Development and application of experimental tools for studying the distribution and dynamics of chromatin modifications

Von der Fakultät 4: Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Goran Kungulovski

aus Resen, Mazedonien

Hauptberichter: Prof. Dr. Albert Jeltsch

Mitberichter: Prof. Dr. Marianne Rots

Tag der mündlichen Prüfung: 22.06.2015

Institut für Biochemie der Universität Stuttgart

2015

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbst verfasst und dabei keine anderen als die angegeben Quellen und Hilfsmittel verwendet habe.

Stuttgart, 7 April 2015

Goran Kungulovski

Table of Contents

Acknowledgements

I am grateful to my mentor Prof. Dr. Albert Jeltsch for his excellent guidance, intellectual support and invaluable influence on my personal and scientific growth.

I am thankful to my non-scientist friends and family, who tolerated (most of the time) my effusive daily babblings about science. On the same note, I am thankful to my scientist friends who tolerated my daily sermons about life, existence, the human state and the Cosmos.

This thesis is dedicated to all the people I like, adulate and love. You know who you are. Thank you for surviving during my periods of blues and immoderate passions, during the times of ebullient delight and untrammelled vehemence.

List of publications and achievements

List of publications and achievements resulting from the doctoral studies

Kungulovski G, Mauser R, Jeltsch A. (2015) "**Isolation of nucleosomes having multiple-modified co-existing histone protein octamers.**" EP15161621.6. Patent application submitted to EPO on 01.04.2015.

Development of a novel method for affinity purification of nucleosomes carrying two co-existing histone modifications at the same time in a single step.

Kungulovski G, Mauser R, Jeltsch A. (2015) "**Specificity Analysis of Histone Modification Specific Antibodies or Reading Domains on Histone Peptide Arrays",** *Methods In Molecular Biology*, 1348, 275-284 doi: 10.1007/978-1-4939-2999-3_24.

Detailed protocol for application and bioinformatic analysis of peptide arrays for specificity profiling of histone modification antibodies or reading domains.

Kungulovski G, Mauser R, Jeltsch A. (2015) **"Affinity reagents for studying histone modifications and guidelines for their quality control"**, *Epigenomics, (in press).*

Review paper, which pinpoints the advantages and disadvantages or using histone PTM antibodies or reading domains, respectively. It also discusses the drawbacks of the current quality control measures and at the same time upgrades the guidelines for rigorous quality control of affinity reagents used in histone PTM studies.

Kungulovski G, Nunna S, Thomas M, Zanger UM, Reinhardt R, Jeltsch A. (2015) "**Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained"**, *Epigenetics & Chromatin,* 8:12 doi: 10.1186/s13072-015-0002-z.

In this paper, we apply zinc finger targeted DNA and histone methylation of an endogenous locus and show for the first time with time-course experiments that the newly established chromatin state is not maintained. This study provides incredible insights into the innerworkings of chromatin.

Kungulovski G, Henry C, Kycia I, Reinhard R, Tamas R, Labhart P, Jurkowska RZ, Jeltsch A. (2014) "**Application of Histone Modification Specific Interaction Domains as an Alternative to Antibodies**", *Genome Research,* 24, 1842–1853 doi: 10.1101/gr.170985.113.

In this study we developed and applied histone modification interacting domains, as a novel, cheap and reproducible alternative to antibodies. We anticipate that this groundbreaking work will have an immense influence and further the development and accuracy of chromatin

biology.

Kycia I, Kudithipudi S, Tamas R, **Kungulovski G,** Dhayalan A, Jeltsch A, (2014). "**The Tudor domain of the PHD finger protein 1 is a dual reader of lysine trimethylation at lysine 36 of histone H3 and lysine 27 of histone variant H3t**", *Journal of Molecular Biology*, 17;426(8):1651-60 doi: 10.1016/j.jmb.2013.08.009.

In this paper we report that the PHD domain of PHF1 (a component of PRC2) binds nucleosomes trimethylated at lysine 36 of canonical histone H3 or trimethylated lysine 27 of histone variant H3t.

List of publications and achievements resulting from work prior to the doctoral studies

Bock I, Kudithipudi S, Tamas R, **Kungulovski G**, Dhayalan A, Jeltsch A. (2011) "**Application of Celluspots peptide arrays for the analysis of the binding specificity of epigenetic reading domains to modified histone tails**", *BMC Biochemistry*, 12:48 doi: 10.1186/1471-2091-12-48.

In this study we show the applicative and combinatorial power of CelluSpots Peptide arrays in specificity profiling analyses of recombinant histone modifications interacting domains.

Zusammenfassung

Alle Zellen eines Organismus tragen dieselbe genetische Information und folgen dennoch im Laufe ihrer Entwicklung einem eigenen transkriptionellen Programm, aus dem phänotypische und funktionelle Unterschiede resultieren. Diese unterschiedlichen Genexpressionprogramme werden durch epigenetische Mechanismen kontrolliert, welche Chromatinmodifikationen wie DNA Methylierung und posttranslationale Histonmodifikationen beinhalten. Der Einfluss von posttranslationalen Chromatinmodifikationen auf Chromatin assoziierte Prozesse, ihre Assoziation mit verschiedenen genomischen Elementen und die Bedeutung ihrer Etablierung und Erhaltung sind von großem wissenschaftlichen Interesse. Der primäre Fokus dieser Arbeit lag auf der Analyse der genomweiten Verteilung von Chromatinmodifikationen sowie deren lokalen Auswirkungen.

Der erste Teil dieser Arbeit behandelt die Entwicklung und Anwendung neuer affinitätsbasierter Reagenzien für die lokale und genomweite Charakterisierung von Histonmodifikationen. Dafür verwendeten wir native und speziell weiterentwickelte rekombinante Proteine, die als natürliche Eigenschaft eine spezifischen Interaktion mit modifizierten Histonen aufweisen. Durch einen präzisen Vergleich mit qualitativ hochwertigen Antikörpern gerichtet gegen dieselben Histonmodifikationen, konnten wir erfolgreich die Funktionalität dieser neuen affinitätsbasierten Reagenzien in Methoden wie dem Western Blot und der Chromatinpräzipitation gekoppelt mit quantitativer PCR oder Next Generation Sequenzierung nach etablierten Kriterien der Qualitätskontrolle überprüfen und aufzeigen. Die entwickelten rekombinanten Affinitätsreagenzien besitzen zudem im Vergleich zu Antikörpern die Vorteile von geringen Produktionskosten in Kombination mit hoher Ausbeute, die Möglichkeit der simplen Modifikation und eine gleichbleibende Qualität, welche vor allem die Reproduzierbarkeit von Daten verbessert.

Der zweite Teil dieser Arbeit beschäftigt sich mit der Aufklärung des Grundmechanismus, mit dem epigenetische Modifikationen auf die Zellentwicklung Einfluss nehmen. Um die Effekte und die Dynamik von Histonmodifikationen und der DNA Methylierung in situ zu studieren, etablierten wir ein Protokoll zur gezielten Promotormethylierung von *VEGF-A* (Vaskularer endothelialer Wachstumsfaktor A). Zu diesem Zweck wurden die Zellen mit einem adenoviralen Vektor infiziert, der für einen Target-Zinkfinger fusioniert mit der katalytischen Domäne eines epigentischern Enzymes wie DNA- oder Histonmethyltransferasen kodiert. Die Etablierung, der lokale Effekt und die Kinetik der eingeführten Chromatinmodifikationen konnten im Detail verfolgt werden. Unsere Daten zeigen, dass lokales Editieren eines Ziellokus seinen initialen Chromatinstatus verändern und als Konsequenz die Transkription modifizieren kann. Der Effekt überdauert jedoch nur eine kurze Zeit, bevor der Chromatinstatus in seinen nativen Zustand zurückfällt.

Zusammenfassend wurden während dieser Arbeit neue, affinitätsbasierte Reagenzien zur Analyse der genomischen Verteilung von Histonmodifikationen erfolgreich entwickelt und angewendet. Nach einem ähnlichen Schema wurde eine Strategie zur gezielten Chromatinmodifizierung entwickelt und angewendet, die es erlaubt die Dynamik, den induzierten Effekt und die Aufrechterhaltung von Chromatinmodifizierungen zu studieren.

Abstract

All cells in a multicellular organism carry the same genetic information, and yet throughout their lifetimes they follow unique transcriptional programs, which lead to phenotypical and functional differences. These differential gene expression programs are enacted by highly coordinated epigenetic mechanisms, which include modifications of chromatin, such as DNA methylation and histone post-translational modifications. Their involvement in chromatin-associated processes, association with different genomic elements and the means of their establishment and maintenance are crucial scientific issues. The primary focus of this study was to shed light on the genome-wide distribution of chromatin modifications and the effects of their local establishment.

First, we focused our efforts into developing and applying novel affinity reagents for local and genome-wide characterization of histone modifications. We made use of native and engineered recombinant proteins that have an intrinsic ability to interact specifically with modified histones. In a rigorous side-by-side comparison with high quality histone modification antibodies, we successfully applied these novel affinity reagents in approaches such as western blot and chromatin precipitation coupled with quantitative PCR or massively parallel sequencing, following established quality control criteria. By this, we validated the feasibility of this strategy. We also discussed in detail the advantages of using recombinant proteins in lieu of antibodies, such as their cheap production with high yield, ease of protein engineering and consistent quality and reproducibility.

Secondly, we wanted to clarify some of the principal mechanisms by which epigenetic modifications operate inside the cell. To this aim, we established and applied an approach based on zinc finger targeted promoter methylation of *VEGF-A* (vascular endothelial growth factor) in order to study the *in vivo* effects and dynamics of histone modifications and DNA methylation. Adenoviral constructs made of the targeting zinc finger were fused with catalytic domains from epigenetic enzymes such as DNA and histone methyltransferases and were used to infect cells. By these means we were able to successfully follow in detail the establishment, effects and kinetics of the installed chromatin modifications. Our data indicate that local chromatin editing of a target locus can change its intial chromatin state and consequently modulate its transcriptional output, albeit for a short period of time, before it returns to its native configuration.

In conclusion, in this body of work we were able to successfully develop and apply novel affinity reagents for studying the distribution of histone modifications. Along the same lines, we also successfully developed and applied a strategy for targeted chromatin editing, which allowed us to study the dynamics of establishment, downstream effects and maintenance of chromatin modifications.

List of abbreviations

3C	Chromosome conformation capture. A method for studying 3D interactions within chromosomes
Ac	Acetyl
ADD	ATRX-Dnmt3a-Dnmt3l
AGO	Argonaute protein family, which plays a central role in RNA silencing processes
ATRX	Alpha-thalassemia mental retardation syndrome
BAM	Binary Alignment Map (format after mapping of DNA sequences)
BED	Browser Extensible Data (format of peak regions Chr Start End)
BEDGRAPH	Display of continuous-valued data in track format preserved in original state
BIGWIG	Means to display dense, continuous data as a graph
CBX	Chromobox
CEAS	Cis-regulatory Element Annotation System
CID	Collision Induced Dissociation
CIDOP	Chromatin Interacting DOmain Precipitation
ChAP-MS	Chromatin Affinity Purification with Mass Spectrometry
Chromatin	A complex mixture of DNA, RNA, histones and non-histone proteins
ChIP	Chromatin ImmunoPrecipitation
Chip	DNA microarrays
Chromodomain	CHRomatin Organization MOdifier domain
COMPASS	COMplex of Proteins ASsociated with Set1
CpG	Cytosine phosphate guanine. Phosphate links the two nucleosides together
C-myc	Regulator gene that codes for a transcription factor
CTCF	CCCTC-binding factor
DAPI	4',6-Diamidino-2-Phenylindole, Dihydrochloride
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease I
Dnmt1	DNA methyltransferase 1
Dnmt3a	DNA methyltransferase 3a

Dnmt3b	DNA methyltransferase 3b
Dnmt3L	DNA methyltransferase 3-like
ENCODE	Encyclopedia of DNA elements
EPL	Expressed Protein Ligation
ETD	Electron Transfer Dissociation
E(Z)	Enhancer of Zeste
FASTQ	Text-based format for storing both a biological sequence (usually nucleotide sequence) and its corresponding quality scores
FPKM	Fragments Per Kilobase per Million
FRET	Förster Resonance Energy Transfer
G9A	Euchromatic histone lysine N-methyltransferase 2
GLP	G9A-like protein
HAT	Histone AcetylTransferase
HEK293	Human Embryonic Kidney fibroblasts 293
HepG2	Hepatocellular carcinoma G2 cells
HMID	Histone Modification Interacting Domain
Hox	Homeobox
HP1	Heterochromatin Protein 1
IGV	Integrative Genomics Viewer
iMEF	Immortalized Mouse Embryonic Fibroblasts
ING	Inhibitor of Growth
LexA	Repressor protein that represses the SOS system in bacteria
LINE	Long Interspersed Elements
MACS	Model-based Analysis of ChIP-seq
Me	Methyl
Me1	Monomethyl
Me2	Dimethyl
Me3	Trimethyl
MNase	Micrococcal nuclease
MLA	Methyl Lysine Analogues
MPP8	M-Phase Phosphoprotein 8
MS	Mass Spectrometry

NPL	Native Chemical Ligation
NcRNA	Noncoding RNA
OD_{600}	Optical Density at 600 nm
PCR	Polymerase Chain Reaction
PICh	Proteomics of Isolated Chromatin
PHD	Plant HomeoDomain
PRC1	Polycomb Repressive Complex 1
PRC2	Polycomb Repressive Complex 2
PTM	Post-Translational Modifications
PWWP	Pro-Trp-Trp-Pro motif-containing protein
qPCR	quantitative PCR
RITS	RNA-Induced Transcriptional Silencing
RNA	Ribonucleic acid
RNAP II	RNA polymerase II
RPM	Reads Per Million
RPKM	Reads Per Kilobase per Million
SAHF	Senescence-Associated Heterochromatin Foci
SAM	S-Adenosyl Methionine
Seq	Massively parallel sequencing
SET	Su(Var)3-9 (Suppressor of variegation 3-9), E(Z), and Trx (Trithorax) domain
SDS-PAGE	Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
SICER	Spatial clusterIng approach for the dentification of ChIP-enriched regions
Suv39H1/2	Histone-lysine N-methyltransferase SUV39H1/2. Suppressor of variegation 3-9 homolog 1/2 (Drosophila)
TAD	Topologically Associating Domains
TALE	Transcription Activator-Like Effector
TET	Ten Eleven Translocation enzyme
tRNA	transport RNA
TSS	Transcription Start Site
TTS	Transcription Termination Site
UHRF1	Ubiquitin-like, containing PHD and RING finger domains 1

UCSC	University of California Santa Cruz
VEGF-A	Vascular Endothelial Growth Factor A
WCX-HILIC	Weak Cation eXchange Hydrophilic Interaction LIquid Chromatography
WIG	Display of continuous-valued data in track format
ZF	Zinc finger protein

1. Introduction

1.1 Epigenetics

The term epigenetics (in the context of "epigenetic landscape") was coined by Conrad Waddington in 1942 as a blend between **epi-** (Greek: επί- above, over, outside of, around) - **genetics**. He defined epigenetics as "the branch of biology, which studies the causal interactions between genes and their products, which bring the phenotype into being" (Waddington, 1942). Waddington's epigenetic landscape was a metaphor for biological development and cellular differentiation, described as a marble rolling down a landscape until it reaches the lowest point of local elevation. The lowest point of local elevation in this case represents the final cell fate, whilst the journey of the marble symbolizes the biological development from totipotency to cellular specialization (Fig. 1).

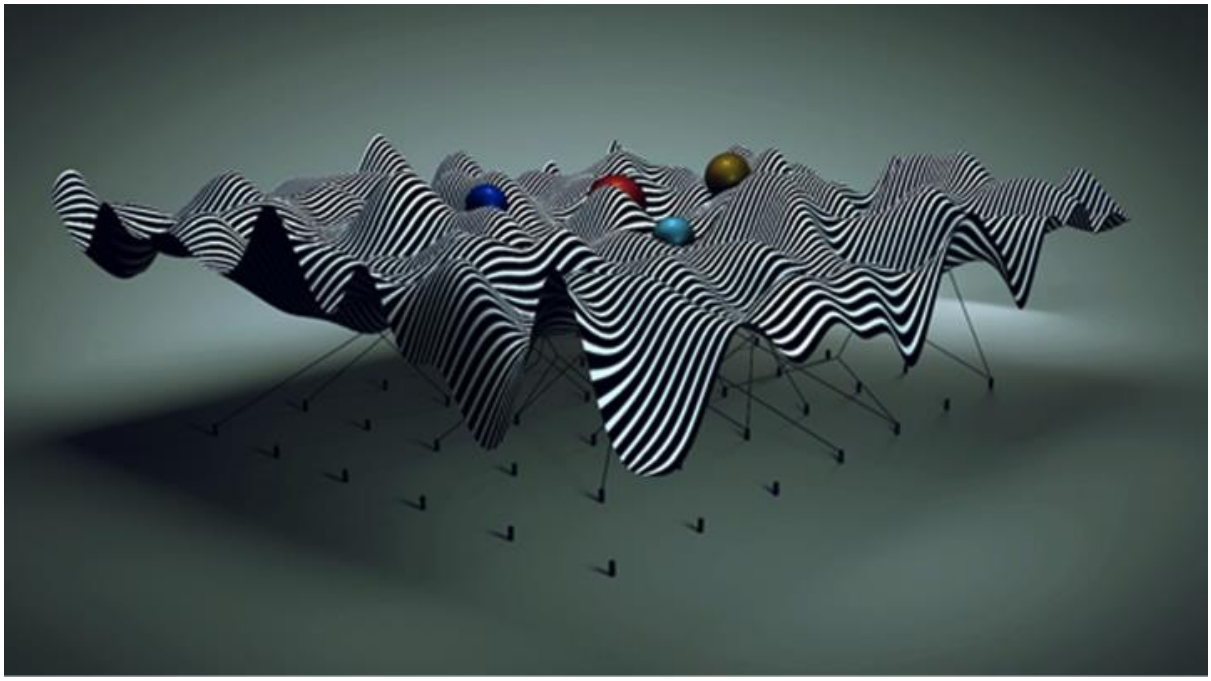

Figure 1. Graphic rendition of the epigenetic landscape, surmised by Waddington. The journey of the marbles represents cellular development, whilst the end points represent the final cell fate, which Waddington at that time thought it was irreversible. The picture is taken from: https://vivomotion.wordpress.com/folio/.

Since then, the term epigenetics has underwent a couple of cycles of evolution and redefinition, and currently the consensus working definition is that "epigenetics is the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequences" (Riggs, 1996) or similarly "an epigenetic trait is a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (Berger et al., 2009). In the latter operational definition of epigenetics, it has been suggested that three major categories of signals are involved in the triggering, establishment, maintenance and heritability of epigenetic states. An "epigenator" is an external environmental signal that can trigger the establishment of an epigenetic pathway. The "epigenetic initiator" is the receiver of

the external signal that can be precisely targeted to modify the local chromatin environment and set the epigenetic state at a local chromatin position. This is followed by the "epigenetic maintainer," which has a principal role of sustaining the initiated epigenetic state (Fig. 2).

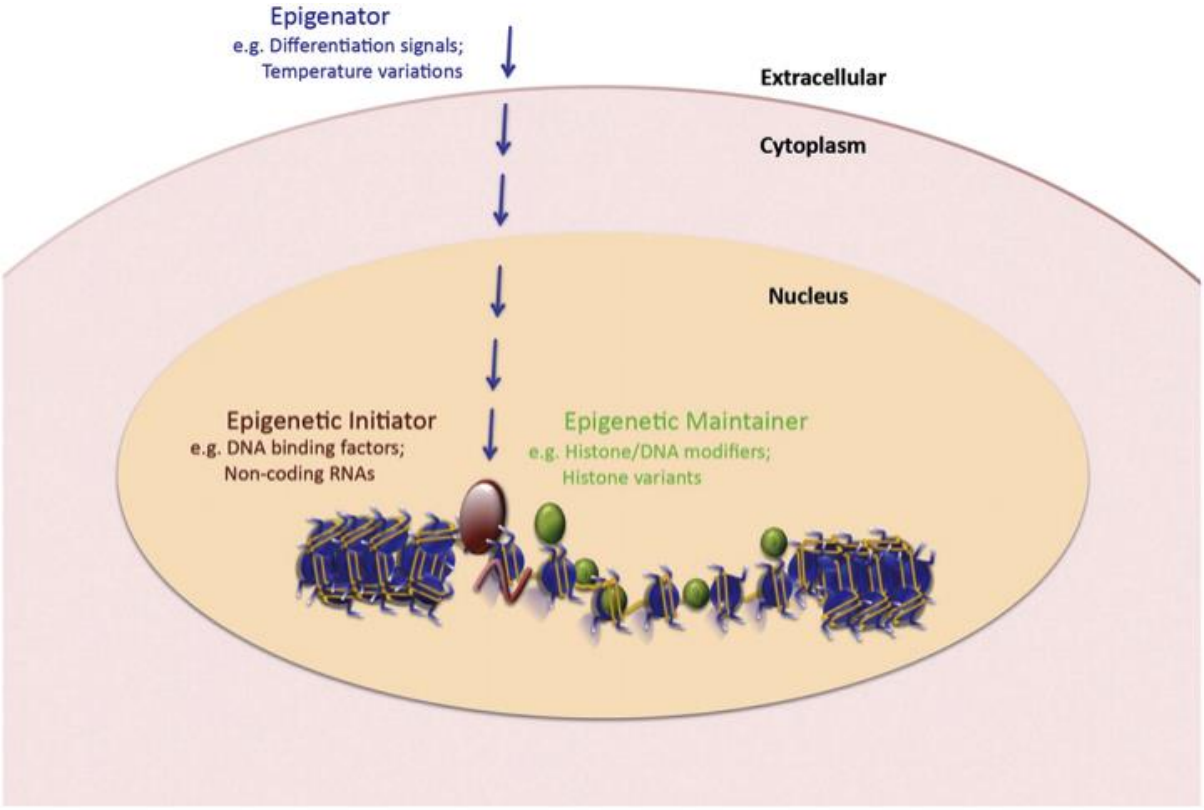

Figure 2. Graphic illustration of the three major signals involved in triggering, establishment, maintenance and heritability of epigenetic states. The figure is taken from (Berger et al., 2009).

Lately, the definition of epigenetics has become less rigid and the term is used more loosely to explain all chromatin-templated phenomena associated with transcription, replication, DNA repair et cetera, regardless of whether they are heritable in the classical sense, transient or a mix of being dynamic and heritable. Epigenetics is also trying to explain the mechanisms by which different cells, from the same organism, harboring the same DNA sequence obtain distinct and unique cellular programs and identities (Goldberg et al., 2007). This expansion is reasonable, if one process controls chromatin biology in non-dividing cells (where inheritance is not possible) but the same process mediates heritable effects in other systems. On the other hand the fashionable designation of any transient regulatory event as "epigenetic" should be avoided.

1.2 Principal organization and structure of chromatin

The genetic information of each and every eukaryotic cell is encoded within the linear sequence of its DNA molecule. In order for DNA to fit inside the miniscule nucleus it has to be condensed and compacted about 10^6 fold of its original length. This is achieved with the help of histone and non-histone proteins, and the resulting mishmash of proteins and nucleic acids is called

chromatin. The smallest building unit of chromatin is the nucleosome, which consists of 147 bp of DNA wrapped around an octamer of highly basic proteins called histones. The histone octamer is composed of two copies of each of the core histone proteins, H2A, H2B, H3 and H4. The final configuration of the nucleosome is made of an H3-H4 heterotetramer and two H2A-H2B heterodimers (Luger et al., 1997) (Fig 3A). Then, individual nucleosomes are annexed consecutively with a linker DNA and form the first level of chromatin organization termed "the 11-nm fiber." This entity of regularly interspaced nucleosomes can be directly observed as "beads on a string structure" in electron micrographs (Fig. 3B).

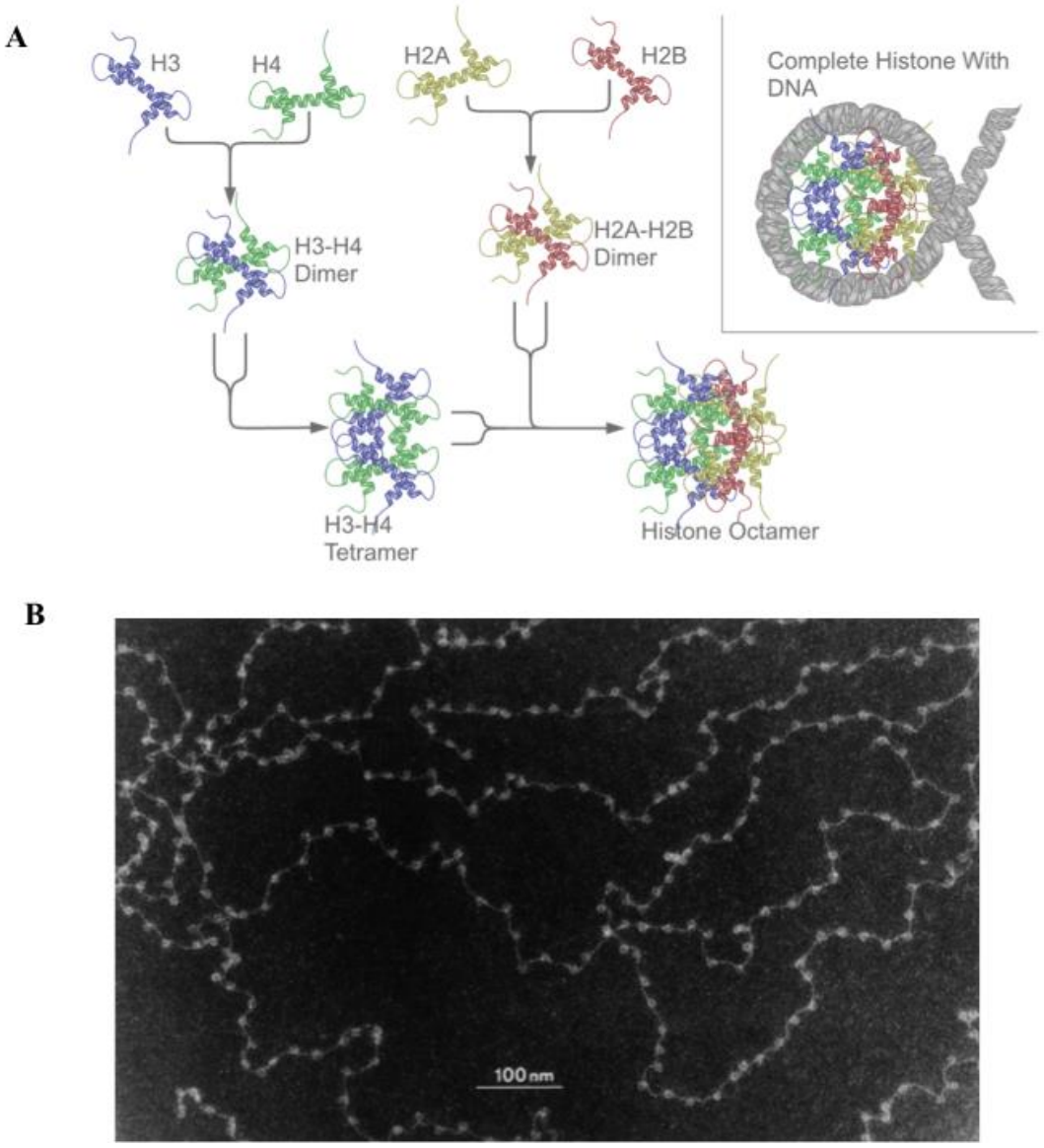

Figure 3. First level of chromatin organization. A) Schematic illustration of the formation of the core histone octamer. Histones H3 and H4 form a heterotetramer, whilst histones H2A and H2B form two heterodimers. The image is licensed under CC BY-SA 3.0 via Wikimedia Commons. B) Beads on a string, 11-nm chromatin structure. Experimental evidence from a darkfield electron micrograph of chromatin spilling out of a chicken erythrocyte nucleus stained with uranyl acetate. The image is taken from (Olins and Olins, 1972).

Upon addition of histone H1 (also called linker histone), the beads-on-a-string structure transforms into a 30 nm fiber, which is the second level of chromatin organization (Robinson

and Rhodes, 2006). The exact structure of the 30 nm fiber or even its mere existence is still hotly debated (Grigoryev and Woodcock, 2012; Maeshima et al., 2010). The two most popular models for the organization of the 30 nm fiber is the so called one-start-helix (solenoid), where individual nucleosomes form a superhelix of nucleosomes that follow each other along the same helical path, with 6-8 nucleosomes per turn (Widom et al., 1985). In an alternative model, a so called two-start-helix is suggested to form a zigzag structure where the linker DNA connects two opposing nucleosomes making each alternate nucleosomes into interaction partners (Dorigo et al., 2004) (Fig. 4A).

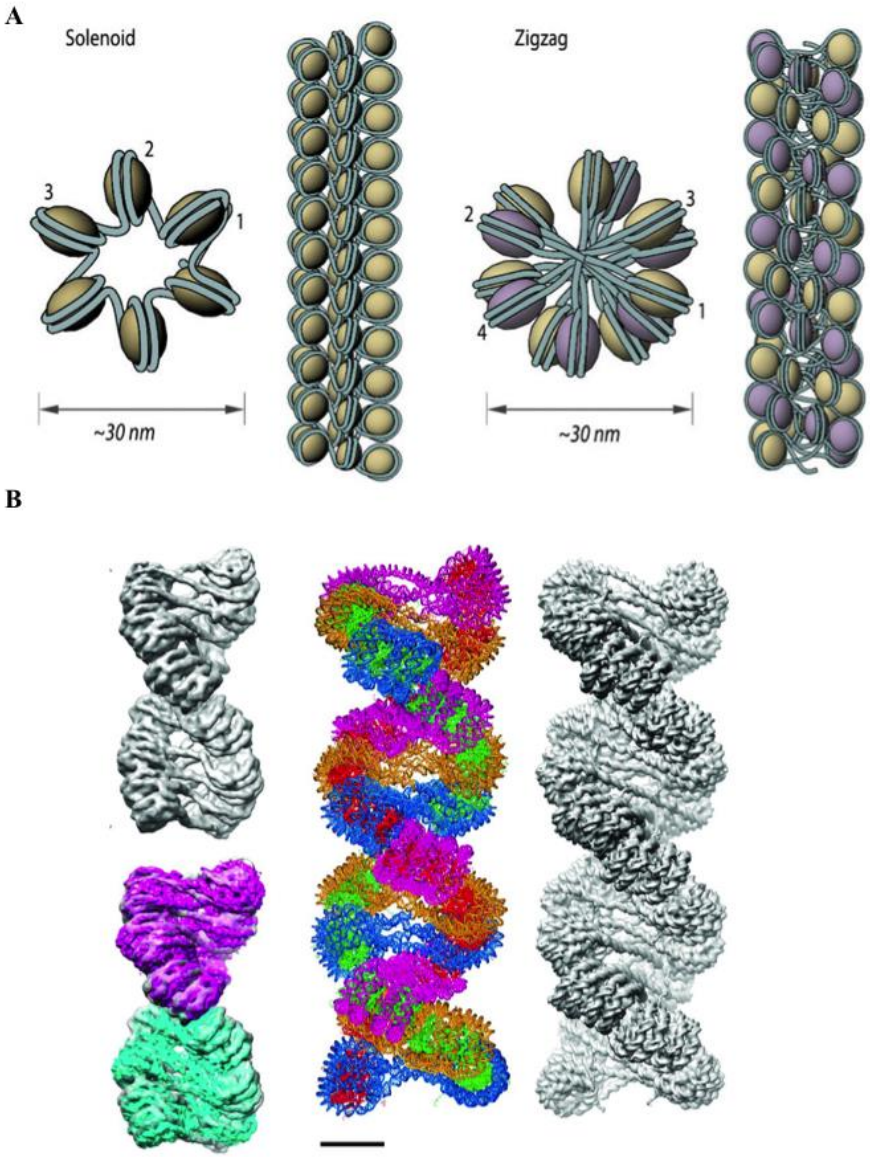

Figure 4. Second level of chromatin organization. A) One-start solenoid model (left) and two-start zigzag model (right) of the 30 nm fiber. The picture is taken from http://www.mechanobio.info/. B) The most recent high resolution structure of a 12-nucleosome fiber indicates that the 30 nm structure complies to the two-start zigzag model and it forms a double helical structure, just like the underlying DNA molecule. The image is taken from (Song et al., 2014).

The discord between the experimental data to which the two models are based on is likely due to differences in the methods used for preparation of the 30 nm fiber. Recently, with the help

of cryogenic electron microscopy the structure a 12-nucleosome string of DNA was determined at 11-angstrom resolution (Song et al., 2014). These data indicate that under the experimental conditions the structure of the 30 nm fiber is a two-start zigzag twisting double helix, just like the DNA molecule it packages within its confines (Fig. 4B). Whether this structure corresponds to the native 30 nm fiber is still an open question (Travers, 2014).

Furthermore, it is speculated that additional higher order structures are generated by inter-fiber interactions and formation of intricate loops (Li and Reinberg, 2011). Recent technological advancements based on proximity dependent formaldehyde crosslinking of any genomic loci (3C methods) (Dekker et al., 2002) have been used for probing the 3D folding of chromosomes and have provided incredible insights into the innerworkings and formation of higher-order chromatin structures (Gibcus and Dekker, 2013). The data obtained within the past few years, suggest a hierarchical structural organization of chromatin, starting with loops of genes and enhancers and ending up with chromatin domains and nuclear compartments. First, experiments have unveiled the presence of gigantic three-dimensional subchromosomal entities called compartment A (gene rich, transcriptionally active, DNase I hypersensitive) and compartment B (gene poor, transcriptionally silent, DNase I insensitive) (Lieberman-Aiden et al., 2009), which are cell type specific. Each compartment is further divided into a string of smaller chromatin domains with sizes of 100 kb to 1 Mb, referred to as topologically associating domains (TADs) (Dixon et al., 2012; Nora et al., 2013; Sexton et al., 2012). These domains have coordinated dynamics of formation and expression during differentiation, and most of the promoter-enhancer interactions happen within a TAD (Shen et al., 2012).

The organization of chromatin requires a balance between stability and flexibility of chromatin states. For example, chromosomal territories and compartments are stable within one cell cycle, but variable from one cell cycle to the other. On the other hand, the interactions within TADs are more stochastic and less stable, but the TADs themselves are reproducible and consistent from one cell cycle to the other, which in turn safeguards the identity of the cell (Gibcus and Dekker, 2013).

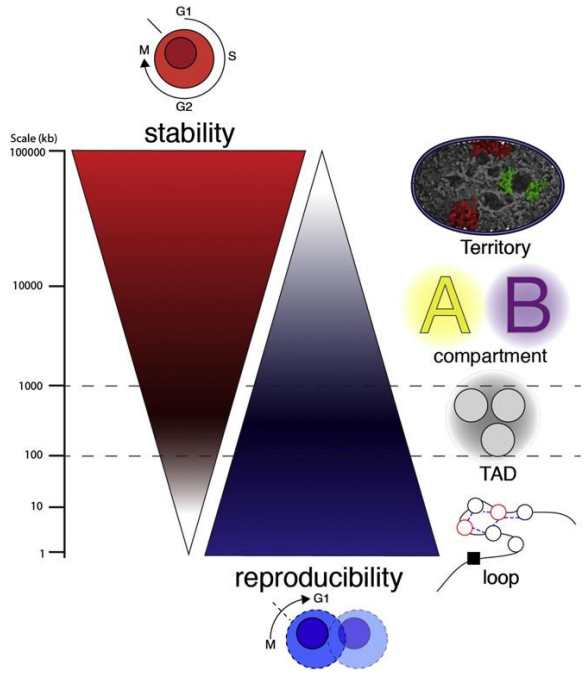

Figure 5. Higher levels of chromatin organization. Chromatin needs a balance between stability and flexibility. The image is taken from (Gibcus and Dekker, 2013).

1.3 Epigenetic mechanisms: DNA methylation, noncoding RNAs

Most epigenetic phenomena can be attributed to three types of signals: DNA methylation, histone post-translational modifications (PTMs) and noncoding RNAs. They all hold crucial regulatory roles in all chromatin-templated processes, such as the establishment and maintenance of transcriptional programs and transcription in general (Shilatifard, 2006), DNA replication (Rivera et al., 2014), genome integrity and DNA repair (Papamichos-Chronakis and Peterson, 2013), chromatin compaction (Shogren-Knaak et al., 2006), X chromosome inactivation, imprinting (Jurkowska and Jeltsch, 2013) et cetera. Aberrant epigenetic signals have been associated with many diseases, including cancer (Greer and Shi, 2012; Portela and Esteller, 2010; Suva et al., 2013; Timp and Feinberg, 2013). For example, models suggest that the epigenetic dysregulation in many cancers leads to epigenetic flexibility and epigenetic drift, resulting in tumor heterogeneity and selection with darwinian mechanisms (Timp and Feinberg, 2013). Similarly, aberrant epigenetic signals are the culprits for imprinting disorders such as Angelman syndrome and Prader-Willi Syndrome (Butler, 2009).

DNA methylation in mammals is referred to as the transfer of a methyl group from S-adenosyl methionine (SAM) to the 5' position of cytosine (Jeltsch, 2002), mostly in a CpG context (Fig. 6A). Around 60-80% of all CpG sites in the human genome are methylated. The genes with heavily methylated CpG islands at their promoters are strongly repressed, whilst gene body methylation is associated with active transcription (Jones, 2012). Also, DNA methylation has a vital role in the suppression of transposable elements and repetitive sequences (Jia et al., 2009; Jin and Robertson, 2013).

Functionally, DNA methylation can affect transcription in two ways: first, 5-cytosine methylation might impede the binding of transcription factors, such as c-myc and CTCF; second, methylated cytosines can serve as landing pads for methyl-binding proteins, which can directly recruit repressor protein complexes, such as histone deacetylases (HDACs) or histone lysine methyltransferases (HKMTs) (Miranda and Jones, 2007).

The enzymes responsible for introducing methyl moieties on target cytosines in mammalian cells are called DNA methyltransferases (Dnmts). Based on their function they can be classified into **maintenance DNA methyltransferases** (Dnmt1) and ***de novo* DNA methyltransferases** (Dnmt3a, Dnmt3b and the inactive Dnmt3l). Dnmt1 is the maintenance methyltransferase and its catalytic domain has an intrinsic preference for hemimethylated DNA (Bashtrykov et al., 2012), which is aided by recruitment from the UHRF1 protein (Zhang et al., 2011). Dnmt3a and Dnmt3b are responsible for the establishment of DNA methylation during development, and their activity is stimulated by Dnmt3l, which itself is catalytically inactive (Denis et al., 2011; Jurkowska et al., 2011). The classical model of establishment of DNA methylaiton (by Dnmt3a/3b) and its maintenance (by Dnmt1 based on its preference for hemimethylated substrates), albeit conceptually elegant, is not fully compatible with the experimental data (Jeltsch and Jurkowska, 2014) (Fig. 6B). Firstly, Dnmt1 has a considerable *de novo* activity on unmethylated substrates *in vitro* and some repetitive sequences *in vivo*, and secondly, Dnmt3a and Dnmt3b are necessary for maintenance of repetitive sequences and LINEs (Chen et al., 2003). Based on this, a new stochastic model has been suggested (Jeltsch and Jurkowska, 2014), where the methylation at each site is determined by the local rates of methylation and demethylation (Shipony et al., 2014). This is dependent on the intrinsic catalytic properties of the Dnmts, but also their targeting and regulation, which is controlled by feedback mechanisms.

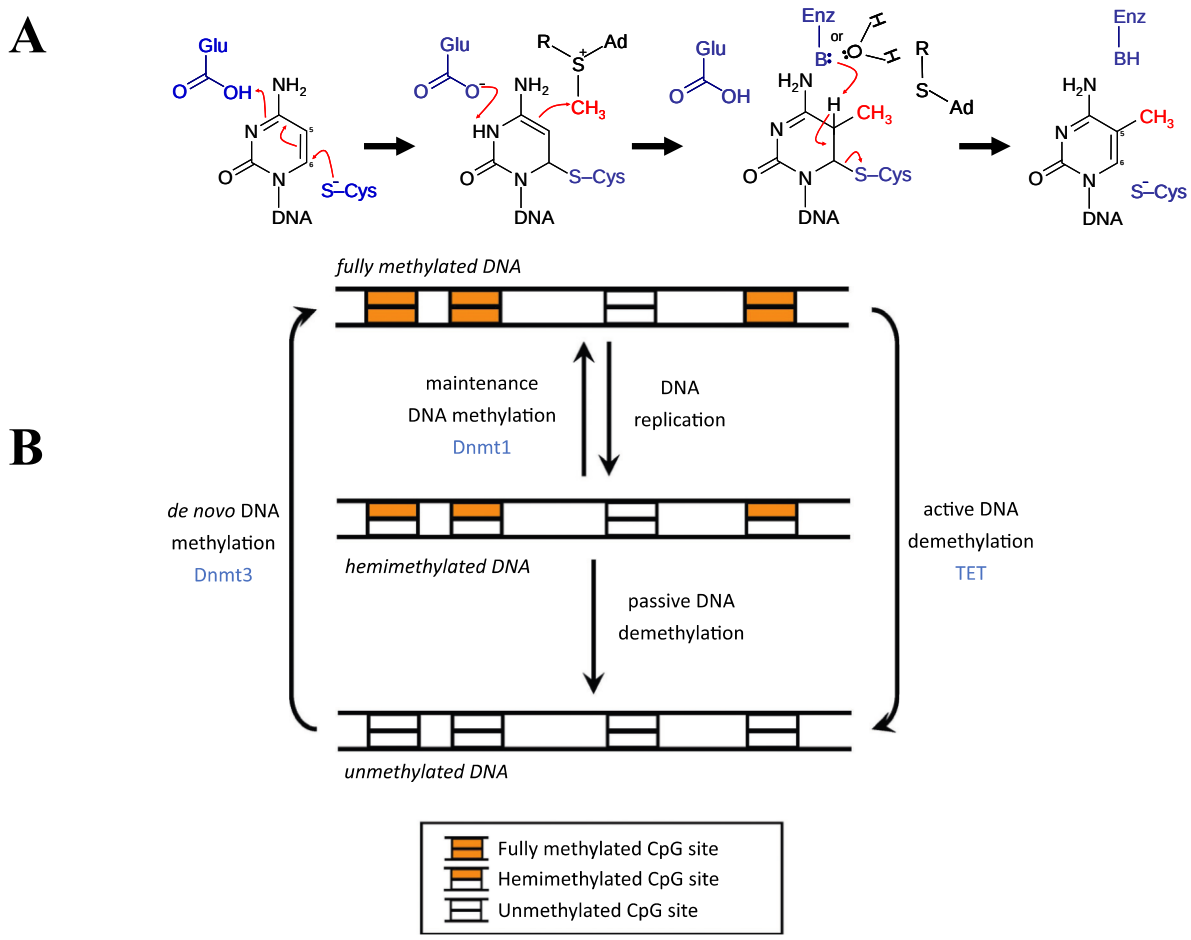

Figure 6. DNA methylation. A) The reaction mechanism of transfer of a methyl group from S-adenosyl methionine to 5-cytosine. The figure is taken from (Jurkowska, 2008). B) The classical mechanism of establishment, removal and maintenance of DNA methylation. The image is taken from (Jeltsch and Jurkowska, 2014).

Recently, noncoding RNAs (ncRNAs) have been catapulted into the limelight of epigenetic research, as an additional layer of biological regulation. They are involved in a vast variety of biological functions, such as X-chromosome inactivation, inactivation of repetitive and parasitic sequences, post-transcriptional regulation of gene expression and enhancer transcription (Cech and Steitz, 2014). They come in many different flavors, such as small noncoding and long noncoding RNA molecules. The former are involved in argonaute (Ago) dependent post-transcriptional control of gene regulation or Ago-dependent recruitment of histone and DNA methyltransferases to complementary RNA scaffolds onto chromatin. The latter are involved in *cis* or *trans* recruitment of chromatin-modifying complexes to specific sites in the genome (Holoch and Moazed, 2015; Mercer and Mattick, 2013).

1.4 Histone modifications and histone variants

Apart from compacting the DNA within the confines of the small nucleus, chromatin is a highly dynamic entity with an immense regulatory power. The building blocks of chromatin are the nucleosomes, which consist of 147 bp of DNA wrapped around highly charged basic histone proteins. Histones carry unstructured N-terminal tails, which protrude from the core nucleosome and are subjected to a staggering array of post-translational modifications (Kouzarides, 2007). Moreover, lately it has become clear that the globular surface of the

nucleosome can also undergo post-translational modifications (Tropberger and Schneider, 2013). The focal point of chromatin regulation is thought to lie within "the histone code" made of histone PTMs, which together with DNA methylation extend the information encoded in the genome (Jenuwein and Allis, 2001). The broad compendium of histone modifications divides into two groups: first, addition of small chemical groups such as acetylation, methylation, phosphorylation and second, addition of larger bulky protein modules, such as ubiquitylation and sumoylation (Bannister and Kouzarides, 2011; Kouzarides, 2007) (Fig. 7). An additional layer of complexity is given by the fact that the ε-amino group of lysines can be mono-, di- and trimethylated, whilst the guanidine-ε-amino group of arginines can be mono- or dimethylated in a symmetric or asymmetric manner (Fig. 8).

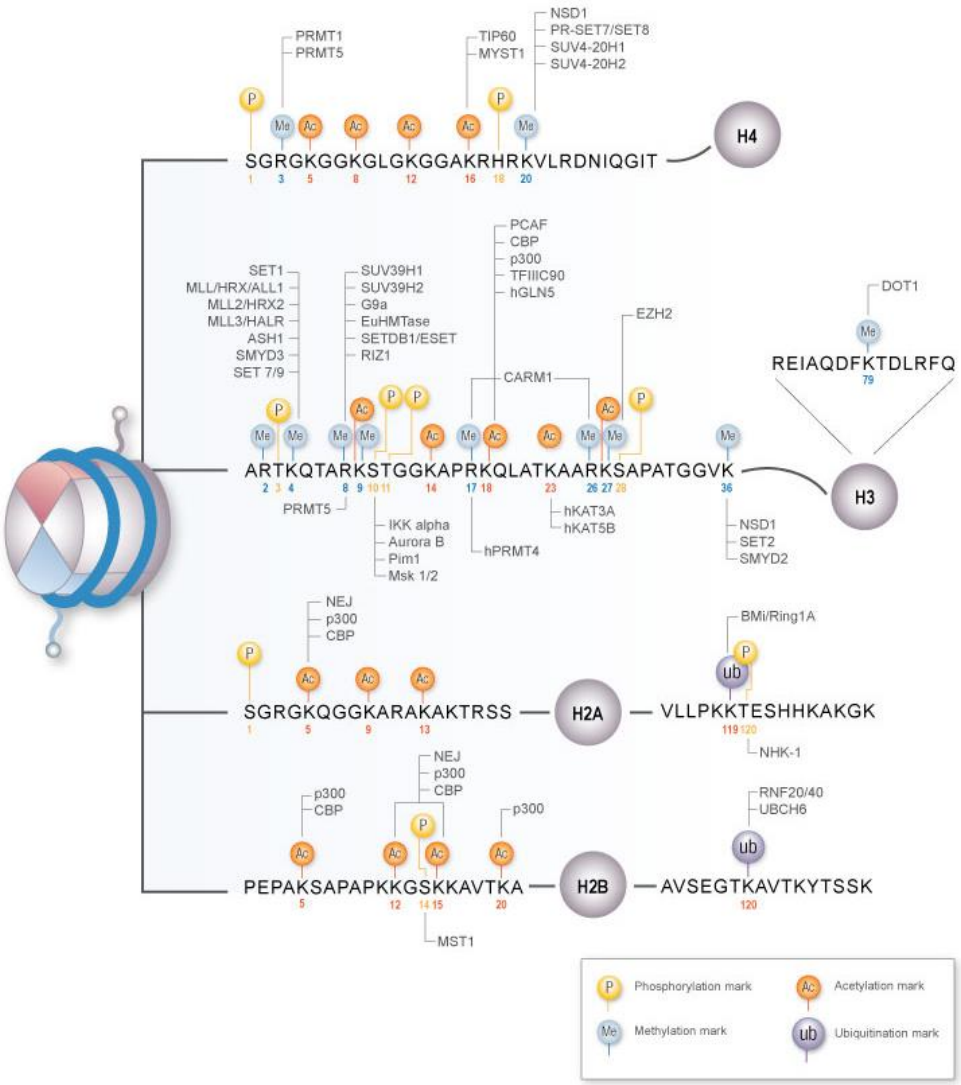

Figure 7. The N-terminal tails of histones are decorated with many post-translational modifications. Most of them occur on the tails of histone H3 and H4. The responsible enzymes are given above each modification. The image is taken from http://www.epitomics.com/ modified from (Margueron et al., 2005).

Functionally, histone PTMs can have three salient effects: firstly, they directly alter the biophysical properties of the nucleosome by disrupting the electrostatic contacts between the histone proteins and DNA, which could lead to destabilization and increased accessibility of chromatin. Secondly, they can affect protein/protein interactions between nucleosomes and

lead to differential condensation states of chromatin (Shogren-Knaak et al., 2006; Song et al., 2014). Thirdly, histone PTMs can serve as specific "recruitment platforms" for chromatin complexes that harbor histone modification interacting domains (HMIDs) or popularly termed "reading domains" (Fig. 9). For instance, methylation is recognized by chromo-like domains of the Royal family (chromo, tudor, MBT) and PHD domains, acetylation is recognized typically by bromodomains, but also PHD and tudor domains and finally phosphorylation is recognized by 14-3-3 proteins (Kouzarides, 2007). Compared to DNA methylation, histone PTMs are more dynamic, with histone methylation being the most stable with lower turnover rate than for instance histone acetylation and phosphorylation (Jenuwein and Allis, 2001). On a similar note, the nucleosomes assosicated with actively transcribed chromatin have the highest turnover rates (Deal and Henikoff, 2010; Dion et al., 2007; Radman-Livaja et al., 2011; Rufiange et al., 2007).

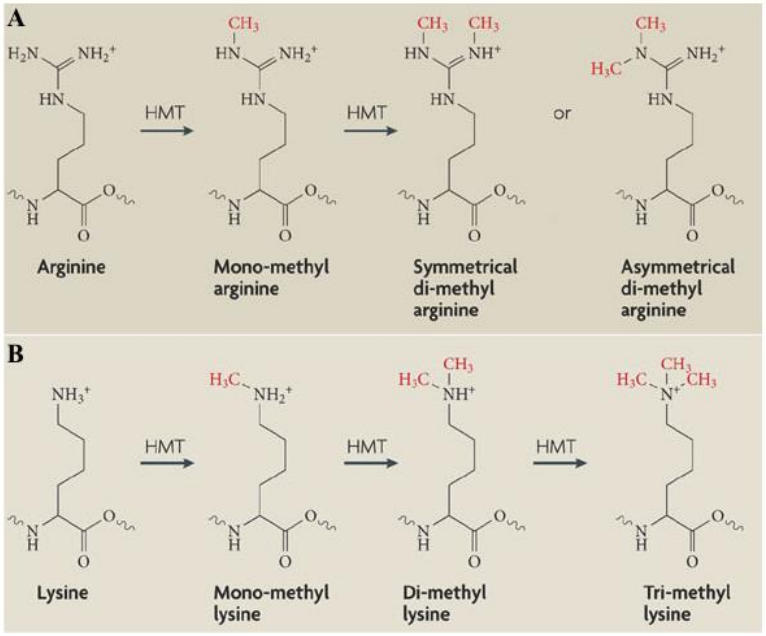

Figure 8. Differential methylation states of arginine and lysine residues within histones. A) Arginine residues can be mono- and dimethylated. Dimethylation occurs either symmetrically or asymmetrically. B) Lysine residues can be mono- di- and trimethylated. The image is adapted from (Klose and Zhang, 2007).

Based on correlative studies, histone modifications belong in two categories, associated with transcriptional activation (H3K4me2/3, H3K36me2/3 and acetylations) or transcriptional repression (H3K9me3, H3K27me3, H4K20me3) (Barski et al., 2007; Kouzarides, 2007). Regardless of whether histone PTMs have a role in transcription, repair or maintenance of chromatin structure, they are indispensable for cellular identity and have to be transmitted correctly with each cell division. Most models predict that histone PTMs can be effectively transmitted only if they exist on numerous adjacent nucleosomes, minimizing the dilution effect (Huang and Jiao, 2012; Margueron and Reinberg, 2010). Namely, one scenario for histone segregation and deposition of "old" and "new" histones after DNA replication is that old H3 and H4 histones are randomly partitioned to both DNA strands as tetramers (Deal and Henikoff, 2010; Xu and Zhu, 2010) while H2A-H2B dimers are released and re-enter from a pool of free histones. The histone PTMs from the randomly distributed pool of the orginal H3-H4 tetramers can be copied to new H3-H4 tetramers by "reader-writer" positive feedback loop (Gaydos et al., 2014; Margueron and Reinberg, 2010). Another scenario is that chromatin modifiers (such as histone methyltransferases) persist on the locus during replication and they are the mediators of epigenetic inheritance that re-establish the histone PTM patterns after cell division (Petruk et al., 2012).

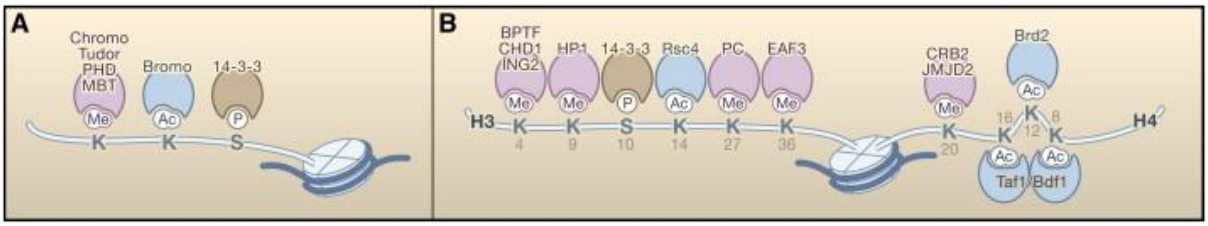

Figure 9. Histone modifications can serve as recruitment platforms for histone modification interacting domains or popularly termed "reading domains." These domains are part of chromatin complexes and they are major targeting platform for epigenetic activities. The image is taken from (Kouzarides, 2007).

1.4.1 Histone variants

Canonical histones H2A, H2B, H3 and H4 are encoded in repeat arrays, expressed only during S-phase and they are deposited in a replication dependent manner. This serves to fill-in the gaps left during replication and package the newly synthetized genome. In contrast to canonical histones, histone variants are specialized histone isoforms, found in single copy genes that are expressed throughout the cell cycle (Chen et al., 2013; Marzluff and Duronio, 2002). They have distinct amino acid sequences from their canonical paralogues, which alter the properties of the nucleosome (Weber and Henikoff, 2014). The deposition of histone variants is not random and their incorporation into chromatin is tightly regulated by different histone chaperones, which target them to specific genomic locations. So far, only histone H2A and H3 variants have been discovered, whereas variants of histone H2B and H4 are yet to be uncovered.

One of the most common histone variants is **H2A.Z**, which comprises around 15% of total H2A with a nonrandom genomic distribution in both eu- and heterochromatin (Fan et al., 2004; Sarcinella et al., 2007). The chromatin remodeller Swr1 acts on nucleosome-depleted regions around promoters (Ranjan et al., 2013) and replaces canonical H2A with H2A.Z, while Ino80 does the opposite (Papamichos-Chronakis et al., 2011). This might explain the H2A.Z promoter distribution in yeast, but not in other model organisms, whereby H2A.Z is also enriched in gene bodies. In metazoans, H2A.Z correlates with gene expression levels (Nekrasov et al., 2012; Weber et al., 2010) and is anticorrelated with DNA methylation (Conerly et al., 2010; Zemach et al., 2010).

The histone variant **H2A.B** is found only in mammals. First, it has been revealed that it is deficient in Barr bodies (and therefore sometimes termed H2A Barr-body deficient or H2A.Bbd) (Chadwick and Willard, 2001). Later, it was also discovered that this variant lacks an acidic patch (therefore sometimes refered to as H2A.Lap1 – lacks acidic patch) (Soboleva et al., 2012). The lack of acidic patch is known to inhibit the formation of compact chromatin (Zhou et al., 2007). The terminology has been unified and this variant is now named H2A.B. Further efforts are needed to fully characterize the function of this histone variant.

MacroH2A is three times the size of canonical H2A because it contains an additional non-histone globular (macro) domain (Chakravarthy et al., 2005). Both, MacroH2A and canonical H2A participate in the formation of heterotypic nucleosomes (Chakravarthy and Luger, 2006). Macro domains are known to bind NAD^+ metabolites (also poly-ADP-ribose) (Han et al., 2011) and in the context of macroH2A, only one of its isoforms is capable of binding NAD^+. The function of macroH2A most likely lies in inactive chromatin states, such is inactivated X chromosome, senescence-associated heterochromatin foci (SAHF) et cetera (Costanzi and Pehrson, 1998; Gamble et al., 2010; Tolstorukov et al., 2012; Zhang et al., 2005).

The **H2A.X** variant is defined by the presence of a carboxy-terminal amino acid motif, with a serine that could be phosphorylated. In the phosphorylated form this variant is known as ϒ-

H2A.X and is one of the best-characterized markers of double-strand breaks (Morrison and Shen, 2005).

In metazoans, the **H3.3** variant differs from H3 by only four to five amino acids (Filipescu et al., 2013), with three of them found in the histone fold domain, playing a crucial role in the replication independent deposition (Ahmad and Henikoff, 2002). H3.3 deposition occurs at highly dynamic nucleosomes and fills-in the gaps left by canonical H3, with the help of the HIRA chaperone (Schneiderman et al., 2012). A newly reported histone H3.3-specific chaperone (DAXX), along with the ATRX chromatin remodeler has been associated with incorporation of H3.3 at telomeric and pericentric heterochromatin (Drane et al., 2010; Goldberg et al., 2010). Interestingly, H3.3/H2A.Z occupy dynamic regulatory elements with high turnover (Jin et al., 2009), which is in line with the sensitive salt-dependent disruption of H3.3/H2A.Z nucleosomes (Jin and Felsenfeld, 2007).

Finally, the **CENP-A** protein (also known as CenH3) is the first discovered centromeric protein, and it was later shown to be a histone (Earnshaw and Rothfield, 1985; Palmer et al., 1991). CENP-A is a component of a nucleosome-like structure endemic for centromeres.

1.4.2 Histone H3 Lysine 4 methylation

Methylation of lysine 4 of histone H3 is one of the most widely studied histone PTMs. It was shown for the first time to be enriched in actively transcribed chromatin in Tetrahymena (Strahl et al., 1999) and later it was demonstrated that H3K4me3 is highly conserved from yeast to mammals (Shilatifard, 2008) and it is considered a landmark of promoters based on its distribution in the close vicinity of 5' of genes (Barski et al., 2007; Guenther et al., 2007) (Fig. 10). This modification is associated with active genes (Santos-Rosa et al., 2002) and transcriptional memory (Ng et al., 2003). In yeast, fruit fly and mammals, H3K4me3 is dependent on the activity of Set1 or Set1-related enzymes. There is a single Set1 enzyme in yeast, which belongs to the COMPASS complex, and sets all H3K4 methylation forms, whilst in fruit fly there are three and in mammals there are at least six Set1-related enzyme such as Set1A, Set1B and MLL1-4, from which only Set1A/B, MLL1 and MLL4 introduce trimethylation (Shilatifard, 2012).

The establishment of H3K4me3 around promoters is dependent on the presence of RNA pol II. The Paf1 complex has multiple roles in transcription and it is known to interact with the modified tail of RNA pol II. It serves as a landing pad for histone modifying activities around promoters, by interacting with Set1/COMPASS, Dot1 (H3K79 methylation) and Rad6/Bre (H2B monoubiquitylation) (Krogan et al., 2003; Wood et al., 2003). Based on these and other studies, it was concluded that Set1 acts downstream of Rad6 dependent H2B monoubiquitination (Dover et al., 2002; Schneider et al., 2005; Sun and Allis, 2002) making the methylation of lysine 4 dependent on H2B monoubiquitylation. H3K4me3 is also associated with the transcription machinery through its capability to recruit the basal transcription complex TFIID via its PHD finger domain (Lauberth et al., 2013). Aside from its association with the transcription machinery, there is an antagonism between H3K4me3 and H3R2me2a and H3R2me2a-setting enzymes and H3K4me3 and DNA methylation (Kirmizis et al., 2007; Rose and Klose, 2014). Mechanistically, H3R2me2a inhibits the Set1 enzyme activity (Kirmizis et al., 2007), while in the case of DNA methylation, H3K4me2/3 prevents the binding of Dnmt3a, 3b and 3l to H3K4me2/3 enriched chromatin. Collectively, H3K4me3 is a beacon for euchromatic states in fruit fly (Filion et al., 2010) and mammalian cells (Ernst et al., 2011).

Similarly to H3K4me3, H3K4me2 is also associated with actively transcribed chromatin states, such as promoters and enhancers (Barski et al., 2007; Ernst et al., 2011). The monomethylation

of lysine 4 has a completely different role from H3K4me3 and is diagnostic of poised or active enhancers (Calo and Wysocka, 2013; Ernst et al., 2011).

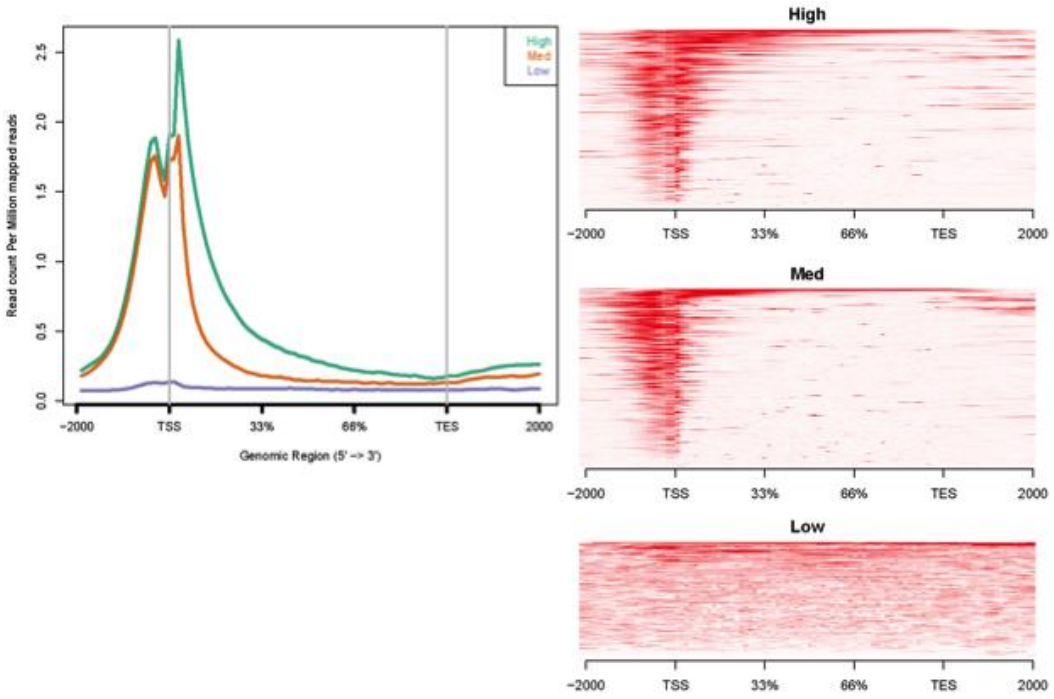

Figure 10. All known genes can be bioinformatically transformed into one metagene with uniform length. Then the cumulative signal obtained over all genes for a histone modification can be plotted on this metagene. Metagene profile of H3K4me3 distribution in gene bodies, showing enrichment around promoters (TSS) of expressed genes. On the left panel there is a composite profile of H3K4me3 distribution dependence on gene expression, and on the right panel the same data are shown in a heatmap format. The image is taken from https://code.google.com/p/ngsplot/.

1.4.3 Histone H3 Lysine 9 methylation

Trimethylation of lysine 9 of histone H3, has been historically the most studied histone mark due to its association with position effect variegation (Tschiersch et al., 1994) and later with the first discovered protein/histone lysine methyltransferase (PKMT) Suv39H1 (Rea et al., 2000). It is predominantly localized in centromeric, telomeric and pericentric stretches of repressed chromatin (Kouzarides, 2007). The first level of mechanistic understanding of H3K9me3 came about from studies, which tried to dissect the establishment of heterochromatin. It was found that the PKMTs, Suv39H1 and Suv39H2 in mammals and Clr4 in fission yeast, cooperate with an H3K9me3 binding protein termed Heterochromatin protein 1 (HP1) in mammals and Swi6 in fission yeast (Bannister et al., 2001; Lachner et al., 2001; Nakayama et al., 2001; Noma et al., 2001). According to this elegant model of heterochromatin maintenance, HP1 binds to H3K9me3 modified nucleosomes and recruits Suv39H, which can further methylate adjacent nucleosomes and continue the cycle of binding and recruitment. In addition, HP1 proteins can interact with one another via their chromoshadow domains, facilitating the spreading of heterochromatin (Fig. 11) and condensation of H3K9me3 tagged chromatin.

In fission yeast there is also a link between the RNAi machinery and H3K9 methylation (Martin and Zhang, 2005). In this case siRNAs are employed to target the RNA-induced transcriptional silencing (RITS) complex to genomic regions, which then binds to nucleosomes with its chromodomain and nucleates the establishment of heterochromatin. This chromatin state is

then propagated and maintained by the Clr4 PKMT and Swi6 H3K9me32 reader (Verdel and Moazed, 2005; Zofall and Grewal, 2006).

The differentially modified moieties of H3K9 are localized in distinct nuclear domains, which can be distinguished by their staining patterns or replication timings in mammals (Rice et al., 2003; Wu et al., 2005). H3K9me3 colocalizes within pericentric heterochromatin, overlapping with DAPI dense spots, while H3K9me2 localizes towards the periphery of the nucleus and around nucleoli. In contrast, H3K9me1 localizes inside euchromatic domain and it is enriched in active promoters (Barski et al., 2007).

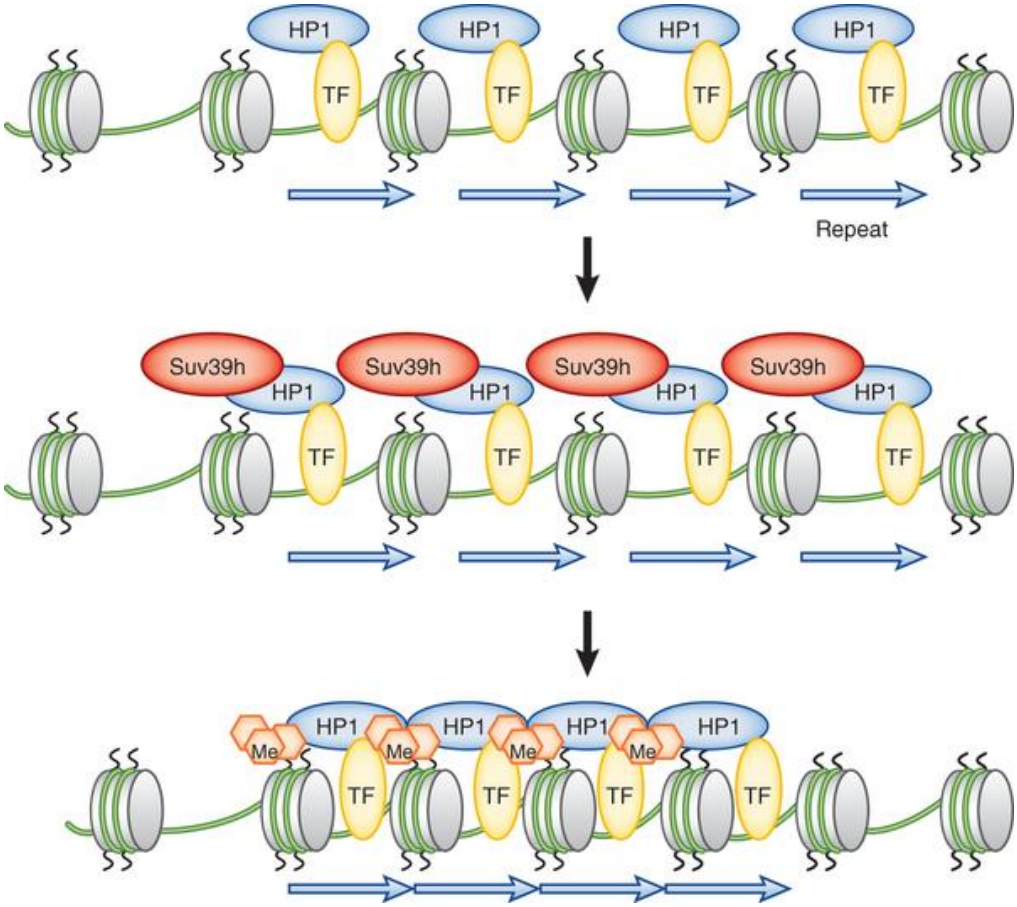

Figure 11. The HP1-Suv39H methylation pathway of heterochromatin maintenance. It is thought that HP1 is recruited to heterochromatin by a targeting protein. Subsequently, HP1 recruits Suv39H, which establishes the H3K9me2/3 modification, which in turn leads to HP1 docking and again recruitment of Suv39H to methylate the adjacent nucleosomes. The picture is taken from (Festenstein and Chan, 2012).

All these different flavours of H3K9 methylation are dependent on non-redundant methyltransferases. Suv39H1/2 are responsible for the majority of H3K9me3 and all methylation at constitutive heterochromatin (Rice et al., 2003), G9A/GLP are responsible for most of the H3K9me2 and some H3K9me1 (Peters et al., 2003), while Prdm3 and Prdm16 are responsible for cytoplasmic H3K9 monomethylation. The involvement of H3K9me3 in repressive chromatin types was supported by studies of chromatin segmentation in fruit fly and mammalian cells (Ernst et al., 2011; Filion et al., 2010).

1.4.4 Histone H3 Lysine 27 methylation

Trimethylation of H3K27 is a classical repressive mark associated with facultative heterochromatin, detected during X-chromosome inactivation (Payer and Lee, 2008; Plath et al., 2003) in mammals, and Hox gene clusters and polycomb domains in mammals and fruit

flies (Schuettengruber et al., 2007; Volkel and Angrand, 2007). The histone methyltransferase responsible for methylation of H3K27 is Enhancer of zeste (EZH), which along with suppressor of zeste 12 (Suz12), embryonic ectoderm development (EED) and RbBP4 builds one of the most famous epigenetic regulators of transcription, the polycomb repressive complex 2 (PRC2). This complex has profound effects on cell identity, differentiation and disease (Di Croce and Helin, 2013; Prezioso and Orlando, 2011; Schuettengruber et al., 2007). A simple recruitment model states that H3K27me3 recruits polycomb repressive complex 1 (PRC1), which has an H3K27me3 binding domain and an ubiquitin transferase activity specific for H2A. The recruitment of PRC1 leads to H2A ubiquitylation and chromatin compaction and further recruitment of PRC2 leading to H3K27me3 spreading. This model was challenged and it was suggested that noncanonical PRC1 complexes (lacking the H3K27me3 interacting module) first set the ubiquitin mark on H2A, which is followed by PRC2 recruitment and trimethylation of H3K27me3 to initiate a polycomb domain (Blackledge et al., 2014) (Fig. 12). The general overlap of H3K27me3 with polycomb domains was reaffirmed in chromatin segmentation studies in fruit fly and mammals (Ernst et al., 2011; Filion et al., 2010).

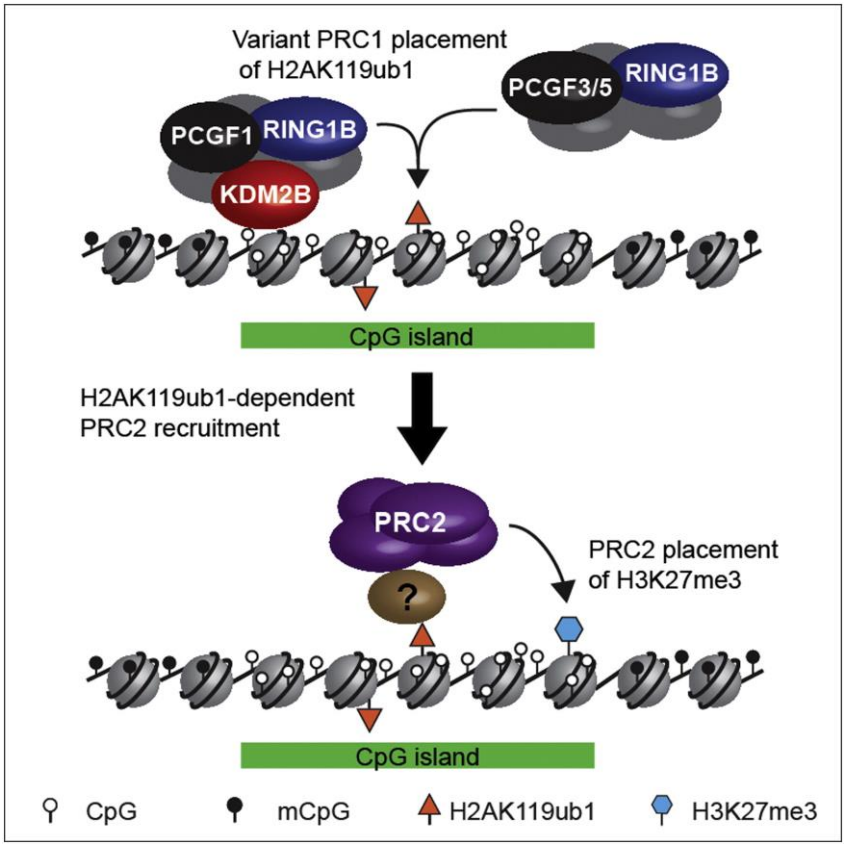

Figure 12. Mode of polycomb recruitment. A noncanonical PRC1 complex ubiquitylates H2A, which leads to PRC2 recruitment. PRC2 trimethylates H3K27, which might lead to canonical PRC1 recruitment via its H3K27me3 binding chromo domain. This mechanism of binding and setting H3K27me3 can further lead to maintenance of polycomb chromatin. The image is taken from (Blackledge et al., 2014).

Moreover, an interesting phenomenon in embryonic stem cells (ESc) is the presence of so-called “bivalent domains” containing H3K4me3 and H3K27me3. These seemingly opposite histone modifications are thought to poise developmental genes for differentiation. Namely, if H3K4me3 is retained, this will lead to activation or vice versa, if H3K27me3 is retained the gene will be silenced (Voigt et al., 2012).

A proteomics study showed that most of the H3K27 residues in the chromatin of mouse ES cells are dimethylated, with an average abundance of H2K27me2 of around 72%, H3K27me1 around 4% and H3K27me3 around 7% (Ferrari et al., 2014). The same study also suggested that EZH is responsible for all methylation states and unveiled the roles of H3K27me1 in transcription, controlled by the presence of H3K36me3 (when H3K36 is trimethylated, H3K27 cannot undergo di- and trimethylation), and the role of H3K27me2 in protection from aberrant H3K27 acetylation of enhancer regions.

1.4.5 Histone H3 Lysine 36 methylation

H3K36 methylation is a repressive mark enriched in the body of expressed genes, with the distribution of its mono-, di- and trimethylated forms shifting from the 5' to the 3' ends of genes (Barski et al., 2007). The histone methyltransferase responsible for H3K36 methylation is called Set2 and it associates with the elongating RNA pol II traversing along the body of actively transcribed genes (Krogan et al., 2003) By this means, the accumulation of H3K36me3 starts at the 3' of the first intron and proceeds, increasing towards the 3' of the gene (Kim et al., 2011).

The role of H3K36me3 could be in defining exons (Kolasinska-Zwierz et al., 2009) and in prevention of cryptic transcriptional starts in the gene body of expressed genes, which are decondensed and thereby accessible for the transcription machinery. The H3K36me3 mark can mediate silencing because it recruits the Rpd3S histone deacetylase complex via its Eaf3 binding domain (Carrozza et al., 2005; Joshi and Struhl, 2005; Wagner and Carpenter, 2012) (Fig. 13). Similarly, the H3K36me3 also serves as a binding platform for recruitment of PWWP domain harboring de novo DNA methyltransferases (Baubec et al., 2015; Dhayalan et al., 2010).

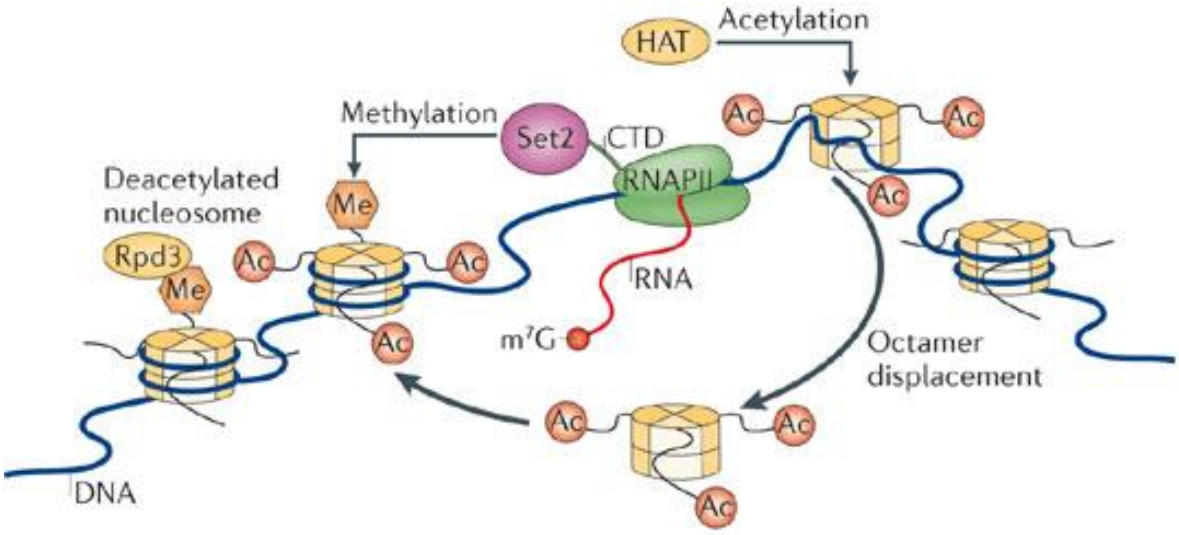

Figure 13. The function of H3K36me3. The elongating RNA pol II recruits Set2, which installs H3K36me3 on gene body nucleosomes. H3K36me3 recruits the Rpd3S complex, which deacetylates nucleosomes in order to prevent cryptic internal transcription. The image is adapted from (Wagner and Carpenter, 2012).

In a genome-wide profiling of chromatin states in fruit fly it was discovered that the chromo domain of MRG15 (homolog of the yeast Eaf3) binds H3K36me3 and it abundantly localizes in genes with broad expression in many cell types (housekeeping genes), while its levels are decreased in genes with specific and more narrow expression (cell type specific genes) (Filion et al., 2010). This indicates that there are at least two types of transcribed chromatin, with

different cellular functions. On a similar note, in a chromatin segmentation study in mammals it was corroborated that H3K36me3 is involved in transcriptional transition and transcriptional elongation (Ernst et al., 2011). Finally, it is thought that H3K36me3 and H3K27me3 are the main antagonists in the establishment of active or repressive chromatin states. This is based on the inverse correlation between H3K36me3 and H3K27me3 in animals and plants (Voigt et al., 2012; Yuan et al., 2011) and the inhibititory effect of H3K36me3 on the histone methyltransferase activity of PRC2 (Voigt et al., 2012).

1.4.5 Histone H3 Lysine 79 methylation

In contrast to the above-mentioned histone modifications, which occur on the N-terminal tails of the histone proteins, the H3K79 residue is located at the lateral surface of the nucleosome and it can be mono-, di- and trimethylated by a single enzyme, called Disruptor of telomeric silencing (Dot1). It was initially reported that H3K79 methylation plays a role in heterochromatin formation and telomeric silencing (Ng et al., 2003; van Leeuwen et al., 2002), DNA damage and cell cycle check-point (Giannattasio et al., 2005; Wysocki et al., 2005). Later, genome-wide localization studies have uncovered the presence of H3K79 methylation in gene coding regions as well (Kouskouti and Talianidis, 2005). Interestingly, H3K79 and H2BK123 lie in a close proximity (Luger et al., 1997) indicating a possibility for crosstalk. Consistent with this view, deletion of the H2B123 monoubiqutylase Rad6 or members of the Paf1 complex prevented H3K79 methylation (Briggs et al., 2002; Ng et al., 2003; Nguyen and Zhang, 2011; Sun and Allis, 2002) indicating that Dot1 works downstream from Rad6.

1.4.6 Histone H4 Lysine 20 methylation

The lysine 20 residue from histone H4 exists in three methylation states, which are associated with diverse processes, such as replication, DNA repair and heterochromatin (Balakrishnan and Milavetz, 2010; Jorgensen et al., 2013). The Set8 enzyme is responsible for H4K20 monomethylation in a cell cycle dependent manner, peaking at G2 (Pesavento et al., 2008), whilst the Suv4-20H1/2 are responsible for H4K20me2/3. The role of H4K20me2 is still elusive but it is likely to be involved in silencing and DNA repair (Jorgensen et al., 2013). The trimethylated form of H4K20 is strongly associated with constitutive heterochromatin and chromatin compaction, with HP1 serving as a recruiting factor for Suv4-20H1/2 (Evertts et al., 2013; Schotta et al., 2004).

1.4.7 Histone acetylation

Different from lysine methylation, acetylation of lysine residues neutralizes the charge of the side chain and thereby weakens the electrostatic contacts between the histone proteins and the DNA (Vettese-Dadey et al., 1996). It was shown in 1999 for the first time that histone lysine acetylation is associated with actively transcribed chromatin (Strahl et al., 1999) and numerous acetylation sites of histones H2A, H2B, H3 and H4 were subsequently discovered (Basu et al., 2009). The structural consequences of histone acetylation leading to open chromatin seem to be dependent on the overall cumulative effect of charge neutralization (Zentner and Henikoff, 2013). In line with this, histone acetylation is typically positively correlated with actively transcribed and dynamic chromatin. It has a short half-life, with rapid cycles of acetylation and deacetylation (Kurdistani et al., 2004; Pokholok et al., 2005). The acetylation reaction is performed by promiscuous histone acetyltransferases (HATs) such as Gcn5 and Esa1, which use acetyl-CoA as a donor of acetyl groups (Berndsen and Denu, 2008). The opposite reaction of deacetylation is performed by histone deacetylases (HDACs), which also display a low substrate specificity (de Ruijter et al., 2003). Apart from the cumulative effect of charge neutralization, histone acetylation in particular at H4K16 can have a direct effect on higher-order chromatin structure by inhibiting the formation of the 30-nm fiber (Horikoshi, 2013; Shogren-Knaak et al., 2006). In this case, the acetylation of H4K16 interferes with the

interaction between the unmodified tail of H4 with an acidic patch on H2A, which results in decondensation. Finally, histone acetylation can provide a binding surface for "reading domains" such as the bromo domains, which are part of *trans* acting factors such as HATs or RNA pol II (Josling et al., 2012; Kouzarides, 2007) (Fig. 14).

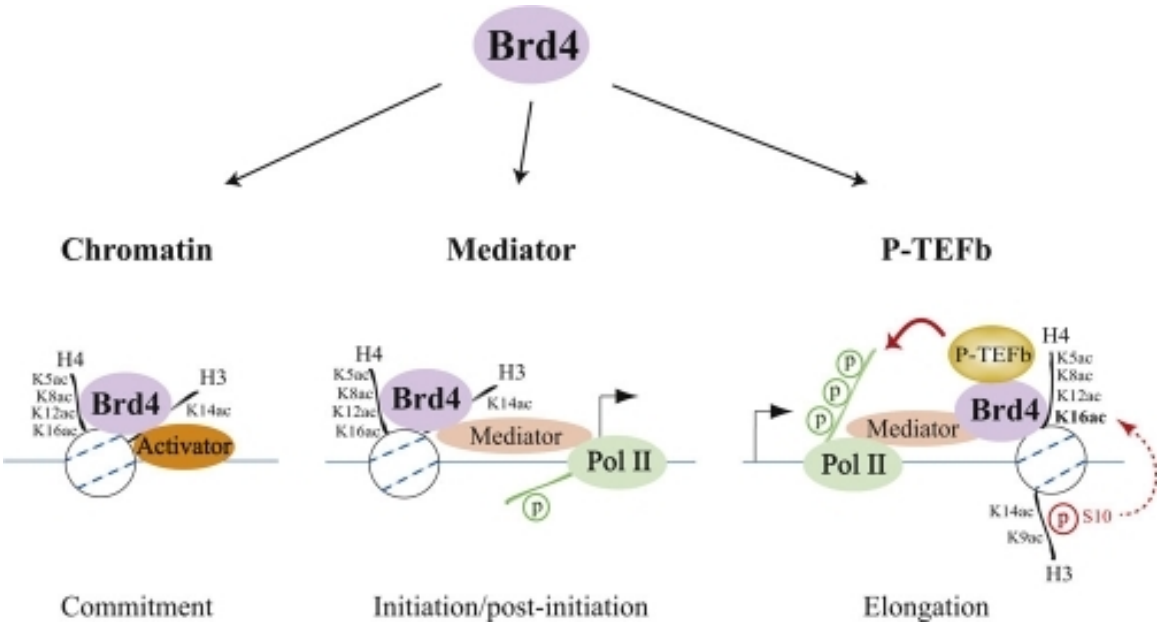

Figure 14. Representative example of recruitment of chromatin complexes to acetylated chromatin by a protein (Brd4) containing a bromo domain. The image is taken from (Chiang, 2009).

1.5 Readout of modified lysines

One of the most prominent roles of histone PTMs is their ability to serve as docking platforms for effector domains, so called histone modification interacting domains (HMIDs) or "reading domains" and their associated complexes. A compilation of HMIDs has been reported to bind to modified or unmodified histone tails (Taverna et al., 2007). The first reader of histone PTM was the bromo domain from PCAF, which recognizes acetyl lysine modifications (Dhalluin et al., 1999). The structure of this domain in complex with an acetylated H4-tail peptide showed a left-handed antiparallel four-helix bundle with a deep and narrow hydrophobic pocket with a hydrogen-bonding capacity, which accommodates the acetyllysine (Owen et al., 2000). Binding specifcitiy is provided by a key asparagine residue forming hydrogen bond with the acetyl group (Mujtaba et al., 2007). Bromo domains are by far the most common acetyllysine interactors, but sometimes PHD and Tudor domains can also bind acetyl marks (Qiu et al., 2012; Jurkowska and Jeltsch, personal communication).

Histone lysine methylation is the most diverse and versatile histone PTM, and as stated above it can be found in a monomethylated, dimethylated or trimethylated state. All methylated lysines are cationic at physiological pH. With incremental addition of methyl moieties, the hydrophobicity increases and the ability for hydrogen bonding decreases, leading to differential biophysical properties that enable a specific readout (Taverna et al., 2007). Most of the methyllysine readers belong to the so-called Royal family of domains (chromo, chromo-barrel, PWWP, MBT, Tudor) (Musselman et al., 2012; Taverna et al., 2007) (Fig. 15). The most notable structural trait of these methyllysine readers is an aromatic pocket, typically consisting of 2-4 aromatic residues, positioned perpendicular to one another, optimal for formation of

cation-π and van der Waals interactions with the methyllysine □-amino group. Another similarity in their topology is the presence of β-strands, which in turn form an incomplete β-barrel. The histone peptide completes the β-barrel structure by providing an extra β-strand. This induced fit is stabilized through side-chain hydrogen bonds and interactions, which impart specificity. This binding mode is typical of heterochromatin protein domains HP1α (CBX5), HP1β (CBX1) HP1ϒ (CBX3) and polycomb Pc protein domain such as CBX2, CBX4, CBX6 and CBX7. The presence of negatively charged amino acid on the walls of the aromatic cage can lead to a preference for lower methylation states (this could be also achieved with a smaller pocket) (Bannister et al., 2001; Fischle et al., 2003; Jacobs and Khorasanizadeh, 2002; Min et al., 2003; Musselman et al., 2012; Nielsen et al., 2002; Taverna et al., 2007). Additionally, based on the binding modes, reading domains can either bind their target peptide via surface groove recognition mode (such as HP1 and binders of higher methylation states) or cavity insertion mode, where the methyllysine peptide is inserted within a deep protein cleft (typical for binders of lower methylation states of lysines).

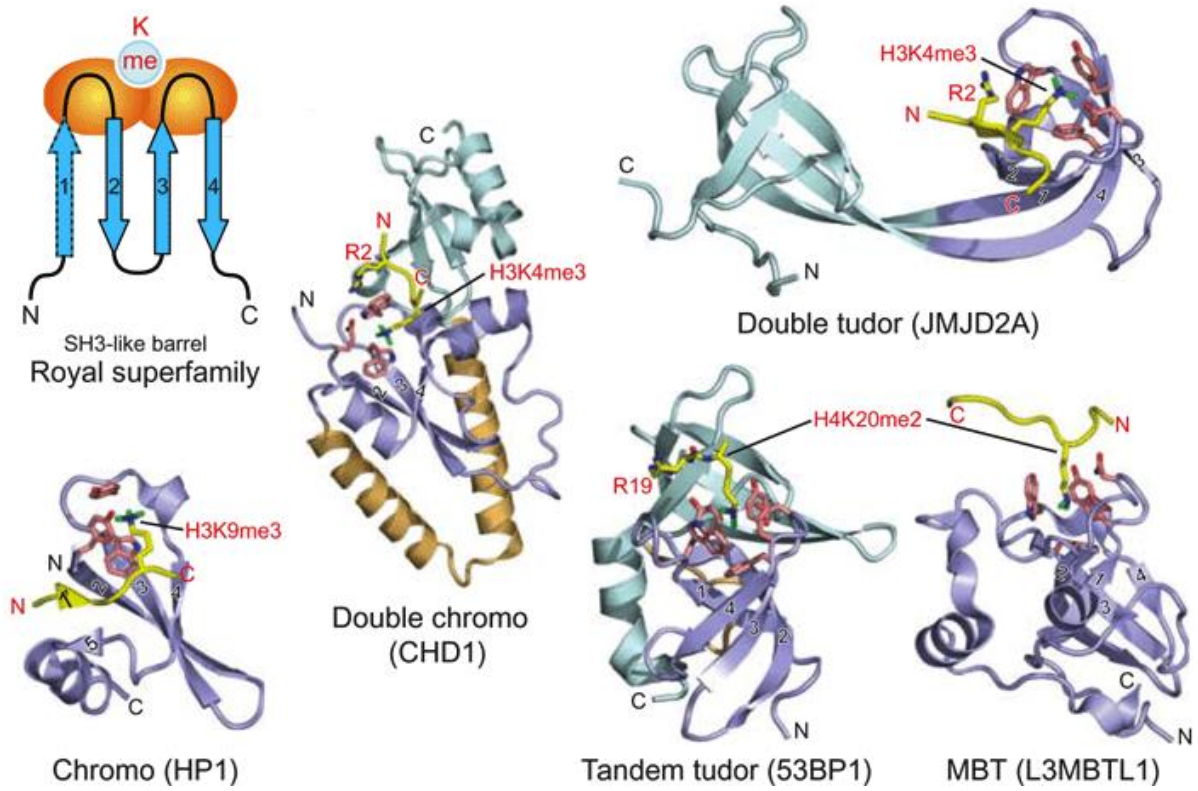

Figure 15. Readout of methyllysine marks by Royal-family modules. The image is adapted from (Taverna et al., 2007).

Plant homeodomain (PHD) containing proteins are another big family of reading domains, whose Cys_4-His-Cys_3 motif is stabilized by zinc ions. They most commonly interact with H3K4 methylated or unmodified peptides, which typically dock onto the surface of the domain and extend the β-sheet of the PHD interior. The N-terminus of the histone peptides forms extensive contacts, but the binding is again mediated by aromatic caging of the trimethylammonium group (Patel and Wang, 2013). Additionally, phosphorylated serines can also be recognized by 14-3-3 proteins, which mediate different downstream effects (Macdonald et al., 2005).

Based on the flurry of structural data, it has become clear that small chemical modification can have immense influence on chromatin by recruiting chromatin interacting complexes or

modifiers. For instance, the H3K4me3 interaction with TAF3 PHD regulates activation and preinitiation complex assembly (Lauberth et al., 2013); the PHD finger domain from ING2 binds to promoters of proliferation genes during upon DNA damage (Shi et al., 2006); and CBX domains bind and recruit the constitutive or facultative heterochromatin machinery. Further structural studies should provide further insight into the function and binding modes of many other reading domains.

1.6 Methods for studying histone modifications

After describing the nature and biological role of histone modificaitons, this chapter will provide a summary of the methods available to study the distribution and function of histone modifications. Also, additional methods, which allow for generation of the required substrates for a biochemical investigation of histone modifying enzymes or readers will be outlined.

1.6.1 Chromatin Immunoprecipitation

The most popular and prevalent method for studying histone modifications is chromatin immunoprecipitation (ChIP) of histone PTMs, which can be used to directly pinpoint the genomic location and distribution of a target histone modification. In a typical ChIP experiment, either micrococcal nuclease (MNase) digested chromatin (native ChIP) or chemically crosslinked chromatin (X-ChIP) sheared by sonication is used as a substrate in a following immunoprecipitation step using an antibody, which specifically binds the modification in question (Gilmour and Lis, 1985). Upon chromatin precipitation, the underlying DNA is purified and analyzed by quantitative PCR (qPCR), DNA microarrays (chip) or massively parallel sequencing (seq). ChIP-qPCR can be employed to quantify the presence of the target histone PTM at selected genomic sites, whilst ChIP-chip or ChIP-seq provide a genome-wide overview of the studied histone PTM (Fig. 16). The distribution and functional roles of many histone modifications have been unveiled by ChIP-seq studies (Barski et al., 2007). The antibody used in the precipitation step has a pivotal role in the validity and reliability of the experiment, since it is the major liaison translating the language of histone modifications into observable properties that can be studied. Therefore, almost all we know about histone modifications has been passed through an antibody as a middleman. Indeed the quality and specificity of the used antibody is of a paramount importance, and proper quality control is an absolute requirement (Bock et al., 2011a; Egelhofer et al., 2011).

1.7.2 Peptide arrays and designer chromatin for studying histone modifying enzymes and readers *in vitro*

Due to the unstructured nature of the N-terminal tails of histones, short peptides reflecting portions of the histone tail have been particularly useful and practical substrates for *in vitro* biochemical dissection of histone mark functionalities. Historically, researchers have used peptide pull-down methods, dot blots and peptide arrays for characterization of antibodies, reading domains and histone modifying enzymes. All of these methods are conceptually related but differ in the number of peptide probes that are used. Peptide pull-down and dot blot experiments have a limited throughput and are often conducted using only a few single modified peptides. Given the combinatorial diversity of natural chromatin modifications, this is not sufficiently informative. Therefore, the application of **peptide array** based methods has been emerging as the method of preference for studying histone binding proteins and enzymes (Bock et al., 2011a; Bock et al., 2011b; Bua et al., 2009; Dhayalan et al., 2010; Garske et al., 2008; Garske et al., 2010; Rathert et al., 2008; Su et al., 2014; Zhang et al., 2010). Peptide arrays offer remarkable combinatorial power, which can be put into use for studying the effects of neighbouring secondary or tertiary histone marks of protein reading domains or antibodies

(Bock et al., 2011a; Bock et al., 2011b). This advantage is also related to the fact that the costs for the synthesis of one peptide are much lower on peptide arrays, as compared to peptide synthesis by conventional methods.

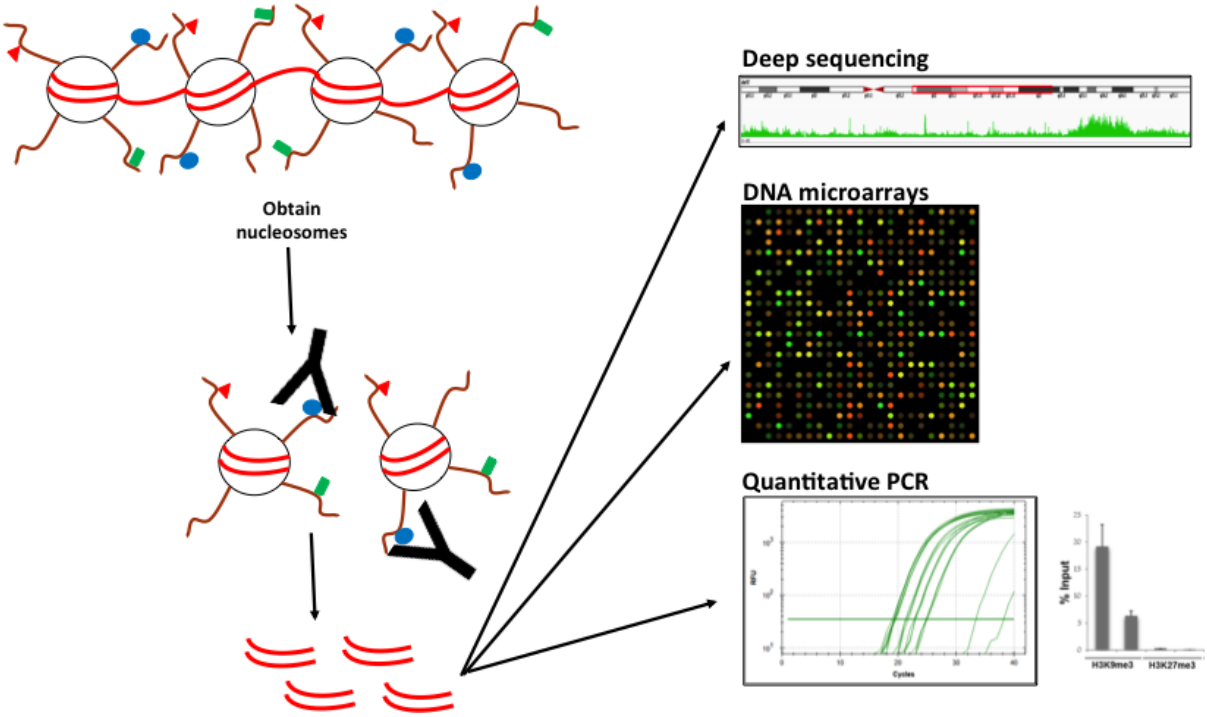

Figure 16. . The principle of native chromatin immunoprecipitation (native-ChIP) coupled with qPCR, DNA microarrays or deep sequencing. The image is adapted from (Kungulovski, submitted 2015).

Although preferential, the mechanistic studies of histone modifications and their effectors in native chromatin systems have been a daunting task due to the high complexity and heterogeneity of the histone code. Therefore, approaches pertaining to reconstituted systems of nucleosomes have been developed. They are well suited for *in vitro* biochemical investigation by virtue of their reduced complexity and defined configuration. Nucleosome core particles can be reconstituted with strong positioning DNA sequences and recombinantly expressed histones via salt gradient exchange (Luger et al., 1999). Recombinantly purified histones are unmodified by default and numerous methods have been developed for site-specific, defined introduction of histone modifications or their mimics into full-length histones. Reconstituted in this way, the nucleosomes with defined modifications can be used to study histone modifying enzymes or histone binding proteins.

1.6.2.1 Native chemical and expressed protein ligation for designer nucleosomes

Chemical synthesis of peptides allows the specific incorporation of modified amino acid residues at any site. However, since the efficiency of chemical peptide synthesis declines for peptides longer than 50 amino acids, this approach cannot be applied for synthesis of longer proteins, including histones. At the same time the recombinant production of (unmodified) proteins is highly efficient for many proteins, especially histones. To introduce modified amino acids into larger proteins (histones) these approaches can be combined and modified peptides generated by peptide synthesis can be attached to unmodified protein chassis, expressed and purified recombinantely. **Native chemical ligation (NCL)** (Dawson et al., 1994) and **expressed protein ligation (EPL)** (Muir et al., 1998) are two complementary methods, which have been prevalently used for such site-specific introduction of histone modifications. In more

detail, NCL is used to connect two peptide fragments with a native peptide bond, one harbouring a C-terminal thioester and the other one having an N-terminal cysteine residue, (any of these can bear the modification of interest). The strategy employed by NCL is useful for the synthesis of small, modified proteins, but it can be challenging for proteins longer than 100 amino acids. To overcome these limitations, the expressed protein ligation (EPL) method has been developed. In this approach a bacterially expressed protein is used as a fusion partner to the modified peptide. One of the most critical aspects of this method is to obtain the C-terminal thioster functionality. This can be achieved by thiolysis of a C-terminally fused intein, which will introduce a C-terminal thioster moiety to the expressed protein of interest that can later react with the N-terminal cysteine of a modified peptide, therefore introducing the desired post-translational modification into the full length protein. This is especially useful in chromatin research, owing to the fact that most of the modifications are concentrated on the short N-terminal tail of the histone protein, which can be easily synthesized as a peptide and ligated to the C-terminal part of the rest of the histone protein. Desulfurization of the cysteine residue to alanine at the expressed protein-peptide intersection can render the histone protein almost scarless (Yan and Dawson, 2001).

1.6.2.2 Introduction of histone modification analogues

Although NCL and EPL based strategies have the potential introduce any post-translational modification into any protein, they demand specialized equipment and appropriate chemical expertise, thwarting broader application of this technology. As a result, a new chemical approach was developed to install **site-specified analogues of methylated lysines** into recombinant proteins (Simon et al., 2007). This method is affordable and allows for rapid generation of large quantities of modified histones, which are chemically and functionally similar to their native counterparts. The introduction of a methyl-lysine analogue (MLA) is dependent on the reactivity of the thiol group of cysteine, which can be alkylated to mono-, di- and trimethylated aminoethylcysteine. The resulting analogues differ from the corresponding modified lysine moiety only in the γ-carbon, which is replaced by a sulphur atom. Recently, the alkylation-based strategy has been expanded to acetyllysine analogues (Huang et al., 2010), methylarginine analogues (Le et al., 2013) and serine modification analogues such as glycosylation and phosphorylation (Chalker et al., 2012). The only requirement for this method is replacement of the targeted amino acid to cysteine and exchange of the all natively occurring cysteine (e.g. C110 in human H3) to alanine via site-directed mutagenesis.

1.6.2.3 Introduction of histone modifications by expansion of the genetic code

Defined post-translational modifications can be introduced to any protein by means of **genetic code expansion**. This is possible by introduction of orthogonal pairs of tRNA-synthetase and tRNA in *Escherichia coli*, which have been evolved to incorporate unnatural or modified amino acids into any protein by utilizing a re-purposed stop codon. This scheme has been brought into effect to incorporate acetyllysine (Neumann et al., 2008) or derivatized methyllysine into recombinant histones. The derivatized methyllysine then can be deprotected upon purification, to yield native mono- or dimethyllysine (Ai et al., 2010; Nguyen et al., 2009; Nguyen et al., 2010).

1.6.2.4 Barcoded reconstituted nucleosomes

Recently, a novel exciting method based on **barcoded reconstituted nucleosomes** was used for detailed examination of recruitment and modulation of histone modification binding domains and enzymes (Nguyen et al., 2014). The conceptual catch of this method is the streamlined production of several different modified reconstituted nucleosomes, all of which bear a unique DNA barcode appended to their positioning sequence that reflects the specific modification state of the assembled nucleosome. These differently modified nucleosomes are then mixed and used for pulldown with a chromatin-binding complex. The DNA-based readout

of this in vitro ChIP method allows for a very sensitive detection of binding affinities and preferences of "reader domains" and antibodies coupled with low material expenditure.

1.6.3 Mass spectrometry

Mass spectrometry (MS) is one of the most powerful, sensitive and versatile analytical techniques that enable an unbiased study of histone marks (Britton et al., 2011). Affinity reagents require *ad hoc* generation and *a priori* knowledge of the targeted histone modifications, whilst MS has the ability to identify completely novel modifications or novel combinations of modifications simultaneously co-occurring on the same nucleosome, taken that sufficient material is provided.

1.6.3.1 Bottom-up, middle-down and top-down mass spectrometry in histone analyses

The traditional method of mass spectrometric study of histone modifications is the bottom-up approach, which revolves around SDS-PAGE separation of native histones and in-gel digestion of the histone of interest with a suitable protease, typically trypsin. Trypsin digests the protein into small peptides and this is preferably followed by liquid chromatography separation and mass spectrometric analyses. The downside of using trypsin for enquiries into the histone code is the vast abundance of arginines and lysines in the histone proteins (trypsin cleaves randomly after unmodified lysines or arginines), which results in a high complexity pool of heterogeneous small peptides. This makes the identification, quantification and association of distinct histone marks a difficult task. Derivatization of the histone protein with propionic anhydride (which propionylates unmodified and monomethylated lysines) prior to trypsinization can alleviate some of the problems and reduce the complexity of the peptide pool (Voigt et al., 2012). Although helpful, this derivatization strategy does not fully solve the difficulty of studying co-occurring histone marks, which in fact, is one of the most outstanding questions in chromatin biology. To do this, the histone modifications must be physically contained on the same peptide, which can be assisted by generation of larger peptides or full proteins. Top-down MS makes use of intact proteins and ensures complete connectivity of PTMs on the nascent protein. Middle-down represents a trade-off between bottom-up and top-down approaches and yields larger protein sizes than bottom-up. This is achieved by the application of endoproteases such as GluC and AspN, which typically cleave histone proteins generating lower complexity fragments (Karch et al., 2013). All three approaches have pros and cons, with bottom-up being the most sensitive one and technically the easiest to perform, while the top-down and middle-down approaches offer better co-association of modifications, but with reduced sensitivity and more demand for sophisticated instruments and expertise (Fig. 17).

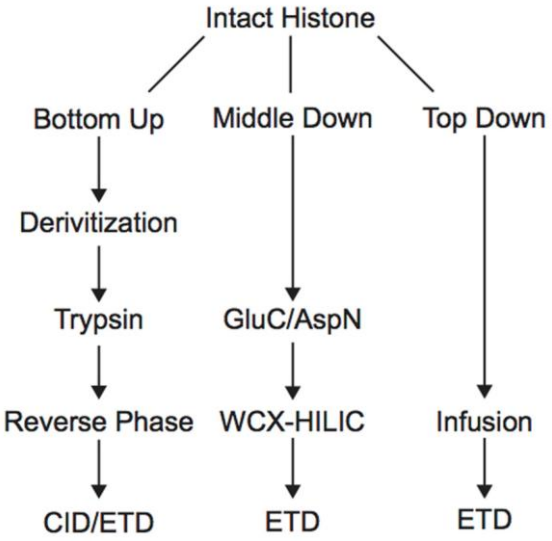

Figure 17. Different mass spectrometry approaches for studying histone modifications. CID and ETD are different methods for fragmentation of peptides. The image is taken from (Karch et al., 2013).

1.6.3.1 Affinity purification of chromatin states coupled to mass spectrometry

Recently, two groups employed ChIP-like approaches to investigate different instances of the "chromatome" (Soldi and Bonaldi, 2013; Su et al., 2014). Namely, they used antibodies or "reader domains" respectively, to precipitate chromatin and instead of deep sequencing the nucleosomally bound DNA, they carried out mass spectrometric profiling of the pulled nucleosomes and discovered new heterochromatin signatures.

Lately, numerous labs have started addressing the idea of locus specific characterization of chromatin states, where DNA or single locus specific factors could be used as handles to purify the full local chromatin state of the targeted locus. The pioneering study utilized a method called PICh (Proteomics of Isolated Chromatin) (Dejardin and Kingston, 2009) in which nucleic acid based baits were used to precipitate and probe the local chromatin interactome of telomeres in Drosophila melanogaster. Another milestone was achieved with the establishment of ChAP-MS, where yeast cells were genetically engineered with a Lex-A binding site upstream of the GAL1 promoter, and then expressed affinity tagged Lex-A protein was used to purify the GAL1 promoter interactome (Byrum et al., 2012). To overcome the necessity for genetic engineering of the targeted locus, the same group (Byrum et al., 2013) implemented a TALE based system to purify the GAL1 interactome. However, these approaches demand immense amount of starting material, and only time will tell whether they can be reliably applied in mammalian systems.

1.6.4 *In vivo* tracking of histone modifications

With the development of novel imaging techniques and the strong pace set by synthetic biology it has become possible to study the dynamics of chromatin states within living cells. A handful of studies have set the foundation for *in vivo* imaging of chromatin, either based on Förster resonance energy transfer (FRET), split protein complementation or fluorescently labeled minimal antibodies. Most of the studies that have employed FRET and complementation based sensors, exploited the specific interaction between "reader domains" and histone modifications in different settings: first, single chain fusion of reader and histone tail were tagged with fluorescent proteins on the C- and N-termini, which upon modification of the histone tail resulted in a shift in FRET (Lin et al., 2004; Lin and Ting, 2004). Second, fluorescently labeled histones and "reader domains" either with full fluorescent protein (FRET upon interaction) or split fluorescent protein (fluorescence upon interaction) were used to directly probe the dynamics of histone modifications or their spatial distribution in live cells, respectively (Ito et

al., 2011; Kanno et al., 2004; Kerppola, 2008; Sasaki et al., 2009). With the development of recombinant and minimal size antibodies (Hattori et al., 2013; Hayashi-Takanaka et al., 2009; Hayashi-Takanaka et al., 2011; Sato et al., 2013) similar approaches are anticipated for development.

2. Principal aims of the study

The post-translational modifications of histones and DNA methylation are one of the most prominent regulatory hubs of chromatin. Their precise genome-wide distribution, association with genomic elements and structures, dynamics of establishment and loss and their functional roles are pressing scientific issues. In the present body of work, firstly we wanted to develop and apply novel investigative tools for studying the genome-wide distribution of histone PTMs. Secondly, we wanted to establish and apply an approach based on zinc finger targeted promoter methylation in order to study the *in vivo* dynamics of histone PTMs and DNA methylation.

2.1 Development and application of histone modification interacting domains (HMIDs) as an alternative to antibodies

Histone PTMs constitute the most diverse indexing mechanism of chromatin. It is of paramount importance to properly elucidate their involvement in the corresponding chromatin-templated regulatory processes. Most of the contemporary experimental approaches are based on mass spectrometry or specific precipitation of histone PTMs with antibodies. Mass spectrometry is a very powerful qualitative and quantitative approach, but so far it has been unable to precisely associate a particular histone PTM with an underlying genomic region. The single experimental tool that could dissect the association of particular genomic loci with their underlying modified histones is based on antibodies. Although they have a central role in chromatin research many antibodies suffer numerous drawbacks, such as insufficient specificity and lot-to-lot variability, which highly undermines the reliability of scientific research. For this reason, we wanted to develop and experimentally validate a cheaper and more reproducible affinity reagent. More precisely, in this work we aimed to test the ability of isolated recombinant histone modification interacting domains (HMIDs) to detect and precipitate distinct types of chromatin (Fig. 18).

2.2 Guidelines for quality control of affinity reagents

The affinity reagent is the middleman between a genuine presence of a histone PTM inside the cell and the data we observe with state-of-the-art molecular biology approaches. On these grounds, proper and rigorous quality control and extensive documentation is a pre-requirement for any affinity reagent. In this work, we aimed to review the current guidelines for antibody quality control set forth by the ENCODE consortium (Egelhofer et al., 2011; Landt et al., 2012) and with the knowledge obtained from (Kungulovski et al., 2014) we modified and updated them in a comprehensive manner.

2.3 Targeted epigenome editing of an endogenous locus with chromatin modifiers

The chief role of epigenetic modifications in many essential cellular processes, along with the plasticity and reversibility of epigenetic states, has inspired the scientific community to focus its efforts on approaches to locally rewrite epigenetics states. Genome-wide studies have provided incredible insight into the innerworkings of chromatin, but most of these data are correlative and do not directly pinpoint the effects, consequences and associations of different epigenetic marks to one another or with other cellular processes, such as transcription.

Moreover, the question of whether histone PTMs are cogs or causes remains hotly debated (Henikoff and Shilatifard, 2011).

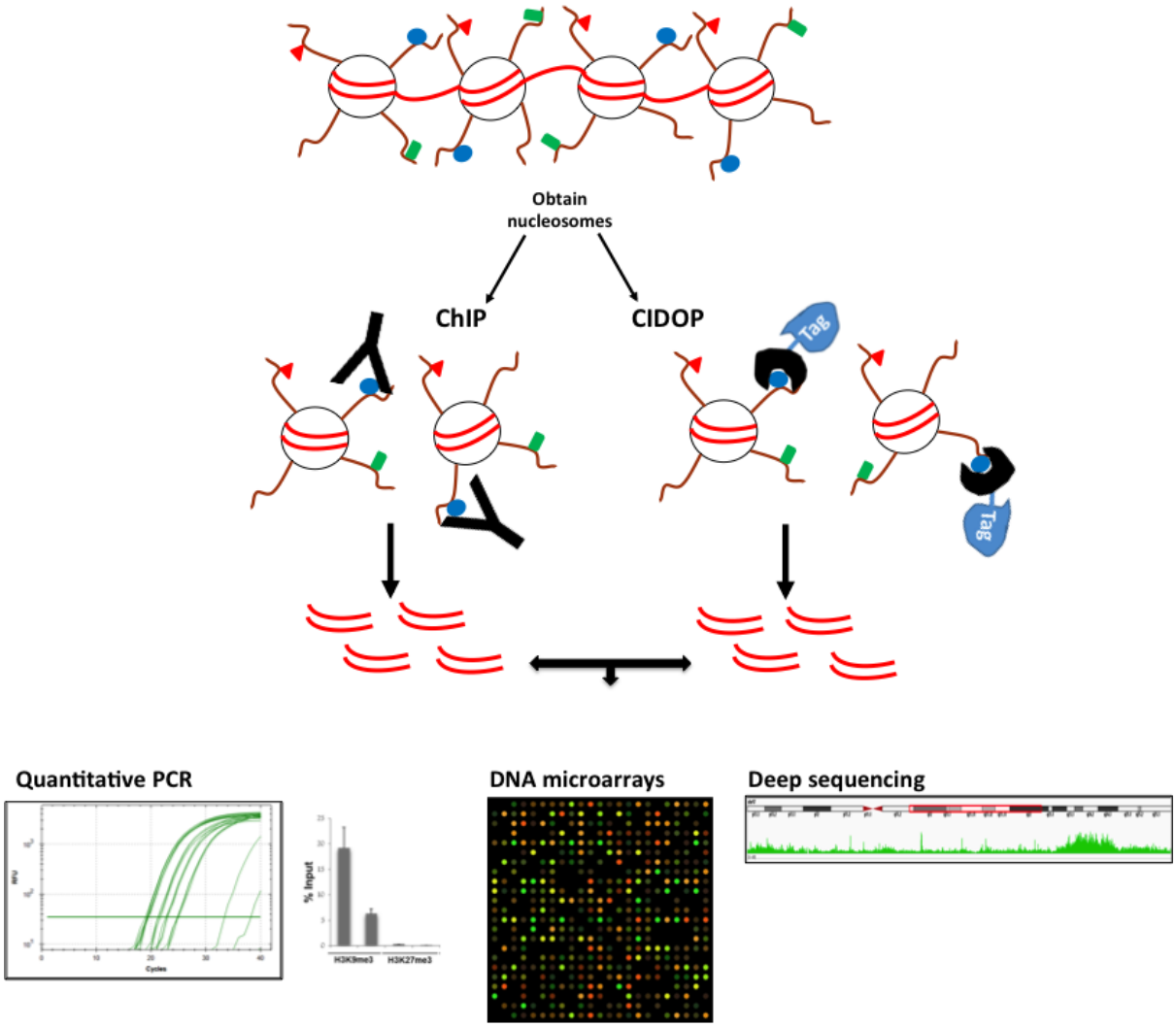

Figure 18. The basic principle of Chromatin ImmunoPrecipitation (ChIP) and the alternative ChromatIn Domain Precipitation (CIDOP) that we aimed to develop in this study. We wanted to test both affinity reagents (antibodies or HMIDs, respectively) to precipitate specifically modified nucleosomes, and then analyse the underlying DNA with approaches such as qPCR, DNA microarrays or deep sequencing. The image has been adapted from (Kungulovski, submitted 2015).

In this study we aimed to elucidate some of the principal mechanisms by which epigenetic modifications operate, by fusing a DNA targeting module (zinc finger that binds the promoter of *VEGF-A*) with a catalytic domain that harbors a defined epigenetic activity (DNA or H3K9 methylation) and introduce it in SKOV3 cells via adenoviral vector (Fig. 19). By these means we wanted to study the dynamics of establishment, effect and loss of epigenetic marks.

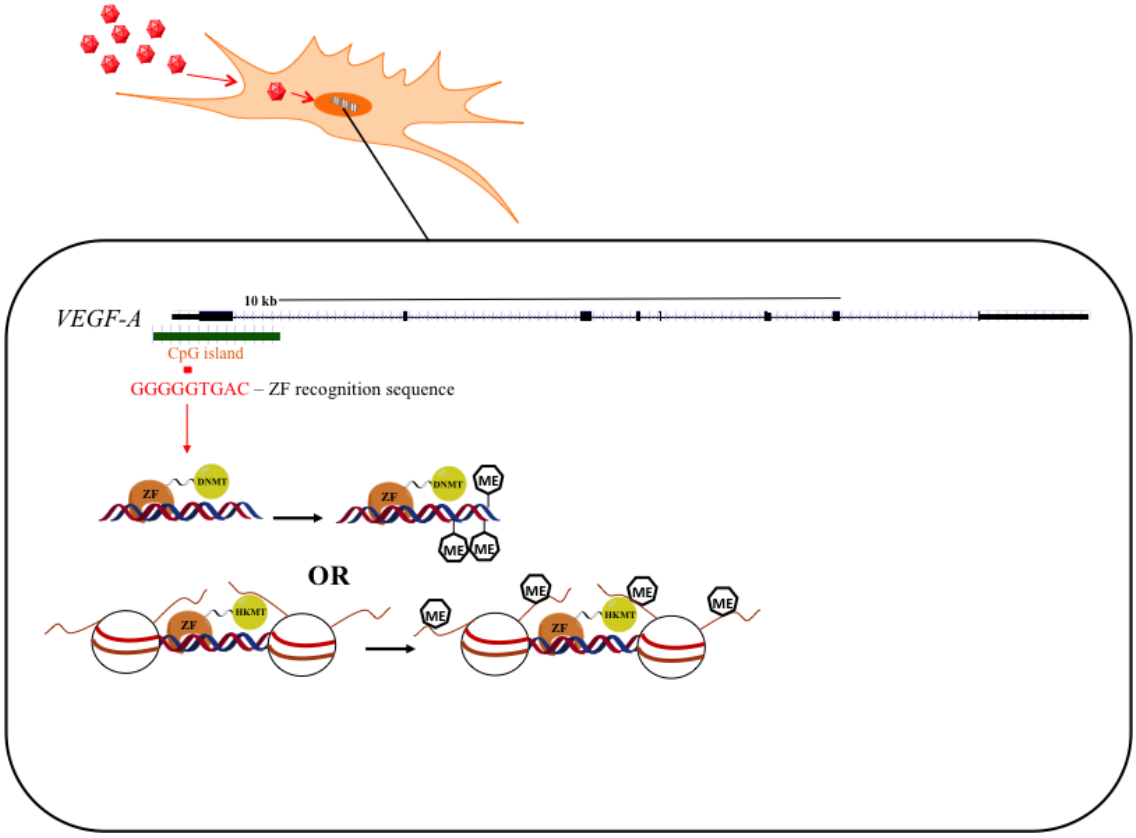

Figure 19. Schematic overview of the strategy that we aimed to use in this study. The image has been adapted from (Kungulovski et al., 2015).

3. Materials and methods

3.1 Application of HMIDs as an alternative to antibodies

3.1.1 Cloning, site-directed mutagenesis and protein purification

Cloning was performed in pGEX-6p-2 (GE Healthcare) vectors using standard ligase dependent molecular biology methods and the site-directed mutagenesis was performed as described in (Jeltsch and Lanio, 2002). The human TAF3 PHD (932-1004) construct was cloned using BamHI and XhoI cloning sites. All GST- tagged proteins were overexpressed and purified essentially as described in (Rathert et al., 2008). For more details regarding cloning sites, sequences et cetera, refer to the Method's section in the manuscript in Appendix 4.

3.1.2 Peptide array analyses

The peptide array was incubated with the GST-tagged protein of interest for 2 hours at room temperature, followed by incubation with primary goat anti-GST antibody and secondary anti-Goat-HRP conjugated antibody. The signal was detected by chemiluminescence. The protocol is described in extensive detail in the manuscript added in Appendix 2 and (Bock et al., 2011b).

3.1.3 Cell culture

HEK293 cells were grown in Dulbecco's modified Eagle's medium with 5% fetal bovine serum 37°C at 5% CO_2 until they reached 90% confluence. HepG2 cells were grown in Dulbecco's modified Eagle's medium with 10% fetal bovine serum at 37°C and 5% CO_2 until they reached 75% confluence. Immortalized MEFs were grown in DMEM supplemented with 10% FBS, L-glutamine, nonessential amino acids, sodium pyruvate, and 100 mM b-mercaptoethanol. For more details regarding inhibitors refer to the Method's section in the manuscript in Appendix 4.

3.1.4 Western and Far Western blot

Native histones were isolated by acid extraction (Shechter et al., 2007) from HEK293 or *Saccharomyces cerevisiae*. Recombinant histones were purchased from New England Biolabs. For more details regarding the protocol and antibodies refer to the Method's section in the manuscript in Appendix 4.

3.1.5 Native chromatin precipitation (ChIP and CIDOP) and quantitative PCR (qPCR)

Nucleosomes were isolated from around 20 million HepG2 cells by micrococcal (MNase) digestion of nuclei and were precipitated with the HMID or antibody of interest. The quantitative PCR assays were performed on a CFX96 Touch or CFX96 Real-Time detection system with SybrGreen or EvaGreen dyes and the percent of precipitated target genomic locus was calculated based on a standard curve. For more details regarding antibodies, HMIDs, concentrations, washing conditions, elution and qPCR analyses refer to the Method's section in the manuscript in Appendix 4.

3.1.6 Deep sequencing and detailed bioinformatic analyses

The ChIP-seq or CIDOP-seq samples were sequenced at the Active Motif or Max Planck-Genome-Center Cologne (MP-GC) facillities. Approximately 20-50 million, 50- or 100-nt sequence reads in a FASTQ format were obtained with Illumina's HiSeq 2000 or 2500 genome sequencer. For more details regarding the specifications and datasets used for analysis refer to the Method's section of Appendix 4. Below are given the explanations and guidelines pertaining to the more general pipeline of sequencing data analysis.

3.1.6.1 Quality control of FASTQ sequences

The modern DNA sequencers can provide millions of raw sequences in one run. The first part of the analysis workflow was to check the quality of the obtained sequences before drawing any biological conclusions. All raw sequences were obtained in a FASTQ format (which is a text-based format for storing nucleotide sequences and their corresponding quality scores) and underwent quality control analysis with the FastQC software. FastQC has a user-friendly graphical interface and the analysis is relatively straightforward. In order to open one or more sequences, the program was ran and File > Open was selected. The files appear on the top of the screen. Typically the size of the files is more than a couple of gigabytes, and it might take a couple of minutes to open them. If more files are opened at the same time, they will be queued and successively processed. On the left hand side there is a main interactive display, which can help to quickly visualize the quality of the data (green tick, orange triangle and red cross indicating normal, slightly abnormal or very unusual data, respectively) (Fig. 20). These suggestions are to be taken in the context of what is expected from the data and in light of this, the intrinsic biases that might be expected. The reports can be saved in File>Save report, containing all the necessary information and graphs. For more information about the modules refer to http://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/. FastQC can be downloaded from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

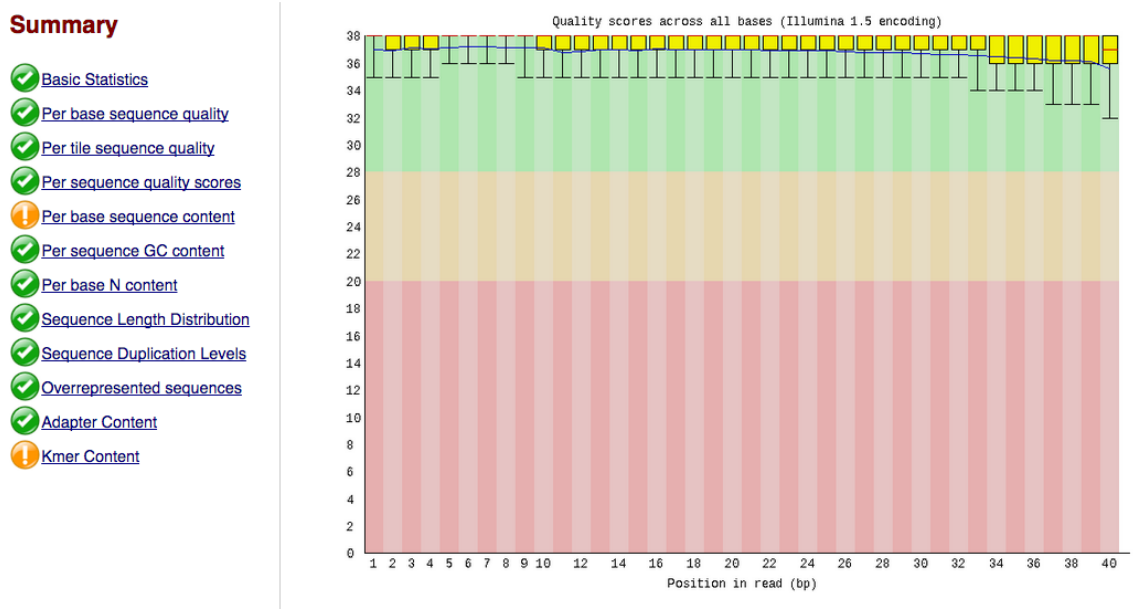

Figure 20. Representative example of a quality report of good data. The interactive display is shown on the left hand site, and an example of per base sequence quality is shown on the right hand side. On the Y-axis the Phred score is labelled (20 = 99% base call accuracy, 30 =99.9% base call accuracy etc) and on the X-axis an example of 40-bp long read is given.

3.1.6.2 ChIP- sequencing: mapping, peak calling, genome coverage maps, annotation to genomic elements, K-means clustering

Mapping: upon confirmation of the quality of data, the next step was mapping of the raw sequences to a reference genome. The reads we obtained were mapped to the human reference genome Hg19 (also known as GRCh37) with the Chipster assembly of bioinformatic tools for high-throughput data analysis (Kallio et al., 2011). First, the data were imported and aligned with Bowtie for single reads (Langmead et al., 2009) with the following parameters:

- Genome: [human genome Hg19]
- Number of mismatches allowed (0, 1, 2, 3): [2]
- Consider mismatches only in the seed region (yes, no): [no]
- Length of the seed region (5-50): [28]
- Allowed total of mismatch qualities (10-100): [70]
- Quality value format used (Sanger, Illumina GA v1.3 or later): [Sanger]
- How many best category hits is a read allowed to have (1, 2, no limit): [2]
- How many valid alignments are reported per read (1, 2, 3): [1]
- Put multireads to a separate file (yes, no): [yes]
- Put unaligned reads to a separate file (yes, no): [yes]

The genome alignment takes between 30-120 min, depending on the number of reads in the FASTQ file. The format of the output files for multireads and unaligned reads is FASTQ, for the index file is bam.bai (allows fast look up of sorted data) and the aligned sequences are in a BAM format (BAM is the compressed binary version of the Sequence Alignment/Map (SAM) format, a compact and indexable representation of nucleotide sequence alignments). The report for the total number of mapped reads is in a TXT format. In cases where more files with raw reads were obtained (which is quite typical), they were aligned separately and then merged with the merge BAM tool from Utilities.

Peak calling: next, in order to find the coordinates of statistically significant enriched genomic regions, peak calling was performed with MACS or MACS2 (Zhang et al., 2008). MACS is typically used when narrow signal distrubution is expected (e.g. transcription factors or H3K4me3) and MACS2 is used when broader signal distribution is expected (e.g. H3K27me3 or H3K9me3). It can be selected from the ChIP-, DNase-, and Methyl-seq menu with the following parameters:

- File format (ELAND, BAM, BED): [BAM]
- Mappable genome size: [hg19]
- Read length (1..200) [50- or 100]
- Bandwidth (1..1000) [75] (this is half of the length of the sequenced fragments)
- P-value cutoff (0..1) [0.00001]
- Peak model (yes, no) [yes]
- Upper M-fold cutoff (1..100) [30]
- Lower M-fold cutoff (1..100) [10]
- Input datasets (treatment data file-this is ChIP, control data file- this is Input)

The most important output file is the one with positive-peaks.tsv (although it is labeled as TSV, this is BED format). BED formated files contain the name of the chromosome, the start and end position of the peak coordinate, and sometimes descriptive statistics such as p-values. MACS was used only for the analyses of H3K4me3 readers.

For broader histone PTMs (H3K9me3, H3K27me3 and H3K36me3) the peak calling algorithm SICER was used (Zang et al., 2009). The BAM files were uploaded to one of the Galaxy based platforms https://usegalaxy.org/ or http://deeptools.ie-freiburg.mpg.de/ (Blankenberg et al., 2010; Ramirez et al., 2014). SICER can be selected from the NGS: peak calling menu and run with the following parameters: window size of 200, fragment size of 150, gap size of 1200 (200 for narrow marks), and statistical threshold value of 0.01 (False Discovery Rate or FDR). The most important output file is the BED file, which is comprised of the genomic coordinates of the detected enriched regions. Peak overlap was performed with the "Intersect" tool from the "Operate on Genomic Intervals" menu from a galaxy-based platform (Galaxy, Cistrome or DeepTools).

Genome coverage maps: for further downstream analyses and browser visualizations, so called density profiles of genome coverage maps were produced. There are numerous ways to prepare density profiles in different file formats, which are interconvertible (BedGraph, Wig and BigWig). For bedGraph profiles, the obtained BAM files were uploaded to https://usegalaxy.org/ and "Create a BedGraph of genome coverage" was selected from the BedTools (Quinlan and Hall, 2010) menu. All bedgraph profiles were normalized to the sample with the highest number of reads. This generated normalized raw bedgraph files with RPM (reads per million). Then, the bedgraph file were converted to bigwig with the "Wig/BedGraph-to-BigWig converter" for the Galaxy menu. The size of the bigwig is smaller and easier to load to genome browsers such as UCSC or IGV, when manual inspection is necessary. Another way to produce density profiles, directly into bigwig format is with the "bamCoverage" tool from DeepTools. Similarly, BAM files have to be uploaded to http://deeptools.ie-freiburg.mpg.de/ and converted into bigwig files with different normalization options, such as normalization to 1x, normalization to own scale factor, no normalization, which is equal to RPM (Reads Per Million) or to RPKM (Reads Per Kilobase per Million). Herein, RPM or RPKM was used.

Annotation to genomic elements and correlation: the BED files with peaks were uploaded to Cistrome http://cistrome.org/ap/root and analyzed with CEAS (Liu et al., 2011; Shin et al., 2009). The CEAS tool can be found in the "Integrative analysis" menu and can be used to annotate to list of peak to different genomic elements, such as promoters, distal elements, intergenic regions et cetera. EpiExplorer (Halachev et al., 2012) was used to annotate peaks to different genomic elements, chromatin segments and chromatin states. In order to calculate the Pearson correlation coefficient of multiple samples in fixed size windows (data in wig or bigwig files) the "Multiple Wiggle Files Correlation" tool from the "Integrative Analysis" menu from Cistrome was used. Another way to calculate the Pearson or Spearman Correlation coefficient of raw aligned data (BAM files) was with the "bamCorrelate" tool from DeepTools. The latter was applied in Appendix 4.

Metagene analyses: another useful way of presenting deep sequencing data is to plot the raw signal (BAM or bigwig) over certain genomic elements (gene bodies, TSS, TTS, enhancers et cetera). This was achieved with "computeMatrix" from DeepTools, which prepares an intermediary file with scores of raw data associated with selected genomic elements (BED files), which can be later plotted with the "profiler" tool. For example, if the goal is to examine the enrichment of the target histone modification over promoters, a BED file comprised of the promoter genomic coordinates needs to be prepared. To do this, a list of genes should be uploaded in a BED format directly to a galaxy-based platform (for example DeepTools) from the UCSC table browser or other source. Then, promoters can be prepared with "Get Flanks" from the "Operate on Genomic Intervals" menu by selecting the following flanking regiins: + 1 kb and -1 kb around the start site.

Heatmap generation: another way to present the distribution of raw data over different genomic intervals is by heatmap visualization. This was achieved with K-means clustering using the SeqMINER computational tool (Ye et al., 2011), which is freely available http://sourceforge.net/projects/seqminer/. The data (BAM format) was uploaded along with a set of genomic coordinates (BED format), which can be peaks from a ChIP experiment, TSS, TTS, enhancers et cetera. The software is then used to compute a density array over a defined window around the reference coordinates (e.g. +5 kb, -5 kb, depending on the analysis), which can be readily visualized. In this way, clusters of overlapping chromatin states can be discovered, which can be used for instance, to show the similarity between two affinity reagents or the overlapping regions of two different antibodies.

3.1.6.2 RNA- sequencing: mapping, assembly of transcripts, ranking of transcripts and other.

The most commonly used package for RNA-seq data analysis is the Tuxedo Suite (Trapnell et al., 2012), which is integrated within Chipster (Kallio et al., 2011). Upon confirmation of the data quality with FastQC, the next step was to map the raw to a reference genome (in our case Hg19). The data were imported to Chipster and mapped with TopHat2 for single reads from the "Alignment" menu, with the following parameters:

- Genome: [human hg19]
- How many hits is a read allowed to have (1-1000000): [20]
- Number of mismatches allowed in final alignment: [2]
- Minimum anchor length (3-1000): [8]
- Maximum number of mismatches allowed in the anchor (0, 1, 2): [0]
- Minimum intron length (10-1000): [70]
- Maximum intron length (1000-1,000,000): [500,000]

After alignment, the obtained BAM file was used for transcript assembly using Cufflinks from the "RNA-seq" menu from Chipster. Cufflinks assembles the alignments into a parsimonious set of transcripts and it estimates the relative abundance of these transcripts based on the number of reads. Cufflinks results in four output files:

1. Transcripts.gtf – contains the assembled isoforms with standardized attributes such as "gene id" and "transcript id"
2. Genes.fpkm_tracking.tsv – contains the estimated gene-level expression values in FPKM tracking format.
3. Isoforms.fpkm_tracking.tsv – contains the estimated isoform-level expression values in FPKM tracking format.
4. Skipped.gtf – contains the skipped loci.

After assembly the transcripts can be ranked based on their FPKM values in Excel or elsewhere and used as a BED file for metagene analyses.

3.2 Targeted epigenome editing

3.2.1 Cloning, site-directed mutagenesis and production of recombinant adenovirus

Cloning was performed in pAdTrackCMV and pAdEasy-1 vectors (purchased from Addgene) using standard ligase dependent molecular biology methods and the site-directed mutagenesis

was performed as described in (Jeltsch and Lanio, 2002). The adenoviruses were produced in HEK293 cells maintained in Dulbecco's modified Eagle's medium with 5% fetal bovine serum 37°C at 5% CO_2 essentially as described in (Luo et al., 2007). The virus was collected and purified by using CsCl density gradient centrifugation and gel filtration. All adenoviral work was performed at the Dr. Margarete Fischer-Bosch Institute for Clinical Pharmacology in Stuttgart, compliant to Biosafety level 2 guidelines and regulations. For more details refer to the Method's section in the manuscript in Appendix 3.

3.2.2 Infection of SKOV3 cells with recombinant adenovirus

SKOV3 cells were obtained from ATCC (American Type Cell culture Collection) and cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, L-glutamine and penicillin/streptomycin. Virus dilutions yielding >95% infection, without affecting the cell's viability, were used. For more details refer to the Method's section in the manuscript in Appendix 3.

3.2.3 Analysis of methylation by bisulfite conversion

Genomic DNA was recovered with QIAmp® columns (QIAGEN) and underwent bisulfite conversion with sodium bisulfite and sodium hydroxide. For more details regarding bisulfite conversion and primers used in this study refer to the Method's section in the manuscript in Appendix 3.

3.2.4 Quantitative PCR and gene expression analysis

The quantitative PCR assays were performed on a CFX96 Connect Real-Time detection system (Bio-Rad) using SsoFast EvaGreen supermix (Bio-Rad). In the nChIP experiments a standard curve was used to calculate the percent of precipitated DNA. For the mRNA transcript quantification analyses the expression *VEGF-A* was quantified using the $2^{-\Delta\Delta CT}$ method (C_T - threshold cycle) (Livak and Schmittgen, 2001). For more details regarding primers, and normalization strategy refer to the manuscript in Appendix 3.

4. Results

The unstructured N-terminal tails of histones stick out from the core nucleosomes and are massively decorated with a multitude of post-translational modifications (Kouzarides, 2007). Along with DNA methylation, they represent major epigenetic means of chromatin regulation. They are engaged in many chromatin transactions and their dysregulation is associated with many diseases. In order to study the functional effects of these chromatin modifications, two approaches were conceived, developed and successfully applied in this body of work. First of all, we developed and applied novel affinity reagents based on histone modification interacting recombinant proteins in chromatin studies. This strategy can be used as a cheap and reproducible alternative to histone modification antibodies. Second, we developed a zinc finger based targeting strategy of chromatin modifiers to an endogenous locus. With this experimental approach, we delved into the innerworkings of chromatin state establishment, studying its kinetics and maintenance.

4.1 Development and application of histone modification interacting domains (HMIDs) as an alternative to antibodies

So far, antibodies have been the only tools in our molecular toolkit, which can be employed for genome-wide studies of histone PTMs with nucleosome and locus-specific resolution. In spite of their historically essential role in the elucidation of the histone "code", antibodies sometimes lack specificity and suffer from batch to batch variability of properties. To develop a cheap alternative to antibodies with invariable properties, we assessed the potential and utility of recombinant histone modification interacting domains (HMIDs) in western blot-like and ChIP-like experiments. The results provided below have been published in *Genome Research* and in this section they will be only concisely summarized. For more details, refer to (Kungulovski et al., 2014), which is added in Appendix 4.

4.1.1 HMIDs as H3K9me3, H3K27me3 and H3K36me3 specific affinity reagents

The following well characterized HMIDs were selected for the initial work: ATRX ADD, MPP8 Chromo, HP1β Chromo, Dnmt3a PWWP, PHF20L1 MBT, Setd7 MORN, TAF3 PHD, Rag2 PHD, TFIID double Bromo, KDM4A double Tudor, PHF1 Tudor, CBX7 Chromo, CBX2 Chromo. All of them have been studied in our lab, and critical discoveries were made pertaining to the roles of Dnmt3a PWWP (Dhayalan et al., 2010), ATRX ADD (Dhayalan et al., 2011) and PHF1 Tudor (Kycia et al., 2014). After initial screening and specificity profiling, a handful of them were selected for a more extensive characterization viz., ATRX ADD, MPP8 Chromo, CBX7 Chromo, Dnmt3a PWWP. The ATRX ADD domain is a hybrid of a GATA-like zinc finger and PHD finger and it binds to H3K9me3 when H3K4 is unmodified (Dhayalan et al., 2011; Iwase et al., 2011). The H3K9me3 binding pocket of ATRX ADD is atypical because it contains only one aromatic residue. The MPP8 Chromo (binding to H3K9me3), CBX7 Chromo (binding to H3K27me3) and Dnmt3a PWWP (binding to H3K36me3) have a typical conformation of Royal family of domains, with three, three and five β-strands, and the target peptide binding antiparallelly (Bock et al., 2011b; Dhayalan et al., 2010; Min et al., 2003; Wu et al., 2011). Throughout this work we performed side-by-side comparison of the properties relevant for application of our HMIDs and validated antibodies (raised against H3K9me3, H3K27me3 and H3K36me3) following ENCODE quality criteria (Egelhofer et al., 2011; Landt et al., 2012).

First, we employed CelluSpots Peptide arrays to examine the binding specificity of both, the antibodies and HMIDs. The CelluSpots peptide arrays have been established in our lab to study the specificity of antibodies and HMIDs (Bock et al., 2011a; Bock et al., 2011b). They were

used here for the first time in a systematic side-by-side comparison of HMIDs and antibodies (Fig.1; Fig. 2A; Fig. 4A; Supplemental Fig. S1; Supplemental Fig. S6 from Appendix 4). All reagents showed minor or major cross-reactivity with related or unrelated histone modifications. In general, the cross-reactivity was more pronounced with antibodies. For a more detailed explanation of the specificity profiles on a per affinity reagent basis, please refer to Appendix 4.

Next, all affinity reagents were tested in western blot experiments with native histones (which are modified and binding is expected) and recombinant histones (which are unmodified and no binding is expected). This is an established quality control check up of chromatin antibodies, which examines the preferential binding to modified histones. All reagents performed robustly and bound only modified histones (Fig. 2B; Fig. 4B and Fig. 5A from Appendix 4). Furthermore, they were successfully tested with nucleosomes isolated from wild type cells or cells depleted from the corresponding histone methyltransferase (Fig. 2C; Fig. 4C and Fig. 5B from Appendix 4). The specific binding to nucleosomes isolated from wild type cells and the notable reduction of signal with nucleosomes isolated from methyltransferase depleted cells were strong indicators for specific binding to the targeted modification.

The standout application of antibodies lies in chromatin immunoprecipitation (ChIP) experiments coupled with readout by qPCR or deep sequencing. For this aim, we investigated the potential of HMIDs to precipitate discrete types of chromatin in native ChIP-like assays, termed CIDOP (Chromatin Interaction DOmain Precipitation). First of all, the antibodies and HMIDs were compared in ChIP-qPCR and CIDOP-qPCR assays, respectively and later these assays were extended with deep sequencing of the recovered DNA (Fig. 3; Fig. 4D-I; Fig. 5D-H from Appendix 4). The chromatin precipitation experiments showed high correlation between corresponding validated antibodies and HMIDs, and proper distribution over expected genomic elements (for example signals obtained with an anti-H3K9me3 antibody, ATRX ADD and MPP8 Chromo were enriched in heterochromatin elements). This was an important result, which demonstrated the proof-of-principle applicability of HMIDs. Similar results were obtained for HMIDs and antibodies specific for H3K36me3 and H3K27me3, respectively.

Finally, we showed that HMIDs are amenable to protein engineering approaches, by altering the specificity of the KDM4A double Tudor domain from di- and trimethylation pan-H3/H4 binding to dimethylation pan-H3/H4 binding (E929A variant) and H4K20me2/3 monospecific readout (D969A variant) (Supplemental Fig. S11 from Appendix 4).

4.1.2 HMIDs as H3K4me3 specific affinity reagents

After clearly demonstrating that HMIDs specific for broad histone marks such as H3K9me3, H3K27me3 and H3K36me3 performed robustly in chromatin experiments (Kungulovski et al., 2014), we sought to develop an H3K4me3 affinity reagent. For this reason we investigated the potential of the TAF3 PHD domain as an H3K4me3 specific affinity reagent in chromatin studies (these results are not yet published). The peptide array specificity profiling of TAF3 PHD showed binding to H3K4me3 peptides (Fig. 21A), whilst far western blot experiments with native and recombinant histones showed preference for modified histones, but also some cross-reactivity with non-histone proteins (Fig. 21B).

CIDOP-qPCR experiments of selected genomic regions known to be either associated with H3K4me3 signal or with other non-target histone modifications (selected based on available ChIP-seq data) indicated a specific precipitation of nucleosomes harboring histone H3 trimethylated at lysine 4 by TAF3 PHD (Fig. 22A). This observation was further validated with side-by-side comparison of TAF3 PHD CIDOP-seq and anti-H3K4me3 ChIP-seq datasets, which showed similar overall distribution of signals (Fig. 22B), overlap of peaks and high

Pearson correlation in promoter regions (Fig. 23A and B). In addition to this, chromatin precipitation with both the recombinant domain and the antibody exhibited strong enrichment in the immediate vicinity of TSS (Fig. 23C and D), which is in line with literature (Barski et al., 2007).

Collectively, the presented data show that TAF3 PHD can be successfully applied as an alternative to anti-H3K4me3 antibodies, albeit under the initial conditions used here with lower sensitivity and lower signal to background ratio than good anti-H3K4me3 antibodies (Fig. 22B; Fig. 23C and D).

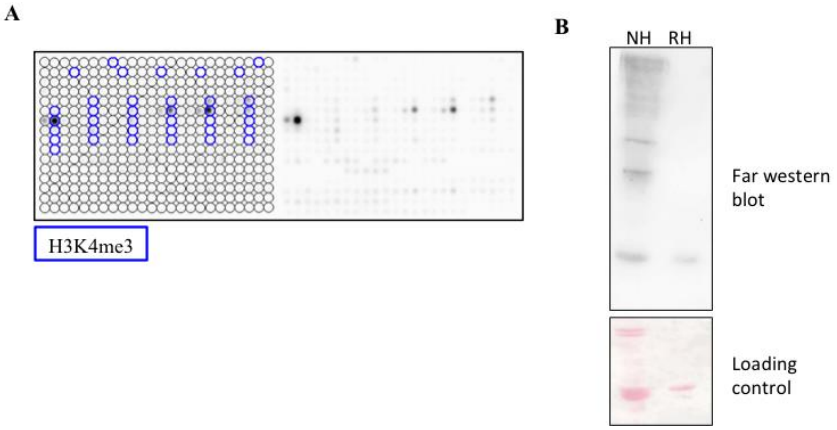

Figure 21. Specificity analyses of TAF3 PHD. A) Peptide array analysis of the specificity of TAF3 PHD. For comparison with anti-H3K4me3 antibodies, refer to (Bock et al., 2011a). B) Western blot analysis using TAF3 PHD with native (NH) and recombinant histones (RH).

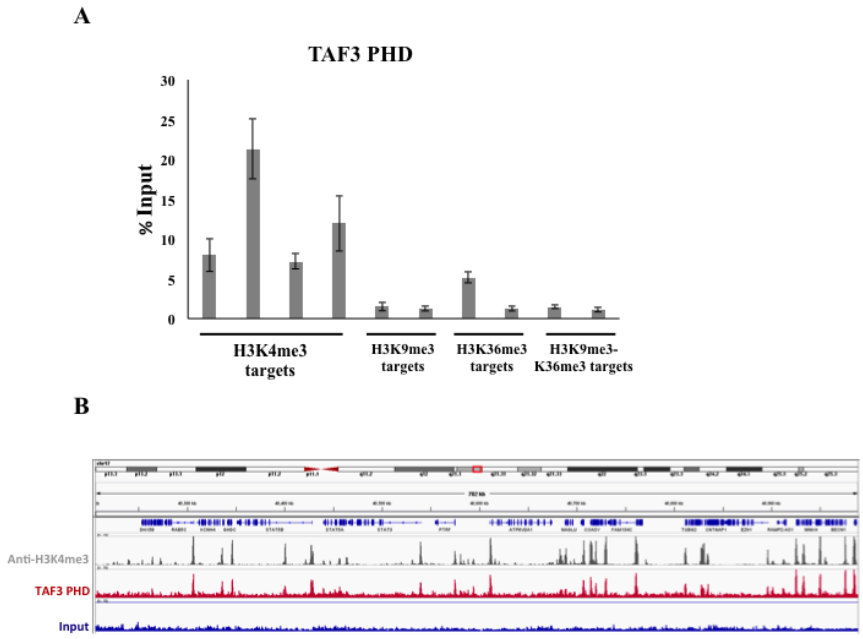

Figure 22. CIDOP and ChIP experiments with H3K4me3 affinity reagents. A) CIDOP-qPCR results of the TAF3 PHD domain using amplicons known to be associated with H3K4me3, H3K9me3, H3K36me3 and H3K9me3-H3K36me3 histone PTMs. All experiments were carried out in triplicates of biological duplicates. Error bars represent the standard errors of the mean. B) Representative genome browser snapshot comparing CIDOP-seq and ChIP-seq results obtained with TAF3 PHD and anti-H3K4me3 antibody (ENCODE).

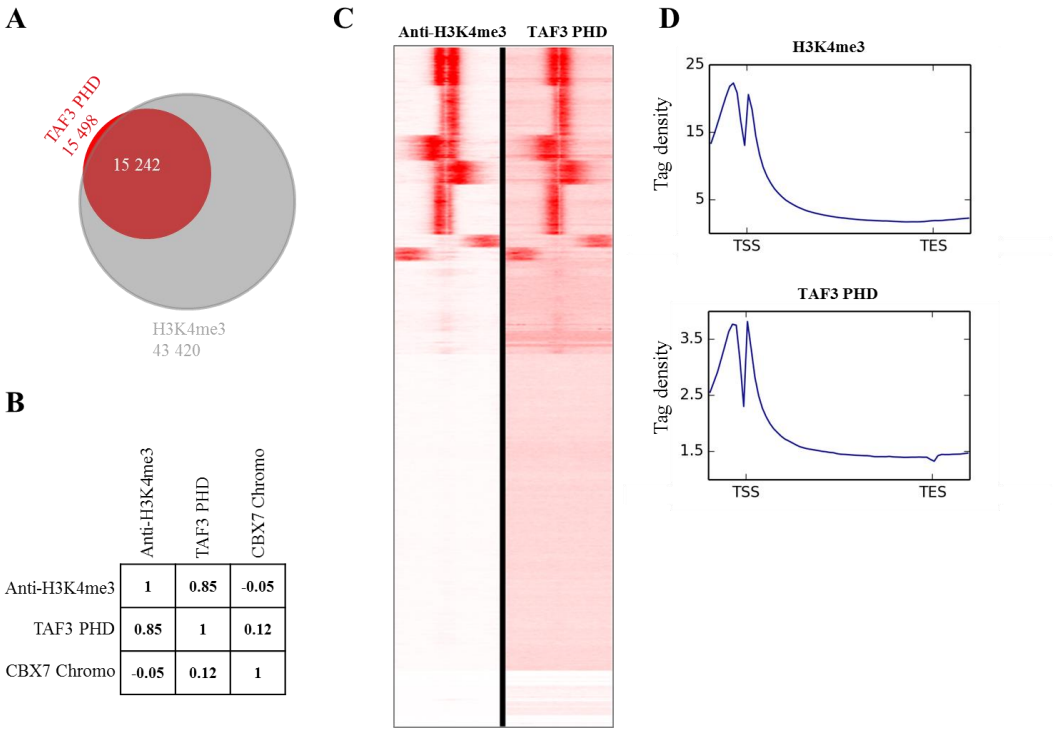

Figure 23. CIDOP-seq and ChIP-seq downstream analyses. A) Venn diagram of the overlap of peak regions between TAF3 PHD and anti-H3K4me3 antibody. The number of unique TAF3 PHD peaks is 256, and the number of unique anti-H3K4me3 peaks is 28178. B) Pearson correlation coefficient of TAF3 PHD, anti-H3K4me3 and control affinity reagent (CBX7 Chromo) within promoters (defined as ±2000 bp from TSS). C) K-means clustering analysis of tag densities from TAF3 PHD and anti-H3K4me3 antibody CIDOP-seq and ChIP-seq datasets around promoters. Tags were collected within 10-kb window around TSS and sorted by K-means clustering (10 clusters). D) Composite profiles showing the average intensity and distibution of TAF3 PHD and anti-H3K4me3 antibody signal plotted over a bioinformatically generated metagene with uniform length.

In conclusion, our proof-of-principle analyses clearly demonstrated that various HMIDs can in general perform at least as good as validated antibodies and can be successfully used alone or in concert with antibodies in genome-wide or locus-specific chromatin studies.

4.2 Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained

DNA methylation and histone 3 lysine 9 methylation are considered as *bona fide* epigenetic modifications, meaning that most researchers would assume that they are stably maintained upon introduction to a selected genomic region. The recent bloom of genome-wide studies has provided a deep but often implicit and merely correlative understanding of the distribution and association of epigenetic marks with one another or with other genomic elements. In contrast, the development and application of more direct and targeted functional studies has been lagging behind. In the present study, we wanted to explore in detail the functional repercussions, stability and dynamics of DNA and H3K9 methylation by targeted epigenome rewriting of the *VEGF-A* locus. To this aim, a zinc finger protein that binds to the promoter of the *VEGF-A* gene was fused with the catalytic domain of Dnmt3a (ZF-Dnmt3a-CD) or GLP (ZF-GLP-CD) and introduced into SKOV3 cells via an adenoviral vector. The results provided below have been published in *Epigenetics & Chromatin* and in this section they will be only concisely

summarized. For more details, refer to (Kungulovski et al., 2015), which is added in Appendix 3.

4.2.1 Establishment of an adenoviral delivery sistem

Previous work from our group has successfully demonstrated targeted DNA methylation in the promoters of the *VEGF-A* and the *EpCam* genes after transient conventional transfection with vectors expressing targeted DNA methyltransferases (Nunna et al., 2014; Siddique et al., 2013). *VEGF-A* was selected as a suitable model system for follow up experiments aiming to study chromatin dynamics and kinetics, because no change in cell proliferation was observed after its repression. In contrast silencing of *EpCam* led to reduced cell proliferation (Nunna et al., 2014; Nunna and Jeltsch, personal communication). Adenoviral delivery allows for almost 100% infection rates, making the data analysis and interpretation easier and more reliable, compared to transient transfection. Upon adenoviral infection, the cargo proteins encoded by the adenoviral vector are expressed for approximately 5 days. In this time period, the the zinc finger protein guided DNA or histone methyltransferase catalytic domains can conduct targeted epigenome editing in the promoter of *VEGF-A*. After expression of approximately 5 days, the virus vector is diluted and eventually lost from the cells, which terminates the production of the targeted methyltransferases. The reason for this is the inability of the virus to propagate inside the cells, because it is lacking the E1 and E3 genes, encoding for capsid proteins. Consequently, since the targeted methyltransferase is lost due to virus dilution, this allows for investigation of the capabilities of the cellular system to maintain the introduced epigenetic signal.

4.2.2 Zinc finger targeted DNA methylation

Upon adenoviral infection of SKOV3 with the ZF-Dnmt3a-CD construct, the changes in DNA methylation and expression of the *VEGF-A* gene were monitored. We observed a gradual elevation of DNA methylation in the gene promoter from day 1 post-infection onwards. The methylation signal peaked at day 5 and then it was gradually lost, reaching basal levels at day 15 (Fig. S2 and Fig. 2A, B and C from Appendix 3). H3K9me3 is a silencing mark, correlated with DNA methylation. To check if the presence of DNA methylation leads to introduction of H3K9me3, we performed native ChIP-qPCR experiments at day 5 but observed no changes in H3K9 trimethylation (Fig. S8A from Appendix 3). Next, we monitored the levels of *VEGF-A* gene expression upon targeted DNA methylation and observed a strong reduction of mRNA levels at day 5, followed by gradual re-activation and return to native levels at day 15 (Fig. 2D from Appendix 3). The specificity of the targeted methylation was confirmed in control experiments with constructs with targeted but catalytically inactive Dnmt3a or constructs harboring only the ZF protein or only the Dnmt3a (Fig. S3 from Appendix 3).

4.2.3 Zinc finger targeted histone H3K9 methylation

Furthermore, we sought to investigate the functional consequences of installing H3K9 di- and trimethylation at the *VEGF-A* promoter. To this aim, we fused the catalytic domain of the GLP H3K9 methyltransferase with the same zinc finger protein. Upon adenoviral delivery, we observed a gradual deposition of H3K9me2 and H3K9me3 in the vicinity of the ZF binding site, which peaked at day 5. In a systematic approach, we also detected massive spreading of H3K9 methylation at least 5 kb upstream and 15 kb downstream from the ZF binding site. Similar to targeted DNA methylation, the levels of H3K9me2 and H3K9me3 started to gradually decrease after day 5 and returned to their basal levels at day 15, albeit with different kinetics of disappearance, suggesting active histone demethylation (Fig. 3 from Appendix 3), which is addressed in the detail in the Discussion section. The deposition of H3K9 methylation resulted in strong gene silencing at day 5, with gene expression returning to its native state at day 15 (Fig. 4A from Appendix 3).

To further investigate the chromatin state refurbishment and cross talk of epigenetic marks, we performed native ChIP-qPCR experiments against H4-acetylation, MNase mapping to check for nucleosome occupancy and bisulfite sequencing to check for DNA methylation. Interestingly, we observed an overload of histone H4 acetylation at day 10, which is the exact time point when the expression of the targeted methyltransferase had declined (Fig. 4B and C from Appendix 3), suggesting a recruitment of histone acetyltransferase to the *VEGF-A* promoter. No changes were detected in nucleosome occupancy or DNA methylation (Fig. S5 and S8B from Appendix 3). The specificity of the targeted H3K9 methylation was confirmed with an array of control experiments with zinc finger fused inactive methyltransferase constructs or constructs harboring only the ZF protein or only the GLP methyltransferase (Fig. S4, S6 and S7 from Appendix 3). In addition, the zinc finger was fused with two different catalytically inactive GLP variants (ZF-GLP-CD C1201A and ZF-GLP-CD ΔNHHC) (Tachibana et al., 2008). These constructs differ in their ability to interact with G9A H3K9 methyltransferase, which is a heterodimeric partner of GLP. The GLP C1201A variant is still able to form heterodimers with G9A but the GLP-CD ΔNHHC is not. We observed slightly stronger residual silencing of *VEGF-A* with the GLP C1201A variant, suggesting that the native G9A enzyme could be recruited to the target site and this has a weak enhancing effect on gene repression. Finally, we confirmed that the loss of DNA and H3K9 methylation signal is not caused by a selective growth disadvantage after *VEGF-A* silencing, illustrated in time-course experiments with silencing and control vectors (Fig. 4C from Appendix 3).

Collectively these data show a direct link between deposition of DNA or H3K9me2/3 and gene silencing. Excitingly, the unexpected lack of stability and inheritance of the established epigenetic state, provides a valuable insight into the complex regulation of chromatin maintenance, which will be further discussed.

5. Discussion

Multicellular organisms consist of tens to hundreds of cell types, which contain the same genetic information but are phenotypically, morphologically and functionally different. Such overt diversity based on identical genomic sequences can be established and maintained only with highly regulated spatial and temporal control of gene expression enacted by so-called epigenetic mechanisms. The epigenetic means of gene expression offer the perfect counterbalance between stable maintenance of cellular identity and plastic responses to environmental changes. DNA methylation and post-translational modifications of histones are fundamental epigenetic modalities and chromatin is embellished with complex patterns of these chromatin modifications (Bannister and Kouzarides, 2011; Jones, 2012; Kouzarides, 2007; Margueron and Reinberg, 2010; Tan et al., 2011). Knowing the prime importance of epigenetic marks, in this body of work we have developed and applied new experimental tools for thorough dissection of the distribution, function and dynamics of chromatin modifications. The following discussion is an attempt to extend some of the ideas conceived in the publications attached in the Appendix and offer a more comprehensive and integrated outlook.

5.1 Application of histone modification interacting domains (HMIDs) as an alternative to antibodies

5.1.1 The problem with histone PTM antibodies

Chromatin immuprecipitation studies of histone PTMs have been the workhorse of chromatin biology and have provided incredible insight in deciphering the complex syntax of the histone modification language. Indeed, the antibody is the indirect mediator that translates the language of the histone code into observable properties that can be studied with molecular biology approaches. Therefore, the reliability of the scientific work in this field is deeply reliant on the quality of the antibody employed for capturing the material. For specific interaction with a defined target histone PTM, the antibody has to have the ability to recognize and discriminate the small, subtle and chemically similar histone modifications. In other words, the antibody has to intrinsically discern the unmodified amino acid residue from all heterotypically modified forms (e.g. acetylation or methylation) and in addition to that it has to distinguish different homotypic modifications (e.g. mono-, di-, and trimethylation). A further complication for proper readout is the hypermodified state of the histone tail, in which secondary modifications might abrogate binding and yield false negative results. Moreover, some histone modifications such as acetylation or methylation of H3K9 and H3K27 lie within an identical amino acid context (an ARKS motif in this case), which can only be distinguished if the antibody also contacts amino acids surrounding this central motif. Apart from the requirement to identify and discriminate minute chemical modifications in a hypermodified environment, which often leads to cross-reactivity, antibodies also suffer from some major production drawbacks due to polyclonality, such as batch dependent variability of specificity and affinity. Knowing these downsides, many labs have raised concerns about the quality of histone PTM antibodies (Bock et al., 2011a; Egelhofer et al., 2011; Fuchs et al., 2011; Hattori et al., 2013; Heubach et al., 2013; Nishikori et al., 2012; Peach et al., 2012).

We corroborated these observations and concerns with peptide array analyses of different antibody lots having the same catalog number (Fig. 1 from Appendix 4) and to tackle this issue, we devised and applied an alternative to antibodies based on recombinant “reading domains” in western blot-like and ChIP-like assays. In side-by-side comparison, we successfully examined the potential of the MPP8 Chromo and ATRX ADD, CBX7 Chromo, and Dnmt3a

PWWP domains to substitute anti-H3K9me3, anti-H3K27me3 and anti-H3K36me3 antibodies, respectively. In support of this concept, affinity-based reagents for enrichment of methylated or unmethylated CpG islands (Blackledge et al., 2012; Cross et al., 1994) or enrichment of methylated non-histone proteins (Moore et al., 2013) have been also used for similar purposes. With this work we extended the arsenal of tools deployed for genome-wide and locus-specific investigation of histone modifications.

5.1.2 Establishment of chromatin precipitation protocols and bioinformatics pipeline

The first goal of this doctoral thesis was to assess the applicability of HMIDs in western blot-like approaches. After successfully profiling the specificity of the recombinant proteins on CelluSpots peptide arrays, we applied the same protocol in far western blot experiments, simply by exchanging the peptide array with nitrocellulose membranes. Following the success of this approach, the next step was to set up the epigenomics work *de novo*.

The chromatin precipitation protocols were conceived by taking into account the stringency of conventional ChIP protocols and the binding affinities of the corresponding HMIDs. Therefore, different buffers with different ionic strengths and detergents were tested for different domains until successful precipitation was achieved. The outcome of chromatin precipitation was examined initially by end-point PCR and later by quantitative PCR. The design of primers was based on available ChIP-seq data and the optimization of the qPCR reactions and protocol were established in the lab for the first time in this work. We note that additional optimization steps including buffer composition, elution and library construction might be necessary for true and commercial application of HMIDs. For more details regarding the final buffer composition, refer to the manuscript added in Appendix 4.

Next, the bionformatic pipeline for CIDOP-seq and ChIP-seq data analysis had to be set up in the lab from scratch. First of all, significant efforts were undertaken in order to establish a bioinformatic workflow comprised of different tools and algorithms, which needed to provide a balance between speed, accuracy and ease of use. For genome alignment, the BWA and bowtie mappers integrated within the Galaxy platform (public server) or the Chipster analysis tool (guest server) were tested. Due to the speed and reliability, the bowtie mapper from Chipster was selected and subsequently used. Next, for determination of enriched regions different peak callers such as MACS (for narrow histone marks) and SICER (for broad histone marks) were used. MACS is available through the command line interface or integrated within graphical interface tools such as Chipster or Galaxy-based platforms (Galaxy, Cistrome or deepTools) while SICER is only available through the command line or Galaxy-based platforms. These algorithms provide files with genomic coordinates, and different algorithms based on BEDtools, EpiExplorer or CEAS were tested and subsequently used for their downstream analyses such as intersection of enriched regions or annotation of peaks to genomic elements. Finally, numerous datasets had to be compared side-by-side and different statistical, metaregion and clustering approaches had to be tested and applied in the final analyses.

5.1.3 Questions of affinity

One of the most relevant questions regarding the robust and reproducible application of HMIDs as an alternative affinity reagent for the characterization of histone modifications is their affinity for the target substrate mark. The K_D values of the HMIDs used in (Kungulovski et al., 2014) are in the high nanomolar to low micromolar range (between 120-140 nM for MPP8 Chromo to H3K9me3 peptides (Bock et al., 2011b; Chang et al., 2011); 1.4 μM for ATRX ADD to H3K9me3 peptides (Dhayalan et al., 2011); 64 μM for Dnmt3a PWWP to H3K36me3 peptides (Dhayalan et al., 2010); 22 μM for CBX7 Chromo to H3K27me3 peptides (Bernstein et al., 2006)). Typical commercial histone PTM antibodies have K_D values, which are in the

low nanomolar to low micromolar range (0.2, 4, 60, 83, 520, 820, 1000, 2200 and 2700 nM (Hattori et al., 2013; Nishikori et al., 2012)). Hence, although the ranges of affinities between HMIDs and antibodies generally overlap, antibodies show slightly stronger affinities. The weaker binding affinity may lead to lower fractional saturation with the target substrate, but this can be compensated with higher concentrations of HMIDs per ChIP experiment as demonstrated by the successful application of HMIDs in my thesis.

5.1.4 Advantages of using recombinant HMIDs

The appeal of using HMIDs in western blot and chromatin precipitation experiments lies firstly in the ease of recombinant protein production in *Escherichia coli* making the technology available to any biochemistry lab. Secondly, most HMIDs are small and have a stable fold, which allows for high production yields of recombinant proteins, making the technology cost-effective, in comparison to limited batches of antibodies. Thirdly, crystal structures are available for many reading domains, which in turn allows for rational protein design and engineering of domains with modulated, altered or cleaner specificities, preparing of binding pocket mutants as optimal matching negative controls and addition of different or multiple affinity tags, which are massively expanding their applicative potential. However, the greatest advantage of the HMIDs technology is the elimination of batch-to-batch variability, which guarantees consistent quality and unlimited availability and potentially represents a break-through discovery for future chromatin research.

The initial observation and discovery of a scientific phenomenon is of paramount importance, but the real merit of the discovery lies in its general reproducibility. Hence, the relevance of a scientific observation can be significantly devalued, if the reagent (in this case antibody) used to generate the data is no longer available. First of all the data cannot be technically reproduced between labs, which will likely thwart independent confirmation of results. In addition, follow up studies, which might be necessary to study the observed phenomenon more thorougly would be impossible. Moreover, important questions arising from the study can no longer be addressed experimentally. All of these concerns are critical for high quality science. Second, the application of reagents with variable properties would have a limited utility in long clinical studies, where the invariability of the experimental approach is of vital importance.

In the case of HMIDs, once the reagent has passed the rigorous quality control pipeline, the obtained documentation will be valid for all of its future lots if the same standard production protocol is applied. Once documented, the primary specificity, effects of adjacent marks and other strengths and weaknesses of the reagent can be taken into consideration in the interpretation of data. The enlisted advantages are also valid for recombinant antibodies (Hattori et al., 2013), but their generation requires significant expertise, which is not readily available in a typical biochemistry lab. Similarly, the application of monoclonal antibodies could also reduce batch dependent variability, but empirically they have been shown to be less robust than polyclonals.

5.1.5 An update of quality control guidelines for histone PTM affinity reagents

Realizing the essential nature of antibodies in chromatin reserach, in the past couple of years, member labs from the ENCODE consortium suggested guidelines for primary and secondary criteria applied for quality control of histone PTM antibodies (Egelhofer et al., 2011; Landt et al., 2012). Inspired by the whole process of development and application of HMIDs in chromatin research, we set out to update, refine and extend the quality metrics for passing grade of an affinity reagent.

The first step for primary quality assessment of antibodies (and now HMIDs) set fourth by ENCODE is binding to a matrix of 43 peptides in dot blot experiments, where at least 75% of

the total signal has to be specific for the cognate peptide. These analyses should reveal the ability of the affinity reagent to accurately recognize the amino acid sequences on the histone peptide (e.g. discrimination of H3K9 and H3K27 marks); to effectively recognize the specific modification (e.g. discrimination of Kme2 and Kme3); and to determine the effects of adjacent marks on the binding affinity. However, peptide pull-down and dot blot experiments have a low throughput, and given the combinatorial diversity and abundance of histone modifications on native chromatin, we strongly advice the application of peptide array based methods which simply feature a larger number of different peptides for specificity analysis (Bock et al., 2011a; Bock et al., 2011b; Bua et al., 2009; Dhayalan et al., 2010; Garske et al., 2008; Garske et al., 2010; Rathert et al., 2008; Su et al., 2014; Zhang et al., 2010). Peptide arrays offer remarkable combinatorial power, and with the evolution and advancement of peptide synthesis they would provide even more valuable insight. For instance, loss of signal in a ChIP experiment could be due to active removal of the probed histone PTM or a negative effect on binding emanating from an adjacent mark. Primary peptide array analysis could aid the interpretation of data and increase the overall scientific accuracy of ChIP and CIDOP studies.

Although peptide based assays offer robustness and convenience, they collectively share common caveats, related to the peptide synthesis technology. As shown in Appendix 4, the specificity data obtained with peptide-based methods do not always reflect the results obtained in chromatin studies. For example, the MPP8 Chromo domain, which only binds H3K9me3 in a chromatin context (Fig. 3 from Appendix 4), exhibited modest cross-reactivity to H3K27me3 peptides on the peptide array (Fig. 2 from Appendix 4). Similarly, CBX7 Chromo bound H3K27me3 and H3K9me3 on the peptide array (Supplemental Fig. S6 from Appendix 6), but displayed a specific binding to H3K27me3 modified nucleosomes and histones (Fig. 5 from Appendix 4). The incongruity between peptide and chromatin based assays could be explained by the short sizes of the peptides, which in turn might lead to promiscuous binding or by the fact that peptides present artificial ends, which do not reflect the termini of the full-length native histone proteins.

In order to circumvent the caveats inherent to peptide based method, we suggest the use of barcoded nucleosomes, whenever possible (Nguyen et al., 2014). The principle of this method lies in the streamlined production of reconstituted nucleosomes with various modification patterns, which harbor a unique DNA barcode that reflects the modification state of the nucleosome. The mix of modified nucleosomes is precipitated with the affinity reagent, the recovered DNA is deep sequenced and quantified, and the detailed precipitation pattern directly reflects the specificity profile of the reagent. With the expansion of the throughput of barcoded libraries, we anticipate that this technology will take over and eventually replace the peptide arrays to a certain extent. However, since peptide array analyses are technically much easier and also less expensive than analyses using barcoded nucleosomes, we maintain that peptide array based methods provide valuable insight and should remain an essential quality control metric.

The second step for primary quality assessment of antibodies (and now HMIDs) set fourth by ENCODE is binding to histone extracts in western blot experiments. The suggestions from ENCODE are that the signal intensity of the specifically bound histone band should constitute at least 50% of the total western-blot signal observed with the histone extract. Moreover, the intensity of the specific band should be at least 10 fold higher than any other band or recombinant unmodified histones (Egelhofer et al., 2011; Landt et al., 2012). In the case of HMIDs, binding pocket mutants should be included in this step to ensure specificity is based on the presence of the modification. We suggest an obligatory extension of this primary quality

control step and inclusion of western blot experiments, where the interaction of the affinity reagents is examined with histone extracts from wild type cells in comparison with histone extracts from cells in which the responsible histone modification enzyme has been deleted or depleted. Depletion of the responsible enzyme should lead to attenuation or loss of signal. This experimental guideline was first suggested by ENCODE (Landt et al., 2012), but only as a secondary characterization step, which in our opinion should be made mandatory.

The third step for primary quality assessment of antibodies (and now HMIDs) set fourth by ENCODE is the detailed analysis of ChIP-seq and CIDOP-seq data. Firstly, the signals in CIDOP/ChIP biological replicates should have a high statistical correlation (e.g. Spearman's Rho of at least 0.8). In addition to this, we suggest to extend the correlative tests and include comparisons between two different affinity reagents (one antibody vs. another or antibody vs. HMIDs) (Egelhofer et al., 2011; Kungulovski et al., 2014) and (Kungulovski, submitted 2015). After passing these stringency criteria, tests such as IDR (Irreproducible Discovery Rate) (Landt et al., 2012), annotation of peaks to genomic elements and enrichment and distribution over expected genomic elements (obtained in similar studies) should provide additional insurance of the reliability of the reagent.

Mass spectrometry is one of the most sensitive and versatile analytical tools for studying histone modifications. In our experience mass spectrometry can be used to back up the enlisted quality control measures, but it should not be a primary quality control metric. First of all, although mass spectrometry can detect enrichment of certain histone modifications, it cannot corroborate the direct interaction of the cognate histone mark with the affinity reagent. Second, the enrichment of a histone modification could be due to cross-reactivity of the affinity reagents or colocalization of the histone marks on the same nucleosomal fragment.

All in all our suggestions for quality control of affinity reagents against histone modifications are along the following lines (Fig. 24):

- Specificity profiling to modified peptides in peptide array experiments using as many peptides with combinatorial modification patterns as possible. The most intense signal should be obtained for the targeted histone modification. Secondary modifications that prevent reagent binding and cross-reactivity leading to binding to non-target peptides need to be identified and documented. If possible, the affinity reagent should be tested against a library of modified barcoded reconstituted nucleosomes. In the case of HMIDs, binding pocket mutants should be included.
- Detection of modified histones in western blot experiments, with sufficient specificity as described above. Loss or attenuation of signal in western blot upon using histones isolated from cells where the corresponding histone modifying enzyme has been depleted. In the case of HMIDs, binding pocket mutants should be included.
- Reproducibility of chromatin precipitation experiments coupled to deep sequencing between replicates or two different affinity reagents and expected distribution and annotation of enriched regions to genomic elements.

5.1.6 Conclusion and perspectives

In this body of work we established a novel affinity and detection reagent based on HMIDs for reproducible characterization of histone post-translational modifications. We presented cogent evidence that HMIDs show unparalleled versatility in many chromatin applications. Similar to antibodies, they show some cross-reactivity on peptide substrates and dependence on adjacent marks. In accord with our work, recently it was demonstrated that HMIDs can also be used for

analysis of chromatin states by mass spectrometry (Su et al., 2014). We maintain that this novel technology holds a genuine applicative potential and could open many novel research avenues. For instance, with proper protein engineering, a researcher could be able to rationally design domains with novel specificities, domains with insensitivities to adjacent marks or domains with dual readout of coexisting histone modifications. The latter is especially noteworthy, because it can provide an easy, single step experimental system for direct investigation of co-occurring histone modifications. It is conceivable to engineer double domains, which would follow AND logic of binding (interaction only when two modifications are present at the same time). The AND logic of binding could be achieved by the biophysical properties of multi-dentate binding to the nucleosomal substrate, which leads to synergistic binding. This means, that the affinity of the double domain towards doubly modified nucleosomes would be stronger than the sum of affinities of the individual single domains towards singly modified nucleosomes. Dual readers of H4K20me3-H3K9me3 chromatin states could provide insight into the formation of heterochromatin or dual readers of H3K4me3-H3K27me3 could significantly aid the characterization of bivalent chromatin in embryonic and cancer cells. These applications were developed in a master thesis carried out under my direct supervision (Mauser, 2013). They are described in a patent application that is part of my doctoral work.

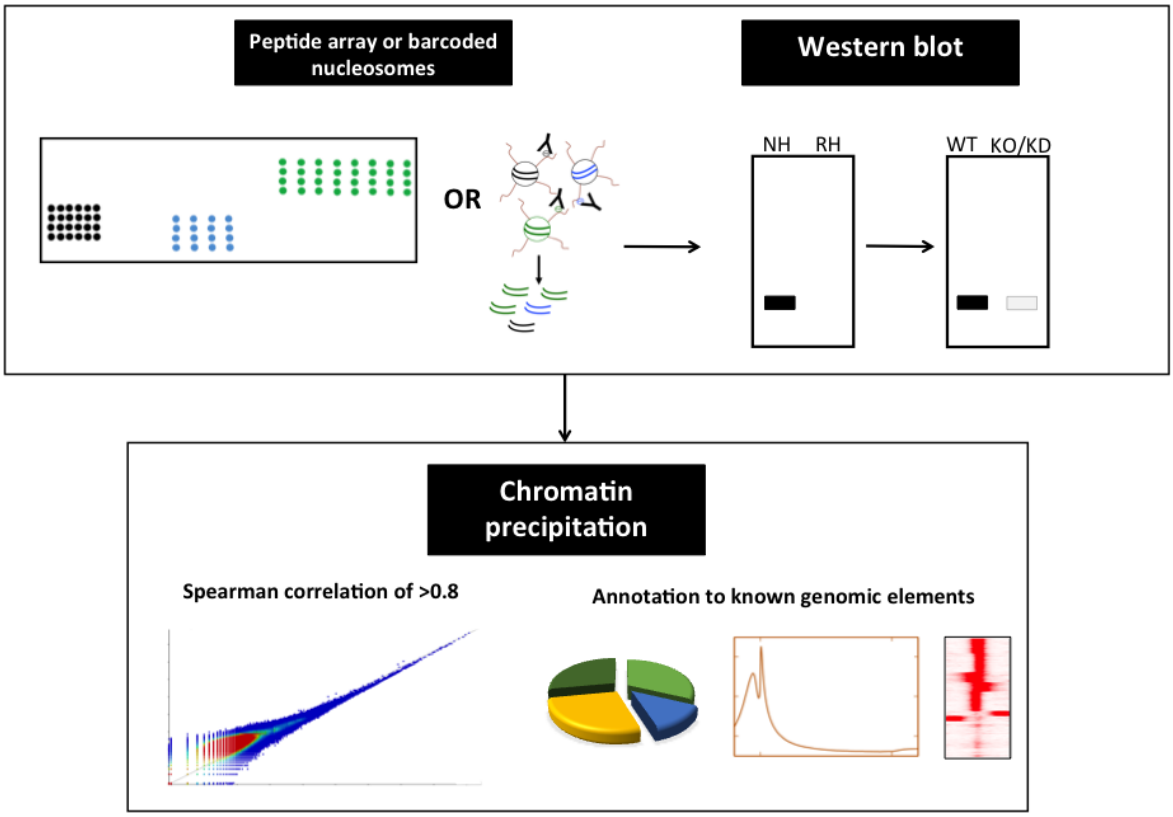

Figure 24. Summary of the guidelines for quality control of affinity reagents against histone PTMs put forward here. The image is taken from Appendix 1.

Finally, due to the hypermodified state of the N-terminal tails, obtaining ultraspecific "ideal" affinity reagents for particular histone modifications is a challenging and perhaps often unreachable goal. The only way to safeguard from spurious data interpretation and misguided conclusions is to establish a proper systematic and meticulous documentation of specificities and affinities of each detection reagent. With the addition of HMIDs into our molecular toolkit, we have broadened the palette of reagents for studying histone marks in a genome-wide scale with locus and nucleosome specific resolution, and solved the problem of batch-to-batch

varibilty of properties and long term availability, which represents an important contribution to the future of chromatin research.

5.2 Rewriting of epigenetic states

The transcriptional program, which orchestrates and maintains cellular identities, is generally controlled via epigenetic mechanisms, such as DNA methylation and histone PTMs. DNA methylation of CpG islands in gene promoters is associated with gene silencing, while DNA methylation within gene bodies correlates with gene expression and splicing (Jones, 2012). Histone modifications such as H3K9me3, H4K20me3 and H3K27me3 are correlated with repressive states with low transcription (constitutive and facultative heterochromatin, respectively). In constrast, histone modifications such as acetylation of histone H3 and H4, H3K4me2/3 and H3K36me2/3 are associated with actively transcribed chromatin. The dominant role of epigenetic signaling in many cellular processes, along with the plasticity of epigenetic states has inspired many approaches aiming to rewrite epigenetic states. One strategy to study the structure, function and dynamics of chromatin states is to locally perturb the native chromatin state of a locus and to follow the changes ensuing in a stepwise and systematic manner. One way to do this is by using chimeric proteins consisting of a targeting DNA binding module (such as zinc finger, transcription activator-like effector (TALE), or inactive CRISPR/Cas9) and an effector protein, which harbors a chromatin modifying activity. In the past decade a flurry of evidence has demonstrated the feasibility of the approach in reporter plasmids, viral DNA and endogenous regions (Chen et al., 2014; de Groote et al., 2012; Falahi et al., 2013; Maeder et al., 2013; Mendenhall et al., 2013; Nunna et al., 2014; Papworth et al., 2006; Ragunathan et al., 2015; Rivenbark et al., 2012; Siddique et al., 2013; Snowden et al., 2002). However, the stability and heritability of such perturbed epigenetic states, especially in mammals, remains elusive.

In our study, we tried to address some of these issues by targeting DNA and histone methylation to the promoter of the *VEGF-A* gene, followed by stepwise analysis of the appearance and disappearance of the epigenetic marks with spatial and temporal resolution. Compared to previous work from our lab (Nunna et al., 2014; Siddique et al., 2013), we significantly improved the efficiency of construct delivery with adenoviral vectors (Nunna et al., 2014). We observed a clear reverse trend between appearance of DNA/H3K9 methylation and reduction of gene expression or disappearance of DNA/H3K9 methylation and elevation of gene expression, which is highy suggestive of a direct effect of these modifications on gene regulation. The lack of preservation of the newly installed chromatin state after cessation of the triggering epigenetic modifier, indicates that the native activating mechanisms are sufficient to overrule the locally silenced state.

5.2.1 Stability of targeted DNA methylation

In contrast to our results, in a previous report the introduction of targeted DNA methylation in the *Maspin* promoter led to stable gene repression over multiple generations (Rivenbark et al., 2012). The dissonance between the two studies could be explained by the chromatin context of the two loci tested, with *Maspin* perhaps being more permissive to silencing than *VEGF-A*. This is supported by the lower level of DNA methylation needed to sufficiently silence *Maspin* (<10-15% methylation level of 2-6 CpGs in the case of *Maspin*, compared to >50% methylation levels of at least 12 CpGs in the case of *VEGF-A*) . Another explanation could lie in the difference of the technical approach employed by both studies. Namely, stable cell lines expressing the chimeric DNA methyltranferase were generated and used in the *Maspin* study,

while in our case adenoviral delivery was used for transient expression of the chimeric DNA methyltransferase. Although it was found that the levels of the chimeric DNA methyltransferase were drastically reduced after several generations in the stable cell lines, it is still conceivable that a small amout of the effector might have been still present in the cells, sufficient for stable maintenance of the signal. On a similar note, the expression of our effector peaked at day 5, which could be sufficient for induction of repression, but insufficient for modulation of the feedback loop that leads to stable maintenance. It stands to reason that more time of effector expression might be necessary for stable propagation of DNA methylation. This could be solved with optimized inducible systems, where the temporal expression of the targeted effector methyltransferase can be controled with an inducer.

5.2.2 Targeted histone H3K9 methylation

A few studies prior to the work reported here, have successfully applied chimeric zinc finger targeted histone methyltransferases in epigenome editing approaches (Falahi et al., 2013; Ragunathan et al., 2015; Snowden et al., 2002). Apart from reaffirming the general validity of this approach, we also systematically monitored the kinetics of the installed H3K9me2/3 modification. We showed that zinc finger targeted installment of H3K9me2/3 led to spreading of the H3K9 methylation signal over around 20 kb from the nucleation site, with no crosstalk to DNA methylation.

Our observations of targeted local heterochromatization and spreading are in accord with previous studies (Hathaway et al., 2012; Ragunathan et al., 2015). In the first study, the authors used a system capable of chemically induced recruitment of the HP1 chromo shadow domain to an artificial promoter. The HP1 chromo shadow domain can recruit the Suv39H1/2 H3K9 trimethyltransferase and propagate methylation. The presence of the chemical inducer for 3 to 7 days led to transient deposition of H3K9me3 and gene silencing, but only after a prolonged stimulus of 4.5 weeks, a gain in DNA methylation was recorded, again indicating that the temporal presence of the trigger stimulus is of critical importance. In the second study, the authors successfully applied TetR fused Clr4 methyltransferase to ectopically induce H3K9me2 heterochromatin domains in *S. pombe*. This interesting study provided an insightful mechanistic overview of the innerworkings of chromatin. First of all, the authors determined that HP1 proteins and histone deacetylases, but not RNAi are needed for establishment of heterochromatin domains, which act downstream of H3K9 methylation. Then, by ChIP-seq experiments it was determined that the ectopically induced heterochromatin was propagated around 40-50 kb from the nucleation site and was maintained for 24 hours (corresponding to around 10 cell divisions). Finally, they achieved long-lasting maintenance of the induced heterochromatin state by removal of a putative demethylase. Interestingly, this is in accord with our results, in which we have observed different kinetics of loss of H3K9me2 (slower than H3K9me3) and H3K9me3. This is highly indicative of an active demethylation processes taking place inside the cells, which would demethylate H3K9me3 via me2 and me1 intermediates. In support of this, our data also suggest that the potential demethylase activity is coupled with enhanced recruitment of histone acetyltransferase machineries, which helps in the re-establishment of the native chromatin state of the locus and its transcriptional output. In summary, our and the work performed in the Ragunathan study demonstrate the intricate interplay between different chromatin machineries and suggest that similar networks are likely to exist in mammalian cells.

5.2.3 Conclusion and pespectives

The data obtained in our experimental system and the studies discussed above, collectively indicate that chromatin states can be successfully edited by setting a trigger "repressive mark" or removing an "activating enhancer mark" (de Groote et al., 2012; Mendenhall et al., 2013) followed by transcriptional modulation. Likewise, it would be curious to examine the functional effects and repercussions of targeted installment of "activating" chromatin marks, such as H3K4me3 or histone acetylation.

The establishment of an ectopic chromatin domain is only the first step of chromatin refurbishment. The above data strongly indicate that maintenance and mitotic inheritance of an "epigenetic" mark, even DNA methylation might not be a trivial task. This is in line with recent observations suggesting that the elegant model of DNA methylation might be an oversimplification (Jeltsch and Jurkowska, 2014; Shipony et al., 2014). Our current understanding of epigenetics argues that epigenetic states are encoded in a network of epigenetic marks and epigenetic factors that cannot be reset easily (Takahashi and Yamanaka, 2006). What can be done to achieve stable epigenetic changes that would be stably propagated and maintained? First, longer-lasting trigger stimuli alone or in concert with histone and/or DNA demethylase inhibitors or siRNAs against demethylases might be able to overcome the tendency of the cell to return to its native state and maintain the introduced epigenetic state. Second, multivalent deposition of functionally related epigenetic functionalities could enhance the effect of the installed epigenetic state and increase the chance of maintaining it. For instance, co-targeting of DNA and histone H3K9me2/3 methyltransferases might have a synergistic effect on setting a stable repressive chromatin state. Similarly, co-targeting of an H3K4me3 demethylase (such as KDM5B) along with DNA or H3K9me2/3 methyltransferase could significantly aid the deconstruction of epigenetic networks and permanent epigenetic effect. In contrast, it might be possible to activate polycomb H3K27me3 repressed genes by the concerted action of the targeted UTX (H3K27 demethylase) and synthetically fused or *in vivo* recruited MLL complexes. Third, the stability of the epigenetic effects might be dependent on the local chromatin context. With the development of novel epigenome targeting technologies, it will be compelling to observe the effects of setting or erasing single epigenetic marks or their combinations in different chromatin contexts. One way to achieve this is by introduction of promiscuous targeting modules fused with different chromatin modifiers. This could be supported by genome-wide investigation of the effects on different genomic regions at different time points coupled with RNA-seq and ChIP-seq. These powerful experiments could drastically assist the dissection of the players involved in the maintenance of chromatin states, such as the presence of crucial histone modifications and the level of initial transcription. Collectively, this approach could provide a valuable insight into the chromatin regulatory logic in native settings and could pave the way into genuine and heritable epigenetic editing.

Based on the above line of reasoning, it is likely that essential cellular processes such as imprinting, heterochromatization or transcription of housekeeping genes are continuously maintained with epigenetic modulators that preserve the already established epigenetic state. An elegant study, has recently unearthed the logic of regulation and behavior of many chromatin modifiers (Keung et al., 2014) with certain modalities able to repress single genes, multiple genes or entire genomic regions with different spatial modes of binding.

Finally, the real question is: what is the value of epigenetic editing? In basic research, it can serve as a valuable experimental tool for studying the behavior and response of the chromatin network to different epigenomic perturbations, as elaborated on above. This would lead to

novel insight with unprecedented level of detail. The knowledge acquired from these fundamental studies of chromatin regulatory logic can be applied in clinical settings for treatment or prevention of cancer and other diseases. For example, the expression levels of tumor-suppressor or oncogenes can be modulated to native levels. Similar approaches can be applied for treatment of viral infections, by targeting viral genes or critical viral receptors or in any other disease state where epigenetic modulation and tinkering is necessary.

All in all, the field of epigenetic editing is still in its infancy. We anticipate that in the next decade, this approach would continue developing with a staggering pace. In line with this, we are optimistic that epigenetic editing would be widely embraced by the scientific community and it will result in many important scientific discoveries.

6. References

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Molecular cell *9*, 1191-1200.

Ai, H.W., Lee, J.W., and Schultz, P.G. (2010). A method to site-specifically introduce methyllysine into proteins in E. coli. Chem Commun (Camb) *46*, 5506-5508.

Balakrishnan, L., and Milavetz, B. (2010). Decoding the histone H4 lysine 20 methylation mark. Critical reviews in biochemistry and molecular biology *45*, 440-452.

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

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

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

Bashtrykov, P., Jankevicius, G., Smarandache, A., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2012). Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain. Chemistry & biology *19*, 572-578.

Basu, A., Rose, K.L., Zhang, J., Beavis, R.C., Ueberheide, B., Garcia, B.A., Chait, B., Zhao, Y., Hunt, D.F., Segal, E., *et al.* (2009). Proteome-wide prediction of acetylation substrates. Proceedings of the National Academy of Sciences of the United States of America *106*, 13785-13790.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schubeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Gene Dev *23*, 781-783.

Berndsen, C.E., and Denu, J.M. (2008). Catalysis and substrate selection by histone/protein lysine acetyltransferases. Current opinion in structural biology *18*, 682-689.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Molecular and cellular biology *26*, 2560-2569.

Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L., Ito, S., Cooper, S., Kondo, K., Koseki, Y., *et al.* (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell *157*, 1445-1459.

References

Blackledge, N.P., Long, H.K., Zhou, J.C., Kriaucionis, S., Patient, R., and Klose, R.J. (2012). Bio-CAP: a versatile and highly sensitive technique to purify and characterise regions of non-methylated DNA. Nucleic Acids Res *40*, e32.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. Current protocols in molecular biology / edited by Frederick M Ausubel [et al] *Chapter 19*, Unit 19 10 11-21.

Bock, I., Dhayalan, A., Kudithipudi, S., Brandt, O., Rathert, P., and Jeltsch, A. (2011a). Detailed specificity analysis of antibodies binding to modified histone tails with peptide arrays. Epigenetics : official journal of the DNA Methylation Society *6*, 256-263.

Bock, I., Kudithipudi, S., Tamas, R., Kungulovski, G., Dhayalan, A., and Jeltsch, A. (2011b). Application of Celluspots peptide arrays for the analysis of the binding specificity of epigenetic reading domains to modified histone tails. BMC biochemistry *12*, 48.

Briggs, S.D., Xiao, T., Sun, Z.W., Caldwell, J.A., Shabanowitz, J., Hunt, D.F., Allis, C.D., and Strahl, B.D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. Nature *418*, 498.

Britton, L.M., Gonzales-Cope, M., Zee, B.M., and Garcia, B.A. (2011). Breaking the histone code with quantitative mass spectrometry. Expert review of proteomics *8*, 631-643.

Bua, D.J., Kuo, A.J., Cheung, P., Liu, C.L., Migliori, V., Espejo, A., Casadio, F., Bassi, C., Amati, B., Bedford, M.T., *et al.* (2009). Epigenome microarray platform for proteome-wide dissection of chromatin-signaling networks. PloS one *4*, e6789.

Butler, M.G. (2009). Genomic imprinting disorders in humans: a mini-review. Journal of assisted reproduction and genetics *26*, 477-486.

Byrum, S.D., Raman, A., Taverna, S.D., and Tackett, A.J. (2012). ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. Cell reports *2*, 198-205.

Byrum, S.D., Taverna, S.D., and Tackett, A.J. (2013). Purification of a specific native genomic locus for proteomic analysis. Nucleic Acids Res *41*, e195.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Molecular cell *49*, 825-837.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., *et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. Cell *157*, 77-94.

Chadwick, B.P., and Willard, H.F. (2001). A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. The Journal of cell biology *152*, 375-384.

References

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

Chakravarthy, S., and Luger, K. (2006). The histone variant macro-H2A preferentially forms "hybrid nucleosomes". J Biol Chem *281*, 25522-25531.

Chalker, J.M., Lercher, L., Rose, N.R., Schofield, C.J., and Davis, B.G. (2012). Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications. Angew Chem Int Ed Engl *51*, 1835-1839.

Chang, Y., Horton, J.R., Bedford, M.T., Zhang, X., and Cheng, X. (2011). Structural insights for MPP8 chromodomain interaction with histone H3 lysine 9: potential effect of phosphorylation on methyl-lysine binding. Journal of molecular biology *408*, 807-814.

Chen, H., Kazemier, H.G., de Groote, M.L., Ruiters, M.H., Xu, G.L., and Rots, M.G. (2014). Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. Nucleic Acids Res *42*, 1563-1574.

Chen, P., Zhao, J., and Li, G. (2013). Histone variants in development and diseases. Journal of genetics and genomics = Yi chuan xue bao *40*, 355-365.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Molecular and cellular biology *23*, 5594-5605.

Chiang, C.M. (2009). Brd4 engagement from chromatin targeting to transcriptional regulation: selective contact with acetylated histone H3 and H4. F1000 biology reports *1*, 98.

Conerly, M.L., Teves, S.S., Diolaiti, D., Ulrich, M., Eisenman, R.N., and Henikoff, S. (2010). Changes in H2A.Z occupancy and DNA methylation during B-cell lymphomagenesis. Genome research *20*, 1383-1390.

Costanzi, C., and Pehrson, J.R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599-601.

Cross, S.H., Charlton, J.A., Nan, X., and Bird, A.P. (1994). Purification of CpG islands using a methylated DNA binding column. Nature genetics *6*, 236-244.

Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B. (1994). Synthesis of proteins by native chemical ligation. Science *266*, 776-779.

de Groote, M.L., Verschure, P.J., and Rots, M.G. (2012). Epigenetic Editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. Nucleic Acids Res *40*, 10596-10613.

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

Deal, R.B., and Henikoff, S. (2010). Catching a glimpse of nucleosome dynamics. Cell Cycle *9*, 3389-3390.

References

Dejardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Denis, H., Ndlovu, M.N., and Fuks, F. (2011). Regulation of mammalian DNA methyltransferases: a route to new mechanisms. EMBO reports *12*, 647-656.

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

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J Biol Chem *285*, 26114-26120.

Dhayalan, A., Tamas, R., Bock, I., Tattermusch, A., Dimitrova, E., Kudithipudi, S., Ragozin, S., and Jeltsch, A. (2011). The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. Human molecular genetics *20*, 2195-2203.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nature structural & molecular biology *20*, 1147-1155.

Dion, M.F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., and Rando, O.J. (2007). Dynamics of replication-independent histone turnover in budding yeast. Science *315*, 1405-1408.

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

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

Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem *277*, 28368-28371.

Drane, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. Genes Dev *24*, 1253-1265.

Earnshaw, W.C., and Rothfield, N. (1985). Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma *91*, 313-321.

Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A., Cheung, M.S., Day, D.S., Gadel, S., Gorchakov, A.A., *et al.* (2011). An assessment of histone-modification antibody quality. Nature structural & molecular biology *18*, 91-93.

References

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., *et al.* (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43-49.

Evertts, A.G., Manning, A.L., Wang, X., Dyson, N.J., Garcia, B.A., and Coller, H.A. (2013). H4K20 methylation regulates quiescence and chromatin compaction. Molecular biology of the cell *24*, 3025-3037.

Falahi, F., Huisman, C., Kazemier, H.G., van der Vlies, P., Kok, K., Hospers, G.A., and Rots, M.G. (2013). Towards sustained silencing of HER2/neu in cancer by epigenetic editing. Molecular cancer research : MCR *11*, 1029-1039.

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

Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stutzer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014). Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Molecular cell *53*, 49-62.

Festenstein, R., and Chan, J.P. (2012). Context is everything: activators can also repress. Nature structural & molecular biology *19*, 973-975.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., *et al.* (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212-224.

Filipescu, D., Szenker, E., and Almouzni, G. (2013). Developmental roles of histone H3 variants and their chaperones. Trends in genetics : TIG *29*, 630-640.

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

Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L., and Strahl, B.D. (2011). Influence of combinatorial histone modifications on antibody and effector protein recognition. Current biology : CB *21*, 53-58.

Gamble, M.J., Frizzell, K.M., Yang, C., Krishnakumar, R., and Kraus, W.L. (2010). The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. Genes Dev *24*, 21-32.

Garske, A.L., Craciun, G., and Denu, J.M. (2008). A combinatorial H4 tail library for exploring the histone code. Biochemistry *47*, 8094-8102.

Garske, A.L., Oliver, S.S., Wagner, E.K., Musselman, C.A., LeRoy, G., Garcia, B.A., Kutateladze, T.G., and Denu, J.M. (2010). Combinatorial profiling of chromatin binding modules reveals multisite discrimination. Nature chemical biology *6*, 283-290.

Gaydos, L.J., Wang, W., and Strome, S. (2014). Gene repression. H3K27me and PRC2 transmit a memory of repression across generations and during development. Science *345*, 1515-1518.

References

Giannattasio, M., Lazzaro, F., Plevani, P., and Muzi-Falconi, M. (2005). The DNA damage checkpoint response requires histone H2B ubiquitination by Rad6-Bre1 and H3 methylation by Dot1. J Biol Chem *280*, 9879-9886.

Gibcus, J.H., and Dekker, J. (2013). The hierarchy of the 3D genome. Molecular cell *49*, 773-782.

Gilmour, D.S., and Lis, J.T. (1985). In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. Molecular and cellular biology *5*, 2009-2018.

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: A landscape takes shape. Cell *128*, 635-638.

Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., *et al.* (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell *140*, 678-691.

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

Grigoryev, S.A., and Woodcock, C.L. (2012). Chromatin organization - the 30 nm fiber. Experimental cell research *318*, 1448-1455.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Halachev, K., Bast, H., Albrecht, F., Lengauer, T., and Bock, C. (2012). EpiExplorer: live exploration and global analysis of large epigenomic datasets. Genome biology *13*, R96.

Han, W., Li, X., and Fu, X. (2011). The macro domain protein family: structure, functions, and their potential therapeutic implications. Mutation research *727*, 86-103.

Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of heterochromatin in living cells. Cell *149*, 1447-1460.

Hattori, T., Taft, J.M., Swist, K.M., Luo, H., Witt, H., Slattery, M., Koide, A., Ruthenburg, A.J., Krajewski, K., Strahl, B.D., *et al.* (2013). Recombinant antibodies to histone post-translational modifications. Nature methods *10*, 992-995.

Hayashi-Takanaka, Y., Yamagata, K., Nozaki, N., and Kimura, H. (2009). Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. The Journal of cell biology *187*, 781-790.

Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T.J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H., Nozaki, N., *et al.* (2011). Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. Nucleic Acids Res *39*, 6475-6488.

Henikoff, S., and Shilatifard, A. (2011). Histone modification: cause or cog? Trends in genetics : TIG *27*, 389-396.

References

Heubach, Y., Planatscher, H., Sommersdorf, C., Maisch, D., Maier, J., Joos, T.O., Templin, M.F., and Poetz, O. (2013). From spots to beads-PTM-peptide bead arrays for the characterization of anti-histone antibodies. Proteomics *13*, 1010-1015.

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nature reviews Genetics *16*, 71-84.

Horikoshi, M. (2013). Histone acetylation: from code to web and router via intrinsically disordered regions. Current pharmaceutical design *19*, 5019-5042.

Huang, H., and Jiao, R. (2012). Roles of chromatin assembly factor 1 in the epigenetic control of chromatin plasticity. Science China Life sciences *55*, 15-19.

Huang, R., Holbert, M.A., Tarrant, M.K., Curtet, S., Colquhoun, D.R., Dancy, B.M., Dancy, B.C., Hwang, Y., Tang, Y., Meeth, K., *et al.* (2010). Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation. Journal of the American Chemical Society *132*, 9986-9987.

Ito, T., Umehara, T., Sasaki, K., Nakamura, Y., Nishino, N., Terada, T., Shirouzu, M., Padmanabhan, B., Yokoyama, S., Ito, A., *et al.* (2011). Real-time imaging of histone H4K12-specific acetylation determines the modes of action of histone deacetylase and bromodomain inhibitors. Chemistry & biology *18*, 495-507.

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P.W., Cochrane, J.C., Allis, C.D., Picketts, D.J., Patel, D.J., Li, H., *et al.* (2011). ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. Nature structural & molecular biology *18*, 769-776.

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

Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem : a European journal of chemical biology *3*, 274-293.

Jeltsch, A., and Jurkowska, R.Z. (2014). New concepts in DNA methylation. Trends in biochemical sciences *39*, 310-318.

Jeltsch, A., and Lanio, T. (2002). Site-directed mutagenesis by polymerase chain reaction. Methods Mol Biol *182*, 85-94.

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

Jia, Y., Lisch, D.R., Ohtsu, K., Scanlon, M.J., Nettleton, D., and Schnable, P.S. (2009). Loss of RNA-dependent RNA polymerase 2 (RDR2) function causes widespread and unexpected changes in the expression of transposons, genes, and 24-nt small RNAs. PLoS genetics *5*, e1000737.

Jin, B., and Robertson, K.D. (2013). DNA methyltransferases, DNA damage repair, and cancer. Advances in experimental medicine and biology *754*, 3-29.

References

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev *21*, 1519-1529.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nature genetics *41*, 941-945.

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

Jorgensen, S., Schotta, G., and Sorensen, C.S. (2013). Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. Nucleic Acids Res *41*, 2797-2806.

Joshi, A.A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Molecular cell *20*, 971-978.

Josling, G.A., Selvarajah, S.A., Petter, M., and Duffy, M.F. (2012). The role of bromodomain proteins in regulating gene expression. Genes *3*, 320-343.

Jurkowska, R.Z. (2008). Biochemical characterization of the mammalian Dnmt3 family of DNA methyltransferases In School of Engineering and Science (Bremen, Germany: Jacobs University Bremen).

Jurkowska, R.Z., and Jeltsch, A. (2013). Genomic imprinting--the struggle of the genders at the molecular level. Angew Chem Int Ed Engl *52*, 13524-13536.

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

Kallio, M.A., Tuimala, J.T., Hupponen, T., Klemela, P., Gentile, M., Scheinin, I., Koski, M., Kaki, J., and Korpelainen, E.I. (2011). Chipster: user-friendly analysis software for microarray and other high-throughput data. BMC genomics *12*, 507.

Kanno, T., Kanno, Y., Siegel, R.M., Jang, M.K., Lenardo, M.J., and Ozato, K. (2004). Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. Molecular cell *13*, 33-43.

Karch, K.R., Denizio, J.E., Black, B.E., and Garcia, B.A. (2013). Identification and interrogation of combinatorial histone modifications. Frontiers in genetics *4*, 264.

Kerppola, T.K. (2008). Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. Annual review of biophysics *37*, 465-487.

Keung, A.J., Bashor, C.J., Kiriakov, S., Collins, J.J., and Khalil, A.S. (2014). Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. Cell *158*, 110-120.

Kim, S., Kim, H., Fong, N., Erickson, B., and Bentley, D.L. (2011). Pre-mRNA splicing is a determinant of histone H3K36 methylation. Proceedings of the National Academy of Sciences of the United States of America *108*, 13564-13569.

References

Kirmizis, A., Santos-Rosa, H., Penkett, C.J., Singer, M.A., Vermeulen, M., Mann, M., Bahler, J., Green, R.D., and Kouzarides, T. (2007). Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. Nature *449*, 928-932.

Klose, R.J., and Zhang, Y. (2007). Regulation of histone methylation by demethylimination and demethylation. Nature reviews Molecular cell biology *8*, 307-318.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nature genetics *41*, 376-381.

Kouskouti, A., and Talianidis, I. (2005). Histone modifications defining active genes persist after transcriptional and mitotic inactivation. The EMBO journal *24*, 347-357.

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

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

Kungulovski, G., Kycia, I., Tamas, R., Jurkowska, R.Z., Kudithipudi, S., Henry, C., Reinhardt, R., Labhart, P., and Jeltsch, A. (2014). Application of histone modification-specific interaction domains as an alternative to antibodies. Genome research *24*, 1842-1853.

Kungulovski, G., Nunna, S., Thomas, M., Zanger, M.U., Reinhardt, R., and Jeltsch, A. (2015). Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. Epigenetics & chromatin *8*.

Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. Cell *117*, 721-733.

Kycia, I., Kudithipudi, S., Tamas, R., Kungulovski, G., Dhayalan, A., and Jeltsch, A. (2014). The Tudor domain of the PHD finger protein 1 is a dual reader of lysine trimethylation at lysine 36 of histone H3 and lysine 27 of histone variant H3t. Journal of molecular biology *426*, 1651-1660.

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

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., *et al.* (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome research *22*, 1813-1831.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.

Lauberth, S.M., Nakayama, T., Wu, X., Ferris, A.L., Tang, Z., Hughes, S.H., and Roeder, R.G. (2013). H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. Cell *152*, 1021-1036.

References

Le, D.D., Cortesi, A.T., Myers, S.A., Burlingame, A.L., and Fujimori, D.G. (2013). Site-specific and regiospecific installation of methylarginine analogues into recombinant histones and insights into effector protein binding. Journal of the American Chemical Society *135*, 2879-2882.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Current opinion in genetics & development *21*, 175-186.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Lin, C.W., Jao, C.Y., and Ting, A.Y. (2004). Genetically encoded fluorescent reporters of histone methylation in living cells. Journal of the American Chemical Society *126*, 5982-5983.

Lin, C.W., and Ting, A.Y. (2004). A genetically encoded fluorescent reporter of histone phosphorylation in living cells. Angew Chem Int Ed Engl *43*, 2940-2943.

Liu, T., Ortiz, J.A., Taing, L., Meyer, C.A., Lee, B., Zhang, Y., Shin, H., Wong, S.S., Ma, J., Lei, Y., *et al.* (2011). Cistrome: an integrative platform for transcriptional regulation studies. Genome biology *12*, R83.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 angstrom 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.

Luo, J., Deng, Z.L., Luo, X., Tang, N., Song, W.X., Chen, J., Sharff, K.A., Luu, H.H., Haydon, R.C., Kinzler, K.W., *et al.* (2007). A protocol for rapid generation of recombinant adenoviruses using the AdEasy system. Nature protocols *2*, 1236-1247.

Macdonald, N., Welburn, J.P., Noble, M.E., Nguyen, A., Yaffe, M.B., Clynes, D., Moggs, J.G., Orphanides, G., Thomson, S., Edmunds, J.W., *et al.* (2005). Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. Molecular cell *20*, 199-211.

Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E., *et al.* (2013). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nature biotechnology *31*, 1137-1142.

Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? Current opinion in cell biology *22*, 291-297.

Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nature reviews Genetics *11*, 285-296.

Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? Current opinion in genetics & development *15*, 163-176.

References

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nature reviews Molecular cell biology *6*, 838-849.

Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. Current opinion in cell biology *14*, 692-699.

Mauser, R. (2013). Generation and characterization of synthetic chromatin interacting domains with double specificity. In Institute of Biochemistry (Stuttgart: University of Stuttgart).

Mendenhall, E.M., Williamson, K.E., Reyon, D., Zou, J.Y., Ram, O., Joung, J.K., and Bernstein, B.E. (2013). Locus-specific editing of histone modifications at endogenous enhancers. Nature biotechnology *31*, 1133-1136.

Mercer, T.R., and Mattick, J.S. (2013). Structure and function of long noncoding RNAs in epigenetic regulation. Nature structural & molecular biology *20*, 300-307.

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

Miranda, T.B., and Jones, P.A. (2007). DNA methylation: the nuts and bolts of repression. Journal of cellular physiology *213*, 384-390.

Moore, K.E., Carlson, S.M., Camp, N.D., Cheung, P., James, R.G., Chua, K.F., Wolf-Yadlin, A., and Gozani, O. (2013). A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. Molecular cell *50*, 444-456.

Morrison, A.J., and Shen, X. (2005). DNA repair in the context of chromatin. Cell Cycle *4*, 568-571.

Muir, T.W., Sondhi, D., and Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. Proceedings of the National Academy of Sciences of the United States of America *95*, 6705-6710.

Mujtaba, S., Zeng, L., and Zhou, M.M. (2007). Structure and acetyl-lysine recognition of the bromodomain. Oncogene *26*, 5521-5527.

Musselman, C.A., Lalonde, M.E., Cote, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. Nature structural & molecular biology *19*, 1218-1227.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

Nekrasov, M., Amrichova, J., Parker, B.J., Soboleva, T.A., Jack, C., Williams, R., Huttley, G.A., and Tremethick, D.J. (2012). Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics. Nature structural & molecular biology *19*, 1076-1083.

Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008). Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. Nature chemical biology *4*, 232-234.

References

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Molecular cell *11*, 709-719.

Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. Genes Dev *25*, 1345-1358.

Nguyen, D.P., Garcia Alai, M.M., Kapadnis, P.B., Neumann, H., and Chin, J.W. (2009). Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones. Journal of the American Chemical Society *131*, 14194-14195.

Nguyen, D.P., Garcia Alai, M.M., Virdee, S., and Chin, J.W. (2010). Genetically directing varepsilon-N, N-dimethyl-L-lysine in recombinant histones. Chemistry & biology *17*, 1072-1076.

Nguyen, U.T., Bittova, L., Muller, M.M., Fierz, B., David, Y., Houck-Loomis, B., Feng, V., Dann, G.P., and Muir, T.W. (2014). Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. Nature methods *11*, 834-840.

Nielsen, P.R., Nietlispach, D., Mott, H.R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A.G., Murzina, N.V., and Laue, E.D. (2002). Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature *416*, 103-107.

Nishikori, S., Hattori, T., Fuchs, S.M., Yasui, N., Wojcik, J., Koide, A., Strahl, B.D., and Koide, S. (2012). Broad Ranges of Affinity and Specificity of Anti-Histone Antibodies Revealed by a Quantitative Peptide Immunoprecipitation Assay. Journal of molecular biology *424*, 391-399.

Noma, K., Allis, C.D., and Grewal, S.I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science *293*, 1150-1155.

Nora, E.P., Dekker, J., and Heard, E. (2013). Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? BioEssays : news and reviews in molecular, cellular and developmental biology *35*, 818-828.

Nunna, S., Reinhardt, R., Ragozin, S., and Jeltsch, A. (2014). Targeted methylation of the epithelial cell adhesion molecule (EpCAM) promoter to silence its expression in ovarian cancer cells. PloS one *9*, e87703.

Olins, D.E., and Olins, A.L. (1972). Physical studies of isolated eucaryotic nuclei. The Journal of cell biology *53*, 715-736.

Owen, D.J., Ornaghi, P., Yang, J.C., Lowe, N., Evans, P.R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A.A. (2000). The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. The EMBO journal *19*, 6141-6149.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proceedings of the National Academy of Sciences of the United States of America *88*, 3734-3738.

References

Papamichos-Chronakis, M., and Peterson, C.L. (2013). Chromatin and the genome integrity network. Nature reviews Genetics *14*, 62-75.

Papamichos-Chronakis, M., Watanabe, S., Rando, O.J., and Peterson, C.L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. Cell *144*, 200-213.

Papworth, M., Kolasinska, P., and Minczuk, M. (2006). Designer zinc-finger proteins and their applications. Gene *366*, 27-38.

Patel, D.J., and Wang, Z. (2013). Readout of epigenetic modifications. Annual review of biochemistry *82*, 81-118.

Payer, B., and Lee, J.T. (2008). X chromosome dosage compensation: how mammals keep the balance. Annual review of genetics *42*, 733-772.

Peach, S.E., Rudomin, E.L., Udeshi, N.D., Carr, S.A., and Jaffe, J.D. (2012). Quantitative Assessment of Chromatin Immunoprecipitation Grade Antibodies Directed against Histone Modifications Reveals Patterns of Co-occurring Marks on Histone Protein Molecules. Molecular & Cellular Proteomics *11*, 128-137.

Pesavento, J.J., Yang, H., Kelleher, N.L., and Mizzen, C.A. (2008). Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. Molecular and cellular biology *28*, 468-486.

Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., *et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Molecular cell *12*, 1577-1589.

Petruk, S., Sedkov, Y., Johnston, D.M., Hodgson, J.W., Black, K.L., Kovermann, S.K., Beck, S., Canaani, E., Brock, H.W., and Mazo, A. (2012). TrxG and PcG proteins but not methylated histones remain associated with DNA through replication. Cell *150*, 922-933.

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science *300*, 131-135.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., *et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nature biotechnology *28*, 1057-1068.

Prezioso, C., and Orlando, V. (2011). Polycomb proteins in mammalian cell differentiation and plasticity. FEBS letters *585*, 2067-2077.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

References

Radman-Livaja, M., Verzijlbergen, K.F., Weiner, A., van Welsem, T., Friedman, N., Rando, O.J., and van Leeuwen, F. (2011). Patterns and mechanisms of ancestral histone protein inheritance in budding yeast. PLoS biology *9*, e1001075.

Ragunathan, K., Jih, G., and Moazed, D. (2015). Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. Science *348*, 1258699.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res *42*, W187-191.

Ranjan, A., Mizuguchi, G., FitzGerald, P.C., Wei, D., Wang, F., Huang, Y., Luk, E., Woodcock, C.L., and Wu, C. (2013). Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. Cell *154*, 1232-1245.

Rathert, P., Dhayalan, A., Ma, H., and Jeltsch, A. (2008). Specificity of protein lysine methyltransferases and methods for detection of lysine methylation of non-histone proteins. Molecular bioSystems *4*, 1186-1190.

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

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Molecular cell *12*, 1591-1598.

Riggs, A.D. (1996). Epigenetic Mechanisms of Gene Regulation, Vol 32 (Cold Spring Harbor Monograph Archive).

Rivenbark, A.G., Stolzenburg, S., Beltran, A.S., Yuan, X., Rots, M.G., Strahl, B.D., and Blancafort, P. (2012). Epigenetic reprogramming of cancer cells via targeted DNA methylation. Epigenetics : official journal of the DNA Methylation Society *7*, 350-360.

Rivera, C., Gurard-Levin, Z.A., Almouzni, G., and Loyola, A. (2014). Histone lysine methylation and chromatin replication. Biochimica et biophysica acta *1839*, 1433-1439.

Robinson, P.J., and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Current opinion in structural biology *16*, 336-343.

Rose, N.R., and Klose, R.J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. Biochimica et biophysica acta *1839*, 1362-1372.

Rufiange, A., Jacques, P.E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. Molecular cell *27*, 393-405.

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.

References

Sarcinella, E., Zuzarte, P.C., Lau, P.N., Draker, R., and Cheung, P. (2007). Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. Molecular and cellular biology *27*, 6457-6468.

Sasaki, K., Ito, T., Nishino, N., Khochbin, S., and Yoshida, M. (2009). Real-time imaging of histone H4 hyperacetylation in living cells. Proceedings of the National Academy of Sciences of the United States of America *106*, 16257-16262.

Sato, Y., Mukai, M., Ueda, J., Muraki, M., Stasevich, T.J., Horikoshi, N., Kujirai, T., Kita, H., Kimura, T., Hira, S., *et al.* (2013). Genetically encoded system to track histone modification in vivo. Scientific reports *3*, 2436.

Schneider, J., Wood, A., Lee, J.S., Schuster, R., Dueker, J., Maguire, C., Swanson, S.K., Florens, L., Washburn, M.P., and Shilatifard, A. (2005). Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. Molecular cell *19*, 849-856.

Schneiderman, J.I., Orsi, G.A., Hughes, K.T., Loppin, B., and Ahmad, K. (2012). Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant. Proceedings of the National Academy of Sciences of the United States of America *109*, 19721-19726.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev *18*, 1251-1262.

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

Sexton, T., Kurukuti, S., Mitchell, J.A., Umlauf, D., Nagano, T., and Fraser, P. (2012). Sensitive detection of chromatin coassociations using enhanced chromosome conformation capture on chip. Nature protocols *7*, 1335-1350.

Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. Nature protocols *2*, 1445-1457.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., *et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M.R., *et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annual review of biochemistry *75*, 243-269.

Shilatifard, A. (2008). Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. Current opinion in cell biology *20*, 341-348.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual review of biochemistry *81*, 65-95.

References

Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. Bioinformatics *25*, 2605-2606.

Shipony, Z., Mukamel, Z., Cohen, N.M., Landan, G., Chomsky, E., Zeliger, S.R., Fried, Y.C., Ainbinder, E., Friedman, N., and Tanay, A. (2014). Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. Nature *513*, 115-119.

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

Siddique, A.N., Nunna, S., Rajavelu, A., Zhang, Y., Jurkowska, R.Z., Reinhardt, R., Rots, M.G., Ragozin, S., Jurkowski, T.P., and Jeltsch, A. (2013). Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. Journal of molecular biology *425*, 479-491.

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.

Snowden, A.W., Gregory, P.D., Case, C.C., and Pabo, C.O. (2002). Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. Current biology : CB *12*, 2159-2166.

Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J. (2012). A unique H2A histone variant occupies the transcriptional start site of active genes. Nature structural & molecular biology *19*, 25-30.

Soldi, M., and Bonaldi, T. (2013). The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components. Molecular & cellular proteomics : MCP *12*, 764-780.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P., and Li, G. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science *344*, 376-380.

Strahl, B.D., Ohba, R., Cook, R.G., and Allis, C.D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proceedings of the National Academy of Sciences of the United States of America *96*, 14967-14972.

Su, Z., Boersma, M.D., Lee, J.H., Oliver, S.S., Liu, S., Garcia, B.A., and Denu, J.M. (2014). ChIP-less analysis of chromatin states. Epigenetics & chromatin *7*, 7.

Sun, Z.W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Suva, M.L., Riggi, N., and Bernstein, B.E. (2013). Epigenetic reprogramming in cancer. Science *339*, 1567-1570.

References

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. The EMBO journal *27*, 2681-2690.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., *et al.* (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell *146*, 1016-1028.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature structural & molecular biology *14*, 1025-1040.

Timp, W., and Feinberg, A.P. (2013). Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. Nature reviews Cancer *13*, 497-510.

Tolstorukov, M.Y., Goldman, J.A., Gilbert, C., Ogryzko, V., Kingston, R.E., and Park, P.J. (2012). Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. Molecular cell *47*, 596-607.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols *7*, 562-578.

Travers, A. (2014). Structural biology. The 30-nm fiber redux. Science *344*, 370-372.

Tropberger, P., and Schneider, R. (2013). Scratching the (lateral) surface of chromatin regulation by histone modifications. Nature structural & molecular biology *20*, 657-661.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. The EMBO journal *13*, 3822-3831.

van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell *109*, 745-756.

Verdel, A., and Moazed, D. (2005). RNAi-directed assembly of heterochromatin in fission yeast. FEBS letters *579*, 5872-5878.

Vettese-Dadey, M., Grant, P.A., Hebbes, T.R., Crane- Robinson, C., Allis, C.D., and Workman, J.L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. The EMBO journal *15*, 2508-2518.

Voigt, P., LeRoy, G., Drury, W.J., 3rd, Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically modified nucleosomes. Cell *151*, 181-193.

Volkel, P., and Angrand, P.O. (2007). The control of histone lysine methylation in epigenetic regulation. Biochimie *89*, 1-20.

References

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. Nature *150*, 563-565.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nature reviews Molecular cell biology *13*, 115-126.

Weber, C.M., Henikoff, J.G., and Henikoff, S. (2010). H2A.Z nucleosomes enriched over active genes are homotypic. Nature structural & molecular biology *17*, 1500-1507.

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

Widom, J., Finch, J.T., and Thomas, J.O. (1985). Higher-order structure of long repeat chromatin. The EMBO journal *4*, 3189-3194.

Wood, A., Krogan, N.J., Dover, J., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J.F., *et al.* (2003). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Molecular cell *11*, 267-274.

Wu, H., Zeng, H., Lam, R., Tempel, W., Amaya, M.F., Xu, C., Dombrovski, L., Qiu, W., Wang, Y., and Min, J. (2011). Structural and histone binding ability characterizations of human PWWP domains. PloS one *6*, e18919.

Wu, R., Terry, A.V., Singh, P.B., and Gilbert, D.M. (2005). Differential subnuclear localization and replication timing of histone H3 lysine 9 methylation states. Molecular biology of the cell *16*, 2872-2881.

Wysocki, R., Javaheri, A., Allard, S., Sha, F., Cote, J., and Kron, S.J. (2005). Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkpoint functions of Rad9. Molecular and cellular biology *25*, 8430-8443.

Xu, M., and Zhu, B. (2010). Nucleosome assembly and epigenetic inheritance. Protein & cell *1*, 820-829.

Yan, L.Z., and Dawson, P.E. (2001). Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. Journal of the American Chemical Society *123*, 526-533.

Ye, T., Krebs, A.R., Choukrallah, M.A., Keime, C., Plewniak, F., Davidson, I., and Tora, L. (2011). seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res *39*, e35.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J Biol Chem *286*, 7983-7989.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952-1958.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science *328*, 916-919.

References

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nature structural & molecular biology *20*, 259-266.

Zhang, J., Gao, Q., Li, P., Liu, X., Jia, Y., Wu, W., Li, J., Dong, S., Koseki, H., and Wong, J. (2011). S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. Cell research *21*, 1723-1739.

Zhang, R., Poustovoitov, M.V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L., *et al.* (2005). Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Developmental cell *8*, 19-30.

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

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

Zhou, J., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. Nature structural & molecular biology *14*, 1070-1076.

Zofall, M., and Grewal, S.I. (2006). RNAi-mediated heterochromatin assembly in fission yeast. Cold Spring Harbor symposia on quantitative biology *71*, 487-496.

Author's contribution

Kungulovski G, Mauser R, Jeltsch A. (2015) "**Isolation of nucleosomes having multiple-modified co-existing histone protein octamers.**" EP15161621.6. Patent application submitted to EPO on 01.04.2015.

G.K and A.J conceived the idea. G.K designed the study and along with A.J guided the work of R.M. G.K performed some of the CIDOP experiments and analyzed all CIDOP-seq experiments. G.K and A.J developed new analytical approaches for assessment of chromatin states comprised of two histone modifications.

Kungulovski G, Mauser R, Jeltsch A. (in press) "**Specificity Analysis of Histone Modification Specific Antibodies or Reading Domains on Histone Peptide Arrays",** *Methods In Molecular Biology*, 1348, 275-284 doi: 10.1007/978-1-4939-2999-3_24.

G.K wrote and prepared the first draft of the manuscript including all figures and provided input on all successive drafts, including the final manuscript.

Kungulovski G, Mauser R, Jeltsch A. (2015) **"Affinity reagents for studying histone modifications and guidelines for their quality control"**, *Epigenomics (in press).*

G.K wrote and prepared the first draft of the manuscript including all figures and provided input on all successive drafts, including the final manuscript.

Kungulovski G, Nunna S, Thomas M, Zanger UM, Reinhardt R, Jeltsch A. (2015) "**Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained"**, *Epigenetics & Chromatin,* 8:12 doi: 10.1186/s13072-015-0002-z.

G.K wrote and prepared the first draft of the manuscript and provided input on all successive drafts, including the final manuscript. G.K performed some of the adenoviral work, the majority of ZF-GLP work and all ChIP experiments. G.K analyzed the ChIP and mRNA data obtained with the ZF-GLP constructs. G.K and A.J conceived the study.

Kungulovski G, Henry C, Kycia I, Reinhard R, Tamas R, Labhart P, Jurkowska RZ, Jeltsch A. (2014) "**Application of Histone Modification Specific Interaction Domains as an Alternative to Antibodies**", *Genome Research,* 24, 1842–1853 doi: 10.1101/gr.170985.113.

G.K wrote and prepared the first draft of the manuscript and provided input on all successive drafts, including the final manuscript. G.K performed or contributed to all experiments. G.K analyzed the vast majority of CIDOP-seq, ChIP-seq and RNA-seq data. G.K and A.J conceived and designed the study.

Bock I, Kudithipudi S, Tamas R, **Kungulovski G**, Dhayalan A, Jeltsch A. (2011) "**Application of Celluspots peptide arrays for the analysis of the binding specificity of epigenetic reading domains to modified histone tails**", *BMC Biochemistry*, 12:48 doi: 10.1186/1471-2091-12-48.

G.K performed the HP1 experiments and contributed to the writing of the manuscript.

Appendix (not included in the published thesis)

Appendix 1

Kungulovski G, Mauser R, Jeltsch A. (2015) **"Affinity reagents for studying histone modifications and guidelines for their quality control"**, *Epigenomics (in press).*

Appendix 2

Kungulovski G, Mauser R, Jeltsch A. (in press) "**Specificity Analysis of Histone Modification Specific Antibodies or Reading Domains on Histone Peptide Arrays",** *Methods In Molecular Biology*, 1348, 275-284 doi: 10.1007/978-1-4939-2999-3_24.

Appendix 3

Kungulovski G, Nunna S, Thomas M, Zanger UM, Reinhardt R, Jeltsch A. (2015) "**Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained"**, *Epigenetics & Chromatin,* 8:12 doi: 10.1186/s13072-015-0002-z.

Appendix 4

Kungulovski G, Henry C, Kycia I, Reinhard R, Tamas R, Labhart P, Jurkowska RZ, Jeltsch A. (2014) "**Application of Histone Modification Specific Interaction Domains as an Alternative to Antibodies**", *Genome Research*, 24, 1842–1853 doi: 10.1101/gr.170985.113.

Appendix 5

Kungulovski G, Mauser R, Jeltsch A. (2015) "**Isolation of nucleosomes having multiple-modified co-existing histone protein octamers.**" EP15161621.6. Patent application submitted to *EPO* on 01.04.2015.