2014

Identification of a Functional Hotspot on Ubiquitin Required for Stimulation of Methyltransferase Activity on Chromatin

Matthew Thomas Holt

IDENTIFICATION OF A FUNCTIONAL HOTSPOT ON UBIQUITIN REQUIRED FOR STIMULATION OF METHYLTRANSFERASE ACTIVITY ON CHROMATIN

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Matthew Thomas Holt

June 2014

IDENTIFICATION OF A FUNCTIONAL HOTSPOT ON UBIQUITIN REQUIRED FOR STIMULATION OF METHYLTRANSFERASE ACTIVITY ON CHROMATIN

Matthew Thomas Holt, Ph.D.

The Rockefeller University 2014

Histone post-translational modifications modulate chromatin structure and function either by directly altering the intrinsic physical properties of the chromatin fiber or by nucleating the recruitment and activity of a host of *trans*-acting nuclear factors. Histone ubiquitylation is one class of histone PTMs where the 76 amino acid protein, ubiquitin, is ligated to the ε-nitrogen of a lysine amino acid residue within the histone substrate. 36 unique ubiquitin sites across the 4 canonical histones have been annotated, with 7 of these sites associated with chromatin-templated processes. One such modification, the ubiquitylation of Histone H2B at lysine 120 (H2B-Ub) is enriched at the 5' end of active genes and has been implicated in transcriptional elongation and chromatin conformation, interacting with over 90 *trans*-acting nuclear factors. Further, it is additionally responsible for the regulation of H3K4 and H3K79 methylations, through the direct stimulation of methyltransferase activity specific to the installation of these marks. Interested in the surface features on ubiquitin required for H2B-Ub stimulation of the human H3K79 methyltransferase, hDot1L, we developed a strategy for the site-specific chemical ligation of ubiquitin to pre-assembled mononucleosomes to greatly expedite structure activity studies of

ubiquitin in a nucleosome context. Accordingly, we synthesized a library of H2B-ligated ubiquitin alanine mutant nucleosomes and tested their ability to stimulate hDot1L-mediated H3K79 methylation. A functional hotspot on Ub that is required for the stimulation of hDot1L activity *in vitro* was identified. We additionally investigated the structural implication of this functional hotspot in the context of nucleosomes and nucleosomal array compaction. Both nucleosomal structural and H2B-Ub-induced impairment of chromatin fiber compaction was not affected by this functional surface. Lastly, this functional hotspot was further tested with the yeast H3 methyltransferases, yDot1 and ySet1C. Surprisingly, this functional surface was dispensable for yDot1-mediated H3K79 methylation both *in vivo* and *in vitro*. Further insights into yDot1-mediated H3K79 methylation was obtained through an alanine scan of the Ub surface in the context of the nucleosome. Interestingly, the hDot1L hotspot within ubiquitin was found to be important for the regulation of ySet1-mediated H3K4 methylation. Collectively, these data delineate the multi-functionality of H2B-Ub as a histone post-translational modification.

To my friends and family,

as none of this would have been possible without.

ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor, Tom Muir, for providing me with the opportunity to conduct research under his guidance and for facilitating my growth as a scientist over the years. Considering my unique situation as a graduate student, I am forever grateful of his continued support. I would also like to thank my faculty committee: David Allis, Seth Darst, and Tarun Kapoor, for their ongoing critical evaluation of my research and for providing valuable feedback as well as generous support in my final years at Rockefeller University. Additionally, I need to thank both the members of the Kapoor and Allis laboratories as this is where I conducted a significant amount of my research. Further, I especially appreciate the entirety of the Muir lab for contributing to nearly every aspect of my research –in the guise of being overly critical during group meetings- and for constantly taking the time to help me to succeed. Over the years I've had the opportunity to form many friendships with people within the Muir Lab as well as my fellow classmates which not only benefited my scientific rigor, but additionally allowed for me to enjoy my time here at Rockefeller University. Lastly, I'll always miss my Tuesday morning struggle of getting from NYC to Princeton and the many times I had to email Tom to tell him I was going to be late because the train jumped the tracks, a oil truck exploded under an underpass or I got on the 'wrong' train. Anyways, see you next Tuesday!

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Ac	Acetylation
act-MN	Activated mononucleosome
APAGE	native 1%, agarose, 1% polyacrylamide gel electrophoresis
BP	base pair
CTCF	CCCTC binding factor
ChIP-seq	Chromatin Immunoprecipitation followed by massively parallel DNA sequencing
CBD	Chitin binding domain
COMPASS	Complex of Proteins Associated with Set1
CSP	Chemical shift perturbations
CTD	C-terminal domain
DNA	Deoxyribonucleic acid
Di	Di-ubiquitylated nucleosomes
DiTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
Dot1	Disruptor of Telomeric Silencing 1
DTT	Dithiolthreitol (DTT)
E	Enhancer
E1	Ubiquitin activating enzyme
E2	Ubiquitin conjugating enzyme
E3	Ubiquitin protein ligase
ENCODE	Encyclopedia of DNA Elements
ESI-MS	Electrospray Ionization Mass Spectrometry
EPL	Expressed protein ligation
FACT	Facilitates Chromatin Transcription
H2B-Ub	H2B ubiquitylation at lysine 120
HMGN2	High mobility group nucleosomal 2
Homo-FRET	homo-fluorescence resonance energy transfer
HP1	Heterochomatin Protein 1
HSQC	Heteronuclear single quantum coherence
ILV	Isoleucine, leucine, and valine amino acid residues
ING	Inhibitor of Growth Protein 2
lncRNA	Long non-coding RNA
Me	Methylation
Me1	Mono-methylation
Me2	Di-methylation
Me3	Tri-methylation
MLL	Mixed Linear Leukemia protein
MMTV	Mouse mammory tumor virus DNA sequence
MN	Mononucleosomes
MNase	Micrococcal nuclease
Mono	mono-ubiquitylated nucleosomes
mPEG	Maleimide polyethylene glycol
NCL	Native chemical ligation
Non	non-modified nucleosomes

PAF	RNA Polymerase II Associated Factor
PAGE	polyacrylamide gel electrophoresis
PAR-	Post-translationally modified after reconstitution
PEV	Position Effect Variegation
PF	Promoter Flanking
Ph	Phosphorylation
PHD	Plant Homeo Domain
PRE	paramagnetic relaxation enhancement
PTM	Post-translational modification
R	Repressed state
RNA	Ribonucleic acid
RP-HPLC	Reversed Phase High-Performance Liquid Chromatography
SAGA	Spt-Ada-Gcn5-Acetyl transferase
SAM	S-Adenosyl-L-methionine
SAR	Structure-activity relationship
SDS	sodium dodecyl sulfate
Set1	Su(var)3-9, Enhancer of zeste, and Trx 1
SPPS	Solid-phase peptide synthesis
SSA	Steady-state anisotropy approach
T	Transcribed state
T_2	Relaxation time
τ_c	Rotational correlational time
TAD	Topologically associated domains
TCEP	*Tris*(2-carboxyethyl)phosphine
TF	Transcription Factor
TNB	5-thio-2-nitro-benzoic acid
TROSY	Transverse relaxation-optimized NMR spectroscopy
TSS	Transcriptional Start Site
Ub	Ubiquitylation
Ubl	Ubiquitin like proteins
WE	Weak Enhancer

Chapter 1. Introduction

1.1. Introduction

Deoxyribonucleic acid (DNA) is the carrier of genetic information from which all life as we know it is based. Encoded by the nucleobases of DNA, genes are functional units that are transcribed into ribonucleic acid (RNA), and may be subsequently translated into proteins.[1, 2] Considering the essentiality of this for all of life's processes, organisms have evolved complex mechanisms for the spatiotemporal regulation of their genes. With the sequencing of the human genome, and the recent Encyclopedia of DNA Elements (ENCODE) project that has served to functionally annotate DNA elements, we have progressed far beyond the simple idea that DNA contains genes and have begun to disentangle the structures and mechanisms behind how genetic information is utilized and propagated.

It is clear that DNA sequence alone is not sufficient to explain certain phenotypic or transcriptional phenomena; rather it's the integration of multiple different inputs that determines the transcriptional fate of a gene. This was first observed in 1930, as Position Effect Variegation (PEV), where the translocation of a gene resulted in its variable expression between cells with otherwise identical DNA sequences.[3, 4] This has also been demonstrated in processes such as X-chromosome inactivation and in multicellular development.[5] In order to understand gene function we must consider the context in which DNA exists.

This is the study of epigenetics, which herein, will be narrowly defined as the study of gene function that is not directly attributed to DNA sequence alone.

1.2. Chromatin Structure and Function

In eukaryotes, DNA is segregated into the nucleus of a cell where it exists in chromosomes, or DNA polymers (human chromosomes are between 50 and 250 megabases) stored in the form of the nucleoprotein complex known as chromatin. During interphase, the nuclear structure resembles a fractal globular state where chromosomes occupy discrete chromosomal territories (Figure 1.1a).[7] Gene-rich chromosome domains localize towards the center of the nucleus and gene poor areas to the nuclear periphery.[6, 7] Recently, proximity-based DNA ligation technologies coupled to bioinformatic analysis of annotated DNA elements (discerned through epigenomic methods, such as Chromatin Immunoprecipitation followed by massively parallel DNA sequencing (ChIP-seq)) have markedly enhanced the resolution to which we can explore the large-scale organization of chromatin.[7-9] Accordingly, we have learned from these studies that macro-chromatin structure correlates well with finer chromatin elements (discussed further below) suggesting that the spatial organization of the nucleus may either influence or be a consequence of gene function.

The fundamental repeating unit of chromatin, at the 100 base pair (bp) scale, is the nucleosome. The nucleosome consists of a ~146 bp DNA duplex wrapped around a histone octamer composed of two copies each of histone H2A, histone

H2B, histone H3, and histone H4 (Figure 1.1b).[10] Atomic resolution crystal structures of the nucleosome reveal that DNA is wrapped in 1.65 turns of a left-handed super helix around the histone octamer in which two H3/H4 dimers form a tetramer and associate with individual H2A/H2B dimers on either side of the nucleosome face.[11] Histones possess a canonical alpha-helical histone fold, which forms the structural scaffolding of the octamer core, and a basic N-terminal tail protruding out from the nucleosome (Figure 1.1b).

Intervening DNA connects nucleosomes and their genomic positions have been mapped to base pair resolution using ChIP-Seq as well as Micrococcal Nuclease (MNase) digestion. Accordingly, nucleosomes are repeated throughout the genome about every 160 to 240 bp and nucleosome positioning is determined through both DNA sequence-dependent and epigenetic mechanisms (Figure 1.1c).[12-14] Nucleosomes have been shown to condense into higher order structures *in vivo*, stabilized by inter-nucleosomal interactions and, in some contexts, through the help of additional proteins (e.g. linker histone H1 and Heterochromatin Protein 1 (HP1).[15-17] *In vitro* studies reveal that contiguous nucleosome arrays fold into regular local structures 30 nm in diameter (so-called 30 nm fibers).[18] However the intricacies of higher order chromatin folding *in vivo* are unclear, as this area remains the focus of ongoing studies.[19]

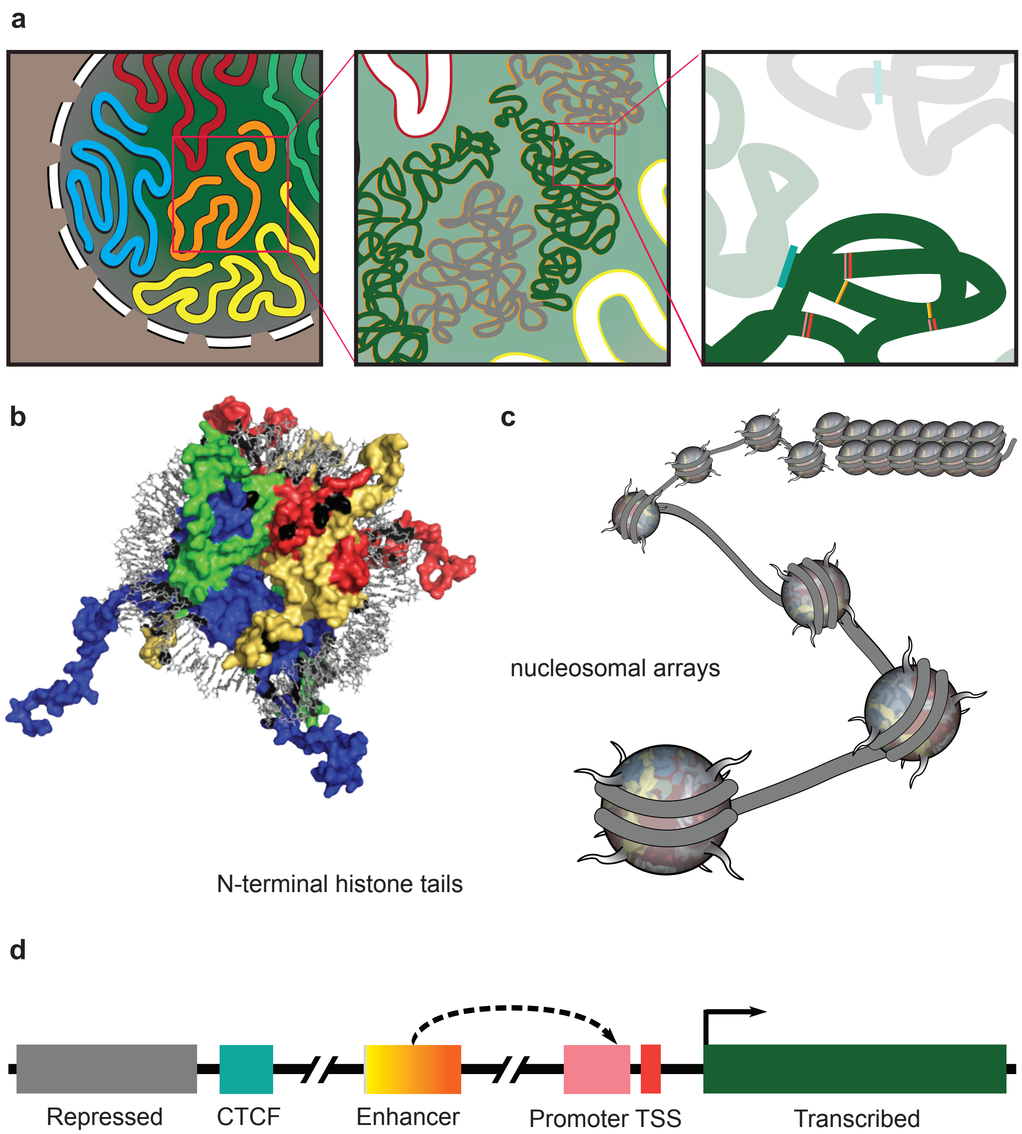


Figure 1.1. Chromatin structure and organization. (a) Chromatin is organized hierarchically where at a nuclear level chromosomes occupy distinct territories (left panel). On the megabase scale chromosomes divide into euchromatin (green line, middle panel) and heterochromatin (grey line, middle panel), which can be further divided into TADs (green line, right panel). TADs display discreet functional elements (functional elements correspond to Table 1.1 colors and elements identified in Figure 1.1d). (b) The structure of the nucleosome (PBD accession code: 1KX5). The N-terminal histone tails protrude out from the nucleosome core. H2A (yellow) H2B (red) H3 (blue) H4 (green) DNA (grey). (c) Nucleosomes show different levels of compaction and are dispersed along the genomic DNA (grey line). (d) The seven major classes of chromatin. Heterochromatin (repressed, gray box) is separated from euchromatin typically through CTCF enriched elements (turquoise box). Enhancers (orange/yellow box) typically interact with promotes (light red box) in trans, which lie adjacent to TSS (red box) and transcribed regions (green box).

Chromatin can be divided into euchromatin and heterochromatin based on regions of high and low gene transcription, respectively. Euchromatin is generally less compact than heterochromatin[15] and one way to subdivide euchromatin is into six categories that serve to explain many functional transcriptional elements of chromatin (Figure 1.1d, Table 1.1).[9] The transcribed (T) state of chromatin includes genomic regions that contain transcribed genes. The Transcriptional Start Site (TSS) state is proximally upstream of the T state where gene transcription is initiated. The Promoter Flanking (PF) state surrounds the TSS state and contains chromatin elements that aid in transcription, including Transcription Factor (TF) recognition motifs. Both Enhancer (E) and Weak Enhancer (WE) elements function in similar ways to the PF, however are distal to their corresponding TSS region. Lastly, the CCCTC binding factor (CTCF) state is enriched in insulator elements, which serve to sequester enhancer regions from non-cognate TSS regions. Conversely, heterochromatin or the Repressed (R) state is generally associated with transcriptionally silent regions of the genome, which is more nucleosome dense. Heterochromatic regions include inactive genes, telomeres, and tandem repeat DNA.[15] Heterochromatin can be further subdivided into functionally distinct types of chromatin that is either transcriptional inactive or transcriptionally repressed.[20]

Table 1.1. The seven major classes of chromatin. This table was reproduced from Consortium, and Bernistein et al 2012.[9] Chromatin classes were determined from integrative analysis of elements identified from the ENCODE project. The colors of the chromatin classes correspond to those used in Figure 1.1d.

Chromatin State	Description	Details	Color
R	Repressed and low activity regions	Enriched in heterochromatin marks and regions showing no or low signal	Grey
CTCF	CTCF-enriched insulator element	Sites of CTCF signal and lack of histone modifications	Turquoise
E	Predicted Enhancer	Enriched for enhancer associated marks	Orange
WE	Predicted Weak enhancer or *cis*-regulatory element	Similar to E state but weaker enrichments	Yellow
PF	Predicted Promoter flanking region	Regions that surround TSS	Light red
TSS	Predicted promoter region including TSS	Enriched in H3K4me3 and RNA Pol II	Red
T	Predicted transcribed region	Enriched for RNA Poll II and transcription associated marks	Green

At a 100 kilobase scale, euchromatin and heterochromatin organize into distinct topologically associated domains (TADs). [8] TADs are defined as regions of DNA that are separated by characteristic boundary elements and where DNA-DNA interactions *in-cis* are predominant. Since gene transcription is coordinated within TADs more so than between TADs –this is not to say gene transcription between TADs is not coordinated- it is posited that TADs contain primarily autonomous gene regulatory networks with relatively little inter-TAD associations.[8] TADs are postulated to be hardwired features of chromosomes and are invariant across cell-types. However, whether a TAD is euchromatic or heterochromatic is cell-type dependent.

On a megabase scale, groups of adjacent TADs organize into two distinct alternating chromatin states, known as open and closed chromatin.[7] Open and closed chromatin states correspond to groups of euchromatic TADs and heterochromatic TADs, respectively. These different chromatin compartments preferentially self-associate; that is, open chromatin preferentially associates with other open chromatin regions, whereas closed chromatin prefers to associate with other closed chromatin domains, creating functionally distinct compartments within the nucleus.[7]

1.2.1. *Chromatin Regulation*

Chromatin architecture does not offer up mechanistic principles for the organization and regulation of chromatin; it is merely descriptive. Fortunately,

much work has focused on the regulation of chromatin structure and function and how this relates to the many *trans*-acting factors that bind to and often modify the chromatin template. Based on this ever expanding body of knowledge, we now know that several biochemical inputs impinge on chromatin to regulate the activity of DNA-templated process such as transcription, principally: (i) DNA sequence-specific chromatin binders; (ii) non-coding RNA; (iii) histone variants; and modification of the (iv) DNA and (v) histones.

DNA sequence-dependent binding of proteins to specific chromatin elements is a firmly established mode of regulation. This is exemplified through the DNA sequence specific recognition of transcription factor proteins such as GCN4 and NF-Y (Figure 1.2a).[21, 22] Transcription factors activate gene transcription and represent ~8% of protein coding genes in the human genome. Other chromatin binders, such as the insulator CTCF, also recognize DNA in a sequence-specific manner as an integral part of their function.[23]

RNA has been shown to play both a sequence-dependent and sequence-independent role in chromatin regulation. Comprising about 60% of genomic coding regions in mammals, RNA has been shown to both recruit and compete away specific factors to/from chromatin.[24] Although the true extent of RNA regulation of chromatin is unknown, mechanisms in which RNA regulates chromatin include: (i) RNA interference; (ii) competition with endogenous RNAs; (iii) direct RNA binding of chromatin modifiers such as the PRC2 complex; and

(iv) RNA-DNA interactions both *in-cis* and *in-trans* as in the case of XIST in X-chromosome inactivation (Figure 1.2b).[25-27]

Nucleosome composition is additionally adjusted through the incorporation of histones variants into the nucleosome. In addition to the 4 canonical histones there are over 20 reported histone variants, which are specific to different chromatin regions and cell-types (Figure 1.2c).[28, 29] For example H2A.Z, a variant of H2A, localizes to gene promoters where it plays a specific role in nucleosomal stability and gene transcription.[28] Further, H2AZ.2.2, a spliced isoform of H2A.Z.2 is specifically enriched in the brain and serves to further destabilize nucleosomes.[28]

Covalent modification of chromatin has emerged as a central component of epigenetic regulation (Figure 1.2c). Both the DNA template and the histone spool around which the nucleic acid is packaged are subject to dynamic chemical modification, the former through methylation on the 5 position of the cytosine base (preferentially in CpG sequences) and latter through the introduction of manifold post-translational modifications (PTMs).[15] These DNA and protein modifications are faithfully propagated and maintained down cell lineages due to the action of dedicated enzymes for their installation and removal.[15, 30] Notably,

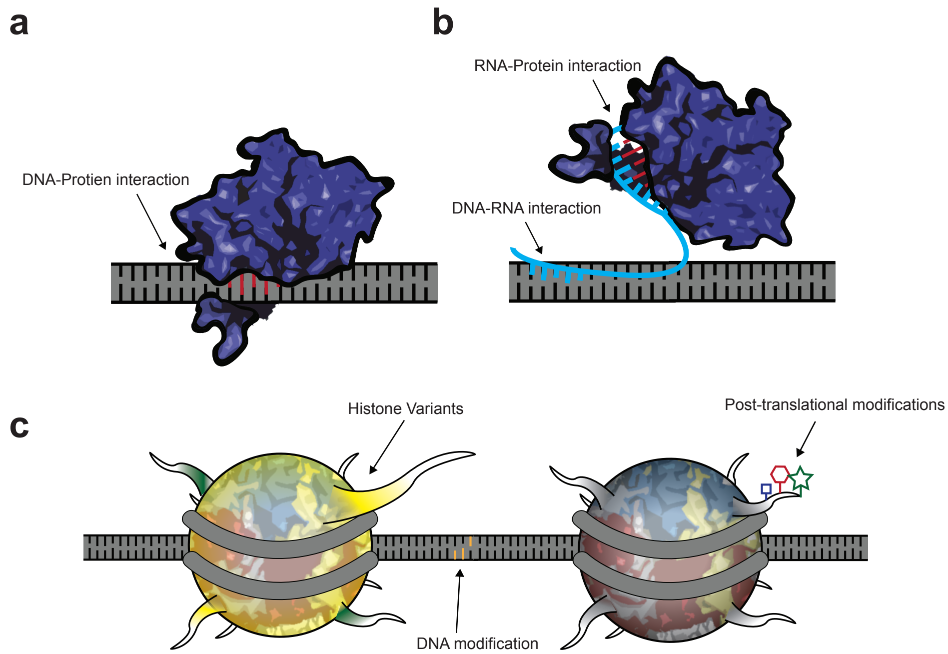


Figure 1.2. Mechanisms of chromatin regulation. (a) DNA sequence-dependent recognition of DNA by a transcription factor. (b) RNA recognition of DNA and tethering of RNA-binding protein. (c) Different 'types' of nucleosomes containing different histone variants (left nucleosome), and post-translational modifications (right nucleosome histone tail). Additionally, the cystosine bases of the DNA can be modified (orange DNA bases).

mis-regulation of these enzymes is frequently associated with disease in humans, for example cancer, underlining the importance of the associated chromatin modifications in the regulation of nuclear processes.[31] While DNA methylation is primarily associated with gene silencing and inactive chromatin generally, histone PTMs have been correlated with both euchromatin and heterochromatin states. Both DNA and histone modifications can function through the recruitment or repulsion of *trans*-acting chromatin binders – for

example, DNA methylation at CpG sequences can interfere with TF binding or can directly recruit transcriptional silencing factors through methyl-CpG binding domains.[15] As noted in the following sections, histone PTMs (which are the focus of this thesis research) provide an amazingly diverse set of 'ingredients' for the regulation of chromatin.

1.2.2. *Histone post-translation modifications*

Histone PTMs are site-specific covalent modifications that are attached to the amino acid side chains of the histone proteins. Modification sites number in the hundreds with many sites showing multiple types of PTMs.[15] Over 15 chemotypes of histone PTMs, or 'marks', have been annotated across the 4 canonical histones.[32, 33] PTMs vary in chemical complexity and composition and include: phosphorylation (ph), acetylation (ac), methylation (me), ubiquitylation (Ub), crotonylation, citrullination, SUMOylation, ADP ribosylation, and proline isomerization. Furthermore, a subset of these modifications can exist in multiple states (e.g. lysine can be mono-methylated (me1), di-methylated (me2) or tri-methylated (me3)). Histone PTMs are dynamic and are installed and removed by a subset of chromatin modifiers.

A recent mass spectrometry study identified 708 unique post-translationally modified histones indicative of a complex combinatorial nucleosomal PTM landscape.[34] Yet another layer of complexity is given by the possibility of asymmetric nucleosome modification *in vivo*, as each nucleosome contains two copies of each canonical histone that can contain very different PTM patterns.[35]

However, while nucleosomes contain multiple histone PTMs, their combinatorial co-occurrence is remarkably segregated as groups of histone PTMs co-occupy with one another.[36, 37] Broadly, histone PTMs can be grouped as either active, i.e. associated with euchromatin, or repressive, i.e. associated with heterochromatin, marks. It is hypothesized that the combinatorial nature of histone PTMs constitute a histone code, a code that is read out by chromatin binders and modifiers to bring about distinct downstream effects.[30]

1.2.3. *Functions of histone post-translational modifications*

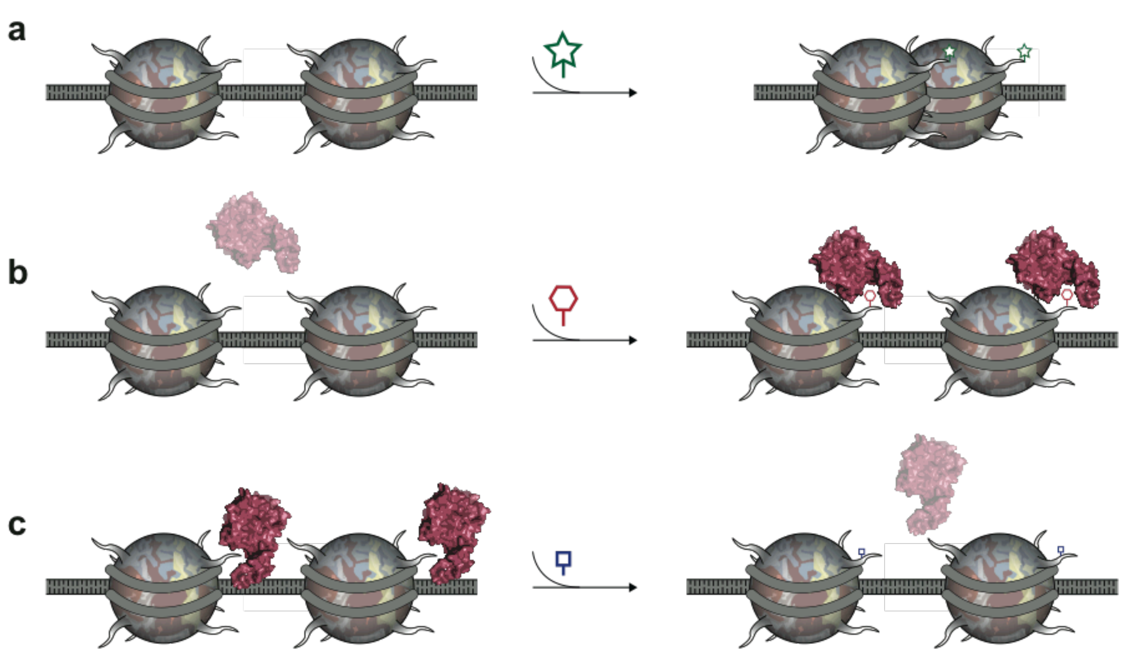


Figure 1.3. The regulation of chromatin by histone PTMs. (a) Histone PTMs modulate DNA/histone and histone/histone interactions altering chromatin structure. (b) Histone PTMs directly recruit chromatin binders (c) Histone PTMs directly block chromatin binders through disruption of protein/chromatin interaction surfaces.

Histone PTMs play a role in nucleosome dynamics including processes in which nucleosomes are translocated, unwrapped, evicted or replaced. Neutralization of the positive charge of lysine upon acetylation, and potentially less well-understood acylations such as crotonylation, suggests a direct physical mechanism for how this is achieved; namely through the disruption of DNA-histone and histone-histone interactions (Figure 1.3a). Similarly, phosphorylation, which installs a negative charge to an otherwise neutrally charged amino acid residue, has been described to interfere with nucleosome wrapping, making chromatin more accessible to DNase I *in vitro.*[38] In addition, many histone PTMs exist on the surface of the nucleosome proximal to histone-DNA contacts and are thought to function by altering intra-nucleosome dynamics.[39, 40] For example, the acetylation of H3K115 and the phosphorylation of H3K118 modifications are adjacent to the nucleosomal DNA and alter nucleosome dynamics.[38, 41, 42]

In addition to directly affecting chromatin compaction, PTMs can regulate chromatin structure and function through more subtle *trans*-acting mechanisms. The readout of histone modifications, such as acetyl lysine, methyl lysine, methyl arginine, and phospho serine, by a variety of specific protein modules, or histone PTM binders, has been well documented.[43, 44] For example, structural studies have demonstrated that the protein 53BP1 specifically recognizes H4K20me2 through its tandem Tudor domain, an interaction integral for recruitment of 53BP1 to sites of DNA damage.[45] Other marks affect nucleosome dynamics upon recruitment of histone PTM binders. H3K9me3 has been shown to recruit HP1,

thereby inducing heterochromatin formation (Figure 1.3b).[46] Additionally, H3K36me recruits chromatin remodelers, enzymes responsible for the translocation of nucleosomes along the DNA, as well as deacetylases highlighting that histone PTMs have the capacity to recruit both chromatin binders and modifiers. Conversely, histone PTMs can actively inhibit binding of certain chromatin modifiers and binders (Figure 1.3c). This has been shown for the Inhibitor of Growth Protein 2 (ING2) Plant Homeo Domain (PHD) finger, which recognizes H3K4me3 with a low micromolar affinity. However, in the presence of H3K4me3 and either H3R2me or H3T3ph a ~20-fold decrease in binding affinity is observed.[47] Accordingly, histone PTMs (or patterns of histone PTMs) have the ability to recruit both chromatin modifiers and binders, initiating further chromatin-templated processes.

1.2.4. *Chromatin Complexes*

Analogous to how nucleosomes contain multiple histone PTMs and histone variants, many chromatin-modifying complexes contain multiple histone PTM binding domains and are post-translationally modified. For example, the chromatin remodeler SWI/SNF complex, an 11-subunit complex comprising multi-histone PTM binding domains, is subjected to a host of post-translational modifications.[48, 49] The combinatorial nature of histone PTM binder subunits, and the dynamic modification of these subunits, in chromatin complexes may serve to recognize chromatin both intra and inter-nucleosomally through spatio-temporal multivalent associations.[50] This has profound implications for the regulation of

chromatin structure and function and thus understanding the mechanistic role that histone PTMs play in chromatin involves understanding both the mark as well as the context in which the mark is recognized and interpreted.

1.2.5. *Histone Ubiquitylation*

Ubiquitylation refers to the covalent attachment of the small (76 amino acids) globular protein Ub, typically to the ε-nitrogen of a lysine residue in a target protein substrate[51, 52]. Ub is a highly conserved protein (96% sequence homology between yeast and human) indicating the protein is under strong evolutionary pressure to retain its surface features. Ub is also a highly stable protein existing in micromolar concentrations inside the cell where it functions both as a signal for protein degradation, as well as in the regulation of protein-protein interactions, modulation of protein activity, and protein localization.[53] Proteolytic processes have been shown to mainly involve UbK48 poly-ubiquitylated substrates (all seven lysines in ubiquitin can themselves be ubiquitylated to generate polymers). Conversely, non-proteolytic processes are thought to occur through other poly-ubiquitin linkages and through mono-ubiquitylation. Ub has been shown to play a role in numerous biological processes and in the context of the nucleus these include gene regulation and DNA damage repair.[54]

Protein ubiquitylation is achieved via a three-enzyme cascade that ultimately results in the formation of an isopeptide linkage between the C-terminus of Ub and the ε-NH_2 of a lysine within the substrate. In this reaction cascade, the C-

terminus of Ub is activated in an ATP-dependent reaction by an Ub activating enzyme (E1), transferred to an ubiquitin conjugating enzyme (E2) that then associates with an Ub protein ligase (E3) before ligation to its substrate. In humans, 2 different non-specific E1 enzymes have been identified.[55] By contrast, the numbers of E2 and E3 ligases in humans are in the dozens and hundreds, respectively, and ultimately substrate specificity in the ubiquitylation system arises from the particular E2/E3 combination used.[56] Importantly, the ubiquitylation mark is reversible and can be removed from a given protein via dedicated de-ubiquitylating enzymes (DUBs) making ubiquitylation a dynamic reversible modification.[57]

Thirty-six ubiquitylation sites have been annotated in the 4 canonical human histones using proteomics methods.[58] Of these sites, 7 histone ubiquitylations have been validated and characterized further.[59-62] By far the two best-studied chromatin ubiquitylation marks occur on histone H2A at lysine 119 (H2AK119-Ub) and histone H2B at lysine 120 (herein referred to simply as H2B-Ub).[60] To the first approximation, H2AK119-Ub and H2B-Ub have opposing effects on chromatin regulation, at least as it pertains to the regulation of transcription.[63] H2AK119-Ub comprises about 5 to 15% of the total pool of H2A in higher eukaryotes and is, in fact, the most abundant histone ubiquitylation mark in higher eukaryotes, although the modification is surprisingly absent in yeast.[64] H2AK119-Ub is primarily associated with transcriptional silencing, but has also been linked to DNA damage repair.[54] By contrast, H2B-Ub, which is the major

focus of this thesis work, is linked to regions of active transcription, but as detailed in the following section, has been linked to many processes.

1.3. H2BK120 ubiquitylation

H2B-Ub has been known for over 40 years and is one of the most studied histone ubiquitylations.[60] The H2B-Ub modification is conserved from yeast to humans where it comprises ~10% and ~1%, of the total H2B in cells, respectively. [65, 66] On a molecular level, H2B-Ub is associated with euchromatin and found mainly towards the 5' end of gene bodies and within intron-exon boundaries of highly expressed skipped exons as well as chromatin boundaries in higher eukaryotes. [67-70] H2B-Ub has been implicated in almost every molecular process ascribed to chromatin with over 90 H2B-Ub interacting factors identified.[71, 72] Processes involving H2B-Ub include: (i) the regulation of transcription elongation,[73-75] (ii) RNA processing and DNA replication,[76-78] (iii) DNA damage response and repair,[79-82] (iv) nucleosome positioning,[83] (v) chromatin segregation,[84] and (vi) the maintenance of chromatin boundaries.[68, 70]

The ubiquitylation machinery responsible for the generation of H2B-Ub is well characterized and is conserved from yeast to humans.[85-89] Rad6A and Rad6B in humans (Rad6 in yeast) are the E2 conjugating enzymes for H2B-Ub with Rad6A being in much greater abundance than Rad6B (also of note, these E2 enzymes are 95% homologous).[86, 87] Interestingly, in yeast, efficient H2B ubiquitylation requires Rad6 S120 phosphorylation which highlights that H2B-Ub, itself, is

regulated through post-translational modifications.[85, 90] RNF20 and RNF40 form the H2B-Ub E3 complex (referred herein as hBre1), which has been suggested to exist as a tetramer containing RNF20/RNF20 and RNF40/RNF40 pairs (in yeast the complex is a homo-tetrameric Bre1 complex). [85, 86] In yeast, Rad6/Bre1 form a higher-order complex with the Large 1 protein, which is thought to regulate both the association of Rad6/Bre1 in gene bodies as well as Ubp8, a H2B-Ub specific DUB.[88, 91]

In yeast, both Ubp8 and Ubp10 have been reported to deubiquitylate H2B-Ub.[73, 92-94] These DUBs act on distinct pools of H2B-Ub, Ubp8 is responsible for targeting active chromatin near the TSS and Ubp10 within the gene body.[69] Ubp8 is a member of the Spt-Ada-Gcn5-Acetyl transferase (SAGA) complex, which is best known for its histone acetyltransferase activity encoded by the integral subunit Gcn5.[95, 96] The SAGA complex in humans has been shown to contain the Ubp8 homolog, USP22, which is thought to be the main H2B-Ub deubiquitylase in humans.[97] Less is known about Ubp10, however it’s thought to be associated with telomeric silencing.[69, 94, 98-100] In humans, exon specific DUBs have been reported and many more H2B-UB specific DUBs have been proposed.[101-104] Moreover, the requirement for efficient ubiquitylation/deubiquitylation of H2B-Ub for optimal transcription has led to the hypothesis that H2B-Ub is transient and cyclic in nature.[73, 105]

1.3.1. *H2B-Ub role in transcription elongation*

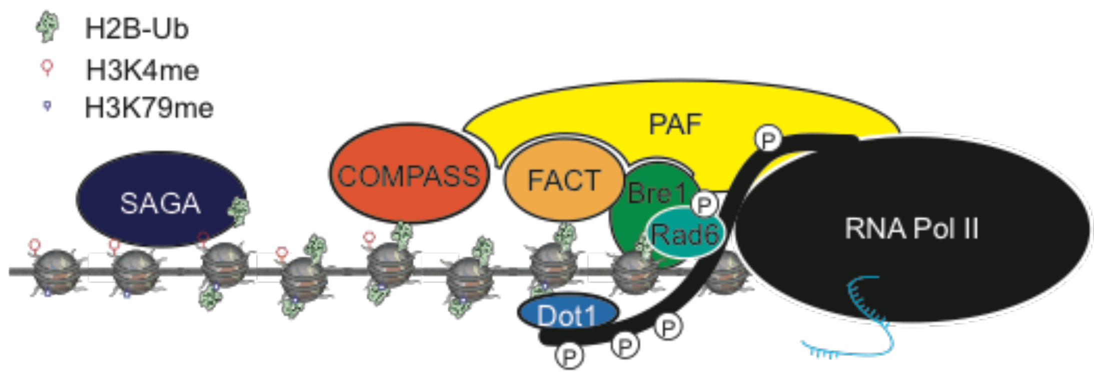


Figure 1.4. The role of H2B-Ub in transcriptional elongation. The H2B-Ub conjugation machinery localizes with RNA Pol II through association with the PAF complex serving to ubiquitylate H2B during transcription. Rad6S120ph is essential for efficient ubiquitylation (The letter P denotes a site of phosphoylation implicated in this pathway). The FACT complex recognizes H2B-Ub facilitating nucleosomal reassembly, potentially through its interactions with chromatin remodelers. COMPASS and Dot1 methylate H3K4 and H3K79 in the presence of H2B-Ub during transcription and are associated with the PAF complex and the S2/S5ph CTD of RNA Pol II, respectively. Further, the Ubp8 containing SAGA complex follows the RNA Pol II complex and removes H2B-Ub.

A growing body of work has focused on the role of H2B-Ub in transcription elongation. In transcription, RNA Pol II transcribes DNA to RNA in a multi-step fashion, a process that requires many components such as activators and co-activators for proper function. The role of H2B-Ub in transcription has been established by the observation that the ubiquitin conjugating enzymes specific for H2B-Ub, Bre1 and Rad6, co-localize with the elongating form of RNA Pol II (Figure 1.4). [85, 86] Partial depletion of RNF20 in humans or yBre1 in yeast

reduces H2B-Ub levels, thereby up- or downregulating different subsets of genes.[69, 106] In biochemically-reconstituted systems, localization of the H2B-Ub conjugation machinery involves recruitment of Bre1 through interactions with the subunit hPAF1 (or yPAF1 in yeast) of the RNA Polymerase II Associated Factor (PAF) complex, and the subsequent recruitment of Rad6 through interactions with Bre1 (Figure 1.4).[85] Further, it has been shown *in vitro* that transcription is required for efficient H2B ubiquitylation.[86] Since the presence of H2B-Ub did not alter the rate of transcription in this system, it was proposed that H2B-Ub might function in transcriptional steps not observable in such a well-defined but simplified system.[86]

H2B-Ub has been suggested to play a role in nucleosome reassembly in the wake of RNA Pol II passage through its functional association with the Facilitates Chromatin Transcription (FACT) complex. FACT is a heterodimer protein complex containing the subunits Spt16 and SSRP1. It is believed to enhance 'nucleosome breathing', thereby stabilizing the open configuration of nucleosomes and facilitating RNA Pol II passage.[107-109] H2B-Ub and Spt16 have been shown to functionally interact and affect chromatin structure through transcription-coupled effects in yeast.[110] This functional association is most likely through a physical interaction between the PAF complex and both Bre1 and the FACT complex, although recently a direct physical interaction between Bre1 and the FACT complex has been reported (Figure 1.4).[74, 111-113] Additionally, H2B-Ub

has been associated with other histone chaperones/remodelers such as CHD1, which has been shown to physically interact with FACT, and SWI/SNF.[71, 111, 114]

1.3.2. *H2B-Ub affects chromatin structure and stability*

It has been proposed that H2B-Ub plays a structural role to facilitate transcription. The identification of functional interactions between H2B-Ub and various histone chaperones and chromatin remodelers is indicative of this activity. The effect of H2B-Ub on chromatin structure and stability has been examined from both a direct and an indirect perspective, using well-defined biochemical systems and *in-vivo* experiments. It has been shown that H2B-Ub has a role in regulating nucleosome occupancy levels and nucleosomal positioning –which could be due to functional interactions between H2B-Ub with histone chaperones/remodelers- in genic regions during transcription.[83] The presence of H2B-Ub appears to correlate with more stable nucleosomes in cell extracts, as both H3 and H2B histones are more easily salt-extracted from yeast chromatin in the absence of H2B-Ub.[115] Presumably, experiments of this type reflect both direct and indirect effects as H2B-Ub co-occupies with many factors responsible for altering chromatin structure and stability.[68]

In well-defined reconstituted systems, H2B-Ub only modestly affects nucleosomal stability under equilibrium conditions.[116] However, H2B-Ub impedes Mg^{2+} mediated compaction of well-defined nucleosomal arrays indicating that the modification can intrinsically affect chromatin structure.[117] Further, this property

was attributed to specific effects of ubiquitin as this affect on chromatin compaction could not be reproduced with the structurally similar ubiquitin-like protein, Hub1.[117] It remains to be seen whether this structural property of H2B-Ub significantly contributes to transcriptional regulation *in vivo*.

1.3.3. *H2B-Ub directly controls H3K4 and H3K79 methylations*

Apart from intrinsic effects, H2B-Ub plays a role in transcription through the *trans*-histone cross-talk between H2B-Ub and H3 methylation. It has been shown that H2B-Ub is critical for both H3K4 methylation and H3K79 methylation, through the stimulation of the dedicated H3K4 and H3K79 methyltransferases, Su(var)3-9, Enhancer of zeste, and Trx 1 (Set1) and Disruptor of Telomeric Silencing 1 (Dot1), respectively.[86, 118, 119] *In vivo*, H2B-Ub has been shown to be dispensable for both monomethylation of these lysines (H3K4me1 and H3K79me1), but necessary for stimulating higher methylation states.[120] *In vitro* studies using chemically reconstituted mono- and polynucleosomes, demonstrated that H2B-Ub enhances the kinetics and efficacy of all methylation states of H3K4 and H3K79.[86, 121] This section will focus on what is known about the mechanistic details of the installation of these marks in relation to H2B-Ub.

1.3.3.1. H2B-Ub directly regulates H3K4 methylation

H3K4 methylation is implicated in transcriptional regulation with H3K4 tri-methylation being associated with the TSS of virtually all active genes.[122] It is postulated that H3K4me3 influences transcription primarily through the

recruitment of transcriptional co-activators (H3K4 recognition domains number in the 20s)[44]. [122, 123] Surprisingly, given the aforementioned *trans*-histone cross-talk, this localization is in stark contrast to H2B-Ub, which is primarily found within active gene bodies (Figure 1.5).[68, 124] H3K4me3 is known to correlate with transcriptional rates, co-occupy with RNA Pol II and histone acetylation; all indicative of gene transcription.[124-127] H3K4me1/2 patterns differ significantly between yeast and humans and, in both organisms, have been hypothesized to have non-overlapping functions. Indeed, these different methylation states often occupy discrete chromatin regions.[128] For example, in humans, enhancers are often marked with H3K4me1 whereas as already noted H3K4me3 localizes with the TSS of genes. [128]

In yeast, the dedicated H3K4 methyltransferase ySet1 contains the canonical Set catalytic domain observed in all known lysine methyltransferases, except for the H3K79 methyltransferase, Dot1. ySet1 exists in an 8 member protein complex, aptly named the Complex of Proteins Associated with Set1 (COMPASS) (hereafter referred to as ySet1C). This complex is required for proper H3K4 methylation activity, as ySet1 is inactive on its own. In humans, there are six known H3K4 methyltransferase complexes.[129] Two of these complexes have compositions similar to ySet1C, including polypeptides similar to ySet1 (hSet1A and hSet1B), and are known as the human COMPASS complexes (referred to herein collectively as hSet1C). The four other complexes that are known to methylate H3K4 contain the Set domain Mixed Linear Leukemia protein (MLL).

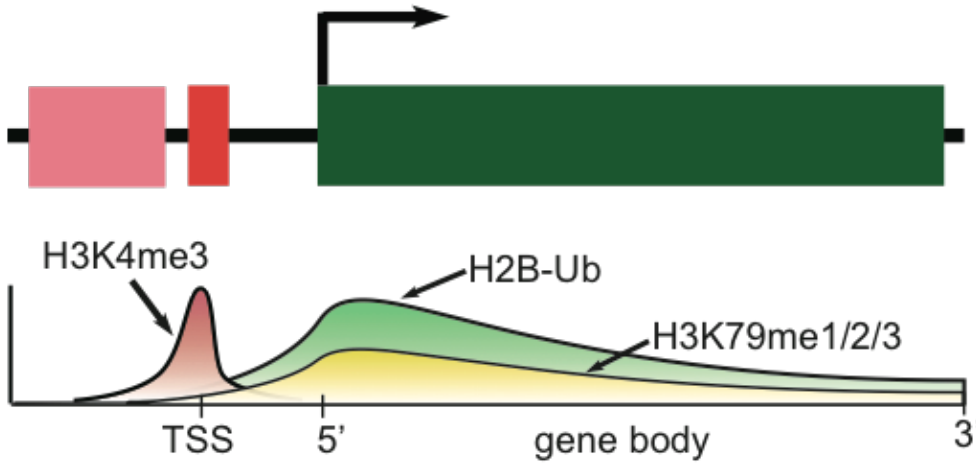


Figure 1.5. Localization of H2B-Ub and related histone PTMs. Gene schematic including promotor (light red), TSS (red), and transcribed (green) regions (top panel). Simplified H2B-Ub/H3K4me3/H3K79me1/2/3 occupancy patterns determined by ChIP seq (bottom panel).[130] H3K4me3 localizes to the TSS of genes whereas both H2B-Ub and H3K79me1/2/3 co-localize to the gene body and are bias to the 5' end (vertical axis represents relative abundance, horizontal axis represents linear DNA segment).

hSet1C is thought to be the primary H3K4me3 methyltransferase in mammalian cells as it is generally more active and less gene-specific than the other reported H3K4 methyltransferases, although all have been shown to have non-redundant roles.[131]

Both ySet1C and hSet1C (collectively referred to as Set1C) are associated with transcriptional elongation through interactions with the PAF complex.[132] It has been hypothesized that, analogous to Rad6/Bre1, ySet1C follows RNA Pol II during transcriptional elongation (Figure 1.4).[64] It has been postulated that Rad6

and Bre1 play a role in this *trans*-histone cross talk through the direct ubiquitylation of the ySet1C containing subunit, Cps35/Swd2.[133] However, *in vitro* studies with homogenously prepared H2B-Ub and recombinantly prepared Set1C have demonstrated a direct role for H2B-Ub in H3K4 methylation in the absence of RNA Pol II and Rad6/Bre1. [86] Nonetheless, the mechanistic details of how H2B-Ub directly stimulates Set1C-mediated H3K4 methylation, including how Set1C reads out H2B-Ub, remains inadequately understood.

1.3.3.2. H2B-Ub directly regulates H3K79 methylation

Analogous to Set1C, H2B-Ub can directly stimulate H3K79 methylation.[118] H3K79 methylation has been linked to transcription, cell cycle regulation, DNA damage/repair, and development, as well as the regulation of heterochromatin.[134-137] H3K79 has been found to exhibit all 3 methylation states (me1, me2, and me3) in yeast, although the existence of H3K79me3 in higher eukaryotes remains unclear. All H3K79 methylations correlate with actively transcribed regions of the genome, and in contrast to H3K4me3, co-occupy with H2B-Ub (~24% overlap between H3K79me3 and H2B-Ub in human embryonic carcinoma (NCCIT) cells) (Figure 1.5).[68] H3K79me2 has been shown to exert a minimal impact on overall nucleosomal structure. Instead, H3K79me2 alters the electrostatic environment in the vicinity of H3K79me2, by adopting a distinct rotameric conformation, and exposing a hydrophobic cavity near H3K79me2.[138] Three chromatin modifiers have been postulated to interact with H3K79me2

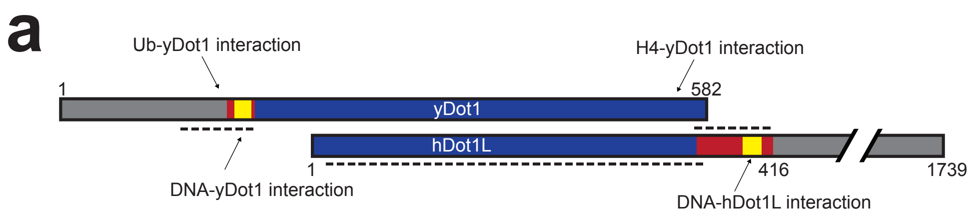
chromatin however whether these interactions occur *in vivo* (or are through direct physical interaction) is a matter of much debate.[45, 139-142]

H3K79 methylation is carried out solely by the methyltransferase Dot1 – note the human version of the enzyme is also referred to as Dot1-like (hDot1L). Knock-down or deletion of Dot1 in human cell lines and yeast, respectively, leads to a loss of H3K79 methylation.[130, 143] Dot1 is the only known histone lysine methyltransferase that contains a non-Set catalytic domain. Indeed, bioinformatic analysis indicates Dot1 is evolutionarily more similar to arginine methyltransferases.[144] The crystal structures of the yDot1 and hDot1L catalytic domains have been solved and show a high degree of structural conservation (Figure 1.6a and b).[145, 146] Interestingly, these crystal structures reveal a S-adenosyl methionine (SAM), the methyl-donating co-factor, binding pocket in the interior of the catalytic core and a lysine binding channel that contains the H3K79 methylation site. This is in contrast to Set domains, where SAM binds in a shallow pocket on the Set domain surface.[129]

The conserved nature of the catalytic domain in the yeast and human versions of the enzyme notwithstanding, the two enzymes diverge in terms of their overall architectures and interacting partners. The yeast version of Dot1 is 582 amino acids in length with the catalytic domain located near the C-terminus of the protein (Figure 1.6a). By contrast the human enzyme is much larger, containing 1739 amino acids, and has the catalytic domain located at the N-terminus. DNA

binding regions crucial to enzymatic activity have been identified in both yDot1 and hDot1L, and localize to the N-terminus of the polypeptide and the C-terminus of the polypeptide, respectively (Figure 1.6a). [145] The reason for this rearrangement is unclear. Additional sequence motifs have been identified in the yeast enzyme that are required for H3K79 methylation (Figure 1.6a), namely; an acidic patch near the C-terminus that physically interacts with a basic patch in histone H4 (particularly H4 R17 and R19),[121, 147] and a region spanning amino acids 101-140 that has been reported to physically interact with ubiquitin.[148] The human enzyme also requires H2B-Ub and the same basic patch in H4 for its stimulation.[121] However, the associated interaction motifs present in the yeast Dot1 are not conserved in hDot1L and so it remains unclear how the human enzyme engages ubiquitylated chromatin.

Both the yeast and human Dot1 enzymes have been shown to interact with various other nuclear factors, besides chromatin. For instance, Cps35, a member of the COMPASS complex that is also ubiquitylated, has been shown to functionally interact with yDot1 and could serve to localize yDot1 to genic regions through ySet1C and the PAF complex (Figure 1.4).[133] hDot1L has been shown to exist in protein complexes, albeit poorly characterized, that contain components critical in the Wnt development pathway, suggesting that in higher eukaryotes hDot1L plays a role in development.[137, 149-151] hDot1L physically interacts with the phosphorylated (S2ph and S5ph) C-terminal domain (CTD) of RNA Pol II, with a basic region corresponding to residues 618 to 627 of hDot1L being required for



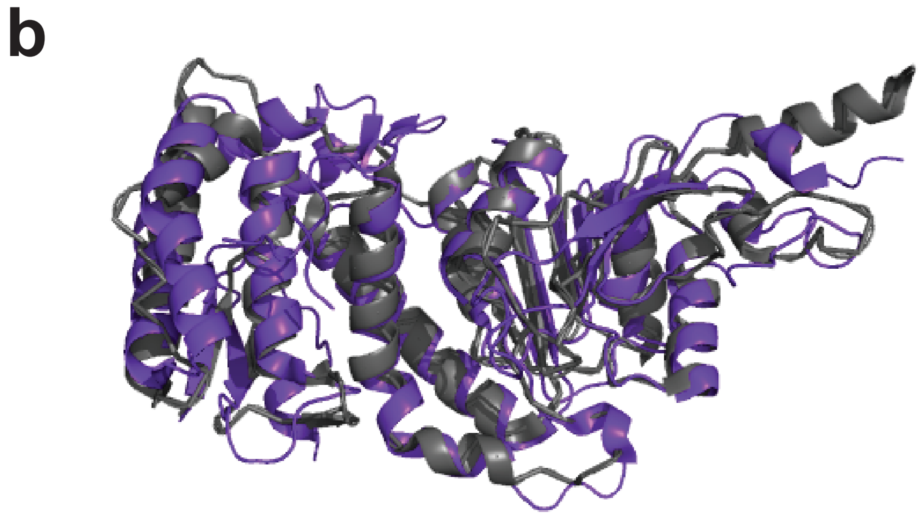


Figure 1.6. Comparison of yDot1 and hDot1L. (a) Schematic comparing the yDot1 and hDot1L domain architectures. The catalytic domain (blue) is conserved between the two methyltransferase whereas the DNA binding domain (yellow) and the unstructured region it localizes to (red) is different in relation to the catalytic domain. (b) Crystal structure comparison of the yDot1 (purple, PBD code: 1U2Z) and hDot1L (gray, PBD code: 1NW3) catalytic cores show a conserved tertiary structure.

this interaction.[130] Further, hDot1L co-localizes with RNA Pol II at TSSs, linking hDot1L and H3K79 methylation to transcription in humans. This led to the hypothesis, similar to the Csp35 example in yeast, that hDot1L follows RNA Pol II throughout transcription, methylating H2B-Ub nucleosomes along the way

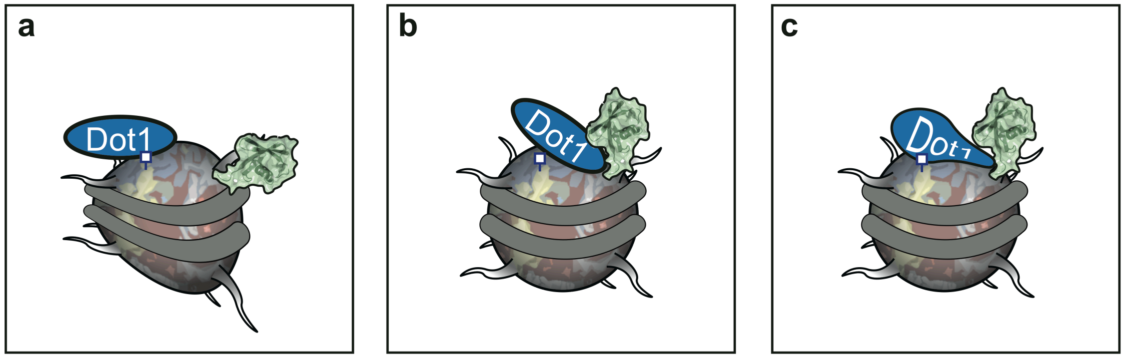


Figure 1.7. Potential mechanisms of intra-nucleosomal hDot1L regulation by H2B-Ub (a) In this scenario, Ub may impart an alternative nucleosome conformation that is required for H3K79 methylation. (b) hDot1L could additionally require H2B-Ub for correct postiioning on the nucleosomal surface. (c) H2B-Ub may allosterically regulate hDot1L as well.

(Figure 1.4). Notably, the same study demonstrated that yDot1 does not interact with the CTD of RNA Pol II.

Importantly, biochemical studies have shown that hDot1L is capable of directly methylating H3K79 in the context of H2B-Ub containing nucleosomes, chromatin arrays, and chromatinized plasmids.[118] Thus, the only strict requirement for H3K79 methylation by Dot1 is the presence of H2B-Ub – presumably the role of additional factors *in vivo* is to help regulate this activity in a spatial-temporal sense. Biochemical studies further reveal that the H2B-Ub/H3K79me crosstalk is intra-nucleosomal, non-cooperative i.e. one H3K79me to one H2B-Ub, and involves specific regions on the ubiquitin surface.[117, 118, 121, 152] However, despite extensive biochemical study in well-defined H2B-Ub systems (discussed further in Chapter 3), it is still unclear how hDot1L functionally recognizes H2B-Ub.

Three non-mutually exclusive mechanistic models for how this could be achieved are: (i) H2B-Ub may allosterically alter the nucleosome and indirectly regulate H3K79 methylation, (ii) H2B-Ub could potentially directly interact with hDot1L and be required for correct hDot1L positioning and/or binding of the nucleosome, and (iii) H2B-Ub could additionally allosterically modify hDot1L leading to H3K79 methylation (Figure 1.7).

1.3.4. *Role of H2B-Ub and hDot1L in cancer*

Cancer is increasingly being recognized as having an epigenetic component that contributes to tumor initiation and progression.[31] Whole genome sequencing of numerous cancer cell lines has revealed persistent somatic mutations within a host of epigenetics regulators, including those that modulate histone PTMs.[31, 153] Given that the biochemical crosstalk between H2B-Ub and H3 lysine methylation is integral to both transcription and DNA damage repair, it should not be surprisingly then that mis-regulation of H2B-Ub and/or H3K79/H3K4me is intimately involved in human cancers.[72] Aberrant levels of H2B-Ub have been shown to result from altered activities of Rad6/Bre1 and DUBs.[72] This in turn leads to the disruption of proper chromatin function and gene expression patterns –specifically in proto-oncogenes- and increases genome instability; all processes that can lead to cancer.

hDot1L has been implicated in the misregulation of chromatin resulting in mixed lineage leukemia.[154] Specifically, 10% of childhood malignancies have been

attributed to translocation events of the MLL1 gene leading to the generation of MLL fusion proteins.[129] Three of these fusion proteins, MLL-AF10, MLL-AF9, and MLL-ENL, interact with hDot1L and have been shown to exist in a functional complex with Dot1.[155] hDot1L recruitment by MLL-AF10 has been shown to result in abnormal H3K79 methylation and to be essential –however not sufficient— for leukemogenesis. Aberrant H3K79 methylation leads to the mis-regulation of HOX gene transcription, which is a hallmark of cancer. Importantly, inhibition of this enzyme using either knock-down or pharmacological methods dramatically slows disease progression in mouse leukemia models.[72]

1.4. Semi-synthetic strategies to make histone PTMs

Analysis of ubiquitylated chromatin in biochemically homogenous systems is critical to elucidate the role that histone ubiquitylation plays in chromatin biology. Although the purification of endogenous ubiquitylated histones is possible, it results in heterogeneous substrates (i.e. the isolated ubiquitylated histones will contain additional PTMs) obfuscating biochemical analysis.[60, 156] Enzymatic ubiquitylation is also feasible, and has been demonstrated for H2BK34, H2BK120, and H2AK119.[86, 157] However, this approach requires the isolation of the cognate E2/E3 enzymes and results in rather inefficient ubiquitylation. Due to the many shortcomings of the aforementioned methods, chemical methods for the incorporation of post-translationally modified histones have gained much traction in the last decade. The main advantage of chemical synthesis is

homogenous substrate preparation, thereby allowing for the precise biochemical analysis of both chromatin structure and function.

1.4.1. *Expressed protein ligation*

One invaluable method to generate histones containing specific PTMs is expressed protein ligation (EPL). This semi-synthesis strategy is based on native chemical ligation (NCL) but employs recombinant production of one or more of the polypeptides, making the synthesis of full-length proteins, such as histones, tractable. In this section, both NCL and EPL will be discussed and their applications to the preparation of chemically defined PTM-modified chromatin substrates, in particular ubiquitylation, will be explored.

Solid-phase peptide synthesis (SPPS) has been employed to study histone acetylation as early as the 1970's.[158] In an early landmark study, the specificity of deacetylase enzymes, which are responsible for removing histone acetylation marks, was investigated with chemically synthesized peptides containing acetyl lysine.[158] This peptide strategy has been applied to other histone PTMs where the PTM-containing amino acid derivative is chemically accessible, i.e. lysine acetylation, lysine methylation, serine/threonine/tyrosine phosphorylation etc.[159] Indeed, this approach has been taken to the point where combinatorial libraries of modified histone peptides have been generated and used to probe a number of biochemical questions such as the binding specificities of various chromatin modifiers.[160] Although histone-derived peptides have proven useful, they have

obvious shortcomings when the process in question requires a more physiologically relevant substrate. For example, biochemical crosstalk pathways, such as the H2B-Ub – H3 methylation examples discussed above, can only be probed using nucleosome substrates.[118] Thus, considerable efforts in recent years have focused on the generation of modified full-length histones and their subsequent reconstitution into physiologically relevant chromatin fragments.

Arguably the most powerful route to modified histones employs the related NCL and EPL methods. Here, the target protein is assembled from a series of fragments, some of which contain the PTM of interest. NCL relies on the condensation in water of a peptide bearing a C-terminal thioester with a second peptide containing an N-terminal 1,2-amino thiol moiety (i.e. cysteine), resulting in a native peptide bond between the two peptides (Figure 1.8).[161] If necessary this process can be conducted in an iterative fashion until the complete protein is assembled. When all the fragments are synthetic in origin the process is referred to as NCL (i.e. total synthesis), whereas when one or other of these building blocks is a recombinant polypeptide, the process is termed EPL (semi-synthesis).[162] Operationally, EPL is well suited to generating modified histones since many of the PTMs are located close to the N- or C-termini of the proteins. Thus, the target histone can be assembled from a suitable modified, typically short synthetic peptide and a longer recombinant fragment (Figure 1.8).[162] Histone ligation sites are typically chosen to be adjacent to an alanine residue so that upon desulfurization (there is only one cysteine throughout the canonical

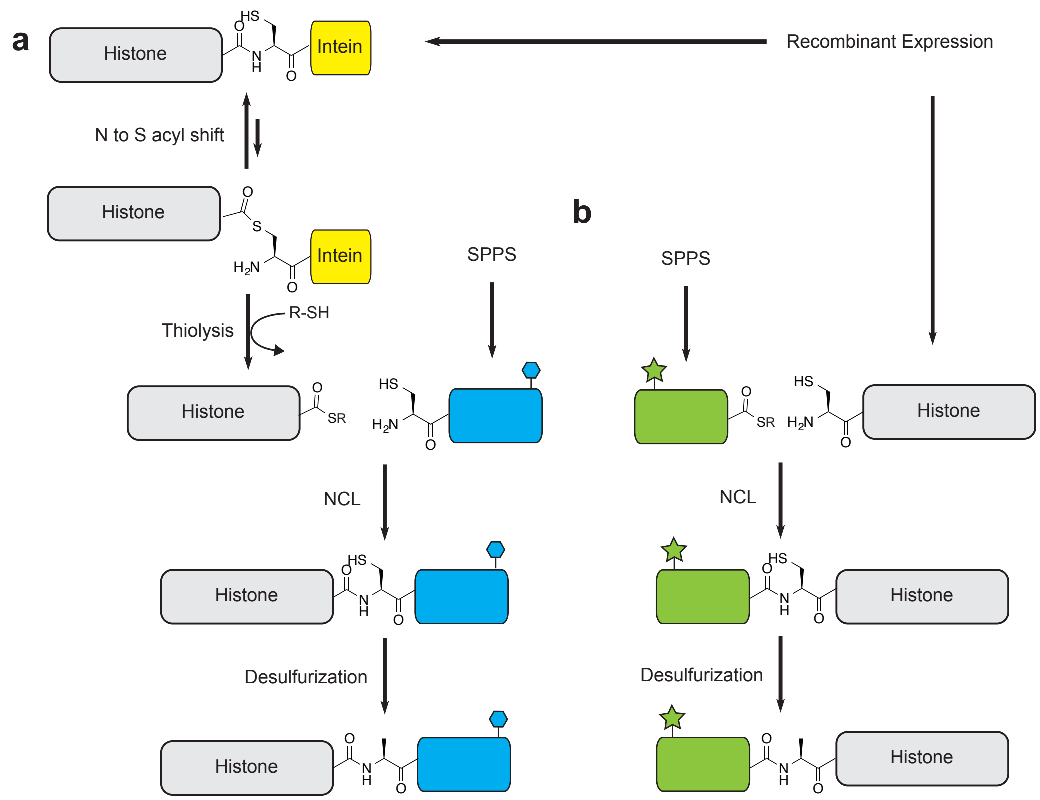


Figure 1.8. EPL applied to the semi-synthesis of histones. (a) The use of inteins affords a histone C-terminal thioester upon thiolysis with a small-molecule thiol (R-SH). Histone PTMs are incorporated into a N-terminal cysteine-containing peptide via SPPS. Ligation of the peptide containing the desired PTM (hexagon) and recombinant histone thioester via NCL produces the native PTM-containing histone upon desulfurization of the cysteine used for ligation resulting in an native alanine at the ligation site. (b) Conversely a N-terminal cysteine containing recombinant histone fragment can be produced by proteoltic processing and ligated with a PTM-containing (star) synthetic peptide thioester. Subsequent desulfurization results in a native N-terminally modified histone.

histones) of the initial EPL produced construct results in the native histone with the PTM (or PTMs) of interest. In other words the entire process is traceless. Additional methods have also been developed to make internal PTMs accessible by using multi-piece ligation strategies.[163] EPL has been applied to study histone PTMs and modified histones in many different contexts.[159]

A key feature of the EPL technology is the use of autoprocessing proteins known as inteins that are involved in a protein editing process known protein splicing.[162] A detailed understanding of the mechanism of intein-mediated protein splicing has allowed the rational design of engineered inteins that can be used to install a C-terminal thioester into essentially any recombinant protein. Operationally, an intein fusion protein is expressed in a suitable cell host and, following purification, is exposed to a small molecule thiol to afford the desired protein thioester derivative, which can then be utilized for semi-synthesis (Figure 1.8a). Furthermore, a subset of inteins are naturally split and conduct splicing *in trans* and have proven more efficient for the generation of protein thioesters.[164] Accordingly, many inteins have been characterized and utilized for the generation of proteins containing C-terminal thioesters.[165]

The first semi-synthesis of H2B-Ub employed the use of a three-piece EPL strategy (Figure 1.9a and b).[118, 166] In this strategy, the peptide corresponding to H2B(117-125) was synthesized with a photolytically removable ligation auxiliary at H2BK120. Two recombinant thioester fragments, an Ub(1-75)-thioester and an

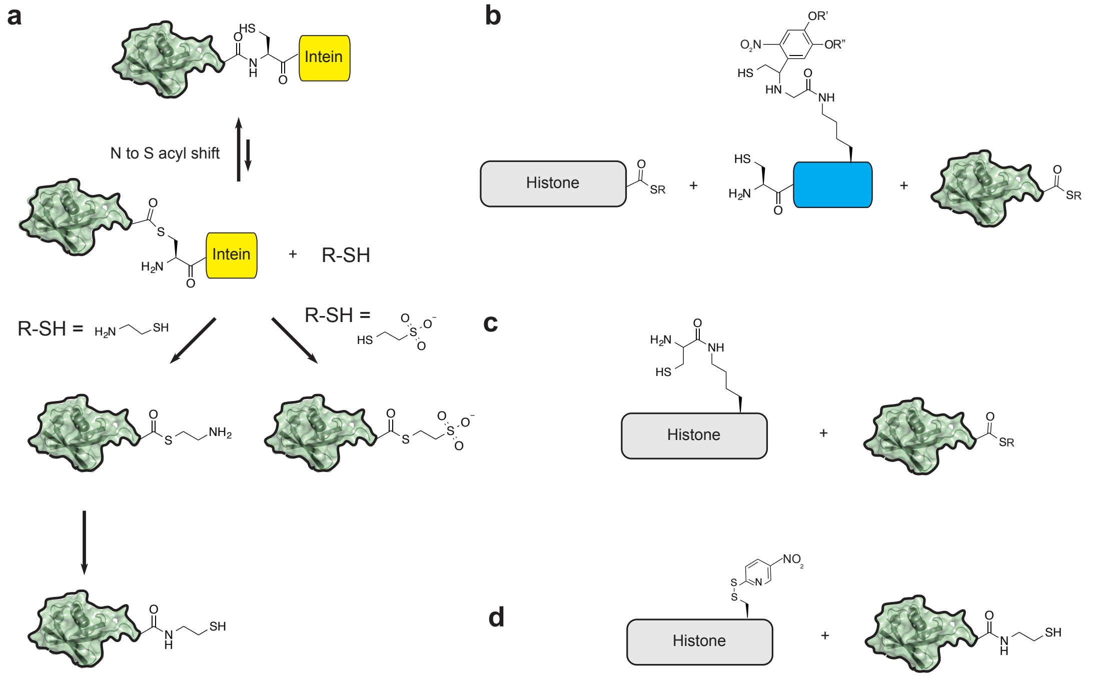


Figure 1.9. Semi-synthesis of ubiquitylated histones. (a) Schmatic of the preparation of ubiquitin. For EPL-based approaches Ub is prepared as a C-terminal thioester through the use of inteins followed by thiolysis with a small-molecule alkyl or aryl thiol (right panel). For disulfide conjugated ubiquitylation, aminoethanethiol is used for intien thiolysis (left panel). Aminoethanethiol subsequently undergoes a spontanous N to S acyl shift resulting in a C-terminal thiol. (b) Retrosynthetic analysis of the preparation of ubiquitylation histones using EPL. The larger histone fragement (grey) is prepared recombantly through the use of inteins and the peptide (blue) is synthesized by SPPS. R = $CH_2CH_2SO_3H$; R'= $CH_2CH_2CH_2C(O)NH_2CH_3$; R" = CH_3 (c) Retrosynthetic analysis of the preparation of ubiquitylation histones using amber codon supression. The full-length histone (grey) is prepared recombinantly through amber codon suppression. R = $CH_2CH_2SO_3H$ (d) Retrosynthetic analysis of the preparation of ubiquitylation histones using asymeteric disulfide chemistry. A histone containing a cysteine mutation (grey) is recombnantly prepared, and subsquently activated with a DTNP. Ubiquitin is prepared as in (a), (left panel).

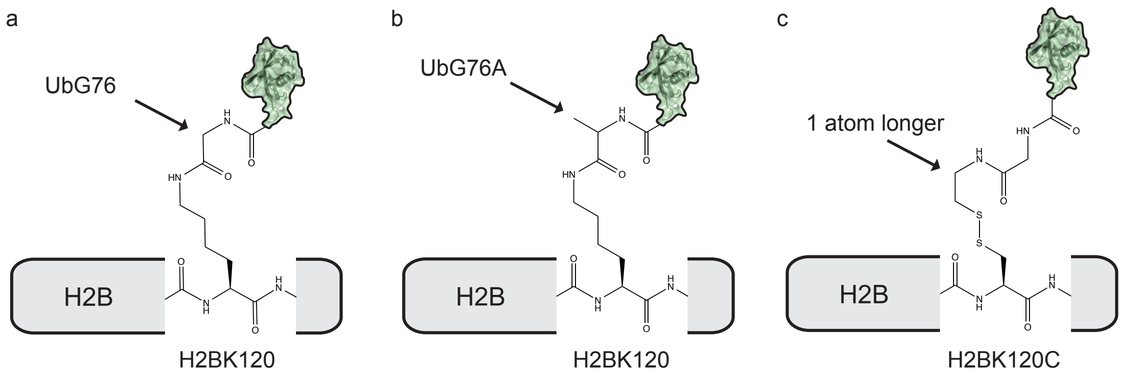


Figure 1.10. Comparison of H2B-Ub ligation junctions. (a) The native iso-peptide linkage between the ε-NH_2 lysine and ubiquitin produced from the photo-auxiliary EPL synthesis. (b) A UbG76A mutation results when using an EPL approach that greatly expedites the synthesis of H2B-Ub (or an amber codon suppresion approach). (c) The ligation junction between H2B and Ub results in a disulfide bond when Ub-SH is ligated to H2BK120C.

H2B(1-116)-thioester, were prepared through thiolysis of the corresponding intein fusions. The Ub(1-75) thioester fragment was first ligated to the K120 position within the H2B(117-125) peptide and following photolytic removal of the ligation auxiliary and a cysteine protecting group, this intermediate was subsequently ligated to the H2B(1-116) thioester. This yielded native H2B-Ub upon desulfurization (Figure 1.9b and Figure 1.10a). The second generation EPL approach expedited the ligation of ubiquitin to the H2B(117-125) peptide in exchange for an UbG76A mutation upon desulfurization (Figure 1.10b).[121] This greatly enhanced the yield whilst shortening the time required to prepare

homogeneous H2B-Ub. Variants of this semi-synthetic route have subsequently appeared in the literature, involving the use of different lysine analogs for the ubiquitin ligation step, as well as N-methylation of the isopeptide bond between ubiquitin and its substrate to prevent deubiquitylation *in vivo*.[167-169]

1.4.2. *Unnatural Amino Acid Mutagenesis*

Ribosomal methods for incorporating PTMs into histones site-specifically have also been reported.[170, 171] This approach relies on the expansion of the genetic code by the appropriation of the amber stop codon (UAG) to insert an unnatural amino acid, such as an amino acid containing a PTM, during mRNA translation. The type of amino acid inserted is achieved using ‘orthogonal’ aminoacyl-tRNA synthetase-tRNA pairs that naturally, or have been evolved to, recognize the unnatural amino acid. This so-called ‘amber-suppression’ method has been employed to incorporate PTMs such as phosphorylation, acetylation, and methylation into recombinant histones.[170, 172]

Although this method is not well suited for the direct incorporation of ubiquitin into a protein, strategies have been developed to incorporate amino acid analogs to facilitate EPL-based ubiquitylation through subsequent steps.[173, 174] Specifically, thia-lysine analogs, containing a 1,2 amino thiol side chain, that are recognized by ‘orthogonal’ aminoacyl-tRNA synthetase-tRNA pairs can be incorporated into proteins, and subsequently ubiquitylated (Figure 1.9c and Figure 1.10b). This method alleviates the need for a multi-piece ligation as the full-length amino thiol-

handle containing protein is recombinantly prepared. However, the yields of recombinant protein obtained by this amber-suppression method can be quite low compared to the semisynthetic method. Thus this method may be best suited for histone ubiquitylations that occur in the middle of the histone, and that are thus more difficult to obtain by semi-synthesis.

1.4.3. *Cysteine derivatization as a route to histone PTMs*

The derivatization of cysteine has been used to incorporate many PTM analogs into proteins including: glycosylation, prenylation, lysine methylation, ubiquitylation, and lysine acetylation.[159] The chemical alkylation of cysteine containing histones was first demonstrated in the semi-synthesis of H3 and H4 containing methyl lysine analogs.[175] These methyl-lysine analogs could be read out by site-specific methyl lysine antibodies, as demonstrated for H3K9me1/2/3, H3K4me3, H3K79me2 and H4K20me3. Additionally it was shown that the chromatin structural protein, HP1, bound to an H3K9me2 methyl lysine analog containing histone similar to the native H3K9me2. Lysine acetylation, such as on H4K16, has also been mimicked using cysteine chemistry.[176, 177] The H4K16ac analog was shown to be recognized by an anti-H4K16ac antibody and, when incorporated into nucleosomal arrays, displayed the expected chromatin decompaction behavior previously ascribed to the native modification.[176]

Cysteine derivatization for the preparation of ubiquitylated histone proteins has also been reported.[152, 178] This employs the use of disulfide chemistry and results in a non-native disulfide linkage between ubiquitin and its substrate. Specifically,

the linkage is 1 atom longer (~2.4 Å) when compared to the native iso-peptide linkage (Figure 1.10c). An intein-fusion strategy is used to generate a ubiquitin derivative containing a C-terminal thiol group, via thiolysis of the ubiquitin-intein construct with aminoethanethiol (Figure 1.9a).[152, 178] This Ub-thiol derivative was then attached to a unique cysteine within the target protein, e.g. H2B bearing a K120C mutation, via a nucleophilic disulfide exchange reaction (Figure 1.9d).[152 178] Importantly, although the disulfide approach generates a non-native linkage that is slightly longer than the normal isopeptide bond, the resulting $H2B_{ss}Ub$ conjugate was found to be indistinguishable from the native H2B-Ub in terms of hDot1L stimulation.[152] Thus, this simple route to ubiquitylated histones offers great promise in teasing apart the function of Ub in chromatin regulation. Indeed, this disulfide strategy has recently been used to install ubiquitin at H2AK15 allowing for the determination that the DNA damage protein 53PB1 recognizes H2BK15-Ub through the hydrophobic patch on ubiquitin (I44).[179]

1.5. Summary and conclusions

Gene regulation is a multifaceted process that has been fine-tuned so as to rely on the integration of multiple genetic and epigenetic factors. The identification of histone PTMs as fundamental to this regulation necessitates a thorough investigation of the mechanisms by which they function. H2B-Ub is one such modification that has been identified in many major chromatin associated processes and is biomedically relevant. Although semi-synthetic methods have allowed progress to be made in understanding the mechanistic details of H2B-

Ub-related processes, many questions remain unanswered, particularly with respect to how H2B-Ub is being recognized in these processes.

My thesis will focus on elucidation of the critical determinants governing Dot1 recognition of Ub and apply what is learned from this study to ySet1C regulation and chromatin compaction. In Chapter 2, we will first explore a more efficient Ub ligation methodology to expedite the study of the H2B-Ub modification. With this improved methodology we will then investigate the structure activity relationship between hDot1L and H2B-Ub in Chapter 3, with a more detailed analysis of our results in Chapter 4. In Chapter 5, we will attempt to complement these mutagenesis studies from a structural perspective both in a nucleosomal and nucleosomal array context. Further, in Chapter 6 we will extend our findings to both ySet1 and yDot1. This thesis is rooted in the belief that, similar to the histone code, the molecular mechanistic contributions of individual histone PTMs can be discerned in well-defined systems and integrated into the biological framework in which they function. In the discussion I will put the findings of my thesis into a more biological context, as well as look forward to what lies ahead in the fascinating world of histone PTMs.

Chapter 2. Site-specific chemical ubiquitylation of the mononucleosome via an asymmetric disulfide approach

2.1. Introduction

Histone ubiquitylation has been implicated in multiple different chromatin-templated processes.[64, 71, 72] Although proteome-wide approaches have been invaluable in the identification of histone ubiquitylation sites and their associated interactome, *in vitro* studies with homogenously prepared substrates are crucial for the further characterization of ubiquitylated protein–protein interactions.[58] This presents a significant challenge for synthetic chemists as a thorough investigation not only requires the synthesis of the ubiquitylated histone, but also ideally involves the synthesis of many ubiquitylated histone mutants in order to explore attendant structure-activity relationships (SAR).

Over the past six years, semi-synthetic strategies for the homogenous preparation of site-specific ubiquitylated histones, has enabled the study of these modified histones in well-defined systems.[118, 121, 152] Methodologies include the preparation of ubiquitylated histones containing a native isopeptide linkage as well as methodologies that introduce non-native linkages near the Ub ligation junction (see section 1.4 for a more detailed discussion) (Figure 1.10). The latter strategies circumvent the complexity of the first reported synthesis, as well as facilitate the preparation of ubiquitylated histones at internal sites. This has

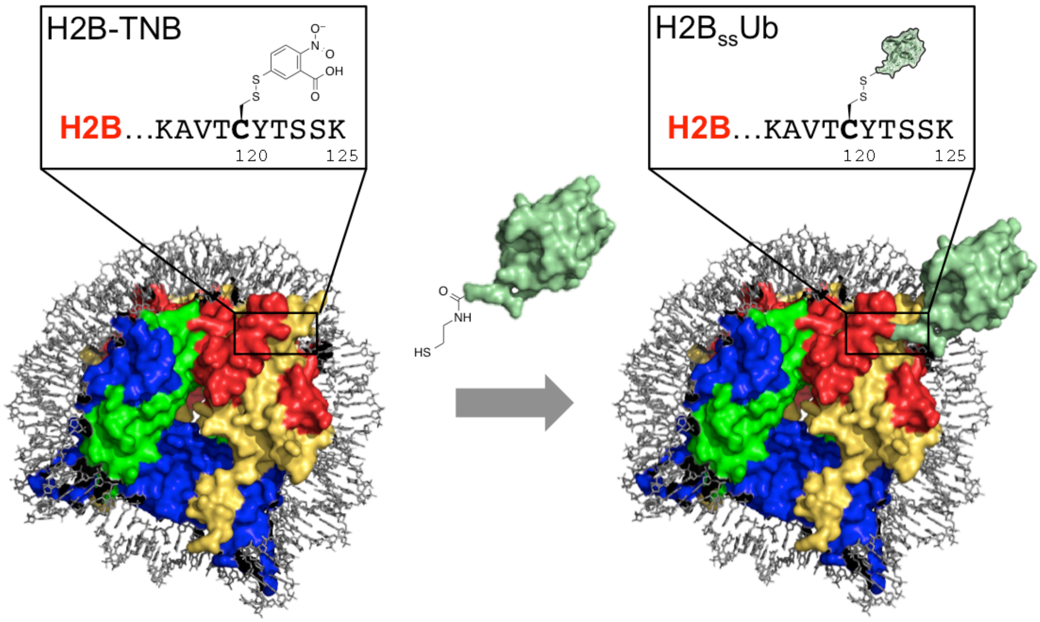


Figure 2.1. Preparation of $H2B_{ss}Ub$ MNs from act-MNs. Scheme of the Ub-SH act-MN ligation strategy. Site-specific ligation of Ub directly to the MN is achieved upon addition of Ub-SH (unmodified Ub shown in green, PDB 1UBQ3) to the act-MN (left panel) (wild-type MN shown, PDB 3LZ0). The H2B-Ub MN is modeled as a composite of Ub and MN structures (right panel). H2A (yellow), H2B (red), H3 (blue), H4 (green) are shown. Insets: close-up of the act-MN and $H2B_{ss}Ub$ modifications. Note as each act-MN contains 2 copies of H2B-TNB the act-MN can be di-ubiquitylated.

served to expedite the synthetic preparation of ubiquitylated histones as a trade off for non-native, but biochemically suitable, substrates.[121, 152] However, despite these improvements, all synthetic routes to date for histone ubiquitylation involve multistep protocols and are performed at the histone level, taking weeks to prepare a single ubiquitylated construct. Further, once synthesized, these ubiquitylated constructs must be incorporated into mononucleosomes (MNs) and/or chromatin substrates, which is both a time consuming and low yielding process.

We were motivated to develop a flexible strategy that would allow for the incorporation of Ub and Ub variants in a fast and facile way to enable SAR studies of the Ub surface in the context of H2BK120 ubiquitylation (H2B-Ub). In order to expedite the overall synthesis of H2B-Ub MNs and significantly reduce the amount of precious materials needed for reconstitution, we decided to test a strategy where H2B is ubiquitylated subsequent to the incorporation into MN substrates. We hypothesized that the site-specific ubiquitylation of H2B could be achieved within the context of the MN based on an asymmetric disulfide strategy that results in an Ub linkage to H2BK120C through a disulfide bond ($H2B_{ss}Ub$) (Figure 2.1, Figure 1.10c). In this strategy, we envisioned the preparation of an activated MN (act-MN) that contains a cysteine activated with 5-thio-2-nitro-benzoic acid (TNB) –in this case at H2BK120C (H2B-TNB). Upon addition of an Ub engineered with a C-terminal thiol (Ub-SH), this act-MN would undergo ubiquitylation to produce $H2B_{ss}Ub$ MNs (Figure 2.1).

2.2. The preparation of Ub-SH using ultra-fast splicing inteins

Our laboratory previously reported a strategy for the preparation of Ub-SH for use in the site-specific ligation to histones activated with DTNP (see section 1.4.3 for more details).[152] In this strategy, Ub was cloned into the commercially available pTBX1 vector that contained the GyrA intein followed by a Chitin Binding domain (CBD). Following affinity purification on chitin beads, thiolysis with aminoethanethiol afforded the desired Ub-SH protein. Similar to the original

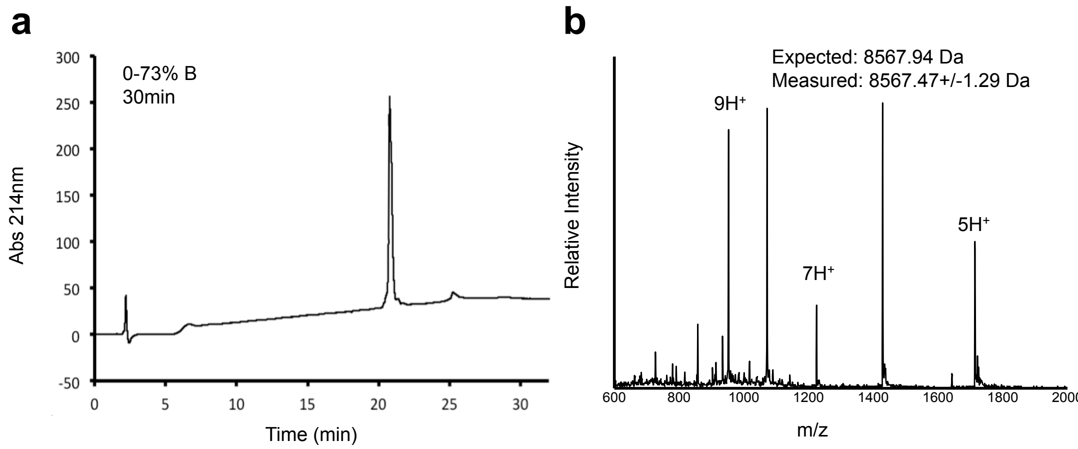


Figure 2.2. Characterization of Ub-SH prepared via an Ub-NPU-His construct (a) C18 analytical RP-HPLC chromatogram of purified Ub-SH. (b) ESI-MS of purified Ub-SH.

synthesis of H2B-Ub, this method was both low-yielding (2mg/L expression) and the preparation was on the time scale of weeks due to inefficient thiolysis and a time consuming purification procedure. Thus, we sought to take advantage of the newly characterized ultra-fast splicing inteins as well as the use of a 6xHis tag for greatly expedited protein affinity purification.[165]

A plasmid containing residues 1-75 of Ub (i.e. lacking the final glycine of the native protein) followed by the ultrafast fused-Npu intein and a 6xHis tag (Ub-Npu-His) was generated and expression of this fusion was optimized in *E. coli* BL21 cells. Additional optimization of both the affinity purification and subsequent thiolysis (with aminoethanethiol) steps was carried out, ultimately yielding Ub-SH in substantially improved amounts relative to the prior procedure (20 mg versus 2 mg of purified protein per liter expression). Moreover, the entire procedure was

shortened to a few days from weeks due to these improvements in the purification and thiolysis steps. For example, the thiolysis step using the fused-Npu intein was complete in 5 hours compared to incomplete thiolysis after 4 days with the GyrA intein. Using this improved procedure, Ub-SH was expressed, purified and lyophilized prior to use. The purity of Ub-SH was confirmed by Reversed Phase High-Performance Liquid Chromatography (RP-HPLC) and Electrospray Ionization Mass Spectrometry (ESI-MS) (Figure 2.2).

2.3. Preparation of an activated H2BK120C MN

We were next interested to see if we could prepare a MN containing the single lysine to cysteine mutant (H2BK120C) activated with TNB to yield an act-MN (Figure 2.3). Cysteine activation with 5,5'-dithiobis-(2-nitrobenzoic acid) (Di-TNB) has been reported, under non-denaturing conditions, which is why we pursued this strategy as opposed to what was previously employed to activate H2BK120C at the histone level.[178] Further, we designed our protocol to include the activation step after histone octamer formation. The rationale behind this was to allow the H2BK120C-containing octamer to be prepared under reducing conditions to prevent the formation of H2B disulfide-linked dimers during the octamer reconstitution process. Moreover, this strategy would allow unreacted Di-TNB (or reacted TNB) to be removed concurrent upon dialysis of the modified octamers with DNA to form act-MNs, thus streamlining the process.

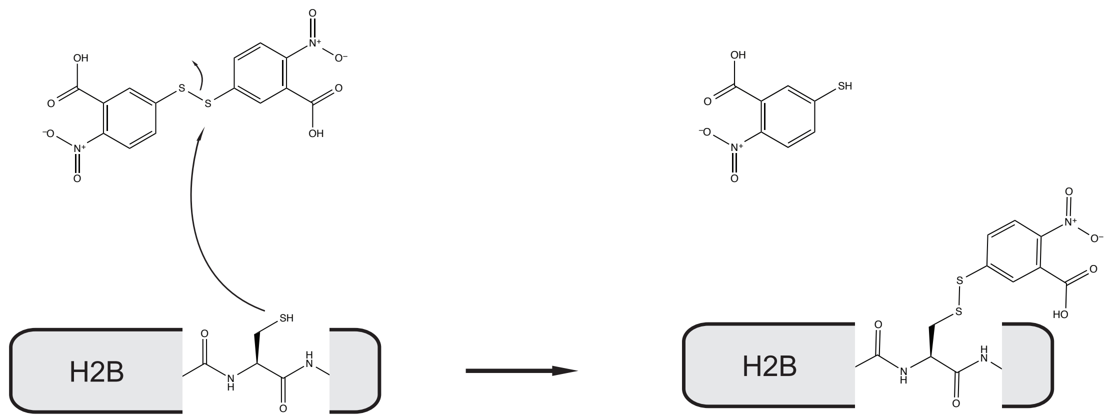


Figure 2.3. The activiation of H2BK120C with Di-TNB. H2BK120C is activated in the presence of Di-TNB through a thiol exchange of cysteine with TNB (left panel). This results in H2B-TNB and free TNB which has a strong absorbance at 412 nm (left panel).

The desired H2BK120C containing histone octamer was reconstituted under reducing conditions, following dialysis of the purified recombinant core histones from a 7M Guanidine HCl containing buffer into a high salt buffer.[180] Importantly, in order to achieve site specificity within the MN, H2BK120C was the only cysteine in any of the four histones employed in this study. As the only naturally occurring cysteine found within the four canonical core histones is H3C110, we used a H3C110S histone point mutant. H3C110 has been mutated previously in $H2B_{ss}Ub$ MNs without any adverse effects.[152] More specifically, this amino acid residue is found in the globular octamer histone core and both a H3C110S and H3C110A mutation have been shown to have no impact on MN function.[152, 179]

Size exclusion chromatography was employed for purification of the octamer and was performed under non-reducing conditions. This ensured removal of the reducing agents used in the octamer assembly. The H2BK120C octamer was then concentrated and a Di-TNB containing buffer was added 1:1 to activate the octamer. Note, this protocol ensured that Di-TNB was present in large molar excess (10 mM) compared to the octamer (10-100 μM). Reaction progression could be visualized photometrically as Di-TNB reacts with free thiol to yield TNB, which has a strong absorbance at 412 nm (Figure 2.3). By this measure, the disulfide exchange reaction to yield the H2B-TNB containing histone octamer was deemed to be complete after 10 minutes. Act-MNs were then formed via the standard semi-gradual dialysis approach, where the '601' strong positioning 153 bp DNA sequence (601 DNA) was mixed with the H2B-TNB containing histone octamer and dialyzed from a high salt buffer into a low salt buffer. Act-MN formation was analyzed via native polyacrylamide gel electrophoresis (PAGE) followed by ethidium bromide staining to ensure properly formed MNs (vida infra).

2.4. Ligation of Ub-SH to the act- MN

We then proceeded to test whether the act-MNs were accessible to ligation with Ub-SH. The ligation reactions were performed at 55^{o}C in order to enhance the kinetics of the conjugation. This temperature (for up to 3 hours) is commonly applied to MNs to correctly shift MN DNA to a homogenous position with respect to the octamer core.[180] Further, to allow for small volumes, parallel sample

preparation (vide infra), and homogenous ligation conditions, we decided to optimize our reaction using PCR tubes in a PCR machine using a 20-fold excess of Ub-SH.

Addition of Ub-SH to the act-MNs resulted in $H2B_{ss}Ub$ MNs as observed via gel shift by sodium dodecyl sulfate (SDS) PAGE followed by Coomassie staining and confirmed by mass spectrometry (Figure 2.4 a and b, respectively). A time course of this ligation was monitored via native PAGE followed by ethidium bromide staining (Figure 2.5a). A mix of non-modified (non), mono-ubiquitylated (mono) and di-ubiquitylated (di) containing MNs were observed after one hour of ligation as each MN contains two copies of each histone (Figure 2.5a). These will

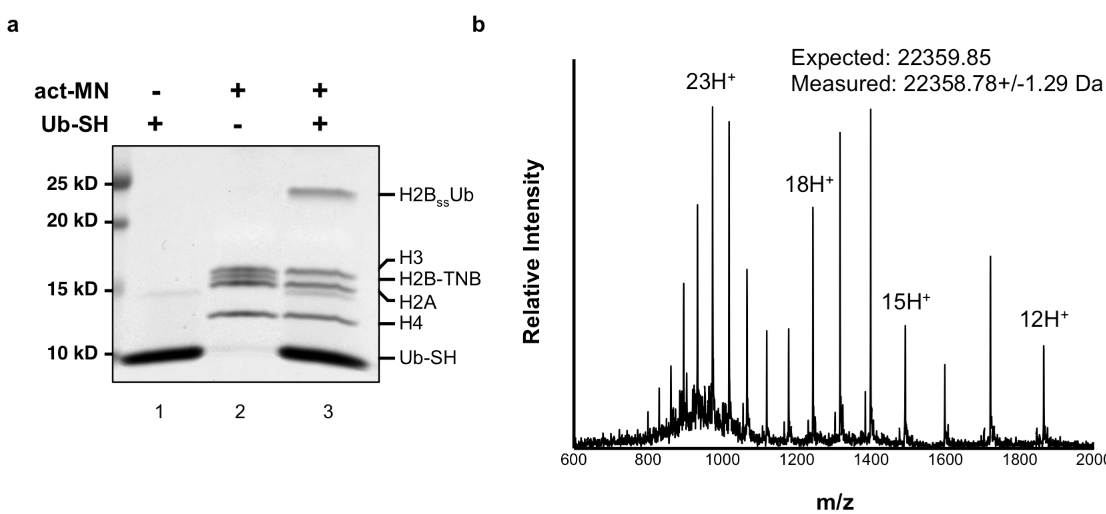


Figure 2.4. SDS-PAGE analysis and mass spectrometry of $H2B_{ss}Ub$ after addition of Ub-SH to act-MNs. (**a**) SDS-PAGE gel of act-MNs before and after addition of Ub-SH visualized by coomassie staining. The appearance of a band that runs at the molecular weight of $H2B_{ss}Ub$ is present when Ub-SH is added to act-MNs (lane 3). (**b**) ESI-MS of purified $H2B_{ss}Ub$ after ligation to act- MNs.

be collectively referred to as Ub-SH act-MNs to indicate both the ~20-fold excess of Ub-SH present *in trans* and the heterogeneity of Ub modification (this is will be critical to differentiate $H2B_{ss}Ub$ prepared via on-nucleosomal ligations (Ub-SH act-MNs) to that of homogenously reconstituted $H2B_{ss}Ub$ MNs).

Native PAGE gel bands (either non, mono, or di-Ub species) stained with ethidium bromide were quantified. Ub-SH act-MN gel bands were normalized to the total amount of Ub-SH act-MN (at the time of 60 min) to determine the absolute amount of each ubiquitylated species. This resulted in the calculation of the abundance of each type of MN (A_x, where x = non, mono, di) to the total amount of Ub-SH act-MN after 60 min:

$$A_{non} = non/total_{60}\ ,\ A_{mono} = mono/total_{60}\ ,\ A_{di} = di/total_{60}$$

Where non, mono, di = quantification of the respective gel band;

$total_{60}$ = non + mono + di at the time = 60 Ub-SH time point

Note that the A_x value is normalized to the absolute abundance relative to Ub-SH act-MNs at T = 60 min and is thus a good way to compare ligations between samples. Further, the total abundance of $H2B_{ss}Ub$ MNs can be calculated by

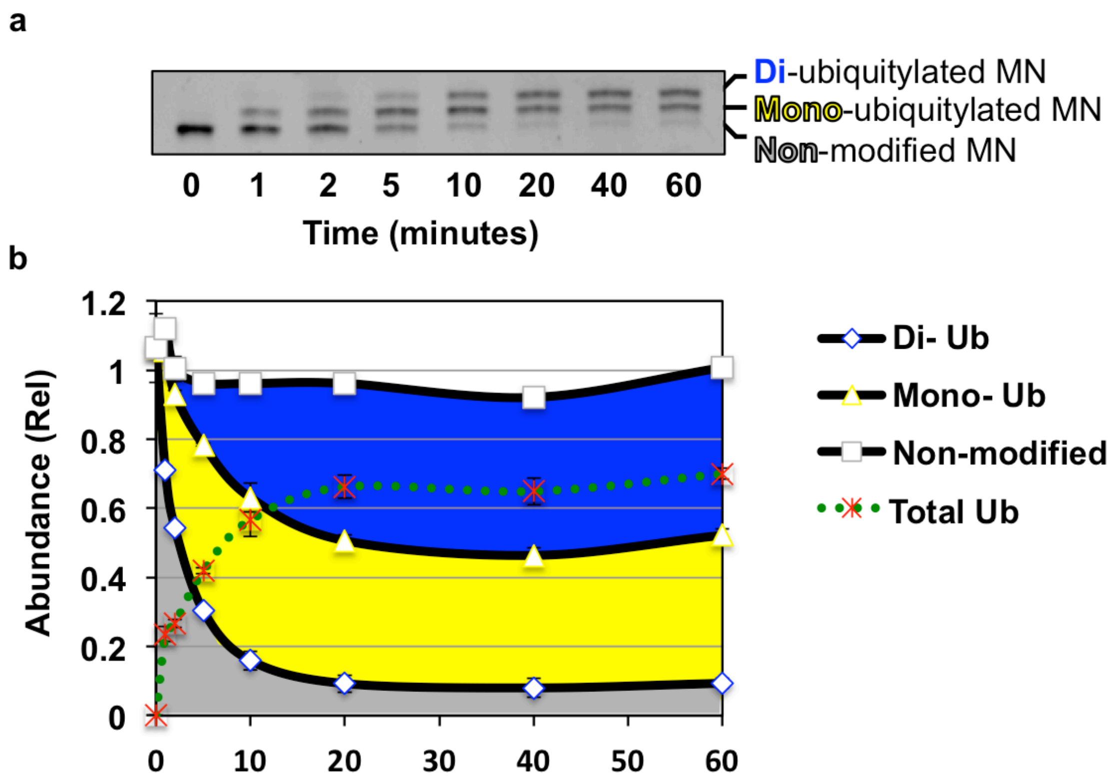


Figure 2.5. Time course analysis of the Ub-SH act-MN conjugation reaction. (a) Ub-SH act-MN ligations were monitored via native PAGE followed by ethidium bromide staining over the course of 1 hour. (b) The progress of the Ub-SH act-MN ligations was monitored by quantification of the gel bands in panel a using image J software. A stacked quantification of all three species (di, mono, non) is shown, relative to total MN abundance at t = 60 minutes (blue, yellow, and gray). The total $H2B_{ss}Ub$ abundance (green dotted line) is calculated to take into account that the mono-ubiquitylated MNs contains one copy of H2BK120C and one copy of $H2B_{ss}Ub$. The reaction progressed to about 70% completion (total $H2B_{ss}Ub$) over the course of 1 hour resulting in a mix of unmodified, mono-, and di-ubiquitylated MN species. Error bars, s.e.m (n = 3).

taking into account that the mono species contained one copy of H2BK120C and one copy of $H2B_{ss}Ub$ per MN:

$$\text{Total } H2B_{ss}Ub = A_{di} + (0.5 \times A_{mono})$$

Quantification of the time course revealed that the reaction went to 70% completion (total $H2B_{ss}Ub$) resulting in approximately an equimolar ratio of di-ubiquitylated to mono-ubiquitylated MNs (A_{mono} and A_{di} values), as each MN contains two copies of H2B (Figure 2.5b). Under the reaction conditions, the reaction appeared to not progress any further after 1 hour and thus all further ligations –and quantifications- were performed at 1 hour.

2.5. Note about terminology used throughout this thesis

In an effort to simplify nomenclature, while remaining informative, we have adopted terminology to account for both the type of H2B-Ub ligation as well as whether the ligations were performed on the histone or MN level. The semi-synthetic method utilized for ubiquitylation will be annotated by the use of H2B-Ub and $H2B_{ss}Ub$, which respectively indicates ubiquitylation via an EPL approach (that results in an UbG76A mutation), or an asymmetric disulfide approach (Figure 1.10b and c, respectively). The MN denotation will be directly preceded by either modified histones, e.g. H2B-Ub MN, or when necessary, to prevent ambiguity, unmodified histones, e.g. H2B MN to delineate what type of MN is being discussed and to indicate the homogeneity of the MN. Histone and MNs

will further be annotated to include species i.e. *Xenopus laevis, Drosophila melanogaster, Saccharomyces cerevisiae*, when appropriate e.g. *x*H2B, *d*H2B, *y*H2B respectively. Lastly, in the case of on-nucleosomal ubiquitylations, which results in heterogeneous MNs, Ub (or Ub variants) will be included for clarification and the MNs will be referred to as act-MNs, e.g. Ub-SH act-MNs. This will serve to further underscore that these MNs mixtures also contain a ~20-fold excess of Ub-SH. All non-annotated substrates will assume the H2BK120 (or K120C) site of ubiquitylation and *x*MNs.

2.6. hDot1L activity is stimulated by Ub-SH act-MNs

Ultimately, we were interested in utilizing act-MNs to facilitate a SAR study of the Ub surface in regard to hDot1L activity. Consequently, it was critical to show that Ub-SH act-MNs were able to stimulate hDot1L activity similarly to those generated by previous reconstitution methods after histone ubiquitylation, i.e. $H2B_{ss}Ub$ MNs. Previous studies demonstrated that the presence of Ub *in trans* neither upregulated hDot1L activity towards H2B MNs or titrated hDot1L activity away from H2B-Ub MNs.[118] Further, an analysis of MNs with different intra-nucleosomal ratios of H2B-Ub to H2B revealed a lack of cooperatively within the MN for H2B-Ub stimulation of hDot1L-mediated H3K79 methylation (discussed in more detail in section 1.3.3.2).[121] Accordingly, we reasoned hDot1L methyltransferase assays could be performed on Ub-SH act-MNs without further purification as only the Ub *in cis*, i.e. $H2B_{ss}Ub$, would influence hDot1L activity.

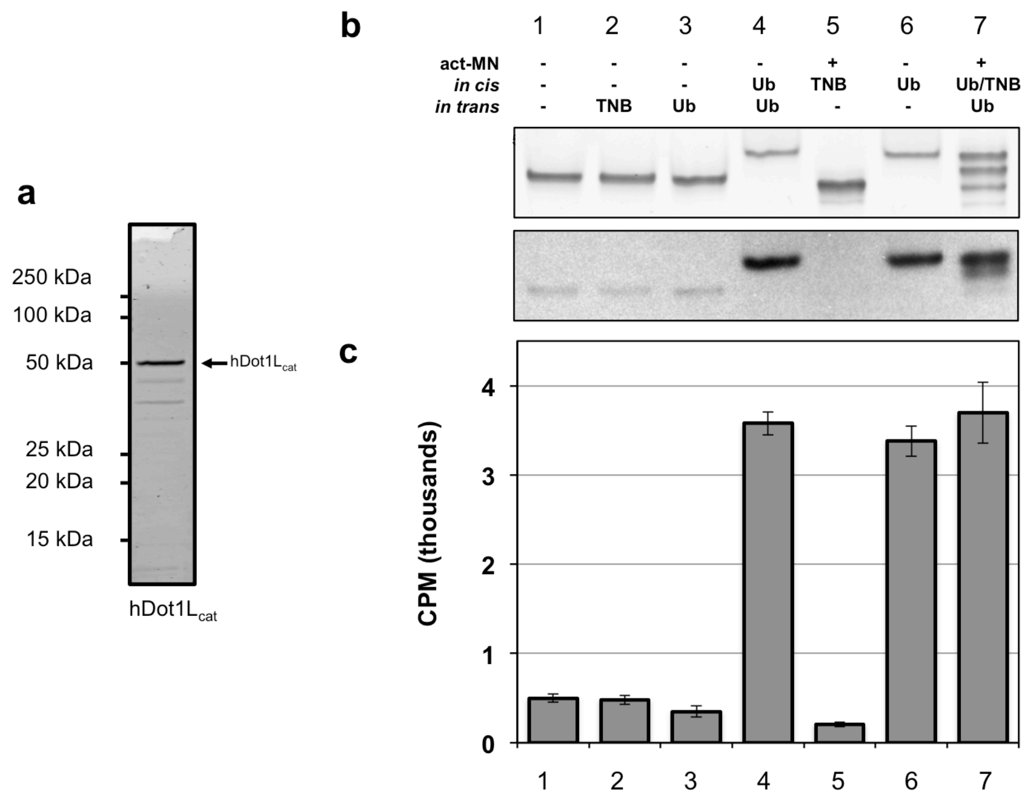


Figure 2.6. hDot1L methyltransferase assays on both act-MNs and MNs with Ub *in cis* and *in trans*. (a) Coomassie stained SDS PAGE gel showing hDot1L(1-416) purified from BL21 cells. (b) ^{3}H-SAM methyltransferase assays were performed on H2B MNs (lanes 1-3), H2B$_{ss}$Ub MNs (lanes 4 and 6), and act-MNs (lanes 5 and 7) in the presence or absence of TNB or Ub-SH *in trans* (as indicated in panel b). MNs were visualized by native-PAGE followed by Sybr Gold staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (bottom panel). (c) Quantification of ^{3}H-methyl incorperation was performed by filter binding assays followed by liquid scintillation counting. ^{3}H-methylation was observed in samples where Ub was present in the reaction *in cis* (Panels (b) and (c), lanes 4, 6, 7) and no additional stimulation was observed when Ub was added *in trans* (middle and bottom panels lane 4 versus lane 6). Error bars, s.e.m. (n = 3).

We chose to test the activity of hDot1L towards Ub-SH act-MNs and compare it to fully reconstituted $H2B_{ss}Ub$ MN substrates. Specifically, the catalytic domain of the H3K79 methyltransferase, hDot1L(1-416) (referred to herein as hDot1L), was utilized (Figure 2.6a). Using ^{3}H-labeled S-Adenosyl-L-methionine (SAM) in the hDot1L methyltransferase assay, we found Ub-SH act-MN were able to activate hDot1L-mediated H3K79 methylation, and showed an expected 1:1 ratio of H2B-Ub to H3K79 methylation as evidenced by the 2-fold difference in fluorography between the di-ubiquitylated and mono-ubiquitylated MNs species (Figure 2.6a, middle panel lane 7). No methylation was observed on non-modified act-MNs (Figure 2.6a and b**,** lane 5). Neither the presence of TNB *in cis*, i.e. H2B-TNB, nor Di-TNB/TNB *in trans* affected hDot1L-mediated H3K79 methylation, indicative of the null effect of TNB toward hDot1L activity (Figure 2.6a and b**,** lanes 2 and 5). Importantly, and consistent with previous reports, Ub-SH *in trans* did not affect hDot1L activity toward H2B MN or $H2B_{ss}Ub$ MNs either in the same reaction i.e. Ub-SH act-MNs, or in homogenously prepared i.e. H2B MNs and $H2B_{ss}Ub$ MNs samples (Figure 2.6a and b**,** lanes 3, 4 and 7).

Moreover, to ensure the methylation reaction did not reach saturation on the time scale we were interested in, we performed a time course (Figure 2.7). We found that a time point of 15 min was optimal for our methyltransferase assays, i.e. the reaction had progressed significantly (and so could be easily monitored) but was far below saturation allowing meaningful comparisons to be made. We also calculated, using densitometry, the extent of ubiquitylation of the Ub-SH act-MNs

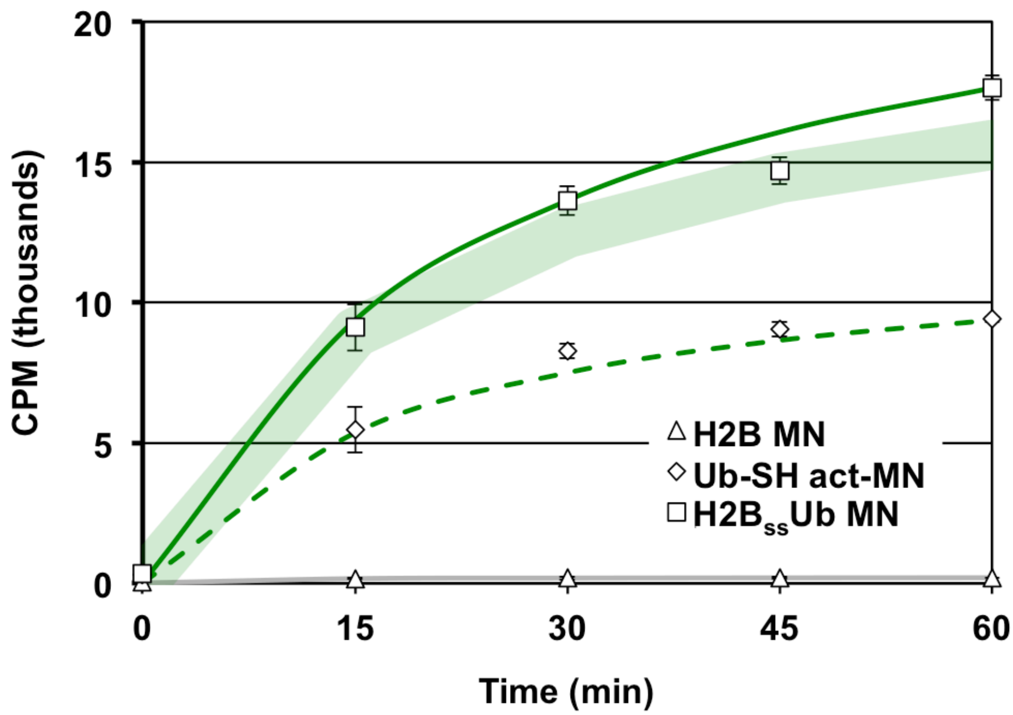


Figure 2.7. hDot1L methyltransferase assays on MNs with Ub made from on-nucleosomal ligations and on-histone reconstitutions. ^{3}H-SAM methyltransferase assays were performed on H2B MNs (grey line, triangle), and H2B$_{ss}$Ub MNs made either by on-histone reconstitution methods (H2B$_{ss}$Ub, green line, square) or on-nucleosomal ligations (Ub-SH act-MN, green dashed line, diamond). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting. By taking into account the extent of H2B$_{ss}$Ub in Ub-SH act-MNs, a fully 100% H2B$_{ss}$Ub stimulation curve was estimated (thick green line). Error bars, s.e.m (n = 3).

and estimated the amount of scintillation counts expected if these MNs contained 100% $H2B_{ss}Ub$ (Figure 2.7, thick green line). This estimation was similar to the activity of hDot1L on $H2B_{ss}Ub$ MNs. Collectively, these assays demonstrated that Ub-SH act-MNs were substrates for hDot1L-mediated methylation and could be used to further probe this *trans*-histone pathway.

2.7. Adaptability to other ubiquitylation sites on the MN

An attractive feature of our strategy is that it should be easily adaptable to any accessible ubiquitylation site on the MN surface via the preparation of a cysteine mutant histone and its subsequent incorporation into an act-MN. This circumvents the need to design completely new ligation junctions and multiple histone expression constructs typical for the preparation of a new protein via EPL semi-synthesis.

To test this concept we generated 4 different act-MNs containing the cysteine mutant histones (H2AK119C, H2BK116C, H2BK108C, and H3K79C) to compare to the H2BK120C act-MN ligation with Ub-SH. All ligation sites we employed are annotated sites of ubiquitylation in human histones, and map to different regions on the MN surface (Figure 2.8a).[58] Upon Ub-SH ligation, it was apparent that all of these act-MNs were capable of being ubiquitylated although to different extents. H2BK116C, H2BK120C and H3K79C act-MNs showed comparable Ub-SH ligation comprising 50-60% of total ligation sites (Figure 2.8a, lanes 3, 5 and 6). H2BK108C and H2AK119C ligated to a lesser degree, 30% and 20%

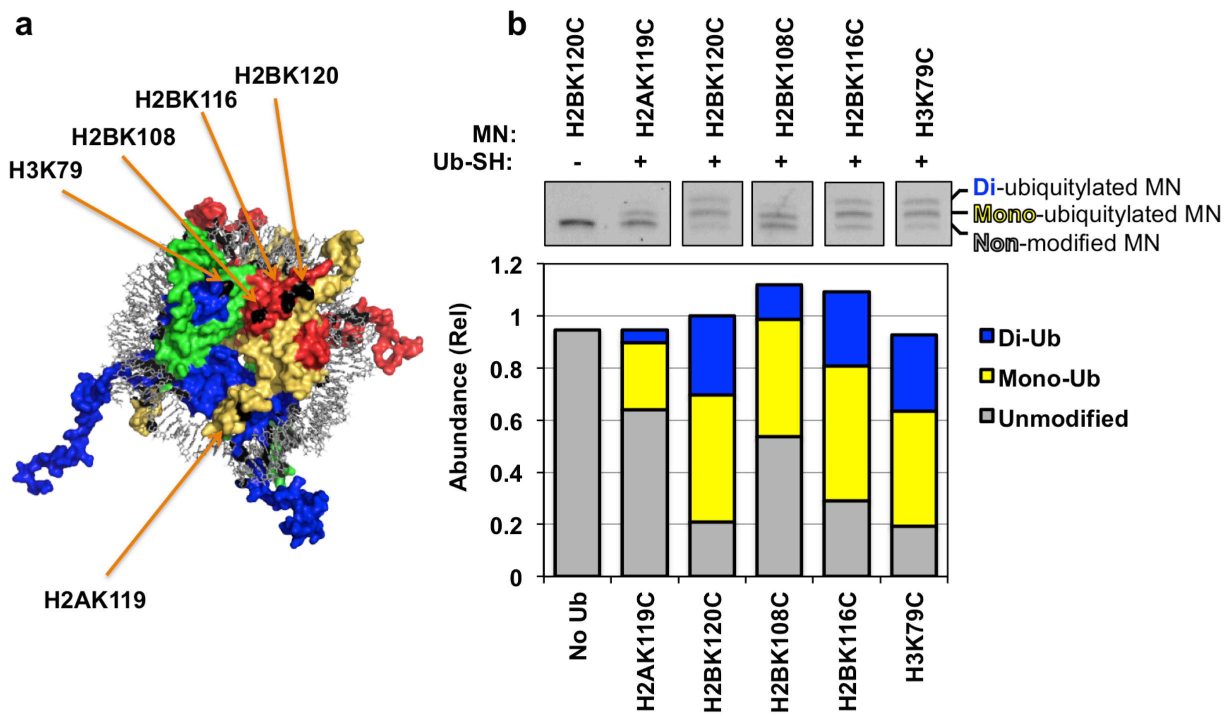


Figure 2.8. Direct Ub-SH ligations to other single cysteine histone mutants after incorporation into nucleosomes. Ub-SH was ligated to act-MN prepared with single cysteine histone mutants. (a) Histone cysteine ligate sites mapped to the MN surface (PDB 3LZ0. (b) Act-MNs were visualized by native-PAGE followed by ethidium bromide staining (top panel). Ub-SH act-MNs ubiquitylated species were quantified by normalizing the absolute amount of each species to the total amount observed for H2BK120C (bottom panel). The extent of the reaction was dependent on the ligation site, however all Ub-SH act-MNs show Ub-SH ligation.

respectively (Figure 2.8, lanes 2 and 4). Nonetheless, this ligation method worked for the attachment of Ub-SH to other regions of the act-MN surface and hence may be of practical use for accessing other histone ubiquitylations in a rapid, high throughput manner.

2.8. Summary and conclusions

In this chapter we explored the use of site-specific asymmetric disulfide exchange as an approach to ubiquitylate pre-assembled MNs containing a TNB activated cysteine, where the TNB activation is performed on the octamer level. The approach is compatible with nucleosomal reconstitution protocols, and achieves regioselectivity by exploiting the ability to introduce a unique cysteine residue into recombinant chromatin. While several methods for the preparation of H2B-Ub substrates have already been described, the method described here is simple, fast and can potentially be used to study other histone ubiquitylations and their associated functions in a high throughput manner. As this method relies on the late stage diversification of the Ub substrate, systematic structure-activity analysis of the system becomes tractable, which is the focus of the next chapter.

Chapter 3. A 13-member ubiquitin alanine mutant library to probe H2B-Ub stimulation of hDot1L activity

3.1. Introduction

The rapid synthesis of $H2B_{ss}Ub$ through the ligation of Ub-SH to act-MNs lends itself to SAR studies of the Ub surface in relation to H2B-Ub processes. In our initial validation of this technique, we tested the stimulation of hDot1L-mediated H3K79 methylation in the context of Ub-SH act-MNs. This showed that on-nucleosome ligations are amenable to hDot1L methyltransferase assays and the stimulation of hDot1L activity is directly attributed to the formation of $H2B_{ss}Ub$ within the Ub-SH act-MN ligation mixtures. In this chapter we utilize Ub-SH act-MN ligations combined with the creation of an Ub alanine mutant library for a comprehensive alanine scan of the ubiquitin surface to identify functional residues in hDot1L-mediated H3K79 methylation.

Compared to smaller PTMs such as acetylation and methylation, Ub is 'information rich' in that it alters the steric and electrostatic properties around its attachment site, as well as presenting a large surface area for the recruitment of binding factors. At 8.5 kDa, Ub is nearly as large as the histone to which it is linked (13.8 kDa in the case of H2B), increasing the MN surface by as much as 4,800 A^2.[181] Structural and biochemical studies of ubiquitin-ligand complexes,

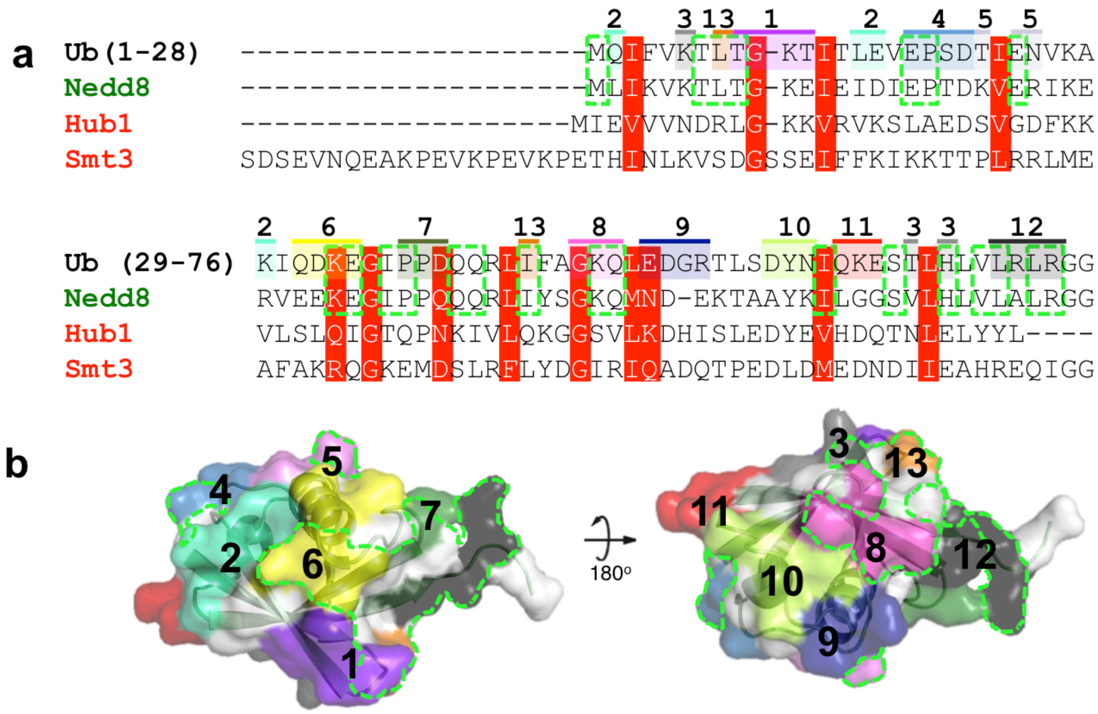


Figure 3.1. Design of the ubiquitin surface mutant library. (a) Ubiquitin and Ubl sequence alignment. Ubiquitin, Nedd8, Hub1 and Smt3 were aligned using the Lalign multiple sequence server. Similar residues shared between all Ubl's and ubiquitin are highlighted in red. Conserved residues between only ubiquitin and Nedd8 are outline in a green dashed box. The surface residues in ubiquitin that were mutated to alanine are shaded. The colors, and the numbers above, indicate the groupings in the 13 surface mutants. (b) Visual representation of the ubiquitin surface mutants mapped on to the ubiquitin structure (PDB 1UBQ3) in their respective color. Homologous regions between ubiquitin and Nedd8, but not Hub1 and Smt3, are showed in a dashed green line.

Table 3.1 Alanine ubiquitin mutants. All alanine mutants are shown and colored as in Figure 3.1. Alanine mutations are bolded in each respective mutant.

Ub	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub1	MQIFVKTL**AAAA**ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub2	M**A**IFVKTLTGKTIT**AA**VEPSDTIENVKA**A**IQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub3	MQIFV**A**TLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKES**A**L**A**LVLRLRG
Ub4	MQIFVKTLTGKTITLEV**AAAA**TIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub5	MQIFVKTLTGKTITLEVEPSD**A**I**AA**VKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub6	MQIFVKTLTGKTITLEVEPSDTIENVKAKI**AAAA**GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub7	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGI**AAA**QQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
Ub8	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFA**AAA**LEDGRTLSDYNIQKESTLHLVLRLRG
Ub9	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQL**AAAA**TLSDYNIQKESTLHLVLRLRG
Ub10	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLS**AAA**IQKESTLHLVLRLRG
Ub11	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNI**AAA**STLHLVLRLRG
Ub12	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV**AAAA**G

including the ubiquitylation of H2AK15, have revealed a canonical binding hotspot on ubiquitin involving a hydrophobic patch centered on L8/I44,[53, 179] however additional interaction surfaces, albeit not in a chromatin context, have been extensively characterized.[182, 183]

Previously, using EPL synthesized H2B-Ub reconstituted into H2B-Ub MNs, it was shown that the mutation of the canonical hydrophobic patch in ubiquitin (L8/I44) to alanine did not affect hDot1L stimulation.[121] A complementary study, using disulfide linked $H2B_{ss}Ub$ reconstituted MNs and involving the attachment of

ubiquitin like proteins (Ubls) to H2BK120C, revealed that ubiquitin functions in a residue specific manner.[121, 152] Nedd8 stimulated hDot1L while two other Ubls, Smt3 and Hub1, did not. These Ubl proteins share the same tertiary structure, but have between 23-55% sequence identity to ubiquitin, leading us to hypothesize that a specific surface shared between ubiquitin and Nedd8 was critical for hDot1L-mediated H3K79 methylation (Figure 3.1a). However, the cumbersome H2B-Ub synthetic approaches used for these studies severely limited the type of study that could be performed on a reasonable time scale. Thus, enabled by our recently developed on-nucleosomal ligation technology, we undertook the first comprehensive SAR study on the H2B-Ub/H3K79me crosstalk phenomenon.

3.2. Synthesis of a H2B-Ub mutant MN library

We hypothesized that stimulation of hDot1L methyltransferase activity required a specific surface region on ubiquitin, i.e. a functional 'hotspot'. To test this, the ubiquitin surface was subdivided into thirteen distinct patches, consisting of two to four solvent exposed residues, and each individual patch was mutated to alanine through site-directed mutagenesis (Figure 3.1). Note, one of these surface alanine mutants, Ub13, centered on the canonical L8/I44 binding hotspot, served as a internal positive control since it has been previously shown not to be involved in hDot1L stimulation.[121]

Preparation of the desired MN library began with the generation of the corresponding library of Ub mutants, each bearing a C-terminal thiol group to facilitate on-nucleosome ligation (Ub_{mut}-SH). Each of these was prepared analogous to the wild-type Ub-SH (as detailed in Chapter 2), i.e. by thiolysis of an Ub_{mut}-intein fusion protein followed by HPLC purification. All Ub_{mut}-SH were well behaved throughout the purification process and displayed similar thiolysis rates and yields compared to Ub-SH. Ub2-SH, due to the mutation of the +2 amino acid to alanine, was additionally missing the initiator +1 methionine (Figure 3.2a and b). This was presumably due to N-terminal processing during expression. All Ub_{mut}-SH proteins were characterized by ESI-MS and stored as lyophilized powders until use in ligation reactions (Figure 3.2). Typically, we obtained 20 mgs of each purified Ub_{mut}-SH protein from an initial 1 L *E. coli* expression system.

With the set of purified Ub-SH proteins in hand, we moved to the parallel assembly of the corresponding MN library. This was carried out by on-nucleosome ligation according to the conditions optimized in Chapter 2. All ligations, proceeded similarly to the wild-type Ub-SH ligation, with efficiencies between 0.51-0.71 as measured by total $H2B_{ss}Ub$ (Figure 3.3). The major exception to this was the Ub12-SH ligation, which proved sluggish and afforded mostly unmodified and mono-ubiquitylated MNs (total $H2B_{ss}Ub$ value of 0.25).

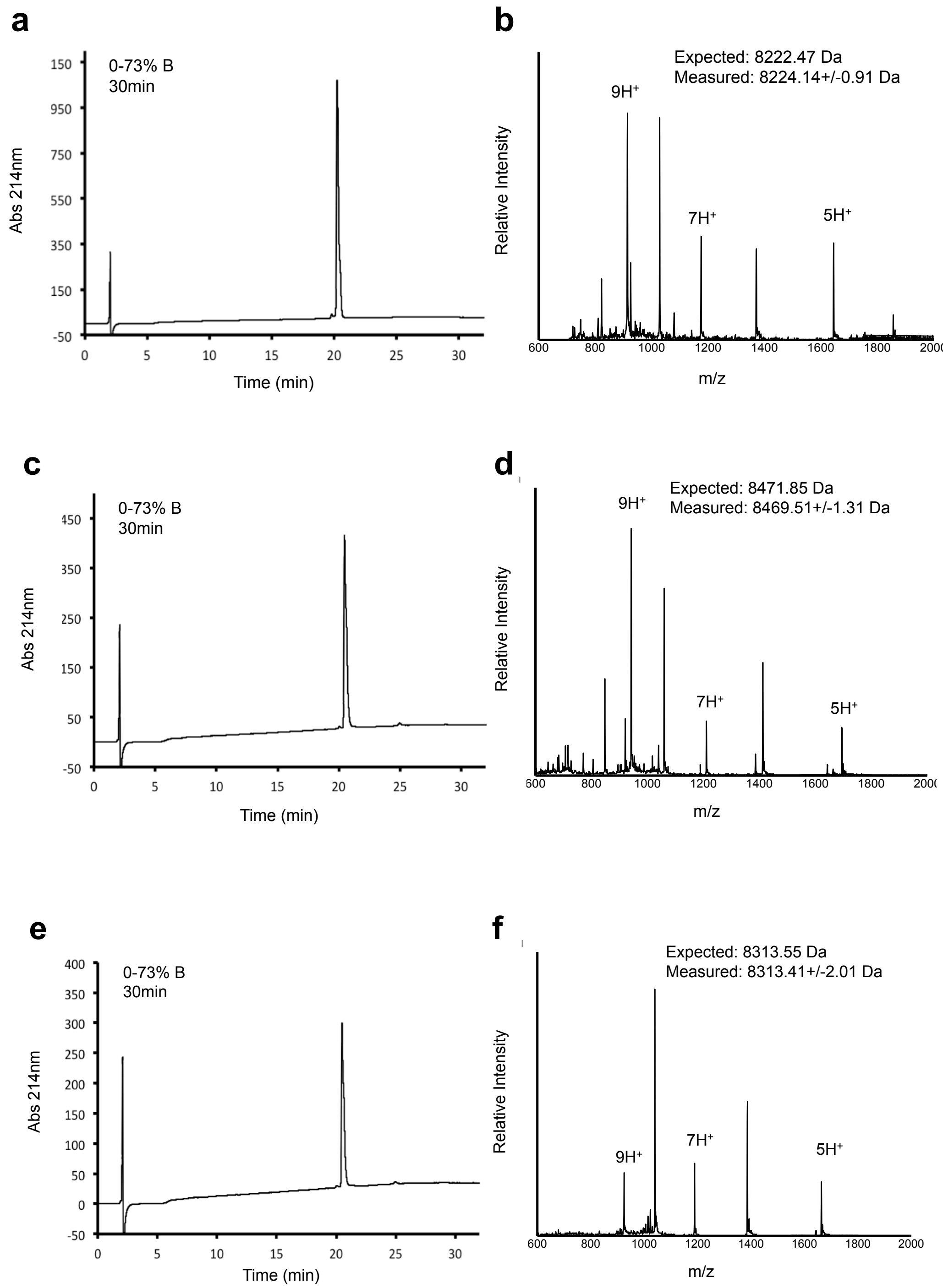


Figure 3.2. HPLC and Mass Spectrometry characterization of select Ub-SHs. (a) C18 analytical RP-HPLC chromatogram of purified Ub2-SH. (b) ESI-MS of purified Ub2-SH. (c) C18 analytical RP-HPLC chromatogram of purified Ub7-SH. (d) ESI-MS of purified Ub7-SH. (e) C18 analytical RP-HPLC chromatogram of purified Ub12-SH. (f) ESI-MS of purified Ub12-SH.

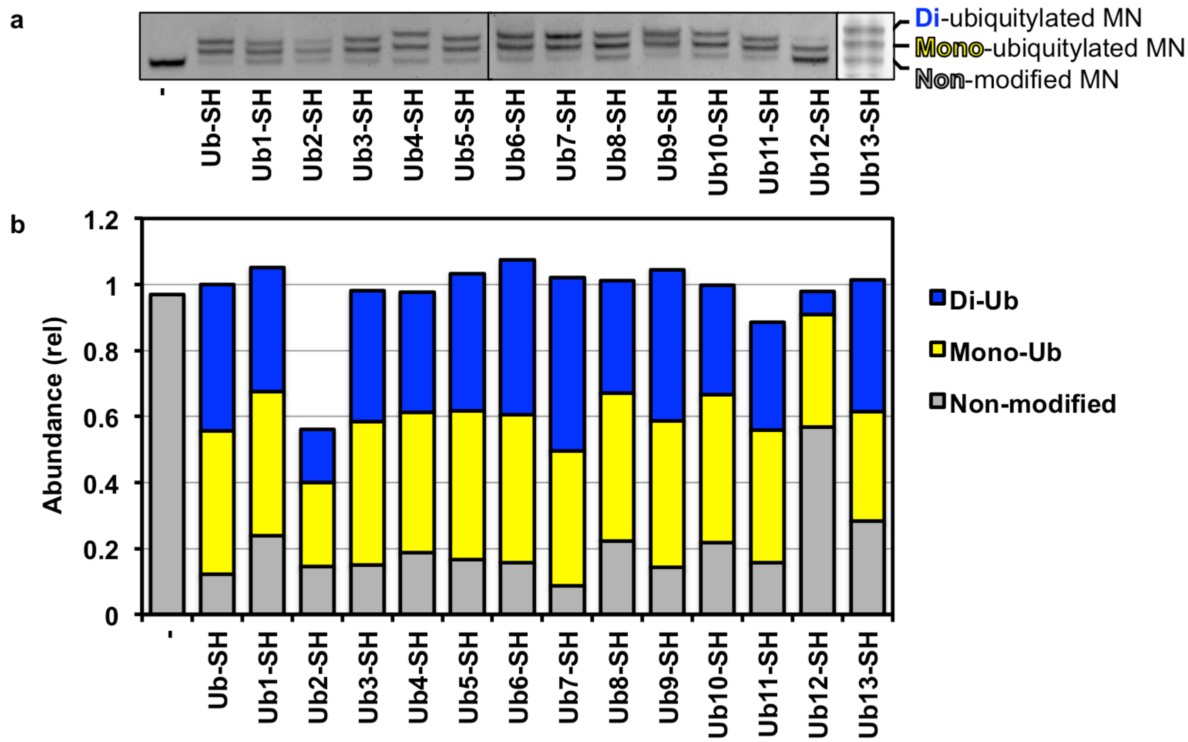


Figure 3.3. Analysis of Ub-SH and Ub_{mut}-SH ligation to act-MNs. (a) MNs were visualized by native PAGE followed by ethidium bromide staining. (b) Quantification of ligation was performed by densitometry and adjusted relative to Ub-SH act-MNs. Abundance, A_{Non} (grey box), A_{Mono} (yellow box), and A_{Di} (blue box) are plotted for each Ub_{mut}-SH. All ligations except for Ub2-SH and Ub12-SH were comparable to Ub-SH.

3.3. hDot1L methyltransferase assays with the Ub-SH act-MN library

With our Ub_{mut}-containing MN library in hand, we explored the Ub surface requirements needed to stimulate hDot1L-mediated H3K79 methylation. Note that since the presence of Ub-SH *in trans* has no effect on hDot1L stimulation (see Chapter 2, Figure 2.6), this initial screening experiment could be performed directly using the crude mixtures from each of the on-nucleosome ligation

reactions, thereby expediting the entire SAR experiment. Accordingly, a series of ^{3}H-SAM hDot1L methyltransferase assays were carried out on the library with act-MNs and Ub-SH act-MNs used as negative and positive controls, respectively (Figure 3.4). Quantitative ^{3}H-methyl incorporation, measured through scintillation counting, was adjusted to take the absolute values of $H2B_{ss}Ub$ (see section 2.4) into account termed Ub_{stim}:

$$Ub_{stim} = CPM/(\text{Total } H2B_{ss}Ub)$$

Where CPM = counts per minute

Ub_{stim} is normalized so that Ub_{stim} of Ub-SH act-MNs = 1

Thus Ub_{stim} was a measure of ^{3}H-methyl incorporation per absolute abundance of each $H2B_{ss}Ub$ contained in each Ub-SH act-MN reaction. After adjusting for differences in Ub-SH act-MNs ubiquitylation it was evident that Ub7 and Ub12 uniquely disrupted hDot1L-mediated H3K79 methylation while all other mutants, including the Ub13 mutant, showed activity comparable to that of wild type ubiquitin (Figure 3.4a, lane 2). Consistent with our initial hypothesis, the residues collectively mutated in Ub7 and Ub12 have 71% sequence identity to Nedd8, and only 14% sequence identity to both Hub1 and Smt3.

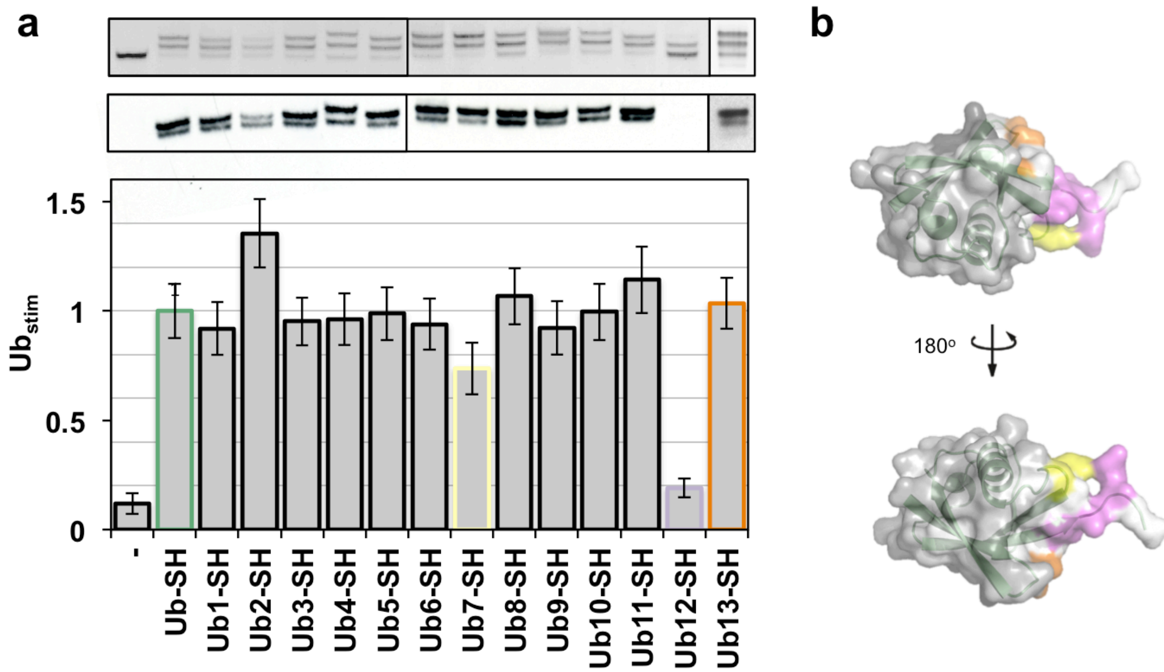


Figure 3.4. Surface features on ubiquitin critical for hDot1L stimulation. (a) hDot1L activity on each of the Ub surface mutants, 1-13. MNs were visualized by native-PAGE followed by ethidium bromide staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting and adjusted to include the extent of Ub-SH ligation, termed Ub_{stim} (bottom panel). Error bars, s.e.m (n = 6). Unmodified act-MNs and Ub-SH act-MNs were included as negative and positive controls, respectively. Only Ub7-SH act-MNs (yellow box) and Ub12-SH act-MNs (purple box) mutants led to a substantial reduction in hDot1L stimulation. The Ub13-SH act-MN mutant (orange box), centered on the canonical hydrophobic hotspot, did not lead to reduction in hDot1L activity, which is consistent with a previous study (see text). (b) Summary of alanine scanning results. Grey residues were tested and did not have an affect on hDot1L stimulation. Residues colored white were not tested. Ub13 (orange), Ub7 (yellow), Ub12 (purple)

3.4. Confirmation that Ub7 and Ub12 disrupt hDot1L activity

To confirm the results of this initial screen, and for further validation of our on-nucleosomal technology to probe SAR's in a nucleosome context, we prepared $H2B_{ss}Ub$ MNs containing the Ub7 and Ub12 mutants via the standard asymmetric disulfide reconstitution protocol, which affords homogeneous substrates. These were subsequently incorporated into octamers and MNs using standard procedures (Figure 3.5a, top panel).[180] Methyltransferase assays with hDot1L established that both mutants have reduced ability to activate the enzyme, with Ub12 having the most profound effect, again consistent with the initial screen (Figure 3.5a and b). Comparison of $H2B_{ss}Ub$ MNs methyltransferase results to that of Ub-SH MNs (through calculation of Ub_{stim} for the homogenously prepared MNs), revealed that the methyltransferase assays were consistent (Figure 3.5b). Interestingly, although not contiguous in primary

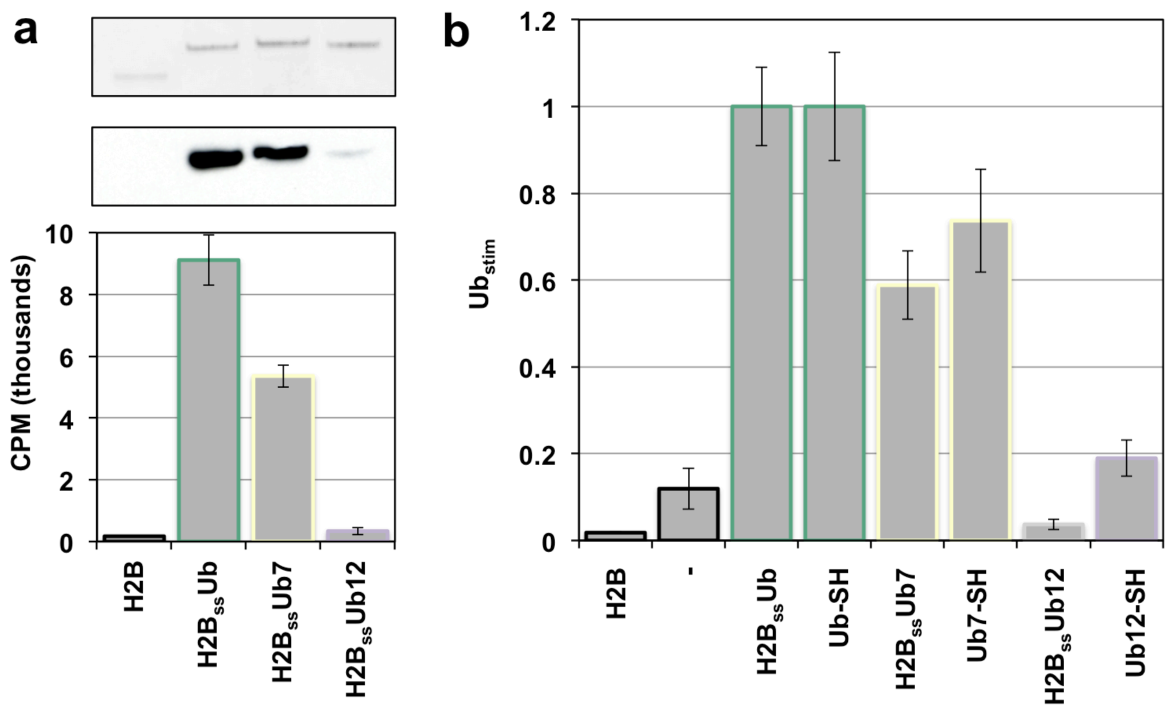


Figure 3.5. Comparison of hDot1L activity on H2B$_{ss}$Ub and Ub-SH act-MNs. (a) The reduced ability of the Ub7 and Ub12 mutants to stimulate hDot1L compared to Ub was evaluated on fully reconstituted (i.e. homogeneous) H2B$_{ss}$Ub MNs. MNs were visualized by native-PAGE followed by ethidium bromide staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting and reported in counts per minute. Error bars, s.e.m (n = 3). (b) Methyltransferase results from both fully reconstituted H2B$_{ss}$Ub MN (all MNs indicated by H2B$_{ss}$Ub$_{mut}$ label) and Ub-SH act-MNs (all MNs indicated with Ub$_{mut}$-SH label) are compared. Scintillation counts from panel a, were adjusted to include the extent of Ub-SH ligation, termed Ub$_{stim}$ and compared to data in Figure 3.4a (bottom panel, see text details). Error = s.e.m. (n = 3-6).

sequence, the residues mutated in Ub7 and Ub12 formed a contiguous patch near the C-terminus of ubiquitin (Figure 3.4b, yellow and purple patches).

3.5. Summary and conclusions

Collectively, in Chapters 2 and 3 we have demonstrated an efficient alanine scanning approach for SAR studies of histone ubiquitylations. This was applied to the study H2B-Ub in the context of hDot1L-mediated H3K79 methylation however this could easily be extended to the other annotated Ub sites within the octamer core or H2B-Ub functions (we'll revisit this idea in Chapter 6). Using this approach we generated 13 unique Ub_{mut}-SH act-MNs and tested their ability to stimulate of hDot1L activity. Ub7 and Ub12 were unable to upregulate hDot1L activity to comparable levels of Ub. This was further confirmed through the synthesis of $H2B_{ss}Ub7$ and $H2B_{ss}Ub12$ and the reconstitution of these constructs into MNs followed by hDot1L assays. As these residues are remarkably close in the tertiary structure of Ub, this chapter indicates that hDot1L functionally recognizes the Ub surface near the C-terminus of Ub. In Chapter 4 we will explore components of this patch further in order to identify the critical determinants of Ub governing hDot1L regulation.

Chapter 4. Investigation of the critical surface on Ub required for hDot1L-mediated H3K79 methylation

4.1. Introduction

In the previous chapter we used an on-nucleosomal ligation approach and tested the capacity of a library of Ub-SH mutants to stimulate hDot1L activity in the context of Ub-SH act-MNs. Additionally, the results from this scan were validated through hDot1L methyltransferase assays on homogeneous $H2B_{ss}Ub7$ MNs and $H2B_{ss}Ub12$ MNs prepared using a multi-step reconstitution approach. In this chapter we will delve deeper into Ub12, as this mutant had the most profound

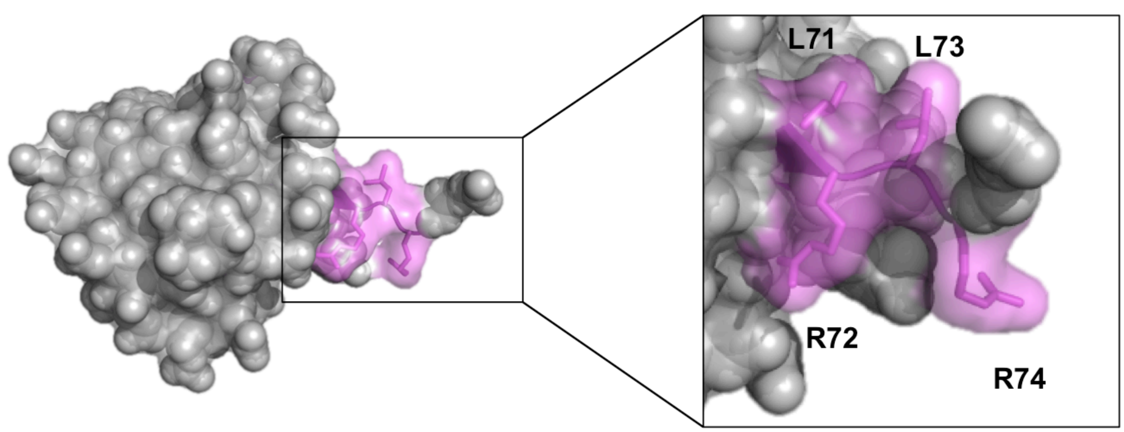


Figure 4.1. The Ub12 mutated patch mapped onto the Ub structure. Structure of Ub (PDB 1UBQ3) with the LRLR patch highlighted in purple. Residues from the LRLR patch are shown as a stick representation whilst other residues are shown as spheres. Inset: Close-up of LRLR amino acid residues.

effect returning hDot1L activity to levels exhibited by unmodified H2B MNs. Ub12 is a quadruple alanine mutant where the contiguous L71 R72 L73 R74 motif near the C-terminus of Ub has been mutated to alanine (Figure 4.1). This region is critical for Ub recognition by many identified E2/E3 ligases, displays structural plasticity, and has been implicated in many known Ub interactions.[182-185] Specifically, the last 6 amino acids of Ub (70-76) are directly involved in ~50% of Ub-protein interactions in all reported Ub-protein complex structures, however within these structures the conformation of Ub(70-76) is not conserved.[183, 186] Moreover, it has been found that all of these residues are essential for yeast viability. Ub engineered with any single alanine point mutant within the LRLR motif, as the sole Ub source in yeast, results in non-viable yeast strains.[184] Accordingly, we hypothesized that, due to the importance of this surface in other known Ub interactions as well as *in vivo*, the mutated residues contained within Ub12 comprise the functional surface of Ub recognized by hDot1L.

In this chapter, we will investigate the critical determinants within the Ub12 patch responsible for reduced act-MN ligation capacity as well as hDot1L activity. We were interested in whether the effect of these mutations was specific for H2BK120C act-MN ligation and hDot1L activity –and thus potentially linked-, or a general property of Ub12-SH. This could be rationalized as LRLR being a conserved structure motif within Ub recognized by hDot1L and/or the MN or Ub12 'mis'-folding respectively. To gain insight into this we subdivided the LRLR motif and employed further mutagenesis to test both hDot1L activity as well as

ligation to the act-MN. Using these results we structurally characterized the Ub mutant that we believe contains the two most central residues on Ub for stimulation of hDot1L activity.

4.2. The aberrant Ub12 ligation to act-MNs is not specific to H2BK120C MNs

The Ub12-SH act-MN conjugation reaction was characterized further in an attempt to determine whether the slow ligation kinetics observed was specific to the H2BK120C attachment site or a general consequence of the Ub12 mutations. We reasoned that the proximity of the mutations in Ub12 to the H2B-Ub ligation junction could be responsible for this effect (these mutations are adjacent to the

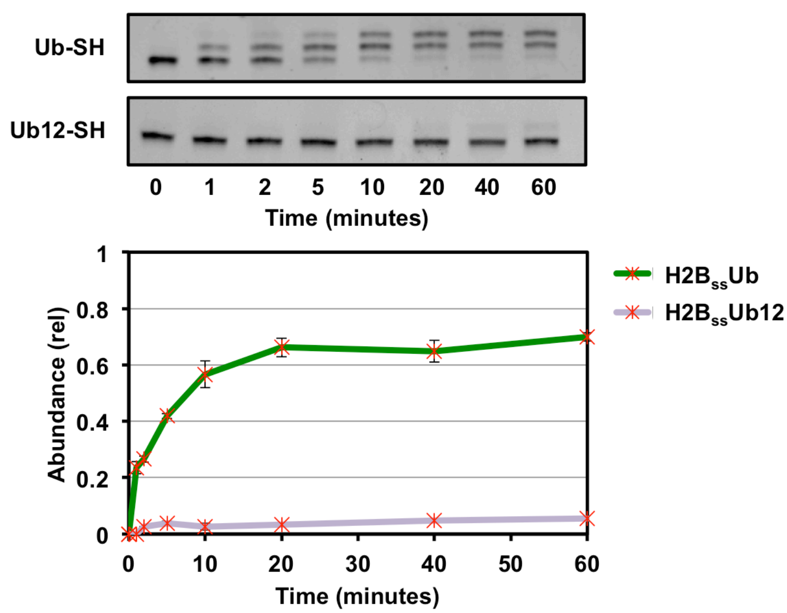


Figure 4.2. Ub12-SH ligation kinetics toward the act-MN. Time course for the on-nucleosome Ub-SH and Ub12-SH ligation to the act-MN. MNs at the indicated time points were visualized by ethidium bromide staining (top panels) and total H2B$_{ss}$Ub and H2B$_{ss}$Ub12 was calculated via densitometry (bottom panel). Error S.E.M. (n = 3).

C-terminal glycines of Ub). Further the mutated patch resulted in the C-terminal amino acid residue sequence of LVAAAAG instead of LVLRLRG, which could conceivably have major structural ramifications as the former is more non-polar. This non-polar mutated patch could either be occluded in the Ub12-SH structure or unable to interact with the highly charged MN surface explaining both the slowed ligation kinetics and the inability of Ub12-SH to stimulate hDot1L activity.

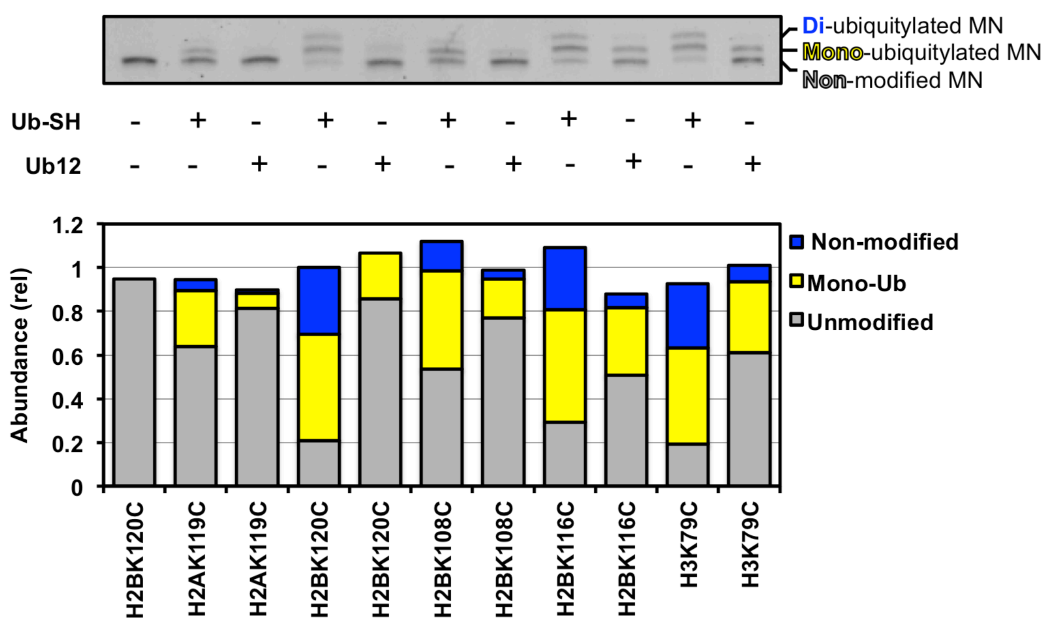


Figure 4.3. Ligation of Ub-SH and Ub12-SH to act-MNs containing different lysine to cysteine mutations. Ub-SH and Ub12-SH were ligated to multiple different act-MNs. MNs were visualized via native PAGE followed by ethidium bromide staining (top panel). Absolute abundance (A_x) (grey, yellow, and blue boxes) was calculated by densitometry. In all different act-MNs, Ub12-SH ligated to the act-MN less so than Ub-SH.

To explore this, we first performed a time course of the ligation of Ub12-SH to the act- MN to compare to that of Ub-SH act-MNs (Figure 4.2). As expected, the ligation of Ub12-SH to the act-MN was extremely slow at all tested time points (Figure 4.2). To determine whether this was specific for the H2BK120C act-MN or a general consequence of the Ub12-SH mutations we tested the ability of Ub12-SH to ligate to act-MNs at different ligation sites (Figure 4.3). Interestingly, we found that the slow reactivity of Ub12-SH was independent of the attachment site; the mutant had a reduced ability to ligate to all positions tested on the MN surface (Figure 4.3). Thus, the poor reactivity of Ub12-SH appeared to be an intrinsic feature of the Ub12-SH protein.

4.3. $H2B_{ss}$Ub12 prevents the stimulation of hDot1L activity through L71A and L73A amino acid residue mutations

As the H2B-Ub ligation junction can be varied substantially with little effect on hDot1L-mediated H3K79 methylation,[152] we reasoned we may be able to tease apart if the inability of $H2B_{ss}$Ub12 to promote H3K79 methylation and the slowed ligation kinetics of this mutant were in some way linked. We chose to subdivide Ub12 into two individual mutants retaining either both arginines (Ub L71/73A, referred to herein as uLL) or both leucines (Ub Arg72/74A, referred to herein as uRR); potentially narrowing this ligation/H3K79 methylation defect into either a

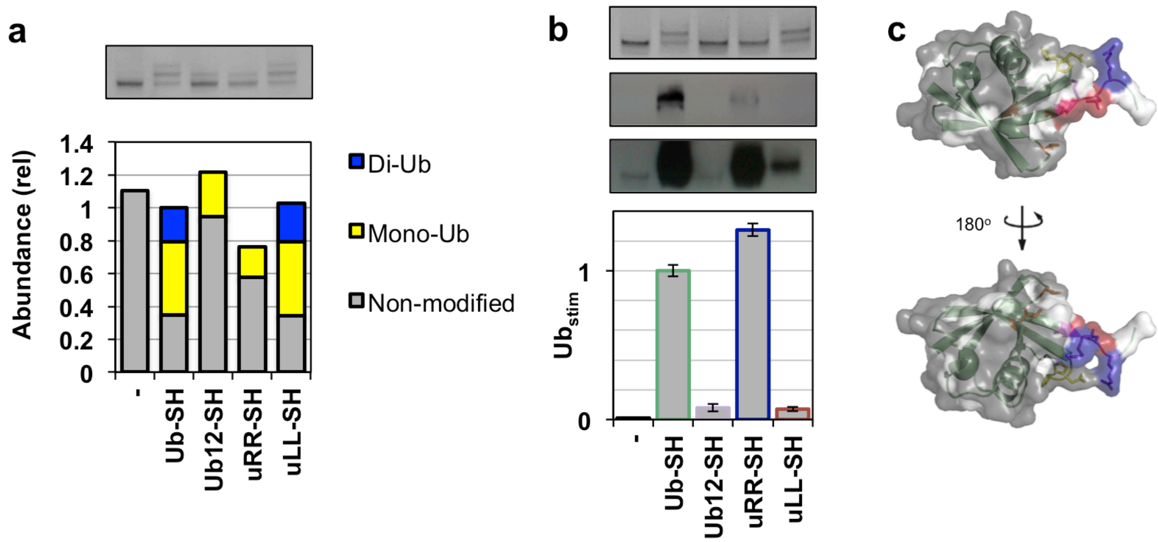


Figure 4.4. The ligation of uRR-SH and uLL-SH to the act-MN and subsequent methyltransferase assays. (a) Ub12 was split into two alanine sub-mutants, uLL and uRR, and Ub-SH act-MN ligations were monitored by native PAGE and visualized by Sybr Gold staining DNA (top panel). Non, mono, and di ubiquitylated species were quantified and the absolute abundance (Gray, yellow, and blue boxes) was calculated. (b) ^{3}H-SAM methyltransferase assays with uLL-SH and uRR act-MNs. MNs were visualized by Sybr Gold staining the DNA (top panel). ^{3}Methyl- incorporation was visualized by fluorography. Two different exposures of 12 hours (top middle panel) and 5 days (bottom middle panel) are shown. Ub_{stim} was calculated based on filter binding assays followed by scintillation counting (bottom panel). (c) uRR (blue) and uLL (red) mutations mapped on to the ubiquitin surface (PBD 1UBQ3).

charge-based or hydrophobicity-based effect. The two Ub12-derived mutants were prepared, ligated to the act-MN and hDot1L methyltransferase assays were performed (Figure 4.4a and b). The uRR mutant showed a ligation defect similar to Ub12, but still promoted robust H3K79 methylation considering the extent of ligation of this mutant (Figure 4.4a and b lane 4). Specifically, uRR-SH ligated to the act-MN similar to Ub12-SH, with both ligations only showing a small fraction of the mono species and no di species, in contrast with what was observed for Ub-SH (Figure 4.4a, lanes 3 and 4 compared to lane 2). However, subsequent to hDot1L methyltransferase assays showed ^{3}H-methyl incorporation for uRR-SH more so than Ub12-SH (Figure 4.4b, lanes 3 and 4). Conversely, the uLL mutant ligated to the MN to a similar extent as wild type Ub (Figure 4.4a, lanes 2 and 5). Surprisingly, considering the extent of uLL-SH ligation compared to uRR-SH, uLL-SH was unable to stimulate hDot1L-mediated H3K79 methylation (Figure 4.4b, lane 5). Taken together, the R72/74A double mutation prevented ligation to the act-MN whilst the L71/73A double mutation was responsible for the inability of these mutants to stimulate hDot1L activity. Thus, the ligation deficiency of Ub12-SH and its inability to activate hDot1L activity are not linked.

4.4. uLL is unable to stimulate hDot1L in multiple contexts[1]

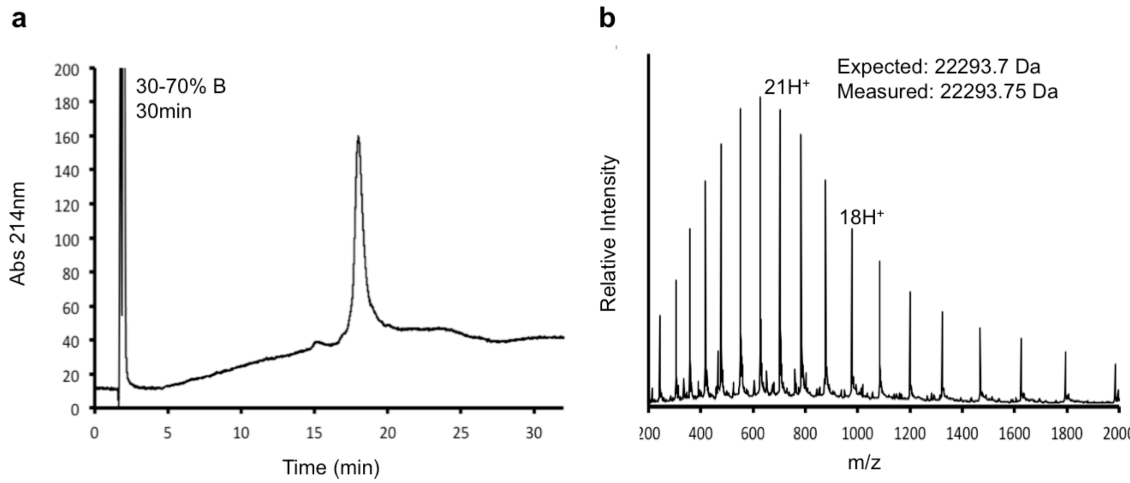


Figure 4.5. Characterization of H2B-uLL prepared via EPL. (a) C18 analytical RP-HPLC chromatogram of purified H2B-uLL. (b) ESI-MS of purified H2B-uLL.

We chose to characterize the uLL mutant further as it contributed to the bulk of the reduction in hDot1L stimulation compared to Ub and Ub12. We synthesized uLL with an isopeptide linkage to H2B (H2B-uLL), in order to determine whether the non-native conjugation chemistry used in the initial analysis (i.e. $H2B_{ss}$uLL in the context of uLL-SH act-MNs) had any combinatorial effect in this crosstalk. H2B-uLL was reconstituted into MNs and hDot1L activity was assayed (Figure 4.5, Figure 4.6a). Importantly, H2B-uLL was unable to stimulate hDot1L in the context of the MN in comparison to H2B-Ub, which was consistent with the

[1] The H2B-Ub and H2B-uLL EPL-based histones were synthesized by Sam Pollock and full-length hDot1L was recombinantly prepared by Beat Fierz in Professor Tom Muir's

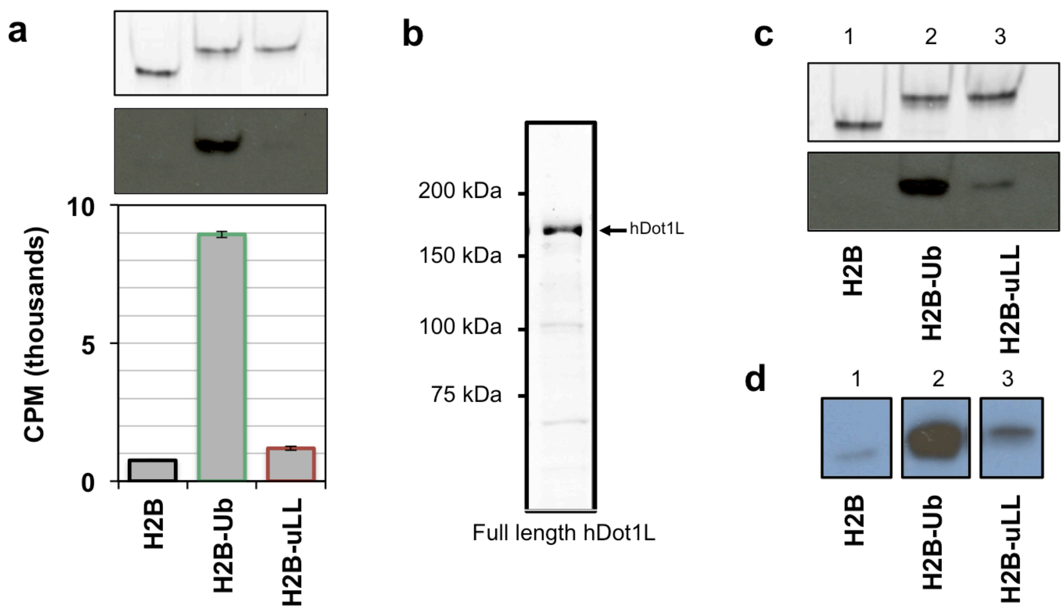


Figure 4.6. hDot1L methyltransferase assays with H2B-uLL containing MNs and chromatinized plasmids. (a) ^{3}H-SAM methyltransferase assays were performed on MNs (visualized by Sybr Gold, top panel) using hDot1L(1-416). ^{3}H-methyl incorporation was visualized by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting (bottom panel). Error bars, s.e.m (n = 3). (b) Coomassie stained SDS PAGE gel showing full length hDot1L purified from insect cells. (c) ^{3}H-SAM methyltransferase assays were performed on MNs (visualized by Sybr Gold, top panel) using full-length hDot1L. ^{3}H-methyl incorporation was visualized by fluorography (bottom panel). (d) H2B-Ub and H2B-uLL containing octamers were reconstituted into a chromatinized plasmid and ^{3}H-SAM methyltransferase assays with full-length hDot1L were performed and analyzed via fluorography.

previous hDot1L assays using the disulfide linked $H2B_{ss}$uLL (Figure 4.6a lane 2 compared to lane 3). Since this experiment employed the catalytic domain of hDot1L, we were keen to see whether the results held true using the full-length hDot1L. Accordingly, a flag tagged full-length hDot1L was expressed in insect cells and purified (see section 8.4.11 for details) (Figure 4.6b). As with the catalytic domain of hDot1L, the full-length enzyme had greatly reduced methyltransferase activity against H2B-uLL MNs compared to H2B-Ub containing MNs (Figure 4.6c). Moreover, we observed the same behavior in the context of a chromatinized plasmid template. Specifically, H2B, H2B-Ub and H2B-uLL containing octamers were individually reconstituted into a plasmid containing 23 copies of the 601 positioning sequence and full-length hDot1L methyltransferase assays were performed (Figure 4.6d). H2B and H2B-uLL were unable to stimulate hDot1L activity to the same extent as H2B-Ub in these substrates (Figure 4.6d, lane 2 compared to lane 3). Taken together, our biochemical data clearly indicates that the region of the Ub surface centered on L71-L73 is essential for stimulation of hDot1L methyltransferase activity.

While it is tempting to conclude that this region is directly involved in the stimulation by, for example, acting as a binding hotspot, it is conceivable that mutation of these two leucine residues leads to a misfolding of Ub and, hence, the disruption of a *bonafide* hotspot elsewhere on protein surface. Thus, it was important to examine the structural integrity of the uLL mutant. For this we chose

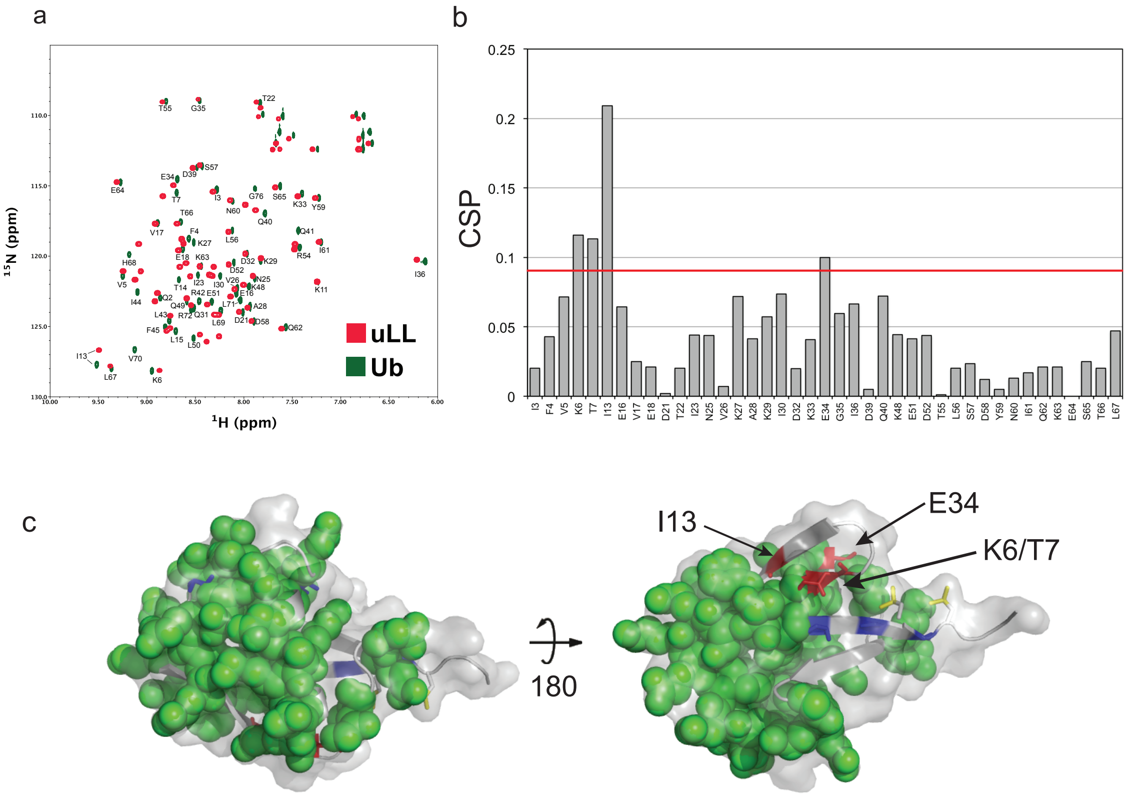


Figure 4.7. HSQC spectrum of Ub and uLL. (a) ^{1}H-^{15}N HSQC spectrum of uniformly ^{15}N labeled uLL was compared to an HSQC spectrum of uniformly ^{15}N labeled Ub. All peaks were annotated from the BMRB databank (accession code 6457). (b) Chemical shift perturbation analysis of annotated peaks assigned in both the Ub and uLL spectra. The averaged, weighted CSP larger than the 10% trimmed mean + 2 standard deviations (red line) was used as a cutoff to identify significant CSPs. The average, weighted CSP was calculated according to CSP = $(\Delta\delta_H^{2} + (\Delta\delta_N{*}0.2)^{2})^{1/2}$, where Δδ is the difference in resonance position between Ub and uLL. Note 0.2 is a weighting factor used to compare ^{1}H and ^{15}N chemical shifts. (c) Visual representation of CSP analysis. Ub is shown as a grey surface representation (PBD 1UBQ3). Residues for which CSPs were calculated are shown as green spheres. Significantly perturbed amino acid residues are labeled and shown in red. Amino acid residues annotated in the Ub spectra and not found in uLL are shown in blue. L71 and L73 are shown in yellow sticks.

solution NMR spectroscopy since Ub has been extensively studied using this method.[183] Indeed, the presence or absence of the Ub fold can be easily ascertained by comparing the amide fingerprint region in the [^{15}N, ^{1}H] heteronuclear single quantum coherence (HSQC) spectrum of the uLL mutant with that of native Ub.

A uniform ^{15}N labeled version of Ub and uLL (both containing a C-terminal carboxylic acid) were prepared for HSQC spectroscopy using established *E. coli* expression methods (see section 8.4.4 for more details). Analysis of the fingerprint region in [^{15}N, ^{1}H] HSQC spectrum of uLL in comparison to Ub revealed that the majority of chemical resonances were conserved between Ub and uLL (we were unable to assign 8 resonances between Ub and uLL) (Figure 4.7a). Further, the majority of the annotated peaks (37 of 41) were similar between both proteins (Figure 4.7a and b). Out of the four residues that were shown to be perturbed, two (K6 and E34) were included in the Ub alanine scan employed in Chapter 3 and did not affect hDot1L activity (Figure 3.4, lanes 5 and 8). All four of these amino acid residues are proximal to L71 and L73 (Figure 4.7c, colored red). Analogously, chemical resonances unidentified in the uLL spectrum, due to lack of proximal chemical resonance, were localized to the vicinity of the L71 and L73 mutations (Figure 4.7c, colored blue). This is indicative that uLL retains a native ubiquitin fold, ruling out the possibility that uLL lacks activity through indirect means by globally disrupting the Ub tertiary

structure. Collectively, these data indicate that stimulation of hDot1L activity is directly mediated by this hydrophobic surface-patch on Ub.

4.5. A Hub1-Ub chimera containing L71/73 is unable to recover hDot1L activity toward H2B$_{ss}$Hub1 MNs

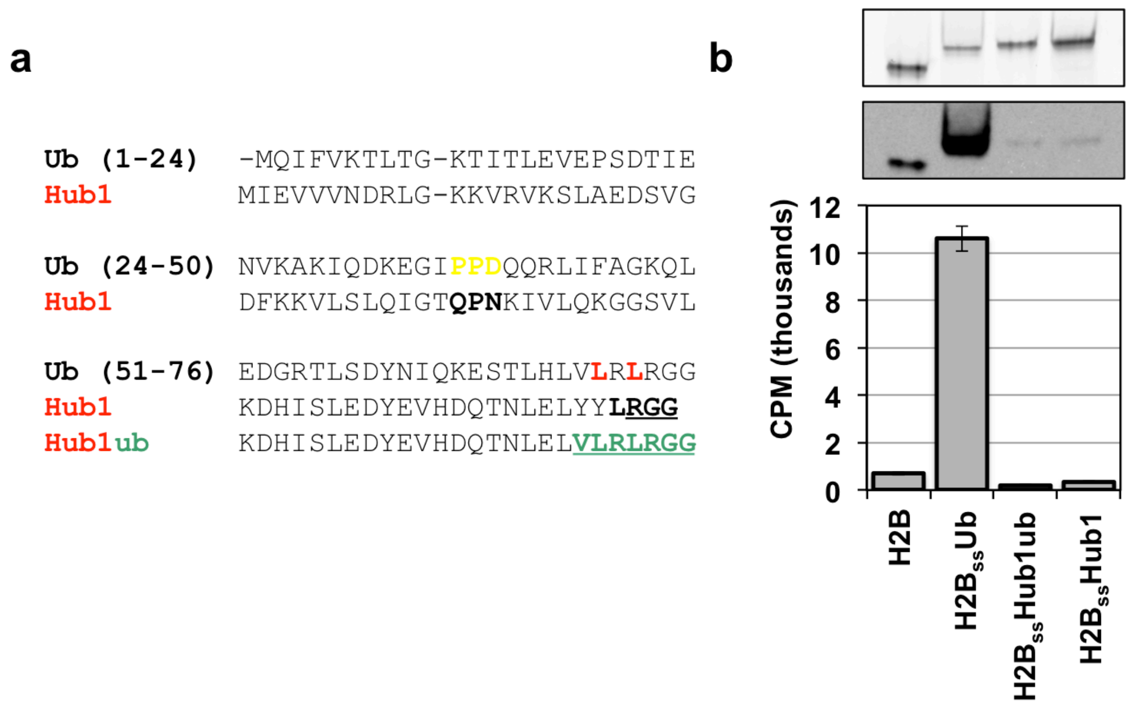


Figure 4.8. hDot1L methyltransferase assays on a H2B$_{ss}$Hub1-Ub chimera. (a) The Hub1ub chimera is shown along side of both Ub and Hub1. As Hub1 and Hub1ub are identical in residues 1-50 these residues are not shown in Hub1ub for simplicity. The last 7 residues of Hub1ub are identical to Ub (bolded green) and include L71 and L73 (bolded red). The Ub7 mutated patch in Ub is shown (bolded yellow). (b) ^{3}H-SAM methyltransferase assays were performed on MNs (visualized by Sybr Gold, top panel). ^{3}H-methyl incorporation was visualized by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting (bottom panel). Error bars, s.e.m (n = 3).

Lastly, we wondered whether we could use the information derived from our SAR studies on Ub, namely the identification of the L71-L73 hotspot, to reverse engineer an ubiquitin-like protein that is normally incapable to stimulating hDot1L. For this we chose Hub1, which, in line with our hypothesis, differs from Ub at its C-terminus and is specifically the missing an LR in the LRLR motif (Figure 4.8a). Further, the construct employed in the original Hub1 study (note the endogenous Hub1 does not contain and amino acid residues RGG (Figure 3.1 compared to Figure 4.8a)) is one amino acid residue short of a homologous C-terminal Ub tail and contains bulky tyrosine residues adjacent to the LRGG motif (Figure 4.8a).

We designed a chimeric protein (Hub1ub) in which the last 7 residues from the C-terminus of Hub1 (residues 69-76) were replaced with that of Ub (Figure 4.8a). The corresponding $H2B_{ss}$Hub1ub construct was synthesized, reconstituted into MNs and methyltransferase assays were performed using the catalytic domain of hDot1L (Figure 4.8b). $H2B_{ss}$Hub1ub MNs failed to stimulate hDot1L, as its activity was comparable to that of unmodified H2B and $H2B_{ss}$Hub1 MNs. Conceivably, our inability to “ubiquitinize” Hub1 through this simple sequence switch might relate to a requirement for additional ‘second shell’ residues that would place the leucine hotspot in the proper structural context. Indeed, examination of the Ub and Hub1 sequences reveals sequence differences in regions that are proximal to the location of the hotspot, for example in the PPD patch that reduced the ability of Ub7 to stimulate hDot1L activity by 50% when mutated to alanine (Figure 3.5 and Figure 4.8a). Thus, installing the hDot1L

hotspot into Hub1 is likely to require much more sophisticated protein-engineering strategy than the somewhat simple one attempted herein.

4.6. Summary and conclusions

An alanine scan of the Ub surface in regard to hDot1L activity from Chapter 3 led to the identification of a contiguous surface patch on Ub critical for hDot1L stimulation. This chapter expanded upon this initial finding and explored Ub12 through mutagenesis to determine the critical components governing the regulation of hDot1L activity. We elucidated that L71 and L73 within Ub were necessary for hDot1L-mediated H3K79 methylation. This was further confirmed in multiple different types of homogenously prepared chromatin substrates and reproduced using the full-length hDot1L polypeptide. This led to the conclusion that H2B-uLL is unable to stimulate hDot1L activity in chromatin contexts beyond our initial simplified system. Moreover, uLL did not show any appreciable affect on the global tertiary structure of Ub. This was directly investigated by NMR and additionally inferred by the ability of uLL-SH to ligate to the act-MN comparably to Ub-SH.

It is interesting that Ub12 could be segregated into either an Ub-SH ligation incompetent (i.e. uRR) or hDot1L stimulation incompetent activity (i.e. uLL) mutant. As uRR involves the mutation of two arginine amino acid resides to alanine in the amino acid sequence LVLRLRGG, we hypothesize that this results in a highly non-polar patch at the C-terminus of Ub that is structurally dissimilar to

that of native Ub (or uLL). This hydrophobic patch could in some way be occluded as to prevent Ub12-SH/uRR-SH ligation to the act-MN. It would be interesting to investigate this further especially considering it is able to stimulate hDot1L mediated H3K79 methylation. Potentially, this region could be 'mis'-folded when uRR/Ub12-SH is re-folded in solution to prevent ligation to the act-MN, however the 'correct' conformation could be selected for by hDot1L and/or the MN in order to stimulate H3K79 methylation through the recognition of L71/73 in the context of the H2B-Ub MN.

It would also be valuable to examine the structural ramifications of uLL more closely. From our HSQC experiments, we observed local perturbations surrounding L71 and L73 in uLL compared to Ub. However, as our experiments detected chemical shifts primarily in the polypeptide backbone, we do not know what role these leucine to alanine mutations play in side chain dynamics surrounding the L71 and L73 sites. Thus it would be interesting to characterize uLL using multi-dimensional NMR experiments to determine side chain dynamics using residual dipolar coupling experiments that have previously been employed to study Ub.[183] Analogously, for the Hub1ub chimera, it would be worthwhile to compare the local dynamics of L71 and L73 in this mutant to that of UbL71/L73. This could differentiate between the hypotheses that Hub1ub is unable to promote H3K79 methylation due to incorrect positioning (or dynamics) of L71/73 or potentially other additional factors, such as an additional global steric/interaction requirement. The validity of comparisons using Hub1 will be

discussed further in section 7.3.2. Although the H2B-uLL MN clearly disrupts hDot1L activity we have not tested whether the Ub7 mutant disrupts hDot1L activity by disrupting the confirmation L71/73 or through independent means, which would be another potential reason why our attempt to ‘ubiquitinize’ Hub1 was unsuccessful.

Although, the NMR experiments suggested herein would be powerful approaches to systemically characterize the findings throughout this chapter, we were interested in the broader consequences of Ub (specifically L71 and L73) in the context of chromatin structure. In Chapter 5 we will take biophysical approaches to investigate the structural ramifications of Ub in the context of H2B-Ub MNs as well as H2B-uLL in the context of chromatin arrays to gain further insight into the role of H2B-Ub in the stimulation of hDot1L activity.

Chapter 5. The biophysical characterization of Ub and uLL in MNs and nucleosomal arrays

5.1. Introduction

Certain histone PTMs can affect chromatin function through the direct alteration of chromatin structure.[30, 117, 138, 187, 188] This has been demonstrated in the context of the MN,[138] as well as in the context of the 30-nm fiber.[187, 188] Specifically, histone PTMs such as lysine methylation and lysine acetylation modulate inter and intra-nucleosomal dynamics through the disruption of histone-DNA and

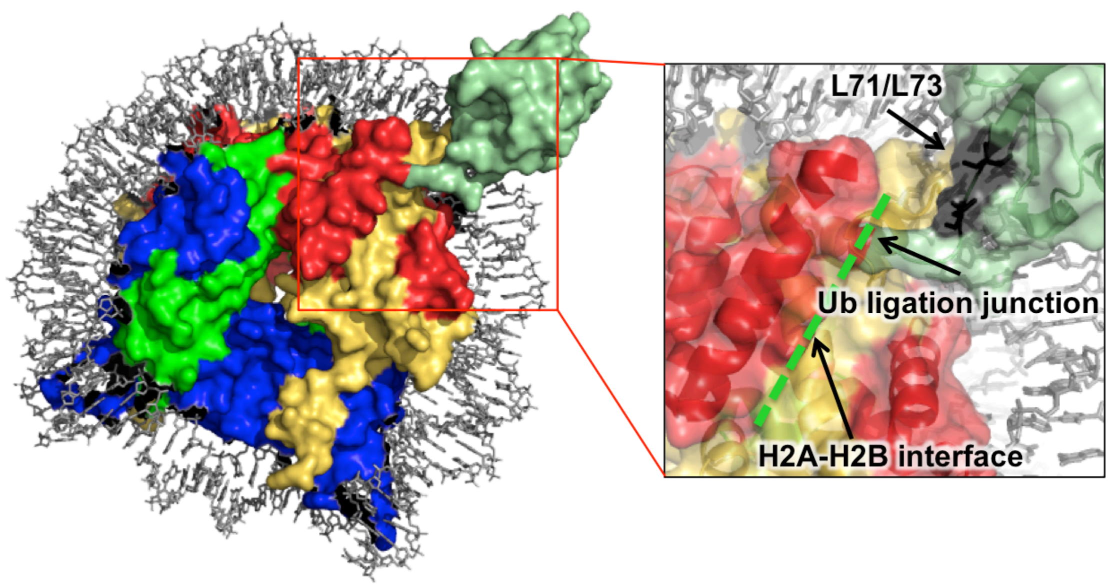


Figure 5.1. Localization of L71 and L73 in relation to the H2A-H2B dimer interface in the MN. The H2B-Ub nucleosome is modeled as a composite of Ub and MN structures (PDB 1UBQ3 and 3LZ0). Histones H2A (yellow), H2B (red), H3 (blue), H4 (green), Ub (light green) are shown. Inset: The L71/73 patch (black) is located near the Ub ligation junction which is in proximity to the hydrophobic H2A-H2B dimer interface.

histone-histone contacts. In this chapter we set out to investigate, using biophysical approaches, whether H2B ubiquitylation alters chromatin structure and, if so, whether the functional hotspot identified in the preceding chapters (i.e. L71 and L73) is involved. As the L71 and L73 hydrophobic side chains are solvent exposed in the Ub structure, we expect these residues to form direct contacts with either the MN surface in the vicinity of H2BK120 or hDot1L to facilitate H3K79 methylation (Figure 5.1). We hypothesized that, given their proximity to the hydrophobic interface between H2A and H2B, L71/73 may orient Ub on the MN surface through histone contacts (Figure 5.1). Moreover, if Ub is indeed interacting with the MN surface, this interaction could be responsible for a myriad of associated H2B-Ub biochemical functions, extending beyond H3K79 methylation.

We were specifically interested in the impact of the L71/73 hydrophobic patch on H2B-Ub chromatin structure in order to gain mechanistic insight into the regulation of hDot1L and to see whether this epitope was conserved in H2B-Ub-mediated chromatin decompaction. This was addressed by: (i) Probing the solvent accessibility of H3K79 in H2B and H2B-Ub containing MNs; (ii) An NMR approach, that specifically detects the isoleucine, leucine, and valine methyl groups of Ub in the context of the MN, and; (iii) Comparison of the compaction properties of $H2B_{ss}uLL$ to $H2B_{ss}Ub$ containing nucleosomal arrays. These experimental approaches suggest that L71 and L73 do not have any appreciable

effect on H2B-Ub chromatin structure in the context of either the MN or nucleosomal arrays.

5.2. The H2B-Ub MNs does not alter the solvent accessibility of H3K79

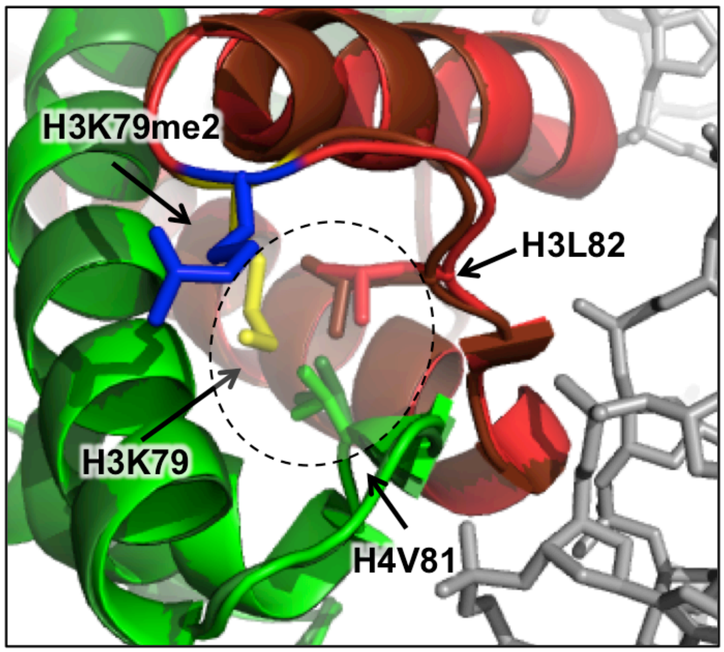


Figure 5.2. Effect of H3K79me2 on the MN structure. The structure of the region surrounding H3K79 in MNs with and without H3K79me2 are compared (PDB code 3C1C and 1KX5 respectively). H4 is shown as a cartoon colored green (Forest dark for H3 MN, green for H3K79me2 MN). H3 is shown as a cartoon colored red (Maroon for H3 MN, red for H3K79me2 MN). Residues H3L82, H4V81, H3K79 (yellow) and H3K79me2 (blue) are shown as stick representations. Dashed circle indicates hydrophobic surface exposed upon local rearrangement of H3K79 when dimethylated. This results in a solvent accessible small hydrophobic cavity surrounding the H3L82 and H4V81 amino acid residues.

There are several different mechanisms by which ubiquitylation of H2B might stimulate methylation of H3K79. One possibility involves an allosteric-type mechanism in which the presence of H2B-Ub changes the local structure around H3K79 making it more accessible to the hDot1L. Indeed, comparison of the x-ray crystal structures of MNs containing unmodified H3K79 and H3K79me2, reveals that the region surrounding H3K79 undergoes a local structural rearrangement upon methylation (discussed further in section 1.3.2) (Figure 5.2).[138] Specifically, H3K79 assumes a different rotameric structure upon methylation that leads to the exposure of a hydrophobic cavity. Given this structural plasticity, we designed an experiment to determine whether H2B-Ub causes a change in H3K79 solvent accessibility in the context of the MN.

We prepared a H3K79C mutant histone (also containing a H3C110A mutation) in order to test the accessibility of H3K79C towards a cysteine reactive probe. Specifically, a maleimide polyethylene glycol (mPEG) molecule was utilized to react with cysteine resulting in a stable thioether linkage between the sulfhydryl group of cysteine and the maleimide of mPEG. The mPEG used had a molecular weight of 6 kDa so that, in the context of the H3K79C MN, this reaction could be monitored via gel shift upon native PAGE followed by silver staining. MN substrates were prepared containing H3K79C, and either H2B or H2B-Ub histones. Cysteine-free MNs that were not ubiquitylated were reconstituted and were used as a control to detect any non-specific mPEG modifications.

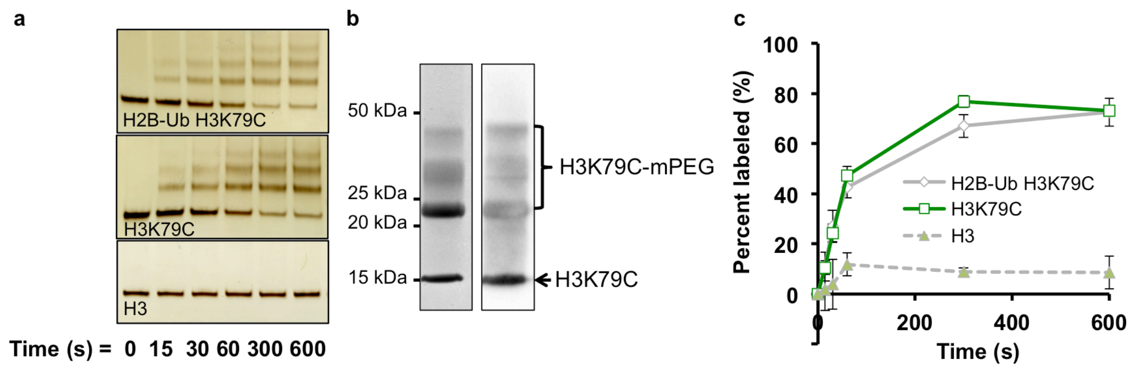


Figure 5.3. The mPEG modification of H3K79C and H3K79 MNs containing H2B or H2B-Ub. (a) A silver stained Native PAGE gel showing a time course of mPEG modified MNs. MNs are indicated in the lower left corner of each panel, and mPEG modification is obsered via gel shift from the unmodifed band at Time = 0. (b) SDS-PAGE (left panel) and western blot analysis (right panel) of H3K79C modified with mPEG. H3K79C was reacted with mPEG similar to MNs and visulized via Coomassie staining subsequent to SDS-PAGE as well as western blot analysis with an antibody directed to H3. A band runing at the molecular weight of H3 (15kDa) is observed, as well as higher MW bands (mPEG modifications) that contain H3 (left panel western blot analysis). (c) The percentage of mPEG modification was analyzed via densitometry of the silver stained naitve PAGE mPEG modification time course (panel a). H2B-Ub MNs (grey line) containing H3K79C showed simlar mPEG modification to that of H3K79C MNs (green line), whereas H3 MN (gray dashed line) were not modified. Error bars, s.e.m (n = 3).

A time course of the mPEG reactions with all three MNs species (H3, H3K79C, and H2B-Ub H3K79C) were performed over the course of 10 minutes and quenched with dithiolthreitol (DTT) (Figure 5.3a). As expected, the cysteine-less MN (H3 MN) showed no gel shift upon mPEG addition and thus mPEG did not non-specifically modify MN substrates under the conditions tested (Figure 5.3a, lower panel). H2B or H2B-Ub MNs containing H3K79C showed multiple mPEG

modification states (Figure 5.3a, upper and middle panels). The majority of the mPEG modification species were mono- or di- modified and the reaction proceeded primarily in an additive fashion, i.e. MNs were modified with mono-mPEG at earlier time points in the reaction, followed by the formation of di-mPEG. Based on the lack of modification of the H3 MNs, we believe the modification states observed beyond di-mPEG (as each H3K79C containing MNs has two cysteine residues) were the result of heterogeneities in the mPEG reagent. This was supported by western blot analysis of the free H3K79C histone subjected to mPEG modification (Figure 5.3b).

Given the uncertainties of the higher mPEG modification states, we quantified the non-modified MN bands by densitometry and calculated the percentage of MNs labeled (the unmodified MN band at Time = 0 correspond to 0% labeled) (Figure 5.3c). By this measure, we observed that H3K79C together with either H2B or H2B-Ub MNs showed a similar degree of mPEG modification at all time points tested (Figure 5.3c). Further, the degree of labeling neared 70% in both contexts, suggesting that H3K79 is quite solvent exposed regardless of the ubiquitylation state. This data argues against the role of H2B-Ub in hDot1L stimulation being related to a dramatic change in the accessibility of the H3K79 side chain.

5.3. Nuclear magnetic resonance spectroscopy reveals little ubiquitin-specific structural perturbations of the H2B-Ub MN[2]

As previous attempts to crystallize H2B-Ub MNs have proven unsuccessful,[189] we took a solution NMR approach to gain structural insights into this system. The approach we took was inspired by a recently reported NMR study that elucidated the interaction of the High mobility group nucleosomal 2 (HMGN2) protein with the MN surface.[190] This study was performed on MN containing wild-type histones from *Drosophila melanogaster* (abbreviated as *d*H2B for example, as opposed to *Xenopus laevis* histones which will be labeled as *x*H2B, etc.) utilizing transverse relaxation-optimized NMR spectroscopy (TROSY). TROSY is a technique for the NMR analysis of large supra-molecular systems, as conventional solution NMR techniques are limited to a molecular size of less than ~50 kDa (the MN is ~200kDa in size) (Figure 5.4).[191] This size constraint results because the transverse relaxation time (T_2) is inversely proportional to the rotational correlational time (τ_c) of the molecule. As T_2 determines the linewidth of the NMR signals, a long τ_c (short T_2) can result in broad, un-interpretable NMR spectra (Figure 5.4b). TROSY exploits destructive interference between T_2 mechanisms to generate spectra with greater signal-to-noise (Figure 5.4c).

[2] Methyl-TROSY experiments were performed in collaboration with Dr. Julianne Kitevski in Lewis Kay's lab at Toronto University.

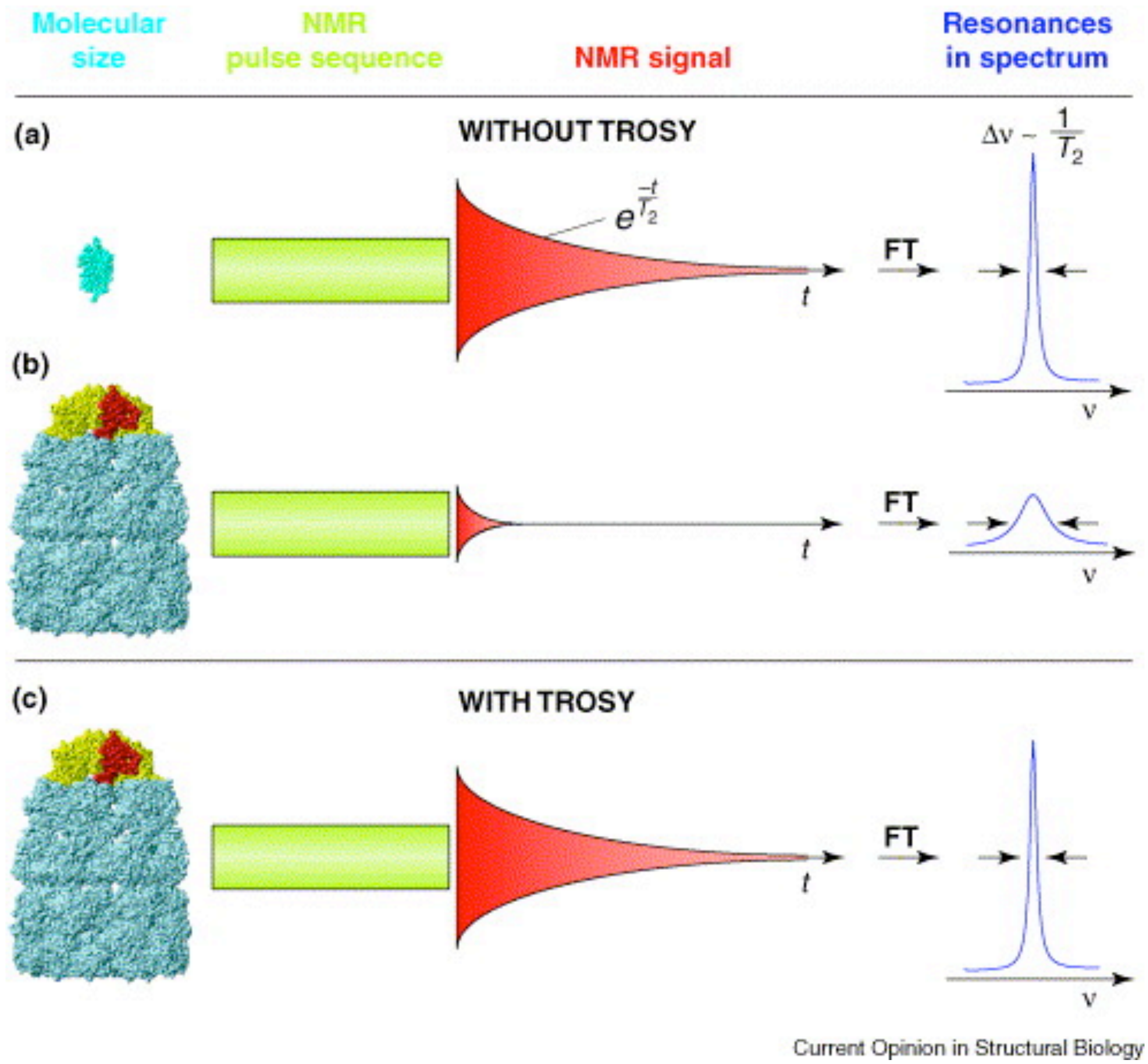


Figure 5.4. The molecular size of a protein molecule influences the NMR signal. This figure is reproduced from Fernadex et al (2003).[192] (a) For a protein below ~50kDa, a short correlational time (τ_c) results in a long T_2 and thus narrow line widths (Δv) of the NMR signal upon Fourier transformation (FT). (b) For larger molecules, e.g the proteasome, the very long τ_c (small T_2) results in the rapid decay of the NMR signal and broad line widths (Δv). (c) With TROSY, the signal can be partially recovered, although sample deuteration is usually required.

The so-called methyl-TROSY variation of the experiment additionally improves the sensitivity by selecting for the slow relaxation coherences of isotopically labeled methyl groups (Figure 5.4c).[193-195] These methyl probes [^{13}C, ^{1}H] are incorporated into fully deuterated protein samples through the metabolic labeling of isoleucine, leucine, and valine (ILV) methyl groups (see section 8.4.5 for more details).[196] In the context of the nucleosome, each histone was ILV labeled separately and incorporated into a MN where the other three histones were deuterated which allowed for better cross-peak resolution. Additionally to annotate ILV methyl resonances, 58 single-point mutant ILV labeled histone constructs were incorporated into MNs and analyzed by Methyl-TROSY.[190] Considering the significant amount of effort that went into assignment of the ILV labeled *Drosophila* histone resonances, we planned to test the effect of H2B-Ub in the context of the *Drosophila* MN.

To elucidate whether Ub was interacting with the MN surface (potentially through the L71/L73 patch), we set out to prepare an ILV-labeled Ub and incorporate it into an H2B$_{ss}$Ub MN NMR sample. However, as the preparation of fully deuterated histones and the subsequent ligation of an ILV labeled Ub-SH (Ub$_{ILV}$-SH) is non-trivial, we first wanted to determine whether structural insights gained from the *d*H2B-Ub MN would be relevant for hDot1L-mediated H3K79 methylation. We did not attempt to express and purify the *Drosophila* Dot1 homologue, Grappa, as this is a poorly characterized 1848 amino acid residue

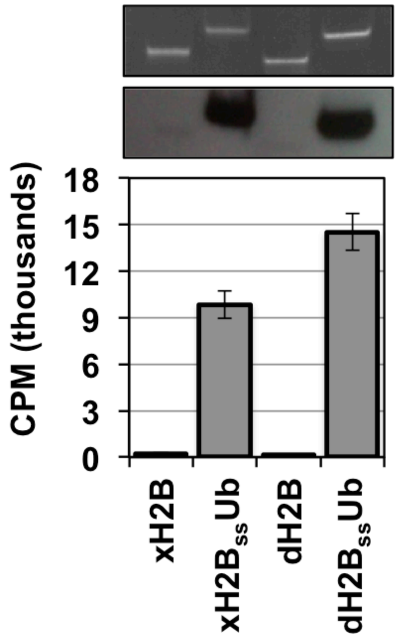


Figure 5.5. hDot1L methyltransferase assays with *x*H2B and *x*H2B$_{ss}$Ub MNs compared to *d*H2B and *d*H2B$_{ss}$Ub MNs. ^{3}H-SAM methyltransferase assays were performed on *x*H2B, *x*H2B$_{ss}$Ub, *d*H2B, and *d*H2B$_{ss}$Ub MNs. MNs were visualized by native-PAGE followed by ethidium staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting (bottom panel), Error S.E.M (n = 3).

polypeptide. However, as Grappa and hDot1L are have 42% sequence identity (57.1% sequence identity in the conserved structured catalytic core hDot1(1-330)) and *d*- and *x*- histones show greater than 71% sequence identity, we hypothesized hDot1L would be stimulated by *d*H2B-Ub MNs and thus any structurally related insights we gained from the *d*H2B-Ub MN could be rationalized toward hDot1L and *x*H2B-Ub MN biochemical results. Ligation of this construct with Ub-SH was performed and the resulting conjugate was

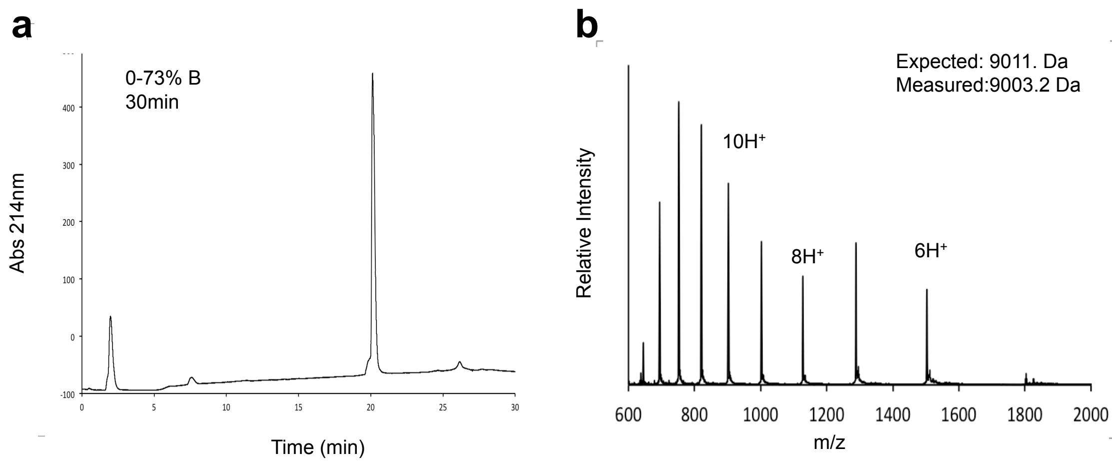


Figure 5.6. Characterization of Ub_{ILV}-SH prepared via a Ub-NPU-6xHis construct. (a) C18 analytical RP-HPLC chromatogram of purified Ub_{ILV}-SH. (b) ESI-MS of purified Ub_{ILV}-SH.

incorporated into *Drosophila* MNs. We employed the same reconstitution and purification protocols for *Drosophila* MNs as for *Xenopus* MNs and the former behaved similarly to the latter. ^{3}H-SAM hDot1L methyltransferase assays were performed on both the *Xenopus* and *Drosphila* MNs (Figure 5.5). As expected, no ^{3}H-methyl incorporation was observed for the non-ubiquitylated MN substrates, whereas hDot1L activity was strongly stimulated upon ubiquitylation of H2B, for both the *Xenopus* and *Drosophila* MNs (Figure 5.5). Based on this result, we reasoned that any structural effects observed for Ub in the context of the *Drosophila* MN could be extended to our previous biochemical analysis of hDot1L-mediated H3K79 stimulation.

We next sought to express and purify a deuterated Ub_{ILV}-SH. As recombinant expression yields in deuterated media are typically far less than in LB media,

recombinant expression in deuterated media of the Ub-NPU-6xHis construct was extensively optimized. The generation of Ub-SH was achieved by thiolysis of the Ub-NPU-6xHis construct subsequent to affinity purification. After intein thiolysis and HPLC purification, we were able to obtain a yield of ~7mgs/liter of pure Ub_{ILV}-SH (Figure 5.6). These yields are suitable for the preparation of an NMR sample however, as this was pursued in parallel with alternative ligation techniques, we did not attempt to incorporate Ub_{ILV}-SH into fully deuterated *d*MNs.

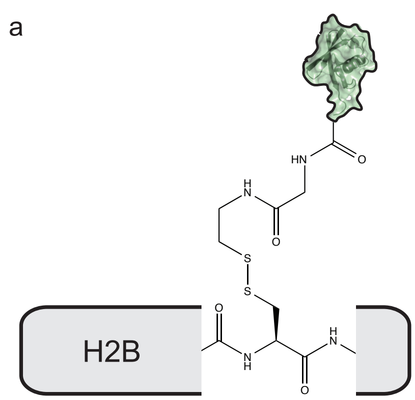


Figure 5.7. Comparasion of $H2B_{ss}Ub$ ligation and $H2B_{ss}UbG76C$ ligation junctions. (a) Ligation of Ub-SH to dH2BK118C results in a ligation junction that is 1 atom longer than the native isopeptide junction (see Figure 1.10 for more details). (b) The recombinant expression of UbG76C and subsequent ligation to H2BK118C, results in a ligation junction similar to Ub-SH with an additional carboxylate group proximal to the disulfide bond.

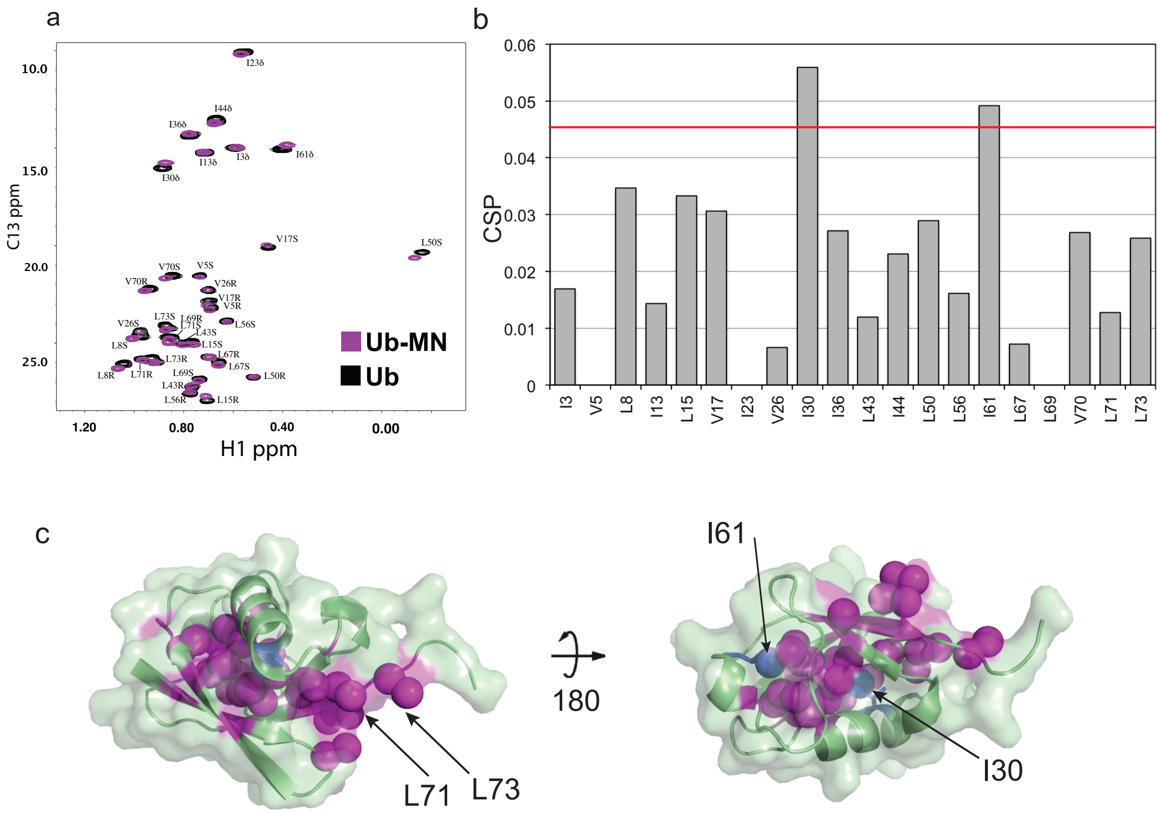


Figure 5.8. Methyl-TROSY of *d*H2B$_{ss}$UbG76C$_{ILV}$ MNs and CSP analysis. (a) Methyl-TROSY spectra of UbG76C$_{ILV}$ incoprorated in MNs (Ub-MN, purple) compared to free UbG76C$_{ILV}$ in solution (Ub, black). Methyl probes for the respective amino acid residues are labeled and include chirality (R or S configuration) when necessary. (b) (^{1}H, ^{13}C) chemical shift perturbation analysis of all residues in UbG76C$_{ILV}$ MNs compared to UbG76C$_{ILV}$ free in solution. Red line indicates the 10% trimmed mean + 2 standard deviations (see section 8.10.3 for analysis details). (c) Visual representation of Ub$_{ILV}$ methyl groups mapped onto the Ub structure (PDB 1UBQ3) as spheres. Methyl groups that displayed chemical shifts similar to Ub$_{ILV}$G76C free in solution (purple), and significantly perturbed methyl groups (blue) are shown. Significantly perturbed amino acid residues and L71 and L73 are labeled.

Alternatively, an Ub_{ILV} construct containing a G76C mutation was prepared for ligation to *d*H2BK118C. Ub_{ILV}G76C was ligated to *d*H2BK118C and a fully deuterated $H2B_{ss}Ub$ mononucleosome was reconstituted. The integrity of this ligation junction, which contains an additional carboxylate has yet to be validated with respect to H2B-Ub-related functions (Figure 5.7). Methyl-TROSY experiments were performed using $dH2B_{ss}Ub_{ILV}$G76C MNs. The cross peaks observed from Ub_{ILV}G76C, incorporated within MNs, was compared with Ub_{ILV}G76C free in solution (Figure 5.8a, purple peaks compared to black resonances). Chemical shift perturbations (CSP) were calculated and 2 residues (I30 and I61) were shown to have significantly different chemical shifts (significant being greater that 2 standard deviations from the 10% trimmed mean) (Figure 5.8b). These residues did not co-localize to a specific region of Ub and instead were dispersed throughout the hydrophobic core of Ub (Figure 5.8b). Additionally, it is clear from these studies that Ub_{ILV} is not interacting with the *Drosophila* MN as to significantly affect the chemical shifts of the annotated L71 and L73 methyl groups (Figure 5.8b). Although these experiments are still in their initial stages, we hypothesize that L71 and L73 do not interact with the MN in any significant way. Indeed, the NMR data are more consistent with the idea that Ub 'hangs off' the MN from its H2B attachment site, as opposed to the modification making intimate contacts (for example through the L71 and L73 hotspot) with the MN.

5.4. L71 and L73 are not involved in H2B-Ub-induced impairment of chromatin fiber compaction

The biochemical and spectroscopic studies in the preceding sections suggest that attachment of Ub to H2B in the nucleosome does not lead to a major structural perturbation of Ub or the MN, at least around H3K79. However, these studies were, by experimental necessity, performed on a MN. Thus, we were keen to see whether the L71/L73 'hotspot' played any role in the known ability of H2B-Ub to impede chromatin compaction.[117] Interestingly, while our laboratory showed that H2B-Ub disrupts higher-order chromatin structure, attachment of Hub1 to the same site did not.[117] This suggests that a specific surface of Ub, rather than its size per se, is required for this physical effect. We of course wondered whether the L71/L73 surface required for hDot1L stimulation might also be required for this structural property of H2B-Ub.

We decided to test this idea using a previously developed homo-fluorescence resonance energy transfer (homo-FRET) steady-state anisotropy approach (SSA) (Figure 5.9).[117] We prepared 12-mer nucleosome arrays containing H2B, $H2B_{ss}Ub$ or $H2B_{ss}uLL$ octamers (Figure 5.9), where the H2A variant used was a H2AN110C histone labeled with fluorescein maleimide (fH2A). Fluorescein was the spectroscopic probe necessary for Homo-FRET. Arrays were formed with a DNA template containing 12 copies of the 177bp '601' positioning sequence (12x601 DNA). ScaI restriction enzyme cut sites were engineered between the

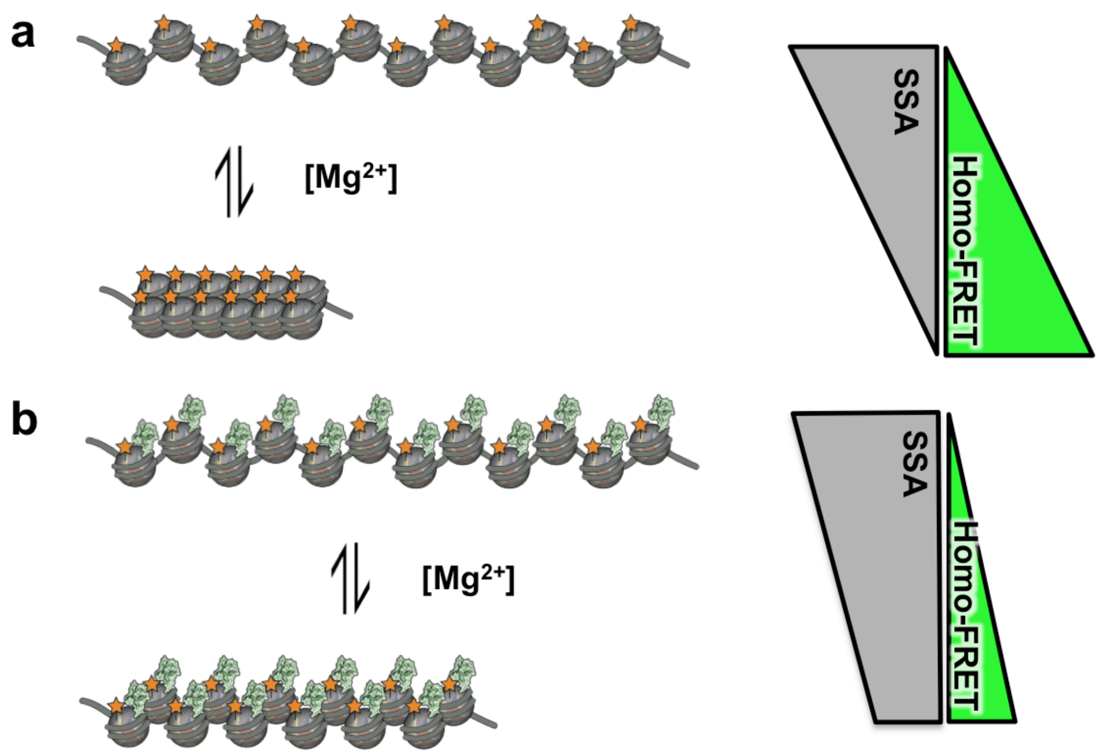


Figure 5.9. Schematic of the Homo-FRET assay used to probe chromatin compaction. (a) 12-mer nucleosome arrays containing fH2A (indicated by orange star) compact upon addition of Mg^{2+}. Compaction correlates with a lower steady-state anisotropy (SSA) signal due to homo-FRET (right panel). (b) 12-mer nucleosome arrays containing H2B-Ub (Ub in green) compact upon addition of Mg^{2+}. Schematic illustrates the fact that the presence of H2B-Ub is known to impede compaction compared to unmodified arrays. This results in a smaller change in SSA for H2B-Ub containing arrays compared to unmodified arrays readout by a smaller change in Homo-FRET (right panels, panel a compared to b).

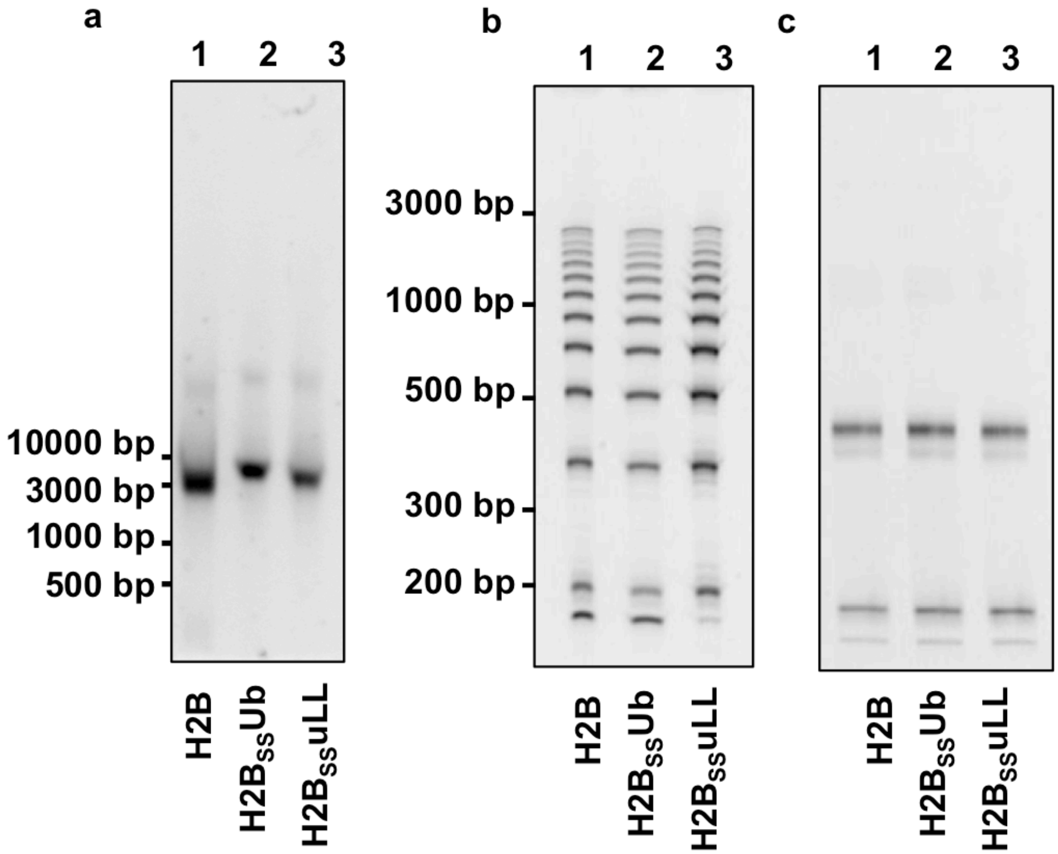


Figure 5.10. Reconstitution of unmodified, H2B$_{ss}$Ub, and H2B$_{ss}$uLL fluorescently labeled chromatin arrays. (a) Sybr gold stained native 1% agarose, 1% polyacrylamide gel electrophoresis (APAGE) analysis of reconstituted H2B (lane 1), H2B$_{ss}$Ub (lane 2), and H2B$_{ss}$uLL (lane 3) nucleosome arrays. (b) Sybr Gold stained native PAGE gel following micrococcal nuclease (MNase) digestion of reconstituted H2B (lane 1), H2B$_{ss}$Ub (lane 2), and H2B$_{ss}$uLL (lane 3) nucleosome arrays. (c) Sybr Gold stained native-PAGE gel showing ScaI restriction enzyme digests of reconstituted H2B (lane 1), H2B$_{ss}$Ub (lane 2), and H2B$_{ss}$uLL (lane 3) nucleosome arrays. The nucleosome band is observed at ~500 bp and a small amount of free 601 DNA is observed around ~200 bp.

‘601’ positioning sequence sites (601 sites) for analysis of full 601-site octamer occupancy.

Nucleosome arrays were formed using a semi-gradual dialysis method similar to MN formation, in the presence of the 155 bp ‘low affinity’ Mouse mammary tumor virus (MMTV) DNA.[117] The addition of MMTV DNA prevents the formation of oversaturated, and biophysically unsuitable, arrays. H2B, $H2B_{ss}Ub$, and $H2B_{ss}uLL$ arrays were formed in parallel and were analyzed through native 1%, agarose, 1% polyacrylamide gel electrophoresis (APAGE) followed by staining with Sybr Gold (Figure 5.10a). All arrays showed the expected single band running near the molecular weight of 3000 bp. Micrococcal nuclease (MNase) digests were performed to determine the positioning of the octamers on the 12x601 DNA template. After MNase digests DNA was extracted and PAGE was performed followed by Sybr Gold staining. 12 distinct DNA bands were observed for all NAs, indicative of 12 well-positioned octamers on the DNA template (Figure 5.10b). Complementary to the MNase digests, Sca1 digests were performed to assess the full occupancy of histone octamers at the 601 sites. ScaI digests were executed in the presence of increased DTT (10mM) to reduce Ub-SH from H2BK120C. This was done in order to simplify the analysis by native PAGE (Figure 5.10c). All nucleosome arrays showed a predominant MN band at the molecular weight of an unmodified H2B MN. Collectively, these analyses are indicative of well-formed, completely octamer-saturated, H2B, $H2B_{ss}Ub$, and $H2B_{ss}uLL$ nucleosome arrays.

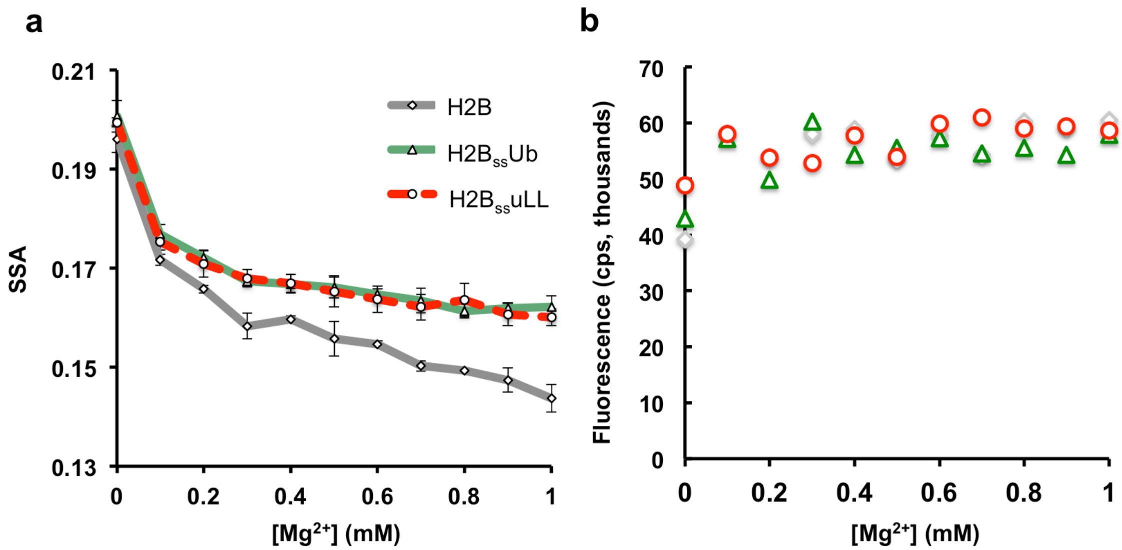


Figure 5.11. Mg^{2+}-induced compaction of H2B, $H2B_{ss}Ub$, and $H2B_{ss}uLL$ nucleosome arrays. (a) SSA as a function of Mg^{2+} for H2B (grey line), $H2B_{ss}Ub$ (red dashed line), and $H2B_{ss}uLL$ (green line) arrays. Error bars, s.e.m (n = 3). (b) Fluorescence emission of H2B (grey diamonds), $H2B_{ss}Ub$ (green triangles), and $H2B_{ss}uLL$ (red circles) arrasys as a function of added Mg^{2+} to the sample. Graph shows vertically polarized fluorescence emission of each sample excited with vertically polarized light.

Homo-FRET SSA was employed to elucidate the Mg^{2+}-induced chromatin compaction behavior of H2B, $H2B_{ss}Ub$ and $H2B_{ss}uLL$ nucleosome arrays (Figure 5.11a). As expected, the presence of $H2B_{ss}Ub$ impeded Mg^{2+} induced compaction of the 12-mer array compared to the H2B (i.e. unmodified) array (Figure 5.11a, green line compared to grey line). Remarkably, given the profound effect on hDot1L methyltransferase activity, mutation of the L71/73 surface had no impact on array compaction (Figure 5.11a, red dashed line). The $H2B_{ss}Ub$ and $H2B_{ss}uLL$ arrays behaved identically upon Mg^{2+} induced compaction, both

compacting less than H2B arrays (Figure 5.11a, green line compared toward dashed red line). As a control to ensure the arrays did not precipitate out of solution during the experiments, the fluorescence emission of the samples was also recorded (Figure 5.11b). All samples showed a similar absolute fluorescence, across all Mg^{2+} concentrations tested, ruling out any complicating effects such as precipitation (Figure 5.11b). This result indicates that the L71/L73 surface in Ub is not required for chromatin decompaction.

5.5. Summary and conclusions

In this chapter we investigated the structural ramifications of the Ub post-translational modification in the context of H2B-Ub chromatin. In the first part of the chapter, the solvent accessibility of H3K79C was probed in H2B and H2B-Ub MNs. We were unable to detect a difference in mPEG labeling leading to the conclusion that H2B-Ub does not drastically alter the local surface surrounding H3K79 in a way that leads to enhanced solvent accessibility. We next took a spectroscopic approach to investigate the Ub structure in the context of the MN. Surprisingly, we observed little difference in methyl-TROSY spectrum of Ub when attached to the MN surface at its canonical site on H2B as compared to the Ub free in solution. While this argues against an intimate interaction between Ub and the MN surface, we caution that these preliminary experiments employed a conjugation strategy (i.e. an UbG76C mutant instead of Ub-SH) different from that used in the biochemical experiments described in the preceding chapters. Although, previous studies in our laboratory indicate a significant tolerance in this

junction region with respect to hDot1L stimulation,[121, 152] it is conceivable that this new structure, particularly the presence of the carboxyl group, disrupts the normal stimulatory behavior of Ub when attached to H2B. Future studies will need to address this possibility. Lastly, we demonstrated that the L71/L73 hydrophobic patch on Ub is not required for the ability of H2B-Ub to disrupt chromatin compaction. Specifically, when incorporated into nucleosome arrays, the uLL mutant did not compact any different from that of Ub.

Collectively, this chapter suggests that when attached to position 120 of H2B, Ub does not dramatically alter the MN structure and, more specifically, that the L71/73 functional 'hotspot' is not involved in modulating higher-order chromatin structure. Since this 'hotspot' is clearly required for stimulation of hDot1L activity (Chapters 3 and 4), we speculate that Ub may be directly interacting with the enzyme itself. It is interesting to note, however, that hDot1L binds to mononucleosomes in both the presence and absence of H2B-Ub.[152] Indeed, ChIP-Seq data indicates that hDot1L does not co-localize with H2B-Ub containing chromatin, but rather is spread throughout euchromatin regions.[121, 130, 145] Thus, we suspect that hDot1L directly binds to H2B-Ub through the surface centered on L71/L73 and that this interaction activates the enzyme, either through a direct allosteric process and/or by repositioning the protein on the nucleosome surface such that a productive (for H3K79 methylation) enzyme-substrate complex is generated. This hypothesis will be discussed further in section 7.4.

Finally, this chapter indicates that Ub, in the context of H2B-Ub, is read out differentially in regard to H2B-Ub functions. As the L71/L73 surface on Ub does not contribute to the effect that H2B-Ub has on chromatin compaction, yet is responsible for the stimulation of hDot1L activity, we can conclude that different surface features of Ub are required for separate H2B-Ub functions. Because of this, we wondered whether L71/73 'hotspot' was specific only to the regulation of hDot1L or a general feature required for stimulation of all H3 methyltransferases regulated by H2B-Ub. This will be addressed in the following chapter through the analysis of the critical determinants in Ub required for both yDot1 and ySet1C methylation.

Chapter 6. Characterization of the yeast Dot1 and Set1 methyltransferases towards H2B-Ub *in vivo* and *in vitro*[3]

6.1. Introduction

The ubiquitylation of histone H2B is conserved from the budding yeast *Saccharomyces cerevisiae* (referred herein as ‘yeast’) to humans, as is its involvement in the regulation of both H3K4 and H3K79 methylation through the direct stimulation of the relevant H3 methyltransferases. In the previous chapters, we have shown that the hydrophobic patch comprising L71/73 of Ub is essential for stimulation of the hDot1L enzyme. This raises the question of whether this same motif is involved in stimulating the yeast version of the methyltransferase, yDot1. On the one hand this might seem likely given the high degree of homology within the catalytic domains of the two enzymes (29.6% identity). However, we note that the two enzymes differ markedly in size and organization outside the catalytic domain (see section 1.3.3.2). In particular, the location of the basic patch responsible for DNA-binding is inverted between the yeast and human enzymes and the yeast enzyme contains a putative ubiquitin interaction domain within its N-terminal region that is not present in hDot1L.[148] Thus, it was far from clear whether the molecular details of the stimulation by H2B-Ub would be the same for the yeast and human Dot1 homologs.[145, 148]

---

[3] The *in vivo* experiments were done in collaboration with Dr. Jung Ae Kim in Professor David Allis’s lab at Rockefeller University. Further, yDot1 and ySet1C constructs were prepared by Dr. Jaehoon Kim in Professor Robert Roeder’s lab at Rockefeller University

In addition to yDot1, we also wanted to see if the L71/73 hotspot on Ub was critical for H3K4 methylation. As H3K4 methylation in humans is catalyzed by multiple Set1/MLL family histone methyltransferases, which exist in multi-subunit complexes and play non-redundant, conditionally Ub-independent roles (see section 1.3.3.1 for a further discussion),[197] we sought to use ySet1C to investigate the effect of uLL on H3K4 methylation. *In vivo*, ySet1C has been shown to require amino acid residues in the C-terminus of H2B (H2BR119 and H2BT122), which are spatially adjacent to the H2B-Ub site (H2BK123 in yeast) and function in the recruitment of ySet1C to chromatin.[198] Conceivably, the proximity of L71/73 in Ub to these C-terminal H2B residues (which are solvent exposed on an α-helix) could form a composite recognition surface between H2B and Ub necessary for ySet1C-mediated H3K4 methylation.

In the following, we explore whether the L71/73 hotspot on Ub is important for H3 methyltransferase activity in yeast. We reasoned that studying the yDot1 methyltransferase would not only complement the work on the human enzyme but would also provide a potential avenue for *in vivo* studies, as yeast offers a more genetically tractable system than mammalian cells. Specifically, we were interested in the development of a genetic system to study the effect of the uLL mutations on yDot1-mediated H3K79me2 *in vivo*. Further, we were keen to investigate whether the same L71/L73 surface on Ub is required for the other histone methylation regulation by H2B-Ub, ySet1C-mediated H3K4 methylation.[86]

Surprisingly, this chapter concludes with the finding that L71/73 surface is critical for ySet1C activity but dispensable for yDot1 stimulation.

6.2. An *in vivo* system to investigate Ub surface features in H2B-Ub

We set out to devise a genetic system that would allow us to ask whether the L71/L73 hydrophobic hotspot on Ub is important for ubiquitin-methylation crosstalk *in vivo*. Ubiquitin is a highly conserved protein, abundant in cells and essential in a range of activities such as protein degradation, localization, and signaling.[53] An alanine scan of ubiquitin in yeast identified all non-viable ubiquitin mutants, including both L71 and L73 residues,[8] obviously complicating the design of a cell-based experiment. In principle, such lethality might be by-passed if the Ub mutant were selectively targeted to chromatin. In thinking of this possibility we were inspired by two observations, namely; (i) that Ub can be moved to position 20 on the N-terminal tail of histone H2A, which is juxtaposed to the native ubiquitylation site in H2B in the nucleosome structure, without disrupting Dot1 activation[152] and, (ii) that deletion of the first 20 residues of H2A in yeast does not abrogate H3K4 and H3K79 methylation.[199] We reasoned that an Ub fusion near the N-terminus of H2A might be able to recover yDot1 activity in yeast when the sole H2B-Ub E2 ligase, Rad6, is deleted and thus no native H2B-Ub or H3K79me2/3 is observed. Specifically, we chose to engineer an Ub-H2A(20-116) chimera to test if we could recover H3K79 methylation in a Rad6 deletion background in yeast (Figure 6.1b).

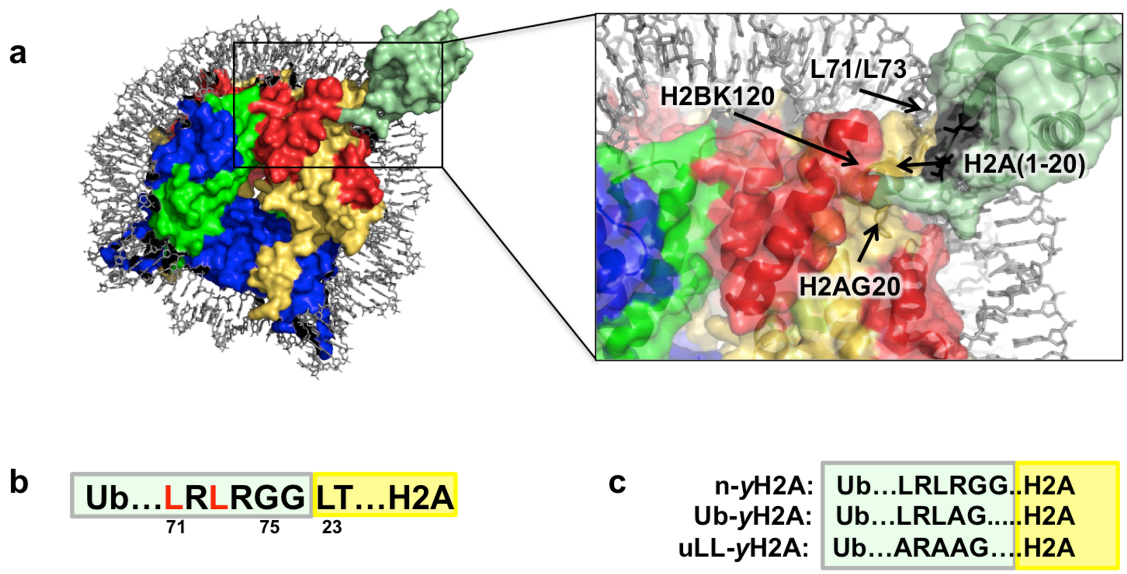


Figure 6.1. The positioning of H2BK120 Ub in the MN structure enables an N-terminal Ub fusion to H2A as a H2B-Ub mimic. (a) Structural model of the H2B-Ub MN, (this model is a composite of the Ub structure, PDB 1UBQ3, and the wild-type MN structure, PDB 3LZ0). Ub (light green), H2A (yellow), H2B (red), H3 (blue), H4 (green), are shown as both cartoon and surface representations. L71 and L73 of Ub are shown as black sticks, and H2BK120 and H2AG20 as spheres. Inset, a close up view of the Ub ligation junction. (b) The primary sequence of a linear Ub-H2A fusion surrounding the Ub-H2A fusion site. L71 and L73 are colored red. (c) All fusions employed in this chapter. Fusions are identical in amino acid sequence except for the residues depicted. Ub-*y*H2A was utlized to prevent de-ubiquitylation of the N-terminally fused Ub moitey.

Initial attempts to incorporate the N-terminally ubiquitylated H2A (n-*y*H2A) into chromatin were unsuccessful in that only a truncated version H2A was observed in histones extracted from yeast chromatin (Figure 6.2b lane 3). We reasoned that this is most likely due to the action of ubiquitin C-terminal hydrolases (or possibly DUBs), which are known to remove Ub from linear fusions in this case the Ub-H2A(20-116) fusion. To prevent this, we mutated the chimera to remove known hydrolase recognition elements in the C-terminus of Ub.[200] Specifically, we replaced R74 with alanine and deleted G75; note, the resulting fusion is referred to as simply Ub-yH2A (Figure 6.1c and Figure 6.2a). Further, as these amino acid residues were dispensable in our analysis of hDot1L-mediated H3K79 methylation, we expected them to have little impact on yDot1 activity. Analysis of yeast expressing Ub-*y*H2A indicated this strategy was successful in that an H2A-containing protein was detected by western blotting at the expected MW of the fusions (~18 kDa) (Figure 6.2a and b, lane 4). Critically, ΔRad6 yeast expressing this optimized Ub-*y*H2A construct exhibited partial recovery of H3K79 methylation (Figure 6.2a and b, lane 4), thereby creating a selective assay system to test the role of uLL *in vivo*. To our knowledge, this is the first instance of an *in vivo* system amenable to mutagenesis of Ub targeted to one specific ubiquitylation mark.

Having established this system, we incorporated L71A and L73A into the Ub-*y*H2A fusion (uLL-*y*H2A) and tested for H3K79 di-methylation (Figure 6.2b, lane 5). uLL-*y*H2A expressed at levels comparable to Ub-*y*H2A (Figure 6.2b, H2A

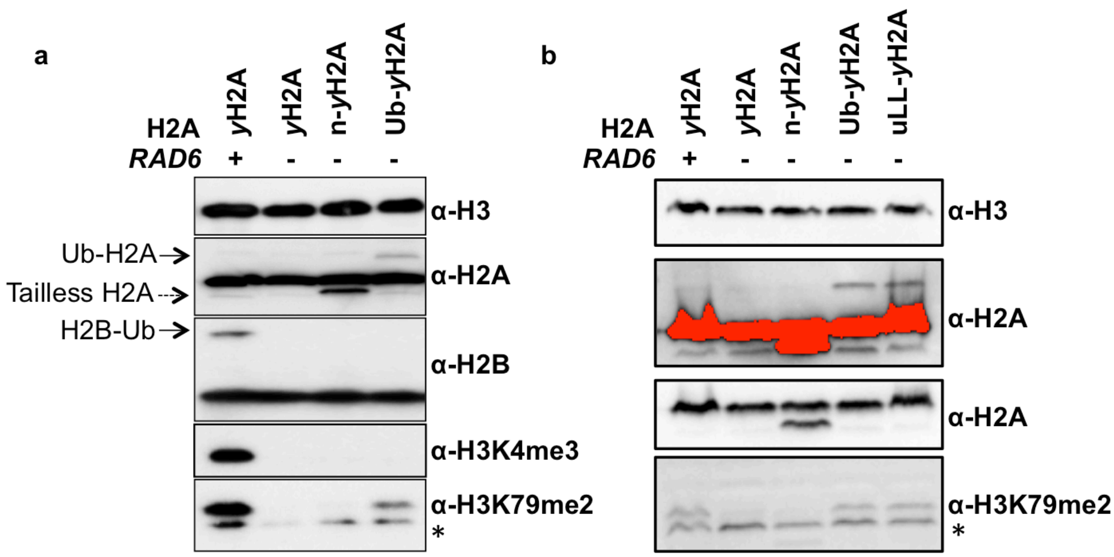


Figure 6.2. An *in vivo* system to test yDot1 activity. (a) The *in vivo* analysis of the effect of Ub fusion mutants on H3 methylation. Western blot analysis of indicated histones as well as the H3 histone modifications, H3K79me2 and H3K4me3, isolated from yeast lysates. Asterisk denotes a non-specfic band. (b) The *in vivo* analysis of the effect of an uLL-*y*H2A fusion mutant on H3K79me2. Western blot analysis of indicated histones and H3K79me2 isolated from yeast lysates. Asterisk denotes a non-specfic band.

upper panel lane 4 compared to lane 5). Surprisingly, given the *in vitro* studies employing hDot1L, H3K79me2 levels in yeast were rescued equally by the Ub-*y*H2A and uLL-*y*H2A (Figure 6.2b, lane 4 compared to lane 5). Thus, the Ub surface centered on L71/L73 appears not be involved the stimulation of yDot1 *in vivo*, at least in the context of our Ub-H2A fusion system.

Expression of Ub-*y*H2A in yeast was unable to recover H3K4me3 by ySet1C (Figure 6.2a, lane 4). This could be due to; (1) additional factors needed for efficient H3K4 methylation, (2) the inability for ySet1C to methylate H3K4 where H2A-Ub has been incorporated into chromatin at non-specific genomic locations, or (3) that the Ub-H2A fusion is not a substrate for ySet1C-mediated H3K4 methylation. *In vitro* experiments with ySet1C and Ub-H2A will be discussed further in section 6.5.

6.3. The L71/L73 surface in not involved in the stimulation of yDot1 activity *in vitro*

Surprised by the ability of the uLL-yH2A chimera to stimulate yDot1-mediated H3K79 methylation in yeast, we next asked whether these findings could be recapitulated in a biochemically-defined system. Accordingly, we set out to reconstitute MNs *in vitro* with the same H2A ubiquitin fusions employed *in vivo*. Unmodified yeast histones (yH2A, yH2B, yH3, and yH4) and histones containing Ub fusions (n-yH2A, Ub-yH2A, and uLL-yH2A) were expressed in *E. coli,* purified and incorporated into MNs using the appropriate 601 nucleosomal DNA. In parallel, a flag-tagged full-length yDot1 was affinity purified from Sf9 cells expressing this construct (Figure 6.3a). ^{3}H-SAM methyltransferase assays were performed on the various yeast MNs containing the various Ub fusions with both yDot1 and hDot1L (Figure 6.3 b and c). To gauge the integrity of the MNs after methylation, the reaction products were separated by native PAGE followed by Sybr Gold DNA staining and fluorography (Figure 6.3 b and c, top and middle

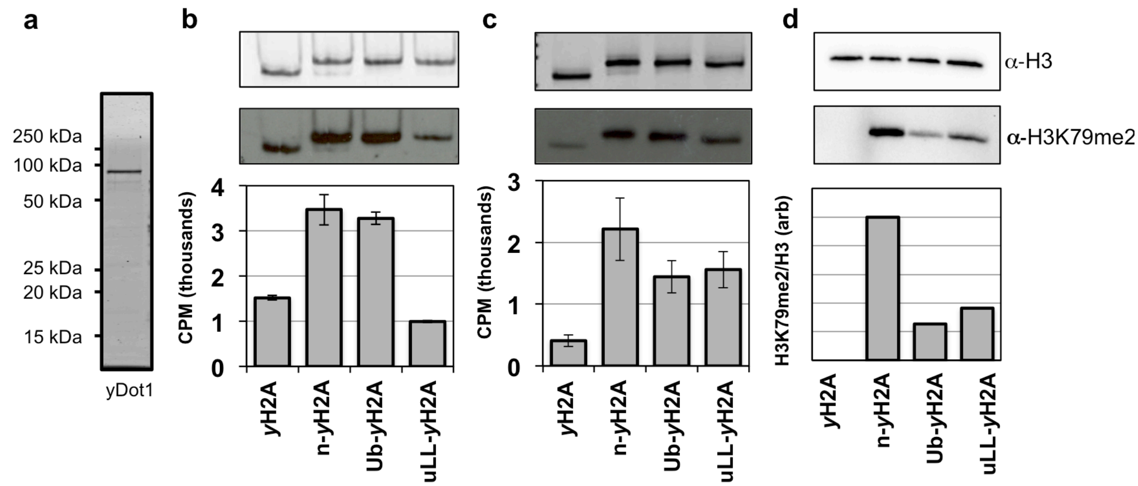


Figure 6.3. hDot1L and yDot1 methyltransferase assays on Ub-*y*H2A fusion containing MNs. (a) Coomassie stained SDS PAGE gel showing yDot1 purified from Sf9 cells. (b) ^{3}H-SAM hDot1L methyltransferase assays on yeast MNs containing yH2A, n-yH2A, Ub-yH2A and uLL-yH2A histones. MNs were visualized by native PAGE followed by Sybr gold staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting (bottom panel). Error bars, s.e.m (n = 3). (c) ^{3}H-SAM yDot1 methyltransferase assays on yeast MNs containing *y*H2A, n-*y*H2A, Ub-*y*H2A and uLL-*y*H2A histones. MNs were visualized by native PAGE followed by Sybr gold staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting (bottom panel). Error bars, s.e.m (n = 3). (d) H3K79me2 western blot analysis of 'cold' SAM yDot1 methyltransferase assays on yeast MNs containing *y*H2A, n-*y*H2A, Ub-*y*H2A and uLL-yH2A histones. MNs were seperated via SDS-PAGE followed by transfer to a PVDF membrane. H3 was blotted to ensure equal loading of MNs (top panel), and H3K79me2 was probed using a H3K79me2 specfic antibody (middle panel). The ratio of H3K79me2 to H3 was quantificated by densitometry (bottom panel).

panels). The extent of bulk methylation was assessed by filter binding assays followed by scintillation counting (Figure 6.3 b and c, lower panel). hDot1L methyltransferase activity was stimulated by both n-yH2A and Ub-yH2A fusions MNs compared to *y*H2A MNs (Figure 6.3b, lanes 2 and 3 compared to lane 1). Consistent with previous SAR studies, uLL-*y*H2A MNs were unable to stimulate hDot1L activity to the same degree as Ub-*y*H2A MNs (Figure 6.3b, lane 3 compared to lane 4). In contrast, yDot1 methyltransferase activity was observed for MNs containing all Ub fusion histones (n-*y*H2A, Ub-*y*H2A and uLL-*y*H2A) (Figure 6.3c, lanes 2, 3 and 4). To further confirm that yDot1 activity on uLL-*y*H2A MNs was comparable to Ub-*y*H2A MNs, yDot1 methyltransferase assays were performed using cold SAM and the extent of H3K79me2 incorporation was probed via western blot analysis (Figure 6.3d). Consistent with both the *in vivo* and ^{3}H-SAM methyltransferase assays, a comparable degree of H3K79me2 was detected in MNs containing Ub-*y*H2A and uLL-*y*H2A MNs (Figure 6.3c and d, lanes 3 and 4). The difference in the degree of stimulation of yDot1 activity between n-*y*H2A and Ub-*y*H2A in the ^{3}H-SAM methyltransferase assay versus western blot analysis is most likely a result of the read-out (Figure 6.3, c versus d): the radioactive experiment follows bulk methylation, i.e. all three H3K79 methylation states (H3K79me1/2/3), whereas the Western blot analysis allows for the detection of H3K79me2.

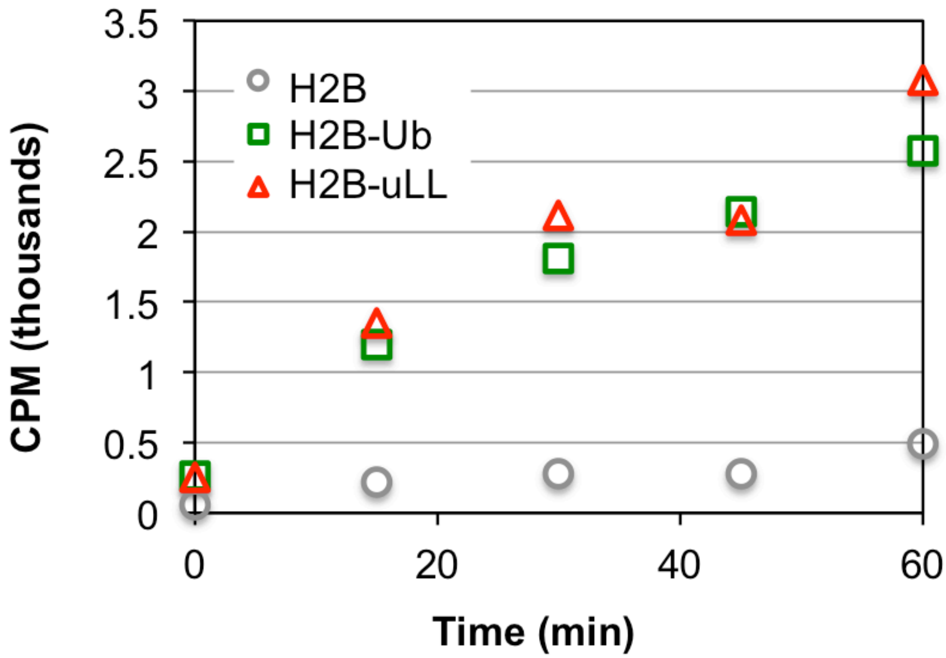


Figure 6.4. A time course analysis of yDot1 activity toward *x*H2B, *x*H2B-Ub, and *x*H2B-uLL MNs. ^{3}H-SAM yDot1 methyltransferase assays were performed on *x*H2B, *x*H2B-Ub, and *x*H2B-uLL MNs over the course of one hour. Quantification of the extent of methylation was performed by filter binding assays followed by liquid scintillation counting.

Lastly, to further show that the L71/L73 surface on Ub was not critical for yDot1 stimulation, we tested yDot1 activity on EPL-generated *x*H2B-uLL MNs, which were used to probe hDot1L activity in Chapter 4 (Figure 6.4). ^{3}H-methyl incorporation into MNs mediated by yDot1 was followed in a time-dependent fashion using *x*H2B, *x*H2B-Ub and *x*H2B-uLL MNs substrates (Figure 6.4). Quantification was performed by scintillation counting as described (Figure 6.4). Within the time course tested, the yDot1 activity toward *x*H2B-Ub MNs mirrored

that of *x*H2B-uLL MNs, demonstrating that the L71/L73 surface is not necessary for stimulation of yDot1 methyltransferase activity (Figure 6.4, green squares compared to red triangles at all time points). Remarkably, although H2B-Ub stimulates yDot1 and hDot1L activity, these methyltransferases diverge in their requirement of the L71/L73 amino acid residues within Ub to function.

6.4. A Ub alanine scan reveals mutants centered on the canonical hydrophobic patch of Ub critical for yDot1 stimulation

Having confirmed that the Ub surface critical for H3K79 methylation is not conserved between yDot1 and hDot1L, we wanted to test the activity of yDot1 toward the Ub-SH act-MN library to identify the surface features of Ub governing yDot1-mediated H3K79 methylation. A library of Ub-SH act-MNs (using *xenopus laevis* histones) was prepared as previously described in Chapter 3. However, we did not test the Ub12 mutant, and instead opted to investigate uLL instead as uLL-SH ligated more efficiently to the act-MN. Ub_{mut}-SH ligations to act-MNs were visualized by native PAGE followed by Sybr Gold DNA staining, and quantified using densitometry (Figure 6.5). All Ub_{mut}-SH ligations proceeded to a comparable extent as Ub-SH with the exception of Ub2-SH, which resulted in a reduced absolute abundance relative to Ub-SH (Figure 6.5, lower panel lane 4). The lower ligation level observed for Ub2-SH is consistent with previous ligation results (see section 3.2). ^{3}H-SAM yDot1 methyltransferase assays were performed on all MNs (Figure 6.5) The assay results were adjusted to take into

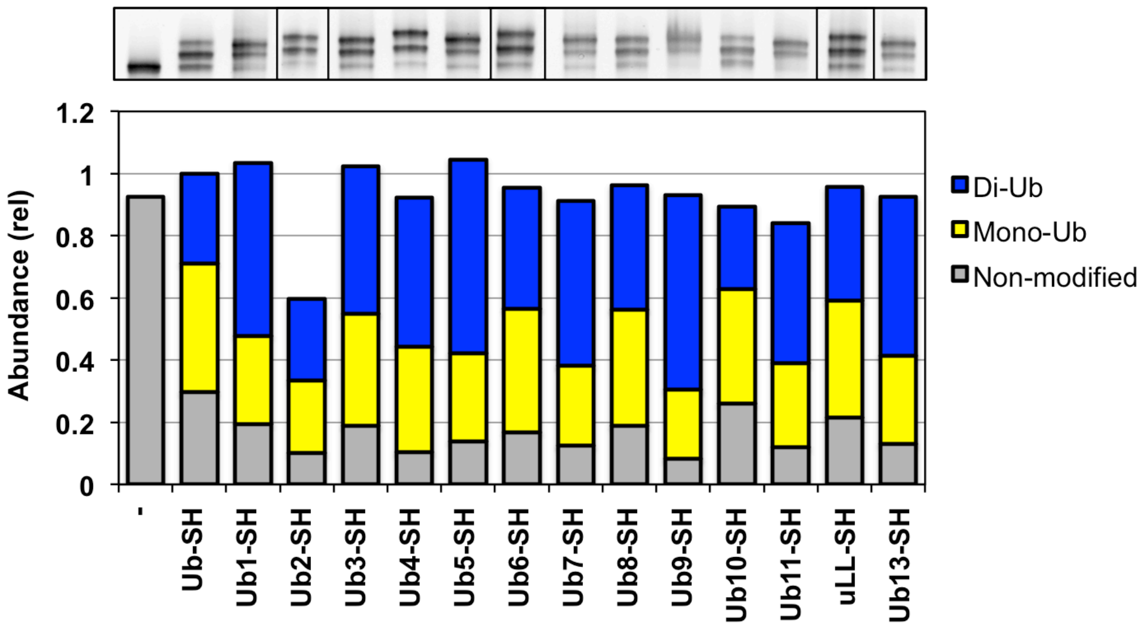


Figure 6.5. Analysis of Ub-SH and Ub$_{mut}$-SH ligation efficiency to act-MNs. MNs were visualized by native PAGE followed by ethidium bromide DNA staining (top panel). Quantification of the on-nucleosome ligation was performed by densitometry and adjusted relative to Ub-SH act-MNs. Non (grey box), mono (yellow box), and di ubiquitylated MNs (blue box) are plotted for each Ub$_{mut}$-SH.

account the relative abundance of Ub$_{mut}$-SH act-MNs and Ub$_{stim}$ was calculated as in section 3.3 (Figure 6.5a). Interestingly, no single Ub mutant completely abrogated yDot1 activity – this should be contrasted with the results for hDot1L where Ub12 (represented in this experiment by uLL) had a profound effect on methylation activity. Rather, multiple Ub$_{mut}$-SH act-MNs partially disrupted yDot1L activity, namely; Ub1, Ub3, Ub7, and Ub13 (Figure 6.6a).

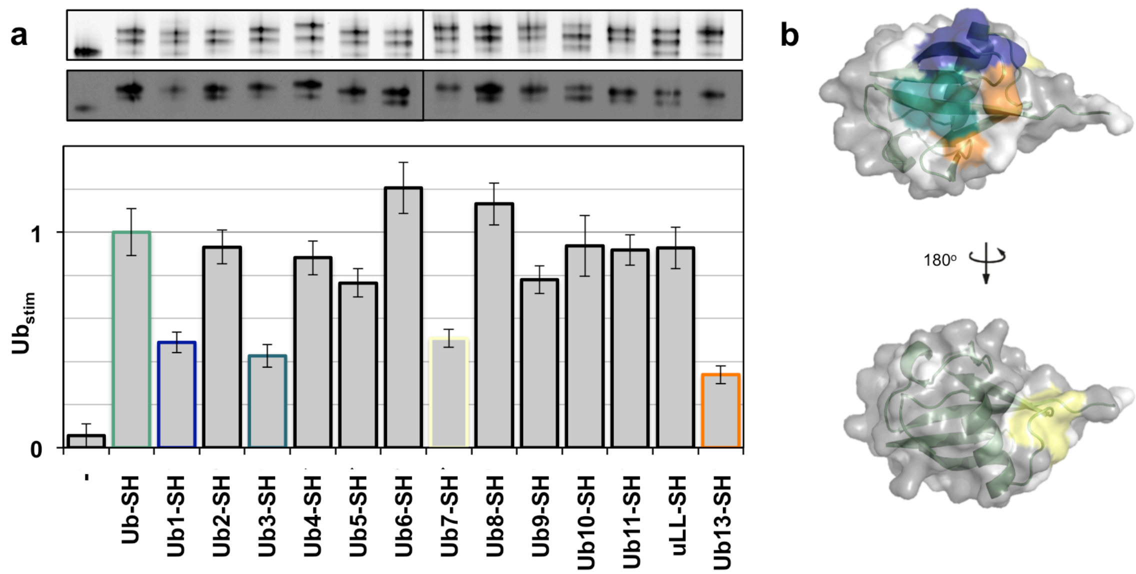


Figure 6.6. Surface features on ubiquitin critical for yDot1 stimulation. (a) yDot1 activity with respect to each Ub_{mut} 1-13 (note that uLL has been subsituted for Ub12). MNs were visualized by native-PAGE followed by Sybr Gold DNA staining (top panel), and ^{3}H-methyl incorporation was probed by fluorography (middle panel). Quantification of methylation was performed by filter binding assays followed by liquid scintillation counting and adjusted to include the extent of Ub-SH ligation, termed Ub_{stim} (bottom panel, see section 3.3). Error bars, s.e.m (n = 3). Ub1 (dark blue box), Ub3 (turquoise box), Ub7 (yellow box) and Ub13 (orange box) mutants led to a partial reduction in yDot1 stimulation. (b) Summary of alanine scanning results. Residues colored in grey did not have an effect on hDot1L stimulation, those colored in white were not tested. Ub1 (dark blue), Ub3 (turquoise), Ub7 (yellow) and Ub13 (orange) are shown. With the exception of Ub7, these mutations are centered around the canional hydrophobic patch on Ub (Ub13).

The only Ub_{mut} that partially reduced stimulation of both hDot1L and yDot1 activity was Ub7 (Figure 6.6a, lane 9). Interestingly, the mutations in Ub1, Ub3, and Ub13 constitute a contiguous surface on Ub (Figure 6.6b). Importantly, this surface is well known to play a role in Ub-protein interactions (i.e. the canonical

hydrophobic Ub surface discussed in section 3.1). Thus, an important avenue for future experiments will be to further validate this finding through additional biochemical experiments (e.g. using fully reconstituted mutant H2B-Ub MNs) as well as *in vivo* studies using the yeast Ub-H2A chimera system described earlier in this chapter.

6.5. The L71/L73 surface of Ub is required for ySet1C stimulation

We next investigated the broader role of the L71/L73 surface patch on Ub by focusing on H3K4 methylation by ySet1C *in vitro*. Our decision to use the 8 subunit yeast complex was driven by several considerations the most important of which was its availability in highly pure form from an insect cell over-expression system. Indeed, this recombinant system has proven extremely powerful for dissecting aspects of the H2B-Ub crosstalk with H3K4me – in particular the role of the various ySet1C subunits in sensing H2B-Ub.[201] We first wanted to test whether yeast MNs containing Ub-*y*H2A were able to upregulate ySet1C methyltransferase activity *in vitro*, particularly since the Ub-*y*H2A chimera was unable to stimulate H3K4 methylation *in vivo*. ySet1C was affinity purified from Sf9 cells co-infected with baculoviruses that express FLAG-Set1 and the seven other untagged ySet1C subunits (Figure 6.7a). We then tested whether

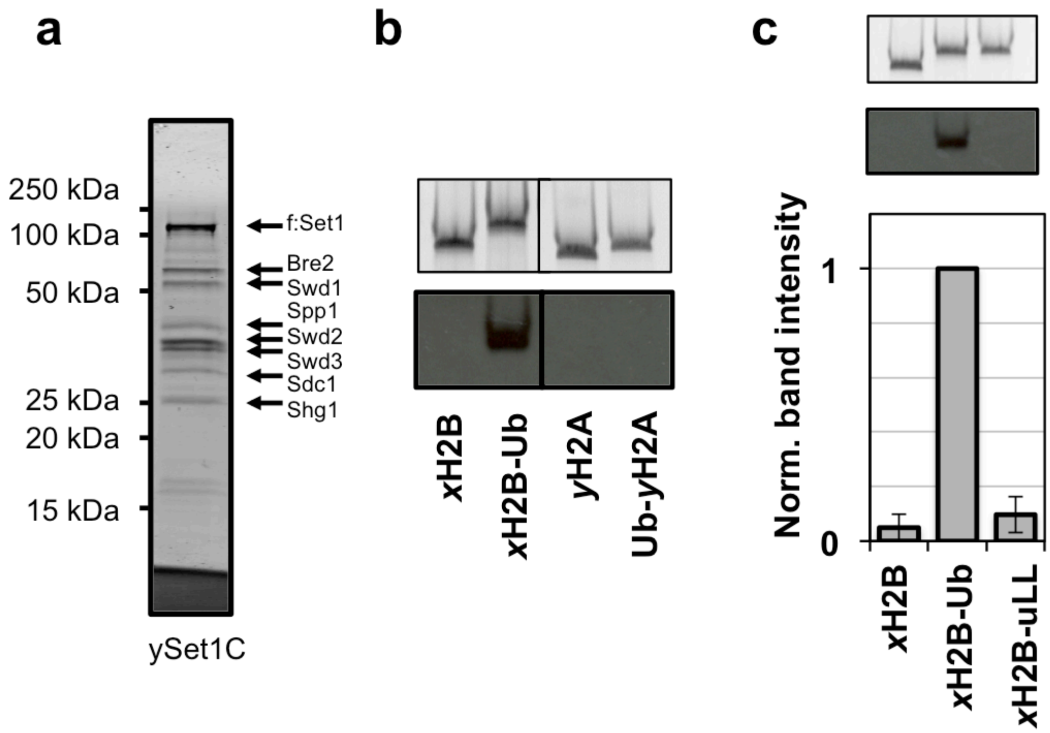


Figure 6.7. ySet1C methyltransferase assays with Ub-yH2A and H2B-Ub MNs. (a) Coomassie stained SDS PAGE gel showing ySet1C purified from Sf9 cells. All subunits are labeled. (b) ^{3}H-SAM ySet1C methyltransferase assays were performed on *x*H2B-Ub and Ub-*y*H2A containing MNs (DNA visualized by Sybr Gold, top panel) using ySet1C. ^{3}H-methyl incorporation was detected by fluorography (bottom panel). (c) ^{3}H-SAM methyltransferase assays were performed on *x*H2B, *x*H2B-Ub and *x*H2B-uLL MNs (DNA visualized by Sybr Gold, top panel) using ySet1C. ^{3}H-methyl incorporation was detected by fluorography (middle panel) and quantified using densitometry (bottom panel). Error bars, s.e.m (n = 3).

Ub-yH2A MNs could stimulate ySet1C activity in a ^{3}H-SAM methyltransferase assay using *x*H2B-Ub MNs as a positive control (Figure 6.7b). Consistent with previous studies[86, 201], unmodified MNs were unable to stimulate ySet1C methyltransferase activity, whereas *x*H2B-Ub containing MNs supported robust methylation (Figure 6.7b, lanes 1 and lane 3 compared to 2). No detectable stimulation of ySet1C was observed using Ub-yH2A MNs under the same conditions (Figure 6.7b, lane 2 compared to lane 4). These results explain why we were unable to rescue H3K4me3 using the Ub-yH2A fusion *in vivo* (Figure 6.2), i.e. the Ub-H2A fusion does not seem to be recognized by ySet1C as an H2B-Ub mimic, whereas this is the case for yDot1.

We next wanted to test whether the L71/L73 surface on Ub was required for the stimulation of ySet1C activity, analogous to hDot1L. ^{3}H-SAM methyltransferase assays were performed on *x*H2B, *x*H2B-Ub, and *x*H2B-uLL MNs (Figure 6.7c). As before, ySet1C was stimulated by xH2B-Ub MNs and showed no activity towards *x*H2B MNs **(**Figure 6.7c, lane 1 compared to lane 2). Surprisingly, and analogous to hDot1L, ySet1C was not stimulated by *x*H2B-uLL MNs (Figure 6.7c, lane 3). Thus, the requirement of UbL71/L73 in H3 methyltransferases activity is conserved between ySet1C and hDot1L, highlighting the importance of this functional hotspot for both ySet1C-mediated H3K4 and hDot1L-mediated H3K79 methylation.

6.6. Summary and conclusions

As H2B-Ub is conserved between yeast and humans and is a prerequisite in both organisms for H3K4 and H3K79 methylations, we explored the surface features of Ub in the H2B-Ub stimulation of yDot1 and ySet1C activity. We designed an *in vivo* system to test the affect the uLL mutant had on yDot1 in yeast. This was done through the generation of an N-terminally fused Ub to a truncated H2A, creating a genetically tractable construct that mimicked the MNs positioning of H2B-Ub. Using this system we were able to partially recover H3K79me2 methylation and elucidate that yDot1 activity was stimulated by Ub-yH2A and uLL-yH2A similarly in an *in vivo* context. Unfortunately, we did not detect H3K4 methylation leading us to conclude, and later confirm *in vitro*, that Ub-yH2A does not upregulate ySet1C activity.

This most surprising finding in this chapter is that the yeast Dot1 enzyme was unaffected by mutation of the L71/L73 surface on Ub, both *in vivo* and *in vitro*. Rather our preliminary SAR data indicates that the yeast enzyme senses a different surface of Ub, one that localizes to the canonical hydrophobic patch on the protein (i.e. centered on residues L8 and I44) that is known to engage a wide-range of target proteins through typically helical ubiquitin interaction domains.[53, 184] As demonstrated in Chapter 3, the hDot1L does not require the canonical hydrophobic patch on Ub for stimulation, but rather senses L71/L73. Thus, there is a clear divergence in how the yeast and human versions of this

methyltransferase readout H2B-Ub. This remarkable finding speaks to information-rich nature of the Ub protein itself; the large size of the modification (say compared to methylation or acetylation) means that it can engage *trans*-acting factors through different non-overlapping surfaces, perhaps even mediating multiple biochemical processes simultaneously. With respect to this, we note that the L71/L73 surface is not required for H2B-Ub mediated structural changes on the chromatin fiber (see Chapter 5) – in other words H2B-Ub appears to be able to simultaneously decompact chromatin and stimulate hDot1L.[117]

We can only speculate as to why the yeast and human versions of Dot1 sense different surfaces of Ub, despite the fact that both enzymes share a highly homologous catalytic core. As noted at the beginning of this chapter, the overall size and domain organization of the enzymes differ markedly. The human enzyme contains 1739 amino acids, as opposed to 582 for the yeast enzyme. The former contains a DNA-binding motif on the C-terminal side of the catalytic domain, whereas the latter contains this motif N-terminal the catalytic domain. Thus, there may be geometric differences in the way these two enzymes engage the nucleosome. Further to this point, biochemical studies have identified a short ubiquitin-binding domain in yDot1 that in not conserved in the human enzyme. It is likely relevant to consider that histone ubiquitylations underwent a rapid expansion from yeast to humans, with only 2 (H2BK123-Ub, and H3-Ub) in yeast, to 36 Ub sites annotated across the 4 histones in mammalian cells.[58] Thus, the

business of discriminating between stimulatory and non-stimulatory histone ubiquitylation sites would seem to be a much more challenging task for hDot1L than its yeast counterpart. Conceivably, this readout may be more selective when the L71/L73 surface of Ub is employed due its spatial proximity to the histone attachment site – this would allow for a composite recognition epitope involving both Ub and the local histone surface. Additional biochemical analyses will be needed to test this idea.

In the last part of this chapter, we asked whether the L71/L73 surface patch on Ub is required for stimulation of the ySet1C H4K4 methyltransferase. We found that Ub fused to an N-terminally truncated version of H2A (Δ1-20) was unable to stimulate H3K4 methylation. Previously studies have shown that the H2A tail is dispensable for H2B-Ub stimulation of ySet1C[199]. Thus, it is likely that the inability of Ub-yH2A to activate ySet1C results from this being an imperfect mimic of H2B-Ub, presumably this enzyme complex is intolerant of the changes in positioning and attachment site associated with the Ub-H2A fusion. Indeed, it is known ySet1C-mediated H3K4 methylation is sensitive to mutation of residues immediately around the Ub attachment site on H2B; a H2B-R119A/T122D construct reduced global levels of chromatin bound ySet1C *in vivo* even in the presence of high levels of H2B-Ub.[198] This finding combined with the results of this chapter, that ySet1C requires UbL71/L73 for H3K4 methylation, suggests that a composite surface between H2B and Ub is required for ySet1C function.

Chapter 7. Discussion and outlook

7.1. Introduction

The ubiquitylation of H2B at lysine 120 is intimately involved in transcription elongation, acting as both a positive regulator of H3K4 and H3K79 methylation, as well as contributing to the maintenance of a less compact local chromatin state.[64, 86, 117, 118] This is achieved through specific properties of the ubiquitin modification, and is not simply due to the steric bulk of adding an 8.5 kDa protein to the nucleosomal surface. The work presented in this thesis describes the biochemical characterization of the surface features of Ub that are sensed by the enzymes responsible for H3K4 and H3K79 methylation, Set1 and Dot1, respectively. Additionally, the structural implication of H2B-Ub was assessed and provided insight into H2B-Ub chromatin structure in regard to H2B-Ub function. Collectively, this work is indicative of Ub as an information rich modification able to orchestrate distinct biochemical functions on chromatin by using non-overlapping surface epitopes. A synthesis of the impact of this work in regard to H2B-Ub function and a conceptual framework for the advancement of histone PTM biology will be addressed below.

7.2. The post-translational modification of pre-assembled nucleosomes.

Several methods for the preparation of Ub substrates have been described, making the biochemical reconstitution of homogenous ubiquitylated histone containing chromatin accessible.[118, 121, 152] A caveat to these existing methods is

the need to prepare H2B-Ub prior to its incorporation into chromatin, resulting in a multi-step process that is both time-consuming and sample intensive. Herein, we greatly simplify the process by showing that Ub can be chemically introduced in a site-specific fashion to a pre-assembled MN. This approach involves an asymmetric disulfide formation reaction, compatible with nucleosomal reconstitution protocols, and achieves regioselectivity by exploiting the ability to introduce a unique cysteine residue into recombinant chromatin. As this method relies on the late stage diversification of the ubiquitin substrate, systematic structure-activity analysis of the system becomes tractable, as demonstrated in Chapter 2.

This approach will be of use to study other histone ubiquitylations and their associated functions in a high throughput manner. Moreover, through combination with the ubiquitin alanine mutant library from Chapter 3, the elucidation of SARs in other ubiquitylated histone contexts becomes feasible. Additionally, the disulfide attachment chemistry employed has the attractive feature that the modification can be removed through simple reduction, thereby offering the potential to study, in a time-resolved way, the functional and structural consequences of PTM removal.

7.2.1. *The application of cysteine chemistry to other histone PTMs*

It may be possible to adapt our on-nucleosomal ligation strategy to give non-reducible thioether linkages through use of alternate thiol-directed chemistries,

such as the conversion of cysteine into dehydroalanine or via thiol-ene chemistry.[202, 203] Alternatively, analogs of other PTMs can, in principle, be introduced chemically into pre-assembled chromatin using known cysteine derivatization routes, for example, lysine methylation,[138] and acetylation.[176] Conceivably, other chemical functionalities utilized in bio-orthogonal reactions (e.g. an azide functionality) could be used instead of cysteine, thereby further expanding the palette of chemistries that can be exploited to install PTMs into pre-assembled chromatin.[204] This could be achieved through the preparation of histones with bio-orthogonal amino acid residues via unnatural amino acid mutagenesis and EPL.[159, 172] Additionally, thiol-directed chemistries could be used in combination with these chemistries, enabling the chemical multi-modification of pre-assembled chromatin.

7.2.2. *The application of split inteins to histone PTMs*

Alternatively, split inteins could be utilized for the incorporation of histone PTMs into a pre-assembled MN. Split inteins are naturally occurring inteins that are split into two fragments (N-intein and C-intein) and direct protein splicing *in trans*.[165] This results in the polypeptide N-terminal to the N-intein being spliced to the polypeptide C-terminal to the C-intein (Figure 7.1a). Histone peptides, containing PTMs, can be synthesized through SPPS, and subsequently incorporated into either the N- or C- intein fragment construct using EPL.[164] As discussed in section 1.4.1, the only ‘trace’ of ligation in the spliced product is the cysteine

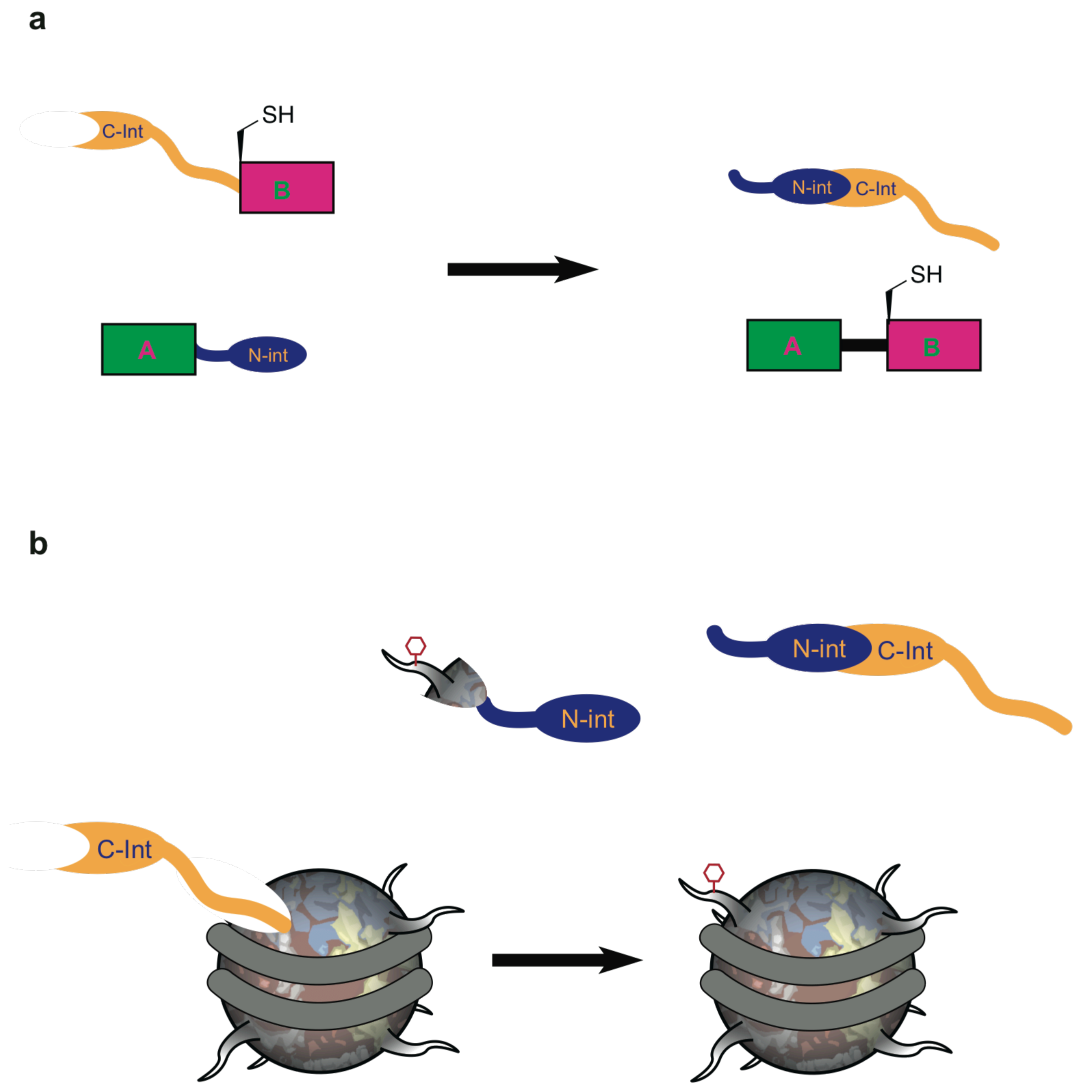


Figure 7.1. The application of split inteins to generate histone PTMs within the nucleosome (a) Split inteins are naturally occuring inteins that splice proteins *in trans*. Fragment A (green box) is ligated to Fragment B (pink box) through *trans* splicing subsequent to the association of the N-intein (labled N-Int) and C-intein (labeled C-Int). Trans splicing requires a cysteine which remains in the final A-B construct (right panel). (b) Through the synthesis of a PTM containing histone peptide construct N-terminal to the N-intein fragment, histone PTMs can be introduced into MNs containing a C-intein histone fragment.

necessary for ligation. This is a requirement for split intein function and we suspect, based on mutations made in nucleosomes in the preceding chapters, that the introduction of a cysteine mutation into the nucleosome will not significantly impact nucleosome structure or function. However, it may also be possible to desulfurize pre-assembled nucleosomes by adapting currently employed histone desulfurization methods to be amendable to non-denaturing conditions.[121]

We envision, much like the ubiquitylation of activated MNs, split inteins could be utilized to install PTMs into pre-assembled chromatin containing histones fusions to either an N or C-intein fragment. The advantage of this approach would be to expedite the incorporation of histone PTMs into the MN, as histone PTM peptides would need only to be ligated to either an N- or C-intein fragment, which then could be directly incorporated into chromatin. Moreover, with the newly discovered ultrafast split inteins these reactions could be performed at extremely low reactant concentrations (the intein fragments have picomolar affinity) and would be complete in a matter of seconds.[165] This would enable the streamlined synthesis of modified MNs and substantially reduce the amount of histone PTM peptide needed (typically yields from octamer reconstitution to nucleosome formation are ~10-20%)[180]. This approach could be utilized to titrate the amount of a histone PTM (or modification state, e.g. me1, me2, and me3) into chromatin, through combination with differentially modified histone tail-intein fragments. Adding an additional layer of complexity, orthogonal split intein fragments could

be incorporated into chromatin, creating a template for the generation of multi-PTM containing MNs.[205]

The use of split inteins may additionally facilitate the investigation of histone PTMs *in vivo*, as the modulation of enzyme activity or mutation of histone PTM sites often have multiple biological outcomes (i.e. chromatin modifiers often have multiple targets and histone modification sites often are modified with multiple types of PTMs).[47] Our lab has recently begun to pursue *in vivo* chromatin modification with engineered split inteins applied to the study of H2B-Ub in mammalian cells.[4] Specifically, this was achieved through the in vivo expression of an H2B(1-116)-N-intein fragment and the subsequent addition of a complementary K120 ubiquitylated C-intein-H2B(117-125)-Ub fragment. Through this technique, we envision that SAR studies, much like the *in vitro* mutagenesis of the Ub surface, can be elucidated in an *in vivo* context similar to the techniques introduced in Chapter 6. Further, the use of split inteins *in vivo* enables the incorporation of biophysical probes and bio-orthogonal handles into native chromatin which will be of great use in a variety of applications such as monitoring chromatin dynamics.[164]

---

[4] This is the work of Dr. Miquel Vila-Perello and Dr. Yael David Shternberg in Professor Tom Muir's lab at Princeton University.

7.2.3. *The post-translational modification of histones within pre-assembled MNs in high throughput systems*

The post-translational modification of pre-assembled MNs, activated for ubiquitylation or through the establishment of methodologies previously discussed (collectively referred to as Post-translationally modified after reconstitution (PAR-)), may be of use in combination with methodologies that take advantage of high-throughput DNA sequencing and mass spectrometry. The Muir laboratory has recently developed a high throughput chromatin biochemistry platform based on the generation and subsequent ChIP-Seq analysis of DNA-barcoded MN libraries containing user-defined DNA and/or histone modifications.[5] Using multiplexed next generation sequencing, it possible to perform thousands of biochemical assays in parallel and deduce, in a quantitative fashion, the binding preferences or substrates preferences of chromatin modifiers. We envision that encoded PAR-MNs could add to the flexibility of this approach. In this strategy, a library of PAR-MNs (containing either single cysteine mutations or split intein fusions) would be uniquely encoded. Depending on the type of library needed, these PAR-MNs could be diversified and pooled. The advantage of this would be that the library could be tailored to the biochemical question and easily adapted, which is not possible with current methodologies. For instance, encoded PAR-MNs could be

[5] This is the work of Dr. Uyen Nguyen in Professor Tom Muir's lab at Princeton University.

ubiquitylated, with either wild-type Ub or Ub mutants, and high-throughput assays could be performed as an alternative way to test both enzyme activity, i.e. pull downs directed towards a chromatin modification (e.g. H3K79me2), or relative binding affinity (e.g. hDot1L), i.e. pull downs directed towards the chromatin binder. Accordingly, PAR-MNs have the potential to greatly expedite powerful approaches to study histone PTMs in well-defined systems.

7.2.4. *The application of PAR-MNs to investigate 'H2B-Ub-like' ubiquitylations*

'H2B-Ub-like' ubiquitylations that can recapitulate specific H2B-Ub functions have been identified and suggested. [22, 157, 206] Specifically, H2BK34 ubiquitylation is able to stimulate H3K4 and H3K79 methylation. This was observed both *in vivo* and *in vitro* in the absence of H2BK120-Ub.[157] Recently, this remarkable finding has been extended to H4K31 ubiquitylation.[206] As these ubiquitylation sites localize to distinct regions within the nucleosome, it is unclear how these histone ubiquitylations perform similar functions (Figure 7.2). Further, the transcription factor NY-F has been structurally characterized in complex with DNA and shown to be homologues to a H2A/H2B dimer containing two histone folds through a tri-meric complex (comprised of NY-y, NY-b, and NY-a).[22] Interestingly, NY-F is ubiquitylated at K138 of the subunit NY-y, which corresponds to that of H2BK120 based on structural alignment (Figure 7.2).[22] Collectively, these studies suggest that, beyond H2B-Ub, both alternate histone ubiquitylation sites, and *trans*-acting

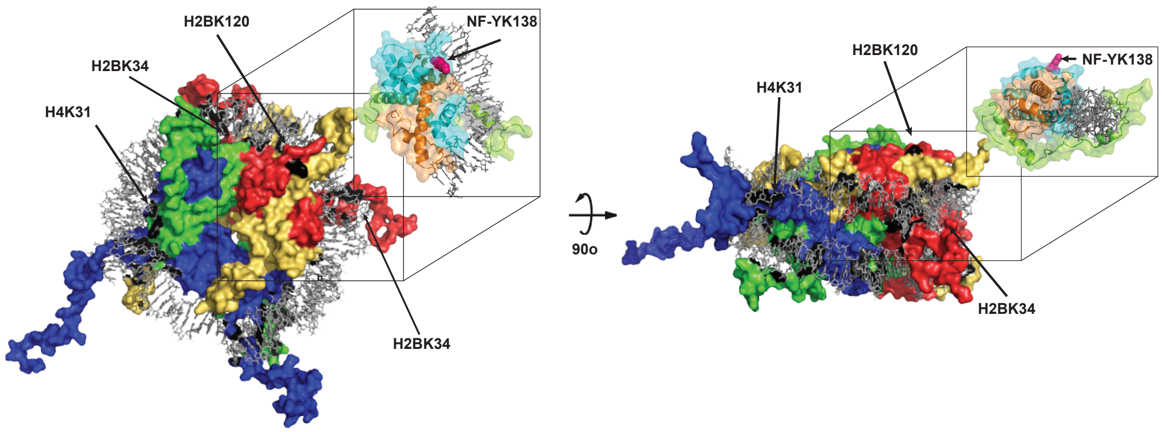


Figure 7.2. Ubiquitylations that recapitulate H2BK120 ubiquitylation functions. Multipe histone ubiquitylation sites (in black) and the transcription factor NY-F, a H2B-Ub PTM mimic, could potentially regulate H3K79 methylation similar to H2BK120-Ub. The MN is shown (PDB 1KX5), H2A (yellow), H2B (red), H3 (blue), H4 (green). All potential ubiquitylation sites that have been shown to regulate hDot1L are labeled. Inset: NY-F in complex with DNA (PDB 4AWL) is shown. The H2BK120 ubiquitylation site analog NY-y K138 is labeled.

chromatin factors may be functionally redundant for H2B-Ub associated functions. Further work is necessary to explore these ubiquitylations and the alanine mutant library employed in Chapter 3 could be employed to test whether the specificity of hDot1L for L71 and L73 is conserved between the different histone ubiquitylation sites.

Recent proteomic studies have identified numerous histone ubiquitylation sites (i.e. in addition to those noted above) whose epigenetic functions are completely unclear.[58] It would be interesting to see which, if any, of these new ubiquitylations

are capable of stimulating H3 methyltransferases (and potentially other chromatin modifiers). This could serve to focus *in vivo* studies on the relevance of such marks. Testing this would be feasible through the expression of histone cysteine mutants and incorporation into MNs using optimized MN reconstitution procedures. Furthermore, these MNs could be encoded and pooled to create a comprehensive library of ubiquitylated nucleosomes, where these nucleosomes could be studied in a competitive fashion. Specifically, methyltransferase assays in high-throughput systems could reveal whether hDot1L is stimulated equally by identified ubiquitylation sites.

7.3. H2BK120 ubiquitylation in chromatin structure

7.3.1. *H2B-Ub nucleosomal structure*

Understanding the intrinsic structural implications of H2B-Ub as a histone PTM is crucial to elucidate its associated processes. In Chapter 5 we took multiple approaches to determine what effect Ub had on chromatin structure and specifically tested whether any observed structural change was mediated by the L71/73 hotspot necessary for H3 methyltransferase regulation. Although much of this work is in its infancy, we have not been unable to detect a significant structural change in the nucleosome caused by Ub or uLL. We hypothesize that ubiquitin does not significantly affect the nucleosome structure and, furthermore, that hDot1L regulation by H2B-Ub is not dependent on a change in nucleosome conformation. This is based on: (i) that only minor perturbations of ILV residues

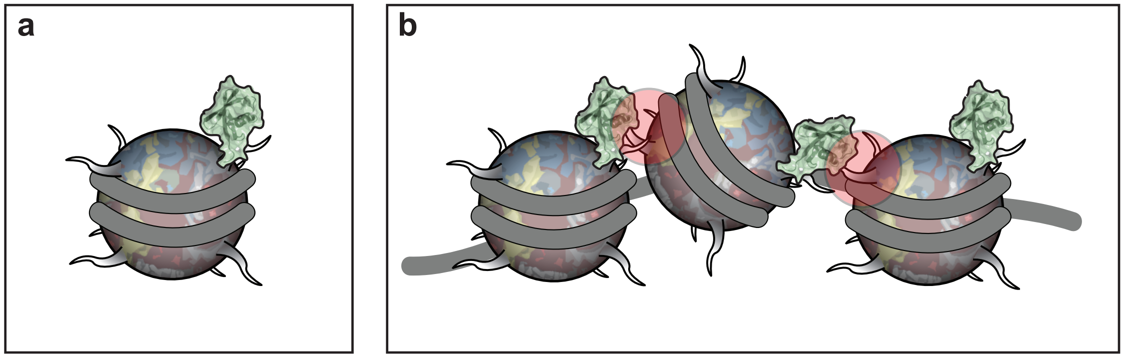


Figure 7.3. The contribution of ubiquitin to nucleosomal and array structure. (a) Little, if any, structural change is observed in the nculeosome upon H2B ubiquitylation. (b) In the context of nucleosomal arrays, ubiquitin could interact inter-nucleosomally to prevent compaction (red circles indicate inter-nucleosomal interactions).

within Ub are observed via methyl-TROSY in H2B-Ub nucleosomes; (ii) mPEG assays failed to detect a difference in H3K79C solvent accessibility; (iii) all 12 permissive alanine Ub mutants (which contain the majority of Ub surface residues) stimulate hDot1L-mediated H3K79 methylation, and; (iv) uLL, which does not stimulate hDot1L activity, was directly observed in the methyl-TROSY analysis. The hypothesis most consistent with the findings of this thesis is that Ub merely 'hangs off' of the H2BK120 surface and does not maintain any specific Ub-nucleosome contacts (Figure 7.3a). This is further substantiated as the effect of H2B-Ub on nucleosomal stability is modest[116] and that H2B-Ub functions are recapitulated by multiple distinct histone ubiquitylations, and thus the spatial orientation of Ub in respect to the nucleosome displays plasticity[152, 157, 206].

Future investigation of Ub in the context of the H2B-Ub nucleosome is warranted and is being further investigated using the methyl-TROSY approach presented in Chapter 5.[6] Additionally, paramagnetic relaxation enhancement (PRE) experiments in the context of the ILV labeled nucleosomes may be of use to obtain Ub-position sampling information on H2BK120-Ub nucleosomes. PRE experiments detect long-range distance information (~5-30nm) and thus a paramagnetic spin label attached to Ub could reveal how Ub is positioned within the nucleosome.[190, 207] Specifically, Ub-SH could be engineered to contain S-Methanethiosulfonylcysteaminyl-EDTA–Mn^{2+} (a paramagnetic spin label) distal to the Ub ligation junction. Proximity of this spin label on ubiquitin to different regions of the nucleosome, would reveal which, if any, specific orientation was favored as the presence of the paramagnetic spin label reduces the peak intensities in the nucleosome in a distance-dependent manner.[190]

7.3.2. *H2B-Ub in chromatin compaction*

Since the mutation of the L71/73 epitope has no impact on the ability of H2B-Ub to inhibit chromatin compaction, it is still unclear how this is achieved by ubiquitin. Conversely, as the compaction properties of a nucleosomal array is intimately intertwined with chromatin structure, we can infer that the structure of H2B-uLL containing arrays is similar to that of H2B-Ub arrays. However, as Hub1ylated nucleosomal arrays show no such alteration in array compaction properties

---

[6] This is the work of Dr. Julianne Kitveski in Professor Lewis Kay's lab at the University of Toronto.

(compared to unmodified arrays), the regulation of chromatin structure by H2B-Ub appears to be sequence specific. Collectively, as H2B-Ub has little, if any, structural effect on the nucleosome, a specific, i.e. not exclusively steric, H2B-Ub inter-nucleosomal interaction in the context of nucleosomal arrays could potentially mediate this effect (Figure 7.3b). The positioning of the H2BK120 site in the context of the tetra-nucleosome x-ray crystal structure and chromatin array models indicates that this may indeed be the case.[117, 208]

We propose that the library of alanine Ub mutants developed in Chapter 3, in combination of pre-assembled activated chromatin arrays could be employed to determine whether a discrete epitope on the Ub surface is responsible for the inhibition of array compaction. Specifically, H2BK120C could be incorporated in octamers, activated with DTNB, and incorporated into 12-mer nucleosomal arrays. Prior to array precipitation (a crucial step in the purification of nucleosomal arrays), these pre-assembled arrays could be ubiquitylated with Ub-SH and the library of Ub_{mut}-SHs. Homo-FRET assays followed by SDS-PAGE analysis of the arrays (to determine the extent of ubiquitylation) could be performed to normalize for the extent of ubiquitylation as it is additive in the context of chromatin compaction.[117] Further, if specific Ub mutants are found that do not alter the chromatin compaction behavior of these arrays, crosslinking studies could be pursued to identify the putative Ub-nucleosome array interaction.

We must be cautious as to attribute too much significance to results using Hub1, however, as Hub1ylated chromatin is neither native nor shown to function in an ‘H2B-Ub-like’ manner. More work to confirm that Hub1 is well folded and able to recapitulate the sterics of Ub, in the context of H2B-Ub chromatin, is needed. Given the inability of a Hub1-Ub chimera to stimulate hDot1L, it is also possible that Hub1 may not be a ‘passive’ Ub-like molecule in the context of the nucleosome but may intrinsically disrupt H2B-Ub processes through other undefined means.

7.4. H2B-Ub regulation of H3 methyltransferase activity

Our structure-activity studies establish a hydrophobic surface on ubiquitin, comprising L71 and L73, that is required for the stimulation of hDot1L-mediated H3K79 methylation as well as ySet1C-mediated H3K4 methylation. Surprisingly, these amino acid residues were not required for yDot1-mediated H3K79 methylation, and in Chapter 6, we were able to identify (although follow up studies are on going) that the canonical hydrophobic surface (L8/I44) on Ub promoted this methylation instead. Taken together, H2B-Ub is able to regulate these three H3 methyltransferases through either the L8/I44 or L71/L73 hydrophobic patches. As these patches are spatially adjacent in the Ub tertiary structure it is surprising –in the case of hDot1L and yDot1- that each patch regulates these enzymes separately (Figure 7.4). Taken together, as hDot1L is not homologous to ySet1C, but is homologous to yDot1, a conserved Ub

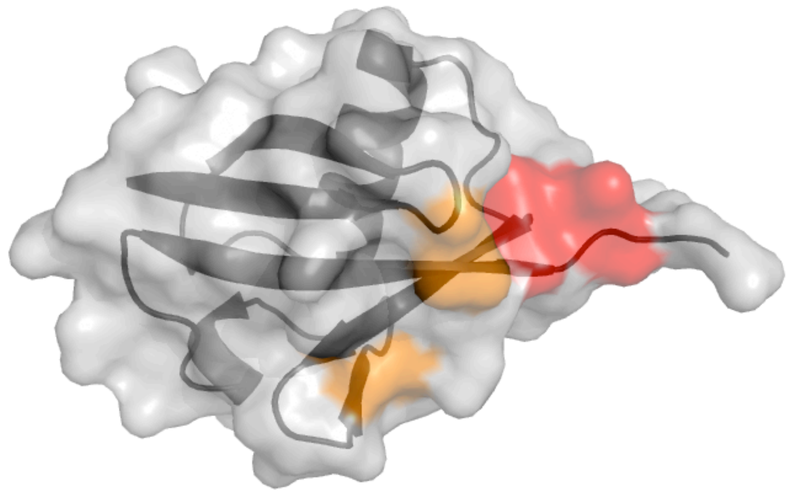


Figure 7.4. The proxmity of the L8/I44 and L71/L73 hydrophobic patches. Both L8/I44 (orange) and L71/L73 (red) hydrophobic patches are adjacent in the tertiary structure of ubiquitin.

recognition element or allosteric mechanism within these proteins is unlikely.[121, 201, 209]

7.4.1. *Insights into the mechanism of hDot1L regulation by H2B-Ub*

The three non-mutually exclusive hypotheses, in section 1.3.3.2, for the regulation of hDot1L activity by H2B-Ub were predicated upon the fact that this regulation occurs intra-nucleosomally (Figure 1.7).[118, 121] We can now readjust our hypothesis to integrate the information contained herein. As we have been unable to detect a significant structural change in the nucleosome upon H2B ubiquitylation (specifically in the L71/73 patch that is required hDot1L activity) we posit that hDot1L forms direct hDot1L-Ub contacts mediated by L71 and L73

within Ub (Figure 1.7 b and c). We note that that hDot1L binds to the nucleosome in the presence of absence of Ub.[152] Thus, the putative hDot1L-Ub interaction would either have to allosterically activate the enzyme and/or simply lead to a re-positioning of hDot1L on the nucleosome surface to facilitate K79 methylation.

Crosslinking studies to detect whether hDot1L interacts directly with the L71/73 surface are underway.[7] Photo-leucine, a diazirine containing cross-linking mimic of leucine, was incorporated in ubiquitin at position L73 and ligated to H2B via an asymmetric disulfide approach. This construct can further be incorporated into MNs, and then cross-linking can be carried out in the presence of hDot1L upon UV radiation. Upon reduction of Ub-SH from the nucleosomes we expect, based on our model, to detect an Ub-hDot1L cross-linked species. Further, if a cross-link between hDot1L is demonstrated, the identification of the interaction site within hDot1L would be crucial to understand the consequences of this recognition.

The modulation of hDot1L binding affinity to the nucleosome by Ub could be elucidated through the determination of binding constants of hDot1L to the H2B and H2B-Ub nucleosomes. Potentially, the contribution of Ub to hDot1L binding the nucleosome could be minor as hDot1L is reported to interact with the unmodified nucleosome similar to the H2B-Ub nucleosome in low resolution gel

[7] This is the work of Catlin Zhou in Professor Tom Muir's lab at Princeton University.

shift assays (presumably through the H4 basic patch and nucleosomal DNA).[121, 145, 147] Detecting a binding difference, if any, would provide much needed insight into how hDot1L engages the nucleosome. Further, the specific contributions of each nucleosomal recognition element could be investigated through mutagenesis. Although, an hDot1L-H2B-Ub MN co-crystal structure would be the most informative to how hDot1L engages the H2B-Ub nucleosome, this is potentially the most challenging undertaking.

7.4.2. *Future avenues in the investigation of hDot1L*

Many facets of hDot1L regulation remain to be explored and are highlighted below as possible areas of investigation:

(i) hDot1L has been shown to interact with long non-coding RNAs (lncRNA) (Figure 7.5).[210] Specifically, the lncRNA PRNCR1 was found to bind to the androgen receptor protein within DNA enhancer elements where its association with hDot1L led to the methylation of the androgen receptor. An hDot1L-PRNCR1 interaction was identified by a mass spectrometry pull down with the in vitro transcribed PRNCR1. It would be interesting to investigate the mechanism by which hDot1L is recruited by PRNCR1, and use hDot1L to pull down and identify other hDot1L interacting lncRNAs. Additionally, methylation of the androgen receptor led to the recruitment of another lncRNA, PCGEM1. As PCGEM1 was found to bind to the methylated androgen receptor, it would be

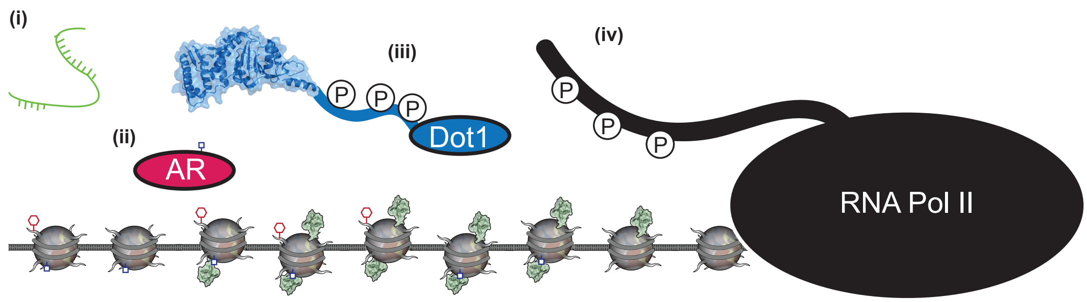


Figure 7.5. The many facets of hDot1L regulation. hDot1L (i) interacts with lncRNAs (green), (ii) methylates the androgen receptor (labeled AR, pink), (iii) is heavily post-translationally modified (Dot1 in blue, P denotes a phosphoylation site), (iv) and interacts with RNA Pol II (black). The catalytic domain of hDot1L is shown in blue (PDB 1NW3) along with the rest of hDot1L (blue oval) which structure has not been determined. Nucleosomes containing histone PTMs are shown as these are chromatin-templated processes.

interesting to investigate whether any lncRNAs interact with the H3K79 methylated nucleosome.[210]

(ii) Additionally, it would be worthwhile to investigate the ability of hDot1L to methylate non-histone substrates (Figure 7.5). Currently, the androgen receptor protein is methylated by hDot1L at K349 and is the only identified non-H3K79 methylation attributed to hDot1L.[210] Genome-wide chromatin methylation profiling in the presence or absence of hDot1L could reveal additional hDot1L substrates.[211] Additionally, it would be interesting to investigate whether the methylation of the androgen receptor, (or additional methylations attributed to hDot1L) is Ub-dependent. As this is presumably a chromatin-templated process, either H2B-Ub or the ubiquitylation of the androgen receptor itself may be

required for this methylation. Interestingly, Ub-proteomic screens have identified multiple ubiquitylation sites within the androgen receptor.[58]

(iii) hDot1L is heavily acetylated and phosphorylated (Figure 7.5).[212] These PTMs have been annotated in the catalytic core of hDot1L, in the region directly C-terminal to the catalytic core of hDot1L required for binding MNs, and near the identified site of hDot1L-RNA Pol II interaction (this is a RNA Pol II CTD S2ph and S5ph based interaction).[130, 212] It would be fascinating to test the effect of these modifications on the modulation of hDot1L-protein and hDot1L-chromatin interactions as well as H3K79 methylation activity. As these modification sites are localized towards the C-terminus of the catalytic domain, a post-translationally modified hDot1L is an ideal candidate for EPL.

(iv) Lastly, given the recently identified RNA Pol II-hDot1L interaction[130] and the ability to reconstitute transcriptional systems with semi-synthetic H2B-Ub (or H2B-uLL), it would be interesting to test what impact the L71/73 surface has in a broader *in vitro* chromatin context i.e. whether this surface is conditionally or absolutely required (Figure 7.5). Interestingly, hSet1C-mediated methylated can be stimulated through the co-activator P300 combined with the transcription factor p53 with methyltransferase activity being indistinguishable in the presence or absence of H2B-Ub.[197] It would be interesting to elucidate whether hDot1L can be activated in the absence of H2B-Ub mimicry, which could be potentially achieved through additional co-factors or histone PTMs.

7.5. Perspective

The ubiquitylation of H2B at K120 is an information rich post-translational modification implicated in both *trans*-histone pathways and the alteration of chromatin structure through discrete epitopes of the ubiquitin surface. Herein, we gained much insight into the many facets of H2B-Ub regulated chromatin processes. However, much remains to be explored in this complex system. The experimental approaches proposed above will enhance our understanding of H2B-Ub functions as well as allow for the investigation of the regulation of hDot1L in a greater detail.

The biochemical analysis of chromatin-templated processes is integral to the understanding the complexity of chromatin functions at the molecular level. Accordingly, the strategies developed within this thesis and discussed herein will serve to expedite the reconstitution of chromatin containing histone PTMs. We anticipate that, in combination with the substantially improved epigenomic methods that quantitatively discern the contribution of these epigenetic marks within the nucleus, the further biochemical analysis of modified chromatin has the potential to substantially expand our knowledge of the underlying mechanisms of chromatin function.

Chapter 8 Materials and Methods

8.1 Materials

Amino acid derivatives, coupling reagents and resins were purchased from Novabiochem (Merck, NJ). The thiol activating reagents 2,2'-dithiobis(5-nitropyridine) (DTNP), 5, 5'-Dithiobis(2-nitrobenzoic acid) (DTNB), and cystamine dihydrochloride were purchased from Sigma-Aldrich Chemical Company (Milwaukee, WI). [^{3}H]-S-adenosyl methionine (^{3}H-SAM), Amplify solution and Sephacryl S-200 resin were obtained from GE Healthcare (Waukesha, WI). S-adenosyl methionine was obtained from New England Biolabs (Ipswich, MA). ^{15}N labeled ammonium chloride and deuterium oxide (99.9%) were purchased from Cambridge Isotopes. All other commonly used chemical reagents and solvents were purchased from Sigma-Aldrich Chemical Company (Milwaukee, WI) or Fischer Scientific (Pittsburgh, PA). Chemically competent DH5alpha, BL21(DE3), and BL21(DE3)pLysS cells were purchased from Novagen (Madison, WI). The pTXB1 vector, restriction enzymes, T4 DNA ligase, chitin resin, and NiNTA resin were obtained from New England BioLabs (Ipswitch, MA). Primer synthesis and gene sequencing were performed by Integrated DNA Technologies (Coralville, IA) and Genewiz (South Plainfeld, NJ), respectively. Criterion 15% Tris-HCl and 5% TBE gels were purchased from BioRad (Hercules, CA). Centricons were from Sartorius (Goettingen, Germany) and dialysis cassettes were from Pierce (Rockford, IL). PCR purification and gel extraction kits were purchased from Qiagen (Valencia, CA).

Table 7.1. List of antibodies used in this thesis.

Antigen	Supplier	Identifier
H3	Abcam	Ab1791
H3K79me2	Abcam	Ab3594
H3K4me3	Abcam	Ab7766
*y*H2A	Active Motif	39235
*y*H2B	Active Motif	39237

8.2 Equipment

Size-exclusion and ion-exchange chromatography were performed on an AKTA FPLC system from GE Healthcare equipped with a P-920 pump and UPC-900 monitor. Analytical reversed-phase HPLC (RP-HPLC) was performed on a Hewlett-Packard 1100 series instrument with a Vydac C18 column (5 micron, 4 x 150 mm), employing 0.1% TFA in water (A), and 90% CH3CN, 0.1% TFA in water (B), as the mobile phases. Typical analytical gradients were 0-73% B over 30 min at a flow rate of 1 mL/min. Preparative HPLC was carried out on a Waters prep LC system comprised of a Waters 2545 Binary Gradient Module and a Waters 2489 UV detector. A Vydac C18 process column (15-20 micron, 50 x 250 mm) or a semi-preparative column (12 micron, 10 mm x 250 mm) was employed at a flow rate of 30 mL/min, or, 4 mL/min, respectively. ESI-MS analysis was conducted on a Sciex API-100 single quadrupole spectrometer or Bruker Daltonics MicrOTOF-Q II mass spectrometer. All protein starting materials and ligation products were analyzed by C18 analytical RP- HPLC and ESI-MS. Scintillation counting was performed on a LKB Wallac 1209 RackBeta Primo

Liquid scintillation counter. All fluorescent measurements were done on a Fluorolog-3 fluorescence spectrometer (HORIBA Jobin Yvon, Edison NJ).

8.3 Cloning

All plasmids were prepared by either site directed mutagenesis or restriction enzyme-free cloning of the construct into a suitable plasmid.[213] Synthetic constructs were purchased directly from Integrated DNA technologies. All previously unreported constructs used throughout this thesis will be discussed.

8.3.1 Ubiquitin

Ubiquitin mutants were prepared using either an ubiquitin GyrA-intein Chitin Binding Domain fusion (Ub-GyrA-CBD) or an ubiquitin NPU intein 6xHis-tag fusion (Ub-NPU-6xHis) protein with the following human ubiquitin sequence:

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY
NIQKESTLHLVLRLRGG

8.3.1.1 Mutagenesis of pUb-GyrA-CBD:

All ubiquitin library mutants were made by site-directed mutagenesis using a pUb plasmid template previously engineered with ubiquitin(1-75) in a pTXB1 vector and mutagenic primers as follows [118]:

pUb1. Forward primer (Ub T9A/G10A/K11A/T12A) 5'-CAT ATG CAG ATC TTC GTG AAG ACT CTG <u>GCT GCT GCG GCC</u> ATC ACT CTC GAA GTG GAG

CCG-3’ and reverse primer (Ub T9A/G10A/K11A/T12A) 5’- CTC GGC TCC ACT TCG AGA GTG ATG GCC GCA GCA GC C AGA GTC TTC ACG AAG ATC TGC ATA TG-3’.

pUb2. Forward primer (Ub Q2A) 5’- CTTTAAGAAGGAGATATACAT ATG GCG ATC TTC GTG AAG ACT CTG AC -3’ and reverse primer (Ub Q2A) 5’ -GT CAG AGT CTT CAC GAA GAT CGC CAT ATG TAT ATC TCC TTC TTA AAG-3’ and forward primer (UbL15A/E16A) 5’- CT CTG ACT GGT AAG ACC ATC ACT GCC GCA GTG GAG CCG AGT GAC ACC ATT GAG -3’ and reverse primer (UbL15A/E16A) 5’ - CTC AAT GGT GTC ACT CGG CTC CAC TGC GGC AGT GAT GGT CTT ACC AGT CAG AG-3’ and forward primer (Ub K29A) 5’- C ACC ATT GAG AAT GTC AAG GCA GCG ATC CAA GAC AAG GAA GGC ATC CC-3’ and reverse primer (Ub K29A) 5’- GGG ATG CCT TCC TTG TCT TGG ATC GCT GCC TTG ACA TTC TCA ATG GTG -3’

pUb3. Forward primer (Ub K6A) 5’- GATATACAT ATG CAG ATC TTC GTG GCG ACT CTG ACT GGT AAG ACC ATC AC-3’ and reverse primer (Ub K6A) 5’- GTG ATG GTC TTA CCA GTC AGA GTC GCC ACG AAG ATC TGC ATA TGT ATA TC-3’ and forward primer (Ub T66A/H68A) 5’- GAC TAC AAC ATC CAG AAA GAG TCC GCC CTG GCC CTG GTA CTC CGT CTC AGA GGT TG-3’ and reverse primer (Ub T66A/H68A) 5’-CAA CCT CTG AGA CGG AGT ACC AGG GCC AGG GCG GAC TCT TTC TGG ATG TTG TAG TC-3’.

pUb4. Forward primer (Ub E18A/P19A/S20A/D21A) 5'- GGT AAG ACC ATC ACT CTC GAA GTG GCG GCG GCT GCC ACC ATT GAG AAT GTC AAG GCA AAG-3' and reverse primer (Ub E18A/P19A/S20A/D21A) 5'-CTT TGC CTT GAC ATT CTC AAT GGT GGC AGC CGC CGC CAC TTC GAG AGT GAT GGT CTT ACC-3'.

pUb5. Forward primer (Ub T22A/E24A/N25A) 5'-CT CTC GAA GTG GAG CCG AGT GAC GCC ATT GCG GCT GTC AAG GCA AAG ATC CAA GAC AAG -3' and reverse primer (Ub T22A/E24A/N25A) 5'-CTT GTC TTG GAT CTT TGC CTT GAC AGC CGC AAT GGC GTC ACT CGG CTC CAC TTC GAG AGT G-3'.

pUb6. Forward primer (Ub Q31A/D32A/K33A/E34A) 5'- C ATT GAG AAT GTC AAG GCA AAG ATC GCA GCC GCG GCA GGC ATC CCT CCT GAC CAG CAG AG -3' and reverse primer (Ub Q31A/D32A/K33A/E34A) 5'- CCT CTG CTG GTC AGG AGG GAT GCC TGC CGC GGC TGC GAT CTT TGC CTT GAC ATT CTC AAT G-3'.

pUb7. Forward primer (Ub P37A/P38A/D39A) 5'- GCA AAG ATC CAA GAC AAG GAA GGC ATC GCT GCT GCC CAG CAG AGG TTG ATC TTT GCT GGG-3' and reverse primer (Ub P37A/P38A/D39A) 5'- CCC AGC AAA GAT CAA CCT CTG CTG GGC AGC AGC GAT GCC TTC CTT GTC TTG GAT CTT TGC-3'.

pUb8. Forward primer (Ub G47A/K48A/Q49A) 5'- GAC CAG CAG AGG TTG ATC TTT GCT GCG GCA GCG CTG GAA GAT GGA CGC ACC CTG TCT G-3' and reverse primer (Ub G47A/K48A/Q49A) 5'-CAG ACA GGG TGC GTC CAT CTT CCA GCG CTG CCG CAG CAA AGA TCA ACC TCT GCT GGT C-3'.

pUb9. Forward primer (Ub E51A/D52A/G53A/R54A) 5'- G TTG ATC TTT GCT GGG AAA CAG CTG GCA GCT GCA GCC ACC CTG TCT GAC TAC AAC ATC C-3' and reverse primer (Ub E51A/D52A/G53A/R54A) 5'- CTG GAT GTT GTA GTC AGA CAG GGT GGC TGC AGC TGC CAG CTG TTT CCC AGC AAA GAT CAA C-3'.

pUb10. Forward primer (Ub D58A/Y59A/N60A) 5'- G CTG GAA GAT GGA CGC ACC CTG TCT GCC GCC GCC ATC CAG AAA GAG TCC ACC CTG CAC C-3' and reverse primer (Ub D58A/Y59A/N60A) 5'- GGT GCA GGG TGG ACT CTT TCT GGA TGG CGG CGG CAG ACA GGG TGC GTC CAT CTT CCA GC-3'.

pUb11. Forward primer (Ub Q62A/K63A/E64A) 5'- CGC ACC CTG TCT GAC TAC AAC ATC GCG GCA GCG TCC ACC CTG CAC CTG GTA CTC CGT C-3' and reverse primer (Ub Q62A/K63A/E64A) 5'- GAC GGA GTA CCA GGT GCA GGG TGG ACG CTG CCG CGA TGT TGT AGT CAG ACA GGG TGC G-3'.

pUb12. Forward primer (UbL71/R72/L73/R74A) 5'- G AAA GAG TCC ACC CTG CAC CTG GTA GCC GCT GCC GCA GGT TGC ATC ACG GGA GAT GCA CTA

G-3' and reverse primer (UbL71/R72/L73/R74A) 5' CTA GTG CAT CTC CCG TGA TGC AAC CTG CGG CAG CGG CTA CCA GGT GCA GGG TGG ACT CTT TC-3'.

8.3.1.2 Mutagenesis of ubiquitin in NPU intein 6XHis tagged vector:

Selected ubiquitin mutants were generated by site-directed mutagenesis using a previously described pUb-NPU-6xHis plasmid template[165] and mutagenic primers as follows**:**

pUb13. Forward primer (UbL8A) 5'-GCA GAT CTT CGT GAA GAC TGC GAC TGG TAA GAC CAT CAC T-3' and reverse primer (Ub L8A) 5'- AGT GAT GGT CTT ACC AGT CGC AGT CTT CAC GAA GAT CTG C -3' and forward primer (UbI44A) 5'-CTG ACC AGC AGA GGT TGG CCT TTG CTG GGA AAC AGC-3' and reverse primer (UbI44A) 5'- GCT GTT TCC CAG CAA AGG CCA ACC TCT GCT GGT CAG-3'.

pULL. Forward primer (UbL71/73A) 5'-C CTG CAC CTG GTA GCC CGT GCC AGA GGT GGT TGT TTA AGC TAT GAA ACG GAA ATA TTG AC-3' and reverse primer (UbL71/73A) 5'-GT CAA TAT TTC CGT TTC ATA GCT TAA ACA ACC ACC TCT GGC ACG GGC TAC CAG GTG CAG G-3'.

pURR. Forward primer (UbR72/74A) 5'- GAG TCC ACC CTG CAC CTG GTA CTC GCT CTC GCA GGT GGT TGT TTA AGC TAT GAA ACG G -3' and

reverse primer (R72/74A) 5' C CGT TTC ATA GCT TAA ACA ACC ACC TGC GAG AGC GAG TAC CAG GTG CAG GGT GGA CTC -3'.

8.3.2 Preparation of Hub1 containing plasmid

A Hub1ub mutant construct was prepared using a Hub1ub-NPU-6xHis protein with the following Hub1ub sequence:

MIEVVVNDRLGKKVRVKSLAEDSVGDFKKVLSLQIGTQPNKIVLQKGGSVLKDHISLED
YEVHDQTNLELVLRLRGG

8.3.2.1 Construction of pHub1ub-NPU-6xhis vector:

pHub1ub. Hub1ub was cloned into a the Ub-NPU-6xHis vector via ligation-independent cloning of the Hubub1 mini-gene ordered from IDT with the sequence:

5'-ATGATTGAAG TCGTCGTTAA TGATCGCCTT GGGAAAAAGG TGCGTGTAAA ATCCTTAGCG GAAGATAGTG TTGGTGATTT TAAAAAAGTG CTGTCTTTAC AGATCGGCAC CCAGCCAAAC AAAATTGTGC TGCAGAAAGG CGGTAGTGTG CTGAAAGATC ACATTTCACT TGAAGACTAC GAAGTTCATG ATCAAACTAA CCTTGAACTG GTACTGCGTC TGCGGGGCGG G-3'

To anneal the N-terminal and C-terminal sequence of the mini-gene to the a linearized Ub-NPU-6xHis pTXB1 vector, PCR was utilized with the following

primers: Forward primer (N-terminus of Hub1ub): 5'-GTT TAA CTT TAA GAA GGA GAT ATA CAT ATG ATC GAA GTT GTT GTG AAT GAT CGC-3', Reverse primer (C-terminus of Hub1ub): 5'-GGT TCA TGA CCA GAC AAA TTT AGA ACT TTA TTA TCT G GGT TCA TGA CCA GAC AAA TTT AGA ACT TGTA CTC AGA CTG AGA GGT GGT TG-3'.

8.3.3 Histones

8.3.3.1 Mutagenesis of histone H3

H3C110S was prepared using the previously described *xenopus laevis* H3.1 construct in a pET3a plasmid as a template[118] using the mutagenic primers: Forward primer (H3C110S) 5'-GCT CTC TTT GAG GAC ACC AAC CTG AGC GCC ATC CAC GCC AAG AGG GTC ACC ATC ATG C-3' and reverse (H3C110S) 5' G CTA GAT GGT GAC CCT CTT GGC GTG GAT GGC GCT CAG GTT GGT GTC CTC AAA GAG TGC-3'.

8.3.3.2 Mutagenesis of Ub-H2A yeast constructs

Ub-H2A fusions were prepared using a wt *y*H2A construct in a pET3a vector as a template protein and the ubiquitin sequence from the Ub-NPU-6xHis vector which resulted in an n-yH2A sequence of:

```
MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY
NIQKESTLHLVLRLRGGLTFPVGRVHRLLRRGNYAQRIGSGAPVYLTAVLEYLAAEILE
LAGNAARDNKKTRIIPRHLQLAIRNDDELNKLLGNVTIAQGGVLPNIHQNLLPKKSAKT
AKASQEL
```

8.3.3.3 Construction of n-yH2A vector:

The following primers were utilized to PCR amplify an linear H2A(20-123)-pET3a vector for the insertion of Ub: Forward primer (directed toward the H2A(20-28) 5'-TTA ACA TTC CCA GTT GGT AGA GTG C-3' Reverse primer (directed toward the 5' end of vector construct): 5'-GGA GAT ATA CAT ATG-3'

Primers to create the Ub(1-76) product: Forward primer (directed toward the N-terminus of Ub) 5' CTT TAA GAA GGA GAT ATA CAT ATG CAG ATC TTC GTG AAG ACT CT-3' Reverse primer (directed toward the C-terminus of Ub and H2A(20-26)): 5'- GCA CTC TAC CAA CTG GGA ATG TTA AAC CCG CGG CAC GGG CTA CCA GGT GCA GGG-3'

pUb-yH2A Forward primer (L71/L73A): 5'- CC CTG CAC CTG GTA CTG CGT CTG GCG GGT TTA ACA TTC CCA GTT GG-3' and reverse primer (L71/L73A): 5'- CC AAC TGG GAA TGT TAA ACC CGC CAG ACG CAG TAC CAG GTG CAG GG-3'

pULL-yH2A Forward primer (L71/L73A): 5'- CC CTG CAC CTG GTA GCC CGT GCC GCG GGT TTA ACA TTC CCA GTT GG-3' and reverse primer (L71/L73A): 5'- CC AAC TGG GAA TGT TAA ACC CGC GGC ACG GGC TAC CAG GTG CAG GG-3'

8.3.4 Mutagenesis of hDot1L(1-416)

hDot1L(1-416)C44S/C74S/C178S (referred as hDot1L in the main text):

```
MGEKLELRLKSPVGAEPAVYPWPLPVYDKHHDAAHEIIETIRWVSEEIPDLKLAMENYV
LIDYDTKSFESMQRLSDKYNRAIDSIHQLWKGTTQPMKLNTRPSTGLLRHILQQVYNHS
VTDPEKLNNYEPFSPEVYGETSFDLVAQMIDEIKMTDDDLFVDLGSGVGQVVLQVAAAT
NSKHHYGVEKADIPAKYAETMDREFRKWMKWYGKKHAEYTLERGDFLSEEWRERIANTS
VIFVNNFAFGPEVDHQLKERFANMKEGGRIVSSKPFAPLNFRINSRNLSDIGTIMRVVE
LSPLKGSVSWTGKPVSYYLHTIDRTILENYFSSLKNPKLREEQEAARRRQQRESKSNAA
TPTKGPEGKVAGPADAPMDSGAEEEKAGAATVKKPSPSKARKKKLNKKGRKMAGRKRGR
PKK
```

A 6xHis-Sumo tagged cysteine-less hDot1L(1-416) construct was prepared and cloned into a pET30 vector prepared via sequence and ligation-independent cloning using a 6XHis-SUMO PCR product and a PCR product prepared from the previously reported hDot1L(1-416) construct [118] where all cysteine residues (C44S/C74S/C178S) were mutated to serine. The mutagenic primers used prepare a cysteine-less hDot1L(1-416):

Forward primer (C44S) 5'- ACC ATC CGA TGG GTC AGT GAA GAA ATC CCG G -3' and reverse, 5'- CCG GGA TTT CTT CAC TGA CCC ATC GGA TGG T -3'. Forward primer (C74S) 5'- AGC ATG CAG AGG CTC AGC GAC AAG TAC AAC C -3' and reverse, 5'- GGT TGT ACT TGT CGC TGA GCC TCT GCA TGC T -3'. Forward primer (C178S) 5'- TGC TGC TGC CAC CAA CAG CAA ACA TCA CTA

TGG -3' and reverse, 5'-CCA TAG TGA TGT TTG CTG TTG GTG GCA GCA GCA-3'.

Primers to create hDot1L(1-416) PCR product: Forward primer (directed toward Sumo-hDot1) 5’-GCT CAC AGA GAA CAG ATT GGT GGT ATG GGG GAG AAG CTG GAG CTG -3’ Reverse primer (Directed toward hDot1L–pET30 vector) 5’-CCG CAA GCG CGG GCG CCC CAA GAA GTA GGA ATT CGA GCT CCG TCG ACA AGC TTG CGG C-3’

8.4 Protein preparation

8.4.1 Preparation of ubiquitin aminoethanethiol analogs by thiolysis of GyrA intein fusions

Ubiquitin aminoethanethiol analogs were prepared as described.[152] E. coli BL21(DE3) cells were transformed with plasmids containing the Ubiquitin-GyrA-Chitin Binding Domain (CBD) fusion or ubiquitin mutant-GyrA-CBD fusions and grown in 6L of Luria-Bertani (LB) media (100ug/L ampicillin) at 37^{o}C until an OD600 of 0.6. Overexpression of the desired proteins was induced by the addition of 0.5 mM IPTG and the cells were grown for an additional 5 h at 30 °C. The cells were then harvested by centrifugation at 10k x *g* for 15 min and the cell-pellet was resuspended in buffer A (50 mM Tris, 200 mM NaCl, 1 mM EDTA, pH 7.2). The cells were lysed by passage through a French Press and the soluble fraction separated from insoluble cellular debris by centrifugation at 18.5-

20k x *g* for 20 min. After filtration through a 0.45 μm filter, supernatants were bound to a 50 mL chitin column, pre-equilibrated with ten column volumes of buffer A, for 2 h at 4 °C. The resin was washed with 20 column volumes of buffer A, followed by 3 column volumes of column buffer B (50 mM Tris, 200 mM NaCl, and 1 mM EDTA, pH 7.5). Ubiquitin and Ub mutants were cleaved from the respective intein-CBD fusions by incubation with 1.5 column volumes of buffer B containing 50 mM of cystamine-dihydrochloride, and 50 mM of Tris(2-carboxyethyl) phosphine, pH 7.75 for 48 h. The eluted proteins, bearing the desired C-terminal aminoethanethiol linker, were subsequently purified by C18 process RP-HPLC employing a gradient of 25-55% B, over 60 min, followed by a second purification by C18 semi-preparative HPLC with a gradient of 25-55% B over 60 min to remove residual cystamine. Typical yields were 4-6mg for each protein. The purest fractions were pooled and analyzed by analytical HPLC, and mass spectrometry.

8.4.2 Preparation of ubiquitin and Hub1 aminoethanethiol analogs by thiolysis of NPU intein 6xHis fusions

E. coli BL21(DE3) cells were transformed with plasmids containing the Ubiquitin-NPU-6XHis fusion or ubiquitin mutant-NPU-His tag fusions and grown in 6L of Luria-Bertani (LB) media (100 ug/L ampicillin) at 37 °C until an OD600 of 0.6. Overexpression of the desired proteins was induced by the addition of 0.5 mM IPTG and the cells were grown for an additional 5 h at 30 °C. The cells were then

harvested by centrifugation at 10k x *g* for 15 min and the cell-pellet was resuspended in buffer B (50 mM Sodium Phosphate, 200 mM NaCl, 1 mM EDTA, pH 6.0). The cells were lysed by passage through a French Press and the soluble fraction separated from insoluble cellular debris by centrifugation at 18.5-20k x *g* for 20 min. After filtration through a 0.45 μm filter, supernatants were bound to a 5 mL Ni-NTA column, pre-equilibrated with ten column volumes of buffer B, for 30 min at 4 °C. The resin was washed with 10 column volumes of buffer B, followed by 10 column volumes of column buffer B containing 20 mM imidazole, and eluted with 10 column volumes of buffer B containing 250 mM imidazole. Following dialysis into buffer B (containing no imidazole), Ub and Ub mutants were cleaved from the respective NPU-6XHis fusions by incubation with 10 column volumes of buffer B containing 50 mM of cystamine dihydrochloride, and 50 mM of Tris(2-carboxyethyl) phosphine, pH 7.2 for 10 h. The eluted proteins, bearing the desired C-terminal aminoethanethiol appendage, were subsequently purified by C18 process RP-HPLC employing a gradient of 25-55% B, over 60 min, followed by a second purification by C18 semi-preparative HPLC with a gradient of 25-55% B over 60 min to remove residual cystamine. Typical yields were 10-30 mgs per liter for each protein. The purest fractions were pooled and analyzed by analytical HPLC, and mass spectrometry.

8.4.3 Preparation of ubiquitin α-thioesters

All constructs were expressed, purified and dialyzed into Buffer B similar to Ub aminoethanethiol constructs and then thiolyzed with mercaptoethane sulfonate

(MES) as referenced in McGinty et al. 2008.[165] Briefly, ubiquitin and Ub mutants were cleaved from the respective NPU intein fusions, after a dialysis step into buffer B, by incubation with 10 column volumes of buffer B containing 80 mM MES and 10 mM of Tris(2-carboxyethyl) phosphine, pH 7.2 for 18 h. The eluted proteins, bearing the desired C-terminal MES a-thioester, were subsequently purified by C18 process RP-HPLC employing a gradient of 25-55% B, over 60 min, and the purest fractions were pooled and analyzed by analytical HPLC, and mass spectrometry. Typical yields were 10-30 mgs per liter per protein.

8.4.4 Preparation of ^{15}N labeled Ub and uLL

E. coli BL21(DE3) cells were transformed with a plasmid encoding the ubiquitin mutant-NPU-His tag fusion and grown in 1L of M9 minimal media containing ^{15}N labeled ammonium chloride (100ug/L ampicillin) at 37 °C until an OD600 of 0.6. Overexpression of the desired proteins was induced by the addition of 0.5 mM IPTG and the cells were grown for an additional 4 h at 37 °C. The Ub-NPU-His fusion was prepared exactly as the aminoethanethiol constructs but then thiolyzed with Dithiothreitol (DTT). Briefly, uLL was cleaved from the respective NPU intein fusion, resulting in a free carboxyl group, after a dialysis step into buffer B, by incubation with 10 column volumes of buffer B containing 80 mM DTT, pH 8.5 for 10 h. The eluted protein was subsequently purified by C18 process RP-HPLC employing a gradient of 25-55% B, over 60 min, and the purest fractions were pooled and analyzed by analytical HPLC, and mass spectrometry. Typical yields were 5 -10mgs per liter per protein.

8.4.5 Preparation of deuterated Ub_{ILV}-SH

E. coli Codon plus cells were transformed with a plasmid encoding the ubiquitin-NPU-6xHis tag fusion and grown in 1L of deuterated M9 minimal media containing 60mg α-ketobutyric acid and 80mg α-keto-3-methylbutyric acid necessary for ILV labeling, and 100ug/L ampicillin at 37 °C until an OD600 of 0.6. Overexpression of the desired proteins was induced by the addition of 0.5 mM IPTG and the cells were grown for an additional 4 h at 37 °C. The Ub_{ILV}-NPU-6xHis fusion was prepared exactly as the aminoethanethiol constructs. The eluted protein was subsequently purified by C18 process RP-HPLC employing a gradient of 25-55% B, over 60 min, and the purest fractions were pooled and analyzed by analytical HPLC, and mass spectrometry. Typical yields were 5-7 mgs per liter per protein.

8.4.6 Expression of recombinant histones

All histones (including the Ub-yH2A fusions) were prepared as previously reported.[118] Untagged histone constructs and Ub-H2A fusions were grown in LB media (100ug/L ampicillin) at 37 °C until an OD600 of 0.6. Overexpression of the desired protein was induced by the addition of 0.5 mM IPTG and the cells were grown for an additional 3 h at 37 °C. The cells were then harvested by centrifugation at 10k x *g* for 15 min and the cell-pellet was resuspended in buffer B (50 mM Sodium Phosphate, 200 mM NaCl, 1 mM EDTA, pH 8.0). The cells were lysed by passage through a French Press and the insoluble fraction was

washed 3 times with buffer B containing 1% triton x and centrifuged at 18.5-20k x *g* for 20 min. Histones were extracted with through the addition of 1 ml DMSO followed by 20 ml of buffer B containing 6M guanidinium hydrochloride (Gn-HCl). Size exclusion chromatography was utilized for the purification under denaturing conditions using an S200 preparatory column. Histone containing fractions from SEC were dialyzed into H_2O further HPLC purified on 30 to 70% B gradient. The purest fractions from HPLC purification were pooled and analyzed by analytical HPLC and mass spectrometry.

8.4.7 Fluorescent labeling of H2A(N110C)

H2AN110C was labeled with fluorescein as previously reported.[117] Briefly, 4 mgs of H2AN110C was dissolved in 2 mL labeling buffer (20 mM Tris, pH 7.8, 6 M Gn-HCl, 200 uM TCEP) and 0.4 mg fluorescein-5-maleimide (0.8 umol, 2.7 equivalents) was added in 50 ul N,N-dimethylformamide (DMF) in the dark. The mixture was stirred for 1 h and quenched by addition of 1 mM 2-mercaptoethanol. fH2A was purified through a 30-70% B HPLC gradient over the course of 1 hour. The purest HPLC fractions were pooled and analyzed by analytical HPLC, and mass spectrometry.

8.4.8 Synthesis of ubiquitylated H2B$_{ss}$Ub constructs

H2B$_{ss}$Ub and H2B$_{ss}$Ub mutants were prepared according to published protocols.[152] In a typical reaction, 2 mgs of DTNP (6.45 mmol) was dissolved in 500 mL of a 3:1 (v/v) acetic acid:water mixture and added to 2mgs H2BK120C

(0.15 μmol, ~20 equivalents). The reaction was allowed to proceed for 12 h at 25 °C, prior to purification by C18 semi-preparative RP-HPLC with a 42-48% B gradient at 50 °C over 45 min. The purest HPLC fractions were pooled and analyzed by analytical HPLC, and mass spectrometry.

For Ub-SH conjugation, 1.5 mg of H2BK120C-TNP (0.108 umol) and 2.0 mgs of Ub-SH or Ub_{mut}-SH (0.232 umol, 2 equivalents) were dissolved in 625 uL of reaction buffer consisting of 1M HEPES, 6 M Gn-HCl, pH 6.93. The reaction was allowed for 1 h at 25°C with continuous shaking. The reaction was purified by C18 semi-preparative RP-HPLC with a 30-70% B gradient at 25 °C over 60 min. The purest HPLC fractions were pooled and analyzed by analytical HPLC, and mass spectrometry.

8.4.9 Preparation of H2B-Ub conjugates by expressed protein ligation

H2B-Ub and H2B-uLL were prepared by an optimized sequential expressed protein ligation procedure described in McGinty et al 2009.[121] Note the ubiquitin in these semisynthetic constructs harbored a G76A mutation.

The sequence corresponding to residues 117–125 of *x*H2B with an A117Thz replacement and a cysteine conjugated to K120 was synthesized as described previously, with some exceptions.[118] Briefly, on pre-loaded Wang resin using manual solid-phase peptide synthesis with an Fmoc N^{α} protection strategy and using 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

(HBTU) for amino-acid activation. Standard ^{t}butyl side-chain protection was used throughout with the following exceptions: the ϵ-amino group of K120 was protected with the Alloc group, and the thiol group of the C117 synthon was protected as Boc-Thz. The ligation auxiliary (cysteine conjugated to K120) was installed on the solid phase as follows: (i) the Alloc group on K120 was deprotected by two successive incubations of the peptidyl-resin in DCM with 24 eq. $PhSiH_3$ and 0.25 eq. $Pd(PPh_3)_4$ for 30 min each; (ii) dry, alloc deprotected peptidyl-resin was re-swelled in DMF and Boc-Cys(Trt)-OH was coupled using standard Fmoc coupling conditions**.** After cleavage from the resin with TFA:TIS:H_2O (95:2.5:2.5) for 3 h, H2B(117—125)A117Thz,CysK120 was purified by RP-HPLC on a semi-preparative scale using a 5–35% B gradient over 60 min.

The ligation reaction between H2B(117-125)A117Thz,CysK120 and Ub(1–75)-MES or uLL(1-75)-MES was performed using conditions similar to those previously employed.[161] 8mM purified H2B peptide and 1.5mM Ub(1-75)-MES or uLL(1-75)-MES was dissolved in 140 µl of buffer containing 6M Gn-HCl, 200 mM sodium phosphate pH 7.0, 25 mM TCEP, 100mM MESNa and incubated at room temperature for 4 hours. The Thz moiety on each H2B peptide was then deprotected by adding 140uL 1:1 solvent A:solvent B, 40uL 3M methoxylamine-HCl, then increasing the pH to 4.0 using 5N NaOH before flushing with argon and letting sit at room temperature overnight. The ligation products were purified using semi-preparative HPLC with a 20–45% B gradient over 60 min, yielding approximately 4.0 mg of pure, lyophilized protein each.

100uL of 6M Gn-HCl, 200mM sodium phosphate pH 7.9, 25mM TCEP, 100mM MESNa was added to 2.5mM Thz deprotected uH2B(117-125) or uLLH2B(117-125) and 2.1mM H2B(1–116)-MES. The pH was adjusted to 7.8 using 2N NaOH before flushing with argon and letting sit at room temperature overnight. 15mM TCEP was then added to the ligation mixtures, respectively, and the exogenous thiol was diluted 625x using a 5000 MWCO spin concentrator and argon-sparged 6M Gn-HCl, 0.2M Pi, pH 7.0 solution (Gua buffer) over a period of several hours. After spinning each solution down to a final concentration of 150uL, 50uL of 750mM TCEP in Gua buffer, 20uL 400mM reduced glutathione, and 16uL 0.2M VA-061 radical initiator in MeOH was added to each. The desulfurization solutions were flushed with argon, and placed at 37^{o} C overnight. H2B-Ub and H2B-uLL were then purified out of the solutions by diluting each 5x with 30%B before performing semi-preparative purification on a 40-60%B in 60min gradient.

The purest fractions from HPLC purification were pooled and analyzed by analytical HPLC, and mass spectrometry, yielding approximately 1mg of final product each.

8.4.10 hDot1L(1-416) preparation

E. coli BL21(DE3) cells were transformed with plasmids containing the 6XHis-SUMO-hDot1L(1-416) fusion and grown in 6L of Luria-Bertani (LB) media (50 ug/L kanamycin) at 37 °C until an OD600 of 0.6. Overexpression of the desired protein was induced by the addition of 0.5 mM IPTG and the cells were grown for

an additional 18 h at 18 °C. The cells were then harvested by centrifugation at 10k x *g* for 15 min and the cell-pellet was resuspended in buffer C (50 mM Tris, 200 mM NaCl, 1 mM EDTA, pH 7.5 at 4 °C). The cells were lysed by passage through a French Press and the soluble fraction separated from insoluble cellular debris by centrifugation at 18.5-20k x *g* for 20 min. After filtration through a 0.45 µm filter, supernatants were bound to a 5 mL Ni-NTA column, pre-equilibrated with ten column volumes of buffer C, for 12 h at 4 °C. The resin was washed with 10 column volumes of buffer C, followed by 10 column volumes each of column buffer C containing 25mM, 50mM, and 100mM imidazole. His-SUMO-hDot1L(1-416) was eluted with 10 column volumes of buffer B containing 250mM imidazole by collecting 1 column volume fractions which were analyzed by SDS-PAGE and the purest fractions were pooled. The His-SUMO tag was cleaved from hDot1L(1-416) at 4 °C by the addition of SUMO protease. Cleavage was monitored by SDS-PAGE, and upon 100% cleavage hDot1L(1-416) was purified by cation exchange chromatography using a High Trap SP FF 5ml column (gradient is 100mM NaCl to 1M NaCl over 10 column volumes). Fractions were analyzed by SDS-PAGE, and the purest fractions were pooled, concentrated and purified further using gel filtration. Fractions were analyzed by analytical SEC FPLC, pooled and stored at -80 °C in 50% glycerol.

8.4.11 Preparation of full-length hDot1L, yDot1, and ySet1C

These enzymes were purified from baculovirus expression system as previously described. [118, 121] Briefly, all enzymes contained a FLAG tag on the N-terminus of

the Dot1 or Set1 methyltransferase. The baculovirus was generated according to the manufacturer's instructions. Sf9 cells were infected with the baculovirus and the resulting cell extracts were subjected to standard purification procedures as described below. Sf9 cells (700 mL at a density of 1 million cells per mL) were infected with fresh virus (7.5 mL) three days prior to collection by centrifugation at 430 g for 5 min. Cells were washed with phosphate buffered saline and centrifuged again. Washed cells were resuspended in 18 mL lysis buffer (20 mM Tris HCl, 500 mM NaCl, 4 mM $MgCl_2$, 0.4 mM EDTA, 20% glycerol, 1 mM PMSF, 2 mM DTT, pH 7.9), and disrupted with 3x10 strokes with a dounce homogenizer with a 10 min break between sets of 10. After removal of cell debris by centrifugation at 22.5 kg for 15 min, the supernatant was adjusted to 0.1% NP-40 and 300 mM NaCl, by dilution with 20 mM Tris HCl containing 10% glycerol,and the resulting solution was incubated with 400 µL M2-agarose for 3.5 h at 4°C. The agarose was washed four times with 10 mL wash buffer (20 mM Tris HCl, 150 mM NaCl, 2 mM $MgCl_2$, 0.2 mM EDTA, 15% glycerol, 0.1% NP-40, 1 mM PMSF, 1mM DTT (this was omitted for hDot1L), pH 7.9). Each wash was followed by centrifugation at 400 g for 1 min. The third wash was left on a rotator for 10 min at 4 °C. Bound methyltrasnferase was eluted with 3x100 µL wash buffer containing 0.5 mg/mL 3x FLAG peptide (Sigma). These enzyme preparations were stored at -80 °C in 50% glycerol.

8.5 Nucleosome reconstitution

8.5.1 Histone octamer formation

Histone octamers were formed as previously described.[117] Briefly, lyophilized histones were resuspended in unfolding buffer (20 mM Tris, pH 7.5, 7 M Gn-HCl), combined in equimolar amounts and dialyzed into refolding buffer (10 mM Tris, pH 7.5, 2 M NaCl, 1 mM ETDA) for 3 x 2 h. The refolded histone octamers were concentrated and purified by size-exclusion chromatography (Superdex 200 10/300). Fractions were analyzed by SDS-PAGE and the fractions containing equivalent amount of all histones were pooled, concentrated, and stored at -20 °C after the addition of 50% glycerol (v/v).

8.5.2 DNA preparation

A 153bp segment containing the 601 DNA sequence was prepared as previously described with slight alterations.[180] A plasmid containing 30 copies of the 147 base pair 601 DNA flanked by EcoRV sites (153 bp in total length) was assembled following according to a general DNA plasmid isolation protocol.[180] The plasmid was prepared via alkaline lysis, digested with EcoRV, and purified from the vector with 10% polyethylene glycol-3500 precipitation on ice followed by centrifugation at 26,00 g for 30 min. The 153bp construct was further purified by isopropanol and ethanol precipitation, centrifuged, resuspended in TE buffer (10 mM Tris pH 7.5, and 1 mM EDTA), and quantified by UV spectroscopy and stored in aliquots at -20 ^{o}C.

For array formation, a plasmid containing 12 copies of a 177 base pair repeat of the 601 nucleosome positioning sequence (12-177-601) flanked by EcoRV sites was purified from a 6 L culture of DH5alpha cells as previously described [180]. The 12-177-601 sequence was obtained by preparative digestion of the plasmid followed by selective precipitation of the fragment with 6% polyethylene glycol-6000 on ice followed by centrifugation at 26,000 g for 30 min. After phenol extraction and ethanol precipitation the DNA was redissolved in TE buffer (10 mM Tris pH 7.5, 1 mM EDTA), quantified by UV spectroscopy and stored in aliquots at -20 °C.

8.5.3 Mononucleosome reconstitution

1) Small-scale nucleosome reconstitution by dilution. Mononucleosomes were formed using a previously described step-wise dilution procedure with slight modifications.[152, 214] Briefly, octamers and 601 DNA were combined in 10 µL high salt (2M NaCl) refolding buffer to a final concentration of 3 µM. After incubation at 37 °C for 15 min, 3.3 µL of dilution buffer 1 (10 mM HEPES, 1 mM EDTA, 0.5 mM PMSF, pH7.9) was added and the temperature was dropped to 30 °C. Further dilutions of 6.7, 5, 3.6, 4.7, 6.7, 10, 30, and 20 µL, respectively, were then performed every 15 min. A final dilution with 100 µL of dilution buffer 2 (10 mM Tris HCl, 1 mM EDTA, 0.1% NP-40, 0.5 mM PMSF, 20% glycerol, pH 7.5) was carried out. After an additional 15 min, the nucleosomes were concentrated using Vivaspin 500 centricons (3-10 kDa MWCO) at 4 °C. Nucleosome formation

was verified by separation on a Criterion 5% TBE gel run in 0.5x TBE, followed by ethidium bromide staining.

2) Large-scale nucleosome reconstitution by dialysis. For act-MNs, large-scale nucleosome reconstitution was performed following procedures similar to those previously reported.[180] Briefly, octamers were mixed with 601 DNA in ratios optimized using the small scale reconstitution protocol described above, at 5 μM DNA concentration, and the NaCl concentration was adjusted to 2 M. The mixture was dialyzed against 200 mL initial buffer (1.4 M KCl, 10 mM Tris HCl, 0.1 mM EDTA, pH 7.5) at 4 °C for 60 min. Semi-gradual dialysis was performed at a flow rate of 1 mL/min diluting with final buffer (10 mM KCl, 10 mM Tris HCl,0.1 mM EDTA) over 12 hours. 1 subsequent dialysis step of at least 2 h in final buffer was performed to ensure that the nucleosomes were fully dialyzed.

8.5.4 Yeast Nucleosome formation

Yeast nucleosomes were formed via the small-scale dilution method but were additionally heat shifted at 55°C for 1 hour subsequent to MN formation to heat shift DNA relative to the octamer core. Nucleosome formation was verified by separation on a Criterion 5% TBE gel run in 0.5x TBE, followed by ethidium bromide staining.

8.5.5 Nucleosomal array formation

Nucleosomal arrays were prepared by the large scale dialysis method used to make nucleosomes however a 12x601 DNA template was used instead of the 153bp 601 DNA.[117] Upon formation, nucleosomal arrays were precipitated from solution with $MgCl_2$ (concentrations were empirically determined) and resuspended in nucleosome array buffer (10 mM KCl, 10 mM Tris HCl, , pH 7.5). Array formation was verified by separation on a APAGE (1% agarose, 1% polyacrylamide) gel run in 0.5x TBE, followed by Sybr Gold staining. Partial micrococcal nuclease digestion and restriction enzyme digests were utilized to assess array quality.[117]

8.5.6 Recombinant chromatin formation

Chromatinized plasmids were assembled using a previously described procedure[215] and plasmid[216]. Micrococcal nuclease digestion was employed to verify equivalent chromatin reconstitution with H2B, H2B-Ub and H2B-uLL.

8.6 Preparation of $H2B_{ss}Ub$ nucleosomes via direct ligation of ubiquitin.

In a typical reaction, reconstituted H2BK120C containing octamers (5 nmol, 50 uL) in octamer formation buffer (10 mM Tris, 2 M NaCl, 1 mM EDTA, pH 7.5) were mixed 1:1 with octamer formation buffer containing 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (50 mmol, 50 uL) and nutated at RT for 10 min. to yield activated (H2B-TNB) octamers. Act-MNs were formed via dialysis (also removing all free DTNB) as described above (see mononucleosome/nucleosome

array reconstitution). The quality of the reconstitution was assessed by separation on a Criterion 5% TBE gel run in 0.5x TBE buffer, followed by staining with ethidium bromide or Sybr Gold.

Ub-SH or Ub-SH mutants (4.25 mg, 0.5 nmol) were added to 50 pmol of act-MNs in a final reaction volume of 10 uL and incubated at 55 °C for 1 hour. The degree of Ub-SH ligation was determine by separation on a Criterion 5% TBE gel run in 0.5x TBE buffer, followed by staining with ethidium bromide. These Ub-SH act-MNs were used in methyltransferase assays without subsequent purification.

8.7 Methyltransferase assays

8.7.1 hDot1L methyltransferase assays

Methyltransferase assays were performed using an optimized protocol as previously described [118]. 1.5 pmol of nucleosomes and 0.2 pmol of hDot1L or 0.3 pmol full-length hDot1L were equilibrated in 20 mL of assay buffer (20 mM Tris 7.9, 140 mM NaCl, 2 mM $MgCl_2$, 1 mM PMSF). 250 nCi of [^{3}H]-S-adenosyl methionine was added and the reaction was incubated at 30 °C for 15 minutes. 15 mL of the reaction mixture was run on a Criterion 5% TBE gel in 0.5x TBE buffer, followed by staining with ethiduim bromide, or Sybr Gold. The gel was incubated in Amplify solution for 15 min. and then dried and visualized by fluorography. The remaining 5 mL of the reaction mixture was spotted on Whatman p81 filter paper and washed three times with $NaHCO_3$ solution (pH 9) and air-dried. Econo F Liquid Scintillation Cocktail (GE healthcare) was then

added, and the samples were counted with a LKB Wallac 1209 RackBeta Primo Liquid scintillation counter. Data presented is the average of 3 to 6 independent experiments ± s.e.m.

8.7.2 yDot1 methyltransferase assays

Methyltransferase assays were performed using an optimized protocol as previously described.[118] 1.5 pmol of nucleosomes and 0.2 pmol of yDot1 was equilibrated in 20 mL of assay buffer (20 mM Tris 7.9, 140 mM NaCl, 2 mM $MgCl_2$, 1 mM PMSF). 250 nCi of [^3H]-S-adenosyl methionine was added and the reaction was incubated at 30 °C for 15 minutes. 15 uL of the reaction mixture was run on a Criterion 5% TBE gel in 0.5x TBE buffer, followed by staining with ethiduim bromide, or Sybr Gold. The gel was incubated in Amplify solution for 15 min. and then dried and visualized by fluorography. The remaining 5 uL of the reaction mixture was spotted on Whatman p81 filter paper and washed three times with $NaHCO_3$ solution (pH 9) and air-dried. Econo F Liquid Scintillation Cocktail (GE healthcare) was then added, and the samples were counted with a LKB Wallac 1209 RackBeta Primo Liquid scintillation counter. Data presented is the average of 3 independent experiments ± s.e.m.

8.7.3 Western blot analysis of yDot1 methyltransferase assays

Methyltransferase assays were performed identical to yDot1 assays using [^3H]-S-adenosyl methionine except 0.8 mM of 'cold' [^1H]-S-adenosyl methionine was utilized. 20uL of the reaction mixture was run on a 15% Tris-HCl gel in 1.0x Tris-

Glycine buffer, followed by semi-dry transfer to a PVDF membrane and blotting with an antibody directed to H3K79me2. This was read out by chemo-luminescence

8.7.4 ySet1 methyltransferase assays

Methyltransferase assays were performed using a slightly altered protocol from previous reports.[201] 3 pmol of nucleosomes and ySet1C (containing 50 ng of Bre2) were equilibrated in 40 mL of HEG buffer (25 mM HEPES pH 7.6, 0.1 mM EDTA, 10% Glycerol). 0.7 mCi of [^{3}H]-S-adenosyl methionine was added and the reaction was incubated at 30 °C for 2 h. 40 uL of the reaction was run on a Criterion 5% TBE gel in 0.5x TBE buffer, followed by staining with Sybr Gold stain. The gel was incubated in Amplify solution for 15 min. and then dried and visualized by fluorography.

8.8 In vivo analysis of H3K79me2

Yeast experiments were performed using the wild-type yeast strain BY4742 and a strain that contained a Rad6 deletion in the BY4742 background (CDAY23). Plasmids were constructed through insertion of the Ub-H2A constructs into the pQQ18 plasmid that contained all H2B, H3, H4 histones. These plasmids contained Leu2 auxotropic markers and CEN sequences.

Plasmids (H2A, n-*y*H2A, Ub-*y*H2A, or uLL-*y*H2A) were transformed into BY4742 (for H2A) and CDAY23 (for H2A, n-*y*H2A, Ub-*y*H2A, or uLL-*y*H2A) and plated

on –Leu dropout plates. Colonies for each type of yeast were selected for and grown in –Leu dropout media overnight to reach confluence. Cells were diluted to an OD600 of 0.2 and grown for an additional ~6 hours until they reached an OD600 of 0.8. Cells were spun down for 20 min at 4^{o}C and lysed by TCA precipitation and bead beating. The TCA protein precipitate was resuspended Laemmli SB buffer heated at 100^{o}C for 10 min and run on a 15% Tris-HCl criterion gel. Western blot analysis was performed using standard protocols and the antibodies indicated in Table 1.

8.9 mPEG assays

Nucleosomes were subjected to mPEGylation according to the following protocol. Nuclesomes (4.5pmol), in small-scale nucleosome reconstitution buffer (see section 8.5.3) were combined 1:1 with mPEG buffer (1mM mPEG 10 mM KCl, 10 mM Tris HCl, pH 7.5) and incubated at 30^{o}C. To quench the reaction, at the indicated times, 30 ul of the reaction was added to 2.5 ul of DTT (10mM) on ice. The reactions were run on a Criterion 5% TBE gel in 0.5x TBE buffer, followed by silver staining to visualize the mPEGylated nucleosomes.

8.10 Biophysical analysis of Ub and H2B-Ub chromatin

8.10.1 NMR Analysis of Ub and uLL

NMR spectra of Ub and uLL in 20 mM Tris, pH 7.5, 10 mM KCl, and 1 mM EDTA concentrated to ~1 mM was recorded at 30°C on either a Varian Inova 800-MHz spectrometer or an A800 Avance III 800MHz spectrometer. All recorded data

were processed using NMRPipe and analyzed using Sparky (http://www.cgl.ucsf.edu/home/sparky/).

8.10.2 Methyl-TROSY of Ub_{ILV} containing nucleosomes

Reconstituted $H2B_{ss}Ub_{ILV}G76C$ MNs with deuterated histones were exchanged into NMR buffer (99% D_2O, 20mM sodium phosphate, pD 6.0) following reconstitution. MNs were concentrated to 100uM. NMR experiments were carried out on an 11.7-T (500 MHz 1H frequency) spectrometer equipped with a room-temperature probe head or a 14.1-T (600 MHz) spectrometer with a cryoprobe at 45^{o}C. Methyl-TROSY spectra were recorded with acquisition times t_1 and t_2 of 30 and 64 ms, respectively. All data were processed using NMRPipe and analyzed with Sparky (http://www.cgl.ucsf.edu/home/sparky/).

8.10.3 Chemical Shift Perturbation analysis

Chemical shift perturbation analysis was conducted similar to the analysis in Kato et al. 2011 and Lee et al. 2014 for methyl-TROSY data and ubiquitin analysis respectively.[190, 217] Briefly, for comparasion of Ub and uLL, HSQC annotated chemical resonances were identified and the chemical shift perturbation (CSP) was calculated according to:

$$CSP = (\Delta\delta_H^2 + (\Delta\delta_N * w_i)^2)^{1/2},$$

where $\Delta\delta_x$ is the difference in peak position between Ub and uLL

w_i = weighting factor to compare chemical shifts (0.2 for ^{15}N)

Spectra were referenced to minimize the 20% trimmed mean of the chemical shift perturbations. The 10% trimmed mean and the standard deviation of this mean was calculated so that the 10% trimmed mean + 2 standard deviations could be used as a cut off to determine whether the chemical shift perturbation of each amino acid residue was significant.

For comparison of Ub and H2B-Ub MN via methyl-TROSY annotated chemical resonances were identified and the chemical shift perturbation of each methyl group was calculated according to:

$$CSP = (\Delta\delta_H^{\,2} + (\Delta\delta_c{*}w_i)^2)^{1/2},$$

where $\Delta\delta_x$ is the difference in peak position

w_i = weighting factor to compare chemical shifts (0.16 for ^{13}C)

Spectra were referenced to minimize the 20% trimmed mean of the chemical shift perturbations. In the case of amino acid residues with 2 annotated methyl groups (for L and V residues) the CSP of each methyl group was averaged to calculate the total CSP for each amino acid residue.[190] The 10% trimmed mean and the standard deviation of this mean was calculated for the amino acid residue CSPs so that the 10% trimmed mean + 2 standard deviations could be used as a cut off to determine whether the chemical shift perturbation of each amino acid residue was significant.

8.10.4 Fluorescence–based chromatin compaction assay

Fluorescence measurements were performed in measurement buffer (10 mM Tris, pH 7.8, 10 mM KCl) containing appropriate amounts of $MgCl_2$ as previously described.[117] For measurements in the absence of Mg^{2+}, the buffer contained 0.1 mM EDTA. All buffers were degassed before use by sonication. Nucleosomal array concentrations were 50 nM (per nucleosome, measured by abs at 260nm) for all measurements. Samples containing different Mg^{2+} concentrations were prepared fresh by mixing an equivalent volume of array stock (100 nM per nucleosome) with measurement buffer containing twice the final amount of $MgCl_2$ and equilibrating for 15 min at RT before measurement. Measurements were performed on a Fluorolog-3 instrument (HORIBA Jobin Yvon) equipped with automated dual polarizers and using a Sub-Micro Fluorometer cell (Starna Cells) with 10-mm path length. The excitation wavelength was 480 nm with a bandwidth of 5 nm, and emission was recorded at 520 nM with 5 nm bandwidth. Six to eight measurements were taken per sample with an integration time of 5 sec for V/V, V/H, H/H and H/V polarizer settings (excitation/emission, V: vertical polarization, H: horizontal polarization). The final anisotropy was calculated by the formula SSA = $(I_{VV} - G \times I_{VH})/(I_{VV} + 2 \times G \times I_{VH})$, with $G = I_{HV}/I_{HH}$. Vertically polarized emission intensity was monitored as a function of [Mg^{2+}] for all samples to ensure no loss of signal. Data presented is the average of 3 independent experiments ± s.e.m.

8.11 Analytical data for proteins and peptides

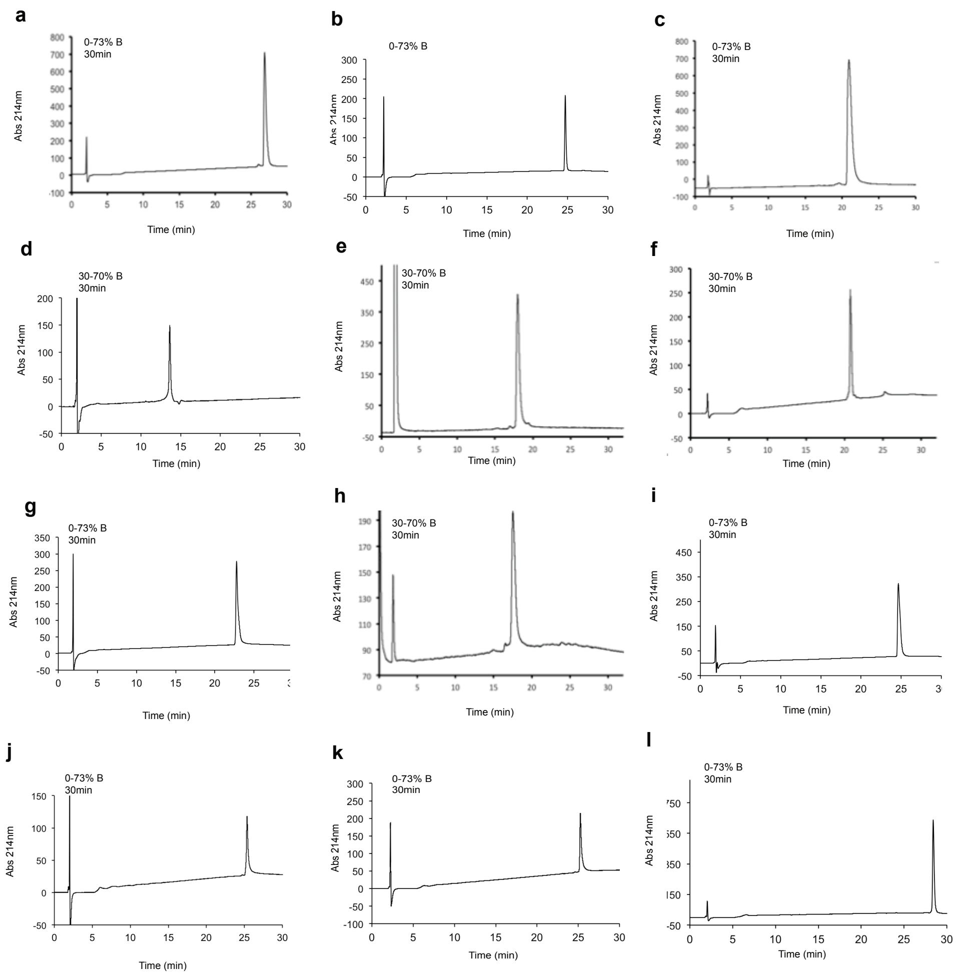


Figure 8.1 RP-HPLC analysis of proteins from Chapter 2. Proteins were run on a analytical C18 column at the indicated gradient and time. (a) Purified H2A (b) Purified H2B (c) Purified H3C110S (d) purified H4 (e) purified H2BK120C (f) purified Ub-SH (g) Purified H2BK120-DTNP (h) Purified $H2B_{ss}Ub$ (i) Purified H2AK119C (j) Purified H2BK108C (k) Purified H2BK116C (l) Purified H3K79C.

Table 8.2. Masses of purfied proteins from Chapter 2

Protein	Expected Mass (Da)	Observed Mass (Da)
H2A	13951.17	13953.81
H2B	13818	13820
H3C110S	15240.75	15240.71
H4	11237	11238
H2BK120C	13791.94	13791.8
Ub-SH	8567.94	8567.47
H2B-DTNP	13947	13946.03
$H2B_{ss}Ub$	22414.9	22413.78
H2AK119C	13925.4	13924.735
H2BK108C	13793	13794
H2BK116C	13793	13794
H3K79C	15215.4	15214.54

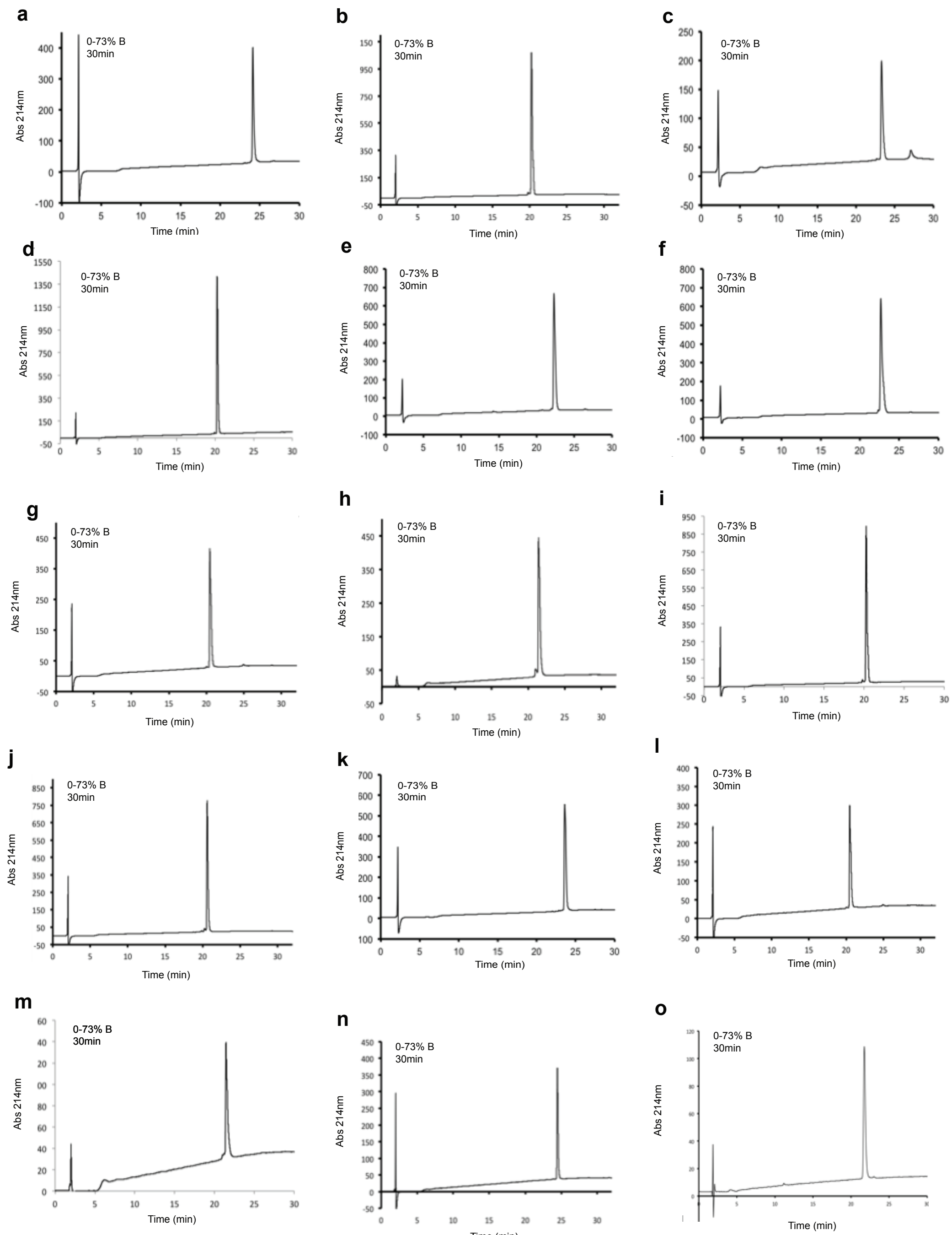


Figure 8.2 RP-HPLC analysis of proteins from Chapter 3. Proteins were run on a analytical C18 column at the indicated gradient and time. (a) Purified Ub1-SH (b) Purified Ub2-SH (c) Purified Ub3-SH (d) Purified Ub4-SH (e) Purified Ub5-SH (f) Purified Ub6-SH (g) Purified Ub7-SH (h) Purified Ub8-SH (i) Purified Ub9-SH (j) Purified Ub10-SH (k) Purified Ub11-SH (l) Purified Ub12-SH (m) Purified Ub13-SH (n) Purified $H2B_{ss}Ub7$ (o) Purified $H2B_{ss}Ub12$

Table 8.2. Masses of purfied proteins from Chapter 3

Protein	Expected Mass (Da)	Observed Mass (Da)
Ub1-SH	8464.85	8463.86
Ub2-SH	8222.47	8224.14
Ub3-SH	8414.75	8414.9
Ub4-SH	8423.85	8424.06
Ub5-SH	8436.85	8437.31
Ub6-SH	8351.75	8350.62
Ub7-SH	8471.85	8469.51
Ub8-SH	8467.85	8465.49
Ub9-SH	8394.85	8394.59
Ub10-SH	8388.85	8388.1
Ub11-SH	8395.85	8394.29
Ub12-SH	8313.55	8313.41
Ub13-SH	8482.76	8482.34
$H2B_{ss}Ub7$	22263.76	22262.51
$H2B_{ss}Ub12$	22104.49	22102.73

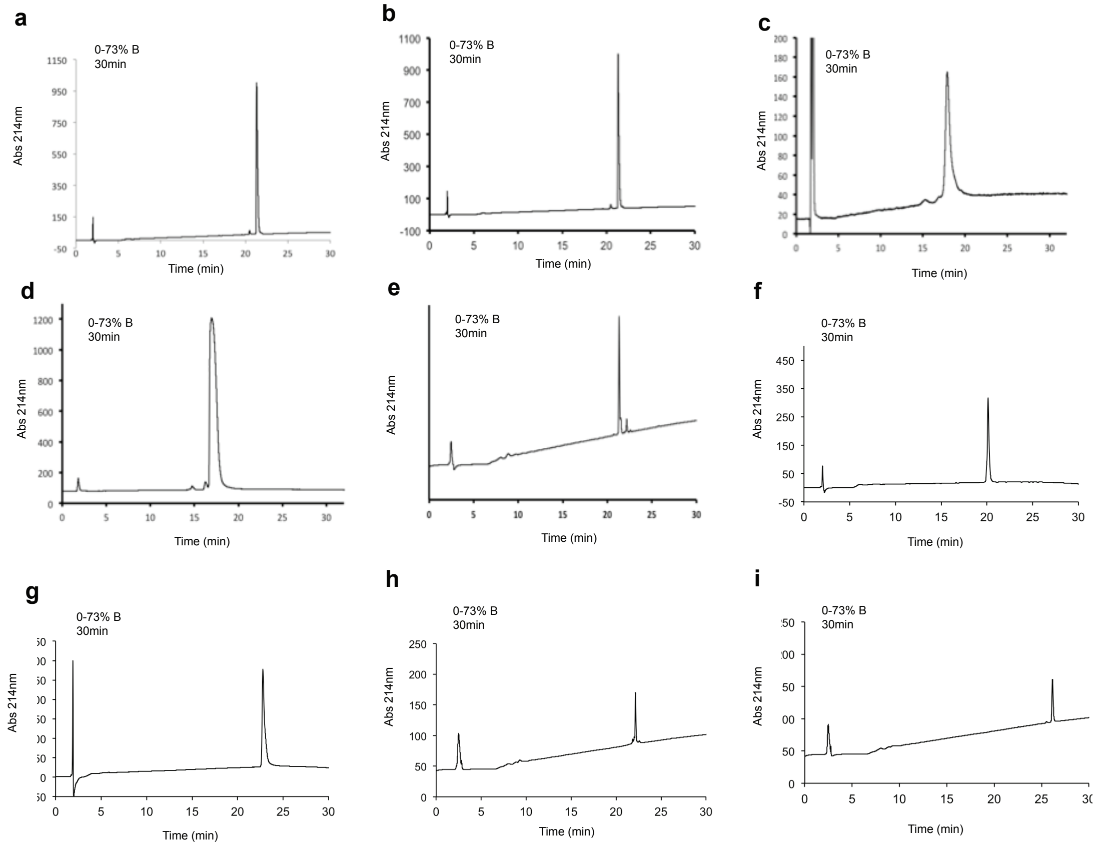


Figure 8.3. RP-HPLC analysis of proteins from Chapter 4. Proteins were run on a analytical C18 column at the indicated gradient and time. (a) Purified uRR-SH (b) Purified uLL-SH (c) Purified H2B-Ub (d) Purified H2B-uLL (e) Purified ^{15}N uLL (f) Purified Hub1-SH (g) Purified H2B$_{ss}$Hub1 (h) Purified Hub1ub-SH (i) Purified H2B$_{ss}$Hub1ub

Table 8.4. Masses of purfied proteins/peptides from Chapter 4

Protein/Peptide	Expected Mass (Da)	Observed Mass (Da)
uRR-SH	8397.71	8400.71
uLL-SH	8483.76	8483.06
H2B(1-116)-Mes	12990.9	12990.3
H2B(117-125)	1116.1	1115.5
Ub75-Mes	8632	8632
uLL-Mes	8547.82	8547.64
H2B(117-125)-Ub	8631.7	8631.7
H2B-Ub	22378	22377.68
H2B(117-125)-uLL	8547.6	8547.5
H2B-uLL	22293.7	22293.75
(^{15}N, ^{1}H) uLL	8584.67	8584.35
Hub1-SH	8586	8588
$H2B_{ss}Hub1$	22376	22379
Hub1ub-SH	8627.07	8626.59
$H2B_{ss}Hub1ub$	22417.98	22416.83

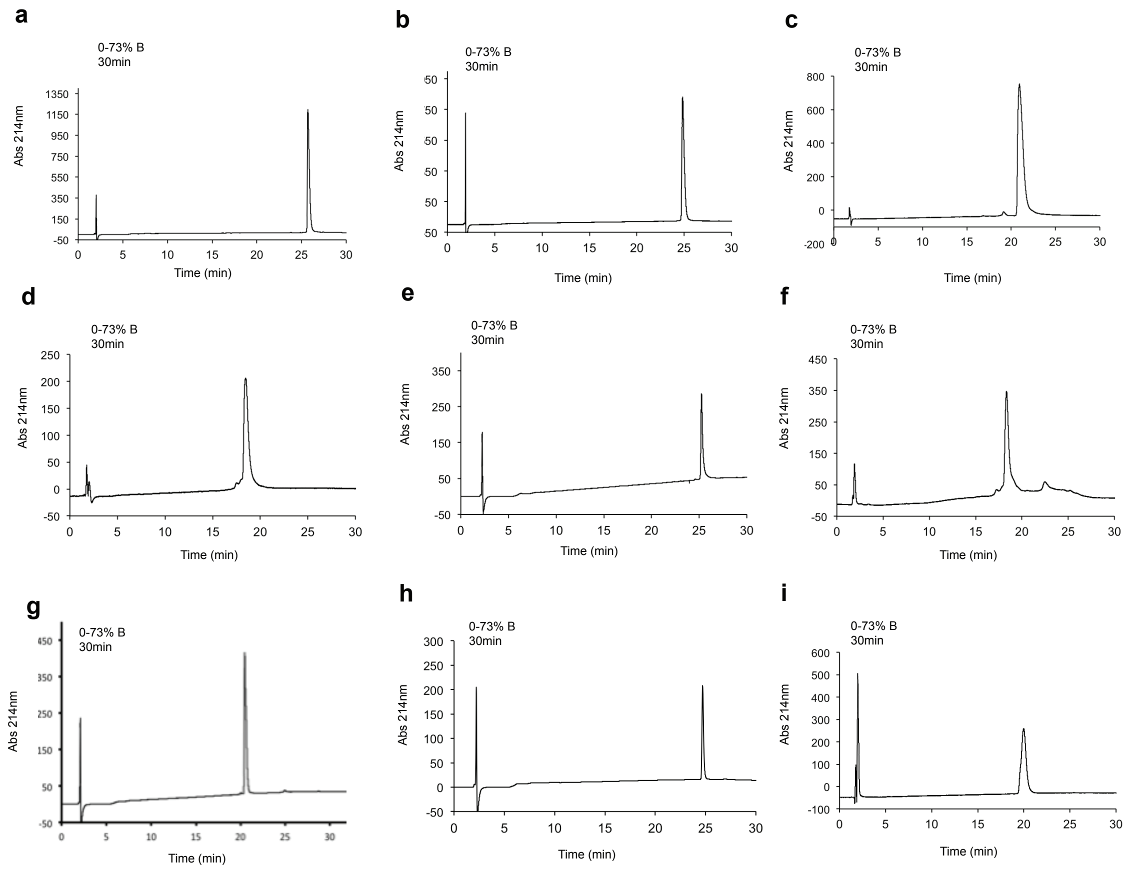


Figure 8.4. RP-HPLC analysis of proteins from Chapter 5. Proteins were run on a analytical C18 column at the indicated gradient and time. (a) Purified dH2A (b) Purified dH2B (c) Purified dH3S110C (d) Purified dH4 (e) Purified dH2BK118C (f) Purified dH2B$_{ss}$Ub (g) Purified Ub$_{ILV}$-SH (h) Purified H2AN110C (i) Purified fH2A

Table 8.5. Masses of purfied proteins from Chapter 5

Protein	Expected Mass (Da)	Observed Mass (Da)
*d*H2A	13232.47	13232.29
*d*H2B	13809.09	13825.52
*d*H3	15240.75	15240.71
*d*H4	11495.51	11493.89
*d*H2BK120C	13785.06	13786.78
d$H2B_{ss}Ub$	22352.43	22356.65
Ub_{ILV}-SH	9011.00	9004.19
H2AN110C	13940	13939.13
fH2A	14368.37	14366.64

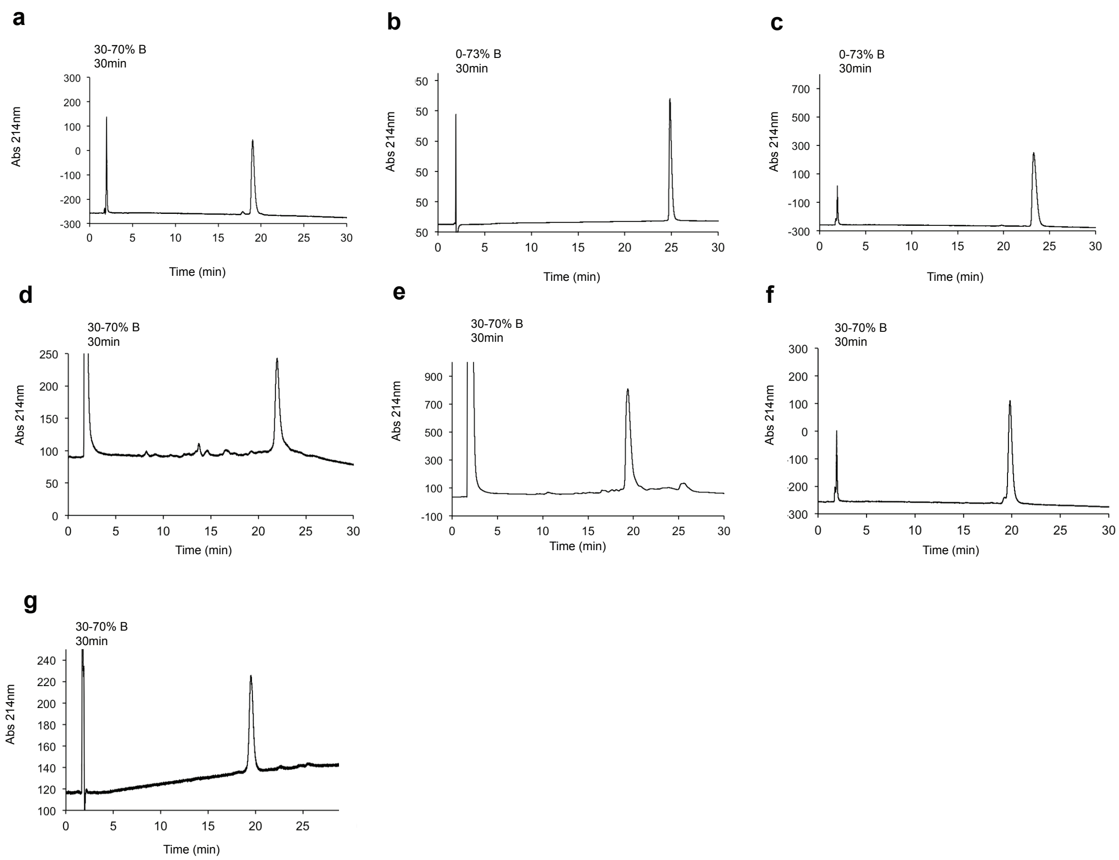


Figure 8.5. RP-HPLC analysis of proteins from Chapter 6. Proteins were run on a analytical C18 column at the indicated gradient and time. (a) Purified yH2A (b) Purified yH2B (c) Purified yH3 (d) Purified yH4 (e) Purified n-yH2A (f) Purified Ub-yH2A (g) Purified uLL-yH2A

Table 8.6. Masses of purfied proteins from Chapter 6

Protein	Expected Mass (Da)	Observed Mass (Da)
*y*H2A	13857.94	13858.41
*y*H2B	14105.01	14105.87
*y*H3	15252.74	15252.36
*y*H4	13989.14	13988.81
n-*y*H2A	20404.60	20403.43
Ub-*y*H2A	20262.44	20261.65
uLL-*y*H2A	20177.27	20176.26

References

[1] Crick, F. H. (1958) On protein synthesis, *Symposia of the Society for Experimental Biology 12*, 138-163.

[2] Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., Emanuelsson, O., Zhang, Z. D., Weissman, S., and Snyder, M. (2007) What is a gene, post-ENCODE? History and updated definition, *Genome Research 17*, 669-681.

[3] Muller, H. J. (1930) Types of visible variations induced by x-rays in Drosophila, *Journal of Genetics 22*, 299-U297.

[4] Hannah, A. (1951) Localization and function of heterochromatin in Drosophila melanogaster, *Advances in Genetics 4*, 87-125.

[5] Gelbart, M. E., and Kuroda, M. I. (2009) Drosophila dosage compensation: a complex voyage to the X chromosome, *Development 136*, 1399-1410.

[6] Bickmore, W. A. (2013) The spatial organization of the human genome, *Annual Review of Genomics and Human Genetics 14*, 67-84.

[7] Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., and Dekker, J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome, *Science 326*, 289-293.

[8] Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature 485*, 376-380.

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

[10] Kornberg, R. D. (1977) Structure of chromatin, *Annual Reveiw of Biochemistry 46*, 931-954.

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

[12] McGhee, J. D., and Felsenfeld, G. (1980) Nucleosome structure, *Annual Reveiw of Biochemistry 49*, 1115-1156.

[13] Yuan, G. C., Liu, Y. J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J., and Rando, O. J. (2005) Genome-scale identification of nucleosome positions in S. cerevisiae, *Science 309*, 626-630.

[14] Valouev, A., Johnson, S. M., Boyd, S. D., Smith, C. L., Fire, A. Z., and Sidow, A. (2011) Determinants of nucleosome organization in primary human cells, *Nature 474*, 516-520.

[15] Allis, C. D., Jenuwein, T, Reinberg, D (2007) *Epigenetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[16] Georgel, P. T., and Hansen, J. C. (2001) Linker histone function in chromatin: dual mechanisms of action, *Biochemistry Cell Biol 79*, 313-316.

[17] Robinson, P. J. J., Fairall, L., Huynh, V. A. T., and Rhodes, D. (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure, *Proceedings of the National Academy of Sciences United States of America 103*, 6506-6511.

[18] Hansen, J. C. (2002) Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions, *Annual Review of Biophysics and Biomolecular Structure 31*, 361-392.

[19] Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014) Chromatin as dynamic 10-nm fibers, *Chromosoma*.

[20] Filion, G. J., van Bemmel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., Brugman, W., de Castro, I. J., Kerkhoven, R. M., Bussemaker, H. J., and van Steensel, B. (2010) Systematic protein location mapping reveals five principal chromatin types in Drosophila cells, *Cell 143*, 212-224.

[21] Keller, W., Konig, P., and Richmond, T. J. (1995) Crystal structure of a bZIP/DNA complex at 2.2 A: determinants of DNA specific recognition, *Journal of Molecular Biology 254*, 657-667.

[22] Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., Vonrhein, C., Moras, D., Romier, C., Bolognesi, M., and Mantovani, R. (2013) Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination, *Cell 152*, 132-143.

[23] Nakahashi, H., Kwon, K. R., Resch, W., Vian, L., Dose, M., Stavreva, D., Hakim, O., Pruett, N., Nelson, S., Yamane, A., Qian, J., Dubois, W., Welsh, S., Phair, R. D., Pugh, B. F., Lobanenkov, V., Hager, G. L., and

Casellas, R. (2013) A genome-wide map of CTCF multivalency redefines the CTCF code, *Cell reports 3*, 1678-1689.

[24] Maher, B. (2012) ENCODE: The human encyclopaedia, *Nature 489*, 46-48.

[25] Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Roder, M., Kokocinski, F., Abdelhamid, R. F., Alioto, T., Antoshechkin, I., Baer, M. T., Bar, N. S., Batut, P., Bell, K., Bell, I., Chakrabortty, S., Chen, X., Chrast, J., Curado, J., Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R., Falconnet, E., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M. J., Gao, H., Gonzalez, D., Gordon, A., Gunawardena, H., Howald, C., Jha, S., Johnson, R., Kapranov, P., King, B., Kingswood, C., Luo, O. J., Park, E., Persaud, K., Preall, J. B., Ribeca, P., Risk, B., Robyr, D., Sammeth, M., Schaffer, L., See, L. H., Shahab, A., Skancke, J., Suzuki, A. M., Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H., Wrobel, J., Yu, Y., Ruan, X., Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T., Reymond, A., Antonarakis, S. E., Hannon, G., Giddings, M. C., Ruan, Y., Wold, B., Carninci, P., Guigo, R., and Gingeras, T. R. (2012) Landscape of transcription in human cells, *Nature 489*, 101-108.

[26] Jeon, Y., and Lee, J. T. (2011) YY1 tethers Xist RNA to the inactive X nucleation center, *Cell 146*, 119-133.

[27] Guttman, M., and Rinn, J. L. (2012) Modular regulatory principles of large non-coding RNAs, *Nature 482*, 339-346.

[28] Yuan, G., and Zhu, B. (2013) Histone variants and epigenetic inheritance, *Biochimica et Biophysica Acta 1819*, 222-229.

[29] Maze, I., Noh, K. M., Soshnev, A. A., and Allis, C. D. (2014) Every amino acid matters: essential contributions of histone variants to mammalian development and disease, *Nature Reviews. Genetics 15*, 259-271.

[30] Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications, *Nature 403*, 41-45.

[31] Suva, M. L., Riggi, N., and Bernstein, B. E. (2013) Epigenetic reprogramming in cancer, *Science 339*, 1567-1570.

[32] Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification, *Cell 146*, 1016-1028.

[33] Kouzarides, T. (2007) Chromatin modifications and their function, *Cell 128*, 693-705.

[34] Tian, Z., Tolic, N., Zhao, R., Moore, R. J., Hengel, S. M., Robinson, E. W., Stenoien, D. L., Wu, S., Smith, R. D., and Pasa-Tolic, L. (2012) Enhanced top-down characterization of histone post-translational modifications, *Genome Biology 13*, R86.

[35] Voigt, P., LeRoy, G., Drury, W. J., 3rd, Zee, B. M., Son, J., Beck, D. B., Young, N. L., Garcia, B. A., and Reinberg, D. (2012) Asymmetrically modified nucleosomes, *Cell 151*, 181-193.

[36] Huff, J. T., Plocik, A. M., Guthrie, C., and Yamamoto, K. R. (2010) Reciprocal intronic and exonic histone modification regions in humans, *Nature Structural & Molecular Biology 17*, 1495-1499.

[37] Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M., and Bernstein, B. E. (2011) Mapping and analysis of chromatin state dynamics in nine human cell types, *Nature 473*, 43-49.

[38] North, J. A., Javaid, S., Ferdinand, M. B., Chatterjee, N., Picking, J. W., Shoffner, M., Nakkula, R. J., Bartholomew, B., Ottesen, J. J., Fishel, R., and Poirier, M. G. (2011) Phosphorylation of histone H3(T118) alters nucleosome dynamics and remodeling, *Nucleic Acids Research 39*, 6465-6474.

[39] Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E. T., Buhler, M., Margueron, R., and Schneider, R. (2013) Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer, *Cell 152*, 859-872.

[40] Tropberger, P., and Schneider, R. (2013) Scratching the (lateral) surface of chromatin regulation by histone modifications, *Nature Structural & Molecular Biology 20*, 657-661.

[41] Simon, M., North, J. A., Shimko, J. C., Forties, R. A., Ferdinand, M. B., Manohar, M., Zhang, M., Fishel, R., Ottesen, J. J., and Poirier, M. G. (2011) Histone fold modifications control nucleosome unwrapping and disassembly, *Proceedings of the National Academy of Sciences of the United States of America 108*, 12711-12716.

[42] Hyland, E. M., Cosgrove, M. S., Molina, H., Wang, D., Pandey, A., Cottee, R. J., and Boeke, J. D. (2005) Insights into the role of histone H3 and histone H4 core modifiable residues in Saccharomyces cerevisiae, *Molecular and Cellular Biology 25*, 10060-10070.

[43] Patel, D. J., and Wang, Z. (2013) Readout of epigenetic modifications, *Annual Reveiw of Biochemistry 82*, 81-118.

[44] Yun, M., Wu, J., Workman, J. L., and Li, B. (2011) Readers of histone modifications, *Cell Research 21*, 564-578.

[45] Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., and Mer, G. (2006) Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair, *Cell 127*, 1361-1373.

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

[47] Garske, A. L., Oliver, S. S., Wagner, E. K., Musselman, C. A., LeRoy, G., Garcia, B. A., Kutateladze, T. G., and Denu, J. M. (2010) Combinatorial profiling of chromatin binding modules reveals multisite discrimination, *Nature Chemical Biology 6*, 283-290.

[48] Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L., and Crabtree, G. R. (1996) Purification and biochemical heterogeneity of the mammalian SWI-SNF complex, *The EMBO Journal 15*, 5370-5382.

[49] Zhao, K., Wang, W., Rando, O. J., Xue, Y., Swiderek, K., Kuo, A., and Crabtree, G. R. (1998) Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling, *Cell 95*, 625-636.

[50] Wu, J. I., Lessard, J., and Crabtree, G. R. (2009) Understanding the words of chromatin regulation, *Cell 136*, 200-206.

[51] Hershko, A., Ciechanover, A., and Varshavsky, A. (2000) Basic Medical Research Award. The ubiquitin system, *Nature Medicine 6*, 1073-1081.

[52] Pickart, C. M. (2001) Mechanisms underlying ubiquitination, *Annual Reveiw of Biochemistry 70*, 503-533.

[53] Komander, D., and Rape, M. (2012) The ubiquitin code, *Annual Reveiw of Biochemistry 81*, 203-229.

[54] Shilatifard, A. (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression, *Annual Reveiw of Biochemistry 75*, 243-269.

[55] Schulman, B. A., and Harper, J. W. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways, *Nature Reviews. Molecular Cell Biology 10*, 319-331.

[56] Harper, J. W., and Schulman, B. A. (2006) Structural complexity in ubiquitin recognition, *Cell 124*, 1133-1136.

[57] Walsh, C. T. (2005) *Posttranslation Modification of Proteins: Expanding Nature's Inventory*, Roberts & Company Publishers, Portland, OR.

[58] Chen, T., Zhou, T., He, B., Yu, H., Guo, X., Song, X., and Sha, J. (2014) mUbiSiDa: A Comprehensive Database for Protein Ubiquitination Sites in Mammals, *PloS one 9*, e85744.

[59] Goldknopf, I. L., and Busch, H. (1977) Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24, *Proceedings of the National Academy of Sciences of the United States of America 74*, 864-868.

[60] West, M. H., and Bonner, W. M. (1980) Histone 2B can be modified by the attachment of ubiquitin, *Nucleic Acids Research 8*, 4671-4680.

[61] Wisniewski, J. R., Zougman, A., Krüger, S., and Mann, M. (2007) Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue, *Molecular Cell Proteomics 6*, 72-87.

[62] Yan, Q., Dutt, S., Xu, R., Graves, K., Juszczynski, P., Manis, J. P., and Shipp, M. A. (2009) BBAP monoubiquitylates histone H4 at lysine 91 and selectively modulates the DNA damage response, *Molecular Cell 36*, 110-120.

[63] Whitcomb, S. J., Fierz, B., McGinty, R. K., Holt, M., Ito, T., Muir, T. W., and Allis, C. D. (2012) Histone monoubiquitylation position determines specificity and direction of enzymatic cross-talk with histone methyltransferases Dot1L and PRC2, *The Journal of Biological Chemistry 287*, 23718-23725.

[64] Weake, V. M., and Workman, J. L. (2008) Histone ubiquitination: triggering gene activity, *Molecular Cell 29*, 653-663.

[65] Zhang, Y. (2003) Transcriptional regulation by histone ubiquitination and deubiquitination, *Genes & Development 17*, 2733-2740.

[66] Osley, M. A. (2004) H2B ubiquitylation: the end is in sight, *Biochimica et Biophysica Acta 1677*, 74-78.

[67] Minsky, N., Shema, E., Field, Y., Schuster, M., Segal, E., and Oren, M. (2008) Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells, *Nature Cell Biology 10*, 483-488.

[68] Jung, I., Kim, S. K., Kim, M., Han, Y. M., Kim, Y. S., Kim, D., and Lee, D. (2012) H2B monoubiquitylation is a 5'-enriched active transcription mark and correlates with exon-intron structure in human cells, *Genome Research 22*, 1026-1035.

[69] Schulze, J. M., Hentrich, T., Nakanishi, S., Gupta, A., Emberly, E., Shilatifard, A., and Kobor, M. S. (2011) Splitting the task: Ubp8 and Ubp10 deubiquitinate different cellular pools of H2BK123, *Genes & Development 25*, 2242-2247.

[70] Ma, M. K., Heath, C., Hair, A., and West, A. G. (2011) Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity, *PLoS Genetics 7*, e1002175.

[71] Shema-Yaacoby, E., Nikolov, M., Haj-Yahya, M., Siman, P., Allemand, E., Yamaguchi, Y., Muchardt, C., Urlaub, H., Brik, A., Oren, M., and Fischle, W. (2013) Systematic identification of proteins binding to chromatin-embedded ubiquitylated H2B reveals recruitment of SWI/SNF to regulate transcription, *Cell Reports 4*, 601-608.

[72] Fuchs, G., and Oren, M. (2014) Writing and reading H2B monoubiquitylation, *Biochimica et Biophysica Acta*.

[73] Henry, K. W., Wyce, A., Lo, W. S., Duggan, L. J., Emre, N. C., Kao, C. F., Pillus, L., Shilatifard, A., Osley, M. A., and Berger, S. L. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8, *Genes & Development 17*, 2648-2663.

[74] Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A., and Reinberg, D. (2006) Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II, *Cell 125*, 703-717.

[75] Weinberger, L., Voichek, Y., Tirosh, I., Hornung, G., Amit, I., and Barkai, N. (2012) Expression noise and acetylation profiles distinguish HDAC functions, *Molecular Cell 47*, 193-202.

[76] Trujillo, K. M., and Osley, M. A. (2012) A role for H2B ubiquitylation in DNA replication, *Molecular Cell 48*, 734-746.

[77] Pirngruber, J., Shchebet, A., Schreiber, L., Shema, E., Minsky, N., Chapman, R. D., Eick, D., Aylon, Y., Oren, M., and Johnsen, S. A. (2009) CDK9 directs H2B monoubiquitination and controls replication-dependent histone mRNA 3'-end processing, *EMBO Reports 10*, 894-900.

[78] Vitaliano-Prunier, A., Babour, A., Herissant, L., Apponi, L., Margaritis, T., Holstege, F. C., Corbett, A. H., Gwizdek, C., and Dargemont, C. (2012) H2B ubiquitylation controls the formation of export-competent mRNP, *Molecular Cell 45*, 132-139.

[79] Chernikova, S. B., Dorth, J. A., Razorenova, O. V., Game, J. C., and Brown, J. M. (2010) Deficiency in Bre1 impairs homologous recombination repair and cell cycle checkpoint response to radiation damage in mammalian cells, *Radiation Research 174*, 558-565.

[80] Moyal, L., Lerenthal, Y., Gana-Weisz, M., Mass, G., So, S., Wang, S. Y., Eppink, B., Chung, Y. M., Shalev, G., Shema, E., Shkedy, D., Smorodinsky, N. I., van Vliet, N., Kuster, B., Mann, M., Ciechanover, A., Dahm-Daphi, J., Kanaar, R., Hu, M. C., Chen, D. J., Oren, M., and Shiloh, Y. (2011) Requirement of ATM-dependent monoubiquitylation of histone H2B for timely repair of DNA double-strand breaks, *Molecular Cell 41*, 529-542.

[81] Nakamura, K., Kato, A., Kobayashi, J., Yanagihara, H., Sakamoto, S., Oliveira, D. V., Shimada, M., Tauchi, H., Suzuki, H., Tashiro, S., Zou, L., and Komatsu, K. (2011) Regulation of homologous recombination by RNF20-dependent H2B ubiquitination, *Molecular Cell 41*, 515-528.

[82] Kari, V., Shchebet, A., Neumann, H., and Johnsen, S. A. (2011) The H2B ubiquitin ligase RNF40 cooperates with SUPT16H to induce dynamic changes in chromatin structure during DNA double-strand break repair, *Cell Cycle 10*, 3495-3504.

[83] Batta, K., Zhang, Z., Yen, K., Goffman, D. B., and Pugh, B. F. (2011) Genome-wide function of H2B ubiquitylation in promoter and genic regions, *Genes & Development 25*, 2254-2265.

[84] Latham, J. A., Chosed, R. J., Wang, S., and Dent, S. Y. (2011) Chromatin signaling to kinetochores: transregulation of Dam1 methylation by histone H2B ubiquitination, *Cell 146*, 709-719.

[85] Kim, J., and Roeder, R. G. (2009) Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast, *The Journal of Biological Chemistry 284*, 20582-20592.

[86] Kim, J., Guermah, M., McGinty, R. K., Lee, J. S., Tang, Z., Milne, T. A., Shilatifard, A., Muir, T. W., and Roeder, R. G. (2009) RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells, *Cell 137*, 459-471.

[87] Robzyk, K., Recht, J., and Osley, M. A. (2000) Rad6-dependent ubiquitination of histone H2B in yeast, *Science 287*, 501-504.

[88] Hwang, W. W., Venkatasubrahmanyam, S., Ianculescu, A. G., Tong, A., Boone, C., and Madhani, H. D. (2003) A conserved RING finger protein required for histone H2B monoubiquitination and cell size control, *Molecular Cell 11*, 261-266.

[89] Wood, A., Krogan, N. J., Dover, J., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J. F., Johnston, M., and Shilatifard, A. (2003) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter, *Molecular Cell 11*, 267-274.

[90] Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2005) The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS, *Molecular Cell 20*, 589-599.

[91] Song, Y. H., and Ahn, S. H. (2010) A Bre1-associated protein, large 1 (Lge1), promotes H2B ubiquitylation during the early stages of transcription elongation, *The Journal of Biological Chemistry 285*, 2361-2367.

[92] Daniel, J. A., Torok, M. S., Sun, Z. W., Schieltz, D., Allis, C. D., Yates, J. R., 3rd, and Grant, P. A. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription, *The Journal of Biological Chemistry 279*, 1867-1871.

[93] Emre, N. C., and Berger, S. L. (2004) Histone H2B ubiquitylation and deubiquitylation in genomic regulation, *Cold Spring Harbor symposia on Quantitative Biology 69*, 289-299.

[94] Gardner, R. G., Nelson, Z. W., and Gottschling, D. E. (2005) Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin, *Molecular and Cellular Biology 25*, 6123-6139.

[95] Lee, K. K., and Workman, J. L. (2007) Histone acetyltransferase complexes: one size doesn't fit all, *Nature Reviews. Molecular Cell Biology 8*, 284-295.

[96] Daniel, J. A., and Grant, P. A. (2007) Multi-tasking on chromatin with the SAGA coactivator complexes, *Mutation Research 618*, 135-148.

[97] Lang, G., Bonnet, J., Umlauf, D., Karmodiya, K., Koffler, J., Stierle, M., Devys, D., and Tora, L. (2011) The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements, *Molecular and Cellular Biology 31*, 3734-3744.

[98] Kahana, A., and Gottschling, D. E. (1999) DOT4 links silencing and cell growth in Saccharomyces cerevisiae, *Molecular and Cellular Biology 19*, 6608-6620.

[99] Emre, N. C., Ingvarsdottir, K., Wyce, A., Wood, A., Krogan, N. J., Henry, K. W., Li, K., Marmorstein, R., Greenblatt, J. F., Shilatifard, A., and Berger, S. L. (2005) Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing, *Molecular Cell 17*, 585-594.

[100] Smith, E., and Shilatifard, A. (2010) The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes, *Molecular Cell 40*, 689-701.

[101] Zhang, Z., Jones, A., Joo, H. Y., Zhou, D., Cao, Y., Chen, S., Erdjument-Bromage, H., Renfrow, M., He, H., Tempst, P., Townes, T. M., Giles, K. E., Ma, L., and Wang, H. (2013) USP49 deubiquitinates histone H2B and regulates cotranscriptional pre-mRNA splicing, *Genes & Development 27*, 1581-1595.

[102] Joo, H. Y., Jones, A., Yang, C., Zhai, L., Smith, A. D. t., Zhang, Z., Chandrasekharan, M. B., Sun, Z. W., Renfrow, M. B., Wang, Y., Chang, C., and Wang, H. (2011) Regulation of histone H2A and H2B deubiquitination and Xenopus development by USP12 and USP46, *The Journal of Biological Chemistry 286*, 7190-7201.

[103] van der Knaap, J. A., Kumar, B. R., Moshkin, Y. M., Langenberg, K., Krijgsveld, J., Heck, A. J., Karch, F., and Verrijzer, C. P. (2005) GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7, *Molecular Cell 17*, 695-707.

[104] Nicassio, F., Corrado, N., Vissers, J. H., Areces, L. B., Bergink, S., Marteijn, J. A., Geverts, B., Houtsmuller, A. B., Vermeulen, W., Di Fiore, P. P., and Citterio, E. (2007) Human USP3 is a chromatin modifier required for S phase progression and genome stability, *Current Biology : CB 17*, 1972-1977.

[105] Xiao, T., Kao, C. F., Krogan, N. J., Sun, Z. W., Greenblatt, J. F., Osley, M. A., and Strahl, B. D. (2005) Histone H2B ubiquitylation is associated with elongating RNA polymerase II, *Molecular and Cellular Biology 25*, 637-651.

[106] Shema, E., Tirosh, I., Aylon, Y., Huang, J., Ye, C., Moskovits, N., Raver-Shapira, N., Minsky, N., Pirngruber, J., Tarcic, G., Hublarova, P., Moyal, L., Gana-Weisz, M., Shiloh, Y., Yarden, Y., Johnsen, S. A., Vojtesek, B., Berger, S. L., and Oren, M. (2008) The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through

selective regulation of gene expression, *Genes & Development 22*, 2664-2676.

[107] Winkler, D. D., and Luger, K. (2011) The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization, *The Journal of Biological Chemistry 286*, 18369-18374.

[108] Formosa, T. (2012) The role of FACT in making and breaking nucleosomes, *Biochimica et Biophysica Acta 1819*, 247-255.

[109] Hondele, M., Stuwe, T., Hassler, M., Halbach, F., Bowman, A., Zhang, E. T., Nijmeijer, B., Kotthoff, C., Rybin, V., Amlacher, S., Hurt, E., and Ladurner, A. G. (2013) Structural basis of histone H2A-H2B recognition by the essential chaperone FACT, *Nature*.

[110] Fleming, A. B., Kao, C. F., Hillyer, C., Pikaart, M., and Osley, M. A. (2008) H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation, *Molecular Cell 31*, 57-66.

[111] Krogan, N. J., Kim, M., Ahn, S. H., Zhong, G., Kobor, M. S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J. F. (2002) RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach, *Molecular and Cellular Biology 22*, 6979-6992.

[112] Squazzo, S. L., Costa, P. J., Lindstrom, D. L., Kumer, K. E., Simic, R., Jennings, J. L., Link, A. J., Arndt, K. M., and Hartzog, G. A. (2002) The Paf1 complex physically and functionally associates with transcription elongation factors in vivo, *The EMBO Journal 21*, 1764-1774.

[113] Oliveira, D. V., Kato, A., Nakamura, K., Ikura, T., Okada, M., Kobayashi, J., Yanagihara, H., Saito, Y., Tauchi, H., and Komatsu, K. (2014) Histone chaperone FACT regulates homologous recombination by chromatin remodeling through interaction with RNF20, *Journal of Cell Science 127*, 763-772.

[114] Lee, J. S., Garrett, A. S., Yen, K., Takahashi, Y. H., Hu, D., Jackson, J., Seidel, C., Pugh, B. F., and Shilatifard, A. (2012) Codependency of H2B monoubiquitination and nucleosome reassembly on Chd1, *Genes & Development 26*, 914-919.

[115] Chandrasekharan, M. B., Huang, F., and Sun, Z. W. (2009) Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability, *Proceedings of the National Academy of Sciences of the United States of America 106*, 16686-16691.

[116] Fierz, B., Kilic, S., Hieb, A. R., Luger, K., and Muir, T. W. (2012) Stability of nucleosomes containing homogenously ubiquitylated H2A and H2B

prepared using semisynthesis, *Journal of the American Chemical Society 134*, 19548-19551.

[117] Fierz, B., Chatterjee, C., McGinty, R. K., Bar-Dagan, M., Raleigh, D. P., and Muir, T. W. (2011) Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction, *Nature Chemical Biology 7*, 113-119.

[118] McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G., and Muir, T. W. (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation, *Nature 453*, 812-816.

[119] Sun, Z. W., and Allis, C. D. (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast, *Nature 418*, 104-108.

[120] Shahbazian, M. D., Zhang, K., and Grunstein, M. (2005) Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1, *Molecular Cell 19*, 271-277.

[121] McGinty, R. K., Kohn, M., Chatterjee, C., Chiang, K. P., Pratt, M. R., and Muir, T. W. (2009) Structure-activity analysis of semisynthetic nucleosomes: mechanistic insights into the stimulation of Dot1L by ubiquitylated histone H2B, *ACS Chemical Biology 4*, 958-968.

[122] Werner, M., and Ruthenburg, A. J. (2011) The United States of histone ubiquitylation and methylation, *Molecular Cell 43*, 5-7.

[123] Ruthenburg, A. J., Li, H., Milne, T. A., Dewell, S., McGinty, R. K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T. W., Patel, D. J., and Allis, C. D. (2011) Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions, *Cell 145*, 692-706.

[124] Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3, *Nature 419*, 407-411.

[125] Ng, H. H., Dole, S., and Struhl, K. (2003) The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B, *The Journal of Biological Chemistry 278*, 33625-33628.

[126] Schubeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J., Bell, S. P., and Groudine, M. (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote, *Genes & Development 18*, 1263-1271.

[127] Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J.,

Lewitter, F., Gifford, D. K., and Young, R. A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast, *Cell 122*, 517-527.

[128] Rosenbloom, K. R., Sloan, C. A., Malladi, V. S., Dreszer, T. R., Learned, K., Kirkup, V. M., Wong, M. C., Maddren, M., Fang, R., Heitner, S. G., Lee, B. T., Barber, G. P., Harte, R. A., Diekhans, M., Long, J. C., Wilder, S. P., Zweig, A. S., Karolchik, D., Kuhn, R. M., Haussler, D., and Kent, W. J. (2013) ENCODE data in the UCSC Genome Browser: year 5 update, *Nucleic Acids Research 41*, D56-63.

[129] Shilatifard, A. (2012) The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis, *Annual Review of Biochemistry 81*, 65-95.

[130] Kim, S. K., Jung, I., Lee, H., Kang, K., Kim, M., Jeong, K., Kwon, C. S., Han, Y. M., Kim, Y. S., Kim, D., and Lee, D. (2012) Human histone H3K79 methyltransferase DOT1L protein [corrected] binds actively transcribing RNA polymerase II to regulate gene expression, *The Journal of Biological Chemistry 287*, 39698-39709.

[131] Wu, M., Wang, P. F., Lee, J. S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. (2008) Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS, *Molecular and Cellular Biology 28*, 7337-7344.

[132] Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation, *Molecular Cell 11*, 721-729.

[133] Lee, J. S., Shukla, A., Schneider, J., Swanson, S. K., Washburn, M. P., Florens, L., Bhaumik, S. R., and Shilatifard, A. (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS, *Cell 131*, 1084-1096.

[134] Takahashi, Y. H., Schulze, J. M., Jackson, J., Hentrich, T., Seidel, C., Jaspersen, S. L., Kobor, M. S., and Shilatifard, A. (2011) Dot1 and histone H3K79 methylation in natural telomeric and HM silencing, *Molecular Cell 42*, 118-126.

[135] Tatum, D., and Li, S. (2011) Evidence that the histone methyltransferase Dot1 mediates global genomic repair by methylating histone H3 on lysine 79, *The Journal of Biological Chemistry 286*, 17530-17535.

[136] Schulze, J. M., Jackson, J., Nakanishi, S., Gardner, J. M., Hentrich, T., Haug, J., Johnston, M., Jaspersen, S. L., Kobor, M. S., and Shilatifard, A. (2009) Linking cell cycle to histone modifications: SBF and H2B

monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation, *Molecular Cell 35*, 626-641.

[137] Mohan, M., Herz, H. M., Takahashi, Y. H., Lin, C., Lai, K. C., Zhang, Y., Washburn, M. P., Florens, L., and Shilatifard, A. (2010) Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom), *Genes & Development 24*, 574-589.

[138] Lu, X., Simon, M. D., Chodaparambil, J. V., Hansen, J. C., Shokat, K. M., and Luger, K. (2008) The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure, *Nature Structural & Molecular Biology 15*, 1122-1124.

[139] Wakeman, T. P., Wang, Q., Feng, J., and Wang, X. F. (2012) Bat3 facilitates H3K79 dimethylation by DOT1L and promotes DNA damage-induced 53BP1 foci at G1/G2 cell-cycle phases, *The EMBO Journal 31*, 2169-2181.

[140] Wysocki, R., Javaheri, A., Allard, S., Sha, F., Cote, J., and Kron, S. J. (2005) Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkpoint functions of Rad9, *Molecular and Cellular Biology 25*, 8430-8443.

[141] Kim, W., Choi, M., and Kim, J. E. (2014) The histone methyltransferase Dot1/DOT1L as a critical regulator of the cell cycle, *Cell Cycle 13*, 726-738.

[142] Sabra, M., Texier, P., El Maalouf, J., and Lomonte, P. (2013) The Tudor protein survival motor neuron (SMN) is a chromatin-binding protein that interacts with methylated lysine 79 of histone H3, *Journal of Cell Science 126*, 3664-3677.

[143] van Leeuwen, F., Gafken, P. R., and Gottschling, D. E. (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core, *Cell 109*, 745-756.

[144] Richon, V. M., Johnston, D., Sneeringer, C. J., Jin, L., Majer, C. R., Elliston, K., Jerva, L. F., Scott, M. P., and Copeland, R. A. (2011) Chemogenetic analysis of human protein methyltransferases, *Chemical Biology & Drug Design 78*, 199-210.

[145] Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R. M. (2003) Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase, *Cell 112*, 711-723.

[146] Sawada, K., Yang, Z., Horton, J. R., Collins, R. E., Zhang, X., and Cheng, X. (2004) Structure of the conserved core of the yeast Dot1p, a

nucleosomal histone H3 lysine 79 methyltransferase, *The Journal of Biological Chemistry 279*, 43296-43306.

[147] Fingerman, I. M., Li, H. C., and Briggs, S. D. (2007) A charge-based interaction between histone H4 and Dot1 is required for H3K79 methylation and telomere silencing: identification of a new trans-histone pathway, *Genes & Development 21*, 2018-2029.

[148] Oh, S., Jeong, K., Kim, H., Kwon, C. S., and Lee, D. (2010) A lysine-rich region in Dot1p is crucial for direct interaction with H2B ubiquitylation and high level methylation of H3K79, *Biochemical and Biophysical Research Communications 399*, 512-517.

[149] Mueller, D., Bach, C., Zeisig, D., Garcia-Cuellar, M. P., Monroe, S., Sreekumar, A., Zhou, R., Nesvizhskii, A., Chinnaiyan, A., Hess, J. L., and Slany, R. K. (2007) A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification, *Blood 110*, 4445-4454.

[150] Mueller, D., Garcia-Cuellar, M. P., Bach, C., Buhl, S., Maethner, E., and Slany, R. K. (2009) Misguided transcriptional elongation causes mixed lineage leukemia, *PLoS Biology 7*, e1000249.

[151] Bitoun, E., Oliver, P. L., and Davies, K. E. (2007) The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling, *Human Molecular Genetics 16*, 92-106.

[152] Chatterjee, C., McGinty, R. K., Fierz, B., and Muir, T. W. (2010) Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation, *Nature Chemical Biology 6*, 267-269.

[153] Lewis, P. W., Muller, M. M., Koletsky, M. S., Cordero, F., Lin, S., Banaszynski, L. A., Garcia, B. A., Muir, T. W., Becher, O. J., and Allis, C. D. (2013) Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma, *Science 340*, 857-861.

[154] Krivtsov, A. V., and Armstrong, S. A. (2007) MLL translocations, histone modifications and leukaemia stem-cell development, *Nature Review Cancer 7*, 823-833.

[155] Slany, R. K. (2009) The molecular biology of mixed lineage leukemia, *Haematologica 94*, 984-993.

[156] Davies, N., and Lindsey, G. G. (1994) Histone H2B (and H2A) ubiquitination allows normal histone octamer and core particle reconstitution, *Biochimica et biophysica acta 1218*, 187-193.

[157] Wu, L., Zee, B. M., Wang, Y., Garcia, B. A., and Dou, Y. (2011) 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 43*, 132-144.

[158] Krieger, D. E., Levine, R., Merrifield, R. B., Vidali, G., and Allfrey, V. G. (1974) Chemical studies of histone acetylation. Substrate specificity of a histone deacetylase from calf thymus nuclei, *The Journal of Biological Chemistry 249*, 332-334.

[159] Fierz, B., and Muir, T. W. (2012) Chromatin as an expansive canvas for chemical biology, *Nature Chemical Biology 8*, 417-427.

[160] Bua, D. J., Kuo, A. J., Cheung, P., Liu, C. L., Migliori, V., Espejo, A., Casadio, F., Bassi, C., Amati, B., Bedford, M. T., Guccione, E., and Gozani, O. (2009) Epigenome microarray platform for proteome-wide dissection of chromatin-signaling networks, *PloS One 4*, e6789.

[161] Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation, *Science 266*, 776-779.

[162] Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: A general method for protein engineering, *Proceedings of the National Academy of Sciences of the United States of America 95*, 6705-6710.

[163] Casadio, F., Lu, X., Pollock, S. B., LeRoy, G., Garcia, B. A., Muir, T. W., Roeder, R. G., and Allis, C. D. (2013) H3R42me2a is a histone modification with positive transcriptional effects, *Proceedings of the National Academy of Sciences of the United States of America 110*, 14894-14899.

[164] Shah, N. H., and Muir, T. W. (2011) Split Inteins: Nature's Protein Ligases, *Israel Journal of Chemistry 51*, 854-861.

[165] Shah, N. H., Dann, G. P., Vila-Perello, M., Liu, Z., and Muir, T. W. (2012) Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering, *Journal of the American Chemical Society 134*, 11338-11341.

[166] Chatterjee, C., McGinty, R. K., Pellois, J. P., and Muir, T. W. (2007) Auxiliary-mediated site-specific peptide ubiquitylation, *Angewandte Chemie 46*, 2814-2818.

[167] Siman, P., Karthikeyan, S. V., Nikolov, M., Fischle, W., and Brik, A. (2013) Convergent chemical synthesis of histone H2B protein for the site-specific ubiquitination at Lys34, *Angewandte Chemie 52*, 8059-8063.

[168] Ajish Kumar, K. S., Haj-Yahya, M., Olschewski, D., Lashuel, H. A., and Brik, A. (2009) Highly efficient and chemoselective peptide ubiquitylation, *Angewandte Chemie 48*, 8090-8094.

[169] Haj-Yahya, M., Eltarteer, N., Ohayon, S., Shema, E., Kotler, E., Oren, M., and Brik, A. (2012) N-methylation of isopeptide bond as a strategy to resist deubiquitinases, *Angewandte Chemie 51*, 11535-11539.

[170] Davis, L., and Chin, J. W. (2012) Designer proteins: applications of genetic code expansion in cell biology, *Nature reviews. Molecular Cell Biology 13*, 168-182.

[171] Kang, T. J., Yuzawa, S., and Suga, H. (2008) Expression of histone H3 tails with combinatorial lysine modifications under the reprogrammed genetic code for the investigation on epigenetic markers, *Chemistry & Biology 15*, 1166-1174.

[172] Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annual Review of Biophysics and Biomolecular Structure 35*, 225-249.

[173] Li, X., Fekner, T., Ottesen, J. J., and Chan, M. K. (2009) A pyrrolysine analogue for site-specific protein ubiquitination, *Angewandte Chemie 48*, 9184-9187.

[174] Nguyen, D. P., Elliott, T., Holt, M., Muir, T. W., and Chin, J. W. (2011) Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation, *Journal of the American Chemical Society 133*, 11418-11421.

[175] Simon, M. D., Chu, F., Racki, L. R., de la Cruz, C. C., Burlingame, A. L., Panning, B., Narlikar, G. J., and Shokat, K. M. (2007) The site-specific installation of methyl-lysine analogs into recombinant histones, *Cell 128*, 1003-1012.

[176] Li, F., Allahverdi, A., Yang, R., Lua, G. B., Zhang, X., Cao, Y., Korolev, N., Nordenskiold, L., and Liu, C. F. (2011) A direct method for site-specific protein acetylation, *Angewandte Chemie 50*, 9611-9614.

[177] Huang, R., Holbert, M. A., Tarrant, M. K., Curtet, S., Colquhoun, D. R., Dancy, B. M., Dancy, B. C., Hwang, Y., Tang, Y., Meeth, K., Marmorstein, R., Cole, R. N., Khochbin, S., and Cole, P. A. (2010) Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation, *Journal of the American Chemical Society 132*, 9986-9987.

[178] Chen, J., Ai, Y., Wang, J., Haracska, L., and Zhuang, Z. (2010) Chemically ubiquitylated PCNA as a probe for eukaryotic translesion DNA synthesis, *Nature Chemical Biology 6*, 270-272.

[179] Fradet-Turcotte, A., Canny, M. D., Escribano-Diaz, C., Orthwein, A., Leung, C. C., Huang, H., Landry, M. C., Kitevski-Leblanc, J., Noordermeer, S. M., Sicheri, F., and Durocher, D. (2013) 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark, *Nature*.

[180] Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., and Luger, K. (2004) Reconstitution of nucleosome core particles from recombinant histones and DNA, *Methods in Enzymology 375*, 23-44.

[181] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *Journal of Molecular Biology 372*, 774-797.

[182] Winget, J. M., and Mayor, T. (2010) The diversity of ubiquitin recognition: hot spots and varied specificity, *Molecular Cell 38*, 627-635.

[183] Lange, O. F., Lakomek, N. A., Fares, C., Schroder, G. F., Walter, K. F., Becker, S., Meiler, J., Grubmuller, H., Griesinger, C., and de Groot, B. L. (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution, *Science 320*, 1471-1475.

[184] Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M., and Hicke, L. (2001) Distinct functional surface regions on ubiquitin, *The Journal of Biological Chemistry 276*, 30483-30489.

[185] Kiel, C., and Serrano, L. (2006) The ubiquitin domain superfold: structure-based sequence alignments and characterization of binding epitopes, *Journal of Molecular Biology 355*, 821-844.

[186] Wlodarski, T., and Zagrovic, B. (2009) Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin, *Proceedings of the National Academy of Sciences of the United States of America 106*, 19346-19351.

[187] Robinson, P. J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R. G., and Rhodes, D. (2008) 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction, *Journal of Molecular Biology 381*, 816-825.

[188] Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions, *Science 311*, 844-847.

[189] McGinty, R. K. (2010) Mechanistic insights into the stimulation of Dot1L -mediated methylation of histone H3 by semisynthetically ubiquitylated histone H2B, pp xi, 195 leaves, Rockefeller University.

[190] Kato, H., van Ingen, H., Zhou, B. R., Feng, H., Bustin, M., Kay, L. E., and Bai, Y. (2011) Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR, *Proceedings of the National Academy of Sciences of the United States of America 108*, 12283-12288.

[191] Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proceedings of the National Academy of Sciences of the United States of America 94*, 12366-12371.

[192] Fernandez, C., and Wider, G. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules, *Current opinion in Structural Biology 13*, 570-580.

[193] Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E., and Kay, L. E. (2003) Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes, *Journal of the American Chemical Society 125*, 10420-10428.

[194] Sprangers, R., and Kay, L. E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR, *Nature 445*, 618-622.

[195] Gelis, I., Bonvin, A. M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., and Kalodimos, C. G. (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR, *Cell 131*, 756-769.

[196] Tugarinov, V., Kanelis, V., and Kay, L. E. (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy, *Nature Protocols 1*, 749-754.

[197] Tang, Z., Chen, W. Y., Shimada, M., Nguyen, U. T., Kim, J., Sun, X. J., Sengoku, T., McGinty, R. K., Fernandez, J. P., Muir, T. W., and Roeder, R. G. (2013) SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53, *Cell 154*, 297-310.

[198] Chandrasekharan, M. B., Huang, F., Chen, Y. C., and Sun, Z. W. (2010) Histone H2B C-terminal helix mediates trans-histone H3K4 methylation independent of H2B ubiquitination, *Molecular and Cellular Biology 30*, 3216-3232.

[199] Zheng, S., Wyrick, J. J., and Reese, J. C. (2010) Novel trans-tail regulation of H2B ubiquitylation and H3K4 methylation by the N terminus of histone H2A, *Molecular and Cellular Biology 30*, 3635-3645.

[200] Amerik, A. Y., Li, S. J., and Hochstrasser, M. (2000) Analysis of the deubiquitinating enzymes of the yeast Saccharomyces cerevisiae, *Biological Chemistry 381*, 981-992.

[201] Kim, J., Kim, J. A., McGinty, R. K., Nguyen, U. T., Muir, T. W., Allis, C. D., and Roeder, R. G. (2013) The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation, *Molecular Cell 49*, 1121-1133.

[202] Chalker, J. M., Lercher, L., Rose, N. R., Schofield, C. J., and Davis, B. G. (2012) Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications, *Angewandte Chemie 51*, 1835-1839.

[203] Valkevich, E. M., Guenette, R. G., Sanchez, N. A., Chen, Y. C., Ge, Y., and Strieter, E. R. (2012) Forging isopeptide bonds using thiol-ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases, *Journal of the American Chemical Society 134*, 6916-6919.

[204] Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proceedings of the National Academy of Sciences of the United States of America 99*, 19-24.

[205] Shah, N. H., Vila-Perello, M., and Muir, T. W. (2011) Kinetic control of one-pot trans-splicing reactions by using a wild-type and designed split intein, *Angewandte Chemie 50*, 6511-6515.

[206] Kim, K., Lee, B., Kim, J., Choi, J., Kim, J. M., Xiong, Y., Roeder, R. G., and An, W. (2013) Linker Histone H1.2 cooperates with Cul4A and PAF1 to drive H4K31 ubiquitylation-mediated transactivation, *Cell Reports 5*, 1690-1703.

[207] Clore, G. M., Tang, C., and Iwahara, J. (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement, *Current Opinion in Structural Biology 17*, 603-616.

[208] Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre, *Nature 436*, 138-141.

[209] Wu, L., Lee, S. Y., Zhou, B., Nguyen, U. T., Muir, T. W., Tan, S., and Dou, Y. (2013) ASH2L regulates ubiquitylation signaling to MLL: trans-regulation of H3 K4 methylation in higher eukaryotes, *Molecular Cell 49*, 1108-1120.

[210] Yang, L., Lin, C., Jin, C., Yang, J. C., Tanasa, B., Li, W., Merkurjev, D., Ohgi, K. A., Meng, D., Zhang, J., Evans, C. P., and Rosenfeld, M. G.

(2013) lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs, *Nature 500*, 598-602.

[211] Wang, R., Islam, K., Liu, Y., Zheng, W., Tang, H., Lailler, N., Blum, G., Deng, H., and Luo, M. (2013) Profiling genome-wide chromatin methylation with engineered posttranslation apparatus within living cells, *Journal of the American Chemical Society 135*, 1048-1056.

[212] Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse, *Nucleic Acids Research 40*, D261-270.

[213] Bryksin, A. V., and Matsumura, I. (2010) Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids, *BioTechniques 48*, 463-465.

[214] Owen-Hughes, T., Utley, R. T., Steger, D. J., West, J. M., John, S., Cote, J., Havas, K. M., and Workman, J. L. (1999) Analysis of nucleosome disruption by ATP-driven chromatin remodeling complexes, *Methods in Molecular Biology 119*, 319-331.

[215] Ito, M., Yuan, C. X., Malik, S., Gu, W., Fondell, J. D., Yamamura, S., Fu, Z. Y., Zhang, X., Qin, J., and Roeder, R. G. (1999) Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators, *Molecular Cell 3*, 361-370.

[216] An, W., Kim, J., and Roeder, R. G. (2004) Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53, *Cell 117*, 735-748.

[217] Lee, S. Y., Pullen, L., Virgil, D. J., Castaneda, C. A., Abeykoon, D., Bolon, D. N., and Fushman, D. (2014) Alanine Scan of Core Positions in Ubiquitin Reveals Links between Dynamics, Stability, and Function, *Journal of Molecular Biology 426*, 1377-1389.