." Blood **90**(5): 1799-1806.

Holland, P. W. (2013). "Evolution of homeobox genes." Wiley Interdiscip Rev Dev Biol **2**(1): 31-45.

Horton, S. J., D. G. Grier, G. J. McGonigle, A. Thompson, M. Morrow, I. De Silva, D. A. Moulding, D. Kioussis, T. R. Lappin, H. J. Brady and O. Williams (2005). "Continuous MLL-ENL expression is necessary to establish a "Hox Code" and maintain immortalization of hematopoietic progenitor cells." Cancer Res **65**(20): 9245-9252.

Hsu, J. Y., Z. W. Sun, X. Li, M. Reuben, K. Tatchell, D. K. Bishop, J. M. Grushcow, C. J. Brame, J. A. Caldwell, D. F. Hunt, R. Lin, M. M. Smith and C. D. Allis (2000). "Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes." Cell **102**(3): 279-291.

Huang da, W., B. T. Sherman, R. Stephens, M. W. Baseler, H. C. Lane and R. A. Lempicki (2008). "DAVID gene ID conversion tool." Bioinformatics **2**(10): 428-430.

Huang, Y., J. Fang, M. T. Bedford, Y. Zhang and R. M. Xu (2006). "Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A." Science **312**(5774): 748-751.

Inbar-Feigenberg, M., S. Choufani, D. T. Butcher, M. Roifman and R. Weksberg (2013). "Basic concepts of epigenetics." Fertil Steril **99**(3): 607-615.

Izpisua-Belmonte, J. C., H. Falkenstein, P. Dolle, A. Renucci and D. Duboule (1991). "Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body." Embo j **10**(8): 2279-2289.

Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33 Suppl**: 245-254.

Jakovcevski, M. and S. Akbarian (2012). "Epigenetic mechanisms in neurological disease." Nat Med **18**(8): 1194-1204.

Jirtle, R. L. and M. K. Skinner (2007). "Environmental epigenomics and disease susceptibility." Nat Rev Genet **8**(4): 253-262.

Jude, C. D., L. Climer, D. Xu, E. Artinger, J. K. Fisher and P. Ernst (2007). "Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors." Cell Stem Cell **1**(3): 324-337.

Kadauke, S. and G. A. Blobel (2013). "Mitotic bookmarking by transcription factors." Epigenetics Chromatin **6**(1): 6.

Kalashnikova, A. A., M. E. Porter-Goff, U. M. Muthurajan, K. Luger and J. C. Hansen (2013). "The role of the nucleosome acidic patch in modulating higher order chromatin structure." J R Soc Interface **10**(82): 20121022.

Kaneko, Y., J. D. Rowley, H. S. Maurer, D. Variakojis and J. W. Moohr (1982). "Chromosome pattern in childhood acute nonlymphocytic leukemia (ANLL)." Blood **60**(2): 389-399.

Karolchik, D., G. P. Barber, J. Casper, H. Clawson, M. S. Cline, M. Diekhans, T. R. Dreszer, P. A. Fujita, L. Guruvadoo, M. Haeussler, R. A. Harte, S. Heitner, A. S. Hinrichs, K. Learned, B. T. Lee, C. H. Li, B. J. Raney, B. Rhead, K. R. Rosenbloom, C. A. Sloan, M. L. Speir, A. S. Zweig, D. Haussler, R. M. Kuhn and W. J. Kent (2014). "The UCSC Genome Browser database: 2014 update." Nucleic Acids Res **42**(Database issue): D764-770.

Kauffman, M. G., S. J. Noga, T. J. Kelly and A. D. Donnenberg (1990). "Isolation of cell cycle fractions by counterflow centrifugal elutriation." Anal Biochem **191**(1): 41-46.

Khorasanizadeh, S. (2004). "The nucleosome: from genomic organization to genomic regulation." Cell **116**(2): 259-272.

Kim, H., K. Kang and J. Kim (2009). "AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2." Nucleic Acids Res **37**(9): 2940-2950.

Klein, A. M., E. Zaganjor and M. H. Cobb (2013). "Chromatin-tethered MAPKs." Curr Opin Cell Biol **25**(2): 272-277.

Klymenko, T. and J. Muller (2004). "The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins." EMBO Rep **5**(4): 373-377.

Kolybaba, A. and A. K. Classen (2014). "Sensing cellular states--signaling to chromatin pathways targeting Polycomb and Trithorax group function." Cell Tissue Res **356**(3): 477-493.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.

Krause, A. and M. Olson (2005). The Basics of R and S-Plus. New York, Springer-Verlag.

Krivtsov, A. V. and S. A. Armstrong (2007). "MLL translocations, histone modifications and leukaemia stem-cell development." Nat Rev Cancer **7**(11): 823-833.

Lai, J. S. and W. Herr (1992). "Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations." Proc Natl Acad Sci U S A **89**(15): 6958-6962.

Lan, F., R. E. Collins, R. De Cegli, R. Alpatov, J. R. Horton, X. Shi, O. Gozani, X. Cheng and Y. Shi (2007). "Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression." Nature **448**(7154): 718-722.

Lange, M., S. Demajo, P. Jain and L. Di Croce (2011). "Combinatorial assembly and function of chromatin regulatory complexes." Epigenomics **3**(5): 567-580.

Latham, J. A. and S. Y. Dent (2007). "Cross-regulation of histone modifications." Nat Struct Mol Biol **14**(11): 1017-1024.

Lau, P. N. and P. Cheung (2011). "Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing." Proc Natl Acad Sci U S A **108**(7): 2801-2806.

Lee, B. M. and L. C. Mahadevan (2009). "Stability of histone modifications across mammalian genomes: implications for 'epigenetic' marking." J Cell Biochem **108**(1): 22-34.

Lee, M. G., J. Norman, A. Shilatifard and R. Shiekhattar (2007). "Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein." Cell **128**(5): 877-887.

Lewis, E. B. (1978). "A gene complex controlling segmentation in *Drosophila*." Nature **276**(5688): 565-570.

Li, B. E., T. Gan, M. Meyerson, T. H. Rabbitts and P. Ernst (2013). "Distinct pathways regulated by menin and by MLL1 in hematopoietic stem cells and developing B cells." Blood **122**(12): 2039-2046.

Li, G., R. Margueron, M. Ku, P. Chambon, B. E. Bernstein and D. Reinberg (2010). "Jarid2 and PRC2, partners in regulating gene expression." Genes Dev **24**(4): 368-380.

Liedtke, M. and M. L. Cleary (2009). "Therapeutic targeting of MLL." Blood **113**(24): 6061-6068.

Liu, H., E. H. Cheng and J. J. Hsieh (2007). "Bimodal degradation of MLL by SCFSkp2 and APCCdc20 assures cell cycle execution: a critical regulatory circuit lost in leukemogenic MLL fusions." Genes Dev **21**(19): 2385-2398.

Liu, H., S. Takeda, R. Kumar, T. D. Westergard, E. J. Brown, T. K. Pandita, E. H. Cheng and J. J. Hsieh (2010). "Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint." Nature **467**(7313): 343-346.

Liu, Y., E. G. Shepherd and L. D. Nelin (2007). "MAPK phosphatases--regulating the immune response." Nat Rev Immunol **7**(3): 202-212.

Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution." Nature **389**(6648): 251-260.

Macdonald, N., J. P. Welburn, M. E. Noble, A. Nguyen, M. B. Yaffe, D. Clynes, J. G. Moggs, G. Orphanides, S. Thomson, J. W. Edmunds, A. L. Clayton, J. A. Endicott and L. C. Mahadevan (2005). "Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3." Mol Cell **20**(2): 199-211.

Macrae, I. J., K. Zhou, F. Li, A. Repic, A. N. Brooks, W. Z. Cande, P. D. Adams and J. A. Doudna (2006). "Structural basis for double-stranded RNA processing by Dicer." Science **311**(5758): 195-198.

Maeshima, K. and M. Eltsov (2008). "Packaging the genome: the structure of mitotic chromosomes." J Biochem **143**(2): 145-153.

Mahadevan, L. C., A. C. Willis and M. J. Barratt (1991). "Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors." Cell **65**(5): 775-783.

Mallette, F. A. and S. Richard (2012). "The fight of the Tudor domain "Royal family" for a broken DNA throne." Cell Cycle **11**(8): 1483-1484.

Margueron, R. and D. Reinberg (2011). "The Polycomb complex PRC2 and its mark in life." Nature **469**(7330): 343-349.

Martinez-Zamudio, R. and H. C. Ha (2012). "Histone ADP-ribosylation facilitates gene transcription by directly remodeling nucleosomes." Mol Cell Biol **32**(13): 2490-2502.

McCoy, C. E., D. G. Campbell, M. Deak, G. B. Bloomberg and J. S. Arthur (2005). "MSK1 activity is controlled by multiple phosphorylation sites." Biochem J **387**(Pt 2): 507-517.

Mellor, J. (2006). "It takes a PHD to read the histone code." Cell **126**(1): 22-24.

Mendenhall, E. M. and B. E. Bernstein (2008). "Chromatin state maps: new technologies, new insights." Curr Opin Genet Dev **18**(2): 109-115.

Miller, T., N. J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst, M. Johnston, J. F. Greenblatt and A. Shilatifard (2001). "COMPASS: a complex of proteins associated with a trithorax-related SET domain protein." Proc Natl Acad Sci U S A **98**(23): 12902-12907.

Milne, T. A., S. D. Briggs, H. W. Brock, M. E. Martin, D. Gibbs, C. D. Allis and J. L. Hess (2002). "MLL targets SET domain methyltransferase activity to Hox gene promoters." Mol Cell **10**(5): 1107-1117.

Mishra, B. P., K. I. Ansari and S. S. Mandal (2009). "Dynamic association of MLL1, H3K4 trimethylation with chromatin and Hox gene expression during the cell cycle." Febs j **276**(6): 1629-1640.

Montavon, T. and D. Duboule (2013). "Chromatin organization and global regulation of Hox gene clusters." Philos Trans R Soc Lond B Biol Sci **368**(1620): 20120367.

Montavon, T. and N. Soshnikova (2014). "Hox gene regulation and timing in embryogenesis." Semin Cell Dev Biol **34**: 76-84.

Muller, F. and L. Tora (2014). "Chromatin and DNA sequences in defining promoters for transcription initiation." Biochim Biophys Acta **1839**(3): 118-128.

Musselman, C. A., M. E. Lalonde, J. Cote and T. G. Kutateladze (2012). "Perceiving the epigenetic landscape through histone readers." Nat Struct Mol Biol **19**(12): 1218-1227.

Nakamura, T., T. Mori, S. Tada, W. Krajewski, T. Rozovskaia, R. Wassell, G. Dubois, A. Mazo, C. M. Croce and E. Canaani (2002). "ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation." Mol Cell **10**(5): 1119-1128.

Naqvi, S., K. J. Martin and J. S. Arthur (2014). "CREB phosphorylation at Ser133 regulates transcription via distinct mechanisms downstream of cAMP and MAPK signalling." Biochem J **458**(3): 469-479.

Ng, H. H., R. M. Xu, Y. Zhang and K. Struhl (2002). "Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79." J Biol Chem **277**(38): 34655-34657.

Niedzialkowska, E., F. Wang, P. J. Porebski, W. Minor, J. M. Higgins and P. T. Stukenberg (2012). "Molecular basis for phosphospecific recognition of histone H3 tails by Survivin paralogues at inner centromeres." Mol Biol Cell **23**(8): 1457-1466.



Nightingale, K. P., S. Gendreizig, D. A. White, C. Bradbury, F. Hollfelder and B. M. Turner (2007). "Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation." J Biol Chem **282**(7): 4408-4416.

Nightingale, K. P., L. P. O'Neill and B. M. Turner (2006). "Histone modifications: signalling receptors and potential elements of a heritable epigenetic code." Curr Opin Genet Dev **16**(2): 125-136.

Ohta, S., L. Wood, J. C. Bukowski-Wills, J. Rappsilber and W. C. Earnshaw (2011). "Building mitotic chromosomes." Curr Opin Cell Biol **23**(1): 114-121.

Olins, A. L. and D. E. Olins (1974). "Spheroid chromatin units (v bodies)." Science **183**(4122): 330-332.

O'Neill, L. P., M. D. VerMilyea and B. M. Turner (2006). "Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations." Nat Genet **38**(7): 835-841.

Orlovsky, K., A. Kalinkovich, T. Rozovskaia, E. Shezen, T. Itkin, H. Alder, H. G. Ozer, L. Carramusa, A. Avigdor, S. Volinia, A. Buchberg, A. Mazo, O. Kollet, C. Largman, C. M. Croce, T. Nakamura, T. Lapidot and E. Canaani (2011). "Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation." Proc Natl Acad Sci U S A **108**(19): 7956-7961.

Pang, D. and D. N. Thompson (2011). "Embryology and bony malformations of the craniovertebral junction." Childs Nerv Syst **27**(4): 523-564.

Papp, B. and J. Muller (2006). "Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins." Genes Dev **20**(15): 2041-2054.

Paro, R. and D. S. Hogness (1991). "The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*." Proc Natl Acad Sci U S A **88**(1): 263-267.

Patel, A., V. Dharmarajan and M. S. Cosgrove (2008). "Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide." J Biol Chem **283**(47): 32158-32161.

Patel, A., V. Dharmarajan, V. E. Vought and M. S. Cosgrove (2009). "On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex." J Biol Chem **284**(36): 24242-24256.

Pengelly, A. R., O. Copur, H. Jackle, A. Herzig and J. Muller (2013). "A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb." Science **339**(6120): 698-699.

Petruk, S., Y. Sedkov, S. Smith, S. Tillib, V. Kraevski, T. Nakamura, E. Canaani, C. M. Croce and A. Mazo (2001). "Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene." Science **294**(5545): 1331-1334.

Phatnani, H. P. and A. L. Greenleaf (2006). "Phosphorylation and functions of the RNA polymerase II CTD." Genes Dev **20**(21): 2922-2936.

Poirier, M. G., E. Oh, H. S. Tims and J. Widom (2009). "Dynamics and function of compact nucleosome arrays." Nat Struct Mol Biol **16**(9): 938-944.

Razin, S. V. and A. A. Gavrilov (2014). "Chromatin without the 30-nm fiber: constrained disorder instead of hierarchical folding." Epigenetics **9**(5): 653-657.

Reber, L., L. Vermeulen, G. Haegeman and N. Frossard (2009). "Ser276 phosphorylation of NF- $\kappa$ B p65 by MSK1 controls SCF expression in inflammation." PLoS One **4**(2): e4393.

- Riggs, A. and T. Porter (1996). Epigenetic Mechanisms of Gene Regulation. CoCSH Monograph Archive. **32**: 29-45.
- Ringner, M. (2008). "What is principal component analysis?" Nat Biotechnol **26**(3): 303-304.
- Robinson, P. J. and D. Rhodes (2006). "Structure of the '30 nm' chromatin fibre: a key role for the linker histone." Curr Opin Struct Biol **16**(3): 336-343.
- Rosenbloom, K. R., C. A. Sloan, V. S. Malladi, T. R. Dreszer, K. Learned, V. M. Kirkup, M. C. Wong, M. Maddren, R. Fang, S. G. Heitner, B. T. Lee, G. P. Barber, R. A. Harte, M. Diekhans, J. C. Long, S. P. Wilder, A. S. Zweig, D. Karolchik, R. M. Kuhn, D. Haussler and W. J. Kent (2013). "ENCODE data in the UCSC Genome Browser: year 5 update." Nucleic Acids Res **41**(Database issue): D56-63
- .
- Rothbart, S. B. and B. D. Strahl (2014). "Interpreting the language of histone and DNA modifications." Biochim Biophys Acta **1839**(8): 627-643.
- Routh, A., S. Sandin and D. Rhodes (2008). "Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure." Proc Natl Acad Sci U S A **105**(26): 8872-8877.
- Ruthenburg, A. J., C. D. Allis and J. Wysocka (2007). "Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark." Mol Cell **25**(1): 15-30.
- Ruthenburg, A. J., W. Wang, D. M. Graybosch, H. Li, C. D. Allis, D. J. Patel and G. L. Verdine (2006). "Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex." Nat Struct Mol Biol **13**(8): 704-712.
- Sanchez, R., J. Meslamani and M. M. Zhou (2014). "The bromodomain: from epigenome reader to druggable target." Biochim Biophys Acta **1839**(8): 676-685.

Santos-Rosa, H., A. Kirmizis, C. Nelson, T. Bartke, N. Saksouk, J. Cote and T. Kouzarides (2009). "Histone H3 tail clipping regulates gene expression." Nat Struct Mol Biol **16**(1): 17-22.

Schneider, I. (1972). "Cell lines derived from late embryonic stages of *Drosophila melanogaster*." J Embryol Exp Morphol **27**(2): 353-365.

Schones, D. E., K. Cui, S. Cuddapah, T. Y. Roh, A. Barski, Z. Wang, G. Wei and K. Zhao (2008). "Dynamic regulation of nucleosome positioning in the human genome." Cell **132**(5): 887-898.

Schones, D. E. and K. Zhao (2008). "Genome-wide approaches to studying chromatin modifications." Nat Rev Genet **9**(3): 179-191.

Schroter, H., G. Maier, H. Ponstingl and A. Nordheim (1985). "DNA intercalators induce specific release of HMG 14, HMG 17 and other DNA-binding proteins from chicken erythrocyte chromatin." Embo j **4**(13b): 3867-3872.

Schuettengruber, B., D. Chourrout, M. Vervoort, B. Leblanc and G. Cavalli (2007). "Genome regulation by polycomb and trithorax proteins." Cell **128**(4): 735-745.

Schumacher, A. and T. Magnuson (1997). "Murine Polycomb- and trithorax-group genes regulate homeotic pathways and beyond." Trends Genet **13**(5): 167-170.

Schwammle, V., C. M. Aspalter, S. Sidoli and O. N. Jensen (2014). "Large scale analysis of co-existing post-translational modifications in histone tails reveals global fine structure of cross-talk." Mol Cell Proteomics **13**(7): 1855-1865.

Schwartz, Y. B., T. G. Kahn, D. A. Nix, X. Y. Li, R. Bourgon, M. Biggin and V. Pirrotta (2006). "Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*." Nat Genet **38**(6): 700-705.

Selenko, P., R. Sprangers, G. Stier, D. Buhler, U. Fischer and M. Sattler (2001). "SMN tudor domain structure and its interaction with the Sm proteins." Nat Struct Biol **8**(1): 27-31.

Shein, H. M. and J. F. Enders (1962). "Transformation induced by simian virus 40 in human renal cell cultures. I. Morphology and growth characteristics." Proc Natl Acad Sci U S A **48**: 1164-1172.

Shi, X., T. Hong, K. L. Walter, M. Ewalt, E. Michishita, T. Hung, D. Carney, P. Pena, F. Lan, M. R. Kaadige, N. Lacoste, C. Cayrou, F. Davrazou, A. Saha, B. R. Cairns, D. E. Ayer, T. G. Kutateladze, Y. Shi, J. Cote, K. F. Chua and O. Gozani (2006). "ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression." Nature **442**(7098): 96-99.

Shiio, Y. and R. N. Eisenman (2003). "Histone sumoylation is associated with transcriptional repression." Proc Natl Acad Sci U S A **100**(23): 13225-13230.

Shimada, M., M. Haruta, H. Niida, K. Sawamoto and M. Nakanishi (2010). "Protein phosphatase 1gamma is responsible for dephosphorylation of histone H3 at Thr 11 after DNA damage." EMBO Rep **11**(11): 883-889.

Shogren-Knaak, M., H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie and C. L. Peterson (2006). "Histone H4-K16 acetylation controls chromatin structure and protein interactions." Science **311**(5762): 844-847.

Sim, S. P. and L. F. Liu (2001). "Nucleolytic cleavage of the mixed lineage leukemia breakpoint cluster region during apoptosis." J Biol Chem **276**(34): 31590-31595.

Simonet, T., R. Dulermo, S. Schott and F. Palladino (2007). "Antagonistic functions of SET-2/SET1 and HPL/HP1 proteins in *C. elegans* development." Dev Biol **312**(1): 367-383.

Slany, R. K. (2005). "When epigenetics kills: MLL fusion proteins in leukemia." Hematol Oncol **23**(1): 1-9.

Slany, R. K. (2009). "The molecular biology of mixed lineage leukemia." Haematologica **94**(7): 984-993.

Smith, C. L. (2001). Mammalian Cell Culture. Current Protocols in Molecular Biology, John Wiley & Sons, Inc.

Soloaga, A., S. Thomson, G. R. Wiggan, N. Rampersaud, M. H. Dyson, C. A. Hazzalin, L. C. Mahadevan and J. S. Arthur (2003). "MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14." Embo j **22**(11): 2788-2797.

Southall, S. M., P. S. Wong, Z. Odho, S. M. Roe and J. R. Wilson (2009). "Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks." Mol Cell **33**(2): 181-191.

Steffen, P. A. and L. Ringrose (2014). "What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory." Nat Rev Mol Cell Biol **15**(5): 340-356.

Steward, M. M., J. S. Lee, A. O'Donovan, M. Wyatt, B. E. Bernstein and A. Shilatifard (2006). "Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes." Nat Struct Mol Biol **13**(9): 852-854.

Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-45.

Struhl, K. (1999). "Fundamentally different logic of gene regulation in eukaryotes and prokaryotes." Cell **98**(1): 1-4.

Suganuma, T. and J. L. Workman (2011). "Signals and combinatorial functions of histone modifications." Annu Rev Biochem **80**: 473-499.

Talbert, P. B. and S. Henikoff (2006). "Spreading of silent chromatin: inaction at a distance." Nat Rev Genet **7**(10): 793-803.

Tamaru, H. (2010). "Confining euchromatin/heterochromatin territory: jumonji crosses the line." Genes Dev **24**(14): 1465-1478.

Tamkun, J. W., R. Deuring, M. P. Scott, M. Kissinger, A. M. Pattatucci, T. C. Kaufman and J. A. Kennison (1992). "brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2." Cell **68**(3): 561-572.

Terranova, R., H. Agherbi, A. Boned, S. Meresse and M. Djabali (2006). "Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll." Proc Natl Acad Sci U S A **103**(17): 6629-6634.

Terrenoire, E., F. McDonald, J. A. Halsall, P. Page, R. S. Illingworth, A. M. Taylor, V. Davison, L. P. O'Neill and B. M. Turner (2010). "Immunostaining of modified histones defines high-level features of the human metaphase epigenome." Genome Biol **11**(11): R110.

Thomson, S., A. L. Clayton and L. C. Mahadevan (2001). "Independent dynamic regulation of histone phosphorylation and acetylation during immediate-early gene induction." Mol Cell **8**(6): 1231-1241.

Tie, F., R. Banerjee, A. R. Saiakhova, B. Howard, K. E. Monteith, P. C. Scacheri, M. S. Cosgrove and P. J. Harte (2014). "Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing." Development **141**(5): 1129-1139.

Tittel-Elmer, M., E. Bucher, L. Broger, O. Mathieu, J. Paszkowski and I. Vaillant (2010). "Stress-induced activation of heterochromatic transcription." PLoS Genet **6**(10): e1001175.

Towbin, H., T. Staehelin and J. Gordon (1992). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979." Biotechnology **24**: 145-149.

Trojer, P. and D. Reinberg (2007). "Facultative Heterochromatin: Is There a Distinctive Molecular Signature?" Molecular Cell **28**(1): 1-13.

Tschopp, P., B. Tarchini, F. Spitz, J. Zakany and D. Duboule (2009). "Uncoupling time and space in the collinear regulation of Hox genes." PLoS Genet **5**(3): e1000398.

Turner, B. M. (2005). "Reading signals on the nucleosome with a new nomenclature for modified histones." Nat Struct Mol Biol **12**(2): 110-112.

Turner, B. M. (2012). "The adjustable nucleosome: an epigenetic signaling module." Trends Genet **28**(9): 436-444.

Tusher, V. G., R. Tibshirani and G. Chu (2001). "Significance analysis of microarrays applied to the ionizing radiation response." Proc Natl Acad Sci U S A **98**(9): 5116-5121.

Unnikrishnan, A., P. R. Gafken and T. Tsukiyama (2010). "Dynamic changes in histone acetylation regulate origins of DNA replication." Nat Struct Mol Biol **17**(4): 430-437.

Vermeulen, L., G. De Wilde, P. Van Damme, W. Vanden Berghe and G. Haegeman (2003). "Transcriptional activation of the NF-kappaB p65 subunit by mitogen- and stress-activated protein kinase-1 (MSK1)." Embo j **22**(6): 1313-1324.

Vermeulen, L., W. Vanden Berghe, I. M. Beck, K. De Bosscher and G. Haegeman (2009). "The versatile role of MSKs in transcriptional regulation." Trends Biochem Sci **34**(6): 311-318.



Wang, P., C. Lin, E. R. Smith, H. Guo, B. W. Sanderson, M. Wu, M. Gogol, T. Alexander, C. Seidel, L. M. Wiedemann, K. Ge, R. Krumlauf and A. Shilatifard (2009). "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 **29**(22): 6074-6085.

Wang, S., G. P. Robertson and J. Zhu (2004). "A novel human homologue of Drosophila polycomblike gene is up-regulated in multiple cancers." Gene **343**(1): 69-78.

Wang, Z., C. Zang, K. Cui, D. E. Schones, A. Barski, W. Peng and K. Zhao (2009). "Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes." Cell **138**(5): 1019-1031.

Weber, C. M. and S. Henikoff (2014). "Histone variants: dynamic punctuation in transcription." Genes Dev **28**(7): 672-682.

West, M. H. and W. M. Bonner (1980). "Histone 2B can be modified by the attachment of ubiquitin." Nucleic Acids Res **8**(20): 4671-4680.

Weth, O. and R. Renkawitz (2011). "CTCF function is modulated by neighboring DNA binding factors." Biochem Cell Biol **89**(5): 459-468.

Wiggin, G. R., A. Soloaga, J. M. Foster, V. Murray-Tait, P. Cohen and J. S. Arthur (2002). "MSK1 and MSK2 are required for the mitogen- and stress-induced phosphorylation of CREB and ATF1 in fibroblasts." Mol Cell Biol **22**(8): 2871-2881.

Woodcock, C. L. and R. P. Ghosh (2010). "Chromatin higher-order structure and dynamics." Cold Spring Harb Perspect Biol **2**(5): a000596.

Wysocka, J., T. Swigut, H. Xiao, T. A. Milne, S. Y. Kwon, J. Landry, M. Kauer, A. J. Tackett, B. T. Chait, P. Badenhorst, C. Wu and C. D. Allis (2006). "A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling." Nature **442**(7098): 86-90.

Xia, Z. B., M. Anderson, M. O. Diaz and N. J. Zeleznik-Le (2003). "MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein." Proc Natl Acad Sci U S A **100**(14): 8342-8347.

Xu, Y. M., J. Y. Du and A. T. Lau (2014). "Posttranslational modifications of human histone H3: an update." Proteomics **14**(17-18): 2047-2060.

Yang, H., A. Ganguly and F. Cabral (2010). "Inhibition of cell migration and cell division correlates with distinct effects of microtubule inhibiting drugs." J Biol Chem **285**(42): 32242-32250.

Yin, H., S. Sweeney, D. Raha, M. Snyder and H. Lin (2011). "A high-resolution whole-genome map of key chromatin modifications in the adult *Drosophila melanogaster*." PLoS Genet **7**(12): e1002380.

Yokoyama, A. and M. L. Cleary (2008). "Menin critically links MLL proteins with LEDGF on cancer-associated target genes." Cancer Cell **14**(1): 36-46.

Yokoyama, A., I. Kitabayashi, P. M. Ayton, M. L. Cleary and M. Ohki (2002). "Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties." Blood **100**(10): 3710-3718.

Yokoyama, A., T. C. Somervaille, K. S. Smith, O. Rozenblatt-Rosen, M. Meyerson and M. L. Cleary (2005). "The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis." Cell **123**(2): 207-218.

Yu, B. D., R. D. Hanson, J. L. Hess, S. E. Horning and S. J. Korsmeyer (1998). "MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis." Proc Natl Acad Sci U S A **95**(18): 10632-10636.

Yu, B. D., J. L. Hess, S. E. Horning, G. A. Brown and S. J. Korsmeyer (1995). "Altered Hox expression and segmental identity in Mll-mutant mice." Nature **378**(6556): 505-508.

Yuan, W., T. Wu, H. Fu, C. Dai, H. Wu, N. Liu, X. Li, M. Xu, Z. Zhang, T. Niu, Z. Han, J. Chai, X. J. Zhou, S. Gao and B. Zhu (2012). "Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation." Science **337**(6097): 971-975.

Zacharias, H. (1995). "Emil Heitz (1892-1965): chloroplasts, heterochromatin, and polytene chromosomes." Genetics **141**(1): 7-14.

Zee, B. M., R. S. Levin, P. A. Dimaggio and B. A. Garcia (2010). "Global turnover of histone post-translational modifications and variants in human cells." Epigenetics Chromatin **3**(1): 22.

Zeisig, B. B., T. Milne, M. P. Garcia-Cuellar, S. Schreiner, M. E. Martin, U. Fuchs, A. Borkhardt, S. K. Chanda, J. Walker, R. Soden, J. L. Hess and R. K. Slany (2004). "Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization." Mol Cell Biol **24**(2): 617-628.

Zeleznik-Le, N. J., A. M. Harden and J. D. Rowley (1994). "11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene." Proc Natl Acad Sci U S A **91**(22): 10610-10614.

Zeng, L., Q. Zhang, S. Li, A. N. Plotnikov, M. J. Walsh and M. M. Zhou (2010). "Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b." Nature **466**(7303): 258-262.

Zhang, G. and S. Pradhan (2014). "Mammalian epigenetic mechanisms." IUBMB Life **66**(4): 240-256.

Zhang, Y., A. Chen, X. M. Yan and G. Huang (2012). "Disordered epigenetic regulation in MLL-related leukemia." Int J Hematol **96**(4): 428-437.

Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li and X. S. Liu (2008). "Model-based analysis of ChIP-Seq (MACS)." Genome Biol **9**(9): R137.

Zheng, C. and J. J. Hayes (2003). "Intra- and inter-nucleosomal protein-DNA interactions of the core histone tail domains in a model system." J Biol Chem **278**(26): 24217-24224.

Ziemin-van der Poel, S., N. R. McCabe, H. J. Gill, R. Espinosa, 3rd, Y. Patel, A. Harden, P. Rubinelli, S. D. Smith, M. M. LeBeau, J. D. Rowley and et al. (1991). "Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias." Proc Natl Acad Sci U S A **88**(23): 10735-10739.

Zink, B. and R. Paro (1989). "In vivo binding pattern of a trans-regulator of homoeotic genes in *Drosophila melanogaster*." Nature **337**(6206): 468-471.

Zippo, A., R. Serafini, M. Rocchigiani, S. Pennacchini, A. Krepelova and S. Oliviero (2009). "Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation." Cell **138**(6): 1122-1136.

Zur Hausen, H. (1967). "Induction of specific chromosomal aberrations by adenovirus type 12 in human embryonic kidney cells." J Virol **1**(6): 1174-1185.