Development of artificial single and double reading domains to analyze chromatin modification patterns

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
Rebekka Mauser
aus Ludwigsburg, Deutschland

Hauptberichter: Prof. Dr. Albert Jeltsch
Mitberichter: Prof. Dr. Roland Kontermann
Tag der mündlichen Prüfung: 1.08.2018

Institut für Biochemie und Technische Biochemie der Universität Stuttgart
2018

Universität Stuttgart

Eidesstattliche Erklärung

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

Stuttgart, den 3.04.2018

Rebekka Mauser

Inhaltsverzeichnis

Zusammenfassung

Der unstrukturierte N-terminale Rest der Histonproteine trägt eine Vielzahl von post-translationalen Modifikationen (PTMs), wie zum Beispiel Methyl-, Acetyl- oder Phosphatgruppen. Diese PTMs beeinflussen die Chromatinstruktur, sowie die Interaktion benachbarter Nukleosome. Des Weiteren dienen die Modifikationen als Bindestellen für Histon-interagierende Domänen. Die Untersuchung von Histon PTMs basiert hauptsächlich auf der Verwendung von Antikörpern, jedoch kamen einige Zweifel über die Spezifität dieser Antikörper und die daraus resultierende Reproduzierbarkeit der Daten auf. Deshalb war es ein Ziel dieser Arbeit, einen alternativen Ansatz zu den bisher verwendeten Antikörpern zu entwickeln. In vorherigen Arbeiten konnte gezeigt werden, dass Histon-Lesedomänen Antikörper auf eine sehr effektive Weise ersetzten können. Als Teil dieser Arbeit wurde die TAF3 PHD Domäne als neue spezifische H3K4me3 Histon-Lesdomäne etabliert.

Peptide Arrays und Far-Western Blot Experimente zeigten eine spezifische Interaktion mit H3K4me3 Modifikationen. Auch ChIP ähnliche Experimente (CIDOP: Chromatin Interacting Domain Precipitation) welche mit quantitativer PCR und Next Generation Sequenzierung ausgewertet wurden, zeigten vergleichbare Resultate wie ein validierter H3K4me3 Antikörper.

Mit der ersten Erwähnung der Histoncode-Hypothese, kam die Frage auf, ob verschiedene Kombinationen von Histonmodifikationen für spezifische biologische Funktionen stehen. Allerdings ist es immer noch eine sehr anspruchsvolle und schwierige Aufgabe, die Koexistenz von Histonmodifikationen am selben Histon genomweit nachzuweisen. Daher war das Hauptziel dieser Arbeit, doppelte Histon-Lesedomänen zu entwickeln, welche durch einen flexiblen Linker verbunden sind, um eine simultane Interaktion mit doppelt modifizierten Histonenden zu erreichen und diese in einem einzelnen CIDOP Experiment zu präzipitieren.

Um dieses neue Konzept zu validieren, wurden die Dnmt3a PWWP Domäne und die MPP8 Chromo Domäne miteinander fusioniert und ihre spezifische Erkennung von H3K36me2/3 und H3K9me3 Histonmodifikationen analysiert. Biochemische Untersuchungen, darunter Peptide Arrays, Far-Western Blot und Western Blot Experimente zeigten, dass beide Domänen weiterhin spezifisch mit ihren Zielmodifikationen interagieren und ebenfalls bevorzugt mit doppelt modifiziertem Chromatin interagieren. Zusätzlich konnte die bevorzugte Bindung an doppelt modifiziertes Chromatin durch Bindestellen-Mutanten und Methyl-Lysine Analogen bestätigt werden. Die neu generierte Doppel-Lesedomäne wurde in CIDOP Experimenten

verwendet, um Genomregionen zu identifizieren, welche mit beiden Histonmodifikationen assoziiert sind. Die Untersuchung der genomweiten Verteilung von H3K36me2/3 und H3K9me3 zeigte die Existenz dieser Doppelmodifikation in humanen Zellen, welche einen neuen Chromatinstatus repräsentiert. Dieser ist mit schwach transkribierten Genen assoziiert und an Bindestellen von ZNF274 und SetDB1 angereichert.

Des Weitern wurde in dieser Arbeit eine neue Analysemethode entwickelt, um die neu generierten Doppel-Lesedomänen effizient charakterisieren zu können. Dafür wurde die natürlich vorkommende Doppel-Lesedomäne des BPTF Proteins verwendet, um das Potenzial dieser neuen Methode zu demonstrieren. BPTF besitzt eine PHD Domäne, welche an H3K4me3 Modifikationen bindet und eine Bromo Domäne, welche mit Acetylgruppen am N-terminalen Ende des H4 Histon interagiert. Eine synergistische Bindung an modifizierte Peptide konnte erfolgreich mit den neuen gemischten Peptide Arrays gezeigt werden. Zudem, wurde diese neue Methode eingesetzt, um weitere Doppel-Lesedomänen zu optimieren.

Im Zuge dieser Arbeit, wurden weitere Doppel-Lesedomänen gegen verschiedene Histonmodifikationskombinationen generiert. Dazu zählen unter anderem, PWWP-ATRX, MPP8 Chromo Domäne-L-Double Tudor und CBX7 Chromo Domäne-L-MPP8 Chromo Domäne, welche auf ihre Doppelspezifität hin untersucht wurden.

Ebenfalls, wurden Doppel-Lesedomänen für die gleichzeitige Interaktion mit DNA Methylierung und Histonmodifikationen erstellt. Zuerst wurde die methyl-DNA Bindedomäne des MBD2 Proteins verwendet, welche allerdings eine sehr starke Interaktion aufwies und dadurch die Bindung an Chromatin dominierte. Deshalb wurde die schwächere, jedoch ebenfalls spezifische methyl-DNA Bindedomäne des MBD1 Proteins verwendet. In ersten Experimenten konnte gezeigt werden, dass diese neuen Fusionskonstrukte eine simultane Interaktion mit Chromatin, welches mit DNA Methylierung und den jeweiligen Histonmodifikationen assoziiert ist, aufweisen.

Zusammenfassend zeigten die Studien mit Doppel-Lesedomänen, dass mit dieser neuen Methode doppelt modifiziertes Chromatin sehr spezifisch präzipitiert werden kann und die genomweite Untersuchung des neu identifizierten bivalenten Chromatinstatus möglich ist. Dies bietet nun eine Grundlage für die weitere Untersuchung vieler verschiedener Histonmodifikationskombinationen. Außerdem kann der kombinatorische Einfluss von Histonmodifikationen auf die Chromatinstruktur analysiert werden und ebenfalls ermöglicht

diese neue Methode, ein besseres Verständnis der biologischen Funktion von verschiedenen Histonmodifikationskombinationen zu gewinnen.

Abstract

The unstructured N-terminal tails of histone proteins carry many different post-translational modifications (PTMs), like methylation, acetylation or phosphorylation. These PTMs can alter the chromatin structure, influence the interaction of adjacent nucleosomes and serve as specific binding sites for histone interacting domains. Currently, the investigation of histone tail PTMs is mainly based on antibodies, however concerns about the specificity of these antibodies and reproducibility of data arouse. Therefore, it was one aim of this thesis to develop alternative approaches to histone tail PTM antibodies. Previous studies already showed that histone modification interacting domains (HiMIDs) can replace histone tail antibodies in a highly effective manner. As part of this work, the TAF3 PHD domain was established as new H3K4me3 specific HiMID. In peptide array binding and Far-western blot assays, the domain showed a specific interaction with H3K4me3 modifications. Also in ChIP like experiments (CIDOP: Chromatin Interacting Domain Precipitation) coupled to qPCR and next generation sequencing, the domain showed a similar performance as validated H3K4me3 antibodies.

With the proposal of the histone code hypothesis the question was raised if combinations of histone modifications carry specific biological functions. However, so far, the experimental analysis of the co-occurrence of histone modification on the same nucleosome in a genome-wide manner is a challenging task. For this reason, the main aim of this work was to develop double reading domains in which two histone reading domains are fused together with a flexible linker to achieve simultaneously readout of dual histone tail modifications in a single CIDOP experiment. To validate the concept, the Dnmt3a PWWP domain and the MPP8 Chromo domain were fused together and their specific recognitions of H3K36me2/3 and H3K9me3 histone tail modifications were analyzed. Biochemical investigations like peptide arrays, Far-western blot and western blot experiments showed that both domains specifically interact with their targets and preferentially interact with double modified chromatin.

Additionally, the preferred interaction with double modified chromatin could be further verified with binding pocket mutants and methyl-lysine analogues. The newly generated double domain was used in chromatin precipitation experiments to identify genome regions where both modifications are present. The genome-wide distribution of the H3K36me2/3-H3K9me3 showed that this combination of histone marks represents a novel bivalent

chromatin state, which is associated with weakly transcribed genes and is enriched for binding sites of ZNF274 and SetDB1.

Also in this work, mixed peptide arrays were introduced as new screening method for the efficient analysis of double reading domains. The naturally occurring double reading domain of the BPTF protein was used to demonstrate the capability of this new screening tool. BPTF contains a PHD domain, which binds to H3K4me3 and a Bromo domain, which interacts with acetyl groups of the H4 tail. Synergistic binding to both peptides was shown using the newly developed mixed peptide arrays. Additionally, in the course of this work mixed peptide arrays were used to optimize several of the designed double reading domains.

Furthermore, some other double reading domains were generated in this work, like PWWP-ATRX, MPP8 Chromo domain-L-double Tudor and CBX7 Chromo domain-L-MPP8 Chromo domain and analyzed for specific dual readout. Also double reading domains with dual specificity for DNA methylation and histone marks were generated. The firstly used methyl-DNA binding domain of the MBD2 protein showed a strong binding, dominating the effect of the HiMIDs. Therefore, the weaker but still specific methyl-DNA binding domain of the MBD1 protein was used. First experiments with this new fusion constructs showed a simultaneously interaction with chromatin which is associated with DNA methylation and histone PTMs.

In summary, the studies with double reading domains showed that with this novel method precipitation of double modified chromatin is possible and that the genome-wide investigation of newly studied bivalent chromatin states is feasible. Therefore, this novel approach makes it possible to analyze many different combinations of histone modifications, investigate their influence on chromatin and gain a deeper understanding of the biological role behind histone tail modification patterns.

List of abbreviations

AA	Amino Acid
AB	Antibody
Ac	Acetyl
ADD	ATRX-Dnmt3a-Dnmt3l
Amp	Ampicillin
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
bp	Base pair
BPTF	Bromo domain PHD finger transcription factor
CBX	Chromobox
CHO	Chinese hamster ovary
CIDOP	Chromatin Interacting Domain Precipitation
ChIP	Chromatin Immunoprecipitation
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
CTD	C-terminal domain
CXXC	CXXC finger protein, zing finger domain-containing protein
ddH_2O	Double distilled water
DNA	Deoxyribonucleic acid
Dnmt1	DNA methyltransferase 1
Dnmt3a	DNA methyltransferase 3a
Dnmt3b	DNA methyltransferase 3b

Dnmt3L	DNA methyltransferase 3-like
dNTP	2'-Desoxynucleotid-5'-triphosphate
DTT	Dithiothreitol
DMEM	Dulbecco's modified Eagle's medium
E.coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic
ENCODE	Encyclopedia of DNA elements
EPL	Expressed protein ligation
EZH2	Enhancer of Zeste 2
FASTQ	Text-based format for storing both a biological sequence (usually nucleotide sequence) and its corresponding quality scores
FRET	Fluorescent Resonance Energy Transfer
FPKM	Fragments per kilobase per Million
G9A	Euchromatic histone lysine N-methyltransferase 2
GLP	G9A-like protein
GST	Glutathione S-transferase
HAT	Histone acetyl-transferase
HDAC	Histone deacetylase
HEK293	Human Embryonic Kidney fibroblasts 293
HepG2	Hepatocellular carcinoma G2 cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HiMID	Histone Modification Interacting Domain
Hox	Homeobox
HP1	Heterochromatin Protein 1
IGV	Integrative Genomics Viewer
IPTG	Isopropyl β-D-thiogalactoside
JMJD2	Jumonji domain
KCl	Potassium chloride

kDa	Kilodalton
LB	Luria-Bertani
m	Meter
MBD	Methyl-CpG-binding domain
MBP	Maltose binding protein
MACS	Model-based Analysis of ChIP-seq
Me	Methyl
Me1	Monomethyl
Me2	Dimethyl
Me3	Trimethyl
MeCP2	Methyl CpG binding protein 2
min	Minute
mL	Milliliter
mM	Millimolar
MNase	Micrococcal nuclease
MLA	Methyl Lysine Analogues
MLL	Mixed lineage leukaemia
MPP8	M-Phase Phosphoprotein 8
MS	Mass Spectrometry
ng	Nanogram
NGS	Next generation sequencing
NH	Native histones
nm	Nanometer
NMR	Nuclear magnetic resonance
NPL	Native Chemical Ligation
OD_{600}	Optical Density at 600 nm
PCR	Polymerase Chain Reaction
PHD	Plant Homeo Domain

PIC	Protease inhibitor cocktail
PKMT	Protein lysine methyltransferase
PMSF	Phenylmethylsulfonyl fluoride
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
rcf	Relative centrifugal force
RH	Recombinant histones
RNA	Ribonucleic acid
RNAP II	RNA polymerase II
RPKM	Reads Per Kilobase per Million
rpm	Rounds per minute
RT	Room temperature (approx. 22 °C)
SAM	S-Adenosyl Methionine
SEM	Standard error of the mean
SETD1A/B	Histone-lysine N-methyltransferase SETDB1A/B
Seq	Massively parallel sequencing
SDS-PAGE	Sodium Dodecyl Sulphate Poly-acrylamide Gel Electrophoresis
SICER	Spatial clustering approach for the identification of ChIP-enriched regions
SUV420	Histone-lysine N-methyltransferase SUV4-20. Specifically methylate lysine 20 of histone H4
SUV39H1/2	Histone-lysine N-methyltransferase SUV39H1/2. Suppressor of variegation 3-9 homolog 1/2 (Drosophila)
TAF3	TATA box-binding protein-associated factor 3
TSA	Trichostatin A
TSS	Transcription Start Site
TTBS buffer	Tween 20 containing Tris buffered saline

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
Å	Angstrom
°C	Degree Celsius
µg	Microgram
µL	Microliter
µM	Micromolar

1 Introduction

1.1 Epigenetics and chromatin structure

Almost all cells in a multicellular organism share the same genetic information, but during development, diverse cell types with distinct functions arise, which are generated by epigenetic mechanisms. The term “Epigenetics” was the first time introduced by Conrad Waddington in 1942 “as the study of the interactions between genes and their products which results in a particular phenotype” (Waddington C.H. 1942).

The interest in the field epigenetic grew rapidly in the ensuing years, which resulted in gradual changes of the definition of epigenetics. Today, the term epigenetics also includes all processes that alters gene functions resulting from chromosomal changes and do not entail changes in DNA sequence (Bonasio et al. 2010; Goldberg et al. 2007; Dupont et al. 2009). The changes in gene expression derive from epigenetic modifications namely DNA modifications, histone tail post-translational modifications, histone variants and ncRNAs, which will also be transmitted to daughter cells and inherited to the next generation (Blake and Watson 2016).

The genetic information of eukaryotic cells is encoded in the DNA sequence and in order to fit with its enormous length of around 2 m into the cell nucleus, with a diameter of only 5 to 10 µm, the DNA needs to be extremely condensed. To achieve this high degree of compaction, histone and non-histone proteins are required, forming together a structure called chromatin. This complex macromolecule fulfils, besides keeping the DNA in a condensed form, further important functions, like preventing DNA damage, allowing mitosis, as well as controlling gene expression and DNA replication (Shanle et al. 2014; Kouzarides 2007).

The smallest unit of the chromatin fibre is the nucleosome which consists of 147 bp of DNA wrapped around one H3-H4 heterotetramer and two H2A-H2B heterodimers that together form the histone octamer (Luger et al. 1997) (Fig. 1A). Individual nucleosomes are connected via linker DNA resulting in a so-called “beads-on-a-string” structure (Fig. 1B). To condense this structure even more linker histones (histone H1) can be added which transform the beads-on-a-string pattern into a higher compact 30 nm fibre (Robinson and Rhodes 2006). Recent studies showed, that condensed chromatin is a highly dynamic structure and that the electrostatic interaction between adjacent nucleosomes can stabilize or weaken the

chromatin fibre formation (Grigoryev and Woodcock 2012; Grigoryev et al. 2009; Kalashnikova et al. 2013; Pepenella et al. 2014).

Due to their high degree of packing some types of chromatin cannot be accessed by the transcriptional apparatus as easily, leading to gene repression. This tightly packed chromatin is referred to heterochromatin, which can be further differentiated into constitutive heterochromatin, found at centromeres, pericentromeric regions, telomeres and repetitive sequences, and facultative heterochromatin, which is associated with developmental genes (Trojer and Reinberg 2007; Saksouk et al. 2015). In contrast to the highly packed heterochromatin, euchromatin is lightly packed and contains actively transcribed genes (Filion et al. 2010).

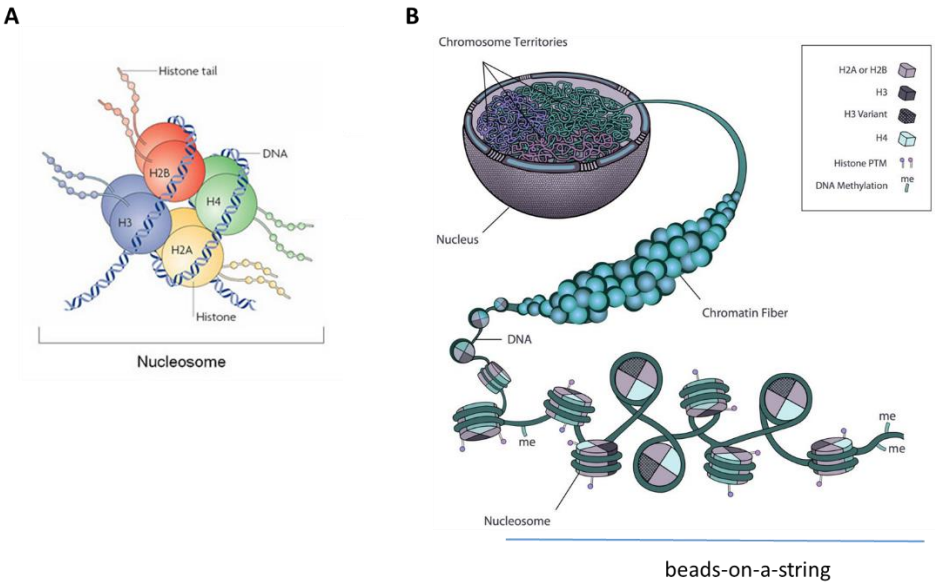


Figure 1. Schematic illustration of the nucleosome and higher order chromatin structure. **A)** Structure of a nucleosome. The DNA is wrapped around the four core histones H2A, H2B, H3 and H4. The core particle is predominantly globular but contains flexible N-terminal tails, which project out of the globular core particle. Figure taken from (Tsankova et al. 2007). **B)** Scheme of the chromatin organisation in the cell. The DNA wrapped around the histone octamer forming the so-called beads-on-a-string structure. Histone tail modification and DNA methylation are indicated. The DNA is compacted into a higher order chromatin fibre structure, which is packed into the nucleus. Figure adapted from (Rosa and Shaw 2013).

1.2 Epigenetic mechanisms

There are three principal players involved in epigenetic processes: non coding RNA (ncRNA), DNA methylation and histone post-translational modifications (PTMs). These epigenetic signals together affect important regulatory functions in all chromatin-templated processes.

1.2.1 DNA methylation

In mammals the DNA is methylated at the C5 position of cytosine residues predominantly in the context of CpG dinucleotides (Shanle et al. 2014). This is associated with gene silencing, X-chromosome inactivation, genomic imprinting and suppresses the activation of repetitive sequences and transposable elements.

In mammals roughly 60 - 80% of all CpG dinucleotides in the genome are methylated in a tissue and cell type specific manner (Suzuki and Bird 2008; Klose and Bird 2006; Jurkowska et al. 2011; Jones 2012; Jeltsch and Jurkowska 2014). DNA methylation of promoter regions leads to gene silencing, whereas gene body methylation is associated with highly expressed genes. For normal development, a correct DNA methylation pattern is crucial and leads to differentiation of specific cellular phenotypes (Jones 2012; Deaton and Bird 2011). Aberrant DNA methylation results in multiple diseases like cancer (Baylin and Jones 2011; Bergman and Cedar 2013; Suva et al. 2013; Hamidi et al. 2015), autoimmune diseases (Javierre et al. 2010) or imprinting disorders (Butler 2009). In addition to methylation at CpG sites, non-CpG methylation occurs, particularly in early embryonic cells and neurons (Jang et al. 2017) but its function is still elusive.

It was long time thought that methylation exclusively occurs on the C5 position of cytosine but there are some hints that methylation can also occur on the N6 position of adenine. This DNA modification is common in prokaryotes and was found in eukaryotes to frequently appear on mRNAs (Heyn and Esteller 2015; Jia et al. 2013). Recently it was shown that methylation of adenine at N6 can also occur in eukaryotes leading to gene activation (Zhang et al. 2015a; Greer et al. 2015; Fu et al. 2015; Sun et al. 2015) but this finding is still controversial (Schiffers et al. 2017).

The DNA methyltransferases (Dnmts) Dnmt1, Dnmt3a and Dnmt3b are responsible for the establishment and maintenance of DNA methylation. These enzymes are highly similar in their C-terminal part, which contains the catalytic domain, but they differ in their N-terminal part

(Fig. 2) (Maresca et al. 2015). The different domains located in the N-terminal part show a variety of functions, by for example guiding their localisation, mediating interaction with proteins, RNAs or post-translational modifications of histone tails or regulating the Dnmt activity.

Dnmt1 is the largest protein of this family. It comprises of a DNA methyltransferase associated protein 1 interacting domain (DMAPD), which interacts with the transcriptional repressor DMAP1, a PCNA binding domain (PBD), a replication foci targeting domain (RFTD), a CXXC domain, which binds unmethylated DNA and two Bromo-adjacent homology domains 1 and 2 (BAH1 and BAH2) that are necessary for folding and chromatin targeting. All these domains and the nuclear localisation signal (NLS) are located in N-terminus of the protein (Song et al. 2012) and the C-terminus represents the conserved catalytic domain.

Dnmt3a and Dnmt3b contain both a PWWP and an ADD domain located in their N-terminal part, which mediate the binding to the H3 tail. The PWWP domain is known to specifically recognize the H3K36me3 modification and to a lower extent also H3K36me2 and the ADD domain (ATRX-Dnmt3-Dnmt3L) interacts with the H3 tail if H3K4 is unmodified (Ooi et al. 2007; Zhang et al. 2010; Dhayalan et al. 2010). It has been shown that H3 tail binding to the ADD domain stimulate the enzymatic activity of Dnmt3a *in vitro* (Li et al. 2011; Guo et al. 2014; Zhang et al. 2010). Structural investigations and kinetic studies showed that the ADD domain can bind to the catalytic domain in an autoinhibitory conformation, where it inhibits the catalytic domain through blocking DNA binding.
However, interaction with the H3 histone tail interrupts the autoinhibitory interaction of the ADD domain with the catalytic domain leading to a large movement of the ADD domain and docking of the ADD domain to another binding site, resulting in an active conformational state of the enzyme (Guo et al. 2014).

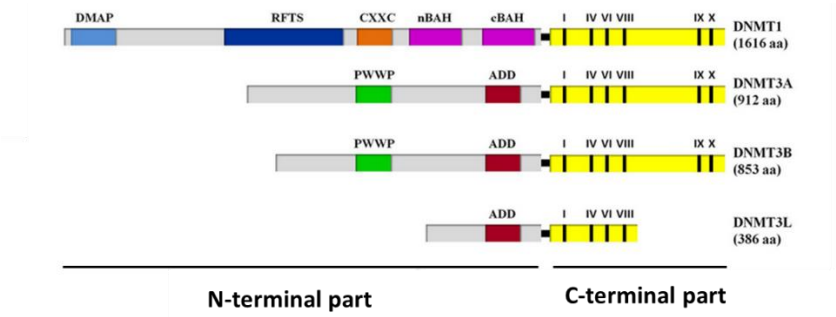


Figure 2. Representation of human Dnmt1, Dnmt3a, Dnmt3b and Dnmt3L. The domain structure of the catalytic C-terminal part and the N-terminal part of the different variants is represented. Figure adapted from (Maresca et al. 2015).

DNA methylation is established during embryonic development by the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b with their catalytic inactive regulatory factor Dnmt3L and maintained in dividing cells mainly catalysed by Dnmt1 (Goll and Bestor 2005). However, evidence is accumulating showing that DNA methylation is a dynamic process where the *de novo* and the maintenance DNA methylation is created by all family members of the Dnmts (Jurkowska and Jeltsch 2016; Riggs and Xiong 2004; Jones and Liang 2009).

1.2.2 Histone tail post-translational modifications (PTMs)

The unstructured N-terminal tails of histones are decorated with a variety of post-translational modifications. These modifications can alter the chromatin structure directly or indirectly as they can serve as binding sites for modification specific histone tail interacting domains, also called reading domains (Fig. 3).

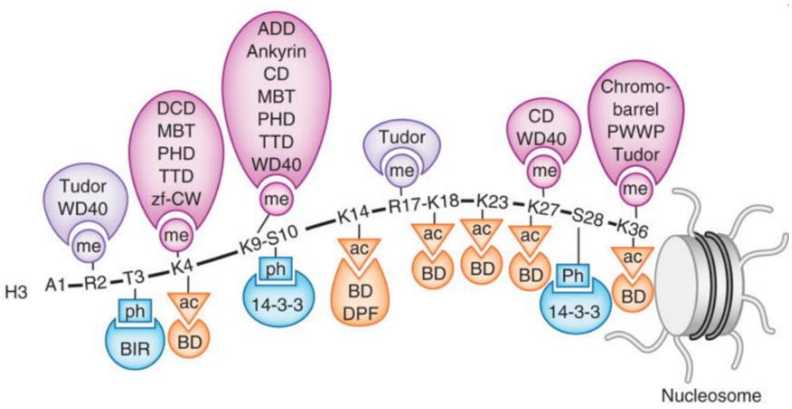


Figure 3. Schematic representation of the recognition of modified histone H3 tail. Different reading domains interact specifically with PTMs like acetylation (ac), methylation (me) and phosphorylation (ph). Figure taken from (Musselman et al. 2012).

Many different histone tail modifications have been described to date, for example lysine residues can be methylated, ubiquitinated, SUMOylated, crotonylated, butyrylated, acetylated or propionylated. Furthermore, arginine residues can carry methylation, citrulllination or ADP-ribosylation marks. Phosphorylation and glycosylation can occur on serine and threonine residues (Strahl and Allis 2000; Bannister and Kouzarides 2011; Kebede et al. 2015; Arnaudo and Garcia 2013; Chen et al. 2007; Tan et al. 2011; Rivera and Ren 2013).

One of the best-studied histone tail PTM is the methylation of lysine residues, a modification that can directly affect gene expression (Greer and Shi 2012). For example, H3K4me3 and H3K79me2 signals are enriched at the transcription start site (TTS) indicating correlation with actively transcribed genes (Fig. 4A and C), whereas H3K36me3 follows a gradient of enrichment peaking at the 3` ends of expressed genes (Fig. 4B). H3K9me3, H4K20me1 and H3K27me3 do not show signals correlated with active genes but they are associated with transcriptional repression (Fig. 4D-F) (Barski et al. 2007).

Methylation and acetylation of histone tails were extensively studied over the past years, whereas not so many details about the function of the other modifications are known. Nevertheless, recent studies tried to couple histone modifications with cellular metabolism and to investigate their effect on the chromatin architecture. For example, propionylation was found to be preferentially enriched at promoters of active genes and be recognized by specific acetylation reading domains (Kebede et al. 2017).

Additionally, the globular histone core was shown to be subjected to numerous modifications such as H2BK40, H3K56, H3K79 and H4R92 methylation. But also acetylation, ubiquitinaylation and phosphorylation were found on the globular domain of histones (Lawrence et al. 2016; Cosgrove et al. 2004; Tjeertes et al. 2009; Feng et al. 2002). The histone core has direct contacts with the DNA and therefore histone core modifications play an essential role in DNA replication and repair (Lawrence et al. 2016). Furthermore, the histone core modifications were shown to crosstalk with histone tail modifications such as the core modification H3K79me3 which is found at genomic regions enriched for H3K4me3 histone tail modifications (Nakanishi et al. 2009).

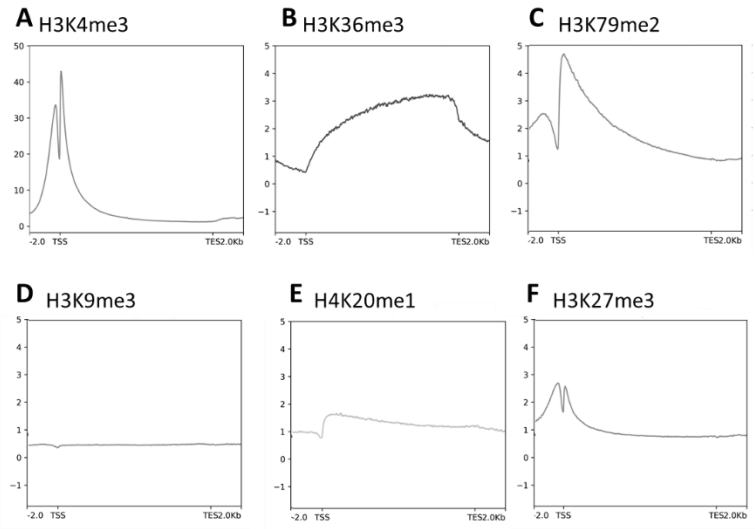
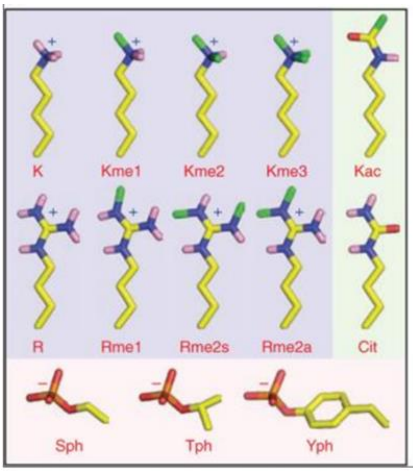


Figure 4. Density profiles of **A)** H3K4me3, **B)** H3K36me3, **C)** H3K79me2, **D)** H3K9me3, **E)** H4K20me1 and **F)** H3K27me3 and in human HepG2 cells. Density profiles are generated ranging from TSS to TES with a window of 10 kb. Flanking sites +/- 2 kb. Density profiles were generated using the DeepTools program plotProfile (Galaxy).

1.2.2.1 Role and function of lysine methylation

Genome-wide studies identified certain regions where particular methylation states of lysine residues are enriched illustrating special biological functions behind every modification state (Martin and Zhang 2005). The ε-amino group of lysine residues can be either mono-, di- or trimetylated (Rice and Allis 2001). With the addition of a methyl-group the cation radius expands and the ability to donate hydrogen bonds decreases (Taverna et al. 2007). Therefore, every state of lysine methylation requires a specific readout by different readers. The two main families of methylated-lysine reading domains identified so far are the Royal superfamily and PHD fingers (Fig 5).



	Reader module	PTM mark
	Bromodomain	Many histone Kac, (Kac)
Royal	Chromodomain	H3K9me2/3, H3K27me2/3
Royal	Double chromodomain	H3K4me1/2/3
Royal	Chromo barrel	H3K36me2/3
Royal	Tudor	(Rme2s)
Royal	Double/tandem tudor	H3K4me3, H4K20me3 H4K20me1/2, (Kme2)
Royal	MBT	H4K20me1/2, H1K26me1/2 H3K4me1, H3K9me1/2
	PHD finger	H3K4me3, H3K4me0 H3K9me3, H3K36me3
	WD40 repeat	H3R2/K4me2, (R, Sph, Tph)
	14-3-3	H3S10ph, H3S28ph, (Sph, Tph)
	BRCT	H2AX-S139ph, (Sph, Tph)

Figure 5. Illustration of the different histone tail PTMs and their corresponding reading domains. **A)** Different post-translationally modified amino acid residues (Yellow, carbon; blue, nitrogen; pink, polar hydrogen; red, oxygen; orange, phosphorus; green, methyl groups). Background shading indicate the different charge of side chains at physiological pH (light blue, positive; pink, negative; light green, uncharged). **B)** Table shows a summary of some important reading domains and the post-translational histone mark associated with the reading domains. Figure taken from (Taverna et al. 2007).

The small zinc fingers are able to read histone tails in a sequence dependent manner influenced by the methylation state of histone tail residues (Li et al. 2006; Wysocka et al. 2006; Shi et al. 2006; Taverna et al. 2006; Nakamura et al. 2007; Matthews et al. 2007; Lange et al. 2008; Zeng et al. 2010). The PHD domain consists of two anti-parallel ß-sheets with a Cys_4-His-Cys_3 motif stabilized by two zinc atoms. The PHD fingers specifically interact with the H3 tail and recognize the target modification via a surface groove where four aromatic residues form a pocket stabilized by cation-π and van der Waals interaction (Li et al. 2006). This binding mode was reported the first time for the PHD domain of the BPTF protein and its aromatic cage is formed by one tryptophan and three tyrosine residues interacting with H3K4me3. The structure of other aromatic cages of PHD domains can vary between two and four aromatic and hydrophobic residues but the tryptophan seems to be invariable.

Most of the known methyl-lysine readers belong to the Royal superfamily. The Royal superfamily includes the Chromo domain, Chromo barrel domain, Tudor domain, PWWP (highly conserved Pro-Trp-Trp-Pro motif) domain and the MBT (malignant brain tumor) domain (Fig. 5) and the binding of reading domains occurs in different binding modes dependent on the degree of the lysine methylation.

For lower methylation states, binding takes place in a cavity-insertion recognition mode. Here the methyl-ammonium group of the methyl-lysine is inserted into a deep protein cleft. Binding

to higher methylation states take place in a surface-groove recognition mode where the binding pocket is wider and more accessible (Taverna et al. 2007). Therefore, the common binding structure of the Royal superfamily members is an aromatic pocket consisting of two to four aromatic residues, which stabilize the interaction with the methyl-lysine via cation-π interaction.

The Royal superfamily domain PWWP (proline-tryptophan-tryptophan-proline motif) occur among others in the DNA methyltransferase Dnmt3a and was shown to specifically interact with H3K36me2 and H3K36me3 modifications (Fig. 6) (Dhayalan et al. 2010).

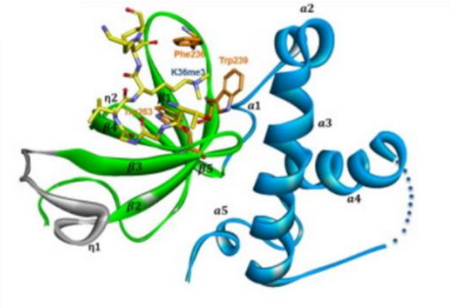


Figure 6. Crystal structure of Dnmt3b PWWP domain in complex with an H3K36me3 peptide. Figure taken from (Rondelet et al. 2016).

1.2.2.2 Histone 3 lysine 4 methylation

Various enzymes set the H3K4 modification but the most important enzyme group is the MLL family comprising SETD1A, SETD1B and MLL 1-4 (Shilatifard 2012).

Depending on the methylation state H3K4 fulfils different functions and it is associated with different chromatin states. For example, the H3K4me1 modification is a mark of poised or active enhancers (Calo and Wysocka 2013; Ernst et al. 2011). In contrast H3K4me2 and H3K4me3 are associated with actively transcribed chromatin regions, like promoters and TSS (Fig. 4A) (Liang et al. 2004; Ernst et al. 2011; Mikkelsen et al. 2007; Schneider et al. 2004).

Different plant homeodomains (PHD) were found to interact specifically with H3K4me3 modifications. H3K4me3 binding mode was reported the first time for the PHD domain of the BPTF protein. Later the PHD domain of the TATA box binding protein associated factor 3 (TAF3), a subunit of the general transcription factor TFIID, were also shown to interact with

H3K4me3 and stimulate the RNA polymerase II-mediated transcription (Fig. 7) (Vermeulen et al. 2007; Vermeulen and Timmers 2010; Lauberth et al. 2013).

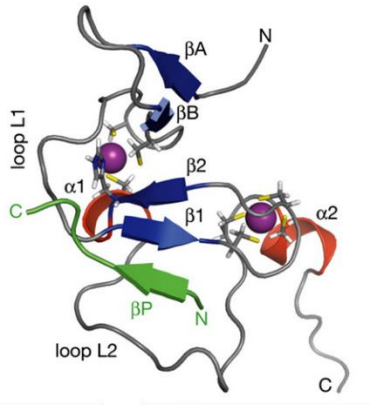


Figure 7. TAF3 PHD domain interacting with an H3K4me3 (green) modified peptide (zinc ion in purple). Figure taken from (van Ingen et al. 2008).

1.2.2.3 Histone 3 lysine 36 methylation

The three members of the NSD family (nuclear receptor binding SET domain) NSD1, NSD2 and NSD3 were shown to mainly catalyse H3K36me1 and H3K36me2 methylation (Morishita and di Luccio 2011), whereas the H3K36me3 modification is uniquely set by the enzyme SETD2. SETD2 binds to the phosphorylated serine residues of the CTD domain of the elongating RNA polymerase II (Martin and Zhang 2005; Wagner and Carpenter 2012). Therefore, the H3K36me3 modification accumulates in the gene body of transcribed genes (Fig. 4B), where it contributes to repression of internal initiation events by recruiting chromatin remodelers like ISWI (Imitation Switch) and CGD (Chromo domain helicase DNA-binding) or the H3K4 lysine specific demethylase (LSD1) (Smolle et al. 2012).

Furthermore, H3K36me3 can be recognized by the Eaf3 domain of the Rpd3S histone deacetylase complex, which leads to the removal of acetylation marks and results in gene repression (Keogh et al. 2005; Carrozza et al. 2005; Joshi and Struhl 2005; Wagner and Carpenter 2012). Also, the PWWP domains of Dnmt3a and Dnmt3b interact with H3K36me2/3, which leads to DNA methylation in gene bodies and gene silencing (Dhayalan et al. 2010; Baubec et al. 2015). Additionally, H3K36me3 was found to be involved in the regulation of alternative splicing (Kolasinska-Zwierz et al. 2009; Saint-André et al. 2011) and it was reported to correlate with the heterochromatin protein 1 (HP1) and H3K9me3 modifications in Drosophila (Riddle et al. 2011).

1.2.2.4 Histone 3 lysine 9 methylation

The H3K9 modification is mainly associated with gene silencing, heterochromatin formation and DNA methylation (Saksouk et al. 2015). Depending on the modification status, lysine 9 of the H3 tail contributes differently to the complex chromatin network. H3K9me2 and H3K9me3 are mostly associated with silenced genes, whereas H3K9me1 is enriched at the transcription start side of active genes (Barski et al. 2007).

There are numerous enzymes know to catalyse the modification on histone 3 lysine 9. For example, PRDM3 and PRDM16 are responsible for the monomethylation of H3K9, G9a and G9a-like protein are mostly responsible for the dimethylation of H3K9, whereas SUV39H1 and SUV39H2 set the trimethylation mark (Rea et al. 2000; Peters et al. 2003; Rice et al. 2003; Mozzetta et al. 2015).

The H3K9me3 modification is a hallmark of constitutive heterochromatin and H3K9me3 was shown to interact with the heterochromatin binding protein1 (HP1) (Martin and Zhang 2005; Krishnan et al. 2011; Becker et al. 2016; Mozzetta et al. 2015). HP1 binds to H3K9me3, recruits the histone lysine methyltransferase SUV39H1, subsequently adjacent nucleosomes will be methylated which leads to further recruitment of SUV39H1 resulting in heterochromatin formation (Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001; Squazzo et al. 2006). Furthermore, SUV420H2 interact with HP1 resulting in an establishment of H3K9me3 and H4K20me3 at pericentromeric heterochromatin leading to gene silencing (Schotta et al. 2004).

The Chromo domain of the MPP8 protein belongs to the Royal superfamily of reading domains recognizing their target modification (H3K9me3) via an aromatic pocket (Fig. 8A). The pocket is formed by three aromatic residues phenylalanine (F59), tryptophan (W80), tyrosine (Y83) and an aspartate (D87). The MPP8 Chromo domain recognizes specifically the residues 5 to 10 of the histone H3 tail (QTARKS). MPP8 was shown to be phosphorylated at several positions including tyrosine (Y83) in the aromatic cage. The negatively charge of the phosphoryl group is thought to enhance the binding of the positively charged methyl group (Chang et al. 2011).

Another domain interacting with H3K9me3 modification is the ATRX (alpha thalassemia/mental retardation X linked) ADD domain. This domain uses a zinc finger to interact with H3K9me3 modifications, but is inhibited by the presence of H3K4me3

modifications (Dhayalan et al. 2011). The ADD domain interacts with the H3K9me3 modification with its N-terminal GATA-like C2C2 zinc finger and its C-terminal C4C4 incomplete zinc finger. Both zinc fingers form a pocket where the H3K9me3 group is inserted (Fig. 8B) (Iwase et al. 2011).

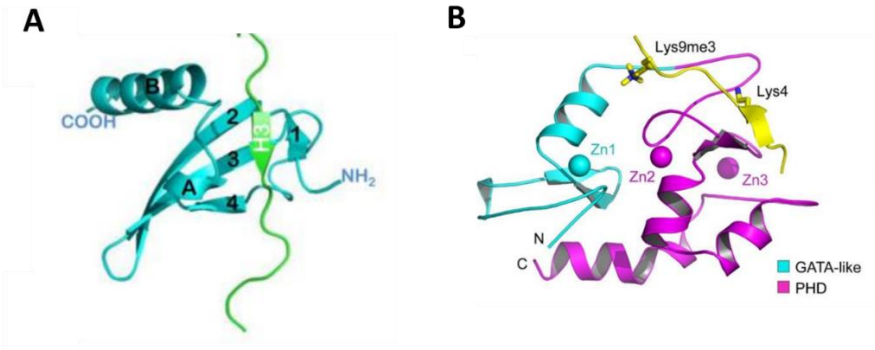


Figure 8. A) MPP8 Chromo domain (cyan) bound to a histone H3 peptide (green). Figure taken from (Chang et al. 2011). **B)** Complex of ATRX ADD domain with H3K9me3 peptide. (ATRX ADD cyan, PHD domain magenta, H3 peptide yellow, zinc ions cyan and magenta). Figure taken from (Iwase et al. 2011).

1.2.2.5 Histone 4 lysine 20 methylation

Additionally to H3K9me3, the H4K20me3 modification is a mark of heterochromatin (Schotta et al. 2004; Gonzalo et al. 2005; Regha et al. 2007) and found to be associated with repressed genes. H4K20me2 and H4K20me3 are set by SUV420H1 and SUV420H2 and are enriched at pericentric heterochromatin, telomeres, imprinted regions, repetitive elements and involved in gene silencing. The methyltransferase SUV420H2 were shown to interact with HP1 isoforms and thereby is recruited to H3K9me3 regions (Schotta et al. 2004).

H4K20me2 is broadly distributed over the genome and shows no specific influence on transcriptional regulation, but it is involved in DNA repair and replication (Jorgensen et al. 2013; Schotta et al. 2004). The H4K20me1 modification is uniquely added by the SET8 mono-methyltransferase and as H3K4me2 involved in DNA repair and DNA replication. It has been shown that the regulation of the H4K20me1 is cell cycle dependent. H4K20me1 accumulates in the S and G2 phase and peaks at the M phase whereas it declines during G1 phase (Pesavento et al. 2008; Oda et al. 2009).

Another reader of the Royal superfamily is the MBT domain (malignant brain tumor) of the Drosophila Polycomb group repressor dSfmbt which interacts with H4K20me1 (Grimm et al.

2009). Interaction with H4K20me2 was observed for the Tudor domains of JMJD2A and 53BP1 and are essential for the regulation of cellular DNA damage response (Tang et al. 2013; Mallette et al. 2012). Additionally, the Tandem Tudor domain of 53BP1 is thought to interact with the H4K20me3 modification (Kachirskaia et al. 2008; Roy et al. 2010).

1.2.2.6 Histone 3 lysine 27 methylation

Like for the other histone tail lysine methylation, the three different methylation states of H3K27 have different biological function. H3K27me1 is found to occur on active genes while the H3K27me2 is broadly distributed over the whole genome and was shown to silence non-cell-type specific enhancers (Barski et al. 2007; Ferrari et al. 2014).
The trimethylation of H3K27 is associated with repressed chromatin and is involved in X-chromosome inactivation and silencing of HOX genes, which are required for proper embryonic development (Fig. 4F) (Rougeulle et al. 2004).

The H3K27me3 modification is introduced by Enhancer of Zeste 2 (EZH2), which is part of the Polycomb repressive complex 2 (PRC2). PRC2 contains beside EZH2 also Suz12, EED, RbBP4 and is responsible for the silencing of developmental and cell differentiation genes (Boyer et al. 2006; Bracken et al. 2006; Payer and Lee 2008; Plath et al. 2003; Schuettengruber et al. 2007; Völkel and Angrand 2007). A recent study showed that the PRC2 complex can introduce all three methylation states of H3K27 (Ferrari et al. 2014). Afterwards, Polycomb repressive complex 1 (PRC1) interacts with H3K27me3, which leads to monoubiquitinylation of H2A on lysine 119 initiating chromatin packing.

The Polycomb proteins CBX2, CBX4, CBX6, CBX7 and CBX8 belong to the Royal superfamily of reading domains and generally prefer trimethylated lysine residues by forming an aromatic cage consisting of three aromatic residues, which clasp around the trimetylated lysine residue. The Polycomb group proteins were the first time identified in Drosophila involved in gene silencing and can be clustered into two main complexes PRC1 and PRC2. The PRC1 protein Chromobox homolog 7 (CBX7) contains an N-terminal Chromo domain which specifically interacts with the repressive H3K27me3 modification (Fig. 9) (Ren et al. 2015).

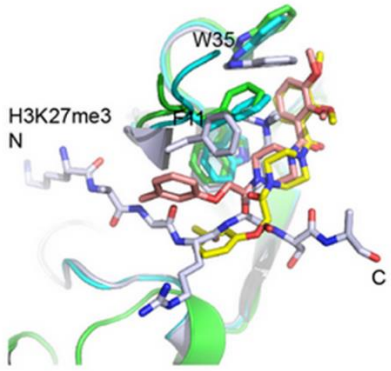


Figure 9. CBX7 Chromo domain bound to H3K27me3 peptide. Figure taken from (Ren et al. 2015).

1.2.2.7 Histone lysine acetylation

In contrast to the diverse functions of histone tail methylation, acetylation is always associated with chromatin decondensation and it is involved in transcriptional activation, DNA replication and repair (Kurdistani and Grunstein 2003).

Histone acetylation was found to occur on the N-terminal tails of H4 (in particular on residue K5, K8, K12 and K16) and of H3 (residue K9, K14, K18, K23 and K27), but also on the globular H3 (at position K56, K64, K115 and K122) and on H2A and H2B (Basu et al. 2009; Lawrence et al. 2016). Many different histone acetyltransferases have been identified, for example Gnc5, which catalyses the acetylation on H3 and H2B, and Esa1, which is responsible for the acetylation on H4 and H2A (Suka et al. 2001). Genome-wide studies have shown that acetylation and deacetylation occur in a rapid cycle (Kurdistani et al. 2004; Pokholok et al. 2005).

Acetylated histones alter the chromatin structure by neutralizing the positive charge of lysine residues leading to weakening of the interaction between DNA and histones. The loss of the strong interaction between DNA and acetylated histones leads to an open chromatin conformation and subsequently gene activation (Vettese-Dadey et al. 1996; Horn and Peterson 2002; Tse et al. 1998). In particular, acetylation of the H3-DNA contact region (globular H3 on position K56, K64, K115 and K122) weakens the interaction between DNA and the histone octamer (Di Cerbo et al. 2014; Yuan et al. 2009; Tropberger et al. 2013; Lawrence et al. 2016).

Histone acetylation was the first identified histone tail modification and was shown to be involved in the regulation of RNA synthesis (Allfrey et al. 1964). Also the first discovered

histone tail reader was the acetyl-lysine reading domain of PCAF (P300/CBP-associated factor), which recognizes acetylation marks on lysine residues of the histone tails with its Bromo domain.

Nuclear magnetic resonance (NMR) spectroscopy studies identified the intramolecular contacts between the Bromo domain of PCAF and the acetylation modification in a hydrophobic binding pocket embedded in a left-handed antiparallel four-helical bundle (Fig. 10A) (Taverna et al. 2007). More details about the specific recognition were discovered in the crystal structure of Gcn5p, the PCAF homolog of *Saccharomyces cerevisiae*, in complex with an acetylated H4 peptide. The acetyl group is insert in a deep and narrow binding pocket and anchored through hydrogen bonds (Taverna et al. 2007; Owen et al. 2000). The specific interaction is mediated by an asparagine residue located in the binding domain of PCAF and Gcn5p, interacting specifically with the acetyl-group of the histone residue (Mujtaba et al. 2007).

The Bromo domain of the transcription factor BPTF was reported to interact in the same way with H4K16ac modifications than the Bromo domains of PCAF and Gcn5p (Ruthenburg et al. 2011). The human TAF1 protein, which is part of TFIID recognizes multiple acetylated residues of the H4 tail with its double Bromo domain. Likewise to PCAF, both Bromo domains of TAF1 TFIID fold in a four-helix bundle, and they interact side by side with the acetyl groups (Taverna et al. 2007).

Other dual acetylation readers are the tandem Bromo domain of Rsc4p a subunit of the yeast chromatin remodelling complex RSC and the double Bromo domain of TAF1. TAF1 is the large subunit of TFIID, which is involved in initiating the assembly of the transcription machinery (VanDemark et al. 2007; Taverna et al. 2007). Tandem Bromo domains appear to fold as a single structural unit, whereas double Bromo domains appear to be relatively independent from each other. Therefore, the tandem Bromo domain of Rsc4p and the double Bromo domain of TAF1 show diverse biological functions.

The tandem Bromo domains of the Rsc4p protein are separated by 20 Å and recognizes mostly H4K5ac and H4K12ac (Fig. 10B). Because the double Bromo domains of the TAF1 fold independently from each other, they are separated by 25 Å and therefore can recognize more flexible modifications apart from each other (Fig. 10C). Rsc4p was identified to interact with H3K14ac and its own K25ac modification (Taverna et al. 2007).

Numerous proteins with multiple Bromo domains exist, illustrating the fact that multiple histone interacting domains can be combined to increase the specificity and binding affinity towards histone tail PTMs (Kostrhon et al. 2017).

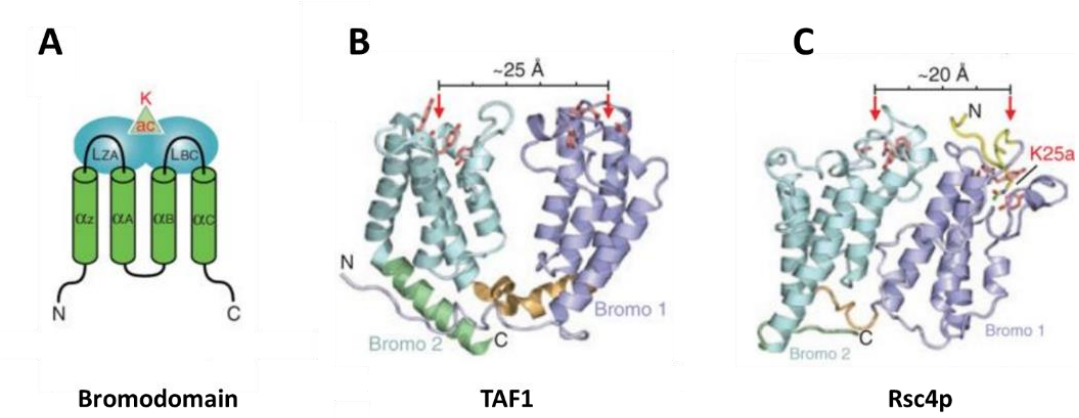


Figure 10. Readout of acetylation modifications by Bromo domains. **A)** Structure of a Bromo domain interacting with acetylation mark (triangle). **B)** Structural differences of double Bromo domain of TAF1 and **C)** Tandem Bromo domain of Rsc4p leading to specific readout of differently modified histone tails. Figure adapted from (Taverna et al. 2007).

These examples clearly show that histone proteins are highly modified and that every modification has a particular influence on the gene expression. However, in addition to this it has been observed that defined combinations of histone tail modifications can also influence the complex chromatin structure and affect gene expression, as described in the next chapter of this thesis.

1.3 Complex patterns of epigenetic modifications

1.3.1 Co-occurrence of histone tail modifications

Chromatin modifications extend the genetic information beyond the DNA sequence and are involved in various chromatin-templated processes. As described before, each post-translational histone modification has specific influences on the chromatin network. In 2000, Strahl and colleagues introduced for the first time the possibility of a “histone code” (Strahl and Allis 2000) hypothesizing that combinations of different chromatin modifications may act synergistically and have specific roles, distinct from the mere sum of the roles of individual marks. Furthermore it was proposed that defined combinations of modifications can be

interpreted by multiple reading domains or proteins (Nightingale et al. 2006; Margueron et al. 2005).

Interaction with combinations of PTMs could increase the readout specificity and affinity towards certain combinations resulting in a highly complex and specific chromatin environment. It was suggested that one modification alone is not sufficient and specific enough to recruit a given complex to a specific place and that the combinatorial interaction might increase the specific regulation of chromatin-templated processes (Ruthenburg et al. 2007).

The histone code hypothesis is strongly supported by the fact that almost all chromatin interacting complexes contain different histone tail PTMs binding domains specific for individual histone tail modifications, which results in a specific combinatorial readout (Ruthenburg et al. 2007; Musselman et al. 2012; Su and Denu 2016).

One well-studied example of a combinatorial histone mark is the co-occurrence of H3K27me3 and H3K4me3 on promoters in embryonic stem cells (ESC). This so called “bivalent” modification keeps the promoter in a poised state which allows a fast responds to cellular signals. Keeping either the repressive H3K27me3 modification or the active H3K4me3 modification leads to the respective gene expression changes (Vastenhouw and Schier 2012; Voigt et al. 2013).

Another example for the co-occurrence of two histone tail modifications, found in mass spectrometry studies, is the co-localization of H3K36me3 and H3K9me3 in gene bodies which is associated with alternative splicing (Kolasinska-Zwierz et al. 2009; Saint-André et al. 2011). One well-studied crosstalk of modifications is the stimulation of H3K4me3 and H3K79me3 by monoubiquitinylation of H2BK120 (H2BK123 in yeast) (Sun and Allis 2002). This crosstalk is conserved from yeast to mammals and it was seen that loss of H2BK120 monoubiquitinylation leads to a global reduction of H3K4me3 (Zhang et al. 2015b; Kim et al. 2005). Additionally, a crosstalk between H2BK34 monoubiquitinylation and H3K4me3 was observed (Wu et al. 2011b).

Moreover, it is well established that H3K4me3 and histone acetylation are both marks associated with active transcription which coexist at promoter and TSS of active transcribed genes, where they can recruit several protein complexes with combined reading domains. For

example, the BPTF protein, a subunit of the chromatin remodelling complex NURF, is recruited to histone tails carrying H3K4me3 and acetylation marks. BPTF contains a plant homeodomain finger (PHD) and a Bromo domain to recognize H3K4me3 and H4K16, respectively (Badenhorst et al. 2002; Li et al. 2006; Ruthenburg et al. 2011).

Many reports demonstrated clearly that H3K9me3 and H3K27me3 are mutually exclusive (Pauler et al. 2008; Cooper et al. 2014; Peters et al. 2003; Saksouk et al. 2014).
Similarly, H3K36me3 and H3K27me3 are the main antagonists in the establishment of active and repressive chromatin states and the inverse correlation between H3K36me3 and H3K27me3 was extensively studied in animals and plants (Voigt et al. 2012; Yuan et al. 2011). Furthermore H3K36me3 was shown to have an inhibitory effect on the histone methyltransferase activity of PRC2, which is responsible for introducing the H3K27me3 modification (Voigt et al. 2012; Barski et al. 2007).

The high diversity of histone tail PTMs and the combinatorial interplay between the modifications make the chromatin network highly specific and complex. Therefore, it is of great importance to understand the complexity of the histone code and gain a deeper understanding of the biological functions and diseases associated with combinatorial chromatin modifications.

1.3.2 Co-occurrence of histone tail modification and DNA methylation

There are several DNA binding proteins known, which connect DNA methylation and histone tail modifications (Lewis et al. 2004; Umlauf et al. 2004a).
For example, the methylated CpG binding proteins MeCP2 and MBD 1-4 as well as the BTB/POZ family members also known as Kaiso or ZBTB 33 (Zinc finger and BTB domain containing protein 33), localize to methylated promoters and recruit effector proteins like histone deacetylases (HDACs) or histone methyltransferases (Kondo 2009).

MeCP2 binds methylated promoter regions and recruits SUV39H1 and SUV39H2 histone methyltransferases which induce H3K9me3 methylation and gene silencing (Fuks et al. 2003; Lunyak et al. 2002). In addition, the MBD1 protein interacts with the H3K9 methyltransferase

SetDB1, therefore playing a role in the maintenance of H3K9 methylation at heterochromatin regions during DNA replication (Sarraf and Stancheva 2004).

DNA methylation as well as the H3K36me3 modification occur in the gene bodies of actively transcribed genes and the PWWP domain of Dnmt3a and Dnmt3b interact with the H3K36me3 modification (Dhayalan et al. 2010). Furthermore, H3K36me3 can recruit histone deacetylases to prevent spurious transcription initiation within gene bodies (Joshi and Struhl 2005; Carrozza et al. 2005; Keogh et al. 2005). In contrast, the H3K4me3 modification is known to inhibit DNA methylation because the ADD domain of the Dnmts interact only with the H3 histone tail if H3K4 is unmodified (Ooi et al. 2007; Zhang et al. 2010).

Another example of the interaction between epigenetic signals is the human ubiquitin-like PHD and RING finger containing 1 (UHRF1) E3 ubiquitin ligase. It consists of multiple domains, which can interact with histone tail modifications and DNA methylation. UHRF1 was shown to be essential for the establishment and maintenance of DNA methylation. The tandem Tudor domain of UHRF1 is responsible for the interaction with H3K9me3 and its SRA (SET and RING associated) domain mediate binding to hemimethylated CpG dinucleotides (Rothbart et al. 2012; Arita et al. 2008; Hashimoto et al. 2008). The interaction of UHRF1 with H3K9me3 modified histone tails leads to an ubiquitination of the H3 tail at lysine 23. This modification further recruits the DNA methyltransferase 1 (Dnmt1) to DNA replication sites for maintaining DNA methylation at this positions (Nishiyama et al. 2013).

These examples clearly demonstrate that epigenetic signals influence each other and that an accurate interaction between all signals is necessary for normal cell development. Furthermore, it is important to have methods, which enable the genome-wide analysis of PTMs to elucidate the complexity of the epigenetic code.

1.4 Methods to study epigenetic modifications

1.4.1 Mass spectrometry

One method to investigate the global occurrence of histone PTMs is mass spectrometry (MS). Different histone modifications add different weight to the histone proteins therefore making it possible to identify histone tail modifications residing on the isolated histone proteins. It is

also possible to determine the co-occurrence of PTMs, if they are located next to each other on the histone protein sequence (Guan et al. 2013; Britton et al. 2011).

For mass spectrometric analysis, the histone of interest is digested using a serine protease, typically trypsin, to specifically hydrolyse peptide bonds at the carboxyl side of lysine and arginine residues. However, histones contain many arginine and lysine residues and trypsin cleavage results in a complex accumulation of small diverse peptides. This makes the analysis and identification of histone PTMs a difficult task. Another possibility is the application of endoproteases like GluC or AspN resulting in longer peptides (Karch et al. 2013). The co-occurrence of histone PTMs is best studied with top-down MS, which uses intact proteins and allows the identification of several histone tail modifications simultaneously (Taverna et al. 2007).

However, mass spectrometry does not provide any information about the genome-wide localisation and distribution of the analyzed histone PTMs. This is an important limitation, because the genome-wide localisation of histone PTMs and the association to particular genome region is of high importance to understand chromatin-templated processes.

1.4.2 Application of reading domains in living cells for global detection of chromatin modifications

Single and double reading domains can be used to specifically detect epigenetic modifications in live cells analyzed by microscopy techniques, which were recently shown in several publications. For example, DNA methylation changes during development and differentiation in mouse ESC cells were analyzed using a fluorescent MBD protein, using a mouse strain containing a knocked-in fluorophore fusions protein and DNA methylation was studied by live-cell imaging (Ueda et al. 2014).

Furthermore, single reading domains can be used to analyze histone tail PTMs in live cells. For example, histone methylation reporters were designed to track histone methylation in single, living cells. To this end, the Chromo domains of HP1 and Polycomb were fused to fluorescent reporters to analyze H3K9me3 and H3K27me3, respectively (Lin et al. 2003). Furthermore, fluorescent resonance energy transfer (FRET) was applied to analyze histone acetylation using Bromo domains (Kanno et al. 2004; Sasaki et al. 2009). Additionally, the Bromo domains of the

human polybromo -1 protein were used to create a recombinant protein sensor to track the H3K14ac modification in HEK293T cells (Sanchez et al. 2017).

A recently published study documented the usage of reading domains to analyze complex chromatin modification patterns in living cells. Therefore, two reading domains have been coupled to a fluorescent reporter to globally visualize bivalent chromatin states in living ESC (Delachat et al. 2018). Genetically encoded chromatin-sensing multivalent probes based on reading domains (cMAPs) have been developed to analyze nucleosomes containing H3K4me3 and H3K27me3 simultaneously. In addition, the loss of double marks upon treatment with small-molecule epigenetic modulators had been investigated. However, these global microscopic studies in general does not provide information about the genomic location of the corresponding modifications.

To receive information about epigenetic modifications at a particular loci, the recently published bimolecular fluorescence complementation approach can be used. This method uses DNA and histone PTM reading domains to study epigenetic signals in living cells at a certain loci. For this, the DNA binding domain of the MBD1 protein and the H3K9me3 reading domain of HP1ß have been individually fused to non-fluorescent fragments of a fluorescent protein and co-expressed in cells. If the epigenetic mark is present at the target DNA sequence, a functional fluorophore is reconstituted. This allows the direct readout of locus-specific epigenetic signals. This method allows the first time the direct detection of DNA methylation and H3K9me3 modification at specific genome sites in living cells. Furthermore, the investigation of dynamic changes is possible using this method (Lungu et al. 2017).

However, different loci can be analyzed with this method only one after each other, because no multiplexing is possible. Furthermore, the genome-wide distribution of modifications can be only studied using ChIP-seq and CIDOP-seq experiments, but these methods disrupt the cell and do not allow to correlate individual epigenetic profiles with particular cellular phenotypes.

1.4.3 Genome-wide analysis of histone PTMs using affinity reagents

Locus specific analysis of chromatin marks is based on a single analytical technology. Modification specific pull down of chromatin with a specific affinity reagents, followed by downstream analysis procedures like locus specific quantitative PCR (qPCR) or next generation sequencing (NGS), allow the locus specific analysis. Affinity reagents can be antibodies (established) or reading domains (developed here). These classes of reagents will be introduced in the next chapters.

Chromatin immunoprecipitation (ChIP) is the key method to analyze the genomic localisation of chromatin modifications. There are two ChIP strategies available differing in their starting chromatin preparation procedure. The cross-linked chromatin immunoprecipitation (X-ChIP) involves the reversible cross-linking of cells with formaldehyde or UV-light (Gilmour and Lis 1984; Kuo and Allis 1999). This method is suitable for the investigation of the genome-wide binding of transcription factors or chromatin interacting proteins. After cross-linking, chromatin extraction is followed by sonication or micrococcal nuclease digestion to obtain chromatin fragments. Antibodies against the proteins of interests are used to selectively precipitate the chromatin fragments with bound protein of interest.

An alternative approaches is native chromatin immunoprecipitation (nChIP). With this method, proteins that are tightly bound to DNA can be investigated. Therefore, nChIP allows the investigation of histone variants and histone PTMs. The interaction of histones and DNA is strong under physiological salt conditions and the chromatin is sonicated or digested using micrococcal nuclease which selectively cuts the linker DNA, leaving the nucleosomes intact (Umlauf et al. 2004b; Pillai and Chellappan 2009).

Both methods have their limitations and advantages. X-ChIP is the only way to study weakly bound proteins, but crosslinking can create artificial interactions. In contrast, nChIP is only suitable for tightly bound proteins but proteins remain in their native state and nucleosomes can get lost during the preparation of the chromatin.

Nevertheless, in this type of experiments, chromatin is precipitated using antibodies specific for the PTM of interest and therefore, the results of such experiments are highly dependent on the antibody quality. Recognizing histone PTMs is challenging for antibodies because of the

minor structural differences especially among methylation states as well as the high sequence similarity surrounding the different modification sites. It has been observed in several studies that histone tail antibodies are often not as specific and reliable as required (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; Kungulovski et al. 2015b, 2014, 2015a).

As an alternative to histone tail antibodies also histone modification interacting domains (HiMIDs) can be used in chromatin precipitation experiments (Kungulovski et al. 2014). Chromatin preparation is done like described before but GST-tagged HiMIDs are used instead of antibodies to precipitate chromatin (Fig. 11).

Regardless of the used method, the precipitated DNA can be investigated with quantitative PCR (ChIP-qPCR/CIDOP-qPCR) or next generation sequencing (ChIP-seq/CIDOP-seq) and conclusions about the genomic localisation and involvement in biological processes can be drawn.

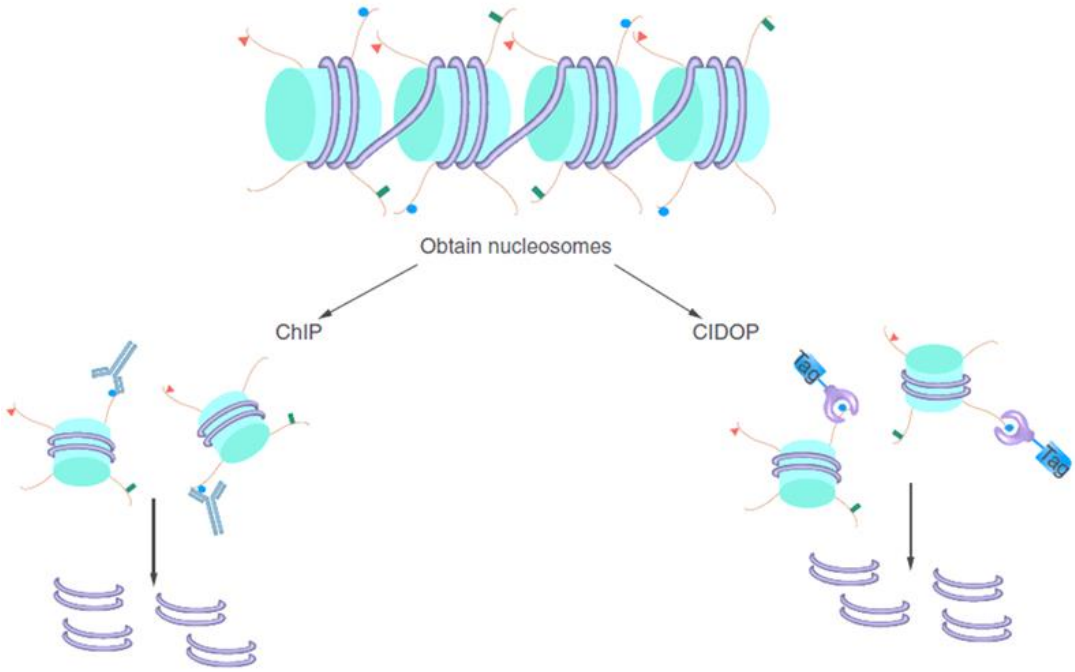


Fig 11. Principle of chromatin precipitation experiments. Chromatin is fragmented and precipitated either with antibodies (ChIP) or with GST-tagged HiMIDs (CIDOP). Isolated DNA can be further investigated using quantitative PCR (qPCR) or next generation sequencing (NGS). ChIP: Chromatin immunoprecipitation, CIDOP: Chromatin interaction domain precipitation, HiMID: histone modification interacting domains. Figure taken from (Kungulovski et al. 2015b).

1.4.3.1 Sequential-ChIP (re-ChIP)

To analyze the co-occurrence of histone tail modifications and their genome-wide localisation, sequential-ChIP (re-ChIP) can be use. For this, two consecutive ChIPs have to be conducted. After performing the first ChIP the recovered chromatin is incubated with a second antibody against another PTM. After the second ChIP the received DNA refers to nucleosomes carrying both investigated PTMs. However, this method requires a lot of starting material and nevertheless the recovered amounts of DNA are quite low, which makes the library preparation for next generation sequencing very difficult. Consequently, the co-occurrence of histone tail modifications are often only analyzed by semi-quantitative or quantitative PCR.

Another approach to study the co-existence of histone modifications in natural chromatin, is the recently developed co-ChIP method (Weiner et al. 2016). This approach uses the technology of barcoding histones after immunoprecipitation to identify dual modified chromatin states. Magnetic beads, each carrying a specific PTM antibody, are incubated with nucleosomes and after immunoprecipitation, nucleosomes are fused to a unique barcode specific for the respective modification. Then, the first antibody is inactivated and the barcoded nucleosomes are released. The pool of barcoded nucleosomes is then incubated with an antibody recognizing the second PTM to analyze nucleosomes carrying the two modifications of interest. The isolated DNA associated with both histone tail modifications can be further analyzed with next generation sequencing.

1.5 Affinity reagents for chromatin immunoprecipitation experiments

1.5.1 Antibodies

Antibodies are extremely important reagents in modern molecular life science with numerous therapeutic-, diagnostic- and research-based applications. Antibodies, also known as immunoglobulins (Igs), are able to recognize specific epitopes and are used by the immune system to identify foreign compounds (Meffre et al. 2000).

These heavy glycoproteins have a characteristic Y-shape, comprising of two heavy and two light chains connected by disulphide bonds. The heavy chain defines the five different isoforms of antibodies, namely IgA, IgD, IgE, IgG, and IgM. Each heavy and light chain contains variable

and constant regions (Fig. 12 and Fig. 13A). The variable regions of the heavy and light chains are responsible for the antigen binding. They are genetically created by V(D)J recombination, a process in which the random combination of different segments of the variable- (V), diversity- (D) and joining- (J) genes leads to a high diversity in sequence and structure of each antibody (Wilson and Stanfield 2014; Roth 2014; Schatz and Swanson 2011; Helmink and Sleckman 2012).

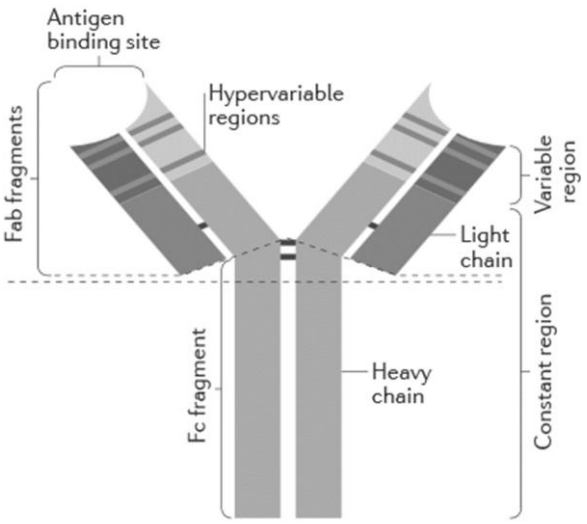


Figure 12. Schematic image of the typical antibody structure comprising two heavy (light grey) and two light chains (dark grey) connected via disulphide bonds. The lower part represents the constant region (Fc), whereas the upper part is variable and responsible for the antigen binding (Fab). Figure adopted from (Crisp et al. 2016).

Antibodies are frequently used in clinical diagnostics, as therapeutic agents and as research reagents. Polyclonal antibodies are widely used but each polyclonal antibody lot is limited and therefore long-term studies are not feasible. Furthermore, reproduction of results is difficult and sometimes not possible, if the polyclonal antibody lot used for a previous study is no longer available. Additionally, the specificity of a polyclonal antibody needs to be validated for every lot individually.

To overcome these limitations, monoclonal antibodies have been developed. Monoclonal antibodies derive from a single immune cell, which is immortalized using the hybridoma technology, where the antibody producing spleen cells from mice or rats is fused with immortal myeloma cells (Köhler and Milstein 1975). Therefore, monoclonal antibodies specifically target one desired epitope. Furthermore, production of one particular antibody is

possible as long as the cell line is maintained. Nevertheless, the production of monoclonal antibodies is more difficult and more expensive than of polyclonal antibodies.

Due to their frequently application, a great interest in the development and engineering of antibodies appeared and great efforts in this direction were undertaken. Therefore, an alternative for mono- and polyclonal antibodies is the generation of recombinant antibodies. These recombinant antibodies are generated by vector construction and transfection into a host cell leading to consistent quality of these antibodies. The specificity and sensitivity can be improved through antibody engineering and no batch-to-batch variability exits, which ensure reproducibility of data and allows long-term studies.

For the recombinant production of antibodies, several different hosts cells can be used depending on the standards the antibodies have to fulfil. For example, the most prominent cell lines for antibody production are Chinese hamster ovary (CHO) cells as well as human embryonic kidney 293 (HEK293) cells or the human embryonic retinal cell line Per.C6. But also *E.coli*, yeast or insect cells can be used to generated recombinant antibodies (Frenzel et al. 2013).

To further increase the production efficiency, smaller antibody fragments have been developed which can be easier produced but at the same time retain the antigen binding capacity of the full-length antibody (Fig. 13B-E). Through versatile techniques of genetic engineering, domains of the light and heavy chain can be combined in different ways to form recombinant antigen-binding fragments (Fab) (Nelson and Reichert 2009). The variable region of heavy and light chains can be fused together via short linker peptides to create a single-chain variable fragment (scFv). Both, the scFv and Fab, are the most frequently used recombinant antibodies today. Furthermore, bispecific antibodies (bsAB) against two or more different specific antigen binding elements or epitopes exist, which are suitable for therapeutic and diagnostic applications (Kontermann and Brinkmann 2015; Weidle et al. 2014; Müller and Kontermann 2010).

Additionally, multimeric antibodies like dia-, tria-, or tetra-bodies or minibodies (miniABs) where the scFv is linked to an oligomerization domains can be efficiently generated (Hudson and Kortt 1999; Hust et al. 2007; Hu et al. 1996) and showing advantages in some therapeutically applications (Holliger and Hudson 2005; Townsend et al. 2006).

The field of antibody engineering is constantly developing to find new strategies to improve the expression, the stability and the affinity of antibodies (Chiu and Gilliland 2016; Hudson and Souriau 2003).

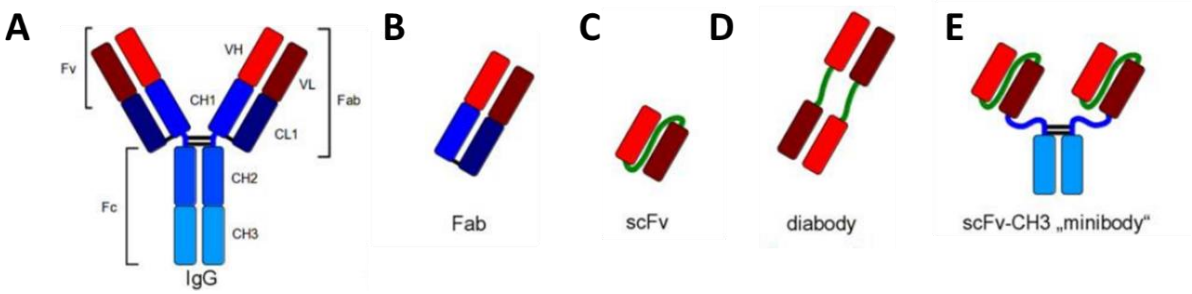


Figure 13. Schematic drawing of different recombinant antibodies. **A)** IgG antibody. **B)** Antigen binding fragment (Fab). **C)** Single chain fragment (scFv). **D)** Diabody, combining two different antigen binding domains in one molecule. **E)** Minibodies (miniABs) scFv is linked to an oligomerization domains. Figure adapted from (Frenzel et al. 2013).

1.5.2 Reading domains

Inside the cell, reading domains of proteins or large protein complexes are known to specifically interact with various PTMs and even a combination of PTMs. These reading domains were shown to successfully operate in chromatin precipitation assays as an alternative for histone tail antibodies because of their reliable and specific interaction with PTMs (Kungulovski et al. 2014). In this study the applicative potential of naturally occurring and engineered histone modification interacting domains (HiMIDs) were shown. It was demonstrated that HiMIDs can successfully replace histone tail antibodies in ChIP like experiment called CIDOP (Chromatin interacting domain precipitation).

Histone modification interacting domains have several advantages over antibodies. Most importantly, they can be easily produced with constant quality in *E.coli* and no batch-to batch variability exists. Also the production costs are much cheaper than for antibodies. Furthermore, the large number of available crystal structures enable the generation of binding pocket mutants, which serve as a perfect negative control. Furthermore, the binding specificity can be altered by systematic mutagenesis (Kungulovski et al. 2014).

The following reading domains were already successfully used as chromatin affinity precipitation reagents; MPP8 Chromo domain, ATRX ADD domain, Dnmt3a PWWP domain, CBX7 Chromo domain and KDM4A double Tudor domain (Kungulovski et al. 2014).

The DNA methyltransferase 3a (Dnmt3a) comprises several domains. One of these domains is the PWWP domain, which belong to the Royal superfamily of typically 2-4 aromatic residues optimal for the interaction with the methyl-lysine ε-amino group. The PWWP domain of Dnmt3a was shown to interact specific with H3K36me2/3 (Bock et al. 2011b; Dhayalan et al. 2010; Yun et al. 2011; Min et al. 2003; Wu et al. 2011a).

Another member of the Royal superfamily is the Chromobox protein homolog 7 (CBX7), which is a component of the Polycomb group proteins. The CBX7 Chromo domain was identified to interact with H3K27me3 (Bernstein et al. 2006b).

The TATA box-binding protein-associated factor 3 (TAF3) harbours a PHD domain which interacts with H3K4me3 and therefore locates the transcription factor TFIID to nucleosomes containing H3K4me3 modifications and stimulates the pre-initiation complex formation with the RNA polymerase II (Vermeulen et al. 2007; Lauberth et al. 2013).

The M-phase phosphoprotein 8 also known as MPP8 mediates the recruitment of the HUSH complex to H3K9me3 and it was shown that the Chromo domain of MPP8 specifically recognize H3K9me3 histone tail modifications (Bock et al. 2011b; Chang et al. 2011).

The large chromatin-associated nuclear protein ATRX belongs to the SNF2 family, which are chromatin remodelling complexes. Besides other domains, it contains an ADD domain interacting specific with H3K9me3 histone tail modification if H3K4 is unmodified (Dhayalan et al. 2011; Iwase et al. 2011).

The double Tudor domain of the lysine-specific demethylase 4A (KDM4A) also known as JMJD2A demethylates specific H3K9me3 and H3K36me3 while it has no activity on mono- or dimethylated residues. In the work of (Kungulovski et al. 2014) the KDM4A double Tudor domain variant D969A was engineered as a specific H4K20me3 reading domain.

However, there are many more HiMIDs known which can be investigated as alternatives to histone modification antibodies and used in chromatin studies.

1.6 Methods for quality control of affinity reagents

1.6.1 Designer nucleosomes

Designer nucleosomes with defined modifications are currently the best available method to validate the specificity of chromatin affinity reagents.

The chemical synthesis of peptides allows the specific incorporation of defined modifications. However, peptide synthesis of peptides longer than 50 aa is not efficient such that it is not possible to synthesize entire proteins with modifications. There are two methods to overcome this problem, native chemical ligation (NLC) and expressed protein ligation (EPL), which allow the generation of histone proteins containing defined PTMs.

Native chemical ligation connects two synthetic peptides via a nucleophilic attack. One peptide has to contain an N-terminal cysteine residue and the second peptide is activated by the presence of a C-terminal thioester. Both peptides can carry the modification of interest, the C-terminal part can also be a recombinant expressed protein. The intermediate product contains a five-membered ring which undergoes a rapid rearrangement resulting in an end product with a native peptide bond at the ligation site (Dawson et al. 1994).

The EPL method exploits the natural occurring protein inteins to prepare a recombinant polypeptide activated with a C-terminal thioester. The recombinant thioester can be ligated to synthetic peptides containing an N-terminal cysteine (Muir et al. 1998). However, these methods include many steps, which need to be validated carefully. Furthermore, specific expertise and special knowhow is required to perform experiment using designer nucleosomes, not every lab can afford. Nevertheless, this method currently is the most powerful approach to study the specificity of chromatin affinity reagents but also of natural reading domains.

Designer nucleosomes are difficult to generate and requires special knowledge and knowhow. An alternative to introduce defined lysine methylation in target proteins including histones are methyl-lysine analogues. This technique uses a chemical approach to install methyl-lysine analogues into recombinant proteins (Simon et al. 2007). However, only modifications on lysine residues can be introduced and the simultaneously introduction of different methylation degrees is not possible with this method.

The target lysine in the protein sequence is replaced by a cysteine and alkylated with an electrophilic ethylamine containing one, two or three methyl groups at the amino group, resulting in (methyl) aminoethylcysteine. Depending on the used chemicals site specific mono-, di- or trimethylated aminoethylcysteine residues are generated simulating the natural lysine methylation (Fig. 14). For the introduction of the trimethylated aminoethylcysteine (2-bromoethyl) trimethyl ammonium bromide is used.

The methyl-lysine analogues has been demonstrated to be functionally and chemically similar to their natural counterparts and are ideal to study the influence of lysine residue methylation on the binding of reading proteins, because it was demonstrated that methyl-lysine analogues are equally bound by many reading domains (though not all of them) (Wilkinson and Gozani 2014).

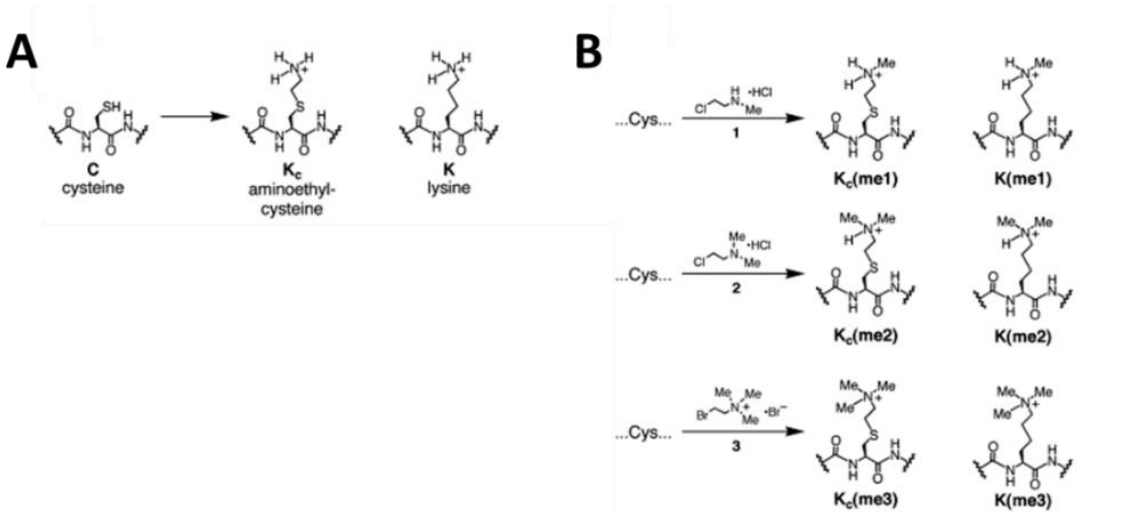


Figure 14. Installation of methyl-lysine analogues into recombinant proteins. **A)** The target lysine needs to be exchanged by a cysteine and the thiolgroup can be converted to a methyl-lysine analogue. **B)** Dependent on the alkylation reagents mono-, di-, or trimethylated lysine analogues can be generated. Figure taken from (Simon et al. 2007).

1.6.2 Peptide arrays

Another approach to validate the specificity of chromatin affinity reagents is presented by peptide arrays. These arrays are technically simpler to generate, than designer nucleosomes.

Short peptides, reflecting parts of the unstructured N-terminal histone tail, are a useful tool to investigate the biochemical behaviour of histone interacting proteins *in vitro*. Therefore, peptide arrays are a good method to study the binding preference of histone tail interacting domains or antibodies. Additionally, the influence of neighbouring marks on the binding specificity of the histone tail interacting domain or antibody can be investigated (Bock et al. 2011a, 2011b; Bua et al. 2009; Dhayalan et al. 2010; Garske et al. 2010, 2008; Rathert et al. 2008a; Su et al. 2014; Zhang et al. 2010). Peptide arrays reflect the natural combinatorial diversity of chromatin modifications but the combinatorial interaction of histone tail interacting proteins bearing multiple domains cannot be studied with peptide arrays.

Detailed validation of the affinity reagents for chromatin immunoprecipitation experiments showed, that histone tails antibodies are not always specific enough to receive significant results, but reading domains showed a higher specificity towards their target modifications.

Aims of the study

Epigenetic signals such as histone modifications and DNA methylation alter the chromatin structure and influence the accessibility of DNA. These epigenetic signals are essential for normal cell differentiation and development and are known to be altered in diseases such as cancer or autoimmune disorders. Hence, it is of fundamental importance to understand how epigenetic modifications are set, read and erased.

There are many different PTMs known to date and due to the high diversity of chromatin modifications, it is likely that epigenetic signals influence each other leading to specific biological processes. Since the hypothesis of the histone code was proposed in 2000, the research on combinatorial histone PTMs raised great interest. Therefore, an important question to answer is how the co-occurrence of PTMs correlates with their interpretation and the biological signals they induce.

So far, it is still an experimentally difficult task to map the co-existence of two or more histone PTMs genome-wide. Therefore, in this work it was planned to generate and validate double reading domains for the simultaneous readout of two histone PTMs to analyze the co-occurrence of distinct histone modifications on the same histone tail.

In previous studies it was already shown that HiMIDs can be reliably used as an alternative to histone PTMs antibodies. These reading domains occur naturally as part of chromatin interacting proteins and are able to specifically interact with defined histone modifications. In this work, the existing technology should be further developed and refined to achieve simultaneous readout of two histone modifications on the same nucleosome. To this end, it was planned to identify additional HiMIDs for important modifications, and couple validated histone interacting domains together to generate several dual combinations of HiMIDs to analyze the simultaneous presence of both mark. Furthermore, it was planned to develop approaches to analyze the specificity of these double reading domains.

Finally, it was intended to use these novel tools to study the genome-wide distribution of dual chromatin marks and to identify new bivalent chromatin states. Thereby, the present study should also document that this new technology is very powerful and it can contribute to the deciphering of the histone code improve our understanding of chromatin-templated processes.

2 Materials and methods

2.1 Cloning, protein purification and site-directed mutagenesis

The sequences encoding for the reading domains were amplified from existing plasmids (Kungulovski et al. 2014) or cDNA prepared from HEK293 or HepG2 cells. The inserts were cloned via Gibson assembly (Gibson et al. 2009) as GST-fusion proteins (vector pGEX-6p-2) or MBP-fusion proteins (vector pMAL-c2X). As linker, the following aa were used to ensure flexibility of each individual reading domain in the context of the fusion protein:
amino acid sequence S S G N S N A N S R G P S F S S G L V P L S L R G S H.

The linker will be abbreviated as "L" in the names of the double reading domains. In all cases when Dnmt3a PWWP was the first domain in the fusion construct, no artificial linker was introduced because the two domains were connected using a natural linker resulting from the sequence between the PWWP domain and the ADD domain of Dnmt3a (see appendix 4.1, Fig. S6 in (Mauser et al. 2017)). Therefore, no additional abbreviation is used for these constructs. All used single and double reading domains prepared in this work are listed in table 1 and table 2, together with additional information.

The fusion proteins were overexpressed in *E.coli* BL21-CodonPlus (Novagen) at 37 °C in LB medium until the culture reached an OD_{600} of 0.6 - 0.8 and induction was done with 1mM IPTG final. The induced culture was shifted to 20 °C and production of protein was carried out overnight. Purification for the GST-tagged and the MBP-tagged proteins was done using affinity chromatography as described in (Rathert et al. 2008b). For proteins, containing the TAF3 PHD domain or the ATRX ADD domain 50 - 150 µM zinc was added to the LB-media and the purification buffers. Binding pocket mutations (table 3) were introduced by site-directed mutagenesis (Jeltsch and Lanio 2002) and validated by restriction analysis and Sanger sequencing.

Table 1: Overview of the single reading domains used in this study, including abbreviations, domain borders and known interaction with histone PTMs.

Name (abbreviation)	Domain	Accession number	Domain boundaries (aa)	Interaction with histone tail modification
TAF3 PHD domain (T)	TAF3 PHD	NP_114129.1.1	856–929	H3K4me3
MPP8 Chromo (MPP8)	MPP8 Chromo	Q99549	59-118	H3K9me3
Dnmt3a PWWP (PWWP)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me3
ATRX ADD (ADD)	ATRX ADD	P46100	159-296	H3K9me3/H3K4unm
CBX7 Chromo (CBX7)	CBX7 Chromo	O95931	11-69	H3K27me3
MBD2 (MBD2)	Methyl-binding domain 2	Q9UBB5	147-220	5mC
MBD1 (MBD1)	Methyl-binding domain 1	Q9UIS9	1-80	5mC

Table 2: Overview of the designed double reading domains including abbreviations, domain borders and known interaction with histone PTMs.

Name (abbreviation)	Domains	Accession number	Domain boundaries (aa)	Interaction with histone tail modification
Dnmt3a PWWP-Linker-MPP8 Chromo (PM)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/me3
	MPP8 Chromo	Q99549	59-118	H3K9me3
MPP8 Chromo-Linker-MPP8 Chromo (MLM)	MPP8 Chromo	Q99549	59-118	H3K9me3
	MPP8 Chromo	Q99549	59-118	H3K9me3
Dnmt3a PWWP-Linker-MPP8 Chromo (PLM)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/me3
	MPP8 Chromo	Q99549	59-118	H3K9me3
Dnmt3a PWWP-Linker-Dnmt3a PWWP (PLP)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/me3
	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/me3
MPP8 Chromo-Linker-double Tudor (MLdT)	MPP8 Chromo	Q99549	59-118	H3K9me3
	JMJD2A Double Tudor	O75164	897-1011	H4K20me3
double Tudor-Linker-MPP8 Chromo (dTLM)	JMJD2A Double Tudor	O75164	897-1011	H4K20me3
	MPP8 Chromo	Q99549	59-118	H3K9me3

Name (abbreviation)	Domains	Accession number	Domain boundaries (aa)	Interaction with histone tail modification
Dnmt3a PWWP-Linker-TAF3 PHD (PT)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
	TAF3 PHD	Q5VWG9	865-915	H3K4me3
TAF3 PHD-Linker-CBX7 Chromo (TLC)	TAF3 PHD	Q5VWG9	865-915	H3K4me3
	CBX7 Chromo	O95931	11-69	H3K27me3
TAF3 PHD-Linker-Dnmt3a PWWP (TLP)	TAF3 PHD	Q5VWG9	865-915	H3K4me3
	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
CBX7 Chromo-Linker-MPP8 Chromo (CLM)	CBX7 Chromo	O95931	11-69	H3K27me3
	MPP8 Chromo	Q99549	59-118	H3K9me3
CBX7 Chromo-Linker-TAF3 PHD (CLT)	CBX7 Chromo	O95931	11-69	H3K27me3
	TAF3 PHD	Q5VWG9	865-915	H3K4me3
Dnmt3a PWWP-ATRX ADD (PA)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
	ATRX ADD	P46100	159-296	H3K9me3/H3K4unm
BPFT PHD domain and Bromodomain (BPTF-PB)	PHD domain	gi:31322942	2743-2907	PHD: H3K4me3
	Bromo domain			Bromo: H4K16ac
MBD2-Linker-MPP8 Chromo (M2LM)	Methyl-binding domain 2	Q9UBB5	147-220	5mC
	MPP8 Chromo	Q99549	59-118	H3K9me3
MBD2-Linker-Dnmt3a PWWP (M2LP)	Methyl-binding domain 2	Q9UBB5	147-220	5mC
	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
Dnmt3a PWWP-Linker-MBD2 (PLM2)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
	Methyl-binding domain 2	Q9UBB5	147-220	5mC
ATRX ADD-Linker-MBD2 (ALM2)	ATRX ADD	P46100	159-296	H3K9me3/H3K4unm
	Methyl-binding domain 2	Q9UBB5	147-220	5mC
CBX7 Chromo-Linker-MBD2 (CLM2)	CBX7 Chromo	O95931	11-69	H3K27me3
	Methyl-binding domain 2	Q9UBB5	147-220	5mC
MPP8 Chromo-Linker-MBD2 (MLM2)	MPP8 Chromo	Q99549	59-118	H3K9me3
	Methyl-binding domain 2	Q9UBB5	147-220	5mC
MBD2-Linker-ATRX ADD (M2LA)	Methyl-binding domain 2	Q9UBB5	147-220	5mC
	ATRX ADD	P46100	159-296	H3K9me3/H3K4unm

Name (abbreviation)	Domains	Accession number	Domain boundaries (aa)	Interaction with histone tail modification
MBD1-Linker-TAF3 PHD (M1LT)	Methyl-binding domain 1	Q9UIS9	1-80	5mC
	TAF3 PHD	Q5VWG9	865-915	H3K4me3
ATRX ADD-Linker-MBD1 (ALM)	ATRX ADD	P46100	159-296	H3K9me3/H3K4unm
	Methyl-binding domain 1	Q9UIS9	1-80	5mC
MBD1-Linker-Dnmt3a PWWP (M1LP)	Methyl-binding domain 1	Q9UIS9	1-80	5mC
	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
Dnmt3a PWWP-Linker-MBD1 (PLM1)	Dnmt3a PWWP	Q9Y6K1	292-350	H3K36me2/3
	Methyl-binding domain 1	Q9UIS9	1-80	5mC
MBD1-Linker-MPP8 Chromo (M1LM)	Methyl-binding domain 1	Q9UIS9	1-80	5mC
	MPP8 Chromo	Q99549	59-118	H3K9me3
MPP8 Chromo-Linker-MBD1 (MLM1)	MPP8 Chromo	Q99549	59-118	H3K9me3
	Methyl-binding domain 1	Q9UIS9	1-80	5mC
TAF3 PHD-Linker-MBD1 (TLM1)	TAF3 PHD	Q5VWG9	865-915	H3K4me3
	Methyl-binding domain 1	Q9UIS9	1-80	5mC

Table 3: Binding pocket mutants used as control in these study, * indicated the mutated domain.

Name	Abbreviation	Binding pocket mutation	Reference
Dnmt3a PWWP*-MPP8 Chromo	P*M	D329A	(Dhayalan et al. 2010)
Dnmt3a PWWP-MPP8 Chromo*	PM*	F59A	(Mauser et al. 2017)
BPFT PHD* domain and Bromodomain	BPTF PHD*	W2765F	(Li et al. 2006)
BPFT PHD domain and Bromodomain*	BPTF Bromo*	F2887A	(Ruthenburg et al. 2011)
CBX7 Chromo-L-MBD2*	CLM2*	R188C	(Scarsdale et al. 2011)
CBX7 Chromo*-L-MBD2	C*LM2	W32A	(Gil et al. 2004)

2.2 Western blot and Far-western blot

Proteins were electrophoresed on 16 - 18% SDS gels and transferred onto a nitrocellulose membrane by a semi-dry western blotting transfer system (Trans-Blot®Turbo™ Bio-Rad, USA) as described in (Mauser et al. 2017) (see appendix 4 of this work).

For Far-western blots of histone PTMs, 2.5 µg of native histones (NH) isolated by acid extraction (Shechter et al. 2007) from HepG2 cells or HEK293 cells and 1 µg of recombinant histone H3 (RH) (New England Biolabs) were separated on a SDS gel. Afterwards, the membrane was incubated overnight in blocking solution. The next day, the membrane was incubated with Interaction buffer (IB) containing 10 nM - 1 µM protein, followed by an incubation with anti-GST antibody and afterwards with anti-goat antibody conjugated with horseradish peroxidase. Finally, the membrane was immersed in ECL solution (Thermo Fisher Scientific) and chemiluminescence was detected on X-ray film or Fusion-SL3500 (Peqlab).

2.3 Dot blot assay

The DNA binding specificity of the MBD2 and MBD1 domains alone and in the fusion constructs was analyzed using dot blot assays. The unmethylated DNA used in this assay was produced by amplification of a CGI of the ZNF280B gene located in the human chromosome 22. The PCR product has a size of 349 bp and contains 44 CpG sites. The methylated DNA was generated by *in vitro* methylation of the unmethylated PCR product using the bacterial M.SssI CpG methyltransferase. The methylated and unmethylated DNA was kindly provided by Dr. Pavel Bashtrykov.

For the binding studies of the DNA methyl-binding domains to DNA, 2 µl of methylated and unmethylated DNA in different dilutions containing 200 ng to 2 pg of DNA were spotted manually on a Amersham Hybond-N+ membrane (GE Healthcare). After air-drying the spots, the membrane was blocked in TTBS with 5 % milk and washed three times with TTBS. Afterwards, the membrane was incubated in IB (100 mM KCl, 20 mM HEPES pH 7.5, 1 mM EDTA, 0.1 mM DTT and 10% glycerol) containing 0.1 µM protein. Subsequently, the membrane was washed three times with TTBS and incubated for 1 h with anti-GST antibody.

Another three washing steps were performed followed by an incubation with anti-goat antibody fused to horseradish peroxidase for 1 h and chemiluminescence was detected on

Fusion-SL3500 (Peqlab Biotechnologie GmbH, Erlangen, Germany) using the SuperSignal West Femto Maximum Sensitivity Substrate kit (Thermo Fisher Scientific).

2.4 Peptide arrays

CelluSpots peptide arrays (Active Motif, Carlsbad, USA) contain 384 peptides from 8 different regions of the N-terminal histone tails (H3 1-19, 7-26, 16-35 and 26-45, H4 1-19 and 11-30, H2A 1-19 and H2B 1-19) with 59 post-translational modifications. The protocol used for protein binding to peptide arrays is described in (Bock et al. 2011a).

The arrays were blocked overnight in TTBS with 5% skim milk and washed the next day two times with 3.5 ml TTBS. Afterwards the CelluSpots arrays were washed once with 3.5 ml IB (100 mM KCl, 20 mM HEPES pH 7.5, 1 mM EDTA, 0.1 mM DTT and 10% glycerol) and incubated with 0.01 - 0.1 µM of GST-tagged or MBP-tagged protein for 2 hours in IB.
Subsequent the CelluSpots peptide arrays were washed three times with 3.5 ml TTBS and incubated with primary anti-GST antibody for 1 h at RT, washed again three times with TTBS and incubated with the secondary anti-goat-HRP conjugated antibody.
The signal was detected by chemiluminescence with Fusion-SL3500 (Peqlab). The CelluSpots peptide arrays were analyzed with the Array-Analyze software which is available at (http://www.activemotif.com/catalog/667.html) and the full annotation of all spots is provided at (https://www.activemotif.com/catalog/668/modified-histone-peptide-array).

2.5 Mixed peptide arrays

The peptides were synthesized as described in (Maisch et al. 2011; Dikmans et al. 2006) via the SPOT method using Fmoc-based chemistry on a MultiPep RSi synthesizer (Intavis AG). Equal amounts of the different peptides (table 4) were mixed and 50 nl where spotted on a glass slide. The procedure was similar to the protocol of the CelluSpots arrays and is described in detail in (Mauser et al. 2018). See appendix 3 of this work.

Table 4: Peptides used in the mixed peptide arrays.

Peptide name (including length of the peptide)	Sequence (with or without modifications)
H3 1-19	A R T K Q T A R K S T G G K A P R K Q
H3 1-19 K9me3	A R T K Q T A R Kme3 S T G G K A P R K Q
H3 1-19 K4me3	A R T Kme3 Q T A R K S T G G K A P R K Q
H3 26-45	R K S A P A T G G V K K P H R Y R P G
H3 26-45 K36me3	R K S A P A T G G V Kme3 K P H R Y R P G
H4 11-30	G K G G A K R H R K V L R D N I Q G I T
H4 11-30 K16ac	G K G G A Kac R H R K V L R D N I Q G I T
H3 16-35	P R K Q L A T K A A R K S A P A T G G
H3 16-35 K27me3	P R K Q L A T K A A R Kme3 S A P A T G G

2.6 H3-GST tagged methyl-lysine analogues and pull downs

Methyl-lysine analogues where introduced in a histone H3 N-terminal tail - GST fusion protein as described in (Simon et al. 2007).

For that, the sequence encoding for the first 60-aa of human histone H3 was cloned with C-terminal GST tag via Gibson assembly into the pGEX-6p-2 vector. All four cysteines in the GST tag were replaced by serine. Then, the target lysine residues were replaced with cysteine and alkylated with (2-bromoethyl) trimethylammonium bromide to the respective trimethyl analogue. For more details, refer to the respective paragraph in (Mauser et al. 2017) appendix 4.2 of this work. To verify the efficiency of conversion MALDI-TOF (matrix-assisted laser desorption ionization-time-of-flight) mass spectrometry was used. For detailed information about MALDI-TOF mass spectrometry see paragraph 2.7.

For the pull down, 25 µg of H3(1-60)-GST containing methyl lysine analogues were incubated with 0.5 µM of MBP-tagged binding domains. The next day pre-washed glutathione sepharose

4B beads (GE Healthcare) where added to the complex. The complex was washed with CIDOP buffers or PB200 for the D3PWWP-M8Chromo*.

Elution was done with LAP and the co-precipitated MBD-tagged binding domains were electrophoresed on an 18% SDS gel. Afterwards, they were transferred on a nitrocellulose membrane and probed with anti-MBP antibody followed by secondary anti-mouse horseradish peroxidase conjugated antibody and the western blot was developed as described above.

2.7 MALDI-TOF mass spectrometry

To verify the efficiency of the H3 analogues conversion, MALDI-TOF (matrix-assisted laser desorption ionization-time-of-flight) mass spectrometry was used. The protein samples (~0.5 mg/ml) were diluted 1:5 in 0.1% TFA (trifluoroacetic acid) and afterwards 2 µl of the protein dilution were mixed with 2 µl 2% TFA and 2 µl 2,5-DHAP matrix (2,5-Dihydroxyacetophenone) until crystallisation was observed. Then, 2 µl of the mix were spotted on a MTP 384 ground steel plate and the droplets were dried at RT. Additionally 2 µl of Protein calibration standard II (Bruker part number 207234) were mixed with 2 µl of 2,5-DHAP matrix until crystallisation was observed and were spotted in the middle of every four sample spots.

The samples were analyzed in the FlexControl program (Bruker) with the acquisition method LP_30-210 kDa. The Laser power was set to 75 % and first the Protein calibration standard II was measured. After a sufficient quality spectrum was obtained, the calibration was applied using AutoCalibration method in the FlexControl-Calibration tab. After successful calibration, approximately 1000 shots were collected for every data point with the same laser power as used for the calibrant. The data was transferred to FlexAnalysis and analyzed. The expected mass shift by one alkylation reaction is + 86 Da and for two alkylation reactions + 172 Da. The calculated theoretical masses are: H3K9C 34417 Da, H3K36C 34417 Da, H3K9C/K36C 34392 Da, H3K9Cme3 34503 Da, H3K36Cme3 34503 Da and H3K9Cme3K36Cme3 34564 Da.

2.8 Cell culture

HepG2 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 20 mM L-glutamine and 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin at 37°C and 5% CO2 until they reached 75% confluence.

2.9 Preparation of nucleosomes

HepG2 cells were grown until they reached 70% confluency. The culture medium of the HepG2 cells used in the BPFT studies was supplemented with 0.1 µM Trichostatin A (TSA) two days before harvesting and 0.1 µM TSA was also added to all isolation buffers.

Mononucleosomes were obtained as described in (Kasinathan et al. 2014). Therefore, around 20 million cells were harvested and the pellet was resuspended in 5 ml TM2 buffer (10mM Tris pH 7.4, 2 mM MgCl2, 0.5 mM PMSF supplemented with EDTA free PIC tablets (Roche)). To lyse the cells the TM2 buffer was supplemented with 0.6 % of NP-40. Lysis took place for 5 min on ice afterwards the nuclei was obtained by centrifugation for 10 min at 1000 rcf. The nuclei were treated with 1mM $CaCl_2$ and to obtain mononucleosomes 1.5 µl Mnase per 100 ng DNA were added (New England Biolabs M0247S) for 5-10 min at 37°C. With 2 mM EGTA the reaction was stopped and the sample was centrifuged at 13 000 rcf for 10 minutes. The supernatant containing the mononucleosomes was collected.

For quality control of the isolated nucleosomes, 500 ng of the DNA and 5 µg of proteins were loaded on an Agarose or 18% SDS gel respectively (Fig. 15).
Additionally, the isolated DNA was analyzed using the Bioanalyzer (LabChip GX II Touch HT, Perkin Elmer). A typical gel image obtained from LabChip analyzing the isolated DNA from nucleosomes is shown in Fig. 15C and the corresponding electropherogram in Fig. 15D.

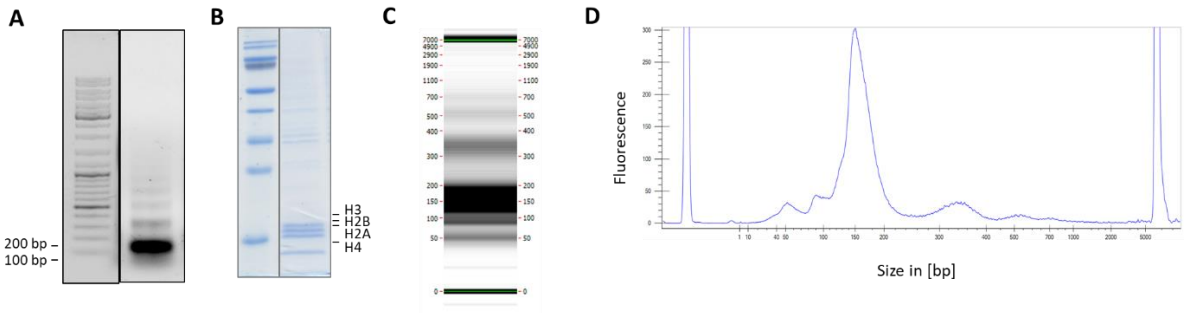


Figure 15. Preparation and control of mononucleosomes. **A)** Typical result of an agarose gel with isolated DNA from nucleosomes, size of 150 bp represents Mononucleosomes, 300 bp Di-nucleosomes and so on. **B)** Typical result of an 18% SDS gel with histone proteins H3 (17 kDa), H2B (16 kDa), H2A (16 kDa) and H4 (11 kDa).**C)** Representation of a gel image generated using LabChip GX II Touch HT (Perkin Elmer) representing the typical result of DNA isolated from nucleosome. Size of 150 bp represents Mononucleosomes, 300 bp Di-nucleosomes and so on. Green lines representing system peaks. **D)** Electropherogram generated using LabChip GX II Touch HT (Perkin Elmer) representing the typical pattern of isolated DNA from nucleosomes. Size of 150 bp represents Mononucleosomes and 300 bp Di-nucleosomes. Peak at the beginning (below 1 bp) and peak at the end (500 bp) representing system peaks.

2.10 Chromatin Interacting Domain Precipitation (CIDOP)

CIDOP protocol was performed like published in (Kungulovski et al. 2014, 2016; Mauser et al. 2017).

To pre-clear the chromatin, 30 µg mononucleosomes isolated from HepG2 cells were added to 20 µl pre-washed beads filled up with DP buffer (16.7 mM Tris-Cl, 167 mM NaCl, 1.1% Triton X-100, 1.2 mM EDTA and protease inhibitor) to 500µl and incubated for 1 h at 4°C with rotation. The pre-clearing step is important to remove unspecific chromatin binding to the beads.

The supernatant was transferred into a new tube and incubated with 0.1 - 0.5 µM of GST-tagged HiMID or antibody (concentration was used according to suppliers protocol for ChIP) overnight. The next day, 40µl pre-washed glutathione sepharose 4B beads or dynabeads Protein G (Invitrogen) were added to the supernatant and incubated for 2 hours. Followed by five washing steps in which the salt conditions were used depending on to the binding strength of the reading domains and elution was done by adding 100 µl elution buffer supplemented with 3 µl proteinase K (NEB # P8107S) for 45 min at room temperature.

Afterwards, the tubes were incubated at +55°C for 60 minutes to improve the activity of the enzyme. The DNA was recovered using the Chromatin IP DNA Purification Kit from Active Motif.

The washing steps for all domains, except Dnmt3a PWWP single domain, PWWP-MPP8* and PWWP-L-PWWP, was carried out under stringent conditions using CIDOP buffers (1x with 1 ml low salt buffer [20mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.1% SDS and 2 mM EDTA], 1x with 1 ml high salt buffer [20 mM Tris-Cl pH 8.0, 500 mM NaCl, 1% Triton X-100, 0.1% SDS and 2 mM EDTA], 1x with 1 ml LiCl buffer [10 mM Tris-Cl pH8.0, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate and 1 mM EDTA] and 2x with 1 ml TE buffer [10 mM Tris-HCl pH 8.0, 1mM EDTA]).

The washing steps for Dnmt3a PWWP single domain, PWWP-MPP8* and PWWP-L-PWWP was carried out under less stringent conditions (3x with 1 ml PB 200 [50 mM Tris-HCl, 200mM NaCl, 1mM EDTA, 0,5 % NP-40, 2 mM DTT] and 2x with 1 ml TE buffer [10 mM Tris-HCl pH 8.0, 1mM EDTA]).

2.11 DNA analysis

2.11.1 qPCR analysis

Analysis of the recovered DNA was done with quantitative PCR (qPCR) assay as described in (Mauser et al. 2017). In brief, SsoFast EvaGreen supermix (Bio-Rad) was used and the fluorescence was detect with a CFX96 Real-Time machine (Bio-Rad). Each sample was analyzed in three technical repeats and a standard curve was generated to calculate the percent of precipitated DNA. Two to three biological repeats were performed for each conditions. The error bars were calculated based on the biological repeats. The primers used for qPCR in this study can be found in table 5. Primers used for qPCR studies were designed based on the UCSC genome browser views using the respective data of the ENCODE database (Consortium 2012).

Table 5: Sequence and genomic localisation of the amplicons used for the qPCR experiments.

Histone modification	Name	Genomic location of amplicon	Amplicon size	Amplicon Sequence 5´-3´
Primers used in (Kungulovski et al. 2016)				
H3K4me3	VEGF-A 5´-3´1	Chr 6, Position 43737363-43737490	128 bp	GCAGTCACTAGGGGGCGCTC CAACGCCCTCAACCCCACA
H3K4me3	VEGF-A 5´-3´2	Chr 6, Position 43738352-43738469	118 bp	ACAGGGGCAAAGTGAGTGAC GCGGTGTCTGTCTGTCTGTC
H3K4me3	VEGF-A 5´-3´3	Chr 6, Position 43739179-43739301	123 bp	CCCCCTCTGTCGTCTTAGGT AGATCGTACGTGCGGTGACT
H3K4me3	VEGF-A 5´-3´4	Chr 6, Position 43752242-43752382	141 bp	GCCCTAACCCCAGCCTTTGTTT GTATCGATCGTTCTGTATCAGTCTTTCC
H3K4me3	PABPC1 1	Chr 8 Position : 101784959-101785065	107 bp	ATCACTCCCAAGAAATTGAGTAGT CACAGTCCTCAAAGCCCAAA
H3K4me3	PABPC1 2	Chr 8 Position : 101793210-101793304	95 bp	GAAGTTCTCTTCTAGATGTTTTTGTCTAG CCCAACCCCAGCAAAAT
H3K4me3	PABPC1 3	Chr 8 , Position: 10180125-101801383	126 bp	TGACACAGAGCCAGAAGTTGTAAAA CATGCATTGATACATCCTGCCTAA
H3K9me3	1	Chr 10, Position: 42384283-42384400	171 bp	GACCATTGGATGATTGCAGTCA FATCGAATGGAAATGAAAGGAGTCA
H3K9me3	2	Chr 5,Position: 2252820-22529032	113 bp	AGAACACCATGGACCACCAG TTTCTGAATTGGTTCTGGGTTT
H3K36me3	1	Chr 5, Position: 150280046-150280158	113 bp	CTGCTCCCATGTCTGCTACA TGGAAGGACTGCAGAGAAAAA
H3K36me3	2	Chr 9, Position: 100402514-100402612	98 bp	TGCTCCTTTTTCCCATCTTTT GCAAAACCAAGTCGAATGCT
Primer used in (Mauser et al. 2017)				
H3K9me3 and H3K36me3	Region 9	Chr 4, Position: 455037-455147	111 bp	CTAGAGCCCAGGCTGTTTTG AAGCAGATTGCAGTGGGAAG
H3K9me3 and H3K36me3	Region 10	Chr 19, Position: 36449278-36449371	94 bp	GCATTGCCACATTCCTTACA GCGGGAAGGCCTTTAGTAGT
H3K9me3 and H3K36me3	Region 11	Chr 19, Position: 37739671-37739765	95 bp	TGGTTTTTCGCCAGTATGAA GTGGGAAGCCGTATGAATGT
H3K9me3 and H3K36me3	Region 12	Chr 19, Position: 40015195-40015290	96bp	GTGGGAAGGCCTTTATTCGT ACGTCTTCCCACATTCCGTA
H3K36me3	Region 5	Chr 12, Position: 6645729-6645847	119 bp	CAATGACCCCTTCATTGACC GGGGGAATACGTGAGGGTAT
H3K36me3	Region 6	Chr 9, Position: 100402514-100402612	99 bp	TGCTCCTTTTTCCCATCTTTT GCAAAACCAAGTCGAATGCT
H3K36me3	Region 7	Chr 19, Position: 38660862-38660965	104bp	AGTGATTGGGAATGGCTTTG AACAGCTGAAATGCCACCTC
H3K36me3	Region 8	Chr 19, Position: 41272007-41272112	106 bp	GGAAAAGGGACCAGAAAAG GGTGCTCCTTGCTTCATCTC
H3K9me3	Region 1	Chr 5, Position: 22891711-22891829	119 bp	TGCATGATGTTTTCCTCAGC ATCTTGCGCAAATGCTCTG

Histone modification	Name	Genomic location of amplicon	Amplicon size	Amplicon Sequence 5´-3´
H3K9me3	Region 2	Chr 19, Position: 37625187-37625309	123 bp	TTGTCACCACTGTCCAGGAA CAGCTGCCTCAGAGACACAC
H3K9me3	Region 3	Chr 5, Position: 22528920-22529032	113 bp	AGAACACCATGGACCACCAG TTTCTGAATTGGTTCTGGGTTT
H3K9me3	Region 4	Chr 19, Position: 40954465-40954565	101 bp	TTCACAGAAACTGCCACTGC GCAGATGAGAAGGCAAGGAA
ZFN274 binding sites with H3K9me3-H3K36me3 signals				
H3K9me3-H3K36me3	Region 20	Chr 19, Position: 52554314-52554418	105 bp	TAATTCACAGCTGGCACGAC TGGCATAGAAGGGATGACCT
H3K9me3-H3K36me3	Region 21	Chr 19, Position: 52840919-52841019	101 bp	TCTGATGTTGTGCCAGGTGT CCTGGAAAGACACAGGAGGA
H3K9me3-H3K36me3	Region 22	Chr 19, Position: 52949774-52949856	83 bp	TGATTTGCAATGGTTGTAGCA AAGCAATCCATGGGTGTAGG
H3K9me3-H3K36me3	Region 24	Chr 19, Position: 53410202-53410301	100 bp	TCAAAAAGCAAAGCTTGCAC GAGGCTTGATTTGCGACTGT
H3K9me3-H3K36me3	Region 25	Chr 19, Position: 53456346-53456442	97 bp	GACCTTCAGTCGGACGTCAT CACACGGAAAGCTTTGTCAC
Control region	Region 23	Chr 19, Position: 53305130-53305236	107 bp	TTCCTTCCTGCATGAGATCC GAGTGCTTCCCCCAGATTTT
Control region	Region 23	Chr 19, Position: 53879224-53879313	90 bp	CTATCCTTCGCCCTAGGTCAC GGGAATCCGAGTCCTCTTGT
Primers used for CLT and TLC				
H3K27me3	HOX11	Chr 7, Position: 2718943-27189325	83 bp	TTTCCACAGCCTTTGCAGGCG TTCAGGCTGCAAGAAGAAGCGGAG
H3K27me3	WNT2	Chr 7, Position: 117322057-117322121	65 bp	GGTGCAAGGAAATTACAGGGC TCCATCTGCCGACTTTCTGG
H3K4me3	1	Chr 7, Position: 103848557-103848665	109 bp	ATCCTCTTCTCGCTGGTCCT TCCATGTCGTCTCCTTAGCC
H3K4me3	2	Chr 8, Position: 103424387-10342449	104 bp	GCCAGCCCTATTACCTGTCA TCGAGTGGAGGACGAGAAG
H3K4me3 and H3K27me3	1	Chr 14, Position: 24836100-24836206	107 bp	CCGGCTTCTGAGTGTTTGTA CTAAATCCCTCCGGTTCAGG
H3K4me3 and H3K27me3	2	Chr 17, Position: 60704674-60704767	94 bp	CGGAAAGTTTTGCAGTTGCT GCCTTCCCTCTCTTCTGTCC
Primers used for MLM and MPP8 Chromo domain				
H3K9me3	M2	Chr 5 , Position : 22891602-22891720	119 bp	TGCATGATGTTTTCCTCAGC ATCTTGCGCAAATGCTCTG
H3K4me3 (low mDNA)	VEGF-A promoter	Chr 6, Position: 43770434-43770635	202 bp	GCTTGCCATTCCCCACTTGAATCG GGTCACTCACTTTGCCCCTGTC
H3K9me3 and mDNA	VEGF-A GB4	Chr 6, Position: 43770615-43770732	118 bp	ACAGGGGCAAAGTGAGTGAC GCGGTGTCTGTCTGTCTGT
H3K9me3, mDNA	Satα	Chr 10, Position: 42384283-42384400	171 bp	GACCATTGGATGATTGCAGTCA ATCGAATGGAAATGAAAGGAGTCA

Histone modification	Name	Genomic location of amplicon	Amplicon size	Amplicon Sequence 5´-3´
H3K9me3	Gene desert 12	Chr12 , Position: 61273939-61274046	108 bp	GCTGTTACTTTTTACAGGGAGTTTTTA ATAAAGCAGGTAAAGGTCCATATTTC
H3K27me3	HOX11	Chr 7, Position: 2718943-27189325	83 bp	TTT CCA CAG CCT TTG CAG GCG TTC AGG CTG CAA GAA GAA GCG GAG
H3K27me3	WNT2	Chr 7, Position: 117322057-117322121	65 bp	GGT GCA AGG AAA TTA CAG GGC TCC ATC TGC CGA CTT TCT GG
H3K36me3	PABCP1 gene body	Chr 8 , Position: 100719854-100719979	126 bp	TGACACAGAGCCAGAAGTTGTAAAA CATGCATTGATACATCCTGCCTAA
Primers used for double domains with dual specificity for methylated DNA and histone PTMs				
meDNA and H3K27me3	1	Chr 1 , Position: 166958559-166958658	100 bp	GTCTGAGGCCAGGAAGTTTG CCGTAGTTCGGGGATCTTCT
meDNA and H3K27me3	2	Chr 2 , Position: 238707647-238707764	118 bp	ACTCCGAACGGCGGTGTA AGAACGACTTTTGCGGTGTC
meDNA and H3K27me3	3	Chr 7, Position: 27135986-27136065	80 bp	ATGAAAGATGAACTGGCGAGA GAGCTCTGGCCCACTGATT
unmDNA and H3K27me3	1	Chr 7, Position: 27136527-27136614	88 bp	CCCCTTTCTCTTTCCCACTC CGCCTCCTACCCCTAAAAAT
unmDNA and H3K27me3	2	Chr 9, Position: 129455650-129455748	99 bp	FP ACTGCCTTTCTGGAGACCAC RP TGGTGGGTTTAGGGGTATGA
unmDNA and H3K27me3	3	Chr 17, Position: 41909883-41909967	85 bp	AGTCGGGCTTCCACTATGC CTTTTTGAGGTCCACGTGCT
meDNA, no H3K27me3	1	Chr 11 Position: 63530599-62530699	101 bp	GAGGATTCTCGGAAAACGTG TTGGGTCCCTGAGCCAGT
meDNA, no H3K27me3	2	Chr 12, Position: 130646580-130646676	97 bp	CCGAACTTCCCGTAACCTC TTGCCTCTCGCTATCCTCTC
meDNA, no H3K27me3	3	Chr 1, Position: 29448729-29448881	153 bp	CGCGTCTCCAGGGTAAGTA GCTGCGGTATGTAAAGCACA
ZNF420 (H3K9me3, low mDNA)		Chr 19, Position: 37134285-37134406	122 bp	TTGTCACCACTGTCCAGGAA AGCTGCCTCAGAGACACAC
FZD10 (mDNA)		Chr 12 , Position 130162035-130162131	97 bp	CCGAACTTCCCGTAACCTC TTGCCTCTCGCTATCCTCTC

2.11.2 CIDOP-seq and ChIP-seq analysis

The precipitated DNA was sent to the Max Planck-Genome-Center (Cologne) and analyzed with the BioAnalyzer (Agilent Technologies, (Santa Clara, USA)). Size selection of DNA fragments around 150 bp via beads was performed by the sequencing facility. The obtained fragments were used for library preparation and approximately 35-90 millions (150 bp read-length) single end reads were generated using Illumina HiSeq 3000 genome sequencer at the Max Planck-Genome-Centre (Cologne).

First, the quality of the raw sequences was controlled using the FastQC program of the Galaxy server (https://usegalaxy.org).

After successful quality control, the reads were mapped to the human reference genome Hg38 using the Chipster or the Galaxy platform (Kallio et al. 2011; Goecks et al. 2010). The data was mapped using Bowtie to obtain only uniquely mapped reads (Langmead et al. 2009).

To obtain statistically significant enriched genomic regions, peak calling was done with Galaxy using the tool MACS (NGS Peak Calling) applying the broad options (Zhang et al. 2008). Peak intersection was performed with BEDtools (Quinlan and Hall 2010). Peak overlap was done with the tool Intersect (Operate on Genomic Intervals) on the Galaxy server. EpiExplorer was used to determine the distribution of peaks, genes per chromosome and overlap with chromatin segments (Halachev et al. 2012).

To create density profiles the BAM files were converted to Bigwig files with DeepTools and the reads were normalized to reads per kilobase per million mapped reads (RPKM) (Ramirez et al. 2014). Visualization of the density profiles was performed with the Integrative Genomics Viewer (IGV) (Robinson et al. 2011).

For the analysis of H3K9me3 and H3K36me2/3 single and double readers, different chromatin states were defined: no H3K9me3 and no H3K36me2/3, H3K9me3-only, H3K36me3-only and overlap of H3K36me3 and H3K9me3. For that the genome was divided in 1 kb or 3 kb bins and the number of normalized (to the highest dataset) reads per million (RPM) was quantified using the SeqMonk software (http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/). The analysis of the overlap of the 3 kb bins was carried out in Excel. To obtain the signal for the 1 kb bin regions, the Filter tool in Galaxy was used and the overlap was identified with the Intersect Interval tool.

To generate the k-means clusters and heatmaps, seqMINER was used (Ye et al. 2011). BAM files and files containing the reference genomic coordinates, like RNA pol II or H3K4me1, were used to create density profile within a defined window. Like this, clusters of overlapping chromatin states can be easily visualized and different datasets can be compared.

Furthermore the BAM file was plotted over defined genomic elements to perform Metagene analysis (DeepTools computeMatrix) (Ramirez et al. 2014). The output file was plotted with the tool profiler. In addition, Spearman correlation of raw data in 10 kb bins were generated in DeepTools.

ChIP-Enrich was used to perform GO analysis of clusters obtained by k-means clustering from before (Welch et al. 2014). Only the first 10-15 categories termed “biological process” were selected for the analysis. For comparison ENCODE (Consortium 2012) data of H3K4me1, ZNF274, SetDB1 and KAP1 (Trim28) were downloaded and also mapped to Hg38 like it was described before. The exact data source are specified in the results section of this work or in the corresponding publications. The ZNF274, SetDB1 and Trim28 peaks were directly downloaded from ENCODE and liftOvered to Hg38 (Kent et al. 2002).

2.11.3 RNA-seq analysis

RNA analysis was done like described in (Mauser et al. 2017). All RNA data in this work was taken from ENCODE (Consortium 2012) produced by Caltech:

wgEncodeCaltechRnaSeqHepg2R1x75dFastqRep1.fastq.gz

wgEncodeCaltechRnaSeqHepg2R1x75dFastqRep2.fastq.gz

Briefly, the single reads were mapped to the Hg38 with TopHat from the Tuxedo Suite package applying the default settings (Trapnell et al. 2012). The transcript of the two replicates was assembled with SeqMonk and the obtained transcript lists from both replicates were merged. Based on their FPKM (fragments per kilobase of exon per million fragments mapped) value the data was sorted and separated in four groups based on their frequency distribution: no expression, low expression-1, low expression-2, medium expression, high expression.

2.11.4 Data access

The data for the TAF3 PHD domain published in (Kungulovski et al. 2016) have been submitted to the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/). Accession number E-MTAB-4103.

The raw input and H3K9me3 CIDOP-seq data were taken from (Kungulovski et al. 2014) and can be found under the accession number E-MTAB-2143.

The data for the PM double domain published in (Mauser et al. 2017) have been submitted to the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/). Accession number E-MTAB-4216.

The raw input, MPP8 Chromo and DNMT3A PWWP data were taken from (Kungulovski et al. 2014) and can be found under the accession number E-MTAB-2143.

2.12 Sequential CIDOP/ChIP

A schematic representation of this method is provided in Fig. 16 and for detail description see appendix 4 of this work.

To test the co-existence of two marks on the same histone tail two consecutive affinity pull down steps were carried out. First, the chromatin was incubated with the GST-tagged MPP8 Chromo domain and elution was done with 40 mM reduced L-Glutathione (Applichem) for 2h at 4°C with rotation. 20 µl of the eluted material were used for later PCR analysis and the rest of the eluted material was incubated with H3K36me3 antibody overnight (concentration was used according to suppliers protocol for ChIP).

The next day, 40 µl of pre-washed Dynabeads Protein G (Invitrogen) were added and after2 h of incubation the complex was washed using CIDOP buffers and the second elution was done by adding 100 µl elution buffer supplemented with 3 µl proteinase K (NEB # P8107S) for 45 min at room temperature. Afterwards, the tubes were shifted to +55°C for 60 minutes to improve the activity of the enzyme.

The DNA was recovered using the Chromatin IP DNA Purification Kit from Active Motif. The recovered DNA of both steps was analyzed by semiquantitative PCR with different primers to detect the presence of the H3K9me3 only, H3K36me3 only or H3K9me3 and H3K36me3 marks. Primers used for the analysis of the sequential ChIP are listed in table 6.

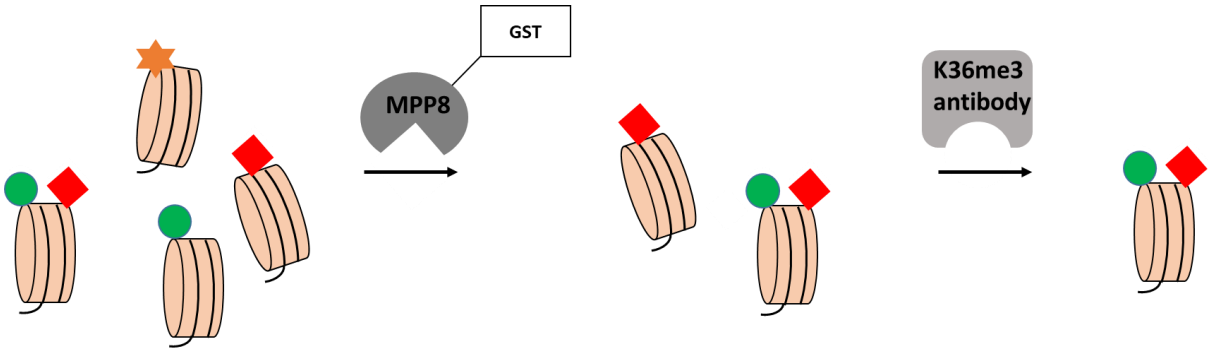


Figure 16. Scheme of the sequential CIDOP/ChIP. First precipitation of H3K9me3 (red square) containing nucleosomes with MPP8 Chromo domain, after eluting from the beads precipitation with H3K36me3 antibody was carried out to obtain only nucleosomes bearing H3K9me3 (red square) and H3K36me3 (green circle) marks. The DNA was isolated and analyzed by semi quantitative or quantitate PCR.

Table 6: Sequence and genomic localisation of the amplicons used for sequential CIDOP/ChIP experiments.

	Name	Genomic location of amplicon	Amplicon size	Amplicon Sequence 5´-3´
H3K9me3	Region 1	Chr 5, Position: 22891711-22891829	119 bp	TGCATGATGTTTTCCTCAGC ATCTTGCGCAAATGCTCTG
H3K36me3	Region 6	Chr 9, Position: 100402514-100402612	99 bp	TGCTCCTTTTTCCCATCTTTT GCAAAACCAAGTCGAATGCT
H3K9me3 and H3K36me3	Region 9	Chr 4, Position: 455037-455147	111 bp	CTAGAGCCCAGGCTGTTTTG AAGCAGATTGCAGTGGGAAG
H3K9me3 and H3K36me3	Region 10	Chr 19, Position: 36449278-36449371	94 bp	GCATTGCCACATTCCTTACA GCGGGAAGGCCTTTAGTAGT

2.13 MBD-Pull down

To compare the affinity for methylated DNA of MBD1 and MBD2 pull down experiments were carried out. Therefore, 1 µg of sonicated DNA (~300-600 bp) was incubated with 8 µg of protein in 200 µl PB buffer with 150 mM salt (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40, 2 mM DTT) overnight at 4 °C with rotation. The next day, 40 µl of pre washed Glutathione Sepharose 4B were added to the DNA-domain complex and binding to the beads was carried out for 2 h at 4 °C with rotation. Afterwards the complex was washed three times with 200 µl of PB buffer with 150 mM salt or PB with 500 mM salt (50 mM Tris-HCl, 150/500 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40, 2 mM DTT). Each time the samples were rotated for 5 min, spun down for 2 min at 2000 rcf at 4°C and the supernatant was discarded. Elution

was done for 15 min with 150 µl of PB 2 M salt at RT. The supernatant was transferred into a new tube and the elution step was repeated. Finally, the DNA was purified using the Chromatin IP DNA purification kit (Machery-Nagel).

2.14 Fluorescence Anisotropy (FA) peptide binding

The binding specificity and the binding constant of BPTF to different fluorescently labelled peptides was done like published in (Mauser et al. 2018). Fluorescence anisotropy was measured using the Jasco FP-8300 Spectrofluorometer at an excitation wavelength of 493 nm and an emission wavelength of 516.6 nm, using slits of 5 nm. For peptide binding to the BPTF wt and its variants, 2 µL aliquots of the respective binding domain (stock concentrations ranging from 15 to 30 µM) were added stepwise to the cuvette and after incubation at RT the anisotropy was measured.

2.15 Antibodies

All antibodies used in the study are listed in table 7.

Table 7: Listed antibodies used in the studies.

Antibody	Catalogue number	Lot number	Dilution
Anti-H3K9me3	Abcam 8898	GR130967-1	1: 1000
		GR3176466-3	1: 1000
Anti-H3K36me3	Abcam 9050	GR260274-1	1: 1000
		GR310541-1	1: 1000
Anti-H3K4me3	Abcam 8580	GR85670-1	1: 1000
		GR224425-1	1: 1000
		GR224576	1: 1000
Anti-H3K27me3	Active Motif 39155,	01613015	1: 1000
Anti-H4 pan-acetyl	Active Motif, 39243	05308001	1: 1000
Anti H3 antibody	Abcam 1791	GR135320-1	1: 1000
		GR103869-1	1: 1000
Anti-GST AB Healthcare	Life Science 27-4577-01	9541184	1 : 5000
Horseradish peroxidase conjugated anti-goat AB	Sigma-Aldrich A4174	071M4767	1: 10 000
Anti-rabbit	GE Healthcare NA934-1Ml	9739638	1 : 5000
Anti-mouse	GE Healthcare HRP NXA931	973269	1 : 5000
Anti-MBP antibody	New England Biolabs E8032	0091208	1 : 10 000

3 Results

The unstructured N-terminal tails of histones carry many post-translational modifications (PTMs) each of which having a distinct influence on gene expression and other chromatin processes (Strahl and Allis 2000; Bannister and Kouzarides 2011). Histone PTMs and DNA methylation represent major epigenetic signals which are essential for normal development (Greer and Shi 2012). Consequently, aberrant modifications are associated with various diseases, such as cancer or developmental defects (Portela and Esteller 2010; Suva et al. 2013; Timp and Feinberg 2013; Chi et al. 2010). Therefore, it is of high interest to study the different chromatin modifications and obtain a better insight into the biological relevance of the different chromatin modifications.

Current methods to decipher the complex histone PTM code rely on antibodies, but it was observed that histone tail antibodies have several drawbacks as described in the next chapter. In order to overcome some of these disadvantages and study the effects of these chromatin modifications in an effective manner, histone modification interacting domains (HiMIDs) were developed as an alternative to antibodies and in the present work, the portfolio of the existing HiMIDs was successfully extended for a specific H3K4me3 reading domain.

Additionally, since the first proposal of the “histone code hypothesis” (Strahl and Allis 2000) great efforts were undertaken to identify the biological relevance behind the numerous possible combinations of histone modifications. Nevertheless, it still remains an experimentally difficult task to analyze the co-occurrence of more than one histone mark at the same time and map the co-existence of two or more histone PTMs genome-wide. To this end, it was the main aim of this work to develop the method of HiMIDs further to generate an analysis tool for the investigation of the co-occurrence of two histone marks as well as co-occurrence of DNA methylation with defined histone modifications. In the following parts, it is described how the novel tools were developed and used in chromatin precipitation assays to gain new insights into the biological function of defined patterns of epigenetic modifications.

3.1 Published results

3.1.1 Implementation of the single reader TAF3 PHD domain as an alternative to H3K4me3 antibodies

Current methods to analyze histone modifications and the investigation of their specific biological roles in chromatin templated processes are all based on antibodies. However, over the past years some drawbacks of antibodies were noticed by several labs (Bock et al. 2011b; Egelhofer et al. 2011; Fuchs et al. 2011; Hattori et al. 2013; Heubach et al. 2013; Nishikori et al. 2012; Peach et al. 2012; Kungulovski et al. 2014; Kungulovski and Jeltsch 2015; Kungulovski et al. 2015b).

Many antibodies show cross-reactivity with related or unrelated marks and the interaction with histone modifications can be influenced by the presence of secondary modifications. In addition, new batches of antibodies often exhibit variability in properties, which leads to a different binding specificity and makes long time studies and reproduction of data not possible if one batch is no longer available. Furthermore, several important histone modifications occur in a very similar amino acids sequence motif like H3K9 and H3K27 (ARKS motif), which is often difficult for antibodies to discriminate. For these reasons, it is urgent to develop reliable alternatives for histone modification antibodies.

It has been shown before that histone modification interacting domains (HiMIDs) can be used as a specific and reliable alternative to histone modification antibodies (Kungulovski et al. 2014). Recombinant HiMIDs can be easily produced in *Escherichia coli* (*E.coli*) and protein engineering is simpler to achieve for histone tail interacting domains than as it is for antibodies. Furthermore, the unlimited availability of recombinant proteins with constant quality will enable long-time application and reproducibility of results (Kungulovski et al. 2014). However, so far HiMIDs are only available for the specific readout of H3K9me2/3, H3K36me2/3, H3K27me3 and H4K20me2/3 histone tail modifications.

To expand the repertoire of reading domains for CIDOP assays, an H3K4me3 affinity reagent based on the PHD domain of the TAF3 protein was developed and validated.

The H3K4me3 modification is highly conserved from mammals to yeast and it is one of the most important activating histone PTMs found mainly on the promoters of active genes

(Shilatifard 2008, 2012). Therefore, H3K4me3 is widely studied and a novel affinity reagent for the specific investigation of this modification is of great interest.

In this study, the TAF3 PHD domain was established as H3K4me3 affinity reagent. To this end, the domain was cloned and purified and its properties were compared to a validated H3K4me3 specific antibody in CelluSpots peptide arrays, Far-western blot assays and CIDOP experiments coupled to qPCR and next generation sequencing.

The results of this study are published in (Kungulovski et al. 2016), see appendix 1 of this work and will be here summarized only briefly.
Initially, the binding specificity of the TAF3 PHD domain was investigated using CelluSpots peptide arrays demonstrating its specific interaction with H3K4me3 modified peptides (Fig. 17A). Next, the ability to bind modified histones was studied in Far-western blot experiments using native and recombinant histones, showing binding of the TAF3 PHD domain and the validated H3K4me3 antibody only to native histones which contain histone PTMs including H3K4me3 (Fig. 17B).

The specific interaction of the TAF3 PHD domain and the H3K4me3 antibody with histones carrying H3K4me3 modifications was verified using nucleosomes isolated from *Saccharomyces cerevisiae* and Set1 knockout cells. In *Saccharomyces cerevisiae* the SET1 enzyme is the only enzyme introducing the H3K4me3 modification and no H3K4me3 modification is present on the histone tails of Set1 knockout cells. Pull down experiments revealed an exclusive binding of the TAF3 PHD domain and the antibody to nucleosomes isolated from wild type *Saccharomyces cerevisiae* but not from SET1 mutant further confirming the specific interaction of both reagents with H3K4me3 modifications (Fig. 17C).

Next, the TAF3 PHD domain was used in CIDOP-qPCR and CIDOP-seq experiments and compared to the results obtained with ChIP experiments with the H3K4me3 antibody. A high correlation of the signals obtained with the TAF3 PHD domain and H3K4me3 antibody in qPCR readout as well as a similar profile of the genome-wide binding of both regents were obtained (Fig. 17D, E). Furthermore, the peaks detected in the genome-wide density profiles observed with both reagents, showed a high correlation (Spearman correlation of about 0.70-0.90), again highlighting the similarity between these two datasets (Fig. 17F). Additionally, a metagene analysis of the density distributions resulted in a similar profile for the PHD domain and the H3K4me3 antibody binding (Fig. 17G).

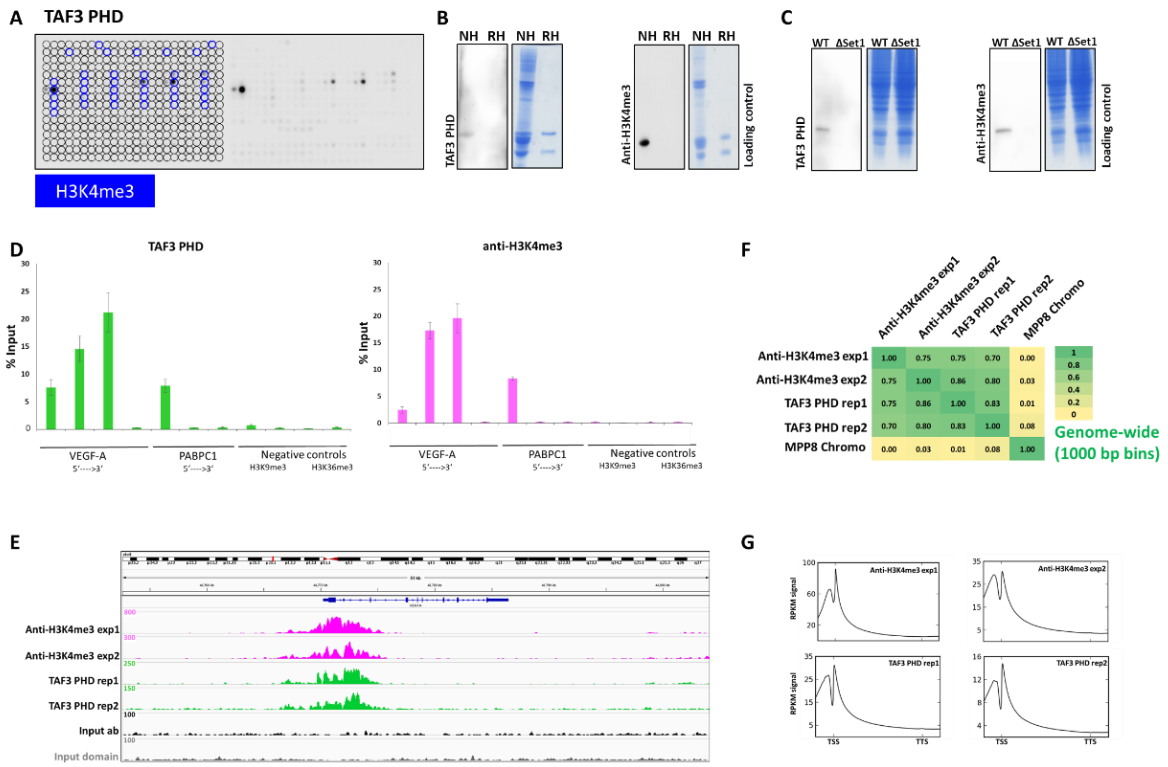


Figure 17. Summary of the main results of the investigation of the single reader TAF3 PHD. **A)** Peptide array binding of TAF3 PHD. **B)** Far-western blot analysis showing binding of the TAF3 PHD domain and H3K4me3 antibody only to modified H3 (NH) and no binding to recombinant histones (RH). **C)** Far-western blot analysis with TAF3 PHD domain and H3K4me3 antibody with lysates isolated from S. cerevisiae wild type (WT (containing H3K4me3)) and Set1 knockout (ΔSet1 (not containing H3K4me3)). **D)** CIDOP-qPCR of TAF3 PHD and ChIP-qPCR of H3K4me3 antibody using amplicons in the VEGF-A, PABPC1 region and unrelated H3K9me3 and H3K36me3 regions as negative control. **E)** Genome browser views of CIDOP-seq and ChIP-seq signals around the VEGF-A promoter region. **F)** Spearman correlation in 1-kb bins genome-wide. **G)** Profile of TAF3 PHD domain and H3K4me3 antibody distribution over metagene fitted in 5-kb gene bodies plus 1-kb flanks from the TSS and TTS. These figure panels were taken from (Kungulovski et al. 2016).

With all these experiments, it was clearly demonstrated that the TAF3 PHD domain can be used as chromatin H3K4me3 affinity reagent and it performs as good as a validated H3K4me3 antibody. Therefore, the new HiMID TAF3 PHD domain can be used as a reliable substitute of H3K4me3 antibodies in chromatin precipitation assays.

Application of the TAF3 domain in chromatin assays

The newly available HiMID TAF3 PHD domain was successfully applied in the *in situ* measurement of nucleosome stability containing H3K4me3 modifications (appendix 2 and 2.1 of this work). The principle of the quantitative image of nuclei after elution with salt/intercalators (QINESIn) assay is that agarose embedded cells are lysed and exposed to salt or DNA intercalating agents like ethidium bromide. Salt exposure affects electrostatic histone-histone and histone-DNA interactions, whereas DNA intercalators unwind and at higher

concentrations overwind the DNA, all leading to the weakening of nucleosomal binding. Afterwards, the remaining chromatin-bound histones were detected using specific histone PTM antibodies or the TAF3 PHD reading domain and analyzed by laser scanning cytometry (LSC). With this method, the association between nucleosome stability and specific PTMs across cell types, differentiation state and during the different cell-cycle phases was investigated. The method was validated in H3K4me3 ChIP-seq experiments by the estimation of chromatin loop relaxation, which is required for nucleosomal destabilization. Furthermore, salt and intercalator induced release from the nucleosomes of different histones was analyzed.

In the course of the development of the method, it was important to investigate if the presence of histone binding antibodies affect the stability of nucleosomes carrying H3K4me3 modifications, and therefore the measurements were also conducted using the recombinant H3K4me3 reading domain of TAF3. The result was compared to the measurement with H3K4me3 antibody and no differences were observed indicating that the assay is reliable (Imre et al. 2017).

3.1.2 Development of mixed peptide arrays to study the combinatorial readout of double reading domains

One central aim of this work was to develop novel single and dual affinity reagents to study histone PTMs. For this, methods to study the specificity of the interaction of antibodies and reading domains with modified histone tails were essential. So far, peptide arrays and designer nucleosomes are the methods of choice to study the interaction of reading domains and antibodies with histone tail modifications (Fierz and Muir 2012).

Designer nucleosomes are recombinant nucleosomes containing specific histone PTMs and like this, they are a powerful tool to investigate the interaction between histone modifications and reading domains (Su and Denu 2016). However, the preparation of designer nucleosomes with defined single or combinatorial modifications demands extensive preparative work and requires a high level of expertise.

Peptide arrays are used successfully for the investigation of histone modification interacting domains and antibodies, but they cannot be used to study the combinatorial readout of histone marks.

Within the scope of this work, a new and efficient method was developed to investigate the specificity of double reading domains towards combination of histone tail modifications using mixed peptide arrays. This technical development has been published in (Mauser et al. 2018), see appendix 3 and 3.1 of this work and the corresponding results will be described here only briefly.

To analyze the interaction of chromatin interacting proteins, bearing double reading domains, with two different histone modifications, unmodified and modified peptides were mixed at equal concentrations, spotted on a glass slide and incubated with the double reading domain. To verify the new approach, the well-studied natural occurring dual reading domain of BPTF was used.

BPTF is a subunit of the NURF chromatin-remodelling complex and it is involved in the regulation of the chromatin structure (Badenhorst et al. 2002). Its double reading domain (BPTF-PB) consists of a Bromo domain, which interacts with H4K16ac and with lower specificity with H4K12ac and H4K20ac, and a plant homeodomain finger (PHD), which binds H3K4me3 (Li et al. 2006; Ruthenburg et al. 2011). First, CelluSpots peptide arrays were used to analyze the specific interaction of each domain of BPTF with modified peptides (appendix 3.1, Fig. S1 and S2). The corresponding binding pocket mutants were used to validate the domain specific readout. As expected, an interaction of BPTF-PB with acetylated H4 peptides was observed confirming the interaction of the Bromo domain with its target modifications. Surprisingly, no binding to peptides carrying the H3K4me3 modifications was observed. To this end, fluorescence anisotropy peptide binding experiments were conducted, which confirmed the specific interaction of the BPTF-PB PHD domain to H3 tail peptides containing the H3K4me3 modification.

Using the newly developed mixed peptide arrays a strong and synergistic binding of BPTF-PB to peptide spots containing a mixture of H3K4me3 and H4K16ac modified peptides was observed (Fig. 18A-C). As control, binding pocket mutants were prepared and also incubated with the mixed peptide arrays, but only weak or no interaction was observed (Fig. 18D). In addition, the synergistic behaviour of histone PTM binding of both domains in BPTF-PB was

investigated in CIDOP experiments. To this end, precipitated nucleosomes were detected by western blot with an anti-H3K4me3 antibody or an anti-H4 pan-acetyl antibody. Interestingly, precipitation of nucleosomes was only observed with the intact double domain but not with variants containing binding pocket mutations in one or both reading domains (Fig. 18E). This is in agreement with previous findings showing a synergistic binding of BPTF to designer nucleosomes containing H3K4me3 and H4K16ac modifications (Li et al. 2006; Ruthenburg et al. 2011).

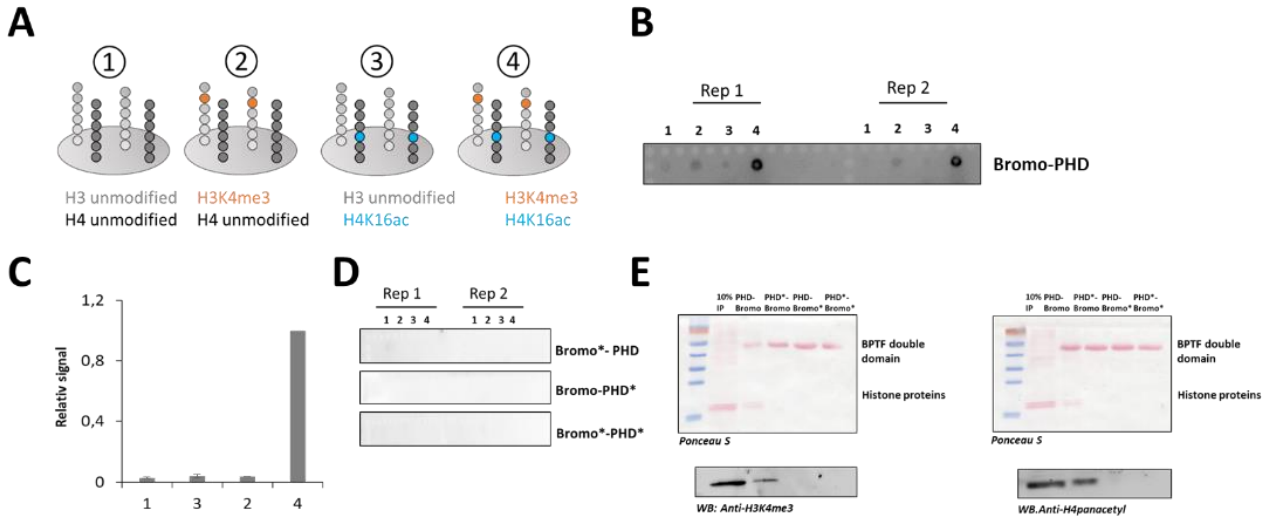


Figure 18. Development of mixed peptide arrays. **A)** Schematic view of the double peptide arrays. Equal amounts of unmodified H3 and H4 peptides (1), H3K4me3 and H4 unmodified (2), H3 unmodified and H4K16ac (3) and H3K4me3 and H4K16ac (4) were spotted on a glass slide. **B)** BPTF-PB domain binding to the double peptide array. The array contains two replicates of the four spots. **C)** Quantitative analysis of the results. Error bar indicate the SEM based on two experiments. **D)** BPTF-PB and its variants containing binding pocket mutations were incubated with the double peptide array. The array contains two replicates of the four spots. The asterisks indicate which domains contain a binding pocket mutation. In agreement with the results of panel B, no binding was observed. **E)** CIDOP of nucleosomes with BPTF-PB and the respective binding pocket mutant variants. Asterisks indicate the mutated domains. A Ponceau S staining represents the loading control. Precipitated nucleosomes were detected with an anti-H3K4me3 or an anti-H4 pan-acetyl antibody. These figure panels were taken from (Mauser et al. 2018).

This new double peptide spot method was also successfully applied to validate the double reading domain PM (see section 3.1.4) and to optimize the double reading domain for specific readout of H3K4me3 and H3K27me3 modifications (refer to section 3.2.1).

This new technology enables a fast and simple investigation of the binding specificity of double reading domains. Additionally, this method can be used as an efficient screening tool for proteins harbouring multiple reading domains. The results obtained in this study illustrate the

power of mixed peptide spot arrays for the analysis and discovery of combinatorial histone PTM readout.

3.1.3 Development of double reading domains with dual specificity towards histone tail modifications

As mentioned above, histone tail modifications play an important role in chromatin-templated processes and great efforts were undertaken to identify and map histone PTMs with the ultimate aim of understanding their biological role. In this context, the identification of chromatin states with multiple modifications and the investigation of their functions are of major interest. Therefore, the investigation of the co-occurrence of histone tail modifications is a major task in the epigenetic field, necessary to decode the complex histone code and understand the biological function of bivalent chromatin states. However, the locus specific analysis of multiple histone PTMs is an experimentally difficult task. To overcome the limitations of existing methods, the single HiMID approach was extended in this work to the application of double HiMIDs. In these studies, two validated histone interacting domains were fused together to facilitate the simultaneous readout of two histone marks on the same nucleosome. This approach was patent on the 30.03.2015 with the following reference number EP 15 161 621.6.

The single domains, used to generate the double domains, were shown before to have specific interaction with defined histone PTMs (table 8).

Readout of multiple histone tail modifications on the same tail is called *cis* readout, whereas interaction with multiple PTMs on different histone tails within the same histone is called *trans* readout (intranucleosomal). The combinatorial readout on two different histone tails on adjacent histones, either directly linked by DNA or in close spatial proximity, is called *trans* internucleosomal (Fig. 19) (Musselman et al. 2012).

In nature, these binding modes are found amongst others, in the DNA methyltransferase Dnmt3a, which interacts preferentially with the H3 tail if K36 is trimethylated and K4 unmethylated (combinatorial readout in *cis*) and in the BPTF protein, which binds to H3K4me3 and acetylated H4 tail simultaneously (combinatorial readout in *trans*).

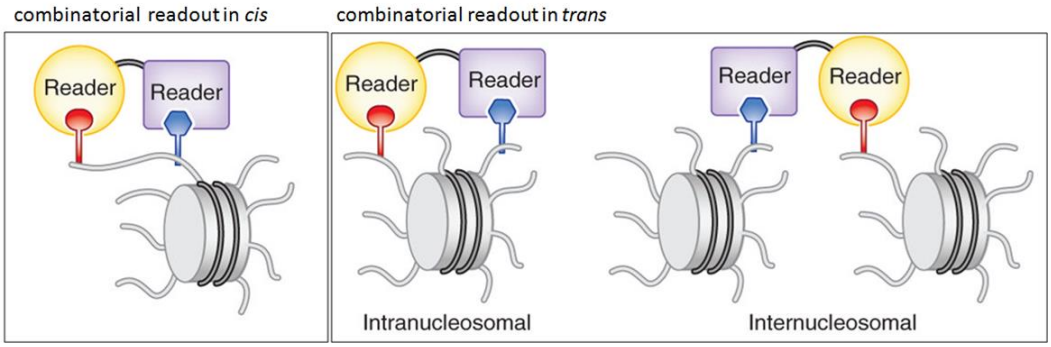


Figure 19. Combinatorial readout of two histone PTMs with double reading domains. *Cis* readout of two modifications on one histone tail or *trans* readout on different tail within one nucleosome (intranucleosomal) or one adjacent histone tails (internucleosomal).Figure is taken from (Musselman et al. 2012).

Table 8: Single domains used in this work for the generation of double reading domains.

HiMID	Modification	Reference
MPP8 Chromo	H3K9me3	(Bock et al. 2011b; Yun et al. 2011)
Dnmt3a PWWP	H3K36me2/3	(Dhayalan et al. 2010; Bock et al. 2011b)
ATRX ADD	H3K9me3 (H3K4unmodified)	(Dhayalan et al. 2011)
TAF3 PHD	H3K4me3	(Vermeulen et al. 2007; Lauberth et al. 2013)
CBX7 Chromo	H3K27me3	(Kaustov et al. 2011)
KDM4A double Tudor D969A	H4K20me3	(Kungulovski et al. 2014)

To unify antibody validation criteria, in 2012 the ENCODE project has proposed consistent quality criteria for chromatin antibodies, which include the binding to differently modified peptides, the ability to bind in western blot experiments to histone extracts and the detailed analyzes of ChIP-seq data (Consortium 2012). Therefore, all newly generated double reading domains (table 2) were validated using a workflow, which is based on these criteria as illustrated in Fig. 20.

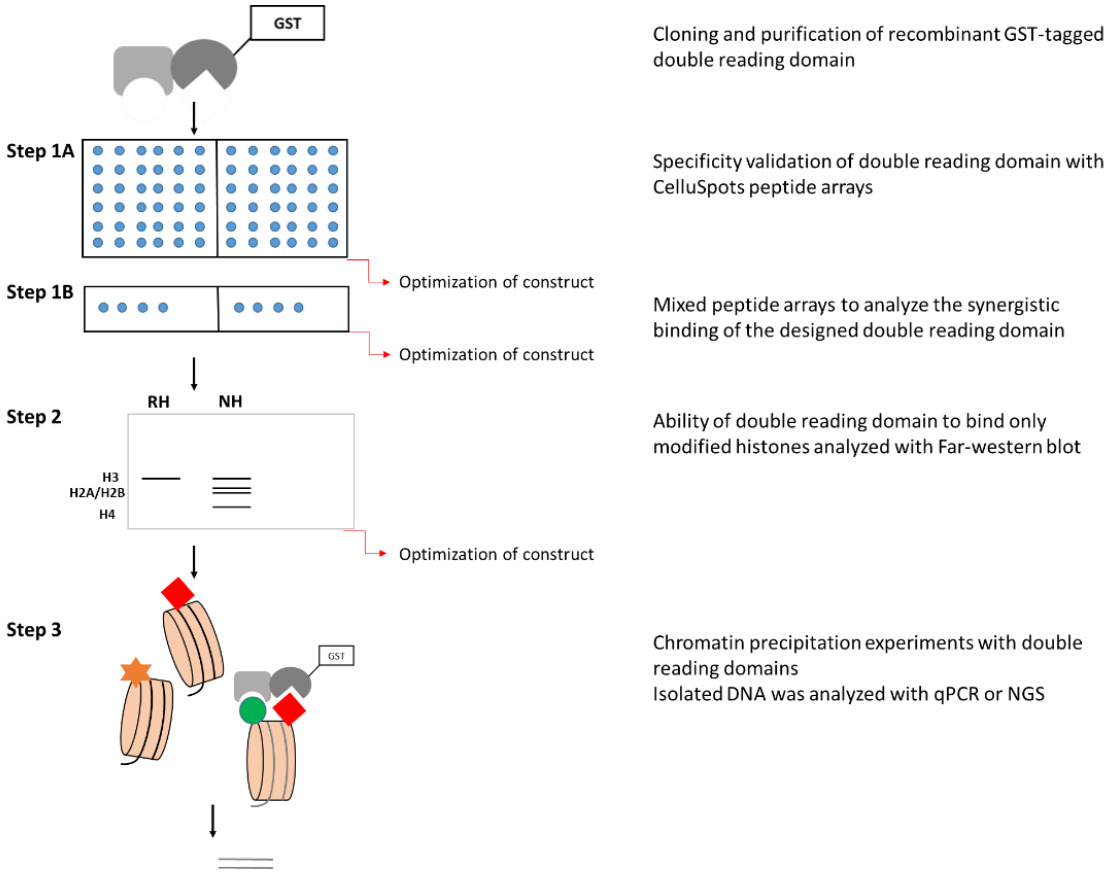


Figure 20. Workflow of specificity validation of the double reading domains. Note that after the development of double peptide arrays, they were included as additional step 1B in this workflow, but it is not available for all designed double reading domains so far.

The binding specificity of double reading domains was first investigated using CelluSpots peptide arrays. The array consists of different modified and unmodified peptides spotted on a glass slide and incubation with the double reading domains allows to determine the binding preference of the fusion proteins. If both domains in combination show their predicted interaction with PTMs, the ability of the double domain to bind to modified histones was analyzed using Far-western blot experiments. Therefore, native histones (NH) which carry PTMs and recombinant histones (RH) without post-translational modifications were separated by an SDS-gel, transferred onto a membrane and incubated with the dual reading domains.

In the case of a specific interaction with modified nucleosomes, chromatin precipitation with isolated mononucleosomes was conducted and the isolated DNA was analyzed for the enrichment of specific genome regions associated with defined histone modifications using qPCR or next generation sequencing to investigate biological functions behind the combination of histone PTMs.

3.1.4 Development of the double reading domain PM with specific binding to H3K9me3 and H3K36me2/3 modified nucleosomes

As a first example of a double reading domain, a fusion construct comprising out of the Dnmt3a PWWP (P) domain, which interacts with H3K36me2/3 marks and the MPP8 Chromo (M) domain, which binds specific to H3K9me3 was generated. This was engineered to allow the study of the co-occurrence of H3K9me3 and H3K36me2/3 modifications on one nucleosome. The co-occurrence of H3K9me3 and H3K36me3 histone tail modifications was previously documented over gene bodies and observed to be associated with alternative splicing (Kolasinska-Zwierz et al. 2009; Saint-André et al. 2011). Furthermore, H3K9me3 and H3K36me3 were found to co-occur frequently in mass spectrometric studies (Young et al. 2009).

The double reading domain PM was validated and used in chromatin precipitation experiments. The obtained data from this study are published in (Mauser et al. 2017), appendix 4 of this work and will be summarized here only briefly.
The PM double domain showed specific binding to H3K36me2/me3 and H3K9me2/me3 in CelluSpots peptide arrays (Fig. 21A). Additionally, binding to H3K27me3 was observed which was already recognized during the validation of the single MPP8 Chromo domain. However, it was shown before that binding of MPP8 Chromo domain to H3K27me3 is a peptide array artefact not appearing in the nucleosomal context (Kungulovski et al. 2014). In addition, the ability of PM to bind exclusive to native histones was shown in Far-western blot assays (Fig. 21B).

To test if the PM double domain specifically precipitates mononucleosomes carrying H3K9me3 and H3K36me2/3, mononucleosomal pull downs were carried out and analyzed by western blot with H3K9me3 and H3K36me3 antibodies. Furthermore, binding pocket mutants where used to confirm the domain specific interaction. Using all these approaches, the specific precipitation of mononucleosomes carrying H3K36me2/3 and H3K9me3 modifications by the PM double reading domain was clearly demonstrated (Fig. 21C).

With the newly developed mixed peptide arrays (see chapter 3.1.2) the new affinity reagent PM was investigated and a stronger interaction of the double reading domain with spots bearing both modifications was observed (Fig. 21D and E.). To study the efficiency of the PM

domain in the genome-wide analysis of PTM patterns, CIDOP-qPCR experiments were conducted and twelve genomic regions (associated with H3K9me3 only, H3K36me2/3 only, H3K9me3 and H3K36me2/3) were analyzed with PM and its binding pocket mutants P*M, PM* (asterisk is indicating the mutated domain) as well as with H3K9me3 and H3K36me3 antibodies.

Clear preferences for binding to doubly modified regions were observed for the PM double domain. In contrast, the binding pocket mutants as well as the antibodies showed binding to regions bearing only one modification irrespective of the presence of the second modification (Fig. 22A).

Furthermore, with sequential CIDOP-ChIP experiments the ability of the double reading domain to precipitate mononucleosomes carrying both histone modifications was confirmed (appendix 4, Fig. 5B).

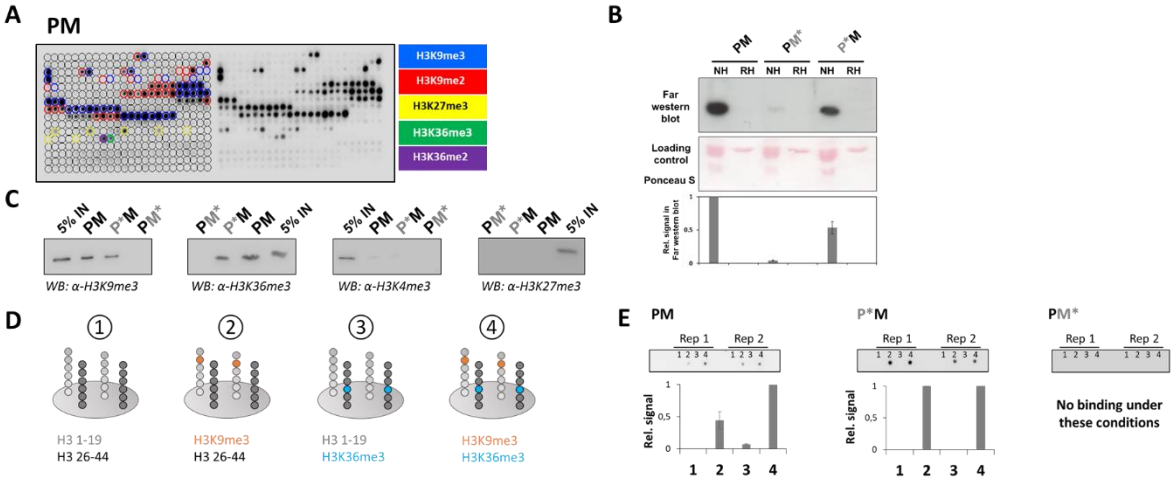


Figure 21. Compilation of the main results of the development of the double reading domain PM. **A)** Peptide array profile. The peptide spots are annotated on the left side of the array. **B)** Far-western blot analysis with PM double domain and the corresponding binding pocket mutants PM* and P*M. Asterisk indicates the mutated domain and the bar diagram shows a quantification of the data based on two repetitions. Error bars indicate the SEM. Binding to native histones (NH) and no binding to recombinant histones (RH) was observed. **C)** Mononucleosomal pull down with PM and the binding pocket mutants analyzed by western blot (WB) with H3K9me3, H3K36me3, H3K4me3 and H3K27me3 antibodies. **D)** Schematic view of the double peptide arrays. (1) unmodified H3 (1–19) and unmodified H3 (26-44), (2) H3K9me3 modified H3 (1–19) and unmodified H3 (26–44), (3) unmodified H3 (1–19) and H3K36me3 modified H3 (26–44), and (4) H3K9me3 modified H3 (1–19) and H3K36me3 modified H3 (26–44) and spotted on a glass slide. **E)** Binding of the PM double domain and the corresponding binding pocket mutants to the double peptide array in two replicates. The bar diagrams show the quantification of the double array based on two repetitions and error bars indicate the SEM. These figure panels were taken from (Mauser et al. 2017).

To study the genome-wide distribution of this dual mark, genome-wide CIDOP-seq studies were conducted with PM and the binding profiles compared with the single domain Dnmt3a PWWP and P*M (asterisk indicate the mutated domain) to identify overlapping genome regions (Fig. 22B). Cluster analysis showed a high similarity between density profiles obtained with H3K36me3 antibodies and Dnmt3a PWWP domain as well as a high correlation of H3K9me3 antibodies and MPP8 Chromo domain binding. In addition, a high correlation of the two PM replicates was observed but only a moderate correlation with the corresponding single domains, indicating that the binding profile of the double domain differs from the mere sum of the profiles of the single domain (Fig. 22C).

All these results clearly show that the new affinity tool is able to precipitate specific nucleosomes carrying simultaneously H3K9me3 and H3K36me2/3 histone modifications and they clearly demonstrate that double reading domains are a powerful tool to identify new dual modified chromatin states.

To elucidate the biological function of this bivalent chromatin state the software EpiExplorer was used and it was shown that the peaks obtained with PM were enriched at defined chromatin segments such as weakly transcribed genes and weak enhancers (Fig. 22D). GO analysis identified the involvement of the corresponding genes in biological processes like cell cycle regulation and metabolism (appendix 4, Fig. 6d, 7d). In addition, an association with the zinc finger Trim 28 SetDB1 pathway was identified (Fig. 22E), which confirmed and extended a previous finding showing the association of the H3K9me3 and H3K36me2/3 single modifications with the Trim 28 SetDB1 pathway (Blahnik et al. 2011).

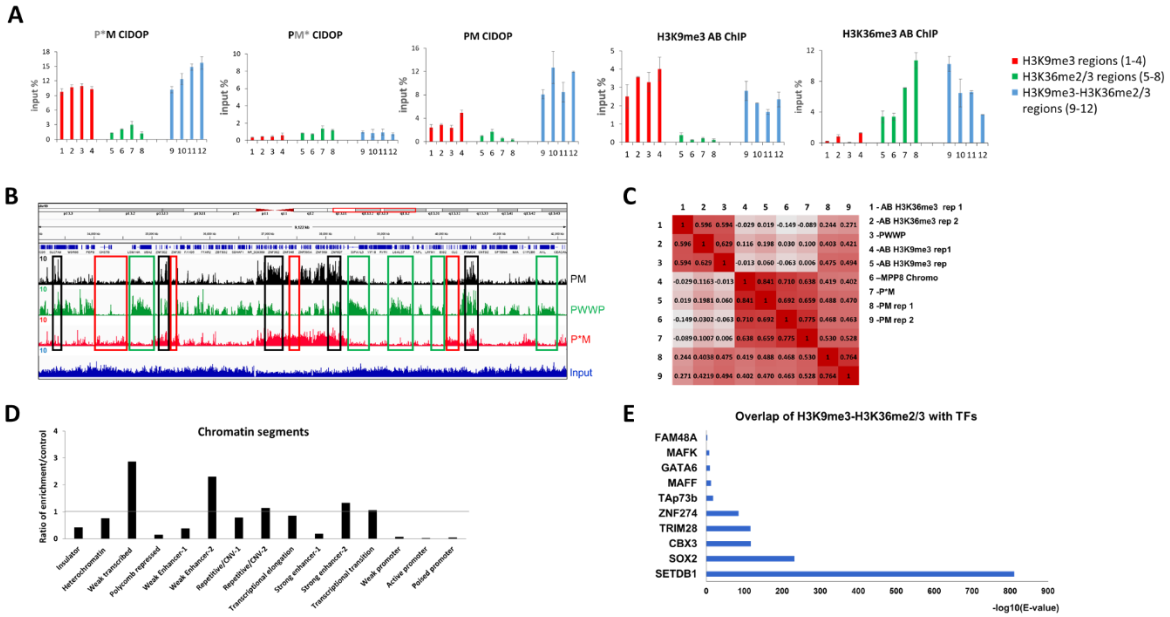


Figure 22. Investigation of the genome-wide distribution of H3K9me3 and H3K36me2/3 modified chromatin. **A)** CIDOP-qPCR of PM double domain and the binding pocket mutants. Twelve different genome regions containing only H3K9me3 (red), only H3K36me3 (green) or both modifications (blue) were analyzed. Additionally, ChIP-qPCR of the same genome regions was performed to confirm the presents or absence of the respective modifications. Two biological replicates of each experiment were conducted and error bars show the SEM. **B)** Genome browser view of the CIDOP-seq signals from PM, Dnmt3a PWWP domain and P*M. Strong PM signal is only observed if both single variant signals are present (black boxes), green boxes show only signal of Dnmt3a PWWP and red boxes only of P*M. **C)** Spearman correlation coefficients of raw read densities in 10-kb bins of PM, Dnmt3a PWWP, P*M, MPP8 Chromo and various antibodies (ChIP-seq data of antibodies is taken from (Kungulovski et al. 2014)). **D)** Peaks obtained from PM are enriched at defined chromatin segments. **E)** Overlap of PM with signals of DNA-binding factors from the Re-map database indicating overlap with the zinc finger Trim28 SetDB1 pathway. These figure panels were taken from (Mauser et al. 2017).

In summary, these data show that with double domains as highly specific affinity tools, one-step precipitation of mononucleosomes carrying two defined modification is possible and the direct mapping of the recovered DNA makes the identification of new chromatin states and the investigation of their biological functions feasible.

Review papers describing advantages of reading domains and their quality control

The investigation of histone post-translational modifications and their biological role is of great interest and their characterization is based on affinity reagents such as antibodies or recombinant reading domains. Because all downstream analyses rely on the quality of these affinity reagents, their careful validation is of major importance. The investigation of reading domains as an alternative to antibodies in chromatin immunoprecipitation experiments lead to an extensive analysis of these reagents, which culminated in the development of guidelines for the specific validation of the affinity reagents. These guidelines, advantage and

disadvantages of using recombinant reading domains and histone tail antibodies and the different available methods to control PTM affinity reagents are reviewed in (Kungulovski et al. 2015b), appendix 5 of this work.

One important method for the specificity analysis of antibodies and recombinant reading domains are MODified™ Histone Peptide Arrays. Their advantages and usage as a powerful validation method for chromatin affinity reagents is described in (Kungulovski et al. 2015a), appendix 6 of this work.

3.2 Unpublished results

3.2.1 Development of double reading domains for the analysis of H3K27me3 and H3K4me3 modified nucleosomes

It was shown in many studies, that the promoters of developmental genes in embryonic stem cells (ESC) simultaneously carry active H3K4me3 modifications and repressive H3K27me3 modifications. These promoters are called bivalent and they reside in a poised state, which facilitates a rapid silencing or activation of the target gene depending on the cellular trigger (Vastenhouw and Schier 2012; Voigt et al. 2013).

To specifically precipitate mononucleosomes carrying both H3K4me3 and H3K27me3, a double reading domain consisting of the CBX7 Chromo domain and the TAF3 PHD domain was generated including a linker (L) between both domains to guarantee flexibility of the newly designed construct. The CBX7 Chromo domain was shown to interact with the H3K27me3 modification (Ren et al. 2015) and the TAF3 PHD domain was shown to bind specifically H3K4me3 modified histone tails (Vermeulen et al. 2007). Both individual domains were validated as chromatin affinity reagent.

The specific interaction of the double reading domain CLT (CBX7 Chromo domain-Linker-TAF3 PHD domain) with the individual H3K4me3 and H3K27me3 marks was confirmed with CelluSpots peptide array (Fig. 23A). Further, the ability to bind only to modified native histones was analyzed using Far-western blot experiments. CLT showed strong interaction with

modified native histones (NH), but also a weak interaction with unmodified recombinant nucleosomes (RH) (Fig. 23B).

Unfortunately, CIDOP-qPCR experiments did not show a clear preference for genomic regions associated with both histone modifications (Fig. 23C), but they revealed a strong preference of CLT for genome regions associated with H3K27me3 irrespective of the presence of H3K4me3.

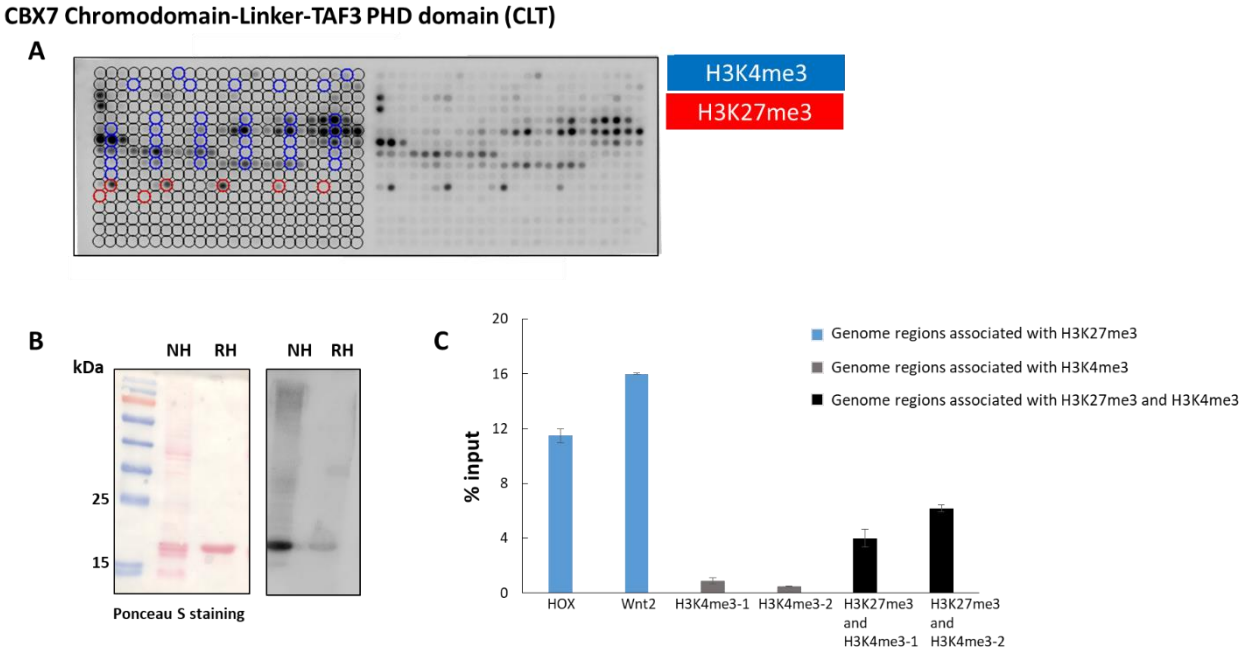


Figure 23. Specificity analysis of CBX7-L-TAF3 PHD (CLT). **A)** CelluSpots array of CLT showing expected results of both reading domains. CBX7 Chromo domain is interacting with H3K27me3 (red) and TAF3 PHD domain interacts with H3K4me3 (blue). **B)** Far-western blot showing stronger signal towards modified native histones (NH) than to unmodified recombinant histone H3 (RH). A Ponceau S staining as loading control is shown on the left. **C)** CIDOP-qPCR for CLT. HOX and Wnt2 are genome regions associated with H3K27me3 modifications (blue). Genome regions associated with H3K4me3 (grey) and regions associated with H3K4me3 and H3K27me3 (black). Two biological replicates of each experiment were conducted and error bars show the SEM.

Because the CIDOP-qPCR experiments showed a strong influence of the CBX7 Chromo domain on the binding of the double reader it was likely that in this configuration the TAF3 PHD domain cannot interact with the histone tail modification H3K4me3 as efficiently as observed for the single domain. Trying to eliminate steric influences, the orientation of the domains were swapped and a TLC double domain was cloned and purified. CelluSpots binding arrays confirmed that binding specificity of both domains was still maintained (Fig. 24A).

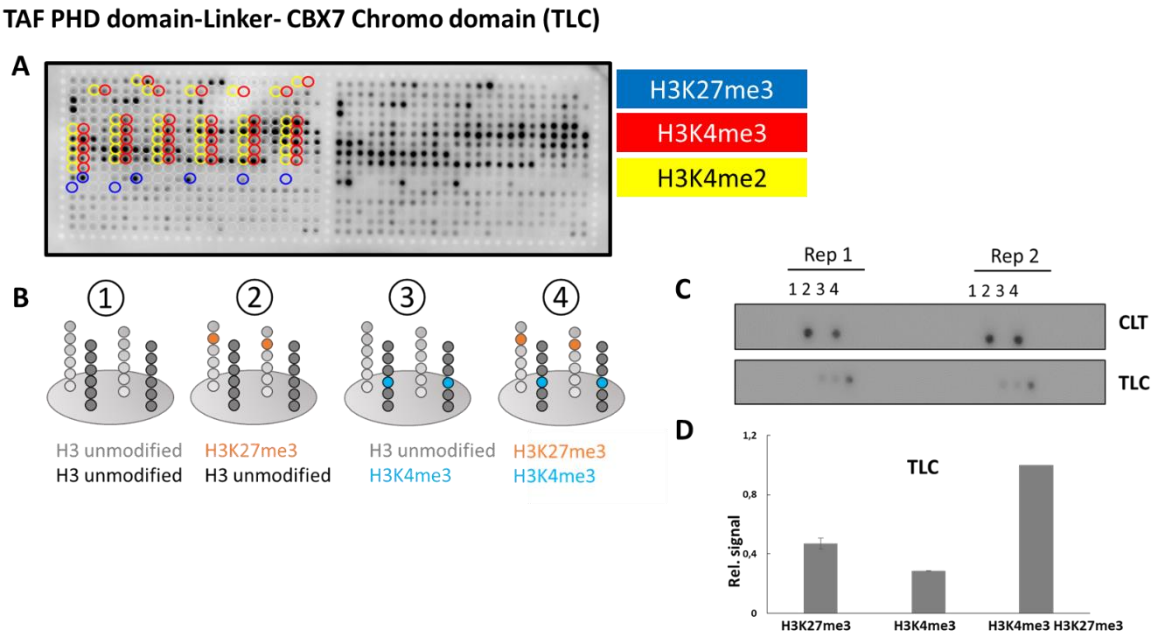


Figure 24. Investigation of TAF3 PHD Linker-CBX7 Chromo double domain (TLC) **A)** Specificity profile of the double reading domain showing interaction with H3K27me3 (blue) and H3K4me2/3 (yellow, red), which is mediated by CBX7 Chromo domain and TAF3 PHD domain respectively. **B)** Schematic illustration of the mixed peptide array containing H3 unmodified peptides (1), H3K27me3 and H3 unmodified (2), H3K4me3 and H3 unmodified (3) and H3K27me3 and H3K4me3 peptides (4). **C)** Probing the mixed arrays with CLT resulted in an almost equal interaction with H3K27me3 alone and the combination of H3K4me3 and H3K27me3 (spot 2 and 4). In contrast, the binding of TLC showed the strongest interaction with H3K27me3 and H3K4me3 peptides (spot 4). **D)** The bar diagram shows a quantification of the data obtained by TLC based on two repetitions. Error bars indicate the SEM.

Next, the ability of the TLC double reading domain to interact simultaneously with both modifications was analyzed using the newly develop method of mixed peptide arrays (Fig. 24B). Strikingly in this assay, the TLC double reading domain showed a stronger interaction with mixes spots containing both types of modified peptides and a weaker interaction with the single modifications (Fig. 24C and D). Additionally, the mixed peptide array experiment with the double reading domain CLT showed a similar binding to peptides carrying H3K27me3 modifications regardless of the second modification and no interaction with H3K4me3 peptides supporting the results obtained in CIDOP-qPCR experiments before (Fig. 24C and 23C).

After optimizing the new double reading domain, CIDOP-qPCR assay were carried out with TLC using mononucleosomes prepared from Hep G2 cells. The recovered DNA was analyzed using specific primers for genome regions carrying H3K27me3 only (light grey), H3K4me3 only (dark grey), or both marks (black). A clear preference towards genome regions associated with both H3K4me3 and H3K27me3 histone modifications was observed (Fig. 25). This is a very promising result suggesting that this domain will be applicable in future.

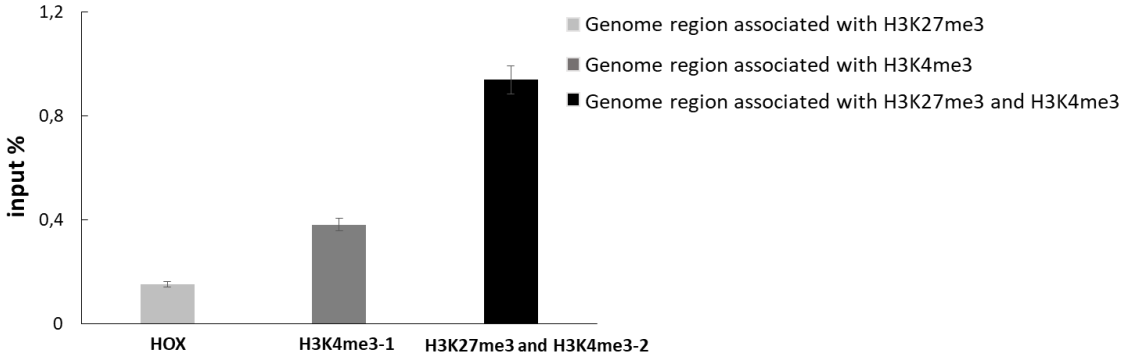


Figure 25. CIDOP-qPCR analysis of TLC. Strongest signals were observed for genome regions associated with both modifications H3K27me3 and H3K4me3 (black) and lower signals for the HOX (H3K27me3, light grey) and the H3K4me3-1 genome region (H3K4me3, dark grey).

3.2.2 Development of the double reading domain Dnmt3a PWWP-TAF3 PHD (PT)

H3K36me3 and H3K4me3 are marks on actively expressed genes and both marks are known to inhibit the Polycomb repressive complex 2 (PRC2). Therefore it would be interesting to investigate the co-occurrence of both modifications with double reading domains (Schmitges et al. 2011).

For that reason the Dnmt3a PWWP (P) domain that binds to H3K36me2/3, and the PHD domain of TAF3 (T) that binds to H3K4me3 were fused together. First, the binding specificity of the new double reading domain was validated using CelluSpots peptide arrays and the expected binding to individual peptides carrying H3K4me3 and H3K36me2/me3 was observed (Fig. 26A). Furthermore, the ability of the double reading domain PT to interact with modified nucleosomes was validated in Far-western blot experiments, where only interaction with native histones was observed and no binding to unmodified histones could be detected (Fig. 26B).

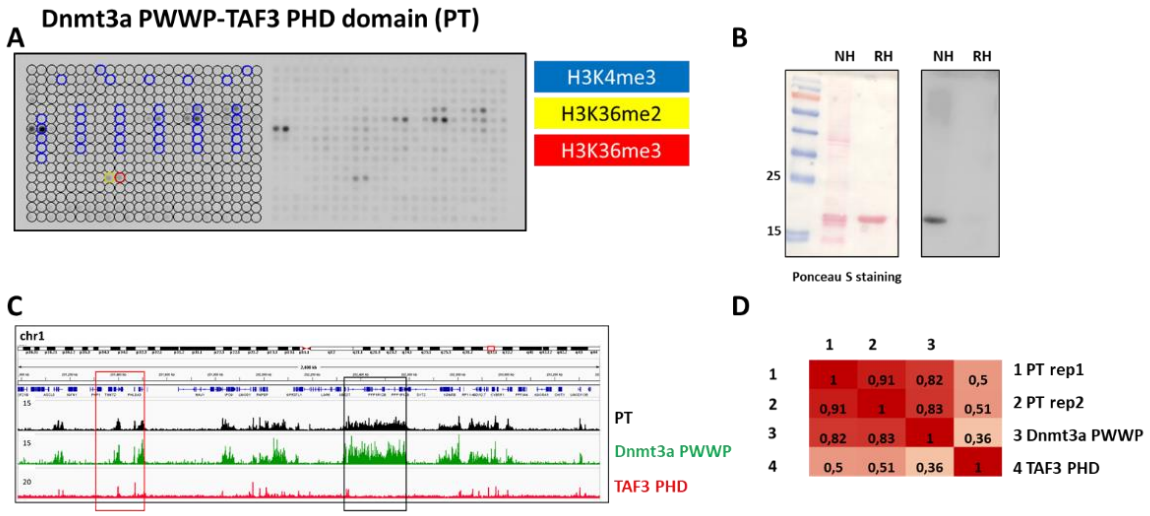


Figure 26. Specificity analysis of Dnmt3a PWWP-TAF3 PHD domain (PT). **A)** CelluSpots array of PT showing the expected results of both reading domains. Dnmt3a PWWP interacts with H3K36me2/3 (yellow, red) and TAF3 PHD interacts with H3K4me3 (blue). **B)** Far-western blot showing exclusive binding to modified native histones (NH) and no binding to recombinant histones (RH). Ponceau S staining as loading control is shown on the left. **C)** CIDOP-seq analysis for PT. Genome browser view of the read density in chromosome 1. Red box shows overleaping signals for PT, Dnmt3a and TAF3 PHD, whereas the black box indicate a region where only signals of PT and Dnmt3a PWWP overlap. **D)** Heatmap of the Spearman correlation coefficients of raw read densities in 5-kb bins of PT replicate 1 and 2, Dnmt3a PWWP and TAF3 PHD domain.

Next, the novel double reading domain PT was used in CIDOP-seq experiments. The obtained reads from PT and its corresponding single domains (Dnmt3a PWWP and TAF3 PHD) were mapped genome-wide and the distribution of both histone modifications was analyzed. The browser views of the genome-wide distribution showed some regions were the signal obtained from both single domains was summed up in the signal of the double reading domain (Fig. 26C red boxes) but also an equal number of regions were only the Dnmt3a PWWP signal contributed to the signal of PT (Fig. 26C black boxes). This result suggests that binding of both domains takes place independently and the PT profile represents the sum of profiles of the individual domains. Among the individual domains, the contribution of the PWWP domain appeared to dominate over the TAF3 domain. This conclusion is also supported by the correlation matrix of the read densities where both replicates of PT showed a higher correlation with PWWP (Spearman correlation of 0.82 and 0.83 (Fig. 26D)) than with the TAF3 PHD domain (Spearman correlation of 0.5).

These results together denote that the TAF3 PHD domain is not fully functional in this conformation to interact with isolated nucleosomes due to folding problems or to steric hindrance as seen before with the CLT double reading domain. It is possible that with a new design of the double reading domain or with optimized conditions a simultaneous readout can

be achieved similarly to the TLC double reading. At the time of these experiments the double peptide arrays methods was not yet available. Double peptide arrays of PT will be carried out in future.

3.2.3 Development of the double reading domain Dnmt3a PWWP-ATRX ADD (PA) for the detection of H3K36me2/3 and H3K9me3 modified nucleosomes

In the beginning of this work, two approaches were tested to design a double reading domain for H3K36me2/3 and H3K9me3 modified nucleosomes by fusing the Dnmt3a PWWP (P) domain, which interacts with H3K36me2/3 modifications, and either the MPP8 Chromo domain (M) or the ATRX ADD (A) domain, which both bind to H3K9me3. Both domains were validated as chromatin PTM affinity reagents.

The specificity validation with PA CelluSpots peptide arrays showed that both individual modifications where bound by the double reading domain as expected. Far-western blot studies revealed only an interaction with native histones but not with recombinant histones (Fig. 27A and B).

Unfortunately, in one attempt PA failed in CIDOP experiments, because the interaction with isolated nucleosomes was too weak and not enough material could be obtained for NGS analysis (data not shown). Therefore the double reading domain as well as the conditions in CIDOP experiments have to be optimized. However, as very good results were obtained with the PM double reading domain, optimization of PA was not further attempted.

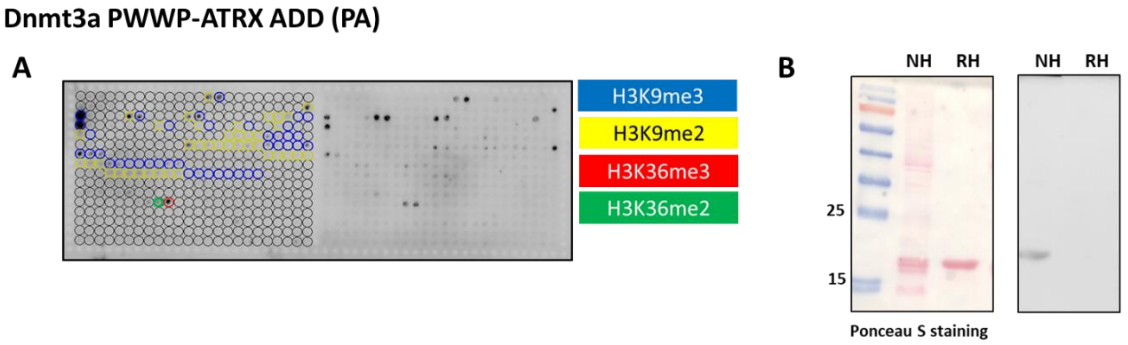


Figure 27. Specificity analysis of Dnmt3a PWWP-ATRX ADD (PA). **A)** CelluSpots array of PA showing the expected results of both reading domains. Dnmt3a PWWP interacts with H3K36me2/3 (green, red) and ATRX ADD interacts with H3K9me2/3 (yellow, blue). **B)** Far-western blot showing exclusive binding to modified native histones (NH) and no binding to recombinant histones (RH). A Ponceau S staining as loading control is shown on the left.

3.2.4 Development of the double reading domain MPP8 Chromo-Linker-double Tudor (MLdT) for the binding of H3K9me3 and H4K20me3 modified nucleosomes

The H3K9me3 modification is a hallmark of pericentromeric chromatin. Another repressive histone modification is the H4K20me3 mark, which is also found at pericentromeric heterochromatin (Fischle et al. 2003; Lachner et al. 2003). H4K20me3 is set by SUV420H1 and SUV420H2 and it was reported previously to be important for the compaction of chromatin, because H4K20me3 promotes the binding of the H4 tail to adjacent nucleosomes.

Furthermore, previous reports suggested that SUV39H introduces the H3K9me3 modification, followed by HP1 interaction which recruits SUV420H resulting in the establishment of H3K9me3 and H4K20me3 at pericentromeric heterochromatin (Schotta et al. 2004).

To analyze the co-occurrence of these two modifications and get a deeper insight into heterochromatin formation, the MPP8 Chromo domain (M), which interacts with H3K9me3, and the KDM4A double Tudor mutant D969A (dT), which binds to H4K20me3, were fused together. The KDM4A double Tudor domain normally interacts with global di- and trimethylation on H3 and H4, but by protein engineering the binding specificity of the KDM4 D969A variant was shifted towards the specific readout of H4K20me3 (Kungulovski et al. 2014). Both domains were validated as chromatin PTM affinity reagents.

As expected, CelluSpots peptide arrays showed binding of MLdT to peptides carrying individual H3K9me3, H3K27me3 and H4K20me3 validating the specificity of the individual

domains (binding of the MPP8 Chromo domain to H3K27me3 is only observed on peptide level (Kungulovski et al. 2014). In addition, MLdT showed exclusive binding to native histones (Fig. 28A and B). The next step of the validation of the MLdT domain would have been a CIDOP experiment analyzed by qPCR with primer pairs specific for regions with H3K9me3 only, H4K20me3 only and both modifications. Unfortunately, the primer design turned out to be extremely difficult due to the close connection of both marks.

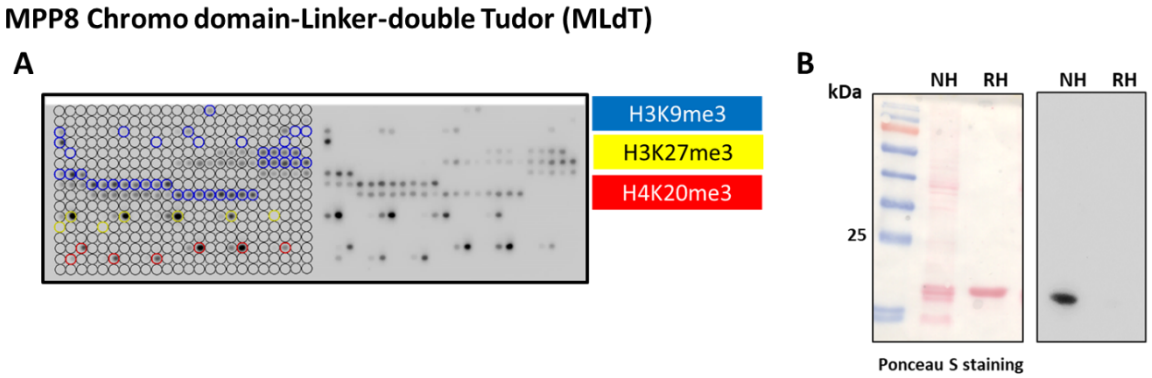


Figure 28. Specificity analysis of the MPP8 Chromo -Linker-double Tudor (MLdT) double domain. **A)** CelluSpots array of MLdT showing expected results of both reading domains. The MPP8 Chromo domain showed interaction with H3K9me3 peptides (blue) as well as an artefact binding on peptide level to H3K27me3 peptides (yellow). The KDM4A double Tudor mutant D969A domain interacted with H4K20me3 peptides (red).**B)** Far-western blot showing exclusive binding to modified native histones (NH) and no binding to recombinant histones (RH). A Ponceau S staining is shown on the left side as loading control.

3.2.5 Development of the double reading domain CBX7 Chromo-Linker-MPP8 Chromo (CLM) for binding of H3K27me3 and H3K9me3 modified nucleosomes

Like H3K9me3, H3K27me3 is associated with repressive chromatin, but it is found on facultative heterochromatin and on inactive or poised promoters (Payer and Lee 2008; Plath et al. 2003). The MPP8 Chromo domain was shown to bind to H3K9me3 and the Chromo domain of CBX7 interacts with H3K27me3. Both domains were validated as chromatin PTM affinity reagents.

To investigate the potential co-occurrence of H3K9me3 and H3K27me3, the double reading domain CLM was generated, consisting of the CBX7 Chromo domain connected via a linker

with the MPP8 Chromo domain. CelluSpots peptide array as well as Far-western blot assay showed the expected binding preference of this double reading domain (Fig. 29A and B).

Unfortunately, in an initial attempt CLM failed in CIDOP-qPCR experiments. No preferred binding to genome regions associated with both modifications, H3K27me3 and H3K9me3, could be observed (data not shown) indicating that with this design of CLM, no specific dual readout of bivalent chromatin is possible and the double domain, as well as the conditions in CIDOP experiments have to be optimized.

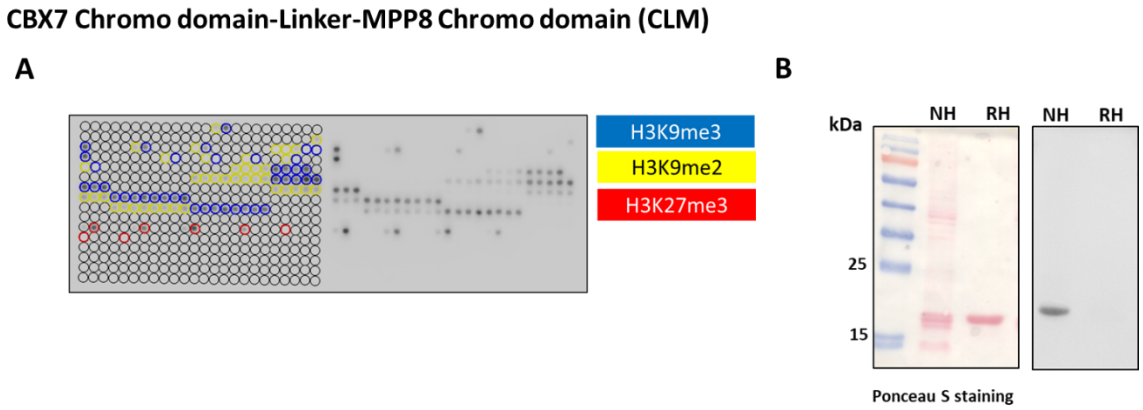


Figure 29. Specificity analysis of the CBX7 Chromo-Linker-MPP8 Chromo (CLM) double domain. **A)** CelluSpots array of CLM showing the expected results of both reading domains. The CBX7 Chromo domain interacted with peptides harbouring H3K27me3 modifications (red). MPP8 Chromo domain showed interaction with H3K9me2/3 peptides (yellow, blue). **B)** Far-western blot showing exclusive binding to modified native histones (NH) and no binding to recombinant histones (RH). A Ponceau S staining is shown on the left side as loading control.

3.2.6 Development of homotypic double reading domains consisting of two identical domains

Nucleosomes consist of two H2A, H2B, H3 and H4 proteins, respectively. Symmetrically modified nucleosomes carry the same modification on both corresponding histone tails, whereas in asymmetrically modified nucleosomes only one histone tail is affected (Fig. 30). It was seen, that many chromatin effector proteins contain multiple chromatin interacting domains for the same or different histone tail PTM (Ruthenburg et al. 2007) and that multiple binding events enhance the binding affinity and also increase specificity (Voigt and Reinberg 2011).

Additionally, certain histone tail modification patterns can also prevent binding of chromatin effector proteins. For example, it was previously shown that symmetric modification of the H3 tail with the K36me3 modification inhibits methylation via PRC2 complex (Voigt et al. 2012).

However, so far no tools are available for the straightforward investigation of the genome-wide distribution of symmetrically and asymmetrically modified histone tails. Therefore, in this work, two homotypic double readers MLM (MPP8 Chromo-Linker-MPP8 Chromo) and PLP (Dnmt3a PWWP-Linker-DNmt3a PWWP) were designed and tested. The idea behind the design was, that the single domain should bind to its target modification regardless of the modification symmetry but that the double reading domain should preferentially interact with symmetrically modified nucleosomes (Fig. 30).

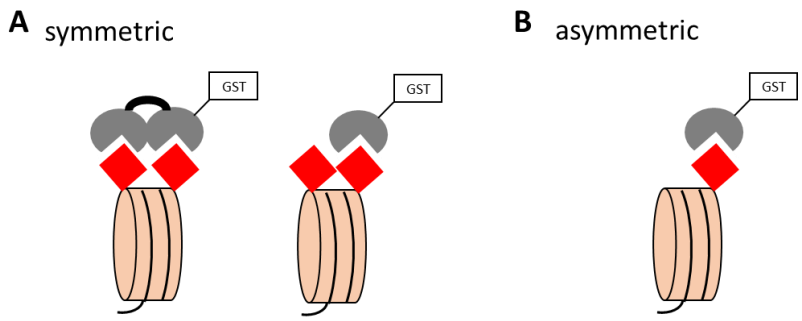


Figure 30. Schematic illustration of symmetrically and asymmetrically modified nucleosomes and the readout with single and double reading domains.

The specific interaction with their predicted histone modifications (MLM to H3K9me3 and PLP to H3K36me2/3) was analyzed and compared to their corresponding single domains in CelluSpots peptide arrays and Far-western blot assays (Fig. 31A and B). Both homotypic double reading domains interacted specifically with H3K9me3 and H3K36me3 modified peptides, respectively. In addition, MLM as well as PLP interacted specifically with native histone but not with recombinant histones (Fig. 31C and D). Both double reading domains exhibited a stronger interaction compared to their single domain counterparts. It is likely that due to the dual readout, the affinity towards double modified chromatin is stronger compared to the separate interaction of the single domain.

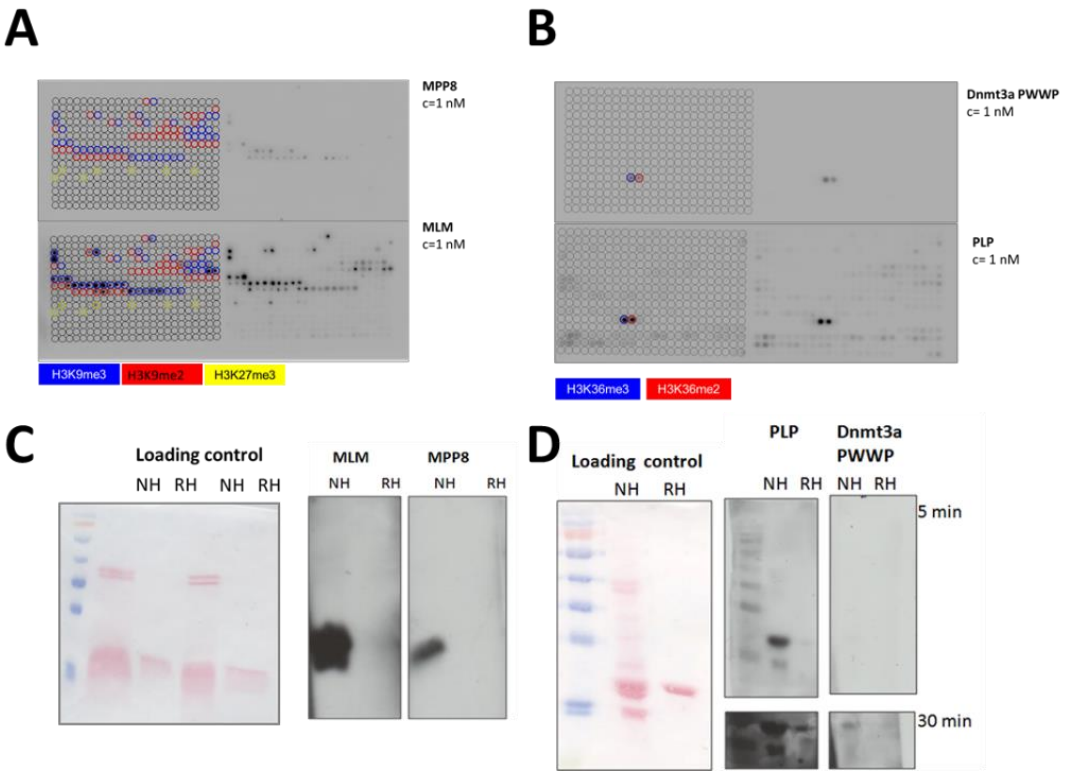


Figure 31. Specificity analysis of MPP8-Linker-MPP8 (MLM) and Dnmt3a PWWP-Linker-Dnmt3a PWWP (PLP) double domain and comparison to the corresponding single domains. **A)** CelluSpots array of MLM showing interaction to peptides carrying H3K9me2/3 modifications (blue, red) and the artefact binding on peptide level to H3K27me3 (yellow). CelluSpots array of MPP8 Chromo domain revealed the same binding specificity. Investigation of the binding profile of PLP in comparison to Dnmt3a PWWP domain showed specific interaction with H3K36me2/3 peptides (red, blue). **B)** Far-western blot showing exclusive binding to modified native histones (NH) and no binding to recombinant histones (RH) for MLM, MPP8 Chromo domain, PLP and Dnmt3a PWWP domain. Both double domains show stronger interaction than the corresponding single domains. Corresponding Ponceau S stained images are shown as loading controls on the left.

Next, the double reading domains were compared to the single domains in pull down experiments analyzed by western blot using an anti-H3 antibody. Interestingly, with increasing amount of salt in the washing buffer, the double reading domains showed a stronger interaction with nucleosomes than their single domain counterparts (Fig. 32A-C). Importantly, no binding to nucleosomes could be observed for the single Dnmt3a PWWP domain under stringent conditions, but the double reading domain PLP was able to interact with nucleosomes under these conditions.

The single MPP8 Chromo domain, as well as the double reading domain MLM domain showed binding towards nucleosomes under mild and stringent conditions, however MLM exhibited a stronger interaction than MPP8 Chromo domain. This is in agreement with previous studies

were MPP8 Chromo showed a K_d of 100 - 200 nM and PWWP has an K_d of 50 - 100 µM (Bock et al. 2011b; Chang et al. 2011).

Strikingly, both double reading domains showed a stronger interaction in comparison to their corresponding single domain, suggesting that the double reading domains interact stronger with nucleosomes due to multi-dentate binding of bivalent modified histone tails.

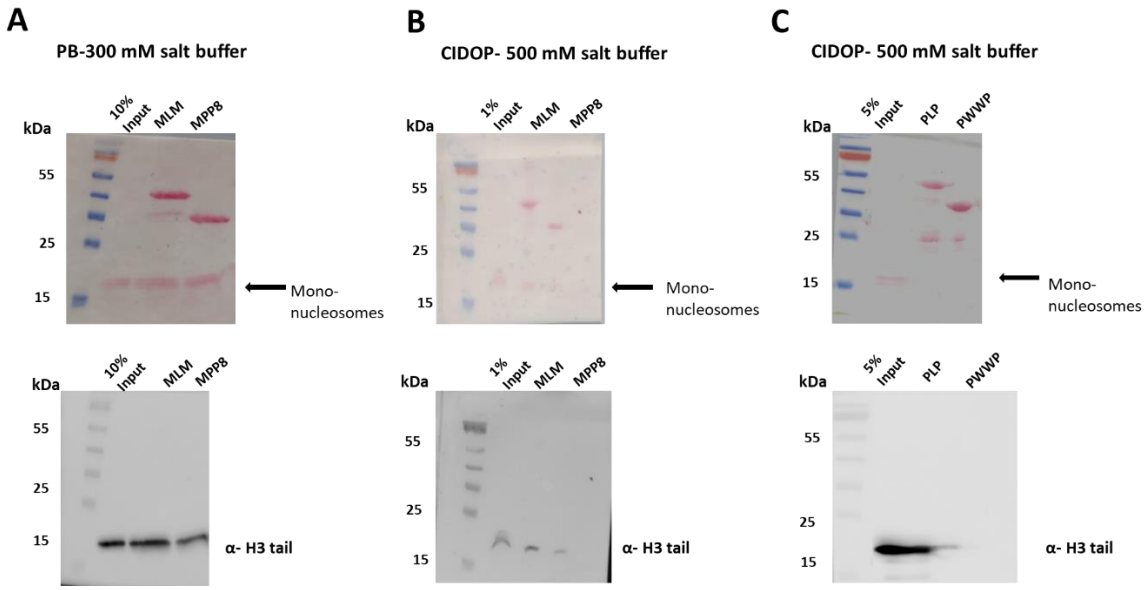


Figure 32. Chromatin pull down studies with homotypic double and single domains. **A)** Pull down with MLM and MPP8 Chromo domain using 300 mM wash buffer. **B)** Pull down with MLM and MPP8 Chromo domain using 500 mM wash buffer and detergents. **C)** Pull down with PLP and Dnmt3a PWWP domain using 500 mM wash buffer and detergents. Pull down was detected by western blot using an H3 antibody. A Ponceau S staining represents the loading control.

CIDOP-qPCR experiments were carried out with the homotypic double reading domains and the corresponding single domains. For this, different genomic regions associated with different histone modifications were analyzed. MLM showed the same specificity as the MPP8 Chromo domain, both binding to genome regions associated with H3K9me3 marks and only background signals for regions associated with unrelated histone marks (Fig. 33A and B). Unfortunately PLP failed in CIDOP-qPCR experiments, because no discrimination between related (H3K36me3 associated genome regions) and unrelated genome regions was observed (data not shown).

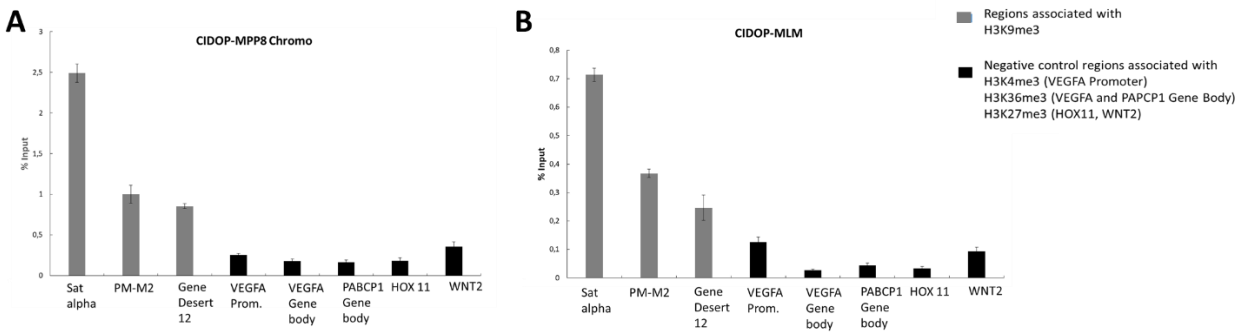


Figure 33. CIDOP-qPCR experiments of MPP8 Chromo domain and MLM double reading domain. **A)** MPP8 Chromo domain. Investigation of the interaction with genome regions associated with H3K9me3 modifications (grey bars) and control regions (black bars). Experiments were conducted three times and error bars indicate the SEM. **B)** MLM. Investigation of the interaction with genome regions associated with H3K9me3 modifications (grey bars) and control regions (black bars). Experiments were conducted three times and error bars indicate the SEM.

In addition, CIDOP-seq was performed for MLM and the single MPP8 Chromo domain. As described before, histone tails can be modified in as symmetric or asymmetric way. Differences in the genome-wide distribution of the double reader MLM and the single MPP8 Chromo domain could arise from symmetrically or asymmetrically modified nucleosomes. Unfortunately, the comparison of the genome-wide distribution did not reveal significant differences in the data obtained from these two CIDOP-seq experiments (data not shown). Therefore, no evidence for the discrimination of asymmetric and symmetric modified nucleosomes could be provided in these experiments.

3.2.7 Generation of double reading domains with specificity for DNA methylation and histone marks

DNA methylation is an important key regulator of epigenetic processes and influences gene expression. Promoter methylation is associated with gene repression, while DNA methylation in gene bodies correlates with active gene expression (Deaton and Bird 2011). Additionally DNA methylation and histone tail modifications are known to affect each other and influence together chromatin-templated processes (Ruthenburg et al. 2007; Bartke et al. 2010).

To analyze the co-occurrence of DNA methylation and defined histone PTMs, new double reading domains with specific interaction for DNA methylation and histone modifications were generated. The DNA binding domain of MBD2 (Bird et al. 1999) was isolated and fused to the validated histone interacting domains MPP8 Chromo domain (interacts with H3K9me3),

Dnmt3a PWWP (binds to H3K36me2/3), ATRX ADD (interacts with H3K9me3 if H3K4 is unmodified) and CBX7 Chromo domain (binds to H3K27me3) generating the following double domains: MBD2-L-MPP8; MPP8-L-MBD2; MBD2-L-ADD; ADD-L-MBD2; CBX7-L-MBD2; PWWP-L-MBD2 and MBD2-L-PWWP. In all of these constructs both domains were connected via a flexible Linker (L).

3.2.7.1 Generation of double reading domains with specificity for DNA methylation and histone modifications using MBD2 as DNA interacting domain

First, the DNA interacting domain of the Methyl-CpG-binding domain 2 (MBD2) protein alone and in the fusion constructs was analyzed for specific interaction with methylated DNA. The DNA interacting domain of the MBD2 protein will be abbreviated as MBD2 in the following parts.

Binding to methylated DNA was analyzed with a dot blot assays using methylated and unmethylated DNA as substrate that were spotted on a Amersham Hybond-N+ membrane in different amounts ranging from 200 ng to 20 pg. MBD2 alone as well as all fusion proteins showed binding to methylated DNA at higher concentrations and, as expected, no binding to unmethylated DNA was observed (Fig. 34A-H).

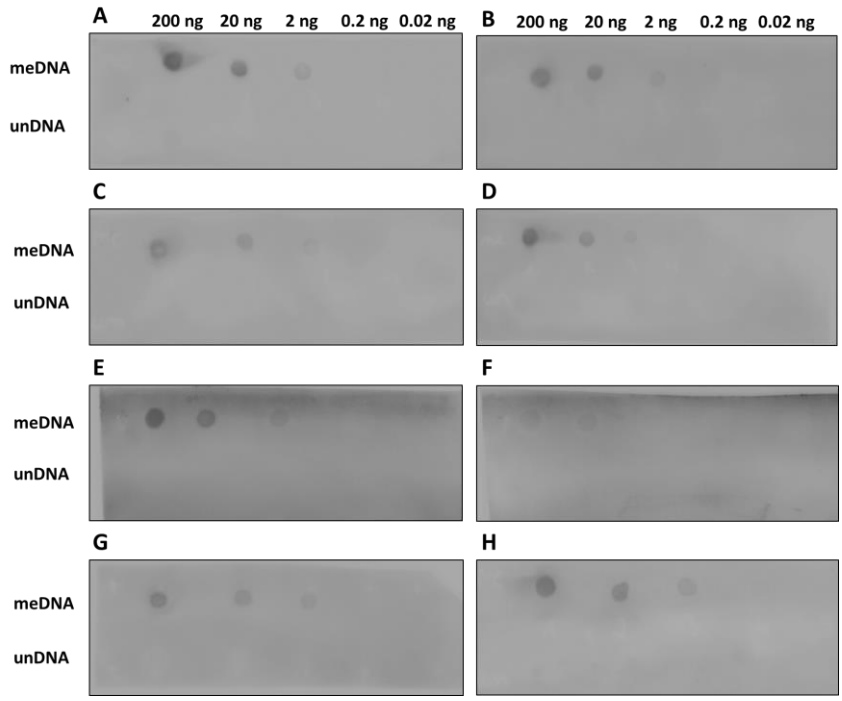


Figure 34. Dot blot experiments of double reading domains with binding specificity towards DNA methylation and histone tail modifications. Methylated DNA and unmethylated DNA were spotted in different amounts ranging from 200 ng to 0.02 ng. Binding of ATRX ADD-Linker-MBD2 **(A)**, MPP8-Linker-MBD2 **(B)**, MBD2-Linker-Dnmt3a PWWP **(C)**, MBD2-Linker-ATRX ADD **(D)**, MBD2-Linker-MPP8 Chromo domain **(E)**, Dnmt3a PWWP-Linker-MBD2 **(F)**, CBX7 Chromo domain-Linker-MBD2 **(G)** and Methyl binding domain of MBD2 protein **(H)**.

Next, the double reading domains were tested in Far-western blot assays. The Far-western blot assay showed binding of all constructs to native histones, additionally also linker histones (size ~35 kDa) were bound, but surprisingly also binding to unmodified recombinant histone H3 was observed (Fig. 35A). Binding studies with the MBD2 domain alone clarified that the binding to linker histones and to recombinant histone H3 derived from the DNA interacting domain. This conclusion was also supported by the fact that the single domain CBX7 Chromo domain showed only interaction with native histones as expected and observed before (Fig. 35B).

Furthermore, the specific binding of the histone interacting domains in the fusion constructs to modified peptides was analyzed using CelluSpots peptide arrays. Unfortunately, most of the double reading domains did not show specific interaction with their target peptides. Out of the tested combinations, only CLM showed to some extent binding to the expected H3 tail modifications, but even these arrays were not very specific (Fig. 35C).

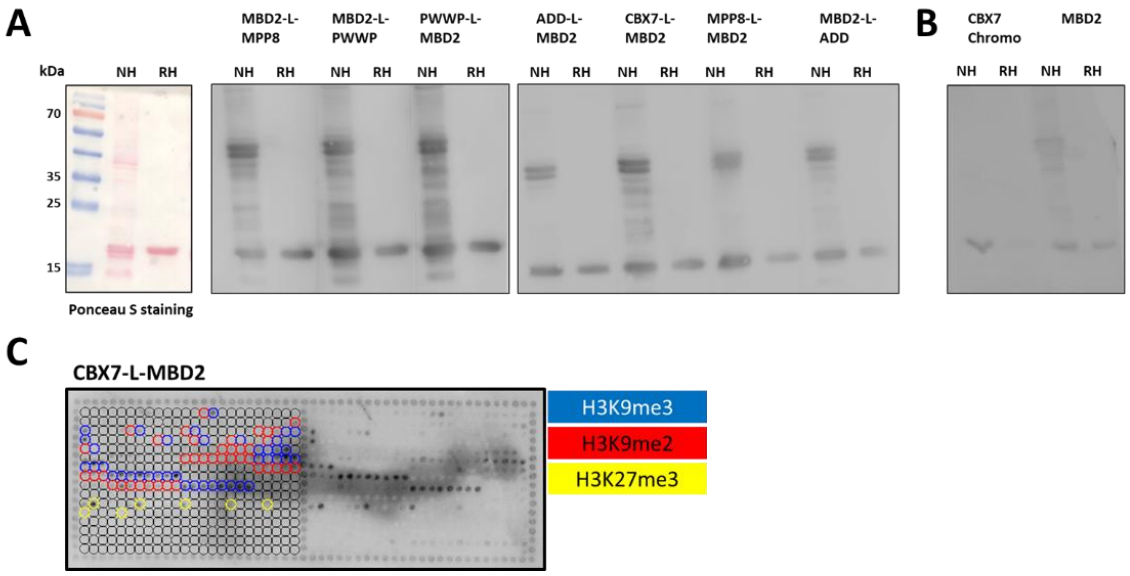


Figure 35. Specificity analysis of double reading domains with binding specificity towards DNA methylation and histone tail modifications. **A)** Far-western blot analysis of MBD2-L-MPP8, MBD2-L-PWWP, PWWP-L-MBD2, ADD-L-MBD2, CBX7-L-MBD2, MPP8-L-MBD2 and MBD2-L-ADD binding to modified native histones (NH) and recombinant histones (RH). A Ponceau S staining of the gel is shown as loading control on the left. **B)** Far-western blot analysis of binding of CBX7 Chromo domain and DNA binding domain of MBD2 to modified native histones (NH) and recombinant histones (RH). **C)** CelluSpots array of CBX7-Linker-MBD2 showing interaction to peptides carrying H3K27me3 (yellow) and H3K9me2/3 modifications (blue, red).

3.2.7.2 Genome wide analysis of the DNA precipitated by double reading domains

Because only CLM showed partially acceptable results, binding pocket mutants were designed for this double reading domain and CIDOP-qPCR experiments were carried out. The precipitated DNA was analyzed using primers specific for genome regions associated with H3K27me3 and unmethylated DNA, methylated DNA without H3K27me3 modifications and the combination of H3K27me3 and methylated DNA (Fig. 36A). The double reading domain showed an equal affinity to genome regions associated with methylated DNA regardless of the presence of the H3K27me3 modification indicating a dominating influence of the MBD2 domain. Furthermore the MBD2 binding pocket mutant with intact CBX7 Chromo domain (CLM*, asterisk indicates the mutated domain) showed no specificity towards H3K27me3 associated regions (Fig. 36B). These data indicate that the CBX7 Chromo domain does not contribute to the precipitation of chromatin and the configuration of this double domain is not functional.

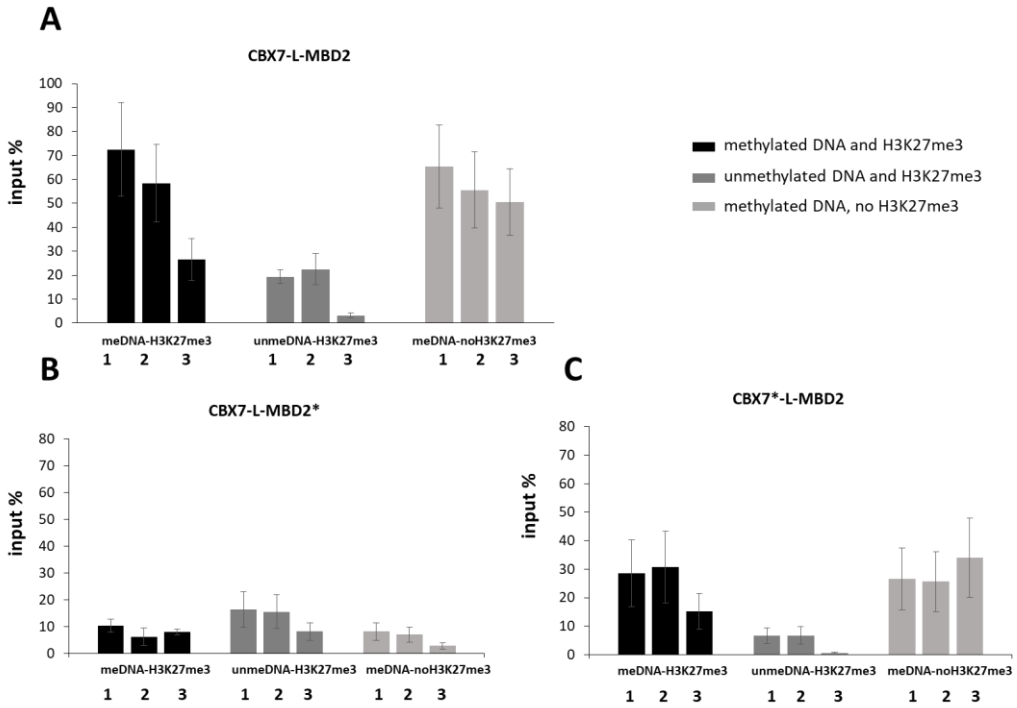


Figure 36. CIDOP-qPCR experiments of nucleosome pull down prepared using CLM **(A)** and the corresponding variants CLM* **(B)** and C*LM (asterisk indicates the mutated domain) **(C)**. Black bars represent the regions associated with DNA methylation and H3K27me3 modifications. Dark grey bars correspond to genome regions associated with H3K27me3 but no DNA methylation and light grey bars represent genome regions associated with methylated DNA but no H2K37me3. Experiments were conducted three times and error bars indicate the SEM.

3.2.7.3 Generation of double reading domains with specificity for DNA methylation and histone modification using MBD1 as DNA interacting domain

To avoid the dominating influence of the MBD2 domain in the fusion constructs, the DNA binding efficiency of the Methyl-binding domain 1 (abbreviated as MBD1 in the following parts) protein was compared to MBD2. DNA pull down experiments were conducted with the DNA binding domains of MBD1 and MBD2 and analyzed via qPCR. Two genome region FZD 10 (a region associated with methylated DNA) and VEGF-A promoter (a region associated with low levels of methylated DNA) were analyzed, and two different washing buffers differing in their salt concentrations (150 mM salt vs. 500 mM salt) were used (Fig. 37A and B). MBD1 showed an overall weaker interaction with methylated DNA. The MBD2 domain showed unspecific binding at 150 mM salt, whereas MBD1 showed still some preference for methylated DNA. At 500 mM salt, overall weaker binding was observed for both domains, but MBD1 as well as MBD2 interacted specifically with methylated DNA. Therefore, the MBD1

domain was fused to the validated HiMIDs (MBD1-L-MPP8; MPP8-L-MBD1; ATRX-L-MBD1; TAF3-L-MBD1; MBD1-L-TAF3; PWWP-L-MBD1; MBD1-L-PWWP) and analyzed for simultaneously readout of the chromatin modifications.

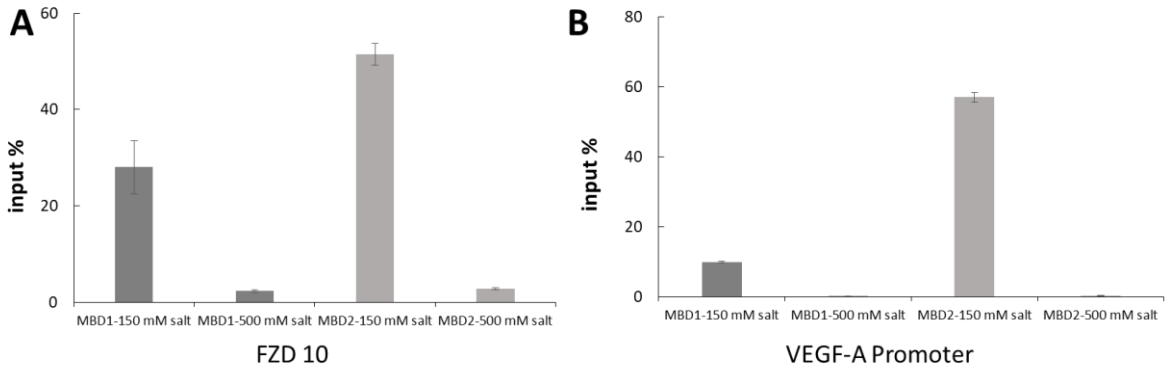


Figure 37. Comparison of pull down efficiency of MBD1 (dark grey bars) and MBD2 (light grey bars) using different salt conditions. **A)** The genome region FZD 10 is associated with methylated DNA. **B)** The VEGFA promoter region is associated with mainly less methylated DNA in HEK293 cells. Experiments were conducted two times and error bars indicate the SEM.

As a next step, binding to methylated DNA was analyzed by dot blot assays, indicating a specific interaction with methylated DNA of MBD1 alone but also in all fusion constructs. However, with MBD1 and MBD1 constructs, signal was only detected with 200 ng methylated DNA, whereas MBD2 was able to interact with much lower amounts of DNA (2 ng) (Fig. 34, Fig. 38).

The double reading domains were all capable of methylated DNA readout, interestingly the double reading domains containing the PHD domain of TAF3 also showed an interaction with unmethylated DNA. This interaction arose from the histone interacting domain, because MBD1 alone and all the other histone interacting double domains did not show binding to unmethylated DNA.

In CelluSpots peptide arrays the, MPP8-L-MBD1, ATRX ADD-L-MBD1 and PWWP-L-MBD1 showed to some extent the predicted peptide binding specificity. MPP8-L-MBD1 showed weak interaction with peptides carrying H3K9me2 and H3K9me3 modifications (Fig. 39A). ATRX ADD-L-MBD1 showed interactions with peptides harbouring H3K9me2 and H3K9me3 modifications (Fig. 39B), whereas PWWP-L-MBD1 showed unspecific binding to almost all

peptides, but the predicted H3K36me2 and H3K36me3 peptides showed a slightly increased signal (Fig. 39C).

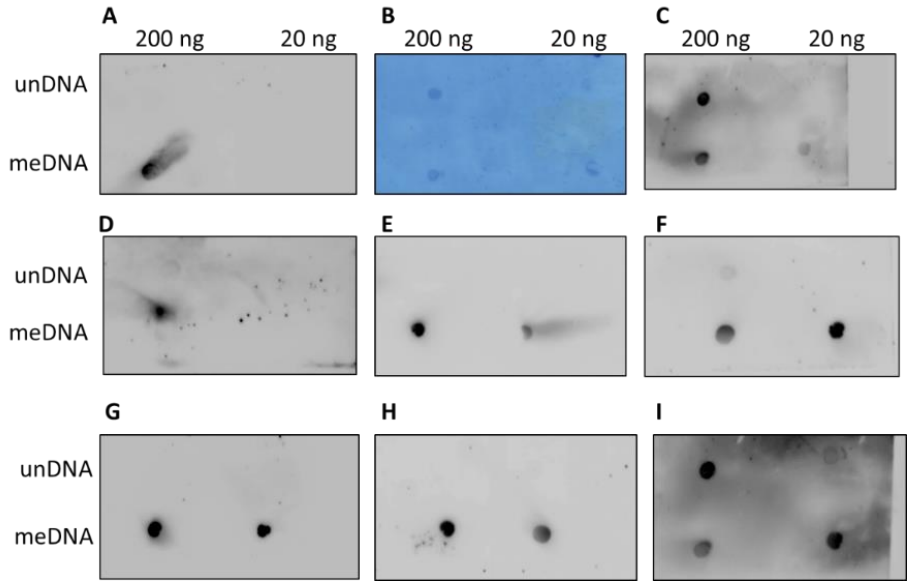


Figure 38. Dot blot experiments of double reading domains with binding specificity towards DNA methylation and histone tail modifications. 200 ng and 20 ng of methylated DNA and unmethylated DNA was spotted on an Amersham Hybond-N+ membrane. Binding of the methyl binding domain of the MBD1 protein **(A)**, control staining with methylene blue **(B)**, MBD1-Linker-TAF3 PHD **(C)**, ATRX ADD-Linker-MBD1 **(D)**, MBD1-Linker-Dnmt3a PWWP domain **(E)**, Dnmt3a PWWP-Linker-MBD1 **(F)**, MBD1-Linker-MPP8 Chromo domain **(G)**, MPP8 Chromo domain-Linker-MBD1 **(H)** and TAF3 PHD domain-Linker-MBD1 **(I)**.

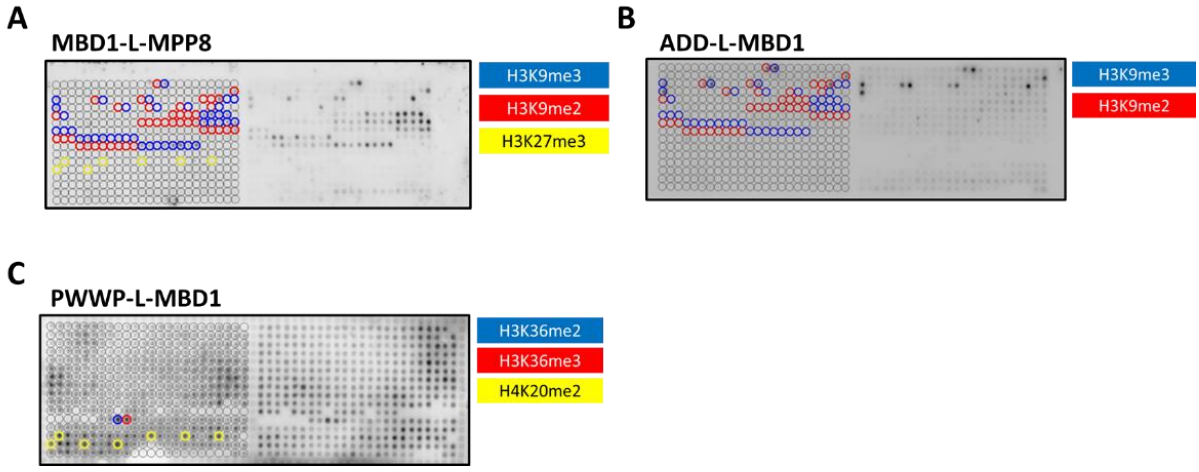


Figure 39. Specificity analysis of double reading domains with binding specificity towards DNA methylation and histone tail modifications. **A)** CelluSpots array of MBD1-Linker MPP8 Chromo domain showing an interaction with peptides harbouring H3K9me3 (blue) and H3K9me2 (red). **B)** Specificity analysis of ATRX ADD-Linker-MBD1 showing interaction with H3K9me2/3 peptides (red, blue). **C)** Dnmt3a PWWP-Linker-MBD1 showed specific binding to H3K36me2/3 (blue, red) and unspecific binding to H4K20me2 (yellow).

CIDOP-qPCR experiments were conducted with the double reading domains MBD1-L-MPP8, MBD1-L-PWWP and ATRX ADD-L-MBD1.

A genome region associated with H3K9me3 and DNA methylation (ZNF420 promoter region) was used for readout. MBD1, MPP8 Chromo domain as well as MBD1-L-MPP8 (M1LM) showed interaction to methylated DNA connected with H3K9me3 modification (Fig. 40A). The same genomic region associated with H3K9me3 and DNA methylation (ZNF420 promoter region) was analyzed for ATRX ADD-L-MBD1 (ALM1). Interestingly, ATRX ADD-L-MBD1 showed a much stronger interaction with methylated DNA and H3K9me3 compared to its single variants (Fig. 40B). A genome region associated with H3K36me3 and DNA methylation (VEGF-A gene body) was used for the analysis of MBD1-L-PWWP, as well as for MBD1 and Dnmt3a PWWP domain pull down efficiency. Also in this experiment, the double reading domain MBD1-L-PWWP showed a much stronger interaction to this genome region than the corresponding single domains (Fig. 40C).

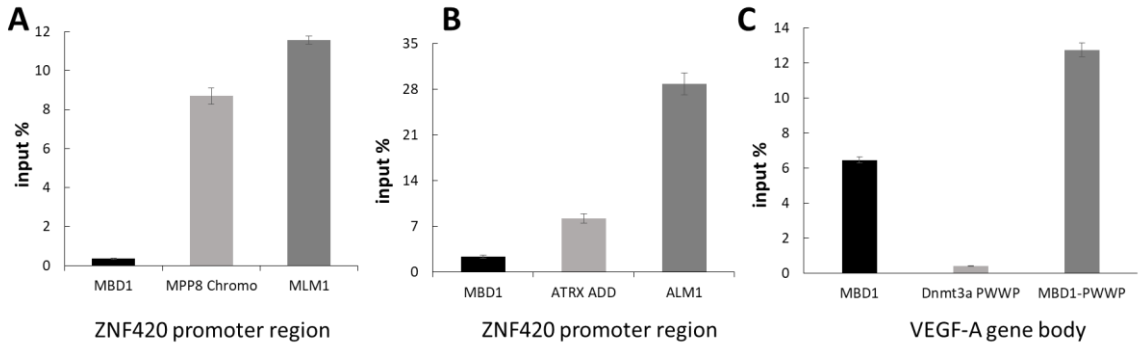


Figure 40. CIDOP-qPCR experiments of MPP8-Linker-MBD1 (MLM1), ATRX ADD-Linker-MBD1 (ALM1) and MBD1-Linker-Dnmt3a PWWP (M1LP). **A)** Genome regions associated with methylated DNA and H3K9me3 (ZNF420 promoter) analyzed after pull down with MLM1 and its corresponding single domains **B)** Genome regions associated with methylated DNA and H3K9me3 (ZNF420 promoter) analyzed after pull down with ALM1 and its corresponding single domains. **C)** Genome regions associated with methylated DNA and H3K36me3 (VEGF-A gene body) analyzed after pull down with M1LP and its corresponding single domains. Experiments were conducted two times and error bars indicate the SEM.

As a next step, it will be necessary to compare the double reading domains to their corresponding binding pocket mutants to investigate if dual binding occurs. Afterwards, double domain binding can be analyzed genome-wide to investigate the genome-wide distribution of the combined occurrence of DNA methylation and the defined histone modifications.

4 Discussion

The impressive diversity of cell types in the human body and their specialized functions cannot be explained just by the DNA sequence. It is based on the presence of epigenetic signals like DNA methylation, ncRNA and histone tail PTMs, which influence gene expression leading to cell type specific gene expression profiles. Moreover, epigenetic signals are necessary for the establishment and maintenance of the complex chromatin structure. Aberrant epigenetic modifications as well as environmental influences can alter gene expression programs leading to various diseases like cancer or autoimmune disorders (Baylin and Jones 2011; Bergman and Cedar 2013; Suva et al. 2013; Hamidi et al. 2015; Javierre et al. 2010; Butler 2009).

Therefore, it is of utmost importance to study epigenetic signals, their influence on gene expression and their biological roles. In the case of histone tail PTMs another challenge is that, they occur in complex patterns where each combination of individual modifications potentially could influence gene expression differently.

However, the biological meaning of such combinatorial histone PTMs is not well understood so far, which is partially due to a lack of available methods. Therefore, one key aim of this thesis was to fill this methodological gap by the generation and validation of recombinant double reading domains to investigate the genome-wide co-occurrence of histone marks in a highly effective and reproducible way.

4.1 Development of histone modification interacting domains (HiMIDs) as an alternative to histone PTM antibodies

To study the influence of histone tail modifications on chromatin structure, their impact on gene expression and their biological functions, histone PTM specific antibodies are typically used as key research tool. Based on the specific interaction of antibodies with certain histone PTMs, chromatin immunoprecipitation experiments are conducted and with the analysis of the isolated DNA, the genome-wide distribution of each mark can be investigated. Therefore, all downstream analysis and conclusions of ChIP experiments depend on the specific recognition of the target PTM by the antibodies used for the underlying pulldown reaction.

However, the recognition of histone tail PTMs by antibodies is challenging, because they have to discriminate minor structural differences in the target sequence. For example, it is necessary to discriminate efficiently between the different methylation states of lysine or arginine residues as well as the modification state of arginine being either symmetrically or asymmetrically dimethylated.

Additionally, some important modifications are embedded in an identical amino acid sequence like the ARKS motif of H3(7-10) and H3(25-28). Both sequences are methylated at the lysine residue and phosphorylated at the serine residue, but the different lysine and serine modifications have distinct roles and they need to be discriminated. Indeed, some doubts about the reliability of histone tail antibodies arose in recent years (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; Kungulovski et al. 2015b, 2015a, 2014).

One point of concern was, and it was seen in detailed studies, that histone tail interacting antibodies showed cross-reactivity to secondary targets. In addition, the recognition of one modification can be influenced by the presence of a second modification nearby, which can prevent antibody binding. Even more critically, the specificity of polyclonal (but also to a lower degree monoclonal) antibodies has been demonstrated to vary from batch to batch. Hence, a specificity analysis for each individual batch would be necessary. Therefore, it is also not always possible to reproduce published data if one batch of antibody is no longer available. Based on this, long-term studies as well as clinical studies are impeded. Furthermore, histone tail PTMs are highly diverse and not for all identified histone tail modification antibodies are available.

The application of monoclonal antibodies reduces the batch depended variations but their production requires higher levels of technology and monoclonal antibodies are significantly more expensive and more difficult to produce than polyclonal antibodies. Furthermore, practical experience showed, that monoclonal histone PTM antibodies are often less powerful than polyclonal antibodies in chromatin precipitation experiments (Busby et al. 2016).

One more advanced solution to overcome the limitation of mono- and polyclonal antibodies are recombinant antibodies. The recombinant production in *E.coli* enables the unlimited availability of the antibody with constant quality. Therefore, they can be used in long-time applications and reproducibility of results is guaranteed.

Indeed, recently it has been shown, that recombinant histone tail PTMs antibodies can be efficiently optimised to enhance their binding specificity (Hattori et al. 2016). In this publication, highly specific and highly affine recombinant H3K4me3 and H3K9me3 antibodies were generated based on antibody-peptide crystal structure. Nevertheless, not many engineered recombinant histone PTMs antibodies exist so far and their generation is a formidable technical challenge that requires high technological expertise and knowledge.

These problems related to the application of histone PTM antibodies combined with their extremely high importance, stimulated our group to develop new analysis tools for the investigation of chromatin modifications. One alternative approach to antibodies could be based on naturally occurring reading domains, which are part of chromatin interacting proteins and specifically interact with defined histone tail modifications. These domains are capable of interacting with their target modification in a highly specific manner and a previous study from our lab has already documented the successful application of HiMIDs in chromatin immunoprecipitation assays (Kungulovski et al. 2014).

HiMIDs can be easily expressed recombinantly in *E.coli*, which is generally possible in almost every biochemical lab. Furthermore, most of the domains are small, resulting in a high production yield and constant quality of the purified proteins. Furthermore, the unlimited availability of recombinant proteins with constant quality enables their long-time application and guarantees reproducibility of results. Another advantage is that for most of the reading domains crystal structures are available allowing the generation of binding pocket mutants as negative controls. Furthermore, based on the crystal structure, protein engineering is possible to enhance or change the binding preference for histone interacting domains, which is also much easier to achieve than for recombinant antibodies (Kungulovski et al. 2014).
This was demonstrated for the KDM4A double Tudor domain whose binding specificity was specifically altered through systematic mutagenesis. Normally KDM4A double Tudor interacts with pan di- and trimethylated H4 tail modifications, but was changed to a H4K20me2/3 specific reader (Kungulovski et al. 2014). Nevertheless, antibodies in general show a higher binding constant towards their targets in comparison to reading domains, which is a clear advantage in certain applicative settings.

In the context of this work, the already existing portfolio of reading domains was extended for a specific H3K4me3 readout. The H3K4me3 modification is one of the most intensively studied

histone tail modifications and screening for a highly specific H3K4me3 reading domain identified the PHD domain of the TAF3 protein as an appropriate candidate (Kungulovski et al. 2016). In previous studies, it was already shown that the PHD domain of the TAF3 protein interacts specifically with H3K4me3 (Vermeulen et al. 2007).

Biochemical specificity analysis using CelluSpots peptide arrays showed that the interaction of the TAF3 PHD domain with peptides carrying H3K4me3 modifications is highly specific and that the reading domain specifically discriminates between modified and unmodified nucleosomes. The profile of CIDOP-qPCR data as well as the genome-wide analysis of TAF3 PHD chromatin binding (CIDOP-seq) showed the same pattern for a validated H3K4me3 antibody and the TAF3 PHD domain. This include, a high correlation of genome-wide signal intensities between the H3K4me3 antibodies (replicate 1 and 2) and the TAF3 PHD domains (replicate 1 and 2) and a similar signal distribution in promoters and CpG islands. Other studies investigating the biological role of TAF3 also demonstrated a genome-wide correlation of TAF3 PHD domain with H3K4me3 (Lauberth et al. 2013).

Nevertheless, the present work showed successfully, the first time, the application of the TAF3 PHD domain in chromatin precipitation assays replacing an H3K4me3 antibody. All results obtained in this study clearly demonstrate that the TAF3 PHD domain can be used as a reliable tool in chromatin immunoprecipitation assays.

This contribution expands the repertoire of validated HiMIDs helping to overcome the limitations and the drawbacks of polyclonal histone tail interacting antibodies. Many more histone tail PTMs interacting domains exist, which bind highly relevant and important PTMs. These include Bromo domains, which recognize acetylated histone tails. For example the Bromo domain of BAZ2B (Bromo domain adjacent to zinc finger domain 2B), which interacts with H3K14ac, the Bromo domain of BRD4 (Bromo domain-containing protein 4), which binds to H3K36ac or the Bromo domain of CREBBP (CREB binding protein), which interacts with H2AK15ac (Filippakopoulos et al. 2012) can be tested as HiMIDs. In addition, a reader for H3K4me1 was identified in the DPF3 (double PHD fingers 3) zinc finger protein of the BAF complex (Local et al. 2018). The PHD2 domain of DPF3 could be isolated and tested for the specific readout of H3K4me1 modifications.

4.2 Development of double reading domains with dual specificity

The remarkable diversity of histone tail PTMs and the biological specificity resulting from the complex chromatin structure has led to the concept of the histone code (Strahl and Allis 2000). It was proposed that distinct combinations of histone tail modifications can have specific meanings that differ from or go beyond the meanings of the underlying individual modifications. This would massively expand the complexity of the chromatin language leading to a plethora of possible specific downstream effects. Enormous efforts were undertaken to analyze the co-occurrence of histone tail PTMs and study their meanings. For this, different approaches exist and continuously new methods are introduced to decode combinatorial histone tail modifications, which will be described in the next chapters.

However, none of them is ideal and it is still a difficult task to precipitate chromatin bearing two distinct modifications and map its genome-wide distribution. Therefore, the identification of new bivalent chromatin states and investigation of their distribution and potential biological function(s) is very challenging.

Mass spectrometry is a powerful analytical tool to study histone tail PTMs as well as the co-occurrence of histone tail PTMs on the same histone tail (Guan et al. 2013; Britton et al. 2011; Taverna et al. 2007). However, mass spectrometry does not provide information about the genomic locations of histone marks and therefore mapping the genome-wide distribution and the analysis of the biological function of dual marks is not possible with this approach.
To provide genomic information about histone tail PTMs chromatin immunoprecipitation (ChIP) assays are typically used. However, this method cannot effectively distinguish whether particular histone modifications observed at the same genomic locus indeed physically co-occur or whether they originate from distinct places like different alleles or even different cells.

To investigate the direct co-occurrence of histone marks, two consecutive ChIPs (re-ChIP) can be conducted. For this, chromatin is precipitated using an antibody against a particular histone mark and the obtained material is then again precipitated with a second antibody against another histone modification. After performing both ChIPs, nucleosomes bearing two histone tail PTMs at the same time can be recovered and the underlying DNA can be analyzed.

Although this method provides genomic information about dual chromatin states, it is technically very difficult to perform, and large amounts of input material are required because of the low yield of the two-step procedure. Therefore high quality library preparation of NGS remains difficult which affects and limits the downstream analysis (Truax and Greer 2012).

Recently an improved re-ChIP protocol has been published called co-ChIP (Weiner et al. 2016). Using this approach already known but also many new bivalent chromatin states like the co-occurrence of H3K9me1 and H3K27ac at super-enhancers were identified.
To perform co-ChIP, chromatin is isolated and incubated with specific histone tail PTMs antibodies and immobilized on magnetic beads. This is followed by barcoding of the nucleosomes. After this step is performed, the first antibody is inactivated and the barcoded nucleosomes are released from the beads, incubated with a secondary antibody and another barcode is attached. The DNA is sequenced to reveal the genomic loci and the corresponding barcode allows the genome-wide identification of corresponding modification patterns. It is important to mention though, that the specificity of the antibodies has to be validated first in order to remain good results (Weiner et al. 2016).

Another recently developed method uses high-throughput single-molecule imaging to identify dual modified nucleosomes. For this, mononucleosomes are isolated and biotinylated oligonucleotide adaptors are ligated to the free linker DNA ends. This enables the direct coupling of the nucleosomes to streptavidin coated slides. The mononucleosomes are then incubated with different fluorescently labelled antibodies (one after the other) and using TIRF microscopy the position and complex modification state of each mononucleosome can be analyzed. To identify the genomic localisation of the mononucleosomes, the histone proteins are eluted by increasing salt concentrations and the DNA sequence associated with this particular modification pattern can be directly analyzed by Illumina sequencing on the slide (Shema et al. 2016).

Despite their usefulness, all these methods rely on the quality of histone PTMs antibodies and as mentioned before (see section 4.1) several drawbacks associated with histone tail PTMs antibodies were observed in several studies (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; Kungulovski et al. 2015b; Kungulovski and Jeltsch 2015; Kungulovski et al. 2014).

Reading domains occur naturally and were identified to occur in numerous proteins. It was seen that almost all chromatin interacting complexes bear multiple histone modification interacting domains for specific readout, which results in the possibility of a combinatorial readout (Su and Denu 2016; Musselman et al. 2012). Combinatorial readout can take place on one histone tail (*cis* readout) or on different histone tails within one nucleosome (*trans* intranucleosomal) or on adjacent nucleosomes (*trans* internucleosomal). Furthermore, dual readout of DNA modification and histone tail PTM can occur (*trans* DNA: Histone) (Fig. 41).

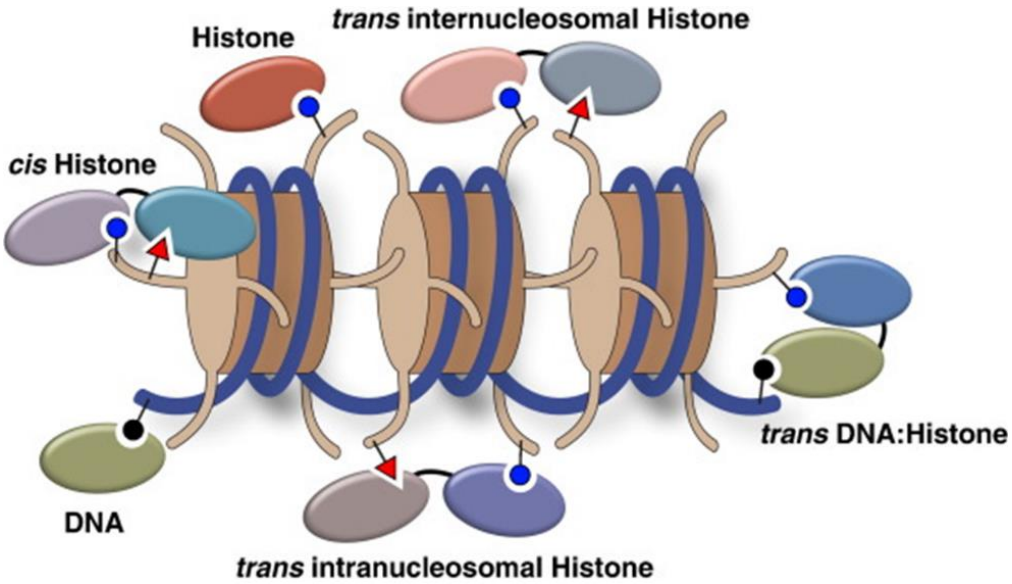


Figure 41. Readout of epigenetic modifications. Single chromatin interacting domains can interact either with one single histone PTMs or with DNA methylation. Double reading domains bind specifically to two histone tail PTMs on one histone tail (*cis* readout) or on two adjacent histone tails (*trans* readout, same nucleosome *trans* internucleosomal, adjacent nucleosomes *trans* intranucleosomal). Furthermore, chromatin interacting domains can bind to histone PTMs and DNA methylation (*trans* DNA: Histone). Figure taken from (Rothbart and Strahl 2014).

Inspired by this, and the fact that reading domains can be easily engineered, the approach of using HiMIDs as alternative to antibodies (Kungulovski et al. 2016, 2014) was further developed to achieve simultaneously readout of two histone PTMs at one nucleosome by an artificial double reading domain. Therefore, the main goal of this doctoral thesis was the generation and validation of double reading domains to achieve simultaneous readout of combinatorial histone PTMs. Followed by this, the identification of new bivalent chromatin states and their genomic distribution and investigation of their function was aimed to gain a deeper insight into the complex chromatin language.

To achieve this goal, two validated reading domains such as MPP8 Chromo domain, Dnmt3a PWWP domain, CBX7 Chromo domain, KDM4 double Tudor domain, ATRX ADD domain or

TAF3 PHD domain were fused together in different combinations including a flexible linker. In this respect, it is most important that the chromatin interaction of the double domains should only lead to an efficient pulldown if both modifications were present, such that the double domains allow the identification of double modified nucleosomes and chromatin regions and not of region that carry only one of the modifications.

The criteria established for histone tail antibody validation as summarised by the ENCODE project were applied for the validation of double reading domains as well (Consortium 2012). These quality criteria include the specific binding to modified peptides, which was analyzed using CelluSpots peptide arrays. Furthermore, they include the analysis of the specific binding to modified nucleosomes, which was tested using Far-western blot assays with native nucleosomes and unmodified recombinant histones. Additionally, the potential to precipitate nucleosomes in chromatin precipitation experiments in a reproducible fashion was tested. The double specificity was validated using mixed peptide spots (a novel method developed during this thesis) and by the analysis of genome-wide modification patterns and comparison with the distributions of the single marks and with sequential CIDOP-ChIP results. Finally, with established double domains the genome-wide investigation of the newly identified bivalent chromatin states was analyzed and linked to specific functions and genome elements.

H3K9me3 is known to be a repressive mark but it was also found to co-occur with RNAPII and H3K9ac modification on active promoters (Squazzo et al. 2006; Wiencke et al. 2008) and on transcribed regions (Vakoc et al. 2005; Miao and Natarajan 2005). The co-occurrence of H3K9me3 and H3K36me3 histone tail modifications has been observed in mass spectrometry studies (Kolasinska-Zwierz et al. 2009; Saint-André et al. 2011). Using the new double domain analysis tool the co-occurrence of these two histone tail modifications was verified (Mauser et al. 2017).

The double reading domain used in this study comprised of the Dnmt3a PWWP (P) domain and the MPP8 Chromo (M) domain and the specific interaction with H3K36me2/3 and H3K9me3 modifications was shown using CelluSpots peptide array and Far-western blot assays. Additionally, the specific interaction with mononucleosomes bearing H3K9me3 and H3K36me2/3 modifications was demonstrated with pull down experiments analyzed via western blot. Using H3K9me3 and H3K36me3 specific antibodies, a clear signal was obtain in the precipitated nucleosomes, whereas after probing the samples with unrelated H3K4me3

and H3K27me3 antibodies no signal was observed. Furthermore, the specificity of the interaction was further confirmed using inactive binding pocket mutants as negative controls in all experiments. Moreover, within the scope of this work, a new method was developed to investigate the dual specificity of the double reading domains using mixed peptide arrays. With this approach, a stronger interaction of the double domains with mixed peptide spots containing both modified peptides was observed when compared to single modified or unmodified peptides also confirming the desired dual readout.

After successful biochemical validation of the new double reading domain, PM as well as the respective binding pocket mutants were tested in CIDOP-qPCR experiments and the double reading domain PM showed clear preferences for genomic regions associated with both H3K9me3 and H3K36me2/3 histone tail modifications in comparison to genomic regions associated with only one of the modifications. The coexistence of both modifications in the pulldown material of the double reading domain was also confirmed in sequential CIDOP-ChIP experiments.

The comparison of the genome-wide distribution pattern of the double reading domain PM with the corresponding mutants identified numerous genome regions where these two histone PTMs occur together. The identified genome regions were analyzed for specific biological functions, which led to the new finding that H3K9me3 and H3K36me2/3 form a new bivalent chromatin state, which is associated with weakly transcribed genes and weak enhancers.

The genes associated with H3K9me3 and H3K36me2/3 modifications are involved in cell cycle regulation as well as in metabolic and signalling pathways. This observation was verified by analysing chromosome 19, where many zinc finger (ZNF) genes are located and a clear co-localisation of ZFN274 with H3K9me3-H3K36me2/3 signals was detected. A previous study found that the 3' end of ZNF exons are enriched for H3K9me3 and H3K36me2/3 (Blahnik et al. 2011) and based on direct evidence these authors concluded that both modifications co-occur. This conclusion is strongly supported by the recent findings with the PM double domain.

The data obtained with the double reading domain PM demonstrate the power of this newly developed analysis tool and its successful application in chromatin precipitation assays. With the newly generated double reading domains, it is possible to precipitate mononucleosomes

carrying two defined modifications in one highly specific step helping to decipher the complex histone code and better understand chromatin templated processes.

4.2.1 Development of double reading domain with dual specificity for H3K4me3 and H3K27me3 modifications

Another well-studied bivalent chromatin state is the co-occurrence of H3K4me3 and H3K27me3 in ESC (Vastenhouw and Schier 2012; Voigt et al. 2013). To develop a double domain that can be used to investigate this bivalent modification state, the CBX7 Chromo (C) domain and the TAF3 PHD (T) domain were fused together to precipitate mononucleosomes carrying H3K27me3 and H3K4me3 simultaneously. The new double reading domain CLT retained the binding specificity of each individual domain as indicated by the CelluSpots peptide array where binding to both modifications was observed.

CLT also bound stronger to native nucleosomes than to recombinant histones indicating its binding preference for modified histones. Despite this, no preference for genome regions associated with both modifications was observed in CIDOP-qPCR, where it was seen that the TAF3 PHD domain did not efficiently contribute to the interaction with nucleosomes in the context of the fusion construct. This was surprising, because the TAF3 PHD domain alone was already shown to work optimally in chromatin immunoprecipitation assays and that the domain is fully comparable with a validated H3K4me3 antibody (Kungulovski et al. 2016). Therefore, it was likely that the TAF3 PHD domain is sterically hindered in this particular construct.

To improve the performance of the CBX7/TAF3 double reading domains the domain order was swapped and a TLC double domain was cloned and purified. The binding preference of CLT and TLC were analyzed with mixed peptide arrays, showing that CLT binds with equal strength to H3K27me3 and H3K27me3/H3K4me3 containing peptides regardless of the second modifications as observed before. In contrast, TLC interacted preferentially with mixed peptides spots containing peptides with H3K27me3 and H3K4me3 modifications. Afterwards, the optimized TLC double reading domain was used in CIDOP-qPCR clearly indicating the preference of TLC for binding to genome regions associated with both modifications. These data illustrate, that the orientation of the domains in the context of the fusion protein plays an important role and therefore it is necessary to investigate the influence of the orientation

on the binding specificity of each newly generated double reading domain. At the same time, this result indicates that the functionality of double reading domains can be improved by protein design.

Subsequently, binding pocket mutants of TLC should be generated and all constructs should be used in CIDOP-seq experiments. Promoters of ESC were shown to carry the specific bivalent H3K4me3-H3K27me3 chromatin state and with the optimized double TLC reading domain and the corresponding binding pocket mutants it should be possible to identify these promoter regions using mononucleosomes isolated from ESC. Bivalent chromatin states, especially in ESC, lead to fine-tuned expression of developmental genes and extend the histone language to enable the highly specific regulation of the chromatin network. The newly developed double reading domain tool will help to further understand this specific aspect of chromatin regulation.

4.2.2 Development of double reading domain with dual specificity for H3K4me3 and H3K36me2/3 modifications

The H3K4me3 and H3K36me3 histone tail modifications were shown to be enriched at actively transcribed genes (Young et al. 2011). However, they are usually considered to be associated with different elements, promoters in the case of H3K4me3 and gene bodies in case of H3K36me3. To investigate the potential co-occurrence of these two modifications, the double reading domain PT (Dnmt3a PWWP-TAF3 PHD) was generated.

PT showed specific binding to the individual histone PTMs in CelluSpots array and to modified histones in Far-western blot experiments. Genome-wide investigation of the DNA precipitated by PT double domain indeed identified genome regions associated with both modifications, a potentially very important finding.
However, also genome regions were isolated, which were only associated with H3K36me2/3 modifications, indicating that the PT double domain is not yet fully functional. Nevertheless, this result indicates that dual regions marked by H3K36me2/3 and H3K4me3 do exist, but for their further investigation, the PT double domain needs to be optimized.

First, with the newly developed mixed peptide arrays the specificity of PT has to be improved. For example, the orientation of the domains can be altered as well as the linker length and sequence. Afterwards, the concentration of the double domain has to be optimized in CIDOP

experiments and more stringent conditions have to be applied in order to enhance specific binding of nucleosomes associated with H3K36me2/3 and H3K4me3 modification.

4.2.3 Outlook for the generation of additional double reading domains.

Preliminary biochemical data were obtained for some additional double reading domains, viz. PA (Dnmt3a PWWP-ATRX ADD), MLdT (MPP8 Chromo domain-Linker-double Tudor KDM4A) and CLM (CBX7 Chromo domain-Linker-MPP8 Chromo domain). All these double domains showed the expected and specific binding in CelluSpots peptide arrays to individual modified peptides and exclusive binding to native nucleosomes. PA was designed as an alternative for PM however, as PM showed convincing results from the beginning on, PA was not further investigated. Next steps would be to confirm the dual specificity with mixed peptide arrays and optimize the double domain if necessary.

In the case of MLdT CIDOP-qPCR experiments could not be conducted, because it turned out to be extremely difficult to find target regions with only H3K9me3, only H4K20me3 or both marks because of the close connection of these two marks. Therefore, the dual specificity towards H3K9me3 and H4K20me3 has to be confirmed with mixed peptide arrays. After dual specificity confirmation, the double reading domain can be used in CIDOP-seq experiments to analyze the co-occurrence of H3K9me3 and H4K20me3 in cells.
In these studies chromatin from SUV39H1/H2 and SUV420H1/H2 double knock out cell lines can be used as control, which are selectively lacking the heterochromatic H3K9me3 and H4K20me3 marks, respectively.
The double reading domain CLM failed in first CIDOP-qPCR experiments, because no preferred binding to genome regions associated with both modifications was observed. Therefore, the specific binding to peptides carrying both modifications has to be optimized using the mixed peptide arrays. Based on the published and these preliminary data provided in this study it is anticipated that many more double reading domains can be established, which will play an important role in the further development of chromatin research.

As an alternative, to study the co-occurrence of histone modification is the usage of bispecific antibodies (Kontermann and Brinkmann 2015), which recognize simultaneously two different epitopes, is conceivable, but this approach has not been tested so far. However, in this case

the specificity of double modified targets as opposed to an interaction with target containing either one or the other modification also needs to be established carefully using similar methods as applied here for the development of double reading domains.

4.2.4 Investigation of homotypic double reading domains

Another feature, which can extend the complexity of the histone code, is the question whether nucleosomes are symmetrically or asymmetrically modified. Previous studies found that histone tails can be modified asymmetrically with H3K27me2/3 and H4K20me1 and the symmetric or asymmetric modification with H3K27me3 differently influenced downstream events. For example, it was observed that the presence of asymmetric H3K27me3 stimulates the activity of PRC2, whereas symmetric H3K27me3 inhibits the binding of PRC2 to these nucleosomes (van Rossum et al. 2012; Voigt et al. 2012).

In different studies, it has been shown, that multivalent binding of chromatin binding proteins to chromatin results in dramatic affinity enhancement due to cooperative binding and additionally specificity while complex binding remain dynamic and susceptible to competition despite the strong interaction (Ruthenburg et al. 2007). For example, the dissociation constants of single Bromo domains are higher (K_d of 100 - 300 µM) than the dissociation constant of double Bromo domains (K_d of 1 - 20 µM), illustrating the fact that multidentate binding lead to a tighter and more specific interaction (Zhou et al. 1999; Mujtaba et al. 2004).

To investigate the presence of symmetrically or asymmetrically modified nucleosomes in chromatin, homotypic double reading domains containing either two MPP8 Chromo domains (MLM), or two Dnmt3a PWWP domains (PLP) were generated. Both homotypic double reading domains showed the expected binding to peptides harbouring H3K9me3 or H3K36me2/3, respectively. Also exclusive binding to native nucleosomes was observed. Strikingly both homotypic double reading domains showed a stronger affinity to chromatin than the corresponding single domains although the single domains were used in the same concentrations. It is likely that the dual readout enhances the binding affinity and specificity towards double modified chromatin, suggesting the presence of large quantities of symmetrically modified nucleosomes of both types.

Despite the initial promising results with CelluSpots, Far-western and pull down experiments, PLP failed in CIDOP-qPCR due to unspecific binding of related and unrelated genome regions.

By contrast, MLM and MPP8 Chromo domain showed the expected preferred interaction with genomic regions associated with H3K9me3 modifications. It was expected that the MPP8 Chromo domain binds to nucleosomes carrying H3K9me3 modifications regardless if they are symmetrically or asymmetrically modified, whereas MLM was expected to specifically interact with symmetric modified nucleosomes due to the higher affinity resulting from both domains. The binding of MLM and MPP8 Chromo domain, was investigated on genome-wide level to identify genome regions where MLM and MPP8 Chromo domain show explicit differences, which would hint towards differently modified (symmetrically or asymmetrically) nucleosomes. Unfortunately, no difference in the genome-wide distribution of MLM and MPP8 Chromo domain was observed.

It could be possible that on nucleosomal level only one domain of MLM was interacting with H3K9me3 modifications, which would explain the lack of any detectable difference with MPP8 binding. To avoid this potential problem, the binding stringency has to be increased in order to create conditions where only the double reading domain is able to precipitate any nucleosomes. Furthermore, it is possible, that the known dimerization of the GST-tag was strong enough to create an artificial double domain from the single domain MPP8. To exclude this, the GST-tag should be exchanged by another tag, which does not dimerize such as the MBP-tag or His-tag. Finally, it is also possible, that asymmetrically H3K9me3 modified nucleosomes do not exists in cells, because no information regarding this is available in literature so far.

4.2.4 Investigation of double reading domains with dual specificity for DNA methylation and histone marks

Histone modifications do not operate alone to regulate the chromatin structure, but they act together and in conjunction with DNA methylation. It has been shown in several studies that histone tail modifications and DNA methylation are bound simultaneously by chromatin interacting proteins influencing the gene expression (Rothbart and Strahl 2014; Bartke et al. 2010).

For example, H3K27me3 was found to co-occur with 5-methylcytosine on silent gene loci, whereas H3K9me3, H4K20me3 and 5-methylcytosine co-exist in heterochromatin (Bernstein et al. 2006a; Sims et al. 2006; Mikkelsen et al. 2007). Furthermore, the DNA methyltransferase

Dnmt3a was shown to interact with the H3 tail preferentially if K36 is trimethylated and H3K4 is unmodified also connecting DNA methylation with histone PTMs (Dhayalan et al. 2010; Zhang et al. 2010; Otani et al. 2009).

In this work, the new approach of using double reading domains with dual specificity against two histone PTMs was extended to investigate co-occurrence of DNA methylation and histone tail PTMs. Therefore, the methyl DNA binding domain of the MBD2 protein was fused to several histone interacting domains. The ability of the newly designed double reading domains to interact with methylated DNA was analyzed using dot blot assays showing that all generated double reading domains specifically interacted with methylated DNA.

However, Far-western blot assays showed binding of all domains to native as well as to recombinant histones. It was likely that this unspecific histone binding originated from the MBD2 domain and this hypothesis was confirmed in additional control experiments. Moreover, the investigation of the binding specificity of the histone tail interacting domains in these fusions constructs using CelluSpots arrays failed almost in all cases. Only CBX7-L-MBD2 showed to some extent the expected specificity profile. However, CIDOP-qPCR experiments with this double reading domain showed a dominating influence of the MBD2 domain on the binding specificity. In summary, all these results suggested that the MBD2 domain is not working in combination with other histone PTM reading domains.

To overcome these difficulties, the methyl DNA binding domain of the MBD1 protein was used in a second round of double domain design. As expected from literature data (Baubec et al. 2013), the MBD1 domain showed a weaker binding to methylated DNA compared to MBD2. The DNA binding specificity of the new double reading domains containing the MBD1 domain was investigated in dot blot assays showing specific interaction to methylated DNA for all generated double reading domains.

Interestingly, the double reading domain containing the PHD domain of TAF3 also showed binding to unmethylated DNA, which has not been reported so far. Based on this finding, it could be possible that the interaction of TAF3 PHD domain with histones carrying H3K4me3 modifications is affected by the surrounding DNA, and it will be interesting to study this effect further for example using designer nucleosomes containing the H3K4me3 modification.

Unfortunately, most of the MBD1 containing double reading domains still showed unspecific peptide binding in CelluSpots arrays. This unspecific binding could result either from an influence of the DNA binding domain or from steric problems related to the domain fusion, which may disrupt the correct function of the reading domains. This problem can be addressed in future study by changing the domain order or increasing the size of the linker between the two domains. So far, only the MBD1-L-MPP8, ATRX-L-MBD1 and MBD1-L-PWWP double domains passed to some extent the first quality steps and were investigated in CIDOP-qPCR experiments.

It is known that DNA methylation and H3K9me3 modifications at promoter lead to repressed genes (Rose and Klose 2014). Therefore, a silenced promoter associated with DNA methylation and H3K9me3 modification (ZNF420 promoter region) was analyzed for the pull down with MBD1-L-MPP8 and its individual single domains. This experiment resulted in the highest signal for the double reading domain. The same trend was observed for the double reading domain ATRX-L-MBD1, were the same genome region enriched for DNA methylation and H3K9me3 modification was analyzed.
DNA methylation as well as H3K36me3 modification in gene bodies is correlated with actively transcribed genes (Lorincz et al. 2004; Sun et al. 2005) and, therefore, the VEGF-A gene body was analyzed for MBD1-L-PWWP in CIDOP-qPCR experiments showing a stronger binding to the target region than its individual single domains.

These results indicate that the combination of the MBD1 DNA binding domain and a histone tail PTM interacting domain leads to an improved binding to the specific target chromatin sites suggesting that with these combinations the co-occurrence of DNA methylation and histone tail PTMs can be investigated. In the future, another generation of double domain could be designed, which are able to interact with unmethylated DNA in context of certain histone tail modifications. For this, a specific reading domain for unmethylated DNA like the CXXC domain of MLL or Dnmt1 (Cierpicki et al. 2010; Pradhan et al. 2008) could be used as fusion partner.

4.3 Conclusions and perspectives

The results of the present study demonstrate the successful application of double reading domains in chromatin analysis. Double domains comprise two fused single domains generating a dual specificity for two histone tail modifications or for DNA methylation together with histone mark. It has been shown in this work that many of these artificial double reading domains specifically interact with chromatin modifications. However, there is still a lot of optimisation work to be done for each individual double reading domain, but this new approach represents an important step towards a better understanding of how chromatin modifications act together and influence the chromatin network.

There are several parameters, which can be improved in further design. For example, the linker length and the linker sequence can be varied, salt concentrations of the used buffers in the different application have to be optimised, the concentrations of domains in the pull down experiments have to be adapted to create specific binding events of the double reading domains and the orientation of domains has to be optimized.

Naturally occurring linkers between reading domains in chromatin binding proteins show a perfect example how the linker sequence and length influence the binding specificity. For example the double Bromo domains of TAF1 and the tandem Bromo domain of Rsc4p showed different binding specificities related to differences in the length of the linker in the particular protein (Taverna et al. 2007). Furthermore, the linker length also influences the readout in *cis* or *trans*, especially *trans* internucleosomal or *trans* intranucleosomal. Therefore, linker optimization could be used to achieve specific readout of *cis* or *trans* occurring double modifications.

The data obtained in this work illustrate how powerful the new method is in order to analyze co-occurrence of histone tail modifications on one nucleosome. One of its important advantages is the simultaneous precipitation of double modified nucleosomes or chromatin in one step allowing the usage of less input material in comparison to re-ChIP experiments and still generate good quality libraries for next generation sequencing. The possibility to engineer the reading domains in many different ways make them versatile tool to analyze the co-occurrence of histone tail modifications. Furthermore, the new double domain tool might help to discover new not yet identified combinations of epigenome modifications and helps to decipher the complex histone code and gain deeper insights into epigenetic processes. In

this context, it is even conceivable to fuse more than two domains together to even extend the complexity of epigenome modification readout.

Publications

1. **Mauser R,** Kungulovski G, Meral D, Maisch D and Jeltsch A (2018) **„Application of mixed peptide arrays to study combinatorial readout of chromatin modifications".** *Biochimie,* doi: 10.1016/j.biochi.2017.11.008

2. **Mauser R*, Kungulovski G*,** Keup C, Reinhardt R and Jeltsch A (2017) **„Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state".** *Epigenetics & Chromatin,* doi: 10.1186/s13072-017-0153-1

3. Imre L, Simándi Z, Horváth A, Fenyőfalvi G, Nánási P, Niaki E, Hegedüs E, Bacsó Z, Weyemi U, Mauser R, Ausio J, Jeltsch A, Bonner W, Nagy L, Kimura H and Szabó G (2017) **„Nucleosome stability measured in situ by automated quantitative imaging".** *Scientific reports,* doi: 10.1038/s41598-017-12608-9

4. **Kungulovski G*, Mauser R*,** Reinhardt R and Jeltsch A (2016) **„Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody".** *Epigenetics & Chromatin,* doi: 10.1186/s13072-016-0061-9

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.

6. **Kungulovski G,** Kycia I, 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.

7. **Kungulovski G,** Mauser R, Jeltsch A. (2015) **„Affinity reagents for studying histone modifications and guidelines for their quality control".** *Epigenomics,* doi: 10.12688/f1000research.726

   * shared first authors

Author´s contribution

- **Mauser R,** Kungulovski G, Meral D, Maisch D and Jeltsch A (2018) **„Application of mixed peptide arrays to study combinatorial readout of chromatin modifications".** *Biochimie,* 146:14-19. doi: 10.1016/j.biochi.2017.11.008. Epub 2017 Nov 11.

RM and AJ designed the experiments. RM with help of DMe and DMa conducted the experiments. RM and AJ analyzed the data and interpreted the results. RM and AJ wrote the manuscript draft.

- **Mauser R,** Kungulovski G**,** Keup C, Reinhardt R and Jeltsch A (2017) **„Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state".** *Epigenetics & Chromatin,* 25;10(1):45. doi: 10.1186/s13072-017-0153-1.

RM, GK and AJ conceived and designed the experiments. RM performed the experiments with contributions from CK and GK. RM, GK and AJ were involved in data analysis and interpretation. RM contributed to the writing of the manuscript.

- Imre L, Simándi Z, Horváth A, Fenyőfalvi G, Nánási P, Niaki E, Hegedüs E, Bacsó Z, Weyemi U, **Mauser R**, Ausio J, Jeltsch A, Bonner W, Nagy L, Kimura H and Szabó G (2017) **„Nucleosome stability measured in situ by automated quantitative imaging".** *Scientific reports,* 7(1):12734. doi: 10.1038/s41598-017-12608-9.

RM cloned and purified the TAF3 PHD domain.

- Kungulovski G, **Mauser R,** Reinhardt R and Jeltsch A (2016) **„Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody".** *Epigenetics & Chromatin,* 9:11. doi: 10.1186/s13072-016-0061-9. eCollection 2016.

AJ, GK and RM conceived and designed the experiments. RM and GK performed the experiments. AJ, GK and RM were involved in data analysis and interpretation. RM contributed to the writing of the manuscript.

- Kungulovski G**,** Kycia I, **Mauser R**, Jeltsch A. (2015) **„Specificity Analysis of Histone Modification Specific Antibodies or Reading Domains on Histone Peptide Arrays".** *Methods Mol Biol*., 1348:275-84. doi: 10.1007/978-1-4939-2999-3_24.

RM contributed to the writing of the manuscript.

- Kungulovski G, **Mauser R**, Jeltsch A. (2015) **„Affinity reagents for studying histone modifications and guidelines for their quality control".** *Epigenomics,* 7(7):1185-96. doi: 10.2217/epi.15.59. Epub 2015 Nov 6.

RM contributed to the writing of the manuscript.

Acknowledgements

I would like to thank everybody who has supported me during my studies and in completing this work.

First of all, I would like to thank my supervisor, Prof. Dr. Albert Jeltsch, who offered me the opportunity to be a part of this working group and I am also thankful for his support, excellent guidance and many great suggestions during my entire work and also during the writing process.

I am also thankful to Prof Dr. Roland Kontermann, who has kindly agreed to be the second co-referee for my PhD Thesis.

I would also like to thank Prof. Dr.-Ing Ralf Takors for being the leader of the committee, Prof Dr. Dieter H. Wolf, Prof. Dr. Stephan Nußberger and Prof. Dr. Monilola Olayioye for accepting to read and review my PhD Thesis.

I am very thankful to Dr. Goran Kungulovski for his supervision, his guidance and support during my Master study and in the first time of my PhD study. He was never tired of talking about science and I learned a lot from him.

I would like to thank all lab members, especially Miru, Sara, Cristiana, Raluca, Aga, Max, Goran, Denis, Micha, Alex and Julian for creating a very welcoming and warm environment at the workplace, for all the discussions, not always science related, and for the great fun we had always together. You turned the daily work to something special.

I would like to thank my students Corinna, Mareike, Tesmin and David, which contributed with their work to my studies.

Moreover, special thanks to Cristiana Lungu for the discussions, her suggestions and for her support during the writing process. Furthermore, I would like to thank Maren and my sister Lea for their support during the writing process.

Last but not least, I am very grateful to all my friends, my boyfriend and my family, especially to my parents, for all their love and permanent support.

References

Allfrey VG, Faulkner R, Mirsky A E. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci U S A* **51**: 786–94. http://www.ncbi.nlm.nih.gov/pubmed/14172992 (Accessed January 19, 2018).

Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M. 2008. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455**: 818–821. http://www.ncbi.nlm.nih.gov/pubmed/18772891 (Accessed January 24, 2018).

Arnaudo AM, Garcia BA. 2013. Proteomic characterization of novel histone post-translational modifications. *Epigenetics Chromatin* **6**: 24. http://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/1756-8935-6-24 (Accessed February 26, 2018).

Badenhorst P, Voas M, Rebay I, Wu C. 2002. Biological functions of the ISWI chromatin remodeling complex NURF. *Genes Dev* **16**: 3186–3198. http://www.ncbi.nlm.nih.gov/pubmed/12502740.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381–395. http://www.ncbi.nlm.nih.gov/pubmed/21321607.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120–124. http://www.ncbi.nlm.nih.gov/pubmed/11242054 (Accessed January 18, 2018).

Barski A, Cuddapah S, Cui K, Roh T-YY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823–837. http://www.ncbi.nlm.nih.gov/pubmed/17512414 (Accessed January 18, 2018).

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. 2010. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143**: 470–484. http://www.ncbi.nlm.nih.gov/pubmed/21029866.

Basu A, Rose KL, Zhang J, Beavis RC, Ueberheide B, Garcia BA, Chait B, Zhao Y, Hunt DF, Segal E, et al. 2009. Proteome-wide prediction of acetylation substrates. *Proc Natl Acad Sci* **106**: 13785–13790. http://www.ncbi.nlm.nih.gov/pubmed/19666589 (Accessed January 18, 2018).

Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schubeler D. 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**: 243–247. http://www.ncbi.nlm.nih.gov/pubmed/25607372.

Baubec T, Ivánek R, Lienert F, Schübeler D. 2013. Methylation-Dependent and -Independent Genomic Targeting Principles of the MBD Protein Family. *Cell* **153**: 480–492. https://www.sciencedirect.com/science/article/pii/S0092867413003334 (Accessed March 20, 2018).

Baylin SB, Jones PA. 2011. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer* **11**: 726–34. http://www.ncbi.nlm.nih.gov/pubmed/21941284 (Accessed January 18, 2018).

Becker JS, Nicetto D, Zaret KS. 2016. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet* **32**: 29–41. http://www.ncbi.nlm.nih.gov/pubmed/26675384.

Bergman Y, Cedar H. 2013. DNA methylation dynamics in health and disease. *Nat Struct Mol Biol* **20**: 274–281. http://www.ncbi.nlm.nih.gov/pubmed/23463312 (Accessed January 18, 2018).

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006a. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**: 315–326. http://www.ncbi.nlm.nih.gov/pubmed/16630819 (Accessed February 6, 2018).

Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD. 2006b. Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol Cell Biol* **26**: 2560–2569. http://www.ncbi.nlm.nih.gov/pubmed/16537902 (Accessed January 23, 2018).

Bird A, Ng H-H, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* **23**: 58–61. http://www.nature.com/doifinder/10.1038/12659 (Accessed February 15, 2018).

Blahnik KR, Dou L, Echipare L, Iyengar S, O'Geen H, Sanchez E, Zhao Y, Marra MA, Hirst M, Costello JF, et al. 2011. Characterization of the contradictory chromatin signatures at the 3' exons of zinc finger genes. *PLoS One* **6**: e17121. http://www.ncbi.nlm.nih.gov/pubmed/21347206.

Blake GE, Watson ED. 2016. Unravelling the complex mechanisms of transgenerational epigenetic inheritance. *Curr Opin Chem Biol* **33**: 101–107. https://www.sciencedirect.com/science/article/pii/S1367593116300849?via%3Dihub (Accessed March 2, 2018).

Bock I, Dhayalan A, Kudithipudi S, Brandt O, Rathert P, Jeltsch A. 2011a. Detailed specificity analysis of antibodies binding to modified histone tails with peptide arrays. *Epigenetics* **6**: 256–263. http://www.ncbi.nlm.nih.gov/pubmed/20962581.

Bock I, Kudithipudi S, Tamas R, Kungulovski G, Dhayalan A, Jeltsch A. 2011b. Application of Celluspots peptide arrays for the analysis of the binding specificity of epigenetic reading domains to modified histone tails. *BMC Biochem* **12**: 48. http://www.ncbi.nlm.nih.gov/pubmed/21884582.

Bonasio R, Tu S, Reinberg D. 2010. Molecular signals of epigenetic states. *Science* **330**: 612–6. http://www.ncbi.nlm.nih.gov/pubmed/21030644 (Accessed March 1, 2018).

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, et al. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**: 349–353. http://www.ncbi.nlm.nih.gov/pubmed/16625203 (Accessed January 18, 2018).

Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. 2006. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* **20**: 1123–1136. http://www.ncbi.nlm.nih.gov/pubmed/16618801 (Accessed January 18, 2018).

Britton L-MP, Gonzales-Cope M, Zee BM, Garcia BA. 2011. Breaking the histone code with quantitative mass spectrometry. *Expert Rev Proteomics* **8**: 631–643. http://www.ncbi.nlm.nih.gov/pubmed/21999833 (Accessed January 19, 2018).

Bua DJ, Kuo AJ, Cheung P, Liu CL, Migliori V, Espejo A, Casadio F, Bassi C, Amati B, Bedford MT, et al. 2009. Epigenome Microarray Platform for Proteome-Wide Dissection of Chromatin-Signaling Networks ed. A. Imhof. *PLoS One* **4**: e6789. http://dx.plos.org/10.1371/journal.pone.0006789 (Accessed January 19, 2018).

Busby M, Xue C, Li C, Farjoun Y, Gienger E, Yofe I, Gladden A, Epstein CB, Cornett EM, Rothbart SB, et al. 2016. Systematic comparison of monoclonal versus polyclonal antibodies for mapping histone modifications by ChIP-seq. *Epigenetics Chromatin* **9**: 49. http://www.ncbi.nlm.nih.gov/pubmed/27826357 (Accessed March 16, 2018).

Butler MG. 2009. Genomic imprinting disorders in humans: a mini-review. *J Assist Reprod Genet* **26**: 477–86. http://www.ncbi.nlm.nih.gov/pubmed/19844787 (Accessed January 18, 2018).

Calo E, Wysocka J. 2013. Modification of Enhancer Chromatin: What, How, and Why? *Mol Cell* **49**: 825–837. https://www.sciencedirect.com/science/article/pii/S1097276513001020 (Accessed January 18, 2018).

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581–592. http://www.ncbi.nlm.nih.gov/pubmed/16286007.

Chang Y, Horton JR, Bedford MT, Zhang X, Cheng X. 2011. Structural insights for MPP8 chromodomain interaction with histone H3 lysine 9: potential effect of phosphorylation on methyl-lysine binding. *J Mol Biol* **408**: 807–814. http://www.ncbi.nlm.nih.gov/pubmed/21419134.

Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng J, Gu W, Zhao Y. 2007. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* **6**: 812–9. http://www.ncbi.nlm.nih.gov/pubmed/17267393 (Accessed February 26, 2018).

Chi P, Allis CD, Wang GG. 2010. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer* **10**: 457–69. http://www.ncbi.nlm.nih.gov/pubmed/20574448 (Accessed March 2, 2018).

Chiu ML, Gilliland GL. 2016. Engineering antibody therapeutics. *Curr Opin Struct Biol* **38**: 163–173. https://www.sciencedirect.com/science/article/pii/S0959440X16300872#bib0525 (Accessed March 16, 2018).

Cierpicki T, Risner LE, Grembecka J, Lukasik SM, Popovic R, Omonkowska M, Shultis DD, Zeleznik-Le NJ, Bushweller JH. 2010. Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. *Nat Struct Mol Biol* **17**: 62–8. http://www.ncbi.nlm.nih.gov/pubmed/20010842 (Accessed March 20, 2018).

Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74. http://www.ncbi.nlm.nih.gov/pubmed/22955616.

Cooper S, Dienstbier M, Hassan R, Schermelleh L, Sharif J, Blackledge NP, De Marco V, Elderkin S, Koseki H, Klose R, et al. 2014. Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep* **7**: 1456–1470. http://www.ncbi.nlm.nih.gov/pubmed/24857660 (Accessed January 18, 2018).

Cosgrove MS, Boeke JD, Wolberger C. 2004. Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* **11**: 1037–1043. http://www.nature.com/articles/nsmb851 (Accessed February 26, 2018).

Crisp SJ, Kullmann DM, Vincent A. 2016. Autoimmune synaptopathies. *Nat Rev Neurosci* **17**: 103–117. http://www.nature.com/articles/nrn.2015.27 (Accessed March 16, 2018).

Dawson PE, Muir TW, Clark-Lewis I, Kent SB. 1994. Synthesis of proteins by native chemical ligation. *Science* **266**: 776–9. http://www.ncbi.nlm.nih.gov/pubmed/7973629 (Accessed January 19, 2018).

Deaton AM, Bird A. 2011. CpG islands and the regulation of transcription. *Genes Dev* **25**: 1010–1022. http://www.ncbi.nlm.nih.gov/pubmed/21576262 (Accessed January 23, 2018).

Delachat AM-F, Guidotti N, Bachmann AL, Meireles-Filho ACA, Pick H, Lechner CC, Deluz C, Deplancke B, Suter DM, Fierz B. 2018. Engineered Multivalent Sensors to Detect Coexisting Histone Modifications in Living Stem Cells. *Cell Chem Biol* **25**: 51–56.e6. https://www.sciencedirect.com/science/article/pii/S2451945617303896#undfig1 (Accessed March 8, 2018).

Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A. 2010. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* **285**: 26114–26120. http://www.ncbi.nlm.nih.gov/pubmed/20547484 (Accessed January 18, 2018).

Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, Ragozin S, Jeltsch A. 2011. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum Mol Genet* **20**: 2195–2203. http://www.ncbi.nlm.nih.gov/pubmed/21421568.

Di Cerbo V, Mohn F, Ryan DP, Montellier E, Kacem S, Tropberger P, Kallis E, Holzner M, Hoerner L, Feldmann A, et al. 2014. Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. *Elife* **3**: e01632. http://www.ncbi.nlm.nih.gov/pubmed/24668167 (Accessed January 18, 2018).

Dikmans A, Beutling U, Schmeisser E, Thiele S, Frank R. 2006. SC2: A Novel Process for Manufacturing Multipurpose High-Density Chemical Microarrays. *QSAR Comb Sci* **25**: 1069–1080. http://doi.wiley.com/10.1002/qsar.200640130 (Accessed January 19, 2018).

Dupont C, Armant DR, Brenner CA. 2009. Epigenetics: definition, mechanisms and clinical perspective. *Semin Reprod Med* **27**: 351–7. http://www.ncbi.nlm.nih.gov/pubmed/19711245 (Accessed January 18, 2018).

Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung M-S, Day DS, Gadel S, Gorchakov AA, et al. 2011. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* **18**: 91–93. http://www.ncbi.nlm.nih.gov/pubmed/21131980 (Accessed January 18, 2018).

Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, 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. http://www.ncbi.nlm.nih.gov/pubmed/21441907.

Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y. 2002. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* **12**: 1052–8. http://www.ncbi.nlm.nih.gov/pubmed/12123582 (Accessed January 18, 2018).

Ferrari KJ, Scelfo A, Jammula S, Cuomo A, Barozzi I, Stützer A, Fischle W, Bonaldi T, Pasini D. 2014. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol Cell* **53**: 49–62. http://www.ncbi.nlm.nih.gov/pubmed/24289921 (Accessed January 18, 2018).

Fierz B, Muir TW. 2012. Chromatin as an expansive canvas for chemical biology. *Nat Chem Biol* **8**: 417–427. http://www.ncbi.nlm.nih.gov/pubmed/22510649.

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **143**: 212–24. http://www.ncbi.nlm.nih.gov/pubmed/20888037 (Accessed January 19, 2018).

Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert J-P, Barsyte-Lovejoy D, Felletar I, Volkmer R, Müller S, Pawson T, et al. 2012. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**: 214–31. http://www.ncbi.nlm.nih.gov/pubmed/22464331 (Accessed March 20, 2018).

Fischle W, Wang Y, Allis CD. 2003. Histone and chromatin cross-talk. *Curr Opin Cell Biol* **15**: 172–183. https://www.sciencedirect.com/science/article/pii/S0955067403000139 (Accessed March 19, 2018).

Frenzel A, Hust M, Schirrmann T. 2013. Expression of recombinant antibodies. *Front Immunol* **4**: 217. http://www.ncbi.nlm.nih.gov/pubmed/23908655 (Accessed February 23, 2018).

Fu Y, Luo G-Z, Chen K, Deng X, Yu M, Han D, Hao Z, Liu J, Lu X, Doré LC, et al. 2015. N6-Methyldeoxyadenosine Marks Active Transcription Start Sites in Chlamydomonas. *Cell* **161**: 879–892. http://www.ncbi.nlm.nih.gov/pubmed/25936837 (Accessed February 22, 2018).

Fuchs SM, Krajewski K, Baker RW, Miller VL, Strahl BD. 2011. Influence of Combinatorial Histone Modifications on Antibody and Effector Protein Recognition. *Curr Biol* **21**: 53–58. http://www.ncbi.nlm.nih.gov/pubmed/21167713 (Accessed January 18, 2018).

Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. 2003. The Methyl-CpG-binding Protein MeCP2 Links DNA Methylation to Histone Methylation. *J Biol Chem* **278**: 4035–4040. http://www.ncbi.nlm.nih.gov/pubmed/12427740 (Accessed January 18, 2018).

Garske AL, Craciun G, Denu JM. 2008. A Combinatorial H4 Tail Library for Exploring the Histone Code †. *Biochemistry* **47**: 8094–8102. http://www.ncbi.nlm.nih.gov/pubmed/18616348 (Accessed January 19, 2018).

Garske AL, Oliver SS, Wagner EK, Musselman CA, LeRoy G, Garcia BA, Kutateladze TG, Denu JM. 2010. Combinatorial profiling of chromatin binding modules reveals multisite discrimination. *Nat Chem Biol* **6**: 283–290. http://www.ncbi.nlm.nih.gov/pubmed/20190764 (Accessed January 19, 2018).

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–345. http://www.nature.com/articles/nmeth.1318 (Accessed January 19, 2018).

Gil J, Bernard D, Martínez D, Beach D. 2004. Polycomb CBX7 has a unifying role in cellular lifespan. *Nat Cell Biol* **6**: 67–72. http://www.ncbi.nlm.nih.gov/pubmed/14647293 (Accessed January 23, 2018).

Gilmour DS, Lis JT. 1984. Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proc Natl Acad Sci U S A* **81**: 4275–9. http://www.ncbi.nlm.nih.gov/pubmed/6379641 (Accessed January 19, 2018).

Goecks J, Nekrutenko A, Taylor J, Galaxy T. 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* **11**: R86. http://www.ncbi.nlm.nih.gov/pubmed/20738864.

Goldberg AD, Allis CD, Bernstein E. 2007. Epigenetics: a landscape takes shape. *Cell* **128**: 635–638. http://www.ncbi.nlm.nih.gov/pubmed/17320500.

Goll MG, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74**: 481–514. http://www.ncbi.nlm.nih.gov/pubmed/15952895.

Gonzalo S, García-Cao M, Fraga MF, Schotta G, Peters AHFM, Cotter SE, Eguía R, Dean DC, Esteller M, Jenuwein T, et al. 2005. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat Cell Biol* **7**: 420–428. http://www.nature.com/articles/ncb1235 (Accessed January 18, 2018).

Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, Aristizábal-Corrales D, Hsu C-H, Aravind L, He C, Shi Y. 2015. DNA Methylation on N6-Adenine in C. elegans. *Cell* **161**: 868–878. http://www.ncbi.nlm.nih.gov/pubmed/25936839 (Accessed February 22, 2018).

Greer EL, Shi Y. 2012. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet* **13**: 343–357. http://www.ncbi.nlm.nih.gov/pubmed/22473383 (Accessed March 2, 2018).

Grigoryev SA, Arya G, Correll S, Woodcock CL, Schlick T. 2009. Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc Natl Acad Sci* **106**: 13317–13322. http://www.ncbi.nlm.nih.gov/pubmed/19651606 (Accessed January 18, 2018).

Grigoryev SA, Woodcock CL. 2012. Chromatin organization — The 30nm fiber. *Exp Cell Res* **318**: 1448–1455. http://www.ncbi.nlm.nih.gov/pubmed/22394510 (Accessed January 18, 2018).

Grimm C, Matos R, Ly-Hartig N, Steuerwald U, Lindner D, Rybin V, Müller J, Müller CW. 2009. Molecular recognition of histone lysine methylation by the Polycomb group repressor dSfmbt. *EMBO J* **28**: 1965–77. http://www.ncbi.nlm.nih.gov/pubmed/19494831 (Accessed February 21, 2018).

Guan X, Rastogi N, Parthun MR, Freitas MA. 2013. Discovery of Histone Modification Crosstalk Networks by Stable Isotope Labeling of Amino Acids in Cell Culture Mass Spectrometry (SILAC MS). *Mol Cell Proteomics* **12**: 2048–2059. http://www.ncbi.nlm.nih.gov/pubmed/23592332 (Accessed January 19, 2018).

Guo X, Wang L, Li J, Ding Z, Xiao J, Yin X, He S, Shi P, Dong L, Li G, et al. 2014. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**: 640–644. http://www.nature.com/doifinder/10.1038/nature13899 (Accessed February 12, 2018).

Halachev K, Bast H, Albrecht F, Lengauer T, Bock C. 2012. EpiExplorer: live exploration and global analysis of large epigenomic datasets. *Genome Biol* **13**: R96. http://www.ncbi.nlm.nih.gov/pubmed/23034089.

Hamidi T, Singh AK, Chen T. 2015. Genetic alterations of DNA methylation machinery in human diseases. *Epigenomics* **7**: 247–265. http://www.ncbi.nlm.nih.gov/pubmed/25942534 (Accessed January 18, 2018).

Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X. 2008. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455**: 826–829. http://www.ncbi.nlm.nih.gov/pubmed/18772888 (Accessed January 24, 2018).

Hattori T, Lai D, Dementieva IS, Montaño SP, Kurosawa K, Zheng Y, Akin LR, Świst-Rosowska KM, Grzybowski AT, Koide A, et al. 2016. Antigen clasping by two antigen-binding sites of an exceptionally specific antibody for histone methylation. *Proc Natl Acad Sci U S A* **113**: 2092–7. http://www.ncbi.nlm.nih.gov/pubmed/26862167 (Accessed January 30, 2018).

Hattori T, Taft JM, Swist KM, Luo H, Witt H, Slattery M, Koide A, Ruthenburg AJ, Krajewski K, Strahl BD, et al. 2013. Recombinant antibodies to histone post-translational modifications. *Nat Methods* **10**: 992–995. http://www.ncbi.nlm.nih.gov/pubmed/23955773 (Accessed January 18, 2018).

Helmink BA, Sleckman BP. 2012. The Response to and Repair of RAG-Mediated DNA Double-Strand Breaks. *Annu Rev Immunol* **30**: 175–202. http://www.ncbi.nlm.nih.gov/pubmed/22224778 (Accessed January 18, 2018).

Heubach Y, Planatscher H, Sommersdorf C, Maisch D, Maier J, Joos TO, Templin MF, Poetz O. 2013. From spots to beads-PTM-peptide bead arrays for the characterization of anti-histone antibodies. *Proteomics* **13**: 1010–1015. http://www.ncbi.nlm.nih.gov/pubmed/23401470 (Accessed January 18, 2018).

Heyn H, Esteller M. 2015. An Adenine Code for DNA: A Second Life for N6-Methyladenine. *Cell* **161**: 710–713. https://www.sciencedirect.com/science/article/pii/S0092867415004389?via%3Dihub (Accessed February 22, 2018).

Holliger P, Hudson PJ. 2005. Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* **23**: 1126–1136. http://www.nature.com/articles/nbt1142 (Accessed January 18, 2018).

Horn PJ, Peterson CL. 2002. MOLECULAR BIOLOGY: Chromatin Higher Order Folding--Wrapping up Transcription. *Science (80-)* **297**: 1824–1827. http://www.ncbi.nlm.nih.gov/pubmed/12228709 (Accessed January 18, 2018).

Hu S, Shively L, Raubitschek A, Sherman M, Williams LE, Wong JY, Shively JE, Wu AM. 1996. Minibody: A novel engineered anti-carcinoembryonic antigen antibody fragment (single-chain Fv-CH3) which exhibits rapid, high-level targeting of xenografts. *Cancer Res* **56**: 3055–61. http://www.ncbi.nlm.nih.gov/pubmed/8674062 (Accessed February 23, 2018).

Hudson PJ, Kortt AA. 1999. High avidity scFv multimers; diabodies and triabodies. *J Immunol Methods* **231**: 177–89. http://www.ncbi.nlm.nih.gov/pubmed/10648937 (Accessed February 23, 2018).

Hudson PJ, Souriau C. 2003. Engineered antibodies. *Nat Med* **9**: 129–134. http://www.ncbi.nlm.nih.gov/pubmed/12514726 (Accessed January 18, 2018).

Hust M, Jostock T, Menzel C, Voedisch B, Mohr A, Brenneis M, Kirsch MI, Meier D, Dübel S. 2007. Single chain Fab (scFab) fragment. *BMC Biotechnol* **7**: 14. http://www.ncbi.nlm.nih.gov/pubmed/17346344 (Accessed February 23, 2018).

Imre L, Simándi Z, Horváth A, Fenyőfalvi G, Nánási P, Niaki EF, Hegedüs É, Bacsó Z, Weyemi U, Mauser R, et al. 2017. Nucleosome stability measured in situ by automated quantitative imaging. *Sci Rep* **7**: 12734. http://www.nature.com/articles/s41598-017-12608-9 (Accessed March 18, 2018).

Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, et al. 2011. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* **18**: 769–776. http://www.ncbi.nlm.nih.gov/pubmed/21666679 (Accessed January 23, 2018).

Jang HS, Shin WJ, Lee JE, Do JT. 2017. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes (Basel)* **8**. http://www.ncbi.nlm.nih.gov/pubmed/28545252 (Accessed March 15, 2018).

Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JI, Rodriguez-Ubreva J, Berdasco M, Fraga MF, O'Hanlon TP, Rider LG, et al. 2010. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* **20**: 170–179. http://www.ncbi.nlm.nih.gov/pubmed/20028698 (Accessed January 18, 2018).

Jeltsch A, Jurkowska RZ. 2014. New concepts in DNA methylation. *Trends Biochem Sci* **39**: 310–318. http://www.ncbi.nlm.nih.gov/pubmed/24947342 (Accessed January 18, 2018).

Jeltsch A, Lanio T. 2002. Site-directed mutagenesis by polymerase chain reaction. *Methods Mol Biol* **182**: 85–94. http://www.ncbi.nlm.nih.gov/pubmed/11768980.

Jia G, Fu Y, He C. 2013. Reversible RNA adenosine methylation in biological regulation. *Trends Genet* **29**: 108–15. http://www.ncbi.nlm.nih.gov/pubmed/23218460 (Accessed February 22, 2018).

Jones PA. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**: 484–492. http://www.ncbi.nlm.nih.gov/pubmed/22641018.

Jones PA, Liang G. 2009. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* **10**: 805–811. http://www.ncbi.nlm.nih.gov/pubmed/19789556 (Accessed January 18, 2018).

Jorgensen S, Schotta G, Sorensen CS. 2013. Histone H4 Lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res* **41**: 2797–2806. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt012 (Accessed January 18, 2018).

Joshi AA, Struhl K. 2005. Eaf3 Chromodomain Interaction with Methylated H3-K36 Links Histone Deacetylation to Pol II Elongation. *Mol Cell* **20**: 971–978. https://www.sciencedirect.com/science/article/pii/S1097276505018083 (Accessed January 18, 2018).

Jurkowska RZ, Jeltsch A. 2016. Mechanisms and Biological Roles of DNA Methyltransferases and DNA Methylation: From Past Achievements to Future Challenges. pp. 1–17, Springer, Cham http://link.springer.com/10.1007/978-3-319-43624-1_1 (Accessed February 26, 2018).

Jurkowska RZ, Jurkowski TP, Jeltsch A. 2011. Structure and Function of Mammalian DNA Methyltransferases. *ChemBioChem* **12**: 206–222. http://www.ncbi.nlm.nih.gov/pubmed/21243710 (Accessed January 18, 2018).

Kachirskaia I, Shi X, Yamaguchi H, Tanoue K, Wen H, Wang EW, Appella E, Gozani O. 2008. Role for 53BP1 Tudor Domain Recognition of p53 Dimethylated at Lysine 382 in DNA Damage Signaling. *J Biol Chem* **283**: 34660–34666. http://www.ncbi.nlm.nih.gov/pubmed/18840612 (Accessed February 15, 2018).

Kalashnikova AA, Porter-Goff ME, Muthurajan UM, Luger K, Hansen JC. 2013. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J R Soc Interface* **10**: 20121022. http://www.ncbi.nlm.nih.gov/pubmed/23446052 (Accessed January 18, 2018).

Kallio MA, Tuimala JT, Hupponen T, Klemela P, Gentile M, Scheinin I, Koski M, Kaki J, Korpelainen EI. 2011. Chipster: user-friendly analysis software for microarray and other high-throughput data. *BMC Genomics* **12**: 507. http://www.ncbi.nlm.nih.gov/pubmed/21999641.

Kanno T, Kanno Y, Siegel RM, Jang MK, Lenardo MJ, Ozato K. 2004. Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol Cell* **13**: 33–43. http://www.ncbi.nlm.nih.gov/pubmed/14731392 (Accessed March 26, 2018).

Karch KR, Denizio JE, Black BE, Garcia BA. 2013. Identification and interrogation of combinatorial histone modifications. *Front Genet* **4**: 264. http://www.ncbi.nlm.nih.gov/pubmed/24391660 (Accessed January 19, 2018).

Kasinathan S, Orsi GA, Zentner GE, Ahmad K, Henikoff S. 2014. High-resolution mapping of transcription factor binding sites on native chromatin. *Nat Methods* **11**: 203–209. http://www.ncbi.nlm.nih.gov/pubmed/24336359.

Kaustov L, Ouyang H, Amaya M, Lemak A, Nady N, Duan S, Wasney GA, Li Z, Vedadi M, Schapira M, et al. 2011. Recognition and Specificity Determinants of the Human Cbx Chromodomains. *J Biol Chem* **286**: 521–529. http://www.ncbi.nlm.nih.gov/pubmed/21047797 (Accessed January 23, 2018).

Kebede AF, Nieborak A, Shahidian LZ, Le Gras S, Richter F, Gómez DA, Baltissen MP, Meszaros G, Magliarelli H de F, Taudt A, et al. 2017. Histone propionylation is a mark of active chromatin. *Nat Struct Mol Biol* **24**: 1048–1056. http://www.ncbi.nlm.nih.gov/pubmed/29058708 (Accessed February 28, 2018).

Kebede AF, Schneider R, Daujat S. 2015. Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS J* **282**: 1658–1674. http://doi.wiley.com/10.1111/febs.13047 (Accessed February 26, 2018).

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* **12**: 996–1006. http://www.ncbi.nlm.nih.gov/pubmed/12045153.

Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, et al. 2005. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**: 593–605. http://www.ncbi.nlm.nih.gov/pubmed/16286008.

Kim J, Hake SB, Roeder RG. 2005. The Human Homolog of Yeast BRE1 Functions as a Transcriptional Coactivator through Direct Activator Interactions. *Mol Cell* **20**: 759–770. http://www.ncbi.nlm.nih.gov/pubmed/16337599 (Accessed February 22, 2018).

Klose RJ, Bird AP. 2006. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* **31**: 89–97. http://www.ncbi.nlm.nih.gov/pubmed/16403636 (Accessed January 18, 2018).

Köhler G, Milstein C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**: 495–497. http://www.nature.com/doifinder/10.1038/256495a0 (Accessed February 23, 2018).

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* **41**: 376–381. http://www.ncbi.nlm.nih.gov/pubmed/19182803.

Kondo Y. 2009. Epigenetic cross-talk between DNA methylation and histone modifications in human cancers. *Yonsei Med J* **50**: 455–63. http://www.ncbi.nlm.nih.gov/pubmed/19718392 (Accessed January 18, 2018).

Kontermann RE, Brinkmann U. 2015. Bispecific antibodies. *Drug Discov Today* **20**: 838–847. http://www.sciencedirect.com/science/article/pii/S135964461500077X?via%3Dihub (Accessed January 19, 2018).

Kostrhon S, Kontaxis G, Kaufmann T, Schirghuber E, Kubicek S, Konrat R, Slade D. 2017. A histone-mimicking interdomain linker in a multidomain protein modulates multivalent histone binding. *J Biol Chem* **292**: 17643–17657. http://www.ncbi.nlm.nih.gov/pubmed/28864776 (Accessed January 19, 2018).

Kouzarides T. 2007. Chromatin Modifications and Their Function. *Cell* **128**: 693–705. http://www.ncbi.nlm.nih.gov/pubmed/17320507 (Accessed January 18, 2018).

Krishnan S, Horowitz S, Trievel RC. 2011. Structure and function of histone H3 lysine 9 methyltransferases and demethylases. *Chembiochem* **12**: 254–263. http://www.ncbi.nlm.nih.gov/pubmed/21243713.

Kungulovski G, Jeltsch A. 2015. Quality of histone modification antibodies undermines chromatin biology research. *F1000Res* **4**: 1160. http://www.ncbi.nlm.nih.gov/pubmed/26834995.

Kungulovski G, Kycia I, Mauser R, Jeltsch A. 2015a. Specificity Analysis of Histone Modification-Specific Antibodies or Reading Domains on Histone Peptide Arrays. In *Methods in molecular biology (Clifton, N.J.)*, Vol. 1348 of, pp. 275–284 http://www.ncbi.nlm.nih.gov/pubmed/26424280 (Accessed March 19, 2018).

Kungulovski G, Kycia I, Tamas R, Jurkowska RZ, Kudithipudi S, Henry C, Reinhardt R, Labhart P, Jeltsch A. 2014. Application of histone modification-specific interaction domains as an alternative to antibodies. *Genome Res* **24**: 1842–1853. http://www.ncbi.nlm.nih.gov/pubmed/25301795.

Kungulovski G, Mauser R, Jeltsch A. 2015b. Affinity reagents for studying histone modifications & guidelines for their quality control. *Epigenomics* **7**: 1185–1196. http://www.futuremedicine.com/doi/10.2217/epi.15.59 (Accessed February 23, 2018).

Kungulovski G, Mauser R, Reinhardt R, Jeltsch A. 2016. Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody. *Epigenetics Chromatin* **9**: 11. http://www.ncbi.nlm.nih.gov/pubmed/27006701.

Kuo M-H, Allis CD. 1999. In Vivo Cross-Linking and Immunoprecipitation for Studying Dynamic Protein:DNA Associations in a Chromatin Environment. *Methods* **19**: 425–433. http://www.ncbi.nlm.nih.gov/pubmed/10579938 (Accessed January 19, 2018).

Kurdistani SK, Grunstein M. 2003. Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* **4**: 276–284. http://www.ncbi.nlm.nih.gov/pubmed/12671650 (Accessed January 18, 2018).

Kurdistani SK, Tavazoie S, Grunstein M. 2004. Mapping Global Histone Acetylation Patterns to Gene Expression. *Cell* **117**: 721–733. http://www.ncbi.nlm.nih.gov/pubmed/15186774 (Accessed January 18, 2018).

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120. http://www.ncbi.nlm.nih.gov/pubmed/11242053 (Accessed January 18, 2018).

Lachner M, O'Sullivan RJ, Jenuwein T. 2003. An epigenetic road map for histone lysine methylation. *J Cell Sci* **116**: 2117–24. http://www.ncbi.nlm.nih.gov/pubmed/12730288 (Accessed March 19, 2018).

Lange M, Kaynak B, Forster UB, Tönjes M, Fischer JJ, Grimm C, Schlesinger J, Just S, Dunkel I, Krueger T, et al. 2008. Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev* **22**: 2370–84. http://www.ncbi.nlm.nih.gov/pubmed/18765789 (Accessed January 18, 2018).

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25. http://www.ncbi.nlm.nih.gov/pubmed/19261174.

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013. H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. *Cell* **152**: 1021–1036. http://www.ncbi.nlm.nih.gov/pubmed/23452851 (Accessed January 18, 2018).

Lawrence M, Daujat S, Schneider R. 2016. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* **32**: 42–56. https://www.sciencedirect.com/science/article/pii/S0168952515001936?via%3Dihub#bib0660 (Accessed February 26, 2018).

Lewis A, Mitsuya K, Umlauf D, Smith P, Dean W, Walter J, Higgins M, Feil R, Reik W. 2004. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet* **36**: 1291–1295. http://www.nature.com/articles/ng1468 (Accessed March 2, 2018).

Li B-Z, Huang Z, Cui Q-Y, Song X-H, Du L, Jeltsch A, Chen P, Li G, Li E, Xu G-L. 2011. Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. *Cell Res* **21**: 1172. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3193484/ (Accessed February 12, 2018).

Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**: 91–95. http://www.ncbi.nlm.nih.gov/pubmed/16728978 (Accessed January 18, 2018).

Liang G, Lin JCY, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA, et al. 2004. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci* **101**: 7357–7362. http://www.ncbi.nlm.nih.gov/pubmed/15123803 (Accessed January 18, 2018).

Lin C-W, Jao CY, Ting AY. 2003. Genetically Encoded Fluorescent Reporters of Histone Methylation in Living Cells. *J Am Chem Soc* **126**: 5982–5983. https://pubs.acs.org/doi/pdf/10.1021/ja038854h (Accessed February 13, 2018).

Local A, Huang H, Albuquerque CP, Singh N, Lee AY, Wang W, Wang C, Hsia JE, Shiau AK, Ge K, et al. 2018. Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat Genet* **50**: 73–82. http://www.nature.com/articles/s41588-017-0015-6 (Accessed March 20, 2018).

Lorincz MC, Dickerson DR, Schmitt M, Groudine M. 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* **11**: 1068–1075. http://www.ncbi.nlm.nih.gov/pubmed/15467727 (Accessed January 18, 2018).

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251–260. http://www.ncbi.nlm.nih.gov/pubmed/9305837.

Lungu C, Pinter S, Broche J, Rathert P, Jeltsch A. 2017. Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites. *Nat Commun* **8**: 649. http://www.ncbi.nlm.nih.gov/pubmed/28935858 (Accessed January 30, 2018).

Lunyak V V., Burgess R, Prefontaine GG, Nelson C, Sze S-H, Chenoweth J, Schwartz P, Pevzner PA, Glass C, Mandel G, et al. 2002. Corepressor-Dependent Silencing of Chromosomal Regions Encoding Neuronal Genes. *Science (80-)* **298**: 1747–1752. http://www.ncbi.nlm.nih.gov/pubmed/12399542 (Accessed January 18, 2018).

Maisch D, Schmitz I, Brandt O. 2011. CelluSpots Arrays as an Alternative to Peptide Arrays on Membrane Supports. *Mini Rev Org Chem* **8**: 132–136. http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1570-193X&volume=8&issue=2&spage=132 (Accessed January 19, 2018).

Mallette FA, Mattiroli F, Cui G, Young LC, Hendzel MJ, Mer G, Sixma TK, Richard S. 2012. RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *EMBO J* **31**: 1865–78. http://www.ncbi.nlm.nih.gov/pubmed/22373579 (Accessed February 15, 2018).

Maresca A, Zaffagnini M, Caporali L, Carelli V, Zanna C. 2015. Dna Methyltransferase 1 Mutations and Mitochondrial Pathology: Is Mtdna Methylated? *Front Genet* **6**: 90. http://journal.frontiersin.org/Article/10.3389/fgene.2015.00090/abstract (Accessed January 19, 2018).

Margueron R, Trojer P, Reinberg D. 2005. The key to development: interpreting the histone code? *Curr Opin Genet Dev* **15**: 163–176. https://www.sciencedirect.com/science/article/pii/S0959437X05000237?via%3Dihub (Accessed March 2, 2018).

Martin C, Zhang Y. 2005. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* **6**: 838–849. http://www.ncbi.nlm.nih.gov/pubmed/16261189.

Matthews AGW, Kuo AJ, Ramón-Maiques S, Han S, Champagne KS, Ivanov D, Gallardo M, Carney D, Cheung P, Ciccone DN, et al. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* **450**: 1106–10. http://www.ncbi.nlm.nih.gov/pubmed/18033247 (Accessed January 18, 2018).

Mauser R, Kungulovski G, Keup C, Reinhardt R, Jeltsch A. 2017. Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state. *Epigenetics Chromatin* **10**: 45. http://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/s13072-017-0153-1 (Accessed January 19, 2018).

Mauser R, Kungulovski G, Meral D, Maisch D, Jeltsch A. 2018. Application of mixed peptide arrays to study combinatorial readout of chromatin modifications. *Biochimie* **146**: 14–19. http://www.ncbi.nlm.nih.gov/pubmed/29133117 (Accessed January 18, 2018).

Meffre E, Casellas R, Nussenzweig MC. 2000. Antibody regulation of B cell development. *Nat Immunol* **1**: 379–385. http://www.nature.com/doifinder/10.1038/80816 (Accessed January 18, 2018).

Miao F, Natarajan R. 2005. Mapping Global Histone Methylation Patterns in the Coding Regions of Human Genes. *Mol Cell Biol* **25**: 4650–4661. http://www.ncbi.nlm.nih.gov/pubmed/15899867 (Accessed February 6, 2018).

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim T-K, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560. http://www.ncbi.nlm.nih.gov/pubmed/17603471 (Accessed January 18, 2018).

Min J, Zhang Y, Xu R-M. 2003. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev* **17**: 1823–1828. http://www.ncbi.nlm.nih.gov/pubmed/12897052 (Accessed January 23, 2018).

Morishita M, di Luccio E. 2011. Cancers and the NSD family of histone lysine methyltransferases. *Biochim Biophys Acta - Rev Cancer* **1816**: 158–163. http://www.ncbi.nlm.nih.gov/pubmed/21664949 (Accessed February 22, 2018).

Mozzetta C, Boyarchuk E, Pontis J, Ait-Si-Ali S. 2015. Sound of silence: the properties and functions of repressive Lys methyltransferases. *Nat Rev Mol Cell Biol* **16**: 499–513. http://www.ncbi.nlm.nih.gov/pubmed/26204160 (Accessed January 18, 2018).

Muir TW, Sondhi D, Cole PA. 1998. Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A* **95**: 6705–10. http://www.ncbi.nlm.nih.gov/pubmed/9618476 (Accessed January 19, 2018).

Mujtaba S, He Y, Zeng L, Yan S, Plotnikova O, Sachchidanand, Sanchez R, Zeleznik-Le NJ, Ronai Z, Zhou M-M. 2004. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol Cell* **13**: 251–63. http://www.ncbi.nlm.nih.gov/pubmed/14759370 (Accessed February 6, 2018).

Mujtaba S, Zeng L, Zhou M-M. 2007. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**: 5521–5527. http://www.ncbi.nlm.nih.gov/pubmed/17694091 (Accessed January 19, 2018).

Müller D, Kontermann RE. 2010. Bispecific Antibodies for Cancer Immunotherapy. *BioDrugs* **24**: 89–98. http://www.ncbi.nlm.nih.gov/pubmed/20199124 (Accessed March 30, 2018).

Musselman CA, Lalonde ME, Cote J, Kutateladze TG. 2012. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* **19**: 1218–1227. http://www.ncbi.nlm.nih.gov/pubmed/23211769.

Nakamura Y, Umehara T, Hamana H, Hayashizaki Y, Inoue M, Kigawa T, Shirouzu M, Terada T, Tanaka A, Padmanabhan B, et al. 2007. Crystal Structure Analysis of the PHD Domain of the Transcription Co-activator Pygopus. *J Mol Biol* **370**: 80–92. http://www.ncbi.nlm.nih.gov/pubmed/17499269 (Accessed January 18, 2018).

Nakanishi S, Lee JS, Gardner KE, Gardner JM, Takahashi Y, Chandrasekharan MB, Sun Z-W, Osley MA, Strahl BD, Jaspersen SL, et al. 2009. Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *J Cell Biol* **186**: 371–7. http://www.ncbi.nlm.nih.gov/pubmed/19667127 (Accessed February 26, 2018).

Nakayama J -i., Rice JC, Strahl BD, Allis CD, Grewal SI. 2001. Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. *Science (80-)* **292**: 110–113. http://www.ncbi.nlm.nih.gov/pubmed/11283354 (Accessed January 18, 2018).

Nelson AL, Reichert JM. 2009. Development trends for therapeutic antibody fragments. *Nat Biotechnol* **27**: 331–337. http://www.nature.com/articles/nbt0409-331 (Accessed March 16, 2018).

Nightingale KP, O'Neill LP, Turner BM. 2006. Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. *Curr Opin Genet Dev* **16**: 125–136. https://www.sciencedirect.com/science/article/pii/S0959437X06000372?via%3Dihub#bib4 (Accessed March 2, 2018).

Nishikori S, Hattori T, Fuchs SM, Yasui N, Wojcik J, Koide A, Strahl BD, Koide S. 2012. Broad Ranges of Affinity and Specificity of Anti-Histone Antibodies Revealed by a Quantitative Peptide Immunoprecipitation Assay. *J Mol Biol* **424**: 391–399. http://linkinghub.elsevier.com/retrieve/pii/S0022283612007875 (Accessed January 18, 2018).

Nishiyama A, Yamaguchi L, Sharif J, Johmura Y, Kawamura T, Nakanishi K, Shimamura S, Arita K, Kodama T, Ishikawa F, et al. 2013. Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502**: 249–253. http://www.nature.com/articles/nature12488 (Accessed February 23, 2018).

Oda H, Okamoto I, Murphy N, Chu J, Price SM, Shen MM, Torres-Padilla ME, Heard E, Reinberg D. 2009. Monomethylation of Histone H4-Lysine 20 Is Involved in Chromosome Structure and Stability and Is Essential for Mouse Development. *Mol Cell Biol* **29**: 2278–2295. http://www.ncbi.nlm.nih.gov/pubmed/19223465 (Accessed January 18, 2018).

Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin S-P, Allis CD, et al. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**: 714–717. http://www.nature.com/articles/nature05987 (Accessed January 19, 2018).

Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. 2009. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep* **10**: 1235–41. http://www.ncbi.nlm.nih.gov/pubmed/19834512 (Accessed January 19, 2018).

Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P, Travers AA. 2000. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J* **19**: 6141–6149. http://www.ncbi.nlm.nih.gov/pubmed/11080160 (Accessed January 19, 2018).

Pauler FM, Sloane MA, Huang R, Regha K, Koerner M V., Tamir I, Sommer A, Aszodi A, Jenuwein T, Barlow DP. 2008. H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res* **19**: 221–233. http://www.ncbi.nlm.nih.gov/pubmed/19047520 (Accessed January 18, 2018).

Payer B, Lee JT. 2008. X Chromosome Dosage Compensation: How Mammals Keep the Balance. *Annu Rev Genet* **42**: 733–772. http://www.ncbi.nlm.nih.gov/pubmed/18729722 (Accessed January 18, 2018).

Peach SE, Rudomin EL, Udeshi ND, Carr SA, Jaffe JD. 2012. Quantitative Assessment of Chromatin Immunoprecipitation Grade Antibodies Directed against Histone Modifications Reveals Patterns of Co-occurring Marks on Histone Protein Molecules. *Mol Cell Proteomics* **11**: 128–137. http://www.ncbi.nlm.nih.gov/pubmed/22442256 (Accessed January 18, 2018).

Pepenella S, Murphy KJ, Hayes JJ. 2014. Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. *Chromosoma* **123**: 3–13. http://www.ncbi.nlm.nih.gov/pubmed/23996014 (Accessed January 18, 2018).

Pesavento JJ, Yang H, Kelleher NL, Mizzen CA. 2008. Certain and Progressive Methylation of Histone H4 at Lysine 20 during the Cell Cycle. *Mol Cell Biol* **28**: 468–486. http://www.ncbi.nlm.nih.gov/pubmed/17967882 (Accessed January 18, 2018).

Peters AHFM, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AAHA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, et al. 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* **12**: 1577–89. http://www.ncbi.nlm.nih.gov/pubmed/14690609 (Accessed January 18, 2018).

Pillai S, Chellappan SP. 2009. ChIP on Chip Assays: Genome-Wide Analysis of Transcription Factor Binding and Histone Modifications. In *Methods in molecular biology (Clifton, N.J.)*, Vol. 523 of, pp. 341–366 http://www.ncbi.nlm.nih.gov/pubmed/19381927 (Accessed January 19, 2018).

Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y. 2003. Role of Histone H3 Lysine 27 Methylation in X Inactivation. *Science (80-)* **300**: 131–135. http://www.ncbi.nlm.nih.gov/pubmed/12649488 (Accessed January 18, 2018).

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, et al. 2005. Genome-wide Map of Nucleosome Acetylation and Methylation in Yeast. *Cell* **122**: 517–527. http://www.ncbi.nlm.nih.gov/pubmed/16122420 (Accessed January 18, 2018).

Portela A, Esteller M. 2010. Epigenetic modifications and human disease. *Nat Biotechnol* **28**: 1057–1068. http://www.nature.com/articles/nbt.1685 (Accessed January 23, 2018).

Pradhan M, Estève P-O, Chin HG, Samaranayke M, Kim G-D, Pradhan S. 2008. CXXC Domain of Human DNMT1 Is Essential for Enzymatic Activity. *Biochemistry* **47**: 10000–10009. http://www.ncbi.nlm.nih.gov/pubmed/18754681 (Accessed March 20, 2018).

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842. http://www.ncbi.nlm.nih.gov/pubmed/20110278.

Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187-91. http://www.ncbi.nlm.nih.gov/pubmed/24799436.

Rathert P, Dhayalan A, Ma H, Jeltsch A. 2008a. Specificity of protein lysine methyltransferases and methods for detection of lysine methylation of non-histone proteins. *Mol Biosyst* **4**: 1186. http://www.ncbi.nlm.nih.gov/pubmed/19396382 (Accessed January 19, 2018).

Rathert P, Dhayalan A, Murakami M, Zhang X, Tamas R, Jurkowska R, Komatsu Y, Shinkai Y, Cheng X, Jeltsch A. 2008b. Protein lysine methyltransferase G9a acts on non-histone targets. *Nat Chem Biol* **4**: 344–346. http://www.ncbi.nlm.nih.gov/pubmed/18438403.

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun Z-W, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593–599. http://www.ncbi.nlm.nih.gov/pubmed/10949293 (Accessed January 18, 2018).

Regha K, Sloane MA, Huang R, Pauler FM, Warczok KE, Melikant B, Radolf M, Martens JHA, Schotta G, Jenuwein T, et al. 2007. Active and Repressive Chromatin Are Interspersed without Spreading in an Imprinted Gene Cluster in the Mammalian Genome. *Mol Cell* **27**: 353–366. http://www.ncbi.nlm.nih.gov/pubmed/17679087 (Accessed January 18, 2018).

Ren C, Morohashi K, Plotnikov AN, Jakoncic J, Smith SG, Li J, Zeng L, Rodriguez Y, Stojanoff V, Walsh M, et al. 2015. Small-molecule modulators of methyl-lysine binding for the CBX7 chromodomain. *Chem Biol* **22**: 161–8. http://www.ncbi.nlm.nih.gov/pubmed/25660273 (Accessed February 15, 2018).

Rice JC, Allis CD. 2001. Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr Opin Cell Biol* **13**: 263–73. http://www.ncbi.nlm.nih.gov/pubmed/11343896 (Accessed January 18, 2018).

Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD. 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* **12**: 1591–8. http://www.ncbi.nlm.nih.gov/pubmed/14690610 (Accessed January 18, 2018).

Riddle NC, Minoda A, Kharchenko P V, Alekseyenko AA, Schwartz YB, Tolstorukov MY, Gorchakov AA, Jaffe JD, Kennedy C, Linder-Basso D, et al. 2011. Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. *Genome Res* **21**: 147–63. http://www.ncbi.nlm.nih.gov/pubmed/21177972 (Accessed January 18, 2018).

Riggs AD, Xiong Z. 2004. Methylation and epigenetic fidelity. *Proc Natl Acad Sci* **101**: 4–5. http://www.ncbi.nlm.nih.gov/pubmed/14695893 (Accessed January 18, 2018).

Rivera CM, Ren B. 2013. Mapping Human Epigenomes. *Cell* **155**: 39–55. http://www.ncbi.nlm.nih.gov/pubmed/24074860 (Accessed March 7, 2018).

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26. http://www.ncbi.nlm.nih.gov/pubmed/21221095.

Robinson PJ, Rhodes D. 2006. Structure of the “30nm” chromatin fibre: A key role for the linker histone. *Curr Opin Struct Biol* **16**: 336–343. http://www.ncbi.nlm.nih.gov/pubmed/16714106 (Accessed January 18, 2018).

Rondelet G, Dal Maso T, Willems L, Wouters J. 2016. Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. *J Struct Biol* **194**: 357–367. http://www.ncbi.nlm.nih.gov/pubmed/26993463.

Rosa S, Shaw P. 2013. Insights into Chromatin Structure and Dynamics in Plants. *Biology (Basel)* **2**: 1378–1410. http://www.mdpi.com/2079-7737/2/4/1378 (Accessed February 16, 2018).

Rose NR, Klose RJ. 2014. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim Biophys Acta - Gene Regul Mech* **1839**: 1362–1372. https://www.sciencedirect.com/science/article/pii/S1874939914000285 (Accessed February 6, 2018).

Roth DB. 2014. V(D)J Recombination: Mechanism, Errors, and Fidelity. *Microbiol Spectr* **2**. http://www.ncbi.nlm.nih.gov/pubmed/26104458 (Accessed January 18, 2018).

Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Barsyte-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, et al. 2012. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* **19**: 1155–60. http://www.ncbi.nlm.nih.gov/pubmed/23022729 (Accessed January 24, 2018).

Rothbart SB, Strahl BD. 2014. Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta - Gene Regul Mech* **1839**: 627–643. http://www.ncbi.nlm.nih.gov/pubmed/24631868 (Accessed February 6, 2018).

Rougeulle C, Chaumeil J, Sarma K, Allis CD, Reinberg D, Avner P, Heard E. 2004. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* **24**: 5475–84. http://www.ncbi.nlm.nih.gov/pubmed/15169908 (Accessed January 18, 2018).

Roy S, Musselman CA, Kachirskaia I, Hayashi R, Glass KC, Nix JC, Gozani O, Appella E, Kutateladze TG. 2010. Structural Insight into p53 Recognition by the 53BP1 Tandem Tudor Domain. *J Mol Biol* **398**: 489–496. http://www.ncbi.nlm.nih.gov/pubmed/20307547 (Accessed February 15, 2018).

Ruthenburg AJ, Li H, Milne TA, Dewell S, McGinty RK, Yuen M, Ueberheide B, Dou Y, Muir TW, Patel DJ, et al. 2011. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**: 692–706. http://www.ncbi.nlm.nih.gov/pubmed/21596426.

Ruthenburg AJ, Li H, Patel DJ, Allis CD, David Allis C, Allis CD. 2007. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**: 983–994. http://www.ncbi.nlm.nih.gov/pubmed/18037899 (Accessed January 23, 2018).

Saint-André V, Batsché E, Rachez C, Muchardt C. 2011. Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. *Nat Struct Mol Biol* **18**: 337–344. http://www.ncbi.nlm.nih.gov/pubmed/21358630 (Accessed January 18, 2018).

Saksouk N, Barth TK, Ziegler-Birling C, Olova N, Nowak A, Rey E, Mateos-Langerak J, Urbach S, Reik W, Torres-Padilla M-E, et al. 2014. Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation. *Mol Cell* **56**: 580–594. http://www.ncbi.nlm.nih.gov/pubmed/25457167 (Accessed January 18, 2018).

Saksouk N, Simboeck E, Déjardin J. 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**: 3. http://www.ncbi.nlm.nih.gov/pubmed/25788984 (Accessed March 2, 2018).

Sanchez OF, Mendonca A, Carneiro AD, Yuan C. 2017. Engineering Recombinant Protein Sensors for Quantifying Histone Acetylation. *ACS Sensors* **2**: 426–435. http://pubs.acs.org/doi/10.1021/acssensors.7b00026 (Accessed March 26, 2018).

Sarraf SA, Stancheva I. 2004. Methyl-CpG Binding Protein MBD1 Couples Histone H3 Methylation at Lysine 9 by SETDB1 to DNA Replication and Chromatin Assembly. *Mol Cell* **15**: 595–605. http://www.ncbi.nlm.nih.gov/pubmed/15327775 (Accessed January 18, 2018).

Sasaki K, Ito T, Nishino N, Khochbin S, Yoshida M. 2009. Real-time imaging of histone H4 hyperacetylation in living cells. *Proc Natl Acad Sci U S A* **106**: 16257–62. http://www.ncbi.nlm.nih.gov/pubmed/19805290 (Accessed March 26, 2018).

Scarsdale JN, Webb HD, Ginder GD, Williams DC, Jr. 2011. Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. *Nucleic Acids Res* **39**: 6741–52. http://www.ncbi.nlm.nih.gov/pubmed/21531701 (Accessed January 23, 2018).

Schatz DG, Swanson PC. 2011. V(D)J Recombination: Mechanisms of Initiation. *Annu Rev Genet* **45**: 167–202. http://www.ncbi.nlm.nih.gov/pubmed/21854230 (Accessed January 18, 2018).

Schiffers S, Ebert C, Rahimoff R, Kosmatchev O, Steinbacher J, Bohne A-V, Spada F, Michalakis S, Nickelsen J, Müller M, et al. 2017. Quantitative LC-MS Provides No Evidence for m6dA or m4dC in the Genome of Mouse Embryonic Stem Cells and Tissues. *Angew Chem Int Ed Engl* **56**: 11268–11271. http://doi.wiley.com/10.1002/anie.201700424 (Accessed March 15, 2018).

Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, et al. 2011. Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Mol Cell* **42**: 330–341. https://www.sciencedirect.com/science/article/pii/S1097276511002875?via%3Dihub (Accessed January 23, 2018).

Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. 2004. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* **6**: 73–77. http://www.ncbi.nlm.nih.gov/pubmed/14661024 (Accessed January 18, 2018).

Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251–1262. http://www.ncbi.nlm.nih.gov/pubmed/15145825 (Accessed January 18, 2018).

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. 2007. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* **128**: 735–745. http://www.ncbi.nlm.nih.gov/pubmed/17320510 (Accessed January 18, 2018).

Shanle EK, Rothbart SB, Strahl BD. 2014. Chromatin biochemistry enters the next generation of code "seq-ing." *Nat Methods* **11**: 799–800. http://www.ncbi.nlm.nih.gov/pubmed/25075906.

Shechter D, Dormann HL, Allis CD, Hake SB. 2007. Extraction, purification and analysis of histones. *Nat Protoc* **2**: 1445–1457. http://www.ncbi.nlm.nih.gov/pubmed/17545981.

Shema E, Jones D, Shoresh N, Donohue L, Ram O, Bernstein BE. 2016. Single-molecule decoding of combinatorially modified nucleosomes. *Science (80-)* **352**: 717–721. http://www.ncbi.nlm.nih.gov/pubmed/27151869.

Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Peña P, Lan F, Kaadige MR, et al. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–9. http://www.ncbi.nlm.nih.gov/pubmed/16728974 (Accessed January 18, 2018).

Shilatifard A. 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr Opin Cell Biol* **20**: 341–8. http://www.ncbi.nlm.nih.gov/pubmed/18508253 (Accessed January 23, 2018).

Shilatifard A. 2012. The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. *Annu Rev Biochem* **81**: 65–95. http://www.ncbi.nlm.nih.gov/pubmed/22663077 (Accessed January 18, 2018).

Simon MD, Chu F, Racki LR, de la Cruz CC, Burlingame AL, Panning B, Narlikar GJ, Shokat KM. 2007. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **128**: 1003–1012. http://www.ncbi.nlm.nih.gov/pubmed/17350582.

Sims JK, Houston SI, Magazinnik T, Rice JC. 2006. A Trans-tail Histone Code Defined by Monomethylated H4 Lys-20 and H3 Lys-9 Demarcates Distinct Regions of Silent Chromatin. *J Biol Chem* **281**: 12760–12766. http://www.ncbi.nlm.nih.gov/pubmed/16517599 (Accessed February 6, 2018).

Smolle M, Venkatesh S, Gogol MM, Li H, Zhang Y, Florens L, Washburn MP, Workman JL. 2012. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat Struct Mol Biol* **19**: 884–892. http://www.ncbi.nlm.nih.gov/pubmed/22922743.

Song J, Teplova M, Ishibe-Murakami S, Patel DJ. 2012. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science* **335**: 709–12. http://www.ncbi.nlm.nih.gov/pubmed/22323818 (Accessed January 19, 2018).

Squazzo SL, O'Geen H, Komashko VM, Krig SR, Jin VX, Jang S, Margueron R, Reinberg D, Green R, Farnham PJ. 2006. Suz12 binds to silenced regions of the genomein a cell-type-specific manner. *Genome Res* **16**: 890–900. http://www.ncbi.nlm.nih.gov/pubmed/16751344 (Accessed January 18, 2018).

Strahl BD, Allis CD. 2000. The language of covalent histone modifications. *Nature* **403**: 41–45. http://www.ncbi.nlm.nih.gov/pubmed/10638745.

Su Z, Boersma MD, Lee J-H, Oliver SS, Liu S, Garcia BA, Denu JM. 2014. ChIP-less analysis of chromatin states. *Epigenetics Chromatin* **7**: 7. http://www.ncbi.nlm.nih.gov/pubmed/24872844 (Accessed January 19, 2018).

Su Z, Denu JM. 2016. Reading the Combinatorial Histone Language. *ACS Chem Biol* **11**: 564–574. http://www.ncbi.nlm.nih.gov/pubmed/26675328 (Accessed January 23, 2018).

Suka N, Suka Y, Carmen AA, Wu J, Grunstein M. 2001. Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol Cell* **8**: 473–9. http://www.ncbi.nlm.nih.gov/pubmed/11545749 (Accessed January 18, 2018).

Sun Q, Huang S, Wang X, Zhu Y, Chen Z, Chen D. 2015. N ^{6} -methyladenine functions as a potential epigenetic mark in eukaryotes. *BioEssays* **37**: 1155–1162. http://doi.wiley.com/10.1002/bies.201500076 (Accessed March 2, 2018).

Sun X-J, Wei J, Wu X-Y, Hu M, Wang L, Wang H-H, Zhang Q-H, Chen S-J, Huang Q-H, Chen Z. 2005. Identification and Characterization of a Novel Human Histone H3 Lysine 36-specific Methyltransferase. *J Biol Chem* **280**: 35261–35271. http://www.ncbi.nlm.nih.gov/pubmed/16118227 (Accessed January 18, 2018).

Sun Z-W, Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**: 104–108. http://www.ncbi.nlm.nih.gov/pubmed/12077605 (Accessed January 18, 2018).

Suva ML, Riggi N, Bernstein BE. 2013. Epigenetic reprogramming in cancer. *Science (80-)* **339**: 1567–1570. http://www.ncbi.nlm.nih.gov/pubmed/23539597.

Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**: 465–476. http://www.ncbi.nlm.nih.gov/pubmed/18463664 (Accessed January 18, 2018).

Tan M, Luo H, Lee S, Jin F, Yang JS, 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. http://www.ncbi.nlm.nih.gov/pubmed/21925322.

Tang J, Cho NW, Cui G, Manion EM, Shanbhag NM, Botuyan MV, Mer G, Greenberg RA. 2013. Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination. *Nat Struct Mol Biol* **20**: 317–25. http://www.ncbi.nlm.nih.gov/pubmed/23377543 (Accessed February 15, 2018).

Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait BT, et al. 2006. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol Cell* **24**: 785–96. http://www.ncbi.nlm.nih.gov/pubmed/17157260 (Accessed January 18, 2018).

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**: 1025–1040. http://www.ncbi.nlm.nih.gov/pubmed/17984965 (Accessed January 18, 2018).

Timp W, Feinberg AP. 2013. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat Rev Cancer* **13**: 497–510. http://www.ncbi.nlm.nih.gov/pubmed/23760024 (Accessed January 23, 2018).

Tjeertes J V, Miller KM, Jackson SP. 2009. Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J* **28**: 1878–89. http://www.ncbi.nlm.nih.gov/pubmed/19407812 (Accessed February 26, 2018).

Townsend S, Finlay WJJ, Hearty S, O'Kennedy R. 2006. Optimizing recombinant antibody function in SPR immunosensing. *Biosens Bioelectron* **22**: 268–274. http://www.ncbi.nlm.nih.gov/pubmed/16487701 (Accessed January 18, 2018).

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**: 562–578. http://www.ncbi.nlm.nih.gov/pubmed/22383036.

Trojer P, Reinberg D. 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol Cell* **28**: 1–13. http://www.ncbi.nlm.nih.gov/pubmed/17936700 (Accessed January 18, 2018).

Tropberger P, Pott S, Keller C, Kamieniarz-Gdula K, Caron M, Richter F, Li G, Mittler G, Liu ET, Bühler M, et al. 2013. Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer. *Cell* **152**: 859–872. http://www.ncbi.nlm.nih.gov/pubmed/23415232 (Accessed January 18, 2018).

Truax AD, Greer SF. 2012. ChIP and Re-ChIP Assays: Investigating Interactions Between Regulatory Proteins, Histone Modifications, and the DNA Sequences to Which They Bind. pp. 175–188, Springer, New York, NY http://link.springer.com/10.1007/978-1-61779-376-9_12 (Accessed March 7, 2018).

Tsankova N, Renthal W, Kumar A, Nestler EJ. 2007. Epigenetic regulation in psychiatric disorders. *Nat Rev Neurosci* **8**: 355–367. http://www.nature.com/articles/nrn2132 (Accessed February 16, 2018).

Tse C, Sera T, Wolffe AP, Hansen JC. 1998. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol Cell Biol* **18**: 4629–38. http://www.ncbi.nlm.nih.gov/pubmed/9671473 (Accessed January 18, 2018).

Ueda J, Maehara K, Mashiko D, Ichinose T, Yao T, Hori M, Sato Y, Kimura H, Ohkawa Y, Yamagata K. 2014. Heterochromatin dynamics during the differentiation process revealed by the DNA methylation reporter mouse, MethylRO. *Stem cell reports* **2**: 910–24. http://www.ncbi.nlm.nih.gov/pubmed/24936475 (Accessed March 26, 2018).

Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, Feil R. 2004a. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**: 1296–1300. http://www.nature.com/articles/ng1467 (Accessed March 2, 2018).

Umlauf D, Goto Y, Feil R. 2004b. Site-Specific Analysis of Histone Methylation and Acetylation. In *Epigenetics Protocols*, Vol. 287 of, pp. 099–120, Humana Press, New Jersey http://www.ncbi.nlm.nih.gov/pubmed/15273407 (Accessed January 19, 2018).

Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. 2005. Histone H3 Lysine 9 Methylation and HP1γ Are Associated with Transcription Elongation through Mammalian Chromatin. *Mol Cell* **19**: 381–391. http://www.ncbi.nlm.nih.gov/pubmed/16061184 (Accessed February 6, 2018).

van Ingen H, van Schaik FMA, Wienk H, Ballering J, Rehmann H, Dechesne AC, Kruijzer JAW, Liskamp RMJ, Timmers HTM, Boelens R. 2008. Structural Insight into the Recognition of the H3K4me3 Mark by the TFIID Subunit TAF3. *Structure* **16**: 1245–1256. https://www.sciencedirect.com/science/article/pii/S0969212608002517#fig1 (Accessed February 16, 2018).

van Rossum B, Fischle W, Selenko P. 2012. Asymmetrically modified nucleosomes expand the histone code. *Nat Struct Mol Biol* **19**: 1064–1066. http://www.nature.com/articles/nsmb.2433 (Accessed January 30, 2018).

VanDemark AP, Kasten MM, Ferris E, Heroux A, Hill CP, Cairns BR. 2007. Autoregulation of the Rsc4 Tandem Bromodomain by Gcn5 Acetylation. *Mol Cell* **27**: 817–828. http://www.ncbi.nlm.nih.gov/pubmed/17803945 (Accessed January 19, 2018).

Vastenhouw NL, Schier AF. 2012. Bivalent histone modifications in early embryogenesis. *Curr Opin Cell Biol* **24**: 374–386. http://www.ncbi.nlm.nih.gov/pubmed/22513113.

Vermeulen M, Mulder KW, Denissov S, Pijnappel WWMP, van Schaik FMA, Varier RA, Baltissen MPA, Stunnenberg HG, Mann M, Timmers HTM. 2007. Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell* **131**: 58–69. http://www.ncbi.nlm.nih.gov/pubmed/17884155 (Accessed January 18, 2018).

Vermeulen M, Timmers HM. 2010. Grasping trimethylation of histone H3 at lysine 4. *Epigenomics* **2**: 395–406. http://www.ncbi.nlm.nih.gov/pubmed/22121900 (Accessed January 18, 2018).

Vettese-Dadey M, Grant PA, Hebbes TR, Crane- Robinson C, Allis CD, Workman JL. 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO J* **15**: 2508–18. http://www.ncbi.nlm.nih.gov/pubmed/8665858 (Accessed January 18, 2018).

Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB, Young NL, Garcia BA, Reinberg D, Reinberg D. 2012. Asymmetrically modified nucleosomes. *Cell* **151**: 181–93. http://www.ncbi.nlm.nih.gov/pubmed/23021224 (Accessed January 18, 2018).

Voigt P, Reinberg D. 2011. Histone Tails: Ideal Motifs for Probing Epigenetics through Chemical Biology Approaches. *ChemBioChem* **12**: 236–252. http://www.ncbi.nlm.nih.gov/pubmed/21243712 (Accessed January 23, 2018).

Voigt P, Tee W-WW, Reinberg D. 2013. A double take on bivalent promoters. *Genes Dev* **27**: 1318–1338. http://www.ncbi.nlm.nih.gov/pubmed/23788621 (Accessed January 18, 2018).

Völkel P, Angrand P-O. 2007. The control of histone lysine methylation in epigenetic regulation. *Biochimie* **89**: 1–20. http://www.ncbi.nlm.nih.gov/pubmed/16919862 (Accessed January 18, 2018).

Waddington C.H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* **150**: 563–565.

Wagner EJ, Carpenter PB. 2012. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol* **13**: 115–126. http://www.ncbi.nlm.nih.gov/pubmed/22266761.

Weidle UH, Kontermann RE, Brinkmann U. 2014. Tumor-Antigen–Binding Bispecific Antibodies for Cancer Treatment. *Semin Oncol* **41**: 653–660. http://www.ncbi.nlm.nih.gov/pubmed/25440609 (Accessed March 30, 2018).

Weiner A, Lara-Astiaso D, Krupalnik V, Gafni O, David E, Winter DR, Hanna JH, Amit I. 2016. Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. *Nat Biotechnol* **34**: 953–961. http://www.ncbi.nlm.nih.gov/pubmed/27454738.

Welch RP, Lee C, Imbriano PM, Patil S, Weymouth TE, Smith RA, Scott LJ, Sartor MA. 2014. ChIP-Enrich: gene set enrichment testing for ChIP-seq data. *Nucleic Acids Res* **42**: e105. http://www.ncbi.nlm.nih.gov/pubmed/24878920.

Wiencke JK, Zheng S, Morrison Z, Yeh R-F. 2008. Differentially expressed genes are marked by histone 3 lysine 9 trimethylation in human cancer cells. *Oncogene* **27**: 2412–2421. http://www.ncbi.nlm.nih.gov/pubmed/17968314 (Accessed February 6, 2018).

Wilkinson AW, Gozani O. 2014. Histone-binding domains: strategies for discovery and characterization. *Biochim Biophys Acta* **1839**: 669–75. http://www.ncbi.nlm.nih.gov/pubmed/24525102 (Accessed March 29, 2018).

Wilson IA, Stanfield RL. 2014. Antibody Structure. In *Antibodies for Infectious Diseases*, Vol. 2 of, pp. 49–62, American Society of Microbiology http://www.asmscience.org/content/book/10.1128/9781555817411.chap3 (Accessed January 18, 2018).

Wu H, Zeng H, Lam R, Tempel W, Amaya MF, Xu C, Dombrovski L, Qiu W, Wang Y, Min J. 2011a. Structural and Histone Binding Ability Characterizations of Human PWWP Domains ed. P. Kursula. *PLoS One* **6**: e18919. http://www.ncbi.nlm.nih.gov/pubmed/21720545 (Accessed January 23, 2018).

Wu L, Zee BM, Wang Y, Garcia BA, Dou Y. 2011b. The RING Finger Protein MSL2 in the MOF Complex Is an E3 Ubiquitin Ligase for H2B K34 and Is Involved in Crosstalk with H3 K4 and K79 Methylation. *Mol Cell* **43**: 132–144. http://www.ncbi.nlm.nih.gov/pubmed/21726816 (Accessed February 22, 2018).

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, et al. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86–90. http://www.ncbi.nlm.nih.gov/pubmed/16728976.

Ye T, Krebs AR, Choukrallah MA, Keime C, Plewniak F, Davidson I, Tora L. 2011. seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* **39**: e35. http://www.ncbi.nlm.nih.gov/pubmed/21177645.

Young MD, Willson TA, Wakefield MJ, Trounson E, Hilton DJ, Blewitt ME, Oshlack A, Majewski IJ. 2011. ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Res* **39**: 7415–27. http://www.ncbi.nlm.nih.gov/pubmed/21652639 (Accessed February 6, 2018).

Young NL, DiMaggio PA, Plazas-Mayorca MD, Baliban RC, Floudas CA, Garcia BA. 2009. High throughput characterization of combinatorial histone codes. *Mol Cell Proteomics* **8**: 2266–2284. http://www.ncbi.nlm.nih.gov/pubmed/19654425.

Yuan J, Pu M, Zhang Z, Lou Z. 2009. Histone H3-K56 acetylation is important for genomic stability in mammals. *Cell Cycle* **8**: 1747–1753. http://www.ncbi.nlm.nih.gov/pubmed/19411844 (Accessed January 18, 2018).

Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. 2011. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem* **286**: 7983–9. http://www.ncbi.nlm.nih.gov/pubmed/21239496 (Accessed January 18, 2018).

Yun M, Wu J, Workman JL, Li B. 2011. Readers of histone modifications. *Cell Res* **21**: 564–78. http://www.ncbi.nlm.nih.gov/pubmed/21423274 (Accessed January 18, 2018).

Zeng L, Zhang Q, Li S, Plotnikov AN, Walsh MJ, Zhou M-M. 2010. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**: 258–62. http://www.ncbi.nlm.nih.gov/pubmed/20613843 (Accessed January 18, 2018).

Zhang G, Huang H, Liu D, Cheng Y, Liu X, Zhang W, Yin R, Zhang D, Zhang P, Liu J, et al. 2015a. N6-Methyladenine DNA Modification in Drosophila. *Cell* **161**: 893–906. http://www.ncbi.nlm.nih.gov/pubmed/25936838 (Accessed February 22, 2018).

Zhang T, Cooper S, Brockdorff N. 2015b. The interplay of histone modifications – writers that read. *EMBO Rep* **16**: 1467. http://www.ncbi.nlm.nih.gov/pubmed/26474904 (Accessed January 19, 2018).

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**. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910041/ (Accessed January 19, 2018).

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137. http://www.ncbi.nlm.nih.gov/pubmed/18798982.

Zhou M-M, Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou M-M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**: 491–496. http://www.ncbi.nlm.nih.gov/pubmed/10365964 (Accessed February 6, 2018).

Appendix

(not included in the published version)

Appendix 1

Kungulovski G, Mauser R, Reinhardt R and Jeltsch A (2016) „**Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody.”** *Epigenetics & Chromatin,* doi: 10.1186/s13072-016-0061-9

Appendix 1.1

Supplement material: Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody.

Appendix 2

Imre L, Simándi Z, Horváth A, Fenyőfalvi G, Nánási P, Niaki E, Hegedüs E, Bacsó Z, Weyemi U, Mauser R, Ausio J, Jeltsch A, Bonner W, Nagy L, Kimura H and Szabó G (2017) **„Nucleosome stability measured in situ by automated quantitative imaging”.** *Scientific reports,* doi: 10.1038/s41598-017-12608-9

Appendix 2.1

Supplement material: Nucleosome stability measured in situ by automated quantitative imaging

Appendix 3

Mauser R, Kungulovski G, Meral D, Maisch D and Jeltsch A (2018) **„ Application of mixed peptide arrays to study combinatorial readout of chromatin modifications”.** *Biochimie,* doi: 10.1016/j.biochi.2017.11.008

Appendix 3.1

Supplement material: Application of mixed peptide arrays to study combinatorial readout of chromatin modifications

Appendix 4

Mauser R, Kungulovski G, Keup C, Reinhardt R and Jeltsch A (2017) **„Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state".** *Epigenetics & Chromatin,* doi: 10.1186/s13072-017-0153-1

Appendix 4.1

Supplement material: Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state

Appendix 4.2

Supplement material: Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state

Appendix 5

Kungulovski G, Mauser R, Jeltsch A. (2015) **"Affinity reagents for studying histone modifications and guidelines for their quality control"**, *Epigenomics,* doi: 10.12688/f1000research.7265.1

Appendix 6

Kungulovski G, Kycia I, 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.