Biochemical analysis of DNA- and protein methyltransferases using recombinant designer nucleosomes

Von der Fakultät 3: Chemie der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

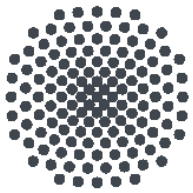
Alexander Bröhm

aus Backnang

Hauptberichter:	Prof. Dr. Albert Jeltsch
Mitberichter:	Prof. Dr. Jörn Lausen
Prüfungsvorsitzender:	Prof. Dr. Andreas Köhn

Tag der mündlichen Prüfung:

30.05.2022



Universität Stuttgart

Institut für Biochemie und technische Biochemie

2022

Seeing, contrary to popular wisdom, isn't believing. It's where belief stops because it isn't needed anymore. – Terry Pratchett

The cosmos is within us. We are made of star-stuff. We are a way for the universe to know itself. – Carl Sagan

Erklärung über die Eigenständigkeit der Dissertation

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Titel

Biochemical analysis of DNA- and protein methyltransferases using recombinant designer nucleosomes

selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe; aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht.

Declaration of Authorship

I hereby certify that the dissertation entitled

Biochemical analysis of DNA- and protein methyltransferases using recombinant designer nucleosomes

is entirely my own work except where otherwise indicated. Passages and ideas from other sources have been clearly indicated.

Alexander Bröhm

Stuttgart, 06.04.2022

Table of contents

Acknowledgements

First of all, I wish to thank Prof. Albert Jeltsch for the opportunity to conduct my PhD studies in his department and for the continuous support and inspiration.

I'm grateful to Prof. Jörn Lausen and Prof. Andreas Köhn for participating in my examination committee.

Furthermore, my thanks go to the PKMT group, especially my former supervisors Srikanth Kudithipudi and Sara Weirich, as well as to Mina, TC, and Maren.

I also wish to thank all students which performed their bachelor and master studies under my supervision: Tabea, Evelin, Franzi, David, and Tabea again.

Of course, my gratitude goes to my present and former colleagues, including our permanent employees: Hans for the exciting discussions, Elisabeth for saving the institute from chaos, Regina for being compassionate, Dragica (Schatzi!), and Branka. Among us PhD students, special thanks goes to the old-timers: Rebekka for being herself, Micha for being the certified nicest guy ever, Franzi for the drinks, Max for being the best handyman buddy, Sabine for the best concerts, Julian for proving that there are cool Badensers, Stefan for the laughs, and Sabrina (Salamander!). Of course I thank our youngsters as well: Anja for sharing the burden of being the lone non-corpsemuncher, Tabea (Seagulls…), Jannis (Mahlzeit), Dimi for sharing his mango, and all the rest of the institute.

Considering the impact her fight for accessible science had on the availability of information in my field of research, I want to thank Alexandra Elbakyan for not giving up!

Furthermore, of course I wish to thank my family for always being there and being the way they are, as well as for their continuous support.

I want to thank my friends, especially Christoph, Philipp and Lisa, for the joyful times.

Finally, I thank my love Sandra for believing in me and for bringing out the best in me.

List of Publications

Bröhm, A., Elsawy, H., Rathert, P., Kudithipudi, S., Schoch, T., Schuhmacher, M.K., Weirich, S., and Jeltsch, A. (2019). Somatic Cancer Mutations in the SUV420H1 Protein Lysine Methyltransferase Modulate Its Catalytic Activity. Journal of molecular biology *431*, 3068-3080. https://doi.org/10.1016/j.jmb.2019.06.021.

Bröhm, A., Schoch, T., Dukatz, M., Graf, N., Dorscht, F., Mantai, E., Adam, S., Bashtrykov, P., and Jeltsch, A. (2022a). Methylation of recombinant mononucleosomes by DNMT3A demonstrates efficient linker DNA methylation and a role of H3K36me3. Communications Biology *5*, 192. https://doi.org/10.1038/s42003-022-03119-z.

Bröhm, A., Schoch, T., Grünberger, D., Khella, M.S., Schuhmacher, M.K., Weirich, S., and Jeltsch, A. (2022b). The H3.3 G34W oncohistone mutation increases K36 methylation by the protein lysine methyltransferase NSD1. Biochimie *198*, 86-91. https://doi.org/10.1016/j.biochi.2022.03.007.

Khella, M.S., **Bröhm, A.**, Weirich, S., and Jeltsch, A. (2020). Mechanistic Insights into the Allosteric Regulation of the Clr4 Protein Lysine Methyltransferase by Autoinhibition and Automethylation. International journal of molecular sciences *21*(22), 8832. https://doi.org/10.3390/ijms21228832.

Schuhmacher, M.K., Rolando, M., **Bröhm, A.**, Weirich, S., Kudithipudi, S., Buchrieser, C., and Jeltsch, A. (2018). The Legionella pneumophila Methyltransferase RomA Methylates Also Non-histone Proteins during Infection. Journal of molecular biology *430*, 1912-1925. https://doi.org/10.1016/j.jmb.2018.04.032.

Rajavelu, A., Lungu, C., Emperle, M., Dukatz, M., **Bröhm, A.**, Broche, J., Hanelt, I., Parsa, E., Schiffers, S., and Karnik, R., et al. (2018). Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. Nucleic acids research *46*, 9044-9056. https://doi.org/10.1093/nar/gky715.

Schuhmacher, M.K., Beldar, S., Khella, M.S., **Bröhm, A.**, Ludwig, J., Tempel, W., Weirich, S., Min, J., and Jeltsch, A. (2020). Sequence specificity analysis of the SETD2 protein lysine methyltransferase and discovery of a SETD2 super-substrate. Communications Biology *3*, 511. https://doi.org/10.1038/s42003-020-01223-6.

Abstract

Epigenetics is the study of heritable changes in phenotype which do not entail a change in the DNA sequence. Different epigenetic mechanisms exist that mediate these changes, the most prominent being DNA methylation and histone modifications. Together, these modifications form stable states of gene activation or repression while also steering the accessibility of whole genomic regions. The complex structure that packages the DNA in cells is called chromatin and beside the DNA comprises a plethora of associated proteins, most importantly the histone proteins. These small, basic proteins form an octameric complex around which ~147 bp of DNA are wound, generating the so-called nucleosome. Nucleosomes serve two functions at once: In addition to DNA compaction, histone proteins play essential roles for gene regulation and chromatin organization through the binding of secondary factors to their N-terminal domains, called histone tails. The histone tails are subject to many different modifications, such as acetylation, methylation, and phosphorylation, which modulate the binding of other chromatin-associated factors and thereby regulate transcription and chromatin states. Likewise, the methylation of DNA alters its properties and modulates gene expression. In order to study the biochemical properties of enzymes involved in the aforementioned processes, usually simple substrates such as peptides or DNA fragments are used. While these experiments are suitable for some scientific questions, they lack the physiological relevance of cellular studies. A good middle ground therefore lies in the use of recombinant nucleosomes, which represent a comparable structure as chromatin in living cells, but still are available in chemically homogeneous form as well as customizable in term of DNA sequence and histone modifications. This study consists of three projects which aimed to employ recombinant mononucleosomes in order to gain new insights into the function of selected DNA- and protein methyltransferases.

In the first project, recombinant nucleosomes were employed to study DNA methylation introduced by the DNA methyltransferase DNMT3A. For this, the nucleosomes were endowed with an extended CpG-rich linker DNA and methylation of these CpGs was detected by next generation bisulfite sequencing. The data revealed that DNMT3A has a distinct preference for the CpGs in the linker region while no methylation was detected in the nucleosome binding region. In comparison with the methylation pattern of free DNA, it was apparent that one distinct CpG site was preferentially methylated in the nucleosome context. This specific site was found to be located in close proximity to the active site of DNMT3A, based on a recently published cryo-EM structure of a DNMT3A/3B3 tetramer bound to a mononucleosome. Furthermore, the effect of histone tail modifications on DNMT3A activity was studied. For nucleosomes containing H3K4me3, a reduction of methylation close to the nucleosome was observed, while

the introduction of H3K36me3 led to an increase in methylation over the entire linker DNA region. Both of these effects are strongly correlated with DNA methylation patterns observed in human cells. For H3K36me3, a role in modulating linker DNA binding by the H3 tails was furthermore unveiled. In this study the nucleosomal linker DNA methylation by DNMT3A was studied in unprecedented detail, which resulted in new insights about the interaction of DNMT3A with nucleosomes as well as an expanded understanding of the role of histone tail modifications in the regulation of DNMT3A activity.

The second project focused on somatic cancer mutants of the SUV420H1 methyltransferase, which uses histones containing H4K20me1 as substrate to generate H4K20me2 and me3. Here it was discovered that most of the eight selected mutants exhibit greatly reduced activity or no activity at all on peptide substrates. SUV420H1 was demonstrated to be much more active on recombinant nucleosomes than on H4 protein, despite the lack of H4K20me1 on the substrate. To study the enzyme activity on the preferred monomethylated substrate in a nucleosomal context, native mononucleosomes isolated from SUV420H1/H2 double knockout cells were employed. On these native nucleosomes, some previously inactive mutants displayed detectable activity. In a cell culture assay using SUV420H1/H2 double knockout cells, reintroduction of the mutant enzymes followed by histone methylation analysis yielded a similar activity profile as observed with recombinant nucleosomes. Based on this comprehensive analysis, the role of the investigated mutations in cancer was considered.

In the third project recombinant nucleosomes were employed to investigate effects of the H3.3 G34W oncohistone mutation. Oncohistones are cancer-associated somatic mutations of the histone proteins, which have come into the focus of research in recent years. The G34W mutation is especially prevalent in patients with giant cell tumor of the bone and was shown to influence the catalytic activity of some methyltransferases which act on H3K36. The data revealed that the H3K36 dimethyltransferase NSD1 has a strong and unexpected preference for substrates containing G34W, which was replicated on peptide and on nucleosome substrates. This preference was considerably stronger on nucleosomes compared to peptides, indicating that nucleosome contacts are essential for the function of this enzyme. Conclusions were drawn from this finding regarding the involvement of the G34W mutation in secondary changes of the epigenome and in carcinogenesis.

Taken together, the application of recombinant nucleosomes as substrate for DNMTs und histone methyltransferases led to novel and important scientific discoveries with medical implications. This underscores the relevance of this experimental approach for future investigations in these research areas.

Zusammenfassung

Epigenetik umfasst die Untersuchung von vererbbaren Änderungen des Phänotyps, denen keine Änderung der DNA-Sequenz zugrunde liegt. Diese Veränderungen werden von unterschiedlichen epigenetischen Mechanismen gesteuert; als bedeutendste gelten hierbei DNA-Methylierung und die Modifikation von Histonproteinen. Durch die Kombination dieser Modifizierungen werden stabile Zustände von aktiver Genexpression oder Genrepression erreicht, sowie die Zugänglichkeit ganzer Genomregionen gesteuert. Die komplexe Organisationsstruktur des Genoms wird als Chromatin bezeichnet und beinhaltet neben der DNA auch eine Vielfalt an assoziierten Proteinen, wobei die Histonproteine zu den bedeutendsten gehören. Diese kleinen, basischen Proteine lagern sich zu einem oktamerischen Komplex zusammen, um den ~147 bp DNA gewickelt sind, wodurch das sogenannte Nukleosom gebildet wird. Nukleosome erfüllen zwei Funktionen: Zusätzlich zu der stattfindenden DNA-Komprimierung spielen die Histonproteine eine essentielle Rolle in der Genregulation und Chromatinstruktur, was durch die Bindung von sekundären Faktoren an ihre N-terminalen Domänen zustande kommt. Diese N-terminalen Histonenden unterliegen verschiedenen Modifikationen wie Acetylierung, Methylierung oder Phosphorylierung, die das Bindeverhalten von anderen chromatinassoziierten Faktoren beeinflussen und dadurch sowohl die Gentranskription als auch die Chromatinstruktur regulieren. Ebenso verändert die Methylierung der DNA deren Eigenschaften und beeinflusst hierdurch die Genexpression. Um die biochemischen Eigenschaften der Enzyme zu untersuchen, die in den genannten Prozessen beteiligt sind, werden oftmals einfache Substrate wie Peptide oder DNA-Fragmente verwendet. Derartige Experimente eignen sich für bestimmte wissenschaftliche Fragestellungen, entbehren jedoch oftmals physiologischer Bedeutung. Daher stellt die Verwendung von rekombinanten Nukleosomen einen guten Mittelweg dar, da diese eine vergleichbare Struktur wie das Chromatin in lebenden Zellen bieten, und dennoch in chemisch homogener Form präparativ zugänglich sowie in Bezug auf die DNA-Sequenz und Histonmodifikationen flexibel anpassbar sind. Diese Arbeit besteht aus drei Projekten, deren Ziel es ist, unter Zuhilfenahme von rekombinanten Nukleosomen neue Erkenntnisse über die Funktion ausgewählter DNA- und Protein Methyltransferasen zu gewinnen.

Im ersten Projekt wurden rekombinante Nukleosome zur Untersuchung der DNA-Methylierung durch die DNA Methyltransferase DNMT3A eingesetzt. Hierzu wurden die Nukleosome mit einer verlängerten, CpG-reichen Linker-DNA-Sequenz ausgestattet, deren Methylierung durch Next Generation Bisulfitsequenzierung detektiert wurde. Hierbei wurde gezeigt, dass DNMT3A die CpGs in der Linkerregion präferiert, wohingegen keine Methylierung in der nukleosomalen DNA beobachtet wurde. Im Vergleich mit dem Methylierungsmuster von freier DNA

wurde zudem beobachtet, dass eine bestimmte CpG-Stelle im nukleosomalen Kontext stärker methyliert wird. Mit Hilfe einer kürzlich publizierten kryo-EM-Struktur eines DNMT3A/3B3 Tetramers, welches an ein Mononukleosom gebunden ist, konnte festgestellt werden, dass sich diese spezifische CpG-Stelle in unmittelbarer Nähe zum katalytischen Zentrum einer der beiden DNMT3A-Untereinheiten befindet. Desweiteren wurde der Effekt verschiedener Histonmodifikationen auf die Methylierungsaktivität von DNMT3A untersucht. Hierbei wurde für Nukleosome mit H3K4me3 eine verringerte CpG-Methylierung nahe am Nukleosome detektiert, während H3K36me3 zu erhöhter DNA-Methylierung über die gesamter Linkerregion führte. Beide Effekte korrelieren mit DNA-Methylierungsmustern in menschlichen Zellen. Zudem wurde erstmals gezeigt, dass die Histonmodifikation H3K36me3 die Interaktion des H3-Endes mit der Linker-DNA beeinflusst. Diese Studie stellt die bis dato umfassendste Analyse der nukleosomalen Linker-DNA-Methylierung durch DNMT3A dar, wodurch neue Erkenntnisse über die Interaktion von DNMT3A mit Nukleosomen gewonnen werden konnten. Zuletzt konnte das Verständnis über die Rolle von Histonmodifikationen in der Regulierung der DNMT3A Aktivität erweitert werden.

Im zweiten Projekt wurden Krebs-assoziierte somatische Mutationen der SUV420H1 Methyltransferase untersucht, welche Histonproteine, die H4K20me1 beinhalten, als Substrat nutzt und zu H4K20me2 und me3 konvertiert. Es zeigte sich, dass die Meisten der acht ausgewählten Mutanten mit Peptidsubstraten keine oder nur sehr geringe katalytische Aktivität aufweisen. Zudem wurde gezeigt, dass SUV420H1 rekombinante Mononukleosome gegenüber dem H4 Protein als Substrat deutlich präferiert, obwohl diese kein H4K20me1 besitzen. Um die Methylierungsaktivität des Enzyms im nukleosomalen Kontext mit dem präferierten monomethylierten Substrat zu testen, wurden native Mononukleosome aus SUV420H1/H2 Knockout-Zellen eingesetzt. Mit diesen nativen Mononukleosomen zeigten auch einige bisher als inaktiv eingestufte Enzymmutanten eine detektierbare Aktivität. Zuletzt wurde durch Zellkulturexperimente gezeigt, dass die Wiedereinführung der mutierten Enzyme in SUV420H1/H2 Knockout-Zellen, gefolgt von einer Analyse des Histonmethylierungsstatus, zu einem vergleichbaren Aktivitätsprofil führt wie es bereits mit den nativen Nukleosomen beobachtet wurde. Basierend auf dieser umfassenden Analyse wurden die Auswirkungen der einzelnen Mutationen in Krebserkrankungen beleuchtet.

Im dritten Projekt wurden rekombinante Nukleosome eingesetzt, um die Effekte der H3.3 G34W Onkohistonmutation zu beleuchten. Als Onkohistone werden somatische Krebs-assoziierte Histonmutationen bezeichnet, die in den letzten Jahren vermehrt erforscht wurden. Die G34W Mutation tritt sehr häufig in Patienten mit dem Riesenzelltumor des Knochens auf und hat die Eigenschaft, die

katalytische Aktivität bestimmter H3K36 Methyltransferasen zu beeinflussen. Es wurde gezeigt, dass die H3K36 Dimethyltransferase NSD1 unerwarteterweise solche Peptide- und Nukleosomsubstrate präferiert, die G34W enthalten. Diese Präferenz war auf Nukleosomen deutlich stärker ausgeprägt als auf Peptiden, was die essenzielle Rolle von Nukleosomkontakten für die enzymatische Aktivität dieses Enzyms illustriert. Aus den gewonnenen Daten konnten Rückschlüsse auf die Beteiligung der G34W-Mutation auf sekundäre Veränderungen im Epigenom und in der Krebsentstehung gezogen werden.

Zusammengefasst konnten durch die Anwendung rekombinanter Nukleosome als Substrat für DNA- und Histon Methyltransferasen neue und bedeutende Erkenntnisse gewonnen werden, was die Relevanz dieses experimentellen Ansatzes für zukünftige Untersuchungen in den betreffenden Forschungsfeldern unterstreicht.

List of figures

List of Abbreviations

ADD	ATRX-DNMT3A-DNMT3B
AdoHcy/SAH	S-Adenoyl-L-homocysteine
AdoMet/SAM	S-Adenosyl-L-methionine
bp	Base pairs
CD	Circular dichroism
CHIP	Chromatin immunoprecipitation
COSMIC	Catalogue of somatic mutations in cancer
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
FPLC	Fast protein liquid chromatography
FRET	Förster resonance energy transfer
GCTB	Giant cell tumor of the bone
GST	Glutathione-S-transferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HKMT	Histone lysine methyltransferase
MALDI-TOF	Matrix-assisted laser desorption/ionization; time-of-flight
NGS	Next generation sequencing
PWWP	Proline-tryptophan-tryptophan-prolin motif
PCR	Polymerase chain reaction
PKMT	Protein lysine methyltransferase
SET	**S**u(var)3-9, **E**(z) and **T**rithorax

TSS	Transcriptional start site

1 Introduction

1.1 Epigenetics

Since its emergence more than 4 billion years ago, life has evolved to tremendous complexity. After the discovery of DNA in the second half of the 19th century (Miescher-Rüsch, 1871), our understanding of the molecular processes that constitute life has grown rapidly, culminating in the complete sequencing of the genomes of entire species. The human genome project, which came to conclusion at the beginning of the new millennium (Venter et al., 2001), revealed that our genomic information comprises more than 3 billion base pairs of DNA encoding more than 20000 genes. This genetic information is present in almost identical form in every cell of the human body. However, the human body consists of about 200 different cell types (Moris et al., 2016), and this functional diversity cannot be explained by the DNA sequence alone. Figuratively speaking, the DNA can be seen as a blueprint for life, but it is missing the instructions for assembly. Specifically, for a long time it was unknown how exactly the DNA is translated into functionally diverse cellular phenotypes. As an explanation for this matter, the geneticist Conrad Hal Waddington proposed a mechanism he described as "Canalization of development and the inheritance of acquired characters" (Waddington, 1942). Later, he invented and established the term "Epigenetics" and described the differentiation of totipotent cells to specific cells types during development on an "epigenetic landscape" (Figure 1A). In this picture, a cell, which is visualized as a marble, travels down a surface composed of several branching valleys, which symbolize the different developmental paths.

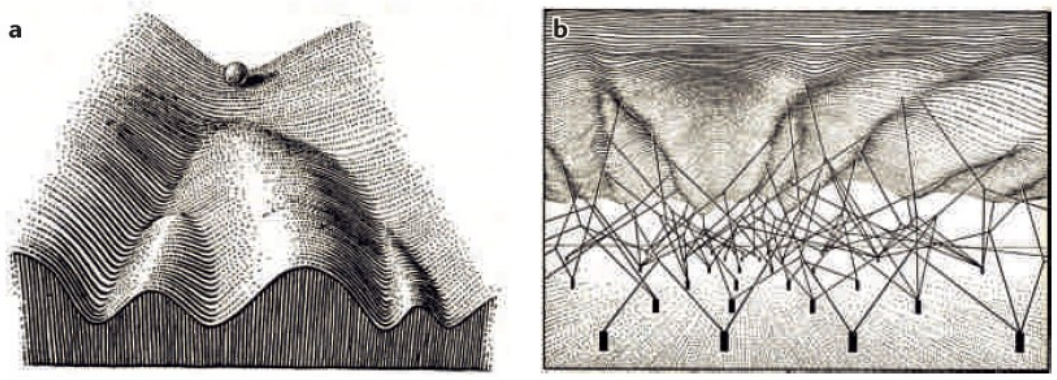


Figure 1: The epigenetic landscape as described by Conrad Waddington. (A) Epigenetic landscape with a marble about to roll down into the branching valleys. (B) The foundation underneath the epigenetic surface is formed by a network of supporting struts. The image was reproduced from Waddingtons work by Moris (Moris et al., 2016).

Importantly, the valley or differentiation state the cell travels into is irreversible but non-random. In order to influence the fate of the cell, a complex network of supporting struts is present under the surface (Figure 1B), which represent the biological factors influencing the epigenetic landscape.

Since these early landscape-models of epigenetics, numerous advances in the biological sciences have refined our understanding to a point where an updated definition of the term "Epigenetics" was needed (Wu and Morris, 2001; Berger et al., 2009). The currently accepted definition is as follows: "An epigenetic trait is a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (Berger et al., 2009). This definition encompasses the understanding that the differentiation states of cells, which are represented by valleys in Waddingtons epigenetic landscape, are described by gene expression states as an epigenomic regulation mechanism. These expression changes do not alter the DNA sequence but instead are the result of modifications to the DNA and DNA-binding proteins, which allow for or repress gene expression.

1.2 Chromatin structure

The human genome comprises 3 billion base pairs, which results in a total length of the DNA approaching 2 meters. In order to fit into the nucleus of a human cell, which has an average diameter of only 10 µm, the DNA has to be drastically compacted. This compaction is achieved by various DNA-binding proteins, which form small-scale complexes as well as large-scale interactions between these units, resulting in a compacted but nevertheless dynamic state. This conglomerate of DNA, DNA-binding proteins, and other regulatory factors is called chromatin. It is worth noting that the contribution of the DNA to the bulk mass of chromatin is surprisingly small: In yeast nuclei, it was found that DNA only comprises ~2.5 % of the isolated material, while up to 80 % was attributed to proteins (Rozijn and Tonino, 1964). This illustrates the complexity of the structural framework of chromatin which is needed to organize the DNA. On the smallest scale, chromatin consists of nucleosome particles, which are comprised of a DNA fragment wrapped around an octameric histone protein complex. Through spooling the DNA into a barrel-like structure, already a sevenfold compaction compared to a linear chain is achieved (Cutter and Hayes, 2015). By interaction between these nucleosome particles, fiber-like structures can be formed, resulting in even further compaction (Robinson and Rhodes, 2006). Ultimately, the condensation of chromatin progresses from this molecular scale up to the chromosomes. A schematic representation of the DNA compaction in chromatin is depicted in Figure 2.

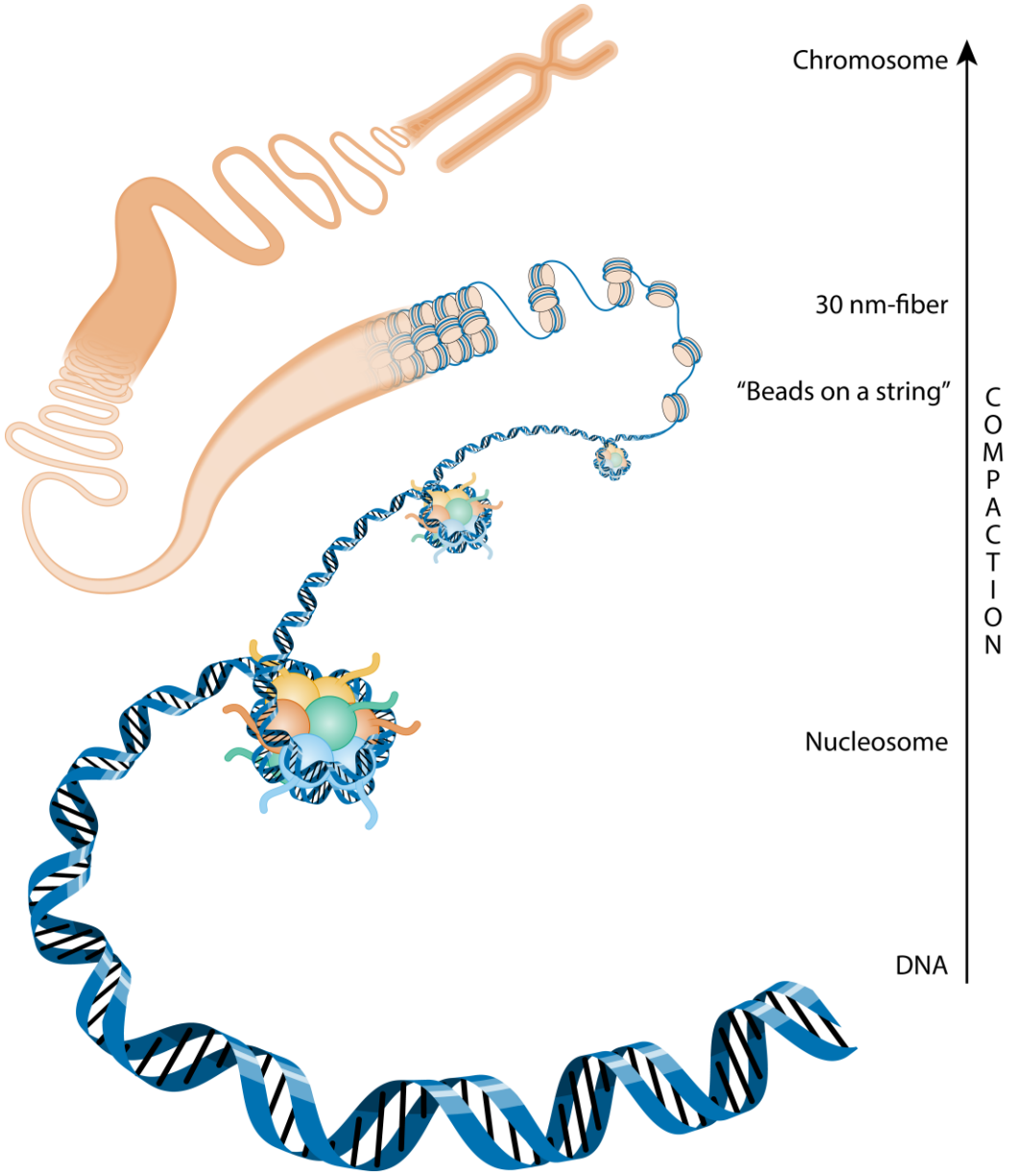


Figure 2: Overview of chromatin structure. The DNA is wrapped around histone octamers to form nucleosomes, which interact and stack to form higher-order structures, ultimately resulting in chromosomes.

The three-dimensional structure of the nucleosome core particle was solved to unprecedented resolution by Luger et al. at the end of the 20th century (Luger et al., 1997). This study allowed for a detailed understanding of the interaction between histone proteins and DNA on an atomic level. The nucleosome consists of an octameric complex contain two copies of each of the core histone proteins H2A, H2B, H3, and H4, with 147 bp of DNA wrapped around (Figure 3A). This

interaction is largely facilitated by the strong interaction between the negatively charged DNA backbone and the basic histone proteins. The resulting complex is relatively flat and cylinder shaped, with a diameter of about twice the height of the cylinder. The globular portion of the octamer complex displays a twofold rotary symmetry, while the whole nucleosome can be considered pseudo-symmetrical. This rotary symmetry is defined by the dyad axis, which goes from the middle of the DNA entry/exit site to the opposite side of the nucleosome.

The individual histones comprise a common fold motif consisting of three helices α1-α3, which allow for the formation of histone dimers between H2A and H2B, or H3 and H4, respectively (Figure 3B,C). Dimers formed by H3 and H4 rapidly self-associate to tetramers by a strong interaction between the C-terminal helices α2 and α3 of histone H3. The histone octamer structure is then completed by addition of two H2A/H2B dimers, which is mostly facilitated by an interaction of the C-terminal helices and loops of H4 and H2B (Figure 3C,D).

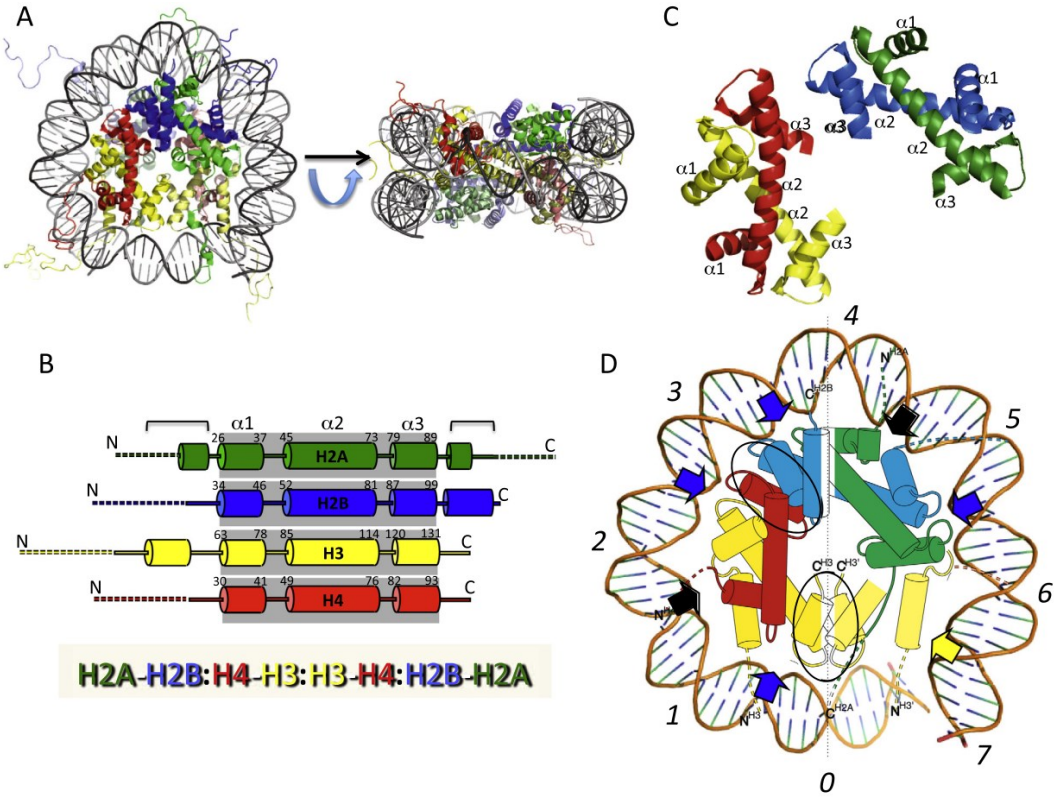


Figure 3: Overview of the nucleosome structure. (A) Top and side view (along the dyad axis) of the nucleosome. (B) Schematic view of the individual histone fold motifs. (C) Partial view of the crystal structure with focus on the interactions between H2A/H2B as well as H3/H4. (D) Schematic view of the nucleosome with histone helix fold motifs and indicated DNA helical turns. The image was taken from (Cutter and Hayes, 2015).

In addition to this globular core domain, the histone proteins each contain an unstructured N-terminal domain, which faces outward of the core nucleosome in the direction of the DNA. Although there can be no single definitive conformation for these domains due to their unstructured nature, it is known that the histone tails contribute to DNA binding and internucleosomal interactions, as well as nucleosome stability (Davey et al., 2002; Pepenella et al., 2014; Li and Kono, 2016). Furthermore, the dynamic nature of the histone tail binding to DNA makes them important factors determining DNA accessibility (Li and Kono, 2016; Lehmann et al., 2020; Ghoneim et al., 2021).

Nucleosomes act as an inherent barrier for cellular processes involving DNA, ranging from protection against nuclease digestion and methylation to large-scale processes like replication and transcription (Felle et al., 2011; Grigoryev and Woodcock, 2012; Kelly et al., 2012). Although the nucleosome is an overall very stable structure under physiological conditions, it is also highly dynamic: On one hand, the hierarchical architecture of the histone octamer allows for exchange of the relatively weakly bound H2A/H2B dimers while the more rigid H3/H4 tetramer keeps the complex together (Das and Tyler, 2013). On the other hand the DNA is also known to transiently unwind and reattach to the octamer, a process which is called “DNA breathing” (Ngo et al., 2015; Winogradoff and Aksimentiev, 2019). These mechanisms contribute to the dynamic nature of chromatin by allowing access to DNA-binding factors (Luger et al., 2012). In addition, external ATP-dependent remodeling factors are able to either induce shifting of the DNA along the nucleosome or completely disassemble the nucleosome by successive removal of histone subunits (Hota and Bruneau, 2016).

Nucleosomes are bound to DNA in semi-regular intervals, which are defined by the DNA sequence. While some DNA regions exhibit only loosely preferred nucleosome binding sites, others display strongly directed positioning (Jansen An and Verstrepen Kevin J., 2011). The nucleosomes are separated by a linker DNA segment of variable length, but usually not exceeding 70 bp (Olins and Olins, 1974). Like the nucleosome-binding sequences, the linker DNA has an intrinsic periodicity, resulting in preferred regular spacing of nucleosome particles (Widom, 1992). The resulting structure, termed "beads on a string" can be further stabilized and condensed by the linker histone H1, which binds to the DNA at the nucleosome entry/exit site at the dyad axis and promotes stacking interactions between nucleosomes (Robinson and Rhodes, 2006). By regular stacking, the nucleosome chain forms a more condensed fiber which is often called the 30 nm fiber due to its approximate diameter of 30 nm (Li and Reinberg, 2011; Zhu and Li, 2016). The 30 nm fiber mostly consists of a two-start zig-zag structure, which can exhibit different degrees of compaction depending on the local linker DNA length and presence of linker histones (Dorigo Benedetta et al., 2004; Routh Andrew et al., 2008). However, precisely ordered fibers were mostly detected using unmodified reconstituted chromatin fragments while cellular chromatin has proven to be much more heterogeneous (Grigoryev and Woodcock, 2012). Nevertheless, the small-scale condensation of nucleosome-containing DNA is sufficiently described by the ensemble of models for 30 nm fibers.

The large-scale compaction of the genetic material into chromatin has puzzled researchers for a long time. Inspired by the regularity of the 30 nm fiber models, a hierarchical folding mechanism has been proposed, which would suggest higher-order chromatin structures that follow a repeated coiling mechanism (Ozer et al., 2015). However, more recent advances in chromatin research have led to the conclusion that chromatin can in fact be largely disordered but still compacted into relevant sub-domains by a liquid-liquid phase separation mechanism involving histones, transcription factors, and chromatin remodeling factors (Erdel and Rippe, 2018; Gibson et al., 2019; Wang et al., 2019; Shakya et al., 2020; Wagh et al., 2021; Rippe, 2022). This newly found perspective also overcomes the most drastic limitation of the hierarchical folding model: In a highly ordered state, accessing a single locus involves dissolving not only the local structure but also the higher-order folds, which makes efficient gene regulation impossible.

1.3 Histone modifications

Histone proteins are of utmost importance for chromatin organization due to the stable nucleosome complex they form together with DNA. However, DNA cannot be compressed and irreversibly archived but instead has to be selectively accessible for cellular processes. One important mechanism for the regulation of chromatin accessibility is the post-translational modification of histone proteins. The earliest discovery of modified amino acids in histones dates back to the 1960s, when Allfrey et al. demonstrated that histones can be acetylated and methylated (Allfrey et al., 1964). Since then, a plethora of post-translational modifications (PTMs) on a similarly large number of histone residues was discovered (Arnaudo and Garcia, 2013; Zhao and Garcia, 2015). Due to the distinct chemical properties of the amino acid side chains, the majority of described modifications are limited to a select group of chemically related amino acids (Bischoff and Schlüter, 2012). The most prominent include lysine and arginine methylation, lysine acetylation, serine and threonine phosphorylation, and lysine ubiquitination. In total, more than 200 different types of modifications were described. Notably however, many of these modifications are believed to serve very specific roles in gene regulation and chromatin organization and are therefore not particularly widespread (Leroy et al., 2013).

Histone modifications can influence chromatin architecture mainly on two different levels. By regulation of intranucleosomal contacts between histones or between histones and DNA, the nucleosome can be essentially destabilized. Most notably, acetylation in the globular fold of histone H3 is able to loosen DNA binding by reversing the positive charge of lysine residues, thereby creating a local repressive force (Neumann et al., 2009; Shimko et al., 2011). This effect is not limited to the globular histone domains, as widespread acetylation was also described on the histone tails, especially in a clustered region on the H4 tail (Zhao and Garcia, 2015). On a more global level, modification of histone residues involved in internucleosomal interactions can lead to altered or impaired chromatin fiber formation, which leads to widespread chromatin decompaction. Such an effect was described for H4K16 acetylation, but other histone PTMs, such as H4K20me3, have similar roles as well (Shogren-Knaak Michael et al., 2006).

As described earlier, the N-terminal histone tail domains play an important role in nucleosome stability and accessibility. It is therefore no surprise that the histone tails are key targets for post-translational modifications, which alter the chemical and sterical properties of the tails and thereby modulate the chromatin structure and condensation, enhance or repress the binding of secondary factors as well as dictate chromatin accessibility. An overview of selected known modification on histone tail domains is depicted in Figure 4. Although many described modifications are omitted here for the sake of readability, it immediately becomes clear that the H3 and H4 N-terminal tails are the main targets for amino acid modifications. This seems natural, because binding of the H3 and H4 tails to DNA is a major determining factor for nucleosome stability and therefore ideal leverage points for influencing nucleosome and chromatin dynamics (Zentner and Henikoff, 2013).

Figure 4: Overview of the most common amino acid modifications on the tail domains of histone proteins. Me: methylation, Ac: acetylation, Ph: phosphorylation, Ub: ubiquitylation.

Considering the plethora of known histone modifications, is stands to reason that these modifications seldom exist isolated from each other. Indeed, decades of research has provided insight into the functional complexity of histone modifications, leading to the understanding that many modifications can influence by each other in a positive or negative manner. This in turn opens up a great number of mechanisms for transcriptional regulation (Bannister and Kouzarides, 2011). Specifically, it was found that the trimethylation of H3K4 and H3K36 as well as histone acetylation are associated with actively transcribed genes, while trimethylation of H3K9, H3K27, and H4K20 is attributed to transcriptional repression. As shown in Figure 5, many of the modifications which directly influence gene expression are located in the promoter region or around the transcriptional start site (TSS) of a gene. Nevertheless, some regulatory mechanisms also rely on a more widespread localization of histone marks including the gene bodies. As mentioned before, histone modifications influence each other in the form of positive or negative feedback, which is orchestrated by reading domains which are part of the enzyme complexes responsible for deposition of the marks (Suganuma and Workman, 2008; Zhang et al., 2015). As a result, stable states of active

transcription or inactivated genes can not only be established, but also maintained and inherited, thereby confirming the epigenetic nature of these modifications.

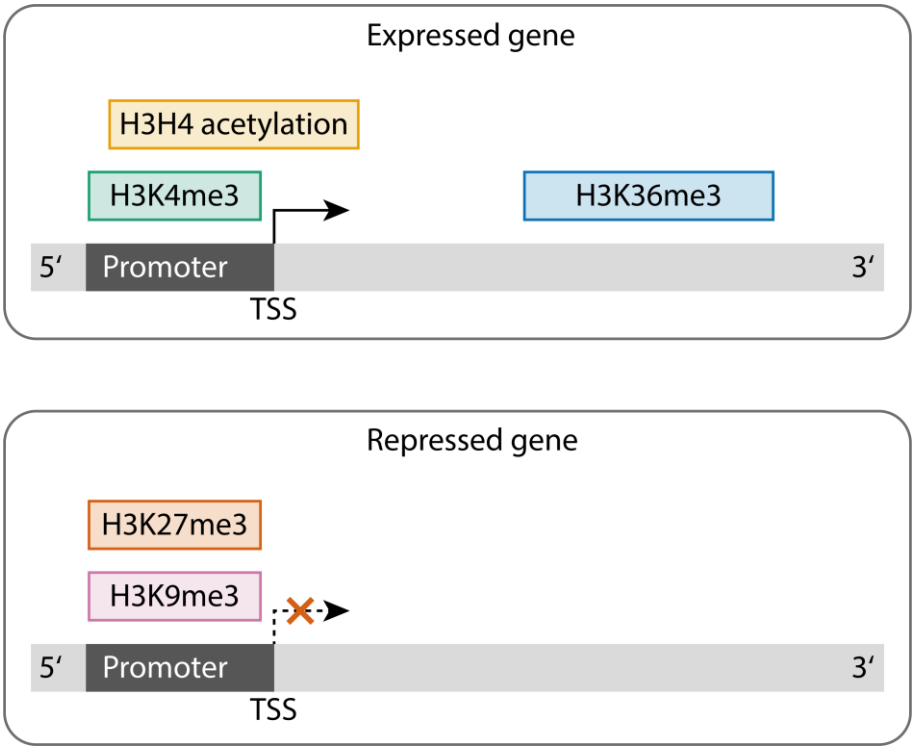


Figure 5: The prevalence of prominent histone modifications on a representative gene. In the actively expressed state, H3K4me3 in combination with histone acetylation is focused on the promoter region while H3K36me3 is spread throughout the gene body. In the repressed state, the inactivating histone marks H3K9me3 and H3K27me3 are present at the promoter region.

Similar to other epigenetic mechanisms, histone modifications play an important role in disease, especially cancer. Studies have shown that histone modifications in cancer can be altered substantially and globally (Seligson et al., 2009). Aberrant regulation of certain modifications can influence disease onset and progression either by direct effects, such as through inadvertent activation of oncogenes, or through unintended recruitment of secondary factors. Through the aforementioned crosstalk mechanisms between different histone marks, mistakes can be maintained and even propagated, thereby resulting in detrimental physiological effects (Audia and Campbell, 2016).

1.3.1 Histone methylation

One of the most common post-translational modifications of histone proteins is methylation of lysine and arginine side chains. For this work, mainly lysine methylation is of importance, and will be discussed in further detail in the following chapter. Lysine methylation was first described in the 1960s by Allfrey et al. and Murray (Allfrey et al., 1964; Murray, 1964). Since then, a large number of histone lysine residues were shown to be modified by methylation (see Figure 4). In concurrence with this finding, a group of histone lysine methyltransferases (HKMTs) was discovered. These enzymes are called SET enzymes according to their conserved catalytic SET domain named after the group of proteins in which it was first described: **S**u(var)3-9, **E**(z) and **T**rithorax. SET enzymes were shown to modulate chromatin dynamics (Jenuwein et al., 1998), however, which of these enzymes catalyze the methylation of specific residues remained elusive. Further research on SET enzymes then revealed the Suv39H1/H2 enzymes as specific H3K9 methyltransferases (Dónal O'Carroll et al.; Rea et al., 2000). Since then, many more SET-domain enzymes with specific functions were discovered (Alvarez-Venegas and Avramova, 2002; Xiao et al., 2003; Dillon et al., 2005; Qian and Zhou, 2006). In parallel, it was also found that the ability to methylate lysine residues in histone proteins is not limited to the SET enzyme family, as demonstrated by the DOT1L protein (Min et al., 2003). In the decade after the discovery of the first histone lysine methyltransferase, a tremendous amount of other HKMTs was described and characterized (Kouzarides, 2007; Greer and Shi, 2012).

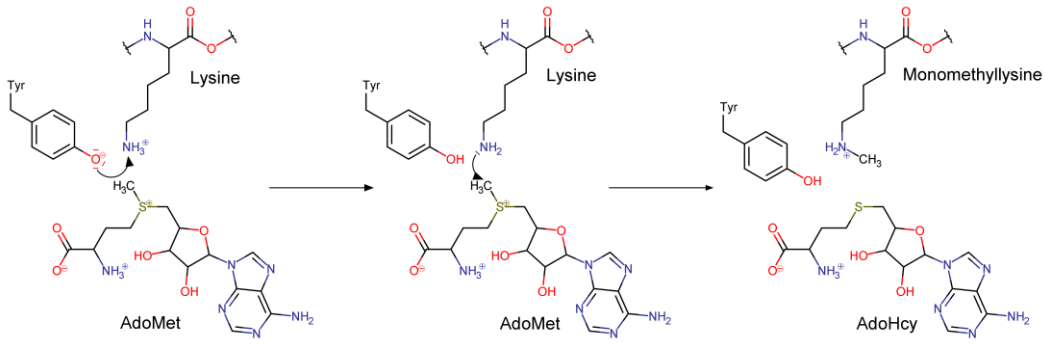


Figure 6: Schematic reaction mechanism of the methylation of a lysine residue by a SET-domain enzyme.

For the SET domain enzymes, the catalytic mechanism of lysine methylation relies on deprotonation of the lysine amine by a conserved tyrosine residue, thereby increasing its nucleophilic properties sufficiently to abstract the sulfur-bound methyl group from the cofactor S-adenosyl-L-methionine (SAM or AdoMet) (Trievel et al., 2002). The mechanism is illustrated in Figure 6. Notably, the binding of the cofactor and the substrate are independent of each other due to the

relatively large distance between the binding pockets of the SET domain. This circumstance allows for a great variety of substrate binding site geometries while preserving a similar catalytic function, thereby enabling the recognition and methylation of specific substrates (Del Rizzo and Trievel, 2014). As a result, the SET domain family comprises a large number of functionally distinct enzymes, which usually focus on a single histone mark (Qian and Zhou, 2006). However, since the scope of potential substrates of these enzymes is largely dictated by their sequence specificity, it is no surprise that methylation of other lysine residues on histones or non-histone proteins is frequently observed (Rathert et al., 2008; Zhang et al., 2012; Kudithipudi et al., 2014a; Schuhmacher et al., 2015; Weirich et al., 2020). The methylation of non-histone proteins has broad implications: For instance, the H4K20 methyltransferase SET8 was shown to methylate a lysine residue on the tumor suppressor p53, which modulates its activity (Shi et al., 2007). This confirms the role of non-histone protein methylation as an active regulatory mechanism involved in cellular function. In fact, given the previously reported methylation of non-histone proteins by putative HKMTs, today most of these enzymes are regarded more generally as protein lysine methyltransferases (PKMTs).

Central to the understanding of lysine methylation is the fact that the terminal amine moiety of lysine can be subject to several methylation steps ranging from mono- to trimethylation, as illustrated in Figure 7. The outstanding feature of these modifications is that they exhibit a tremendous functional diversity depending on the location and state of methylation (Martin and Zhang, 2005; Gelato and Fischle, 2008).

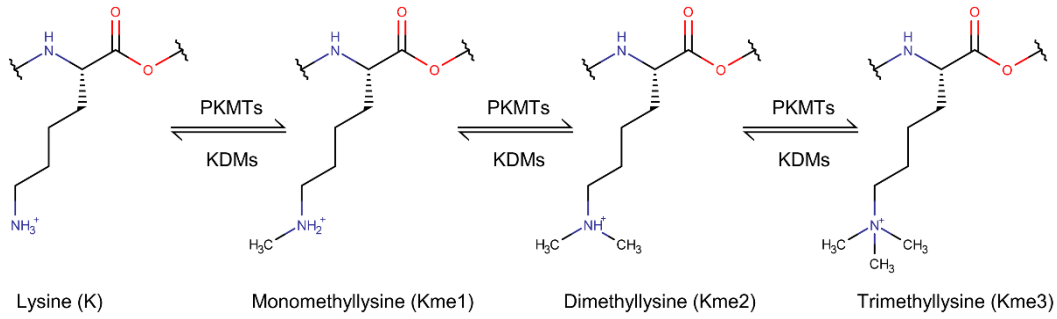


Figure 7: Scheme of the different methylation states of lysine, ranging from the unmethylated to the trimethylated side chain.

As mentioned earlier, histone methylation can enable or inhibit the binding of transcription factors or other epigenetic effectors by interaction with reading domains of these proteins, as well as modulate chromatin dynamics by altering nucleosome stability. Due to the great importance of histone methylation in gene regulation, its dysregulation can be associated with cancer and other diseases

(Seligson et al., 2009). One major cause for the aberrant regulation of histone methylation is an impaired function of the responsible PKMTs, which can be caused by mutations or translocations (Bennett et al., 2017). Mutations in PKMTs have been found in a great variety of human cancers and are mostly associated with loss of function, thereby leading to reduced levels of the respective methylated histone mark (Kudithipudi and Jeltsch, 2014). More rarely, these mutations can also lead to a gain of function or even to a change in functional properties of the enzyme (Majer et al., 2012; Weirich et al., 2015). Consequently, it is of great interest to gain a deeper understanding of the PKMTs are involved in cancer and to find novel therapeutic approaches.

In the following sections, the function and responsible PKMTs of several methylated histone marks, which are of interest for the present study, will be described in further detail.

1.3.2 H3K36 methylation

Methylation of the H3K36 site is an important regulatory mark involved in a number of cellular processes (Li et al., 2019). It is generally associated with actively transcribed genes (Lucio-Eterovic et al., 2010; Wagner and Carpenter, 2012) and essential for development (Rayasam et al., 2003). Furthermore, it plays a role in DNA repair and thereby in genome integrity (Li et al., 2013). By crosstalk with other epigenetic effectors, H3K36 trimethylation acts in an antagonistic manner to the EZH2-mediated deposition of the repressive H3K27me3 mark, which aids in the segregation of active and repressed genetic regions. However, the lower methylation states of H3K36 and H3K27 were also observed to coexist in the same regions, which indicates a more complex switch mechanism between active, inactive and poised chromatin states (Streubel et al., 2018). Another important mechanism facilitated by H3K36 methylation is the crosstalk with DNA methylation (Choufani et al., 2015). This crosstalk is mediated by the PWWP-domain of the *de novo* DNA methyltransferases, which recognize the di- and trimethylated states of H3K36 and methylate adjacent DNA regions (Dhayalan et al., 2010; Dukatz et al., 2019; Weinberg et al., 2019). By this mechanism, H3K36me3 present in gene bodies effectively recruits DNA methylation to these regions, which prevents incorrect transcription initiation (Neri et al., 2017). Dysregulation of H3K36me3 is a major risk factor in disease, most notably in Sotos syndrome (Kurotaki et al., 2002), but also in leukemia (Rosati et al., 2002) and other cancers (Berdasco et al., 2009).

The enzymes responsible for the deposition of H3K36 methylation can be broadly classified into two groups of mono- and dimethyltransferases on one hand and trimethyltransferases on the other hand (Li et al., 2019). The most important enzymes involved in mono- and dimethylation of H3K36 in mammals are the

members of the NSD family, NSD1, NSD2, and NSD3. The final trimethylation step is exclusively carried out by the SETD2 enzyme (Pfister et al., 2014; Li et al., 2016). The involvement of these enzymes in disease was extensively described earlier, demonstrating that loss of function leads to loss of H3K36me2/3 and further perturbations of H3K27me3 as well as DNA methylation. The NSD3 protein is specifically affected by a translocation event, which creates a chimeric fusion with NUP98 and is associated with acute myeloid leukemia (Rosati et al., 2002). Loss of function of NSD1 is furthermore considered an essential factor in the Sotos overgrowth syndrome (Kurotaki et al., 2002; Rio et al., 2003; Tatton-Brown and Rahman, 2004; Tatton-Brown et al., 2005; Berdasco et al., 2009).

For the NSD2 and NSD3 enzymes, recently cryo-EM structures were published, which reveal the interaction of the enzymes with a nucleosome particle in unprecedented detail (Li et al., 2021; Sato et al., 2021). In these structures, it can be seen that the NSD protein unwraps around 20 bp of nucleosomal DNA and tightly binds to the resulting space between histone octamer and DNA. A binding mode which so severely disrupts the nucleosome structure was unanticipated and suggests an even more complicated mechanism of NSD recognition and catalytic activity, since it multiplies the number of potential influential factors due to the many intricate interactions with its substrate.

1.3.3 H4K20 methylation

Methylation of the H4K20 mark is the most well documented histone mark on histone 4 and serves a variety of purposes depending on the methylation state (Wang and Jia, 2009). Generally, H4K20 methylation is associated with constitutive heterochromatin, especially in the telomere and centromere regions (Balakrishnan et al., 2010; Jørgensen et al., 2013). Monomethylation is essential for chromatin structure and compaction (Oda et al., 2009), while dimethylation plays an important role in DNA damage response (Tuzon et al., 2014). Trimethylation of H4K20 is a repressive mark and appears mostly in heterochromatic regions (Kourmouli et al., 2004). It was shown to modulate the higher-order chromatin structure, which results in enhanced condensation (Lu et al., 2008).

In mammals, there are three important enzymes which catalyze the methylation of H4K20: The monomethyltransferase SET8 (Fang et al., 2002; Couture et al., 2005) and the two SUV420 paralogs H1 and H2 (Schotta et al., 2004). The SUV420 enzymes prefer monomethylated H4K20 as substrate and are thus largely dependent on SET8 activity to generate H4K20me1 (Wu et al., 2013a; Southall et al., 2014). The importance of SET8 activity was confirmed by studies of SET8 knockout mouse embryos, which displayed loss of H4K20me1, coinciding with

reduced higher methylation states, as well as embryonic lethality (Jørgensen et al., 2007; Oda et al., 2009). The same study also demonstrated the catastrophic effect of H4K20me1 loss on cell cycle progression. In healthy cells, high levels of H4K20me2 and me3 are present in heterochromatic regions during G1 phase, while H4K20me1 is limited to specific regions. With the incorporation of new, unmodified histones during chromatin recondensation in S phase, most H4K20 methylation is lost and later reintroduced by SET8 as monomethylation during G2 phase. During mitosis, this widespread monomethylated state is maintained and converted to H4K20me2 and me3 by the SUV420 enzymes in early G1 phase (Jørgensen et al., 2013).

The SUV420 enzymes are recruited to chromatin by heterochromatin protein 1 (HP1) binding to the repressive H3K9me3 mark, thereby facilitating stable chromatin silencing (Schotta et al., 2004). The two SUV420 paralogs have overlapping, but also distinctive functions. While SUV420H2 expression is limited to certain tissues, SUV420H1 is ubiquitously expressed during development and in adult tissues (Schotta et al., 2008). The H4K20 methylation introduced by SUV420H1 also facilitates DNA damage signaling during double-strand breaks by recruiting p53-binding protein 1 (53BP1) (Botuyan et al., 2006b; Tuzon et al., 2014), a mechanism that was shown to be impeded by employing a SUV420H1 inhibitor (Bromberg et al., 2017). Furthermore, SUV420H1 seems to be especially prone to mutation, with several mutations being described in cancers (Feinberg et al., 2016). Loss of the SUV420-mediated H4K20me3 is also associated with increased invasion and poor prognosis in breast cancer (Yokoyama et al., 2014), which demonstrates the importance of this enzyme.

1.4 DNA methylation

DNA methylation is the second important chromatin modification aside the modification of histone proteins. As the name implies, here the DNA bases themselves are subject to modification, as opposed to DNA-binding proteins. The earliest description of methylated DNA dates back to the middle of the 20th century, when Rollin Hotchkiss separated DNA bases extracted from cells by paper chromatography (Hotchkiss, 1948). In these experiments, he observed an additional fraction of unknown composition running close to the spot of cytosine, which he dubbed "epi-cytosine", and which was later identified as 5-methylcytosine. Through comprehensive whole-genome studies we know today that DNA methylation in human cells mainly occurs as 5-methylcytosine, with approximately 4-6 % of all cytosine residues being methylated (Lister et al., 2009). However, there exists a strong sequence preference for a CpG context, although CpG is by far the least common of all CpX combinations. Non- CpG methylation also occurs to a certain degree in human cells, but strongly dependent on the cell type (Schultz et al., 2015). CpG sites also are unevenly distributed throughout the genome: While most parts of the genome including the gene bodies only display sparse CpG density, the promoter regions and transcriptional start sites show an overrepresentation of CpG sites (Saxonov et al., 2006).

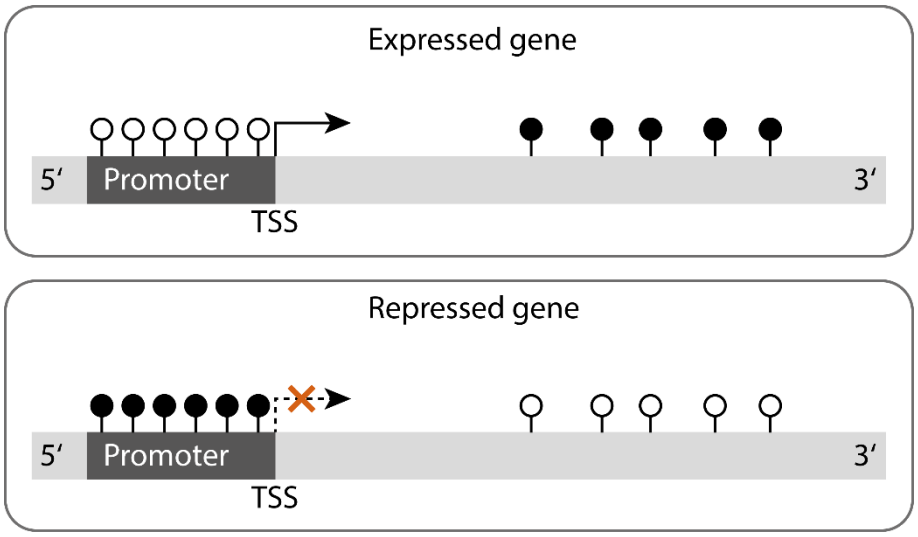


Figure 8: Schematic model of the distribution of methylated CpG sites. In actively expressed genes, the CpG islands at the transcriptional start sites are unmethylated and some methylation is present in the gene body. In stably repressed genes, the CpG island is densely methylated, while the gene body is largely unmethylated.

It is at these so-called CpG islands that the most important regulatory function of DNA methylation takes effect: Through modulated binding to other chromatin-associated factors (Patel, 2016), DNA methylation induces a strong and stable repression of the affected regions (Razin and Cedar, 1991; Deaton and Bird, 2011; Jones, 2012). However, although DNA methylation at the TSS site is solely associated with repression of gene transcription, the function of DNA methylation occurring in the gene body is not entirely clarified. The currently established model suggests that although gene body methylation has little effect on the expression of the target gene itself, it serves to repress the spurious transcription initiation at alternative start sites and thereby aids flawless transcription elongation (Neri et al., 2017; Teissandier and Bourc'his, 2017). A visual representation of the distribution of CpG methylation at actively transcribed or repressed genes is given in Figure 8.

The silencing mechanism relying on DNA methylation of promoters and transcriptional start sites is of a very stable and long-term manner and therefore often associated with large-scale processes that rely on these qualities. These processes include the X-chromosome inactivation, where a permanent repression is imposed on one copy of the female sex chromosomes (Cotton et al., 2015). Another common occurrence of this long-term silencing is on the heterochromatic regions at telomers and centromers, as well as on repeats and transposable elements, which make up a sizeable fraction of the genome (Jones, 2012). Furthermore, DNA methylation plays a critical role in development (Chen et al., 2003; Smith and Meissner, 2013) including genomic imprinting and X-chromosome inactivation (Li et al., 1993; Razin and Cedar, 1994).

1.4.1 DNA methyltransferases and their function

The enzymes responsible for the deposition of DNA methylation are called DNA methyltransferases. In human cells, the family of DNA methyltransferases comprises two structurally and functionally distinct groups. First, the so-called maintenance methyltransferase DNMT1 and second, the *de novo* methyltransferases DNMT3A and DNMT3B (Cheng, 1995; Okano et al., 1999). As outlined in Figure 9, the human DNMTs can be characterized according to their domain compositions. The N-terminal part of the enzymes contains regulatory and targeting domains responsible for interactions with DNA, histone tails, and other factors, as well as the regulation of the DNMT activity. The C-terminal part comprises the catalytically active site, which is formed by conserved amino acid motifs. Another part of the *de novo* DNA methylation machinery is also depicted in Figure 9: A truncated version of the DNMT3 paralogs, named DNMT3L, which lacks part of the regulatory domain as well as part of the catalytic motifs necessary for methylation activity, rendering it catalytically inactive.

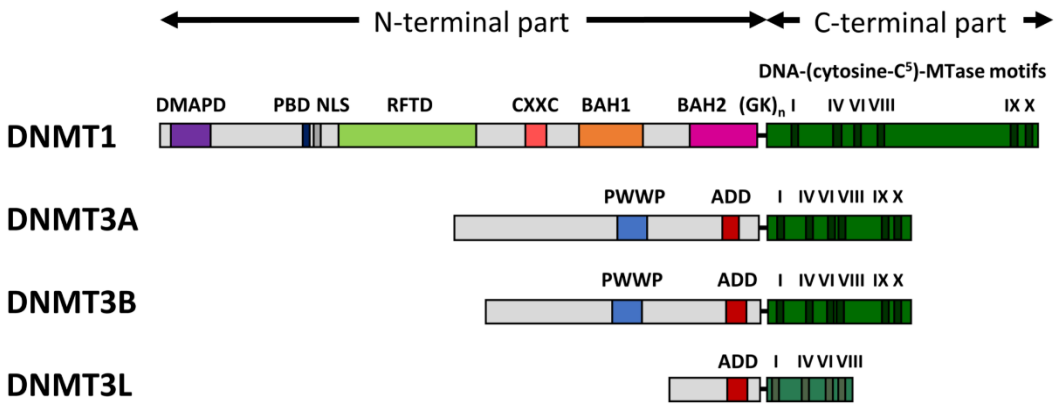


Figure 9: Schematic domain composition of the human DNA methyltransferases DNMT1, DNMT3A and 3B, as well as the catalytically inactive variant DNMT3L. The motifs of the catalytic domain are indicated by roman numerals (Jeltsch and Jurkowska, 2016).

Nevertheless, it was reported that DNMT3L has as strong stimulatory effect on DNMT3 activity (Chedin et al., 2002). A crystal structure of a DNMT3A-DNMT3L complex bound to DNA shed some light on the molecular mechanisms of this stimulation: It was found that DNMT3A does not act as a monomer but instead forms a heterotetramer consisting of two central DNMT3A subunits accompanied by two DNMT3L subunits bound to the sides of the complex (Jia et al., 2007). In this conformation, the whole complex is bound to DNA in a way that places the catalytic sites of the two inner DNMT3A subunits close to the DNA, which enables efficient methylation.

The mechanism of cytosine methylation at position C-5 is depicted in Figure 10. As can be seen, a deprotonated cysteine residue from the catalytic motif IV serves as a nucleophile and attacks the pyrimidine ring at position C-6, which is accompanied by the shift of an electron pair and protonation of N-3 by a glutamic acid residue from motif VI. Then, as N-3 is deprotonated again, the activated C-5 can abstract a methyl group from the cofactor AdoMet. Finally, the C-5 position is deprotonated and the electron pairs shifts back to form a double bond to C-6, while the bond to the motif IV cysteine is broken. Like with protein methyltransferases, the cofactor AdoMet is used during the reaction and emerges as AdoHcy.

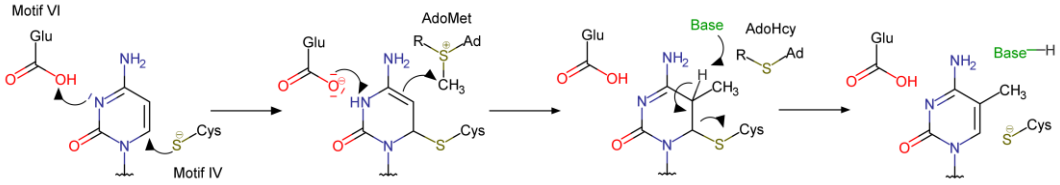


Figure 10: Reaction mechanism of cytosine methylation by DNMT enzymes with the cofactor AdoMet.

1.4.2 Dynamics of DNA methylation by DNMTs

Numerous studies have tried to uncover the precise cellular roles of the human DNMTs. As first proposed in 1975 by Holliday, Pugh, and Riggs, enzymes with two different activities are responsible for the establishment and maintenance of DNA methylation of putative palindromic motifs: These enzymes are classified either as *de novo* enzyme or maintenance enzyme and have no overlapping functions (Holliday and Pugh, 1975; Riggs, 1975). Later, it was found that DNA methylation mainly occurs on palindromic CpG sites, where both cytosines are methylated. However, the methylation on one strand would be lost during DNA replication, leaving the daughter DNA strand completely unmethylated. Repeated replication steps would therefore result in a loss of DNA methylation patterns. The newly discovered DNMT enzymes displayed activities according to the model proposed by Holliday, Pugh, and Riggs: The function of DNMT1 in this setting is the recognition of hemimethylated CpG sites and performing methylation of the unmethylated cytosine to restore the previous pattern, hence the term maintenance methyltransferase. The DNMT3 enzymes on the other hand are the only enzymes able to act on unmethylated DNA and are solely responsible for the methylation of both cytosines of unmethylated CpG sites to establish new methylation patterns. However, the strict separation of these enzymes into two distinct groups proved to

be an oversimplification and was not supported by advanced experimental procedures (Jones and Liang, 2009; Jeltsch and Jurkowska, 2014).

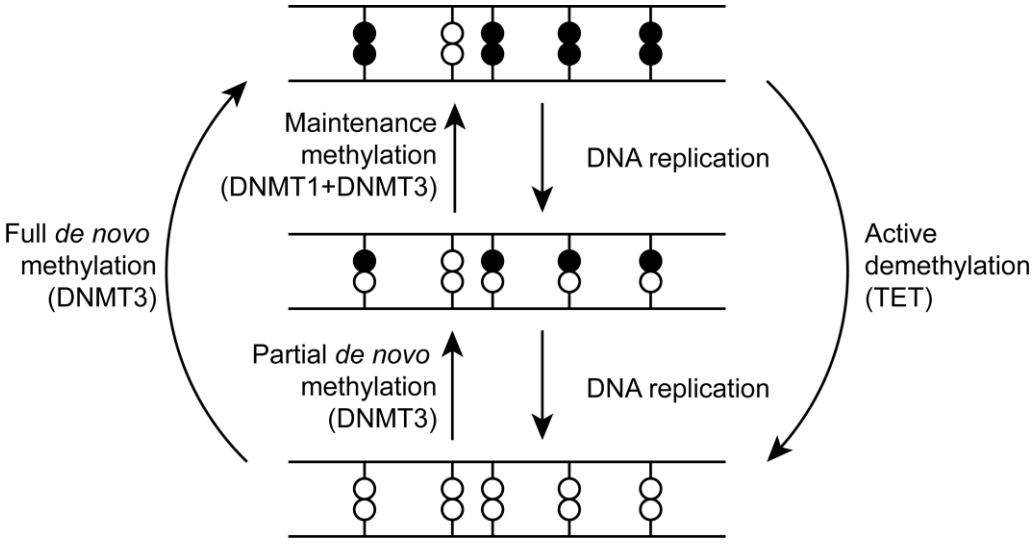


Figure 11: Schematic model of the DNA methylation and demethylation mechanismns as well as the roles of DNA methyltransferases.

The lines between maintenance and *de novo* enzymes have since then blurred, as both enzyme groups have been shown to be involved in both processes. For instance, DNMT1 can play a role in *de novo* methylation by completing the methylation pattern in the case where only one cytosine per CpG site is methylated by DNMT3 (Egger Gerda et al., 2006). Likewise, the DNMT3 enzymes were also shown to be involved in maintenance DNA methylation by remaining bound to nucleosomal DNA after replication (Dodge et al., 2005; Jones and Liang, 2009). Finally, the discovery of active DNA demethylation by the ten-eleven-translocation (TET) enzymes (Rasmussen and Helin, 2016) has completed the model of DNA methylation dynamics as it is depicted in Figure 11. In summary, the methylation of CpG sites is a highly dynamic process and can be described as the sum of all de- and remethylation processes, for which the equilibrium can reach any state between stable methylation or stable demethylation (Jeltsch and Jurkowska, 2014).

1.4.3 Regulation of DNMT3 enzymes

As shown in Figure 9, the *de novo* DNA methyltransferases DNMT3A and DNMT3B carry two important regulatory domains in their N-terminal region, which facilitate targeting and binding of the enzymes to DNA. The first of these domains is the PWWP domain, which was initially described as a weak DNA binding domain (Qiu et al., 2002). The name originates from the one-letter code of the amino acids proline and tryptophan which make up the sequence motif,

although this sequence is only weakly conserved. In fact, the DNMT3 PWWP domain actually has the sequence SWWP. The important feature of the PWWP motif is its ability to bind the histone tail at the di- and trimethylated lysine 36 (H3K36me2/3) (Dhayalan et al., 2010). This binding is facilitated by interactions of the trimethylammonium group with an aromatic cage formed mostly by the two tryptophan residues (Rondelet et al., 2016; Wang et al., 2020). The PWWP domain mediates targeting of DNMT3 to heterochromatic regions (Chen et al., 2004; Ge et al., 2004) and contributes to subnuclear localization of the enzyme (Dukatz et al., 2019). Another important aspect of this interaction is the mechanism described in section 1.3.2: To prevent spurious transcription initiation, DNMT3 is recruited to gene bodies by the H3K36me3 mark and deposits DNA methylation there (Neri et al., 2017).

The second relevant domain of DNMT3, the ADD domain, plays a role in both regulation of enzyme activity and targeting. It is a unique cysteine-rich domain found only in the proteins ATRX (Gibbons et al., 1997) as well as DNMT3A and DNMT3B and was therefore named after these proteins. The ADD domain is able to bind the histone 3 tail at position 4 (H3K4), but only in its unmethylated state (Otani et al., 2009; Zhang et al., 2010). This property explains the mutual exclusiveness of H3K4me3, an activating histone mark, and DNA methylation, a repressive mark, in the promoter region of the same gene (see Figure 5 and Figure 8). Additionally, the binding of the unmethylated H3 tail serves a direct function in enzyme activation: Without a bound H3 tail, the ADD domain takes up an autoinhibitory conformation, essentially blocking the catalytic site. Upon H3 tail binding, a hinge-like conformational change occurs, which folds away the entire ADD domain and results in unimpeded access to the catalytic site (Guo et al., 2015; Jeltsch and Jurkowska, 2016). A crystal structure showing this allosteric enzyme activation is depicted in Figure 12. Consequently, the ADD domain fulfills two roles at once: first, it targets the enzyme to specific regions and then only activates the enzyme once the correct substrate is bound.

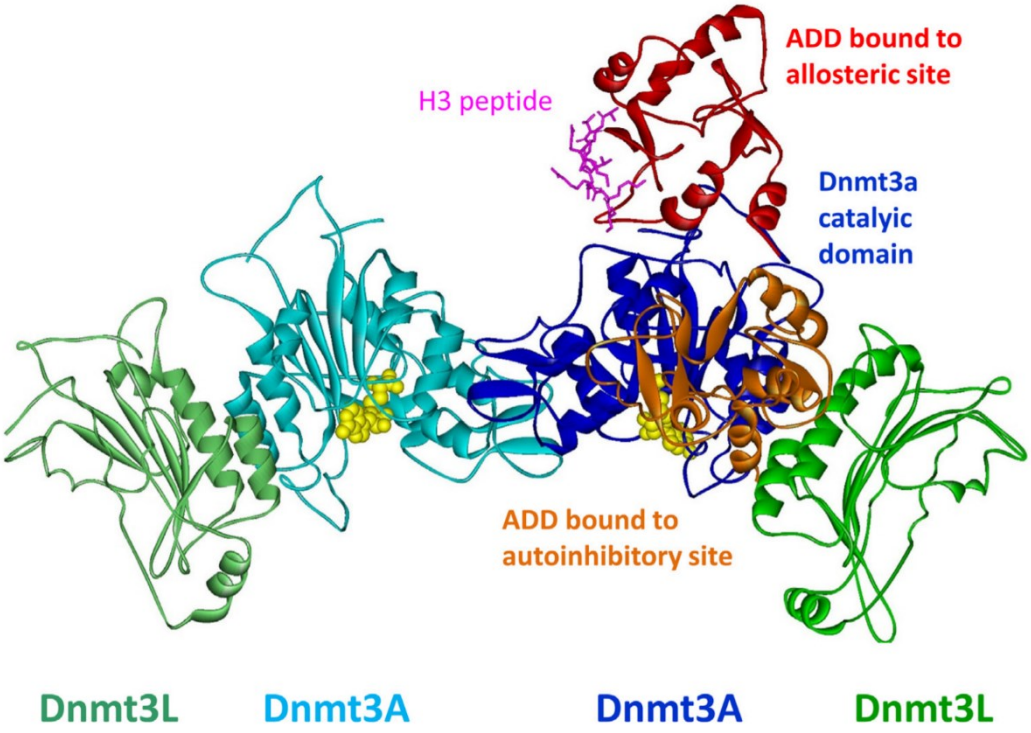


Figure 12: Crystal structure of a heterotetrameric DNMT3A/DNMT3L complex. The conformational change between autoinhibition by the ADD domain and the active conformation achieved upon H3-tail binding is displayed in orange and red, respectively (Jeltsch and Jurkowska, 2016).

2 Principal aims of this study

This thesis consists of three main projects, which aimed to gain novel insights into the function of DNA and histone methyltransferases. These projects share one common feature, which is the application of recombinant mononucleosomes as an *in vitro* study tool. For many *in vitro* experiments, short substrates such as DNA fragments in the case of DNA methyltransferases or peptides in the case of histone methyltransferase are frequently used in biochemical studies. There substrates are sufficient for screening purposes but are sometimes unable to fully capture the relevant molecular interactions with the enzyme. Nucleosomes however represent a physiologically relevant substrate for many chromatin-modifying enzymes including the ones studied here. A graphical representation of the projects which are part of this thesis is depicted in Figure 13.

In the first project, the methylation of nucleosomal linker DNA by the DNA methyltransferase DNMT3A was investigated. Prior to this project, a cryo-EM structure depicting a DNMT3A/DNMT3B3 heterotetramer in complex with a nucleosome particle was published (Xu et al., 2020). This structure has revealed a distinct binding mode of the enzyme complex to the acidic patch of the nucleosome, which results in an orientation towards the linker DNA. Since methylation of nucleosomal linker DNA has not yet been studied adequately at base-pair resolution, the aim of this project was to gain an understanding of linker DNA methylation by DNMT3A in the context of the published structural data.

To this end, it was aspired to generate recombinant mononucleosome particles with a suitable linker DNA fragment, which contains CpG sites at the same positions as in the cryo-EM structure. These nucleosomes were to be used for *in vitro* methylation by the full length DNMT3A2 enzyme (one of the DNMT3A isoforms) and the catalytic domain of DNMT3A, termed DNMT3AC, and subsequent bisulfite sequencing to demonstrate the feasibility of this approach. Furthermore, an experimental and analytical pipeline for NGS library generation and analysis had to be established. In order to gain even more relevant data in context with the published structure, it was attempted to establish a chromatography workflow to selectively purify DNMT3AC/DNMT3B3C heterotetramers, which were to be used in analogous nucleosome methylation experiments. Furthermore, the effect of the H3K4me3 and H3K36me3 histone modifications on the catalytic activity of DNMT3A was to be investigated, since these modifications play a critical role in recruitment and activation of the enzyme by interaction with the ADD and PWWP domain. To this end, it was necessary to prepare the respective modified nucleosomes by chemical conversion of a previously introduced cysteine to a trimethyllysine analog. The experiments with these modified nucleosomes were to be conducted with DNMT3A2 and DNMT3AC, which lacks the regulatory ADD and PWWP domains, in order to be able to draw conclusions about the

involvement of these domains in linker DNA methylation and H3 tail interaction. In summary, this study aimed to gather new insight into nucleosomal DNA methylation by DNMT3A in the context of previously reported structural data.

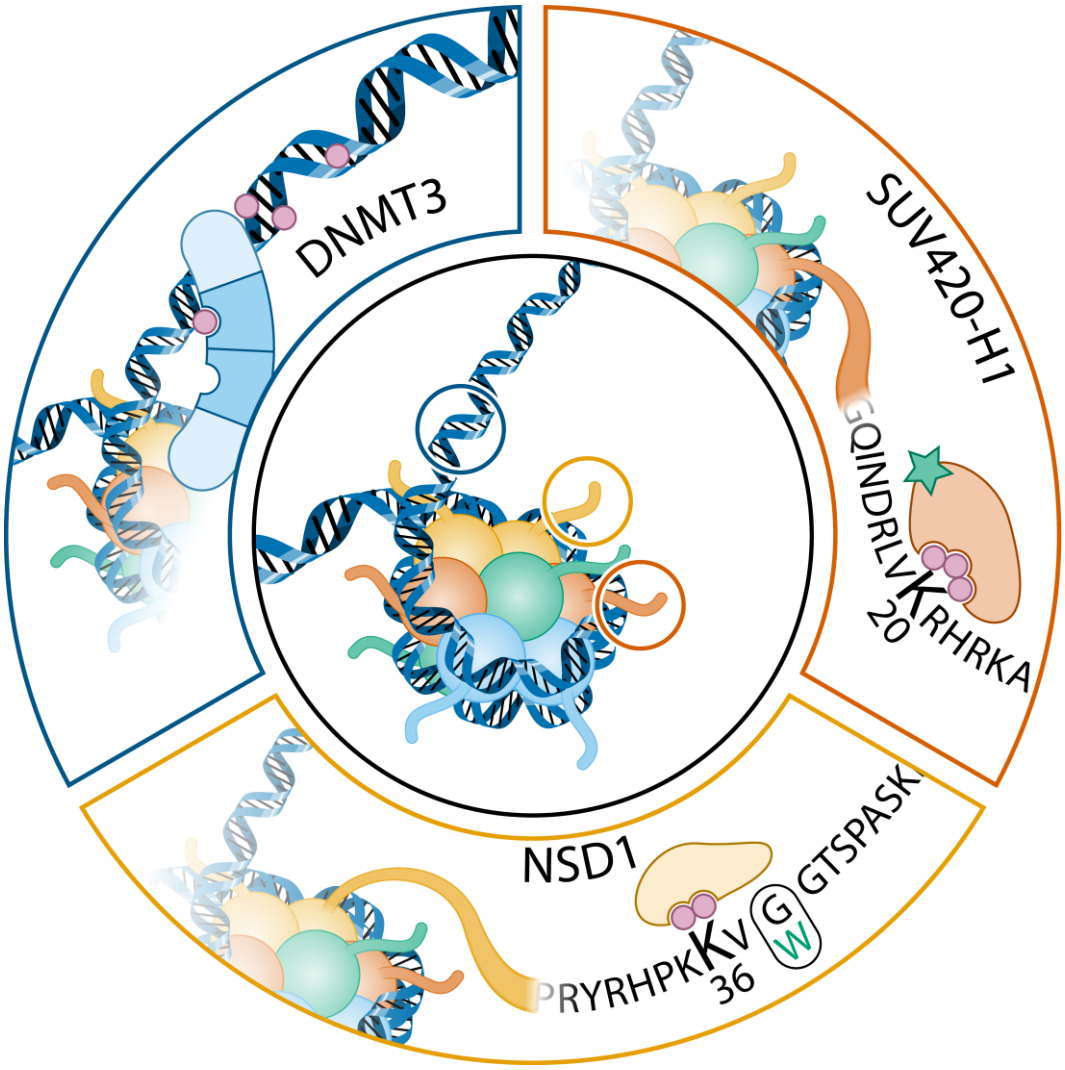


Figure 13: Graphical summary of the research projects in this thesis. All projects are connected by the usage of recombinant mononucleosomes and cover the methylation of linker DNA by DNMT3A, the methylation of H4K20 by mutant SUV420H1 variants, and the methylation of the H3.3 G34W oncohistone by NSD1.

The second project of this thesis dealt with the H4K20 methyltransferase SUV420H1, one of the three major enzymes responsible for modification of this site in addition to SUV420H2 and SET8. Point mutations in this enzyme are frequently associated with a variety of human cancers, a matter that that might be explained by aberrant function of this enzyme and resulting epigenomic dysregulation.

Therefore, it was the aim of this project to gain an understanding of how cancer-associated mutations affect the catalytic activity of SUV420H1. To this end, a multi-step approach was contrived: First, the methylation activity of SUV420H1 was to be tested using peptides as substrate, which potentially also allows for binding studies by employing peptide pulldowns. Second, different experiments with nucleosomes were devised. First, since the workflow for the reconstitution of recombinant mononucleosomes was established in parallel to this work, these nucleosomes were to be investigated in methylation experiments in comparison with H4 protein. Furthermore, native mononucleosomes isolated from SUV420H1/H2 dko cells, which represent an optimal substrate for SUV420H1 due to their putative extensive H4K20me1 methylation state, were planned to be tested as a physiologically relevant substrate for further *in vitro* studies of SUV420H1 mutants. Lastly, it was aimed to employ the same SUV420H1/H2 dko cells to study the catalytic activity of SUV420H1 mutants in cells. Collectively, this study aimed to investigate the effects of selected cancer-associated mutants of SUV420H1 on the methylation of peptide and nucleosome substrates as well as in cells in order to gain novel insights into the functional role of the affected residues and the potential pathological mechanisms of the mutations.

The third and final project of this thesis focused on the H3.3 G34W oncohistone mutation, which has come into the focus of research in the recent years. The G34W mutation is a distinctive factor in giant cell tumors of the bone (GCTB), as it is found in around 90 % of the affected patients. The H3.3 histone variant is a vulnerable target for mutations, because only two genetic copies of the proteins exist in contrast to the ten copies of H3.1. Furthermore, H3.3 is expressed independent of the cell cycle and serves distinctive roles is genome regulation. Due to the close proximity of the G34W mutation to the important K36 site, combined with the severity of the glycine to tryptophan exchange, it stands to reason that G34W could affect the methylation of K36 and thereby cause epigenomic disturbances. Therefore, it was the aim of this project to decipher how the key H3K36 methyltransferases NSD1, NSD2, and SETD2 are influenced by the G34W mutation. In order to investigate this, methylation experiments using modified peptides as well as recombinant nucleosomes containing H3.3 with or without the G34W mutation were to be used in methylation experiments with the aforementioned enzymes. Depending on the observed effect of the G34W mutation on the catalytic activity of these enzymes, a model of how this mutation leads to global epigenomic changes was to be developed.

3 Results

3.1 Linker DNA methylation of recombinant mononucleosomes by DNMT3A

Nucleosome positioning is an important regulatory factor in DNA methylation, because the *de novo* DNA methyltransferases DNMT3A and DNMT3B are known to mainly act on linker DNA in a nucleosomal context (Takeshima et al., 2008; Felle et al., 2011). Furthermore, histone modifications control and modulate the binding and activity of the DNMT3 complexes to nucleosomal DNA by interaction with their ADD and PWWP domain (Dhayalan et al., 2010; Zhang et al., 2010; Dukatz et al., 2019). In a recent fundamental study, the cryo-EM structure of DNMT3A/3B3 complex bound to a mononucleosome was published (PDB entry: 6PA7) (Xu et al., 2020). In this setting, is was shown that one of the outer DNMT3B3 subunits binds to the nucleosome disc face via two arginine residues in the acidic patch, while the inner DNMT3A subunits form a bridge-like structure towards the linker DNA near the nucleosome dyad axis. Due to the close sequence similarity between DNMT3A and DNMT3B, it can be assumed that a DNMT3A homotetramer forms a similar structure. However, methylation of nucleosomal DNA by DNMT3A has not been studied at a level sufficient to put it in context with the cryo-EM structure.

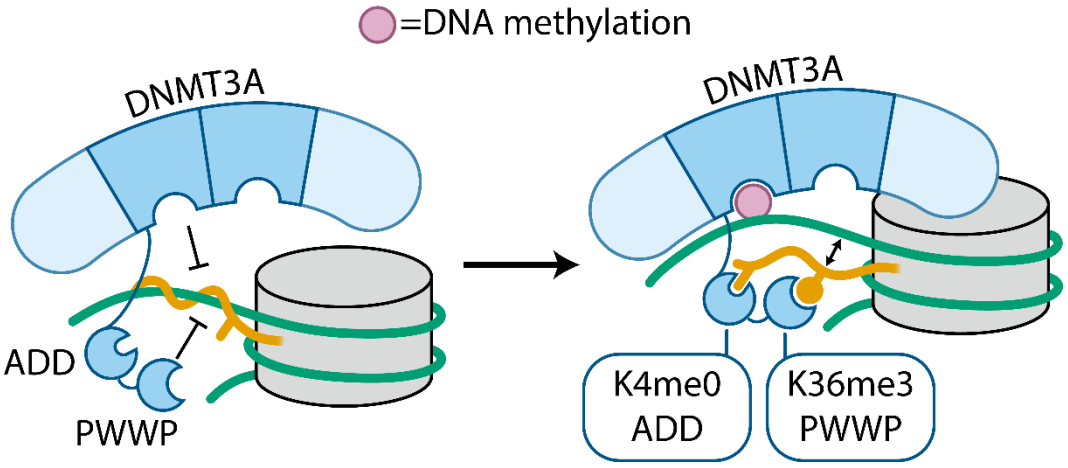


Figure 14: Graphical summary of the findings in this study. It was shown the DNMT3A preferentially methylates nucleosomal linker DNA, specifically one, according to structural data, conveniently positioned CpG site. Furthermore, a inhibitory effect of H3K4me3 and a stimulatory effect of H3K36me3 was observed. The latter could be in part attributed to binding dynamics between the H3 tail and the linker DNA.

In this study, the methylation of nucleosomal DNA by different DNMT3 constructs was investigated. First, the full length DNMT3A2, which is one of the murine DNMT3A isoforms. Then, the catalytic domain of DNMT3A, termed DNMT3AC.

Finally, the complex of the catalytic domains of DNMT3A and the DNMT3B isoform DNMT3B3 (DNMT3AC/DNMT3B3C) were investigated. Additionally, the effect of the H3K$_C$4me3 and H3K$_C$36me3 trimethyllysine analogs was evaluated. As depicted in Figure 14, methylation was observed primarily on linker DNA with specific CpG site methylation in agreement with the cryo-EM structure. Additionally, an inhibiting effect of H3K$_C$4me3 and a stimulating effect of H3K$_C$36me3, which was independent of the DNMT3A PWWP domain, were observed. These findings were published in the journal *Communications biology* (Bröhm et al., 2022a)(Appendix 1) and will be summarized here only.

3.1.1 Investigation of nucleosomal DNA methylation

In order to study the methylation of nucleosomal DNA by DNMT3A, first suitable nucleosome substrates were generated as described in chapter 5.6. To this end, a DNA fragment was used which incorporates a CpG-rich linker DNA in addition to the Widom 601 binding sequence. Specifically, the linker DNA contains two CpG sites in the same relative position as the fragment in the DNMT3A/3B3-mononucleosome structure. A sequence alignment of the DNA fragment used in this study and other relevant sequences is shown in Figure 15A. Using the amplified 240 bp fragment, different mononucleosomes were reconstituted, which were either unmodified or contained the H3K$_C$4me3 or H3K$_C$36me3 trimethyllysine analog. Additional data regarding trimethyllysine analog generation, histone octamer refolding, and nucleosome reconstitution are shown in Appendix 1. These mononucleosomes were then used for DNA methylation experiments as described in chapter 5.7. The workflow for DNA methylation and subsequent library preparation for Illumina NGS is depicted in Figure 15B. Briefly, in the first step excess free DNA was cleaved with the MluI restriction enzyme and the DNA methylation was subsequently carried out. Then, all DNA-bound proteins were digested by proteinase K treatment and the resulting free DNA was purified and used for bisulfite conversion. The converted DNA was then amplified in two PCR steps to add barcodes, indices, and adaptors for Illumina paired-end sequencing. Control experiments related to this process are shown in Appendix 1. The NGS sequencing data was analyzed using a local instance of the Galaxy server as described in chapter 5.7.

In the methylation experiments with DNMT3A2, a characteristic methylation pattern of the individual CpG sites was observed: The first six CpG sites residing in the linker DNA were preferably methylated with methylation levels ranging from 20 % to 60 %, while only miniscule methylation was detected over the nucleosome binding sequence (Figure 15C). In control experiments with nucleosome-free DNA, methylation was more widespread and observed over the complete sequence. The

same characteristic nucleosome DNA methylation pattern was also observed with DNMT3AC and with nucleosomes modified with H3K$_C$4me3 or H3K$_C$36me3. The methylation level differences between CpG sites on the linker DNA is furthermore fitting previous work on the flanking sequence preference of DNMT3A (Appendix 1).

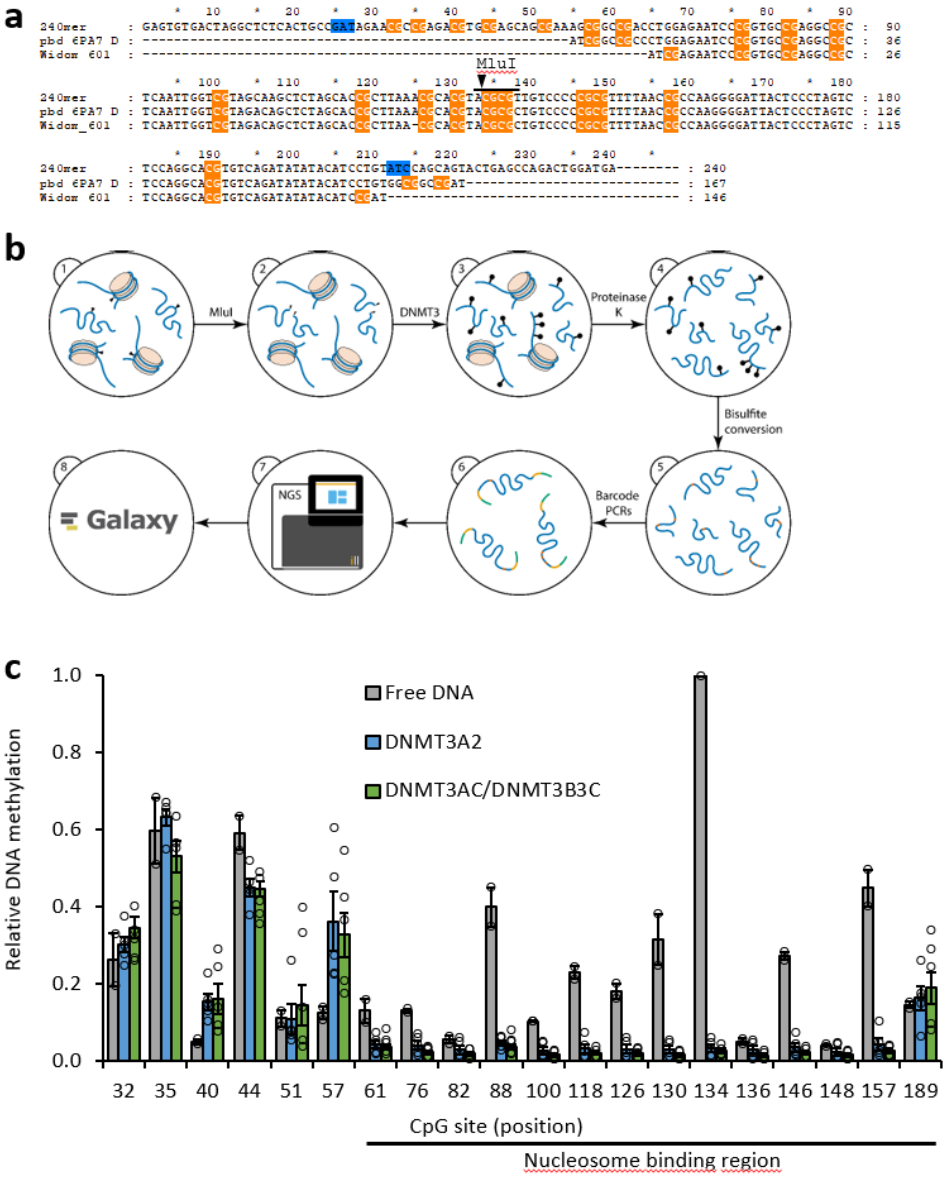


Figure 15: Investigation of nucleosomal DNA methylation. (A) Sequence alignment of the nucleosomal DNA used in this study, the fragment used in the PDB structure 6PA7, and the canonical Widom 601 sequence. (B) Schematic workflow for nucleosome methylation and NGS. The image was taken from (Bröhm et al., 2022).

3.1.2 Linker DNA methylation by DNMT3A2 and DNMT3AC/3B3C heterotetramer is in agreement with the cryo-EM structure

As described earlier, the cryo-EM structure of a heterotetrameric DNMT3A/3B3 complex has revealed direct binding of one of the outer DNMT3B3 subunits to the nucleosome and the orientation of the complex towards the linker DNA. In this orientation, the DNMT3A subunit closer to the nucleosome is positioned too far away from the DNA to be able to methylate any CpG at this site. The second, more distal DNMT3A subunit however is directly anchored to the DNA and its catalytic site is well within a distance that allows for base flipping (Figure 16A). Therefore, it would be plausible to assume that this DNMT3A subunit is able to methylate the DNA at this site. The linker DNA used here contains two CpG sites at the same position as in the cryo-EM structure, one of which is located precisely near the active site of the distal DNMT3A subunit (site 57, Figure 16A). As depicted in Figure 16B, the observed methylation levels at CpG 57 and 61 are in agreement with the proposed behavior of the enzyme based on the cryo-EM structure: While site 61 shows only a low methylation level comparable to that of free DNA, site 57 shows significantly increased methylation compared to free DNA, which cannot be explained by sequence preference. However, uncertainties remain because it cannot be ruled out that DNMT3A2 alone behaves differently than the DNMT3A/3B3 complex in the cryo-EM structure.

While the nucleosome methylation experiments using only DNMT3A2 should also be relevant due to high amino acid sequence similarity (Appendix 1) between DNMT3A and DNMT3B, it was nevertheless attempted to accurately replicate these findings using a similar heterotetrameric complex. To this end, a double-tag purification strategy was employed to selectively isolate complexes of DNMT3AC/3B3C with the desired stoichiometric composition. The purification procedure is detailed in chapter 5.1.1 and in Appendix 1. As depicted in Figure 15C and D as well as Figure 16B, methylation experiments using the purified heterotetramer complexes showed a pattern which closely resembles that of DNMT3A2. Ultimately, the relative CpG preferences of both DNMT3A2 and the DNMT3AC/3B3C complex agree very well with the cryo-EM structure.

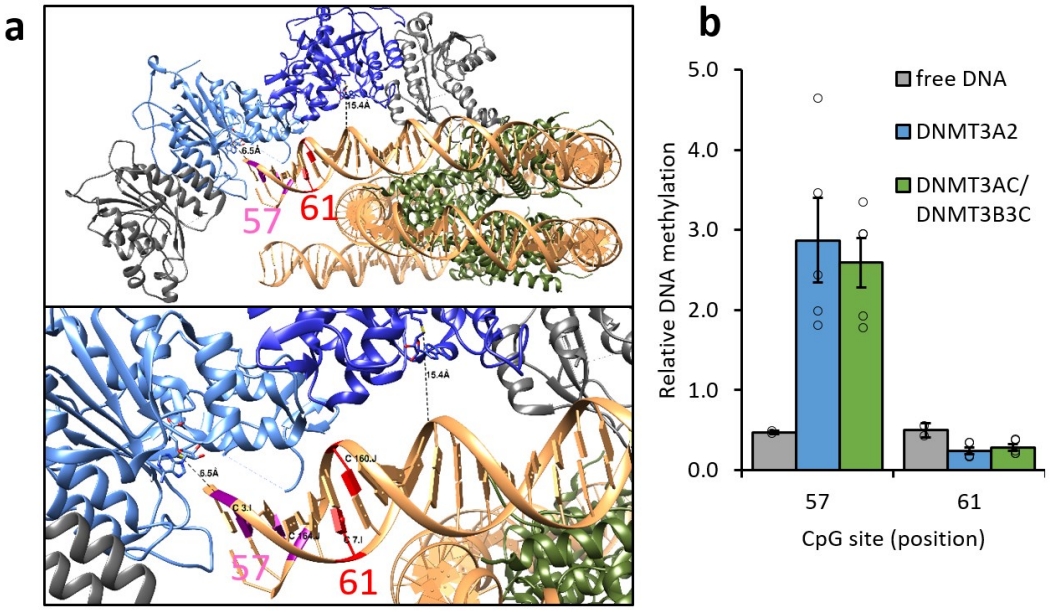


Figure 16: Methylation levels of CpG sites 57 and 61 in context of the cryo-EM structure. (A) Position of CpG site 57 and 61 in the cryo-EM structure. (B) Relative DNA methylation of the two sites compared between free DNA and nucleosomes methylated with DNMT3A2. The image was taken from (Bröhm et al., 2022).

3.1.3 Modulation of DNMT3A methylation activity by histone modifications

Next, the influence of histone modifications on the activity of DNMT3A was to be investigated. The H3K4me0-binding ADD domain and the H3K36me2/3-binding PWWP domain are involved in targeting and regulation of DNMT3A. Consequently, introduction of H3K4me3 into nucleosomes should result in reduced catalytic activity, since the ADD domain cannot bind to the H3 tails anymore and the enzyme is less likely to assume a catalytically active state. To test this hypothesis, nucleosomes containing the trimethyllysine analog H3K$_C$4me3 were generated, methylated with DNMT3A2 in competition with unmodified nucleosomes and processed analogous to previous experiments.

As shown in Figure 17A, the methylation pattern of H3K$_C$4me3-modified nuclesosomes was generally similar with most methylation observed on the first six CpG sites and almost no methylation in the nucleosome binding sequence. However, some differences can be observed: Some CpG sites close to the nucleosome show decreased methylation levels upon H3K$_C$4me3 introduction, while the outermost CpG site shows a slight increase in methylation. This effect is visualized in Figure 17B, where the ratio between methylation levels of modified and unmodified nucleosomes is shown for each CpG in the linker region. Overall, this

observation is in agreement with the hypothesized effect of H3K$_C$4me3: Since ADD binding to the H3 tail is inhibited by H3K$_C$4me3, methylation efficiency decreases close to the nucleosome. The enzyme instead accumulates on the free linker DNA, which leads to a slight increase in methylation observed on the most distal CpG site.

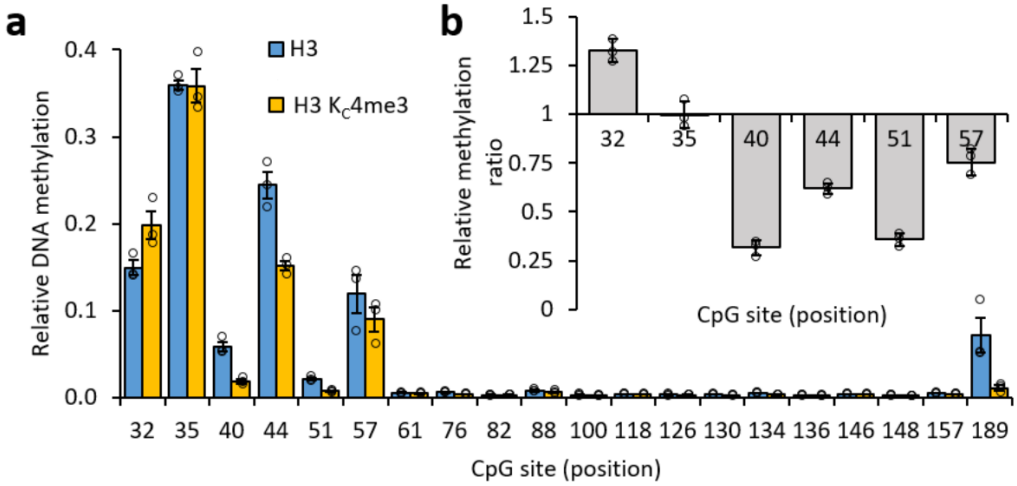


Figure 17: Effect of H3K$_C$4me3 on DNMT3A2 methylation. (A) Relative CpG methylation by DNMT3A2 of unmodified and H3K$_C$4me3 nucleosomes. (B) Ratio of relative methylation levels between modified and unmodified nucleosomes for the first six CpG sites. The image was adapted from (Bröhm et al., 2022).

In a similar approach as with H3K$_C$4me3, the H3K$_C$36me3 trimethyllysine analog was introduced into nucleosomes and methylation reactions in competition with unmodified nucleosomes were performed with DNMT3A2. The result can be seen in Figure 18A: Overall, the same methylation pattern as in previous experiments could be observed, but the H3K$_C$36me3-modified nucleosomes showed a significantly increased methylation level at the three most strongly methylated CpGs (32, 35, and 40) of about 1.6-fold (p-value 1.15×10^{-4}). This stimulation can be attributed to the binding of the PWWP domain to trimethylated H3K36 leading to improved recruitment of DNMT3A2 to these substrates.

The same experiment was also repeated using DNMT3AC, which comprises only the catalytic domain without ADD and PWWP domain. It was expected that the stimulatory effect of H3K$_C$36me3 should be alleviated in this setting, however, the opposite was observed: As shown in Figure 18B, methylation of the linker DNA by DNMT3AC was strongly increased for all CpG sites. In comparison with DNMT3A2, where the stimulation was mostly observed with the CpG sites 32, 35, and 40, methylation level of the same sites was increased by a factor of 3.7 (p-value 2.32×10^{-4}). Since the PWWP domain is not present in this context, the strong

stimulation of the activity of DNMT3AC must be attributed to a mechanism unrelated to the K36me2/3-PWWP interaction.

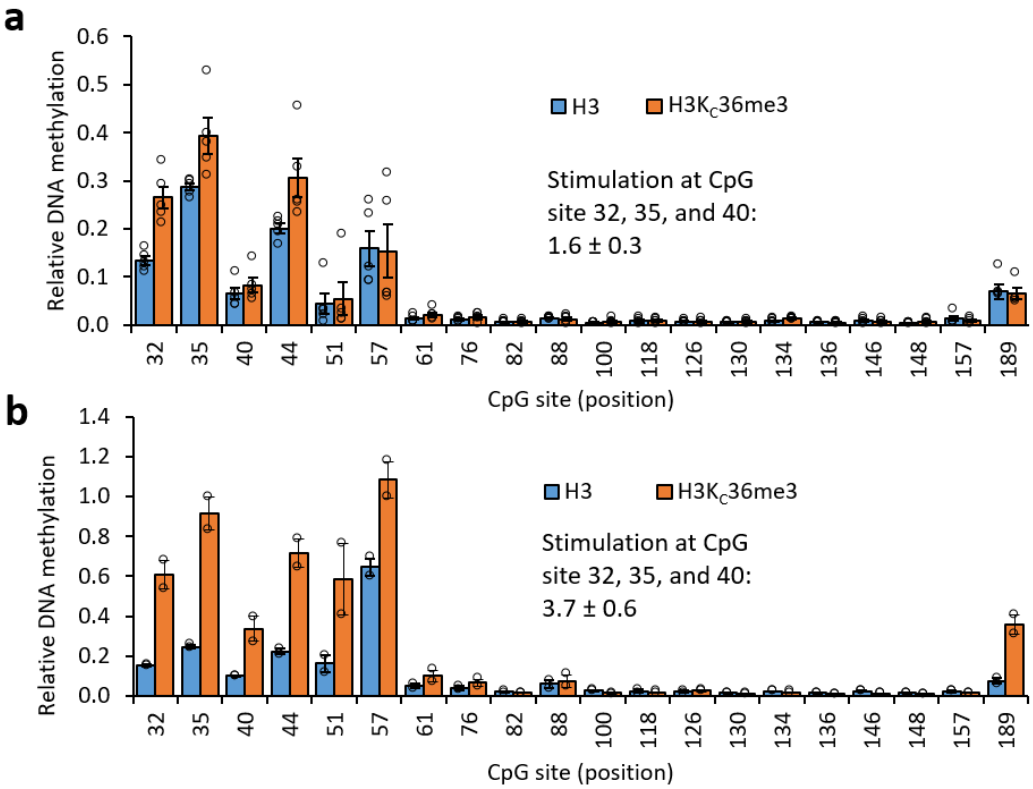


Figure 18: Methylation of unmodified and H3K$_C$36me3 containing nucleosomes by DNMT3A2 and DNMT3AC. (A) Relative CpG methylation of modified and unmodified nucleosomes by DNMT3A2. (B) Relative CpG methylation of modified and unmodified nucleosomes by DNMT3AC. The image was taken from (Bröhm et al., 2022).

3.1.4 Binding dynamics of the H3 tail is influenced by H3K$_C$36me3

The puzzling finding that H3K$_C$36me3 stimulates DNA methylation by DNMT3AC without interaction with the PWWP domain could be explained by weakened binding of the H3 tail to the linker DNA, which would increase the accessibility of the linker DNA for methylation. In order to support this hypothesis, the conformational dynamics of the H3 tail upon introduction of H3K$_C$36me3 were investigated by fluorescence spectroscopy. To this end, a tryptophan residue was introduced into the H3 tail at position 15 (A15W) by site-directed mutagenesis and corresponding nucleosomes with and without H3K$_C$36me3 were generated. Since the A15W residue comprises the only tryptophan residue in the whole nucleosome, it can be used as an H3-tail specific fluorophore for the following fluorescence spectroscopy experiments. It was expected that intercalation of tryptophan into DNA would quench fluorescence due to the tight interactions with DNA bases. Therefore, increased fluorescence should be observed with nucleosomes containing H3K$_C$36me3 if it is really a determining factor in the binding of the H3 tail to the linker DNA.

As can be seen in Figure 19, after careful and accurate adjustment of their concentration as described in detail in the attached publication (Appendix 1), the fluorescence spectra of nucleosomes with and without H3K$_C$36me3 showed an increased fluorescence of the H3K$_C$36me3-containing nucleosomes by a factor of 1.5. Consequently, the previously observed effect of increased methylation activity of DNMT3AC can be attributed to the hypothesized weakening of the H3 tail-DNA interaction.

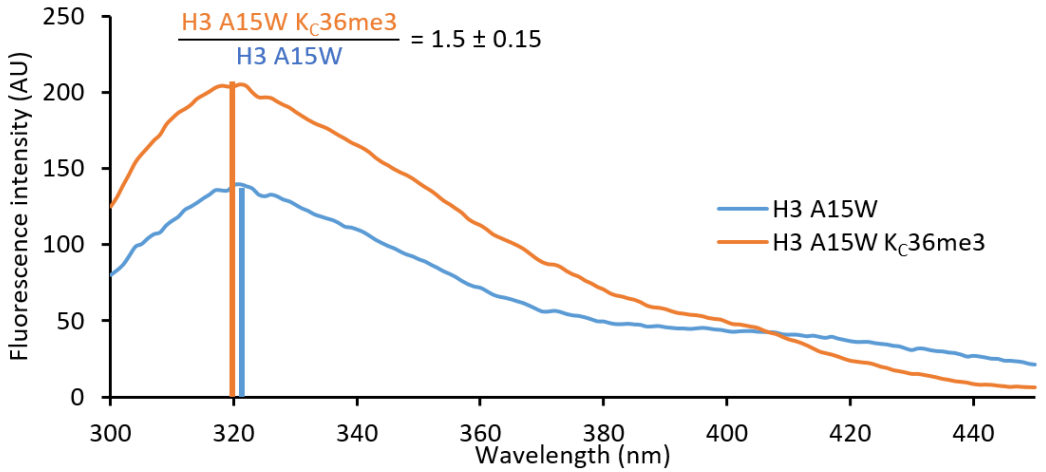


Figure 19: Fluorescence spectra of nucleosomes containing the A15W mutation either with or without H3KC36me3. The image was taken from (Bröhm et al., 2022).

3.2 Cancer mutants of the SUV420H1 methyltransferase modulate its catalytic activity

Aberrant function of PKMTs by loss or gain of function is a frequent driving factor in cancer and other diseases (Martin and Zhang, 2005; Greer and Shi, 2012). In this study, the effect of several somatic cancer mutations of the SUV420H1 methyltransferase on its catalytic activity was evaluated using peptide substrates as well as nucleosomal substrates *in vitro* and in cells. A graphical summary of the resulting findings is depicted in Figure 20. The eight mutant proteins of SUV420H1 can be divided in three categories according to their activity on different substrates: First, variants, which retain catalytic activity, albeit reduced, in all tested experimental setups. Category two variants did not exhibit activity on protein substrates but were able to methylate nucleosome substrates *in vitro* and in cells. Proteins of the last category showed no detectable activity on any tested substrate. These findings were published in the *Journal of molecular biology* (Appendix 2) and are only summarized here (Bröhm et al., 2019).

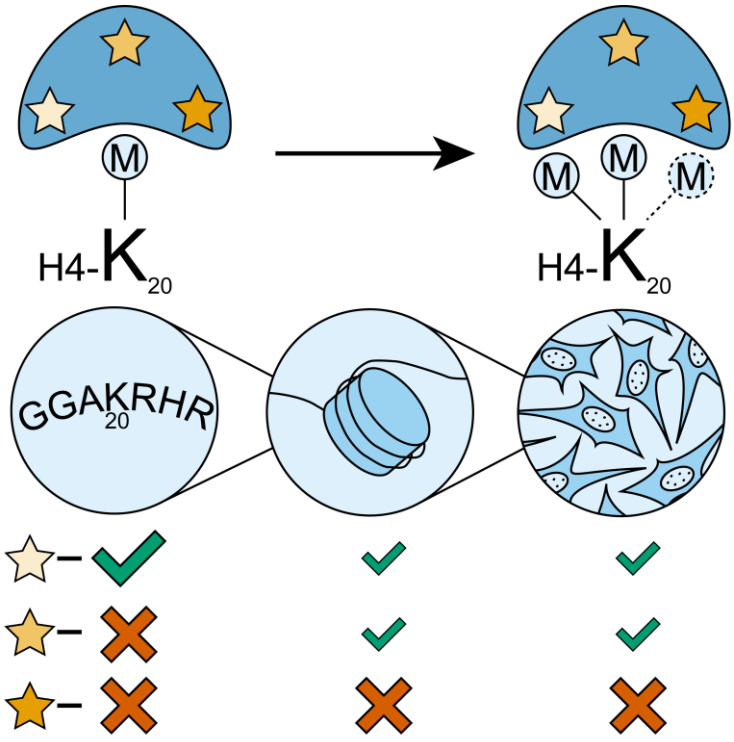


Figure 20: Graphical summary of the findings regarding the catalytic activity of SUV420H1 mutants. The enzyme mutants (indicated by stars) are categorized according to their catalytic activity on different substrates. The image was taken from (Bröhm et al., 2019).

3.2.1 Selection and isolation of SUV420H1 mutant proteins

To identify potentially relevant mutant variants of the SUV420H1 protein, first the list of known somatic cancer mutations was obtained from the COSMIC (catalogue of somatic mutations in cancer) database (Forbes et al., 2017) and classified by the type of mutation (Figure 21A). It was apparent that the majority of mutations do not result in clear loss-of-function protein variants, which would be assumed for nonsense and frameshift mutations. Instead, nearly three quarters of the known mutations are missense mutations, which could modulate protein function due to amino acid exchanges. Further analysis of the mutations using the cBioportal database (Cerami et al., 2012; Gao et al., 2013) revealed that mutations of the SUV420H1 protein are frequently associated with several human cancers such as colon, prostate, lung, esophagus, ovarian, and endometrial tumors (Figure 21B). Eight mutated residues were selected for further analysis: E238V, D249N, S255F, K258E, A269V, S283L, S304Y and E320K.

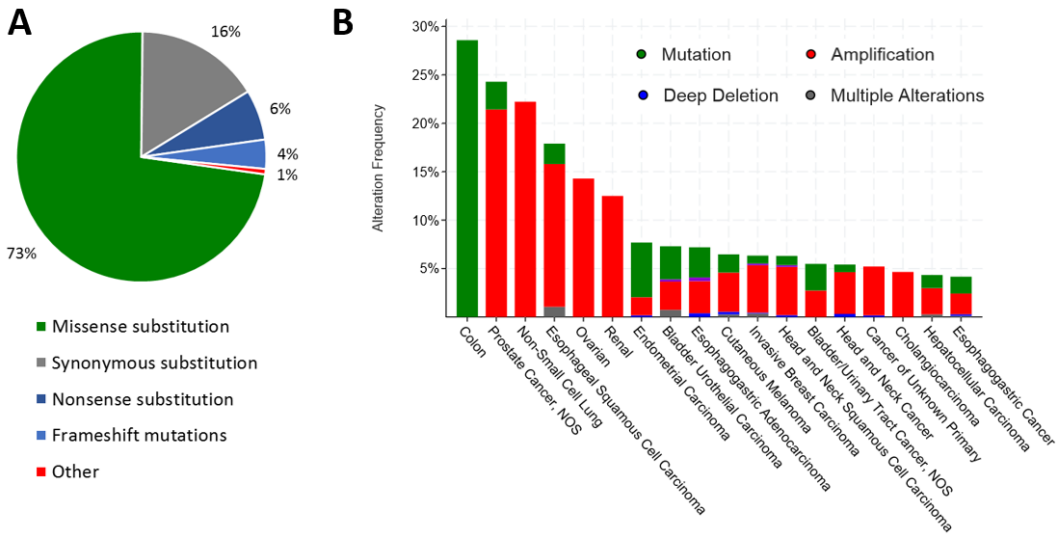


Figure 21: Overview of point mutations occurring in the SUV420H1 protein. (A) Classification of mutations based on their effect on the protein sequence. (B) Occurrence of mutations in various human cancer cells. The image was taken from (Bröhm et al., 2019).

In order to investigate the effect of the selected mutations on the catalytic activity of SUV420H1, the expression construct for the murine SUV420H1 catalytic domain was used to generate the respective mutant constructs by site-directed mutagenesis, which were subsequently overexpressed and purified as detailed in chapter 5.1.

3.2.2 Methylation of histone H4 peptide substrates by SUV420H1 mutants

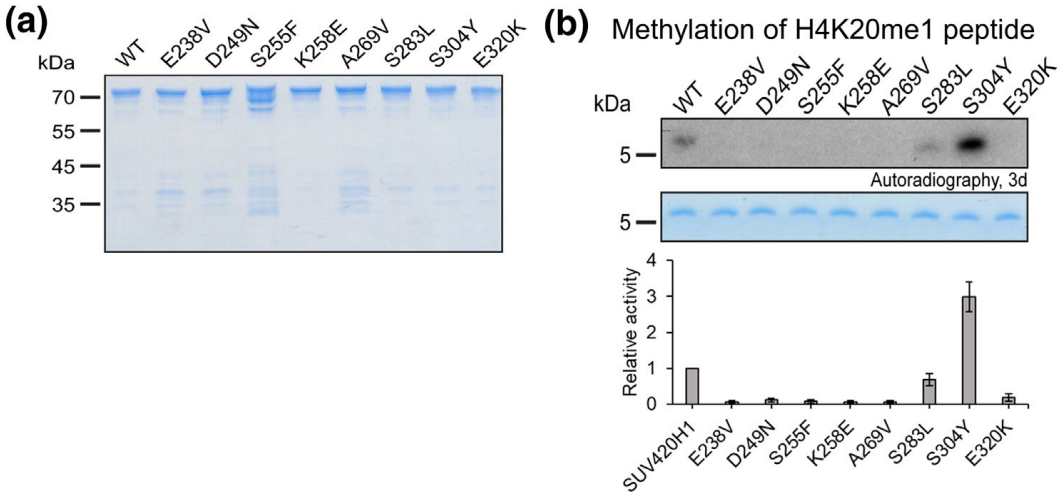


Figure 22: Investigation of SUV420H1 mutant activity on peptide substrates. (A) Purified SUV420H1 variant proteins. (B) Methylation of H4K20me1 peptide detected by autoradiography and quantification of three independent replicates. The image was taken from (Bröhm et al., 2019).

Successful purification of each of the SUV420H1 mutants was confirmed by SDS PAGE (Figure 22A), and the folding was confirmed by CD spectroscopy (Appendix 2). The enzymes were then used to perform radioactive methylation experiments with the H4 peptide (amino acids 12-29) containing K20 in monomethylated form, which presents the favored substrate methylation state of SUV420H1. As shown in Figure 22B, most of the mutant proteins did not exhibit detectable activity on this peptide, except for the S283L and S304Y variants. The S304Y mutant stands out because it displayed about threefold stimulated activity compared to the wild type. The same experiment was also conducted with unmodified (H4K20me0) and dimethylated (H4K20me2) substrate peptides. Here, a similar activity pattern was observed with a generally reduced activity being the only significant difference to the H4K20me1 peptide (Appendix 2). Additionally, the binding affinity of the H4K20me1 peptide to the different proteins was evaluated by peptide pulldown experiments and subsequent detection of the pulled protein by western blot against GST. In these experiments, only the K258E mutant showed a significantly reduced pulldown efficiency compared to the wild type while all other mutants were unaffected (Appendix 2).

3.2.3 Catalytic activity of SUV420H1 mutants on nucleosome substrates

Short peptides or even full-length histone proteins do not present an ideal substrate for PKMTs because additional contacts to the nucleosome complex are formed in cells, which enable and modulate their catalytic function. In a methylation experiment with recombinant H4 protein and recombinant nucleosomes, a greatly increased activity of SUV420H1 on the nucleosome substrate was detected (Figure 23A and B). This demonstrates the value of nucleosome substrates, however, the unmodified nucleosomes used here do not carry the H4K20me1 modification and are therefore not the ideal substrate. To test the activity of the selected SUV420H1 mutants on nucleosomes with H4K20me1, the SUV420H1/H2 double knockout (dko) MEF cells were utilized. Without the enzymes which di- and trimethylate H4K20, the H4K20me1 state accumulates in these cells, making nucleosomes isolated from this source a very suitable substrate to use. Hence, for follow-up *in vitro* methylation experiments, native nucleosomes extracted from these cells were employed as described in section 5.5 and Appendix 2. As shown in Figure 23C and D, the catalytic activity of the SUV420H1 mutants on native nucleosomes follows a slightly different pattern compared to peptide substrates. While the S304Y mutant shows activity comparable to the wild type, the previously inactive mutants E238V, D249N, and E320K also display a detectable, although reduced, activity.

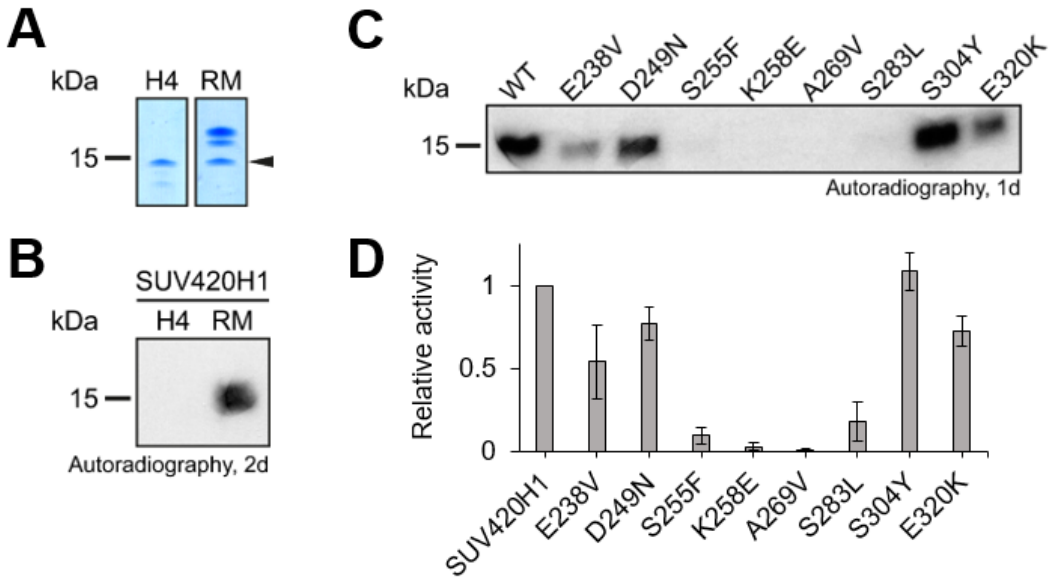


Figure 23: Methylation of nucleosome substrates by SUV420H1 *in vitro*. (A) Loading control of recombinant H4 protein and recombinant mononucleosomes. (B) Methylation of the same protein amounts as in (A) by SUV420H1. (C) Methylation of native nucleosomes extracted from SUV420H1/H2 dko MEF cells. (D) Quantification of three independent replicates of (C). The image was taken from (Bröhm et al., 2019).

3.2.4 Catalytic activity of SUV420H2 mutants in cells

To investigate the catalytic activity of the SUV420H1 mutants in cells, stable cell lines were generated from MEF dko cells by viral transduction. After several days of induction with doxycycline, the cells were collected and the histones extracted. The amount of trimethylated H4K20 was then evaluated by western blot with a specific antibody as shown in Figure 24. Despite an overall very low expression level of the virally transduced proteins (see Appendix 2), a substantial increase in H4K20me3 levels could be detected in comparison to untreated dko cells and even comparable to wild type cells. In comparison with the wild type SUV420H1, the E238V, D249N, and S255F mutants showed a moderate to strong reduction of H4K20me3 and a similar expression level. The K258E, A269V, and S283L mutants showed greatly reduced H4K20me3 with about 60 % expression compared to the wild type. The S304Y and E320K mutants also displayed moderately decreased H4K20me3 levels, however at the same time showed a similarly reduced expression. In summary, the mutants which displayed the lowest H4K20me3 levels in cells (S255F, K258E, A269V, and S283L) are the same which already showed a severe loss of activity on nucleosome substrates.

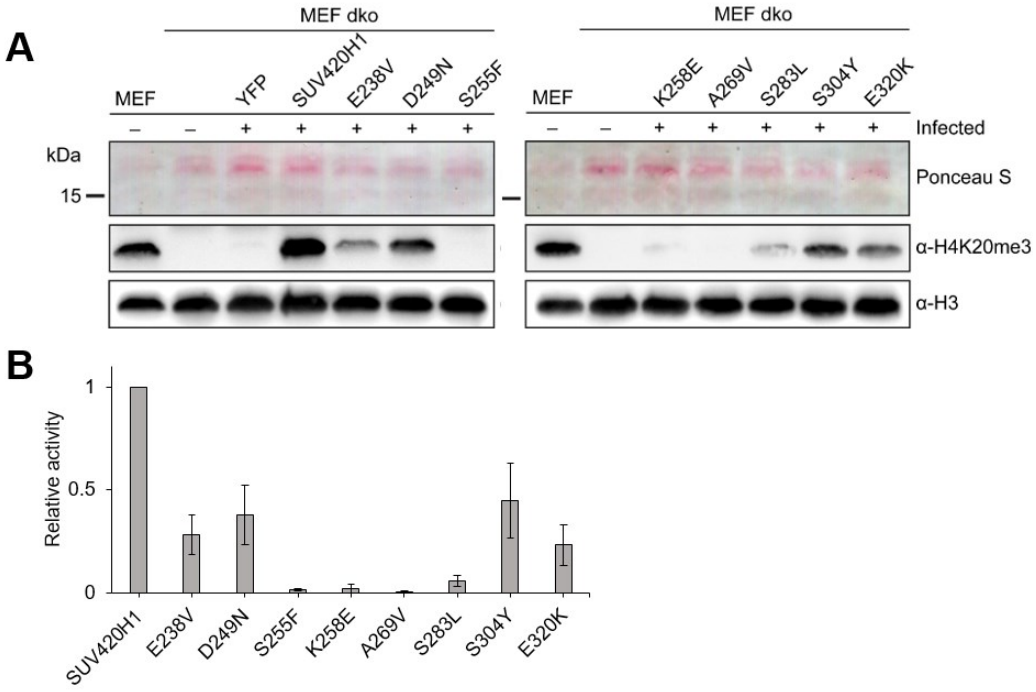


Figure 24: Detection of histone methylation by SUV420H1 mutants in MEF dko cells. (A) Western blots against H4K20me3 for each SUV420H1 variant. (B) Quantification of three independent replicates of (A). The image was taken from (Bröhm et al., 2019).

3.3 Stimulation of NSD1 H3K36 methylation by the H3.3 G34W oncohistone mutation

In this study, an emerging factor in epigenomic (dys-)regulation was investigated, which is oncomutations of histone variants. Specifically, the G34W mutation of the histone H3.3 variant was identified as a unique feature of more than 90 % of patients affected by the giant cell tumor of the bone (GCTB) (Amary et al., 2017; Lüke et al., 2017). This mutation has an influence on the methylation levels of the adjacent H3K27 and H3K36 marks as well as other regulatory processes, leading to altered epigenome regulation and impaired differentiation (Shi et al., 2018; Khazaei et al., 2020; Lutsik et al., 2020). How the G34W mutation induces these changes is unknown and therefore subject of this study. The close proximity of the G34W mutation to the H3K36 site combined with the drastic nature of the exchange from glycine to the bulky tryptophan implies a direct effect on the function of H3K36 methyltransferases. Indeed, previous studies reported an inhibition of the SETD2 methyltransferase, however no alteration of NSD2 activity (Jain et al., 2020). Here, in addition to the aforementioned proteins, the effect of the G34W mutation on the catalytic activity of the NSD1 methyltransferase was investigated. These findings were published in the journal *Biochimie* (Appendix 3) and are only summarized here (Bröhm et al., 2022b).

To investigate the effect of the G34W mutation on the NSD1 protein, the catalytic domain of the murine NSD1 was purified and used for methylation experiments with peptides containing H3K36 (H3 27-43), which carried the H3.1, H3.3, or H3.3 G34W sequence. Here, it was observed that NSD1 displays only miniscule catalytic activity on the H3.1 peptide but robust activity on H3.3. Unexpectedly, the catalytic activity on H3.3 G34W peptide was increased by a factor of 2.3 compared to unmodified H3.3 (Appendix 3).

In order to confirm this initially intriguing finding with a more physiologically relevant substrate, the same methylation experiments were repeated using recombinant nucleosomes. To this end, the G34W mutation was introduced into the H3.3 expression construct and the resulting histone proteins were used to reconstitute recombinant nucleosomes containing either H3.1, H3.3, or H3.3 G34W. The process of nucleosome reconstitution and relevant control experiments are further detailed in chapter 5.6 and in Appendix 3. Initial methylation experiments with these nucleosomes showed a comparable result as previously observed with peptides, with little activity on H3.1 but increased activity on H3.3 G34W (Figure 25A). Due to the drastic differences in signal intensity, only a rough quantification of all three nucleosome variant could be performed (Figure 25B-C). Therefore, additional experiments were carried out using only H3.3 and H3.3 G34W nucleosomes and adjusted reaction times to gain comparable signal intensities. As

a result, a strong stimulation of NSD1 by G34W by a factor of 6.3 could be detected (Figure 25D-E).

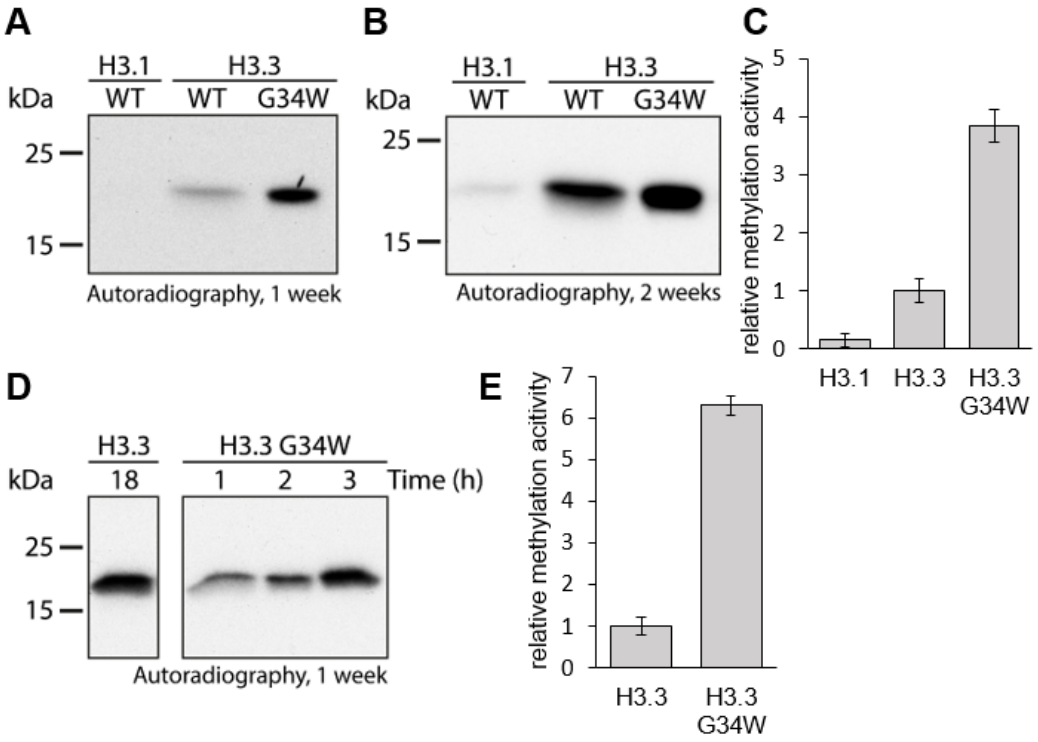


Figure 25: Nucleosome methylation experiments with NSD1. (A) Autoradiography of the methylation of nucleosomes containing H3.1, H3.3, or H3.3 G34W histones by NSD1. (B) Long exposure of (A). (C) Quantification of three independent replicates of (A). Error bars represent the SEM. (D) Autoradiography of the methylation of H3.3 and H3.3 G34W nucleosomes at different time points. (E) Quantification of three independent replicates of (D). Error bars represent the SEM. The image was taken from (Bröhm et al., 2022b).

To evaluate this result in context with previously reported data on SETD2 and NSD2, further experiments were conducted with these enzymes and the same recombinant nucleosomes. For SETD2, a moderate inhibition of catalytic activity by G34W was observed, which is in agreement with previous data (Appendix 3). The NSD2 protein, which was reported to be unaffected by G34W, showed a moderately increase methylation activity, which may be attributed to methodological differences, because, in the previous study, native nucleosomes isolated from cells were used while this study employed recombinant nucleosomes with defined composition (Appendix 3). In summary, the data presented here clearly documents a strong and specific stimulatory effect of the G34W mutation on NSD1 activity on nucleosome substrates (Figure 26).

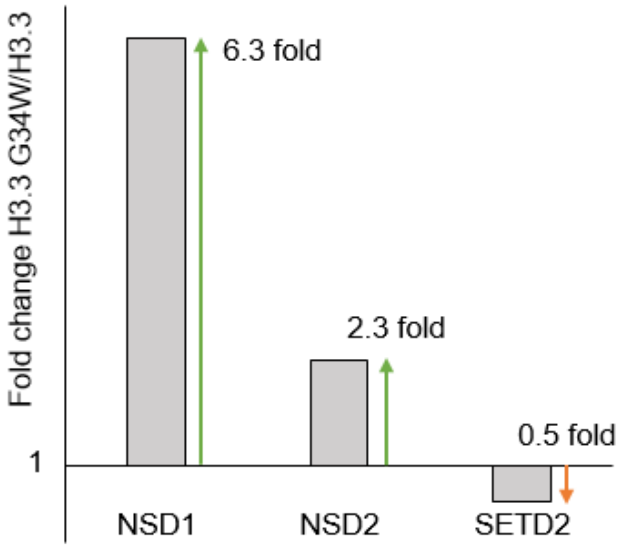


Figure 26: The effect of the G34W mutation on the catalytic activity of the H3K36 methyltransferases NSD1, NSD2, and SETD2. The image was taken from (Bröhm et al., 2022b).

4 Discussion

4.1 Linker DNA methylation of recombinant mononucleosomes by DNMT3A

The data presented in this study report a novel, comprehensive *in vitro* study of nucleosomal DNA methylation by DNMT3A. It was found that the nucleosomal DNA itself is largely protected against methylation by DNMT3A2, DNMT3AC, or DNMT3AC/3B3C. Instead, all tested enzymes primarily acted on the linker DNA, where the highest methylation activity could be detected. This behavior was expected for the present experimental setup and is in agreement with previous studies regarding DNMT3A and DNMT1 (Felle et al., 2011). The last CpG site in the nucleosome binding sequence displayed a notable exception insofar as it consistently showed increased methylation activity comparable to linker DNA CpGs. This can be explained by the inherent asymmetric binding mode of the Widom 601 DNA sequence to the histone octamer, which leads to transient unwrapping of the DNA and therefore to easier accessibility from the right DNA side (Ngo et al., 2015).

One central finding of the experiments with DNMT3A3 and DNMT3AC/3B3C is that the methylation levels of the linker DNA CpG sites at position 57 and 61 are in very good agreement with the cryo-EM structure of the DNMT3A/3B3-nucleosome complex (Xu et al., 2020). In this structure, the heterotetramer complex consisting of two inner DNMT3A subunits and two outer DNMT3B3 subunits is anchored to the nucleosome particle by an interaction of DNMT3B with the acidic patch on the disk face of the histone octamer. The whole complex is oriented directly towards the linker DNA near the dyad axis, which places the CpG site 57 in an ideal position to be methylated by the distal DNMT3A subunit. CpG site 61, however, is located in the interface region between the DNMT3A subunits and is therefore not expected to be methylated. Furthermore, the proximal DNMT3A subunit, while being oriented directly over the linker DNA, also is unlikely to methylate CpG sites located here, because it forms a bridge-like structure with a too large distance to the DNA. Hence, the experiments conducted with DNMT3A2 and DNMT3AC/3B3C in this study perfectly agree with the assumptions based on the cryo-EM structure, with high methylation observed on CpG 57 and low methylation on CpG site 61. Therefore, the cryo-EM structure can be interpreted as showing a catalytically competent complex conformation. Furthermore, the presumption that DNMT3A2 forms homotetramers similar in structure to DNMT3A/3B3 is vindicated since no significant differences were visible between DNMT3A2 and DNMT3AC/3B3C.

The linker DNA fragment in this study extended considerably further than the 10 bp sequence which was resolved in the cryo-EM structure. Notably, the strongest

methylation was observed on the more distal CpG sites, which can be explained by different mechanisms: Firstly, since the methylation preference at these sites correlates very well with the previously described flanking sequence preference of DNMT3A, it is possible that the outer linker DNA (most apart from the nucleosome) is bound and methylated like free DNA without contacts of DNMT3A to the nucleosome as described earlier (Zhang et al., 2018). Furthermore, it is possible that a second DNMT3 tetramer binds to the extended linker DNA as shown recently (Emperle et al., 2021). Lastly, is also seems possible that the distal CpG sites can be reached by the nucleosome-bound DNMT3 tetramer by bending and shifting of the linker DNA and conformational changes of the tetramer. The precise mechanism of distal linker DNA methylation will be the subject of further study. Potentially, Förster resonance energy transfer (FRET) experiments could be employed to monitor the DNMT3A-nucleosome distance while methylation takes place using substrates with different CpG locations. To this end, the FRET probes would be placed on the enzyme and the nucleosome. With increasing CpG distance from the nucleosome core, the behavior of the FRET signal would then hint about the DNMT3 binding dynamics: Strong methylation of distal CpGs and strong FRET signal would suggest a DNA bending mechanism while low FRET signal would imply binding and methylation of DNA independent of nucleosomal contacts.

Considering the nucleosome-proximal linker DNA methylation, further work is needed to evaluate the role of DNMT3A-nucleosome binding. For instance, the binding of the DNMT3A tetramer to the acidic patch on the histone octamer can be investigated by mutation of the responsible basic amino acids on DNMT3A, which should abolish the high methylation level of CpG 57. Similarly, the amino acids forming the acidic patch on the octamer could be mutated to disrupt the DNMT3A contact. Furthermore, the methylation of the well described DNMT3A/DNMT3L heterotetramer, which is known to display enhanced methylation activity compared to homomeric DNMT3A, can be tested using nucleosome substrates. Since the outer DNMT3L subunits lack the amino acid residues which confer the interaction with the octamer acidic patch in the cryo-EM structure, it would be expected that binding is lost and the CpG site 57 shows a methylation level comparable to free DNA. However, DNMT3L might either form a different contact to the nucleosome particle or enhance DNMT3A activity in such a way that nucleosomal binding is essentially unneeded due to increased catalytic activity.

Another important aspect of this work is the study of the influence of histone modifications on DNA methylation in a nucleosomal context. For the full-length DNMT3A to efficiently methylate DNA, its autoinhibitory conformation has to be relieved by binding of the ADD domain to the H3 tail, which is hindered by the

presence of H3K4me3 (Guo et al., 2015). As expected, the introduction of the H3K$_C$4me3 trimethyllysine analog led to a decreased methylation on the linker DNA in this study, however this was seen only for the nucleosome-proximal sites. This likely is the result of the aforementioned binding mode of the DNMT3A tetramer to the nucleosome: While the described structure allows for simultaneous binding of the H3 tail and the linker DNA, binding to the free linker DNA would most likely not include H3 tail binding, thereby alleviating the inhibitory effect of H3K4me3.

Additionally, the effect of H3K$_C$36me3 on the PWWP-mediated binding and methylation of the linker DNA by DNMT3A2 and DNMT3AC was studied. Unexpectedly, a stimulatory effect on linker DNA methylation was observed not only for the full length DNMT3A2 but also for DNMT3AC, which lacks the ADD and PWWP domain. Any effect observed with DNMT3AC is therefore independent of PWWP-H3K$_C$36me3 binding.

Previous studies using a wide range of methods ranging from NMR, fluorescence experiments, and molecular dynamics have demonstrated that the histone tails, albeit usually not resolved in crystal structures of the nucleosome, play a critical role in nucleosome stability and binding dynamics of external factors (Ikebe et al., 2016; Li and Kono, 2016; Gatchalian et al., 2017; Morrison et al., 2018, 2018; Lehmann et al., 2020; Ghoneim et al., 2021). Especially the H3 N-terminal tail was observed to engage in frequent transient interactions with the linker DNA including wrapping around the negatively charged DNA backbone. These interactions certainly limit the accessibility of the linker DNA and might, in addition to preventing DNA from winding off from the nucleosome, represent a control mechanism to limit binding of other factors. Thus, it stands to reason that trimethylation of H3K36, which lies at the base of the H3 N-terminal tail, has a detrimental effect on the interaction of the H3 tail with the DNA due to obstruction of hydrogen bond formation and sterical constraints. In fluorescence spectroscopy experiments, it was confirmed here that H3K36me3 results in a weakened interaction between H3 tail and linker DNA, which could explain the stimulation of DNA methylation by H3K36me3 observed with DNMT3AC. While DNMT3A2 could capture the H3 tail with its ADD and PWWP domains while transiently unbound, this mechanism is not available for DNMT3AC. Therefore, this enzyme is even more dependent on the H3K36me3-mediated increased availability of the linker DNA.

Based on the insights gained in this study, a new light is also shed on the evolutionary conserved interplay between H3K36me3 and gene body DNA methylation of actively transcribed genes. Deposition of H3K36me3 leads to repression of spurious transcription initiation by recruitment of HDACs and DNMTs (Wagner and Carpenter, 2012; Neri et al., 2017). DNMT-mediated CpG

methylation in gene bodies of expressed genes is conserved in eukaryotes (Feng et al., 2010; Zemach et al., 2010; Mendoza et al., 2019), but it is unknown why it has such a strong and widespread connection to H3K36me3. The data presented here suggest that H3K36me3 and DNA methylation are evolutionary connected by the inherent properties of the H3K36me3 modification which perhaps only later gained its signaling function. Specifically, the ability to weaken H3-tail binding to the DNA and thereby stimulate transcription may have been the initial function of H3K36me3. To fulfill this role, enzymes introducing H3K36me3 are delivered to gene bodies of active genes by the elongating RNA polymerase. This was later supported by silencing mechanisms to prevent erroneous transcription initiation mediated by binding proteins that recruit silencing factors to H3K36me3. At the same time, DNA methylation was enhanced by the better accessibility of the DNA leading to deposition of DNA methylation in H3K36me3 marked regions. Thereby, DNA methylation in gene bodies became connected to H3K36me3 and only later this liaison was further supported by specific recruitment of DNMTs to H3K36me3 by interaction with reading domains such as their PWWP domain. In agreement with this model, it is known that contrary to vertebrate and invertebrate DNMTs, the DRM family of plant DNA methyltransferases do not contain PWWP domains (Jeltsch, 2010; Bhattacharyya et al., 2020) but gene bodies in plants still are enriched by DNA and H3K36 methylation (Bewick and Schmitz, 2017; Pajoro et al., 2017; Liu et al., 2019). This suggests that the signaling connection between H3K36me3 and DNA methylation via the PWWP domain evolved after the separation of the plant and animal kingdoms.

The observation that the H3K36me3 modification enhances DNMT3 methylation activity by decreasing the affinity of the H3 tail to the linker DNA raises the question if this is a common mechanism of histone tail modifications or limited to specific cases. In the conformational studies of the linker DNA presented here, the effect of the H3K36 trimethylation on H3 tail binding was confirmed by fluorescence spectroscopy, utilizing an artificially introduced tryptophan residue at position 15. It can therefore be reasonably assumed that disruption of H3 tail binding at the base of the H3 tail, which is very close to the K36 site, weakens the binding of the whole histone tail to the linker DNA. It is however unknown at this point if a given site of perturbation leads only to decreased binding of the distal tail segment or the whole histone tail. This question is relevant for other common histone modifications on the H3 tail: Considering H3K27me3 for instance, one might assume a similar result of fluorescence spectroscopy experiments with tryptophan at position 15 if binding of the distal segment is impaired. For H3K9me3 however, it seems more plausible that H3 tail binding to the linker DNA is only weakly impaired due to the long nucleosome-proximal segment, which can interact normally with the linker DNA. In order to address these questions, fluorescence experiments with tryptophan fluorophores at different positions in the H3 tail have

to be carried out to precisely determine the effect of each histone tail modification on the conformational dynamics of the H3 tail. Furthermore, the H3 tail is not the only histone tail known to interact with the linker DNA. Due to its position near the DNA entry/exit sites, the H4 tail also forms transient interactions with nucleosomal and linker DNA, which raises the question if the H4K20me3 mark has a similar influence on these interactions as H3K36me3. Likewise, the G34W mutation, which was investigated in the context of NSD1-mediated H3K36 methylation, might influence the H3 tail dynamics in a similar way as H3K36me3 or even enhance the effect: First, the G34W exchange itself represents a drastic change of the sterical properties of the H3 tail, which might already affect the interactions with the linker DNA. Second, the observed stimulation of NSD1 activity (as described in chapter 3.3) leads to increased H3K36me2 levels, which in turn disrupts the H3 tail binding to linker DNA.

Finally, it has to be considered that the mononucleosomes employed in this study do not represent the physiological chromatin state. The question remains how DNMT3A acts in the context of oligo- or polynucleosomes, which are a better mimic of the chromatin environment in cells. The added complexity of internucleosomal interactions opens up a variety of possibilities: DNMT3A tetramers could bind to nucleosomes as shown by the cryo-EM structure while linker DNA fragments are methylated independently, or the linker DNA might be flexible enough to be accessed by nucleosome-bound DNMT3A. Furthermore, the known property of DNMT3A to self-associate beyond the tetramer structure into oligomers (Jurkowska et al., 2011) might allow the enzyme to bridge multiple DNA strands from neighboring nucleosomes. At the same time, one Tetramer or higher DNMT3A aggregate may interact with two neighboring nucleosomes. In order to study the behavior of DNMT3A on oligonucleosomes, suitable substrates for *in vitro* studies would have to be generated, which are optimized for NGS library generation. Combined with different linker lengths, this would allow the precise study of DNA methylation in a context which very closely mimics the cellular chromatin environment.

4.2 Cancer mutants of the SUV420H1 methyltransferase modulate its catalytic activity

In this study, the effect of eight amino acid exchanges in the SUV420H1 protein, which were identified as somatic mutations in cancer patients, was investigated using peptide and nucleosome substrates as well as in cells. It was found that all mutations affect the catalytic activity of SUV420H1 to a varying degree, indicating their potential role in tumorigenesis. The mutants E238V, D249N, S255F, K258E, A269V, S283L, S304Y, and E320K could be classified in three categories based on

their behavior on different substrates. Furthermore, the opposing effects of some mutants on the catalytic activity suggest that the investigated mutations play different roles in carcinogenesis depending on the tumor type.

Mutated SUV420H1 enzymes of the first group (S255F, K258E, A269V) displayed greatly reduced or even undetectable activity on both peptide and nucleosome substrates. Based on structural analysis, this effect could be explained by disrupted interactions with AdoMet (A269V) and impaired peptide recognition (K258E, S255F). For K258E, a reduced peptide binding was furthermore confirmed by peptide pulldown experiments. These group of mutations was reported to occur in breast carcinoma (S255F) and uterine endometrioid carcinoma (K258E, A269V). These tumors might be affected by the loss of H4K20me3 through the resulting perturbation of heterochromatin and possible aberrant upregulation of oncogenes. Such a mechanism would classify these mutations as epigenetic driver mutations, since the epigenomic dysregulation caused by the mutations directly promotes cancer progression.

The second group of mutations comprised the E238V, D249N, and E320K mutants, which showed greatly reduced activity on peptide substrates but partially recovered activity on nucleosome substrates. Structural analysis showed that these residues likely are involved in peptide (E238V, D249N) and cofactor binding (E320K). This suggests that although the affected amino acids are directly or indirectly relevant for substrate or cofactor binding, the disruption caused by the respective mutations can be rescued by additional contacts of the enzyme to nucleosomes. As seen in methylation experiments with H4 protein and nucleosomes, the overall activity of SUV420H1 is drastically increased with nucleosome substrates, indicating that the nucleosome contact is essential for proper function. The mutations of this group all occur in adenocarcinoma of the lung, rectum, and colon (see Appendix 2). Since the catalytic activity of these mutant enzymes is only moderately reduced compared to the wild type, it is likely that these mutations do not act as isolated driver mutations but are accompanied by other factors. A true driver function might be plausible in the case of a change in substrate specificity, which could lead to aberrant targeting of the mutant enzymes to tumor suppressor genes and subsequent silencing, but this had not been detected in any of these cases, at least not at peptide level.

The last group of mutations, S283L and S304Y, displayed a reversed effect compared to group 2: Here, the strongest activity was observed with peptide substrates and moderately reduced activity was detected on nucleosomes. For S283L, structural modeling suggested that this residue forms a nucleosome contact, which might affect the methylation efficiency. In case of S304Y, which seems to play a role in conformational dynamics of the enzyme, the mutation might affect activation of SUV420H1 upon nucleosome binding. This finding suggests that

SUV420H1 can adopt an autoinhibited state, which is allosterically resolved by nucleosome contacts. Such mechanisms are already known for other chromatin-modifying enzymes and might constitute a widespread regulatory instrument relevant to avoid undesired enzymatic activity at off-targets by a kind of double-filter control (Jeltsch and Jurkowska, 2016; Zucconi and Cole, 2017; Kim et al., 2019). The mutations of this group were reported in adenocarcinoma of the lung, colon, and rectum (see Appendix 2). The contribution of these mutations to carcinogenesis is unclear, but, similar to group 2 mutations, it is unlikely that they act as isolated driver mutations. The proposed influence of the S304Y mutation on the enzyme activation mechanism however raises the question if the substrate preference is also affected, which might result in aberrant methylation and silencing of tumor suppressors. To answer this question, more structural data of the enzyme beyond the available SET domain crystal structures is needed to identify the parts of the enzyme that are involved in its regulation.

Taken together, the data presented here highlight the importance of nucleosomal contacts of the SUV420H1 enzyme and suggest a role of some residues in allosteric activation. At the time of conduction of this research, the chemical modification of expressed histone proteins to generate methyllysine analogs was not yet established. The opportunity to employ native nucleosomes from SUV420H1/H2 double knock-out cells was therefore critical, since these cells accumulate the H4K20me1 mark, which is the desired substrate of the SUV420H1 enzyme. Considering the advances made in the course of other studies, any further similar experiments would certainly be carried out using recombinant nucleosomes specifically carrying the $H4K_C20me1$ analog, provided the analog is accepted by SUV420H1 as substrate.

Furthermore, the compelling finding that almost all mutations studied here led to a reduction of catalytic activity implies a suppressive role of SUV420H1 in tumorigenesis. This interpretation is supported by the known relevance of SUV420H1 for DNA repair (Botuyan et al., 2006a; Hsiao and Mizzen, 2013; Tuzon et al., 2014; Bromberg et al., 2017). Loss of this function might promote the accumulation of other mutations, which is a characteristic feature of many cancer types. Lastly, since H4K20me3 is a key heterochromatic modifications, its loss is detrimental to heterochomatin stability and therefore results in genomic instability, which is a hallmark of nearly all tumor cells. H4K20me3 loss is associated with poor prognosis in breast cancer, which highlights the importance of proper SUV420H1 function (Yokoyama et al., 2014).

In order to further study the role of SUV420H1 and the present mutations in carcinogenesis, experiments with cell lines derived from patient samples could be considered. CHIP-sequencing could reveal the precise regions where H4K20 methylation is dysregulated and which other changes of histone marks are co-occurring. This could reveal other factors involved in disease progression.

Furthermore, comparison with knockout cells could show if the effect on cancer is simply due to reduced H4K20 methylation or due to aberrant targeting of the mutated enzyme to other regions such as tumor suppressors. Since cancers typically accumulate many passenger mutations due to genomic instability which are not directly involved in tumor progression, it is possible that some of the investigated mutations fall into this category. If a mutation was successfully be identified as a driver mutation with aberrant methylation activity, previous studies which investigated inhibitors of the SUV420 enzymes might be used as a basis to develop a selective inhibitor for the mutant (Bromberg et al., 2017).

4.3 Stimulation of NSD1 H3K36 methylation by the H3.3 G34W oncohistone mutation

In this study, a strong stimulatory effect of the histone H3.3 G34W mutation on the catalytic activity of the NSD1 H3K36 methyltransferase was observed. In addition, the previous observation of inhibition of the SETD2 enzyme was reproduced (Jain et al., 2020). With the NSD2 enzyme however, a slightly increased activity was observed which stands in contrast to previously reported data indicating that this enzyme is unaffected by G34W (Jain et al., 2020). This observation might be explained by methodological differences: While the aforementioned study used native nucleosomes extracted from cells, here recombinant nucleosomes with a defined composition were employed. The affinity purification employed to isolate nucleosomes from cells could for instance also yield asymmetric nucleosomes containing both unmodified H3.3 and H3.3 G34W, or the histones could already carry other modifications which affect methylation experiments. The increased activity of NSD1 on G34W-containing substrates is in agreement with previous data, which showed increased activity of NSD1 on peptides containing G34F or G34Y, thereby indicating a preference for bulky aromatic residues (Kudithipudi et al., 2014b). However, due to the lower yield of tryptophan in peptide synthesis as compared with other amino acids, for the tryptophan-containing spots no accurate quantification of the NSD1 activity on these peptides could be achieved in the previous study. The furthermore detected preference of histone H3.3 over H3.1 is most likely attributed to the A31S exchange in the H3.3 protein.

The here discovered stimulation of NSD1 activity by the G34W mutation has widespread potential implications for the previously observed epigenomic alterations caused by this mutation. Specifically, the simultaneous stimulation of the dimethyltransferases NSD1 and NSD2 and inhibition of the trimethyltransferase SETD2 should lead to an increase in H3.3 K36me2 in cells. This in turn leads to an increased interaction with the DNMT3A-PWWP domain

and might result in a more stable anchoring of DNMT3A to H3.3 K36me2 genomic regions (Weinberg et al., 2019). Ultimately, this would mean that less free DNMT3A is available for genomic DNA methylation, which might provide an explanation for the global 20 % reduction of DNA methylation observed in the presence of G34W (Lutsik et al., 2020). This hypothesis is supported by an orthogonal study where loss of H3K36me3 binding due to a mutated DNMT3A-PWWP domain led to a global increase of DNA methylation (Sendzikaite et al., 2019). Another mechanism potentially involved is the H2AK119ub1-mediated targeting of DNMT3A to polycomb domains (Weinberg et al., 2021), which could be disrupted by increased H3K36me2-mediated PWWP targeting, leading to loss of methylation at the regions containing H3K27me3 and H2AK119ub. Moreover, an antagonistic interplay between H3K36me2 and EZH2-mediated H3K27me3 is known, which suggests that G34W could result in reduced H3K27me3 levels (Streubel et al., 2018; Finogenova et al., 2020). Since H3K27me3 plays important roles in gene regulation and differentiation, this might provide an explanation for the impaired differentiation in a G34W context. (Weinberg et al., 2021)

Considering the growing number of oncohistone mutations uncovered in the recent years (Nacev et al., 2019), the G34W mutation displays a remarkable behavior. Two of the most prominent histone mutations which have been described as oncohistones are the H3K27M and H3K36M mutations (Sarthy et al., 2020; Rajagopalan et al., 2021). These mutations have drastic consequences due to a combination of several factors: First, the common feature of these mutations is that they directly affect a modified histone mark. Second, the methionine residue is not repelled but instead strongly bound by the catalytic SET domain of the respective methyltransferases EZH2 and NSD1/2, as well as SETD2. For SETD2 specifically, structural data has confirmed that the K36M mutation leads to a strong anchoring of the whole enzyme complex to nucleosomes (Liu et al., 2021). As a result, although only a fraction of the available H3 protein in the cell carries the respective mutation, its prominent inhibitory effect on specific histone methyltransferases leads to genome-wide loss of the affected histone marks, namely H3K27me3 and H3K36me2/3. The global epigenomic dysregulation arising from this event is a driving factor in tumorigenesis (Mohammad and Helin, 2017). Despite being a very recent field of research, efforts during the last decade have led to the identification of many more oncogenic histone mutation not just limited to H3 (Nacev et al., 2019; Amatori et al., 2021; Flaus et al., 2021). The molecular mechanisms underlying the oncogenic effect of these mutations beyond the prominent inhibitory effects of the K27M and K36M mutations are just beginning to be unraveled. In this context, the G34W mutation is remarkable for several reasons: It mainly occurs in H3.3, which is incorporated into chromatin in a replication-independent manner, and furthermore, in contrast to H3.1, originates only from two genes, which means that a large fraction of the available protein is affected. Most importantly however,

the strong stimulatory effect of the G34W mutation on the NSD1 H3K36 methyltransferases stands in contrast to expectations and to the mechanisms underlying other oncohistone mutations. How exactly the epigenomic perturbations caused by increased H3K36me2 levels is connected to the cancers which exhibit the G34W mutations remains an interesting question for further study.

4.4 Recombinant nucleosomes as a universal *in vitro* study tool

The data presented here demonstrate the value of recombinant nucleosomes as substrates for the investigation of histone and DNA methylation *in vitro*. As seen in several cases, experiments carried out on isolated peptide or DNA fragments often do not capture the full extent of relevant interactions between modifying enzymes and their substrates. With the emergence of structure elucidation by cryo-EM, a rapidly growing number of enzymes in complex with nucleosome particles have been described (Xu et al., 2020; Li et al., 2021; Sato et al., 2021).

In the case of DNMT3A, for which so far only crystal structures of the tetrameric complex bound to DNA existed, the discovery of a direct anchoring to the nucleosome particle was unexpected. The implications of the described nucleosomal contacts are diverse: First, the orientation of the tetramer complex in the direction of the linker DNA helps to understand how the linker DNA is targeted with the aid of the H3-tail interactions of the ADD and PWWP domains. Second, the interaction of the outermost DNMT subunits with the acidic patch on top of the histone octamer highlights the importance of this conserved interaction site.

Considering the recently published structures of the NSD2 and NSD3 enzymes, it can be seen that these enzymes do not only bind on the surface of the nucleosome or to their H3 tail substrate, but instead displace the DNA in a way that up to 20 bp are unwrapped from the histone octamer. Based on the high amino acid sequence similarity of NSD1 to NSD2 and NSD3 and the fact that all three enzymes methylate the same K36 lysine residues, it is very likely that NSD1 employs a very similar mechanism. The resulting interaction network with the histones and DNA on one hand provides an excellent model for the H3K36 specificity of these enzymes. On the other hand, it immediately becomes clear that no model system other than nucleosomes can recapitulate the complex way the NSD enzymes interact with their substrate. It can be expected that for many other chromatin-modifying enzymes where only structures in complex with short peptides are available so far, future elucidation of nucleosome-bound conformations will unveil their mode of action in a previously unparalleled way.

The recently developed understanding of histone tail-DNA interactions add another layer of complexity to the interactions of external factors with nucleosomes. It was shown here that histone modifications can disrupt the H3 tail binding to the

nucleosomal and linker DNA, making it more accessible. This confirms that nucleosomal interactions are highly dynamic and dependent on multiple factors. Since for many proteins an allosteric regulation of activity is emerging, the entire spectrum of substrate interactions is needed to fully understand their function, which can only be achieved with nucleosomes.

A central feature of the recombinant nucleosome substrates used in the studies presented here is the versatility achieved by the incorporation of methyllysine analogs. During the investigation of nucleosomal DNA methylation by DNMT3A, the H3K$_C$4me3 and H3K$_C$36me3 analogs were used. These trimethyllysine analogs are chemically identical to lysine except that one carbon atom in the side chain is replaced by sulfur. While it is generally possible that the properties of this residue are thereby changed under some conditions, in the overwhelming majority of scenarios it likely represents a very good mimic of the "true" trimethyllysine, since most contacts with this residue are formed with the terminal amine. For H3K$_C$36me3, binding of peptides containing this analog to certain PWWP domains was already reported previously (Li et al., 2013; Mauser et al., 2017; Wang et al., 2020). The chemical alkylation of cysteine used here can furthermore be adjusted to yield other modified lysine analogs, such as mono- or dimethylated states. Other modifications such as lysine acetylation are also accessible by employing a UV-initiated radical mechanism (Li et al., 2011). However, the scope of accessible modifications is still relatively limited. Additional considerations regarding chemical modification based on the nucleophilic cysteine side chain have to be taken into account: Naturally, the desired modification is incorporated into all present cysteine residues. In principle, this opens up the possibility of simultaneous dual modifications. For the modified recombinant H3 proteins used in the studies presented here, this circumstance entails further mutagenesis steps in order to change the two cysteine s C96 and C110 to serine, which might not be applicable in other experimental setups.

An alternative, even more versatile approach for the introduction of histone modifications into recombinant nucleosomes is native chemical ligation of a truncated recombinant histone to a synthetic modified peptide (Dawson Philip E. et al., 1994; Shogren-Knaak and Peterson, 2003). This technique has the advantage of a nearly unrestrained number of combinations regarding the introduced modifications, which is only limited by the available modified amino acids for peptide synthesis. However, several drawbacks are also inherent to this method: First, large scale synthesis of peptides carrying specific modifications is drastically more expensive than chemical modification of expressed proteins. Furthermore, efficient lab-scale peptide synthesis is usually limited to ~20 amino acids, therefore the introduction of H3K36 modifications would require linking of two peptides, further complicating the process.

The present studies employed only mononucleosomes with a linker DNA fragment. However, some molecular mechanisms, especially those affecting compaction or opening of chromatin, can only be captured in longer fragments containing at least several nucleosome particles. By linking of differently modified nucleosomes to a chain, even trans-effects between individual nucleosomes can be studied as previously reported (Wu et al., 2013b). Regarding DNA methylation, this approach is currently only constrained by the limited read length of the Illumina paired-end sequencing, which cannot cover more than one linker DNA segment between nucleosomes. Extension of the employed DNA fragment to include several nucleosomes also increases the potential significance of linker histones. In principle, the linker histone H1 can be bound to previously reconstituted nucleosomes, or might already be included in the reconstitution mixture. Regarding mononucleosome dynamics, the effect of H1 binding is presumably mostly limited to strengthened DNA binding. However, due to its position at the dyad axis, binding and activity of any enzymes accessing the linker DNA could also be impaired. On longer nucleosomal arrays, the previously described functions of H1 in chromatin fiber compaction would come into effect, thereby potentially opening up novel ways to study this mechanism. The largely unstructured C- and N-terminal tails of H1 likely form similar contacts with the linker DNA as described for the H3 and H4 tails, since the amino acid composition especially of the N-terminal H1 tail is closely related to that of the H3 tail (Hao et al., 2021). How these interactions affect the binding dynamics of external factors such as PKMTs or DNMTs remains to be elucidated. If H1 tail binding to linker DNA is as pronounced as for H3, the association of other factors to linker DNA might be further prohibited. Since histone modifications like acetylation, methylation, and phosphorylation were also described on the H1 protein, it seems natural that the binding dynamics of its tail are fine-tuned by these modifications. Due to the position of the H1 protein on the dyad axis of the nucleosome, the distance to the core particle is relatively large, which suggests that the regulation of H1 binding to linker DNA is largely independent of core histone modifications and instead is modulated in parallel.

5 Materials and methods

5.1 Cloning, expression, and purification of recombinant proteins

The wild type plasmids encoding for the Suv420H1, NSD1, DNMT3AC, and DNMT3B3C proteins were obtained from coworkers. The constructs for the wild type histone proteins and the full length SUV420H1 were received from collaborators. In order to generate the Suv420H1 and SUV402H1 cancer mutants as well as the H3.1 and H3.3 variants, point mutations were introduced by PCR-based site-directed mutagenesis as described (Jeltsch and Lanio, 2002). An overview of the plasmids used in this work is given in Table 1. For the overexpression of proteins, the plasmids were transformed into BL21-CodonPlus cells (Stratagene) which were grown at 37 °C to an OD_{600} of 0.6-0.8. The protein expression was induced by addition of IPTG to a concentration of 1 mM and continued at reduced temperature of 16°C-20°C overnight. The cells were harvested by centrifugation at 4500 rcf and stored at -20°C. To isolate the respective protein, cells were disrupted by sonication and the cell debris was separated from the lysate by centrifugation at 40000 rcf. The lysate was then passed over the respective affinity tag binding beads (Ni-NTA agarose for His6-tag, Qiagen; Glutathione agarose 4B for GST-tag, Machery Nagel) equilibrated in sonication buffer. The beads were washed with at least 30 column volumes of sonication buffer and the bound proteins were eluted with sonication buffer containing the eluting agent (300 mM Imidazole for His6-tag; 40 mM reduced glutathione for GST-tag). Fractions were collected, pooled according to concentration and dialysed against dialysis buffer in order to remove the eluting agent. The proteins were the either flash frozen in liquid N_2 and stored at -80°C or further dialysed against dialysis buffer containing 65 % glycerol and stored at -20°C.

Table 1: Overview of the plasmids used in this thesis with information about domain boundaries, used vectors, and protein tags used for affinity purification or for flow cytometry.

Sequence of interest	Boundaries	Vector backbone	Protein tag	Accession number
Suv420H1 WT	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 E238V	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 D249N	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 S255F	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 K258E	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 A269V	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 S283L	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 S304Y	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
Suv420H1 E320K	9-392 (Cat. domain)	pDEST15	GST	Q3U8K7
SUV420H1 WT	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 E238V	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 D249N	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 S255F	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 K258E	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 A269V	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 S283L	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 S304Y	10-885	pSIN-TRE3G	EYFP	Q4FZB7
SUV420H1 E320K	10-885	pSIN-TRE3G	EYFP	Q4FZB7
H2A	1-130	pET21a	-	P0C0S8
H2B	1-126	pET21a	-	P62807
H3.1 WT	1-136	pET21a	-	P68431
H3.1 K4C	1-136	pET21a	-	P68431
H3.1 K36C	1-136	pET21a	-	P68431
H3.3 WT	1-136	pET21a	-	P84243
H3.3 G34W	1-136	pET21a	-	P84243
H4	1-103	pET21a	-	P62805
DNMT3AC	612-912 (Cat. domain)	pMAL	MBP	Q9Y6K1
DNMT3B3C	558-859 (Cat. domain)	pET28a	His6	O88509
Dnmt3a2 WT	221-908	pET28a	His6	O88508
NSD1	1700–1987	pGEX-6P2	GST	O88491

5.1.1 Copurification of DNMT3 heterotetramers

For the copurification of heterotetramers containing DNMT3AC and DNMT3B3C a double-tag affinity purification was conducted according to the scheme depicted in Figure 27. Briefly, both proteins were overexpressed independently as described in chapter 5.1, and the cell lysates were combined. The mixed lysate was then passed over amylose beads (New England Biolabs) and the bound protein was cleaved from the beads using TEV protease which was prepared in the lab. This was done to remove the bulky MBP-tag from the DNMT3AC subunits. The eluate was then passed over Ni-NTA beads from which the purified heterotetramers were eluted using sonication buffer containing 300 mM imidazole. Further details of the procedure are described in Appendix 1.

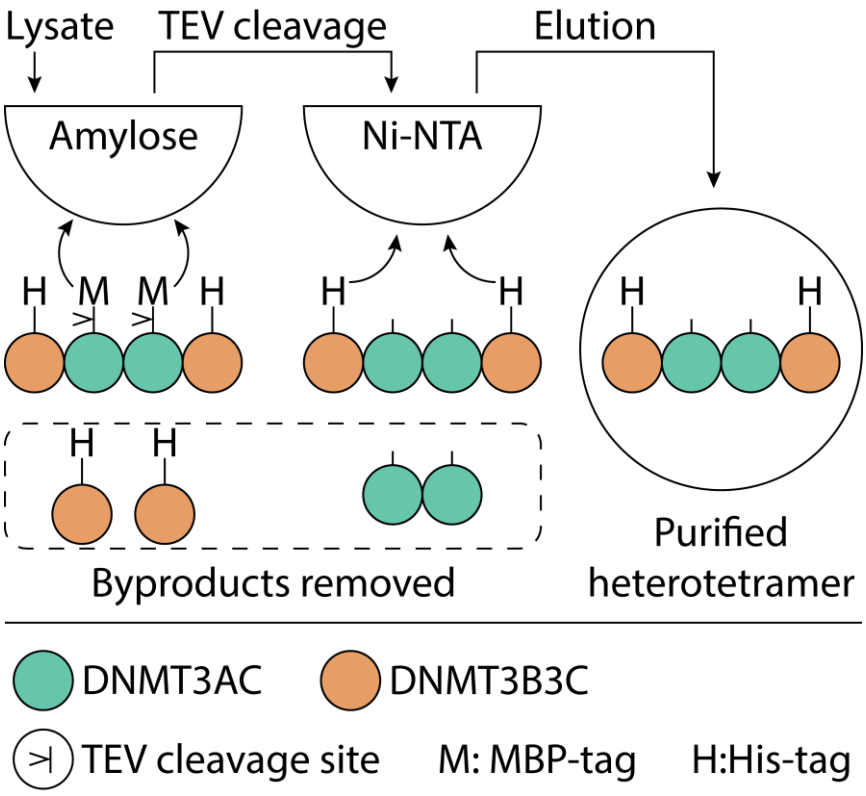


Figure 27: Scheme describing the double-tag purification process of heterotetramer complexes containing DNMT3AC and DNMT3B3C. Monomeric His6-tagged proteins are washed out in the first step and any proteins which don't form a complex with the His6-tagged proteins are washed out in the second step.

5.2 Radioactive methylation experiments with peptides, proteins and nucleosomes

For *in vitro* methylation experiments, enzymes were incubated in suitable buffer together with the substrate in presence of radioactively (^{3}H)-labelled AdoMet (Perkin Elmer) for 3 h at temperatures ranging from 20 °C to 37 °C. The reactions were then stopped by addition of SDS gel loading buffer and incubation at 95 °C for 10 min. Tris- or tricine-based SDS polyacrylamide electrophoresis was used to resolve the desired bands and the gels were dried under vacuum. The dried gels were then imaged using photosensitive films at -80°C in the dark. Details regarding enzyme and substrate concentrations are described in Appendix 2 and Appendix 3.

5.3 Circular dichroism spectroscopy

The folding of the purified proteins was confirmed by circular dichroism spectroscopy using a J-815 spectrophotometer (Jasco) at 20 °C. The spectra were collected at a wavelength range between 190 nm and 250 nm using a 0.1 mm cuvette and a scanning speed of 100 nm/min. Further details are described in Appendix 2.

5.4 Cell culture experiments

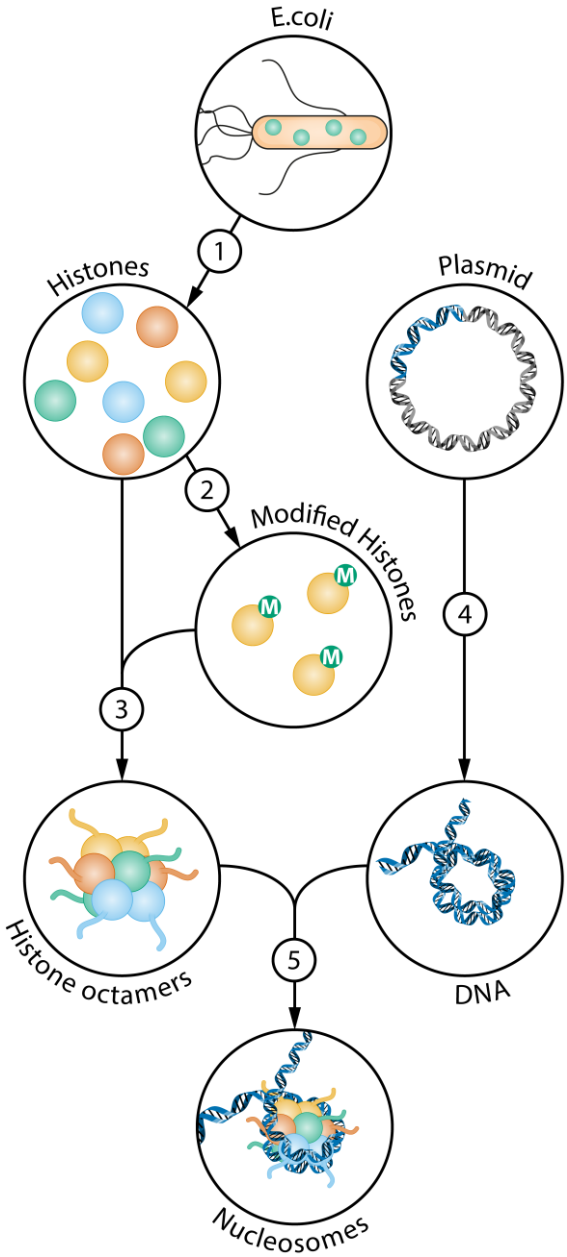
Wild type and SUV420H1/SUV420H2 double knockout mouse embryonic fibroblast (MEF) cells were cultured in DMEM media supplemented with 10% fetal bovine serum. 2 mM L-Glutamine, 1X penicillin/streptomycin, 1 mM sodium pyruvate, 1X non-essential amino acid solution, and 0.1 mM β-mercaptoethanol. Stable cell lines containing a SUV420H1 variant were generated by viral transduction of the respective construct. This was done by transfection of PlatE packaging cells with the construct by calcium phosphate co-precipitation and subsequent infection of the MEF cells with the collected retrovirus particles. Afterwards, cells were selected using puromycin treatment over 5 days and the introduced gene was induced by addition of doxycycline for 4 days. The induction efficiency was evaluated by flow cytometry (MACSQuant VYB, Miltenyi Biotec). Further details are described in Appendix 2.

5.5 Isolation of native mononucleosomes

Native mononucleosomes were isolated from mouse embryonic fibroblast (MEF) cell basically as described (Kasinathan et al., 2014). Briefly, MEF cells were lysed with NP-40 buffer (10 mM Tris–HCl pH 7.4, 2 mM MgCl2, 500 µM PMSF, 0.6% NP-40), and the extracted nuclei were digested with micrococcal nuclease (MNase) for 5 min in the presence of 1 mM CaCl2. The digestion was stopped by addition of 2 mM EGTA, and the mononucleosomes were isolated from the nuclei by treatment with 0.1% Triton X-100 and 300 mM NaCl. Sufficient chromatin

digestion was confirmed by gel electrophoresis. Finally, the nucleosomes were flash frozen in liquid N_2 and stored at −80 °C.

5.6 Reconstitution of recombinant nucleosomes



The process of nucleosome reconstitution was based on previously published protocols (Luger et al., 1999; Klinker et al., 2014) with some adaptations. An overview of the procedure is given in Figure 28. First, the individual histone proteins were overexpressed and purified by cation exchange chromatography. Then, some of the histone proteins were used to generate site-specific trimethyllysine analogs by chemical alkylation. All four individual histones were then refolded into stable octamers and isolated by size exclusion chromatography. The DNA fragment needed for the final nucleosome assembly was amplified by PCR and the reconstitution was conducted by salt gradient dialysis. Further details are described in the following in Appendix 1 and Appendix 3.

Figure 28: Scheme describing the workflow of recombinant nucleosome assembly, consisting of histone purification and modification, octamer refolding, DNA amplification and the final nucleosome reconstitution.

5.6.1 Cloning, expression, and purification of Histone proteins

The wild type full length histone constructs of H3.1 and H3.3 were first modified by site-specific mutagenesis to include the desired point mutations for the following installation of trimethyllysine analogs and for other studies. These mutations were H3.1 K4C and K36C as well as H3.3 G34W. In all H3.1 constructs, additionally the C96S and C110S exchanges were introduced to prevent incorporation of the trimethyllysine analog at these sites while still keeping a chemically and topologically similar amino acid residue. Likewise, in all H3.3 constructs the C110S exchange was introduced.

The individual histone proteins were overexpressed in a manner similar to other proteins as described in chapter 5.1. For the purification by cation exchange chromatography, a previously described protocol was used with some adaptations (Klinker et al., 2014). First, the cells were resuspended in denaturing sodium acetate-urea (SAU) buffer (NaOAc/HCl pH 7.5, 1 mM EDTA, 5 mM β-mercaptoethanol, 200 mM NaCl, 6 M urea) and disrupted by sonication. The lysate was cleared by centrifugation at 40000 rcf and passed through a 0.45 µM syringe filter. Next, the lysate was loaded on a 5 ml SP HP strong cation exchange column (GE Healthcare) using an NGC FPLC system (BioRad). Unbound proteins were washed out with several column volumes of SAU buffer and the bound histones were eluted using a salt gradient form 200 mM to 800 mM in SAU buffer. Fractions were collected, pooled according to purity, and dialysed against two charges of distilled water to remove all buffer components. Finally, the solution was dried overnight in a vacuum centrifuge and the dried proteins were stored at 4 °C.

5.6.2 Installation of trimethyllysine analogs

The chemical modification of the previously introduced cysteine residues by alkylation was carried out essentially as described (Simon et al., 2007). Briefly, the proteins were dissolved in alkylation buffer (1 M HEPES pH 7.8, 4 M guanidinium chloride, 10 mM methionine) and incubated for 1 h at 37 °C under reducing conditions (20 mM DTT). Subsequently the alkylation was started by addition of 100 mg/ml 2-bromoethyltrimethylammoniumbromide and the reaction mixture was incubated a further 2 h at 37 °C. Afterwards, the reaction was quenched by addition of 50 µl/ml β-mercaptoethanol and the sample was again dialysed against water, dried under vacuum and stored at 4 °C. Successful conversion was confirmed by MALDI-TOF spectrometry.

5.6.3 Refolding of histone octamers

The refolding of individual histone proteins into octamer complexes was conducted as described earlier (Luger et al., 1999). First, all histones were dissolved separately

in unfolding buffer (20 mM Tris/HCl pH 7.5, 7 M guanidinium chloride, 5 mM DTT) and their concentrations were determined spectrophotometrically. The proteins were then combined in a molar ration of 1 (H3, H4) to 1.2 (H2A, H2B) and dialysed against two charges of refolding buffer overnight (10 mM Tris/HCl pH 7.5, 1 mM EDTA, 2 M NaCl, 5 mM β-mercaptoethanol). To isolate the octamers from partially assembled histone complexes the sample was loaded on a Superdex 200 16/600 PG size exclusion column (GE Healthcare) and eluted isocratically using refolding buffer. Fractions were collected, pooled according to purity and concentrated tenfold using spin filters with a cutoff of 30 kDa (Merck Millipore). The samples were then flash frozen in liquid N_2 and stored at -80 °C.

5.6.4 Reconstitution of mononucleosomes

To generate a suitable DNA fragment for nucleosome binding, the Widom 601 sequence (Lowary and Widom, 1998) was cloned into a TOPO-TA vector together with a linker DNA sequence containing several CpG sites. This sequence was amplified in great quantities by PCR using barcoded primers for later discrimination of nucleosome variants. The DNA and histone octamer were combined in different molar ratios, ranging from equimolar to twofold excess of octamer. The samples were the transferred to Slider-a-Lyzer microdialysis devices (ThermoFisher) and dialysed against high salt buffer (10 mM Tris/HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 1 mM DTT) which was continuously replaced by 5 volumes of low salt buffer (same composition but 250 mM NaCl) over the course of 24 h. Subsequently the samples were dialysed overnight against storage buffer (10 mM Tris/HCl pH 7.5, 1 mM EDTA, 1 mM DTT, 20 % glycerol), flash frozen in liquid N_2, and stored at -80 °C. Successful reconstitution was confirmed by EMSA (electrophoretic mobility shift assay, see Appendix 1- Appendix 3).

5.7 Nucleosome DNA methylation and Library preparation

For the methylation of nucleosomes by DNA methyltransferases, first the residual free DNA in the sample was digested with MluI (NEB) for 60 min at 37 °C in NEB Cutsmart buffer. Afterwards, the respective DNMT was added to the mixture in NEB Cutsmart buffer containing 10 mM EDTA and 25 µM AdoMet (PerkinElmer) and the reaction was allowed to proceed for 2 h at 37 °C. The reaction was stopped and all DNA-bound proteins were digested by addition of proteinase K and further incubation at 37 °C for 60 min. The resulting unbound DNA was purified from the reaction mixture using the Nucleospin Gel and PCR cleanup kit (Macherey-Nagel). Subsequent bisulfite conversion of the methylated DNA was performed using the EZ DNA Methylation-Lightning kit (Zymo Research). Methylation of free DNA was conducted the same way using 15 µM DNA. For competitive nucleosome

methylation experiments, each nucleosome variant was independently treated with DpnI and subsequently combined for the methylation reaction.

Sample-specific barcodes and indices were added to the bisulfite-converted DNA by in a two-step PCR process. In the first PCR, the DNA was amplified using barcoded primers and the HotStartTaq Polymerase (Qiagen), and the resulting 321 bp fragment was purified using the Nucleospin Gel and PCR cleanup kit (Macherey-Nagel). In the second PCR step, adaptors and indices required for sequencing were added by amplification with the respective primers and the Phusion polymerase (ThermoFisher). The final 390-bp product was purified and used for Illumina paired end 2x250 bp sequencing. Datasets were analyzed using a local instance of the Galaxy bioinformatics server (Afgan et al., 2018). Sequence reads were trimmed with the Trim Galore! Tool (developed by Felix Krueger at the Babraham Institute) and subsequently paired using PEAR (Zhang et al., 2014). The reads were filtered according to the expected DNA length using the Filter FASTQ tool and mapped to the corresponding reference sequence using bwameth to determine the percentage of methylated CpGs (Blankenberg et al., 2010; Pedersen et al., 2014). All statistical analyses were done in Microsoft Excel 2016.

6 References

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., and Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic acids research *46*, W537-W544. https://doi.org/10.1093/nar/gky379.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proceedings of the National Academy of Sciences of the United States of America *51*, 786-794. https://doi.org/10.1073/pnas.51.5.786.

Alvarez-Venegas, R., and Avramova, Z. (2002). SET-domain proteins of the Su(var)3-9, E(z) and Trithorax families. Gene *285*, 25-37. https://doi.org/10.1016/S0378-1119(02)00401-8.

Amary, F., Berisha, F., Ye, H., Gupta, M., Gutteridge, A., Baumhoer, D., Gibbons, R., Tirabosco, R., O'Donnell, P., and Flanagan, A.M. (2017). H3F3A (Histone 3.3) G34W Immunohistochemistry: A Reliable Marker Defining Benign and Malignant Giant Cell Tumor of Bone. Am J Surg Pathol *41*, 1059-1068. https://doi.org/10.1097/pas.0000000000000859.

Amatori, S., Tavolaro, S., Gambardella, S., and Fanelli, M. (2021). The dark side of histones: genomic organization and role of oncohistones in cancer. Clinical epigenetics *13*, 71. https://doi.org/10.1186/s13148-021-01057-x.

Arnaudo, A.M., and Garcia, B.A. (2013). Proteomic characterization of novel histone post-translational modifications. Epigenetics & chromatin *6*, 24. https://doi.org/10.1186/1756-8935-6-24.

Audia, J.E., and Campbell, R.M. (2016). Histone Modifications and Cancer. Cold Spring Harbor perspectives in biology *8*, a019521. https://doi.org/10.1101/cshperspect.a019521.

Balakrishnan, L., Gefroh, A., and Milavetz, B. (2010). Histone H4 lysine 20 mono- and tri-methylation define distinct biological processes in SV40 minichromosomes. Cell cycle (Georgetown, Tex.) *9*, 1320-1332. https://doi.org/10.4161/cc.9.7.11123.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395. https://doi.org/10.1038/cr.2011.22.

Bennett, R.L., Swaroop, A., Troche, C., and Licht, J.D. (2017). The Role of Nuclear Receptor-Binding SET Domain Family Histone Lysine

Methyltransferases in Cancer. Cold Spring Harbor perspectives in medicine *7*. https://doi.org/10.1101/cshperspect.a026708.

Berdasco, M., Ropero, S., Setien, F., Fraga, M.F., Lapunzina, P., Losson, R., Alaminos, M., Cheung, N.-K., Rahman, N., and Esteller, M. (2009). Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. Proceedings of the National Academy of Sciences *106*, 21830. https://doi.org/10.1073/pnas.0906831106.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes & Development *23*, 781-783. https://doi.org/10.1101/gad.1787609.

Bewick, A.J., and Schmitz, R.J. (2017). Gene body DNA methylation in plants. Current opinion in plant biology *36*, 103-110. https://doi.org/10.1016/j.pbi.2016.12.007.

Bhattacharyya, M., De, S., and Chakrabarti, S. (2020). Origin and Evolution of DNA methyltransferases (DNMT) along the tree of life: A multi-genome survey. bioRxiv, 2020.04.09.033167. https://doi.org/10.1101/2020.04.09.033167.

Bischoff, R., and Schlüter, H. (2012). Amino acids: chemistry, functionality and selected non-enzymatic post-translational modifications. Journal of proteomics *75*, 2275-2296. https://doi.org/10.1016/j.jprot.2012.01.041.

Blankenberg, D., Gordon, A., Kuster, G. von, Coraor, N., Taylor, J., and Nekrutenko, A. (2010). Manipulation of FASTQ data with Galaxy. Bioinformatics *26*, 1783-1785. https://doi.org/10.1093/bioinformatics/btq281.

Botuyan, M.V., Lee, J., Ward, I.M., Kim, J.-E., Thompson, J.R., Chen, J., and Mer, G. (2006a). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. Cell *127*, 1361-1373. https://doi.org/10.1016/j.cell.2006.10.043.

Botuyan, M.V., Lee, J., Ward, I.M., Kim, J.-E., Thompson, J.R., Chen, J., and Mer, G. (2006b). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. Cell *127*, 1361-1373. https://doi.org/10.1016/j.cell.2006.10.043.

Bröhm, A., Elsawy, H., Rathert, P., Kudithipudi, S., Schoch, T., Schuhmacher, M.K., Weirich, S., and Jeltsch, A. (2019). Somatic Cancer Mutations in the SUV420H1 Protein Lysine Methyltransferase Modulate Its Catalytic Activity. Journal of molecular biology *431*, 3068-3080. https://doi.org/10.1016/j.jmb.2019.06.021.

Bröhm, A., Schoch, T., Dukatz, M., Graf, N., Dorscht, F., Mantai, E., Adam, S., Bashtrykov, P., and Jeltsch, A. (2022a). Methylation of recombinant mononucleosomes by DNMT3A demonstrates efficient linker DNA methylation and a role of H3K36me3. Communications Biology *5*, 192. https://doi.org/10.1038/s42003-022-03119-z.

Bröhm, A., Schoch, T., Grünberger, D., Khella, M.S., Schuhmacher, M.K., Weirich, S., and Jeltsch, A. (2022b). The H3.3 G34W oncohistone mutation increases K36 methylation by the protein lysine methyltransferase NSD1. Biochimie 198, 86-91. https://doi.org/10.1016/j.biochi.2022.03.007.

Bromberg, K.D., Mitchell, T.R.H., Upadhyay, A.K., Jakob, C.G., Jhala, M.A., Comess, K.M., Lasko, L.M., Li, C., Tuzon, C.T., and Dai, Y., et al. (2017). The SUV4-20 inhibitor A-196 verifies a role for epigenetics in genomic integrity. Nature Chemical Biology *13*, 317-324. https://doi.org/10.1038/nchembio.2282.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., and Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discovery *2*, 401-404. https://doi.org/10.1158/2159-8290.CD-12-0095.

Chedin, F., Lieber, M.R., and Hsieh, C.L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proceedings of the National Academy of Sciences of the United States of America *99*, 16916-16921. https://doi.org/10.1073/pnas.262443999.

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. Molecular and cellular biology *24*, 9048-9058. https://doi.org/10.1128/MCB.24.20.9048-9058.2004.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Molecular and cellular biology *23*, 5594-5605. https://doi.org/10.1128/mcb.23.16.5594-5605.2003.

Cheng, X. (1995). Structure and function of DNA methyltransferases. Annual review of biophysics and biomolecular structure *24*, 293-318. https://doi.org/10.1146/annurev.bb.24.060195.001453.

Choufani, S., Cytrynbaum, C., Chung, B.H.Y., Turinsky, A.L., Grafodatskaya, D., Chen, Y.A., Cohen, A.S.A., Dupuis, L., Butcher, D.T., and Siu, M.T., et al. (2015). NSD1 mutations generate a genome-wide DNA methylation signature. Nature Communications *6*, 10207. https://doi.org/10.1038/ncomms10207.

Cotton, A.M., Price, E.M., Jones, M.J., Balaton, B.P., Kobor, M.S., and Brown, C.J. (2015). Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. Human molecular genetics *24*, 1528-1539. https://doi.org/10.1093/hmg/ddu564.

Couture, J.-F., Collazo, E., Brunzelle, J.S., and Trievel, R.C. (2005). Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes & Development *19*, 1455-1465. https://doi.org/10.1101/gad.1318405.

Cutter, A.R., and Hayes, J.J. (2015). A brief review of nucleosome structure. FEBS letters *589*, 2914-2922. https://doi.org/10.1016/j.febslet.2015.05.016.

Das, C., and Tyler, J.K. (2013). Histone exchange and histone modifications during transcription and aging. Biochimica et biophysica acta *1819*, 332-342. https://doi.org/10.1016/j.bbagrm.2011.08.001.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution††We dedicate this paper to the memory of Max Perutz who was particularly inspirational and supportive to T.J.R. in the early stages of this study. Journal of molecular biology *319*, 1097-1113. https://doi.org/10.1016/S0022-2836(02)00386-8.

Dawson Philip E., Muir Tom W., Clark-Lewis Ian, and Kent Stephen B. H. (1994). Synthesis of Proteins by Native Chemical Ligation. Science *266*, 776-779. https://doi.org/10.1126/science.7973629.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes & Development *25*, 1010-1022. https://doi.org/10.1101/gad.2037511.

Del Rizzo, P.A., and Trievel, R.C. (2014). Molecular basis for substrate recognition by lysine methyltransferases and demethylases. Biochimica et biophysica acta *1839*, 1404-1415. https://doi.org/10.1016/j.bbagrm.2014.06.008.

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. https://doi.org/10.1074/jbc.M109.089433.

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome biology *6*, 1-10.

Dodge, 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. The Journal of biological chemistry *280*, 17986-17991. https://doi.org/10.1074/jbc.M413246200.

Dónal O'Carroll, Harry Scherthan, Antoine H. F. M. Peters, Susanne Opravil, Andrew R. Haynes, Götz Laible, Stephen Rea, Manfred Schmid, Angelika Lebersorger, and Martin Jerratsch, et al. Isolation and Characterization ofSuv39h2, a Second Histone H3 Methyltransferase Gene That Displays Testis-Specific Expression.

Dorigo Benedetta, Schalch Thomas, Kulangara Alexandra, Duda Sylwia, Schroeder Rasmus R., and Richmond Timothy J. (2004). Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. Science *306*, 1571-1573. https://doi.org/10.1126/science.1103124.

Dukatz, M., Holzer, K., Choudalakis, M., Emperle, M., Lungu, C., Bashtrykov, P., and Jeltsch, A. (2019). H3K36me2/3 Binding and DNA Binding of the DNA Methyltransferase DNMT3A PWWP Domain Both Contribute to its Chromatin Interaction. Journal of molecular biology *431*, 5063-5074. https://doi.org/10.1016/j.jmb.2019.09.006.

Egger Gerda, Jeong Shinwu, Escobar Sonia G., Cortez Connie C., Li Tony W. H., Saito Yoshimasa, Yoo Christine B., Jones Peter A., and Liang Gangning (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 *103*, 14080-14085. https://doi.org/10.1073/pnas.0604602103.

Emperle, M., Bangalore, D.M., Adam, S., Kunert, S., Heil, H.S., Heinze, K.G., Bashtrykov, P., Tessmer, I., and Jeltsch, A. (2021). Structural and biochemical insight into the mechanism of dual CpG site binding and methylation by the DNMT3A DNA methyltransferase. Nucleic acids research *49*, 8294-8308. https://doi.org/10.1093/nar/gkab600.

Erdel, F., and Rippe, K. (2018). Formation of Chromatin Subcompartments by Phase Separation. Biophysical journal *114*, 2262-2270. https://doi.org/10.1016/j.bpj.2018.03.011.

Fang, J., Feng, Q., Ketel, C.S., Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P., Simon, J.A., and Zhang, Y. (2002). Purification and Functional Characterization of SET8, a Nucleosomal Histone H4-Lysine 20-Specific Methyltransferase. Current Biology *12*, 1086-1099. https://doi.org/10.1016/S0960-9822(02)00924-7.

Feinberg, A.P., Koldobskiy, M.A., and Göndör, A. (2016). Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nature reviews. Genetics *17*, 284-299. https://doi.org/10.1038/nrg.2016.13.

Felle, M., Hoffmeister, H., Rothammer, J., Fuchs, A., Exler, J.H., and Langst, G. (2011). Nucleosomes protect DNA from DNA methylation in vivo and in vitro. Nucleic acids research *39*, 6956-6969. https://doi.org/10.1093/nar/gkr263.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., and Halpern, M.E., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. Proceedings of the National Academy of Sciences *107*, 8689. https://doi.org/10.1073/pnas.1002720107.

Finogenova, K., Bonnet, J., Poepsel, S., Schäfer, I.B., Finkl, K., Schmid, K., Litz, C., Strauss, M., Benda, C., and Müller, J. (2020). Structural basis for PRC2 decoding of active histone methylation marks H3K36me2/3. eLife *9*, e61964. https://doi.org/10.7554/eLife.61964.

Flaus, A., Downs, J.A., and Owen-Hughes, T. (2021). Histone isoforms and the oncohistone code. Current opinion in genetics & development *67*, 61-66. https://doi.org/10.1016/j.gde.2020.11.003.

Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., and Ponting, L., et al. (2017). COSMIC: somatic cancer genetics at high-resolution. Nucleic acids research *45*, D777-D783. https://doi.org/10.1093/nar/gkw1121.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., and Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science signaling *6*, pl1. https://doi.org/10.1126/scisignal.2004088.

Gatchalian, J., Wang, X., Ikebe, J., Cox, K.L., Tencer, A.H., Zhang, Y., Burge, N.L., Di, L., Gibson, M.D., and Musselman, C.A., et al. (2017). Accessibility of the histone H3 tail in the nucleosome for binding of paired readers. Nature communications *8*, 1489. https://doi.org/10.1038/s41467-017-01598-x.

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. The Journal of biological chemistry *279*, 25447-25454. https://doi.org/10.1074/jbc.M312296200.

Gelato, K.A., and Fischle, W. (2008). Role of histone modifications in defining chromatin structure and function. Biological chemistry *389*, 353-363. https://doi.org/10.1515/BC.2008.048.

Ghoneim, M., Fuchs, H.A., and Musselman, C.A. (2021). Histone Tail Conformations: A Fuzzy Affair with DNA. Trends in biochemical sciences *46*, 564-578. https://doi.org/10.1016/j.tibs.2020.12.012.

Gibbons, R.J., Bachoo, S., Picketts, D.J., Aftimos, S., Asenbauer, B., Bergoffen, J., Berry, S.A., Dahl, N., Fryer, A., and Keppler, K., et al. (1997). Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. Nat Genet *17*, 146-148. https://doi.org/10.1038/ng1097-146.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell *179*, 470-484.e21. https://doi.org/10.1016/j.cell.2019.08.037.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nature reviews. Genetics *13*, 343-357. https://doi.org/10.1038/nrg3173.

Grigoryev, S.A., and Woodcock, C.L. (2012). Chromatin organization — The 30nm fiber. Experimental Cell Research *318*, 1448-1455. https://doi.org/10.1016/j.yexcr.2012.02.014.

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., and Li, G., et al. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature *517*, 640-644. https://doi.org/10.1038/nature13899.

Hao, F., Kale, S., Dimitrov, S., and Hayes, J.J. (2021). Unraveling linker histone interactions in nucleosomes. Current Opinion in Structural Biology *71*, 87-93. https://doi.org/10.1016/j.sbi.2021.06.001.

Holliday, R., and Pugh, J.E. (1975). DNA Modification Mechanisms and Gene Activity During Development: Developmental clocks may depend on the enzymic modification of specific bases in repeated DNA sequences. Science *187*, 226-232.

Hota, S.K., and Bruneau, B.G. (2016). ATP-dependent chromatin remodeling during mammalian development. Development (Cambridge, England) *143*, 2882-2897. https://doi.org/10.1242/dev.128892.

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, an d nucleosides by paper chromatography. J. biol. Chem *175*, 315-332.

Hsiao, K.-Y., and Mizzen, C.A. (2013). Histone H4 deacetylation facilitates 53BP1 DNA damage signaling and double-strand break repair. Journal of molecular cell biology *5*, 157-165. https://doi.org/10.1093/jmcb/mjs066.

Ikebe, J., Sakuraba, S., and Kono, H. (2016). H3 Histone Tail Conformation within the Nucleosome and the Impact of K14 Acetylation Studied Using Enhanced Sampling Simulation. PLoS computational biology *12*, e1004788. https://doi.org/10.1371/journal.pcbi.1004788.

Jain, S.U., Khazaei, S., Marchione, D.M., Lundgren, S.M., Wang, X., Weinberg, D.N., Deshmukh, S., Juretic, N., Lu, C., and Allis, C.D., et al. (2020). Histone H3.3 G34 mutations promote aberrant PRC2 activity and drive tumor progression. Proceedings of the National Academy of Sciences *117*, 27354-27364. https://doi.org/10.1073/pnas.2006076117.

Jansen An, and Verstrepen Kevin J. (2011). Nucleosome Positioning in Saccharomyces cerevisiae. Microbiology and Molecular Biology Reviews *75*, 301-320. https://doi.org/10.1128/MMBR.00046-10.

Jeltsch, A. (2010). Phylogeny of methylomes. Science *328*, 837-838. https://doi.org/10.1126/science.1190738.

Jeltsch, A., and Jurkowska, R.Z. (2014). New concepts in DNA methylation. Trends in biochemical sciences *39*, 310-318. https://doi.org/10.1016/j.tibs.2014.05.002.

Jeltsch, A., and Jurkowska, R.Z. (2016). Allosteric control of mammalian DNA methyltransferases - a new regulatory paradigm. Nucleic acids research *44*, 8556-8575. https://doi.org/10.1093/nar/gkw723.

Jeltsch, A., and Lanio, T. (2002). Site-directed mutagenesis by polymerase chain reaction. Methods in molecular biology *182*, 85-94. https://doi.org/10.1385/1-59259-194-9:085.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cellular and molecular life sciences : CMLS *54*, 80-93. https://doi.org/10.1007/s000180050127.

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. https://doi.org/10.1038/nature06146.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews. Genetics *13*, 484-492. https://doi.org/10.1038/nrg3230.

Jones, P.A., and Liang, G. (2009). Rethinking how DNA methylation patterns are maintained. Nature reviews. Genetics *10*, 805-811. https://doi.org/10.1038/nrg2651.

Jørgensen, S., Elvers, I., Trelle, M.B., Menzel, T., Eskildsen, M., Jensen, O.N., Helleday, T., Helin, K., and Sørensen, C.S. (2007). The histone methyltransferase SET8 is required for S-phase progression. The Journal of cell biology *179*, 1337-1345. https://doi.org/10.1083/jcb.200706150.

Jørgensen, S., Schotta, G., and Sørensen, C.S. (2013). Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. Nucleic acids research *41*, 2797-2806. https://doi.org/10.1093/nar/gkt012.

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. https://doi.org/10.1074/jbc.M111.254987.

Kasinathan, S., Orsi, G.A., Zentner, G.E., Ahmad, K., and Henikoff, S. (2014). High-resolution mapping of transcription factor binding sites on native chromatin. Nature methods *11*, 203-209. https://doi.org/10.1038/nmeth.2766.

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 research *22*, 2497-2506. https://doi.org/10.1101/gr.143008.112.

Khazaei, S., Jay, N. de, Deshmukh, S., Hendrikse, L.D., Jawhar, W., Chen, C.C.L., Mikael, L.G., Faury, D., Marchione, D.M., and Lanoix, J., et al. (2020). H3.3 G34W Promotes Growth and Impedes Differentiation of Osteoblast-Like Mesenchymal Progenitors in Giant Cell Tumor of Bone. Cancer Discovery *10*, 1968. https://doi.org/10.1158/2159-8290.cd-20-0461.

Kim, J.-A., Kwon, M., and Kim, J. (2019). Allosteric Regulation of Chromatin-Modifying Enzymes. Biochemistry *58*, 15-23. https://doi.org/10.1021/acs.biochem.8b00894.

Klinker, H., Haas, C., Harrer, N., Becker, P.B., and Mueller-Planitz, F. (2014). Rapid Purification of Recombinant Histones. PLoS ONE *9*. https://doi.org/10.1371/journal.pone.0104029.g001.

Kourmouli, N., Jeppesen, P., Mahadevhaiah, S., Burgoyne, P., Wu, R., Gilbert, D.M., Bongiorni, S., Prantera, G., Fanti, L., and Pimpinelli, S., et al. (2004).

Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. Journal of cell science *117*, 2491-2501. https://doi.org/10.1242/jcs.01238.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705. https://doi.org/10.1016/j.cell.2007.02.005.

Kudithipudi, S., and Jeltsch, A. (2014). Role of somatic cancer mutations in human protein lysine methyltransferases. Biochimica et biophysica acta *1846*, 366-379. https://doi.org/10.1016/j.bbcan.2014.08.002.

Kudithipudi, S., Lungu, C., Rathert, P., Happel, N., and Jeltsch, A. (2014a). Substrate specificity analysis and novel substrates of the protein lysine methyltransferase NSD1. Chemistry & Biology *21*, 226-237. https://doi.org/10.1016/j.chembiol.2013.10.016.

Kudithipudi, S., Lungu, C., Rathert, P., Happel, N., and Jeltsch, A. (2014b). Substrate Specificity Analysis and Novel Substrates of the Protein Lysine Methyltransferase NSD1. Chemistry & Biology *21*, 226-237. https://doi.org/10.1016/j.chembiol.2013.10.016.

Kurotaki, N., Imaizumi, K., Harada, N., Masuno, M., Kondoh, T., Nagai, T., Ohashi, H., Naritomi, K., Tsukahara, M., and Makita, Y. (2002). Haploinsufficiency of NSD1 causes Sotos syndrome. Nature genetics *30*, 365-366.

Lehmann, K., Felekyan, S., Kühnemuth, R., Dimura, M., Tóth, K., Seidel, C.A.M., and Langowski, J. (2020). Dynamics of the nucleosomal histone H3 N-terminal tail revealed by high precision single-molecule FRET. Nucleic acids research *48*, 1551-1571. https://doi.org/10.1093/nar/gkz1186.

Leroy, G., Dimaggio, P.A., Chan, E.Y., Zee, B.M., Blanco, M.A., Bryant, B., Flaniken, I.Z., Liu, S., Kang, Y., and Trojer, P., et al. (2013). A quantitative atlas of histone modification signatures from human cancer cells. Epigenetics & chromatin *6*, 20. https://doi.org/10.1186/1756-8935-6-20.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.

Li, F., Allahverdi, A., Yang, R., Lua, G.B.J., Zhang, X., Cao, Y., Korolev, N., Nordenskiöld, L., and Liu, C.-F. (2011). A direct method for site-specific protein acetylation. Angewandte Chemie (International ed. in English) *50*, 9611-9614. https://doi.org/10.1002/anie.201103754.

Li, F., Mao, G., Tong, D., Huang, J., Gu, L., Yang, W., and & Li, G.M. (2013). The Histone Mark H3K36me3 Regulates Human DNA Mismatch Repair through Its Interaction with MutSα. Cell *153*, 590-600. https://doi.org/10.1016/j.cell.2013.03.025.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Current opinion in genetics & development *21*, 175-186. https://doi.org/10.1016/j.gde.2011.01.022.

Li, J., Ahn, J.H., and Wang, G.G. (2019). Understanding histone H3 lysine 36 methylation and its deregulation in disease. Cellular and molecular life sciences : CMLS *76*, 2899-2916. https://doi.org/10.1007/s00018-019-03144-y.

Li, J., Duns, G., Westers, H., Sijmons, R., van den Berg, A., and Kok, K. (2016). SETD2: an epigenetic modifier with tumor suppressor functionality. Oncotarget *7*, 50719-50734. https://doi.org/10.18632/oncotarget.9368.

Li, W., Tian, W., Yuan, G., Deng, P., Sengupta, D., Cheng, Z., Cao, Y., Ren, J., Qin, Y., and Zhou, Y., et al. (2021). Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases. Nature *590*, 498-503. https://doi.org/10.1038/s41586-020-03069-8.

Li, Z., and Kono, H. (2016). Distinct Roles of Histone H3 and H2A Tails in Nucleosome Stability. Scientific reports *6*, 31437. https://doi.org/10.1038/srep31437.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., and Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322. https://doi.org/10.1038/nature08514.

Liu, B., Liu, Y., Wang, B., Luo, Q., Shi, J., Gan, J., Shen, W.-H., Yu, Y., and Dong, A. (2019). The transcription factor OsSUF4 interacts with SDG725 in promoting H3K36me3 establishment. Nature Communications *10*, 2999. https://doi.org/10.1038/s41467-019-10850-5.

Liu, Y., Zhang, Y., Xue, H., Cao, M., Bai, G., Mu, Z., Yao, Y., Sun, S., Fang, D., and Huang, J. (2021). Cryo-EM structure of SETD2/Set2 methyltransferase bound to a nucleosome containing oncohistone mutations. Cell discovery *7*, 32. https://doi.org/10.1038/s41421-021-00261-6.

Lowary, P.T., and Widom, J. (1998). New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-directed Nucleosome Positioning. Journal of molecular biology *276*, 19±42.

Lu, X., Simon, M.D., Chodaparambil, J.V., Hansen, J.C., Shokat, K.M., and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. Nature structural & molecular biology *15*, 1122-1124. https://doi.org/10.1038/nsmb.1489.

Lucio-Eterovic, A.K., Singh, M.M., Gardner, J.E., Veerappan, C.S., Rice, J.C., and Carpenter, P.B. (2010). Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. Proceedings of the National Academy of Sciences *107*, 16952. https://doi.org/10.1073/pnas.1002653107.

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nature reviews. Molecular cell biology *13*, 436-447. https://doi.org/10.1038/nrm3382.

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

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Methods in enzymology *304*.

Lüke, J., Baer, A. von, Schreiber, J., Lübbehüsen, C., Breining, T., Mellert, K., Marienfeld, R., Schultheiss, M., Möller, P., and Barth, T.F.E. (2017). H3F3A mutation in giant cell tumour of the bone is detected by immunohistochemistry using a monoclonal antibody against the G34W mutated site of the histone H3.3 variant. Histopathology *71*, 125-133. https://doi.org/10.1111/his.13190.

Lutsik, P., Baude, A., Mancarella, D., Öz, S., Kühn, A., Toth, R., Hey, J., Toprak, U.H., Lim, J., and Nguyen, V.H., et al. (2020). Globally altered epigenetic landscape and delayed osteogenic differentiation in H3.3-G34W-mutant giant cell tumor of bone. Nature Communications *11*, 5414. https://doi.org/10.1038/s41467-020-18955-y.

Majer, C.R., Jin, L., Scott, M.P., Knutson, S.K., Kuntz, K.W., Keilhack, H., Smith, J.J., Moyer, M.P., Richon, V.M., and Copeland, R.A., et al. (2012). A687V EZH2 is a gain-of-function mutation found in lymphoma patients. FEBS letters *586*, 3448-3451. https://doi.org/10.1016/j.febslet.2012.07.066.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nature reviews. Molecular cell biology *6*, 838-849. https://doi.org/10.1038/nrm1761.

Mauser, R., Kungulovski, G., Keup, C., Reinhardt, R., and Jeltsch, A. (2017). Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state. Epigenetics & Chromatin *10*, 1-19. https://doi.org/10.1186/s13072-017-0153-1.

Mendoza, A. de, Lister, R., and Bogdanovic, O. (2019). Evolution of DNA Methylome Diversity in Eukaryotes. Journal of molecular biology. https://doi.org/10.1016/j.jmb.2019.11.003.

Miescher-Rüsch, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen.

Min, J.R., Feng, Q., Li, Z.H., Zhang, Y., and Xu, R.M. (2003). Structure of the Catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase.

Mohammad, F., and Helin, K. (2017). Oncohistones: drivers of pediatric cancers. Genes & Development *31*, 2313-2324. https://doi.org/10.1101/gad.309013.117.

Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. Nature Reviews Genetics *17*, 693-703. https://doi.org/10.1038/nrg.2016.98.

Morrison, E.A., Bowerman, S., Sylvers, K.L., Wereszczynski, J., and Musselman, C.A. (2018). The conformation of the histone H3 tail inhibits association of the BPTF PHD finger with the nucleosome. eLife *7*. https://doi.org/10.7554/eLife.31481.

Murray, K. (1964). The occurrence of iε-N-methyl lysine in histones. Biochemistry *3*, 10-15.

Nacev, B.A., Feng, L., Bagert, J.D., Lemiesz, A.E., Gao, J., Soshnev, A.A., Kundra, R., Schultz, N., Muir, T.W., and Allis, C.D. (2019). The expanding landscape of 'oncohistone' mutations in human cancers. Nature *567*, 473-478. https://doi.org/10.1038/s41586-019-1038-1.

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. Nature *543*, 72-77. https://doi.org/10.1038/nature21373.

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. Molecular Cell *36*, 153-163. https://doi.org/10.1016/j.molcel.2009.07.027.

Ngo, T.T., Zhang, Q., Zhou, R., Yodh, J.G., and Ha, T. (2015). Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. Cell *160*, 1135-1144. https://doi.org/10.1016/j.cell.2015.02.001.

Oda, H., Okamoto, I., Murphy, N., Chu, J., Price, S.M., Shen, M.M., Torres-Padilla, M.E., Heard, E., and Reinberg, D. (2009). Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. Molecular and cellular biology *29*, 2278-2295. https://doi.org/10.1128/MCB.01768-08.

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. https://doi.org/10.1126/science.183.4122.330.

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. https://doi.org/10.1038/embor.2009.218.

Ozer, G., Luque, A., and Schlick, T. (2015). The chromatin fiber: multiscale problems and approaches. Current Opinion in Structural Biology *31*, 124-139. https://doi.org/10.1016/j.sbi.2015.04.002.

Pajoro, A., Severing, E., Angenent, G.C., and Immink, R.G.H. (2017). Histone H3 lysine 36 methylation affects temperature-induced alternative splicing and flowering in plants. Genome biology *18*, 102. https://doi.org/10.1186/s13059-017-1235-x.

Patel, D.J. (2016). A Structural Perspective on Readout of Epigenetic Histone and DNA Methylation Marks. Cold Spring Harbor perspectives in biology *8*, a018754. https://doi.org/10.1101/cshperspect.a018754.

Pedersen, B.S., Eyring, K., De, S., Yang, I.V., and Schwartz, D.A. (2014). Fast and accurate alignment of long bisulfite-seq reads. arXiv *arXiv:1401.1129*.

Pepenella, S., Murphy, K.J., and Hayes, J.J. (2014). Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. Chromosoma *123*, 3-13. https://doi.org/10.1007/s00412-013-0435-8.

Pfister, S.X., Ahrabi, S., Zalmas, L.-P., Sarkar, S., Aymard, F., Bachrati, C.Z., Helleday, T., Legube, G., La Thangue, N.B., and Porter, A.C.G., et al. (2014). SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. Cell reports *7*, 2006-2018. https://doi.org/10.1016/j.celrep.2014.05.026.

Qian, C., and Zhou, M.-M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cellular and molecular life sciences : CMLS *63*, 2755-2763. https://doi.org/10.1007/s00018-006-6274-5.

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. Nature structural biology *9*, 217-224. https://doi.org/10.1038/nsb759.

Rajagopalan, K.N., Chen, X., Weinberg, D.N., Chen, H., Majewski, J., Allis, C.D., and Lu, C. (2021). Depletion of H3K36me2 recapitulates epigenomic and phenotypic changes induced by the H3.3K36M oncohistone mutation. Proceedings of the National Academy of Sciences of the United States of America *118*. https://doi.org/10.1073/pnas.2021795118.

Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. Genes & Development *30*, 733-750. https://doi.org/10.1101/gad.276568.115.

Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008). Protein lysine methyltransferase G9a acts on non-histone targets. Nature Chemical Biology *4*, 344-346. https://doi.org/10.1038/nchembio.88.

Rayasam, G.V., Wendling, O., Angrand, P.-O., Mark, M., Niederreither, K., Song, L., Lerouge, T., Hager, G.L., Chambon, P., and Losson, R. (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. The EMBO Journal *22*, 3153-3163. https://doi.org/10.1093/emboj/cdg288.

Razin, A., and Cedar, H. (1991). DNA methylation and gene expression. Microbiological reviews *55*, 451-458.

Razin, A., and Cedar, H. (1994). DNA methylation and genomic imprinting. Cell *77*, 473-476. https://doi.org/10.1016/0092-8674(94)90208-9.

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

Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. Cytogenetic and Genome Research *14*, 9-25.

Rio, M., Clech, L., Amiel, J., Faivre, L., Lyonnet, S., Le Merrer, M., Odent, S., Lacombe, D., Edery, P., and Brauner, R., et al. (2003). Spectrum of NSD1

mutations in Sotos and Weaver syndromes. Journal of Medical Genetics *40*, 436. https://doi.org/10.1136/jmg.40.6.436.

Rippe, K. (2022). Liquid-Liquid Phase Separation in Chromatin. Cold Spring Harbor perspectives in biology *14*. https://doi.org/10.1101/cshperspect.a040683.

Robinson, P.J.J., and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Current Opinion in Structural Biology *16*, 336-343. https://doi.org/10.1016/j.sbi.2006.05.007.

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. https://doi.org/10.1016/j.jsb.2016.03.013.

Rosati, R., La Starza, R., Veronese, A., Aventin, A., Schwienbacher, C., Vallespi, T., Negrini, M., Martelli, M.F., and Mecucci, C. (2002). NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). Blood *99*, 3857-3860. https://doi.org/10.1182/blood.v99.10.3857.

Routh Andrew, Sandin Sara, and Rhodes Daniela (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proceedings of the National Academy of Sciences *105*, 8872-8877. https://doi.org/10.1073/pnas.0802336105.

Rozijn, T., and Tonino, G. (1964). Studies on the yeast nucleus: I. The isolation of nuclei. Biochimica et Biophysica Acta (BBA) - Specialized Section on Nucleic Acids and Related Subjects *91*, 105-112. https://doi.org/10.1016/0926-6550(64)90174-4.

Sarthy, J.F., Meers, M.P., Janssens, D.H., Henikoff, J.G., Feldman, H., Paddison, P.J., Lockwood, C.M., Vitanza, N.A., Olson, J.M., and Ahmad, K., et al. (2020). Histone deposition pathways determine the chromatin landscapes of H3.1 and H3.3 K27M oncohistones. eLife *9*. https://doi.org/10.7554/eLife.61090.

Sato, K., Kumar, A., Hamada, K., Okada, C., Oguni, A., Machiyama, A., Sakuraba, S., Nishizawa, T., Nureki, O., and Kono, H., et al. (2021). Structural basis of the regulation of the normal and oncogenic methylation of nucleosomal histone H3 Lys36 by NSD2. Nature Communications *12*, 6605. https://doi.org/10.1038/s41467-021-26913-5.

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. Proceedings of the National Academy of Sciences of the United States of America *103*, 1412-1417. https://doi.org/10.1073/pnas.0510310103.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes & Development *18*, 1251-1262. https://doi.org/10.1101/gad.300704.

Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callén, E., Celeste, A., Pagani, M., Opravil, S., and La Rosa-Velazquez, I.A. de, et al. (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes & Development *22*, 2048-2061. https://doi.org/10.1101/gad.476008.

Schuhmacher, M.K., Kudithipudi, S., Kusevic, D., Weirich, S., and Jeltsch, A. (2015). Activity and specificity of the human SUV39H2 protein lysine methyltransferase. Biochimica et biophysica acta *1849*, 55-63. https://doi.org/10.1016/j.bbagrm.2014.11.005.

Schultz, M.D., He, Y., Whitaker, J.W., Hariharan, M., Mukamel, E.A., Leung, D., Rajagopal, N., Nery, J.R., Urich, M.A., and Chen, H., et al. (2015). Human body epigenome maps reveal noncanonical DNA methylation variation. Nature *523*, 212-216. https://doi.org/10.1038/nature14465.

Seligson, D.B., Horvath, S., McBrian, M.A., Mah, V., Yu, H., Tze, S., Wang, Q., Chia, D., Goodglick, L., and Kurdistani, S.K. (2009). Global levels of histone modifications predict prognosis in different cancers. The American journal of pathology *174*, 1619-1628. https://doi.org/10.2353/ajpath.2009.080874.

Sendzikaite, G., Hanna, C.W., Stewart-Morgan, K.R., Ivanova, E., and Kelsey, G. (2019). A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice. Nature communications *10*, 1884. https://doi.org/10.1038/s41467-019-09713-w.

Shakya, A., Park, S., Rana, N., and King, J.T. (2020). Liquid-Liquid Phase Separation of Histone Proteins in Cells: Role in Chromatin Organization. Biophysical journal *118*, 753-764. https://doi.org/10.1016/j.bpj.2019.12.022.

Shi, L., Shi, J., Shi, X., Li, W., and Wen, H. (2018). Histone H3.3 G34 Mutations Alter Histone H3K36 and H3K27 Methylation In Cis. Journal of molecular biology *430*, 1562-1565. https://doi.org/10.1016/j.jmb.2018.04.014.

Shi, X., Kachirskaia, I., Yamaguchi, H., West, L.E., Wen, H., Wang, E.W., Dutta, S., Appella, E., and Gozani, O. (2007). Modulation of p53 function by SET8-mediated methylation at lysine 382. Molecular Cell *27*, 636-646. https://doi.org/10.1016/j.molcel.2007.07.012.

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. Journal of molecular biology *408*, 187-204. https://doi.org/10.1016/j.jmb.2011.01.003.

Shogren-Knaak, M.A., and Peterson, C.L. (2003). Creating Designer Histones by Native Chemical Ligation. In Methods in Enzymology : Chromatin and Chromatin Remodeling Enzymes, Part A (Academic Press), pp. 62–76.

Shogren-Knaak Michael, Ishii Haruhiko, Sun Jian-Min, Pazin Michael J., Davie James R., and Peterson Craig L. (2006). Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. Science *311*, 844-847. https://doi.org/10.1126/science.1124000.

Simon, M.D., Chu, F., Racki, L.R., La Cruz, C.C. de, Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. Cell *128*, 1003-1012. https://doi.org/10.1016/j.cell.2006.12.041.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nature reviews. Genetics *14*, 204-220. https://doi.org/10.1038/nrg3354.

Southall, S.M., Cronin, N.B., and Wilson, J.R. (2014). A novel route to product specificity in the Suv4-20 family of histone H4K20 methyltransferases. Nucleic acids research *42*, 661-671. https://doi.org/10.1093/nar/gkt776.

Streubel, G., Watson, A., Jammula, S.G., Scelfo, A., Fitzpatrick, D.J., Oliviero, G., McCole, R., Conway, E., Glancy, E., and Negri, G.L., et al. (2018). The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells. Molecular Cell *70*, 371-379.e5. https://doi.org/10.1016/j.molcel.2018.02.027.

Suganuma, T., and Workman, J.L. (2008). Crosstalk among Histone Modifications. Cell *135*, 604-607. https://doi.org/10.1016/j.cell.2008.10.036.

Takeshima, H., Suetake, I., and Tajima, S. (2008). Mouse Dnmt3a preferentially methylates linker DNA and is inhibited by histone H1. Journal of molecular biology *383*, 810-821. https://doi.org/10.1016/j.jmb.2008.03.001.

Tatton-Brown, K., Douglas, J., Coleman, K., Baujat, G., Cole, T.R.P., Das, S., Horn, D., Hughes, H.E., Temple, I.K., and Faravelli, F., et al. (2005). Genotype-Phenotype Associations in Sotos Syndrome: An Analysis of 266 Individuals with NSD1 Aberrations. The American Journal of Human Genetics *77*, 193-204. https://doi.org/10.1086/432082.

Tatton-Brown, K., and Rahman, N. (2004). Clinical features of NSD1-positive Sotos syndrome. Clinical Dysmorphology *13*.

Teissandier, A., and Bourc'his, D. (2017). Gene body DNA methylation conspires with H3K36me3 to preclude aberrant transcription. The EMBO Journal *36*, 1471-1473. https://doi.org/10.15252/embj.201796812.

Trievel, R.C., Beach, B.M., Dirk, L., Houtz, R.L., and Hurley, J.H. (2002). Structure and Catalytic Mechanism of a SET Domain Protein Methyltransferase.

Tuzon, C.T., Spektor, T., Kong, X., Congdon, L.M., Wu, S., Schotta, G., Yokomori, K., and Rice, J.C. (2014). Concerted activities of distinct H4K20 methyltransferases at DNA double-strand breaks regulate 53BP1 nucleation and NHEJ-directed repair. Cell reports *8*, 430-438. https://doi.org/10.1016/j.celrep.2014.06.013.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., and Holt, R.A., et al. (2001). The sequence of the human genome. Science *291*, 1304-1351. https://doi.org/10.1126/science.1058040.

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

Wagh, K., Garcia, D.A., and Upadhyaya, A. (2021). Phase separation in transcription factor dynamics and chromatin organization. Current Opinion in Structural Biology *71*, 148-155. https://doi.org/10.1016/j.sbi.2021.06.009.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nature reviews. Molecular cell biology *13*, 115-126. https://doi.org/10.1038/nrm3274.

Wang, H., Farnung, L., Dienemann, C., and Cramer, P. (2020). Structure of H3K36-methylated nucleosome-PWWP complex reveals multivalent cross-gyre binding. Nature structural & molecular biology *27*, 8-13. https://doi.org/10.1038/s41594-019-0345-4.

Wang, L., Gao, Y., Zheng, X., Liu, C., Dong, S., Li, R., Zhang, G., Wei, Y., Qu, H., and Li, Y., et al. (2019). Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. Molecular Cell *76*, 646-659.e6. https://doi.org/10.1016/j.molcel.2019.08.019.

Wang, Y., and Jia, S. (2009). Degrees make all the difference: the multifunctionality of histone H4 lysine 20 methylation. Epigenetics *4*, 273-276. https://doi.org/10.4161/epi.4.5.9212.

Weinberg, D.N., Papillon-Cavanagh, S., Chen, H., Yue, Y., Chen, X., Rajagopalan, K.N., Horth, C., McGuire, J.T., Xu, X., and Nikbakht, H., et al. (2019). The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. Nature *573*, 281-286. https://doi.org/10.1038/s41586-019-1534-3.

Weinberg, D.N., Rosenbaum, P., Chen, X., Barrows, D., Horth, C., Marunde, M.R., Popova, I.K., Gillespie, Z.B., Keogh, M.-C., and Lu, C., et al. (2021). Two competing mechanisms of DNMT3A recruitment regulate the dynamics of de novo DNA methylation at PRC1-targeted CpG islands. Nat Genet *53*, 794-800. https://doi.org/10.1038/s41588-021-00856-5.

Weirich, S., Kudithipudi, S., Kycia, I., and Jeltsch, A. (2015). Somatic cancer mutations in the MLL3-SET domain alter the catalytic properties of the enzyme. Clinical epigenetics *7*, 36. https://doi.org/10.1186/s13148-015-0075-3.

Weirich, S., Schuhmacher, M.K., Kudithipudi, S., Lungu, C., Ferguson, A.D., and Jeltsch, A. (2020). Analysis of the Substrate Specificity of the SMYD2 Protein Lysine Methyltransferase and Discovery of Novel Non-Histone Substrates. Chembiochem : a European journal of chemical biology *21*, 256-264. https://doi.org/10.1002/cbic.201900582.

Widom, J. (1992). A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells. Proceedings of the National Academy of Sciences of the United States of America *89*, 1095-1099. https://doi.org/10.1073/pnas.89.3.1095.

Winogradoff, D., and Aksimentiev, A. (2019). Molecular Mechanism of Spontaneous Nucleosome Unraveling. Journal of molecular biology *431*, 323-335. https://doi.org/10.1016/j.jmb.2018.11.013.

Wu, C., and Morris, J.R. (2001). Genes, genetics, and epigenetics: a correspondence. Science (New York, N.Y.) *293*, 1103-1105. https://doi.org/10.1126/science.293.5532.1103.

Wu, H., Siarheyeva, A., Zeng, H., Lam, R., Dong, A., Wu, X.-H., Li, Y., Schapira, M., Vedadi, M., and Min, J. (2013a). Crystal structures of the human histone H4K20 methyltransferases SUV420H1 and SUV420H2. FEBS letters *587*, 3859-3868. https://doi.org/10.1016/j.febslet.2013.10.020.

Wu, L., Lee, S.Y., Zhou, B., Nguyen, U.T., Muir, T.W., Tan, S., and Dou, Y. (2013b). ASH2L Regulates Ubiquitylation Signaling to MLL: trans-Regulation of H3 K4 Methylation in Higher Eukaryotes. Molecular Cell *49*, 1108-1120. https://doi.org/10.1016/j.molcel.2013.01.033.

Xiao, B., Wilson, J.R., and Gamblin, S.J. (2003). SET domains and histone methylation. Current Opinion in Structural Biology *13*, 699-705. https://doi.org/10.1016/j.sbi.2003.10.003.

Xu, T.H., Liu, M., Zhou, X.E., Liang, G., Zhao, G., Xu, H.E., Melcher, K., and Jones, P.A. (2020). Structure of nucleosome-bound DNA methyltransferases DNMT3A and DNMT3B. Nature *586*, 151-155. https://doi.org/10.1038/s41586-020-2747-1.

Yokoyama, Y., Matsumoto, A., Hieda, M., Shinchi, Y., Ogihara, E., Hamada, M., Nishioka, Y., Kimura, H., Yoshidome, K., and Tsujimoto, M., et al. (2014). Loss of histone H4K20 trimethylation predicts poor prognosis in breast cancer and is associated with invasive activity. Breast cancer research : BCR *16*, R66. https://doi.org/10.1186/bcr3681.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science *328*, 916-919. https://doi.org/10.1126/science.1186366.

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nature structural & molecular biology *20*, 259-266. https://doi.org/10.1038/nsmb.2470.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614-620. https://doi.org/10.1093/bioinformatics/btt593.

Zhang, T., Cooper, S., and Brockdorff, N. (2015). The interplay of histone modifications - writers that read. EMBO reports *16*, 1467-1481. https://doi.org/10.15252/embr.201540945.

Zhang, X., Wen, H., and Shi, X. (2012). Lysine methylation: beyond histones. Acta biochimica et biophysica Sinica *44*, 14-27. https://doi.org/10.1093/abbs/gmr100.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., and 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. https://doi.org/10.1093/nar/gkq147.

Zhang, Z.M., Lu, R., Wang, P., Yu, Y., Chen, D., Gao, L., Liu, S., Ji, D., Rothbart, S.B., and Wang, Y., et al. (2018). Structural basis for DNMT3A-mediated de novo DNA methylation. Nature *554*, 387-391. https://doi.org/10.1038/nature25477.

Zhao, Y., and Garcia, B.A. (2015). Comprehensive Catalog of Currently Documented Histone Modifications. Cold Spring Harbor perspectives in biology *7*, a025064. https://doi.org/10.1101/cshperspect.a025064.

Zhu, P., and Li, G. (2016). Structural insights of nucleosome and the 30-nm chromatin fiber. Current Opinion in Structural Biology *36*, 106-115. https://doi.org/10.1016/j.sbi.2016.01.013.

Zucconi, B.E., and Cole, P.A. (2017). Allosteric regulation of epigenetic modifying enzymes. Current opinion in chemical biology *39*, 109-115. https://doi.org/10.1016/j.cbpa.2017.05.015.

7 Appendix

Appendix 1 (not included in the published thesis)

Bröhm, A., Schoch, T., Dukatz, M., Graf, N., Dorscht, F., Mantai, E., Adam, S., Bashtrykov, P., and Jeltsch, A. (2022a). Methylation of recombinant mononucleosomes by DNMT3A demonstrates efficient linker DNA methylation and a role of H3K36me3. Communications Biology *5*, 192. https://doi.org/10.1038/s42003-022-03119-z.

Appendix 2 (not included in the published thesis)

Bröhm, A., Elsawy, H., Rathert, P., Kudithipudi, S., Schoch, T., Schuhmacher, M.K., Weirich, S., and Jeltsch, A. (2019). Somatic Cancer Mutations in the SUV420H1 Protein Lysine Methyltransferase Modulate Its Catalytic Activity. Journal of molecular biology *431*, 3068-3080. https://doi.org/10.1016/j.jmb.2019.06.021.

Appendix 3 (not included in the published thesis)

Bröhm, A., Schoch, T., Grünberger, D., Khella, M.S., Schuhmacher, M.K., Weirich, S., and Jeltsch, A. (2022b). The H3.3 G34W oncohistone mutation increases K36 methylation by the protein lysine methyltransferase NSD1. Biochimie 198, 86-91. https://doi.org/10.1016/j.biochi.2022.03.007.