Histone methylation by PRC2 is inhibited by active chromatin marks

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch–Naturwissenschaftlichen Fakultät

der Universität Basel

von

Frank W. Schmitges

aus Deutschland

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Dr. Susan Gasser, Dr. Nicolas Thomä, Dr. Wolfgang Fischle.

Basel, den 20. September 2011

Prof. Dr. Susan Gasser
(Fakultätsverantwortliche)

Prof. Dr. Martin Spiess
(Dekan)

Table of Contents

Abbreviations

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AEBP2	adipocyte enhancer-binding protein 2
AG	AGAMOUS
Ala	alanine
Arg	arginine
Ash1	absent, small or homeotic 1
at	Arabidopsis thaliana
ATP	Adenosine-5'-triphosphate
BRCT	BRCA1 C Terminus
Caf-1	chromatin assembly factor 1
Cbx	chromobox homolog
CDK	cyclin dependent kinase
ChIP	chromatin immuno precipitation
CLF	CURLY LEAF
CpG	cytosine-guanine dinucleotide
dm	Drosophila melanogaster
DNA	deoxyribonucleic acid
Dnmt1	DNA methyltransferase 1
E(z)	enhancer of zeste
EAF3	Esa1p-associated factor 3
EED	embryonic ectodermal development
EMF2	EMBRYONIC FLOWER2
ES cell	embryonic stem cell
Esc	extra sex combs
FIE	FERTILIZATION-INDEPENDENT ENDOSPERM
FIS2	FERTILIZATION INDEPENDENT SEED2
FLC	flowering locus C
Fog-1	friend of GATA protein 1
FP	fluorescence polarization
Gln	glutamine
GST	glutathione S-transferase
H3K27me3	histone H3 lysine 27 trimethylation
HAT	histone acetylase
HDAC	histone deacetylase
HMTase	histone methyltransferase
HOTAIR	HOX antisense intergenic RNA
HP1	heterochromatin protein 1
HPLC	high-performance liquid chromatography
ICE	imprint control element
ITC	isothermal titration calorimetry
Jarid2	jumonji, AT rich interactive domain 2
Jmj	Jumonji
kDa	kilodalton
LSD1	lysine specific demethylase 1
Lys	lysine
me1/me2/me3	mono-/di-/trimethylation
MEA	MEDEA

MLL	mixed lineage leukemia
MSI1	MULTICOPY SUPPRESSOR OF IRA1
MSK	mitogen- and stress-activated protein kinase
ncRNA	non-coding RNA
NEB	Nurf55 binding epitope
NSD	nuclear receptor SET domain-containing
NuRD	nucleosome remodeling and histone deacetylase
Nurf55	nucleosome remodeling factor 55
Oct4	octamer binding transcription factor 4
Pc	polycomb
PcG	polycomb group
Pcl	Polycomb-like
Ph	polyhomeotic
PHD	plant homeo domain
PHF1	PHD finger protein 1
Pho	pleiohomeotic
PRC1/2	polycomb repressive complex 1/2
PRE	polycomb response element
PRMT	protein arginine methyltransferase
Psc	posterior sex combs
RbAp46/48	retinoblastoma-associated protein 46/48
RBBP4/7	retinoblastoma binding protein 4/7
RING	really interesting new gene
RNA	ribonucleic acid
Rnf1/2	Ring finger protein 1/2
Ser	serine
SET	Suv39h, E(z), Trithorax
SH	Src homology
Sir	silent information regulatory
STM	SHOOTMERISTEMLESS
Su(z)12	suppressor of zeste 12
SWN	SWINGER
TAF3	TATA binding protein associated factor 3
Tet	Ten-eleven translocation
TEV	tobacco etch virus
TFIID	transcription factor IID
Thr	threonine
thrxG	trithorax group
TRE	trithorax response element
Tyr	tyrosine
VRN2	VERNALIZATION2
Xic	X-chromosomal inactivation center
Xist	X-inactivation specific transcript

1. Summary

Chromatin modifiers serve as regulatory switches that control the cell cycle, maintain pluripotency and drive differentiation and development. Positive feedback mechanisms help to pass on transcriptional information from one generation of cells to the next one. The polycomb repressive complex 2 (PRC2) is responsible for methylation of histone H3 at lysine 27, a typical mark of repressive chromatin. *De novo* methylation requires sequence-specific recruitment factors. In contrast, propagation and inheritance of the H3K27me3 mark after replication relies on a self-maintaining feedback loop: direct interaction of PRC2 with existing H3K27me3 marks triggers an allosteric stimulation of the methyltransferase activity and results in efficient modification of new histones that have been incorporated in repressive chromatin regions.

In this study we present an inhibitory mechanism that limits the spread of H3K27 methylation and protects active chromatin by breaking the positive feedback loop. PRC2 is allosterically inhibited by nucleosomes carrying active chromatin modifications such as H3K4me3 or H3K36me2/3. The mechanism is conserved in mammals, flies and even plants. In addition, plants have distinct PRC2 subcomplexes and can modulate their specificity by the choice of the Su(z)12 homologue. Furthermore, we have identified Nurf55 as another histone binding module in the PRC2 complex that recognizes unmodified histone H3 but not H3K4me3.

Taken together, H3K27 methylation presents itself as a typical bistable switch. It is driven by the positive feedback loop in PRC2 activation and limited by active mark inhibition. Numerous chromatin modifying complexes recognize their own products and positive feedback loops are a common mechanism. We postulate that all these complexes need an additional inhibitory switch that prevents spreading of histone modifications over the entire genome.

2. Introduction

2.1 Chromatin

Eukaryotic DNA is condensed and combined with proteins to form a higher order structure. The hierarchic packaging allows storage of meters of DNA in the limited space of the nucleus. On the first level of organization the DNA double strand is wrapped around nucleosomes, the fundamental repeat unit of chromatin. Nucleosomes consist of the highly conserved histone proteins H2A, H2B, H3 and H4. Two molecules of each of these histones form the core nucleosome particle with approximately 146 base pairs of DNA wrapped around it (Kornberg, 1974; Luger et al, 1997). The spacing between nucleosome particles varies between species and chromatin regions. The linker histone H1 binds DNA at the entry and exit of the nucleosome and assists in further compaction of the "beads-on-a-string" structure into the 30nm fiber (Li & Reinberg, 2011; Robinson et al, 2006). Higher-order compaction of chromatin is not so well understood but it is clear that its regulation is crucial for DNA replication and transcriptional control. DNA condensation restricts the access of DNA binding proteins. The localized activity of chromatin remodeling complexes keeps chromatin structure dynamic and allows control of transcription factor and polymerase complex binding (Li et al, 2007). Using the energy of ATP hydrolysis these chromatin remodelers temporarily unwrap the DNA to increase accessibility or to change the position of nucleosomes (Flaus & Owen-Hughes, 2004; Smith & Peterson, 2005). In addition, the removal or exchange of histones and even entire histone octamers has been described, painting the picture of a highly dynamic system (Workman, 2006). Chromatin organization is also influenced by covalent modifications on histones and DNA (Li & Reinberg, 2011). These modifications can affect DNA condensation either directly (by changing the charge of a molecule) or indirectly by interaction with chromatin modifying complexes.

2.2 Epigenetics

The field of epigenetics encompasses mechanisms that regulate gene expression without changes in the DNA sequence (Bird, 2007). The term was coined in 1942 by Conrad Waddington to describe the interaction of genes with their surroundings that influences gene expression and contributes to a phenotype. In higher organisms different cell types vary

greatly in their gene expression patterns although almost all their cells share exactly the same genomic information. Thus, mechanisms are required that regulate gene expression independently of the DNA sequence. Importantly, epigenetic mechanisms have also been shown to be responsible for processes that cannot be explained by classic genetics such as imprinting and X inactivation (Lyon, 1999; Reik & Walter, 2001).

One of the central features of epigenetic mechanisms is the accurate inheritance of gene expression patterns - dependent on the cell type and not on the DNA sequence. However, the term "epigenetics" is also commonly used for that transcriptional information which is erased during the cell cycle and requires sequence specific reestablishment after cell division. A second feature of epigenetic regulation is the ability to change expression patterns quickly and stably. These highly regulated mechanisms allow dynamic changes in gene expression and are the basis of development.

2.3 Mechanisms of epigenetic regulation

Epigenetic mechanisms affect transcriptional activation of genes in many different ways that do not require a change in the DNA sequence. Covalent modification of DNA itself or histone proteins can either directly affect chromatin structure and thereby accessibility of the DNA, or influence transcription by interaction with chromatin binding proteins. Mammalian DNA can be methylated on cytosines in the context of a CpG dinucleotide. Histones are known to be modified on over 60 residues, the most prominent modifications include methylation, acetylation, phosphorylation and ubiquitination (Kouzarides, 2007). Most of the modification sites are located on the histone tails which are accessible while the nucleosomal core is occluded by DNA. Analogous to the genetic code, an epigenetic code has been suggested (Jenuwein & Allis, 2001). While early models tried to assign functions to single histone modifications, it has since become evident that the marks have to be understood in their chromatin context: the role of a certain modification can vary depending on other modifications in its neighborhood and on the chromatin region it is found in Lee et al, 2010. Similarly, the original view of static epigenetic landscapes with permanent marks had to be corrected. More and more enzymes have been identified that specifically remove modifications, therefore the dynamics of epigenetic modifications are not limited by the exchange of histones. The correct epigenetic patterns are essential not only for development but also for the adult organism. Malfunction of the epigenetic machinery has been linked to

several diseases such as immunodeficiency, centromeric instability, facial anomalies syndrome (ICF), fragile X syndrome, and the development of cancer (Robertson & Wolffe, 2000; Tsai & Baylin, 2011).

2.3.1 DNA methylation

In mammalian cells, DNA methylation is found at the carbon-5 position of cytosines (therefore referred to as 5mC), mainly in the context of cytosine-guanine dinucleotides (CpGs). The addition of the methyl group is catalyzed by a family of enzymes called DNA methyltransferases (DNMTs) (Goll & Bestor, 2005). Cytosine methylation has been linked to gene silencing (Bird & Wolffe, 1999), triggering chromatin reorganization via methyl-binding proteins (Wade, 2001). Furthermore, it has been described that methylation can sterically hamper the binding of transcription factors to their recognition sites (Tate & Bird, 1993) and interfere with specific recognition of histone marks (Bartke et al, 2010). CpG dinucleotides are underrepresented in much of the genome but are often found in high density in sections of 500 to 4000 base pairs in the proximal promoter regions of genes. These sections have been termed CpG-islands. Methylated cytosines are distributed in a non-random fashion in genomic DNA. While mostly methylated in the bulk of DNA, cytosines in CpG-islands are often found unmethylated, thereby allowing gene expression.

In mammals three enzymes are responsible for DNA methylation (Jurkowska et al, 2011). DNA methyltransferase 1 (Dnmt1) has a preference for hemimethylated DNA and is found at replication foci pointing towards a role as a maintenance methyltransferase. In contrast Dnmt3a and Dnmt3b are *de novo* methyltransferases, known for establishment of methylation patterns during development. Interestingly, knockout studies have found a high degree of redundancy between the three enzymes. DNA demethylation is believed to mostly rely on passive demethylation by incorporation of unmethylated cytosines during replication. However, specific enzymes have been identified as well that allow active demethylation. In plants the DNA repair pathway involving DEMETER and DEMETER-like proteins has been shown to be involved in demethylation (Gehring et al, 2009). Active demethylation also plays a role in early mammalian development. The cytidine deaminase AID has been implicated in demethylation during primordial germ cell development (Popp et al, 2010).

5-methylcytosine can be further modified by hydroxylation to 5-hydroxymethylcytosine (5hmC) (Veron & Peters, 2011). This conversion is catalyzed by

proteins of the Ten-eleven translocation (Tet) family. 5hmC is found enriched on exons and correlates with gene transcription (Pastor et al, 2011; Williams et al, 2011; Wu et al, 2011). It is possible that 5hmC is an intermediate state during active DNA demethylation. Interestingly, Tet1 targets overlap significantly with polycomb targets and depletion of Tet1 indicates a direct role of 5hmC in recruitment of the polycomb repressive complex 2 (PRC2) (Wu et al, 2011). Although the mechanisms are not fully understood so far these results indicate a dual role for 5hmC in activation and repression of transcription.

2.3.2 Histone methylation

Histones can be methylated at lysine or arginine residues. Histone methyltransferases are among the most specific epigenetic modifiers, often specialized for only one residue. They can add one or two methyl groups on arginine and up to three on lysine residues (Sims et al, 2003). In contrast to other modifications, methylation is not believed to have a direct influence on chromatin structure. The marks are rather specifically recognized by chromatin binding modules that are connected to modifying complexes or remodeling machines. This allows histone methylation - depending on methylation state, position and combination with other epigenetic marks - to function either in transcriptional activation or silencing (Lee et al, 2010; Yun et al, 2011). Remodeling requires in many cases the sequential modification of nucleosomes, where each step is associated with a specific mark. A classic example is the H3K27me3 mark, involved in polycomb mediated silencing. Deposited by polycomb repressive complex 2 (PRC2), it is known to recruit in a canonical pathway the second polycomb complex PRC1 which in turn ubiquitinates histone H2A at lysine 119 (see below).

Arginine methylation is found as monomethylation, symmetrical dimethylation and asymmetrical dimethylation. Although arginine readers are by far not as well understood as lysine readers it is known that some domains can distinguish between the three methylation states. The ADD domain of Dnmt3a recognizes only symmetrically dimethylated H4R3, thereby linking this modification to DNA methylation and transcriptional repression (Zhao et al, 2009). Studies in yeast have found different distributions of H3R2me1 and H3R2me2, indicating a role for monomethylation in gene activation and dimethylation in repression (Kirmizis et al, 2009).

Lysine methylation is carried out by a family of histone methyltransferases (HMTases) that typically harbor a SET domain (named after the three HMTases Suv39h, Enhancer of

Zeste and Trithorax). They catalyze the transfer of the methyl group from the donor S-adenosyl-L-methionine to the target residue (Yeates, 2002). The exception is the Dot1 methyltransferase, a non-SET domain methyltransferase specific for H3K79 (Min et al, 2003a). Arginine methylation is catalyzed by protein arginine methyltransferases (PRMTs) which are divided into three groups: Type 1 is responsible for asymmetrical dimethylation, type 2 for symmetrical dimethylation and type 3 only catalyzes monomethylation (Bedford & Clarke, 2009).

In contrast to other histone modifications, methylation was for a long time believed to be a static mark, only reversible by exchange of histones. This dogma was dismissed upon the recent discovery of a family of demethylases, able to specifically remove methyl-groups from certain lysine residues (Shi & Whetstine, 2007). This family contains a Jumonji C (JmjC) domain which is the catalytic center for demethylation. While lysine demethylation was quickly established as a key player in development and disease, the role and mechanisms of arginine demethylation remain more controversial (Bedford & Clarke, 2009). The Jumonji domain protein Jmjd6 was originally identified as an arginine demethylase (Chang et al, 2007). More recent work, however, suggests that Jmjd6 is actually a lysine-hydroxylase (Webby et al, 2009).

As mentioned above the decoding of transcriptional information from lysine methylation highly depends on specialized chromatin readers. Lysine methylation is recognized by domains of the royal family (chromo, tudor, malignant brain tumor (MBT)) as well as plant homeodomains (PHD) and WD40 domains (Kouzarides, 2007). These domains usually form an aromatic cage in the binding site and some are able to distinguish between mono-, di- and trimethylation (Yun et al, 2011). In contrast, readers of unmethylated lysine usually rely on formation of intermolecular hydrogen bonds. Their binding pockets cannot accept methylated residues due to spatial restrictions. The readout of methylation can have different effects. On some residues the methylation state simply tunes the affinity for one binding partner, e.g. the affinity of the histone deacetylase complex Rpd3S is highest for trimethylated H3K36me3 but gradually reduced for lower methylation states (Li et al, 2009). Other residues act as a switch, which can recruit different adaptors depending on the methylation state. For example in the recruitment of the HMTase Set9, Pdp1 binds H4K20me1 but is replaced by Crb2 on H4K20me2 (Wang & Jia, 2009). As flanking regions are often important for the interaction with chromatin readers, the cross-talk of modifications in close proximity can also influence binding. The PHD domain of TAF3, part of the basal

transcription factor TFIID, binds H3K4me3 with high affinity. Asymmetrical arginine methylation on H3R2 has been shown to interfere with this interaction, thereby implicating a methyl-methyl switch in the recruitment of the transcription factor TFIID (van Ingen et al, 2008).

The role of histone modifications has been extensively studied by chromatin immunoprecipitation (ChIP) experiments that identify DNA sequences associated with certain marks or combinations of marks. Although the function of methylation marks is highly complex and often context-dependent, some marks are fairly well understood. Methyl marks on lysines H3K4, H3K36 and H3K79 for example are generally associated with active transcription (Barski et al, 2007; Pokholok et al, 2005; Saunders et al, 2006); methylation of H3K9, H3K27 and H4K20 are markers of a repressive chromatin state.

H3K4 methylation

High levels of H3K4 trimethylation are found in discrete zones at the 5' regions of almost all active genes. Only *Hox* gene clusters show continuous methylation patterns. In general, H3K4me3 strongly correlates with transcription rate, polymerase II occupancy and histone acetylation (Ruthenburg et al, 2007a). While dimethylation in yeast is rather spread throughout genes and associated with active and poised genes, in vertebrates it mostly colocalizes with H3K4me3, in proximity of active genes. High levels of H3K4me1 have been found on enhancers of active genes (Hon et al, 2009). A special role for H3K4 methylation has been described in the poised state of bivalent domains which are discussed later (Bernstein et al, 2006). In addition, H3K4 methylation has also been reported in CpG-rich promoters independent of transcriptional activity (Guenther et al, 2007; Roh et al, 2006). It has been suggested that H3K4 methylation protects these regions from DNA methylation (Ooi et al, 2007). In yeast only one HMTase, Set1, is known to methylate H3K4. In mammals at least ten complexes are responsible: six HMTases of the MLL family, ASH1, SET7/9, SMYD3, and Meisetz (Marmorstein, 2003; Ruthenburg et al, 2007a). These complexes are mostly not redundant and vary in their expression patterns. H3K4 methylation is coupled to further chromatin modification mechanisms. Readers for this mark have been found in complexes involved in ATP-dependent chromatin remodeling, histone acetylation and histone methylation and demethylation. Reading modules are found in the royal super family, the PHD-finger superfamily and WD40-proteins (Bienz, 2006; Maurer-Stroh et al, 2003). As common feature with other methyl-lysine readers, they share an aromatic cage that recognizes the modified lysine residue. It remains controversial whether H3K4 methylation is the cause

or rather the result of active transcription. PHD-fingers in transcription factors that bind H3K4me3 indicate a role in facilitating transcription (Vermeulen et al, 2007). On the other hand H3K4 HMTases have been found associated with the RNA polymerase II, which has been interpreted as a memory for recent transcriptional activity (Ng et al, 2003).

H3K27 methylation

H3K27 di- and trimethylation is associated with facultative heterochromatin and located around the transcription start site (Margueron & Reinberg, 2011). Similar to H3K4 methylation, larger domains are found on *Hox* gene clusters. Monomethylation is associated with constitutive heterochromatin, but also found in the gene body in actively transcribed regions. In contrast to H3K4 methylation only PRC2 complexes (discussed in detail below) have been found responsible for H3K27 methylation, with the exception of a viral HMTase (Margueron & Reinberg, 2011; Montgomery et al, 2005). H3K27me3 is recognized by the chromodomain of Pc and by the WD40 propeller of Esc (see below). The role of H3K27me2 is not so well understood. It could simply present an intermediate product of H3K27 trimethylation or protect H3K27 from acetylation (Tie et al, 2009).

H3K36 methylation

H3K36 methylation is another typical mark found in actively transcribed chromatin (Bannister & Kouzarides, 2011). It displays a broader distribution within the gene body, with H3K36me2 starting downstream of the H3K4me3 peak and H3K36me3 enriched in the 3' region of active genes (Bell et al, 2008; Santos-Rosa et al, 2002). Set2 is the only H3K36 HMTase in yeast. In mammals Set2 is responsible for trimethylation while the NSD proteins NSD1-3 are able to dimethylate H3K36 (Qiao et al, 2011; Strahl et al, 2002). H3K36 methylation activity has also been reported for ASH1 (Tanaka et al, 2007). In yeast Set2 has been shown to interact with the RNA polymerase II complex (Krogan et al, 2003a; Krogan et al, 2003b). Methylated H3K36 is bound by the chromodomain of EAF3 which results in recruitment of a deacetylase (Kouzarides, 2007). Deacetylation in the gene body after transcription has been suggested to prevent transcription initiation from cryptic start sites and to stabilize chromatin (Carrozza et al, 2005).

2.3.3 Histone acetylation

Acetylation of histones is found on lysine residues. In contrast to histone methylation acetylation has a strong, direct influence on chromatin structure by affecting the contact

between DNA and histones. The positive charge of lysine residues is important for the interaction of histones with the negatively charged DNA phosphate backbone. Neutralization of the positive charge on lysines by acetylation allows the controlled weakening of this interaction and increases accessibility of the DNA. Even acetylation of a single residue has been shown to alter the compaction level of a nucleosomal array (Shogren-Knaak et al, 2006). The increased mobility of nucleosomes in acetylated regions facilitates the passage of polymerases and leads to increased binding of non-histone proteins (Wolffe & Hayes, 1999; Workman & Kingston, 1998). Histone acetylation is therefore generally associated with transcriptional activation.

In addition to its role in modulating the electrostatic interactions with DNA, histone acetylation can be specifically recognized by bromodomains and tandem PHD domains (Yun et al, 2011). In bromodomains the binding pocket is formed by the inter-helical loops of the helix bundle. In PHD12 the binding site is on the β-sheet of the PHD domain. In all cases the acetylated lysine residue intercalates into a narrow, hydrophobic pocket. The acetyl amide is positioned by hydrogenbonding with amino acids at the bottom of the pocket. Compared to the often highly specific readers of histone methylation, bromo and tandem PHD domains have broader specificity and are usually able to bind multiple sites on several histones. This lack of specificity stems from less defined contacts with the flanking histone regions. Similar to the readers of histone acetylation the enzymes depositing and removing this mark, histone acetyltransferases (HATs) and deacetylases (HDACs), respectively, often show low substrate specificity and redundant functions (Yun et al, 2011).

2.3.4 Histone phosphorylation

Phosphorylation is a very common modification on non-histone proteins and a multitude of binding domains have been identified. Interestingly, only few examples have been identified so far for phosphorylated histone tails. The BRCT domain of the mediator of DNA damage checkpoint protein 1 (MDC1) has been shown to bind to phosphorylated serine at the C-terminus of the histone variant H2AX in the context of DNA repair (Stucki et al, 2005). Similarly, members of the 14-3-3 family of regulator proteins have been reported to bind phosphorylated H3S10 (Macdonald et al, 2005). Another interesting role for histone phosphorylation is found in the cross-talk with other chromatin marks: phosphorylation of H3S10 interferes with recognition of H3K9 methylation by HP1, thereby controlling

heterochromatin formation (Fischle et al, 2005). Analogously, the interaction of the demethylase LSD1 with methylated H3K4 is prevented by H3T6 phosphorylation (Metzger et al, 2010). Histone phosphorylation by mitogen- and stress-activated kinases (MSKs) has been demonstrated to assist in gene activation in response to multiple signaling pathways. Phosphorylation of H3S28 by MSKs results in displacement of polycomb proteins and eventually in derepression of PcG target genes (Gehani et al, 2010). Although specific binding domains for phosphorylated histones are less common, these marks play an important role in regulating the access to neighboring residues.

2.3.5 Histone ubiquitination

Ubiquitination has been found on all four core histones and it has been suggested to function in both gene activation and silencing as well as in DNA repair (Muratani & Tansey, 2003). The ubiquitin moiety is much larger than the other modifications and histones can undergo mono- and poly-ubiquitination (Geng & Tansey, 2008). Reading modules for ubiquitin usually recognize only a hydrophobic patch on the ubiquitin or a region near the C-terminus. Strikingly, so far no specific reader for histone ubiquitination could be identified and many aspects of its mechanisms in transcriptional control are still very enigmatic (Yun et al, 2011). In yeast the ubiquitination machinery for H2B has been shown to interact with the transcription apparatus (Xiao et al, 2005) and ubiquitination of H2B is essential for deposition of active methylation marks on H3K4 and H3K79 (Shahbazian et al, 2005; Sun & Allis, 2002). Furthermore proteases that remove ubiquitin from H2B have been suggested to provide directionality for different stages in transcription (Wyce et al, 2007). Ubiquitination of H2A by the polycomb complex PRC1 was for a long time thought to directly trigger chromatin compaction and silencing. Recent result, however, show that ubiquitination is not essential for *HOX* gene repression, raising more questions about its function (Eskeland et al, 2010).

2.3.6 The role of non-coding RNAs in transcriptional control

The finding that large parts of the genome are transcribed but not translated into protein has triggered investigations of the function of the resulting non-coding RNAs (ncRNAs). In addition to the role as tRNAs and structural functions, ncRNAs have been found to be involved in central events of transcriptional control.

The most prominent role of ncRNAs in silencing is found in female mammals where one X chromosome is stably inactivated for dosage compensation (Wutz, 2011). This process is triggered in early embryogenesis by transcription of the Xist RNA from the Xic locus on the X chromsome. The RNA is recruited to the target chromosome and triggers a cascade of events leading to the formation of facultative heterochromatin. The inactivation also involves repressive histone and DNA methylation and propagation of the silenced state is independent of Xist in somatic cells.

Another link between histone modification and ncRNAs is found in polycomb mediated silencing. The RNA HOTAIR located in the *HOXC* locus has been reported to bind to the PRC2 complex and is required for polycomb silencing *in vivo* (Rinn et al, 2007). More recently HOTAIR was found to act as a scaffold by linking the PRC2 complex to the H3K4 specific demethylase LSD1 (Tsai et al, 2010a). Thus, it has been suggested to assist H3K27 methylation by targeting and removing of competing histone marks. Similarly, the ncRNA KCNQ1OT1 interacts with PRC2 and is implicated in the silencing of the KCNQ1 domain (Pandey et al, 2008). The ncRNA ANRIL, on the other hand, interacts with the chromodomain of CBX7 in PRC1 and is involved in regulation of the tumor supressor INK4b-ARF-INK4a (Yap et al, 2010). Furthermore, ncRNAs such as Ev1xas and Hoxb5/6as also regulate trithorax activity by association with MLL1 (Dinger et al, 2008).

Non-coding RNAs are also involved in genomic imprinting where genes are expressed in a parent-of-origin-specific manner (Reik & Walter, 2001). The imprinted allele is silenced allowing gene expression specifically from the other allele, either inherited from the mother or in other cases from the father. The clusters of imprinted genes contain ncRNAs that upon expression repress transcription from the opposite strand (Koerner et al, 2009). Imprinting is based on DNA methylation of an imprint control element (ICE) that is established during gametogenesis and controls transcription of the surrounding cluster. In addition, also repressive histone modifications have been implicated in ICE regulation.

2.4 Transcriptional control by polycomb and trithorax proteins

The development of higher organisms is marked by the specialization of cells from a totipotent state to highly differentiated cell types. This process requires a tightly regulated, timed and localized control system that allows the expression of only the appropriate set of genes in a given tissue at a certain point of time. The most fundamental decisions about the

proper placement of the segments in early embryogenesis are controlled by transcription factors from the *Hox* gene clusters. Hox proteins can act as transcriptional activators or repressors and they contain a homeodomain that is used for DNA binding. A single Hox protein can activate a whole network of genes for one segment and at the same time repress genes specific for other segments (Grimaud et al, 2006).

The expression of *Hox* genes in the appropriate cell lineages is established early by the graded expression of transcription factors and morphogens. Later in development these factors are replaced by proteins of the polycomb group (PcG) and trithorax group (trxG) that propagate cell type specific *Hox* gene patterns (Grimaud et al, 2006). In general, PcG proteins are known to be transcriptional repressors while trxG proteins are activators of *Hox* gene expression. Both groups form several distinct multi-protein complexes that have either histone methyltransferase activity or ubiquitin ligase activity. The maintenance of these histone marks over generations of cell divisions presents an epigenetic memory and is key to the understanding of differentiation.

2.4.1 Polycomb group complexes

The polycomb protein (Pc) was identified in 1947 in a genetic screen in *Drosophila*. *Pc* mutant male flies displayed ectopic sex combs on the second and third leg. Subsequent studies defined the polycomb group as genes that showed *Pc*-like phenotypes in mutational studies (Grimaud et al, 2006). On the other hand trithorax proteins were identified by their ability to counteract PcG activity in homeotic gene regulation.

Biochemical purification has confirmed the existence of two distinct PcG complexes. The first complex, PRC1, consists in *Drosophila* of Polycomb (Pc), Polyhomeotic (Ph), Posterior Sex Combs (Psc) and the Ring-domain protein dRing. dRing is known to be the catalytic center of the complex and ubiquitinates histone H2A at lysine 119. The PRC2 complex is a histone methyltransferase specific for H3K27, composed around the SET-domain protein Enhancer of zeste (E(z), EZH2 in mammals). The other core subunits include the zinc-finger protein suppressor of zeste 12 (Su(z)12, mammalian SUZ12) and the WD40 proteins Extra sex combs (Esc, mammalian EED) and Nurf55 (mammalian RbAp46/48) (Lund & van Lohuizen, 2004; Nekrasov et al, 2005).

E(z) and its homologues are the only eukaryotic methyltransferases known that are specific for H3K27. In addition to its SET domain, E(z) harbors an interaction site for Esc close to its N-termius and two SANT domains (Han et al, 2007). E(z) is not catalytically active by itself but requires a minimal complex including Su(z)12 and Esc (Nekrasov et al, 2005). EZH1 is a second mammalian homologue of E(z) with high sequence homology to EZH2. It is also specific for H3K27, mainly catalyzing monomethylation and to a smaller extent di- and trimethylation. Its role is still controversial. Knockout of *Ezh2* in ES cells resulted in global loss of H3K27me2/3. However, a small number of crucial genes were repressed by polycomb complexes containing Ezh1 (Shen et al, 2008). Thus a redundant function has been suggested on a subset of EZH2 target genes. In addition, another recent study has suggested a role for EZH1 in chromatin compaction independent of its catalytic activity (Margueron et al, 2008).

Su(z)12 is known to contribute to nucleosome binding and at the same time it is essential for stability and activity of E(z) (Ketel et al, 2005; Nekrasov et al, 2005). It has a zinc-finger domain that has been implicated in DNA and RNA binding as well as a VEFS domain, named after the homologues VRN2, EMF2, FIS2 and Su(z)12. Mutations in the VEFS domain indicate that it might constitute a binding interface of Su(z)12 with E(z) (Ketel et al, 2005).

Interestingly, Nurf55 is the only protein within the PRC2 complex that is not exclusively found in polycomb complexes but is also a core component of other chromatin modifiers involved in histone acetylation as well as nucleosome assembly and remodeling. It was first identified in mammals as ubiquitous binding partner for the retinoblastoma protein (Qian et al, 1993). Like Esc, it belongs to the WD40 family, a group of adaptor proteins with a large spectrum of substrates ranging from peptides to nucleic acids and even sugars (Xu & Min, 2011). WD40 proteins such as Esc and WDR5 have been found to be specific readers of histone modifications. Although Nurf55 is a core component of the PRC2 complex and essential for H3K27 methylation *in vivo* (Anderson et al, 2011) its mode of operation is not understood. Recently, binding of Nurf55 to helix 1 of histone H4 was reported. However, biochemical evidence suggests that H4 binding is relevant in the context of histone acetylation but not for PRC2-mediated methylation (Song et al, 2008).

The size of a stoichiometric four-component PRC2 complex in *Drosophila* would be approximately 270 kDa. Gelfiltration analysis of biochemically purified complexes, however, showed a peak at ~600 kDa for the bulk of E(z) proteins (Ng et al, 2000; Tie et al, 2001)

indicating the presence of more than one copy of the core proteins per complex or association with additional PRC2 binding proteins. The role and stoichiometry of distinct subcomplexes is not completely understood so far. A number a proteins directly associated with core PRC2 proteins that are responsible for targeting and modulation of enzymatic activity have been reported that affect H3K27 methylation *in vitro* and *in vivo*. The PcG protein Polycomb-like (Pcl, mammalian PHF1) was reported to be required for high levels of H3K27 methylation (Nekrasov et al, 2007). Other proteins such as the catalytically inactive Jumonji domain protein Jarid2 and the zinc-finger proteins Jing/AEBP2 and Pho/YY1 have been implicated in targeting of PRC2 (Brown et al, 2003; Kim et al, 2009; Pasini et al, 2010).

2.4.2 Polycomb proteins in plants

Both polycomb complexes are highly conserved and homologues are even found in plants. Studies of PRC2 in *Arabidopsis* have identified three different homologues of E(z) (MEDEA (MEA), SWINGER (SWN) and CURLY LEAF (CLF)) and three homologues of Su(z)12 (EMBRYONIC FLOWER2 (EMF2), VERNALIZATION2 (VRN2) and FERTILIZATION-INDEPENDENT SEED2 (FIS2)). Together with the Esc and Nurf55 homologues FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) and MULTICOPY SUPPRESSOR OF IRA1 (MSI1) these proteins are found in different combinations as functionally distinct complexes that are essential for female gametophyte, endosperm and embryonic development (Pien & Grossniklaus, 2007). Mutation of several PcG proteins leads e.g. to endosperm development in the absence of fertilization. CLF and SWN appear to be functionally redundant and are the main E(z) homologues during vegetative growth, while MEA is functioning during gametogenesis and seed development. The Su(z)12 homologue FIS2 associates with MEA; EMF2 and VRN2 form distinct complexes with CLF or SWN which differ in their targets. EMF2-complexes control flower development by repression of AGAMOUS (AG) and SHOOTMERISTEMLESS (STM) (Schubert et al, 2006). VRN2 on the other hand is involved in the initiation of flowering after the winter. The Flowering Locus C (FLC) suppresses flowering during cold. In spring, a process called vernalization triggers VRN2-PRC2 dependent silencing of FLC and thereby the transition to flowering (Schmitz & Amasino, 2007). The distinct PRC2 subcomplexes in plants are not fully characterized on a biochemical level and more information is necessary to understand the contribution of the different homologues.

2.4.3 Mechanisms of polycomb-mediated silencing

PRC1 and PRC2 have been shown to cooperate in a stepwise mechanism in order to silence *Hox* genes. PRC2-mediated H3K27 methylation at the target region creates specific binding sites for the chromodomain of Pc (Fischle et al, 2003b; Min et al, 2003b), thereby recruiting PRC1 (**Figure 1**). This relationship is supported by studies in flies and mammalian cells reporting a colocalization of PRC1 and PRC2 and showing that depletion of PRC2 also leads to a loss of PRC1 at the target site (Boyer et al, 2006; Cao et al, 2002). Moreover an artificial increase of H3K27 methylation by knock down of the H3K27 specific demethylase UTX resulted in increased PRC1 recruitment (Lee et al, 2007).

Studies in ES cells deficient in PRC1 activity (*Rnf2* mutant), PRC2 activity (*Eed* mutant) or both have demonstrated that silencing of most target genes requires the coordinated activity of both complexes. Recent results, however, might hint at redundant or even independent functions of the two complexes. Mapping of PRC1 and PRC2 distribution in ES cells has revealed PRC2-only regions (Ku et al, 2008). Furthermore, ES cells deficient in either PRC1 or PRC2 were able to differentiate, while simultaneous loss of both complexes abrogated differentiation (Leeb et al, 2010). A possible explanation might be the observed direct interaction of PRC1 with the DNA-binding protein Pho (Mohd-Sarip et al, 2006). Furthermore X inactivation in *Eed*-deficient embryos did not require PRC2 and Xist RNA was able to recruit PRC1 independent of H3K27me3 (Kalantry & Magnuson, 2006; Schoeftner et al, 2006). Additional examples for PRC2-independent recruitment have been found in experiments with *Suz12*-deficient ES cells and *Ezh2* mutant zygotes lacking H3K27me3 (Pasini et al, 2007; Puschendorf et al, 2008). Taken together, PRC1 is in most cases targeted by PRC2-mediated H3K27 methylation. Alternative recruitment in certain situations might however allow PRC1 function in the absence of PRC2.

Both methylation of H3K27 and monoubiquitination of H2AK119 are typical marks of repressive chromatin. However, the mechanisms that lead downstream to chromatin compaction and silencing are poorly understood. For a long time ubiquitination of H2AK119 was thought to play a central role in the polycomb silencing pathway. Recent results have challenged this view. While PRC1 itself is essential for gene repression (Wang et al, 2004), a catalytically inactive complex is still able to induce silencing independent of H2A ubiquitination at *Hox* loci (Eskeland et al, 2010). Therefore, a more direct role of PRC1 has been suggested, e.g. by blocking the transcription machinery. PRC1 does not interfere with

transcription initiation, instead transcription elongation is a much more likely target. Poised RNA polymerase II in bivalent domains has been reported to be activated upon depletion of PRC1 in ES cells (Stock et al, 2007). Similarly, poised RNA polymerase II has also been found on polycomb targets in *Drosophila* (Chopra et al, 2009).

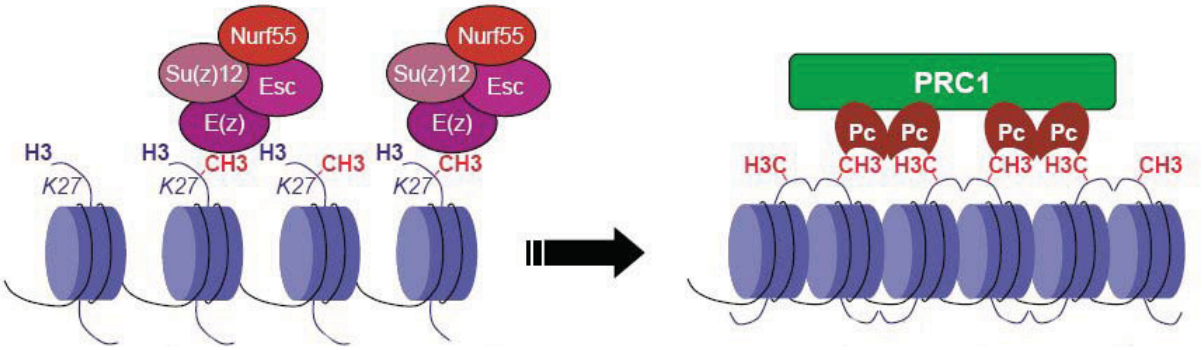


Figure 1: Cooperation of PRC1 and PRC2 in silencing. The PRC2 complex is responsible for trimethylation of H3K27. This mark serves as binding platform for Pc, thereby recruiting the PRC1 complex and triggering chromatin compaction and silencing (Lund & van Lohuizen, 2004).

Another mode of action could include nucleosome remodeling. *In vitro* studies have shown that PRC1 is able to compact nucleosomal arrays, thereby making them refractory to SWI/SNF-class remodelers (Shao et al, 1999). At the same time, chromatin compaction could also contribute to the interference with transcriptional elongation described above. The relationship between PRC1 and nucleosome remodelers *in vivo* is controversial but in summary, results indicate a role in limiting the access of proteins necessary for transcription to the silenced chromatin.

2.4.4 Targeting of polycomb complexes

Polycomb complexes are able to bind nucleosomes and show activity *in vitro*. However, both complexes do not have DNA binding modules that would allow site-specific targeting. PcG proteins are ubiquitously expressed; however, their recruitment to target sites varies greatly between cell types. Thus, initiation of Polycomb silencing *in vivo* requires the association with additional DNA binding factors.

PRC2 targeting in *Drosophila*

In *Drosophila* two classes of cis-regulatory DNA elements have been identified as initiator elements and maintenance elements in transcriptional regulation. During the very first steps of development the initiator elements control gene expression based on the local concentration of segmentation gene products that varies throughout the embryo. Once the homeotic gene patterns have been established the positional information is maintained by the concerted action of polycomb and trithorax complexes after decay of the segmentation gene products (Ringrose & Paro, 2007). Maintenance elements called polycomb and trithorax response elements (PREs/TREs) which are only a few hundred base pairs long are required for the recruitment of these complexes. PREs and TREs function as epigenetic memory that retains the transcriptional state of the associated genes through many cell divisions in the absence of the original activating or repressing factors. The epigenetic state of these sequences is, however, reversible, thereby allowing to switch the transcriptional state during development.

Several proteins have been identified as mediators between PcG and trxG complexes and their response elements (Ringrose & Paro, 2004). The best studied adaptors include Pho and Pho-like proteins, which are involved in PcG silencing as well as Zeste, Gaf and Pipsqueak which can bind to PREs and TREs.

In contrast to other DNA recognition motifs PREs and TREs are not easy to identify as they are not binding sites for a single transcription factor, but rather a family of recruitment sites for a whole set of adaptor proteins. Therefore known PRE/TRE sequences share common mechanistic features but little sequence homology and lack any kind of consensus sequence. Some attempts to predict PREs and TRES have used an alignment-independent algorithm (Ringrose et al, 2003). Predicted sites were compared to experimentally identified sites showing that many sites were found at least with weak scores by the algorithm. However, some sites seem to require DNA looping and cooperative binding to a second (non-PRE/TRE) site or the combination with weak unspecific DNA binding. Interestingly, studies on synthetic PREs and TREs suggest that additional binding sites for *Hox* regulators such as Dsp1 are required for polycomb and trithorax proteins to function (Dejardin et al, 2005). Other proteins have been associated as well with a role for certain PREs and TREs. Their importance for polycomb and trithorax systems on a global level needs to be evaluated though (Ringrose & Paro, 2007). In any case, the identification of additional factors might help to explain the cell type-specific regulation of target genes.

PRC2 targeting in mammals

Although thousands of polycomb and trithorax target sites have been mapped, the search for specific recognition motifs in the mammalian genome turned out to be more difficult. While binding of PcG proteins in flies is very localized, it is found in much broader domains in mammals (Boyer et al, 2006). The recognition and recruitment process seems to be much more complex than in *Drosophila,* but similar factors seem to play a key role. Knockdown of YY1, the mammalian homologue of Pho, was shown to reduce H3K27me3 at target sites (Caretti et al, 2004). Other studies have suggested a role for the transcription factors Oct4, Nanog and Sox2 in recruitment, as they have been found to co-occupy a subset of PRC2 targets in ES cells (Boyer et al, 2006; Lee et al, 2006). However, no functional link has been provided for this model and co-occupancy might simply result from redundant but independent silencing pathways. Another study reported a high correlation of PRC2 binding and CpG islands, suggesting a potential role of CpG binding proteins in targeting (Ku et al, 2008). Furthermore, the Jumonji domain protein Jarid2 has been identified in PRC2 complexes and a role in H3K27 methylation *in vivo* has been demonstrated (Pasini et al, 2010). The AIRD domain of Jarid2 has a broad specificity for different DNA sequences. A model has therefore been suggested that combines the low specificity of different DNA binding modules to cooperatively achieve high specificity for polycomb targeting.

The involvement of both H3K27 methylation and ncRNAs in X inactivation and imprinting triggered the idea that RNAs could also play a role in polycomb targeting (Mercer et al, 2009). The ncRNA Xist involved in the initiation of X inactivation has been shown to bind PRC2 *via* a repetitive element called repA (Zhao et al, 2008). A similar physical interaction with PRC2 has been described for the ncRNA Kcnq1ot1 essential for imprinted silencing (Pandey et al, 2008) and HOTAIR, a ncRNA that regulates H3K27 methylation *in vivo* and connects PRC2 with the demethylase LSD1 (Rinn et al, 2007; Tsai et al, 2010a). Further research is likely to identify more PRC2 associated RNAs and to increase our understanding of their role in polycomb targeting.

In summary the targeting of PcG and trxG complexes in *Drosophila* is guided by the existence of response elements at the target sites and specific sets of adaptor proteins leading to very localized occupancy. The situation in mammalian cells is more complex and different targeting mechanisms might be responsible for different silencing phenomena. Thus, and due to the lack of a consensus sequence only very few response elements have been reported in mammals (Sing et al, 2009).

2.4.5 Polycomb silencing as an epigenetic memory

Transcriptional information in the form of histone marks faces the problem that each cycle of DNA replication and cell division is associated with the incorporation of fresh nucleosomes and thereby a thinning out of the existing marks. Stable transcriptional repression over many generations of cell divisions as observed in polycomb targets therefore requires ways to faithfully propagate this information. In order to maintain cell type-specific expression profiles the underlying mechanisms must be mostly independent of the DNA sequence.

The finding that PRC2 is able to bind its own product has led to different models of the propagation of the H3K27me3 mark during replication. Transcription of PRC2 proteins is initiated at the beginning of S-phase (Bracken et al, 2003; Muller et al, 2001). The PRC2 complexes localize with sites of replication to modify newly incorporated histones. A first study found that a trimeric complex of EZH2, SUZ12 and EED is required for binding of H3K27me3 (Hansen et al, 2008). The observation that after initial targeting PRC2-mediated repressive chromatin can be maintained independently of additional targeting factors implicated a self-maintaining feedback loop. It was suggested that PRC2 is directly recruited by H3K27me3 modified histone tails on the same nucleosome, on neighboring nucleosomes or even on the opposite DNA strand at the replication fork. Further structural and functional studies of PRC2 have challenged this recruitment-based model. EED was found to bind H3K27me3 independently of other PRC2 components and thereby trigger allosteric stimulation of the EZH2 methyltransferase activity (Margueron et al, 2009). Mutations in the *Drosophila* homologue demonstrated the importance of this interaction for PRC2 activity *in vivo*. This mechanism supports a model in which chromatin interactions of other PRC2 components are recruiting the complex to the replication site. Binding of EED to H3K27me3 on neighboring nucleosomes is mainly responsible for enhancing the methyltransferase activity in regions where repressive chromatin marks are present.

In contrast to inheritance of the PRC2 mark, little is known so far about H3K27me3-independent propagation of PRC1 ubiquitination. Removal of H2A ubiquitination is necessary at the onset of mitosis to allow progression of the cell cycle (Joo et al, 2007). The PRC1 complex has been reported to stay associated with chromatin during S- and M-phase, opening the possibility for direct re-ubiquitination (Francis et al, 2009; Puschendorf et al, 2008). However, there is so far no experimental evidence for this step and it is possible that new complexes have to be recruited guided by PRC2 activity.

2.4.6 Polycomb in pluripotency and differentiation

Reproduction and development form a cycle in all higher organisms that is marked by differentiation and reprogramming of cells and requires a specific and highly dynamic regulation of transcription. The gametes that fuse to form the zygote are among the most specialized cell types with a unique gene expression pattern. Fusion of the two cells triggers a program that resets most of the transcriptional information and leads to the totipotent state of embryonic stem cells. But already in the early pre-implantation embryo this process is reversed and some cells start to specialize. Stem cells have the ability to maintain a pluripotent state for the entire life of the organism, while at the same time offer the possibility to regenerate certain tissues by differentiation into specialized cells upon the correct signals. Epigenetic mechanisms offer the perfect set of tools to control this process. The transcriptional information in form of epigenetic modifications can be faithfully inherited but also provides means to either transiently or permanently change expression patterns. By their nature epigenetic mechanisms do not change the DNA sequence so that genomic information can be inherited without changes. The importance of epigenetic regulation during development is reflected in the drastic changes of DNA and histone modifications during this process (Santos & Dean, 2004).

The central role of polycomb proteins in development is demonstrated by studies showing that knockouts lead to severe malformations or even death in flies (Grimaud et al, 2006). Knockout experiments in mice deleting any core PRC2 component or Rnf2 (mammalian homologue of *Drosophila* dRing) from the PRC1 complex all result in embryonic lethality (Faust et al, 1995; O'Carroll et al, 2001; Voncken et al, 2003). Knockouts of other PRC1 members lead to very mild phenotypes - most likely due to the existence of several homologues and a high degree of redundancy.

Mapping of PcG protein occupancy in ES cells has found them associated with many key developmental regulators that are not expressed in stem cells or only at very low levels (Boyer et al, 2006; Lee et al, 2006). Differentiation of stem cells is marked by the specific activation of certain subsets of these genes, suggesting an important role for polycomb proteins at the threshold to differentiation. This is supported by results of PcG knockouts in ES cells: while maintenance of pluripotency strictly speaking does not require fully functional PRC1 or PRC2, knockout cells show expression of typcial differentiation markers and are prone to spontaneous differentiation (Chamberlain et al, 2008; Leeb & Wutz, 2007; Pasini et

al, 2007; Shen et al, 2008; van der Stoop et al, 2008). Surprisingly, upon *in vitro* differentiation these knockout cells cannot give rise to the full range of cell types since they fail to repress pluripotency genes and to activate the appropriate set of differentiation markers. Double knockout of *Rnf1* and *Rnf2* (and thereby complete loss of PRC1 function) even interrupts stem cell proliferation, again hinting at a PRC2-independent function of PRC1 (Endoh et al, 2008).

Bivalent domains are an interesting phenomenon found in ES cells where large domains are kept in a transcriptionally poised state by simultaneous presence of active H3K4me3 and repressive H3K27me3 marks (Bernstein et al, 2006). These regions are pre-loaded with poised RNA polymerase and are rapidly activated upon loss of H3K27me3 or permanently silenced by removal of H3K4me3. Bivalent domains have also been identified to a smaller extend in other cell types (Mikkelsen et al, 2007; Mohn et al, 2008) but bivalency seems to be a hallmark of pluripotency and declines with increasing commitment of a cell. Bivalency is not limited to mammals. The FLC locus in *Arabidopsis* has also been reported to carry active and repressive marks simultaneously (Jiang et al, 2008).

Polycomb proteins have also been shown to act within the regulatory network of the key transcription factors Oct4, Sox2 and Nanog. These pluripotency factors are able to promote stemness and repress differentiation while propagating their own as well as each other's transcription level in a positive feedback loop (Boyer et al, 2005; Catena et al, 2004; Kuroda et al, 2005; Okumura-Nakanishi et al, 2005; Rodda et al, 2005). Polycomb complexes have been found at many silenced targets of this network. Knockdown of Oct4 reduces recruitment of PRC1 at these sites while loss of PRC1 does not affect Oct4 (Endoh et al, 2008). Thus, polycomb silencing seems to function downstream of the Oct4/Sox2/Nanog network.

Taken together, these observations show the two faces of polycomb mediated repression. Long term silencing by faithful inheritance of histone modifications is crucial for the maintenance of a differentiated state in somatic cells. Regulation of the balance between self-renewal and differentiation in stem cells in contrast requires a more plastic and flexible role that allows rapid changes in gene expression. This is achieved by the combination of repressive and active chromatin marks in bivalent domains and by the use of a transcription factor-driven feedback loop as master switch in pluripotency.

2.5 Scope of the thesis

Recent studies have shed light on structural and functional aspects of the PRC2 complex. We have started to gain better understanding of the subunit contributions and the targeting mechanisms involved. It remains, however, enigmatic how the individual subunits of the PRC2 core complex affect methyltransferase activity of the E(z) SET domain.

At the time when I started my PhD studies the role of the PRC2 core components in histone methylation and nucleosome binding had been characterized *in vitro* (Nekrasov et al, 2005). No structural information was available and no data about the molecular mechanisms that regulate the activity of the complex. The finding that PRC2 activity is stimulated by binding to its own product, H3K27me3, constitutes a hallmark in the understanding of its regulation (Hansen et al, 2008; Margueron et al, 2009). This positive feedback loop allows inheritance of the PRC2 mark after DNA replication, in the absence of sequence specific targeting factors. However, it also became quickly evident that such a powerful activation mechanism requires counteractive measures to ensure the long-term integrity of clearly defined active and repressive domains in chromatin. We therefore set out to look for further interactions of PRC2 with chromatin marks and to study their influence of H3K27 methylation.

Numerous studies have reported direct interaction of histone H3 with PRC2 component Nurf55 (Beisel et al, 2002; Hansen et al, 2008; Song et al, 2008; Wysocka et al, 2006). At the same time, Nurf55 has been shown to interact with histone H4 and Fog-1 (Lejon et al, 2011; Song et al, 2008). However, these interactions are relevant for the role of Nurf55 in histone acetylation and nucleosome remodeling but not in the context of PRC2. Taken together these findings triggered us to investigate the molecular basis of histone H3 recognition by Nurf55 and to eventually understand how its substrate specificity is determined.

In our approach to understand the regulation and substrate specificity of PRC2 we have combined structural and biochemical studies. Based on our results and available in vivo data we have expanded the current model for the inheritance of repressive chromatin marks.

3. Results

3.1 Histone methylation by PRC2 is inhibited by active chromatin marks

(published manuscript)

Molecular Cell
Article

Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks

Frank W. Schmitges,[1,6] Archana B. Prusty,[2,6] Mahamadou Faty,[1] Alexandra Stützer,[3] Gondichatnahalli M. Lingaraju,[1] Jonathan Aiwazian,[1] Ragna Sack,[1] Daniel Hess,[1] Ling Li,[4] Shaolian Zhou,[4] Richard D. Bunker,[1] Urs Wirth,[5] Tewis Bouwmeester,[5] Andreas Bauer,[5] Nga Ly-Hartig,[2] Kehao Zhao,[4] Homan Chan,[4] Justin Gu,[4] Heinz Gut,[1] Wolfgang Fischle,[3] Jürg Müller,[2,7,*] and Nicolas H. Thomä[1,*]

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland
[2]Genome Biology Unit, EMBL Heidelberg, Meyerhofstrasse 1, D-69117 Heidelberg, Germany
[3]Laboratory of Chromatin Biochemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany
[4]China Novartis Institutes for Biomedical Research, Lane 898 Halei Road, Zhangjiang, Shanghai, China
[5]Novartis Institutes for Biomedical Research, CH-4002 Basel, Switzerland
[6]These authors contributed equally to this work
[7]Present address: Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany
*Correspondence: muellerj@biochem.mpg.de (J.M.), nicolas.thoma@fmi.ch (N.H.T.)
DOI 10.1016/j.molcel.2011.03.025

SUMMARY

The Polycomb repressive complex 2 (PRC2) confers transcriptional repression through histone H3 lysine 27 trimethylation (H3K27me3). Here, we examined how PRC2 is modulated by histone modifications associated with transcriptionally active chromatin. We provide the molecular basis of histone H3 N terminus recognition by the PRC2 Nurf55-Su(z)12 submodule. Binding of H3 is lost if lysine 4 in H3 is trimethylated. We find that H3K4me3 inhibits PRC2 activity in an allosteric fashion assisted by the Su(z)12 C terminus. In addition to H3K4me3, PRC2 is inhibited by H3K36me2/3 (i.e., both H3K36me2 and H3K36me3). Direct PRC2 inhibition by H3K4me3 and H3K36me2/3 active marks is conserved in humans, mouse, and fly, rendering transcriptionally active chromatin refractory to PRC2 H3K27 trimethylation. While inhibition is present in plant PRC2, it can be modulated through exchange of the Su(z)12 subunit. Inhibition by active chromatin marks, coupled to stimulation by transcriptionally repressive H3K27me3, enables PRC2 to autonomously template repressive H3K27me3 without overwriting active chromatin domains.

INTRODUCTION

Polycomb (PcG) and trithorax group (trxG) proteins form distinct multiprotein complexes that modify chromatin. These complexes are conserved in animals and plants and are required to maintain spatially restricted transcription of HOX and other cell fate determination genes (Henderson and Dean, 2004; Pietersen and van Lohuizen, 2008; Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007). PcG proteins act to repress their target genes while trxG protein complexes are required to keep the same genes active in cells where they must be expressed.

Among the PcG protein complexes, Polycomb repressive complex 2 (PRC2) is a histone methyl-transferase (HMTase) that methylates Lys27 of H3 (H3K27) (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2004; Müller et al., 2002). High levels of H3K27 trimethylation (H3K27me3) in the coding region generally correlate with transcription repression (Cao et al., 2008; Nekrasov et al., 2007; Sarma et al., 2008). PRC2 contains four core subunits: Enhancer of zeste [E(z), EZH2 in mammals], Suppressor of zeste 12 [Su(z)12, SUZ12 in mammals], Extra-sex combs [ESC, EED in mammals] and Nurf55 [Rbbp4/RbAp48 and Rbbp7/RbAp46 in mammals] (reviewed in Schuettengruber et al., 2007; Wu et al., 2009). E(z) is the catalytic subunit; it requires Nurf55 and Su(z)12 for nucleosome association, whereas ESC is required to boost the catalytic activity of E(z) (Nekrasov et al., 2005). Recent studies reported that ESC binds to H3K27me3 and that this interaction stimulates the HMTase activity of the complex (Hansen et al., 2008; Margueron et al., 2009; Xu et al., 2010). The observation that PRC2 is able to bind to the same modification that it deposits led to a model for propagation of H3K27me3 during replication. In this model, recognition of H3K27me3 on previously modified nucleosomes promotes methylation of neighboring nucleosomes that contain newly incorporated unmodified histone H3 (Hansen et al., 2008; Margueron et al., 2009). However, it is unclear how such a positive feedback loop ensures that H3K27 trimethylation remains localized to repressed target genes and does not invade the chromatin of nearby active genes.

In organisms ranging from yeast to humans, chromatin of actively transcribed genes is marked by H3K4me3, H3K36me2, and H3K36me3 modifications: while H3K4me3 is tightly localized at and immediately downstream of the transcription start site, H3K36me2 peaks adjacently in the 5′ coding region and H3K36me3 is specifically enriched in the 3′ coding region (Bell et al., 2008; Santos-Rosa et al., 2002). Among the trxG proteins that keep PcG target genes active are the HMTases Trx and Ash1, which methylate H3K4 and H3K36, respectively (Milne et al., 2002; Nakamura et al., 2002; Tanaka

et al., 2007). Studies in *Drosophila* showed that Trx and Ash1 play a critical role in antagonizing H3K27 trimethylation by PRC2, suggesting a crosstalk between repressive and activating marks (Papp and Müller, 2006; Srinivasan et al., 2008).

In this study we investigated how PRC2 activity is modulated by chromatin marks typically associated with active transcription. We found that the Nurf55 WD40 propeller binds the N terminus of unmodified histone H3 and that H3K4me3 prevents this binding. In the context of the tetrameric PRC2 complex, we find that H3K4me3 and H3K36me2/3 (i.e., both H3-K36me2 and H3-K36me3) inhibit histone methylation by PRC2 in vitro. Dissection of this process by using fly, human, and plant PRC2 complexes suggests that the Su(z)12 subunit is important for mediating this inhibition. PRC2 thus not only contains the enzymatic activity for H3K27 methylation and a recognition site for binding to this modification, but it also harbors a control module that triggers inhibition of this activity to prevent deposition of H3K27 trimethylation on transcriptionally active genes. PRC2 can thus integrate information provided by pre-existing histone modifications to accurately tune its enzymatic activity within a particular chromatin context.

RESULTS

Structure of Nurf55 Bound to the N Terminus of Histone H3

Previous studies reported that Nurf55 alone is able to bind to histone H3 (Beisel et al., 2002; Hansen et al., 2008; Song et al., 2008; Wysocka et al., 2006) but not to a GST-H3 fusion protein (Verreault et al., 1998). By using fluorescence polarization (FP) measurements, we found that Nurf55 binds the very N terminus of unmodified histone H3 encompassing residues 1–15 ($H3_{1-15}$) with a K_D of ~0.8 ± 0.1 μM but does not bind to a histone $H3_{19-38}$ peptide (Figure 1A). Crystallographic screening resulted in the successful cocrystallization of Nurf55 in complex with an $H3_{1-19}$ peptide. After molecular replacement with the known structure of Nurf55 (Song et al., 2008), the initial $mF_o - DF_c$ difference map showed density for H3 residues 1–14 in both Nurf55 molecules in the crystallographic asymmetric unit. Figures 1B–1E show the structure of $H3_{1-19}$ bound to *Drosophila* Nurf55, refined to 2.7 Å resolution (R/R_{free} = 20.1% and 25.0%, Table 1; Figure S1A, available online). The H3 peptide binds to the flat surface of the Nurf55 WD40 propeller (Figure 1B), subsequently referred to as the canonical binding site (*c*-site) (Gaudet et al., 1996). The H3 peptide is held in an acidic pocket (Figures 1C and 1E) and traverses the central WD40 cavity in a straight line across the propeller (Figure 1B).

Nurf55 binds the H3 peptide by contacting H3 residues Ala1, Arg2, Lys4, Ala7, and Lys9. Each of these residues forms side-chain specific contacts with the Nurf55 propeller (Figures 1D and 1E). The bulk of the molecular recognition is directed toward H3 Arg2 and Lys4. Ala1 sits in a buried pocket with its α-amino group hydrogen bonding to Nurf55 Asp252, which recognizes and fixes the very N terminus of histone H3. The neighboring Arg2 is buried deeper within the WD40 propeller fold, with its guanidinium group sandwiched by Nurf55 residues Phe325 and Tyr185 (Figure 1D). H3 Lys4 binds to a well-defined surface pocket on Nurf55 located on blade 2, near the central cavity of the propeller. Its ε-amino group is specifically coordinated by the carboxyl groups of Nurf55 residues Glu183 and Glu130 and through the amide oxygen of Asn132 (Figure 1E). Lys9 is stabilized by hydrophobic interactions on the WD40 surface while having its ε-amino group held in solvent-exposed fashion (Figure 1D). Ser10 of histone H3 marks the beginning of a turn that inverses the peptide directionality. Histone H3 residues Thr11–Lys14 become progressively disordered and are no longer specifically recognized. No interpretable density was observed beyond Lys14. Taken together, Nurf55 specifically recognizes an extended region of the extreme N terminus of histone H3 (11 residues long, 700 $Å^2$ buried surface area) in the canonical ligand binding location of WD40 propeller domains.

Structure of the Nurf55-Su(z)12 Subcomplex of PRC2

The H3-Nurf55 structure prompted us to investigate how Nurf55 might bind histone tails in the presence of Su(z)12, its interaction partner in PRC2 (Nekrasov et al., 2005; Pasini et al., 2004). As a first step we mapped the Nurf55-Su(z)12 interaction in detail by carrying out limited proteolysis experiments on reconstituted *Drosophila* PRC2, followed by isolation of a Nurf55-Su(z)12 subcomplex. Mass spectrometric analysis and pull-down experiments with recombinant protein identified Su(z)12 residues 73–143 [hereafter referred to as $Su(z)12_{73-143}$] as sufficient for Nurf55 binding (Figures S1C and S1D).

Crystals were obtained when *Drosophila* Nurf55 and Su(z)12 residues 64–359 were set up in the presence of 0.01% subtilisin protease (Dong et al., 2007). After data collection, the structure was refined to a maximal resolution of 2.3 Å (Table 1). Molecular replacement with Nurf55 as search model provided clear initial $mF_o - DF_c$ difference density for a 13 amino acid-long Su(z)12 fragment spanning Su(z)12 residues 79–91 (Figures 2A–2C). The final model was refined to 2.3 Å (R/R_{free} = 17.5%/20.9%) and verified by simulated annealing composite-omit maps (Figure S1B). The portion of Su(z)12 involved in Nurf55 binding will henceforth be referred to as the Nurf55 binding epitope (NBE). The Su(z)12 binding site on Nurf55 is located on the side of the propeller between the stem of the N-terminal α helix (α1) and the PP loop (Figures 2A and2B). Binding between Su(z)12 and Nurf55 occurs mostly through hydrophobic interactions in an extended conformation. The interaction surface between Nurf55 and the NBE is large for a peptide, spanning around 800 $Å^2$. Sequence alignment between Su(z)12 orthologs reveals that the NBE is highly conserved (53% identity and 84% similarity) in animals and in plants (Figure 2E). With the exception of Su(z)12 Arg85, the majority of the conserved Su(z)12 NBE residues engage in hydrophobic packing with Nurf55 (Figures 2B and 2C). Together with the Su(z)12 VEFS domain and the C_2H_2 zinc finger (C5 domain) (Birve et al., 2001), the NBE constitutes the only identifiable motif in Su(z)12 found conserved in all Su(z)12 orthologs.

The NBE binding site on Nurf55 has previously been shown to be occupied by helix 1 of histone H4 (Figure 2D) (Murzina et al., 2008; Song et al., 2008), an epitope not accessible in assembled nucleosomes (Luger et al., 1997). Nurf55 binds H4 and the Su(z)12 NBE epitope in a different mode, and importantly, with opposite directionality (Figure 2D). The detailed comparison of the Nurf55-Su(z)12 structure with that of H4 bound to Nurf55

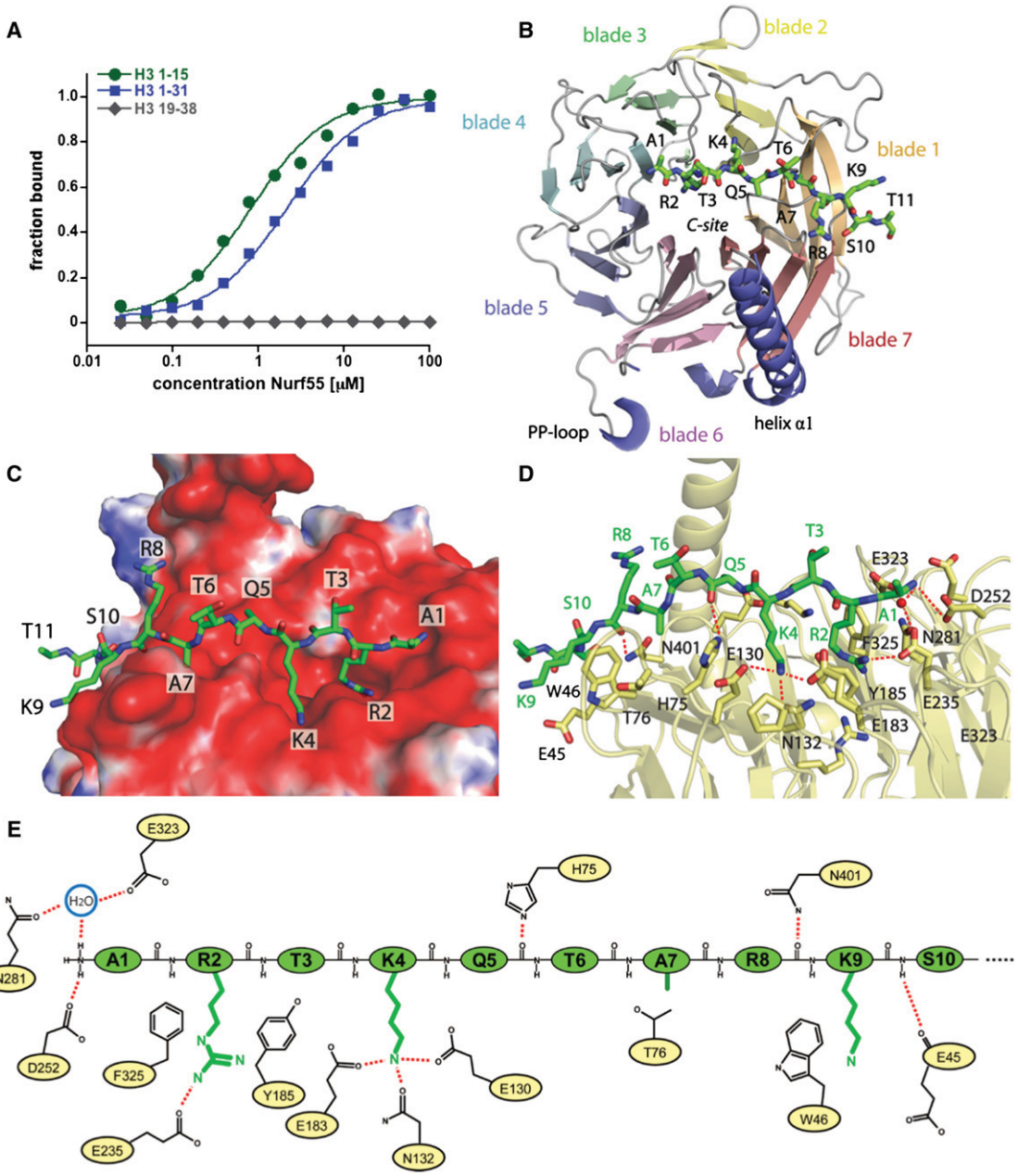


Figure 1. Crystal Structure of Nurf55 in Complex with a Histone $H3_{1–19}$ Peptide

(A) Nurf55 binds to an $H3_{1–15}$ peptide with an affinity of ~0.8 ± 0.1 μM as measured by FP. It has similar affinity for an $H3_{1–31}$ peptide (2.2 ± 0.2 μM) but no binding can be detected to an $H3_{19–38}$ peptide.

(B) Ribbon representation of Nurf55-$H3_{1–19}$. Nurf55 is shown in rainbow colors and $H3_{1–19}$ is depicted in green. The peptide is bound to the *c*-site of the WD40 propeller.

(C) Electrostatic surface potential representation (−10 to 10 kT/e) of the *c*-site with the H3 peptide shown as a stick model in green.

(D) Close-up of the *c*-site detailing the interactions between Nurf55 (yellow) and the $H3_{1–19}$ peptide (green), with a water molecule shown as a red sphere.

(E) Schematic representation of interactions between the $H3_{1–19}$ peptide (green) and Nurf55 (yellow).

strongly suggests that binding of Su(z)12 (NBE) and of H4 (helix 1) are mutually exclusive (Figure 2D). We therefore refer to the Su(z)12 and H4 binding site on Nurf55 as the *S/H*-site.

Su(z)12 fragments that include the NBE have poor solubility by themselves and generally require Nurf55 coexpression for solubilization. However, we were able to measure binding of a chemically synthesized $Su(z)12_{75–93}$ peptide to Nurf55 by isothermal titration calorimetry (ITC) and found that the peptide was bound with a K_D value of 6.7 ± 0.3 μM in a 1:1 stoichiometry (Figure 2F). Pull-down experiments with recombinant protein and

Table 1. Crystallographic Data and Refinement Statistics

	Nurf55 – Su(z)12	Nurf55 – $H3_{1–19}$
Space Group	$P2_12_12_1$	$P2_12_12_1$
Unit Cell Dimensions		
a, *b*, *c* (Å)	53.03, 87.19, 99.54	55.97, 88.15, 204.02
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	25.0 – 2.3 (2.38-2.30)[a]	37.7 – 2.7 (2.8-2.7)[a]
Percent complete	99.2 (92.0)[a]	96.6 (90.6)[a]
Redundancy	15.6 (12.4)[a]	6.4 (5.3)[a]
R_{sym}	0.074 (0.338)[a]	0.091 (0.428)[a]
$I/\sigma I$	43.3 (6.2)[a]	15.6 (4.6)[a]
Resolution (Å)	2.3	2.7
Number of reflections	20984	28271
R_{work}/R_{free}	0.175/0.209	0.201/0.250
Number of atoms	3315	6197
B-Factors	29.8	46.9
Protein	29	47.0
Water	38.3	38.1
RMS Deviations		
Bond lengths (Å)	0.008	0.003
Bond angles (°)	1.124	0.752

[a] The values for the data in the highest resolution shell are shown in parentheses.

streptavidin beads suggest that Su(z)12 residues 94–143 harbor an additional Nurf55 binding site not visible in the structure (Figure S1E). $Su(z)12_{144–359}$, lacking the N-terminal 143 residues, no longer binds to Nurf55. The NBE (residues 79–93) and the region adjacent to the NBE (residues 94–143) are thus required for stable interaction with Nurf55. The extended NBE was found enriched after limited proteolysis and in subsequent gel filtration runs coupled with quantitative mass spectrometry (Figure S1C). As the NBE was the only fragment visible after structure determination, we conclude that it represents the major Su(z)12 interaction epitope for Nurf55 binding.

The Nurf55-Su(z)12 Complex Binds to Histone H3

In order to study the potential interdependence of the identified Nurf55 binding sites we compared binding of Nurf55 and Nurf55-Su(z)12 to the histone H3 N terminus. FP experiments showed similar affinities for binding of a histone $H3_{1–15}$ peptide to Nurf55 (K_D ~0.8 ± 0.1 μM; Figure 1A) and a Nurf55-$Su(z)12_{73–143}$ complex (K_D ~0.6 ± 0.1 μM; Figure 2G). Importantly, mutation of Nurf55 residues contacting H3 via its *c*-site drastically reduced binding to an $H3_{1–15}$ peptide (Figure S2A), demonstrating that the Nurf55-$Su(z)12_{73–143}$ complex indeed binds the $H3_{1–15}$ peptide through the *c*-site. We conclude that the presence of Su(z)12 is compatible with Nurf55 binding to H3 via its *c*-site and that the two binding interactions are not interdependent.

The observation that the Su(z)12 NBE occupies the same Nurf55 pocket that was previously shown to bind to helix 1 of histone H4 prompted us to test whether the $Su(z)12_{73–143}$-Nurf55 complex could still bind to histone H4. We performed pull-down experiments with a glutathione S-transferase (GST) fusion protein containing histone $H4_{1–48}$ (Murzina et al., 2008) and found that H4 stably interacted with isolated Nurf55 but not with $Su(z)12_{73–143}$-Nurf55 (Figure 2H). In PRC2, the presence of Su(z)12 in the Nurf55 *S/H*-site therefore precludes binding to helix 1 of histone H4.

H3 Binding by Nurf55-Su(z)12 Is Sensitive to the Methylation Status of Lysine 4

We next investigated how posttranslational modifications of the H3 tail affect binding to the Nurf55-$Su(z)12_{73–143}$ complex. Modifications on H3 Arg2, Lys9, and Lys14 did not change affinity of Nurf55-Su(z)12 for the modified $H3_{1–15}$ peptide (Figures S2B and S2D). In contrast, peptides that were mono-, di-, or trimethylated on Lys4 were bound with significantly reduced affinity exhibiting K_D values of 17 ± 3 μM (H3K4me1), 24 ± 3 μM (H3K4me2), and >70 μM (H3K4me3), respectively (Figure 2I). The FP binding data were independently confirmed by ITC measurements (Figures S2C–S2F). Together, these findings are in accord with the structural data, which show that H3K9 and H3K14 are being held with their ε-amino moiety solvent-exposed, while the H3K4 side chain is tightly coordinated (Figure 1E). The additional methyl groups on the H3K4 ε-amino group are expected to progressively decrease affinity because of increased steric clashes within the H3K4 binding pocket.

H3K27 Methylation by PRC2 Is Inhibited by Histone H3K4me3 Marks

We then examined the effect of H3K4me3 modifications, which are no longer retained by Nurf55-Su(z)12, on the catalytic activity of PRC2. In a first set of experiments, we determined PRC2 steady-state parameters on histone $H3_{1–45}$ peptide substrates that were either unmodified or methylated at Lys 4. We observed similar K_M values for H3 and H3K4me3 peptides of 0.84 ± 0.21 μM and 0.36 ± 0.07 μM, respectively (Figure 3A), and similar K_M values for SAM (5.42 ± 0.65 μM for H3 and 10.04 ± 1.56 μM for H3K4me3). The turnover rate constant k_{cat}, however, was 8-fold reduced in the presence of H3K4me3: 2.53 ± 0.21 min^{-1} for unmodified H3 and 0.32 ± 0.08 min^{-1} in the presence of H3K4me3 (Figure 3A). While substrate binding is largely unaffected, turnover is thus severely inhibited in the presence of H3K4me3. This behavior, which results in a k_{cat}/K_M specificity constant of 7.8 × 10^3 $M^{-1}s^{-1}$ (unmodified H3) compared to 0.53 × 10^3 $M^{-1}s^{-1}$ (H3K4me3), is consistent with heterotrophic allosteric inhibition of the PRC2 HMTase triggered by the presence of the H3K4me3.

To investigate the effect of the H3K4me3 modification on PRC2 activity in the context of nucleosomes, we reconstituted mononucleosomes with a trimethyllysine analog (MLA) at Lys4 in H3 (referred to as H3Kc4me3; Figure S3A) (Simon et al., 2007). We found that total H3K27 methylation (measured by incorporation of ^{14}C-labeled methyl groups) was substantially impaired on H3Kc4me3-containing nucleosomes compared to wild-type nucleosomes (Figures S3B and S3C). We used western blot analysis to monitor how levels of H3K27 mono-, di-, and trimethylation were affected by the H3Kc4me3 modification. While H3K27me1 formation was reduced by more than 50% on H3Kc4me3 nucleosomes compared to unmodified nucleosomes (Figures 3B and 3C), H3K27 dimethylation and

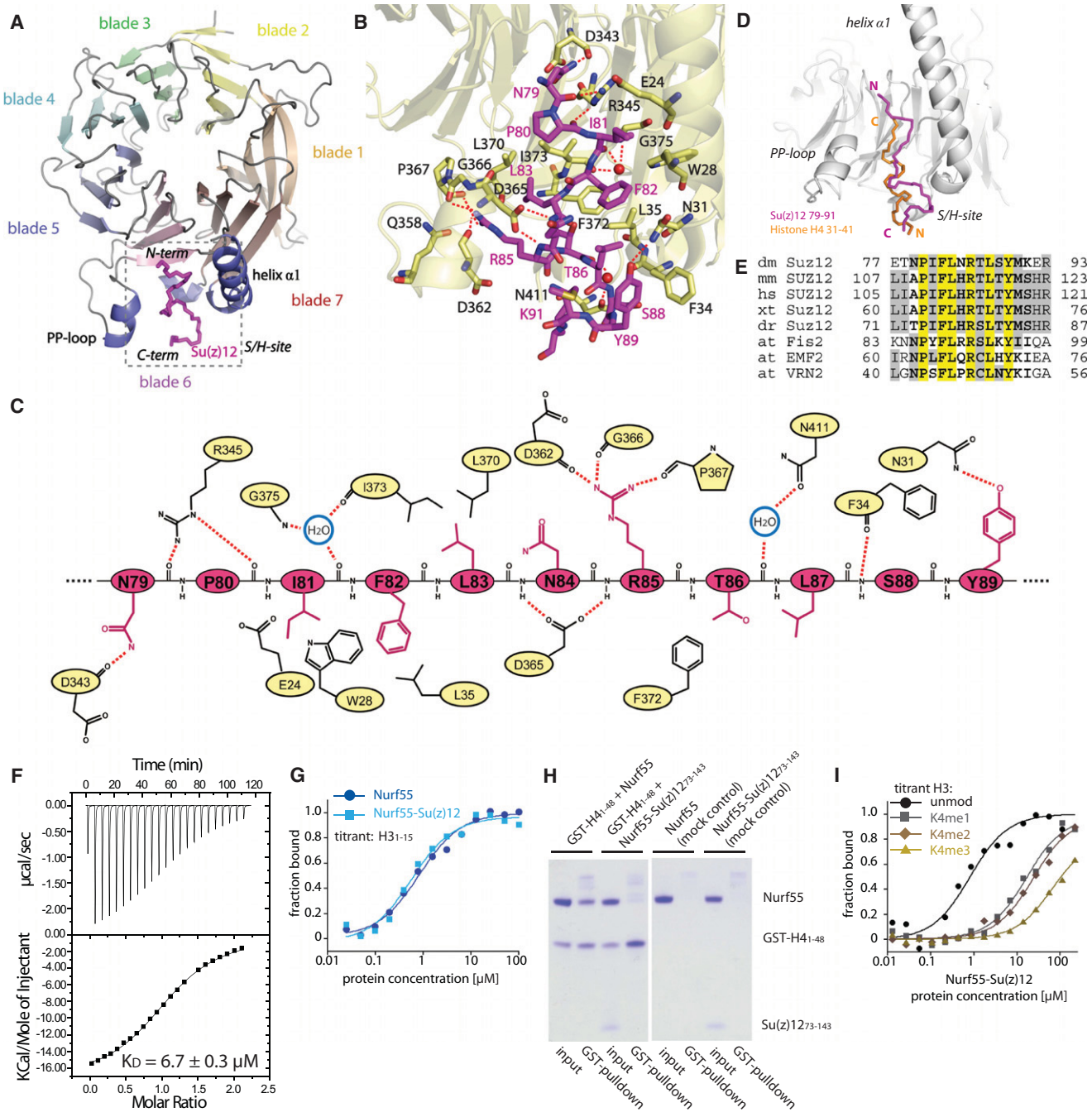


Figure 2. Crystal Structure and Characterization of Nurf55 in Complex with the Su(z)12 Binding Epitope for Nurf55

(A) Ribbon representation of Nurf55-Su(z)12. Nurf55 (rainbow colors) depicts the WD40 domain nomenclature and Su(z)12 is shown in magenta. The *S/H*-site is marked by a dashed box.

(B) Detailed interactions of Su(z)12 (magenta) with the *S/H*-site (yellow). Water molecules are depicted as red spheres.

(C) Schematic representation of interactions between Su(z)12 (magenta) and Nurf55 (yellow).

(D) Overlay of the backbone trace of Su(z)12 (magenta) and the H4 helix α1 (orange) (Song et al., 2008) in the *S/H*-site.

(E) Alignment of the Su(z)12 NBE with sequences from *Drosophila melanogaster* (dm, Q9NJG9), mouse (mm, NP_954666), human (hs, AAH15704), *Xenopus tropicalis* (xt, BC121323), zebrafish (dr, BC078293), and the three *Arabidopsis thaliana* (at) homologs Fis2 (ABB84250), EMF2 (NP_199936), and VRN2 (NP_567517). Identical residues are highlighted in yellow.

(F) ITC profile for binding of a $Su(z)12_{75–93}$ peptide to Nurf55. Data were fitted to a one-site model with stoichiometry of 1:1. The derived K_D value is 6.7 ± 0.3 μM.

(G) Binding of $H3_{1–15}$ to Nurf55 (0.8 ± 0.1 μM) and Nurf55-$Su(z)12_{73–143}$ (0.6 ± 0.1 μM) measured by FP.

A

Substrate	Apparent Km of SAM (μM)	Apparent Km of peptide (nM)	k_{cat} (min^{-1})	k_{cat}/Km ($M^{-1}S^{-1}$)
$H3_{1-45}$ -biotin	5.42 ± 0.65	836 ± 207	2.53 ± 0.21	7.8×10^3
$H3K4me3_{1-45}$ -biotin	10.04 ± 1.56	355 ± 74.6	0.32 ± 0.08	0.53×10^3

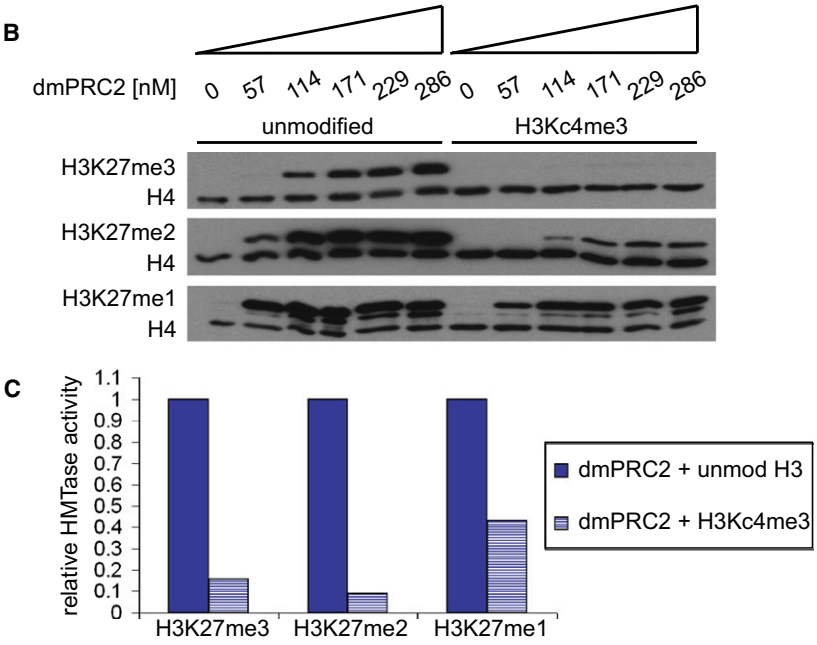


Figure 3. HMTase Activity of PRC2 Is Inhibited by H3K4me3 Marks

(A) HMTase assay with PRC2 and $H3_{1-45}$-biotin peptides measuring the concentration of SAH produced by the enzymatic reaction. When an H3K4me3-modified peptide is used, the specificity constant (k_{cat}/K_M) is drastically reduced, indicative of heterotrophic allosteric inhibition.

(B) Western blot-based HMTase assay by using recombinant *Drosophila* mononucleosomes (571 nM) and increasing amounts of PRC2. HMTase activity was monitored with antibodies against H3K27me1, H3K27me2, or H3K27me3 as indicated; in each case the membrane was also probed with an antibody against unmodified histone H4 to control for equal loading and western blot processing. Deposition of K27 di- and trimethylation is drastically reduced when nucleosomes are used that carry a H3Kc4me3 modification.

(C) Quantification of HMTase activity of *Drosophila* PRC2 (286 nM) on unmodified and H3Kc4me3-modified nucleosomes by quantitative western blotting.

trimethylation were impaired by more than 80% by using H3Kc4me3 nucleosomes (Figure 3C). In order to ascertain that inhibition of PRC2 is indeed due to trimethylation of the amino group in the lysine side chain, and not due to the use of the MLA, we performed HMTase assays on H3K4me3-containing nucleosomes generated by native peptide ligation (Shogren-Knaak et al., 2003) and on H3Kc4me0 and H3K4A nucleosomes. H3K27 mono-, di-, and trimethylation was comparably inhibited on H3K4me3 and on H3Kc4me3-containing nucleosomes, but was not affected by H3Kc4me0 and H3K4A (Figures S3D and S3E). We conclude that H3K4me3 specifically inhibits PRC2-mediated H3K27 methylation with the most pronounced inhibitory effects observed for H3K27 di- and trimethylation.

We next tested whether the H3K4me3 modification affects PRC2 nucleosome binding. In electrophoretic mobility shift assays (EMSA), we found that PRC2 binds unmodified or H3Kc4me3-modified nucleosomes with comparable affinity (Figure S4A). Even though binding of Nurf55 to the N terminus of histone H3 is almost 100-fold reduced by H3K4me3 (Figure 2I and Figure S2F), interaction of the Nurf55 *c*-site with H3K4 does not seem to make a detectable contribution to nucleosome binding by PRC2 in this assay. Consistent with the allosteric mechanism of H3K4me3 inhibition that we had observed in the peptide assays (Figure 3A), inhibition of the PRC2 HMTase activity by H3K4me3-containing nucleosomes is not caused by impaired nucleosome binding, but is rather the consequence of reduced catalytic turnover.

H3K4me3 Needs to Be Present on the Same Tail as K27 to Inhibit PRC2

We then assessed whether inhibition of the PRC2 HMTase activity by H3K4me3 requires the K4me3 mark to be located on the substrate nucleosome (in *cis*), or whether it could also be triggered if the H3K4me3 modification was provided on a separate peptide (in *trans*). We performed HMTase assays on unmodified oligonucleosomes in the presence of increasing amounts of a histone $H3_{1-15}$ peptide trimethylated at K4 ($H3_{1-15}$-K4me3) (Figure 4A). Addition of the $H3_{1-15}$-K4me3 peptide did not affect PRC2 HMTase activity at peptide concentrations as high as ~200 μM. When testing $H3_{1-19}$-unmodified peptide in controls at comparable concentrations, we did observe concentration-dependent PRC2 inhibition (Figure 4A), probably because of substrate competition at large peptide excess. As H3K4me3-

(H) GST pull-down assay with recombinant GST-$H4_{1-48}$ and Nurf55 and Nurf55-Su(z)12_{73-143} proteins. GST-$H4_{1-48}$ is able to bind Nurf55 alone but in the Nurf55-Su(z)12_{73-143} complex the binding site is occupied by Su(z)12 (left panel). Control pull-downs with GST beads and either Nurf55 or Nurf55-Su(z)12_{73-143} alone showed no unspecific binding (right panel).

(I) Binding of different $H3_{1-15}$ peptides to Nurf55-Su(z)12_{73-143} measured by FP. While unmodified H3 is bound with 0.8 ± 0.1 μM affinity, methylation of Lys 4 drastically reduces binding affinity (17 ± 3 μM for K4me1, 24 ± 3 μM for K4me2, and >70 μM for K4me3).

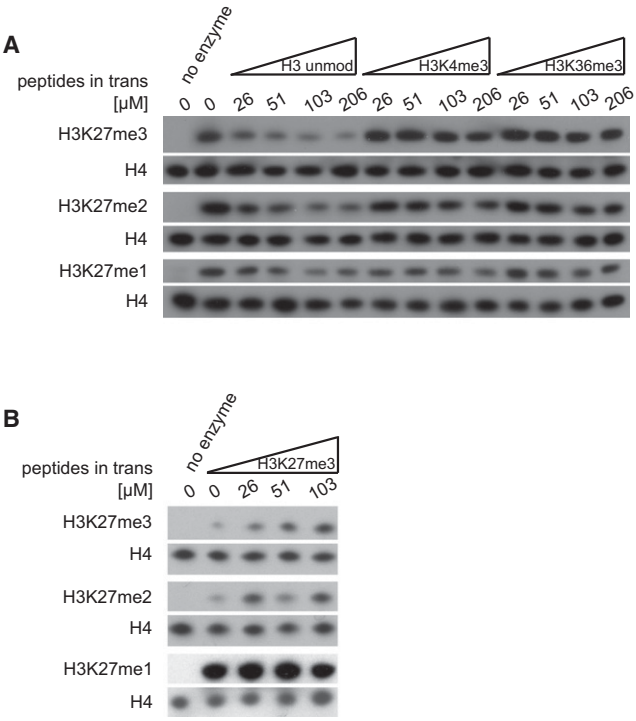


Figure 4. PRC2 Activity Is Not Inhibited by H3K4me3 Peptides in *trans*

(A) Western blot-based HMTase assay by using unmodified 4-mer oligonucleosomes (36 nM) and increasing amounts of H3 peptides added in *trans*. Enzyme concentration was kept constant at 86 nM. Western blots were processed as described in Figure 3B. HMTase activity is inhibited by an unmodified $H3_{1-19}$ peptide (left), but not by H3K4me3- or H3K36me3-modified peptides.

(B) HMTase assay with H3Kc4me3-modified oligonucleosomes (36 nM), 86 nM PRC2, and H3K27me3 peptide in *trans*. Western blots were processed as described in Figure 3B. HMTase activity of PRC2 can be stimulated by the H3K27me3 peptide even on inhibiting substrate leading to increased levels of H3K27 di- and trimethylation.

modified peptides did not show this competitive behavior, we conclude that PRC2 is not inhibited by H3K4me3 in *trans* and that H3K4me3 and unmodified H3 peptides are probably bound to PRC2 in a different fashion. Analogously, we saw no inhibition when testing the effect of H3K4me3 in *trans* by using peptides as substrates (Figure S4B**)**. Taken together, our findings strongly argue that H3K4me3 only inhibits PRC2 if present on the same tail that contains the H3K27 target lysine (in *cis*).

Previous studies reported that addition of H3K27me3 peptides in *trans* enhances H3K27 methylation of oligonucleosomes by human PRC2 through binding to the EED WD40 domain (Margueron et al., 2009; Xu et al., 2010). We tested whether addition of H3K27me3 peptides in *trans* would stimulate H3K27 methylation by PRC2 on H3Kc4me3-modified nucleosomes. We observed that the inhibitory effect of H3Kc4me3-containing nucleosomes can, at least in part, be overcome through addition of high concentrations of H3K27me3 peptides (Figure 4B and Figure S4C). PRC2 is therefore able to simultaneously integrate inhibitory (H3K4me3) and activating (H3K27me3) chromatin signatures and adjust its enzymatic activity in response to the surrounding epigenetic environment.

PRC2 Inhibition of H3K4me3 Is Conserved in Mammalian PRC2

Our results with *Drosophila* PRC2 prompted us to investigate to what extent inhibition by H3K4me3 is an evolutionarily conserved mechanism. H3K27 methylation by human and mouse PRC2 on nucleosome substrates carrying H3Kc4me3 modifications was also strongly inhibited, comparable to the inhibition observed for *Drosophila* PRC2 (Figure 5A and Figure S5A).

The Su(z)12 Subunit Codetermines Whether PRC2 Is Inhibited by H3K4me3

In *Arabidopsis thaliana*, three different E(z) homologs combined with three Su(z)12 homologs have been described. The distinct PRC2 complexes in plants harboring the different E(z) or Su(z)12 subunits are implicated in the control of distinct developmental processes during *Arabidopsis* development (He, 2009). In this study we focused on PRC2 complexes containing the E(z) homolog CURLY LEAF (CLF). We expressed and reconstituted the *Arabidopsis* PRC2 complex comprising CLF, FERTILIZATION INDEPENDENT ENDOSPERM (FIE, a homolog of ESC), EMBRYONIC FLOWER 2 (EMF2, a homolog of Su(z)12), and MULTICOPY SUPPRESSOR OF IRA (MSI1, a homolog of Nurf55). We found that CLF indeed functions as a H3K27me3 HMTase (Figure 5B). Moreover, H3K27 methylation by the CLF-FIE-EMF2-MSI1 complex on nucleosome arrays containing H3Kc4me3 was inhibited (Figure 5B) in a manner comparable to human or *Drosophila* PRC2.

We next tested a related *Arabidopsis* PRC2 complex again composed of CLF, FIE, and MSI1 but containing the Su(z)12 homolog vernalization 2 (VRN2) instead of EMF2. The VRN2 protein is specifically implicated as a repressor of the FLC locus, thereby controlling flowering time in response to vernalization (reviewed in Henderson and Dean, 2004). The CLF-FIE-MSI1-VRN2 complex was active on unmodified nucleosomes but, strikingly, it was not inhibited on H3Kc4me3-modified nucleosomes (Figure 5C). Substitution of a single subunit (i.e., EMF2 by VRN2) thus renders the complex nonresponsive to the H3K4 methylation state. While PRC2 inhibition by H3K4me3 appears hardwired in mammals and flies, in which only a single-Su(z)12 ortholog is present, *Arabidopsis* inhibition can be enabled or disabled through exchange of the Su(z)12 homolog.

The Su(z)12 C Terminus Harboring the VEFS Domain Is the Minimal Su(z)12 Domain Required for Activation and Active Mark Inhibition

The importance of the Su(z)12 subunit in active mark H3K4me3 inhibition prompted us to map the Su(z)12 domains required for inhibition. Previous findings showed that E(z) or E(z)-ESC in the absence of Su(z)12 is enzymatically inactive (Nekrasov et al., 2005). Moreover, the *VEFS* domain (Birve et al., 2001) was found to be the major E(z) binding domain (Ketel et al., 2005). We reconstituted mouse PRC2 complexes containing EZH2, EED, and either SUZ12 C_2H_2 domain + VEFS (residues 439–741) or SUZ12 VEFS alone (residues 552–741). Both of these minimal complexes were active in HMTase assays on nucleosomes (Figure 5D) but with lower activity than that of the full PRC2 complex. We therefore focused on formation of

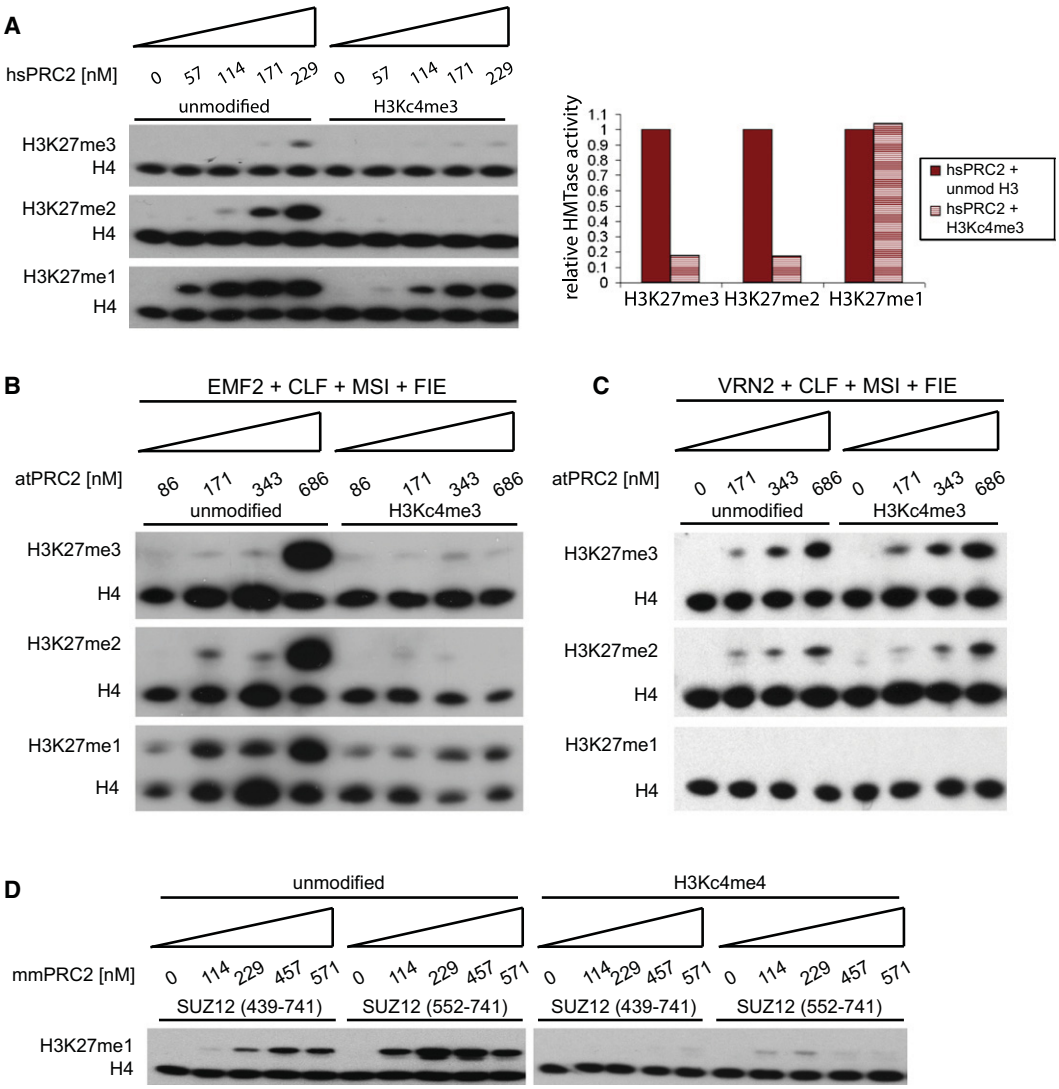


Figure 5. Sensitivity of PRC2 toward H3K4me3 Is Conserved in Mammals and Plants

(A) Left panel: Western blot-based HMTase assay with human PRC2 complex on unmodified and H3Kc4me3-modified mononucleosomes (571 nM). Western blots were processed as described in Figure 3B. Similar to the fly complex, human PRC2 is less active on H3Kc4me3 nucleosomes and H3K27 di- and trimethylation is severely hampered. Right panel: Quantification of HMTase activity of human PRC2 (286 nM) on unmodified and H3Kc4me3-modified nucleosomes by quantitative western blotting.

(B) HMTase assay with a plant PRC2 complex composed of EMF2, CLF, MSI1, and FIE on 4-mer oligonucleosomes (36 nM). Western blots were processed as described in Figure 3B. This complex is also sensitive to H3Kc4me3 and H3K27 di- and trimethylation on modified nucleosomes is severely reduced.

(C) HMTase assay with a plant PRC2 complex composed of VRN2, CLF, MSI1, and FIE on 4-mer oligonucleosomes (36 nM). Western blots were processed as described in Figure 3B. In contrast to the EMF2-containing complex, the VRN2 complex is insensitive toward H3Kc4me3 and has comparable activity on unmodified and modified nucleosomes.

(D) HMTase assay with mouse EZH2-EED-SUZ12 complexes containing truncated SUZ12 constructs (439–741 and 552–741) on 4-mer oligonucleosomes (36 nM). Western blots were processed as described in Figure 3B. Only the monomethylation signal is shown because the truncated SUZ12 complexes have poor HMTase activity. Both complexes remain inhibited by the H3Kc4me3 modification.

H3K27me1 as the strongest readout. When testing H3Kc4me3-containing nucleosomes, we found that both SUZ12 fragments rendered the complex sensitive to H3K4me3, resulting in the inhibition of H3K27 monomethylation (Figure 5D). The SUZ12 C terminus (including the VEFS domain) is thus the minimal SUZ12 fragment required for PRC2 binding, HMTase activity, and inhibition by H3K4me3.

Transcriptionally Active H3K36me2/3 Nucleosome Marks Also Inhibit PRC2

Finally, we asked whether other methylation marks associated with transcriptionally active chromatin but located outside the Nurf55-Su(z)12 binding site inhibit PRC2. We found that H3K27 methylation by PRC2 was also inhibited on H3Kc36me3 or H3Kc36me2 nucleosomes (Figure 6 and Figure S6). No

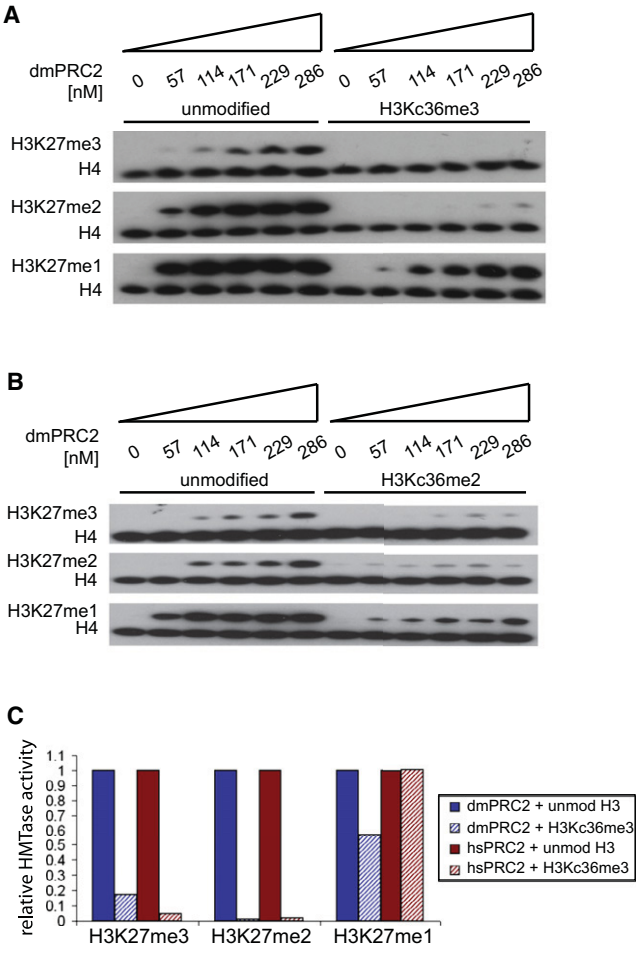


Figure 6. PRC2 Activity Is Also Inhibited by H3K36me2/3

(A) Western blot-based HMTase assay with *Drosophila* PRC2 complex on unmodified and H3Kc36me3-modified mononucleosomes (571 nM). Western blots were processed as described in Figure 3B. Similar to H3Kc4me3, H3Kc36me3 also inhibits the PRC2 complex.

(B) Same as in (A), western blot-based HMTase assay with *Drosophila* PRC2 complex, but with H3Kc36me2-modified nucleosomes.

(C) Quantitative western blot analysis estimating the HMTase activity of *Drosophila* (blue) and human (red) PRC2 complex (286 nM) on unmodified and H3Kc36me3-modified mononucleosomes.

significant inhibition was observed on control nucleosomes containing H3K36A or H3Kc36me0 lysine analog control (Figure S3E). As in the case of H3K4me3, inhibition by H3K36me2/3 could not be triggered by addition of the modified peptides, but required the modification to be present on the nucleosome containing the Lys27 substrate site (in *cis*) (Figure 4A). Moreover, we found that fly and human PRC2 were inhibited comparably. In the case of *Arabidopsis* PRC2, EMF2-containing complexes were inhibited by H3K36me3, whereas VRN2-containing complexes were not (Figure S6).

In conclusion, we find that three distinct histone methylation marks that are present in the coding region of actively transcribed genes—H3K4me3, H3K36me2, and H3K36me3—act as universal inhibitors of PRC2 complexes in both animals and plants.

DISCUSSION

Understanding how histone modification patterns are propagated during cell division is essential for understanding the molecular basis of epigenetic inheritance. Trimethylation of H3K27 by PRC2 has emerged as a key step in generating transcriptionally repressed chromatin in animals and plants. This study investigates how PRC2 recognizes the H3 tail and responds to H3-associated marks of active chromatin. Our crystallographic analyses reveal the molecular basis for $H3_{1-14}$ recognition by the Nurf55-Su(z)12 module of PRC2 and demonstrate that H3 tails carrying K4me3 are no longer recognized by Nurf55-Su(z)12. In the context of the whole PRC2 complex, H3K4me3 triggers allosteric inhibition of PRC2, a process that requires H3K4me3 to be present on the same histone molecule containing the substrate Lys27. We also observed PRC2 inhibition by H3K36me2/3. PRC2 inhibition by active chromatin marks (H3K4me3 and H3K36me2/3) is conserved in PRC2 complexes reconstituted from humans, mouse, flies, and plants.

The Role of Nurf55 in PRC2

Minimal PRC2 complexes lacking Nurf55 retain partial catalytic activity and are inhibited by H3K4me3 (Figure 5D). H3K4me3, once free of Nurf55, is thereby able to trigger PRC2 inhibition. We favor a model where Nurf55-Su(z)12 serves in sequestration and release of histone H3. We propose that the release of the H3 tail from Nurf55-Su(z)12 is required, but not sufficient, to induce H3K4me3 inhibition as it needs to trigger allosteric inhibition in conjunction with Su(z)12 and the E(z) SET domain (see below). Unmodified H3, H3K9me3, or H3R2me-modified tails, on the other hand, remain sequestered and are shielded from other chromatin factors. These sequestered marks are also not expected to interfere with PRC2 regulation. In line with this prediction, we observed that H3K9me3, which remained bound to Nurf55-Su(z)12, also did not interfere with PRC2 activity in vitro (Figure S5C). In vitro, binding of Nurf55 to the N terminus of H3 was not critical for the overall nucleosome binding affinity of PRC2 under our assay conditions (Figures S4A and S5B). However, small differences in PRC2 affinity amplified by large chromatin arrays could skew PRC2 recruitment toward sites of unmodified H3K4. Additionally, the Nurf55 interaction might play a more subtle role in positioning the complex correctly on nucleosomes.

Su(z)12 Mediates Inhibition by Active Marks in Conjunction with the E(z) SET Domain

Our in vitro findings suggest that active chromatin mark inhibition by PRC2 is largely governed through allosteric inhibition of the PRC2 HMTase activity thereby limiting processivity of the enzyme. We defined a minimal trimeric PRC2 subcomplex that retains both activity and H3K4me3/H3K36me2/3 inhibition. This minimal complex consists of ESC, an E(z) fragment that comprises the ESC binding region at the N terminus, the Su(z)12 binding domain in the middle (Hansen et al., 2008), and the C-terminal catalytic domain, and the Su(z)12 C terminus harboring the VEFS domain (Figure 5D). The importance of Su(z)12 is underlined by our findings on the *Arabidopsis* PRC2 complexes that revealed that active mark inhibition is determined by the choice of Su(z)12 subunit (i.e., inhibition with

EMF2, but not with VRN2) (Figures 5B and 5C). As the extent of methylation inhibition and the domains required for inhibition were similar for H3K4me3 and H3K36me2/3 (Figures 3 and 6), we hypothesize that both peptides function through a related mechanism allosterically affecting E(z) SET domain processivity with the help of Su(z)12. Further structural studies are required to reveal how these active marks are recognized and how this recognition is linked to inhibition of the E(z) SET domain.

PRC2 Mediates Crosstalk between Active and Repressive Histone Methylation Marks

H3K27me3 recognition by PRC2 has been reported to recruit and stimulate PRC2 (Hansen et al., 2008; Margueron et al., 2009), a mechanism implicated in creating and maintaining the extended H3K27me3 domains at target genes in vivo (Bernstein et al., 2006; Pan et al., 2007; Papp and Müller, 2006; Schwartz et al., 2006). Such positive feedback, however, necessitates a boundary element curtailing the expansion of H3K27me3. Our results suggest that actively transcribed genes (i.e., marked with H3K4me3 and H3K36me2/3) that flank domains of H3K27me3 chromatin may represent such boundary elements. In conjunction with H3K27me3-mediated stimulation, this provides a model how PRC2 could template domains of H3K27me3 chromatin during replication without expanding H3K27me3 domains into the chromatin of active genes. The inhibitory circuitry present in PRC2, however, does not function as a binary ON/OFF switch. PRC2 is able to integrate opposing H3K4me3 and H3K27me3 modifications into an intermediary H3K27 methylation activity (Figure 4B).

PRC2-Mediated Interdependence of H3K4me3/H3K36me3 versus H3K27me3 In Vivo

The crosstalk between H3K4me3 and H3K36me2/3 versus H3K27me3 has been extensively studied in vivo. Specifically, HOX genes in developing *Drosophila* larvae, or in mouse embryos, show mutually exclusive H3K27me3 and H3K4me3 domains that correlate with transcriptional OFF and ON states, respectively (Papp and Müller, 2006). In *Drosophila*, maintenance of HOX genes in the ON state critically depends on the trxG regulators Trx and Ash1, which methylate H3K4 and H3K36, respectively (Smith et al., 2004; Tanaka et al., 2007). At the *Ultrabithorax* (*Ubx*) gene, lack of Ash1 results in PRC2-dependent H3K27me3 deposition in the coding region of the normally active gene and the concomitant loss of *Ubx* transcription (Papp and Müller, 2006). Similarly, in the *Arabidopsis Flowering Locus C* (*FLC*), CLF-dependent deposition of H3K27me3 reduces H3K4me3 levels (Jiang et al., 2008), while deletion of the H3K4me3 demethylase *FLD* increases H3K4me3 levels and concomitantly diminishes H3K27me3 levels (Yu and Michaels, 2010). Our results provide a simple mechanistic explanation for these observations in plants and flies. We propose that H3K4 and H3K36 modifications in the coding region of active PcG target genes function as barriers that limit H3K27me3 deposition by PRC2.

Depositing Epigenetic Signatures-Knowing How to Stop?

We note that a number of HMTase complexes contain histone mark recognition domains that bind the very same mark that is deposited by their catalytic domain (Collins et al., 2008; Zhang et al., 2008; Chang et al., 2010). While this positive feedback loop guarantees the processivity of histone mark deposition, it also requires a control mechanism that avoids excessive spreading of marks. The direct inhibition of HMTases by histone marks, as seen for PRC2, may offer a paradigm of how excessive processivity can be counteracted in other HMTases.

Active Mark Inhibition Can Be Deactivated

Arabidopsis VRN2 is implicated in the control of the *FLC* locus after cold shock (reviewed in He, 2009). *FLC* is a bivalent locus containing both repressive H3K27me3 and active H3K4me3 marks (Jiang et al., 2008). In a VRN2-dependent fashion, H3K27me3 levels increase at *FLC* during vernalization. We find that while EMF2-containing PRC2 complexes are sensitive to H3K4me3 and H3K36me3, their VRN2-containing counterparts are not. In response to environmental stimuli plant PRC2 H3K4me3/H3K36me3 inhibition can thus be switched OFF (or ON). This offers the possibility that inhibition in animal PRC2 could also be modulated either by posttranslational modification of SUZ12 or by association with accessory factors.

PRC2 Inhibition and Coexistence of Active and Repressive Histone Methylation Marks

Quantitative mass spectrometry analyses of posttranslational modifications on the H3 N terminus in HeLa cells found no evidence for significant coexistence of H3K27me3 with H3K4me3 on the same H3 molecule (Young et al., 2009). Similarly, the fraction of H3 carrying both H3K27me3 and H3K36me3 was reported to be extremely low (~0.078%), while H3K27me3 and H3K36me2 coexist on ~1.315% of H3 molecules. However, H3K27me3/H3K4me3 and H3K27me3/H3K36me2/3 bivalent domains have been reported to exist in embryonic stem cells, and they have been implicated to exist on the same nucleosome (Bernstein et al., 2006). Given that PRC2 is inhibited by active methylation marks, how then could such bivalent domains be generated? We envisage two main possibilities. First, PRC2 inhibition in vivo could be alleviated by specific posttranslational modifications on PRC2 in embryonic stem cells (see in plants, VRN2). Second, H3K27me3 could be deposited prior to modification of H3K4 or H3K36. According to this view, one would have to postulate that the HMTases depositing H3K4me3 or H3K36me2/3 can work on nucleosomes containing H3K27me3. In support of this view, we found that H3K36 methylation by an NSD2 catalytic fragment is not inhibited by H3K27me3 marks on a peptide substrate (data not shown).

In summary, we found that mammalian and fly PRC2 complexes are not only activated by H3K27me3 as was recently reported (Margueron et al., 2009; Xu et al., 2010), but they are also inhibited by H3K4me3 and H3K36me2/3. PRC2, as a single biochemical entity, can thus integrate the information provided by histone modifications with antagonistic roles in gene regulation. While the biological network overseeing crosstalk between active and repressive chromatin marks in vivo probably extends beyond PRC2, including other chromatin modifiers such as histone demethylases (e.g., see Yokoyama et al., 2010), we identified a regulatory logic switch in PRC2 that intrinsically

separates active and repressive chromatin domains. Given the dynamic nature of the nucleosome template that makes up eukaryotic chromosomes, this circuitry probably equips PRC2 with the necessary precision to heritably propagate a repressed chromatin state.

EXPERIMENTAL PROCEDURES

Protein Purification and Crystallization

PRC2 proteins were recombinantly expressed and purified as described in Scrima et al. (2008). All crystals were grown at 20–25°C by the hanging drop vapor diffusion method. For Nurf55-Su(z)12 crystallization, a coexpressed $Nurf55_{1–418}$ and $Su(z)12_{73–143}$ complex at 16 mg/ml was incubated for 10 min with 0.01% subtilisin prior to setup. Subsequently, 1 μl drops of protein solution were mixed with 1 μl of reservoir solution containing 100 mM potassium acetate and 2.1 M ammonium sulfate. For Nurf55-$H3_{1–19}$ crystallization, Nurf55 protein was incubated with a ~5-fold molar excess of the $H3_{1–19}$ peptide for 30 min prior to crystallization. One microliter drops of a 22 mg/ml protein solution were mixed with 1 μl of reservoir solution containing 100 mM sodium citrate (pH 5.4), 200 mM ammonium acetate, and 23% PEG 3350. A detailed description of the experimental procedures is available in the Supplemental Experimental Procedures.

Binding Experiments

ITC and FP measurements were carried out as described in Grimm et al. (2007) and Jacobs et al. (2004), respectively. A detailed description of the experimental procedures is available in the Supplemental Experimental Procedures.

Nucleosome Assembly and Histone Methyltransferase Assay

Recombinant histones were expressed and purified (Luger et al., 1999) and site-specific methylation reactions were carried out as described in Simon (2010) and Simon et al. (2007). Histone octamers were reconstituted and purified as described in Luger et al. (1999) and assembled into nucleosomes by using the 601 sequence (Thåström et al., 1999) during stepwise salt dilution by dialysis. For the HMTase reaction mononucleosomes or 4-mer oligonucleosomes were incubated with the indicated amounts of PRC2 for 2 hr at 25°C in the presence of S-adenosyl methionine and reaction products were analyzed by western blotting. A detailed description of the experimental procedures is available in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The coordinates for the Nurf55-H3 and the Nurf55-Su(z)12 complex have been deposited in the PDB under ID codes 2YBA and 2YB8, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.molcel.2011.03.025.

ACKNOWLEDGMENTS

We thank Dirk Schübeler, Antoine Peters, and Susan Gasser for helpful discussions. We thank Norman Kairies and Andrea Scrima for crystallographic support, Maja Koehn and Victoria McParland for advice on peptide synthesis and purification, and Vladimir Rybin for help with ITC measurements. We are also grateful to Ernest Laue for providing us with the plasmid for the GST-H4 construct. Funding in the laboratory of N.H.T. is provided by Association for International Cancer Research grant 10-0292 and the Novartis Research Foundation. F.W.S. gratefully acknowledges funding from the Schering Foundation and the Studienstiftung des deutschen Volkes. A.B.P., N.L.-H., and J.M. were supported by European Molecular Biology Laboratory.

Received: June 14, 2010
Revised: December 8, 2010
Accepted: March 18, 2011
Published: May 5, 2011

REFERENCES

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. (2002). Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature *419*, 857–862.

Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schübeler, D. (2008). Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. Mol. Cell. Biol. *28*, 3401–3409.

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

Birve, A., Sengupta, A.K., Beuchle, D., Larsson, J., Kennison, J.A., Rasmuson-Lestander, A., and Müller, J. (2001). Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. Development *128*, 3371–3379.

Cao, R., Wang, H., He, J., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2008). Role of hPHF1 in H3K27 methylation and Hox gene silencing. Mol. Cell. Biol. *28*, 1862–1872.

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

Chang, P.Y., Hom, R.A., Musselman, C.A., Zhu, L., Kuo, A., Gozani, O., Kutateladze, T.G., and Cleary, M.L. (2010). Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. J. Mol. Biol. *400*, 137–144.

Collins, R.E., Northrop, J.P., Horton, J.R., Lee, D.Y., Zhang, X., Stallcup, M.R., and Cheng, X. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nat. Struct. Mol. Biol. *15*, 245–250.

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

Dong, A., Xu, X., Edwards, A.M., Chang, C., Chruszcz, M., Cuff, M., Cymborowski, M., Di Leo, R., Egorova, O., Evdokimova, E., et al; Midwest Center for Structural Genomics; Structural Genomics Consortium. (2007). In situ proteolysis for protein crystallization and structure determination. Nat. Methods *4*, 1019–1021.

Gaudet, R., Bohm, A., and Sigler, P.B. (1996). Crystal structure at 2.4 angstroms resolution of the complex of transducin betagamma and its regulator, phosducin. Cell *87*, 577–588.

Grimm, C., de Ayala Alonso, A.G., Rybin, V., Steuerwald, U., Ly-Hartig, N., Fischle, W., Müller, J., and Müller, C.W. (2007). Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein Sex comb on midleg. EMBO Rep. *8*, 1031–1037.

Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. Nat. Cell Biol. *10*, 1291–1300.

He, Y. (2009). Control of the transition to flowering by chromatin modifications. Mol Plant *2*, 554–564.

Henderson, I.R., and Dean, C. (2004). Control of Arabidopsis flowering: the chill before the bloom. Development *131*, 3829–3838.

Jacobs, S.A., Fischle, W., and Khorasanizadeh, S. (2004). Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. Methods Enzymol. *376*, 131–148.

Jiang, D., Wang, Y., Wang, Y., and He, Y. (2008). Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components. PLoS ONE *3*, e3404.

Ketel, C.S., Andersen, E.F., Vargas, M.L., Suh, J., Strome, S., and Simon, J.A. (2005). Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Mol. Cell. Biol. *25*, 6857–6868.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol. Cell *14*, 183–193.

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

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol. Biol. *119*, 1–16.

Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, W.J., 3rd, Voigt, P., Martin, S.R., Taylor, W.R., De Marco, V., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762–767.

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

Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell *111*, 197–208.

Murzina, N.V., Pei, X.Y., Zhang, W., Sparkes, M., Vicente-Garcia, J., Pratap, J.V., McLaughlin, S.H., Ben-Shahar, T.R., Verreault, A., Luisi, B.F., and Laue, E.D. (2008). Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. Structure *16*, 1077–1085.

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

Nekrasov, M., Klymenko, T., Fraterman, S., Papp, B., Oktaba, K., Köcher, T., Cohen, A., Stunnenberg, H.G., Wilm, M., and Müller, J. (2007). Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. EMBO J. *26*, 4078–4088.

Nekrasov, M., Wild, B., and Müller, J. (2005). Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep. *6*, 348–353.

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

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

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. *23*, 4061–4071.

Pietersen, A.M., and van Lohuizen, M. (2008). Stem cell regulation by polycomb repressors: postponing commitment. Curr. Opin. Cell Biol. *20*, 201–207.

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

Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., and Reinberg, D. (2008). Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Mol. Cell. Biol. *28*, 2718–2731.

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

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

Schwartz, Y.B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. Nat. Rev. Genet. *8*, 9–22.

Scrima, A., Konícková, R., Czyzewski, B.K., Kawasaki, Y., Jeffrey, P.D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N.P., and Thomä, N.H. (2008). Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. Cell *135*, 1213–1223.

Shogren-Knaak, M.A., Fry, C.J., and Peterson, C.L. (2003). A native peptide ligation strategy for deciphering nucleosomal histone modifications. J. Biol. Chem. *278*, 15744–15748.

Simon, M.D. (2010). Installation of site-specific methylation into histones using methyl lysine analogs. In Current Protocols in Molecular Biology, Ausubel, F.M., et al., eds. (Hoboken, NJ: John Wiley & Sons), Unit 21.18.1–21.18.10.

Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. Cell *128*, 1003–1012.

Smith, S.T., Petruk, S., Sedkov, Y., Cho, E., Tillib, S., Canaani, E., and Mazo, A. (2004). Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. Nat. Cell Biol. *6*, 162–167.

Song, J.J., Garlick, J.D., and Kingston, R.E. (2008). Structural basis of histone H4 recognition by p55. Genes Dev. *22*, 1313–1318.

Srinivasan, S., Dorighi, K.M., and Tamkun, J.W. (2008). Drosophila Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA polymerase II. PLoS Genet. *4*, e1000217.

Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D., and Kitajima, S. (2007). Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene *397*, 161–168.

Thåström, A., Lowary, P.T., Widlund, H.R., Cao, H., Kubista, M., and Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. J. Mol. Biol. *288*, 213–229.

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

Wu, J.I., Lessard, J., and Crabtree, G.R. (2009). Understanding the words of chromatin regulation. Cell *136*, 200–206.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86–90.

Xu, C., Bian, C., Yang, W., Galka, M., Ouyang, H., Chen, C., Qiu, W., Liu, H., Jones, A.E., Mackenzie, F., et al. (2010). Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc. Natl. Acad. Sci. U. S. A. *107*, 19266–19271.

Yokoyama, A., Okuno, Y., Chikanishi, T., Hashiba, W., Sekine, H., Fujiki, R., and Kato, S. (2010). KIAA1718 is a histone demethylase that erases repressive histone methyl marks. Genes Cells *15*, 867–873.

Young, N.L., DiMaggio, P.A., Plazas-Mayorca, M.D., Baliban, R.C., Floudas, C.A., and Garcia, B.A. (2009). High throughput characterization of combinatorial histone codes. Mol. Cell. Proteomics *8*, 2266–2284.

Yu, X., and Michaels, S.D. (2010). The Arabidopsis Paf1c complex component CDC73 participates in the modification of FLOWERING LOCUS C chromatin. Plant Physiol. *153*, 1074–1084.

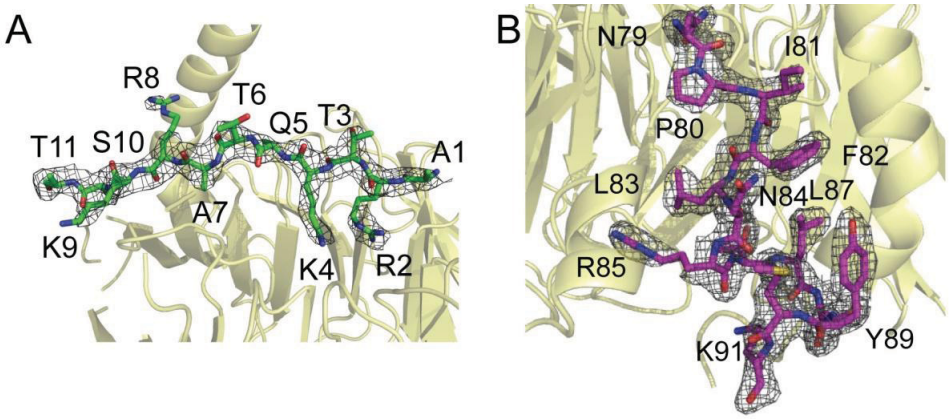
Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat. Struct. Mol. Biol. *15*, 381–388.

Molecular Cell, Volume *42*

Supplemental Information

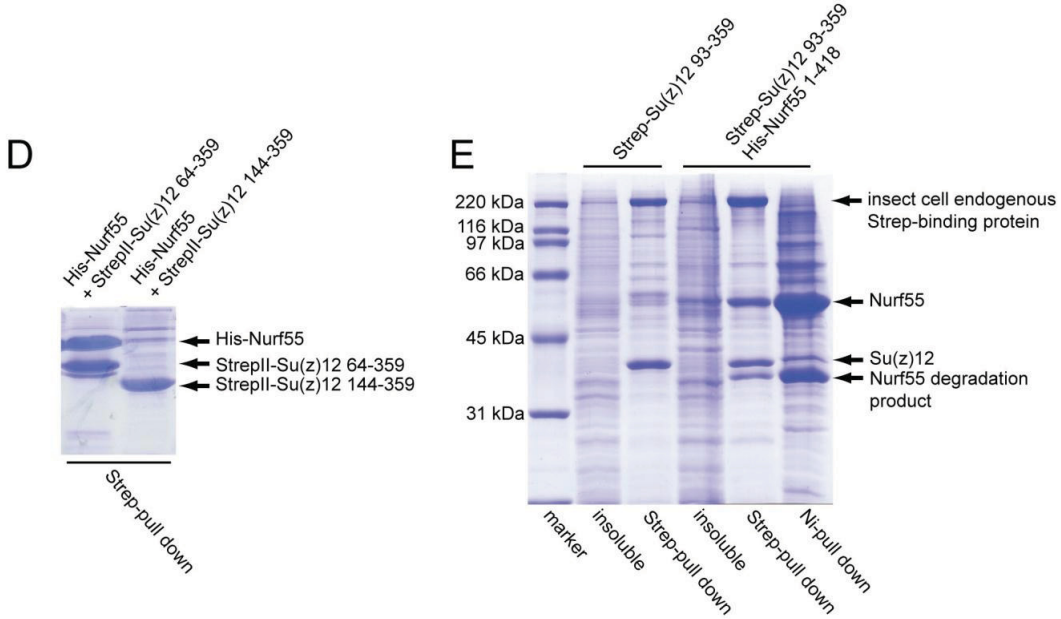
Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks

Frank W. Schmitges, Archana B. Prusty, Mahamadou Faty, Alexandra Stützer, Gondichatnahalli M. Lingaraju, Jonathan Aiwazian, Ragna Sack, Daniel Hess, Ling Li, Shaolian Zhou, Richard D. Bunker, Urs Wirth, Tewis Bouwmeester, Andreas Bauer, Nga Ly-Hartig, Kehao Zhao, Homan Chan, Justin Gu, Heinz Gut, Wolfgang Fischle,, Jürg Müller, and Nicolas H. Thomä



C

PEPTIDE SEQUENCE for Su(z)12	PEPTIDE START POSITION (aa)	LOG(2) FOLD CHANGE (GF FRACTION/CONTROL)
LNGHQQEQELFLQAFEKPTQIYR	48-70	no iTRAQ quantification
ISFQVNSMLESITQK	102-116	-3,4
SEAVSQNYLHVIYDSLHEK	117-135	no iTRAQ quantification
KDSTLDFQELLSK	176-188	-9,4
DSTLDFQELLSK	177-188	-6,5
MFGSELILYEK	256-266	-10,0
SSGFITEGEYEAMLQPLNSTSIK	267-289	-7,6
FHLTLSNEQLPEMISAPELQR	323-343	-7,4
FNFTYQPAGSGAR	437-449	-9,9
ELDIDSEGESDPLWLR	542-557	-7,2
TIQMIDEFSDVNEGEKELMK	560-579	-2,2
TIQMIDEFSDVNEGEK	560-575	-1,6
YAAGQELMQR	652-661	-2,3
YWLDVGMHK	669-677	-0,6
GSAASSPGVQSK	728-739	-5,0

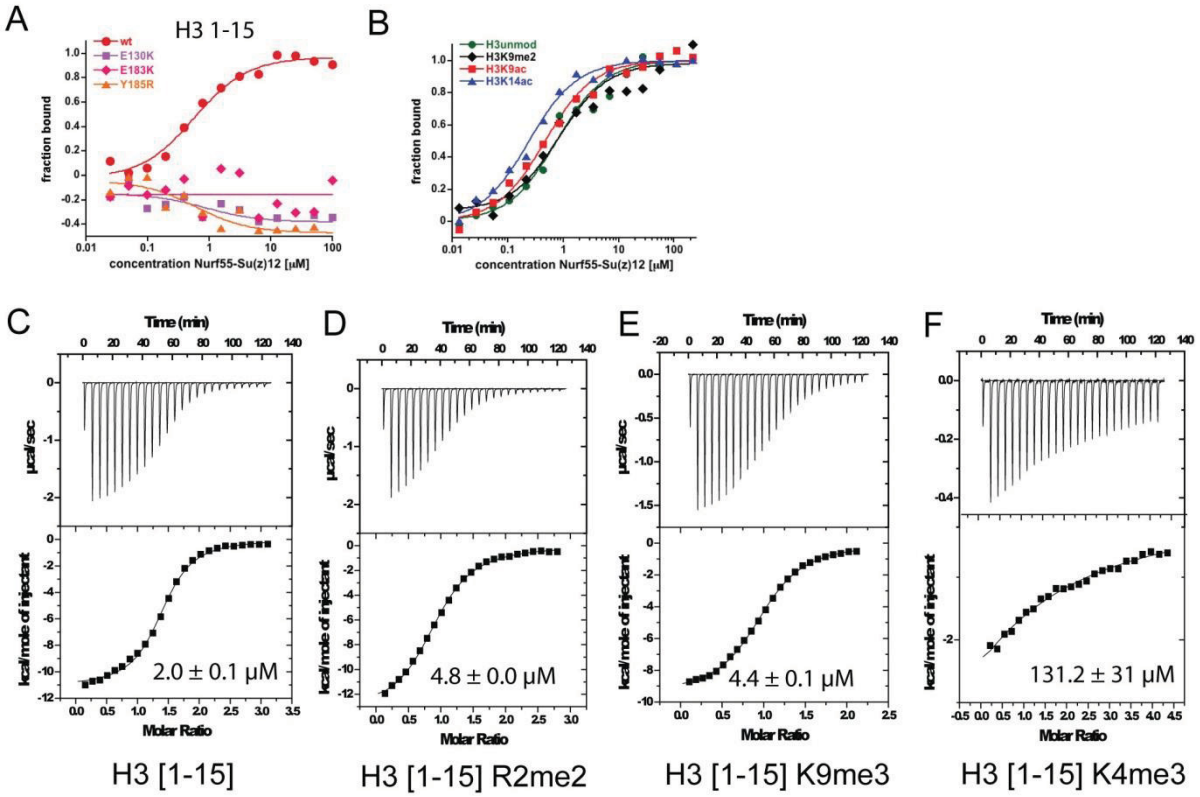


Supplemental Figure 1

Figure S1. Crystal Structures of Nurf55-H3 and Nurf55-Su(z)12 Complexes
(A) Simulated annealing composite omit map (2mFo-DFc) of Nurf55-H3. $H3_{1-19}$ peptide (green) in the *c*-site of Nurf55, contoured at *1* σ.
(B) Simulated annealing composite omit map (2mFo-DFc) of Nurf55-Su(z)12- $Su(z)12_{79-91}$ (magenta) in the Nurf55 S/H-site (yellow), contoured at *1* σ.
(C) Unbiased mapping of Su(z)12 epitopes involved in Nurf55 binding.
The tetrameric PRC2 core complex was subjected to limited proteolysis using 1% subtilisin for 1h at room temperature. The digestion was stopped by addition of 0.1 mM PMSF and the digest immediately injected on a 25 ml Superdex200 column equilibrated with 50 mM HEPES pH 7.4, 200 mM NaCl, 0.1 mM DTT. Upon injection of the digest, 500 µl fractions were collected and rapidly frozen in liquid nitrogen. We then sampled fractions on the basis of spectral counting, and in-gel tryptic fingerprinting using a largely non-quantitative mass spectrometry approach (data not shown). This allowed us to identify those fractions enriched in Nurf55, Su(z)12 and E(z). In order to determine which Su(z)12 fragments remained stably bound to Nurf55 following subtilisin digestion, we focused on the fraction that retained the highest sum of spectral counts for Nurf55. As this fraction eluted at an approximate molecular weight of ~50-60 kDa, we assumed that the WD40 fold of Nurf55 is largely intact (as expected for a folded domain) and associated with smaller peptide fragments derived from Su(z)12. We note that ~35-40 kDa fragments of E(z) were present in this fraction as well.

We then performed a two-channel iTRAQ quantitative mass spectrometric anlysis in order to quantify which Su(z)12 fragments are left following digest and gelfiltration. Protein fragments from the gelfiltration fractions were fully digested with trypsin and extracted peptides labeled with iTRAQ reagent 114 (Invitrogen). PRC2 core complex not subjected to limited proteolysis, denoted as 'control', was processed in the identical manner and labeled with iTRAQ reagent 117. Both tryptic fractions were combined and analyzed by liquid chromatography–tandem mass spectrometry essentially as described in Huang et al., 2009. Relative fold changes were calculated for all peptides identified with >1 spectra and a probability score of >0.98 (PeptideProphet). Low iTRAQ values are indicative for peptides that are under-represented in the gelfiltration fraction respective to the control sample and we assume that those are not bound to Nurf55. Peptides that were not accessible to iTRAQ quantification are labeled as "no iTRAQ quantification". Similarly peptides digested by subtilisin on one end of the peptide, while being trypsin digested on the other were ignored in this analysis due to the absence of reference spectra.

When examining iTRAQ fold-changes we find two portions of Su(z)12 above a threshold of log(2)= -4 (highlighted in red). These comprise the Su(z)12 C-terminus including the VEFS domain, and the extended NBE. We attribute the presence of the Su(z)12 C-terminus to binding to E(z) fragments that co-migrated in this fraction. We independently confirmed binding of these Su(z)12 C-terminal fractions to E(z) through binding experiments in the absence of Nurf55 (data not shown; see *also* Ketel et al., 2005. The enrichment of the Su(z)12 N-terminus comprising the NBE (peptides found cover Su(z)12 residues 48 to 135), is in line with our mapping data (Figure S1D) and is in accordance with the Su(z)12-Nurf55 structure (Figure 2). The most significant fold-enrichment was observed at Su(z)12 residues 102-116, immediately C-terminal of the NBE. We note that the NBE is directly sampled through Su(z)12 peptide 40-70, which was found only in the digested sample. Based on the nature of the mass-spectrometry based approach we cannot make statements about regions in Su(z)12 not covered in peptide space. The enrichment of the Su(z)12 NBE was also observed when full analytical Asp-N digest of the subtilisin digest gel-filtration fraction was used (data not shown).
(D) His-tagged Nurf55 and StrepII-tagged $Su(z)12_{64-359}$ or $Su(z)12_{144-359}$ were co-expressed in insect cells. While $Su(z)12_{64-359}$ interacts with Nurf55, the $Su(z)12_{144-359}$ construct is not able to bind anymore.
(E) His-tagged Nurf55 and StrepII-tagged $Su(z)12_{93-359}$ were co-expressed in insect cells. Although this Su(z)12 construct lacks the NBE, the two proteins co-purify in a pull-down assay using Strep-Tactin resin, suggesting the presence of additional binding sites between Nurf55 and Su(z)12.



Supplemental Figure 2

Figure S2. Measurement of Binding of H3 Peptides to Nurf55-Su(z)12 by Fluorescence Polarization and Isothermal Titration Calorimetry

(A) Fluorescence polarization measurements of Nurf55-Su(z)12$_{3\text{-}143}$ complexes carrying different mutations in the Nurf55 binding site for H3. The Nurf55 mutations E130K, E183K, Y185R (squares) result in loss of binding to a H3$_{1\text{-}15}$ peptide, compared to the wild-type Nurf55-Su(z)12$_{3\text{-}143}$ control (circles).

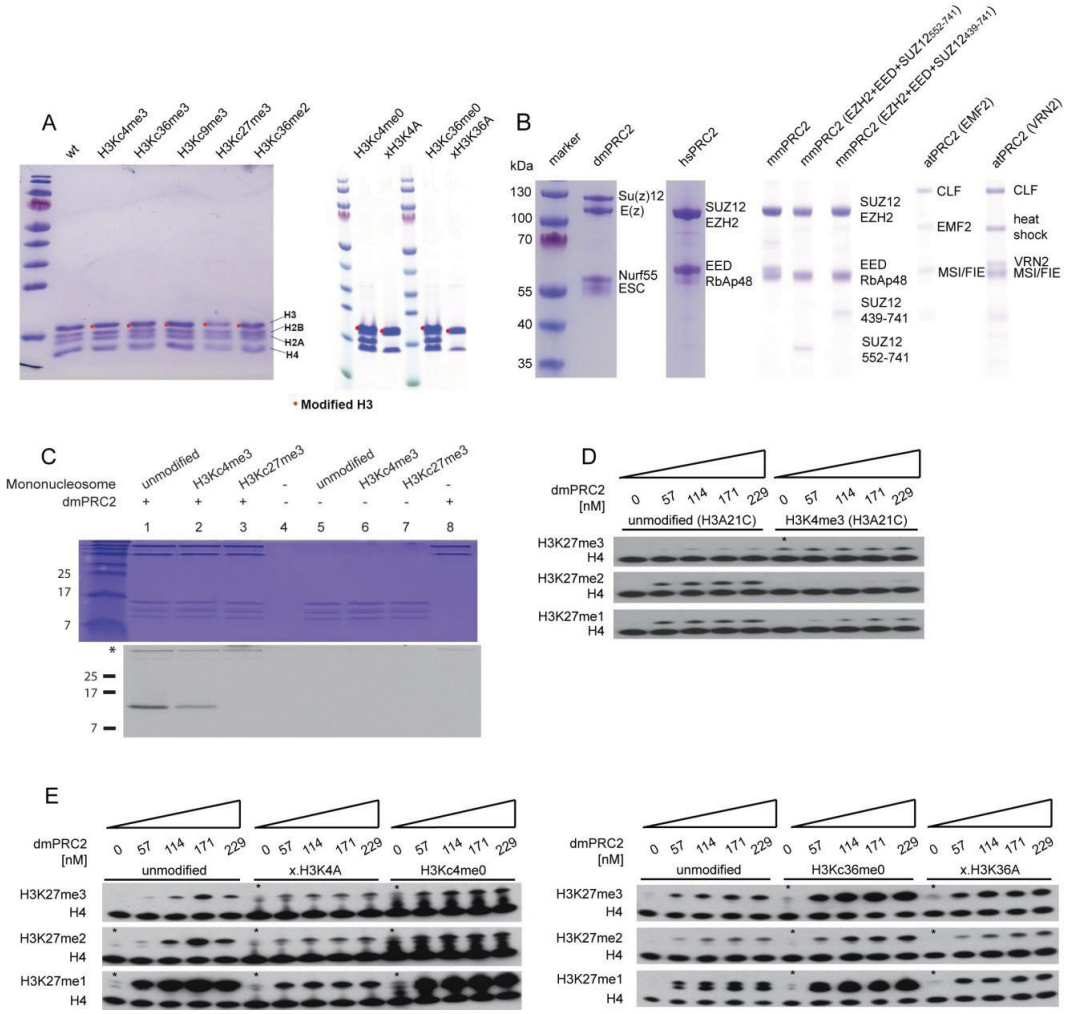
(B) Fluorescence polarization assay showing that H3$_{1\text{-}20}$ peptides carrying either Lys9me2, Lys9ac or Lys14ac modification bind with comparable affinities as seen for the unmodified peptide.

(C) ITC profile for the binding of an unmodified H3$_{1\text{-}15}$ peptide to Nurf55-Su(z)12$_{73\text{-}143}$. Data were fitted to a one-site model with stoichiometry of 1:1. The derived K_D value is 2.0 ± 0.1 µM.

(D) ITC profile for the binding of a H3$_{1\text{-}15}$R2me2 peptide (ω-NG,NG-asymmetric dimethylarginine (aDMA)) to Nurf55-Su(z)12$_{73\text{-}143}$. Data were fitted to a one-site model with stoichiometry of 1:1. The derived K_D value is 4.8 ± 0.0 µM.

(E) ITC profile for the binding of a H3$_{1\text{-}15}$-K9me3 peptide to Nurf55-Su(z)12$_{73\text{-}143}$. Data were fitted to a one-site model with stoichiometry of 1:1. The derived K_D value is 4.5 ± 0.1 µM.

(F) ITC profile for the binding of a H3$_{1\text{-}15}$-K4me3 peptide to Nurf55-Su(z)12$_{73\text{-}143}$. Data were fitted to a one-site model with stoichiometry of 1:1. The derived K_D value is 131.2 ± 31 µM.



Supplemental Figure 3

Figure S3. Analysis of MLA-Containing Histone H3

(A) Coomassie staining of *Drosophila* (wt, H3Kc4me3, H3Kc36me3, H3Kc9me3, H3Kc27me3, H3Kc36me2, H3Kc4me0, H3Kc36me0) and *Xenopus* (H3K4A, H3K36A) containing octamers run on an SDS-polyacrylamide gel after purification on a Superdex200 size exclusion column.

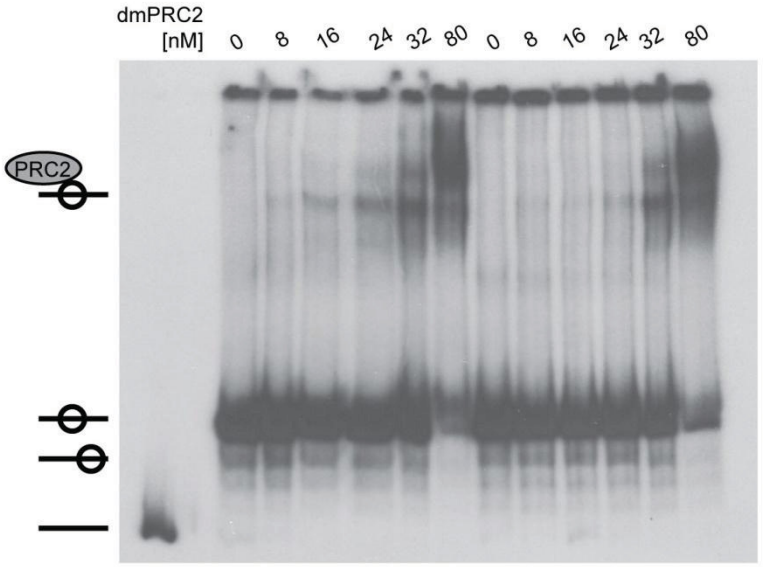
(B) Coomassie staining of recombinant PRC2 complexes (from left to right): *Drosophila* (dm) PRC2, human (hs) PRC2, mouse (mm) wild-type and SUZ12 deletion complexes (552-741 and 439-741), *Arabidopsis thaliana* (at) EMF2 and VRN2 containing PRC2.

(C) *Drosophila* PRC2 HMTase assay using ^{14}C-labled SAM and recombinant mononucleosomes as substrate. On unmodified nucleosomes (lane 1) incorporation of ^{14}C-SAM can be detected. HMTase activity is significantly reduced when H3Kc4me3 modified nucleosomes are used (lane 2). No incorporation is observed with H3Kc27me3 nucleosomes (lane 3) or in reactions without PRC2 (lanes 4-7). The high molecular weight signal results from automethylation of the PRC2 complex.

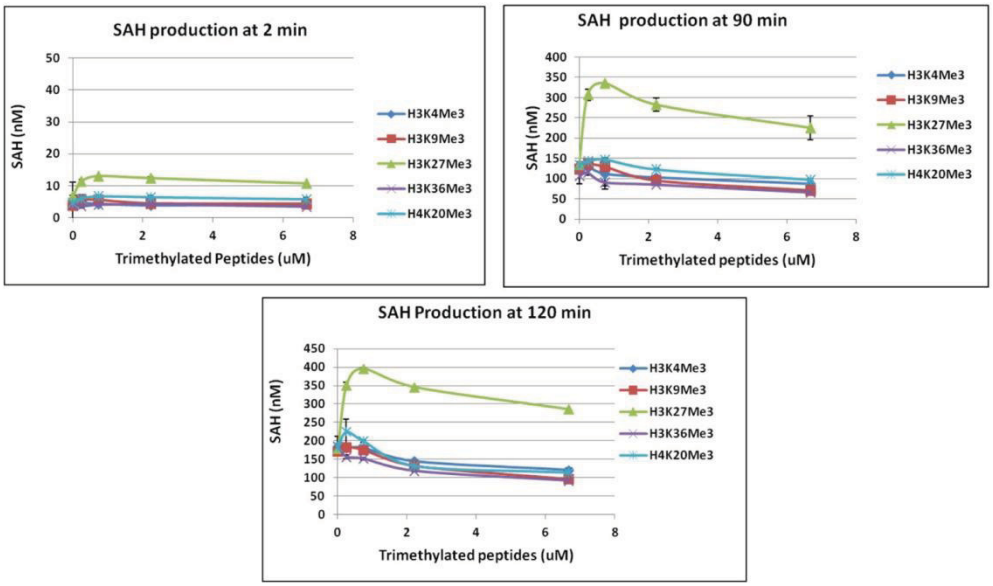
(D) HMTase assay with *Drosophila* PRC2 complex on peptide ligated mononucleosomes. Recombinant *Xenopus* mononucleosomes (571 nM) containing unmodified histone H3 (H3A21C) or H3K4me3 (A21C) generated by native peptide ligation were used as substrates in HMTase assays with the indicated amounts of *Drosophila* PRC2. Reactions were analyzed

as described in Figure 3B. The asterisk indicates cross-reactivity of the H3K27me3 antibody with the H3K4me3 (A21C) histone; note that the same signal is present in lanes 6-10.
(E) HMTase assay with *Drosophila* mononucleosomes (571 nM) containing H3K4A, H3Kc4me0, H3K36A, or H3Kc36me0 histones. Western blots were processed as described in Figure 3B. The asterisk indicates the cross-reactivity of the H3K27me1/2/3 antibody to the histone H3. The modifications show no significant inhibition of HMTase activity, indicating that the effects seen with H3Kc4me3 and H3Kc36me3 nucleosomes are specific and not due to the MLA modification.

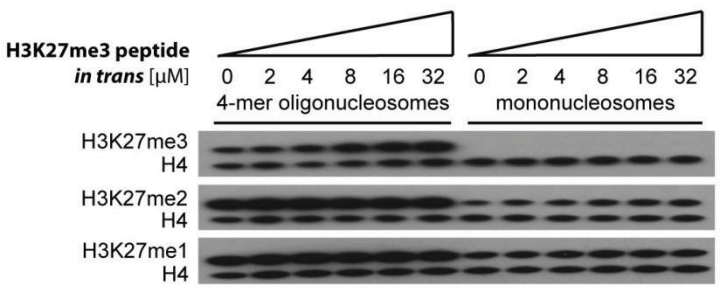
A



B



C



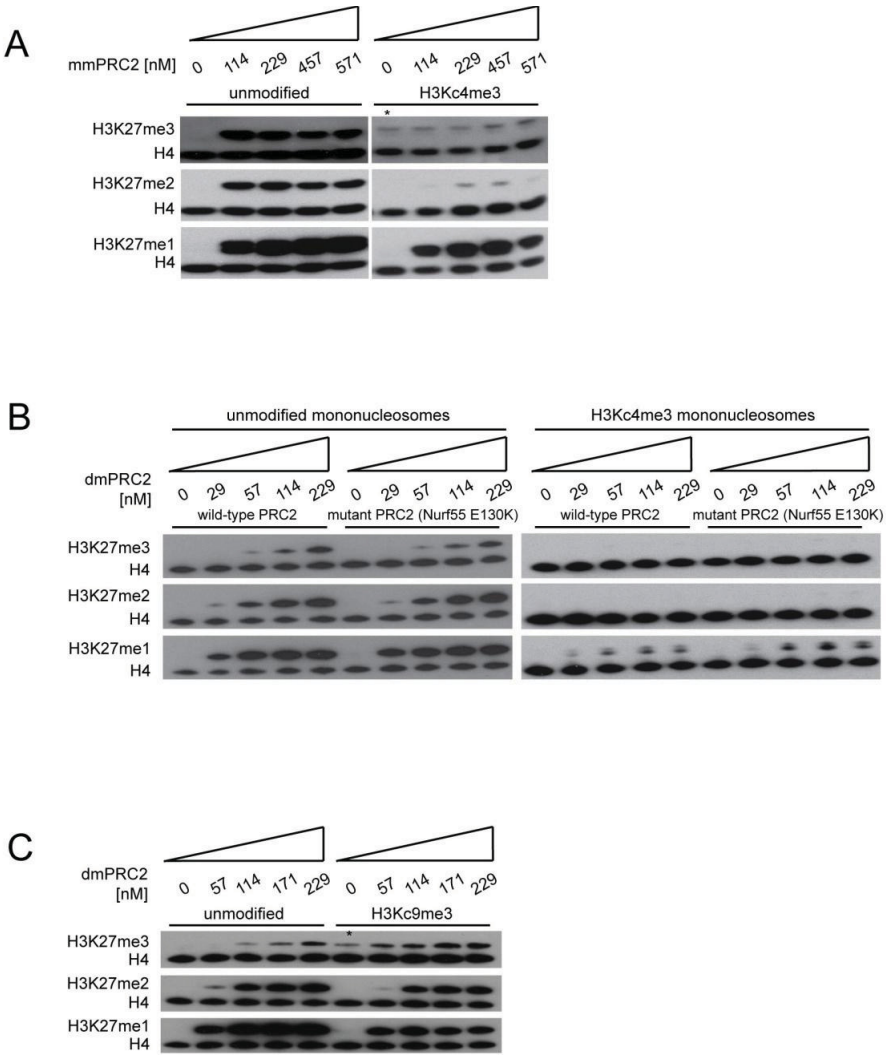
Supplemental Figure 4

Figure S4. Electrophoretic Mobility Shift Assay (EMSA) and HMTase Assay with PRC2 and Recombinant Mononucleosomes

(A) Unmodified and H3Kc4me3 modified mononucleosomes were incubated with increasing amounts of PRC2 and analyzed on a 4% native polyacrylamide gel. A specific PRC2 nucleosome gel-shift was observed in a concentration dependent manner. Unmodified and H3Kc4me3 modified nucleosomes exhibited comparable affinities, indicating that the H3Kc4me3 modification does not affect the overall affinity of PRC2 for the nucleosome.

(B) Peptide based HMTase assay with modified peptides added *in trans*. HMTase activity of human PRC2 complex on an unmodified $H3_{21-44}$ peptide was measured by determination of SAH concentration produced by the enzymatic reaction. Differently modified peptides ($H3_{1-21}$-K4me3-biotin, $H3_{1-21}$-K9me3-biotin, $H3_{21-44}$-K27me3-biotin, $H3_{21-44}$-K36me3-biotin, $H4_{10-30}$-K20me3) were added *in trans* to the reaction mix. While H3K27me3 resulted in stimulation of HMTase activity, H3K4me3 and H3K36me3 modified peptides had little effect when provided *in trans*. All modified peptides used in this study are subject to a low level of inhibition with the unmodified H3 substrate peptide, when used at high concentration. This also occurs with an H4 peptide used as control and we conclude that this inhibition is unspecific. This effect became more pronounced as the reaction progressed, and hence the H3K27 substrate concentration decreased. The observed curve shapes are identical for all peptides, including H3K27me3. In the case of H3K27me3, however, inhibition happens on top of the previously observed strong activation.

(C) *Drosophila* PRC2 HMTase assay with H3K27me3 peptide addition *in trans*. Unmodified recombinant *Drosophila* 4-mer oligonucleosome array (left) or mononucleosomes (right) (571 nM nucleosomes in either case) were used as substrates in HMTase assays with 86 nM *Drosophila* PRC2 in the presence of the indicated amounts of H3K27me3 peptide. The reactions were analyzed as described in Figure 3B. The H3K27me3 peptide stimulates H3K27 trimethylation on oligonucleosomes, and some minor stimulation of H3K27 dimethylation is also observed on mononucleosomes. PRC2 activity on oligonucleosomes is generally higher than on mononucleosomes. Therefore, olignucleosomes but not mononucleosomes become trimethylated in the absence of peptide addition even at low amounts of PRC2 (86 nM).



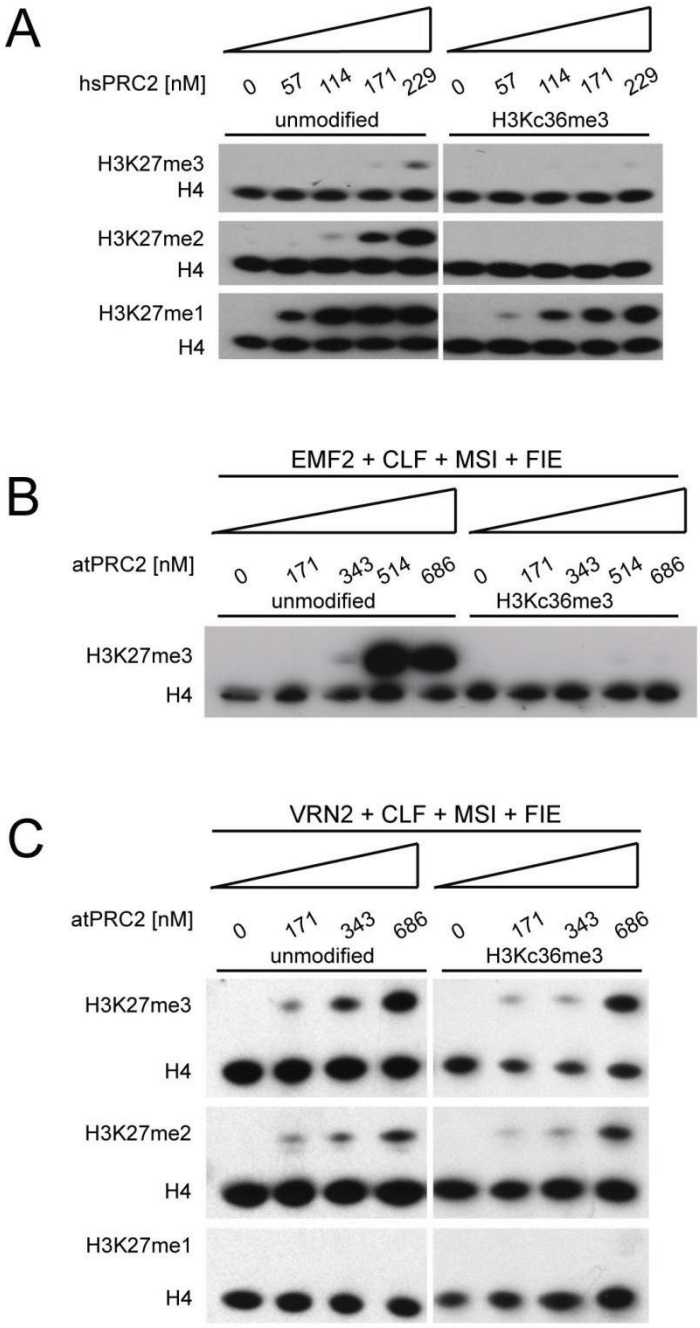
Supplemental Figure 5

Figure S5. HMTase Assays with Mouse PRC2, Nurf55 Mutants, and H3Kc9-Modified Histones

(A) HMTase assay with mouse PRC2 complex and recombinant *Drosophila* 4mer oligonucleosomes (36 nM). The reactions were analyzed as described in Figure 3B. While mono-, di- and trimethylation are observed on unmodified nucleosomes, monomethylation is reduced on H3Kc4me3 modified nucleosomes and no di- or trimethylation can be detected.

(B) PRC2 containing mutant Nurf55 disabled in H3-binding shows wild-type HMTase activity. Recombinant *Drosophila* mono-nucleosomes (571 nM) containing unmodified H3 were used as substrates in HMTase assays with the indicated amounts of *Drosophila* PRC2 containing either wild-type Nurf55 or Nurf55^{E130K}. Reactions were analyzed as described in Figure 3B. The mutant and wild-type complexes mono-, di- and trimethylate H3K27 with comparable efficiency on unmodified nucleosomes and are inhibited to an equal extent on H3K4me3 modified nucleosomes.

(C) HMTase assay with *Drosophila* PRC2 on unmodified and H3Kc9me3 modified mononucleosomes (571 nM). Western blots were processed as described in Figure 3B. The complex shows comparable HMTase activity on both substrates. The asterisk indicates cross reactivity of the H3K27me3 antibody with the H3Kc9me3 histone.



D

Substrate	Apparent Km of SAM (μM)	Apparent Km of peptide (nM)	k_{cat} (min^{-1})	k_{cat}/Km ($M^{-1}S^{-1}$)
$H3_{1-45}$ -biotin	5.42 ± 0.65	836 ± 207	2.53 ± 0.21	7.8×10^3
$H3K36me3_{1-45}$ -biotin	4.84 ± 0.81	392 ± 69.8	2.04 ± 0.15	7.02×10^3

Supplemental Figure 6

Figure S6. Sensitivity towards H3K36me3 Is Conserved in Human and Plant Complexes
HMTase assay with unmodified and H3Kc36me3 modified nucleosomes. Human PRC2 (A) and a plant EMF2-containing complex (B) are inhibited by Kc36me3 comparable to the *Drosophila* complex. The plant VRN2-containing complex (C), in contrast, is insensitive and shows comparable activity on unmodified and H3Kc36me3 modified nucleosomes. In the case of human PRC2 571 nM mononucleosomes and for plant complexes 36 nM 4-mer oligonucleosomes were used as substrate.
(D) HMTase assay with *Drosophila* PRC2 using an $H3_{1-45}$-biotin peptide as substrate. The reaction was quantified by measurement of SAH produced. When a $H3_{1-45}$K36me3 modified peptide is used, no decrease in enzymatic turnover is observed, unlike with what was observed with a $H3_{1-45}$K4me3 peptide (Figure 3A). It is not uncommon for enzymes working on chromatin substrates to show specificity only on chromatin, whilst not being able to function on isolated histone tail peptides. In those cases it was proposed that epitopes on the nucleosome are required to guarantee proper positioning of the histone tail respective to the active site.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Protein Expression and Purification

All full-length PRC2 constructs were expressed in High Five insect cells utilizing the Bac-to-Bac system (Invitrogen) according to the manufacturer's instructions and purified by sequential column based chromatographic methods. Full-length E(z) and Su(z)12 were cloned into vectors with an N-terminal StrepII-tag, Nurf55, ESC and Su(z)12_{73-143} into vectors with an N-terminal His_6-tag. Histone $H4_{1-48}$ and mouse SUZ12 deletion constructs 439-741 and 552-741 were expressed as GST-tagged proteins in BL21(DE3) cells. Mutations in Nurf55 were introduced using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene).

For infection 2 x 1 L of High Five insect cells (4 x 10^6 cells/ml) were harvested by centrifugation (1,500 rpm, 4°C,15 min) and resuspended in 15 ml of Baculovirus (P2 or P3 amplification state) per liter. The cells were incubated for one hour at room temperature prior to resuspension in 4 x 1 L Sf-900 II medium that was supplemented with Penicilin and Streptomycin (100 µg/ml). The cultures were grown in 2.5 L Fernbach flasks for 48 hrs at 27°C while shaking (150 rpm). BL21(DE3) cells were induced with 0.5 mM IPTG at OD 0.6 and grown over night at 20°C.

The cells were harvested by centrifugation (1,500 rpm, 4°C, 15 min) and the pellets resuspended in lysis buffer (50 mM Tris, pH 8.0, 200 mM NaCl, 1 mM PMSF, 0.1% TritonX 100, 0.25 mM TCEP; approximately 200 ml total volume). Cells were disrupted by sonication (8-10 times 20 sec, 60%, on ice) followed by ultracentrifugation (40,000 rpm, 4°C, 1.5 hrs, Rotor: 45Ti). The supernatant contains the soluble protein and was used for subsequent purification by chromatographic methods.

For affinity purification 25 ml NiNTA (Sigma), 25ml Glutathione Sepharose (GE Healthcare) or 8 ml Strep-Tactin MacroPrep (IBA) were equilibrated first with lysis buffer. The supernatant from the ultracentrifugation step was subsequently applied to the column. A washing step was carried out with around 10 column volumes (CV) washing buffer (50 mM Tris, pH 8.0, 200 mM NaCl, 0.25 mM TCEP) to remove bound proteins. His_6-tagged proteins were eluted with washing buffer containing 100 mM imidazole, GST-tagged proteins with 10 mM Glutathione. For StrepII-tagged proteins elution buffer contained 2.5 mM D-Desthiobiotin (IBA).

For ion-exchange chromatography Poros 50HQ resin (GE healthcare) was packed into a HR16/5 (10 ml CV) column and chromatography was carried out with the ÄKTA FPLC. The column was equilibrated with 4 CV of low salt buffer, the sample diluted to around 80 mM NaCl with dilution buffer and loaded on the column with a flow-rate of 4 ml/min. A washing step was performed with 5 CV of low salt buffer, followed by elution with a linear gradient from 0mM to 700 mM NaCl over 15 CV with a fraction size of 5 ml. For crystallization proteins were further purified by size exclusion chromatography on a Superdex200 column. Relative UV absorption was recorded during the procedure and subsequent analysis of peak fractions by SDS-PAGE verified the fractions containing the proteins of interest. Prior to the following gel filtration step, the fractions containing the proteins were pooled and concentrated, followed by flash-freezing in liquid N_2 and storage at -80°C.

Crystallization and Structure Determination

Crystals were grown at 20-25°C by the hanging drop vapor diffusion method. For Nurf55-Su(z)12 crystallization, a co-expressed $Nurf55_{1-418}$ and Su(z)12_{73-143} complex at 16 mg/ml was incubated for 10 min with 0.01% subtilisin prior to set-up. Subsequently 1 µl drops of protein solution were mixed with 1 µl of reservoir solution containing 100 mM potassium

acetate and 2.1 M ammonium sulfate. Crystals were harvested after 2 days and flash frozen in crystallization buffer supplemented with 25% glycerol. For Nurf55-$H3_{1-19}$ crystallization, Nurf55 protein was incubated with a ~5 molar excess of the $H3_{1-19}$ peptide for 30 min prior to crystallization. 1 µl drops of a 22 mg/ml protein solution were mixed with 1 µl of reservoir solution containing 100 mM sodium citrate, pH 5.4; 200 mM ammonium acetate; 23% PEG 3350. For data collection crystals were cryoprotected in a solution containing 100 mM sodium citrate, pH 5.4; 200 mM ammoniumacetate; 40% PEG 3350 and subsequently flash frozen in liquid N_2. Native data sets were collected at the Swiss Light Source, beamline X10SA, Paul Scherrer Institut, Villigen, Switzerland. Reflection data were indexed, integrated, and scaled using the HKL2000 package (Otwinowski and Minor, 1997). Nurf55-Su(z)12 and Nurf55-$H3_{1-19}$ crystallized in space group $P2_12_12_1$, with a single and two complexes located in the asymmetric unit, respectively. The structures were solved by molecular replacement using Phaser (McCoy et al., 2007) with our previously determined 1.8 Å structure of Nurf55 as a search model (data not shown). Iterative cycles of model building and refinement were performed using Coot (Emsley and Cowtan, 2004) and Phenix (Adams et al., 2010) or BUSTER (Bricogne et al., 2009). The refined structures were verified using simulated annealing composite-omit maps as implemented in CNS 1.2 with an omit ratio of 5%. Structural Figures in this manuscript were generated with PyMol (DeLano Scientific, http://www.pymol.org).

Fluorescence Polarization Measurements

The following peptides were used for fluorescence polarization measurements: $H3_{1-15}$, ARTKQTARKSTGGKA-fluorescein (the fluorescein moiety was introduced site-specifically at the C-terminus during peptide synthesis) $H3_{1-31}$ (fluorescein)ARTKQTARKSTGGKAPRKQLATKAARKSAPA. The $H3_{1-31}$ peptide was chemically labeled post peptide synthesis using NHS-fluorescein; single labeled species were purified by RP-HPLC. FP assays were essentially carried out and analyzed as described before (Jacobs et al., 2004). Titration series of protein in 10 µl 50mM HEPES pH 7.4, 150 mM NaCl, 0.25 mM TCEP with a final concentration of fluorescinated peptide of 1 nM were read multiple times in 384 well plates on a Plate Chameleon II plate reader (HIDEX Oy). Multiple readings and independent titration series were averaged after data normalization.

Isothermal Titration Calorimetry (ITC)

Isothermal Titration Calorimetry was carried out using a VP-ITC Microcal calorimeter (Microcal, Northhampton, MA, USA). The Nurf55 protein was dialyzed overnight against ITC buffer (20 mM Tris-Cl pH 8.0, 150 mM NaCl and 2 mM β-mercaptoethanol) before the titration. The synthesized $Su(z)12_{75-93}$ (RHETNPIFLNRTLSYMKER) peptide was purified by reverse-phase HPLC in the presence of trifluoroacetic acid, the dried peptide was treated with freshly prepared 20 mM ammonium bicarbonate solution, lyophilized and resuspended in ITC buffer. Experiments were carried out at 25°C. During titration, 12 µl aliquots of 427 µM of Su(z)12 peptide were injected into a solution of 41.1 µM Nurf55 protein at time intervals of 5 minutes. For Nurf55 and Nurf55-Su(z)12 subcomplex binding to histone H3 tail peptides, 40-50 µM protein in ITC buffer was titrated against ~1 mM histone peptide at 25°C. ITC data was corrected for the heat of dilution by subtracting the mixing enthalpies from titrant solution injections into protein-free ITC buffer. ITC data were analyzed using the program provided by the manufacturer (Origin version 5.0), using a one-site binding model with a stoichiometry of 1:1.

Peptide-Based HMTase Assay

For the assay 7 nM PRC2 were incubated with 1 µM $H3_{1-45}$ (unmodified or carrying H4K4me3 or H3K36me3 modifications) or $H3_{21-44}$ substrate together with 2 µM SAM in reaction buffer (20 mM Tris pH 8.5, 0.5 mM DTT, 0.1% BSA, 0.01% Triton). To determine

HMTase activity the concentration of SAH generated by the enzymatic reaction was measured by LC-MS/MS using d4-SAH as internal standard. All experiments were performed on an AB Sciex (Foster City, CA, USA) API 4000 triple quadrupole mass spectrometer configured with a Shimadzu (Nakagyo-ku, Kyoto, Japan) LC-20AD liquid chromatograph. A gradient separation of SAH was carried out on a Chromolith RP-18e column (50 x 2 mm) with 0.1% formic acid in water as solvent A and 0.1% formic acid in acetonitrile as solvent B at a flow rate of 0.8 ml/min. MRM transition ions for detecting SAH and d4-SAH were *m/z* 385 -> 136 and *m/z* 389 -> 136, respectively with the same collision energy of 28 eV.

Histone Expression, Purification, Methylation, and Octamer Assembly

Drosophila histones were expressed in BL21(DE3) PlysS *E.coli* and purified through a Hi-Trap Sepharose SP-FF ion exchange column. Histone H3 used in the different methylation reactions carried a background mutation of C110A and K→C mutations at position 4, 9 or 36 on the N-terminal tail (Kc4, Kc9, and Kc36, respectively). The site-specific methylation reactions were carried out as described in Simon, 2010 and Simon et al., 2007 and the efficiency of the reaction was analyzed by Q-TOF mass spectrometry analysis. The octamers were reconstituted by mixing and refolding the core histones in equimolar ratio and purified using a Superdex200 size exclusion column (Luger et al., 1999).

Nucleosome Assembly and Histone Methyl-Transferase Assay

For mononucleosome assembly, histone octamers were assembled onto a 201 bp '601' DNA template (Thastrom et al., 1999) in the ratio of 1.2 µg octamer per 1 µg DNA in a volume of 10 µl with 2 M NaCl, 10 mM Tris-Cl pH 8.0, 0.1mM EDTA, 10 mM β-mercaptoethanol. After incubating at room temperature for 30 minutes, the nucleosomes were dialyzed successively against 1.2 M NaCl, 1 M NaCl, 0.8 M NaCl, 0.6 M NaCl for two hours each at 4^{o}C and overnight against 0.1 M NaCl buffer. For 4-mer oligonucleosome assembly, octamers were assembled on the 4 repeats of the 201 bp '601' DNA in the ratio of 1.4 µg octamer per 1 µg DNA. The same protocol as for mononucleosome assembly was followed. The assembled nucleosomes were pre-warmed for 1 hour at 37^{o}C prior to starting the HMTase reaction. The assay was carried out in a reaction volume of 35 µl, where mononucleosomes or 4-mer oligonucleosomes (as indicated in the figure) were incubated with the indicated concentration of the purified PRC2 complex and 2150 pmol S-adenosylmethionine for 2 hrs at 25^{o}C in reaction buffer (10 mM HEPES pH 7.9, 100 mM NaCl, 0.25 mM EDTA (pH 8.0), 10 mM DTT, 2.5 mM $MgCl_2$, 5% Glycerol). The reaction was stopped by adding 10 µl of 4xSDS-loading buffer and boiling the sample for 5 min. 3 µl of this sample were used per western blot. Western blot analysis was carried out using antibodies specific to H3K27me1 (Millipore #07-448), H3K27me2 (Millipore #07-452) and H3K27me3 (Millipore #07-449) in 1:3000 dilution. α-H4 (Abcam #ab10156) was used at 1:100,000 as a loading control. For quantitative western blotting, the blots were incubated in 1:1000 diluted Goat anti-rabbit Alexa 546 antibody (Invitrogen A-11010) for 2 hrs. Blots were washed and scanned using the fuji FLA7000 phosphoimager. The quantification then was done using the Multigauge software.

Eletrophoretic Mobility Shift Assay

Histone octamers were assembled onto a 201bp '601' DNA template in the ratio of 1.2 µg octamer per 1 µg DNA (100,000 cpm of ^{32}P-dCTP labeled '601' DNA was mixed with 1 µg of unlabeled DNA prior to assembly) in a volume of 10 µl with 2M NaCl, 10 mM Tris-Cl pH 8.0, 0.1 mM EDTA, 10 mM β-mercaptoethanol. After incubating at room temperature for 30 minutes, a step-wise salt dilution with 10 mM Tris-Cl pH 8.0, 0.1 mM EDTA containing 2 mg/ml BSA was carried out till the final salt concentration of 100 mM NaCl was reached. The nucleosomes were pre-warmed at 37^{o}C for 2 hrs. 0.35 pmol of nucleosomes were incubated with 0 - 2 pmol of PRC2 complex in 25 µl binding reactions (10 mM Tris-Cl pH 8.0, 100 mM

NaCl, 10% glycerol) for 1 hr at 25°C. Binding reactions were then analyzed on native 4% 0.5X Tris–borate 60:1 polyacrylamide gels. Gels were fixed in 10% methanol, 10% acetic acid for 15 min, washed with dH_2O, vacuum dried and exposed for autoradiography.

SUPPLEMENTAL REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. *66*, 213-221.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Smart, O.S., Vonrhein, C., and Womack, T.O. (2009). BUSTER, version 2.8.0. Cambridge, United Kingdom: Global Phasing Ltd.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. *60*, 2126-2132.

Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S.*, et al.* (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature *461*, 614-620.

Jacobs, S.A., Fischle, W., and Khorasanizadeh, S. (2004). Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. Methods Enzymol. *376*, 131-148.

Ketel, C.S., Andersen, E.F., Vargas, M.L., Suh, J., Strome, S., and Simon, J.A. (2005). Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Mol. Cell. Biol. *25*, 6857-6868.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol. Biol. *119*, 1-16.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods Enzymol. *276*, 307-326.

Simon, M.D. (2010). Installation of site-specific methylation into histones using methyl lysine analogs. Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al *Chapter 21*, Unit 21 18 21-10.

Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. Cell *128*, 1003-1012.

Thastrom, A., Lowary, P.T., Widlund, H.R., Cao, H., Kubista, M., and Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. J. Mol. Biol. *288*, 213-229.

3.2 The structure of Nurf55 bound to residues 1 to 30 of histone H3 - context-dependent specificity and plasticity of Nurf55

Frank W. Schmitges[1], Richard D. Bunker[1], Alexandra Stützer[2], Wolfgang Fischle[2], Nicolas H. Thomä[1]‖

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4051 Basel, Switzerland.

[2]Max Planck Institute for Biophysical Chemistry, Laboratory of Chromatin Biochemistry, Am Fassberg 11, D-37077 Göttingen, Germany.

‖to whom correspondence should be addressed: Nicolas.Thoma@fmi.ch

Abstract

The Nurf55 subunit is found in more than six chromatin modifier complexes. Its detailed molecular role remains unclear at present. Here we present the crystal structure of Nurf55 in complex with the first 30 residues of the very N-terminus of histone H3. The H3 tail binds through the canonical WD40 ligand binding interface (*c-site*), supported by a binding site located on the side of the propeller (*S/H-site*). Binding of the full H3 tail is mutually exclusive with all previously characterized Nurf55 substrates (Fog-1, Su(z)12 and histone H4). Nurf55 binds these histone and non-histone substrates with comparable affinity. We propose a model whereby Nurf55 serves in the specific recognition of structurally and functionally unrelated substrates. The absolute specificity of Nurf55 is not a property of the subunit itself, but depends on the molecular environment Nurf55 operates in.

Introduction

Drosophila Nurf55 (also known as p55, RbAp48/46 or RBBP4/7 in mammals, and MSI1 in plants) is a member of the WD40 protein family. WD40 proteins are typically ligand binding modules involved in the specific recognition of proteins, sugar and nucleic acids (Xu & Min, 2011). Nurf55 is a particularly promiscuous family member and highly conserved from yeast to man. It is known to serve in histone methyltransferase complexes including the polycomb repressive complex 2 (PRC2) (Cao et al, 2002; Czermin et al, 2002; Kuzmichev et al, 2002; Müller et al, 2002), the nucleosome assembly factor Caf-1 (Tyler et al, 1996; Verreault et al, 1996), ATP-dependent remodeling within the NURF and NuRD complexes (Hassig & Schreiber, 1997; Kuzmichev et al, 2002; Xue et al, 1998; Zhang et al, 1998; Zhang et al, 1999), histone acetylation catalyzed by Hat1 (Parthun et al, 1996) and histone deacetylation within the HDAC1 complex (Taunton et al, 1996). In addition, RbAp48/46 co-purifies with human HJURP and centromeric $CenH3^{Cenp\text{-}A}$ and has been implicated as histone chaperone in $CenH3^{Cenp\text{-}A}$ deposition (Dunleavy et al, 2009; Foltz et al, 2009), with its ortholog Mis16 serving in a related pathway in *Saccharomyces pombe* (Hayashi et al, 2004; Pidoux et al, 2009). Interestingly, the exact nature of the histone substrate acted on by Nurf55 differs significantly: while Caf-1 binds H3-H4 tetramers, Hat1 targets H3-H4-dimers or -tetramers, and PRC2 and NuRD complexes function on fully assembled nucleosomes (Suganuma et al, 2008).

Considering the multitude of Nurf55-containing complexes and the different kinds of histone substrates it is not surprising that Nurf55 possesses at least two substrate recognition sites. Recent structural studies have identified two major binding interfaces on Nurf55: histone H4 is bound *via* a pocket formed by helix α1 and the PP-loop on the side of the WD40 propeller (Murzina et al, 2008; Nowak et al, 2011; Song et al, 2008). The same binding site (hereafter referred to as *S/H-site*) is used to attach Nurf55 to the PRC2 core complex, by interacting with a conserved region in the N-terminus of Su(z)12 (Schmitges et al, 2011). Binding of Su(z)12 and histone H4 to the *S/H-site* is mutually exclusive. A second binding site is provided by the canonical WD40 ligand binding interface located on the flat face of the WD40 propeller (*c-site*). This site has been shown to bind to the N-terminus of the GATA-1 cofactor Fog-1 (Lejon et al, 2011) and the unmodified N-terminal tail (residues 1-19) of histone H3 (Schmitges et al, 2011). The binding affinity for histones and non-histones to the *c-site* is ~1μM, a range typically observed for epigenetic reader domains. Interestingly, the

Nurf55 *c-site* is unable to recognize peptides carrying an H3K4me3 modification thereby serving as a reader sensitive to epigenetic marks (Schmitges et al, 2011).

While mutational studies clearly demonstrate a biological role of Nurf55 in Hat1 (Song et al, 2008), NuRD (Lejon et al, 2011) and PRC2 function (Anderson et al, 2011), it is enigmatic how Nurf55 can serve as a specific binding platform for each of those very different substrates. No common substrate consensus has been derived for Nurf55. In fact, Nurf55 is frequently assumed to function as a structural component of these complexes.

Here we present the structural basis of Nurf55-mediated recognition of the entire histone H3 N-terminal tail. The 2.1 Å structure shows Nurf55 bound to residues 1-42 of histone H3 peptide (hereafter referred to as $H3_{1-42}$), occupying both the *S/H-site* and the *c-site*. This new binding mode is not compatible with the binding of H4, Su(z)12, Fog-1, or $H3_{1-15}$. We present evidence that H3 can be bound in two registers, with residue H3K4 at the *c-* or *S/H-site,* respectively. The data presented suggest that binding of the different Nurf55 substrates, and hence the different binding modes, are mutually exclusive each depicting a different working mode of Nurf55. Specificity is provided by the complex partners surrounding it. While Nurf55 is promiscuous, having a broad R(x)K substrate consensus for its *c-site*, it nevertheless recognizes different substrates in a specific fashion.

Results

Histone H3 is able to bind to the Nurf55 *c-site* and *S/H-site*

The previously reported structure of Nurf55-$H3_{1-19}$ detailed the specific interactions between the first 11 amino acids of histone H3 and the Nurf55 *c-site*. In the structure H3 residues Arg2 and Lys4 are held in an acidic binding pocket close to the central cavity of the propeller (Schmitges et al, 2011). Consequently, we observed that a Nurf55-Su(z)12 complex carrying a N132K mutation in the acidic *c-site* pocket, was unable to bind to a synthetic histone $H3_{1-31}$ peptide (**Figure 1**). However, $Nurf55_{N132K}$ alone, i.e. with its *S/H-site* unoccupied, surprisingly bound the $H3_{1-31}$ peptide with an affinity similar to that of wild-type Nurf55. This prompted us to investigate the possibility of different binding modes for H3 peptides.

The structure of histone H3 residues 1 to 42 bound to Nurf55

We co-crystallized Nurf55 in the presence of a recombinant histone $H3_{1-42}$ peptide (**Table1**). The structure was refined to 2.1 Å resolution showing clear density mF_o-DF_c for H3 residues 1 to 32 (**Supplemental Information 1**). The recombinant peptide still contains a linker sequence of eight amino acids on the N-terminus that is not visible in the structure. The remaining amino acids from residues 33 to 42 are disordered. In contrast to the previously published Nurf55-$H3_{1-19}$ model, $H3_{1-42}$ occupies both *S/H-site* and *c-site* simultaneously (**Figure 2A-D**). Binding proceeds with a concomitant change in binding register placing residues R26 and K27 at the *c-site* (**Figure 2A-C**). Binding of $H3_{1-42}$ to Nurf55 induces an alpha-helical conformation in H3 residues 3 to 10 (designated $H3_{\alpha0}$). The most extensive interactions between $H3_{\alpha0}$ and Nurf55 are formed *via* R8 contacting the C-terminal part of the PP-loop where it hydrogen bonds to the Nurf55 backbone residues L370, P367, G366 and D362 (**Figure 2D**).

Hydrophobic interactions mediate substrate binding in the S/H-site

Helix $H3_{\alpha0}$ occupies the same binding cleft as the previously mapped binding epitope of H4 and the Su(z)12 Nurf55-binding epitope (NBE, **Figure 3A-D**). The helical structures of H3 and H4 are well superimposable, but differ from the conformation of the Su(z)12 peptide backbone (**Figure 3B**). Compared to the histone helices Su(z)12 NBE is shifted towards the helix α1 of Nurf55, stabilized through additional hydrophobic interactions of its side chains. Importantly, the binding modes differ in their orientation. While $H3_{1-42}$ and H4 have their N-

terminus close to the bottom of the WD40 propeller and continue up towards the *c-site*, the Su(z)12 NBE is bound with opposite directionality. The location of the H3R8 guanidinium group is reminiscent of that of Arg85 in Su(z)12 and Arg39 in H4 (**Figure 3B**). Apart from that the three binding epitopes have little in common and no consensus sequence can be determined. Of note are the superimposable side chains of H3K4 and H3K9 on the hand and H4R35 and H4R40 on the other hand which are, however, all solvent exposed. The bulk of interactions is mediated by hydrophobic interactions in all three structures, further supported by hydrogen bonding of Nurf55 with the peptide backbones. In the Nurf55-$H3_{1-42}$ structure the side chains of Arg2 and Lys4, which form the majority of contacts in the Nurf55-$H3_{1-19}$ structure are solvent exposed. Residues 15 to 22 form a loop at the top edge of the propeller, packing against Nurf55 helix α1 before connecting the *c-* with the *S/H-site* (**Figure 2B**). Stabilization of this loop is provided by H3R17 contacting the backbone oxygen of H3 residues Ala15 and Thr22, thereby clamping the loop in a tight 90° turn.

<u>The plasticity of the c-site allows different binding registers for the histone H3 tail</u>

$H3_{1-42}$ residues Thr22 to Thr32 cross the propeller at the c-site in a path and orientation similar to that seen for $H3_{1-19}$ and Fog-1 (**Figure 3A**). In contrast to $H3_{1-19}$, however, $H3_{1-42}$ positions residues Arg26 and Lys27 in the *c-site*. This R(x)K consensus sequence is also shared with the Fog-1 peptide: the ε-amino group of $H3_{1-42}$K27/$H3_{1-19}$K4 and the guanidinium-group of $H3_{1-42}$R26/$H3_{1-19}$R2 are found in similar positions as those of Fog-1 Arg3 and Lys4 (**Figure 3A**). Differences in spacing due to the interspersed Thr3 in $H3_{1-19}$ are off-set in $H3_{1-42}$ and Fog-1 by the arginine residue projecting away from the lysine, while in contrast the side chains of Arg2 and Lys4 are oriented in parallel. The peptide backbone of $H3_{1-19}$ stays at the propeller surface, keeping the T3 residue solvent exposed, while $H3_{1-42}$ and Fog-1 are buried deeper in the binding pocket. The pocket engaged by the amino terminus of $H3_{1-19}$ is occupied by the ε-amino group of $H3_{1-42}$K23 and by the guanidinium-group of Arg2 of Fog-1. Apart from the R(x)K motif the three binding epitopes share no real consensus sequence (**Figure 3E**). Gln5 and Lys9 of $H3_{1-19}$ are superimposible with Gln6 and Arg10 of Fog-1, respectively. However, these residues are solvent exposed and interactions outside of the R(x)K site are mainly mediated by the peptide backbone. Thus, various peptides, with very different primary sequences, are bound at the *S/H-* and *c-sites* with a considerable degree of plasticity.

Based on our work, two different binding registers have been observed with H3 peptides of different lengths. The two binding modes have been validated by binding experiments *in vivo* and we hypothesize that both exist in parallel *in vivo* but in the context of different complexes. The $H3_{1-19}$ mode is the only possible way to bind H3 while the *S/H-site* is occupied: fluorescence polarization (FP) experiments have demonstrated that $H3_{1-19}$ and $H3_{1-31}$ can interact with a Nurf55-Su(z)12 complex; an $H3_{19-38}$ peptide, however, showed no binding (Schmitges et al, 2011) indicating that the H3R26/K27 interaction with the c-*site* is not sufficient for stable binding. On the other side, the $H3_{1-42}$ binding register described in this study is the only possible mode when binding of H3R2/K4 to the *c-site* is impeded by (i) mutations or (ii) methylation of H3K4: (i) $H3_{1-31}$ is able to bind $Nurf55_{N132K}$ alone but not in the context of the Su(z)12 complex (**Figure 1**), again demonstrating the importance of the *S/H-site* for this interaction. (ii) $H3_{1-28}$K4me3 still binds to Nurf55 if the *S/H-site* is available (Nowak et al, 2011).

Discussion

Nurf55 is a versatile H3/H4 chromatin binder

Recent structural studies have described two binding interfaces on the Nurf55 WD40 domain, each with 3 potential binding partners: histone H4 (Murzina et al, 2008; Nowak et al, 2011; Song et al, 2008), Su(z)12 (Schmitges et al, 2011) and histone $H3_{1-42}$ (this study) are recognized through the Nurf55 *S/H-site*, while histones $H3_{1-19}$ (Schmitges et al, 2011), $H3_{1-42}$ (this study) and Fog-1 (Lejon et al, 2011) bind to the *c-site*. Additional binding sites most likely exist on the β-propeller connecting Nurf55 to the core of the six known Nurf55-containing complexes. The affinities for those binding partners that utilize the *c-site* lie in the low micromolar range (see $H3_{1-19}$ and Fog-1 studies), while those solely occupying the *S/H-site* (H4 and Su(z)12) appear to be bound slightly tighter (Nowak et al, 2011; Schmitges et al, 2011).

Mutations of Nurf55 prove functional relevance of Nurf55 in different complexes

Several mutational studies have investigated the role of Nurf55 both *in vitro* and *in vivo*: disruption of the H4-interaction has been reported to interfere *in vitro* with histone acetylation in the context of the Hat1 complex (Song et al, 2008). The association of Nurf55 with deacetylation activity was observed in co-immunoprecipitation experiments using *Drosophila* embryos (Tyler et al, 1996). Loss of Su(z)12 binding leads to dissociation of Nurf55 from the PRC2 core complex (Song et al, 2008), and knockout of Nurf55 in *Drosophila* results in reduction of H3K27me3, characteristic for PRC2 malfunction (Anderson et al, 2011). In general, mutational studies in flies and plants (Anderson et al, 2011; Hennig et al, 2003) reported pleiotropic phenotypes reflecting the various biological roles of Nurf55. Together these findings raise the question how Nurf55 is able to recruit the correct binding partner for any given function.

Determinants of Nurf55 substrate specificity

Histone H4 helix 1 has significantly higher affinity for the *S/H-site* compared to the H3 N-terminal tail. When Nurf55 is confronted with H3/H4 dimers (or tetramers) competing for the *S/H-site*, the H4-*S/H-site* interaction is expected to dominate. Yet H4 helix 1 is not accessible in fully assembled nucleosomes (Luger et al, 1997). Therefore, Nurf55 in the PRC2 complex, which targets assembled nucleosomes, binds to H3. In the PRC2 complex Su(z)12

constitutively occupies and blocks the *S/H-site* (Schmitges et al, 2011) and thereby enforces a register where H3K4 is bound to the c-site. If both *c*- and *S/H-sites* are available, however, Nurf55 is expected to bind the entire H3 tail of assembled nucleosomes using both *c*- and *S/H-sites* (**Figure 3D**).

An additional level of specificity of Nurf55 binding may be provided by cooperative interactions between other complex components and its substrate: in the NuRD complex both Nurf55/RbAp48 and MTA-1 simultaneously bind to different Fog-1 molecules in the same region (Lejon et al, 2011). Yet, Fog-1 binding to Nurf55/RbAp48 is not observed in the absence of MTA-1. A single Nurf55-substrate interaction (i.e. Nurf55-Fog-1; Nurf55-H3) may therefore by itself be too weak. Additional weak interactions (e.g. Fog-1-MTA-1 complex, Su(z)12-chromatin) may be needed to specifically recruit the complex to its target site. Each Nurf55 binding mode thereby represents a different "life" of Nurf55 determined by the macromolecular context it functions in.

Context-dependent specificity of WD40 propellers, a more general property?

While most WD40 proteins are regarded as specific recognition modules serving in binding of only one type of substrate, we wondered whether other examples of context specificity exist. The epigenetic reader WDR5 has been identified as a core component of the MLL1 complex (Wysocka et al, 2005) but is also associated with various histone acetyltransferase complexes (Suganuma et al, 2008). The *c-site* of this WD40 propeller was first shown to recognize histone H3 (Couture et al, 2006; Han et al, 2006; Ruthenburg et al, 2006; Schuetz et al, 2006). Later the same binding site was found to interact with the MLL1 SET-domain (Patel et al, 2008; Song & Kingston, 2008). Although these peptides show a more pronounced primary sequence consensus than that seen for Nurf55 substrates, they nevertheless differ significantly. Yet, MLL1 and H3 are recognized by WDR5 with similar affinity. In analogy to Nurf55, we thus hypothesize that other WD40 proteins, such as WDR5, exist whose specificity is context-dependent. Their binding properties are expected to provide a low affinity binding platform, which can be easily customized for targeting and regulatory purposes in a given complex.

Experimental Procedures

Protein Purification and Crystallization

PRC2 proteins were recombinantly expressed and purified as described (Schmitges et al, 2011). Recombinant histone $H3_{1-42}$ was expressed with an N-terminal 6xHis-tag. The tag was removed by cleavage with Tobacco Etch Virus (TEV) protease leaving an eight amino acid linker (GAHMSGRM-). Crystals were grown at 20–25°C by the hanging drop vapor diffusion method. Nurf55 protein was incubated with a ~5-fold molar excess of recombinant $H3_{1-42}$ peptide for 30 min prior to crystallization. 1µl drops of a 16 mg/ml protein solution were mixed with 1 µl of reservoir solution containing 100 mM MES, pH 6.5 and 40% PEG200. A detailed description of the experimental procedures is available in the Supplemental Experimental Procedures.

Fluorescence polarization measurements

For fluorescence polarization measurements the $H3_{1-31}$ peptide (fluorescein-ARTKQTARKSTGGKAPRKQLATKAARKSAPA) was chemically labeled post peptide synthesis using NHS-fluorescein; single-labeled species were purified by RP-HPLC. FP assays were essentially carried out and analyzed as described before (Jacobs et al, 2004). Titration series of protein in 10 µl 50 mM HEPES pH 7.4, 150 mM NaCl, 0.25 mM TCEP with a final concentration of fluorescinated peptide of 1 nM were read multiple times in 384 well plates on a Plate Chameleon II plate reader (HIDEX Oy). Multiple readings and independent titration series were averaged after data normalization.

Figure Legends

Figure 1: Binding of Nurf55 to a histone $H3_{1-31}$ peptide quantified by FP.

Wildtype Nurf55 binds to a histone $H3_{1-31}$ peptide with an affinity of 2.2 ± 0.2 μM (black). A N132K mutation in the Nurf55 *c-site* (blue) has only little effect while the same mutation in the context of the Nurf55-Su(z)12_{73-143} complex (red) results in a loss of binding.

Figure 2: Crystal structure of Nurf55 in complex with a histone $H3_{1-42}$ peptide.

(A) Ribbon representation of Nurf55-$H3_{1-42}$. Nurf55 is shown in rainbow colors and $H3_{1-42}$ is depicted in cyan. The peptide occupies both the *S/H-* and the *c-site* of the WD40 propeller.

(B) Nurf55 *S/H-site* with the $H3_{1-42}$ peptide (cyan) bound between the helix α1 and the PP-loop.

(C) Histone $H3_{1-42}$ residues K23, R26 and K27 are bound to the Nurf55 *c-site*.

(D) Schematic representation of interactions between Nurf55 (yellow) and histone $H3_{1-42}$ (cyan) in the *S/H-site* (top panel) and the *c-site* (bottom panel).

Figure 3: Comparison of different Nurf55 binding modes.

(A) Upper panel: Overlay of histone $H3_{1-19}$ (green) and $H3_{1-42}$ (cyan) bound to the Nurf55 *c-site*. Lower panel: Overlay of Fog-2 (pink) and $H3_{1-42}$ (cyan) bound to the Nurf55 *c-site*. Note that the arginine and lysine residues of $H3_{1-19}$ are bound in a similar fashion but in a different register compared to the other peptides.

(B) Overlay of histone $H3_{1-42}$ (cyan), Su(z)12_{79-91} (magenta) and histone $H4_{31-41}$ (orange) in the Nurf55 *S/H-site*. All three epitopes contain prominent arginine residues that are recognized by the same binding pocket.

(C) Histone $H3_{1-19}$ (green) binds to the *c-site* while Su(z)12_{79-91} (magenta) binds to the *S/H-site*. The binding sites are independent and both peptides can bind at the same time.

(D) Histone $H3_{1-42}$ occupies both the *S/H-* and the *c-site*, the binding epitopes are connected by a loop. This binding mode is not compatible with Su(z)12 or histone H4 binding.

(E) Sequence alignment of the peptides that bind to the Nurf55 *c-site*. Apart from the R(x)K motif (red) no consensus sequence can be determined.

Table 1: Crystallographic data and refinement statistics

	Nurf55 – H3$_{1-42}$
Space group	$P2_1$
Unit cell dimensions	
***a*, *b*, *c* (Å)**	65.57 58.97 67.00
β (°)	99.36
Resolution range (Å)	66.1 – 1.9 (1.909-1.900) [a]
Percent complete	96.7 (74.0) [a]
Redundancy	5.0 (3.7) [a]
R_{sym}	0.066 (0.547) [a]
$I / \sigma I$	17.5 (2.5) [a]
Resolution (Å)	1.9
No. reflections	38506
R_{work} / R_{free}	0.165 / 0.191
No. atoms	3645
***B*-factors (Å^2)**	
Protein	36.0
Water	41.9
R.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.120

[a]The values for the data in the highest-resolution shell are shown in parentheses.

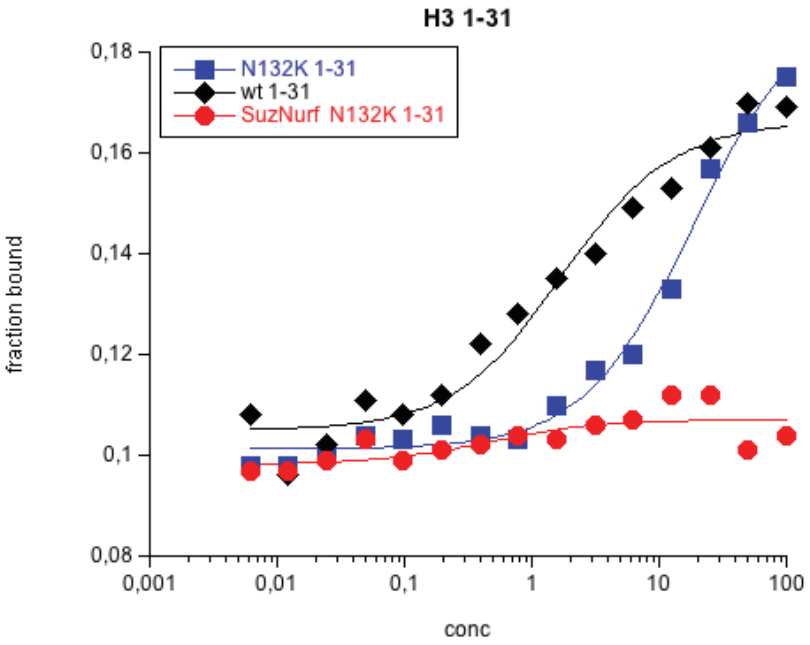

H3 1-31
N132K 1-31
wt 1-31
SuzNurf N132K 1-31
fraction bound
0,18
0,16
0,14
0,12
0,1
0,08
0,001
0,01
0,1
1
10
100
conc

Figure 1

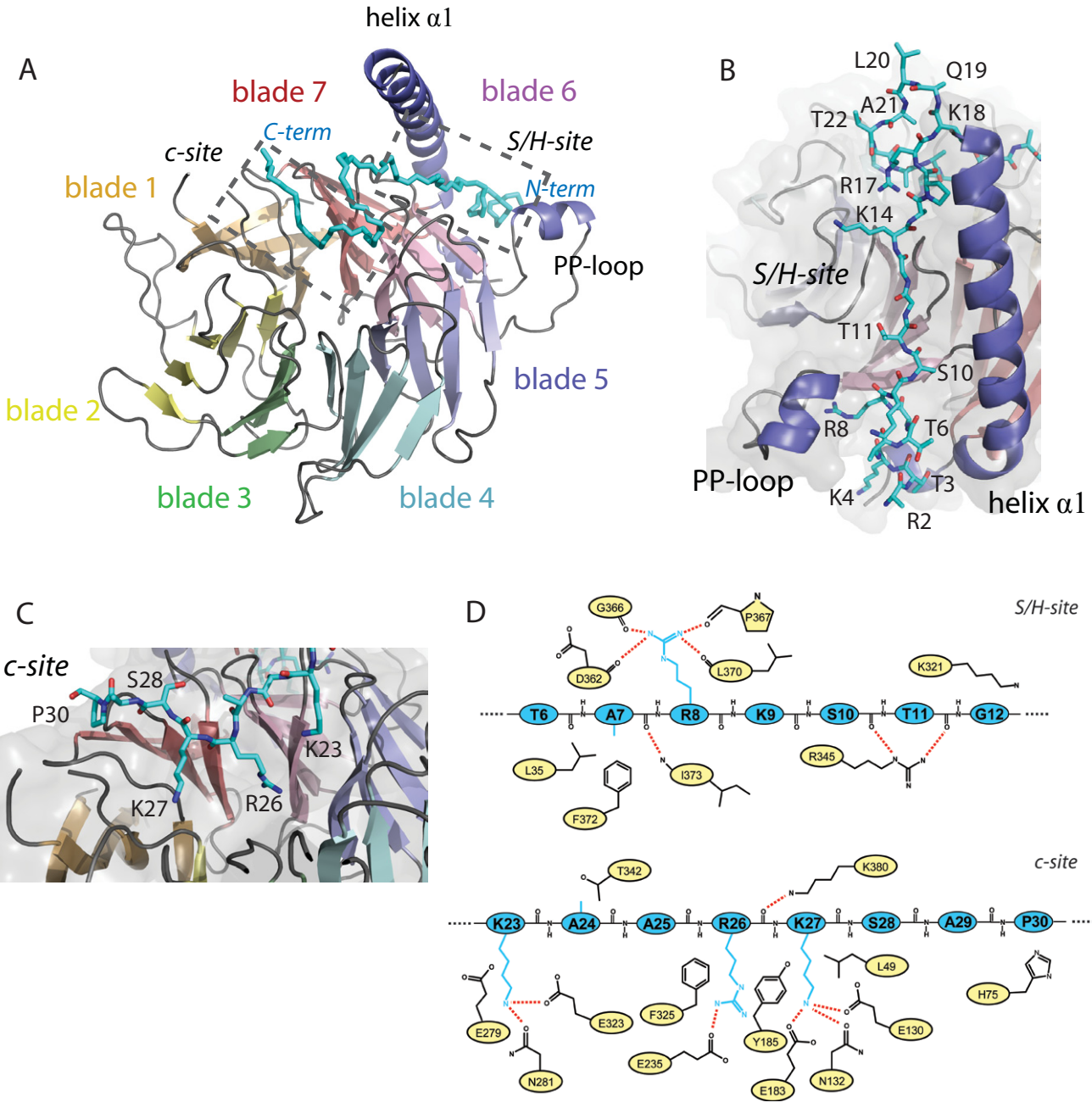

A
helix α1
blade 7
blade 6
C-term
S/H-site
c-site
blade 1
N-term
PP-loop
blade 5
blade 2
blade 3
blade 4
B
L20
Q19
A21
T22
K18
R17
K14
S/H-site
T11
S10
R8
T6
PP-loop
K4
T3
R2
helix α1
C
c-site
S28
P30
K23
K27
R26
D
S/H-site
G366
P367
D362
L370
K321
T6
A7
R8
K9
S10
T11
G12
L35
I373
R345
F372
c-site
T342
K380
K23
A24
A25
R26
K27
S28
A29
P30
L49
H75
E279
E323
F325
E130
Y185
E235
N132
E183
N281

Figure 2

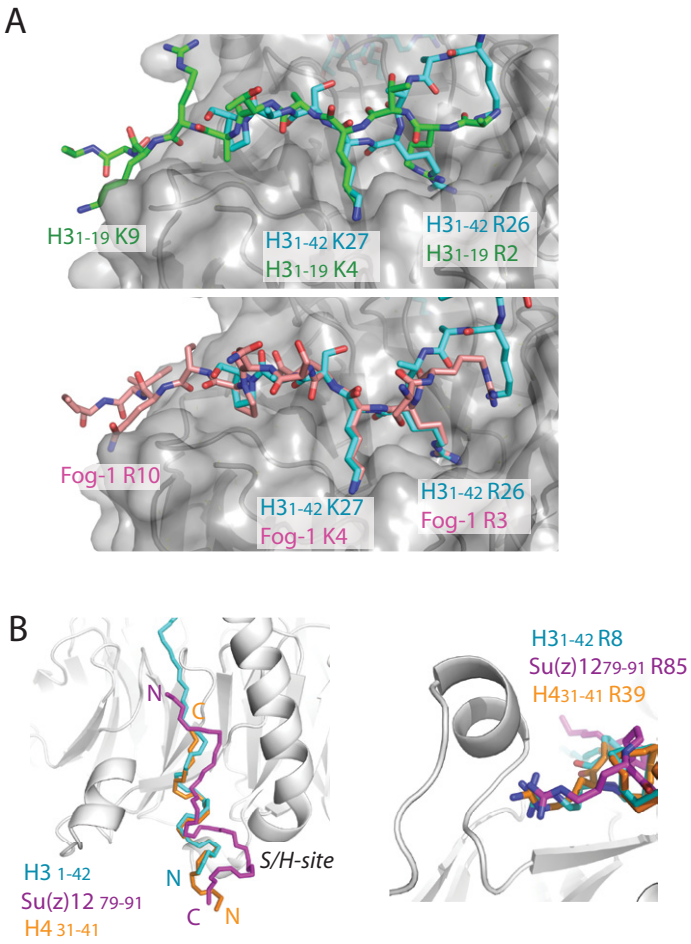

A
H31-19 K9
H31-42 K27
H31-19 K4
H31-42 R26
H31-19 R2
Fog-1 R10
H31-42 K27
Fog-1 K4
H31-42 R26
Fog-1 R3
B
N
C
S/H-site
H3 1-42
Su(z)12 79-91
H4 31-41
N
C
N
H31-42 R8
Su(z)1279-91 R85
H4 31-41 R39
C
H3 1-19
Su(z)12 79-91
D
H3 1-42
E
Histone H3 1- artkqtarkst
Histone H3 23- kaark sapstgg
Fog-1 1- msrrkqsnprqi

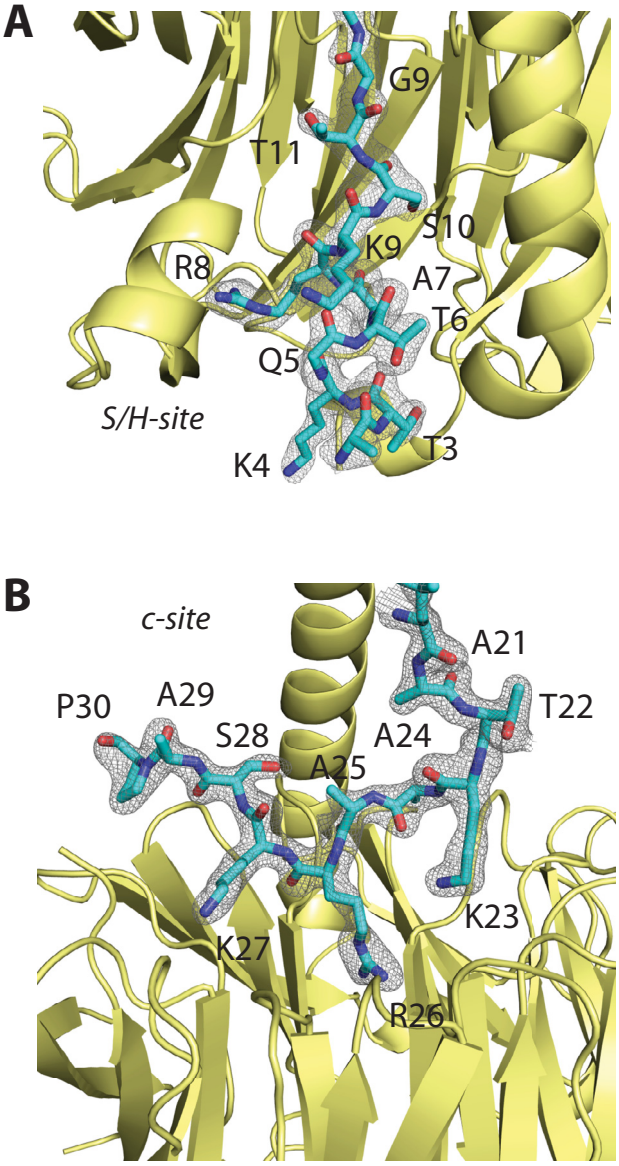
Figure 3

Supplemental Figure 1

Crystal structures of Nurf55-H3$_{1-42}$

(A) Simulated annealing composite omit map (2mFo-DFc) of Nurf55-H3$_{1-42}$. The H3$_{1-42}$ peptide (cyan) in the *S/H*-site of Nurf55 is contoured at *1* σ.

(B) Simulated annealing composite omit map (2mFo-DFc) of Nurf55-H3$_{1-42}$. The H3$_{1-42}$ peptide (cyan) in the *c*-site of Nurf55 is contoured at *1* σ.


A
G9
T11
S10
K9
R8
A7
T6
Q5
S/H-site
K4
T3
B
c-site
A21
P30
A29
T22
S28
A24
A25
K23
K27
R26

5. General Discussion

5.1 Bistability

Biological systems are frequently regulated through bi- or multistable switches. These regulatory systems allow to populate one of two or more basic states. Biological switches execute fundamental decisions about life and death of a cell (Bagci et al, 2006; Cui et al, 2008; Ho & Harrington, 2010), cell division (Ingolia, 2005), cell fate and differentiation (Yan et al, 2009) and neuronal signaling (Lee & Heckman, 1998; Malashchenko et al, 2011). While the switch stably retains a given state for prolonged periods of time, it needs to respond swiftly to appropriate external stimuli, triggering the transition into another local energetic minimum. The activation energy required to overcome the barrier between these different ground states allows simultaneous integration of positive and negative signals. In addition, it provides the system with a certain degree of memory, as it will retain the current state even in the absence of the initiating stimuli, until a new signal reverses the *status quo*.

5.2 Components of bistable systems

Positive feedback loops are a central element of all bistable switches (Ferrell, 2002). They ensure that no intermediate states can exist. Once a threshold is reached, the signal amplification will enforce adaptation of a new ground state. A ground state represents a local minimum, which is separated from other local minima through a kinetic barrier. The positive feedback loop allows to overcome this barrier in a unidirectional fashion, and thus prevents rapid fluctuations between multiple states. The signal threshold further acts as a filter, which blocks activation through small stimuli making the system more inert. The positive feedback loop, however, also comes with inherent problems: continuous signal amplification is expected to result in signal explosion (Wilhelm, 2009). A circuit breaker is thus required delimiting an otherwise endless amplification. My work on the PRC2 complex has revealed such a circuit breaker in the process of the epigenetic inheritance of transcriptional signatures. In this discussion I will first review classical switches in cell biology and classical epigenetic switches, and then contrast those with our findings regarding the activation and inhibition of the PRC2 complex.

5.3 Classical switches in cell signaling and apoptosis

5.3.1 Unidirectionality in cell biological switches

The best-characterized biological feedback loops are found in the cell cycle. They safeguard the transition between the different phases of the cell cycle and halt transition through cell division until processes such as replication and repair are properly initiated and completed. Importantly these feedback loops ensure that the cell cycle only proceeds in one direction; a reversal would lead to serious chromosomal abnormalities. The center of cell cycle control consists of a network of cyclins and cyclin-dependent kinases (CDKs). Cyclins are the regulatory subunits which bind to the kinases and by phosphorylation activate or inhibit downstream target proteins. The onset of mitosis is regulated by phosphorylation and dephosphorylation of the kinase Cdc2. Phosphorylation of Cdc2 by the kinase Wee1 keeps it in an inactive state. Upon activation, Cdc2 is able to phosphorylate and thereby inactivate its own inhibitor Wee1 and at the same time to activate the phosphatase Cdc25 which keeps Cdc2 in an active state. Cdc2 activity is essential for DNA replication initiation and centrosome duplication. Once active, Cdc2 is able to counteract its own inactivation, thereby creating a positive feedback that prevents the cycle from slipping back into interphase (Sha et al, 2003). The filter and circuit breaker in this system are both found in form of the Cdc2 co-activator cyclin B. The concentration of cyclin B is a crucial threshold for Cdc2 activation in the first place. Once activated, the positive feedback loop of Cdc2 is limited by a negative feedback where Cdc2 drives degradation of cyclin B *via* phosphorylation of Fizzy.

5.3.2 Thresholding in biological systems

The programmed cell death, or apoptosis, is a key process in development, regeneration and protection from cancer and infections. Its initiation is tightly controlled by decision-making circuits that allow the integration of pro- and antiapoptotic signals in the formation of the apoptosome (Bagci et al, 2006). In the canonical pathway the permeability of the mitochondrial membrane constitutes an important threshold that is modulated by pro- and anti-apoptotic proteins (Green & Reed, 1998). While this filter tolerates low levels of DNA damage or toxins that do not pose a threat to the cell, significant damage causes mitochondrial depolarization. The release of cytochrome c and pro-apoptotic proteins from the mitochondria eventually marks the point of no return and the start of the degradation of the cell. Positive

feedback loops execute the decision downstream of mitochondrial depolarization and prevent abortion of the cell death program. Small amounts of cytochrome c trigger calcium release from the ER, which in turn leads to the mass exodus of cytochrome c from the mitochondria (Boehning et al, 2003).

5.3.3 Feedback present in enzymatic complexes

Positive feedback loops have also been described for enzymes serving in the regulation of enzymatic activity and processivity. The kinase Hck is a member of the Src family of tyrosine-protein kinases that plays an important role in the transmission of cellular responses to extracellular stimuli (Brown & Cooper, 1996). In addition to its Src kinase domain Hck harbors the regulatory SH2 and SH3 domains. The SH2 domain of Hck has been shown to bind specifically to sites phosphorylated by the Hck kinase domain (Pellicena et al, 1998). The creation of high affinity SH2 binding sites by kinase activity increases the local Hck concentration, constituting a positive feedback that promotes the processive phosphorylation of target proteins.

5.4 Positive feedback loops in transcriptional regulation

How do these regulatory circuits apply to the regulation of gene expression? Gene expression is a tightly regulated and highly dynamic process. Under steady-state conditions the transcriptional state is faithfully maintained over many generations of cell division, while activation, or silencing, of genes occurs quickly in response to specific cues, e.g. during differentiation or in response to changing environmental conditions. Bistable switches have clear advantages in transcriptional regulation as they supply a long-term memory that can be defined and maintained at the onset of differentiation (Fischle et al, 2003a; Grunstein, 1998; Micheelsen et al, 2010; Sneppen et al, 2008). Chromatin modifications are deposited by large multi-protein complexes (reviewed in Taverna et al, 2007; Wang & Patel, 2011; Yun et al, 2011) and the transcriptional state appears to strongly correlate with the type of modification present in a chromatin region (Jenuwein & Allis, 2001). DNA replication is a potentially destabilizing event for epigenetic regulation: the histones of the original DNA strand are distributed on to the daughter strands and new histones are incorporated to fill in the gap (Annunziato, 2005). The regulatory machinery has to ensure that histone modifications are

copied quickly and faithfully to the daughter strand thus preventing dilution and eventual loss of the epigenetic information. Based on this study and others, a very attractive hypothesis is beginning to emerge, where the marks present on old histones are used as templates: activating modifier complexes are targeted to active chromatin marks, while repressing modifiers are targeted to silenced regions. Positive feedback loops would thereby be provided with the necessary signal to faithfully replicate epigenetic signatures (**Figure 2**). In the following sections I will review our present understanding of this process, focusing on our current knowledge concerning the inheritance of biological states and the involvement of positive feedback loops in this process.

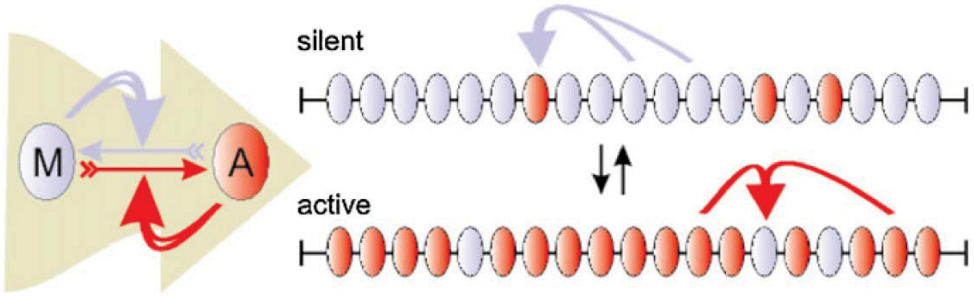


Figure 2: Model for propagation of epigenetic marks. Nucleosomes with repressive chromatin modifications help to maintain silent chromatin (light blue, M=modified), active chromatin marks (red, A=anti-modified) in contrast consolidate an active chromatin state (taken from Micheelsen et al, 2010).

5.5 Establishment and maintenance of heterochromatin in yeast and mammals

Saccharomyces cerevisiae has long served as a model for inheritance and spreading of heterochromatin. Silencing of the mating type locus is mediated by Sir proteins. Initial binding of a Sir protein complex at the silencer leads to deacetylation of neighboring H3 and H4 tails by the NAD-dependent histone deacetylase Sir2 (Hoppe et al, 2002; Rusche et al, 2002; Sedighi & Sengupta, 2007). The hypoacetylated histone tails provide high affinity binding sites for Sir3 and Sir4 proteins which in turn recruit further Sir2 deacetylase molecules and can lead to the sequential spreading of heterochromatin domains. The requirement for specific initiaton factors such as Sir1 and the origin recognition complex can be understood as the filter that prevents spontaneous formation of heterochromatin (Fox et al, 1997). The spreading is driven independently of these factors by a feedback loop in which the Sir proteins create their own recruitment platform. Limitation of the spread is, however, easy in this system due to the short range of Sir2 deacetylation. Short patches of nucleosome-

depleted DNA, or the localized activity of a histone acetylase, have been suggested to deprive the Sir complex of its binding platform and to serve thereby as a circuit breaker. This is in agreement with observations that silencing by Sir proteins is tightly limited to the target domain (Rusche et al, 2002). Spreading of the Sir2,3,4 complex and associated histone deacetylation is additionally counteracted at telomeres by the action of telomeric proteins, in a fashion that is believed to be independent of chromatin marks.

Epigenentic switches also regulate the maintenance of heterochromatin in mammals. Constitutive heterochromatin is enriched in trimethylated H3K9 (H3K9me3), which in mammals is mediated by the SUV39H1 and SUV39H2 enzymes (Peters et al, 2001). Members of the HP1 family are mediators of gene silencing known to bind heterochromatin *via* the specific interaction of their chromodomain with the histone H3K9me3 mark (Bannister et al, 2001; Fischle et al, 2003b; Lachner et al, 2001). SUV39H1 itself interacts with HP1, suggesting a self-maintaining positive feedback loop for the modification of new histones in heterochromatin after DNA replication. However, although H3K9me3 is required for targeting of HP1, it does not seem to be sufficient but requires a structural RNA of unknown origin (Maison & Almouzni, 2004). By analogy with silencing by Sir proteins, such targeting mechanisms are important filters that prevent feedback loops in the wrong location. Once initiated, H3K9 methylation can be still influenced by other chromatin modifications. Phosphorylation of histone H3 at Thr3 or Tyr41 has been shown to inhibit binding of HP1 and can therefore act as circuit breakers (Dawson et al, 2009; Fischle et al, 2005).

5.6 Allosteric activation of chromatin modifiers

While most chromatin modifying complexes are regulated by marks that are already present on a recruitment level only few examples have been described for allosteric stimulation. Dot1L is a non-SET methyltransferase specific for H3K79 with a role in transcriptional silencing. Ubiquitination of histone H2b has been shown to correlate with H3K79 methylation *in vivo* (Briggs et al, 2002; Ng et al, 2002) and *in vitro* studies with chemically ubiquitinated H2b indicate a direct allosteric activation of Dot1L (McGinty et al, 2008). Additional examples for allosteric control of chromatin modifiers have recently been described in the crosstalk of histone marks and DNA methylation. The *de novo* DNA methyltransferase Dnmt3a was found to be allosterically activated by unmethylated H3K4 (Li et al, 2011). A role for H3K9 methylation in DNA methylation was already established earlier (Fuks et al,

2003; Smallwood et al, 2007) but only H3K4 has been reported to stimulate methylation in an allosteric fashion.

5.7 Reversing chromatin modifications

In epigenetic switches filters usually safeguard the start of positive feedback loops. In addition the system also includes means to modify and interrupt feedback loops after their initiation as a "repair" function. Specific enzymes such as demethylases, deacetylases and phosphatases have been identified for many chromatin modifications that are able to directly erase these marks, thus reversing the signal. An alternative way to erase epigenetic modifications has been described recently: gene activation can be achieved by enzymatic clipping of modified histone tails (Santos-Rosa et al, 2009). With their activity these enzymes contribute to the drastic changes in the epigenetic landscape during development. They can modulate the threshold of filters or act as circuit breakers of feedback loops.

Furthermore, specific reversal of histone marks has been shown to enhance positive feedback loops. Mathematic modeling of bistable systems has underlined the importance of chromatin modifiers cooperating with complexes that remove competing modifications in the same area (Dodd et al, 2007). In *Drosophila* the histone H3K9 methyltransferase Su(var)3-9 has been found associated with the histone deacetylase HDAC1 (Czermin et al, 2001). The deacetylation activity was found to be essential for the methylation of pre-acetylated histone tails and thereby for the establishment of heterochromatin. Similar functional associations have been described for the MLL2 complex responsible for H3K4 methylation and the H3K9 specific demethylase JMJD2B (Shi et al, 2011). In summary the cooperation of different chromatin modifiers is used to amplify feedback loops and to overcome boundaries originating from opposing epigenetic marks.

5.8 PRC2 possesses the required circuitry making it a bistable switch

In this study we examined how these regulatory principles apply to the PRC2 complex. Recent structural and functional insights indicate that the methyltransferase activity of PRC2 is influenced by existing chromatin marks: (a) PRC2 is largely inefficient in the absence of a positive feedback loop boosting its enzymatic activity, (b) it is sensitive to boundaries of

repressive chromatin that limit the spreading of its marks and (c) it is able to integrate positive and negative signals making the switch more robust.

Initiation of polycomb-mediated silencing is a complex process that requires targeting factors and the presence of response elements in the DNA (Ringrose & Paro, 2004; Ringrose & Paro, 2007). I will focus my analysis on the propagation of polycomb marks after DNA replication that is regulated by a self-maintaining feedback cycle.

5.8.1 Positive feedback and inhibition in the regulation of the PRC2 complex

The PRC2 subunit EED has been shown to bind H3K27me3, the product of PRC2 and a typical mark of repressive chromatin (Margueron et al, 2009). The interaction with H3K27me3 on neighboring nucleosomes stimulates the methyltrasferase activity of the EZH2 SET domain and results in a positive feedback loop that allows effective deposition of H3K27me3 marks in repressive chromatin (Hansen et al, 2008; Margueron et al, 2009).

Like other positive feedback loops PRC2-driven H3K27 methylation requires a circuit breaker as safeguard. For PRC2 this implies preventing the spread of repressive chromatin into transcriptionally active domains. The results presented in this study provide a model for protection of chromatin from PRC2 methylation via direct allosteric inhibition by active chromatin marks. Histone H3K4me3 and H3K36me2/3 modifications inhibit PRC2 activity on the same nucleosome, probably even specifically on the same histone tail.

5.8.2 Propagation of H3K27me3 after DNA replication

After DNA replication the old nucleosomes carrying modifications are distributed stochastically between the two daughter strands and chromatin modifiers face the problem that new histones have to be modified correctly as fast as possible to avoid loss of epigenetic information and eventually misregulation of transcription. The PRC2 complex is able to solve this problem by using the existing marks on old nucleosomes as template for H3K27 methylation of new histones. Since PRC2 is rather ineffective in the absence of stimulation through H3K27me3 binding (Margueron et al, 2009) regions that are completely devoid of this mark will not be methylated. This activity filter protects chromatin from unspecific *de novo* methylation. In contrast, PRC2 has relatively high methylation activity in domains

carrying H3K27me3 marks on inherited nucleosomes. These areas can therefore be fast and efficiently modified, driven by the active feedback loop and without the requirement of additional targeting. Boundary regions between repressive and active chromatin present a problem: pushed forward by the active feedback loop in adjacent repressive chromatin PRC2 could invade active regions and step by step H3K27me3 could spread over the entire genome. Based on our findings the active marks H3K4me3 and H3K36me2/3 present a protective barrier that is able to break the feedback loop when it reaches active chromatin. Old nucleosomes carrying an active mark are directly protected by the allosteric inhibition of PRC2. Since activation by H3K27me3 is assumed to be able to bridge gaps of several nucleosomes without stimulation, newly synthesized, unmodified histones between protected nucleosomes could still be targeted. However, H3K4 and H3K36 methylation marks help to "thin out" stimulating signals by gradually decreasing the local concentration of H3K27me3 modified tails and eventually breaking the feedback loop (**Figure 3**). This mechanism could be supported by additional factors: swift maintenance of H3K4 and H3K36 in active chromatin after replication would further decrease potential substrate nucleosomes for PRC2. Localized activity of H3K27 specific demethylases on the other hand could remove invasive marks and help to retain the border between active and repressive chromatin (Lee et al, 2007).

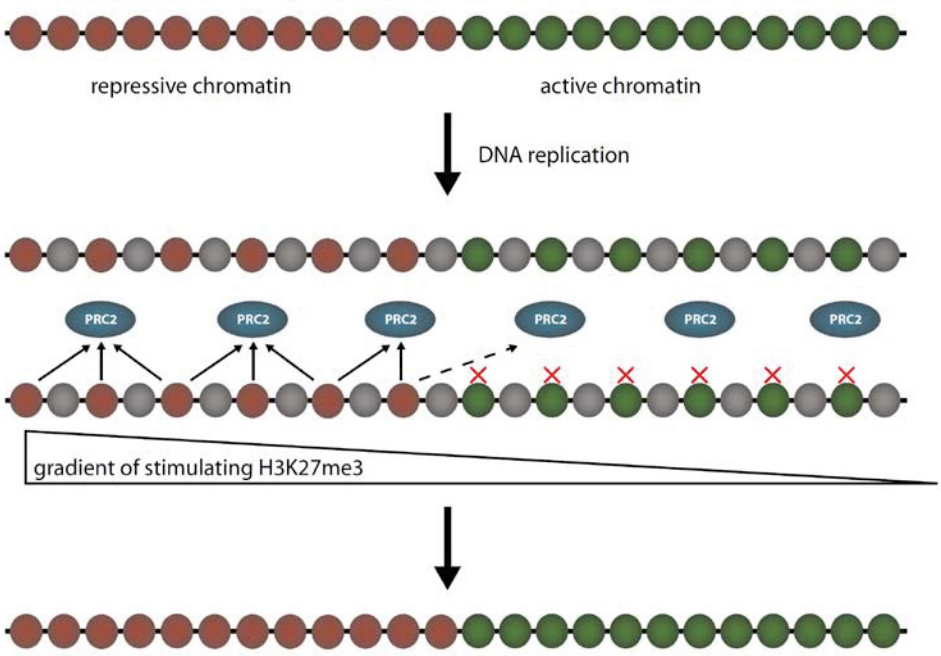


Figure 3: Protection of active chromatin. Nucleosomes carrying repressive chromatin marks are depicted in red, nucleosomes with active chromatin marks in green. Inhibition of PRC2 by H3K4me3 and H3K36me2/3 results in a gradual reduction of stimulating H3K27me3 marks in border regions between active and repressive chromatin.

5.8.3 Long-range and short-range interactions in PRC2 regulation

In epigenetic bistability the distance over which signals can be amplified determines the sensitivity of an enzymatic complex to interspersed unmodified areas or nucleosomes with opposing marks. Regulatory switches that depend on short-range interactions such as Sir protein-mediated heterochromatin formation require a high feedback-to-noise ratio (Dodd et al, 2007). On the other side, these systems are easy to control with very localized boundary elements as effective circuit breakers. PRC2 must be regulated by a mixture of short- and long-range interactions. The importance of long-range signal amplification has been implicated by mathematical modeling of H3K27 dynamics during vernalization in *Arabidopsis* (Angel et al, 2011). Models utilizing only short-range interactions were unsuccessful to stably maintain a silenced state, due to the high noise levels in the system (active demethylation, histone exchange, unmodified nucleosomes). What these long-range interactions are in molecular terms is presently unclear, although interactions facilitated through (i) looping of the DNA, (ii) tight packaging in heterochromatin, or (iii) diffusion and re-association of modifying complexes in proximity to their recruitment site have been discussed (Angel et al, 2011). Importantly, the amplification distance has to be limited in order for boundary elements to be effective. Thus, also polycomb stimulation must decrease gradually with the distance from repressive chromatin, otherwise whole domains with inhibitory marks could be overcome.

5.8.4 Formation of repressive chromatin is amplified by cooperation with demethylases

Recently PRC2 has been found to associate with histone demethylases that remove competing modifications. The H3K4me3 demethylase RBP2 (Jarid1a) has been shown to physically interact with the PRC2 complex and to assist in polycomb mediated repression in ES cells (Pasini et al, 2008). Another demethylase, LSD1, specific for H3K4me2, has been found in complex with PRC2 and the long intergenic noncoding RNA HOTAIR (Tsai et al, 2010a) where the RNA serves as a scaffold connecting the two complexes. This interaction with different demethylases can affect PRC2 regulation in two ways: on one hand it can create more substrates for PRC2 by demethylation of H3K4, thereby enhancing the feedback loop. On the other hand it has been suggested that it serves in further long-range interactions, since PRC2 could be recruited to RBP2 and Jarid1a target sites (Angel et al, 2011). The cooperation also allows dynamic changes by effective silencing of active chromatin. In any case this

combined mechanism has to be limited and tightly regulated to prevent uncontrolled H3K27me3 spreading.

5.8.5 A potential role for Nurf55 in further long-range interactions

In addition to allosteric inhibition mediated by the Su(z)12 VEFS domain and activation *via* EED/Esc further regulatory modules within the PRC2 core complex cannot be ruled out. As demonstrated in this study PRC2 component Nurf55 binds to histone H3 and binding is sensitive to H3K4 methylation. Although our *in vitro* experiments could not determine any functional influence of Nurf55 on PRC2 activity, mutations of Nurf55 in flies were recently reported to show reduced levels of H3K27me3 (Anderson et al, 2011). It is therefore possible that Nurf55 is involved in long range contacts which played no role in our assays using mononucleosomes or 4-mer nucleosomal arrays. Cooperative binding by different subunits of chromatin modifying complexes has been postulated to increase substrate specificity (Lejon et al, 2011) and interaction of a complex with more than one nucleosome simultaneously has been shown to contribute strongly to bistability (Dodd et al, 2007). With the SET domain of E(z), the WD40 propellers of Esc and Nurf55 as well as the VEFS domain of Su(z)12 the PRC2 complex contains four domains that have been reported to interact with histone tails so far. Taken together the combination of short-range and long-range interactions allows PRC2 a precise readout of its chromatin environment.

5.8.6 Allosteric regulation allows fine-tuning of PRC2 activity

We have found that PRC2 inhibition can be overcome by simultaneous H3K27me3 stimulation. PRC2 methylation therefore presents itself not as a simple on/off-switch but as a delicate sensor able to integrate positive and negative signals and to tune its activity accordingly. This is made possible by the allosteric nature of its regulation. Unlike PRC2, other chromatin modifiers are usually rather controlled on the level of recruitment. This kind of regulation relying on binding affinities offers a stable off-switch but does not allow incorporation of multiple inputs. This makes sense for enzymes with high activity that processively modify their substrate. PRC2 in contrast is considered to have low efficiency; it does not move along the DNA to methylate successively neighboring nucleosomes. Thus, for

maintenance of repressive chromatin its recruitment can be rather unspecific while modulation of its activity in contrast provides a high degree of control.

5.8.7 PRC2 inhibition and bivalency - a problem for bistability?

In differentiated tissues H3K27 methylation is for the most part mutually exclusive with active histone marks and their dynamics fit to the model of a typical bistable switch. However, studies in stem cells have found active and repressive histone marks to exist simultaneously in the same chromatin regions, so called bivalent domains that keep genes in a poised state (Bernstein et al, 2006). Currently it is controversial to which extent bivalency exists on the same histone tail (Young et al, 2009) and different mechanisms to establish bivalent domains have been suggested (Schmitges et al, 2011). The above described feedback loops and control mechanisms are optimized for the exclusiveness of either active or repressive marks. Changes to this system necessary for bivalent marks could reduce the effectiveness of the feedback loops or on the other hand could result in bivalency all over the genome. Thus, it must be assumed that establishment and safe propagation of bivalent domains demands additional regulatory mechanisms in order to sustain pluripotency.

5.9 Diversity of histone modifications and combinatorial readout

When covalent modifications of histone tails were found to have an effect on gene expression a simple histone code was proposed that associated each modification with either gene activation or repression (Jenuwein & Allis, 2001). While tissue-specific gene expression could be governed this way by a small number of modifications this simple model of epigenetic regulation had to be substantially revised: new findings over the past decade have shown that in addition to controlling tissue-specific gene expression histone modifications are involved in cellular processes as varied as DNA replication, stem cell maintenance, DNA repair, X-chromosome inactivation, epigenetic inheritance, development, apoptosis, and even memory storage. To cope with this wealth of functions writing and reading of histone modifications have to be much more sophisticated. Although some marks are typically found either in active chromatin or repressive chromatin our current understanding is that histone modifications can have very different effects depending on the context, i.e. on the simultaneous presence or absence of other histone marks. An updated, "combinatorial code"

has been suggested which links the transcriptional state to different combinations of histone modifications (Fischer et al, 2008). However, this model still views histone modifications as simple on/off switches and does not explain the necessity for a multitude of interacting modifications.

Histone modifications can either have a direct effect on chromatin structure (e.g. by neutralizing or adding a charge on histone tails) or affect the chromatin state by recruiting effector complexes including chromatin remodelers, chromatin modifiers and components of the DNA repair machinery (Suganuma & Workman, 2011). The latter requires recognition of histone tails and their modifications by specialized protein domains such as bromodomains, chromodomains, PHD domains, tudor domains, ankyrin repeats, MBT domains, WD40 repeats, BRCT domains and 14-3-3 domains (a list of histone modifications recognized by these domains is provided in Suganuma & Workman, 2011). The effector complexes often consist of a sizeable number of subunits and can harbor multiples of these histone reader domains which allows a combinatorial readout of different modifications. Thereby one complex can integrate the information from different inputs.

Over the past decade more and more cases of crosstalk between histone modifications have been described (Bannister & Kouzarides, 2011; Fischle et al, 2003a; Suganuma & Workman, 2011; Wang & Patel, 2011). In principle, crosstalk can be divided into two categories: (1) in some cases one histone modification depends directly on the presence or absence of another one when the complex catalyzing the reaction contains a histone reader domain. (2) The downstream signal of a histone modification can be affected by crosstalk when the effector complex recognizes two or more modified histone tails.

5.9.1 Influence of multiple modifications on complex recruitment

Recruitment of effector complexes to specific histone modifications is one of the fundamental mechanisms of epigenetic regulation. Many chromatin marks are, however, widespread and are often individually recognized by a large number of chromatin readers (e.g. more than 10 proteins are known to bind H3K4me3, most of them stimulating trancription but some inhibiting it (Ruthenburg et al, 2007b)). The distinct tissue- and time-dependent localization patterns of chromatin readers on the genome cannot be explained by a simple model where one complex binds only one histone mark. By increasing or decreasing the affinity of chromatin readers for certain combinations of histone marks (relative to the affinity for the

individual marks) combinatorial readout ensures that effector complexes are preferentially recruited to a small subset of very specific target sites in a time- and tissue-specific context.

The first case of combinatorial readout of histone modifications was reported for the double bromodomain of TAF1, a component of the TFIID complex (Jacobson et al, 2000). TAF1 has a preference for diacetylated histone H4, its affinity for peptides with multiple acetylated sites is 7-27 fold higher than for peptides with only one acetylated lysine. Thereby, TAF1 recruits the TFIID complex specifically to hyperacetylated genes. Another example for a protein recruited via two histone reader modules is TRIM24 containing a PHD domain specific for unmodified histone H3K4 and a bromodomain binding acetylated lysines (Tsai et al, 2010b). Similar to TAF1 the co-operative binding of the PHD domain and the bromodomain drastically increases the affinity for sites with the preferred combination of histone marks. Recruitment can also be negatively influenced by chromatin marks. HP1 is targeted by binding of its chromodomain to H3K9me3. Binding of the chromodomain and thereby recruitment of HP1 is inhibited by phosphorylation of neighboring H3S10 (Bannister et al, 2001; Fischle et al, 2003b; Lachner et al, 2001).

5.9.2 Functional regulation by chromatin readers

While some systems are based on a simple recruitment model, combinatorial readout must have a more complex or even different function in other situations. BPTF, a protein involved in ATP-dependent chromatin remodeling, contains a PHD domain binding H3K4me3 and a bromodomain binding acetylated H4 (Li et al, 2006; Wysocka et al, 2006). While the bromodomain binds to H4 peptides carrying 1 of 3 acetyl-lysine residues with comparable affinity, at the nucleosomal level binding of the PHD domain to H3K4me3 restricts specificity of the bromodomain to H4K16ac resulting in increased binding affinity (Rando, 2012; Ruthenburg et al, 2011). While the exact mechanism is not fully understood it has been suggested that H3K4me3 binding restrains the possible binding orientation of H4 peptides favoring H4K16 for the bromodomain. For other complexes the binding to histone marks does not affect localization on the genome but instead steps distal to recruitment. The histone deacetylase complex Rpd3S has been shown to bind specifically to H3K36me3 via its Eaf3 subunit. Surprisingly, localization of Rpd3S is, however, unaffected by the deletion of the chromodomain in Eaf3. Instead, recruitment of Rpd3S to active genes is mediated by RNAPII. The interaction of the chromodomain of Eaf3 and H3K36me3 plays a functional

role, it is essential for the activation of Rpd3S activity by H3K36me3 (i.e. distal to recruitment) (Drouin et al, 2010).

The PRC2 complex takes a special position among regulators of chromatin modifications. Its enzymatic activity is regulated allosterically by at least two separate subunits that recognize the histone mark deposited by PRC2 itself (H3K27me3) as well as marks typically found in active chromatin (H3K4me3 and H3K36me2/3). In addition, we have identified Nurf55 as a third histone reader domain within PRC2, however without allosteric effect on histone methylation. With regard to the dynamics of methylation on lysines 4, 27 and 36 of histone H3 it is not surprising to find a very complex regulatory system in PRC2. These modifications safeguard the expression patterns of some of the most important regulators of development and have to undergo a very tightly-timed change in early differentiation. The mechanisms discussed above ensure three basic requirements: (1) the maintenance of a repressive chromatin state within stable boundaries, (2) protection of active chromatin from PRC2 activity, and (3) rapid spreading of repressive chromatin once a change has been initiated by external stimuli. Each function individually could be controlled by a single chromatin reader but the necessary versatility of PCR2 would be impossible to attain without the integration of positive and negative input signals. The special role of PRC2 in epigenetic regulation is also demonstrated by the fact that it is to date the only complex known to methylate histone H3K27. In contrast, most other histone marks can be deposited by several complexes (H3K4, e.g., is methylated by at least 9 different enzymes (Kouzarides, 2007)). However, it would not be surprising to find that complexes working in concert with PRC2 underlie a similarly complex regulation. The MLL complex, an important counterpart of the polycomb system, is already known to bind its own product, H3K4me3, via its PHD3 domain and this binding is required for the transcriptional maintenance functions of MLL (Chang et al, 2010). Additional regulation of MLL comes from binding to another protein, CyP33, which by triggering a cis/trans proline isomerization in MLL exposes its own binding site. MLL also contains a bromodomain. Its function is unclear so far, but MLL might thereby integrate the input of other chromatin marks.

5.9.3 The advantages of multivalent binding

Cooperative binding of several domains or subunits to different histone marks first of all dramatically increases the affinity of a complex for its target site. The increase is higher than

the simple sum of the individual affinities due to a reduced loss of entropy when the binding complex is already pre-assembled (Krishnamurthy et al, 2006; Ruthenburg et al, 2007a). In addition, the presence of multiple interaction sites increases the local concentration of binding partners and thereby causes an increase of the rate of association (a principle also underlying the annealing of DNA strands) (Ptashne & Gann, 1997).

While a high-affinity interaction could also be provided by a single binding site the advantage of having multiple binding sites lies in the increase of regulatory potential (Ruthenburg et al, 2007b). Deposition and removal of multiple marks requires different complexes each of which can add another layer of regulation. Competition for individual binding sites is a superior way to modulate effector complex binding dynamically as it can be triggered faster than a change of chromatin marks. Finally, the positioning of the modifications and the spacing between them adds another variable that can increase specificity. In summary, multivalent binding provides great specificity and high affinity coupled with a susceptibility to competition.

5.9.4 The histone code - more than just an on/off switch

Already a small number of histone modifications would be sufficient to store the information about the transcriptional state of a chromatin region, therefore the vast number of potential histone modification sites and the use of combinatorial readout seems to be unnecessary for the sole purpose of transcriptional regulation. However, histone modifications serve more than to store simple on/off information. On one hand, integration of multiple signals can function as failsafe when the reader is set up for an all-or-nothing response. One the other hand, when complexes react in a gradual fashion to the input, integration allows the fine-tuning of their output. In systems that act as bistable switches additional input signals can help to stabilize the current state or allow an even faster transition to the other state (Angel et al, 2011).

Importantly, combinatorial readout increases the amount of information that can be stored on a limited number of histone tail residues exponentially: while a single modification can either be present or absent there are four possible outcomes when it is combined with the information of a second modification (present/present, absent/absent, present/absent, absent/present). This enables chromatin to contain not only information about the current transcriptional state, but also to store information about past transcriptional events, to prepare

the cell for drastic changes during differentiation (see bivalency) or to allow a differentiated response depending on different external stimuli. Finally, combinatorial readout of histone modifications has even been suggested in memory formation (Wood et al, 2006).

5.10 Conclusion

The balance between pluripotency and differentiation is regulated by multiple bistable switches. The propagation of and crosstalk between active and repressive chromatin marks is a prominent example for an epigenetic switch and a key player in differentiation. It is based on the ability of complexes to recognize existing modifications in their chromatin environment and fine-tune their activity accordingly. The positive feedback of H3K27me3 on PRC2 methyltransferase activity provides a powerful mechanism for propagation and spreading of repressive chromatin. On the other hand our study presents a way to protect active chromatin regions by inhibition of PRC2 via H3K4me3 and H3K36me2/3 marks.

Regulation of modifying complexes via recognition of their own product by a domain distinct from their catalytic domain appears to be a common mechanism. Particularly histone methyltransferases are using this kind of direct feedback loop to propagate their marks. By analogy with PRC2 the H3K9 specific methyltransferases G9a and GLP bind their H3K9me1 and H3K9me2 products *via* ankyrin repeat domains (Collins & Cheng, 2010) and recognition of H3K9 methylation *via* a chromodomain is essential for the function of the H3K9 specific methyltransferase Clr4 (Zhang et al, 2008). The MLL complex recognizes its H3K4me3 mark through its PHD3 domain, which has been shown to be essential for MLL function *in vivo* (Chang et al, 2010). On the other hand methyltransferase complexes have also been found to respond to opposing marks. The active H3K4me3 mark inhibits not only PRC2 but also the H3K9 methyltransferase SETDB1 (Binda et al, 2010). The PRC2 complex might stand out for its allosteric regulation, but histone methylation in general appears to be driven by positive feedback loops formed by modifiers binding their own product. While this mechanism provides an effective way for propagation of one mark, we postulate that in order to maintain domains of active and repressive chromatin all complexes need a switch that is able to break the positive feedback loop.

The interaction between H3K27me3, H3K4me3 and H3K36me2/3 is just one example for crosstalk between different histone modifications. These modifications are involved in a multitude of different processes, some are even associated with activities causing opposite

effects, so only a combinatorial readout by the effectors can cope with the complexity of the systems. Multivalent binding can result in significant affinity enhancement, additional specificity and a much more flexible, i.e. dynamic interaction. The combinatorial readout functions as failsafe in all-or-nothing responses and allows, on the other hand, the fine-tuning of systems that react in a gradual fashion. In addition, the amount of information that can be stored increases exponentially with the number of binding partners.

5.11 Outlook

Identification of an amazing multitude of histone modifications and the unraveling of their complex functional interactions has been achieved by the combination of biochemical experiments and elucidation of the structures of the involved proteins. The next step is to verify that the observed mechanisms also work *in vivo*.

Beyond the verification of the *in vitro* results *in vivo* the door to new research areas has been opened by the work on histone modifications. There is a close relationship between histone modifications, particularly histone methylation and DNA methylation. It will be important to examine how the "silencing arm" of the histone modification system and DNA methylation are integrated and whether there is a "division of labor" between the two systems.

Another promising area of future research in the field of histone modifications are the dynamics of these modifications. Originally it was thought that most of the histone modifications are static, even irreversible. But in the meantime for virtually all modifications enzymes have been identified that can reverse the modifications and it was found that the turnover rate varies depending on the histone, the residue modified, the kind of modification and the neighboring modifications. It will be of utmost interest to elucidate how the histone modification pattern changes, e.g. during the cell cycle, in differentiation/development and in response to changing environmental conditions.

A defective histone modification machinery can lead to various diseases, e.g. cancer and autoimmune, cardiovascular and neurological diseases (epimutation-induced diseases). It is safe to predict that further unraveling of the histone interactome will lead to the identification of additional diseases caused by the derailment of the histone modification system. As a consequence, interference with the histone modification systems offers therapeutic opportunities for a wide spectrum of disease.

References

Anderson AE, Karandikar UC, Pepple KL, Chen Z, Bergmann A, Mardon G (2011) The enhancer of trithorax and polycomb gene Caf1/p55 is essential for cell survival and patterning in Drosophila development. *Development* **138:** 1957-1966

Angel A, Song J, Dean C, Howard M (2011) A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476:** 105-108

Annunziato AT (2005) Split decision: what happens to nucleosomes during DNA replication? *The Journal of biological chemistry* **280:** 12065-12068

Bagci EZ, Vodovotz Y, Billiar TR, Ermentrout GB, Bahar I (2006) Bistability in apoptosis: roles of bax, bcl-2, and mitochondrial permeability transition pores. *Biophys J* **90:** 1546-1559

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

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120-124

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

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143:** 470-484

Bedford MT, Clarke SG (2009) Protein arginine methylation in mammals: who, what, and why. *Mol Cell* **33:** 1-13

Beisel C, Imhof A, Greene J, Kremmer E, Sauer F (2002) Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. *Nature* **419:** 857-862

Bell O, Conrad T, Kind J, Wirbelauer C, Akhtar A, Schubeler D (2008) Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. *Mol Cell Biol* **28:** 3401-3409

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125:** 315-326

Bienz M (2006) The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci* **31:** 35-40

Binda O, LeRoy G, Bua DJ, Garcia BA, Gozani O, Richard S (2010) Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9: regulation of lysine methyltransferases by physical interaction with their substrates. *Epigenetics* **5:** 767-775

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

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

Boehning D, Patterson RL, Sedaghat L, Glebova NO, Kurosaki T, Snyder SH (2003) Cytochrome c binds to inositol (1,4,5) trisphosphate receptors, amplifying calcium-dependent apoptosis. *Nature cell biology* **5:** 1051-1061

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122:** 947-956

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441:** 349-353

Bracken AP, Pasini D, Capra M, Prosperini E, Colli E, Helin K (2003) EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. *The EMBO journal* **22:** 5323-5335

Briggs SD, Xiao T, Sun ZW, Caldwell JA, Shabanowitz J, Hunt DF, Allis CD, Strahl BD (2002) Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* **418:** 498

Brown JL, Fritsch C, Mueller J, Kassis JA (2003) The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development* **130:** 285-294

Brown MT, Cooper JA (1996) Regulation, substrates and functions of src. *Biochim Biophys Acta* **1287:** 121-149

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

Caretti G, Di Padova M, Micales B, Lyons GE, Sartorelli V (2004) The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev* **18:** 2627-2638

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, Workman JL (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123:** 581-592

Catena R, Tiveron C, Ronchi A, Porta S, Ferri A, Tatangelo L, Cavallaro M, Favaro R, Ottolenghi S, Reinbold R, Scholer H, Nicolis SK (2004) Conserved POU binding DNA sites in the Sox2 upstream enhancer regulate gene expression in embryonic and neural stem cells. *The Journal of biological chemistry* **279:** 41846-41857

Chamberlain SJ, Yee D, Magnuson T (2008) Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* **26:** 1496-1505

Chang B, Chen Y, Zhao Y, Bruick RK (2007) JMJD6 is a histone arginine demethylase. *Science* **318:** 444-447

Chang PY, Hom RA, Musselman CA, Zhu L, Kuo A, Gozani O, Kutateladze TG, Cleary ML (2010) Binding of the MLL PHD3 Finger to Histone H3K4me3 Is Required for MLL-Dependent Gene Transcription. *J Mol Biol*

Chopra VS, Hong JW, Levine M (2009) Regulation of Hox gene activity by transcriptional elongation in Drosophila. *Current biology : CB* **19:** 688-693

Collins R, Cheng X (2010) A case study in cross-talk: the histone lysine methyltransferases G9a and GLP. *Nucleic Acids Res* **38:** 3503-3511

Couture JF, Collazo E, Trievel RC (2006) Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat Struct Mol Biol* **13:** 698-703

Cui J, Chen C, Lu H, Sun T, Shen P (2008) Two independent positive feedbacks and bistability in the Bcl-2 apoptotic switch. *PloS one* **3:** e1469

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

Czermin B, Schotta G, Hulsmann BB, Brehm A, Becker PB, Reuter G, Imhof A (2001) Physical and functional association of SU(VAR)3-9 and HDAC1 in Drosophila. *EMBO Rep* **2:** 915-919

Dawson MA, Bannister AJ, Gottgens B, Foster SD, Bartke T, Green AR, Kouzarides T (2009) JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature* **461:** 819-822

Dejardin J, Rappailles A, Cuvier O, Grimaud C, Decoville M, Locker D, Cavalli G (2005) Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. *Nature* **434:** 533-538

Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, Askarian-Amiri ME, Ru K, Solda G, Simons C, Sunkin SM, Crowe ML, Grimmond SM, Perkins AC, Mattick JS (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* **18:** 1433-1445

Dodd IB, Micheelsen MA, Sneppen K, Thon G (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* **129:** 813-822

Drouin S, Laramee L, Jacques PE, Forest A, Bergeron M, Robert F (2010) DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. *PLoS Genet* **6:** e1001173

Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137:** 485-497

Endoh M, Endo TA, Endoh T, Fujimura Y, Ohara O, Toyoda T, Otte AP, Okano M, Brockdorff N, Vidal M, Koseki H (2008) Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135:** 1513-1524

Eskeland R, Leeb M, Grimes GR, Kress C, Boyle S, Sproul D, Gilbert N, Fan Y, Skoultchi AI, Wutz A, Bickmore WA (2010) Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol Cell* **38:** 452-464

Faust C, Schumacher A, Holdener B, Magnuson T (1995) The eed mutation disrupts anterior mesoderm production in mice. *Development* **121:** 273-285

Ferrell JE, Jr. (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol* **14:** 140-148

Fischer JJ, Toedling J, Krueger T, Schueler M, Huber W, Sperling S (2008) Combinatorial effects of four histone modifications in transcription and differentiation. *Genomics* **91:** 41-51

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438:** 1116-1122

Fischle W, Wang Y, Allis CD (2003a) Binary switches and modification cassettes in histone biology and beyond. *Nature* **425:** 475-479

Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S (2003b) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* **17:** 1870-1881

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

Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW (2009) Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137:** 472-484

Fox CA, Ehrenhofer-Murray AE, Loo S, Rine J (1997) The origin recognition complex, SIR1, and the S phase requirement for silencing. *Science* **276:** 1547-1551

Francis NJ, Follmer NE, Simon MD, Aghia G, Butler JD (2009) Polycomb proteins remain bound to chromatin and DNA during DNA replication in vitro. *Cell* **137:** 110-122

Fuks F, Hurd PJ, Deplus R, Kouzarides T (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31:** 2305-2312

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

Gehring M, Reik W, Henikoff S (2009) DNA demethylation by DNA repair. *Trends in genetics : TIG* **25:** 82-90

Geng F, Tansey WP (2008) Polyubiquitylation of histone H2B. *Mol Biol Cell* **19:** 3616-3624

Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74:** 481-514

Green DR, Reed JC (1998) Mitochondria and apoptosis. *Science* **281:** 1309-1312

Grimaud C, Negre N, Cavalli G (2006) From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. *Chromosome Res* **14:** 363-375

Grunstein M (1998) Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell* **93:** 325-328

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130:** 77-88

Han Z, Guo L, Wang H, Shen Y, Deng XW, Chai J (2006) Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD-40 protein WDR5. *Mol Cell* **22:** 137-144

Han Z, Xing X, Hu M, Zhang Y, Liu P, Chai J (2007) Structural basis of EZH2 recognition by EED. *Structure* **15:** 1306-1315

Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K (2008) A model for transmission of the H3K27me3 epigenetic mark. *Nature cell biology* **10:** 1291-1300

Hassig CA, Schreiber SL (1997) Nuclear histone acetylases and deacetylases and transcriptional regulation: HATs off to HDACs. *Curr Opin Chem Biol* **1:** 300-308

Hayashi T, Fujita Y, Iwasaki O, Adachi Y, Takahashi K, Yanagida M (2004) Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* **118:** 715-729

Hennig L, Taranto P, Walser M, Schonrock N, Gruissem W (2003) Arabidopsis MSI1 is required for epigenetic maintenance of reproductive development. *Development* **130:** 2555-2565

Ho KL, Harrington HA (2010) Bistability in apoptosis by receptor clustering. *PLoS Comput Biol* **6:** e1000956

Hon GC, Hawkins RD, Ren B (2009) Predictive chromatin signatures in the mammalian genome. *Hum Mol Genet* **18:** R195-201

Hoppe GJ, Tanny JC, Rudner AD, Gerber SA, Danaie S, Gygi SP, Moazed D (2002) Steps in assembly of silent chromatin in yeast: Sir3-independent binding of a Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation. *Mol Cell Biol* **22:** 4167-4180

Ingolia N (2005) Cell cycle: bistability is needed for robust cycling. *Current biology : CB* **15:** R961-963

Jacobs SA, Fischle W, Khorasanizadeh S (2004) Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. *Methods Enzymol* **376:** 131-148

Jacobson RH, Ladurner AG, King DS, Tjian R (2000) Structure and function of a human TAFII250 double bromodomain module. *Science* **288:** 1422-1425

Jenuwein T, Allis CD (2001) Translating the histone code. *Science* **293:** 1074-1080

Jiang D, Wang Y, Wang Y, He Y (2008) Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components. *PloS one* **3:** e3404

Joo HY, Zhai L, Yang C, Nie S, Erdjument-Bromage H, Tempst P, Chang C, Wang H (2007) Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* **449:** 1068-1072

Jurkowska RZ, Jurkowski TP, Jeltsch A (2011) Structure and function of mammalian DNA methyltransferases. *Chembiochem* **12:** 206-222

Kalantry S, Magnuson T (2006) The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. *PLoS Genet* **2:** e66

Ketel CS, Andersen EF, Vargas ML, Suh J, Strome S, Simon JA (2005) Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol Cell Biol* **25:** 6857-6868

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

Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, Green RD, Kouzarides T (2009) Distinct transcriptional outputs associated with mono- and dimethylated histone H3 arginine 2. *Nat Struct Mol Biol* **16:** 449-451

Koerner MV, Pauler FM, Huang R, Barlow DP (2009) The function of non-coding RNAs in genomic imprinting. *Development* **136:** 1771-1783

Kornberg RD (1974) Chromatin structure: a repeating unit of histones and DNA. *Science* **184:** 868-871

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

Krishnamurthy VM, Estroff LM, Whitesides GM (2006) *A comprehensive review of the thermodynamics of multivalency.* in *Fragment-based Approaches in Drug Discovery* (eds Jahnke W, Erlanson DA) 11-53 (Wiley-VCH, Weinheim, 2006)

Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A (2003a) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11:** 721-729

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003b) Methylation of histone

H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **23:** 4207-4218

Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander ES, Koseki H, Bernstein BE (2008) Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4:** e1000242

Kuroda T, Tada M, Kubota H, Kimura H, Hatano SY, Suemori H, Nakatsuji N, Tada T (2005) Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol* **25:** 2475-2485

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

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

Lee JS, Smith E, Shilatifard A (2010) The language of histone crosstalk. *Cell* **142:** 682-685

Lee MG, Villa R, Trojer P, Norman J, Yan KP, Reinberg D, Di Croce L, Shiekhattar R (2007) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318:** 447-450

Lee RH, Heckman CJ (1998) Bistability in spinal motoneurons in vivo: systematic variations in rhythmic firing patterns. *J Neurophysiol* **80:** 572-582

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125:** 301-313

Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A (2010) Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev* **24:** 265-276

Leeb M, Wutz A (2007) Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *The Journal of cell biology* **178:** 219-229

Lejon S, Thong SY, Murthy A, AlQarni S, Murzina NV, Blobel GA, Laue ED, Mackay JP (2011) Insights into association of the NuRD complex with FOG-1 from the crystal structure of an RbAp48.FOG-1 complex. *The Journal of biological chemistry* **286:** 1196-1203

Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* **128:** 707-719

Li B, Jackson J, Simon MD, Fleharty B, Gogol M, Seidel C, Workman JL, Shilatifard A (2009) Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *The Journal of biological chemistry* **284:** 7970-7976

Li BZ, Huang Z, Cui QY, Song XH, Du L, Jeltsch A, Chen P, Li G, Li E, Xu GL (2011) Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. *Cell Res* **21:** 1172-1181

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

Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442:** 91-95

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389:** 251-260

Lund AH, van Lohuizen M (2004) Polycomb complexes and silencing mechanisms. *Curr Opin Cell Biol* **16:** 239-246

Lyon MF (1999) X-chromosome inactivation. *Current biology : CB* **9:** R235-237

Macdonald N, Welburn JP, Noble ME, Nguyen A, Yaffe MB, Clynes D, Moggs JG, Orphanides G, Thomson S, Edmunds JW, Clayton AL, Endicott JA, Mahadevan LC (2005) Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Mol Cell* **20:** 199-211

Maison C, Almouzni G (2004) HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol* **5:** 296-304

Malashchenko T, Shilnikov A, Cymbalyuk G (2011) Six types of multistability in a neuronal model based on slow calcium current. *PloS one* **6:** e21782

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, Pirrotta V, Reinberg D, Gamblin SJ (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461:** 762-767

Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, Dynlacht BD, Reinberg D (2008) Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* **32:** 503-518

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

Marmorstein R (2003) Structure of SET domain proteins: a new twist on histone methylation. *Trends Biochem Sci* **28:** 59-62

Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F, Ponting CP (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci* **28:** 69-74

McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453:** 812-816

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10:** 155-159

Metzger E, Imhof A, Patel D, Kahl P, Hoffmeyer K, Friedrichs N, Muller JM, Greschik H, Kirfel J, Ji S, Kunowska N, Beisenherz-Huss C, Gunther T, Buettner R, Schule R (2010) Phosphorylation of histone H3T6 by PKCbeta(I) controls demethylation at histone H3K4. *Nature* **464:** 792-796

Micheelsen MA, Mitarai N, Sneppen K, Dodd IB (2010) Theory for the stability and regulation of epigenetic landscapes. *Phys Biol* **7:** 026010

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X,

Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553-560

Min J, Feng Q, Li Z, Zhang Y, Xu RM (2003a) Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* **112:** 711-723

Min J, Zhang Y, Xu RM (2003b) Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev* **17:** 1823-1828

Mohd-Sarip A, van der Knaap JA, Wyman C, Kanaar R, Schedl P, Verrijzer CP (2006) Architecture of a polycomb nucleoprotein complex. *Mol Cell* **24:** 91-100

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

Montgomery ND, Yee D, Chen A, Kalantry S, Chamberlain SJ, Otte AP, Magnuson T (2005) The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol* **15:** 942-947

Muller H, Bracken AP, Vernell R, Moroni MC, Christians F, Grassilli E, Prosperini E, Vigo E, Oliner JD, Helin K (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev* **15:** 267-285

Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111:** 197-208

Muratani M, Tansey WP (2003) How the ubiquitin-proteasome system controls transcription. *Nat Rev Mol Cell Biol* **4:** 192-201

Murzina NV, Pei XY, Zhang W, Sparkes M, Vicente-Garcia J, Pratap JV, McLaughlin SH, Ben-Shahar TR, Verreault A, Luisi BF, Laue ED (2008) Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. *Structure* **16:** 1077-1085

Nekrasov M, Klymenko T, Fraterman S, Papp B, Oktaba K, Kocher T, Cohen A, Stunnenberg HG, Wilm M, Müller J (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J* **26:** 4078-4088

Nekrasov M, Wild B, Müller J (2005) Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep* **6:** 348-353

Ng HH, Robert F, Young RA, Struhl K (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11:** 709-719

Ng HH, Xu RM, Zhang Y, Struhl K (2002) Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *The Journal of biological chemistry* **277:** 34655-34657

Ng J, Hart CM, Morgan K, Simon JA (2000) A Drosophila ESC-E(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC. *Mol Cell Biol* **20:** 3069-3078

Nowak AJ, Alfieri C, Stirnimann CU, Rybin V, Baudin F, Ly-Hartig N, Lindner D, Muller CW (2011) Chromatin-modifying complex component Nurf55/p55 associates with histones H3, H4 and Polycomb Repressive Complex 2 subunit Su(z)12 through partially overlapping binding sites. *The Journal of biological chemistry*

O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T (2001) The polycomb-group gene Ezh2 is required for early mouse development. *Mol Cell Biol* **21:** 4330-4336

Okumura-Nakanishi S, Saito M, Niwa H, Ishikawa F (2005) Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *The Journal of biological chemistry* **280:** 5307-5317

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448:** 714-717

Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* **32:** 232-246

Parthun MR, Widom J, Gottschling DE (1996) The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* **87:** 85-94

Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K (2007) The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* **27:** 3769-3779

Pasini D, Cloos PA, Walfridsson J, Olsson L, Bukowski JP, Johansen JV, Bak M, Tommerup N, Rappsilber J, Helin K (2010) JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* **464:** 306-310

Pasini D, Hansen KH, Christensen J, Agger K, Cloos PA, Helin K (2008) Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev* **22:** 1345-1355

Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley GQ, Liu XS, Ecker JR, Milos PM, Agarwal S, Rao A (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473:** 394-397

Patel A, Dharmarajan V, Cosgrove MS (2008) Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *The Journal of biological chemistry* **283:** 32158-32161

Pellicena P, Stowell KR, Miller WT (1998) Enhanced phosphorylation of Src family kinase substrates containing SH2 domain binding sites. *The Journal of biological chemistry* **273:** 15325-15328

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, Sibilia M, Jenuwein T (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107:** 323-337

Pidoux AL, Choi ES, Abbott JK, Liu X, Kagansky A, Castillo AG, Hamilton GL, Richardson W, Rappsilber J, He X, Allshire RC (2009) Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. *Mol Cell* **33:** 299-311

Pien S, Grossniklaus U (2007) Polycomb group and trithorax group proteins in Arabidopsis. *Biochim Biophys Acta* **1769:** 375-382

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122:** 517-527

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

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

Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K, Brykczynska U, Kolb C, Otte AP, Koseki H, Orkin SH, van Lohuizen M, Peters AH (2008) PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* **40:** 411-420

Qian YW, Wang YC, Hollingsworth RE, Jr., Jones D, Ling N, Lee EY (1993) A retinoblastoma-binding protein related to a negative regulator of Ras in yeast. *Nature* **364:** 648-652

Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu RM (2011) The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *The Journal of biological chemistry* **286:** 8361-8368

Rando OJ (2012) Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Curr Opin Genet Dev* **22:** 148-155

Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2:** 21-32

Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **38:** 413-443

Ringrose L, Paro R (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134:** 223-232

Ringrose L, Rehmsmeier M, Dura JM, Paro R (2003) Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. *Dev Cell* **5:** 759-771

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129:** 1311-1323

Robertson KD, Wolffe AP (2000) DNA methylation in health and disease. *Nat Rev Genet* **1:** 11-19

Robinson PJ, Fairall L, Huynh VA, Rhodes D (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci U S A* **103:** 6506-6511

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P (2005) Transcriptional regulation of nanog by OCT4 and SOX2. *The Journal of biological chemistry* **280:** 24731-24737

Roh TY, Cuddapah S, Cui K, Zhao K (2006) The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci U S A* **103:** 15782-15787

Rusche LN, Kirchmaier AL, Rine J (2002) Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. *Mol Biol Cell* **13:** 2207-2222

Ruthenburg AJ, Allis CD, Wysocka J (2007a) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* **25:** 15-30

Ruthenburg AJ, Li H, Milne TA, Dewell S, McGinty RK, Yuen M, Ueberheide B, Dou Y, Muir TW, Patel DJ, Allis CD (2011) Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145:** 692-706

Ruthenburg AJ, Li H, Patel DJ, Allis CD (2007b) Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8:** 983-994

Ruthenburg AJ, Wang W, Graybosch DM, Li H, Allis CD, Patel DJ, Verdine GL (2006) Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. *Nat Struct Mol Biol* **13:** 704-712

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

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407-411

Santos F, Dean W (2004) Epigenetic reprogramming during early development in mammals. *Reproduction* **127:** 643-651

Saunders A, Core LJ, Lis JT (2006) Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* **7:** 557-567

Schmitges FW, Prusty AB, Faty M, Stutzer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, Bunker RD, Wirth U, Bouwmeester T, Bauer A, Ly-Hartig N, Zhao K, Chan H, Gu J, Gut H, Fischle W, Muller J, Thoma NH (2011) Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Mol Cell* **42:** 330-341

Schmitz RJ, Amasino RM (2007) Vernalization: a model for investigating epigenetics and eukaryotic gene regulation in plants. *Biochim Biophys Acta* **1769:** 269-275

Schoeftner S, Sengupta AK, Kubicek S, Mechtler K, Spahn L, Koseki H, Jenuwein T, Wutz A (2006) Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J* **25:** 3110-3122

Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, Goodrich J (2006) Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J* **25:** 4638-4649

Schuetz A, Allali-Hassani A, Martin F, Loppnau P, Vedadi M, Bochkarev A, Plotnikov AN, Arrowsmith CH, Min J (2006) Structural basis for molecular recognition and presentation of histone H3 by WDR5. *EMBO J* **25:** 4245-4252

Sedighi M, Sengupta AM (2007) Epigenetic chromatin silencing: bistability and front propagation. *Phys Biol* **4:** 246-255

Sha W, Moore J, Chen K, Lassaletta AD, Yi CS, Tyson JJ, Sible JC (2003) Hysteresis drives cell-cycle transitions in Xenopus laevis egg extracts. *Proc Natl Acad Sci U S A* **100:** 975-980

Shahbazian MD, Zhang K, Grunstein M (2005) Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. *Mol Cell* **19:** 271-277

Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W, Kingston RE (1999) Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98:** 37-46

Shen X, Liu Y, Hsu YJ, Fujiwara Y, Kim J, Mao X, Yuan GC, Orkin SH (2008) EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* **32:** 491-502

Shi L, Sun L, Li Q, Liang J, Yu W, Yi X, Yang X, Li Y, Han X, Zhang Y, Xuan C, Yao Z, Shang Y (2011) Histone demethylase JMJD2B coordinates H3K4/H3K9 methylation and

promotes hormonally responsive breast carcinogenesis. *Proc Natl Acad Sci U S A* **108:** 7541-7546

Shi Y, Whetstine JR (2007) Dynamic regulation of histone lysine methylation by demethylases. *Mol Cell* **25:** 1-14

Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311:** 844-847

Sims RJ, 3rd, Nishioka K, Reinberg D (2003) Histone lysine methylation: a signature for chromatin function. *Trends in genetics : TIG* **19:** 629-639

Sing A, Pannell D, Karaiskakis A, Sturgeon K, Djabali M, Ellis J, Lipshitz HD, Cordes SP (2009) A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138:** 885-897

Smallwood A, Esteve PO, Pradhan S, Carey M (2007) Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev* **21:** 1169-1178

Smith CL, Peterson CL (2005) ATP-dependent chromatin remodeling. *Curr Top Dev Biol* **65:** 115-148

Sneppen K, Micheelsen MA, Dodd IB (2008) Ultrasensitive gene regulation by positive feedback loops in nucleosome modification. *Mol Syst Biol* **4:** 182

Song JJ, Garlick JD, Kingston RE (2008) Structural basis of histone H4 recognition by p55. *Genes Dev* **22:** 1313-1318

Song JJ, Kingston RE (2008) WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *The Journal of biological chemistry* **283:** 35258-35264

Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nature cell biology* **9:** 1428-1435

Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, Allis CD (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* **22:** 1298-1306

Stucki M, Clapperton JA, Mohammad D, Yaffe MB, Smerdon SJ, Jackson SP (2005) MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* **123:** 1213-1226

Suganuma T, Pattenden SG, Workman JL (2008) Diverse functions of WD40 repeat proteins in histone recognition. *Genes Dev* **22:** 1265-1268

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

Sun ZW, Allis CD (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418:** 104-108

Tanaka Y, Katagiri Z, Kawahashi K, Kioussis D, Kitajima S (2007) Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* **397:** 161-168

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

Taunton J, Hassig CA, Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272:** 408-411

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14:** 1025-1040

Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, Diaz MO, Scacheri PC, Harte PJ (2009) CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* **136:** 3131-3141

Tie F, Furuyama T, Prasad-Sinha J, Jane E, Harte PJ (2001) The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development* **128:** 275-286

Tsai HC, Baylin SB (2011) Cancer epigenetics: linking basic biology to clinical medicine. *Cell Res* **21:** 502-517

Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY (2010a) Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329:** 689-693

Tsai WW, Wang Z, Yiu TT, Akdemir KC, Xia W, Winter S, Tsai CY, Shi X, Schwarzer D, Plunkett W, Aronow B, Gozani O, Fischle W, Hung MC, Patel DJ, Barton MC (2010b) TRIM24 links a non-canonical histone signature to breast cancer. *Nature* **468:** 927-932

Tyler JK, Bulger M, Kamakaka RT, Kobayashi R, Kadonaga JT (1996) The p55 subunit of Drosophila chromatin assembly factor 1 is homologous to a histone deacetylase-associated protein. *Mol Cell Biol* **16:** 6149-6159

van der Stoop P, Boutsma EA, Hulsman D, Noback S, Heimerikx M, Kerkhoven RM, Voncken JW, Wessels LF, van Lohuizen M (2008) Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. *PloS one* **3:** e2235

van Ingen H, van Schaik FM, Wienk H, Ballering J, Rehmann H, Dechesne AC, Kruijzer JA, Liskamp RM, Timmers HT, Boelens R (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure* **16:** 1245-1256

Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131:** 58-69

Veron N, Peters AH (2011) Epigenetics: Tet proteins in the limelight. *Nature* **473:** 293-294

Verreault A, Kaufman PD, Kobayashi R, Stillman B (1996) Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* **87:** 95-104

Voncken JW, Roelen BA, Roefs M, de Vries S, Verhoeven E, Marino S, Deschamps J, van Lohuizen M (2003) Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci U S A* **100:** 2468-2473

Wade PA (2001) Methyl CpG-binding proteins and transcriptional repression. *BioEssays : news and reviews in molecular, cellular and developmental biology* **23:** 1131-1137

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431:** 873-878

Wang Y, Jia S (2009) Degrees make all the difference: the multifunctionality of histone H4 lysine 20 methylation. *Epigenetics* **4:** 273-276

Wang Z, Patel DJ (2011) Combinatorial readout of dual histone modifications by paired chromatin-associated modules. *The Journal of biological chemistry* **286:** 18363-18368

Webby CJ, Wolf A, Gromak N, Dreger M, Kramer H, Kessler B, Nielsen ML, Schmitz C, Butler DS, Yates JR, 3rd, Delahunty CM, Hahn P, Lengeling A, Mann M, Proudfoot NJ, Schofield CJ, Bottger A (2009) Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science* **325:** 90-93

Wilhelm T (2009) The smallest chemical reaction system with bistability. *BMC Syst Biol* **3:** 90

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473:** 343-348

Wolffe AP, Hayes JJ (1999) Chromatin disruption and modification. *Nucleic Acids Res* **27:** 711-720

Wood MA, Hawk JD, Abel T (2006) Combinatorial chromatin modifications and memory storage: a code for memory? *Learn Mem* **13:** 241-244

Workman JL (2006) Nucleosome displacement in transcription. *Genes Dev* **20:** 2009-2017

Workman JL, Kingston RE (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* **67:** 545-579

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y (2011) Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473:** 389-393

Wutz A (2011) Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet* **12:** 542-553

Wyce A, Xiao T, Whelan KA, Kosman C, Walter W, Eick D, Hughes TR, Krogan NJ, Strahl BD, Berger SL (2007) H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Mol Cell* **27:** 275-288

Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121:** 859-872

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442:** 86-90

Xiao T, Kao CF, Krogan NJ, Sun ZW, Greenblatt JF, Osley MA, Strahl BD (2005) Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol Cell Biol* **25:** 637-651

Xu C, Min J (2011) Structure and function of WD40 domain proteins. *Protein Cell* **2:** 202-214

Xue Y, Wong J, Moreno GT, Young MK, Cote J, Wang W (1998) NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell* **2:** 851-861

Yan SJ, Zartman JJ, Zhang M, Scott A, Shvartsman SY, Li WX (2009) Bistability coordinates activation of the EGFR and DPP pathways in Drosophila vein differentiation. *Mol Syst Biol* **5:** 278

Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38:** 662-674

Yeates TO (2002) Structures of SET domain proteins: protein lysine methyltransferases make their mark. *Cell* **111:** 5-7

Young NL, DiMaggio PA, Plazas-Mayorca MD, Baliban RC, Floudas CA, Garcia BA (2009) High throughput characterization of combinatorial histone codes. *Mol Cell Proteomics* **8:** 2266-2284

Yun M, Wu J, Workman JL, Li B (2011) Readers of histone modifications. *Cell Res* **21:** 564-578

Zhang K, Mosch K, Fischle W, Grewal SI (2008) Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* **15:** 381-388

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

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

Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322:** 750-756

Zhao Q, Rank G, Tan YT, Li H, Moritz RL, Simpson RJ, Cerruti L, Curtis DJ, Patel DJ, Allis CD, Cunningham JM, Jane SM (2009) PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat Struct Mol Biol* **16:** 304-311

Acknowledgments

I would like to thank Nicolas Thomä for giving me the opportunity to conduct my PhD studies in his lab and for his great supervision over the past years. He always took his time to answer questions and to discuss the project, I never felt alone with a problem.

I would like to thank the members of my PhD committee, Wolfgang Fischle, Susan Gasser and Antoine Peters. My meetings with them and the feedback for the preparation of the manuscript have been very helpful and encouraging.

I would like to express my thanks to all current and past members of the Thomä lab. Especially, I thank "Raju" Gondichatnahalli Manjappa Lingaraju, Jonathan Aiwazian and Richard Bunker for their work on the PRC2 project. I also want to thank the very first generation of lab members, Mahamadou Faty, Andrea Scrima and Markus Vogel for their help and support over the past years.

I would like to thank our collaborators Archana Prusty, Nga Ly-Hartig and Jürg Müller from the EMBL Heidelberg for establishing the HMTase assays and performing the ITC measurements, and also Alexandra Stützer and Wolfgang Fischle from the MPI Göttingen for the FP measurements. I also would like to thank all members of the PRC2 team at Novartis, Ling Li, Shaolian Zhou, Urs Wirth,Tewis Bouwmeester, Andreas Bauer, Kehao Zhao, Homan Chan, Justin Gu.

I acknowledge all the excellent facilities at the FMI. Especially, I thank Ragna Sack, Daniel Hess and Dominique Klein for the mass spectrometry analysis, as well as Heinz Gut and Norman Kairies for crystallographic support.

I would like to thank Eszter Posfai and my father, Claus Schmitges for helpful comments on my thesis.

I gratefully acknowledge funding from the Schering Stiftung as well as support by the Studienstiftung des deutschen Volkes.

I thank Eszter, my family and friends for supporting me during my PhD studies.