

Aus dem Adolf-Butenandt-Institut, Lehrstuhl: Molekularbiologie
im Biomedizinischen Centrum München (BMC)

der

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Vorstand: Prof. Dr. Peter B. Becker

und der

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Analysis of the multivalent binding properties of the novel H2A.Z interactor PWWP2A



Dissertation zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät der Ludwig-Maximilians-Universität München

Ramona Spitzer

aus Günzburg

2018

**Mit Genehmigung der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München**

Betreuerin: Prof. Dr. Sandra B. Hake

Zweitgutachter: Prof. Dr. Andreas Ladurner

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 14.09.2018

Eidesstattliche Versicherung

Ich, Ramona Spitzer, erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

*“Analysis of the multivalent binding properties of the novel H2A.Z interactor
PWWP2A”*

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe. Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den14.09.2018.....

.....
(Ramona Spitzer)

Acknowledgements

First, I would like to thank my supervisor Prof. Dr. Sandra B. Hake for the opportunity to work as a PhD student in her lab. Sandra, you always pushed my thesis with constant input and ideas and it was always fun to discuss my project or to defend my point of view. You gave me the opportunity to improve myself in giving talks, discussing and defending scientific issues and strengthening my personality. On top, you always had a sympathetic ear for me, no matter if it was about scientific questions or personal problems. Thank you so much for letting me be a part of the amazing “HAKE-GIRLS”.

Next, I want to thank Dr. Catherine Regnard (“PWWP”-Team) and Dr. Felix Müller-Planitz for being members of my thesis advisory committee. There was always an “open door” to discuss questions and to get helpful advice for my PhD project.

I am happy for the opportunity to work in the Biomedical Center (BMC) and being part of a great “chromatin-network-community” and the IRTG1064 (thanks Elizabeth for being such a great coordinator). There was no question, method or reagent which could not be answered, learned or “borrowed”. I want to thank all former members of the Hake-Lab and of course our former master student Lucia (sushi & rice-balls-4-ever). It was a pleasure to work together with you and I really enjoyed your company. Furthermore, I want to thank Stephi Link (the crazy cat lady) for being the other 50% of the Hake-Pair (LMU-Crew). It was always fun to work together with you and discuss all the gossip, which was floating around. I think with you, it is not possible to miss anything new or secret in the lab community.

I was also lucky to share the lab with the former and current FMP’s (especially Sabrina, Ashish and Petra), where I got constantly input for “*in vitro*” questions like purification (Äkta), assemblies and cloning strategies or just having a little chit-chat.

In addition, I am grateful that I could experience collaborative work with outstanding scientists here in Munich and all over the world: Stephanie Link, Moritz Völker-Albert, Axel Imhof, Masahiko Harata, Thomas Burgold, Brian Hendrich, Mariam Sana, Mario Torrado del Rey and Joel Mackay.

Last but not least I want to thank my mum Elisabeth and my dad Anton for always standing behind me, my brothers Jonathan and Fabian and my whole family and friends for their constant support and motivating words. A special thanks goes to my better half Tobi who was always at my side, motivating me when things weren’t running well and pushing me to go on. Thank you!

*In silent remembrance of Karl Bader who always believed in science and technical progress
and Agnes Müller, my beloved grandmother, who is greatly missed here.*

TABLE OF CONTENTS

Summary	1
Zusammenfassung.....	3
1 Introduction	5
1.1 The structure of chromatin	5
1.2 Alteration of chromatin structure	8
1.2.1 Posttranslational modifications of histones.....	11
1.2.2 Histone variants.....	13
1.2.3 H2A variants	15
1.3 The histone variant H2A.Z	17
1.4 H2A.Z binding proteins	21
1.5 PWWP domain-containing proteins	25
1.6 Objectives	27
2 Material and Methods	28
2.1 Materials.....	28
2.1.1 Technical devices.....	28
2.1.2 Consumables	30
2.1.3 Chemicals	31
2.1.4 Kits, enzymes and markers.....	33
2.1.5 Antibodies	34
2.1.6 Plasmids.....	35
2.1.7 Oligonucleotides.....	36
2.1.8 Bacterial strains and cell lines	38
2.1.9 Software	38
2.1.10 Buffers and solutions.....	39
2.2 Cell biological methods.....	41
2.2.1 Mammalian cell cultivation	41
2.2.2 Transient transfection of mammalian cells.....	42
2.3 Molecular biological methods	42
2.3.1 Agarose gel electrophoresis	42
2.3.2 Cloning of PWWP2A truncations.....	42
2.3.3 Site-directed mutagenesis.....	43

2.4	Biochemical methods	44
2.4.1	SDS-polyacrylamide gel electrophoresis (SDS-PAGE)	44
2.4.2	Coomassie staining of SDS-PAGE gels	44
2.4.3	Immunoblotting	44
2.4.4	Mononucleosome preparation of HeLa Kyoto cells.....	45
2.4.5	Purification of MNase digested DNA.....	46
2.4.6	Purification of recombinant GST-tagged PWWP2A and PWWP2A truncations	47
2.4.7	Identification of histone modifications by Mass-Spectrometry (MS).....	50
2.4.8	Reconstitution of recombinant mononucleosomes	52
2.4.9	<i>In vitro</i> binding assays with recombinant PWWP2A constructs	58
2.4.10	NMR spectroscopy	62
2.5	Bioinformatics.....	62
2.5.1	MS Data analysis.....	62
3	Results.....	64
3.1	PWWP2A is a direct nucleosome interaction protein	64
3.2	PWWP2A's internal region divides labor – general nucleosome binding and specific H2A.Z recognition	67
3.3	H2A.Z's C-terminus affects binding to PWWP2A in a HK cell-derived system.....	70
3.4	Point mutations within IC impair binding to H2A and H2A.Z.....	73
3.5	Single histone tail deletions do not impair binding of IN <i>in vitro</i>	77
3.6	IN interacts with free DNA and binds nucleosomal linker DNA	79
3.7	PWWP2A's PWWP domain directly interacts with recombinant mononucleosomes..	83
3.8	The PWWP domain interacts with free nucleic acids and nucleosomal linker DNA.	85
3.9	Combination of S and PWWP domain interacts with <i>in vivo</i> -derived H3K36me3-mononucleosomes	91
3.10	PWWP2A interacts with γ -H2A.X mononucleosomes.....	102
4	Discussion.....	106
4.1	PWWP2A is a multivalent chromatin binder and specifically recognizes H2A.Z	106
4.2	A combination of S and PWWP interacts in a sequence- and methylation state-specific manner with H3K36me3-nucleosomes	112
4.3	PWWP2A's C-terminus interacts with the modified histone variant γ -H2A.X.....	115
4.4	Summary and future perspectives	117
	ABBREVIATIONS.....	122
	BIBLIOGRAPHY.....	125

APPENDIX	155
CURRICULUM VITAE	158

Summary

The constraint space of the eukaryotic cell nucleus requires the specific packing of DNA with basic histone proteins into chromatin, with the nucleosome as its basic subunit. The nucleosome itself consists of approximately 147 bp of DNA wrapped around a histone octamer, which in turn is composed of two copies of each core histone H2A, H2B, H3 and H4. In order to control DNA-related processes like replication or transcription, the packaged DNA has to be always accessible. One possibility to achieve or prevent access to DNA sequences is the replacement of canonical histones with histone variants, which can lead to altered chromatin structure, different posttranslational histone modifications or the interaction with distinct, specific effector proteins. To date, several human histone H2A variants are known including controversial H2A.Z. In order to explain the different functions of the essential histone variant H2A.Z, our lab performed pull-down experiments with GFP-tagged H2A.Z and discovered a novel H2A.Z-interaction protein – PWWP2A. PWWP2A is a vertebrate-specific protein, involved in cell cycle control and essential for eye and brain development in frogs. Interestingly, PWWP2A possesses multiple domains, with the predicted to be disordered internal region (I) and the eponymous PWWP domain being the most interesting. In my PhD thesis, I aimed to investigate the multivalent binding properties of PWWP2A's distinct domains – especially in regard of H2A.Z – and to determine whether it interacts with further members of the H2A variant family. First, I showed PWWP2A's direct interaction with H2A.Z-containing mononucleosomes and its ability to distinguish between canonical H2A and H2A.Z *in vitro*. In order to identify the domains of PWWP2A being able to interact with chromatin and H2A.Z, pull-down experiments with HeLa Kyoto cell-derived mononucleosomes or competitive Electrophoretic Mobility Shift Assays (cEMSA) with *in vitro* reconstituted mononucleosomes were performed. I demonstrated that PWWP2A harbors at least two interaction domains – the internal region (I) and the C-terminal PWWP domain, whereby the internal region can be further subdivided into an N- (IN) and C-terminal part (IC). Here, IN mediates general nucleosome binding via DNA interaction and IC confers to H2A.Z-specific nucleosome recognition. Moreover, I showed that the C-terminal PWWP domain interacts with free nucleic acids and nucleosomal flanking linker DNA. Furthermore, I unraveled that the combination of the serine rich stretch S and PWWP is crucial to recognize H3K36me3-nucleosomes in a methylation state- and site-specific manner. Additionally, I demonstrated PWWP2A being an interactor of another member of the histone H2A variant family – γ -H2A.X.

SUMMARY

Altogether, I gained insights how PWWP2A is able to establish a strong chromatin association and specific interaction with H2A.Z, where it might act as an adaptor to recruit further chromatin interactors. Furthermore, the association of PWWP2A with histone mark H3K36me3 and γ -H2A.X-nucleosomes indicates next to H2A.Z-nucleosome binding additional interaction modes, which might be crucial for the regulation of cellular mechanisms like DNA damage response. Hence, future experiments are indispensable to further investigate PWWP2A's various functions at distinct chromatin moieties and may help to understand how the complex interplay of different chromatin features may work.

Zusammenfassung

Der begrenzte Raum des eukaryotischen Zellkerns erfordert die spezifische Verpackung der DNA mit basischen Histonproteinen in Chromatin, wobei das Nukleosom als grundlegende Untereinheit dient. Das Nukleosom selbst besteht aus circa 147 bp DNA, die um ein Histonoktamer gewickelt ist, welches wiederum aus zwei Kopien jedes Kernhistons H2A, H2B, H3 und H4 besteht. Um DNA-relevante Prozesse wie Replikation oder Transkription zu steuern, muss die verpackte DNA jederzeit zugänglich sein. Eine Möglichkeit, die den Zugang zu DNA-Sequenzen ermöglicht oder verhindert, ist der Austausch kanonischer Histone durch Histon-Varianten, was zu einer veränderten Chromatin-Struktur, verschiedenen posttranslationalen Histon-Modifikationen oder der Interaktion mit unterschiedlichen, spezifischen Effektor-Proteinen führen kann. Bis heute sind mehrere humane Histon H2A-Varianten bekannt, mitunter die kontroverse Variante H2A.Z. Um die verschiedenen Funktionen der essentiellen Histonvariante H2A.Z zu erklären, führte unser Labor Pull-Down-Experimente mit GFP-markiertem H2A.Z durch und entdeckte ein neuartiges H2A.Z-Interaktionsprotein – PWWP2A. PWWP2A ist ein vertebraten-spezifisches Protein, das an der Zellzykluskontrolle beteiligt ist und für die Augen- und Gehirnentwicklung bei Fröschen essentiell ist. Interessanterweise besitzt PWWP2A mehrere Domänen, wobei die vermutlich ungeordnete interne Region (I) und die namensgebende PWWP-Domäne am interessantesten sind. Das Ziel meiner Dissertation war die Untersuchung, der durch PWWP2A's verschiedenen Domänen vermittelten, multivalenten Bindungseigenschaften – insbesondere im Hinblick auf H2A.Z – und herauszufinden, ob es mit weiteren Mitgliedern der H2A-Variantenfamilie interagiert. Zuerst zeigte ich PWWP2A's direkte Interaktion mit H2A.Z-haltigen Mononukleosomen und seine Fähigkeit, zwischen kanonischem H2A und H2A.Z *in vitro* zu unterscheiden. Um die Domänen von PWWP2A zu identifizieren, die in der Lage sind, mit Chromatin und H2A.Z zu interagieren, wurden Pull-Down-Experimente mit aus HeLa Kyoto-Zellen isolierten Mononukleosomen oder kompetitive Electrophoretic Mobility Shift Assays (cEMSA) mit *in vitro* rekonstituierten Mononukleosomen durchgeführt. Ich habe gezeigt, dass PWWP2A mindestens zwei Interaktionsdomänen beherbergt - die interne Region (I) und die C-terminale PWWP-Domäne, wobei die interne Region weiter in einen N- (IN) und C-terminalen Teil (IC) unterteilt werden kann. IN vermittelt dabei eine generelle Bindung an Nukleosomen, indem es mit DNA interagiert und IC eine spezifische Erkennung von H2A.Z. Außerdem habe ich gezeigt, dass die C-terminale PWWP-Domäne mit freien Nukleinsäuren

ZUSAMMENFASSUNG

und nukleosomaler flankierender Linker-DNA interagiert. Des Weiteren konnte ich klären, dass die Kombination aus der Serin-reichen Region S und PWWP entscheidend ist, um H3K36me3-Nukleosomen in einer Methylierungszustands- und ortsspezifischen Weise zu erkennen. Zusätzlich habe ich bewiesen, dass PWWP2A mit einem weiteren Mitglied der Histon H2A-Variantenfamilie interagiert – γ -H2A.X.

Insgesamt habe ich Einblicke gewonnen, wie PWWP2A in der Lage ist, eine starke Assoziation mit Chromatin und eine spezifische Interaktion mit H2A.Z auszubilden, wo es als Adapter fungieren könnte, um weitere Chromatin-Interaktoren zu rekrutieren. Die Assoziation von PWWP2A mit der Histon-Modifikation H3K36me3 und γ -H2A.X-Nukleosomen zeigt, dass neben der H2A.Z-Nukleosomenbindung noch zusätzliche Interaktionsmodi vorhanden sind, die für die Regulation von zellulären Mechanismen wie der DNA-Schadensantwort entscheidend sein könnten. Zukünftige Experimente zur Untersuchung von PWWP2A's verschiedenen Funktionen an unterschiedlichen Chromatin-Einheiten sind daher unerlässlich und können helfen, das komplexe Zusammenspiel verschiedener Chromatin-Eigenschaften zu verstehen.

1 Introduction

1.1 The structure of chromatin

The most important function of DNA is to carry the genetic code, the major sequence information to generate proteins that make up every organism [1]. The DNA of each single human cell has a length of almost 2 m [2, 3] (3.2×10^9 nucleotides) but fits easily into the cell nucleus with a diameter around 6 μm , thus being geometrically comparable with packing 40 km of wafer-thin fine wire into a tennis ball [4]. Moreover, since the human genome got sequenced in 2001 [5, 6] it became obvious, that the genetic code itself cannot explain each cellular process [7] such as the differentiation of multicellular organisms from a unique DNA sequence present in each single cell. At this point, chromatin enters the stage. DNA is not packed as a naked molecule into the eukaryotic nucleus, but is assembled as a nucleoprotein complex – the so-called chromatin – which is composed of a basic repeating subunit, the nucleosome [8, 9]. The nucleosome itself consists of a disc-shaped nucleosome core particle (NCP), composed of two copies of each histone H2A, H2B, H3 and H4 and 145-147 bp of superhelical, left-handed DNA wrapped 1.75 times around the octamer in a left-handed manner [3, 10, 11] (**Figure 1.1.1**).

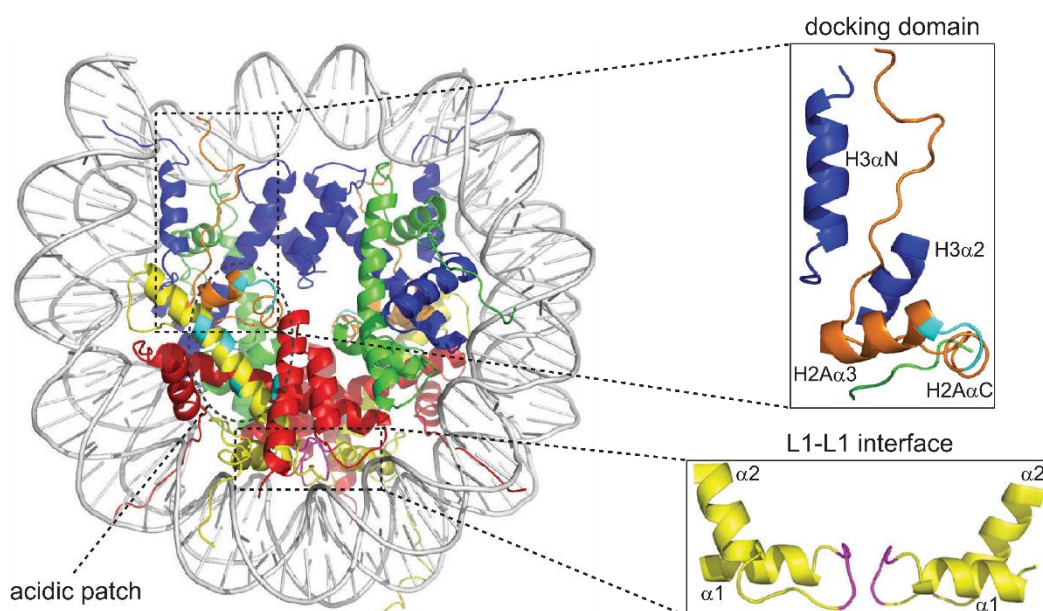


Figure 1.1.1 The nucleosome crystal structure. Nucleosomal crystal structure as described by Luger et al. [10]. H2A is depicted in yellow, H2B in red, H3 in blue, H4 in green and DNA in grey. Docking domain (orange), acidic patch (cyan) and L1-L1 interface (magenta) are shown as blow-ups on the right. Figure reprinted from [12], with permission of Oxford University Press.

INTRODUCTION

The four “core” histones can be functionally and structurally divided into two parts: the histone tail and the histone fold [2]. The N-terminal core histone tail is largely unstructured and consists of up to 35 residues, which corresponds to almost 20% of its total amino acids. The core histone fold domain (HFD) in turn consists of 80-90 amino acid residues and exhibits an additional, unique secondary structure mediating interactions between histones [13] and DNA [14]. Further, the different HFDs of all core histones are highly similar within their structural “ $\alpha 1$ -L1- $\alpha 2$ -L2- $\alpha 3$ ” motif, where each HFD forms three α -helices ($\alpha 1$ - $\alpha 3$) connected by two loops (L1 and L2) [10]. In solution, histones obligatorily form heterodimers, whereby H2A only pairs with H2B and H3 only pairs with H4 [2]. Two histone H3-H4 pairs form via a 4-helix bundle a (H3-H4)₂ tetramer in a twisted W-shaped arrangement. The histone octamer is formed either in presence of DNA or in high-salt conditions (2 M) by the interaction of two H2A-H2B dimers with one (H3-H4)₂ tetramer [13, 15]. Therefore, H2B is connected via a four-helix bundle with H4 and a further interaction profile is formed between the (H3-H4)₂ tetramer and H2A’s docking domain [2]. Finally, a 147 bp DNA fragment – its minor groove always facing the basic surface of histones – interacts base-unspecific with the octamer, thereby forming the nucleosome [2, 8, 10]. NCPs, connected via linker DNA to the neighbored nucleosome core, form the primary so-called nucleosomal ‘beads-on-a-string’ structure [16-18]. The secondary chromatin structure is created by further organizing the nucleosomal arrays into a highly condensed 30 nm fiber by the linker histones H1 and H5 [19, 20] whereas the tertiary structure is formed by large-scale configurations that build the entire chromosome [17]. Focusing on the 30 nm fiber, extensive biochemical and biophysical studies were carried out but the existence of a 30 nm fiber *in vivo* is still under discussion [21]. Nevertheless, there exist two competing models for its structure: the solenoid model and the zigzag model [17] (**Figure 1.1.2**). The one-start helix or solenoid model proposes that neighbored nucleosomes are tightly coiled up and connected via variable linker DNA lengths to a compact structure [20, 22]. In turn, the two-start helix or zigzag model suggests that nucleosomes are arranged with straight linker DNA connecting oppositely placed nucleosomes with each other [23, 24].

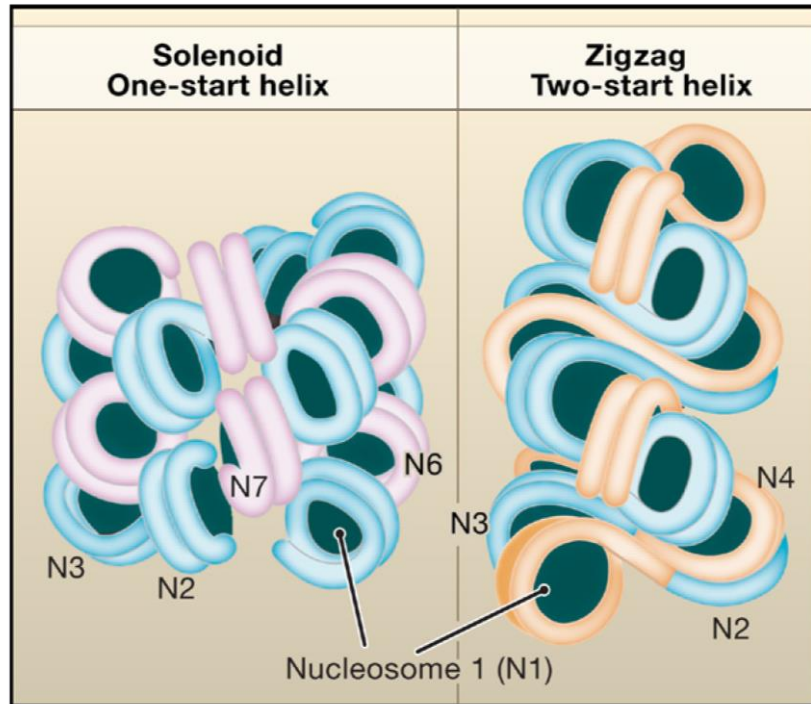


Figure 1.1.2 Two current models of the 30 nm chromatin fiber. (Left) Solenoid model of the 30 nm fiber modeled without DNA, proposed by [25]. Helical gyres are shown in magenta and blue, respectively. (Right) Zigzag model of the 30 nm fiber with linker DNA criss-crossing between NCPs, proposed by [26]. Helical gyres are shown in orange and blue, respectively. Figure reprinted from [17], with permission from Elsevier.

To understand the specific regulation of gene expression, recent technological advances fostered the research on certain chromatin states that repress or activate gene activity [27]. Cytologically, chromatin can be partitioned into two major states: heterochromatin and euchromatin [27, 28]. Heterochromatin mainly consists of repressive, condensed chromatin regions and harbors repetitive DNA elements [28, 29] whereas euchromatin is active, gene-rich and less condensed [30]. It is assumed, that the switch between eu- and heterochromatin is fundamental for the accessibility of DNA thus enabling repressive or active transcriptional states [27].

Since the main focus of this thesis is set on the work with human cells, I refer to the situation in mammals in the following sections, unless otherwise stated.

1.2 Alteration of chromatin structure

As mentioned in 1.1, chromatin is partitioned into two major states – euchromatin and heterochromatin. Heterochromatin can be additionally subdivided into a constitutive and a facultative form, whereas constitutive heterochromatin describes a permanent, transcriptionally silenced and facultative heterochromatin just a temporarily silenced state [31]. However, regarding the lacking explanation for distinct cellular processes and differentiation events only based on DNA [7], it was obvious that there has to be a connection between the biological activity and chromatin structure [32]. Already in 1942, Conrad Waddington pointed out that epigenetics is the causal mechanism that brings the genotype into phenotype [33], whereas today's widely accepted opinion describes epigenetics as the study of genomic alterations without changing the DNA sequence [34]. This definition is based on heritable, cell identity-defining instructions that regulate how, when and whether genetic information will be transcribed and translated [35]. In general, chromatin structure and thus DNA packaging dynamics are altered by distinct, but interconnected key players like DNA methylation, non-coding RNAs (ncRNAs), nuclear architecture, ATP-dependent chromatin remodelers, post-translational modifications (PTM) of histones or the incorporation/replacement of histone variants [36].

The chemical modification of DNA by methylating the fifth position of cytosine (5mC) is one possibility to alter chromatin structure [37, 38]. The covalent modification occurs at 60-80% of palindromic, symmetrical CpG sites [39-41], is associated with direct or indirect transcription silencing and has fluctuating levels for different cell types, gene states (active versus silenced) and genetic locations (gene body versus promoter) [42]. A direct way to repress gene transcription is either to prevent binding of transcription factors (TFs) to their recognition motif [43] or the recruitment of methyl-CpG domain-binding proteins fostering local chromatin condensation [44, 45]. However, recent studies also imply that methylated cytosines might play a role in gene activation [46]. Moreover, 5mC is linked to biological functions such as nucleosome positioning [47, 48], splicing [49] or genomic and chromosomal stability [50, 51]. In general, the modification is established and maintained by three conserved DNA methyltransferases (DNMT1, DNMT3A and DNMT3B) [52, 53]. DNMT3A and DNMT3B are fundamental for *de novo* cytosine methylation [53], whereas DNMT1 is responsible for the maintenance of CpG methylations and the modification of newly

INTRODUCTION

synthesized CpGs during DNA replication [54]. Interestingly, the absence of DNMT1 results in decreasing 5mC levels [55] pointing towards a specific demethylation of CpG sites [56]. Indeed, members of the Ten Eleven Translocation (TET) dioxygenase family (TET1, TET2 and TET3) were shown to catalyze the demethylation of DNA [55] by oxidizing 5mC into three distinct forms: 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [57-59]. The stable modification 5hmC [60, 61] is associated with promoters and gene bodies of highly expressed genes [62, 63], whereas 5fC plays a role in transcription regulation [64] and 5caC reduces the substrate specificity of RNA polymerase II [65].

Second, the structure of chromatin can be influenced by long noncoding RNAs (lncRNAs) constituting a heterogenic class of RNAs [66]. To distinguish lncRNAs from small noncoding RNAs [67] they are defined as “non-protein-coding transcripts larger than 200 nucleotides” [68]. However, there is no standard classification and lncRNA nomenclature is rather based on features like transcription origin, molecular function or tissue specificity [69]. Surprisingly, lncRNAs are evolutionary poorly conserved [70]. Nevertheless, their function (recruitment of chromatin/histone modifying enzymes, nucleosome positioning modulation) can be maintained via a high structural conservation [71-73]. One of the most popular functions of lncRNAs is the inactivation of the X-chromosome in female mammals [74, 75]. There, the important lncRNA Xist is sufficient to silence X-chromosome linked genes [76-78] by coating the inactive X [79]. However, in addition to their role in dosage compensation, lncRNAs seem to be involved in distinct cancer types [80] thus being in focus of recent research.

In addition to DNA methylation and lncRNAs, chromatin structure can be also altered by the presence of architectural proteins. These proteins guarantee via intra- and inter-nucleosome interactions the accessibility of eukaryotic DNA by influencing the conformation of the chromatin fiber (see also **Figure 1.1.2**) [81]. In the genome, architectural proteins are generally localized at architectural protein binding sites (APBSs) [82] that can be subdivided, dependent on the number of architectural and associated accessory proteins, in low-occupancy or high-occupancy APBS [83]. Thereby, the specific binding of architectural proteins to different binding sites leads to fundamental roles in genome organization and function [82] like 3D genome organization or the mediation of functional chromatin interactions [83]. Amongst others, the CCCTC-binding factor (CTCF) – a zinc finger protein – is the best characterized

INTRODUCTION

architectural protein in vertebrates [84]. In general, CTCF-binding sites are often associated with different methylation states of DNA (see above) [85], but also DNA methylation-independent binding sites exist, pointing towards CTCF's complex functionality like enhancer blocking, transcriptional pausing, alternative mRNA splicing or as boundary element [84]. To control biological functions such as gene expression, architectural proteins like CTCF facilitate the contact between different genome regions by bringing them into close proximity [84]. Moreover, to prevent the interaction of adjacent domains, topologically associating domains (TADs) are enriched for clusters of CTCF binding sites [86] thereby establishing TAD borders [87]. In addition to its function as TAD border, CTCF was also identified as boundary element for lamin-associated domains (LADs), which are linked to the nuclear lamina [84, 88]. Although it is tempting to assume CTCF being the main contributor of border formation, it is unlikely that this is one of its main functions [84].

The compact organization of chromatin within the cell nucleus has to be still enough flexible to enable access of enzymes and regulatory factors to DNA [89]. One possibility to render the accessibility of nucleosomal DNA is ATP-dependent chromatin remodeling [90]. Multi-protein chromatin remodeler complexes are remarkably diverse [91], harbor an ATPase subunit that enables chromatin remodeling by ATP-hydrolysis [92] and up to 20 non-catalytic subunits that are indispensable for the regulation of chromatin remodeling activities [91]. The catalytic ATPase subunit is highly conserved and flanked by defined domains, therefore remodelers can be divided into four families: SWI/SNF (switch/sucrose non-fermentable), ISWI (imitation switch), CHD (chromodomain helicase DNA-binding protein) and INO80 (inositol-requiring 80) [91, 93, 94]. Members of the SWI/SNF family possess an N-terminal HSA (helicase-SANT) domain that recruits actin-related proteins and a C-terminal bromodomain, which is known to interact with acetylated histone lysines [95]. This family is capable of sliding and/or evicting nucleosomes [96] and plays a role as suppressive complex in pancreatic cancer [97]. ATPases of the ISWI family are defined by the presence of a C-terminally located SANT domain neighbored to a SLIDE domain, which together recognize nucleosomes and bind to unmodified H4 tails and DNA [92]. ISWI family complexes are involved in transcriptional repression [98], nucleosome positioning and gene silencing [99]. CHDs harbor two chromo domains and an N-terminal ATPase domain [94, 100]. Chd1, a member of the CHD family, co-localizes with promoters of actively transcribed genes and prevents the accumulation of heterochromatic regions [101]. Members of the INO80 group exhibit helicase activity via an unique ATPase split

INTRODUCTION

domain [102] and play a crucial role in DNA double strand break repair and protein recruitment to DNA damage sites in human cells [103].

The following chapters explain two further possibilities of chromatin regulation in more detail.

1.2.1 Posttranslational modifications of histones

As described in 1.2 not only DNA is covalently modified but also histones that build up the entire nucleosome. These PTMs occur in a large number of different variations: methylation, glycosylation, ubiquitination, propionylation, acetylation, deimination, phosphorylation, ADP-ribosylation, formylation, crotonylation and SUMOylation [104-106]. Especially the intrinsically disordered and highly flexible N-terminal histone tail is subjected to PTMs [10], but also the histone core can be posttranslationally modified [107]. More than a decade ago, T. Jenuwein and D. Allis already proposed that histone modifications do not occur individually but in different combinations thus generating a specific "histone code" [108]. Hence, it would be possible to control biological processes and mechanisms with far-reaching consequences for distinct cell fates [108]. In general, the modification of histones is very dynamic, strongly influences the genomic structure and is regulated by highly specific enzymes [104, 109]. "Writers" such as (1) histone acetyltransferases, (2) histone methyltransferases, (3) kinases and (4) ubiquitin ligases catalyze posttranslational modifications like (1) acetylations (lysine), (2) methylations (arginine or lysine), (3) phosphorylations (tyrosine, threonine or serine) and (4) ubiquitinations (lysine), whereas "erasers" like histone deacetylases, demethylases, phosphatases or ubiquitin proteases remove them [110-113]. The functional influence of PTMs on chromatin can be either direct by changing chromatin structure or indirect by recruiting effector proteins [110]. These epigenetic regulators are so-called "readers", which harbor specific domains such as e.g. tudor, MBT (Malignant Brain Tumor), chromo, WD40, bromo and PWWP domains that can interact with specific modifications [112]. Proteins of the tudor family bind to methylated substrate residues (arginine or lysine) thereby regulating RNA metabolism, small RNA pathways or DNA damage response and function as adaptor proteins to facilitate the assembly of macromolecular complexes [114]. MBT domain-containing proteins interact with mono- and di-methylated lysines on histone H3 and H4 tails, are associated with neurological tumors in humans and play a crucial role in gene regulatory pathways [115]. Members of the chromo domain-containing family bind to di- and trimethylated lysines (Kme2, Kme3) and are involved in DNA degradation, transcription and

INTRODUCTION

chromatin structure [116-118], whereas the β -propeller structured WD40 domain family interacts with methylated H3 lysines and participates in signal transduction, cytoskeleton assembly, RNA processing and cell division [119, 120]. Bromo domain-harboring proteins in contrast recognize acetylated lysines (Kac) and have important roles in gene expression and cancer development [121, 122]. Last but not least PWWP domain-containing proteins interact with methylated lysines on the H3 or H4 tail and have an important role in chromatin-associated processes like DNA repair or transcriptional regulation (for more detailed information about PWWP domains see section 1.5) [123]. Altogether, the above-mentioned chromatin readers in concert with further PTM-mediated mechanisms define inaccessible (inactive) and accessible (active) chromatin states by modulating the global chromatin environment [113]. Shortly, inactive chromatin is characterized by modifications such as H3K27 trimethylation (H3K27me₃) in combination with Polycomb-repressive complexes 1 and 2 (PRC1 and PRC2) as well as H3K9 di- and trimethylation and H4K20me₃ together with heterochromatin binding protein 1 α (HP1 α) [124, 125]. Transcriptionally active chromatin on the other hand is associated with a plethora of posttranslational histone modifications like acetylations of multiple lysine residues on H3 and H4 and methylations on H3 lysine 4, 36 and 79 [124]. Thereby, tri-methylated H3 lysine 36 (H3K36me₃) is found within the body of actively transcribed genes [126]. The methylation of mono- and di-methylated H3 lysine 36 (H3K36me₁ and H3K36me₂) is catalyzed by the nuclear SET domain (NSD)-containing enzymes NSD1, NSD2 and NSD3, whereby H3K36me₂ is further established by methyltransferases such as ASH1L [127, 128]. On the other hand, SET domain-containing 2 (SETD2)/huntingtin interacting protein B (HYPB) is responsible for the trimethylation of H3K36 [129, 130] and is recruited via hyperphosphorylated RNA polymerase II (RNAPII) [131]. H3K36me₃ mainly serves as binding platform [132], especially via the direct interaction with PWWP-domain containing proteins such as mismatch recognition enzyme hMutS α (hMSH2-hMSH6) [133] or developmental gene regulatory PC4 and SF2 interacting protein 1 (Pcip1) [134]. Via the recruitment of binding proteins to gene bodies, H3K36me₃ is able to regulate a plethora of biological mechanisms including transcriptional elongation, repression of cryptic transcripts, chromatin accessibility and DNA damage response signaling [127, 129, 133, 135].

All in all, PTMs and their interplay with additional chromatin components is crucial for the control of cellular processes [136]. Therefore it is important to further study the functional

consequences of combinatorial PTMs and their crosstalks to provide a better understanding of these, almost all biological processes underlying, regulatory mechanisms [136, 137].

1.2.2 Histone variants

In addition to PTMs, the exchange of the canonical core histones H2A, H2B, H3 and the linker histone H1 (to date no H4 variants have been found in humans and other higher eukaryotes) with specific histone variants modifies the nucleosomal composition and can further regulate the global chromatin landscape [138-140]. In general, the incorporation of histone variants creates specialized nucleosomes, enhances chromatin plasticity and complexity and can result in profound alterations of the chromatin structure thereby influencing multiple biological processes [12]. Canonical histones and the according variants differ in some basic aspects. First, the multiple canonical histone genes are expressed during S-phase whereas the expression of most histone variant genes takes during the whole cell cycle place [141, 142]. Second, canonical histone genes lack introns and give rise to a unique 3' stem-loop mRNA structure [141, 143]. By contrast, pre-mRNAs of most histone variants often contain introns thus providing the opportunity to generate alternative splice isoforms [12, 141]. Furthermore, some histone variant mRNAs have a poly-adenylated (poly(A)) tail [142], whereas canonical histone mRNAs miss this specific pattern [12]. Last, histone variants are non-allelic isoforms of their canonical counterparts and differ markedly from their primary sequences and sometimes additionally in secondary and/or tertiary conformations [141]. These differences in protein sequence can be either minimal (few individual amino acid substitutions) or far more tremendous (exchange of entire domains) [142, 144]. Concluding, histone proteins possess distinct structural and functional properties, hence it is imperative to study the composition of canonical histones together with histone variants in nucleosomes to gain a better understanding of their complex interplay [12].

Histone variants and their canonical counterparts do not only differ in expression timing, mRNA structure and at protein levels, but also in their distribution within the linear genome [141]. In contrast to canonical histones, which are rather equally distributed, histone variant incorporation and eviction is highly specific and dependent on remodeling and chaperone complexes [141, 145]. Histone chaperones are crucial for several steps in nucleosome formation [146] and bind in a direct manner to newly synthesized histones to regulate their degradation or stability [147]. In general, some chaperones have important roles in histone

INTRODUCTION

trafficking (after their synthesis), where they regulate their transport from cytoplasm to the cell nucleus [148]. Others affect PTMs by regulating the interaction with the corresponding enzymes [143] or promote the formation of nucleosomes by facilitating histone interactions [149, 150]. On the other hand, ATP-dependent remodelers eject or slide already assembled nucleosomes thus organizing chromatin structure and promoting the exchange of H2A-H2B dimers with distinct variant dimers [151].

More than 80% of all nucleosomal linker DNA binds at least one protein of the histone H1 family, which represents the largest group of histone variants with its four germ-line and seven somatic (mammalian) specific variants [152, 153]. Besides their high primary sequence homology and evolutionary conservation, H1 variants differ in some aspects including euchromatin and heterochromatin distribution or chromatin binding affinity [153-156]. In addition they play an important role in chromatin compaction and transcriptional regulation [157].

To date, solely three, functional uncharacterized, *bona fide* H2B variants (testis-specific hTSH2B, H2BFWT and spH2B) are known in the human genome, although more than 15 H2B genes have been identified, yet [158]. In 2012, Santoro and Dulac identified in mice an additional, fourth H2B variant – the olfactory, neuron-specific H2BE [159]. hTSH2B differs in its amino acid sequence from its canonical counterpart H2B mainly in N-terminal regions [158], but its biological function still remains elusive [160]. The primate-specific variant H2BFWT shares just 45% homology with canonical H2B, but was found to act similar in a nucleosomal background [161]. Further, it is hypothesized to act as an epigenetic marker modulating telomeric identity [158]. spH2B has been identified in a telomere-associated complex and varies from somatic H2B in biochemical parameters including electrophoretic mobility [162]. The newly identified neuron-specific H2BE variant is incorporated into H2B-containing chromatin and has a direct effect on the expression level of genes and the persistence of olfactory sensory neurons [159].

Currently, two canonical H3 (H3.1 and H3.2) and a plethora of H3 variants (H3.3, H3t, H3.5, H3.X, H3.Y, CENP-A (CenH3), H3.6, H3.7 and H3.8) are known [138, 163-166]. H3.3 is encoded by two independent human genes (*H3F3A* and *H3F3B* [167]) and differs from its canonical counterpart H3.2 in four amino acids [168]. Remarkably, these differences in primary sequence lead to interactions of H3.3 with distinct chaperones and remodelers compared to

INTRODUCTION

H3.1 and H3.2 [169, 170]. Whereas H3.1 and H3.2 deposition is catalyzed by the histone chaperone CAF1 (Chromatin-Assembly Factor1) [170], H3.3 is incorporated either by the chaperone complex HIRA at actively transcribed gene regions [171] or by the death-domain associated protein/ α Thalassemia/Mental Retardation X-linked (DAXX/ATRX) chaperone complex at repetitive heterochromatic regions. Interestingly, the N-terminal tail of H3.3 harbors a unique amino acid – S31 – that is phosphorylated close to centromeric regions in metaphase HeLa cells [172]. H3t and H3.5 are both testis-specific [164], whereby H3t plays an important role in chromatin reorganization and early spermatogenesis [166, 173] and H3.5 may influence DNA synthesis [174]. The primate-specific histone H3 variant H3.X was identified together with H3.Y in our lab [175]. Both variants are expressed in different parts of the brain and are stably incorporated into the nucleosome. Interestingly, external stress stimuli like starvation and overgrowth of cells can enhance the expression of H3.Y in osteosarcoma-derived cancer cells. H3.Y regulates cell cycle-associated gene expression and cell proliferation [175]. The kinetochore- and centromere-associated histone variant protein CENP-A serves as binding platform for the assembly of kinetochores during cell division and regulates chromosomal stability [176, 177]. Furthermore, the overexpression and mislocalization of CENP-A results in higher invasiveness and bad prognosis in several cancer types [177]. The recently identified H3 variants H3.6, H3.7 and H3.8 are known to exist in many human tissues and were so far just studied in *in vitro* nucleosome assemblies without pointing towards any specific biological function [165].

1.2.3 H2A variants

In 1977, two different mammalian isoforms of canonical H2A (H2A.1 and H2A.2) differing in amino acid position 51 were discovered [178] and over the last decades a plethora of histone H2A variants could be added to the growing list of human H2A variants [12]. In addition to canonical H2A, eight further variants exist in humans: the universal variants H2A.X and H2A.Z (H2A.Z.1, H2A.Z.2.1 and H2A.Z.2.2) as well as H2A.Bbd and macroH2A (macroH2A1.1, macroH2A1.2 and macroH2A2). Generally, the highest variation of H2A variants can be observed within their C-termini, ranging from differences in amino acid composition and alterations in C-terminal length.

In the 1980s, H2A.X was identified together with H2A.Z [179]. H2A.X is evolutionary highly conserved, but very variable in its cell levels (2-25% of the H2A pool in mammals [180]) and

INTRODUCTION

deposited throughout the genome by a yet unknown deposition machinery [181]. H2A.X is stated as genomic “histone guardian” [180, 182] and is famous for its well-studied modification on S139 (S139ph), which is established upon DNA damage [180, 182]. In its phosphorylated state, H2A.X is also called γ -H2A.X, because it was initially discovered after treating mammalian cells with gamma irradiation [180, 183]. γ -H2A.X serves at DNA damage sites as recruiting platform for repair proteins, chromatin remodelers and chromatin-modifying enzymes [181]. In turn, γ -H2A.X is stabilized at DNA lesions via the methylation of lysine residues such as H3K9, H3K27 and H3K36 [184]. However, in addition to its role in double-strand-break repair (DSB), γ -H2A.X may help to keep broken DNA ends together through accumulating specialized repair proteins [185]. Due to its important roles in DNA damage repair (DDR) it is not quite surprising that mutations in or deletions of H2A.X have been detected in many tumors [181].

One of the most rapid evolving histone variants is testis-specific H2A.Bbd (Barr body deficient) [186, 187], which is mammalian-specific [186, 188] and shares on the protein level only 50% identity with canonical H2A [189]. Strikingly, H2A.Bbd lacks, due to its shortened primary sequence, huge parts of the docking domain, the C-terminal tail and the acidic patch that is needed for nucleosome-protein interactions [186]. The consequence of these missing structures is therefore a reduction in nucleosome stability and a more opened chromatin structure [186, 190-192]. Additionally, H2A.Bbd incorporation has similar structural effects as shown for highly dynamic acetylated nucleosomes [186] and could be an alternative, acetylation-independent mechanism to loosen chromatin structure.

MacroH2A was discovered in 1992 [193] and is encoded by two different genes: *H2AFY* (macroH2A1) and *H2AFY2* (macroH2A2) [139]. *H2AFY* can be alternatively spliced and gives rise to two protein isoforms: macroH2A.1.1 and macroH2A.1.2 [194]. MacroH2A members are fascinating proteins due to their exceptional tripartite domain architecture: the N-terminal histone region is linked to a non-histone macro domain through a H1-like linker leading to a unique, three-fold larger protein structure than H2A [193]. Interestingly, macroH2A.1.1 but not macroH2A.1.2, can interact with specific metabolites of nicotinamide adenine dinucleotide (NAD) including poly(ADP-ribose) [195-197] suggesting a unique function for macroH2A.1.1 [12]. However, several studies propose a broad range of biological roles for macroH2A such as epigenome establishment and maintenance, embryonic and adult stem cell

INTRODUCTION

differentiation, somatic cell reprogramming and tumor suppression, depending on isoform and chromatin location [12, 195, 198].

Since H2A.Z is one of the main actors in this work; its function and structure will be discussed in more detail in section 1.3.

1.3 The histone variant H2A.Z

H2A.Z, an almost universal variant [12], was first described in the 1980s [179] and is known to have evolved early during evolution [199]. H2A.Z shares only ~60% identity with canonical H2A while it is ~90% conserved between different eukaryotes pointing towards specific, not exchangeable functions within the H2A family [200] (**Figure 1.3.1**). Indeed, H2A.Z is crucial in many organisms like *X. laevis* (frog) [201], *M. musculus* (mouse) [202], *T. thermophile* (tetrahymena) [203] and *D. melanogaster* (fruit fly) [204] but not in *S. pombe* (fission yeast) and *S. cerevisiae* (budding yeast) that show severe growth phenotypes upon knock-out [205, 206].

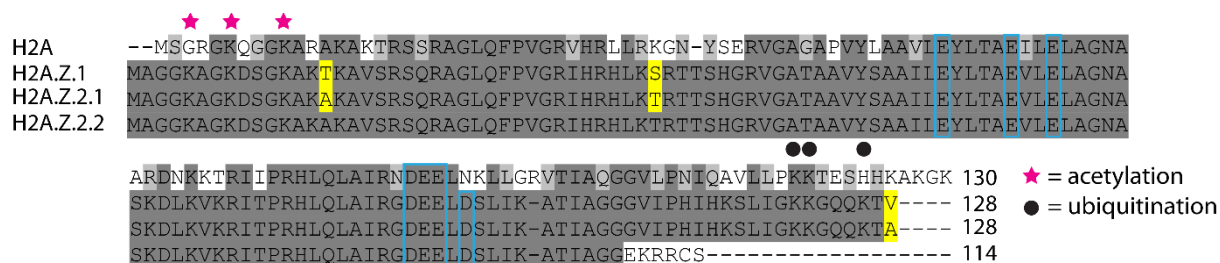


Figure 1.3.1 H2A and H2A.Z isoforms amino acid sequence alignment. Amino acid sequence alignment of human H2A and human H2A.Z isoforms (H2A.Z.1, H2A.Z.2.1 and H2A.Z.2.2). Amino acid sequences of H2A [P0C0S8], H2A.Z.1 [P0C0S5], H2A.Z.2.1 [Q71UI9-1] and H2A.Z.2.2 [Q71UI9-2] were aligned using the web browser-based MUSCLE multiple sequence alignment tool (default settings). Identical amino acids are depicted in dark grey, similar amino acid residues in light grey and changes are highlighted in white. The three amino acids differing between H2A.Z.1 and H2A.Z.2.1 are illustrated in yellow and acidic patch residues are boxed in cyan. Possible modification sites of H2A.Z variants are marked with stars (acetylation, pink) and dots (ubiquitination, black). Figure based on [12].

In vertebrates, H2A.Z is encoded by two non-allelic, intron-containing genes: *H2AFZ* (H2A.Z.1) and *H2AFV* (H2A.Z.2), which both give rise to polyadenylated mRNAs and to H2A.Z proteins that differ in just three amino acids [12, 207] (**Figure 1.3.1**). Notably, *H2AFV* can be alternatively spliced resulting in the expression of two subforms: H2A.Z.2.1 (H2A.Z.2, vertebrate-specific) and H2A.Z.2.2 (primate-specific) [12, 190]. Especially in drosophila, the

INTRODUCTION

essential parts for H2A.Z function lay in its C-terminal region (M6 and M7) [208]. In concordance, also in yeast (*S. cerevisiae*) M6 fulfills important functions and is e.g. essential for the interaction with the Swi2/Snf2-related ATPase-1 (SWR-1) remodeling complex [209] or H2A.Z deposition [210]. In addition, the discovery of H2A.Z.2.2, which in its C-terminus lacks 14 amino acids and harbors a unique stretch of six amino acids (**Figure 1.3.1**), provided further fundamental insights in H2A.Z's C-terminal functions [190, 211]. Compared to other H2A.Z members, the shorter H2A.Z.2.2 isoform remains soluble and is not completely associated with chromatin and less tightly incorporated in nucleosomes [190, 211].

At the turn of the millennium, the crystal structure of the H2A.Z-containing nucleosome was solved and revealed, despite the huge differences in primary sequence, a highly structural similarity to canonical H2A nucleosomes [212]. This structure also unraveled alterations in two regions that might be responsible for contrary observations in nucleosome stability. Indeed, many studies addressed this issue with contrasting results: H2A.Z seems to have both, stabilizing and destabilizing effects [12, 200]. Suto et al. reported, that H2A.Z's L1-L1 interface forms more contacts between H2A.Z- than H2A-dimers resulting in stabilized H2A.Z-homodimers [212], but also the presence of heterotypic H2A-H2A.Z nucleosomes could be already shown in different model organisms [213-217]. Moreover, the destabilizing effects of H2A.Z were also observed when it teams up with histone H3 variant H3.3 in the same nucleosome [218]. Under physiological salt conditions it was shown that nucleosomes prepared from native chromatin having both variants incorporated are less stable than those containing only H3.3 [219]. Surprisingly, in contrast to its destabilizing function observed in mononucleosomes, H2A.Z incorporation is crucial for the formation of higher chromatin structures [220]. Compared to canonical H2A, H2A.Z has an extended acidic patch with an increased affinity to the H4 tail [221]. This strong electrostatic interaction with H4 results in an increased intrafibre folding and a higher compaction of secondary structures [221, 222]. Further, HP1 α prefers highly folded H2A.Z- over H2A-containing fibers and enhances intrafibre folding [223], highlighting the function of histone variants in the establishment of chromatin structures by recruiting further chromatin structure modifying proteins [12].

Like other histone variants, H2A.Z can be posttranslationally modified resulting in different functional consequences [224, 225]. Mammalian H2A.Z can be chemically altered by four distinct mechanisms: acetylation, monoubiquitylation, SUMOylation and monomethylation at specific lysines [12, 226]. In general, H2A.Z can be acetylated on lysine 4 (H2A.ZK4ac), 7

INTRODUCTION

(H2A.ZK7ac) and 11 (H2A.ZK11ac) (**Figure 1.3.1**) and is reported to be associated with promoter regions of active genes [227, 228], confers to an open nucleosome conformation and has destabilizing effects [229]. In addition, the acetylation of H2A.Z at promoter regions is suggested to activate oncogenes thus deregulating gene expression and foster tumor progression [226]. In addition to N-terminal acetylations, H2A.Z can be further modified by C-terminal monoubiquitinations on lysine 120, 121 and 126 that are associated with inactive X chromosomes in female mammals [230]. H2A.Z SUMOylation in turn has been implicated to be involved in DNA repair in yeast (*S. cerevisiae*) [231] and monomethylation of H2A.ZK7 was suggested to play a role in transcriptional repression in embryonic stem cells (*M. musculus*) [232].

The deposition or/and eviction of H2A.Z is dependent on multifactor ATP-dependent chromatin remodeling complexes, whereby genome-wide studies helped to unravel possible target sites [233]. In *S. cerevisiae*, the 14 subunit SWI/SNF-related SWR1 complex (SWR1) is crucial for the incorporation of H2A.Z-H2B dimers into chromatin [234-236]. To do so, SWR1 deposits H2A.Z-H2B into nucleosomes that flank the nucleosome free region (NFR) of transcription start sites (TSS) and well distributes it all over the genome (H2A.Z is present in 5% of nucleosomes but 65% of all genes) [237]. In higher eukaryotes like humans, the SWR1 complex has two counterparts named p400/NuA4/TIP60 and SRCAP (SWI2/SNF2-related CBP activator protein) [233, 238, 239] where each of them harbors (working context dependent) histone replacement abilities *in vivo* and *in vitro* [233]. Additionally, Papamichos-Chronakis et al. discovered that a specific remodeling complex – INO80 – facilitates the eviction of H2A.Z in an ATP-dependent manner after transcriptional induction in *S. cerevisiae* [233, 240], but it still remains elusive if the mammalian INO80 complex may have a similar role in higher eukaryotes [241]. However, Mao and co-workers could identify Anp32e as a mammalian, H2A.Z-specific chaperone facilitating the eviction of H2A.Z at transcription start sites [242].

Interestingly, H2A.Z is not uniformly distributed all over the genome suggesting three non-mutually exclusive mechanisms [12]. First, H2A.Z is undirected incorporated and afterwards actively evicted from non-target sites; second, H2A.Z is incorporated in an active manner at specific sites by special target factors [243] and third, the localization of H2A.Z is dependent on deviating stabilities of homotypic/heterotypic H2A.Z-containing compared with canonical nucleosomes [216]. Regardless how H2A.Z is incorporated into chromatin, the variant is crucial for a plethora of biological functions like transcription regulation, DNA repair, cell cycle

INTRODUCTION

progression, genome stability, telomere integrity, chromosome segregation, mitosis, brain memory formation, heterochromatin formation and just recently discovered epithelial-mesenchymal transition [12, 181, 200, 225, 244-246].

In 1986, Allis et al. initially suggested that H2A.Z has a role in transcription regulation [247]. The invention of genome-wide assays like chromatin immunoprecipitation sequencing (ChIP-seq) enabled to further dissect the function of H2A.Z in transcription. It could be elucidated that H2A.Z is present around the TSS at gene promoters in human [248, 249], fly [250], plants [251], yeast [252, 253] and mouse [254] and further regulatory elements like insulators or enhancers [200]. Remarkably, H2A.Z is enriched at both, transcriptionally active and silenced genes whereby the context-dependent underlying mechanism has remained elusive [255]. In human, mouse and yeast the -1 and +1 nucleosomes (nucleosomes that flank the NDR) at TSS are occupied by H2A.Z [200]. But why is the incorporation of H2A.Z at TSS so important for transcriptional regulation? *In vitro* studies have demonstrated that the +1 nucleosome forms a significant barrier to transcription [256, 257] by preventing the transit of RNA polymerase II (RNAPII) [258]. Notably, the incorporation of H2A.Z leads to decreased RNAPII stalling, hinting towards a function in dissolving the high-energy barrier [255]. The underlying mechanism might be the labialization of the +1 nucleosome by H2A.Z incorporation, thereby modulating the kinetics and transcriptional elongation of RNAPII and passively regulating transcription [255, 258].

Moreover, current literature demonstrates that H2A.Z plays controversial roles in DNA damage response. Studies in yeast showed increased genomic instability and sensitivity to DNA-damage upon H2A.Z knock-out [240, 259] and suggest an important role for H2A.Z in DNA damage response [231, 260]. Furthermore, Xu et al. reports H2A.Z is actively incorporated at DSB by the p400/NuA4/TIP60 complex in an ATPase-dependent manner [260], thereby influencing chromatin structure and recruiting further, essential DSB repair proteins to sites of DNA damage. By contrast, the Price group demonstrated that H2A.Z is removed by histone chaperone Anp32e at DNA double strand breaks, which promotes DNA repair and nucleosome reorganization [261].

Besides transcription regulation and DNA repair, H2A.Z also affects cell cycle progression [262]. The directed expression of genes in yeast during the cell cycle requires chromatin remodeling (Swi/Snf) and modifying (SAGA) complexes and the presence of unique histone variants like H2A.Z [263]. Thereby, H2A.Z is crucial for the induction of cell cycle-regulatory

INTRODUCTION

proteins during S phase and the regulation of cyclin genes [263] like cyclin D1, that is involved in G1-S progression and a plethora of biological functions [264].

Compared to the large number of biological functions that H2A.Z performs within the cell, it is not surprising that even the slightest changes in H2A.Z levels can lead to different types of cancer [181]. A number of microarray studies, aiming to identify different gene expression patterns in human cancer, revealed increased H2A.Z levels in different tumor types like colorectal, lung, bladder and breast [265-268]. Further, global H2A.Z levels are upregulated in melanoma cell lines [269] and a direct role in cancer could be detected in breast and prostate cancer [181]. However, H2A.Z.2 is a key driver of malignant melanoma and correlates with a bad prognosis for patients, whereas the absence of H2A.Z.2 sensitizes melanoma cells to targeted chemotherapies [270]. In detail, H2A.Z.2 regulates the transcription of highly expressed E2F target genes, which show a distinct pattern of H2A.Z occupancy. The unique function of H2A.Z.2 is further the promotion and maintenance of E2F1, BRD2 and histone acetylation levels to promote transcription of cell cycle regulating factors. In fact, BRD2 is crucial for the cell cycle and controls together with E2F1 the expression of cyclins and further regulatory genes [271-273]. In addition, BRD2 may recruit distinct transcription factors to initiate oncogenic gene transcription in several cancer types [270]. To date, metastatic melanoma cannot be effectively treated by current therapies [274], but the combinatorial depletion of H2A.Z.2 and BRD2 (BET protein) is an effective tool in inducing cell death and sensitizing cells to distinct therapies [270].

Summarizing, H2A.Z has a variety of biological functions and is involved in the development of different types of cancer - therefore it is very important to study the biological mechanism of this variant in order to gain a better fundamental understanding and to develop possible therapies against different diseases.

1.4 H2A.Z binding proteins

To gain deeper insights in the underlying mechanisms of H2A.Z's multiple biological functions, we and others analyzed the interactome of the histone variant in more detail [275, 276]. Therefore, Draker et al. used human embryonic kidney cells (293T) and transfected them with either FLAG-tagged H2A.Z or H2A (control) [275] and analyzed co-purified proteins by applying liquid chromatography–mass spectrometry/mass spectrometry (LC-MS/MS). There, Draker

INTRODUCTION

and co-workers found 21 top interactors specifically enriched for H2A.Z, where most of them are associated with chromatin and are involved in transcription and chromatin organization. In general, they identified members of the H2A.Z remodeling complex SRCAP and protein readers like PHD finger-containing or PWWP domain-containing proteins. Also our group aimed to identify the interactome specifically enriched on H2A.Z compared to H2A. Therefore, Sebastian Pünzeler (a former PhD student of our lab) together with the group of Matthias Mann (MPI Munich) performed label-free quantitative mass spectrometry (lf-qMS) of mononucleosomes derived from human HeLa Kyoto (HK) and metastatic melanoma SK-mel147 cell lines (stably expressing GFP-tagged H2A or H2A.Z) [276]. There, more than 40 H2A.Z-specific binding proteins were detected (**Figure 1.4.1**). In line with the findings of Draker and co-workers [275], we also identified members of the chaperone complex SRCAP, chromatin-modifying complexes, PHD finger-containing and PWWP-domain-containing proteins [276], but also subunits of the p400/TIP60 chaperone complex and many additional chromatin binders. Interestingly, in comparison to SRCAP we could only enrich the two largest subunits of the p400/TIP60 complex suggesting rather a transient interaction with H2A.Z e.g. while removing together with Anp32e the variant upon DNA damage [261]. Moreover, the detection of histone-modifying complexes such as mixed lineage leukemia (MLL), which was initially identified in human leukemia [277-280], is a target of chromosomal translocation [281] and catalyzes the histone modification H3K4me3 [282], supports the idea that H2A.Z acts as a recruitment platform for distinct proteins [282]. In concordance with that hypothesis, we additionally identified histone-modifying enzymes like the demethylases PHF2 and KDM2A [276]. PHF2 harbors an N-terminal plant homeodomain (PHD) binding to H3K4me2/3 and a Jumonji (JmjC) domain, involved in removing repressive methyl marks (H3K9me1) [283, 284]. KDM2A also harbors a JmjC domain and demethylates mono- and dimethylated Lys 36 of histone H3 (H3K36me1/2) but not trimethylated Lys 36 [285, 286], that only serves as an inactive substrate [287].

INTRODUCTION

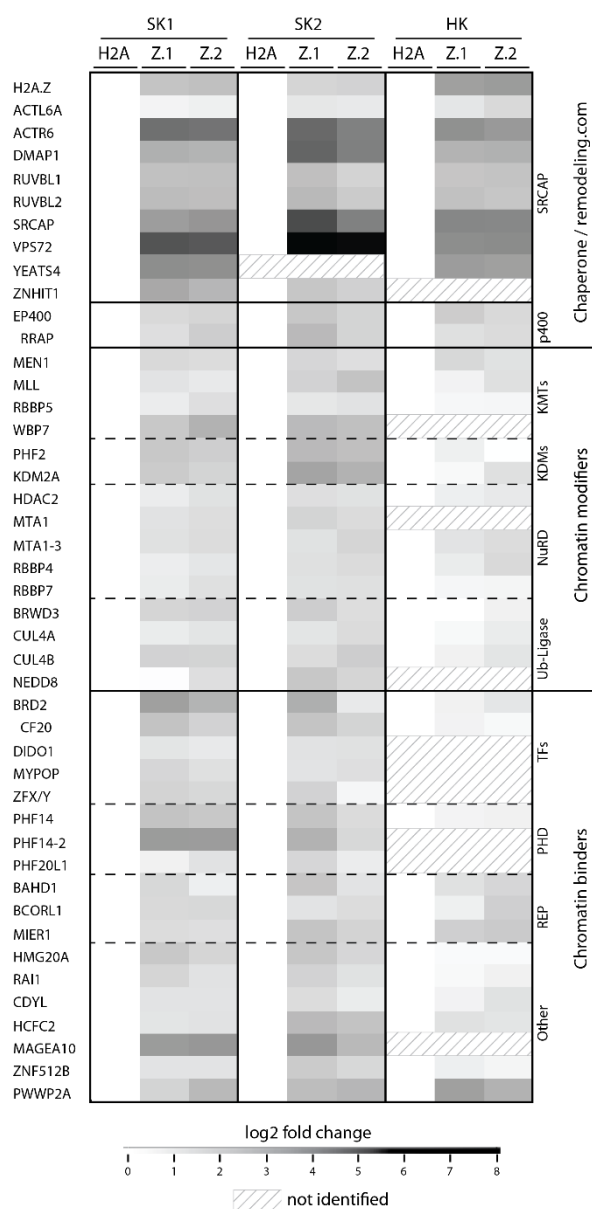


Figure 1.4.1 Heatmap of H2A.Z.1- and H2A.Z.2-nucleosome-specific binding proteins. Pull-down experiments with mononucleosomes derived from HK (HK) and SK-mel147 cells (SK1 and SK2) (each stably expressing GFP-H2A, GFP-H2A.Z.1 (Z.1) and GFP-H2A.Z.2 (Z.2)) were analyzed by label-free qMS-based proteomics and normalized to H2A. SK1 and SK2 represent two biological replicates and were reproduced in HK cells. Log₂-fold change was hierarchically clustered and plotted as heatmap (Perseus). Log₂ LFQ intensities (fold-change for GFP-H2A.Z.1, GFP-H2A.Z.2 vs GFP_H2A pull-downs) are illustrated for 44 identified proteins specifically enriched for GFP-H2A.Z. White color: no enrichment; black color: strong enrichment in comparison to GFP-H2A. Figure reprinted from [276], with permission of John Wiley and Sons.

Our lab decided to further investigate the vertebrate-specific, PWWP-domain containing protein PWWP2A, which harbors a unique domain structure and was already discovered in previous H2A.Z-pull-down studies [275]. PWWP2A has a molecular mass of approximately 82 kDa and consists of 755 amino acids (aa). The UniProt and NCBI databases predict one

INTRODUCTION

canonical PWWP2A protein and two additional isoforms with most differences located within their C-terminus. PWWP2A consists of two proline-rich regions “P1” (aa 61 - 146) and “P2” (aa 240 - 291), an internal region “I” [276] (aa 292 - 574) and one serine-rich stretch “S” (aa 575 - 632). In addition, PWWP2A belongs to the PWWP domain family of proteins, as it harbors a C-terminal PWWP domain “PWWP” (aa 655 - 715) (**Figure 1.4.2**).

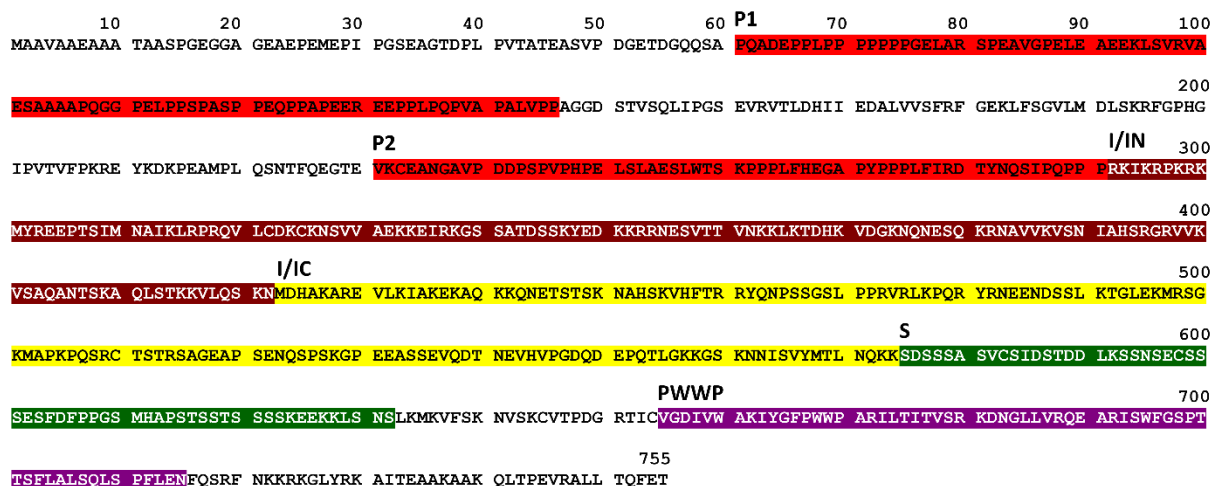


Figure 1.4.2 Primary structure of human PWWP2A. PWWP2A consists of 755 amino acids and possesses two proline-rich regions P1 (aa 61 - 146) and P2 (aa 240 - 291) (red), an internal region I (aa 292 - 574), which can be further subdivided into IN (brown) and IC (yellow), one serine-rich stretch S (aa 575 - 632) (green) and a C-terminal PWWP domain PWWP (aa 655 - 715) (purple). NCBI reference sequence NP_001124336.1 (NM_001130864.1) or Uniprot identifier Q96N64-1.

PWWP2A locates together with H2A.Z at +1 and -1 nucleosomes at the TSS of highly transcribed genes [276]. Moreover, PWWP2A does not contain any known enzymatic domains and might therefore act as an adapter between chromatin-modifying proteins and H2A.Z-containing chromatin, which is in line with the finding that it is not required for H2A.Z occupancy and deposition. Like described above, PWWP2A recognizes H2A.Z-containing nucleosomes at the TSS of actively transcribed genes, but is depleted at non-promoter sites. However, the exact underlying positioning mechanism remains elusive. Further, the depletion of PWWP2A leads to the deregulation of gene expression programs and has profound consequences on cellular mechanisms. In HK cells the knock-down of PWWP2A leads to an increase in mitotic cells, especially for those in prometaphase. PWWP2A-depleted cells stay for up to 10 h in mitosis (control cells <1.5 h) while jumping back and forth between two states (prometaphase and metaphase) and cannot arrange all chromosomes at the equatorial plate.

INTRODUCTION

Furthermore, the depletion of PWWP2A results in developmental defects in frogs (*X. laevis* and *X. tropicalis*). During the developmental stage of the tadpole, the loss of PWWP2A affected head differentiation and severely impaired proper eye development. These problems upon PWWP2A depletion in frogs are caused by its fundamental role in neural crest cell differentiation and/or migration during early developmental stages. Nevertheless, the strong interaction of PWWP2A with H2A.Z-containing chromatin, thus initiating multiple biological functions and processes, is likely achieved by its multi-domainous structure and is the main topic of this thesis.

The special and unique features of PWWP domains are explained in more detail in the next chapter.

1.5 PWWP domain-containing proteins

The PWWP domain was named after the structural motif “proline (P) – tryptophan (W) – tryptophan (W) – proline (P)” (90-130 amino acids) that was first discovered in the PWWP-domain containing Wolf-Hirschhorn syndrome candidate protein 1 (WHSC1) [288] and is involved in chromatin binding. Thus far, more than 20 different human PWWP domain-containing proteins have been identified [123]. Despite its high conservation, some variations of the PWWP-motif can occur like SWWP in DNMT3A/B [289] or PHWP in HDGF [290]. However, PWWP domains that contain a proline residue as first amino acid show higher stability and are less prone to aggregation [291]. In general, PWWP domains belong together with Tudor, plant Agenet, chromo and MBT domains to the Tudor domain “Royal Family” [292], that divergently developed during evolution but share the same ancestor, structurally characterized by three β -strands [293]. To date, several studies described the three-dimensional structures of PWWP domains including those of ZMYND11, PSIP1, DNMT3A, DNMT3B or BRPF1 [294-298]. Interestingly, all these PWWP domains fold in a similar manner: an N-terminal β -barrel and a C-terminal α -helical substructure [289]. The N-terminally located β -barrel is conserved and consists of five antiparallel β -strands (β 1 – β 5), whereby the PWWP motif is located at the interface of the β 1 – β 2 arch [289]. In contrast, the C-terminal α -helices (two to six) are highly variable [289, 298]. In general, the stability of the PWWP domain structure is based on polar bonds and intramolecular hydrogen interactions, which occur between residues in the same area and different substructures [289, 296, 298-301]. Due to

INTRODUCTION

their structural features, PWWP domains fulfill distinct tasks such as protein-protein interactions, DNA binding or the recognition of posttranslational histone modifications [298]. Indeed, the PWWP domain of DNMT3A directly interacts with the inhibitory factor SALL3 thereby reducing DNMT3A-mediated CpG island methylation [302] and mutations in the PWWP domain of DNMT3B result in a reduced interaction with SUMO E3 ligase PIAS1 [303]. Another member of PWWP domain-containing protein family is the transcriptional coactivator PSIP1 (PC4 and SFRS1 interacting protein 1) that was shown to interact via its PWWP domain with DNA and chromatin, crucial for HIV-1 integration [304, 305]. Likewise, the eukaryotic mismatch repair protein MSH6 is able to bind non-specifically to double stranded DNA with its N-terminal PWWP domain [306]. However, for most PWWP domain-containing proteins the interaction with DNA takes not place in a specific or sequence selective manner but rather via electrostatic contacts between the phosphate backbone of DNA and the highly positively charged surface formed by arginine and lysine residues in the PWWP motif [293] (**Figure 1.5.1**).

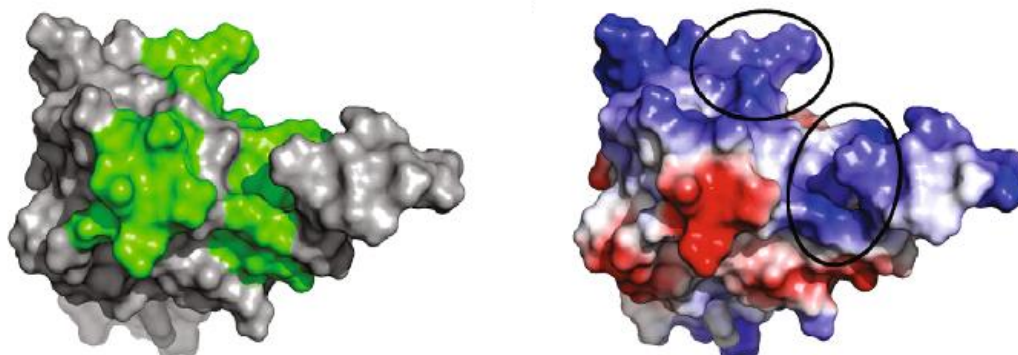


Figure 1.5.1 DNA binding interface of HDGF PWWP domain. Residues, directly involved in the formation of the DNA interaction surface are highlighted in green (left). Electrostatic distribution of the HDGF PWWP surface structure demonstrates an accumulation of highly positively charged residues (blue). Figure reprinted from [293], with permission of Springer Nature.

In addition, the similarities within structure and sequence of PWWP domains and other “Royal family” member domains suggest a possible role in the recognition of posttranslational modifications of histones [293]. Indeed, in 2009 the first PWWP domain-containing protein being able to recognize H4K20me1 was discovered [307] with many further following [293]. Furthermore, a high-throughput mass spectrometry approach unraveled PWWPs ability as a histone H3 lysine 36 trimethylation interaction partner [308], hinting towards a functional role

INTRODUCTION

as a putative H3K36me3 sensor [293]. The interaction with posttranslational histone modification is established by three conserved aromatic residues [294-298, 300, 309] forming an aromatic cage in which the trimethyl ammonium group fits in perfectly [293]. Altogether, PWWP domains harbor several contiguous interfaces for binding histones in general and DNA and posttranslational histone modifications in particular [295, 300] thus harboring multiple binding modes with chromatin.

1.6 Objectives

In my PhD thesis, I aimed to gain a deeper insight in the multivalent chromatin binding modes of PWWP2A, a novel H2A.Z interaction protein we recently identified via label-free quantitative mass spectrometry (lf-qMS) [276]. I applied *in vitro* binding studies such as competitive Electrophoretic Mobility Shift Assays (cEMSAs) to investigate the properties of the PWWP2A internal region, its eponymous C-terminal PWWP domain and their combinatorial nucleosome binding effects. In detail, I determined how the N-terminal and C-terminal parts of the internal region are able to specifically recognize chromatin in general and H2A.Z-containing mononucleosomes in particular. Furthermore, I analyzed the binding features of PWWP2A's PWWP domain in terms of nucleic acid interaction and its ability to bind to recombinant as well as cell-derived human mononucleosomes. Moreover, I determined whether the PWWP domain is able to specifically recognize certain posttranslational histone modifications. Here, I performed pull-down experiments with recombinantly expressed PWWP2A deletion proteins followed by label-free proteomics. Further, I tried to unravel whether PWWP2A is additionally able to interact with other members of the H2A histone variant family, underlining its unique and multivalent binding features.

2 Material and Methods

2.1 Materials

2.1.1 Technical devices

Description	Supplier
-20°C Freezer	Beko, Liebherr
-80°C Freezer	Thermo Scientific
4°C Fridge	Liebherr
37°C Incubator (bacteria)	Binder
37°C Incubator (mammalian cells)	New Brunswick
Äkta Pure	GE Healthcare
Agarose Gel Chamber	neoLab
Autoclave (PX-150)	Systec
CASY Cell Counter	Innovatis
Centrifuges	Beckman XPN-80 (SW-41 rotor) Beckman JXN-26 (JA-10, JA-25.50 rotors) Eppendorf 5424R Eppendorf 5810R Heraeus Cryofuge 6000i Thermo Scientific Pico17 Thermo Scientific X3R
ChemiDoc System	BioRad
Developer Machine Curix 60	AGFA
Gel Documentation System	Peqlab
Hood	Berner
Gradient Master	Biocomp
Incubation Shaker	Eppendorf, Infors
Licor	Odyssey

MATERIAL AND METHODS

Microscope	Leica DMIL LED
Microwave	Severin
Magnetic Stirrer	Heidolph, IKA
MilliQ-Water System	Millipore
PCR Cycler	PE applied Biosystem
pH-Meter	Mettler Toledo
PhosphorImager FLA3000	Fuji
Pipetboy Accu-Jet® pro	Brand
Pipettes	Gilson
Power Supply	BioRad
Protein Gel Chamber	Invitrogen, SERVA
Pump Minipuls	Gilson
Rotating Wheel	Neolab, Stuart, VWL
Scales	Kern, Sartorius
Scanner	Epson
Shaker	IKA, New Brunswick, Roth
Sonifier MD-250	Branson
Spectrophotometer (DS-11)	Denovix
Speed-Vac	Eppendorf
Typhoon FLA9500	GE Healthcare
Thermomixer	Eppendorf 5436 ThermoMixer C
Trans Blot SD Semi-dry Transfer Cell	BioRad
UV Spectrophotometer	Thermo Scientific
Vortex Genie 2	Bachofer

MATERIAL AND METHODS

2.1.2 Consumables

Description	Supplier
1.5 ml and 2 ml Reaction Tubes	Greiner, Sarstedt
1.5 ml Low Binding Tubes	Sarstedt
Amicon Ultra 0.5 ml, 4 ml, 15 ml	Millipore
Cling Film	Saran
Cassettes 1.0 mm	Life technologies
Cell Culture Plates	Sarstedt
Centrifuge Tubes Polyalloma	Beckman
Combi Tips Plus	Eppendorf
Filter Paper Whatman 3MM	Whatman
Filter Tips	Biozym, Gilson
Glass Pipettes 5 ml, 10 ml	Hirschmann
Glassware	Schott
GST Trap HP Column	GE healthcare
Inoculation Loops	Sarstedt
Laboratory Film	Parafilm
MaXtract High Density Column	Qiagen
Membrane filter nitrocellulose	Roth
Pasteur Pipettes	Brand
PCR Reaction Tubes	Greiner
Petridishes	Greiner
pH Indicator Stripes	Merck
Pipette Tips	Biozym, Greiner, Sarstedt
Protein Gel Cassettes	Invitrogen
Protein Native Gels (precast)	SERVA
Protein SDS Gels (precast)	SERVA
Protran Nitrocellulose Transfer Membrane	Whatman

MATERIAL AND METHODS

Slide-A-Lyzer Mini 7K 0.1 ml	Life technologies
Spectra Por (Dialysis membrane)	Roth
Sterican Needles	Braun
Syringes 3 ml, 5 ml, 10 ml	Braun
Water, PCR-grade	Roche
Water, RNase free	Ambion
X-ray Films	Fujifilm

2.1.3 Chemicals

Description	Supplier
Acetic Acid	Sigma
Agar (LB)	SERVA
Agarose	Bio & Sell
Ammonium-Bicarbonate (NH_4HCO_3)	Sigma
Ampicillin	Roth
Aprotinin	Genaxxon
APS	Roth
Bromophenol Blue	Sigma
BSA 98%	Sigma
Complete Protease Inhibitor Cocktail Tablets	Roche
Coomassie Brilliant Blue	Sigma
Developer	AGFA
Disodium-Hydrogen-Phosphate (Na_2HPO_4)	Sigma
DMSO	Sigma
DNA Oligonucleotides	Sigma
dNTP Mix	NEB
DTT	Roth
DMEM	Sigma

MATERIAL AND METHODS

ECL Western Blotting Detection Reagents	Amersham
EDTA	Sigma
EGTA	Sigma
Ethanol, absolute	VWR
Ethidiumbromide	Sigma
FCS Dialyzed	Sigma
Fixer	AGFA
Glutathione Sepharose Beads	GE healthcare
Glycerol	VWR
β -Glycerolphosphate	Sigma
Glycin	VWR
Guanidine	Sigma
HEPES	VWR, SERVA, Promega
IPTG	Roth
Isoamylalcohol	Merck
Leupeptin	Genaxxon
Magnesium Chloride	VWR
Methanol	Sigma
Milkpowder	Heirler Cenovis
Monopotassium Phosphate (KH_2HPO_4)	Sigma
NP-40	Sigma
OptiMEM	Life technologies
PEG8000	Promega
Pepstatin	Genaxxon
PMSF	Genaxxon, Sigma
Potassium Chloride	VWR, Sigma
Potassium Hydroxide	Sigma
Rotiphorese Acrylamide/Bisacrylamide Mix	Roth
Rotiphenol	Roth

MATERIAL AND METHODS

SDS	SERVA
Sodium Acetate	Merck
Sodium Chloride	SERVA
Sodium Hydroxide	Sigma
TBE (5x)	VWR
TEMED	Roth
TFA	Sigma
TSA	Sigma
Tris	Invitrogen
Triton-X-100	Sigma
Trypsin/EDTA (cell culture)	Sigma
Trypsin (MS)	Promega
Tween 20	Sigma
Urea	Life technologies
X-tremeGENE	Roche
Zeomycin	Invitrogen

2.1.4 Kits, enzymes and markers

Description	Supplier
100 bp DNA Marker	NEB
1 kb DNA Marker	NEB
Dpn1	NEB
Gel Extraction Kit	Qiagen
Mini-, Midi-, Maxi-, Gigaprep Kit	Macherey Nagel
Micrococcal Nuclease	Sigma
PCR Purification Kit	Macherey Nagel, Qiagen
peqGOLD Protein Marker IV, V	Peqlab
Phusion R DNA Polymerase	Finnzymes

MATERIAL AND METHODS

Pfu Turbo DNA Polymerase	Stratagene/Agilent
Taq DNA Polymerase	NEB

2.1.5 Antibodies

2.1.5.1 Primary antibodies

Table 2.1 Primary antibodies

Name	Supplier	Host	Catalogue number	Application	Dilution
α -FLAG	Sigma	mouse	F3165	WB	1:5.000
α -GST	E. Kremmer	rat	Klon 6c9 R-2A	WB	1:500
α -H2A	Abcam	rabbit	ab13923	WB	1:250
α -H2A	Active Motif	rabbit	39209	WB	1:1.000
α -H2A.X	Millipore	rabbit	07-627	WB	1:1.000
α -H2A.X S139ph	Millipore	rabbit	07-164	WB	1:1.000
α -H2A.Z	Abcam	rabbit	ab4174	WB	1:3.000
α -H3	Abcam	rabbit	ab1791	WB	1:30.000
α -H3	Active Motif	mouse	61475	WB	1:5.000
α -H3K4me3	Active Motif	mouse	61379	WB	1:500
α -H3K4me3	Diagenode	rabbit	C1541.0003	WB	1:1.000
α -H3K9me3	Diagenode	rabbit	C15410056	WB	1:1.000
α -H3K27me3	Diagenode	rabbit	C15410069	WB	1:1.000
α -H3K36me1	Active Motif	rabbit	61351	WB	1:1.000
α -H3K36me2	Active Motif	rabbit	39255	WB	1:1.000
α -H3K36me3	Active Motif	rabbit	61102	WB	1:4.000
α -H3K79me2	Abcam	rabbit	ab177184	WB	1:1.000
α -H3K79me3	Diagenode	rabbit	C15310068	WB	1:1.000
α -H4	Abcam	rabbit	ab10158	WB	1:1.000
α -H4K20me3	Abcam	rabbit	ab9053	WB	1:1.000
α -PWWP2A	Novus Biologicals	rabbit	NBP2-13833	WB	1:500

2.1.5.2 Secondary antibodies

Table 2.2 Secondary antibodies

Name	Supplier	Catalogue number	Application	Dilution
α -mouse HRP	VWR	NA931-1ml	WB	1:10.000
α -rabbit HRP	VWR	NA934-1ml	WB	1:10.000
α -rat HRP	VWR	NA935	WB	1:10.000
α -rabbit IRDye 800CW	LI-COR Biosciences	926-32211	WB	1:10.000
α -mouse IRDye 680RD	LI-COR Biosciences	926-68070	WB	1:10.000

MATERIAL AND METHODS

2.1.6 Plasmids

Table 2.3 Plasmids

Name	Source	Description	Resistance
pDONR	Life technologies	Gateway cloning	Zeo
pUC18 (147 bp DNA)	Müller-Planitz Lab	Expression of 147 bp DNA fragment	Amp
pUC18 (187 bp DNA)	Müller-Planitz Lab	Expression of 187 bp DNA fragment	Amp
pGEX6P1-GST	Amersham	Expression of GST in competent <i>E. coli</i>	Amp
pGEX6P1-GST -PWWP2A_FL -PWWP2A_P1 -PWWP2A_I -PWWP2A_IN -PWWP2A_IC -PWWP2A_I_S -PWWP2A_I_S_PWWP -PWWP2A_S_PWWP -PWWP2A_PWWP -PWWP2A_PWWP_short -PWWP2A_S_long -PWWP2A_S_short -PWWP2A_S_PWWP_F666A -PWWP2A_S_PWWP_W669A -PWWP2A_S_PWWP_W695A -PWWP2A_I_R461Q -PWWP2A_I_R475C	Hake Lab	Expression GST-tagged (N-terminal) PWWP2A deletion constructs in <i>E. coli</i>	Amp
p3xFLAG-CMV10-H2A.ZΔC	Harata Lab	Expression of FLAG-tagged	Amp

MATERIAL AND METHODS

		H2A.ZΔC in HeLa Kyoto cells	
p3xFLAG-CMV10-H2A	Harata Lab	Expression of FLAG-tagged H2A in HeLa Kyoto cells	Amp
p3xFLAG-CMV10-H2A.Z	Harata Lab	Expression of FLAG-tagged H2A.Z in HeLa Kyoto cells	Amp

2.1.7 Oligonucleotides

Sequences of oligonucleotides are always listed from 5' to 3' end.

2.1.7.1 Oligonucleotides for GATEWAY cloning

Table 2.4 GATEWAY cloning

Name	Sequence	T _M in °C	Description
PWWP2A_S_short_f	GGGGACAAGTTTGTACAAAAAAGCAG GCTTAACTTCTGACTCTTCCAGTGCTTC AGTGTG	85.4	Cloning of S_short into pGEX6P1
PWWP2A_S_short_r	GGGGACCACTTTGTACAAGAAAGCTG GGTCTCAGGAATTACTGAGCTTTTTCT CTTCCTTTG	85.4	
PWWP2A_S_short_f	GGGGACAAGTTTGTACAAAAAAGCAG GCTTAACTTCTGACTCTTCCAGTGCTTC AGTGTG	85.4	Cloning of S_long into pGEX6P1
PWWP2A_S_long_r	GGGGACCACTTTGTACAAGAAAGCTG GGTCTCAACATATGGTCCTGCCATCTG GTGTG	89.2	
PWWP2A_PWWP_short_f	GGGGACAAGTTTGTACAAAAAAGCAG GCTTAACTGTAGGGGACATTGTTTGG GCCA	86.6	Cloning of PWWP_short into pGEX6P1
PWWP2A_PWWP_short_r	GGGGACCACTTTGTACAAGAAAGCTG GGTCTCACGTTTCAAAGTGTGCAACA AAGCC	87.9	

2.1.7.2 Oligonucleotides for site-directed mutagenesis

Table 2.5 Site-directed mutagenesis

Name	Sequence	T _m in °C	Mutation
PWWP2Amut461for	CAAAAGTCCATTTACACGTCAATATCAGAAT CCTAGCTCAGG	74.1	R461Q

MATERIAL AND METHODS

PWWP2Amut461rev	CCTGAGCTAGGATTCTGATATTGACGTGTGAA ATGGACTTTTG	74.1	
PWWP2Amut475for	TCCCTTCCACCCCGGGTTTGTAAAACACAG AGGTAC	81.6	R475C
PWWP2Amut475rev	GTACCTCTGTGGTTTTAAACAAACCCGGGGTG GAAGGGA	81.6	
F666A_FOR	CAAGATATATGGCGCCCTTGGTGGCCAG	70.9	F666A
F666A_REV	CTGGCCACCAAGGGGCGCCATATATCTTG	70.9	
W669A_FOR	GGCTTCCCTGGGCGCCAGCCCGTATTC	73.9	W669A
W669A_REV	GAATACGGGCTGGCGCCCAAGGGAAGCC	73.9	
W695A_FOR	GAGGCCCGTATTTACGCGTTTGGGTCTCC	70.9	W695A
W695A_REV	GGAGACCCAAACGCTGAAATACGGGCCTC	70.9	

2.1.7.3 Oligonucleotides for EMSA

Table 2.6 RNA and DNA

Name	Sequence	Length	Description
601_RNA_rv	CCAAUUGAGCGGCCUCGGCACCGGG	25	EMSA
601RNA_25bp_cy3	CCCGGUGCCGAGGCCGCUCAAUUGG	25	EMSA
F-601 25bp Cy5	CCCGGTGCCGAGGCCGCTCAATTGG	25	EMSA
R-601 25bp	CCAATTGAGCGGCCTCGGCACCGGG	25	EMSA
R-601 50bp Cy5	CCCGGTGCCGAGGCCGCTCAATTGGTCGT AGACAGCTCTAGCACCGCTTA	50	EMSA
R-601 50bp	TAAGCGGTGCTAGAGCTGTCTACGACCAAT TGAGCGGCCTCGGCACCGGG	50	EMSA
R-601 75bp Cy5	CCCGGTGCCGAGGCCGCTCAATTGGTCGT AGACAGCTCTAGCACCGCTTAAACGCACGT ACGCGCTGTCCCCGC	75	EMSA
R-601 75bp	GCGGGGACAGCGCGTACGTGCGTTTAAG CGGTGCTAGAGCTGTCTACGACCAATTGA GCGGCCTCGGCACCGGG	75	EMSA

2.1.7.4 Oligonucleotides for PCR-based DNA preparation

Table 2.7 Primer for PCR-based DNA preparation

Name	Sequence	Length	Description
Cy3_147bp_F	Cy3-CTGGAGAATCCCGGTGCCG	19	147 bp DNA fragment (0 bp linker DNA)
Cy5_147bp_F	Cy5-CTGGAGAATCCCGGTGCCG	19	147 bp DNA fragment (0 bp linker DNA)

MATERIAL AND METHODS

147bp_R	ACAGGATGTATATATCTGACACGTGCCTGG	30	147 bp DNA fragment (0 bp linker DNA)
Cy3_187bp_F	Cy3-GGACCCTATACGCGGCCG	18	187 bp DNA fragment (20 bp linker DNA)
Cy5_187bp_F	Cy5-GGACCCTATACGCGGCCG	18	187 bp DNA fragment (20 bp linker DNA)
187bp_R	GGTCGCTGTTCAATACATGCACAGG	25	187 bp DNA fragment (20 bp linker DNA)

2.1.8 Bacterial strains and cell lines

2.1.8.1 *E. coli* strains

Table 2.8 Bacterial strains

Strain	Genotype	Supplier
BL21-CodonPlus (DE3)-RIL	B F ⁻ <i>ompT</i> hsdS (r _B ⁻ , m _B ⁻) <i>dcm</i> ⁺ Tet ^r gal λ /DE3) <i>endA</i> The [<i>argU ileY leuW Cam</i> ^r]	Stratagene
DH5α	F ⁻ Φ80 <i>lacZ</i> ΔM15 Δ(<i>lacZYA-argF</i>) _{U169} <i>recA1 endA1 hsdR17</i> (r _K ⁻ , m _K ⁺) <i>phoA supE44</i> λ ⁻ <i>thi-1 gyrA96 relA1</i>	Genentech

2.1.8.2 Human cell lines

Table 2.9 Human cell lines

Cell line	Origin	Source
HeLa Kyoto (HK)	Cervical cancer	H. Leonhardt, LMU Munich

2.1.9 Software

Table 2.10 Software and web-browser-based programs

Application	Software
Image Processing	Adobe Photoshop CS5
	Adobe Illustrator CS5
	Image Studio Lite Ver. 5.2
Sequence alignment	Clustal Omega (web browser-based)

MATERIAL AND METHODS

Computation of physical and chemical parameters for a given protein	ProtParam, ExPASy (web browser-based)
---	---------------------------------------

2.1.10 Buffers and solutions

Name	Reagent	
Ampicillin Stock Solution	100 mg/ml	Ampicillin (1.000x)
Coomassie Fixing Solution	10%	Acetic acid (v/v)
	50%	EtOH (v/v)
Coomassie Staining Solution	7.5%	Acetic acid (v/v)
	5.0%	EtOH (v/v)
Coomassie Blue Solution	0.25%	Coomassie Brilliant Blue (w/v)
	100%	EtOH (v/v)
Ethidium Bromide Stock Solution	10 mg/ml	Ethidium bromide (20.000x)
HEPES (adjust pH to 7.6 or 7.9 with NaOH)	1M	HEPES
5x Laemmli Loading Buffer (adjust pH to 6.8 with HCl)	250 mM	Tris-HCl pH 7.6
	10%	SDS (w/v)
	30%	Glycerol (v/v)
	0.5 M	DTT
	0.02%	Bromphenol blue (w/v)
Laemmli Running Buffer	25 mM	Tris-HCl pH 7.6
	192 mM	Glycine
	0.1%	SDS (w/v)
LB Agar Plates	1.5%	LB Agar (w/v)
LB Medium	1.0%	Tryptone (w/v)
	0.5%	Yeast extract (w/v)
	1.0%	NaCl (w/v)
PBS	140 mM	NaCl

MATERIAL AND METHODS

	2.7 mM	KCl
	10 mM	Na ₂ HPO ₄
	1.8 mM	KH ₂ PO ₄
Protease inhibitor mix	1.0 mg/ml	Aprotinin (in H ₂ O)
	1.0 mg/ml	Leupeptin (in H ₂ O)
	0.7 mg/ml	Pepstatin A (in EtOH)
	0.2 M	PMSF (in isopropanol)
	1.0 M	DTT (in 10 mM sodium acetate pH 5.2)
Native PAGE	xx%	Polyacrylamide (v/v)
	0.9%	Ammonium persulfate (v/v)
	0.1%	TEMED (v/v)
TBE	45 mM	Tris-HCl pH 7.6
	45 mM	Boric acid
	1 mM	EDTA
TE	10 mM	Tris-HCl pH 7.8
	0.1 mM	EDTA
Transfer Buffer	48 mM	Tris-HCl pH 7.6
	39 mM	Glycine
	0.0375%	SDS (w/v)
	20%	Methanol (v/v)

2.2 Cell biological methods

2.2.1 Mammalian cell cultivation

2.2.1.1 *Thawing of mammalian cells*

Frozen HeLa Kyoto cells (storage in liquid nitrogen; cryo vials) were thawed for 2-3 min in a 37°C water bath. Afterwards, cells were resuspended in 1 ml of Dulbecco's modified Eagle medium (DMEM) supplemented with 1% penicillin/streptomycin (P/S) and 10% fetal calf serum (FCS) and transferred to a 15 ml falcon tube. Resuspension medium was removed after pelleting cells at 1.200 rpm (tabletop centrifuge) and 5 ml of fresh DMEM were added. HeLa Kyoto cells were transferred to a 10 cm plate and additional 5 ml of fresh DMEM were added to a final volume of 10 ml.

2.2.1.2 *Cultivation and storage of mammalian cells*

Adherent HeLa Kyoto cells were kept in DMEM at 37°C and 5% CO₂ in a humidified incubator. Usually, cells were cultivated in 10 cm tissue culture plates (10 ml DMEM) and passaged to 15 cm tissue culture plates (20 ml DMEM) if needed for experiments. Every second to third day (due to doubling time and confluence) cells were split in a ratio of 1:10 or 1:20. Therefore, old and yellowish medium was sucked off and cells were washed with sterile PBS to completely remove residual medium. To detach adherent cells from tissue culture plates, 1 ml (10 cm plate) or 2 ml (15 cm plate) of trypsin/EDTA were added and cells were incubated for 5 min at 37°C. Trypsinized cells were resuspended in fresh medium and superfluous cell suspension was discarded and replaced by fresh medium. HeLa Kyoto cells were frozen at regular intervals for storage. Here, trypsinized cells were resuspended in FCS + 10% DMSO and stored in liquid nitrogen. Cells in culture were regularly (every 4-8 weeks) exchanged with freshly thawed cells.

2.2.1.3 *Counting mammalian cells and preparation for experiments*

Cell number and viability were determined with the CASY cell counter (Innovatis). HeLa Kyoto cells were trypsinized, resuspended in DMEM and the required amount of cells was transferred to a 15 ml falcon tube. Afterwards, cells were pelleted at 1.200 rpm (tabletop centrifuge), washed once with PBS and subsequently used for experiments or stored at -80°C.

2.2.2 Transient transfection of mammalian cells

For transient transfection of HeLa Kyoto cells with plasmid DNA, cells were grown up to 50% confluence in 15 cm tissue culture plates. 1.5 ml OptiMEM (Life technologies), 20 µg plasmid DNA and 60 µl X-tremeGENE (Roche) were mixed and incubated for 15-30 min at room temperature (RT). Next, medium of adherent cells was changed (10 ml DMEM + 600 µg/ml G-418 sulfate liquid (PAA)) and transfection mix was added drop by drop. Cells were kept under selection for 2-3 days, harvested and pelleted for experiments.

2.3 Molecular biological methods

2.3.1 Agarose gel electrophoresis

DNA fragments of PCR-based DNA preparations, micrococcal nuclease (MNase) digests or cloning reactions were checked on native agarose gels. Gels were prepared with 1% agarose (w/v) in TBE buffer and the intercalating reagent ethidium bromide was added to a final concentration of 0.5 µg/ml to visualize DNA samples. Gels were run in TBE buffer for 1 h at 120 V and detected with a gel documentation system (Peqlab).

2.3.2 Cloning of PWWP2A truncations

Different N-terminal GST-tagged PWWP2A truncation constructs were generated by using a GATEWAY-cloning strategy, with the PWWP2A full-length construct (pGEX6P1-GST-PWWP2A_fl) serving as initial template DNA. For the PCR reaction, template (1 ng), 1x HF buffer (NEB), dNTP mix (0.2 mM, NEB), primers (100 nM, see section 2.1.7.1) and Phusion polymerase (1U) (NEB) were mixed and GATEWAY PCR program (see below) was run. Amplified products were checked on a 1% agarose gel and purified with the “PCR clean-up Gel extraction kit” (Macherey-Nagel) according to the manufacturer’s instructions. The entry clone was generated by incubating 100 ng of purified PCR product with 150 ng pDONR and BP (attB x attP reaction) clonase enzyme mix II (life technologies) at 25°C for 1h. The reaction was stopped by adding 2 µg/µl Proteinase K (life technologies) (37°C, 10 min). Positive clones were selected by transforming BP reaction product in DH5α competent *E.coli* cells and growing them on zeomycin agar plates (overnight (ON), 37°C). Clones were picked, inoculated in 6 ml LB-zeomycin and DNA was isolated with a NucleoSpin Plasmid Miniprep kit (Macherey-Nagel)

MATERIAL AND METHODS

according to the manufacturer's instructions. Then, 100 ng of the entry clone were mixed with 150 ng/ μ l of destination vector (pGEX6P1) and LR (attL x attR reaction) clonase enzyme mix II (life technologies) (25°C, 1h). The reaction was stopped by adding 2 μ g/ μ l Proteinase K and LR reaction product was transformed in DH5 α competent *E.coli* cells and bacteria were grown on ampicillin agar plates (ON, 37°C). Positive clones were inoculated in 6 ml LB-ampicillin and DNA was isolated with a Miniprep kit (Macherey-Nagel). Subsequent sequencing (Eurofins MWG Operon) revealed positive clones.

GATEWAY PCR program

3 min	98°C	
30 sec	98°C	} 30 cycles
30 sec	58°C	
1 min	72°C	
7 min	72°C	

2.3.3 Site-directed mutagenesis

Aromatic cage (S_PWWP_F666A, S_PWWP_W669A and S_PWWP_W695A) and internal region (I_R461Q and I_R475C) point mutant constructs were generated by site-directed mutagenesis. Therefore, template DNA (pGEX6P1-GST-S_PWWP or pGEX6P1-GST-I; 10 ng), primers (100 nM, see section 2.1.7.2), dNTP mix (0.2 mM, NEB), 1x Pfu buffer and Pfu Hotstart turbo DNA polymerase (1U) (Agilent) were mixed and PCR program (see below) was run. Following, 20 μ l of PCR reaction were mixed with 1 μ l Dpn1 (NEB) and digested for 2h at 37°C. Next, 10 μ l of Dpn1 digested PCR product was transformed in DH5 α competent *E.coli* cells and grown on amp-agar plates. Colonies were picked, inoculated in 6 ml LB-amp and DNA was isolated with a Miniprep kit (Macherey-Nagel). Subsequent sequencing (Eurofins MWG Operon) revealed positive clones.

MATERIAL AND METHODS

Site-directed mutagenesis PCR program

	5 min	95°C	
	50 sec	95°C	} 19 cycles
	50 sec	58°C	
1 min per kb of construct	xx min	68°C	
	7 min	68°C	

2.4 Biochemical methods

2.4.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated due to their molecular weight (kDa) using precast polyacrylamide gradient gels (4-20%, SERVA). Peq GOLD Protein Marker IV and V served as size markers to distinguish between different proteins. Prior to loading, protein samples were denatured in Laemmli loading buffer at 95°C for 5 min. Gels were run in Laemmli running buffer at 150-180 V until marker bands were well separated. Subsequently, gels were either stained with Coomassie solution or used for immunoblotting.

2.4.2 Coomassie staining of SDS-PAGE gels

To visualize separated proteins, polyacrylamide gels were incubated for 20 min at RT in fixing solution. Next, fixing solution was exchanged with 10 ml staining solution and 1 ml Coomassie blue solution. Gels were incubated for 30 min at RT and destained in VE water until bands got visible. Gels were scanned (V700 PHOTO, Epson) with the following settings: transparency mode, 300 dpi resolution, 16-32bit HDR grey scale, tiff-format.

2.4.3 Immunoblotting

A semidry blotting device (Bio-Rad) was used to blot SDS-PAGE-separated proteins on nitrocellulose membranes (Amersham). Whatman papers (3 mm), nitrocellulose membranes and gels were soaked for 5 min in transfer buffer and assembled in the following way: two Whatman papers, nitrocellulose membrane, gel, two Whatman papers. Proteins were transferred on the membrane with 250 mA (1 blot) or 300 mA (>1 blot) for 1 h 5 min. Next, membranes were blocked for 1 h at RT in either PBS (Fluorophores, Licor) or PBS-T (PBS + 0.1%

MATERIAL AND METHODS

Tween) (ECL, x-ray) containing 5% milk powder (w/v). The incubation with primary, protein specific antibody was done in PBS-T (5% milk) ON at 4°C (see section 2.1.5 for antibody dilutions). After 3 x 10 min washing steps with PBS-T, membranes were incubated with secondary antibody in either PBS-T (Fluorophores, Licor) or PBS-T + 5% milk powder (ECL, x-ray) for 1 h at RT. For the detection with a developer machine (AGFA), blots were covered for 2 min with ECL solution (RT). X-ray films (Fujifilm) were exposed in a dark room and due to signal intensity; incubation time of x-ray films with membrane took 1 sec – 15 min. For the visualization of signals with the Licor machine, membranes were dried between two Whatman papers for 30 min at RT and scanned in the corresponding channel (700 nm mouse, 800 nm rabbit).

2.4.4 Mononucleosome preparation of HeLa Kyoto cells

HeLa Kyoto and transiently transfected cells were grown in 15 cm tissue culture plates and prepared as explained in section 2.2.1.2 and 2.2.1.3.. 1×10^7 cells were aliquoted into a 15 ml falcon tube and washed once with PBS. All buffers contained protein inhibitor mix and used tubes had low-binding affinity for proteins. Cell pellets were resuspended in 1.25 ml PBS containing 0.03% Triton X-100 and were incubated for 10 min at 4°C on a rotating wheel to lyse the plasma but not nuclear membrane. Following, nuclei were pelleted for 5 min at 2.000 rpm at 4°C (Eppendorf 5810 R) and washed once with cold 1.25 ml PBS on ice. Supernatant was removed and pelleted nuclei were resuspended in 125 μ l ice-cold EX100 buffer per 1×10^7 cells on ice. Furthermore, the CaCl_2 concentration, important for subsequent MNase digestion, was adjusted to 2 mM (1 μ l of 250 mM stock solution). Lyophilized micrococcal nuclease (MNase, Sigma) was resuspended in sterile Milli-Q water to a final enzyme activity of 1 U/ μ l, aliquoted, shock-frozen in liquid nitrogen and finally stored at -80°C. For MNase digestion of isolated nuclei, 1.5 U of enzyme was added and incubated at 26°C for 30 min. The reaction was stopped by adding 10 mM EGTA (2.5 μ l of 500 mM stock solution) and centrifuged for 30 min, 15.000 rpm at 4°C (Eppendorf 5424 R). The supernatant (S1, early chromatin), containing almost exclusively mononucleosomes, was transferred to a fresh low-binding tube and pooled (>1 fraction). 25 μ l of S1 mononucleosomes were kept as “input” and boiled in Laemmli buffer for 5 min at 95°C. Another 25 μ l were used to analyze mononucleosomal DNA. The rest of the pooled mononucleosomal fractions were mixed with

MATERIAL AND METHODS

glutathione sepharose bead linked GST-tagged PWWP2A truncation constructs for immunoprecipitations (see section 2.4.6).

EX100 buffer (storage at 4°C)

	10 mM	HEPES pH 7.6
	100 mM	NaCl
	1.5 mM	MgCl ₂
	0.5 mM	EGTA
	10%	Glycerol (v/v)
	10 mM	β-Glycerol phosphate
Add prior to use:	1 mM	DTT
		Protease inhibitor mix (1:1.000)

2.4.5 Purification of MNase digested DNA

The digestion degree of MNase was checked by DNA extraction. Therefore, 175 µl of 5 mM Tris-HCl (pH 8.0) were mixed with 25 µl S1 supernatant and 200 µl phenol. Following, 200 µl of chloroform:isoamylalcohol (ratio 24:1) were added and transferred to an empty, pre-spun (13.000 rpm, 30 sec, tabletop centrifuge) maXtract tube (Qiagen). The separation of organic from aqueous phase was done by a centrifugation step at 11.000 rpm (tabletop centrifuge) for 1 min at RT. The upper aqueous phase was transferred to a fresh tube and 2 µl glycogen (20 mg/ml) and 20 µl sodium acetate (3 M) were added. To precipitate DNA, 500 µl of 100% ice-cold EtOH were added and incubated for at least 20 min at -20°C. Subsequently, the solution was centrifuged for 30 min, 15.000 rpm (Eppendorf 5424R) at 4°C and washed once with cold 70% EtOH. After centrifuging for 15 min, 15.000 rpm (Eppendorf 5424R) at 4°C supernatant was removed and pellet was air dried until ethanol was evaporated. The pellet was resuspended in 20 µl 5 mM Tris-HCl (pH 8.0) and concentration was checked with Denovix photometer (Peqlab). The analysis of DNA length was done by running 1 µg of DNA on a 1% agarose gel (1 h, 100 V).

MATERIAL AND METHODS

2.4.6 Purification of recombinant GST-tagged PWWP2A and PWWP2A truncations

2.4.6.1 Protein purification on beads

For approaches like electrophoretic mobility shift assays (EMSAs), purified GST-tagged PWWP2A constructs were indispensable. For this purpose, 50 μ l of competent BL21 *E. coli* cells were transformed with 50 ng of pGEX6P1-PWWP2A construct (see section 2.1.6) and incubated for 30 min on ice. After heat shocking bacteria at 42°C for 45 sec they were cooled down on ice for 2 min and resuspended in 100 μ l LB medium (free of antibiotics). After an incubation of 1 h at 37°C, bacteria were plated on ampicillin (stock: 100 mg/ml, dilution: 1:1.000) agar plates and incubated ON at 37°C to grow single colonies. The next day, one clone was inoculated by adding it to 8 ml LB medium supplemented with ampicillin (100 μ g/ml) and grown ON to a pre-culture while shaking at 37°C, 150 rpm (INFORS shaker). 2 ml of that culture were used to inoculate 200 ml LB medium containing 100 μ g/ml ampicillin. Then, cultures were incubated at 37°C, 150 rpm (INFORS shaker) and grown until an OD₆₀₀ of 0.5 – 0.6 was reached. Next, bacteria were cooled down to 4°C to stop further growth. After inducing recombinant protein expression in bacteria with 0.3 mM IPTG (stock 1 M) and adding additional 50 μ g/ml ampicillin, cultures were incubated ON for 16 h – 18 h while shaking at 150 rpm (INFORS shaker). Following, the overnight culture was spun down for 30 min, 4.000 rpm (Heraeus Cryofuge 6000i) at 4°C and bacterial pellet was either shock frozen in liquid nitrogen or stored at -80°C. Protein purification was carried out on ice or 4°C, using ice-cold buffers and protease inhibitor mix in every step. Cell lysis took place by resuspending cells in 10 ml lysis buffer and sonicating them for 2 min with the Branson Sonifier (30% output, 1 sec on, 1 sec off). The separation of proteins from cell debris was examined by centrifuging lysates for 30 min at 4°C and 15.000 rpm (Beckman JXN-26). In the meantime, glutathione sepharose beads (GE) were equilibrated by washing them 2 x 10 min with PBS while rotating at 4°C. Therefore, 200 μ l of slurry were transferred to a 15 ml falcon tube and 5 ml of cold PBS were added. Centrifugation of beads was always limited to 2.000 rpm to prevent the loss of bound protein. Clear supernatant, containing bacterial lysates, was added to washed beads and incubated for 2 h at 4°C on a rotating wheel. Then, beads were pelleted by centrifugation (5 min, 4°C, 2.000 rpm, tabletop centrifuge) and supernatant was removed. Beads with bound proteins were resuspended in 1 ml lysis buffer, transferred to low protein binding tubes and washed for 5 min at 4°C on a rotating wheel. After removing lysis buffer, two washing steps

MATERIAL AND METHODS

with 1 ml of PBS followed. Next, beads were resuspended in storage buffer and 100 µl were taken for checking protein purification extent on a SDS-PAGE. Thus, beads were spun down and resuspended in 20 µl Laemmli buffer and boiled for 5 min at 95°C. 10 µl were loaded on a 4-20% gradient gel and run for 1 h at 180 V. Remaining protein on beads was aliquoted and stored at -80°C after shock freezing them in liquid nitrogen.

2.4.6.2 *Protein purification followed by elution of recombinant GST-tagged proteins from GST-trap column*

For several experimental approaches eluted recombinant GST-protein is necessary. All steps including cell lysis followed by centrifugation were carried out as described in 2.4.6.1. After the separation of soluble protein from cell debris, supernatant was loaded on a 5 ml GST-trap column (GE) and purified with an Äkta pure system (GE) with the following settings: alarm system pressure 0.5 MPa, Δcolumn pressure 0.3 MPa, system flow 0.5 ml/min-0.8 ml/min (dependent on column). Thereby, GST-tagged proteins were bound to a GST-trap column and washed with lysis buffer until all parameters of the HPLC reached again their base line levels. Next, proteins were eluted by using elution buffer and were fractionated into a 96 well fractionation plate. Purified, protein-containing fractions were identified by loading 10 µl of eluate with 5 µl of 60% glycerol on a 4-20% SDS-PAGE with subsequent Coomassie-staining. Fractions with eluted GST-protein were combined and dialyzed while stirring against 1 l of dialysis buffer (Spectra/Por membranes, MWCO 10 kDa) at 4°C for 3-4 h. Next, dialysis buffer was exchanged and GST-proteins were dialyzed ON at 4°C while stirring. GST-protein was concentrated using Amicon Ultra centrifugation tubes (10 k) to a final volume of 1/10. GST-protein concentration was determined by using a Denovix photometer (Peqlab) and GST-proteins were stored at -80°C after shock freezing in liquid nitrogen.

Lysis buffer (sterile filtered) (storage at 4°C)

	1x	PBS
	500 mM	NaCl
Add prior to use:	protease inhibitor mix (1:1.000)	

Storage buffer (storage at 4°C)

	1x	TE
--	----	----

MATERIAL AND METHODS

50 mM	NaCl
10%	Glycerol (v/v)

Elution buffer (sterile filtered, prepare fresh)

30 mM	L-glutathione
1x	PBS

Adjust pH: add NaOH to a final pH of 7.4

Dialysis buffer (storage at 4°C)

150 mM	NaCl
25 mM	HEPES pH 7.6
2 mM	MgCl ₂
10%	Glycerol (v/v)
0.1 mM	EDTA
0.05%	NP-40 (v/v)

Table 2.11 Recombinant GST-PWWP2A full length and truncations. MW, expected molecular weight.

PWWP2A_	Amino acid stretch	MW incl. GST-tag (kDa)
full-length (fl)	1 - 755	110
P1	1 - 240	53
I	292 - 574	60
IN	292 - 422	43
IC	423 - 574	45
I_S	292 - 654	69
I_S_PWWP	292 - 755	80
S_PWWP	575 - 755	51
PWWP	633 - 755	45
PWWP_short	655 - 755	42
S_long	575 - 654	39
S_short	575 - 632	37
I_R461Q	292 - 574	60
I_R475C	292 - 574	60
S_PWWP_F666A	575 - 755	51
S_PWWP_W669A	575 - 755	51
S_PWWP_W695A	575 - 755	51

MATERIAL AND METHODS

2.4.7 Identification of histone modifications by Mass-Spectrometry (MS)

2.4.7.1 MNase Immunoprecipitations

Mononucleosomes from HeLa Kyoto cells were prepared as described in 2.4.4. All buffers were supplemented with protease inhibitor mix and all reactions were carried out in low protein binding tubes. Mononucleosomes from 1×10^7 cells were immunoprecipitated ON with 25 μ l slurry glutathione sepharose bead-linked GST-tagged PWWP2A full-length and truncation constructs while rotating at 4°C. Beads were spun down (2.000 rpm, 4°C, tabletop centrifuge) and supernatant was discarded. Following, beads were washed 2 x 10 min with 1 ml wash buffer 1 and 2 x 10 min with 1 ml wash buffer 2 while rotating at 4°C. After supernatant was discarded, beads were boiled in 50 μ l Laemmli loading buffer for 5 min at 95°C. Immunoprecipitations were stained with Coomassie (brilliant blue G-250) (MS) after they were separated by SDS-PAGE (4-20% gradient gel, SERVA).

Wash buffer 1 (storage at 4°C)

	10 mM	HEPES pH 7.6
	300 mM	NaCl
	1.5 mM	MgCl ₂
	0.1%	NP-40 (v/v)
Add prior to use:	1 mM	DTT
		protease inhibitor mix (1:1.000)

Wash buffer 2 (storage at 4°C)

	10 mM	HEPES pH 7.6
	300 mM	NaCl
	1.5 mM	MgCl ₂
Add prior to use:	1 mM	DTT
		protease inhibitor mix (1:1.000)

Coomassie blue solution (MS)

	50%	Methanol (v/v)
	10%	Acetic acid (v/v)
	0.1%	Coomassie brilliant blue (w/v)

MATERIAL AND METHODS

Destaining solution (MS)

10%	Acetic acid (v/v)
-----	-------------------

2.4.7.2 Sample preparation

Separated protein fractions in the molecular weight range of histones (15-23 kDa) were excised as single band/fraction. Destaining of gel pieces was done in 50% acetonitrile/50 mM ammonium bicarbonate. Tryptic cleavage was prevented by propionylating lysine residues with 2.5% propionic anhydride (Sigma) in ammonium bicarbonate (pH 7.5) for 30 min at RT. Thus, only unmodified and monomethylated lysines were chemically modified with a propionyl-group, whereas lysines with other side chain modifications were not. Next, proteins were digested with 200 ng trypsin (Promega) in 50 mM ammonium bicarbonate and supernatant was desalted by carbon Top-Tips (Glygen) and C18-Stagetips (reversed-phase resin) according to the manufacturer's instructions. Finally, dried peptides were resuspended in 17 μ l of 0.1% TFA.

2.4.7.3 LC-MS analysis of histone modifications

LC-MS analysis and all quantifications of histone modifications were done by Moritz Völker-Albert (AG Imhof). Sample separation was done on a C18 home-made column (C18RP Reasil-Pur AQ, 120 x 0.075 mm x 2.4 μ m, 100 Å , Dr. Maisch, Germany) for 5 μ l sample volume with a gradient from 5% B to 30% B (solvent A 0.1% FA in water, solvent B 80% CAN, 0.1% FA in water). The optimal flow rate is at 300 nl/min (Ultimate 3000 RSLC Thermo-Fisher, San Jose, CA) over 32 min. Samples were directly sprayed into a Q-Exactive HF mass spectrometer (Thermo-Fisher Scientific). The identification and quantification of specific fragment ions of N-terminal peptides of human histone 3.1 and histone 4 proteins was done in PRM operation mode of the mass spectrometer. There, the mass spectrometer switched between one survey scan and 9 MS/MS acquisitions of the m/z values in an automatic way. M/z values are described in the inclusion list containing modifications, fragmentation conditions and precursor ions (**Appendix Table A.1**). MS spectra (survey full scan) from m/z 250-800 were acquired with a resolution of 15.000 to a target value of 2×10^5 , fragmented at 27% normalized collision energy, isolation window 0.7 m/z and maximum IT 60 ms. Used mass

MATERIAL AND METHODS

spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250°C.

2.4.7.4 LC-MS/MS analysis

A reversed-phase liquid chromatography on an EASY-nLC 1000 or 1200 system (Thermo Fisher Scientific, Odense, Denmark) coupled to a Q Exactive plus or HF mass spectrometer (Thermo Fisher Scientific) was used for the analysis of peptides. 50 cm (length) HPLC columns with an inner diameter of 75 µm were “in-house” packed with ReproSil-Pur 120 C18-AQ 1.9 µm particles (Dr. Maisch GmbH, Germany). A linear gradient of 120 or 140 min (total run time and washout) and a two buffer system (buffer A⁺⁺ and buffer B⁺⁺) were used to separate peptide mixtures. The data-dependent top 10 or top 15 operation mode was used for the mass spectrometer. Fragmentation of peptides occurred by higher energy collisional dissociation (HCD) with a normalized collision energy of 27.

Buffer A⁺⁺ (prepare fresh)

0.1%	Formic acid (v/v)
------	-------------------

Buffer B⁺⁺ (prepare fresh)

80%	Acetonitrile (v/v)
-----	--------------------

0.1%	Formic acid (v/v)
------	-------------------

2.4.8 Reconstitution of recombinant mononucleosomes

2.4.8.1 Octamer assembly

Lyophilized, recombinant human H2A, H2A.Z, H2B, H3 and H4 histone proteins were ordered from Hataichanok Scherman (www.histonesource.com). All buffers were sterile filtered and stored at 4°C until usage. Lyophilized histones were resuspended in unfolding buffer with a final concentration of 4 mg/ml and unfolded at RT for 3 h while rotating. After 30 min of unfolding, purity and presence of similar histone concentrations were determined by checking 5 µg and 10 µg (diluted in 10 µl urea buffer) of each histone on a 16% SDS-PAGE gel (175 V, 90 min, SERVA) followed by Coomassie-staining. After unfolding, histones (H2A, H2B, H3, H4 or H2A.Z, H2B, H3, H4) were combined in a 2 ml tube. Thereby, the concentration of

MATERIAL AND METHODS

H2A/H2A.Z/H2B is always 1.2-fold the concentration of H3/H4, because H2A/H2B or H2A.Z/H2B-dimers are easier to separate from octamers than H3/H4-tetramers. Next, histone mix was carefully transferred to a special, in refolding buffer-moistened membrane (MWCO 3.5 kDa, spectrumlabs) and dialyzed at 4°C while stirring with the following intervals: 2 h/0.5 l, ON/1 l, 2 h/0.5 l in fresh refolding buffer. Surplus histones can be stored at -80°C after shock freezing in liquid nitrogen. Octamers were separated from dimers and tetramers by size exclusion (gel filtration) using a Superdex2000 column (GE) linked to the Äkta pure system (GE). Here, higher molecular-weighted octamers were eluted later than dimers and tetramers and could be separately fractionated (**Figure 2.4.1**). Next, octamer-containing fractions were centrifuged for 20 min, 15.000 rpm (Eppendorf 5424R) at 4°C and checked on a 16% SDS-PAGE gel (SERVA). After combining all octamer-containing fractions, octamers were concentrated (Amicon ultra (Merck/Millipore), MWCO 10 kDa) to a final volume of 1/10, aliquoted (low protein binding tubes) and stored at -80°C after shock freezing.

Unfolding Buffer (storage at 4°C)

	7 M	Guanidium-HCl
	20 mM	Tris-HCl pH 7.5
Add prior to use:	10 mM	DTT
Adjust pH:	add NaOH to a final pH of 7.5	

Refolding Buffer (storage at 4°C)

	2 M	NaCl
	10 mM	Tris-HCl pH 7.5
	1 mM	EDTA
Add prior to use:	5 mM	DTT
Adjust pH:	add NaOH to a final pH of 7.5	

Urea Buffer

	20 mM	Tris-HCl pH 7.5
	8 M	Urea
Adjust pH:	add NaOH to a final pH of 7.5	

MATERIAL AND METHODS

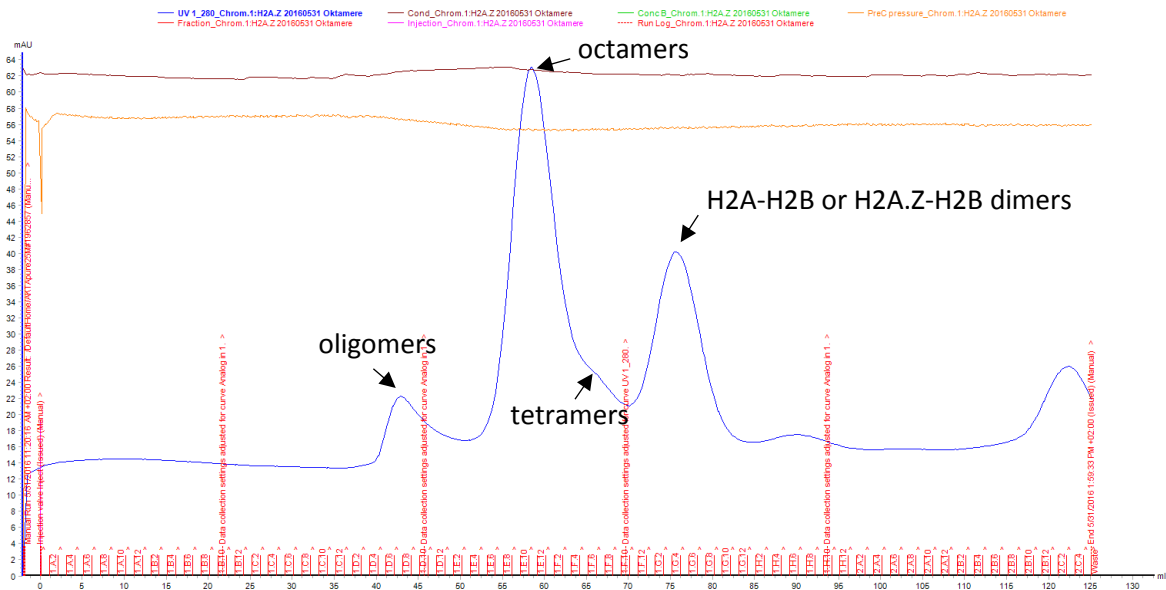


Figure 2.4.1 Gel filtration elution profile of recombinant octamers. Recombinant assembled octamers were separated using a Superdex2000 column (GE) and the Äcta pure system (GE). The y-axis represents the milli-absorption-unit (mAu) and the x-axis the retention volume (ml). mAu measurement occurs by UV light (280 nm). Red numbering on x-axis (1A1, 1A2, 1A3,...) represents the well number of a 96-well fractionation plate. Oligomers elute first (40-50 ml) followed by octamers (51-63 ml), tetramers (64-70 ml) and H2A-H2B or H2A.Z-H2B dimers (71-85 ml). Fractions containing only octamers were subsequently combined, concentrated and stored at -80°C .

2.4.8.2 PCR-based DNA preparation

For the reconstitution of recombinant mononucleosomes, fragments of 147 bp (0 bp linker) and 187 bp (20 bp linker) DNA were required (for primer information see also 2.1.7.4). 1 ml Taq buffer (10x, NEB), 200 μl dNTPs (10 mM, NEB), 50 μl forward primer (100 μM), 50 μl reverse primer (100 μM), 10 μl template (see also 2.1.6) (400 ng/ μl), 50 μl Taq Polymerase (22 U/ μl , NEB) and 8640 μl water (distilled, deionized) were mixed and split to 192 x 50 μl aliquots in PCR reaction tubes. After the PCR run (see below for PCR program information), all products were combined in a 15 ml falcon tube and polyethylene glycol (PEG)-precipitated to remove template DNA (can be circumvented if small DNA fragment is used as template). Therefore, PCR products were split into 1 ml aliquots and 500 μl of PEG buffer (final PEG concentration of 10% (v/v)) were added. After 20 min centrifugation (15.000 rpm, RT, tabletop centrifuge) supernatant was transferred into an oakridge centrifugation tube and sample volume was determined by pipetting. Primer pairs, Taq-Polymerase and dNTPs were removed by adding 1 volume of isopropanol (RT) to PEG supernatant and DNA was pelleted for 30 min, 15.000 rpm (Beckmann JXN-26, JA25-50 rotor) at 25°C . Next, supernatant was decanted and pellet was

MATERIAL AND METHODS

washed with 70% of ice-cold ethanol. After an incubation step of 30 min at -20°C, washed DNA was centrifuged for 30 min, 15.000 rpm (Beckmann JXN-26, JA25-50 rotor) at 4°C and supernatant was removed. DNA pellet was air-dried and finally dissolved in a total volume of 500 µl TE. DNA concentration and quality were determined by using a Denovix photometer (Peqlab) and/or an agarose gel followed by ethidium bromide staining.

Table 2.12 PCR-based DNA fragments and annealing temperature

Fragment	Primer forward	Primer reverse	T _{annealing}
0/0 bp linker	147bp_F	147bp_R	59°C
20/20 bp linker	187bp_F	187bp_R	58°C

PCR-based DNA preparation program

	4 min	95°C	
T _{annealing}	30 sec	95°C	} 30 cycles
	30 sec	xx°C	
	30 sec	68°C	
Per kb of construct	30 sec	68°C	
	7 min	68°C	

PEG buffer

30%	PEG 8000 (w/v)
30 mM	MgCl ₂

2.4.8.3 Reconstitution of mononucleosomes

For the reconstitution of recombinant H2A- or H2A.Z-containing, single histone tail-deleted and H3Kc36me3- or H3Kc36me0-containing mononucleosomes (H3Kc36: (methyl-) lysine analog at position 36 of H3) with or without linker DNA, either a salt and glycerol gradient-based or salt dilution-based reconstitution method was applied.

2.4.8.3.1 Salt and glycerol gradient-based reconstitution of mononucleosomes

The general idea of salt and glycerol gradient-based reconstitution of mononucleosomes is the attachment of DNA to the basic surface of histones via salt-gradient dialysis (2M KCl to 200 mM KCl) and the separation of mononucleosomes from free DNA via glycerol gradient

MATERIAL AND METHODS

ultra-centrifugation. For that workflow, all used buffers were sterile-filtered and stored at 4°C until usage. Usually it is not known, what the optimal ratio of DNA to octamer for the assembly of mononucleosomes is. Hence, different ratios with a total sample volume of 20 µl and DNA concentration of 200 ng/µl must be initially checked in test assemblies. Possible octamer/DNA ratios in 2M KCl, 20 mM Tris-HCl (pH 7.7) and 10 mM DTT are: 1:1, 1.3:1, 1.8:1 and 2.1:1. Special slide-a-lyzer tubes (Thermo Fischer Scientific) were dipped in Mononuc-2000 buffer and octamer/DNA mix was added. Next, tubes were placed in a styropor-swimmer and salt gradient dialysis was done by using a minipuls pump (Gilson) for 24 h, 0.3 rpm at 4°C. Mononuc-2000 buffer was gradually exchanged with Mononuc-0 buffer to a final salt concentration of 200 mM. Following, samples were transferred to 1.5 ml low protein binding tubes, quickly spun down (13.000 rpm, 5 min, 4°C, Eppendorf 5424R) and supernatant was transferred to a fresh tube. After measuring DNA concentrations with a Denovix photometer (Peqlab), mononucleosomes were additionally checked on a 4.5% native PAA (polyacrylamide) gel. Therefore, 200 ng of DNA were resuspended in dialysis buffer to a final volume of 10 µl. Further, 1/10 (v/v) of glycerol (2,33 µl of 60% glycerol stock solution) were added. Gel was pre-run for 1 h at 100 V and pockets were flushed before sample loading. DNA only (control) and mononucleosomes were loaded on the gel and run for 90 min at 100 V. DNA was visualized with ethidium bromide staining (5 µl of 20 mg/ml stock solution in 100 ml Mononuc-0 buffer) or with the Typhoon FLA9500 (GE) scanner (fluorophoric labels like Cy3 and Cy5 must be present). Thus, the best octamer/DNA ratio for the “real assembly” can be determined and over-assemblies can be avoided. Following, the optimal octamer/DNA ratio approach was prepared with a maximum volume up to 500 µl. Instead of slide-a-lyzer tubes, a spectra/ Por®1 Dialysis membrane (MWCO 6-8 kDa, flat width 10 mm, diameter 6,4 mm) was used. Before sample loading, the dialysis membrane was first moistened for 10 min in Mononuc-2000 buffer. After salt-gradient dialysis (minipuls pump (Gilson), 24 h, 0.3 rpm, 4°C), Mononuc-0 buffer was exchanged and samples were kept ON in fresh buffer at 4°C while stirring. Next, mononucleosomes were separated from free DNA by glycerol gradient centrifugation. Therefore, a glycerol gradient was prepared by undercoating 5.5 ml of 10% glycerol buffer with 5.5 ml of 30% glycerol buffer and using the master cycler (biocamp) with the following settings: 2.25 min, 81.5°, 11 rpm, UP rotation. Samples were added drop-wise on top of the glycerol gradient and centrifuged for 19 h, 34.000 rpm at 4°C (Beckmann XPN-80, SW-41 rotor). Mononucleosome-containing fractions were collected in 1.5 ml low protein-binding

MATERIAL AND METHODS

tubes and checked on a native 4.5% PAA gel. After gel was pre-run for 1h at 100 V, 15 μ l of mononucleosomes were loaded and run for 90 min at 110 V. Mononucleosomes were checked either by ethidium bromide staining or with the Typhoon FLA9500. Fractions, that contained the desired mononucleosomal assembly were pooled and concentrated at 1.000 rpm (Eppendorf 5810R), 4°C, using an Amicon concentration falcon (MWCO 10 kDa) until a final volume of approximately 100-120 μ l was reached. Mononucleosomes were stored at 4°C for at least one year.

10x Mononuc-0 (storage at 4°C)

	200 mM	Tris-HCl pH 7.7
	1 mM	EDTA pH 8.0
Add prior to use:	1 mM	DTT

1x Mononuc-2.000 (storage at 4°C)

	20 mM	Tris-HCl pH 7.7
	0.1 mM	EDTA pH 8.8
	2 M	KCl
Add prior to use:	1 mM	DTT

30% Glycerol buffer (storage at 4°C)

	20 mM	Tris-HCl pH 7.4
	0.1 mM	EDTA pH 8.0
	30%	Glycerol (v/v)

10% Glycerol buffer (storage at 4°C)

	20 mM	Tris-HCl pH 7.4
	0.1 mM	EDTA pH 8.0
	10%	Glycerol (v/v)

MATERIAL AND METHODS

2.4.8.3.2 Serial dilution-based reconstitution of mononucleosomes

The reconstitution of mononucleosomes via serial dilution is in comparison to section 2.4.8.3.1 a quick assembly method where less protein is required. Nevertheless, separating free DNA from mononucleosomes is not possible with that method. Buffers were always freshly prepared and stored at RT. Reactions, containing 0.85 µg of octamer and 1 µg of DNA, were set up in initial dilution buffer with a final salt concentration of 2 M NaCl and a total volume of 10 µl. Initially, the reaction was incubated for 15 min at 37°C and subsequently transferred to 30°C. Then, samples were diluted every 15 min at 30°C with the following initial dilution buffer volumes: 3.3 µl, 6.7 µl, 5.0 µl, 3.6 µl, 4.7 µl, 6.7 µl, 10 µl, 30 µl and 20 µl. Finally, mononucleosomes were resuspended in 100 µl of final dilution buffer and incubated for 15 min at 30°C. The assembly was checked by loading 10 µl of reconstituted mononucleosomes and 10 ng of free DNA (control) on a 5% native PAA gel. Concentrations were determined by quantification with Image studio lite version 5.2 (Licor). Mononucleosomes were stored at 4°C for at least one year.

Initial dilution buffer (prepare fresh)

10 mM	HEPES pH 7.9
1 mM	EDTA pH 8.0
0.5 mM	PMSF

Final dilution buffer (prepare fresh)

10 mM	Tris-HCl pH 7.5
5 mM	EDTA pH 8.0
20%	Glycerol (v/v)
0.1%	NP-40 (v/v)
0.5 mM	PMSF
100 µg/ml	BSA fraction V (w/v)

2.4.9 *In vitro* binding assays with recombinant PWWP2A constructs

Recombinantly expressed GST-PWWP2A full-length and deletion constructs were kept on glutathione sepharose beads after purification and were used for *in vitro* binding assays with

MATERIAL AND METHODS

recombinant mononucleosomes or with mononucleosomes derived from HeLa Kyoto cells. Protease inhibitor mix was added to all buffers. Protein interactions between recombinant PWWP2A constructs and mononucleosomes were analyzed by SDS-PAGE followed by Coomassie-staining or immunoblotting.

2.4.9.1 *Binding assays with HeLa Kyoto cell-derived mononucleosomes*

Mononucleosomes were prepared from HeLa Kyoto cells as described in 2.4.4.. 25 µl of recombinant GST-tagged PWWP2A_{fl} or PWWP2A deletion constructs, bound to glutathione sepharose beads, were thawed on ice and washed 2 x 5 min with PBS. After PBS was removed, proteins were incubated ON with mononucleosomes derived from 1 x 10⁷ cells at 4°C while rotating. Then, beads were pelleted by mild centrifugation (2.000 rpm, 4°C, tabletop centrifuge) and supernatant was discarded. For all washing steps (4 x 10 min, 4°C), a reduced EX100 buffer (150 mM or 300 mM NaCl) was used. The first two washes were done with reduced EX100 supplemented with 0.1% NP-40 and the last two washes with reduced EX100 without NP-40. Proteins were eluted by boiling them in 50 µl Laemmli buffer for 5 min at 95°C. Protein samples were checked by SDS-PAGE followed by Coomassie-staining or immunoblotting.

Reduced EX100 (storage at 4°C)

	10 mM	HEPES pH 7.6
	150 mM	NaCl
	1.5 mM	MgCl ₂
Add prior to use:	1 mM	DTT
		protease inhibitor mix (1:1.000)

2.4.9.2 *Binding assays with recombinant mononucleosomes*

In addition to HeLa Kyoto cell-derived mononucleosomes, recombinantly expressed and *in vitro* reconstituted mononucleosomes were used for interaction studies. Here, 25 µl of glutathione sepharose bead linked protein were incubated ON with 1.5 µg of recombinant mononucleosomes in 500 µl EX100 buffer at 4°C on a rotating wheel. Washing steps, elution and analysis of protein interactions were conducted as described in 2.4.9.1..

MATERIAL AND METHODS

2.4.9.3 Electrophoretic Mobility Shift Assay (EMSA)

Recombinantly expressed and reconstituted mononucleosomes, coupled with fluorescent Cy3- or Cy5-labels were used for *in vitro* binding studies with eluted, GST-tagged PWWP2A full-length and deletion constructs. All working steps were carried out on ice and all buffers were ice-cooled. Gels were always pre-run in 0.2x TBE for 1h at 110 V and pockets were flushed with running buffer.

2.4.9.3.1 EMSAs with nucleic acids

Protein-DNA interactions were examined by applying EMSAs. Recombinant GST-PWWP2A proteins were diluted in protein dialysis buffer to a final concentration of 5 μ M. Next, according to experimental settings and conditions, an initial protein concentration (e.g. 800 nM) was set up by mixing concentrated GST-protein stock with EMSA binding buffer. In order to test several protein concentrations for their binding affinities, a 1:1 dilution series was performed by combining 10 μ l of previous protein concentration with 10 μ l of EMSA binding buffer. Single-stranded DNA or RNA containing a 601 sequence (DNA: CCCGGTGCCGAGGCCGCTCAATTGG; RNA: CCCGGUGCCGAGGCCGCUCAAUUGG) and labeled at the 5' end either with Cy5 or Cy3 were hybridized by mixing them 1:1 with their complementary DNA or RNA oligonucleotides. To hybridize nucleic acids, they were boiled at 95°C for 5 min and slowly cooled down ON. Double- and single-stranded DNA or RNA (concentration 50 μ M) were diluted to 20 nM in EMSA binding buffer and 7.5 μ l of nucleic acids were added to 7.5 μ l of GST-PWWP2A construct dilutions. Binding reactions were performed on ice for 25 min and analyzed on an 8% native PAA gel (110 V, 105 min, on ice, running buffer: 0.2x TBE) and following visualized with the Typhoon FLA9500 (GE).

EMSA Binding Buffer RNA/DNA (storage at 4°C)

	1x	TE
	50 mM	NaCl
	10%	Glycerol (v/v)
	0.1%	BSA (v/v)
	0.03 ng/ μ l	ytRNA
Add prior to use:	0.5 mM	DTT

MATERIAL AND METHODS

2.4.9.3.2 Competitive EMSAs (cEMSAs)

Proteins can have different affinities for distinct binding partners. In order to clarify the preference of certain GST-protein truncations for a respective prey, competitive EMSAs (cEMSAs) were applied. Differently Cy-labeled (Cy3 or Cy5) recombinant mononucleosomes were combined and each diluted to 20 nM in EMSA binding buffer. 7.5 µl of serial protein dilutions (see section 2.4.9.3.1) were mixed with 7.5 µl of differently Cy-labeled recombinant mononucleosomes and incubated for 25 min on ice. Analysis of protein interaction was done running a 5% PAA gel (110 V, 105 min, on ice, running buffer: 0.2x TBE) followed by Typhoon FLA9500 (GE) detection.

EMSA Binding Buffer GST-PWWP2A constructs (storage at 4°C)

	1x	TE
	50 mM	NaCl
	10%	Glycerol (v/v)
	0.1%	BSA (v/v)
Add prior to use:	0.5 mM	DTT

Table 2.13 Differently Cy-labeled recombinant mononucleosomes

Type	linker length	Label	Species	Source
WT	20 bp/20 bp	Cy5	homo sapiens	www.histonesource.com
WT	0 bp/0 pb	Cy3	homo sapiens	www.histonesource.com
H2A.Z-containing	20 bp/20 bp	Cy3	homo sapiens	www.histonesource.com
H2A.Z-containing	0 bp/0 pb	Cy5	homo sapiens	www.histonesource.com
H3Kc36me0	20 bp/20 bp	Cy3	Xenopus laevis	Michaela Smolle lab
H3Kc36me3	20 bp/20 bp	Cy5	Xenopus laevis	Michaela Smolle lab
H3Kc36me0	0 bp/0 bp	Cy3	Xenopus laevis	Michaela Smolle lab
H3Kc36me3	0 bp/0 bp	Cy5	Xenopus laevis	Michaela Smolle lab
H2A TL	20 bp/20 bp	Cy3	homo sapiens	www.histonesource.com
H2A TL	0 bp/0 bp	Cy5	homo sapiens	www.histonesource.com

MATERIAL AND METHODS

H2A.Z TL	20 bp/20 bp	Cy5	homo sapiens	www.histonesource.com
H2B TL	20 bp/20 bp	Cy3	homo sapiens	www.histonesource.com
H2B TL	0 bp/0 bp	Cy5	homo sapiens	www.histonesource.com
H3 TL	20 bp/20 bp	Cy3	homo sapiens	www.histonesource.com
H3 TL	0 bp/0 bp	Cy5	homo sapiens	www.histonesource.com
H4 TL	20 bp/20 bp	Cy3	homo sapiens	www.histonesource.com
H4 TL	0 bp/0 bp	Cy5	homo sapiens	www.histonesource.com

2.4.10 NMR spectroscopy

NMR spectroscopy was done in collaboration with the Mackay lab (School of Life and Environmental Sciences, University of Sydney, New South Wales 2006, Australia). Competent *E.coli* Rosetta (DE3) pLysS cells were transformed with either S_PWWP WT or point mutant construct encoding plasmids and cultured at 37°C while shaking. Log-phase cultures (OD 0.4–0.8) were subsequently induced by adding 0.3 mM IPTG and incubated for 18 h at 18°C while shaking. Cell pellets containing overexpressed proteins were lysed by shock freezing (liquid N₂) followed by sonication and purified using GSH affinity chromatography. GST (N-terminal) was removed by incubating GSH elution fractions with HRV-3C proteases (Mackay lab) at 4°C. Fractions were concentrated and one-dimensional ¹H NMR spectra recorded on a Bruker 600 MHz NMR spectrometer at 25 °C.

2.5 Bioinformatics

2.5.1 MS Data analysis

Data were analyzed by Moritz Völker-Albert (Imhof group, Molecular Biology, LMU Munich, Germany) applying Skyline (version 3.6) [310]. Double and triple charged peptide masses were used for extracted ion chromatograms (XICs). Integrated, manually selected peak values (Total Area MS1) were exported as csv.file for further calculations. The ratio of structural modified peptides to the sum of all isotopical similar peptides defined the percentage of each modification within the same peptide. Total Area MS1 value was used to determine the relative abundance of a detected modified peptide as percentage of overall peptide. Three

MATERIAL AND METHODS

unique MS2 fragment ions were used to quantify coeluting isobaric modifications. Averaged integrals of these ions were used to determine their respective contribution to the isobaric MS1 peak (e.g. H3K27me2K36me1 and H3K36me3). The percentage value of each modification was normalized to the percentage value of “input”. Values were plotted with Perseus software version 1.5.1.6 with euclidean clustering and subsequent visualization in heatmaps [311]. For the analysis in Perseus default settings were used.

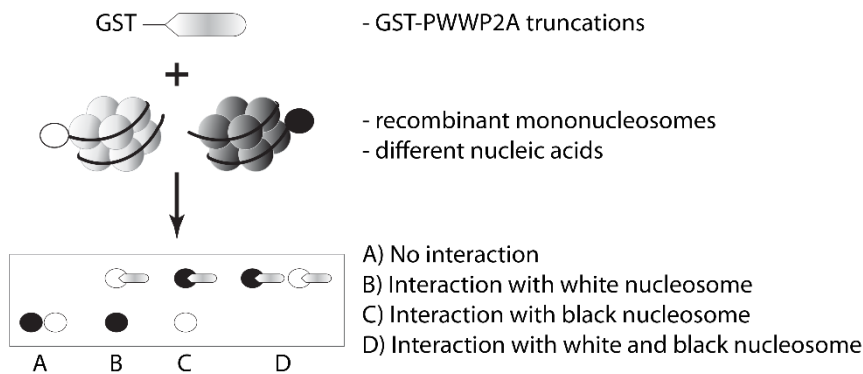
3 Results

3.1 PWWP2A is a direct nucleosome interaction protein

To analyze the multivalent binding properties of the recently in our lab identified H2A.Z interactor PWWP2A [276] in detail, an *in vitro* competitive electrophoretic mobility shift assay (cEMSA) system (**Figure 3.1.1**), based on differently fluorescent labeled tags (**Figure 3.1.1A**), was established. In this approach, Cy3- or Cy5-labeled 147 or 187 bp DNA fragments were wrapped around different octamers to obtain mononucleosomes with distinct properties like incorporated histone variants, single histone tail deletions, missing linker DNA, present linker DNA or special modifications like H3Kc36me3 (**Figure 3.1.1B**) which were finally distinguishable in a laser-based detection system (Typhoon FLA9500). The different fluorescent tags were stimulated by special excitation wavelengths of certain lasers (excitation wavelength Cy3: 550 nm and Cy5: 650 nm) and enabled to perform cEMSAs with 1:1 mixtures of different mononucleosome combinations and increasing concentrations of the desired recombinant GST-tagged PWWP2A proteins. Thereby, GST-PWWP2A full-length or truncation proteins can favor *in vitro* a specific bait and clear statements about preferences and affinities can be made. Additionally, in case of free DNA being present after recombinant mononucleosome reconstitutions, cEMSA mixtures always contain same amounts of free nucleic acids for each construct instead of having different concentrations when separated experimental approaches are performed and are thus better comparable. Hence, using the described *in vitro* cEMSA system is a convincing procedure to elucidate PWWP2A's various interaction profiles together with a plethora of differently recombinantly reconstituted mononucleosomes.

RESULTS

A



B

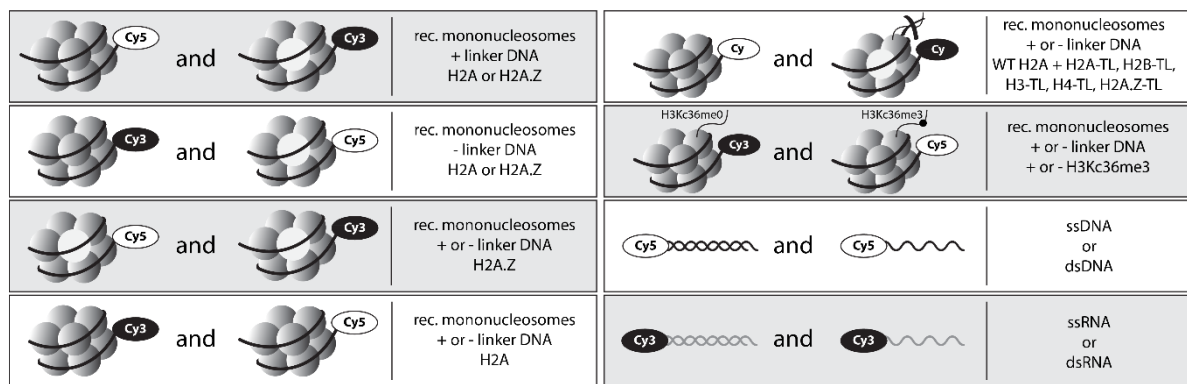


Figure 3.1.1 Schematic depiction of EMSA/cEMSA and different recombinant mononucleosomes or nucleic acids. GST-tagged PWWP2A and protein truncation specificities for distinct recombinant mononucleosomes and nucleic acids were analyzed applying an *in vitro* EMSA/cEMSA approach. (A) Basic information on the cEMSA approach. GST-PWWP2A and protein truncations were incubated with a mix of differently Cy3- or Cy5-labeled recombinant mononucleosomes and visualized after native PAGE with a laser-based system (Typhoon FLA9500). A: no shift, no interaction; B: shift of white but not black nucleosome, specific interaction with white nucleosome; C: shift of black but not white nucleosome, specific interaction with black nucleosome; D: shift of black and white nucleosome, interaction with both nucleosomes. (B) Schematic overview of used recombinant mononucleosomes and nucleic acids.

Recently, we have shown that PWWP2A is specifically enriched for GFP-H2A.Z over GFP-H2A using If-qMS [276]. cEMSA were applied to clarify whether PWWP2A's interaction with chromatin in general and H2A.Z in particular is direct or indirect and whether it discriminates *in vitro* between canonical H2A and variant H2A.Z as well as different properties of individually reconstituted mononucleosomes. To do so, recombinant human histones H2A, H2A.Z, H2B, H3 and H4 (**Figure 3.1.2A**) were assembled to either H2A- or H2A.Z-containing octamers (**Figure 3.1.2B**) and finally reconstituted to mononucleosomes by salt-gradient dialysis [312,

RESULTS

313] using 147 bp or 187 bp DNA fragments (Widom 601-sequence [314]) with different fluorescent tags.

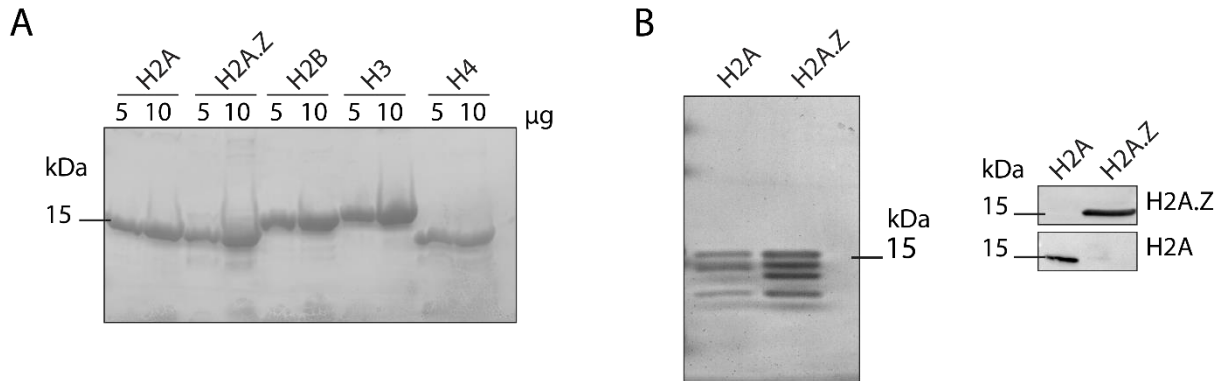


Figure 3.1.2 Recombinant canonical histones and histone variant H2A.Z and respective octamers. (A) Coomassie-stained SDS-PAGE gel of purified recombinant canonical human histones H2A, H2B, H3 and H4 and histone variant H2A.Z. Canonical histones and histone variant H2A.Z were ordered from: www.histonesource.com (B) Coomassie-stained SDS-PAGE gel of H2A- or H2A.Z-containing octamers (left). Immunoblotting with commercially available antibodies against H2A or H2A.Z (right) independently confirms the respective incorporated H2A variant in assembled octamers.

GST-tagged PWWP2A full-length protein (**Figure 3.1.3A**) was purified from transformed, competent BL21 *E.coli* cells, eluted from glutathione sepharose beads (**Figure 3.1.3B**) and subsequently used in cEMSAs. To test whether GST-PWWP2A binds to recombinant mononucleosomes and distinguishes between H2A and H2A.Z *in vitro*, cEMSAs with H2A- and H2A.Z-containing mononucleosomes containing 20 bp of linker DNA (20 – Θ – 20) and increasing GST-protein concentrations (0 nM - 200 nM) were performed (**Figure 3.1.3C**). Indeed, GST-PWWP2A was able to shift H2A- and H2A.Z-containing mononucleosomes, whereby a stronger binding for H2A.Z- in comparison to H2A-nucleosomes was observed at a GST-PWWP2A protein concentration of 60 nM (**Figure 3.1.3C**), indicating a direct discrimination between H2A and H2A.Z. Moreover, to investigate whether nucleosome properties like DNA length influence the direct binding strength of GST-PWWP2A to recombinant mononucleosomes, cEMSAs using H2A.Z-containing mononucleosomes without (0 – Θ – 0) or with (20 – Θ – 20) linker DNA were performed (**Figure 3.1.3D**). Notably, mononucleosomes containing linker DNA were preferably bound at GST-PWWP2A concentrations of 150 nM and 200 nM in comparison to mononucleosomes without linker DNA (**Figure 3.1.3D**), suggesting an additional binding modus to linker DNA.

RESULTS

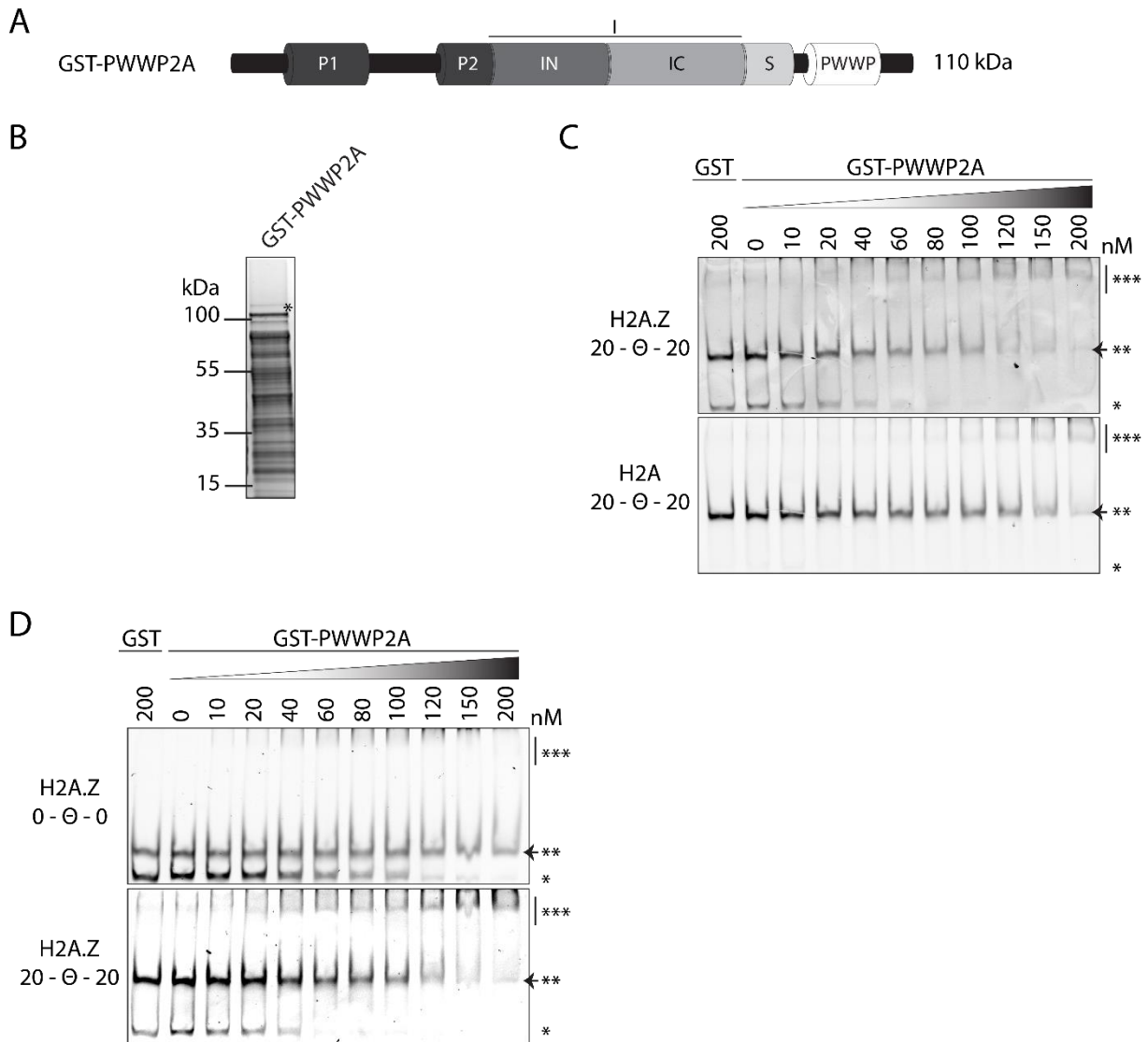


Figure 3.1.3 PWWP2A is a direct nucleosome binder and distinguishes between H2A and H2A.Z *in vitro*. (A) Schematic depiction of N-terminal GST-tagged PWWP2A full-length with a predicted molecular weight of 110 kDa. P1 and P2: proline rich regions, IN: N-terminal part of internal region, IC: C-terminal part of internal region, I: combination of IN and IC, S: serine-rich stretch, PWWP: Proline – Tryptophan – Tryptophan – Proline domain. (B) Coomassie-stained SDS-PAGE gel of purified and eluted GST-PWWP2A used for cEMSA. *marks purified GST-PWWP2A. Additional bands are GST-protein degradation products. (C) Representative cEMSA (n=2) using 15 nM of recombinant human H2A.Z- (top) or H2A- (bottom) containing mononucleosomes with linker DNA (20–Θ–20) incubated with indicated increasing concentrations of GST-PWWP2A. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-PWWP2A complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-PWWP2A complex is formed. (D) Representative cEMSA (n=2) similar to (C) using 15 nM of recombinant H2A.Z-containing mononucleosomes without (0–Θ–0) (top) and with (20–Θ–20) (bottom) linker DNA incubated with indicated increasing concentrations of GST-PWWP2A.

3.2 PWWP2A's internal region divides labor – general nucleosome binding and specific H2A.Z recognition

After having determined PWWP2A's ability of recombinant mononucleosome interaction, the next step was to find out which domains of PWWP2A mediate the association with chromatin

RESULTS

in more detail. Nina Kronbeck, a former master student of our lab, previously identified the internal region (I), which is located between the proline-rich region P2 and the serine-rich stretch S as well as the C-terminal PWWP domain being sufficient to pull-down HK cell-derived mononucleosomes by using several protein deletion constructs. Thus, I focused on these domains to evaluate how their precise interaction with nucleosomes works. To find out, which part is contributing to the pull-down of HK cell-derived mononucleosomes, the internal region was further subdivided into an N- (aa 292 - 422) and a C-terminal part (aa 423 - 574) (**Figure 3.2.1A**). To dissect which part leads to the specific recognition of H2A.Z, HK cell-derived and MNase-digested mononucleosomes (**Figure 3.2.1B**) were incubated with recombinant, on glutathione-sepharose bead linked GST-I, GST-IN and GST-IC and immunoprecipitations (IPs) were either visualized by Coomassie-staining or immunoblotting. Intriguingly, GST-IN efficiently pulled-down nucleosomes as indicated by SDS-PAGE separated histone bands, whereas GST-IC showed only low pull-down efficiency (**Figure 3.2.1C**). Interestingly, immunoblots using an H2A.Z-specific antibody revealed a strong enrichment in H2A.Z signal intensity for GST-IC but not GST-IN (**Figure 3.2.1C**) pull-downs. Moreover, the H2A.Z signal intensity of GST-I displayed a combination of both GST-IN as well as GST-IC (**Figure 3.2.1C**) pull-down approaches, proposing different binding functions of I's subparts.

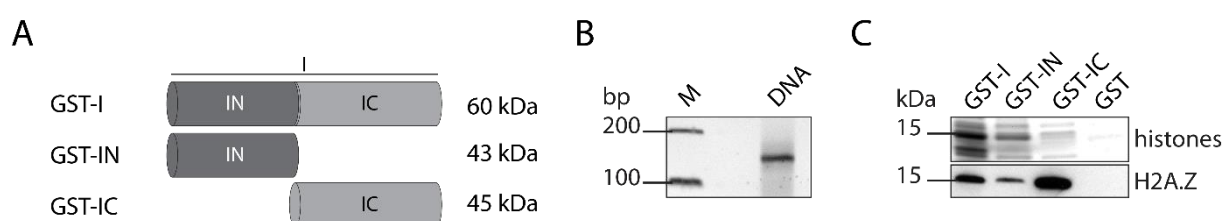


Figure 3.2.1 The N-terminal part of the internal region efficiently binds nucleosomes in general, while the C-terminal part shows a specificity for H2A.Z-nucleosomes. (A) Schematic depiction of GST-tagged internal region GST-I (60 kDa, top), GST-IN (43 kDa, middle) and GST-IC (45 kDa, bottom). (B) MNase-digested mononucleosomes were prepared from HK cell nuclei. DNA extraction followed by subsequent agarose gel electrophoresis and ethidium bromide staining revealed almost pure mononucleosomes (~150 bp). (C) Pull-downs of recombinant GST-I, GST-IN or GST-IC (n=2) with HK-derived mononucleosomes. GST served as negative control. Precipitated nucleosomes as indicated by histone bands (top) are detected with Coomassie Blue staining after SDS-PAGE, and H2A.Z levels (bottom) in immunoblots with commercially available antibodies. Notice enriched signal intensity in H2A.Z for GST-IC in comparison to GST-IN.

To determine, whether IN is able to directly bind to mononucleosomes in general and IC to H2A.Z-containing mononucleosomes in particular or whether the interaction is rather

RESULTS

indirectly mediated by posttranslational histone modifications or PWWP2A-associated interaction factors, *in vitro* binding assays (cEMSAs) were performed. Therefore, mononucleosomes with H2A- or H2A.Z-containing 187 bp of either Cy3- or Cy5-labeled DNA fragments were assembled (see also section 3.1) and recombinant GST-I, GST-IN or GST-IC (**Figure 3.2.1A**) were expressed in BL21 *E.coli* cells and purified with help of a glutathione-sepharose column and the Äcta Pure system (GE) (**Figure 3.2.2A**). Remarkably, already at a GST-I protein concentration of 10 nM more than half of all H2A.Z-containing nucleosomes were bound, whereas H2A-containing nucleosomes were shifted to a lesser extent at same protein concentrations (**Figure 3.2.2B**). To shift equal H2A-nucleosome amounts as observed for GST-I and H2A.Z-containing nucleosomes at 10 nM, a GST-I protein concentration of minimum 40 nM was needed, indicating a preference of GST-I for H2A.Z-nucleosomes *in vitro* (**Figure 3.2.2B**). GST-IN showed no difference in H2A.Z- or H2A-nucleosome binding and shifted both nucleosome types equally well at all protein concentrations (10 nM, 20 nM, 40 nM and 60 nM). This is also in line with previous results using cell-derived mononucleosomes, where GST-IN showed no enrichment in H2A.Z-nucleosomes, suggesting rather a general binding mechanism (**Figure 3.2.2B**). Interestingly, GST-IC bound already at low protein concentrations (20 nM) almost exclusively to H2A.Z-containing mononucleosomes. In contrast, H2A-nucleosomes were only slightly shifted at high GST-IC concentrations (60 nM), hinting towards IC's function in specific H2A.Z-recognition (**Figure 3.2.2B**). Moreover, comparing the distinct nucleosome binding profiles of GST-I to GST-IN and GST-IC, the internal region as a whole displays a combination of its two subparts (**Figure 3.2.2B**).

RESULTS

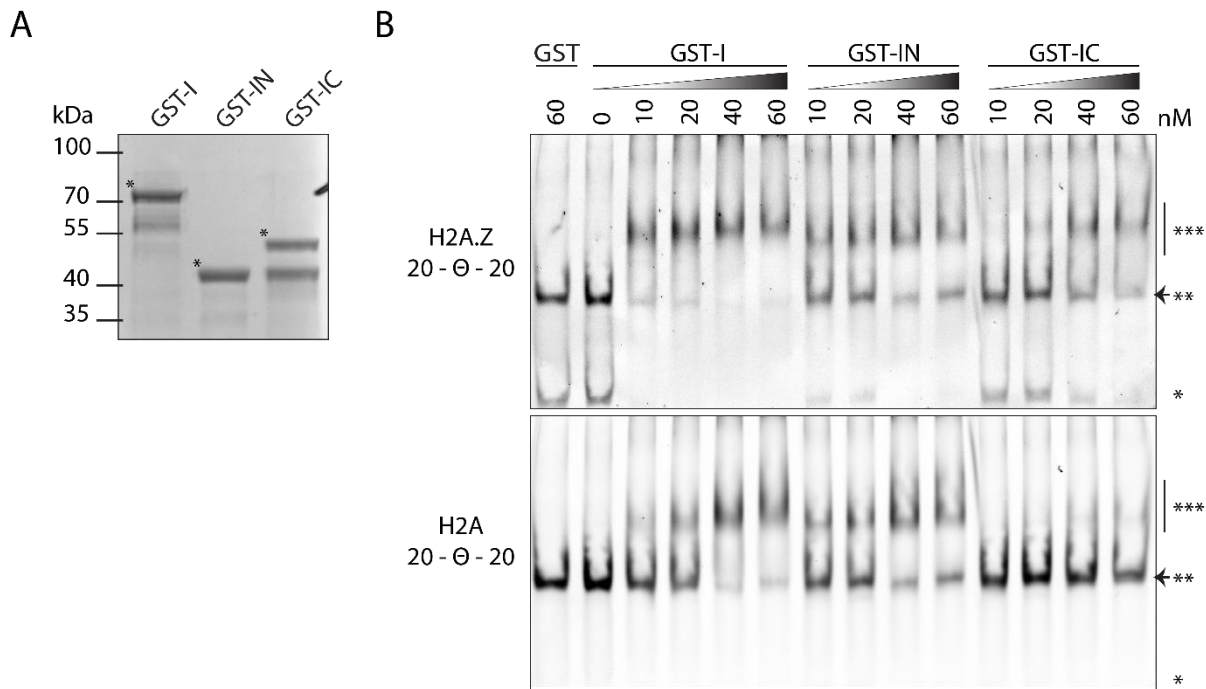


Figure 3.2.2 The internal region and its subparts directly bind to recombinant mononucleosomes. (A) Coomassie-stained SDS-PAGE gel of purified GST-I, GST-IN and GST-IC used for cEMSA. *marks purified GST-tagged constructs. Additional bands are GST-protein degradation products. **(B)** Representative cEMSA (n=3) using 15 nM of recombinant human H2A.Z- (top) or H2A- (bottom) containing mononucleosomes with linker DNA (20-θ-20) incubated with indicated increasing concentrations of GST-I, GST-IN and GST-IC. GST served as negative control. *free DNA, **nucleosome, ***nucleosome-GST-construct complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-protein complex is formed.

In summary, (1) IN and (2) IC divide their labor into (1) general nucleosome recognition without variant preference and (2) specific H2A.Z recognition thus contributing to a strong and direct interaction with chromatin.

3.3 H2A.Z's C-terminus affects binding to PWWP2A in a HK cell-derived system

The previous data indicate a strong and specific interaction between the C-terminal part of PWWP2A's internal region and the histone variant H2A.Z. H2A.Z's C-terminus harbors a so-called acidic patch that has been proposed to provide a binding platform for H2A.Z-specific interaction proteins [212]. In order to determine the part of H2A.Z mediating IC specificity, HK cells were transiently transfected with FLAG-tagged chicken H2A, H2A.Z or C-terminal deleted H2A.Z (H2A.ZΔC) (kind gift of our collaborator Masahiko Harata (Graduate School of

RESULTS

Agricultural Science/Faculty of Agriculture, Tohoku University, Sendai, Japan), constructs described in [315]) (**Figure 3.3.1A**). Next, pull-downs of recombinant GST-PWWP2A (**Figure 3.3.3A**) and mononucleosomes prepared from either transiently FLAG-H2A-, -H2A.Z- or -H2A.ZΔC-expressing HK cells were carried out. GST-PWWP2A was able to pull-down each FLAG-histone protein complex as observed in anti-FLAG immunoblots (**Figure 3.3.1B**). In line with previous findings, GST-PWWP2A pull-downs showed enriched immunoblot signaling for FLAG-H2A.Z compared to FLAG-H2A (**Figure 3.3.1B, C**). Interestingly, GST-PWWP2A pull-down approaches with FLAG-H2A.ZΔC, which lacks nine unique C-terminal amino acids that usually keep the extended acidic patch intact, showed a decreased immunoblot signal intensity compared to FLAG-H2A.Z but similar intensities like FLAG-H2A (**Figure 3.3.1B, C**). In conclusion, H2A.Z's C-terminal histone tail, including parts of the extended acidic patch, are partially needed for the interaction with PWWP2A.

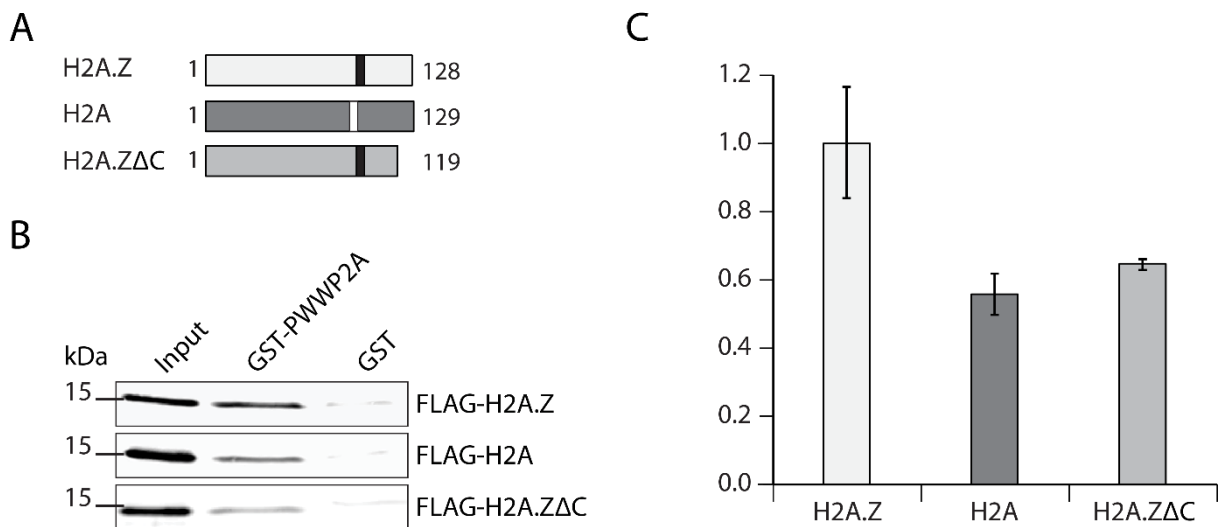


Figure 3.3.1 C-terminal H2A.Z histone tail depletion impairs PWWP2A nucleosome binding. (A) Schematic depiction of chicken H2A, H2A.Z and H2A.Z C-terminal deleted (H2A.ZΔC) proteins. Acidic patch of H2A is indicated with white and extended acidic patch of H2A.Z with black boxes. (B) GST-PWWP2A and GST were incubated with mononucleosomes derived from transiently transfected (FLAG-constructs) HK cells. Binding efficiency and variant specificity were tested by SDS-PAGE separation followed by immunoblots with commercially available anti-FLAG antibody. GST served as negative control. Shown is one representative blot of four independent biological replicates. (C) Quantification of immunoblot signal intensities shown in (B) using Image Studio Lite Ver 5.2 (LI-COR). Error bars indicate standard error of the mean (SEM) of four independent biological replicates. H2A.Z was normalized to 1.0, while H2A and H2A.ZΔC are illustrated in x-fold change compared to H2A.Z. Y-axis: immunoblot signal intensity, x-axis: distinct histone proteins.

RESULTS

After I could demonstrate the importance of H2A.Z's C-terminus for binding to PWWP2A in a cell-derived system, I wondered whether the interaction is mediated in a direct manner or rather indirect by additional factors like modifications or further interaction components. Therefore, cEMSA using recombinant mononucleosomes containing either H2A.Z wildtype (WT) or H2A.Z tail-depleted (tail-less, TL) (**Figure 3.3.2A**) incubated with indicated increasing amounts of GST-PWWP2A (0 nM - 160 nM) and GST-IC (0 nM - 80 nM) were carried out. GST-PWWP2A bound both, H2A.Z WT- and H2A.Z TL-containing mononucleosomes equally well, whereby band shifts were only observed at GST-protein concentrations of 80 nM and 160 nM (**Figure 3.3.2B**). In concordance, GST-IC bound to both nucleosome types with similar binding affinities, whereby protein-protein complex formations were already detected for GST-IC protein concentrations of 40 nM (instead of minimum 80 nM observed for PWWP2A full-length) (**Figure 3.3.2C**). In line with PWWP2A full-length shifts, the H2A.Z-specific domain IC showed no difference in recombinant H2A.Z WT- or H2A.Z TL-nucleosome binding, indicating no direct interaction of PWWP2A with H2A.Z's seven unique C-terminal amino acids.

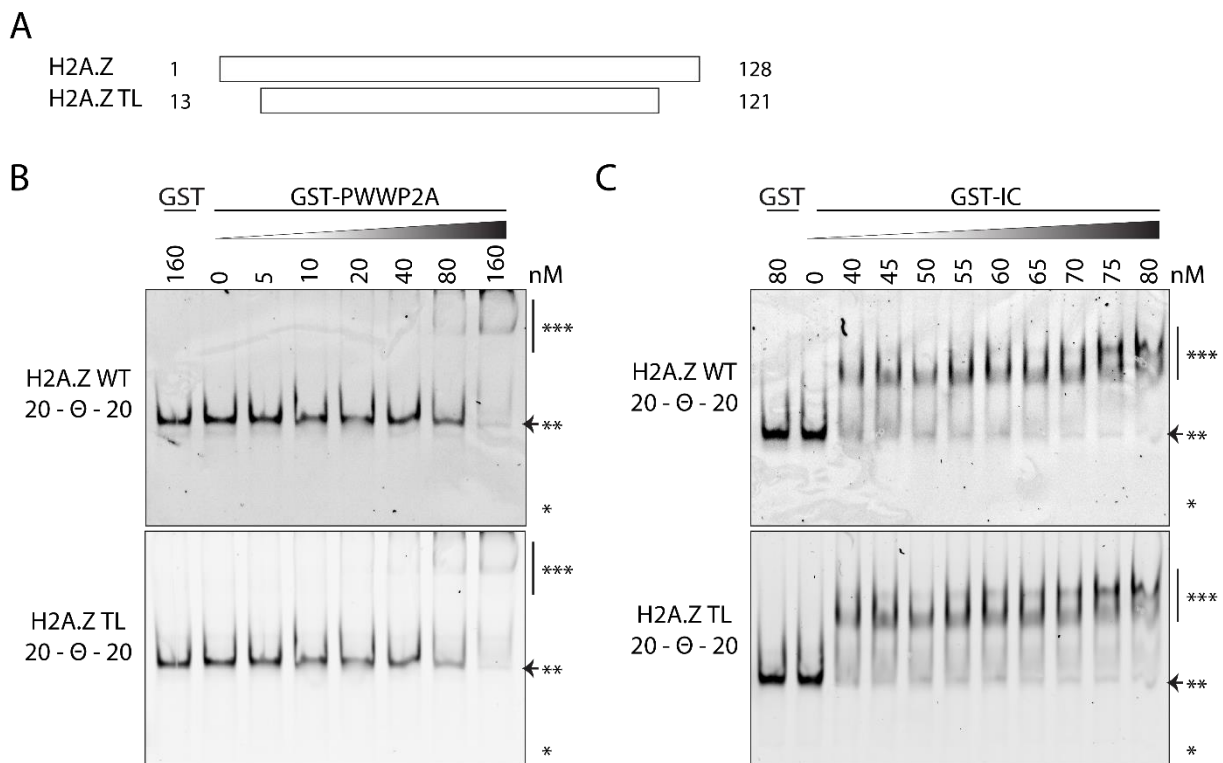


Figure 3.3.2 H2A.Z tails are not required for *in vitro* binding of nucleosomes to PWWP2A and IC. (A) Schematic depiction of human H2A.Z WT (top) as well as N- and C-terminal tail depleted H2A.Z tail-less mutant (TL, bottom). TL octamers were ordered from: www.histonesource.com (**B**) Representative cEMSA (n=2) using 15 nM of recombinant human H2A.Z WT- (top) or H2A.Z TL- (bottom) containing

RESULTS

mononucleosomes with linker DNA (20–0–20) incubated with indicated increasing concentrations of GST-PWWP2A. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-PWWP2A complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-PWWP2A complex is formed. (C) Representative cEMSA (n=2) similar to (B) using 15 nM of recombinant human H2A.Z WT- (top) or H2A.Z TL- (bottom) containing mononucleosomes with linker DNA (20–0–20) incubated with indicated increasing concentrations of GST-IC.

In summary, H2A.Z's nine unique C-terminal amino acids play a crucial role in the binding to PWWP2A in a cell-derived system, whereas *in vitro* neither for PWWP2A full-length nor its IC domain alone an effect upon N- and C-terminal H2A.Z tail-depletion was observed.

3.4 Point mutations within IC impair binding to H2A and H2A.Z

The previous data discussed the influence of H2A.Z's C-terminus on the binding to PWWP2A. Here, I focused on PWWP2A's IC domain and possible amino acid residues that could mediate the specific recognition of H2A.Z. Since other proteins are known to be involved in specific protein-protein interactions due to conserved domains and amino acid residues [316], I initially checked the evolutionary conservation of the IC domain. The web browser-based alignment tools MUSCLE and LALIGN were used to determine the degree of evolutionary conservation between different organisms (mammals and amphibians). Indeed, IC is highly conserved between different organisms (**Figure 3.4.1A**) and shows high amino acid identities when the human primary sequence (*Homo sapiens*) was compared to mouse (*Mus musculus*, 84.1%), cow (*Bos taurus*, 91.4%), chicken (*Gallus gallus*, 88.8%) and frogs (*Xenopus laevis* and *Xenopus tropicalis*, each 74.3%) (**Figure 3.4.1B**), suggesting a conserved function of IC within different species.

RESULTS

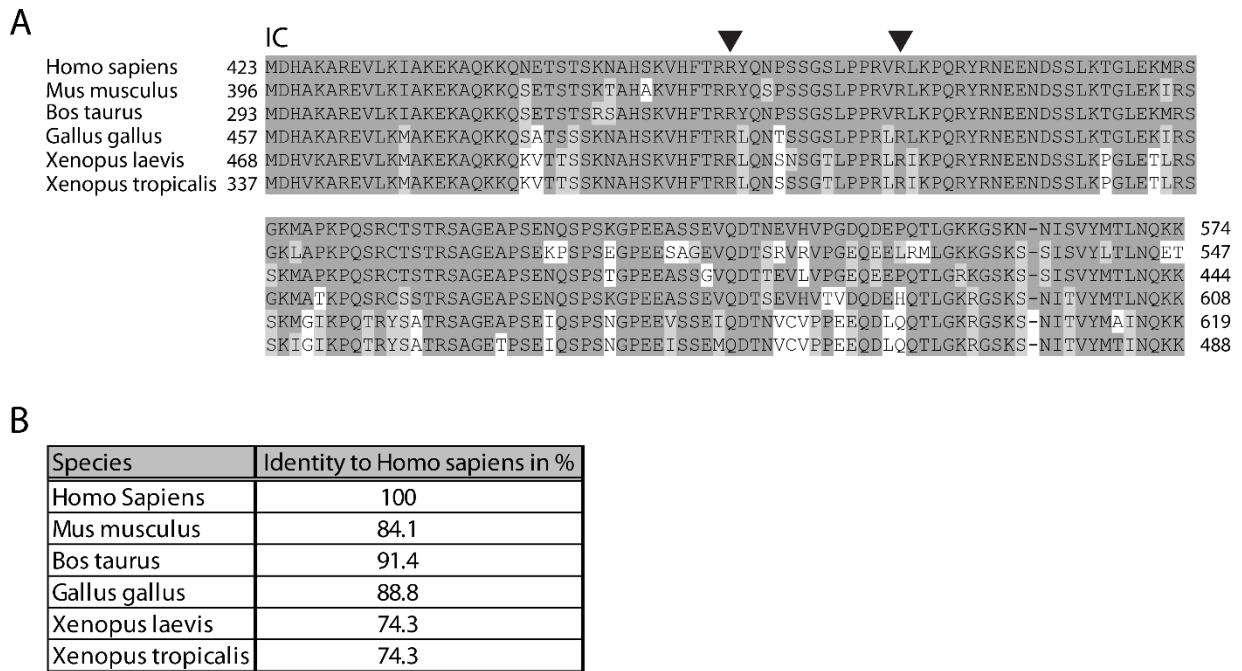


Figure 3.4.1 PWWP2A's IC region is highly conserved between different organisms. (A) Amino acid sequence alignment of the C-terminal part (IC) of PWWP2A's internal region from the indicated species. IC sequences of *Homo sapiens* [Q96N64], *Mus musculus* [Q69Z61], *Bos taurus* [F1MKS1], *Gallus gallus* [XP_015149148.1], *Xenopus laevis* [XeXenL6RMv1.0002517m, Xenbase] and *Xenopus tropicalis* [XM_002940175, Xenbase] were aligned using the web browser-based MUSCLE multiple sequence alignment tool (default settings). Identical amino acids are depicted in dark grey, similar amino acid residues in light grey and changes are highlighted in white. Black arrowheads mark conserved amino acid residues that were found to be frequently mutated in several types of cancer (www.cbioportal.org) **(B)** Amino acid identities of indicated species in comparison to *Homo sapiens*. Calculations and alignments were carried out using the web browser-based LALIGN tool on the SIB Bioinformatics Resource Portal based on [317].

In order to identify amino acid residues that might be necessary for H2A.Z recognition, screens of the cancer database cBioPortal (www.cbioportal.org) revealed two point mutations within the IC region of PWWP2A that were frequently mutated in different types of tumors [318, 319] (**Figure 3.4.2**). Arginine at position 461 is exchanged to glutamic acid (R461Q) and was found in rectal adenocarcinoma, uterine endometrioid carcinoma and metastatic melanoma whereas arginine at position 475 is exchanged to cysteine (R475C) and was only identified in metastatic melanoma [318, 319]. Chiara Vardabasso from Emily Bernsteins group (Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA) already demonstrated that deregulated H2A.Z levels play a role in melanoma development and drug response [320]. Hence, I wondered whether the substitutions of R461 or R475 lead to an impaired binding of PWWP2A to H2A.Z and might therefore be one new contributor in cancer development.

RESULTS

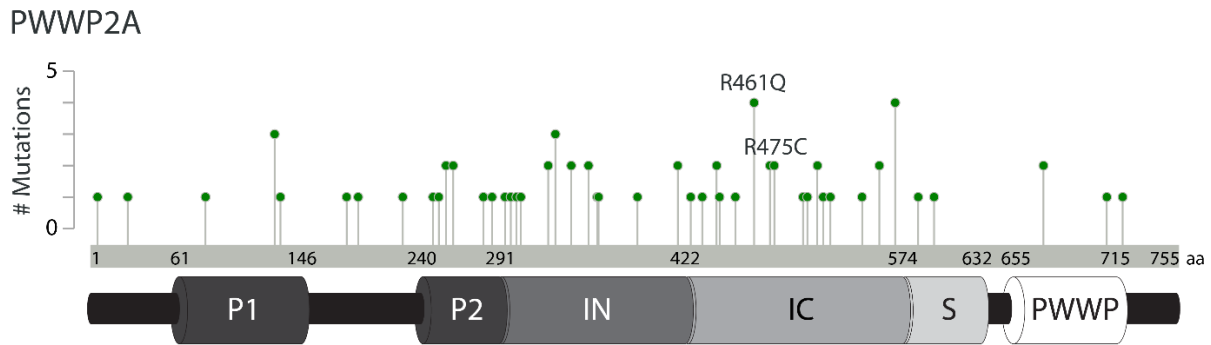


Figure 3.4.2 PWWP2A's IC domain harbors residues that are frequently mutated in cancer. Different cancer types show single point mutations within the human PWWP2A protein. The Y-axis displays the number of mutations detected in distinct cancers for different studies. Five independent studies revealed amino acid exchanges for aa 461 screening >3000 participants, whereas aa 475 was found to be mutated in two independent studies screening >300 participants. The X-axis displays amino acid positions within PWWP2A, whereby R461Q and R475C are located in IC. Green mutation diagram circles indicate missense mutations. Schematic was taken from: www.cbioportal.org and slightly modified.

To test, whether the exchange of these two conserved amino acid residues (**Figure 3.4.1A**) leads to an altered binding of the internal region to H2A.Z, GST-I containing R461Q or R475C (generated by site-directed mutagenesis) (**Figure 3.4.3A**) and wildtype GST-I were expressed in competent BL21 *E.coli* cells. Subsequent, proteins were purified and eluted by using a glutathione sepharose column and the Äcta Pure System (GE) (**Figure 3.4.3B**). Following, purified GST-tagged proteins (0 nM - 100 nM) were incubated with recombinant human H2A- and H2A.Z-containing mononucleosomes (20 – 0 – 20) and cEMSA were performed. Like previously reported, GST-I bound to H2A.Z-nucleosomes already at low GST-protein concentrations (40 nM), whereby all H2A.Z-containing nucleosomes were shifted at 80 nM. By contrast, higher GST-I protein concentrations were required (between 60 nM and 80 nM) to shift H2A-containing nucleosomes (**Figure 3.4.3C**). Compared to GST-I, the point mutant GST-I_R461Q showed similar H2A.Z-nucleosome binding. Interestingly, GST-I_R461Q appears to loose binding to H2A-nucleosomes (compare with shift at 80 nM for GST-I), while the second mutant GST-I_R475C showed highly decreased binding for H2A.Z- as well as H2A-containing mononucleosomes for all indicated concentrations (40 nM - 100 nM), suggesting both amino acid residues being important for nucleosome interaction *in vitro*.

RESULTS

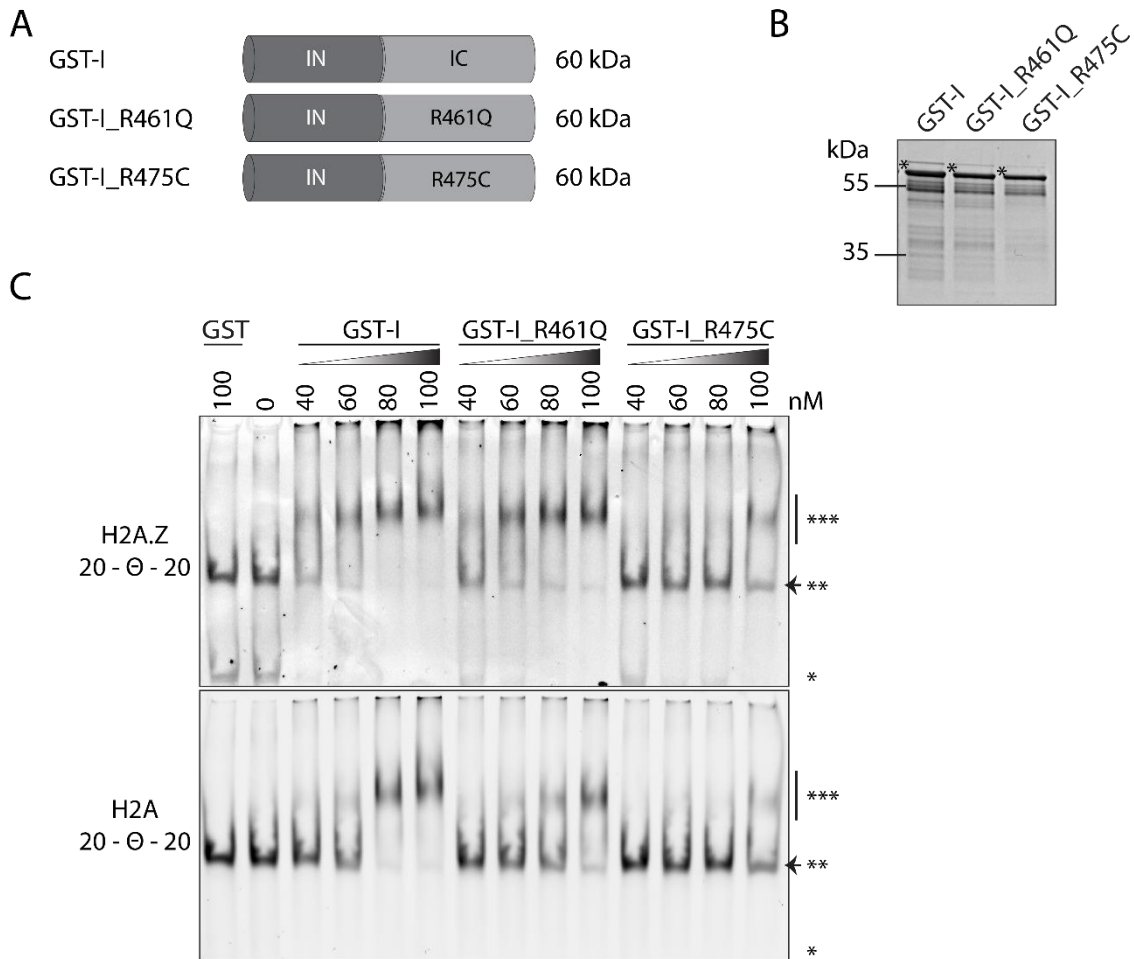


Figure 3.4.3 Recombinant single IC point mutants show impaired binding to recombinant mononucleosomes. (A) Schematic depiction of GST-tagged internal region GST-I (60 kDa, top), GST-I_R461Q (60 kDa, middle) and GST-I_R475C (60 kDa, bottom). (B) Coomassie-stained SDS-PAGE gel of purified GST-I, GST-I_R461Q and GST-I_R475C used for cEMSA. *marks purified GST-tagged constructs. Additional bands are GST-protein degradation products. (C) Representative cEMSA (n=2) using 15 nM of recombinant human H2A.Z- (top) or H2A- (bottom) containing mononucleosomes with linker DNA (20- Θ -20) incubated with indicated increasing concentrations of GST-I, GST-I_R461Q and GST-I_R475C. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-protein complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-protein complex is formed.

In conclusion, while nine unique amino acid residues of H2A.Z's C-terminus mediate the binding to PWWP2A in a cell-derived system, two amino acid residues in PWWP2A's internal region, observed to be mutated in different tumors, contribute to the highly affine interactions with H2A- and H2A.Z-containing mononucleosomes *in vitro*.

RESULTS

3.5 Single histone tail deletions do not impair binding of IN *in vitro*

The previous results unraveled IC as a direct H2A.Z recognition domain with specific amino acid residues being involved in this interaction. Next, I analyzed the binding mechanism of the direct but unspecific nucleosome-interaction domain IN. In general, chromatin-binding proteins can have different mechanisms mediating nucleosome interactions [293, 297, 316]. Hence, I wondered whether the histones and/or the DNA itself might contribute to IN's interaction with nucleosomes in general. As many proteins recognize the flexible tails of nucleosomal histones rather than the hardly accessible histone core regions [113], a series of different recombinant, single histone tail-lacking mononucleosomes ("tail-less", TL) (**Figure 3.5.1A, B**) was tested in cEMSA. Here, distinct TL nucleosomes were assembled using a serial dilution step-based protocol (see also section 2.4.8.3.2) with either 147 bp (0 – Θ – 0) or 187 bp (20 – Θ – 20) Cy-labeled DNA fragments (Widom sequence) [314].

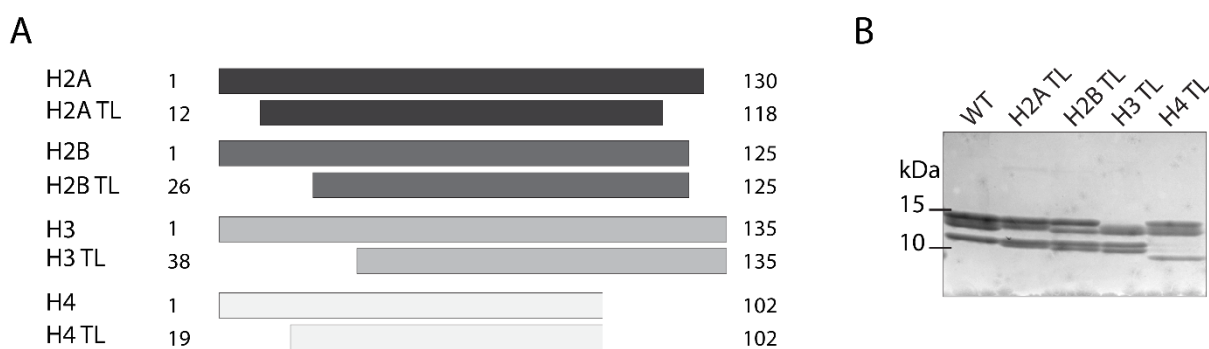


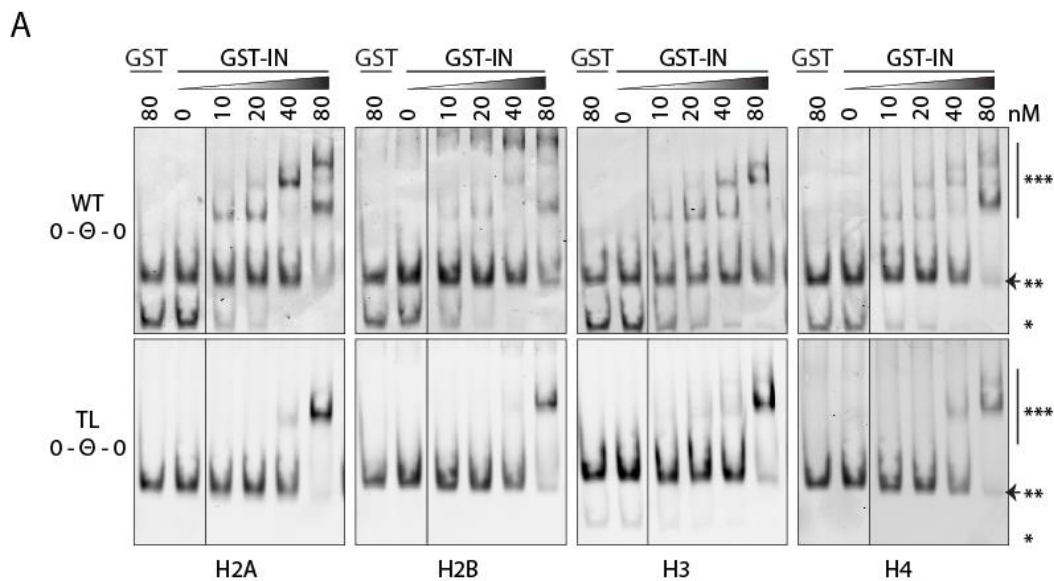
Figure 3.5.1 Single histone tail deleted octamers. (A) Schematic depiction of human H2A, H2B, H3 and H4 wildtype (top) as well as tail-depleted H2A, H2B, H3 and H4 tail-less (TL, bottom) mutants. Note that H2A is N- and C-terminally deleted whereas all other canonical histones are only lacking parts of the N-terminus. (B) Coomassie-stained SDS-PAGE gel of H2A TL, H2B TL, H3 TL and H4 TL-containing octamers in comparison to H2A wildtype octamer. TL octamers and wildtype histones were ordered from: www.histonesource.com

Next, increasing amounts of recombinant and purified GST-tagged IN (0 nM - 80 nM) (**Figure 3.2.2A**) were incubated with WT or different single TL mononucleosomes without (**Figure 3.5.2A**) or with (**Figure 3.5.2B**) linker DNA. GST-IN bound to all recombinant linker-less mononucleosomes, regardless if single histone tails were depleted or not, whereby a minimum protein concentration of 80 nM was sufficient (**Figure 3.5.2A**). Interestingly, similar observations were made for mononucleosomes containing linker DNA (**Figure 3.5.2B**). Here,

RESULTS

GST-IN also interacted with all TL nucleosomes, but reduced GST-protein concentrations (10 nM) were necessary to shift linker DNA-containing nucleosomes compared to those lacking it (80 nM) (**Figure 3.5.2B**), concluding no single histone tail is required for the interaction of GST-IN with recombinant mononucleosomes. Interestingly, GST-IN showed better binding to nucleosomes lacking the H3 histone tail for both, mononucleosomes with or without flanking DNA (**Figure 3.5.2A, B**). Interestingly, biochemical studies of tail-less mononucleosomes have revealed that removal of the N-terminal H3 histone tail can increase the accessibility of nucleosomal DNA [321], suggesting a role of IN as DNA binding module. In line with this finding, additional bands were detected for cEMSA with free DNA being present, assuming a putative binding of IN to free nucleic acids (**Figure 3.5.2A**).

For this reason, IN was subsequently tested in further *in vitro* studies to find out whether it is able to bind directly to DNA.



RESULTS

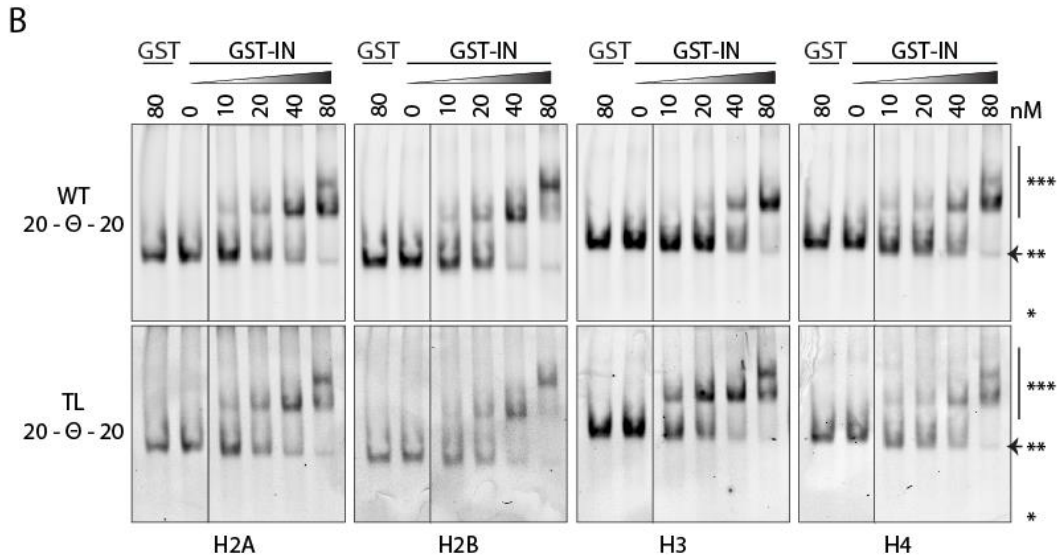


Figure 3.5.2 Single histone tails are not required for the binding of IN to mononucleosomes *in vitro*. (A) Representative cEMSA (n=2) using 15 nM of recombinant human wildtype mononucleosomes (WT, top) or mononucleosomes lacking single histone tails (TL, bottom) without linker DNA (0-θ-0) incubated with indicated increasing concentrations of GST-IN. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-IN complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-IN complex is formed. (B) Representative cEMSA (n=2) similar to (A) using 15 nM of recombinant human wildtype mononucleosomes (WT, top) or mononucleosomes lacking single histone tails (TL, bottom) with linker DNA (20-θ-20) incubated with indicated increasing concentrations of GST-IN.

3.6 IN interacts with free DNA and binds nucleosomal linker DNA

The previous results demonstrated that single histone tails are not required for the interaction of IN with mononucleosomes. In addition, band shifts of free nucleic acids were observed, leading to the question whether IN is a DNA binding domain. Sequence alignments showed no homology to other protein DNA binding motifs, but high levels of conservation between amphibians and mammals (Figure 3.6.1A). In fact, IN of mouse, cow and chicken have an amino acid identity of more than 90%, whereas amphibians (*X. tropicalis* and *X. laevis*) are more diverse in their primary sequence compared to mammals, but still show identities of more than 70%, pointing towards a conserved function of this protein region (Figure 3.6.1B).

RESULTS

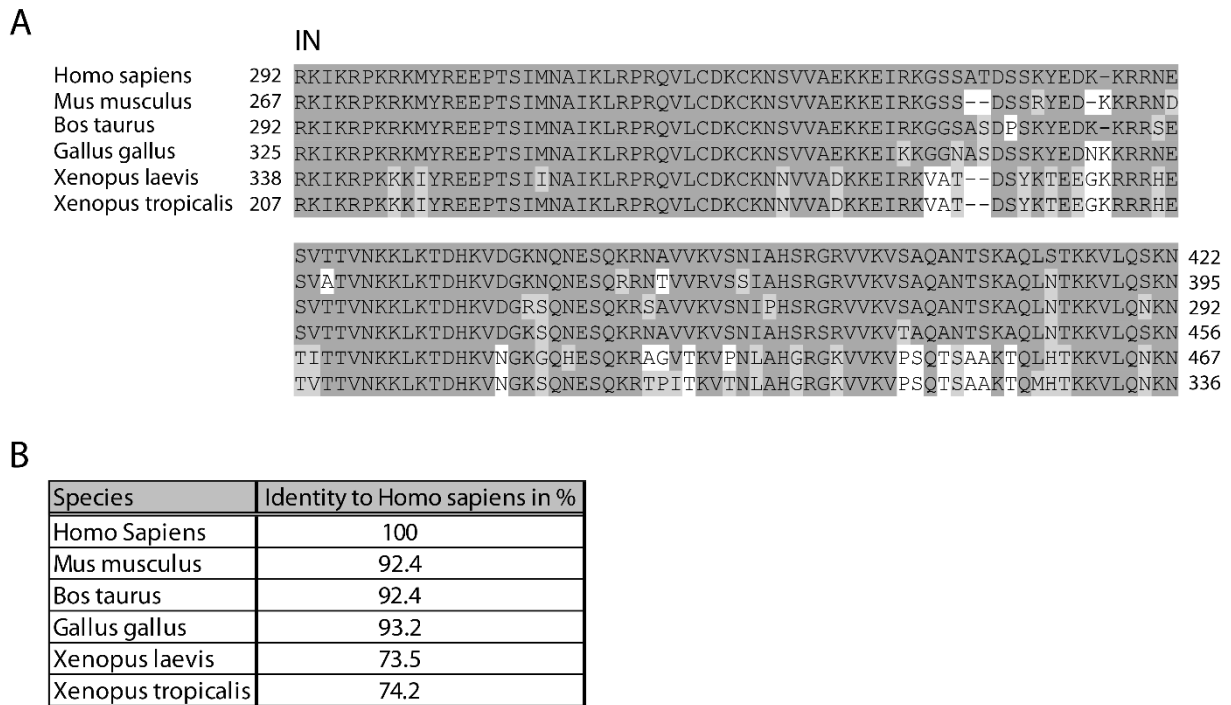


Figure 3.6.1 PWWP2A's IN region is highly conserved between different organisms. (A) Amino acid sequence alignment of the N-terminal part (IN) of PWWP2A's internal region from the indicated species. IN sequences of *Homo sapiens* [Q96N64], *Mus musculus* [Q69Z61], *Bos taurus* [F1MKS1], *Gallus gallus* [XP_015149148.1], *Xenopus laevis* [XeXenL6RMv1.0002517m, Xenbase] and *Xenopus tropicalis* [XM_002940175, Xenbase] were aligned using the web browser-based MUSCLE multiple sequence alignment tool (default settings). Identical amino acids are depicted in dark grey, similar amino acid residues in light grey and changes are highlighted in white. **(B)** Amino acid identities of indicated species in comparison to *Homo sapiens*. Calculations and alignments were carried out using the web browser-based LALIGN tool on the SIB Bioinformatics Resource Portal based on [317].

In addition to the degree of evolutionary conservation, the isoelectric point (PI) of GST-IN and further PWWP2A constructs (GST-PWWP2A, GST-IC and GST-PWWP) was calculated (**Figure 3.6.2A**). As there are acidic and basic amino acid residues within a protein, the net surface charge is pH-dependent [322]. The PI therefore indicates the pH at which a molecule or protein equals a surface charge of zero [323]. In case of having a pH lower than the PI under acidic conditions, basic amino acids are ionized while the ionization of carboxyl groups is inhibited resulting in an overall positive net charge of the protein surface [322]. In return, the protein receives a negative charge under basic conditions as carboxyl groups are ionized and amino groups remain neutral. Therefore, proteins with similar/identical PIs show an equal binding potential due to their surface charge dependent on the surrounding pH. GST-IN as well as GST-PWWP2A, GST-IC and GST-PWWP possess all PI values around 9 and the proportion of positively charged amino acids ranges similarly between ~14-18% (**Figure 3.6.2A**). To clarify,

RESULTS

whether GST-IN is able to bind DNA, EMSAs with 187 bp of Cy5-labeled DNA fragments were performed. Surprisingly only IN but not IC was able to efficiently bind DNA (**Figure 3.6.2B**) although they have similar PIs and harbor a certain amount of positively charged amino acid residues (**Figure 3.6.2A**).

A

Construct	% of positively charged residues	Isoelectric point
GST-PWWP2A	13.9	8.67
GST-IN	18.1	9.47
GST-IC	15.5	8.92
GST-PWWP	14.5	9.08

B

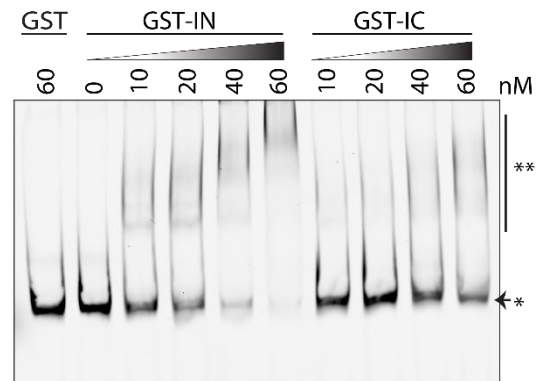


Figure 3.6.2 IN binds to free DNA. (A) Biochemical properties of different indicated PWWP2A protein regions. Computation of various physical and chemical parameters for indicated proteins were carried out applying the web-browser based ProtParam tool of the SIB Bioinformatics Resource Portal (www.expasy.org) on individually entered protein sequences. (B) Representative EMSA (n=2) using Cy5-tagged 187 bp dsDNA fragment and indicated concentrations of GST-IN and GST-IC. *free dsDNA, **DNA-GST-protein complex. Arrowhead marks loss of band signal intensity when DNA-GST-protein complex is formed.

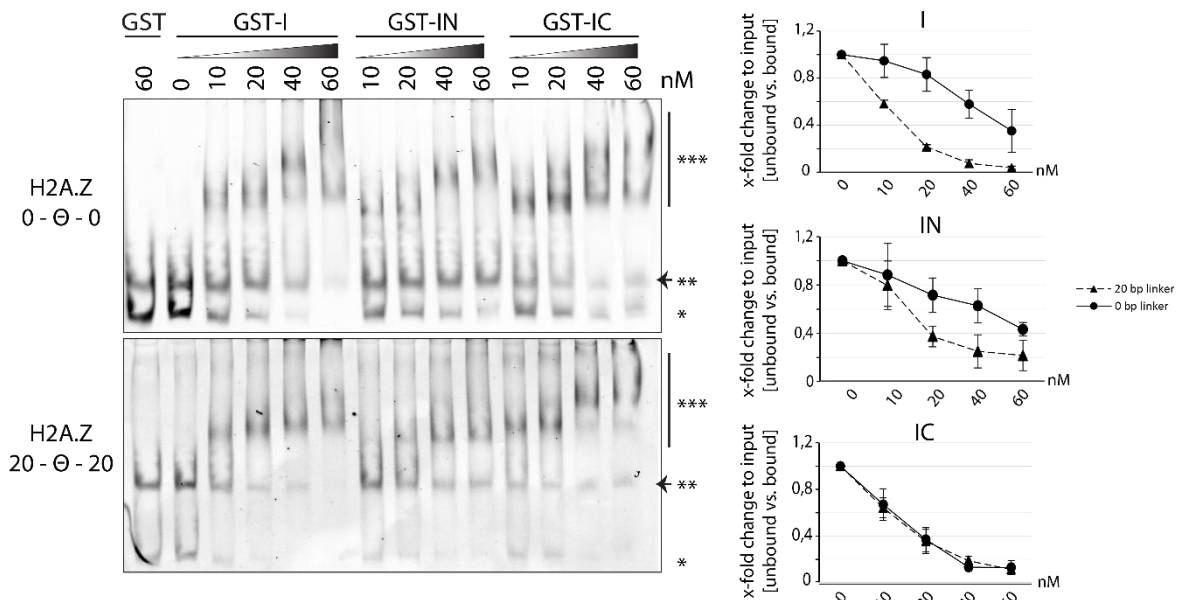
After I demonstrated the binding properties of IN to free DNA, the next step was to elucidate whether IN can also recognize histone-wrapped nucleosomal DNA, since other proteins such as pioneer transcription factors are known to bind inaccessible chromatin regions [324]. To do so, cEMSA with indicated increasing amounts of recombinant GST-I, GST-IN and GST-IC (0 nM - 60 nM) (**Figure 3.2.2A**) together with H2A.Z- or H2A-containing mononucleosomes were performed (**Figure 3.6.3A, B**). In each cEMSA, mononucleosomes displayed or lacked a 20 bp linker 601-Widom DNA sequence [314]. Interestingly, GST-I shifted more than half of the H2A.Z-nucleosomes containing 20 bp linker DNA already at a minimum GST-protein concentration of 10 nM, while linker-less mononucleosomes required therefore higher protein concentrations (60 nM) (**Figure 3.6.3A**). GST-IN showed similar results in H2A.Z-nucleosome binding like observed for GST-I, whereby linker-containing mononucleosomes were already bound at a GST-protein concentration of minimum 40 nM (**Figure 3.6.3A**). In contrast, linker-less H2A.Z-mononucleosomes were only slightly bound at concentrations of

RESULTS

40 nM and 60 nM (**Figure 3.6.3A**). Moreover, GST-IC showed no differences in certain nucleosome-interactions and bound equally well to linker-containing and linker-less mononucleosomes at minimum protein concentrations of 10 nM (**Figure 3.6.3A**). In addition, similar results were observed for all recombinant GST-tagged proteins (I, IN and IC) applying the same approach with H2A- instead of H2A.Z-nucleosomes (**Figure 3.6.3B**), suggesting IN is rather interacting with free linker DNA, than inaccessible, nucleosomal wrapped DNA as e.g. observed for pioneer transcription factors [324].

In summary, the IN domain is a strong and direct nucleosome binding region mediating the interaction with chromatin mostly via free flanking DNA.

A



RESULTS

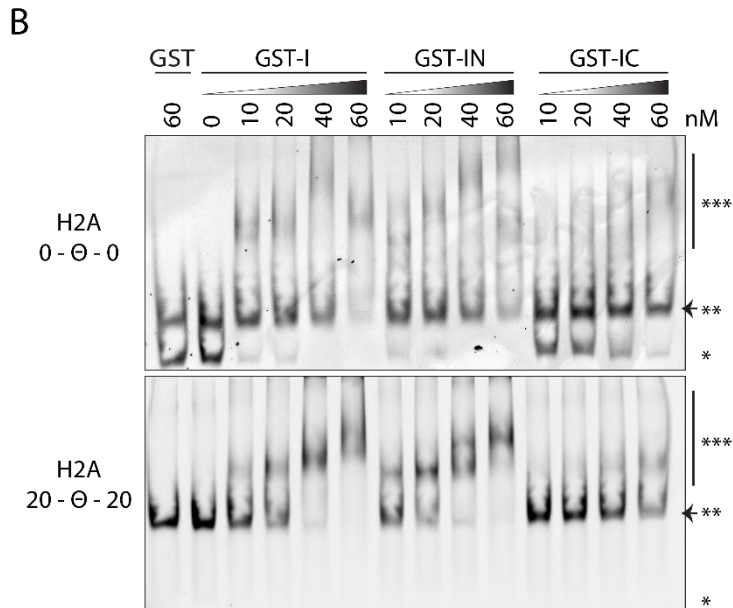


Figure 3.6.3 Nucleosomal linker DNA affects binding of I and IN. (A) Left: Representative cEMSA (n=3) using 15 nM of recombinant human linker-less H2A.Z (0-Θ-0, top) or linker-containing H2A.Z (20-Θ-20, bottom) mononucleosomes incubated with indicated increasing concentrations of GST-I, GST-IN and GST-IC. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-protein complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-protein complex is formed. Right: Quantification of signal intensities by analyzing bound versus unbound mononucleosomes (**) using Image Studio Lite Ver 5.2 (LI-COR). Error bars indicate standard error of the mean (SEM) of three independent replicates. (B) Representative cEMSA (n=2) similar to (A) using 15 nM of recombinant human linker-less H2A (0-Θ-0, top) or linker-containing H2A (20-Θ-20, bottom) mononucleosomes incubated with indicated increasing concentrations of GST-I, GST-IN and GST-IC.

3.7 PWWP2A's PWWP domain directly interacts with recombinant mononucleosomes

Besides the internal region, that is able to specifically interact with H2A.Z and nucleosomes in general by recognition of free/linker DNA, PWWP2A harbors another chromatin interaction domain in its very C-terminus – PWWP. PWWP domains of other PWWP domain-containing proteins are known to have distinct binding mechanisms to chromatin and are usually weakly conserved among different proteins [289, 293]. However, PWWP2A's PWWP domain shows high levels of evolutionary conservation between mammals and amphibians (Figure 3.7.1A) and harbors amino acid sequence identities up to 96.7% compared to the human protein (Figure 3.7.1B). Furthermore, previous experiments have already demonstrated its ability to pull-down HK cell-derived mononucleosomes [276].

RESULTS

A

		PWWP	
Homo sapiens	655	VGDIVWAKIYGFPWWPARILTIIVSRKDNGLLVRQEARISWFGSPTTSFLALSQLSPFLEN	715
Mus musculus	630	VGDIVWAKIYGFPWWPARILTIIVSRKDNGLLARQEARISWFAASPTTSSLALSQLSPFLEN	690
Bos taurus	653	VGDIVWAKIYGFPWWPARILTIIVSRKDSGILVRQEARISWFGSPTTSFLALSQLSPFLEN	713
Gallus gallus	687	VGDIVWAKIYGFPWWPARILTIIVSRKDNGLLVRQEARISWFGSPTTSLALSQLSPFLEN	747
Xenopus laevis	684	VGDIVWAKIYGFPWWPARILAIITIVSRKDNGLLVRQEARISWFGSPTTSFLALSQLAAPFLEN	744
Xenopus tropicalis	684	VGDIVWAKIYGFPWWPARILAIITIVSRKDTGILLVRQEARISWFGSPTTSFLALSQLTTPFLEN	744

B

Species	Identity to Homo sapiens in %
Homo Sapiens	100
Mus musculus	95.1
Bos taurus	96.7
Gallus gallus	96.7
Xenopus laevis	93.4
Xenopus tropicalis	95.1

Figure 3.7.1 PWWP2A's PWWP domain is highly conserved between different organisms. (A) Amino acid sequence alignment of PWWP2A's C-terminal PWWP domain from the indicated species. PWWP domain sequences of *Homo sapiens* [Q96N64], *Mus musculus* [Q69Z61], *Bos taurus* [F1MKS1], *Gallus gallus* [XP_015149148.1], *Xenopus laevis* [XeXenL6RMv1.0002517m, Xenbase] and *Xenopus tropicalis* [XM_002940175, Xenbase] were aligned using the web browser-based MUSCLE multiple sequence alignment tool (default settings). Identical amino acids are depicted in dark grey, similar amino acid residues in light grey and changes are highlighted in white. **(B)** Identities of indicated species in comparison to *Homo sapiens*. Calculations and alignments were carried out using the web browser-based LALIGN tool on the SIB Bioinformatics Resource Portal based on [317].

To determine whether PWWP2A's PWWP domain is able to interact directly with recombinant mononucleosomes in general and whether it distinguishes between H2A and H2A.Z *in vitro*, GST-PWWP (**Figure 3.7.2A**) was expressed in competent BL21 *E.coli* cells, purified using a glutathione sepharose column (**Figure 3.7.2B**) and subsequently used in cEMSA. Recombinant mononucleosomes containing either canonical H2A (20 – Θ – 20) or histone variant H2A.Z (20 – Θ – 20) were therefore incubated with increasing concentrations (0 nM - 300 nM) of GST-PWWP. Intriguingly, GST-PWWP concentrations of 50 nM were sufficient to shift mononucleosomes containing either H2A or H2A.Z, whereby a GST-protein concentration of 250 nM resulted in the binding of almost all free mononucleosomes (**Figure 3.7.2C**). In addition, GST-PWWP bound equally well to H2A- or H2A.Z-containing mononucleosomes, suggesting the domain to have none preference for histone content and rather general nucleosome interaction features.

RESULTS

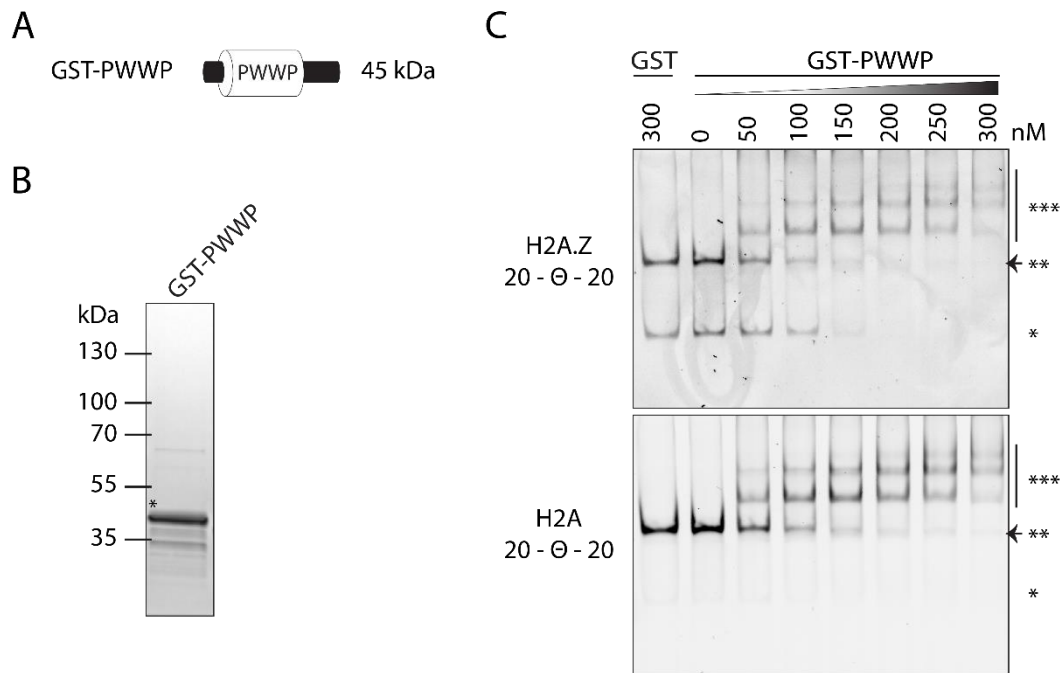


Figure 3.7.2 PWWP2A's PWWP domain directly binds to recombinant H2A and H2A.Z mononucleosomes. (A) Schematic depiction of N-terminal GST-tagged PWWP domain with a predicted molecular weight of 45 kDa. (B) Coomassie-stained SDS-PAGE gel of purified and eluted GST-PWWP used for cEMSA. *purified GST-PWWP. Additional bands are GST-protein degradation products. (C) Representative cEMSA (n=2) using 15 nM of recombinant human H2A.Z- (top) or H2A- (bottom) containing mononucleosomes with linker DNA (20-θ-20) incubated with indicated increasing concentrations of GST-PWWP. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-PWWP complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-PWWP complex is formed.

3.8 The PWWP domain interacts with free nucleic acids and nucleosomal linker DNA

After having determined the ability of PWWP2A's PWWP domain to interact with recombinant mononucleosomes, I wondered how this domain can associate with chromatin. PWWP domains are known to recognize histone posttranslational modifications and/or nucleic acids like DNA [293, 325, 326]. To see whether the PWWP domain displays a potential DNA-interaction surface profile, Sebastian Pünzeler modeled the 3D structure of the PWWP domain of PWWP2A [276] based on the PWWP domain solution structure of PWWP2B [123], that is a close homolog to PWWP2A (**Figure 3.8.1**). Here, PSIP1's PWWP domain represents the experimentally established and published electrostatic surface potential (ESP) of the PWWP domain [294] and serves as comparison to PWWP2A's PWWP domain. Indeed, the solvent-

RESULTS

accessible surface revealed an electrostatic potential due to the enrichment in basic residues, suggesting a possible DNA-binding ability (**Figure 3.8.1**).

In silico

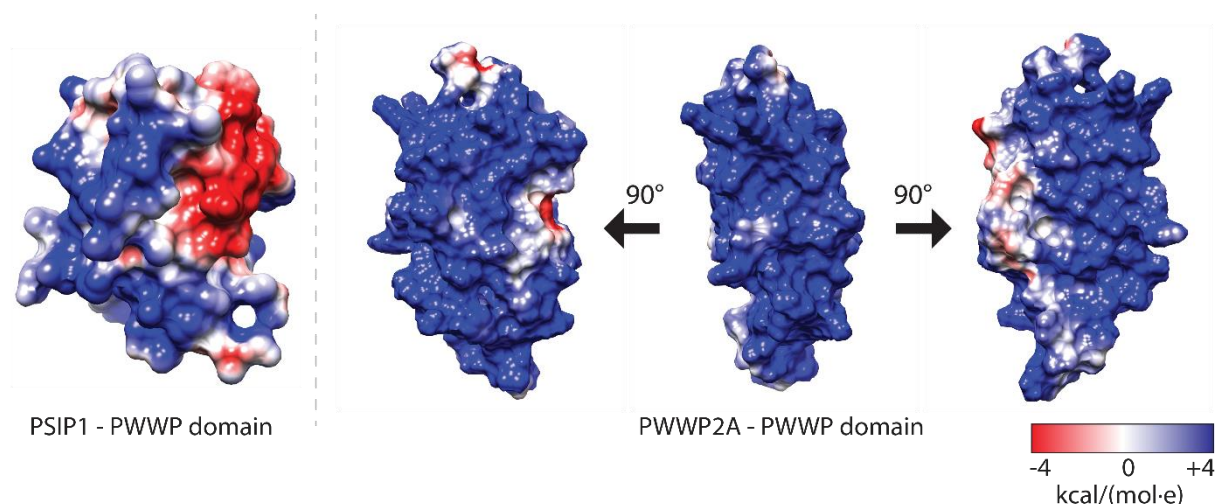


Figure 3.8.1 PWWP2A's surface potential is highly positive charged. Left: the electrostatic surface potential (ESP) of the already published PSIP1 DNA-binding PWWP domain [294] was computed and served as a positive control. Right: ESP model of PWWP2A's PWWP domain was based on the published 3D structure of its close homolog PWWP2B [123]. Shown are three representative views of the surface potential of PWWP2A's PWWP domain. ESP color values are in units of kcal/(mol·e) at 298K. Red color indicates negatively charged surface potential and blue color a positively charged surface potential. Figure adapted and reprinted from [276], with permission of John Wiley and Sons.

To determine a possible DNA and/or RNA interaction of PWWP2A's PWWP domain or any preference for certain nucleic acids, GST-tagged recombinant PWWP protein (**Figure 3.7.2B**) was incubated with distinct nucleic acids (ssDNA, dsDNA, ssRNA and dsRNA) in EMSAs. First, it was investigated whether the PWWP domain interacts in general with nucleic acids. Here, 75 bp dsDNA (50 nM) were incubated with indicated increasing concentrations of recombinant GST-PWWP (0 nM - 4500 nM) (**Figure 3.8.2A**). GST-PWWP showed DNA binding upon a minimum protein concentration of 500 nM, whereby a maximum GST-protein concentration of 4500 nM was not sufficient to shift all free nucleic acids (**Figure 3.8.2A**), suggesting only a low affine interaction of PWWP with nucleic acids in general and dsDNA in particular. To investigate, whether PWWP2A's PWWP domain preferably binds a certain type of nucleic acids (ss or ds, RNA or DNA) or whether the interaction is mainly achieved via its electrostatic surface potential [294, 295, 327], 25 bp ss/ds RNA/DNA (15 nM) were incubated with indicated increasing concentrations of GST-PWWP (0 nM - 1200 nM) (**Figure 3.8.2B**).

RESULTS

Notably, GST-PWWP shifted all kinds of nucleic acids equally well, whereby minimum GST-protein concentrations of 200 nM were sufficient to enable nucleic acid binding (**Figure 3.8.2B**). Moreover, at protein concentrations of 800 nM and 1200 nM all free nucleic acids (ssRNA, dsRNA, ssDNA and dsDNA) were bound by recombinant GST-PWWP (**Figure 3.8.2B**). Taken together, high GST-protein concentrations of PWWP2A's PWWP domain are required to shift different nucleic acids without any preference, pointing towards a low-affine, electrostatically-mediated nucleic acid-interaction.

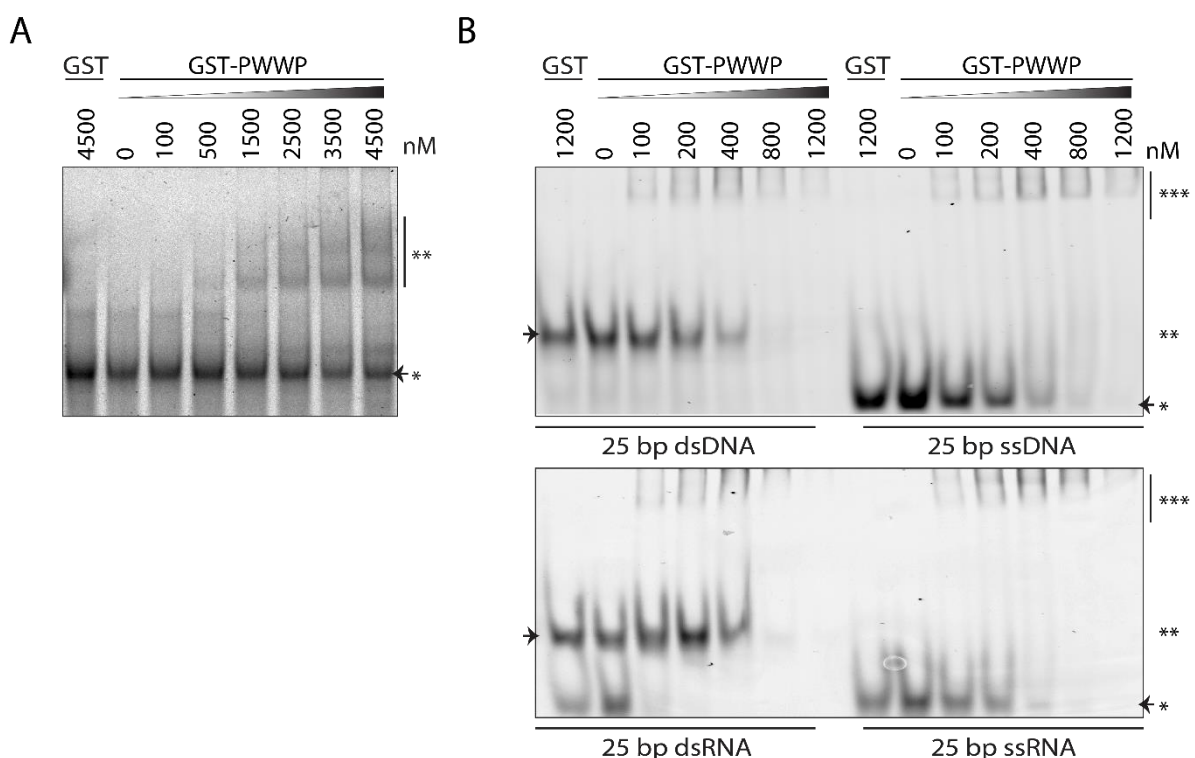


Figure 3.8.2 PWWP2A's PWWP domain interacts with nucleic acids. (A) Representative EMSA (n=2) using Cy5-tagged 75 bp dsDNA fragment and indicated concentrations of GST-PWWP. *free dsDNA, **DNA-GST-PWWP complex. (B) Representative EMSAs (n=2) using Cy5-tagged 25 bp dsDNA and ssDNA as well as Cy3-tagged 25 bp dsRNA and ssRNA fragments and indicated increasing concentrations of GST-PWWP. *free ssDNA/ssRNA, **free dsDNA/dsRNA ***nucleic acid-GST-PWWP complex.

The previous results indicate a general nucleic acid binding potential of the PWWP domain. Following, the ability of PWWP2A's PWWP domain to recognize wrapped nucleosomal DNA was investigated. Therefore, increasing concentrations of recombinant GST-PWWP (0 nM - 350 nM) were incubated (1) with recombinant H2A.Z-containing mononucleosomes with and without linker DNA (**Figure 3.8.3A**) as well as (2) recombinant H2A.Z- and H2A-containing

RESULTS

mononucleosomes lacking flanking DNA (**Figure 3.8.3B**) followed by subsequent cEMSA. GST-PWWP bound mononucleosomes containing 20 bp flanking DNA at a minimum GST-protein concentration of 50 nM, while mononucleosomes lacking linker DNA required at least 6x more GST-PWWP (300 nM) to get slightly shifted (**Figure 3.8.3A**). To evaluate, whether the absence of linker-containing mononucleosomes increases binding to mononucleosomes lacking linker DNA, GST-PWWP was incubated with both, linker-less H2A.Z- and H2A-nucleosomes. In line with the previous result, high GST-PWWP concentrations (250 nM - 300 nM) were required to shift linker-less mononucleosomes (**Figure 3.8.3B**), indicating a strong preference for linker DNA and just extremely low affine interaction with linker-less mononucleosomes.

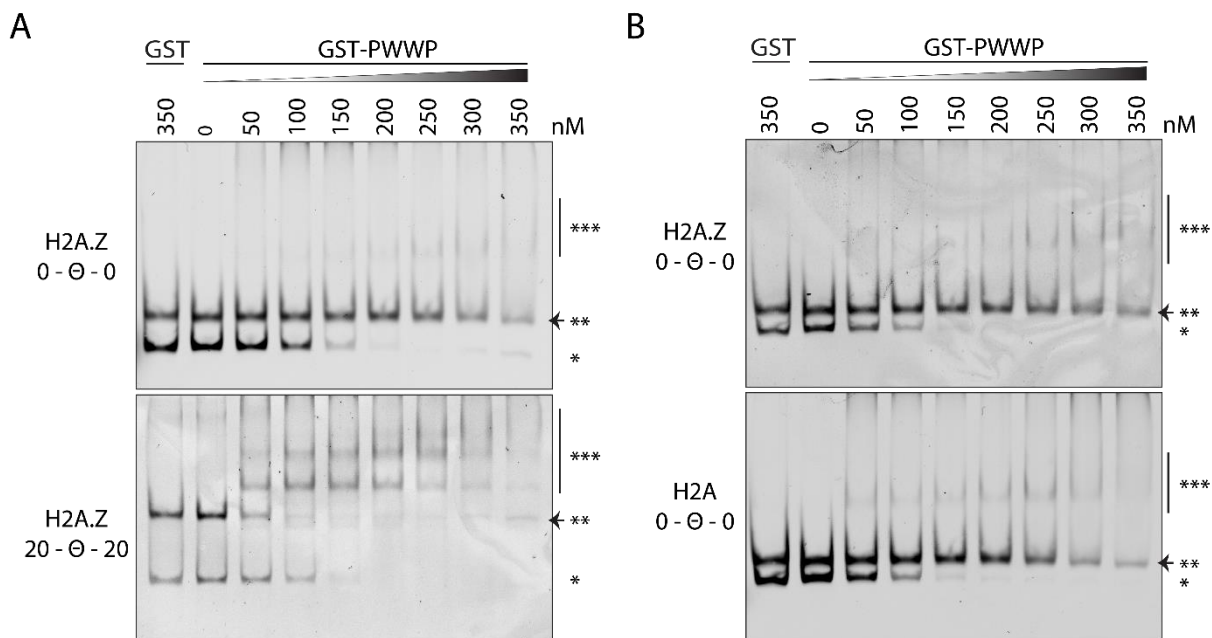
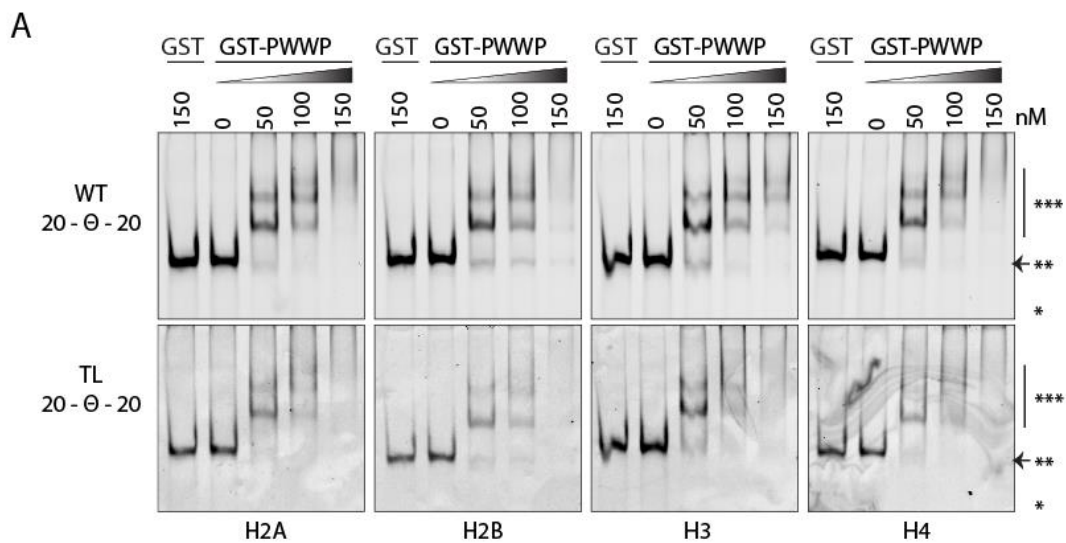


Figure 3.8.3 PWWP interacts with nucleosomal linker DNA. (A) Representative cEMSA (n=2) using 15 nM of recombinant human linker-less H2A.Z (0-Θ-0, top) or linker-containing H2A.Z (20-Θ-20, bottom) mononucleosomes incubated with indicated increasing concentrations of GST-PWWP. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-PWWP complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-PWWP complex is formed. (B) Representative cEMSA (n=2) similar to (A) using 15 nM of recombinant human H2A.Z- (top) or H2A- (bottom) containing mononucleosomes without linker DNA (0-Θ-0) incubated with indicated increasing concentrations of GST-PWWP.

Notably, PWWP2A's PWWP domain requires linker DNA to achieve a good nucleosome association. However, still some binding to mononucleosomes lacking flanking-DNA was observed and albeit very low, I wanted to determine whether this is a real effect and whether

RESULTS

further nucleosome components such as histone tails contribute to it. To test whether the flexible N-terminal histone tails might mediate this weak interaction, recombinant purified GST-PWWP (0 nM - 400 nM) was incubated with H2A-containing wildtype mononucleosomes and either H2A TL, H2B TL, H3 TL or H4 TL mononucleosomes lacking linker DNA or containing 20 bp on each side of the mononucleosome. GST-PWWP bound to linker DNA-containing WT and TL mononucleosomes in a similar manner, whereby a GST-protein concentration of 50 nM was sufficient to shift more than 50% of all nucleosomes (**Figure 3.8.4A**). In line with previous results, linker-less mononucleosomes were shifted only at high GST-PWWP concentrations (400 nM), whereby WT or TL linker-less nucleosomes were bound to a similar extent (**Figure 3.8.4B**). In conclusion, the small binding effect of GST-PWWP to linker-less mononucleosomes is probably real, but at least not mediated via single N-terminal histone tails.



RESULTS

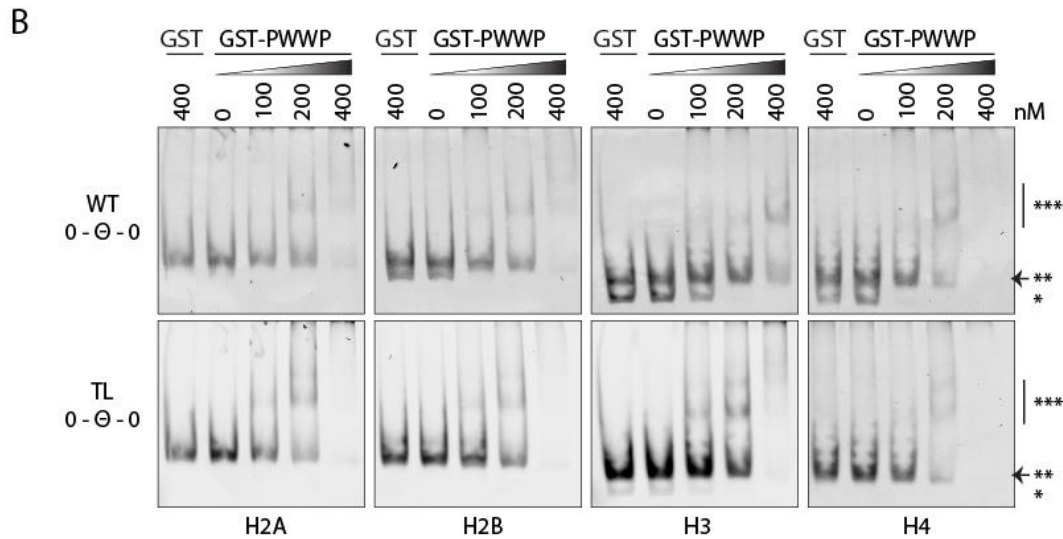


Figure 3.8.4 Single histone tails are not required for the binding of PWWP to mononucleosomes *in vitro*. (A) Representative cEMSA (n=2) using 15 nM of recombinant human wildtype mononucleosomes (WT, top) or mononucleosomes lacking single histone tails (TL, bottom) with linker DNA (20–0–20) incubated with indicated increasing concentrations of GST-PWWP. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-PWWP complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-PWWP complex is formed. (B) Representative cEMSA (n=2) similar to (A) using 15 nM of recombinant human wildtype mononucleosomes (WT, top) or mononucleosomes lacking single histone tails (TL, bottom) without linker DNA (0–0–0) incubated with indicated increasing concentrations of GST-PWWP.

The previous results clearly show that PWWP2A has two DNA binding domains, the N-terminal part of the internal region and its C-terminal PWWP domain. Next, I aimed to investigate whether the two domains have a combined effect in terms of their DNA binding properties. Therefore, EMSAs with 187 bp Cy5-labeled dsDNA and respective concentrations of recombinant purified GST-I_S_PWWP, GST-IN and GST-PWWP (0 nM, 25 nM, 50 nM, 100 nM and 200 nM) (**Figure 3.8.5A, B**) were performed. Remarkably, GST-I_S_PWWP bound the whole amount of 187 bp dsDNA with a very low GST-protein concentration of 25 nM (**Figure 3.8.5C**). In contrast, GST-IN concentrations of 100 nM and GST-PWWP concentrations of more than 200 nM were required to achieve a comparable binding effect as observed for GST-I_S_PWWP, indicating a combinatorial binding effect of both DNA-binding domains.

RESULTS

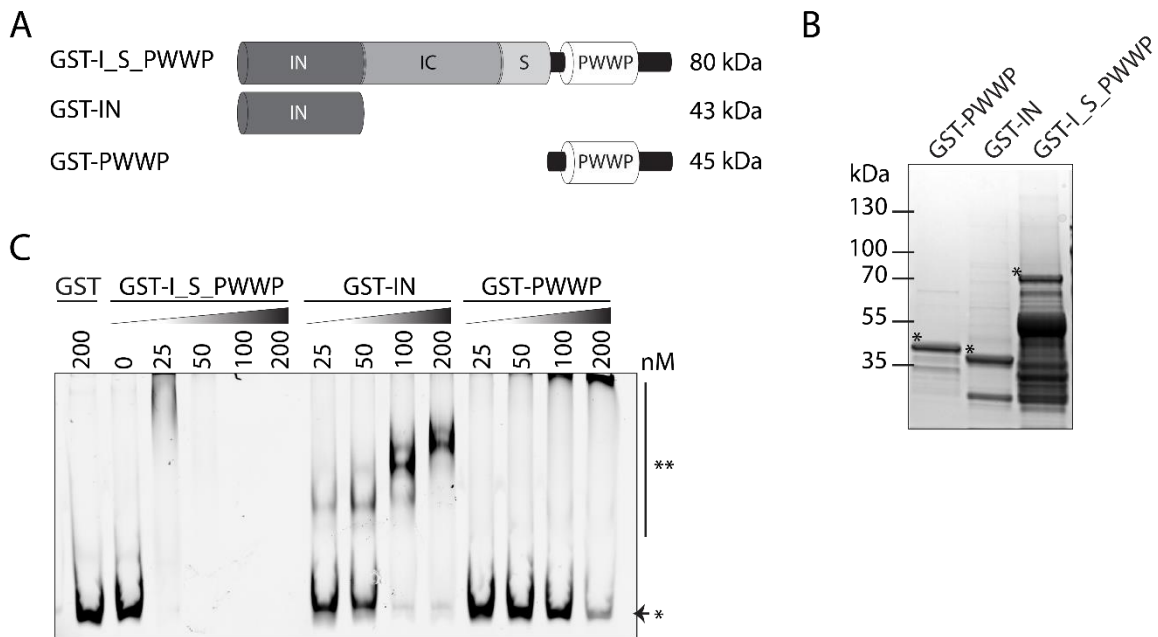


Figure 3.8.5 PWWP2A's IN and PWWP domains show a combinatorial DNA binding effect. (A) Schematic depiction of PWWP2A's N-terminal GST-tagged I_S_PWWP (80 kDa), IN (43 kDa) and PWWP domain (45 kDa). (B) Coomassie-stained SDS-PAGE gel of purified recombinant GST-I_S_PWWP, GST-IN and GST-PWWP used for cEMSA. *purified GST-protein. (C) Representative EMSA (n=3) using Cy5-tagged 185 bp dsDNA fragment and indicated concentrations of GST-I_S_PWWP, GST-IN and GST-PWWP. *free dsDNA, **DNA-GST-protein complex.

In summary, PWWP2A has at least two DNA interaction domains – PWWP and IN, whereby both domains in combination lead to an additive effect in nucleic acid binding strength. Moreover, PWWP and IN preferably interact with free nucleosomal flanking DNA without requiring free histone tails for their chromatin association.

3.9 Combination of S and PWWP domain interacts with *in vivo*-derived H3K36me3-monomucleosomes

In addition to general nucleic acid interaction, many PWWP domains are able to recognize specific posttranslational modifications (PTMs) of histones with three conserved amino acid residues that form an aromatic cage [294-298, 300, 328] with the recognition of H3K36me3 being most common [296]. Interestingly, PWWP2A is a member of the DNMT3A-related PWWP domain family, which consists besides DNMT3A also of DNMT3B and its close homolog PWWP2B [123]. Both *de novo* methyltransferases DNMT3A and DNMT3B catalyze DNA methylation on chromatin [325] and are known to recognize the posttranslational histone

RESULTS

modification H3K36me3 [329, 330]. Primary sequence alignments of PWWP2A's PWWP domain with those of PWWP2B, DNMT3A and DNMT3B revealed, that all of them harbor identically conserved amino acid residues (one phenylalanine and two tryptophanes) (**Figure 3.9.1A**). Here, the three conserved aromatic residues of PWWP2A's PWWP domain are at position 666 (F), 669 (W) and 695 (W). Additionally, *in silico* modeling of PWWP2A's PWWP domain (done by Sebastian Pünzeler) indicates that F666, W669 and W695 might form an aromatic cage thus possibly being able to recognize histone PTMs (**Figure 3.9.1B**).

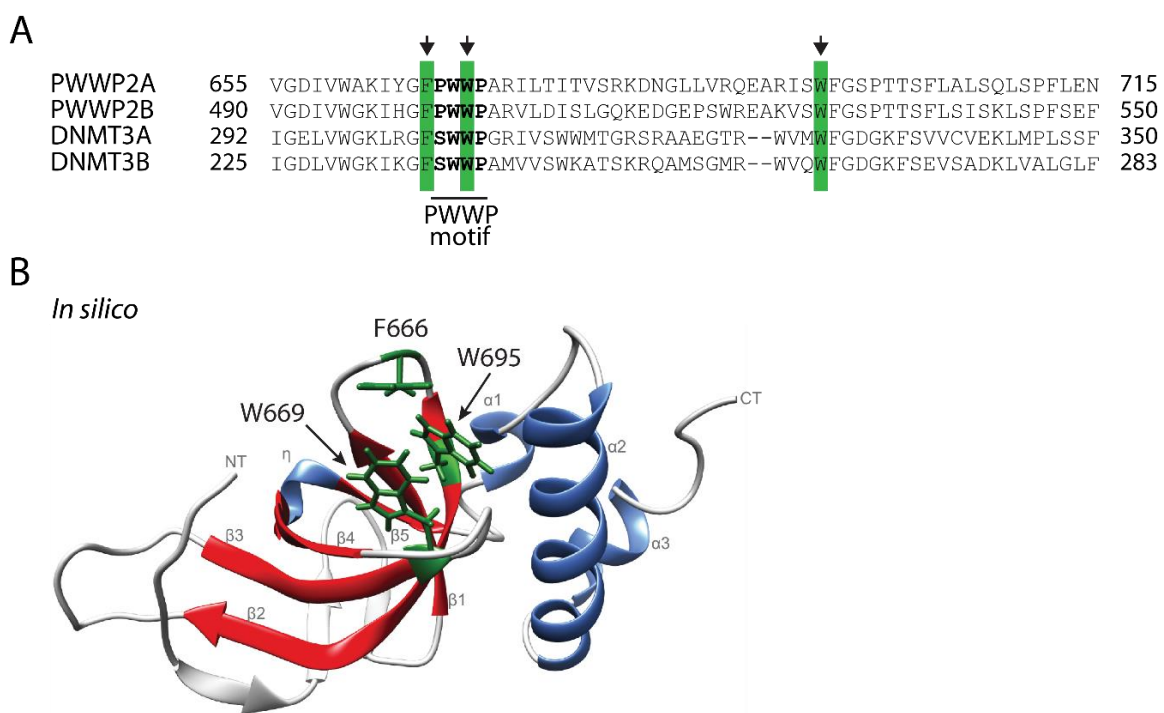
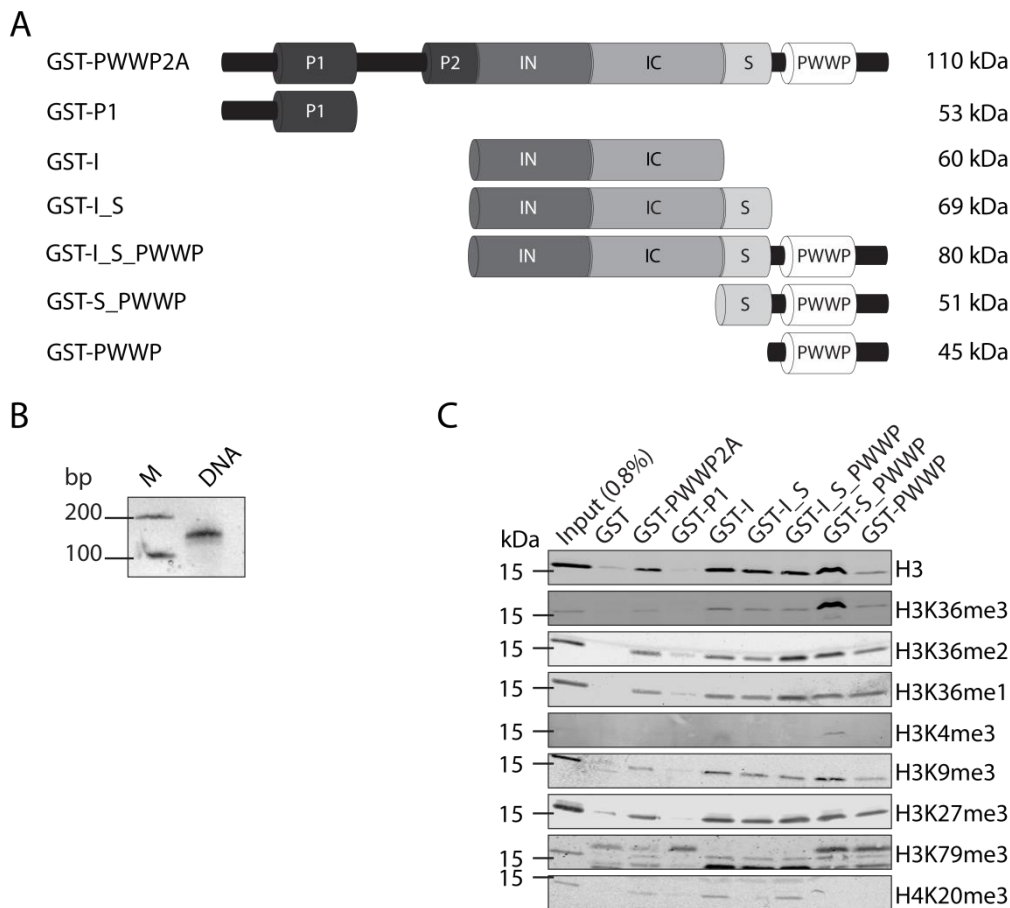


Figure 3.9.1 PWWP2A's PWWP domain possesses three conserved amino acid residues, which might form an aromatic cage. (A) Primary amino acid sequence alignment of PWWP domains of human PWWP2A [Q96N64], PWWP2B [Q6NUJ5], DNMT3A [Q9Y6K1] and DNMT3B [Q9UBC3]. Alignments were performed using the web browser-based tool MUSCLE. Amino acid residues shown to form an aromatic cage are highlighted in green and indicated with black arrowheads. Conserved PWWP motif is displayed in bold letters. (B) *In silico* secondary structure of PWWP domain modeled with the web browser-based tool iTASSER and visualized with Chimera (1.8.0.). α -helices (α 1 - α 3) and η -helix are depicted in blue, β -barrels (β 1 - β 5) in red and the three conserved amino acid residues forming the aromatic cage (F666, W669 and W695) are illustrated in green and stick mode. NT = N-terminus, CT = C-terminus

To investigate whether PWWP2A's PWWP domain can recognize certain PTMs on histone tails, recombinant GST-tagged PWWP2A full-length protein or PWWP2A truncations (P1, I, I_S,

RESULTS

I_S_PWWP, S_PWWP, PWWP) (**Figure 3.9.2A**) were incubated with MNase-digested mononucleosomes derived from HK cells (**Figure 3.9.2B**) and subsequently tested for specific PTM enrichment in immunoblots (**Figure 3.9.2C**). Notably, only a combination of the serine rich stretch S and the PWWP domain (S_PWWP) was able to efficiently enrich H3K36me3-mononucleosomes in comparison to other histone modifications (**Figure 3.9.2C, D**). Moreover, S_PWWP exclusively pulled-down mononucleosomes enriched in H3K36me3 but not mono- or dimethylated H3 lysine 36 (H3K36me1 and H3K36me2) (**Figure 3.9.2C, D**), proposing a site-specific and methylation state-dependent interaction.



RESULTS

D

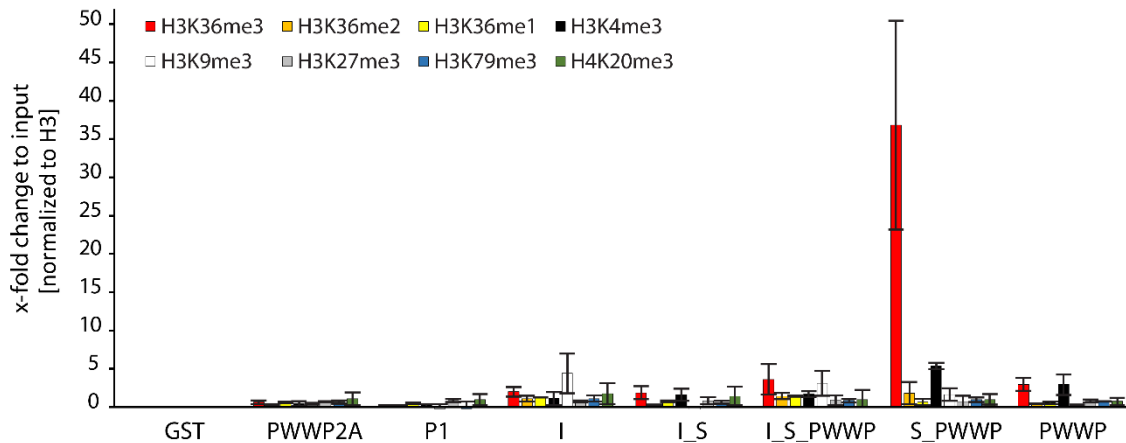


Figure 3.9.2 S_PWWP specifically interacts with H3K36me3-mononucleosomes. (A) Schematic depiction of N-terminal GST-tagged PWWP2A (110 kDa), P1 (53 kDa), I (60 kDa), I_S (69 kDa), I_S_PWWP (80 kDa), S_PWWP (51 kDa) and PWWP (45 kDa). (B) MNase-digested mononucleosomes were prepared from HK cell nuclei. DNA extraction followed by subsequent agarose gel electrophoresis and ethidium bromide staining revealed almost pure mononucleosomes (~150 bp). (C) Immunoblots (n=3) of pull-downs of recombinant GST-tagged PWWP2A full-length and truncations (P1, I, I_S, I_S_PWWP, S_PWWP, PWWP) incubated with HK-derived mononucleosomes. GST served as negative control. Histone modifications and histones (H3) are detected with specific commercially available antibodies. Notice increased signal intensity of H3K36me3 in S_PWWP pull-downs. (D) Data quantification done for three biological replicates for each PTM (n=3) by analyzing signal intensities using Image Studio Lite Ver 5.2 (LI-COR). Data are means of three biological replicates and are shown as x-fold change to input. Error bars depict the standard error of the mean (SEM).

In conclusion, I could show that a combination of PWWP2A's S and PWWP domain is able to specifically enrich mononucleosomes with the modification H3K36me3. Next, I wondered whether the three conserved aromatic cage residues are required for this interaction, hinting towards PWWP2A also forming an aromatic cage structure. Therefore, single point mutants having the respective amino acid residue changed to alanine (F666A, W669A and W695A) in GST-tagged S_PWWP were generated by site-directed mutagenesis (**Figure 3.9.3A**). Moreover, mutant proteins were tested for proper folding in nuclear magnetic resonance spectroscopy (NMR) by our collaboration partner Joel Mackay (School of Life and Environmental Sciences, University of Sydney, New South Wales 2006, Australia) (**Figure 3.9.3B**). Indeed, all point mutants displayed well-defined folds, since signals between 0 and 0.5 ppm could be detected.

RESULTS

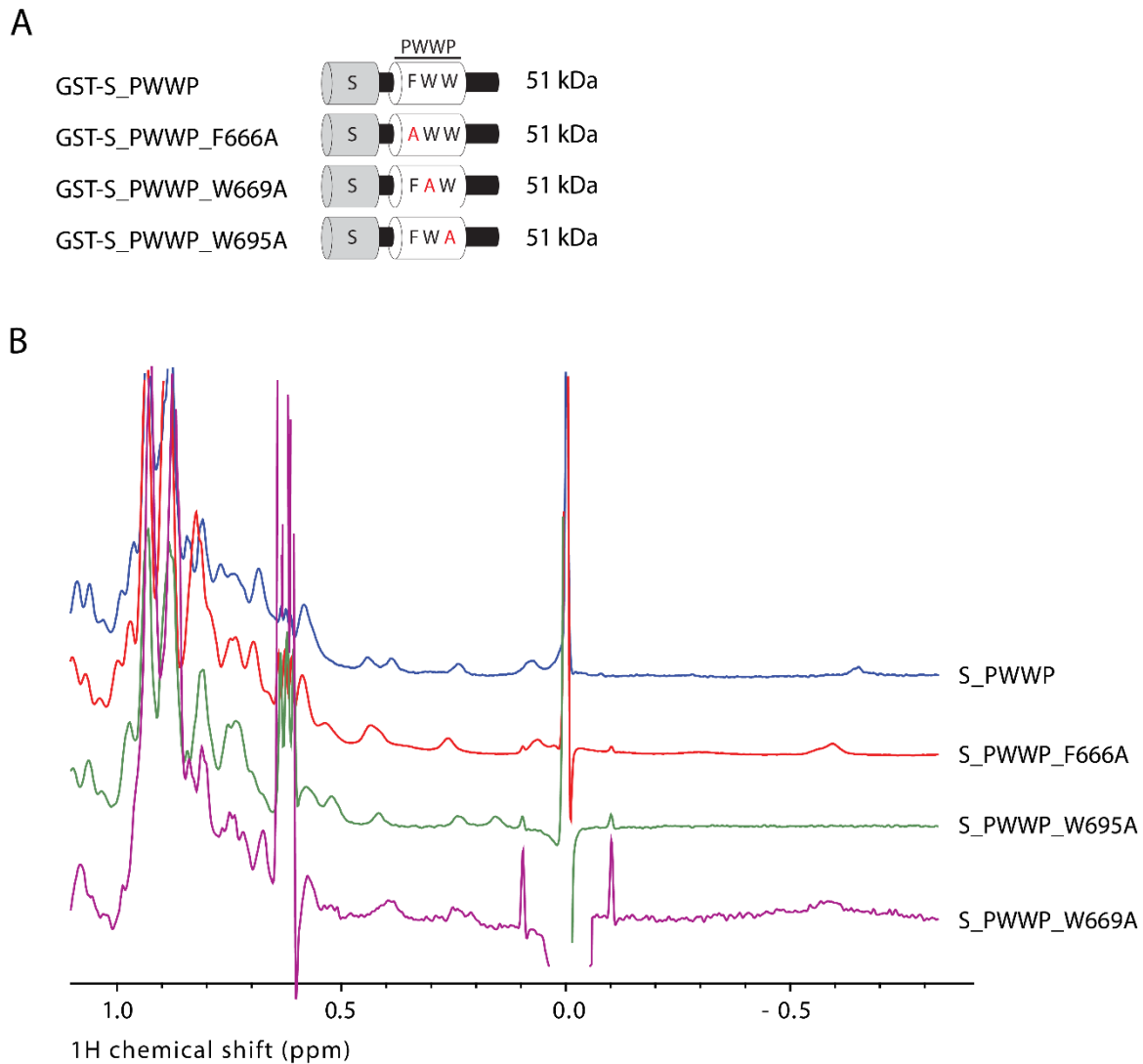


Figure 3.9.3 Aromatic cage mutants are properly folded. (A) Schematic depiction of N-terminal GST-tagged S_PWWP, S_PWWP_F666A, S_PWWP_W669A and S_PWWP_W695A with a molecular weight of 51 kDa. Single amino acid exchanges of conserved aromatic cage residues to alanine are highlighted in red. (B) Methyl region of one-dimensional (1D) ^1H NMR spectra of wildtype S_PWWP (blue) and aromatic cage mutants (S_PWWP_F666A (orange), S_PWWP_W669A (purple), and S_PWWP_W695A (green)). The presence of signal peaks between 0.0 and 0.5 ppm indicates well-defined folds and form stability of every mutant.

Next, to determine whether the form-stable and well folded point mutants lose their interaction with mononucleosomes containing H3K36me₃, recombinant GST-tagged S_PWWP, S_PWWP_F666A, S_PWWP_W669A and S_PWWP_W695A were incubated with MNase-digested, HK-derived mononucleosomes (Figure 3.9.4A) and pull-downs were analyzed by immunoblotting (Figure 3.9.4B). Interestingly, for two of the three single point mutants (GST-S_PWWP_W669A and GST-S_PWWP_W695A) the interaction with H3K36me₃ was almost completely lost upon amino acid exchange (Figure 3.9.4B, C). Moreover,

RESULTS

decreased immunoblot signal intensities were also observed for H3 after S_PWWP_W6669A and S_PWWP_W695A pull-downs, suggesting H3K36me3 interaction is a main mediator of S_PWWP mononucleosome binding. In contrast, exchange of phenylalanine to alanine (S_PWWP_F666A) did not dramatically affect the binding to this histone modification (**Figure 3.9.4B, C**), indicating that two of three aromatic-cage forming amino acid residues are required for H3K36me3-nucleosome interaction.

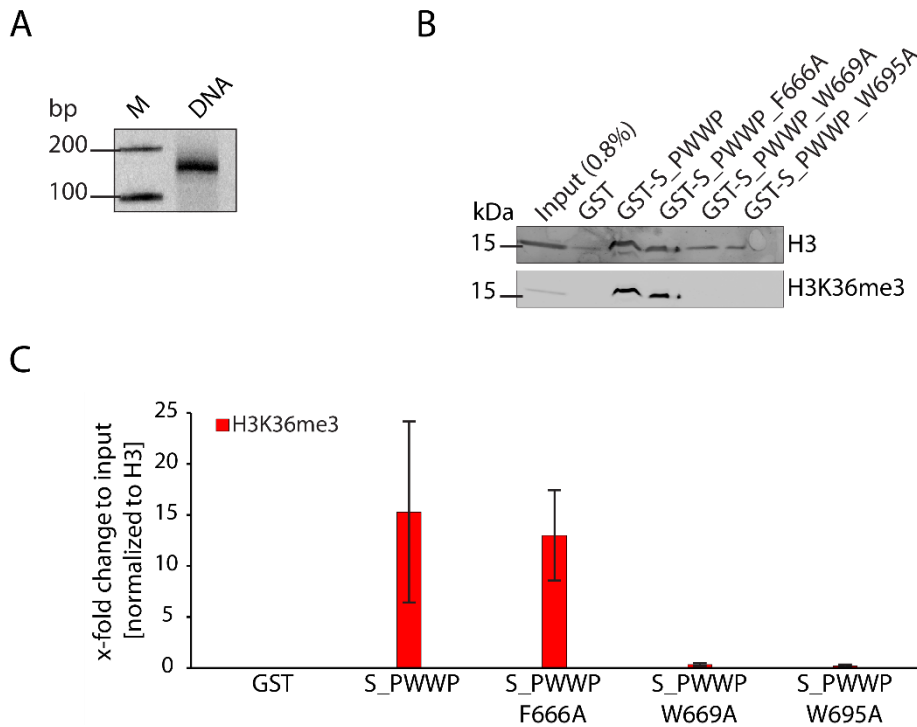


Figure 3.9.4 S_PWWP interaction with H3K36me3 mononucleosomes is dependent on aromatic cage residues W669 and W695. (A) MNase-digested mononucleosomes were prepared from HK cell nuclei. DNA extraction followed by subsequent agarose gel electrophoresis and ethidium bromide staining revealed almost pure mononucleosomes (~150 bp). (B) Immunoblots of pull-downs of recombinant GST-tagged S_PWWP and single aromatic cage point mutants (S_PWWP_F666A, S_PWWP_W669A and S_PWWP_W695A) incubated with HK-derived mononucleosomes. GST served as negative control. Histone modification H3K36me3 and histones (H3) are detected with specific commercially available antibodies. (C) Data quantification done for three biological replicates for each PTM (n=3) by analyzing signal intensities using Image Studio Lite Ver 5.2 (LI-COR). Data are means of three biological replicates and are shown as x-fold change to input. Error bars depict the standard error of the mean (SEM).

In addition to immunoblot analysis, label-free quantitative mass spectrometry (lf-qMS) was performed in collaboration with Moritz Völker-Albert and Axel Imhof (Molecular Biology, Zentrallabor für Proteinanalytik (ZfP), Biomedical Center, LMU, Munich). This independent and unbiased method was applied to additionally investigate the interaction with histone

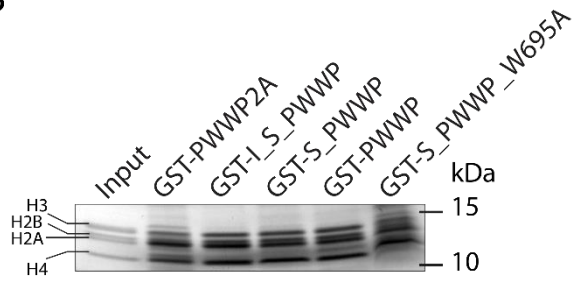
RESULTS

modifications and to test those, where no commercial antibodies are available (e.g. unmodified amino acid residues). Therefore, immunoprecipitations of recombinant, GST-tagged PWWP2A full-length and truncations (I_S_PWWP, S_PWWP, PWWP and S_PWWP_W695A) (**Figure 3.9.5A**) incubated with HK-derived mononucleosomes were performed. Input and pulled-down nucleosomes/histones were separated by SDS-PAGE, stained with Coomassie and all histone bands were excised together (**Figure 3.9.5B**). Following, excised mix of all histones was destained with acetonitrile (ACN), acetylated, digested with trypsin (Sigma) and loaded on C18/carbon stage tip to remove remaining contaminants. Next, samples were injected in mass spectrometer and obtained data were analysed with Skyline (version 3.6) [310]. Interestingly, the lf-qMS approach revealed a strong enrichment in H3K36me3-mononucleosomes after S_PWWP pull-downs and respective reduction upon S_PWWP point mutation (S_PWWP_W695A), what nicely verified previous findings (**Figure 3.9.5C**). In addition, enriched signal intensities were observed for H3K27me1H3K36me3 peptides, most likely due to co-immunoprecipitations of H3K36me3-nucleosomes with recombinant GST-S_PWWP. Unexpectedly, GST-tagged S_PWWP pull-downs followed by MS analysis (**Figure 3.9.5C**) or immunoblotting with specific antibodies (**Appendix Figure A1A, B**) revealed an enrichment in nucleosomes containing the posttranslational histone modification H3K79me2. This indicates either two independent binding mechanisms [298] or co-immunoprecipitations due to close proximity in distinct cellular situations [331-334].



RESULTS

B



C

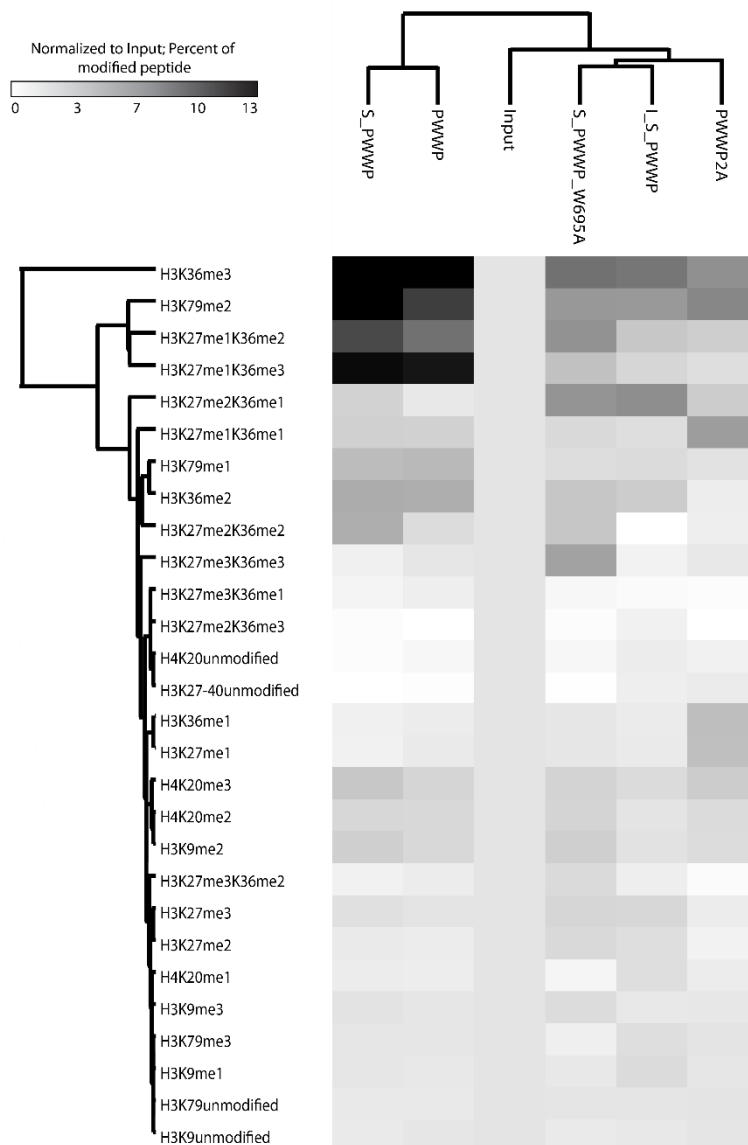


Figure 3.9.5 PWWP2A's S_PWWP interacts with H3K36me3- and H3K79me2-mononucleosomes. (A) Schematic depiction of N-terminal GST-tagged PWWP2A (110 kDa), I_S_PWWP (80 kDa), S_PWWP (51 kDa), PWWP (45 kDa) and S_PWWP_W695A (51 kDa). (B) GST-tagged, recombinant PWWP2A full-length, I_S_PWWP, S_PWWP, PWWP and S_PWWP_W695A were incubated with mononucleosomes derived from HK cells, separated by SDS-PAGE and subsequently stained with Coomassie Blue solution. Excised histone bands were cut out, destained, acetylated and trypsin-digested before they were loaded on C18/Carbon stage tips to remove final contaminants. Samples were separated (M/z) by MS and analyzed with Skyline. (C) If-qMS based heatmap of the abundance of multiple histone

RESULTS

modifications from pull-down assays described in (B), normalized to input (n=2). Scale-bar: percent of identified and modified peptide.

Pull-down experiments followed by MS or immunoblotting could identify that the combination of S and PWWP is needed to specifically pull-down H3K36me3-containing mononucleosomes in a HK cell-derived system. In order to gain insights whether this interaction is solely mediated by PWWP2A's S_PWWP domain or whether additional factors like histone variants as e.g. observed for ZMYND11, which needs histone H3 variant H3.3 being present to efficiently bind to H3K36me3 [297] might play a role, cEMSAs were carried out. Therefore, recombinant and chemically modified histone octamers (*X. laevis*) containing either H3Kc36me0 or H3Kc36me3 (H3Kc36: (methyl-) lysine analog at position 36 of H3, whereby H3Kc36me3 mimics trimethylation) (kind gift of Michaela Smolle, **Figure 3.9.6A**) were reconstituted with either 147 bp or 187 bp Widom-based DNA fragments [314] and incubated with recombinant GST-tagged PWWP2A or deletion proteins. First, low salt conditions were tested using recombinant GST-PWWP (0 nM - 400 nM) (**Appendix Figure A2A**), GST-S_PWWP (0 nM - 200 nM) (**Appendix Figure A2B**) or GST-PWWP2A (0 nM - 240 nM) (**Appendix Figure A2C**) incubated with linker-containing H3Kc36me0 and H3Kc36me3-nucleosomes in cEMSAs. GST-PWWP was able to shift both, nucleosomes with or without trimethylation-mimic mark at minimum GST-protein concentrations of 50 nM, whereby almost half of them were shifted at 200 nM and all nucleosomes at a final GST-PWWP concentration of 400 nM (**Appendix Figure A2A**). For GST-tagged S_PWWP only a GST-protein concentration of 50 nM was required to shift approximately half of all mononucleosomes (H3Kc36me0- and H3Kc36me3-containing), whereby 200 nM instead of 400 nM, as observed for GST-PWWP, were sufficient to bind all of them (**Appendix Figure A2B**). In contrast, to bind H3Kc36me0- and H3Kc36me3-containing nucleosomes to GST-PWWP2A, a minimum GST-protein concentration of 80 nM was needed, whereby at 240 nM all mononucleosomes were shifted, regardless if the trimethylation-mimic mark was present or not (**Appendix Figure A2C**). Sometimes unspecific binding can lead to misinterpretations. Hence, I tested whether higher stringencies (300 mM NaCl instead of 150 mM NaCl) or different linker DNA lengths might lead to any *in vitro* binding preference. Therefore, recombinant H3Kc36me3- and H3Kc36me0-containing nucleosomes, wrapped with either 147 bp or 187 bp DNA fragments were incubated with indicated increasing concentrations of recombinant GST-tagged

RESULTS

S_PWWP (**Figure 3.9.6B, C**) (0 nM - 250 nM) and a final NaCl concentration of 300 mM. In line with previous findings, binding of S_PWWP to linker-less H3Kc36me0- and H3Kc36me0-nucleosomes required higher protein concentrations (250 nM) than binding to linker-containing H3Kc36me0- and H3Kc36me0-nucleosomes (100 nM) (**Figure 3.9.6D, E**). In addition, H3Kc36me0- and H3Kc36me0-nucleosomes were bound in a similar manner, also demonstrated in cEMSA band signal quantifications (bound vs unbound) (**Figure 3.9.6D, E**). In conclusion, neither the presence or absence of linker DNA nor higher stringencies (300 mM) revealed a direct preference for H3Kc36me3-nucleosomes *in vitro*, hinting towards additional factors that might be required to achieve a strong association *in vivo*.

RESULTS

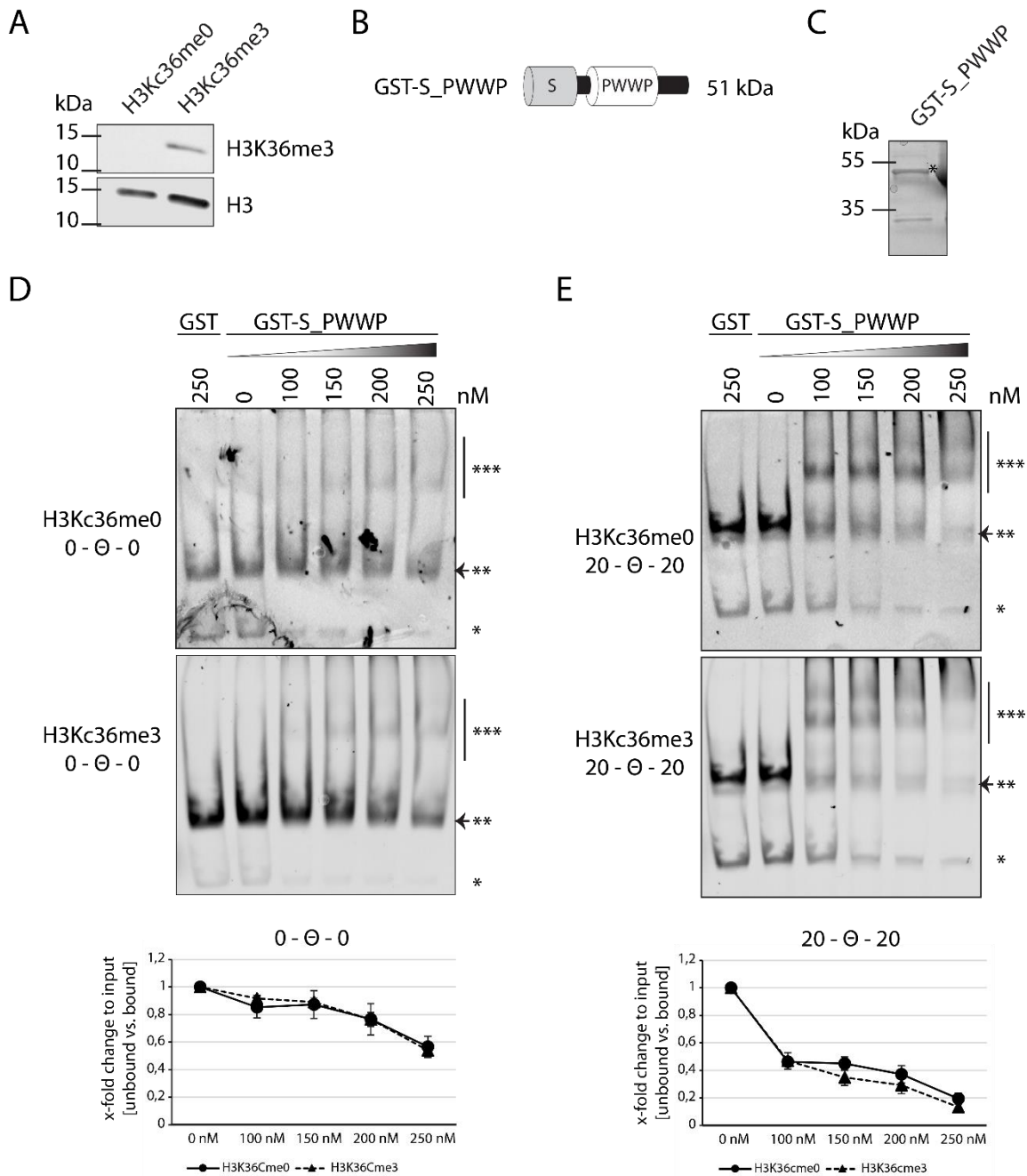


Figure 3.9.6 S_PWWP does not distinguish between H3K36me0 and H3K36me3 *in vitro*. (A) Immunoblots with commercially available antibodies for H3K36me3 and H3 (loading control) confirmed the absence or presence of trimethyl modification of H3K36, respectively. (B) Schematic depiction of N-terminal GST-tagged S_PWWP (51 kDa). (C) Coomassie-stained SDS-PAGE gel of purified and eluted GST-S_PWWP used for cEMSA. *purified GST-protein. Additional bands are GST-protein degradation products. (D) Top: Representative cEMSA (n=3) using 15 nM of recombinant frog linker-less (0-θ-0) H3K36me0 or H3K36me3-containing mononucleosomes incubated with indicated increasing concentrations of GST-S_PWWP. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-protein complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-S_PWWP complex is formed. Bottom: Quantification of signal intensities by analyzing bound versus unbound mononucleosomes (**), using Image Studio Lite Ver 5.2 (LI-COR). Error bars indicate standard error of the mean (SEM) of three independent replicates. (E) Representative cEMSA (n=3) similar to (D) using 15 nM of recombinant frog 20 bp linker-containing (20-θ-20) H3K36me0 or H3K36me3 mononucleosomes incubated with indicated increasing concentrations of GST-S_PWWP. GST served as negative control.

RESULTS

In summary, S_PWWP specifically enriches H3K36me3-containing mononucleosomes that were derived from human cells, whereby a combination of the serine rich stretch S and the PWWP domain is required. Moreover, S_PWWP is not able to distinguish *in vitro* between H3K36me0- and H3K36me3 chemically modified recombinant mononucleosomes.

3.10 PWWP2A interacts with γ -H2A.X mononucleosomes

In the previous sections, I demonstrated that PWWP2A possesses multiple binding mechanisms to different chromatin moieties with distinct domains. Surprisingly, I realized that for the GST-tagged PWWP2A truncation proteins S_PWWP and PWWP neither histone variant H2A.Z (due to lack of H2A.Z-specific recognition domain IC) nor canonical histone H2A was enriched in pull-down experiments with HK-derived mononucleosomes (**Figure 3.10.1A**). Hence, I wondered which member of the histone H2A family might be incorporated instead of H2A.Z or H2A in pull-down experiments with GST-S_PWWP and GST-PWWP. Besides canonical histone H2A and histone variant H2A.Z, the H2A family consists in addition of H2A.Bbd, macroH2A and H2A.X [12]. H2A.Bbd is mostly a testis-specific variant [186, 187, 189] and heterochromatic macroH2A has main functions in female X-inactivation [335, 336] or embryonic stem cell differentiation [337]. In contrast, H2A.X is found at DNA double strand breaks and gets specifically modified upon DNA damage [183]. Since we are working with HK cells (and not in testis (H2A.Bbd) or embryonic stem cells (macroH2A)), I decided to initially test whether the DNA damage response-related variant H2A.X [183] is enriched in GST-tagged S_PWWP and PWWP pull-down experiments. To do so, different GST-tagged PWWP2A full-length and deletion proteins were incubated with MNase-digested, HK-derived mononucleosomes and subsequently visualized by immunoblotting. Anti-H2A.X immunoblots of GST-S_PWWP and GST-PWWP HK-derived mononucleosome pull-downs did not show an enrichment in H2A.X signal intensity in comparison to other PWWP2A truncation proteins (GST-tagged PWWP2A, P1, I, I_S and I_S_PWWP) (**Figure 3.10.1B**). Nevertheless, only pull-downs followed by anti-H2A.X immunoblots of GST-tagged S_PWWP and PWWP showed a clear double band signal running slightly higher than the H2A.X signal. Interestingly, H2A.X can be modified e.g. by phosphorylation, whereby the most popular is the modification at position 139 [183]. Hence I wondered, whether GST-tagged S_PWWP and PWWP proteins enrich the on S139 phosphorylated version of H2A.X (γ -H2A.X) in HK cell-derived pull-down approaches.

RESULTS

Notably, immunoblots with commercially available antibodies against γ -H2A.X revealed that both proteins, GST-S_PWWP and GST-PWWP, accumulated mononucleosomes containing γ -H2A.X, whereby other PWWP2A proteins (GST-tagged PWWP2A, P1, I, I_S and I_S_PWWP) did not (**Figure 3.10.1B**). In conclusion, GST-tagged S_PWWP and PWWP proteins are able to specifically enrich γ -H2A.X-nucleosomes, but not unmodified H2A.X-nucleosomes, and might link PWWP2A to DNA damage events.

Next, I wanted to investigate how this interaction is established in more detail. In previous experiments, I determined S_PWWP's ability to enrich H3K36me3-containing mononucleosomes. Interestingly, H3K36me3 can also occur at DNA damage sites and is involved in double-strand break repair [338] as shown for γ -H2A.X [339]. Hence I wondered, whether γ -H2A.X is pulled-down in combination with the aromatic cage-dependent enrichment of H3K36me3-nucleosomes. To test this hypothesis, GST-tagged single aromatic cage point mutant and wildtype proteins (S_PWWP, S_PWWP_F666A, S_PWWP_W669A and S_PWWP_W695A) were incubated with HK-derived mononucleosomes. Immunoblot analysis with specific antibodies revealed, that in line with previous experiments two of three point mutant proteins (S_PWWP_W669A and S_PWWP_W695A) showed a loss of H3K36me3-containing mononucleosome interaction (**Figure 3.10.1C**). Intriguingly, γ -H2A.X band signal intensities displayed constant levels for all proteins (S_PWWP, S_PWWP_F666A, S_PWWP_W669A and S_PWWP_W695A), indicating no co-immunoprecipitation with H3K36me3-containing nucleosomes but rather an independent binding mechanism (**Figure 3.10.1C**).

Following, I wanted to determine the region of S_PWWP being sufficient to mediate the interaction with γ -H2A.X-containing nucleosomes. To do so, different GST-tagged PWWP2A truncation proteins (S_short, S_long, S_PWWP, PWWP and PWWP_short (**Figure 3.10.1D**)) were incubated with HK-derived mononucleosomes and subsequently analyzed in immunoblots by using commercially available antibodies. Notably, the serine rich stretch on its own (S_short and S_long) was not able to enrich γ -H2A.X-containing mononucleosomes in pull-down approaches (**Figure 3.10.1E**). Remarkably, GST-tagged PWWP_short was sufficient to specifically pull-down mononucleosomes containing γ -H2A.X (**Figure 3.10.1E**), suggesting a serine-rich stretch independent recognition of DNA damage-related γ -H2A.X by PWWP2A's very C-terminus by a yet unknown binding mechanism.

RESULTS

In summary, γ -H2A.X-containing mononucleosomes are specifically enriched in pull-downs with PWWP2A's C-terminal half, whereby the presence of the serine-rich stretch is not required and PWWP_short is sufficient. Moreover, the interaction is not mediated by co-immunoprecipitations with H3K36me3-containing mononucleosomes, suggesting an independent binding mechanism to γ -H2A.X-nucleosomes in a direct or indirect manner.

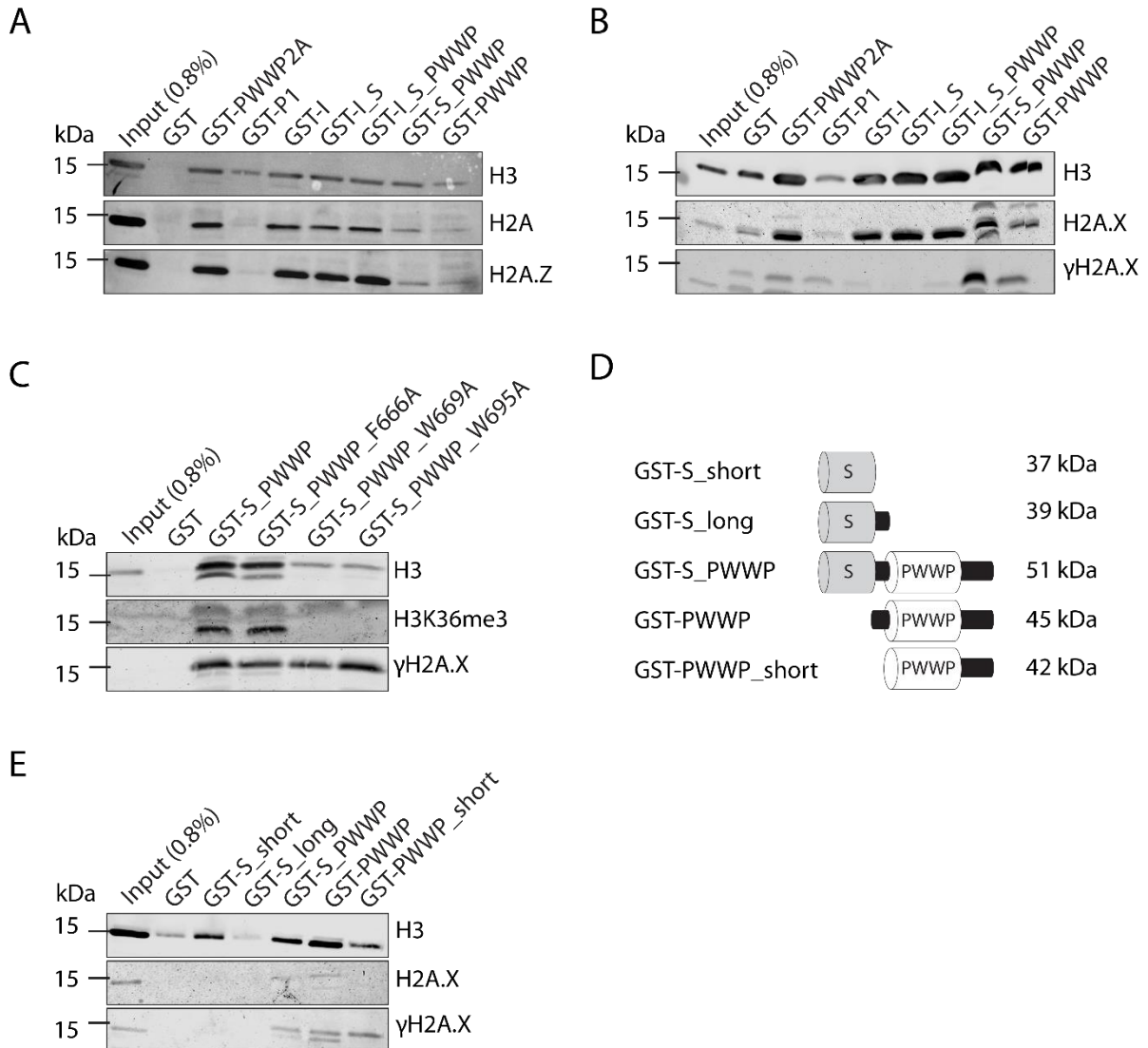


Figure 3.10.1 PWWP2A's C-terminus interacts with nucleosomes containing the modified histone variant γ -H2A.X independent of the aromatic cage. (A) Pull-downs of GST-tagged recombinant PWWP2A and indicated truncation proteins (P1, I, I_S, I_S_PWWP, S_PWWP and PWWP) with HK cell-derived mononucleosomes. GST served as negative control. Precipitated histone H2A family members (H2A and H2A.Z) and H3 are detected in immunoblots using commercially available antibodies. (B) Similar to (A) using commercially available antibodies against H2A.X, phosphorylated γ -H2A.X (S139ph) and H3. Notice double band for S_PWWP and PWWP in H2A.X immunoblot. (C) Similar to (A) using GST-tagged recombinant wildtype S_PWWP and single aromatic cage point mutant proteins (S_PWWP_F666A, S_PWWP_W669A, and S_PWWP_W695A). H3K36me3, γ -H2A.X and H3 are detected in immunoblots by using commercially available antibodies. Notice decreased signal intensity in H3K36me3 binding for S_PWWP_W669A and S_PWWP_W695A in comparison to wildtype S_PWWP

RESULTS

but constant levels of γ -H2A.X. (D) Schematic depiction of N-terminal GST-tagged S_short (37 kDa), S_long (39 kDa), S_PWWP (51 kDa), PWWP (45 kDa) and PWWP_short (42 kDa). (E) Similar to (B) using GST-tagged recombinant proteins (S_short, S_long, S_PWWP, PWWP and PWWP_short).

Summary

All in all, I demonstrated that PWWP2A directly binds to chromatin with at least four distinct domains – IN, IC, S_PWWP and PWWP. PWWP2A's internal region divides its labor in general nucleosome interaction and specific H2A.Z-recognition, whereby IC strongly binds to recombinant mononucleosomes and discriminates *in vitro* between H2A and H2A.Z. Moreover, two amino acid residues (aa 461 and aa 475) are crucial for the interaction with H2A- and H2A.Z-nucleosomes *in vitro*. However, I additionally showed nine unique amino acids at the C-terminus of the histone variant H2A.Z playing an important role for the interaction with PWWP2A in a cell-derived system. The N-terminal part of the internal region contributes to general recognition of nucleosomes by strongly binding free DNA and preferring mononucleosomes with flanking DNA. Furthermore, I determined that the C-terminally located PWWP domain directly binds to chromatin via interaction with free nucleic acids like DNA or RNA and prefers mononucleosomes containing flanking DNA. In addition, for both DNA binding domains IN and PWWP, single histone tail deletions did not negatively affect the binding strength to recombinant mononucleosomes. Further, I unraveled that a combination of S and PWWP is required to recognize cell-derived H3K36me3-containing mononucleosomes in a methylation state- and site-specific manner. Nevertheless, S_PWWP was not able to discriminate *in vitro* between H3Kc36me0 and H3Kc36me3-containing mononucleosomes. Last but not least I demonstrated that PWWP2A interacts with yet another member of the histone H2A family – γ -H2A.X. There, already a shortened version of the recombinant GST-PWWP construct (GST-PWWP_short) was sufficient to interact with γ -H2A.X without requiring the presence of the serine-rich stretch like detected for H3K36me3-nucleosome interaction. To date, the detailed underlying binding mechanism of PWWP2A to its newly identified interaction partner γ -H2A.X and any functional consequence of this interaction is still elusive, and will be further analyzed in future experiments.

4 Discussion

PWWP2A is a novel vertebrate-specific H2A.Z interactor we could recently identify in pull-down experiments via If-qMS [276]. Genome-wide mapping further revealed that PWWP2A co-localizes with H2A.Z-containing nucleosomes incorporated at TSS of highly transcribed genes. In HeLa Kyoto cells, the knockdown of PWWP2A deregulates several hundred genes and affects mitosis. However, PWWP2A seems not to be an H2A.Z-specific chaperone or chaperone-complex member since it does not influence H2A.Z occupancy upon depletion. Aiming to unravel the detailed underlying binding mechanisms of PWWP2A with chromatin, it was first investigated whether the interaction is achieved in a direct or indirect manner (**Figure 3.1.3**). Notably, PWWP2A is a direct interaction protein of histone variant H2A.Z and possesses due to its distinct domains multiple and independent binding modes to enable a stable and specific association with chromatin.

4.1 PWWP2A is a multivalent chromatin binder and specifically recognizes H2A.Z

PWWP2A consists of several domains that independently contribute to a directed and specific interaction with chromatin. Since large proteins are difficult to study and protein domains are assumed to be autonomous units, single domain analyses are usually applied [340]. Nina Kronbeck already demonstrated that PWWP2A harbors at least two independent chromatin interaction domains – the predicted to be unstructured internal region I and the C-terminally located PWWP domain [276], whereby the internal region can be further subdivided into an N- (IN) and C- (IC) terminal part. cEMSAs revealed a direct chromatin association for both parts of the internal region, where IN generally interacts with mononucleosomes regardless if H2A or H2A.Z is incorporated and IC specifically binds to H2A.Z *in vitro* (**Figure 3.2.2**). It has been noted for several other H2A.Z-recognition proteins that they interact with certain amino acid residues of histone variant H2A.Z [242, 341]. YL1, a highly conserved Swc2 homolog expressed in higher eukaryotes [342, 343] as well as Anp32e, a chaperone for H2A.Z [344], are reported to specifically recognize the α C cleft of H2A.Z's C-terminal half [242, 341]. Interestingly, I could demonstrate that also PWWP2A's IC domain interacts via the C-terminus of H2A.Z (**Figure 3.3.1**). Thereby, only a short, nine amino acids comprising C-terminal flexible region, that

DISCUSSION

extends out of the H2A.Z nucleosomal core and is conserved in vertebrates, is sufficient for a specific binding of IC to H2A.Z. As PWWP2A interacts equally well with nucleosomes containing either H2A.Z.1 or H2A.Z.2 [276], that differ within their nine C-terminal amino acids only in the last residue, PWWP2A's interaction motif seems to be the C-terminal H2A.Z histone tail amino acid sequence "GKKGQQKT". Nonetheless, the extended acidic patch, unique for H2A.Z in comparison to H2A [221], and crucial for proper targeting, deposition and recruitment of interaction proteins [212] is not the sole mediator for PWWP2A nucleosome binding supporting the independent binding modes of its multiple domains to chromatin. Surprisingly, cEMSA using N- (aa 1 - 12) and C- (aa 122 - 128) terminal tail depleted H2A.Z-containing mononucleosomes incubated with recombinant PWWP2A or its IC domain only, couldn't confirm the previous results *in vitro* (**Figure 3.3.2**). Interestingly, PWWP2A's possible interaction motif harbors lysine residues that can be ubiquitylated [230] thereby regulating the localization and activity of diverse cellular proteins [345]. Specific H2A.Z ubiquitination levels are for example required for the association of H2A.Z with transcriptionally silenced facultative heterochromatin [230] or the regulation of androgen receptor-linked genes [346]. Moreover, it is known that histone PTMs like ubiquitinations are capable of creating new protein interaction sites or abrogating protein-protein contacts [347]. Thus, depleting the nine unique C-terminal amino acids of H2A.Z's histone tail might result in the loss of specific ubiquitinations, which are usually crucial to prevent binding of distinct proteins [230, 347]. Now, these interaction proteins might be able to bind H2A.Z upon missing ubiquitination marks and decrease the accessibility of IC's target sequence or specific H2A.Z interaction motif. This is also in line with our *in vitro* data, but suggests rather an indirect loss of IC – H2A.Z interaction in HeLa Kyoto cells due to occupied target sites. Moreover, another possibility to explain these contradictory findings is that *in vitro* reconstituted mononucleosomes lack only seven instead of nine unique C-terminal amino acids. Thus, the rather N-terminally located residues "G120" and "K121" of the "GKKGQQKT" motif would be crucial to establish the specific interaction between H2A.Z and IC, regardless if H2A.Z's N-terminal tail is present or not.

Next, I was interested to determine which amino acid residues within IC are required to specifically attract H2A.Z-containing mononucleosomes. Notably, the IC domain harbors two amino acid residues (R461 and R475) frequently found to be mutated in different cancer types as identified by cBioPortal [318, 319]. Interestingly, the internal domain responded extremely

DISCUSSION

sensitive to these point mutations by displaying reduced binding affinity for H2A.Z- and surprisingly also for H2A-monomucleosomes *in vitro* (**Figure 3.4.3**). Non-synonymous amino acid exchanges are responsible for many gene lesions that result in human diseases [348]. Moreover, several amino acid substitutions change the secondary structure of proteins thereby affecting the attraction of distinct interaction partners [349] or destroy general protein function [350]. Hence, impaired binding of the point mutated internal region to mononucleosomes might be a result of varied secondary protein structures or due to loss of protein function upon amino acid exchange. Especially the second internal region point mutant “I_R475C” showed decreased binding to H2A.Z-containing mononucleosomes. Deregulation of H2A.Z is known to be involved in many different cancer types like malignant melanomas [265, 320], whereby the correlation between distinct H2A.Z levels and poor prognosis for patient survival quickly became obvious [351]. Notably, PWWP2A has no predicted enzymatic activity therefore rather acting as possible binding hub for distinct H2A.Z- interactors, whereby our group already identified PWWP2A-specific interaction proteins by If-qMS (Link & Spitzer et al., manuscript under review). Possibly, upon amino acid exchange within I, PWWP2A is not able to establish a strong and specific association with H2A.Z-containing nucleosomes anymore. This might lead to an impaired protein (complex) recruitment by PWWP2A to H2A.Z sites, possibly affecting the surrounding chromatin structure or/and gene regulation. Interestingly, this theory is also in line with previous findings, where we demonstrated several hundred genes being deregulated upon PWWP2A knock-down [276]. In conclusion, point mutations within PWWP2A’s internal region possibly impair PWWP2A’s function as recruitment platform for specific H2A.Z interaction proteins, which in turn leads to the deregulation of several hundred genes, most likely also oncogenes. Furthermore, a possible deregulation of oncogenes could be the reason, why mutations within the internal region were found to be correlated with distinct cancer types [318, 319].

As mentioned before, PWWP2A consists of distinct domains whereby not all of them have predictable structures, e.g. the internal region. Usually, multidomain proteins consist of distinct protein domains connected by linker peptides thereby providing interdomain interactions, flexibility or biological function [352]. In the past, the unstructured linker regions had been vastly underestimated, because the widespread opinion was that only a fix 3D structure is leading to proper protein function [352]. However, recent studies show that disordered protein regions have functionally significant tasks [353]. Also PWWP2A’s IN and IC

DISCUSSION

region are predicted to be disordered and connect the proline-rich region P2 with the serine-rich stretch S. Interestingly, I showed that also IN and IC fulfill important functions, whereby IC plays a role in H2A.Z-recognition and IN in general nucleosome binding. Experiments using recombinant IN and TL mononucleosomes (**Figure 3.5.2**) demonstrated that IN does not require the presence of single histone tails to achieve general mononucleosome interaction. However, the absence of histone H3 tail facilitated binding to nucleosomes most likely via feasible access to flanking DNA [321, 354], suggesting a putative role of IN as DNA binder. Notably, also other disordered proteins regions are known to interact with nucleic acids such as DNA [355] mostly mediated by nonspecific, electrostatic interactions [352]. Those regions often possess many positively charged residues like arginines that insert into the minor groove and mediate DNA binding [352, 356]. Notably, also IN is enriched in positively charged lysines and arginines and was found to efficiently bind to free and nucleosomal linker DNA (**Figure 3.6.2** and **Figure 3.6.3**), most likely via electrostatic surface potential or certain DNA sequence motifs. In addition, the specific binding of IC to H2A.Z-nucleosomes possibly results in well-defined secondary protein structures thereby mediating the contact of IN to nucleosomal linker DNA. Furthermore, crystallization of the internal region without or with H2A.Z-containing nucleosomes could clarify whether these domains are folded upon nucleosome binding or not and how they are exactly oriented at the nucleosome.

However, maybe there is a more complex mechanism behind IN and specific DNA recognition since PWWP2A harbors another, rather general DNA interaction motif (discussed later). Notably, nucleosomes are condensed and folded into higher order chromatin structures [357]. DNA interaction proteins and especially transcriptions factors must bind to their target sequences and enable cellular processes including transcription or repair [324]. Thereby, so-called pioneer transcription factors form a specific class of proteins being able to access DNA target sites, when others cannot [324]. Interestingly, single “common” transcription factors mostly cannot properly bind to nucleosomal target DNA *in vitro* [358, 359], whereby IN showed very strong binding to recombinant nucleic acids. In concordance, *in vitro* DNA binding, strong DNA association and the presence of an additional but rather general DNA interaction domain, indicates a possible role for IN as pioneer factor. Nevertheless, further experiments are required to unravel whether IN is a possible pioneer transcription factor interacting with specific DNA motifs or rather a general DNA binder via electrostatic surface interactions.

DISCUSSION

In addition to the H2A.Z-interaction domain IC and DNA-recognition part IN, PWWP2A possesses another important chromatin interaction motif, the C-terminally located and eponymous PWWP domain. Interestingly, PWWP2A's PWWP domain displays a positively charged surface (**Figure 3.8.1**) pointing towards a possible function in DNA-binding. This is in agreement with previous studies showing that positively charged surfaces of PWWP domains are required for DNA interaction [reviewed in [293]]. Numerous PWWP domains have been demonstrated to exhibit the ability of DNA-binding including those of PSIP1 [294, 295, 360], DNMT3A [361] and the bromo zincfinger-PWWP domain of ZMYND11 [328]. cEMSA with PWWP2A's PWWP domain and different types of nucleic acids revealed a low affinity of the domain towards distinct DNA and RNA fragments (**Figure 3.8.2**), hinting towards an unspecific, electrostatically-mediated interaction. This is in line with other publications, where PWWP domains are shown to bind DNA just nonspecifically (no sequence selectivity) and with a wide range of different affinities, varying from low nanomolar to high micromolar [reviewed in [293]]. For example, the PWWP domain of HDGF cannot distinguish between A/T and C/G-rich sequences [327] and that of DNMT3A exhibits similar to PWWP2A's PWWP domain no preference for double-stranded DNA sequences [361]. In addition, also the PWWP domains of PSIP1 and ZMYND11 were shown to interact with DNA only in a nonspecific manner [294, 295, 300, 328]. Furthermore, PWWP2A's PWWP domain showed better complex-formation, detected by well-defined band shifts, for 75 bp dsDNA compared to 25 bp nucleic acid fragments (**Figure 3.8.2**), suggesting several PWWP domain proteins interacting with one single DNA fragment. This is in concordance with the finding that binding of DNA is mostly linked to DNA size, because multiple binding events can occur simultaneously [294, 295, 300, 327]. All in all, the contact of PWWP2A's PWWP domain to nucleic acids seems to occur rather nonspecific, most likely via electrostatic contacts of the negatively charged phosphate backbone and the basic surface of the domain.

Moreover, PWWP2A's PWWP domain preferred mononucleosomes containing 20 bp of linker DNA compared to linker-less nucleosomes, indicating the interaction of the PWWP domain with mononucleosomes is mostly mediated by binding to free flanking DNA (**Figure 3.8.3**). However, independent of flexible histone tails (**Figure 3.8.4**), the PWWP domain also showed an extremely weak interaction with mononucleosomes lacking flanking DNA, possibly suggesting a low interaction with wrapped nucleosomal DNA. This is in agreement with published structural data, where the PWWP domain was shown to contact the two wrapped

DISCUSSION

DNA duplexes at two distinct interaction areas [294, 295, 300]. Here, each PWWP binding interface displays a high degree of positively charged residues, which are complementary to the negatively charged phosphate backbone of nucleosomal DNA [294, 295, 300]. Concluding, the observation that PWWP2A's PWWP domain binds to distinct nucleic acids equally well and prefers free nucleic acids like free nucleosomal linker DNA (**Figure 3.8.3**), but does not require flexible histone tails (**Figure 3.8.4**), points towards a sequence-unrelated and rather electrostatic interaction.

Interestingly, the combination of both exhibited DNA binding domains (IN and PWWP) showed an additive effect in DNA binding strength (**Figure 3.8.5**). Possibly, IN acts as/similar to pioneer transcription factor(s), and initially binds to specific DNA sequences [324]. Following, this might enable the rather sequence-independent, electrostatic interaction of PWWP2A's PWWP domain with free linker DNA. Thus, the combination of both domains is most likely required to establish a strong interaction with nucleosomal nucleic acids to support PWWP2A's remarkably strong association with H2A.Z-containing chromatin.

The biological function of H2A.Z has been excessively studied and revealed its role in transcription regulation, DNA repair, mitosis and many other cellular processes [12]. Recent studies concentrated on H2A.Z's impact on transcription [362, 363], where the histone variant is involved in both activating as well as repressive regulation of cellular transcription [245]. Hu et al. already hypothesized that H2A.Z might act as a binding hub for different chromatin-modifying complexes whose attraction depends on additional factors and specific chromatin surroundings [282]. The direct interaction of H2A.Z with PWWP2A might result in special, PWWP2A-dependent biological functions within the cell. Due to PWWP2A's unique domains, which enable the interaction with many different chromatin features, PWWP2A is able to achieve an extremely strong association with chromatin, also observed in FRAP experiments [276]. Moreover, PWWP2A co-localizes with H2A.Z-containing nucleosomes at both promoters of highly transcribed genes [276] and regulatory regions (Link & Spitzer et al., manuscript under review). Further, several hundred genes are deregulated upon PWWP2A depletion [276], indicating a possible role in transcription regulation. Notably, PWWP2A does not exhibit any predicted enzymatic activity, but when identifying PWWP2A's interactome, many H2A.Z binders were found (Link & Spitzer et al., manuscript under review). Therefore, it

is highly possible that PWWP2A acts as binding hub or recruitment platform for H2A.Z-specific interactors thus helping to regulate H2A.Z-dependent transcription.

4.2 A combination of S and PWWP interacts in a sequence- and methylation state-specific manner with H3K36me3-nucleosomes

PWWP domains of different proteins show structural and sequential similarities, suggesting a potential role in the recognition of posttranslational modifications of histone residues [reviewed in [293]]. Interestingly, high-throughput mass spectrometry screenings identified the PWWP domain mainly as a reader of histone 3 lysine 36 trimethylation (H3K36me3) [308]. Notably, amino acid sequence alignments (**Figure 3.9.1**) revealed that also PWWP2A harbors three conserved aromatic residues, which possibly form an aromatic cage, suggesting a putative role as H3K36me3 sensor. Immunoblot analysis demonstrated, a strong enrichment for H3K36me3- but not H3K36me2- and H3K36me1-containing nucleosomes, indicating a sequence- and methylation state-specific interaction (**Figure 3.9.2**).

Interestingly, the specific enrichment of H3K36me3-containing mononucleosomes was observed in immunoblots only by a combination of the serine-rich stretch S and the PWWP domain, whereas MS data showed the same effect also for the PWWP domain alone. This can be explained by higher sensitivity of the MS approach, indicating a combination of S and PWWP is required to achieve the exceptional strong interaction with H3K36me3-containing mononucleosomes observed in immunoblots. Notably, it has been reported for several other proteins including ZMYND11 and further chromatin interactors like PHD fingers, that the power of a combinatorial readout by pairing different modules is crucial in the context of nucleosome/PTM recognition and non-histone partner targeting [297, 364]. ZMYND11 was reported to require at least three modules (bromo, zinc-finger and PWWP domain) to form a stable composite pocket [297]. In concordance, PWWP2A's PWWP domain alone might be not sufficient for specific H3K36me3 recognition. The reason might be that the PWWP domain possibly requires the adjacent serine-rich stretch to form a stable and functional aromatic cage structure, which is only then able to establish a strong interaction with H3K36me3-containing nucleosomes. Although the recognition of this PTM, mainly found at gene bodies of actively transcribed genes and involved in cryptic transcript initiation suppression [365, 366], is detected for recombinant S_PWWP, constructs that additionally include the H2A.Z-specific

DISCUSSION

recognition part IC (PWWP2A full-length and I_S_PWWP) are not capable of binding efficiently to H3K36me3-nucleosomes. This observation is in line with our previous results, where we could show in nChIP-seq that full-length PWWP2A mainly co-localizes with H2A.Z-containing nucleosomes at both promoters of highly transcribed genes [276] and regulatory regions, whereby only minor levels are found at H3K36me3-containing gene bodies (Link & Spitzer et al., manuscript under review). Possibly, either the specific interaction of IC with H2A.Z dominates over S_PWWP-H3K36me3 interaction due to different binding affinities or an auto-inhibitory element within the internal region might block the S_PWWP domain. This auto-inhibition could be regulated by dynamic covalent modifications of PWWP2A's internal region that modulate and coordinate its ability to interact with these two mutually exclusive chromatin moieties. Interestingly, it has been reported for other non-histone proteins including tumor suppressor p53 that modifications like lysine acetylation and/or phosphorylations are crucial for the regulation of proper protein function dependent on certain cellular events [367]. Notably, also the I domain contains a number of amino acid residues displaying possible sites that are proposed to be subject to lysine acetylation and also several possible phosphorylation sites as identified at PhosphoSitePlus® (www.phosphosite.org).

Interestingly, single point mutations of PWWP2A's aromatic cage residues, that did not affect proper protein folding, resulted in the abrogation of H3K36me3-nucleosome interaction for two of three mutants (**Figure 3.9.4**). This is in agreement with other publications that report single aromatic cage residues usually abolish the interaction with methylated histone peptides [296, 297, 307, 328, 368-370]. There, e.g. ZMYND11 lost specific recognition of H3K36me3 without displaying inappropriate folding, upon aromatic cage mutations [297]. Moreover, Gatchalian and co-workers found that single aromatic cage point mutations might but not have to result in loss of H3K36me3 interaction [371]. In concordance with this finding, this leads to the conclusion that PWWP2A most likely forms an aromatic cage with either only two residues being crucial for histone PTM recognition or all of them, with one displaying no loss of interaction upon amino acid exchange.

Wen et al. demonstrated that ZMYND11 is one of the first known PWWP-domain containing proteins that specifically recognizes a posttranslational histone modification on a histone variant [297]. There, ZMYND11 interacts only exceptionally strong with the posttranslational histone mark H3K36me3, if it is located at the histone tail of H3 variant H3.3. The reason for

DISCUSSION

this lies in H3.3's primary sequence, which differs from those of canonical H3.1/2. Analysis of H3.3 protein structure revealed, that the variant harbors a specific serine residue on position 31 (S31) and an adjacent threonine (T32). Here, ZMYND11 initially encapsulates S31 and T32 by a bromo-zinc finger PWWP pocket to ensure strong nucleosome binding. Following H3.3-interaction, ZMYND11's aromatic cage is now able to additionally recognize H3.3K36me3. Remarkably, the combination of H3.3 and K36me3 (H3.3K36me3) resulted in an eightfold higher binding affinity compared to canonical H3.1K36me3. Additionally, solely FLAG-tagged histone H3.3 but not H3.1 showed strong enrichment for ZMYND11 in pull-down approaches. Together, Wen and co-workers demonstrate that the H3.3-specific residue S31 is indispensable for the interaction of ZMYND11 with the posttranslational histone mark H3K36me3 [297]. Interestingly, cEMSA demonstrated that GST-tagged S_PWWP was not able to distinguish between recombinant mononucleosomes containing either H3K36me0 or H3K36me3 *in vitro* (**Figure 3.9.6**), indicating that the modification alone is not sufficient to establish a strong association with S_PWWP. Moreover, performing pull-down experiments I could already exclude a possible correlation between S_PWWP, H3K36me3 and histone variant H3.3, since no enrichment in H3.3-containing nucleosomes was observed upon immunoprecipitation with S_PWWP and HK-derived mononucleosomes (data not shown). However, keeping ZMYND11 in mind, PWWP2A might also need the presence of other histone variants (not H3.3) or, alternatively, modifications on itself to achieve a strong H3K36me3-nucleosome interaction. Summing up, the exact binding mechanism of PWWP2A's S_PWWP domain with the posttranslational histone modification H3K36me3 remains yet elusive. Hence, future experiments need to address the questions whether possibly incorporated histone variants or, alternatively, modifications could contribute to a strong association with H3K36me3-nucleosomes.

It is known, that single PTMs or domains can recruit complexes to defined target sites within the genome [108, 372]. For example, (1) different readers recognize unique histone marks; (2) complexes with contradictory tasks can share identical binding motifs and (3) single domains can interact with different PTMs [308, 372, 373]. Therefore, emerging evidence indicates that the "one domain – one mark mechanism" [372] might be inadequate to regulate the complex interplay of different PTMs in a cellular system. This finally proposed the existence of a complex histone code [108, 374], which postulates that posttranslational modifications recruit specific proteins to distinct chromatin sites and that those are finally

DISCUSSION

determining the functional output of the PTM. Hence, this system represents a crucial regulatory mechanism with far-reaching impact on chromatin-related processes [108]. Interestingly, pull-downs followed by If-MS (**Figure 3.9.5**) or immunoblotting (**Appendix Figure A1**) using recombinantly S_PWWP revealed a specific enrichment of H3K79me2-containing mononucleosomes, suggesting either an independent binding of both modifications or co-immunoprecipitating one modification with the other. The first hypothesis is also in line with literature, because the interaction with distinct methylation marks was already detected for PWWP domain-harboring proteins including bromo and plant homeodomain (PHD) finger-containing protein 1 (BRPF1) and 2 (BRPF2) [298]. In addition, distinct combinations of histone marks such as lysine H3-methylation or -acetylation, crucial for transcriptional regulation, are known [375]. There, histone mark combinations such as H3K4me3, H3K9ac and H3K14ac at promoters or H3K36me3 and H3K79me2 in gene bodies characterize transcriptionally active gene regions [331-334]. This supports the second hypothesis, that S_PWWP possibly co-immunoprecipitates H3K79me2-nucleosomes together with its most likely main PTM target H3K36me3. In summary, PWWP2A might either interact with one PTM like H3K36me3 thereby also pulling-down H3K79me2-nucleosomes since the modifications occur together in actively transcribed gene bodies [331-334] or with both modifications independently. Thereby, H3K36me3- and/or H3K79me2-containing nucleosomes possibly recruit PWWP2A transiently to gene bodies (Link & Spitzer et al., manuscript under review), where its main target H2A.Z is largely depleted [276]. However, the exact biological function PWWP2A might fulfill at gene bodies is still unclear. Therefore, PWWP2A's detailed recruitment mechanism to gene body regions containing H3K36me3/H3K79me2-nucleosomes and possibly therefrom resulting functions have to be addressed in future studies.

4.3 PWWP2A's C-terminus interacts with the modified histone variant γ -H2A.X

PWWP2A's internal region IC is a strong and specific interactor of H2A.Z-containing nucleosomes. Hence, constructs missing this protein part lack the ability of enriching H2A.Z-nucleosomes in pull-down experiments. Remarkably, S_PWWP and PWWP pull-down approaches were not only depleted of H2A.Z-, but also H2A-nucleosomes, suggesting the interaction with another member of the H2A histone variant family (**Figure 3.10.1**).

DISCUSSION

Immunoprecipitations of recombinant S_PWWP and PWWP with cell-derived mononucleosomes revealed the interaction with histone variant γ -H2A.X, which is the phosphorylated (S139ph) form of H2A.X and involved in DNA damage repair [180, 182]. Interestingly, PWWP2A's C-terminus showed only enrichment in γ -H2A.X- but not H2A.X-containing nucleosomes, proposing the phosphorylation at position 139 of H2A.X is a main mediator of PWWP2A - γ -H2A.X interaction. This is also in concordance with literature, where it was shown that PTMs are fundamental for fine-tuning the interaction with certain interaction proteins [376]. To date it is unclear, how the interaction between PWWP2A and γ -H2A.X is established in detail and whether it is direct or indirect. Nonetheless, the specific phosphorylation mark of variant γ -H2A.X seems to be crucial for PWWP2A recruitment. In conclusion, PWWP2A might be either directly recruited via S139ph interaction using a binding motif in its C-terminus (**Figure 3.10.1**) or it interacts with γ -H2A.X-specific interactors. Interestingly, If-qMS pull-downs (method published in [276]) using GFP-PWWP2A revealed that amongst many other proteins also MDC1 was enriched (Link & Spitzer et al., under review), which is known to be a γ -H2A.X-specific interactor [377]. Notably, PWWP2A does not possess any of the currently known phosphoserine/threonine-binding features such as 4-3-3 domains, WW domains, BRCT and fork-head-associated (FHA) domains or WD40 repeats [378, 379], proposing rather an indirect interaction with γ -H2A.X via variant-specific binders or containing a novel, yet uncharacterized phosphor-binding module. In contrast to PWWP2A, MDC1 harbors two phosphor-specific recognition domains in its N- (FHA domain) and C-terminus (BRCT) and was already demonstrated to be a direct interactor of S139ph on H2A.X [378]. Interestingly, DNA damage response-related proteins including tumor suppressor p53 are reported to get activated by certain modifications such as phosphorylation or acetylation upon DNA damage [367]. Like already mentioned above, also PWWP2A harbors several target sites for potential acetylations or phosphorylations (www.phosphosite.org). Probably PWWP2A is phosphorylated upon DNA damage-causing cellular events such as stress, irradiation, starvation or contact inhibition and might be a potential target of MDC1. Moreover, it has been shown that MDC1 interacts only via its C-terminal BRCT domain with S139ph [378]. Hence, MDC1's N-terminally located FHA domain might be a possible interaction partner of phosphorylated PWWP2A thereby recruiting the latter to foci of DNA damage and high γ -H2A.X levels [180, 182]. This is also in agreement with the finding that phosphorylated proteins can establish distinct protein-protein interaction networks [380].

DISCUSSION

Furthermore, the interaction of PWWP2A and γ -H2A.X is not dependent on H3K36me3-containing nucleosomes or the aromatic cage, because no decrease in γ -H2A.X immunoblot signal intensity was observed upon aromatic cage point mutations (**Figure 3.10.1**). This suggests rather an aromatic cage-unrelated interaction of PWWP2A with γ -H2A.X-nucleosomes and an independent recognition of H3K36me3-containing nucleosomes.

4.4 Summary and future perspectives

While direct binding of H2A.Z-containing nucleosomes via the internal region seems to be the underlying mechanism for PWWP2A's recruitment to promoters of highly transcribed genes [276] and regulatory regions (Link & Spitzer et al., manuscript under review), the interactions of S_PWWP with H3K36me3 and PWWP2A's C-terminus with γ -H2A.X are still mechanistically and functionally puzzling. Nevertheless, the following proposed model should give an overview of the current state of knowledge and is mainly based on the results of this thesis.

Model

The model proposes three independent binding mechanisms for PWWP2A with H2A.Z-, H3K36me3- or γ -H2A.X-containing nucleosomes (**Figure 4.4.1**).

PWWP2A is mainly found at promoters of highly transcribed genes [276] as well as regulatory regions (Link & Spitzer et al., manuscript under review), where it co-localizes with histone variant H2A.Z. To achieve an exceptional strong association with H2A.Z-containing chromatin, PWWP2A establishes multivalent binding mechanisms via its multiple domains. The specific and strong interaction with histone variant H2A.Z is achieved via the IC domain and supported by various additional, but rather nonspecific interactions of the adjacent IN and PWWP domain. The IN as well as PWWP domain harbor both DNA binding properties due to positively charged residues and show in combination even a stronger association with DNA. Since PWWP2A does not have any predicted enzymatically active domain, but several hundred genes are deregulated upon PWWP2A depletion [276], it is proposed to act as a binding hub for regulatory proteins to histone variant H2A.Z. This was also supported by the identification of PWWP2A's interactome, where also many H2A.Z binders were found (Link & Spitzer et al.,

DISCUSSION

manuscript under review). Besides PWWP2A's important role at H2A.Z-containing promoters, it is further found, albeit at low levels, in H3K36me3-enriched gene bodies (Link & Spitzer et al., manuscript under review). The interaction with H3K36me3-nucleosomes is rather transient, whereby the internal region additionally complicates the binding to them due to its high affinity for H2A.Z. The posttranslational histone modification mark H3K36me3 is known to have important functions in actively transcribed genomic regions, mRNA splicing and transcription elongation [126, 381, 382] and can act as binding or recruitment platform for certain interaction proteins or complexes [372]. Probably, before being able to be recruited to H3K36me3-nucleosomes, PWWP2A must be initially modified by adding or removing distinct modifications. Thereby, a possible auto-inhibitory effect of the internal region might be abolished and following, PWWP2A is able to recognize H3K36me3-nucleosomes in an aromatic cage-dependent manner. However, to achieve a strong and specific association with H3K36me3-nucleosomes, PWWP2A's S_PWWP domain possibly requires additional, yet unknown factors like distinct PTMs or histone variants. After PWWP2A bound to H3K36me3-nucleosomes it might recruit further interactors to these gene body regions. PWWP2A's third chromatin target is γ -H2A.X, the modified version of H2A.X which is phosphorylated (S139ph) upon DNA damage [180, 183]. Notably, PWWP2A does not harbor any of the currently known phosphoserine/threonine-binding domains. Therefore, the interaction with γ -H2A.X is rather achieved indirectly. A possible underlying mechanism might be the interaction with γ -H2A.X specific interactor MDC1 [378], which is also part of PWWP2A's interactome (Link & Spitzer et al., manuscript under review). MDC1 is initially recruited by MOF-mediated H4K16 acetylation to foci of DNA damage where its direct interaction partner γ -H2A.X is located [378, 383]. Possibly, similar to other proteins, also PWWP2A might be modified upon DNA damage [367]. Hence, PWWP2A maybe is also modified within the internal region (similar to H3K36me3 recruitment) to abolish its auto-inhibitory function in response to DNA damage. In addition, PWWP2A might be phosphorylated within its last C-terminal amino acids (PWWP_short, aa 655 - 755) thereby displaying a possible target for MDC1. In turn, MDC1, which is already known to act as recruitment platform for additional proteins [384] might recruit PWWP2A to sites of DNA damage. There, it might help to mediate the recruitment of factors such as DNA damage- or DNA repair-related protein complexes.

In summary, PWWP2A predominantly might act as a binding hub for additional, site-specific (H2A.Z-, H3K36me3- and γ -H2A.X-containing chromatin regions) interaction proteins. In turn,

DISCUSSION

PWWP2A's recruitment to specific genomic regions might be regulated by different histone variants (H2A.Z and γ -H2A.X) or posttranslational histone modifications (H3K36me3) as well as auto-regulatory modifications (e.g. phosphorylations and acetylations).

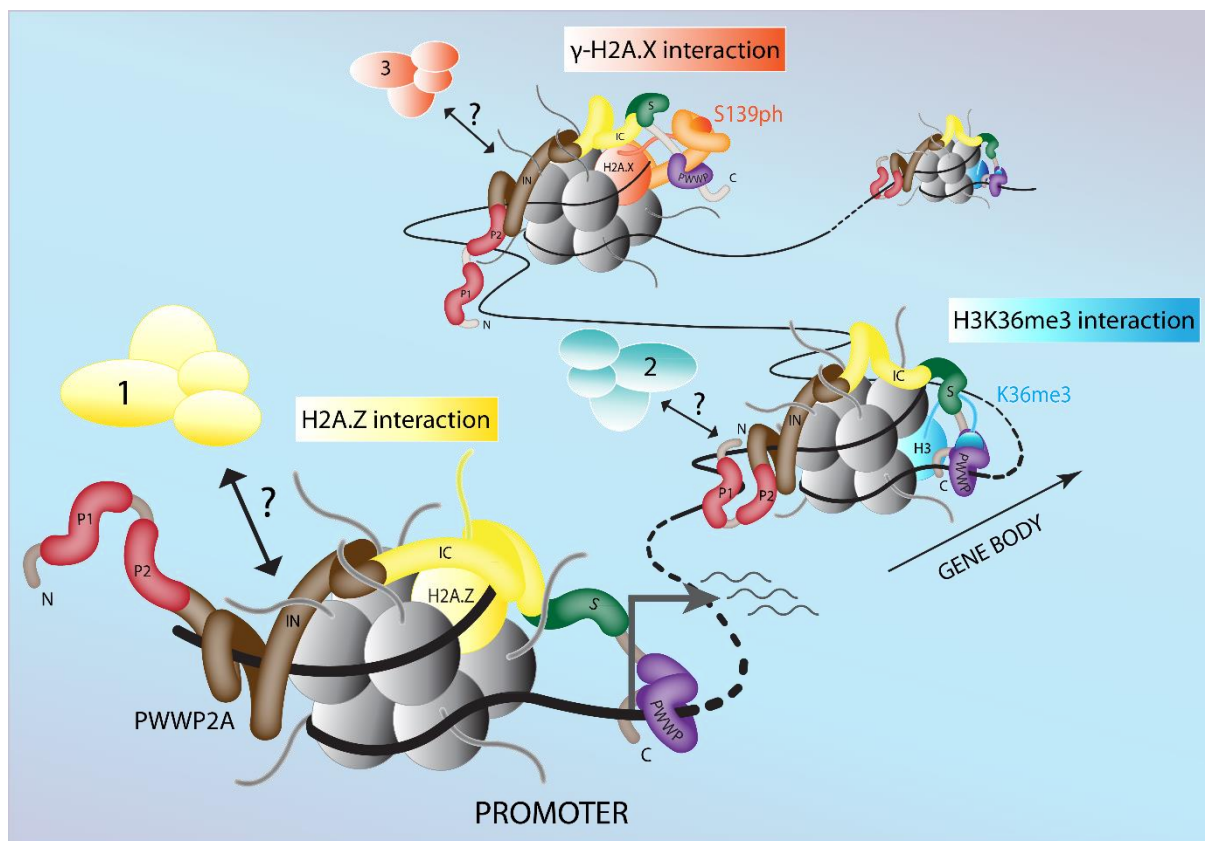


Figure 4.4.1 PWWP2A is independently recruited to distinct chromatin moieties. PWWP2A is able to interact with distinct chromatin regions due to its multiple domains. It weakly binds to H3K36me3-enriched gene body regions with its S_PWWP domain in an aromatic cage-dependent manner. Strong chromatin association is further established by PWWP2A's IN and PWWP domain, which recognize free nucleosomal flanking DNA. In addition, PWWP2A binds strongly to H2A.Z-containing promoters as well as regulatory regions, whereby IC mediates specific H2A.Z recognition. Moreover, PWWP2A interacts with histone variant γ -H2A.X-containing nucleosomes using its C-terminal half (PWWP_short). On chromatin, PWWP2A most likely acts as adapter for different nucleosome components (H2A.Z, H3K36me3 and γ -H2A.X) and chromatin-modifying complexes.

Future perspectives

In the future, it would be interesting to analyze the binding of IC to H2A.Z in more in detail. Are H2A.Z's amino acid residues G120 and K121 the main mediators of the interaction with IC or is the reduction in H2A.Z Δ C binding to PWWP2A rather indirect? To test, whether the aforementioned residues are the main actors for IC attraction, specifically mutated recombinant mononucleosomes have to be generated. Here, either recombinant mononucleosomes lacking nine instead of seven C-terminal amino acid residues or

DISCUSSION

nucleosomes having H2A.Z G120 and K121 point mutated e.g. to alanine would help to investigate IC's interaction with H2A.Z. In addition, it would be interesting whether an exchange of H2A's C-terminal tail with that of H2A.Z (harboring IC's interaction motif) might enhance the binding of IC to H2A.

I could show, that two amino acid residues are important for proper nucleosome binding within the IC domain. Amino acid residue R461 and R475 are crucial to establish a strong association with recombinant mononucleosomes. Crystallography of wildtype internal region and single amino acid substituted internal region mutants together with H2A.Z-containing mononucleosomes would reveal (1) how the exact binding of IC to H2A.Z works, (2) if binding of IC influences structural features of IN and (3) how the binding is impaired upon mutation. In addition to R461 and R475, further mutated residues within the IC region could be already identified in distinct cancer types (H425Y, A435T, E437K, T448A, P472S, E495Q, R498Q, K505N, R509H, R514H, E536D, A548S and R559I, identified by www.cbioportal.com). Therefore, it would be interesting to investigate, whether those point mutations also result in decreased interactions with recombinant nucleosomes and to distinguish the residue(s) contributing most to H2A.Z-recognition. Notably, the group of Zheng Zhou is highly interested in crystallizing PWWP2A bound to a H2A.Z-containing mononucleosome in order to get the specific 3D structure of the complex. Thereby, the 3D structure of PWWP2A would help to understand the complex binding mechanisms of PWWP2A's multiple domains.

Additionally, I identified another histone variant binding partner of PWWP2A. To date it is unclear, whether the interaction of PWWP2A with γ -H2A.X-nucleosomes is mediated in a direct or indirect manner. In order to investigate the underlying binding mechanism, it is crucial to reconstitute recombinant mononucleosomes containing either H2A, H2A.X or γ -H2A.X and to perform cEMSA. Here, the phosphorylation on S139 of recombinant γ -H2A.X can be generated either by phosphokinases or, more specifically, by single amino acid substitution (S139D) since aspartate acts as phospho-mimicry (e.g. described in [385]). Using the different types of recombinant mononucleosomes and GST-tagged PWWP_short protein, we could determine whether the interaction with γ -H2A.X-containing nucleosomes is achieved directly, like for H2A.Z, or if another protein is needed to mediate the interaction. If the interaction between PWWP2A and γ -H2A.X-containing nucleosomes is indirect, the intermediary factor has to be elucidated. Interestingly, pull-downs of GFP-PWWP2A followed

DISCUSSION

by If-qMS already demonstrated that MDC1 could be a potential candidate, since it interacts with PWWP2A and is known to be a direct binder of γ -H2A.X. Initially, it would be quite easy to test, whether MDC1 is enriched in the same PWWP2A pull-downs as γ -H2A.X. Here, C-terminal GST-tagged PWWP2A truncation proteins (S_short, S_long, S_PWWP, PWWP and PWWP_short) could be incubated with HK cell-derived mononucleosomes and subsequently analyzed in immunoblots using specific antibodies for MDC1 and γ -H2A.X. If these approaches reveal similar enrichments for MDC1, a next step is to test whether the interaction is direct or indirect. To test a possible direct interaction, GST-tagged PWWP_short could be immunoprecipitated with recombinant, e.g. FLAG-tagged MDC1 and following analyzed in immunoblots.

Moreover, it would be interesting to see, whether PWWP2A goes to sites of DNA damage *in vivo*. Therefore, HK cells could be treated with the DNA damage-causing agent aphidicolin (polymerase inhibitor that reversibly blocks DNA replication [386]) followed by immunofluorescence stainings (PWWP2A, γ -H2A.X and MDC1). Co-localization of PWWP2A with γ -H2A.X (and MDC1) would indicate PWWP2A's recruitment to sites of DNA damage. Furthermore, the same assay could be repeated upon MDC1 knock-down. In case PWWP2A is recruited by MDC1 to DNA damage regions, PWWP2A would not anymore co-localize with γ -H2A.X-foci upon MDC1 knock-down.

In summary, I demonstrated that PWWP2A is an exceptional strong chromatin interactor using at least five different domains (IN, IC, S_PWWP, PWWP) to allow binding of six different chromatin features: H2A.Z, free linker DNA, histones, H3K36me3 (H3K79me2), nucleic acids and γ -H2A.X.

ABBREVIATIONS

ABBREVIATIONS

%	percent
(v/v)	(volume/volume)
(w/v)	(weight/volume)
~	approximately
°C	degree celsius
µg	micro gram
µm	micro meter
3D	3-dimensional
5caC	5-carboxylcytosine
5fC	5-fluorcytosine
5mC	5 methylcytosine
A	ampere
aa	amino acid
APBS	architectural protein binding sites
APS	ammonium persulfate
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
bp	base pair(s)
BSA	bovine serum albumine
cEMSA	competitive Electrophoretic Mobility Shift Assay
cm	centimeter
CO ₂	carbon dioxide
CpG	cytosine phosphate guanine
CTCF	CCCTC-binding factor
DDR	DNA damage repair
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyl transferase
dNTP	deoxynucleotide
DSB	double strand break
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EGTA	egtazic acid
ESP	electrostatic surface potential
FCS	fetal calf serum
Fe ²⁺	ferrum
G1	gap 1

ABBREVIATIONS

G2	gap 2
HCl	hydrochloric acid
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HFD	histone fold domain
HK	HeLa Kyoto
IP	immunoprecipitation
IPTG	isopropyl β -D-1-thiogalactopyranoside
K	potassium
Kb	kilo base
kDa	kilo Dalton
km	kilometer
kV	kilo volt
L1	loop 1
LB	lysogeny broth
lf-qMS	label free quantitative mass spectrometry
lncRNAs	long non-coding ribonucleic acids
m	meter
M	marker
M	molar
Mnase	micrococcal nuclease
m/z	mass to charge
mA	milliampere
MgCl ₂	magnesium chloride
Min	minute
ml	milliliter
mM	millimolar
mRNAs	micro ribonucleic acids
MWCO	molecular weight cut off
NaCl	sodium chloride
NCP	nucleosome core particle
NDR	nucleosome depleted region
ng	nanogram(s)
NH ₄ HCO ₃	ammonium bicarbonate
nm	nanometer
nM	nanomolar
NP-40	nonyl phenoxypolyethoxyethanol
PAA	poly acrylamide
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline tween
PCR	polychain reaction
PEG	polyethylen glycole
pH	potentia hydrogenii

ABBREVIATIONS

PMSF	phenylmethylsulfonyl fluoride
P/S	penicillin/streptomycin
PTM	posttranslational modification
PWWP	(P) proline - (W) tryptophane - (W) tryptophane - (P) proline
rev	reverse
RNA	ribonucleic acid
Rpm	rounds per minute
RT	room temperature
SANT	"Swi3, Ada2, N-Cor, and TFIIB"
SDS	sodium Dodecyl Sulfate
SDS-PAGE	sodium Dodecyl Sulfate polyacrylamide gel electrophoresis
Sec	second
TBE	tetrabromoethane
TE	tris-EDTA
TEMED	tetramethylethylenediamine
TF	transcription factor
TFA	trifluoroacetic acid
TL	tail-less
T _m	melting temperature
TSS	transcription start site
V	volt
VE	voll entsalzt
WD40	(W) tryptophan - (D) aspartic acid repeat
WT	wildtype

BIBLIOGRAPHY

1. Crick, F.H., et al., *General nature of the genetic code for proteins*. Nature, 1961. **192**: p. 1227-32.
2. Andrews, A.J. and K. Luger, *Nucleosome structure(s) and stability: variations on a theme*. Annu Rev Biophys, 2011. **40**: p. 99-117.
3. Fussner, E., R.W. Ching, and D.P. Bazett-Jones, *Living without 30nm chromatin fibers*. Trends Biochem Sci, 2011. **36**(1): p. 1-6.
4. Alberts, B., et al., *Chromosomal DNA and Its Packaging in the Chromatin Fiber*. . Molecular Biology of the Cell. 4th edition 2002: New York: Garland Science.
5. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
6. Venter, J.C., et al., *The sequence of the human genome*. Science, 2001. **291**(5507): p. 1304-51.
7. Marstrand, T.T. and J.D. Storey, *Identifying and mapping cell-type-specific chromatin programming of gene expression*. Proc Natl Acad Sci U S A, 2014. **111**(6): p. E645-54.
8. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA*. Science, 1974. **184**(4139): p. 868-71.
9. Oudet, P., M. Gross-Bellard, and P. Chambon, *Electron microscopic and biochemical evidence that chromatin structure is a repeating unit*. Cell, 1975. **4**(4): p. 281-300.
10. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 Å resolution*. Nature, 1997. **389**(6648): p. 251-60.
11. Richmond, T.J. and C.A. Davey, *The structure of DNA in the nucleosome core*. Nature, 2003. **423**(6936): p. 145-50.
12. Bonisch, C. and S.B. Hake, *Histone H2A variants in nucleosomes and chromatin: more or less stable?* Nucleic Acids Res, 2012. **40**(21): p. 10719-41.
13. Arents, G., et al., *The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix*. Proc Natl Acad Sci U S A, 1991. **88**(22): p. 10148-52.

BIBLIOGRAPHY

14. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution*. J Mol Biol, 2002. **319**(5): p. 1097-113.
15. Eickbush, T.H. and E.N. Moudrianakis, *The histone core complex: an octamer assembled by two sets of protein-protein interactions*. Biochemistry, 1978. **17**(23): p. 4955-64.
16. Olins, A.L. and D.E. Olins, *Spheroid chromatin units (v bodies)*. Science, 1974. **183**(4122): p. 330-2.
17. Tremethick, D.J., *Higher-order structures of chromatin: the elusive 30 nm fiber*. Cell, 2007. **128**(4): p. 651-4.
18. Zhu, P. and G. Li, *Structural insights of nucleosome and the 30-nm chromatin fiber*. Curr Opin Struct Biol, 2016. **36**: p. 106-15.
19. Li, G. and D. Reinberg, *Chromatin higher-order structures and gene regulation*. Curr Opin Genet Dev, 2011. **21**(2): p. 175-86.
20. Robinson, P.J. and D. Rhodes, *Structure of the '30 nm' chromatin fibre: a key role for the linker histone*. Curr Opin Struct Biol, 2006. **16**(3): p. 336-43.
21. Maeshima, K., S. Hihara, and M. Eltsov, *Chromatin structure: does the 30-nm fibre exist in vivo?* Curr Opin Cell Biol, 2010. **22**(3): p. 291-7.
22. Finch, J.T. and A. Klug, *Solenoidal model for superstructure in chromatin*. Proc Natl Acad Sci U S A, 1976. **73**(6): p. 1897-901.
23. Williams, S.P., et al., *Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length*. Biophys J, 1986. **49**(1): p. 233-48.
24. Woodcock, C.L., et al., *A chromatin folding model that incorporates linker variability generates fibers resembling the native structures*. Proc Natl Acad Sci U S A, 1993. **90**(19): p. 9021-5.
25. Robinson, P.J., et al., *EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure*. Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6506-11.
26. Schalch, T., et al., *X-ray structure of a tetranucleosome and its implications for the chromatin fibre*. Nature, 2005. **436**(7047): p. 138-41.

BIBLIOGRAPHY

27. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control*. Nat Rev Genet, 2016. **17**(8): p. 487-500.
28. Grewal, S.I. and S.C. Elgin, *Heterochromatin: new possibilities for the inheritance of structure*. Curr Opin Genet Dev, 2002. **12**(2): p. 178-87.
29. Elgin, S.C., *Heterochromatin and gene regulation in Drosophila*. Curr Opin Genet Dev, 1996. **6**(2): p. 193-202.
30. Tamaru, H., *Confining euchromatin/heterochromatin territory: jumonji crosses the line*. Genes Dev, 2010. **24**(14): p. 1465-78.
31. Trojer, P. and D. Reinberg, *Facultative heterochromatin: is there a distinctive molecular signature?* Mol Cell, 2007. **28**(1): p. 1-13.
32. Weintraub, H. and M. Groudine, *Chromosomal subunits in active genes have an altered conformation*. Science, 1976. **193**(4256): p. 848-56.
33. Waddington, C.H., *The epigenotype*. Endeavour, 1942. **1**: p. 18-20.
34. Riggs, A.D., R.A. Martienssen, and V.E.A. Russo, *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, 1996. **32**: p. 1-4.
35. Probst, A.V., E. Dunleavy, and G. Almouzni, *Epigenetic inheritance during the cell cycle*. Nat Rev Mol Cell Biol, 2009. **10**(3): p. 192-206.
36. Bonisch, C., et al., *Chromatin proteomics and epigenetic regulatory circuits*. Expert Rev Proteomics, 2008. **5**(1): p. 105-19.
37. Feng, S., S.E. Jacobsen, and W. Reik, *Epigenetic reprogramming in plant and animal development*. Science, 2010. **330**(6004): p. 622-7.
38. Tirado-Magallanes, R., et al., *Whole genome DNA methylation: beyond genes silencing*. Oncotarget, 2017. **8**(3): p. 5629-5637.
39. Ramsahoye, B.H., et al., *Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a*. Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5237-42.
40. Ziller, M.J., et al., *Genomic distribution and inter-sample variation of non-CpG methylation across human cell types*. PLoS Genet, 2011. **7**(12): p. e1002389.

BIBLIOGRAPHY

41. Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development*. Nat Rev Genet, 2013. **14**(3): p. 204-20.
42. Yin, Y., et al., *Impact of cytosine methylation on DNA binding specificities of human transcription factors*. Science, 2017. **356**(6337).
43. Stadler, M.B., et al., *DNA-binding factors shape the mouse methylome at distal regulatory regions*. Nature, 2011. **480**(7378): p. 490-5.
44. Klose, R.J. and A.P. Bird, *Genomic DNA methylation: the mark and its mediators*. Trends Biochem Sci, 2006. **31**(2): p. 89-97.
45. Schubeler, D., *Function and information content of DNA methylation*. Nature, 2015. **517**(7534): p. 321-6.
46. Bahar Halpern, K., T. Vana, and M.D. Walker, *Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development*. J Biol Chem, 2014. **289**(34): p. 23882-92.
47. Chodavarapu, R.K., et al., *Relationship between nucleosome positioning and DNA methylation*. Nature, 2010. **466**(7304): p. 388-92.
48. Jimenez-Useche, I., et al., *DNA methylation effects on tetra-nucleosome compaction and aggregation*. Biophys J, 2014. **107**(7): p. 1629-36.
49. Lev Maor, G., A. Yearim, and G. Ast, *The alternative role of DNA methylation in splicing regulation*. Trends Genet, 2015. **31**(5): p. 274-80.
50. Gopalakrishnan, S., et al., *DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions*. Hum Mol Genet, 2009. **18**(17): p. 3178-93.
51. Law, J.A. and S.E. Jacobsen, *Establishing, maintaining and modifying DNA methylation patterns in plants and animals*. Nat Rev Genet, 2010. **11**(3): p. 204-20.
52. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality*. Cell, 1992. **69**(6): p. 915-26.
53. Okano, M., et al., *Assignment of cytosine-5 DNA methyltransferases Dnmt3a and Dnmt3b to mouse chromosome bands 12A2-A3 and 2H1 by in situ hybridization*. Cytogenet Cell Genet, 1999. **86**(3-4): p. 333-4.

BIBLIOGRAPHY

54. Sharif, J., et al., *The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA*. Nature, 2007. **450**(7171): p. 908-12.
55. An, J., A. Rao, and M. Ko, *TET family dioxygenases and DNA demethylation in stem cells and cancers*. Exp Mol Med, 2017. **49**(4): p. e323.
56. Delatte, B., R. Deplus, and F. Fuks, *Playing TETris with DNA modifications*. EMBO J, 2014. **33**(11): p. 1198-211.
57. He, Y.F., et al., *Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA*. Science, 2011. **333**(6047): p. 1303-7.
58. Ito, S., et al., *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine*. Science, 2011. **333**(6047): p. 1300-3.
59. Ko, M., et al., *Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2*. Nature, 2010. **468**(7325): p. 839-43.
60. Mitrea, C., et al., *Integrating 5hmC and gene expression data to infer regulatory mechanisms*. Bioinformatics, 2017.
61. Penn, N.W., et al., *The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid*. Biochem J, 1972. **126**(4): p. 781-90.
62. Bhattacharyya, S., et al., *Genome-wide hydroxymethylation tested using the HELP-GT assay shows redistribution in cancer*. Nucleic Acids Res, 2013. **41**(16): p. e157.
63. Lister, R., et al., *Global epigenomic reconfiguration during mammalian brain development*. Science, 2013. **341**(6146): p. 1237905.
64. Raiber, E.A., et al., *5-Formylcytosine alters the structure of the DNA double helix*. Nat Struct Mol Biol, 2015. **22**(1): p. 44-49.
65. Kellinger, M.W., et al., *5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription*. Nat Struct Mol Biol, 2012. **19**(8): p. 831-3.
66. Boon, R.A., et al., *Long Noncoding RNAs: From Clinical Genetics to Therapeutic Targets?* J Am Coll Cardiol, 2016. **67**(10): p. 1214-1226.
67. Mercer, T.R., M.E. Dinger, and J.S. Mattick, *Long non-coding RNAs: insights into functions*. Nat Rev Genet, 2009. **10**(3): p. 155-9.

BIBLIOGRAPHY

68. Cao, J., *The functional role of long non-coding RNAs and epigenetics*. Biol Proced Online, 2014. **16**: p. 11.
69. Ma, L., V.B. Bajic, and Z. Zhang, *On the classification of long non-coding RNAs*. RNA Biol, 2013. **10**(6): p. 925-33.
70. Struhl, K., *Transcriptional noise and the fidelity of initiation by RNA polymerase II*. Nature Structural & Molecular Biology, 2007. **14**(2): p. 103-105.
71. Bohmdorfer, G. and A.T. Wierzbicki, *Control of Chromatin Structure by Long Noncoding RNA*. Trends Cell Biol, 2015. **25**(10): p. 623-32.
72. Johnsson, P., et al., *Evolutionary conservation of long non-coding RNAs; sequence, structure, function*. Biochim Biophys Acta, 2014. **1840**(3): p. 1063-71.
73. Pang, K.C., M.C. Frith, and J.S. Mattick, *Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function*. Trends Genet, 2006. **22**(1): p. 1-5.
74. Froberg, J.E., L. Yang, and J.T. Lee, *Guided by RNAs: X-Inactivation as a Model for lncRNA Function*. Journal of Molecular Biology, 2013. **425**(19): p. 3698-3706.
75. Lyon, M.F., *Gene Action in X-Chromosome of Mouse (Mus Musculus L)*. Nature, 1961. **190**(477): p. 372-&.
76. Luikenhuis, S., A. Wutz, and R. Jaenisch, *Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells*. Molecular and Cellular Biology, 2001. **21**(24): p. 8512-8520.
77. Stavropoulos, N., N.F. Lu, and J.T. Lee, *A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice*. Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(18): p. 10232-10237.
78. Wutz, A. and R. Jaenisch, *A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation*. Molecular Cell, 2000. **5**(4): p. 695-705.
79. Clemson, C.M., et al., *XIST RNA paints the inactive X chromosome at interphase: Evidence for a novel RNA involved in nuclear chromosome structure*. Journal of Cell Biology, 1996. **132**(3): p. 259-275.
80. Huarte, M., *The emerging role of lncRNAs in cancer*. Nat Med, 2015. **21**(11): p. 1253-61.

BIBLIOGRAPHY

81. McBryant, S.J., V.H. Adams, and J.C. Hansen, *Chromatin architectural proteins*. *Chromosome Res*, 2006. **14**(1): p. 39-51.
82. Van Bortle, K., et al., *Insulator function and topological domain border strength scale with architectural protein occupancy*. *Genome Biol*, 2014. **15**(6): p. R82.
83. Gomez-Diaz, E. and V.G. Corces, *Architectural proteins: regulators of 3D genome organization in cell fate*. *Trends Cell Biol*, 2014. **24**(11): p. 703-11.
84. Ong, C.T. and V.G. Corces, *CTCF: an architectural protein bridging genome topology and function*. *Nat Rev Genet*, 2014. **15**(4): p. 234-46.
85. Wang, H., et al., *Widespread plasticity in CTCF occupancy linked to DNA methylation*. *Genome Res*, 2012. **22**(9): p. 1680-8.
86. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions*. *Nature*, 2012. **485**(7398): p. 376-380.
87. Nora, E.P., et al., *Spatial partitioning of the regulatory landscape of the X-inactivation centre*. *Nature*, 2012. **485**(7398): p. 381-5.
88. Guelen, L., et al., *Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions*. *Nature*, 2008. **453**(7197): p. 948-51.
89. Becker, P.B. and W. Horz, *ATP-dependent nucleosome remodeling*. *Annu Rev Biochem*, 2002. **71**: p. 247-73.
90. Becker, P.B. and J.L. Workman, *Nucleosome remodeling and epigenetics*. *Cold Spring Harb Perspect Biol*, 2013. **5**(9).
91. Langst, G. and L. Manelyte, *Chromatin Remodelers: From Function to Dysfunction*. *Genes (Basel)*, 2015. **6**(2): p. 299-324.
92. Clapier, C.R. and B.R. Cairns, *The biology of chromatin remodeling complexes*. *Annu Rev Biochem*, 2009. **78**: p. 273-304.
93. Eisen, J.A., K.S. Sweder, and P.C. Hanawalt, *Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions*. *Nucleic Acids Res*, 1995. **23**(14): p. 2715-23.
94. Flaus, A., et al., *Identification of multiple distinct Snf2 subfamilies with conserved structural motifs*. *Nucleic Acids Res*, 2006. **34**(10): p. 2887-905.

BIBLIOGRAPHY

95. Filippakopoulos, P. and S. Knapp, *The bromodomain interaction module*. FEBS Lett, 2012. **586**(17): p. 2692-704.
96. Wilson, B.G. and C.W.M. Roberts, *SWI/SNF nucleosome remodellers and cancer*. Nature Reviews Cancer, 2011. **11**(7): p. 481-492.
97. Shain, A.H., et al., *Convergent structural alterations define SWItch/Sucrose NonFermentable (SWI/SNF) chromatin remodeler as a central tumor suppressive complex in pancreatic cancer*. Proc Natl Acad Sci U S A, 2012. **109**(5): p. E252-9.
98. Strohner, R., et al., *Recruitment of the nucleolar remodeling complex NoRC establishes ribosomal DNA silencing in chromatin*. Molecular and Cellular Biology, 2004. **24**(4): p. 1791-1798.
99. Li, J., G. Langst, and I. Grummt, *NoRC-dependent nucleosome positioning silences rRNA genes*. EMBO J, 2006. **25**(24): p. 5735-41.
100. Sims, J.K. and P.A. Wade, *SnapShot: Chromatin remodeling: CHD*. Cell, 2011. **144**(4): p. 626-626 e1.
101. Gaspar-Maia, A., et al., *Chd1 regulates open chromatin and pluripotency of embryonic stem cells*. Nature, 2009. **460**(7257): p. 863-U97.
102. West, S.C., *Processing of recombination intermediates in vitro*. Bioessays, 1990. **12**(4): p. 151-4.
103. Gospodinov, A., et al., *Mammalian Ino80 Mediates Double-Strand Break Repair through Its Role in DNA End Strand Resection*. Molecular and Cellular Biology, 2011. **31**(23): p. 4735-4745.
104. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications*. Cell Res, 2011. **21**(3): p. 381-95.
105. Ng, M.K. and P. Cheung, *A brief histone in time: understanding the combinatorial functions of histone PTMs in the nucleosome context*. Biochem Cell Biol, 2016. **94**(1): p. 33-42.
106. Xu, Y.M., J.Y. Du, and A.T. Lau, *Posttranslational modifications of human histone H3: an update*. Proteomics, 2014. **14**(17-18): p. 2047-60.
107. Jack, A.P. and S.B. Hake, *Getting down to the core of histone modifications*. Chromosoma, 2014. **123**(4): p. 355-71.

BIBLIOGRAPHY

108. Jenuwein, T. and C.D. Allis, *Translating the histone code*. Science, 2001. **293**(5532): p. 1074-1080.
109. Henikoff, S. and A. Shilatifard, *Histone modification: cause or cog?* Trends Genet, 2011. **27**(10): p. 389-96.
110. Berger, S.L., *The complex language of chromatin regulation during transcription*. Nature, 2007. **447**(7143): p. 407-12.
111. Deshaies, R.J. and C.A. Joazeiro, *RING domain E3 ubiquitin ligases*. Annu Rev Biochem, 2009. **78**: p. 399-434.
112. Helin, K. and D. Dhanak, *Chromatin proteins and modifications as drug targets*. Nature, 2013. **502**(7472): p. 480-8.
113. Kouzarides, T., *Chromatin modifications and their function*. Cell, 2007. **128**(4): p. 693-705.
114. Pek, J.W., A. Anand, and T. Kai, *Tudor domain proteins in development*. Development, 2012. **139**(13): p. 2255-66.
115. Bonasio, R., E. Lecona, and D. Reinberg, *MBT domain proteins in development and disease*. Semin Cell Dev Biol, 2010. **21**(2): p. 221-30.
116. Bannister, A.J., et al., *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain*. Nature, 2001. **410**(6824): p. 120-4.
117. Tran, H.G., et al., *The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor*. EMBO J, 2000. **19**(10): p. 2323-31.
118. Kim, J., et al., *Tudor, MBT and chromo domains gauge the degree of lysine methylation*. EMBO Rep, 2006. **7**(4): p. 397-403.
119. Stirnimann, C.U., et al., *WD40 proteins propel cellular networks*. Trends Biochem Sci, 2010. **35**(10): p. 565-74.
120. Suganuma, T., S.G. Pattenden, and J.L. Workman, *Diverse functions of WD40 repeat proteins in histone recognition*. Genes Dev, 2008. **22**(10): p. 1265-8.
121. Fujisawa, T. and P. Filippakopoulos, *Functions of bromodomain-containing proteins and their roles in homeostasis and cancer*. Nat Rev Mol Cell Biol, 2017. **18**(4): p. 246-262.

BIBLIOGRAPHY

122. Winston, F. and C.D. Allis, *The bromodomain: a chromatin-targeting module?* Nat Struct Biol, 1999. **6**(7): p. 601-4.
123. Qin, S. and J. Min, *Structure and function of the nucleosome-binding PWWP domain.* Trends Biochem Sci, 2014. **39**(11): p. 536-47.
124. Ciabrelli, F. and G. Cavalli, *Chromatin-driven behavior of topologically associating domains.* J Mol Biol, 2015. **427**(3): p. 608-25.
125. Felsenfeld, G. and M. Groudine, *Controlling the double helix.* Nature, 2003. **421**(6921): p. 448-53.
126. Kolasinska-Zwierz, P., et al., *Differential chromatin marking of introns and expressed exons by H3K36me3.* Nat Genet, 2009. **41**(3): p. 376-81.
127. Wagner, E.J. and P.B. Carpenter, *Understanding the language of Lys36 methylation at histone H3.* Nat Rev Mol Cell Biol, 2012. **13**(2): p. 115-26.
128. Zhu, L., et al., *ASH1L Links Histone H3 Lysine 36 Dimethylation to MLL Leukemia.* Cancer Discovery, 2016. **6**(7): p. 770-783.
129. Edmunds, J.W., L.C. Mahadevan, and A.L. Clayton, *Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation.* Embo Journal, 2008. **27**(2): p. 406-420.
130. Pfister, S.X., et al., *SETD2-Dependent Histone H3K36 Trimethylation Is Required for Homologous Recombination Repair and Genome Stability.* Cell Reports, 2014. **7**(6): p. 2006-2018.
131. Venkatesh, S. and J.L. Workman, *Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development.* Wiley Interdiscip Rev Dev Biol, 2013. **2**(5): p. 685-700.
132. Jha, D.K., et al., *SET-ting the stage for DNA repair.* Nat Struct Mol Biol, 2014. **21**(8): p. 655-7.
133. Li, F., et al., *The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha.* Cell, 2013. **153**(3): p. 590-600.
134. Pradeepa, M.M., et al., *Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing.* PLoS Genet, 2012. **8**(5): p. e1002717.

BIBLIOGRAPHY

135. Li, L. and Y.S. Wang, *Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair*. Journal of Biological Chemistry, 2017. **292**(28): p. 11951-11959.
136. Csizmok, V. and J.D. Forman-Kay, *Complex regulatory mechanisms mediated by the interplay of multiple post-translational modifications*. Curr Opin Struct Biol, 2017. **48**: p. 58-67.
137. Sims, R.J., 3rd and D. Reinberg, *Processing the H3K36me3 signature*. Nat Genet, 2009. **41**(3): p. 270-1.
138. Talbert, P.B., et al., *A unified phylogeny-based nomenclature for histone variants*. Epigenetics Chromatin, 2012. **5**: p. 7.
139. Talbert, P.B. and S. Henikoff, *Histone variants--ancient wrap artists of the epigenome*. Nat Rev Mol Cell Biol, 2010. **11**(4): p. 264-75.
140. Sancho, M., et al., *Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth*. PLoS Genet, 2008. **4**(10): p. e1000227.
141. Buschbeck, M. and S.B. Hake, *Variants of core histones and their roles in cell fate decisions, development and cancer*. Nature Reviews Molecular Cell Biology, 2017. **18**(5): p. 299-314.
142. Marzluff, W.F., E.J. Wagner, and R.J. Duronio, *Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail*. Nat Rev Genet, 2008. **9**(11): p. 843-54.
143. Cakmakci, N.G., et al., *SLIP1, a factor required for activation of histone mRNA translation by the stem-loop binding protein*. Molecular and Cellular Biology, 2008. **28**(3): p. 1182-1194.
144. Marzluff, W.F., et al., *The human and mouse replication-dependent histone genes*. Genomics, 2002. **80**(5): p. 487-98.
145. Mattioli, F., S. D'Arcy, and K. Luger, *The right place at the right time: chaperoning core histone variants*. Embo Reports, 2015. **16**(11): p. 1454-1466.
146. Burgess, R.J. and Z.G. Zhang, *Histone chaperones in nucleosome assembly and human disease*. Nature Structural & Molecular Biology, 2013. **20**(1): p. 14-22.
147. Cook, A.J., et al., *A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3-H4 in the histone supply chain*. Mol Cell, 2011. **44**(6): p. 918-27.

BIBLIOGRAPHY

148. Campos, E.I., et al., *The program for processing newly synthesized histones H3.1 and H4*. Nature Structural & Molecular Biology, 2010. **17**(11): p. 1343-U201.
149. Elsasser, S.J., *A common structural theme in histone chaperones mimics interhistone contacts*. Trends in Biochemical Sciences, 2013. **38**(7): p. 333-336.
150. Nakatani, Y., et al., *Two distinct nucleosome assembly pathways: Dependent or independent of DNA synthesis promoted by histone H3.1 and H3.3 complexes*. Cold Spring Harbor Symposia on Quantitative Biology, 2004. **69**: p. 273-280.
151. Narlikar, G.J., R. Sundaramoorthy, and T. Owen-Hughes, *Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes*. Cell, 2013. **154**(3): p. 490-503.
152. Bustin, M., F. Catez, and J.H. Lim, *The dynamics of histone H1 function in chromatin*. Mol Cell, 2005. **17**(5): p. 617-20.
153. Starkova, T.Y., et al., *Post-translational modifications of linker histone H1 variants in mammals*. Phys Biol, 2017. **14**(1): p. 016005.
154. Ponte, I., J.M. Vidal-Taboada, and P. Suau, *Evolution of the vertebrate H1 histone class: evidence for the functional differentiation of the subtypes*. Mol Biol Evol, 1998. **15**(6): p. 702-8.
155. Talasz, H., et al., *In vitro binding of H1 histone subtypes to nucleosomal organized mouse mammary tumor virus long terminal repeat promotor*. J Biol Chem, 1998. **273**(48): p. 32236-43.
156. Th'ng, J.P., et al., *H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain*. J Biol Chem, 2005. **280**(30): p. 27809-14.
157. Izzo, A., et al., *The genomic landscape of the somatic linker histone subtypes H1.1 to H1.5 in human cells*. Cell Rep, 2013. **3**(6): p. 2142-54.
158. Bernstein, E. and S.B. Hake, *The nucleosome: a little variation goes a long way*. Biochem Cell Biol, 2006. **84**(4): p. 505-17.
159. Santoro, S.W. and C. Dulac, *The activity-dependent histone variant H2BE modulates the life span of olfactory neurons*. Elife, 2012. **1**: p. e00070.
160. Montellier, E., et al., *Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B*. Genes Dev, 2013. **27**(15): p. 1680-92.

BIBLIOGRAPHY

161. Boulard, M., et al., *The NH2 tail of the novel histone variant H2BFWT exhibits properties distinct from conventional H2B with respect to the assembly of mitotic chromosomes*. Mol Cell Biol, 2006. **26**(4): p. 1518-26.
162. Gineitis, A.A., et al., *Human sperm telomere-binding complex involves histone H2B and secures telomere membrane attachment*. Journal of Cell Biology, 2000. **151**(7): p. 1591-1597.
163. Kwak, H.G., T. Suzuki, and N. Dohmae, *Global mapping of post-translational modifications on histone H3 variants in mouse testes*. Biochem Biophys Rep, 2017. **11**: p. 1-8.
164. Maze, I., et al., *Every amino acid matters: essential contributions of histone variants to mammalian development and disease*. Nat Rev Genet, 2014. **15**(4): p. 259-71.
165. Taguchi, H., et al., *Crystal Structure and Characterization of Novel Human Histone H3 Variants, H3.6, H3.7, and H3.8*. Biochemistry, 2017. **56**(16): p. 2184-2196.
166. Tachiwana, H., et al., *Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(23): p. 10454-10459.
167. Albig, W., et al., *A solitary human H3 histone gene on chromosome 1*. Hum Genet, 1996. **97**(4): p. 486-91.
168. Ederveen, T.H., I.K. Mandemaker, and C. Logie, *The human histone H3 complement anno 2011*. Biochim Biophys Acta, 2011. **1809**(10): p. 577-86.
169. Eitoku, M., et al., *Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly*. Cell Mol Life Sci, 2008. **65**(3): p. 414-44.
170. Tagami, H., et al., *Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis*. Cell, 2004. **116**(1): p. 51-61.
171. Lewis, P.W., et al., *Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres*. Proc Natl Acad Sci U S A, 2010. **107**(32): p. 14075-80.
172. Hake, S.B., et al., *Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes*. Proc Natl Acad Sci U S A, 2005. **102**(18): p. 6344-9.

BIBLIOGRAPHY

173. Ueda, J., et al., *Testis-Specific Histone Variant H3t Gene Is Essential for Entry into Spermatogenesis*. Cell Rep, 2017. **18**(3): p. 593-600.
174. Shiraishi, K., et al., *Roles of histone H3.5 in human spermatogenesis and spermatogenic disorders*. Andrology, 2018. **6**(1): p. 158-165.
175. Wiedemann, S.M., et al., *Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y*. J Cell Biol, 2010. **190**(5): p. 777-91.
176. Black, B.E. and D.W. Cleveland, *Epigenetic centromere propagation and the nature of CENP-a nucleosomes*. Cell, 2011. **144**(4): p. 471-9.
177. Shrestha, R.L., et al., *Mislocalization of centromeric histone H3 variant CENP-A contributes to chromosomal instability (CIN) in human cells*. Oncotarget, 2017. **8**(29): p. 46781-46800.
178. Franklin, S.G. and A. Zweidler, *Non-allelic variants of histones 2a, 2b and 3 in mammals*. Nature, 1977. **266**(5599): p. 273-5.
179. West, M.H. and W.M. Bonner, *Histone 2A, a heteromorphous family of eight protein species*. Biochemistry, 1980. **19**(14): p. 3238-45.
180. Rogakou, E.P., et al., *DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139*. J Biol Chem, 1998. **273**(10): p. 5858-68.
181. Vardabasso, C., et al., *Histone variants: emerging players in cancer biology*. Cell Mol Life Sci, 2014. **71**(3): p. 379-404.
182. Fernandez-Capetillo, O., et al., *H2AX: the histone guardian of the genome*. DNA Repair (Amst), 2004. **3**(8-9): p. 959-67.
183. Rogakou, E.P., et al., *Megabase chromatin domains involved in DNA double-strand breaks in vivo*. Journal of Cell Biology, 1999. **146**(5): p. 905-915.
184. Tiwari, M., S. Parvez, and P.K. Agrawala, *Role of some epigenetic factors in DNA damage response pathway*. AIMS Genetics, 2017. **4**(1): p. 69-83.
185. Palla, V.V., et al., *gamma-H2AX: Can it be established as a classical cancer prognostic factor?* Tumour Biol, 2017. **39**(3): p. 1010428317695931.
186. Eirin-Lopez, J.M., T. Ishibashi, and J. Ausio, *H2A.Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way*. FASEB J, 2008. **22**(1): p. 316-26.

BIBLIOGRAPHY

187. Ishibashi, T., et al., *H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis*. Nucleic Acids Research, 2010. **38**(6): p. 1780-1789.
188. Tolstorukov, M.Y., et al., *Histone Variant H2A.Bbd Is Associated with Active Transcription and mRNA Processing in Human Cells*. Molecular Cell, 2012. **47**(4): p. 596-607.
189. Chadwick, B.P. and H.F. Willard, *A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome*. Journal of Cell Biology, 2001. **152**(2): p. 375-384.
190. Bonisch, C., et al., *H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization*. Nucleic Acids Res, 2012. **40**(13): p. 5951-64.
191. Doyen, C.M., et al., *Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome*. EMBO J, 2006. **25**(18): p. 4234-44.
192. Okuwaki, M., et al., *Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein I*. Mol Cell Biol, 2005. **25**(23): p. 10639-51.
193. Pehrson, J.R. and V.A. Fried, *MacroH2A, a core histone containing a large nonhistone region*. Science, 1992. **257**(5075): p. 1398-400.
194. Rasmussen, T.P., et al., *Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing*. Nucleic Acids Res, 1999. **27**(18): p. 3685-9.
195. Han, W.D., X.L. Li, and X.B. Fu, *The macro domain protein family: Structure, functions, and their potential therapeutic implications*. Mutation Research-Reviews in Mutation Research, 2011. **727**(3): p. 86-103.
196. Kustatscher, G., et al., *Splicing regulates NAD metabolite binding to histone macroH2A*. Nature Structural & Molecular Biology, 2005. **12**(7): p. 624-625.
197. Timinszky, G., et al., *A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation*. Nature Structural & Molecular Biology, 2009. **16**(9): p. 923-U41.
198. Douet, J., et al., *MacroH2A histone variants maintain nuclear organization and heterochromatin architecture*. J Cell Sci, 2017. **130**(9): p. 1570-1582.
199. Thatcher, T.H. and M.A. Gorovsky, *Phylogenetic analysis of the core histones H2A, H2B, H3, and H4*. Nucleic Acids Res, 1994. **22**(2): p. 174-9.

BIBLIOGRAPHY

200. Zlatanova, J. and A. Thakar, *H2A.Z: view from the top*. Structure, 2008. **16**(2): p. 166-79.
201. Iouzalén, N., J. Moreau, and M. Mechali, *H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A*. Nucleic Acids Res, 1996. **24**(20): p. 3947-52.
202. Faast, R., et al., *Histone variant H2A.Z is required for early mammalian development*. Curr Biol, 2001. **11**(15): p. 1183-7.
203. Liu, X., B. Li, and GorovskyMa, *Essential and nonessential histone H2A variants in Tetrahymena thermophila*. Mol Cell Biol, 1996. **16**(8): p. 4305-11.
204. van Daal, A. and S.C. Elgin, *A histone variant, H2AvD, is essential in Drosophila melanogaster*. Mol Biol Cell, 1992. **3**(6): p. 593-602.
205. Carr, A.M., et al., *Analysis of a histone H2A variant from fission yeast: evidence for a role in chromosome stability*. Mol Gen Genet, 1994. **245**(5): p. 628-35.
206. Jackson, J.D. and M.A. Gorovsky, *Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants*. Nucleic Acids Res, 2000. **28**(19): p. 3811-6.
207. Dryhurst, D., et al., *Characterization of the histone H2A.Z-1 and H2A.Z-2 isoforms in vertebrates*. BMC Biol, 2009. **7**: p. 86.
208. Clarkson, M.J., et al., *Regions of variant histone His2AvD required for Drosophila development*. Nature, 1999. **399**(6737): p. 694-7.
209. Wu, W.H., et al., *Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange*. Nat Struct Mol Biol, 2005. **12**(12): p. 1064-71.
210. Jensen, K., et al., *Histone H2A.Z acid patch residues required for deposition and function*. Mol Genet Genomics, 2011. **285**(4): p. 287-96.
211. Wrattling, D., et al., *A Conserved Function for the H2A.Z C Terminus*. Journal of Biological Chemistry, 2012. **287**(23): p. 19148-19157.
212. Suto, R.K., et al., *Crystal structure of a nucleosome core particle containing the variant histone H2A.Z*. Nat Struct Biol, 2000. **7**(12): p. 1121-4.
213. Chakravarthy, S., et al., *Structural characterization of histone H2A variants*. Cold Spring Harb Symp Quant Biol, 2004. **69**: p. 227-34.

BIBLIOGRAPHY

214. Luk, E., et al., *Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome*. Cell, 2010. **143**(5): p. 725-36.
215. Viens, A., et al., *Analysis of human histone H2AZ deposition in vivo argues against its direct role in epigenetic templating mechanisms*. Mol Cell Biol, 2006. **26**(14): p. 5325-35.
216. Weber, C.M., J.G. Henikoff, and S. Henikoff, *H2A.Z nucleosomes enriched over active genes are homotypic*. Nat Struct Mol Biol, 2010. **17**(12): p. 1500-7.
217. Nekrasov, M., et al., *Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics*. Nat Struct Mol Biol, 2012. **19**(11): p. 1076-83.
218. Jin, C. and G. Felsenfeld, *Nucleosome stability mediated by histone variants H3.3 and H2A.Z*. Genes Dev, 2007. **21**(12): p. 1519-29.
219. Henikoff, S., *Labile H3.3+H2A.Z nucleosomes mark 'nucleosome-free regions'*. Nat Genet, 2009. **41**(8): p. 865-6.
220. Chen, P., et al., *H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin*. Genes Dev, 2013. **27**(19): p. 2109-24.
221. Fan, J.Y., et al., *H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding*. Mol Cell, 2004. **16**(4): p. 655-61.
222. Fan, J.Y., et al., *The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states*. Nat Struct Biol, 2002. **9**(3): p. 172-6.
223. Greaves, I.K., et al., *H2A.Z contributes to the unique 3D structure of the centromere*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(2): p. 525-530.
224. Sevilla, A. and O. Binda, *Post-translational modifications of the histone variant H2AZ*. Stem Cell Res, 2014. **12**(1): p. 289-95.
225. Thambirajah, A.A., et al., *New developments in post-translational modifications and functions of histone H2A variants*. Biochem Cell Biol, 2009. **87**(1): p. 7-17.
226. Valdes-Mora, F., et al., *Acetylated histone variant H2A.Z is involved in the activation of neo-enhancers in prostate cancer*. Nat Commun, 2017. **8**(1): p. 1346.
227. Bruce, K., et al., *The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken*. Nucleic Acids Res, 2005. **33**(17): p. 5633-9.

BIBLIOGRAPHY

228. Valdes-Mora, F., et al., *Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer*. *Genome Res*, 2012. **22**(2): p. 307-21.
229. Ishibashi, T., et al., *Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome*. *Biochemistry*, 2009. **48**(22): p. 5007-17.
230. Sarcinella, E., et al., *Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin*. *Mol Cell Biol*, 2007. **27**(18): p. 6457-68.
231. Kalocsay, M., N.J. Hiller, and S. Jentsch, *Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break*. *Mol Cell*, 2009. **33**(3): p. 335-43.
232. Binda, O., et al., *SETD6 monomethylates H2AZ on lysine 7 and is required for the maintenance of embryonic stem cell self-renewal*. *Epigenetics*, 2013. **8**(2): p. 177-83.
233. Billon, P. and J. Cote, *Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance*. *Biochim Biophys Acta*, 2013. **1819**(3-4): p. 290-302.
234. Kobor, M.S., et al., *A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin*. *PLoS Biol*, 2004. **2**(5): p. E131.
235. Krogan, N.J., et al., *A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1*. *Mol Cell*, 2003. **12**(6): p. 1565-76.
236. Mizuguchi, G., et al., *ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex*. *Science*, 2004. **303**(5656): p. 343-8.
237. Neves, L.T., et al., *The histone variant H2A.Z promotes efficient cotranscriptional splicing in *S. cerevisiae**. *Genes Dev*, 2017. **31**(7): p. 702-717.
238. Gevry, N., et al., *p21 transcription is regulated by differential localization of histone H2A.Z*. *Genes Dev*, 2007. **21**(15): p. 1869-81.
239. Ruhl, D.D., et al., *Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes*. *Biochemistry*, 2006. **45**(17): p. 5671-7.
240. Papamichos-Chronakis, M., et al., *Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity*. *Cell*, 2011. **144**(2): p. 200-13.

BIBLIOGRAPHY

241. Conaway, R.C. and J.W. Conaway, *The INO80 chromatin remodeling complex in transcription, replication and repair*. Trends Biochem Sci, 2009. **34**(2): p. 71-7.
242. Mao, Z., et al., *Anp32e, a higher eukaryotic histone chaperone directs preferential recognition for H2A.Z*. Cell Res, 2014. **24**(4): p. 389-99.
243. Hardy, S. and F. Robert, *Random deposition of histone variants: A cellular mistake or a novel regulatory mechanism?* Epigenetics, 2010. **5**(5): p. 368-72.
244. Domaschenz, R., et al., *The Histone Variant H2A.Z Is a Master Regulator of the Epithelial-Mesenchymal Transition*. Cell Rep, 2017. **21**(4): p. 943-952.
245. Marques, M., et al., *Reconciling the positive and negative roles of histone H2A.Z in gene transcription*. Epigenetics, 2010. **5**(4): p. 267-72.
246. Zovkic, I.B., et al., *Histone H2A.Z subunit exchange controls consolidation of recent and remote memory*. Nature, 2014. **515**(7528): p. 582-6.
247. Allis, C.D., et al., *hvl1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes*. J Biol Chem, 1986. **261**(4): p. 1941-8.
248. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-37.
249. Schones, D.E., et al., *Dynamic regulation of nucleosome positioning in the human genome*. Cell, 2008. **132**(5): p. 887-98.
250. Schauer, T., et al., *CAST-ChIP maps cell-type-specific chromatin states in the Drosophila central nervous system*. Cell Rep, 2013. **5**(1): p. 271-82.
251. Zilberman, D., et al., *Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks*. Nature, 2008. **456**(7218): p. 125-9.
252. Albert, I., et al., *Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome*. Nature, 2007. **446**(7135): p. 572-6.
253. Raisner, R.M., et al., *Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin*. Cell, 2005. **123**(2): p. 233-48.
254. Conerly, M.L., et al., *Changes in H2A.Z occupancy and DNA methylation during B-cell lymphomagenesis*. Genome Res, 2010. **20**(10): p. 1383-90.

BIBLIOGRAPHY

255. Subramanian, V., P.A. Fields, and L.A. Boyer, *H2A.Z: a molecular rheostat for transcriptional control*. F1000Prime Rep, 2015. **7**: p. 01.
256. Nock, A., et al., *Mediator-regulated transcription through the +1 nucleosome*. Mol Cell, 2012. **48**(6): p. 837-48.
257. Rhee, H.S. and B.F. Pugh, *Genome-wide structure and organization of eukaryotic pre-initiation complexes*. Nature, 2012. **483**(7389): p. 295-301.
258. Weber, C.M., S. Ramachandran, and S. Henikoff, *Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase*. Mol Cell, 2014. **53**(5): p. 819-30.
259. Morillo-Huesca, M., et al., *The SWR1 histone replacement complex causes genetic instability and genome-wide transcription misregulation in the absence of H2A.Z*. PLoS One, 2010. **5**(8): p. e12143.
260. Xu, Y., et al., *Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair*. Mol Cell, 2012. **48**(5): p. 723-33.
261. Gursoy-Yuzugullu, O., M.K. Ayrapetov, and B.D. Price, *Histone chaperone Anp32e removes H2A.Z from DNA double-strand breaks and promotes nucleosome reorganization and DNA repair (vol 112, pg 7507, 2015)*. Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(28): p. E3750-E3750.
262. Cosma, M.P., *Ordered recruitment: gene-specific mechanism of transcription activation*. Mol Cell, 2002. **10**(2): p. 227-36.
263. Dhillon, N., et al., *H2A.Z functions to regulate progression through the cell cycle*. Mol Cell Biol, 2006. **26**(2): p. 489-501.
264. Dalvai, M., et al., *H2A.Z-dependent crosstalk between enhancer and promoter regulates cyclin D1 expression*. Oncogene, 2013. **32**(36): p. 4243-51.
265. Dunican, D.S., et al., *Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization*. Oncogene, 2002. **21**(20): p. 3253-7.
266. Hua, S., et al., *Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression*. Mol Syst Biol, 2008. **4**: p. 188.

BIBLIOGRAPHY

267. Rhodes, D.R., et al., *Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression*. Proc Natl Acad Sci U S A, 2004. **101**(25): p. 9309-14.
268. Zucchi, I., et al., *Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis*. Proc Natl Acad Sci U S A, 2004. **101**(52): p. 18147-52.
269. Kapoor, A., et al., *The histone variant macroH2A suppresses melanoma progression through regulation of CDK8*. Nature, 2010. **468**(7327): p. 1105-9.
270. Vardabasso, C., et al., *Histone Variant H2A.Z.2 Mediates Proliferation and Drug Sensitivity of Malignant Melanoma*. Mol Cell, 2015. **59**(1): p. 75-88.
271. Denis, G.V., et al., *Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines*. J Proteome Res, 2006. **5**(3): p. 502-11.
272. Denis, G.V., et al., *RING3 kinase transactivates promoters of cell cycle regulatory genes through E2F*. Cell Growth Differ, 2000. **11**(8): p. 417-24.
273. Sinha, A., D.V. Faller, and G.V. Denis, *Bromodomain analysis of Brd2-dependent transcriptional activation of cyclin A*. Biochem J, 2005. **387**(Pt 1): p. 257-69.
274. Lito, P., N. Rosen, and D.B. Solit, *Tumor adaptation and resistance to RAF inhibitors*. Nat Med, 2013. **19**(11): p. 1401-9.
275. Draker, R., et al., *A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation*. PLoS Genet, 2012. **8**(11): p. e1003047.
276. Punzeler, S., et al., *Multivalent binding of PWWP2A to H2A.Z regulates mitosis and neural crest differentiation*. EMBO J, 2017. **36**(15): p. 2263-2279.
277. Thirman, M.J., et al., *Rearrangement of the MLL gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations*. N Engl J Med, 1993. **329**(13): p. 909-14.
278. Tkachuk, D.C., S. Kohler, and M.L. Cleary, *Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias*. Cell, 1992. **71**(4): p. 691-700.

BIBLIOGRAPHY

279. Wiesmann, U.N., S. DiDonato, and N.N. Herschkowitz, *Effect of chloroquine on cultured fibroblasts: release of lysosomal hydrolases and inhibition of their uptake*. *Biochem Biophys Res Commun*, 1975. **66**(4): p. 1338-43.
280. Ziemin-van der Poel, S., et al., *Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias*. *Proc Natl Acad Sci U S A*, 1991. **88**(23): p. 10735-9.
281. Popovic, R. and N.J. Zeleznik-Le, *MLL: how complex does it get?* *J Cell Biochem*, 2005. **95**(2): p. 234-42.
282. Hu, G., et al., *H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation*. *Cell Stem Cell*, 2013. **12**(2): p. 180-92.
283. Horton, J.R., et al., *Structural basis for human PHF2 Jumonji domain interaction with metal ions*. *J Mol Biol*, 2011. **406**(1): p. 1-8.
284. Wen, H., et al., *Recognition of histone H3K4 trimethylation by the plant homeodomain of PHF2 modulates histone demethylation*. *J Biol Chem*, 2010. **285**(13): p. 9322-6.
285. Tanaka, Y., et al., *JmjC enzyme KDM2A is a regulator of rRNA transcription in response to starvation*. *EMBO J*, 2010. **29**(9): p. 1510-22.
286. Tsai, C.L., Y. Shi, and J.A. Tainer, *How substrate specificity is imposed on a histone demethylase--lessons from KDM2A*. *Genes Dev*, 2014. **28**(16): p. 1735-8.
287. Cheng, Z., et al., *A molecular threading mechanism underlies Jumonji lysine demethylase KDM2A regulation of methylated H3K36*. *Genes Dev*, 2014. **28**(16): p. 1758-71.
288. Stec, I., et al., *WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphia gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma*. *Hum Mol Genet*, 1998. **7**(7): p. 1071-82.
289. Qiu, C., et al., *The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds*. *Nat Struct Biol*, 2002. **9**(3): p. 217-24.
290. Sue, S.C., et al., *Solution structure and heparin interaction of human hepatoma-derived growth factor*. *J Mol Biol*, 2004. **343**(5): p. 1365-77.
291. Hung, Y.L., et al., *The First Residue of the PWWP Motif Modulates HATH Domain Binding, Stability, and Protein-Protein Interaction*. *Biochemistry*, 2015. **54**(26): p. 4063-74.

BIBLIOGRAPHY

292. Maurer-Stroh, S., et al., *The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains*. Trends Biochem Sci, 2003. **28**(2): p. 69-74.
293. Rona, G.B., E.C.A. Eleutherio, and A.S. Pinheiro, *PWWP domains and their modes of sensing DNA and histone methylated lysines*. Biophys Rev, 2016. **8**(1): p. 63-74.
294. Eidahl, J.O., et al., *Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes*. Nucleic Acids Res, 2013. **41**(6): p. 3924-36.
295. van Nuland, R., et al., *Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain*. Epigenetics Chromatin, 2013. **6**(1): p. 12.
296. Vezzoli, A., et al., *Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1*. Nat Struct Mol Biol, 2010. **17**(5): p. 617-9.
297. Wen, H., et al., *ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression*. Nature, 2014. **508**(7495): p. 263-8.
298. Wu, H., et al., *Structural and histone binding ability characterizations of human PWWP domains*. PLoS One, 2011. **6**(6): p. e18919.
299. Nameki, N., et al., *Solution structure of the PWWP domain of the hepatoma-derived growth factor family*. Protein Sci, 2005. **14**(3): p. 756-64.
300. Qiu, Y., et al., *Solution structure of the Pdp1 PWWP domain reveals its unique binding sites for methylated H4K20 and DNA*. Biochem J, 2012. **442**(3): p. 527-38.
301. Slater, L.M., M.D. Allen, and M. Bycroft, *Structural variation in PWWP domains*. J Mol Biol, 2003. **330**(3): p. 571-6.
302. Shikauchi, Y., et al., *SALL3 interacts with DNMT3A and shows the ability to inhibit CpG island methylation in hepatocellular carcinoma*. Mol Cell Biol, 2009. **29**(7): p. 1944-58.
303. Park, J., et al., *DNA methyltransferase 3B mutant in ICF syndrome interacts non-covalently with SUMO-1*. J Mol Med (Berl), 2008. **86**(11): p. 1269-77.
304. Botbol, Y., et al., *Chromatinized templates reveal the requirement for the LEDGF/p75 PWWP domain during HIV-1 integration in vitro*. Nucleic Acids Res, 2008. **36**(4): p. 1237-46.
305. Shun, M.C., et al., *Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity*. J Virol, 2008. **82**(23): p. 11555-67.

BIBLIOGRAPHY

306. Laguri, C., et al., *Human mismatch repair protein MSH6 contains a PWWP domain that targets double stranded DNA*. *Biochemistry*, 2008. **47**(23): p. 6199-207.
307. Wang, Y., et al., *Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein*. *Mol Cell*, 2009. **33**(4): p. 428-37.
308. Vermeulen, M., et al., *Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers*. *Cell*, 2010. **142**(6): p. 967-80.
309. Wang, J., et al., *Crystal structure of human BS69 Bromo-ZnF-PWWP reveals its role in H3K36me3 nucleosome binding*. *Cell Res*, 2014. **24**(7): p. 890-3.
310. MacLean, B., et al., *Skyline: an open source document editor for creating and analyzing targeted proteomics experiments*. *Bioinformatics*, 2010. **26**(7): p. 966-8.
311. Tyanova, S., et al., *The Perseus computational platform for comprehensive analysis of (prote)omics data*. *Nat Methods*, 2016. **13**(9): p. 731-40.
312. Lee, K.M. and G. Narlikar, *Assembly of nucleosomal templates by salt dialysis*. *Curr Protoc Mol Biol*, 2001. **Chapter 21**: p. Unit 21 6.
313. Schnitzler, G.R., *Isolation of histones and nucleosome cores from mammalian cells*. *Curr Protoc Mol Biol*, 2001. **Chapter 21**: p. Unit 21 5.
314. Lowary, P.T. and J. Widom, *New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning*. *J Mol Biol*, 1998. **276**(1): p. 19-42.
315. Kusakabe, M., et al., *Genetic complementation analysis showed distinct contributions of the N-terminal tail of H2A.Z to epigenetic regulations*. *Genes to Cells*, 2016. **21**(2): p. 122-135.
316. Goto, Y., et al., *Pds5 Regulates Sister-Chromatid Cohesion and Chromosome Bi-orientation through a Conserved Protein Interaction Module*. *Curr Biol*, 2017. **27**(7): p. 1005-1012.
317. Huang, X.Q. and W. Miller, *A Time-Efficient, Linear-Space Local Similarity Algorithm*. *Advances in Applied Mathematics*, 1991. **12**(3): p. 337-357.
318. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data*. *Cancer Discov*, 2012. **2**(5): p. 401-4.
319. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal*. *Sci Signal*, 2013. **6**(269): p. p11.

BIBLIOGRAPHY

320. Vardabasso, C., S.B. Hake, and E. Bernstein, *Histone variant H2A.Z.2: A novel driver of melanoma progression*. *Mol Cell Oncol*, 2016. **3**(2): p. e1073417.
321. Iwasaki, W., et al., *Contribution of histone N-terminal tails to the structure and stability of nucleosomes*. *FEBS Open Bio*, 2013. **3**: p. 363-9.
322. Kopaciewicz, W., et al., *Retention Model for High-Performance Ion-Exchange Chromatography*. *Journal of Chromatography*, 1983. **266**(Aug): p. 3-21.
323. Audain, E., et al., *Accurate estimation of isoelectric point of protein and peptide based on amino acid sequences*. *Bioinformatics*, 2016. **32**(6): p. 821-7.
324. Zaret, K.S. and J.S. Carroll, *Pioneer transcription factors: establishing competence for gene expression*. *Genes & Development*, 2011. **25**(21): p. 2227-2241.
325. Chen, T., N. Tsujimoto, and E. Li, *The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin*. *Mol Cell Biol*, 2004. **24**(20): p. 9048-58.
326. Yang, J. and A.D. Everett, *Hepatoma-derived growth factor binds DNA through the N-terminal PWWP domain*. *BMC Mol Biol*, 2007. **8**: p. 101.
327. Lukasik, S.M., et al., *High resolution structure of the HDGF PWWP domain: a potential DNA binding domain*. *Protein Sci*, 2006. **15**(2): p. 314-23.
328. Wang, J.C., et al., *Crystal structure of human BS69 Bromo-ZnF-PWWP reveals its role in H3K36me3 nucleosome binding*. *Cell Research*, 2014. **24**(7): p. 890-893.
329. Baubec, T., et al., *Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation*. *Nature*, 2015. **520**(7546): p. 243-U278.
330. Dhayalan, A., et al., *The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation*. *Journal of Biological Chemistry*, 2010. **285**(34): p. 26114-26120.
331. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. *Cell*, 2007. **129**(4): p. 823-837.
332. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types*. *Nature*, 2011. **473**(7345): p. 43-U52.

BIBLIOGRAPHY

333. Nguyen, A.T. and Y. Zhang, *The diverse functions of Dot1 and H3K79 methylation*. Genes & Development, 2011. **25**(13): p. 1345-1358.
334. Steger, D.J., et al., *DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells*. Molecular and Cellular Biology, 2008. **28**(8): p. 2825-2839.
335. Mermoud, J.E., et al., *Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation*. J Cell Biol, 1999. **147**(7): p. 1399-408.
336. Rasmussen, T.P., et al., *Dynamic relocalization of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation*. J Cell Biol, 2000. **150**(5): p. 1189-98.
337. Mermoud, J.E., et al., *Centrosomal association of histone macroH2A1.2 in embryonic stem cells and somatic cells*. Exp Cell Res, 2001. **268**(2): p. 245-51.
338. Aymard, F., et al., *Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks*. Nat Struct Mol Biol, 2014. **21**(4): p. 366-74.
339. Pinto, D.M. and A. Flaus, *Structure and function of histone H2AX*. Subcell Biochem, 2010. **50**: p. 55-78.
340. Han, J.H., et al., *The folding and evolution of multidomain proteins*. Nat Rev Mol Cell Biol, 2007. **8**(4): p. 319-30.
341. Liang, X., et al., *Structural basis of H2A.Z recognition by SRCAP chromatin-remodeling subunit YL1*. Nat Struct Mol Biol, 2016. **23**(4): p. 317-23.
342. Ranjan, A., et al., *Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement*. Cell, 2013. **154**(6): p. 1232-45.
343. Wu, W.H., et al., *N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex*. J Biol Chem, 2009. **284**(10): p. 6200-7.
344. Gursoy-Yuzugullu, O., M.K. Ayrapetov, and B.D. Price, *Histone chaperone Anp32e removes H2A.Z from DNA double-strand breaks and promotes nucleosome reorganization and DNA repair*. Proc Natl Acad Sci U S A, 2015. **112**(24): p. 7507-12.
345. Hicke, L., *Protein regulation by monoubiquitin*. Nat Rev Mol Cell Biol, 2001. **2**(3): p. 195-201.

BIBLIOGRAPHY

346. Draker, R., E. Sarcinella, and P. Cheung, *USP10 deubiquitylates the histone variant H2A.Z and both are required for androgen receptor-mediated gene activation*. *Nucleic Acids Research*, 2011. **39**(9): p. 3529-3542.
347. Hunter, T., *The age of crosstalk: phosphorylation, ubiquitination, and beyond*. *Mol Cell*, 2007. **28**(5): p. 730-8.
348. Krawczak, M., et al., *Human gene mutation database - A biomedical information and research resource*. *Human Mutation*, 2000. **15**(1): p. 45-51.
349. Hong, S.Y., J.E. Oh, and K.H. Lee, *Effect of D-amino acid substitution on the stability, the secondary structure, and the activity of membrane-active peptide*. *Biochemical Pharmacology*, 1999. **58**(11): p. 1775-1780.
350. Ng, P.C. and S. Henikoff, *SIFT: predicting amino acid changes that affect protein function*. *Nucleic Acids Research*, 2003. **31**(13): p. 3812-3814.
351. Monteiro, F.L., et al., *Expression and functionality of histone H2A variants in cancer*. *Oncotarget*, 2014. **5**(11): p. 3428-43.
352. Gao, Y.F., et al., *Charged residues in the H-NS linker drive DNA binding and gene silencing in single cells*. *Proceedings of the National Academy of Sciences of the United States of America*, 2017. **114**(47): p. 12560-12565.
353. Wang, C., V.N. Uversky, and L. Kurgan, *Disordered nucleome: Abundance of intrinsic disorder in the DNA- and RNA-binding proteins in 1121 species from Eukaryota, Bacteria and Archaea*. *Proteomics*, 2016. **16**(10): p. 1486-1498.
354. Mutskov, V., et al., *Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding*. *Molecular and Cellular Biology*, 1998. **18**(11): p. 6293-6304.
355. Malaga, F., et al., *A genetic and physical study of the interdomain linker of E-Coli AraC protein-a trans-subunit communication pathway*. *Proteins-Structure Function and Bioinformatics*, 2016. **84**(4): p. 448-460.
356. Joshi, R., et al., *Functional specificity of a Hox protein mediated by the recognition of minor groove structure*. *Cell*, 2007. **131**(3): p. 530-543.

BIBLIOGRAPHY

357. Schwarz, P.M. and J.C. Hansen, *Formation and Stability of Higher-Order Chromatin Structures - Contributions of the Histone Octamer*. Journal of Biological Chemistry, 1994. **269**(23): p. 16284-16289.
358. Adams, C.C. and J.L. Workman, *Binding of Disparate Transcriptional Activators to Nucleosomal DNA Is Inherently Cooperative*. Molecular and Cellular Biology, 1995. **15**(3): p. 1405-1421.
359. Steger, D.J. and J.L. Workman, *Stable co-occupancy of transcription factors and histones at the HIV-1 enhancer*. Embo Journal, 1997. **16**(9): p. 2463-2472.
360. Singh, D.P., et al., *DNA binding domains and nuclear localization signal of LEDGF: contribution of two helix-turn-helix (HTH)-like domains and a stretch of 58 amino acids of the N-terminal to the trans-activation potential of LEDGF*. J Mol Biol, 2006. **355**(3): p. 379-94.
361. Purdy, M.M., C. Holz-Schietinger, and N.O. Reich, *Identification of a second DNA binding site in human DNA methyltransferase 3A by substrate inhibition and domain deletion*. Arch Biochem Biophys, 2010. **498**(1): p. 13-22.
362. Shin, H., et al., *Transcriptional regulation mediated by H2A.Z via ANP32e-dependent inhibition of protein phosphatase 2A*. Biochim Biophys Acta, 2018.
363. Sura, W., et al., *Dual Role of the Histone Variant H2A.Z in Transcriptional Regulation of Stress-Response Genes*. Plant Cell, 2017. **29**(4): p. 791-807.
364. Li, Y.Y. and H.T. Li, *Many keys to push: diversifying the 'readership' of plant homeodomain fingers*. Acta Biochimica Et Biophysica Sinica, 2012. **44**(1): p. 28-39.
365. Carrozza, M.J., et al., *Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription*. Cell, 2005. **123**(4): p. 581-592.
366. Keogh, M.C., et al., *Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex*. Cell, 2005. **123**(4): p. 593-605.
367. Yang, X.J. and E. Seto, *Lysine acetylation: codified crosstalk with other posttranslational modifications*. Mol Cell, 2008. **31**(4): p. 449-61.
368. Guo, R., et al., *BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing*. Mol Cell, 2014. **56**(2): p. 298-310.
369. Maltby, V.E., et al., *Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin*. Mol Cell Biol, 2012. **32**(17): p. 3479-85.

BIBLIOGRAPHY

370. Smolle, M., et al., *Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange*. Nat Struct Mol Biol, 2012. **19**(9): p. 884-92.
371. Gatchalian, J., et al., *An aromatic cage is required but not sufficient for binding of Tudor domains of the Polycomb-like protein family to H3K36me3*. Epigenetics, 2015. **10**(6): p. 467-473.
372. Yun, M.Y., et al., *Readers of histone modifications*. Cell Research, 2011. **21**(4): p. 564-578.
373. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription*. Cell, 2007. **128**(4): p. 707-719.
374. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications*. Nature, 2000. **403**(6765): p. 41-45.
375. Papait, R., et al., *Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(50): p. 20164-20169.
376. Xu, Y. and K.C. Chou, *Recent Progress in Predicting Posttranslational Modification Sites in Proteins*. Current Topics in Medicinal Chemistry, 2016. **16**(6): p. 591-603.
377. Xie, A.Y., A. Kwok, and R. Scully, *Role of mammalian Mre11 in classical and alternative nonhomologous end joining*. Nature Structural & Molecular Biology, 2009. **16**(8): p. 814-U31.
378. Stucki, M., et al., *MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks (vol 123, pg 1213, 2005)*. Cell, 2008. **133**(3): p. 549-549.
379. Yaffe, M.B. and A.E.H. Elia, *Phosphoserine/threonine-binding domains*. Current Opinion in Cell Biology, 2001. **13**(2): p. 131-138.
380. Korkuc, P. and D. Walther, *Towards understanding the crosstalk between protein post-translational modifications: Homo- and heterotypic PTM pair distances on protein surfaces are not random*. Proteins-Structure Function and Bioinformatics, 2017. **85**(1): p. 78-92.
381. Nimura, K., et al., *A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome*. Nature, 2009. **460**(7252): p. 287-91.
382. Schmidt, C.K. and S.P. Jackson, *On your mark, get SET(D2), go! H3K36me3 primes DNA mismatch repair*. Cell, 2013. **153**(3): p. 513-5.

BIBLIOGRAPHY

383. Li, X.Z., et al., *MOF and H4 K16 Acetylation Play Important Roles in DNA Damage Repair by Modulating Recruitment of DNA Damage Repair Protein Mdc1*. *Molecular and Cellular Biology*, 2010. **30**(22): p. 5335-5347.
384. Bick, G., et al., *Coordination of the recruitment of the FANCD2 and PALB2 Fanconi anemia proteins by an ubiquitin signaling network*. *Chromosoma*, 2017. **126**(3): p. 417-430.
385. Normanno, D., et al., *Mutational phospho-mimicry reveals a regulatory role for the XRCC4 and XLF C-terminal tails in modulating DNA bridging during classical non-homologous end joining*. *Elife*, 2017. **6**.
386. Ikegami, S., T. Taguchi, and M. Ohashi, *Aphidicolin Prevents Mitotic Cell-Division by Interfering with Activity of DNA Polymerase-Alpha*. *Nature*, 1978. **275**(5679): p. 458-460.

APPENDIX

APPENDIX

Table A.1 PRM inclusion list.

List of precursor masses for specified histone modifications and according processing parameters for PRM analysis.

Mass [m/z]	CS [z]	Polarity	Start [min]	End [min]	(N)CE	(N)CE type	Comment
380,72	2	Positive	20	35	30	NCE	H3_3_8_K4_p_light
387,73	2	Positive	30	40	30	NCE	H3_3_8_K4_me1_light
366,72	2	Positive	20	35	32	NCE	H3_3_8_K4_me2_light
373,73	2	Positive	20	35	32	NCE	H3_3_8_K4_me3_light
539,31	3	Positive	40	50	29	NCE	H3_27_40_K27K36K37_m1pp+H3_27_40_K27K36K37_pm1p_light
529,98	3	Positive	38	48	30	NCE	H3_27_40_K27K36K37_m2m1p+H3_27_40_K27K36K37_pm3p_light
534,65	3	Positive	38	48	30	NCE	H3_27_40_K27S28K36K37_me3me1p+H3_27_40_K27S28K36K37_me1me3p_light
525,31	3	Positive	38	44	30	NCE	H3_27-40_K27K36K37_pme2p+H3_27-40_K27K36K37_me2pp_light
520,65	3	Positive	35	40	32	NCE	H3_27-40_K27K36K37_me2me3p+H3_27-40_K27K36K37_me3me2p_light
525,32	3	Positive	35	50	32	NCE	H3_27-40_K27K36K37_me3me3p_light
543,99	3	Positive	40	50	30	NCE	H3_27-40_K27K36K37_me1me1p_light
544,65	3	Positive	40	50	29	NCE	H3.3_27_40_K27K36K37_m1pp+H3_27_40_K27K36K37_pm1p_light
535,31	3	Positive	38	48	30	NCE	H3.3_27_40_K27K36K37_m2m1p+H3_27_40_K27K36K37_pm3p_light
539,99	3	Positive	38	48	30	NCE	H3.3_27_40_K27S28K36K37_me3me1p+H3_27_40_K27S28K36K37_me1me3p_light
530,64	3	Positive	38	44	30	NCE	H3.3_27-40_K27K36K37_pme2p+H3_27-40_K27K36K37_me2pp_light
525,98	3	Positive	35	40	32	NCE	H3.3_27-40_K27K36K37_me2me3p+H3_27-40_K27K36K37_me3me2p_light
530,65	3	Positive	35	50	32	NCE	H3.3_27-40_K27K36K37_me3me3p_light
543,99	3	Positive	40	50	30	NCE	H3.3_27-40_K27K36K37_me1me1p_light
682,36	2	Positive	42	55	32	NCE	H3_73-83_K79_me2_light
689,37	2	Positive	40	50	32	NCE	H3_73-83_K79_me3_light
286,20	2	Positive	30	45	30	NCE	H4_20-23_K20_p_light
293,21	2	Positive	30	45	30	NCE	H4_20-23_K20_me1_light
272,20	2	Positive	18	30	32	NCE	H4_20-23_K20_me2_light
279,21	2	Positive	18	30	32	NCE	H4_20-23_K20_me3_light

APPENDIX

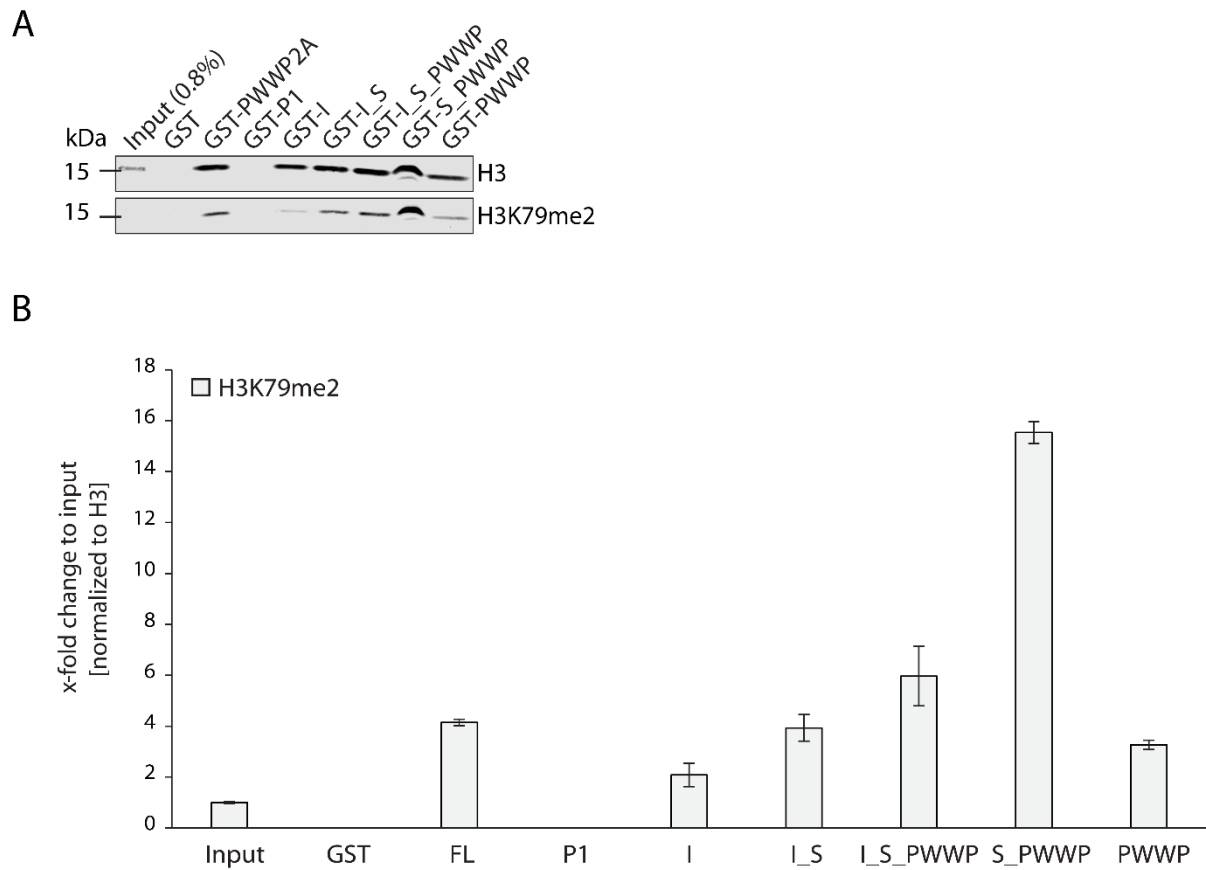


Figure A1: S_PWWP enriches H3K79me2-containing mononucleosomes. (A) Pull-downs (n=3) of recombinant PWWP2A full-length and truncations (P1, I, I_S, I_S_PWWP, S_PWWP, PWWP) with HK-derived mononucleosomes. GST served as negative control. Precipitated histone modifications and histones (H3) are detected in immunoblots using commercially available antibodies. Notice increased signal intensity in H3K79me2 binding for S_PWWP in comparison to remaining constructs. (B) Data quantification was done for three biological replicates for each PTM (n=3) by analyzing signal intensities using Image Studio Lite Ver 5.2 (LI-COR). Data shown are means and error bars depict the standard error of the mean (SEM).

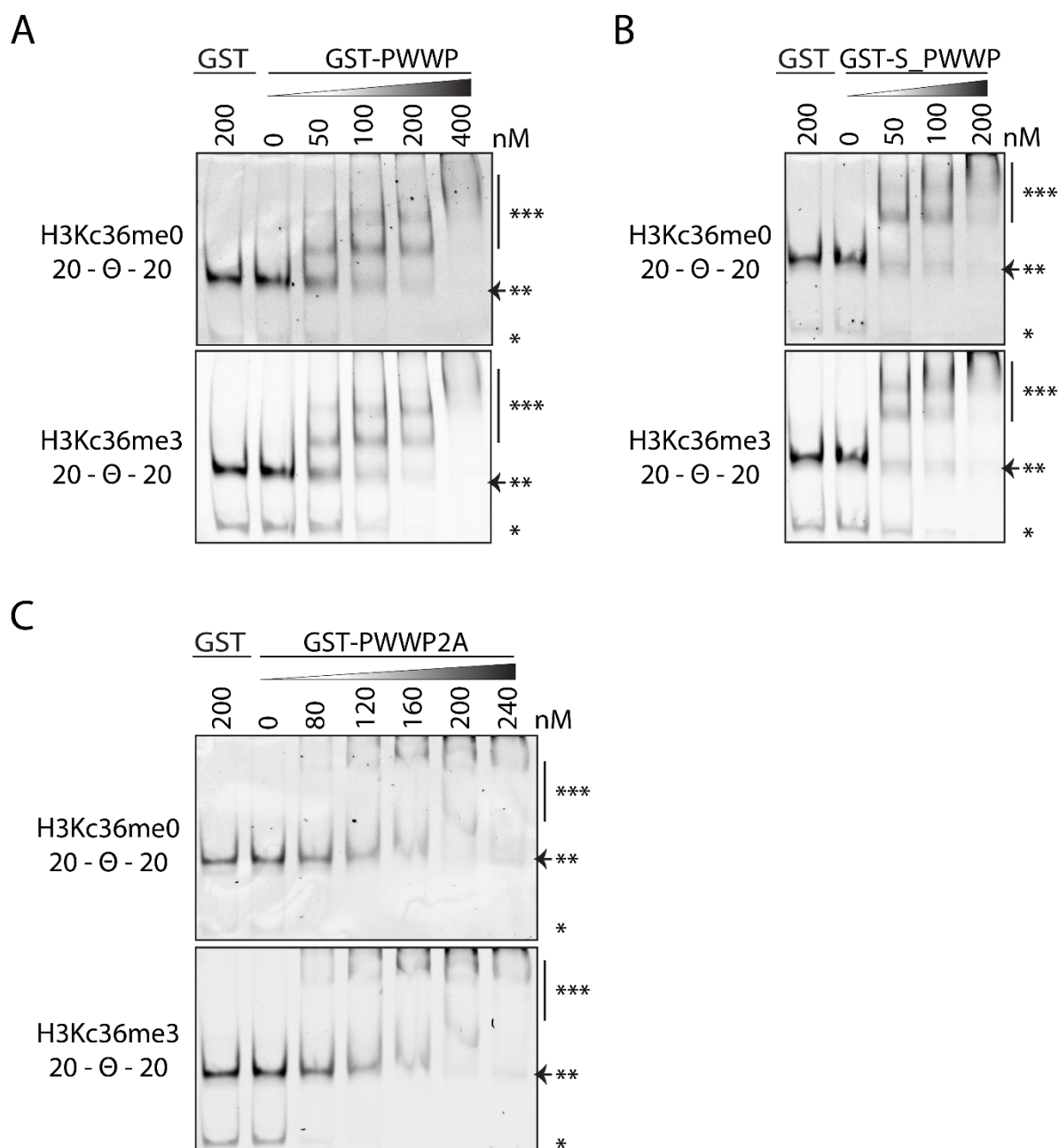


Figure A2: Neither PWWP2A nor its truncations distinguish between H3Kc36me0 and H3Kc36me3 *in vitro*. (A) Representative cEMSA (n=2) using 15 nM of recombinant frog 20 bp linker-containing (20- Θ -20) H3Kc36me0- or H3Kc36me3-mononucleosomes incubated with indicated increasing concentrations of GST-PWWP. GST served as negative control. *free DNA, **nucleosome, ***nucleosome GST-protein complex. Arrowhead marks loss of band signal intensity when mononucleosome-GST-PWWP complex is formed. (B) Representative cEMSA (n=2) similar to (A) using 15 nM of recombinant frog 20 bp linker-containing (20- Θ -20) H3Kc36me0- or H3Kc36me3-mononucleosomes incubated with indicated increasing concentrations of GST-S_PWWP. (C) Representative cEMSA (n=2) similar to (A) using 15 nM of recombinant frog 20 bp linker-containing (20- Θ -20) H3Kc36me0- or H3Kc36me3-mononucleosomes incubated with indicated increasing concentrations of GST-PWWP2A.

CURRICULUM VITAE

CURRICULUM VITAE
