Chromatin affinity purification coupled with mass spectrometry identifies novel histone ubiquitylation interactors

PhD Thesis

in partial fulfilment of the requirements for the degree
Doctor of Philosophy
in the Molecular Biology Graduate Program
at the Georg August University Göttingen,
Faculty of Biology

submitted by

Stefan-Sebastian David

born in
Iasi, Romania

2018

First Referee: Prof. Dr. Wolfgang Fischle, Chromatin Biochemistry Group, Max-Planck-Institute for Biophysical Chemistry, Göttingen
Second Referee: Prof. Dr. Patrick Cramer, Department of Molecular Biology, Max-Planck-Institute for Biophysical Chemistry, Göttingen

Extended PhD Thesis Committee
Prof. Dr. Claudia Höbartner, Institute for Organic and Biomolecular Chemistry, Georg-August-University, Göttingen
Prof. Dr. Hening Urlaub, Bioanalytical Mass Spectrometry Group, Max-Planck-Intsitute for Biophysical Chemistry, Göttingen
Prof. Dr. Andre Fischer, German Centre for Neurodegenerative Diseases, Göttingen
Dr. Alexis Caspar Faesen, Biochemistry of Signal Dynamics Group, Max-Planck-Institute for Biophysical Chemistry, Göttingen

Date of submission of the PhD thesis: February 16th, 2018
Date of oral examination: April 16th, 2018

Herewith I declare, that I prepared the PhD Thesis "Chromatin affinity purification coupled with mass spectrometry identifies novel histone ubiquitylation interactors" on my own and with no other sources and aids than quoted.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to acknowledge the moral, physical and scientific support I had from my doctor father for the entire duration of my PhD thesis. Prof. Wolfgang Fischle has been more than a mentor to me. Given the many hardships I was confronted with, it would have been impossible to finish the degree without his support. I appreciated discussing at length with him hypotheses, experiments and results. Most of the work presented throughout the thesis originated from ideas that were developed during such conversations.

I would like to acknowledge the support from the second referee of my PhD thesis. Prof. Patrick Cramer has welcomed me into his group during the third year of my doctorate study and facilitated the conduction of many experiments that are presented in the thesis. I would also like to thank the students and scientists in his group with whom I enjoyed talking about science and culture and discussing experiments on a day to day basis with. In particular, I want to thank Dr. Seychelle Vos and Franz Fischer, who taught me to use insect cells and Ute Neef, who performed all the insect cell work in the fourth year of my doctorate study.

I appreciate the support of Prof. Henning Urlaug and the staff in the Bionalytical Mass Spectrometry research group who facilitated the conduction of all the mass spectrometry experiments presented in the thesis.

I want to thank Prof. Ashraf Brick and his co-workers Dr. Malikanti Seenu and Dr. Guy Kamnesky from the Technion Insitute in Haifa, Israel for their help with the synthesis of the ubiquitylated histones.

I would like to thank Prof. Claudia Hoebartner, who has been part of the thesis committee since the very begining and has been supportive and constructive throughout all the meetings. I thank Prof. Andre Fischer and Dr. Alex Faesen who will be part of the extended thesis advisory committee and wil read the PhD thesis and participate in the defense.

I am thankful to all the past and current members of the Chromatin Biochemistry research group. My PhD developed on important concepts, setups and techniques which were first described in this group. In particular, I thank Dr. Miroslav Nikolov, who guided my first steps in chromatin biochemistry and mass spectrometry as a supervisor during my Master's thesis. I want to acknowledge the support of Dr. Maria Tauber and Dr. Yogesh Ostwal who have been very helpful and knowledgeable for the entire duration of my doctorate degree. I want to acknowledge the significant contribution of my friend Sebastian Burk and my student Ezgi Akidil who provided invaluable moral and physical support in the last year of my PhD. Ezgi Akidil purified all recombinant proteins from bacteria and insect cells which were used to understand the functional signnificance of SCML2 recruitment to chromatin.

I want to thank the International Max Planck Research School for Molecular Biology who financially supported me and a student helper to be able to complete the PhD degree.

Abstract

Histone ubiquitylation is the post-translational covalent attachment of ubiquitin to histones. Ubiquitylation of histones regulates chromatin-templated processes through a combination of structural and functional effects. To date, several ubiquitylation sites have been mapped on all histones, including residues from the flexible histone tails and residues from the structured core domains. It is not yet clear if the different ubiquitylation marks are read separately of each other or if they act in synchrony. How ubiquitylation of histones is recognised by the nuclear proteome remains a standing question in the field of chromatin biochemistry. To address this question, we prepared site-specifically ubiquitylated histones by native chemical ligation and incorporated them into homogeneous nucleosomal arrays. Histone H2B ubiquitylated at lysine 120 (H2BK120ub), H2BK34ub and H3K23ub nucleosomal arrays were used in chromatin affinity purification experiments coupled with mass spectrometry to find the proteins that were enriched by ubiquitylation. We showed that the different ubiquitylation marks generally recruited unique proteins and protein complexes and that ubiquitylated histones needed to be embedded within nucleosomes for recognition by the nuclear proteome. Three separate ubiquitylated constructs were prepared for the N-terminal ubiquitylation of histone H3: H3K18ub, H3K23ub and H3K18/23ub2. On all templates, DNA methyltrasnferase I (DNMT1) bound directly the H3 ubiquitylation marks and recruited ubiquitin specific protease 7 (Usp7) and Sex comb on midleg-like 2 (SCML2) to the modified chromatin. We showed that Usp7 could remove all three H3 ubiquitylation marks. SCML2 stimulated Usp7's deubiquitylation activity when mononucleosomes and nucleosomal arrays were used as substrates. We showed that SCML2 competed with DNMT1 for Usp7 binding and proposed that SCML2 stimulated Usp7 by stabilising an activating conformation in the enzyme directly on chromatin. We suggest that SCML2 controls DNMT1 recycling from H3 ubiquitylated chromatin to insure faithful maintenance of DNA methylation across the genome. Our work establishes native chemical ligation as an efficient method to prepare site-specifically ubiquitylated histones and affirms chromatin affinity purification coupled with mass spectrometry as a reliable tool for identification of novel interactors of modified chromatin templates.

Chapter 1

Introduction

1.1 Packaging of genetic information

Genetic information is stored in polymers of deoxyribonucleic acid (DNA) [1]. The diploid human genome is made of 46 DNA polymers which contain approximately 6 billion base pairs (bp) [2]. 10 bp of DNA form a complete turn of a right-handed helix with a pitch of 34 Å. If connected end-to-end, the 46 human DNA polymers would stretch to reach 2 m in length. To fit within the eukaryotic nucleus, the genetic material is stored in a specialised nucleoprotein complex referred to as chromatin [3], whose repeating unit is the nucleosome [4].

A nucleosome contains two copies each of four proteins which organise 145-147 bp of DNA [5]. Consecutive nucleosomes are connected by stretches of linker DNA, which in humans vary from 20 to 60 bp in length [7]. Histones are the architectural proteins that shape the nucleosome. Histones are small, highly basic and share a common fold referred to as the handshake motif (Figure 1.1A) [8]. The handshake motif consists of three α-helices which are connected by two short loops. This motif stabilises the interaction of histone H2A with histone H2B and of histone H3 with histone H4 (Figure 1.1B). In addition to the handshake motif, for stabilization of the H3/H4 tetramers, a four helix bundle forms between the C-terminal helices of the two H3 copies (Figure 1.1B). A similar four helix bundle interaction is formed between the C-terminal helices of histone H2B and H4. The interaction between H2B and H4 is important for positioning of the two H2A/H2B dimers under the H3/H4 tetramer to form a histone octamer (Figure 1.1B).

A nucleosome core particle organises 145-147 bp of DNA around the histone octamer into a left-handed DNA superhelix (Figure 1.2). The nucleosome core particle is shaped as a flat disk or a wedge with a height of approximately 100 Å and a length of 60 Å at its base and 25 Å at its apex [5] (Figure 1.2). The two H2A and H2B dimers are positioned in the lower halves of the two DNA gyres that make up the DNA superhelix. The H3 and H4 tetramer is positioned in the upper halves of the gyres, on each side of the nucleosome dyad axis. Nucleosomes are pseudo-symmetric complexes and present two identical surfaces on either side of the nucleosome dyad (Figure 1.2).

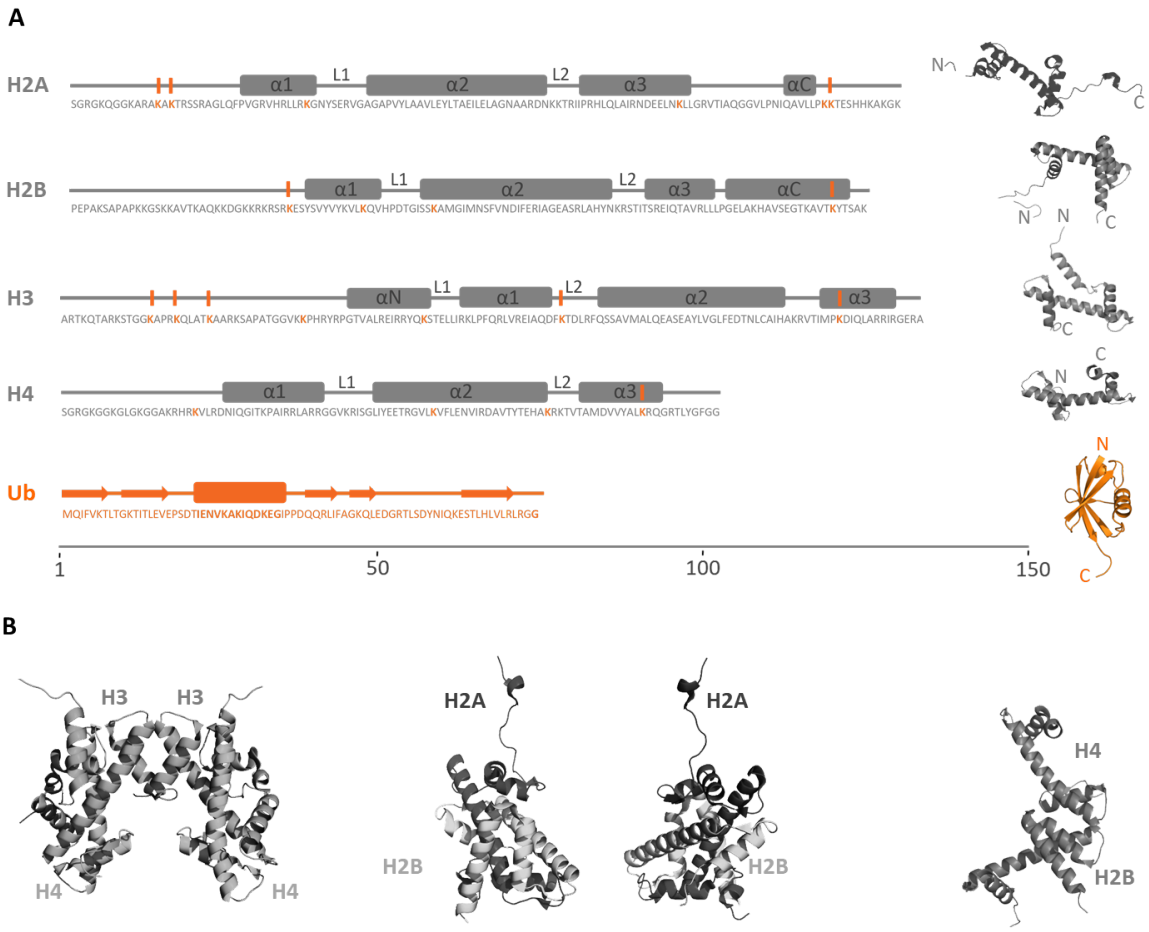


Figure 1.1: Structural details of suboctameric histone complexes. (A) Primary and secondary structure representations of ubiquitin and core histones. Histone cores share a common fold composed of three central α-helices (1-3) connected by two short loops (L1-2). Ubiquitin has a characteristic β-grasp fold containing a five-stranded β-sheet. Ubiquitin attaches to target lysine residues (orange prints, orange bars) through its C-terminal diglycine anchor. (B) The histone fold allows histone H2A to pair with histone H2B and form a dimer stabilised through hydrophobic interactions. Similarly, H3 and H4 pair to form a heterotetramer. A four-helix bundle forms between the two C-terminal helices of H2B and H4 and is responsible for the stabilisation of higher order hexameric and octameric histone assemblies. Structural details were adapted from published nucleosome (PDB 1AOI [5]) and ubiquitin (PDB 1UBQ [6]) crystal structures.

Genetic information is packaged into strings of nucleosomes. Chromatin is not uniform, as the architecture of the nucleosome would suggest. It assumes different folds in different parts of the nucleus. Several biological processes are only targeted to specialised chromatin areas. Genes are expressed in a tissue-specific manner. Chromatin is a dynamic environment, which integrates metabolic signals from within and environmental queues from outside of the cell. The collection of all processes that shape chromatin to regulate gene expression, with no direct interference to the DNA sequence, is referred to as epigenetic modifications.

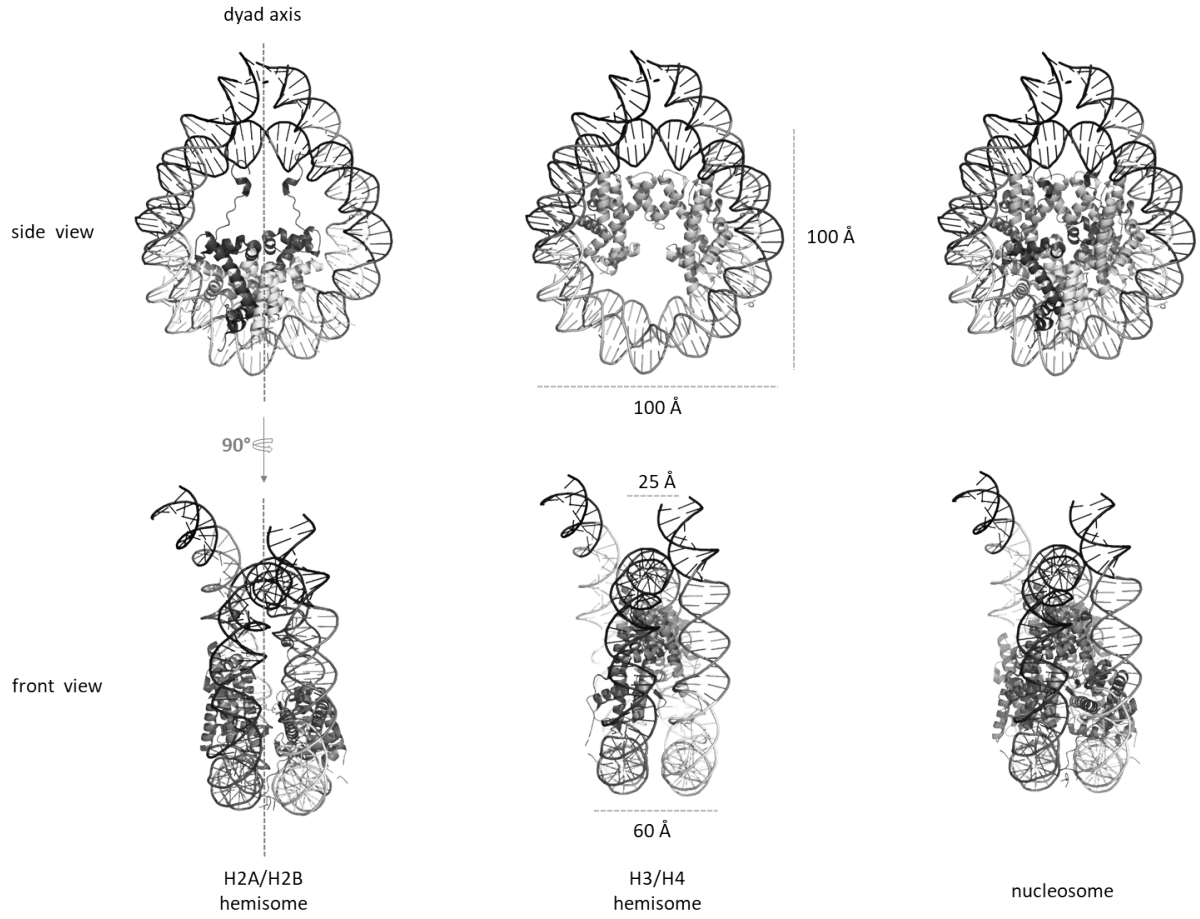


Figure 1.2: Nucleosome architecture. Representation of a 177 bp DNA template assembled around a histone octamer as visualised from the front and from the side. The globular domains of the two H2A/H2B dimers fit within the lower half of the nucleosome core particle, each heterodimer being positioned in planes parallel to the nucleosome dyad axis. The globular domains of the H3/H4 tetramer are positioned in the upper half of the nucleosome, diagonally with respect to the H2A/H2B dimer planes. Structural details were adapted from the published tetranucleosome crystal structure (PDB 1ZBB [9]).

1.1.1 Epigenetic modifications

The human embryo contains all genetic information that will be passed on to each of the more than 200 types of cells that make up the fully developed man. During development, cells differentiate and acquire lineage- and tissue-specific gene expression profiles. For long it was believed that development was one-directional, following the path from pluripotentcy to terminal differentiation [10]. Transcription factors are proteins that bind specific DNA sequences and control gene expression, maintaining thus cellular identity [11]. When scientists treated terminally differentiated cells with pluripotency-specific transcription factors, they managed to reprogram the differentiated cells back to pluripotency [12]. While the DNA sequence did not change in the process, chromatin must have reorganised in the reprogrammed cells to allow for the expression of the factors that control pluripotency.

Chromatin's flexibility during differentiation and de-differentiation (reprogramming) results from reversible chemical and physical changes to the DNA or to the histones.

In cytosine-guanine (CpG) dinucleotides, the carbon 5 of the cytosine base can be methylated. An estimated 80% of all CpG dinucleotides are methylated in the genome [13], [14]. Most of these methylation sites are found in noncoding, repetitive or transposable DNA elements. Methylation of DNA is read by dedicated protein complexes which possess additional enzymatic activities that silence chromatin. This mechanism is seen as an adaptation to protect the small part of the coding genome from the invading transposable elements or repetitive DNA sequences [15]. Promoter DNA elements contain high densities of CpG dinucletodies, which form CpG islands. Differently than the repetitive or the non-coding regions of the genome, most CpG islands are kept unmethylated [16]. This allows for uninterrupted transcription of house-keeping genes and for the regulation of tissue-specific gene expression profiles. DNA methylation is a modification that in mammals is erased pre-implantation and needs to be set up de novo afterwards [17]. Maintenance DNA methylation refers to the copying of already set CpG methylation marks during development [18]. Methylation of DNA is not reversible *per se*, but dedicated mechanisms have evolved for the conversion and removal of this mark from the genome [17].

Histones are subjected to many post-translational modifications [19]. The side-chains of amino acids can be modified by small hydrophobic groups (methylation), by small negatively charged groups (acetylation, phosporylation), by small negatively charged metabolites (biotinylation, butirylation, crotonylation, malonylation, propionylation) and by small polar sugars (glycosylation) [19], [20], [21]. These reversible modifications alter the chemical properties of the marked amino acids. In addition to them, histones can also be modified by conjugation of entire proteins: ubiquitin, SUMO and NEDD [20], [21]. In constrast to the small modifications, the later reversible marks are thought to also induce physical (structural) changes to the chromatin template [22], [23]. The combined chemical and physical modifications of histones make up a histone code [24] which is read, by specialised protein complexes that coordinate downstream signaling events [20], [21]. Histone post-translational modifications are reversible and specialised proteins have evolved to erase them [25], [26].

In addition to the chemical modification of cysotine (the 5th base) and the post-translational modification of histones (the histone code), chromatin flexibility is regulated by remodeling of nucleosome positioning, by incorporation of non-cannonical (variant) histones and by expression of non-coding RNAs [27], [28], [29]. All these epigenetic mechanisms are interconnected and work to relay cellular or environmental signals to the chromatin template.

1.1.2 Cross-talk between DNA and histone modifications

Dedicated protein complexes have evolved to recognise and coordinate modifications on the DNA template with modifications on the histone octamer core, such that chromatin functions as a whole. As previously explained, at CpG sites, DNA may be found unmethylated, hemi-methylated or fully methylated.

Unmethylated CpG sites are recognized by proteins that contain a zinc finger cysteine-x-x-cysteine (ZF-CxxC) domain [30], [31]. CxxC finger protein 1 (CFP1) binds unmethylated CpG islands and recruits the SET1 methyltransferase which modifies the lysine 4 of histone H3 (H3K4) [30]. Trimethylation of H3K4 (H3K4me3) directs the transcription machinery to the 5' ends of active genes. Lysine demethylase 2A (KDM2A) also contains a ZF-CxxC domain which binds unmethylated CpG islands [31]. Targeted KDM2A removes the H3K36me2 marks, which are found in every third histone across the genome [31]. CFP1 and KDM2A coordinate at CpG islands to mark these as H3K4me3-enriched and H3K36me2-depleted, transcriptionally poised regions [32].

Hemi-methylated DNA is recognised by proteins containing set and ring associated (SRA) domains [33]. At replication forks, ubiquitin like domain containing plant homeodomain and really interesting new gene finger 1 (UHRF1) recognises using its SRA domain hemi-methylated CpG sites [33], [34]. This stimulates the enzyme to ubiquitylate histone H3 at lysine 23 [35]. The ubiquitylation mark is read by DNA methyltransferase I (DNMT1), which copies the CpG methylation pattern from the parental onto the daughter DNA strand [35], [36], [37].

Fully methylated CpG sites are recognised by methyl-binding domain (MBD) containing proteins. MBD containing protein 2 (MeCP2) binds fully methylated CpG sites [38] and recruits a complex of histone methltransferases and histone deacetylases to chromatin [39]. The histone methyltransferases deposit the silencing mark H3K9me3 which is subsequently read by heterochromatin protein 1 (HP1) [40]. This interaction creates a positive feedback loop which attracts more methyltransferases and deacetylases at the marked sites to propagate heterochromatin spreading [41].

1.2 Histone ubiquitylation

1.2.1 The serendipitous discovery of histone ubiquitylation

In 1975 a hormone which had the ability to induce B-cell differentiation was isolated from bovine thymus [42]. The hormone was later found in several other tissues and in extracts originating from yeast, plant and animal cells. This led the authors to refer to it as ubiquitous immunopoietic polypeptide. In parallel, the non-histone chromosomal protein A24 was purified from bovine thymus with the belief that it may regulate gene expression. Protein A24 was present in nucleosomes and had an amino acid sequence similar to that of histone H2A, but was larger than the later and had a unique architecture, containing two N-termini [43], [44]. Independently, the ATP-dependent proteolytic factor-1 (APF-1) was shown to be conjugated to proteins before these undergo degradation in rabbit reticulocyte extracts [45], [46]. The three separate discoveries where connected in 1980 when it was shown that the ubiquitous immunopoietic polypeptide, the H2A conjugate and APF-1 all share the same amino acid sequence [47]. This led to the formation of the new research field of histone ubiquitylation.

1.2.2 Mapping of histone ubiquitylation sites

Soon after the discovery of ubiquitylated H2A, it was found that histone H2B was also ubiquitylated. It was approximated that 10% of histone H2A and 1-2% of histone H2B undergo ubiquitylation [48], [49]. If identification of the H2A attachment site was possible already in 1977, the difficulty to produce sufficient amounts of ubiquitylated H2B delayed the mapping of its conjugation site for one more decade [44], [50]. Later, it was found that core histone H3, linker histone H1 and variant histones H2A.X and H2A.Z are also ubiquitylated [51], [52], [53]. The identification of their conjugation sites had to wait for the development of highly sensitive peptide sequencing techniques which came with the improvement of ionisation, selection and detection methodologies used in mass spectrometry (Figure 1.1)(Table 1.1) [54], [55], [56], [57], [19]. While the global mass spectrometry investigations provided comprehensive lists with putative histone ubiquitylation sites, it took careful mutagenic approaches to map the involvement of the novel modifications in individual biological processes (Table 1.1) [58], [59], [60], [61].

1.2.3 Properties of ubiquitylated histones

Histone ubiquitylation referes to the covalent attachment of ubiquitin, through its C-terminus, to histone proteins [44]. All core and linker histones can be ubiquitylated (Figure 1.1)(Table 1.1). Ubiquitylation sites are found both in the flexible tails and in the structured histone fold domains (Figure 1.1). Ubiquitin attachment increases the mass of histones by 8.5 kDa, which is more than half of the mass of each individual histone. Ubiquitin contains 76 amino acids, of which six conserved lysine residues are involved in poly-ubiquitin chain formation [62]. With one notable exception [59], histones are monoubiquitylated. While generally polar, the ubiquitin surface contains a hydrophobic patch, centered on isoleucine I44. This serves as a binding platform for many ubiquitin readers [6], [63], [64]. Ubiquitylation of histones is a reversible process, which is controlled by ubiquitylation/deubiquitylation cycles [65], [66]. The ubiquitin deposited on histones is removed by ubiquitin C-terminal hydrolases and recycled directly in the nucleus [67]. Monoubiquitylation adds up to the complexity of the histone code by creating new possibilities for hierarchical signaling and cross-talk with other epigenetic modifications.

1.2.4 Recognition of histone ubiquitylation

Ubiquitin is recognised by specialised proteins (ubiquitin readers) which bind either to the I44 hydrophobic patch or ubiquitin's C-terminal tail (in particular the two arginines R72 and R74) [63]. When small variations to the substrate are presented to the ubiquitin readers, as is the case with ubiquitylation of histones, discrimination between different ubiquitylation marks is achieved by dedicated ubiquitin interaction motifs (UIMs), which fold either as single elongated alpha helices (Rad18, RNF169, DNMT1) [68], [37] or disordered loops or fingers (Usp7; 53BP1, UBP8, DNMT1) [69], [70], [71], [72].

The evolved UIMs contain several polar and negatively charged residues as well as a number of hydrophobic amino acids. While the hydrophobic core of the motif binds the I44 patch

of ubiquitin, the charged residues bind its basic side areas. Some UIMs contain additional amino acids which probe the surface of the nucleosome (Figure 1.3). In the case of the transcription factor 53BP1 and the E3 ligase RNF169, their respective UIMs contact both the ubiquitin I44 patch and the nucleosome acidic patch [71], [68]. In the case of the ubiquitin specific protease UBP8, its annotated UIM is anchored on the I44 patch but also contacts the two C-terminal arginines of ubiquitin, probing the isopeptide linkage [72]. In the case of DNMT1's RFTS domain, which is able to bind two ubiquitin groups at once, the UIM also contacts the underlying histone tail in its entirety [37]. The strength of the association and the specificity of the interaction between the reader and ubiquitin may be fine-tuned by the presence of adaptor subunits which contact both the reader and the nucleosome. In the case of UBP8, the yeast SAGA - associated factor Sgf11 serves as an adaptor protein which contacts both the reader and the nucleosome, at its acidic patch [72].

Even though all known histone ubiquitylation readers display a common binding pattern, they share no consensus primary amino acid sequence in their ubiquitin interaction motifs. Identification of novel histone ubiquitylation readers using *in silico* algorithms is limited by the small number of known histone ubiquitylation-specific UIMs and the difficulty to assign reliable defining features within them. Identification of new histone ubiquitylation readers remains a challenging endeavour.

1.2.5 Representative histone ubiquitylation marks

This thesis aimed to prepare several site-specifically ubiquitylated histones in order to find out how these post-translational modifications are read by the nuclear proteome. In particular, we wanted to resolve how different proteins and protein complexes, including direct histone ubiquitylation readers, are interacting with the modified histone constructs. To understand how the histone ubiquitylation readout is achieved we aimed to find the proteins that bind the modified histones and compare these with the proteins that bind modified mononucleosomes or modified chromatin fibers.

More than two dozen histone ubiquitylation marks have already been described in the literature (Figure 1.1)(Table 1.1). To select the ubiquitylation sites to be used in the systematic interactome analysis, several parameters were considered: the topology of the modification within the nucleosome core particle, previous documented association of the modification with a chromatin-templated process and the readiness to prepare the modification by expressed protein ligation. The thesis aimed to investigate if the different modifications recruit similar or unique proteins. The identification of the histone ubiquitylation interactors would help undestand how correct recognition of the modification is achieved and how this recognition contributes to the specificity of the downstream signalling events.

H2AK119ub was the first histone ubiquitylation site described in the literature [44]. This mark was the first post-translational modification found on histones and also the very first ubiquitylated protein found in cells [73]. H2AK119ub is presently seen as a heterochromatic modification. H2A ubiquitylated tetranucleosomes were used to describe a cross-talk between the polycomb repressive complexes PRC1 and PRC2 [74]. This connection was shown

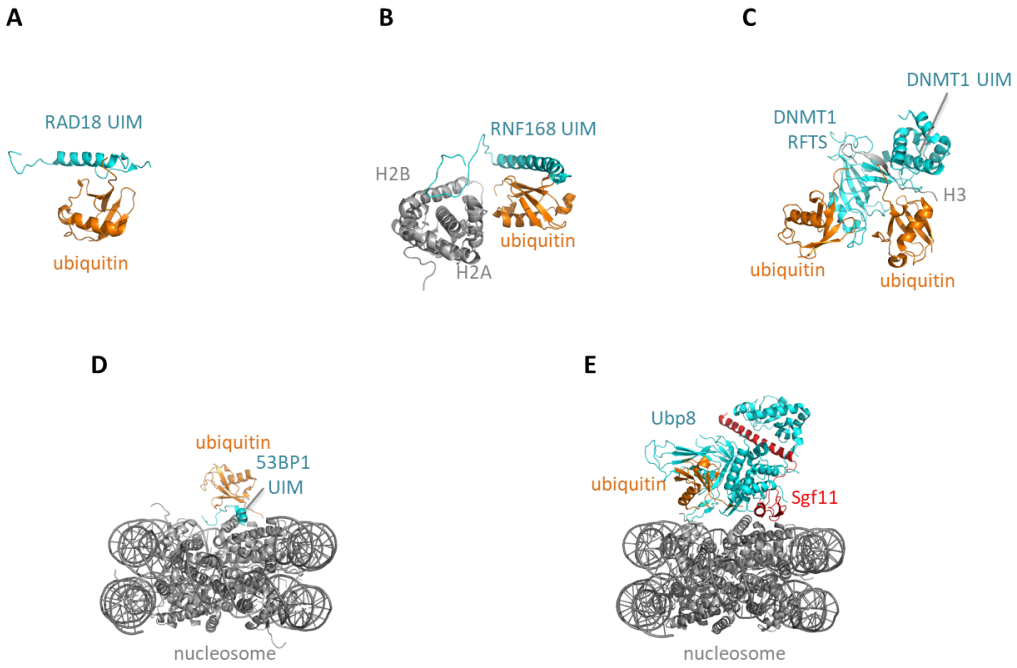


Figure 1.3: Recognition of histone ubiquitylation marks.(A) Recognition of ubiquitin by RAD18 requires a single helix ubiquitin interaction motif (UIM) (PDB 5VF0 [68]). (B) RNF168 uses an elongated UIM to recognise ubiquitin conjugated to a H2A/H2B heterodimer (PDB 5VEY [68]). (C) The RFTS domain of DNMT1 (containing a single helix UIM) binds two ubiquitin molecules conjugated to the N-terminal tail of histone H3 (PDB 5WVO [37]). (D) A single helix UIM in 53BP1 recognises both ubiquitin and the nucleosomal acid patch (PDB 5KGF [71]). (E) Sgf11 stabilises the interaction of Ubp8 with ubiuquitin and the nucleosome by contacting both Ubp8 and the nucleosomal acidic patch (PDB 4ZUX [72]).

to be needed for heterochromatin spreading by formation of a positive feedback loop between H2AK119ub and H3K27me3 at unmethylated CpG islands (promoters) [75], [76]. H2AK119 is found on the C-terminal tail of histone H2A, which protrudes outside of the nucleosome core particle, around the nucleosome dyad axis (Figure 1.4).

H2BK120ub was the second histone ubiquitylation mark described in the literature [49]. This modification is presently associated with transcriptionally active chromatin. Already in 1989, using the *Tetrahymena* model organism, H2BK120ub was found enriched in this organism's active macronucleus. Removal of the yeast ubiquitin conjugating enzyme RAD6 abolished H3K4me3, a methylation mark that had been associated with active chromatin [77], [78]. H2BK120ub facilitated *in vitro* RNA Polymerase II transcript elongation in the presence of the histone chaperone FACT and the regulatory PAF complex [79]. H2BK120ub activated *in vitro* the histone methyltransferase responsible for deposition of H3K4me3 [80]. H2BK120ub localised at actively transcribed gene bodies in live cells [81], [82]. The H2BK120ub interactome contained proteins involved in transcription elongation and RNA editing [83]. Finally, levels of H2BK120ub have been shown to respond rapidly to external stimuli, such as hormones, cytokines and differentiation cues [82], [84], [85], [86]. H2BK120 is located in H2B's

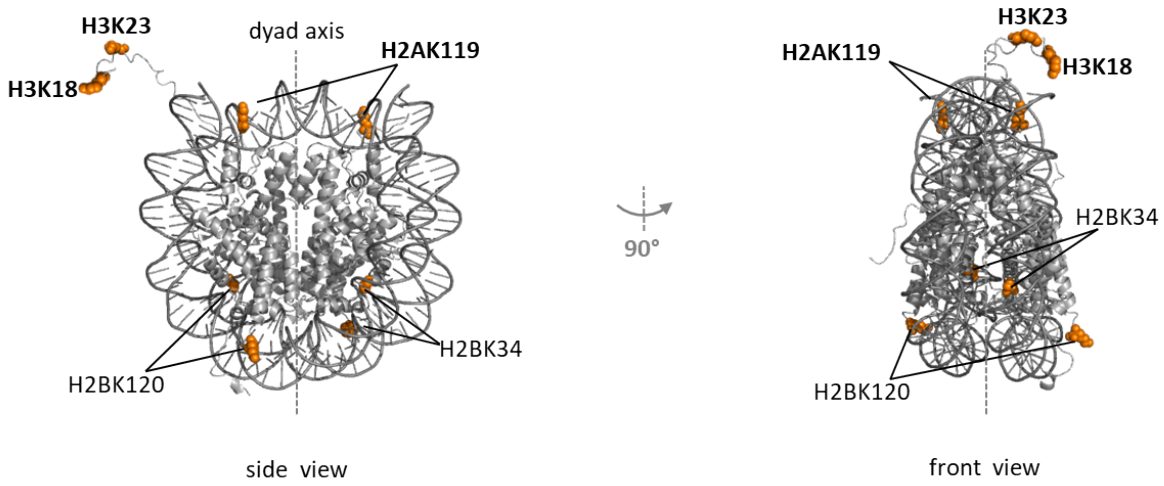


Figure 1.4: Topology of ubiquitylation sites. Different views of a 147 bp nucleosome core particle. The ubiquitylation sites prepared in this thesis (atoms highlighted as orange spheres) are located proximal to the dyad axis within (H2BK34) and just above (H2AK119) of the two DNA gyres, on the outer lower half of the nucleosome core (H2BK120) and away from it on one of the histone tails (H3K18; H3K23). Structural details were adapted from the published nucleosome crystal structure (PDB 1AOI [5]).

C-terminal α-helix, on a ridge close to the nucleosomal acidic patch (Figure 1.4).

H2BK34ub was discovered recently using modern mass spectrometry sequencing techniques. H2BK34ub was enriched from histone extracts obtained from neuronal cultures [55]. This ubiquitylation mark was also enriched by affinity purification using an antibody raised against the di-glycine anchor of ubiquitin [87]. The MSL-MOF E3 ubiquitin ligase complexes was proposed to be responsible for writing this modification [58]. It was suggested that H2BK34ub may stimulate H3K4 and H3K79 methylation in a mechanism that resembles the H2BK120ub transcriptional activation [58]. H2BK34 is located on the N-terminal tail of histone H2B which is buried inside the nucleosome core, between the two DNA gyres (Figure 1.4).

H3K23ub was recently associated with DNA replication [35]. The modification was found to be written by UHRF1 and shown to affect the localisation of DNMT1 to replication foci. With improved sequencing techniques and the identification of the E2 conjugating enzyme it was shown that, besides H3K23, also residues H3K14 and H3K18 were targeted by ubiquitylation [35], [36], [88]. Additional evidence suggested that the ubiquitylation mark was in fact a two-mono ubiquitylation signal where H3K14 and H3K18 or H3K18 and H3K23 act in concert for DNMT1 recruitment [37]. These ubiquitylation marks are located on the N-terminal tail of histone H3, which projects outside from the nucleosome, away from the dyad axis (Figure 1.4).

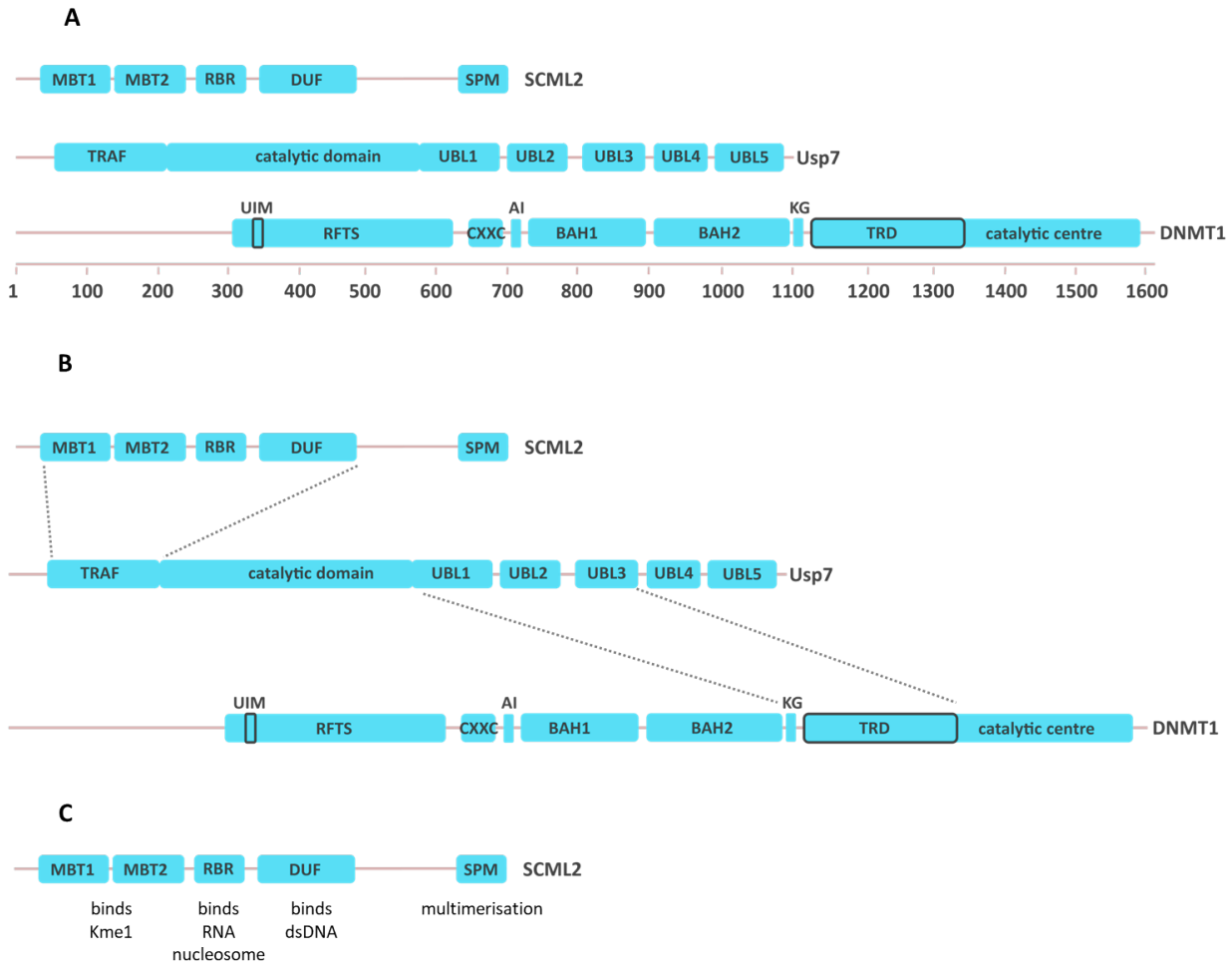


Figure 1.5: Primary structure of SCML2, Usp7 and DNMT1. (A) Schematic annotation of domains in Sex comb on midleg-like 2 (SCML2), Ubiquitin-specific protease 7 (Usp7) and DNA methyltransferase 1 (DNMT1). (B) Previously known interaction surfaces connecting SCML2 with Usp7 and Usp7 with DNMT1 are displayed as dotted lines. (C) Previously documented roles of individual domains in SCML2. MBT = malignant brain tumor; RBR = RNA-binding region; DUF = domain of unknown function; SPM = sex comb on midleg and polyhomeotic; TRAF = TNF receptor associated factor; UBL = ubiquitin like domain; UIM = ubiquitin interaction motif; RFTS = replication foci-targeting sequence; CXXC = cysteine XX cysteine; AI = autoinhibitory linker; BAH = bromo-adjacent homology domain; KG = lysine glycine rich linker; TRD = target-recognition domain; Kme1 = monomethylated lysine; dsDNA = double-stranded DNA.

1.3 Maintenance DNA methylation

After DNA replication, the methylation pattern of CpG dinucleotides needs to be copied from the parental DNA strand onto the daughter DNA strand. This process insures that areas of the genome which need to be silenced (repetitive sequences and transposable elements) are clearly marked. Maintenance DNA methylation differs from *de novo* DNA methylation in the fact that the substrate is already hemi-methylated. DNA methyltransferase I (DNMT1)

is the enzyme which performs most of the maintenance methylation [18].

Previously, targeting of DNMT1 to replication foci was shown to be mediated by the replication foci targeting sequence (RFTS) domain at the protein's N-terminus [89]. This domain was shown to be important for DNMT1's interaction with UHRF1 [90], [91]. Through this interaction, UHRF1, which contains a set and ring associated (SRA) domain, that recognises hemi-methylated CpG sites, was proposed to target DNMT1 for methylation.

Following the discovery that UHRF1 ubiquitylates H3K14/K18/K23 at hemi-methylated CpG sites it was suggested and then shown that DNMT1 is stimulated by these modifications [35], [36], [88], [37]. DNMT1 bound H3 ubiquitylation using an ubiquitin interaction motif (UIM) within its RFTS domain [36], [37].

This strong interaction brought into question the fact that DNMT1 needs to be recycled from the H3 ubiquitylated chromatin. It was suggested that the ubiquitin specific protease 7 (Usp7), which was known to interact both with UHRF1 and DNMT1 is responsible for DNMT1's recycling (Figure 1.5B) [90], [91], [92]. Inhibition of Usp7 or removal of the protease from frog extracts was shown to prolong DNMT1's recruitment to H3 ubiquitylated chromatin, which resulted in a loss of global DNA methylation [93]. It is currently unresolved how Usp7 is recruited to the modified chromatin and what activates the protease to remove the H3 ubiquitylation mark.

The isolated catalytic domain of Usp7 was previously shown to have a one hundred fold lower activity in comparison with the activity of the full-length (FL) enzyme [94]. FL Usp7 was proposed before to reversibly shift between an inactive and an active conformation, for which the presence of the ubiquitin like UBL4,5 domains was essential [95]. A peptide at the C-terminus of Usp7 was shown to fold back onto the catalytic domain to stimulate the enzyme, suggesting that large conformational changes, affecting all five ubiquitin-like domains of Usp7, are needed for enzymatic function [70].

The deubiquitylase was previously shown to be recruited to H2AK119ub and H2BK120ub marked nucleosomes [74], [96]. The activity of the enzyme was stimulated on H2BK120ub by guanosine 5'-monophosphate synthase (GMPS), which served as an allosteric regulator that shifted the Usp7 equilibrium towards an active conformation [96], [95].

Besides GMPS, DNMT1 and UHRF1, Usp7 was previously shown in a large scale proteomics study to bind a number of other nuclear factors [97]. The polycomb group protein Sex combs on midleg-like 2 (SCML2) was one of the additional targets. More recently, SCML2 was co-immunoprecipitated with Usp7 from live cells and shown recombinantly to interact with the enzyme [98].

It is currently unknown if SCML2 plays a role in the control of maintenance DNA methylation. SCML2 is a multidomain human homolog of *Drosophila* Scm (Figure 1.5A) [99]. The protein uses the different domains to receive various signals from its surrounding environment. SCML2 binds long noncoding RNA and is thought to be recruited to chromatin

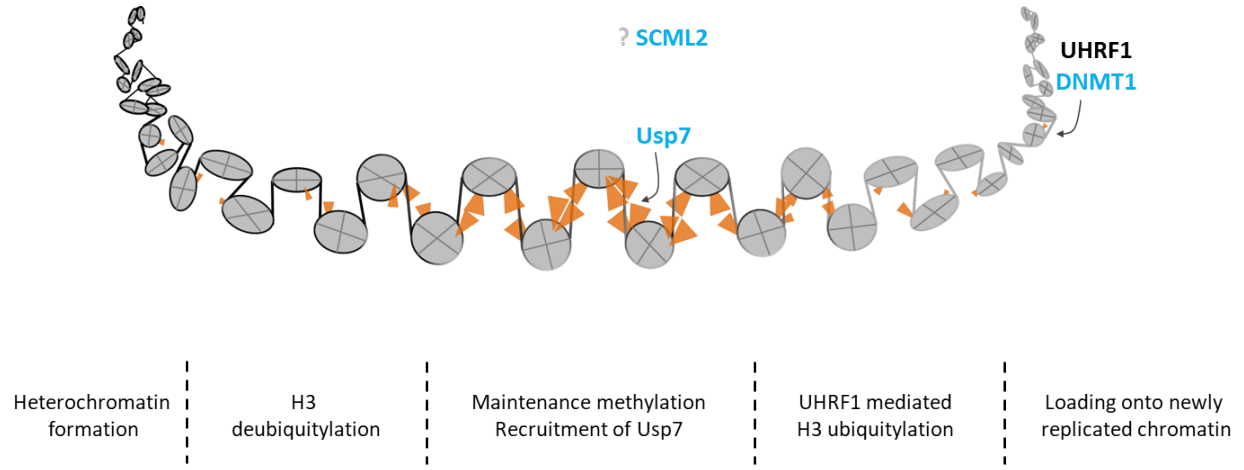


Figure 1.6: Current model of the events happening on the H3 ubiquitylated chromatin. UHRF1, DNMT1 and PCNA form a complex that travels with the DNA replication fork. UHRF1 senses hemi-methylated CpG sites and ubiquitylates histone H3. DNMT1 binds H3 ubiquitylated nucleosomes, is activated by this modification and converts hemi-methylated DNA to fully methylated DNA. Usp7 interacts with UHRF1 and DNMT1 and removes H3 ubiquitylation to facilitate DNMT1's recycling. DNMT1 = DNA methyltransferase 1; UHRF1 = ubiquitin-like containing plant homeodomain and really interesting new gene finger 1; PCNA = proliferating cell nuclear antigen; hemi-meCpG = hemimethyalted DNA; meCpG = fully methylated DNA.

through this interaction [100]. In vitro, the RBR domain of SCML2 binds at the same time HOTAIR lncRNA and a nucleosome core particle [100](Figure 1.5C). SCML2 binds DNA weakly with its DUF domain (Figure 1.5C) [101]. The second MBT domain of SCML2 recognises monomethylated lysine residues, in particular H4K20me1 (Figure 1.5C) [102]. SCML2 is expressed in two isoforms. The long SCML2 isoform, SCML2A, differs from the shorter one, SCML2B, in the presence of the C-terminal SPM multimerisation domain (Figure 1.5B) [103]. In *Drosophila*, the Scm SPM domain is thought to play a role in chromatin silencing [104]. SCML2 may achieve this silencing by interacting with itself or with SPM domains from other polycomb group proteins [105], [106], [100].

Since GMPS could stimulate Usp7, it may be possible that also DNMT1, UHRF1 or SCML2 which interact with the protease, control its activity. How Usp7 coordinates with DNMT1 and UHRF1 to control their recycling from the H3 ubiquitylated chromatin DNA replication and where SCML2 interferes in this cross-talk remains unresolved (Figure 1.6).

1.4 Protein engineering

The attempts to purify ubiquitylated histones from tissue or cellular extracts using traditional chromatography techniques have been very successful in obtaining the more abundant H2A and H2B ubiquitylated species [44], [49]. Ubiquitylated histones purified in this way were used to assemble ubiquitylated nucleosomes [107] and ubiquitylated nucleosomal arrays [108] to test if the presence of ubiquitin influences the nucleosome structure and the higher order folding. These protocols were not efficient at enriching for other less repre-

sented ubiquitylated constructs. In addition, they yielded products with additional post-translational modifications, which could interfere with the effects caused by ubiquitylation. These limitations forced researchers to look for alternative ways to prepare ubiquitylated histones.

The identification of the first E2 conjugating [77] and E3 ligase [109] enzyme pair allowed for the development of *in vitro* ubiquitylation assays. Histone H2B was modified at lysine K120 directly on a chromatin template to promote ongoing transcription [79]. The same approach was later used with other E2/E3 enzyme pairs on recombinant histones or nucleosomes to generate H2BK34ub and H2AK13ub/K15ub or H2AK119ub [58], [59], [74]. Even though quite efficient in producing H2BK120ub and H2AK119ub, this method fell short for H2BK34ub and H2AK13ub/K15ub as those ubiquitylation reactions yielded also unspecific products.

In parallel, site-directed mutagenesis of ubiquitin and histone H2B has been used for preparation of H2BK120ub through a disulfide linkage [22]. This approach produced a non-native cystine that was incompatible with reducing buffers.

To circumvent the sensitive nature of the disulfide linkage and address all other previous limitations, several semi-synthetic approaches for histone ubiquitylation have been described (Figure 1.7A) [80], [110], [111], [61], [112]. The most successful strategy, which leaves a native isopeptide linkage, was native chemical ligation and was independently adapted to histones in the laboratories of Tom Muir and Ashraf Brick [80], [110].

1.4.1 Expressed protein ligation

Expressed protein ligation is a novel protein engineering tool that merges recombinant protein production with peptide chemistry to facilitate site-specific incorporation of difficult post-translational modifications. At the centre of the native chemical ligation strategy lies the specific chemical reaction between a thiol and a thioester [113]. As neither ubiquitin nor any of the core histones contains a thiol-bearing amino acid, both reactive moieties needed for native chemical ligation can be artificially added at the desired locations. To react ubiquitin with target peptides or proteins by means of chemical ligation, a thioester is inserted at its C-terminus and an unnatural lysine-mimic amino acid, containing a thiol on its penultimate carbon, is introduced at the target location in the (poly)peptide chain. The first modification is achieved using intein splicing protocols [114], [80], the second modification is incorporated using genetic code expansion or peptide chemistry [115], [80], [110].

Inteins are proteins with self-splicing activity [116]. Premature proteins, like precursor messenger RNAs, contain external protein domains referred to as exteins and internal protein domains known as inteins [117]. The catalytic properties of inteins lie in their horseshoe-like shape which causes their N and C termini to be only 10 Å apart [118]. The splicing reaction leads to the removal of the intein and the ligation of the N- and C-exteins. This is achieved through the cleavage of two amide bonds and the formation of a novel peptide bond. The reactions occur spontaneously and do not require any cofactor or energy source [119]. Purifi-

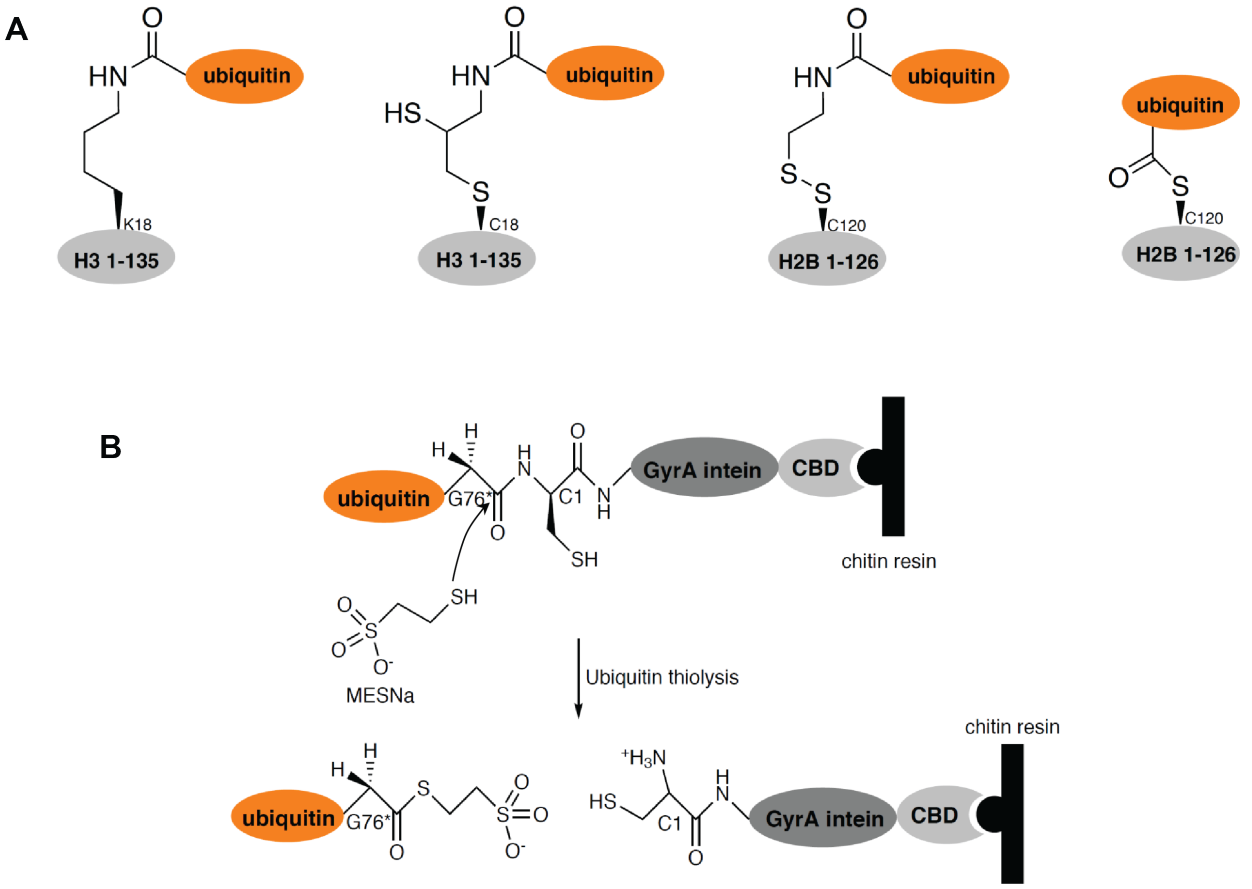


Figure 1.7: Ubiquitylation of histones by expressed protein ligation. (A) Non-enzymatic attachment of ubiquitin to histones was achieved using several different strategies: isopeptide linkage (this thesis), thiirane linkage [112], disulfide linkage [22] and thioester linkage [111]. (B) Ubiquitin thioester was prepared in this thesis by sodium mercaptoethanesulfonate (MESNa) treatment of a purified ubiquitin-GyrA intein-chitin binding domain fusion construct.

cation of ubiquitin and histone C-terminal thioesters relies on thiol-mediated intein cleavage (Figure 1.7B) [113], [114].

1.5 Mass spectrometry and discovery proteomics

1.5.1 Peptide sequencing

Mass spectrometry is a technique that calculates the mass of molecules as a function of their charge state. Positive ion mass spectrometry is the analysis of molecules in an acidic environment. In proteomic research, positive ion mass spectrometry is used to ionise the amino acids lysine and arginine which have high pKa values. The most common analytes used in proteomic research are peptides, whose mass and charge can be extracted from the ions that fall within the scanning range of most mass analysers. The peptide mass refers to

the sum of the individual masses of amino acids and accounts for the loss of 18 Da during formation of amide bonds. The total peptide charge is the sum of all lysine and arginine residues and accounts for the presence of an ammonium ion ($NH3^+$) at the N-terminus and of a carboxylic acid (COOH) at the C-terminus.

Mass analysers separate peptide ions based on their mass-over-charge (m/z) ratios. Inside the mass analyser, peptides form ionisation patterns (charge states) of different m/z ratios. A minimum of two charge states is needed to identify the mass of a peptide. The charge z of a single ion from the ionisation pattern, is calculated using the ion's less intense isotopic peaks given by the naturally-occuring isotope ^{13}C. This isotope is present in all proteins at a concentration of 1% and adds 1 Da per carbon to the peptide ion. The isotopic peaks have m/z ratio of (m+1)/z. The difference between the ^{12}C m/z and the isotopic ^{13}C (m+1)/z is used to calculate the charge z. The mass analyser collects m/z ratios from a minimum of two consecutive charge states and uses this information to deconvolute the peptide mass m.

The peptide mass calculated in the first mass analyser is an important parameter, but it is insufficient for identification of the amino acid sequence of the peptide [120]. Modern mass spectrometers contain two (or even three) mass analysers connected in series. Between the first and the second mass analyser, there is a fragmentation cell which uses high energy to break peptide bonds [121]. The parental ion is fragmented into product ions which are analysed in the second mass analyser. The product ion series (b and y ion series) is made of smaller peptide ions which have lost amino acids from the N- and the C-termini of the parental peptide (Figure 1.8). Using the parental peptide mass and several product ions masses, a number of amino acids from the N- and the C-termini of the proteins can be assigned [120].

Peptides could in theory be "sequenced" relying solely on the product ion masses. However, in a complex mixture of peptide ions, a complete series of product ions is rarely observed in the spectra collected from the second mass analyser. As such, the information provided by the two mass analysers is complemented with the use of peptide databases that are prepared *in silico* [122], [123], [124]. These databases provide a search space for peptide sequencing algorithms which match the measured product ions with the synthetic peptide library. The database includes the amino acid sequences of the proteins that are present in the analyte and account for the use of proteases (most often trypsin), which are used to generate the peptides.

Several software packages were developed to extract the data obtained from the mass spectrometer using synthetic peptide databases as search spaces [122], [125].

1.5.2 The use of mass spectrometry in discovery proteomics

The high resolution capacity of modern mass spectrometers allows for the use of isotopic tags (chemical tags, chemical labels and stable isotopes of amino acids) [126], [127], [128]. Mass analysers separate between proteins that were exposed to the isotopic tag and those which were not based on the differences in the parental ion masses. As such, stable isotopes of the amino acids lysine and arginine (which are present in all the peptides digested by

trypsin) [129] can be used to label a cell's entire proteome [128]. Stable isotope labeling of amino acids in cell culture (SILAC) facilitates the identification of proteins that change their abundance with respect to a particular treatment applied to cells [128]. The technique allows also for the preparation of labeled cellular or subcellular extracts that can be used in affinity purification experiments outside of the cellular environment [130]. SILAC nuclear extracts were previously used in affinity purification experiments to address which nuclear interactors bind modified histone tails [131]. The technique was also used in chromatin affinity purification (ChAP) experiments, to compare the factors which bound modified nucleosomal arrays to the interactors that were previously enriched with modified histone peptides [132]. Recently, chromatin affinity purification coupled with mass spectrometry (ChAP-MS) has been used to find the nuclear interactors that bind to ubiquitylated nucleosomal arrays [83], [74].

To distinguish between false positive and true positive identifications, chromatin affinity purification was designed to contain a forward and a reverse biochemical experiment (Figure 3.11A) [132]. In the forward experiment, the modified chromatin is incubated with SILAC-labeled (heavy) nuclear extract and the unmodified chromatin is incubated with unlabeled (light) nuclear extract. The eluates from the two purifications are mixed such that the pooled forward eluate contains both light and heavy proteins. In the reverse experiment, the modified chromatin is incubated with light nuclear extract and the unmodified chromatin is incubated with heavy nuclear extract. The pooled reverse eluate contains both light and heavy proteins. True positive identifications refers to the factors which are enriched in the forward experiment and depleted in the reverse experiment.

The pooled eluate from the forward experiment is separated by SDS-PAGE to reduce sample complexity and proteins are digested by trypsinisation (Figure 1.8). This step relies on the assumption that both the light and the heavy proteins are digested to the same extent by trypsin. Peptides originating from the forward or the reverse experiment are separated prior to detection using high performance liquid chromatography (HPLC). HPLC separation of corresponding heavy and light peptides is assumed not to be influenced by their mass difference. After elution from the HPLC column, peptides are ionised and analysed, assuming that neither the ionisation ability, nor the detection in the two mass analysers is affected by the peptide pair mass difference. Pairs of heavy and light parental ions are analysed in the first mass spectrometer (Figure 1.8). The ratio between the intensity of the heavy parental ion and the intensity of its corresponding light parental ion is referred to as the H/L ratio. An H/L ratio higher than 1 suggests that the heavy protein was more abundant than the light protein. An H/L ratio higher than 1 in the forward experiment indicates that the heavy protein preferred the modified chromatin template over the unmodified control.

Assignment of enriched factors was previously done using a cutoff value that was chosen based on the distribution of all H/L ratios [131], [132], [83], [74]. More recently, statistical analysis was introduced to measure reproducibility of biological experiments and technical measurements [133]. Mass spectrometry is thus coupled with affinity purification to discover new proteins that are enriched by a particular modification.

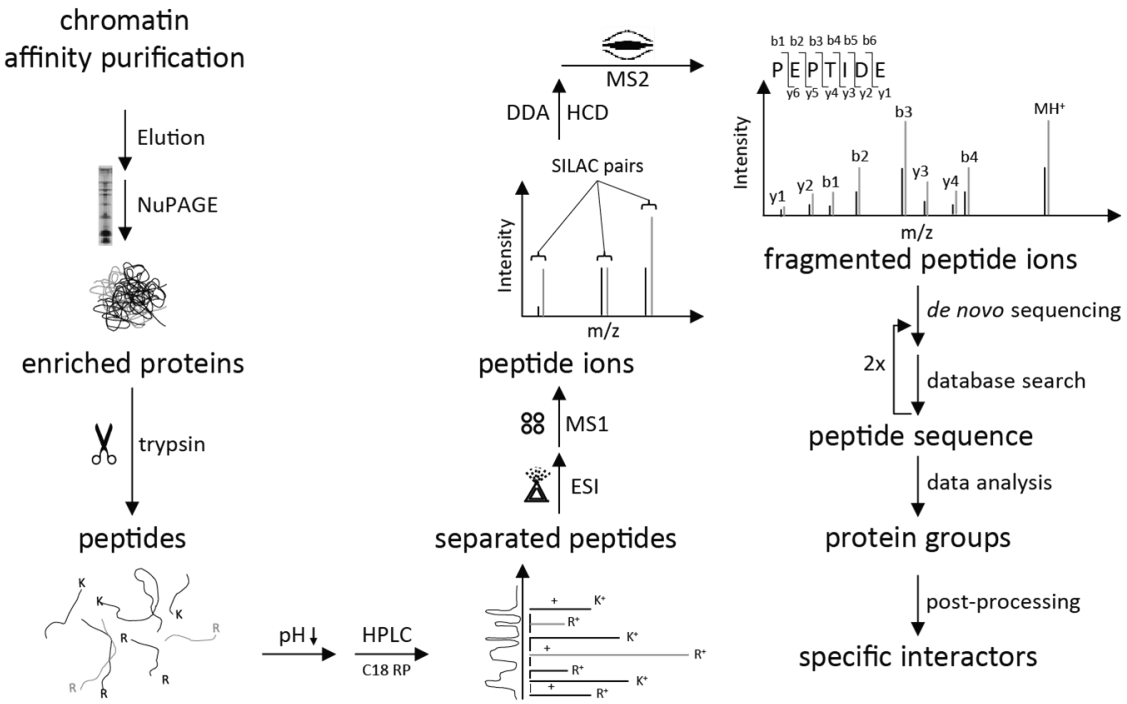


Figure 1.8: Identification of chromatin interactors by mass spectrometry. Enriched proteins from the forward or reverse chromatin affinity purification experiments are separated according to their molecular weight by polyacrylamide gel electrophoresis. The gel is sliced to reduce sample complexity and proteins from each gel slice are digested with trypsin. Peptides are eluted from the gel slices in an acidic environment and separated by high performance liquid chromatography on a reverse phase C18 column according to their hydrophobicity index. Peptides are injected into the mass spectrometer by electrospray ionisation and separated according to their mass-over-charge (m/z) ratio using a quadrupole mass analyser. SILAC pairs (light and heavy peptide ions) are resolved based on their different m/z values [128]. Most intense peptide precursor ions are selected in data-dependent acquisition mode and fragmented by high-energy collision dissociation to give rise to series of product ions. The product ions are measured in an orbitrap which serves as the second mass analyser. *De novo* sequencing algorithms (MaxQuant [125]) rely on species-specific databases of tryptic digests (Andromeda [124]) and use the m/z ratio of the product ions to reconstruct peptide sequences. Reconstructed peptides are assembled into protein groups and the corresponding summed heavy and light peptide intensities are scored to obtain heavy-over-light (H/L) ratios. Statistical post-processing is performed on the reported H/L ratios (Perseus [133]) to quantitatively assess modification-specific chromatin enrichment.

1.5.3 Crosslinking mass spectrometry

Crosslinking mass spectrometry is emerging as a technology that allows for the identification of protein-protein interaction surfaces which can be used to assist the reconstruction of low resolution structures [134], [135], [136], [137]. Crosslinkers form covalent bonds with target amino acid residues. Bifunctional crosslinkers interact with two amino acids residues from within the same protein or from two different proteins. Stable as well as flexible interactions can be detected using crosslinkers with different arm lengths (distance between the two functional groups). Bis(sulfosuccinimidyl)suberate BS3 is an amine-specific crosslinker

that is used to stabilise medium range interactions. Its two homophilic functional groups are separated by a spacer of 11.4 Å [138].

Mass spectrometry can be used to analyse crosslinked peptides in order to identify the crosslinked residues [139]. Most challenging in the identification of crosslinked sites is the generation of specific peptide databases. Standard peptide libraries used in *de novo* sequeucing algorithms, are not useful because crosslinked peptides contain two separate and non-adjacent amino acid sequences which are covalently attached. To sequence crosslinked peptides, special databases and search algorithms are generated [140]. The identification of crosslinked sites needs to account for the primary amino acid composition of the crosslinked proteins, the enzyme used to prepapre the peptides, all crosslinkable sites (every lysine in the case of BS3) and the mass of the crosslinker (572.43 Da). Special software packages are dedicated to the analysis of crosslinked peptides [140].

1.6 Objectives of the PhD thesis

Histone ubiquitylation is a post-translational modification which introduces chemical and physical changes to chromatin. Several histones residues are modified by ubiquitylation. Distinct histone ubiquitylation marks are generally associated with unique biological functions. We hypothesise that ubiquitylation of histones creates unique chromatin environments which recruit specialised proteins and protein complexes that control different chromatin-templated processes.

Our understanding of the mechanism through which histone ubiquitylation marks are recognised by the nuclear proteome is limited presently by the availability of tools to study the modification in a living organism or in an *in vitro* system. To understand how different protein factors are recruited to ubiquitylated histones, these can be used directly in affinity purification experiments. Since ubiquitylated histones cannot be expressed recombinantly, the production of ubiquitylated histones relies on the development of *in vitro* enzymatic systems specific for every single modification. A more general strategy needs to be designed to be able to produce ubiquitylate histones. Using ubiquitylated histones in affinity purification experiments may increase the identification of novel direct histone ubiquitylation readers, but would fall short in identifying interactors whose enrichment depends on additional features present on the (modified) chromatin fibers. A standardised protocol which makes use of chromatinised modified histones as affinity purification matrices needs to be developed.

Previously in our laboratory, native chemical ligation was employed to produce semi-synthetic H2BK120ub which was then chromatinised and used to find the nuclear proteins that associate with this modification [132]. This study set up the systematic analysis of histone ubiquitylation readers which will be described in this PhD thesis. The main objective of the PhD thesis is to develop the tools to efficiently ubiquitylate histone residues and reliably identify the nuclear proteins which recognise these modifications. To test the mapping strategy, a subset of enriched interactors will be analysed recombinantly in the context of the modified histone which was used for their identification.

The first aim of the PhD thesis is to develop a general strategy to prepare ubiquitylated histones and incorporate them into uniform nucleosomal arrays.

The second aim of the PhD thesis is to map the nuclear proteins which bind distinct ubiquitylated histones and find what chromatin features are needed in the readout process.

The third aim of the PhD thesis is to focus onto the proteins enriched by the N-terminal ubiquitylation of histone H3 marks and test the specificity of these interactors recombinantly. We focused on the description of the interactions between DNMT1, Usp7, SCML2 and the H3 ubiquitylated chromatin in order to understand the regulation of Usp7's activity.

Table 1.1: Histone ubiquitylation sites. Histone ubiquitylation writers, readers and erasers. Biological functions.

Histone	Residue	E2 and E3 enzymes	Readers	Deubiquitylating enzymes	Biological function
H2A	K13/15	RNF168; RNF8 mono/poly-Ub [59]	53BP1 [141]	Usp16 [142]; Usp3 [143]	DNA repair
	K36				
	K95				
	K118/119	Ring1A/B [144]	JARID2 [145]	Usp21 [146]; MYSM1 [147]; Bap1 [148]	heterochromatin spreading
	K127/129	BRCA1 [149]			DNA repair
H2B	K34	MSL1/2 [58]			transcription elogation
	K46				
	K57				
	K108				
	K116				
	K120	RNF20/40 [150]; Mdm2 [151]; ARID1 [152]	Ubp8 [72]	Usp7 [96]; Usp12 [153]; Usp22 [154]	transcription elogation
H3	K14/18/23	UHRF1 [35]	DNMT1 [35]	Usp7 [93]	maintenance methylation
	K36	NEDD4 [155]			oncogenic transcription
	K56				
	K79	Rag1 [156]; CUL4A/B poly-Ub [157]			V(D)J recombination
	K122	CUL4A/B [60]			chromatin assembly
H4	K31				
	K59				
	K77				
	K91	BBAP [158]			DNA damage response

Chapter 2

Materials and Methods

2.1 Instruments and Equipment

Table 2.1: List of common instruments and equipment

Instrument	Company	Addrees
AEKTA Explorer/Purifier	GE Healthcare	Buckinghamshire (UK)
Balances	Sartorius	Goettingen (DE)
Centrifuge Sorvall Evolution RC	Thermo Scientific	Braunschweig (DE)
Centrifuge 5415R/5810R	Eppendorf	Hamburg (DE)
Electrophoresis power supplies	Bio-Rad	Mnchen (DE)
Ultimate 3000 NanoLC	Thermo Fisher Scientific	Braunschweig (DE)
Qexactive HF	Thermo Fischer Scientific	Bremen (DE)
Mini-PROTEAN Tetra PAGE cell	Bio-Rad	Muenchen (DE)
Molecular Imager Gel Doc	Bio-Rad	Muenchen (DE)
NanoDrop ND-1000	Peqlab	Erlangen (DE)
Peristaltic pump	Ismatec	Glattburgg (CH)
pH meter	Metler-Toledo	Giesen (DE)
Prominence HPLC	Shimadzu	Kyoto (JP)
Sonication bath SONOREX Super	BANDELIN Electronic	Berlin (DE)
Sorval SLC1500 rotor	Thermo Scientific	Braunschweig (DE)
SpeedVac Savant SPD121P	Thermo Scientific	Braunschweig (DE)
Sub-Cell agarose gel electrophoresis	Bio-Rad	Muenchen (DE)
Thermo F8-6x1000y rotor	Thermo Scientific	Braunschweig (DE)
Thermocycler epgradientS	Eppendorf	Hamburg (DE)
Thermomixer Comfort	Eppendorf	Hamburg (DE)
Turbo-Blot system	Bio-Rad	Muenchen (DE)
Water bath TW12	Julabo	Selbach (DE)
XCell Sure Lock Mini NuPAGE cell	Invitrogen	Karlsruhe (DE)

2.2 Chemicals and Reagents

Bacterial strains

Escherichia coli XL-1 Blue [159]
genotype: F' proAB lacIqZΔM15 Tn10 Tetr recA1 endA1 gyrA96 thi-1 hsdR17 supE44
resistance: tetracyclin

Escherichia coli DH10BacCre [160]
genotype: F mcrA (mrr-sdRMS-mcrBC) 80lacZ M15 lacX74 recA1 endA1 araD139 (ara, leu)7697 galU galK rpsL nupG/bMON14272/pMON7124 pBADZ His-Cre
resistance: tetracyclin, kanamycin

Escherichia coli BL21 (DE3)-RIL [161]
genotype: F ompT hsdS (rB mB) dcm+ Tetr gal l (DE3) endA Hte [argU ileY leuW Camr]
resistance: chloramphenicol

Escherichia coli C2925 [162]
genotype: ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10) TetS endA1 rspL136 (StrR) dam13::Tn9 (CamR) xylA-5 mtl-1 mcrB1
resistance: streptomyicin, chloramphenicol

Genotype abbreviations: [163]

Insect cell lines

Spodoptera frugiperda Sf9 [164]
cell line: IPLB-Sf-21-AE, Thermo Fischer Scientific
insect cell medium: Sf900 II SFM medium, Thermo Fischer Scientific

Trichoplusia ni Hi5 [165]
cell line: BTI-TN-5B1-4, Expression Systems
insect cell medium: ESF 921 medium, Expression Systems

Human cell lines

HeLa clone S3: human cervical adenocarcinoma [166]

Table 2.2: List of common chemicals

Product	Ordering number	Supplier
4-(2-Hydroxyethyl)-1-piperazineethanesulfonate	75277-39-3	VWR
Acetic acid	1000632511	Merck
Acetonitrile	1000292500	Merck
Agarose	9012-36-6	Serva
Boric acid	1001625000	Merck
Bovine serum albumin (BSA)	9048-46-8	Sigma-Aldrich
Bromophenol blue	34725-61-6	Serva
Bis(sulfosuccinimidyl)suberate (BS3)	21585	Thermo Scientific
Dithiothreitol (DTT)	3483-12-3	Merck
Deoxynucleotide-5'-phosphate mix	L540.2	Roth
Ethanol	1009832511	Merck
Ethidium bromide	1239-45-8	Merck
Ethylendiamine tetraacetate (EDTA)	6381-92-6	Roth
Formic acid	64-18-6	VWR
Glycerol	1040922500	Merck
Guanidine hydrochloride	G3272-1KG	Sigma-Aldrich
Hydrochloric acid (37%)	1003172500	Merck
Iodacetamide	I6125-100G	Sigma-Aldrich
Magnesium chloride ($MgCl_2$)	M8266-1KG	Sgima-Aldrich
Methanol	1060092511	Merck
Non-fat dry milk powder	70166-500G	Sigma-Aldrich
Ortho-Phosphoric acid	1005731000	Merck
Polyethylene glycol 6000 (PEG 6000)	8074911000	Merck
Potassium chloride (KCl)	1049330500	Merck
Potassium dihydrogen phosphate (KH_2PO_4)	3904.1	Roth
Potassium hydrogen phosphate (K_2HPO_4)	T875.2	Roth
S-(5-adenosyl)-L-methionine (SAM)	A2408-25MG	Sigma-Aldrich
S-(5-adenosyl)-L-homocysteine (SAH)	A9384-10MG	Sigma-Aldrich
Sodium acetate	6773.2	Roth
Sodium chloride (NaCl)	1064045000	Merck
Sodium dodecyl sulfate (SDS)	1.13760.1000	VWR
Sodium hydroxide (NaOH)	1064621000	Merck
Sodium mercaptoethanesulfonate (MESNA)	M1511-25G	Sigma-Aldrich
Trifluoroacetic acid (TFA)	91707-250ML-M	Sigma-Aldrich
Tris(hydroxymethyl)amino ethane (Tris base)	4855.3	Roth
Triton X-100	1086431000	Merck
Tween-20	P1379-1L	Sigma-Aldrich
Urea	1084875000	Merck
Ubiquitin-vinyl-sulfone	U-212	Boston Biochem
Water	1153332500	Merck

Table 2.3: List of plasmids

Plasmids	Promoter	Resistance	Supplier
pET-22b_H2A	T7	ampicillin	Prof. Wolfgang Fischle
pET-22b_H2B	T7	ampicillin	Prof. Wolfgang Fischle
pET-22b_H3	T7	ampicillin	Prof. Wolfgang Fischle
pET-22b_H4	T7	ampicillin	Prof. Wolfgang Fischle
pET-22b_H2A E61/64A	T7	ampicillin	cloned
pET-3a_H3	T7	ampicillin	Prof. Wolfgang Fischle
pET-3a_H3Δ1-24A25C	T7	ampicillin	cloned
pTXB1_Ub	T7	ampicillin	Dr. Shira Albeck
pTXB1_H2AΔ113-129	T7	ampicillin	cloned
pUC18_12x200x601		ampicillin	Prof. Wolfgang Fischle
pUC18_52x187x601		ampicillin	Prof. Wolfgang Fischle
pUC18_16x145x601		ampicillin	Prof. Song Tab
pFB_DNMT1	polyhedrin	ampicillin, gentamycin	Prof. Albert Jeltsch
pFB_DNMT1ΔRFTS	polyhedrin	ampicillin, gentamycin	cloned
pFB_DNMT1ΔUIM	polyhedrin	ampicillin, gentamycin	cloned
pFB_Usp7	polyhedrin	ampicillin, gentamycin	Prof. Lori Frappier
pFB_Usp7 C223A	polyhedrin	ampicillin, gentamycin	cloned
pFB_Usp7ΔTRAF	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔMBT	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔRBR	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔDUF	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔlongSPM	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔpreSPM	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔSPM	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔN	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔC	polyhedrin	ampicillin, gentamycin	cloned
pFB_SCML2ΔRBR-DUF	polyhedrin	ampicillin, gentamycin	cloned
pGEX_SCML2	T7	ampicillin	cloned
pGEX_SCML2ΔMBT	T7	ampicillin	cloned
pGEX_SCML2ΔRBR	T7	ampicillin	cloned
pGEX_SCML2ΔDUF	T7	ampicillin	cloned
pGEX_SCML2ΔlongSPM	T7	ampicillin	cloned
pGEX_SCML2ΔpreSPM	T7	ampicillin	cloned
pGEX_SCML2ΔSPM	T7	ampicillin	cloned
pGEX_SCML2ΔN	T7	ampicillin	cloned
pGEX_SCML2ΔC	T7	ampicillin	cloned
pGEX_SCML2ΔRBR-DUF	T7	ampicillin	cloned
pGEX_RBR-DUF	T7	ampicillin	cloned
pGEX_TRAF	T7	ampicillin	cloned

Table 2.4: List of oligonucleotides

Oligonucleotide	Sequence	Supplier
*Eco*RI-btn-fwd	5' - GGGGGGGGATCCGGGGGGG[phos] - 3'	Sigma-Aldrich
*Eco*RI-btn-rev	5' - [phos]AATTCCCCCCCGGATCCCCCCCC[biotin] - 3'	Sigma-Aldrich
*Not*I-btn-fwd	5' - GGCCGGGGGGATCCGGGGG[phos] - 3'	Sigma-Aldrich
*Not*I-btn-rev	5' - [phos]CCCCCGGATCCCCCC[biotin] - 3'	Sigma-Aldrich

Table 2.5: List of antibodies

Antibody	Dilution	Ordering number	Supplier
rabbit-anti-H2B	1:1000	ab1790	Abcam
rabbit-anti-H3	1:5000	ab1791	Abcam
rabbit-anti-PCNA	1:500	sc-56 (PC10)	Santa Cruz
rabbit-anti-SCML2	1:500		Prof. Danny Reinberg
mouse-anti-SCML2	1:250	sc-271097 (F-7)	Santa Cruz
mouse-anti-UHRF1	1:1000	sc-373750 (H-8)	Santa Cruz
rabbit-anti-Usp7	1:1000	A300-033A-M	Bethyl
rabbit-anti-DNMT1	1:500	A300-041A-M	Bethyl
rabbit-anti-ACACA	1:250	C83B10	Cell Signalling Technology
swine-anti-rabbit	1:5000	P0399	Agilent
goat-anti-mouse	1:5000	P0447	Agilent
streptavidin-HRP	1:5000	P0397	Agilent

Table 2.6: Molecular biology enzymes

Enzyme	Ordering number	Supplier
*Ava*I	R0152L	New England Biolabs
*Bam*HI	R3136L	New England Biolabs
*Ban*I	R0118S	New England Biolabs
*Bfu*CI	R0636	New England Biolabs
*Dde*I	R0175L	New England Biolabs
*Dpn*I	R0176L	New England Biolabs
*Eco*RI	R3101L	New England Biolabs
*Hae*II	R0107L	New England Biolabs
*Not*I	R3189L	New England Biolabs
*Nde*I	R0111L	New England Biolabs
*Sau3*AI	R0169L	New England Biolabs
Antarctic phosphatase	M0289L	New England Biolabs
T4 DNA ligase	M0202T	New England Biolabs
*M.Sss*I methyltransferase	M0226L	New England Biolabs
Pfu polymerase	EP0501	Thermo Fischer Scientific
Ribonuclease A	9001-99-4	Sigma-Aldrich

Table 2.7: Kits and other reagents

Product	Ordering number	Supplier
Infusion Cloning Kit	639648	Takara Bio company
Gel extraction kit	740609.250	Macherey Nagel
Plasmid Purification kit	740588.250	Macherey Nagel
Gigaprep kit	740593	Macherey Nagel
Chitin resin	S6651L	New England Biolabs
Magnesphere streptavidin beads	Z5481	Promega
Glutathione agarose	16101	Thermo Fischer Scientific
GFP-trap resin	gta-10	ChromoTek
Dynabeads sheep-anti-rabbit	11203D	Thermo Fischer Scientific
Dynabeads goat-anti-mouse resin	11033	Thermo Fischer Scientific
TALON magnetic beads	635636	Takara Bio company
Anti-HA agarose	11815016001	Sigma-Aldrich
Q sepharose	17-5072-01	GE Healthcare
SP sepharose	17-5073-01	GE Healthcare
Superdex S200 10/300 GL	17-5175-01	GE Healthcare
HiTrap SP	17-1151-01	GE Healthcare
HiTrap Q	17-1153-01	GE Healthcare
Resource Q	17-1177-01	GE Healthcare
Reprosil-Pur120 C18 5μm resin	r15-aq	Dr. Maisch GmbH
Reverse phase C4	214TP54	Grace-Vydac
Amberlite	000020275	Sigma-Aldrich
NuPAGE gels	NP0321BOX	Thermo Fischer Scientific
NuPAGE loading dye	NP0007	Thermo Fischer Scientific
NuPAGE reducing agent	NP0004	Thermo Fischer Scientific
NuPAGE antioxidant	NP0005	Thermo Fischer Scientific
SeeBlue Plus2 Protein Standard	LC5925	Invitrogen
Isopropyl β-D-thiogalactoside	I6758-10G	Sigma-Aldrich
Silde-A-Lyser Mini	69552	Thermo Fischer Scientific
Spectrapor 3 Dialysis Membrane	132720	Spectrum Laboratories
X-Gal	15520034	Thermo Fischer Scientific
Ampicillin sodium salt	A9518-25G	Sigma-Aldrich
Gentamycin sulfate	G1914-250MG	Sigma-Aldrich
LB medium	X965.3	Roth
2xYT medium	X966.3	Roth
Ammonium bicarbonate	09830-1KG	Simga-Aldrich
ECL plus	RPN2132	Amersham GE Healthcare

2.3 Preparation of nuclear extracts

HeLa S3 light and SILAC-labelled (Arg-6, Lys-4) nuclear extracts to be used in mass spectrometry experiments were prepared by Dr. Miroslav Nikolov and Dr. Nadin Zimmerman using the Dignam protocol [167] and stored at 10 mg/mL concentration at -80 °C.

The Dignam protocol for preparation of fresh HeLa S3 nuclear extracts was modified to be used in western blot analysis. 20 mL packed cells were swollen in 100 mL hypotonic buffer (10 mM HEPES-NaOH pH 7.9, 10 mM NaCl) for 10 minutes on ice. Swollen cells were transferred in a glass homogeniser and lysed by manual douncing over 20 strokes. Lysed cells were centrifuged at 3000xg for 15 minutes at 4 °C to yield the cytosolic (supernatant) and the nuclear (pellet) fractions. The nuclear pellet was resuspended in 40 mL high salt buffer (25 mM HEPES-NaOH pH 7.9, 10% glycerol, 420 mM NaCl) and transfered to a glass homogeniser. Nuclei were lysed by manual douncing over 20 strokes. Lysed nuclei were incubated for 30 minutes at 4 °C at 200 rpm on a stirring plate. The nuclear extract was obtained after centrifugation of the lysed nuclei at 13000xg for 30 minutes at 4 °C. The nuclear extract was dialysed for 2 hours at 4 °C against storage buffer (25 mM HEPES-NaOH pH 7.9, 10% glycerol, 150 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 0.1% Triton-X 100). The dialysed nuclear extract was frozen in liquid nitrogen and stored at -80 °C.

2.4 Molecular cloning

Site-directed mutagenesis was performed in order to introduce amino acid mutations, insert short tags (His- and HA-tags) and for deletion of domains in the SCML2 insect cell and bacteria vectors. For site-directed mutagenesis, oligonucleotides were designed according to the quick change protocol [168]. In addition, oligonucleotides with longer complementary regions to the vector backbone were prepared for the mutants where the quick change protocol did not yield any clones. Mutants were obtained by polymerase chain reaction (PCR) (Table 2.8; Table 2.9). 20 μL from the PCR reactions were digested with *Dpn*I for 1 hour at 37 °C to remove the methylated parental plasmid. The digestion reaction was transformed into the chemically competent XL-1 Blue *Escherichia coli* bacteria. Positive colonies were screened by colony PCR and sequenced using primers against the vector backbone.

GST fusion constructs were prepared by infusion cloning according to the manufacturer protocol [169]. pGEX.6P1 plasmid was linearised with *Eco*RI and purified from an agarose gel using a gel extraction kit. Inserts were amplified by PCR and contained 5' and 3' overhangs complementary to the linearised plasmid (Table 2.10). After PCR, the insert was purified from an agarose gel. In the infusion (homologous recombination) reaction, 50 ng linearised plasmid was incubated with an equimolar amount of insert DNA in the presence of 1 μL of infusion cloning mix for 15 minutes at 50 °C. The reaction was then placed on ice for 2 minutes and transformed into the chemically competent XL-1 Blue *Escherichia coli* strain. Positive colonies were screened by colony PCR and further sequenced using primers against the pGEX vector backbone.

The H2AΔ113-129-intein construct was prepared by fusion PCR. Synthesis of the insert required three separate PCR steps. First, H2AΔ113-129 was amplified from pET-3a_H2A plasmid with a 3' end extension complementary to the 5' region of the GyrA intein sequence. Second, GyrA_CBD was amplified from pTXB1_Ub plasmid with a 5' extension end complementary to the 3' region of truncated H2A histone. Third, the forward primer of the first PCR and the reverse primer of the second PCR where used to amplify a full-length construct starting from an equimolar mix of the previous two PCR products. The final PCR product was digested with *Nde*I and *Bam*HI, purified by gel extraction and incubated in a three-fold excess molar ratio with a pTXB1 linearised vector backbone. The mixture was ligated with T4 DNA ligase for 1 hour at room temperature and transformed into the chemically competent XL-1 *Escherichia coli* strain. Positive colonies were screened by colony PCR and further sequenced using primers against the pTXB1 vector backbone.

Table 2.8: *Pfu* PCR reactions mix

Reagent	Concentration	Volume [μL]
Pfu buffer with $MgCl_2$	10x	5
dNTPs mix	10 mM	1
primer mix	10 μM	1
DNA template	x ng/μL	x
Pfu Polymerase	2.5U/μL	1
H_2O		42 - x
Total		50

Table 2.9: Site-directed mutagenesis thermomixer setup

Cycles	Process	Temperature [°C]	Time [sec]
1x	denaturation	95	180
35x	denaturation	95	30
	annealing	60	30
	elongation	72	480
1x	elongation	72	600

Table 2.10: *Pfu* PCR thermomixer setup

Cycles	Process	Temperature [°C]	Time [sec]
1x	denaturation	95	180
35x	denaturation	95	30
	annealing	55-65	30
	elongation	72	60
1x	elongation	72	600

2.5 Preparation of recombinant proteins

2.5.1 Purification of human histones from bacteria

pET-3a plasmids containing *Xenopus laevis* histone expression constructs or pET-22b plasmids containing human histone expression constructs were transformed by heat shock in BL21 (DE3)-RIL *Escherichia coli*. Single colonies were inoculated in 100 mL 2xYT medium and grown overnight at 37 °C, at 140 revolutions per minute (rpm) shaking. The preculture was diluted 1:100 in 2L 2xYT medium and grown at 37 °C until an OD_{600} value of 0.4 - 0.6. Protein synthesis was induced with 0.5 mM IPTG. Bacteria were harvested three hours after the induction by centrifugation at 6000xg, for 12 minutes, at 4 °C. Bacteria were resuspended in lysis buffer (10 mL buffer/ 1L bacterial culture) (100 mM NaCl, 10 mM Tris-HCl pH 7.5) and lysed by passing them three times through a French press at 1000-1500 psi. The lysate was clarified by centrifugation at 25000xg for 30 minutes, at 4 °C. The presence of the expressed protein in the lysate pellet (inclusion bodies) or in the lysate supernatant was checked by SDS-PAGE.

Inclusion bodies were washed twice with wash buffer (20 mL buffer/ 1L culture) (100 mM NaCL, 10 mM Tris-HCl pH 7.5, 1% Triton X-100) until complete resuspension of the pellet and centrifuged at 25000xg for 30 minutes at 4 °C. The inclusion bodies were washed once more with lysis buffer to remove the detergent, centrifuged and stored at -20 °C. The clean inclusion bodies were solubilized in 1mL dimethylsulfoxide (DMSO). Proteins were extracted from the solubilised inclusion bodies in 30 mL unfolding buffer (7 M guanidinium hydrochloride, 10 mM Tris-HCl pH 7.5, 10 mM DTT) for 1 hour under stirring at room temperature. The extract was clarified by centrifugation at 25000xg for 30 minutes at 4 °C. An additional 10 mL unfolding buffer was added to the pellet to fully extract proteins from the inclusion bodies. The pooled extract was dialysed three times against 2 L SAU 200 (7M urea, 200 mM NaCl, 20 mM NaAc pH 5.2, 2mM DTT, 1 mM EDTA), centrifuged and loaded onto a combination of Q and SP sepharose anion-cation exchange columns.

Expressed histones which remained in the supernatant of the bacterial lysate were recovered by acid extraction. Hydrochloric acid was added to a concentration of 0.5 M in the supernatant, which was stirred for 30 min at room temperature and centrifuged at 25000xg for 30 min at 4 °C. The supernatant containing the expressed histones was collected and incubated with 5 volumes of cold acetone overnight at -20 °C. The precipitated histones were collected by centrifugation and washed to complete resuspension in 1 volume of acetone. The acetone was removed by centrifugation and the protein pellet was air dried and stored at -20 °C. The protein pellet was solubilised by incubation with 1 mL DMSO for 30 minutes at room temperature. Histones were extracted from the solubilised pellet in an unfolding buffer (7 M guanidinium hydrochloride, 10 mM Tris-HCl pH 7.5, 10 mM DTT) under stirring, for 1 hour at room temperature. The histone extract was dialysed three times against 2 L SAU 200, centrifuged to remove aggregates and applied onto a combination of Q-Sepharose and SP-Sepharose columns.

After loading of the histone extracts, the Q column was disconnected. Unbound proteins

were removed by washing with 2 column volumes (CV) of SAU 200 buffer. Histones were eluted with SAU 600 buffer (7M urea, 600 mM NaCl, 20 mM NaAc pH 5.2, 2mM DTT, 1 mM EDTA) (0.5 CV 30% SAU 600 followed by a 2 CV 30-100% SAU 200 - SAU 600 gradient). Peak fractions were collected, dialysed thrice against 2 L 2 mM DTT, lyophilsed and stored at -80 °C. Protein purity was inspected using SDS-PAGE and ESI-MS. The urea used in the buffers SAU 200 and SAU 600 was deionised by passing it three times through a 3 L amberlite column.

2.5.2 Purification of C-terminal thioester constructs

Ubiquitin thioester

pTXB1.Ub plasmid was transformed into BL-21 chemically competent cells and single colonies were inoculated overnight at 37 °C in 100 mL 2xYT medium. The preculture was diluted 1:100 in 2 L 2xYT, brought overday to OD_{600} 0.6 - 0.8 at 37 °C, induced with 0.5 mM IPTG and cultured overnight at 18 °C and 140 rpm. The bacteria were pelleted by centrifugation at 6000xg for 12 minutes at 4 °C. Bacterial cells were lysed by passing them three times through the French press in lysis buffer (200 mM NaCl, 20 mM HEPES-NaOH pH 6.5, 1 mM EDTA). The lysate was pelleted by centrifugation at 25000xg for 45 minutes at 4 °C. The supernatant was collected and filtered through 0.8 μm filters before incubating overnight with pre-equilibrated (lysis buffer) chitin resin (10 mL beads slurry/2 L bacterial culture). The bound fraction was washed four times with 30 mL lysis buffer/ 10 mL resin by rotating the slurry for 30 minutes at 4 °C. Ubiquitin thioester was generated by incubating the chitin resin with cleavage buffer (20 mL buffer/10 mL resin) (200 mM NaCl, 100 mM MES-Na, 20 mM HEPES-NaOH pH 6.5, 1 mM EDTA) on a rotating wheel for up to one week at room temperature. Cleavage reactions were collected every 24 hours, centrifuged at 25000xg for 10 minutes and stored at 4 °C. The cleavage fractions were dialysed three times against 2 L deionized water and centrifuged to remove aggregates. The dialysed reaction soup was centrifuged at 25000xg for 10 minutes, lyophilised and resuspended in 5% acetonitrile (ACN) and 0.1% TFA. The concentrated ubiquitin thioester solution was purified by HPLC on a C4 column. The sample was applied onto the column with a 0.5 mL/min flow of 5% ACN and 0.1%TFA. Ubiquitin thioester was separated from the protein cleavage soup using a linear 90 minutes gradient of 5-95% ACN and 0.1% TFA. Pure fractions were dialysed three times against 0.5 L deionized water, lyophilised and stored at -80 °C. Protein purity was inspected by SDS-PAGE and ESI-MS.

H2AΔ113-129 thioester

The histone construct was expressed using the protocol described above for the preparation of the ubiquitin fusion construct. Bacteria were lysed by passing them three times through the French press in lysis buffer (200 mM NaCl, 20 mM HEPES pH 6.5, 1 mM EDTA), centrifuged and filtered. The filtered lysate was incubated with pre-equilibrated chitin resin (10 mL beads slurry/ 2 L bacterial culture) overnight at 4 °C on a rotating wheel. The bound fraction was washed four times with 30 mL lysis buffer/ 10 mL resin by rotating the slurry for 30 minutes at 4 °C. H2A thioester was obtained by incubating the chitin resin with cleavage

buffer (20 mL buffer/ 10 mL resin)(200 mM NaCl, 100 mM MES-Na, 20 mM HEPES pH 6.5, 1 mM EDTA) over 48-64 hours. Cleavage fractions were collected every 24 hours, dialysed three times against 2 L deionized water, centrifuged at 25000xg for 10 minutes to remove aggregates, lyophilised and stored at -20 °C. Lyophilised proteins were resuspended in 0.5 mL DMSO and incubated at room temperature for 30 minutes. 30 mL unfolding buffer (7 M guanidinium hydrochloride, 10 mM Tris pH 7.5) per 10L histone culture were added to unfold the proteins from the cleavage reaction. The unfolded polypeptides were dialysed three times against 2 L SAU 200 buffer (7 M urea, 20 mM NaAc pH 5.2, 200 mM NaCl, 1 mM EDTA). Dialysed proteins were centrifuged and loaded onto anion-cation exchange Q and SP sepharose columns. The proteins which bound to the SP-Sepharose column were eluted with a gradient of SAU 200 - SAU 600 (7 M urea, 20 mM NaAc pH 5.2, 600 mM NaCl, 1 mM EDTA) (0.5 column volumes with 30% SAU 600, then 2 column volumes with a linear 30-100% SAU200 - SAU 600 gradient). Peak fractions were collected and dialysed three times against 2 L deionized water, centrifuged at 25000xg for 10 minutes, lyophilised and stored at -80 °C. Protein purity was inspected by SDS-PAGE and ESI-MS.

2.5.3 Purification of GST-tagged proteins

pGEX plasmids containing full-length or truncated human SCML2 expression constructs were transformed by heat shock in BL21 (DE3)-RIL *Escherichia coli*. Single colonies were inoculated in 100 mL 2xYT medium and grown at 37 °C overnight on 140 rpm shake. The preculture was diluted 1:100 in 2 L 2xYT medium and grown at 37 °C until an OD_{600} value of 0.6 - 0.8. Expontential growth was stopped by incubating the cultures for 10 min on an ice bath. Protein synthesis was induced with 0.1 mM IPTG. Bacteria were grown overnight at 18 °C and harvested by centrifugation at 6000xg, for 15 minutes, at 4 °C. Bacteria were resuspended in lysis buffer (10 mL buffer/ 1 L bacterial culture) (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 2 mM DTT, 1 mM EDTA) and lysed by passing them three times through a French press at 1000-1500 psi. The lysate was clarified by centrifugation at 25000xg for 30 minutes, at 4 °C. The lysate supernatant was incubated overnight with pre-equilibrated glutathione agarose (1 mL resin/ 1 L bacterial culture). Unbound proteins were discarded and the beads were washed three times in wash buffer (750 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 2 mM DTT, 1 mM EDTA) (10 mL wash buffer/1 mL resin), washed once more in lysis buffer and eluted with 10 mL elution buffer (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 10 mM glutathione reduced, 2 mM DTT, 1 mM EDTA). The eluted proteins were dialysed for 2 hours against dialysis buffer I (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 2 mM DTT, 1 mM EDTA) and overnight against dialysis buffer II (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 50% glycerol, 2 mM DTT, 1 mM EDTA). Protein purity was inspected by SDS-PAGE. The recombinant proteins were aliquoted and stored at -20 °C.

To further increase purity, the proteins were run over a 1 mL SP anion-cation exchange column. Protein aliquouts were dialysed against SP Buffer A (250 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 2 mM DTT, 1 mM EDTA), loaded onto the SP column and eluted over 5 CV with a linear (0-100%) gradient with SP buffer B (750 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 2 mM DTT, 1 mM EDTA). Pure fractions were collected, dialysed

overnight against dialysis buffer II and stored at -20 °C.

GST-tagged full-length SCML2 and GST-tagged SCML2 truncations, as well as GST-tagged isolated domains were prepared according to this purification protocol.

2.5.4 Purification of His-tagged proteins from insect cells

Recombinant proteins were produced in insect cells using the Bac-to-Bac expression system [170]. Coding sequences of the recombinant proteins were cloned in pFastBacHT donor plasmids which were amplified in the electrocompetent DH10BacY *Escherichia coli* strain. Transformation efficiency into the competent cells was controlled using the gentamycin resistance on the donor plasmid. Transposition efficiency onto the baculovirus shuttle vector was controlled using blue-white colony selection. Bacteria were transformed with 1 μg plasmid DNA by electroporation (1 pulse, 25F, 1.8kV) using 0.1 cm electroporation cuvettes. After 1 minute incubation on ice, bacteria were resuspended in 1 mL LB medium and grown overnight at 37 °C on shake. Screening was performed by plating several ten-fold dilutions of the overgrown bacteria on LB-agar plates containing 20 μg/mL gentamycin, 150 μg/mL x-Gal and 1 mM IPTG. White colonies were restreaked on fresh plates. Five mL single-colony inoculates were grown overnight to obtain the recombinant bacmid, which was purified by isopropanol precipitation of the neutralised bacterial lysate. The lysate was incubated with 70% isopropanol overnight at -20 °C and the precipitated bacmid was collected by centrifugation at 13200xg at 4 °C for 30 minutes. The supernatant was removed and the bacmid pellet was washed twice (without resuspension) with 500 μL 70% ethanol. The clean bacmid was collected by centrifugation at 13200xg for 10 minutes at 4 °C and stored at -20 °C in 30 μL 70% ethanol.

The bacmid pellet was soaked for 15 minutes at room temperature in 20 μL deionised water. The dissolved DNA was diluted in 200 μL *Spodoptera frugiperda* Sf9 insect cell culture medium and incubated with 100 μL 1x transfection agent for 1 hour at room temperature. Three mL 1 million Sf9 insect cells/ mL were infected with 150 μL transfection-ready bacmid and incubated over 48 hours at 27 °C. Transfection efficiency was monitored using fluorescence microscopy which detected the expression of the YFP marker. The insect cell culture medium was collected and stored at 4 °C as v0 virus. Twenty-five mL 0.7 million Sf9 insect cells/ mL were infected with 500 μL vo virus and grown at 27 °C on 40 rpm shake in the dark. Cell viability, mean cell diameter and cell count were checked every 24 hours to monitor the progress of the viral infection. Cultures were diluted daily to 0.7 million cells/ mL until the day after proliferation arrest (DPA) when the viability dropped bellow 80%. Insect cells were pelleted 24 hours after the DPA at 300xg for 10 minutes at room temperature. Pellets were inspected for expression and solubility of the recombinant proteins and the culture media was collected and stored at 4 °C as v1 viruses. Two hundered mL 0.7 million *Trichoplusia ni* Hi5 insect cells/ mL were infected with 500 μL v1 virus and grown at 40 rpm shake at 27 °C in the dark. Cell viability, mean cell diameter and cell count were checked every 24 hours to monitor the progress of the viral infection. Insect cells were pelleted 48 hours after reaching the proliferation arrest. Insect cells were pelleted at 300xg for 20 minutes at 4 °C and dry pellets were frozen in liquid nitrogen, then stored at -80 °C.

Insect cell pellets were resuspended in lysis buffer (10 mL buffer/ 1 L insect cells culture)(300 mM NaCl, 50 mM Tris-HCl pH 7.5, 30 mM imidazole pH 7.5, 10% glycerol) and lysed by sonication with 20% intensity in 0.5 seconds ON, 1.5 seconds OFF cycles for 3 minutes. The lysate was clarified by centrifugation at 25000xg for 30 minutes, at 4 °C. The lysate supernatant was incubated for 3 hours with pre-equilibrated Ni-NTA agarose beads (2 mL resin/ 600 mL bacterial culture). Unbound proteins were discarded and the beads were washed three times in wash buffer (750 mM NaCl, 50 mM Tris-HCl pH 7.5, 30 mM imidazole pH 7.5, 10% glycerol) (10 mL wash buffer/1 mL resin), one more time in lysis buffer and then eluted with 10 mL elution buffer (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 300 mM imidazole pH 7.5). Eluted proteins were dialysed for 2 hours against dialysis buffer I (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 10% glycerol, 2 mM DTT, 1 mM EDTA) and overnight against dialysis buffer II (300 mM NaCl, 50 mM Tris-HCl pH 7.5, 50% glycerol, 2 mM DTT, 1 mM EDTA). Protein purity was inspected by SDS-PAGE. Proteins were aliquoted and stored at -20 °C.

His-tagged full-length DNMT1, Usp7 and SCML2 as well as several His-tagged DNMT1, Usp7 and SCML2 truncations were prepared according to this purification protocol.

2.6 Preparation of biotinylated DNA templates

Mononucleosomal DNA templates were prepared from plasmids containing repeats of the 601 positioning sequence: pUC18_16X145 and pUC18_52x187. Nucleosomal array templates were prepared from a pUC18_12x200 vector. Plasmids were transformed by heat shock in methylation deficient (dam-/dcm-) C2925 *Escherichia coli*, which were plated on 100 mg/mL ampicillin resistance plates. Single colonies were inoculated in 100 mL LB medium and grown overday at 37 °C. Precultures were diluted 1:100 into 2 L LB each and grown overnight at 37 °C under 140 rpm shake. Bacteria were centrifuged at 6000xg, for 12 minutes, at 4 °C and plasmid DNA was purified from the pellet using a DNA GigaPrep extraction kit. DNA was stored in 10 mM Tris-HCl pH 8.0 at -20 °C.

Mononucleosomal templates were obtained from the plasmid DNA by digesting the repeats into monomeric units using the *Ava*I and *Not*I restriction enzymes. Nuclesomal array templates were obtained by digesting the plasmid backbone into short fragments (leaving the inserts intact) using the *Eco*RI, *Dde*I, *Bfu*CI and *Hae*II restriction enzyme mix. Generally, 1 mg plasmid DNA was incubated overnight with 300 units of each of the restriction enzymes. Complete digestions were checked by agarose gel electrophoresis. Separation of the monomeric nucleosomal DNA from the intact backbone or of the long chromatin templates from the plasmid backbone fragments was achieved by polyethyleneglycol (PEG) precipitation [171].

Digestion reactions were stopped by heat-inactivation of the enzymes at 65 °C for 10 minutes. The digest mixtures were centrifuged for 10 minutes at 13200xg, at 4 °C to remove protein aggregates. Supernatants were transferred to new microreaction tubes, where the starting

salt concentration of the precipitation mix was adjusted to 0.5 M NaCl. The precipitation mix was pre-incubated at 37 °C for 30 minutes. PEG 6000 was added drop-wisely to the reaction. The starting PEG concentration was 5% (weight/volume). The mix was incubated for 15 minutes at room temperature and centrifuged for 10 minutes, at room temperature, at 13200xg. The resulting supernatant was transferred to a new tube, where additional PEG was added in consecutive steps to reach 6, 7 or 8 % PEG concentrations. The pellets of all these fractions were cleaned with 70% ethanol by brief vortexing followed by a 5 minutes, 13200xg room temperature centrifugation step. The clean pellets were air-dried and then resuspended in 10 mM Tris.HCl pH 8.0. The quality of the DNA preparation was assessed by agarose gel electrophoresis.

DNA templates were dephosphorylated at their 5' ends using 0.1 units/ μg DNA Antarctic Phosphatase. The dephosphorylation reaction was stopped by heat-inactivation of the enzyme at 65 °C for 10 minutes. Biotinylated oligos complementary to the *Eco*RI or *Not*I digestion sites were ordered with phosphorylated 3' ends. A ten-fold molar excess of the biotinylated oligonucleotides was incubated with the dephosphorylated template DNA, in the presence of T4 DNA ligase (10 units/ng DNA) in the appropriate buffer for 2 hours at room temperature. The ligation reactions were supplemented with 1 mM ATP and incubated for 1 hour more at room temperature. The reactions were stopped by heat-inactivaton of the ligase at 65 °C for 10 minutes. The DNA templates were purified from the unligated oligonucleotides by PEG precipitation. DNA was stored in 10 mM Tris-HCl pH 8.0 at -20 °C. Biotinylation efficiency of nucleosome and chromatin DNA templates was checked by agarose gel electrophoresis and by dot blot, using a Streptavidin-HRP conjugate probe.

2.7 Histone octamer assembly

Lyophilised histones were dissolved in 1-2 mL unfolding buffer (7 M guanidinium hydrochloride, 10 mM Tris-HCl pH 7.5, 10 mM DTT) for 1 hour on the rotating wheel at 4 °C. Histone concentrations were calculated based on their theoretical molecular weights and extinction coefficients. Histones H2A, H2B, H3 and H4 were mixed in equimolar amounts and diluted to a 1 mg/mL with unfolding buffer. The histone octamer mixture was dialysed three times against 2 L RB high buffer (2 M NaCl, 10 mM Tris-HCl pH 7.5, 2 mM DTT, 1 mM EDTA). Histone octamers were collected from the dialysis bag and centrifuged for 10 minutes at 13200xg, at 4 °C to remove aggregates. The supernatant was loaded onto a Superdex 200 gel filtration column. Protein complexes of various molecular weights were eluted using 1 CV of RB high buffer. Protein fractions were collected and checked by SDS-PAGE for equimolar amounts of histones. Octamer-containing fractions were pooled and stored in 50% glycerol at -20 °C or used directly for reconstitution of chromatin or mononucleosomes.

2.8 Mononucleosome and chromatin reconstitution

To assemble mononucleosomes and nucleosome arrays, histone octamers were mixed with the positioning DNA templates at an optimal molar ratio. The ratio was determined in small

scale reactions by titrating increasing histone octamer amounts to constant DNA amounts. The reconstitution octamer:DNA ratios ranged between 0.7:1 and 1.5:1. The starting salt concentration of the DNA fragments was adjusted to 2 M NaCl. Histone octamers were used fresh (in 2M NaCl) or dialysed overnight at 4 °C from the 50 % glycerol stocks against 500 mL RB high (2 M NaCl, 10 mM Tris-HCl pH 7.5, 2 mM DTT, 1 mM EDTA). Dialysed octamers were concentrated by ultrafiltration at 13200xg at 4 °C to reach a concentration of 0.5 - 1 mg/mL. Reconstitution reactions were set up at 4 °C in 500 μL dialysis tubes with 3500 Da molecular weight cutoff. The dialysis tubes were set on a floating disk in a chamber filled with 500 mL RB high buffer. The salt concentration in the dialysis chamber was decreased under continuous stirring over 36 hours by a peristaltic pump which exchanged the 0.5 L RB high for 2 L RB low (20 mM NaCl, 10 mM Tris-HCl pH 7.5, 2 mM DTT, 1 mM EDTA). Reconstituted templates were collected from the dialysis buttons and stored at 4 °C. Reconstitution efficiency was checked by electrophoresis on 0.5% (chromatin) or 1% (nucleosomes) agarose gels, which were run in 0.2x TB (0.178 M Tris, 0.178 M boric acid) buffer at 4 °C, for 2 hours at 100 V. Gels were stained after the run with 1 μg/ mL ethidium bromide (EtBr) for 15 minutes and destained in deionised water for further 15 minutes.

2.8.1 Quality control of reconstituted chromatin

To test nucleosomal positioning, 500 ng chromatin was incubated with 0.5 units restriction enzymes (*Ava*I; *Not*I; *Ban*I) and the appropriate buffer in 20 μL reactions. Digestion proceeded at 37 °C for 2 hours. The digested DNA products were loaded on a 1.5% native agarose gel which preserved histone:DNA interactions. The gel was run at 120 V for 1 hour and stained after the run in EtBr.

To test nucleosomal occupancy, 2.5 μg chromatin was incubated with 4 miliunits/μL micrococcal nuclease in a 500 μL reaction. The reaction was performed in a reaction buffer containing 10 mM Tris-HCl pH 7.5, 5 mM NaCl and 2 mM $CaCl_2$ at room temperature. One hundret μL aliquots were collected from the digestion reaction after 0, 10, 30 and 60 seconds. Full digestion to mononucleusomal product was achieved after 30 minutes. The digestion reactions were stopped by mixing them with an equal volume of binding buffer from a PCR cleanup kit. Digests were purified on spin columns and eluted in 20 μL water. Digests were run on 1.5% native agarose gels at 120 V for 1 hour and stained after the run in EtBr.

2.9 Interaction experiments

2.9.1 Histone affinity purification

Magnetic goat-anti-rabbit beads were used to capture a rabbit antibody which recongnises the C-terminus of histone H3. In an affinity purification reaction 10 μL pre-equilibrated beads were mixed with 1 μg antibody and 10 μg histone in the presence of 1 mg/mL BSA in pull-down buffer (10 mM HEPES-NaOH pH 7.9, 150 mM NaCl, 5% glycerol, 1 mM DTT, 0.5 mM EDTA, 0.5 mM PMSF, 0.1% Triton X-100). Immobilisation was performed overnight

at 4 °C on the rotating wheel. Unbound material was removed by washing with pull-down buffer three times, over 5 minutes, on the rotating wheel. 75 pmol recombinant proteins or 2 mg HeLa nuclear extract was added to each pull-down reaction and allowed to incubate for 1 hour. Unbound nuclear extract was removed by washing with pull-down buffer three times, over 5 minutes, on the rotating wheel. The beads from each reaction were collected by centrifugation at 3000xg for 3 min at 4 °C and resuspended in SDS-PAGE loading buffer. Samples were incubated for 10 minutes at 70 °C and loaded on gradient 4-12% SDS-PAGE gels. In the experiments using recombinant proteins, detection was achieved by Coomassie staining of the gels. In the experiments where nuclear extracts were used, bound proteins were transferred from the gels to PVDF membranes with constant 2.3 A for 8 minutes using a semidry transfer system. The PVDF membrane was blocked for 1 hour in 5% milk in PBST at room temperature and incubated overnight at 4 °C with the appropriate antibodies. Unbound primary antibodies were removed with three 5 minutes washes of PBST. The blots were incubated with appropriate secondary antibodies for 1 hour at room temperature. Unbound secondary antibodies were removed with three 5 minutes washes of PBST. Proteins recruited in the pull-downs were detected using the ECL chemiluminescent substrate.

Magnetic HisPur beads were used to capture the His.H2B and His.H2BK120ub constructs. An adjusted pull-down buffer was used for this experiment (10 mM HEPES-NaOH pH 7.9, 150 mM NaCl, 5% glycerol, 0.5 mM PMSF, 0.1% Triton X-100).

Anti-HA affinity matrix was used for immobilization of the HA.H2B and HA.H2BK34ub constructs. The experiment was performed with unchanged buffer conditions.

2.9.2 Chromatin affinity purification

Streptavidin paramagnetic beads were used for immobilisation of the biotinylated chromatin and mononucleosome templates. Affinity purification reactions were performed with recombinant proteins or with HeLa nuclear extracts. Bound proteins were analysed by Coomassie staining, by western blot or by mass spectrometry.

In the experiments where recombinant proteins were used, 20 μL pre-equilibrated streptavidin beads were incubated with 1 μg chromatin and 75 pmol recombinant proteins and 1 mg/mL BSA in pull-down buffer (10 mM HEPES-NaOH pH 7.9, 150 mM NaCl, 5% glycerol, 1 mM DTT, 0.5 mM EDTA, 0.5 mM PMSF, 0.1% Triton X-100) for one hour at 4°C on the rotating wheel. The chromatin affinity purification protocol used the washing and elution steps described above for the histone affinity purification protocol.

For western blot detection, 20 μL pre-equilbrated streptavidin beads were incubated with 2 μg chromatin and 1 mg/mL BSA in pull-down buffer overnight, at 4°C, on the rotating wheel. Unbound chromatin was removed by washing with pull-down buffer three times, over 5 minutes, on the rotating wheel. 2 mg HeLa nuclear extract was added to each pull-down reaction and allowed to incubate for 1 hour. Washing and elution steps were performed as described before for the histone affinity purification protocol. Proteins were transferred onto PVDF membranes using semi-dry transfer. Western blot detection was performed as

described before for the histone affinity purification protocol.

For mass spectrometry detection, 100 μL pre-equilibrated streptavidin beads were incubated with 20 μg chromatin and 1 mg/mL BSA in pull-down buffer overnight, at 4 °C, on the rotating wheel. Unbound chromatin was removed by washing with pull-down buffer three times, over 5 minutes on the rotating wheel. Ten mg unlabeled or SILAC-labeled (Arg-6, Lys-4) HeLa nuclear extracts were added to either the unmodified or the modified chromatin samples in SILAC label-swap experiments [132]. Unbound nuclear extract was removed by washing with pull-down buffer three times, over 5 minutes, on the rotating wheel. The beads from each reaction were collected dry and resuspended in NuPAGE loading buffer. Samples were incubated for 10 minutes at 70 °C and loaded on 4-12% gradient NuPAGE gels.

For detection of crosslinked peptides by mass spectrometry, 20 μL pre-equilibrated streptavidin beads were incubated with 4 μg unmodified chromatin, 75 pmol recombinan full-length His.Usp7 and GST.RBR-DUF and 1 mg/mL BSA in pull-down buffer for one hour, at 4°C, on the rotating wheel. Washing and elution steps were performed as described before for the histone affinity purification protocol. Chromatin was digested to mononucleosomes with 100 units *Ava*I. Intra- and inter-molecular interactions were stabilised with 1 mM Bis(sulfosuccinimidyl)suberate S3 (dissolved in pull-down buffer) treatment. Crosslinking reactions were performed in a thermomixer for 30 minutes at room temperature on 300 rpm shake and stopped by boiling for 10 minutes at 95 °C. The crosslinked products were separated on 4-12% gradient NuPAGE gels.

2.9.3 GST and YFP affinity purification

Glutathione agarose resin was used to immoblise GST-tagged full-length or trunctated SCML2 constructs in affinity purification (protein-protein interaction) experiments. GFP-trap (agarose conjugated with an antibody raised against GFP) was used to immobilise YFP-tagged FL or truncated DNMT1 constructs. 20 μL pre-equilibrated glutathione resin or GFP-trap resin were incubated with 75 pmol recombinant proteins and 1 mg/mL BSA in pull-down buffer (same as the one used in the chromatin pull-down experiments) for one hour, at 4 °C, on the rotating wheel. Washing and elution steps were performed as described before for the histone and chromatin pull-down experiments. The bound proteins were eluted by incubation for 10 minutes at 70 °C, loaded on 4-12% gradient NuPAGE gels and visualised by Coomassie staining of the gels.

2.9.4 Gel shift assays

To map the interaction between SCML2 and the nucleosome, 15 pmol unmodified nucleosomes (containing 147, 171 or 187 bp of template DNA) were incubated with full-length or truncated SCML2 proteins. SCML2 was either kept stoichiometric to the nucleosomes or was added in excess to the nucleosomes (three-fold titration series). SCML2 and the nucleosome were incubated in a thermomixer at 16 °C, at 300 rpm for 1 hour. SCML2-nucleosome complexes were separated on 1% agarose gels at 120 V, for 1 hour and visualised by EtBr staining of the gel after the run.

2.9.5 Deubiquitylation experiments

Three hundret ng H3K18ub chromatin or mononucleosomes were incubated with 0.75 pmol Usp7 and increasing concentrations of full-length or truncated SCML2 or DNMT1 constructs. Titration series were performed in five-fold steps. For deubiquitylation of histones, 30 ng H3K18ub histone were incubated with 0.08 pmol Usp7 in the presence of increasing amounts of full length SCML2. 145 bp unmethylated or fully methylated DNA was added to the histone deubiquitylation experiment in a equimolar ratio to the starting SCML2 concentration.

2.10 Mass spectrometry analysis

2.10.1 Peptide preparation

Peptides were prepared according to published protocols [172]. To reduce sample complexity in the pull-down reactions, each lane was sliced into 2x13 bands which were chopped down with a scalpel into 1x1 mm^2 pieces. Unless otherwise stated, all incubation steps were performed in a thermomixer at 800 rpm, for 15 minutes, at 26 °C. Gel pieces were washed for 5 minutes in 300 μL deionised water and centrifuged at 3000 rpm for 1 minute. The supernatant was removed and the gel pieces were dehydrated by incubating with 300 μL acetonitrile. Acetonitrile was removed by pipetting and the gel pieces were completely dried in a vacuum concentrator. To reduce disulfide bridges, the gel pieces were incubated with 200 μL 10 mM DTT for 50 minutes at 56 °C. The pieces were centrifuged and the reducing solution was removed. Gel pieces were dehydrated further with 300 μL acetonitrile. The acetonitrile was removed and the gel pieces were incubated for 20 minutes with 200 μL 55 mM iodacetamide to alkylate the reduced cysteines. Gel pieces were incubated for further 15 minutes with 300 μL 100 mM ammonium bicarbonate. The alkylating solution was removed and the gel pieces were dehydrated twice in 300 μL acetonitrile. Gel pieces were completely dried in a vacuum concentrator. The dried gel pieces were rehydrated in 200 μL digestion buffer (42 mM ammonium bicarbonate, 4.2 mM $CaCl_2$, 0.125 units/μL trypsin. The digestion reaction was allowed to proceed overnight at 37 °C at 800 rpm. The digested peptides were centrifuged at 3000xg for 1 minute and 100 μL deionised water were added on top to cover the gel slices. 200 μL acetonitrile were added on top to shrink the gel pieces at 37 °C for 15 minutes at 800 rpm. The extract was collected into a new microreaction tube. 100 μL 5% formic acid was added to the shrunk gel pieces and incubated for 15 minutes at 37 °C at 800 rpm. The gel pieces were shrunk with 100 μL more acetonitrile for 15 minutes. The mixture was centrifuged at 300xg for 1 minute and added over the extract from the previous step. The mixed peptide extract was dried to completion by vacuum centrifugation. Peptides were stored dry at -20 °C. Before usage on the liquid chromatographer, the peptides were resuspended in 20 μL 5% acetonitrile and 0.1% formic acid, sonicated for 3 minutes in a water bath, centrifuged at 13200xg for 10 minutes and transferred to chromatography vials.

2.10.2 Liquid chromatography and mass spectrometry

The processed peptides were analyzed on a Q-Exactive Plus hybrid Quadrupole-Orbitrap mass spectrometer coupled to a NanoLC pump. Samples were loaded with an autosampler

and concentrated using a 10 μL/min flow of solvent A (5% acetonitrile, 0.1% formic acid) on a reversed phase C18 precolumn (0.15mm ID x 20mm self-packed with Reprosil-Pur 120 C18-AQ 5μm). Peptides were separated at 60 °C on a reversed phase nanoflow C18 column (0.075mm ID x 200mm self-packed with Reprosil-Pur 120 C18-AQ, 3μm). Peptides were eluted from the column under a 0.3 μL/min flow using a 5 - 44% gradient of solvent A solvent B (80% acetonitrile, 0.1% formic acid) over 90 minutes. Eluted peptides were ionised by electron spray ionisation in the positive ion mode. Full scan MS1 spectra were acquired in the 350 - 1600 m/z range at a resolution of 70000. The 20 most intense peaks from the survey scan were selected for fragmentation with Higher-energy Collisional Dissociation (HCD, 15 of normalised collision energy). Product ions MS2 spectra were acquired in the 200 - 2000 m/z range at a resolution of 15000. Forward and reverse experiment samples were measured in triplicates.

Raw files were imported into MaxQuant [173], [125], where they were grouped into forward and reverse experiments. Ion intensity quantitation was performed from the MS1 spectra. The ion multiplicity level was set to 2 and the heavy amino acid isotopes were indicated: Arg-6, Lys-4. The digestion mode was set to specific and trypsin was indicated as the chosen protease with two allowed missed cleavages per peptide. The minimum peptide length was set to 7 amino acids and razor peptides were included into the search algorithm. The peptide search database was generated from the reviewed human proteome [174]. The decoy search database was generated by reversing each polypeptide sequence from the human proteome database. Peptide sequencing was performed using the Andromeda-search engine incoporated in MaxQuant [124]. The false discovery rate for both the peptide spectrum match and the protein groups were set to 0.01. Carbamidomethylation of cysteine residues was set as fixed modification for all peptides. Acetylation of protein N-termini and oxidation of methionine residues were included in the search as variable modfiications.

2.10.3 Data analysis

MaxQuant output protein groups tables were imported as txt files into Perseus [175], [133]. Contaminant and hits from the decoy reverse database were manually removed from the protein groups list. Entries with less than 4 valid values (from the 6 normalised heavy/light abundance ratios) were excluded. Valid values were subjected either to one- or to two-sample student's t-tests. For the one sample test, the mean of all six measurements (combining the foward and the reverse experiment) was compared to the zero mean of the entire protein groups. For the two-sample t-test, the mean of the three forward experiment replicates was compared to the mean of the three reverse experiment replicates. A Perseus build in permutation based algorithm was used to control the identifications of the two tests. The algorithm's false discovery rate was 0.01 and its backrgound correction value S0 = 2. Perseus output tables were imported as csv files into R [176]. To obtain the interactome plots, heavy/light enrichment ratios from the forward experiment were plotted against the inverse of the enrichment ratios in the reverse experiment. Separately, to obtain volcano plots, the t-test statistic was plotted against the t-test difference to display the statistical relevance of the enriched interactors. Associations between the signifcantly enriched proteins within the different interactomes were displayed using an online protein-protein interaction platform [177].

For analysis of the crosslinking mass spectrometry data, raw files were converted to mgf format and imported into pLink [178], [140]. A database which contained the amino acid sequences of all proteins used in the crosslinking experiment was manually prepared. BS3 was indicated as a conventional crosslinker. Trypsin was selected as the protease used for digestion. The maximum number of missed cleavage sites was set to 3 sites and the peptide length limits were set between 6 and 60 amino acids. Acetylation of protein N-termini and oxidation of methionine residues were included in the search as variable modfiications. The false discovery rate of the search algorithm was set to 0.05. Identified crosslinks were accepted if they were present in a minimmum of two spectra. Inter- and intramolecular crosslinks were displayed on the primary amino acid sequences of the proteins present in the crosslinking experiment using xiNET [179], [180]. Crosslinkins were displayed on the nucleosome and Usp7 cystal structures using PyMOL [181].

Chapter 3

Results

3.1 Preparation of uniformly ubiquitylated nucleosomes and nucleosomal arrays

3.1.1 Preparation of unmodified human histones

Histone ubiquitylation marks have been found in organisms throughout all kingdoms of life. Considering the well-studied example of H2BK120ub, its functional relevance appears to be conserved from yeast to human. To prepare nucleosomal arrays, we needed to decide to use histones from a single organism. On one hand, the use of yeast as a model organism has been valuable in the early years of histone ubiquitylation because of the ease with which recombination techniques were used to manipulate the yeast genome. On the other hand, frog histones have traditionally been used in the field of chromatin biochemistry due to their convenient production in high amounts and purity. However, the most comprehensive annotation of histone ubiquitylation marks has been curated from human cells and tissues. With the advent of new genome editing technologies and the prospect of improved purification strategies, it became desirable to switch from the yeast and frog histones towards the human histones and use these as building blocks for the ubiquitylated nucleosomal arrays.

To prepare human histones two separate schemes were used. On one hand, the standard protocol for purification of recombinant frog histones from bacterial overexpressions was applied [182]. On the other hand, a protocol describing the extraction of histones from cultured human cells was used [183]. For standardisation, the two protocols were combined into a single scheme which takes advantage of the convenient overexpression of proteins in bacteria (Figure 3.1A). After centrifugation of the bacterial lysate, presence of histones in the lysate pellet or supernatant was inspected by SDS PAGE gel electrophoresis. Histones present in the pellet (inclusion bodies) were purified with the frog histone protocol. Histones found in the supernatant were purified with the acid (HCl) extraction protocol. Unmodified human histones H2A and H2B were prepared using acid extraction (Figure 3.1B). Unmodified human histones H3 and H4 were prepared from inclusion bodies. This scheme was applied also for preparation of tagged HA.H2B and His.H2B or N-terminally truncated H3 and H2A histones.

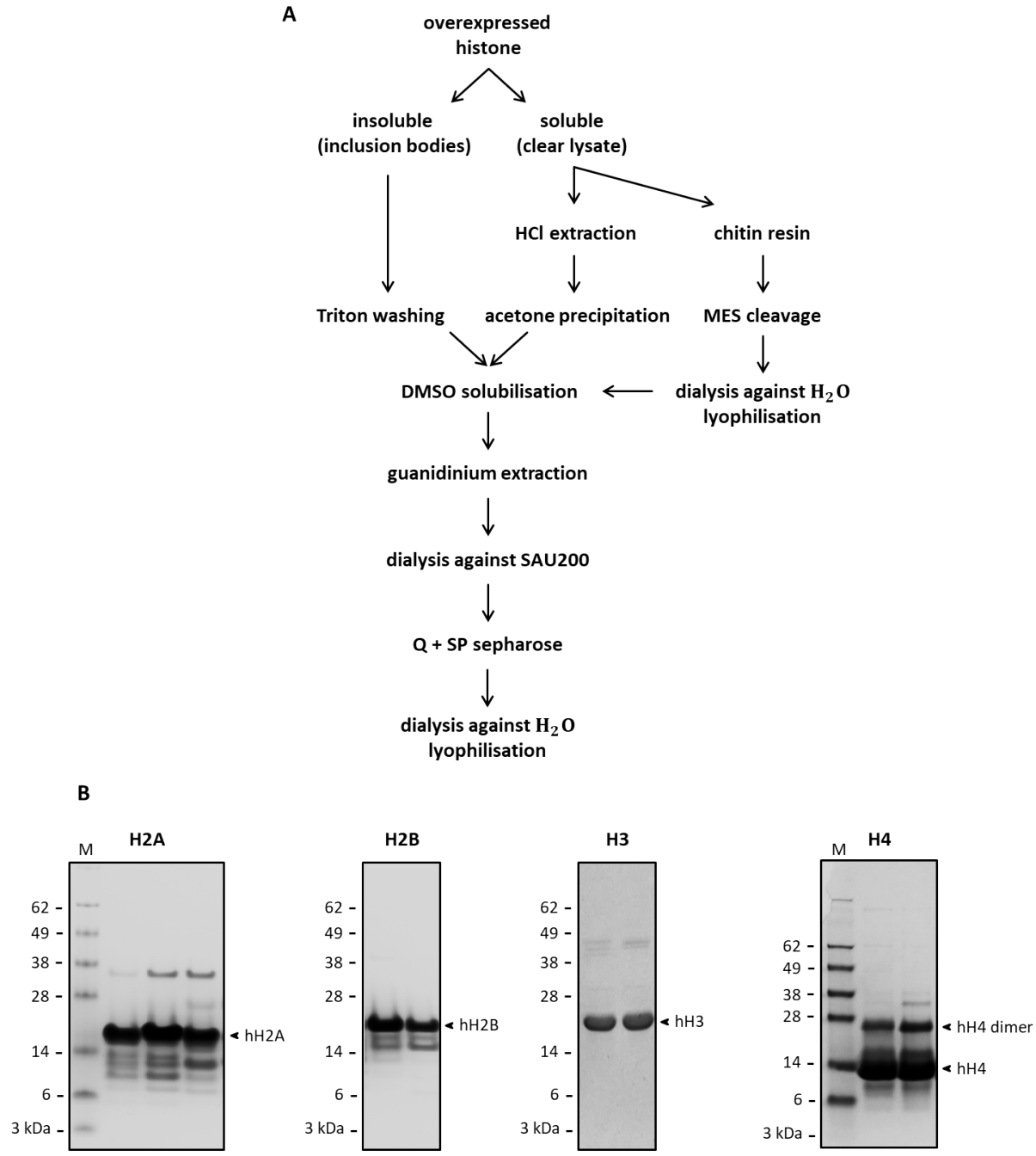


Figure 3.1: Preparation of recombinant human histones. (A) Schematic protocols used to purify recombinant human histones. Soluble proteins were purified by acid extraction, insoluble protein were prepared from inclusion bodies as per the standard published protocol [182]. Histone-intein fusion constructs were enriched on a chitin resin, making use of the C-terminal chitin binding domain (B) Coomassie-stained SDS-PAGE gels of purified unmodified human histones.

3.1.2 Preparation of ubiquitylated human histones

Expressed chemical ligation makes uses of the advantageous protein expression in bacterial cultures and the specificity of the chemical reaction between thiols and thioesters. Since only one human histone (H3.1) contains two cysteines - which have been mutated to serine and alanine here - this strategy could efficiently be applied for ligation of ubiquitin to histones.

Ubiquitylation of histone H3

Semi-synthesis of ubiquitylated H3 histone required three components. Recombinant truncated H3 histone and recombiannt ubiquitin thioester were prepared from bacteria. The synthetic peptide was made in the laboratory of Prof. Ashraf Brik (Figure 3.4).

Purification of H3 Δ1-24 A25C was performed according to the standardised histone purification protocol described above. Two protein peaks were recorded in the elution from the SP

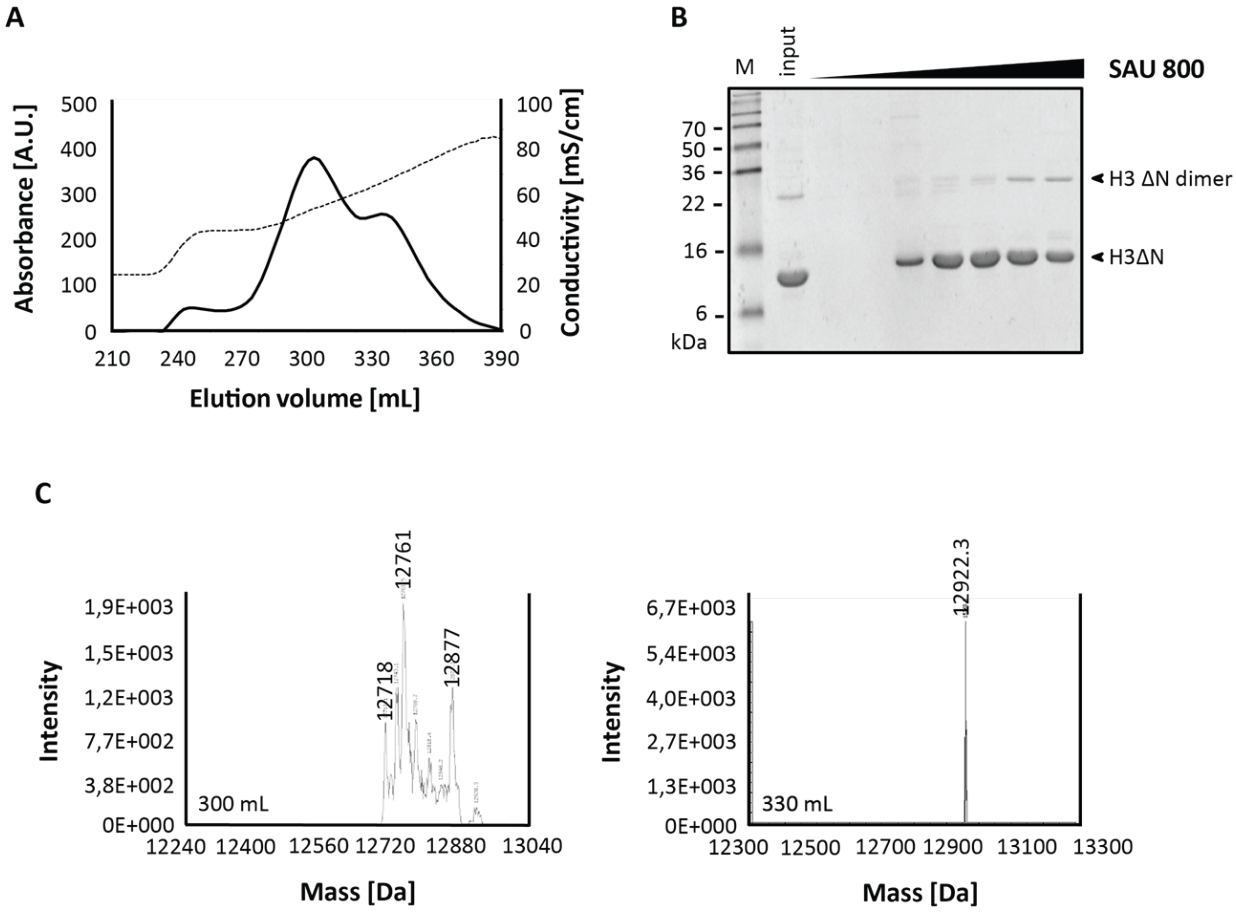


Figure 3.2: Preparation of recombinant truncated H3. (A) Purification profile of the truncated histone collected from the SP sepharose column. (B) Coomassie-stained SDS-PAGE gel of fractions collected from the SP-sepharose column. (C) Electron-spray ionisation mass spectrometry (ESI-MS) analysis of the first and second protein peaks (fractions collected at 300 and 330 mL, respectively). M = molecular weight marker

sepharose column (Figure 3.2A). SDS-PAGE gel electrophoresis indicated that both peaks contained proteins of similar molecular weight (Figure 3.2B). For clarification of the identity of the proteins in the two peaks, aliquots from both fractions were measured by native electron spray ionisation mass spectrometry (ESI-MS)(Figure 3.2C). The sample collected at 300 mL contained a mixture of ions with molecular weight (MW) values between 12718 Da and 12877 Da. The sample collected at 330 mL contained a single ion with a MW of 12922 Da. The theoretical MW of the histone fragment with cysteine on the first position was 12790 Da. The theoretical MW of the histone fragment including methionine on the first position was 12922 Da. The protein collected from the first peak contained thus histone suitable for native chemical ligation, whose N-terminal cysteine had however suffered modifications during expression in the bacterial host. To activate the cysteine, the recombinant histone collected from the peak eluting at 300 mL was treated with methoxylamine in the laboratory of Prof. Ashraf Brik.

Preparation of ubiquitin thioester was adapted from a protocol obtained from the laboratory of Shira Albeck from the Weizmann Institute of Science and Technology in Rehovot, Israel.

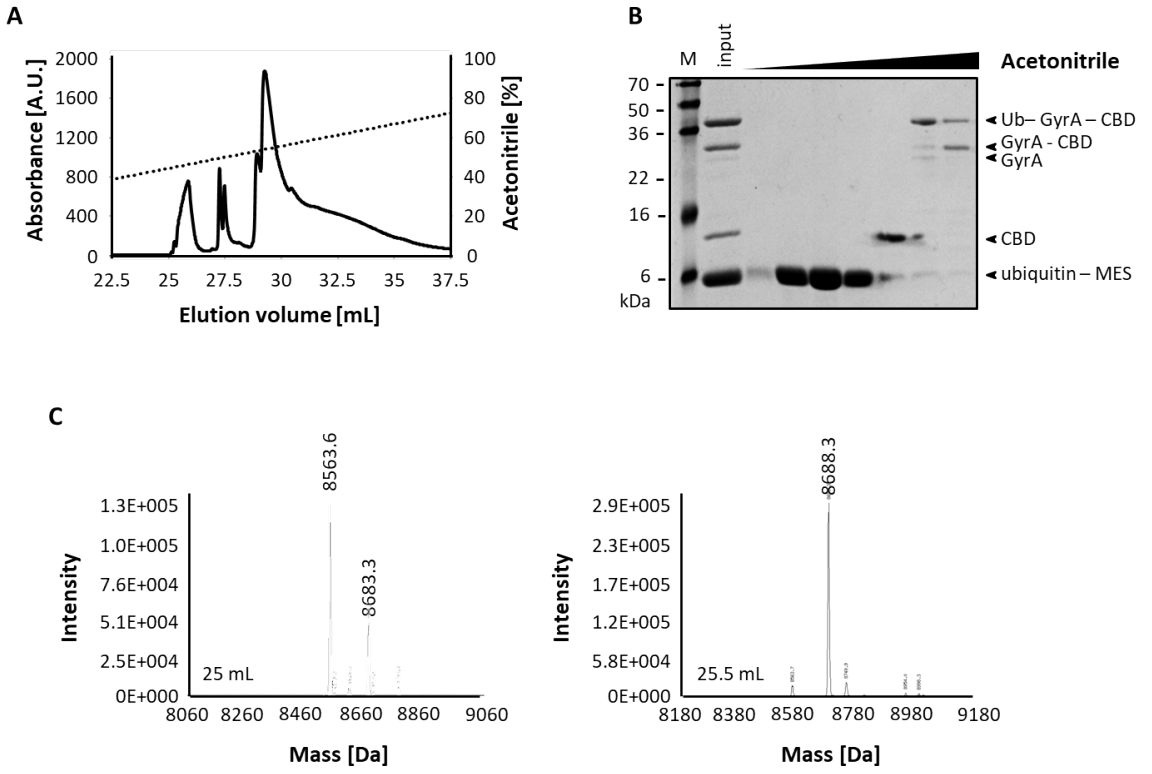


Figure 3.3: Preparation of recombinant ubiquitin thioester. (A) Purification profile of ubiquitin from the GyrA intein cleavage reaction collected from a C4 high performance liquid chromatography (HPLC) column. (B) Coomassie-stained SDS-PAGE gel of representative HPLC fractions collected from the C4 column. (C) ESI-MS analysis of the fractions containing ubiquitin (8563.6 Da) and ubiquitin thioester (8688.3 Da). CBD = chitin binding domain; MES = mercaptoethanesulfonate. M = molecular weight marker

Ubiquitin was expressed as a fusion construct containing an intein derived from the *E. coli* DNA Gyrase A (GyrA) and a chitin binding domain (CBD) used as affinity tag. The fusion construct was affinity purified using chitin beads and treated with sodium meracaptoethane-sulfonate (MES-Na) to trigger a intein-catalysed cleavage reaction (Figure 1.7B; Figure 3.3). The cleavage reaction was separated on a high performance liquid chromatography (HPLC) column, which yielded several protein peaks at 214 nm (Figure 3.3A). Ubiquitin was purified from the fractions eluting between 25 and 26.5 mL (Figure 3.3B). ESI-MS analysis indicated that the 25 mL aliquout contained ubiquitin without the C-terminal MES thioester (8563 Da) (Figure 3.3C). The remaining fractions contained intact ubiquitin thioester (8688 Da) and were used in native chemical ligation reactions in the laboratory of Prof. Ashraf Brik.

The synthetic peptide used in the native chemical ligation reaction containted on positions K18, K23 or both on K18 and K23 protected thiollysines (Figure 3.4). Two separate ligation reactions were performed to connect the three fragments. First, the truncated histone was reacted with the synthetic peptide. Second, after thiollysine deprotection, the full-length histone was reacted with the ubiquitin thioester. The cysteine inserted on position 25 to aid the ligation reaction was converted to alanine. Ubiquitylated histones were purified from the ligation reactions by HPLC. The masses of the pure constructs, as measured with ESI-MS, were 23768, 23768 and 32314 Da for H3K18ub, H3K23ub and K3K18/23ub2, respectively.

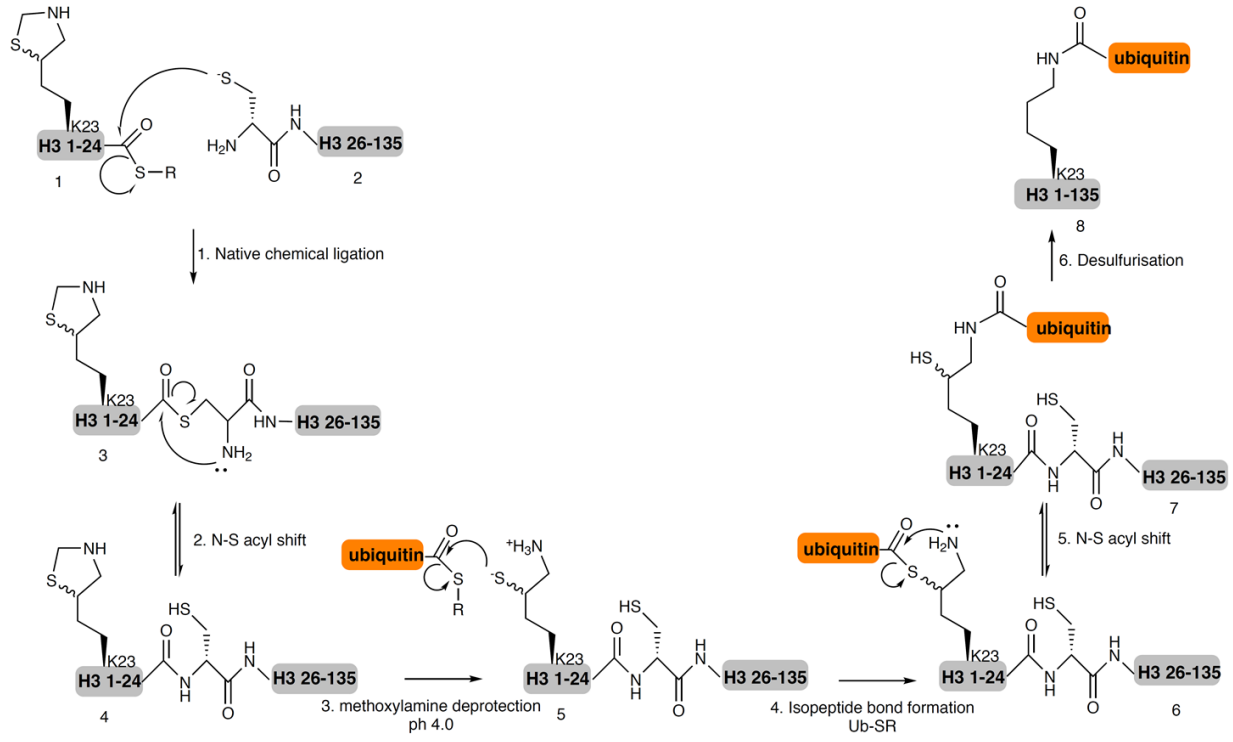


Figure 3.4: Schematic representation of the synthesis of H3K23ub by native chemical ligation. Synthetic H3 peptide and subsequent native chemical ligation reactions were performed in the laboratory of Prof. Ashraf Brik at the Technion in Hiafa, Israel.

Using this strategy, in collaboration with Prof Ashraf Brik's laboratory, milligrams of ubiquitylated histones with native isopeptide linkage were produced for the three H3ub analogs.

Ubiquitylation of histone H2A

In parallel to the production of ubiquitylated H3, truncations were made for other histone constructs: the N-terminus of histone H2A (Δ1-20), the C-terminus of histone H2A (Δ113-129) and the C-terminus of histone H4 (Δ89-102). These were prepared in order to attach ubiquitin to H2A at positions K13 or K15 and K119 as well as to H4 at position K91.

Histone H2A Δ1-20 A21C was extracted from the bacterial lysate according to the protocol used to prepare unmodified human H2A. The N-terminal truncation of H2A presented, similarly to H3 Δ1-24 A25C (Figure 3.2A), two peaks in the eluate collected from the SP sepharose column. The protein collected from the second peak contained the N-terminal methionine (MW = 12099 Da). The protein collected from the first peak had a chemically modified cysteine at the N-terminus, which needed subsequent activation for chemical ligation.

Purification of the H2A and H4 C-terminal truncations was performed on a scheme combining elements from the protocols used to prepare ubiquitin thioester and unmodified human

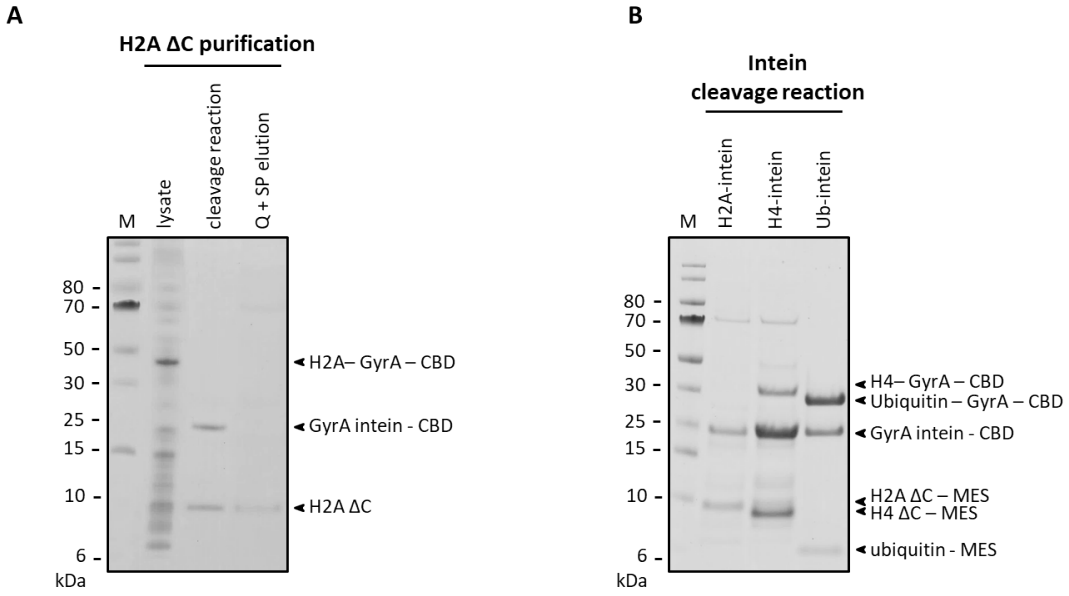


Figure 3.5: Preparation of recombinant H2A and H4 thioesters. (A) Coomassie-stained SDS-PAGE gel of representative aliquots illustrating the expression of the H2A-GyrA intein-CBD fusion, the intein-catalysed cleavage reaction and the truncated H2A eluate collected from the SP sepharose column. (B) Coomassie-stained SDS-PAGE gel of GyrA-intein catalysed cleavage reactions of H2A-intein, H4-intein and ubiquitin-intein fusions in the presence of mercaptoethanesulfonate (MES). Both H2A (1-112) and H4 (1-88) constructs present C-terminal truncations.

histones (Figure 3.1A, side loop). Both C-terminal truncations were found in the supernatant after centrifugation of the bacterial lysate and were first purified as GyrA-CBD fusion constructs on chitin beads (Figure 3.5A). The truncated histones were released from the fusion construct after addition of MES-Na (Figure 3.5B). HPLC purification of these truncations was not as efficient as in the case of the ubiquitin thioester. To separate H2A Δ113-129 MES from the cleavage reaction, the protein was applied to a combination of Q and SP sepharose anion-cation exchange columns. The pure truncated H2A harboring a thioester at the C-terminus had a molecular weight of 12244 Da as measured by ESI-MS (theoretical MW = 12243 Da).

Following the successful collaboration on the native chemical ligation of the H3 ubiquitylation marks, semi-synthetic H2AK119ub was prepared from recombinant H2A Δ113-129 MES, ubiquitin MES and a chemically synthesized peptide with cysteine on position 113 and thiolysine on position 119. After all purification steps, the measured MW of the histone analog was 22455 Da. One milligram of site-specifically ubiquitylated H2A with a native isopeptide linkage was prepared in the laboratory of Prof. Ashraf Brik.

3.1.3 Assembly of ubiquitylated histone octamers

Ubiquitylated histones were incorporated into nucleosomes or nucleosomal arrays by mixing them with unmodified histones into an equimolar 1:1:1:1 mixture. In this way, a H2BK120ub histone octamer contained two copies of the ubiquitylated H2B and two copies each of the unmodified human histones. Purification of the histone octamers from suboctameric assemblies was achieved by size exclusion chromatography (Figure 3.6A). Three separate peaks were generally recorded on such a chromatograph. The first peak, eluting around 6-7 mL, represented aggregated histones. The second peak, eluting around 10-12 mL, contained histone octamers as indicated by the stoichiometric amounts of the individual histones on the SDS-PAGE gel (Figure 3.6B). The third peak, eluting around 14 mL, contained dimers of H2A and H2B.

Semi-synthetic histone His.H2BK120ub and the fully synthetic histone HA.H2BK34ub, which had previously been produced in the laboratory of Ashraf Brik, both contained affinity purification tags (Kumar et al., 2009; Siman et al., 2013 [110]; [184]). Control His.H2B and HA.H2B proteins were prepared using the standardised human histone purification protocol described above (Figure 3.1A). Tagged unmodified and ubiquitylated H2B octameric assemblies were prepared for assembly into nucleosomal arrays (Figure 3.6C). Similarly, H3 ubiquitylated at lysines K18, K23 or K18 and K23 was also incorporated into octameric assemblies. The unmodified control of these assemblies needed no additional tag on histone H3 (Figure 3.6D).

3.1.4 Preparation of biotinylated DNA templates

Histone octamers were assembled on DNA templates to create nucleosomes and nucleosomal arrays (also referred to as chromatin fibers). Templates of human, frog or viral origin may be used in the process. Histone octamers assemble well onto these templates and reconstitute

naturally occurring chromatin fibers. Such *in vitro* reconstitutions are succeptible however to spontanous disruptions due to the low affinity of the underlying DNA sequence for the histone octamer [185]. It is for this reason that synthetic DNA templates with high nucleosomal positioning properties have artificially been evolved [186]. The 601 Widom positioning sequence of 147 bp (or the shorter 145 bp) was used in this thesis either as such or with additional linker DNA to create symmetric (187 bp) or asymmetric (171 bp) DNA templates. For nucleosomal arrays, a construct containing twelve repeats of a 200 bp sequence embedding the 147 bp Widom positioning sequence was used.

To be able to use the nucleosomes and nucleosomal arrays in affinity purification experiments the DNA templates were prepared in house in high purity, after which they were biotinylated. Production of the DNA templates required amplification using high copy number plasmids. The inserts containing the nuclesomal positioning sequences were separated from the plasmid backbone using restriction enzymes (Figure 3.7). Preparation of 171 bp DNA required the use of *Ava*I and *Not*I which produced the short DNA template and an intact plasmid

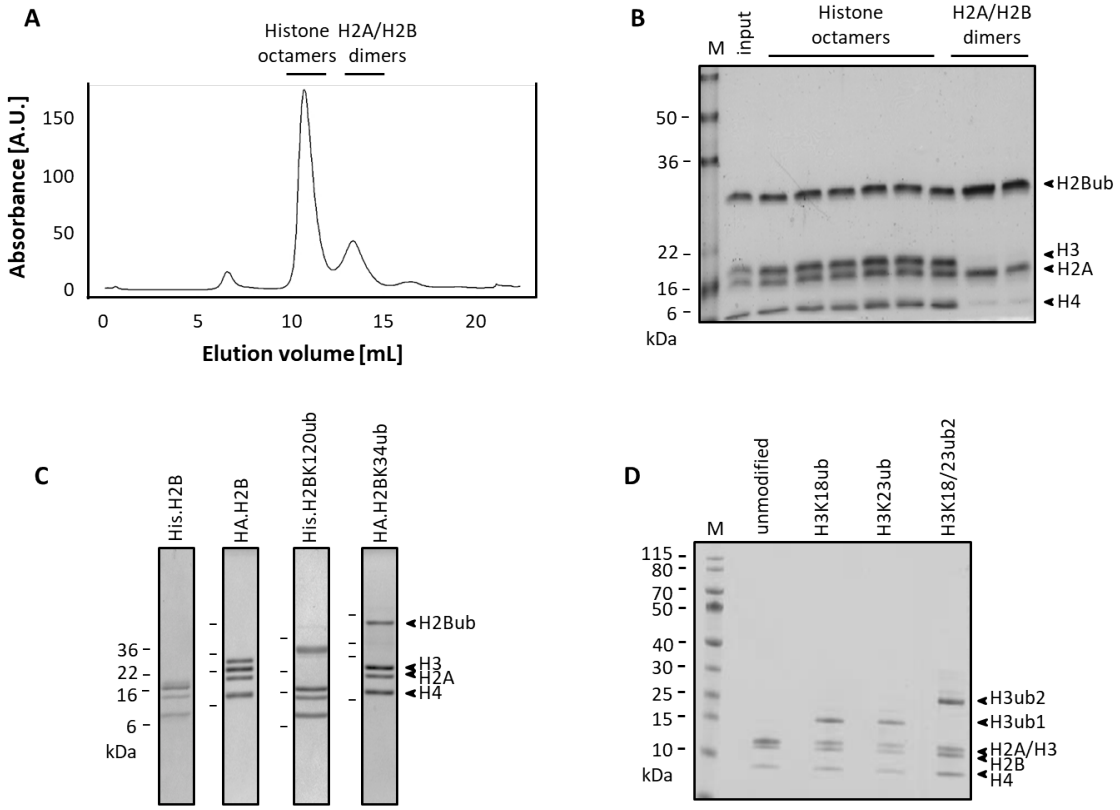


Figure 3.6: Assembly of histone octamers. (A) Elution profile of H2BK120ub-containing histone octamers from a Superdex S200 size exclusion column. (B) Coomassie-stained SDS-PAGE gel of fractions collected from the elution profile. (C) Coomassie-stained SDS-PAGE gel of representative fractions of histone octamer assemblies containing His.H2B, HA.H2B, His.H2BK120ub and HA.H2BK34ub histones. (D) Coomassie-stained SDS-PAGE gel of representative fractions of histone octamer assemblies containing unmodified, H3K18ub, H3K23ub and H3K18/23ub2 histones.

backbone (Figure 3.7A). Preparation of 12x200 DNA required the combined use of *Eco*RI, *Hae*II, *Dde*I and *Bfu*CI which digested the plasmid backbone into small fragments and released an intact 2400 bp DNA template (Figure 3.7D). Both inserts were separated from the corresponding digestion reactions by gradient PEG precipitation. To prepare the DNA templates for biotinlylation, these were first dephosphorylated at their 5' ends using Antarctic Phosphatase. This was performed to prevent self-ligation during the subsequent ligation reaction (Figure 3.7B). Biotinylation was performed with biotinylated oligos (phosphorylated at their 3' ends), specific for the 171 and for the 12x200 templates, in the presence of T4 DNA ligase (Figure 3.7B, Figure 3.7D). The incorporation of the biotin tag into the 187 and 12x200 DNA templates was verified by dot blot with a streptavidin-HRP probe (Figure 3.7C, Figure 3.7E).

3.1.5 Reconstitution of mononucleosomes and nucleosomal arrays

Unmodified or ubiquitylated histone octamers were assembled onto the biotinylated bio-187 or bio-12x200 synthetic DNA templates using salt gradient dialysis. Generally, titration series with histone octamer:template DNA ratios of 0.9:1 - 1.5:1 were performed prior to large scale reconstitutions. Most mononucleosome reconstitution reactions required octamer:DNA ratios of 1.3:1 and 1.5:1 to obtain fully saturated constructs, devoid of hemisomes (subnucleosome assemblies) or of free unbound template DNA. Most chromatin reconstitutions required octamer:DNA ratios of 1.1:1 and 1.3:1 to obtain full saturation, which was examined by micrococcal nuclease and restriction enzyme digestions.

Using biotinylated 187 bp DNA and octameric assemblies containing His.H2BK120ub and HA.H2BK34ub, uniformly ubiquitylated mononucleosomes were prepared for identification of novel nuclear interactors in affinity purification experiments (Figure 3.8B). In parallel, control biotinylated mononucleosomes were prepared from the corresponding His.H2B and HA.H2B octameric aseemblies.

Using biotinylated 12x200 bp DNA and octameric assemblies containing unmodified, H3K18ub, H3K23ub and H3K18/23ub2 or His.H2B, HA.H2B, His.H2BK120ub and HA.H2BK34ub uniformly ubiquitylated nucleosomal arrays were reconstituted for identification of novel nuclear interactors in chromatin affinity puritication experiments (Figure 3.8C, Figure 3.8E). Additional chromatin templates containing a mutated H2A histone (H2A E61A E64A, H2Amt) were prepared later to better understand the readout of H3 ubiquitylated histone by a group of proteins identified during the chromatin affinity purification experiments (Figure 3.8D).

3.1.6 Quality control of reconstituted nucleosomes and nucleosomal arrays

To obtain reproducible ubiquitylated nucleosome and chromatin templates, these were reconstituted to full saturation of the DNA template with histone octamers. Mononucleosomes saturation was examined by native agarose gel electrophoresis. The presence of unbound template DNA or the unwanted assembly of subnucleosome complexes was resolved by overtitration of histone octamers into the reconstitution reaction. Fully saturated mononucle-

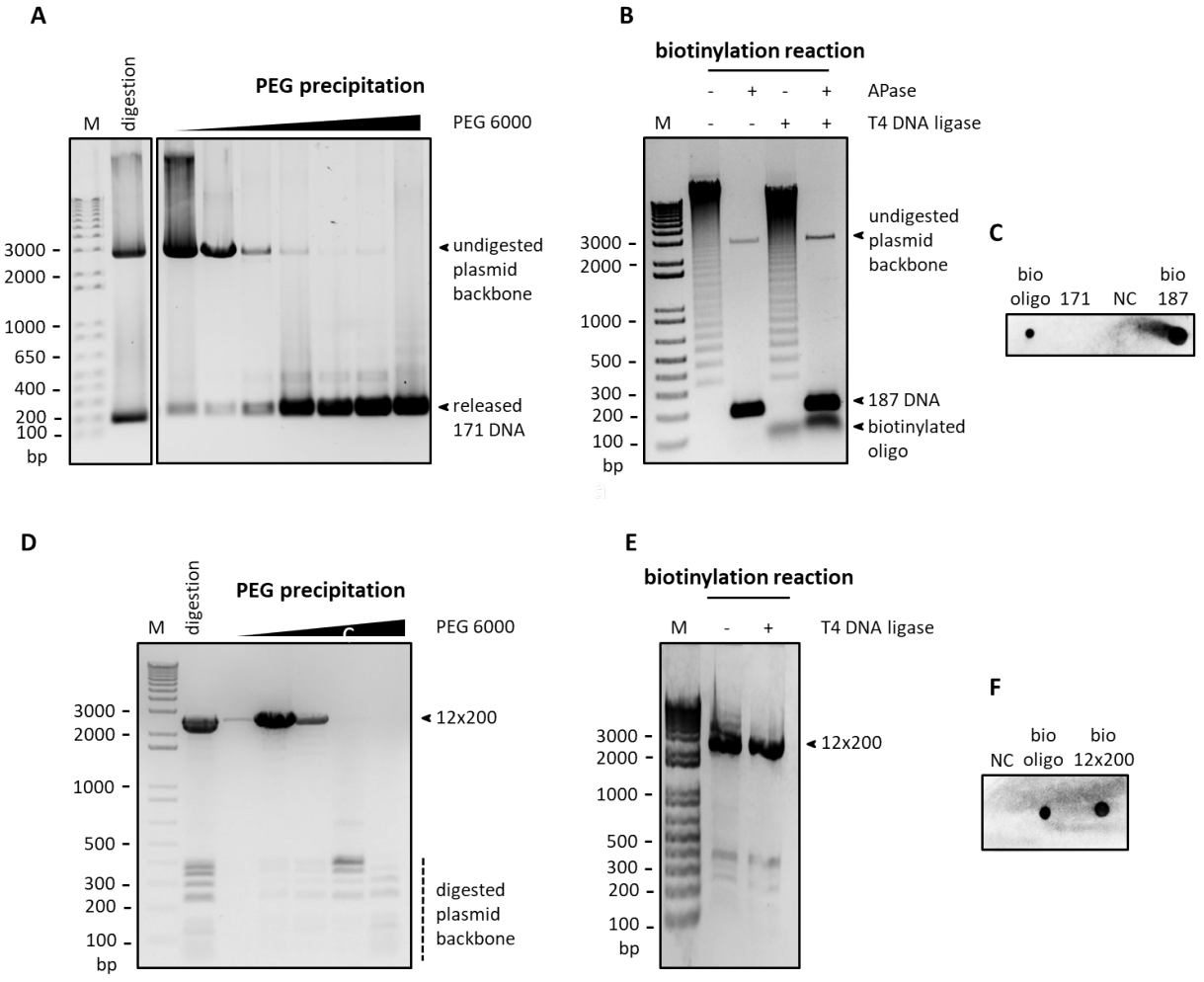


Figure 3.7: Preparation of DNA templates for mononucleosome and chromatin reconstitution. (A) Ethidium bromide (EtBr)-stained agarose gel of a pUC18-52x187 plasmid digested with *Ava*I and *Not*I restriction ezymes and gradient PEG 6000 purification of the resulting 171 bp DNA insert away from the intact plasmid backbone. (B) EtBr-stained agarose gel of dephosporylation (Antarctic phosphatase) and biotinylation (T4 DNA ligase) reactions on the 171 bp DNA. (C) Dot blot using Streptavidin-HRP conjugate to detect biotinylated DNA templates. (D) EtBr-stained agarose gel of a pUC18-12x200 plasmid digested with *Ecor*I, *Hae*II, *Dde*I and *Bfu*CI restriction ezymes and gradient PEG 6000 purification of the resulting 12x200 insert from the digested plasmid backbone. (E) EtBr-stained agarose gel of the biotinylation reaction of the 12x200 template. (F) Dot blot using Streptavidin-HRP conjugate to detect biotinylated DNA templates. NC = negative control. After biotinitylation, the mononucleosomal DNA template was 187 bp long. M = DNA ladder marker [bp]

osomes appear on an agarose gel as a single band that is shifted from 187 bp to 400-500 bp (Figure 3.8C). Since octamer composition may change the surface properties of the nucleosomes (charge, folding), different mononucleosomes travel slightly differently through a native gel.

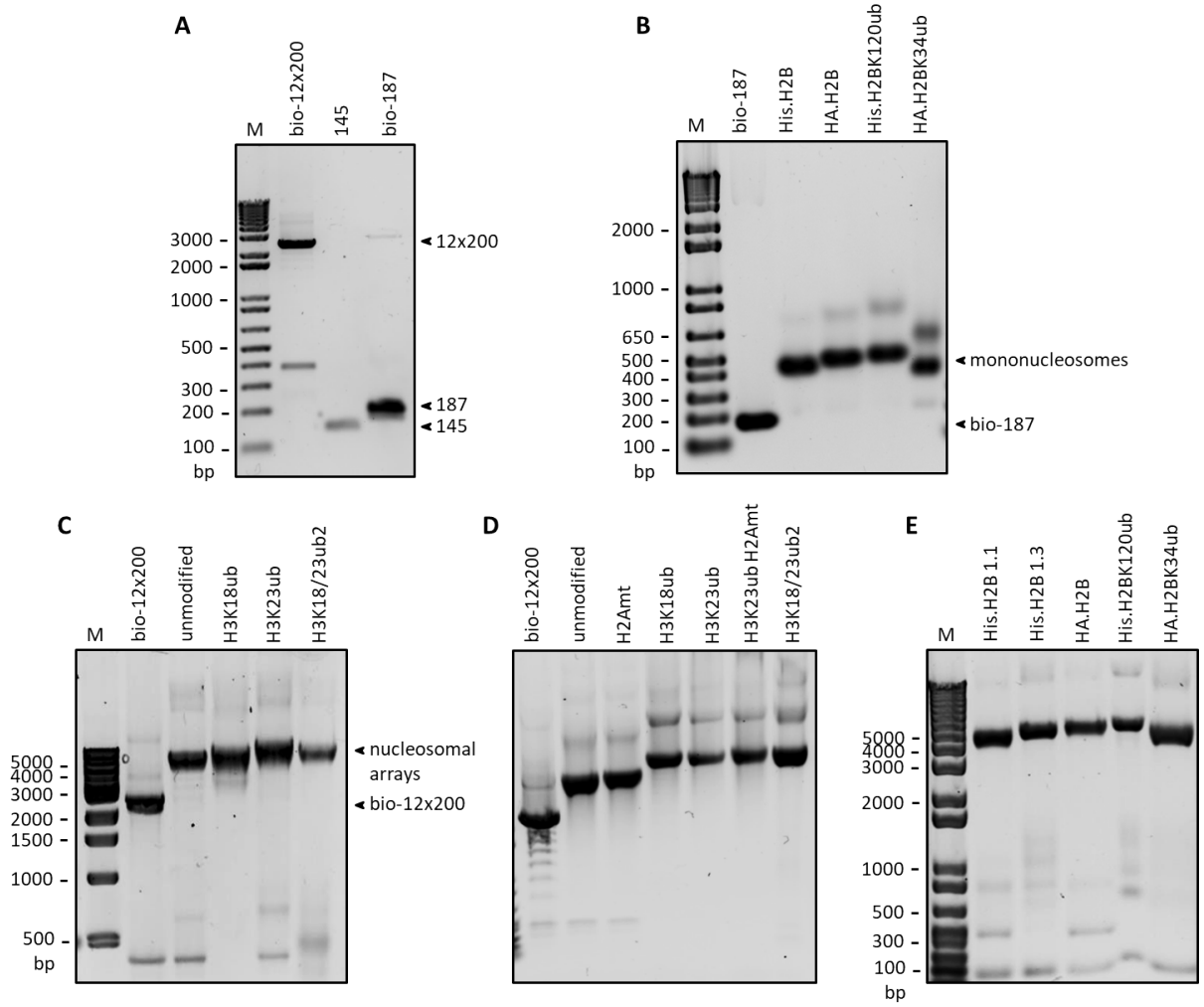


Figure 3.8: Reconstitution of mononucleosomes and nucleosomal arrays. (A) EtBr-stained gel of the pure DNA templates used in the reconstitution reactions. (B) EtBr-stained gel of biotinylated mononucleosomes, including the His.H2B, HA.H2B, His.H2BK120ub and HA.H2BK34ub reconstitutions used in affinity purification experiments for mass spectrometry analysis. (C) EtBr-stained gel of biotinylated nucleosomal arrays, including unmodified, H3K18ub, H3K23ub and H3K18/23ub2 reconstitutions used in affinity purification experiments for mass spectrometry analysis. (D) EtBr-stained gel of biotinylated nucleosomal arrays, including unmodified, H2Amt, H3K18ub, H3K23ub, H3K23ub H2Amt and H3K18/23ub2 reconstitutions used in affinity purification experiments with recombinant proteins. (E) EtBr-stained gel of biotinylated nucleosomal arrays, including unmodified, His.H2B, HA.H2B, His.H2BK120ub and HA.H2BK34ub reconstitutions used in affinity purification experiments for mass spectrometry analysis.

Reconstitution of nucleosomal arrays was investigated with respect to two different parameters. The first parameter, nucleosome positioning, refers to the ability of the DNA sequence to prevent deviation (sliding or unwrapping) of a nucleosome from a set point (dyad axis) along the DNA sequence. If nucleosomes were misplaced, the desired uniform distribution conferred by the core positioning sequences would be skewed, which would result in the for-

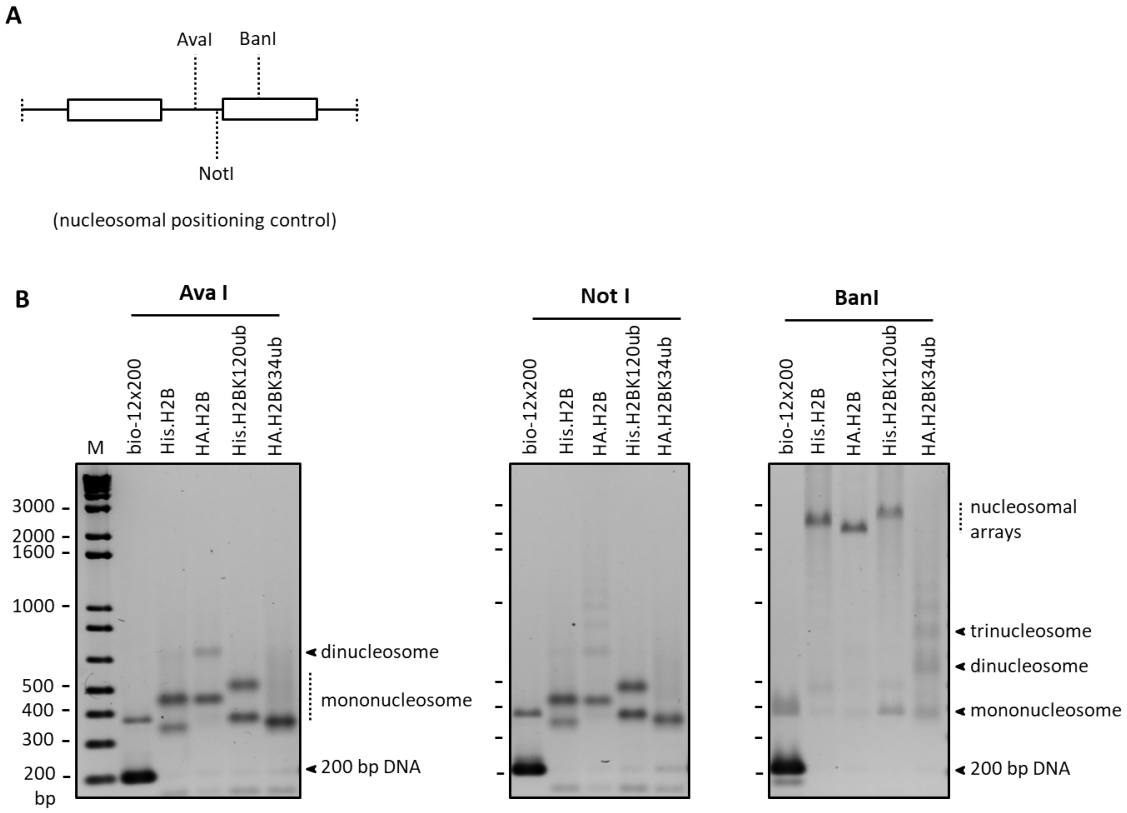


Figure 3.9: Nucleosomal positioning control on reconstituted nucleosomal arrays. (A) Schematic annotation of cleavage sites for *Ava*I, *BanI*I and *Not*I restriction enzymes with respect to two consecutive nucleosomal positioning sequences. (B) EtBr-stained gel of enzymatic digests of nucleosomal arrays including His.H2B, HA.H2B, His.H2BK120ub and HA.H2BK34ub-containing templates.

mation of unequal nucleosome spacing across the array. The second parameter, nucleosome occupancy, refers to the local density of nucleosomes within a given chromatin template. The nucleosome arrays used in this thesis contained twelve positioning sequences and were thus able to position strongly twelve nucleosomes. An undersaturated nucleosome array would contain less than twelve assembled nucleosomes.

Nucleosome positioning was addressed by restriction enzyme digestion, using enzymes that cleave unprotected DNA (Figure 3.9A). Restriction enzyme *Ava*I cleaves the 12x200 DNA template 34 bp upstream of the start of the core positioning sequence (108 bp away from the dyad axis). Restriction enzyme *Not*I cleaves the DNA template 7 bp upstream of the core positioning sequence (81 bp from the dyad axis). Restriction enzyme *Ban*I cleaves the DNA template 13 bp downstream of the core positioning sequence (61 bp from the dyad axis).

Upon treatment with *Ava*I, the unmodified and ubiquitylated H2B chromatin arrays were fully digested to products which run to a front corresponding of 400 - 500 bp (Figure 3.9B). This indicated that the linker DNA was available for digestion along the entire length of

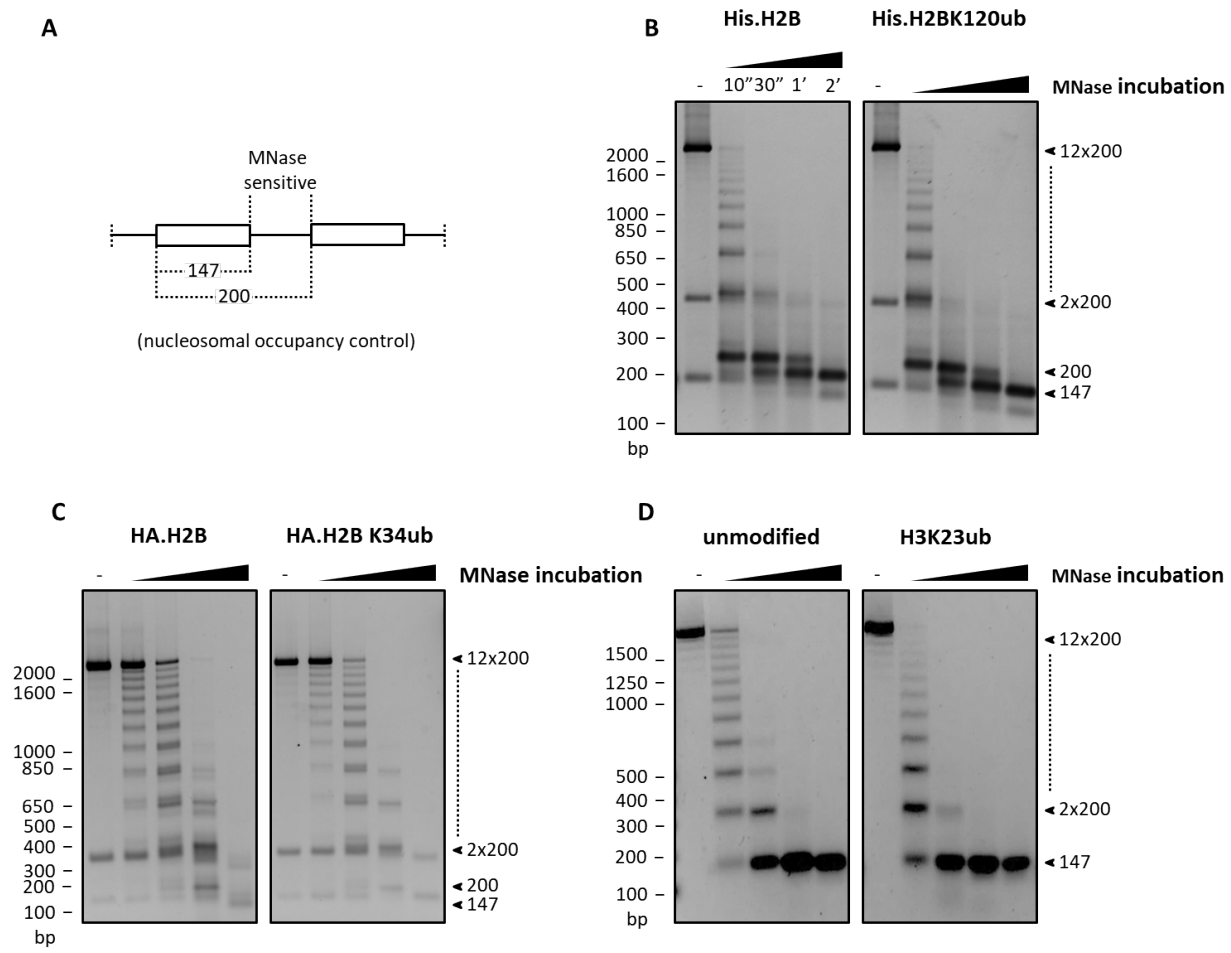


Figure 3.10: Nucleosomal occupancy control on reconstituted nucleosomal arrays. (A) Schematic representation of the micrococcal nuclease (MNase) sensitive region between two consecutive nucleosomal positioning sequences. (B) EtBr-stained gel of time-course MNase digests of nucleosomal arrays containing His.H2B and His.H2BK120ub. (C) EtBr-stained gel of time-course MNase digests of nucleosomal arrays containing HA.H2B and HA.H2BK34ub. (D) EtBr-stained gel of time-course MNase digests of nucleosomal arrays containing unmodified and H3H23ub.

the nucleosome array. Upon treatment with *Not*I, the unmodified and ubiquitylated H2B chromatin arrays were fully digested to products which run to a front corresponding of 400 - 500 bp. This suggested that nucleosomes were not placed outside of the core positioning sequences. Together, the digestion reactions indicate that the nucleosomes were properly positioned after reconstitution.

Upon treatment with *Ban*I, the His.H2B, HA.H2B and His.H2BK120ub chromatin fibers show protection against enzymatic digestion. In the case of HA.H2BK34ub, this protection is lost and several digestion products (mono-, di-, trinucleosomes) are observerd on the native agarose gel. This is not an effect of nucleosomal positioning, but rather a consequence of spontaneous nucleosomal breathing [187]. H2BK34ub lowers the energy needed to brake the histone-DNA contacts at the two extremeties of the nucleosomal positioning sequence.

Nucleosome occupancy was addressed experimentally by digestion with micrococcal nuclease (MNase). In a MNase digestion experiment, unmodified chromatin is processed in parallel with ubiquitylated chromatin. Unprotected DNA (linker DNA or unbound nucleosomal positioning sequence) is sensitive to MNase digestion (Figure 3.10A). All chromatin templates assayed produced twelve distinct cleavage bands in the first time-point of the digestion reaction (Figure 3.8B, Figure 3.10C, Figure 3.10D). Second, upon treatment with either *Ava*I, *Not*I or *Ban*I, there was no evidence of unbound positioning sequences, which would appear as 200 bp DNA fragments on the native agarose gel (Figure 3.10B, Figure 3.10C). Together with the observation that MNase digests produced twelve cleavage products, this indicated that all nucleosome arrays that were produced for the chromatin affinity purification experiments were fully saturated.

The micrococcal nuclease digestion rate was slightly faster in the case of ubiquitylated nucleosomal arrays (Figure 3.10B, Figure 3.10C, Figure 3.10D). This was not an effect of nucleosome occupancy since the 147 bp or 200 bp digestion end products accumulate to the same extent in both unmodified and ubiquitylated arrays. This observation may have resulted from a higher order folding of ubiquitylated chromatin fibers that differed from that of the unmodified controls. Ubiquitylation of histones likely opened chromatin fibers locally, making the linker DNA more sensitive to MNase treatment.

3.2 Mapping of nuclear proteins recognising ubiquitylated histones

3.2.1 Chromatin affinity purification - mass spectrometry

To isolate protein complexes that bind modified histones, an affinity purification approach was previously developed in our laboratory [132]. This technique uses biotinylated nucleosomal arrays as an affinity tag for nuclear proteins. After purification, the proteins which are recruited to chromatin are identified by mass spectrometry. Chromatin affinity purification - mass spectrometry (ChAP-MS) enriches for proteins from nuclear extracts (Figure 3.11B). To separate false-positives from true-positive interactors, ChAP-MS is combined with stable isotope labeling of amino acids in cell culture (SILAC) and uses extracts that have been prepared from human cells grown in media that have been supplemented with heavy arginine (+6 Da, ^{13}C) and lysine (+4 Da, ^{2}H) isotopes.

A regular ChAP-MS scheme contains a forward and a reverse experiment (Figure 3.11A). In the forward experiment, unmodified chromatin is incubated with light nuclear extract and modified chromatin is incubated with heavy nuclear extract. In the reverse experiments, labels are swapped such that the unmodified chromatin is incubated with heavy nuclear extract and the modified chromatin is incubated with light nuclear extract. Pooled eluates from the forward experiment are analysed by mass spectrometry. Peptide intensities from proteins enriched from the heavy or the light extract are measured and the resulting ratios between heavy and light protein groups (H/L) are sorted on a logarithmic scale (Figure 3.11C). In parallel, the eluates from the reverse experiment are also analysed by mass spectrometry. The H/L ratios are inverted and sorted as in the forward experiment (Figure3.11D). The distributions of H/L ratios identified in the two experiments follow closely a normal distribution of 0 mean and 1 standard deviation (Figure 3.11C, Figure3.11D). If most measured protein groups fall within the boundaries of the normal distribution, several proteins are found on either side of the distribution mean. Enriched outliers have a positive H/L ratio. Excluded outliers have a negative H/L ratio.

To separate false-positive from true-positive outliers the two H/L distributions were plotted against each other (Figure 3.11E). Such an interactome plot separates all measured protein groups into four quadrants. The top left quadrant included ubiquitin, histones and proteins that bind strongly to chromatin with no regard for the modification. The bottom left quadrant included proteins that are reproducibly excluded from modified chromatin. The bottom right quadrant identified interactors that have been enriched during the preparation of the heavy extract and bind chromatin with no regard for the modification. The top right quadrant included protein groups that have been reproducibly enriched on the modified chromatin.

Until recently, true-positive and true-negative outliers were selected from the top right and bottom left quadrants based on an arbitrary threshold or cutoff value which was applied to both the forward H/L and the reverse L/H ratio distributions (Figure 3.11F). Since the

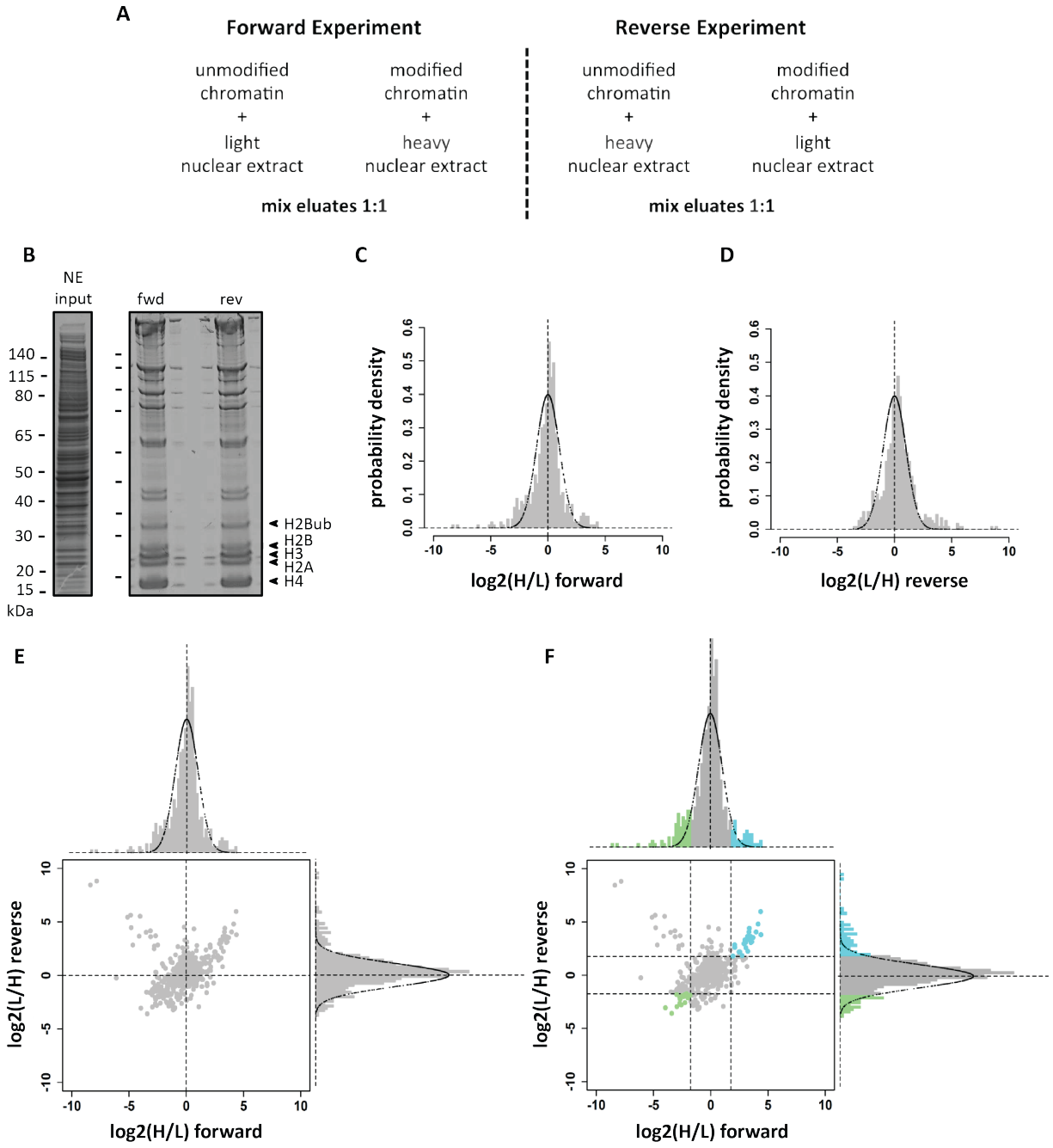


Figure 3.11: Chromatin affinity purification - mass spectrometry. (A) Schematic representation of experiments performed for isolation of nuclear proteins that recognise modified chromatin. (B) Input nuclear extract and eluted proteins from forward (fwd) and reverse (rev) biochemical experiments. (C) Histogram of heavy/light (H/L) ratios for proteins identified in the forward experiment, with normal distribution plotted on top. (D) Histogram of L/H ratios for proteins indetified in the reverse experiment. (E) Intersection of the H/L distributions from the forward and reverse experiments. (F) Highlighted enriched (cyan) and excluded (light green) proteins in the two H/L distributions and their intersection.

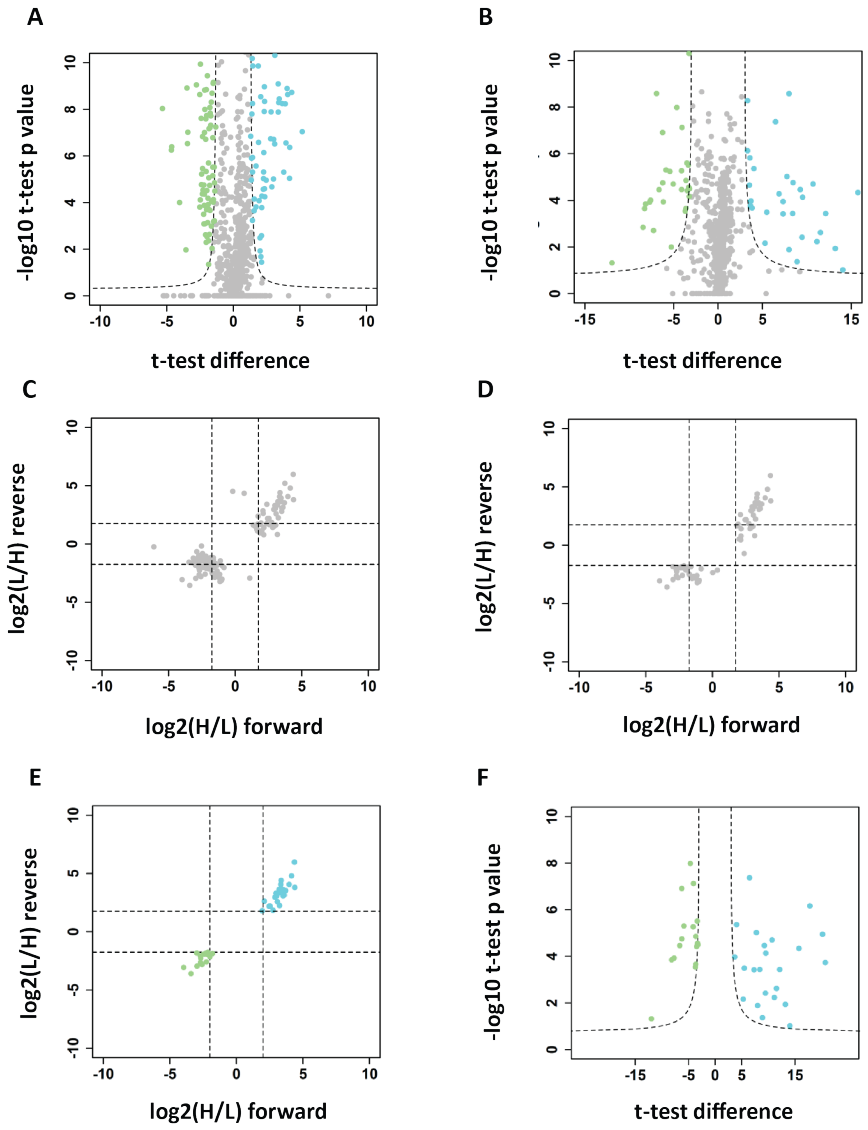


Figure 3.12: Statistical analysis of ChAP-MS datasets. (A) One-sample student's t-test analysis of the grouped H/L ratios from the forward and reverse experiemnts. The dashed parabolas deliniate the test's significance threshold (B) Two-sample student's t-test analysis of the separate mean H/L ratios from the forward and reverse experiments. The dashed parabolas represent the significance treshold. (C) Representation of the one-sample t-test significant outliers on the interactome plot. (D) Representation of the two-sample t-test significant outliers on the interactome plot. (E) Representation of significant reproducible outliers on an interactome plot. (F) Representation of significant reproducible outliers on a volcano plot. Enriched outliers are highlighted in cyan, excluded outliers are highlighter in light green.

affinity purification experiments are measured thrice in the mass spectrometer, the reproducibility of the measurements can be statistically quantified [133]. In order to identify significant outliers the triplicate measurements from the forward and the reverse experiments were subjected to student's t-test analysis. Two types of analyses can be performed.

A one-sample t-test statistical analysis pools the information from the forward and the reverse experiments in one dataset. The test calculates the overall mean of the three H/L ratios from the forward distribution and the three inverse L/H ratios from the reverse distribution and compares it to the zero mean of the entire dataset. The one-sample t-test calculates if the pooled mean is significantly different from the zero mean and plots the difference values against the corresponding p values (Figure 3.12A). The one-sample t-test is an indicator of measurement reproducibility across the forward and the reverse datasets.

A two-sample student's t-test calculates if the difference between the mean of the three H/L ratios of a protein identified in the forward experiment is significantly different from the mean H/L ratios of the same protein identified in the reverse experiment. The difference between the mean H/L ratios of each protein group measured in the two experiments is plotted against the corresponding t-test p value (Figure 3.12B). The two-sample t-test is an indicator of experimental difference between the forward and the reverse datasets.

The p values in the two t-test analyses were calculated using a permutation based algorithm whose false discovery rate (FDR) value was set to 0.01. Additionally, a S0 = 2 constant was added to increase pooled sample variance (decrease background noise) during calculations of both the t-test statistics. These parameters specified the significance threshold of the two analyses.

The significant outliers identified by the statistical analyses did not all agree with the previously set H/L cutoff values. Both t-test analyses were more permissive in the identifiaction of enriched factors than the set H/L cutoff. In the one-sample t-test analysis, many interactors which defied the null hypothesis had low H/L enrichment or depletion ratios (Figure 3.12C). This selection, included all enriched or depleted factors, which were selected based on the H/L cutoff, but had a relatively high identification background. In the two-sample t-test, the comparison between the measurements in the forward and reverse experiments, improved the identification confidence (Figure 3.12D). This being said, some significantly enriched interactors were still only enriched in one of the H/L ratio distributions and some significantly excluded interactors were only excluded from one of the H/L ratio distributions. This happened because the two-sample t-test evaluated if the difference between the mean H/L ratios from the forward and the reverse experiments was statistically significant. The two-sample t-test did not check for strict reproducibility of the two experiments, that is if the inverse of the reverse experiment matched the forward experiment, which was controlled by the one-sample t-test. The two-sample t-test was superior in its identification confidence to the one-sample t-test, but had some limitations with regard to the reproducibility of the forward and reverse biochemical experiments.

The t-test analysis was complemented by the previously set H/L cutoff values (Figure 3.12E, Figure 3.12F). The intersection of the two-sample t-test with the H/L distribution thresholds was chosen to insure that biochemically reproducible enriched or excluded outliers were also statistically significant. Throughout the thesis, the thresholds were set to a log2 (H/L) value of 1.75 in the forward experiment, a log2 (L/H) value of 1.75 in the reverse experiment and a

two-sample t-test FDR value of 0.01 and S0 value of 2. Two parameters were thus arbitrary set for confident identification of statiscally enriched interactors: the H/L cutoff and the S0 constant. The final interactome (H/L ratio distributions) and volcano (t-test statistics) plots presented in the thesis contained all measured protein groups and highlighted the significantly enriched outliers. To this end, the interactome plots focused on the representation of the top right quadrant and the volcano plots display only the positive t-test significance area.

3.2.2 Histone ubiquitylation interactome mapping

The interactome of H2BK120ub

Previously in our laboratory, using chromatin affinity purification coupled with quantitative mass spectrometry, the interactome of H2BK120ub was presented [83]. In that study, ubiquitylation recruited to chromatin the RNA polymerase, proteins that control transcription elongation, the small nuclear RNA-processing integrator complex and the switch/sucrose nonfermentable chromatin remodeling complex.

To understand the molecular requirements for histone ubiquitylation readout in general and for H2BK120ub readout in particular, we repeated the affinity purification scheme mentioned above with the addition of ubiquitylated H2BK120ub histone and ubiquitylated H2BK120ub mononucleosomes as modified templates (Figure 3.11A, Figure 3.13). The two additional affinity matrices were selected to provide information with regard to the requirement of inter- and trans-nucleosomal contacts as well as additional interaction surfaces (DNA gyres, linker DNA and unmodified histones) for readout of the ubiquitylation mark.

In the histone affinity purification experiment, four proteins were reproducibly enriched (Figure 3.13A, Figure 3.13B). Nuclesomal assembly protein 1-like 1 and 4 (NAP1L1, NAP1L4), nucleolar (upstream) transcription factor (UBTF) and the Treacher-Collins ribosome biogenesis factor 1 (TCOF1) discriminated ubiquitylated H2B histone from the unmodified control.

In the mononucleosome affinity purification experiment, a group of fourteen proteins were reproducibly enriched (Figure 3.13C, Figure 3.13D). Four subunits of the integrator complex (INTS 1, INTS3, INTS6 and INTS12) and the RNA polymerase II subunit A (POLR2A) were identified as strong readers. Zinc-finger containing proteins ZMYNDB, ZNF687, structural maintenance of chromosomes 1A and 3 (SMC1A, SMC3) and the E3 protein ligase ring finger 169 (RNF169) also prefered the ubiquitylated mononucleosome matrix to the unmodified control.

In the chromatin affinity purification experiment, thirty-five proteins were reproducibly enriched (Figure 3.13E, Figure 3.13F). In agreement with the published H2BK120ub chromatin interactome [83], eleven subunits of the integrator complex (INTS1-INTS10, INTS12) were recruited to the modified chromatin. Three subunits of the RNA polymerase machinery (POLR2B, POLR2C, POLR2E) and several subunits controlling transcription elongation (NELFB, NELFC/D, SUPT5H) were also enriched. DNA excision repair proteins 1 and 4 (ERCC1 and ERCC4) as well as the Zn-finger proteins ZMYNDB, ZNF609 and ZNF687

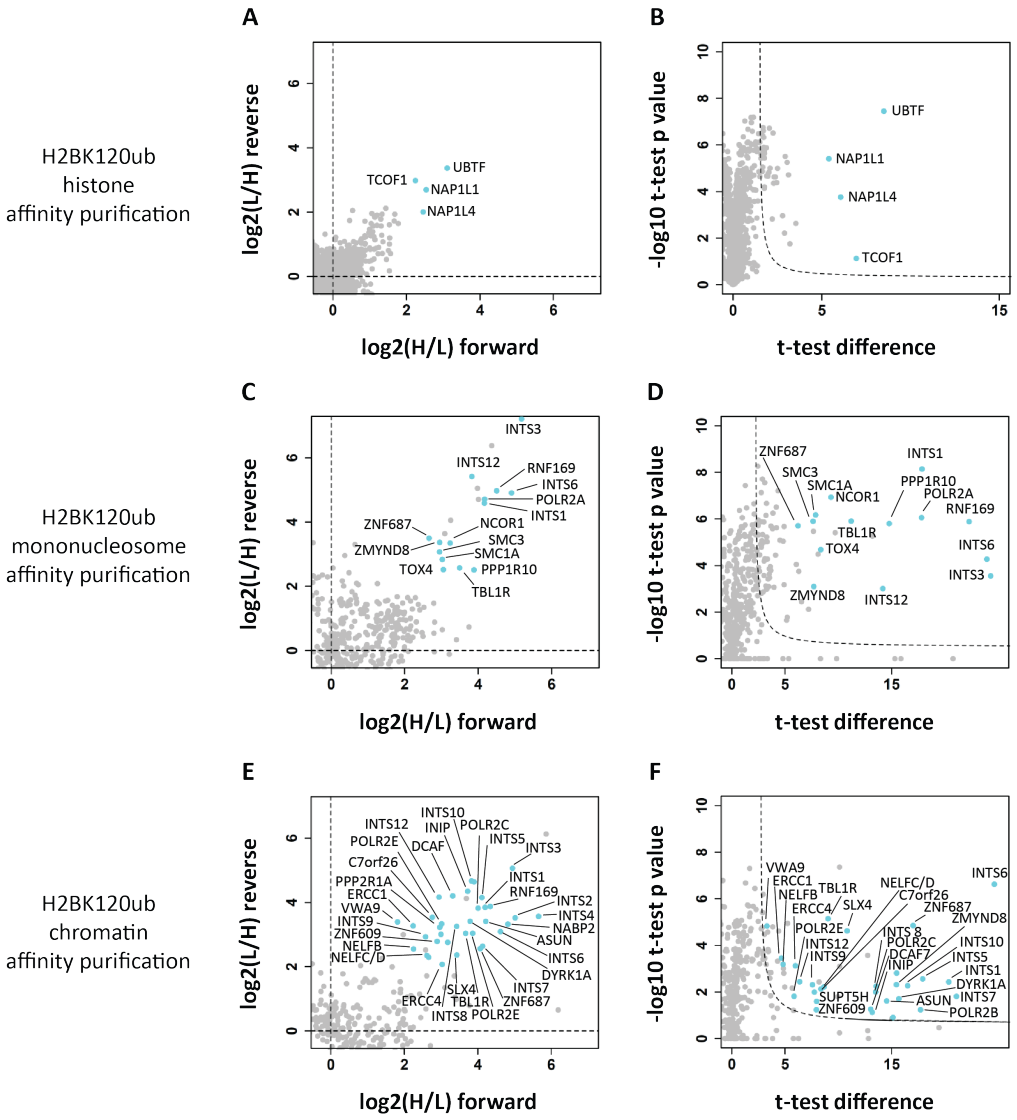


Figure 3.13: ChaP-MS analysis of H2BK120ub. (A) Representation of enriched interactors on an interactome plot from a histone affinity purification experiment. (B) Representation of enriched interactors on a volcano plot from a histone affinity purification experiment. (C) Representation of enriched interactors on an interactome plot from a mononucleosome affinity purification experiment. (D) Representation of enriched interactors on a volcano plot from a mononucleosome affinity purification experiment. (E) Representation of enriched interactors on an interactome plot from a chromatin affinity purification experiment. (F) Representation of enriched interactors on a volcano plot from a chromatin affinity purification experiment.

and the E3 ligase RNF169 preferred the ubiquitylated chromatin array over the unmodified control.

There was no overlap between the factors that were enriched using histone affinity purification and the factors that bound the mononucleosome or the chromatin templates. None

of the proteins enriched with the histone matrix werre present in the mononucleosome or the chromatin sets. None of the proteins enriched on the mononucleosome or the chromatin matrices were found within the histone set.

There was overlap between the mononucleosome and chromatin affinity purification experiments. Several subunits of the integrator complex, the RNA polymerase and several Zn-finger proteins were found in both experiments. The momonucleosome matrix recruited however only a fraction of the factors enriched on the chromatina arrays.

The nucleosomal arrays provide thus a more complex and larger interaction surface, which translates during the affinity purification into more binding events. The biotiniylated nucleosomal arrays create efficient affinity matrices, where the nuclear interactors' need for inter- and trans-nucleosomal contacts is statisfied and where the presence of ubiquitin is readily discriminated.

The interactome of H2BK34ub

Histone and chromatin affinity purification experiments were also used to identify proteins that recognise H2BK34ub.

Using a ubiquitylated histone matrix, the H2BK34ub modification enriched for a number of proteins with seemingly unrelated functions: the phospholipase A2 activating protein (PLAA), the transcriptional regular SUB1, the solute carrier family member 4A1 (SLC4A1), the ADP ribosylation factor like 6 (ARL6), and the Zn-finger RNA binding protein 2 (ZRANB2) (Figure 3.14A, Figure 3.14B).

Using a ubiquitylated chromatin matrix, the H2BK34ub mark reproducibly enriched for six proteins, including the DNA Polymerase E3 (POLE3), the chromatin accessibility complex 1 (CHRAC1), a subunit of the INO80 chromatin remodeling complex and the actin related protein 5 (ACTR5) (Figure 3.14C, Figure 3.14D).

Similarly as for the H2BK120ub modification, there was no overlap between the histone and the chromatin affinity purification datasets collected for H2BK34ub.

None of the interactors enriched on the H2BK120ub histone, mononucleosome or chromatin matrices was identified within the H2BK34ub-specific datasets.

The interactome of H3K18ub and H3K23ub

The chromatin affinity purification mass spectrometry technique was also used to identify the proteins which recognise the H3K18ub, H3K23ub and H3K18/23ub2 modifications.

The proteins purified on the H3K18ub template were more widely distributed than the other two H3 ubiquitylated templates (Figure 3.15A, Figure 3.15B). After applying the statistical

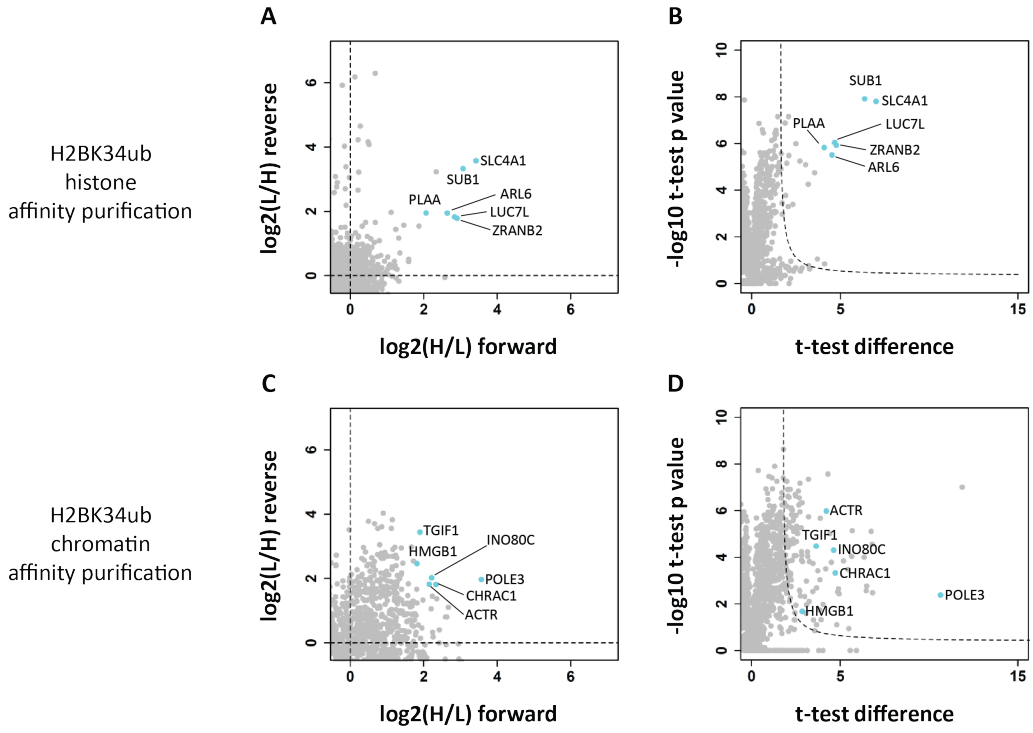


Figure 3.14: ChaP-MS analysis of H2BK34ub. (A) Representation of enriched interactors on an interactome plot from a histone affinity purification experiment. (B) Representation of enriched interactors on a volcano plot from a histone affinity purification experiment. (C) Representation of enriched interactors on an interactome plot from a chromatin affinity purification experiment. (D) Representation of enriched interactors on a volcano plot from a chromatin affinity purification experiment.

analysis and the enrichment threshold, twenty-eight proteins were found reproducibly enriched. These included as strongest interactors the DNA methyltransferase 1 (DNMT1), the ubiquitin specific protease 7 (Usp7), the Sex comb on midleg-like 2 (SCML2), the acetyl-coA carboxylase alpha (ACACA) and the ubiquitin specific protease 3 (Usp3). Other enriched interactors included Zn-finger proteins ZBTB1, ZBTB14 and ZBTB44, histone deacetylase components SIN3A and associated subunits SAP30 and SAP 130, transcription regulators Forkhead Box 1 and 2 (FOXK1, FOXK2) as well as MAX dimerization protein MLX and interacting proteins MLXIP and MLXIPL with transcription factor properties.

Using the H3K23ub nucleosomal array, six proteins were enriched by affinity purification (Figure 3.15C, Figure 3.15D). Similarly to the H3K18ub template, DNMT1, Usp7, SCML2 and ACACA were among the enriched factors. In addition, the DNA sensor 5-hydroxymethylcytosine binding embryonic stem cell specific (HMCES) and the G2/M phase specific E3 ubiquitin ligase G2E3 also preferred the ubiquitylated template to the unmodified control. Using the double-modified H3K18/23ub2 nucleosomal array, six proteins were enriched by affinity purification (Figure 3.15E, Figure 3.15F). As in the H3K18ub and the H3K23ub ex-

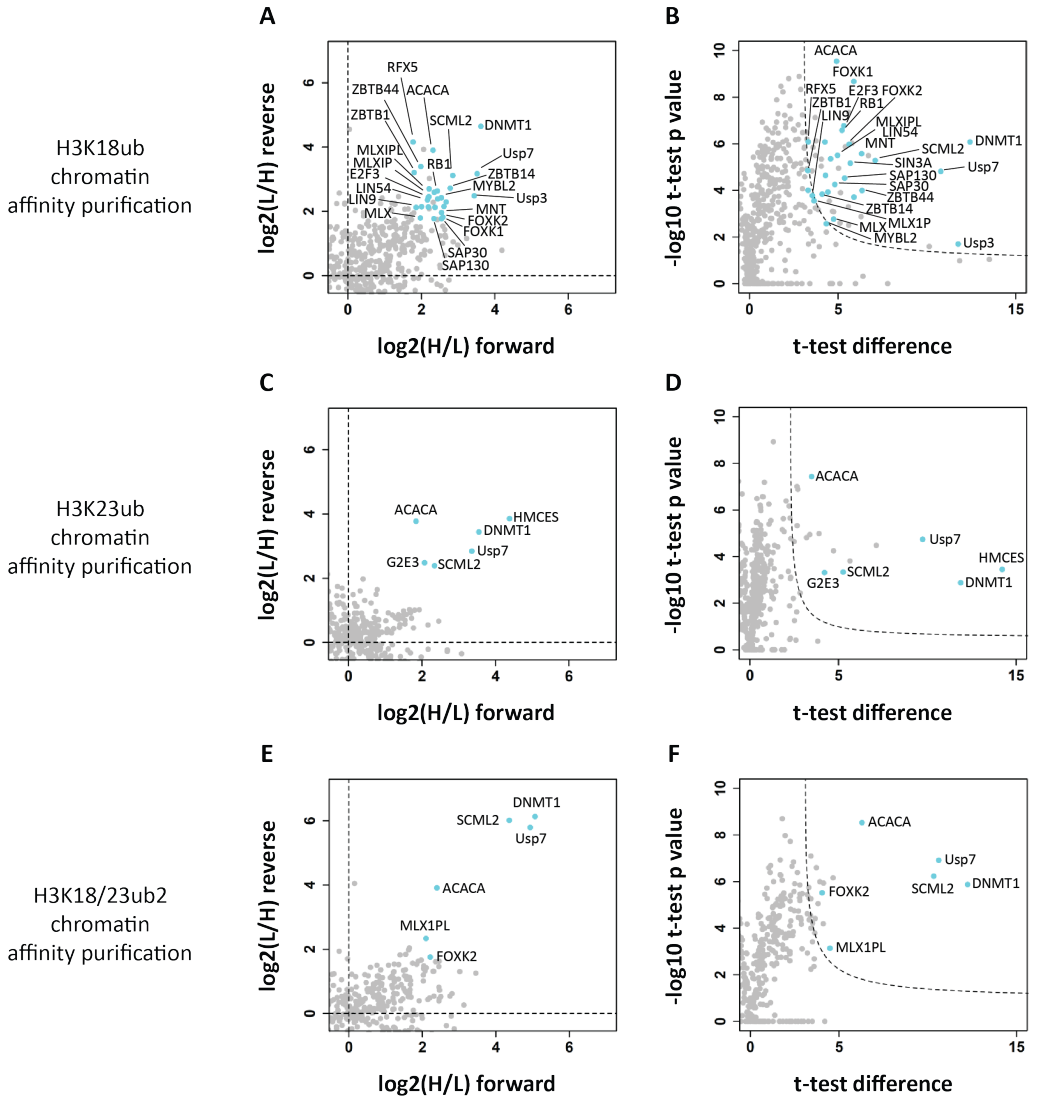


Figure 3.15: ChaP-MS analysis of H3 ubiquitylaion. (A) Representation of enriched interactors on an interactome plot from a H3K18ub chromatin affinity purification experiment. (B) Representation of enriched interactors on a volcano plot from a H3K18ub chromatin affinity purification experiment. (C) Representation of enriched interactors on an interactome plot from a H3K23ub chromatin affinity purification experiment. (D) Representation of enriched interactors on a volcano plot from a H3K23ub chromatin affinity purification experiment. (E) Representation of enriched interactors on an interactome plot from a H3K18/23ub2 chromatin affinity purification experiment. (F) Representation of enriched interactors on a volcano plot from a H3K18/23ub2 chromatin affinity purification experiment.

periments, this purification enriched for DNMT1, Usp7, SCML2 and ACACA. In addition, it recruited MLX1PL and FOXK2, which were also found on the H3K18ub affinity matrix.

There was overlap between the three H3 ubiquitylated affinity matrices. DNMT1, Usp7,

SCML2 and ACACA were the highest enriched factors in all datasets. There was no overlap between these three interactomes and the interactomes collected for the H2BK120ub or H2BK34ub chromatin templates.

Complete lists of enriched factors are presented in Table 3.1, Table 3.2 and Table 3.3. For each interactome, the tables list three separate identifications: the two-sample t-test analysis, the application of the H/L enrichment cutoff and the combined use of the H/L enrichment cutoff with the t-test statistical analysis which has been presented in the thesis.

3.2.3 Network analysis of the histone ubiquitylation interactome maps

Some proteins identified in the histone, mononucleosome and chromatin affinity purification experiments are part of large annotated complexes which perform known biochemical functions. This is the case of the RNA polymerase machinery and the integrator complex found in the H2BK120ub interactome. Other proteins form smaller complexes involved in processes such as DNA repair (ERCC1, ERCC4, SLX4), transcription elongation control (NELFB, NELFCD, SUPT5H) or maintenance DNA methylation (DNMT1 and Usp7).

There are however many factors which are not part of any of these two types of complexes. Such proteins may still form weaker interactions with the defined core protein complexes, they may form interactions with other isolated proteins or they may contact the ubiquitylated chromatin templates directly. To connect proteins from within individual interactomes, STRING analysis was performed (Figure 3.16) [177]. An online STRING analysis curates information about a given set of proteins (one interactome) from several publically available databases and links proteins that share common features. STRING analysis collects experimental evidence (biochemical interaction, coexpression analysis), genetic evidence (fusion events, neighboring genes) and evidence derived from text-mining from publically available abstracts.

The STRING analysis indicated that in the H2BK120ub mononucleosome list of enriched factors, the negative receptor corepressor 1 (NCOR1) connected transducing beta like1 X-linked receptor 1 (TBL1XR1) to the RNA polymerase A subunit (Figure 3.16B). Similarly, in the H2BK34ub chromatin interactome, the STRING analysis pointed out that the weak association between INO80C and CHRAC1 connected ACTR5 and POLE3 in a process that may involve actin polymerisation and DNA replication (Figure 3.16E). Using STRING analysis, the protein core formed by the integrator complex in the H2BK120ub chromatin interactome accommodated also additional factors such as the cleavage and polyadenylation specific factor 3L (CPSF3L), the human asunder homolog (ASUN), the von Willebrand factor A9 (VWA9), and the nucleic acid binding protein 2 (NABP2) (Figure 3.16C). Further text-mining revealed that CPS3FL, ASUN and VWA9 were in fact subunits of the integrator complex: INTS11, INTS13 and INTS14. The STRING analysis facilitated the identification of all 14 subunits of the integrator complex within the enriched proteins purified from the H2BK120ub chromatin affinity matrix.

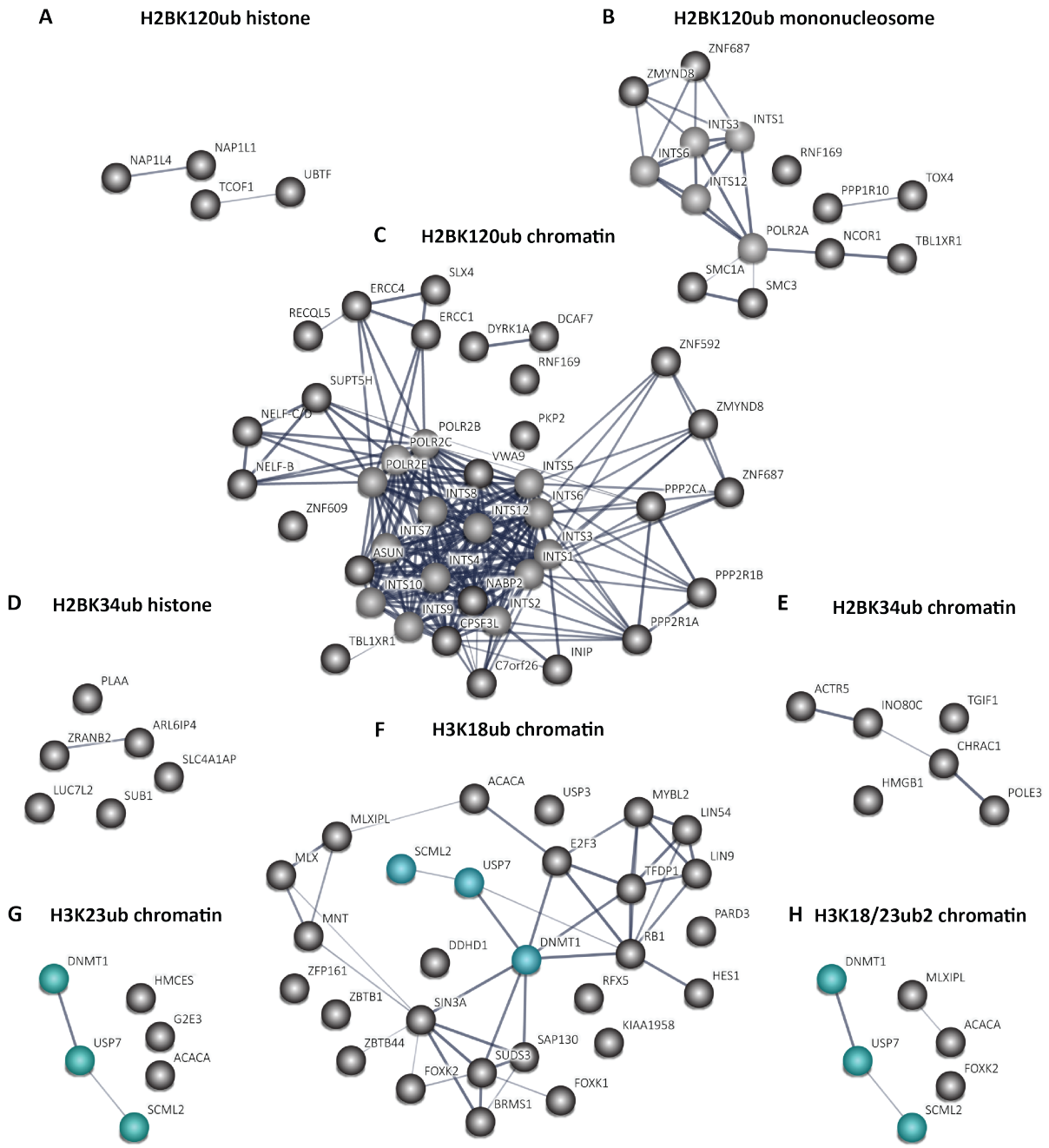


Figure 3.16: STRING analysis of significantly enriched interactors. (A) Interaction network of the H2BK120ub histone affinity purification dataset. (B) Interaction network of the H2BK120ub mononucleosome affinity purification dataset. (C) Interaction network of the H2BK120ub chromatin affinity purification dataset. (D) Interaction network of the H2BK34ub histone affinity purification dataset. (E) Interaction network of the H2BK34ub chromatin affinity purification dataset. (F) Interaction network of the H3K18ub chromatin affinity purification dataset. (G) Interaction network of the H3K23ub chromatin affinity purification dataset. (H) Interaction network of the H3K18/23ub2 chromatin affinity purification dataset. The DNMT1-Usp7-SCML2 association in the three H3 ubiquitylation interactomes is highlighted in cyan.

With regard to the H3 ubiquitylation datasets, the STRING analysis connected DNMT1 and Usp7 with SCML2 in all three interactomes (Figure 3.16F, Figure 3.16G, Figure 3.16H). The strength of the association was stronger between DNMT1 and Usp7 and weaker between Usp7 and SCML2. DNMT1 was placed at the centre of the STRING association network in the H3K18ub dataset, implying that it may serve to recognise the H3 ubiquitylated chromatin and recruit subsequently additional proteins and protein complexes.

There was no direct linkage between ACACA and any of the other strongly enriched interactors (Figure 3.16G, Figure 3.16H). ACACA could however indirectly be connected to DNMT1 through MLXIPL (and SIN3A) or through the transcription factor E2F3 which accumulates in the cell at the end of S phase (Figure 3.16F). In the H3K23ub dataset, no known association between SCML2, Usp7 or DNMT1 with the hydroxymethylation sensor HMCES or the ubiquitin ligase G2E3 could be assigned (Figure 3.16G).

The STRING analysis also facilitated the identification of two additional complexes recruited on the H3K18ub chromatin template (Figure 3.16F). First, a histone deacetylation complex formed and included SIN3A, SUDS3, SAP30 (BRMS1) and SAP130 with possible assoaciation of FOXK1 and FOXK2. Second, a heterochromatinisation complex was revealed including subunits of the dimerisation partner, RB-like, E2F and multivalval class B (DREAM) complex: LIN9 and LIN54 as components of the muvB complex, E2F3, transcription factor dimerisation protein 1 (TFDP1) and retinoblastoma 1 (RB1).

The Zn-finger containing Usp3 did not connect with any other enriched factor from the H3K18ub interactome list, which might suggest that it could contact the ubiquitylated chromatin directly (Figure 3.16F).

Table 3.1: Enriched proteins across all affinity purification experiments. T-test significance.

H2B K120ub histone	H2B K120ub mono	H2B K120ub chromatin	H2B K34ub histone	H2B K34ub chromatin	H3 K18ub chromatin	H3 K23ub chromatin	H3 K18/23ub2 chromatin
UBTF	INTS3	INTS3	SLC4A1	POLE3	DNMT1	PHF14	DNMT1
NAP1L1	INTS6	INTS4	SUB1	TGIF1	USP3	RPL7	USP7
TCOF1	RNF169	INTS2	ZRANB2	PGK1	USP7	DNMT1	SCML2
NAP1L4	POLR2A	CPSF3L	LUC7L2	FSCN1	MAX	RPL12	ACACA
LEO1	INTS1	INTS10	ARL6IP4	PPA1	SCML2	RPL4	ARID4B
RPL24	PPP1R10	INTS5	SRRM2	HMGB1	MYC	RPL7A	MLXIPL
PAF1	INTS12	RNF169	PLAA	NME2	ZBTB14	RPL6	SIN3A
CDC73	RBM15	BP2	DDX46	INO80C	MNT	ACACA	MNT
RPS3A	TBL1XR1	INIP	SF3B1	CHRAC1	WDR43	G2E3	FOXK1
GNL2	KIAA1429	INTS1	PHF5A	TPI1	HES1	PELP1	FOXK2
CSNK2A1	NCOR1	POLR2C	PUF60	ACTR5	FOXK1	USP7	SAP130
WDR61	ZC3H13	INTS6	UBTF	PRDX6	NFRKB	HMCES	ZBTB14
RPL13	TOX4	ASUN	U2AF1	FUBP1	SIN3A	ANKRD32	ZBTB21
CWC22	BCOR	DCAF7	U2AF2	SET	FOXK2	WDR18	ARID4A
DIEXF	SMC1A	DYRK1A	PCMT1	PPIA	HMBOX1	RAD18	TONSL
ABCF1	ZMYND8	INTS12	BCAR1	PSAT1	ACTR5	TEX10	SAP30
RPL18	SMC3	ZMYND8	TK1	KMT2B	GTF2IRD1	SCML2	LM
RPL17	NOP2	INTS7	CHCHD2	ZFP64	SAP130		SMC1A
TCEB3	G3BP1	POLR2B	ACTA1	BAZ1A	E2F3		SMC3
RPL35	DDX21	INTS8	ANXA2	RTN4	RB1		KMT2B
ZFR	RAD18	ZNF687		EEF1D	L3MBTL2		
RPL32	ZNF687	POLR2E		EIF4A1	H2AFV/Z		
BRD2	MYBBP1A	C7orf26		CFL1	EP400		
RPL13A	KDM2B	PPP2R1A		PSMA5	SUDS3		
RPL4	RSL1D1	INTS9		EIF5A	ACACA		
RPL7	T10	TBL1XR1		AHK	SAP30		
RPL23A	CPSF1	SLX4		TFAP4	MYBL2		
NPM1	HNRNPU	ZNF609		INO80B	MLXIPL		
RPL18A	UBTF	ERCC1		HSP90AA1	ACTR8		
RPL29	GNL2	PPP2CA		EIF3A	LIN54		
RPL35A	PABPC1	VWA9		MLXIP	MLX		
RPL26	NIPBL	SUPT5H		PSMD11	TFDP1		
NCL	NCL	RAD18		ASS1	KIAA1958		
RPL5	SRRT	RECQL5		POLR1C	ZBTB44		
RPL27	CDC5L	ERCC4		STRAP	TRRAP		
CSNK2A2	DDX23	NELFB		CLTC	KLF5		
CSNK2B	DDX5	POLR2A		POLR2H	RAD18		
EIF3B	MATR3	ANKRD32		ASH2L	KMT2B		
RPL10A	GCN1L1	POLR2G		HSP90AB1	INO80		
RPL7A	EFTUD2	BCOR		MEN1	BEND3		
RPL8	ABCF1	KP1		CBX3	MLXIP		
RPL6	PRPF6	PCGF1		PSMC2	ARNTL		
RPL3	THOC5	TOX4		BAG6	LIN9		
RPS14	RRP12	USP7		LDHB	BCOR		

Table 3.2: Enriched proteins across all affinity purification experiments. H/L cutoff selection.

H2B K120ub histone	H2B K120ub mono	H2B K120ub chromatin	H2B K34ub histone	H2B K34ub chromatin	H3 K18ub chromatin	H3 K23ub chromatin	H3 K18/23ub2 chromatin
UBTF	INTS3	PKP2	SLC4A1	POLE3	DNMT1	DNMT1	DNMT1
NAP1L1	INTS6	INTS3	SUB1	TGIF1	USP3	ACACA	USP7
TCOF1	RNF169	INTS4	ZRANB2	HMGB1	USP7	G2E3	SCML2
NAP1L4	INTS10	INTS2	LUC7L2	INO80C	SCML2	USP7	ACACA
	POLR2A	CPSF3L	ARL6IP4	CHRAC1	ZBTB14	HMCES	MLXIPL
	INTS1	INTS10	CTR9	ACTR5	MNT	SCML2	FOXK2
	POLR2C	INTS5	PLAA		HES1		
	POLR2B	RNF169			FOXK1		
	PPP1R10	NABP2			SIN3A		
	INTS12	INIP			FOXK2		
	TBL1XR1	INTS1			SAP130		
	ZNF592	POLR2C			E2F3		
	NCOR1	ZNF592			RB1		
	PDS5B	INTS6			SUDS3		
	TOX4	ASUN			ACACA		
	SMC1A	DCAF7			SAP30		
	ZMYND8	DYRK1A			BRMS1		
	SMC3	INTS12			MYBL2		
	ZNF687	ZMYND8			PARD3		
		INTS7			ARID4A		
		POLR2B			MLXIPL		
		INTS8			LIN54		
		ZNF687			MLX		
		POLR2E			TFDP1		
		C7orf26			KIAA1958		
		PPP2R1A			ZBTB44		
		INTS9			MLXIP		
		TBL1XR1			DDHD1		
		SLX4			LIN9		
		ZNF609			RFX5		
		ERCC1			YWHAE		
		PPP2CA			ZBTB1		
		VWA9			TLE3		
		SUPT5H					
		PPP2R1B					
		NELFCD					
		RECQL5					
		ERCC4					
		NELFB					

Table 3.3: Enriched proteins across all affinity purification experiments. Intersection.

H2B K120ub histone	H2B K120ub mono	H2B K120ub chromatin	H2B K34ub histone	H2B K34ub chromatin	H3 K18ub chromatin	H3 K23ub chromatin	H3 K18/23ub2 chromatin
UBTF	INTS3	INTS3	SLC4A1A	POLE3	DNMT1	DNMT1	DNMT1
NAP1L1	INTS6	INTS4	SUB1	TGIF1	USP3	ACACA	USP7
TCOF1	RNF169	INTS2	ZRANB2	HMGB1	ACACA	G2E3	SCML2
NAP1L4	POLR2A	CPSF3L	LUC7L2	INO80C	DDHD1	USP7	ACACA
	INTS1	INTS10	ARL6IP4	CHRAC1	SCML2	HMCES	MLXIPL
	PPP1R10	INTS5	CTR9	ACTR5	RFX5	SCML2	FOXK2
	INTS12	RNF169	PLAA		USP7		
	TBL1XR1	BP2			ZBTB14		
	NCOR1	INIP			ZBTB44		
	TOX4	INTS1			E2F3		
	SMC1A	POLR2C			ZBTB1		
	ZMYND8	INTS6			FOXK2		
	SMC3	ASUN			MNT		
	ZNF687	DCAF7			MYBL2		
		DYRK1A			MLXIPL		
		INTS12			RB1		
		ZMYND8			FOXK1		
		INTS7			LIN54		
		POLR2B			PARD3		
		INTS8			TFDP1		
		ZNF687			MLX		
		POLR2E			SIN3A		
		C7orf26			SUDS3		
		PPP2R1A			HES1		
		INTS9			KIAA1958		
		TBL1XR1			SAP130		
		SLX4			BRMS1		
		ZNF609			LIN9		
		ERCC1					
		PPP2CA					
		VWA9					
		SUPT5H					
		RECQL5					
		ERCC4					
		NELFB					

3.3 DNMT1, Usp7 and SCML2 cross-talk on the H3 ubiquitylated chromatin

3.3.1 DNMT1 recruits Usp7 and SCML2 *ex vivo* to H3 ubiquitylated chromatin

Chromatin affinity purification coupled with mass spectrometry indicated that DNMT1, Usp7, SCML2 and ACACA bind the three different H3 ubiquitylated templates (Figure 3.15). STRING analysis pointed out that DNMT1 may be found at the center of the H3K18ub interactome, which may help to recruit additional proteins and protein complexes to the modified chromatin. DNMT1 was shown to have a strong association with Usp7, which in turn was found to associate weakly with SCML2 (Figure 3.16). No direct association was evident between ACACA and any of DNMT1, Usp7 or SCML2, but a couple of possible indirect paths from ACACA towards DNMT1 were suggested during the STRING analysis.

To validate the mass spectrometry identifications and the statistical enrichment analysis upon which the STRING maps were drawn, fresh HeLa nuclear extracts were prepared and a series of western blot (WB) analyses was performed (Figure 3.17).

WB analysis was performed with antibodies against H3 histone, proliferating cell nuclear antigen (PCNA), SCML2, ubiquitin-like containing plant homeodomain and really interesting new gene finger 1 (UHRF1), Usp7, DNMT1 and ACACA to validate both antibody specificity and preparation of the nuclear extract (Figure 3.17A). PCNA and UHRF1 were also investigated even though they were not enriched in any of the interactomes. Previous biochemical evidence linked PCNA loading at the replication fork and UHRF1's ability to catalyse the ubiquitylation of the H3 N-terminal histone tail with DNMT1 recruitment [188], [189], [35], [36].

Ubiquitylated histone affinity matrices were created by immobilising unmodified or ubiquitylated histones with the help of a C-terminal specific anti-H3 antibody to a magnetic resin (Figure 3.17B). The fresh nuclear extract was incubated with the histone matrices and the eluates were subjected to western blot analysis. The analysis indicated that during incubation with the nuclear extract, DNMT1 was enriched on the H3K18ub and H3K18/23ub2 and, to a lower extent, to the H3K23ub histone templates. Both Usp7 and UHRF1 bound unspecifically to the mock control, making it difficult to argue whether their binding to the affinity matrices was ubiquitylation-specific. SCML2 did not bind any of the matrices.

Six different chromatin templates were used in affinity purification experiments (Figure 3.17C). Next to the control unmodified chromatin and an acidic patch H2A mutant (H2Amt) chromatin, four ubiquitylated templates were prepared. Besides the three templates used in the mass spectrometry identification experiments, a H3K23ub H2Amt template was also prepared. The two H2A mutant chromatin templates were added to understand if the nucleosome acidic patch were a putative interaction surface for any of the factors enriched on the H3 ubiquitylated chromatin.

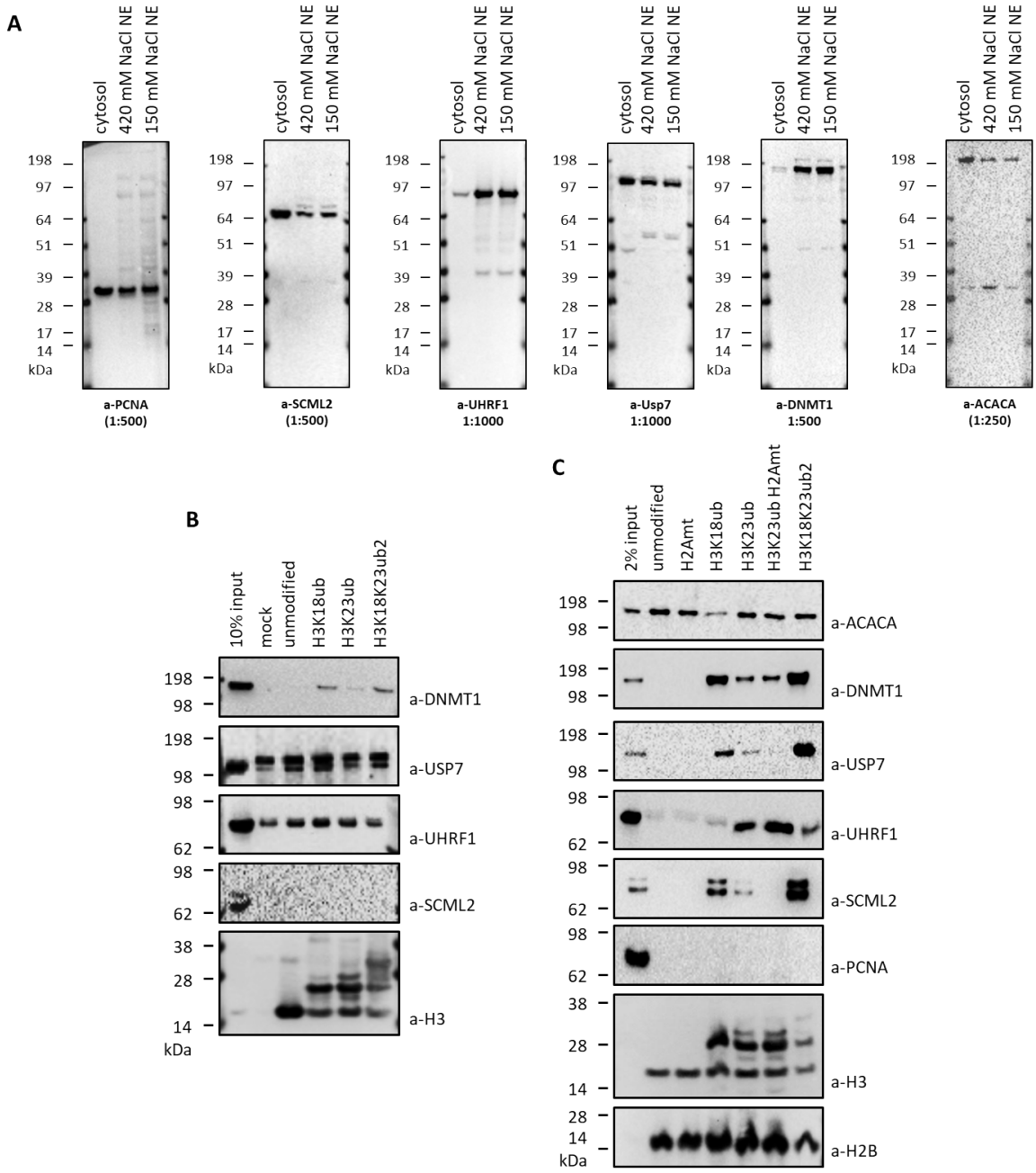


Figure 3.17: Western blot validation of the H3ub interactors identified by mass spectroemtry (A) Preparation of fresh HeLa nuclear extracts and validation of primary antibodies used for detection of H3ub interactors. (B) Histone affinity purification from HeLa nuclear extract using unmodified H3, H3K18ub, H3K23ub, and H3K18/23ub2 histones. Western blots analysis of the elution fractions from the different affinity purification experiments was performed for Usp7, SCML2 and H3. (C) Chromatin affinity purification from HeLa nuclear extract using biotinylated chromatin containing unmodified, H2Amt, K18ub, K23ub, K23ub H2Amt and K18/23ub2 histones. Western blots analysis of the elution fractions from the different affinity purification experiments was performed for ACACA, DNMT1, Usp7, UHRF1, SCML2, PCNA, H3 and H2B.

There was noticeable deubiquitylation of the chromatin arrays during the incubation with the nuclear extract. This was indicated in the WB analysis by the accumulation of free H3 histone in the H3K18ub, H3K23ub and H3K23ubH2Amt samples and by the conversion of the H3ub2 mark to H3ub1 and free H3 histone in the H3K18/23ub2 sample. There was no evidence for binding of PCNA to any of the chromatin affinity purification matrices. Both the short and the long isoform of SCML2 were enriched on the ubiquitylated chromatin arrays, with preference towards H3K18/23ub2 and H3K18ub. SCML2 was not recovered on the H3K23ubH2Amt array. UHRF1 bound preferentially to the chromatin templates containing H3K23ub. Usp7 was enriched strongest on the double-modified H3K18/23ub2, and the H3K18ub arrays. Usp7 recruitment, like that of SCML2 was lost in the H3K23ub H2Amt sample. DNMT1 bound strongest to H3K18/23ub2 and to H3K18ub, and associated weakly with H3K23ub and H3K23ubH2Amt. ACACA bound all chromatin templates, forming a weaker association with the H3K18ub chromatin array.

The western blot analyses validated the mass spectrometric findings that SCML2, Usp7 and DNMT1 were specific for the ubiquitylated chromatin matrices. ACACA's binding to the mock control makes it difficult to argue if the protein prefers ubiquitylated over unmodified chromatin. The western blot analyses also suggested that PCNA is not involved in the events occurring on the H3ub chromatin and that UHRF1 bound ubiquitylated chromatin, specifically H3K23ub and doubly modified H3K18/23ub2 chromatin, which was not observed during the mass spectrometry analysis. The use of the K23ubH2Amt arrays showed that the chromatin targeting of both Usp7 and SCML2 was sensitive to the mutation of the nucleosome acidic patch.

Using histones as affinity tags, only DNMT1 was specifcally enriched by H3 ubiquitin. Using chromatin arrays as affinity tags, SCML2, UHRF1, Usp7 and DNMT1 were all preferentially recruited to the ubiquitulated templates. This suggested that SCML2, UHRF1 and Usp7 needed additional surfaces on the nucleosome which were not present on H3 ubiquitin to be able to interact with the modified histone.

Inhibition of DNA methylation and histone deubiquitylation

To better understand the order of recruitment of H3ub interactors to chromatin, nuclear extracts were treated with inhibitors to separate the two main events occuring on this template (Figure 1.6; Figure 3.18). First, DNA methylation was blocked with S-adenosyl homocysteine (SAH) treatment [190]. Second, deubiquitylation was slowed using ubiquitin vinyl sulfone (UVS) treatment [93].

Treatments were performed in parallel to untreated controls to minimise experimental variations in terms of recruitment reproducibility. The western blot analysis performed on the eluates of the SAH-treated samples showed that the deubiquitylation of histone H3 was unaffected by the DNA methylation block (Figure 3.17C, Figure 3.18A). SCML2 recruitment to ubiquitylated chromatin was lower in the treated samples than in the untreated controls. Both isoforms were affected equally by the SAH treatment, an observation which held on all

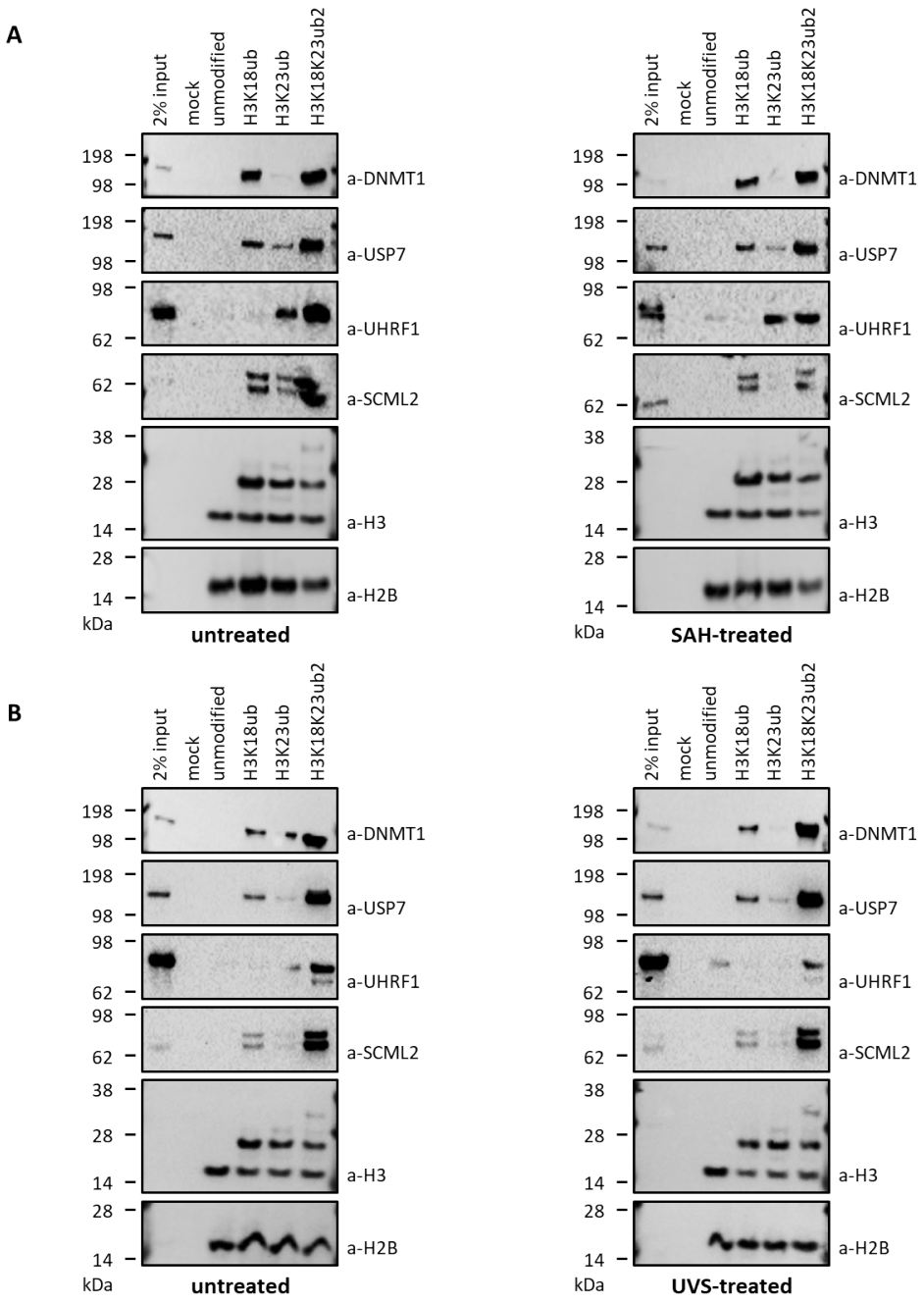


Figure 3.18: Inhibition of DNA methylation and H3 deubiquitylation in the HeLa nuclear extract affects recruitment of several factors to H3ub chromatin. Chromatin affinity purification from HeLa nuclear extract using biotinylated nucleosomal arrays containing unmodified, K18ub, K23ub and K18/23ub2 histones. Western blots analysis of the elution fractions from the different affinity purification experiments was performed for DNMT1, Usp7, UHRF1, SCML2, H3 and H2B. (A) Inhibition of DNA methylation was achieved with 100 μM S-adensoyl Homocysteine (SAM) treatment. (B) Inhibition of global deubiquitylation was performed by treatment of the nuclear extract with 10 μg/mL recombinant ubiquitin vinyl sulfone (UVS) treatment.

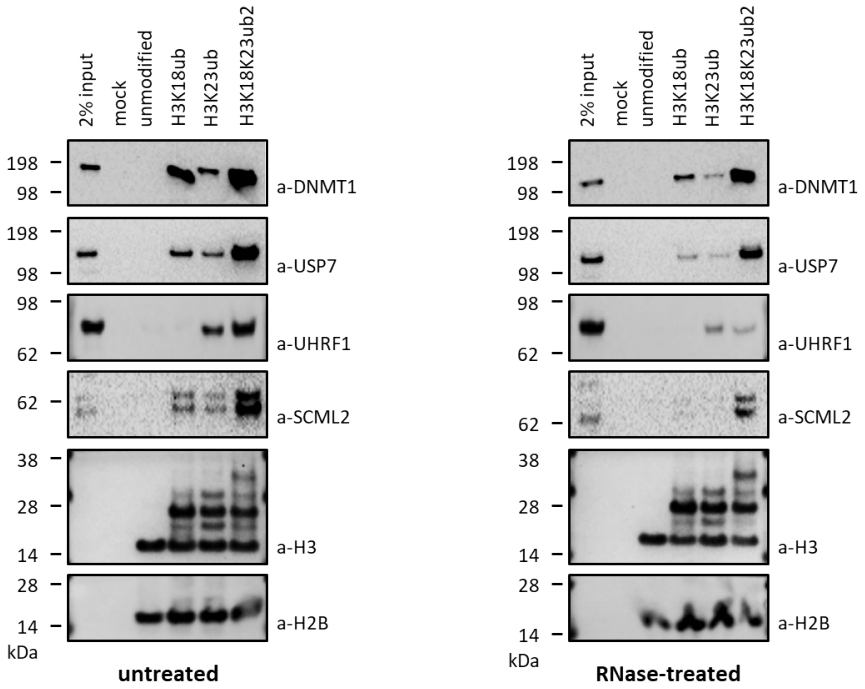


Figure 3.19: Degradation of RNA from nuclear extract results in decreased binding of several interactors to chromatin. Chromatin affinity purification from HeLa nuclear extract using biotinylated nucleosomal arrays containing unmodified, K18ub, K23ub and K18/23ub2 histones. Western blots analysis of the elution fractions from the different affinity purification experiments was performed for DNMT1, Usp7, UHRF1, SCML2, H3 and H2B. Degradation of RNA was performed by treating the nuclear extract with 50 μg/mL recombinant heat-inactivated ribonuclease A (RNase A).

ubiquitylated chromatin templates. UHRF1 recruitment to H3K23ub and H3K18/23ub2 was not affected by the SAH treatment. Similarly, DNMT1 and Usp7 did not show differences in recruitment after the treatment to inhibit DNA methylation.

To inhibit Usp7 activity, exracts were treated with UVS prior and during incubation with the chromatin templates (Figure 3.18B). As shown in the western blot analysis, H3 deubiquitylation remained largely unaffected after UVS treatment. Similarly, SCML2 and DNMT1 recruitment to all three ubiquitylated templates did not change. Interestingly, UHRF1 recruitment from the treated extracts suffered some modifications. UHRF1 bound unmodified chromatin and was recruited less to the H3K18/23ub2 template. Usp7 was recruited in higher amounts to the H3K23ub and the H3K18/23ub2 templates. Neither the SAH nor the UVS treatment of the nuclear extracts resulted in clear inhibition of H3 deubiquitylation.

SCML2 and UHRF1 are sensitive to the removal of nuclear RNA

SCML2 has previously been shown to have the ability to bind RNA, a property which was thought to play an essential role in its chromatin targeting [100]. To address if SCML2 recruitment to chromatin may be mediated by RNA, the nuclear extract was treated with

ribonucleas A (RNase A) prior to the chromatin affinity purification (Figure 3.19). While SCML2 recruitment decreased after RNase treatment, so did the recruitment of DNMT1 and to a lesser extent that of Usp7. Interestingly, UHRF1 revocery on the H3K23ub and H3K18/23ub2 chromatin templates was also affected by the RNase treatment. Deubiquitylation proceeded slower in the H3K18ub and H3K18/23ub2 samples in comparison to the untreated controls. This observation correlated with the decreased recruitment of DNMT1, Usp7 and SCML2 onto these templates.

The treatments applied to the nuclear extract did not provide clear answers with respect to the order of recruitment of the different factors. However, the recovery of SCML2 in the SAH treatment experiment and the recovery of SCML2 and UHRF1 in the RNase treatment experiment were affected by the induced perturbations. This suggests that their recruitment and the release to and from the modified chromatin may be also regulated by mechanisms other than hierarchical protein-protein interactions.

3.3.2 Preparation of recombinant SCML2 truncations

To tackle the order of events on the H3 ubiquitylated chromatin from a different angle, DNMT1, Usp7 and SCML2 were produced recombinantly. While the interaction between DNMT1 and Ups7 is very well documented in the literature [92], the interaction between Usp7 and SCML2, as suggested by the STRING association (Figure 3.16) needed further biochemical characterisation [98].

Recombinant Usp7 was produced in insect cells and purified by affinity chromatography using an N-terminal His-tag (Figure 3.20D). Recombinant SCML2 was produced in bacteria or insect cells and purified by affinity chromatography using N-terminal GST- or His-tags (Figure 3.20B, (Figure 3.20C)). Several SCML2 truncations were presented throughout the thesis. To map the interaction between SCML2 and Usp7, several truncations were made in SCML2 which included SCML2's previously annotated domains: malignant brain tumor domains (MBT), RNA-binding region (RBR), domain of unknown function (DUF) and sex comb on midleg and polyhomeotic (SPM) (Figure 3.20A). With the exception of the bacterial DUF domain deletion, recombinant proteins were prepared in high amounts and purity both from the bacterial and the insect cell expression systems.

3.3.3 SCML2 recruits Usp7 to ubiquitylated chromatin arrays

RBR-DUF region connects SCML2 with Usp7

It has previously been shown that SCML2 interacts with Usp7 *in vitro* [98]. There the region between the MBT and DUF domains of SCML2 was shown to bind to the tumor necrosis facor receptor associated factor (TRAF) domain of Usp7. To better define the regions in SCML2 responsible for its interaction with Usp7, GST-tagged full-length (FL) SCML2 or SCML2 truncations were immobilised on a glutathione resin and FL His.Usp7 was added to each mixture (Figure 3.21). Usp7 did not bind the affinity matrix by itself (Figure 3.21A). Usp7 did not bind the matrix when incubated with GST recombinant protein (Fig-

ure 3.21C). FL SCML2 recovered Usp7 on the beads in 1:1 stoichiometry. None of the three initial SCML2 truncations tested could lose Usp7 recovery on the GST beads (Figure 3.21A).

A similar experiment was performed using His-tagged SCML2 constructs which were expressed in insect cells. Instead of the glutathione affinity resin, an antibody specific for SCML2's C-terminus was used to co-immunoprecipitate SCML2 and Usp7 (Figure 3.21B). The co-immunoprecipitation experiment suggested that the deletion of the DUF domain resulted in a substoichiometric recovery of Usp7. The DUF domain was thus found to be important, but insufficient for the full interaction between SCML2 and Usp7. Additional interaction surfaces in SCML2 were responsible for its binding to Usp7.

Two additional SCML2 truncations were made, which separated the protein in two halves. The first half contained the MBT domains and the RBR region (ΔC) and the second half contained the DUF and the SPM domains (ΔN). GST affinity purification experiments showed that both the ΔN and the ΔC truncations recruited substoichiometric amounts of

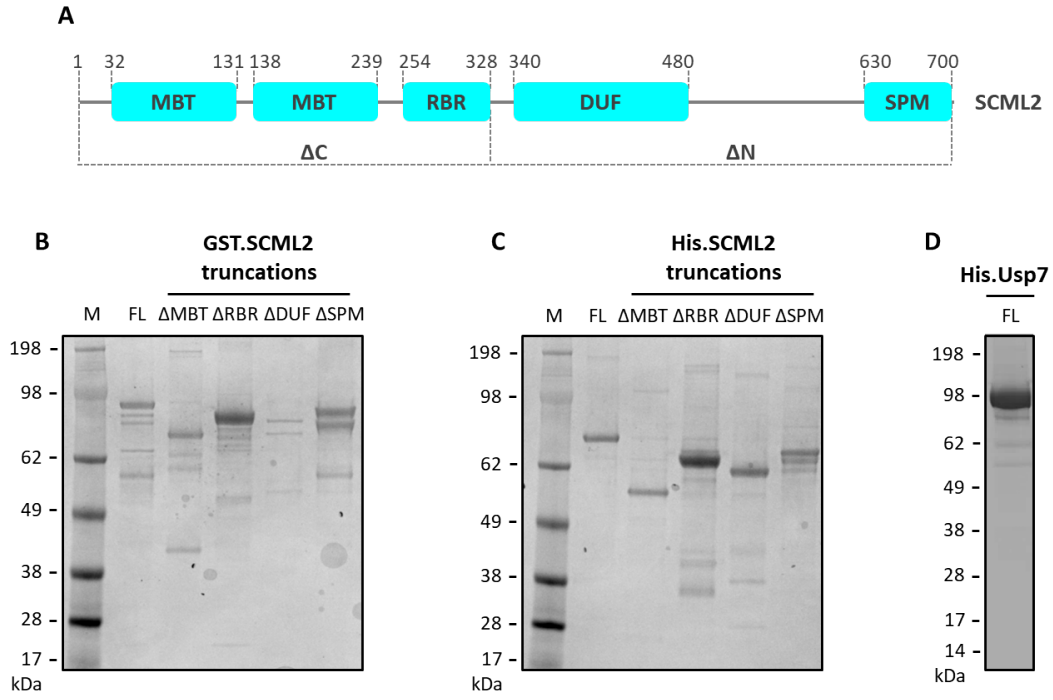


Figure 3.20: Purification of recombinant Usp7 and SCML2 proteins. (A) Schematic annotation of the domains and regions present in Sex-comb on midleg-like 2 (SCML2). (B) Coomassie-stained SDS-PAGE gel of affinity-purified full-length (FL) and truncated GST-SCML2 proteins from bacterial overexpressions. (C) Coomassie-stained SDS-PAGE gel of affinity-purified FL and truncated His-SCML2 proteins from insect cells infections. (D) Coomassie-stained SDS-PAGE gel of affinity-purified FL His.Usp7 protein from insect cells infections. MBT = malignant brain tumor; RBR = RNA-binding region; DUF = domain of unknown function; SPM = sex-comb on midleg and polyhomeotic.

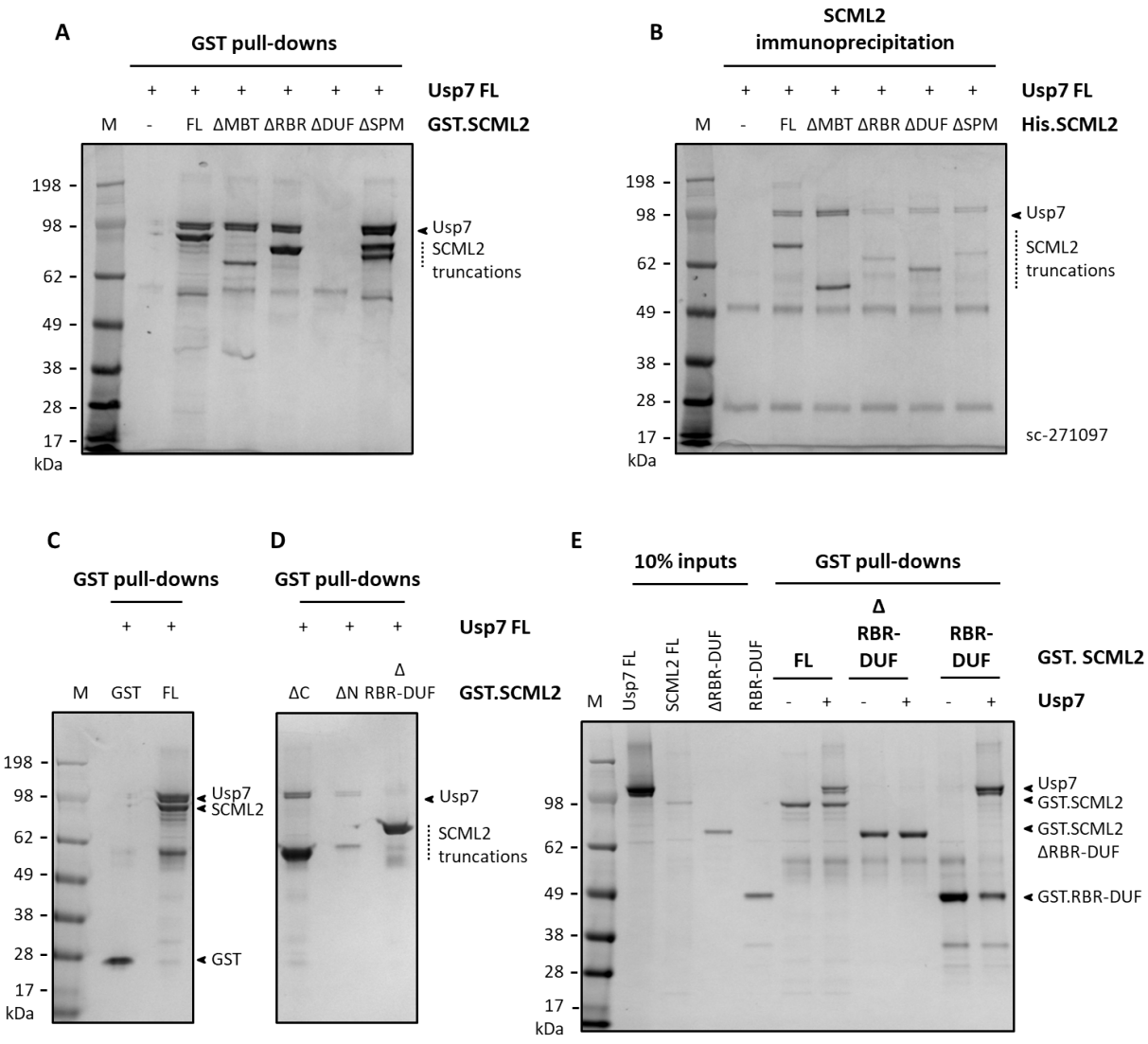


Figure 3.21: Mapping of SCML2 interaction surface responsible for Usp7 binding. (A) Coomassie-stained SDS-PAGE gel of GST affinity purifications of GST-tagged FL or truncated SCML2 proteins with FL Usp7. (B) Coomassie-stained SDS-PAGE gel of co-immunoprecipitation experiment using a commercial antibody against the C-terminus of SCML2 (sc271097) with FL or truncated His.SCML2 proteins and FL Usp7. (C) Coomassie-stained SDS-PAGE gel of GST affinity purification of free GST and GST-tagged FL SCML2 with FL Usp7. (D) GST-pull-downs of additional SCML2 truncations with FL Usp7. (E) Coomassie-stained SDS-PAGE gel of GST pull-downs of FL SCML2, SCML2 ΔRBR-DUF and GST.RBR-DUF constructs with FL Usp7.

Usp7 to the beads, indicating that there were interaction surfaces in both halves of the protein (Figure 3.21D). Knowing that the DUF domain was one of the contact points between SCML2 and Usp7, the two domains connecting the two SCML2 halves, namely the RBR region and the DUF domain, were truncated together. GST-SCML2 ΔRBR-DUF could not bind Usp7, suggesting that SCML2 required both domains for a stable interaction with Usp7.

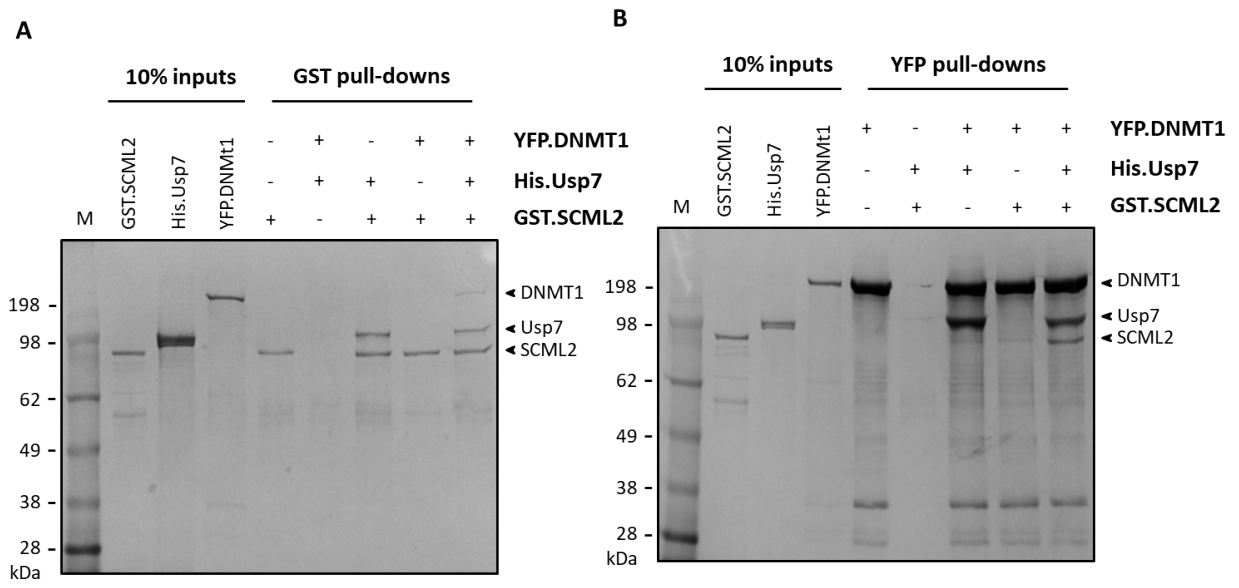


Figure 3.22: Interaction of SCML2 with Usp7 and DNMT1. (A) Coomassie-stained SDS-PAGE gel of GST affinity purifications of FL GST.SCML2 with FL His.Usp7 and/or FL His.YFP.DNMT1. (B) Coomassie-stained SDS-PAGE gel of YFP affinity purifications of FL His.YFP.DNMT1 with His Usp7 and or GST.SCML2.

To test if the RBR-DUF region was sufficient for the interaction between SCML2 and Usp7, the fragment was produced recombinantly with a N-terminal GST tag. Using the previously described affinity purification strategy, FL SCML2 recovered stoichiometric amounts of Usp7, His.SCML2 ΔRBR-DUF lost the interaction with Usp7 and GST.RBR-DUF recovered the lost interaction fully (Figure 3.21E).

The STRING analysis of the three H3 ubiquitylated chromatin arrays pointed out that SCML2 may be connected to DNMT1 through its association with Usp7. Two affinity purification schemes were used to test this hypothesis (Figure 3.22). First, FL GST.SCML2 was incubated with FL Usp7, FL DNMT1 or both Usp7 and DNMT1 on the glutathione resin. While SCML2 could not recruit DNMT1 when the two were incubated together, it did so, even though in a substoichiometric manner, in the presence of Usp7 (Figure 3.22A). In the reverse experiment, YFP-DNMT1 was incubated with FL SCML2, FL Usp7 or both SCML2 and Usp7 on a GFP-trap affinity matrix. While DNMT1 could not recruit SCML2 on its own, it did so in the presence of Usp7 (Figure 3.22B). SCML2, Usp7 and DNMT1 could be recovered together on beads but in substoichiometric amounts.

RBR region links SCML2 to mononucleosomes

The RBR domain of SCML2 was shown to bind RNA, a mononucleosome and a ternary complex containing both RNA and mononucleosome [100]. In parallel, the DUF domain of SCML2 was also shown to bind dsDNA, irrespective of its methylation status [101].

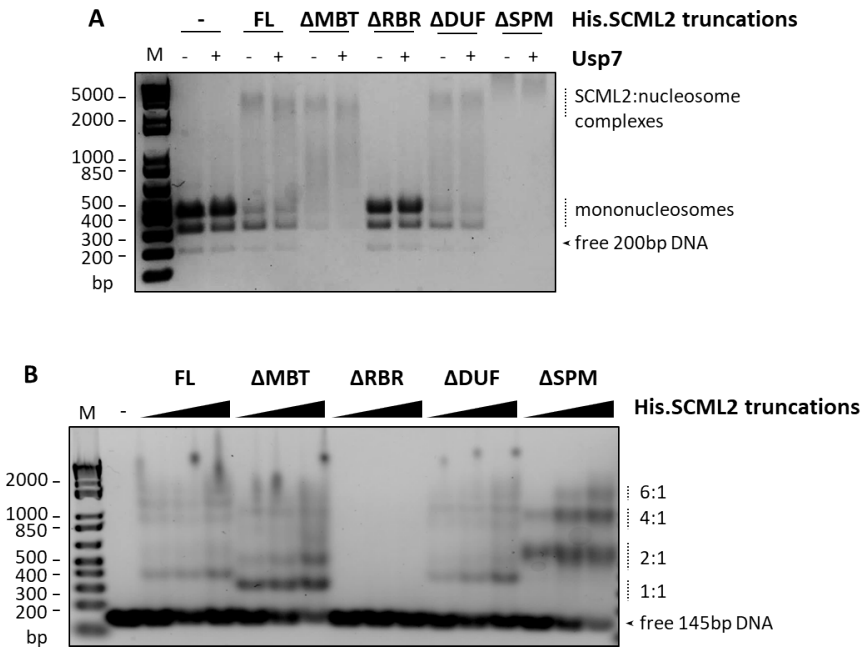


Figure 3.23: The RBR region links SCML2 to nucleosomes. (A) EtBr-stained native agarose gel documenting the interaction between an undersaturated 200 bp nucleosome and FL as well truncated His-tagged SCML2 constructs. FL Usp7 was added in alternate reactions. (B) EtBr-stained native agarose gel of a 145 bp DNA construct incubated with FL or truncated His.SCML2.

To understand which of the two domains was important for SCML2 chromatin targeting, we used the His.SCML2 truncations produced for the Usp7 interaction experiments in native gel shift assays (Figure 3.23A). Unsaturated 200 bp mononucleosomes were incubated with different SCML2 constructs in the presence or absence of Usp7. All SCML2 constructs bound the mononucleosome as indicated by their corresponding shifts in the native agarose gel, with the exception of SCML2ΔRBR. The addition of Usp7 did not produce noticeable supershifts, although the resolution of the native agarose might not have been optimal to separate that interaction.

Since SCML2 has the ability to bind not only mononucleosomes, but also dsDNA, the different SCML2 truncations were incubated in increasing concentration to DNA of 145 bp (Figure 3.23B). All SCML2 constructs, with the exception of SCML2ΔRBR, bound the DNA and shifted it. This pointed out that the SCML2 interaction with the nucleosomes might have been mediated by the linker DNA or by the DNA wrapped around the histone octamer core. If the FL SCML2 and SCML2ΔMBT and SCML2ΔDUF constructs bound the DNA in regular steps (the equivalent of 150 bp), the SCML2 construct lacking SPM bound the DNA at twice that interval. Likely, several SCML2 molecules bound the same DNA template. However it might be possible that one SCML2 protein bound two DNA molecules at once.

SCML2 binds nucleosomal DNA. To test if SCML2 prefers mononucleosomes with linker over linker-less nucleosome core particles, unmodified nucleosomes were reconstituted on 145

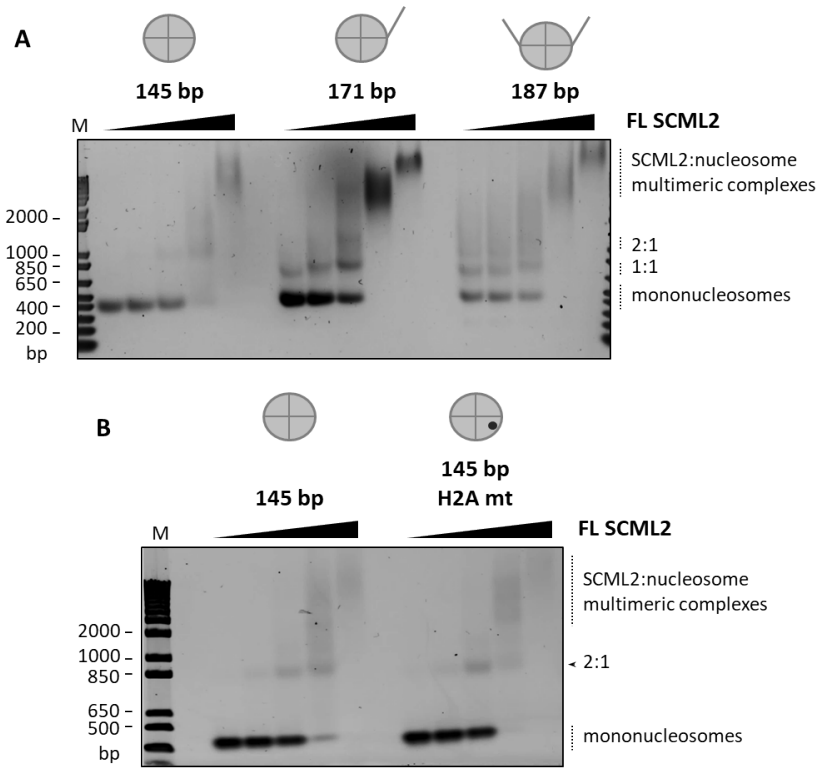


Figure 3.24: SCML2 binding to mononucleosomes is affected by linker DNA. (A) EtBr-stained native agarose gel of nucleosomes assembled on DNA templates of 145, 171 and 187 bp which were incubated with increasing amounts of FL SCML2. (B) EtBr-stained native agarose gel of unmodified or H2A mutant (H2A E61A/E64A) nucleosomes incubated with FL SCML2.

bp, 171 bp and 187 bp DNA templates. The 171 bp DNA is an asymmetric template, with 16 bp of linker DNA present downstream of the core positioning sequence. The 187 bp DNA is a symmetric template, with linker DNA present on both sides of the nucleosome. Increasing amounts of SCML2 were incubated with the unmodified nucleosome templates to obtain full saturation (Figure 3.24A). SCML2 bound all nucleosome templates, with a slight preference for the asymmetric 171 bp construct. Interestingly, while SCML2 bound to the nucleosome with linker DNA in the equivalent of 300 bp steps, the 145 bp nucleosome was shifted to twice that distance in the presence of SCML2. Bearing in mind the observation that SCML2ΔSPM bound free DNA at intervals twice as large as the FL protein, it could be that the linker DNA might be needed for the SPM domain to induce a conformational change in the full-length protein that allowed two interaction surfaces from within the RBR domain to contact the nucleosome core.

SCML2 binds nucleosomes in the absence of linker DNA. To test if SCML2 binds on the nucleosome core, two mutations were prepared in the nucleosome acidic patch. 145 bp nucleosome core particles were reconstituted with unmodified or H2A mutant octamers which were previously shown to deter binding of proteins to the nucleosome [72]. SCML2 was able

to bind a 145 bp nucleosome core particle with a mutated acidic patch (Figure 3.24B), . Both the unmodified and the mutant nucleosome core particles were shifted at twice the distance observed with the 171 bp and 187 bp constructs.

SCML2 and DNMT1 recruit Usp7 to unmodified chromatin

Whilst FL SCML2 has the ability to bind nucleosomes, Usp7 does not (Figure 3.23A). SCML2 interacts in 1:1 stoichiometry with Usp7 and may be able through this interaction to recruit Usp7 to nucleosomes or to nucleosomal arrays. Chromatin affinity purification assays were performed on unmodified chromatin using recombinant FL Usp7 and FL SCML2 as well as all previously described truncations used to map the interaction between the two proteins (Figure 3.25A). By itself, Usp7 could not bind nucleosomal arrays. FL SCML2 could recover Usp7 to chromatin in 1:1 stoichiometry. Deletion of the MBT, preSPM or the SPM domains had no effect on the equimolar recovery of SCML2 and Usp7 to chromatin. SCML2ΔRBR was shown not to be able to bind free DNA or interact with nucleosomes (Figure 3.23A, Figure 3.23B). In consequence, removal of the RBR domain resulted in the loss of SCML2 and Usp7 binding to nucleosomal arrays. Additionally, SCML2ΔC and SCML2ΔRBR-DUF, which lacked the RBR domain, could not bind to the nucleosomal arrays and were not able to recruit Usp7 to chromatin. Interestingly, the SCML2ΔC truncation, which bound nucleosomal arrays, showed a severe loss of Usp7-mediated recruitment, pointing out the necessity of a continuous RBR-DUF surface for stable interaction between SCML2 and Usp7. To verify if the RBR-DUF region was sufficient for SCML2-mediated Usp7 chromatin targeting, FL SCML2, SCML2ΔRBR-DUF and GST.RBR-DUF constructs were used in similar chromatin pull-down experiments with FL Usp7 (Figure 3.25B). GST.RBR-DUF rescued the Usp7 recruitment lost with the SCML2ΔRBR-DUF truncation.

SCML2 interacts with Usp7 in 1:1 stoichiometry and can recruit it to chromatin. DNMT1 has the ability to bind both nucleic acid and interact with Usp7 (Figure 3.22B). To test if DNMT1 can target Usp7 to chromatin, chromatin pull-downs were performed as before with FL recombinant SCML2 and Usp7 or FL DNMT1 and FL Usp7 (Figure 3.26A). Usp7 could not by itself bind the nucleosomal arrays, but when incubated with either SCML2 or DNMT1 it was recruited in 1:1 stoichiometry to chromatin. Interestingly, when incubated with both SCML2 and DNMT1, Usp7 mananged to stabilise an equimolar trimeric complex on the unmodified chromatin, which was not the case when the recombinant proteins were mixed in the absence of chromatin (Figure 3.22). To test if the observed complex was a stable trimeric interaction or a mixture of two distinct complexes, where Usp7 is recruited to chromatin separately by SCML2 or DNMT1, increasing amounts of SCML2 were titrated into a 1:1 Usp7:DNMT1 complex on unmodified chromatin (Figure 3.26B). Addition of SCML2 recovered more Usp7 on the chromatin, but the effect was not controlled by the SCML2 concentration, but rather by the availability of DNMT1. Excess SCML2 bound chromatin separate of Usp7 and DNMT1 with formation of SCML2 homodimers. The reverse experiment, in which increasing amounts of DNMT1 were titrated into an equimolar SCML2:Usp7 complex on unmodified chromatin, pointed out that excess DNMT1 also bound chromatin separate of the other two recombinant factors (Figure 3.26C). We deduced that SCML2, Usp7 and DNMT1 formed a stable trimeric complex on unmodified nucleosomal arrays.

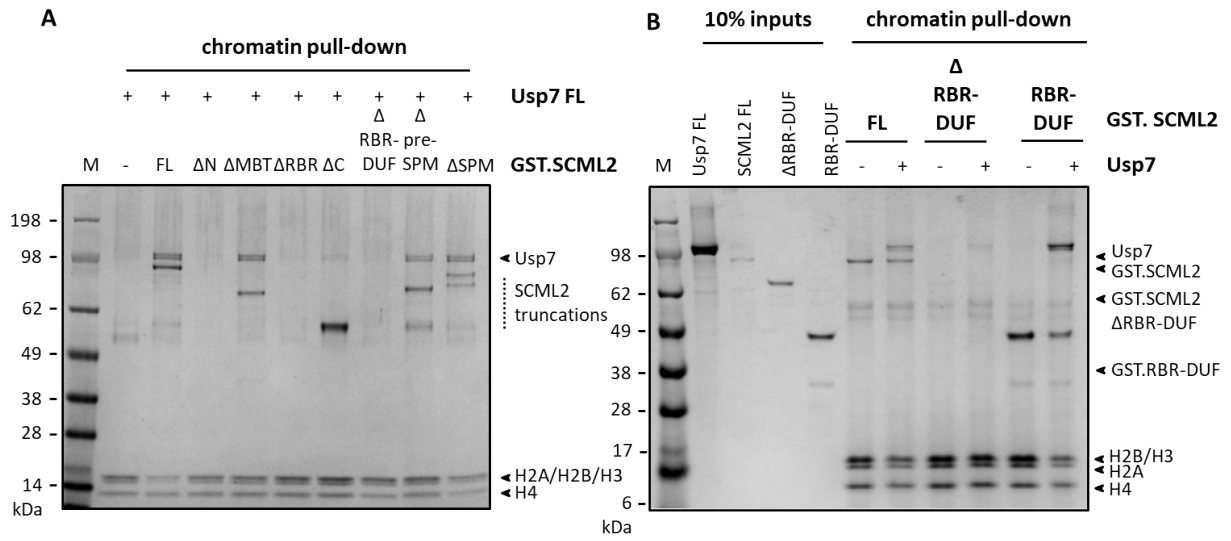


Figure 3.25: SCML2 recruits Usp7 to chromatin (A) Coomassie-stained SDS-PAGE gel of chromatin affinity purifications using biotinylated unmodified nucleosome arrays with GST.SCML2, His.Usp7 and complete series of His.SCML2 truncations. (B) Coomassie-stained SDS-PAge of chromatin affinity purifications using biotinylated unmodified nucleosome arrays with GST.SCML2, His.Usp7 and FL or ΔRBR-DUF or GST.RBR-DUF constructs.

SCML2 and DNMT1 recruit Usp7 to H3 ubiquitylated chromatin

Recombinant SCML2, Usp7 and DNMT1 form a stable complex on unmodified chromatin (Figure 3.26). The initial chromatin affinity enrichment protocols indicated however that SCML2, Usp7 and DNMT1 were enriched on H3 ubiquitylated over unmodified chromatin arrays in both the mass spectrometry and the western blot analyses (Figure 3.15; Figure 3.17C).

To address if SCML2, DNMT1 and Usp7 are recruited specifically to ubiquitylated chromatin, the recombinant proteins were incubated alone or in various combinations with six different types of chromatin arrays. These included unmofidied, H2Amt, ubiquitylated H3K18ub, H3K23ub and H3K18/23ub2 as well as H3K23ub H2Amt chromatin templates (Figure 3.27). SCML2 bound all six chromatin templates, with decreased recovery on the two H2A mutant matrices (Figure 3.27A). Usp7 was not able to form stable interactions with any chromatin templates and was not recruited by any of the templates. Usp7 may have transiently bound the modified chromatin templates as these underwent partial deubiquitylation during incubation with the protease (Figure 3.27B). DNMT1 bound all chromatin templates, with decreased recovery on the H2A mutant, but strong enrichment on all ubiquitylated templates. DNMT1 was enriched strongest on the doubly modified H3K18/23ub2 chromatin, then on the H3K18ub and then on the H3K23ub templates (Figure 3.27C). When incubated together with SCML2, Usp7 was recovered on all chromatin templates. There was no preference for the Usp7:SCML2 complex for one or the other matrix. In the presence of SCML2, Usp7 deubiquitylated all modified chromatin arrays to completion (Figure 3.27D). In the presence of DNMT1, Usp7 was recovered on all chromatin arrays. Both recombinant

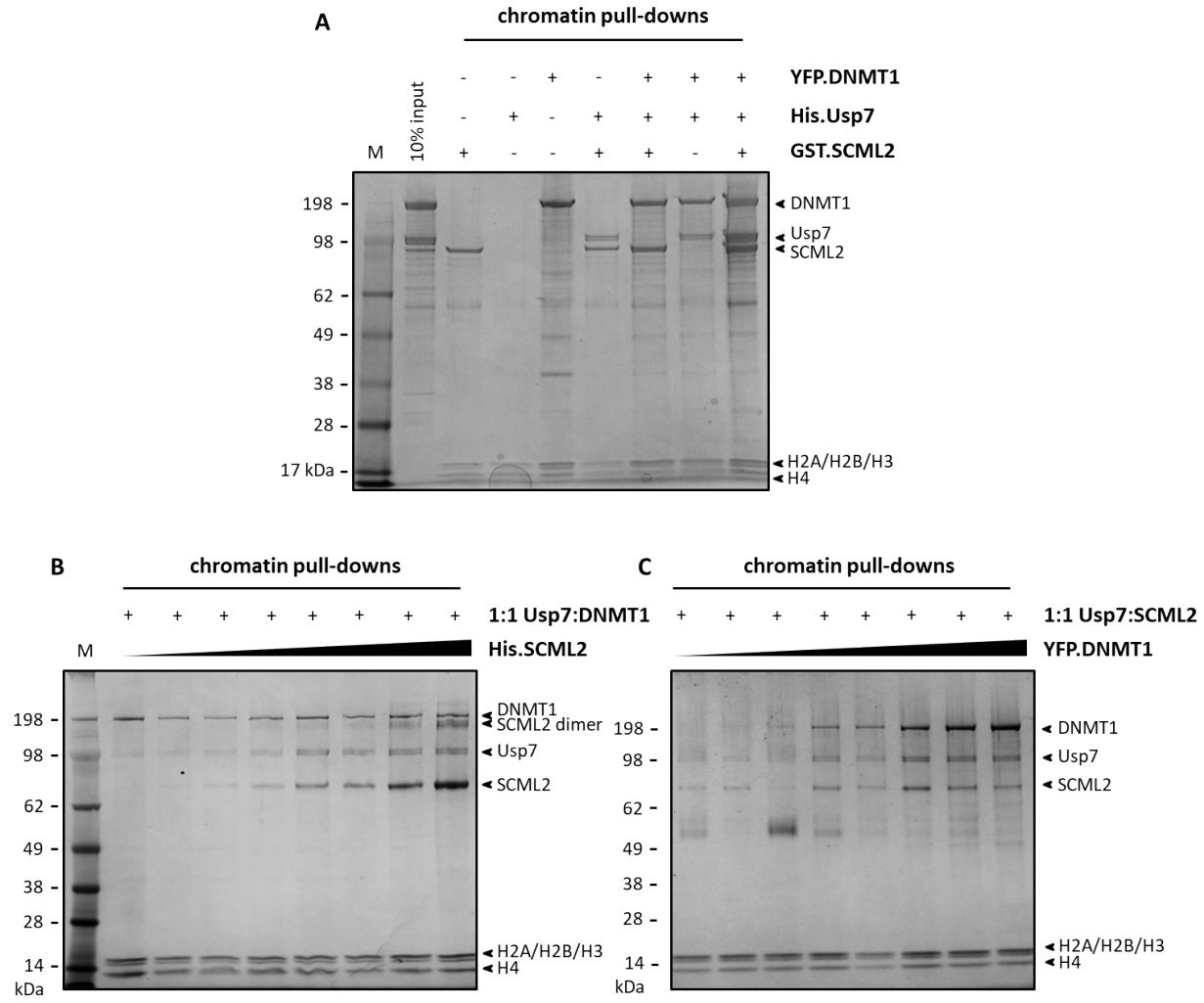


Figure 3.26: SCML2 and DNMT1 recruit Usp7 to chromatin. (A) Coomassie-stained SDS-PAGE gel of chromatin affinity purifications using biotinylated unmodified nucleosome arrays and GST.SCML2, His.Usp7 and His.YFP.DNMT1. (B) Coomassie-stained SDS-PAGE gel of chromatin affinity purifications using unmodified nucleosome arrays with stoichiometric His.Usp7 and His.YFP.DNMT1 and increasing concentraitons of His.SCML2. (C) Coomassie-stained SDS-PAGE gel of chromatin affinity purifications using unmodified nucleosome arrays with stoichiometric His.Usp7 and His.SCML2, but increasing concentraitons of His.YFP.DNMT1.

proteins show decreased recruitment to the H2A mutant template and increased binding to the H3K18/23ub2 matrix. DNMT1 did not influence Usp7's deubiquitylation activity (Figure 3.27E). When incubated together, DNMT1, Usp7 and SCML2 were all recovered on the six chromatin arrays. There was no clear preference for ubiquitylated versus unmodified chromatin templates, though all recombinant proteins were recruited less to the acidic patch mutant. Within the given experimental conditions, in the presence of DNMT1, SCML2 could not stimulate Usp7's deubiquitylation function (Figure 3.27F).

SCML2's interaction with Usp7 proved important in targeting Usp7 to H3 ubiquitylated

chromatin. Interestingly, SCML2 stimulated Usp7's deubiquitylation function and this effect was lost when DNMT1 was also present in the reaction mixture. To better understand how the SCML2 activation was brought about, several SCML2 truncations were incubated with FL Usp7 on ubiquitylated chromatin in affinity purification experiments (Figure 3.28). Removal of the RBR domain was shown before to lose SCM2 binding to DNA, mononucleosome and unmodified chromatin arrays (Figure 3.23, Figure 3.25). SCML2 ΔRBR could not bind any ubiquitylated template and could not target Usp7 to chromatin. All ubiquitylated arrays remained unchanged after incubation with the Usp7 protease (Figure 3.28A). Removal of the DUF domain was shown to be important for contacting Usp7 and for targeting it to unmodified nucleosomal arrays (Figure 3.23, Figure 3.25). SCML2 ΔDUF targeted Usp7 to ubiquitylated chromatin arrays in substoichiometric amounts, but this targeting was sufficient to activate the enzyme which deubiquitylated all of the H3K18ub and H3K18/23ub2 and parts of the H3K23ub templates. This suggested that the removal of the H3K18ub mark may have been more prone than that of H3K23ub (Figure 3.28B).

Deletion of the entire RBR-DUF region was previously shown to lose both the interaction between SCML2 and Usp7 and SCML2's chromatin binding completely (Figure 3.21E, Figure 3.25A). During incubation of SCML2ΔRBR-DUF with Usp7 on the ubiquitylated chromatin arrays, no stable interaction could be preserved and none of the two proteins could be recovered from the beads (Figure 3.28C). Also, the modified nuclesomal arrays suffered marginal deubiquitylation. Upon incubation with GST.RBR-DUF chromatin targeting to the modified nucleosomal arrays of Usp7 was restored. All modified templates were completely deubiquitylated. Recovery of GST.RBR-DUF and Usp7 was considerably lower on the H2A mt chromatin arrays, suggesting that additional domains in SCML2 may have been important for SCML2 stabilisation on a mutated nucleosomal surface (Figure 3.28D).

It has previously been reported that SCML2 binds to the TRAF domain of Usp7. To test if this interaction is also relevant in the context of H3 deubiquitylation, chromatin affinity purification experiments were performed with FL SCML2 and Usp7ΔTRAF. Deletion of the TRAF domain resulted in the loss of SCML2-mediated recruitment of Usp7 to any of the six chromatin arrays. As a result, the enzyme was unable to catalyse the deubiquitylation reactions (Figure 3.28E).

Together, the chromatin affinity pull-down experiments performed with recombinant proteins showed that SCML2 stimulated Usp7 deubiquitylation by interacting with the enzyme and targeting it to chromatin.

3.3.4 SCML2 stimulates Usp7 on H3 ubiquitylated chromatin

To better understand the mechanism of SCML2-induced activation of Usp7 function, the chromatin pull-down experiments presented above were complememnted with a series of deubiquitylation experimetns performed on various substrates ranging from free histones to assembled chromatin arrays.

To test the degree of SCML2-induced Usp7 stimulation, several substrates, including H3K18ub

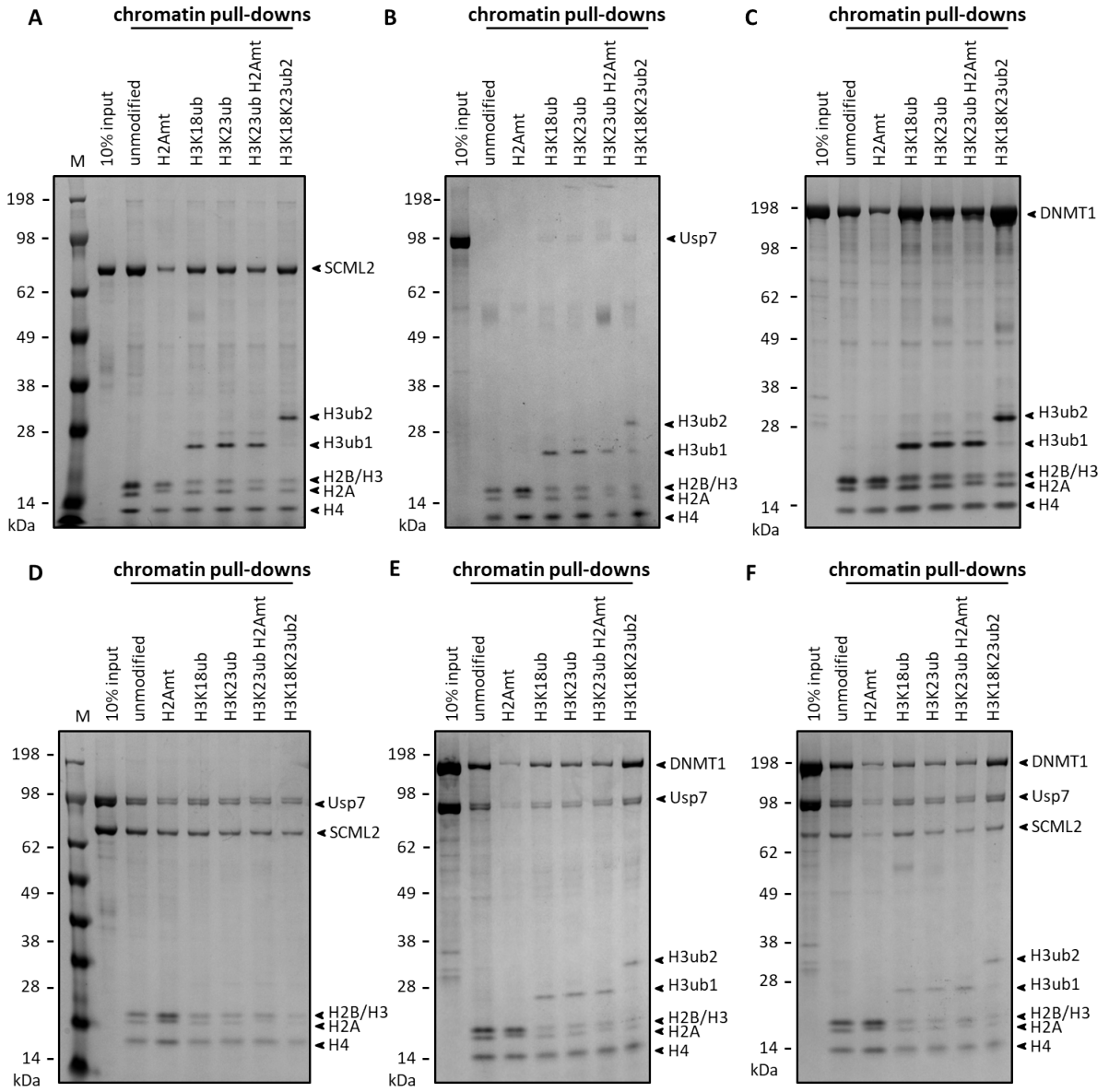


Figure 3.27: SCML2 and DNMT1 recruit Usp7 to H3 ubiquitylated chromatin templates. Coomassie-stained SDS-PAGE gels of chromatin affinity purifications documenting the interaction of SCML2, Usp7 and DNMT1 with ubiquitylated chromatin arrays. Biotinylated nucleosomal arrays containing unmodified, H2Amt, H3K18ub, H3K23ub, H3K23ub H2Amt and H3K18/23ub2 histones were incubated with recombinant His.SCML2 (A), His.Usp7 (B), His.YFP.DNMT1 (C), His.Usp7 and His.SCML2 (D), His.Usp7 and His.DNMT1 (E) and His.SCML2 and His.Usp7 and His.DNMT1 (F).

histone, mononucleosomes and chromatin arrays were incubated with constant amounts of Usp7 and increasing amounts of SCML2 (Figure 3.29A). Usp7 stimulation was observed in the case of the mononucleosome assembled on the symmetric 187 bp DNA template and in the case of the nucleosomal array. Interestingly, SCML2 could not stimulate Usp7 function when H3K18ub histone was used as a substrate.

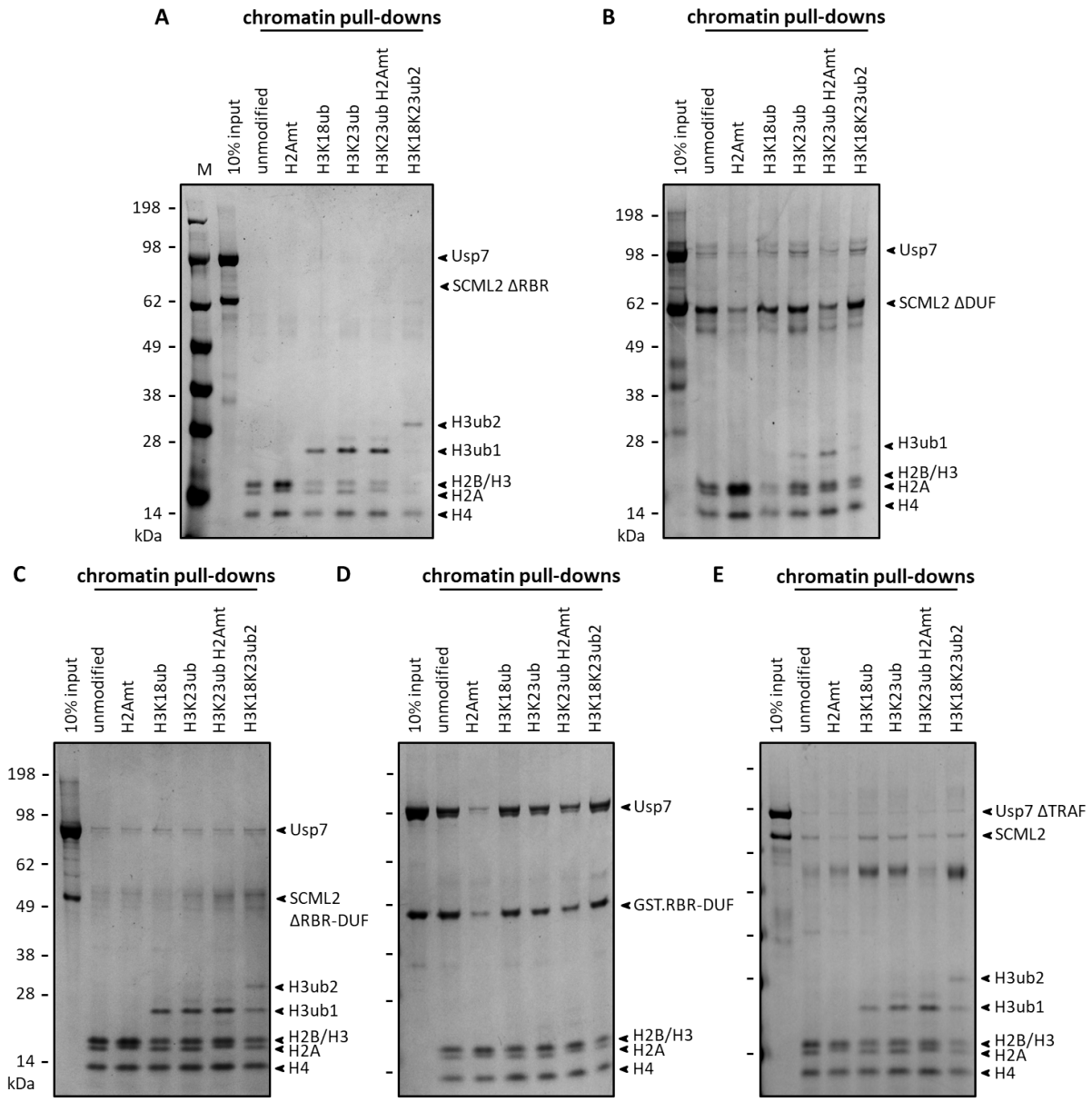


Figure 3.28: SCML2 stimulates Usp7 by targeting it to chromatin. Coomassie-stained SDS-PAGE gels of chromatin affinity purifications documenting SCML'2 stimulation of Usp7 deubiquitylation activity. Biotinylated nucleosomal arrays containing unmodified, H2Amt, H3K18ub, H3K23ub, H3K23ub H2Amt and H3K18/23ub2 histones and FL His.Usp7 were incubated with recombinant His.SCML2 ΔRBR (A), His.SCML2 ΔDUF (B), His.SCML2 ΔRBR-DUF (C), FL His.Usp7 with GST.RBR-DUF (D) and His.Usp7 ΔTRAF with FL His.SCML2 (E).

SCML2 may activate Usp7 by inducing a conformational change in the enzyme upon binding to its TRAF domain. Since direct stimulation on the histone substrate was not achieved by SCML2, linker DNA, as a constituent of nucleosomes but not of free histones was added to the deubiquitylation reaction. Free DNA (unmodified and fully methylated at CpG dinucleotide sites) was used to induce a conformational change in SCML2, which in turn might

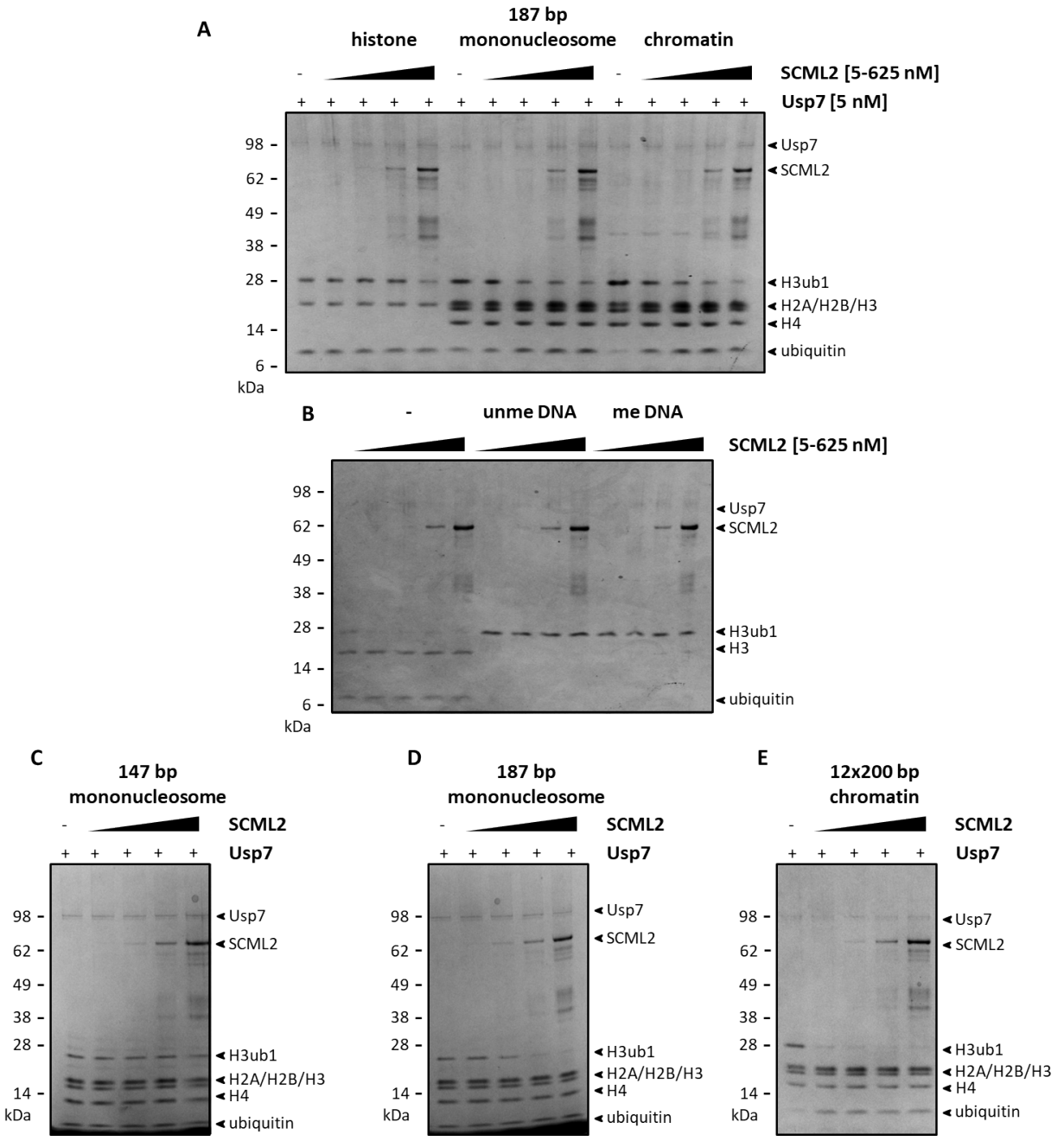


Figure 3.29: SCML2 requires a nucleosome containing linker DNA for stimulation of Usp7. Coomassie-stained SDS-PAGE gels documenting deubiquitylation reactions. (A) Constant Usp7 and increasing amounts of FL His.SCML2 were used to deubiquitylate H3K18ub histone, 187 bp mononucleosome or chromatin substrates. (B) H3K18ub histone deubiquitylation reaction in the presence of constant Usp7 with stoichiometric unmethylated or methylated DNA and increasing amounts of FL His.SCML2. Deubiquitylation experiments performed with constant FL Usp7 and increasing amounts of SCML2 on H3K18ub 147 bp mononucleosome (C), 187 bp mononucleosome (D) and chromatin (E) substrates.

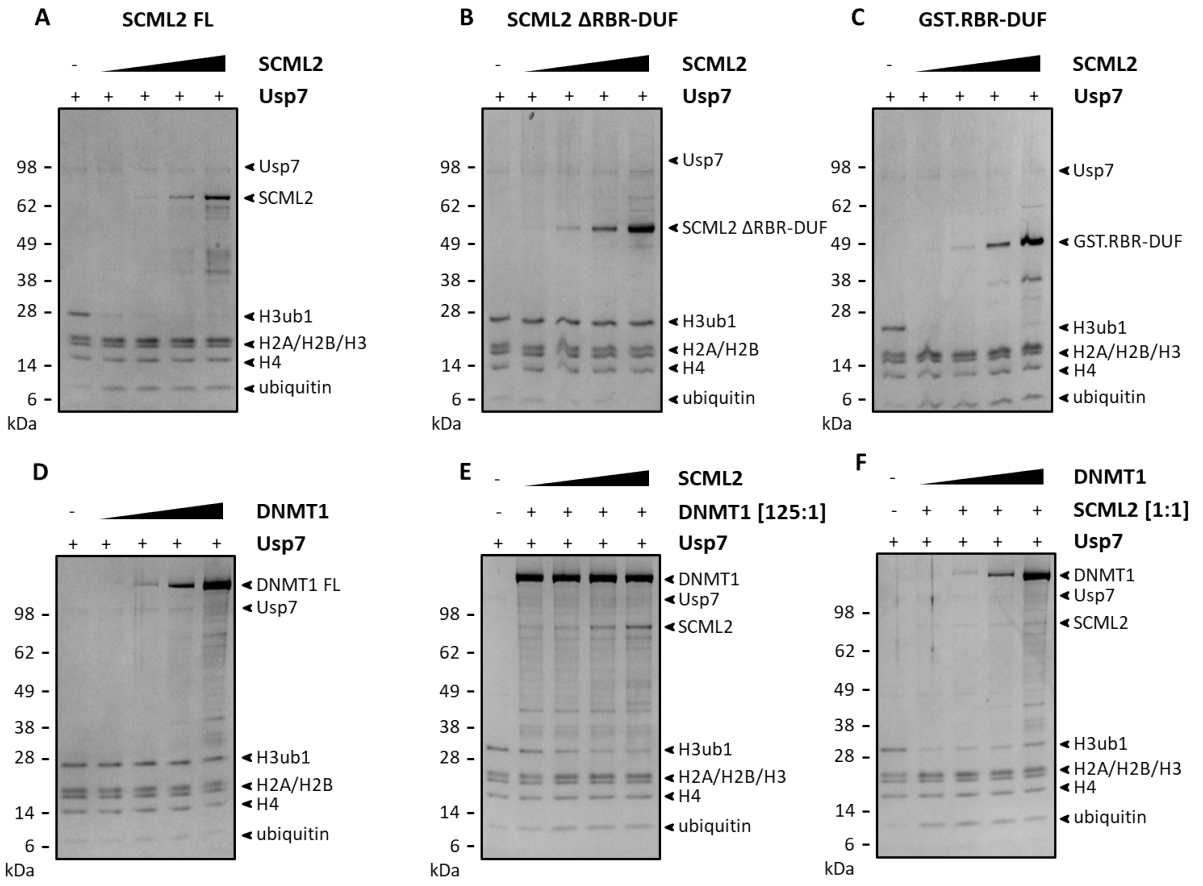


Figure 3.30: DNMT1 inhibits SCML2's stimulation of Usp7 deubiquitylation function. Coomassie-stained SDS-PAGE gels documenting chromatin deubiquitylation reactions. Constant Usp7 and increasing amounts of FL His.SCML2 (A), His.SCML2 RBR-DUF (B), GST.RBR-DUF (C) and FL DNMT1 were used to convert H3K18ub chromatin to H3 and free ubiquitin (D). (E) Chromatin deubiquitylation reactions with constant Usp7 and overstoichioemtric DNMT1 in the presence of increasing amounts of FL His.SCML2. (E) Chromatin deubiquitylation reactions with constant Usp7 and constant understoichiometric FL His.SCML2 in the presence of increasing amounts of FL DNMT1.

have activated Usp7 (Figure 3.29A). Surprisingly, addition of DNA to the deubiquitylation reaction did not stimulate Usp7 function, but completely blocked the enzyme's activity. Even at high SCML2 molar excess (which should have buffered the DNA), the enzyme remained largely inactive. Arguably, DNA directly inhibited Usp7 enzyme.

Free DNA was able to inhibit Usp7 function. This set the posibility that unprotected linker DNA within a chromatin array may also inhibit Usp7. To address this, linker-less 147 bp nucleosome, symmetric linker 187 bp nucleosomes and chromatin arrays were subjected to deubiquitylation reactions in the presence of increasing SCML2 concentrations. Even though SCML2 bound all templates, it could only stimulate Usp7 activity on the nucleosome with linker DNA and on the chromatin array (Figure 3.29D, Figure 3.29E). It could be that SCML2 needed to bind the linker DNA to activate Usp7's function. By doing so, SCML2

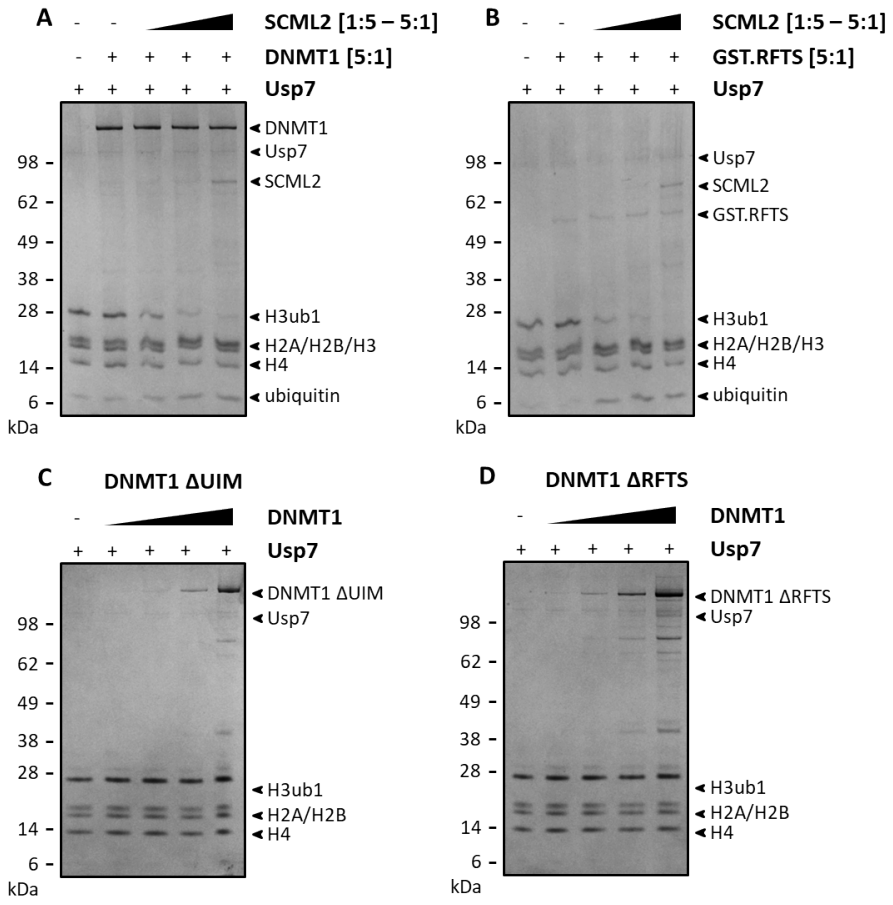


Figure 3.31: DNMT1 does not hinder Usp7 function by binding the ubiquitylation mark Coomassie-stained SDS-PAGE gels documenting chromatin deubiquitylation reactions. (A) Constant Usp7 and DNMT1 (A) or GST.RFTS (B) were added to the deubiquitylation reaction in the presence of increasing amounts of FL His.SCML2.. Chromatin deubiquitylation reactions with constant Usp7 and increasing amounts of DNMT1ΔUIM (C) and DNMT1ΔRFTS (D). UIM = ubiquitin interaction motif; RFTS = replication foci targeting sequence.

might have performed two tasks: it underwent an activating conformational change and buffered the free DNA from inhibiting the Usp7 enzyme.

Deletion of the RBR-DUF region was shown to be essential for Usp7 binding and its recruitment to ubiquitylated chromatin. SCML2 could not stimulate Usp7 activity when missing the RBR-DUF region (Figure 3.30B). Incubation of Usp7 with recombinant RBR-DUF rescued this function, showing that the two domains contained all the necessary information for stimulation of the deubiquitylating enzyme in chromatin context (Figure 3.30C).

DNMT1 binds Usp7 and recruits it to chromatin, but it does not stimulate its deubiquitylation function. When incubated with Usp7 and SCML2, DNMT1 inhibited the stimulatory effect of SCML2 (Figure 3.27). On experiments performed on unmodified chromatin, DNMT1 and SCML2 formed a stable trimeric complex with Usp7 (Figure 3.26). The trimeric complex

could be static in that all proteins are locked in immobile conformations, but also dynamic where Usp7 might be able to transition from a DNMT1-bound to a SCML2-bound conformation. In the later case, an increased local concentration of either DNMT1 or SCML2 would shift the Usp7 conformational equilibrium. To test that hypothesis, Usp7 was incubated with overstoichiometric DNMT1 concentrations on H3K18ub chromatin arrays and SCML2 was titrated over a wide range of concentrations. While at low SCML2 concentrations, DNMT1 prevented SCML2's stimulatory effect of Usp7 function (Figure 3.30A), at higher SCML2 concentrations, there was an increased deubiquitylation activity. In the reverse experiment, SCML2 was kept at understoichiometric amounts with respect to Usp7 and DNMT1 was added in increasing amounts to the deubiquitylation reaction. Less deubiquitylation was observed as increasing amounts of DNMT1 were being titrated in the reaction. Likely, DNMT1 inhibits the SCML2 stimulation of Usp7's deubiquiylation activity.

Whilst the observed inhibition of SCML2 by DNMT1 may be explained by Usp7's promiscuous behaviour, it could also be that DNMT1, which directly binds H3K18ub [37], protects the mark from deubiquitylation. To test this hypothesis, FL DNMT1 or the replication foci targeting sequence (RFTS) domain, responsible for H3ub binding, were titrated into deubiquitylation reactions in the presence of SCML2 (Figure 3.31A, Figure 3.31B). While there was some inhibition of SCML2 stimulation in the reactions where the recombinant RFTS domain was added, the effect was not as pronounced as in the reaction where FL DNMT1 was added. To test this hypothesis further, the RFTS domain and the ubiquitin interaction motif (UIM) within RFTS were removed from DNMT1. The deletion constructs were used in deubiquitylation reactions (Figure 3.30C, Figure 3.30D). Neither the RFTS nor the UIM deletions could stimulate Usp7 activity on H3K18ub nucleosomal arrays, suggesting that blocking or protection of the H3ub mark was not the mechanism by which DNMT1 failed to activate Usp7 or inhibit the SCML2 stimulatory effect.

3.3.5 SCML2 positions Usp7 at the N-terminus of histone H3

To better understand how SCML2 activates Usp7 on chromatin, crosslinking mass spectrometry analysis was performed (Figure 3.32). Recombinant FL His.Usp7 and recombinant GST.RBR-DUF were incubated with unmodified chromatin on streptavidin beads. Excess His.Usp7 and RBR-DUF were removed during the chromatin pull-down. The GST.RBR-DUF construct was prefered over FL SCML2 to avoid SPM-mediated multimerisation of SCML2 during protein-protein crosslinking (Figure 3.26B) [106]. GST.RBR-DUF was shown to be necessary and sufficient for SCML2's interaction with Usp7, for SCML2 and Usp7 chromatin targeting and for SCML2 stimulation of Usp7 deubiquitylation activity (Figure 3.21E, Figure 3.30C, Figure 3.28D). Stable as well as transient interactions between GST.RBR-DUF, FL His.Usp7 and the histones making up the unmodified chromatin were covalently linked at the end of the affinity purification experiment. The crosslinker used was BS3, a homobyfunctional crosslinker with a spacer of 11.4 Å that links lysine residues found in close proximity to each other.

Besides the known interaction with the TRAF domain of Usp7 (Figure 3.28F) [98], RBR-DUF formed several crosslinks with the ubiquitin like domains UBL1, UBL2, UBL3, UBL4

and UBL5 (Figure 3.32A). Most crosslinks originated from the RBR region and contacted the UBL2-3 domains. The RBR region also contacted the nucleosome at the N-termini of histones H2B, H3 and H4. The Usp7 catalytic domain contacted the nucleosome on the N-terminal tail of histone H3 (residues K15 and K37) and the C-terminal tail of histone H4. Additional contacts were detected between UBL4-5 domains and the N-terminus of histone H2B. The extreme C-terminus of Usp7 formed intramolecular crosslinks with Usp7's catalytic domain and the UBL1 domain, in what was previously shown to constitute an activating conformational change [70]. GST.RBR-DUF recruited Usp7 to the nucleosome and positioned it in the right orientation to mount a subsequent deubiquitylation attack onto the N-terminal tail of histone H3 (Figure 3.32B).

Previously, it was shown that ubiquitin specific protease adaptors bind the nucleosome at its acidic patch [72]. In the crosslinking mass spectrometry experiment, GST.RBR-DUF bound the nucleosome around the DNA gyres proximal to the N-terminus of histone H3. Only two crosslinking sites (H2BK112; H3K56) were found proximal to the nucleosomal acidic patch. As suggested before, SCML2 may bind the nucleosome on two distinct surfaces. (Figure 3.17, Figure 3.24A, Figure 3.27A, Figure 3.28D).

The known interation between SCML2 and the TRAF domain of Usp7 was only indicated by two crosslinks (Figure 3.32A), one to the RBR domain, one to the DUF domain. Arguably, this was a consequence of the tight interaction between RBR-DUF and TRAF, which created a solvent-inaccesible interface. GST.RBR-DUF's crosslinking sites on Usp7 mapped primarily to the UBL1-5 domains (Figure 3.32C). Surprisingly, the crosslinking sites coincided with known interaction surfaces between Usp7 and the polybasic region (PBR) of UHRF1 and between Usp7 and the bromo-adjacent homology (BAH) domains of DNMT1 (Figure 3.32D, Figure 3.32E) [91], [92].

The crosslinking mass spectrometry experiments suggested the formation of a cross-talk or competition between DNMT1, SCML2 (and UHRF1) for interaction with the UBL1-2-3 domains of Usp7. Previously, we showed that DNMT1 inhibits SCML2 stimulation of Usp7 (Figure 3.30, Figure 3.31). During the deubiquitylation experiment, Usp7 likely switched between the DNMT1-bound state and the SCML2-bound state which caused conformational changes with functional implications in the enzyme. SCML2 may be needed on the H3ub chromatin to lower the activation energy for the Usp7-catalysed deubiquitylation reaction, to foster the recycling of the H3ub-bound DNMT1. To test this hypothesis, di-ubiquitylated chromatin arrays were preincubated to saturation with recombinant DNMT1 and then subjected to deubiquitylation in the presence or absence of GST.RBR-DUF over a time course of four hours (Figure 3.33). After chromatin afinity purification, Usp7 partially converted the H3ub2 mark to H3ub1 and decreased marginally the DNMT1 concentration on the chromatin array. In the presence of excess GST.RBR-DUF, more Usp7 was recruited to chromatin which resulted in more conversion of H3ub2 to H3ub1 and free H3 and less recovery of DNMT1 on the biotinylated arrays. This finding is in agreement with the hypothesis that SCML2 is needed for recycling of DNMT1 by facilitating the removal of the H3ub mark from chromatin arrays.

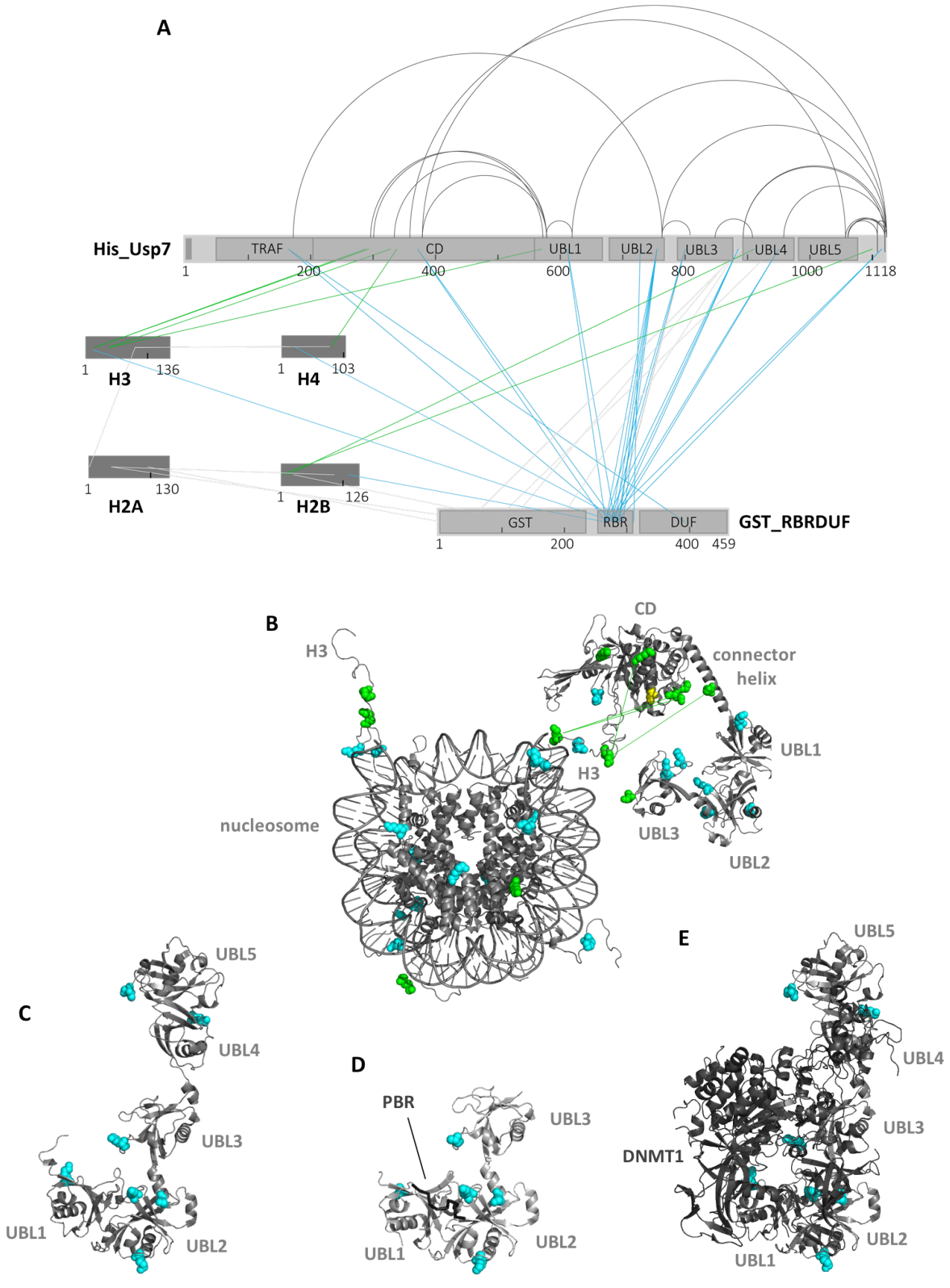

A
His_Usp7
TRAF
CD
UBL1
UBL2
UBL3
UBL4
UBL5
1
200
400
600
800
1000
1118
H3
1
136
H4
1
103
H2A
1
130
H2B
1
126
GST
RBR
DUF
GST_RBRDUF
1
200
400
459
B
H3
CD
connector
helix
UBL1
H3
nucleosome
UBL3
UBL2
C
UBL5
UBL4
UBL3
UBL1
UBL2
D
PBR
UBL3
UBL1
UBL2
E
UBL5
UBL4
UBL3
DNMT1
UBL2
UBL1

Figure 3.32 *(preceding page)*: SCML2 adapts Usp7 to H3ub chromatin. (A) Crosslinking mass spectrometry analysis of Usp7 and GST.RBR-DUF on unmodified chromatin arrays. Intermolecular crosslinks are depicted with straight lines. Intramolecular crosslinks are depicted as curves. SCML2 crosslinks to the nucleosome or to Usp7 are colored in cyan. Usp7 crosslinks to the nucleosome are colored in green. (B) Representation of the observed crosslinks on the nucleosome and Usp7 crystal structures. SCML2 crosslinked sites on Usp7 and on the nucleosome are highlighted as cyan spheres. The SCML2 crosslinks on the nuclesome are displayed for each of the two copies of H2B, H3 and H4. Usp7 crosslinks on the nucleosome and nucleosome crosslinks on Usp7 are highlighted as green spheres. The interaction between the catalytic domain and the H3 tail is represented by straight green lines. Usp7 catalytic centre is depicted in yellow. (C) Representation of SCML2 crosslinks on the UBL1-2-3-4-5 crystal structure. (D) SCML2 crosslinks on the UBL1-2-3 domains coincide with the interaction surface occupied by the UHRF1's PBR peptide. (E) SCML2 crosslinks on the UBL1-2-3-4-5 coincide with both of DNMT1's binding surfaces onto Usp7. Structural information was obtained from crystal structures with PDB accession numbers 1KX5 (mononucleosome), 5FWI (Usp7 CD and UBL1-2-3), 5C6D (Usp7 UBL1-2-3 with UHRF1 PBR) and 4YOC (Usp7 UBL1-2-3-4-5 with DNMT1) [191], [192], [91], [92].

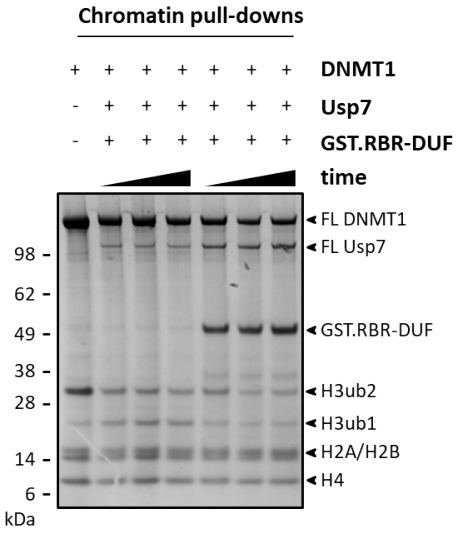


Figure 3.33: SCML2 controls DNMT1 residence time on H3 ubiquitylated chromatin. Coomassie-stained SDS-PAGE gel of a time-course chromatin pull-down experiment with preincubated FL DNMT1 in the presence of equimolar FL Usp7 or equimolar FL Usp7 and overstoichiometric GST.RBR-DUF. Samples were collected 1, 2 and 4 h after addition of Usp7 or Usp7 and GST.RBR-DUF.

Chapter 4

Discussion

4.1 Histone ubiquitylation confers chromatin unique properties

We hypothesised that ubiquitylation of histones creates distinct chromatin environments which differ from unmodified environments in the nuclear proteins and protein complexes which are attracted to them. To address this hypothesis, site-specifically ubiquitylated histones were incorporated in chromatin arrays which were used in affinity purification experiments to map the nuclear interactomes associated with them. The interactomes of all modified chromatin templates differered from the control chromatin and the interactomes of all modified chromatin templates differed between themselves. This suggests that ubiquitylation of histones confers different properties to the underlying conjugation sites which specify these for unique downstream processes. We propose that the readout of the different histone ubiquitylation marks coroborates with parallel chemical and physical changes on chromatin and on the associated interactors to create unique signalling hierarchies which translate into specific biological functions.

4.1.1 ChAP-MS highlights the requirements for histone ubiquitylation readout

Recognition of ubiquitin conjugation to target proteins occurs through dedicated reader proteins, which contain specialised ubiquitin binding domains (UBDs). Based on their architectural features, UBDs are divided into two main groups [64]. The first group of domains consists of alpha-helices and includes the elongated single helix of the ubiquitin interaction motif (UIM), the three-helix bundle of ubiquitin associated domain (UBA), coupling of ubiquitin to ER degradation (CUE) and GGA and TOM1 (GAT) as well as the eight-helix fold of Vps, HRF and STAM (VHS). The second group of ubiquitin binding domains consists of folded beta-strands and includes the four-stranded antiparallel beta-sheet of Npl4 Zn Finger (NZF) and the beta-propeller of tryptophan aspartate 40 (WD40). We suggest that, for correct recognition of ubiquitylated histones, the reader must contain a dedicated ubiquitin binding domain which contacts ubiquitin, but also several distinguishing features of the nucleosome core particle.

The chromatin affinity purification experiments presented in this thesis found a large number of proteins which associated with the different histone ubiquitylation marks (Table 3.3). To date, only a limited number of these proteins have been structurally described to the extent that their respective ubiquitin binding domains are well annotated. This is the case of Rad18, Rnf169 and DNMT1, where recognition is achieved by specialised UIMs [68], [37]. An intersection of all UBDs known to date with a database containing all annotated protein domains within the identified interactors yielded a few additional proteins which may have the ability to bind ubiquitin. This is the case of PLAA, DCAF7 and TBL1XR1, which contain WD40 domains, and the case of Usp3, ZMYNDB, Znf592, ZnF609, ZnF687, which contain Zn fingers.

We have shown that ubiquitin is interpreted differently by the nuclear proteome when it is incorporated on a histone protein than when it is part of a mononucleosome or chromatin array. Neither the RNA polymerase nor the integrator complex could be recruited on the H2BK120ub histone, but they were both enriched on ubiquitylated mononucleosomes or ubiquitylated chromatin arrays (Figure 3.13). None of the proteins enriched on the H2BK34ub chromatin array were previously found on the H2BK34ub histone template (Figure 3.14). Neither Usp7 nor SCML2 was recruited to any of the H3 ubiquitylated histones, but they were both enriched onto all of the H3 ubiquitylated chromatin arrays (Figure 3.17B, Figure 3.17C, Figure 3.14). The additional features present on the more complex templates allows for the formation of more specific and more stable interaction surfaces where both dedicated UIMs and adaptor proteins are likely to play a role.

We found that each histone ubiquitylation mark generally recruited a distinct set of proteins (Table 3.3). When confronting the literature, there were some notable exceptions to this observation. We showed that DNMT1 was enriched on three H3 ubiquitylation templates (Figure 3.17C, Figure 3.15). In the literature, DNMT1 also appeared enriched on H2AK119ub mononucleosomes [74]. We showed that Rad18 and RNF169 were enriched on the H2BK120ub mononucleosome and chromatin templates (Figure 3.13C, Figure 3.13E). In the literature, both Rad18 and RNF169 have been associated with the DNA-damage response H2AK13/15ub mark [68]. We reported that Usp7 was found enriched on the H3ub chromatin arrays. In the literature Usp7 was previously shown to control the ubiquitylation levels of both H2BK120ub and H2AK119ub [96], [98], [193]. We reported that Usp7 can also deubiquitylate H3K18/23ub chromatin (Figure 3.29).

DNMT1, Rad18 and RNF169 have specialised ubiquitin interaction motifs (Figure 1.3) [68], [37]. This allows them to directly recognise ubiquitin embedded in a nucleosomal context. Binding of DNMT1 to both H2AK119ub and H3K14/18/23ub may be a consequence of the topological proximity of the flexible H3 and H2A N- and C-terminal tails at the nucleosomal dyad axis. Similarly, both Rad18 and RNF 169 exchanged H2AK13/15ub for H2BK120ub arguably because of their spatial proximity at the DNA gyres with superhelical location SHL +4 and SHL -5. Finally, for Usp7 to discriminate between H2AK119ub, H2BK120ub and H3K14/18/23ub, its targeting depends largely on ubiquitin readers (DNMT1) or adaptor proteins: GMPS for H2BK120ub and SCML2 for H3K18/23ub (Figure 3.25, Figure 3.32) [96], [95].

4.1.2 The histone ubiquitylation interactomes reveal complex signalling events

Only a small number of the proteins that were found in the different mass spectrometry analyses contain ubiquitin binding domains or ubiquitin interaction motifs. Similarly, few interactors are involved directly in the metabolism of the histone ubiquitylation marks. We argue that several intermolecular interactions, biochemical reactions and structural changes occur in parallel to the histone ubiquitylation readout process and that these may account for the enrichment of the remaining interacting proteins.

First, most proteins identified in the mass spectrometry analysis are part of annotated protein complexes. As such, readout of the ubiquitylated histone is expected to occur through one dedicated subunit which recruits the remaining complex to the marked site. By extension, some proteins and protein complexes are only recruited to the modified chromatin as part of secondary binding events. The interaction between the RNA polymerase II and the NELF and DSIF complexes, which are involved in transcription elongation, as well as the association of the polymerase with ERCC1/ERCC4/SLX4, which are part of the nucleotide excision repair machinery, are well documented in the literature (Figure 3.13, Figure 3.16C) [194], [195]. Recruitment of the transcription elongation factors or of the DNA repair machinery may not have been observed in the absence of the RNA Polymerase II.

Second, since the nuclear extract contains not only proteins and nucleic acid molecules, but also a limited yet complete set of small soluble cofactors, several biochemical reactions may occur on the ubiquitylated chromatin templates during incubation with the nuclear extract. These reactions may affect the chemistry of the histone proteins, of the underlying DNA template or of the proteins that are recruited to the modified chromatin arrays. DNMT1 has previously been shown to have increased methylation activity in the presence of ubiquitylated H3 [37]. It is possible that the ubiquitylated chromatin arrays also undergo DNA methylation during incubation with the extract, which would explain the enrichemnt of a heterochromatin formation complex on the H3K18ub template. Acetylation of DNMT1 has been shown before to control its stability [92], [93]. Deletion or mutation of DNMT1's lysine residues in its KG linker resulted in the loss of the interaction between Usp7 and DNMT1 [92], [93]. This suggested that acetylation of DNMT1's KG linker hinders its interaction with Usp7. The Sin3a-SAP130-SUDS3 co-repressor deacetylation complex was enriched in the H3K18ub affinity purification experiment (Figure 3.15, Figure 3.16F) and opens the possibility that DNMT1 may be deacetylated in the course of the nuclear extract incubation reaction. This may affect DNMT1's surface chemistry in such a way that DNMT1-mediated Usp7 recruitment to the H3 ubiquitylated chromatin is favored. These observations indicate that the H3ub chromatin and the associated interactors do not form a static environment, but a dynamic stage where several reactions take place.

Third, the presence of twenty-four ubiquitin molecules per chromatin array may have affected the array's folding properties. Ubiquitylated chromatin templates may be more accessible than unmodified fibers to proteins which bind the linker DNA or the surface of the nucleosome core. Histone ubiquitylation in general and the H2BK120ub modification in particular

have been shown *in vitro* to induce changes in the way chromatin folds [22], [196]. These differences were not attributed to the position of ubiquitin, but to the presence of an acidic patch on the surface of ubiquitin [23]. This implies that all histone ubiquitylation marks may have the ability to induce conformational changes to the marked chromatin fibers. We showed that Zn finger containing proteins as well as Swi3, Ada2, N-Cor, and TFIIIB (SANT) domain containing proteins, which probe the nucleic acid and the surface of the histone octamers, were recovered in the histone ubiquitylation datasets. Their recovery may have resulted from the larger accessibility to the chromatin array induced by histone ubiquitylation.

Fourth, it is likely that a number of interactors contain yet to be annotated ubiquitin interaction motifs. Rad18, RNF168 and RNF169 are E3 ligases which associate with DNA repair and were found to target H2A's N-terminal K13 and K15 residues [59], [197]. Recent structural investigation of the ubiquitylation mechanism pointed out the presence of UIMs in Rad18 and RNF169 [68]. Along the same lines, following the observation that JARID2 acts as a reader which recruits the PRC2 complex to H2AK119 ubiquitylated chromatin, the presence of a UIM at the JARID2's N-terminus was mapped [74], [75], [145]. Similarly, after observing that DNMT1 recognises H3K18 and H3K23 ubiquitylation, the UIM within its RFTS was also identified [35], [36]. Further biochemical experiments will be needed to define which subunits of the multiprotein complexes described in this thesis are responsible for direct recognition of the various ubiquitin marks.

4.2 SCML2 fine-tunes H3 deubiquitylation during maintenance DNA methylation

In addition to the readout process, each ubiquitylation reaction requires dedicated protein complexes for proper setup and removal of the modification. As such, with the exception of the ubiquitylation sites found in close proximity to each other (as it is the case with H2AK13 and H2AK15, or with H2AK118 and H2AK119 or with H3K14 and H3K18 and H3K23), each modification is deposited by a dedicated set of E2 and E3 conjugating and ligating enzymes (Table 1.1). Along the same lines, different enzymes catalyse the removal of distinct ubiquitylation marks and increased specificity for these proteases is achieved with the help of adaptor proteins. For deubiquitylation of H2BK120ub, Ubp8 requires the adaptor Sgf11 [154], [198], [72] and Usp7 requires the adaptor GMPS [96]. For deubiquitylation of H2A residues K13 and K15, Usp16 requires the adaptor Herc2 [142]. Similarly, for deubiquitylation of H2AK119ub, BRCA-1 associated protein 1 (BAP1/Calypso) protease requires the adaptor protein additional sex combs (ASX) [148]. Our results indicate that SCML2 has a similar regulatory role in the context of H3 ubiquitylation. We propose that SCML2 regulates the timing of events on the H3 ubiquitylated chromatin by inducing an activating conformational change in Usp7, which removes the H3 ubiquitylation mark to promote recycling of DNMT1 from fully-methylated CpG sites.

4.2.1 SCML2 activates Usp7 for deubiquitylation of histone H3

Previously, SCML2 has been shown to mediate the interaction between Usp7 and PCGF4. This interaction was shown to be important *in vivo* for Usp7 driven deubiquitylation of the E3 ubiquitin ligase RING1B [98]. In the same article, recombinant SCML2 was incubated with Usp7, but no direct stimulation of Usp7's deubiquitylation activity could be recorded using a minimal deubiquitylation substrate [98]. The authors argued in that study that no stimulation was observed, because the substrate used in their experiments was not "physiologically relevant". We prepared ubiquitylated H3 histone with a native isopeptide linkage and provided evidence that the interaction between SCML2 and Usp7 is important for Usp7's deubiquitylation activity. We propose that the observed activation is mediated by an allosteric effect caused by SCML2 on Usp7 directly on the H3 ubiquitylated chromatin.

We have shown that Usp7 can by itself deubiquitylate H3 histones, mononucleosomes and chromatin arrays (Figure 3.29). This was achieved because the protease recognises ubiquitin through a dedicated ubiquitin binding domain, which is part of its larger catalytic domain [69]. We reported that Usp7 deubiquitylated free histones more readily than chromatin arrays (Figure 3.29A). The difference in activation energy was lowered on chromatin and mononucleosomes containing linker DNA by Usp7's interaction with SCML2 (Figure 3.29D; Figure 3.29E). As suggested before, SCML2 stimulated Usp7 function in the presence of a physiologically relevant substrate [98].

We showed that SCML2 could not stimulate Usp7 when ubiuitylated linker-less nucleosomes or free histones were used as substrates (Figure 3.29A, Figure 3.29C). The observation that SCML2 cannot stimulate Usp7 on a linker-less nucleosome may be explained by the fact that only one of the two SCML2 nucleosome binding surfaces within the RBR domain was available to bind the nucleosome in the absence of linker DNA (Figure 3.23, Figure 3.24). SCML2's inability to stimulate Usp7 on a free histone may be explained by the fact that SCML2 does not bind either unmodified or ubiquitylated H3 histone (Figure 3.17B) to increase Usp7's local concentration onto these substrates. We argue that SCML2 binding to Usp7 is insufficient for activation of the protease. Likely, allostery alone is insufficient for Usp7 activation by SCML2.

We found that SCML2 recruited Usp7 to chromatin arrays and stimulated its deubiquitylation activity on both H3K18ub, H3K23ub and H3K18/23ub2 (Figure 3.27D). Deletion of the RBR-DUF domain, which was essential to recover a complete interaction between SCML2 and Usp7 resulted in the loss of chromatin recruitment and the inability to deubiquitylate any of the chromatin templates (Figure 3.28C). Deletion of the RBR domain, which we found to be essential for SCML2 targeting to chromatin, showed the same effects (Figure 3.28A). These experiments suggest that SCML2 binding to chromatin is essential for stimulation of Usp7's deubiquitylation activity.

We found that a SCML2 mutant which lacks the DUF domain, whose interaction with Usp7 is decreased but is still able to bind chromatin, exhibited an Usp7 stimulatory effect comparable to that of the FL protein, even though it recovered substoichiometric amounts of the

protease to the chromatin templates (Figure 3.28B). This observation suggests that SCML2 activates Usp7 through a mechanism more complex than simple recruitment to the substrate.

We presented that DNMT1 can also target Usp7 to chromatin, but that this association does not result in stimulation of deubiquitylation activity (Figure 3.30D, Figure 3.27E). Removal of the RFTS domain of DNMT1 or of the UIM within the RFTS domain, which was previously shown to bind ubiquitylated H3 with nanomolar affinity [37], could not stimulate Usp7 activity (Figure 3.31C, Figure 3.31D). We argue that Usp7 targeting to chromatin is insufficient for stimulation of its deubiquitylation activity.

We found that the N-terminal tail of histone H3 crosslinks to residues proximal to the active centre of Usp7 when the protease is incubated with unmodified chromatin in the presence of SCML2 (Figure 3.32). This suggests that SCML2 positions Usp7's catalytic domain close to the isopeptide linkages of ubiquitin to the H3 histone tail. We propose that Usp7's activity on the H3 ubiquitylated chromatin templates is stimulated by SCML2 through a mechanism which combines chromatin targeting with allosteric activation.

4.2.2 SCML2 mediates the cross-talk between DNMT1, UHRF1 and Usp7

Shortly after DNA replication, UHRF1 recognises hemi-methylated CpG marks [33], [34]. Upon this recognition, UHRF1 ubiquitylates the N-terminal tail of histone H3 which recruits DNMT1 that pastes the methylation pattern from the parental DNA strand onto the daughter strand [35]. The timing and the mechanism of Usp7 and SCML2 recruitment to the H3 ubiquitylated chromain is not yet clear. We propose that DNMT1 recruits Usp7 and SCML2 (through its interaction with Usp7) to the modified chromatin.

Using western blot analysis of chromatin affinity purification experiments from nuclear extracts, we have shown that, differently than the recombinant proteins, neither SCML2 nor DNMT1 could bind unmodified chromatin (Figure 3.17C). In addition to that, the three H3 ubiquitylated chromatin arrays showed different recruitment patterns for DNMT1, Usp7, UHRF1 and SCML2. On one hand, DNMT1, Usp7 and SCML2 reproducibly preferred the H3K18ub and H3K18/23ub2 over the H3K23ub chromatin arrays (Figure 3.17C, Figure 3.18, Figure 3.19). On the other hand, UHRF1 bound better to the H3K23ub and H3K18/23ub2 arrays (Figure 3.17C, Figure 3.18, Figure 3.19).

In agreement with the previously described binding constants calculated for the RFTS domain in the presence of ubiquitylated H3 histone peptides, DNMT1 prefered the H3K18ub and H3K18/K23ub2 over the H3K23ub modification (Figure 3.17C) [37]. In that study, the affinity of the RFTS domain for H3K18ub was twenty fold higher than that for H3K23ub. The affinity of the RFTS domain for H3K18/23ub2 was forty fold higher than that for H3K23ub. We observed similar differences in enrichment in the affinity purification experiments performed with nuclear extracts (Figure 3.17C).

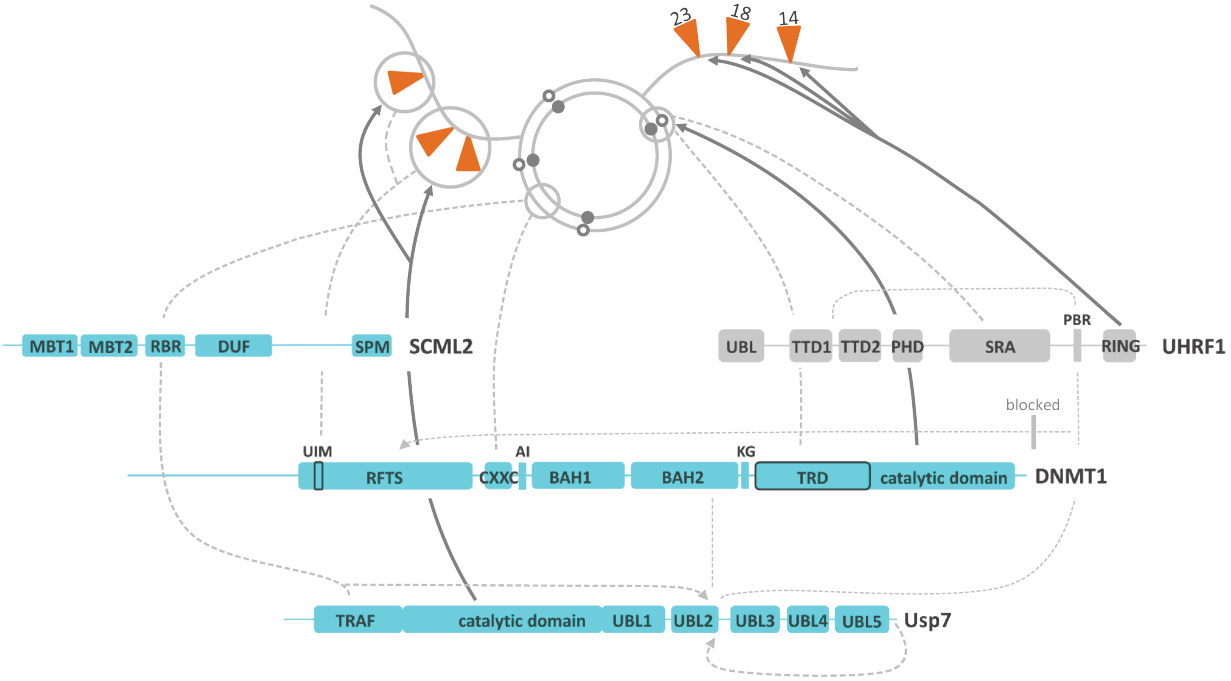


Figure 4.1: Cross-talk between the H3ub-specific interactors. UHRF1 recognition of hemimethylated CpG sites (open and full circles) results in deposition of ubiquitin (orange triangle) on the N-terminal tail of histone H3 at lysine residues K14, K18 and K23. DNMT1 recognition of H3 ubiquitin (large open circle) stimulates methylation of the hemimethylated CpG sites. Recruitment of Usp7 to chromatin relies on its interaction with SCML2, DNMT1 or UHRF1. SCML2 stimulates Usp7 to deubiquitylate the N-terminal tail of H3 histone. Catalytic activities are represented with thick arrows. Recognition events on the nucleosome are depicted with dashed lines that end in open circles. Inter and intra molecular interactions between the H3ub interactors are depicted with dashed lines. Stimulatory interactions end with arrowheads.

The high recruitment of DNMT1 in the case of the H3K18/23ub2 arrays was correlated with the high enrichment of Usp7 and SCML2. Similarly, lower recruitment of DNMT1 in the case of the H3K23ub arrays was correlated with lower enrichment of Usp7 and SCML2. It is unlikely that SCML2 recognised the ubiquitylated chromatin separately of DNMT1, since SCML2 did not show enrichment levels of equal intensity on all three ubiquitylated arrays (Figure 3.17C). More than that, SCML2 could not be recovered on ubiquitylated histones from nuclear extracts (Figure 3.17B) and recombinant SCML2 did not show any preference for ubiquitylated chromatin over the unmodified control templates (Figure 3.27A). We propose that SCML2 is recruited to the the modified chromatin through its interaction with Usp7 which binds DNMT1.

Differently than Usp7 and SCML2, UHRF1's enrichment on chromatin arrays did not depend on that of DNMT1. UHRF1's PBR region contacts DNMT1 at the RFTS domain (Figure 4.1) [90]. It is possible that, when bound by the ubiquitylated H3, the RFTS domain of DNMT1 was inaccessible for PBR binding, which resulted in the loss of the interaction between DNMT1 and UHRF1 (Figure 4.1). Since UHRF1 does not contain a known ubiquitin binding motif and was not enriched in the mass spectrometry analysis, it remains unexplained

how it bound specifically onto the H3K23ub-containing arrays. Of note, the chromatin arrays used throughout this thesis are prepared to be devoid of CpG or H3K9 methylation at the beginning of each experiment, which could have been recognised by UHRF1's SRA and TTD domains.

UHRF1's binding to the H3K23ub arrays suggested however two probable scenarios. First, it may be that the E3 ligase is stimulated by this modification to further ubiquitylate other neighbouring lysine residues on the N-terminal tail of histone H3. In this case, the H3K23ub-containing chromatin would, like the better studied examples of H2BK120ub and H2AK119ub, experience ubiquitylation/deubiquitylation cycles [65], [66]. These cycles would insure a stepwise, rather than a linear removal of ubiquitin (Figure 4.2). Second, if UHRF1 contacts Usp7 on the H3K23ub-containing arrays, it may also be possible that in addition to its activity on the H3 histone tail, UHRF1 ubiquitylates Usp7, controlling thus the protease's activity. The opposite, namely that Usp7 controls UHRF1's autoubiquitylation activity, thus the protein's stability, has already been shown [91].

When bound by H3 ubiquitin, the RFTS domain of DNMT1 cannot interact with UHRF1 (Figure 4.1). UHRF1-mediated-targeting of DNMT1 to new CpG sites is prevented. When DNMT1 binds ubiquitylated nucleosomes, its KG linker interacts with the UBL1-5 domains of Usp7 and recruit the protease to the marked nucleosomes. This interaction limits the availability of Usp7 for UHRF1 regulation. Presently, additional information is needed to resolve the problem of DNMT1 recycling which would liberate Usp7 from the marked nucleosomes to control UHRF1 activity in order to drive the maintenance methylation process forward. Acetylation of the KG linker of DNMT1 may regulate Usp7 removal from the marked nucleosomes [92]. DNMT1 acetylation would need to be coordinated with two additional chemicmal modifications: methylation of the marked CpG sites and full deubiquitylation of histone H3. We propose that removal of histone H3 is the trigger that initiates the recycling process and that SCML2, through its interaction with Usp7 mediates the cross-talk between UHRF1, Usp7 and DNMT1.

The PBR region of UHRF1 contacts the UBL1-2-3 domains of Usp7 [91]. The three UBL domains form also the largest interaction surface between Usp7 and DNMT1 [92]. Crosslinking mass spectrometry analysis of Usp7 and SCML2 in the presence of unmodified chromatin (Figure 3.32), suggests that SCML2 also contacts Usp7 at its UBL1-2-3 domains, separate of the known interaction with the TRAF domain (Figure 3.28E) [98]. Usp7 constitutes the central interaction node of the H3 ubiquitylated chromatin network. SCML2, UHRF1 and DNMT1 compete for UBL1-2-3 binding. The deubiquitylation experiments suggested that SCML2 stimulates Usp7 (Figure 4.1). Crosslinking analysis of Usp7 in complex with SCML2 and chromatin indicates that SCML2 stabilises Usp7's active conformation (Figure 3.32). First, in the presence of SCML2, the C-terminal tail of Usp7 flips onto the back of the catalytic domain of the enzyme to activate it (Figure 3.32A) [70]. Second, the catalytic center of Usp7 is positioned by SCML2 in the right orientation and close to the N-terminal tail of histone H3, which protrudes outside of the nucleosome (Figure 3.32B). We propose that, through its ability to stimulate H3 deubiquitylation, SCML2 resolves the cross-talk between UHRF1, Usp7 and DNMT1 in order to accelerate their recycling from chromatin.

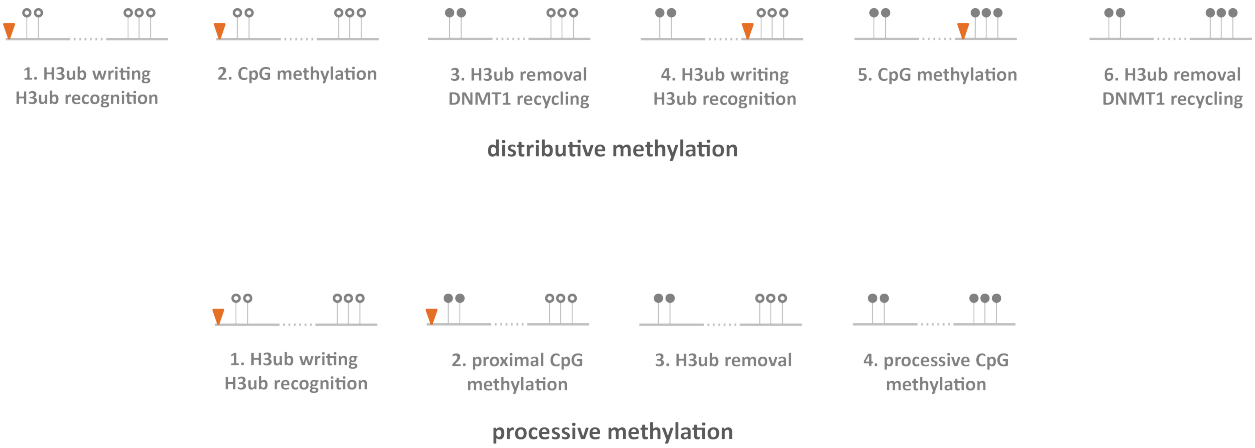


Figure 4.2: Adaptation of DNMT1 to a distributive methylation mechanism. Ubiquitylation of histone H3 (orange triangle) marks hemi-methylated CpG sites (open circles) for DNMT1 recognition. The enzyme methylates the CpG sites (full circles) proximal to the tagged nucleosome and is recycled from chromatin by Usp7-mediated removal of H3 ubiquitin. In a distributive methylation mechanism, a new group of hemimethylated CpG sites from within the same CG-rich region needs to be specified by the addition of a new H3 ubiquitylation mark. In a processive methylation mechanism, after removal of the initial H3 ubiquitylation mark, the methyltransferase moves along the substrate to methylate also the distal CpG sites from within the same CG-rich region without the need of another H3 ubiquitylation mark.

4.2.3 SCML2 controls DNMT1 residence time on H3 ubiquitylated chromatin

The diploid human genome contains approximately 56 million CG/CpG dinucleotides [199]. An estimated 80% of the CpG sites are methylated or undergo oxidation reactions downstream of cytosine methylation [13], [14]. Each HeLa cell contains an average of 25000 DNMT1 protein copy numbers [200]. After DNA replication, the methylation status of each CpG dinucleotide needs to be correctly inherited by the daughter strand. Sequencing analysis of nascent DNA strands suggests that the vast majority (70%) of all CpG sites across the genome are being processed by DNMT1 within the first hour after replication [201], [202]. DNMT1 needs thus to be recycled in an efficient manner to be able to scan in such a short time most of the genome.

On average, each nucleosome from human cells contains two CpG sites. Recruitment of all 25000 DNMT1 molecules by H3 ubiquitylation would cover in the absence of a recycling mechanism only the first 50000 methylation sites. CpG dinucleotides are however not normally distributed across the genome, but often found close to each other in regions that span several hundreds of base pairs. Such areas include methylation free CpG islands in promoter regions and repetitive, highly methylated regions, like transposons (viral DNA elements) and centromeric or satellite DNA sequences [15]. The fact that methylated CpG dinucleotides are not normally distributed throughout the genome, but compacted together in CpG righ regions facilitates DNMT1 recycling.

H3 ubiquitylation recruits DNMT1 to hemi-methylated CpG sites and stimulates DNMT1's methylation activity [35], [37]. Recycling of DNMT1 is thought to occur through Usp7-mediated removal of the H3 ubiquitylation marks [93]. It may seem surprising that Usp7, which removes the ubiquitylation mark, is recruited to the H3ub chromatin primarily in a DNMT1-dependent manner (Figure 3.17C, Figure 3.18, Figure 3.19). We propose that Usp7 (and SCML2) are needed for DNMT1 recycling from H3 ubiquitylated chromatin.

First, the tight binding of the RFTS domain of DNMT1 with the H3 ubiquitylation marks makes it difficult for the recycling process to rely solely on DNMT1's association and dissociation kinetics [37], [93].

Second, the spatial proximity of (SCML2) and Usp7 to DNMT1 and UHRF1 streamlines the genome-wide maintenance DNA methylation process by removing the ubiquitylation mark right after DNA methylation. We reported that irrespective of the levels of Usp7 recruitment, the extent of H3 deubiquitylation was similar on the H3K18ub, H3K23ub and H3K18/23ub2 chromatin arrays (Figure 3.17C, Figure 3.18, Figure 3.19). Along the same lines, the different enrichment levels of DNMT1 or those of SCML2 or UHRF1 did not correlate with the extent of deubiquitylation on the three modified chromatin arrays. We showed that SCML2, Usp7 and DNMT1 form a stoichiometric complex when incubated on unmodified chromatin *in vitro* (Figure 3.26B, Figure 3.26C). Likely, the H3ub interactors do not function independently of each other, but they act in synchrony for efficient maintenance methylation. Recognition of hemi-methylated CpG sites is followed by ubiquitylation of histone H3. This is succeded by methylation of the CpG sites, which is then followed by deubiquitylation of H3 ubiquitin (Figure 4.2, top panel).

Third, DNMT1's recycling by Usp7 could also be seen as an adaptation of the methyltrasnferase to a processive behavior, where the enzyme methylates consecutive CpG sites in CpG-rich sequences without dissociation from the substrate (Figure 4.2, bottom panel). A processive methylation mechanism would be more efficient than a distributive mechanism, where DNMT1 would need to be recycled after each molecular encounter [203], [204]. In the scenario where DNMT1 assumes a processive behavior, the enzyme would make use of its two domains which have the ability to bind DNA (the CXXC domain and the target recognition domain) and move along the substrate according to a "random walk model" [204]. *In vitro*, methylation experiments performed on linear DNA templates show that DNMT1 does not exchange strands or skip target CpG sites during the methylation process [205], [203], [206]. Sequencing of nascent DNA strands suggests the formation of homogeneous methylation landscapes as opossed to accumulation of locally hypermethylated areas [201]. *In vivo*, processive methylation may be facilitated by chromatin remodeler activities which "melt" nucleosomal histone-DNA contacts that may obstruct the movement of the methyltransferase along the substrate [207]. Finally, possible premature removal of the ubiquitylation mark by Usp7, that is before DNMT1 finishes methylating the daughter strand, may be dealt with by UHRF1 which proofreads if hemi-methylated CpG sites are left unprocessed behind.

4.2.4 SCML2 integrates complementing queues from surrounding chromatin

Previously, it was shown that SCML2 integrates a large number of biochemical queues including post-translational modifications of histones, the presence of RNA molecules, accessibility to DNA and the protein's readiness for multimerisation [104], [102], [103], [100], [101]. We propose that SCML2 chromatin targeting relies primarily on the interaction of the RBR domain with the nucleosome. First, we have shown that the RBR domain of SCML2 is essential for binding to free DNA, mononucleosomes and chromatin arrays (Figure 3.23). Targeting of SCML2 to chromatin *in vitro* does not require RNA, as previously suggested [100]. Second, we have observed that removal of the DUF domain does not abrogate SCML2's binding to DNA or to a nucleosome core particle (Figure 3.23). If the DUF domain is required for DNA binding, as previously suggested in the literature [101], then such binding events are secondary to the main RBR-directed interaction. Third, we have shown that *in vitro*, the RBR-DUF region of SCML2 is sufficient to bind Usp7, recruit it to chromatin and stimulate its deubiquitylation function.

While the work presented in this thesis suggests that the RBR-DUF region suffices to fine-tune Usp7's activity and streamline the maintenance DNA methylation process, it is not excluded that SCML2 may need to integrate additional queues from the surrounding chromatin environment during DNA replication for proper stimulation of Usp7. *In vivo*, mutation of key residues in the second MBT domain of the *Drosophila* Scm results in embryonic lethality [102]. Removal of the SPM domain of Scm prevents heterochromatin spreading [104]. It was previously suggested that this effect was caused by SCML2's SPM-mediated interaction with the polyhomeotic-like protein 1 (PHC1), for incorporation into the PRC1 complex [106]. We have shown that SCML2 may have the ability to form homotypic multimers (Figure 3.26). While the SPM-mediated association with the polycomb complex is not excluded by our observation, one needs to account for the fact that the SPM domain may be unavailable when involved in homotypic interactions.

We showed that UHRF1 and SCML2 chromatin targeting is sensitive to the degradation of nuclear RNA (Figure 3.19). Both UHRF1 and SCML2 have been proposed to bind RNA before [100], [208]. No connection has to date been documented between SCML2 and UHRF1. It is thus possible that SCML2 and UHRF1 bound the same RNA or two separate RNA molecules.

The long noncoding RNA UHRF1 protein-associated transcript (UPAT) was shown to stabilise UHRF1 by limiting its auto-ubiquiytlation activity [208]. UHRF1 was found enriched on the H3K23ub-containing chromatin (Figure 3.17C, Figure 3.18, Figure 3.19) which suggested that its ubiquitylation activity may still play an important role after setup of this mark. We hypothesise that RNA stabilization of UHRF1 on the H3K23ub chromatin is important to maintain the appropriate balance between ubiquitylation and deubiuquitylation of the H3 tail on the freshly replicated chromatin to drive the maintenance methylation process forward in an error-free manner (Figure 4.3).

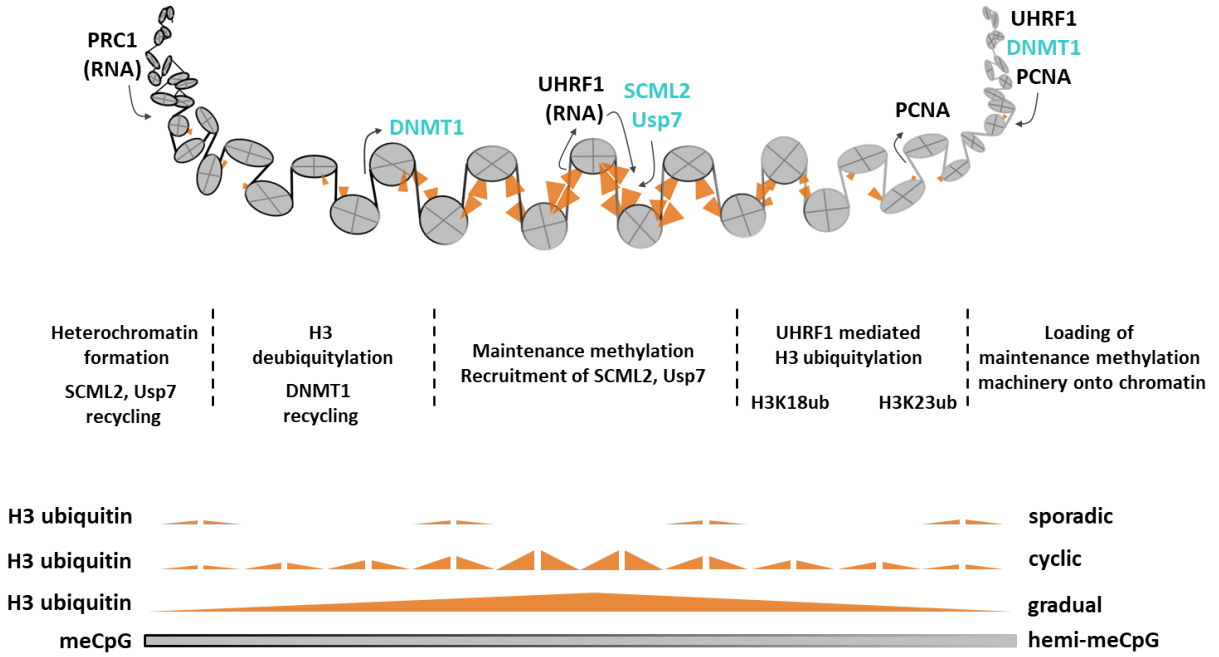


Figure 4.3: Updated order of events on the H3ub chromatin. The recruitment of nuclear proteins to the H3ub chromatin is controlled by the cross-talk between DNA methylation and histone ubiquitylation. Hemi-methylated DNA is converted to fully methylated DNA, in a linear irreversible process. Differently than methylation, ubiquitin is first deposited, after which it is erased. As such, three ubiquitylation/deubiquitylation patterns may form: a gradual pattern, which follows closely the concentration of UHRF1 and DNMT1; a cyclic pattern, where Usp7 and UHRF1 coordinate to establish an optimal amount of H3 ubiquitylation; a sporadic pattern, where a high local density of CpG sites signals to SCML2, Usp7 and UHRF1 to adapt DNMT1 for processive methylation. Timing and efficiency of the recycling of H3ub interactors may constitute the driving force of all the events occurring on the H3ub chromatin fiber and as such a combination of the suggested ubiquitylation/deubiquitylation patterns may form. RNA may be involved in the regulation of UHRF1 stability and SCML2-mediated recruitment of complexes dedicated to heterochromatin formation.

The long noncoding HOX transcript antisense RNA (HOTAIR) was proposed to recruit SCML2 to PRC1-targetted chromatin [100]. We have shown that the RBR-DUF region of SCML2 is sufficient *in vitro* for SCML2 stimulation of Usp7 function on chromatin (Figure 3.28). Previously, the isolated RBR domain of SCML2 was found to be able to simultaneously bind a nucleosome and a short fragment of the HOTAIR transcript [100]. We argue that RNA is not needed for the direct binding of SCML2 to the H3 ubiquitylated chromatin, but that SCML2 may exist inside the nuclues in ribonucleoprotein complexes whose interaction with chromatin is stabilised by RNA, which may act as intermolecular scaffolds. Along this line of argumentation, RNA may also be needed for recruitment of polycomb complexes to SCML2 coated chromatin. In this event, SCML2 would mark nucleosomes for setup and spreading of heterochromatic marks after the maintenance DNA methylation process is complete (Figure 4.2) [209].

Conclusions and Perspectives

Histone ubiquitylation is the first post-translational modification which was identified on histones. The modification changes the physical and chemical properties of the marked histone. Until recently, using traditional endogenous purification protocols or *in vitro* ubiquitylation strategies, only a handful of ubiquitylated histones could be prepared. The study of low abundance histone ubiquitylation marks only became possible recently with the advent of expressed protein ligation. Using this method, we presented the preparation of H2AK119ub, H3K18ub, H3K23ub and H3K18/23ub2 histones which were successfully incorporated into nucleosomal arrays. The optimised strategy which we described for preparation of H2AK119ub and H3K18/23ub can be expanded to synthesise several other site-specific ubiquitylation marks on any of the core histone proteins.

Several histone, mononucleosome and chromatin affinity purification experiments were coupled with quantitative mass spectrometry to understand the requirements of histone ubiquitylation readouts. First, we showed that, for recognition of the ubiquitylation mark beyond the direct reader, histones need to be imbedded minimally within a nucleosome. Second, we showed that the different ubiquitylation marks generally recruit unique proteins and protein complexes to the nucleosomal arrays. These observations suggest that, in the presence of the nuclear extract, in parallel to the readout process, several biochemical transactions and structural changes occur, which specify the identity of the interacting proteins and the biological process they are involved in. The mass spectrometry analyses found several novel ubiquitylation-specific factors. In addition to this, network analyses suggested many novel associations between the enriched proteins. Further experiments would be needed to dissect the roles of the newly identified factors, the functional relevance of the suggested protein associations and the importance of the accompanying biochemical reactions in relaying signalling events downstream.

To verify the mass spectrometry identifications, we focused on the N-terminal ubiquitylation of histone H3, where we described the reproducible enrichment of DNMT1, Usp7 and SCML2 on the modified chromatin. We showed that DNMT1 recruits Usp7 and SCML2 to the marked chromatin, where SCML2 stimulates Usp7's activity, an effect which we found to be inhibited by DNMT1. Using crosslinking mass spectrometry we showed that SCML2 and DNMT1 compete for the same binding surface on Usp7 and found that SCML2 positions Usp7 close to the N-terminal tail of histone H3. We proposed that SCML2 controls DNMT1 recycling from H3 ubiquitylated chromatin by stimulating Usp7's deubiquitylation activity. To test this hypothesis, we suggest two complementary lines of experiments. On one hand,

the levels of ubiquitylated H3 and the levels of CpG methylation in live cells would need to be measured in the presence and absence (transient knock-down or knock-out) of SCML2. In addition, nuclear extracts devoid of SCML2 may be used to further study Usp7's activity on recombinant H3 ubiquitylated chromatin. On the other hand, an *in vitro* methylation assay that can distinguish between a processive and a distributive DNMT1 methylation behaviour can be developed. Such an assay makes use of ubiquitylated nucleosomal arrays, where the density of ubiquitin and the positioning of the ubiquitylated nucleosome(s) can be controlled. The influence of Usp7 and/or SCML2 on DNMT1 activity can directly be monitored as a function of the methylation throughput. The necessity of additional nucleosomal remodeling or histone chaperoning for DNMT1 activity, can also be experimentally addressed in such an assay.

Bibliography

[1] A. D. Hershey and M. Chase, "Independent functions of viral protein and nucleic acid in growth of bacteriophage," *Journal of General Physiology*, vol. 36, no. 1, pp. 39–56, 1952.

[2] I. H. G. S. Consortium, "Finishing the euchromatic sequence of the human genome," *Nature*, vol. 431, p. 931, 2004.

[3] W. Flemming, "Beitraege zur kenntniss der zelle und ihrer lebenserscheinungen," *Archiv für mikroskopische Anatomie*, vol. 20, pp. 1–86, Dec 1878.

[4] R. D. Kornberg, "Chromatin structure: A repeating unit of histones and DNA," *Science*, vol. 184, no. 4139, pp. 868–871, 1974.

[5] K. Luger, A. W. Mder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, "Crystal structure of the nucleosome core particle at 2.8 å resolution," *Nature*, vol. 389, pp. 251–260, 1997.

[6] S. Vijay-Kumar, C. E. Bugg, and W. J. Cook, "Structure of ubiquitin refined at 1.8å resolution," *Journal of Molecular Biology*, vol. 194, no. 3, pp. 531 – 544, 1987.

[7] A. Valouev, S. M. Johnson, S. D. Boyd, C. L. Smith, A. Z. Fire, and A. Sidow, "Determinants of nucleosome organization in primary human cells," *Nature*, vol. 474, p. 516, 2011.

[8] G. Arents and E. N. Moudrianakis, "The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization," *Proceedings of the National Academy of Sciences*, vol. 92, no. 24, pp. 11170–11174, 1995.

[9] T. Schalch, S. Duda, D. F. Sargent, and T. J. Richmond, "X-ray structure of a tetranucleosome and its implications for the chromatin fibre," *Nature*, vol. 436, p. 138, 2005.

[10] C. H. Waddington, "The strategy of the gene," *Allen and Unwin*, 1957.

[11] X. Chen, H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang, J. Jiang, Y.-H. Loh, H. C. Yeo, Z. X. Yeo, V. Narang, K. R. Govindarajan, B. Leong, A. Shahab, Y. Ruan, G. Bourque, W.-K. Sung, N. D. Clarke, C.-L. Wei, and H.-H. Ng, "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells," *Cell*, vol. 133, no. 6, pp. 1106–1117, 2008.

[12] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *Cell*, vol. 126, no. 4, pp. 663–676, 2006.

[13] F. Eckhardt, J. Lewin, R. Cortese, V. K. Rakyan, J. Attwood, M. Burger, J. Burton, T. V. Cox, R. Davies, T. A. Down, C. Haefliger, R. Horton, K. Howe, D. K. Jackson, J. Kunde, C. Koenig, J. Liddle, D. Niblett, T. Otto, R. Pettett, S. Seemann, C. Thompson, T. West, J. Rogers, A. Olek, K. Berlin, and S. Beck, "DNA methylation profiling of human chromosomes 6, 20 and 22," *Nature Genetics*, vol. 38, p. 1378, 2006.

[14] M. Yu, G. C. Hon, K. E. Szulwach, C.-X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J.-H. Min, P. Jin, B. Ren, and C. He, "Base-resolution analysis of 5-Hydroxymethylcytosine in the mammalian genome," *Cell*, vol. 149, no. 6, pp. 1368 – 1380, 2012.

[15] D. Takai and P. A Jones, "Comprehensive analysis of CpG islands in human chromosomes 21 and 22.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 3740–3745, 2002.

[16] A. Bird, M. Taggart, M. Frommer, O. J. Miller, and D. Macleod, "A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA," *Cell*, vol. 40, no. 1, pp. 91 – 99, 1985.

[17] S. C. Wu and Y. Zhang, "Active DNA demethylation: many roads lead to Rome," *Nature reviews. Molecular cell biology*, vol. 11, p. 607620, September 2010.

[18] A. Jeltsch and R. Z. Jurkowska, "New concepts in DNA methylation," *Trends in Biochemical Sciences*, vol. 39, no. 7, pp. 310–318, 2014.

[19] H. Huang, B. R. Sabari, B. A. Garcia, C. D. Allis, and Y. Zhao, "Snapshot: histone modifications," *Cell*, vol. 159, p. 458, 2014.

[20] S. B. Rothbart and B. D. Strahl, "Interpreting the language of histone and DNA modifications," *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1839, no. 8, pp. 627 – 643, 2014. Molecular mechanisms of histone modification function.

[21] M. M. Mueller and T. W. Muir, "Histones: At the crossroads of peptide and protein chemistry," *Chemical Reviews*, vol. 115, no. 6, pp. 2296–2349, 2015.

[22] B. Fierz, C. Chatterjee, R. McGinty, M. Bar-Dagan, D. Raleigh, and M. T.W., "Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction," *Nature Chemical Biology*, vol. 7, no. 2, pp. 113–119, 2011.

[23] G. T Debelouchina, K. Gerecht, and T. W Muir, "Ubiquitin utilizes an acidic surface patch to alter chromatin structure," *Nature chemical biology*, vol. 13, 11 2016.

[24] B. D. Strahl and C. D. Allis, "The language of covalent histone modifications," *Nature*, vol. 403, p. 41, 2000.

[25] Y. Shi and Y. ichi Tsuakada, "Epigenetics. Chapter 2: The discvoery of histone demethylases," *Cold Spring Harbour Laboratory press*, vol. 2, p. 13, 2014.

[26] E. Seto and M. Yoshida, "Epigenetics. Chapter 5: Erasers of histone acetylation: the histone deacetylase enzymes," *Cold Spring Harbour Laboratory press*, vol. 2, p. 143, 2014.

[27] S. Henikoff and M. Smith, "Epigenetics. Chapter 20: Histone variants and epigenetics," *Cold Spring Harbour Laboratory press*, vol. 2, p. 529, 2014.

[28] P. Bekker and J. Workmann, "Epigenetics. Chapter 21: Nucleosome remodelling and epigenetics," *Cold Spring Harbour Laboratory press*, vol. 2, p. 555, 2014.

[29] J. Rinn, "Epigenetics. Chapter 2: lncRNAs:linking RNA to chromatin," *Cold Spring Harbour Laboratory press*, vol. 2, p. 18, 2014.

[30] J. P. Thomson, P. J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A. R. W. Kerr, A. Deaton, R. Andrews, K. D. James, D. J. Turner, R. Illingworth, and A. Bird, "CpG islands influence chromatin structure via the CpG-binding protein CFP1," *Nature*, vol. 464, p. 10821086, April 2010.

[31] N. P. Blackledge, J. C. Zhou, M. Y. Tolstorukov, A. M. Farcas, P. J. Park, and R. J. Klose, "CpG islands recruit a histone H3 lysine 36 demethylase," *Molecular Cell*, vol. 38, no. 2, pp. 179 – 190, 2010.

[32] J. T. Neil Blackledge and P. Skene, "Epigenetics. Chapter 2: CpG island chromatin is shaped by recruitment of ZF-CxxC proteins," *Cold Spring Harbour Laboratory press*, vol. 2, p. 30, 2014.

[33] M. Bostick, J. K. Kim, P.-O. Esteve, A. Clark, S. Pradhan, and S. E. Jacobsen, "UHRF1 plays a role in maintaining DNA methylation in mammalian cells," *Science*, vol. 317, no. 5845, pp. 1760–1764, 2007.

[34] H. Hashimoto, J. R. Horton, X. Zhang, M. Bostick, S. E. Jacobsen, and X. Cheng, "The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix," *Nature*, vol. 455, p. 826, 2008.

[35] A. Nishiyama, L. Yamaguchi, J. Sharif, Y. Johmura, K. S. S. Kawamura, Takeshi-and Nakanishi, K. Arita, T. Kodama, F. Ishikawa, H. Koseki, and M. Nakanishi, "Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication," *Nature*, vol. 502, 2013.

[36] W. Qin, P. Wolf, N. Liu, M. Link, Stephanieand Smets, I. P. G. H. D. Mastra, Federica Laand Forn, K. Fellinger, F. Spada, I. M. Bonapace, A. Imhof, H. Harz, and H. Leonhardt, "DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination," *Nature Communications*, vol. 4, 2013.

[37] S. Ishiyama, A. Nishiyama, Y. Saeki, K. Moritsugu, D. Morimoto, L. Yamaguchi, N. Arai, R. Matsumura, T. Kawakami, Y. Mishima, H. Hojo, S. Shimamura, F. Ishikawa, S. Tajima, K. Tanaka, M. Ariyoshi, M. Shirakawa, M. Ikeguchi, A. Kidera, I. Suetake, K. Arita, and M. Nakanishi, "Structure of the Dnmt1 reader module complexed with a unique two-mono-ubiquitin mark on histone H3 reveals the basis for DNA methylation maintenance," *Molecular Cell*, vol. 68, no. 2, pp. 1097–2765, 2017.

[38] J. D. Lewis, R. R. Meehan, W. J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein, and A. Bird, "Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA," *Cell*, vol. 69, no. 6, pp. 905 – 914, 1992.

[39] F. Fuks, P. J. Hurd, D. Wolf, X. Nan, A. P. Bird, and T. Kouzarides, "The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation," *Journal of Biological Chemistry*, vol. 278, no. 6, pp. 4035–4040, 2003.

[40] A. J. Bannister, P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, and T. Kouzarides, "Selective recognition of methylated lysine 9 on histone h3 by the hp1 chromo domain," *Nature*, vol. 410, p. 296, 2001.

[41] C. Maison and G. Almouzni, "HP1 and the dynamics of heterochromatin maintenance," *Nature Reviews Molecular Cell Biology*, vol. 5, p. 296, 2004.

[42] G. Goldstein, M. Scheid, U. Hammerling, D. Schlesinger, H. Niall, and E. Boyse, "Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 72, p. 1115, January 1975.

[43] I. L. Goldknopf, C. W. Taylor, R. M. Baum, L. C. Yeoman, M. O. Olson, A. W. Prestayko, and H. Busch, "Isolation and characterization of protein A24, a "histone-like" non-histone chromosomal protein.," *Journal of Biological Chemistry*, vol. 250, no. 18, pp. 7182–7187, 1975.

[44] I. L. Goldknopf and H. Busch, "Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24," *Proceedings of the National Academy of Sciences*, vol. 74, no. 3, pp. 864–868, 1977.

[45] A. Ciehanover, Y. Hod, and A. Hershko, "A heat-stable polypeptide component of an atp-dependent proteolytic system from reticulocytes," *Biochemical and Biophysical Research Communications*, vol. 81, no. 4, pp. 1100 – 1105, 1978.

[46] A. Hershko, A. Ciechanover, H. Heller, A. L. Haas, and I. A. Rose, "Proposed role of atp in protein breakdown: conjugation of protein with multiple chains of the polypeptide of atp-dependent proteolysis," *Proceedings of the National Academy of Sciences*, vol. 77, no. 4, pp. 1783–1786, 1980.

[47] K. D. Wilkinson, M. K. Urban, and A. L. Haas, "Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes.," *Journal of Biological Chemistry*, vol. 255, no. 16, pp. 7529–32, 1980.

[48] I. L. Goldknopf and H. Busch, "N-bromosuccinimide fragments of protein A24 (uH2A): An implication that ubiquitin is the precursor of conjugation in vivo," *Biochemical and Biophysical Research Communications*, vol. 96, no. 4, pp. 1724 – 1731, 1980.

[49] M. H.P. West and W. M. Bonner, "Histone H2B can be modified by attachment of ubiquitin," *Nucleic acids research*, vol. 8, pp. 4671–80, 11 1980.

[50] A. Thorne, P. Sautiere, G. Briand, and C. Crane-Robinson, "The structure of ubiquitinated histone H2B," *EMBO Journal*, vol. 6, no. 4, pp. 1005–1010, 1987.

[51] B. Nickel, C. Allis, and J. Davie, "Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin," *Biochemistry*, vol. 28, pp. 958–63, 03 1989.

[52] H. Y. Chen, J.-M. Sun, Y. Zhang, J. R. Davie, and M. L. Meistrich, "Ubiquitination of histone H3 in elongating spermatids of rat testes," *Journal of Biological Chemistry*, vol. 273, no. 21, pp. 13165–13169, 1998.

[53] A. Pham and F. Sauer, "Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in drosophila," vol. 289, pp. 2357–60, 10 2000.

[54] D. S. Kirkpatrick, C. Denison, and S. P. Gygi, "Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics," *Nature cell biology*, vol. 7, p. 750757, August 2005.

[55] R. Y. Tweedie-Cullen, J. M. Reck, and I. M. Mansuy, "Comprehensive mapping of post-translational modifications on synaptic, nuclear, and histone proteins in the adult mouse brain," *Journal of Proteome Research*, vol. 8, no. 11, pp. 4966–4982, 2009.

[56] J. M. R. Danielsen, K. B. Sylvestersen, S. Bekker-Jensen, D. Szklarczyk, J. W. Poulsen, H. Horn, L. J. Jensen, N. Mailand, and M. L. Nielsen, "Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level," *Molecular and cellular proteomics*, vol. 10, p. M110.003590, March 2011.

[57] N. D. Udeshi, P. Mertins, T. Svinkina, and S. A. Carr, "Large-scale identification of ubiquitination sites by mass spectrometry," *Nature protocols*, vol. 8, p. 1950, 2013.

[58] L. Wu, B. M. Zee, Y. Wang, B. A. Garcia, and Y. Dou, "The RING finger protein MSL2 in the MOF complex is an E3 ubiquitin ligase for H2B K34 and is involved in crosstalk with H3 K4 and K79 methylation," *Molecular Cell*, vol. 43, no. 1, pp. 132 – 144, 2011.

[59] F. Mattiroli, J. H.A., W. J. Vissers, van Dijk, I. Pauline, C. Elisabetta, V. Wim, J. A., Marteijn, and S. Titia K., "RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling," *Cell*, vol. 150, no. 6, pp. 1182–1195, 2012.

[60] J. Han, H. Zhang, H. Zhang, Z. Wang, H. Zhou, and Z. Zhang, "A Cul4 E3 ubiquitin ligase regulates histone hand-off during nucleosome assembly," *Cell*, vol. 155, no. 4, pp. 817 – 829, 2013.

[61] J. Li, Q. He, Y. Liu, S. Liu, S. Tang, C. Li, D. Sun, X. Li, M. Zhou, P. Zhu, G. Bi, Z. Zhou, J.-S. Zheng, and C. Tian, "Chemical synthesis of K34-ubiquitylated H2B for nucleosome reconstitution and single-particle cryo-electron microscopy structural analysis," *ChemBioChem*, vol. 18, no. 2, pp. 176–180, 2017.

[62] D. Komander and M. Rape, "The ubiquitin code," *Annual Review of Biochemistry*, vol. 81, no. 1, pp. 203–229, 2012.

[63] J. M. Winget and T. Mayor, "The diversity of ubiquitin recognition: Hot spots and varied specificity," *Molecular Cell*, vol. 38, no. 5, pp. 627 – 635, 2010.

[64] K. Husnjak and I. Dikic, "Ubiquitin-binding proteins: Decoders of ubiquitin-mediated cellular functions," *Annual Review of Biochemistry*, vol. 81, no. 1, pp. 291–322, 2012. PMID: 22482907.

[65] V. M. Weake and J. L. Workman, "Histone ubiquitination: Triggering gene activity," *Molecular Cell*, vol. 29, no. 6, pp. 653 – 663, 2008.

[66] B. Schuettengruber, H.-M. Bourbon, L. Di Croce, and G. Cavalli, "Genome regulation by Polycomb and Trithorax: 70 years and counting," *Cell*, vol. 171, no. 1, p. 34, 2017.

[67] A. Hershko and A. Ciechanover, "The ubiquitin system," *Annual Review of Biochemistry*, vol. 67, no. 1, pp. 425–479, 1998.

[68] B. Hu, Qi, M. Victoria, G. Cui, D. Zhao, and G. Mer, "Mechanisms of ubiquitin-nucleosome recognition and regulation of 53BP1 chromatin recruitment by RNF168/169 and RAD18," *Molecular Cell*, vol. 66, no. 4, pp. 473–487, 2017.

[69] M. Hu, P. Li, M. Li, W. Li, T. Yao, J.-W. Wu, W. Gu, R. E. Cohen, and Y. Shi, "Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde," *Cell*, vol. 111, no. 7, pp. 1041 – 1054, 2002.

[70] L. Rouge, T. Bainbridge, M. Kwok, R. Tong, P. DiLello, I. Wertz, T. Maurer, J. Ernst, and J. Murray, "Molecular understanding of USP7 substrate recognition and C-terminal activation," *Structure*, vol. 24, no. 8, pp. 1335–1345, 2016.

[71] M. D. Wilson, S. Benlekbir, A. Fradet-Turcotte, A. Sherker, J.-P. Julien, A. McEwan, S. M. Noordermeer, F. Sicheri, J. L. Rubinstein, and D. Durocher, "The structural basis of modified nucleosome recognition by 53BP1," *Nature*, vol. 536, 2016.

[72] M. T. Morgan, M. Haj-Yahya, A. E. Ringel, P. Bandi, A. Brik, and C. Wolberger, "Structural basis for histone h2b deubiquitination by the saga dub module," *Science*, vol. 351, no. 6274, pp. 725–728, 2016.

[73] A. Varshavsky, "The early history of the ubiquitin field," *Protein Science*, vol. 15, no. 3, pp. 647–654, 2006.

[74] R. Kalb, S. Latwiel, H. I. Baymaz, P. W. T. C. Jansen, C. W. Mller, M. Vermeulen, and J. Mller, "Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression," *Nature Structural and Molecular Biology*, vol. 21, 2014.

[75] S. Cooper, M. Dienstbier, R. Hassan, L. Schermelleh, J. Sharif, N. P. Blackledge, V. De Marco, S. Elderkin, H. Koseki, R. Klose, A. Heger, and N. Brockdorff, "Targeting Polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment," *Cell Reports*, vol. 7, pp. 1456–1470, 2014.

[76] N. P. Blackledge, A. M. Farcas, T. Kondo, H. W. King, J. F. McGouran, L. L. Hanssen, S. Ito, S. Cooper, K. Kondo, Y. Koseki, T. Ishikura, H. K. Long, T. W. Sheahan, N. Brockdorff, B. M. Kessler, H. Koseki, and R. J. Klose, "Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and Polycomb domain formation," *Cell*, vol. 157, no. 6, pp. 1445 – 1459, 2014.

[77] S. Jentsch, J. P. McGrath, and A. Varshavsky, "The yeast DNA repair gene RAD6 ecodes a ubiquitin-conjugating enzyme," *Nature*, vol. 329, p. 131, 1996.

[78] Z.-W. Sun and C. D. Allis, "Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast," *Nature*, vol. 418, pp. 104–108, 2002.

[79] R. Pavri, B. Zhu, G. Li, P. Trojer, S. Mandal, A. Shilatifard, and D. Reinberg, "Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II," *Cell*, vol. 125, no. 4, pp. 703 – 717, 2006.

[80] R. McGinty, J. Kim, C. Chatterjee, and M. T. R.G. Roeder, "Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation.," *Nature*, vol. 453, no. 7196, pp. 812–816, 2008.

[81] N. Minsky, E. Shema, Y. Field, M. Schuster, E. Segal, and M. Oren, "Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells," *Nature Cell Biology*, vol. 10, p. 483, 2008.

[82] E. Shema, I. Tirosh, Y. Aylon, J. Huang, C. Ye, N. Moskovits, N. Raver-Shapira, N. Minsky, J. Pirngruber, G. Tarcic, P. Hublarova, L. Moyal, M. Gana-Weisz, Y. Shiloh, Y. Yarden, S. A. Johnsen, B. Vojtesek, S. L. Berger, and M. Oren, "The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression," *Genes and Development*, vol. 22, no. 19, pp. 2664–2676, 2008.

[83] E. Shema-Yaacoby, M. Nikolov, M. Haj-Yahya, P. Siman, E. Allemand, Y. Yamaguchi, C. Muchardt, H. Urlaub, A. Brik, M. Oren, and W. Fischle, "Systematic identification of proteins binding to chromatin-embedded ubiquitylated H2B reveals recruitment of SWI/SNF to regulate transcription," *Cell Reports*, vol. 4, no. 3, pp. 601 – 608, 2013.

[84] T. Prenzel, Y. Begus-Nahrmann, F. Kramer, M. Hennion, C. Hsu, T. Gorsler, C. Hintermair, D. Eick, E. Kremmer, M. Simons, T. Beissbarth, and S. A. Johnsen, "Estrogen-dependent gene transcription in human breast cancer cells relies upon proteasome-dependent monoubiquitination of histone H2B," *Cancer Research*, vol. 71, no. 17, pp. 5739–5753, 2011.

[85] O. Tarcic, I. S. Pateras, T. Cooks, E. Shema, J. Kanterman, H. Ashkenazi, H. Boocholez, A. Hubert, R. Rotkopf, M. Baniyash, E. Pikarsky, V. G. Gorgoulis, and M. Oren, "RNF20 links histone H2B ubiquitylation with inflammation and inflammation-associated cancer," *Cell Reports*, vol. 14, no. 6, pp. 1462 – 1476, 2016.

[86] O. Karpiuk, Z. Najafova, F. Kramer, M. Hennion, C. Galonska, A. Knig, N. Snaidero, T. Vogel, A. Shchebet, Y. Begus-Nahrmann, M. Kassem, M. Simons, H. Shcherbata, T. Beissbarth, and S. A. Johnsen, "The histone H2B monoubiquitination regulatory pathway is required for differentiation of multipotent stem cells," *Molecular Cell*, vol. 46, no. 5, pp. 705 – 713, 2012.

[87] S. A. Wagner, P. Beli, B. T. Weinert, M. L. Nielsen, J. Cox, M. Mann, and C. Choudhary, "A proteome-wide, quantitative survey of itin vivo ubiquitylation sites reveals widespread regulatory roles," *Molecular and Cellular Proteomics*, vol. 10, no. 10, 2011.

[88] J. S. Harrison, E. M. Cornett, D. Goldfarb, P. A. DaRosa, Z. M. Li, F. Yan, B. M. Dickson, A. H. Guo, D. V. Cantu, L. Kaustov, P. J. Brown, C. H. Arrowsmith, D. A. Erie, M. B. Major, R. E. Klevit, K. Krajewski, B. Kuhlman, B. D. Strahl, and S. B. Rothbart, "Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by fUHRF1," *eLife*, vol. 5, p. e17101, sep 2016.

[89] H. Leonhardt, A. W. Page, H.-U. Weier, and T. H. Bestor, "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei," *Cell*, vol. 71, no. 5, pp. 865 – 873, 1992.

[90] P. Bashtrykov, G. Jankevicius, R. Z. Jurkowska, S. Ragozin, and A. Jeltsch, "The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism," *Journal of Biological Chemistry*, vol. 289, no. 7, pp. 4106–4115, 2014.

[91] Z.-M. Zhang, S. B. Rothbart, D. F. Allison, Q. Cai, J. S. Harrison, L. Li, Y. Wang, B. D. Strahl, G. G. Wang, and J. Song, "An allosteric interaction links USP7 to deubiquitination and chromatin targeting of UHRF1," *Cell Reports*, vol. 12, no. 9, pp. 1400 – 1406, 2015.

[92] J. Cheng, H. Yang, J. Fang, L. Ma, R. Gong, P. Wang, Z. Li, and Y. Xu, "Molecular mechanism for USP7-mediated DNMT1 stabilization by acetylation," *Nature communications*, vol. 6, p. 7023, 05 2015.

[93] L. Yamaguchi, A. Nishiyama, T. Misaki, Y. Johmura, J. Ueda, K. Arita, K. Nagao, C. Obuse, and M. Nakanishi, "Usp7-dependent histone H3 deubiquitylation regulates maintenance of DNA methylation," *Scientific Reports*, vol. 7, no. 1, 2017.

[94] A. Fernndez-Montalvn, T. Bouwmeester, G. Joberty, R. Mader, M. Mahnke, B. Pierrat, J.-M. Schlaeppi, S. Worpenberg, and B. Gerhartz, "Biochemical characterization of USP7 reveals post-translational modification sites and structural requirements for

substrate processing and subcellular localization," *FEBS Journal*, vol. 274, no. 16, pp. 4256–4270, 2007.

[95] A. C. Faesen, A. M. Dirac, A. Shanmugham, H. Ovaa, A. Perrakis, and T. K. Sixma, "Mechanism of USP7/HAUSP activation by its C-terminal ubiquitin-like domain and allosteric regulation by GMP-Synthetase," *Molecular Cell*, vol. 44, no. 1, pp. 147–159, 2011.

[96] J. A van der Knaap, P. k. Bajpe, Y. Moshkin, K. Langenberg, J. Krijgsveld, A. Heck, F. Karch, and C. Peter Verrijzer, "GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7," *Molecular Cell*, vol. 17, pp. 695–707, 04 2005.

[97] M. E. Sowa, E. J. Bennett, S. P. Gygi, and J. W. Harper, "Defining the human deubiquitinating enzyme interaction landscape," *Cell*, vol. 138, no. 2, pp. 389–403, 2009.

[98] E. Lecona, V. Narendra, and D. Reinberg, "Usp7 cooperates with scml2 to regulate the activity of prc1," *Molecular and Cellular Biology*, vol. 35, no. 7, pp. 1157–1168, 2015.

[99] E. Montini, G. Buchner, C. Spalluto, G. Andolfi, A. Caruso, J. T. den Dunnen, D. Trump, M. Rocchi, A. Ballabio, and B. Franco, "Identification of SCML2, a second human gene homologous to thedrosophila Sex comb on midleg(Scm): A new gene cluster on Xp22," *Genomics*, vol. 58, no. 1, pp. 65 – 72, 1999.

[100] R. Bonasio, E. Lecona, V. Narendra, P. Voigt, F. Parisi, Y. Kluger, and D. Reinberg, "Interactions with RNA direct the Polycomb group protein SCML2 to chromatin where it represses target genes," *eLife*, vol. 3, p. e02637, jul 2014.

[101] I. Bezsonova, "Solution nmr structure of the dna-binding domain from scml2 (sex comb on midleg-like 2)," *Journal of Biological Chemistry*, vol. 289, no. 22, pp. 15739–15749, 2014.

[102] C. Grimm, A. G. de Ayala Alonso, V. Rybin, U. Steuerwald, N. Ly-Hartig, W. Fischle, J. Müller, and C. W. Müller, "Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein Sex comb on midleg," *EMBO reports*, vol. 8, no. 11, pp. 1031–1037, 2007.

[103] E. Lecona, L. A. Rojas, R. Bonasio, A. Johnston, O. Fernndez-Capetillo, and D. Reinberg, "Polycomb protein SCML2 regulates the cell cycle by binding and modulating CDK/CYCLIN/p21 complexes," *PLOS Biology*, vol. 11, pp. 1–21, 12 2013.

[104] R. R. Roseman, K. Morgan, D. R. Mallin, R. Roberson, T. J. Parnell, D. J. Bornemann, J. A. Simon, and P. K. Geyer, "Long-range repression by multiple polycomb group (PcG) proteins targeted by fusion to a defined DNA-binding domain in Drosophila," *Genetics*, vol. 158, no. 1, pp. 291–307, 2001.

[105] A. J. Peterson, D. R. Mallin, N. J. Francis, C. S. Ketel, J. Stamm, R. K. Voeller, R. E. Kingston, and J. A. Simon, "Requirement for Sex comb on midleg protein interactions in Drosophila Polycomb group repression," *Genetics*, vol. 167, no. 3, pp. 1225–1239, 2004.

[106] C. A. Kim, M. R. Sawaya, D. Cascio, W. Kim, and J. U. Bowie, "Structural organization of a Sex-comb-on-midleg/Polyhomeotic copolymer," *Journal of Biological Chemistry*, vol. 280, no. 30, pp. 27769–27775, 2005.

[107] B. E. Nickel and J. R. Davie, "Structure of polyubiquitinated histone H2A," *Biochemistry*, vol. 28, no. 3, pp. 964–968, 1994.

[108] S. C. Moore, L. Jason, and J. Ausi, "The elusive structural role of ubiquitinated histones," *Biochemistry and Cell Biology*, vol. 80, no. 3, pp. 311–319, 2002.

[109] A. Wood, N. J. Krogan, J. Dover, J. Schneider, J. Heidt, M. A. Boateng, K. Dean, A. Golshani, Y. Zhang, J. F. Greenblatt, M. Johnston, and A. Shilatifard, "Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter," *Molecular Cell*, vol. 11, no. 1, pp. 267 – 274, 2003.

[110] K. AjishKumar, M. Haj-Yahya, D. Olschewski, H. Lashuel, and A. Brik, "Highly efficient and chemoselective peptide ubiquitylation," *Angewandte Chemie International Edition*, vol. 48, no. 43, pp. 8090–8094, 2009.

[111] S. Machida, S. Sekine, Y. Nishiyama, N. Horikoshi, and H. Kurumizaka, "Structural and biochemical analyses of monoubiquitinated human histones H2B and H4," *Open Biology*, vol. 6, no. 6, 2016.

[112] T. Kawakami, Y. Mishima, H. Hojo, and I. Suetake, "Synthesis of ubiquitylated histone H3 using a thiirane linker for chemical ligation," *Journal of Peptide Science*, vol. 23, no. 7-8, pp. 532–538, 2017. PSC-16-0162.R1.

[113] P. Dawson, T. Muir, I. Clark-Lewis, and S. Kent, "Synthesis of proteins by native chemical ligation," *Science*, vol. 266, pp. 776–9, 12 1994.

[114] S. Chong, F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus, and M.-Q. Xu, "Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element," *Gene*, vol. 192, no. 2, pp. 271 – 281, 1997.

[115] S. Virdee, P. B. Kapadnis, T. Elliott, K. Lang, J. Madrzak, D. P. Nguyen, L. Riechmann, and J. W. Chin, "Traceless and site-specific ubiquitination of recombinant proteins," *Journal of the American Chemical Society*, vol. 133, no. 28, pp. 10708–10711, 2011. PMID: 21710965.

[116] R. Hirata and Y. Anraku, "Mutations at the putative junction sites of the yeast VMA1 protein, the catalytic subunit of the vacuolar membrane H+-ATPase, inhibit its processing by protein splicing," *Biochemical and Biophysical Research Communications*, vol. 188, no. 1, pp. 40 – 47, 1992.

[117] F. Perler, E. Davis, G. Dean, F. Gimble, W. Jack, N. Neff, C. Noren, J. Thorner, and M. Belfort, "Protein splicing elements: inteins and exteins: a definition of terms and recommended nomenclature," *Nucleic acids research*, vol. 22, p. 11251127, April 1994.

[118] B. W. Poland, M.-Q. Xu, and F. A. Quiocho, "Structural insights into the protein splicing mechanism of PI-SceI," *Journal of Biological Chemistry*, vol. 275, no. 22, pp. 16408–16413, 2000.

[119] J. Binschik and H. D. Mootz, "Chemical bypass of intein-catalyzed NS acyl shift in protein splicing," *Angewandte Chemie International Edition*, vol. 52, no. 15, pp. 4260–4264, 2013.

[120] H. Steen and M. Mann, "The ABCs (and XYZs) of peptide sequencing," *Nature Reviews Molecular Cell Biology*, vol. 5, p. 699, 2004.

[121] R. S. Johnson, S. A. Martin, K. Biemann, J. Stults, and J. Throck. Watson, "Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: Differentiation of leucine and isoleucine," *Analytical Chemistry*, vol. 59, pp. 2621–5, 12 1987.

[122] D. N. Perkins, D. J. C. Pappin, D. M. Creasy, and J. S. Cottrell, "Probability-based protein identification by searching sequence databases using mass spectrometry data," *Electrophoresis*, vol. 20, no. 18, pp. 3551–3567, 1999.

[123] J. Peng, J. E. Elias, C. C. Thoreen, L. J. Licklider, and S. P. Gygi, "Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LCMS/MS) for large-scale protein analysis: the yeast proteome," *Journal of Proteome Research*, vol. 2, no. 1, pp. 43–50, 2003.

[124] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, and M. Mann, "Andromeda: A peptide search engine integrated into the maxquant environment," *Journal of Proteome Research*, vol. 10, no. 4, pp. 1794–1805, 2011.

[125] J. Cox and M. Mann, "Maxquant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification," *Nature Biotechnology*, vol. 26, pp. 1367–72, 12 2008.

[126] P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, and D. J. Pappin, "Multiplexed protein quantitation in *saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents," *Molecular and Cellular Proteomics*, vol. 3, no. 12, pp. 1154–1169, 2004.

[127] J.-L. Hsu, S.-Y. Huang, N.-H. Chow, and S.-H. Chen, "Stable-isotope dimethyl labeling for quantitative proteomics," *Analytical Chemistry*, vol. 75, no. 24, pp. 6843–6852, 2003.

[128] S.-E. Ong, B. Blagoev, I. Kratchmarova, D. Bach Kristensen, H. Steen, A. Pandey, and M. Mann, "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics," *Molecular and Cellular Proteomics*, vol. 1, pp. 376–86, 06 2002.

[129] J. V. Olsen, S.-E. Ong, and M. Mann, "Trypsin cleaves exclusively C-terminal to arginine and lysine residues," *Molecular Cellular Proteomics*, vol. 3, no. 6, pp. 608–614, 2004.

[130] J. A Ranish, E. Yi, D. M Leslie, S. Purvine, D. Goodlett, J. Eng, and R. Aebersold, "The study of macromolecular complexes by quantitative proteomics," *Nature genetics*, vol. 33, pp. 349–55, 04 2003.

[131] M. Vermeulen, H. C. Eberl, F. Matarese, H. Marks, S. Denissov, F. Butter, K. K. Lee, J. V. Olsen, A. A. Hyman, H. G. Stunnenberg, and M. Mann, "Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers," *Cell*, vol. 142, no. 6, pp. 967 – 980, 2010.

[132] M. Nikolov, A. Sttzer, K. Mosch, A. Krasauskas, S. Soeroes, H. Stark, H. Urlaub, and W. Fischle, "Chromatin affinity purification and quantitative mass spectrometry defining the interactome of histone modification patterns," *Molecular and Cellular Proteomics*, vol. 10, no. 11, 2011.

[133] S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann, and J. Cox, "The Perseus computational platform for comprehensive analysis of (prote)omics data," *Nature Methods*, 2016.

[134] P. Singh, A. Panchaud, and D. R. Goodlett, "Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique," *Analytical Chemistry*, vol. 82, no. 7, pp. 2636–2642, 2010.

[135] Z. A. Chen, A. Jawhari, L. Fischer, C. Buchen, S. Tahir, T. Kamenski, M. Rasmussen, L. Lariviere, J.-C. Bukowski-Wills, M. Nilges, P. Cramer, and J. Rappsilber, "Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry," *The EMBO Journal*, vol. 29, no. 4, pp. 717–726, 2010.

[136] K. Lasker, F. Foerster, S. Bohn, T. Walzthoeni, E. Villa, P. Unverdorben, F. Beck, R. Aebersold, A. Sali, and W. Baumeister, "Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach," *Proceedings of the National Academy of Sciences*, vol. 109, no. 5, pp. 1380–1387, 2012.

[137] C. Ciferri, G. C. Lander, A. Maiolica, F. Herzog, R. Aebersold, and E. Nogales *Molecular architecture of human polycomb repressive complex 2*, vol. 1, p. e00005, oct 2012.

[138] http://www.thermofisher.com/order/catalog/product/21580, February 2018.

[139] J. Seebacher, P. Mallick, N. Zhang, J. S. Eddes, R. Aebersold, and M. H. Gelb, "Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing," *Journal of Proteome Research*, vol. 5, no. 9, pp. 2270–2282, 2006.

[140] B. Yang, Y.-J. Wu, M. Zhu, S.-B. Fan, J. Lin, K. Zhang, S. Li, H. Chi, Y.-X. Li, H.-F. Chen, S.-K. Luo, Y.-H. Ding, L.-H. Wang, Z. Hao, L.-Y. Xiu, S. Chen, K. Ye, S.-M. He, and M.-Q. Dong, "Identification of cross-linked peptides from complex samples," *Nature Methods*, vol. 9, p. 904, 2012.

[141] A. Fradet-Turcotte, M. D. Canny, C. Escribano-Diaz, A. Orthwein, C. C. Y. Leung, H. Huang, M.-C. Landry, J. Kitevski-LeBlanc, S. M. Noordermeer, F. Sicheri, and D. Durocher, "53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark," *Nature*, vol. 499, p. 50, 2013.

[142] Z. Zhang, H. Yang, and H. Wang, "The histone H2A deubiquitinase USP16 interacts with HERC2 and fine-tunes cellular response to DNA damage," *Journal of Biological Chemistry*, 2014.

[143] F. Nicassio, N. Corrado, J. H. Vissers, L. B. Areces, S. Bergink, J. A. Marteijn, B. Geverts, A. B. Houtsmuller, W. Vermeulen, P. P. D. Fiore, and E. Citterio, "Human USP3 is a chromatin modifier required for S phase progression and genome stability," *Current Biology*, vol. 17, no. 22, pp. 1972 – 1977, 2007.

[144] H. Wang, L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, and R. Jones, "Role of histone H2A ubiquitination in Polycomb silencing," *Nature*, vol. 431, pp. 873–8, 11 2004.

[145] S. Cooper, A. Grijzenhout, E. Underwood, K. Ancelin, T. Zhang, T. Nesterova, B. Anil-Kirmizitas, A. Bassett, S. Kooistra, K. Agger, K. Helin, E. Heard, and N. Brockdorff, "Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2," *Nature Communications*, vol. 7, p. 13661, 11 2016.

[146] T. Nakagawa, T. Kajitani, S. Togo, N. Masuko, H. Ohdan, Y. Hishikawa, T. Koji, T. Matsuyama, T. Ikura, M. Muramatsu, and T. Ito, "Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation," *Genes and Development*, vol. 22, no. 1, pp. 37–49, 2008.

[147] M. Hochstrasser, "Ubiquitin-dependent protein degradation," *Annual Review of Genetics*, vol. 30, pp. 405–439, 1997.

[148] J. C. Scheuermann, A. G. de Ayala Alonso, K. Oktaba, N. Ly-Hartig, R. K. McGinty, S. Fraterman, M. Wilm, T. W. Muir, and J. Mueller, "Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB," *Nature*, vol. 465, p. 243, 2010.

[149] R. Kalb, D. L. Mallery, C. Larkin, J. T. Huang, and K. Hiom, "BRCA1 is a histone H2A-specific ubiquitin ligase," *Cell Reports*, vol. 8, no. 4, pp. 999 – 1005, 2014.

[150] B. Zhu, Y. Zheng, A.-D. Pham, S. S. Mandal, H. Erdjument-Bromage, P. Tempst, and D. Reinberg, "Monoubiquitination of human histone H2B: The factors involved and their roles in HOX gene regulation," *Molecular Cell*, vol. 20, no. 4, pp. 601 – 611, 2005.

[151] N. Minsky and M. Oren, "The RING domain of MDM2 mediates histone ubiquitylation and transcriptional repression," *Molecular Cell*, vol. 16, no. 4, pp. 631 – 639, 2004.

[152] X. Li, P. Trojer, T. Matsumura, J. E Treisman, and N. Tanese, "Mammalian SWI/SNF-A subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B," *Molecular and cellular biology*, vol. 30, pp. 1673–88, 04 2010.

[153] H.-Y. Joo, A. Jones, C. Yang, L. Zhai, A. D. Smith, Z. Zhang, M. B. Chandrasekharan, Z.-w. Sun, M. B. Renfrow, Y. Wang, C. Chang, and H. Wang, "Regulation of histone H2A and H2B deubiquitination and *xenopus* development by USP12 and USP46," *Journal of Biological Chemistry*, vol. 286, no. 9, pp. 7190–7201, 2011.

[154] K. W. Henry, A. Wyce, W.-S. Lo, L. J. Duggan, N. T. Emre, C.-F. Kao, L. Pillus, A. Shilatifard, M. A. Osley, and S. L. Berger, "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8," *Genes and Development*, vol. 17, no. 21, pp. 2648–2663, 2003.

[155] X. Zhang, B. Li, A.-H. Rezaeian, X. Xu, P.-C. Chou, G. Jin, F. Han, B.-S. Pan, C.-Y. Wang, J. Long, A. Zhang, C.-Y. Huang, F.-J. Tsai, C.-H. Tsai, C. Logothetis, and H.-K. Lin, "H3 ubiquitination by NEDD4 regulates H3 acetylation and tumorigenesis," *Nature Communications*, vol. 16, p. 14799, 04 2017.

[156] Z. Deng, H. Liu, and X. Liu, "Rag1-mediated ubiquitylation of histone H3 is required for chromosomal V(D)J recombination," *Cell Research*, vol. 25, p. 181, 2015.

[157] G. Li, T. Ji, J. Chen, Y. Fu, L. Hou, Y. Feng, T. Zhang, T. Song, J. Zhao, Y. Endo, H. Lin, X. Cai, and Y. Cang, "CRL4D/CAF8 ubiquitin ligase targets histone H3K79 and promotes H3K9 methylation in the liver," *Cell Reports*, vol. 18, no. 6, pp. 1499 – 1511, 2017.

[158] Q. Yan, S. Dutt, R. Xu, K. Graves, P. Juszczynski, J. P. Manis, and M. A. Shipp, "BBAP monoubiquitylates histone H4 at lysine 91 and selectively modulates the DNA damage response," *Molecular Cell*, vol. 36, no. 1, pp. 110 – 120, 2009.

[159] https://www.chem-agilent.com/pdf/strata/200249.pdf, February 2018.

[160] http://www.thermofisher.com/order/catalog/product/18290015, February 2018.

[161] https://www.neb.com/products/c2527-bl21de3-competent-e-coli#Product%20Information, February 2018.

[162] https://www.neb.com/products/c2925-dam-dcm-competent-e-coli, March 2014.

[163] https://openwetware.org/wiki/E._coli_genotypes, February 2018.

[164] http://www.thermofisher.com/order/catalog/product/11496015, February 2018.

[165] https://expressionsystems.com/product/insect-cells/, February 2018.

[166] https://www.lgcstandards-atcc.org/products/all/CCL-2.2.aspx?geo_country=de, February 2018.

[167] J. D. Dignam, R. M. Lebovitz, and R. G. Roeder, "Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei.," 1983.

[168] https://www.genomics.agilent.com/primerDesignProgram.jsp, February 2018.

[169] http://www.clontech.com/US/Products/Cloning_and_Competent_Cells/Cloning_Resources/Online_In-Fusion_Tools, February 2018.

[170] http://tools.thermofisher.com/content/sfs/manuals/bactobac_man.pdf, February 2018.

[171] J. Lis and R. Schleif, "Size fractionation of double-stranded DNA by precipitation with polyethylene glycol," *Nucleic Acids Research*, vol. 2, no. 3, pp. 383–389, 1975.

[172] J. Havli, M. Mann, A. Shevchenko, H. Tomas, and J. V. Olsen, "In-gel digestion for mass spectrometric characterization of proteins and proteomes," *Nature Protocols*, vol. 1, 2008.

[173] http://www.coxdocs.org/doku.php?id=maxquant:start, February 2018.

[174] http://www.uniprot.org/, February 2018.

[175] http://www.coxdocs.org/doku.php?id=perseus:start, February 2018.

[176] https://cran.r-project.org/bin/windows/base/, February 2018.

[177] https://string-db.org/, February 2018.

[178] http://pfind.ict.ac.cn/software/plink/, February 2018.

[179] http://crosslinkviewer.org/, February 2018.

[180] C. W. Combe, L. Fischer, and J. Rappsilber, "xiNET: Cross-link network maps with residue resolution," *Molecular Cellular Proteomics*, vol. 14, no. 4, pp. 1137–1147, 2015.

[181] https://pymol.org/2/, February 2018.

[182] K. Luger, T. J. Rechsteiner, and T. J. Richmond, "Preparation of nucleosome core particle from recombinant histones," in *Chromatin*, vol. 304 of *Methods in Enzymology*, pp. 3 – 19, Academic Press, 1999.

[183] D. Shechter, H. Dormann, C. Allis, and S. Hake, "Extraction, purification and analysis of histones," *Nature Protocols*, vol. 2, no. 6, pp. 1445–1457, 2007.

[184] P. Siman, S. V. Karthikeyan, M. Nikolov, W. Fischle, and A. Brik, "Convergent chemical synthesis of histone H2B protein for the site-specific ubiquitination at Lys34," *Angewandte Chemie International Edition*, vol. 52, no. 31, pp. 8059–8063, 2013.

[185] J. S. Godde and A. P. Wolffe, "Disruption of reconstituted nucleosomes. the effect of particle concentration, MgCl2 and KCl conentration, the histone tails and temperature.," *Journal of Biological Chemistry*, vol. 270,, pp. 27399–27402, 1995.

[186] P. Lowary and J. Widom, "New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning," *Journal of Molecular Biology*, vol. 276, no. 1, pp. 19 – 42, 1998.

[187] R. Blossey and H. Schiessel, "The dynamics of the nucleosome: thermal effects, external forces and ATP," *FEBS Journal*, vol. 278, no. 19, pp. 3619–3632, 2011.

[188] J. Sharif, M. Muto, S.-i. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano, and H. Koseki, "The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA," *Nature*, vol. 450, 2007.

[189] E. Hervouet, L. Lalier, E. Debien, M. Cheray, A. Geairon, H. Rogniaux, D. Loussouarn, S. A. Martin, F. M. Vallette, and P.-F. Cartron, "Disruption of dnmt1/pcna/uhrf1 interactions promotes tumorigenesis from human and mice glial cells," *PLOS ONE*, vol. 5, pp. 1–14, 06 2010.

[190] W. J. Lee, J.-Y. Shim, and B. T. Zhu, "Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids," *Molecular Pharmacology*, vol. 68, no. 4, pp. 1018–1030, 2005.

[191] C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond, "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9å resolution.," *Journal of Molecular Biology*, vol. 319, no. 5, pp. 1097 – 1113, 2002.

[192] R. Q. Kim, W. J. van Dijk, and T. K. Sixma, "Structure of USP7 catalytic domain and three UBL-domains reveals a connector α-helix with regulatory role," *Journal of Structural Biology*, vol. 195, no. 1, pp. 11 – 18, 2016.

[193] M. Luo, J. Zhou, N. A. Leu, C. M. Abreu, J. Wang, M. C. Anguera, D. G. de Rooij, M. Jasin, and P. J. Wang, "Polycomb protein SCML2 associates with USP7 and counteracts histone H2A ubiquitination in the XY chromatin during male meiosis," *PLOS Genetics*, vol. 11, pp. 1–17, 01 2015.

[194] Y.-H. Ping and T. M. Rana, "DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation," *Journal of Buichemistry*, vol. 276, no. 16, pp. 12951–12958, 2001.

[195] P. C. Hanawalt and G. Spivak, "Transcription-coupled DNA repair: two decades of progress and surprises," *Nature Reviews Molecular Cell Biology*, vol. 9, p. 958, 2008.

[196] C. Chatterjee, R. K. McGinty, B. Fierz, and T. W. Muir, "Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation," *Nature Chemical Biology*, vol. 6, p. 287, 2010.

[197] M. Gatti, S. Pinato, E. Maspero, P. Soffientini, S. Polo, and L. Penengo, "A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase," *Cell cycle (Georgetown, Tex.)*, vol. 11, p. 25382544, July 2012.

[198] N. L. Samara, A. B. Datta, C. E. Berndsen, X. Zhang, T. Yao, R. E. Cohen, and C. Wolberger, "Structural insights into the assembly and function of the SAGA deubiquitinating module," *Science*, vol. 328, no. 5981, pp. 1025–1029, 2010.

[199] J. C. Rogers, "DNA methylation: Molecular biology and biological significance," *Cell*, vol. 73, no. 3, p. 429, 1993.

[200] N. Nagaraj, J. R. Wisniewski, T. Geiger, J. Cox, M. Kircher, J. Kelso, S. Paabo, and M. Mann, "Deep proteome and transcriptome mapping of a human cancer cell line," *Molecular Systems Biology*, vol. 7, no. 1, pp. 548–n/a, 2011. 548.

[201] J. Charlton, T. L. Downing, Z. D. Smith, H. Gu, K. Clement, R. Pop, V. Akopian, S. Klages, D. P. Santos, A. Tsankov, B. Timmermann, M. J. Ziller, E. Kiskinis, A. Gnirke, and A. Meissner, "Global delay in nascent strand DNA methylation," vol. 25, 03 2018.

[202] C. Xu and V. G. Corces, "Nascent DNA methylome mapping reveals inheritance of hemimethylation at CTCF/cohesin sites," *Science*, vol. 359, no. 6380, pp. 1166–1170, 2018.

[203] R. Goyal, R. Reinhardt, and A. Jeltsch, "Accuracy of DNA methylation pattern preservation by the DNMT1 methyltransferase," vol. 34, pp. 1182–8, 02 2006.

[204] A. Jeltsch, "On the enzymatic properties of DNMT1: Specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme," vol. 1, pp. 63–6, 04 2006.

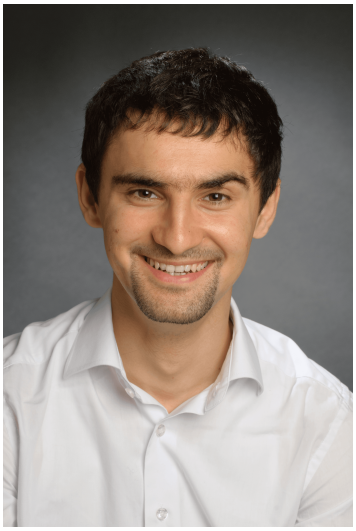
[205] A. Hermann, R. Goyal, and A. Jeltsch, "The DNMT1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites," *Journal of Biological Chemistry*, vol. 279, no. 46, pp. 48350–48359, 2004.

[206] G. Vilkaitis, I. Suetake, S. Klimaauskas, and S. Tajima, "Processive methylation of hemimethylated CpG sites by mouse DNMT1 DNA methyltransferase," *Journal of Biological Chemistry*, vol. 280, no. 1, pp. 64–72, 2005.

[207] A. Schrader, T. Gross, V. Thalhammer, and G. Laengst, "Characterization of DNMT1 binding and DNA methylation on nucleosomes and nucleosomal arrays," *PLOS ONE*, vol. 10, pp. 1–22, 10 2015.

[208] K. Taniue, A. Kurimoto, H. Sugimasa, E. Nasu, Y. Takeda, K. Iwasaki, T. Nagashima, M. Okada-Hatakeyama, M. Oyama, H. Kozuka-Hata, M. Hiyoshi, J. Kitayama, L. Negishi, Y. Kawasaki, and T. Akiyama, "Long noncoding RNA UPAT promotes colon tumorigenesis by inhibiting degradation of UHRF1," *Proceedings of the National Academy of Sciences*, vol. 113, no. 5, pp. 1273–1278, 2016.

[209] L. Ringrose, "Noncoding RNAs in Polycomb and Trithorax regulation: A quantitative perspective," *Annual Review of Genetics*, vol. 51, no. 1, pp. 385–411, 2017.



Stefan-Sebastian David

Christophorusweg 14, 37075, Göttingen
+49 (0)176 87795195
stefansebastian.david@yahoo.com

Education

2018	**Graduate studies** Chromatin Biochemistry Laboratory MPI for Biophysical Chemistry *Georg August University Göttingen*
2014	**Master of Science. GPA: 1.6** IMPRS for Molecular Biology *Georg August University Göttingen*
2012	**Bachelor of Science. GPA: 1.6** Biochemistry and Cell Biology Earth and Space Sciences *Jacobs University Bremen*
2009	**Diplomă de Bacalaureat. GPA: 9.5** *Costache Negruzzi College Iaşi*

Scholarships and Awards

2014	**PhD Fellowship** *Max Planck Society*
2012	**Masters Fellowship** *Max Planck Society*
2011	**Internship Stipend** *Jacobs University Bremen*
2010	**Erasmus Stipend** *Jacobs University Bremen*
2009	**Excellence Diploma** *Romanian Ministry for Education*

Research Experience

April 2014 – April 2018

PhD Thesis
Systematic characterisation of histone ubiquitylation readout

I have completed my PhD in the laboratory of Prof. Wolfgang Fischle, at the Max-Planck-Institute for Biophysical Chemistry. During my thesis, I prepared several ubiquitylated histone constructs and assembled them into chromatin fibers. I used mass spectrometry to find out which nuclear proteins interact with the ubiquitylated chromatin templates and tested the validity of the mass spectrometry identifications by focusing on an a specific protein complex which was enriched on one of the modified chromatin templates.

September 2013 – March 2014

Master's Thesis
Preparation of ubiquitylated histone octamers

During my Master's thesis in the laboratory of Prof. Wolfgang Fischle, at the Max-Planck-Institute for Biophysical Chemistry, I generated truncated histone thioesters as substrates for native chemical ligation. I assembled the ubiquitylated constructs into modified histone octamers.

March 2013 – May 2013

Laboratory rotation
Structural investigations of histone tail folding using FRET

As part of my laboratory rotation, I incorporated unnatural amino acids into histone H2B and used click chemistry to modify pairs of amino acids with compatible fluorophores for FRET experiments.

August 2011 – January 2012

Bachelor Thesis
Recombinant production of histone reader domains

During my Bachelor's thesis in the laboratory of Prof. Albert Jeltsch at the Jacobs University Bremen, I prepared a number of histone post-translational modification reader domains and analysed their binding specificities on spotted histone tail peptide arrays.

January 2012 – June 2012

Bachelor Thesis
Distribution of nitrogen-fixing cyanobacteria in the WTNA

As part of my Bachelor's thesis in the laboraoty of Dr. Rachel Foster at the Max-Planck-Institute for Marine Microbiology, I exracted microbial DNA from water samples and quantified the abundance of the *nifH* gene copy numbers to assess the ability of cyanobacteria to fix nitrogen.

Summer 2011

Summer Internship
Atomic Force Microscopy of microbial symbioses

During my summer internship at the Max-Planck-Insitute for Marine Microbiology, I characterised the intimate interaction between a filamentuous cyanobacterium and its seagrass symbiotic partner. I used AFM to visualise the contact sites which the two organisms were using to exchange nutrients.