Regulation and Readout of Mammalian DNA Methylation

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

Cristiana-Elena Lungu

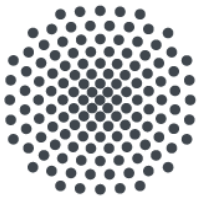
geboren am 18.05.1988 in Constanța, Rumänien

Hauptberichter: Prof. Dr. Albert Jeltsch

Mitberichterin: Prof. Dr. Monilola Olayioye

Prüfungsvorsitzender: Prof. Dr. Markus Morrison

Tag der mündlichen Prüfung : 19.07.2017



Institut für Biochemie und Technische Biochemie

(ehem. Institut für Biochemie) der Universität Stuttgart

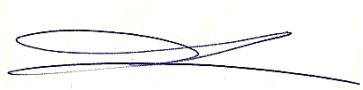
2018

Eidesstattliche Erklärung

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

Declaration of Authorship

I hereby certify that this Dissertation is entirely my own work, apart from where otherwise indicated. Passages and ideas from other sources have been clearly indicated.



Cristiana-Elena Lungu

6th June 2018, Stuttgart

Table of Contents

Acknowledgements

I am grateful to everybody who supported me throughout my doctoral work. In particular, I am thankful to Prof. Dr. Albert Jeltsch, who fortified my scientific enthusiasm and supported my intellectual growth throughout the last 8 years of my scientific career. I am grateful for the freedom and trust I was given during my doctoral studies and for having had the opportunity to explore so many exciting areas of research.

I would also like to thank Prof. Dr. Monilola Olayioye and Prof. Dr. Markus Morrison for taking to time to be co-referees of my PhD thesis.

I am thankful to the Carl Zeiss Stiftung for offering me a PhD stipend to support my work on the epigenetic biosensors project, as well as to all collaborators that were involved in this doctoral work.

A special thanks goes to my MSc students, Sabine and Sebastian. I learnt a lot while working with you and I hope we personally and scientifically grew together.

I would like to thank all my colleagues, scientist and non-scientist friends for their moral, social and intellectual support during this period. You know who you are. There is nothing more precious than time and I am happy that we have spent this together.

Last but not least, my deepest gratitude goes to my family for having the immense strength of transmitting its support and love over a distance of 2000 km. I am proud of being your daughter. I am also grateful to my second family here in Germany, for warmly welcoming me in their lives and making Altbach my second home. Most of all, my love and gratitude goes to Max, for putting up with all of my scientific and non-scientific worries and making me a better person.

This work is dedicated to all ‘naïve’ researchers that still do science for the sake of knowing.

List of publications and achievements

● **Lungu**, **C**., Muegge, K., Jeltsch, A., & Jurkowska, R. Z. (2015). An ATPase-deficient variant of the SNF2 family member HELLS shows altered dynamics at pericentromeric heterochromatin. *Journal of molecular biology*, 427(10), 1903-1915.

This manuscript deals with the cellular and biochemical investigation of the role played by ATP hydrolysis in the interaction between the putative chromatin remodeler HELLS and chromatin.

● Rajavelu, A.[#], **Lungu**, **C**.[#], Hanelt, I., Parsa, A., Karnik, R., Gu, H., Carell, T., Meissner, A. Jurkowska, R. Z. & Jeltsch, A. Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Manuscript submitted for review.* [#]co-shared first authors

In this study, MeCP2 was identified as a strong and direct interactor of DNMT3A and the interaction interfaces were biochemically mapped to the TRD of MeCP2 and the ADD domain of DNMT3A. This interaction was found to result in a concentration-dependent inhibition of DNMT3A activity *in vitro* and in cell lines overexpressing MeCP2. Biochemical investigation revealed that the TRD inhibits the enzymatic activity through an allosteric mechanism and that this effect can be relieved by unmodified histone H3. By performing genome methylome analysis in *Mecp2* knock-out mice brains, both hypo- and hypermethylated regions were found, overlapping with MeCP2 binding sites. The data summed up in this study reveal a novel mechanistic insight into the regulation and targeting of DNMT3A activity and propose that the enzyme is under the combined control of MeCP2 and H3 tail modifications.

● Jeltsch, A., **Lungu**, **C**., Rajavelu, A., Jurkowska, R.Z. (2016) DNA methyltransferase inhibitors for Rett syndrome therapy. EP16156521. Registered on 19.02.2016.

This patent is based on the findings reported in Rajavelu, Lungu *et al*., and proposes the use of DNA methylation inhibitors to compensate for the DNMT3A missregulation caused in the Rett syndrome by mutations in *MECP2*.

● **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. *Nature Communications*, 8(1), 649-649..

In this manuscript, the development and application of a novel set of epigenetic biosensors is described, with which DNA methylation and H3K9me3 marks could be directly visualized for the first time with locus specific resolution in the nucleus of living mammalian cells.

Zusammenfassung

Die faszinierende phänotypische und funktionelle Vielfalt der unterschiedlichen Zelltypen in Säugetieren lässt sich zu einem großen Teil auf epigenetische Signale zurückführen. Diese steuern zusammen mit der DNA Sequenz die Genexpressionsprogramme der Zelle. Eines der wichtigsten epigenetischen Signale ist die DNA Methylierung, ihre besondere Bedeutung wurde bereits in frühen genetischen Studien erkannt. Trotz jahrzehntelanger Forschung ist ein umfassendes Verständnis der Steuerung und Regulation der Aktivität von DNA-Methyltransferasen bisher nicht erreicht worden. Diese Arbeit umfasst drei Forschungsansätze mit dem ultimativen Ziel, mechanistische und methodische Einblicke in die Etablierung und den Erhalt von DNA-Methylierungsmustern zu erzielen.

Im ersten Projekt dieser Arbeit wurde eine Kombination von biochemischen und zellulären Methoden dazu genutzt, um die zelluläre Rolle des potentiellen Chromatin Remodelers HELLS, eines essentiellen Cofaktors für DNA Methyltransferasen in Zellen, zu untersuchen. Unter Verwendung von Chromatin Fraktionierungs Assays und Mikroskopie basierenden Techniken war es mir möglich zu zeigen, dass die ATPase Aktivität von HELLS für dessen hohe Mobilität im Zellkern sowie für dessen Fähigkeit zur Dissoziation von dicht gepacktem Chromatin verantwortlich ist. Darüber hinaus wurde auch festgestellt, dass das H3K9me3-Netzwerk eine wichtige Rolle bei der Interaktion von HELLS mit Heterochromatin spielt. Zusammenfassend legt diese Studie den ersten Beweis dafür dar, dass die Hydrolyse von ATP für die Wechselwirkung zwischen HELLS und Chromatin wichtig ist. Dieses Ergebnis legt nahe, dass der schnelle Austausch von HELLS an repetitiven DNA Sequenzen dazu führen könnte das die Rekrutierung von epigenetischen Enzymen wie beispielsweise DNA Methyltransferasen ermöglicht wird. Dies wiederum hätte eine lokale Stabilisierung von reprimierenden Proteinkomplexen an Heterochromatin zur Konsequenz.

Im zweiten Projekt dieser Arbeit wurde die mögliche Wechselwirkung zwischen der *de novo* DNA Methyltransferase DNMT3A und dem 5mC Leseprotein MeCP2 untersucht. Aufbauend auf früheren Daten unseres Labors, welche eine direkte Interaktion zwischen der TRD Domäne von MeCP2 und der ADD Domäne von DNMT3A zeigten,

die eine inhibierende Wirkung auf die enzymatische Aktivität von DNMT3A *in vitro* hat, konnte ich zeigen, dass beide Proteine auch im Gehirn von Mäusen miteinander interagieren; des Weiteren konnte ich in Zelllinien die MeCP2 stabil überexpremieren den Inhibitionseffekt nachweisen. Außerdem war es mir möglich durch die Verwendung von DNMT3A Varianten, die in einer Konformation arretiert sind und ein neues Werkzeug zur Untersuchung der allosterischen Regulation dieses Enzymes darstellen, den Mechanismus der Inhibierung von DNMT3A durch MeCP2 zu erklären. Mit diesem Ansatz konnte ich zeigen, dass MeCP2 die allosterisch geschlossene, inaktive Konformation von DNMT3A stabilisiert, ein Effekt welcher durch die Zugabe von unmodifiziertem Histon H3 aufgehoben werden konnte. Diese Ergebnisse wurden unterstützt durch eine Bisulfit-Sequenzierungsanalyse des Methyloms von Gehirnzellen eines *Mecp2* Knockout Mausmodells. Zusammenfassend geben die Ergebnisse dieses Projektes einen einmaligen Einblick in die Regulation der Aktivität von DNMT3A und sie unterstützen ein Modell, in welchem das Enzym unter der kombinierten Kontrolle von MeCP2 und den N-terminalen Modifikationen von Histon H3 steht. Dementsprechend fungiert MeCP2 entweder als Repessor oder als Aktivator der DNA Methylierung.

Im dritten Projekt dieser Arbeit, lag der Fokus schließlich auf der Entwicklung und Anwendung einer neuen Methode, welche es zum ersten Mal möglich machen sollte, epigenetische Modifikationen ortsspezifisch und in lebenden Säugetierzellen zu visualisieren. Dieser dringende technologische Bedarf wurde durch die Entwicklung modularer epigenetischer Biosensoren für die Lebendzellmikroskopie gedeckt, welche auf dem Prinzip der Fluoreszenz Komplementation basieren. In diesen Werkzeugen, wurde die hohe Sequenzspezifität von Ankerproteinen wie ZFs, TALEs und CRISPR/Cas9 mit der großen Vielseitigkeit von Chromatin-Lesedomänen als natürliche Detektormodule für die Erkennung von DNA Methylierung und Histon 3 Lysin 9 Trimethylierung kombiniert. Mit diesem Ansatz war es mir zum ersten Mal möglich, beide Modifikationen gezielt an spezifischen endogenen DNA Sequenzen in unterschiedlichen Maus und humanen Zelllinien zu detektieren. Des Weiteren konnte ich ortsspezifische Signaländerungen dieser Modifikationen nach der Zugabe von Inhibitoren oder der Induktion epigenetischer Enzyme beobachten. Diese Werkzeuge können entweder in ihrer gegenwärtigen Form oder in Kombination mit den momentanen Weiterentwicklungen im Bereich des Genom Targeting und der Mikroskopie unser Verständnis dafür erweitern, wie spezifisch epigenetische Signale,

wie beispielsweise DNA Methylierung, während der embryonalen Entwicklung und der Entstehung von Krankheiten, gesetzt, gelöscht und aufrechterhalten werden.

Zusammengenommen zeigen die Ergebnisse dieser Doktorarbeit wie die synergetische Nutzung biochemischer und zellulärer Methoden es ermöglicht, tiefe Einblicke in das epigenetische Signalnetzwerk, welches die Methylierung von DNA bei Säugetieren reguliert, zu erhalten.

Abstract

The mesmerizing phenotypical and functional diversity of mammalian cell types is to a large extent attributed to epigenetic signals that work together with the DNA sequence to determine gene expression programs. DNA methylation is one of the most important types of epigenetic signals and its paramount role was recognized in early genetic studies. Still, even after decades of active research, a comprehensive understanding of the mechanisms that regulate the chromatin targeting and activity of DNA methyltransferases has not been achieved. In this work, three main directions of research were undertaken, with the ultimate goal of shedding mechanistic and methodological insights into the generation and maintenance of DNA methylation patterns.

In the first project of this thesis, a combination of biochemical and cellular experiments was used to assess the cellular role of the putative chromatin remodeler HELLS, an essential cofactor of DNA methyltransferases. By employing chromatin fractionation assays and microscopy-based techniques, I could show that the ATPase activity of HELLS is necessary for the high nuclear mobility of the protein and its ability to get released from compacted chromatin sites. In addition, the H3K9me3 pathway was also found to play an important role in the exchange of HELLS at heterochromatin. Taken together, this work provides the first evidence for a role of ATP hydrolysis in the association between HELLS and chromatin and hints at a model where the fast exchange of HELLS at repetitive DNA sequences might enhance the local recruitment of epigenetic enzymes, such as DNA methyltransferases (DNMTs). This could subsequently lead to the local stabilization of silencing complexes at heterochromatin.

In the second project of this thesis, the putative interaction between the *de novo* DNA methyltransferase DNMT3A and the 5mC-reading protein MeCP2 was addressed. By building on previous data from our laboratory, which documented a direct interaction between the TRD domain of MeCP2 and the ADD domain of DNMT3A, causing an inhibition of DNMT3A activity *in vitro*, I could show that these proteins also interact in the mouse brain and the inhibitory effect of this interaction is also observed in stable cells lines overexpressing MeCP2. Furthermore, by using conformationally locked DNMT3A variants as novel tools to study the allosteric regulation of this enzyme, I

could elucidate the mechanism of the inhibition of DNMT3A by MeCP2. Accordingly, I found that MeCP2 stabilizes an allosterically closed conformation of DNMT3A, an effect that could be successfully relieved by addition of unmodified histone H3. These results were supported by whole genome bisulfite brain methylome analysis of a *Mecp2* knockout mouse model. Collectively, the findings derived from this project offer unprecedented insights into the regulation of DNMT3A activity and propose a model in which the enzyme is under the combined control of MeCP2 and H3 tail modifications. Accordingly, depending on the modification status of the H3 tail at target sites, MeCP2 can act as either a repressor or activator of DNA methylation.

Finally, in the third project of this thesis, the focus was placed on the development and application of a novel method that would enable for the first time the locus-specific visualization of epigenetic modifications in living mammalian cells. This urgent and unmet technological need was solved by developing a set of modular fluorescence complementation-based epigenetic biosensors for live cell microscopy applications. In these tools, the high DNA sequence specificity of engineered anchor proteins such as ZFs, TALEs, and CRISPR/Cas9 proteins, was combined with the great versatility of chromatin reading domains as natural detector modules of DNA methylation and histone 3 lysine 9 trimethylation. With this approach, I could detect both of these marks for the first time, at defined, endogenous DNA sequences in different mouse and human cell lines. Furthermore, I could follow the changes in the levels of these epigenetic modifications with locus-specific resolution after treatment with epigenetic inhibitors or the induction of epigenetic enzymes. It is anticipated that either in their present form or in combination with the ongoing developments in genomic targeting and microscopy technologies, these tools will greatly improve our understanding of how specific epigenetic signals, like DNA methylation, are set, erased and maintained during embryonic development or onset of disease.

Taken together, the results of this doctoral thesis demonstrate how a synergistic use of biochemical and cellular methods allows to derive deep insights into the epigenetic signaling network centered around the regulation of mammalian DNA methylation.

List of abbreviations

5-aza-dC	5-aza-2'-deoxycytidine
ADD	ATRX-DNMT3-DNMT3L
AdoHyc	S-Adenosyl-L-homocysteine
AdoMet	S-Adenosyl-L-methionine
ADP	adenosine diphosphate
AT-hook	adenine-thymine-hook
ATP	adenosine triphosphate
ATPase domain	adenosinetriphosphatase domain
BiAD	bimolecular anchor detector
BiFC	bimolecular fluorescence complementation
BS treatment	bisulfite
C5V	Cerulean-5-Venus
CD	catalytic domain
CDKL5	cyclin dependent kinase-like 5
CFP	cyan fluorescent protein
CG sites	cytosine-guanine dinucleotides
CGI	CG island
ChIP	chromatin immunoprecipitation
ChIP-seq	chromatin immunoprecipitation followed by sequencing
chromodomain	chromatin organization modifier
CLR4	cryptic loci regulator 4
CMV	cytomegalovirus
CoREST	co-repressor for element-1-silncing transcription factor
CRISPR	Clustered regularly interspaced short palindromic repeats
Cas9	CRISPR associated protein 9
dCas9	nuclease deficient Cas9
Cryo-EM	cryo-electron microscopy
CXXC domain	cysteine rich domain; Cys-X-X-Cys
DAPI	4′,6-Diamidin-2-phenylindol
DARPin	designed ankyrin repeat protein

DDM1	decrease in DNA methylation 1
DIM-5	decrease in DNA methylation 5
DMR	differentially methylated region
DNase	deoxyribonuclease
DNMT	DNA (cytosine-5)-methyltransferase
DRAQ5	1, 5– bis{[2-(di- methylamino) ethyl] amino}- 4, 8-dihydroxyanthracene-9, 10-dione
E12.5	embryonic day 12.5 of mouse development
ECL	enhanced chemiluminescence
EHMT	euchromatin histone methyltransferase
ESC	embryonic stem cell
ESET	ERG-associated protein with SET domain
Fab	fragment antigen binding
FACT	facilitates chromatin transcription
FISH	fluorescence *in situ* hybridization
FRAP	Fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
GATA finger	zinc finger protein binding to GATA DNA sequences
GFP	Green fluorescent protein
GKT box	ATP-binding motif amino acid motif in chromatin remodelers
GLP	G9a-like protein
GluC protease	proteinase which selectively cleaves peptide bonds C-terminal to glutamic acid residues
GST	glutathione S-transferase
H3K36me3	histone H3 tri-methylated on lysine 36
H3K4me3	histone H3 tri-methylated on lysine 4
H3K9me3	histone H3 tri-methylated on lysine 9
HCT116	human colorectal carcinoma cell line
HDAC	histone de-acetylase
HELLS	helicase, lymphoid specific
HIMD	histone modification interacting domain
HP1	heterochromatin protein 1

HRP	horseradish peroxidase
ICM	inner cell mass
IFP1.4	infrared fluorescent protein 1.4
iMEF	immortalized mouse embryonic fibroblasts
IPTG	isopropyl-β-D-thiogalactopyranosid
KMT	lysine methyltransferase
LADs	Lamin B1 attachment domains
LC-ESI-MS/MS	liquid chromatography electrospray ionization tandem mass spectrometry
LSH	lymphoid specific helicase
LSM	laser scanning microscope
MBD	methyl CG-binding domain
MBD1	methyl CG-binding domain protein 1
MBP	maltose-binding protein
mCG	C5 methylated CG dinucletide
MeCP2	methyl CG binding protein 2
mintbody	modification-specific intracellular antibody
MS	mass spectrometry
MTase	methyltransferase
N-CoR	nuclear receptor corepressor 1
Ni-NTA	Ni^{2+} nitrilotriacetic acid agarose
NLS	nuclear localization sequence
NMR	nuclear magnetic resonance
PASG	proliferation-associated SNF2-like gene
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PEI	polyethylenimine
PEV	position effects variegation
PGC	primordial germ cell
PHD finger	plant homeodomain finger
PLA	proximity ligation assay
PRDM9	positive-regulatory domain zinc finger protein 9

pre-crRNA	precursor CRISPR RNA
PTM	post-translational modification
qPCR	quantitative polymerase chain reaction
RFP	Red fluorescent protein
RVD	repeat-variable di-residue
SAM	S-Adenosyl-L-methionine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SET domain	Su(var)3-9, Enhancer of Zeste, Trithorax domain
SETDB1	SET domain bifurcated 1
sgRNA	single guide RNA
Sin3A	Swi-independent 3A
SMARCA6	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 6
SMRT	silencing mediator for retinoid and thyroid receptors
SNF2	switch/sucrose non-fermentable
SRA domain	SET-and-Ring finger-associated domain
SUMO	small ubiquitin-like modifier
SUV39DKO	SUV39H1/H2$^{-/-}$ double knock-out cells
SUV39H	suppressor of variegation 3-9 homolog
TALE	transcription activator-like effector
TET	ten-eleven translocation
TetR	tetracycline repressor
tracrRNA	trans-activating CRISPR RNA
TRD	transcriptional repressor domain
TRF1	telomeric repeat factor 1
UHRF1	ubiquitin like with PHD and RING finger domains 1
VenC	C-terminal part of Venus needed for BiFC
VenN	N-terminal part of Venus needed for BiFC
WGBS	whole genome bisulfite sequencing
YFP	yellow fluorescent protein
ZF protein	zinc finger protein

1 Introduction

1.1 Foundations of Epigenetics

The development of a multicellular organism, with its constituting tissues and organs, is a reproducible cascade of events, which involves the coordination of two main processes: an increase in cellular mass as well as a phenotypic and functional differentiation of the expanding cellular populations (Moris et al., 2016). This arising functional diversity is mesmerizing taking into account that nearly all cells of the organism share the same genetic code and have a common gene repertoire (~ 20,000 in humans). One of the main molecular contributors to this cellular heterogeneity is epigenetics. The term was coined in 1942 by Waddington, and was defined as 'the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being' (Waddington, 2012). This description was used to explain changes in phenotype, for which little mechanistic understanding was available (Allis and Jenuwein, 2016; Waddington, 1942). This theoretical framework was further elaborated to describe embryonic tissue development and differentiation and took shape in an illustration referred to as 'Waddington's epigenetic landscape' (**Figure 1a**). In this, a cell is represented by a marble beginning to roll off at the top of a mountain valley and following one of the several existing paths down the landscape. At the end of its trajectory, the marble comes to rest into one of the depicted local energy minima, representing one defined differentiation state. The further the cell travels down the branching points, the more restricted its differential potential becomes. Importantly, the trajectory taken by the cell is non-random. This is guided by specific gene regulatory networks, which chart the shape of the epigenetic landscape, and determine the path of the marble at the branching points (**Figure 1b**).

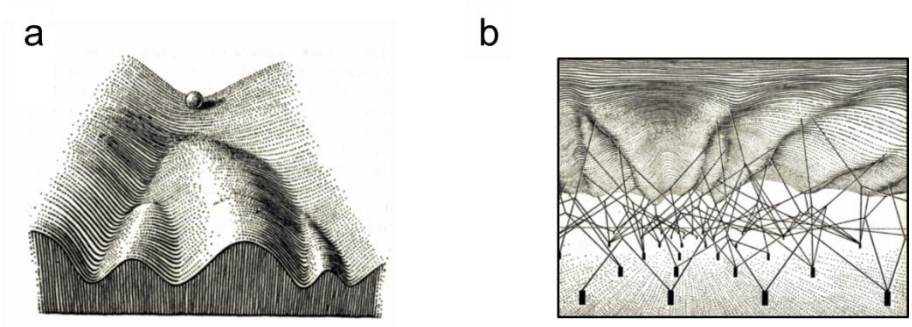


Figure 1| Waddington's epigenetic landscape. a) In Waddington's illustration, the cell, depicted as a marble, travels down the landscape by opting for a specific set of branching points, which stand for developmental decision events. At the end of the journey the cell has reached its full differentiation potential. **b)** The landscape is underpinned by gene regulatory networks, depicted as pegs under the hills and valleys. Both images were adapted after (Moris et al., 2016)

Today, we would argue that Waddington's landscape is a metaphor for how alternative cellular programs can be realized based on epigenetic principles.

With the booming development of the field, the definition of the term epigenetics was further refined and gradually narrowed to ''the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence.'' (Wu and Morris, 2001). Noteworthy, by contrast to Waddington's definition, the new conceptualization makes heritability a necessary part of the epigenetic notion (Deans and Maggert, 2015).

We now know that epigenetic signals fundamentally control cellular diversity by transducing the inheritance of gene expression patterns. This is achieved by regulating the local interpretation of the genetic code and not by modifying the underlying DNA sequence (Allis and Jenuwein, 2016). As a result, even if all cells share the same genetic code including the common repertoire of genes, only a subset of genes will be expressed in a particular cell type at a given developmental stage. Three main categories of signals were described to be essential for the establishment of a stably inheritable epigenetic state (**Figure 2**) (Berger et al., 2009). At the top of the cascade, the 'Epigenator' is found. This is a signal that originates from the extracellular environment and is able to activate an intracellular pathway. It acts as a priming factor for the 'Epigenetic Initiator', which responds by targeting a specific chromosomal locus. Here it is able to instruct a change in the local epigenetic environment. This newly

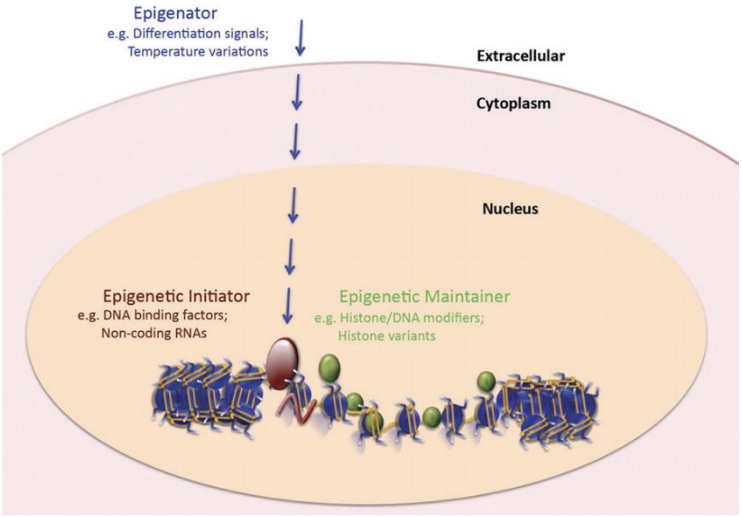


Figure 2| The epigenetic pathway. The epigenetic pathway relies on three categories of signals, which contribute to the establishment of a stably inherited epigenetic state. The 'Epigenator' (blue) originates from the environment and triggers the signaling cascade. In response to this, the 'Epigenetic initiator' (red), which has DNA sequence specificity, targets a specific chromosomal locus where it instructs changes in the epigenetic landscape. The inheritance of these changes is taken over by the 'Epigenetic Maintainer' (green). The image was taken from (Berger et al., 2009).

enforced state is finally sustained by ‘Epigenetic Maintainers’. These factors do not have absolute DNA sequence specificity and are dependent on initiators for locus-specific recruitment. Hence, while maintainers are indispensable for sustaining the local epigenetic state, they are not capable of initiating it.

1.2 Chromatin organization: regulator of the genetic code

Each human cell contains approximately 2 meters of DNA if stretched end-to-end; yet the nucleus, the organelle where this genetic information is stored, is only about 10 µm in diameter. A back-of-the-envelope calculation reveals that the human genetic code must be condensed by circa 10^5-fold in order to fit into the nucleus. To resolve this immense spatial constrain, the genome is tightly compacted by a dedicated set of specialized and highly conserved proteins, which are able to bind to and fold DNA into structures with increasingly higher levels of organization (**Figure 3**) (Li and Zhu, 2015). This hierarchically packed nucleoprotein complex is referred to as chromatin (from the Greek ‘khroma’, colored, because of its staining properties (Filion et al., 2011)). This structure establishes a key control mechanism of gene expression in eukaryotes (Luger et al., 2012). The outmost importance of chromatin architecture and compaction is immediately evident from the dry matter composition of the yeast interphase nucleus, which consists of about 70-80% protein, 20-30% RNA and only ~2% DNA (Rozijn and Tonino, 1964).

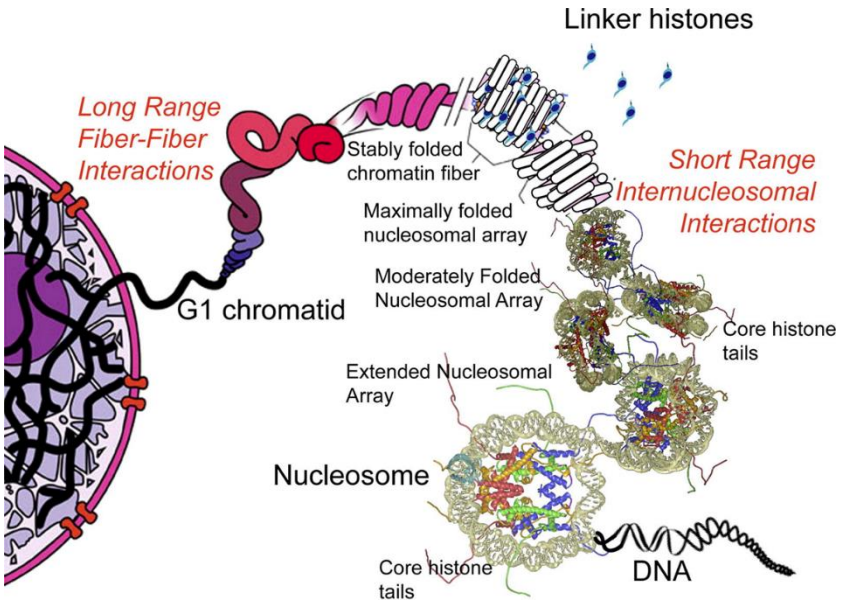


Figure 3| Higher order packaging of the genetic material within the nucleus of a eukaryotic cell. The DNA is serially folded intro structures with increasingly higher orders of compaction, depicted here in a schematic manner. At low-compaction levels, atomic structures are shown. The image was taken from (Chakravarthy et al., 2005)**.**

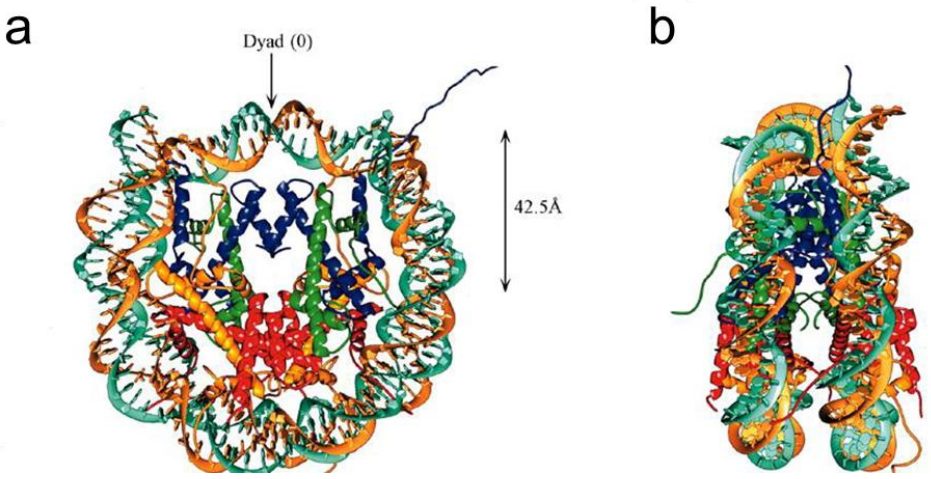


Figure 4| The structure of the nucleosome core particle. a) View of the core particle down the DNA superhelix axis. Each type of histone protein is color-coded and shown as ribbon traces. **b**) View of the core particle perpendicular to the superhelical axis. The image was taken from (Dutnall and Ramakrishnan, 1997)**.**

The basic building block of the chromatin fiber is the nucleosome (**Figure 4**). This elicits an initial ~ 7-fold linear compaction of genomic DNA and consists out of 147 bp DNA wrapped in 1, 7 superhelical turns around an octamer containing two copies of each of the four core histone proteins: H2A, H2B, H3, H4 (Luger et al., 1997). The high affinity of histone proteins for DNA give rise to a highly stable structure that restricts the access of DNA-binding proteins and acts as an inherent barrier for nuclear processes such as transcription, DNA replication, and DNA repair (Grigoryev and Woodcock, 2012). An initial high-resolution X-ray structure of the nucleosome core revealed that two main phosphates per strand per helical turn are responsible for making direct contact between the DNA backbone and the core histone proteins. The DNA helix forms contacts primarily with the paired-loop and paired-end-of-helix elements in the histone fold domains. Highly conserved, arginine residues point towards the minor groove to facilitate DNA positioning, bending and superhelical shape (Davey et al., 2002; Luger et al., 2012; Luger et al., 1997).

Interestingly, in spite of this stability, the nucleosomal structure is also highly dynamic. *In vitro* and in living cells, nucleosomes were shown to be subjected to thermal fluctuations that induce a spontaneous partial unwrapping of the compacted DNA, in a process known as ‘chromatin breathing’. This is a short-lived structural state during which the last 10-20 bps of DNA that make contacts to the histone octamer, dissociate, thereby transiently exposing the underlying DNA sequence for binding by high-affinity DNA-binding factors (Luger et al., 2012). In addition to these passive events, the accessibility of the nucleosomal sequence can also be actively regulated by multi-

subunit ATP-dependent chromatin remodeling complexes. These enzymes use the energy derived from ATP hydrolysis to alter the nucleosome structure or positioning, thereby regulating the access of transcription factors to their cognate DNA sites (Hota and Bruneau, 2016).

Neighboring nucleosomes are connected by 10-70 bps of linker DNA in a 'beads-on-a-string' highly flexible polymer (**Figure 5a**) (Olins and Olins, 1974). Chromatin architectural proteins, such as the linker histone H1, further promote and stabilize the condensation of this 10-nm array into an ordered and highly important secondary structure: the 30-nm fiber (Li and Reinberg, 2011). This serves as a basis for additional loops or hierarchical coiled structures that mature in the form of single chromosomes (**Figure 3**) (Razin and Gavrilov, 2014). Despite considerable efforts during the last three decades, however, the precise molecular organization of the 'bead-on-a-string' arrays into the secondary structure of chromatin, the 30-nm fiber, remained a matter of intense debate (Grigoryev and Woodcock, 2012; Lieberman-aiden et al., 2009; Rao et al., 2014; Travers, 2014).

The 30-nm fiber was first visualized by electron microscopy, upon Mg^{2+} addition to chromatin released from chicken erythrocytes (Razin and Gavrilov, 2014). Subsequent X-ray diffraction studies of the released chromatin proposed that the 30-nm fiber is organized as a solenoid or a 'one-start' helix. Here the nucleosomes coil around a central cavity with ~ 6 nucleosomes per turn, so that each nucleosome in the fiber interacts with its fifth and sixth neighbor nucleosomes (**Figure 5b**) (Finch and Klug, 1976). Follow up electron microscopy studies, using chromatin that has been snap-frozen immediately after its release from the nucleus, lead the formulation of an alternative model: the zig-zag or 'the two-start' helix. Here, each nucleosome binds to its second neighbor (**Figure 5b**) (Woodcock et al., 1984). In spite of this progress, the heterogeneous properties of nucleosomes in native chromatin with different sequences/linker lengths and different histone compositions/modifications make it difficult to precisely deconvolute the detailed structure of chromatin fibers (Zhu and Li, 2016).

The need of chromatin with defined composition was addressed by the development of new methods that allowed the *in vitro* assembly of designer chromatin fibers using unmodified, recombinant histones and regular tandem repeats of unique nucleosome positioning DNA sequences (Dorigo et al., 2004). This advancement facilitated the

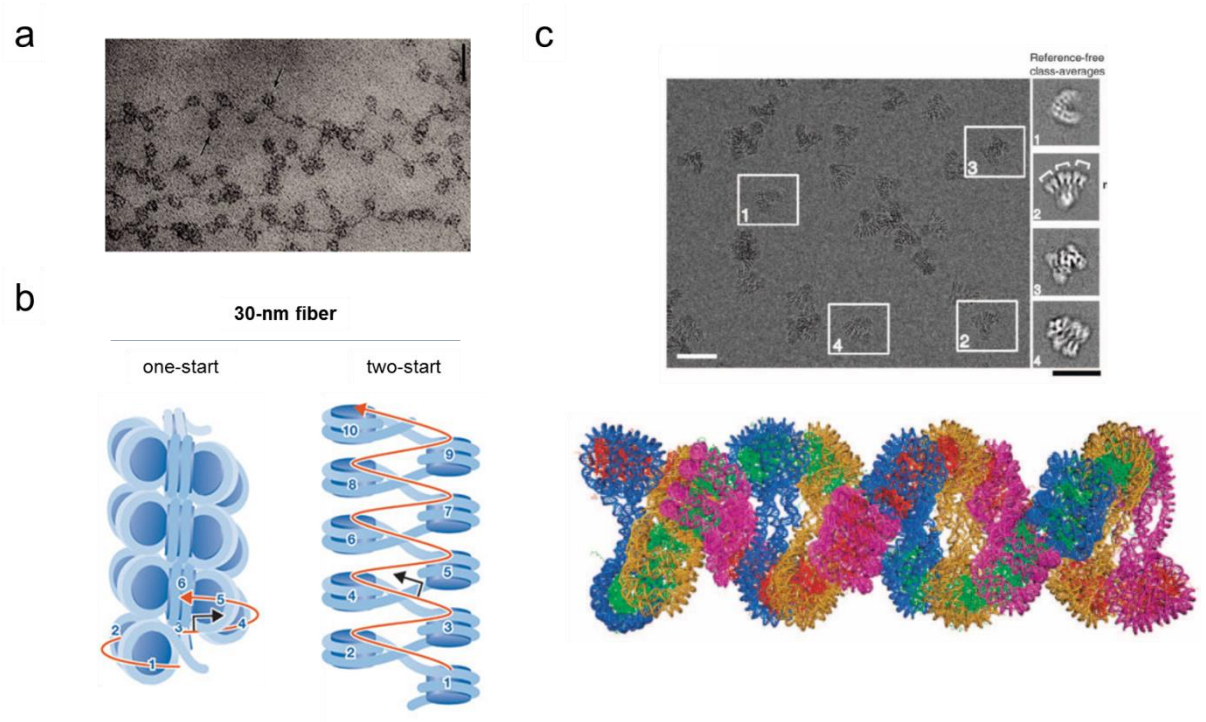


Figure 5| Arrangement of nucleosomes into chromatin fibers with increasingly high levels of compaction. a) Low-ionic strength chromatin spread showing the 'beads on a string' 10-nm fiber. The Image was taken from (Olins and Olins, 2003). **b)** Schematic representation of the one-start (left) and two-start (right) 30-nm fiber. The numbers denote the nucleosomes in the arrays, while the red line indicates the path of the DNA. The image was adapted after (Quenet et al., 2012). **c)** Top: a representative cryo-EM micrograph of 30-nm fibers reconstituted on 12 x 187 bp DNA. The zoom-ins display four selected unsupervised classification generated areas. Bottom: a pseudo-atomic model build by directly stacking the cryo-EM structure of the fibers reconstituted on 12 x 187 bp DNA, on top of each other to form a continuous fiber. These images were taken from (Song et al., 2014).

dissection of the key parameters that are important for the assembly of the chromatin fiber (Zhu and Li, 2016). Using this system, Song and colleagues have recently provided fundamental structural insight into the composition of the 30-nm fiber (Song et al., 2014). By employing 3D cryo-EM on *in vitro* reconstituted arrays of 12 nucleosomes, assembled in the presence of linker histone H1 and with different nucleosome repeat lengths, the authors succeeded for the first time to unequivocally identify the path followed by the linker DNA and pinpoint the arrangement of the individual nucleosomes within the 30-nm fiber. These structures showed a histone H1 dependent left-handed twist of repeating tetranucleosomal structural subunits, within which the four nucleosomes zigzag back and forth with a straight linker DNA (**Figure 5c**). These study resolved the fundamental issue as to whether the 30-nm fiber is built as a one-start or a two-start model, and revealed that the chromatin fiber is double-helical, like the DNA it packages (Song et al., 2014). This biochemical work recently received support from three studies showing that nucleosomes are organized into discrete ''nucleosome clutches'' or ''tetranucleosomal folding motifs'' along the

chromatin fiber also *in vivo* (Grigoryev et al., 2016; Hsieh et al., 2015; Ricci et al., 2015). The functional role of the tetranucleosomal unit was recently addressed by Li and colleagues (2016), using biophysical and genome-wide analysis. With these techniques, the authors revealed that the tetranucleosome is a stable secondary structure during hierarchical *in situ* condensation of chromatin fibers and that its stability is regulated by the histone chaperone FACT (Li et al., 2016). Taken together, the data included in this study indicate that the tetranucleosome might provide an additional level of gene regulation beyond the nucleosome.

Although the higher order hierarchical folding of the chromatin fiber beyond the 30-nm scale, remains a matter of intense dispute, the major factors that affect chromatin organization have been elucidated. These epigenetic signals are DNA methylation, histone posttranslational modifications (PTMs) and non-coding RNAs (ncRNAs). These landmarks, together with ATP-dependent chromatin remodeling complexes lay the basis of an interconnected and highly dynamic epigenetic landscape and contribute to a fascinating increase in the information storage capacity of the genetic code. In the following sections, a more detail view on DNA and histone modifications as well as on chromatin remodeling complexes will be provided.

1.3 Histone modifications and histone modifying enzymes

1.3.1 Histone modification types

Pioneering experiments performed by Allfrey, using acetate-2-C^{14} and methionine-methyl-C^{14} as precursors, revealed already in the mid-1960s that histone proteins can be post-translationally modified and in particular acetylated (Allfrey et al., 1964). Since then, an astonishingly large number of post-translational modification (PTMs) types have been identified and mapped, such as lysine acetylation and methylation, phosphorylation, ubiquitination (Ub), ADP-ribosylation, arginine methylation, citrullination, and SUMOylation (**Figure 6**) (Arnaudo and Garcia, 2013). Noteworthy, while modern mass spectrometry methods have discovered more than 200 different flavors of histone modifications, not all of these PTMs are equally abundant indicating that for some, the functional roles might be rather discrete (Leroy et al., 2013; Rothbart and Strahl, 2014). Out of the 20 standard amino acids coded by the ribosome, lysine stands out as the hotspot for enzymatic PTMs. This essential α-amino acid is usually located in the outer hydrophilic surface of proteins, where its flexible side-chain that contains a terminal ε-amino group, is solvent exposed and forms ionic interactions and

hydrogens bonds in proteins, DNA and by interacting with water molecules. Lysine is not only the most modified amino acid, but also the one which is subjected to the widest range of PTMs (**Figure 6**) (Andrews et al., 2016; Bischoff and Schlüter, 2012; Mann and Jensen, 2003). Among these, acetylation and methylation of lysine residues have emerged as two of the key modulators of chromatin-templated processes and will be addressed in further detail in the following sections.

A first indication of the regulatory role of histone PTMs was provided by the high-resolution X-ray structure of the nucleosome. This revealed that histone proteins have disordered N and C-terminal tails, which protrude out of the nucleosomal core and can make contacts with neighboring nucleosomes (**Figure 4**) (Luger et al., 1997). This lead

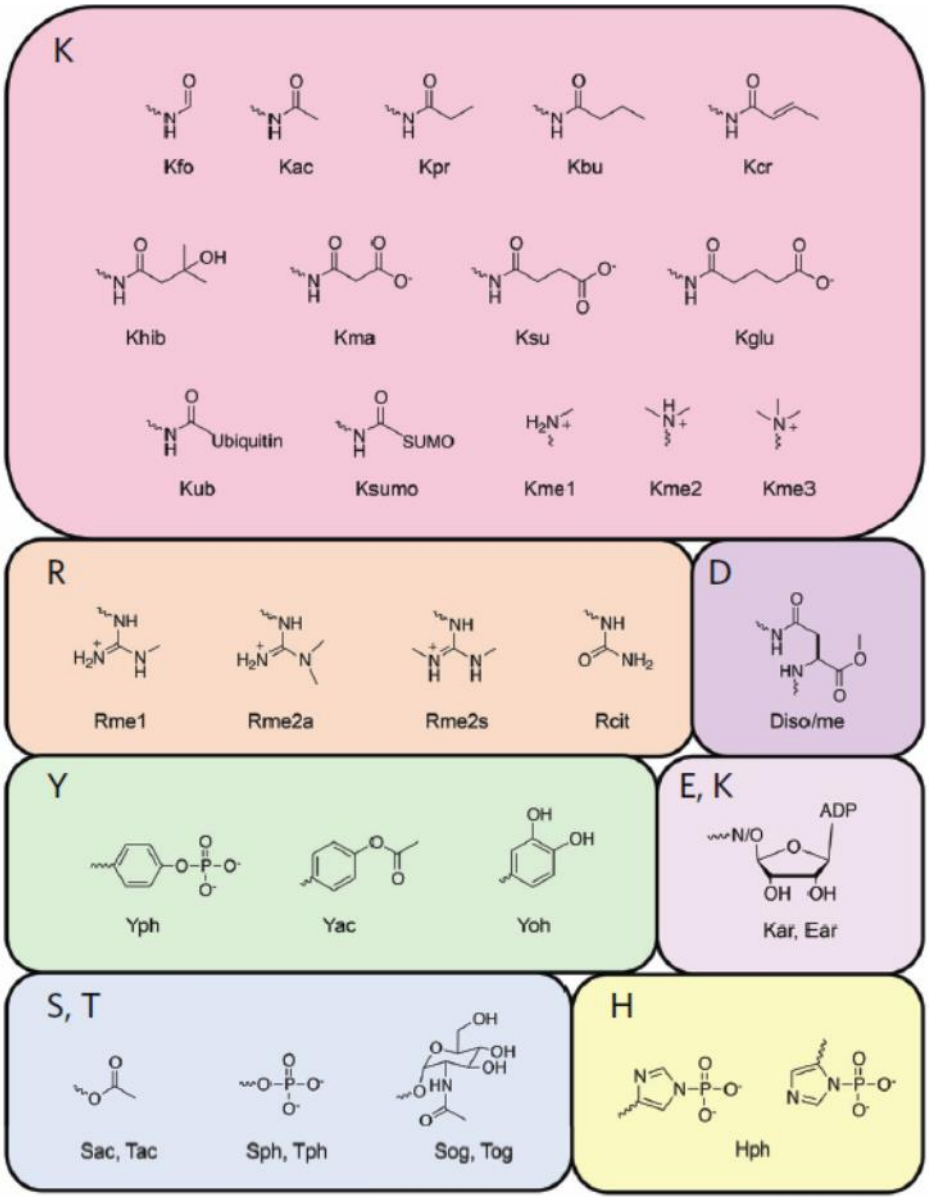


Figure 6| Modifications found on histone proteins. fo, formylation; ma, malonylation; su, succinylation; glu, glutarylation; ub, ubiquitination; cit, citrullination; oh, hydroxylation; ar, ADP ribosylation; og, O-GlcNAcylation. The image was taken from (Andrews et al., 2016).

to the hypothesis that the post-translational modification of histone proteins might mediate inter-nucleosomal interactions and thereby influence chromatin organization (Bannister and Kouzarides, 2011). A plethora of follow up investigations have demonstrated that PTM of histone proteins can indeed directly regulate histone-histone and histone-DNA contacts, particularly when occurring in the histone fold or the globular domains (Cosgrove and Wolberger, 2005; Cosgrove et al., 2004; Rothbart and Strahl, 2014). For instance, elegant *in vitro* experiments employing semi-synthetic or genetically modified nucleosomes showed that the acetylation of the octamer core on histone H3 at lysines 56,122 and 155, weakens histone-DNA interactions leading to increased nucleosome mobility and DNA unwrapping (Manohar et al., 2009; Neumann et al., 2009; Shimko et al., 2011; Simon et al., 2011). Notably, acetylation is not limited to the histone core, but can also be found on numerous lysines residues within the histone tail, such as H3K9, H3K14, H3K18, H4K5, H4K8, and H4K12. This hyperacetylation reduces the positive charge of histone proteins, disrupting the electrostatic interactions between histone and DNA (Bannister and Kouzarides, 2011; Kouzarides, 2007). Apart from the additive charge neutralization effect, histone acetylation, in particular at H4K16, can also have a direct effect on higher-order chromatin structure. This mark perturbs the interaction between the unmodified tail of H4 and the acidic patch on H2A, inhibiting the formation of the 30-nm fiber (Horikoshi et al., 2013; Shogren-Knaak et al., 2006). In addition to the direct physical effect of some PTMs on chromatin structure, a large body of work demonstrated that the high regulatory power of PTMs resides in their ability to recruit protein effectors (also referred to as histone binding domains or readers of PTMs) (Musselman et al., 2012). These specialized proteins are able to dock onto specific histone PTMs and direct defined downstream events on chromatin. This regulatory mechanism is particularly potent in the case of lysine methylation. Unlike other modifications, such as acetylation, methylation does not lead to a change in the overall charge of the lysine. Instead, it alters the hydrophobic character and the size of the modified residue. This change in properties is detected by PTM readers and translated into downstream biological effects. Among the large number of histone PTMs, lysine methylation is one of the most versatile types of chemical modifications. This is because the flexible terminal ε-amino group can accompany 4 different chemical states. As such, the residue can be found either in an unmodified form (me0) or accommodating up to three methyl groups (me1, me2, and me3). Depending on the methylation index and the position of the

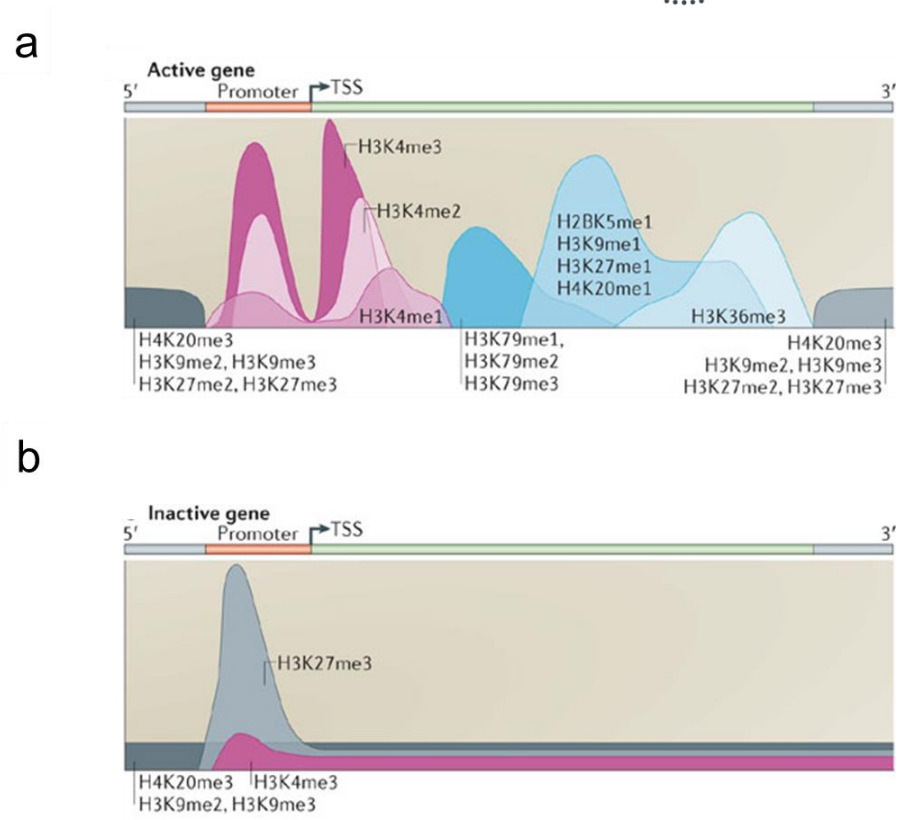


Figure 7| Overview of the main types of histone modifications and their distribution over a gene locus. A representative distribution is shown for both a transcriptionally active (**a**) and a silenced (**b**) gene. The image was taken from (Kooistra and Helin, 2012). Importantly, the quality of the plotted profiles heavily depend on the specificity the used antibodies used in the assay (Kungulovski et al., 2015a).

modified lysine within the histone tail, the epigenetic mark can have different effects on gene regulation (Martin and Zhang, 2005). The canonical sites for lysine methylation comprise six residues on histone H3 (K4, K9, K26, K27, K36, and K79), K20 of histone H4 and K26 of histone H1. Apart from K79, these are all located in the N-terminal tail of histone proteins (Musselman et al., 2012).Interestingly, modifications of different lysines can have opposing regulatory roles. For instance, H3K4me3 and H3K36me3 marks are strongly associated with active genes, when found at promoter or within gene bodies, respectively (**Figure 7a**). By contrast, di and trimethylation at position K9 or K27 are marks that abundantly decorate silenced genes (**Figure 7b**) (Zhou et al., 2011). Due to its high relevance for the present work, the pathways responsible for setting, reading and maintaining of the H3K9me3 modification will be covered in depth, in the next section.

1.3.2 The H3K9me3 pathway and its involvement in constitutive heterochromatin formation

The H3K9me3 mark has important regulatory roles in the repression of both genic and intergenic regions of metazoan genomes (Kim and Kim, 2012). This modification can

be introduced by several enzymes, including SETDB1 (or ESET), SUV39H1, SUV39H2, EHMT1 (GLP), and EHM2 (G9A) (Kim and Kim, 2012). Among these, SETDB1 is highly expressed in embryonic stem (ES) cells and involved in maintaining the pluripotency and self-renewal properties of ES cells. While EHMT1/EHM2 are active at genic regions found in open chromatin, broadly referred to as euchromatin, the SUV39H1 enzyme together with its homologue SUV39H2 are involved in the establishment and maintenance of H3K9me3 at constitutive heterochromatin (Kim and Kim, 2012; Peters et al., 2001; Tachibana et al., 2002; Yuan et al., 2009). This type of highly compacted chromatin is mainly formed at the gene-poor regions around the centers or at the ends of mammalian chromosomes, rich in tandemly repeated DNA sequences. The presence of this repressive histone mark at this sites is essential for preventing harmful genomic rearrangements provoked by spontaneous recombination events of the repetitive DNA (Saksouk et al., 2015).

Su(var) genes were initially discovered by genetic screens on centromeric position effects in *Drosophila melanogaster* (Reute and Spierer, 1992) and *Saccharomyces pombe* (Allshire et al., 1995). The products of *Su(var)* genes suppress position effects variegation (PEV), and were reported to be involved in the formation of silent chromatin domains (Wallrath, 1998). SUV39H1 was the first identified histone-specific lysine methyltransferase (KMT) (Rea et al., 2000). The enzyme belongs to the family of conserved S-adenosyl-L-methionine (AdoMet) dependent KMTs. It possesses a conserved, catalytically active SET domain, surrounded by a pre-and post-SET domain important for regulation the catalytic activity (Dillon et al., 2005). The crystal structure of the *S. pombe* CLR4 homologue (Min et al., 2002) as well as other members of the family (Dillon et al., 2005), revealed that the SET domain forms a β fold that contains a series of curved β-strands that form a knot-like structure, packed together with the pre-and post-SET domains. Characteristic for SET-domain proteins is the fact that the binding site for the histone substrate and the cofactor AdoMet are located on opposite faces of the SET domain. The target lysine is inserted into a deep and narrow hydrophobic channel that runs through the core of the domain, such that the target nitrogen is positioned in close proximity to the methyl donor (**Figure 8**) (Dillon et al., 2005). Biochemical work using the *Neurospora crassa* DIM-5 homolog, demonstrated that this unusual arrangement is important for the processive methylation of the target lysine by the SET superfamily members (Zhang et al., 2003).

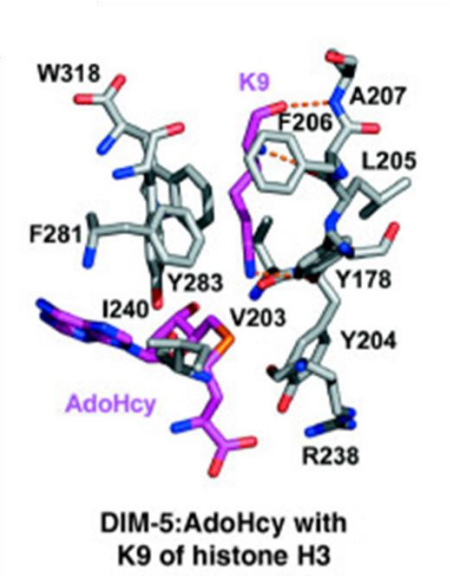


Figure 8| Structure of the DIM-5 SET domain, in complex with the H3K9me3 peptide and cofactor byproduct S-adenosyl-L-homocysteine (AdoHcy). The peptide and AdoHcy are depicted in pink. Hydrogen bonds are shown as dashed lines. The image was taken from (Dillon et al., 2005).

In addition to the catalytically active SET domain at the C terminus, SUV39H enzymes also contain a chromodomain in their N-terminal part. The crystal structure of the chromodomain of SUV39H1 showed that it is able to recognize and bind to K9me2/me3-modifications on histone H3 (Wang et al., 2012). Fluorescence polarization assays revealed an intermediate affinity binding with a Kd of 20±4 µM for H3K9me3, and a Kd of 29±12 µM for H3K9me2. No binding to histone H3K9me0/1 was observed (Wang et al., 2012). Like in other methyllysine readers, the modified lysine is recognized by the hydrophobic compact core within the chromodomain, that builds an aromatic cage consisting out of residues V45, L48, Y60, V62, W64, Y67, L80, I85, and L86. All of these amino acids are also conserved in the chromodomain family (Wang et al., 2012). With the chromodomain crystallization, it became apparent that within the SUV39H1 enzyme two important functions are merged: the setting and the read-out of repressive H3K9me2/3 marks. The importance of this connection was underlined by mutation studies reporting that either deletion of the chromodomain or point mutation of the conserved amino acids, W64A or Y67A, impaired the activity of the enzyme even if its catalytic SET domain was intact (Chin et al., 2006). In the following section, the biological relevance of this connection will be addressed, with a particular focus on pericentromeric heterochromatin, as one of the major target sites of SUV39H enzymes.

1.3.3 The involvement of the H3K9me3 pathway in the formation and maintenance of heterochromatin at pericentromeres

Constitute heterochromatin is formed at gene-poor genomic regions, which are believed to be conserved between cell types. This is conceptually and mechanistically contrasting to the second branch of repressed chromatin, facultative heterochromatin, which involves silencing of genes that have to be repressed upon environmental cues (Saksouk et al., 2015). The bulk of constitutive chromatin forms at pericentromeric regions. These areas of the genome are abundant in repetitive DNA sequences. The sequence of these repeats greatly varies between different chromosomes as well as between organisms, indicating that the organization, rather than a specific DNA motif, is important for the pericentromeric function (**Figure 9**) (Saksouk et al., 2015). For instance, mice centromeres consist mainly of minor satellites, while pericentromeres are built out of major satellite repeats. By contrast, in humans, centromeres consist predominantly of α satellites, while pericentromeres have chromosome-specific satellite sequences, which include satellite I, II and III motifs (Saksouk et al., 2015). Regardless of their DNA sequence, the unifying theme is that due to their repetitive content, these genomic sites need be repressed. Indeed, in several abnormal situations, like cancer, defects in the heterochromatic signaling pathway can result in deleterious chromosomal rearrangements, involving pericentromeric regions (Ehrlich, 2003). The importance of these genomic elements is further underlined by proteomic studies, discovering around 50 different chromatin regulators that are enriched at pericentromeric heterochromatin and responsible for maintaining its compact structure (Saksouk et al., 2015; Soldi and Bonaldi, 2013).

The role of SUV39H enzymes in the regulation of pericentromeric heterochromatin was highlighted by genetic studies. These revealed a selective depletion of H3K9me3 at

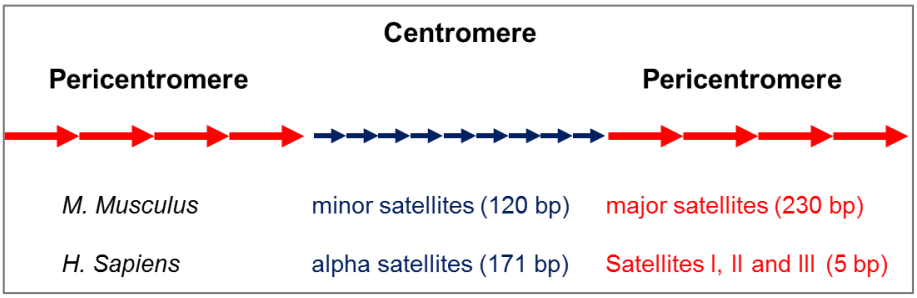


Figure 9| Organization of constitutive heterochromatin in *M. Musculus and H. Sapiens*. Schematic representation of repeat composition of centromeric and pericentromeric chromatin domains. The approximate length of the different repetitive elements is indicated. The image was adapted after (Saksouk et al., 2015)

pericentromeric regions but not at other genomic sites marked by this modification, which hinted towards a specific targeting mechanism. The loss of the silencing mark was accompanied by a local increase in transcription (James and Elgin, 1986; Saksouk et al., 2015). While *Suv39*$^{-/-}$ embryos survive the initial developmental stages, indicating that the enzyme does not play a role in early development, they do show some lethality in the later embryonic stages as well as decreased viability. This is molecularly underpinned by abnormal segregation effects of chromosomes followed by disruption of spermatogenesis and increased cancer risk (Peters et al., 2001).

While the steps that are involved in the initial recruitment of SUV39H to chromatin, remain to be elucidated, the cascade of events responsible for the sustained maintenance of H3K9me3 at pericentromeric sites is clearer. This relies on the interplay between SUV39H and heterochromatin protein 1 (HP1), a small but very versatile chromosomal protein. HP1 contains a chromodomain that is able to bind to H3K9me3 with higher affinity than the corresponding chromodomain of SUV39H (**Figure 10a**) (Bannister et al., 2001; Jacobs et al., 2002). The binding pocket for the N-methyl group becomes ordered upon peptide binding and consists of three conserved aromatic amino acids, Y24, W45, and Y48, forming a so-called aromatic cage into which the ammonium group is inserted (**Figure 10b**). The crystal structure of the chromodomain in complex with the methylated histone peptide revealed that the H3 tail binds through an induced-fit mechanism and adopts an extended β-strand-like conformation. The peptide lies antiparallel with two β regions of the chromodomain,

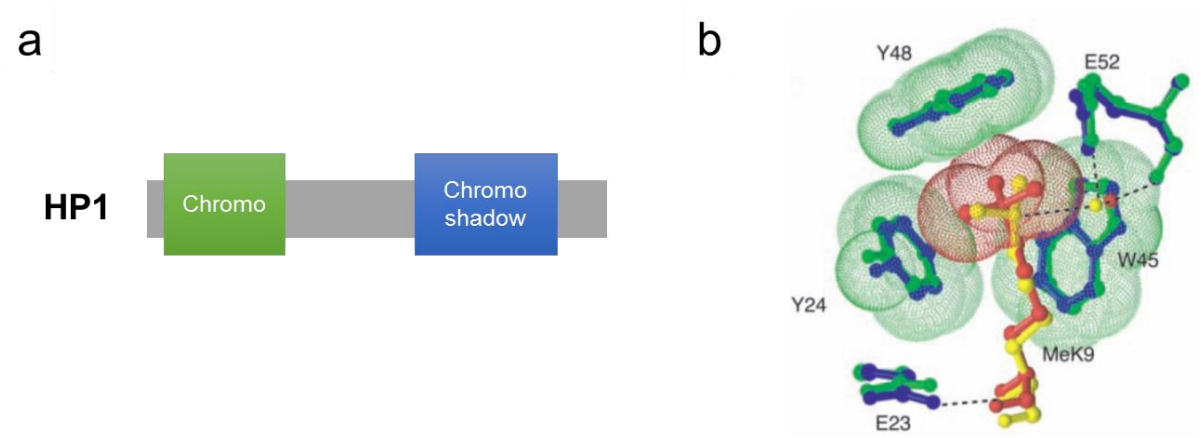


Figure 10| HP1, a specific reader of H3K9me3. a) Schematic representation of the HP1 domain composition. The chromodomain, responsible for H3K9me3 binding, and the chromoshadow domain, important for homodimerization, are depicted in green and blue, respectively. **b)** Stereogram showing the K9me2 (yellow) and K9me3 (red) – modified H3 peptides, in complex with the HP1 chromodomain (blue and green, respectively). Dashed lines indicate the van der Waals contacts between the domain and the peptide. The image was taken from (Jacobs et al., 2002).

completing a three-stranded β sheet and the overall β-sandwitch fold (Jacobs et al., 2002; Nielsen et al., 2002).

The initial localization of HP1 at pericentromeric heterochromatin might dependent on its interaction with RNA. Indeed, sense-oriented repeat transcripts were shown to bind to the HP1 hinge domain (Maison et al., 2002, 2011; Mozzetta et al., 2015; Muchardt et al., 2002). Through direct protein-protein interaction between the HP1 chromoshadow domain and the SUV39H N-terminus, the enzyme is recruited to chromatin (Schotta et al., 2002; Yamamoto and Sonoda, 2003). This is followed by the introduction of novel H3K9me3 marks which further stabilize the binding of SUV39H through its own chromodomain. This results in an increased recruitment of HP1, the association of which can be stabilized through the formation of homodimers via its chromoshadow domain (Yamamoto and Sonoda, 2003). This positive feedback loop results in self-propagating cycles of H3K9me3, causing the spreading of the heterochromatin domain (**Figure 11**) (Wang et al., 2014). The importance of the feedback loop between HP1 and SUV39H becomes apparent from mutational and loss of function studies. Accordingly, in *Suv39h1h2*$^{-/-}$ double knockout (SUV39DKO) cells, the enrichment of HP1 at pericentromeric heterochromatin was dramatically reduced (Peters et al., 2001), while reintroduction of a HP1 interacting-deficient SUV39H in these cells, resulted in diminished accumulation of the enzyme at pericentromeres, and low recovery of H3K9me3 (Muramatsu et al., 2016). This effect was dependent on the expression levels of the enzyme, suggesting that HP1 is not the only factor responsible for SUV39H recruitment at these sites.

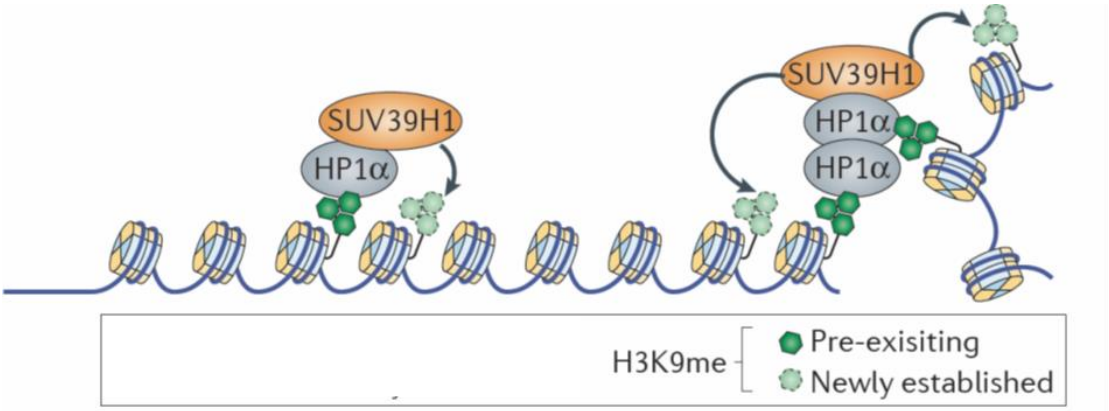


Figure 11| SUV39H-mediated spreading of the H3K9me3 mark. The stable HP1-SUV39H1 complex, subsequentially reads and sets the silencing trimethyl mark on H3K9, in a positive feedback loop that leads to the formation of broad heterochromatic domains and spreading of the silencing signal. This process is supported by the ability of HP1 to form homodimers, through its chromoshadow domain. The image was adapted after (Mozetta et al., 2015).

DNA methylation is another essential epigenetic signal that regulates chromatin-templated processes, in addition to the posttranslational modification of histone tails. In the following chapter, the biological role of this mark, as well as the setting, control, and readout of the signal will be covered in more detail.

1.4 DNA methylation

DNA methylation was first discovered in 1948, by Hotchkiss who used paper chromatography to investigate the purine and pyrimidine base composition of calf thymus DNA (Hotchkiss, 1948). Since then, the existence and role of this initially mysterious mark have become objects of intense investigation. This modification occurs predominantly on the fifth carbon of cytosine bases, 5mC, and can be found in eukaryotes ranging from fungi to vertebrates. While, the significance and function of DNA methylation vary greatly between organism, a large body of work has uncovered that for mammals, DNA methylation is a vital modification, which plays essential roles in embryonic development and adult tissue homeostasis (Messerschmidt et al., 2014).

In mammals, DNA methylation occurs predominantly at palindromic CG sites, on both DNA strands. Out of the 56 million CG sites, 60-80% are methylated, which corresponds to 4-6% of all cytosines (Laurent et al., 2010; Lister et al., 2009). Noteworthy, modified CG sites are mutagenic hotspots, since their hydrolytic deamination can give rise to TG mismatches (Jurkowska et al., 2011). By contrast, spontaneous deamination of unmethylated cytosine bases occurs 2-4 times more slowly *in vitro*, and since this gives rise to uracil, an unnatural DNA base, it can be more readily recognized and corrected by dedicated repair systems, in cells (Krokan et al., 2000; Shen et al., 1994). The mutagenic potential of C5 methylated cytosine led to a selective depletion of CG sites from mammalian genomes during molecular evolution. For instance, in the human genome CG dinucleotides are underrepresented by a factor of 5 in comparison to other dinucleotide combinations (Jurkowska et al., 2011). While being globally underrepresented in inter- and intra-genic regions, CG dinucleotides are abundant at repetitive DNA elements gene regulatory elements. For instance, in gene promoters, they cluster in CG islands (CGI), 500-2000 bp long regions that have a GC content above 50% and a ratio of observed-to-expected number of CpG dinucleotides above 0.6 (Gardiner-Garden and Frommer, 1987; Takai and Jones, 2004). 5mC displays a bimodal distribution. As such, while 60%-90% of ‘single’ CG sites are methylated (depending on the cell type), CGI remain

predominantly hypomethylated (Deaton and Bird, 2011; Messerschmidt et al., 2014). CG islands are found in the promoter regions of about 70% of all human genes including most housekeeping and tissue-specific genes (Saxonov et al., 2006). The survival of these CG-rich clusters in the mammalian genome might be explained by the fact that in the germline, CGIs have no or little methylation, thereby posing little mutagenic drift risk at this stage (Shen et al., 2007; Weber et al., 2007).

The methylation status of CG sites changes dynamically during cellular differentiation. While CG-dense promoters are unmethylated in the germline, some progressively gain methylation during development. This enforces a stable gene repressive state as the cells become increasingly committed (Borgel et al., 2010; Meissner et al., 2008). Repetitive elements such as pericentromeric repeats are also abundantly decorated with 5mC. At these sites, the mark works in concert with the H3K9me3 pathway to repress their latent transcriptional activity, and ensure proper chromosome alignment and segregation (Messerschmidt et al., 2014). Apart from its regulatory role at CGIs and repetitive elements, DNA methylation plays important roles in processes such as X-chromosome inactivation in females, parent-of-origin specific gene expression through imprinting, alternative splicing and enhancer activity (Anastasiadou et al., 2011; Aran et al., 2013; Bartolomei, 2009; Ferguson-Smith, 2011; Hellman and Chess, 2007; Jones, 2012).

1.4.1 The mammalian DNA methylation machinery

DNA methylation is introduced by the family of DNA methyltransferases (DNA MTases or DNMTs) which comprises 4 active enzymes DNTM1, DNMT3A, DNMT3B, DNMT3C and the related but inactive regulatory factor DNMT3L (**Figure 12**) (Jurkowska et al., 2011). The recently identified DNMT3C evolved via a duplication of *Dnmt3b* in rodent genomes and is found in male germ cells (Barau et al., 2016). From the structural composition, mammalian MTAses can be broadly subdivided into two parts: a large multidomain N-terminal part with regulatory functions, and a C-terminal catalytic part. The latter is conserved between prokaryotic and eukaryotic cytosine C5 DNA MTases and accommodates the active center of the enzyme with 10 conserved amino acid motifs that are diagnostic for all C5 DNMTs (Cheng, 1995; Jurkowska et al., 2011). The transfer of the methyl group from the cofactor AdoMet to the C5 position of the cytosine occurs within the AdoMet-dependent MTase fold. This consists of a mixed seven-stranded sheet, formed by 6 parallel β strands and the seventh strand inserted

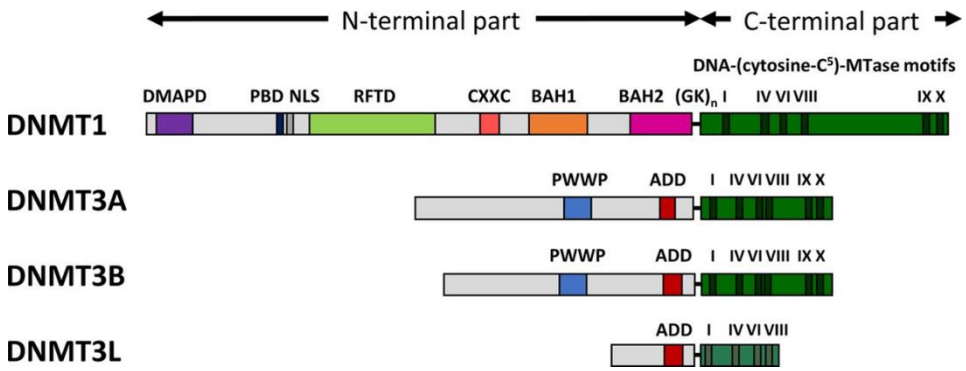


Figure 12| Domain structure of mammalian DNMT proteins. The proteins are broadly split into two parts: the N-terminal part, with regulatory roles, and the C-terminal, containing the motifs important for catalysis. Each characterized domain is correspondingly colored and annotated. The roman numbers were used to annotate the catalytic motifs. The image was taken from (Jeltsch and Jurkowska, 2016).

anti-parallelly between strands 5 and 6. Motifs I and X are involved in cofactor binding, motifs IV, VI and VIII are responsible for catalysis, while the non-conserved region between motifs VIII and IX, plays a role in DNA recognition and specificity (Cheng, 1995; Jeltsch, 2002, 2006; Jurkowska et al., 2011). DNMT3L has a crippled catalytic domain, lacking motifs IX and X, which renders the protein enzymatically inactive (Bourc'his et al., 2001). One of the most interesting structural and mechanistic feature of DNMTs is their mechanism of modifying the cytosine base, which became apparent with the first crystal structures of C5 bacterial methyltransferases HhaI and M. HaeIII, in complex with DNA (Cheng, 1995; Cheng et al., 1993; Jeltsch, 2002; Reinisch et al., 1995; Roberts, 1995). This is conserved in mammalian MTase, and involves base flipping of the target cytosine out of the DNA helix and its burying in a hydrophobic pocket in the active center (Jeltsch, 2002, 2006). The transfer of the methyl group involves a nucleophilic attack of the catalytic cysteine located in the PCQ motif, leading to the formation of a covalent bond between the enzyme and the cytosine target base (**Figure 13**). This is followed by the addition of the methyl group to C5, deprotonation and finally, cleavage of the covalent bond between the enzyme and the DNA (Jeltsch, 2006; Jurkowska et al., 2011). Although the addition of a single methyl group does not change the Watson-Crick paring of the modified base, the presence of the hydrophobic group within the major groove of the DNA, was reported to lead to a subtle bending and twisting effect in crystal structures of methylated DNA oligonucleotides (Tippin and Sundaralingam, 1997). Cytosine methylation was also shown to influence the sequence dependence of nucleosome occupancy and an interplay between nucleosome positioning and the activity of DNMTs has been observed (Chodavarapu et al., 2010; Collings et al., 2013; Kelly et al., 2012; Portella et al., 2013). While the

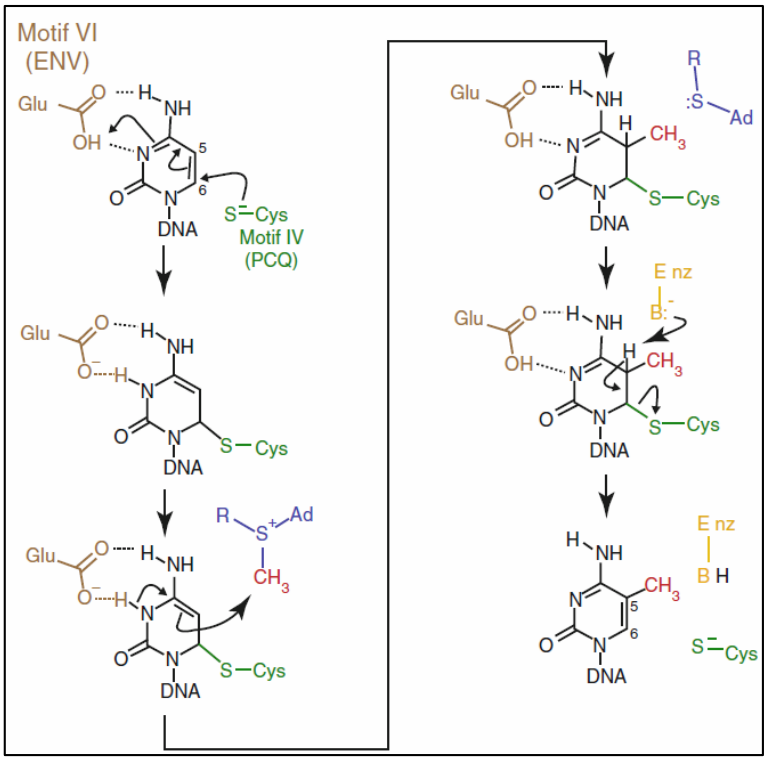


Figure 13| Chemistry of the DNA methylation reaction. The image was taken from (Jurkowska and Jeltsch, 2016)

5mC signal is chemically not as versatile as the PTMs of histone proteins, its involvement in gene regulation is multifactorial (**Figure 14**). This is dependent on the number, status and genomic position of the modified CG sites (Reddington et al., 2013). The clearest downstream effect of 5mC remains its influence on the recruitment of methylation-sensitive proteins and on the binding of transcription factors (Machado et al., 2015; Patel, 2016).

1.4.2 The classical model of DNA methylation inheritance

The four active DNA methyltransferases have important, non-overlapping functions. As such, the initial methylation patterns are set by DNMT3A and DNMT3B, which are called *de novo* DNMTs and use unmethylated DNA as substrate (**Figure 15**) (Chédin, 2011; Okano et al., 1999). *Dnmt3a* is maternally provided and is predominantly expressed in oocytes and early preimplantation embryos. The enzyme has important roles in establishing the differential methylation patterns at imprinting control regions (ICRs) in male and female gametes (Kaneda et al., 2004; Kato et al., 2007; Messerschmidt et al., 2014). *Dnmt3b* is transcribed upon zygotic gene activation and is mostly expressed at the blastocyst stage predominantly in the epiblast lineage (Watanabe et al., 2002). The fact that the two *de novo* enzymes have partial non-

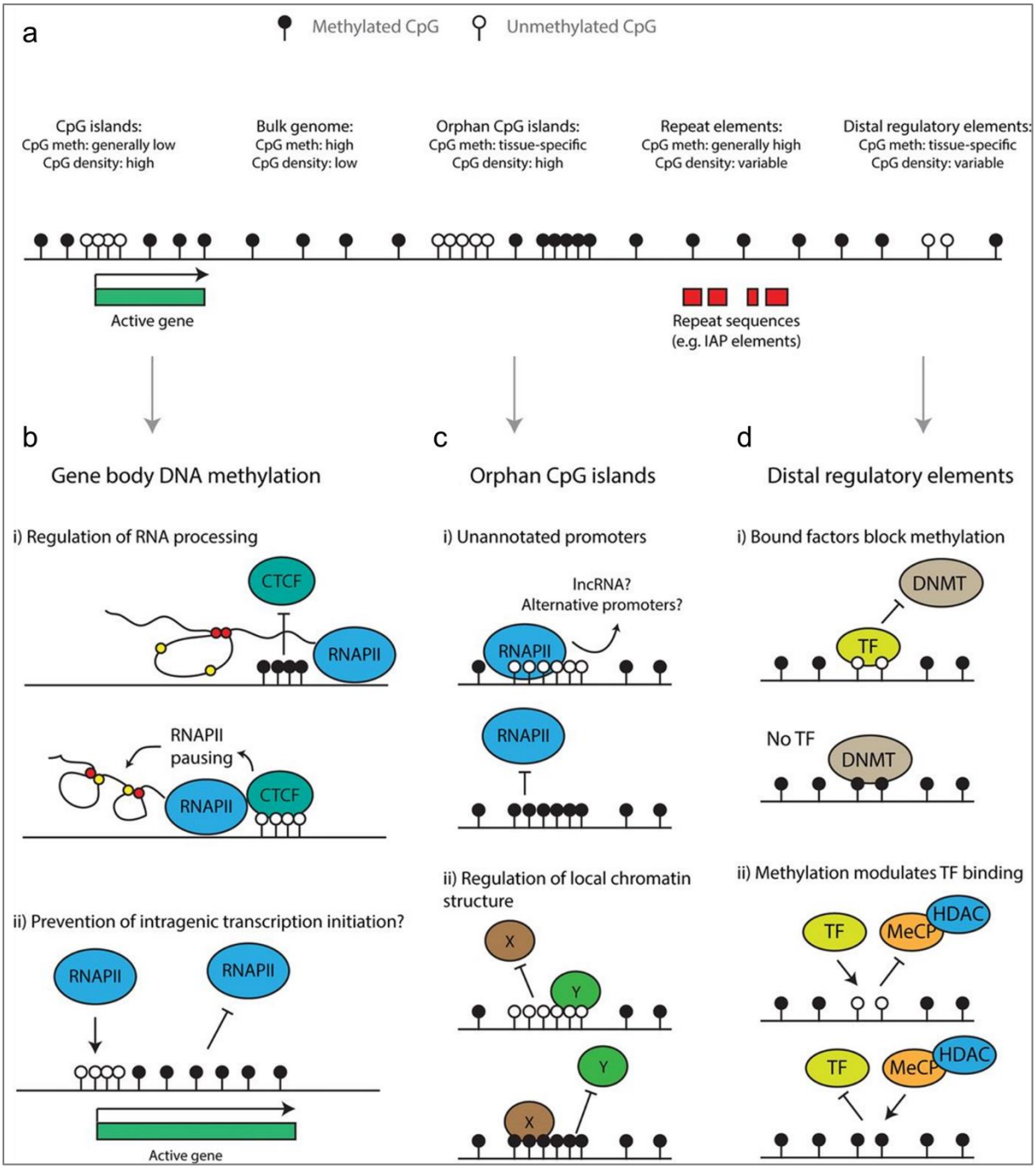


Figure 14| Effects of DNA methylation on gene regulation. a) Schematic representation of the different types of CG methylation patterns found in the mammalian genome. **b**) Gene body methylation has been linked to an effect on RNA processing (by excluding the methylation sensitive CTCF protein) and intragenic transcription initiation. **c**) Involvement of DNA methylation at orphan CGI, CGI that are not associated with the 5'end of annotated genes. Here, the 5mC signal was proposed to block abnormal transcription initiation and differentially recruit chromatin regulators. **d**) DNA methylation of distal regulatory elements is proposed to modulate transcription factor binding. The image was taken from (Reddington et al., 2013).

overlapping functions became apparent from knock-out studies. While deletion of *Dnmt3b* is embryonically lethal, *Dnmt3a* knockouts are partially viable (Okano et al., 1999). The activity of both DNMT3A, in particular at ICR, in gametes, is supported through the interaction partner DNMT3L (Bourc'his et al., 2001). The importance of this

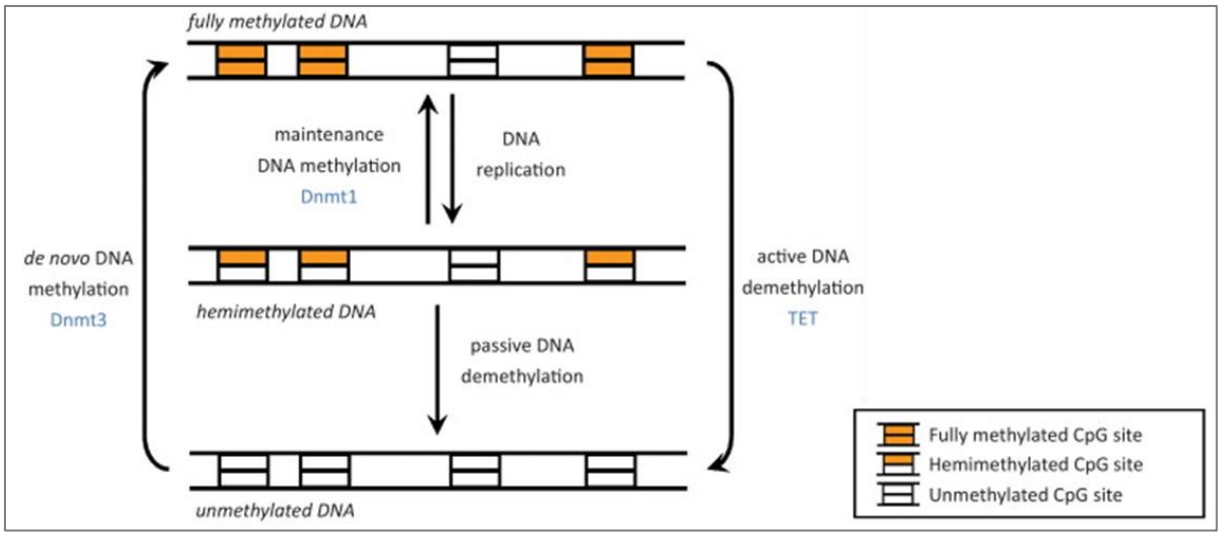


Figure 15| The classical site-specific DNA methylation inheritance model. In the initial stage, the *de novo* methylation machinery sets the 5mC mark on the unmethylated DNA. This pattern is accurately maintained during each DNA replication cycle by Dnmt1. Active demethylation can occur though the action of TET enzymes, which reprogram the DNA methylation patterns. The image was adapted after (Jeltsch and Jurkowska, 2014)

crosstalk is demonstrated by the fact that *Dnmt3l* knockout mice lack *de novo* methylation in the germline, which causes sterility in males and embryonic lethality of maternal null-derived embryos (Bourc'his et al., 2001; Messerschmidt et al., 2014).

Later in development, the perpetuation of the initial methylation pattern is taken over by the maintenance methyltransferase DNMT1 (**Figure 15**). This is important since after each round of DNA replication, hemimethylated DNA is created, where only the parental strand carries the 5mC mark. DNMT1 uses hemimethylated DNA as substrate and is responsible for restoring the methylation pattern on the newly synthesized daughter strand. It has a higher affinity for the hemimethylated template which is essential for restricting the methylation activity to maintenance functions and limiting its *de novo* methylation potential (Goyal et al., 2006; Hermann et al., 2004; Song et al., 2012). The expression of DNMT1 is under tight temporal control. As such, its expression is activated by cell-cycle-dependent transcription factors in S phase, which leads to a high expression in mitotic cells (Kishikawa et al., 2003). The enzyme is recruited to replication foci through interaction with its PCNA-interacting binding partner, UHRF1. UHRF1 specifically attracts DNMT1 to the parental strands and orients the enzyme towards the newly synthesized daughter strand (Avvakumov et al., 2008; Bostick et al., 2007; Sharif and Koseki, 2011; Sharif et al., 2007). In addition to its recruitment role, the protein was also reported to stimulate the activity of DNMT1 by a factor of 5 *in vitro* and increase the specificity of DNMT1 for hemimethylated CG sites (Bashtrykov et al., 2014). Deletion of mouse *Dnmt1* is lethal at gastrulation, this

resulting out of a dramatic global loss of DNA methylation (Brown and Robertson, 2007; Kurihara et al., 2008). Interestingly, loss of UHRF1 cases similar defects (Sharif and Koseki, 2011; Sharif et al., 2007). By suppressing the maintenance MTase activity, passive DNA methylation takes place, which results in around 50% loss of 5mC with each new DNA replication round. In addition to this, active demethylation through the enzymatic activity of TET enzymes was also documented (**Figure 15**) (Ito et al., 2011a; Véron and Peters, 2011). Active demethylation was proposed to have regulatory roles on DNA methylation, for instance by preventing erroneous methylation of CG islands (Williams et al., 2012).

1.4.3 A revised model of DNA methylation inheritance

A significant body of work has recently demonstrated that despite its elegance, the classical inheritance model of DNA methylation, where DNMT3A and 3B set the initial 5mC pattern that is subsequently copied by DNMT1, is an oversimplification (Jeltsch and Jurkowska, 2014). Instead, DNA methylation is most appropriately described by a dynamic process of ongoing methylation and demethylation events in which the functions of DNMT1 and DNMT3 enzymes partially overlap (Jeltsch and Jurkowska, 2014; Jones and Liang, 2009; Riggs and Xiong, 2004). The proposal of this revised model was catalyzed by several experimental observations that were not in agreement with the classical site-specific DNA maintenance concept.

First, while the classical inheritance model would predict that all cells of one tissue display identical DNA methylation patterns, bisulfite conversion followed by sequencing to elucidate the methylation status of individual CG sites, did not confirm this hypothesis. The analysis revealed that the average methylation density profile of a particular DNA region is preserved, instead (Jones and Liang, 2009; Zhang et al., 2009). This outcome is in agreement with biochemical studies showing that while DNMT1 has a 10-40-fold preference for hemimethylated sites, this does not suffice for an accurate copying of the methylation status of the around 56 million CG sites in the human genome (Bashtrykov et al., 2012a, 2012b; Jeltsch and Jurkowska, 2014; Song et al., 2012).

A second argument for the revised model is based on genetic studies in mice and mammalian cell lines, showing that deletion of DNMT3A and DNMT3B leads to loss of methylation at repetitive elements, even in the presence of a fully functional DNMT1 (Chen et al., 2003; Dodget et al., 2005; Egger et al., 2006).These results indicate that

DNMT3A and 3B are involved in the maintenance of DNA methylation levels as well. Recently, cytosine-C5 methylation at asymmetric sequences, particularly in CA context, was discovered in human stem cells, germ cells and differentiated neuronal progenitors (Guo et al., 2014c, 2014d, Lister et al., 2009, 2013; Pinney, 2014). This modification can be introduced by DNMT3A, enzyme with high expression in the tissues that contain the most abundant levels of non-CG methylation (Arand et al., 2012; Shirane et al., 2013). Since template strand information cannot be used for non-palindromic methylated sequences, the stable presence of CA methylation in the human genome is a strong indicator for the permanent need for the *de novo* activity of DNMT3A in certain cell types (Jeltsch and Jurkowska, 2014).

Finally, the discovery of TET enzymes, raised awareness about the continuous requirement for active *de novo* MTase to counteract active DNA demethylation (Arand et al., 2012; Métivier et al., 2008; Williams et al., 2012).

Together, these four key arguments support the revision of the classical model for DNA methylation inheritance and highlight the importance of regulatory factors in the targeting and control of the DNA methylation machinery (Jeltsch and Jurkowska, 2014). Moreover, with this refined concept, corrections of errors that occur during the setting and maintenance of DNA methylation patterns can be explained.

1.4.4 Waves of DNA methylation during development

The dynamic changes in DNA methylation profiles are essential for re-establishing the cellular identity in the early embryo. While mammalian sperm cells show high levels of DNA methylation, the oocytes display a progressive increase in 5mC levels with cellular maturation (Guo et al., 2014b; Smith et al., 2014). Mature mouse sperm contains 80-90% overall CG methylation, while the maternal genome displays lower global methylation levels of ~ 40% (Messerschmidt et al., 2014). There are two dramatic waves of global DNA methylation reprogramming during development (**Figure 16**). The first takes place shortly after fertilization, where the methylomes of both sperm and oocyte are erased in order to provide a blank environment for setting the methylation profiles required for the pluripotency of the blastocyst inner cell mass (ICM) (Ciernia and LaSalle, 2016). During this process, the heavily methylated paternal genome undergoes predominantly active demethylation by TET3, while the maternal genome loses its 5mC levels mostly through passive, replication-dependent dilution (Guo et al., 2014a; Oswald et al., 2000; Shen et al., 2014). Interestingly, despite this

radical reprogramming, the imprinted genes, and some repeats remain methylated during this stage (Jurkowska and Jeltsch, 2013; Moore and Haig, 1991). Following implantation, the ICM starts to acquire tissue-specific differentiation properties, which are molecularly coupled with a sharp increase in 5mC levels across the genome (Messerschmidt et al., 2014).

The second and most dramatic round of epigenetic reprogramming occurs during gametogenesis (**Figure 16**). Here, 5mC is globally removed, including at parental-inherited imprints. This hard reset occurs predominantly through passive demethylation facilitated by the silencing of the *de novo* DNA methylation machinery, nuclear exclusion of the DNMT1 interactor, UHRF1, and shortening of the cell cycle to 16h (Chuva De Sousa Lopes et al., 2008; Hajkova et al., 2002; Kagiwada et al., 2012; Sasaki and Matsui, 2008). Following embryonic day E12.5 sex-specific development of mouse precursor germ cells (PGCs) ensues which occurs with different kinetics between female and male precursor germ cells (PGCs) (Kagiwada et al., 2012). While by E16.5 male PGCs have achieved 50% global DNA methylation, female germ cells

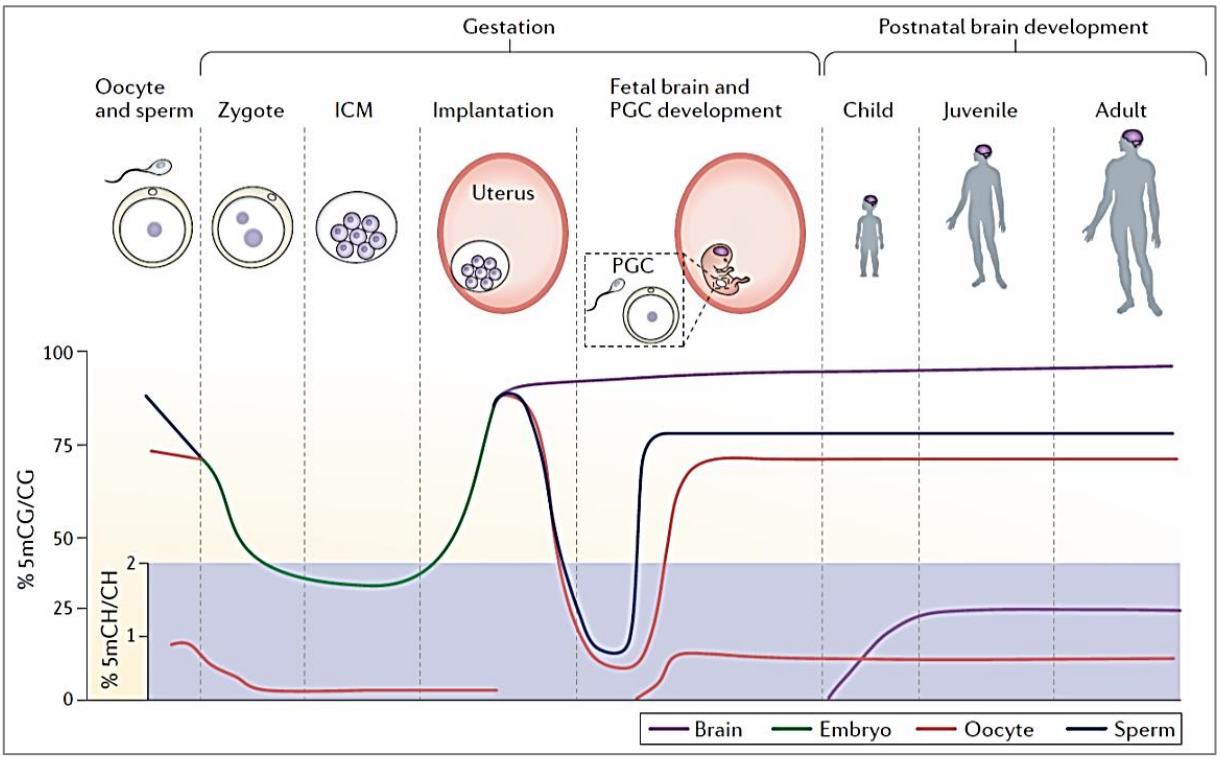


Figure 16| The dynamics of the DNA methylation landscape across human development. The changes occurring in the levels of DNA methylation present in CG (5mCG) as well as in non-CG (5mCH) context, are shown on separate axis. The two global waves of DNA methylation reprogramming occurring at implantation and during the development of PGCs are correspondingly annotated and color-coded. In addition, the methylome profiles of the postnatal human bran, at defined developmental stages, are displayed. The image was taken from (Ciernia and LaSalle, 2016).

display low levels of 5mC. In females, remethylation of gametes is initiated at birth during the growth phase of the oocyte (Kota and Feil, 2010; Saitou and Yamaji, 2012).

DNA methylation profiles are also highly dynamic across the lifespan of the organism with most differences occurring in the transition from fetal to early postnatal life (**Figure 16**) (Ciernia and LaSalle, 2016). For instance, by comparing 25 fetal to 300 postnatal samples, a large change in 5mC levels were discovered in the methylome of the human frontal cortex, in-between fetal and postnatal stages (Jaffe et al., 2016). A global restructuring of the epigenetic landscape could be observed, with > 50% of CG sites of the over 6,000 differential methylation regions (DMRs) and 896 multiple-kilobase regions being affected (Jaffe et al., 2016). These strong changes in DNA methylation were largely attributed to the alteration in the cellular composition of the brain in-between the two analyzed time-points, with fetal cortex containing more progenitor-like cells, while postnatal tissue consisting predominantly out of linage-specified cells (Ciernia and LaSalle, 2016; Jaffe et al., 2016).

The combined control of the targeting and activity of the DNA methylation machinery is essential for a precise spatiotemporal regulation of the function of DNMT enzymes, and thereby for the accurate setting and maintenance of the cellular methylation profiles (Jeltsch and Jurkowska, 2016). Due to its high relevance for the present work, the targeting and regulation of DNMT3 enzymes will be covered in more detail in the following section. Particular emphasis will be placed on the spatiotemporal regulation of DNMT3 enzymes *via* interaction with chromatin and other proteins (Jeltsch and Jurkowska, 2016).

1.4.5 Control of DNA methylation

1.4.5.1 Chromatin-guided spatiotemporal control of DNMT3 enzymes

The regulatory N-terminal part of the enzymatically active *de novo* DNA methyltransferase 3A and 3B takes part in the chromatin targeting of the proteins (Dhayalan et al., 2010; Ge et al., 2004; Noh et al., 2015; Zhang et al., 2010). Two distinct functional domains are present in the N-terminal part: the ADD (ATRX-DNMT3-DNMT3L) and the PWWP (Pro-Trp-Trp-Pro) domain. As indicated by the name, the ADD domain is also found in the catalytically inactive factor DNMT3L. This protein does not contain a PWWP domain, however (Jurkowska et al., 2011). N-terminally to the PWWP domain, both DNMT3A and DNMT3B contain a non-conserved region that was

shown to bind DNA and contribute to the tight association of the enzymes to chromatin (Suetake et al., 2011). While the definite role of this region awaits further elucidation, based on the low sequence identity between the two proteins (only 10.9%), it is tempting to speculate that this area might be involved in differential genomic targeting of DNMT3A and DNMT3B activities (Jeltsch and Jurkowska, 2016; Rondelet et al., 2016).

The PWWP domain is a member of the 'Royal' superfamily of domains that recognize methylated lysines on the histone tail through a conserved aromatic cage (Qin and Min, 2014; Rondelet et al., 2016). In the case of DNMT3 enzymes, this domain was reported to specifically bind to H3K36me3, a mark associated with gene body methylation in ES cells and the elongation phase of transcription (Baubec et al., 2015; Dhayalan et al., 2010; Hahn et al., 2011; Lee and Shilatifard, 2007). Structurally, the domain consists of an antiparallel β-barrel-like fold formed by five β-strands (β1–β5). A short 3_{10} helix is found between the strand β4 and β5 (η2) and a short insertion motif is present between strands β2 and β3 (η1). Within the conserved Pro-Trp-Trp-Pro motif, the first proline is replaced by a serine in both, DNMT3A and DNMT3B (Rondelet et al., 2016). Recently, the first position of the Pro-Trp-Trp-Pro motif has been shown to regulate the PWWP domain stability and oligomerization (Hung et al., 2015). The importance of the PWWP for the chromatin targeting of the DNMTs is demonstrated by the decreased methylation levels observed at human pericentromeric repeats, when the H3K36me3 binding is abolished through an S270P exchange in the DNMT3B aromatic pocket. Furthermore, generation of the corresponding PWWP pocket mutant within DNMT3A, revealed a partial loss in the heterochromatin localization of the enzyme, upon its overexpression in mouse fibroblasts (Chen et al., 2004; Dhayalan et al., 2010). Crystal structure of the DNMT3B PWWP domain, in complex with the H3K36me3 peptide (Rondelet et al., 2016), revealed that the trimethylammonium group is inserted into a Phe-Trp-Trp-Asp cage stabilized by intermolecular hydrogen-bonding interactions between the backbone of the peptide and the domain. In addition to the H3K36me3 binding, the PWWP domain of both DNMT3A and DNMT3B was reported to show a relatively weak non-specific interaction with DNA, through a basic surface adjacent to the histone binding site (Purdy et al., 2010; Qiu et al., 2002). Structural superposition with the previously studied LEDGF PWWP domain indicated that the positively charged basic patch is positioned on the DNA wrapped around the histone core and might form contacts with the phosphate backbone of the DNA helix.

This might further enhance the selectivity of the PWWP for H3K36me3-modified nucleosomal particles (Rondelet et al., 2016). The central role of K36 methylation in targeting of DNMT3 enzymes was recently confirmed in yeast and mouse male germline cells (Jeltsch and Jurkowska, 2016; Morselli et al., 2015). Here, a positive correlation between H3K36me3 and DNA methylation by DNMT3B was observed. Moreover, the analysis revealed a strong preference of DNMT3B for methylating non-nucleosomal linker DNA. This is in line with biochemical data, demonstrating that wild-type DNMT3A is more active on native chromatin by comparison to the PWWP-pocket mutant, and that linker DNA is the preferred substrate when using *in vitro* reconstituted chromatin (Dhayalan et al., 2010; Zhang et al., 2010).

In addition to the PWWP domain, DNMT3 proteins also contain an ADD domain. This structure comprises three subdomains: a GATA-like zinc finger, a PHD finger with cross-braced topology and a long C-terminal α-helix that stems out of the PHD finger subdomain and forms extensive hydrophobic contacts with the GATA finger (Jeltsch and Jurkowska, 2016). The ADD domain is essential for the regulation of DNMT3 proteins, as it was shown to specifically readout H3 tails unmethylated at K4 and serve as a protein-protein interaction platform (Guo et al., 2015c; Ooi et al., 2007; Otani et al., 2009; Zhang et al., 2010). Binding is disrupted by any large modification on K4 such as di-and trimethylation or acetylation (Noh et al., 2015; Otani et al., 2009). Recognition of the unmodified histone H3 tail has drastic effects on the activity and chromatin targeting of the enzyme, with the mechanism responsible for this regulatory effect being recently elucidated in a seminal structural study by Guo *et al.* (**Figure 17**) (Guo et al., 2015c). The authors have uncovered that the ADD and the catalytic domains (CD) of DNMT3A adopt an autoinhibitory conformation, with the two domains folding into two individual structural modules connected by a linker that is packed against a hydrophobic surface of the CD domain (**Figure 17a** and **b**). A loop region formed by the aa 526-531 extends out of the ADD domain and is inserted in the CD domain into a pocket built mainly of hydrophobic residues. This conformation brings three acidic residues (D529/D530/D531) and three hydrophobic residues (Y526/Y528/Y533) from the side of the ADD domain in close proximity to the hydrophobic pocket (**Figure 17**). This interaction blocks the DNA-binding affinity of DNMT3A, consequentially inhibiting catalysis. Interestingly, the addition of the H3K4me0 peptide could allosterically relieve this autoinhibitory effect, through binding to the ADD domain at the ADD-CD interface, disrupting the intramolecular interaction

and causing of a large conformational rearrangement within the enzyme (**Figure 17c** and **e**) (Guo et al., 2015c; Jeltsch and Jurkowska, 2016). Noteworthy, a stimulatory effect on the enzymatic activity could only be observed for H3K4me0 and not for H3K4me3-modified peptides (Guo et al., 2015c). This biochemical observation is in line with several genome-wide studies, where an anti-correlation between H3K4me3 and DNA methylation was observed (Hodges et al., 2009; Meissner et al., 2008; Messerschmidt et al., 2014; Weber et al., 2007). H3K4me3 is found in 75% of all human active gene promoters in several cell types and has a key role in mammalian gene expression (Bernstein et al., 2012). The mark abundantly decorates CGIs, which are present at many gene promoters and are usually free of DNA methylation (Edwards et al., 2010; Meissner et al., 2008; Mikkelsen et al., 2007). This indicates that H3K4me3 and 5mC are antagonistic epigenetic signals and the crosstalk between these two pathways is in part mediated by the ADD domain of DNMT3 enzymes. This hypothesis was recently experimentally confirmed by exogenous introduction of DNMT3B into *Saccharomyces cerevisiae.* The budding yeast is a good model system for this type of analysis, since it lacks endogenous DNA methylation while possessing the conserved

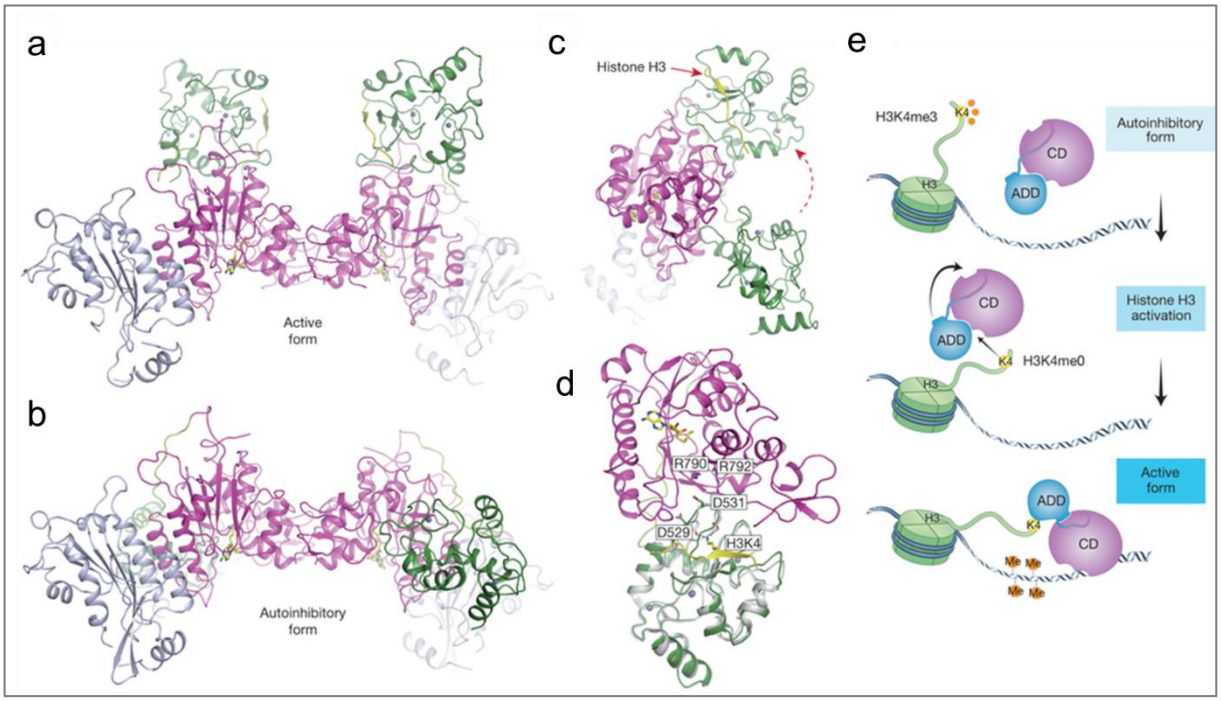


Figure 17| Allosteric control of DNMT3A activity by the H3K4me0 tail. a) Ribbon representation of the active, open conformation of the ADD (blue)-CD(pink)-C^{DNMT3L}(green), in complex with the unmodified H3 peptide (yellow). The autoinhibitory form, obtained in the absence of the peptide, is shown for reference in **b. c)** Superimposition of the active and autoinhibitory form of the ADD-CD-C^{DNMT3L} complex, highlighting the large conformational change triggered by the H3K4me0 binding. **d**) Ribbon representation of the ADD-CD structure, documenting the residues which are important for mediating the contact between the two domains. e) Allosteric activation model of DNM3A by the unmodified histone H3, as proposed by Guo et al (2015c). The image was taken from (Guo et al., 2015c)

histone sequences and many of the PTMs found in higher eukaryotes (Hu et al., 2009). In this setup, the authors observed an anti-correlation between introduced 5mC and H3K4me3 levels (Morselli et al., 2015). Reciprocally, deletion of *set1*, the product of which is responsible for setting the H3K4me3 mark in yeast, led to a spreading of DNA methylation into promoter regions (Morselli et al., 2015). Consistent with this, expression of a DNMT3A variant, with an engineered ADD domain that was rendered insensitive to H3K4 methylation, resulted in ectopic DNA methylation of genes associated with early embryonic lineages and subsequent abnormalities in ESC differentiation (Noh et al., 2015). Interestingly, hypermethylation at H3K4me3/CGI loci was less efficient (< 30%) than at intergenic regions (> 60%), suggesting that there are additional factors that contribute to the protection of gene promoters from DNA methylation (Noh et al., 2015). Together, these results highlight the importance of a regulated DNMT3 activity for proper cellular differentiation and development.

Noteworthy, the cross-talk between the H3K4me3 pathway and DNA methylation is bi-directional, with hypermethylated loci being more resistant to gene reactivation. This was recently demonstrated in an elegant synthetic biology approach, where the SET domain of human PRDM9, was fused to custom-made DNA binding domains, for locus-specific trimethylation of H3K4 (Cano-Rodriguez et al., 2016). Here, the authors have observed that targeting PRDM9 to promoters that carry abundant DNA methylation, resulted in only transient re-expression of the gene locus. By contrast, genes that carried no or low DNA methylation were more readily activated (Cano-Rodriguez et al., 2016).

1.4.5.2 Regulation of DNMT3 enzymes by interacting proteins

In addition to reading the H3K4 modification status, the ADD domain of DNMT3 enzymes serves as an interaction knob with several proteins involved in chromatin regulation. Accordingly, the ADD domain of DNMT3A was reported to directly interact with members of the H3K9me3 silencing pathway such as SUV39H1 (Fuks et al., 2003a) and SETDB1 (Li et al., 2006). In the case of SETDB1, this interaction appears to be essential for the methylation and consequent repression of a set of promoters in cancer cells. Interactions with the H3K9me3-reading proteins HP1α and β as well as with the histone deacetylase HDAC1 were also reported (Fuks et al., 2003a). Together, these data propose a model in which the chromatin targeting of the DNA methylation machinery occurs to a large extent through interaction with other important

epigenetic silencers that through their concerted action enforce a highly repressive local chromatin environment. This hypothesis is supported by experimental data stemming from several knockout studies. Analysis of embryonic stem cells isolated from *Suv39h1/h2*$^{-/-}$ mice revealed reduced levels of DNA methylation at major satellite repeats. This effect appeared to be genomic-context specific, as other classes of repetitive and parasitic elements such as minor satellites and C-type retroviruses showed unaffected DNA methylation levels (Arand et al., 2012; Du et al., 2015a; Lehnertz et al., 2003). Similarly, knockout of *Setdb1* also resulted in selective depletion of DNA methylation at a number of germline-specific genes (Karimi et al., 2011).

The PWWP domain was also reported to function as a protein interaction platform, in addition to its readout of H3K36me3-modified tails. Here, the ZHX1 (zinc-fingers and homeobox protein 1) was reported to interact with DNMT3B and enhance DNA methylation-dependent transcriptional repression (Kim et al., 2007). Noteworthy, protein interactors can also negatively influence the enzymatic activity of DNMT3s. A member of this category of regulators is the SAL-like 3 (SALL3) protein, which was shown to bind to the PWWP domain of DNMT3A. Through this interaction, the chromatin association as well as the enzymatic activity of DNMT3A were diminished (Shikauchi et al., 2009).

1.4.5.3 The chromatin remodeler HELLS

The most mysterious regulator of DNMT3 enzymes remains arguably the putative chromatin remodeler HELLS (helicase, lymphoid-specific) (Briones and Muegge, 2012). The involvement of this protein in the control of DNA methylation in mammals became clear in early knock-out studies, where thin layer chromatography quantification revealed 50% loss in the global levels of 5mC in the DNA isolated from *Hells*$^{-/-}$ mice. Detailed investigation showed that affected sites were highly heterogeneous, comprising mainly repetitive elements but also single-copy sequences like the imprinted region of the H19 gene, these effects being detectable in a broad array of tissue types (Briones and Muegge, 2012; Dennis et al., 2001). The effect of the protein on DNA methylation was even earlier discovered in *Arabidopsis thaliana* in a forward genetic screen for factors involved in the maintenance of DNA methylation. Here, the lack of the *A. thaliana* DDM1 (decrease in DNA methylation 1) protein, led to a 70% reduction in genomic DNA methylation levels, progressively correlated with an increase in mutations and stable epialleles at dispersed sited in the genome, which

eventually culminated in morphological abnormalities (Jeddeloh et al., 1999; Vongs et al., 1993). The gene encoding for the murine DDM1 homolog, HELLS, was first cloned by degenerative PCR to identify novel helicase superfamily members in T-cell precursors (Jarvis et al., 1996). Although its expression was initially characterized in lymphoid cells, follow-up investigations revealed that the protein is present in many tissues, with elevated expression in actively dividing cell types (Briones and Muegge, 2012; Geiman and Muegge, 2000; Geiman et al., 1998; Jarvis et al., 1996). The human gene was cloned from human leukemic cells and it is known as *LSH* (Lymphoid-specific helicase), *HELLS* (helicase, lymphoid-specific), *PASG* (proliferation associated gene) or *SMARCA6* (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 6) (Briones and Muegge, 2012; Lee et al., 2000). The important and non-redundant role of HELLS is highlighted by the embryonic lethal phenotype displayed by the *Hells*$^{-/-}$ mice (Geiman et al., 2001; Sun et al., 2004). These animals show severe developmental defects manifested by kidney necrosis, reduced embryonal growth, abnormal regulation of developmental important genes such as the Hox cluster, incomplete differentiation of ES cells in *in vitro* cultures and impaired hematopoiesis (Fan et al., 2008; Geiman et al., 2001; Sun et al., 2004; Yan et al., 2003a).

Based on the domain composition, DDM1, and its homolog HELLS were classified as a member of the SNF2 family of chromatin remodelers, proteins which contain a domain that is homologues to the helicase-like ATPase domain of the *Saccharomyces cerevisiae* Snf2 protein (**Figure 18**) (Laurent et al., 1992). This Snf2 domain comprises two tandem RecA-like folds and seven conserved helicase-related sequence motifs (Eisen et al., 1995; Flaus et al., 2006; Meehan et al., 2001). While Snf2 family members lack canonical helicase activity and are thereby not able to separate nucleic acid strands, they are able to induce ATP-dependent torsional strain to DNA, which

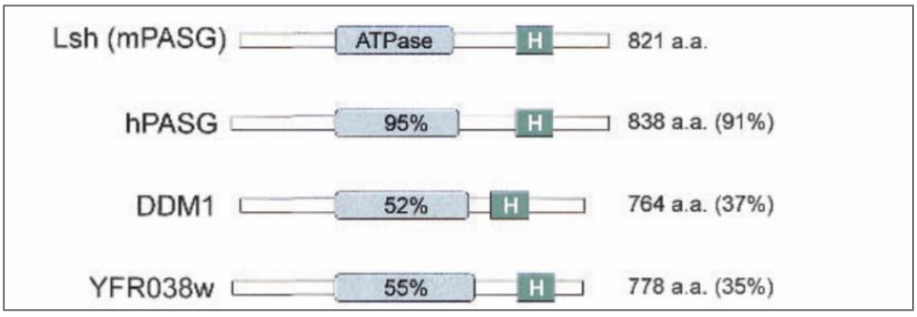


Figure 18| Schematic representation of the SNF2-like ATPases that are homologous to mouse LSH. These are human PASG, *Arabidopsis Thaliana* DDM1, and *Saccharomyces cerevisiae* YFR038w. The percentage of amino acid identity for both the ATPase domain as well as for the full length proteins were correspondingly annotated. The image was adapted after (Meehan et al., 2001).

provides the necessary force to remodel nucleosomes (Ryan and Owen-Hughes, 2011). Depending on the family member, this can have different effects on the local structure of chromatin. The action of Snf2 proteins can result in sliding or exchange of nucleosome, or disassembly of nucleosome histone content (Ryan and Owen-Hughes, 2011). This large diversity of outputs might be controlled by accessory domains and/or interacting proteins outside the Snf2 domain. Indeed, many Snf2 proteins were found to reside in large multi-protein complexes (Clapier and Cairns, 2009). Through this, the activity of Snf2 proteins can be modulated or recruited to different chromatin compartments.

By contrast to other members of the family, native HELLS was found to be predominantly present as a free monomer in nuclear extract of human cells, as analyzed by size exclusion chromatography (Myant and Stancheva, 2008). This indicates that HELLS might be incorporated in unstable or not very abundant complexes. Indeed, alternative methods such as co-immunoprecipitation have uncovered that HELLS can participate in a metastable repressive protein complex that contains DNMT1, DNMT3B, HDAC1 and/or HDAC2 (Myant and Stancheva, 2008; Zhu et al., 2006). *In vitro* pull-down experiments with recombinant proteins have further pinned down the region of HELLS through which the interactions take place, as well as the order of assembly of the complex. Accordingly, the N-terminal portion of HELLS (amino acids 1 to 226), containing a predicted coiled-coiled motif, was shown to be necessary and sufficient for co-immunoprecipitating the two DNMTs. While the interaction with DNMT3B was direct, the association with DNMT1 was mediated by the *de novo* methyltransferase. DNMT1 was responsible for the further recruitment of HDAC enzymes (Myant and Stancheva, 2008). In line with these data, *Hells* knockdown in mouse cells was associated with an increase in histone acetylation and H3K4me3 levels on repetitive elements, coupled with a depletion in DNA methylation and an elevated number of transcripts produced from satellite, Sine and Line1 repeats (Huang et al., 2004). This upregulation of repetitive elements might be in part responsible for the genomic instability observed for the *Hells*$^{-/-}$ cells and animals (Briones and Muegge, 2012).

Despite its drastic and widespread effects on the DNA methylation levels, mechanistic insight into how HELLS controls the DNMT machinery is sparse. Through analogy to the better characterized *A. thaliana* homologue DDM1, the current view proposes a

chromatin scanning role for HELLS, where the protein continuously samples the nuclear environment to open target chromatin regions, which pose accessibility challenges to the activity of DNMT enzymes (**Figure 19**) (Briones and Muegge, 2012; Brzeski and Jerzmanowski, 2003; Myant and Stancheva, 2008; Zemach et al., 2013). Indeed, unlike DNMT1 that works on nucleosome-depleted newly replicated DNA, the DNMT3 enzymes must work on chromatinized template (Chédin, 2011). High resolution mapping of *de novo* methylation events in reconstituted chromatin systems revealed that by comparison to linker DNA, nucleosomal DNA is largely devoid of CG methylation when subjected to *in vitro* methylation (Felle et al., 2011). This result is supported by cellular data, documenting an increase of *de novo* methylation with CG density, and an exclusion from nucleosomes (Baubec et al., 2015). Together, these studies indicate that stably positioned nucleosomes pose an inaccessible barrier for the *de novo* methylation machinery and that efficient methylation requires the activity of chromatin remodelers for dynamic nucleosome repositioning or disruption (Termanis et al., 2016). In cells, this effect might be particularly strong at highly compacted sites such as pericentromeric heterochromatin. This hypothesis was supported by work in *Arabidopsis*, where the remodeler DDM1 was found to promote the access of the DNA methylation machinery at highly compacted DNA sequences that are enriched in the linker histone H1 (**Figure 19**) (Zemach et al., 2013). While the requirement of active chromatin remodeling for efficient DNA methylation is indisputable, the chromatin remodeling activity of mammalian HELLS remains a matter of intense debates. Mechanistic investigations have been challenging, because unlike DDM1, recombinant HELLS is not able to reposition nucleosomes *in vitro* and displays a poor but DNA-stimulated ATPase activity (Brzeski and Jerzmanowski, 2003; Burrage et al., 2012). The functionality and importance of ATP hydrolysis for the function of HELLS has been

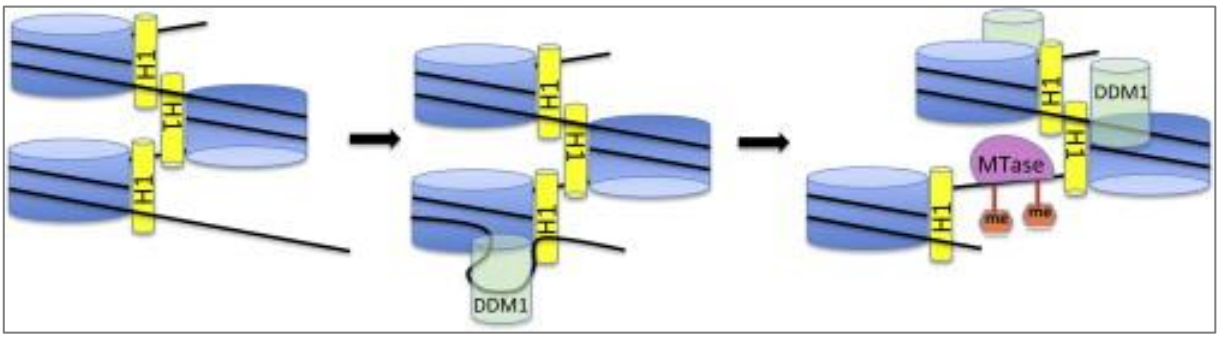


Figure 19| The chromatin remodeler DDM1 facilitates the access of DNMTs to compacted chromatin substrates. The model proposes that through its action DDM1 is capable of counteracting the thigh association between the linker histone H1 (yellow) and the nucleosome, and though this generate accessible DNA templates that can be methylated by DNMT enzymes The image was taken from (Pikaard, 2013).

hinted at by *in vivo* studies, where deletion of HELLS exons 10-12, which leads to the truncation of the catalytic Snf2 domain, resulted in DNA methylation defects in mice, as well as premature aging and upregulated expression of senescence-associated markers (Sun et al., 2004). It remains nevertheless unclear, whether the ATPase activity of HELLS together with its associated chromatin remodeling function directly contribute to this wide range of effects.

According to its classification HELLS contains the signature motifs of SNF2 enzymes, an ATP binding pocket encompassing Walker A and B motifs (Flaus et al., 2006). Within the conserved GKT box, the protein contains a lysine residue (K237), which was shown to be essential for ATP binding and hydrolysis, in several other canonical chromatin remodelers (Deuring et al., 2000; Rowbotham et al., 2011; Shen et al., 2000). Exchanging this residue for glutamine was reported to generate an HELLS variant, which was not longer able to promote efficient phosphorylation of the histone variant H2AX, and through this impair the efficient repair of DNA double-strand breaks in mammalian cells (Burrage et al., 2012). Despite this biological insight, the effect that the lack of ATPase activity has on the chromatin association properties of HELLS remained a matter of speculation. In this work, I took advantage of the HELLS K237Q variant to investigate whether ATP hydrolysis plays a role in the interaction between HELLS and highly compacted chromatin in cells.

1.4.6 DNA methylation readout

60-80% of all CG dinucleotides in the mammalian genome are C5-methylated. While this modification does not dramatically alter the structure of the DNA helix, it exerts its effect predominantly by serving as a docking site for 5mC-binding proteins that in turn act as recruiters for chromatin-modifying complexes (Fournier et al., 2012; Patel, 2016). There are several groups of proteins that can interpret and locally modulate DNA methylation patterns. These include the family of methyl-CG-binding domain (MBD) proteins that recognize fully methylated CG dinucleotides, the Kaiso family of zinc finger (ZF) proteins, which preferentially bind methylated CG sites within longer specific sequences, and the SET-and-Ring finger-associated (SRA) domain family, whose members recognize hemimethylated DNA sites generated after replication (Du et al., 2015b; Filion et al., 2006; Hendrich and Bird, 1998; Liu et al., 2013b; Prokhortchouk et al., 2001; Unoki et al., 2004). The MBD and the ZF binding domains, dock onto the major groove of the DNA and form specific hydrogen bonds to

discriminate between the DNA bases (Buck-Koehntop et al., 2012; Ho et al., 2008; Liu et al., 2012; Scarsdale et al., 2011; Zou et al., 2012). By contrast to this, mammalian SRA proteins, recognize the hemimethylated CG site through base flipping of the modified base out of the double helix and into a hydrophobic binding cage (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2009; Liu et al., 2013b; Rajakumara et al., 2011). Owing to the high relevance for this work, in the following section, particular emphasis will be placed on the MBD family, the members of which are critical regulators of the transcriptional state of the epigenome (Du et al., 2015b).

MBD proteins mainly act as transcriptional repressors, although depending on the chromatin context, activator roles were also reported (see below). To date, the family consists of eleven proteins, identified predominantly through sequence similarity, and known to contain a MBD domain. Apart from the conserved 70-85 amino acid-long MBD, which has the ability to bind single symmetrically methylated CG dinucleotides, each member of the family contains unique accessory domains. This reflects the heterogeneous and non-overlapping functions of the family members (Du et al., 2015b). For instance, while some proteins contain a transcriptional repression domain (TRD), which serves as a protein-protein interaction platform, other members, such as MBD1, are equipped with domains that can recognize unmethylated DNA sequences (Tachibana et al., 2002). Two members of the MBD group will be covered in depth in the following sections: the founding member of the family, methyl-CG-binding protein 2 (MeCP2), and the largest member of the family, methyl binding protein 1 (MBD1).

1.4.6.1 The chromatin regulator MeCP2

1.4.6.1.1 MeCP2: protein structure

MeCP2 was discovered in 1992 by Bird and co-workers in a search for proteins capable of binding to methylated CG dinucleotides (Lewis et al., 1992). Follow-up *in vitro* investigations revealed that the DNA-binding affinity of the protein is enhanced when the sequence flanking the methylated CG site is A/T rich (Klose et al., 2005). The structural basis of this preference was found to lie within the three AT-hook motif sequences that are present in MeCP2 (Baker et al., 2013; Lewis et al., 1992; Lyst et al., 2016). These are short DNA-binding motifs that use an intrinsic RGRP consensus sequence to form contacts with the wide minor groove of AT-rich sequences (Hendrich and Bird, 1998; Ragione et al., 2016). Recent biochemical and cellular work has demonstrated that within MeCP2, AT-hook 1, exhibits the most robust DNA binding,

with AT-hook 2 being much weaker and AT-hook 3 having an interrupted and probably dysfunctional RGRP motif (Lyst et al., 2016).

Noteworthy, while the AT-hooks stabilize the association of MeCP2 with chromatin, the MBD domain remains the primary interaction surface (Ragione et al., 2016). The preference of the MBD domain for methylated over unmethylated DNA is 2-3-fold for both the murine and the human protein (Fraga et al., 2003; Nikitina et al., 2007a). The X-ray structure of the DNA-bound domain revealed that the MBD-DNA interface is stabilized by the hydration of the methyl group on at the cytosine C5 position. In addition to this domain, MeCP2 contains a TRD that mediates interactions with co-repressor complexes, a nuclear localization signal (NLS), and a C-terminal domain involved in binding of DNA and other proteins (Bienvenu and Chelly, 2006; Ebert et al., 2013; Ragione et al., 2016).

Human MeCP2 was purified to homogeneity and its structure was characterized in great depth by ultracentrifugation, circular dichroism and limited protease digestion (Adams et al., 2007; Ragione et al., 2016). This analysis highlighted the structural peculiarities of the protein, with around 60% of MeCP2 being intrinsically disordered and displaying random coil-like hydrodynamic properties (**Figure 20**) (Adams et al., 2007; Hite et al., 2009). These results indicated that MeCP2 might heavily rely on protein interactors for adopting a folded structure and provided a basis for understanding the *in vitro* and *in vivo* multifunctionality of this important chromatin

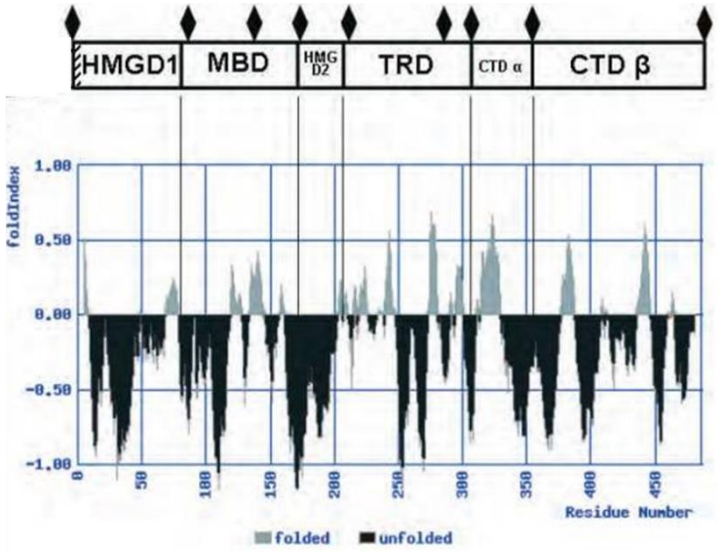


Figure 20| MeCP2 domain organization and folding properties. The domains that make up the MeCP2 protein as identified by partial proteolysis (diamonds) are correspondingly annotated. In the lower part of the figure, a FoldIndex plot is shown, documenting predicted disorderd (black) and ordered (gray) regions. The image was adapted after (Hite et al., 2009).

regulator (Adams et al., 2007; Ragione et al., 2016). The increasing number of reports that uncover novel and diverse MeCP2 interactors come to support this view.

1.4.6.1.2 MeCP2 interacts with multiple partners and has ambivalent effects on transcription

Probably the best-characterized member of the list of interactors is the Sin3A/HDACs co-repressor complex that interacts with MeCP2 through the TRD domain. This interaction is responsible for recruiting the complex to DNA, leading to the deacetylation of neighboring nucleosomes and finally suppression of gene repression. The early elucidation of this cascade of events largely contributed to the traditional classification of MeCP2 as a transcriptional repressor (Jones et al., 1998; Nan et al., 1998). In addition to deacetylation, a link between MeCP2 and trimethylation of H3K9me3 was observed in some neuronal target genes. Here, a complex that includes CoREST (co-repressor for element-1-silncing transcription factor) locally recruits Sin3A/Suv39H1 for cumulative gene silencing by combined deacetylation and methylation of histone tails (Ballas et al., 2005; Fuks et al., 2003b). Sin3A-independent complexes were also discovered, where the corepressors SMRT (silencing mediator for retinoid and thyroid receptors) and N-CoR (nuclear receptor corepressor 1) were characterized as interaction partners (Kokura et al., 2001). Interestingly, the missense mutation R306C occurring in the MeCP2 TRD, which is the second most frequent mutation occurring in the neurodevelopmental Rett syndrome, abolishes the interaction between MeCP2 and the NCoR-SMRT co-repressor complex in *in vitro* assays. This is associated with the development of a Rett phenotype in heterozygous transgenic mice (Baker et al., 2013; Ebert et al., 2013; Heckman et al., 2014; Lyst et al., 2013). This outcome underlines the importance of an intact interaction network centered around the chromatin regulator MeCP2.

The traditional view of MeCP2 as a *bona fide* transcriptional repressor has been challenged by a series of recent reports (Ragione et al., 2016). The protein is most abundantly expressed in the adult brain where it reaches near histone-octamer levels and plays critical roles in chromatin organization together with linker histone H1. The protein levels steadily increase during the differentiation of neuronal cells in mice, rats and humans, where it seems to be more important for the maturation and maintenance of neurons, rather than for cell fate decisions (Cohen et al., 2003; Guy et al., 2011; Skene et al., 2010). Its high expression in the brain provides a molecular basis for the

broad range of neurodevelopmental diseases associated with MeCP2 mutations (Katz et al., 2016) (see below). Surprisingly, transcriptional profiles of the brain isolated from *Mecp2*-null mice have revealed only subtle changes of gene expression levels in comparison to wild-type animals (Tudor et al., 2002). Nevertheless, gene expression profiling of specific brain areas (i.e. hypothalamus and cerebellum), have uncovered dramatic changes in gene expression for thousands of genes. Strikingly, most of these appeared to be downregulated in the absence of MeCP2, indicating that the protein may also function as a transcriptional activator (Ben-Shachar et al., 2009; Chahrour et al., 2008; Sugino et al., 2014). This multivalent effect of MeCP2 could be attributed to the fact that the protein interacts with co-factors that, depending on the chromatin environment, can have either transcriptional repressive and activating roles (Ragione et al., 2016). Together, these findings highlight the importance of cell-type-specific analysis for a full mechanistically understanding of the functions of this versatile protein.

1.4.6.1.3 Chromatin distribution of MeCP2

Since its discovery, it has been known that MeCP2 is particularly enriched at mouse chromocenters (Lewis et al., 1992). Binding to this type of highly compacted chromatin is dependent on DNA methylation and requires the intact MBD (Nan et al., 1996; Yang et al., 2016). Here the protein appears to play important roles in the clustering of centromeres into groups of increasing size during differentiation. These cellular observations are in line with biochemical work documenting the ability of MeCP2 to compact chromatin fibers *in vitro* (Brero et al., 2005; Georgel et al., 2003; Nikitina et al., 2007a, 2007b). The importance of MeCP2 for differentiation-induced chromocenters clustering became apparent from studies using MeCP2-null neurons. These cells exhibited more abundant and smaller chromocenters that unlike for the wild-type counterparts failed to cluster during differentiation and following neuronal depolarization (Singleton et al., 2011). The direct involvement of MeCP2 in regulating the structure of chromatin indicates that the protein may rather act as a global modulator of transcription, rather than a gene-specific transcriptional repressor (Singleton et al., 2011; Skene et al., 2010).

The global role of MeCP2 is supported by data coming from chromatin immunoprecipitation (ChIP)-on-chip assays in the human SH-SY5Y neuronal cell line, revealing that 60% of the MeCP2 binding sites lie in intergenic regions. Interestingly,

when found at gene promoters, MeCP2 appeared to function rather as a transcriptional activator. In line with this, only 2,2% of the highly methylated genes displayed MeCP2 binding, contrasting with the 5mC-guided transcriptional repressive role proposed for the protein (Yasui et al., 2007). Accordingly, in the brain, more than half of MeCP2 molecules were found to reside in open chromatin, and MeCP2 overexpression led to an upregulation of gene expression (Guy et al., 2011). The abundance of MeCP2 in the brain, as well as the high number of available methylated cytosine, make it improbable that the protein acts a classical 'point-and-shoot' transcriptional regulator. Instead, the locus-specific action of MeCP2 is most probably under the combined control of PTMs and protein interactors (Guy et al., 2011).

In a seminal study, Lister *et al.*, have discovered that in both the mouse and human brain, the DNA methylation landscape is dynamically reconfigured during development (Lister et al., 2013) (**Figure 16**). By performing high resolution methylome analysis of the mammalian frontal cortex at key developmental stages, the authors found that unlike other differentiated cell types, such as glial cells, neurons show a progressive DNMT3A-driven increase of non-CG methylation (CH methylation or mCH, in which H = A, C, or T). In the adult mouse and human neuronal genome half of the methylated cytosines were present at non-CG sites, in particular at CA dinucleotides. Interestingly, non-CG methylation was depleted from neuronally expressed genes, suggesting an involvement of this modification in gene silencing (Lister et al., 2013). The contribution of DNMT3A to the setting of non-CG patterns in the brain was highlighted by a conditional knockout study, documenting that upon DNMT3A deletion, mCA levels dramatically dropped. By contrast, CG methylation remained largely unaffected (Gabel et al., 2015). Parallel to the increase in non-CG methylation, methylated CG-sites appear to be converted in a Tet-dependent manner into hydroxymethylated dinucleotides (hmCG), particularly at transcriptionally active genes (Kinde et al., 2015; Lister et al., 2013). Indeed, hmC levels were found to be ~ 10 fold enriched in the brain by comparison to ES cells, accounting for 40% of all modified CG sites in the cerebellum (Gabel et al., 2015; Guo et al., 2014c; Lister et al., 2013). This modification was shown to be enriched in the body of actively transcribed genes and depleted around the transcriptional start site (Kinde et al., 2015; Lister et al., 2013).

Noteworthy, this major restructuring of the brain methylome including the increase in hydroxymethylation and non-CG methylation, is closely correlated with the buildup of

MeCP2 in neurons (Skene et al., 2010). The hypothesis that MeCP2 might have a broader binding specificity and could be involved in the readout of mCH and hmCH, was consequentially considered by many groups, the investigations of which gave rise to conflicting results (Feng et al., 2010; Gabel et al., 2015; Guo et al., 2014c; Hashimoto et al., 2012; Khrapunov et al., 2014; Lister et al., 2013; Meehan et al., 1992; Mellén et al., 2012; Spruijt et al., 2013; Valinluck et al., 2004; Xie et al., 2012). Motivated by the large amount of disparate data, Greenberg and co-workers recently formulated a unifying conceptual framework to explain the binding of MeCP2 to brain chromatin (Kinde et al., 2015). At the core of their model, lies the observation that conversion of methyl-cytosine into hydroxymethyl-cytosine impairs MeCP2 binding, selectively when it occurs CG context. By contrast, hydroxymethylated CA sites are bound as efficiently as mCA dinucleotides (Gabel et al., 2015). The authors propose that the wave of hmCG appearance in the mammalian brain would serve as a 'kick-off' signal to redistribute MeCP2 from the previously high-affinity mCG sites to the newly methylated CA dinucleotides (Kinde et al., 2015). Further work will be needed to explore the mechanisms behind this process.

1.4.6.1.4 The involvement of MeCP2 in disease

As described above, while MeCP2 is found in many tissues, the protein is most abundant in the mature brain (Shahbazian et al., 2002). Direct protein quantification revealed that there are around 16 million molecules of MeCP2 in adult neurons, with almost an order of magnitude less in glial cells, and 30-fold less in the liver (Skene et al., 2010). In line with its role in neuronal maturation and not neurogenesis, MeCP2 levels are quite low at birth and, in mouse, display a dramatic increase in the first 3 post-natal weeks (Kishi and Macklis, 2004; Skene et al., 2010). By this time, neurogenesis is largely complete, but synaptogenesis is ongoing (Guy et al., 2011). Due to its high abundance in the brain, it is not surprising that misregulation of MeCP2, either through mutations or aberrant protein expression levels, was linked to several neurodevelopmental diseases.

The *MECP2* gene is found on the X-chromosome, indicating that *MECP2* mutations can have sex-specific outcomes (Adler et al., 1995). Indeed, mutations that severely affect the MeCP2 function are usually early lethal in males, whereas the mosaicism that arises due to random X-chromosome inactivation, leads to milder phenotypes in females (Amir et al., 1999; Schüle et al., 2008). Mutations affecting *MECP2* are most

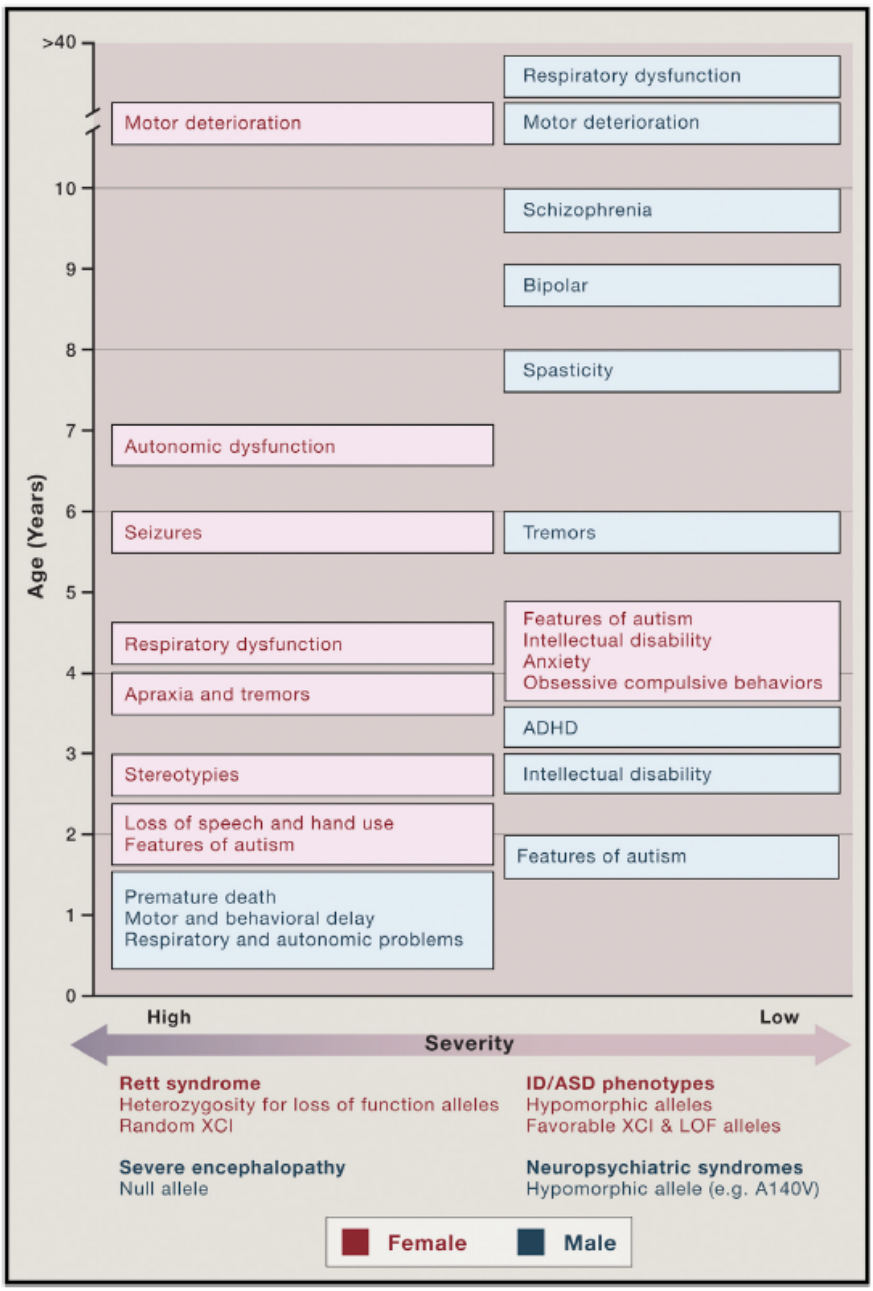


Figure 21| Clinical phenotypes associated with mutations in the X-linked *MECP2* gene. The most common phenotypes associated with deficiencies in *MECP2* are documented in the two panels. A severe phenotype is more often associated with males (blue), since they have only one X chromosome. In some cases, preferential inactivation of the X chromosome that carries an intact *MECP2*, can also lead to severe defects in females (red). Nevertheless, females generally express hypomorphic alleles, displaying milder defects. The chronological order of symptom appearance was annotated on the Y axis. The image was taken from (Zoghbi, 2016).

often associated with the neurodevelopmental Rett syndrome with ~95% of individuals with classical Rett phenotype carry mutations in *MECP2* (Smeets et al., 2012).

The Rett syndrome is characterized by an apparently normal postnatal development until 6-18 months of age, after which severe neurodevelopmental deficits start to appear. These include microcephaly, deterioration of motoric skills, features of autistic

behavior and intellectual disability (Zoghbi, 2016). The most common phenotypical changes associated with the development of the syndrome, together with their gender-specific peculiarities are depicted in **Figure 21**. Noteworthy, since the syndrome can be tracked down to mutations that affect only a single gene, it provides a complete environment for investigating the function of MeCP2, from its involvement in disease appearance to the pathological endpoint (Guy et al., 2011; Zoghbi, 2016). The self-sufficiency of the syndrome is highlighted by the finding that reintroduction of *Mecp2* in adulthood, can rescue the Rett-like phenotype developed by the *Mecp2* null mice (Lombardi et al., 2015).

At the molecular level, over 80 unique missense and 140 unique truncating mutations have been identified in girls with Rett syndrome (Christodoulou et al., 2003; Lombardi et al., 2015). As shown in **Figure 22**, where the most occurring mutations are documented, these are distributed over the whole protein, and cluster particularly in the MBD, hinting at a molecular connection between the 5mC signal and the development of the neurodevelopmental syndrome (Guy et al., 2011; Lombardi et al., 2015).

Interestingly, despite the fact that both DNMT3A and MECP2 are highly expressed in the brain, where they play important functions, and that the two proteins are members of the same epigenetic pathway, the potential interplay between them has not been addressed to date. In the present work, the molecular crosstalk between DNMT3A, the writer of DNA methylation, and MeCP2, the reader of DNA methylation, was investigated by a combinatorial approach involving *in vitro* and cellular analysis.

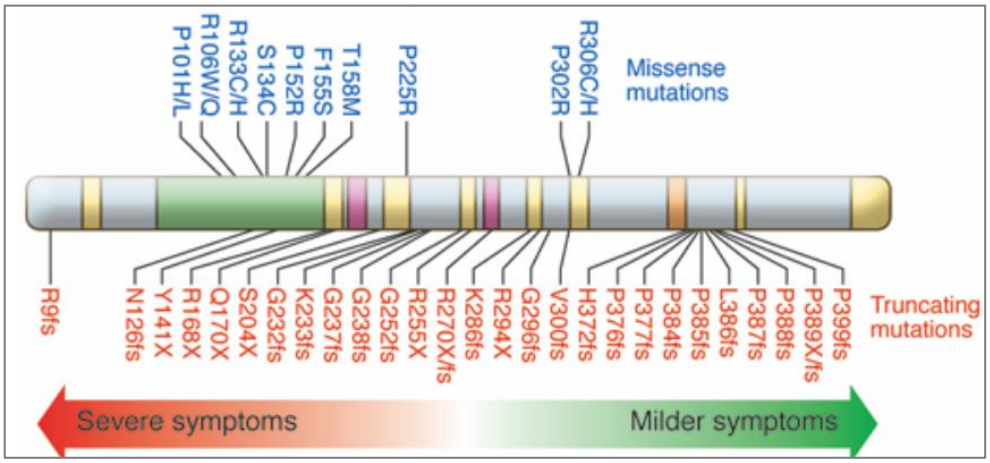


Figure 22| Established Rett-causing mutations in the X-linked *MECP2* gene. Missense mutations were annotated in blue, while truncating mutations are shown in orange. The MBD domain is color coded in green. Truncating mutations present before R270 are associated with more severe symptoms by comparison to the more C-terminally located mutations. The image was taken from (Zoghbi, 2016).

1.4.6.2 The methyl-CG-binding domain protein 1 (MBD1)

MBD1 stands out of the MBD family, owing to its unique and highly complex domain composition. The protein contains three specific CXXC domains, a transcriptional repression domain, and a canonical methyl binding domain. Interestingly, while the first two CXXC domains support the binding of MBD1 to methylated DNA, the third enables MBD1 to dock onto unmethylated genomic sites (Jørgensen et al., 2004; Li et al., 2015a). The dual-affinity of the protein, might be responsible for the involvement of MBD1 in different epigenetic pathways, from DNA methylation to modification of histone tails (Fujita et al., 2000, 2003; Li et al., 2015a; Nakao et al., 2001; Ng et al., 2000; Sarraf and Stancheva, 2004).

The N-terminally located MBD remains the main interface that targets MBD1 to methylated CG islands of tumor suppressors and imprinted genes, where the TRD further recruits chromatin repressive complexes and induces gene silencing (Li et al., 2015a). The first insights into the mechanistic recognition of the methylated CG dinucleotide were provided by an NMR-based structure of the MBD in complex with a methylated DNA oligo (Ohki et al., 2001). This revealed that four β-strands and an α-helix build the central domain carrying two loops which project into the major groove and make contacts with the DNA (Ohki et al., 2001; Patel, 2016). More precisely, the hairpin-like loop L1 is responsible for forming contacts with the major groove of the DNA and for interacting with the methylated cytosine on one of the DNA strands

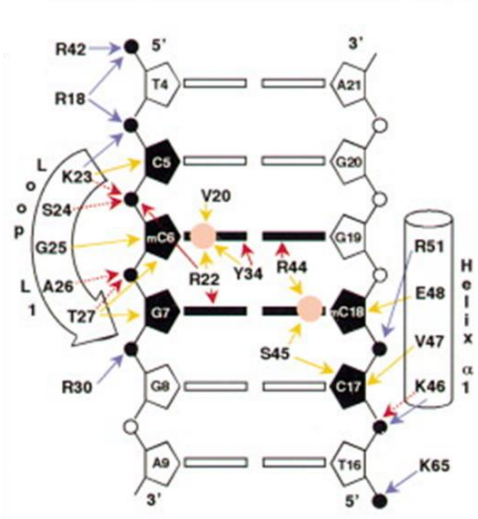


Figure 23| Recognition of the symmetrically methylated CG site by the MBD of MBD1. Schematic representation of the protein-DNA contacts formed between the MBD domain and the DNA helix. The methylated CG sites are denoted with pink circles. The DNA bases that interact with the protein are shown in black. The interactions between the MBD and the DNA helix are depicted with arrows as follows: hydrophobic (yellow), hydrogen bond (red), and electrostatic (blue). Red dashed arrows indicate hydrogen bonds involving backbone amide groups. The image was taken from (Ohki et al., 2001)

(**Figure 23**). The recognition of the 5mC on the complementary strand is more complex and involves loop L2 and a part of the α-helix that builds the central protein core. Here, the loop forms base-specific contacts in the major groove, whereas the helical segment makes contacts with the sugar-phosphate backbone. The methyl group is embedded in a hydrophobic pocket formed by 5 crucial residues: V20, R22, Y34, R44, and S45. Interestingly, this patch is able to access the strand-specific information of the methylated cytosine with the involved amino acids forming distinct interactions with each of the two 5mC (**Figure 23**) (Ohki et al., 2001; Patel, 2016). The importance of these residues is underlined by gel shift experiments showing that exchange of any of these amino acids to A compromised DNA binding (Ohki et al., 2001). These protein variants also displayed altered cellular localization patterns, when used in transient transfection of mouse fibroblasts, documenting the high dependency of MBD1 on an intact MBD domain for the proper recognition of 5mC sites. Importantly, the expression of an MBD1 variant lacking the MBD domain, but with intact CXXC domains, also displayed a diffuse localization and dramatically reduced enrichment at methylated sites (Fujita et al., 2000).

Noteworthy, in addition to identifying the contact points that are necessary for 5mC recognition, the crystal structure also revealed that the MBD displays a small protein-DNA interface, that is essentially restricted to the 5mC/5mC binding site (Ohki et al., 2001). This suggests that the domain is able to access the methylation status of the cytosine base within a narrow area of the major groove and without extending into the minor groove (Ohki et al., 2001; Patel, 2016). The authors speculate that this recognition modus might allow the protein to easily bind to nucleosomal DNA, without encountering steric hindrances from the N-terminal tails of histone H2A, H2B, and H3, which protrude into the minor groove of the DNA helix. Therefore, unlike other DNA proteins that need to embrace the DNA for binding, the MBD of MBD1 can also dock on nucleosomal-wrapped DNA. Importantly, due to the small recognition motif, with the absence of apparent base-protein contacts outside of the CG dinucleotide sequence, the MBD appears not to be biased in the recognition of particular DNA sequences (**Figure 23**) (Ohki et al., 2001). Follow up biochemical experiments have indicated that the domain might have a preference for TCGCA or TGCGCA motifs, although here a longer construct was used for gel shift experiments (Clouaire et al., 2010). The high affinity of the protein for methylated cytosines, together with its low sequence bias, made the MBD a suitable tool for protein engineering approaches, in the search for

improved affinity reagents that recognize 5mC-rich genomic sites (see below) (Yu et al., 2010). In this study, the high specificity and affinity of the MBD domain for methylated CG dinucleotides was exploited to design novel epigenetic sensors that specifically recognize methylated genomic sites in living cells.

1.5 Understanding epigenomic phenomena with state of the art profiling methods

Unlike the genetic code, epigenetic modifications are metastable, with different marks having different degrees of stability and variability (Bheda and Schneider, 2014; Daxinger and Whitelaw, 2010; Heard and Martienssen, 2014; Kouzarides, 2007). While some epigenetic signatures are stably inherited across many cell divisions and even generations, others are rapidly erased (Bheda and Schneider, 2014). This heterogeneity is essential for the generation of a flexible and functional interface between the genome and the environment (Feil and Fraga, 2012; Herceg and Vaissière, 2011). Importantly, while epigenetic landscapes are indispensable for the functional specialization of cells during differentiation, they are not identical within a population of cells belonging to the same cell type (Chang et al., 2008; Landan et al., 2012; Schwartzman and Tanay, 2015; Shalek et al., 2013). This high cell-to-cell variation together with the heterogeneous changes that appear in the epigenetic landscape over time poses several challenges to a complete understanding of how epigenetic states are set, maintained, and erased in both normal development and pathogenesis (Bheda and Schneider, 2014). Indeed, a large body of work documented that deregulation of epigenetic cascades is intimately linked to disease development, in particular, cancer. The high frequency and early appearance of alterations in the epigenetic landscape in malignancy have placed the understanding of epigenetic signals at the core of biomarker discovery, therapeutic intervention and prevention (Lima et al., 2010; Lister and Ecker, 2009; Umer and Herceg, 2013). This has targeted the development of increasingly refined methods for profiling epigenetic modifications, which was also one of the aims of this work. Therefore, in the following sections, an overview of the main techniques that enable the detection of histone modifications and DNA methylation will be provided. Inherent advantages and limitations will be described.

1.5.1 Methods for profiling histone modifications

Analysis of histone modifications relies on two main workhorses: mass spectrometry (MS) and chromatin immunoprecipitation (ChIP). Each of these two techniques has massively contributed to the development of the chromatin field and comes with a unique set of advantages and drawbacks.

1.5.1.1 Comprehensive mapping of histone modifications with mass spectrometry

MS-based proteomic approaches have the unmatched power to provide an unbiased and comprehensive map of histone PTMs. Derivations of this technique have facilitated the discovery of new histone modifications, aided in elucidating the interplay between different histone marks and provided an accurate measure of the quantitative changes in transitions between distinct functional states (Britton et al., 2011; Garcia, 2009; Plazas-Mayorca et al., 2009; Sidoli et al., 2012; Soldi et al., 2014; Zee et al., 2011). In MS, a post-translational modification is detected as an increase in the mass of the peptide between the theoretical and the experimentally measured mass. Histone PTM analysis by MS can be performed *via* three main experimental workflows: bottom-up, middle-down and top-down. The bottom-up approach involves the enzymatic digestion of the histone proteins into small peptides, which are separated *via* chromatography before MS analysis, sequencing, and quantification (Chait et al., 2006). For this procedure, trypsin cleavage is routinely used, due to its high specificity for lysines and arginines (Olsen et al., 2004). In the case of histone proteins, which are lysine/arginine-rich, this methodology comes nevertheless with two disadvantages: on the one hand, the enzymatic digestion leads to the generation of many very small peptides, which can pose problems for the chromatography run, as well as for the MS analysis; on the other hand, trypsin cleavage was shown to be blocked by modified arginine or lysines, giving rise to quantification issues (Karch et al., 2013). To circumvent these caveats, middle and top-down approaches were developed, involving rarely-cutting proteases such as AspN and GluC, and analysis of undigested proteins, respectively. These techniques are particularly useful for studying the co-occurrence of two or more PTMs (Karch et al., 2013). The major impact of the development of MS techniques is attested by the around 200 distinct histone modification types, which were discovered and mapped, with MS methods, in the last 10 years (Arnaudo and Garcia, 2013). The major limitation of MS proteomic approaches is the fact that they lack DNA-sequence

information, thereby they are conceptually not able to provide locus-specific epigenetic information.

1.5.1.2 Locus-specific analysis of histone modifications with chromatin immunoprecipitation

The analysis of histone modifications with locus-specific resolution can be achieved through chromatin enrichment-based methods. These generally involve cellular lysis, followed by nuclease- or sonication-based chromatin fragmentation and precipitation of the nucleosomes of interest with affinity probes. To extract the genomic sequence information, the DNA associated with the immunoprecipitated histones is isolated and subjected to analysis by either quantitative PCR (qPCR) or high-throughput sequencing, depending on the amount of information required (Kimura, 2013).

Two main types of affinity reagents have been used so far to retrieve modified nucleosomes: antibodies and recombinant proteins based on histone modification interacting domains (HIMDs) (Kungulovski et al., 2015a). Out of these, PTM-specific antibodies have been the traditional reagents employed for translating the language of histone marks into experimental observations. Recently, the reliability and quality of these reagents have however started to be questioned, as large lot-to-lot variability in specificity and binding affinity were observed (Baker, 2015; Bordeaux et al., 2010; Kungulovski et al., 2014). Availability of highly specific affinity reagents is particularly critical in the chromatin field, where histone modifications that participate in different biological pathways, can occur on amino acid motifs that are very similar in sequence (Kungulovski et al., 2014). To circumvent the caveats associated with antibodies, engineered histone modification interaction domains have been developed (Kungulovski et al., 2014). They exploit the inherent discriminative power encoded in chromatin readers, domains that are part of natural occurring chromatin effectors. These tools were successfully employed to study the genomic distribution of H3K9me3, H3K27me3, H3K4me0 and H3K4me3 modifications (Kungulovski et al., 2014, 2016; Su et al., 2014). Among the advantages that recombinant binders have over antibodies are their constant quality and easiness to assess unspecific binding by generating mutants in the histone binding pocket. In addition, combinatorial histone marks can be more readily studied by using naturally occurring multi-domains or synthetically-fused single binders. This aspect remains so far difficult to address with antibody-based technology (Fuchs et al., 2011; Su and Denu, 2016).

Regardless of the reagent used for affinity enrichment, ChIP-based methods, as well as derivation thereof, significantly contributed to the generation of comprehensive maps of the epigenetic makeup for many cell types and tissues. The resulting near-complete catalogues of functional elements available for many model organisms, provide a powerful screening platform for elements that are altered in malignancy (Bernstein et al., 2012; Consortium et al., 2015; Farh et al., 2015; Seumois et al., 2014; Yue et al., 2014).

Noteworthy, while these methods provide fundamental insights into the mechanisms responsible for the setting and maintenance of epigenetic patterns, they are very often used on chromatin isolated from a population of cells, therefore averaging cell-to-cell heterogeneities (Bheda and Schneider, 2014). For improved assessment and understanding of the diversities occurring in-between epigenetic landscape, single-cell methods started to be developed.

1.5.1.3 Genome-wide single-cell mapping of histone modifications: the next generation

By building up on the advent of sequencing technologies, current methods enable the profiling of gene expression or genome organization down to the level of individual cells (Angermueller et al., 2016; Burton et al., 2013; Cusanovich et al., 2015; Levsky et al., 2002; Nagano et al., 2013; Ståhlberg and Bengtsson, 2010). Whereas DNA methylation can be recovered from the genomic material directly, profiling of histone modifications is performed indirectly, by chromatin immunoprecipitation followed by sequencing (ChIP-seq). This procedure enriches for the modification of interest 5-100-fold over the genomic background, implying that a large proportion of the sequenced reads are nonspecific. To account for this, the immunoprecipitated material is usually sequenced with a high depth to distinguish positive signals from the background by statistical analysis (Schwartzman and Tanay, 2015). Applying the same approach to single cells remains problematic, as, without the statistical pooling, there is no reliable way to distinguish real signals from false positives. Moreover, when working with only one genome copy, there is a large risk of losing signal due to technical problems and not because the modification of interest is lacking. This can lead to false negative results. For this reason, methods that allow the genome-wide single-cell profiling of histone modifications are still under development (Schwartzman and Tanay, 2015).

1.5.1.4 Live-cell imaging of histone modifications

As described in the introduction, an important feature of epigenetic modifications is their inheritance. Despite the booming development of available detection methods, this aspect remained so far difficult to access, since most of the techniques require cell fixation or lysis. To achieve a bird's-eye view on the dynamics of the epigenetic landscape, live cell microscopy-based methods were developed. These typically rely on *in vivo* fluorescence reporters, to visualize gene expression and inheritance of expression patterns in single live cells (Bheda and Schneider, 2014).

1.5.1.4.1 Detection of histone modifications via Fabs and mintbodies

To monitor histone modifications in living cells, two types of probes have been developed (**Figure 24**). The first approach relies on the direct detection of endogenous histone modifications using fluorescently-labeled modification-specific antigen binding fragments (Fabs) (**Figure 24a**) or genetically encoded modification-specific intracellular antibody probes (mintbodies) (**Figure 24b**) (Hayashi-Takanaka et al., 2009, 2011; Rothbauer et al., 2006; Sato et al., 2013, 2016; Suzuki et al., 2016). Fab fragments are generated through protease digestion from the whole immunoglobulin G (IgG). For microscopy applications, these small fragments are then labeled with fluorescent dyes before being mechanically loaded or injected into living cells, where they enter the nucleus by passive diffusion through the nuclear pores (**Figure 24a**) (Kimura and Yamagata, 2014; Rader, 2009). Residence time measurements by fluorescence recovery after photobleaching (FRAP), revealed that Fabs bind their

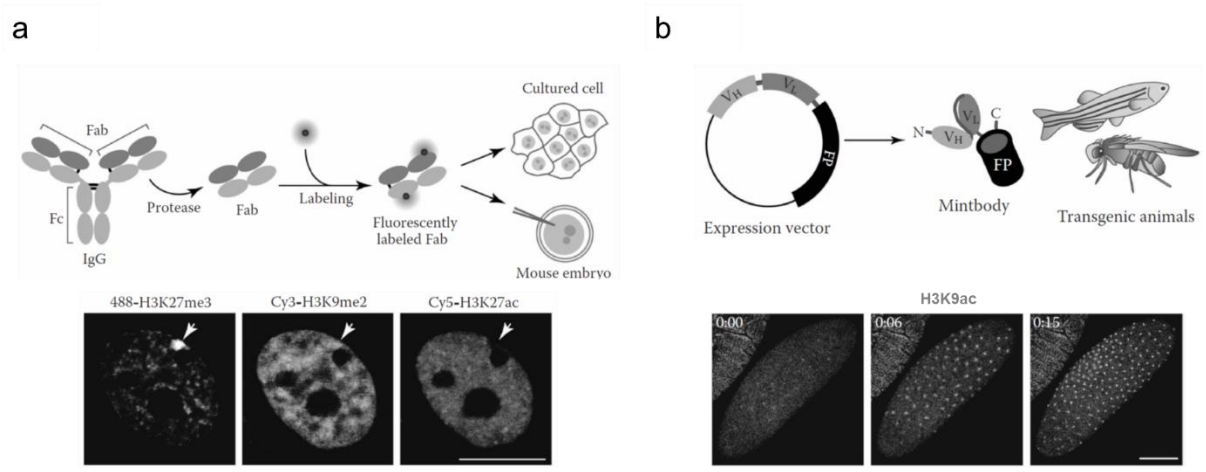


Figure 24| Live cell imaging of histone modifications with Fab fragments (a) and mintbodies (b). a) Schematic representation of the Fab labeling technology. The lower part of the figure documents the application of fluorescently-labeled Fab fragments for the simultaneous detection of three distinct histone marks, in the nuclei of living mouse cells. **b)** Schematic representation of the mintbody approach. The lower panel documents the application of a H3K9ac-mintbody to monitor the increase in the levels of the histone mark, in live Drosophila embryos. The figure was adapted after (Shen, 2015) .

targets for a short time ranging from <1 to >10 s depending on the histone modification, suggesting that these tools do not dramatically disturb cellular processes, even when loaded at high concentrations (Hayashi-Takanaka et al., 2011). Using this technique, H3K9 and H3K27 acetylation levels could be visualized in mouse preimplantation embryos (Hayashi-Takanaka et al., 2011). Triple labeling of H3K27me3, H3K9me2 and H3K27ac was also achieved (**Figure 24a**). A disadvantage of the Fab technology is that the fragments are diluted during cell divisions, making long-term imaging as well as living animal experiments infeasible (Bheda and Schneider, 2014). To overcome this, genetically encoded mintbodies were developed. These can be constitutively expressed *in vivo* as fusions with fluorescent proteins. Using this method, the dynamics of H4K20me1 during the cell cycle and X chromosome inactivation could be successfully visualized (Sato et al., 2016). Although providing the first glimpses in the visualization of chromatin marks, this approach can result in high signal-to-noise ratio caused by the fluorescence of expressed but unbound mintbody probes (Bheda and Schneider, 2014). This limitation can become be particularly severe, the less abundant the histone modification under investigation is.

1.5.1.4.2 Detection of the activity of epigenetic enzymes via FRET biosensors

To circumvent the high background issue, a second approach was developed, relying on Förster resonance energy transfer (FRET) probes (Ito et al., 2011b; Lin et al., 2004; Nakaoka et al., 2016; Sasaki et al., 2009). With these sensors, the balance between modifying and de-modifying histone enzymes can be monitored (**Figure 25a** and **b**). The tools resort to intramolecular FRET between a donor and an acceptor fluorescent protein, which are typically incorporated in a tripartite system consisting out of a FRET

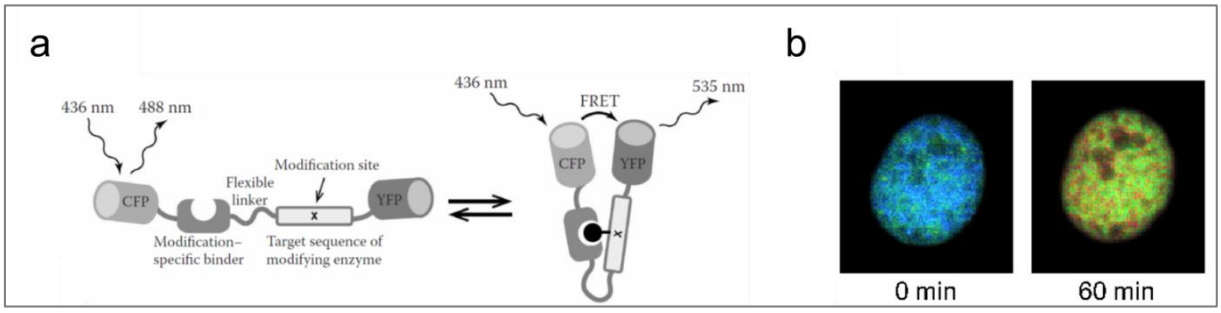


Figure 25| FRET-based detection of histone tail modifications in living cells. a) Schematic representation of a tripartite FRET sensor, where a linker encoding for the target sequence of a histone modifying enzyme, is used to separate a CFP-tagged modification-specific reader from YFP. Successful modification of the linker sequences triggers a large intramolecular conformational change that brings the CFP and the YFP fluorophore in a FRET-productive spatial proximity. The figure was taken from (Shen, 2015) .**b**) Fluorescence microscopy image documenting an increase in FRET signal for a sensor designed to readout the acetylation levels of histone H4. The right-side image was taken 60 min after treating the cells with the histone deacetylase inhibitor Trichostatin A (Sasaki et al., 2009).

donor-fused modification-specific binder, a flexible linker containing an amino acid sequence motif to be modified by the investigated enzyme and a FRET acceptor. The enzymatic modification of the target sequence motif is recognized by the modification-specific binder, initiating an intramolecular structural change that brings the donor and acceptor fluorophores in close spatial proximity, thereby giving rise to an increase in the FRET signal. The advantage of the FRET readout is that the signal is generated only when the two fluorophores are very close to each other, and particular spatial constraints are met (Vogel et al., 2006). Accordingly, an output signal is only produced if the central linker peptide gets modified, limiting the fluorescence background. With these tools, the levels of H4K12 acetylation could be successfully monitored in real-time, by using the bromodomain of BRD2 (bromodomain-containing 2) as a modification-specific reader (Ito et al., 2011b). In addition, the increase in H4 acetylation could be monitored in living cells, upon treatment with a histone deacetylase inhibitor (**Figure 25b**).

Importantly, while these methodologies offer *in vivo* expression and reversible, low-background advantages, they can be only used to monitor the global distribution of histone modifications and have no locus-specific resolution (Bheda and Schneider, 2014).

1.5.2 Methods for profiling DNA methylation

A large number of experimental studies have demonstrated that deregulation of DNA methylation is intimately linked to many human malignancies, most notably cancer. Here, CG islands of tumor suppressor genes were reported to usually gain DNA methylation, while repetitive elements such as heterochromatic DNA repeats become frequently hypomethylated (Baylin, 2005; Baylin and Jones, 2011; Taberlay and Jones, 2011). These changes may promote the growth, survival, and progression of the invasive tumors, thereby fostering malignancy (Baylin and Jones, 2011). This, together with the dynamic changes that occur in DNA methylation levels during development, have placed the mapping of the 5mC mark at the core of novel method development.

1.5.2.1 Methods to map 5mC in lysed cellular material

Standard molecular biology techniques such as hybridization and sequencing, are not able to discriminate between methylated and unmethylated cytosine bases (Umer and

Herceg, 2013). Owing to this, DNA methylation profiling requires that the sample DNA undergoes certain pre-treatments, that can preserve and are sensitive to the modification status of the 5mC group. These include DNA digestion with methylation-sensitive restriction nuclease, affinity enrichment with 5mC-specific probes, MS, or bisulfite (BS) treatment (Umer and Herceg, 2013). The latter is the working horse of methylome analyses and involves the chemical treatment of the isolated DNA with sodium bisulfite. The key concept of the method is that this treatment leads to selective deamination of unmethylated cytosines to uracil. By contrast, the conversion of methylated bases has a much lower rate, and when the treatment is performed under optimal conditions, it leaves the modified bases unaffected. In subsequent PCR amplification, the uracil-derived from unmodified cytosines is replaced with thymine, while the methylated base leaves a cytosine footprint (Frommer et al., 1992). Coupling this method with the development of high-throughput sequencing has led to the generation of whole-genome DNA methylation landscapes, at single base resolution, for a broad range of cell types and conditions (Darst et al., 2010; Li and Tollefsbol, 2011).

Like the profiling of histone tails marks, methods routinely used for mapping 5mC suffer from two main limitations: the analysis of population-based averages and limited resolution power for the assessment of dynamic epigenetic changes, as these methods cannot be applied in living cells (Bheda and Schneider, 2014). Several new approaches were developed to address these issues. As such, the first single-cell methylome was published in 2014 and was based on a joined effort between the Reik and Kelsey labs (Smallwood et al., 2014). This novel platform involves isolation of single cells, followed by lysis and BS treatment for combined DNA fragmentation and chemical conversion of unmethylated cytosines to thymine. Sequencing adaptors are sequentially ligated to the resulting fragments, in a 5-times repeated process, to maximize the number of tagged DNA strands and generate multiple copies of the starting material. With this strategy, the authors were able to successfully assess the epigenetic heterogeneity that arises in-between embryonic stem cells, upon transitions in the culturing conditions (Smallwood et al., 2014). This pioneering report laid the basis for a plethora of follow up method development studies which contributed fundamental insights into the high heterogeneity of epigenetic landscapes (Farlik et al., 2015; Guo et al., 2015a, 2015b).

1.5.2.2 Live-cell imaging of global DNA methylation levels

By analogy to histone marks, microscopy-based methods were also developed for accessing the dynamics of the 5mC modification. This epigenetic mark comes however with the caveat that it is best exposed to antibody binding when DNA is single stranded. To this end, the cellular material needs to be subjected to acidic denaturation and high-temperature hybridization, which make antibody-based approaches incompatible with live cell imaging (Jørgensen et al., 2006). For this reason, natural protein reading domains are more often exploited as a tool for the live-cell tracking of DNA methylation by comparison to the monitoring of histone modifications, where Fab fragments and mintbodies are more frequently used. The most established reader of DNA methylation for live-cell applications is the MBD of MBD1. As described above, this domain has the advantage that it recognizes 5mC in the native double-stranded DNA context, and it has a high specificity for symmetrically methylated CG dinucleotides (Jørgensen et al., 2006). The development of this protein for live cell applications was ignited by a pioneering study where the authors were able to visualize for the first time, the dynamics of DNA methylation levels in preimplantation mouse embryos, upon injecting them with mRNA encoding for a GFP-tagged fusion of the MBD domain (Yamazaki et al., 2007). With this single-cell technology, the group found that centromeric repetitive sequences are hypomethylated in germ cells by comparison to somatic cells

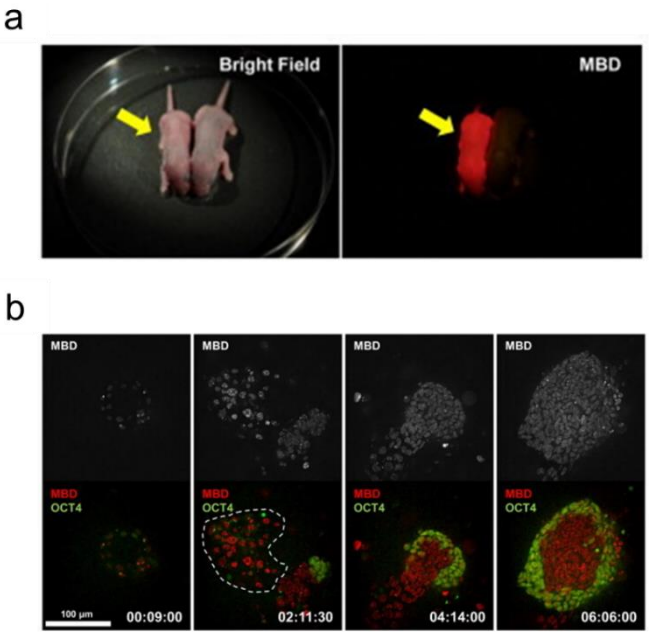


Figure 26| Generation and usage of the methyl*RO* mouse reporter model. a) Neonate MethylRO mice (left) have similar phenotype to wild-type mice (right) and show whole-body red fluorescence. **b)** Time lapse imaging of embryonic stem cells during the differentiation process (tracked with OCT4-GFP). A strong increase in the MBD-RFP signal intensity is observed with time (indicated at the bottom right as day: hour: minute). The images were taken from (Ueda et al., 2014).

(Yamagata et al., 2007). The low toxicity and high specificity of this device was recently exploited for the development of a DNA methylation mouse reporter. Here, the authors generated a knock-in animal to ubiquitously express a RFP-fused MBD protein (**Figure 26a**). With this, the mobility and pattern of heterochromatin, as well as the DNA methylation signal intensity could be dynamically tracked during cellular differentiation (**Figure 26b**) (Ueda et al., 2014). Although useful, this approach only reports on the global level of DNA methylation and lacks locus-specific resolution.

1.5.2.3 Visualization of chromatin marks with single locus resolution

While there are now many methods that enable the tracking of epigenetic modifications at various scales, no technological platform enables the detection of epigenetic marks at the level of single loci, in living cells. The paramount significance of the insight that could be achieved through such technologies, motivated several groups to develop methods that attempt to satisfy these criteria (Bheda and Schneider, 2014). For instance, locus-specific visualization of histone modifications was achieved by combining *in situ* hybridization with a proximity ligation assay (Gomez et al., 2013). To achieve DNA sequence specificity, biotinylated complementary probes were used for hybridization. The biotin moiety was then recognized by a secondary antibody, to which a DNA oligo is coupled. Similarly, the histone modification was recognized *via* an antibody, to which a DNA oligo is linked. When the two probes bind in close spatial proximity, the two oligos can be used for several rolling-circle amplification cycles, to generate a product that can be labeled with fluorescent DNA probes. With this, the H3K4me2 status could be visualized at specific promoters of active genes in single smooth muscle cells, in mammalian tissue samples (Gomez et al., 2013). Although this method provides an elegant solution for the low signal associated with single-copy genomic loci, it is not compatible with live cell imaging, thereby providing only a snapshot of the dynamic epigenetic landscape.

An *in* situ hybridization-based approach also recently enabled the sequence-specific visualization of DNA methylation at single cell level (Li et al., 2013). Here the authors took advantage of the differential reactivity of modified *vs.* unmodified cytosine for interstrand complex formation with osmium and bipyridine-containing nucleic acids. For this chemical treatment, the cellular material needed to be nevertheless subjected to fixation, thereby compromising the applicability of the method in living cells and preventing the assessment of dynamic changes in the 5mC mark, within single nuclei.

The most comprehensive solution for a method that allows dynamic imaging of locus-specific DNA methylation changes was recently proposed by Jaenisch and co-workers (Stelzer and Jaenisch, 2016; Stelzer et al., 2015). The authors designed a novel readout system based on inserting a reporter construct consisting out of the minimal promoter of the imprinted *Snrpn* gene driving the expression of a fluorescent protein, downstream of the genomic locus of interest. Since the *Snrpn* gene contains somatic DMRs, the methylation of the promoter of interest can extend into the *Snrpn* promoter and translate into a reduction of the fluorescent signal. This sequence was used to drive the expression of a fluorophore, which served as a signal amplifier (**Figure 27a**). With this tool, the authors could for the first time, dynamically visualize the methylation status of several endogenous single copy loci (Stelzer et al., 2015). While this method holds great promise for mechanistic studies with potentially broad implications for the field (**Figure 27b**), it still comes with a set of limitations. Among the most important is the fact that the methylation readout is indirect, as it depends on the spreading of DNA methylation into the *Snrpn* DMRs, which may occur with locus-specific kinetics. Moreover, the endogenous genomic locus needs to be synthetically modified in order

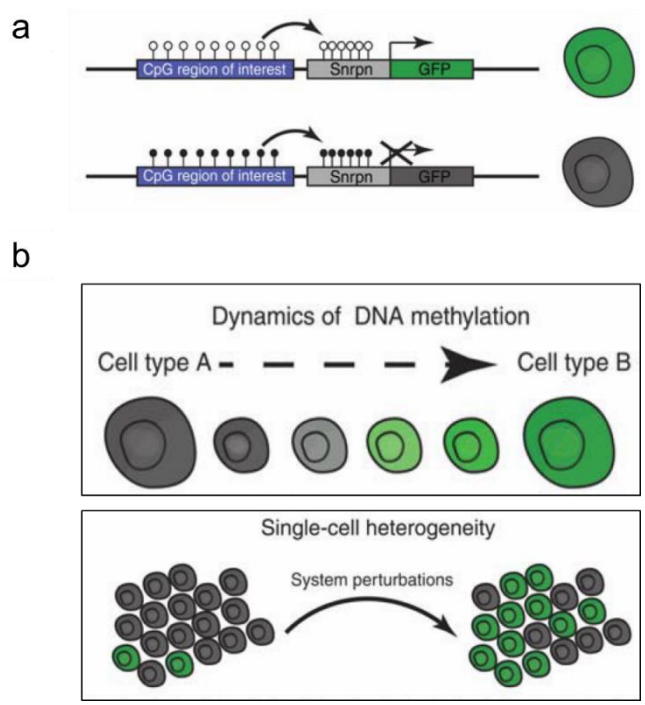


Figure 27| Principle and potential applications of the methylation reporter system developed by (Stelzer et al., 2015). a) Schematic representation documenting the concept of the DNA methylation sensor. Accordingly, DNA methylation of the neighboring CGI spreads into the *Snrpn* gene promoter, leading to epigenetic silencing and no detectable fluoresce. By contrast, unmethylated CGI are associated with a positive fluorescent signal. **b)** Potential application areas for the sensor described in a) involve imaging of the dynamics of DNA methylation during cellular differentiation (top panel) or studying cell-to-cell heterogeneity upon perturbation (lower panel). The images were taken from (Stelzer and Jaenisch, 2016).

to introduce the fluorescence reporter and the methylation kinetics measurements are biased by the half-life of the reporter fluorophore. To solve these issues, the development of methods that enable a direct targeting of genomic sites and facilitate the readout of their epigenetic status in live cells is imperative. This technological revolution is dependent on programmable tools that enable the live-cell visualization of user-defined DNA sequences. In the following section, technologies that enable locus-specific genomic targeting will be described. Particular emphasis will be put on their application for live cell imaging approaches.

1.6 Custom DNA-binding domains for imaging of user-defined genomic sites

There are three main types of powerful tools that were developed for the sequence-specific interrogation of genomic sites: zinc-finger proteins (ZFs), transcription activator-like effectors (TALEs) and the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas-based RNA-guided systems (Chen et al., 2016a; Gaj et al., 2013; Kungulovski and Jeltsch, 2016).

1.6.1 Cys_2-His_2 zinc finger proteins

The Cys_2-His_2 zinc finger domain is among the most common types of eukaryotic DNA-binding motifs. An individual ZF domain consists of 30 amino acids that are arranged in a conserved ββα configuration. Residing on the surface of the α-helix and the loop, there are several important amino acids that engage in contacts with the major groove of the DNA duplex. Typically, each ZF domain recognizes a 3 bps-long motif, with varying levels of selectivity (Beerli and Barbas, 2002; Gaj et al., 2013; Mandell and Barbas, 2006). ZF proteins have laid the basis of the genome targeting revolution. This was initiated by the realization that individual zinc finger modules can be used to generate custom-made modular arrays, through which longer DNA sequences can be recognized. This technological advancement was facilitated by the structure-based discovery of a highly conserved linker sequence that enables the construction of synthetic arrays of 3-6 zinc-finger proteins with a footprint of 9-18 bps. Follow up biochemical investigations have characterized the relationship between protein length and specificity and critically inspected the potential crosstalk between neighboring ZF modules when arranged in one array (Liu et al., 1997; Segal et al., 2003). Importantly, with the extension of the target sequence length to 18 bps, this technology allowed for the first time, the specific targeting of unique, user-defined sequences in the human

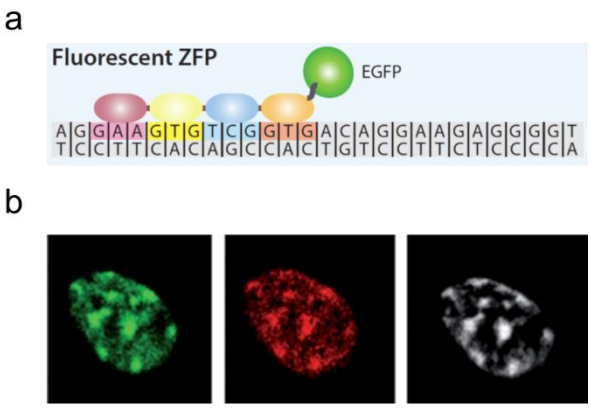


Figure 28| Application of ZF proteins for live-cell labeling of endogenous genomic sites. a) Schematic representation of a EGFP-fused modular zinc finger array, constructed by the fusion of 4 individual zinc finger modules. The nucleotide triples recognized by each module are shown in the same color as the corresponding ZF. The image was taken from (Chen et al., 2016a). **b**) Fluoresce microscopy image showing the live cell labeling of mouse major satellite repeats with a polydactyl EGFP-fused ZF (green). The targeting specificity of the DNA-binding protein was validated by co-staining of the nucleus with DRAQ5 (red). Due to its bias for AT-rich sequences the dye is preferentially enriched at the ZF target sites. This is highlighted in the merged image (white). The image was taken from (Lindhout et al., 2007)

genome (Beerli et al., 1998, 2000). In the past decade, the number of methodologies used for assembly of zinc finger arrays has increased dramatically, hinting at the large impact that this technology has had on genome-targeting applications (Gaj et al., 2013). Custom-made arrays can now be commercially bought, allowing the investigators to bypass the time-consuming process of zinc-finger construction and validation (Gaj et al., 2013).

The application of the zinc-finger technology for *in vivo* chromosome labeling was achieved with some delay and was pioneered by the group of Bert J. van der Zaal (**Figure 28a**). Together with his co-workers, he developed the first synthetic protein with which an endogenous DNA sequence could be labeled and tracked in living cells (**Figure 28b**) (Lindhout et al., 2007). The GFP-fused protein was designed to recognize a 9bp-long sequence within murine major satellite repeats and enabled the first visualization of these sites in living cells (Lindhout et al., 2007).

1.6.2 TALE proteins

TALEs, are proteins naturally secreted by the pathogenic bacteria *Xanthomonas* upon infection of plant hosts. The proteins recognize specific DNA sequences, about 20 bases long, *via* a large DNA-recognition domain composed of a series of 33-35-amino acid repeat domains, each of them specifically recognizing a single base pair. The specificity of each module resides in two hypervariable amino acids, known as the

repeat-variable di-residues (RVDs) (Deng et al., 2012; Mak et al., 2012). Like their ZF predecessors, individual TALE modules can be linked together into custom-made modular arrays that can be used for targeting of defined genomic loci. This technology has two main advantages over ZFs: there is no need to re-engineer the linker between repeats to construct long array, and since each module recognizes only a single base, there is a larger flexibility in the design of target sites, by comparison to ZFs that are triplet-confined (Gaj et al., 2013). Indeed, TALE arrays capable of binding 7-34 bp-long motifs could be successfully constructed (Garg et al., 2012). TALE systems are an attractive technological platform since the nucleotide binding code that links the RVD type to specified nucleotides is known (Doyle et al., 2012). This design nevertheless, comes with a severe drawback that impairs its applicability: since the individual TALE modules are genetically very similar, cloning them into large arrays poses an elevated technical challenge owing to their highly repetitive DNA. To overcome this issue, several methods have been developed to enable the error-prone assembly of TALE modules into custom-made arrays (Briggs et al., 2012; Cermak et al., 2011; Reyon et al., 2012; Schmid-Burgk et al., 2013).

The first locus-specific imaging technology based on TALE proteins stemmed out of the group of Torres-Paddila (Miyanari et al., 2013). Here, the authors established a live cell imaging method to visualize endogenous repetitive genomic sequences in living mouse cells (**Figure 29a** and **b**). Fussing fluorescent proteins to TALEs of different sequence specificities enabled the spatial tracking of user-defined sequences in the

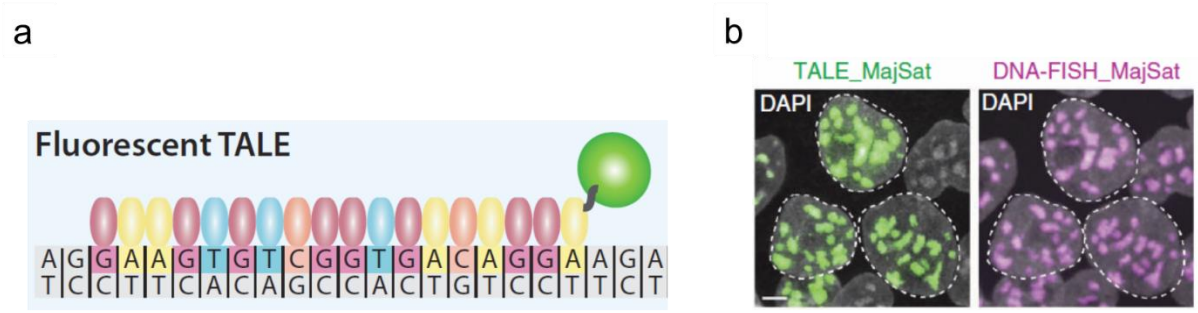


Figure 29| Application of TALE proteins for live-cell labeling of endogenous genomic sites. a) Schematic representation of a EGFP-fused TALE protein, designed to recognize an 18 bp-long DNA sequence. The RVD modules responsible for nucleobase specificity were colored similarly to the nucleotide they recognize. The image was taken from (Chen et al., 2016a). **b)** Fluoresce microscopy image documenting the high DNA sequence specificity of a GFP-TALE protein (green) designed to bind to mouse major satellites. To independently validate the identity of the target sequence, the TALE-transfected cells were subjected to fluorescent in situ hybridization (FISH) with a probe targeting major satellites (red). The green and the red signals appear to localize at the same genomic regions, highlighting the high specificity of the TALE construct. The image was taken from (Miyanari et al., 2013).

nucleus of cultured mouse cells, as well as in a living organism. The generation of stable cell lines and live animals expressing these devices underlined the low toxicity of the proteins and their applicability. Furthermore, by taking advantage of single-nucleotide polymorphism to differentially label parental chromosomes within on single nucleus, this study elegantly demonstrated the high specificity of the TALE technology. This pioneering work catalyzed the design and live cell microscopy application of TALEs with other sequence preferences (Fujimoto et al., 2016; Ma et al., 2013; Yuan and O'Farrell, 2016; Yuan et al., 2014).

A caveat associated with the usability of TALE proteins for cellular applications is brought about by the large size of the protein, which might perturb the local chromatin environment of the labeled locus. This was assessed for a TALE designed to bind mouse satellite repeats, where histone H3 occupancy, as well as the level of trimethylated H3K9, were found to remain unaffected (Miyanari et al., 2013). However, the effect that TALEs might have when used to label protein-coding single copy genes as well as regulatory DNA sequences such as promoters and enhancers, remains to be elucidated (Chen et al., 2016a).

1.6.3 The CRISPR/Cas9 system

Recently, the CRISPR/Cas system has emerged as a potent alternative to ZF and TALE proteins for genome targeting applications. The system is naturally occurring in bacteria and archaea, where it provides acquired immunity against invading foreign DNA *via* RNA-guided DNA cleavage (Wiedenheft et al., 2012). CRISPR/Cas systems consist of 2 main components: a set of *cas9* genes organized in operon(s) and the CRISPR array(s) comprising short repetitive elements DNA elements, separated by unique 30-40 bp-long sequences (spacers), which are derived from viral or plasmid origin (Marraffini, 2015). The CRISPR/Cas9-mediated adaptive immunity occurs in three steps. First, a small sequence of the invading DNA is inserted as a spacer motif into the CRISPR array. This is followed by the transcription of the locus to generate precursor RNAs (pre-crRNAs), which then undergo maturation to produce individual crRNAs, consisting of a repeat part and an invader-targeting spacer part. These crRNAs are finally recognized by nuclease-competent Cas9 proteins, which are guided to sites complementary to the crRNA spacer sequences to cleave the foreign nucleic acid (Doudna and Charpentier, 2014; Jiang and Doudna, 2015; Marraffini, 2015; van der Oost et al., 2014). While the cascade of events that triggers the acquisition of new

spacers is not completely elucidated, chromosomal acquisition data from *E. coli,* indicated that a large contributor may be the nucleolytic processing of dsDNA brakes generated during the replication of the virus or the plasmid (Levy et al., 2015; Marraffini, 2015). This immunization through spacer acquisition enables the bacterial population to rapidly acquire resistance to invading viruses (Marraffini, 2015).

There are 3 main types of CRISPR-Cas systems (I, II and III), each using a distinct molecular mechanism for the recognition and cleave off the target DNA sequence. While types I and III require a large complex of Cas proteins for crRNA-guided targeting, type II is based on only one protein for RNA-guided DNA recognition and cleavage. This property is one of the key features that contributed to the extensive application of the type II CRISPR-Cas system in genomic targeting approaches (Doudna and Charpentier, 2014; Gasiunas et al., 2012; Hale et al., 2009; Haurwitz et al., 2010; Jinek et al., 2012). The high usability of the CRISPR-Cas system was influenced by three main findings. First, trans-activating crRNAs (tracrRNA), which are trans-encoded upstream of the type II CRISPR-Cas locus in *Streptococcus pyogenes*, were found to be essential for the maturation of crRNAs (Deltcheva et al., 2011). Second, the *S.pyogenes* Cas9 protein was shown to be a dual-RNA-guided DNA endonuclease. In follow-up engineering approaches, this requirement could be simplified, as the dual tracrRNA:crRNA was successfully merged into a single guide RNA (sgRNA) that retained 2 critical properties: the 20-nucleotide sequence at the 5' end of the sgRNA, essential for Watson-Crick base pairing to the DNA site, and the double-stranded stem-loop structure at the 3' end of the guide sequence, which is bound by Cas9. (Deltcheva et al., 2011; Jinek et al., 2012). Third, with the successful identification of the residues responsible for DNA cleavage activity in Cas9, a nuclease-deficient version of the protein (deactivated Cas9 or dCas9) could be generated (Gasiunas et al., 2012; Jinek et al., 2012). The latter finding catalyzed a technological revolution in the field of genomic targeting as it provided a programmable RNA-guided DNA binding protein. Although other DNA targeting devices such as ZF and TALE proteins were already established at this time, the CRISPR/Cas9 technology was rapidly accepted and extensively implemented due to its unsurpassed flexibility. The main advancement is that while the RNA-guided CRISPR/Cas9 complexes can be easily re-targeted by designing new sgRNA, ZF and TALE repurposing involves *de novo* protein engineering, which can be very tedious and time-consuming (Chen et al., 2016a; Kungulovski and Jeltsch, 2016).

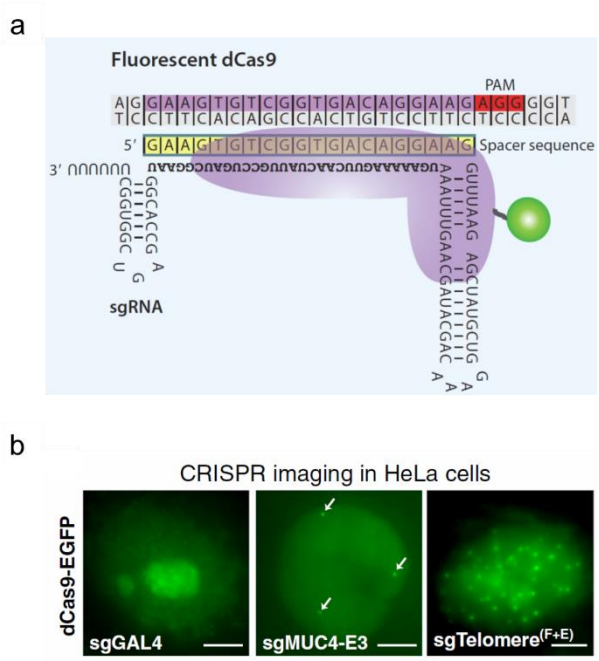


Figure 30| Application of the CRISPR/dCas9 technology for live-cell labeling of endogenous genomic sites. a) Schematic representation of a EGFP-fused dCas9 protein, describing its mechanism of sequence recognition. The dCas9 protein as well as the target site sequence are color coded in purple. The image was taken from (Chen et al., 2016a). **b)** Fluoresce microscopy image documenting the high flexibility of the CRISPR/dCas9 system. The EGFP-tagged protein displays a focal accumulation at 3 sites, when used to targeted single copy genes (middle panel), and multiple loci, when combined with a sgRNA recognizing repetitive telomeric sequences (right-hand pannel). The targeting specificity was validated by using a control guide RNA targeting the irrelevant GAL4 locus, where no distinct localization patterns were obtained. The image was taken from (Chen et al., 2013).

The first live cell imaging application of the CRISPR/Cas9 technology was developed by the groups of Huang and Qi (Chen et al., 2013). The authors used an optimized GFP-tagged dCas9 protein to selectively visualize specific genomic sequences in living human cells (**Figure 30a**). In addition to imaging of repetitive telomeric sequences, this approach was the first to enable the successful visualization of non-repetitive genomic sites (**Figure 30b**). To this end, a cocktail of 73 distinct sgRNAs was used to target the dCas9 protein in a tiled manner along the *MUC4* gene locus and amplify the fluorescent signal. This strategy highlights the unmatchable flexibility of the CRISPR/Cas9 technology, as a comparable ZF or TALE-based approach would have involved the *de novo* design and validation of 73 different DNA binding proteins (Chen et al., 2013). Following this elegant proof-of-concept paper, the Cas9 labeling approach has been successfully employed by other groups, to observe the dynamics of specific chromosomal loci in living mouse embryonic and cultured human cells (Anton et al., 2014; Chen and Huang, 2014; Knight et al., 2015).

Importantly, by contrast to ZF and TALE proteins, the association of the dCas9-sgRNA complex with the target genomic site is based on inducing a local double-stranded DNA unwinding that might have negative effects on the local chromatin environment, for instance by disturbing nucleosomal positioning or binding of transcription factors (Chen et al., 2016a; Sternberg et al., 2014). To evaluate this effect, the dynamics of dCas9-labeled telomeres was compared to the movement of these genomic sites when detected by fluorophore-tagged telomere-binding proteins such as TRF1. This comparison revealed that the dynamics of the telomeres as well as the integrity of telomere shelterin complexes were unperturbed, indicating that the dCas9-labeling technology has no detrimental effects on the local chromatin environment, at least not at the resolution of the experiment (Chen et al., 2013). Assessment of other genomic sites showed nevertheless that some sequences, such as enhancers, are more sensitive to dCas9 binding indicating that the influence of this technology on the local chromatin environment, has to be more intensively assessed on a locus-by-locus basis (Chen et al., 2013; Gilbert et al., 2014a, 2014b).

All in all, the recent years have been marked by a boost in the availability of custom-made DNA-targeting devices, which can be used to specifically detect and track endogenous genomic sites in living mammalian cells. While each of these tools comes with a unique set of advantages and limitations, they all enable the ideally unperturbed visualization of endogenous genomic loci, without the need of exogenously modify these sites. In this work, this technological progress was used as a basis for the generation of novel epigenetic biosensors, with which the status and dynamics of specific epigenetic signals can be assessed at the level of endogenous genomic loci in living cells.

2 Principal aims of the study

Since its discovery in 1948 by Hotchkiss, DNA methylation has become a topic of intense and exciting investigations. Although the importance of this mark was highlighted by early genetic studies and decades of active research, the epigenetic pathways centered around the 5mC signal are far from being fully understood (Jurkowska and Jeltsch, 2016). For instance, although mammalian DNMTs were identified almost 2 decades ago, the mechanisms that are responsible for controlling the chromatin targeting and enzymatic activity of these machineries are still poorly understood (Jurkowska and Jeltsch, 2016; Li et al., 1992; Okano et al., 1998).

It was the aim of this work, to explore three main directions of research, with the ultimate goal of shedding mechanistic and methodological insights into the generation and maintenance of DNA methylation patterns.

2.1 Investigation of the role played by ATP hydrolysis in the interaction of HELLS with chromatin

In the first project, I aimed to study the role played by ATP hydrolysis in regulating the interaction (and implicitly the chromatin remodeling function) between chromatin and the putative remodeler HELLS. This protein is a not well-understood member of the DNA methylation pathway and the knock-out of *Hells* was reported to lead to 50-70% reduction in global DNA methylation levels (Dennis et al., 2001). Currently, it is hypothesized that HELLS is necessary for facilitating the access of the DNA methylation machinery to sites of compacted chromatin, although an ATP-dependent or chromatin remodeling role could not be convincingly demonstrated for HELLS so far (Burrage et al., 2012; Dennis et al., 2001). The aim of the current study was to use the ATPase-deficient HELLS variant developed by Burrage et al. (2012) in order to systematically compare the chromatin association properties of this mutant with that of wild-type HELLS in a cellular-based approach. Insights from this work should contribute to a better understanding of how the association between HELLS and chromatin mediates the drastic effects that the protein has on the global DNA methylation patterns.

2.2 Investigation of the regulatory effects of MeCP2 on DNMT3A activity

The aim of the second project was to address the interaction between the *de novo* DNA MTase DNMT3A and the 5mC reading protein, MeCP2. This study was based on previous findings from our lab documenting a direct interaction between MeCP2, and an inhibitory effect on DNMT3A activity *in vitro*. The goals of the present work were to test whether this interaction can be also detected at the level of endogenously expressed proteins, in the mouse brain, and if the inhibitory effect observed on DNMT3A activity *in vitro* is recapitulated in a cellular context, upon overexpression of MeCP2 in cell culture, as well as in the mouse brain, where both DNM3A and MeCP2 are highly expressed. Furthermore, by taking into account the recently reported allosteric regulation of DNMT3 enzymes by interacting proteins, a third goal was to explore if the inhibition of DNMT3A *via* MeCP2 also occurs through an allosteric mechanism. Understanding the functional cross-talk between the readers and writers of DNA methylation could advance our knowledge of the feedback loops that are in place to control the activity of DNA methyltransferases.

2.3 Development of epigenetic sensors for locus-specific detection of epigenetic marks in living mammalian cells

In the third project, the focus was placed on the development and application of a novel method that would enable for the first time the locus-specific visualization of epigenetic modifications, in particular, DNA methylation, in living mammalian cells. This scientific strive was fueled by the fact that although the importance of epigenetic heterogeneity and dynamics is extensively acknowledged in the field of Epigenetics, no methods are currently available to enable the dynamic tracking of the status of target epigenetic modifications, with locus specific resolution (Bheda and Schneider, 2014). The aim of this project was to resolve this urgent and unmet technological demand by developing an array of modular epigenetic biosensors for live-cell microscopy applications. In its design and implementation, this technological platform was to be based on a combination between the recent advancements in the field of genomic targeting (Chen et al., 2016a) and the high epigenetic mark specificity of naturally-occurring protein

reading domains (Kungulovski et al., 2015b; Yamazaki et al., 2007). For microscopy-based visualization, it was planned to test and compare two types of fluorescent outputs, namely BiFC and FRET. The most robust system out of these was to be selected and incorporated into several sensors detecting different epigenetic marks and DNA sequences. The successful design and implementation of this novel technological platform should significantly contribute to a better understanding of how specific epigenetic signatures are set, erased and maintained during cellular development and pathogenesis.

3 Materials and methods

Most of the methodology used in the course of this work is described in detail in the Method sections of **Appendix 1-3**. Techniques with high relevance for the results described here are summarized in a project-specific manner in the paragraphs bellow. Particular emphasis is put on the description of unpublished methodology.

3.1 ATP hydrolysis regulates chromatin release of HELLS

3.1.1 Cloning of mammalian expression vectors, tissue culture, and transfections

The sequence encoding for the full-length mouse cells was sub-cloned from pGeneLSHF (Yan et al., 2003a) as an N-terminal fusion into pEYFP-C1 and pECFP-C1 as described in the Methods section of **Appendix 1**. The K237Q mutation in the HELLS ATPase domain was generated *via* site-directed mutagenesis using the megaprimer mutagenesis protocol described in (Jeltsch and Lanio, 2002). The identity of both constructs was confirmed by sequencing covering the whole open reading frame.

NIH3T3, HEK293 (American Type Culture Collection), iMEF and Suv39H1/H2 DKO (a kind gift of Dr. T. Jenuwein) were grown at 37°C in a humidified atmosphere, 5% CO2 using Dulbecco's modified Eagle's medium high glucose supplemented with 10% heat - inactivated calf serum. The iMEF-specific extra supplements are described in the Methods section of **Appendix 1**. Transient transfections with either Fugene HD (Promega) or jetPRIME (peqlab) were performed according to the manufacture's recommendations. Transfected cells were cultured for 24 – 72 h before being used for downstream analysis as described in the Methods section of **Appendix 1**. For mammalian protein purification, HEK293 cells were transfected with linear polyethyleneimine MW 250,000 (PEI) (Polysciences, Inc.) at a ratio DNA: PEI of 1: 3 (w/w). The reagent preparation and transfection procedure were implemented as described in (Longo et al., 2013).

3.1.2 Chromatin fractionation assay and western blot analysis

To assess the chromatin interaction strength of wild-type and ATPase-deficient HELLS, NIH3T3 cells were subjected 72 h after transfection to a biochemical extraction protocol adapted after (Yan et al., 2003a) and described in more detail in the Methods

section of **Appendix 1**. Briefly, the harvested cells were first re-suspended in cytoskeleton (CSK) buffer containing 0.5% Triton X-100 (vol/vol) to solubilize proteins that only loosely associate with chromatin. Following a 3 min centrifugation step at 5 000 x g, the resulting pellet was digested with RNase-free DNase I in CSK buffer with NaCl adjusted to 50 mM, to elute proteins that are tighter bound to chromatin. Finally, the nuclear matrix pellet remaining after a 3 min centrifugation step at 5 000 x g was solubilized in 8M urea, to release tight-binding chromosomal proteins. Equal volumes of protein extracts were then separated on a 15% SDS-PAGE by electrophoresis and transferred onto a nitrocellulose membrane. Following primary antibodies were used for detection: EYFP (Clontech, 632592), Lamin B (Santa Cruz, sc-6217) and HP1 (Active Motif, 39979). Detection of horseradish peroxidase-coupled antibodies was performed by enhanced chemiluminescence with Western Lightning Plus ECL reagent (PerkinElmer) according to the manufacturer's instructions. The resulting signal was visualized and analyzed as described in the Methods section of **Appendix 1**.

3.1.3 Preparation and imaging of fixed cells

For inspecting the cellular localization patterns of wild-type and ATPase-deficient HELLS, transfected NIH3T3 cells were fixed with 4% formaldehyde solution as described in the Methods section of **Appendix 1**. For 2% paraformaldehyde solution fixation, the original 4% stock was diluted in PBS shortly before fixation. Imaging was performed on a LSM 710 Zeiss confocal microscope, the imaging conditions are specified in the Methods section of **Appendix 1**.

3.1.4 Live cell imaging and fluorescence recovery after photobleaching

To examine the associate of wild-type and ATPase-deficient HELLS in the absence of formaldehyde cross-linking, live cell imaging was performed. The kinetic behavior of the proteins was analyzed by FRAP, as described in more detail in the Methods section of **Appendix 1**. Briefly, for FRAP experiments the cells were seeded on 35 mm Fluorodish cell culture dishes (World Precision Instruments). 24 to 48 hours after transfection, the growth media was replaced with imaging media containing no phenol red. Live cell imaging was performed on a LSM 710 Zeiss confocal microscope equipped with a Plan-Apochromat 63x/1.40 Oil DIC M27 objective and a XL-LSM 710 S1 incubation chamber for temperature and CO_2 control. For bleaching, a circular region with a diameter of 1 µm was chosen, positioned to cover one of the visible pericentromeric heterochromatic foci. To accommodate for differences in protein

mobility, for wild-type HELLS, images were recorded at an interval of 300 msec. For the HELLS K237Q variant, the image recording time was increased to 1 second, to minimize photobleaching and account for the slower protein exchange rate. The FRAP work-flow included 9 pre-bleach images, 5 bleaching cycles with 100% laser intensity, and acquisition of up to 100 post-bleach images. During acquisition, the 514 nm laser line was turned down to 4% to minimize photobleaching and phototoxicity. Image analysis was performed by extracting the mean pixel intensity of the bleached area, over time and normalizing against fluctuations in laser intensity by dividing against the average mean fluorescence intensity of the 9 pre-bleached images. The signal was scaled according to the fluorescence decrease observed in an unbleached area of the same cell. To derive the half recovery time, $t_{1/2}$, and the mobile fraction percentage, a double-component exponential model was used. Further details on data acquisition and analysis are provided in the Method section of **Appendix 1**.

3.1.5 Recombinant HELLS expression and purification

For 6xHis tag-mediated affinity chromatography, full-length HELLS was sub-cloned from the mammalian expression vector described above into the pET-28a (+) vector (Novagen), as a C-terminal 6xHis fusion. The BL21-CodonPlus(DE3)-RIL *Escherichia coli* protein expression strain was used for recombinant protein production. For this, the cells were grown in LB medium at 37 °C until an OD (600 nm) of 0.6 – 0.7. Protein expression was induced by addition of 1 mM of isopropyl-β-D-thiogalactoside (IPTG) followed by shifting the culture to 16 °C shaking at 160 rpm for 12 h. To support protein expression and folding, a final concentration of 5 mM $MgSO_4$ and 50 µM $ZnCl_2$ was present in the growth medium.

For purification, the harvested cells were resuspended in sonication buffer (20 mM Tris pH 6.8, 500 mM KCl, 20 mM imidazole, 0.1% β-mercaptoethanol, 10% glycerol) including protease inhibitor (Sigma) and lysed by sonication with 20 cycles, 15 s, 30% power (BANDELIN SONOPLUS). The lysate was next cleared by centrifugation at 18,000 rpm for 1 h 20 min and applied to a pre-equilibrated Ni-NTA column (Genaxxon bioscience). To remove impurities, the column was washed with 250 mL of wash buffer consisting of sonication buffer supplemented with 80 mM imidazole. Finally, the bound protein was eluted by increasing the final imidazole concentration to 300 mM. To remove the imidazole and decrease the salt amount, the eluted fraction with the highest concentration was dialyzed for 3 h against dialysis buffer (20 mM Tris pH 6.8, 200 mM

KCl, 0.1% β-mercaptoethanol, 10% glycerol). The purified protein was aliquoted and kept at -80 °C for long term storage. To assess the quality of the isolated protein, 10 µL recombinant HELLS were run on a 16% SDS-PAGE which was transferred to nitrocellulose membrane for detection of His-tagged proteins. For this purpose, an anti-his mouse monoclonal primary antibody (Roche Diagnostics, Cat. No. 11922416001), followed by an HRP-coupled anti-mouse secondary antibody (GE Healthcare, Car No. NA931V) were used.

3.1.6 Isolation of HELLS-EYFP from HEK293 cells via GFP-Trap®

72 h after transient transfection, HEK293 cells were harvested by scraping and resuspended in lysis buffer (20 mM HEPES, 500 mM KCl, 0.5% NP-40, 0.2 mM PMSF). To facilitate the release of nuclear proteins, the lysate was passed periodically through a 26G needle, during the 30 min lysis time. The total lysate was cleared by centrifugation at 20,000 rpm for 15 min, followed by 1:4 dilution with lysis buffer lacking NP-40 and incubation with 30 µL GFP-Trap® agarose slurry (chromotek). Following a 2 h binding time, the beads were washed with lysis buffer. The bound proteins were finally released by boiling for 10 min at 95°C in 50 µL 2x SDS-sample buffer. The quality of the eluted proteins was assessed by SDS-PAGE gel stained with colloidal Coomassie BB.

3.2 Allosteric control of DNMT3A by MeCP2

3.2.1 Cloning, bacterial expression, and recombinant protein purification

For *in vitro* pull-downs and methylation assays, the TRD domain of MeCP2 (aa 170-325) was cloned into pGEX-6P2 (GE Healthcare) as an N-terminal GST fusion. The ADD domain of DNMT3A (aa 472-610) was, in turn, cloned into pMAL-c2X (New England Biolabs) as an N-terminal MBP fusion. Site-directed mutagenesis was performed according to the protocol presented in (Jeltsch and Lanio, 2002) and as described in the Methods section of **Appendix 2**.

The expression and purification conditions of TRD MeCP2, ADD DNMT3A, and DNMT3A2 were described in detail in the Methods section of **Appendix 2**. In brief, the BL21-CodonPlus(DE3)-RIL expression *E.coli* strain was used for protein production. Protein expression was induced by addition of 1 mM IPTG when the culture reached an OD600 of 0.6 – 0.8. The expression was routinely performed in LB medium, at 20 °C. A final concentration of 5 mM MgSO4 and 50 µM ZnCl2 (for ADD and DNMT3A2)

was used in the growth medium to support cellular growth and the folding of Zn^{2+} -binding proteins. Cells were lysed by sonication and the pre-cleared supernatant was applied on a pre-equilibrated GST-sepharose 4B (GE Healthcare), Ni-NTA (Genaxxon Bioscience) or amylose column (New England Biolabs), depending on the affinity tag of the expressed protein (see the Methods section of **Appendix 2**). Bound proteins were eluted with sonication buffer supplemented with 10 mM reduced glutathione, 300 mM imidazole or 10 mM maltose, respectively. The quality of the purified proteins was routinely assessed by SDS-PAGE gel stained with colloidal Coomassie BB.

3.2.2 GST pull-down experiments

To map the interaction interface between DNMT3A and MeCP2 and assess the strength of the observed binding, GST pull-downs were performed with recombinantly expressed proteins, as described in the Methods section of **Appendix 2**. Routinely, 20 µL of Glutathione-Sepharose 4B beads, pre-equilibrated with interaction buffer (25 mM Tris pH 8.0, 100 mM KCl, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP-40, 200 µM PMSF) were used to immobilize 10-15 µg of the different GST-tagged proteins. Following washing, the putative His- or MBP-tagged interaction partner was added (10-15 µg) and the complex was allowed to form for 1 h at 4 °C with mixing end over end. Unspecific/week binding interactors were next removed by washing the beads with interaction buffer containing 300-600 mM KCl. Finally, the beads were re-suspended in SDS gel loading buffer and incubated for 10 min at 95 °C to release bound complexes. The resulting supernatant was loaded on a 12 % SDS-PAGE gel. Proteins were detected by Western blotting or Coomassie BB staining as indicated. Some experiments were conducted in the presence of recombinant hHistone H3.1 (M2503S, NEB), as detailed in the results section of **Appendix 2**.

3.2.3 Endogenous co-immunoprecipitation assay

To test if the interaction between DNMT3A and MeCP2 can also occur outside of the *in vitro* assays and cellular overexpression conditions, immunoprecipitation of endogenous DNMT3A was performed from mouse brain extracts (see detailed description in the Methods section of **Appendix 2**). In brief, whole brains from 16-week old C57Bl/N female mice were homogenized by grinding on liquid N2, followed by douncing in NP-40 lysis buffer with a tight pestle and sonication with EpiShear as described in the Methods section of **Appendix 2**. To facilitate the extraction of chromatin-associated proteins, the lysate was digested with Benzonase (Novagen) for

2h rotating at 4°C. Following pre-clearing with Protein A Sepharose CL-4B (GE Healthcare), 2.5 mg lysate were incubated overnight with either 15 µg anti-Dnmt3a antibody (sc-2070, Santa Cruz) or non-related rabbit IgG control (ab9106, Abcam). The antibody-bound complexes were retrieved by incubation with Protein A Sepharose beads, for 6 h, rotating at 4°C. After washing, the complexes were eluted by boiling in SDS gel loading buffer. The samples were next analyzed by western blot with an anti-MeCP2 monoclonal antibody (Cell Signaling), followed by incubation with an HRP-linked anti-rabbit IgG light chain specific antibody (Jackson ImuunoResearch).

3.2.4 *In vitro* DNA methylation activity assay

To assess the influence of the TRD domain on the DNMT3A2 activity, *in vitro* methylation assays with various biotinylated DNA substrates and labeled SAM were performed, following the protocol of (Roth and Jeltsch, 2000) and as described in detail in the Methods section of **Appendix 2**. Routinely, 1 µM 30mer oligonucleotide DNA and 0.76 µM [methyl-3H]-AdoMet (PerkinElmer Life Sciences) were added to the reaction mixture and the samples incubated at 37 °C. The concentration of DNMT3A2, TRD and DNMT3 CD used in the assays are described in the Methods section of **Appendix 2**. To follow the time-course of the methylation reaction, aliquots of 2 µL were taken from the reaction mixtures in duplicates at time points between 2 and 30 min and applied to one well of an avidin-coated microplate. The reaction was stopped by addition of excess of unlabeled SAM in buffer containing 500 mM NaCl. To release the bound DNA, the unspecific nuclease from *Serratia marcescens* was used. Finally, the released radioactivity was measured using liquid scintillation counting and the average count per minute of the duplicates was plotted against time. Linear regression was used to obtain the slopes of the initial linear parts of the time courses.

3.2.5 Generation of stable cell lines and global DNA methylation analysis of HCT116 DNMT1 hypomorph cells

To study the effect of MeCP2 on DNMT3A mediated methylation in cells, the HCT116 DNMT1 hypomorphic cell line was used (kindly provided by Prof. Bert Vogelstein, HHMI, USA), with reduced levels of DNA methylation (Rhee et al., 2002). As described in more detail in the Method section of **Appendix 3**, the cells were grown in McCoy's 5A medium (Gibco) supplemented with 10% heat-inactivated calf serum and 2 mM L-glutamine (Sigma), at 37 °C in a saturated humidity atmosphere containing 5% CO_2. For stable cell line generation, the cells were transiently transfected with either pEYFP-

MeCP2 or pEYFP, followed by drug selection with 0.5 mg/mL G418 (Sigma-Aldrich). Single clones were hand-picked and propagated in culture for circa 2 months before genomic DNA isolation. LC-ESI-MS/MS analysis of DNA methylation in genomic DNA was performed as described in (Bashtrykov et al., 2014; Schiesser et al., 2013).

3.2.6 Isolation of genomic DNA for whole genome bisulfite sequencing and data analysis

To analyze the effect of MeCP2 on DNA methylation in a biologically relevant setup, the methylome of the brain isolated from wild-type and MeCP2 KO mice was compared (for a detailed description see the Methods section of **Appendix 2**). To this end, the brain of one 6-8-week old wild type (C57BL/6J) and MeCP2 KO (B6.129P2(C)-Mecp2m1.1Bird/J) male mouse, respectively, was obtained from the Jackson Laboratory (Bar Harbor, ME USA). The tissue was mechanically disrupted by grinding in a mortar under liquid nitrogen, followed by Proteinase K digestion at 55°C for 4 hours. The genomic DNA was next extracted with the PureLink® Genomic DNA Mini Kit (Invitrogen™), according to the instructions of the manufacturer. The integrity of the isolated material was assessed by agarose gel electrophoresis.

WGBS library preparation and analysis was performed in cooperation with the group of Alexander Meissner (Harvard), following the procedures described in the Methods section of **Appendix 2**. The library was sequenced for 100-bp paired-end reads on an Illumina HiSeq 2500 sequencer, and the reads were aligned to the mm9 version of the mouse genome. To identify differentially methylated regions, the data was binned into 1kb tiles and the methylation was calculated as the coverage-weighted mean of methylation at CG sites within the tile.

3.2.7 *In vitro* CDKL5 kinase assay

To test if the TRD domain of MeCP2 can be phosphorylated by CDKL5, *in vitro* kinase reactions were set up, containing 15 µg wild-type or R306C GST-tagged TRD, 0.125 µg recombinant CDKL5 (abcam, ab159606) and 2.5 µCi of [γ-^{32}P]-ATP. The proteins were incubated for 1 h at 30°C in kinase assay buffer (25 mM MOPS, pH 7.2, 12.5 mM ß-glycerol-phosphate, 25 mM $MgCl_2$, 5 mM EGTA and 2 mM EDTA, 0.25 mM DTT) according to the instructions of the manufacturer. The reactions were stopped by boiling in SDS gel loading buffer, followed by separation on 12% SDS-PAGE. The phosphorylation signal was detected by autoradiography.

3.3 Locus-specific visualization of chromatin modifications in living mammalian cells

3.3.1 Cloning of BiAD anchor and detector modules for BiFC-based epigenetic sensors

Details on the constructs used for the BiFC-based readout of locus-specific epigenetic modifications are found in the Methods sections of **Appendix 3**. In brief, three anchor modules were used for specific binding of target genomic sequences: a three-fingered ZF array for the recognition of a 9-bp long sequence within mouse major satellite repeats (5'-GGC GAG GAA-3') (Lindhout et al., 2007), a TALE protein for the recognition of a 20-bp long sequence located within human alpha satellites (5'-TAG ACA GAA GCA TTC TCA GA-3') (Ma et al., 2013) and a CRISPR/dCas9 system coupled with a sgRNA recognizing the alpha satellite repeats in the pericentromeric region of chromosome 9 (5'-TGG AAT GGA ATG GAA TGG AA 3') (Ma et al., 2015).

To detect the 5mC mark, the methyl binding domain of human MBD1 (aa 1-113) was amplified out of HEK293 according to the borders described in (Yamagata, 2010). For the recognition of H3K9me3, the chromodomain of mouse HP1β (aa 1-76) was isolated out of NIH3T3 cDNA. Binding pocket mutants were generated by megaprimer mutagenesis according to the protocol described in (Jeltsch and Lanio, 2002).

To construct the BiFC-based biosensors, the anchor and detector modules were fused via Gibson assembly N- or C-terminally to the split Venus fragments (split site aa 210) as described in the Methods section of **Appendix 3**. To enable immunofluorescence and western blot detection a 3xFLAG tag was fused to all constructs. A 14 to 18aa-long GS-rich linker was used to separate the anchor/detector modules from the split fluorophore.

3.3.2 Cloning of BiAD anchor and detector modules for FRET-based epigenetic sensors

To detect H3K9me3 levels at mouse major satellite repeats, the HP1β chromodomain and the three-finger ZF described above were fused N-terminally to the mVenus and mCerulean FRET fluorophores, respectively (Addgene plasmids no. 27794 and no. 27796; (Koushik et al., 2006)). Cloning of the individual domains was performed using the Gibson assembly® mix (New England Biolabs®). A GS-rich 18-aa long flexible linker was used to separate the domains from the fluorophore. While the MBD domain contained an endogenous nuclear localization sequence, the SV40 NLS was used to

target the HP1β chromodomain to the nucleus. As a negative control for the FRET measurements, the SV40 NLS was fused to mVenus and used in the FRET assay instead of the ZF-Venus construct. As a positive control the Cerulean-5-Venus plasmid was used (Addgene plasmid no. #26394), where the two fluorophores are directly linked by a 5 aa-long linker (Koushik et al., 2006). All fusion constructs were expressed under the control of a CMV promoter.

3.3.3 Cloning of epigenetic modification enzymes

The cloning of the enzymes used for rescuing the DNA methylation and H3K9me3 levels is described in more detail in the Methods section of **Appendix 3**. Briefly, a CMV promoter was used to drive the expression of full-length wild-type of catalytically inactive C1226A human DNMT1. The enzyme was fused LSSmKate2 (Addgene plasmid no. 54795) for microscopy-based detection. For cloning the dox-inducible SUV39H1 expression, the CDS encoding for the full-length mouse enzyme was cloned into the pSIN-TRE3G-PGK-Puro-IRES-rtTA3 backbone (Fellmann et al., 2013). Both the wild type and the catalytically inactive H324L variant were cloned as mRuby2 fusions for microscopy-based detection (Addgene plasmid no. 54768).

3.3.4 Cell lines

The cell culture conditions used in this project are described in detail in the Methods section of **Appendix 3.** Cell lines used during the course of this work are: HEK293, NIH3T3 (American Type Culture Collection), wild-type and *Suv39h1h2*$^{-/-}$ iMEFs (gift of Prof. Thomas Jenuwein, MPI Freiburg), *p53*$^{-/-}$ and *p53*$^{-/-}$/*Dnmt1*$^{-/-}$ iMEFs (gift of Prof. Howard Cedar, Institute for Medical Research Israel-Canada), wild-type and HCT116 DNMT1 hypomorphic cells (kindly provided by Prof. Bert Vogelstein, HHMI, USA).

3.3.5 BiFC assay, data collection and analysis

The constructs, amounts of plasmids as well as transfection reagents used for the BiFC assay are described in detail in **Appendix 3**, Methods and Supplementary Tables 1-4. Routinely, a plasmid encoding for the mRuby2 red fluorophore was transfected in parallel to the BiAD modules to facilitate the identification of successfully transfected cells. pcDNA3.1 (invitrogen™) was used as a carrier plasmid in the cases where the plasmid amount used for the modules and mRuby2 did not sum up to the total amount recommended by the manufacturers of the transfection reagent. The complementation signal was imaged 24-48 h after transient transfection with the BiAD modules. For this,

the slides were fixed for 10 min at RT with 4% formaldehyde and finally mounted in ProLong™ Diamond antifade (invitrogen™).

The slides were imaged on a LSM 710 Zeiss confocal microscope equipped with a Plan-Apochromat 63 ×/1.40 Oil DIC M27 objective. The laser excitation wavelengths, as well as emission collection windows, are indicated in **Appendix 3**, Supplementary Table 5. For enhanced sensitivity, the BiFC channel signal was routinely directed to a QUASAR 34-channel photomultiplier unit (Carl Zeiss). Image analysis was performed in ImageJ 1.51a. The number of cells counted for each experimental set-up as well as the number of biological repeats and classification of the BiFC signals was documented in **Appendix 3**, Supplementary Table 6.

3.3.6 Detection of H3K9me3 levels at mouse major satellite sequences with FRET by acceptor photobleaching

For the FRET assay, NIH3T3 cells were seeded on high precision no.1.5 (170± 5 µm) glass slides (Carl Zeiss) and transfected with Fugene HD (Promega) according to the instructions of the manufacturer. Routinely 1.5 µg total DNA and a DNA: Fugene HD ratio of 1:4 were used. The DNA mixture consisted out of 750 ng plasmid encoding for the mVenus-tagged protein and 750 ng plasmid encoding for the mCerulean-tagged protein. For the fusion construct, mCerulean-5-mVenus 1500 ng plasmid were used. 24 h after transient transfection with the corresponding plasmids, the cells were washed with $PBS^{Ca2+Mg2+}$ and fixed for 10 min at room temperature with 2 % freshly-prepared formaldehyde solution. This was followed by a 30 min washing time with $PBS^{Ca2+Mg2+}$: This long washing time was important since it was previously shown that formaldehyde fixation quences mVenus and mCerulean differentially, and 20 min are needed for fluorescence recovery to reach steady state (Chen et al., 2006). The cells were mounted into freshly prepared Mowiol (Harlow and Lane, 2006) and kept at 8°C for long term storage. Before imaging, the slides were allowed to equilibrate at room temperature for at least 2 h to prevent movement artifacts caused by thermal fluctuations.

The acceptor photobleaching workflow was adapted after (Verveer et al., 2006). Here it was important that the mounting media does not contain fluorescence stabilizers such that the mVenus photobleaching kinetics is not influenced. The slides were imaged on a LSM 710 Zeiss confocal microscope equipped with a Plan-Apochromat 63 ×/1.40 Oil DIC M27 objective. The equipment was allowed to equilibrate for at least

30 min before performing the measurements such that the 514 nm laser line is stabilized. For mVenus bleaching, 8 iterations with the 514 nm laser power set at 80% were performed. For enhanced sensitivity, the mCerulean fluorescence (excited with the 405 nm diode) was routinely directed to the QUASAR 34-channel photomultiplier unit (Carl Zeiss). Each of the two fluorophores was imaged before and after mVenus bleaching. The 16-bit images were analyzed in ImageJ 1.51a where the fluorescence signal of the donor channel was visualized using the Fire lookup table (LUT) for improved detection of the fluorescence intensity in-between the imaging time points.

3.3.7 5-aza-dC treatment

To assess the dynamic range of BiAD sensor 2 to changes in DNA methylation levels, HEK293 were depleted of DNA methylation by 5-aza-dC (Sigma-Aldrich®, cat. no. A3656) treatment as described in the Methods section of **Appendix 3**. In brief, the treatment was performed over a period of 3 days at a final drug concentration of 2 µM in the cell culture medium. The drug was dissolved in 50% acetic acid and replaced on a daily basis. As a control, an equal volume of solvent was added to the cell culture medium. The efficiency of the demethylation treatment was assessed by digestion of genomic DNA isolated from the treated cells, with the 5mC-inhibited enzyme HpaII (New England Biolabs®). For the DNA methylation time course experiment, the cells were transfected fresh 2 days prior to the imaging time point (Methods section, **Appendix 3**).

3.3.8 Generation and assay implementation of the Tet-SUV39H1-mRuby2 doxycycline-inducible iMEFs

To generate a suitable cellular model system for assessing the dynamic range of BiAD sensor 4, doxycycline-inducible cells lines were created by retroviral infection of the parental Suv39DKO iMEFs, with viruses delivering transgenes encoding for either the SUV39H1 or the catalytically inactive H342L mutant, both fused to mRuby2. A detailed description of the infection and drug selection procedure is provided in the Methods section of **Appendix 3**. To test the applicability of the BiAD sensor 4, the expression of the SUV39H1 variants was induced with 1 µg mL^{-1} doxycycline, 48 h before the cells were transfected with the biosensor modules. The cells were cultured in the presence of dox for another 48 h prior fixation for imaging.

4 Results

4.1 An ATPase-deficient variant of the SNF2 family member HELLS shows altered dynamics at pericentromeric heterochromatin

The easy access of the DNA methylation machinery to the target DNA, as it is embedded within the context of tightly wrapped chromatin, is essential for both mammalian embryonic development and adult tissues homeostasis. While the maintenance methyltransferase DNMT1 predominantly methylates the newly replicated DNA, which is mostly nucleosome-depleted, the *de novo* DNMT enzymes must work on a chromatinized substrate (Felle et al., 2011). Several experimental approaches have highlighted the severe accessibility challenge posed to the DNA methylation machinery by stably positioned nucleosomes and densely compacted chromatin. Accordingly, high-resolution methylome profiling performed upon reintroduction of DNMT3 enzymes in *Dnmt* triple knockout embryonic stem cells, revealed a predominant accumulation of methylated CG sites in linker DNA, and their exclusion from nucleosomal-bound DNA sequences (Baubec et al., 2015). This kinetic data set is supported by several *in vitro* and *in vivo* studies aimed at elucidating the effect that nucleosome positioning has on DNA methylation (Felle et al., 2011; Kelly et al., 2012).

Among all genomic sites that carry the 5mC mark, the tightly compacted constitutive chromatin found at mammalian pericentromeres poses arguably the highest accessibility challenge for the *de novo* DNMT machinery (Huang et al., 2004). Pericentromeric regions are the embodiment of two antagonistic requirements. On the one hand, these sites need to be compacted in a dense heterochromatic structure, to prevent genomic instability caused by the spontaneous recombination of their highly repetitive DNA content (Soldi and Bonaldi, 2013). On the other hand, pericentromeric DNA needs to be transiently exposed to allow the spatial access to a whole artillery of chromatin complexes, responsible for maintaining pericentromeric repeats in a silent epigenetic state. The transition between these two antagonizing chromatin environments is under the control of chromatin remodeling enzymes, proteins that use the energy derived from ATP hydrolysis to modulate the interaction between the DNA and the core histone proteins, thereby facilitating the access of epigenetic factors to

the chromatinized template (Hota and Bruneau, 2016; Ryan and Owen-Hughes, 2011).

The putative chromatin remodeler HELLS is an important regulator of mammalian DNA methylation. The deep involvement of this protein in the DNA methylation pathway is highlighted by the 50-70% drop in global 5mC levels detected in *Hells*$^{-/-}$ cells. Importantly, repeat elements such as major satellites are among the sites that are most affected by the loss of HELLS, suggesting that this protein is important for facilitating the access of the DNA methylation machinery to sites of highly compacted chromatin (Dunican et al., 2013; De La Fuente et al., 2006; Yu et al., 2014). While this model recently received experimental support, as the *Arabidopsis* HELLS homolog DDM1 was found to be required for the methylation of linker H1-rich heterochromatic sites, a nucleosome remodeling activity has so far not been demonstrated for the mammalian protein HELLS (Zemach et al., 2013). Therefore, it is still unclear if and how a putative chromatin remodeling property of HELLS, together with its associated ATPase activity, contribute to the biological function of the protein. In this project, an ATPase-deficient HELLS variant was used to investigate the of ATP hydrolysis in regulating the interaction between HELLS and chromatin. The results of this work were summed up in a paper published in the Journal of Molecular Biology and attached here as **Appendix 1.** The main findings of this study are succinctly outlined in the following section.

Although baculovirus-expressed HELLS failed to remodel nucleosomal arrays *in vitro*, the protein contains the conserved Walker A and B motifs, which are part of the ATP binding pocket of SNF2 enzymes (Burrage et al., 2012; Flaus et al., 2006). Targeted exchange of a conserved K residue within the GKT box to A or Q, was previously reported to inactivate the ATPase function of different remodelers, and generate an HELLS variant that was no longer able to promote efficient phosphorylation of the histone variant H2AX under DNA damage conditions (Burrage et al., 2012; Richmond and Peterson, 1996; Rowbotham et al., 2011). However, the effect of this mutation on the association between HELLS and chromatin has not been investigated to date.

Wild-type HELLS was previously reported to localize to the nuclear compartment, where the protein was found to tightly associate with chromatin in actively proliferating cells. This interaction was dependent on the cell cycle and the confluence of the culture, with confluent G0-arrested cells having the least protein bound to chromatin

(Yan et al., 2003a). To assess whether the ATPase inactivating mutation influences the strength of chromatin binding by HELLS, I performed biochemical fractionation assays with G0-arrested NIH3T3 cells, 72 h after transient transfection with either the wild-type or the ATPase-deficient HELLS protein. For this, the lysed cells were extracted step-wise with three buffers that differentially solubilize proteins depending on their chromatin binding strength (**Figure 31**) (Yan et al., 2003a). In line with previous reports showing a loose association of HELLS with G0 chromatin, the wild-type protein was predominantly detected in the Triton X solubilized fraction (around 60% of the total protein amount). By contrast to this, the ATPase-deficient mutant displayed a tighter association with chromatin, with only ~ 30% of the protein being released upon Triton X extraction and > 60% being detected in the urea-denatured pellet (**Figure 31** and **Appendix 1, Figure 1b**). The reliability of the cellular fractionation protocol was confirmed by inspecting the extraction pattern of endogenous HP1β and Lamin B proteins, which served as markers for the DNase-released and the urea-solubilized fraction, respectively (**Figure 31**). As a reference for the Triton X-released fraction, a separate extraction on cells transfected with an EYFP-NLS fusion was performed. This also revealed the expected release of the chromatin-inert fluorophore in the first extraction step (**Appendix 1, Figure S1a**). Together, these results that were confirmed in 3 independent biological repeats (**Appendix 1, Figure 1b**), suggest that a functional ATPase domain is important for the release of HELLS from chromatin.

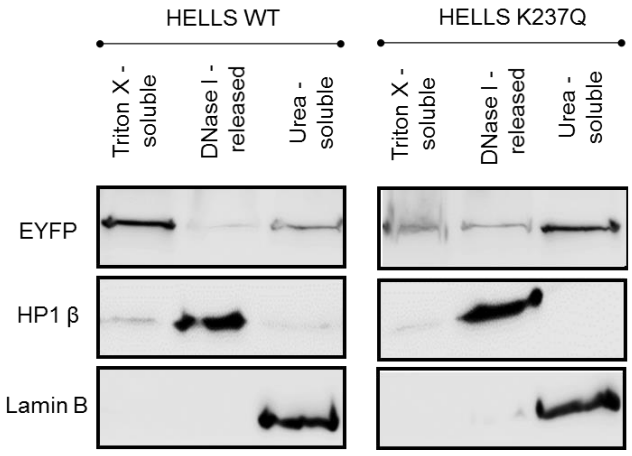


Figure 31| The wild-type and the ATPase-deficient HELLS display differential association with chromatin. Western blot detection of exogenous HELLS in fractions isolated from NIH3T3 cells, 72 h post transient transfection with either the EYFP-tagged wild-type HELLS (left-hand panel) or the EYFP-tagged ATPase deficient variant (right-hand panel). To assess the chromatin association strength of the two proteins, transfected cells were subjected to a stepwise extraction procedure, with buffers containing Triton X, DNase I and urea, respectively. Endogenous HP1β and Lamin B were used as references for the DNase I-released and urea-solubilized fraction, respectively. This figure was taken from Figure 1, **Appendix 1**.

Prompted by the outcome of the cellular fractionation assay, fluorescence microscopy was next employed to inspect whether these differences in chromatin association are also reflected in the cellular localization of the wild-type and the ATPase-deficient HELLS. To this end, mouse fibroblasts were transfected with the fluorophore-tagged proteins and imaged either in formaldehyde-fixed state (**Appendix 1, Figure 2**) or under live-cell conditions (**Appendix 1, Figure 3**). These experiments revealed that the mutant HELLS was predominantly associated with pericentromeric heterochromatin repeats, as detected by co-staining with the nuclear dies DAPI (for fixed cells) and DRAQ5 (for living cells). In contrast, the wild-type HELLS showed a rather fine-granular pattern and less frequent colocalization with mouse chromocenters (**Appendix 1**, **Figure 2a, b** and **d**; **Figure 3**; **Figure S3**). Interestingly, the proportion of cells displaying pericentromeric-localized wild-type HELLS varied depending on the imaging conditions. Accordingly, in formaldehyde-fixed cells, this localization was observed for only 4% of the counted population, whereas under live cell imaging conditions this increased to 59% (**Appendix 1**, **Figure 2d** and **Figure 3**). This variable localization was all the more surprising since pericentromeres were repeatedly shown to be the major site of HELLS activity (Huang et al., 2004; Yan et al., 2003a). By contrast to wild-type HELLS, the K237Q variant displayed a strong pericentromeric localization in > 80% of the transfected cells, regardless of wheatear the imaging was performed on fixed or living cells. (**Appendix 1**, **Figure 2d** and **Figure 3**). Importantly, in co-transfection experiments using wild-type ECFP-HELLS and EYFP-HELLS K237Q, the difference in localization between the two proteins was still maintained under chemical fixation, indicating that this is not an artifact of the imaging setup (**Appendix 1**, **Figure 2c**). Taken together, these results demonstrate that the localization of wild-type HELLS is sensitive to formaldehyde fixation and that the ATPase-deficient variant has an altered interaction with the tightly compacted pericentromeric chromatin.

Noteworthy, a similar formaldehyde dependent alteration of cellular localization patterns was previously reported for other nuclear proteins and was shown to be dependent on the nuclear mobility of these factors (Schmiedeberg et al., 2009). Accordingly, rapidly moving proteins may not be as efficiently fixed at their target sites by comparison to less mobile proteins. The fact that the localization of the K237Q variant was not drastically affected by the fixation process hints at the different mobility of wild-type and ATPase-deficient HELLS at pericentromeric heterochromatin. To test

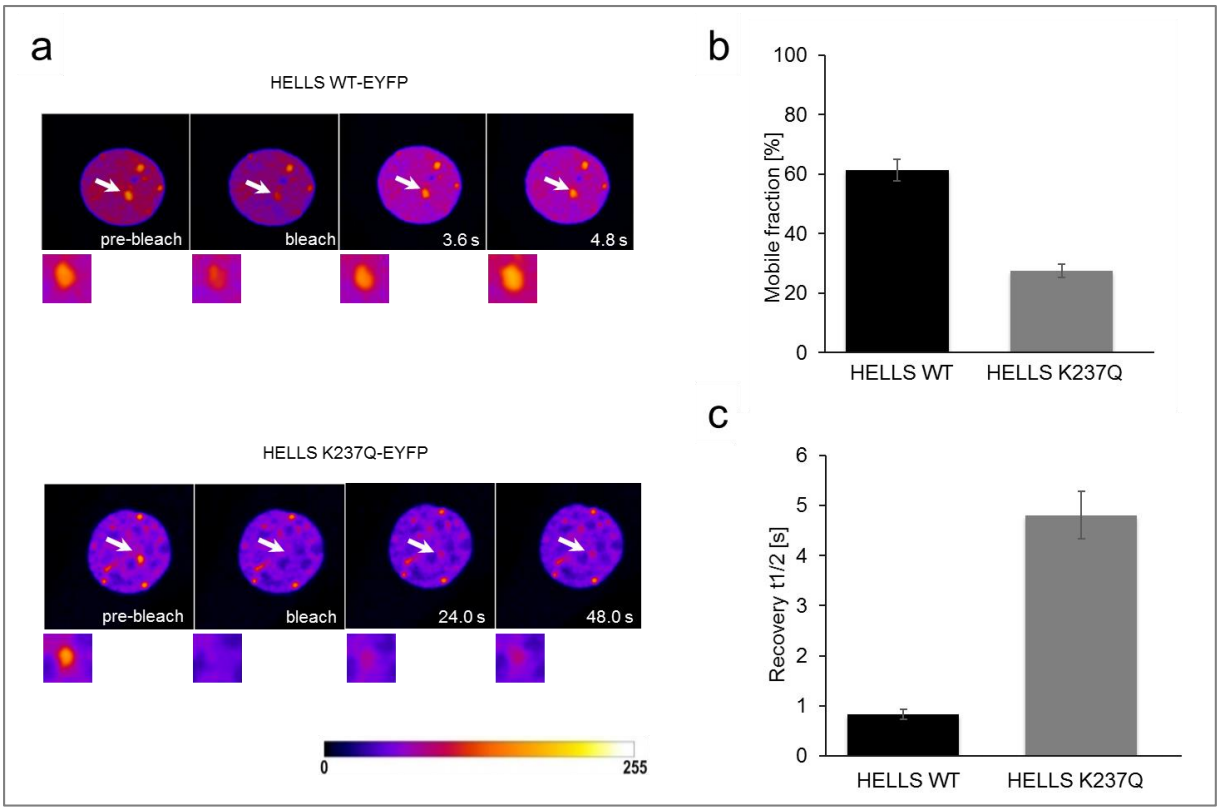


Figure 32| The wild-type and the ATPase-deficient HELLS show distinct dynamics at pericentromeric heterochromatin. a) Representative time frames of exemplary FRAP series (pseudo colored) for the wild-type (top panel) and mutant HELLS (bottom panel). Images were taken at the indicated time points after the bleach pulse. The bleached chromocenters were indicated with the white arrows. For clarity, a zoom-in of the bleached area was included in the lower part of each time-point. The panel was adapted after Figure 4, **Appendix 1**. **b)** Average % mobile fraction determined for the wild-type and ATPase-deficient HELLS. **c)** Average fluorescence half recovery time determined for the EYFP-tagged wild-type and ATPase-deficient HELLS. Error bars represent standard error of the mean based on measurements performed on 9 cells for wild-type HELLS and on 11 cells for the mutant protein. Panels b) and c) were taken from Figure 5, **Appendix 1**.

this hypothesis, FRAP experiments were next performed on cells transfected with either the wild-type or the ATPase-deficient HELLS. Since the differences in localization were the strongest at chromocenters, these sites were chosen to evaluate the mobility of the two protein variants. Strikingly, the two EYFP-tagged variants displayed a dramatically different bleaching behavior in the FRAP assay. Accordingly, while the K237Q HELLS was bleached with a high efficiency of 64%, the fluorescence signal was only reduced by about 34% for the wild-type protein (**Figure 32a** and **Appendix 1**, **Figure 2c**). In addition, the post-bleaching recovery of the fluorescence signal was distinct between the wild-type and the ATPase-deficient HELLS. While for the wild-type protein, the fluorescent signal recovered rapidly, in the case of the K237Q mutant, this occurred over a longer time scale. (**Figure 32a** and **Appendix 1**, **Figure 4a** and **b**). Quantitative analysis of the FRAP data according to a 2-component exponential fit model revealed a 2-fold increase in the mobile protein fraction and a

Table 1| Mobility parameters determined for the wild-type and the ATPase-deficient HELLS based on the FRAP measurements

mobile fraction [%]		recovery $t_{1/2}$ [s]	
HELLS WT	HELLS K237Q	HELLS WT	HELLS K237Q
61.3 ± 3.6	27.5 ± 2.1	0. 8 ± 0.1	4.8 ± 0.5

circa 5-fold shorter half recovery time ($t_{1/2}$) for the wild-type HELLS, by comparison to the ATPase-deficient mutant (**Figure 32b** and **c** and **Table 1**). Together, these results indicate that the inability of HELLS to hydrolyze ATP slows down the exchange of the protein at compacted pericentromeric sites.

Although it is known that pericentromeric repeats are the major sites of HELLS activity, the chromatin signals that trigger the recruitment of the protein to these sequences are not clear (Yan et al., 2003a). Since these sites are rich in H3K9me3, the localization of HELLS to chromocenters was next inspected in cells that do not express the SUV39H1/H2 proteins (Suv39DKO), enzymes responsible for setting the heterochromatic H3K9me2/3 mark (Maison et al., 2002). Interestingly, in these cells, the wild-type HELLS displayed a twofold increased enrichment at pericentromeres by comparison to its transfection in wild-type mouse fibroblast cells (89% *vs* 43%). An effect on the localization of the ATPase-deficient mutant could not be observed (**Appendix 1**, **Figure 6b** and **c**). These data indicate that while a defect in the H3K9me3 cascade does not influence the recruitment of HELLS to chromatin, this pathway might modulate the release of the protein from chromocenters.

Collectively, the data provided by this study highlight for the first time the importance of a functional ATPase module for the interaction between HELLS and highly compacted chromatin and suggests that an intact H3K9me3 pathway may be involved in the release of the protein from these heterochromatic sites.

To further dissect the functional difference between the wild-type and the K237Q HELLS variant, defined *in vitro* biochemical assays need to be employed. Since baculovirus-purified HELLS was reported to lack remodeling activity *in vitro* (Burrage et al., 2012), I took advantage of the EYFP-tagged HELLS constructs to purify both protein variants from mammalian HEK293 cells. By comparison to bacterial and insect systems, mammalian cells are preferable for the expression of proteins of mammalian origin, since they provide a full set of post-translational modifications that can support protein folding and activity (Wurm, 2004). That PTMs are important for the folding of recombinant HELLS is highlighted by the fact that unlike the bacterially-expressed

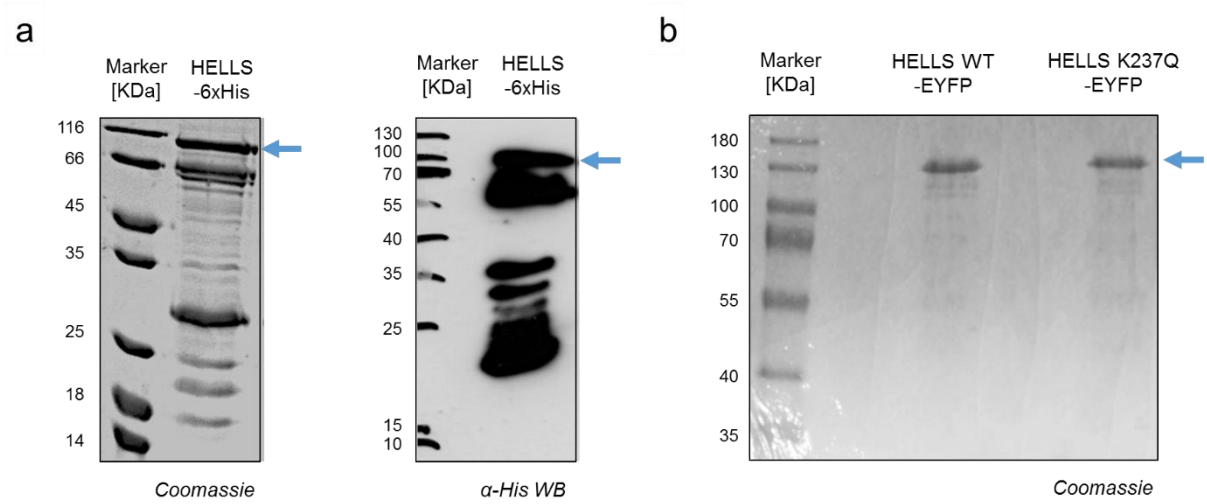


Figure 33| Expression of HELLS in mammalian cells results in a better quality protein preparation. a) Coomassie-stained SDS-PAGE gel (left panel) and α-His western blot detection (right panel) of 6xHis-labeled HELLS isolated out of the BL21-CodonPlus(DE3)-RIL *Escherichia coli* protein expression strain. The expected size of the full-length HELLS-6xHis protein was annotated with a blue arrow. As visible on the Coomassie-stained gel, recombinant HELLS could not be obtained with good quality. The western blot detection of 6xHis signals revealed that recombinant HELLS underwent severe degradation during the expression and/or purification process. **b**) Coomassie-stained SDS-PAGE documenting the good quality of the EYFP-tagged wild-type and K237Q HELLS proteins, after their expression and EYFP-mediated isolation out of HEK293 cells. The expected size of the full-length EYFP-HELLS was annotated with a blue arrow.

protein, which could only be obtained with a poor quality (**Figure 33a**), the mammalian expression system enabled me to purify the full-length HELLS to homogeneity and with little degradation (**Figure 33b**). Similar high-quality preparations could be obtained for both the wild-type and the ATPase-deficient variant (**Figure 33b**). These proteins can now be used to further dissect the importance of ATPase activity on the function of HELLS and assess whether a functional ATPase module affects the recruitment and activity of the DNMT machinery.

4.2 Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2

Accurate targeting and control of the mammalian DNA methylation machinery are essential for the setting and maintenance of the DNA methylation patterns (Jeltsch and Jurkowska, 2016). Recently, a seminal structural study dramatically contributed to our understanding of the chromatin-mediated control of DNMT3A activity (Guo et al., 2015c). The authors found that the N-terminally located ADD domain can bind to the catalytic domain of the DNMT3A enzyme at two distinct sites: an allosteric open, and an inhibitory closed site (**Figure 17a** and **b**, in the Introduction section). Depending on the interface, different amino acids were shown to be responsible for making contacts between the ADD and the CD. For instance, Y526 contacts the CD in the allosteric but not in the autoinhibitory conformation, whereas D531 contacts the CD in the autoinhibitory but not in the allosteric conformation. The binding of the ADD in the closed conformation was shown to block the access of the DNA to the active center and impair catalysis (**Figure 17b**). Interestingly, binding of unmodified histone H3 to the ADD domain triggered a large structural change, stabilizing the catalytically active conformation of the enzyme (**Figure 17c**). Based on these findings, the authors proposed a model in which the chromatin recruitment is temporally uncoupled from the activation of DNMT3 proteins, to ensure that the enzymes methylate the underlying DNA only when the neighboring nucleosomes carry appropriately modified histone tails (Guo et al., 2015c). Taking into account that the ADD domain was shown to function as a protein-protein interaction platform, it is conceivable that other proteins, apart from histone H3, may also affect the activity of DNMT3 enzymes through an allosteric mechanism (Jeltsch and Jurkowska, 2016). Experimental support for this hypothesis is nevertheless lacking to date.

The biological role of DNA methylation is to a large extent mediated by methyl binding proteins (Fournier et al., 2012). Among these, the chromosomal protein MeCP2 is one of the most prominent actors (Lewis et al., 1992; Meehan et al., 1992). Originally discovered due to its strong binding to symmetrically methylated CG, MeCP2 was also recently reported to recognize methylcytosine in a non-CG context (Gabel et al., 2015). Functionally, the protein is involved in many biological pathways where its influence on transcription was shown to be both negative and positive (Ragione et al., 2016). These

ambivalent effects might be in part mediated by the interaction of MeCP2 with proteins that depending on the chromatin environment can have either transcriptional repressive or activating roles (Ragione et al., 2016). Indeed, apart from the MBD domain, MeCP2 contains a TRD domain shown to mediate the interaction with other proteins and in particular with co-repressor complexes (Ballas et al., 2005; Fuks et al., 2003b). The critical biological role of this 5mC binding protein is highlighted by the expression levels of MeCP2 in the adult brain, where the protein is nearly as abundant as the histone octamers (Skene et al., 2010). In this tissue, MeCP2 was recently reported to be involved in controlling the methylation-induced repression of long neuronal genes (Gabel et al., 2015). This was attributed to the mCA reader function of the protein. Mutations in MeCP2 are frequently associated with the occurrence of the neurodevelopmental Rett syndrome (Smeets et al., 2012).

Interestingly, while both DNMT3 and MeCP2 are part of the same epigenetic pathway and are highly expressed in adult neurons, where they play important roles, the functional crosstalk between these two central epigenetic players has not been investigated to date (Guy et al., 2011). In this project, the possibility of a direct interaction between DNMT3A and MeCP2 was explored, and its mechanistic effects on DNMT3A activity were assessed. The resulting data were summarized in the form of a manuscript that has been submitted for review (attached here as **Appendix 2**). The main findings of this study are succinctly outlined in the following section.

Previous work from our group has uncovered that recombinantly purified MeCP2 and DNMT3A can directly interact, and that this association has a negative effect on the activity of DNMT3A activity as based on *in vitro* methylation assays using the radioactively-labeled SAM cofactor (**Appendix 2, Figure 2a** and **Figure 3a**). The interfaces responsible for the interaction were mapped through *in vitro* pull-downs with serially truncated proteins. These proposed the ADD domain of DNMT3A and the TRD domain of MeCP2 as the two main interacting points (**Appendix 2**, **Figure 2b-d**). I validated the self-sufficiency of these domains for the interaction by performing pull-downs with GST-tagged TRD and MBP-tagged ADD domains, where a direct binding between the two proteins could be detected (**Figure 34a**). The strength of their association was tested by performing the assay at 300 mM and 600 mM KCl. The detected binding was observed to be equally strong under both buffer conditions, highlighting the tight and specific association between the TRD and the ADD domain.

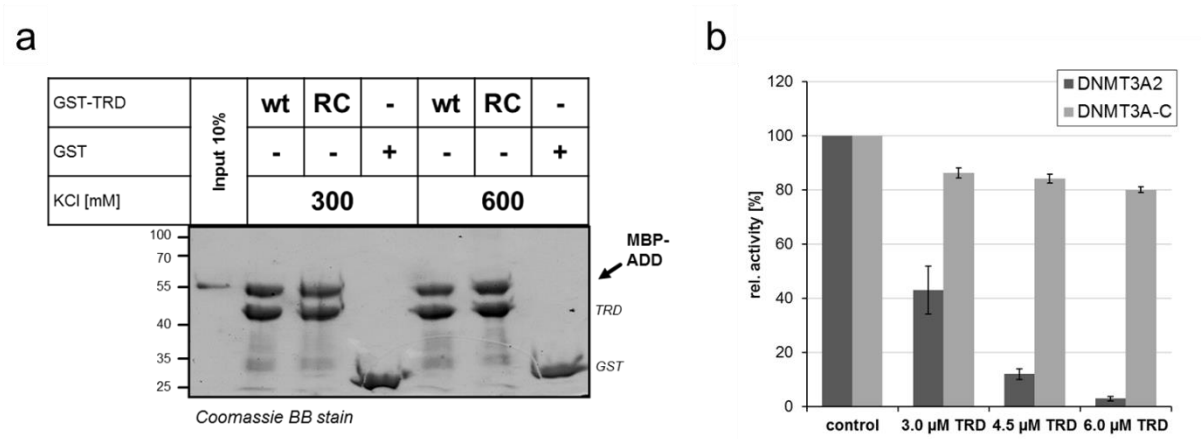


Figure 34| MeCP2 strongly interacts with DNMT3A *in vitro* has a concentration-dependent inhibitory effect on the activity of the enzyme. a) GST-pulldown with recombinant GST-TRD and MBP-ADD (the size of which was annotated with an arrow), performed under low and high salt conditions. The Rett TRD 306C mutant was included in parallel to the wild-type protein. The Coomassie-stained SDS-PAGE gel documents a strong and salt-resistant direct interaction between the TRD and ADD domains. GST was used as a negative control, instead of the TRD-GST domain fusion. This panel was taken from Figure 3, **Appendix 2**. **b)** Inhibition of DNMT3A activity by addition of increasing amounts of TRD using an unmethylated 30 bp-long substrate, with a central CG site. Control refers to a reaction without added TRD. All error bars indicate the standard error of the mean for minimum 2 biological replicates. This panel was taken from Figure 4, **Appendix 2**.

In addition to the wild-type TRD, the Rett R306C variant was also included in the assay. However, this protein was found to bind to the ADD domain equally well, indicating that this mutation does not perturb the interaction between the two domains. An equivalent pull-down was performed with the TRD domain and DNMT3A2, a naturally occurring shorter isoform of DNMT3A that has a truncated N-terminal part but still contains the ADD and PWWP domains (Chen et al., 2002). Also here, increasing the stringency of the pull-down by using 600 mM KCl did not weaken the interaction between DNMT3A2 and the TRD of MeCP2 indicating that the interaction between the two proteins is not mediated by electrostatic interactions (**Appendix 2**, **Figure 3c**). Finally, to confirm that the two proteins are also able to form a complex *in vivo*, co-immunoprecipitation of endogenous DNMT3A from mouse brain protein extracts was performed. Also here MeCP2 could be specifically detected in the pulled-down material, but not in the IgG control (**Appendix 2**, **Figure 2f**). Collectively, these results show that MeCP2 is a novel, strong and biological relevant interactor of DNMT3A.

Previous data from our lab have linked the association between MeCP2 and DNMT3A to a ~ 50% decrease in enzymatic activity (**Appendix 2**, **Figure 4a-c**). To inspect whether this effect depends on the concentration of TRD used in the assay, I performed *in vitro* methylation kinetics with increasing amounts of the TRD domain. Indeed, a strong inhibitory effect (> 95%) could be observed when using a 2.4 excess of TRD

(**Figure 34b**). By contrast, the non-interacting DNMT3A CD, which was used as negative control, displayed only a minor decrease in activity indicating that the inhibition of DNMT3A by TRD is not caused by competition for the DNA substrate or direct effects on the CD (**Figure 34b**). Importantly, kinetics measurements performed with a DNMT3B variant, created to mimic the borders of DNMT3A2, showed a comparable 50% inhibitory effect on the activity of the protein, in the presence of 3 µM TRD (**Appendix 2**, **Figure S9**). This suggests that MeCP2 affects the activity of DNMT3A and DNMT3B through the same mechanism.

Finally, to test whether this strong *in vitro* effects are also recapitulated in a cellular context, I derived cell lines stably expressing the YFP-tagged full-length MeCP2 under the control of a constitutive CMV promoter. Following clonal selection and a 2-month culture period, genomic DNA was isolated and subjected to LC-MS/MS to quantify the global 5mC levels. The mass spectrometry analysis was performed in collaboration with the group of Dr. Thomas Carell, at the Ludwig-Maximilans University in Munich. To better track the potential changes in DNA methylation upon introduction of MeCP2, the HCT116 DNMT1 hypomorphic cell line was used for the experiments. These cells contain an impaired DNMT1 activity and display a drop of about 20% in the global DNA methylation levels (Egger et al., 2006; Rhee et al., 2002). In line with the *in vitro* data, a roughly twofold decrease in the global methylation levels could be detected in the genomic DNA isolated from two independent clones expressing EYFP-MeCP2. This effect was MeCP2-specific as a control cell line expressing only EYFP did not show a significant change in the global DNA methylation levels (**Appendix 2, Figure 7a**). Together, these data indicate that MeCP2 functions as a strong inhibitor of DNMT3 activity both *in vitro* and in cells.

Having clearly demonstrated that MeCP2 and DNMT3A are strong and direct interactors and that their association leads to a concentration-dependent inhibition of DNMT3A activity, the mechanistic of this effect was next addressed. As mentioned above, the interaction between MeCP2 and DNMT3A was found to be mediated *via* the ADD domain of the enzyme. This was recently reported to allosterically modulate the activity of DNMT3A (Guo et al., 2015c). To elucidate whether the association with the TRD domain influences the DNMT3A activity through an allosteric mechanism, two conformationally locked variants of the enzyme were prepared by site-directed amino acid exchange in the ADD domain (**Figure 35a**). To this end, the D531 residue was

exchanged to R to generate a DNMT3A variant that preferentially adopts an autoinhibitory conformation. To obtain a protein variant where the allosterically open conformation form is preferred, the Y536E exchange was introduced. Importantly, these mutations were performed in the context of the full-length DNMT3A2 enzyme to enable the analysis of the crosstalk between the ADD and the CD domains in the presence of the MeCP2 TRD. *In vitro* kinetic assays with the DNMT3 variants revealed that the D531R mutant selectively lost the inhibition by MeCP2 (**Figure 35b**). This finding suggests that through its interaction, the TRD domain allosterically stabilizes the autoinhibitory conformation of DNMT3A, causing an inhibitory effect on the enzymatic activity. Since the unmodified histone H3 was previously found to act as an allosteric activator of DNMT3A, kinetics, and pull-down experiments were performed in the presence of recombinant H3 to test whether this can relive the TRD-mediated inhibition. Indeed, both assays revealed that upon addition of unmodified H3, the interaction between the TRD and ADD was weakened, and the inhibitory effect lost (**Figure 35c** and **Appendix 2**, **Figure 5c**). Together, these results indicate that the binding of H3 to the ADD domain can disrupt the TRD interaction and relieve the TRD-mediated inhibition of DNMT3A.

Finally, to study the influence of MeCP2 on DNA methylation in a biological relevant tissue, whole-genome bisulfite sequencing (WGBS) on genomic DNA isolated from the

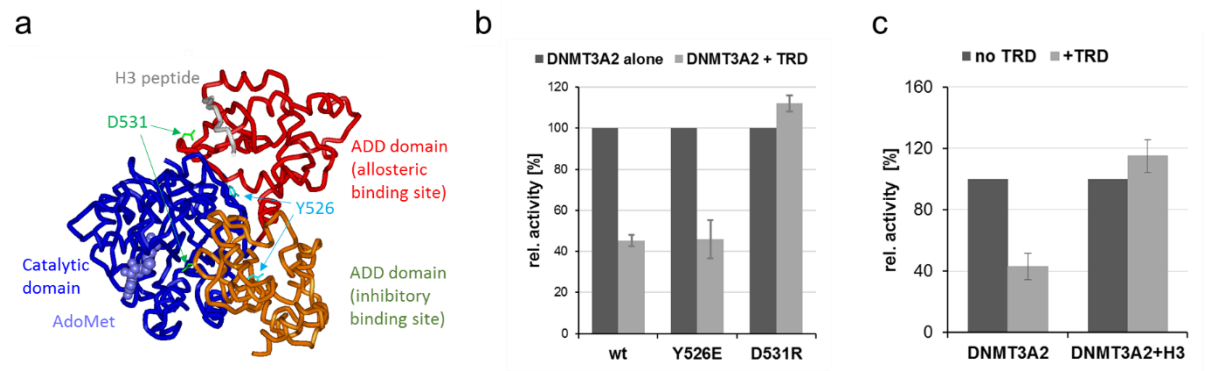


Figure 35| MeCP2 allosterically inhibits the activity of DNMT3A and this inhibition can be relieved by the unmodified histone H3. a) Model of the structure of a DNMT3A fragment comprising ADD (red or orange) and catalytic domain (blue). The ADD domain can bind to the catalytic domain at two interfaces, at an allosteric site (ADD domain colored in red, Y526 is at the interface), which is stabilized by binding of the H3 tail, and at an inhibitory binding site (ADD domain colored in orange, D531 is at the interface). The image was taken from Figure 1, **Appendix 2**. **b**) Inhibition of DNMT3A2 wild-type (wt) and its conformationally locked variants by TRD. The inhibition is lost in the D531R variant carrying a mutation, which destabilizes the autoinhibitory conformation. **c**) DNA methylation activities of DNMT3A2 and DNMT3A2 bound to H3 peptide (amino acid sequence 1-19, 25 µM) in the absence and presence of TRD (3 µM) showing the loss of TRD inhibition in the presence of H3 peptide. Error bars represent standard error of the mean for 2 biological replicates. Panels b) and c) were taken from Figure 5, **Appendix 2**.

whole brain of wild-type and *Mecp2* knockout (KO) mice was performed (Guy et al., 2001). Sequencing and bioinformatic analysis were performed in collaboration with the group of Dr. Alexander Meissner, at the Harvard Stem Cell Institute, USA. Unlike the overexpression experiments in HCT116 cells, no global effects on the DNA methylation levels could be observed in the brain of *Mecp2* KO animals. This is in line with published whole-brain gene expression analysis where only minor changes could be detected upon *Mecp2* knockout (Tudor et al., 2002). To investigate if the absence of MeCP2 is rather connected with locus-specific methylation effects, the data were binned in 1 Kb tiles to identify differentially methylated regions (DMRs). This analysis revealed that the site-specific effects of MeCP2 loss were ambivalent. Accordingly, 85 hypomethylated regions were found, with 24 showing more than 50% reduction in methylation levels; 84 hypermethylation DMRs were discovered, 23 displaying more than 50% increase in methylation (**Appendix 2, Figures 7b and c**). Importantly, 48% (41/85) of all hypomethylated DMRs and 63% (53/84) of all hypermethylated DMRs were found to be associated with MeCP2 binding sites (statistically expected were <10%) (**Appendix 2, Figure 7d**). This indicates that the observed changes in DNA methylation are indeed a direct effect of MeCP2 loss.

In summary, the data provided by this study propose MeCP2 as a novel and direct interactor of DNMT3. The discovered interaction was found to negatively affect the activity of the enzyme, through an allosteric mechanism. Importantly, the strong inhibitory effect could be relieved by addition of unmodified histone H3. Furthermore, WGBS on brain tissue isolated from *Mecp2* KO animals showed that MeCP2 has a punctuate but strong effect on DNA methylation levels. It is anticipated that the *in vivo* effect of this interaction might be more accentuated if specific brain areas such as the hypothalamus and cerebellum were analyzed, as here the lack of MeCP2 was reported to have stronger effects on gene expression (Ben-Shachar et al., 2009; Chahrour et al., 2008; Sugino et al., 2014).

In a recent study, Ebert and colleagues discovered that the interaction between MeCP2 and its interacting partner NCoR is modulated by activity-dependent phosphorylation of T308, within the MeCP2 TRD (Ebert et al., 2013). Importantly, the exchange of T308 to A in mice was shown to lead to the abnormal expression of activity-regulated genes and cause Rett-like symptoms. Molecularly, the alanine substitution was proposed to destroy a phosphorylation basophilic motif. Although the

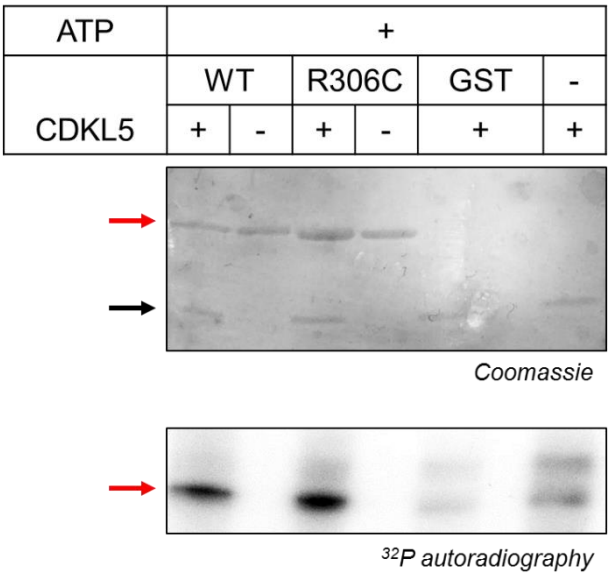


Figure 36| CDKL5 phosphorylates the TRD domain of MeCP2 *in vitro*. Coomassie-stained SDS-PAGE gel (top panel) and corresponding autoradiography (bottom panel) after *in vitro* phosphorylation of wild-type (WT), R306C TRD-GST and GST with $^{\gamma\text{-}32}$P ATP and CDKL5. To exclude that the phosphorylation signal is caused by kinases co-purifying with the GST-tagged proteins, matching reactions with ATP, but without CDKL5 were performed. The red arrow denotes the expected size of the TRD-GST variants, while the black one denotes the position of CDKL5. The gel was exposed for 20 min to film, before development.

kinase responsible for this modification was not identified, this paper proposed phosphorylation as a key chemical modification, through which the interaction between MeCP2 and other proteins can be regulated.

A kinase that is strongly associated with the MeCP2 pathway is the S/T cyclin-dependent kinase-like 5 (CDKL5). This is a nuclear protein whose expression in the nervous system was found to overlap with that of MeCP2, during neural maturation and synaptogenesis. Moreover, mutations in *CDKL5* were reported to give rise to a Rett-like phenotype (Scala et al., 2005; Tao et al., 2004; Weaving et al., 2004). CDKL5 was found to interact with and phosphorylate MeCP2 *in vitro,* although the phosphorylation site was not identified so far (Mari et al., 2005). By analogy to the T308 site it might, however, be that, if occurring on the TRD domain, CDKL5 phosphorylation may mediate the interaction between MeCP2 and DNMT3A.

To test if, the TRD domain of MeCP2 is modified by CDKL5, an *in vitro* phosphorylation assay with recombinant proteins and ^{32}P-labeled ATP was performed. Taking into account that the R306C exchange is found in an S-rich ESSI**R**SVQE sequence motif, and thereby may act similarly as the T308A exchange, the wild type TRD domain as well the Rett R306C variant were included in the assay. As documented in **Figure 36**, a strong and specific phosphorylation activity of CDKL5 towards the TRD domain could

be detected. The equally strong *in vitro* modification of the TRD R306C variant indicated nevertheless that this amino acid exchange did not negatively influence the CDKL5 target motif. The position of the modified residue as well as its potential effect on the association between the TRD and other interactors, such as DNMT3A, awaits further investigations.

4.3 Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites

The sequence-specific investigation of chromatin marks is mandatory for a locus-resolved understanding of epigenetic signaling cascades. To underpin the functional outcome of epigenetic modifications, as they are incorporated in highly dynamic and heterogeneous networks, methods that enable the live cell detection of epigenetic marks are essential. This technological development needs to satisfy two main requirements. First, epigenetic modifications have to be specifically detected at the level of endogenous, unmodified genomic sequences. This is important for a comprehensive understanding of the function of epigenetic marks, as they are set, maintained and erased in the native chromatin environment and not when incorporated in artificially introduced reporter sequence or on genetically engineered genomic loci (Hathaway et al., 2012; Stelzer et al., 2015). Second, owing to their high dynamics and large cell-to-cell heterogeneity, epigenetic modifications need to be assessed at the level of single living cells (Bheda and Schneider, 2014; Levsky et al., 2002; Meissner et al., 2008; Shalek et al., 2013; Stelzer et al., 2015). This is important for the accurate deconvolution of the dynamics and heritability of defined epigenetic states (Bheda and Schneider, 2014; Farlik et al., 2015). Despite their large impact on the development in the field of Epigenetics, methods that simultaneously satisfy these two important criteria are so far not available (Bheda and Schneider, 2014). In this project, this urgent and unmet technological need was addressed by developing an array of modular fluorescence complementation-based epigenetic sensors for live-cell microscopy applications. The results of this work were summarized in a manuscript submitted for review and attached here as **Appendix 3**. The main findings of this study are succinctly outlined in the following section.

To achieve locus-specific readout of target epigenetic modifications in living cells, a set of modular bimolecular fluorescence (BiFC)-based sensors was designed. These consisted of pairwise combinations between an anchor module, for DNA sequence-

specific recognition, and a detector module, specifically binding to a defined chromatin modification. For simultaneous access of DNA sequence and epigenetic information, both of these devices were fused with complementary non-fluorescent N- and C-terminal fragments of the monomeric Venus fluorophore (VenN and VenC, respectively) (Morell et al., 2009; Ohashi and Mizuno, 2014). The working hypothesis was that if the epigenetic modification of interest is present at or next to the target genomic locus, binding of the anchor and detector modules in close spatial proximity would lead to the reconstitution of a functional Venus fluorophore that would emit a stable, microscopically-trackable fluorescent signal (**Figure 37a**). This fluorescence complementation approach has several advantages over traditional co-localization microscopy and was already successfully employed by other groups to visualize various types of intermolecular interactions in living cells and even whole animals (Hudry et al., 2011; Kerppola, 2006, 2008; Kodama and Hu, 2012). Accordingly, while co-localization microscopy is limited by diffraction to a resolution of 200 nm, fluorescence complementation can only occur if both parts approach each other at intermolecular distances of 7 nm or less, dramatically improving the genomic resolution with which the epigenetic mark can be detected (Yu et al., 2013). To assess the chromatin mark-dependency of the novel BiAD (**Bi**molecular **A**cceptor **D**onor) sensors, binding pocket mutations in the detector module were used (**Figure 37a**), as well as

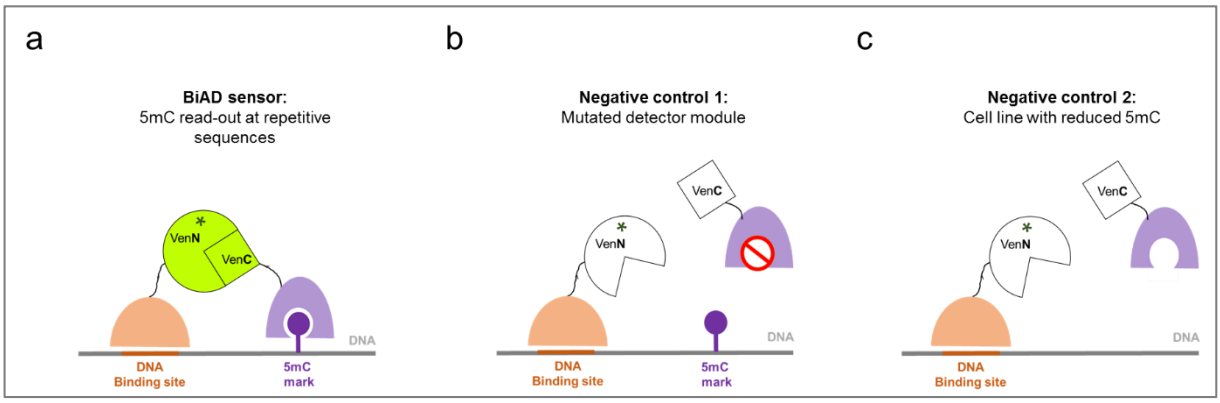


Figure 37| Concept of the BiAD sensors for site-specific detection of DNA methylation. a) Schematic diagram depicting the concept of the BiAD sensors designed for locus-specific readout of DNA methylation. Simultaneous readout of DNA sequence (orange line) and epigenetic mark (purple lollypop) is achieved by transfecting mammalian cells with an engineered anchor module (shown in orange) for sequence specific DNA binding, and a detector module (shown in purple) specifically recognizing methylated CG sites. Both of these devices are fused to complementary N- and C- terminal parts of Venus. If transfected in a cell line where the DNA sequence of interest is methylated, the docking of the anchor and detector modules in close spatial proximity will lead to the generation of a stable fluorescent signal. By contrast, inactivating the methyl binding pocket of the 5mC reader (**b**) or transfecting the intact modules in a cell line with globally reduced 5mC levels (**c**) is expected to give rise to lower fluorescence signals, as the detector module cannot longer stably dock onto chromatin. All panels of this figure were taken from Figure 1, **Appendix 3**.

cell lines with globally reduced levels of the investigated epigenetic mark (**Figure 37b**). In both settings, a strong reduction in the formation of the BiFC signal was expected, owing to the impaired docking of the detector module onto chromatin.

To target endogenous genomic loci, different types of engineered DNA-binding domains, a viz. Zinc-finger, a TAL effector and a CRISPR-dCas9 system, were selected due to their high sequence specificity (Chen et al., 2016a; Kungulovski and Jeltsch, 2016). As detector domains, I took advantage of naturally occurring chromatin reading domains that bind specifically to defined epigenetic modifications (Kungulovski et al., 2015b; Musselman et al., 2012). Accordingly, the MBD of MBD1 was employed for the detection of 5mC, while the chromodomain of HP1β (HP1 CD) was used for the recognition of H3K9me3 (Bannister et al., 2001; Hendrich and Bird, 1998; Jacobs et al., 2002; Li et al., 2015a; Ohki et al., 2001).

To establish the complementation system, the BiAD sensor 1 was designed to readout 5mC found at major satellites of mouse pericentromeric heterochromatin. These sites were used as a proof-of-concept chromatin environment since they harbor large amounts of DNA methylation and form highly abundant tandem repeat arrays, that can be easily microscopically visualized by DAPI-staining (**Appendix 3, Supplementary Fig. 1a**). For the detection of major satellites, a zinc finger protein (ZF) designed to specifically bind the 5'-GGCGAGGAA-3' motif within these sequences was used (Lindhout et al., 2007). As documented in **Figure 38a**, co-transfection of mouse fibroblasts with the VenN-fused-anchor together with the VenC-fused MBD detector gave rise to strong fluorescence signals that were specifically formed at satellites repeat sequences as concluded based on the co-localization of the BiFC signal with DAPI foci (**Appendix 3**, **Fig. 2b,** upper panel). The sensor displayed an excellent signal reconstitution with around 90% of the counted transfected cells showing productive BiFC fluorescence (**Appendix 3**, **Fig. 2d**). To test the 5mC specificity of the BiAD sensor, mouse fibroblasts were triple transfected with the BiAD modules and an MBD-Cerulean construct. This experiment revealed that the BiFC signal was only formed at sites that were also bound by MBD-Cerulean, highlighting the 5mC specificity of the sensor (**Appendix 3**, **Fig. 2b,** lower panel). To further validate this result, the methyl binding pocket of the MBD detector module was inactivated by site-exchanging of the conserved R44 residue to Q (Ohki et al., 2001). Implementation of this protein variant in the BiFC assay resulted in a dramatic reduction in the intensity of the reconstituted

fluorescence signal (**Figure 38b**, **Appendix 3**, **Fig. 2c, d** and **Supplementary Figs. 3b, 6**). As a second negative control, the sensor was transfected in *Dnmt1*$^{-/-}$ mouse fibroblasts, cells that do not express the maintenance methyltransferase DNMT1, and consequentially display a severe reduction of global DNA methylation levels (Casas-Delucchi et al., 2012). In line with the previous 5mC specificity data, the fluorescence complementation signal obtained with the BiAD sensor was strongly reduced also in this case (**Figure 38c**, **Appendix 3**, **Fig. 2c, d** and **Supplementary Figs. 3b, 8, 9**). Collectively, these results demonstrate the feasibility of the BiAD concept and indicate that the BiAD sensor 1 can be used to directly monitor the status of 5mC at repetitive DNA sequences in murine cells.

Motivated by this outcome, the functionality of the system was next challenged by approaching a chromatin environment in which the anchor and detector modules do not co-localize optically. To this end, BiAD sensor 2 was designed for the specific detection of DNA methylation at alpha-satellite sequences found around human centromeres. Taking advantage of the modularity of the system, the already validated MBD module used for DNA methylation readout could be re-integrated into the human sensor. As anchor module, a TALE protein designed to specifically recognize the pan-centromere target sequence 5'-TAGACAGAAGCATTCTCAGA-3' was employed (Ma et al., 2013).

Taking into account that human chromocenters are not as spatially defined as the murine ones, several variants of the detector and anchor modules where generated, in which the position of the VenN and VenC fragments was swapped around the core

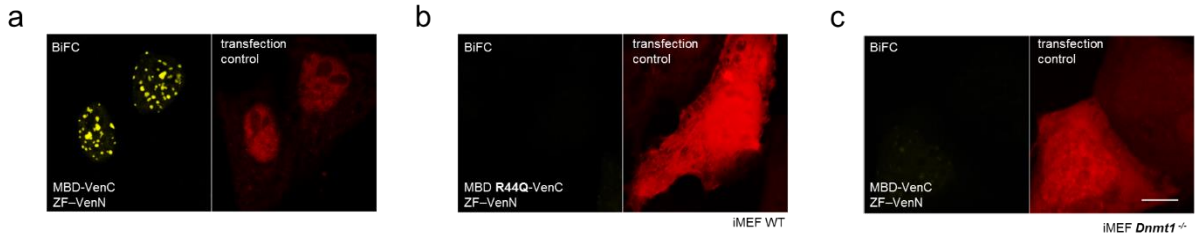


Figure 38| Development of the BiAD sensor 1 for specific detection of DNA methylation at mouse major satellite repeats. a) Representative fluorescence microscopy image of NIH3T3 cells captured 48 h after transfection with the modules of the BiAD sensor 1. The bimolecular fluorescence complementation (BiFC) signal is shown in yellow. **b**) Representative fluorescence microscopy images documenting the 5mC-validation of the BiFC signals. The BiFC signal was lost with the MBD R44Q 5mC binding pocket mutant. **c**) Representative fluorescence microscopy images documenting the 5mC-validation of the BiFC signals. The BiFC signal was lost in cells with globally reduced DNA methylation levels. A plasmid encoding for the red fluorophore mRuby2 was used as a transfection control. All panels of this figure were taken from Figure 2, **Appendix 3**.

modules. In addition, for increased protein flexibility, anchor variants with a longer linker separating the split fluorophore from the TALE protein were designed (**Appendix 3**, **Supplementary Fig. 12**). The anchor-detector pair displaying the highest signal-to-background ratio was selected for downstream specificity analysis, which demonstrated that alike BiAD sensor 1, this system displayed a high specificity for both the DNA sequence and the 5mC mark (**Appendix 3**, **Fig. 3**). Remarkably, despite the less compacted state of the pericentromeric chromatin in human cells in comparison to mouse cells, the optimized BiAD sensor 2 resulted in strong Venus reconstitution in circa 75% of the transfected cells (**Appendix 3**, **Fig. 3d**).

Encouraged by the good performance of the BiAD sensor 2, the dynamic range of this tool was next explored. For this, two approaches were undertaken.

In the first strategy, the sensor was transfected into the HCT116 DNMT1 hypomorphic cells with reduced global methylation, to test whether this change in DNA methylation is reflected in the BiFC yield of the sensor (Egger et al., 2006; Rhee et al., 2002). Indeed, transfection of the BiAD modules in this cellular background resulted in a 4-fold reduction in the number of cells showing a strong and spotty BiFC signal by comparison to the wild-type cells (**Appendix 3**, **Fig. 4a** and **c**). Furthermore, the rescue of DNA methylation by re-introduction of an active DNMT1 protein in the hypomorphic cell line was successfully mirrored by a three-fold increase in the number of cells showing a strong spotty BiFC signal (**Appendix 3**, **Fig. 4c**). Importantly, no signal increase was observed when a catalytically inactive DNMT1 variant was used instead (**Appendix 3**, **Fig. 4b** and **c**). Together, these results provide strong evidence that the BiAD sensor can be used to detect changes in DNA methylation at alpha satellites in live human cells.

In a second approach, the BiAD sensor 2 was used to monitor the changes occurring at alpha satellite sequences upon treatment of HEK293 cells with the strong DNA methylation inhibitor 5-aza-2'-deoxycytidine (5-aza-dC) (Fahy et al., 2012). In agreement with published literature and methylation-sensitive restriction digestion of genomic DNA isolated from the 5-aza-dC-treated cells, a strong decrease in the BiFC signal at 3 days after 5-aza-dC treatment and also at 2 days after drug removal could be detected (**Appendix 3**, **Fig. 4d, e**). Importantly, the effects of the drug were reversible, the signal of the sensor displaying a strong increase at 7 days post drug removal. After 2 weeks of recovery, a further increase in the number of cells with strong

BiFC signals was noticed (**Appendix 3**, **Fig. 4d, e**). A comparable re-methylation kinetics was previously observed for Alu-repeats (Velicescu et al., 2002). Collectively, these data show that this BiAD sensor can be used to study locus-specific changes caused by epigenetic drugs in living cells.

Motivated by the excellent performance of the BiAD approach, I aimed to increase the locus specific resolution of the detection system by incorporating the highly flexible nuclease-deficient Sp-dCas9 protein and its cognate sgRNA, to specifically target a pericentromeric sequence located only on human chromosome 9. As for the other tools, the newly generated BiAD sensor 3 also demonstrated an excellent 5mC and DNA-sequence specificity, as replacing the wild-type MBD for the MBD R44Q variant, as well as absence of the sgRNA, lead to a dramatic reduction in the number of cells displaying strong BiFC signals (**Appendix 3**, **Fig. 5c, d** and **Supplementary Fig. 19**). To evaluate the dynamic range of the novel BiAD sensor 3, 5-aza-dC was used to induce global DNA demethylation in the HEK293 cells. After 3 days of drug treatment, a ten-fold decrease in the number of strong BiFC-positive cells was observed (**Appendix 3, Fig. 5e, f**). All in all, these data clearly demonstrate that dCas9 can function as an anchor module and be incorporated into novel, 5mC-specific BiAD sensors.

Having successfully developed several biosensors for the readout of DNA methylation at different target regions, I next aimed to develop a BiAD system to monitor the status of histone tail marks at endogenous genomic loci in living cells. As such, the BiAD sensor 4 was constructed, to specifically recognize the H3K9me3 marks that abundantly decorate mouse major satellite repeat sequences (Saksouk et al., 2015). To this end, the ZF module employed in BiAD sensor 1, was used as an anchor for the recognition of major satellite sequences, while the chromodomain of HP1β (HP1 CD) was employed as a specific H3K9me3 binder (**Figure 39a**). To test the chromatin mark specificity of the readout Suv39DKO cells were used, where H3K9me3 levels are globally reduced. Here, the stable docking onto chromatin of the HP1 CD should be impaired and the formation of the BiAD signal was expected to less efficient (**Figure 39b**).

It was expected that the readout of histone modifications imposes greater challenges on productive BiFC conformations due to the larger distance between the DNA binding site of the anchor module and the binding site of the detector module on the highly

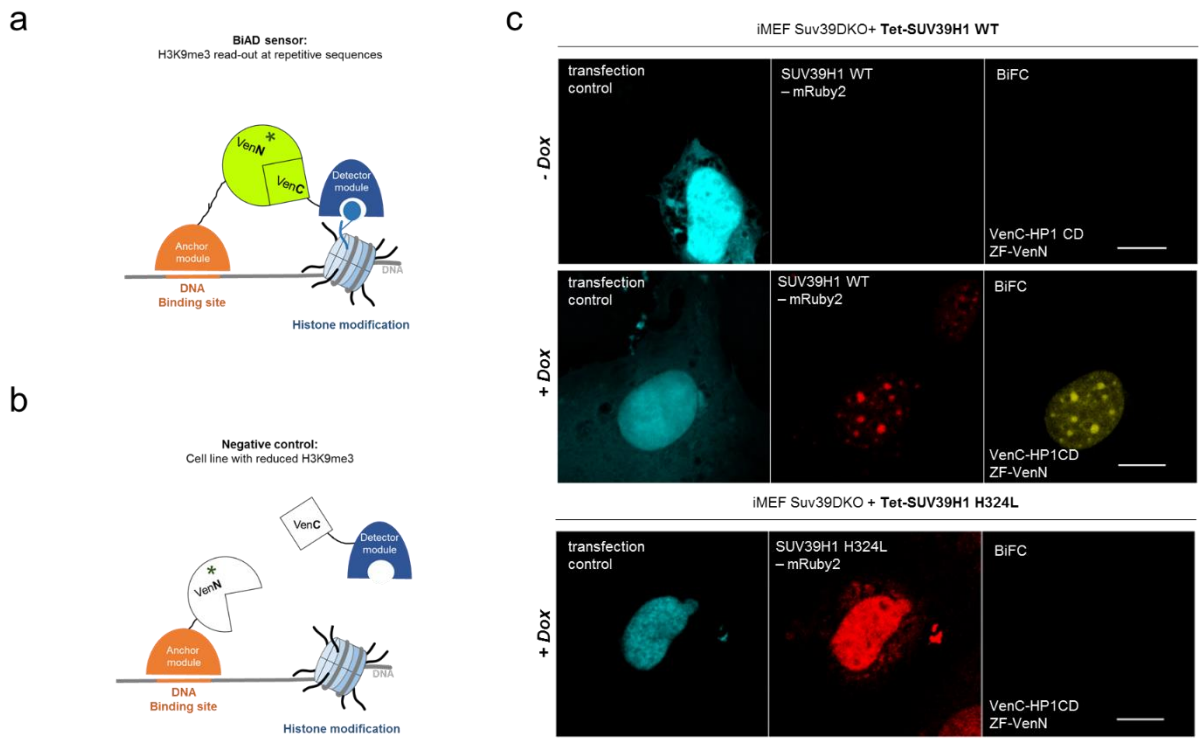


Figure 39| Concept of the BiAD sensors for site-specific detection of H3K9me3 and development and application of BiAD sensor 4 to visualize dynamic changes in H3K9me3 levels at major satellite sequences. a) Schematic diagram depicting the concept of the BiAD sensor. Simultaneous readout of DNA sequence (orange line) and epigenetic mark (blue lollypop) information is achieved by transfecting mammalian cells with an engineered anchor module (shown in orange) for sequence specific DNA binding, and a detector module (shown in blue) specifically recognizing H3K9me3 marks. Both of these devices are fused to complementary N- and C- terminal parts of Venus. If transfected in a cell line where the nucleosome neighboring the DNA sequence of interest carry H3K9me3 marks, the docking of the anchor and detector modules in close spatial proximity will lead to the generation of a stable fluorescent signal. By contrast, transfecting the intact modules in a cell line with globally reduced H3K9me3 levels **(b)** is expected to give rise to lower fluorescence signals, as the detector module cannot longer stably dock onto chromatin. **c)** Application of the BiAD sensor to detect changes in H3K9me3 levels after dox induced expression of SUV39H1 in Suv39DKO iMEF cells. A BiFC signal could only be observed when the sensor was transfected in cells expressing SUV39H1 (middle panel). No BiFC signal could be detected in cells expressing the catalytically inactive H32L SUV39H1 mutant (bottom panel), or in the absence of dox (top panel).

flexible histone tails. For this reason, several versions of the two modules were generated, where the position of the split fluorophores was shuffled around the core modules (**Appendix 3**, **Supplementary Fig. 22**). The pair displaying the highest signal-to-background ratio was further selected for downstream specificity analysis using H3K9me3 antibody co-staining (**Appendix 3**, **Figure 6a**) and transfection of the modules in Suv39DKO iMEF cells (**Appendix 3**, **Supplementary Fig. 21a, b**). These experiments highlighted the H3K9me3 specificity of the novel sensor. The DNA sequence specificity was assessed by DAPI co-staining of BiAD-transfected cells, where the expected co-localization between the BiFC signal and mouse chromocenters was observed (**Appendix 3**, **Fig. 6a**). Remarkably, despite the higher

mobility of histone tails, approx. 85% of the transfected cells displayed a strong BiFC signal (**Appendix 3, Fig. 6c**). This was comparable to what was obtained for the 5mC readout at these sites and documents the general applicability of BiAD sensors.

Having been able to monitor the loss of H3K9me3 levels at major satellite repeats, I next tested whether the sensor can also be used to read out an experimentally induced gain in H3K9me3 at these sites. For this, stable Suv39DKO iMEF cells were generated, that express doxycycline (dox) inducible SUV39H1. The catalytically inactive (H324L) enzyme variant was included as a negative control. Remarkably, transfection of the BiAD modules in the wild-type SUV39H1-expressing cells 4 days after dox induction, gave rise to strong BiFC signals in around 65% of the counted transfected cells (**Figure 39c**, **Appendix 3, Fig. 6b, c**). This corresponded to a 10 fold increase in the number of cells with strong BiFC fluorescence, by comparison to the signal obtained in the parent Suv39DKO iMEF cells. Importantly, this was directly dependent on the induction and activity of the SUV39H1 enzyme, with the induction of the SUV39H1 H324L variant not leading to an increase in fluorescence reconstitution (**Figure 39c** and **Appendix 3, Fig. 6c**). These data clearly demonstrate that the novel BiAD sensor can be used to specifically readout increasing H3K9me3 levels at major satellite repeats in live cells.

In conclusion, in this study, a technological platform is presented allowing for the first time the direct assessment of the epigenetic status of defined genomic loci in living cells. The specificity and portability of the developed tools were assessed with several negative controls in several mammalian cell lines. Furthermore, with these devices, dynamic changes that occur in epigenetic marks upon drug treatment or induction of epigenetic enzymes, could be tracked for the first time in live cells and with locus specific resolution.

While BiFC-based sensors are robust, highly sensitive and compatible with live cell imaging applications, this approach suffers from one inherent limitation: the production of the fluorescence signal reflects the binding of the anchor and detector modules at the target site, but not their dissociation from the genomic locus (Buntru et al., 2016). Furthermore, some reports argued that the complementation of the fluorophore is irreversible, although this is still a matter of debate, and most probably depends on the type of proteins to which the split fluorophore is attached (Cole et al., 2007; Demidov et al., 2006; Guo et al., 2005; Morell et al., 2009). An alternative strategy to visualize dynamic protein-protein interactions relies of FRET. This is a distance dependent

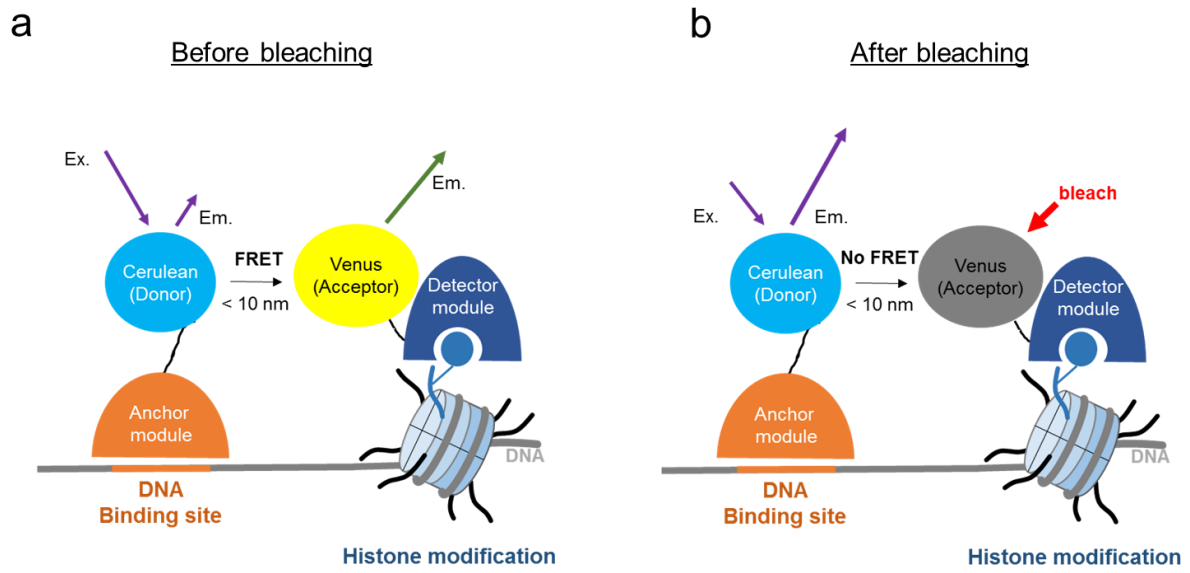


Figure 40| Concept of FRET-based BiAD approach for the detection of H3K9me3 marks at mouse major satellite sequences. The anchor module (ZF) was fused to the mCerulean, FRET donor fluorophore (in cyan), while the detector module (HP1CD) was fused to the mVenus, FRET acceptor fluorophore (in yellow). If the nucleosome found next to the DNA sequence of interest is decorated with H3K9me3 marks, the anchor and donor modules will dock in close spatial proximity, bringing the attached donor and acceptor fluorophores in a FRET-favorable distance. Accordingly, due to the transfer of energy from donor to the acceptor molecule, the fluorescence signal emitted by the donor (purple arrow) is decreased (**a**). By contrast, if the chromophore of the Venus acceptor is destroyed by bleaching, the signal emitted by the FRET donor increases (**b**).

process in which, through dipole–dipole interactions, an excited fluorophore molecule (the donor) transfers energy non-radiatively to another fluorophore molecule (the acceptor), resulting in acceptor emission (Buntru et al., 2016; Lakowicz, 2006). Since FRET relies only on the spatial proximity of the donor and acceptor fluorophores, the measured signal reflects both the association and dissociation of the interacting partners at the genomic target. For this reason, FRET-based approaches are more frequently used than BiFC techniques to study dynamic protein-protein interaction events (Piston and Kremers, 2007). Nevertheless, because FRET signals are usually small, this readout requires sophisticated measurement techniques and is prone to data analysis artifacts (Hirata and Kiyokawa, 2015; Jares-Erijman and Jovin, 2003; Lindenburg and Merkx, 2014).

To test whether a FRET-based readout is compatible with the BiAD approach, I next focused on the detection of H3K9me3 marks that abundantly decorate the major satellite repeats of mouse chromocenters (Saksouk et al., 2015). Owing to the high sensitivity observed with the BiFC-based readout, the anchor and detector modules used to construct the BiAD sensor 4, were re-purposed for a FRET-based measurement. To this end, the ZF anchor, recognizing major satellite sequences, was

fused to the mCerulean FRET donor fluorophore. The HP1 CD, binding specifically to H3K9me3 marks, was fused to the mVenus FRET acceptor fluorophore. To measure FRET, the acceptor photobleaching method was selected (**Figure 40**). This relies on evaluating the intensity of the signal emitted by the donor chromophore before (**Figure 40a**) and after (**Figure 40b**) the acceptor is disrupted by illumination with a high - intensity laser line (König et al., 2006). The working concept is that if the two chromophores are in proximity at steady state, the acceptor photobleaching would lead to an increased donor fluorescence, as the chromophore accepting the energy transferred by the donor is destroyed. By contrast, if the donor and acceptor do not meet the geometrical constraints required for FRET to occur, no change in the donor fluorescence is expected to occur between the two measurement points (Karpova et al., 2003; Vogel et al., 2006). While acceptor photobleaching is less suitable for live cell measurements, owing to phototoxicity effects associated with high-intensity illumination, this measurement method is among the most robust used to readout FRET. In addition, it requires fewer controls and is not as prone to technical and analysis errors as other FRET measurement methods such as the direct measurement of acceptor emission and fluorescence lifetime imaging (Busnelli et al., 2013; Van Munster et al., 2005).

To establish this method, a construct encoding the Cerulean and Venus fluorophores fused together through a 5 amino acid long linker (C5V) was used (Koushik et al., 2006). This fusion construct was shown to have a high average FRET efficiency of circa 30%, serving as a FRET reference standard (Koushik et al., 2006). As documented in **Figure 41a**, upon bleaching the acceptor fluorophore, a visible increase in the donor fluorescence was obtained. By contrast, performing the same experiment on cells transfected with HP1CD-Cerulean and NLS-Venus, where no interaction is expected, did not lead to a measurable increase in the mCerulean fluorescence (**Figure 41b**). Taken together, these results indicate that the setup of the FRET assay is functional as it allows a good discrimination between true interactors and proteins that meet by chance due to overexpression. In spite of this, applying these settings to measure the FRET signal produced upon co-transfection of the HP1CD-Cerulean and ZF-Venus modules, revealed no visible differences in-between the two measurement time points (**Figure 41c**). Here, based on the results obtained with BiAD sensor 4 (**Figure 39c**) as well as published literature, a strong increase in the fluorescence of the chromocenter-localized donor was expected (Saksouk et al., 2015). Collectively,

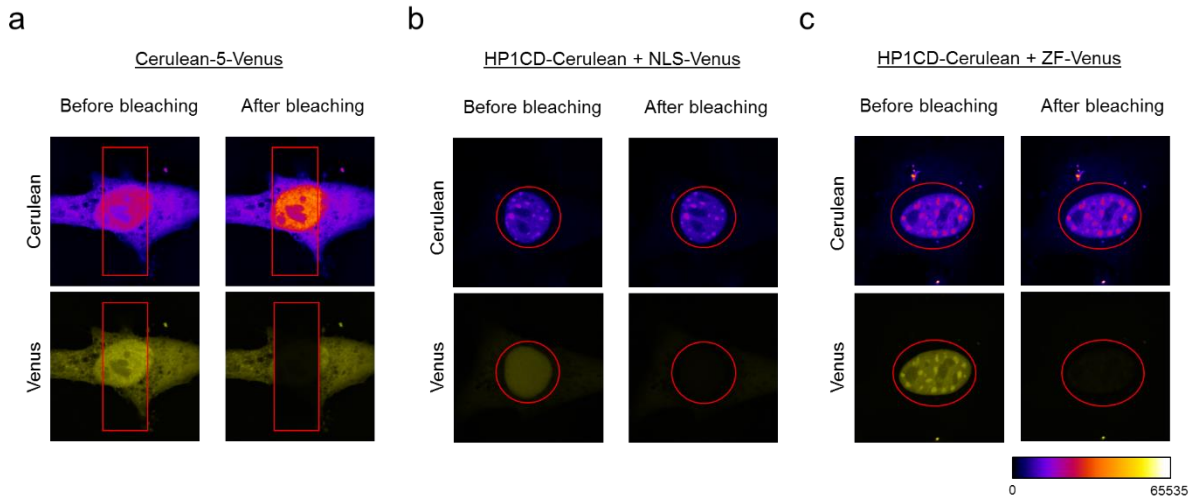


Figure 41| Establishment of the acceptor photobleaching FRET measurement technique for the readout of H3K9me3 marks at mouse major satellite repeats. a) Representative FRET measurement documenting an increase in the fluorescence of the Cerulean donor (pseudocolored), upon Venus bleaching. The Cerulean-5-Venus FRET standard was used for this measurement. **b)** Representative FRET measurement documenting no increase in the fluorescence of the HP1CD-Cerulean signal upon bleaching the NLS-fused acceptor fluorophore. **c)** Representative readout of the FRET effect produced upon co-transfection of mouse cells with constructs encoding for HP1CD-Cerulean and ZF-Venus. No visible change in the signal intensity of the Cerulean donor could be observed, in-between the two imaging time points. In all panels, the area selected for bleaching was marked with red. The fluorescence signal of the donor was pseudocolored to enhance visibility. The 16 bit LUT scale is included in the lower right corner of the figure.

these data highlight the problematics of FRET-based readouts and underline the clear superiority of BiFC approaches.

To further improve the applicability of the BiAD sensors, mVenus could be next replaced with the engineered monomeric *Deinococcus radiodurans* infrared fluorescent protein IFP1.4. Indeed, this was shown to be capable of reversible fluorescence complementation and was already successfully used to detect spatiotemporal dynamics of protein-protein interactions (Tchekanda et al., 2014). Due to its spectral properties, this protein is also suitable for whole-body imaging, opening up new avenues for the applicability of the sensors (Tchekanda et al., 2014).

5 Discussion

In my work I was aiming to shed light onto three major open questions in the field of Molecular Epigenetics, namely, how chromatin remodelers facilitate the access of epigenetic factors to chromatin, how epigenetic enzymes are controlled and regulated to establish and maintain defined epigenetic states and how the epigenetic state of a live cell can be dynamically detected. The main outcomes that resulted from each of these areas of investigation, together with the new directions of research that were opened by these findings, are discussed below in a project-specific manner.

5.1 HELLS, a mysterious chromatin remodeler

5.1.1 ATP hydrolysis regulates the release of HELLS from chromatin

The role of HELLS as an important regulator of constitutive heterochromatin was discovered in genetic studies in mice as early as in 2001 (Dennis et al., 2001), shortly after the discovery of DNMT3 enzymes (Okano et al., 1999) and histone lysine methyltransferases (Rea et al., 2000). Follow-up investigations revealed that the protein is enriched at major satellite sequences, embedded in pericentromeric heterochromatin, as well as minor satellites, found at centromeres (Huang et al., 2004; Muegge, 2005). Depletion of *Hells* in mice was documented to lead to transcriptional reactivation of satellite sequences, culminating in extensive DNA damage and incomplete chromosomal synapsis (Huang et al., 2004; De La Fuente et al., 2006). This effect was largely attributed to the involvement of HELLS in the DNA methylation pathway. Indeed, knock-out of *Hells* in murine cells was reported to cause a global reduction in the DNA methylation levels, the strongest affected sites being repeat sequences, such as retroviral-like sequences (Dunican et al., 2013; Yu et al., 2014). In addition, recently HELLS was described to control cytosine methylation in a nuclear compartment that partially overlaps with lamin B1 attachment regions (Yu et al., 2014). Lamin B1 attachment domains (LADs) were identified based on the physical interaction of DNA sequences with the nuclear membrane protein, lamin B1, and shown to be important for controlling the organization of the genome and for providing a repressed chromatin environment (Guelen et al., 2008; Kim et al., 2011; Zheng et al., 2015). Interestingly, LADs were found to be rich in repeat elements with > 60% of repeats being located here, and for some types of satellite sequences and endogenous retroviral elements the frequency surpassing 80% (Yu et al., 2014).

Owing to its high sequence homology to SNF2 family members, which are ATP-driven chromatin remodelers that can translocate along the DNA and disrupt protein-DNA contacts, HELLS was proposed to be a chromatin remodeler and have critical roles in modulating the structure of compacted chromatin (Briones and Muegge, 2012; Meehan et al., 2001; Ryan and Owen-Hughes, 2011). Its plant homolog DDM1 was one of the first proteins in *Arabidopsis*, for which the knockout led to massive and global loss of DNA methylation, at a time when the DNA methylation field was still in its infancy (Vongs et al., 1993). HELLS is thought to be required for opening up the chromatin fiber at heterochromatic sites to provide developmental-controlled access to the DNA methylation machinery (Briones and Muegge, 2012; Pikaard, 2013; Zemach et al., 2013). These findings were supported by the discovery that HELLS and DNMT3B are direct interactors (Myant and Stancheva, 2008). Nevertheless, while the *A. thaliana* HELLS homologues, DDM1, was found to be able to remodel chromatin *in vitro* (Brzeski and Jerzmanowski, 2003; Stroud et al., 2013) and mediate the access of DNMTs to linker H1-containg chromatin *in vivo* (Pikaard, 2013; Zemach et al., 2013), an ATP hydrolysis-dependent activity as well as a chromatin remodeling function has not been convincingly shown for HELLS to date. In this study, I took advantage of an ATPase-deficient HELLS variant K237Q described by (Burrage et al., 2012) to address the contribution of ATP hydrolysis to the interaction between HELLS and chromatin.

Investigating the nuclear localization of wild-type and ATPase-deficient HELLS with fluorescent microscopy, revealed that both proteins accumulated at DAPI-dense pericentromeric heterochromatic foci, in line with the extensively-documented role of HELLS at these repetitive sites (Huang et al., 2004; De La Fuente et al., 2006; Muegge, 2005; Yu et al., 2014; Zhu et al., 2006). Surprisingly, a strong discrepancy in the localization of wild-type HELLS was observed between live and formaldehyde-fixed cells, with the chemical treatment leading to a more than 10 fold reduction in the number of cells with pericentromeric-localized HELLS. Notably, similar differences in protein localization between live and fixed cells have been previously documented by Schmiedeberg and colleagues for the chromosomal protein MeCP2 (Schmiedeberg et al., 2009). Based on FRAP measurements and comparison of the localization patterns between live and fixed cells for several MeCP2 truncation variants, the authors found that proteins which reside on chromatin for less than 5 s cannot be efficiently captured by the formaldehyde cross-link chemistry. A formaldehyde-mediated eviction of transcription factors from the chromatin of mitotic cells has also been recently observed

in both plant and mammalian cells (Li et al., 2015b; Teves et al., 2016). Based on the localization of the transcription factor Sox2, Teves and colleagues proposed a mechanistic model to explain this formaldehyde-induced fixation artifact. Accordingly, as formaldehyde molecules cross the cell membranes, they rapidly cross-link the pool of free TFs, creating a steep fixation gradient. This creates a sink effect, where crosslinking of free molecules prevents their binding to DNA, while the DNA-bound TFs can still dissociate since protein-DNA cross-links take longer to form than protein-protein cross-links. Noteworthy, this fixation artifact was shown to be dose-dependent, with high formaldehyde concentrations causing a steeper fixation gradient and a stronger depletion of the transcription factor from chromatin (Teves et al., 2016). Importantly, a similar concentration-dependent effect could be observed also in the localization of wild-type HELLS. Accordingly, using 2% formaldehyde instead of 4%, resulted in an increased number of cells where HELLS displayed a spotty localization (**Appendix 1**, **Figure S2a**). Taken together, these results raise concerns about the use of formaldehyde as the chemical of choice for the cross-linking highly mobile nuclear proteins, and suggest that HELLS is dynamically interacting with heterochromatin. This hypothesis was confirmed by biochemical chromatin fractionation experiments with transfected cells, as well as FRAP measurements at chromocenters. These revealed that wild-type HELLS can be easily stripped off chromatin with detergent and that the protein rapidly exchanges at chromocenters between the chromatin-bound and unbound state, with a short half recovery time of only 0.8 s (**Table 1**, Results section). While quantitative measurements of the mobility parameters for chromatin remodeling proteins are still sparse in literature, this value is in the range of what has been previously reported for other chromatin remodelers such as the human ISWI family members, SNF2H/SNF2L and Acf1 (Erdel et al., 2010).

Interestingly, by contrast to the wild-type protein, the ATPase-deficient HELLS variant did not display strong localization differences between fixed and live cells. The enrichment of the protein at DAPI-dense pericentromeres indicates that an intact ATPase activity is not required for the recruitment of HELLS to chromatin. Stimulated by the different sensitivity to formaldehyde fixation observed for the mutant HELLS, FRAP measurements were performed. These revealed that in contrast to the highly mobile wild-type HELLS, the ATPase-deficient mutant displayed an increased chromatin residence time and a slower exchange rate at pericentromeric sites (**Table 1**). This result was in line with biochemical fractionation experiments documenting an

enrichment of the HELLS K237Q variant in the urea-solubilized chromatin fraction. Taken together, these experiments indicate that a functional ATPase domain is important for the release of HELLS from tightly compacted chromatin sites mediating the dynamic interaction of HELLS with chromatin.

While the involvement of HELLS at pericentromeric heterochromatin has been extensively documented, the specific signal that leads to the recruitment of the protein to these sites has not been identified so far. Chromatin structure has been previously shown to be involved in the recruitment of HELLS to pericentromeres. Accordingly, disruption of higher-order chromatin organization by treatment with histone deacetylase inhibitors was reported to cause the dissociation of HELLS from chromatin (Yan et al., 2003a). Today we know that HDAC inhibitors do not only alter the compaction of chromatin but are also involved in the de-acetylation of non-histone proteins and were reported to disturb the interaction of HDAC enzymes with co-repressor complexes (Ropero and Esteller, 2007). Therefore, the effect of HDAC inhibition on the localization of HELLS may be more complex. Since pericentromeric heterochromatin is known to be highly enriched in the H3K9me3 repressive modification, the involvement of this histone mark in the recruitment of HELLS to chromatin was tested (Aagaard et al., 1999; Lehnertz et al., 2003). Comparing the localization of the protein between wild-type and *Suv39H1/H2* double-KO MEFs revealed an increase in the pericentromeric localization of HELLS in Suv39DKO cells. By contrast, the ATPase-deficient HELLS displayed a strong association with chromocenters independent of the presence of the H3K9me3 modification. Importantly, these results are supported by a recent mass spectrometry-based study, documenting an enrichment of HELLS at the pericentromeric heterochromatin of Suv39DKO mouse embryonic stem cells (Saksouk et al., 2014). The differences observed in the localization of wild-type HELLS indicate that while the H3K9me3 mark is not important for the recruitment of HELLS to chromatin, an intact H3K9me3 pathway is needed for the efficient release of the protein from pericentromeric sites. Further investigations will be needed to elucidate whether the H3K9me3 mark affects the chromatin release of HELLS directly, or indirectly via other members of the pathway.

Collectively, the data presented in this study document for the first time the importance of an intact ATPase activity for the interaction between HELLS and chromatin, and support the chromatin remodeling function proposed for this protein. Interestingly, the

ATPase activity was found to be crucial for the release of HELLS from tightly compacted pericentromeric sites, but not for the recruitment of the protein to chromatin. Furthermore, experiments in Suv39DKO cells revealed that the release of HELLS from chromocenters is a multifactorial process, where in addition to the active ATPase module, an intact H3K9me3 pathway also plays a role.

5.1.2 A model for the recruitment and release of HELLS from chromatin

The results obtained in this project are in agreement with a continuous sampling mechanism proposed to explain how some chromatin remodelers are able to rapidly identify their target sites among the 30 million of nucleosomes found in the mammalian nucleus (Erdel et al., 2010; Längst and Manelyte, 2015). According to this, remodelers would freely diffuse throughout the nucleus with most of their binding events being transient and unproductive (Längst and Manelyte, 2015). During this time, the proteins would continuously sample the nuclear space, screening for high-affinity nucleosomes (‘good’ substrates) that can be converted, through remodeling, into low-affinity binding sites (‘bad’ substrates). This conversion would trigger the release of the remodelers from chromatin (**Figure 42**). Inactivating the ATPase activity of these proteins would lead to a preferential enrichment of the remodelers at their target nucleosomes, since in the absence of remodeling activity these sites would remain high affinity ‘good’ substrates (**Figure 42**). Such a mechanism was previously observed for the Isw2 chromatin remodeler and is in good agreement with the FRAP data obtained here for HELLS (Erdel et al., 2010; Gelbart et al., 2005; Rippe et al., 2007).

In addition to their short residence times and rapid diffusion, the efficient target searching mechanism of some remodelers was also shown to be supported by a transient upregulation of protein expression levels. Also for HELLS an increase in protein amount was detected at the onset of S phase (Geiman and Muegge, 2000). Interestingly, a fist wave of RNAPII-dependent transcription of major satellites was observed in the late G1 phase of the cell cycle, with a peak in early S phase (Lu and Gilbert, 2007). As this transcription occurs just before the replication of pericentromeres and HELLS was already shown to be able to bind RNA transcripts, it is tempting to speculate that these pericentromeric RNAs might trigger the initial recruitment of HELLS to chromatin and support the role of the protein in heterochromatin assembly (Wang et al., 2015). Noteworthy, a mild upregulation of pericentromeric transcription was also reported in Suv39DKO cells (Peters et al.,

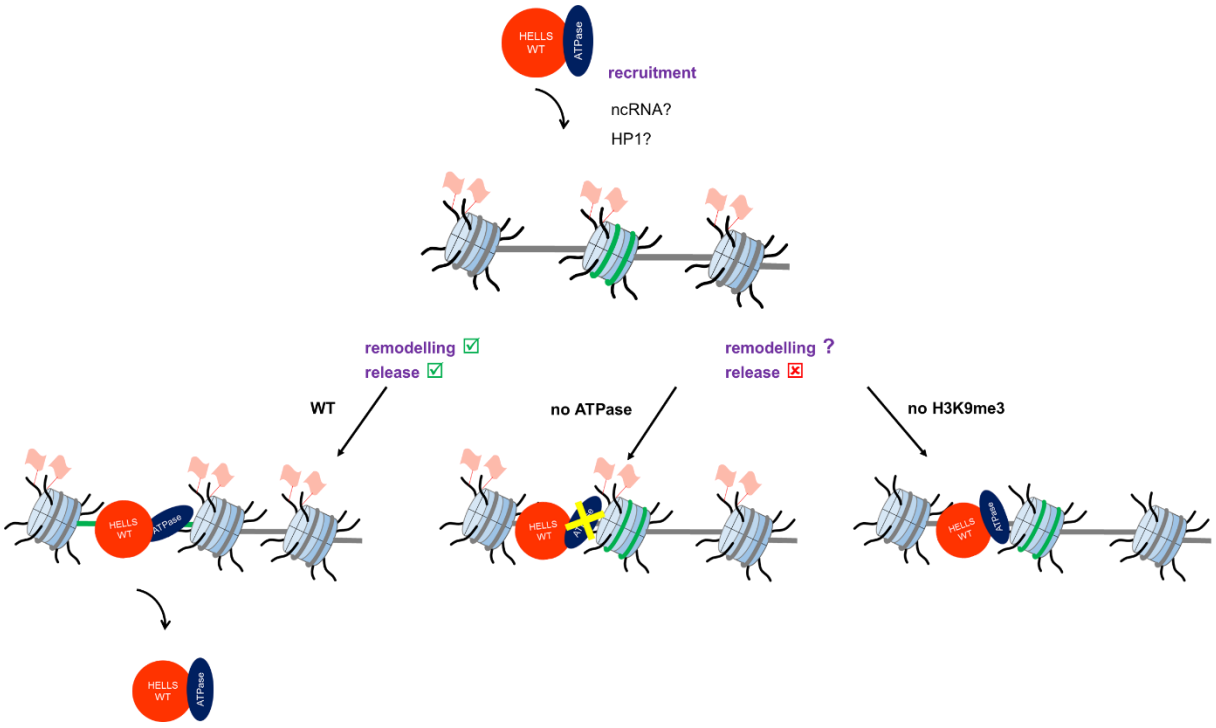


Figure 42| Model for chromatin-mediated recruitment and release of HELLS. Chromatin environments where the nucleosomes carry appropriately-modified histones (in this case H3K9me3 marks, shown as flags) mediate the recruitment of HELLS. This process could be supported by additional factors such as ncRNAs or members of the H3K9me3 pathways, like HP1. Left-handed panel: once recruited, the ATPase activity of HELLS may be stimulated by histone tail modifications. This will result in the exposure of the nucleosome-restricted DNA sequence (shown in green) and eviction of HELLS from chromatin. By contrast, in the absence of ATPase activity (middle panel) or an intact H3K9me3 pathway (right-handed panel) the release of the protein from chromatin is inhibited. The HELLS ATPase module is shown in dark blue, while the rest of the protein is depicted in red.

2003). These transcripts might stabilize the association of HELLS with chromatin and contribute to the increased accumulation of the protein at chromocenters, as observed in the microscopy experiments reported here. Once recruited at these sequences, the absence of an intact H3K9me3 environment would lead to the local accumulation of HELLS at these sites. Importantly, the fact that the ATPase-deficient HELLS was not observed to be affected by the deletion of the H3K9me3 pathway, suggests that the chromatin environment is important for the activation and not for the recruitment of the remodeling activity. This hypothesis is supported by a report documenting that HELLS is detected in a complex with HP1 by immunoprecipitation performed after formaldehyde crosslinking (Yan et al., 2003a). Collectively, the results presented in the present work together with the ncRNA based-recruitment mechanism suggested by literature (Wang et al., 2015), hint at a two-step model for the activation of HELLS. This involves the RNA-mediated recruitment of the protein to chromatin, where depending on the local epigenetic context, the remodeling activity can be initiated (**Figure 42**). This may prevent the protein to rapidly diffuse away from substrates that carry

appropriately-modified nucleosomes and may propagate the downstream recruitment of epigenetic silencers.

In parallel with continuous sampling mechanisms, remodeling complexes were also shown to be engaged in specific recruitment or immobilization at specific nuclear compartments (Längst and Manelyte, 2015). This was previously documented for the chromatin remodelers Brg1 and BRM, where dexamethasone treatment of the cells led to the concentration of these proteins in a single spot in the nucleus (Johnson et al., 2008). Noteworthy, also in this work, a tight association with lamin B was observed for a subpopulation of HELLS-transfected cells (**Appendix 1, Figure 1** and **3**). Since LADs are enriched in repetitive DNA sequences, which are HELLS target regions, it is enticing to speculate that the recruitment of the protein to a stable nuclear landmark such as the lamina, would support the efficient silencing of the lamina-associated chromatin compartments. Here, through its remodeling activity, HELLS might open the highly compacted chromatin sites and recruit the silencing machinery (see below). Comparing the dynamics of HELLS at the nuclear lamina and pericentromeric heterochromatin might offer deeper insights into the HELLS-mediated regulation of these heterochromatin sites.

5.1.3 Chromatin remodeling-dependent regulation of DNMT3 enzymes *via* HELLS

Although in this work the involvement of the HELLS ATPase activity in regulating the activity of DNA methyltransferases was not directly addressed, the rapid exchange observed for HELLS at pericentromeric heterochromatin and its ATP hydrolysis-dependent release suggest that the enzyme might mediate the recruitment of DNMTs at compacted chromatin sites such as chromocenters and LADs. Engagement with appropriately modified nucleosomes could lead to the local stabilization of HELLS at these sequences and recruitment of silencing complexes. Exposing the nucleosome-constrained DNA sequence though remodeling might further positively enforce the association of the silencing machinery with the remodeled locus. By contrast, the high mobility of the freely diffusing HELLS would prevent the local nucleation of repressive complexes at loci that do not require epigenetic silencing. Consistent with the recruitment hypothesis, recent experiments from the group of Kathrin Muegge showed that the remodeling activity of HELLS is required in embryonic stem cells for *de novo* methylation of repetitive elements and for promoting the stable association of DNTM3B

with these sequences (Ren et al., 2015). This effect was later reported to be independent of the differentiation stage of the *Hells*$^{-/-}$ cells used for the rescue experiments, as similar results were obtained upon exogenous expression of HELLS in differentiated mouse fibroblasts (Termanis et al., 2016). These data show that, provided the expression level of DNMT3 enzymes is high enough, HELLS dependent *de novo* DNA methylation can occur in somatic cells in the absence of exogenous signals. Importantly, in both somatic and embryonic stem cells, the ability of HELLS to rescue the DNA methylation levels was dependent on its ATP binding and hydrolysis property (Ren et al., 2015; Termanis et al., 2016).

Interestingly, *Hells* knock-out was documented to lead to the accumulation of H3K4me3 marks at satellite repeats (Dunican et al., 2013; Yan et al., 2003b). Taking into account that the enzymatic activity of DNMT3A and 3B is inhibited by K4me3-modified H3 tails (Zhang et al., 2010), it is tempting to speculate that the severe reduction of DNA methylation observed at repetitive sequences upon *Hells* knock-out is not only due to the spatial inaccessibility of these sites but also to the local increase in histone modifications that antagonize DNA methylation. This hypothesis could be tested by employing synthetic chromatin substrates to dissect the effect of different histone modifications on the activity of DNMT3 enzymes, in the absence and presence of HELLS. Alternatively, cellular experiments could be performed by resorting to inhibitors of MLL activity to reduce H3K4me3 levels and test whether HELLS can rescue DNA methylation at these sites.

5.2 Spatiotemporal control of DNMT3A activity by protein interactors

5.2.1 MeCP2 allosterically regulates DNMT3A activity

During the past decade, a large body of work has accumulated, indicating that the global DNA methylation patterns are highly dynamic and are the product of ongoing *de novo* methylation and demethylation events (Jeltsch and Jurkowska, 2014). This insight has brought about a revision of the elegant textbook DNA maintenance model ,which was proposing that DNA methylation patterns are introduced by *de novo* DNMT3 enzymes during early development and maintained by DNMT1 after every replication cycle (Jeltsch and Jurkowska, 2014). The revised concept describes DNA methylation as a dynamic stochastic model in which the 5mC modification status at

each site is determined by the local activity of DNMTs, TET enzymes (which are involved in DNA demethylation) and the rate of DNA replication (Jeltsch and Jurkowska, 2014). Importantly, this conceptual update places chromatin and protein interactors at the core of targeting and regulation of DNMT activity.

The dynamic DNA methylation landscape plays particularly important roles in non-dividing cells, such as terminally differentiated neurons (Edwards et al., 2010; Heyward and Sweatt, 2015; Weaver and Bartolomei, 2014). Here, in the absence of cell division and DNA replication, DNA methylation profiles can only be controlled through a tight regulation of DNA methylating and demethylating enzymes. Despite their outmost importance, the regulatory processes that govern the targeting and activity of DNMT enzymes are not clear at present. Insights into the cascade of regulatory events would be beneficial for understanding how DNA methylation patterns are set and maintained both in healthy development and during pathogenesis (Bergman and Cedar, 2013; Sharma et al., 2016). In this work, I took a closer look at the regulation of DNMT3 methyltransferases, in particular DNMT3A, by the chromosomal protein MeCP2. This project was based on previous results from our laboratory, documenting a direct interaction between the ADD domain of DNMT3 enzymes and the TRD domain of MeCP2, which was shown to translate into a strong inhibition of DNMT3A activity *in vitro*.

By employing a combination of *in vitro* and cellular assays, I was successful in elucidating the mechanism through which this concentration-dependent inhibition occurs and demonstrated that the strong interaction between DNMT3A and MeCP2 is not an *in vitro* artifact, but can be also detected in mouse whole brain lysates. This tissue was used for investigation since it displays high levels of DNMT3A and MeCP2 (Gabel et al., 2015). Using conformationally locked DNMT3A variants as a novel tool to investigate DNMT3A regulation, I could show that through its binding the TRD domain stabilizes the autoinhibitory conformation of the enzyme and inhibits catalysis on a broad range of substrates. Importantly, this strong interaction and its resulting inhibition were found to be overcome by the presence of unmodified H3 N-terminal peptide, indicating that the binding of the TRD and unmodified H3 to the ADD domain are mutually exclusive.

To investigate the crosstalk between MeCP2 and DNMT3A *in vivo*, whole-genome bisulfite sequencing (WGBS) was performed on genomic DNA isolated from the whole

brain of adult wild-type and *Mecp2* KO mice, respectively (Guy et al., 2001). The overlap in symptoms between the *Mecp2* KO animals and Rett syndrome patients makes these KO mice a disease-relevant model system. For this experiment, 6-week old homozygous null males were chosen, owing to their earlier symptom onset and the homogenous population of MeCP2-negative cells (Guy et al., 2011). Interestingly, unlike the global effects observed upon overexpression of MeCP2 in the HCT116 hypomorphic cells, WGBS analysis revealed only punctate changes in the DNA methylation landscape of the brain of *Mecp2* KO animals. Moreover, overexpression of MeCP2 in HCT116 cells resulted in a dramatic two-fold reduction of global DNA methylation levels, in agreement with the inhibitory function of the protein. By contrast, WGBS analysis of the *Mecp2* KO brain revealed both, hyper- and hypomethylated, DMRs, indicating that the MeCP2-DNMT3A interaction is more complex in the brain. Importantly, around half of the discovered DMRs displayed a change of more than 50% in methylation levels averaged over 1kb tiles and overlapped with MeCP2 binding sites, as extracted from a recent high-resolution ChIP-seq data set from olfactory epithelium (Rube et al., 2016) suggesting that they are directly related to the loss of MeCP2. These data suggest that MeCP2 has a local, but strong effect on the brain methylome.

The discrepancies observed between the cellular and *in vivo* experiments, could be explained by the different expression levels of MeCP2 between the two systems. In addition, whereas the effects of MeCP2 overexpression were analyzed in clonally-derived homogenous cell lines, the brain consists of a heterogeneous mixture of different cell types that express MeCP2 at different levels and where the regulatory roles of the protein may be more complex (Guy et al., 2011). In line with this, several studies already documented the large differences in MeCP2 expression between mature neurons and glial cells, with the latter cell type being more abundant and expressing relatively low amounts of MeCP2 (Ballas et al., 2009; Maezawa and Jin, 2010; Maezawa et al., 2009). This heterogeneity may obscure the MeCP2-mediated DNA methylation effects in the WGBS analysis performed here. In line with this, transcriptional profiling studies using whole brain from *Mecp2*-null mice revealed only subtle changes in gene expression, whereas analysis of defined sub-brain regions such as the cortex, cerebellum and hypothalamus discovered up to thousands of significant gene expression changes, depending on the area selected for investigation (Ben-Shachar et al., 2009; Urdinguio et al., 2008). Still, in most cases the fold change

in the expression of the affected genes was rather mild, indicating that MeCP2 may rather act as a fine-tuner of gene expression (Ben-Shachar et al., 2009).

Interestingly, comparing the gene expression patterns in the hypothalamus of *Mecp2*-null and *MECP2*-Tg mice, which express 2-3 more MeCP2 than the wild-type animals, revealed that most genes were down-regulated in *Mecp2*-null and upregulated in *MECP2*-Tg mice (Ben-Shachar et al., 2009). This proposes an activating role of MeCP2 in gene expression and is in line with the inhibitory effect of MeCP2 on DNA methylation that is observed in the present work. Noteworthy, gene ontology (GO) analysis of the transcripts altered in cerebella of models of MeCP2 disorders found that genes upregulated by MeCP2 were more likely to be associated with specific processes and GO terms, by comparison, to downregulated genes. This is important since it suggests a closer functional relatedness between activated target genes compared with repressed genes. The fact that some of the categories that are enriched in the upregulated targets include transmission of nerve impulse, dendrite development, and neuroblast proliferation, support an important role in the pathogenesis of MeCP2 disorders. Noteworthy, these results could only be obtained by analyzing gene expression data stemming from isolated brain sub-sections and not when performed on single-cell derived wild-type and mutant *MeCP2* expressing fibroblast clones, patient T lymphocytes, *Mecp2*-null mouse whole brain, post-mortem brain tissue from Rett patients or SH-SY5Y neuronal cell lines (Ben-Shachar et al., 2009; Colantuoni et al., 2001; Delgado et al., 2006; Kriaucionis et al., 2006; Peddada et al., 2006; Traynor et al., 2002). This highlights the importance of appropriate model systems in dissecting the complex roles played by MeCP2 in gene regulation. Based on the evidence described above, it is expected that analyzing the methylomes of brain sub-regions where the absence of MeCP2 was already shown to cause more dramatic effects, might lead to more pronounced effects and providing a more defined picture of the role that MeCP2 plays in the regulation of DNMT3A *in vivo*. Since different types of neurons are interspersed through different brain regions, neuronal cultures that have been prepared from Rett patient cells using induced pluripotent stem cell technology, might provide a better defined study system for this purpose (Avior et al., 2016; Hotta et al., 2009; Marchetto et al., 2010).

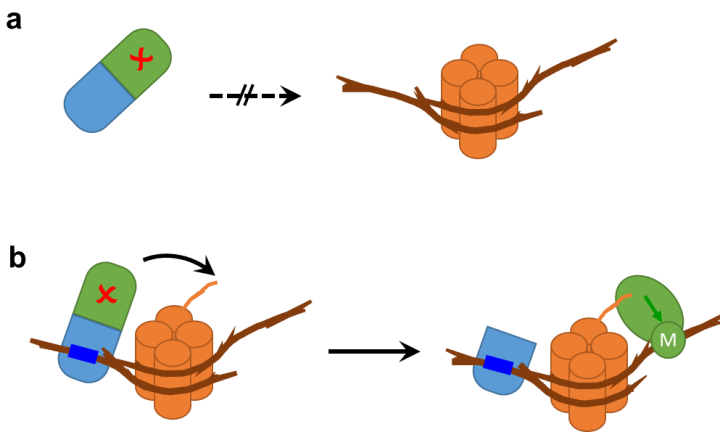


Figure 43. MeCP2 has a dual role in the regulation and targeting of DNMT3A. a) Binding of MeCP2 (blue) to DNMT3A (green) inactivates the methyltransferase and prevents untargeted activity (red x). **b)** After MeCP2 binding to DNA at a genomic locus with appropriately modified H3 tail, H3 tail binding to DNMT3A leads to the dissociation of the DNMT3A-MeCP2 complex and subsequent DNA methylation. This can trigger more MeCP2 binding and by this initiate a positive feedback mechanism. This image was taken from **Appendix 2**.

In summary, the data stemming from this project propose a model in which DNMT3A is under the combined control of MeCP2-mediated targeting and inhibition and the modification state of histone H3 tail at genomic target sites (**Figure 43**). On the one hand, MeCP2 was found to act as an inhibitor of DNA methylation at some target sites in the brain, and after overexpression in a global manner in tissue culture (**Appendix 2**, **Figure 8A**). On the other hand, at repetitive genomic sites such as satellite repeats and a subset of gene promoters, MeCP2 could also recruit DNMT enzymes (**Appendix 2**, **Figures 6** and **7**). Depending on the modifications status of histone tails at the genomic locus, the TRD-mediated inhibition can then be either relieved or maintained. Such a mechanism would protect H3K4me3-rich promoters from abnormal *de novo* methylation while creating a positive 5mC writing-reading feedback loop at sites where the nucleosomes carry appropriately modified histone tails. The model resulting from this study, supports and mechanistically explains the previous findings proposing both positive and negative roles for the chromosomal protein MeCP2 on gene transcription.

5.2.2 Further directions for dissecting the MeCP2-DNMT3A circuit

While in the work documented here, I focused on analyzing the effect that the TRD domain of MeCP2 has on the activity of DNMT3A, it would be interesting to next explore the cross-talk between these two main epigenetic players, with full-length proteins. This would be important for understanding the role played by the MBD domain of MeCP2 in regulating the cross-talk between hydroxymethylated DNA substrates and DNMT3A activity at non-CG sites. As recently proposed by Greenberg and co-workers,

the major accumulation of hydroxymethylated CG sites during brain development might serve as a ‘kick off’ signal to re-distribute MeCP2 from high-affinity CG sites to the newly methylated CA dinucleotides (Kinde et al., 2015). Taking into account that on HpaII pre-methylated DNA substrates the MeCP2-mediated DNM3A inhibition was maintained (**Appendix 2**, **Figure 4A**) it would be interesting to test if converting these sites to hydroxymethylated CGs would have any effect on DNMT3A activity at non-CG dinucleotides.

In addition to the modification status of the histone H3 tail, the DNMT3A-MeCP2 interaction could be regulated by other factors as well. For instance, MeCP2 was reported to be extensively phosphorylated in the brain, and loss of phosphorylation at specific sites was shown to have negative effects on mammalian brain maturation, learning and plasticity (Damen and Heumann, 2013). In this work, the S/T kinase CDKL5 was found to phosphorylate the TRD domain of MeCP2. Investigating the effect that this PTM has on the interaction between MeCP2 and other proteins, including DNMT3A, may offer deeper insights into the complex regulatory roles of this chromosomal protein. To this end, the CDKL5 phosphorylated residue/s could be identified via mass spectrometry, and phosphorylation mimics of MeCP2 could be expressed in neuronal cell culture to test the sensitivity of interacting partners to the phosphorylation state of the identified residue. Noteworthy, CDKL5 was shown to be involved in a negative feedback loop with MeCP2, where increasing MeCP2 protein levels by overexpression or stimulation of cells with serotonin or cocaine, was found to lead to a MeCP2-dependent repression of the *Cdkl5* gene locus (Carouge et al., 2010; Damen and Heumann, 2013). It, would therefore, be interesting to test if in turn, the CDKL5-mediated phosphorylation has any effects on the stability and functions of MeCP2.

5.3 Epigenetic biosensors for locus-specific visualization of epigenetic modifications in living cells

5.3.1 Development and validation of locus and modification-specific BiAD sensors

Covalent modifications of DNA and histone proteins regulate chromatin structure and play central roles in chromatin-templated biological processes such as transcription and DNA repair. Unlike the genome, epigenetic signatures are metastable, with

different epigenetic signals having different degrees of variability and stability. This results in a large heterogeneity in epigenetic profiles in-between cells, even when these are part of an apparently homogenous population (Bheda and Schneider, 2014). This non-genetic variation may be responsible, for instance, for the variability in responses observed between cells when these are stimulated with the same signal, and its analysis would improve our understanding of how epigenetic states are maintained and inherited during cellular differentiation and disease development (Tirado-Magallanes et al., 2016; Yu et al., 2016).

Epigenetic heterogeneity is especially important to detect in cells that cannot be isolated in abundance. For this reason, rare cell types such as the oocyte and zygote were among the first objects of single-cell epigenomics (Guo et al., 2013, 2014b). An important observation derived from these studies was that individual pronuclei of the same cell type are highly heterogeneous, even if inspected at the same time point after intracytoplasmic sperm injection into oocytes (Guo et al., 2014b). A large heterogeneity in DNA methylation landscapes was also observed in ESCs cultured in different types of media (Guo et al., 2013; Smallwood et al., 2014). This was attributed to the dynamic balance between the primed and naïve states of ESCs cultured under serum conditions, by comparison to the rather homogenous population obtained in 2i media (Smallwood et al., 2014). Mapping epigenetic heterogeneity at the single cell level is equally important for instance, for tumor samples, where this can provide a valuable insight into cell type-specific deregulation of epigenetic pathways (Tirado-Magallanes et al., 2016). In addition to underpinning epigenetic heterogeneity via single cell analysis methods, equally critical are tools that allow the assessment of the temporal order in which epigenetic changes occur during cellular reprogramming, the role and importance of each regulatory step, as well as the dynamic connection between epigenetic changes and morphological or physiological alterations.

Since epigenetic signals regulate chromatin functions in a genomic context-specific manner, analyzing their changes with locus-specific resolution is imperative for an operative understating of epigenetic regulatory cascades. The genomic context importance of epigenetic marks is, for instance, evident in diseases such as cancer. Here, selected CpG islands for example of tumor suppressor genes usually gain DNA methylation, while repetitive sequences such as heterochromatic DNA repeats, dispersed retrotransposons, and endogenous retroviral elements, become frequently

hypomethylated (Baylin and Jones, 2011; Bergman and Cedar, 2013; Gal-Yam et al., 2008; Keshet et al., 2006). Stochastic epigenetic errors were also found to be associated with imprinting disorders or *in vitro* fertilization (Kalish et al., 2014; de Waal et al., 2014).

Based on the arguments listed above, it is clear that methods that allow a dynamic readout of epigenetic signals at endogenous genomic loci, inside the nucleus of living cells, are essential for a better understanding of epigenetic mechanisms (Bheda and Schneider, 2014). In this project, this urgent and unmet technological demand was addressed by developing a set of novel BiFC-based epigenetic biosensors for live cell microscopy applications. In these tools, the high sequence specificity of engineered binding domains (ZF, TALE, and CRISPR/dCas9) was combined with the chromatin mark-specific recognition of epigenetic reading domains (MBD for 5mC readout; HP1 chromodomain for H3K9me3 readout). Simultaneous readout of the DNA sequence and the epigenetic mark was achieved by fusion of the anchor and detector modules with complementary parts of the split Venus fluorophore. Accordingly, if the genomic locus of interest carries the candidate epigenetic modification, the docking of the anchor and detector modules in close spatial proximity would lead to the reconstitution of an optically trackable fluorophore. With these tools, the status of 5mC and H3K9me3 marks at endogenous genomic sites in living cells could be directly detected for the first time, and the dynamic change of these marks upon drug treatment and induction of epigenetic enzymes could be followed over time.

Several controls were included to validate the locus and modification specificity of the developed sensors. After ensuring that the functionality of the individual anchor and detector modules was not influenced by their fusion to the complementary Venus parts, the target site specificity of the sensors was validated by sequence-specific DNA staining with synthetic dyes or by co-transfection of the full-fluorophore tagged anchor modules together with the corresponding BiAD systems. The epigenetic mark specificity of the tools was confirmed by employing binding pocket mutants in the detector modules, or cell lines with globally reduced levels of DNA methylation or H3K9me3 marks.

With this technology, the site-specific increase of DNA methylation and H3K9me3 levels could be effectively monitored after exogenous expression of the responsible active epigenetic enzymes in the corresponding knock-out cell lines. These results

highlight the portability and specificity of the novel BiAD sensors, and demonstrate that the dynamic range of these tools covers the extent of biological relevant changes affecting the 5mC and H3K9me3 marks.

5.3.2 Advantages of the BiAD approach over existing technologies

To date, these novel sensors are the only available tools that enable the locus-specific detection of epigenetic modifications in the nucleus of living mammalian cells. Other biochemical methods that allow the locus-resolved mapping of DNA methylation or histone modifications depend on isolated DNA or chromatin as input material, thereby being unsuitable for studying dynamic epigenetic cascades (Bheda and Schneider, 2014). Also, other microscopy-based approaches such as *in situ* hybridization and proximity ligation assays (PLA) make use of fixed cells (Ginart et al., 2016; Gomez et al., 2013; Li et al., 2013). Recently, a live-cell method was developed that uses the expression of a genomically inserted fluorescent reporter gene to measure the methylation state of the adjacent DNA (Stelzer et al., 2015). While this method was successfully applied to track locus-specific changes of DNA methylation in living cells, it comes with a set of disadvantages. Among the most important are the fact that the signal depends on the efficient spreading of DNA methylation and demethylation from the gene locus into the promoter of the reporter, which may vary between genomic sites and may not happen under certain conditions at all. Moreover, the method requires the engineering of the endogenous DNA locus, which may disturb the native chromatin state. In addition, the fluorescence-based readout limits the temporal resolution of the measurement because, after activation of the reporter gene, it needs to be transcribed, the mRNA processed, transported and translated. Furthermore, the reporter protein needs to maturate and accumulate to visible levels. By contrast, the BiAD sensors developed here allow a direct readoutout of the epigenetic marks present at the target epigenetic locus, without the need to engineer the genomic site. Furthermore, with this method, the spatial information is retained as the fluorescent signal is formed directly at the target locus and does not freely diffuse within the cell. Preserving the spatial relationship of epigenetic signals is important for understanding the connection between epigenetic marks and the spatiotemporal dynamics of the DNA sequences that these decorate and to enable the later parallel readout of several signals.

One of the clear advantages of the BiAD technology is its modularity, as demonstrated by the successful generation of ZF, TALE and CRISPR/dCas9-based biosensors in which the anchors could be successfully combined with two distinct detector modules. Although so far the system was only applied to visualize epigenetic modifications of repetitive genomic sites, implementation of the CRISPR/dCas9 anchor in novel sensors could facilitate the visualization of epigenetic marks at single copy loci. In line with this, Chen and colleagues demonstrated that a cocktail consisting of 26 ~ 36 sgRNAs tiling a unique locus site was enough to detect the target locus by conventional confocal microscopy (Chen et al., 2013). While this would suffice to visualize the DNA methylation status of a heavily methylated, long CG islands, more effort needs to be put into further reducing the number of required sgRNAs, before shorter regulatory elements can be addressed. This could be accomplished by taking advantage of the recent developments in microscopy technologies and imaging processing methods (Coltharp et al., 2014; Kusumi et al., 2014; Mishin et al., 2015). For instance, in a recent study Chen et al. (2016) combined photoactivated localization microscopy (PALM) with fluorescence complementation to detect protein complex in live cells with nanometer resolution and single molecule sensitivity (Chen et al., 2016b). Noteworthy, by comparison to co-localization microscopy, the BiFC approach not only improves the spatial resolution of the signal, but it also results in lower background fluorescence since it requires the successful complementation of two non-fluorescent fluorophore parts before a visible signal is generated. By having a controlled expression of the individual BiAD modules, random fluorophore reconstitution events and thereby signal leakiness could be further minimized.

5.3.3 Considerations on the design of BiAD modules

In addition to a resorting to a comprehensive toolbox of versatile anchor domains, two distinct detector modules were used in this study to detect two central epigenetic marks. Accordingly, the MBD of MBD1 was employed for the detection of 5mC, while the chromodomain of HP1β (HP1CD) was used for the recognition of H3K9me3 marks (Bannister et al., 2001; Hendrich and Bird, 1998; Jacobs et al., 2002; Li et al., 2015a; Ohki et al., 2001). There are several advantages that these domains possess over other types of recombinant binders, such us immunoglobulin-derived fragments or DARPins, which are used in cell biology applications (Helma et al., 2015). Importantly, since both of the affinity reagents used in this study are part of naturally occurring nuclear mammalian proteins it might be that their synthesis, folding, and degradation

would pose a smaller burden for the cells, by comparison to synthetically designed proteins. To be kept in mind at the sensor selection/design stage was that both full-length MBD1 and HP1β are multi-domain proteins that were already shown to take part in several main epigenetic pathways (Fujita et al., 2003; Hiragami-Hamada et al., 2016; Liu et al., 2013a; Mattout et al., 2015; Reese et al., 2003; Sarraf and Stancheva, 2004). To minimize the potential competition between the detector domains and the endogenous factors, the proteins were trimmed down to their core MBD and chromodomain respectively, by removing the domains which were not essential for their incorporation in the BiAD systems. This was particularly critical for HP1β, since the protein was reported to directly interact via its chromoshadow domain with the SUV39H N-terminus (Schotta et al., 2002; Yamamoto and Sonoda, 2003). By making use only of the chromodomain of HP1β, it was ensured that that the recovery in BiFC signal observed upon induction of active SUV39H1 in *Suv39h1h2*$^{-/-}$ iMEFs was caused by an increase of H3K9me3 levels at major satellite repeats, and not by the SUV39H1-mediated recruitment of HP1 to these sequences. Development of novel affinity reagents to target epigenetic modifications needs to be built on a careful selection of protein domains, to keep the perturbation of endogenous epigenetic cascades minimal. Since for some epigenetic readers it was documented that their affinity for the target modification is enhanced by the presence of accessory protein domains or PTMs, it is possible that stripping the reader down to its core chromatin binding domain might produce a device with low binding affinity (Badugu et al., 2005; Canzio et al., 2011, 2014; Maison et al., 2011). This aspect needs to be tested on a case-by-case basis and may be alleviated by selection of protein domains that are biochemically well characterized and are amenable to engineering approaches. This may not only result in chromatin binders with improved chromatin binding properties, but also lead to the development of affinity tools with novel specificities. For instance, in an early proof-of-principle study the binding preference of the PHD finger-containing protein BPTF could be successfully altered from H3K4me3 to H3K4me2, by a single amino-acid exchange in the aromatic cage of the PHD finger (Li et al., 2007). Implementation of this methodology can result in the generation of affinity reagents that can specifically discriminate between modification states that are chemically only minutely different from each other, but which are involved in epigenetic signaling cascades with distinct biological outcomes. Taking into account that these modifications are present on peptides with identical amino-acid sequence, generating specific affinity reagents

through modifying naturally occurring chromatin domains is expected to be superior to antibody-based approaches.

In addition to a high specificity for the targeted chromatin mark, other important parameters affecting the selection of detector devices used in this study was cytotoxicity and binding strength. These aspects are particularly critical for the successful implementation of the detector devices into functional BiAD sensors since the modules need to display a chromatin association that is stable enough to promote BiFC pairing, while in the same time only minimally perturb the local epigenetic landscape and ongoing chromatin-templated cellular processes. The MBD was selected for 5mC readout since the protein was shown to have a rapid turnover rate at mouse chromocenters (as revealed by FRAP), and it was successfully used to generate stably expressing cells lines and a DNA methylation mouse reporter (Ueda et al., 2014). Similarly, FRAP measurements also documented a high mobility of HP1β on chromatin, with a half recovery time $t_{1/2}$ of ~ 4s (Manukyan and Singh, 2014). The results of the present work indicate that selection of detector modules with a chromatin binding kinetics similar to what was reported for HP1β and the MBD of MBD1, is beneficial for high-efficiency BiFC outputs. Of note, since chromatin readers usually display pools with different kinetic behaviors, a detailed kinetic modeling is required to be able to pinpoint the optimal mobility parameters of the most suitable detector modules (Manukyan and Singh, 2014). Importantly, due to their generally lower binding affinity and transient residence time, chromatin readers are expected to interfere less with cellular processes by comparison to mintbodies and nanobodies, which bind cellular structures tighter and for a longer time (Helma et al., 2015; Kungulovski et al., 2014).

Arguably, one the strongest advantage of the technological platform provided in this work stems out of using reading domains as detector modules. For instance, in proximity ligation assays, one antibody is used for the recognition of the biotin molecule attached to the probe that binding to the DNA sequence, while another antibody is needed for the recognition of the histone mark. Owing to this, PLA assays have a restricted in applicability since there is only a limited number of animal sources can be used for antibody production. By contrast, the only restriction of the BiAD system is that the reconstituted fluorophores need to be spectrally separable. This strongly

increases the flexibility of the approach and lays the basis of a platform that can be used for multiplexed assessment of distinct epigenetic marks at several genomic loci.

5.3.4 Interdisciplinary applications of the BiAD technology

Metabolic reprogramming resulting either directly or indirectly from oncogenic mutations is the hallmark of cancer development (Hanahan and Weinberg, 2011; Pavlova and Thompson, 2016). An increasingly large body of work has reinforced the view that metabolic features are highly heterogeneous, with each cancer cell likely exhibiting different metabolic profiles. This heterogeneity is to a large extent influenced by the epigenetic landscape and environmental conditions (Gao et al., 2016). Oncogenic transformation involved not only a massive rewiring of the metabolic profile but also dramatic alterations in the cellular behavior. Accordingly, in their transition from epithelial to mesenchymal state, cells lose their apico-basal polarity and tight contacts and gain an irregular, rounded morphology characterized by an increased number of filopodia and lamellipodia (Campbell and Casanova, 2016). While epigenomic profiling has significantly improved our understanding of cancer development, methods that allow simultaneous access to the epigenetic, metabolic and morphological information are currently lacking.

Owing to the fact that the BiAD tools are compatible with live cell applications, these sensors can set the basis of a holistic methodology approach. Accordingly, the BiAD modules could be used in combination with glucose or hypoxia sensors to investigate whether the alterations in metabolic profiles translate into changes of epigenetic marks at candidate genomic loci (Danhier et al., 2015; Song et al., 2016). Importantly, the deregulation of metabolic pathways correlated to cancer progression was also frequently reported to be a direct modulator of the activity of epigenetic enzymes. For instance, in a recent study, methionine restriction was found to alter the metabolism of cellular SAM and impact gene expression reprogramming through histone methylation (Mentch and Locasale, 2016; Mentch et al., 2015). Overexpression of NNMT (nicotinamide N-methyltransferase) was also observed in a variety of cancers such as the lung, liver, kidney, bladder and colon, where it is thought to support tumor growth by acting as a methylation sink and diverting methyl groups away from other methylation processes (Gao et al., 2016; Roeßler et al., 2005; Ulanovskaya et al., 2013). Based on these lines of evidence, combining the BiAD system with recently developed SAM sensors would provide a novel imaging platform that is expected to

connect with unprecedented insight, alterations in the one-carbon metabolism pathway with the methylation status of gene regulatory elements (You et al., 2015). Furthermore, by taking advantage of the recent technological revolution in 3D culture models, like organotypic cultures, the BiAD sensors could be used to correlate epigenetic changes with the alterations in cellular morphology that take place during oncogenesis (David et al., 2016; Golovko et al., 2015; Schweiger and Jensen, 2016).

5.3.5 Current limitations of the BiAD approach

Although the BiAD approach does not require the genetic modification of the target locus, the perturbing effects that this technology might have on the local chromatin environment are still a matter that requires deeper analysis. Conversely, the effects that the local geometry of the chromatin environment might have on the strength of the BiFC signal are not fully clear. For instance, studies in which the nuclease active Cas9 protein was used for gene editing applications, revealed that Cas9 cutting was more promiscuous *in vitro* than *in vivo* (Cho et al., 2014). This could be explained by chromatin blockage of the accessibility of many off-target sites *in vivo* and suggests that the local chromatin environment may play important roles in the target selection of Cas9 (Horlbeck et al., 2016; Kuscu et al., 2014; O'Geen et al., 2015; Wu et al., 2014). In line with this, single-molecule imaging to track fluorescently labeled dCas9 in living mammalian cells revealed that Cas9 searches for target sites predominantly by rapid three-dimensional diffusion. Interestingly, the movement of the protein appeared to be slowed-down within compacted pericentromeric heterochromatin domains (Knight et al., 2015). These results indicate that Cas9 search efficiency is not eliminated in heterochromatic regions, in line with the high BiFC yields observed in the present work when dCas9 was targeted to alpha satellite sequences. Nevertheless, the study by Knight et al. (2015) raises a cautionary point about applying the BiAD technology to compare the modification status of genomic loci found within chromatin domains that have drastically different degrees of compaction. Accordingly, discrepancies in signal intensities between different DNA sites might be connected with different levels of the epigenetic mark or with the distinct search kinetics of the anchor devices within differentially compacted chromatin environments. This issue could be alleviated by visualizing the target searching behavior of full fluorophore-tagged Cas9 at the two different binding sites and adjusting the time given for the BiFC signal to form such that the imaging is performed at steady state.

An aspect which is more difficult to address is whether the status of the epigenetic modifications that is detected with the BiAD sensors truly reflects the native chromatin state at the target locus or if this is rather a reaction of the epigenetic landscape to the tools employed for readout. This issue might be particularly problematic when using large proteins such as TALEs and dCas9 as anchors. While in a recent study, TALEs developed to bind repetitive sequences found in *Drosophila* heterochromatic domains were documented to have no influence of the chromatin architecture of the target loci, the effects of dCas9 binding remain controversial (Yuan and O'Farrell, 2016). For instance, by designing sgRNAs that target dCas9 to 16 regions with low chromatin accessibility in mouse ESCs, Barkal et al. (2016) found that dCas9 binding resulted in an increased chromatin accessibility at the surrounding target sites (Barkal et al., 2016). The opening was strongest within 20 bps of the target sites, and an increase in DNAse sensitivity could be detected up to 100 bps away. Although the increase in DNAase hypersensitivity was mild, of only 1.7 fold, these results highlight that in spite of the utility of dCas9 fusion proteins in a panoply of applications, including imaging, the effects that the protein has on the local chromatin environment need to be assessed on a locus-by-locus basis by a combinatorial methodology approach.

The ultimate aim of the BiAD approach developed in this work is to enable the real-time tracking of locus-specific epigenetic marks within the nucleus of single living cells, during cellular differentiation, pathogenesis or alterations in the cellular environment. Although so far the sensors could be successfully used to visualize locus-specific changes of epigenetic modifications upon drug treatment and overexpression of epigenetic enzymes, the tools had to be transfected freshly before each imaging time-point. While this approach enables a better temporal resolution than ChIP-based methodologies, it does not yet allow the epigenetic tracking of individual cells during cellular transitions. For this, stable cell lines could be generated to express the sensors either constitutively or under inducible promoters. This direction of development is supported by the already successful implementation of the MBD to generate a DNA methylation reporter mouse allowing to visualize the dynamics of heterochromatin during cellular differentiation (Ueda et al., 2014). In addition, for an improved real-time imaging of locus-specific epigenetic changes, the Venus fluorophore, for which the reversibility of the complemented signal is still debated, could be replaced with IFP1.4 (Kodama and Hu, 2012; Tchekanda et al., 2014; Yu et al., 2013). This protein was engineered to be capable of reversible fluorescence complementation and was already

successfully employed to analyze the spatiotemporal dynamics of hormone-induced signaling complexes in mammalian cells at nanometer resolution (Tchekanda et al., 2014). Furthermore, due to its fluorescence in the infrared part of the spectrum, this protein is also suitable for whole-body imaging, opening up new avenues for the applicability of the BiAD sensors.

5.3.6 Conclusions

All in all, in this project a novel technological platform was designed, established and applied to enable for the first time the visualization of the status of epigenetic modifications at individual genomic loci in living cells. Two types of fluorescent readout, namely BiFC and FRET, were tested during the course of this work. While in both cases the methodological workflow could be successfully established in the laboratory, the BiFC approach demonstrated a superior performance and sensitivity and was thereby selected as the readout of choice. With these novel tools, site-specific changes that occur upon treatment of the cells with epigenetic drugs or overexpression of epigenetic enzymes could be monitored in living cells. Based on the results obtained in this work, stable cell lines and animals that can stably express the BiAD sensors could be generated. This will enable the locus-specific detection of epigenetic changes during development, transgenerational inheritance, cellular reprogramming, drug treatment or onset of disease. It is anticipated that either in their current form or through combination with the recent developments in gene targeting and microscopy technologies, these tools will contribute to a better understanding of how specific epigenetic signatures are set, erased and maintained during cellular development and pathogenesis.

5.4 Final conclusions and outlook

Taken together, the results presented in this doctoral thesis demonstrate how a balanced combination of biochemical and cellular methods can be used to derive deeper insights into chromatin-templated and epigenetic cellular processes. My work was addressing three major questions in the field of Chromatin Biology and Epigenetics, namely, how epigenetic factors can get access to the chromatin through the support of chromatin remodelers, how epigenetic enzymes are controlled and regulated to establish and maintain defined epigenetic states and how the epigenetic state of a live cell can be detected.

In the first topic of research, the importance of an intact ATPase activity for the interaction between HELLS and chromatin could be revealed. The existence of an ATP-dependent cycle of HELLS binding and release from heterochromatin points, for the first time, toward an ATP-dependent remodeling activity of the enzyme and opens up new avenues of research to understand the role that HELLS dynamics plays in the maintenance of repressive chromatin states.

In the second topic of research, a strong and direct interaction between the 5mC-reading chromosomal protein MeCP2 and DNMT3A could be mapped *in vitro* and confirmed in brain tissue. Mechanistic dissection of the inhibitory effect resulting from this interaction revealed that DNMT3A localization and activity are under the combined control of MeCP2 and H3 modifications and that depending on the modification status of the H3 tail at the target sites, MeCP2 can act as either repressor or activator. Taken together the results derived from this study provide a unifying framework that can successfully explain that activating and repressive roles previously documented for MeCP2. In addition, these data add protein interactors to the list of molecules that are able to regulate the activity of DNMT3 enzymes through an allosteric mechanism. Collectively, these findings enforce the role of allosteric control at the center of DNMT3 targeting and regulation and highlight the importance of characterizing novel DNMT3 interactors and their effects on the enzymatic activity.

In the third topic of research, a novel readout method was developed and applied to visualize for the first time the dynamic status of 5mC and H3K9me3 marks with locus specific resolution in live mammalian cells. This technological platform successfully solves an urgent and unmet need for tools that enable the live cell tracking of epigenetic modifications at the level of endogenous genomic loci and satisfies three

main criteria. First, the method is compatible with live cell imaging, therefore preserving the temporal dimension and allowing the dynamic tracking of epigenetic signals. Second, the microscopy-based readout enables the single cell dissection of epigenetic profiles and a better understanding of cell-to-cell heterogeneities in relation to the cellular environment. Third, by taking advantage of the recent development in genomic targeting technologies, the epigenetic status of endogenous DNA sequences can now be assessed at an unprecedented resolution. Based on all these, it is anticipated that either in its current form or through combination with the recent developments in gene targeting and microscopy technologies, this novel toolbox will greatly contribute to a better understanding of how specific epigenetic signatures are set, erased and maintained during embryonic development or the onset of disease.

6 References

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, a, Lebersorger, a, Singh, P.B., et al. (1999). Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J. *18*, 1923–1938.

Adams, V.H., McBryant, S.J., Wade, P.A., Woodcock, C.L., and Hansen, J.C. (2007). Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. J. Biol. Chem. *282*, 15057–15064.

Adler, D.A., Quaderi, N.A., Brown, S.D.M., Chapman, V.M., Moore, J., Tate, P., and Disteche, C.M. (1995). The X-linked methylated DNA binding protein, Mecp2, is subject to X inactivation in the mouse. Mamm. Genome *6*, 491–492.

Allfrey, V., Faulkner, R., and Mirsky, A. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. … Sci. United States … *315*, 786–794.

Allshire, R.C., Nimmo, E.R., Ekwall, K., Javerzat, J.P., and Cranston, G. (1995). Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. Genes Dev. *9*, 218–233.

Amir, R.E., Van den Veyver, I.. B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat. Genet. *23*, 185–188.

Anastasiadou, C., Malousi, A., Maglaveras, N., and Kouidou, S. (2011). Human epigenome data reveal increased CpG methylation in alternatively spliced sites and putative exonic splicing enhancers. DNA Cell Biol. *30*, 267–275.

Andrews, F.H., Strahl, B.D., and Kutateladze, T.G. (2016). Insights into newly discovered marks and readers of epigenetic information. Nat. Chem. Biol. *12*, 662–668.

Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X., Krueger, F., Smallwood, S.A., Ponting, C.P., Voet, T., et al. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat. Methods *13*, 229–232.

Anton, T., Bultmann, S., Leonhardt, H., and Markaki, Y. (2014). Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR / Cas system. Nucleus *5*, 163–172.

Aran, D., Sabato, S., and Hellman, A. (2013). DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. Genome Biol. *14*, R21.

Arand, J., Spieler, D., Karius, T., Branco, M.R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V., et al. (2012). In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. PLoS Genet. *8*.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature *455*, 818–821.

Arnaudo, A.M., and Garcia, B.A. (2013). Proteomic characterization of novel histone post-translational modifications. Epigenetics Chromatin *6*, 24.

Avior, Y., Sagi, I., and Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Mol. Cell Biol. *17*, 170–182.

Avvakumov, G. V, Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H., and Dhe-Paganon, S. (2008). Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature *455*, 822–825.

Badugu, R., Yoo, Y., Singh, P.B., and Kellum, R. (2005). Mutations in the heterochromatin protein 1 (HP1) hinge domain affect HP1 protein interactions and chromosomal distribution. Chromosoma *113*, 370–384.

Baker, M. (2015). Antibody anarchy: A call to order. Nature *527*, 545–551.

Baker, S.A., Chen, L., Wilkins, A.D., Yu, P., Lichtarge, O., and Zoghbi, H.Y. (2013). An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. Cell *152*, 984–996.

Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C., and Mandel, G. (2005). REST and its corepressors

mediate plasticity of neuronal gene chromatin throughout neurogenesis. Cell *121*, 645–657.

Ballas, N., Lioy, D.T., Grunseich, C., and Mandel, G. (2009). Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology. Nat. Neurosci. *12*, 311–317.

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

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

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F., and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science *354*, 909–912.

Barkal, A.A., Srinivasan, S., Hashimoto, T., Gifford, D.K., and Sherwood, R.I. (2016). Cas9 functionally opens chromatin. PLoS One *11*.

Bartolomei, M.S. (2009). Genomic imprinting: Employing and avoiding epigenetic processes. Genes Dev. *23*, 2124–2133.

Bashtrykov, P., Ragozin, S., and Jeltsch, A. (2012a). Mechanistic details of the DNA recognition by the Dnmt1 DNA methyltransferase. FEBS Lett. *586*, 1821–1823.

Bashtrykov, P., Jankevicius, G., Smarandache, A., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2012b). Specificity of dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain. Chem. Biol. *19*, 572–578.

Bashtrykov, P., Jankevicius, G., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2014). The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism. J. Biol. Chem. *289*, 4106–4115.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243–247.

Baylin, S.B. (2005). DNA methylation and gene silencing in cancer. Nat. Clin. Pract. Oncol. *2*, S4–S11.

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. Nat. Rev. Cancer *11*, 726–734.

Beerli, R.R., and Barbas, C.F. (2002). Engineering polydactyl zinc-finger transcription factors. Nat. Biotechnol. *20*, 135–141.

Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F. (1998). Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc. Natl. Acad. Sci. U. S. A. *95*, 14628–14633.

Beerli, R.R., Dreier, B., and Barbas 3rd, C.F. (2000). Positive and negative regulation of endogenous genes by designed transcription factors. Proc Natl Acad Sci U S A *97*, 1495–1500.

Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, C.A., and Zoghbi, H.Y. (2009). Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. Hum. Mol. Genet. *18*, 2431–2442.

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

Bergman, Y., and Cedar, H. (2013). DNA methylation dynamics in health and disease. Nat Struct Mol Biol *20*, 274–281.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Bheda, P., and Schneider, R. (2014). Epigenetics reloaded: The single-cell revolution. Trends Cell Biol. *24*, 712–723.

Bienvenu, T., and Chelly, J. (2006). Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. Nat. Rev. Genet. *7*, 415–426.

Bischoff, R., and Schlüter, H. (2012). Amino acids: chemistry, functionality and selected non-enzymatic

post-translational modifications. J. Proteomics *75*, 2275–2296.

Bordeaux, J., Welsh, A.W., Agarwal, S., Killiam, E., Baquero, M.T., Hanna, J.A., Anagnostou, V.K., and Rimm, D.L. (2010). Antibody validation. Biotechniques *48*, 197–209.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nat. Genet. *42*, 1093–1100.

Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760–1764.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science *294*, 2536–2539.

Brero, A., Easwaran, H.P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., and Cardoso, M.C. (2005). Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. J. Cell Biol. *169*, 733–743.

Briggs, A.W., Rios, X., Chari, R., Yang, L., Zhang, F., Mali, P., and Church, G.M. (2012). Iterative capped assembly: Rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res. *40*.

Briones, V., and Muegge, K. (2012). The ghosts in the machine: DNA methylation and the mystery of differentiation. Biochim. Biophys. Acta - Gene Regul. Mech. *1819*, 757–762.

Britton, L.-M.P., Gonzales-Cope, M., Zee, B.M., and Garcia, B. a (2011). Breaking the histone code with quantitative mass spectrometry. Expert Rev. Proteomics *8*, 631–643.

Brown, K.D., and Robertson, K.D. (2007). DNMT1 knockout delivers a strong blow to genome stability and cell viability. Nat. Genet. *39*, 289–290.

Brzeski, J., and Jerzmanowski, A. (2003). Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. J. Biol. Chem. *278*, 823–828.

Buck-Koehntop, B.A., Stanfield, R.L., Ekiert, D.C., Martinez-Yamout, M.A., Dyson, H.J., Wilson, I.A., and Wright, P.E. (2012). Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. Proc. Natl. Acad. Sci. U. S. A. *109*, 15229–15234.

Buntru, A., Trepte, P., Klockmeier, K., Schnoegl, S., and Wanker, E.E. (2016). Current approaches toward quantitative mapping of the interactome. Front. Genet. *7*.

Burrage, J., Termanis, a., Geissner, a., Myant, K., Gordon, K., and Stancheva, I. (2012). SNF2 family ATPase LSH promotes phosphorylation of H2AX and efficient repair of DNA double-strand breaks in mammalian cells. J. Cell Sci. 5524–5534.

Burton, A., Muller, J., Tu, S., Padilla-Longoria, P., Guccione, E., and Torres-Padilla, M.E. (2013). Single-Cell Profiling of Epigenetic Modifiers Identifies PRDM14 as an Inducer of Cell Fate in the Mammalian Embryo. Cell Rep. *5*, 687–701.

Busnelli, M., Mauri, M., Parenti, M., and Chini, B. (2013). Analysis of GPCR dimerization using acceptor photobleaching resonance energy transfer techniques. Methods Enzymol. *521*, 311–327.

Campbell, K., and Casanova, J. (2016). A common framework for EMT and collective cell migration. Development *143*, 4291–4300.

Cano-Rodriguez, D., Gjaltema, R.A.F., Jilderda, L.J., Jellema, P., Dokter-Fokkens, J., Ruiters, M.H.J., and Rots, M.G. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. Nat. Commun. *7*, 12284.

Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Mol. Cell *41*, 67–81.

Canzio, D., Larson, A., and Narlikar, G.J. (2014). Mechanisms of functional promiscuity by HP1 proteins. Trends Cell Biol. *24*, 377–386.

Carouge, D., Host, L., Aunis, D., Zwiller, J., and Anglard, P. (2010). CDKL5 is a brain MeCP2 target gene regulated by DNA methylation. Neurobiol. Dis. *38*, 414–424.

Casas-Delucchi, C.S., Van Bemmel, J.G., Haase, S., Herce, H.D., Nowak, D., Meilinger, D., Stear, J.H., Leonhardt, H., and Cardoso, M.C. (2012). Histone hypoacetylation is required to maintain late replication

timing of constitutive heterochromatin. Nucleic Acids Res. *40*, 159–169.

Cermak, T., Doyle, E., and Christian, M. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids … *39*, e82.

Chahrour, M., Jung, S.Y., Shaw, C., Zhou, X., Wong, S.T.C., Qin, J., and Zoghbi, H.Y. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. Science *320*, 1224–1229.

Chait, B.T., Aebersold, R., Mann, M., Han, X., Jin, M., Breuker, K., McLafferty, F.W., Chait, B.T., Kent, S.B.H., Zubarev, R.A., et al. (2006). Chemistry. Mass spectrometry: bottom-up or top-down? Science *314*, 65–66.

Chakravarthy, S., Park, Y.-J., Chodaparambil, J., Edayathumangalam, R.S., and Luger, K. (2005). Structure and dynamic properties of nucleosome core particles. FEBS Lett. *579*, 895–898.

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature *453*, 544–547.

Chédin, F. (2011). The DNMT3 family of mammalian de novo DNA methyltransferases. Prog. Mol. Biol. Transl. Sci. *101*, 255–285.

Chen, B., and Huang, B. (2014). Imaging genomic elements in living cells using CRISPR/Cas9. Methods Enzymol. *546*, 337–354.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479–1491.

Chen, B., Guan, J., and Huang, B. (2016a). Imaging Specific Genomic DNA in Living Cells. Annu. Rev. Biophys. *45*, 1–23.

Chen, H., Puhl, H.L., Koushik, S. V, Vogel, S.S., and Ikeda, S.R. (2006). Measurement of FRET efficiency and ratio of donor to acceptor concentration in living cells. Biophys. J. *91*, L39-41.

Chen, M., Liu, S., Li, W., Zhang, Z., Zhang, X.-E.X., Zhang, X.-E.X., and Cui, Z. (2016b). Three-Fragment Fluorescence Complementation Coupled with Photoactivated Localization Microscopy for Nanoscale Imaging of Ternary Complexes. ACS Nano *10*, 8482–8490.

Chen, T., Ueda, Y., Xie, S., and Li, E. (2002). A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. J. Biol. Chem. *277*, 38746–38754.

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

Chen, T., Tsujimoto, N., and Li, E. (2004). The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. Mol Cell Biol *24*, 9048–9058.

Cheng, X. (1995). Structure and function of DNA methyltransferases. Annu. Rev. Biophys. Biomol. Struct. *24*, 293–318.

Cheng, X., Kumar, S., Posfai, J., Pflugrath, J.W., and Roberts, R.J. (1993). Crystal structure of the Hhal DNA methyltransferase complexed with S-adenosyl-l-methionine. Cell *74*, 299–307.

Chin, H.G., Patnaik, D., Estève, P.O., Jacobsen, S.E., and Pradhan, S. (2006). Catalytic properties and kinetic mechanism of human recombinant Lys-9 histone H3 methyltransferase SUV39H1: Participation of the chromodomain in enzymatic catalysis. Biochemistry *45*, 3272–3284.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. *24*, 132–141.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y. V, Chen, P.-Y.Y., Stroud, H., Yu, Y., Hetzel, J. a, Kuo, F., Kim, J., Cokus, S.J., et al. (2010). Relationship between nucleosome positioning and DNA methylation. Nature *466*, 388–392.

Christodoulou, J., Grimm, A., Maher, T., and Bennetts, B. (2003). RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. Hum. Mutat. *21*, 466–472.

Chuva De Sousa Lopes, S.M., Hayashi, K., Shovlin, T.C., Mifsud, W., Surani, M.A., and McLaren, A.

(2008). X chromosome activity in mouse XX primordial germ cells. PLoS Genet. *4*.

Ciernia, A.V., and LaSalle, J. (2016). The landscape of DNA methylation amid a perfect storm of autism aetiologies. Nat. Rev. Neurosci.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu. Rev. Biochem. *78*, 273–304.

Clouaire, T., de las Heras, J.I., Merusi, C., and Stancheva, I. (2010). Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. Nucleic Acids Res. *38*, 4620–4634.

Cohen, D.R.S., Matarazzo, V., Palmer, A.M., Tu, Y., Jeon, O.H., Pevsner, J., and Ronnett, G. V. (2003). Expression of MeCP2 in olfactory receptor neurons is developmentally regulated and occurs before synaptogenesis. Mol. Cell. Neurosci. *22*, 417–429.

Colantuoni, C., Jeon, O.H., Hyder, K., Chenchik, A., Khimani, A.H., Narayanan, V., Hoffman, E.P., Kaufmann, W.E., Naidu, S., and Pevsner, J. (2001). Gene expression profiling in postmortem Rett Syndrome brain: differential gene expression and patient classification. Neurobiol. Dis. *8*, 847–865.

Cole, K.C., McLaughlin, H.W., and Johnson, D.I. (2007). Use of bimolecular fluorescence complementation to study in vivo interactions between Cdc42p and Rdi1p of Saccharomyces cerevisiae. Eukaryot. Cell *6*, 378–387.

Collings, C.K., Waddell, P.J., and Anderson, J.N. (2013). Effects of DNA methylation on nucleosome stability. Nucleic Acids Res. *41*, 2918–2931.

Coltharp, C., Yang, X., and Xiao, J. (2014). Quantitative analysis of single-molecule superresolution images. Curr. Opin. Struct. Biol. *28*, 112–121.

Consortium, R.E., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Cosgrove, M.S., and Wolberger, C. (2005). How does the histone code work? Biochem. Cell Biol. *83*, 468–476.

Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat. Struct. Mol. Biol. *11*, 1037–1043.

Cusanovich, D. a, Daza, R., Adey, A., Pliner, H. a, Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Epigenetics. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914.

Damen, D., and Heumann, R. (2013). MeCP2 phosphorylation in the brain: From transcription to behavior. Biol. Chem. *394*, 1595–1605.

Danhier, P., Krishnamachary, B., Bharti, S., Kakkad, S., Mironchik, Y., and Bhujwalla, Z.M. (2015). Combining Optical Reporter Proteins with Different Half-lives to Detect Temporal Evolution of Hypoxia and Reoxygenation in Tumors. Neoplasia *17*, 871–881.

Darst, R.P., Pardo, C.E., Ai, L., Brown, K.D., and Kladde, M.P. (2010). Bisulfite sequencing of DNA. Curr. Protoc. Mol. Biol. *Chapter 7*.

David, C.J., Huang, Y.H., Chen, M., Su, J., Zou, Y., Bardeesy, N., Iacobuzio-Donahue, C.A., and Massagué, J. (2016). TGF-?? Tumor Suppression through a Lethal EMT. Cell *164*, 1015–1030.

Daxinger, L., and Whitelaw, E. (2010). Transgenerational epigenetic inheritance: More questions than answers. Genome Res. *20*, 1623–1628.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25*, 1010–1022.

Delgado, I.J., Kim, D.S., Thatcher, K.N., LaSalle, J.M., and Van den Veyver, I.B. (2006). Expression profiling of clonal lymphocyte cell cultures from Rett syndrome patients. BMC Med. Genet. *7*, 61.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

Demidov, V. V, Dokholyan, N. V, Witte-Hoffmann, C., Chalasani, P., Yiu, H.-W., Ding, F., Yu, Y., Cantor, C.R., and Broude, N.E. (2006). Fast complementation of split fluorescent protein triggered by DNA

hybridization. Proc. Natl. Acad. Sci. U. S. A. *103*, 2052–2056.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y., and Yan, N. (2012). Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. Science (80-.). *335*, 720–723.

Dennis, K., Fan, T., Geiman, T., Yan, Q., and Muegge, K. (2001). Lsh, a member of the SNF2 family, is required for genome-wide methylation. Genes Dev. *15*, 2940–2944.

Deuring, R., Fanti, L., Armstrong, J. a, Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M., et al. (2000). The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. Mol. Cell *5*, 355–365.

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

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. *6*, 227.

Dodget, J.E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., and Li, E. (2005). Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. J. Biol. Chem. *280*, 17986–17991.

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

Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science (80-.). *346*, 1258096–1258096.

Doyle, E.L., Booher, N.J., Standage, D.S., Voytas, D.F., Brendel, V.P., Vandyk, J.K., and Bogdanove, A.J. (2012). TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. Nucleic Acids Res. *40*.

Du, J., Johnson, L.M., Jacobsen, S.E., and Patel, D.J. (2015a). DNA methylation pathways and their crosstalk with histone methylation. Nat. Rev. Mol. Cell Biol. *16*, 519–532.

Du, Q., Luu, P.-L., Stirzaker, C., and Clark, S.J. (2015b). Methyl-CpG-binding domain proteins: readers of the epigenome. Epigenomics *7*, 1051–1073.

Dunican, D.S., Cruickshanks, H.A., Suzuki, M., Semple, C.A., Davey, T., Arceci, R.J., Greally, J., Adams, I.R., and Meehan, R.R. (2013). Lsh regulates LTR retrotransposon repression independently of Dnmt3b function. Genome Biol. *14*, R146.

Dutnall, R.N., and Ramakrishnan, V. (1997). Twists and turns of the nucleosome: tails without ends. Structure *5*, 1255–1259.

Ebert, D.H., Gabel, H.W., Robinson, N.D., Kastan, N.R., Hu, L.S., Cohen, S., Navarro, A.J., Lyst, M.J., Ekiert, R., Bird, A.P., et al. (2013). Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. Nature *499*, 341–345.

Edwards, J.R., O'Donnell, A.H., Rollins, R.A., Peckham, H.E., Lee, C., Milekic, M.H., Chanrion, B., Fu, Y., Su, T., Hibshoosh, H., et al. (2010). Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. Genome Res. *20*, 972–980.

Egger, G., Jeong, S., Escobar, S.G., Cortez, C.C., Li, T.W.H., Saito, Y., Yoo, C.B., Jones, P.A., and Liang, G. (2006). Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proc. Natl. Acad. Sci. U. S. A. *103*, 14080–14085.

Ehrlich, M. (2003). The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. Clin. Immunol. *109*, 17–28.

Eisen, J.A., Sweder, K.S., and Hanawalt, P.C. (1995). Evolution of the SNF2 family of proteins: Subfamilies with distinct sequences and functions. Nucleic Acids Res. *23*, 2715–2723.

Erdel, F., Schubert, T., Marth, C., Längst, G., and Rippe, K. (2010). Human ISWI chromatin-remodeling complexes sample nucleosomes via transient binding reactions and become immobilized at active sites. Proc. Natl. Acad. Sci. U. S. A. *107*, 19873–19878.

Fahy, J., Jeltsch, A., and Arimondo, P.B. (2012). DNA methyltransferase inhibitors in cancer: a chemical and therapeutic patent overview and selected clinical studies. Expert Opin. Ther. Pat. *22*, 1–16.

Fan, T., Schmidtmann, A., Xi, S., Briones, V., Zhu, H., Suh, H.C., Gooya, J., Keller, J.R., Xu, H., Roayaei,

J., et al. (2008). DNA hypomethylation caused by Lsh deletion promotes erythroleukemia development. Epigenetics *3*, 134–142.

Farh, K.K.-H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W.J., Beik, S., Shoresh, N., Whitton, H., Ryan, R.J.H., Shishkin, A.A., et al. (2015). Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature *518*, 337–343.

Farlik, M., Sheffield, N.C., Nuzzo, A., Datlinger, P., Schönegger, A., Klughammer, J., and Bock, C. (2015). Single-Cell DNA Methylome Sequencing and Bioinformatic Inference of Epigenomic Cell-State Dynamics. Cell Rep. *10*, 1386–1397.

Feil, R., and Fraga, M.F. (2012). Epigenetics and the environment: emerging patterns and implications. Nat. Rev. Genet. *13*, 97–109.

Felle, M., Hoffmeister, H., Rothammer, J., Fuchs, A., Exler, J.H., and Längst, G. (2011). Nucleosomes protect DNA from DNA methylation in vivo and in vitro. Nucleic Acids Res. *39*, 6956–6969.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D.Y., Barbosa, I.A.M., Kwon, J.S., Guan, Y., et al. (2013). An optimized microRNA backbone for effective single-copy RNAi. Cell Rep. *5*, 1704–1713.

Feng, J., Zhou, Y., Campbell, S.L., Le, T., Li, E., Sweatt, J.D., Silva, A.J., and Fan, G. (2010). Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. Nat. Neurosci. *13*, 423–430.

Ferguson-Smith, A.C. (2011). Genomic imprinting: the emergence of an epigenetic paradigm. Nat. Rev. Genet. *12*, 565–575.

Filion, G.J.P., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E., and Defossez, P.-A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. Mol. Cell. Biol. *26*, 169–181.

Finch, J.T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. U. S. A. *73*, 1897–1901.

Flaus, A., Martin, D.M.A., Barton, G.J., and Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic Acids Res. *34*, 2887–2905.

Fournier, A., Sasai, N., Nakao, M., and Defossez, P.A. (2012). The role of methyl-binding proteins in chromatin organization and epigenome maintenance. Brief. Funct. Genomics *11*, 251–264.

Fraga, M.F., Ballestar, E., Montoya, G., Taysavang, P., Wade, P.A., and Esteller, M. (2003). The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. Nucleic Acids Res. *31*, 1765–1774.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc. Natl. Acad. Sci. U. S. A. *89*, 1827–1831.

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

Fujimoto, S., Sugano, S.S., Kuwata, K., Osakabe, K., and Matsunaga, S. (2016). Visualization of specific repetitive genomic sequences with fluorescent TALEs in Arabidopsis thaliana. J. Exp. Bot. *67*, 6101–6110.

Fujita, N., Shimotake, N., Ohki, I., Chiba, T., Saya, H., Shirakawa, M., and Nakao, M. (2000). Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. Mol. Cell. Biol. *20*, 5107–5118.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003). Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. J. Biol. Chem. *278*, 24132–24138.

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

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003b). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J. Biol. Chem. *278*, 4035–4040.

Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett

syndrome. Nature *522*, 89–93.

Gaj, T., Gersbach, C.A., and Barbas, C.F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. *31*, 397–405.

Gal-Yam, E.N., Egger, G., Iniguez, L., Holster, H., Einarsson, S., Zhang, X., Lin, J.C., Liang, G., Jones, P.A., Tanay, A., et al. (2008). Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. Proc. Natl. Acad. Sci. U. S. A. *105*, 12979–12984.

Gao, X., Reid, M.A., Kong, M., and Locasale, J.W. (2016). Metabolic interactions with cancer epigenetics. Mol. Aspects Med.

Garcia, B.A. (2009). Mass spectrometric analysis of histone variants and post-translational modifications. Front. Biosci. *1*, 142–153.

Gardiner-Garden, M., and Frommer, M. (1987). CpG Islands in vertebrate genomes. J. Mol. Biol. *196*, 261–282.

Garg, A., Lohmueller, J.J., Silver, P.A., and Armel, T.Z. (2012). Engineering synthetic TAL effectors with orthogonal target sites. Nucleic Acids Res. *40*, 7584–7595.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U. S. A. *109*, E2579-86.

Ge, Y.Z., Pu, M.T., Gowher, H., Wu, H.P., Ding, J.P., Jeltsch, A., and Xu, G.L. (2004). Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. J. Biol. Chem. *279*, 25447–25454.

Geiman, T.M., and Muegge, K. (2000). Lsh, an SNF2/helicase family member, is required for proliferation of mature T lymphocytes. Proc. Natl. Acad. Sci. U. S. A. *97*, 4772–4777.

Geiman, T.M., Durum, S.K., and Muegge, K. (1998). Characterization of gene expression, genomic structure, and chromosomal localization of Hells (Lsh). Genomics *54*, 477–483.

Geiman, T.M., Tessarollo, L., Anver, M.R., Kopp, J.B., Ward, J.M., and Muegge, K. (2001). Lsh, a SNF2 family member, is required for normal murine development. Biochim. Biophys. Acta - Gen. Subj. *1526*, 211–220.

Gelbart, M.E., Bachman, N., Delrow, J., Boeke, J.D., and Tsukiyama, T. (2005). Genome-wide identification of Isw2 chromatin-remodeling targets by localization of a catalytically inactive mutant. Genes Dev. *19*, 942–954.

Georgel, P.T., Horowitz-Scherer, R.A., Adkins, N., Woodcock, C.L., Wade, P.A., and Hansen, J.C. (2003). Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. J. Biol. Chem. *278*, 32181–32188.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014a). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647–661.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Gloria, A., Torres, S.E., Stern-ginossar, N., Brandman, O., Whitehead, H., Doudna, J.A., et al. (2014b). CRISPR-Mediated Modular RNA-Guided Regualtion of Transcription in Eukaryotes. Cell *154*, 442–451.

Ginart, P., Kalish, J.M., Jiang, C.L., Yu, A.C., Bartolomei, M.S., and Raj, A. (2016). Visualizing allele-specific expression in single cells reveals epigenetic mosaicism in an H19 loss-of-imprinting mutant. Genes Dev. *30*, 567–578.

Golovko, D., Kedrin, D., Yilmaz, Ö.H., and Roper, J. (2015). Colorectal cancer models for novel drug discovery. Expert Opin. Drug Discov. *10*, 1217–1229.

Gomez, D., Shankman, L.S., Nguyen, A.T., and Owens, G.K. (2013). Detection of histone modifications at specific gene loci in single cells in histological sections. Nat. Methods *10*, 171–177.

Goyal, R., Reinhardt, R., and Jeltsch, A. (2006). Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. Nucleic Acids Res. *34*, 1182–1188.

Grigoryev, S.A., and Woodcock, C.L. (2012). Chromatin organization - The 30 nm fiber. Exp. Cell Res. *318*, 1448–1455.

Grigoryev, S.A., Bascom, G., Buckwalter, J.M., Schubert, M.B., Woodcock, C.L., and Schlick, T. (2016). Hierarchical looping of zigzag nucleosome chains in metaphase chromosomes. Proc. Natl. Acad. Sci.

U. S. A. *113*, 1238–1243.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951.

Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.P., Hu, B., et al. (2014a). Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell *15*, 447–458.

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., et al. (2015a). The transcriptome and DNA methylome landscapes of human primordial germ cells. Cell *161*, 1437–1452.

Guo, H., Zhu, P., Wu, X., Li, X., Wen, L., and Tang, F. (2013). Single-Cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res. *23*, 2126–2135.

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., et al. (2014b). The DNA methylation landscape of human early embryos. Nature *511*, 606–610.

Guo, H., Zhu, P., Guo, F., Li, X., Wu, X., Fan, X., Wen, L., and Tang, F. (2015b). Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. Nat. Protoc. *10*, 645–659.

Guo, J.U., Su, Y., Shin, J.H., Shin, J.H., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., et al. (2014c). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat Neurosci *17*, 215–222.

Guo, W., Chung, W.Y., Qian, M., Pellegrini, M., and Zhang, M.Q. (2014d). Characterizing the strand-specific distribution of non-CpG methylation in human pluripotent cells. Nucleic Acids Res. *42*, 3009–3016.

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015c). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature *517*, 640–644.

Guo, Y., Rebecchi, M., and Scarlata, S. (2005). Phospholipase Cbeta2 binds to and inhibits phospholipase Cdelta1. J. Biol. Chem. *280*, 1438–1447.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, a (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat. Genet. *27*, 322–326.

Guy, J., Cheval, H., Selfridge, J., and Bird, A. (2011). The Role of MeCP2 in the Brain. Annu. Rev. Cell Dev. Biol. *27*, 631–652.

Hahn, M.A., Wu, X., Li, A.X., Hahn, T., and Pfeifer, G.P. (2011). Relationship between gene body DNA methylation and intragenic H3K9ME3 and H3K36ME3 chromatin marks. PLoS One *6*.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. *117*, 15–23.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. Cell *139*, 945–956.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: The next generation. Cell *144*, 646–674.

Harlow, E., and Lane, D. (2006). Mounting samples in gelvatol or mowiol. CSH Protoc. *2006*.

Hashimoto, H., Horton, J.R., Zhang, X., and Cheng, X. (2009). UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications. Epigenetics *4*, 8–14.

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res. *40*, 4841–4849.

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

Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J.A. (2010). Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. Science (80-.). *329*, 1355–1358.

Hayashi-Takanaka, Y., Yamagata, K., Nozaki, N., and Kimura, H. (2009). Visualizing histone modifications in living cells: Spatiotemporal dynamics of H3 phosphorylation during interphase. J. Cell Biol. *187*, 781–790.

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

Heard, E., and Martienssen, R.A. (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. Cell *157*, 95–109.

Heckman, L.D., Chahrour, M.H., and Zoghbi, H.Y. (2014). Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. Elife *2014*.

Hellman, A., and Chess, A. (2007). Gene body-specific methylation on the active X chromosome. Science *315*, 1141–1143.

Helma, J., Cardoso, M.C., Muyldermans, S., and Leonhardt, H. (2015). Nanobodies and recombinant binders in cell biology. J. Cell Biol. *209*, 633–644.

Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol. Cell. Biol. *18*, 6538–6547.

Herceg, Z., and Vaissière, T. (2011). Epigenetic mechanisms and cancer an interface between the environment and the genome. Epigenetics *6*, 804–819.

Hermann, A., Goyal, R., and Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. J. Biol. Chem. *279*, 48350–48359.

Heyward, F.D., and Sweatt, J.D. (2015). DNA Methylation in Memory Formation Emerging Insights. Neurosci. *21*, 475–489.

Hiragami-Hamada, K., Soeroes, S., Nikolov, M., Wilkins, B., Kreuz, S., Chen, C., De La Rosa-Velázquez, I.A., Zenn, H.M., Kost, N., Pohl, W., et al. (2016). Dynamic and flexible H3K9me3 bridging via HP1β dimerization establishes a plastic state of condensed chromatin. Nat. Commun. *7*, 11310.

Hirata, E., and Kiyokawa, E. (2015). Future Perspective of Single-Molecule FRET Biosensors and Intravital FRET Microscopy. Biophys. J.

Hite, K.C., Adams, V.H., and Hansen, J.C. (2009). Recent advances in MeCP2 structure and function. Biochem. Cell Biol. *87*, 219–227.

Ho, K.L., McNae, I.W., Schmiedeberg, L., Klose, R.J., Bird, A.P., and Walkinshaw, M.D. (2008). MeCP2 Binding to DNA Depends upon Hydration at Methyl-CpG. Mol. Cell *29*, 525–531.

Hodges, E., Smith, A.D., Kendall, J., Xuan, Z., Ravi, K., Rooks, M., Zhang, M.Q., Ye, K., Bhattacharjee, A., Brizuela, L., et al. (2009). High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res. *19*, 1593–1605.

Horikoshi, N., Kumar, P., Sharma, G.G., Chen, M., Hunt, C.R., Westover, K., Chowdhury, S., and Pandita, T.K. (2013). Genome-wide distribution of histone H4 Lysine 16 acetylation sites and their relationship to gene expression. Genome Integr. *4*, 3.

Horlbeck, M.A., Witkowsky, L.B., Guglielmi, B., Replogle, J.M., Gilbert, L.A., Villalta, J.E., Torigoe, S.E., Tjian, R., and Weissman, J.S. (2016). Nucleosomes impede cas9 access to DNA in vivo and in vitro. Elife *5*.

Hota, S.K., and Bruneau, B.G. (2016). ATP-dependent chromatin remodeling during mammalian development. Development *143*, 2882–2897.

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. J. Biol. Chem. *175*, 315–332.

Hotta, A., Cheung, A.Y.L., Farra, N., Vijayaragavan, K., Séguin, C. a, Draper, J.S., Pasceri, P., Maksakova, I. a, Mager, D.L., Rossant, J., et al. (2009). Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. Nat. Methods *6*, 370–376.

Hsieh, T.H.S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108–119.

Hu, J.-L., Zhou, B.O., Zhang, R.-R., Zhang, K.-L., Zhou, J.-Q., and Xu, G.-L. (2009). The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. Proc. Natl. Acad. Sci. *106*, 22187–22192.

Huang, J., Fan, T., Yan, Q., Zhu, H., Fox, S., Issaq, H.J., Best, L., Gangi, L., Munroe, D., and Muegge, K. (2004). Lsh, an epigenetic guardian of repetitive elements. Nucleic Acids Res. *32*, 5019–5028.

Hudry, B., Viala, S., Graba, Y., and Merabet, S. (2011). Visualization of protein interactions in living Drosophila embryos by the bimolecular fluorescence complementation assay. BMC Biol. *9*, 5.

Hung, Y.L., Lee, H.J., Jiang, I., Lin, S.C., Lo, W.C., Lin, Y.J., and Sue, S.C. (2015). The First Residue of the PWWP Motif Modulates HATH Domain Binding, Stability, and Protein-Protein Interaction. Biochemistry *54*, 4063–4074.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011a). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science (80-.). *333*, 1300–1303.

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

Jacobs, S.A., Khorasanizadeh, S., Harp, J.M., Hanson, B.L., Timm, D.E., Bunick, G.J., Luger, K., Maeder, A.W., Richmond, R.K., Sargent, D.F., et al. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science *295*, 2080–2083.

Jaffe, A.E., Gao, Y., Deep-Soboslay, A., Tao, R., Hyde, T.M., Weinberger, D.R., and Kleinman, J.E. (2016). Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat.Neurosci. *19*, 40–47.

James, T.C., and Elgin, S.C. (1986). Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. Mol. Cell. Biol. *6*, 3862–3872.

Jares-Erijman, E. a, and Jovin, T.M. (2003). FRET imaging. Nat. Biotechnol. *21*, 1387–1395.

Jarvis, C.D., Geiman, T., Vila-Storm, M.P., Osipovich, O., Akella, U., Candeias, S., Nathan, I., Durum, S.K., and Muegge, K. (1996). A novel putative helicase produced in early murine lymphocytes. Gene *169*, 203–207.

Jeddeloh, J. a, Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. *22*, 94–97.

Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem *3*, 274–293.

Jeltsch, A. (2006). Molecular enzymology of mammalian DNA methyltransferases. In Current Topics in Microbiology and Immunology, pp. 203–225.

Jeltsch, A., and Jurkowska, R.Z. (2014). New concepts in DNA methylation. Trends Biochem. Sci. *39*, 310–318.

Jeltsch, A., and Jurkowska, R.Z. (2016). Allosteric control of mammalian DNA methyltransferases - a new regulatory paradigm. Nucleic Acids Res. gkw723.

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

Jiang, F., and Doudna, J.A. (2015). The structural biology of CRISPR-Cas systems. Curr. Opin. Struct. Biol. *30*, 100–111.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (80-.). *337*, 816–822.

Johnson, T.A., Elbi, C., Parekh, B.S., Hager, G.L., and John, S. (2008). Chromatin remodeling complexes interact dynamically with a glucocorticoid receptor-regulated promoter. Mol. Biol. Cell *19*, 3308–3322.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Jones, P. a, and Liang, G. (2009). Rethinking how DNA methylation patterns are maintained. Nat. Rev. Genet. *10*, 805–811.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. *19*, 187–191.

Jørgensen, H.F., Ben-Porath, I., and Bird, A.P. (2004). Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. Mol. Cell. Biol. *24*, 3387–3395.

Jørgensen, H.F., Adie, K., Chaubert, P., and Bird, A.P. (2006). Engineering a high-affinity methyl-CpG-binding protein. Nucleic Acids Res. *34*.

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

Jurkowska, R.Z., and Jeltsch, A. (2016). Mechanisms and Biological Roles of DNA Methyltransferases and DNA Methylation: From Past Achievements to Future Challenges. Adv. Exp. Med. Biol. *945*, 1–17.

Jurkowska, R.Z., Jurkowski, T.P., and Jeltsch, A. (2011). Structure and function of mammalian DNA methyltransferases. Chembiochem *12*, 206–222.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., and Saitou, M. (2012). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. EMBO J. *32*, 340–353.

Kalish, J.M., Jiang, C., and Bartolomei, M.S. (2014). Epigenetics and imprinting in human disease. Int. J. Dev. Biol. *58*, 291–298.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900–903.

Karch, K.R., DeNizio, J.E., Black, B.E., and Garcia, B.A. (2013). Identification and interrogation of combinatorial histone modifications. Front. Genet. *4*.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., et al. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mescs. Cell Stem Cell *8*, 676–687.

Karpova, T.S., Baumann, C.T., He, L., Wu, X., Grammer, A., Lipsky, P., Hager, G.L., and McNally, J.G. (2003). Fluorescence resonance energy transfer from cyan to yellow fluorescent protein detected by acceptor photobleaching using confocal microscopy and a single laser. J. Microsc. *209*, 56–70.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and Sasaki, H. (2007). Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum. Mol. Genet. *16*, 2272–2280.

Katz, D.M., Bird, A., Coenraads, M., Gray, S.J., Menon, D.U., Philpot, B.D., and Tarquinio, D.C. (2016). Rett Syndrome: Crossing the Threshold to Clinical Translation. Trends Neurosci. *39*, 100–113.

Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P., and Jones, P.A. (2012). Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res. *22*, 2497–2506.

Kerppola, T.K. (2006). Visualization of molecular interactions by fluorescence complementation. Nat. Rev. Mol. Cell Biol. *7*, 449–456.

Kerppola, T.K. (2008). Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. Annu. Rev. Biophys. *37*, 465–487.

Keshet, I., Schlesinger, Y., Farkash, S., Rand, E., Hecht, M., Segal, E., Pikarski, E., Young, R. a, Niveleau, A., Cedar, H., et al. (2006). Evidence for an instructive mechanism of de novo methylation in cancer cells. Nat. Genet. *38*, 149–153.

Khrapunov, S., Warren, C., Cheng, H., Berko, E.R., Greally, J.M., and Brenowitz, M. (2014). Unusual characteristics of the DNA binding domain of epigenetic regulatory protein MeCP2 determine its binding specificity. Biochemistry *53*, 3379–3391.

Kim, J., and Kim, H. (2012). Recruitment and biological consequences of histone modification of H3K27me3 and H3K9me3. ILAR J. *53*, 232–239.

Kim, S.H., Park, J.H.J., Choi, M.C., Kim, H.P., Park, J.H.J., Jung, Y., Lee, J.H., Oh, D.Y., Im, S.A., Bang, Y.J., et al. (2007). Zinc-fingers and homeoboxes 1 (ZHX1) binds DNA methyltransferase (DNMT) 3B to

enhance DNMT3B-mediated transcriptional repression. Biochem. Biophys. Res. Commun. *355*, 318–323.

Kim, Y., Sharov, A. a., McDole, K., Cheng, M., Hao, H., Fan, C.-M., Gaiano, N., Ko, M.S.H., and Zheng, Y. (2011). Mouse B-Type Lamins Are Required for Proper Organogenesis But Not by Embryonic Stem Cells. Science (80-.). *334*, 1706–1710.

Kimura, H. (2013). Histone modifications for human epigenome analysis. J. Hum. Genet. *58*, 439–445.

Kimura, H., and Yamagata, K. (2014). Visualization of epigenetic modifications in preimplantation embryos. In Nuclear Reprogramming: Methods and Protocols: Second Edition, pp. 127–147.

Kinde, B., Gabel, H.W., Gilbert, C.S., Griffith, E.C., and Greenberg, M.E. (2015). Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. Proc. Natl. Acad. Sci. U. S. A. *112*, 6800–6806.

Kishi, N., and Macklis, J.D. (2004). MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. Mol. Cell. Neurosci. *27*, 306–321.

Kishikawa, S., Murata, T., Ugai, H., Yamazaki, T., and Yokoyama, K.K. (2003). Control elements of Dnmt1 gene are regulated in cell-cycle dependent manner. Nucleic Acids Res. Suppl. 307–308.

Klose, R.J., Sarraf, S.A., Schmiedeberg, L., McDermott, S.M., Stancheva, I., and Bird, A.P. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Mol. Cell *19*, 667–678.

Knight, S.C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L.B., Bosanac, L., Zhang, E.T., El Beheiry, M., Masson, J.-B., Dahan, M., et al. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. Science *350*, 823–826.

Kodama, Y., and Hu, C.D. (2012). Bimolecular fluorescence complementation (BiFC): A 5-year update and future perspectives. Biotechniques *53*, 285–298.

Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001). The Ski Protein Family Is Required for MeCP2-mediated Transcriptional Repression. J. Biol. Chem. *276*, 34115–34121.

König, P., Krasteva, G., Tag, C., König, I.R., Arens, C., and Kummer, W. (2006). FRET-CLSM and double-labeling indirect immunofluorescence to detect close association of proteins in tissue sections. Lab. Invest. *86*, 853–864.

Kooistra, S.M., and Helin, K. (2012). Molecular mechanisms and potential functions of histone demethylases. Nat. Rev. Mol. Cell Biol. *13*, 297–311.

Kota, S.K., and Feil, R. (2010). Epigenetic Transitions in Germ Cell Development and Meiosis. Dev. Cell *19*, 675–686.

Koushik, S. V, Chen, H., Thaler, C., Puhl, H.L., and Vogel, S.S. (2006). Cerulean, Venus, and VenusY67C FRET reference standards. Biophys. J. *91*, L99–L101.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Kriaucionis, S., Paterson, A., Curtis, J., Guy, J., MacLeod, N., and Bird, A. (2006). Gene Expression Analysis Exposes Mitochondrial Abnormalities in a Mouse Model of Rett Syndrome. Mol. Cell. Biol. *26*, 5033–5042.

Krokan, H.E., Nilsen, H., Skorpen, F., Otterlei, M., and Slupphaug, G. (2000). Base excision repair of DNA in mammalian cells. FEBS Lett. *476*, 73–77.

Kungulovski, G., and Jeltsch, A. (2016). Epigenome Editing: State of the Art, Concepts, and Perspectives. Trends Genet. *32*, 101–113.

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

Kungulovski, G., Mauser, R., and Jeltsch, A. (2015a). Affinity reagents for studying histone modifications & guidelines for their quality control. Epigenomics *7*, 1185–1196.

Kungulovski, G., Mauser, R., and Jeltsch, A. (2015b). Affinity reagents for studying histone modifications & guidelines for their quality control. Epigenomics *7*, 1185–1196.

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.

Kurihara, Y., Kawamura, Y., Uchijima, Y., Amamo, T., Kobayashi, H., Asano, T., and Kurihara, H. (2008). Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. Dev. Biol. *313*, 335–346.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat Biotechnol *32*, 677–683.

Kusumi, A., Tsunoyama, T.A., Hirosawa, K.M., Kasai, R.S., and Fujiwara, T.K. (2014). Tracking single molecules at work in living cells. Nat. Chem. Biol. *10*, 524–532.

De La Fuente, R., Baumann, C., Fan, T., Schmidtmann, A., Dobrinski, I., and Muegge, K. (2006). Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells. Nat Cell Biol *8*, 1448–1454.

Lakowicz, J.R. (2006). Principles of fluorescence spectroscopy.

Landan, G., Cohen, N.M., Mukamel, Z., Bar, A., Molchadsky, A., Brosh, R., Horn-Saban, S., Zalcenstein, D.A., Goldfinger, N., Zundelevich, A., et al. (2012). Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. Nat. Genet. *44*, 1207–1214.

Längst, G., and Manelyte, L. (2015). Chromatin remodelers: From function to dysfunction. Genes (Basel). *6*, 299–324.

Laurent, B.C., Yang, X., and Carlson, M. (1992). An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family. Mol. Cell. Biol. *12*, 1893–1902.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Sung, K.W.K., Rigoutsos, I., Loring, J., et al. (2010). Dynamic changes in the human methylome during differentiation. Genome Res. *20*, 320–331.

Lee, J.S., and Shilatifard, A. (2007). A site to remember: H3K36 methylation a mark for histone deacetylation. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *618*, 130–134.

Lee, D.W., Zhang, K., Ning, Z.Q., Raabe, E.H., Tintner, S., Wieland, R., Wilkins, B.J., Kim, J.M., Blough, R.I., and Arceci, R.J. (2000). Proliferation-associated SNF2-like gene (PASG): A SNF2 family member altered in leukemia. Cancer Res. *60*, 3612–3622.

Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr. Biol. *13*, 1192–1200.

Leroy, G., Dimaggio, P.A., Chan, E.Y., Zee, B.M., Blanco, M.A., Bryant, B., Flaniken, I.Z., Liu, S., Kang, Y., Trojer, P., et al. (2013). A quantitative atlas of histone modification signatures from human cancer cells. Epigenetics Chromatin *6*, 1.

Levsky, J.M., Shenoy, S.M., Pezo, R.C., Singer, R.H., Alwine, J.C., Liang, P., Pardee, A.B., Velculescu, V.E., Zhang, L., Vogelstein, B., et al. (2002). Single-cell gene expression profiling. Science *297*, 836–840.

Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., and Sorek, R. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature *520*, 505–510.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to Methylated DNA. Cell *69*, 905–914.

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

Li, Y., and Tollefsbol, T.O. (2011). DNA methylation detection: Bisulfite genomic sequencing analysis. Methods Mol. Biol. *791*, 11–21.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, H., Rauch, T., Chen, Z.X., Szabó, P.E., Riggs, A.D., and Pfeifer, G.P. (2006). The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. J. Biol. Chem. *281*, 19489–19500.

Li, H., Fischle, W., Wang, W., Duncan, E.M., Liang, L., Murakami-Ishibe, S., Allis, C.D., and Patel, D.J. (2007). Structural Basis for Lower Lysine Methylation State-Specific Readout by MBT Repeats of L3MBTL1 and an Engineered PHD Finger. Mol. Cell *28*, 677–691.

Li, L., Chen, B.F., and Chan, W.Y. (2015a). An epigenetic regulator: Methyl-CpG-binding domain protein 1 (MBD1). Int. J. Mol. Sci. *16*, 5125–5140.

Li, M.-W., Zhou, L., and Lam, H.-M. (2015b). Paraformaldehyde Fixation May Lead to Misinterpretation of the Subcellular Localization of Plant High Mobility Group Box Proteins. PLoS One *10*, e0135033.

Li, W., Chen, P., Yu, J., Dong, L., Liang, D., Feng, J., Yan, J., Wang, P.-Y., Li, Q., Zhang, Z., et al. (2016). FACT Remodels the Tetranucleosomal Unit of Chromatin Fibers for Gene Transcription. Mol. Cell *64*, 120–133.

Li, Y., Miyanari, Y., Shirane, K., Nitta, H., Kubota, T., Ohashi, H., Okamoto, A., and Sasaki, H. (2013). Sequence-specific microscopic visualization of DNA methylation status at satellite repeats in individual cell nuclei and chromosomes. Nucleic Acids Res. *41*.

Lieberman-aiden, E., Berkum, N.L. Van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science (80-.). *326*, 289–293.

Lima, S.C.S., Hernandez-Vargas, H., and Herceg, Z. (2010). Epigenetic signatures in cancer: Implications for the control of cancer in the clinic. Curr. Opin. Mol. Ther. *12*, 316–324.

Lin, C.W., Jao, C.Y., and Ting, A.Y. (2004). Genetically Encoded Fluorescent Reporters of Histone Methylation in Living Cells. J. Am. Chem. Soc. *126*, 5982–5983.

Lindenburg, L., and Merkx, M. (2014). Engineering genetically encoded FRET sensors. Sensors (Switzerland) *14*, 11691–11713.

Lindhout, B.I., Fransz, P., Tessadori, F., Meckel, T., Hooykaas, P.J.J., and van der Zaal, B.J. (2007). Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. Nucleic Acids Res. *35*.

Lister, R., and Ecker, J.R. (2009). Finding the fifth base: Genome-wide sequencing of cytosine methylation. Genome Res. *19*, 959–966.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Lister, R., Mukamel, E.A.A., Nery, J.R.R., Urich, M., Puddifoot, C.A.A., Johnson, N.D.D., Lucero, J., Huang, Y., Dwork, A.J.J., Schultz, M.D.D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science (80-.). *341*, 1237905.

Liu, C., Teng, Z.Q., McQuate, A.L., Jobe, E.M., Christ, C.C., von Hoyningen-Huene, S.J., Reyes, M.D., Polich, E.D., Xing, Y., Li, Y., et al. (2013a). An Epigenetic Feedback Regulatory Loop Involving MicroRNA-195 and MBD1 Governs Neural Stem Cell Differentiation. PLoS One *8*.

Liu, Q., Segal, D.J., Ghiara, J.B., and Barbas, C.F. (1997). Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. Proc. Natl. Acad. Sci. U. S. A. *94*, 5525–5530.

Liu, Y., Toh, H., Sasaki, H., Zhang, X., and Cheng, X. (2012). An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev. *26*, 2374–2379.

Liu, Y., Zhang, X., Blumenthal, R.M., and Cheng, X. (2013b). A common mode of recognition for methylated CpG. Trends Biochem. Sci. *38*, 177–183.

Lombardi, L.M., Baker, S.A., and Zoghbi, H.Y. (2015). MECP2 disorders: From the clinic to mice and back. J. Clin. Invest. *125*, 2914–2923.

Longo, P.A., Kavran, J.M., Kim, M.-S., and Leahy, D.J. (2013). Transient Mammalian Cell Transfection with Polyethylenimine (PEI). pp. 227–240.

Lu, J., and Gilbert, D.M. (2007). Proliferation-dependent and cell cycle-regulated transcription of mouse pericentric heterochromatin. J. Cell Biol. *179*, 411–421.

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

Lyst, M.J., Ekiert, R., Ebert, D.H., Merusi, C., Nowak, J., Selfridge, J., Guy, J., Kastan, N.R., Robinson, N.D., de Lima Alves, F., et al. (2013). Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. Nat. Neurosci. *16*, 898–902.

Lyst, M.J., Connelly, J., Merusi, C., and Bird, A. (2016). Sequence-specific DNA binding by AT-hook motifs in MeCP2. FEBS Lett. *590*, 2927–2933.

Ma, H., Reyes-Gutierrez, P., and Pederson, T. (2013). Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. Proc. Natl. Acad. Sci. U. S. A. *110*, 21048–21053.

Ma, H., Naseri, A., Reyes-Gutierrez, P., Wolfe, S.A., Zhang, S., and Pederson, T. (2015). Multicolor CRISPR labeling of chromosomal loci in human cells. Proc. Natl. Acad. Sci. U. S. A. *112*, 3002–3007.

Machado, A.C.D., Zhou, T., Rao, S., Goel, P., Rastogi, C., Lazarovici, A., Bussemaker, H.J., and Rohs, R. (2015). Evolving insights on how cytosine methylation affects protein-DNA binding. Brief. Funct. Genomics *14*, 61–73.

Maezawa, I., and Jin, L.-W. (2010). Rett syndrome microglia damage dendrites and synapses by the elevated release of glutamate. J. Neurosci. *30*, 5346–5356.

Maezawa, I., Swanberg, S., Harvey, D., LaSalle, J.M., and Jin, L.-W. (2009). Rett syndrome astrocytes are abnormal and spread MeCP2 deficiency through gap junctions. J. Neurosci. *29*, 5051–5061.

Maison, C., Bailly, D., Peters, A.H.F.M., Quivy, J.-P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. Nat. Genet. *30*, 329–334.

Maison, C., Bailly, D., Roche, D., Montes de Oca, R., Probst, A. V, Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J.-P., et al. (2011). SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. Nat. Genet. *43*, 220–227.

Mak, A.N.-S., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. Science *335*, 716–719.

Mandell, J.G., and Barbas, C.F. (2006). Zinc Finger Tools: Custom DNA-binding domains for transcription factors and nucleases. Nucleic Acids Res. *34*.

Mann, M., and Jensen, O.N. (2003). Proteomic analysis of post-translational modifications. Nat. Biotechnol. *21*, 255–261.

Manohar, M., Mooney, A.M., North, J.A., Nakkula, R.J., Picking, J.W., Edon, A., Fishel, R., Poirier, M.G., and Ottesen, J.J. (2009). Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. J. Biol. Chem. *284*, 23312–23321.

Manukyan, M., and Singh, P.B. (2014). Epigenome rejuvenation: HP1β mobility as a measure of pluripotent and senescent chromatin ground states. Sci. Rep. *4*, 4789.

Marchetto, M.C.N., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of rett syndrome using human induced pluripotent stem cells. Cell *143*, 527–539.

Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R., Scala, E., Longo, I., Grosso, S., Pescucci, C., et al. (2005). CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. Hum. Mol. Genet. *14*, 1935–1946.

Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. Nature *526*, 55–61.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell Biol. *6*, 838–849.

Mattout, A., Aaronson, Y., Sailaja, B.S., Raghu Ram, E. V, Harikumar, A., Mallm, J.-P., Sim, K.H., Nissim-Rafinia, M., Supper, E., Singh, P.B., et al. (2015). Heterochromatin Protein 1β (HP1β) has distinct functions and distinct nuclear distribution in pluripotent versus differentiated cells. Genome Biol. *16*, 213.

Meehan, R., Lewis, J.D., and Bird, A.P. (1992). Characterization of MECP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Res. *20*, 5085–5092.

Meehan, R.R., Pennings, S., and Stancheva, I. (2001). Lashings of DNA methylation, forkfuls of chromatin remodeling. Genes Dev. *15*, 3231–3236.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell *151*, 1417–1430.

Mentch, S.J., and Locasale, J.W. (2016). One-carbon metabolism and epigenetics: Understanding the specificity. Ann. N. Y. Acad. Sci. *1363*, 91–98.

Mentch, S.J., Mehrmohamadi, M., Huang, L., Liu, X., Gupta, D., Mattocks, D., Gómez Padilla, P., Ables, G., Bamman, M.M., Thalacker-Mercer, A.E., et al. (2015). Histone Methylation Dynamics and Gene Regulation Occur through the Sensing of One-Carbon Metabolism. Cell Metab. *22*, 861–873.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. *28*, 812–828.

Métivier, R., Gallais, R., Tiffoche, C., Le Péron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G., et al. (2008). Cyclical DNA methylation of a transcriptionally active promoter. Nature *452*, 45–50.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Min, J., Zhang, X., Cheng, X., Grewal, S.I.S., and Xu, R.-M. (2002). Structure of the SET domain histone lysine methyltransferase Clr4. Nat. Struct. Biol. *9*, 828–832.

Mishin, A.S., Belousov, V. V., Solntsev, K.M., and Lukyanov, K.A. (2015). Novel uses of fluorescent proteins. Curr. Opin. Chem. Biol. *27*, 1–9.

Miyanari, Y., Ziegler-Birling, C., and Torres-Padilla, M.-E. (2013). Live visualization of chromatin dynamics with fluorescent TALEs. Nat. Struct. Mol. Biol. *20*, 1321–1324.

Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet. *7*, 45–49.

Morell, M., Ventura, S., and Avilés, F.X. (2009). Protein complementation assays: Approaches for the in vivo analysis of protein interactions. FEBS Lett. *583*, 1684–1691.

Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. Nat Rev Genet *17*, 693–703.

Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S., et al. (2015). In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. Elife *2015*.

Mozzetta, C., Boyarchuk, E., Pontis, J., and Ait-Si-Ali, S. (2015). Sound of silence: the properties and functions of repressive Lys methyltransferases. Nat. Rev. Mol. Cell Biol. *16*, 499–513.

Muchardt, C., Guillemé, M., Seeler, J.S., Trouche, D., Dejean, A., and Yaniv, M. (2002). Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1α. EMBO Rep. *3*, 975–981.

Muegge, K. (2005). Lsh, a guardian of heterochromatin at repeat elements. Biochem. Cell Biol. *83*, 548–554.

Van Munster, E.B., Kremers, G.J., Adjobo-Hermans, M.J.W., and Gadella, T.W.J. (2005). Fluorescence resonance energy transfer (FRET) measurement by gradual acceptor photobleaching. J. Microsc. *218*, 253–262.

Muramatsu, D., Kimura, H., Kotoshiba, K., Tachibana, M., and Shinkai, Y. (2016). Pericentric H3K9me3 formation by HP1 interaction-defective histone methyltransferase Suv39h1. Cell Struct. Funct.

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

Myant, K., and Stancheva, I. (2008). LSH cooperates with DNA methyltransferases to repress transcription. Mol. Cell. Biol. *28*, 215–226.

Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature *502*, 59–64.

Nakao, M., Matsui, S.I., Yamamoto, S., Okumura, K., Shirakawa, M., and Fujita, N. (2001). Regulation of transcription and chromatin by methyl-CpG binding protein MBD1. In Brain and Development, p.

Nakaoka, S., Sasaki, K., Ito, A., Nakao, Y., and Yoshida, M. (2016). A Genetically Encoded FRET Probe to Detect Intranucleosomal Histone H3K9 or H3K14 Acetylation Using BRD4, a BET Family Member. ACS Chem. Biol. *11*, 729–733.

Nan, X., Tate, P., Li, E., and Bird, A. (1996). DNA methylation specifies chromosomal localization of MeCP2. Mol. Cell. Biol. *16*, 414–421.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386–389.

Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. Mol. Cell *36*, 153–163.

Ng, H.H., Jeppesen, P., and Bird, a (2000). Active repression of methylated genes by the chromosomal protein MBD1. Mol. Cell. Biol. *20*, 1394–1406.

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

Nikitina, T., Ghosh, R.P., Horowitz-Scherer, R.A., Hansen, J.C., Grigoryev, S.A., and Woodcock, C.L. (2007a). MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. J. Biol. Chem. *282*, 28237–28245.

Nikitina, T., Shi, X., Ghosh, R.P., Horowitz-Scherer, R. a, Hansen, J.C., and Woodcock, C.L. (2007b). Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin. Mol. Cell. Biol. *27*, 864–877.

Noh, K.M., Wang, H., Kim, H.R., Wenderski, W., Fang, F., Li, C.H., Dewell, S., Hughes, S.H., Melnick, A.M., Patel, D.J., et al. (2015). Engineering of a Histone-Recognition Domain in Dnmt3a Alters the Epigenetic Landscape and Phenotypic Features of Mouse ESCs. Mol. Cell *59*, 89–103.

O'Geen, H., Henry, I.M., Bhakta, M.S., Meckler, J.F., and Segal, D.J. (2015). A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. Nucleic Acids Res. *43*, 3389–3404.

Ohashi, K., and Mizuno, K. (2014). A novel pair of split venus fragments to detect protein-protein interactions by in vitro and in vivo bimolecular fluorescence complementation assays. Methods Mol. Biol. *1174*, 247–262.

Ohki, I., Shimotake, N., Fujita, N., Jee, J.G., Ikegami, T., Nakao, M., and Shirakawa, M. (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. Cell *105*, 487–497.

Okano, M., Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat. Genet. *19*, 219–220.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

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

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. Nat. Rev. Mol. Cell Biol. *4*, 809–814.

Olsen, J. V, Ong, S.-E., and Mann, M. (2004). Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. Mol. Cell. Proteomics *3*, 608–614.

Ooi, S.K.T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714–717.

van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. *12*, 479–492.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. EMBO Rep *10*, 1235–1241.

Patel, D.J. (2016). A structural perspective on readout of epigenetic histone and DNA methylation marks. Cold Spring Harb. Perspect. Biol. *8*.

Pavlova, N.N., and Thompson, C.B. (2016). The Emerging Hallmarks of Cancer Metabolism. Cell Metab. *23*, 27–47.

Peddada, S., Yasui, D.H., and LaSalle, J.M. (2006). Inhibitors of differentiation (ID1, ID2, ID3 and ID4) genes are neuronal targets of MeCP2 that are elevated in Rett syndrome. Hum. Mol. Genet. *15*, 2003–2014.

Peters, A.H.F.M., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323–337.

Peters, A.H.F.M., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A.H.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003). Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin. Mol. Cell *12*, 1577–1589.

Pikaard, C.S. (2013). Methylating the DNA of the most repressed: Special access required. Mol. Cell *49*, 1021–1022.

Pinney, S. (2014). Mammalian Non-CpG Methylation: Stem Cells and Beyond. Biology (Basel). *3*, 739–751.

Piston, D.W.D.W., and Kremers, G.-J.G.J. (2007). Fluorescent protein FRET: the good, the bad and the ugly. Trends Biochem. Sci. *32*, 407–414.

Plazas-Mayorca, M.D., Zee, B.M., Young, N.L., Fingerman, I.M., LeRoy, G., Briggs, S.D., and Garcia, B.A. (2009). One-pot shotgun quantitative mass spectrometry characterization of histones. J. Proteome Res. *8*, 5367–5374.

Portella, G., Battistini, F., and Orozco, M. (2013). Understanding the Connection between Epigenetic DNA Methylation and Nucleosome Positioning from Computer Simulations. PLoS Comput. Biol. *9*.

Prokhortchouk, A., Hendrich, B., Jørgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev. *15*, 1613–1618.

Purdy, M.M., Holz-Schietinger, C., and Reich, N.O. (2010). Identification of a second DNA binding site in human DNA methyltransferase 3A by substrate inhibition and domain deletion. Arch. Biochem. Biophys. *498*, 13–22.

Qin, S., and Min, J. (2014). Structure and function of the nucleosome-binding PWWP domain. Trends Biochem. Sci. *39*, 536–547.

Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. *9*, 217–224.

Quenet, D., McNally, J.G., and Dalal, Y. (2012). Through thick and thin: the conundrum of chromatin fibre folding in vivo. EMBO Rep. *13*, 943–944.

Rader, C. (2009). Overview on concepts and applications of fab antibody fragments. Curr. Protoc. Protein Sci.

Ragione, F. Della, Vacca, M., Fioriniello, S., Pepe, G., and Esposito, M.D. (2016). MECP2 , a multi-talented modulator of chromatin architecture. Brief. Funct. Genomics 1–12.

Rajakumara, E., Law, J.A., Simanshu, D.K., Voigt, P., Johnson, L.M., Reinberg, D., Patel, D.J., and Jacobsen, S.E. (2011). A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo. Genes Dev. *25*, 137–152.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn,

A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

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

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

Reddington, J.P., Pennings, S., and Meehan, R.R. (2013). Non-canonical functions of the DNA methylome in gene regulation. Biochem. J. *451*, 13–23.

Reese, B.E., Bachman, K.E., Baylin, S.B., and Rountree, M.R. (2003). The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1. Mol. Cell. Biol. *23*, 3226–3236.

Reinisch, K.M., Chen, L., Verdine, G.L., and Lipscomb, W.N. (1995). The crystal structure of HaeIII methyltransferase convalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. Cell *82*, 143–153.

Ren, J., Briones, V., Barbour, S., Yu, W., Han, Y., Terashima, M., and Muegge, K. (2015). The ATP binding site of the chromatin remodeling homolog Lsh is required for nucleosome density and de novo DNA methylation at repeat sequences. Nucleic Acids Res. *43*, 1444–1455.

Reute, G., and Spierer, P. (1992). Position effect variegation and chromatin proteins. BioEssays *14*, 605–612.

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., and Joung, J.K. (2012). FLASH assembly of TALENs for high-throughput genome editing. Nat. Biotechnol. *30*, 460–465.

Rhee, I., Bachman, K.E., Park, B.H., Jair, K., Yen, R.-W.C., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W., et al. (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature *416*, 552–556.

Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145–1158.

Richmond, E., and Peterson, C.L. (1996). Functional analysis of the DNA-stimulated ATPase domain of yeast SW12/SNF2. Nucleic Acids Res. *24*, 3685–3692.

Riggs, A.D., and Xiong, Z. (2004). Methylation and epigenetic fidelity. Proc. Natl. Acad. Sci. *101*, 4–5.

Rippe, K., Schrader, A., Riede, P., Strohner, R., Lehmann, E., and Längst, G. (2007). DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. Proc. Natl. Acad. Sci. U. S. A. *104*, 15635–15640.

Roberts, R.J. (1995). On base flipping. Cell *82*, 9–12.

Roeßler, M., Rollinger, W., Palme, S., Hagmann, M.L., Berndt, P., Engel, A.M., Schneidinger, B., Pfeffer, M., Andres, H., Karl, J., et al. (2005). Identification of nicotinamide N-methyltransferase as a novel serum tumor marker for colorectal cancer. Clin. Cancer Res. *11*, 6550–6557.

Rondelet, G., Dal Maso, T., Willems, L., and 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.

Ropero, S., and Esteller, M. (2007). The role of histone deacetylases (HDACs) in human cancer. Mol. Oncol. *1*, 19–25.

Roth, M., and Jeltsch, a (2000). Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. Biol. Chem. *381*, 269–272.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. Biochim. Biophys. Acta - Gene Regul. Mech. *1839*, 627–643.

Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., Backmann, N., Conrath, K., Muyldermans, S., Cardoso, M.C., et al. (2006). Targeting and tracing antigens in live cells with fluorescent nanobodies. Nat. Methods *3*, 887–889.

Rowbotham, S.P., Barki, L., Neves-Costa, A., Santos, F., Dean, W., Hawkes, N., Choudhary, P., Will, W.R., Webster, J., Oxley, D., et al. (2011). Maintenance of Silent Chromatin through Replication Requires SWI/SNF-like Chromatin Remodeler SMARCAD1. Mol. Cell *42*, 285–296.

Rube, H.T., Lee, W., Hejna, M., Chen, H., Yasui, D.H., Hess, J.F., LaSalle, J.M., Song, J.S., and Gong, Q. (2016). Sequence features accurately predict genome-wide MeCP2 binding in vivo. Nat. Commun. *7*, 11025.

Ryan, D.P., and Owen-Hughes, T. (2011). Snf2-family proteins: Chromatin remodellers for any occasion. Curr. Opin. Chem. Biol. *15*, 649–656.

Saitou, M., and Yamaji, M. (2012). Primordial germ cells in mice. Cold Spring Harb. Perspect. Biol. *4*.

Saksouk, N., Barth, T.K., 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.

Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin *8*, 3.

Sarraf, S.A., and 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.

Sasaki, H., and Matsui, Y. (2008). Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nat Rev Genet *9*, 129–140.

Sasaki, K., Ito, T., Nishino, N., Khochbin, S., and Yoshida, M. (2009). Real-time imaging of histone H4 hyperacetylation in living cells. Proc. Natl. Acad. Sci. U. S. A. *106*, 16257–16262.

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

Sato, Y., Kujirai, T., Arai, R., Asakawa, H., Otsuki, C., Harikoshi, N., Yamagata, K., Ueda, J., Nagase, T., Haraguchi, T., et al. (2016). A genetically encoded probe for live-cell imaging of H4K20 monomethylation. J. Mol. Biol.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. U. S. A. *103*, 1412–1417.

Scala, E., Ariani, F., Mari, F., Caselli, R., Pescucci, C., Longo, I., Meloni, I., Giachino, D., Bruttini, M., Hayek, G., et al. (2005). CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. J. Med. Genet. *42*, 103–107.

Scarsdale, J.N., Webb, H.D., Ginder, G.D., and Williams, D.C. (2011). Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. Nucleic Acids Res. *39*, 6741–6752.

Schiesser, S., Pfaffeneder, T., Sadeghian, K., Hackner, B., Steigenberger, B., Schröder, A.S., Steinbacher, J., Kashiwazaki, G., Höfner, G., Wanner, K.T., et al. (2013). Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine. J. Am. Chem. Soc. *135*, 14593–14599.

Schmid-Burgk, J.L., Schmidt, T., Kaiser, V., Höning, K., and Hornung, V. (2013). A ligation-independent cloning technique for high-throughput assembly of transcription activator–like effector genes. Nat. Biotechnol. *31*, 76–81.

Schmiedeberg, L., Skene, P., Deaton, A., and Bird, A. (2009). A Temporal Threshold for Formaldehyde Crosslinking and Fixation. PLoS One *4*, e4636.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. *21*, 1121–1131.

Schüle, B., Armstrong, D.D., Vogel, H., Oviedo, A., and Francke, U. (2008). Severe congenital encephalopathy caused by MECP2 null mutations in males: Central hypoxia and reduced neuronal dendritic structure. Clin. Genet. *74*, 116–126.

Schwartzman, O., and Tanay, A. (2015). Single-cell epigenomics: techniques and emerging applications. Nat. Rev. Genet. *16*, 716–726.

Schweiger, P.J., and Jensen, K.B. (2016). Modeling human disease using organotypic cultures. Curr. Opin. Cell Biol. *43*, 22–29.

Segal, D.J., Beerli, R.R., Blancafort, P., Dreier, B., Effertz, K., Huber, A., Koksch, B., Lund, C. V., Magnenat, L., Valente, D., et al. (2003). Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. Biochemistry *42*, 2137–2148.

Seumois, G., Chavez, L., Gerasimova, A., Lienhard, M., Omran, N., Kalinke, L., Vedanayagam, M., Ganesan, A.P. V, Chawla, A., Djukanović, R., et al. (2014). Epigenomic analysis of primary human T cells reveals enhancers associated with TH2 memory cell differentiation and asthma susceptibility. Nat. Immunol. *15*, 777–788.

Shahbazian, M.D., Antalffy, B., Armstrong, D.L., and Zoghbi, H.Y. (2002). Insight into Rett syndrome: MeCP2 levels display tissue-and cell-specific differences and correlate with neuronal maturation. Hum. Mol. Genet. *11*, 115–124.

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature *498*, 236–240.

Sharif, J., and Koseki, H. (2011). Recruitment of Dnmt1: Roles of the SRA protein Np95 (Uhrf1) and other factors. Prog. Mol. Biol. Transl. Sci. *101*, 289–310.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. a, Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908–912.

Sharma, A., Klein, S.S., Barboza, L., Lohdi, N., and Toth, M. (2016). Principles Governing DNA Methylation during Neuronal Lineage and Subtype Specification. J. Neurosci. *36*, 1711–1722.

Shen, Y. (2015). Optical Probes in Biology. Jin Zhang, Sohum Mehta, and Carsten Schultz (Eds.). CRC Press, Boca Raton, FL; 2015, 511 pages. ISBN 1466510110.

Shen, J. cheng, Rideout, W.M., and Jones, P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Res. *22*, 972–976.

Shen, L., Kondo, Y., Guo, Y., Zhang, J., Zhang, L., Ahmed, S., Shu, J., Chen, X., Waterland, R.A., and Issa, J.P.J. (2007). Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet. *3*, 2023–2026.

Shen, L., Inoue, A., He, J., Liu, Y., Lu, F., and Zhang, Y. (2014). Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell *15*, 459–470.

Shen, X., Mizuguchi, G., Hamiche, a, and Wu, C. (2000). A chromatin remodelling complex involved in transcription and DNA processing. Nature *406*, 541–544.

Shikauchi, Y., Saiura, A., Kubo, T., Niwa, Y., Yamamoto, J., Murase, Y., and Yoshikawa, H. (2009). SALL3 interacts with DNMT3A and shows the ability to inhibit CpG island methylation in hepatocellular carcinoma. Mol. Cell. Biol. *29*, 1944–1958.

Shimko, J.C., North, J.A., Bruns, A.N., Poirier, M.G., and Ottesen, J.J. (2011). Preparation of fully synthetic histone H3 reveals that acetyl-lysine 56 facilitates protein binding within nucleosomes. J. Mol. Biol. *408*, 187–204.

Shirane, K., Toh, H., Kobayashi, H., Miura, F., Chiba, H., Ito, T., Kono, T., and Sasaki, H. (2013). Mouse Oocyte Methylomes at Base Resolution Reveal Genome-Wide Accumulation of Non-CpG Methylation and Role of DNA Methyltransferases. PLoS Genet. *9*.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science (80-.). *311*, 844–847.

Sidoli, S., Cheng, L., and Jensen, O.N. (2012). Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry. J. Proteomics *75*, 3419–3433.

Simon, M., North, J.A., Shimko, J.C., Forties, R.A., Ferdinand, M.B., Manohar, M., Zhang, M., Fishel, R., Ottesen, J.J., and Poirier, M.G. (2011). Histone fold modifications control nucleosome unwrapping and disassembly. Proc Natl Acad Sci U S A *108*, 12711–12716.

Singleton, M.K., Gonzales, M.L., Leung, K.N., Yasui, D.H., Schroeder, D.I., Dunaway, K., and LaSalle, J.M. (2011). MeCP2 is required for global heterochromatic and nucleolar changes during activity-dependent neuronal maturation. Neurobiol. Dis. *43*, 190–200.

Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Andrews, R., and Bird, A.P. (2010). Neuronal MeCP2 Is Expressed at Near Histone-Octamer Levels and Globally Alters the Chromatin State. Mol. Cell *37*, 457–468.

Smallwood, S. a, Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S.R., Stegle, O., Reik, W., and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat. Methods *11*, 817–820.

Smeets, E.E.J., Pelc, K., and Dan, B. (2012). Rett syndrome. Mol. Syndromol. *2*, 113–127.

Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. Nature *511*, 611–615.

Soldi, M., and Bonaldi, T. (2013). The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components. Mol. Cell. Proteomics *12*, 764–780.

Soldi, M., Bremang, M., and Bonaldi, T. (2014). Biochemical systems approaches for the analysis of histone modification readout. Biochim. Biophys. Acta *1839*, 657–668.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P., and Li, G. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science (80-.). *344*, 376–380.

Song, J., Teplova, M., Ishibe-Murakami, S., and Patel, D.J. (2012). Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science *335*, 709–712.

Song, Y., Su, D., Shen, Y., Liu, H., and Wang, L. (2016). Design and preparation of open circuit potential biosensor for in vitro and in vivo glucose monitoring. Anal. Bioanal. Chem.

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W.T.C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., et al. (2013). Dynamic readers for 5-(Hydroxy)methylcytosine and its oxidized derivatives. Cell *152*, 1146–1159.

Ståhlberg, A., and Bengtsson, M. (2010). Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. Methods *50*, 282–288.

Stelzer, Y., and Jaenisch, R. (2016). Monitoring dynamics of DNA methylation at single-cell resolution during development and disease. In Cold Spring Harbor Symposia on Quantitative Biology, pp. 199–206.

Stelzer, Y., Shivalila, C.S., Soldner, F., Markoulaki, S., and Jaenisch, R. (2015). Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution. Cell *163*, 218–229.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.

Stroud, H., Greenberg, M.V.C., Feng, S., Bernatavichute, Y. V., and Jacobsen, S.E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell *152*, 352–364.

Su, Z., and Denu, J.M. (2016). Reading the Combinatorial Histone Language. ACS Chem. Biol. *11*, 564–574.

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

Suetake, I., Mishima, Y., Kimura, H., Lee, Y.-H., Goto, Y., Takeshima, H., Ikegami, T., and Tajima, S. (2011). Characterization of DNA-binding activity in the N-terminal domain of the DNA methyltransferase Dnmt3a. Biochem. J. *437*, 141–148.

Sugino, K., Hempel, C.M., Okaty, B.W., Arnson, H.A., Kato, S., Dani, V.S., and Nelson, S.B. (2014). Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes. J Neurosci *34*, 12877–12883.

Sun, L.Q., Lee, D.W., Zhang, Q., Xiao, W., Raabe, E.H., Meeker, A., Miao, D., Huso, D.L., and Arceci, R.J. (2004). Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG. Genes Dev. *18*, 1035–1046.

Suzuki, M.M., Takagi, C., Miura, S., Sakane, Y., Suzuki, M.M., Sakuma, T., Sakamoto, N., Endo, T., Kamei, Y., Sato, Y., et al. (2016). In vivo tracking of histone H3 lysine 9 acetylation in Xenopus laevis during tail regeneration. Genes to Cells *21*, 358–369.

Taberlay, P.C., and Jones, P.A. (2011). DNA methylation and cancer. Prog. Drug Res. *67*, 1–23.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. *16*, 1779–1791.

Takai, D., and Jones, P.A. (2004). Origins of Bidirectional Promoters: Computational Analyses of Intergenic Distance in the Human Genome. Mol. Biol. Evol. *21*, 463–467.

Tao, J., Van Esch, H., Hagedorn-Greiwe, M., Hoffmann, K., Moser, B., Raynaud, M., Sperner, J., Fryns, J.-P.J., Schwinger, E., Gécz, J., et al. (2004). Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation. Am. J. Hum. Genet. *75*, 1149–1154.

Tchekanda, E., Sivanesan, D., and Michnick, S.W. (2014). An infrared reporter to detect spatiotemporal dynamics of protein-protein interactions. Nat. Methods *11*, 641–644.

Termanis, A., Torrea, N., Culley, J., Kerr, A., Ramsahoye, B., and Stancheva, I. (2016). The SNF2 family ATPase LSH promotes cell-autonomous *de novo* DNA methylation in somatic cells. Nucleic Acids Res. gkw424.

Teves, S.S., An, L., Hansen, A.S., Xie, L., Darzacq, X., and Tjian, R. (2016). A dynamic mode of mitotic bookmarking by transcription factors. Elife *5*.

Tippin, D.B., and Sundaralingam, M. (1997). Nine polymorphic crystal structures of d(CCGGGCCCGG), d(CCGGGCCm5CGG), d(Cm5CGGGCCm5CGG) and d(CCGGGCC(Br)5CGG) in three different conformations: effects of spermine binding and methylation on the bending and condensation of A-DNA. J. Mol. Biol. *267*, 1171–1185.

Tirado-Magallanes, R., Rebbani, K., Lim, R., Pradhan, S., and Benoukraf, T. (2016). Whole genome DNA methylation: beyond genes silencing. Oncotarget.

Travers, A. (2014). The 30-nm fiber redux. Sci. (New York, NY) *344*, 370–372.

Traynor, J., Agarwal, P., Lazzeroni, L., and Francke, U. (2002). Gene expression patterns vary in clonal cell cultures from Rett syndrome females with eight different MECP2 mutations. BMC Med. Genet. *3*, 12.

Tudor, M., Akbarian, S., Chen, R.Z., and Jaenisch, R. (2002). Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. Proc. Natl. Acad. Sci. U. S. A. *99*, 15536–15541.

Ueda, J., Maehara, K., Mashiko, D., Ichinose, T., Yao, T., Hori, M., Sato, Y., Kimura, H., Ohkawa, Y., and Yamagata, K. (2014). Heterochromatin dynamics during the differentiation process revealed by the DNA methylation reporter mouse, methylRO. Stem Cell Reports *2*, 910–924.

Ulanovskaya, O.A., Zuhl, A.M., and Cravatt, B.F. (2013). NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. Nat. Chem. Biol. *9*, 300–306.

Umer, M., and Herceg, Z. (2013). Deciphering the epigenetic code: an overview of DNA methylation analysis methods. Antioxid. Redox Signal. *18*, 1972–1986.

Unoki, M., Nishidate, T., and Nakamura, Y. (2004). ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene *23*, 7601–7610.

Urdinguio, R.G., Lopez-Serra, L., Lopez-Nieva, P., Alaminos, M., Diaz-Uriarte, R., Fernandez, A.F., and Esteller, M. (2008). Mecp2-null mice provide new neuronal targets for rett syndrome. PLoS One *3*.

Valinluck, V., Tsai, H.H., Rogstad, D.K., Burdzy, A., Bird, A., and Sowers, L.C. (2004). Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). Nucleic Acids Res. *32*, 4100–4108.

Velicescu, M., Weisenberger, D.J., Gonzales, F.A., Tsai, Y.C., Nguyen, C.T., and Jones, P.A. (2002). Cell division is required for de Novo methylation of CpG islands in bladder cancer cells. Cancer Res. *62*, 2378–2384.

Véron, N., and Peters, A.H.F.M. (2011). Epigenetics: Tet proteins in the limelight. Nature *473*, 293–294.

Verveer, P.J., Rocks, O., Harpur, A.G., and Bastiaens, P.I.H. (2006). Imaging protein interactions by FRET microscopy: FRET measurements by acceptor photobleaching. CSH Protoc. *2006*.

Vogel, S., Thaler, C., and Koushik, S. (2006). Fanciful FRET. Sci STKE *re2*, 1–8.

Vongs, a, Kakutani, T., Martienssen, R. a, and Richards, E.J. (1993). Arabidopsis thaliana DNA methylation mutants. Science *260*, 1926–1928.

de Waal, E., Mak, W., Calhoun, S., Stein, P., Ord, T., Krapp, C., Coutifaris, C., Schultz, R.M., Bartolomei, M.S., Waal, E. De, et al. (2014). In vitro culture increases the frequency of stochastic epigenetic errors at imprinted genes in placental tissues from mouse concepti produced through assisted reproductive technologies. Biol. Reprod. *90*, 22.

Wallrath, L.L. (1998). Unfolding the mysteries of heterochromatin. Curr. Opin. Genet. Dev. *8*, 147–153.

Wang, J., Lawry, S.T., Cohen, A.L., and Jia, S. (2014). Chromosome boundary elements and regulation of heterochromatin spreading. Cell. Mol. Life Sci. *71*, 4841–4852.

Wang, R., Shi, Y., Chen, L., Jiang, Y., Mao, C., Yan, B., Liu, S., Shan, B., Tao, Y., and Wang, X. (2015). The ratio of FoxA1 to FoxA2 in lung adenocarcinoma is regulated by LncRNA HOTAIR and chromatin remodeling factor LSH. Sci. Rep. *5*, 17826.

Wang, T., Xu, C., Liu, Y., Fan, K., Li, Z., Sun, X., Ouyang, H., Zhang, X., Zhang, J., Li, Y., et al. (2012). Crystal Structure of the Human SUV39H1 Chromodomain and Its Recognition of Histone H3K9me2/3. PLoS One *7*.

Watanabe, D., Suetake, I., Tada, T., and Tajima, S. (2002). Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. Mech. Dev. *118*, 187–190.

Weaver, J.R., and Bartolomei, M.S. (2014). Chromatin regulators of genomic imprinting. Biochim. Biophys. Acta - Gene Regul. Mech. *1839*, 169–177.

Weaving, L.S.L., Christodoulou, J., Williamson, S.S.L., Friend, K.L., McKenzie, O.L.D., Archer, H., Evans, J., Clarke, A., Pelka, G.J., Tam, P.P.P.L., et al. (2004). Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. Am. J. Hum. Genet. *75*, 1079–1093.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457–466.

Wiedenheft, B., Sternberg, S.H., and Doudna, J. a. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331–338.

Williams, K., Christensen, J., and Helin, K. (2012). DNA methylation: TET proteins-guardians of CpG islands? EMBO Rep. *13*, 28–35.

Woodcock, C.L.F., Frado, L.L.Y., and Rattner, J.B. (1984). The higher-order structure of chromatin: Evidence for a helical ribbon arrangement. J. Cell Biol. *99*, 42–52.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol *32*, 670–676.

Wurm, F.M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol. *22*, 1393–1398.

Xie, W., Barr, C.L., Kim, A., Yue, F., Lee, A.Y., Eubanks, J., Dempster, E.L., and Ren, B. (2012). Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. Cell *148*, 816–831.

Yamagata, K. (2010). DNA methylation profiling using live-cell imaging. Methods *52*, 259–266.

Yamagata, K., Yamazaki, T., Miki, H., Ogonuki, N., Inoue, K., Ogura, A., and Baba, T. (2007). Centromeric DNA hypomethylation as an epigenetic signature discriminates between germ and somatic cell lineages. Dev. Biol. *312*, 419–426.

Yamamoto, K., and Sonoda, M. (2003). Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1. Biochem Biophys Res Commun *301*, 287–292.

Yamazaki, T., Yamagata, K., and Baba, T. (2007). Time-lapse and retrospective analysis of DNA methylation in mouse preimplantation embryos by live cell imaging. Dev. Biol. *304*, 409–419.

Yan, Q., Cho, E., Lockett, S., and Muegge, K. (2003a). Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. Mol. Cell. Biol. *23*, 8416–8428.

Yan, Q., Huang, J., Fan, T., Zhu, H., and Muegge, K. (2003b). Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. EMBO J. *22*, 5154–5162.

Yang, Y., Kucukkal, T.G., Li, J., Alexov, E., and Cao, W. (2016). Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. ACS Chem. Biol. *11*, 2706–2715.

Yasui, D.H., Peddada, S., Bieda, M.C., Vallero, R.O., Hogart, A., Nagarajan, R.P., Thatcher, K.N., Farnham, P.J., and Lasalle, J.M. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. Proc. Natl. Acad. Sci. U. S. A. *104*, 19416–19421.

You, M., Litke, J.L., and Jaffrey, S.R. (2015). Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. Proc. Natl. Acad. Sci. U. S. A. *112*, E2756-65.

Yu, L., Takeda, K., and Markoff, L. (2013). Protein-protein interactions among West Nile non-structural proteins and transmembrane complex formation in mammalian cells. Virology *446*, 365–377.

Yu, V.W.C., Yusuf, R.Z., Oki, T., Wu, J., Saez, B., Wang, X., Cook, C., Baryawno, N., Ziller, M.J., Lee, E., et al. (2016). Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. Cell *167*, 1310–1322.e17.

Yu, W., McIntosh, C., Lister, R., Zhu, I., Han, Y., Ren, J., Landsman, D., Lee, E., Briones, V., Terashima, M., et al. (2014). Genome-wide DNA methylation patterns in LSH mutant reveals de-repression of repeat elements and redundant epigenetic silencing pathways. Genome Res. *24*, 1613–1623.

Yu, Y., Blair, S., Gillespie, D., Jensen, R., Myszka, D., Badran, A.H., Ghosh, I., and Chagovetz, A. (2010). Direct DNA methylation profiling using methyl binding domain proteins. Anal. Chem. *82*, 5012–5019.

Yuan, K., and O'Farrell, P.H. (2016). TALE-light imaging reveals maternally guided, H3K9me2/3-independent emergence of functional heterochromatin in Drosophila embryos. Genes Dev. *30*, 579–593.

Yuan, K., Shermoen, A.W., and O'Farrell, P.H. (2014). Illuminating DNA replication during Drosophila development using TALE-lights. Curr. Biol. *24*, R144-5.

Yuan, P., Han, J., Guo, J., Orlov, Y.L., Huss, M., Loh, Y.H., Yaw, L.P., Robson, P., Lim, B., and Ng, H.H. (2009). Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev. *23*, 2507–2520.

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature *515*, 355–364.

Zee, B.M., Young, N.L., and Garcia, B. a (2011). Quantitative proteomic approaches to studying histone modifications. Curr. Chem. Genomics *5*, 106–114.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell *153*, 193–205.

Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U., and Cheng, X. (2003). Structural basis for the product specificity of histone lysine methyltransferases. Mol. Cell *12*, 177–185.

Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T.P., Bock, C., Santacruz, D., Ragozin, S., Reinhardt, R., Groth, M., Walter, J., et al. (2009). DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. PLoS Genet *5*, e1000438.

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

Zheng, X., Kim, Y., and Zheng, Y. (2015). Identification of lamin B-regulated chromatin regions based on chromatin landscapes. Mol. Biol. Cell *26*, 2685–2697.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7–18.

Zhu, P., and Li, G. (2016). Structural insights of nucleosome and the 30-nm chromatin fiber. Curr. Opin. Struct. Biol. *36*, 106–115.

Zhu, H., Geiman, T.M., Xi, S., Jiang, Q., Schmidtmann, A., Chen, T., Li, E., and Muegge, K. (2006). Lsh

is involved in de novo methylation of DNA. EMBO J. *25*, 335–345.

Zoghbi, H.Y. (2016). Rett Syndrome and the Ongoing Legacy of Close Clinical Observation. Cell *167*, 293–297.

Zou, X., Ma, W., Solov'Yov, I.A., Chipot, C., and Schulten, K. (2012). Recognition of methylated DNA through methyl-CpG binding domain proteins. Nucleic Acids Res. *40*, 2747–2758.

7 Author's contribution

● **Lungu**, **C**., Muegge, K., Jeltsch, A., & Jurkowska, R. Z. (2015). An ATPase-deficient variant of the SNF2 family member HELLS shows altered dynamics at pericentromeric heterochromatin. *Journal of molecular biology*, 427(10), 1903-1915.

C.L. designed the study with contributions from R.Z.J. and A.J.. C.L. established the chromatin fractionation and FRAP assays and carried out all experimental work. R.Z.J. performed the cloning of the mammalian EYFP/ECFP-HELLS expression vectors and initiated the collaboration with K.M. who provided the template constructs encoding for murine HELLS. C.L wrote and prepared the first draft of the manuscript, including all figures, and provided input on all successive drafts, including the final manuscript. R.Z.J. and A.J. guided the work.

● Rajavelu, A.[#], **Lungu**, **C**.[#], Hanelt, I., Parsa, A., Karnik, R., Gu, H., Carell, T., Meissner, A. Jurkowska, R. Z. & Jeltsch, A. Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Manuscript submitted for review*. [#]co-shared first authors.

A.R., A.J., C.L, and R.Z.J. devised the project and analyzed the data. C.L. and A.R. conducted the biochemical assays with contributions from I.H.. C.L. and A.R. performed the fluorescence microscopy experiments. C.L. performed the cell culture and biochemical work for the data shown in Fig. 7. E.P. and T.C. performed the LC-ESI-MS/MS. H.G. made the WGBS libraries, R.K. performed the analysis and interpreted the results with A.M. All authors contributed to data interpretation and discussion, read and approved the final manuscript.

● Jeltsch, A., **Lungu**, **C**., Rajavelu, A., Jurkowska, R.Z. (2016) DNA methyltransferase inhibitors for Rett syndrome therapy. EP16156521. Registered on 19.02.2016.

A.J. conceived the idea. C.L. and A.R. performed the experimental work. A.J. and R.Z.J guided the work.

● **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. *Nature Communications*, 8(1), 649-649..

C.L. and A.J. conceived and designed the study. C.L. and S.P. performed the experimental work. P.R. provided critical experience, constructs and cell lines for retroviral infection. B.J. performed the genome mining analysis for TALE binding sites. C.L. along with A.J. guided the work of S.P. C.L wrote and prepared the first draft of the manuscript including all figures and provided input on all successive drafts, including the final manuscript.

8 Appendix

8.1 Appendix 1 (not included in the published thesis)

Lungu, **C**., Muegge, K., Jeltsch, A., & Jurkowska, R. Z. (2015). An ATPase-deficient variant of the SNF2 family member HELLS shows altered dynamics at pericentromeric heterochromatin. *Journal of molecular biology*, 427(10), 1903-1915.

8.2 Appendix 2

Rajavelu, A.[#], **Lungu**, **C**.[#], Hanelt, I., Parsa, A., Karnik, R., Gu, H., Carell, T., Meissner, A. Jurkowska, R. Z. & Jeltsch, A. Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Manuscript submitted for review*. [#]co-shared first authors.

8.3 Appendix 3 (not included in the published thesis)

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. *Nature Communications*, 8(1), 649-649..

8.4 Appendix 4

Jeltsch, A., **Lungu**, **C**., Rajavelu, A., Jurkowska, R.Z. (2016) DNA methyltransferase inhibitors for Rett syndrome therapy. EP16156521. Registered on 19.02.2016.

Appendix 2

Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2

Arumugam Rajavelu[1,2,#], Cristiana Lungu[1,#], Ines Hanelt[1], Michael Dukatz[1], Edris Parsa[3], Rahul Karnik[4], Hongcang Gu[4], Thomas Carell[3], Alexander Meissner[4], Renata Z. Jurkowska[1,5] & Albert Jeltsch[1,*]

[1] Institute of Biochemistry, Faculty of Chemistry, University Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart, Germany

[2] Current address: Rajiv Gandhi Center for Biotechnology (RGCB), Trivandrum 695014, Kerala, India

[3] Center for Integrated Protein Science (CiPSM) at the Department of Chemistry, Ludwig-Maximilians-University, Butenandtstr. 5-13, 81377 Munich, Germany

[4] Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA; Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA.

[5] Current address: BioMed X Innovation Center, Im Neuenheimer Feld 583, D-69120 Heidelberg, Germany

[#] These authors contributed equally to the work

* Corresponding author:
Phone: (+)49 711 685 64390
Fax: (+)49 711 685 64392
E-mail: albert.jeltsch@ibc.uni-stuttgart.de
http://www.ibc.uni-stuttgart.de/

Running title: MeCP2 binds and regulates DNMT3A

Abstract

Despite their importance in mammalian development, the mechanisms controlling the generation of genomic methylation patterns are still poorly understood. Here, we identify MeCP2 as a direct and strong interactor of DNMT3A and map interaction to the the ADD domain of DNMT3A and the TRD domain of MeCP2. We observe a strong inhibition of DNMT3A by MeCP2 *in vitro* and a global reduction of DNA methylation after overexpression of MeCP2 in human cells. By using conformationally locked DNMT3A variants, we show that binding of MeCP2 allosterically stabilizes the autoinhibitory conformation of DNMT3A. This interaction and its resulting inhibition could be relieved by histone H3 binding to DNMT3A. We performed whole genome methylome analysis in MeCP2 knock-out mice brain and identified both hypo- and hypermethylated regions overlapping with MeCP2 binding sites. Hypermethylation was mainly observed at CpG promoters, suggesting that MeCP2 contributes to the protection of these regions from methylation in brain. Our data suggest that DNMT3A localization and activity are under the combined control of MeCP2 and H3 tail modifications and, depending on the modification status of the H3 tail at target sites, MeCP2 can act as either a repressor or activator.

Keywords: DNA methylation / DNMT3A / MeCP2 / enzyme regulation / epigenetics / Rett syndrome

Introduction

In mammals, DNA methylation is established during embryogenesis by the *de novo* DNA methyltransferases DNMT3A and DNMT3B (1,2). The correct establishment and maintenance of DNA methylation patterns depends on the accurate targeting and regulation of these enzymes (3). Mammalian DNMTs comprise two parts, a large multi-domain N-terminal part and a C-terminal catalytic domain (CD), which is conserved between eukaryotic and prokaryotic cytosine-C5 DNA methyltransferases (4,5). The N-terminal parts of DNMTs guide their nuclear and sub-nuclear localization and mediate interactions with other proteins, DNA and chromatin, thereby regulating the catalytic activity. The N-terminal parts of DNMT3A and DNMT3B contain two defined sub-domains (Fig. 1A), a PWWP domain, an H3K36me3 binder (6,7), and an ADD domain, which binds the N-terminus of histone H3 if K4 does not carry modifications larger than monomethylation (8-11). Structural and biochemical work has shown that the ADD domain of DNMT3A binds to the CD at two distinct interfaces (Fig. 1B) (12), an autoinhibitory site in which ADD blocks catalytic activity, and an allosteric site where binding does not lead to inhibition. Different ADD residues are contacting the CD in the two conformations: Y526 only contacts the CD in the allosteric conformation and D531 only contacts the CD in the autoinhibitory conformation, while it interacts with the H3 peptide in the allosteric conformation. Binding of the H3 peptide to the ADD domain consequently favors the allosteric conformation and activates DNMT3A (10,12,13). Several studies have recently shown that the activity of DNMT3 enzymes is regulated by H3 tail interaction not only *in vitro,* but also in cells (7,14,15).

The biological role of DNA methylation is mediated by proteins, which specifically bind to DNA carrying methylated cytosines (16). One important reader of DNA methylation is the chromatin regulator MeCP2 (Fig. 1A) (17-20). It is the founding member of a group of proteins containing a so called methyl-binding domains (MBD), which bind to DNA in a methylation specific manner (21). MeCP2 binds methylated DNA with a preference for CpG dense islands (19,22), but its binding strength is strongly influenced by the flanking sequence of the methylated CpGs (23,24) and it also binds methylcytosine in a non-CpG context (25,26). In addition to the MBD domain, MeCP2 contains a Transcriptional Repression Domain (TRD), which serves as a protein recruitment platform and shows methylation-independent DNA binding (20,27). MeCP2 is known to interact with various transcriptional repressors and co-repressors, including HDACs, DNMT1, ATRX and Sin3A (28-31). Functionally, MeCP2 is involved in numerous cellular processes, like methylation-induced gene repression (in particular of long genes), control of repetitive elements, chromatin compaction and looping and

splice site regulation (20,26,27,32-35). In addition to its role in gene silencing, gene expression studies performed in specific brain sub-regions showed altered expression levels of hundreds of genes after loss of MeCP2, most of which upregulated by MeCP2 (36-38) indicating that MeCP2 can function as both a gene activator and repressor depending on the genomic context (39).

Both DNMT3A and MeCP2 are highly expressed in neurons (32,40) (EBI expression atlas http://www.ebi.ac.uk/gxa/) and they have important functions in the brain. DNMT3A has been implicated in neuromuscular control, synaptic plasticity, learning and memory (41-43). MeCP2 functions as structural protein in neurons and forms a specific type of chromatin, which is depleted of histone H1 (20). It plays an essential role in brain plasticity (32) and inactivating mutations of the X-linked *mecp2* gene were shown to cause Rett syndrome, a severe neurodevelopmental disease associated with developmental disorders and autism-like symptoms in females(27,31).

Since it has been observed frequently that readers of epigenetic modifications interact with enzymes setting these modifications (Suppl. Text 1) and DNMT3A and MeCP2 both were known to bind to pericentromeric heterochromatin, we investigated their potential interaction in this work. We observed a strong binding of MeCP2 to DNMT3A *in vitro*, in cells, and in the mouse brain and mapped the interaction interface to the TRD domain of MeCP2 and the ADD domain of DNMT3A, respectively. Binding of MeCP2 to DNMT3A strongly inhibited its enzymatic activity by favoring the autoinhibitory conformation of DNMT3A. Interestingly, this interaction was disrupted and the DNMT3A inhibition relieved by addition of unmodified H3. These results indicate that MeCP2 can function as a negative regulator of global DNA methylation by reducing untargeted DNA methylation. Conversely, after binding of the MeCP2-DNMT3A complex to MeCP2 target sites marked by appropriately modified H3 tails, H3 binding to DNMT3A can disrupt the interaction with MeCP2 and activate DNMT3A. This will initiate a positive feedback loop, in which increased DNA methylation will lead to increased MeCP2 binding. In conclusion, our data show that MeCP2 interaction controls DNMT3A activity and localization and that this interaction has dual and opposing roles in the control of DNA methylation patterns by either inhibiting or stimulating DNA methylation activity, depending on the local chromatin context.

Materials and methods

Generation of DNMT3A2 mutants and MeCP2 domain constructs

The murine DNMT3A and MeCP2 proteins and protein domains were prepared as indicated in Fig. 1A. The MeCP2 domains and mutants proteins were cloned in pGEX-6P2 vector using BamHI and XhoI cloning sites. All constructs were verified by DNA sequencing. The site directed mutagenesis were carried out by rolling circle PCR using a primer carrying point mutation (44). The presence of the mutations was confirmed by restriction marker analysis and by DNA sequencing.

Expression and purification of MeCP2 proteins and DNMT3A proteins

The MeCP2 domains and mutants proteins were expressed *E. coli* BL-21 cells. Cells were cultivated in LB medium at 37 °C while shaking until an OD (600 nm) of 0.6 – 0.7 was reached. Then, protein expression was induced by addition of 1 mM of isopropyl-β-D-thiogalactoside (IPTG) and the culture was incubated at 18 °C shaking at 200 rpm overnight. The cells were harvested by centrifugation (15 min at 4600 rpm) and the pellet resuspended in sonication buffer (20 mM HEPES pH 7.5, 500 mM KCl, 1 mM EDTA, 1 mM DTT, 10% glycerol) including protease inhibitor (Sigma). The cells were lysed by sonication and centrifuged at 18000 rpm for 1 h to prepare a clear lysate, which was applied on a GST-sepharose column (GE Healthcare). After washing with sonication buffer, the protein was eluted with sonication buffer containing 50 mM reduced glutathione and dialyzed first against dialysis buffer I (20 mM HEPES pH 7.5, 200 mM KCl, 1 mM EDTA, 1 mM DTT, 10% glycerol) for three hours, then against dialysis buffer II containing 60% glycerol overnight. The purified proteins were analyzed on 12% SDS-PAGE gel stained with colloidal Coomassie BB. The murine DNMT3A2 and DNMT3A-C proteins were expressed and purified as described previously (45-47). Since all DNMT3A structures were annotated with numbers for the human proteins, we use human numbering here. The residue numbers corresponding to human Q527, D528, and D531 in murine DNMT3A are Q523, D524, and D527, respectively.

GST pull-down experiments

For GST pull-down experiments, 20 µL of Glutathione-Sepharose 4B beads were washed with 200 µL of interaction buffer (25 mM Tris pH 8.0, 100 mM KCl, 5 mM MgCl2, 10 % Glycerol, 0.1 % NP40, 200 µM PMSF). The beads were incubated for 1 hour at 4 °C with 10-15 µg of the different GST-tagged proteins, washed three times with interaction buffer and incubated with His- or MBD-tagged proteins (15 µg) for 1 hour at 4 °C with shaking. Then, the beads were washed three times with wash buffer containing high salt (25 mM Tris pH 8.0, 5 mM MgCl2, 300 mM KCl, 10 % Glycerol, 0.1 % NP40, 200 µM PMSF). The interaction of DNMT3A ADD and MeCP2 TRD domains was also investigate using buffer containing 600 mM KCl. Finally, the beads were resuspended in SDS gel loading buffer and incubated for 10 min at 95 °C. After centrifugation of the beads at 14,000 rpm for 10 min the supernatant was loaded on a 12 % SDS PAGE gel. Proteins were detected by Western blotting or Coomassie BB staining as indicated. Some experiments were conducted in the presence of recombinant histone H3.1 (M2503S, NEB), as detailed in the results section.

Co-immunoprecipitation assay

For co-immunoprecipitation of DNMT3A and MeCP2, pcDNA DNMT3A (expressing myc tagged DNMT3A) and pEYFP-MeCP2 plasmids were co-transfected in HEK293 cells. The pEYFP plasmid was used as control. After 48 h, the cells were collected and the cell pellets stored at -80 °C. The cells were lysed as recommended by the GFP trap protocol (ChromTek). Using GFP trap YFP tagged MeCP2 was pulled-down and the complex washed with buffer (10 mM Tris/Cl pH 7.5, 0.5 mM EDTA, 0.5% NP-40, 200 mM NaCl). The MeCP2 and DNMT3A proteins were separated on a 12 % SDS-PAGE and transferred to nitrocellulose membrane. To detect DNMT3A, the blot was probed with anti-myc antibody (Santa Cruz, 1:1000 dilution) for 1 h at room temperature.

For immunoprecipitation of endogenous DNMT3A and MeCP2, whole brains from sixteen-week old C57Bl/N female mice were used. Following mechanical disruption, the tissue was lysed following a published protocol (48) with some modifications. Three brains were homogenized in NP-40 lysis buffer (10 mM HEPES, pH 7.9, 3 mM $MgCl_2$, 10 mM KCl, 10 mM NaF, 1 mM Na_3VO_4, 0.5 mM DTT, 0.5% NP-40, 1× complete EDTA-free protease inhibitor cocktail (Roche)), by douncing 30× with a tight pestle, and pelleted at 1,000g. Lysates were next diluted 1:1 with Benzonase buffer (10 mM HEPES, pH 7.9, 3 mM $MgCl_2$, 280 mM NaCl,

0.2 mM EDTA, 10 mM NaF, 1 mM Na_3VO_4, 0.5 mM DTT, 0.5% NP-40) supplemented with 1× complete EDTA-free protease inhibitor cocktail (Roche) and sonicated with EpiShear (Active Motif) for 2 min 30 sec (15 sec ON, 30 sec OFF cycles, 20% power, 3.2 mm microtip). The homogenate was then digested with 500 units of Benzonase (Novagen) for 2 h rotating at 4 °C. Chromatin proteins were separated by centrifugation at 17.000g for 20 min at 4 °C. For each pull-down, 2.5 mg lysate were pre-cleared for 1 h at 4 °C with 20 mg Protein A Sepharose CL-4B (GE Healthcare), followed by overnight incubation with 15 µg anti-Dnmt3a antibody (sc-2070, Santa Cruz). For negative control, an equivalent amount of non-related rabbit IgG anti-myc (ab9106, Abcam) antibody was used. The antibody-bound proteins were immobilized to 100 mg Protein A Sepharose CL-4B, blocked in 10% BSA (Roth) for 6 h rotating at 4 °C. After five washes with immunoprecipitation buffer, the immune complexes were eluted from the beads by boiling in 100 µL Laemmli sample buffer. The samples were next analyzed by western blot as described above. For detection, anti-MeCP2 monoclonal primary antibody (#3456, Cell Signaling) was used, followed by incubation with HRP-linked anti-rabbit IgG light chain specific secondary antibody (211-032-171, Jackson ImmunoResearch). Western lighting *Ultra* (Perkin Elmer) was used as ECL HRP substrate.

Fluorescence microscopy

For sub-nuclear localization studies, NIH3T3 cells were seeded on glass slides and transfected with plasmids expressing CFP and YFP tagged DNMT3L, DNMT3A-ADD and MeCP2 using Fugene HD (Promega) according to the manufacture´s instructions. After 48 h, the cells were fixed with 4% formaldehyde, mounted in Mowiol and Z stacks images were collected using a Zeiss LSM 710 confocal microscope.

Substrates used for DNA methylation

The following oligonucleotide substrates were used for DNA methylation assays: a biotinylated unmethylated 30mer containing one CpG site (um30mer: TTG CAC TCT CCT CCC GGA AGT CCC AGC TTC / Bt-GAA GCT GGG ACT TCC GGG AGG AGA GTG CAA), the same sequence hemimethylated at the CpG site with the methylation in the lower DNA strand (hm30mer: TTG CAC TCT CCT CCC GGA AGT CCC AGC TTC / Bt-GAA GCT GGG ACT TC^{m}C GGG AGG AGA GTG CAA), a biotinylated hemimethylated 30mer with optimized

flanks for methylation with DNMT3A (49) (hmF30mer: GAA GCT GGA CAG TAC GTC AAG AGA GTG CAA / Bt-TTG CAC TCT CTT GA^{m}C GTA CTG TCC AGC TTC) and a non-CpG substrate (nonCpG: GAA GCT GGT CCA TT^{m}C GAT GAT GGA GTG CAA / Bt-TTG CAC TCC ATC AT^{m}C GAA TGG ACC AGC TTC). The oligonucleotides were annealed by heating to 86 °C for 5 min and slowly cooling down to ambient temperature. In addition, a biotinylated 585mer DNA substrate obtained by PCR was used which contains 8 HpaII sites (CCGG) and 45 CpG sites (um585mer). The 585mer was amplified from lambda DNA by PCR using the following primers Bt-GAA GGA CAA CCT GAA GTC CAG GTTG and GTG TAT GAC CAC CAG AGC CTT TTGC and purified by PCR purification kits (Qiagen). To prepare partially methylated 585mer (pm585mer), the DNA was methylated with M.HpaII (NEB) following the protocol of the provider and afterwards purified by PCR purification kits. Successful pre-methylation at HpaII sites was confirmed by HpaII (NEB) restriction digestion.

DNA methylation activity assay

The avidin-biotin microplate DNMT activity assay was used to monitor the activity of different DNMT3A variants in the methylation of biotinylated DNA substrates, basically as described (50). Each well of the microplate was coated with 1 µg of Avidin dissolved in 100 µL of 100 mM $NaHCO_3$ (pH 9.6) and incubated overnight at 4 °C. Before starting the assay, the wells were washed five times with 200 µL of 1xPBST (140 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM K_2HPO_4, 0.05% v/v Tween 50, pH 7.2) containing 0.5 M NaCl. The reaction mixtures were prepared containing 2.5 µM DNMT3A2 or DNMT3A-C and 3 µM MeCP2 (or any of its domains) in methylation buffer (20 mM HEPES pH 7.2, 1 mM EDTA, 50 mM KCl, 1.25 mg/ml bovine serum albumin). For the control reactions without MeCP2, the same volume of dialysis buffer II was added instead of the MeCP2. The reaction mixtures were incubated on ice for 20 minutes and the wells of the plate were filled with 5 µL of 0.5 M unlabeled AdoMet (Sigma) dispensed in 35 µL 1x PBST/0.5 M NaCl. Then, 1 µM 30mer oligonucleotide DNA or 100 nM of 585mer DNA and 0.76 µM [methyl-3H]-AdoMet (PerkinElmer Life Sciences) were added to the reaction mixture and the samples incubated at 37 °C. In order to follow the time course of the reaction, aliquots of 2 µL were removed from the reaction mixtures in duplicates at time points between 2’ and 30’ and applied to one well of the microplate where the incorporation of labeled AdoMet was quenched by an excess of unlabeled AdoMet. This mixture was incubated while slightly shaking for 30 minutes to 1 hour. 200 µL of 1 x PBST and 0.5 M NaCl were used to wash the wells five times. 40 µL of unspecific nuclease from

Serratia marcescens was added, diluted 1:2500 in 50 mM Tris/HCl pH 8.0, 5 mM $MgCl_2$ were added per well and the mixture was incubated for 30 to 60 minutes with slight shaking. Finally, the released radioactivity was measured using liquid scintillation counting and the average count per minute of the duplicates was plotted against time. Linear regression was used to obtain the slopes of the initial linear parts of the time courses. The data are reported as averages and SEM of at least two independent experiments.

DNA methylation analysis in HCT116 DNMT1 hyphomorph cells

To study the effect of MeCP2 on DNMT3A mediated methylation in cells, we used HCT116 DNMT1 hyphomorphic cells (kindly provided by Prof. Bert Vogelstein, HHMI, USA), which have a reduced DNA methylation (51,52). HCT116 cells were cultivated in McCoy's 5A medium (Gibco cat no: 16600) supplemented with 10% heat-inactivated calf serum and 2 mM L-glutamine (Sigma), at 37 °C in a saturated humidity atmosphere containing 5% CO_2. For stable cell line generation, pEYFP-NLS or pEYFP-MeCP2 plasmids were transfected into HCT116 cells using jetPRIME (peqlab), following standard protocols. At 24 hours post-transfection the cells were subjected to selection with 0.5 mg/mL G418 (Sigma-Aldrich). For the pEYFP-MeCP2-transfected cells, stable clonal populations were derived by clonal expansion of single cell dilutions. For pEYFP-NLS, a mixed clonal population was derived. The cultures, together with untransfected control cells, were passaged for 2 months before genomic DNA isolation. LC-ESI-MS/MS analysis of DNA methylation in genomic DNA was performed as described (53,54).

WGBS library preparation and analysis

200 ng of genomic DNA was fragmented using a Covaris S2 for 6 min according to the following program: duty cycle 5%; intensity 5; cycle per burst 200. The sheared DNA was purified using the DNA Clean & Concentrator kit from Zymo Research per the manufacturer's recommendations. Bisulfite conversion of DNA was then conducted using the EZ DNA Methylation-Gold kit (Zymo Research) followed the instructions, eluting to 15 µl low TE buffer. To minimize continual loss during storage, the converted DNA was immediately processed for generating WGBS libraries using the Accel-NGS Methyl-Seq DNA library kit (Swift Biosciences) following the manufacturer's protocol. The libraries were sequenced for

100-bp paired-end reads on an Illumina HiSeq 2500 sequencer. Reads were aligned to the mm9 build of the mouse genome using bsmap (55). Methylation calls for CpGs were assigned using mcall from the MOABS package (56). Methylation for each 1kb tile was calculated as the coverage-weighted mean of methylation at CpGs within the tile. Differentially methylated tiles were identified by using a coverage weighted t-test (57) with the methylation values of the CpGs within the tile as samples. Q-values were calculated using the R qvalue package (58). The overlap of DMR regions with MeCP2 binding sites was determined using MeCP2 ChIP-seq data set GEO GSE71126 (59). Enrichments for genomic features were calculated by first counting the number of DMRs overlapping each feature set (including partial overlaps), and then using –log10(p-value) from Fisher's exact test as the degree of enrichment.

Results

DNMT3 proteins interact with MeCP2

To test whether DNMT3A and MeCP2 interact, we first performed GST-pull-down assays using recombinant murine full-length (FL) proteins (Fig. 1A). MeCP2 could be obtained with good purity after generating a version that lacked the N-terminal unstructured domain (MeCP2ΔN) (Suppl. Fig. 1). As shown in Fig. 2A, a robust pull-down of DNMT3A by GST-MeCP2ΔN could be detected. Similar experiments demonstrated the interaction of GST-MeCP2ΔN with DNMT3A2 (Fig. 2B, Suppl. Fig. 2), a naturally occurring isoform of DNMT3A (60). Furthermore, MeCP2ΔN binds to DNMT3L, a non-catalytic member of the DNMT3 family with important regulatory functions (Fig. 2C, Suppl. Fig. 2). However, no interaction of MeCP2ΔN with the C-terminal domain of DNMT3L (DNMT3L-C) could be detected (Fig. 2D, Suppl. Fig. 3) suggesting that the interaction is mediated by the ADD domain. Pull-down experiments with DNMT3A-C were inconclusive, because this protein showed unspecific binding to the GST beads.

We next tested the interaction of DNMT3A and MeCP2 in HEK293 cells by transient co-expression of Myc-tagged DNMT3A and EYFP-tagged MeCP2 followed by EYFP targeted purification. As shown in Fig. 2E, co-purified Myc-DNMT3A was detected after co-expression with EYFP-MeCP2, but not with the EYFP control. To exclude potential artifacts related to protein over-expression, we finally performed immuno-precipitation of endogenous DNMT3A from mouse brain protein extracts. As shown in Fig. 2F, MeCP2 could be specifically detected

in the pulled-down material, but not in the IgG control. Together, these results demonstrate that DNMT3A and MeCP2 interact *in vitro,* after transient expression in cells and in brain extracts.

Mapping of the MeCP2-DNMT3 interface

The fact that MeCP2 interacts with the full-length DNMT3 proteins, but not with DNMT3L-C suggested that MeCP2 binds to the ADD domain of DNMT3A and DNMT3L. To test this hypothesis, we performed pull-down assays with GST-tagged N-terminal domains of DNMT3A (Suppl. Fig. 3) as baits and a GST-cleaved MeCP2ΔN as prey. As documented in Fig. 3A, MeCP2ΔN showed a specific interaction with the ADD domain of DNMT3A, but not with its NTD or PWWP domains. To dissect which part of MeCP2 is responsible for the interaction with DNMT3A, we performed pull-downs with GST-tagged MeCP2 domains (Suppl. Fig. 4). As shown in Fig. 3B, interaction was detectable only for the TRD domain. The MeCP2 TRD also showed interaction with DNMT3A2, which was stable even under high salt conditions indicating that it is not driven by electrostatic interactions (Fig. 3C). Pull-downs using the isolated MeCP2-TRD and DNMT3A-ADD domains confirmed the direct and strong interaction of these domains (Fig. 3D). Since MeCP2 is mutated in Rett syndrome, we also tested the effect of the R306C Rett mutation in the TRD domain of MeCP2 on the interaction with DNMT3A, but did not observe any change when compared with the wild type TRD (Fig. 3D).

The interaction with MeCP2 inhibits the catalytic activity of DNMT3A

To elucidate the function of the interaction between MeCP2 and DNMT3A, we measured the *in vitro* rates of DNA methylation by DNMT3A2 in the presence of MeCP2 or its TRD domain with 6 different DNA substrates: 1) An unmethylated 30mer oligonucleotide (um30mer). 2) A hemimethylated state (hm30mer). 3) A hemimethylated 30mer with an optimized flank for DNMT3A (hmF30mer) (49). 4) A 585mer PCR fragment (um585mer, Suppl. Fig. 5). 5) The 585mer PCR fragment pre-methylated at HpaII sites (pm585mer). 6) A 30mer oligonucleotide non-CpG substrate.

Using 2.5 µM DNMT3A2 and 3 µM MeCP2, we consistently observed that the interaction of MeCP2 with DNMT3A2 resulted in about 40-60% reduction in DNMT3A2 enzymatic activity with the unmethylated substrates (Fig. 4A). A similar reduction in activity was observed with

DNMT3B2, a truncated form of DNMT3B corresponding to DNMT3A2 (Suppl. Fig. 9). The activity of DNMT3A2 was further reduced by about 80% with methylated substrates, which can be attributed to a better binding of MeCP2 to methylated DNA via its MBD domain. We speculated that binding of MeCP2 to pre-methylated DNA might target DNMT3A2. To test this hypothesis, a partially methylated 585mer DNA substrate was prepared by methylation of the DNA substrate with the HpaII methyltransferase, which methylates CG sites in a CCGG context, but leaves the remaining CG sites unmethylated (Suppl. Fig. 5). However, even on this substrate, we observed inhibition of the activity of DNMT3A2 by MeCP2 (Fig. 4A, pm585mer).

As a control, we used DNMT3A-C, which lacks the ADD domain and therefore cannot interact with MeCP2, and observed that MeCP2 did not inhibit its activity (Fig. 4B). In addition, we tested whether the isolated TRD domain also inhibits the activity of DNMT3A2 and observed 40-70% inhibition with the different substrates (Fig. 4C). We have discovered previously that DNMT3A also methylates cytosine in non-CpG sites, preferably in a CA context (61). Non-CpG methylation has recently been detected in considerable amounts in human ES cells and neurons and it was connected to DNMT3A activity (25,62-64). We, therefore, also investigated the influence of MeCP2 on the non-CpG methylation activity of DNMT3A2 using a 30mer oligonucleotide substrate which contains one already methylated CpG site, such that additional methylation could only occur at non-CpG sites. As with the other DNA substrates, we observed a similar inhibition of DNMT3A2 activity by TRD (Fig. 4C, non-CpG 30mer), indicating that CpG and non-CpG methylation are equally inhibited by the TRD interaction.

Finally, we tested the activity of DNMT3A2 in the presence of increasing amounts of TRD and observed that the methyltransferase activity was strongly inhibited (>95%) using a 2.4 fold excess of TRD (6 µM with 2.5 µM DNMT3A2) (Fig. 4D). As a control, the same experiments were conducted with DNMT3A-C, but here only a very weak reduction of activity was observed (Fig. 4D). This result indicates that the inhibition of DNMT3A2 by TRD is not caused by competition for the DNA substrate, which is an important control, since TRD has been reported to bind DNA (65,66). In summary, our results indicate that the interaction between the TRD domain of MeCP2 and the ADD domains of DNMT3A and DNMT3B results in a direct and very strong inhibition of their DNA methylation activity at CpG and non-CpG sites.

The TRD domain inhibits DNMT3A2 activity by an allosteric mechanism

Next, we aimed to elucidate the mechanism of DNMT3A inhibition by the TRD domain of MeCP2. As described above, structural studies showed that the ADD domain of DNMT3A binds to the CD at an allosteric and an autoinhibitory site (12). Binding of the H3 peptide to the ADD domain stabilizes the allosteric conformation and thereby activates DNMT3A (12,13) (Fig. 5A). To investigate the mechanism of the repression of DNMT3A by the TRD domain, we prepared DNMT3A2 variants containing mutations at Y526 or D531 in the ADD domain, two critical residues involved in the two binding sites at the CD (Fig. 1B, Suppl. Fig. 7).

In order to selectively disrupt or strongly destabilize one of the two DNMT3A conformations, we designed mutations that would cause large changes of the physiochemical properties of the corresponding amino acids. Y526E was introduced to disrupt the allosteric and D531R to disrupt the autoinhibitory conformation. After confirming that both mutants still interact with MeCP2-TRD (Suppl. Fig. 8), we investigated if these conformationally locked DNMT3A variants still respond to the presence of the TRD. As shown in Fig. 5B, the inhibitory effect of MeCP2 was specifically lost in the D531R variant that cannot adopt the autoinhibitory conformation. This finding suggests that MeCP2 reduces the activity of DNMT3A by an allosteric mechanism, in which TRD binding favors the autoinhibitory conformation. To investigate if TRD and histone H3 binding to the ADD domain influence each other, we conducted pull-down experiments using GST-TRD and DNMT3A2 in the presence of increasing concentrations of recombinant histone H3. As shown in Fig. 5C, addition of histone H3 abolished the ADD-TRD interaction, suggesting that binding is mutually exclusive and no ternary complex is formed. We next conducted DNA methylation experiments with DNMT3A2 and DNMT3A2 pre-incubated with the unmodified H3 (1-19) peptide in the absence and presence of TRD, and again observed that the inhibitory effect of the TRD domain was lost in the presence of the H3 peptide (Fig. 5D). This indicates that the binding of H3 to the ADD domain can disrupt the TRD interaction and relieve the TRD-mediated inhibition of DNMT3A.

MeCP2 influences the sub-nuclear localization of the DNMT3A ADD domain and DNMT3L

To study the cellular effect of the interaction of DNMT3 proteins with MeCP2, we investigated the sub-nuclear localization of both proteins and protein domains. DNMT3A (47,67) (and references therein) and MeCP2 (19,48) are both known to localize to DAPI-stained heterochromatin, which forms characteristic spots in mouse NIH3T3 cells. This natural co-

localization precluded direct studies of the mutual effect of both proteins on their sub-nuclear localization. However, DNMT3L was shown to have an almost homogenous nuclear distribution (67). Expression of both DNMT3L and MeCP2 with fluorescence tags confirmed the published localization patterns (Fig. 6A). Interestingly, co-expression of DNMT3L with MeCP2 led to a clear redistribution of DNMT3L towards the heterochromatic spots (Fig. 6B). This finding confirms the intracellular interaction of DNMT3L and MeCP2 and indicates that interaction with MeCP2 changes the sub-nuclear localization of DNMT3L. Next, we performed localization studies with the fluorescently tagged DNMT3A-ADD domain, which showed a diffuse nuclear localization (Fig. 6C). Similarly to DNMT3L, after co-expression with CFP-tagged MeCP2, the ADD domain showed a preferential localization to heterochromatic spots, indicating that MeCP2 targets it to heterochromatin (Fig. 6D). These results confirm that the ADD domain interacts with MeCP2 in cells and demonstrate that binding of MeCP2 recruits DNMT3L and DNMT3A-ADD to pericentromeric heterochromatin.

MeCP2 overexpression reduces DNA methylation in HCT116 cells

To study the effect of MeCP2 on the DNA methylation activity of DNMT3A in human cells, we used the HCT116 DNMT1 hypomorphic cell line, which contains a truncated DNMT1 with reduced activity, but active copies of DNMT3A and DNMT3B (51,52). Because of the impaired maintenance DNA methylation activity, these cells have an about 20% reduced amount of DNA methylation and show increased levels of hemimethylation. Due to the reduced activity of DNMT1, the DNA methylation in these cells is more dependent on the activity of DNMT3A and DNMT3B, making this cell line a suitable model system to study the effect of the inhibition of DNMT3A by MeCP2. After generating stable cells lines expressing EYFP-fused MeCP2 or EYFP as control for two months (Suppl. Fig. 6), genomic DNA was isolated and the global levels of 5-methylcytosine were quantified by LC-MS/MS. As shown in Fig. 7A, we observed a roughly twofold decrease in the global methylation of DNA isolated from two independent clones expressing EYFP-MeCP2. Since DNMT1 interaction with MeCP2 did not cause a reduction in catalytic activity (29), this result indicates that overexpression of MeCP2 reduces the activity of DNMT3 enzymes in cells.

Genome-wide bisulfite methylation analysis in brain tissue from MeCP2 KO mice reveals hyper- and hypomethylated DMRs

To study the influence of MeCP2 on DNA methylation in a biologically relevant context, we performed whole genome bisulfite sequencing (WGBS) on genomic DNA isolated from the whole brain of wild type and MeCP2 KO mice (68). The overlap in symptoms between the MeCP2 KO animals and Rett syndrome patients, makes these KO mice a disease-relevant model system. A global analysis of the WGBS data revealed similar levels of global DNA methylation (77%) for both samples. To assess whether the absence of MeCP2 leads to locus-specific changes in DNA methylation, we binned the data in 1kb tiles and identified differentially methylated regions (DMRs, see Methods).

We observed hypermethylated and hypomethylated DMR regions, indicating that the loss of MeCP2 can lead to both gain and loss of DNA methylation (Fig. 7B and C). Although the number of DMRs in WGBS was limited, the methylation changes were rather large, because out of 85 regions that lost methylation in the KO sample, 24 showed more than 50% reduction in methylation levels. In turn, out of 84 identified hypermethylated DMRs, 23 displayed more than 50% increase in methylation. Together, these data indicate that the chromatin regulator MeCP2 affects DNA methylation in a punctuate but notable manner. To check if the identified DMRs are sites of MeCP2 binding, we next used published MeCP2 high resolution ChIP-seq data from main olfactory neuroepithelia (59). We observed that 48% (41/85) of all hypomethylated DMRs and 63% (53/84) of all hypermethylated DMRs were associated with MeCP2 binding sites (statistically expected were <10%) (Fig. 7D). The highly significant enrichment of MeCP2 binding sites in both groups of DMRs, in particular in the hypermethylated DMRs, suggests that the observed methylation changes are direct effects of the loss of MeCP2 in the KO tissue. To determine whether the affected DMRs cluster with particular genomic features, we next analyzed their association with other chromatin elements (Fig. 7E). Next, we analyzed whether the affected DMRs correlate with any genetic and chromatin regulatory elements, showing that regions with gain in DNA methylation in the MeCP2 KO are enriched in gene promoters, in particular high CpG density promoters HCPs. By contrast, hypomethylated DMRs were moderately enriched in CGI and HCPs but not at low CpG density promoters (LCPs) (Fig. 7E). These data indicate that MeCP2 controls DNA methylation at a subset of CGIs and HCPs.

Discussion

During the past decade compelling experimental evidence has accumulated, indicating that DNA methylation patterns are highly dynamic and result from ongoing *de novo* methylation and demethylation events (3). This dynamic landscape plays particularly important roles in non-diving cells, such as terminally differentiated neurons (69-71). In the absence of cell division and DNA replication, the DNA methylation profiles in these cells are only controlled through a tight regulation of the targeting and activity of DNA methylating and demethylating enzymes. However, despite their importance, the regulatory processes of DNMTs are not yet fully understood. In this work, we took a closer look at the DNMT3 methyltransferases, factors that play essential roles in mammalian development and disease (1,2,72). We show that DNMT3 proteins directly and strongly interact with the chromatin regulator MeCP2, an important reader of 5mC and 5hmC, *in vitro* with recombinant proteins, in mammalian cell lines and by co-immunoprecipitation in brain lysates and that this interaction is mediated by the TRD domain of MeCP2 and the ADD domain of DNMT3 proteins. Based on the fact that MeCP2 is highly expressed in neurons and functions mainly in this cell type, we investigated whether the DNMT3A-MeCP2 interaction likely plays an important role in controlling DNA methylation patterns in the brain.

By employing *in vitro* methyltransferase assays using recombinant proteins and a variety of DNA substrates, we observed an almost complete, concentration-dependent inhibitory effect caused by MeCP2 binding to DNMT3A (Fig. 4D). Inhibition on both CpG and CpA substrates was obtained. DNMT3B activity was comparably reduced, indicating a conserved mode of action. Using conformationally locked DNMT3A variants as a novel tool to investigate DNMT3A regulation, we show that the inhibition of DNMT3A by MeCP2 occurs by an allosteric mechanism, in which binding of MeCP2 stabilizes the autoinhibitory conformation of DNMT3A (Fig. 1B and 5A). Interestingly, binding of the unmodified H3 N-terminal tail peptide to the ADD domain of DNMT3A was shown to have the opposite effect, by precluding the autoinhibitory conformation and leading to the activation of DNMT3A (12,13). We mechanistically addressed this cross talk and show that binding of H3 and TRD to DNMT3A are mutually exclusive and the MeCP2-mediated inhibition of DNMT3A2 can be overcome by addition of the unmodified H3 tail peptide.

To understand the effects of the MeCP2 on the function of DNMT3A *in vivo*, we have performed WGBS in the mouse brain, a tissue where MeCP2 and DNMT3A have important biological roles. Although these experiments are limited by the fact that relevant brain regions

or cell types and critical developmental stages may not be sampled, differences in DNA methylation were observed. We observed that MeCP2 knock-out leads to both hyper- and hypomethylation of DNA at specific 1kb tiles indicating that MeCP2 has both positive and negative effects on DNMT3A activity. These results are in agreement with the documented role of MeCP2 as both gene activator and repressor. The overall number of DMRs in WGBS was limited, which is in agreement with the dominant role of DNMT1 for maintenance methylation and the fact that DNMT1 is not influenced by MeCP2 (29). However, the methylation changes at the affected regions were rather large; around half of the discovered DMRs showed a change of more than 50% in methylation levels averaged over the 1kb tiles. Regions gaining DNA methylation in the MeCP2 KO are enriched in gene promoters, in particular HCPs, while hypomethylated DMRs were moderately enriched in CGI and HCPs. Overall, our data indicate that MeCP2 controls DNA methylation at a subset of CGIs and HCPs, which is particularly relevant, because expression of genes with HCPs is often regulated by DNA methylation (73,74). Interestingly, while WGBS demonstrated that MeCP2 functions as repressor and stimulator of DNA methylation at a rather small number of sites in mouse brain tissue, DNA methylation analysis after overexpression of MeCP2 in HCT116 DNMT1 hypomorphic cells resulted in a much more prominent global decrease in methylation. The stronger inhibition observed in the HCP116 cells can be explained by the larger contribution of the DNMT3 enzymes to overall methylation levels in this cell line.

In summary, our data suggest a model in which DNMT3A is under the combined control of MeCP2-mediated targeting and inhibition and the modification state of histone H3 tail at genomic target sites, unravelling one part of the intricate regulatory network, which controls DNA methylation. On the one hand, the interaction with MeCP2 inhibits DNMT3A activity and reduces DNA methylation at a subset of HCP gene promoters in the mouse brain and after overexpression of MeCP2 in a global manner in tissue culture (Fig. 8A). On the other hand, at defined CGIs and a subset of HCP promoters, MeCP2 can act as a recruiter of DNMT3A. As shown by our biochemical data, this leads to a relief of the allosteric inhibition of DNMT3A by the interaction of the DNMT3A ADD domain with appropriately modified histone tails, which do not contain activating marks. Therefore, the specific delivery of DNMT3A to such regions by MeCP2 can target DNA methylation (Fig. 8B). Afterwards, the elevated DNA methylation will increase the methylcytosine-dependent MeCP2 recruitment to these loci, initiating a positive feedback loop, which can contribute to the stable maintenance of methylation at these sites. In neurons, this process may be further supported by the non-CpG (mainly CpA) methylation introduced by DNMT3A which is bound by MeCP2 as well.

Funding

This work was supported by the DFG (Je 252/10-1 to AJ), by fellowships of the Carl Zeiss foundation to CL and RZJ and by the New York Stem Cell Foundation (AM) and NIH grant R01DA036898 (AM).

Acknowledgements

We are grateful to Sylke Lutz and Roland Kontermann (Institute of Cell Biology and Immunology, University Stuttgart) for providing some of the animal brain materials and to Dieter Wolf (Institute of Biochemistry, University Stuttgart) for providing the anti-rabbit light chain specific antibody. We also thank Ingo Amm and Nicole Berner (Institute of Biochemistry, University Stuttgart) for technical advice on the endogenous co-immunoprecipitation protocol and Benjamin Hackner (Department of Chemistry and Pharmacy, Ludwig-Maximilians-Universität München) for help with the LC-MS measurements. We are very grateful to the Central Facility for Advanced Microscopy of the Stuttgart Research Center Systems Biology at the University of Stuttgart, for providing access to the laser scanning microscope.

Author contributions

AJ, AR, RJ and CL devised the project and analyzed the data. CL and AR conducted the biochemical assays with contributions from IH. CL and AR performed the fluorescence microscopy experiments. CL performed the cell culture and biochemical work for the data shown in Fig. 7. MD contributed to the experiments with DNMT3B. EP and TC performed the LC-ESI-MS/MS. HG made the WGBS libraries, RK performed the analysis and interpreted the results with AM. All authors contributed to data interpretation and discussion, read and approved the final manuscript.

Conflict of interest

None declared.

References

1. Jurkowska, R.Z., Jurkowski, T.P. and Jeltsch, A. (2011) Structure and function of mammalian DNA methyltransferases. *Chembiochem : a European journal of chemical biology*, **12**, 206-222.
2. Bergman, Y. and Cedar, H. (2013) DNA methylation dynamics in health and disease. *Nature structural & molecular biology*, **20**, 274-281.
3. Jeltsch, A. and Jurkowska, R.Z. (2014) New concepts in DNA methylation. *Trends in biochemical sciences*, **39**, 310-318.
4. Jeltsch, A. (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem : a European journal of chemical biology*, **3**, 274-293.
5. Cheng, X. (1995) Structure and function of DNA methyltransferases. *Annual review of biophysics and biomolecular structure*, **24**, 293-318.
6. Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S. and Jeltsch, A. (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *The Journal of biological chemistry*, **285**, 26114-26120.
7. Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A. and Schubeler, D. (2015) Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*, **520**, 243-247.
8. Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D. *et al.* (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, **448**, 714-717.
9. Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M. and Shirakawa, M. (2009) Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO reports*, **10**, 1235-1241.
10. 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 research*, **38**, 4246-4253.
11. Rondelet, G., Dal Maso, T., Willems, L. and Wouters, J. (2016) Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. *Journal of structural biology*, **194**, 357-367.
12. Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G. *et al.* (2015) Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature*, **517**, 640-644.
13. Li, B.Z., Huang, Z., Cui, Q.Y., Song, X.H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E. and Xu, G.L. (2011) Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. *Cell research*, **21**, 1172-1181.
14. Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S. *et al.* (2015) In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *eLife*, **4**, e06205.
15. Noh, K.M., Wang, H., Kim, H.R., Wenderski, W., Fang, F., Li, C.H., Dewell, S., Hughes, S.H., Melnick, A.M., Patel, D.J. *et al.* (2015) Engineering of a Histone-Recognition Domain in Dnmt3a Alters the Epigenetic Landscape and Phenotypic Features of Mouse ESCs. *Molecular cell*, **59**, 89-103.
16. Long, H.K., Blackledge, N.P. and Klose, R.J. (2013) ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. *Biochemical Society transactions*, **41**, 727-740.
17. Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905-914.

18. Meehan, R.R., Lewis, J.D. and Bird, A.P. (1992) Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic acids research*, **20**, 5085-5092.
19. Nan, X., Tate, P., Li, E. and Bird, A. (1996) DNA methylation specifies chromosomal localization of MeCP2. *Molecular and cellular biology*, **16**, 414-421.
20. Ausio, J., de Paz, A.M. and Esteller, M. (2014) MeCP2: the long trip from a chromatin protein to neurological disorders. *Trends in molecular medicine*, **20**, 487-498.
21. Klose, R.J. and Bird, A.P. (2006) Genomic DNA methylation: the mark and its mediators. *Trends in biochemical sciences*, **31**, 89-97.
22. Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature genetics*, **19**, 187-191.
23. Klose, R.J., Sarraf, S.A., Schmiedeberg, L., McDermott, S.M., Stancheva, I. and Bird, A.P. (2005) DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Molecular cell*, **19**, 667-678.
24. Hansen, J.C., Ghosh, R.P. and Woodcock, C.L. (2010) Binding of the Rett syndrome protein, MeCP2, to methylated and unmethylated DNA and chromatin. *IUBMB life*, **62**, 732-738.
25. Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G. *et al.* (2014) Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature neuroscience*, **17**, 215-222.
26. Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H. and Greenberg, M.E. (2015) Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature*, **522**, 89-93.
27. Lyst, M.J. and Bird, A. (2015) Rett syndrome: a complex disorder with simple roots. *Nature reviews. Genetics*, **16**, 261-275.
28. Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P. and Kouzarides, T. (2003) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *The Journal of biological chemistry*, **278**, 4035-4040.
29. Kimura, H. and Shiota, K. (2003) Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. *The Journal of biological chemistry*, **278**, 4806-4812.
30. Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M.J., Shu, X., Kriaucionis, S. and Bird, A. (2007) Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 2709-2714.
31. Bienvenu, T. and Chelly, J. (2006) Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nature reviews. Genetics*, **7**, 415-426.
32. Guy, J., Cheval, H., Selfridge, J. and Bird, A. (2011) The role of MeCP2 in the brain. *Annual review of cell and developmental biology*, **27**, 631-652.
33. Hite, K.C., Adams, V.H. and Hansen, J.C. (2009) Recent advances in MeCP2 structure and function. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **87**, 219-227.
34. Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K. and Gage, F.H. (2010) L1 retrotransposition in neurons is modulated by MeCP2. *Nature*, **468**, 443-446.
35. Adkins, N.L. and Georgel, P.T. (2011) MeCP2: structure and function. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **89**, 1-11.
36. Chahrour, M., Jung, S.Y., Shaw, C., Zhou, X., Wong, S.T., Qin, J. and Zoghbi, H.Y. (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science*, **320**, 1224-1229.
37. Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, C.A. and Zoghbi, H.Y. (2009) Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. *Human molecular genetics*, **18**, 2431-2442.
38. Sugino, K., Hempel, C.M., Okaty, B.W., Arnson, H.A., Kato, S., Dani, V.S. and Nelson, S.B. (2014) Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long

genes. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **34**, 12877-12883.

39. Della Ragione, F., Vacca, M., Fioriniello, S., Pepe, G. and D'Esposito, M. (2016) MECP2, a multi-talented modulator of chromatin architecture. *Briefings in functional genomics*.
40. Feng, J., Chang, H., Li, E. and Fan, G. (2005) Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. *Journal of neuroscience research*, **79**, 734-746.
41. Nguyen, S., Meletis, K., Fu, D., Jhaveri, S. and Jaenisch, R. (2007) Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan. *Developmental dynamics : an official publication of the American Association of Anatomists*, **236**, 1663-1676.
42. Feng, J., Zhou, Y., Campbell, S.L., Le, T., Li, E., Sweatt, J.D., Silva, A.J. and Fan, G. (2010) Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nature neuroscience*, **13**, 423-430.
43. Morris, M.J., Adachi, M., Na, E.S. and Monteggia, L.M. (2014) Selective role for DNMT3a in learning and memory. *Neurobiology of learning and memory*, **115**, 30-37.
44. Jeltsch, A. and Lanio, T. (2002) Site-directed mutagenesis by polymerase chain reaction. *Methods Mol Biol*, **182**, 85-94.
45. Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A. and Cheng, X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, **449**, 248-251.
46. Jurkowska, R.Z., Anspach, N., Urbanke, C., Jia, D., Reinhardt, R., Nellen, W., Cheng, X. and Jeltsch, A. (2008) Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic acids research*, **36**, 6656-6663.
47. Rajavelu, A., Jurkowska, R.Z., Fritz, J. and Jeltsch, A. (2012) Function and disruption of DNA methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic acids research*, **40**, 569-580.
48. Ebert, D.H., Gabel, H.W., Robinson, N.D., Kastan, N.R., Hu, L.S., Cohen, S., Navarro, A.J., Lyst, M.J., Ekiert, R., Bird, A.P. *et al.* (2013) Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. *Nature*, **499**, 341-345.
49. Jurkowska, R.Z., Siddique, A.N., Jurkowski, T.P. and Jeltsch, A. (2011) Approaches to enzyme and substrate design of the murine Dnmt3a DNA methyltransferase. *Chembiochem : a European journal of chemical biology*, **12**, 1589-1594.
50. Roth, M. and Jeltsch, A. (2000) Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biological chemistry*, **381**, 269-272.
51. Rhee, I., Bachman, K.E., Park, B.H., Jair, K.W., Yen, R.W., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W. *et al.* (2002) DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, **416**, 552-556.
52. Egger, G., Jeong, S., Escobar, S.G., Cortez, C.C., Li, T.W., Saito, Y., Yoo, C.B., Jones, P.A. and Liang, G. (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 14080-14085.
53. Schiesser, S., Pfaffeneder, T., Sadeghian, K., Hackner, B., Steigenberger, B., Schroder, A.S., Steinbacher, J., Kashiwazaki, G., Hofner, G., Wanner, K.T. *et al.* (2013) Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine. *Journal of the American Chemical Society*, **135**, 14593-14599.
54. Bashtrykov, P., Rajavelu, A., Hackner, B., Ragozin, S., Carell, T. and Jeltsch, A. (2014) Targeted mutagenesis results in an activation of DNA methyltransferase 1 and confirms an autoinhibitory role of its RFTS domain. *Chembiochem : a European journal of chemical biology*, **15**, 743-748.
55. Xi, Y. and Li, W. (2009) BSMAP: whole genome bisulfite sequence MAPping program. *BMC bioinformatics*, **10**, 232.

56. Sun, D., Xi, Y., Rodriguez, B., Park, H.J., Tong, P., Meong, M., Goodell, M.A. and Li, W. (2014) MOABS: model based analysis of bisulfite sequencing data. *Genome biology*, **15**, R38.
57. Bland, J.M. and Kerry, S.M. (1998) Statistics notes. Weighted comparison of means. *BMJ*, **316**, 129.
58. Storey, J.D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9440-9445.
59. Rube, H.T., Lee, W., Hejna, M., Chen, H., Yasui, D.H., Hess, J.F., LaSalle, J.M., Song, J.S. and Gong, Q. (2016) Sequence features accurately predict genome-wide MeCP2 binding in vivo. *Nat Commun*, **7**, 11025.
60. Chen, T., Ueda, Y., Xie, S. and Li, E. (2002) A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *The Journal of biological chemistry*, **277**, 38746-38754.
61. Gowher, H. and Jeltsch, A. (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *Journal of molecular biology*, **309**, 1201-1208.
62. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M. *et al.* (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, **462**, 315-322.
63. Arand, J., Spieler, D., Karius, T., Branco, M.R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V. *et al.* (2012) In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS genetics*, **8**, e1002750.
64. Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D. *et al.* (2013) Global epigenomic reconfiguration during mammalian brain development. *Science*, **341**, 1237905.
65. Ghosh, R.P., Nikitina, T., Horowitz-Scherer, R.A., Gierasch, L.M., Uversky, V.N., Hite, K., Hansen, J.C. and Woodcock, C.L. (2010) Unique physical properties and interactions of the domains of methylated DNA binding protein 2. *Biochemistry*, **49**, 4395-4410.
66. Baker, S.A., Chen, L., Wilkins, A.D., Yu, P., Lichtarge, O. and Zoghbi, H.Y. (2013) An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell*, **152**, 984-996.
67. Jurkowska, R.Z., Rajavelu, A., Anspach, N., Urbanke, C., Jankevicius, G., Ragozin, S., Nellen, W. and Jeltsch, A. (2011) Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: heterochromatic localization and role of Dnmt3L. *The Journal of biological chemistry*, **286**, 24200-24207.
68. Guy, J., Hendrich, B., Holmes, M., Martin, J.E. and Bird, A. (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nature genetics*, **27**, 322-326.
69. Heyward, F.D. and Sweatt, J.D. (2015) DNA Methylation in Memory Formation: Emerging Insights. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*.
70. Weaver, I.C. (2014) Integrating early life experience, gene expression, brain development, and emergent phenotypes: unraveling the thread of nature via nurture. *Advances in genetics*, **86**, 277-307.
71. Shin, J., Ming, G.L. and Song, H. (2014) DNA modifications in the mammalian brain. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **369**.
72. Yang, L., Rau, R. and Goodell, M.A. (2015) DNMT3A in haematological malignancies. *Nature reviews. Cancer*, **15**, 152-165.
73. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature genetics*, **39**, 457-466.
74. Rauch, T.A., Zhong, X., Wu, X., Wang, M., Kernstine, K.H., Wang, Z., Riggs, A.D. and Pfeifer, G.P. (2008) High-resolution mapping of DNA hypermethylation and hypomethylation in lung

cancer. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 252-257.

Figures and figure legends

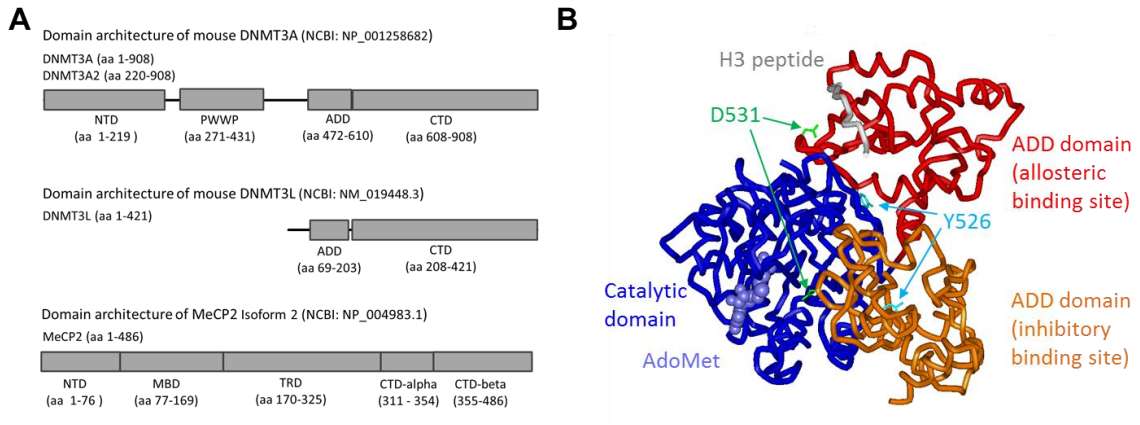


Fig. 1: Structures of DNMT proteins and MeCP2. A) Domain architecture of mouse DNMT3A, DNMT3L and MeCP2 and domain boundaries used in this work. B) Model of the structure of a DNMT3A fragment comprising ADD (red or orange) and catalytic domain (blue). The ADD domain can bind to the catalytic domain at two interfaces, at an allosteric site (ADD domain colored in red, Y526 is at the interface), which is stabilized by binding of the H3 tail, and at an inhibitory binding site (ADD domain colored in orange, D531 is at the interface). The image was generated using pdb files 4U7P and 4U7T (12).

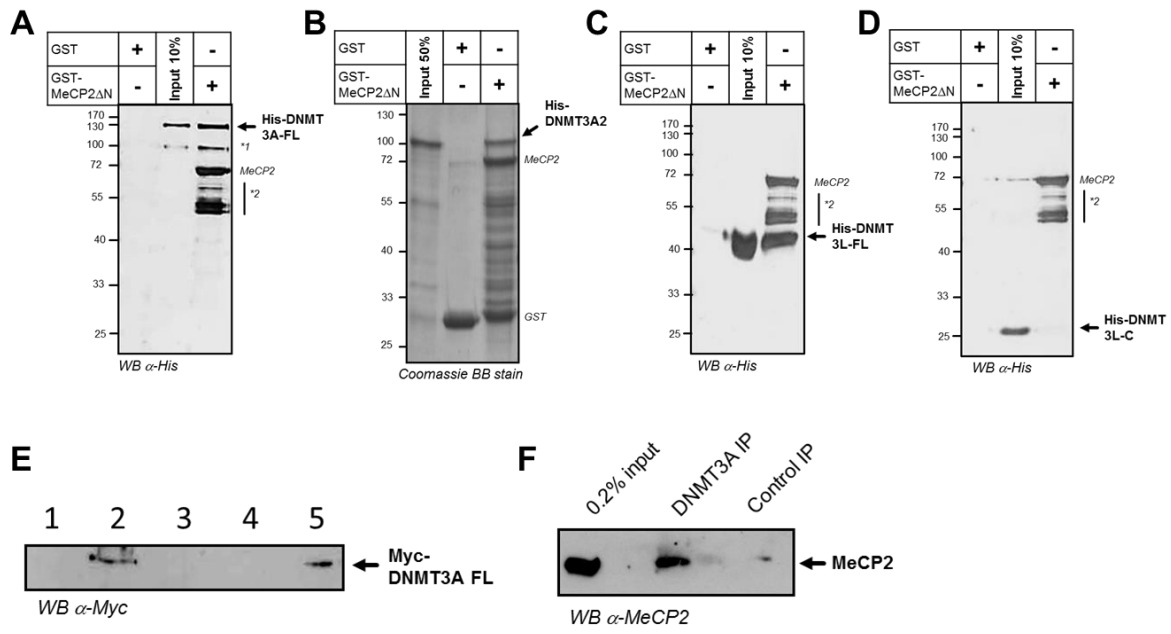


Fig. 2: GST pull-down experiments with DNMT3 proteins and MeCP2. A) Pull-down of His-DNMT3A by GST tagged N-terminal truncated MeCP2 (MeCP2ΔN, residues 104-486). Note that MeCP2 carries an endogenous His-tag. The bands labelled with *1 and *2 correspond to degraded DNMT3A and MeCP2ΔN, respectively. B) Pull-down of His-DNMT3A2 by GST tagged N-terminal truncated MeCP2 (MeCP2ΔN, residues 104-486). C) Pull-down of His-DNMT3L by GST tagged MeCP2ΔN. D) Pull-down of His-tagged C-terminal domain of DNMT3L by GST tagged MeCP2ΔN showing absence of interaction. E) Pull-down of Myc-tagged DNMT3A by EYFP-tagged MeCP2 after transient co-expression of both proteins in human HEK293 cells. Lane 1: α-YFP pulldown after co-expression of EYFP and Myc-DNMT3A. Lane 2: α-YFP pulldown after co-expression of EYFP-MeCP2 and Myc-DNMT3A. Lane 3: α-YFP pulldown after expression of EYFP-MeCP2. Lane 4: empty lane. Lane 5: 10% input of lane 2. F) Pull-down of endogenous MeCP2 by immunoprecipitation of endogenous DNMT3A using mouse brain protein extracts. An immunoprecipitation using α-Myc antibody was conducted as control.

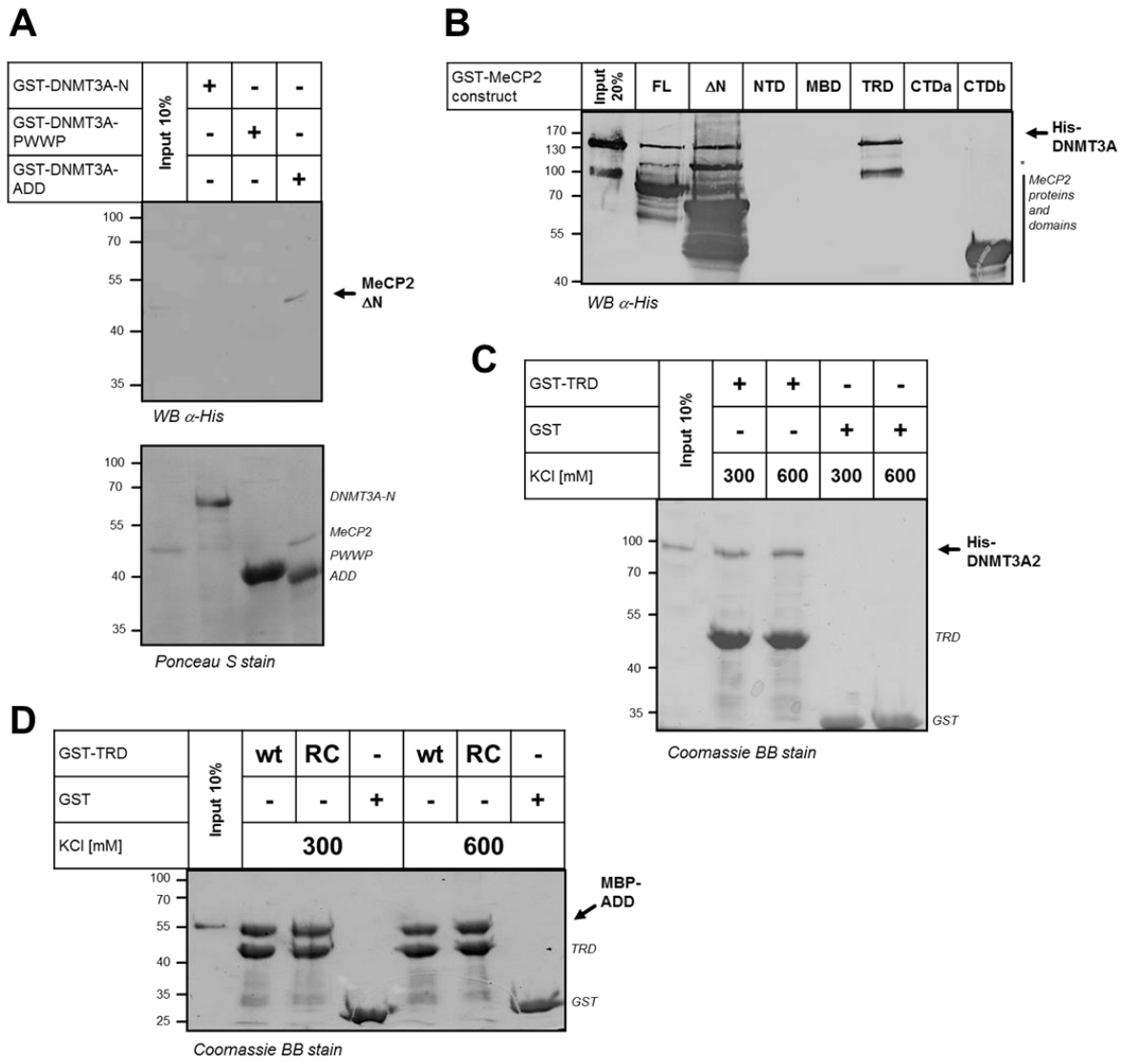


Fig. 3: Mapping of the DNMT3A-MeCP2 interaction. A) Pull-down of MeCP2ΔN with different GST-tagged DNMT3A domains, showing an interaction with the ADD domain. B) Pull-down of DNMT3A with different GST-tagged MeCP2 domains showing interaction with the TRD domain. Note that MeCP2 contains an endogenous His-tag in its CTDb domain. C) Pull-down of His-DNMT3A2 with GST-tagged TRD domain. Efficient interaction could be observed even in the presence of 600 mM KCl. D) Pull-down of MBP-tagged ADD domain by GST-tagged TRD domain (wt) or TRD domain containing the R306C Rett mutation (RC).

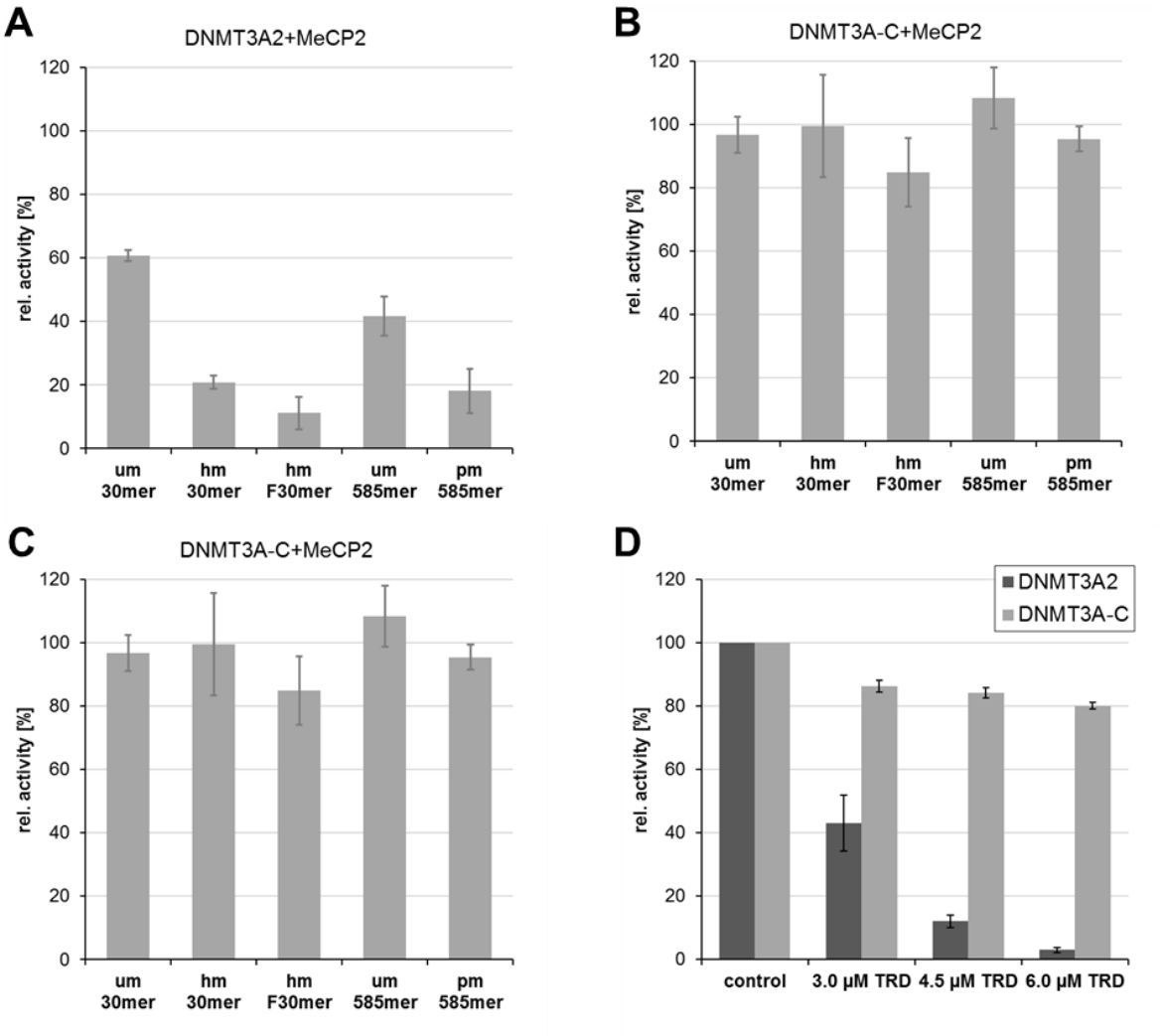


Fig. 4: Inhibition of DNMT3A2 by addition of MeCP2 (A and B) or MeCP2 TRD domain (C and D). In A-C) 2.5 µM DNMT3A2 or DNMT3A-C was used and 3 µM MeCP2 or TRD domain were added. Identical control reaction without addition of MeCP2 were used to calculate relative activities. Different DNA substrates were used as indicated. D) Inhibition of DNMT3A2 by addition of increasing concentrations of TRD using the um30mer substrate. Control refers to a reaction without added TRD. All error bars indicate the SEM based on two independent experiments.

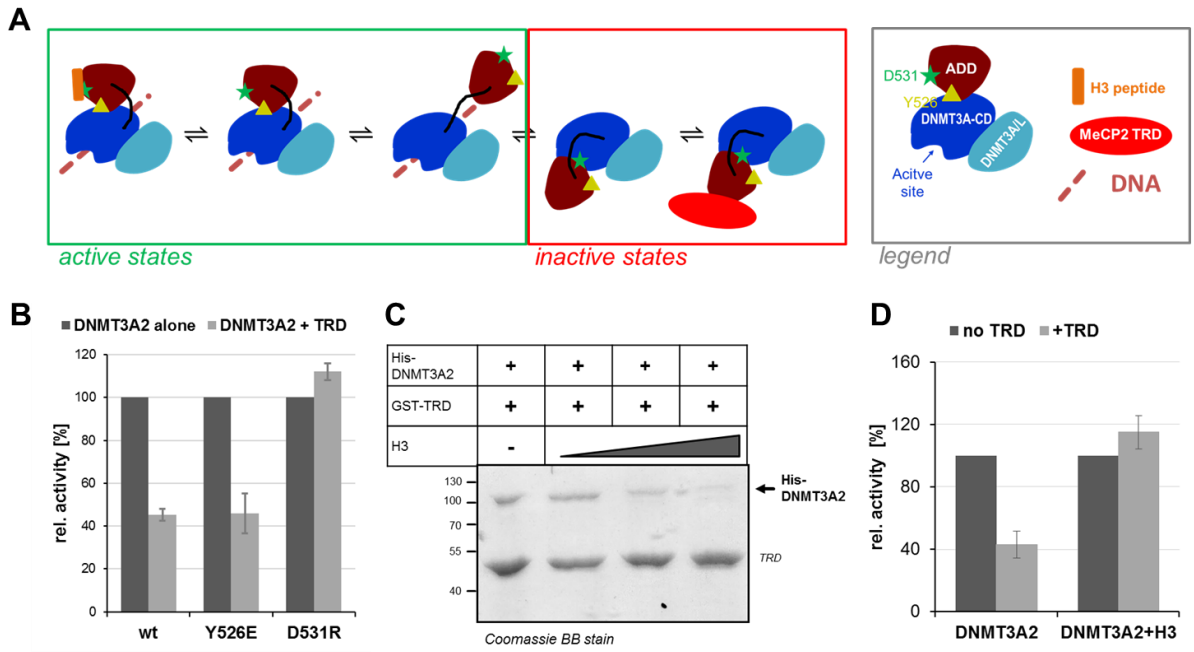


Fig. 5: Mechanism of the inhibition of DNMT3A by the TRD domain of MeCP2. A) Schematic picture of the conformational changes of DNMT3A showing that the ADD domain can bind to the catalytic domain (CD) at an allosteric site in which DNMT3A is active. The allosteric conformation is stabilized by Y526 indicated by a yellow triangle. Binding of the H3 peptide (orange) to the ADD domain stabilizes the allosteric binding mode (12,13). ADD can also bind to CD at an autoinhibitory site, which renders DNMT3A inactive. The autoinhibitory conformation is stabilized by D531 indicated by a green asterisk. Binding of MeCP2 to the ADD domain of DNMT3A via its TRD domain stabilizes the autoinhibitory binding mode, leading to inhibition of DNMT3A. B) Inhibition of Dnmt3a2 wild-type (wt) and its conformational variants by TRD. The inhibition is lost in the D531R variant carrying a mutation, which prevents the autoinhibitory conformation. Error bars indicate the SEM based on three independent experiments. C) Pull-down of His-DNMT3A2 (0.25 µM) by GST-TRD in the presence of increasing concentrations of recombinant histone H3 (0, 0.26, 1.3 and 3.9 µM) indicating that the H3 binding to DNMT3A2 interferes with the TRD interaction. D) DNA methylation activities of DNMT3A2 and DNMT3A2 bound to H3 peptide (amino acid sequence 1-19, 25 µM) in the absence and presence of TRD (3 µM) showing the loss of TRD inhibition in the presence of H3 peptide. Error bars indicate the SEM based on three independent experiments.

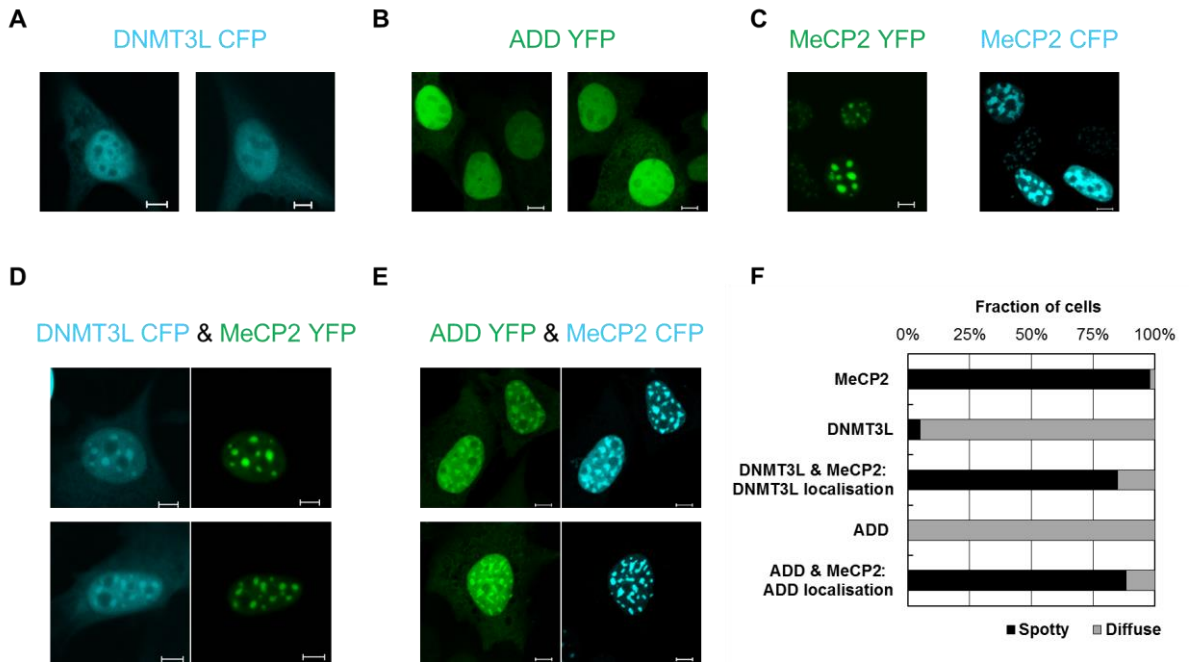


Fig. 6: Cellular localization of fluorescently tagged DNMT3L, ADD domain and MeCP2 after transient expression in NIH3T3 cells. A) DNMT3L shows a homogenous nuclear distribution (67). B) The YPF-ADD domain of DNMT3A shows a homogenous nuclear distribution. C) YFP and CFP tagged MeCP2 localizes to defined foci, which correspond to heterochromatic DAPI dense regions (19,48). D) DNMT3L is recruited to heterochromatic foci when co-expressed with MeCP2. E) The DNMT3A ADD domain is recruited to heterochromatic foci when co-expressed with MeCP2. E) Fraction of cells showing spotty and diffuse localization patterns in the experiments shown in panels A-E (based on analysis of >20 individual cells in each case). See also Suppl. Fig. 10.

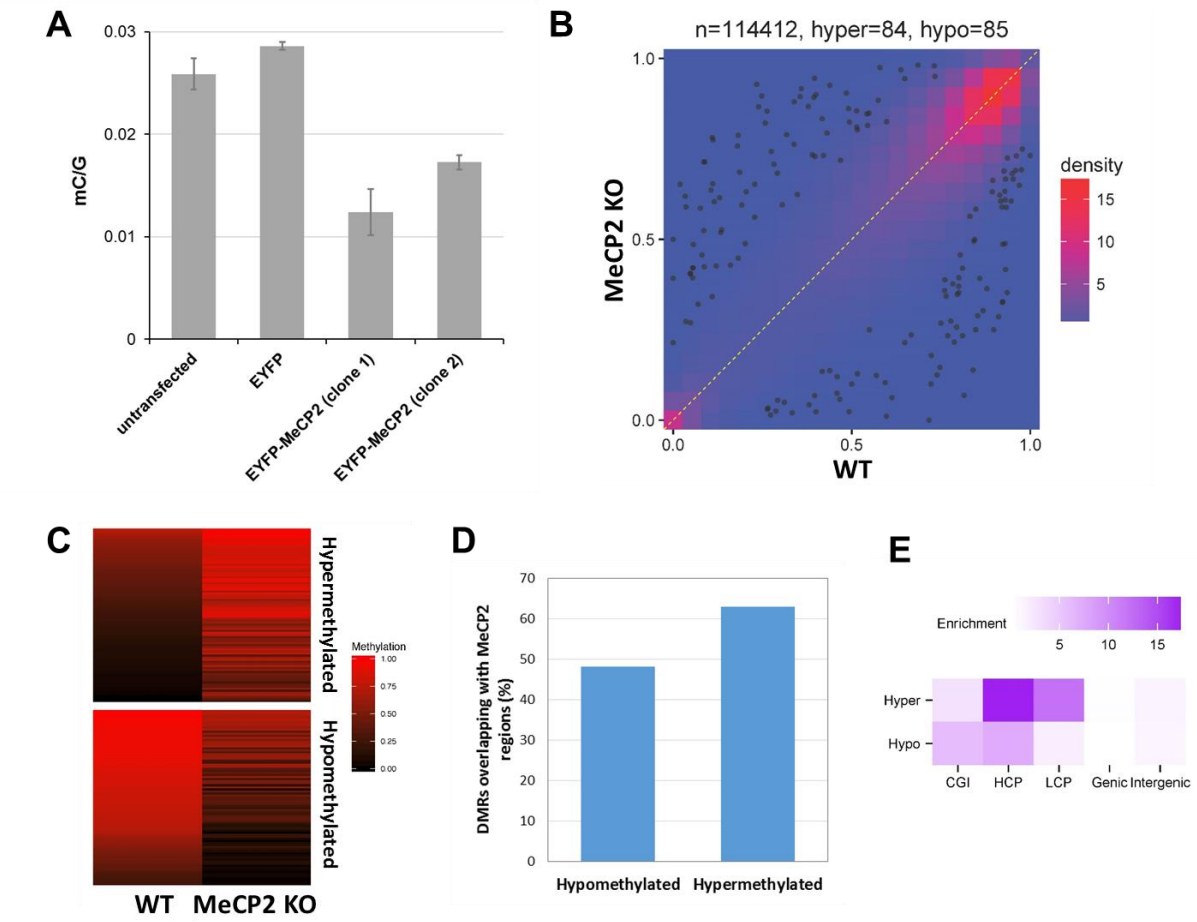


Fig. 7: Effect of MeCP2 on DNA methylation in cells. A) Global DNA methylation levels after expression of MeCP2 in human HCT116 cells containing a DNMT1 hypomorphic allele. Different clonal populations expressing MeCP2 showed an approximately 2-fold reduction of global DNA methylation. The error bars represent the SEM based on three technical repeats. B-E) Results of whole genome bisulfite methylation analysis in mouse MeCP2 KO brain tissue. B) Two dimensional density plot of methylation levels at nonrepetitive 1kb genomic tiles in KO and WT tissues. The DMRs are indicated by black dots. C) Heatmap of the methylation levels of nonrepetitive 1kb DMRs in KO tissue. D) Overlap of hyper- and hypomethylated DMRs with MeCP2 binding site regions. E) Enrichment of selected genomic features at hyper- and hypomethylated DMRs (CGI, CpG islands; HCP, high CpG density promoters; LCP, low CpG density promoters).

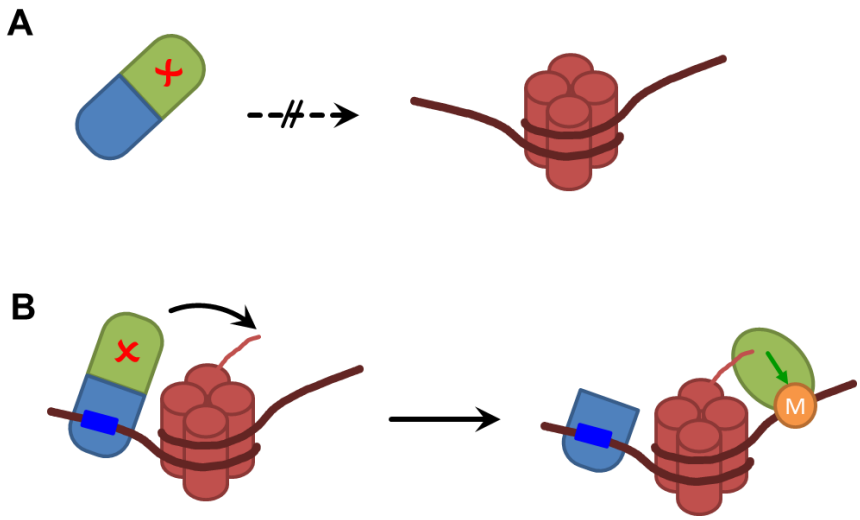


Fig. 8: Model of the dual role of MeCP2 in the regulation and targeting of DNMT3A. A) Binding of MeCP2 (blue) to DNMT3A (green) inactivates the methyltransferase and prevents untargeted activity (upper scheme). B) After MeCP2 binding to DNA at a genomic locus with appropriately modified H3 tail, H3 tail binding to DNMT3A leads to the dissociation of the DNMT3A-MeCP2 complex and subsequent DNA methylation. This can trigger more MeCP2 binding and by this initiate a positive feedback mechanism.

Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2

Arumugam Rajavelu[1,2,#], Cristiana Lungu[1,#], Ines Hanelt[1], Michael Dukatz[1], Edris Parsa[3], Rahul Karnik[4], Hongcang Gu[4], Thomas Carell[3], Alexander Meissner[4], Renata Z. Jurkowska[1,5] & Albert Jeltsch[1,*]

[#] These authors contributed equally to the work

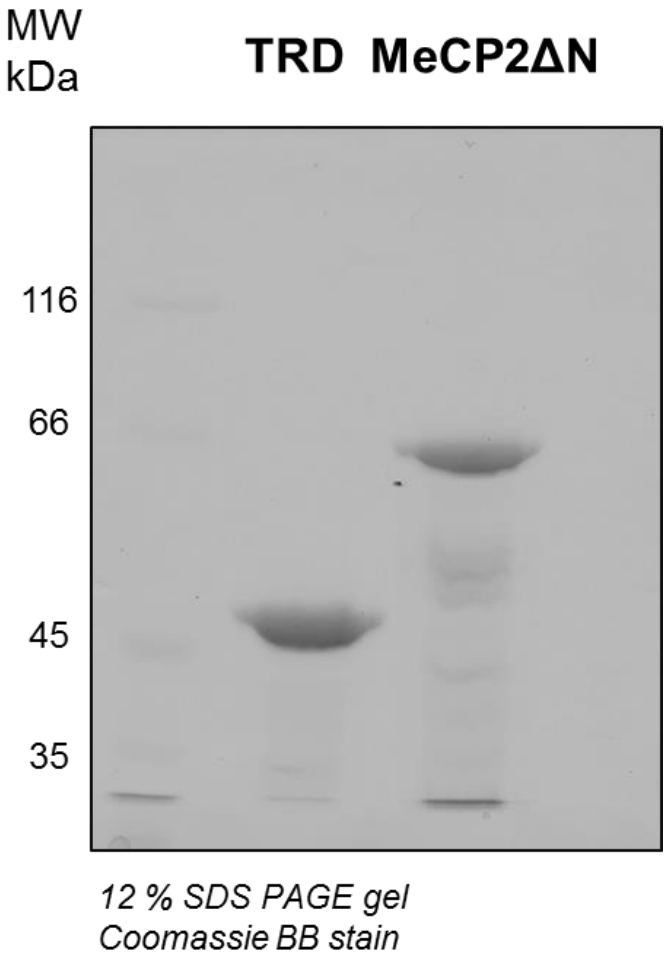
Supplemental information

Supplemental Text 1. Literature examples of the interaction of epigenetic readers and writers.

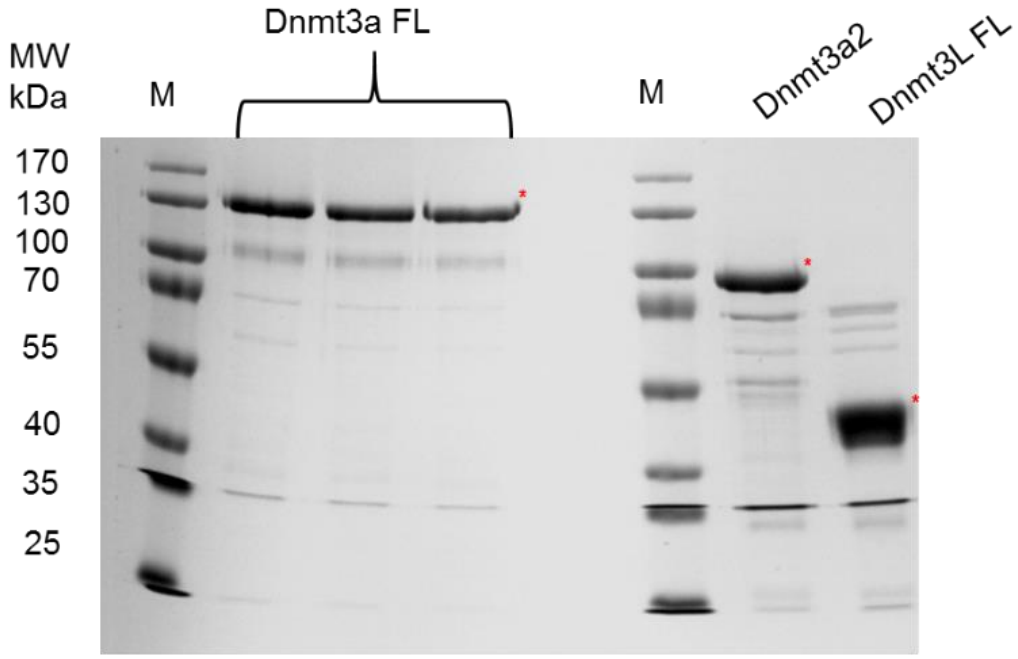
Previous studies have shown that readers of epigenetic marks often interact with enzymes which introduce the same mark. For example, HP1 which reads H3K9me2/3 interacts with SUV39H1/H2 (an H3K9 PKMT) (Schotta et al., 2002), G9a (an H3K9 PKMT) contains an ankyrin repeat domain which binds to H3K9me2 (Collins et al., 2008), and EZH2 (an H3K27 PKMT) forms a complex with EED, which is a K27 methylation reader (Margueron et al., 2009). Similarly, MeCP2, which binds methylated DNA interacts with DNMT1 (Kimura and Shiota, 2003), and UHRF1, which binds hemimethylated DNA with its SRA domain (Achour et al., 2008; Arita et al., 2008; Hashimoto et al., 2008), interacts with DNMT1 as well (Bostick et al., 2007; Sharif et al., 2007) and stimulate its methyltransferase activity by an allosteric mechanism (Bashtrykov et al., 2014; Berkyurek et al., 2014).

Supplemental Figures 1-8

Suppl. Fig. 1: Purification of GST tagged N-terminally truncated MeCP2 and GST-TRD domain proteins. The image shows examples of the purified proteins loaded on SDS polyacrylamide gels and stained with Coomassie BB.

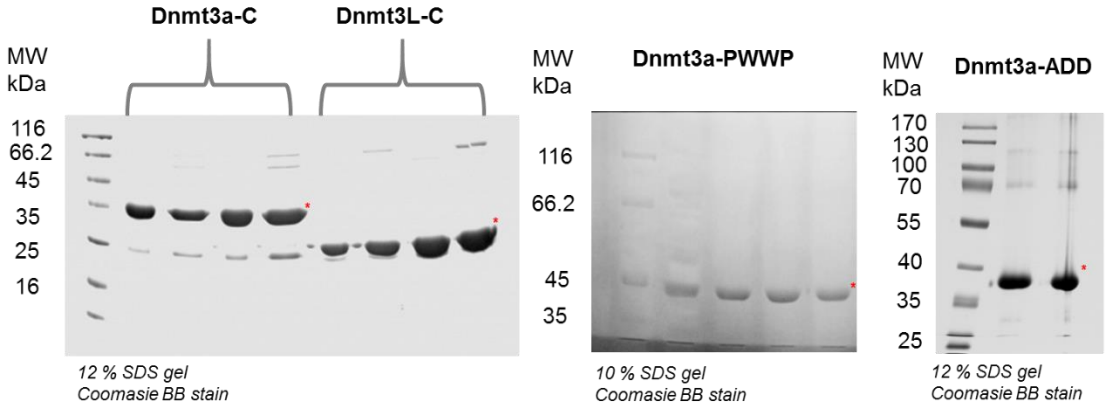


Suppl. Fig. 2: Purification of Dnmt3 proteins. The image shows examples of the purified proteins loaded on SDS polyacrylamide gels and stained with Coomassie BB.

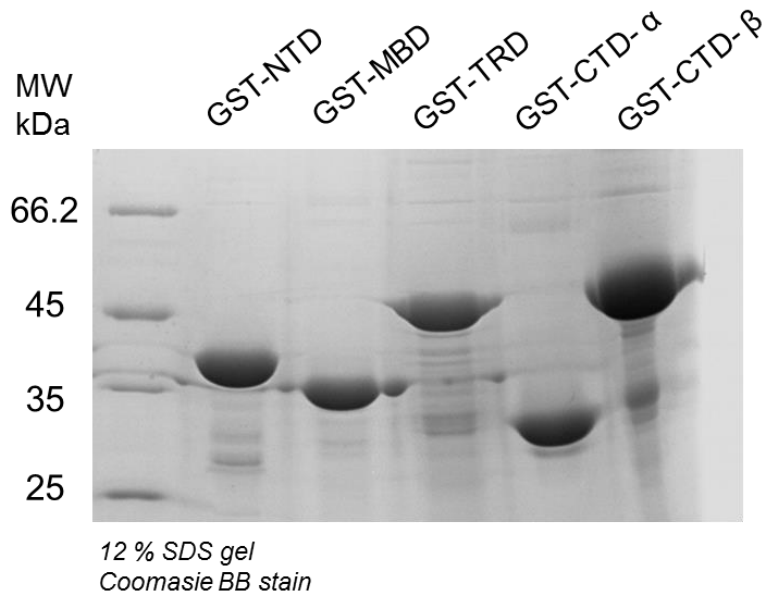


12 % SDS PAGE gel
Coomassie BB stain

Suppl. Fig. 3: Purification of Dnmt3 domains. The image shows examples of the purified proteins loaded on SDS polyacrylamide gels and stained with Coomassie BB.



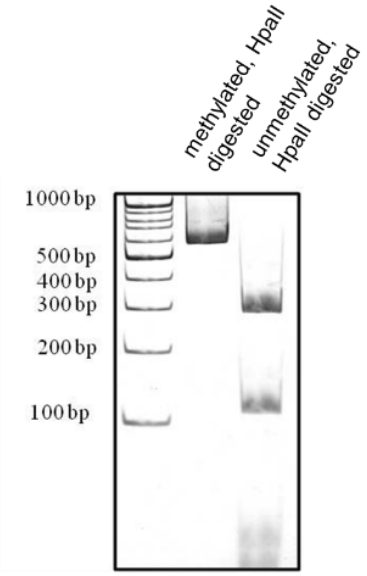
Suppl. Fig. 4: Purification of MeCP2 domains. The image shows examples of the purified proteins loaded on SDS polyacrylamide gels and stained with Coomassie BB.



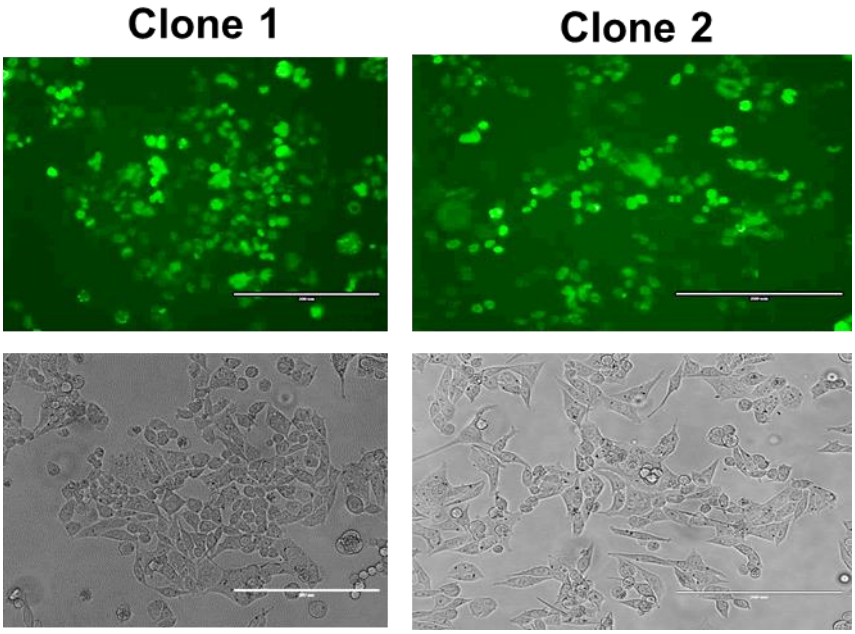
Suppl. Fig. 5: Sequence of the 585mer (left part) and HpaII digestion of unmethylated and pre-methylated 585mer. Samples were loaded on an agarose gel and stained with Ethidium bromide.

585mer Sequence

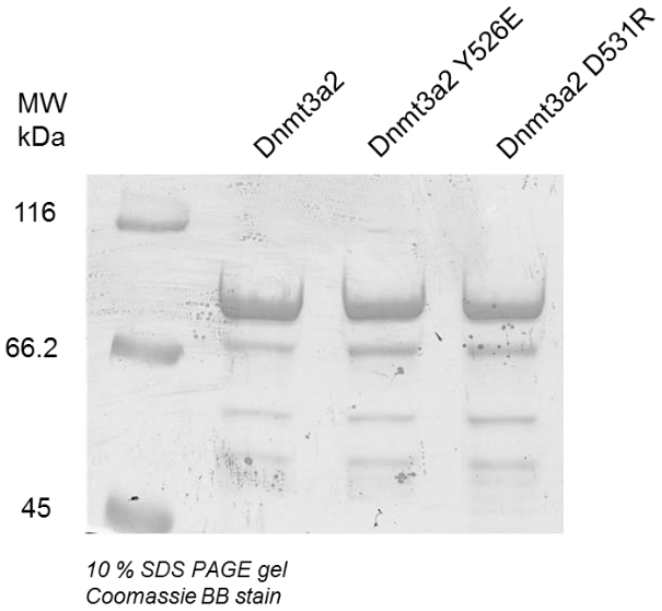
GGTAAAGGAAGCAGTAAGGGGCATACCCCGCGCGAAG
CGAAGGACAACCTGAAGTCCACGCAGTTGCTGAGTGT
GATCGATGCCATCAGCGAAGGGCCGATTGAAGGTCCG
GTGGATGGCTTAAAAAGCGTGCTGCTGAACAGTACGC
CGGTGCTGGACACTGAGGGGAATACCAACATATCCGG
TGTCACGGTGGTGTTCCGGGCTGGTGAGCAGGAGCAG
ACTCCGCCGGAGGGATTTGAATCCTCCGGCTCCGAGA
CGGTGCTGGGTACGGAAGTGAAATATGACACGCCGAT
CACCCGCACCATTACGTCTGCAAACATCGACCGTCTG
CGCTTTACCTTCGGTGTACAGGCACTGGTGGAAACCA
CCTCAAAGGGTGACAGGAATCCGTCGGAAGTCCGCCT
GCTGGTTCAGATACAACGTAACGGTGGCTGGGTGACG
GAAAAAGACATCACCATTAAGGGCAAAACCACCTCGC
AGTATCTGGCCTCGGTGGTGATGGGTAACCTGCCGCC
GCGCCCGTTTAATATCCGGATGCGCAGGATGACGCCG
GACAGCACCACAGACCAGCTGCAGAACAAA



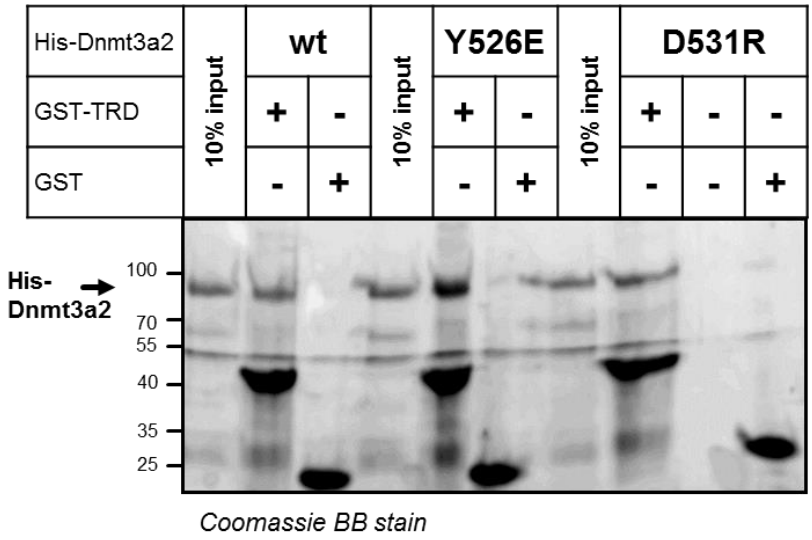
Suppl. Fig. 6: Images of the HCT116 cell lines with stable expression of EYFP-MeCP2. Scale bar: 200 µm.



Suppl. Fig. 7: Purification of Dnmt3 mutants. The image shows examples of the purified proteins loaded on SDS polyacrylamide gels and stained with Coomassie BB.

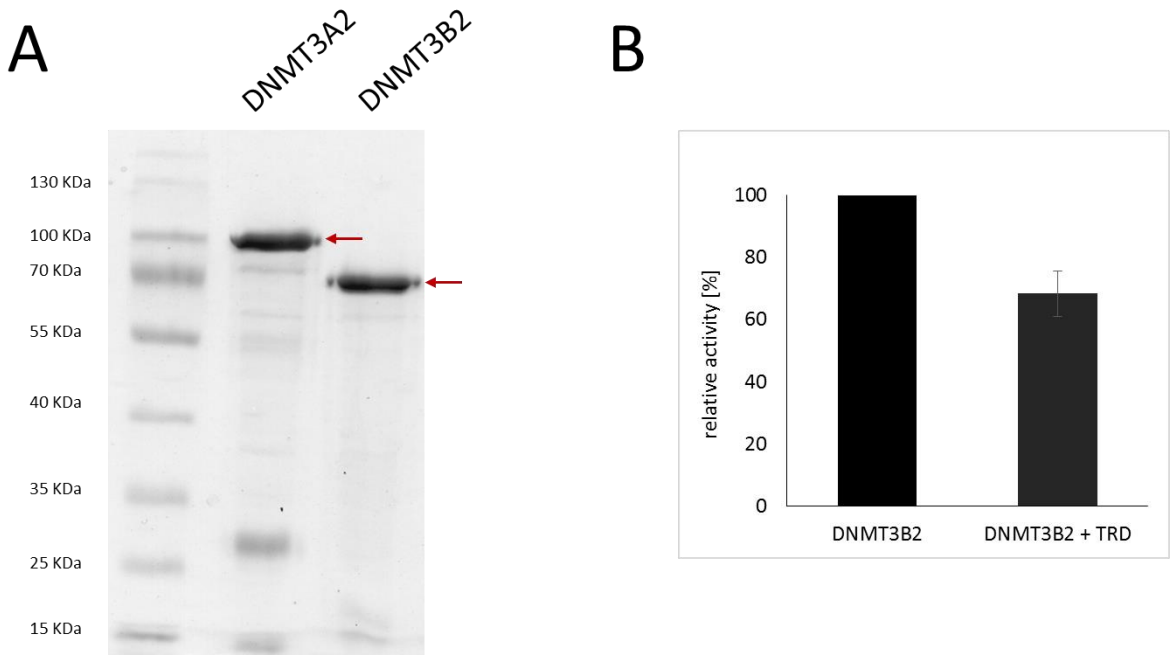


Suppl. Fig. 8: Pulldown of His-Dnmt3a2 wild type and mutants by GST-tagged TRD showing that the Dnmt3a2 Y526E and D531R mutants still interact with TRD.

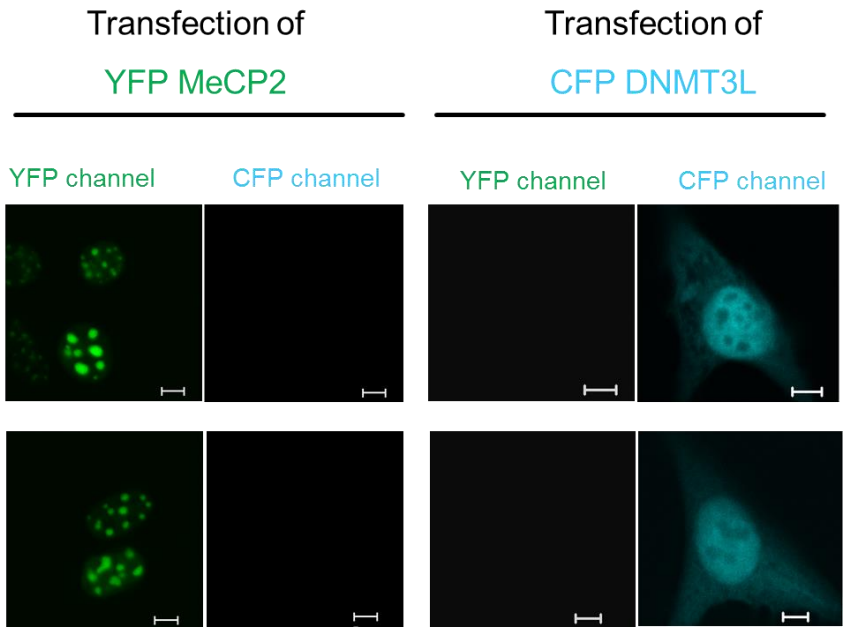


Coomassie BB stain

Suppl. Fig. 9: Inhibition of DNMT3B2 by MeCP2 TRD. The DNMT3B2 is a truncated form of DNMT3B cloned to correspond to DNMT3A2. A) Coomassie-stained SDS-PAGE of the purified DNMT3A2 and DNMT3B2. The expected size of the full-length proteins is indicated with red arrows. Both of the protein preps show comparable high purity. B) Enzymatic activity assay with DNMT3B2 in the absence or presence of the TRD domain of MeCP2. The activity of DNMT3B2 in the absence of TRD was normalized to 100%. The error bar represents standard error of the mean for three independent experiments.



Suppl. Fig. 10: Absence of crosstalk between the YFP and CFP channels in fluorescence microscopy. YFP tagged MeCP2 and CFP tagged DNMT3L were individually transfected into NIH3T3 cells and images were captures in the YFP and CFP channels.



Supplemental References

Achour, M., Jacq, X., Ronde, P., Alhosin, M., Charlot, C., Chataigneau, T., Jeanblanc, M., Macaluso, M., Giordano, A., Hughes, A.D., *et al.* (2008). The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. Oncogene *27*, 2187-2197.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature *455*, 818-821.

Bashtrykov, P., Jankevicius, G., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2014). The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism. The Journal of biological chemistry *289*, 4106-4115.

Berkyurek, A.C., Suetake, I., Arita, K., Takeshita, K., Nakagawa, A., Shirakawa, M., and Tajima, S. (2014). The DNA methyltransferase Dnmt1 directly interacts with the SET and RING finger-associated (SRA) domain of the multifunctional protein Uhrf1 to facilitate accession of the catalytic center to hemi-methylated DNA. The Journal of biological chemistry *289*, 379-386.

Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760-1764.

Collins, R.E., Northrop, J.P., Horton, J.R., Lee, D.Y., Zhang, X., Stallcup, M.R., and Cheng, X. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nature structural & molecular biology *15*, 245-250.

Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E., and Cheng, X. (2008). The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature *455*, 826-829.

Kimura, H., and Shiota, K. (2003). Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. The Journal of biological chemistry *278*, 4806-4812.

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

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. The EMBO journal *21*, 1121-1131.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., *et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Appendix 4

TLB GmbH · Ettlinger Straße 25 · D-76137 Karlsruhe, Germany

Technologie-Lizenz-Büro
der Baden-Württembergischen Hochschulen GmbH

Universität Stuttgart
Institut für Biochemie
Frau Cristiana Lungu MSc.
Pfaffenwaldring 55
70569 Stuttgart

Ettlinger Straße 25
76137 Karlsruhe, Germany
Telefon +49 721 79004-0
Telefax +49 721 79004-79

info@tlb.de · www.tlb.de

Ihr Zeichen: 0521.9-12/1188

24.02.2016

Unser Zeichen: 15/068TLB-si

Patentanmeldung zur Erfindungsmeldung „Einsatz von DNA Methyltransferase Inhibitoren zur Behandlung des Rett Syndroms (Rett-DNMTi)“

Sehr geehrte Frau Lungu,

die Patentanmeldung in Europa zu o. g. Erfindungsmeldung wurde inzwischen eingereicht. Die Patentanmeldung wird beim Patentamt unter dem Aktenzeichen EP 16 156 521.3 geführt. Sie trägt den Titel: „Dnmt inhibitors for Rett syndrome therapy“. Anmeldetag ist der 19.02.2016.

Eine Kopie der eingereichten Patentanmeldung haben wir Ihnen für Ihre Unterlagen beigelegt.

Mit freundlichen Grüßen

---

i. A. Monika Meister
TLB-Sekretariat

Anlage
Patentanmeldung