

**Genome-wide analysis of chromatin modification patterns
and their functional associations with major cellular
processes in *Saccharomyces cerevisiae***

by

Julia Maria Schulze

Diplom Biologin, Friedrich-Schiller Universität Jena, Germany, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES
(Genetics)

The University Of British Columbia
(Vancouver)

October 2010

© Julia Maria Schulze, 2010

Abstract

Chromatin is a nucleoprotein complex packaging DNA inside the cell nucleus and is of crucial relevance for genome regulation. Its structure is highly dynamic undergoing post-translational modifications, replacement by histone variants, and ATP-dependent remodelling. My dissertation aims to better understand the regulation of chromatin by studying the structure and function of a chromatin modifier, mapping chromatin modifications at a genome-wide scale, and linking modification patterns to cellular functions in the model organism *Saccharomyces cerevisiae*.

Multiple chromatin modifying and transcription complexes contain a YEATS domain, and their misregulation has been implicated in the development of cancer. This study recognizes the evolutionary conservation of the YEATS domain from yeast to human, presents its structural composition as well as its function in depositing the histone variant H2A.Z.

Besides histone variants, histone modifications determine chromatin structure and often co-occur in certain genomic regions. Histone H3 lysine 79 methylation is one such modification that adds an additional level of complexity being either mono-, di- or trimethylated (H3K79me1, me2, me3). In this work, I show that these methylation states are functionally not redundant as previously proposed, and that H3K79 di- and trimethylation are associated with different regions of the genome. In contrast to H3K79me3, H3K79me2 marks M/G1 cell cycle regulated genes and its levels change during the cell cycle.

The trigger for trimethylation of H3K79, as well as for trimethylation of lysine 4 on histone H3 (H3K4me3), is monoubiquitination of histone H2B (H2BK123ub). The map of H2BK123ub that I present herein demonstrates its role as upstream

regulator for H3K79me3 and H3K4me3 on a genome-wide scale.

Removal of the transient H2BK123ub mark is facilitated by the deubiquitinases Ubp8 and Ubp10. I herein reveal that they mainly act at distinct genomic loci. While Ubp8 removes H2BK123ub at sites enriched for H3K4me3, Ubp10 functions at those marked by H3K79me3.

Finally, my thesis describes chromatin signatures of multiple modifications, and finds the combination of H2BK123ub, H3K4me3, and H3K36me3 to be specific for highly transcribed genes, including those containing introns. In this context, I show evidence that link the histone H2B ubiquitin ligase Bre1 to mRNA splicing.

Table of Contents

Abstract	ii
Table of Contents	iv
List of Tables	viii
List of Figures	ix
Acknowledgements	xii
Dedication	xiv
Co-authorship Statement	xv
1 Introduction	1
1.1 Chromatin – History and implications	2
1.2 The structural hierarchy of chromatin	2
1.3 The dynamic modifications of chromatin	4
1.3.1 The histone variant H2A.Z	6
1.3.2 Histone acetylation	9
1.3.3 Histone methylation	10
1.3.4 Histone monoubiquitination	14
1.4 Domains as writer, reader, and eraser modules	15
1.5 Transcriptional regulation in the context of chromatin	17
Bibliography	22

2	Asf1-like structure of the conserved Yaf9 YEATS domain and role in H2A.Z deposition and acetylation	48
2.1	Introduction	48
2.2	Results	50
2.2.1	YEATS domain function was conserved from yeast to human	50
2.2.2	The YEATS domain of Yaf9 adopted an immunoglobulin fold	50
2.2.3	Structural features of the YEATS domain	53
2.2.4	Yaf9 shared structural and biochemical properties with the Asf1 histone chaperone	54
2.2.5	Conserved residues on the YEATS domain surface were important for Yaf9 function	57
2.2.6	Yaf9 YEATS domain functioned in both SWR1-C and NuA4	64
2.3	Discussion	68
2.4	Experimental procedures	71
2.4.1	Yeast strains, plasmids, and yeast techniques	71
2.4.2	Analytical-scale affinity purifications	79
2.4.3	Protein expression and purification	79
2.4.4	Crystallization and structure determination	80
2.4.5	GST-fusion protein purifications and histone binding assays	81
2.4.6	Chromatin association and histone acetylation assays . . .	82
2.4.7	Genome-wide ChIP-on-chip	82
2.4.8	Data analysis	83
	Bibliography	84
3	Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation	90
3.1	Introduction	90
3.2	Results	93
3.2.1	Swi4 and Swi6 were required for normal levels of H3K79 di- but not trimethylation	93
3.2.2	Swi4 and Swi6 were not required for H2B monoubiquitination	95

3.2.3	Establishment of H3K79 dimethylation was cell-cycle dependent	97
3.2.4	Genome-wide localizations of H3K79me2 and H3K79me3 were mutually exclusive	100
3.2.5	Differential association of H3K79 di- and trimethylation with promoters and ORFs	103
3.2.6	H3K79 di- and trimethylation associated with genes that tend to be transcriptionally less active	104
3.2.7	M/G1-regulated genes were significantly enriched for H3K79me2	106
3.2.8	Genome-wide association of H3K79 di- but not trimethylation was altered during the cell cycle	108
3.2.9	H3K79me2 was found in intergenic regions and covered autonomously replicating sequences in G2/M-arrested cells	110
3.2.10	Swi4-regulated genes were significantly enriched for H3K79me2 in their promoter region during G2/M phase .	113
3.2.11	Genome-wide colocalization of H2BK123 monoubiquitination with H3K79 tri- but not dimethylation	116
3.2.12	Nrm1 and Whi3 were required for the loss of H3K79 dimethylation in the S phase of the cell cycle	120
3.3	Discussion	120
3.4	Experimental procedures	124
3.4.1	Yeast strains	124
3.4.2	Global proteomic analysis of histone modifications	124
3.4.3	RT-PCR	124
3.4.4	Cell synchronization and FACS analysis	125
3.4.5	ChIP and genome-wide ChIP-on-chip	125
3.4.6	Data analysis	125
3.4.7	Analysis of published genome-wide expression and binding data sets	126
3.4.8	Statistics	126
	Bibliography	127

4	Genome-wide analysis of H2B ubiquitination and deubiquitination in the context of histone H3 methylation	133
4.1	Introduction	133
4.2	Results	136
4.2.1	Genome-wide association of H2BK123 monoubiquitination and histone H3K4- and H3K79 trimethylation	136
4.2.2	Site-specific removal of H2BK123ub by its deubiquitinases Ubp8 and Ubp10	141
4.2.3	Genome-wide combinatorial pattern analysis of H2BK123ub and different H3 methylation states	144
4.2.4	Specific gene regions were characterized by distinct histone H3 methylation and H2B monoubiquitination patterns	146
4.2.5	H2BK123ub, H3K4me3, and H3K36me3 were associated with highly transcribed genes	151
4.2.6	Functional genomic analysis of intron-containing genes	152
4.3	Discussion	157
4.4	Experimental procedures	163
4.4.1	Yeast strains	163
4.4.2	Chromatin immunoprecipitation and genome-wide ChIP-on-chip	163
4.4.3	Data analysis	164
4.4.4	Total RNA extraction, cDNA synthesis and array hybridization	165
	Bibliography	167
5	Discussion	176
	Bibliography	185

List of Tables

Table 1.1	Responsible enzymes for covalent histone modifications and variants studied in this dissertation	20
Table 2.1	Data collection, phasing analysis, and refinement statistics . . .	52
Table 2.2	Yeast strains used in this study	72
Table 2.3	<i>YAF9</i> mutants	76
Table 2.4	<i>yaf9</i> mutant phenotypes	77
Table 4.1	H2B monoubiquitination and H3 methylation were mainly associated with genes transcribed by RNA polymerase II	145
Table 4.2	Histone modification patterns occurred with different frequencies	150
Table 4.3	Combinations of histone modifications were associated with genes bound by certain transcription factors	156
Table 4.4	Yeast strains used in this study	164

List of Figures

Figure 1.1	Chromatin-modifying complexes associated with transcription in <i>S. cerevisiae</i>	19
Figure 2.1	YEATS domain function was conserved from yeast and human	51
Figure 2.2	Secondary structure and amino acid conservation of YEATS domain	55
Figure 2.3	Yaf9 YEATS domain structure was similar to histone chaperone Asf1	58
Figure 2.4	Similarity between the topology of Yaf9 YEATS domain and histone chaperone Asf1	60
Figure 2.5	Mutations in conserved surface residues in the Yaf9 YEATS domain affected protein function	61
Figure 2.6	Growth comparison of Yaf9 mutants that did not have phenotypes nor genetic interaction with Asf1	63
Figure 2.7	Protein levels and overexpression of Yaf9 mutants	64
Figure 2.8	Yaf9 YEATS domain was required for H2A.Z chromatin deposition and acetylation	65
Figure 2.9	Yaf9 YEATS domain was required for H2A.Z chromatin deposition	66
Figure 2.10	Yaf9 YEATS domain mutants bound to histones H3 and H4 <i>in vitro</i>	70
Figure 3.1	GPS reveals that <i>SWI4</i> and <i>SWI6</i> are required for H3K79 di- but not trimethylation	94
Figure 3.2	Role of <i>SWI4</i> and <i>SWI6</i> in H3K79 methylation	96

Figure 3.3	Cell-cycle regulation of H3K79 dimethylation	98
Figure 3.4	High-resolution profile of H3K79me2 and H3K79me3 across the yeast genome with global occupancy analysis	101
Figure 3.5	H3K79me2 and H3K79me3 antibodies were specific	103
Figure 3.6	H3K79me2 was enriched at promoters	105
Figure 3.7	Functional characterization of H3K79-dimethylated genes . . .	107
Figure 3.8	Genome-wide H3K79me2 and H3K79me3 profile in G2/M- arrested cells	109
Figure 3.9	Genome-wide characteristics of H3K79me2-enriched genes in G2/M	111
Figure 3.10	Key cell cycle regulators and histone genes were enriched for H3K79me2	114
Figure 3.11	Genome-wide H3K79me3 profile in G1-arrested cells was similar to H3K79me3 profile in G2/M- arrested cells	115
Figure 3.12	Genome-wide profile of H2B monoubiquitination demon- strates an association with H3K79 tri, but not dimethylation . .	118
Figure 4.1	Genome-wide association of H2BK123 monoubiquitination and H3K4- and H3K79 trimethylation	137
Figure 4.2	Average profiles of H2BK123ub, H3K4me3, H3K36me3, H3K79me3, and H3K79me2 enriched ORFs	139
Figure 4.3	Enrichment of histone modifications across all genes	140
Figure 4.4	Site-specific removal of H2BK123ub by Ubp8 and Ubp10 . . .	141
Figure 4.5	Genome-wide combinatorial pattern analysis of H2BK123ub and different H3 methylation states	143
Figure 4.6	Specific gene regions were characterized by distinct histone H3 methylation and H2B monoubiquitination patterns	147
Figure 4.7	H2BK123ub, H3K4me3, and H3K36me3 were associated with highly transcribed genes	148
Figure 4.8	Intron-containing genes were significantly enriched for co- occurrence of H2BK123ub, H3K4me3, and H3K36me3	153
Figure 4.9	Spliceosome recruitment to chromatin was not affected in <i>bre1</i> deletion strain	155

Figure 4.10	H2BK123ub overlapped strongest with H3K4me3 and H3K79me3 in ORFs with higher transcriptional frequencies .	159
Figure 4.11	Model depicting the circuitry of H2BK123 monoubiquitination and its downstream marks H3K4me3 and H3K79me3 . .	161

Acknowledgements

First of all, I want to thank my supervisor Dr. Michael Kobor. It has been an honour to be his first Ph.D. student. I appreciate all his contributions of ideas, time, and funding to make my Ph.D. time productive and successful. The enthusiasm and joy he has for research was truly motivational for me, even during tough times. He has taught me how to do good science and push the limits at the cutting-edge of the field.

For valuable help during the last years, I would also like to thank my supervisory committee members: Dr. Phil Hieter, Dr. Matthew Lorincz, Dr. Eldon Emberly, and Dr. Raphael Gottardo for their time, interest, and helpful comments. A special thanks merits our collaborator Dr. Ali Shilatifard who opened many scientific doors for me. The level of excitement and energy he has is incredible.

Thanks to all members of the Kobor team for making the lab such a great place to be at. Thanks for all the discussions on science, the supportive team spirit, and especially for many fun times. In addition, I truly enjoyed the wonderful work environment at the CMMT, the team play between labs at the scientific level and beyond especially in the battle for the CMMT cup.

The Genetics Program made the Ph.D. process straightforward and really enjoyable due to ideals and efforts of Dr. Hugh Brock. Thanks to him and Monica Deutsch for support with all organizational needs.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by a fellowship of the German Academic Exchange Service (DAAD) as well as by a graduate scholarship of the Child and Family Research Institute (CFRI).

Lastly, I would like to thank my family for all their love and encouragement, especially my parents who supported me in all my pursuits even if it meant to be far away from home.

And most of all I would like to thank Thomas, the most amazing person in my life. His loving, supportive, and patient attitude encourages me every single day, and this thesis would not have been possible without his tremendous support in all aspects. I am deeply grateful and I am looking forward to the next steps on our journey through this great life. Thank you!

*To my parents
and grandparents*

Co-authorship Statement

A version of Chapter 2 in this dissertation was published in Proceedings of the National Academy of Sciences (PNAS) in December 2009. As equal first author, I was involved in formulating research questions, research design, data collection, data analysis, and manuscript writing. Together with a graduate student in the laboratory, Alice Wang, I performed the experiments for Figure 2.1, 2.3 c–f, 2.5 d–e, 2.6–2.10. Emmanuel Skordalakes performed the experiments for Figure 2.2, 2.3 a–b, 2.4, 2.5 a–c.

A version of Chapter 3 was published in Molecular Cell in September 2009. My contribution to this manuscript as equal first author included research design, data collection, data analysis, literature review, and manuscript writing. I performed all the experiments for Figure 3.4–3.11, 3.12 b–e, and analyzed the data together with a computer science graduate student, Thomas Hentrich. The experiments for Figure 3.1–3.3, 3.12 a, f, g were performed by Jessica Jackson, Shima Nakanishi, and Jennifer Gardner.

Chapter 4 is based on a manuscript currently in preparation. I performed the experiments for all the figures, and, together with Thomas Hentrich, analyzed the data. As first author, I was also responsible for literature review, formulating the resulting model, and writing the paper.

Chapter 1

Introduction

In all living organisms, DNA is the primary storage entity of genetic information and the fundamental template of our heredity. To accommodate its approximate length of two metres into a eukaryotic nucleus of only a few microns, DNA is tightly coiled around proteins to organize it into a structure called chromatin (Kornberg, 1977). Chromatin biology yields a vibrant balance between genome packaging and genome access. The compaction into repeating units, the nucleosomes, allows efficient storage of DNA but also challenges the access of enzymes to DNA during many cellular processes (Kornberg and Lorch, 1999). Therefore, nucleosomes must be stable yet highly dynamic entities, maintaining DNA in a condensed form while allowing controlled access to the genetic information. After being primarily considered to have a structural role during the last century, we are just beginning to understand the additional layer of genome regulation that chromatin offers (Ehrenhofer-Murray, 2004). Moreover, chromatin can adopt stable and potentially heritable alternative states and therefore became a main focus in the field of epigenetics (Bird, 2007). The term epigenetics refers to heritable changes in gene expression that do not involve changes in DNA sequence and in a broad sense bridges genotype and phenotype. The success story of chromatin started about two decades ago with the gradual appreciation of its dynamic nature and influence on cellular functions, including gene transcription, DNA replication, repair, and recombination (Allis et al., 2007).

1.1 Chromatin – History and implications

Walther Flemming first suggested the term chromatin (Greek: *chrōma*, colour) for the stainable fibrous scaffold in the nucleus around 1880 (Flemming, 1882). During Flemming's lifetime, Friedrich Miescher and Albrecht Kossel laid the groundwork for the characterization of chromatin components, and the first detailed descriptions of DNA and histones emerged from their biochemical studies (Miescher, 1871; Kossel, 1911; 1928).

The first half of the twentieth century, primarily brought about discoveries in the rising field of genetics, while understanding the structure and functions of chromatin was largely neglected. It took several decades until studies of chromatin re-entered the mainstream focusing first on describing its basic components. Using electron microscopy, chromatin was visualized as 'beads on a string' (Olins and Olins, 1974), and based on nuclease digestion the fundamental uniformity of chromatin was postulated (Kornberg, 1974). The discovery of the nucleosome then revolutionized the perception of chromatin: in contrast to the initial view of DNA being coated by histones, it was now considered to be wrapped around the histone octamer on the outside of the nucleosome (Oudet et al., 1975). Rather than being a mere packaging layer obscuring access to the DNA, this finding suggested that nucleosomes are an integral part of chromatin modulating DNA accessibility.

Over the last two decades, it became increasingly clear that chromatin is not a static set of equally spaced 'beads on a string' as they appear in electron micrographs (Olins and Olins, 1974). Instead, chromatin is dynamically modified, exhibits high plasticity and hence has active roles in regulating processes such as transcription, replication, and DNA repair (Ehrenhofer-Murray, 2004; Kouzarides, 2007). Current work in the field is directed towards the understanding of these chromatin modifications, their combinatorial complexity, and how they influence cellular functions (Rando, 2007).

1.2 The structural hierarchy of chromatin

Chromatin is dynamically packaged and organized in multiple levels of organization: first, the nucleosome with DNA wrapped around an octamer of histones;

second, the organization of nucleosomes into so-called higher-order domains; and third, the spatial arrangement of chromatin within the nuclear space (Misteli, 2007). On the first level, each nucleosome contains 146 base pairs of DNA wrapped around a histone octamer composed of two heterodimers of core histones H2A and H2B and a tetramer of H3 and H4 (Arents et al., 1991; Luger et al., 1997).

On the second level, chromatin folds into several layers of higher-order structures, such as the 10 nm, 30 nm and 60 nm fibres, that eventually form a chromosome (Woodcock, 2006). Historically, chromatin has been classified as either euchromatin or heterochromatin, originating from the nuclear staining patterns observed with cytological dyes that visualize DNA. During interphase, heterochromatic regions such as telomeres and centromeres stain intensely and are highly compact, whereas the largely decondensed euchromatin stains less. However, the classification into eu- and heterochromatin is rather simplified, since chromatin can adopt many more structural and functional forms (Misteli, 2007).

Another element that contributes to the higher-order organization of chromatin are chromatin loops, which range from smaller kilobase-sized ones to larger loops comprising hundreds of kilobases to bring distant chromatin elements into close proximity (van Driel et al., 2003; Fraser, 2006). Smaller loops often fold active genes in *S. cerevisiae* and in humans, bringing their 5' start and 3' end in physical proximity facilitating the interaction of 3' end-processing factors with the transcription machinery (O'Sullivan et al., 2004; Ansari and Hampsey, 2005; Martin et al., 2005). Larger loops bring distant regulatory elements such as enhancers into proximity of promoters, or even spatially join multiple potentially distant genes (van Driel et al., 2003).

This degree of nuclear arrangement also contributes to the third level of chromatin structure: the spatial arrangement of chromatin within the nuclear space. Chromatin looping combined with the spatial and temporal organization of cellular machineries, creates sub-nuclear compartments for processes such as transcription, DNA replication and repair (Misteli, 2005).

Despite the dynamics of chromatin (Gasser, 2002), each chromosome fills a distinct and compact volume in the interphase nucleus (Cremer and Cremer, 2001; Visser et al., 2000; Duan et al., 2010). During interphase, specialized chromoso-

mal domains such as telomeres, centromeres and other heterochromatic loci, localize to the nuclear periphery (Cockell and Gasser, 1999; Andrulis et al., 1998; Reddy et al., 2008; Finlan et al., 2008), a mechanism which has been proposed to promote transcriptional repression. However, not all genes at the nuclear periphery are necessarily transcriptionally repressed. A large number of *S. cerevisiae* genes for example are repositioned to the nuclear periphery after they have been transcriptionally activated (Brickner, 2010). It has been proposed that this nuclear relocation upon activation is associated with transcriptional memory, but the detailed mechanisms of this phenomenon are still controversial (Brickner, 2009; Halley et al., 2010).

During metaphase of mitosis and meiosis when chromosomes segregate into daughter cells, the most condensed form of chromatin is formed which can be observed by microscopy. This highest level of compaction requires a tremendous restructuring of chromatin and is accompanied by a range of specialized chromatin modifications, including phosphorylation of histone H4 serine 1 (H4S1ph) and histone H3 serine 10 (H3S10ph) (Xu et al., 2009; Govin and Berger, 2009).

1.3 The dynamic modifications of chromatin

All these levels of chromatin organization influence cellular processes and have to be elucidated and integrated by different approaches to fully understand genome regulation. In my dissertation, I present work focusing on the plasticity and modifications at the nucleosome level. Each histone protein in a nucleosome has a globular, highly structured core domain and an unstructured N-terminal tail that protrudes out of the nucleosomal core (Arents et al., 1991; Luger et al., 1997). A large number of histone amino acid residues have been reported to date to be post-transcriptionally modified through acetylation, methylation, ubiquitination, sumoylation, phosphorylation, ADP-ribosylation, biotinylation, and proline isomerization (Bhaumik et al., 2007; Kouzarides, 2007). Initially, modifications at the protruding N-terminal tails were discovered, but many core and C-terminal residues were later found to be modified as well.

Histone modifications can affect chromatin structure by charge-dependent alterations on the surface of nucleosomes, a mechanism discussed further below (Simp-

son, 1978; Ausio and van Holde, 1986; Ura et al., 1997; Ahn et al., 2005). Post-translational modifications (PTMs) may occur individually or in combination on the same nucleosome, thereby generating another layer of complexity. Histone modifications may also influence each other through regulatory crosstalk (Latham and Dent, 2007). Ample evidence suggests that one modification can trigger the occurrence of other ones, either in *cis* on the same histone, or in *trans* across histones and/or nucleosomes, leading to temporal dependencies of observable combinations (Suganuma and Workman, 2008).

It has even been proposed that modifications act together as a ‘histone code’ that is read by effector proteins to trigger downstream functions (Strahl and Allis, 2000; Turner, 2000; Taverna et al., 2007; Jenuwein and Allis, 2001). The ‘epigenetic code’ is an extension of this hypothesis suggesting that histone modifications represent a predictive and heritable code to regulate gene expression during differentiation and development (Turner, 2000; 2007). However, these hypotheses are still controversial, not due to the lack of supporting findings, but rather because more evidence is required to confirm the universality of a histone code (Sims and Reinberg, 2008).

Another fundamental mechanism to alter the chromatin structure is provided by chromatin remodelling enzymes, which use the energy gained from ATP hydrolysis to move, eject, or restructure nucleosomes (Clapier and Cairns, 2009). These remodellers can function to define distinct chromatin regions, to provide DNA accessibility during transcription and DNA repair, or to reassemble chromatin during DNA replication.

One specific class of chromatin remodellers is specialized to replace canonical histones with histone variants at distinct genomic regions. These variants differ from their canonical paralogues in primary amino acid sequence, are enriched in specialized chromatin regions and hence may have consequences and implications for cellular processes (Talbert and Henikoff, 2010). The H2A and H3 family variants, including H2A.Z, H2A.X, H3.3, and CENP-A, are highly conserved and occur almost universally in eukaryotes, suggesting that they may play important roles in chromatin biology (Talbert and Henikoff, 2010).

In Chapter 2 of this dissertation, I present work that advances the understanding of the regulation of histone variant H2A.Z in *Saccharomyces cerevisiae*.

1.3.1 The histone variant H2A.Z

H2A.Z diverged from the core histone H2A prior to the first diversification of eukaryotic lineages (Malik and Henikoff, 2003) and is more conserved between species than H2A itself (Thatcher and Gorovsky, 1994; Iouzalén et al., 1996; Jiang et al., 1998). After the discovery of histone variants in mammalian cells (West and Bonner, 1980), H2A.Z was characterized as being expressed independently of DNA replication, in contrast to the S phase-coupled expression of core histones (Wu and Bonner, 1981).

H2A.Z is deposited into chromatin by the ATP-dependent complex SWR1-C, the first complex identified able to deposit histone variants in *S. cerevisiae* (Kobor et al., 2004; Krogan et al., 2003; Mizuguchi et al., 2004). SWR1-C is named after its catalytic subunit Swr1, which contains a ATPase/helicase domain that is homologous to members of the Swi2/Snf2 chromatin-remodelling family. In budding yeast, SWR1-C shares four of its subunits with the NuA4 histone acetyltransferase (HAT) complex (Kobor et al., 2004; Krogan et al., 2003). Consistent with the evolutionary conservation of H2A.Z, SWR1-C and NuA4 are also conserved in eukaryotes. In higher organisms, the SRCAP (SNF2-related CREB-binding protein activator protein) complex is the functional and structural counterpart of SWR1-C (Cai et al., 2005; Ruhl et al., 2006; Wong et al., 2007; Eissenberg et al., 2005). In addition, SWR1-C and NuA4 seem to have fused together into a single complex, Tip60, potentially to account for the complexity of higher eukaryotes (Doyon and Cote, 2004; Lu et al., 2009). Tip60 was initially found as a homologue of NuA4 (Kimura and Horikoshi, 1998) and its primary function is to acetylate histones, but it can also exchange H2A.Z *in vitro* (Doyon and Cote, 2004; Kusch et al., 2004; Cai et al., 2003; Gévry et al., 2007).

H2A.Z is essential for viability in many organisms, such as *Tetrahymena thermophila* (Liu et al., 1996), *Drosophila melanogaster* (van Daal and Elgin, 1992; Clarkson et al., 1999), *Xenopus leavis* (Iouzalén et al., 1996; Ridgway et al., 2004), and *Mus musculus* (Faast et al., 2001). In mammalian cells, H2A.Z has an es-

sential role during development and is crucial for embryonic stem cell differentiation (Creyghton et al., 2008). In *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, loss of H2A.Z is tolerated, but leads to several phenotypes, including transcriptional defects and chromosome loss (Carr et al., 1994; Jackson and Gorovsky, 2000; Santisteban et al., 2000; Adam et al., 2001; Laroche and Gaudreau, 2003; Krogan et al., 2004). In humans, H2A.Z has been associated with cancer progression and is over-expressed in colorectal and breast metastatic cells (Dunican et al., 2002; Rhodes et al., 2004; Zucchi et al., 2004; Hua et al., 2008; Svtelisl et al., 2010).

Given the impact of H2A.Z on human health and disease, extensive research aims to better understand the physiological role of H2A.Z by deciphering the mechanisms of its functions in different cellular processes. For instance, treating cells that lack H2A.Z with DNA damaging agents results in a severe phenotypic impairment, but the precise mechanistic function of H2A.Z during DNA repair is still not clear (Schleker et al., 2009). Recently it was shown, that in response to DNA damage, SWR1-C, NuA4, and the chromatin-remodelling complex INO80 are recruited to double-stranded breaks partially through physical interactions with phosphorylated H2A.X, the key histone modification during DNA repair (Downs et al., 2004; Morrison and Shen, 2009; van Attikum et al., 2007). As part of the repair attempt, H2A.Z is transiently deposited into chromatin at the damaged site (Kalocsay et al., 2009). Another study could only detect H2A.Z at the double-strand break (DSB) when INO80 was impaired, indicating that INO80 removes H2A.Z from DSBs as it does for H2A.X (van Attikum et al., 2007). Furthermore, sumoylation of H2A.Z at its C-terminus appears to be essential to anchor the DSB to the nuclear periphery (Kalocsay et al., 2009). Despite these interesting evidence, the cooperation of multiple chromatin-modifying complexes at the DNA damage site are just being elucidated, and the exact function of H2A.Z during this process is still controversial.

Furthermore, the physiological importance of H2A.Z may be explained by its roles in heterochromatin boundary element formation (Meneghini et al., 2003) suppression of antisense RNAs (Zofall et al., 2009), and antagonizing DNA methylation in plants and animals (Zilberman et al., 2008; Zemach et al., 2010), but how it func-

tions mechanistically in these processes is largely unknown.

In addition, despite the high level of amino acid sequence conservation of H2A.Z, it is not clear how conserved its roles, such as regulation gene expression, are across eukaryotes. For example, in *T. thermophila*, H2A.Z is found exclusively in the transcriptionally active macronucleus but absent in the transcriptionally inert micronucleus (Allis et al., 1980), suggesting a positive role of H2A.Z in transcription. Consistent with this conclusion, impairing H2A.Z in budding yeast results in an activation defect of the inducible genes *GALI* (Adam et al., 2001; Santisteban et al., 2000), *PHO5* (Santisteban et al., 2000) and *PUR5* (Larochelle and Gaudreau, 2003) as well as the cell-cycle regulated genes *CLN5* and *CLB2* (Dhillon et al., 2006), and some heat shock responsive genes (Zhang et al., 2005). Advances in mapping techniques such as combining chromatin immunoprecipitation with microarrays (ChIP-on-chip) and with high-throughput sequencing (ChIP-Seq) have provided high-resolution maps of H2A.Z localization on a genome scale (Albert et al., 2007; Guillemette et al., 2005; Li et al., 2005; Raisner et al., 2005; Zhang et al., 2005). These studies in *S. cerevisiae* showed that H2A.Z is enriched in many promoters and 5'ends of genes within euchromatic regions, particularly in nucleosomes flanking a nucleosome depleted region (NDR) as further discussed below. In contrast to its proposed role in gene activation, H2A.Z has a clear trend to associate with infrequently transcribed and inducible genes under repressed or basal-expression conditions (Guillemette et al., 2005; Li et al., 2005). In budding yeast, loci specific H2A.Z deposition seems to be controlled by a particular pattern of histone tail acetylation, but also by a DNA sequence-coded deposition signal which contains the binding site of the transcription factor Reb1 (Raisner et al., 2005).

In human cells no clear transcriptional association was found, as H2A.Z localizes to enhancers and gene promoters of both active and inactive genes (Barski et al., 2007; John et al., 2008). Knockdown of H2A.Z in fibroblasts resulted in derepression of the p53-regulated *p21* gene in the non-induced state (Gévry et al., 2007), suggesting that H2A.Z has a role in repressing transcription.

Taken together, it seems that H2A.Z has distinct and sometimes contrasting functions in regulating gene expression that could be explained by a potential role in

poising promoters for either rapid transcriptional activation or repression. Since gene promoters contain regulatory sites for both transcriptional activators and repressors, keeping these sequences amenable for rapid binding by an open chromatin structure could be the key function of H2A.Z. It has also been suggested that PTMs of H2A.Z could play a role in distinguishing its contrasting functions (Draker and Cheung, 2009). H2A.Z can be acetylated at multiple residues in its N-terminus by Esa1 and Gcn5, components of the HAT complexes NuA4 and SAGA, respectively (Babiarz et al., 2006; Keogh et al., 2006; Millar et al., 2006). H2A.Z lysine 14 acetylation is highly correlated with transcriptionally active promoters, whereas its unacetylated form is associated with transcriptionally inactive promoters (Millar et al., 2006). Interestingly, H2A.Z is acetylated after being integrated into chromatin, indicating a potential downstream role for its acetylation (Babiarz et al., 2006; Keogh et al., 2006). H2A.Z acetylation was recently shown to have a role in gene induction, but the function of H2A.Z in *GALI* transcriptional memory are still controversial (Halley et al., 2010; Brickner et al., 2007).

Besides being acetylated, H2A.Z can be sumoylated and monoubiquitinated (Kalocsay et al., 2009; Sarcinella et al., 2007). As mentioned above, sumoylation of H2A.Z is important during the DNA damage response in *S. cerevisiae* (Kalocsay et al., 2009). The monoubiquitinated form of H2A.Z is found in regions of facultative heterochromatin, and is associated with the polycomb-mediated silencing pathway in mammalian cells (Sarcinella et al., 2007). Taken together, to fully unravel the many faces of H2A.Z biology, it is essential to not only understand its deposition patterns but also the function of its PTMs (Draker and Cheung, 2009). My work presented in Chapter 2, aims to reveal the circuitry of H2A.Z deposition and acetylation.

1.3.2 Histone acetylation

Compared to histone variants, the diversity of PTMs on core histones is even greater allowing a higher complexity to define chromatin neighbourhoods that enable diverse cellular functions (Millar and Grunstein, 2006). Transcription was first functionally associated with histone acetylation and methylation almost fifty years ago, and particularly acetylation was shown to have a positive effect on

mRNA biosynthesis (Allfrey et al., 1964). However, it was not until the purification and sequencing of the first transcription-linked histone acetyl-transferase and deacetylase—a homolog of the acetyltransferase Gcn5 from *Tetrahymena* and the mammalian deacetylase Rpd3 (Brownell et al., 1996; Taunton et al., 1996)—that the research fields of transcription and chromatin merged.

HATs are a class of enzymes that use acetyl-CoA as a substrate to add a single acetyl group to the ϵ -amino group of histone lysine's. A variety of HATs has been characterized to date (Lee and Workman, 2007), including NuA4, the only essential HAT in *S. cerevisiae*, which shares four of its subunits with the H2A.Z-deposition complex SWR1-C (Allard et al., 1999). Histone acetylation has become a hallmark of active transcription, because the level of acetylation in the promoter and 5' end of genes correlates with their transcriptional frequency (Liu et al., 2005; Pokholok et al., 2005; Wang et al., 2008). Like most PTMs, acetylation is reversible and eukaryotic genomes encode multiple histone deacetylases (HDACs) that remove acetyl groups from histones (Yang and Seto, 2007). HATs and HDACs do not only modify histones, but are also known to target many other non-histone proteins (Kouzarides, 2000). For instance, while NuA4 was first identified through its ability to acetylate histone H4, it also acetylates critical metabolic enzymes regulating glucose metabolism and life span (yi Lin et al., 2009). Furthermore, proteins with a large impact on human health and disease, such as the tumour suppressor p53, are targets of HATs and HDACs (Gu and Roeder, 1997; Kouzarides, 2000), making HATs and HDACs promising therapeutic targets for treatment of cancer and other diseases (Yang and Seto, 2007).

1.3.3 Histone methylation

Methylation of arginine or lysine residues is another prominent histone PTM (Murray, 1964). Lysines are mono-, di-, or trimethylated by either the SET domain- or non-SET domain-containing histone methyltransferases (HMTases) (Shilatifard, 2006). In histone H3, lysines 4, 9, 27, 36, and 79 are major sites of methylation, while for histone H4, lysine methylation is restricted to residue 20 (Lan and Shi, 2009). Methylation is a relatively stable histone mark, leading to the long lasting belief that this mark is non-reversible and therefore might constitute a crucial

component of epigenetic inheritance. However, like acetylation, methylation is reversible and can be removed by a group of enzymes called histone demethylases (Shilatifard, 2006). Nevertheless, histone methylation is still proposed to function as a memory mark during cell state inheritance (Radman-Livaja et al., 2010; Muramoto et al., 2010; Lan and Shi, 2009).

In general, PTMs and their associated functions are highly conserved between eukaryotes. Exceptions include the repressive methylation marks on histone H3K9, H3K27, and H4K20, which are found in most eukaryotes but have not been detected in *S. cerevisiae* (Lachner et al., 2004). One reason for this may lay in the overall differences in genome organization between species (Kazazian, 2004). For example, higher eukaryotes contain large expansions of silenced repetitive and noncoding sequences, while budding yeast has a more ‘open’ genome with very few regions that are constitutively silent, such as the telomeres and the mating-type loci. As a result, *S. cerevisiae* lacks most of the genome silencing mechanisms, such as RNAi interference, repressive histone lysine methylation marks, and DNA methylation that higher eukaryotes developed. In fact, *S. cerevisiae* uses a unique silencing mechanism primarily based on SIR proteins to repress its few heterochromatic domains (Rusche et al., 2003; Gasser and Cockell, 2001).

In contrast, *S. pombe* that diverged from *S. cerevisiae* around 400 million years ago (Sipiczki, 2000) shares many features of genome organization and silencing mechanisms with higher eukaryotes. Despite the limited amount of repressed chromosomal regions (centromeres, telomeres, and mating-type loci) in *S. pombe*, it contains the repressive histone lysine marks H3K9 and H4K20, as well as RNAi interference (Grewal and Elgin, 2007).

Unlike these evolutionary differences of silencing mechanisms, histone marks primarily associated with active chromatin, including methylation of lysine 4 on histone H3 (H3K4me) are well conserved. From yeast to human, H3K4 trimethylation predominantly marks the 5’ end of active genes and correlates positively with RNA polymerase II (RNAPII) occupancy as well as histone acetylation (Bernstein et al., 2002; 2005; Krogan et al., 2003a; Ng et al., 2003; Santos-Rosa et al., 2002; Schneider et al., 2004; Schubeler et al., 2004; Pokholok et al., 2005).

Dimethylation of H3K4 is broadly associated with coding regions of active as well

as transcriptionally poised genes in *S. cerevisiae* (Ng et al., 2003; Santos-Rosa et al., 2002; Pokholok et al., 2005), whereas in vertebrates, it partially co-occurs with H3K4 trimethylation at transcriptional start sites but is also independently associated with other regions (Schneider et al., 2004; Bernstein et al., 2005). Set1 is the only known H3K4 methylase in *S. cerevisiae*, and is only active within the COMPASS complex, which associates with the early elongating RNAPII via the polymerase II-associated factor 1 (Paf1) complex (Miller et al., 2001; Krogan et al., 2002; Roguev et al., 2001; Wood et al., 2007; Briggs et al., 2001; Nagy et al., 2002). In higher eukaryotes, there are more than half a dozen Set1-related proteins. These include hSet1A and B as well as the MLL (Mixed Lineage Leukemia) proteins, which are also found in COMPASS-like complexes capable of methylating H3K4 (Shilatfard, 2008). The importance of MLL in cancer biology is highlighted by the association of its translocations with the pathogenesis of hematological malignancies such as acute myeloid and lymphoid leukemia (AML and ALL) (Krivtsov and Armstrong, 2007).

Methylation of histone H3 lysine 36 (H3K36) is another chromatin modification known to be associated with transcription. In *S. cerevisiae*, it marks the 3' ends of active genes (Pokholok et al., 2005) and is established by the SET domain-containing protein Set2 (Strahl et al., 2002). During transcriptional elongation, Set2 associates with serine 2-phosphorylation within the C-terminal domain (CTD) of RNAPII (Li et al., 2002; 2003; Xiao et al., 2003; Schaft et al., 2003; Krogan et al., 2003b), and partial deletion of the CTD results in significant global loss of H3K36 methylation. The human Set2 homolog, Setd2/HYPB (huntingtin interacting protein B), also interacts with phosphorylated RNAPII and, like other SET-domain containing proteins, plays an important role in embryo development (Sun et al., 2005; Hu et al., 2010). Another Set2 homolog in humans is the leukemia oncoprotein NSD1 (Rayasam et al., 2003), which is also capable of methylating H3K36. In *S. cerevisiae*, H3K36 methylation is further known to suppress cryptic transcription initiation in coding regions by recruiting the HDAC Rpd3S, which deacetylates histones in open reading frames (ORFs) after transcription (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005; Li et al., 2007b).

Recently, the connection between histone modifications and transcription was ex-

tended to the transcription-coupled process of mRNA splicing. Evidence suggests a causal relationship between H3K4 and H3K36 methylation and mRNA splicing in higher eukaryotes (Schwartz and Ast, 2010). My dissertation will, among other aspects, elaborate on the connection between histone modifications and mRNA splicing in *S. cerevisiae*.

Lysine 79, a conserved residue located on the surface of the histone H3 globular domain, was one of the first core lysines shown to be methylated (van Leeuwen et al., 2002; Lu et al., 2008). The responsible enzyme, Dot1, catalyzes mono-, di-, and trimethylation of H3K79 and is evolutionally conserved. So far, Dot1 is the only known histone lysine methyltransferase lacking a SET domain, but instead it has a conserved catalytic core similar to arginine methyltransferases (Sawada et al., 2004; Min et al., 2003). Dot1 was originally discovered in a high-copy suppressor screen for factors affecting telomeric gene silencing in *S. cerevisiae* (Singer et al., 1998). Subsequently, it was proposed that loss of Dot1 or mutations of H3K79 abolish silencing by reducing the association of Sir2 and Sir3, key heterochromatic proteins, with silenced regions (Lacoste et al., 2002; Ng et al., 2002; van Leeuwen et al., 2002). Besides its role in silencing, Dot1 is involved in the DNA damage response to maintain genomic integrity. In particular, it is required for the meiotic recombination checkpoint as well as for G1/S and intra-S DNA damage checkpoints in vegetative cells (San-Segundo and Roeder, 2000; Giannattasio et al., 2005; Wysocki et al., 2005). However, the exact roles of H3K79 methylation in DNA repair or telomeric silencing are still not fully understood. This uncertainty can largely be attributed to the controversy of whether different methylation states of H3K79 have similar or distinct functions. In Chapter 3 of my dissertation, I show that H3K79 methylation states are mutually exclusive across the genome and are linked to different cellular functions.

The relevance of proper H3K79 methylation for human health is supported by the fact that Dot1 interacts with MLL fusion partners and has been directly associated with leukemogenesis through the misregulation of HOX genes (Okada et al., 2005). Furthermore, in human cancerous cells the MLL-AF4 fusion protein interacts with Dot1 causing abnormal H3K79 methylation at MLL target genes (Krivtsov et al., 2008).

1.3.4 Histone monoubiquitination

Besides acetylation and methylation, ubiquitination of histone lysines plays a key role in regulating gene expression. While acetyl- and methyl-groups are small molecular tags, monoubiquitin is a 76 amino acid protein conjugated to proteins in a cascade of three enzymatic processes (Hochstrasser, 1996). First, ubiquitin has to be activated by an ATP-dependent reaction catalyzed by an E1 enzyme. Next, it gets conjugated to a cysteine residue by an E2 enzyme. Lastly, it is transferred to the target lysine by an E3 ubiquitin-protein ligase.

In *S. cerevisiae*, H2B is ubiquitinated at the highly conserved lysine 123, which is located in the C-terminal tail protruding from the nucleosome core (Luger et al., 1997; White et al., 2001). In budding yeast, the E2-enzyme Rad6 acts in concert with the E3 ubiquitin ligase Bre1 to monoubiquitinate H2B (Hwang et al., 2003; Wood et al., 2003a). Besides H2B monoubiquitination, Rad6 plays a role in DNA damage repair and protein degradation pathways in conjunction with other E3 ligases including Rad18, Ubr1, or Rad5 (Osley et al., 2006).

As one of the first examples for trans-tail histone cross-talk, H2B monoubiquitination was shown to be essential for the establishment of both H3K4- as well as H3K79 trimethylation (Dover et al., 2002; Krogan et al., 2003a; Sun and Allis, 2002; Wood et al., 2003b; Nakanishi et al., 2009). This regulatory pathway is unidirectional as mutations that eliminate either H3K4 or H3K79 methylation have no effect on H2B ubiquitination levels. In this unidirectional crosstalk, monoubiquitin of H2B is postulated to present a molecular ‘tag’ that is recognized by the HMTs Set1 and Dot1, recruiting both to methylate H3K4 and H3K79, respectively (Jeltsch and Rathert, 2008).

Furthermore, ample evidence suggests that H2B monoubiquitination plays a role in transcriptional regulation during initiation and elongation (reviewed in Weake 2008). H2B monoubiquitination is a highly transient mark that is removed by the ubiquitin-specific proteases Ubp8 and Ubp10 (Weake and Workman, 2008). Ubp8 is a subunit of the Spt-Ada-Gcn5-acetyltransferase (SAGA) complex, and the integrity of SAGA is required for Ubp8 deubiquitination activity (Daniel et al., 2004; Henry et al., 2003; Ingvarsdottir et al., 2005; Lee et al., 2005; Powell et al., 2004). In contrast, Ubp10 acts independently of SAGA and is associated with telomeric

silencing (Emre et al., 2005; Gardner et al., 2005). Until now, the reason for the existence of two H2B deubiquitinases and whether they act on similar or different pools of H2B ubiquitin was unclear. In Chapter 4 of my dissertation, I show that Ubp8 and Ubp10 remove H2BK123ub at different genomic loci and associate with H3K4 and H3K79 trimethylation, respectively.

1.4 Domains as writer, reader, and eraser modules

Genome-wide studies showed that PTMs occur in distinct regions, form chromatin neighbourhoods, and exhibit patterns. However, there is still a gap in the mechanistic understanding of how the various modifications are translated into specific cellular outputs. Two non-mutually exclusive models, called *direct* and *effector-mediated*, have been proposed to explain how chromatin patterns influence biological processes (Taverna et al., 2007; Strahl and Allis, 2000; Jenuwein and Allis, 2001). The *direct model* proposes modifications to directly affect the chromatin structure by modulating histone-DNA or histone-histone contacts, such as histone acetylation or phosphorylation weakening favourable interactions between the basic histone proteins and the negatively charged DNA resulting in a more open chromatin structure (Simpson, 1978; Ausio and van Holde, 1986; Ura et al., 1997; Ahn et al., 2005). Similarly, incorporation of H2A.Z into chromatin is thought to weaken nucleosome stability, facilitating rapid exchange during transcriptional activation or repression. However, the results on the precise nature of these effects are still very much controversial (Zlatanova and Thakar, 2008).

In the *effector-mediated model*, chromatin modifications provide binding sites that recruit chromatin-modifying complexes to pass on upstream information to downstream signal cascades (Cheung et al., 2000; Ruthenburg et al., 2007). Typically, these effector complexes comprise multiple subunits and contain protein interaction domains that recognize specific PTMs. This model is inspired by the common theme in biological systems where signalling modules operate in a triad as *writer*, *eraser*, and *reader* (Kouzarides, 2000; Rossman et al., 2005). Many of the identified chromatin-modifying complexes conform with this theme and have a subunit with an enzymatic domain to add (write) or remove (erase) a certain chromatin modification. Writers are represented by histone acetyltransferases (HATs), his-

tone methyltransferases (HMTs), and ubiquitin ligases. A well-studied example of a writer is the SET domain, an evolutionarily conserved, 130 amino acid sequence motif found in a variety of complexes that methylate histones (Shilatifard, 2008; 2006). Its name is derived from the *Drosophila* proteins Su(var)3–9, Enhancer of zeste [E(z)], and trithorax (SET), since it was originally identified in polycomb group (PcG), trithorax group (trxG), and Su(var) genes (Tschiersch et al., 1994; Jones and Gelbart, 1993; Stassen et al., 1995). Eraser functionality can be found in histone deacetylases (HDACs), histone demethylases (HDMTs), and deubiquitinating enzymes (DUBs).

Complexes containing writer or eraser subunits are either directly or indirectly recruited to specific chromatin regions through interaction with sequence-specific DNA binding factors or chromatin readers (Taverna et al., 2007). Biochemical and biophysical studies have identified a large number of readers containing conserved protein domains that specifically bind histone PTMs dependent on both the type and position of the modification within the histone (Taverna et al., 2007). The bromodomain is one well-characterized chromatin reader domain that recognizes lysine acetylation (Dhalluin et al., 1999; Jacobson et al., 2000). This domain, named after the Brahma protein in *Drosophila*, is highly conserved and found in many chromatin and transcription regulators like the HAT Gcn5, the chromatin remodelling complex RSC (remodel the structure of chromatin), as well as in Bdf1, a subunit of both the H2A.Z-deposition complex SWR1-C and the general transcription factor TFIID (Yang, 2004).

Several domains are known to bind distinct methylated histone lysine residues such as chromodomains, plant homeodomain (PHD)-finger domains, WD40 domains, Tudor domains, and MBT domains (Lachner et al., 2001; Huyen et al., 2004; Wysocka et al., 2005; 2006; Kim et al., 2006; Li et al., 2006; Peña et al., 2006; Shi et al., 2006).

Besides these binding domains, chromatin-modifying complexes may contain other highly conserved yet functionally uncharacterized modules. One of them is the YEATS domain found in many chromatin and transcription regulators (Schulze et al., 2009). Its name is derived from the first described YEATS domain-containing proteins, Yaf9, ENL, AF9, Taf14, Sas5 (Masson et al., 2003), although more than

one hundred YEATS-containing proteins have been identified to date. Interestingly, ENL, AF9 as well as GAS41 are human YEATS domain-containing proteins linked to cancer biology (Schulze et al., 2009).

In *S. cerevisiae*, Taf14, Sas5, and Yaf9 are the only proteins that have a YEATS domain. Yaf9 is one of the four shared subunits between SWR1-C and NuA4 and has roles in chromosome segregation, DNA repair, and telomeric silencing (Schulze et al., 2009). Despite the importance of the YEATS domain for biological processes and cancer biology, very little is known about its structure and function. Recently, our laboratory solved the structure of the Yaf9 YEATS domain. The work presented herein describes its functional conservation from yeast to human and its role in H2A.Z deposition at specific promoters. This study further led to the hypothesis that the YEATS domain acts as a chromatin reader through its distinct acetyl-lysine binding pocket (Schulze et al., 2010).

Often, multiple PTM reader modules act together within the same protein or within the same chromatin-modifying complex. The combination of several reader domains in one complex allows for simultaneous reading of individual marks and recognition of patterns (Ruthenburg et al., 2007). In this context, histone tails may act as signalling platforms that integrate the upstream information represented by the individual marks. The complexity of modifications on these platforms allows the storage of more information than individual marks and, hence, may generate greater specificity for the recruitment of chromatin-modifying modules to particular sites.

1.5 Transcriptional regulation in the context of chromatin

Transcription of protein coding genes provides an ideal example to describe the dynamic and combinatorial character of chromatin modifications and their impact on a fundamental cellular process (Li et al., 2007a). As visualized in Figure 1.1, page 19, various histone marks cooperate to generate the plasticity on the chromatin level that is required for the multiple steps of transcription.

During mRNA synthesis, the transcriptional machinery, consisting of RNA polymerase II (RNAPII), the general transcription factors (GTFs), and additional regu-

latory factors, requires access to the DNA template. Since chromatin tightly packages DNA, it generally has an inhibitory effect on transcription as early work already revealed (Wasylyk et al., 1979). Similar to classical bacterial repressors, such as the lac and bacteriophage repressors, nucleosomes block access to DNA binding sites (Kornberg and Lorch, 2007). Thus, the default state of chromatin keeps genes inactive and requires activating mechanisms to remodel chromatin and facilitate transcription. This hypothesis is supported by nucleosome maps in yeast showing that promoters of infrequently transcribed and inducible genes tend to have a higher density of nucleosomes than constitutively active genes, which contain more open, nucleosome-depleted promoters (Cairns, 2009). These open promoters comprise a ~150 bp long nucleosome-depleted region (NDR) directly upstream of the transcription start site (TSS), which bears key *cis*-regulatory sequences. NDRs are sometimes misleadingly described as nucleosome-free, since current techniques are often unable to detect highly dynamic and unstable nucleosomes. However, these regions rather show a gradient of depletion and still contain nucleosomes including H3.3/H2A.Z double variants (Jin et al., 2009). Though not exclusive to, almost all NDR promoters have H2A.Z enriched in both flanking nucleosomes (Raisner et al., 2005; Guillemette et al., 2005). Because transcription factors recognize specific DNA sequences it is likely that in order to bind *cis*-regulatory elements they must compete with nucleosomes. As expected, high nucleosome turnover rates occur in many promoters (Dion et al., 2007), which might facilitate the binding of specific transcription factors. Subsequently, transcription factors may recruit chromatin-remodelling complexes, which slide or eject nucleosomes to establish an open chromatin environment for the transcriptional machinery (Cairns, 2009).

Overcoming the nucleosomal barrier also depends on complexes such as Mediator and SAGA, acting as scaffolds on which factors of the transcription machinery and RNAPII can assemble (Weake and Workman, 2010). These steps eventually lead to the formation of the pre-initiation complex (PIC) at the promoter (Roeder, 2005; Ptashne and Gann, 1997). After PIC formation, a series of sequential chromatin modifications are necessary for RNAPII to start initiation, clear the promoter, and begin transcription elongation: In the initial stage, Bre1 together with Rad6 are

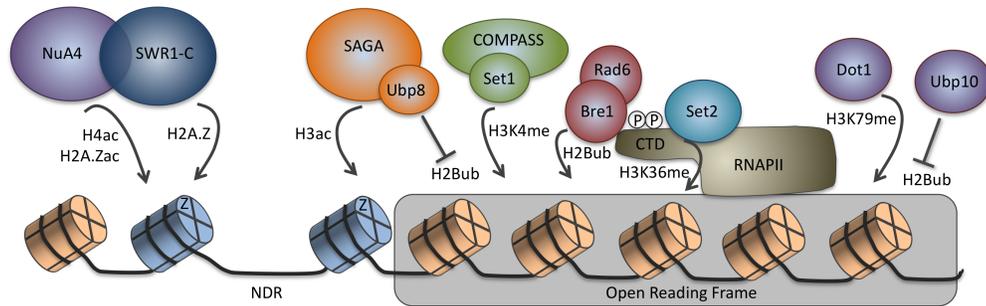


Figure 1.1: Chromatin-modifying complexes associated with transcription in *S. cerevisiae*. Depiction of multiple chromatin-modifying complexes studied in this dissertation that dynamically influence chromatin during transcription. NuA4 and SWR1-C bind to gene promoters to acetylates histones H4 and deposit H2A.Z, respectively. In addition, the SAGA complex acetylates histones H3, and its subunit Ubp8 removes the H2B monoubiquitin, which was established earlier by Rad6/Bre1. The transient H2B monoubiquitin mark is required for H3K4 trimethylation mediated by the Set1/COMPASS complex. During transcription elongation, Rad6/Bre1 as well as Set2 travel with the RNAPII to establish H2B monoubiquitin as well as H3K36 trimethylation across the open reading frame (ORF), respectively. H2B monoubiquitin likely recruits Dot1 to trimethylate H3K79, and as my dissertation shows H2B monoubiquitin is subsequently removed by Ubp10 in these regions.

recruited during transcription initiation to monoubiquitinate histone H2B (Weake and Workman, 2008), which is an essential step for subsequent trimethylation of H3K4 by the Set1-containing COMPASS complex. In the next step, deubiquitination of H2B by Ubp8 triggers the recruitment of Ctk1, which is involved in the sequential phosphorylation of RNAPII CTD (Wyce et al., 2007). In a next step critical for RNAPII promoter release, phosphorylation of the CTD recruits the H3K36 methyltransferase Set2, which travels with the elongating RNAPII to methylate H3K36 at the body and 3' end of genes (Weake and Workman, 2008). Rad6/Bre1 also associate with RNAPII during transcription elongation to ubiquitinate H2B across the ORF. H2BK123ub then likely serves as a signal for Dot1 recruitment and trimethylation of H3K79.

Table 1.1: Responsible enzymes for covalent histone modifications and variants studied in this dissertation. Hs: *Homo sapiens*; Sc: *Saccharomyces cerevisiae*

Modification/Variant	Writer/Remodeler	Eraser
H2A.Z	SWR1-C (Sc), SRCAP (Hs), TIP60 (Hs)	possibly removed through histone turn-over
H2A.Zac	Esa1 (Sc), Gcn5 (Sc)	Hda1 (Sc)
H3K4me	Set1/COMPASS (Sc), SET7/SET9 (Hs), MLL (Hs), Smyd3 (Hs)	Jhd2 (Sc), LSD1 (Hs), JARID1 (Hs)
H3K36me	Set2 (Sc), SETD2/HYPB (Hs), NSD1 (Hs)	Jhd1 (Sc), JHDM1 (Hs), JMJD2C (Hs)
H3K79me	Dot1 (Sc), DOT1L (Hs)	unknown to date
H2BK123ub	Rad6/Bre1 (Sc), HR6A/B (Hs), RNF20/RNF40 (Hs)	Ubp8 (Sc), Ubp10 (Sc), USP22 (Hs), USP7 (Hs)

Besides being a transient mark essential for optimal transcription and the establishment of H3K4- and H3K79 trimethylation, very little is known about the interplay between H2BK123ub and other chromatin modifications involved in transcription initiation and elongation. In my dissertation, I mapped the localization of these marks genome-wide and comprehensively compared them to learn more about the complex circuitry of chromatin modifications during transcription and other cellular processes.

Moreover, most of the chromatin modifications studied in this work are involved in maintaining the boundary between hetero- and euchromatin. The histone variant H2A.Z, methylation of H3K4 and H3K79, as well as monoubiquitination of H2B are euchromatic marks and are proposed to act as boundary factors to prevent the spread of the heterochromatic Sir proteins in *S. cerevisiae* (Meneghini et al., 2003; van Leeuwen et al., 2002; Sun and Allis, 2002; Nislow et al., 1997). However, the cooperation of these marks is certainly not limited to the hetero- and euchromatin

boundary. Work presented in my dissertation aims to broaden the understanding of their cooperation by studying their circuitry on a genome-wide scale. A list of all studied histone modifications/variants and their responsible enzymes is shown in Table 1.1.

In Chapter 2 of this dissertation, I present work that advances the understanding of the regulation of histone variant H2A.Z by the Yaf9 YEATS domain. This work recognizes the evolutionary conservation of the YEATS domain from yeast to human and presents its structural composition and function in depositing the histone variant H2A.Z.

In Chapter 3, I study the different methylation states of H3K79 and show that H3K79 di- and trimethylation are associated with different regions of the genome. In contrast to H3K79me3, H3K79me2 marks M/G1 cell cycle-regulated genes and its levels change during the cell cycle. Furthermore, I present the first genome-wide map of H2BK123ub in *S. cerevisiae* and show that it co-localizes with H3K79me3, but not H3K79me2, consistent with its role as an upstream regulator for H3K79me3.

In Chapter 4, I extend the genome-wide study on H2BK123ub and comprehensively compare its localization to its downstream modifications, H3K4me3 and H3K79me3. In addition, I focus on the role of the two H2B deubiquitinases Ubp8 and Ubp10. I reveal that they mainly act on distinct genomic loci: while Ubp8 removes H2BK123ub at sites enriched for H3K4me3, Ubp10 functions at those marked by H3K79me3. Finally, my thesis describes chromatin signatures of H2BK123ub and H3 methylation marks, and finds the combination of H2BK123ub, H3K4me3, and H3K36me3 to be specific for highly transcribed genes, including those containing introns. In this context, I show evidence that link the histone H2B ubiquitin ligase Bre1 to mRNA splicing.

Bibliography

- Adam, M., Robert, F., Larochelle, M. and Gaudreau, L. (2001). H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. *Mol Cell Biol* *21*, 6270–9. → pages 7, 8
- Ahn, S.-H., Cheung, W. L., Hsu, J.-Y., Diaz, R. L., Smith, M. M. and Allis, C. D. (2005). Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in *S. cerevisiae*. *Cell* *120*, 25–36. → pages 5, 15
- Albert, I., Mavrich, T. N., Tomsho, L. P., Qi, J., Zanton, S. J., Schuster, S. C. and Pugh, B. F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* *446*, 572–6. → pages 8
- Allard, S., Utley, R. T., Savard, J., Clarke, A., Grant, P., Brandl, C. J., Pillus, L., Workman, J. L. and Côté, J. (1999). NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *The EMBO journal* *18*, 5108–19. → pages 10
- Allfrey, V. G., Faulikner, R. and Mirsky, A. E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci USA* *51*, 786–94. → pages 10
- Allis, C. D., Glover, C. V., Bowen, J. K. and Gorovsky, M. A. (1980). Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, *Tetrahymena thermophila*. *Cell* *20*, 609–17. → pages 8
- Allis, C. D., Jenuwein, T., Reinberg, D. and Caparros, M. L. (2007). *Epigenetics*. 1st edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. → pages 1

- Andrulis, E. D., Neiman, A. M., Zappulla, D. C. and Sternglanz, R. (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* 394, 592–5. → pages 4
- Ansari, A. and Hampsey, M. (2005). A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. *Genes Dev* 19, 2969–78. → pages 3
- Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci USA* 88, 10148–52. → pages 3, 4
- Ausio, J. and van Holde, K. E. (1986). Histone hyperacetylation: its effects on nucleosome conformation and stability. *Biochemistry* 25, 1421–8. → pages 5, 15
- Babiarz, J., Halley, J. and Rine, J. (2006). Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in *Saccharomyces cerevisiae*. *Genes Dev* 20, 700–10. → pages 9
- Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–37. → pages 8
- Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. and Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci USA* 99, 8695–700. → pages 11
- Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J., Gingeras, T. R., Schreiber, S. L. and Lander, E. S. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169–81. → pages 11, 12

- Bhaumik, S. R., Smith, E. and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14, 1008–16. → pages 4
- Bird, A. (2007). Perceptions of epigenetics. *Nature* 447, 396–8. → pages 1
- Brickner, D. G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P.-C., Widom, J. and Brickner, J. H. (2007). H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. *PLoS Biol* 5, e81. → pages 9
- Brickner, J. H. (2009). Transcriptional memory at the nuclear periphery. *Curr Opin Cell Biol* 21, 127–33. → pages 4
- Brickner, J. H. (2010). Transcriptional memory: staying in the loop. *Curr Biol* 20, R20–1. → pages 4
- Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., Winston, F. and Allis, C. D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev* 15, 3286–95. → pages 12
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y. and Allis, C. D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843–51. → pages 10
- Cai, Y., Jin, J., Florens, L., Swanson, S., Kusch, T., Li, B., Workman, J., Washburn, M., Conaway, R. and Conaway, J. (2005). The mammalian YL1 protein is a shared subunit of the TRRAP/TIP60 histone acetyltransferase and SRCAP complexes. *J Biol Chem* 280, 13665–70. → pages 6
- Cai, Y., Jin, J., Tomomori-Sato, C., Sato, S., Sorokina, I., Parmely, T., Conaway, R. and Conaway, J. (2003). Identification of new subunits of the multiprotein mammalian TRRAP/TIP60-containing histone acetyltransferase complex. *J Biol Chem* 278, 42733–6. → pages 6

- Cairns, B. R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature* 461, 193–8. → pages 18
- Carr, A. M., Dorrington, S. M., Hindley, J., Phear, G. A., Aves, S. J. and Nurse, P. (1994). Analysis of a histone H2A variant from fission yeast: evidence for a role in chromosome stability. *Mol Gen Genet* 245, 628–35. → pages 7
- Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M. P. and Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581–92. → pages 12
- Cheung, P., Allis, C. D. and Sassone-Corsi, P. (2000). Signaling to chromatin through histone modifications. *Cell* 103, 263–71. → pages 15
- Clapier, C. R. and Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annu Rev Biochem* 78, 273–304. → pages 5
- Clarkson, M. J., Wells, J. R., Gibson, F., Saint, R. and Tremethick, D. J. (1999). Regions of variant histone His2AvD required for *Drosophila* development. *Nature* 399, 694–7. → pages 6
- Cockell, M. and Gasser, S. M. (1999). Nuclear compartments and gene regulation. *Curr Opin Genet Dev* 9, 199–205. → pages 4
- Cremer, T. and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2, 292–301. → pages 3
- Creyghton, M. P., Markoulaki, S., Levine, S. S., Hanna, J., Lodato, M. A., Sha, K., Young, R. A., Jaenisch, R. and Boyer, L. A. (2008). H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* 135, 649–61. → pages 7
- Daniel, J. A., Torok, M. S., Sun, Z.-W., Schieltz, D., Allis, C. D., Yates, J. R. and Grant, P. A. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* 279, 1867–71. → pages 14

- Dhalluin, C., Carlson, J., Zeng, L., He, C., Aggarwal, A. and Zhou, M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399, 491–6. → pages 16
- Dhillon, N., Oki, M., Szyjka, S. J., Aparicio, O. M. and Kamakaka, R. T. (2006). H2A.Z functions to regulate progression through the cell cycle. *Mol Cell Biol* 26, 489–501. → pages 8
- Dion, M. F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N. and Rando, O. J. (2007). Dynamics of replication-independent histone turnover in budding yeast. *Science (New York, NY)* 315, 1405–8. → pages 18
- Dover, J., Schneider, J., Tawiah-Boateng, M., Wood, A., Dean, K., Johnston, M. and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 277, 28368–71. → pages 14
- Downs, J. A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., Kron, S. J., Jackson, S. P. and Côté, J. (2004). Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol Cell* 16, 979–90. → pages 7
- Doyon, Y. and Cote, J. (2004). The highly conserved and multifunctional NuA4 HAT complex. *Curr Opin Genet Dev* 14, 147–54. → pages 6
- Draker, R. and Cheung, P. (2009). Transcriptional and epigenetic functions of histone variant H2A.Z. *Biochem Cell Biol* 87, 19–25. → pages 9
- Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., Shendure, J., Fields, S., Blau, C. A. and Noble, W. S. (2010). A three-dimensional model of the yeast genome. *Nature* 465, 363–7. → pages 3
- Dunican, D. S., McWilliam, P., Tighe, O., Parle-McDermott, A. and Croke, D. T. (2002). Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. *Oncogene* 21, 3253–7. → pages 7

- Ehrenhofer-Murray, A. E. (2004). Chromatin dynamics at DNA replication, transcription and repair. *Eur J Biochem* 271, 2335–49. → pages 1, 2
- Eissenberg, J. C., Wong, M. and Chrivia, J. C. (2005). Human SRCAP and *Drosophila melanogaster* DOM are homologs that function in the notch signaling pathway. *Mol Cell Biol* 25, 6559–69. → pages 6
- Emre, N. C. T., 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. *Mol Cell* 17, 585–94. → pages 15
- Faast, R., Thonglairoam, V., Schulz, T. C., Beall, J., Wells, J. R., Taylor, H., Matthaiei, K., Rathjen, P. D., Tremethick, D. J. and Lyons, I. (2001). Histone variant H2A.Z is required for early mammalian development. *Curr Biol* 11, 1183–7. → pages 6
- Finlan, L. E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J. R. and Bickmore, W. A. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet* 4, e1000039. → pages 4
- Flemming, W. (1882). *Zellsubstanz, Kern und Zelltheilung*. F C W Vogel, Leipzig. → pages 2
- Fraser, P. (2006). Transcriptional control thrown for a loop. *Curr Opin Genet Dev* 16, 490–5. → pages 3
- 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. *Mol Cell Biol* 25, 6123–39. → pages 15
- Gasser, S. M. (2002). Visualizing chromatin dynamics in interphase nuclei. *Science (New York, NY)* 296, 1412–6. → pages 3
- Gasser, S. M. and Cockell, M. M. (2001). The molecular biology of the SIR proteins. *Gene* 279, 1–16. → pages 11

- Gévry, N., Chan, H. M., Laflamme, L., Livingston, D. M. and Gaudreau, L. (2007). p21 transcription is regulated by differential localization of histone H2A.Z. *Genes Dev* 21, 1869–81. → pages 6, 8
- Giannattasio, M., Lazzaro, F., Plevani, P. and Muzi-Falconi, M. (2005). The DNA damage checkpoint response requires histone H2B ubiquitination by Rad6-Bre1 and H3 methylation by Dot1. *J Biol Chem* 280, 9879–86. → pages 13
- Govin, J. and Berger, S. L. (2009). Genome reprogramming during sporulation. *Int J Dev Biol* 53, 425–32. → pages 4
- Grewal, S. I. S. and Elgin, S. C. R. (2007). Transcription and RNA interference in the formation of heterochromatin. *Nature* 447, 399–406. → pages 11
- Gu, W. and Roeder, R. G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* 90, 595–606. → pages 10
- Guillemette, B., Bataille, A., Gevry, N., Adam, M., Blanchette, M., Robert, F. and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol* 3, e384. → pages 8, 18
- Halley, J. E., Kaplan, T., Wang, A. Y., Kobor, M. S. and Rine, J. (2010). Roles for H2A.Z and Its Acetylation in GAL1 Transcription and Gene Induction, but Not GAL1-Transcriptional Memory. *PLoS Biol* 8, e1000401. → pages 4, 9
- Henry, K. W., Wyce, A., Lo, W.-S., Duggan, L. J., Emre, N. C. T., 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 Dev* 17, 2648–63. → pages 14
- Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. *Annu Rev Genet* 30, 405–39. → pages 14
- Hu, M., Sun, X.-J., Zhang, Y.-L., Kuang, Y., Hu, C.-Q., Wu, W.-L., Shen, S.-H., Du, T.-T., Li, H., He, F., Xiao, H.-S., Wang, Z.-G., Liu, T.-X., Lu, H., Huang,

- Q.-H., Chen, S.-J. and Chen, Z. (2010). Histone H3 lysine 36 methyltransferase Hypb/Setd2 is required for embryonic vascular remodeling. *Proc Natl Acad Sci USA* *107*, 2956–61. → pages 12
- Hua, S., Kallen, C. B., Dhar, R., Baquero, M. T., Mason, C. E., Russell, B. A., Shah, P. K., Liu, J., Khramtsov, A., Tretiakova, M. S., Krausz, T. N., Olopade, O. I., Rimm, D. L. and White, K. P. (2008). Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. *Mol Syst Biol* *4*, 188. → pages 7
- Huyen, Y., Zgheib, O., Ditullio, R. A., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S. and Halazonetis, T. D. (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* *432*, 406–11. → pages 16
- Hwang, W., Venkatasubrahmanyam, S., Ianculescu, A., Tong, A., Boone, C. and Madhani, H. (2003). A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell* *11*, 261–6. → pages 14
- Ingvarsdottir, K., Krogan, N. J., Emre, N. C. T., Wyce, A., Thompson, N. J., Emili, A., Hughes, T. R., Greenblatt, J. F. and Berger, S. L. (2005). H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol Cell Biol* *25*, 1162–72. → pages 14
- Iouzalen, N., Moreau, J. and Méchali, M. (1996). H2A.ZI, a new variant histone expressed during *Xenopus* early development exhibits several distinct features from the core histone H2A. *Nucleic Acids Res* *24*, 3947–52. → pages 6
- Jackson, J. D. and Gorovsky, M. A. (2000). Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. *Nucleic Acids Res* *28*, 3811–6. → pages 7
- Jacobson, R., Ladurner, A., King, D. and Tjian, R. (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* *288*, 1422–5. → pages 16

- Jeltsch, A. and Rathert, P. (2008). Putting the pieces together: histone H2B ubiquitylation directly stimulates histone H3K79 methylation. *Chembiochem* 9, 2193–5. → pages 14
- Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* (New York, NY) 293, 1074–80. → pages 5, 15
- Jiang, W., Guo, X. and Bhavanandan, V. P. (1998). Histone H2A.F/Z subfamily: the smallest member and the signature sequence. *Biochem Biophys Res Commun* 245, 613–7. → pages 6
- Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K. and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* 41, 941–5. → pages 18
- John, S., Sabo, P. J., Johnson, T. A., Sung, M.-H., Biddie, S. C., Lightman, S. L., Voss, T. C., Davis, S. R., Meltzer, P. S., Stamatoyannopoulos, J. A. and Hager, G. L. (2008). Interaction of the glucocorticoid receptor with the chromatin landscape. *Mol Cell* 29, 611–24. → pages 8
- Jones, R. S. and Gelbart, W. M. (1993). The *Drosophila* Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol Cell Biol* 13, 6357–66. → pages 16
- Joshi, A. A. and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* 20, 971–8. → pages 12
- Kalocsay, M., Hiller, N. J. and Jentsch, S. (2009). Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. *Mol Cell* 33, 335–43. → pages 7, 9
- Kazazian, H. H. (2004). Mobile elements: drivers of genome evolution. *Science* (New York, NY) 303, 1626–32. → pages 11

- Keogh, M., Mennella, T., Sawa, C., Berthelet, S., Krogan, N., Wolek, A., Podolny, V., Carpenter, L., Greenblatt, J., Baetz, K. and Buratowski, S. (2006). The *Saccharomyces cerevisiae* histone H2A variant Htz1 is acetylated by NuA4. *Genes Dev* 20, 660–5. → pages 9
- Keogh, M.-C., Kurdistani, S. K., Morris, S. A., Ahn, S. H., Podolny, V., Collins, S. R., Schuldiner, M., Chin, K., Punna, T., Thompson, N. J., Boone, C., Emili, A., Weissman, J. S., Hughes, T. R., Strahl, B. D., Grunstein, M., Greenblatt, J. F., Buratowski, S. and Krogan, N. J. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123, 593–605. → pages 12
- Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y. and Bedford, M. T. (2006). Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep* 7, 397–403. → pages 16
- Kimura, A. and Horikoshi, M. (1998). Tip60 acetylates six lysines of a specific class in core histones in vitro. *Genes Cells* 3, 789–800. → pages 6
- Kobor, M., Venkatasubrahmanyam, S., Meneghini, M., Gin, J., Jennings, J., Link, A., Madhani, H. and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* 2, E131. → pages 6
- Kornberg, R. and Lorch, Y. (2007). Chromatin rules. *Nat Struct Mol Biol* 14, 986–988. → pages 18
- Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science (New York, NY)* 184, 868–71. → pages 2
- Kornberg, R. D. (1977). Structure of chromatin. *Annu Rev Biochem* 46, 931–54. → pages 1
- Kornberg, R. D. and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* 98, 285–94. → pages 1

- Kossel, A. (1911). Ueber die chemische Beschaffenheit des Zellkerns. *Munchen Med. Wochenschrift* 58, 65–69. → pages 2
- Kossel, A. (1928). The protamines and histones. Longmans, Green and Company, New York. → pages 2
- Kouzarides, T. (2000). Acetylation: a regulatory modification to rival phosphorylation? *The EMBO journal* 19, 1176–9. → pages 10, 15
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. → pages 2, 4
- Krivtsov, A., Feng, Z., Lemieux, M., Faber, J., Vempati, S., Sinha, A., Xia, X., Jesneck, J., Bracken, A., Silverman, L., Kutok, J., Kung, A. and Armstrong, S. (2008). H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* 14, 355–68. → pages 13
- Krivtsov, A. V. and Armstrong, S. A. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* 7, 823–33. → pages 12
- Krogan, N., Baetz, K., Keogh, M., Datta, N., Sawa, C., Kwok, T., Thompson, N., Davey, M., Pootoolal, J., Hughes, T., Emili, A., Buratowski, S., Hieter, P. and Greenblatt, J. (2004). Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. *Proc Natl Acad Sci U S A* 101, 13513–8. → pages 7
- Krogan, N., Keogh, M., Datta, N., Sawa, C., Ryan, O., Ding, H., Haw, R., Pootoolal, J., Tong, A., Canadien, V., Richards, D., Wu, X., Emili, A., Hughes, T., Buratowski, S. and Greenblatt, J. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell* 12, 1565–76. → pages 6
- Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M. and Shilatifard, A. (2002). COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *The Journal of biological chemistry* 277, 10753–5. → pages 12

- 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. (2003a). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* 11, 721–9. → pages 11, 14
- Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S. and Greenblatt, J. (2003b). Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* 23, 4207–18. → pages 12
- Kusch, T., Florens, L., Macdonald, W., Swanson, S., Glaser, R., Yates, J., Abmayr, S., Washburn, M. and Workman, J. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* 306, 2084–7. → pages 6
- Lachner, M., O’Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–20. → pages 16
- Lachner, M., Sengupta, R., Schotta, G. and Jenuwein, T. (2004). Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. *Cold Spring Harb Symp Quant Biol* 69, 209–18. → pages 11
- Lacoste, N., Utley, R., Hunter, J., Poirier, G. and Cote, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. *J Biol Chem* 277, 30421–4. → pages 13
- Lan, F. and Shi, Y. (2009). Epigenetic regulation: methylation of histone and non-histone proteins. *Sci China, C, Life Sci* 52, 311–22. → pages 10, 11
- Larochelle, M. and Gaudreau, L. (2003). H2A.Z has a function reminiscent of an activator required for preferential binding to intergenic DNA. *The EMBO journal* 22, 4512–22. → pages 7, 8

- Latham, J. A. and Dent, S. Y. R. (2007). Cross-regulation of histone modifications. *Nat Struct Mol Biol* 14, 1017–1024. → pages 5
- Lee, K. K., Florens, L., Swanson, S. K., Washburn, M. P. and Workman, J. L. (2005). The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol Cell Biol* 25, 1173–82. → pages 14
- Lee, K. K. and Workman, J. L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* 8, 284–95. → pages 10
- Li, B., Carey, M. and Workman, J. L. (2007a). The role of chromatin during transcription. *Cell* 128, 707–19. → pages 17
- Li, B., Gogol, M., Carey, M., Pattenden, S. G., Seidel, C. and Workman, J. L. (2007b). Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* 21, 1422–30. → pages 12
- Li, B., Howe, L., Anderson, S., Yates, J. R. and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *The Journal of biological chemistry* 278, 8897–903. → pages 12
- Li, B., Pattenden, S., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J. and Workman, J. (2005). Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. *Proc Natl Acad Sci U S A* 102, 18385–90. → pages 8
- Li, H., Ilin, S., Wang, W., Duncan, E., Wysocka, J., Allis, C. and Patel, D. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442, 91–5. → pages 16
- Li, J., Moazed, D. and Gygi, S. P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *The Journal of biological chemistry* 277, 49383–8. → pages 12
- Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N. and Rando, O. J. (2005). Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol* 3, e328. → pages 10

- Liu, X., Li, B. and GorovskyMA (1996). Essential and nonessential histone H2A variants in *Tetrahymena thermophila*. *Mol Cell Biol* 16, 4305–11. → pages 6
- Lu, P. Y. T., Lévesque, N. and Kobor, M. S. (2009). NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and components. *Biochem Cell Biol* 87, 799–815. → pages 6
- Lu, X., Simon, M., Chodaparambil, J., Hansen, J., Shokat, K. and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol* 15, 1122–4. → pages 13
- Luger, K., Mader, A., Richmond, R., Sargent, D. and Richmond, T. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–60. → pages 3, 4, 14
- Malik, H. S. and Henikoff, S. (2003). Phylogenomics of the nucleosome. *Nat Struct Biol* 10, 882–91. → pages 6
- Martin, M., Cho, J., Cesare, A. J., Griffith, J. D. and Attardi, G. (2005). Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. *Cell* 123, 1227–40. → pages 3
- Masson, I. L., Yu, D., Jensen, K., Chevalier, A., Courbeyrette, R., Boulard, Y., Smith, M. and Mann, C. (2003). Yaf9, a novel NuA4 histone acetyltransferase subunit, is required for the cellular response to spindle stress in yeast. *Mol Cell Biol* 23, 6086–102. → pages 16
- Meneghini, M. D., Wu, M. and Madhani, H. D. (2003). Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* 112, 725–36. → pages 7, 20
- Miescher, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen. *Hoppe-Seyler, med. chem. Unters.* 4, 441–460. → pages 2
- Millar, C., Xu, F., Zhang, K. and Grunstein, M. (2006). Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev* 20, 711–22. → pages 9

- Millar, C. B. and Grunstein, M. (2006). Genome-wide patterns of histone modifications in yeast. *Nat Rev Mol Cell Biol* 7, 657–66. → pages 9
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F. and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci USA* 98, 12902–7. → pages 12
- 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, 711–23. → pages 13
- Misteli, T. (2005). Concepts in nuclear architecture. *Bioessays* 27, 477–87. → pages 3
- Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787–800. → pages 3
- Mizuguchi, G., Shen, X., Landry, J., Wu, W., Sen, S. and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303, 343–8. → pages 6
- Morrison, A. J. and Shen, X. (2009). Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat Rev Mol Cell Biol* 10, 373–84. → pages 7
- Muramoto, T., Müller, I., Thomas, G., Melvin, A. and Chubb, J. R. (2010). Methylation of H3K4 is required for inheritance of active transcriptional states. *Curr Biol* 20, 397–406. → pages 11
- Murray, K. (1964). The occurrence of epsilon-N-methyl lysine in histones. *Biochemistry* 3, 10–5. → pages 10
- Nagy, P. L., Griesenbeck, J., Kornberg, R. D. and Cleary, M. L. (2002). A trithorax-group complex purified from *Saccharomyces cerevisiae* is required for methylation of histone H3. *Proc Natl Acad Sci USA* 99, 90–4. → pages 12

- Nakanishi, S., Lee, J. S., Gardner, K. E., Gardner, J. M., Imai Takahashi, Y., Chandrasekharan, M. B., Sun, Z.-W., Osley, M. A., Strahl, B. D., Jaspersen, S. L. and Shilatifard, A. (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *J Cell Biol* *186*, 371–7. → pages 14
- Ng, H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y. and Struhl, K. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev* *16*, 1518–27. → pages 13
- Ng, H. H., Robert, F., Young, R. A. and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* *11*, 709–19. → pages 11, 12
- Nislow, C., Ray, E. and Pillus, L. (1997). SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol Biol Cell* *8*, 2421–36. → pages 20
- Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V., Su, L., Xu, G. and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. *Cell* *121*, 167–78. → pages 13
- Olins, A. L. and Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science (New York, NY)* *183*, 330–2. → pages 2
- Osley, M. A., Fleming, A. B. and Kao, C.-F. (2006). Histone ubiquitylation and the regulation of transcription. *Results Probl Cell Differ* *41*, 47–75. → pages 14
- O’Sullivan, J. M., Tan-Wong, S. M., Morillon, A., Lee, B., Coles, J., Mellor, J. and Proudfoot, N. J. (2004). Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* *36*, 1014–8. → pages 3

- Oudet, P., Gross-Bellard, M. and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4, 281–300. → pages 2
- Peña, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., Zhao, R. and Kutateladze, T. G. (2006). Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442, 100–3. → pages 16
- Pokholok, D., Harbison, C., Levine, S., Cole, M., Hannett, N., Lee, T., Bell, G., Walker, K., Rolfe, P., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. and Young, R. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 122, 517–27. → pages 10, 11, 12
- Powell, D. W., Weaver, C. M., Jennings, J. L., McAfee, K. J., He, Y., Weil, P. A. and Link, A. J. (2004). Cluster analysis of mass spectrometry data reveals a novel component of SAGA. *Mol Cell Biol* 24, 7249–59. → pages 14
- Ptashne, M. and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* 386, 569–77. → pages 18
- Radman-Livaja, M., Liu, C. L., Friedman, N., Schreiber, S. L. and Rando, O. J. (2010). Replication and active demethylation represent partially overlapping mechanisms for erasure of H3K4me3 in budding yeast. *PLoS Genet* 6, e1000837. → pages 11
- Raisner, R., Hartley, P., Meneghini, M., Bao, M., Liu, C., Schreiber, S., Rando, O. and Madhani, H. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* 123, 233–48. → pages 8, 18
- Rando, O. J. (2007). Global patterns of histone modifications. *Curr Opin Genet Dev* 17, 94–9. → pages 2
- Rayasam, G. V., Wendling, O., Angrand, P.-O., Mark, M., Niederreither, K., Song, L., Lerouge, T., Hager, G. L., Chambon, P. and Losson, R. (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *The EMBO journal* 22, 3153–63. → pages 12

- Reddy, K. L., Zullo, J. M., Bertolino, E. and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452, 243–7. → pages 4
- Rhodes, D. R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A. and Chinnaiyan, A. M. (2004). Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* 101, 9309–14. → pages 7
- Ridgway, P., Brown, K. D., Rangasamy, D., Svensson, U. and Tremethick, D. J. (2004). Unique residues on the H2A.Z containing nucleosome surface are important for *Xenopus laevis* development. *The Journal of biological chemistry* 279, 43815–20. → pages 6
- Roeder, R. G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett* 579, 909–15. → pages 18
- Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. and Stewart, A. F. (2001). The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *The EMBO journal* 20, 7137–48. → pages 12
- Rossmann, K. L., Der, C. J. and Sondek, J. (2005). GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nat Rev Mol Cell Biol* 6, 167–80. → pages 15
- Ruhl, D. D., Jin, J., Cai, Y., Swanson, S., Florens, L., Washburn, M. P., Conaway, R. C., Conaway, J. W. and Chrivia, J. C. (2006). Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry* 45, 5671–7. → pages 6
- Rusche, L. N., Kirchmaier, A. L. and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu Rev Biochem* 72, 481–516. → pages 11

- Ruthenburg, A., Li, H., Patel, D. and Allis, C. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* 8, 983–94. → pages 15, 17
- San-Segundo, P. and Roeder, G. (2000). Role for the silencing protein Dot1 in meiotic checkpoint control. *Mol Biol Cell* 11, 3601–15. → pages 13
- Santisteban, M. S., Kalashnikova, T. and Smith, M. M. (2000). Histone H2A.Z regulates transcription and is partially redundant with nucleosome remodeling complexes. *Cell* 103, 411–22. → pages 7, 8
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–11. → pages 11, 12
- Sarcinella, E., Zuzarte, P. C., Lau, P. N. I., Draker, R. and Cheung, P. (2007). Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol* 27, 6457–68. → pages 9
- 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, 43296–306. → pages 13
- Schaft, D., Roguev, A., Kotovic, K. M., Shevchenko, A., Sarov, M., Shevchenko, A., Neugebauer, K. M. and Stewart, A. F. (2003). The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Res* 31, 2475–82. → pages 12
- Schleker, T., Nagai, S. and Gasser, S. M. (2009). Posttranslational modifications of repair factors and histones in the cellular response to stalled replication forks. *DNA Repair (Amst)* 8, 1089–100. → pages 7
- Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C. and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* 6, 73–7. → pages 11, 12

- Schubeler, D., MacAlpine, D., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D., O'Neill, L., Turner, B., Delrow, J., Bell, S. and Groudine, M. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 18, 1263–71. → pages 11
- Schulze, J. M., Wang, A. Y. and Kobor, M. S. (2009). YEATS domain proteins: a diverse family with many links to chromatin modification and transcription. *Biochem Cell Biol* 87, 65–75. → pages 16, 17
- Schulze, J. M., Wang, A. Y. and Kobor, M. S. (2010). Reading chromatin: Insights from yeast into YEATS domain structure and function. *Epigenetics : official journal of the DNA Methylation Society* 5. → pages 17
- Schwartz, S. and Ast, G. (2010). Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing. *The EMBO journal* 29, 1629–36. → pages 13
- Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M. R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B. R., Ayer, D. E., Kutateladze, T. G., Shi, Y., Côté, J., Chua, K. F. and Gozani, O. (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* 442, 96–9. → pages 16
- Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* 75, 243–69. → pages 10, 11, 16
- Shilatifard, A. (2008). Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr Opin Cell Biol* 20, 341–8. → pages 12, 16
- Simpson, R. T. (1978). Structure of chromatin containing extensively acetylated H3 and H4. *Cell* 13, 691–9. → pages 4, 15

- Sims, R. J. and Reinberg, D. (2008). Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol* 9, 815–20. → pages 5
- Singer, M. S., Kahana, A., Wolf, A. J., Meisinger, L. L., Peterson, S. E., Goggin, C., Mahowald, M. and Gottschling, D. E. (1998). Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* 150, 613–32. → pages 13
- Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? *Genome Biol* 1, REVIEWS1011. → pages 11
- Stassen, M. J., Bailey, D., Nelson, S., Chinwalla, V. and Harte, P. J. (1995). The *Drosophila trithorax* proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech Dev* 52, 209–23. → pages 16
- Strahl, B. D. and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41–5. → pages 5, 15
- Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z.-W., Bone, J. R., Caldwell, J. A., Mollah, S., Cook, R. G., Shabanowitz, J., Hunt, D. F. and Allis, C. D. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* 22, 1298–306. → pages 12
- Suganuma, T. and Workman, J. L. (2008). Crosstalk among Histone Modifications. *Cell* 135, 604–7. → pages 5
- Sun, X.-J., Wei, J., Wu, X.-Y., Hu, M., Wang, L., Wang, H.-H., Zhang, Q.-H., Chen, S.-J., Huang, Q.-H. and Chen, Z. (2005). Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *The Journal of biological chemistry* 280, 35261–71. → pages 12
- Sun, Z. and Allis, C. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418, 104–8. → pages 14, 20

- Svotelis, A., Gévry, N., Grondin, G. and Gaudreau, L. (2010). H2A.Z overexpression promotes cellular proliferation of breast cancer cells. *Cell Cycle* 9, 364–70. → pages 7
- Talbert, P. B. and Henikoff, S. (2010). Histone variants—ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* 11, 264–75. → pages 5
- Taunton, J., Hassig, C. A. and Schreiber, S. L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science (New York, NY)* 272, 408–11. → pages 10
- Taverna, S., Li, H., Ruthenburg, A., Allis, C. and Patel, D. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 14, 1025–40. → pages 5, 15, 16
- Thatcher, T. H. and Gorovsky, M. A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res* 22, 174–9. → pages 6
- Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G. and Reuter, G. (1994). The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *The EMBO journal* 13, 3822–31. → pages 16
- Turner, B. M. (2000). Histone acetylation and an epigenetic code. *Bioessays* 22, 836–45. → pages 5
- Turner, B. M. (2007). Defining an epigenetic code. *Nat Cell Biol* 9, 2–6. → pages 5
- Ura, K., Kurumizaka, H., Dimitrov, S., Almouzni, G. and Wolffe, A. P. (1997). Histone acetylation: influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression. *The EMBO journal* 16, 2096–107. → pages 5, 15
- van Attikum, H., Fritsch, O. and Gasser, S. M. (2007). Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *The EMBO journal* 26, 4113–25. → pages 7

- van Daal, A. and Elgin, S. C. (1992). A histone variant, H2AvD, is essential in *Drosophila melanogaster*. *Mol Biol Cell* 3, 593–602. → pages 6
- van Driel, R., Fransz, P. F. and Verschure, P. J. (2003). The eukaryotic genome: a system regulated at different hierarchical levels. *J Cell Sci* 116, 4067–75. → pages 3
- van Leeuwen, F., Gafken, P. and Gottschling, D. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* 109, 745–56. → pages 13, 20
- Visser, A. E., Jaunin, F., Fakan, S. and Aten, J. A. (2000). High resolution analysis of interphase chromosome domains. *J Cell Sci* 113 (Pt 14), 2585–93. → pages 3
- Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M. Q. and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40, 897–903. → pages 10
- Wasylyk, B., Thevenin, G., Oudet, P. and Chambon, P. (1979). Transcription of in vitro assembled chromatin by *Escherichia coli* RNA polymerase. *J Mol Biol* 128, 411–40. → pages 18
- Weake, V. M. and Workman, J. L. (2008). Histone ubiquitination: triggering gene activity. *Mol Cell* 29, 653–63. → pages 14, 19
- Weake, V. M. and Workman, J. L. (2010). Inducible gene expression: diverse regulatory mechanisms. *Nat Rev Genet* 11, 426–37. → pages 18
- West, M. H. and Bonner, W. M. (1980). Histone 2A, a heteromorphous family of eight protein species. *Biochemistry* 19, 3238–45. → pages 6
- White, C. L., Suto, R. K. and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *The EMBO journal* 20, 5207–18. → pages 14

- Wong, M. M., Cox, L. K. and Chrivia, J. C. (2007). The chromatin remodeling protein, SRCAP, is critical for deposition of the histone variant H2A.Z at promoters. *The Journal of biological chemistry* 282, 26132–9. → pages 6
- Wood, A., Krogan, N., Dover, J., Schneider, J., Heidt, J., Boateng, M., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J., Johnston, M. and Shilatifard, A. (2003a). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* 11, 267–74. → pages 14
- Wood, A., Schneider, J., Dover, J., Johnston, M. and Shilatifard, A. (2003b). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *The Journal of biological chemistry* 278, 34739–42. → pages 14
- Wood, A., Shukla, A., Schneider, J., Lee, J. S., Stanton, J. D., Dzuiba, T., Swanson, S. K., Florens, L., Washburn, M. P., Wyrick, J., Bhaumik, S. R. and Shilatifard, A. (2007). Ctk complex-mediated regulation of histone methylation by COMPASS. *Mol Cell Biol* 27, 709–20. → pages 12
- Woodcock, C. L. (2006). Chromatin architecture. *Curr Opin Struct Biol* 16, 213–20. → pages 3
- Wu, R. S. and Bonner, W. M. (1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. *Cell* 27, 321–30. → pages 6
- Wyce, A., Xiao, T., Whelan, K. A., Kosman, C., Walter, W., Eick, D., Hughes, T. R., Krogan, N. J., Strahl, B. D. and Berger, S. L. (2007). H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Mol Cell* 27, 275–88. → pages 19
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H. and Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121, 859–72. → pages 16
- Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., Wu, C. and Allis, C. D. (2006). A

PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442, 86–90. → pages 16

Wysocki, R., Javaheri, A., Allard, S., Sha, F., Cote, J. and Kron, S. (2005). Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkpoint functions of Rad9. *Mol Cell Biol* 25, 8430–43. → pages 13

Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., Hall, M. C., Borchers, C. H. and Strahl, B. D. (2003). Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev* 17, 654–63. → pages 12

Xu, D., Bai, J., Duan, Q., Costa, M. and Dai, W. (2009). Covalent modifications of histones during mitosis and meiosis. *Cell Cycle* 8, 3688–94. → pages 4

Yang, X.-J. (2004). Lysine acetylation and the bromodomain: a new partnership for signaling. *Bioessays* 26, 1076–87. → pages 16

Yang, X.-J. and Seto, E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26, 5310–8. → pages 10

yi Lin, Y., ying Lu, J., Zhang, J., Walter, W., Dang, W., Wan, J., Tao, S.-C., Qian, J., Zhao, Y., Boeke, J. D., Berger, S. L. and Zhu, H. (2009). Protein acetylation microarray reveals that NuA4 controls key metabolic target regulating gluconeogenesis. *Cell* 136, 1073–84. → pages 10

Zemach, A., McDaniel, I. E., Silva, P. and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science (New York, NY)* 328, 916–9. → pages 7

Zhang, H., Roberts, D. N. and Cairns, B. R. (2005). Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* 123, 219–31. → pages 8

Zilberman, D., Coleman-Derr, D., Ballinger, T. and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 456, 125–9. → pages 7

- Zlatanova, J. and Thakar, A. (2008). H2A.Z: view from the top. *Structure* *16*, 166–79. → pages 15
- Zofall, M., Fischer, T., Zhang, K., Zhou, M., Cui, B., Veenstra, T. D. and Grewal, S. I. S. (2009). Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature* *461*, 419–22. → pages 7
- Zucchi, I., Mento, E., Kuznetsov, V. A., Scotti, M., Valsecchi, V., Simionati, B., Vicinanza, E., Valle, G., Pilotti, S., Reinbold, R., Vezzoni, P., Albertini, A. and Dulbecco, R. (2004). Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis. *Proc Natl Acad Sci USA* *101*, 18147–52. → pages 7

Chapter 2

Asf1-like structure of the conserved Yaf9 YEATS domain and role in H2A.Z deposition and acetylation*

2.1 Introduction

Histones are the major protein constituent of chromatin and can be modified in several fundamental ways, including the addition of posttranslational modifications, ATP-dependent chromatin remodelling, and incorporation of histone variants (Kusch and Workman, 2007). H2A.Z, encoded by the *HTZ1* gene in *Saccharomyces cerevisiae*, is an H2A variant with roles in transcriptional repression and activation, chromosome segregation, DNA replication and repair, and heterochromatin-euchromatin boundary formation (Guillemette and Gaudreau, 2006). H2A.Z is incorporated into nucleosomes by the conserved ATP-dependent chromatin remodelling complex SWR1-C, the first complex discovered to be dedicated

*A version of this chapter has been published. Wang AY, Schulze JM, Skordalakes E, Gin JW, Berger JM, Rine J, Kobor MS. (2009). Asf1-like structure of the conserved Yaf9 YEATS domain and role in H2A.Z deposition and acetylation. *Proc Natl Acad Sci USA*. 106, 21573-8. See Co-authorship Statement on page xv for details about my contributions.

to variant histone deposition (Kobor et al., 2004; Krogan et al., 2003; Mizuguchi et al., 2004). SWR1-C shares four subunits with the NuA4 histone H4 acetyltransferase complex, which among other substrates, also acetylates H2A.Z to restrict spreading of heterochromatin and to prevent chromosome missegregation (Babiarz et al., 2006; Keogh et al., 2006; Millar et al., 2006). One of the shared subunits, Yaf9, is important for cellular response to spindle stress, proper DNA repair and metabolism, H2A.Z chromatin deposition and acetylation, and histone H4 acetylation at telomere-proximal genes (Keogh et al., 2006; Masson et al., 2003; Wu et al., 2005; Zhang et al., 2004).

Yaf9 contains an evolutionarily conserved YEATS domain found in more than a hundred different eukaryotic proteins (Schulze et al., 2009). In human, GAS41 is the closest relative of Yaf9 based on sequence identity, domain organization, and its presence in human SRCAP and TIP60 complexes, which respectively correspond to yeast SWR1-C and NuA4 (Masson et al., 2003; Zhang et al., 2004). Yaf9 and GAS41 have an N-terminal YEATS domain followed by a conserved region of unknown function called the A-box and a C-terminal coiled-coil domain (Zhang et al., 2004). Interestingly, YEATS domain proteins have several connections to cancer. First, GAS41 is highly amplified in human glioblastomas and astrocytomas (Fischer et al., 1997) and is required for repression of the p53 tumour suppressor pathway during normal cellular proliferation (Park and Roeder, 2006). Second, two other YEATS domain containing proteins, ENL and AF9, are frequent fusion partners of the mixed-lineage leukemia protein in leukemia (Daser and Rabbitts, 2004).

This study established the conservation of YEATS domain function from yeast to human. To better understand the physical organization of the YEATS domain, we determined the structure of this region from Yaf9 to 2.3 Å resolution. Interestingly, the Yaf9 YEATS domain structure was highly similar to the structure of the Asf1 histone chaperone, a congruence that extended to an ability of Yaf9 to bind histones H3 and H4 *in vitro*. Using structure-function analysis in yeast, we found that the YEATS domain was required for Yaf9 function, H2A.Z deposition at specific promoters, and H2A.Z acetylation.

2.2 Results

2.2.1 YEATS domain function was conserved from yeast to human

The primary sequence similarity between human GAS41 and yeast Yaf9 predicted a functional conservation. We tested whether any of the three domains (YEATS domain, A-box and coiled-coil) of GAS41 could substitute for its Yaf9 counterpart in yeast. Genes encoding for hybrid proteins, each carrying a distinct module from either GAS41 or Yaf9 (Figure 2.1 (a), page 51), were tested for their ability to complement growth phenotypes caused by loss of *YAF9*. Strains lacking Yaf9 are sensitive to chemical stressors including formamide, hydroxyurea (HU), and benomyl (Schulze et al., 2009). Hybrid proteins carrying the YEATS domain of GAS41 and at least the coiled-coil region of Yaf9 rescued most of the growth defects caused by loss of Yaf9 (Figure 2.1 (b), page 51). In contrast, hybrid proteins carrying the GAS41 coiled-coil and the Yaf9 YEATS domain were unable to support growth of strains lacking Yaf9, irrespective of the source of the A-box (Figure 2.1 (b), page 51). Expression of the human *GAS41* construct alone was unable to complement any of the phenotypes caused by loss of *YAF9* (Figure 2.1 (b), page 51), as was expected from the GAS41-Yaf9 hybrid results. The function of the GAS41-Yaf9 hybrid proteins reflected their incorporation into SWR1-C and NuA4. Functional TAP-GAS41-Yaf9 hybrids and full length TAP-Yaf9 protein all co-purified with SWR1-C subunits Swr1, Swc2 and Swc3, as well as NuA4 subunits Eaf1, Epl1 and Tra1 (Figure 2.1 (c), page 51). In contrast, much less SWR1-C and NuA4 co-purified with hybrids containing the GAS41 coiled-coil or with full-length TAP-GAS41.

2.2.2 The YEATS domain of Yaf9 adopted an immunoglobulin fold

To better understand the function of the evolutionary conserved Yaf9 YEATS domain, we determined its high-resolution structure using X-ray crystallography. The crystallized fragment (Yaf9 amino acids 8–171 determined by mass spectrometry) contained the YEATS domain and A-box, suggesting that these components of Yaf9 form a stable structural element, with a less-structured linker connecting to the coiled-coil domain.

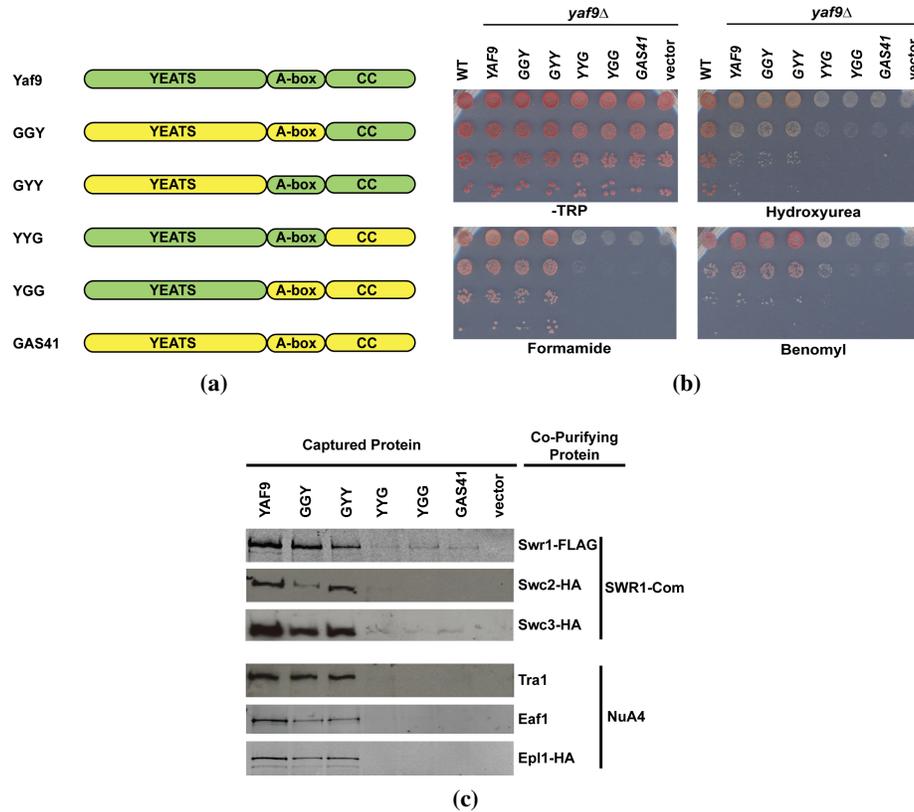


Figure 2.1: YEATS domain function was conserved from yeast and human. (a): Schematic representation of Yaf9-GAS41 hybrid proteins constructed with modules originating from Yaf9 in green and GAS41 in yellow. The nomenclature indicates the origin of a given module where *G* refers to GAS41 and *Y* to Yaf9, in order of YEATS domain, A-box, and coiled-coil. For simplicity, the N-terminal TAP-tag present in all constructs was omitted. (b): GAS41 YEATS domain conferred resistance to genotoxic stress in yeast when fused to Yaf9 coiled-coil. Tenfold serial dilutions of *yaf9Δ* strains carrying the indicated plasmids were plated and incubated on CSM plates with the following concentrations of chemicals: 2 % formamide, 30 $\mu\text{g/ml}$ benomyl and 75 mmol/l HU. (c): GAS41 YEATS domain fused to Yaf9 coiled-coil was incorporated into SWR1-C and NuA4. Analytical-scale affinity purifications of TAP-Yaf9-GAS41 hybrids from cells containing affinity-tagged versions of three representative SWR1-C and three representative NuA4 subunits were performed and immunoblotted for the indicated proteins.

The Yaf9 YEATS domain crystallized in the space group P6₅22. The structure of the selenomethionine-substituted protein was determined by multi-wavelength anomalous dispersion, and refined to a final resolution of 2.3 Å (Table 2.1, page 52). Three copies of the YEATS domain were present in the asymmetric unit. The final model included residues 8–119 and 143–169 for each of the three protomers, and was refined to an R_{work}/R_{free} of 21.8/25.8% with no residues in disallowed regions of Ramachandran space (Table 2.1, page 52). The three protomers were highly similar to each other, displaying an average $rmsd \approx 0.25$ Å across all atoms.

Table 2.1: Data collection, phasing analysis, and refinement statistics.

$^{\S}R_{sym} = \sum |I_i - \langle I_i \rangle| / \sum |I_i|$ where I_i is the scaled intensity of the i th measurement and $\langle I_i \rangle$ is the mean intensity for that reflection. $^{\dagger}R_{work} = ||F_o| - |F_c|| / |F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. $^{\ddagger}R_{free}$ was calculated as R_{work} using 5% of randomly selected data not included in refinement.

Yaf9 _(8–171)	Native	Se-Derivative	
	λ	Se- λ 1	Se- λ 2
Wavelength, Å	1.11587	0.9795	1.0199
Space group	P6522	P6522	
Cell dimensions, Å	84.725, 288.212	84.582, 286.531	84.636, 286.639
Resolution, Å	50–2.3 (2.42–2.3)	50–2.90 (3.02–2.90)	50–2.90 (3.02–2.9)
Redundancy	6.3 (5.0)	3.8 (3.8)	3.8 (3.8)
Completeness, (%)	95.8 (92.8)	100 (100)	100 (100)
R_{sym} , (%) [§]	7.5 (36.8)	13.5 (43.1)	12.3 (41.4)
$I/\sigma(I)$	16.2 (4.4)	8.7 (2.9)	9.2 (3.0)
Phasing Analysis			
Resolution		40-2.9	
No. sites		6	
Mean figure of merit		0.36	
Refinement Statistics			
Resolution, Å		20-2.3	
R_{work} , (%) [†]		21.8	
R_{free} , (%) [‡]		25.8	
RMSD bonds, Å		0.009	
RMSD angles, °		1.143	

Table 2.1 – continued

Yaf9 _(8–171)	Native	Se-Derivative	
	λ	Se- λ 1	Se- λ 2
No. protein atoms		3671	
No. water atoms		200	
Ramachandran % (no res.)			
Most favored		89.7	
Allowed		10	
Gen. allowed		0.3	

The YEATS domain folded into an elongated β -sandwich consisting of eight antiparallel β -strands capped on one end by two short α -helices (Figure 2.2 (a), (b), page 55). Structural comparison of the Yaf9 YEATS domain against the protein databank using the DALI server showed that it adopted an Immunoglobulin (Ig) fold (Holm and Sander, 1995). Ig folds are common macromolecular interaction modules, and are found in proteins with a variety of cellular functions (Bork et al., 1994). The YEATS domain was structurally homologous to the Ig folds of a broad number of factors (over 200 hits with Z-scores smoothly varying from 3–6), with no evidence of any particular standout among fold homologs.

2.2.3 Structural features of the YEATS domain

Our structure revealed three interesting physical features of the YEATS domain. One feature was the presence of a highly conserved cleft located on the end of the Ig fold opposite the two capping helices (Figure 2.2 (b), (d), page 56).

This region was composed of three surface-exposed loops emanating from the core β -sheet region. The first was a His-Thr-His (HTH) triad, which preceded strand β 2 and was invariant in Yaf9/GAS41 family members, but not in Taf14 and Sas5, the two other YEATS domain-containing proteins in yeast (Schulze et al., 2009). The other two loops included a Leu-His-X-Ser/Thr-Tyr/Phe (LHx(S/T)(Y/F)) pentad that connected strands β 3 and β 4, and a Gly-Trp-Gly sequence wedged between strands β 5 and β 6. Both of these segments were invariant across all YEATS proteins and constituted the defining signature sequence motifs of the clade.

The presence of this conservation within the external loops of an Ig fold was intriguing, as these regions often form the principal surface used by such proteins to engage client molecules. A second feature consisted of a relatively shallow groove near the N- and C-termini of the YEATS domain (coloured yellow in Figure 2.2 (d), page 56).

Formed in part by the capping helices, the groove was relatively nonpolar, and displayed only a modest degree of surface amino acid conservation. The groove was distinguished, however, by the presence of a narrow, but deep, hydrophobic pocket that extruded $\sim 7\text{--}8$ Å down into the core of the β -sheet region (Figure 2.2 (d), page 56, also see Figure 2.5 (a), page 61).

The hydrophobic character of the groove and pocket was conserved among all YEATS domains. Moreover, in the crystal, the groove served to bind the extended N-terminal arm of an adjacent YEATS domain protomer, such that the N-terminal tail formed an extra strand of the $\beta 4\text{-}3\text{-}6\text{-}7$ sheet, in *trans*. This interaction was recapitulated between the three YEATS domains present in the asymmetric unit to form a trimeric array of protomers with rotational, three-fold non-crystallographic symmetry (Figure 2.2 (c), page 56).

Although the YEATS fragment we crystallized was monomeric in solution (as determined by gel filtration and dynamic light scattering), the surface features of the groove and its ability to associate with the N-terminus of a partner molecule suggested that the groove may also constitute a peptide binding surface.

The third feature, between the cleft and putative peptide-binding groove, was a region rich in conserved charged residues, particularly basic amino acids. Although the charged patch did not have any notable structural features, such as a deep depression or any particularly marked curvature, it was one of the most electropositive surfaces on the YEATS domain (Figure 2.2 (b), page 56).

2.2.4 Yaf9 shared structural and biochemical properties with the Asf1 histone chaperone

In the course of analyzing the Yaf9 YEATS domain, we noted a marked similarity to the histone chaperone Asf1 (Figure 2.3 (a), (b), page 58). Like Yaf9, the core region of Asf1 is also predicated on an Ig fold. Moreover, of the several different topological groups of Ig fold categorized to date, both proteins belong to the same

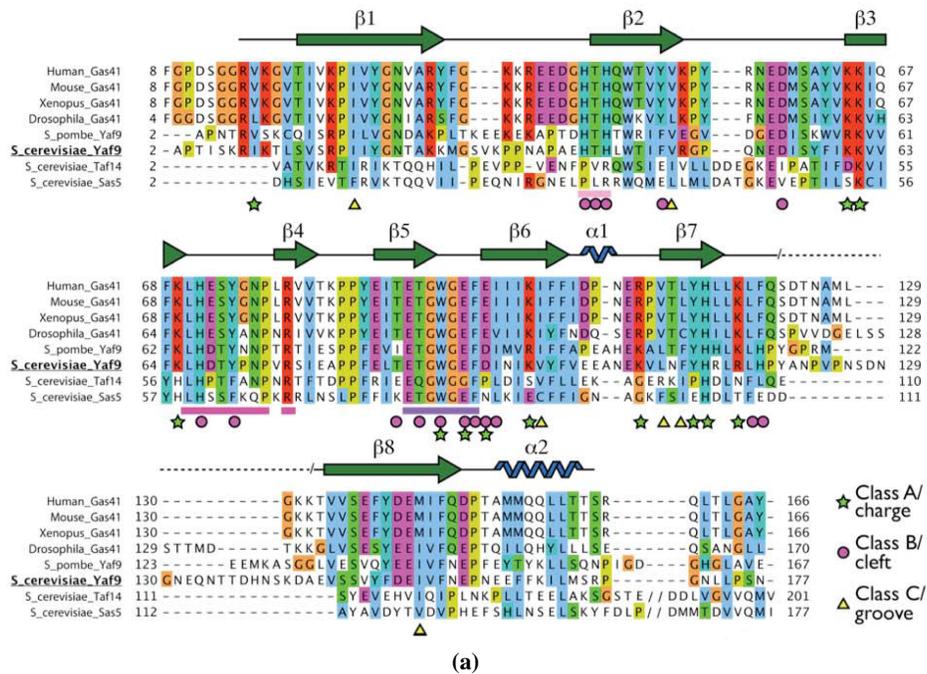


Figure 2.2: Secondary structure and amino acid conservation of YEATS domain. (a): Sequence alignment of YEATS domains showing amino acid conservation. Secondary structure elements and specific classes of mutations are labeled.

‘switched’ Ig subclass (Figure 2.4, page 60). Although the surface features of the two proteins are relatively distinct, for example, Asf1 lacks the deep conserved cleft formed by apical loops of the Yaf9 Ig fold, there were some intriguing similarities beyond the level of fold topology. In particular, Asf1 binds short peptide segments using the edges of its β -sheets, such as with the Hir1 B-box of CAF1 or the C-terminal tail of histone H4 (English et al., 2006; Natsume et al., 2007; Antczak et al., 2006). In our structure, the Yaf9 YEATS domain used one of its β -sheets and a hydrophobic groove to associate with the N-terminal peptide of an adjoining protomer in a similar fashion (Figure 2.2 (c), page 56).

To test whether the structural similarity between the Yaf9 YEATS domain and Asf1 was reflected in a genetic interaction, we used an assay for which double mutant haploid segregants were obtained from a diploid strain with only one copy of *ASF1*

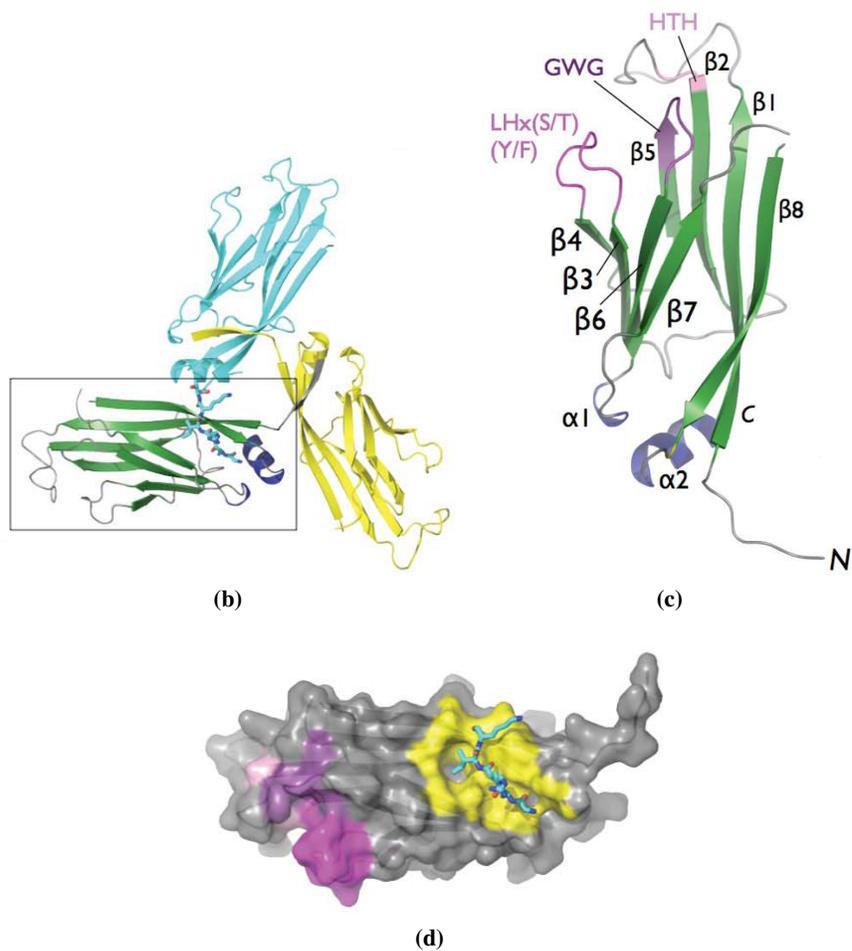


Figure 2.2: Secondary structure and amino acid conservation of YEATS domain. (b): Ribbon diagram of the Yaf9 YEATS domain. Structural elements are labeled and conserved sequence motifs highlighted. This and all other molecular graphics figures were generated using PYMOL (DeLano, 2002). (c): View of the three YEATS domains protomers (green, cyan, and yellow) as related by noncrystallographic symmetry. The extended N-terminal tail of each domain docks into the hydrophobic groove of a neighbouring molecule, forming a pseudocontinuous sheet with strand $\beta 7$ across the domain. (d): Surface view of the Yaf9 YEATS domain (boxed region in (c)) highlighting the conserved cleft motifs (purple, magenta, and pink), and hydrophobic groove (yellow). The N-terminus of an adjacent protomer is shown as cyan sticks.

and *YAF9*. The *asf1Δyaf9Δ* double mutants were viable but grew dramatically more slowly than either single mutant, particularly at a higher temperature or on media with low levels of HU (Figure 2.3 (c), page 58). Consistent with *ASF1* being required for acetylation of H3K56 (Recht et al., 2006), the *yaf9ΔH3K56R* double mutant had a synthetic growth defect (Figure 2.3 (d), page 59). Moreover, this interaction was likely due to Yaf9's role in H2A.Z deposition but not H2A.Z acetylation, as the *htz1ΔH3K56R* double mutant also exhibited a synergistic growth defect while the *htz1K3,8,10,14RH3K56R* double mutant did not (Figure 2.3 (d), page 59).

Given the diversity of folds involved in chromatin remodelling, it was unexpected to find that the Yaf9 YEATS domain was a structural homolog of Asf1, a chaperone for histones H3-H4. Consistent with the structural similarity between the Yaf9 YEATS domain and Asf1, *in vitro* protein-interaction assays showed that GST-Yaf9 bound to histones H3 and H4 (Figure 2.3 (e), page 59). These associations were not likely to reflect non-specific binding to basic charged proteins as Yaf9 did not bind H2B.

2.2.5 Conserved residues on the YEATS domain surface were important for Yaf9 function

To better understand the role of the YEATS domain *in vivo*, we constructed three classes of mutant Yaf9 proteins. Each mutant carried multiple amino acid substitutions of conserved residues on the surface of one of the three structural features. Class A mutants were in the charged surface region of the protein (*yaf9-2*, *yaf9-4*, and *yaf9-24*). Class B mutants were in the conserved cleft (*yaf9-1*, *yaf9-3*, *yaf9-34*, along with *yaf9-23* and various dissections of it: *yaf9-26*, *yaf9-27*, *yaf9-28*, *yaf9-29*, *yaf9-30*, *yaf9-31*, *yaf9-32* and *yaf9-33*). Class C mutants were in the groove that associated with the N-terminal segment of Yaf9 with potential to be a peptide binding groove (*yaf9-17*, *yaf9-18*, *yaf9-19*, *yaf9-20*, *yaf9-21*, and *yaf9-22*). The locations of the altered residues in the proteins encoded by *yaf9* alleles are shown in Figure 2.2 (a) and 2.5 (a), page 61, with an exact description in Table 2.3, page 76.

Yaf9 YEATS domain mutants were tested for their ability to complement the sensitivity of *yaf9Δ* strains to genotoxic agents formamide, HU and benomyl. Drug-sensitivity phenotypes were conferred by mutations in two of the three con-

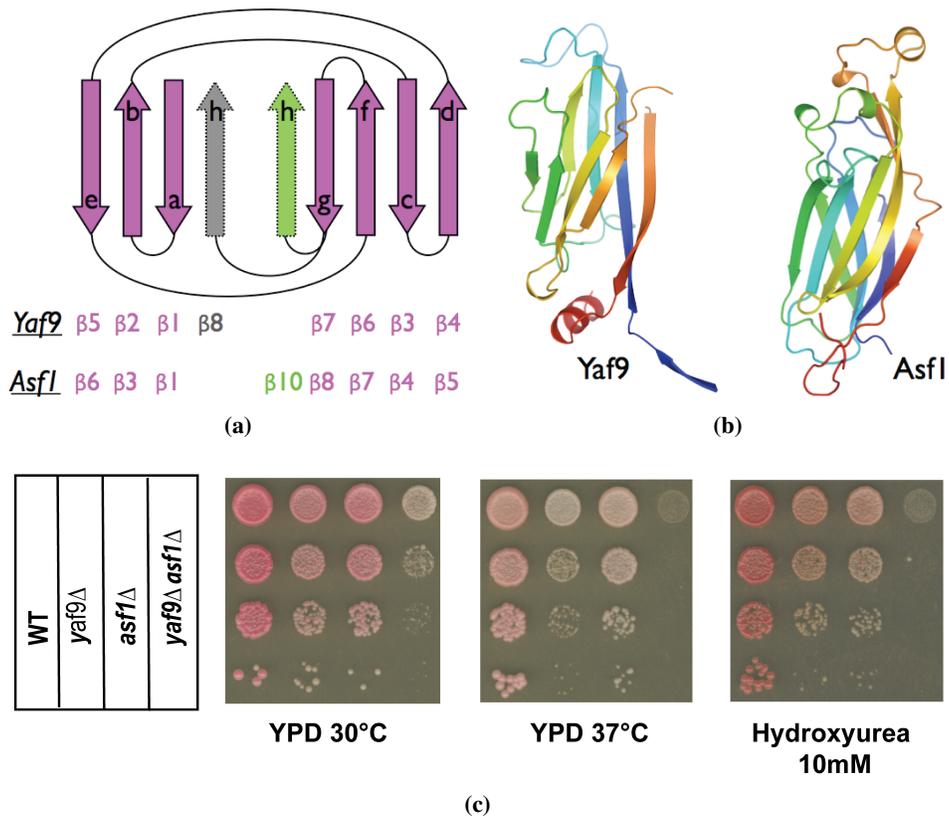


Figure 2.3: Yaf9 YEATS domain structure was similar to histone chaperone Asf1. (a): Topology diagram of the Yaf9 YEATS domain and Asf1 switched-type Ig folds. Conserved β -strands are shown as purple arrows. A swap of the last strand has occurred between the two proteins, pairing the ‘h’ strand of Yaf9 and Asf1 with either the ‘a-b-e’ sheet (gray) or ‘g-f-c-d’ sheet (green). The corresponding β -strands (Daganzo et al., 2003) are shown. (b): Structural comparison of the Yaf9 YEATS domain and Asf1 core. The ribbon diagram is rainbow coloured from the N-terminus (blue) to the C-terminus (red) of each protein. (c): *YAF9* and *ASF1* interacted genetically. Cells lacking both *ASF1* and *YAF9* were hypersensitive to high temperature (37 °C) and HU.

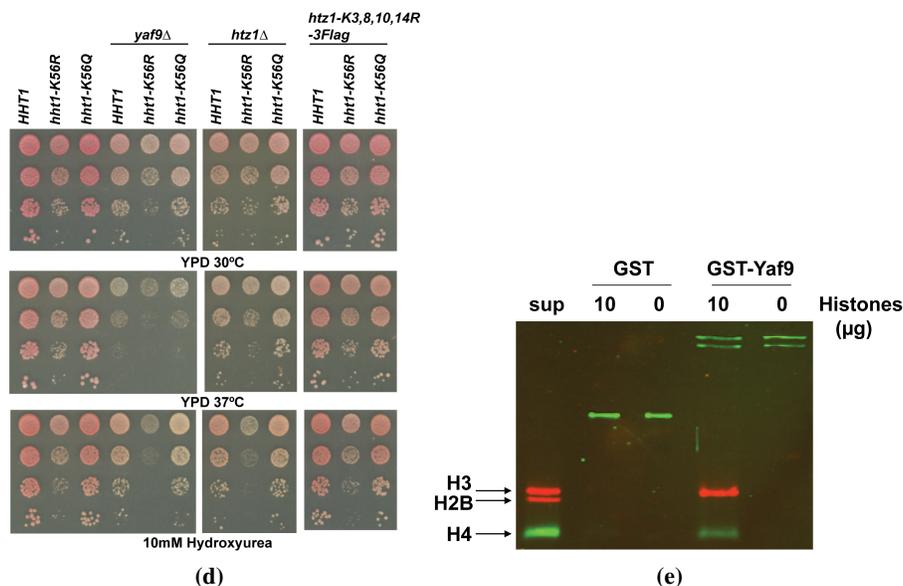


Figure 2.3: Yaf9 YEATS domain structure was similar to histone chaperone Asf1. (d): *YAF9* and *HTZ1* interacted genetically with H3K56R. Cells expressing H3K56R in combination with either *yaf9*Δ or *htz1*Δ were hypersensitive to HU. Cells expressing H3K56R *htz1-K3,8,10,14R* showed no obvious synthetic interaction. (e): Yaf9 bound to histones H3 and H4 *in vitro*. Calf thymus histones bound to the indicated amounts of purified GST or GST-Yaf9 from yeast as judged by immunoblotting with α -H3, α -H4, and α -H2B histone antibodies.

served areas to varying degrees, with Class A allele *yaf9-4* and Class B alleles *yaf9-23*, *yaf9-27*, *yaf9-28* having growth defects similar to those of *yaf9*Δ (Figure 2.5 (d), page 62). The *yaf9-1* strain had growth defects similar to those of *yaf9*Δ on formamide and HU, but grew similar to wild type on benomyl, while *yaf9-3* and *yaf9-34* strains had growth defects only on formamide and grew comparable to wild type on HU and benomyl.

These results hinted at regional specialization of function, with *yaf9-1* being defective in processes leading to formamide and HU sensitivity, whereas the other Class B alleles *yaf9-3* and *yaf9-34* were defective in a process leading only to formamide sensitivity. Strains with mutations in the putative peptide-binding pocket (Class C) had no discernable phenotypes (Figure 2.5 (d), page 62), suggesting this

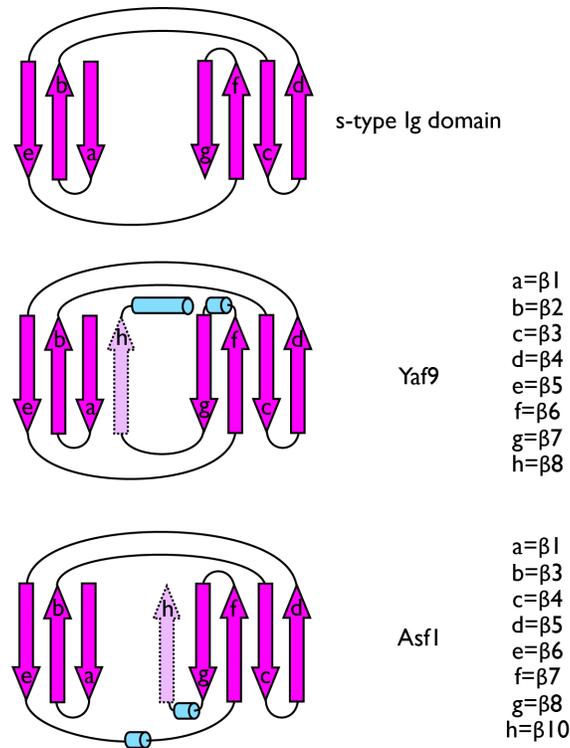


Figure 2.4: Similarity between the topology of Yaf9 YEATS domain and histone chaperone Asf1. Topology diagram of the switched-type Ig folds as well as the Yaf9 YEATS domain and Asf1. Conserved β -strands are shown as purple arrows, and α -helices are shown in cyan cylinders. The two proteins have a swap of the last strand, pairing the ‘h’ strand of Yaf9 and Asf1 with either the ‘a-b-e’ sheet or ‘g-f-c’ sheet, respectively.

feature played no discernable role in the Yaf9-dependent functions tested here. In addition to the 20 alleles shown in Figure 2.5 (d), page 62, and Figure 2.6 (a), page 63, 13 alleles with mutations in other conserved residues were tested without revealing noticeable phenotypes (Table 2.4, page 77).

The amount of Yaf9 mutant protein was similar to the level in wild type strains in most mutants, including the *yaf9-1* and *yaf9-3* strains (Figure 2.7 (a), page 64), further supporting the view that the different phenotypes resulted from qualitative rather than quantitative differences in Yaf9 function. Overexpression experiments of alleles encoding Yaf9 protein with normal or reduced levels showed that the

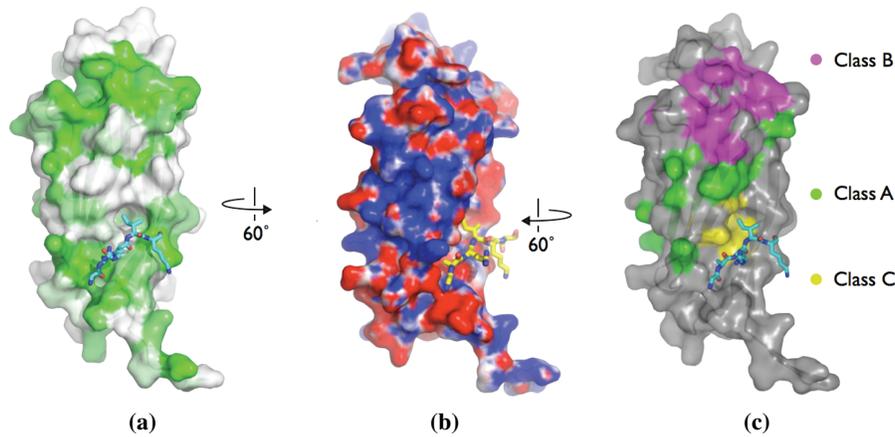


Figure 2.5: Mutations in conserved surface residues in the Yaf9 YEATS domain affected protein function. (a): Yaf9₍₈₋₁₇₁₎ surface representation showing YEATS residue conservation as mapped by Consurf (dark green = invariant, light green = conserved, white = variable) (Glaser et al., 2003), based on an alignment of more than 30 Yaf9 and GAS41 orthologs (Sas5 and Taf14 were excluded). The peptide (cyan sticks)/hydrophobic groove interaction seen between Yaf9 protomers is highlighted. (b): Surface stereorepresentation of Yaf9₍₈₋₁₇₁₎ showing charge distribution (red = negative, blue = positive, contoured at +/- 5 k_BT by APBS) (Baker et al., 2001). The view is rotated approximately 60° to that shown in (a). (c): Surface of Yaf9₍₈₋₁₇₁₎ showing the where the three mutant classes map with respect to the structure and each other. The orientation is the same as in (a).

phenotypes of these mutants were not solely due to reduced protein level and that many of the *yaf9* alleles behaved as hypomorphs (Figure 2.7 (a), page 64).

The enhanced growth defect of the *yaf9Δ asf1Δ* double mutant was contributed by the YEATS domain, as a subset of *yaf9* alleles were unable to complement the loss of *YAF9* in an *asf1Δ* strain and failed to restore growth on HU and higher temperature (37 °C) (Figure 2.5 (e), page 62). Therefore, as the structural similarity implied, it was indeed the YEATS domain, and in particular the charged surface region and cleft of Yaf9, whose function was linked to Asf1. Interestingly, the subset of *yaf9* alleles with lower protein levels (*yaf9-4*, *yaf9-23* and *yaf9-28*) consistently

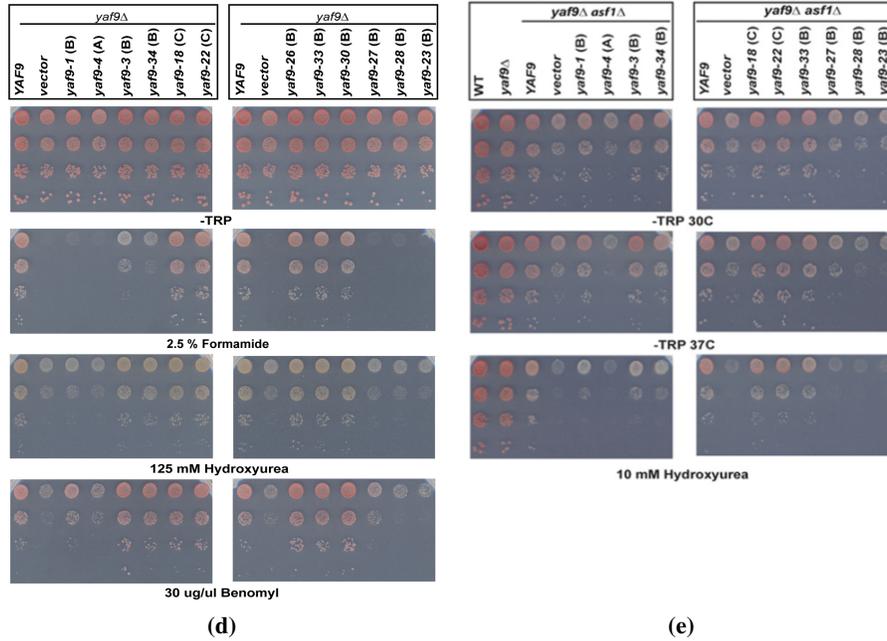


Figure 2.5: Mutations in conserved surface residues in the Yaf9 YEATS domain affected protein function. (d): Cells with *yaf9* alleles of Classes A and B were sensitive to genotoxic stressors. (e): Yaf9 YEATS domain was involved in similar process with Asf1. The alleles carrying mutations in the charged surface or conserved cleft were unable to rescue growth defects of the *yaf9Δasf1Δ* double mutant. Note that the *yaf9-4*, *yaf9-23* and *yaf9-28* alleles caused approximately a tenfold greater growth retardation of the *yaf9Δasf1Δ* strain than empty vector alone in cells grown at 37 °C.

caused approximately a 10-fold greater growth retardation of the *asf1Δ yaf9Δ* strain than empty vector alone in strains grown at 37 °C (Figure 2.5 (e), page 62). Thus, although the absence of *ASF1* magnified the growth defect of *yaf9Δ* strains, it magnified the growth defect of the *yaf9-4*, *yaf9-23* and *yaf9-28* mutants even further.

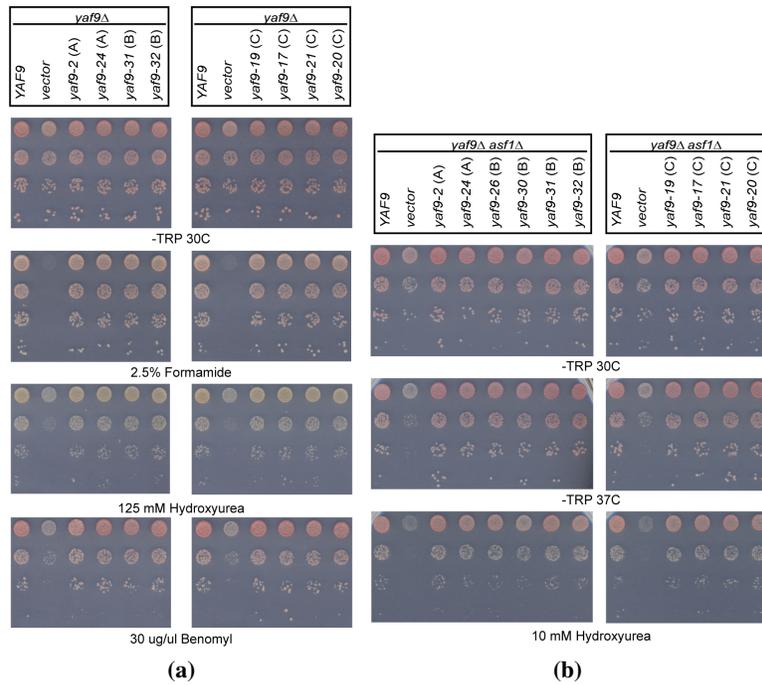


Figure 2.6: Growth comparison of Yaf9 mutants that did not have phenotypes nor genetic interaction with Asf1. (a): Cells with the indicated *YAF9* mutations were not sensitive to genotoxic stressors. Ten-fold serial dilution from stationary overnight cultures of *yaf9* Δ strains carrying the indicated plasmids were plated and incubated at 30 °C for 2–3 days. CSM plates lacking tryptophan with the indicated concentrations of chemicals were used. (b): Growth comparison of Yaf9 YEATS domain mutants that were not involved in the shared function with Asf1. Tenfold serial dilution from stationary overnight cultures of *yaf9* Δ *asf1* Δ double deletion strains carrying different *yaf9* mutant alleles were plated and incubated at 30 °C, 37 °C, and in the presence of 10 mmol/l hydroxyurea for 2–3 days.

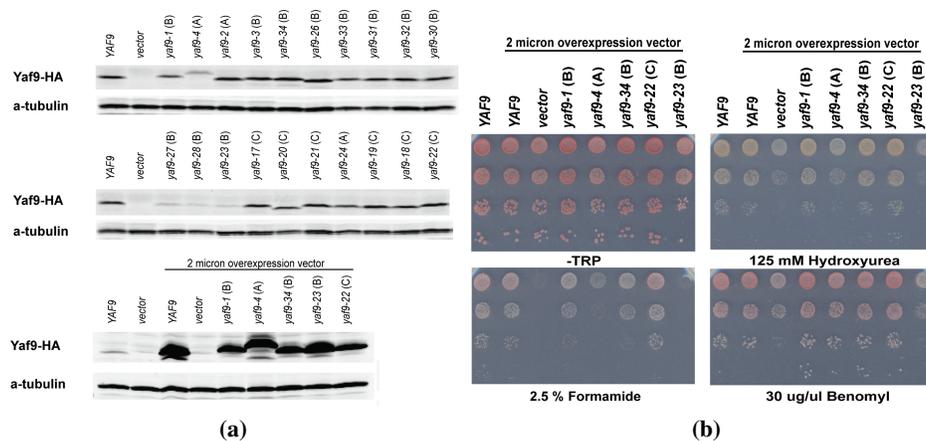
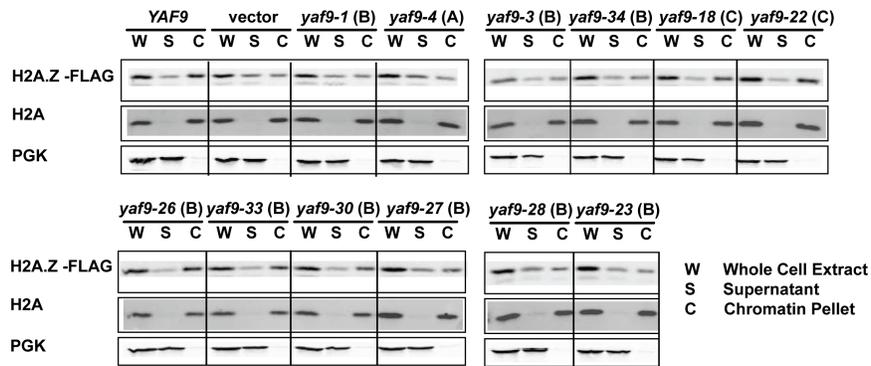


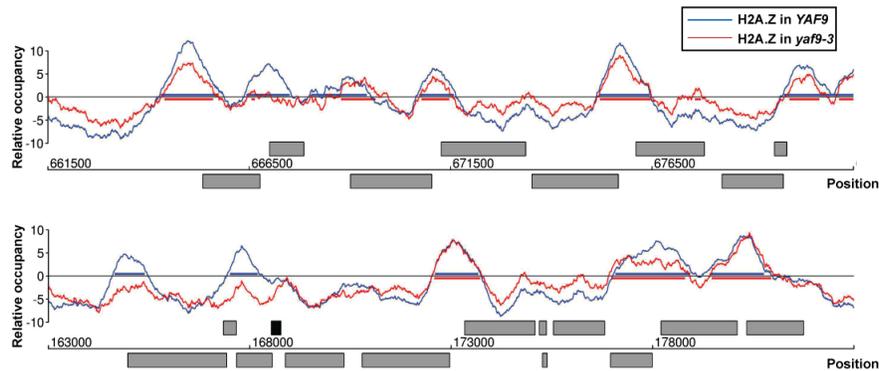
Figure 2.7: Protein levels and overexpression of Yaf9 mutants. (a): Protein levels of Yaf9 with mutations in conserved residues were similar to wild-type except for the protein encoded by the *yaf9-4*, *yaf9-27*, *yaf9-28*, and *yaf9-23* alleles which had somewhat lower levels. Note that *yaf9-4* also migrates slower through the SDS/PAGE gel due to its altered charge properties. Overexpression of the mutant Yaf9 proteins restores their expression levels to that of wild type. Shown is an immunoblot analysis of yeast whole cell extracts, in which levels of Yaf9-HA were compared to each other. An antibody against α -tubulin was used as loading control. (b): Overexpression of mutant Yaf9 proteins revealed their hypomorphic nature. The indicated Yaf9 alleles were overexpressed from the constitutive *GPD1* promoter on a multicopy plasmid and tested for their sensitivity to genotoxic stressors, similar to Figure 2.7 (a).

2.2.6 Yaf9 YEATS domain functioned in both SWR1-C and NuA4

To examine the role of the Yaf9 YEATS domain in histone variant H2A.Z biology, we assayed the activity of *yaf9* mutants biochemically. Bulk chromatin fractionation assays showed that, as expected for cells lacking a component of the SWR1-C, *yaf9* Δ strains had reduced H2A.Z levels in the chromatin pellet and increased levels in the non-chromatin supernatant fraction as compared to wild type (Figure 2.8 (a), page 65). In contrast, the level of H2A in chromatin and the level of Pgc1 in supernatant, were unchanged. Strains with Yaf9 mutations located in the charged surface area (Class A allele *yaf9-4*) or in the conserved cleft (Class B alleles *yaf9-1*,



(a)



(b)

Figure 2.8: Yaf9 YEATS domain was required for H2A.Z chromatin deposition and acetylation. (a): Yaf9 YEATS domain mutant strains had decreased H2A.Z-FLAG chromatin deposition. W, whole cell extract; S, supernatant; C, chromatin pellet. The relative amount of H2A.Z-FLAG in each fraction was determined by immunoblotting in the different strains. Antibodies against histone H3 and Pgc1 were used as loading controls for chromatin pellet and supernatant, respectively. (b): ChIP-on-chip profiles of H2A.Z in *YAF9* and *yaf9-3* strains. Sample genomic positions for chromosomes 4 and 8 were plotted along the x-axis against the relative occupancy of H2A.Z on the y-axis. ORFs are indicated as light gray rectangles above the x-axis for Watson genes and below the x-axis for Crick genes. ARS are indicated as dark gray rectangles. Regions considered enriched above a certain threshold are shown as coloured bars on the x-axis.

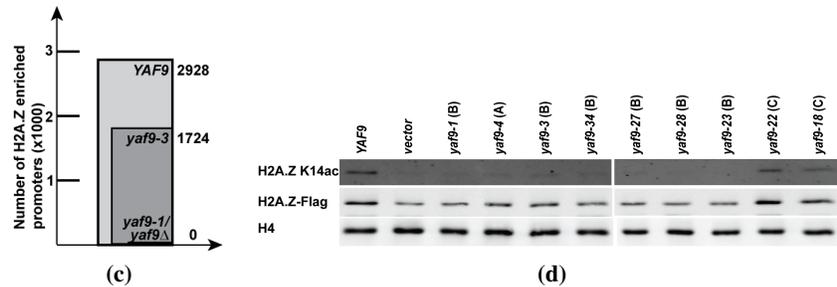


Figure 2.8: Yaf9 YEATS domain was required for H2A.Z chromatin deposition and acetylation. (c): Mutation in Yaf9 YEATS domain resulted in loss of H2A.Z at specific promoters. Shown is the number of H2A.Z enriched promoters determined by ChIP-on-chip in *YAF9* and *yaf9-3* strains. Both *yaf9Δ* and *yaf9-1* strains had low immunoprecipitation efficiency resulting in no detection of H2A.Z enriched regions above background level. (d): Mutations in the Yaf9 YEATS domain resulted in decreased H2A.Z K14ac in chromatin. Chromatin extracts from the bulk chromatin association assays shown in (a) were immunoblotted with an antibody against H2A.Z K14ac. Antibodies against FLAG and H4 were used as loading controls.

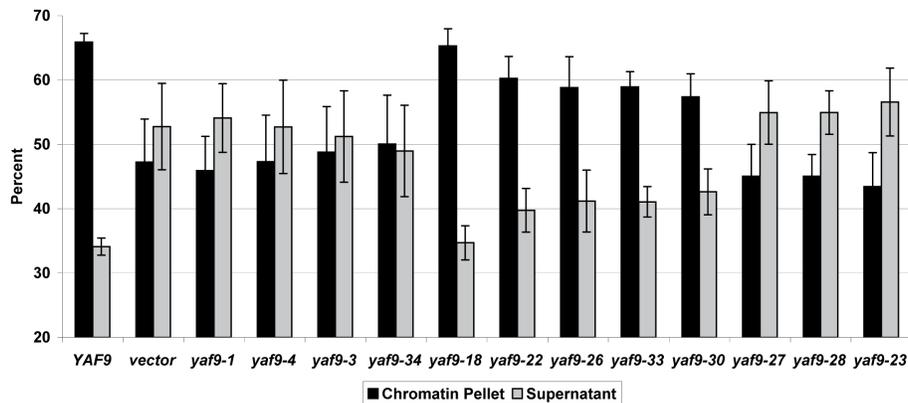


Figure 2.9: Yaf9 YEATS domain was required for H2A.Z chromatin deposition. Quantitation of four independent sets of chromatin association assays, similar to the ones presented in Figure 2.8 (a). Shown is the average percentage of H2A.Z-FLAG present in chromatin and supernatant as a function of the sum of chromatin and supernatant for each respective strain after normalization to the loading controls, i. e. chromatin pellet of H2A and supernatant of Pgk1.

yaf9-23, *yaf9-27*, *yaf9-28*, and *yaf9-34*) had less H2A.Z in the chromatin pellet, comparable to the *yaf9* Δ strain (Figure 2.8 (a), page 65, and 2.9, page 66). Therefore, residues in the conserved charged surface (Class A) and cleft (Class B) of the Yaf9 YEATS domain were important for H2A.Z deposition.

To further explore the conclusions derived from these crude chromatin association assays, the requirement of the Yaf9 YEATS domain for H2A.Z deposition at specific promoters was determined by chromatin immunoprecipitation (ChIP)-on-chip. H2A.Z occupancy was compared between *YAF9* and the *yaf9-1* and *yaf9-3* mutants, which were chosen based on their differences in growth phenotypes (Figure 2.5 (d)). Consistent with previous studies, H2A.Z was present at 2928 promoters in wild type strains, using our enrichment criteria (Figure 2.8 (b), page 65, (c), page 66) (Guillemette and Gaudreau, 2006). In contrast, H2A.Z-ChIP efficiency in both *yaf9* Δ and *yaf9-1* mutants was equal to a non-antibody control resulting in background peaks (Figure 2.8 (c), page 66), as reported previously for other SWR1-C subunits (Li et al., 2005).

Interestingly, in the *yaf9-3* mutant, H2A.Z was specifically lost at about one third of promoters but was present at almost wild type levels at the other two-thirds (Figure 2.8 (b), page 65, (c), page 66). Further supporting the functional link between Yaf9 and Asf1, promoters reported to contain H3K56ac (Rufiange et al., 2007) preferentially lost H2A.Z in the *yaf9-3* mutant ($p < 10^{-8}$). However, gene ontology (GO) analysis of the subsets of genes that either lost or kept H2A.Z in the *yaf9-3* mutant, revealed no significant association with any GO molecular function.

Acetylation of H2A.Z on K14 is catalyzed, at least in part, by NuA4 following H2A.Z deposition into chromatin (Babiarz et al., 2006; Keogh et al., 2006; Millar et al., 2006). Yaf9 is required for H2A.Z acetylation *in vivo*, however, it is not known whether this dependency solely reflects Yaf's contribution to SWR1-C, or Yaf9's contribution to NuA4 function, or both (Keogh et al., 2006).

To determine the role of the Yaf9 YEATS domain in H2A.Z acetylation, we used the chromatin-associated H2A.Z fraction that remained in each *yaf9* mutant strain (Figure 2.8 (a)) and measured acetylation at K14. This approach allowed for the evaluation of NuA4-dependent activities of Yaf9 mutants uncoupled from their

H2A.Z deposition effects. Strains carrying *yaf9* alleles in the charged surface region (*yaf9-4*) or in the conserved cleft (*yaf9-1*, *yaf9-3*, *yaf9-23*, *yaf9-27*, *yaf9-28*, and *yaf9-34*) exhibited reduced H2A.Z K14 acetylation, similar to that of *yaf9* Δ strains (Figure 2.8 (d), page 66).

2.3 Discussion

YEATS domain proteins are found in many protein complexes involved in chromatin biology and are linked to cancers in humans. Despite these intriguing connections, little is known about the specific functions of YEATS domains, or their interaction partners and structure. Here, we demonstrate that there is a functional conservation between the YEATS domain of yeast Yaf9 and human GAS41. The extent of incorporation of the YEATS domain hybrid proteins into SWR1-C and NuA4 closely paralleled their ability to function in place of Yaf9. Furthermore, these results confirmed earlier findings that the coiled-coil of Yaf9 is required for interactions with NuA4 (Zhang et al., 2004), and extended these findings to define the coiled-coil region to also be important for Yaf9 to associate with SWR1-C. Likewise, the GAS41 coiled-coil is required for interaction with the human TIP60 and SRCAP complexes (Park and Roeder, 2006). The ability of the human GAS41 YEATS domain to complement in yeast showed that YEATS domain function was evolutionary conserved and implied that its structure was very similar.

The first structure of a YEATS domain reported here revealed that this domain in Yaf9 consisted of an Ig fold with three distinct conserved surface features, of which at least two were important for function. The strong defects in H2A.Z chromatin deposition caused by mutations in Yaf9 established that the YEATS domain was required for the proper activity of SWR1-C. Promoter-specific measurements of H2A.Z occupancy in two *yaf9* mutants suggested that different amino acid changes in the YEATS domain had distinct effects on H2A.Z deposition. The phenotypically moderate *yaf9-3* mutant surprisingly lost H2A.Z at a particular set of promoters, while retaining normal levels at the remaining promoters. This contrasted with the more severe *yaf9-1* mutant which completely lost H2A.Z at all promoters, similar to the *yaf9* Δ . Furthermore, the YEATS domain was important for H2A.Z K14ac by NuA4 in chromatin, which most likely occurs after H2A.Z has been deposited

by SWR1-C (Babiarz et al., 2006; Keogh et al., 2006). A common function of Yaf9 as part of the module of subunits shared between SWR1-C and NuA4 (Yaf9, Swc4, Arp4 and Act1) was consistent with recent data suggesting that this module has similar roles in both complexes (Wu et al., 2008).

From a conceptual viewpoint, the results from the mutant analysis aimed at determining the functional relevance of the three conserved YEATS domain features were counter-intuitive. *A priori*, one might assume that a protein such as Yaf9, which must function in two different complexes, would be even more constrained than a protein that acts on its own or in only one complex and hence might be more vulnerable to mutation. However, of the 33 mutant alleles that we constructed, 26 had no discernable phenotype despite the 40 mutated amino acids among these 26 mutants. Indeed, no phenotype was observed from single amino acid substitutions even though these were conserved residues. This was an unexpectedly low frequency of phenotypes, especially considering the functional conservation of the YEATS domain from yeast Yaf9 to human GAS41 reported here. Of the seven mutants with phenotypes, four exhibited reduced protein levels (*yaf9-4*, *yaf9-23*, *yaf9-27* and *yaf9-28*), and three (*yaf9-1*, as well as *yaf9-3* and *yaf9-34*; differing by the I94M substitution) had normal protein levels. These results had practical significance for forward genetic studies suggesting that even for non-essential subunits of protein complexes like SWR1-C, discovering their function from mutant screens could be unexpectedly difficult.

Given the lack of amino acid sequence similarity, it was unexpected to discover that the Yaf9 YEATS domain was a structural homolog of the histone chaperone Asf1. Similar to Asf1, Yaf9 bound to histones H3-H4 *in vitro*. The enhanced growth defect of the *yaf9Δ asf1Δ* mutant indicated an involvement of Yaf9 and Asf1 in a similar process. More precisely, we determined that loss of Asf1-dependent H3K56 acetylation caused enhanced growth defects in the absence of either Yaf9 or H2A.Z, but had no effect in the absence of H2A.Z acetylation. This suggested that it was primarily Yaf9's SWR1-C dependent role in H2A.Z deposition and not its role in NuA4-mediated H2A.Z acetylation that was affected by loss of Asf1. The lack of interaction with the N-terminal lysines of H2A.Z contrasts with the previously reported requirement for Asf1 and H3K56 acetylation in cells with mu-

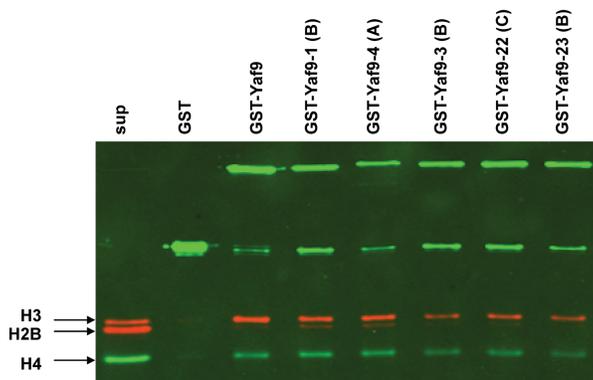


Figure 2.10: Yaf9 YEATS domain mutants bound to histones H3 and H4 *in vitro*. Yaf9 mutants interacted with histones H3 and H4 *in vitro*. Calf thymus histones were bound to the indicated amounts of purified GST or GST-Yaf9 carrying the *yaf9-1*, *yaf9-4*, *yaf9-3*, *yaf9-22* and *yaf9-23* alleles.

tations in the N-terminal lysines of H3 or H4 (Li et al., 2008).

The similar process that Asf1 and Yaf9 ordinarily facilitate might be reflected in our finding that three *yaf9* alleles with lower protein levels had greater growth retardation than the *yaf9Δ asf1Δ* strain with empty vector. It would appear that the three mutant proteins interfered with whatever process Asf1 and Yaf9 converge on. A possible explanation may involve nucleosome assembly and disassembly by Asf1, presumably through its ability to bind to H3-H4 dimer intermediates (Li et al., 2008; Williams et al., 2008). SWR1-C performs a conceptually similar function, disassembling H2A-H2B dimers from nucleosomes and reassembling them with H2A.Z-H2B dimers (Mizuguchi et al., 2004). Perhaps in the absence of Asf1, SWR1-C promotes a more extensive dismantling and reassembly of nucleosomes in a Yaf9-dependent way. If so, the paradoxical phenotype of *yaf9-4*, *yaf9-23* and *yaf9-28* being more defective than *yaf9Δ* in cells lacking Asf1 might reflect their competence in only the disassembly reaction, leaving chromatin in a more compromised state than if Yaf9 were completely absent.

We note that histone binding might be a common feature of YEATS domains since human ENL, through its YEATS domain, specifically binds to histones H3 and

H1, without having affinity for H4, H2A and H2B (Zeisig et al., 2005). However, our data suggested that additional relevant targets must exist to facilitate YEATS domain-dependent activities, at least in the case of Yaf9, because mutations in the YEATS domain of Yaf9 that affected H2A.Z chromatin deposition still bound to histones H3 and H4 (Figure 2.10, page 70). This eliminates a simple model by which diminished interaction of Yaf9 with H3 or H4 results in loss of Yaf9-dependent H2A.Z chromatin deposition, and raises the need for more detailed binding studies.

In summary, this study for the first time established the structure of the YEATS domain, its evolutionary conservation, a close structural and functional relationship with Asf1, as well as a differential requirement of the Yaf9 YEATS domain for H2A.Z deposition at specific genes.

2.4 Experimental procedures

2.4.1 Yeast strains, plasmids, and yeast techniques

All strains used in this study are listed in Table 2.2, page 72, and were created using standard yeast genetic techniques (Ausubel, 1987). Complete deletion of genes and integration of a triple HA-tag (Longtine et al., 1998) or triple FLAG tag (Gelbart et al., 2001) in frame at the 3' end of genes were achieved using the one-step gene integration of PCR-amplified modules (Longtine et al., 1998). Plasmid shuffling experiments were performed using 5-FOA as a counter-selecting agent for the *URA3* plasmid p(*URA3, CEN, ARS, HHT2-HHF2*), and shuffling in plasmids containing histone H3K56 mutations (*TRP1, CEN, hht2-HHF2*).

The *YAF9* gene, including 300 bp up- and downstream of the ORF, was PCR-amplified from genomic DNA and cloned into the centromeric vector plasmid pRS314 (*TRP1*). The plasmid was subsequently modified by integrating the triple HA-tag downstream of the gene using site-directed mutagenesis using the adapted protocol of the QuickChange method (Stratagene). Mutations in specific amino acids of *YAF9* were generated by site-directed mutagenesis in a similar manner as described above following the protocol of the manufacturer (all primers used are

available upon request). Multiple rounds of mutagenesis were conducted for mutations that are at large distances apart in the primary sequence. All mutations were confirmed by DNA sequencing and are listed in Table 2.3, page 76. To determine the expression level of the mutated Yaf9 proteins, yeast whole cell extracts were prepared using the NaOH extraction protocol as previously described (Kushnirov, 2000). Immunoblotting was performed using anti-HA (Applied Biological Materials) and anti- α -tubulin (Sigma) antibodies.

YAF9 and human *GAS41* were cloned into a Gateway Entry vector (Invitrogen). For this purpose, *GAS41* sequences were amplified by PCR from a cDNA clone generously provided by Ivan Still (Roswell Park Cancer Institute) (Lauffart et al., 2002). Four *GAS41*-Yaf9 hybrids (YGG [Yaf9 1-118, *GAS41* 122-227]; YYG [Yaf9 1-171, *GAS41* 161-227]; GGY [*GAS41* 1-160, Yaf9 172-226]; GYY [*GAS41* 1-121, Yaf9 119-226]) were constructed in the respective Gateway Entry vectors by amplifying DNA sequences corresponding to individual segments and replacing the corresponding parental segments using a modified insertion mutagenesis by the QuickChange method described above (details available upon request). The parental *YAF9* and *GAS41* sequences, along with the four hybrid constructs were subsequently transposed into the Gateway destination vector *pVV221*, encoding an N-terminal TAP-tag under the control of the *tetO* promoter, which is constitutively active in the absence of tetracycline (Mullem et al., 2003), using standard procedures.

Table 2.2: Yeast strains used in this study. All strains were constructed for this work (details available upon request) or from the laboratory collections, except for DDY1810 (source, D. Drubin). All strains are in the W303 background, except for DDY1810, which is S288C.

Name	Relevant Genotype
DDY1810	<i>MATA leu2 ura3-52 trp1 prb1-1122 pep4-3 pre1-451</i>
MKY5	<i>MATα ade2-1 can1-100 his3-11 leu2-3,112 trp1-1 ura3-1 LYS2</i>
MKY7	<i>MATA ade2-1 can1-100 his3-11 leu2-3,112 trp1-1 ura3-1 LYS2</i>
MKY844	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221, 5'TAP-YAF9]</i>
MKY845	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221]</i>
MKY846	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221, 5'TAP-GAS41]</i>
MKY847	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221, 5'TAP-GGY]</i>

Table 2.2 – continued

Name	Relevant Genotype
MKY848	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221, 5'TAP-GYY]</i>
MKY849	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221, 5'TAP-YYG]</i>
MKY850	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pVV221, 5'TAP-YGG]</i>
MKY227	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221]</i>
MKY228	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221, 5'TAP-YAF9]</i>
MKY229	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221, 5'TAP-GAS41]</i>
MKY230	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221, 5'TAP-GGY]</i>
MKY231	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221, 5'TAP-GYY]</i>
MKY232	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221, 5'TAP-YYG]</i>
MKY233	MKY5, <i>yaf9Δ::HIS5 SWC2-3HA::KanMX6 [pVV221, 5'TAP-YGG]</i>
MKY235	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221]</i>
MKY236	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221, 5'TAP-YAF9]</i>
MKY237	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221, 5'TAP-GAS41]</i>
MKY238	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221, 5'TAP-GGY]</i>
MKY239	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221, 5'TAP-GYY]</i>
MKY240	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221, 5'TAP-YYG]</i>
MKY241	MKY5, <i>yaf9Δ::HIS5 SWC3-3HA::KanMX6 [pVV221, 5'TAP-YGG]</i>
MKY284	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221]</i>
MKY285	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221, 5'TAP-YAF9]</i>
MKY286	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221, 5'TAP-GAS41]</i>
MKY287	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221, 5'TAP-GGY]</i>
MKY288	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221, 5'TAP-GYY]</i>
MKY289	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221, 5'TAP-YYG]</i>
MKY290	MKY5, <i>yaf9Δ::HIS5 SWR1-3Flag::KanMX6 [pVV221, 5'TAP-YGG]</i>
MKY888	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221]</i>
MKY889	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221, 5'TAP-YAF9]</i>
MKY890	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221, 5'TAP-GAS41]</i>
MKY891	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221, 5'TAP-GGY]</i>
MKY892	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221, 5'TAP-GYY]</i>
MKY893	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221, 5'TAP-YYG]</i>
MKY894	MKY5, <i>yaf9Δ::HIS5 EPL1-3HA::KanMX6 [pVV221, 5'TAP-YGG]</i>
MKY256	MKY7, <i>asf1Δ::HIS5</i>
MKY257	MKY7, <i>yaf9Δ::KanMX6</i>
MKY258	MKY7, <i>yaf9Δ::KanMX6 asf1Δ::HIS5</i>
MKY874	DDY1810, <i>[pYEX-4T3, GST-YAF9]</i>

Table 2.2 – continued

Name	Relevant Genotype
MKY875	DDY1810, [<i>pYEX-4T3</i> , <i>GST</i>]
MKY876	DDY1810, [<i>pYEX-4T3</i> , <i>GST-yaf9-1</i>]
MKY877	DDY1810, [<i>pYEX-4T3</i> , <i>GST-yaf9-3</i>]
MKY878	DDY1810, [<i>pYEX-4T3</i> , <i>GST-yaf9-4</i>]
MKY879	DDY1810, [<i>pYEX-4T3</i> , <i>GST-yaf9-22</i>]
MKY880	DDY1810, [<i>pYEX-4T3</i> , <i>GST-yaf9-23</i>]
MKY819	MKY5, [<i>pRS314</i>]
MKY820	MKY5, <i>yaf9Δ::HIS5 p[CEN, ARS, TRP1, YAF9-3HA]</i>
MKY821	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, YAF9-3HA]</i>
MKY822	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314]</i>
MKY823	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-1-3HA]</i>
MKY824	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-2-3HA]</i>
MKY825	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-3-3HA]</i>
MKY826	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-4-3HA]</i>
MKY827	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-17-3HA]</i>
MKY828	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-18-3HA]</i>
MKY829	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-19-3HA]</i>
MKY830	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-20-3HA]</i>
MKY831	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-21-3HA]</i>
MKY832	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-22-3HA]</i>
MKY833	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-23-3HA]</i>
MKY834	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-24-3HA]</i>
MKY835	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-26-3HA]</i>
MKY836	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-27-3HA]</i>
MKY837	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-28-3HA]</i>
MKY838	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-29-3HA]</i>
MKY839	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-30-3HA]</i>
MKY840	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-31-3HA]</i>
MKY841	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-32-3HA]</i>
MKY842	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-33-3HA]</i>
MKY843	MKY7, <i>yaf9Δ::HIS5 HTZ1-3Flag::KanMX6 [pRS314, yaf9-34-3HA]</i>
MKY881	MKY5, <i>yaf9Δ::HIS5 p[2μ, TRP1, pGPD::YAF9-3HA::HIS5]</i>
MKY882	MKY5, <i>yaf9Δ::HIS5 p[2μ, TRP1, pGPD]</i>
MKY883	MKY5, <i>yaf9Δ::HIS5 p[2μ, TRP1, pGPD::yaf9-1-3HA]</i>
MKY884	MKY5, <i>yaf9Δ::HIS5 p[2μ, TRP1, pGPD::yaf9-4-3HA]</i>

Table 2.2 – continued

Name	Relevant Genotype
MKY885	MKY5, <i>yaf9</i> Δ:: <i>HIS5 p[2μ, TRP1, pGPD::yaf9-22-3HA]</i>
MKY886	MKY5, <i>yaf9</i> Δ:: <i>HIS5 p[2μ, TRP1, pGPD::yaf9-23-3HA]</i>
MKY887	MKY5, <i>yaf9</i> Δ:: <i>HIS5 p[2μ, TRP1, pGPD::yaf9-34-3HA]</i>
MKY851	MKY258, [<i>pRS314, YAF9-3HA</i>]
MKY852	MKY258, [<i>pRS314</i>]
MKY853	MKY258, [<i>pRS314, yaf9-1-3HA</i>]
MKY854	MKY258, [<i>pRS314, yaf9-2-3HA</i>]
MKY855	MKY258, [<i>pRS314, yaf9-3-3HA</i>]
MKY856	MKY258, [<i>pRS314, yaf9-4-3HA</i>]
MKY857	MKY258, [<i>pRS314, yaf9-17-3HA</i>]
MKY858	MKY258, [<i>pRS314, yaf9-18-3HA</i>]
MKY859	MKY258, [<i>pRS314, yaf9-19-3HA</i>]
MKY860	MKY258, [<i>pRS314, yaf9-20-3HA</i>]
MKY861	MKY258, [<i>pRS314, yaf9-21-3HA</i>]
MKY862	MKY258, [<i>pRS314, yaf9-22-3HA</i>]
MKY863	MKY258, [<i>pRS314, yaf9-23-3HA</i>]
MKY864	MKY258, [<i>pRS314, yaf9-24-3HA</i>]
MKY865	MKY258, [<i>pRS314, yaf9-26-3HA</i>]
MKY866	MKY258, [<i>pRS314, yaf9-27-3HA</i>]
MKY867	MKY258, [<i>pRS314, yaf9-28-3HA</i>]
MKY868	MKY258, [<i>pRS314, yaf9-29-3HA</i>]
MKY869	MKY258, [<i>pRS314, yaf9-30-3HA</i>]
MKY870	MKY258, [<i>pRS314, yaf9-31-3HA</i>]
MKY871	MKY258, [<i>pRS314, yaf9-32-3HA</i>]
MKY872	MKY258, [<i>pRS314, yaf9-33-3HA</i>]
MKY873	MKY258, [<i>pRS314, yaf9-34-3HA</i>]
MKY1015	<i>MATα (hhf1-hht1)Δ::HphMX(hhf2-hht2)Δ::NatMX htz1-K3,8,10,14R-3Flag::KanMX [pRS314, HHT2-HHF2]</i>
MKY1016	<i>MATα (hhf1-hht1)Δ::HphMX(hhf2-hht2)Δ::NatMX htz1-K3,8,10,14R-3Flag::KanMX [pRS314, hht2K56R-HHF2]</i>
MKY1017	<i>MATα (hhf1-hht1)Δ::HphMX(hhf2-hht2)Δ::NatMX htz1-K3,8,10,14R-3Flag::KanMX [pRS314, hht2K56Q-HHF2]</i>
MKY1018	<i>MATA (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 yaf9::KAN [pRS314, HHT2-HHF2]</i>

Table 2.2 – continued

Name	Relevant Genotype
MKY1019	<i>MATA (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 yaf9::KAN [pRS314, hht2K56R-HHF2]</i>
MKY1020	<i>MATA (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 yaf9::KAN [pRS314, hht2K56Q-HHF2]</i>
MKY1021	<i>MATα (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 htz1::KAN [pRS314, HHT2-HHF2]</i>
MKY1022	<i>MATα (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 htz1::KAN [pRS314, hht2K56R-HHF2]</i>
MKY1023	<i>MATα (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 htz1::KAN [pRS314, hht2K56Q-HHF2]</i>
MKY1024	<i>MATα (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 [pRS314, HHT2-HHF2]</i>
MKY1025	<i>MATα (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 [pRS314, hht2K56R-HHF2]</i>
MKY1026	<i>MATα (hhf1-hht1)Δ::LEU2 (hhf2-hht2)Δ::HIS3 [pRS314, hht2K56Q-HHF2]</i>

Table 2.3: *YAF9* mutants.

Name	Mutations	Class
<i>yaf9-2</i>	Y112A	A
<i>yaf9-4</i>	K60E, K61E, K65E, H67A, K97E, K107E, H113A, R116E	A
<i>yaf9-6</i>	H39A, T40A, Y70A	A
<i>yaf9-7</i>	K65A, H67A, T69A, Y70A	A
<i>yaf9-8</i>	E86A, T87A, W89A, E91A, F92A	A
<i>yaf9-9</i>	L117A, H118A	A
<i>yaf9-10</i>	W89D	A
<i>yaf9-11</i>	F92D	A
<i>yaf9-12</i>	W89D, F92D	A
<i>yaf9-13</i>	K60E, K61E, K65E, H67A	A
<i>yaf9-14</i>	K60E, K61E, K65E, H67A, K97E, K107E	A
<i>yaf9-15</i>	I9D	A
<i>yaf9-16</i>	2-10Δ, T11M	A

Table 2.3 – continued

Name	Mutations	Class
<i>yaf9-24</i>	Y112E	A
<i>yaf9-1</i>	H39A, T40A, H41A, H67A, Y70A, W89A, F92A, L117A, H118A	B
<i>yaf9-3</i>	W89K E91K D93K	B
<i>yaf9-23</i>	H39W, F46D, D54L, T85D, F92W	B
<i>yaf9-26</i>	H39W	B
<i>yaf9-27</i>	H39W, F46D	B
<i>yaf9-28</i>	H39W, F46D, T85D, F92W	B
<i>yaf9-30</i>	F92W	B
<i>yaf9-31</i>	D54L	B
<i>yaf9-32</i>	T85D	B
<i>yaf9-33</i>	F46D	B
<i>yaf9-34</i>	W89K E91K D93K I94M	B
<i>yaf9-17</i>	L109D	C
<i>yaf9-18</i>	V47M, V98I	C
<i>yaf9-19</i>	I18W	C
<i>yaf9-20</i>	I153W	C
<i>yaf9-21</i>	F111W	C
<i>yaf9-22</i>	V47M, V98I, F111W	C
<i>yaf9-25</i>	V98I	C
<i>yaf9-29</i>	V47M, V70H, V98I, F111W	C

Table 2.4: *yaf9* mutant phenotypes. +++, growth comparable with wild type; –, indicates significant growth defect; + and ++, intermediate phenotypes. *Phenotypes on both 37 °C and hydroxyurea were included as there was no difference between the two conditions.

Allele	Class	Growth under the following conditions			
		Formamide	Hydroxyurea	Benomyl	<i>yaf9Δasf1Δ</i> *
YAF9	wild type	+++	+++	+++	+++
<i>yaf9Δ</i>	deletion	–	+	+	+
<i>yaf9-2</i>	A	+++	+++	+++	+++
<i>yaf9-4</i>	A	–	+	+	+

Table 2.4 – continued

Allele	Class	Growth under the following conditions			
		Formamide	Hydroxyurea	Benomyl	<i>yaf9Δasf1Δ</i> *
<i>yaf9-6</i>	A	+++	+++	+++	+++
<i>yaf9-7</i>	A	+++	+++	+++	+++
<i>yaf9-8</i>	A	+++	+++	+++	+++
<i>yaf9-9</i>	A	+++	+++	+++	+++
<i>yaf9-10</i>	A	+++	+++	+++	+++
<i>yaf9-11</i>	A	+++	+++	+++	+++
<i>yaf9-12</i>	A	+++	+++	+++	+++
<i>yaf9-13</i>	A	+++	+++	+++	+++
<i>yaf9-14</i>	A	+++	+++	+++	+++
<i>yaf9-15</i>	A	+++	+++	+++	+++
<i>yaf9-16</i>	A	+++	+++	+++	+++
<i>yaf9-24</i>	A	+++	+++	+++	+++
<i>yaf9-1</i>	B	–	+	++	++
<i>yaf9-3</i>	B	+	+++	+++	++
<i>yaf9-34</i>	B	+	+++	+++	++
<i>yaf9-23</i>	B	–	+	+	–
<i>yaf9-26</i>	B	+++	+++	+++	+++
<i>yaf9-27</i>	B	–	+	+	–
<i>yaf9-28</i>	B	–	+	+	–
<i>yaf9-30</i>	B	+++	+++	+++	+++
<i>yaf9-31</i>	B	+++	+++	+++	+++
<i>yaf9-32</i>	B	+++	+++	+++	+++
<i>yaf9-33</i>	B	+++	+++	+++	+++
<i>yaf9-17</i>	C	+++	+++	+++	+++
<i>yaf9-18</i>	C	+++	+++	+++	+++
<i>yaf9-19</i>	C	+++	+++	+++	+++
<i>yaf9-20</i>	C	+++	+++	+++	+++
<i>yaf9-21</i>	C	+++	+++	+++	+++
<i>yaf9-22</i>	C	+++	+++	+++	+++

Table 2.4 – continued

Allele	Class	Growth under the following conditions			
		Formamide	Hydroxyurea	Benomyl	<i>yaf9Δasf1Δ</i> *
<i>yaf9-25</i>	C	+++	+++	+++	+++
<i>yaf9-29</i>	C	+++	+++	+++	+++

2.4.2 Analytical-scale affinity purifications

Coprecipitation assays were performed as described previously (Kobor et al., 2004). Briefly, yeast cells were lysed in TAP-IP Buffer (50 mmol/l Tris [pH 7.8], 150 mmol/l NaCl, 1.5 mmol/l MgAc, 0.15 % NP-40, 1 mmol/l DTT, 10 mmol/l NaPPI, 5 mmol/l EGTA, 5 mmol/l EDTA, 0.1 mmol/l Na₃VO₄, 5 mmol/l NaF, CompleteTM Protease inhibitor cocktail) using acid-washed glass beads and mechanically disrupting using a bead beater (BioSpec Products, Bartlesville, Oklahoma, United States). TAP-tagged fusion proteins were captured using IgG sepharose beads (Amersham Biosciences), and subsequently washed in TAP-IP buffer. Captured material was analyzed by immunoblotting and co-purifying proteins were detected with anti-HA (Applied Biological Materials), anti-FLAG M2 (Sigma), Tra1 and Eaf1 antibodies (generous gifts from J. Workman and J. Coté, respectively). Bands were visualized using the Odyssey Infrared Imaging System (Licor).

2.4.3 Protein expression and purification

An N-terminally hexa-histidine tagged Yaf9 construct spanning amino acids 8–171 was cloned into a pET28b derivative and expressed in *E. coli* BL21 (DE3)-RIL cells by IPTG induction (1 mmol/l for 6 h at 37 °C, added after cells reached an A₆₀₀ of 0.5).

Selenomethionine-labeled protein was made similarly, with additions to the previously described protocol (Van Duyne et al., 1993). Cells were harvested by centrifugation, resuspended in lysis buffer containing 25 mmol/l Tris-Cl [pH 7.5], 0.5 mol/l KCl, 10 % glycerol, 10 mmol/l imidazole, 5 mmol/l β-mercaptoethanol and 1 mg/ml lysozyme. Cells were lysed by sonication and insoluble material

removed by centrifugation. The supernatant was passed over a Nickel column (Pharmacia AKTA FPLC) and eluted using a gradient of 10–0.5 mol/l imidazole. Fractions containing Yaf9 were identified by SDS-PAGE, concentrated to 1 ml and treated with TEV protease overnight at 4 °C to cleave the His-tag. The TEV-cleaved protein was passed over a Ni column to remove the His-tag and the flow-through, thereby concentrating the sample to 1 ml. The protein sample was purified on a size exclusion column (Sephacryl S200) equilibrated with 50 mmol/l Tris (pH 7.5), 10 % glycerol, 0.5 mol/l KCl and 1 mmol/l TCEP. Fractions containing pure, monodisperse protein (as indicated by SDS-PAGE and dynamic light scattering) were concentrated to 10 mg/ml using ultrafiltration (Centriprep-10, Millipore).

2.4.4 Crystallization and structure determination

The purified Yaf9 YEATS domain was crystallized by microbatch under paraffin oil. The stock protein solution was dialyzed in 10 mmol/l Tris, 100 mmol/l KCl and 1 mmol/l TCEP, pH 7.5 at 4 °C prior to crystallization. One volume of protein was mixed with one volume of crystallization solution, which contained 10 % PEG 3350, 100 mmol/l Na-Tartrate and 20 % glycerol. Crystals appeared in 3–5 days and grew to average dimensions (100×100×20 µm) in a week. For cryoprotection, crystals were introduced into 15 % PEG 3350, 100 mmol/l Na-Tartrate and 25 % glycerol, pH 7.5 slowly, in five successive steps and then flash frozen in liquid nitrogen.

All data were collected at beam line 8.3.1 at the advanced light source (MacDowell et al., 2004). Native data were collected to 2.3 Å resolution. Initial phases were calculated to 2.9 Å resolution using a two-wavelength MAD dataset from selenomethionine-substituted crystals, and the phases extended to 2.3 Å resolution in DM (Cowtan, 1994). The model was built in O (Jones T. A., 1997) and the structure was refined to a final $R_{work} = 21.8$ and $R_{free} = 25.8$ using REFMAC5 (Murshudov et al., 1997) with TLS refinement in the last steps. Superpositions of Yaf9 with Asf1 were performed using LSQMAN (Kleywegt, 1999), and sequence alignments using MAFFT (Katoh et al., 2005) and JALVIEW (Waterhouse et al., 2009).

2.4.5 GST-fusion protein purifications and histone binding assays

GST and GST-*YAF9* plasmids for protein over-expression in yeast were kindly provided by Robert Slany (Bittner et al., 2004). GST-fusion proteins were expressed and purified in yeast (YEXpress system, Clontech) according to the instructions of the manufacturer. Cultures were grown in SC-URA with 0.5 mmol/l copper sulfate for 4 h to induce protein expression. Cell pellets were lysed in Extraction Buffer (EB: 50 mmol/l Tris-Cl [pH 7.5], 1 mmol/l EDTA, 4 mmol/l MgCl₂, 5 mmol/l DTT, 10 % glycerol, 150 mmol/l NaCl, and protease inhibitor cocktail) using acid washed glass beads and mechanically disrupted using a bead beater. Whole cell extracts were incubated with glutathione sepharose beads (GE Healthcare), the beads were washed with Wash Buffer (50 mmol/l Tris-Cl [pH 7.5], 4 mmol/l MgCl₂, 1 mmol/l DTT, 10 % glycerol, 150 mmol/l NaCl) and the purified GST-fusion protein was eluted with 7 mg/ml glutathione (Sigma) in EB. The purified protein was dialyzed in GST Binding Buffer (50 mmol/l HEPES [pH 7.5], 100 mmol/l potassium acetate, 20 mmol/l magnesium acetate, 5 mmol/l EGTA, 1 μmol/l DTT, 10 % glycerol, 0.5 % NP-40, 1 mmol/l PMSF, 0.03 % Triton X-100 and protease inhibitor cocktail).

The histone binding assay was performed essentially as previously described with minor modifications (Matangkasombut and Buratowski, 2003). Briefly, 10 μg of purified GST-fusion proteins were incubated with 20 μl of 0.5 mg/ml calf thymus histones (Sigma) in GST Binding Buffer (50 mmol/l HEPES [pH 7.5], 100 mmol/l potassium acetate, 20 mmol/l magnesium acetate, 5 mmol/l EGTA, 1 μmol/l DTT, 10 % glycerol, 0.5 % NP-40, 1 mmol/l PMSF, 0.03 % Triton X-100 and protease inhibitor cocktail) for 45 min at 37 °C, then subsequently loaded onto glutathione-agarose (Sigma) and incubated overnight at 4 °C. The supernatant was saved for input and the beads were washed in GST Binding Buffer with increasing NaCl concentration (i. e. 0.2 mol/l, 0.4 mol/l, 0.8 mol/l, 1 mol/l NaCl). The input and material bound to the beads were subjected to SDS-PAGE and immunoblotted using anti-GST (Applied Biological Materials), anti-H3 (Abcam), anti-H4 (Abcam), and anti-H2B (Upstate) antibodies. Immunoblots were analyzed using the Odyssey Infrared Imaging System (Licor).

2.4.6 Chromatin association and histone acetylation assays

The chromatin association assay was adapted from a protocol previously described with minor modifications (Liang and Stillman, 1997). Yeast cells were incubated in Pre-Spheroblast Buffer (100 mmol/l PIPES/KOH [pH 9.4], 10 mmol/l DTT, 0.1 % sodium azide) for 10 min at room temperature, and spheroblasted with 20 mg/ml Zymolyase-100T (Seikagaku Corporation) in Spheroblast Buffer (50 mmol/l KPO₄ [pH 7.5], 0.6 mol/l Sorbitol, 10 mmol/l DTT) at 37 °C for 15 min. Spheroblasts were washed with Wash Buffer (50 mmol/l HEPES/KOH [pH 7.5], 100 mmol/l KCl, 2.5 mmol/l MgCl₂, 0.4 mol/l sorbitol), resuspended in equal volume of EB (50 mmol/l HEPES/KOH [pH 7.5], 100 mmol/l KCl, 2.5 mmol/l MgCl₂, 1 mmol/l DTT, 1 mmol/l PMSF, and protease inhibitor cocktail), and lysed with 1 % Triton X-100. Whole cell extract (WCE) was saved, and the remaining lysate was separated into chromatin pellet (Pellet) and supernatant (SUP) fractions by centrifugation through EBSX (EB+ 0.25 % Triton X-100 and 30 % sucrose). WCE, Pellet, and SUP were subjected to SDS-PAGE and immunoblotted with anti-FLAG M2 (Sigma), anti-H2A (Upstate) and anti-Pgk (Molecular Probes) antibodies. Immunoblots were scanned with the Odyssey Infrared Imaging System (Licor). The average intensities of bands were deduced using the Odyssey V3.0 software.

H2A.Z-Flag was normalized against H2A in chromatin and Pgk1 in supernatant. The enrichment ratios were calculated by dividing the normalized intensity by the total intensity (sum of chromatin and supernatant) for quantification (see Figure 5B). This ratio was averaged for four independent assays.

To determine the relative amounts of H2A.Z K14 acetylation, the chromatin fractions were subjected to SDS-PAGE, immunoblotted with anti-H2A.Z K14ac (Upstate), anti-FLAG M2 (Sigma), and anti-H4 (Abcam) antibodies, and scanned with the Odyssey Infrared Imaging System (Licor).

2.4.7 Genome-wide ChIP-on-chip

ChIP was performed as described previously, using the adapted linear amplification method that involves two rounds of T7 RNA polymerase amplification (van Bakel et al., 2008). In brief, yeast cells (500 ml) were grown in SC-TRP media

to an OD600 of 0.8–0.9 and were crosslinked with 1 % formaldehyde for 20 min before chromatin was extracted. The chromatin was sonicated (Bioruptor, Diagenode; Sparta, NJ: 10 cycles, 30 s on/off, high setting) to yield an average DNA fragment of 500 bp. FLAG antibody (Sigma, F3165) for H2A.Z pull-down (4 μ l) was coupled to 60 μ l of protein A magnetic beads (Invitrogen). After reversal of the crosslinking and DNA purification, the immunoprecipitated and input DNA were amplified to about 6 μ g aRNA using T7 RNA polymerase in two rounds. Samples were labeled with biotin, and the immunoprecipitated and input samples were hybridized to two Affymetrix 1.0R *S. cerevisiae* microarrays, which are comprised of over 3.2 million probes covering the complete genome. Probes (25-mer) are tiled at an average of 5 bp resolution, creating an overlap of approximately 20 bp between adjacent probes.

2.4.8 Data analysis

We used the *Model-based Analysis of Tiling-arrays* (MAT) algorithm to reliably detect enriched regions (Johnson et al., 2006; Droit et al., 2010). MAT was applied to corresponding immunoprecipitated and input sample arrays, and the probe behaviour model was estimated by examining the signal intensity, sequence, and copy number of all probes on an array. After probe behaviour model fitting, the residuals between the model and observation were normally distributed and centred at 0. MAT uses a score function to identify regions of ChIP enrichment, which allows robust *p*-value and false-discovery-rate calculations. MAT scores were calculated for all probes using a 300 bp sliding window.

Annotations for ORFs and ARS were derived from the SGD database. An ORF was termed enriched if at least 50 % of all probes had a MAT score above a threshold of 1.5. Promoters were defined as enriched if 50 % of all probes 500 bp upstream of the transcriptional start site were above the MAT score cutoff. Promoters which overlap with ORFs of other genes were not considered. To compare our data sets with H3K56ac ChIP-on-chip data from previously published studies, we used the supplemental data available from Rufiange et al. (Rufiange et al., 2007). In order to explore this statistical significance between the H2A.Z and H3K56ac ChIP-on-chip data, we used the hypergeometric test as described (Tavazoie et al., 1999).

Bibliography

- Antczak, A., Tsubota, T., Kaufman, P. and Berger, J. (2006). Structure of the yeast histone H3-ASF1 interaction: implications for chaperone mechanism, species-specific interactions, and epigenetics. *BMC Struct Biol* 6, 26. → pages 55
- Ausubel, F. (1987). *Current protocols in molecular biology*. → pages 71
- Babiarz, J., Halley, J. and Rine, J. (2006). Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in *Saccharomyces cerevisiae*. *Genes Dev* 20, 700–10. → pages 49, 67, 69
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J. and McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98, 10037–41. → pages 61
- Bittner, C., Zeisig, D., Zeisig, B. and Slany, R. (2004). Direct physical and functional interaction of the NuA4 complex components Yaf9p and Swc4p. *Eukaryot Cell* 3, 976–83. → pages 81
- Bork, P., Holm, L. and Sander, C. (1994). The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol* 242, 309–20. → pages 53
- Cowtan, K. (1994). 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31, p34–38. → pages 80
- Daganzo, S. M., Erzberger, J. P., Lam, W. M., Skordalakes, E., Zhang, R., Franco, A. A., Brill, S. J., Adams, P. D., Berger, J. M. and Kaufman, P. D. (2003). Structure and function of the conserved core of histone deposition protein Asf1. *Curr Biol* 13, 2148–58. → pages 58
- Daser, A. and Rabbitts, T. (2004). Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis. *Genes Dev* 18, 965–74. → pages 49
- DeLano, W. (2002). The PyMOL Molecular Graphics System. <http://www.pymol.org>. → pages 56

- Droit, A., Cheung, C. and Gottardo, R. (2010). rMAT—an R/Bioconductor package for analyzing ChIP-chip experiments. *Bioinformatics* 26, 678–9. → pages 83
- English, C., Adkins, M., Carson, J., Churchill, M. and Tyler, J. (2006). Structural basis for the histone chaperone activity of Asf1. *Cell* 127, 495–508. → pages 55
- Fischer, U., Heckel, D., Michel, A., Janka, M., Hulsebos, T. and Meese, E. (1997). Cloning of a novel transcription factor-like gene amplified in human glioma including astrocytoma grade I. *Hum Mol Genet* 6, 1817–22. → pages 49
- Gelbart, M., Rechsteiner, T., Richmond, T. and Tsukiyama, T. (2001). Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Mol Cell Biol* 21, 2098–106. → pages 71
- Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E. and Bental, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* 19, 163–4. → pages 61
- Guillemette, B. and Gaudreau, L. (2006). Reuniting the contrasting functions of H2A.Z. *Biochem Cell Biol* 84, 528–35. → pages 48, 67
- Holm, L. and Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem Sci* 20, 478–80. → pages 53
- Johnson, W., Li, W., Meyer, C., Gottardo, R., Carroll, J., Brown, M. and Liu, X. (2006). Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci U S A* 103, 12457–62. → pages 83
- Jones T. A., K. M. (1997). O – the manual, Version 5.1. Uppsala University Uppsala, Sweden. → pages 80
- Katoh, K., ichi Kuma, K., Toh, H. and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33, 511–8. → pages 80

- Keogh, M., Mennella, T., Sawa, C., Berthelet, S., Krogan, N., Wolek, A., Podolny, V., Carpenter, L., Greenblatt, J., Baetz, K. and Buratowski, S. (2006). The *Saccharomyces cerevisiae* histone H2A variant Htz1 is acetylated by NuA4. *Genes Dev* 20, 660–5. → pages 49, 67, 69
- Kleywegt, G. J. (1999). Experimental assessment of differences between related protein crystal structures. *Acta Crystallogr D Biol Crystallogr* 55, 1878–84. → pages 80
- Kobor, M., Venkatasubrahmanyam, S., Meneghini, M., Gin, J., Jennings, J., Link, A., Madhani, H. and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* 2, E131. → pages 49, 79
- Krogan, N., Keogh, M., Datta, N., Sawa, C., Ryan, O., Ding, H., Haw, R., Pootoolal, J., Tong, A., Canadien, V., Richards, D., Wu, X., Emili, A., Hughes, T., Buratowski, S. and Greenblatt, J. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell* 12, 1565–76. → pages 49
- Kusch, T. and Workman, J. (2007). Histone variants and complexes involved in their exchange. *Subcell Biochem* 41, 91–109. → pages 48
- Kushnirov, V. (2000). Rapid and reliable protein extraction from yeast. *Yeast* 16, 857–60. → pages 72
- Lauffart, B., Howell, S., Tasch, J., Cowell, J. and Still, I. (2002). Interaction of the transforming acidic coiled-coil 1 (TACC1) protein with ch-TOG and GAS41/NuBI1 suggests multiple TACC1-containing protein complexes in human cells. *Biochem J* 363, 195–200. → pages 72
- Li, B., Pattenden, S., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J. and Workman, J. (2005). Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. *Proc Natl Acad Sci U S A* 102, 18385–90. → pages 67

- Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A. and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* *134*, 244–55. → pages 70
- Liang, C. and Stillman, B. (1997). Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in *cdc6* mutants. *Genes Dev* *11*, 3375–86. → pages 82
- Longtine, M., McKenzie, A., Demarini, D., Shah, N., Wach, A., Brachat, A., Philippsen, P. and Pringle, J. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* *14*, 953–61. → pages 71
- MacDowell, A. A., Celestre, R. S., Howells, M., McKinney, W., Krupnick, J., Cambie, D., Domning, E. E., Duarte, R. M., Kelez, N., Plate, D. W., Cork, C. W., Earnest, T. N., Dickert, J., Meigs, G., Ralston, C., Holton, J. M., Alber, T., Berger, J. M., Agard, D. A. and Padmore, H. A. (2004). Suite of three protein crystallography beamlines with single superconducting bend magnet as the source. *J Synchrotron Radiat* *11*, 447–55. → pages 80
- Masson, I. L., Yu, D., Jensen, K., Chevalier, A., Courbeyrette, R., Boulard, Y., Smith, M. and Mann, C. (2003). Yaf9, a novel NuA4 histone acetyltransferase subunit, is required for the cellular response to spindle stress in yeast. *Mol Cell Biol* *23*, 6086–102. → pages 49
- Matangkasombut, O. and Buratowski, S. (2003). Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol Cell* *11*, 353–63. → pages 81
- Millar, C., Xu, F., Zhang, K. and Grunstein, M. (2006). Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev* *20*, 711–22. → pages 49, 67
- Mizuguchi, G., Shen, X., Landry, J., Wu, W., Sen, S. and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* *303*, 343–8. → pages 49, 70

- Mullem, V. V., Wery, M., Bolle, X. D. and Vandenhaute, J. (2003). Construction of a set of *Saccharomyces cerevisiae* vectors designed for recombinational cloning. *Yeast* 20, 739–46. → pages 72
- Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53, 240–55. → pages 80
- Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M. and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* 446, 338–41. → pages 55
- Park, J. and Roeder, R. (2006). GAS41 is required for repression of the p53 tumor suppressor pathway during normal cellular proliferation. *Mol Cell Biol* 26, 4006–16. → pages 49, 68
- Recht, J., Tsubota, T., Tanny, J., Diaz, R., Berger, J., Zhang, X., Garcia, B., Shabanowitz, J., Burlingame, A., Hunt, D., Kaufman, P. and Allis, C. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc Natl Acad Sci U S A* 103, 6988–93. → pages 57
- Rufiange, A., Jacques, P., Bhat, W., Robert, F. and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol Cell* 27, 393–405. → pages 67, 83
- Schulze, J. M., Wang, A. Y. and Kobor, M. S. (2009). YEATS domain proteins: a diverse family with many links to chromatin modification and transcription. *Biochem Cell Biol* 87, 65–75. → pages 49, 50, 53
- Tavazoie, S., Hughes, J., Campbell, M., Cho, R. and Church, G. (1999). Systematic determination of genetic network architecture. *Nat Genet* 22, 281–5. → pages 83
- van Bakel, H., van Werven, F., Radonjic, M., Brok, M., van Leenen, D., Holstege, F. and Timmers, H. (2008). Improved genome-wide localization by ChIP-chip

using double-round T7 RNA polymerase-based amplification. *Nucleic Acids Res* 36, e21. → pages 82

Van Duyne, G., Standaert, R., Karplus, P., Schreiber, S. and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* 229, 105–24. → pages 79

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–91. → pages 80

Williams, S., Truong, D. and Tyler, J. (2008). Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc Natl Acad Sci U S A* 105, 9000–5. → pages 70

Wu, W., Alami, S., Luk, E., Wu, C., Sen, S., Mizuguchi, G., Wei, D. and Wu, C. (2005). Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nat Struct Mol Biol* 12, 1064–71. → pages 49

Wu, W.-H., Wu, C.-H., Ladurner, A., Mizuguchi, G., Wei, D., Xiao, H., Luk, E., Ranjan, A. and Wu, C. (2008). N-terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex. *J. Biol. Chem.* 284, M808830200. → pages 69

Zeisig, D., Bittner, C., Zeisig, B., Garcia-Cuellar, M., Hess, J. and Slany, R. (2005). The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. *Oncogene* 24, 5525–32. → pages 71

Zhang, H., Richardson, D., Roberts, D., Utley, R., Erdjument-Bromage, H., Tempst, P., Cote, J. and Cairns, B. (2004). The Yaf9 component of the SWR1 and NuA4 complexes is required for proper gene expression, histone H4 acetylation, and Htz1 replacement near telomeres. *Mol Cell Biol* 24, 9424–36. → pages 49, 68

Chapter 3

Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation*

3.1 Introduction

In eukaryotic cells, genomic DNA is packaged by the histone proteins, forming the fundamental repeating unit of chromatin, the nucleosome (Luger et al., 1997). Several residues within the histone tails and some within the histone core can be altered by posttranslational modifications, including acetylation, phosphorylation, ubiquitination, and methylation (Shilatifard, 2006).

Two classes of histone lysine methylase have been identified to date. The first class

*A version of this chapter has been published. Schulze JM, Jackson J, Nakanishi S, Gardner JM, Hentrich T, Haug J, Johnston M, Jaspersen SL, Kobor MS, Shilatifard A. (2009). Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation. *Mol Cell.* 35, 626-41.

See Co-authorship Statement on page xv for details about my contributions.

includes the SET-domain-containing histone methylases capable of methylating many histones on different residues with varying biological outcomes (Shilatifard, 2006). The second class, the non-SET domain methylases, consists of only one member, Dot1 (disrupter of telomeric silencing 1), which is capable of methylating histone H3 on lysine 79 (H3K79). H3K79 is one of the conserved core residues located in a loop within the histone-fold domain that can be methylated (Lu et al., 2008). Originally identified as a gene affecting the silencing of gene expression near telomeres in *S. cerevisiae*, Dot1 is an evolutionally conserved enzyme that catalyzes mono-, di-, and trimethylation of H3K79 (me1, me2, and me3, respectively) (Lacoste et al., 2002; Ng et al., 2002; van Leeuwen et al., 2002). We and others have demonstrated that, similar to H3K4 methylation, efficient trimethylation of H3K79 by Dot1 requires monoubiquitination of K123 of histone H2B, catalyzed by the Rad6/Bre1 protein complex (Dover et al., 2002; Krogan et al., 2003; Shilatifard, 2006; Sun and Allis, 2002; Wood et al., 2003).

Dot1 methylates H3K79 only in the context of nucleosomes and can be regulated by several sites within other histones (Altaf et al., 2007; Onishi et al., 2007).

The mammalian Dot1 has been directly implicated in leukemogenesis through the misregulation of *HOX* genes (Okada et al., 2005). Furthermore, it was shown that an MLL-AF4 fusion causes ectopic recruitment of Dot1 and the aberrant methylation of H3K79 at MLL target genes (Krivtsov et al., 2008). These studies point the way to the clinical significance of H3K79 methylation and to the pathogenesis of human cancer. We know very little about the functions of the different methylation states of H3K79, and the results of different studies are contradictory. H3K79me2 is suggested to be a mark of transcriptionally active genes in *Drosophila* (Schubeler et al., 2004) and in mammalian cells (Im et al., 2003; Martin and Zhang, 2005; Miao and Natarajan, 2005). However, a different study found that H3K79me1 and H3K79me2 did not have a significant preference toward either active or silent genes (Barski et al., 2007). Furthermore, H3K79me3 did not have a correlation with either active or silent genes in yeast (Pokholok et al., 2005), but is associated with transcriptionally repressed genes in human cells (Barski et al., 2007).

One of the major mechanisms driving cell-cycle regulation is the orchestrated transcriptional program of gene expression. Genome-wide studies of cell cycle-

regulated transcripts in budding yeast indicate that more than 800 genes are periodically expressed (Spellman et al., 1998). More than 300 of those, including the G1 cyclins, *CLN1* and *CLN2*; the S phase cyclins, *CLB5* and *CLB6*; and many genes involved in DNA synthesis show peak expression in G1/S. Two heterodimeric transcription factor complexes are required primarily for activation of gene expression at the G1/S phase transition of the cell cycle: SBF (SCB-binding factor) and MBF (MCB-binding factor) (Breedon, 1996; Harbison et al., 2004; Iyer et al., 2001; Simon et al., 2001). While the SBF complex is composed of Swi4 and Swi6, which bind a repeated upstream regulatory sequence SCB (Swi4,6-dependent cell cycle box), MBF consists of Mbp1 and Swi6, binding to another regulatory sequence MCB (Mlu1-dependent cell-cycle box) in a broad range of targets. Swi4 and Mbp1 are the specific DNA-binding factors, whereas the shared subunit Swi6 has no DNA binding activity but acts as a regulator (Andrews and Herskowitz, 1989; Breedon and Nasmyth, 1987; Koch et al., 1993). Gene activation by SBF and MBF is subject to a variety of regulatory mechanisms, including phosphorylation of Swi6, cell cycle-regulated *SWI4* expression, and control of DNA binding. Despite the detailed knowledge of transcription factors involved in cell-cycle control in budding yeast, fundamental questions about the role of chromatin structure in this process have not been answered so far. One exception involves the tightly timed transcription of canonical histone genes, which is restricted to S phase in *S. cerevisiae* (Hereford et al., 1981).

To define the role of di- and trimethylation of H3K79, we screened the entire yeast gene deletion collection to identify proteins required for the establishment of different forms of H3K79 methylation by Dot1. We found that Swi4 and Swi6 were required for dimethylation but not for trimethylation of H3K79, suggesting a link between cell-cycle progression and chromatin modification at this site. Consistent with this possibility, we found that levels of H3K79me2 were decreased in the G1 phase and peaked in G2/M. Furthermore, we showed that the loss of H3K79me2 in G1 phase requires Nrm1 (negative regulator of MBF targets 1) and Whi3. Nrm1 is a corepressor of MBF-regulated gene expression (de Bruin et al., 2006), and Whi3 is an RNA-binding protein involved in cell-cycle control, and its loss results in the acceleration of the expression of genes controlled by the SBF and MBF

complexes. Interestingly, high-resolution genome-wide mapping of H3K79 methylation by chromatin immunoprecipitation (ChIP) followed by microarray analysis (ChIP-on-chip) showed that H3K79 di- and trimethylation were mutually exclusive and reside in different regions of the genome. Further supporting the link of H3K79 dimethylation to the cell cycle, we found M/G1 cell cycle-regulated genes as well as Swi4/6-bound promoters to be significantly enriched for H3K79 di- but not trimethylation.

Since monoubiquitination of histone H2B is required for H3K79 trimethylation, we have generated polyclonal antibodies to monoubiquitinated H2B on K123. Employing ChIP-on-chip analysis, we demonstrated that H2B monoubiquitination was detected on sites marked by H3K79me3, but not on cell cycle-regulated genes and sites enriched for H3K79me2.

Together, our studies provide a link between Rad6/Bre1-dependent H2B monoubiquitination and H3K79me3 and its exclusion from H3K79me2 modified sites and, for the first time, establish a connection between H3K79me2, the SBF and MBF transcription complexes, and the cell cycle.

3.2 Results

3.2.1 Swi4 and Swi6 were required for normal levels of H3K79 di- but not trimethylation

To identify proteins required for establishment of H3K79me2, we surveyed extracts of the collection of nonessential gene deletion mutants of *S. cerevisiae* for this modification by protein blotting, a method we call GPS (global proteomics in *S. cerevisiae*) (Schneider et al., 2004). This biochemical screen revealed that the loss of either Swi4 or Swi6 resulted in a decrease in the level of H3K79me2 (Figure 3.1 (a)–(b), page 95), without affecting either H3K79me3 or the H3K4me3 levels (Figure 3.1 (c)–(f), page 94). This defect was specifically due to the *swi4* Δ and *swi6* Δ deletions, as both the *MAT α* and *MAT β* versions of the *swi4* Δ and *swi6* Δ mutants in the deletion collection had a reduced level of H3K79me2 (Figure 3.1 (g), page 95). Furthermore, the histone H3K79me2 defect cosegregated with the *swi4* Δ ::*kanMX* mutation in spores from a cross of this mutant to a

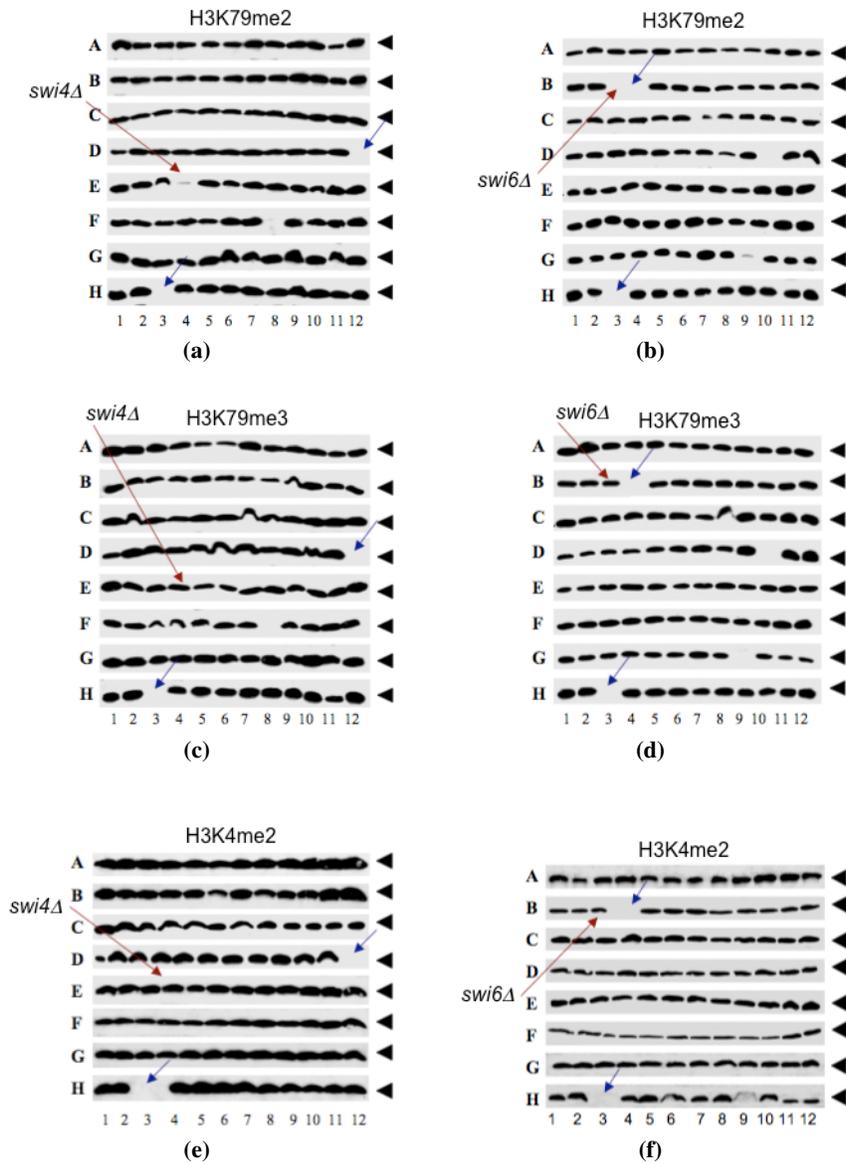


Figure 3.1: GPS reveals that *SWI4* and *SWI6* are required for H3K79 di- but not trimethylation. (a)–(d): Cell extracts from each of the nonessential yeast gene deletion mutants were subjected to SDS-PAGE, western blotted, and probed with antibodies specific for H3K79 dimethylation ((a) and (b)) or trimethylation ((c) and (d)). (e) and (f): H3K4 dimethylation was used as a control. Red arrows represent *swi4Δ* ((a), (c), and (e)) and *swi6Δ* ((b), (d), and (f)). Blue arrows indicate empty wells that serve as plate markers.

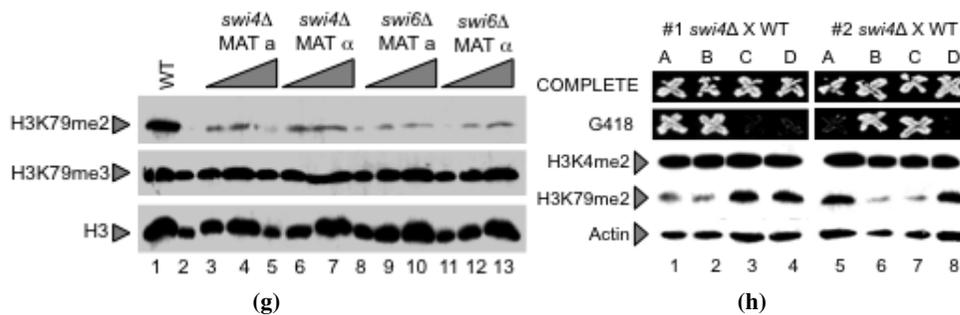


Figure 3.1: GPS reveals that *SWI4* and *SWI6* are required for H3K79 di- but not trimethylation. (g): Protein extracts from two different mating strains of *swi4Δ* and *swi6Δ* were analyzed as in (a)–(f). Anti-acetyl-H3 was probed for a loading control. (h): The *swi4::kanMX* mutant from the deletion collection was mated with a wild-type strain, and spores of the resulting tetrads were dissected and scored for the KanR and H3K79 methylation pattern.

SWI4 (wild-type) strain (Figure 3.1 (h), page 95). The defect was complemented by a plasmid carrying either *SWI4* or *SWI6* (Figure 3.2 (a) and (b), page 96).

Swi4 and Swi6 comprise the SBF complex that binds to the SCB sequence and regulates expression of several genes involved in cell-cycle progression (Harrington and Andrews, 1996). The possibility that the H3K79 dimethylation defect of the *swi4Δ* and *swi6Δ* mutants was due to reduced expression of *DOT1* was ruled out by our observation that *DOT1* expression is not affected in those mutants (Figure 3.2 (c)). Furthermore, *DOT1* does not appear to have SCB sequences in its promoter, but the *DOT1* gene is under transcriptional control with peak expression at the G1/S transition (Spellman et al., 1998). We also observed that loss of either Swi4 or Swi6 did not alter recruitment of Dot1 to chromatin, suggesting that they do not affect H3K79me2 via Dot1 recruitment, but rather alter some other aspect of dimethylation control.

3.2.2 Swi4 and Swi6 were not required for H2B monoubiquitination

Histone H2B monoubiquitination, catalyzed by the Rad6/Bre1 protein complex, is required for di- and trimethylation of H3K4 and trimethylation of H3K79 cat-

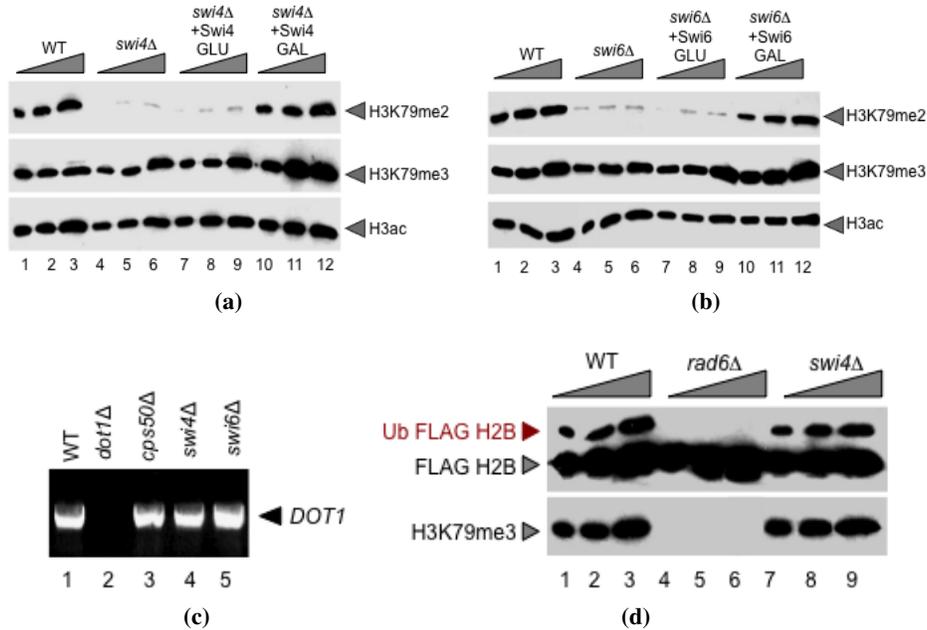


Figure 3.2: Role of *SWI4* and *SWI6* in H3K79 methylation. (a) and (b): Cell extracts were subjected to SDS-PAGE, blotted to a membrane, and probed with H3K79 dimethyl- or trimethyl-specific antibodies. Anti-acetyl-H3 was probed as a loading control. H3K79 dimethylation could be rescued by a plasmid carrying either *SWI4* (a) or *SWI6* (b) under control of the *GAL1* promoter. (c): RT-PCR was performed using cDNA made from RNA extracted from *SWI4* as well as *SWI6* deletion strains via reverse transcriptase. Primers specific for *DOT1* were used to detect its mRNA. (d): The effect of the loss of *SWI4* on histone H2B monoubiquitination levels. Highly purified acid-extracted histones from wild-type strains (strains containing FLAG-tagged H2B as the only source of histone H2B) or strains deleted for *SWI4* in the same background were analyzed by western blotting using (upper panel) antibodies against the FLAG epitope or (middle and lower panels) polyclonal antibodies raised against trimethylated H3K79. As indicated by the red arrows, monoubiquitinated H2B is the slower-migrating and the ubiquitinated species of histone H2B is the faster-migrating form. *rad6Δ* cells lack the E2 ubiquitin-conjugating enzyme required for H2B ubiquitination.

alyzed by COMPASS and Dot1, respectively (Dover et al., 2002; Hwang et al., 2003; Sun and Allis, 2002; Wood et al., 2003). We entertained the possibility that the methylation defect of *swi4* Δ and *swi6* Δ mutants was due to their lack of H2B monoubiquitination. However, we found normal levels of monoubiquitinated histone H2B (tagged with the FLAG epitope) in *swi4* Δ (Figure 3.2 (d), page 96) and *swi6* Δ mutants. Therefore, Swi4 and Swi6 did not affect H3K79 methylation indirectly by regulating the level of H2B monoubiquitination.

3.2.3 Establishment of H3K79 dimethylation was cell-cycle dependent

Due to the important role of Swi4 and Swi6 cell-cycle progression, we next tested whether H3K79 methylation fluctuated during the cell cycle. Flow cytometric analysis of DNA content in *dot1* Δ mutants suggested that H3K79 methylation is important for normal entry into S phase. Compared to wild-type cells, *dot1* Δ mutants displayed a statistically significant increase in the percentage of G1 cells (Figure 3.3 (a), page 98).

To examine whether H3K79me₂ is cell-cycle regulated, cells were synchronized in G1 using the mating pheromone α factor, and cell-cycle progression as well as H3K79me₂ levels were monitored (Figure 3.3 (b) and (c), page 98). Levels of H3K79 dimethylation were low in G1-arrested cells, but increased as cells entered the S phase 15–30 min following release from α factor, as measured by flow cytometric analysis of DNA (Figure 3.3 (b) and (c), page 98). H3K79 dimethylation peaked at G2/M around 60 min and decreased at 105 min at the completion of mitosis (Figure 3.3 (b) and (c), page 98). In contrast, the H3K79me₃ mark, as well as the acetylation of H3, was present throughout the cell cycle with no obvious variation in signal intensity (Figure 3.3 (c), page 98).

To exclude the possibility that cell cycle-dependent changes in H3K79 dimethylation were the result of a global turnover of histones during DNA replication, we used a strain containing the replication initiation factor *CDC6* under the regulatable *GALI* promoter (Figure 3.3 (d)–(f), page 99).

In glucose, *CDC6* expression is repressed and DNA replication does not occur, yet other cell-cycle events, such as Cdk activation/inactivation, securin degradation,

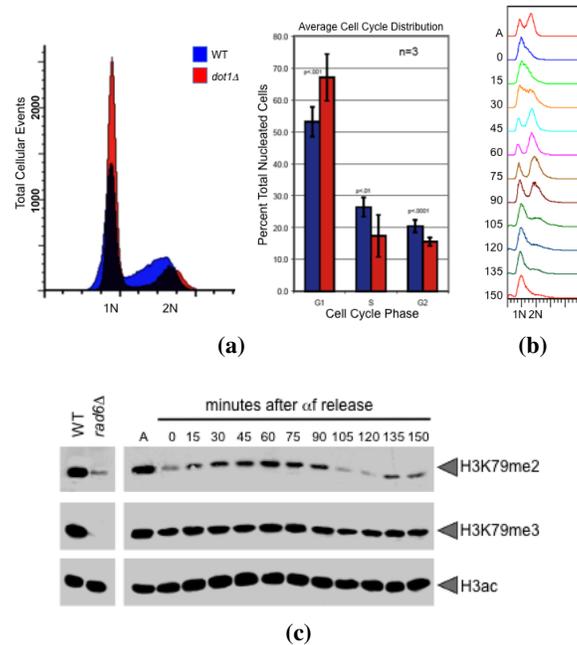


Figure 3.3: Cell-cycle regulation of H3K79 dimethylation. (a): Cells lacking *DOT1* accumulate in G1. Wild-type and *dot1Δ* cells were grown to mid-log phase at 30 °C in YPD, and DNA content was analyzed by flow cytometry. The percentage of cells in G1 with 1N DNA content, in S phase with intermediate DNA content, and in G2/M with 2N DNA content was quantitated using three independent samples. Error bars represent SD of the mean. (b) and (c): H3K79 dimethylation increases during S phase. Wild-type cells (SLJ001) were arrested in G1 with a factor at 30 °C for 3 h. They were released into fresh YPD at 30 °C, and a factor was added back when small buds appeared to prevent cells from entering the next cell cycle. Samples were taken at the indicated times following release from a factor; asynchronous cells (a) were also collected. DNA content was analyzed by flow cytometry to estimate cell-cycle position (b). Entry into S phase begins around 15–30 min based on the shift of the 1N DNA peak, while cells have entered G2/M approximately 45–60 min after a factor release. By 90 min, cells have exited M phase and rearrested in G1. Cell cycle-dependent modification of H3 was analyzed by immunoblotting with anti-acetyl-H3 and anti-H3K79me2 and anti-H3K79me3 antibodies (c). Equal protein concentrations were loaded in all lanes as judged by total H3 levels.

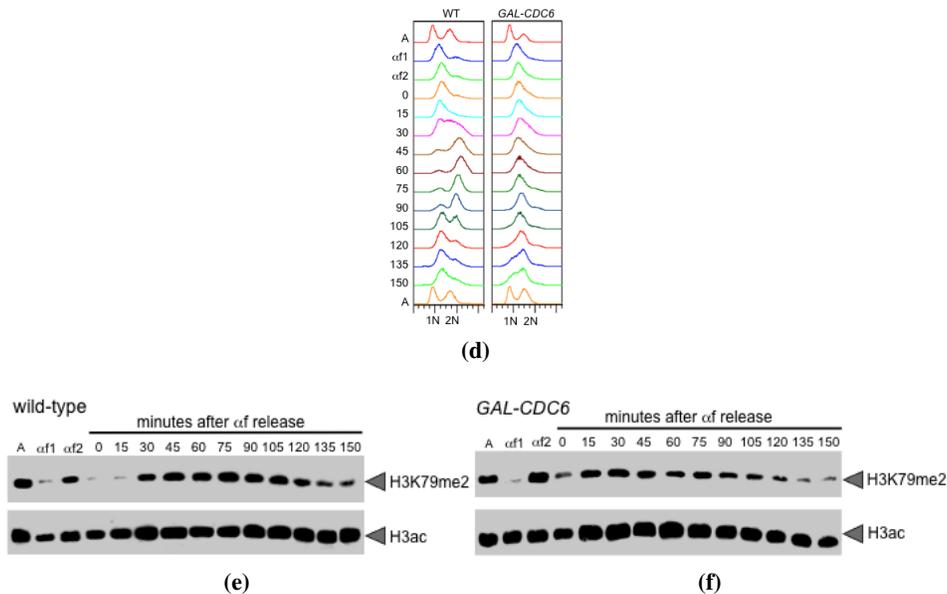


Figure 3.3: Cell-cycle regulation of H3K79 dimethylation. (d)–(f): DNA replication is not required for cell-cycle oscillations in H3K79 dimethylation. Wild-type and *GAL-CDC6* cells were grown in YEP containing 2% raffinose and 4% galactose at 30 °C, then arrested in G1 with α factor (α f1). Cells were released from G1 into YEP containing 2% raffinose and 4% galactose to allow *CDC6* expression to initiate DNA replication. After 20 min, cells were transferred into YPD to repress *GAL-CDC6*, and α factor was added at 45 min to arrest cells in G1 (α f2). Cells were then released from G1 in glucose-containing media to analyze cell-cycle progression in cells lacking Cdc6 protein. α factor was added back when small buds appeared to prevent cells from entering the next cell cycle. Samples were taken at the indicated times following release from the second α factor release; asynchronous cells (a) were also collected. DNA content was analyzed by flow cytometry (d). While wild-type cells replicate DNA (30 min), progress through mitosis, and rearrest in G1, similar to our results in (b), cells lacking Cdc6 protein do not undergo DNA replication and form a 2N DNA peak. Peak drift in this sample is likely due to mitochondrial DNA, since cells continue to increase in size. Cell cycle-dependent modification of H3 was analyzed by immunoblotting with anti-H3K79me2 and anti-acetyl-H3 antibodies in wild-type (e) and *GAL-CDC6* (f) cells. Equal protein concentrations were loaded in all lanes as judged by total H3 levels.

and spindle formation, still occur (Biggins and Murray, 2001; Stern and Murray, 2001). Lack of tension at kinetochores activates the spindle checkpoint in budding yeast (Biggins and Murray, 2001). The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. Cdc6 is an unstable protein whose *de novo* synthesis in G1 is important for the onset of the S phase and for preventing a ‘reductional’ anaphase in the budding yeast *S. cerevisiae* (Piatto et al., 1995).

H3K79 dimethylation also oscillates during the cell cycle in *GAL-CDC6* cells that fail to replicate DNA just as it does in wild-type cells that undergo S phase (Figure 3.3 (d) and (e), page 99).

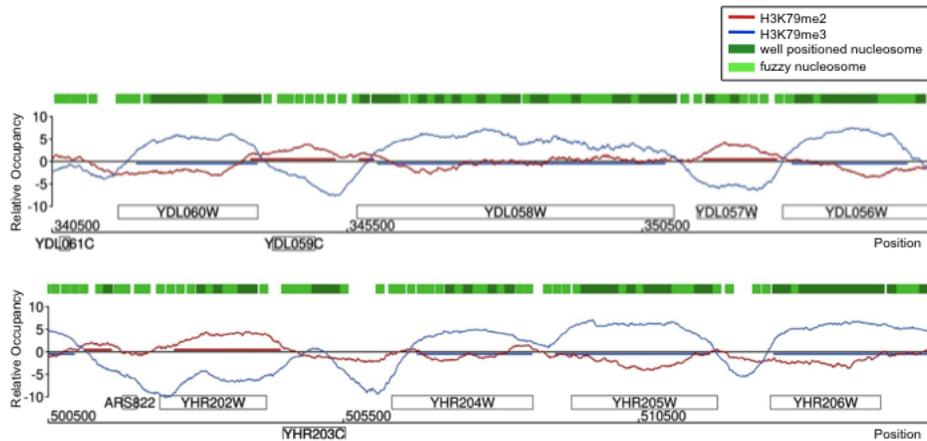
Thus, cell-cycle changes in H3K79me2 are likely regulated at the level of its establishment and/or removal and may be important for progression through the cell cycle.

3.2.4 Genome-wide localizations of H3K79me2 and H3K79me3 were mutually exclusive

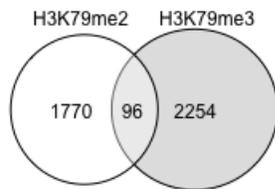
Little is known about the distinct functions of H3K79me2 and H3K79me3 marks in the cell, and a redundancy in their roles was proposed in budding yeast (Fredriks et al., 2008; Shahbazian et al., 2005). To better understand the relationship of H3K79 di- and trimethylation as well as elucidate a possible role for H3K79me2 in cell-cycle progression, a comprehensive genome-wide map of these modifications was established via ChIP-on-chip experiments using high-resolution tiling microarrays. Protein-DNA complexes containing either di- or trimethylated forms of histone H3K79 were specifically immunoprecipitated with antibodies against H3K79me2 and H3K79me3, respectively.

In order to reliably detect enriched regions, we applied an adapted version of the Model-based Analysis of Tiling arrays algorithm (MAT) (Johnson et al., 2006; Droit et al., 2010), comparing signal intensities between ChIP and genomic DNA to calculate the protein-binding profile. Spearman rank correlation coefficients of $r = 0.9$ in average indicated high reproducibility and robustness of the performed replicates.

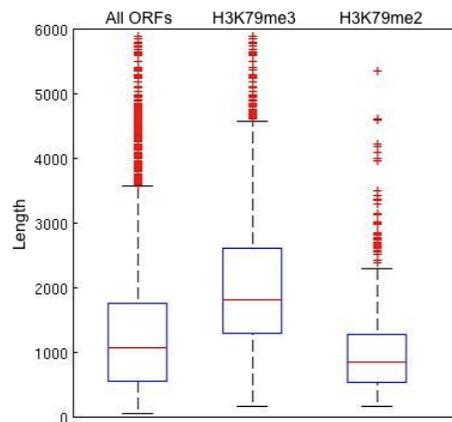
Intriguingly, we found that H3K79me2 and H3K79me3 were localized to different regions of the genome and had distinct and mutually exclusive patterns on



(a)



(b)



(c)

Figure 3.4: High-resolution profile of H3K79me2 and H3K79me3 across the yeast genome with global occupancy analysis. (a): H3K79me2 and H3K79me3 profiles. Sample genomic positions for chromosomes 4 and 8 were plotted along the x -axis against the relative occupancy of H3K79me2 and H3K79me3 on the y -axis. ORFs are indicated as rectangles above the axis for Watson genes and below the axis for Crick genes. Green boxes represent HMM-predicted well-positioned and fuzzy nucleosome positions derived from Lee et al., 2007. (b): Venn diagram comparing the number of H3K79me2- and H3K79me3-enriched ORFs. (c): Average lengths of H3K79me3- and H3K79me2-enriched genes. Boxplots show the lengths of 6576 ORFs in yeast as well as the lengths of H3K79me3- and H3K79me2-enriched ORFs.

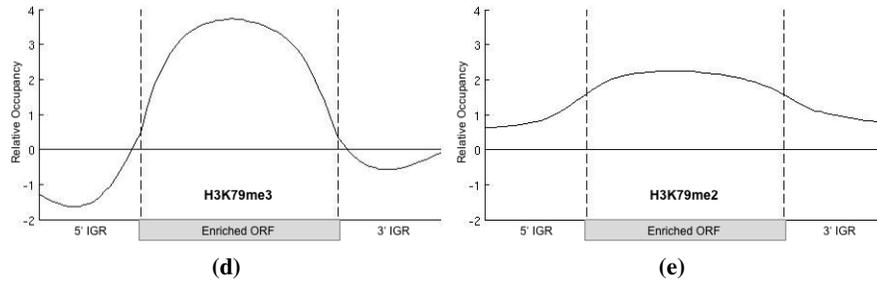


Figure 3.4: High-resolution profile of H3K79me2 and H3K79me3 across the yeast genome with global occupancy analysis. (d) and (e): Average profile of H3K79me3- (d) and H3K79me2- (e) enriched ORFs. A gene was considered to be enriched if at least 50 % of its ORF was covered by the modification. ORFs were aligned according to their translational start and stop sites, similar to an approach by the Young lab (Pokholok et al., 2005). Each ORF was divided into 40 bins of equal length, probes were assigned accordingly, and average enrichment values were calculated for each bin. Probes in promoter regions (500 bp upstream of transcriptional start site) and 3' UTR (500 bp downstream of stop site) were assigned to 20 bins, respectively. The average enrichment value for each bin was plotted.

chromatin (Figure 3.4 (a), page 101). In total, H3K79me2 covered ~22 % and H3K79me3 ~35 % of the genome with only 2 % overlap, suggesting that these H3K79 methyl marks were associated with distinct genomic regions.

To compare our localization data with known genome-wide nucleosome occupancy data, we overlaid our H3K79 methylation profiles with nucleosome position data predicted by a Hidden Markov Model (HMM) (Lee et al., 2007) (Figure 3.4 (a), page 101). Despite the slight differences in resolution, we found enriched regions colocalizing with regions of known nucleosome occupancy.

To further ensure that our profiles truly reflected specific H3K79 methylation marks recognized by the two antibodies, control experiments were performed in a strain lacking the H3K79 methyltransferase Dot1, in which H3K79 methylation is completely eliminated. The genome-wide control profiles showed randomly scattered background peaks and a trend toward occupancy of repetitive regions (Figure 3.5,

page 103), demonstrating that the antibodies were specific for their respective H3K79 methylation state.

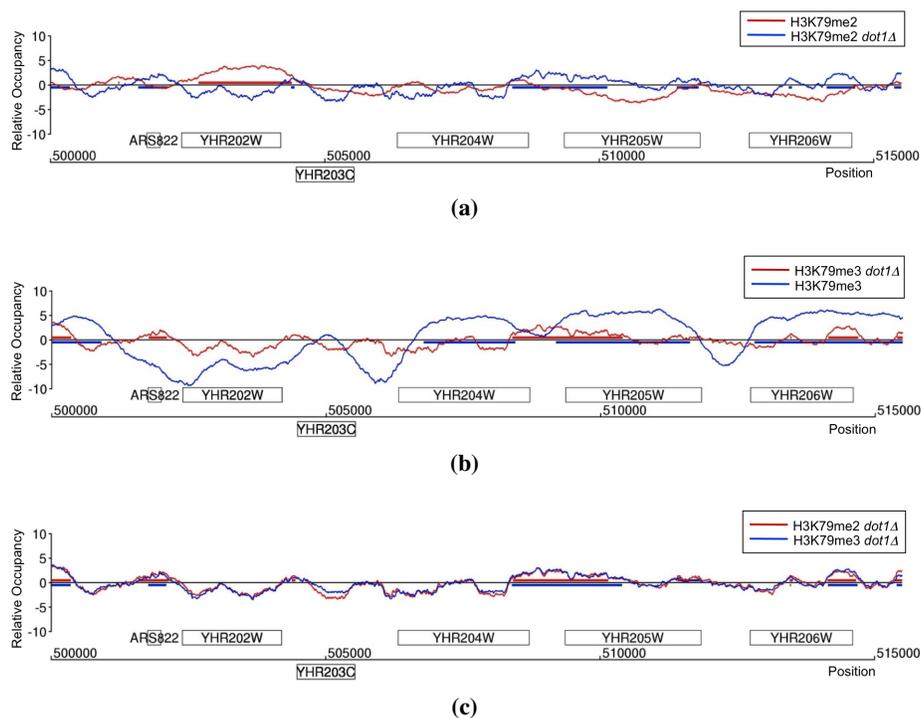


Figure 3.5: H3K79me2 and H3K79me3 antibodies were specific. (a): Comparison of H3K79me2 profiles in wildtype and *dot1* deletion strain. In the *dot1* deletion strain, H3K79me2 was eliminated and the ChIP-on-chip profile showed only background peaks. (b): Comparison of H3K79me3 profiles in wildtype and *dot1* deletion strain. In *dot1* deletion strain H3K79me3 was eliminated, and the ChIP-on-chip profile showed only background peaks. (c): Comparison of H3K79me2 and H3K79me3 background profiles from *dot1* deletion strain. The background was independent of the antibody.

3.2.5 Differential association of H3K79 di- and trimethylation with promoters and ORFs

Having established the detailed maps of H3K79 di- and trimethylation, we wanted to understand the general features of the occupied regions. Genome-wide analy-

sis revealed that H3K79me2 and H3K79me3 covered 1866 and 2350 of 6576 total open reading frames (ORFs), respectively. As expected, the two sets of ORFs enriched with either H3K79 di- or trimethylation overlapped in very few genes (Figure 3.4 (b), page 101).

Interestingly, H3K79me3-enriched ORFs were longer (median 1815 bp) whereas H3K79me2-enriched ORFs were shorter (median 848 bp) relative to the average ORF (median 1067 bp) (Figure 3.4 (c), page 101).

To visualize the average profile of ORFs marked with H3K79me3 and H3K79me2, all enriched ORFs were aligned according to the location of translation initiation and termination sites, similar to an earlier published analysis (Pokholok et al., 2005) (Figure 3.4 (d) and (e), page 102). Consistent with previous studies (Pokholok et al., 2005), H3K79me3 was uniformly enriched within the protein-coding region of genes (Figure 3.4 (d), page 102). In contrast, H3K79me2-enriched ORFs showed that H3K79 dimethylation was not only found in the protein-coding regions of genes, but also covered their promoter region (Figure 3.4 (e), page 102). In general, genes enriched for H3K79me2 in their promoter were also enriched for H3K79me2 in their ORFs (Figure 3.6 (a), page 105), but in addition, there were genes whose promoter region was occupied without extending into the downstream ORF (Figure 3.6 (b), page 105). Overall, H3K79me3 was found in only a few promoters, whereas H3K79me2 covered promoter regions more frequently.

Promoters comprise a fraction of the intergenic region that we defined as regions that do not encode protein. Consistent with the higher occupancy of H3K79me2 in promoter regions, we found that H3K79me2 covered ~20 % of the intergenic regions, whereas H3K79me3 covered less than 4 %.

3.2.6 H3K79 di- and trimethylation associated with genes that tend to be transcriptionally less active

In order to examine the correlation between H3K79 di- and trimethylation and gene expression, enriched genes were assigned into five different classes according to their transcription rate (Holstege et al., 1998), and a composite occupancy profile for each class was determined (Figure 3.7 (a) and (b), page 107). Genes enriched for H3K79 dimethylation had a tendency to be present at higher levels

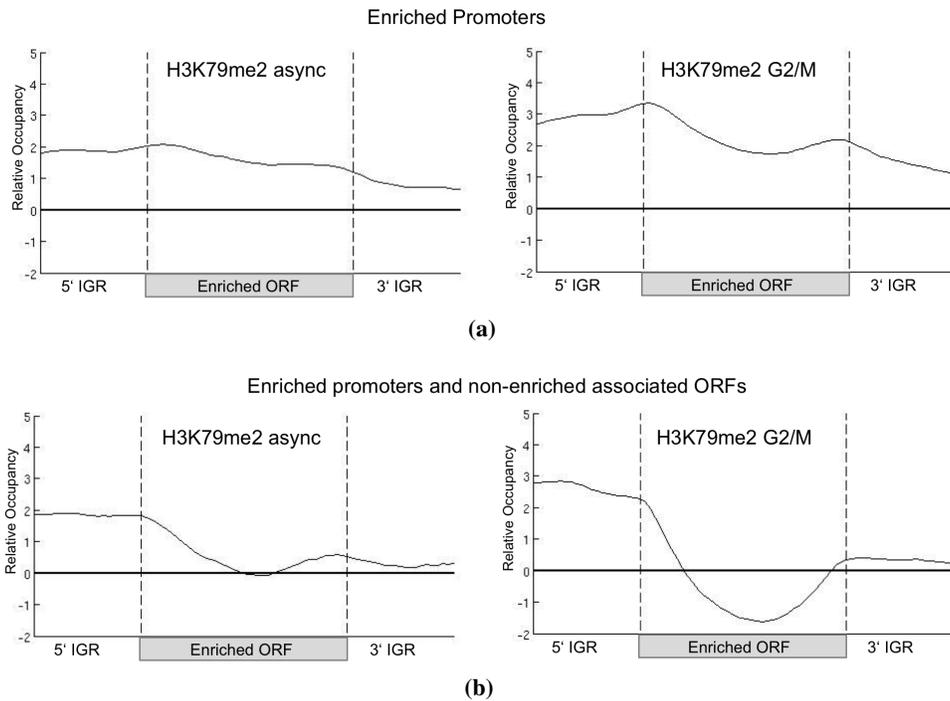


Figure 3.6: H3K79me2 was enriched at promoters. Average profiles of H3K79me2 enriched promoters in asynchronous and G2/M-arrested cells. Promoters were considered to be enriched if the 250 bp region upstream of their transcriptional start site was dimethylated and their promoter did not overlap with any other gene. The average profile was calculated as described in Figure 3.4 (d), page 102. (a): Average profile of genes with H3K79me2 enriched promoters in asynchronous (578 out of 6576) and G2/M arrested cells (1782 out of 6576). (b): Average profile of genes with H3K79me2 enriched promoters and non-enriched associated ORFs in asynchronous (197 out of 6576) and G2/M arrested cells (641 out of 6576).

in transcriptionally less active genes mainly toward their 5' end and promoter region (Figure 3.7 (b), page 107). Consistent with previous findings (Pokholok et al., 2005), no clear correlation to transcriptional activity of genes and enrichment for H3K79 trimethylation was found (Figure 3.7 (a), page 107). Genes with lower expression levels were enriched for H3K79 di- and trimethylation with similar ratios (Figure 3.7 (c), page 107). An exception, however, was the small group of the most highly expressed genes. These genes had very low levels of H3K79 di- and trimethylation in their promoters as well as in their ORFs.

Gene expression has been reported to correlate inversely with nucleosome occupancy in promoters (Lee et al., 2007), so it was not surprising to find low H3K79 di- and trimethylation in these promoters. However, coding regions of highly expressed genes have been shown to be more occupied by nucleosomes than lower expressed genes (Lee et al., 2007). Therefore, ORFs of highly expressed genes are devoid of H3K79 di- and trimethylation despite the dense occupancy with nucleosomes.

3.2.7 M/G1-regulated genes were significantly enriched for H3K79me2

Because of the cell-cycle dependence of H3K79me2, we tested if H3K79 dimethyl-marked genes were regulated during the cell cycle. In budding yeast, ~800 genes change their transcriptional profile and peak in certain stages of the cell cycle (Spellman et al., 1998). We compared ORFs enriched for H3K79 di- and trimethylation with the different classes of cell cycle-regulated genes and asked if the overlap was significant using a hypergeometric test (Tavazoie et al., 1999). Indeed, we found that M/G1-regulated genes were significantly enriched for H3K79me2, but were not enriched for H3K79me3 (Figure 3.7 (d) and (e), page 108). In contrast, genes regulated in the G2 phase showed a significant occupancy with H3K79me3, but are not marked by H3K79me2 (Figure 3.7 (d) and (e), page 108). These results suggest that H3K79 methylation is not random, but rather that it is regulated in conjunction with progression through the cell cycle and might be involved in periodic transcription of genes during distinct cell-cycle phases.

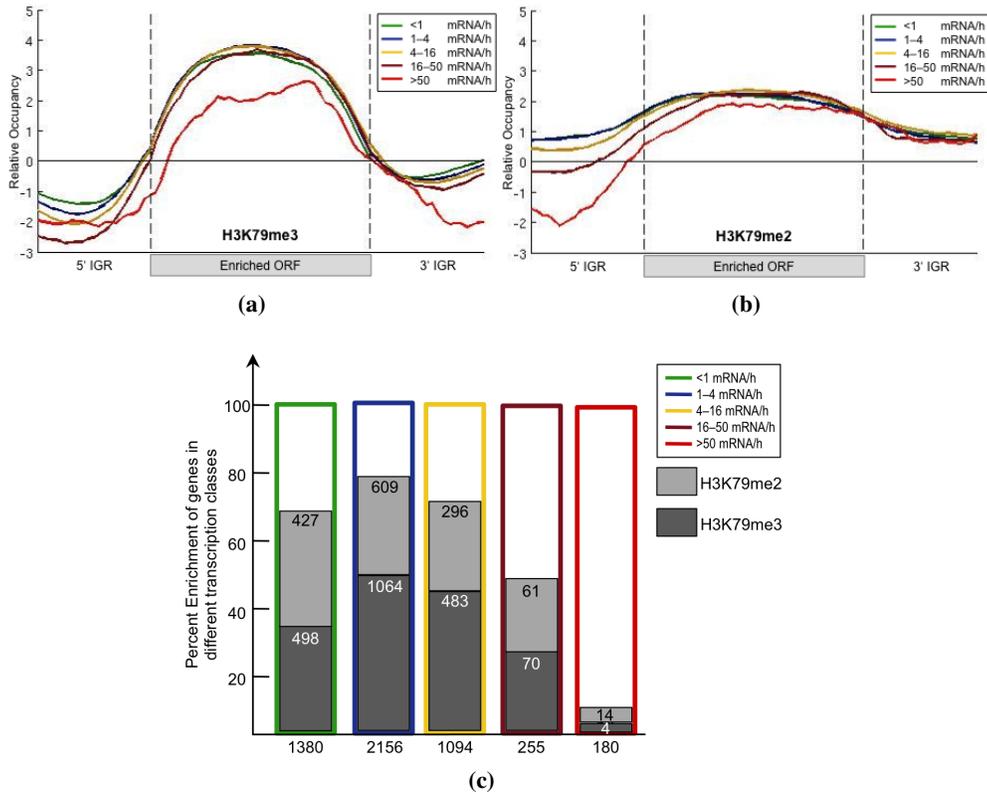


Figure 3.7: Functional characterization of H3K79-dimethylated genes. (a) and (b): Average profiles of H3K79me3- (a) and H3K79me2- (b) enriched ORFs according to transcriptional activity. All genes for which information was available ((Holstege et al., 1998; Pokholok et al., 2005)) were divided into five classes according to their transcriptional rate. Average gene profiles were computed and plotted as described in Figure 3.4 (d), page 102. (c): Percent enrichment of H3K79 di- and trimethylated ORFs in different transcriptional classes. As before, genes were divided into five classes according to their transcriptional activity ((Holstege et al., 1998)), and the percent overlap with H3K79 di- and trimethylated ORFs was plotted.

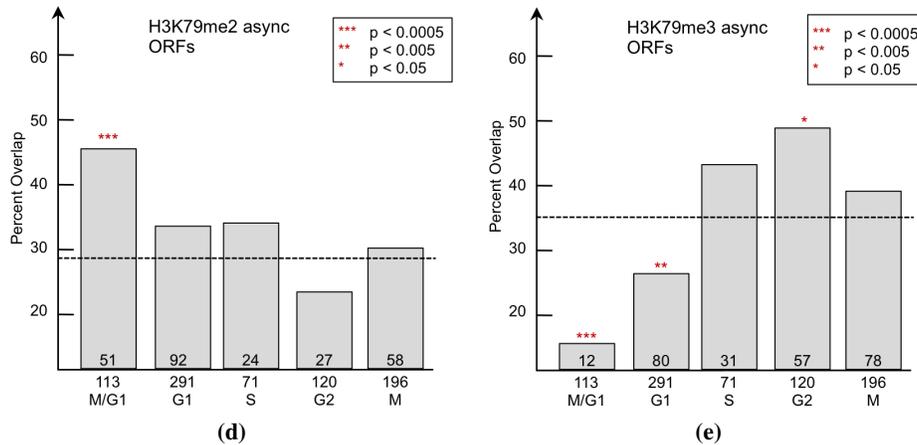


Figure 3.7: Functional characterization of H3K79-dimethylated genes. (d) and (e): Overlap of H3K79me2- and H3K79me3-enriched ORFs with transcriptionally regulated genes for each cell-cycle stage (Spellman et al., 1998). Numbers below the *x*-axis represent total number of genes with periodic transcription. Numbers above the *x*-axis represent the overlap of these genes with those enriched for H3K79me2 and H3K79me3. The percentage of the overlap was plotted on the *y*-axis. H3K79me2-enriched ORFs in asynchronous cells are shown in (d). Expected by chance are 28 % (1866 H3K79me2-enriched ORFs out of 6576 total). H3K79me3-enriched ORFs in asynchronous cells are shown in (e). Expected by chance are 36 % (2350 H3K79me3-enriched ORFs out of 6576 total). The *p*-values were calculated using the hypergeometric test.

3.2.8 Genome-wide association of H3K79 di- but not trimethylation was altered during the cell cycle

Our findings suggested that bulk levels of H3K79me2, but not H3K79me3, changed during the progression of the cell cycle, with reduced levels in G1 and elevated levels in G2/M. To test if levels of H3K79me2 fluctuated at the level of single genes during the cell cycle, we performed ChIP-on-chip assays using chromatin of nocodazole-arrested yeast cells and compared it to asynchronous cells. In nocodazole, cells are arrested in the G2/M phase of the cell cycle and should have the highest level of H3K79me2, while asynchronous cells have a mixed distribution

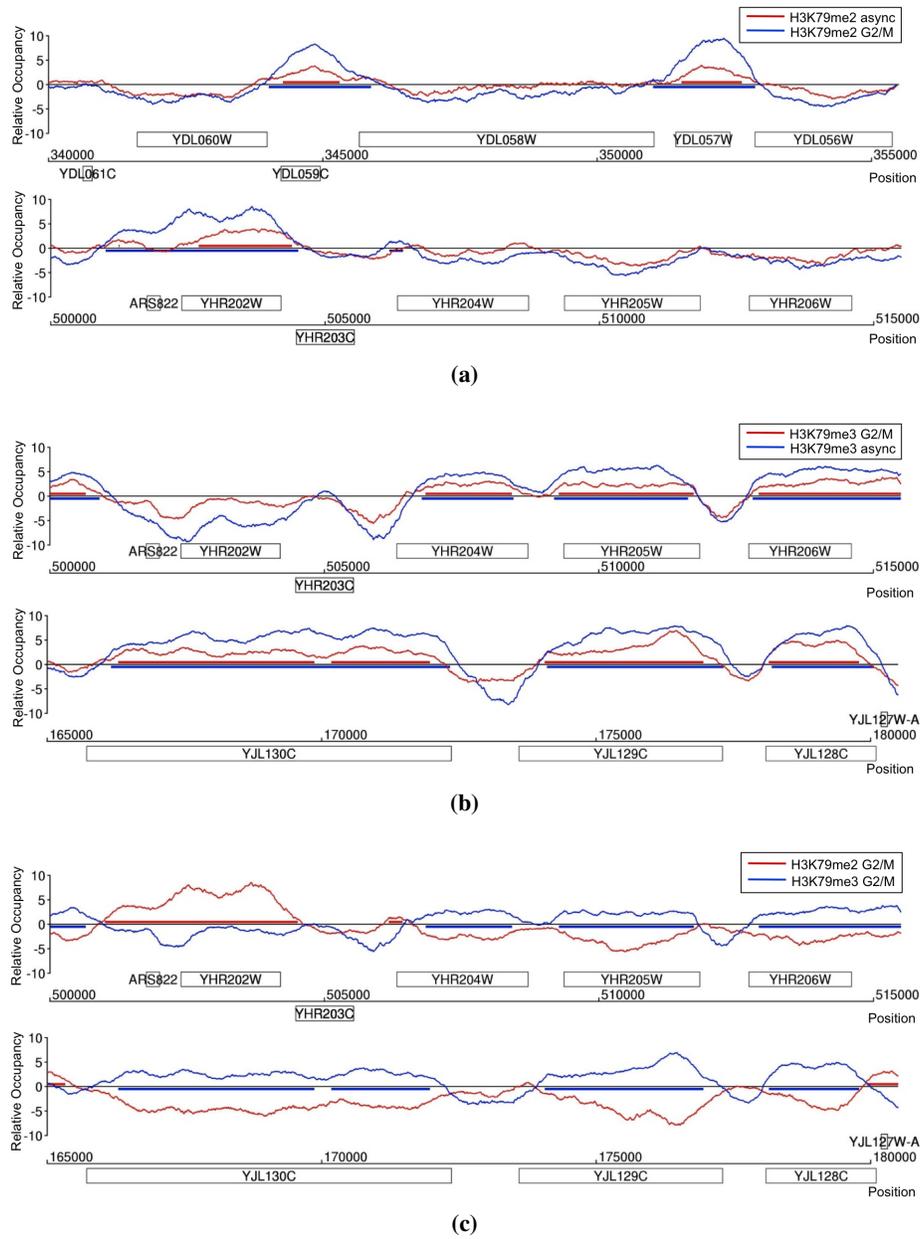


Figure 3.8: Genome-wide H3K79me2 and H3K79me3 profile in G2/M-arrested cells. (a): Comparison of genome-wide H3K79me2 profiles in asynchronous and G2/M-arrested cells. (b): H3K79me3 profiles in asynchronous and in G2/M-arrested cells. (c): H3K79me2 and H3K79me3 profiles of cells arrested in G2/M.

of G1, S, and G2/M phase cells. While H3K79me2 profiles in asynchronous and G2/M-arrested cells were overall similar, with a Spearman rank correlation coefficient of $r = 0.69$ (Figure 3.8 (a), page 109), a detailed analysis revealed important differences between them. In asynchronous cells, ORFs enriched for H3K79me2 significantly overlapped with genes whose expression is regulated in M/G1.

Interestingly, ORFs enriched for H3K79me2 in G2/M-arrested cells significantly overlapped not only with M/G1- but also with G1-regulated genes (Figure 3.9 (a), page 111). Moreover, this effect expanded into the promoter region of genes, since promoter regions of M/G1- and G1-regulated genes were also significantly enriched for H3K79me2 in G2/M-arrested cells (Figure 3.9 (b), page 111). This result suggests that M/G1- and G1-regulated genes are marked in their ORF and promoters by H3K79me2 during cell-cycle stages (G2/M) when these genes are inactive.

In contrast to H3K79me2, global levels of H3K79me3 were not altered during the cell cycle (Figure 3.3 (c), page 98). In order to confirm this observation on a single-gene level, the H3K79me3 profile in G2/M-arrested cells was determined. As expected, the profiles of the asynchronous and G2/M-arrested cells were similar, occupied the same regions, and correlated with a Spearman rank correlation coefficient of $r = 0.91$ (Figure 3.8 (b), page 109). Consistent with the observations in asynchronous cells, H3K79 di- and trimethylation had distinct and mutually exclusive patterns on chromatin in the G2/M phase and overlapped in only 1% of the genome (Figure 3.8 (c), page 109).

3.2.9 H3K79me2 was found in intergenic regions and covered autonomously replicating sequences in G2/M-arrested cells

To further characterize similarities and differences between H3K79me2 in asynchronous and G2/M-arrested cells, we concentrated on typical genomic features. It is known that nucleosome occupancy does not exhibit large, global variation between cell-cycle phases (Hogan et al., 2006), and observed changes in H3K79 dimethylation pattern during the cell cycle were most likely not due to global changes in nucleosome occupancy.

We visualized the average profile of all genes occupied by H3K79me2 in G2/M-arrested cells. Similar to asynchronous cells, we found that H3K79 dimethylation was not only found in the protein-coding regions of genes, but also covered their

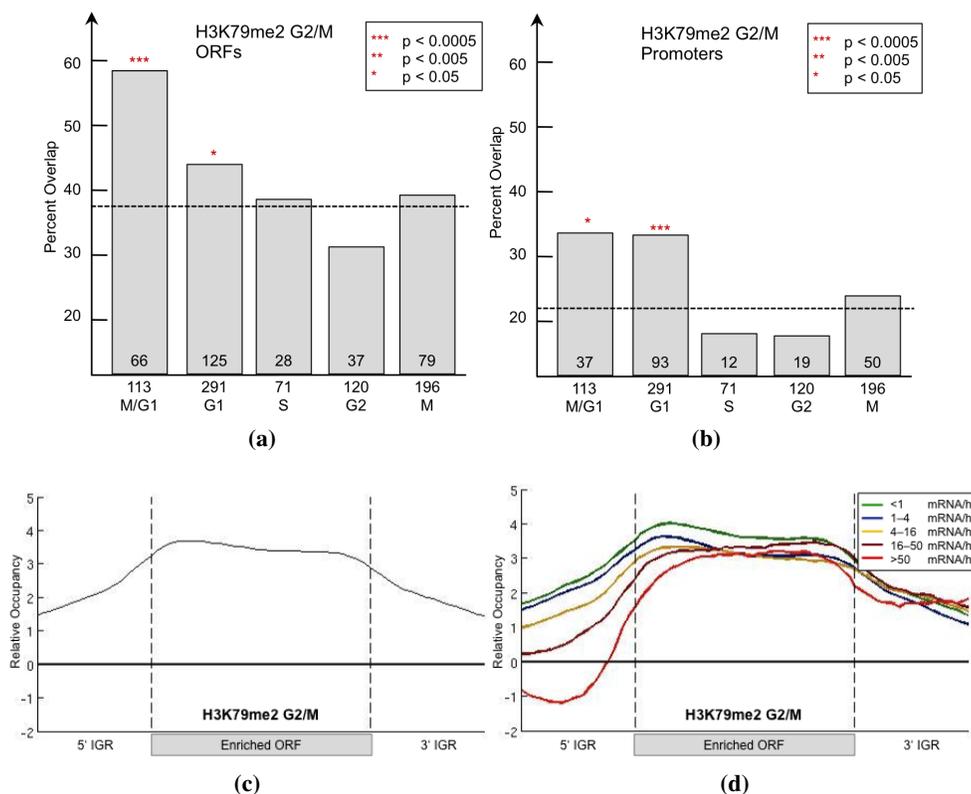
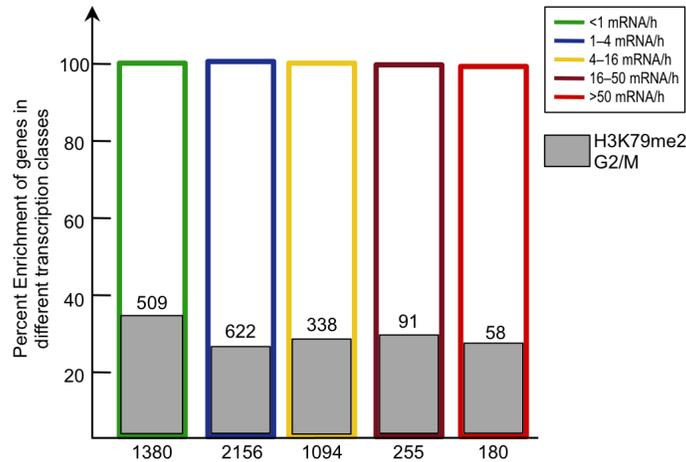
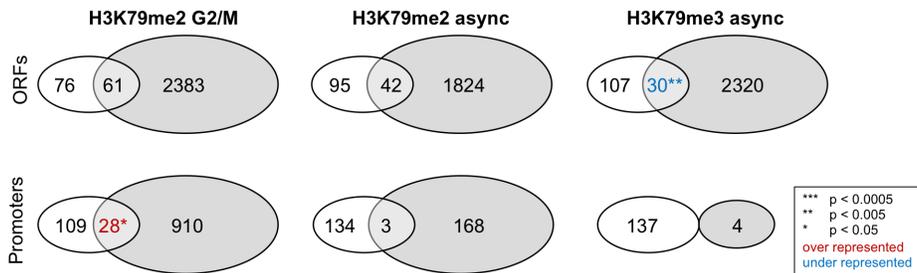


Figure 3.9: Genome-wide characteristics of H3K79me2-enriched genes in G2/M. (a) and (b): Overlap of H3K79me2-enriched ORFs in G2/M-arrested cells with transcriptionally regulated genes for each cell-cycle stage (Spellman et al., 1998). H3K79me2-enriched ORFs in G2/M-arrested cells are shown in (a). Expected by chance are 38% (2444 H3K79me2-enriched ORFs out of 6576 total). H3K79me2-enriched promoters in G2/M are shown in (b). Expected by chance are 23% (1483 H3K79me2-enriched promoters out of 6576 total). (c): Average profile of H3K79me2-enriched ORFs in G2/M-arrested cells. The profile for the average enriched ORF was determined as explained in Figure 3.4 (d), page 102. (d): Average profile of H3K79me2-enriched ORFs in G2/M-arrested cells according to their transcriptional activity. The profile was determined as described in Figure 3.7 (a), page 107.



(e)



(f)

Figure 3.9: Genome-wide characteristics of H3K79me2-enriched genes in G2/M. (e): Percent enrichment of H3K79-dimethylated ORFs in different transcriptional classes. As before, genes were divided into five classes according to their transcriptional activity (Holstege et al., 1998), and the percent overlap with H3K79-dimethylated ORFs in G2/M-arrested cells was plotted. (f): Venn diagram showing overlap of the 137 Swi4-bound genes (Iyer et al., 2001) with H3K79me2- and H3K79me3-enriched ORFs and promoters, respectively. H3K79me2-enriched promoters in G2/M-arrested cells showed significant overlap with Swi4-bound genes. H3K79me3-enriched ORFs and promoters show significant underrepresentation of Swi4-bound genes. Promoters were called enriched when 450 bp upstream of the ORF were covered by the methyl mark.

promoters (Figure 3.9 (c), page 111, and 3.6), page 105. The trend toward occupancy of less-transcribed genes was weaker in G2/M-arrested cells compared to the asynchronous data set (Figure 3.9 (e), page 112), but the relative height of the H3K79 dimethylation profile followed precisely the decreasing order of transcriptional activity (Figure 3.9 (d), page 111). Furthermore, H3K79me2 was enriched more frequently in the promoter regions in the G2/M-arrested cells compared to the asynchronous cells (Figure 3.6, page 105). Not only promoters were enriched more frequently with H3K79me2 in G2/M-arrested cells, but also intergenic regions in general. Indeed, we found them to be covered to ~50 % in G2/M-arrested cells compared to ~20 % in the asynchronous data set.

To further characterize additional chromosomal features for their enrichment with H3K79 di- and trimethylation, we focused on origins of replication (ARSs for autonomously replicating sequences) and centromeres. Intriguingly, we found that ARSs were significantly enriched (131 out of 274 ARSs, $p < 3 \cdot 10^{-14}$) for H3K79me2 in the G2/M phase. In contrast, low ARS occupancy of H3K79 di- and trimethylation (17 and 3 ARSs out of 274) was detected in asynchronous cells. The significant overlap of ARS with H3K79me2 in G2/M-arrested cells indicates a potential role of H3K79me2 in maintaining ARSs in their inactive state. In yeast, the single centromeric nucleosome contains a specialized H3 variant, Cse4, in place of canonical histone H3. Consistent with replacement of H3 at the centromere, neither H3K79me2 nor H3K79me3 antibodies enriched for CEN sequences in either the asynchronous or G2/M data set. The differences in association of H3K79 di- and trimethylation with telomeres and the rDNA locus will be reported elsewhere.

3.2.10 Swi4-regulated genes were significantly enriched for H3K79me2 in their promoter region during G2/M phase

Genome-wide analysis of SBF-binding sites by ChIP-on-chip using the DNA-binding subunit Swi4 revealed that the SBF binds to promoters of genes expressed in G1/S (Harbison et al., 2004; Iyer et al., 2001; Simon et al., 2001). A comparison of H3K79me2-enriched promoters to the 137 promoters bound by Swi4 (Iyer et al., 2001) gave a significant overlap in G2/M-arrested cells (Figure 3.9 (e), page 112). Curiously, this overlap was not significant for H3K79me2-enriched promoters and

ORFs in asynchronous cells, perhaps indicating that H3K79me2 is a consequence of SBF-binding/transcription earlier in the cell cycle. In contrast, ORFs and promoters enriched with H3K79me3 showed significantly lower overlap than expected by chance and no overlap with Swi4-bound genes, respectively (Figure 3.9 (f), page 112). This analysis showed that a significant number of SBF-regulated genes were H3K79 dimethylated in their promoter during G2/M. Moreover, most known

Gene	Async		G2/M		SWI4 binding	Functional Category
	ORFs	Promoters	ORFs	Promoters		
MCM1	+		+			Cell cycle control
CLN1			+	+		Cell cycle control
CLN2	+		+	+	+	Cell cycle control
CLB5	+		+	+		Cell cycle control
CLB6	+		+	+	+	Cell cycle control
NDD1					+	Cell cycle control
SWI4					+	Cell cycle control
HHT2	+		+		+	Chromatin regulation
HHO1	+		+		+	Chromatin regulation
HTZ1	+		+		+	Chromatin regulation

Figure 3.10: Key cell cycle regulators and histone genes were enriched for H3K79me2. Table with characteristic cell cycle regulators and histone genes, showing the enrichment of those for Swi4 binding (Iyer et al., 2001) and H3K79me2 occupancy.

cell-cycle key regulators of G1/S transition, some of them bound by the SBF complex, were clearly enriched for H3K79me2 (Figure 3.10, page 114). These included the G1 cyclins, *CLN1* and *CLN2*, and the S phase cyclins, *CLB5* and *CLB6*. Furthermore, several histone genes (*HHT2*, *HTZ1*, and *HHO1*) were marked with H3K79me2 (Figure 3.10, page 114). Intriguingly, H3K79me2 mainly marks these genes during the cell-cycle stage (G2/M) in which these genes are inactive.

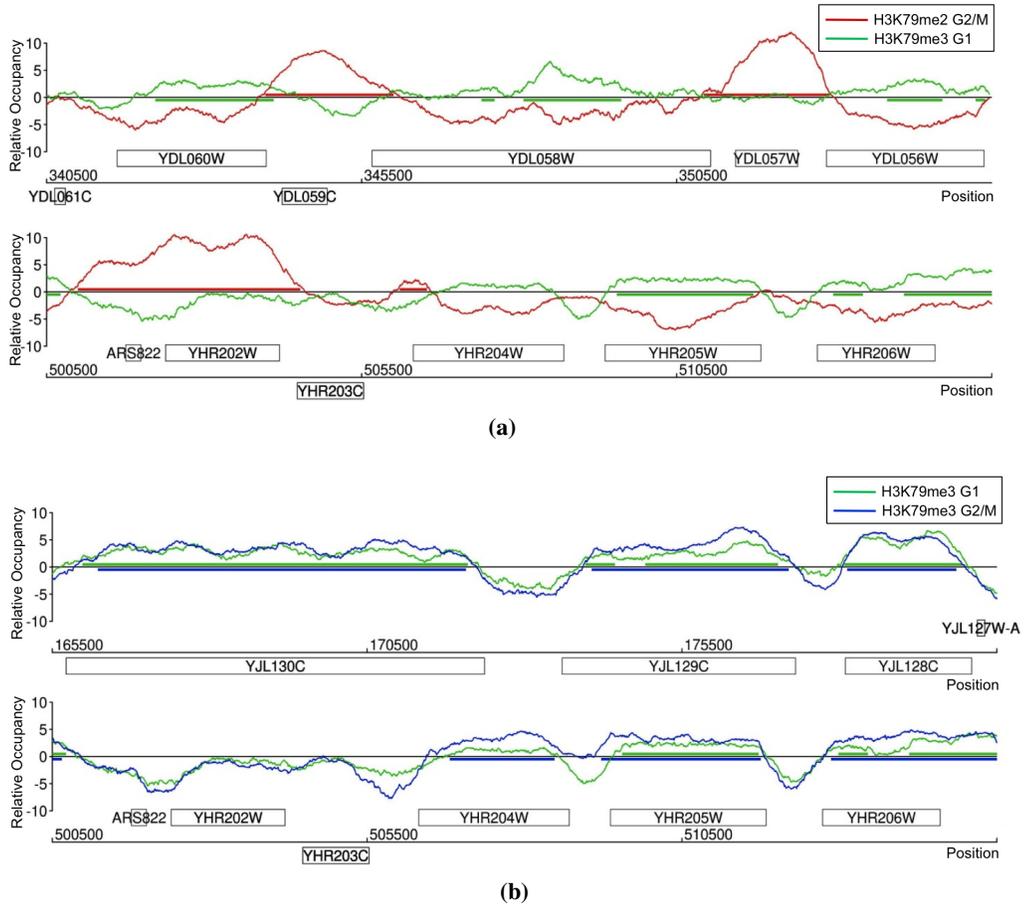


Figure 3.11: Genome-wide H3K79me3 profile in G1-arrested cells was similar to H3K79me3 profile in G2/M- arrested cells. (a): Partial genome plot comparing H3K79me3 profile in G1-arrested cells with H3K79me2 profile in G2/M-arrested cells. (b): Partial genome plot comparing H3K79me3 profiles in G1-arrested with G2/M-arrested cells.

3.2.11 Genome-wide colocalization of H2BK123 monoubiquitination with H3K79 tri- but not dimethylation

Given our results demonstrating that the patterns of H3K79 di- and trimethylation on chromatin are mutually exclusive (Figure 3.4 (b), page 101), we next sought to understand the mechanism by which a single-enzyme Dot1 can distinguish between sites of di- versus trimethylation. The appearance of H3K79me2 after replication in S phase could be explained by either the *de novo* establishment of the methyl mark or the demethylation of an existing H3K79me3 mark resulting in H3K79 dimethylation. The latter would require regions that were dimethylated in G2/M to be trimethylated during G1/S transition. We ruled out this possibility, because ChIP-on-chip of H3K79me3 in G1-arrested cells showed that regions that are dimethylated in G2/M were not trimethylated in G1 (Figure 3.11 (a), page 115), and the profiles of H3K79 trimethylation were similar in G1- and G2/M-arrested cells with a Spearman rank correlation coefficient of $r = 0.82$ (Figure 3.11 (b), page 115).

Based on these observations, it seems to us that demethylation of H3K79me3 may not be the main method by which yeast cells regulate the pattern of H3K79me2 and H3K79me3; however, it remains possible that the trimethyl state could be achieved transiently and quickly removed in G1. Given our data, we hypothesize that H3K79 di- and trimethylation are established independently, and additional factors control the distribution of di- versus trimethylation.

One major candidate is monoubiquitination of lysine 123 on histone H2B mediated by Rad6/Bre1, which is known to be required for proper H3K79 trimethylation. Since ChIP-on-chip studies indicated that the patterns of H3K79 di- and trimethylation are mutually exclusive and have ~2 % overlap throughout the yeast genome, we hypothesize that the pattern of H2B monoubiquitination could control distribution of H3K79 di- versus trimethylation on chromatin. Based on this hypothesis, we predict that factors required for H3K79me3 should function through the regulation of H2B monoubiquitination.

To address this possibility, we set our biochemical screen (GPS) to identify factors that are solely required for H3K79me3 establishment with no effect on H3K79me2. Our survey of extracts of the collection of gene deletion mutants for the presence

of H3K79me₃ resulted in the identification of several known and novel factors, including those that are required for proper H2B monoubiquitination and H3K79 trimethylation.

To further test the above hypothesis that H2BK123 monoubiquitination determines the pattern of H3K79 trimethylation and colocalizes with H3K79 tri- but not dimethylation, we developed a polyclonal antibody specifically recognizing monoubiquitinated H2BK123 (Figure 3.12 (a), page 118). Employing this antibody, we determined a comprehensive genome-wide map of H2BK123 ubiquitination via ChIP-on-chip (Figure 3.12 (b), page 118). To ensure the specific enrichment of H2BK123ub over unmodified H2B, nonspecific binding was blocked with H2B peptide during the immunoprecipitation step. In addition, H2BK123ub enrichment data were normalized using an *H2BK123A* mutant profile obtained from an identical ChIP-on-chip experiment.

Intriguingly, we found that H2BK123ub colocalized with H3K79 trimethylation at many genomic loci, but showed distinct and mutually exclusive patterns with H3K79 dimethylation (Figure 3.12 (b), page 118). Overall, high-confidence regions colocalized with H3K79 trimethylation, and only a very minor fraction overlapped with H3K79 dimethylation (Figure 3.12 (c)). As expected from this analysis, very similar distributions existed at the ORF level, where 812 out of 2350 H3K79 trimethylated ORFs were marked by H2BK123 ubiquitination, but only 45 out of 1866 H3K79-dimethylated ORFs were enriched for H2BK123 ubiquitination (Figure 3.12 (d), page 119).

Since H3K79 dimethylation and H2B monoubiquitination were mutually exclusive on a genome-wide level, we predicted the link of H3K79me₂ to genes expressed specifically during the cell cycle to be independent of H2B monoubiquitination pattern. Indeed, cell cycle-regulated genes, especially those regulated in M/G₁, were not marked by H2B ubiquitination (Figure 3.12 (e), page 119).

Taken together, these findings suggest that the regulation of H2BK123 monoubiquitination is linked to H3K79 tri- but not dimethylation and could play role in distinguishing the genome-wide establishment of H3K79 di- versus trimethylation.

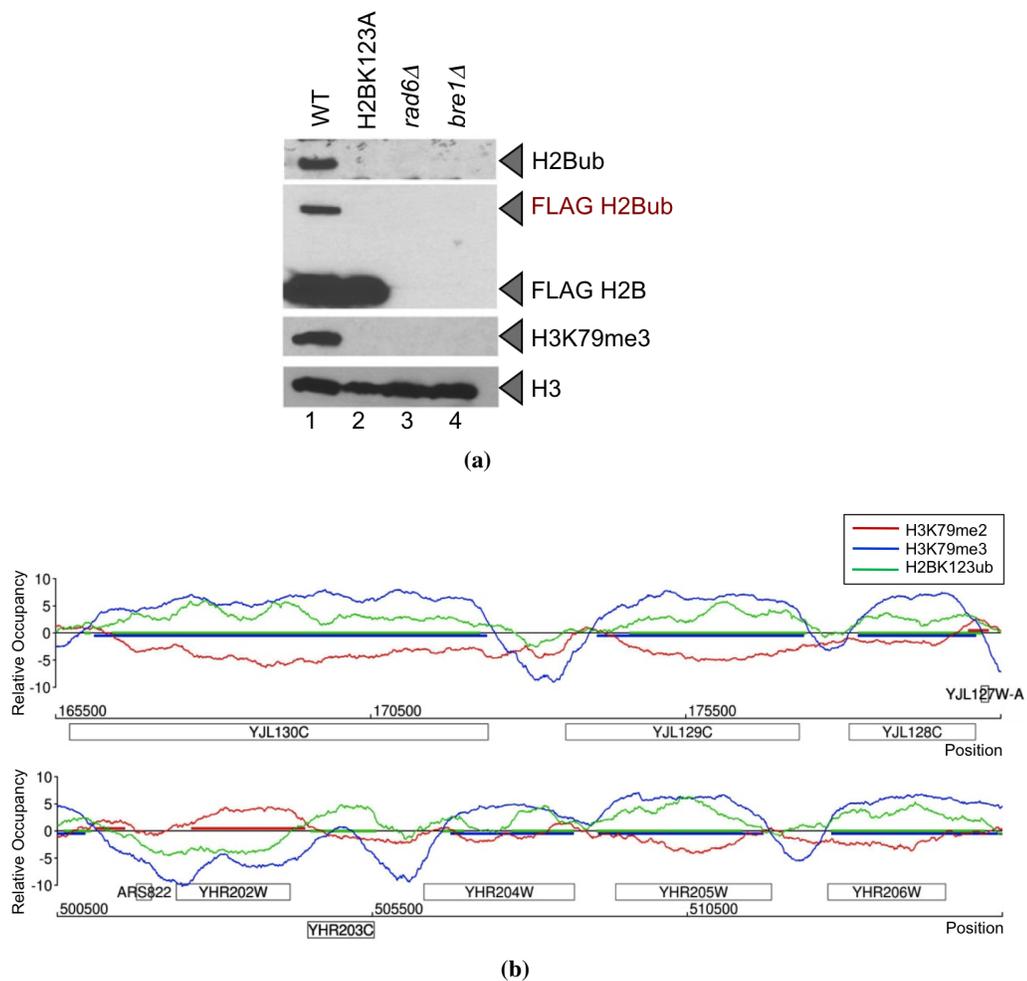


Figure 3.12: Genome-wide profile of H2B monoubiquitination demonstrates an association with H3K79 tri, but not dimethylation. (a): Development of polyclonal antibodies specific to monoubiquitinated H2B. H2BUBq-specific antibodies generated in rabbit were affinity-purified and used for testing extracts from strains either carrying either FLAG::H2B (lane 1) or FLAG::H2BK123R (lane 2) or wild-type H2B in strains deleted for *RAD6* (lane 3) or *BRE1* (lane 4). (b): Overlay of H2BK123ub, H3K79me2, and H3K79me3 profiles. Sample genomic positions for chromosomes 8 and 10 were plotted along the x-axis against the relative occupancy of the indicated histone modifications on the y-axis. ORFs are indicated as rectangles above the axis for Watson and below the axis for Crick strand.

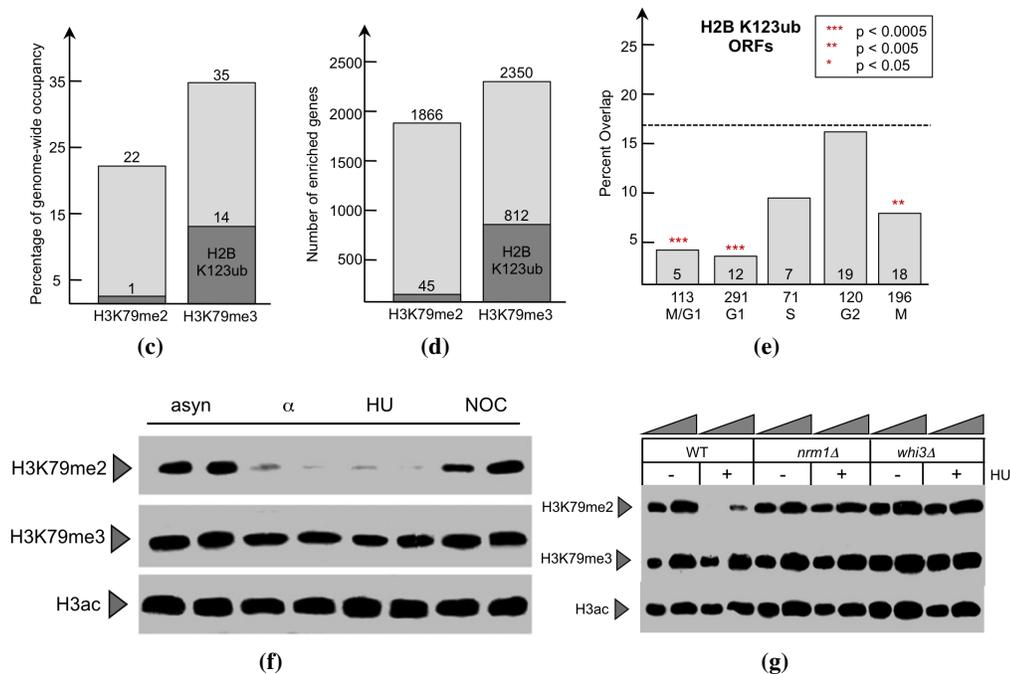


Figure 3.12: Genome-wide profile of H2B monoubiquitination demonstrates an association with H3K79 tri, but not dimethylation. (c): Diagram summarizing the percentage of genome-wide occupancy of H3K79me2 and H3K79me3 and their overlap with H2BK123ub. (d): Diagram illustrating the overlap of H3K79me2 and H3K79me3 enriched with H2BK123ub-enriched genes. (e): Overlap of H2BK123ub-ubiquitinated ORFs with transcriptionally regulated genes for each cell-cycle stage (Spellman et al., 1998). Numbers below and above the *x*-axis represent the total number of genes in each cell-cycle class and their overlap with H2B-ubiquitinated genes, respectively. The percentage of overlap is plotted on the *y*-axis with a dashed line indicating the percentage expected by chance. The *p*-values were calculated using the hypergeometric test. (f): Yeast cell arrest in the G1 and S phases of the cell cycle. To better understand the role of factors responsible for the implementation/removal of H3K79 dimethylation during G1/S stages of cell cycle, we performed a biochemical screen with the entire collection of viable yeast gene deletion mutants arrested with HU. (g): This screen resulted in the identification of Nrm1 and Whi3, which are required for exit from the G1 phase and passage through start of the cell cycle, respectively, as factors required for the removal of dimethylated H3K79, further linking H3K79 methylation status to the cell cycle.

3.2.12 Nrm1 and Whi3 were required for the loss of H3K79 dimethylation in the S phase of the cell cycle

Our data indicate that the establishment of the H3K79 dimethylation pattern requires the Swi4/Swi6 complex, and this mark is removed during G1 and S phase of the cell cycle. Cells arrested in G1 with α factor or in early S phase with hydroxyurea contain low levels of H3K79me2 compared to asynchronous cells or cells arrested in the G2/M phase with nocodazole (Figure 3.12 (f), page 119). To better understand the molecular processes required for the establishment of the H3K79me2 mark and how this posttranslational modification is regulated during certain cell-cycle stages, we performed a biochemical screen using the viable yeast gene deletion mutants arrested in the early S phase with hydroxyurea. Extracts from the arrested collection were prepared, and levels of H3K79me2 were analyzed by western blotting. By arresting cells in hydroxyurea, we eliminated mutants from our screen that simply had low or high levels of H3K79me2 due to the fact that they were enriched in a particular cell-cycle stage.

We identified several factors, including Nrm1 and Whi3, as required for the cell-cycle pattern of H3K79me2 during early S phase (Figure 3.12 (g), page 119). Nrm1 functions as a corepressor with the MBF transcription factor that is required for exit from the G1 phase of the cell cycle (de Bruin et al., 2006). Whi3 is an RNA-binding protein involved in cell-cycle control, and its loss results in the acceleration of the expression of genes controlled by the SBF and MBF complexes. Identification of both Nrm1 and Whi3 in our biochemical screen as specific regulators of H3K79me2 further links the H3K79 methylation pattern to cell-cycle control.

3.3 Discussion

One major aspect of cell-cycle control is the regulatory network of interconnected transcriptional activators (Simon et al., 2001). Often, transcription factors function specifically during one cell-cycle stage and control the expression of transcriptional activators for the subsequent one, thereby forming a feed-forward regulatory circuit to ensure an ordered progression through cell division. In addition to transcriptional regulation, cell-cycle progression also employs proteolysis, phosphorylation, localization, and other regulatory mechanisms to ensure completion

of one event before entry into the next. The role of chromatin modifications in transcription control has been intensively studied in recent years. However, little is known about cell cycle-dependent changes in chromatin structure and the role that chromatin alterations play in normal cell-cycle progression and transcriptional regulation.

Here, we report five fundamental findings, which point toward a connection between the chromatin modification H3K79 methylation and the cell cycle. First and foremost, employing a GPS approach (Schneider et al., 2004), we found that *swi4* Δ and *swi6* Δ mutants are defective in establishing H3K79me₂. Swi4 and Swi6 compose the SBF transcription factor that regulates the expression of genes involved in the G1/S-transition. Second, we showed that global levels of H3K79me₂ fluctuate during the cell cycle, in contrast to H3K79me₃, which was not altered. Third, ChIP-on-chip analysis revealed an overlap between the genes bound by H3K79me₂ and the SBF, suggesting coordination between cell cycle-dependent transcription and dimethylation. Consistent with this idea, H3K79me₂-bound genes were those that showed increased expression during the M/G1 phase of the cell cycle. Fourth, analysis of the entire yeast deletion collection arrested in the G1/S phase pointed to a role for Nrm1 and Whi3, which both impact the transcription by the SBF, and a related transcription complex, MBF, in the process of H3K79me₂ establishment/removal. Finally, we have demonstrated via genome-wide studies that H2BK123 monoubiquitination pattern colocalizes with H3K79 tri- but not dimethylation, indicating that the regulation of H2BK123 monoubiquitination pattern could distinguish the genome-wide establishment of H3K79 di- versus trimethylation.

Based on these observations, we propose the following model of how H3K79me₂ and cell-cycle control relate to one another. Dephosphorylation of Swi6 and synthesis of Swi4 during mitosis allows binding of SBF to its target genes in the late M/early G1 phase of the cell cycle. However, inhibitory factors such as Whi5 prevent transcriptional activation at these targets until early G1, after cells have reached a critical size threshold. At this point in the cell cycle, known as START in budding yeast, activation of the G1 cyclin-dependent kinase Cln3-Cdk1 inside the nucleus renders the SBF functionally active, and a positive feedback loop results in amplification of SBF- as well as MBF-dependent transcription and synthesis of

genes required for progression through G1 and entry into the S phase. The SBF is inactivated upon entry into the S phase through phosphorylation of Swi6, which disrupts its binding to Swi4 and results in its export to the cytoplasm. Without Swi6, Swi4 can no longer bind to DNA. Our observations that levels of H3K79me2 increase during the S phase and remain high during the G2/M phase, combined with our genome-wide location analysis showing that H3K79me2-occupied genes overlap extensively with those expressed specifically in the G1 phase (or bound by Swi4), suggest that H3K79me2 marks cell cycle-specific genes during G2/M. Whether H3K79me2 is a consequence of gene inactivation or if it actively causes transcriptional inactivation remains to be determined. Here, we provide evidence that the modification is created *de novo*—that is, through the addition of two methyl groups to an unmodified H3K79—and is not generated by demethylation of an existing H3K79me3 residue.

How does Swi4/Swi6 control H3K79me2 oscillations? Based on our GPS screen showing that *swi4* Δ and *swi6* Δ mutants are unable to establish the H3K79me2 mark (Figure 3.1, page 95, and 3.2, page 96), it is formally possible that the SBF could directly establish dimethylation. However, sequence and motif analysis of Swi4 and Swi6 do not indicate that either protein contains enzymatic domains that might catalyze dimethylation. Instead, the SBF might activate genes that are required for H3K79me2 formation. One candidate could be *DOT1*. SBF could regulate H3K79me2 establishment through the regulation of *DOT1*, but we were able to rule out this possibility by showing that Dot1 expression (Figure 3.2 (c), page 96) is not altered in *swi4* and/or *swi6* deletion mutants.

The SBF could also control H3K79me2 by recruiting factors required for establishment of this modification to target genes. This scenario is supported by our observation that H3K79me2 is higher at cell cycle-regulated ORFs and promoters that are also bound by Swi4 and is consistent with our genome-wide CHIP-on-chip studies. It also raises the interesting situation where transcription in the preceding G1 could establish a chromatin mark that could be stably inherited and recognized in subsequent cell cycles via an epigenetic mechanism. The cell-cycle oscillation of H3K79me2 could also solve an outstanding question in the field of how SCB-less genes are regulated via the SBF complex. Cross and colleagues have re-

ported a considerable overlap in SBF and MBF target genes as well as cell cycle-dependent genes whose expression requires one or more of these factors (Bean et al., 2005). However, many of these identified genes lack the canonical SCB- and MCB-binding sites in their promoter elements. The existence of H3K79me2 in the promoter regions of SBF-bound genes and their possible propagation through epigenetic mechanisms could, in part, explain how these genes are regulated during the cell cycle.

Why are the patterns of H3K79 di- and trimethylation mutually exclusive? Our ChIP-on-chip analysis of H3K79 methylation has indicated that there is only ~2% overlap between H3K79me2 and H3K79me3 throughout the yeast genome (Figure 3.4 (b), page 101). We found this observation to be quite unexpected. Histone H2B monoubiquitination pattern may in part explain this observation, since H3K79me3 is dependent on H2B monoubiquitination. Therefore, the pattern of H2B monoubiquitination could control distribution of H3K79 di- versus trimethylation on chromatin. Indeed, via our GPS, we have identified several factors required for the proper establishment of H3K79 tri- but not dimethylation and have demonstrated that all of these factors exert their H3K79me3 regulatory activity through the regulation of H2B monoubiquitination levels. Furthermore, our genome-wide studies have demonstrated that the pattern of H2B monoubiquitination appears to be excluded from cell cycle-regulated genes as well as genes enriched for H3K79me2, but to be associated with H3K79me3-containing regions.

Previous analysis of H3K79 methylation during the *S. cerevisiae* cell cycle using an antibody that recognized both H3K79 di- and trimethylation suggested that this modification increased weakly during the late S phase and remained constant during the G2/M phase (Zhou et al., 2006). Our results extend this finding, since we were able to distinguish between the di- and trimethylated forms of H3K79. In contrast to previous findings that implicated di- and trimethylation of H3K79 in redundant functions (Frederiks et al., 2008; Shahbazian et al., 2005), our results clearly show they have distinct, nonoverlapping roles. ChIP-on-chip assays revealed mutually exclusive patterns of H3K79 di- and trimethylation across the yeast genome. In addition, factors influencing H3K79me2 levels had no effect on H3K79me3 levels, and only H3K79me2 levels were cell-cycle regulated. Mitosis-

specific dimethylation of H3K79 in human cells and of the synonymous residue H3K76 in *Trypanosoma brucei* have also been observed (Feng et al., 2002; Janzen et al., 2006). Given the similarity of cell-cycle regulation across species and the evolutionary conservation of H3K79 methylation, it seems probable that dimethylation, transcriptional control, and cell-cycle progression are also coupled in higher eukaryotes as they appear to be in yeast. With this study, we reveal a link between histone modifications and cell-cycle control. The discovery of a functional relationship between H3K79me₂ and the transcriptional cell-cycle activator SBF gives insight into the complex network of cell-cycle control and chromatin regulation.

3.4 Experimental procedures

3.4.1 Yeast strains

Strains used for GPS studies were obtained from the deletion collection mutants. Strains used for cell-cycle studies in Figure 3.3, page 99, were derived from BY4741 (*a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) or from W303 (*a bar ura3-1 trp1-1 ade2-1 his3-11-15 leu2-3,112 can1-100*).

3.4.2 Global proteomic analysis of histone modifications

GPS analyses were carried out as described previously (Schneider et al., 2004) with the use of antibodies specific for dimethylated lysine 79 (04-835; Millipore; Lake Placid, NY) or trimethylated lysine 79 (ab2621; Abcam; Cambridge, UK) and for dimethylated lysine 4 of histone H3 or acetylated histone H3.

3.4.3 RT-PCR

Cells were grown in YPD to an OD₆₀₀ of 0.8–1.0. RNA was extracted by resuspending the cells in 50 ml of TES buffer (10 mmol/l Tris, 10 mmol/l EDTA, 0.5 % SDS) and 500 ml of 65 °C acidic phenol. Cells were then placed at 65 °C for 1 h and vortexed every 10 min. After phenol-chloroform extraction of the cell extract, RNA was precipitated with 100% ethanol and treated with RNase-free DNase I. One microgram of RNA was used in a reverse transcriptase reaction (Clontech;

Palo Alto, CA). cDNA (5 μ l) was used in a 50 μ l PCR with 26 cycles of amplification.

3.4.4 Cell synchronization and FACS analysis

α factor was used at a final concentration of 1 mg/ml. The *GAL-CDC6* experiment was carried out as described (Biggins and Murray, 2001). Analysis of DNA content by flow cytometry was carried out as previously described.

3.4.5 ChIP and genome-wide ChIP-on-chip

ChIP and genome-wide location analyses were performed as described previously, using the adapted linear amplification method that involves two rounds of T7 RNA polymerase amplification (van Bakel et al., 2008). In brief, yeast cells (500 ml) were grown in a rich medium to an OD600 of 0.8–0.9 and were cross-linked with 1% formaldehyde for 20 min before chromatin was extracted. The chromatin was sonicated (Bioruptor, Diagenode; Sparta, NJ: 10 cycles, 30 s on/off, high setting) to yield an average DNA fragment of 500 bp. H3K79me2 or H3K79me3 antibodies (4 μ l) were coupled to 60 μ l of protein A magnetic beads (Invitrogen). After reversal of the crosslinking and DNA purification, the immunoprecipitated and input DNA were amplified to about 6 μ g aRNA using T7 RNA polymerase in two rounds. Samples were labeled with biotin, and the immunoprecipitated and input samples were hybridized to two Affymetrix 1.0R *S. cerevisiae* microarrays, which are comprised of over 3.2 million probes covering the complete genome. Probes (25-mer) are tiled at an average of 5 bp resolution, creating an overlap of approximately 20 bp between adjacent probes.

3.4.6 Data analysis

We used the *Model-based Analysis of Tiling-arrays* (MAT) algorithm to reliably detect enriched regions (Johnson et al., 2006; Droit et al., 2010). MAT was applied to corresponding immunoprecipitated and input sample arrays, and the probe behaviour model was estimated by examining the signal intensity, sequence, and copy number of all probes on an array. After probe behaviour model fitting, the residuals between the model and observation were normally distributed and centred at 0.

MAT uses a score function to identify regions of ChIP enrichment, which allows robust p -value and false-discovery-rate calculations. MAT scores were calculated for all probes using a 300 bp sliding window. The MAT scores, as a measure for relative enrichment, were visualized along the whole genome using custom-written scripts in the language and statistics environment R.

Annotations for ORFs, ARSs, and centromeres were derived from the SGD database. An ORF was termed enriched if at least 50% of all probes had a MAT score above a threshold of 1.5. Promoters were defined as enriched if all probes 300 bp upstream of the transcriptional start site were above the MAT score cutoff. Promoters which overlap with ORFs of other genes were not considered. ARSs were defined to be enriched when they were completely covered by the methyl mark. Nucleosome position data were derived from Lee et al., 2007.

3.4.7 Analysis of published genome-wide expression and binding data sets

For transcriptional analysis, information about transcriptional frequencies for all genes was used from Holstege et al., 1998 and Young Lab, 2005. To examine the overlap between H3K79-methylated genes and cell cycle-regulated genes, the Spellman et al., 1998 analysis was chosen, because it incorporates several types of yeast cell-cycle expression data. To compare our data sets with SBF-binding data from previously published genome-wide localization analyses, we used the supplemental data available from Iyer et al., 2001. Their `Figure 3_data.xls` file was used to identify SBF-bound and SBF + MBF-bound ORFs. For all data sets, any duplicate ORFs and ORFs with highly repetitive sequences were removed, and the systematic yeast gene names were used. MATLAB scripts were used to compare the different data sets.

3.4.8 Statistics

In order to explore the statistical significance of the overlap between different sets of genes, we used the hypergeometric test as described (Tavazoie et al., 1999).

Bibliography

- Altaf, M., Utley, R., Lacoste, N., Tan, S., Briggs, S. and Cote, J. (2007). Interplay of chromatin modifiers on a short basic patch of histone H4 tail defines the boundary of telomeric heterochromatin. *Mol Cell* 28, 1002–14. → pages 91
- Andrews, B. and Herskowitz, I. (1989). The yeast SWI4 protein contains a motif present in developmental regulators and is part of a complex involved in cell-cycle-dependent transcription. *Nature* 342, 830–3. → pages 92
- Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–37. → pages 91
- Bean, J. M., Siggia, E. D. and Cross, F. R. (2005). High functional overlap between MluI cell-cycle box binding factor and Swi4/6 cell-cycle box binding factor in the G1/S transcriptional program in *Saccharomyces cerevisiae*. *Genetics* 171, 49–61. → pages 123
- Biggins, S. and Murray, A. (2001). The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. *Genes Dev* 15, 3118–29. → pages 100, 125
- Breeden, L. (1996). Start-specific transcription in yeast. *Curr Top Microbiol Immunol* 208, 95–127. → pages 92
- Breeden, L. and Nasmyth, K. (1987). Cell cycle control of the yeast HO gene: cis- and trans-acting regulators. *Cell* 48, 389–97. → pages 92
- de Bruin, R., Kalashnikova, T., Chahwan, C., McDonald, W., Wohlschlegel, J., Yates, J., Russell, P. and Wittenberg, C. (2006). Constraining G1-specific transcription to late G1 phase: the MBF-associated corepressor Nrm1 acts via negative feedback. *Mol Cell* 23, 483–96. → pages 92, 120
- Dover, J., Schneider, J., Tawiah-Boateng, M., Wood, A., Dean, K., Johnston, M. and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 277, 28368–71. → pages 91, 97

- Droit, A., Cheung, C. and Gottardo, R. (2010). rMAT—an R/Bioconductor package for analyzing ChIP-chip experiments. *Bioinformatics* 26, 678–9. → pages 100, 125
- Feng, Q., Wang, H., Ng, H., Erdjument-Bromage, H., Tempst, P., Struhl, K. and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* 12, 1052–8. → pages 124
- Frederiks, F., Tzouros, M., Oudgenoeg, G., van Welsem, T., Fornerod, M., Krijgsveld, J. and van Leeuwen, F. (2008). Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. *Nat Struct Mol Biol* 15, 550–7. → pages 100, 123
- Harbison, C., Gordon, D., Lee, T., Rinaldi, N., Macisaac, K., Danford, T., Hannett, N., Tagne, J., Reynolds, D., Yoo, J., Jennings, E., Zeitlinger, J., Pokholok, D., Kellis, M., Rolfe, P., Takusagawa, K., Lander, E., Gifford, D., Fraenkel, E. and Young, R. (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature* 431, 99–104. → pages 92, 113
- Harrington, L. and Andrews, B. (1996). Binding to the yeast SwI4,6-dependent cell cycle box, CACGAAA, is cell cycle regulated in vivo. *Nucleic Acids Res* 24, 558–65. → pages 95
- Hereford, L., Osley, M., Ludwig, T. and McLaughlin, C. (1981). Cell-cycle regulation of yeast histone mRNA. *Cell* 24, 367–75. → pages 92
- Hogan, G., Lee, C. and Lieb, J. (2006). Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. *PLoS Genet* 2, e158. → pages 110
- Holstege, F., Jennings, E., Wyrick, J., Lee, T., Hengartner, C., Green, M., Golub, T., Lander, E. and Young, R. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717–28. → pages 104, 107, 112, 126
- Hwang, W., Venkatasubrahmanyam, S., Ianculescu, A., Tong, A., Boone, C. and Madhani, H. (2003). A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell* 11, 261–6. → pages 97

- Im, H., Park, C., Feng, Q., Johnson, K., Kiekhäfer, C., Choi, K., Zhang, Y. and Bresnick, E. (2003). Dynamic regulation of histone H3 methylated at lysine 79 within a tissue-specific chromatin domain. *J Biol Chem* 278, 18346–52. → pages 91
- Iyer, V., Horak, C., Scafe, C., Botstein, D., Snyder, M. and Brown, P. (2001). Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409, 533–8. → pages 92, 112, 113, 114, 126
- Janzen, C., Hake, S., Lowell, J. and Cross, G. (2006). Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in *Trypanosoma brucei*. *Mol Cell* 23, 497–507. → pages 124
- Johnson, W., Li, W., Meyer, C., Gottardo, R., Carroll, J., Brown, M. and Liu, X. (2006). Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci U S A* 103, 12457–62. → pages 100, 125
- Koch, C., Moll, T., Neuberg, M., Ahorn, H. and Nasmyth, K. (1993). A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase. *Science* 261, 1551–7. → pages 92
- Krivtsov, A., Feng, Z., Lemieux, M., Faber, J., Vempati, S., Sinha, A., Xia, X., Jesneck, J., Bracken, A., Silverman, L., Kutok, J., Kung, A. and Armstrong, S. (2008). H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* 14, 355–68. → pages 91
- 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. *Mol Cell* 11, 721–9. → pages 91
- Lacoste, N., Utley, R., Hunter, J., Poirier, G. and Cote, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. *J Biol Chem* 277, 30421–4. → pages 91

- Lee, W., Tillo, D., Bray, N., Morse, R., Davis, R., Hughes, T. and Nislow, C. (2007). A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* 39, 1235–44. → pages 101, 102, 106, 126
- Lu, X., Simon, M., Chodaparambil, J., Hansen, J., Shokat, K. and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol* 15, 1122–4. → pages 91
- Luger, K., Mader, A., Richmond, R., Sargent, D. and Richmond, T. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–60. → pages 90
- Martin, C. and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6, 838–49. → pages 91
- Miao, F. and Natarajan, R. (2005). Mapping global histone methylation patterns in the coding regions of human genes. *Mol Cell Biol* 25, 4650–61. → pages 91
- Ng, H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y. and Struhl, K. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev* 16, 1518–27. → pages 91
- Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V., Su, L., Xu, G. and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. *Cell* 121, 167–78. → pages 91
- Onishi, M., Liou, G., Buchberger, J., Walz, T. and Moazed, D. (2007). Role of the conserved Sir3-BAH domain in nucleosome binding and silent chromatin assembly. *Mol Cell* 28, 1015–28. → pages 91
- Piatti, S., Lengauer, C. and Nasmyth, K. (1995). Cdc6 is an unstable protein whose de novo synthesis in G1 is important for the onset of S phase and for preventing a 'reductional' anaphase in the budding yeast *Saccharomyces cerevisiae*. *The EMBO journal* 14, 3788–99. → pages 100

- Pokholok, D., Harbison, C., Levine, S., Cole, M., Hannett, N., Lee, T., Bell, G., Walker, K., Rolfe, P., Herbolzheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. and Young, R. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* *122*, 517–27. → pages 91, 102, 104, 106, 107
- Schneider, J., Dover, J., Johnston, M. and Shilatifard, A. (2004). Global proteomic analysis of *S. cerevisiae* (GPS) to identify proteins required for histone modifications. *Methods Enzymol* *377*, 227–34. → pages 93, 121, 124
- Schubeler, D., MacAlpine, D., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D., O'Neill, L., Turner, B., Delrow, J., Bell, S. and Groudine, M. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* *18*, 1263–71. → pages 91
- Shahbazian, M., Zhang, K. and Grunstein, M. (2005). Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. *Mol Cell* *19*, 271–7. → pages 100, 123
- Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* *75*, 243–69. → pages 90, 91
- Simon, I., Barnett, J., Hannett, N., Harbison, C., Rinaldi, N., Volkert, T., Wyrick, J., Zeitlinger, J., Gifford, D., Jaakkola, T. and Young, R. (2001). Serial regulation of transcriptional regulators in the yeast cell cycle. *Cell* *106*, 697–708. → pages 92, 113, 120
- Spellman, P., Sherlock, G., Zhang, M., Iyer, V., Anders, K., Eisen, M., Brown, P., Botstein, D. and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* *9*, 3273–97. → pages 92, 95, 106, 108, 111, 126
- Stern, B. M. and Murray, A. W. (2001). Lack of tension at kinetochores activates the spindle checkpoint in budding yeast. *Curr Biol* *11*, 1462–7. → pages 100

- Sun, Z. and Allis, C. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418, 104–8. → pages 91, 97
- Tavazoie, S., Hughes, J., Campbell, M., Cho, R. and Church, G. (1999). Systematic determination of genetic network architecture. *Nat Genet* 22, 281–5. → pages 106, 126
- van Bakel, H., van Werven, F., Radonjic, M., Brok, M., van Leenen, D., Holstege, F. and Timmers, H. (2008). Improved genome-wide localization by ChIP-chip using double-round T7 RNA polymerase-based amplification. *Nucleic Acids Res* 36, e21. → pages 125
- van Leeuwen, F., Gafken, P. and Gottschling, D. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* 109, 745–56. → pages 91
- Wood, A., Krogan, N., Dover, J., Schneider, J., Heidt, J., Boateng, M., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J., Johnston, M. and Shilatifard, A. (2003). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* 11, 267–74. → pages 91, 97
- Zhou, H., Madden, B., Muddiman, D. and Zhang, Z. (2006). Chromatin assembly factor 1 interacts with histone H3 methylated at lysine 79 in the processes of epigenetic silencing and DNA repair. *Biochemistry* 45, 2852–61. → pages 123

Chapter 4

Genome-wide analysis of H2B ubiquitination and deubiquitination in the context of histone H3 methylation*

4.1 Introduction

In eukaryotic cells, histone proteins play a central role in packaging the genetic information imbedded within the DNA into distinct active and repressed neighbourhoods. These neighbourhoods are specified by histone variants, ATP-dependent chromatin remodelling, and various posttranslational modifications of histones, the latter including methylation, ubiquitination, acetylation, and others (Shilatifard, 2006; Strahl and Allis, 2000; Workman and Kingston, 1998)). More than one hundred different histone modifications are known to date, often occurring in various combinations across several adjacent nucleosomes or even on a single nucleosome (Bhaumik et al., 2007; Kouzarides, 2007). Frequently, these co-occurring histone marks are functionally linked to regulate different chromatin-

*Manuscript in preparation. See Co-authorship Statement on page xv for details about my contributions.

mediated processes such as transcription, DNA replication, or DNA repair (Suganuma and Workman, 2008).

Sets of specific histone modifications not only occur and function in specific chromatin neighbourhoods, but also regulate each other through a process referred to as crosstalk. Histone crosstalk has been demonstrated to play a central role in many biological processes including stem cell self renewal and development (Smith and Shilatifard, 2009). Besides coordinating histone-modifying enzymes in the same protein complex, crosstalk is manifested by an initial histone mark triggering another modification, likely through regulating the activity of the enzyme responsible for the creation of the downstream mark (Lee et al., 2007a; Suganuma and Workman, 2008).

One intriguing example is the crosstalk between H2BK123 monoubiquitination (H2BK123ub) and two specific methylation marks on histone H3. H2BK123ub is a transient mark established during transcription initiation and elongation by the Rad6/Bre1 ubiquitin ligase complex (Dover et al., 2002; Hwang et al., 2003; Robzyk et al., 2000; Wood et al., 2003a).

Before being removed by its deubiquitinases Ubp8 and Ubp10 (Daniel et al., 2004; Gardner et al., 2005; Henry et al., 2003; Sanders et al., 2002; Emre et al., 2005), H2BK123ub serves as a causal signal for the addition of the more stable methylation marks H3K4me3 and H3K79me3 (Briggs et al., 2002; Dover et al., 2002; Krogan et al., 2003; Nakanishi et al., 2009; Ng et al., 2002; Sun and Allis, 2002; Wood et al., 2003a).

Lysine residues can be mono-, di-, or trimethylated, with each methylation state having distinct roles (Li et al., 2009; Schulze et al., 2009). Histone methyltransferases have critical roles in development and have been linked to cancer and other human diseases (Bhaumik et al., 2007).

In *S. cerevisiae*, the three histone methyltransferases Set1/COMPASS (KMT2), Set2 (KMT3), and Dot1 (KMT4), collectively are responsible for all histone lysine methylations, which thus far have been found only on histone H3 residues H3K4, H3K36, and H3K79 (Shilatifard, 2006).

Histone H3 methylation is present at distinct regions of the genome, and in some cases appears to be tightly correlated with gene expression. Set1/COMPASS,

methylates H3K4 at the 5' end of actively transcribed genes (Bernstein et al., 2002; Briggs et al., 2001; Krogan et al., 2003; Miller et al., 2001; Ng et al., 2003; Santos-Rosa et al., 2002). In contrast, methylation of H3K36 by Set2 is associated with later stages of transcriptional elongation (Strahl et al., 2002), consistent with Set2 being recruited by serine 2 phosphorylated RNAPII CTD. The link between transcription and Dot1-dependent methylation of H3K79 is less clear. Dot1 catalyzes the addition of H3K79 di- and trimethylation to distinct and largely mutually exclusive chromatin regions (Krogan et al., 2003; Schulze et al., 2009; Wood et al., 2003a). H3K79 dimethylation is under the control of the cell-cycle transcription factor Swi4/6 and oscillates during the cell cycle.

Besides their roles during transcription, yeast H3 methylation marks have broader functions in cellular homeostasis, and are involved in processes like telomeric silencing, cell cycle checkpoint control and DNA repair (Humpal et al., 2009; Lacoste et al., 2002; Shilatifard, 2006; van Leeuwen et al., 2002). Furthermore, methylation of H3 lysine 4 and 36 have been spatially and mechanistically linked to the biogenesis of messenger RNA, including splicing, mRNA capping, and polyadenylation in higher eukaryotes (Perales and Bentley, 2009). This is perhaps not surprising given that many RNA processing reactions in eukaryotic cells occur co-transcriptionally on the chromatin template.

While it is clear that H2BK123ub is functionally linked to H3K4 and H3K79 methylation, its overall role in establishing these marks across the genome has not been elucidated due to the lack of proper reagents and a genome-wide profile. Here, we present the first genome-wide map of H2B123ub and a comprehensive comparison of its localization to that of H3 methylation marks in budding yeast, obtained on the same high-resolution platform. Expanding the previously reported general co-localization of H3K79 trimethylation with H2BK123ub, we found that H2BK123ub associated with both its downstream marks H3K4me3 and H3K79me3 on a genome-wide scale.

Significantly, our studies provide evidence that H3K4me3 is linked to deubiquitination of H2B by Ubp8, whereas H3K79me3 is similarly connected to Ubp10. Furthermore, our global genomic analysis demonstrates that monoubiquitination of histone H2B co-occurs with distinct combinations of histone H3 methylation marks

in promoters, 5'end, middle, and 3'end of genes. Specific combinations of those were associated with transcription factor binding, high transcriptional frequencies, and intron-containing genes. Consistent with the latter, the H2B ubiquitin ligase Bre1 was required for optimal splicing of highly transcribed genes.

4.2 Results

4.2.1 Genome-wide association of H2BK123 monoubiquitination and histone H3K4- and H3K79 trimethylation

Indicative of a central role in histone crosstalk, H2BK123 monoubiquitination (H2BK123ub) is required for the establishment of both H3K4- and H3K79 trimethylation (H3K4me3 and H3K79me3) (Figure 4.1 (a), page 137). Besides serving as an upstream regulator of these histone modifications, very little is known about the genomic distribution of H2BK123ub and the spatial relationship to its downstream H3 methylation marks. To bridge this gap, we developed a polyclonal antibody specifically recognizing K123-monoubiquitinated H2B (see Chapter 3 Figure 3.12 (a), page 118), and established comprehensive genome-wide maps of H2BK123ub and H3 methylation marks using ChIP-on-chip on high-resolution tiling arrays.

As shown for illustrative purposes in Figure 4.1 (b), page 137, these modifications were present at distinct regions with H2BK123ub and H3K79me3 primarily covering the entire coding sequence of genes while H3K4me3 mainly covered the 5'end. To investigate the distribution of these marks in more detail, we aligned all enriched ORFs according to the coding start and end sites and determined their average occupancy profile, using an approach applied previously (Pokholok et al., 2005). As evident from the exemplary genomic regions shown in Figure 4.1 (b), page 137, and in agreement with studies in human cells (Minsky et al., 2008), H2BK123ub covered the promoter region of genes and extended into the entire coding sequence (Figure 4.2, page 139). Consistent with earlier findings, H3K4me3 marked the 5'end of ORFs while H3K79me3 covered the entire coding sequence (Figure 4.2, page 139) (Pokholok et al., 2005; Schulze et al., 2009). To examine the spatial relationship between H2BK123ub and its downstream marks more closely,

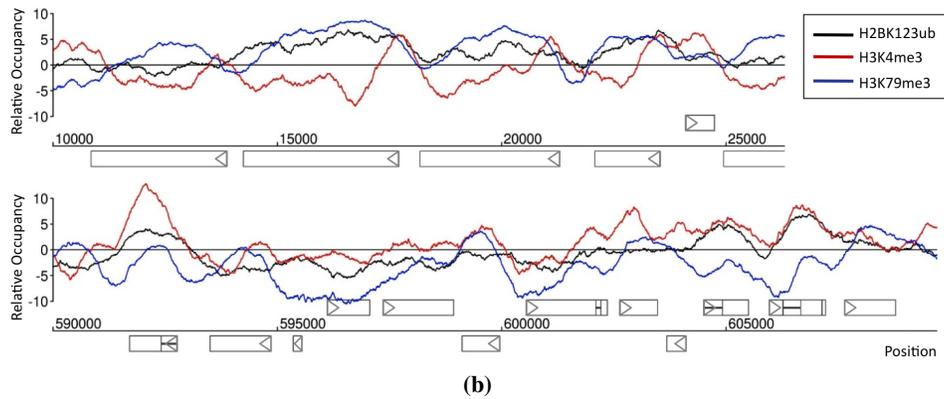
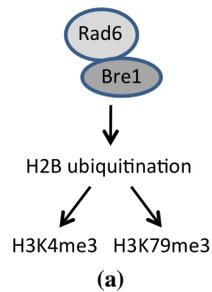


Figure 4.1: Genome-wide association of H2BK123 monoubiquitination and H3K4- and H3K79 trimethylation. (a): Schematic of H2BK123 ubiquitin cross-regulatory pathway in *S. cerevisiae*. Histone H2B lysine 123 is ubiquitinated by Rad6/Bre1 and leads to a downstream trimethylation of histone H3 lysine 4 and 79. (b): Overlay of H2BK123ub, H3K4me3 and H3K79me3 ChIP-on-chip profiles. Sample genomic positions were plotted along the *x*-axis against the relative occupancy of H2BK123ub (black), H3K4me3 (red) and H3K79me3 (blue) on the *y*-axis. ORFs are indicated as rectangles, above and below the axis for Watson and Crick genes, respectively. ARS are shown as filled grey boxes.

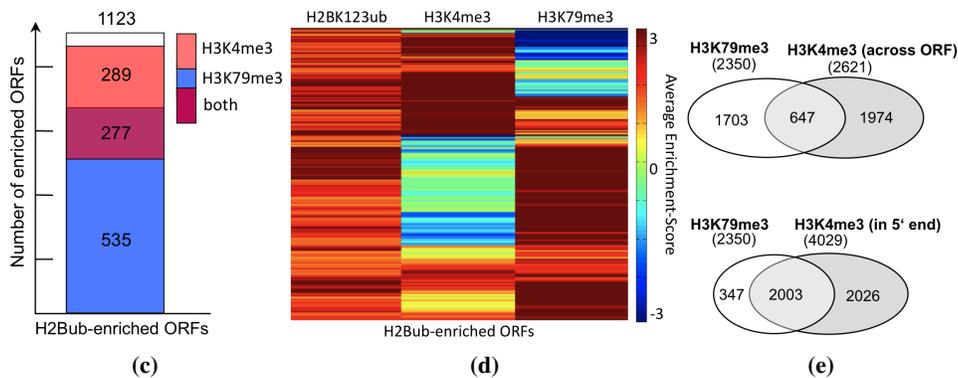


Figure 4.1: Genome-wide association of H2BK123 monoubiquitination and H3K4- and H3K79 trimethylation. (c): All H2B monoubiquitinated ORFs were associated with either H3K4- or H3K79 trimethylation. Diagram illustrating the overlap of H2BK123ub enriched ORFs with H3K4me3 and H3K79me3 enriched ORFs. (d): ORFs enriched for H2BK123ub and their association with H3K4me3, and H3K79me3. Heatmap colour-coding the average enrichment scores of H3K4me3 and H3K79me3 for all H2BK123ub-enriched ORFs. Each row corresponds to an ORF enriched for H2BK123ub, and each column to a particular modification. Rows were hierarchically clustered and plotted. (e): Almost all H3K79me3 enriched ORFs were marked by H3K4me3 in their 5' end. Venn diagrams comparing the number of H3K4me3 and H3K79me3 enriched ORFs as well as the number of H3K4me3 enriched 5' ends with H3K79me3 enriched ORFs.

enriched ORFs were compared to find co-occurrences of H2BK123ub, H3K4me3, and H3K79me3. Genome-wide, most ORFs enriched for H2BK123ub were also marked with H3K4me3, H3K79me3 or both (Figure 4.1 (c), page 138), confirming that these marks were tightly linked.

To further dissect this spatial relationship, average enrichment scores for H3K4me3 and H3K79me3 in H2BK123ub-enriched ORFs were calculated and hierarchically clustered, revealing a clear co-localization of H2BK123ub with H3K4me3, H3K79me3 or both (Figure 4.1 (d), page 138). Genome-wide, H3K4me3 and H3K79me3 were mainly associated with different groups of ORFs (Figure 4.1 (e), page 138).

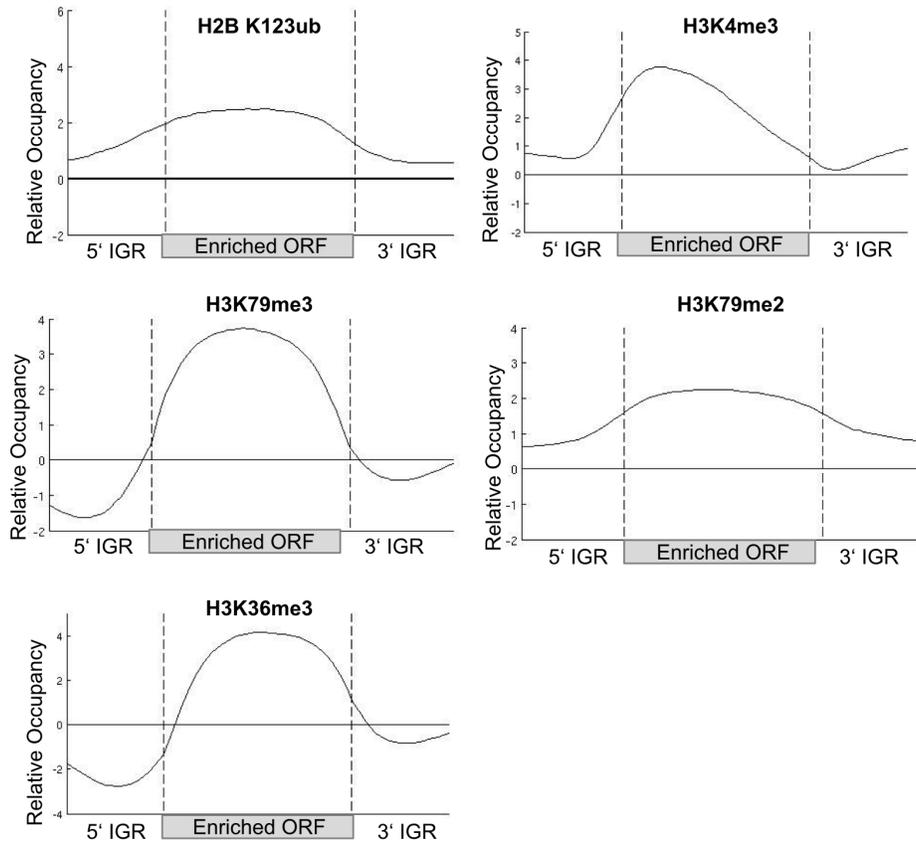


Figure 4.2: Average profiles of H2BK123ub, H3K4me3, H3K36me3, H3K79me3, and H3K79me2 enriched ORFs. H3K4me3 and H3K36me3 peaked in the 5'end and 3'end of genes, respectively. H2BK123ub, H3K79me3, and H3K79me2 were enriched across the ORF. A gene was considered enriched if at least 50% of its ORF was covered by the modification. ORFs were aligned according to their translational start and stop sites similar to a previous approach (Pokholok et al., 2005). Each ORF was divided into 40 bins of equal length, probes were assigned accordingly, and average enrichment values were calculated for each bin. Probes in promoter regions (500 bp upstream of transcriptional start site) and 3'UTR (500 bp downstream of stop site) were assigned to 20 bins, respectively. The average enrichment value for each bin was plotted.

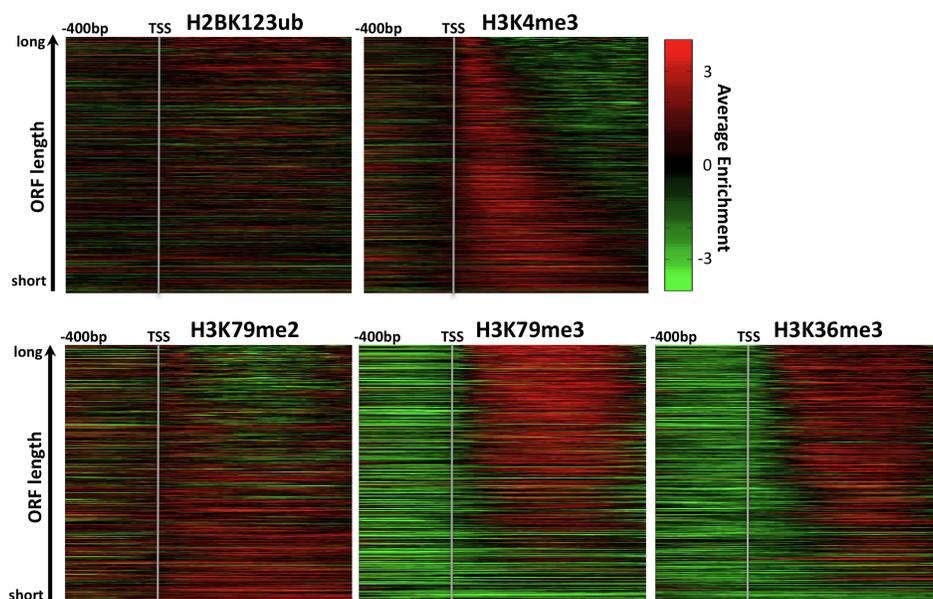


Figure 4.3: Enrichment of histone modifications across all genes. Heatmap of H2BK123ub, H3K4me3, H3K36me3, H3K79me3, and H3K79me2 occupancy across genes sorted by their length and aligned by their TSS. The normalized ChIP-on-chip signals were binned into 20 bp increments for promoter regions from the TSS to 400 bp upstream. The same was done for genic regions except that transcripts were divided into 40 equal sized bins (and thus represent a percentage of the gene length, rather than absolute distance). The average enrichment value for each bin was colour-coded and plotted. The list of transcription start end sites for 5015 genes was derived from Lee et al., 2007b.

However, as discussed in more detail below, H3K4me3 was enriched in the 5' end of two-thirds of all genes, and was also found in the 5' end of almost all H3K79me3 marked genes (Figure 4.1 (e), page 138). To illustrate the distribution of the tested histone modifications, all genes were divided into an equal number of bins, averaged and colour-coded for their enrichment score per bin and aligned by their transcription start site (Figure 4.3, page 140). When sorted by gene length, it becomes apparent that H3K4me3 covers the 5' end of all genes. This demonstrates the limitations of ORF-based analysis and emphasizes the necessity to divide genes into segments when determining enrichment as it is shown below.

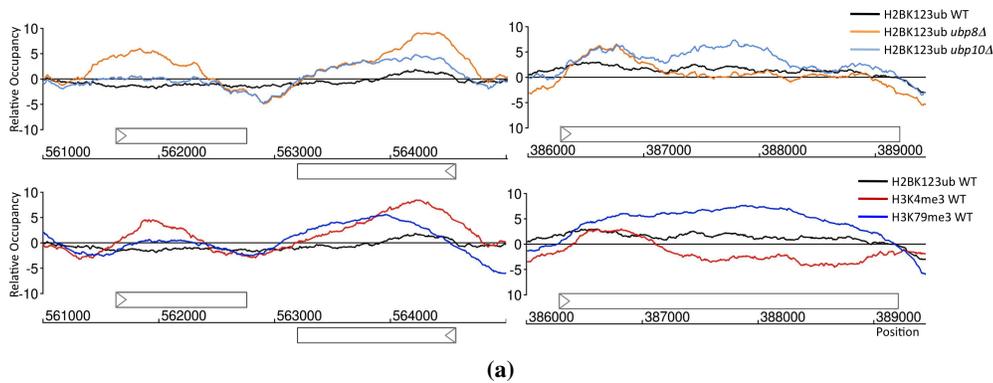


Figure 4.4: Site-specific removal of H2BK123ub by Ubp8 and Ubp10. (a): Overlay of H2BK123ub ChIP-on-chip profiles in wild-type, *ubp8*, and *ubp10* deletion strains. In the upper panel, sample genomic regions were plotted along the *x*-axis against the relative occupancy of H2BK123ub in wild-type (black), *ubp8* deletion (orange), and *ubp10* deletion (light blue) on the *y*-axis. For comparison, profiles of H3K4me3 (red) and H3K79me3 (blue) are shown in the lower panel. ORFs are indicated as rectangles above and below the *x*-axis for Watson and Crick genes, respectively.

4.2.2 Site-specific removal of H2BK123ub by its deubiquitinases Ubp8 and Ubp10

H2BK123ub is a highly transient chromatin modification, which is removed by the ubiquitin proteases Ubp8 and Ubp10. To test if these two enzymes are involved in the removal of ubiquitin from similar or distinct genomic regions, we established genome-wide maps of H2BK123ub in strains lacking either Ubp8 or Ubp10. As expected and shown for illustrative purposes in Figure 4.4 (a), page 141, levels of H2BK123ub increased at many genomic sites upon removal of its deubiquitinases. To qualitatively analyze this increase across the entire genome, ORFs enriched for H2BK123ub in wild type and the two deletion strains were compared (Figure 4.4 (b), page 142).

Strikingly, deletion of either Ubp8 or Ubp10 caused H2BK123ub levels to increase selectively in distinct subsets of ORFs (Figure 4.4 (b), page 142), indicating that Ubp8 and Ubp10 removed H2BK123ub at different sites. Given the requirement of H2BK123ub for the establishment of both H3K4 and H3K79 trimethylation, we

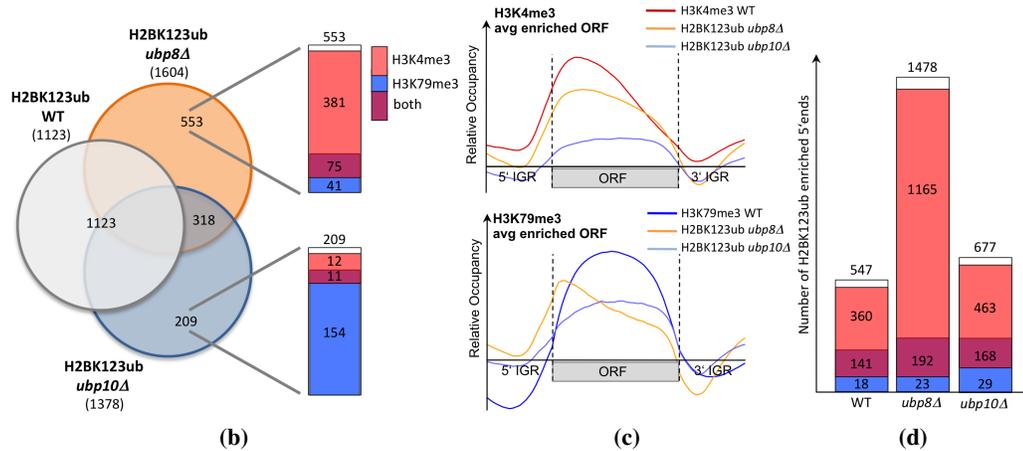


Figure 4.4: Site-specific removal of H2BK123ub by Ubp8 and Ubp10. (b): ORFs deubiquitinated by either Ubp8 or Ubp10 were strongly associated with H3K4me3 and H3K79me3, respectively. Venn diagrams comparing the number of ORFs enriched for H2BK123ub in wild-type as well as *ubp8* and *ubp10* deletion strains. For ORFs enriched for H2BK123ub only in the *ubp8* or *ubp10* deletion strain, the overlap with the number of ORFs either marked by H3K4me3 (red), H3K79me3 (blue) or both (purple) is shown. (c): Distribution of H2BK123ub in *ubp8* and *ubp10* deletion strains resembled average enrichment profile of H3K4me3 and H3K79me3, respectively. Average enrichment profiles of H2BK123ub in *ubp8* and *ubp10* deletion strains for H3K4me3 and H3K79me3 enriched ORFs. Each ORF was divided into 40 bins of equal length, probes were assigned accordingly, and average enrichment values were calculated for each bin. Probes in promoter regions (500 bp upstream of transcriptional start site) and 3'UTR (500 bp downstream of stop site) were assigned to 20 bins, respectively. The average enrichment value for each bin was plotted. (d): 5' ends of genes showed the strongest increase in H2BK123 monoubiquitination after removal of Ubp8. Diagram illustrating the overlap of H2BK123ub enriched 5' ends with H3K4me3 and H3K79me3 in wild-type, *ubp8* and *ubp10* deletion strains.

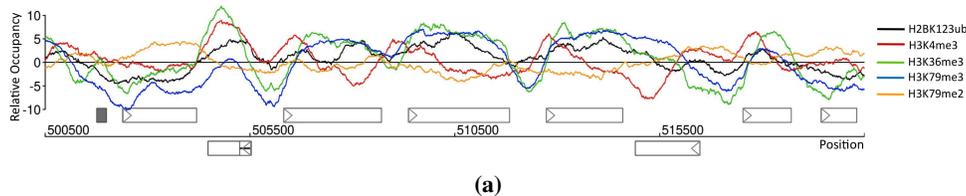


Figure 4.5: Genome-wide combinatorial pattern analysis of H2BK123ub and different H3 methylation states. (a): High-resolution profiles of examined histone modifications. Overlay of H2BK123ub (black), H3K4me3 (red), H3K36me3 (green), H3K79me3 (blue) and H3K79me2 (orange) ChIP-on-chip profiles. Sample genomic position for chromosome 8 was plotted along the x -axis against the relative occupancy of the histone modifications on the y -axis. ORFs are indicated as rectangles, above the axis for Watson genes and below the axis for Crick genes. ARS are shown as filled grey boxes.

examined ORFs exclusively deubiquitinated by either Ubp8 or Ubp10 for their association with H3K4- and H3K79 trimethylation (Figure 4.4 (b), page 142). Interestingly, and as apparent from the sample genomic regions in Figure 4.4 (a), page 141, ORFs deubiquitinated by Ubp8 were primarily marked by H3K4me3 whereas ORFs deubiquitinated by Ubp10 were mainly marked by H3K79me3 (Figure 4.4 (b), page 142). Comparison of the average distribution of H2BK123ub in the *ubp8* and *ubp10* deletion strains in all H3K4me3 and H3K79me3 enriched ORFs showed a similar trend (Figure 4.4 (c), page 142). H2BK123ub removed by Ubp8 peaked towards the 5' end of H3K4me3 enriched ORFs similar to H3K4me3 itself, whereas H2BK123ub removed by Ubp10 covered the entire H3K79me3 enriched ORFs comparable to the distribution of H3K79me3 itself. Interestingly, H2BK123ub in the *ubp8* deletion strain even peaked at the 5' end of H3K79me3 enriched ORFs (Figure 4.4 (c), page 142).

As mentioned above, most H3K79me3 enriched ORFs had H3K4me3 in their 5' end, but were not considered enriched for H3K4me3 since this mark only covered part of the ORF. Therefore, we specifically focused the analysis on the 5' end of genes, and found a strong increase of H2BK123ub specifically in the *ubp8* deletion strain that strongly associated with H3K4me3 (Figure 4.4 (d), page 142). This hints at a complexity not captured by the analysis on the ORF level, and we there-

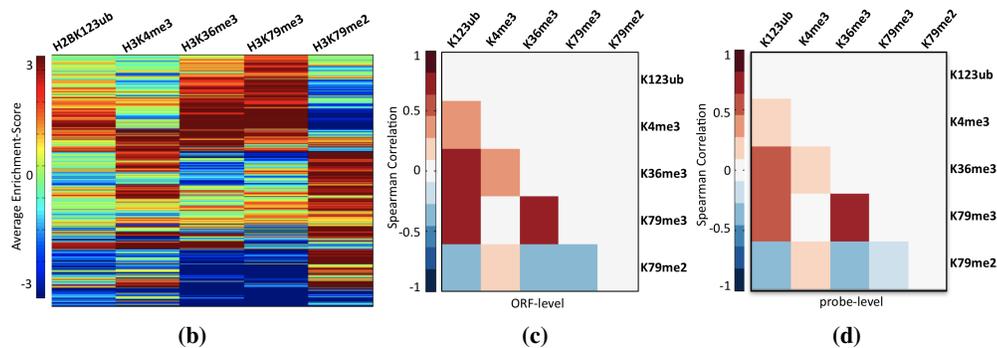


Figure 4.5: Genome-wide combinatorial pattern analysis of H2BK123ub and different H3 methylation states. (b): H2BK123ub and H3 methylation of different lysines occurred in distinct patterns. Heatmap visualizing the average enrichment of H2BK123ub, H3K4me3, H3K36me3, H3K79me3, and H3K79me2 for all known yeast ORFs. Columns represent modifications, rows correspond to ORFs. Each row colour-codes the average enrichment score for a particular ORF across all modifications on the spectrum from red indicating enrichment to blue for depletion. The modification patterns for all ORFs were hierarchically clustered and plotted. (c): Histone modifications showed different degree of correlations. Spearman correlation matrix for ChIP-on-chip profiles of individual histone modifications in all ORFs is shown, with red for high correlation and blue representing anticorrelation. (d): Genome-wide correlations were similar on a 5 bp level. Same as (c) except that Spearman correlation matrix was calculated on a 5 bp probe level.

fore describe a more detailed analysis below.

Taken together, these findings suggested that Ubp8 and Ubp10 remove ubiquitin from H2B at different chromatin neighbourhoods and are associated with the establishment of H3K4- and H3K79 trimethylation, respectively.

4.2.3 Genome-wide combinatorial pattern analysis of H2BK123ub and different H3 methylation states

Histone H3K36 trimethylation (H3K36me3) is a key chromatin modification associated with transcription elongation, a function also attributed to

H2BK123ub (Weake and Workman, 2008). We generated ChIP-on-chip profiles of H3K36 trimethylation to examine its relationship with H2BK123ub, the H2BK123ub-linked H3K4me3 and H3K79me3, as well as the independent H3K79me2 mark. As previously shown, H3K36me3 on average marked nucleosomes at the middle and 3'end of the ORF (Figure 4.2, page 139) (Pokholok et al., 2005). Figure 4.5 (a), page 143 shows a representative genomic region with profiles for all five modifications studied in this work.

To identify co-occurring modifications and associated patterns genome-wide, the average enrichment score for each mark and for all ORFs was calculated and clustered. As visualized by the heatmap, the combinatorial complexity of histone marks was more extensive than the sample region in Figure 4.5 (a), page 143, hinted at (Figure 4.5 (b), page 144). Most prominently, we found again that most of the H2BK123ub enriched ORFs were marked with H3K4me3 or H3K79me3. Interestingly, H2BK123ub also coincided with the elongation mark H3K36me3, which in addition strongly overlapped with H3K79me3. Finally, and consistent with our previously published findings Schulze et al. (2009), the clustering clearly separated ORFs enriched for H3K79me2 and H3K79me3, and showed that H3K79 tri- but not dimethylation was associated with H2BK123ub.

Table 4.1: H2B monoubiquitination and H3 methylation were mainly associated with genes transcribed by RNA polymerase II. Overlap of investigated histone modifications with different genomic features like RNA polymerase I (rRNA), II (ORFs) and III (tRNA) transcribed regions, centromeres (CENs) or autonomously replicating sequence (ARS). rRNAs, tRNAs, ARS, and CENs were called associated when 100 % of underlying probes had an enrichment score above a threshold of 1.5. ORFs were called enriched when at least 50 % underlying probes had an enrichment score above a threshold of 1.5. [†]excluding mitochondrial chromosome.

Category	Total [†]	H2BK123ub	H3K4me3	H3K36me3	H3K79me3	H3K79me2
rRNA (Pol I)	25	0	1	0	0	0
ORF (Pol II)	6567	1498	2621	2714	2350	1866
tRNA (Pol III)	275	0	4	0	0	0
ARS	274	4	9	3	3	17
CEN	16	0	1	0	0	0

To get a more quantitative picture of the relationships between histone marks, their pair-wise Spearman correlation was calculated. As expected from the heatmap, H2BK123ub correlated positively with its downstream marks H3K4me3 and H3K79me3 as well as with the elongation mark H3K36me3 (Figure 4.5 (c), page 144). Remarkably, H3K79me2 correlated negatively with all modifications except H3K4me3. Since these correlations were seen at the probe level as well, the conclusions drawn for coding regions can be extended to the entire genome (Figure 4.5 (d), page 144). Furthermore, we found that all histone modifications investigated in this study were mainly associated with genes transcribed by RNA polymerase II, and absent from genomic regions transcribed by RNA polymerase I or III (Table 4.1, page 145).

4.2.4 Specific gene regions were characterized by distinct histone H3 methylation and H2B monoubiquitination patterns

So far, our analysis was based on averaging the enrichment of each modification across the ORF, revealing their connections at the gene level. However, given that some histone modifications only cover part of the ORF (Figure 4.2, page 139), we reasoned that partitioning each gene in segments might provide a more detailed view. Therefore, we partitioned each gene into four segments: promoter (300 bp upstream of the coding start), 5'CDS (300 bp downstream of the coding start), middle of CDS (300 bp around the centre of each ORF) and 3'CDS (300 bp upstream of the coding end) and calculated the number of genes enriched for specific histone modifications for each segment (Figure 4.6 (a), page 147). Consistent with the average profile (Figure 4.2, page 139), H2BK123ub was enriched in all four regions, covering between 15% and 25% of all genes. H3K4me3 had a more skewed profile with almost 70% of all genes being enriched at the 5' end of the coding sequence, but only between 15% and 30% in the other three regions. For about 40% of all genes, H3K36me3 and H3K79me3 were found predominantly in the middle and 3' end (Figure 4.6 (a), page 147).

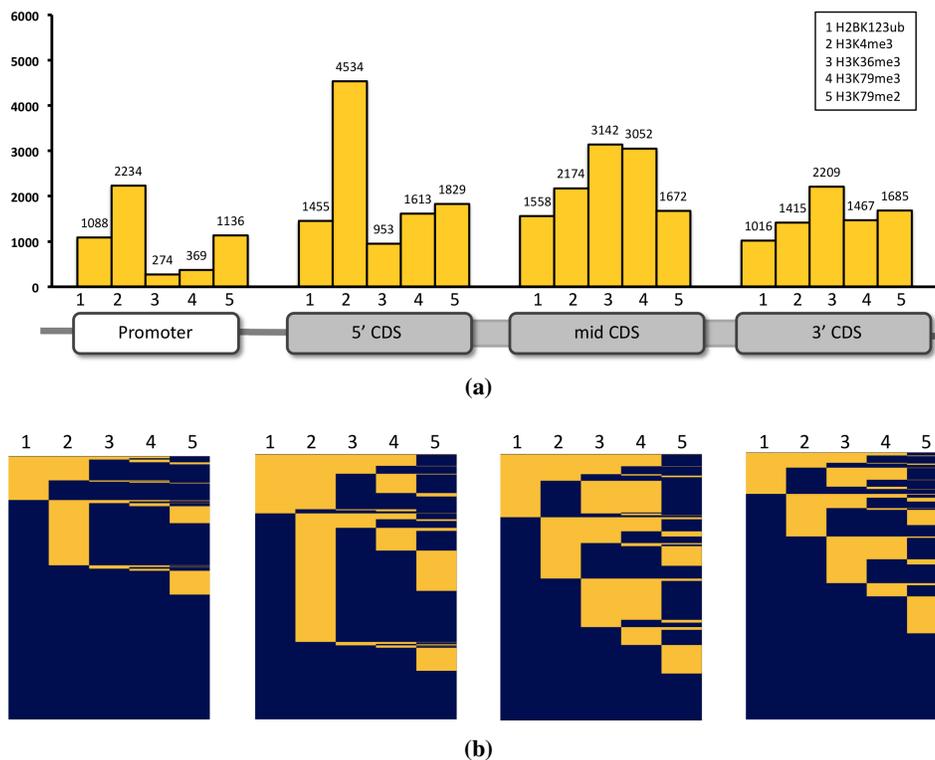
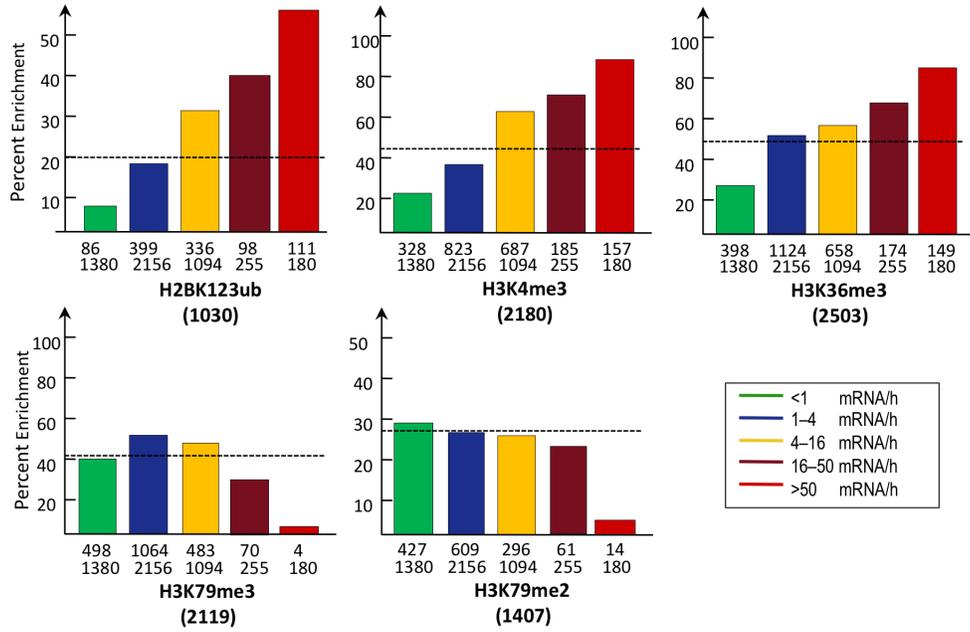


Figure 4.6: Specific gene regions were characterized by distinct histone H3 methylation and H2B monoubiquitination patterns. (a): Specific parts of a gene were differently enriched for certain modifications. Genes were separated into the following parts: promoter region (300 bp upstream of the coding region), 5' CDS (300 bp downstream CDS start), middle CDS (300 bp around centre of ORF), and 3' CDSs (300 bp upstream of CDS end). For each histone modification, the number of enriched genes was plotted on the y-axis. (b): Binary heatmap comparing histone modification patterns in different gene parts. For the four gene segments introduced in (a), each ORF was tested for its enrichment for the different histone modifications and either labeled enriched (yellow) or not enriched (blue). Columns correspond to specific modifications, rows to ORFs. Rows are independently sorted by combination of patterns.



(a)

Figure 4.7: H2BK123ub, H3K4me3, and H3K36me3 were associated with highly transcribed genes. (a): H2BK123ub, H3K4me3, and H3K36me3 were enriched in highly transcribed genes whereas H3K79me3 and H3K79me2 tended to be association with less transcriptionally active genes. ORFs enriched for different histone modification were grouped into five classes according to their transcriptional frequency (Holstege et al., 1998) (Transcriptome 2005). Overlaps of modifications and transcription classes are shown. A gene was considered enriched if at least 50 % of its ORF was covered by a modification. Dashed horizontal lines represent percent overlap expected by chance. Numbers below x -axis represent enriched ORFs and total number of ORFs in each transcription class. Numbers in parenthesis below histone modification indicate total number of enriched ORFs for that modification.

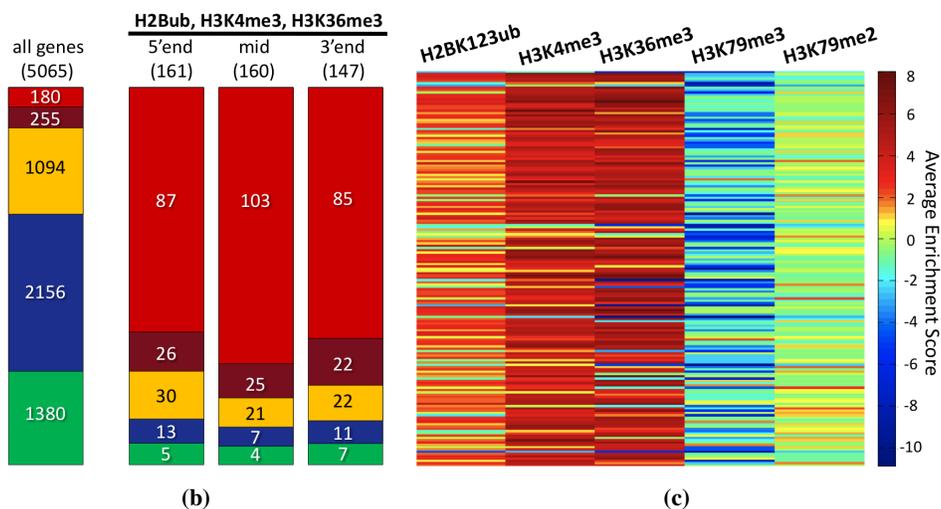


Figure 4.7: H2BK123ub, H3K4me3, and H3K36me3 were associated with highly transcribed genes. (b): Genes marked by a triple combination of H2BK123ub, H3K4me3, and H3K36me3 across their ORF were highly transcribed. Diagram comparing the distribution of transcriptional frequency in all yeast genes with the distribution for the segments of genes with the triple histone modification pattern of H2BK123ub, H3K4me3, and H3K36me3 in 5'CDS, mid-CDS, and 3'CDS, respectively. Colour-codes for the transcriptional classes are identical to the ones in (Figure 4.7 (a), page 148). Numbers indicate overlap of each sub-group of genes with the different transcriptional classes. (c): Almost all highly transcribed genes tended to be enriched for H2BK123ub, H3K4me3, and H3K36me3. For 180 genes with a transcriptional frequency of more than 50 transcripts per hour, the average enrichment for different histone modifications was determined and plotted as heatmap. Rows represent genes and columns enrichment for histone modifications. Rows were sorted by transcriptional frequency of the associated genes.

Table 4.2: Histone modification patterns occurred with different frequencies. For all possible histone modification patterns of H2BK123ub, H3K4me3, H3K36me3, H3K79me3, and H3K79me2 in promoter, 5'CDS, middle CDS and 3'CDS the actual number of genes with each pattern is specified. Groups for which a significant overlap with transcription factor binding was found are labeled with an asterisk.

Profile	Prom	5'CDS	mid CDS	3'CDS
no modification	3117	1230	1146	2128
H2BK123ub	*413	49	45	114
H3K4me3	*1060	*1267	320	284
H3K36me3	32	24	*112	*460
H3K79me3	46	48	*379	*275
H3K79me2	*599	*566	*711	*909
H2BK123ub, H3K4me3	*402	*315	81	67
H2BK123ub, H3K36me3	10	6	19	*168
H3K4me3, H3K36me3	33	*176	*164	*162
H3K4me3, H3K79me3	*70	*482	47	30
H3K4me3, H3K79me2	*423	*996	*480	*373
H3K36me3, H3K79me3	36	43	*959	533
H3K36me3, H3K79me2	7	17	*71	*125
H2BK123ub, H3K4me3, H3K36me3	*35	*184	*180	*162
H2BK123ub, H3K4me3, H3K79me3	*69	*468	*34	18
H2BK123ub, H3K4me3, H3K79me2	44	*79	*47	39
H2BK123ub, H3K36me3, H3K79me3	24	24	786	*290
H3K4me3, H3K36me3, H3K79me3	31	141	*298	79
H3K4me3, H3K36me3, H3K79me2	7	*37	*128	91
H2BK123ub, H3K4me3, H3K36me3, H3K79me3	41	*272	*288	*67

After having examined each modification separately, we determined their combinations within each gene segment by assigning a binary value for each modification depending on whether it was present or not within a segment, thereby allowing us to reduce the complexity of the analysis. Out of the 32 combinations possible for five histone methylation and ubiquitination marks per segment, some occurred frequently whereas others were found only on very few genes (Figure 4.6 (b), page 147, and Table 4.2, page 150). Promoters were mostly enriched for just one modification, including H2BK123ub ($n = 413$), H3K4me3 ($n = 1060$), or

H3K79me2 ($n = 599$), and contained only two combinations of marks H2BK123ub and H3K4me3 ($n = 402$), and H3K4me3 and H3K79me2 ($n = 423$) (Table 4.2, page 150). While H3K4me3 was the most dominant histone modification present at the 5' end of genes, it frequently was present together with other marks, including the triple combinations of H2BK123ub, H3K4me3, and H3K36me3 ($n = 184$) or H2BK123ub, H3K4me3, and H3K79me3 ($n = 468$). Furthermore, a group of genes was characterized by the presence of four marks, those being H2BK123ub, H3K4me3, H3K36me3, and H3K79me3 ($n = 300$). A similar prevalence of specific modifications combinations thereof was also found in the middle segment and 3' end of ORFs (Table 4.2, page 150).

4.2.5 H2BK123ub, H3K4me3, and H3K36me3 were associated with highly transcribed genes

In order to examine the correlation between H2BK123ub and different H3 methylations with gene expression, we determined the occurrence of each modification in five classes of genes sorted by their transcriptional frequency (Holstege et al., 1998). The presence of H2BK123ub, H3K4me3, or H3K36me3 individually in ORFs correlated with higher transcriptional frequencies whereas H3K79me2 as well as H3K79me3 did not significantly correlate with any particular transcriptional frequency, consistent with previous findings (Minsky et al., 2008; Pokholok et al., 2005; Schulze et al., 2009) (Figure 4.7 (a), page 148).

Focusing on the dominant histone mark combinations described above, we found that ORFs belonging to the group of highly transcribed genes were significantly enriched for the triple combination of H2BK123ub, H3K4me3, and H3K36me3, whereas this combination was almost completely absent in the four other transcriptional classes (Figure 4.7 (b), page 149). Surprisingly, this particular triple combination was present across the entire ORF in these genes (Figure 4.7 (b), page 149), despite the general trend of H3K4me3 to occupy the 5' end and H3K36me3 to occupy the 3' end of genes. To get a detailed picture of the histone mark combinations in highly transcribed genes, we colour-coded all enrichment scores for this relatively small group of 180 genes (Figure 4.7 (c), page 149). As obvious from the heatmap, H2BK123ub, H3K4me3, and H3K36me3 were present together at

these genes whereas H3K79 di- and trimethylation were almost completely absent (Figure 4.7 (c), page 149).

To further investigate the relation between chromatin signatures and transcriptional regulation, we tested whether certain patterns of histone modifications derived from our combinatorial analysis were correlated with binding by specific transcription factors. We applied a hypergeometric test to determine all chromatin patterns that were significantly associated with transcription factors, derived from a previously published catalogue (MacIsaac et al., 2006). Using this approach, many statistically significant associations were identified for each gene segment and its associated chromatin marks, summarized in Table 4.3, page 156. Several examples illustrated that this list likely reflected connections of biological relevance. First, the transcription factor Fhl1, a key transcriptional regulator of genes encoding ribosomal proteins, was connected well with the triple combination of H2BK123ub, H3K4me3, and H3K36me3 and highly transcribed genes (Rudra et al., 2005). Second, whereas we previously reported that the transcription factor Swi4 binds to H3K79 dimethylated genes (Schulze et al., 2009), our combinatorial approach now enabled us to further dissect this connection to reveal that H3K79me2 was associated with Swi4 in combination with additional histone marks (Table 4.3, page 156).

4.2.6 Functional genomic analysis of intron-containing genes

Given that in *S. cerevisiae*, around half of all highly transcribed genes belong to the small group of intron-containing genes, we analyzed the histone modification patterns of this particular class of genes. Interestingly, intron-containing genes were significantly enriched for H2BK123ub, H3K4me3, and H3K36me3 but lacked H3K79 di- and trimethylation (Figure 4.8 (a), page 153). These specific histone modifications were primarily associated with the highly transcribed fraction of all intron-containing genes, as shown in the heatmap sorted by transcriptional frequency (Figure 4.8 (b), page 153).

To test for a causal relationship of specific histone modifications in splicing we measured splicing efficiency in *bre1*, *set1*, and *set2* deletion strains, thereby eliminating H2BK123 ubiquitination, H3K4 methylation and H3K36 methylation re-

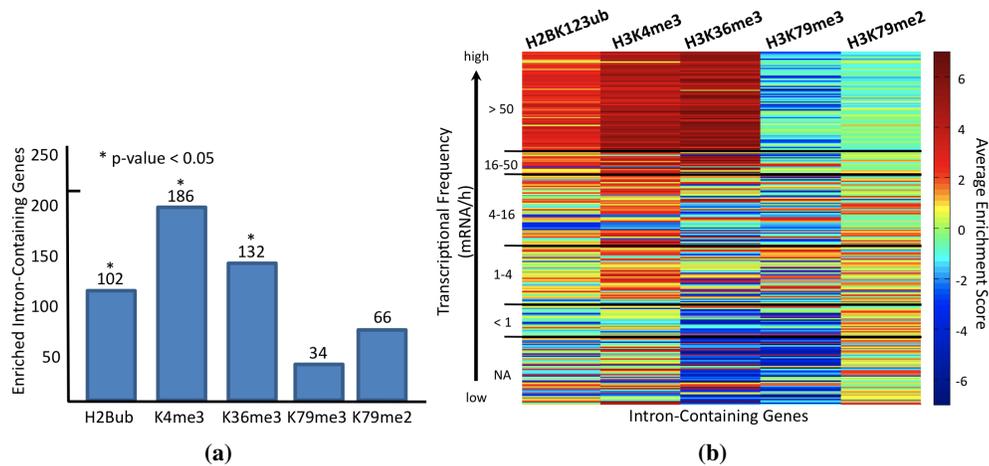


Figure 4.8: Intron-containing genes were significantly enriched for co-occurrence of H2BK123ub, H3K4me3, and H3K36me3. (a): Intron-containing genes were significantly enriched for H2BK123ub, H3K4me3, and H3K36me3 but not H3K79 di- or trimethylation. Diagram showing the number of intron-containing genes enriched for the different histone modifications. (b): H2BK123ub, H3K4me3, and H3K36me3 together mark highly transcribed intron-containing genes. For all intron-containing genes average ORF enrichment scores (intron and exon) for different histone modifications were calculated and plotted as a heatmap. Rows represent genes and columns enrichment for histone modifications. Rows were sorted by transcriptional frequency of the associated genes.

spectively from chromatin. To this end, we compared the ratio of pre- and mature mRNA in spliced genes derived from cDNA from each strain hybridized to high-resolution tiling arrays, using a strategy developed by others (Huber et al., 2006). Strikingly, deletion of *bre1* caused a small but reproducible increase of intronic RNA without changing the transcript level of the exon, indicating an accumulation of unspliced pre-mRNA (Figure 4.8 (c), page 154). To compare splicing efficiencies globally, the ratios of intronic and exonic regions in mutant strains versus wild-type were sorted by transcriptional frequency (Figure 4.8 (d), page 154). As shown in the heatmap, loss of the H2BK123 monoubiquitin ligase Bre1 compromised splicing mainly in highly transcribed genes especially the ones encoding

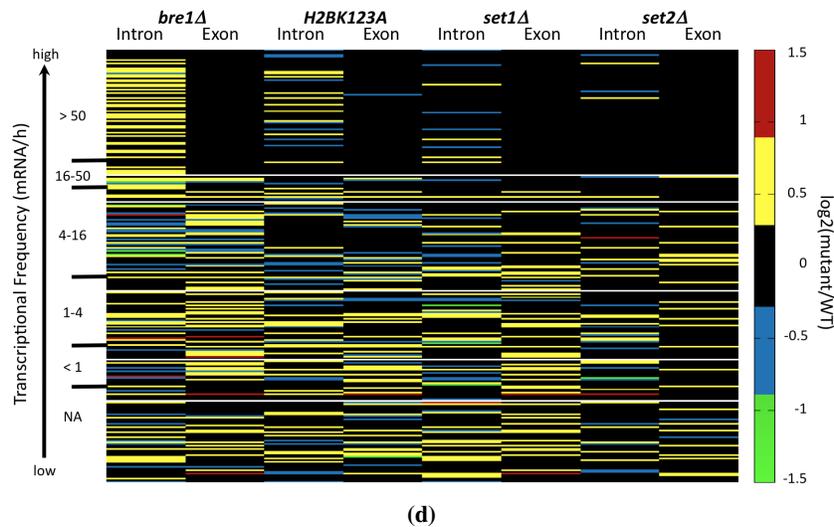
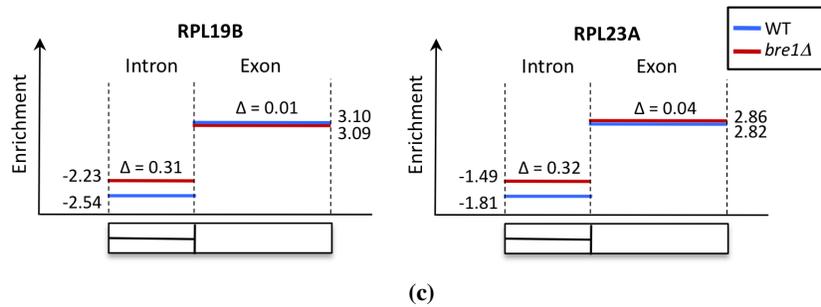


Figure 4.8: Intron-containing genes were significantly enriched for co-occurrence of H2BK123ub, H3K4me3, and H3K36me3. (c): Two representative examples of ribosomal genes showing splicing defects in *bre1* deletion strain. For wild-type and *bre1* deletion strains the average expression level for introns and exons are illustrated. While the expression levels of introns in *bre1* deletion increased slightly, the exons expression levels remained unaltered. Illustration is non-proportional. (d): Heatmap comparing splicing changes in *bre1*, *set1*, *set2* deletion and *H2BK123A* mutant strains. For all spliced genes \log_2 ratios of mutant over wild-type were calculated for introns and exons, respectively. Ratios were colour-coded in a heatmap, in which rows represent intron-containing genes sorted by their transcriptional frequency, and columns represent ratios for each mutant over wild-type.

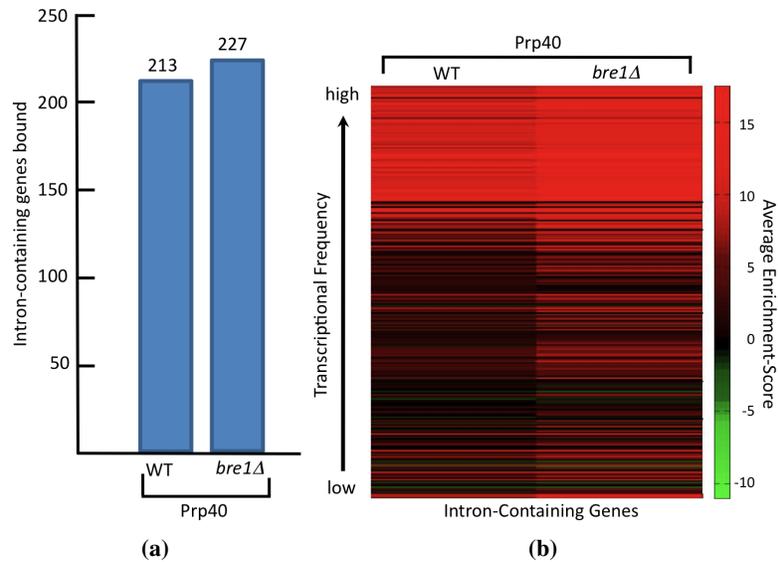


Figure 4.9: Spliceosome recruitment to chromatin was not affected in *bre1* deletion strain. (a): Binding of spliceosome subunit Prp40 to intron-containing genes was not altered in a *bre1* deletion strain. Numbers of spliced genes bound by Prp40 as measured by ChIP-on-chip is compared between wild-type and *bre1* deletion strain. (b): Intron-containing genes were not affected for Prp40 binding to chromatin in *bre1* deletion strain. Heatmap comparing binding of the spliceosome subunit Prp40 to spliced genes in wild-type and *bre1* deletions strains. Rows represent spliced genes and were sorted by their transcriptional frequency.

ribosomal proteins, consistent with H2BK123ub preferential location to the same set of genes. In contrast, the *H2BK123A*, *set1Δ*, and *set2Δ* strains had almost no measurable defects on splicing, indicating that Bre1 might affect splicing of highly transcribed genes independently of its role in regulating H2BK123ub or H3K4me3. Furthermore, the defects in pre-mRNA splicing caused by loss of *bre1* were likely due to spliceosome-recruitment independent steps as the association of spliceosome subunit Prp40 with intron-containing genes was similar in wild-type and *bre1Δ* strain (Figure 4.9, page 155).

Table 4.3: Combinations of histone modifications were associated with genes bound by certain transcription factors. For all histone modification patterns in promoters, 5'CDS, middle CDS and 3'CDS their overlap with transcription factor bound genes was calculated. Listed are those transcription factors with significant occurrences ($p < 0.01$).

Profile	Transcription Factors
Promoter	
H2BK123ub H3K4me3	REB1, ABF1, SFP1, FHL1, RAP1 CIN5, PHD1, HAC1, XBP1, NRG1, SKN7
H3K79me2	FKH2, ABF1, RAP1
H2BK123ub, H3K4me3	RPN4, CIN5, REB1, ABF1, MSN2, MSN4, SOK2, UME6, MOT3
H2BK123ub, H3K79me3 H3K4me3, H3K79me2	REB1 CIN5, REB1, FKH2, ABF1, MBP1, ROX1, MSN2, SOK2, FHL1, SWI5, SWI4, UME6, RAP1, SKN7
H2BK123ub, H3K4me3, H3K79me3	ABF1, ADR1
H2BK123ub, H3K4me3, H3K36me3	FHL1, RAP1
5'CDS	
H3K4me3 H3K79me2	PHO2, MBP1, YHP1, HAP3 ABF1, ROX1, SOK2, YAP6, YAP7, NRG1
H2BK123ub, H3K4me3 H3K4me3, H3K36me3	RPN4, REB1, ABF1 CIN5, GCR1, GCR2, GLN3, ACE2, MSN2, FHL1, NDD1, SWI4, SWI6, STE12, UME6, HAP5, RAP1, MOT3
H3K4me3, H3K79me3	SOK2, FHL1, YAP6, SWI4, RAP1
H3K4me3, H3K79me2	FHL1
H2BK123ub, H3K4me3, H3K36me3	SFP1, GAT3, FHL1, YAP5, RAP1
H2BK123ub, H3K4me3, H3K79me3	REB1, ABF1, FHL1, AFT2, SWI4, NRG1, SKN7
H2BK123ub, H3K4me3, H3K79me2	ABF1
H3K4me3, H3K36me3, H3K79me2	FKH2, INO2, SWI4
H2BK123ub, H3K4me3, H3K36me3, H3K79me3	RPN4, CIN5, REB1, CBF1, ABF1

Table 4.3 – continued

Profile	Transcription Factors
Middle CDS	
H3K36me3	MBP1
H3K79me2	CIN5, REB1, ABF1, STB5, YAP6, XBP1, NRG1, SKN7
H3K4me3, H3K36me3	GCR1 GCR2, FHL1
H3K4me3, H3K79me2	FHL1
H3K36me3, H3K79me3	RTG3, FKH2, FHL1, MCM1
H3K36me3, H3K79me2	SWI4
H2BK123ub, H3K4me3, H3K36me3	SFP1, GAT3, FHL1, YAP5, HAP5, RAP1
H2BK123ub, H3K4me3, H3K79me3	REB1, ABF1, ROX1, MSN2, SOK2, YAP6, SKN7
H2BK123ub, H3K4me3, H3K79me2	ABF1
H3K4me3, H3K36me3, H3K79me3	BAS1, LEU3
H3K4me3, H3K36me3, H3K79me2	SWI4, SWI6
H2BK123ub, H3K4me3, H3K36me3, H3K79me3	RPN4, ABF1, NRG1
3'CDS	
H3K36me3	GCN4
H3K79me3	FHL1
H3K79me2	ABF1, RAP1
H2BK123ub, H3K36me3	ABF1
H3K4me3, H3K36me3	GCR1, SFP1, FHL1, RAP1
H3K4me3, H3K79me2	ABF1
H3K36me3, H3K79me2	CIN5
H2BK123ub, H3K4me3, H3K36me3	SFP1, GAT3, FHL1, YAP5, MET31, HAP5, RAP1
H2BK123ub, H3K36me3, H3K79me3	ABF1
H3K4me3, H3K36me3, H3K79me3	SPT2
H2BK123ub, H3K4me3, H3K36me3, H3K79me3	ABF1

4.3 Discussion

Histone proteins are subject to a striking variety of covalent modifications, with some of those marks being linked by cross-regulatory networks. For example, H2BK123ub, a mark with roles in transcription initiation and elongation, is re-

quired via a 'trans-tail' process for trimethylation of H3K4 and H3K79 (Briggs et al., 2002; Dover et al., 2002; Nakanishi et al., 2009; Schulze et al., 2009; Sun and Allis, 2002). In this study, we utilized high-resolution maps of H2BK123ub and its downstream marks H3K4me3 and H3K79me3 to reveal a tight linkage of these marks on a genome-wide scale. Importantly, the histone deubiquitinases Ubp8 and Ubp10 removed H2BK123ub at different chromatin neighbourhoods and were linked to H3K4- and H3K79 trimethylation, respectively. Furthermore, H2B monoubiquitin frequently co-localized with distinct histone H3 methylation patterns, which were associated with transcription-factor binding, transcriptional frequencies, and intron-containing genes. Consistent with the latter, the H2B monoubiquitin ligase Bre1 was required for proper mRNA splicing.

It was not until almost thirty years after the discovery of monoubiquitination of H2B and H2A in transcriptionally active chromatin (Davie and Murphy, 1990; Levinger and Varshavsky, 1982; Nickel et al., 1989) that the dynamics of H2B monoubiquitination during transcription initiation and elongation started to emerge (Shilatifard, 2006; Weake and Workman, 2008).

As monoubiquitinated H2B is a highly transient histone mark that is rapidly turned over after triggering the establishment of the more stable H3K4 or H3K79 trimethylation marks, each nucleosome carrying those marks should have previously been marked by H2B monoubiquitination. Consistent with this assumption, we found a strong overlap of H3K4- and H3K79 trimethylated regions with monoubiquitinated H2BK123.

The fact that the overlap we see is not complete is consistent with the transient nature of the monoubiquitin mark and hence implies that we should detect mainly those regions that are often monoubiquitinated. In agreement with this hypothesis, H2BK123ub was frequently found in highly transcribed genes. Extending this concept, out of all H3K4me3 marked genes, we found those that also carried H2BK123ub to be the most highly transcribed ones (Figure 4.10, page 159). The connection to transcription was evident even for H3K79 trimethylated genes, which usually have lower transcriptional rates, as those that were also enriched for H2BK123ub belonged to the more frequently transcribed ones in that group (Figure 4.10, page 159).

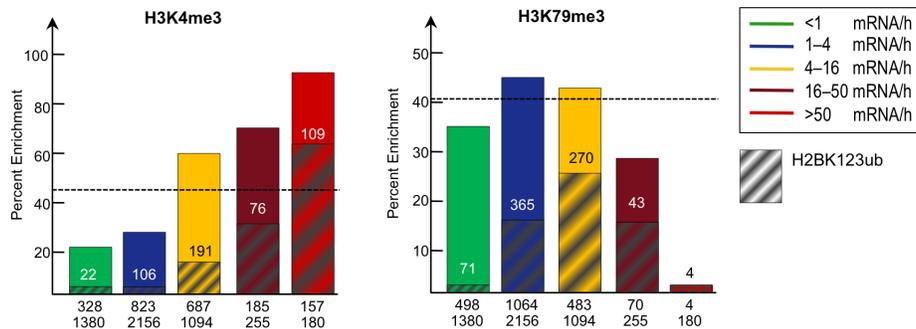


Figure 4.10: H2BK123ub overlapped strongest with H3K4me3 and H3K79me3 in ORFs with higher transcriptional frequencies. As in Figure 4.7 (a), page 148, ORFs enriched for H3K4me3 and H3K79me3 were grouped into five classes according to their transcriptional frequency (Holstege et al., 1998) (Transcriptome 2005). In addition, the sections of these ORFs also marked by H2BK123ub are indicated as gray dashed boxes.

Ubp8 and Ubp10 are the known ubiquitin proteases that both remove monoubiquitin from histone H2B. Consistent with data from protein blotting experiments and single gene studies, we found a significant increase in the number of regions marked by H2B ubiquitin when either Ubp8 or Ubp10 is lacking. Importantly, our data suggested that Ubp8 and Ubp10 act upon separate pools of ubiquitinated H2B, rather than regulating H2B monoubiquitination within the same chromatin regions. Specifically, whereas Ubp8 removed monoubiquitinated H2B predominantly from genes that were H3K4 trimethylated, Ubp10 did so from genes that were H3K79 trimethylated.

These observations suggest that Ubp8 and Ubp10 alone or within their complexes can recognize H3K4 and H3K79 trimethylation, respectively, and catalyze deubiquitinations on specific genomic loci. This molecular insight demonstrates the specific regulatory role associated with the deubiquitination machinery at different regions of chromatin and also is in line with previous findings that the DUBs are the major molecular regulators within eukaryotic cells (D'Andrea, 2010).

Our previous studies demonstrated that there are two modes (monoubiquitination-dependent and monoubiquitination-independent) of regulation for H3K4 trimethy-

lation by Set1/COMPASS in yeast (Nakanishi et al., 2009). Not only we identified factors and sites within histone to be required for H2B monoubiquitination and H3K4 methylation, we also identified residues such as H2BR119 when mutated resulted in an increase in H2B monoubiquitination but a specific loss in H3K4 trimethylation with no effect on H3K79 methylation (Nakanishi et al., 2009). These observations are in line with our report here indicating separate functions of Ubp8 and Ubp10 in establishing H3K4- and H3K79 trimethylation, respectively. In further agreement with our data, H3K4 trimethylation of the *GALI-10* UAS is reduced during gene activation in a *ubp8* deletion strain (Daniel et al., 2004). In contrast to Ubp8, less is known about the biological role H2B deubiquitination by Ubp10. Ubp10 has a role in telomeric silencing (Emre et al., 2005; Gardner et al., 2005) and it was initially found in a screen for high copy disruptors of this process and therefore first named Dot4 (Singer et al., 1998). Interestingly, the H3K79 methyltransferase Dot1 was found in the same screen indicating a potential connection of the H2B ubiquitin protease Ubp10 and Dot1. Here, we propose that Ubp10 and Dot1 might be acting together not only at telomeres but more profoundly on a genome-wide scale. Specifically, Ubp10 removed H2B monoubiquitin from genes that are H3K79 trimethylated, which tend to be those with longer ORFs (Schulze et al., 2009).

Integrating our findings with the existing literature, we propose the following model that distinguishes the roles of H2BK123ub in transcription initiation and elongation (Figure 4.11, page 161). Initially, Rad6/Bre1 get recruited to promoters through interactions with activators and monoubiquitinate H2B. H2B ubiquitin and surrounding residues, in turn, provide a molecular ‘tag’ attracting Cps35, a subunit of the Set1/COMPASS complex, to trimethylate H3K4me3 (Lee et al., 2007a). Finally, Ubp8, as part of the SAGA complex (Köhler et al., 2010; Samara et al., 2010), removes the bulky H2B monoubiquitin group and H3K4me3 serves as a memory mark for recent transcriptional initiation (Gerber and Shilatifard, 2003; Krogan et al., 2003; Muramoto et al., 2010; Ng et al., 2003).

After initiation, Rad6/Bre1 stays associated with the elongating form of RNA polymerase II and monoubiquitinates H2B throughout the ORF. This provides a molecular tag recognized by Dot1, which is stimulated to specifically trimethylate

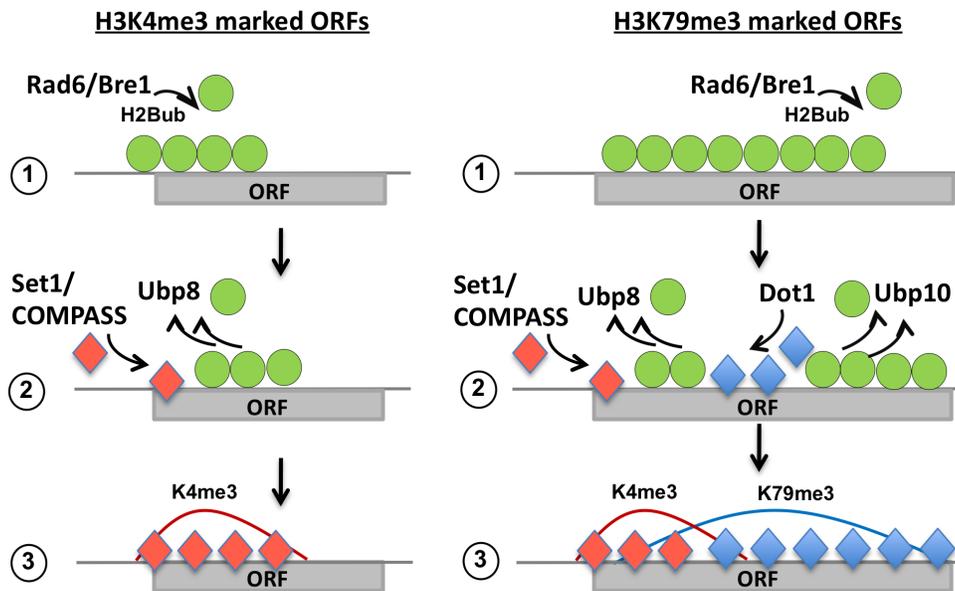


Figure 4.11: Model depicting the circuitry of H2BK123 monoubiquitination and its downstream marks H3K4me3 and H3K79me3. In the left panel, H3K4me3 marked genes: first, H2B is monoubiquitinated by Rad6/Bre1 resulting in the recruitment of Set1/COMPASS to trimethylate H3K4me3. In the last step, H2BK123ub is removed by Ubp8. In the right panel, H3K79me3 marked genes: first, H2B is monoubiquitinated by Rad6/Bre1 resulting in the recruitment of Dot1 to trimethylate H3K79me3. In the last step, H2BK123ub is removed by Ubp10. Note that most ORFs enriched for H3K79me3 are also marked by H3K4me3 in their 5' end, which is established similar to the process depicted in the left panel.

H3K79 (Jeltsch and Rathert, 2008). Owing to its transient nature, the H2B ubiquitin mark is then removed by Ubp10.

In agreement with this model, monoubiquitinated H2B plays an important role during transcription elongation. Not only are the transcription elongation complexes PAF and BUR required for the establishment of H2B ubiquitination (Krogan et al., 2003; Wood et al., 2003b; 2005), but monoubiquitinated H2B also functions cooperatively with the FACT subunit Spt16 in nucleosome reassembly during transcription elongation (Fleming et al., 2008), likely through stabilizing nucleo-

somes (Chandrasekharan et al., 2009). Furthermore, even though H2B ubiquitin is not required for H3K36me3 establishment, we found a strong genome-wide correlation between these marks. This is interesting given that H3K36 trimethylation acts as a key signal for nucleosome reassembly through the recruitment of the Rpd3S histone deacetylase complex (Carrozza et al., 2005; Joshi and Struhl, 2005; Li et al., 2007).

Lastly, we found almost all H3K79 trimethylated genes to also be marked by H3K36me3. This raises the possibility that H2BK123ub, along with H3K36- and H3K79 trimethylation, serves as an important chromatin mark for long genes, as these depend on an extensive re-assembly of chromatin during transcription elongation.

In contrast, in the small group of the most highly transcribed genes, monoubiquitinated H2B was localized in a typical combination with H3K4- and H3K36 trimethylation. Genes belonging to this group are relatively short and often contain introns thus undergoing mRNA splicing. Similar to other eukaryotes, splicing is a crucial mRNA processing step in yeast even if only a minor fraction of the budding yeast genes contain introns. However, their transcripts constitute more than a quarter of the cellular mRNA pool since many encode highly transcribed ribosomal proteins (Ares et al., 1999). Splicing, like other RNA-processing reactions, occurs co-transcriptionally in a chromatin context although the relationship between chromatin structure and pre-mRNA processing is poorly understood (Perales and Bentley, 2009). In human cells, H3K4 trimethylation facilitates pre-mRNA maturation via bridging of the spliceosomal components and Chd1 to actively transcribed genes (Sims et al., 2007). Furthermore, H3K36me3 is associated with exons (Spies et al., 2009) and has a causal role in the regulation of alternative splicing (Luco et al., 2010).

Here, we demonstrated that despite the significant association of H2B monoubiquitination, H3K4- and H3K36 trimethylation with intron-containing genes, splicing is not impaired when these modifications are eliminated in yeast. In contrast, deletion of Bre1 reduces splicing efficiency and therefore this effect is likely mediated through an as of yet unknown target of Bre1. Complementary to our finding, the catalytic subunit of the SAGA complex, Gcn5, is required for spliceosome assem-

bly during transcription in yeast (Gunderson and Johnson, 2009).

However, Bre1's impact on mRNA splicing could also be caused by a function of Bre1 that is not associated with chromatin regulation. For example, it is possible that Bre1 indirectly affects intron levels by degradation of unspliced mRNAs. This hypothesis needs to be tested and future investigations have to elucidate the role of Bre1 in mRNA splicing.

4.4 Experimental procedures

4.4.1 Yeast strains

S. cerevisiae strains used in this study are listed in Figure S4. Complete deletion of genes and integration of FLAG tag (Gelbart et al., 2001) in frame was achieved using the one-step gene integration of PCR-amplified modules (Longtine et al., 1998).

4.4.2 Chromatin immunoprecipitation and genome-wide ChIP-on-chip

Chromatin immunoprecipitation and genome-wide analyses were performed as described previously, using the adapted linear amplification method that involves two rounds of T7 RNA polymerase amplification (van Bakel et al., 2008). In brief, 500 ml of yeast cells were grown in a rich medium to an OD₆₀₀ 0.8–0.9 and were cross-linked with 1% formaldehyde for 20 min before chromatin was extracted. The chromatin was sonicated (Bioruptor, Diagenode: 10 cycles, 30 s on/off, high setting) to yield an average DNA fragment of 500 bp. For immunoprecipitation, 4 µl of H3K4me₃, H3K79me₂, H3K79me₃ (Shilatifard lab) or H3K36me₃ (Abcam ab9050) antibodies were coupled to 60 µl of protein A magnetic beads (Invitrogen). For H2B ubiquitin, 8 µl of H2BK123ub antibody was combined with 4 µl of H2B peptide (Shilatifard lab) to prevent unspecific binding. Prp40 FLAG was immunoprecipitated with anti-FLAG M2 antibody (Sigma F3165). After reversal of the crosslinking and DNA purification, the immunoprecipitated and input DNA were amplified to about 6 µg RNA using T7 RNA polymerase in two rounds. Samples were labeled with biotin, and the immunoprecipitated and input sample were

hybridized to two Affymetrix 1.0R *S. cerevisiae* microarrays, which are comprised of over 3.2 million probes covering the complete genome.

Table 4.4: Yeast strains used in this study. Listed are strains used in this study for ChIP-on-chip as well as gene expression experiments.

Name	Relevant Genotype
SLJ001	<i>MATα bar1 ade2-1 leu2-3,112 his3-11,13 trp1-1 ura3-1 can1-100</i>
YSN545	<i>(hta1-htb1)Δ::LEU2 (hta2-htb2)Δ::TRP1 [pZS145 HTA1-Flag-HTB1 WT CEN HIS3] MATα his3Δ200 leu2Δ1 ura3-52 trp1Δ63 lys2-128Δ</i>
YSN763	<i>(hta1-htb1)Δ::LEU2 (hta2-htb2)Δ::TRP1 [pSN423 HTA1-Flag-HTB1-K123A CEN HIS3] MATα his3Δ200 leu2Δ1 ura3-52 trp1Δ63 lys2-128Δ</i>
YSN4	<i>ubp8::KanMX6 MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>
YSN18	<i>ubp10::NatMX (hta1-htb1)Δ::LEU2, (hta2-htb2)Δ::TRP1 [pSAB6 (HTA1-HTB1-URA3)] MATα his3Δ200 leu2Δ1 ura3-52 trp1Δ63 lys2-128Δ</i>
MKY5	<i>MATα ade2-1 can1-100 his3-11 leu2-3,112 trp1-1 ura3-1 LYS2</i>
MKY927	<i>bre1::HYG MATα ade2-1 can1-100 his3-11 leu2-3,112 trp1-1 ura3-1 LYS2</i>
GLY256	<i>set1::HIS5 MATα ade2-1 can1-100 his3-11 leu2-3,112 trp1-1 ura3-1 LYS2</i>
AWY462	<i>set2::NatMX MATα ade2-1 can1-100 his3-11 leu2-3,112 trp1-1 ura3-1 LYS2</i>

4.4.3 Data analysis

As previously described, we used the MAT (Model-based Analysis of Tiling-arrays) algorithm to reliably detect enriched regions (Johnson et al., 2006; Droit et al., 2010). Each ChIP was normalized against the input except for H2BK123ub, which was normalized against a mock ChIP using an *H2BK123A* strain. MAT scores are calculated from all probes within a 300 bp sliding window and returns a MAT score for each probe. The MAT scores, as a measure for relative enrichment, were visualized along the whole genome using custom-written scripts in the statistics environment R. Given the transient nature of the H2BK123ub mark, we determined a high confidence data set intersecting the to two replicates. For all other marks, the Spearman correlation of replicates was to $r > 0.9$, and we based

our analysis on one representative data set.

Annotations for ORFs and ARSs were derived from the SGD database. An open reading frame (ORF) was termed enriched if at least 50 % of all probes had a MAT score above a threshold of 1.5. For a more comprehensive analysis, genes were separated into the following parts: promoter region (300 bp upstream of the coding region), 5'CDS (300 bp downstream CDS start), middle CDS (300 bp around centre of ORF), and 3'CDS (300 bp upstream of CDS end). These regions were defined as enriched if the average MAT score in that region was above the cutoff of 1.5. The list of transcription start end sites for 5015 genes was derived from Lee et al., 2007b.

To determine co-occurring histone marks the average enrichment scores for each ORF were calculated, and hierarchically clustered using the clustergram function of the Statistics toolbox in MATLAB 2009a. For the comparison with transcriptional activity of genes, each gene was assigned to the transcriptional frequency measured by Holstege et al., 1998 (Transcriptome 2005). To better characterize histone modification patterns, all occurring groups were compared to previously published lists of genes bound by specific transcription factors (MacIsaac et al., 2006). In order to explore the statistical significance of the overlap between these groups, we used the hypergeometric test (Tavazoie et al., 1999) and set a cutoff of $p < 0.001$.

4.4.4 Total RNA extraction, cDNA synthesis and array hybridization

Strains were grown in 100 ml of rich yeast-extract/peptone/dextrose media to mid-exponential phase. Total RNA was isolated by hot phenol extraction. RNA was treated with RNase-free DNaseI (Ambion's Turbo DNA-free Kit) for 30 min at 37 °C according to the manufacturer's instructions and subsequently reverse transcribed to single-stranded cDNA for microarray hybridization. Each 200 µl reverse transcription reaction was comprised of 20 µg of total RNA, 1.7 µg of random primers, 2.5 µg of Oligo(dT)12-18 primer, 0.5 mmol/l dNTPs, and 3000 units of SuperScript III (Invitrogen). After the RNA and primers were denatured for 10 min at 70 °C, the synthesis was carried out at 37 °C and 42 °C for 1 h each, followed by reverse transcriptase inactivation at 70 °C for 10 min. Afterwards, the RNA

was degraded with 1/3 volume of 1 mol/l NaOH incubated at 65 °C for 30 min, and an addition of 1/3 volume of 1 mol/l HCl was used to neutralize the solution before cleanup. cDNA samples were further applied to the MinElute Reaction Cleanup column (Qiagen), with the buffers from the QIAquick Nucleotide Removal Kit (Qiagen) and eluted in 30 µl elution buffer and quantified. Purified cDNA (about 6 µg) was fragmented with DNaseI (Invitrogen), labeled with Biotin-11-ddATP (Perkin-Elmer) and then hybridized to custom Affymetrix tiling arrays (PN 520055), which contain both strands of the *S. cerevisiae* genome tiled with a 8 bp resolution (David et al., 2006). The \log_2 perfect match (PM) probe intensities from each array were background corrected and calibrated using the DNA reference normalization method described in Huber et al., 2006. The average of the probe signals within the boundaries of each intron and exon for all spliced genes was calculated and the ratios between mutant and wild-type strain were determined.

Bibliography

- Ares, M., Grate, L. and Pauling, M. H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* *5*, 1138–9. → pages 162
- Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. and Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci USA* *99*, 8695–700. → pages 135
- Bhaumik, S. R., Smith, E. and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* *14*, 1008–16. → pages 133, 134
- Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., Winston, F. and Allis, C. D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev* *15*, 3286–95. → pages 135
- Briggs, S. D., Xiao, T., Sun, Z.-W., Caldwell, J. A., Shabanowitz, J., Hunt, D. F., Allis, C. D. and Strahl, B. D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* *418*, 498. → pages 134, 158
- Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M. P. and Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* *123*, 581–92. → pages 162
- Chandrasekharan, M. B., Huang, F. and Sun, Z.-W. (2009). Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *Proc Natl Acad Sci USA* *106*, 16686–91. → pages 162
- D'Andrea, A. D. (2010). Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med* *362*, 1909–19. → pages 159

- Daniel, J. A., Torok, M. S., Sun, Z.-W., Schieltz, D., Allis, C. D., Yates, J. R. and Grant, P. A. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* 279, 1867–71. → pages 134, 160
- David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C. J., Bofkin, L., Jones, T., Davis, R. W. and Steinmetz, L. M. (2006). A high-resolution map of transcription in the yeast genome. *Proc Natl Acad Sci USA* 103, 5320–5. → pages 166
- Davie, J. R. and Murphy, L. C. (1990). Level of ubiquitinated histone H2B in chromatin is coupled to ongoing transcription. *Biochemistry* 29, 4752–7. → pages 158
- Dover, J., Schneider, J., Tawiah-Boateng, M., Wood, A., Dean, K., Johnston, M. and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 277, 28368–71. → pages 134, 158
- Droit, A., Cheung, C. and Gottardo, R. (2010). rMAT—an R/Bioconductor package for analyzing ChIP-chip experiments. *Bioinformatics* 26, 678–9. → pages 164
- Emre, N. C. T., 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. *Mol Cell* 17, 585–94. → pages 134, 160
- 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. *Mol Cell* 31, 57–66. → pages 161
- 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. *Mol Cell Biol* 25, 6123–39. → pages 134, 160

- Gelbart, M., Rechsteiner, T., Richmond, T. and Tsukiyama, T. (2001). Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Mol Cell Biol* 21, 2098–106. → pages 163
- Gerber, M. and Shilatifard, A. (2003). Transcriptional elongation by RNA polymerase II and histone methylation. *The Journal of biological chemistry* 278, 26303–6. → pages 160
- Gunderson, F. Q. and Johnson, T. L. (2009). Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. *PLoS Genet* 5, e1000682. → pages 163
- Henry, K. W., Wyce, A., Lo, W.-S., Duggan, L. J., Emre, N. C. T., 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 Dev* 17, 2648–63. → pages 134
- Holstege, F., Jennings, E., Wyrick, J., Lee, T., Hengartner, C., Green, M., Golub, T., Lander, E. and Young, R. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717–28. → pages 148, 151, 159, 165
- Huber, W., Toedling, J. and Steinmetz, L. M. (2006). Transcript mapping with high-density oligonucleotide tiling arrays. *Bioinformatics* 22, 1963–70. → pages 153, 166
- Humpal, S. E., Robinson, D. A. and Krebs, J. E. (2009). Marks to stop the clock: histone modifications and checkpoint regulation in the DNA damage response. *Biochem Cell Biol* 87, 243–53. → pages 135
- Hwang, W., Venkatasubrahmanyam, S., Ianculescu, A., Tong, A., Boone, C. and Madhani, H. (2003). A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell* 11, 261–6. → pages 134
- Jeltsch, A. and Rathert, P. (2008). Putting the pieces together: histone H2B ubiquitylation directly stimulates histone H3K79 methylation. *Chembiochem* 9, 2193–5. → pages 161

- Johnson, W., Li, W., Meyer, C., Gottardo, R., Carroll, J., Brown, M. and Liu, X. (2006). Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci U S A* *103*, 12457–62. → pages 164
- Joshi, A. A. and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* *20*, 971–8. → pages 162
- Köhler, A., Zimmerman, E., Schneider, M., Hurt, E. and Zheng, N. (2010). Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell* *141*, 606–17. → pages 160
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693–705. → pages 133
- 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. *Mol Cell* *11*, 721–9. → pages 134, 135, 160, 161
- Lacoste, N., Utley, R., Hunter, J., Poirier, G. and Cote, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. *J Biol Chem* *277*, 30421–4. → pages 135
- Lee, J., Shukla, A., Schneider, J., Swanson, S., Washburn, M., Florens, L., Bhau-mik, S. and Shilatifard, A. (2007a). Histone crosstalk between H2B monoubiq-uitination and H3 methylation mediated by COMPASS. *Cell* *131*, 1084–96. → pages 134, 160
- Lee, W., Tillo, D., Bray, N., Morse, R., Davis, R., Hughes, T. and Nislow, C. (2007b). A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* *39*, 1235–44. → pages 140, 165
- Levinger, L. and Varshavsky, A. (1982). Selective arrangement of ubiquitinated and D1 protein-containing nucleosomes within the *Drosophila* genome. *Cell* *28*, 375–85. → pages 158

- Li, B., Gogol, M., Carey, M., Pattenden, S. G., Seidel, C. and Workman, J. L. (2007). Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* 21, 1422–30. → pages 162
- Li, B., Jackson, J., Simon, M. D., Fleharty, B., Gogol, M., Seidel, C., Workman, J. L. and Shilatifard, A. (2009). Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *The Journal of biological chemistry* 284, 7970–6. → pages 134
- Longtine, M., McKenzie, A., Demarini, D., Shah, N., Wach, A., Brachat, A., Philippsen, P. and Pringle, J. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14, 953–61. → pages 163
- Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M. and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* 327, 996–1000. → pages 162
- MacIsaac, K. D., Wang, T., Gordon, D. B., Gifford, D. K., Stormo, G. D. and Fraenkel, E. (2006). An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics* 7, 113. → pages 152, 165
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F. and Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci USA* 98, 12902–7. → pages 135
- 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. *Nat Cell Biol* 10, 483–8. → pages 136, 151
- Muramoto, T., Müller, I., Thomas, G., Melvin, A. and Chubb, J. R. (2010). Methylation of H3K4 Is required for inheritance of active transcriptional states. *Curr Biol* 20, 397–406. → pages 160

- Nakanishi, S., Lee, J. S., Gardner, K. E., Gardner, J. M., Imai Takahashi, Y., Chandrasekharan, M. B., Sun, Z.-W., Osley, M. A., Strahl, B. D., Jaspersen, S. L. and Shilatifard, A. (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *J Cell Biol* *186*, 371–7. → pages 134, 158, 160
- Ng, H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y. and Struhl, K. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev* *16*, 1518–27. → pages 134
- Ng, H. H., Robert, F., Young, R. A. and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* *11*, 709–19. → pages 135, 160
- Nickel, B. E., Allis, C. D. and Davie, J. R. (1989). Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochemistry* *28*, 958–63. → pages 158
- Perales, R. and Bentley, D. (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* *36*, 178–91. → pages 135, 162
- Pokholok, D., Harbison, C., Levine, S., Cole, M., Hannett, N., Lee, T., Bell, G., Walker, K., Rolfe, P., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. and Young, R. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* *122*, 517–27. → pages 136, 139, 145, 151
- Robzyk, K., Recht, J. and Osley, M. A. (2000). Rad6-dependent ubiquitination of histone H2B in yeast. *Science (New York, NY)* *287*, 501–4. → pages 134
- Rudra, D., Zhao, Y. and Warner, J. R. (2005). Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. *EMBO J* *24*, 533–42. → pages 152

- Samara, N. L., Datta, A. B., Berndsen, C. E., Zhang, X., Yao, T., Cohen, R. E. and Wolberger, C. (2010). Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science (New York, NY)* 328, 1025–9. → pages 160
- Sanders, S. L., Jennings, J., Canutescu, A., Link, A. J. and Weil, P. A. (2002). Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol Cell Biol* 22, 4723–38. → pages 134
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–11. → pages 135
- Schulze, J., Jackson, J., Nakanishi, S., Gardner, J., Hentrich, T., Haug, J., Johnston, M., Jaspersen, S., Kobor, M. and Shilatifard, A. (2009). Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation. *Mol Cell* 35, 626–41. → pages 134, 135, 136, 145, 151, 152, 158, 160
- Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* 75, 243–69. → pages 133, 134, 135, 158
- Sims, R. J., Millhouse, S., Chen, C.-F., Lewis, B. A., Erdjument-Bromage, H., Tempst, P., Manley, J. L. and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* 28, 665–76. → pages 162
- Singer, M. S., Kahana, A., Wolf, A. J., Meisinger, L. L., Peterson, S. E., Goggin, C., Mahowald, M. and Gottschling, D. E. (1998). Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* 150, 613–32. → pages 160
- Smith, E. and Shilatifard, A. (2009). Developmental biology. Histone cross-talk in stem cells. *Science (New York, NY)* 323, 221–2. → pages 134

- Spies, N., Nielsen, C. B., Padgett, R. A. and Burge, C. B. (2009). Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* 36, 245–54. → pages 162
- Strahl, B. D. and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41–5. → pages 133
- Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z.-W., Bone, J. R., Caldwell, J. A., Mollah, S., Cook, R. G., Shabanowitz, J., Hunt, D. F. and Allis, C. D. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* 22, 1298–306. → pages 135
- Suganuma, T. and Workman, J. L. (2008). Crosstalk among Histone Modifications. *Cell* 135, 604–7. → pages 134
- Sun, Z. and Allis, C. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418, 104–8. → pages 134, 158
- Tavazoie, S., Hughes, J., Campbell, M., Cho, R. and Church, G. (1999). Systematic determination of genetic network architecture. *Nat Genet* 22, 281–5. → pages 165
- van Bakel, H., van Werven, F., Radonjic, M., Brok, M., van Leenen, D., Holstege, F. and Timmers, H. (2008). Improved genome-wide localization by ChIP-chip using double-round T7 RNA polymerase-based amplification. *Nucleic Acids Res* 36, e21. → pages 163
- van Leeuwen, F., Gafken, P. and Gottschling, D. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* 109, 745–56. → pages 135
- Weake, V. M. and Workman, J. L. (2008). Histone ubiquitination: triggering gene activity. *Mol Cell* 29, 653–63. → pages 145, 158
- Wood, A., Krogan, N., Dover, J., Schneider, J., Heidt, J., Boateng, M., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J., Johnston, M. and Shilatifard, A. (2003a). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* 11, 267–74. → pages 134, 135

- Wood, A., Schneider, J., Dover, J., Johnston, M. and Shilatifard, A. (2003b). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *The Journal of biological chemistry* 278, 34739–42. → pages 161
- 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. *Mol Cell* 20, 589–99. → pages 161
- Workman, J. L. and Kingston, R. E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* 67, 545–79. → pages 133

Chapter 5

Discussion

Chromatin represents an important layer of genome regulation and can be described as a genome-organizing platform (Allis et al., 2007). Research in the chromatin field has intensified dramatically over the past decade and showed that many cellular processes are associated with chromatin modifications (Ehrenhofer-Murray, 2004). Although these efforts led to the identification of many chromatin players, we are still at the early stages of grasping the complexity of chromatin-mediated genome regulation.

My dissertation contributes to the understanding of chromatin biology by characterizing the structure and function of a chromatin modifier domain, by elucidating distinct roles of H3K79 methylation states, by identifying genome-wide patterns of H2B monoubiquitination, by revealing the distinct roles of the H2B deubiquitinases Ubp8 and Ubp10, as well as by characterizing chromatin patterns associated with highly transcribed genes, including those containing introns.

In Chapter 2, I present work studying the structure and function of the Yaf9 YEATS domain. It has no known catalytic activity but a characteristic binding pocket, which may bind to acetylated histones. I present evidence that the YEATS domain can indeed bind histones H3 and H4, consistent with its proposed role as a chromatin reader. It is tempting to speculate which histone acetylation mark it may bind more specifically (Schulze et al., 2010). Interestingly, the YEATS domain and the histone chaperone Asf1 are structurally very similar, and *ASF1* and *YAF9* inter-

act genetically, suggesting that they might share a common function. It is known that Asf1 and the associated acetylation of histone H3 lysine 56 (H3K56ac) are essential for checkpoint recovery after DNA repair, leading to inactivation of the DNA damage checkpoint (Recht et al., 2006; Chen et al., 2008). Similarly, NuA4 and SWR1-C, both containing Yaf9 as a subunit, are involved in DNA damage response and get recruited to DNA damage sites together with Ino80 (Bird et al., 2002; Downs et al., 2004; Papamichos-Chronakis et al., 2006; van Attikum et al., 2007). Arp4, another subunit shared between SWR1-C, NuA4, and INO80, is crucial for recruiting these complexes to damaged sites likely through direct physical interaction with phosphorylated H2A.X (Downs et al., 2004). Interestingly, when Arp4 is eliminated the association of Yaf9 with SWR1-C is impaired (Bittner et al., 2004; Szerlong et al., 2008; Wu et al., 2009). It has not been tested to date if Yaf9 might be involved in recruiting SWR1-C to DNA damage sites itself. It is possible that Yaf9, through its YEATS domain, may recognize a specific chromatin mark at the damaged site. Given the structural and genetic connection between Yaf9 and Asf1 as well as its dependent mark H3K56ac, Yaf9 could potentially bind to H3K56ac during the DNA repair process. While this hypothesis is consistent with data presented herein, further work is necessary to test this role of Yaf9 *in vitro* and *in vivo*. For example, mammalian tethered catalysis (MTeC) utilizing a fusion protein affinity tag in tandem with a histone peptide sequence and the reader domain, in this case the YEATS domain, could be used to identify the specific binding target of the YEATS domain (Spektor and Rice, 2009).

Another potential binding target of the Yaf9 YEATS domain is suggested by its role in transcription as part of both NuA4 and SWR1-C. The current model suggests that acetylation of histone H4 by NuA4 is required for deposition of H2A.Z by SWR1-C into chromatin, and subsequently NuA4 acetylates H2A.Z (Lu et al., 2009).

The bromodomain of Bdf1 plays an important role in recruiting SWR1-C to chromatin (Zhang et al., 2005; Raisner et al., 2005; Wu et al., 2009). Despite this essential role of the SWR1-C bromodomain, the YEATS domain of Yaf9 could potentially participate in the binding to acetylated H4 at specific promoters, thereby recruiting SWR1-C and NuA4 to deposit and acetylate H2A.Z at particular sites,

respectively. This speculation is supported by the fact that impairing Yaf9 YEATS domain function leads to loss of H2A.Z at specific promoters as shown in Chapter 2 of this dissertation. However, additional work is necessary to specifically measure the binding affinity of the YEATS domain to acetylated histones and determine its precise binding target. For instance, the binding of SWR1-C or NuA4 to specific chromatin sites could be measured in the presence and absence of Yaf9 or its YEATS domain by ChIP-on-chip to reveal its role in targeting these complexes to chromatin.

Although the YEATS domain may bind one particular histone modification, its binding to chromatin may be context dependent. In general, multivalent assemblages of domains are emerging as crucial interpreters of chromatin modification patterns (Ruthenburg et al., 2007). Thus, it is likely that the combination of all reader modules, like YEATS, bromo- or chromodomains, in given chromatin-modifying complexes such as NuA4 and SWR1-C are responsible for targeting them to specific chromatin neighbourhoods either during transcription or DNA repair. Therefore, the next step is to study chromatin readers in combination to more specifically reveal their biologically relevant targets.

Chromatin neighbourhoods are characterized by a large number of different histone modifications (Kouzarides, 2007). To date, numerous histone modifications have been identified and associated with cellular functions, but some fundamental principles still have to be identified. For instance, as I discuss in Chapter 3, many lysine residues of histones have different methylation states, and it has been a long standing question in the field whether these states are functionally distinct.

For methylation of H3K79, it was postulated that all three states are functionally redundant (Frederiks et al., 2008; Shahbazian et al., 2005). This conclusion was primarily based on genetic evidence indicating that all three methyl states function in telomeric silencing (Frederiks et al., 2008). In my dissertation, I challenge this view and provide evidence in Chapter 3 that H3K79 methylation states have different functions. Using state-specific methylation antibodies, I show that H3K79me2 and H3K79me3 are associated with mutually exclusive regions of the genome, a fundamental prerequisite to be functionally distinct.

To establish these mutually exclusive patterns, the sole enzyme Dot1 requires a

mechanism to specifically establish each state at certain regions in the genome. Dot1 is thought to be a non-processive enzyme, which binds H3K79, adds a single methyl residue, but has to release H3K79 before rebinding and catalyzing an additional methylation step (Frederiks et al., 2008). Generating different methylation sites likely requires mechanisms to either constrain Dot1 from certain regions, as it has been proposed for Sir3 (Fingerman et al., 2007), or to enable the region-specific binding of Dot1, for example through the binding of a particular histone modification. Such a potential modification is monoubiquitination of H2BK123, which is required to establish H3K79me3 specifically (Nakanishi et al., 2009). Indeed, as I show in Chapter 3, H2BK123ub co-localizes with H3K79me3, but not H3K79me2. Furthermore, *in vitro* reconstitution of the methylation reaction suggests that ubiquitin induces a conformational change to the Dot1-nucleosome complex which in turn stimulates Dot1's catalytic activity (Jeltsch and Rathert, 2008; McGinty et al., 2008). Therefore, H2BK123 monoubiquitin might be the primary determinant of H3K79me3 interacting with Dot1 to distinguish genomic regions.

While fundamentally important, this finding raises another question: How is H2B monoubiquitin itself established at certain genomic regions? In principle, there are several possible mechanisms for the recruitment of the H2B monoubiquitin ligase Rad6/Bre1 to target genes: through association of sequence-specific DNA binding factors, via association with the basal transcriptional machinery, or by binding chromatin through histone modification readers. It might also be a combination of these mechanisms, and it remains to be determined how H2BK123 monoubiquitin is established at certain genes and excluded from others. Nevertheless, my work represents an important step forward in understanding the methylation states of H3K79 and determining their functionally distinct roles.

In Chapter 3, I further describe the potential functions of the H3K79 methylation states. H3K79me2 is cell cycle regulated and marks ORFs and promoters of genes that are expressed specifically in G1/S phase. Interestingly, global H3K79me2 levels are highest during G2/M phase when G1/S genes are inactive. Therefore, H3K79me2 might help repress these genes during their inactive phase. To test this causal relationship, experiments are necessary to measure the gene expression of these genes during G2/M phase in the absence of H3K79me2.

As for H3K79 trimethylation, no specific function has been attributed so far and remains to be addressed by future research. The genome-wide map of H3K79me3 described in this work will help determine these functions. I show that H3K79me3 is associated with one third of all genes in budding yeast and that these genes tend to be infrequently transcribed. By comparing multiple histone modification maps across the genome, I found that H3K79me3 strongly co-localizes with H3K36 trimethylation, a key mark for transcription elongation. This strong association opens the door for new experiments to test the interplay between H3K36 and H3K79 methylation and will help to determine their functions.

How do the differences of H3K79 methylation states described herein relate to other histone lysine methylations? What is known about functional differences of methylation states for H3K36 and H3K4? For H3K36 in *S. cerevisiae*, it is not clear to date if its different methylation states are associated with different cellular functions (Li et al., 2009; Youdell et al., 2008). H3K36 trimethylation predominantly marks the 3' end of genes and positively correlates with their transcriptional activity (Pokholok et al., 2005), but as I show in Chapter 4 it is also found in many ORFs with lower transcriptional activity. H3K36 dimethylation is found within the ORF as well and correlates with the overall on/off state of transcription (Rao et al., 2005). The importance of different H3K36 methylation states is evident in higher organisms, which contain more than one enzyme to establish different H3K36 methylation states (Bell et al., 2007; Edmunds et al., 2008; Xu et al., 2008). In mouse and flies, HYPB/Setd2, the yeast Set2 homologue, controls H3K36 trimethylation, whereas *Drosophila* MES4 regulates mono- and dimethylation of H3K36 (Bell et al., 2007; Edmunds et al., 2008).

For H3K4, different methylation states occur in distinct patterns in budding yeast and higher eukaryotes (Pokholok et al., 2005; Schneider et al., 2004; Bernstein et al., 2005). In *S. cerevisiae*, H3K4me3 marks 5' end of ORFs, whereas H3K4me2 is enriched in the body of the ORF, and H3K4me1 is found predominantly at the 3' end of genes (Pokholok et al., 2005; Santos-Rosa et al., 2002; Ng et al., 2003). In vertebrates, H3K4 di- and trimethylation partially co-occur at transcriptional start sites (Schneider et al., 2004), but are generally distributed differently (Schneider et al., 2004; Bernstein et al., 2005). Moreover, ample evidence pointing towards

the importance of different methyl states is given by the specificity of reader modules recognizing distinct methyl states of H3K4 methylation. For instance, the PHD-finger of BPTF, a subunit of the nucleosome remodelling complex NURF, and the PHD fingers of the tumour suppressor ING family preferentially bind to H3K4me3, with up to 10-fold greater affinity compared to H3K4me2 and another 10-fold decrease from H3K4me2 to H3K4me1 (Li et al., 2006; Peña et al., 2006; Taverna et al., 2006). In contrast, the WD repeat domain 5 (WDR5) of the MLL-family complexes preferentially binds to the dimethyl state of H3K4 (Wysocka et al., 2005). Although the precise recognition mechanisms to achieve this kind of specificity are still controversial and have to be elucidated, the existence of methyl state specific readers highlights the importance of minor molecular differences of histone marks.

Taken together, the work presented herein strengthens the hypothesis that different methylation states of histone lysines are likely associated with different functions and contribute to the overall complexity of histone modifications.

Histone modifications do not necessarily occur in isolation, but often co-occur with other modifications (Rando, 2007). A major interest in the field is to identify the complexity of these modification patterns.

In Chapter 4, I describe the genome-wide maps of histone H2B ubiquitination and H3 methylation in *S. cerevisiae*, and compared them to find characteristic patterns. Of all possible combinations, some were characteristic for specific genomic loci: For instance, highly transcribed genes were exclusively and significantly marked by the combination of H2BK123ub, H3K4me3 and H3K36me3. As most highly transcribed genes in *S. cerevisiae* contain introns, this triple combination also characterized many spliced genes. In agreement with this association, in human cells, alternative splicing has been causally linked to some of these histone modifications (Schwartz and Ast, 2010). In contrast, in Chapter 4, I show that eliminating these histone modifications individually does not significantly affect splicing in *S. cerevisiae*. This difference between yeast and human cells, could be due to the lack of alternative splicing in *S. cerevisiae*. Interestingly, the lack of direct functional causality of single histone modifications has also been reported for other processes such as transcription. In most cases, mutating histone modifiers leads

to minor changes in the expression of genes associated with the particular histone mark (personal communication with Frank Holstege).

As a result, it is important to discuss correlations between histone modifications and cellular functions in greater detail, and to carefully distinguish between simple associations versus causal relationships.

One explanation for the observed discrepancy between association and causality may be that the conditions used to measure cellular outcomes were not optimal to detect the function of a particular chromatin modification. For instance, while H2A.Z has been associated with transcriptional regulation (Draker and Cheung, 2009), impairing H2A.Z leads to only minor effect on gene expression when tested under ideal growth conditions (Meneghini et al., 2003; Kobor et al., 2004). However, when gene expression changes are measured under conditions that disturb H2A.Z function such as heat shock or the absence of fatty acids, more dramatic effects can be measured (Zanton and Pugh, 2006; Wan et al., 2009). Therefore, it is important to choose the right conditions to measure cellular outcomes in order to draw conclusions about causal relationships.

Furthermore, it is probably the complex circuitry between different histone modifications in a neighbourhood that profoundly influences cellular functions (Latham and Dent, 2007; Suganuma and Workman, 2008; Ruthenburg et al., 2007). To test this hypothesis, effects on cellular outputs such as transcription or mRNA splicing have to be measured in mutants impairing multiple modifications. Thereby, it will become possible to learn more about complex patterns of histone modifications that act in combination.

When investigating the combinatorial effects of chromatin modifications, one must address the question of how many marks can co-occur on one nucleosome or a given histone. In Chapter 3 and 4, my results suggest that certain histone modifications co-occur within the same nucleosome. However, a limitation of the chromatin modification maps described herein is their resolution, which does not reach nucleosome level. To improve the resolution, the DNA needs to be treated with micrococcal nuclease (MNase) instead of being sheared by sonication (O'Neill and Turner, 2003). Moreover, the technical nature of ChIP-on-Chip as well as ChIP-Seq requires large numbers of cells and consequently yields an average pro-

file across the population. Carrier ChIP (CCIP) reduces the amount of cells required per experiment to as few as 100 cells by spiking the sample with genomic DNA (O'Neill et al., 2006). Ideally, genome-wide mapping has to be done in a single cell, but is technically still not possible as of the writing of this dissertation.

Once we reach nucleosome-resolution in a single cell, we can decipher whether the crosstalk between H2B monoubiquitination and H3K4- and H3K79 trimethylation occurs within the same nucleosome or across adjacent nucleosomes. However, we still would not be able to reveal whether modifications, such as H3K4-, H3K36- and H3K79 trimethylation, occur on the same or on different histone H3 molecules within the same nucleosome. Single-molecule methods have to be developed for mapping all modifications on all histones within a nucleosome at once to examine how they influence one another. A step forward in this direction was made recently by increasing the sensitivity of mass spectroscopic approaches, which allows the identification of co-occurring modifications on the same histone molecule (Young et al., 2010; Taverna et al., 2007). Moreover, advances in protein chemistry allow the *in vitro* composition of semi-synthetic histones with multiple modifications to investigate their mechanistic roles and cooperation, work that may eventually help to test the histone code hypothesis (Chatterjee and Muir, 2010). Furthermore, it is important to keep in mind that histone modifications are highly dynamic, and we have to capture these dynamics in order to understand the sequential logic within their combinatorial patterns. ChIP-on-chip, ChIP-Seq, and mass spectroscopy provide only temporal snapshots of chromatin. One way to capture the dynamics is by arresting cells at certain cell cycle stages and comparing the resulting profiles, as I did for the experiments in Chapter 3, or by capturing certain chromatin states under inducible conditions such as the heat shock response (Zanton and Pugh, 2006). Moreover, new techniques like CATCH-IT (covalent attachment of tags to capture histones and identify turnover) for measuring native histone turnover are promising leaps forward for chromatin research (Deal et al., 2010). For now, however, it remains technically highly difficult to study the dynamics even at nucleosome resolution.

Besides improving our understanding of chromatin at the histone level by zooming into the realm of single molecules, another challenge of chromatin biology is to

zoom out multiple orders of magnitude and understand the nuclear architecture of chromatin. We are still far from a detailed picture of how the genetic material is spatially organized within the nucleus and how dynamic its localization is (Misteli, 2007). However, advances in this direction have been made by using live-cell imaging and chromosome conformation capture assay such as 3C, 4C, 5C, or ChIA-Pet (Zhao et al., 2006; Simonis et al., 2006; Dekker et al., 2002; Vassetzky et al., 2009), but plenty of aspects await further research.

Understanding the nuclear organization of chromatin will also reveal how the nuclear machineries are spatially and temporally coordinated to allow regulation of multiple, sometimes genomically distant, genes. The organization of the genetic information and the factors associated with its regulation seems far from random (Misteli, 2007). Within the nucleus, they create sub-nuclear compartments that seem ideal for genome regulatory mechanisms. Establishing maps of histone modifications as I present herein, and, for example, studies of transcriptional activity, have to go hand in hand with maps of higher-order chromatin structure and spatial position of chromatin to understand how they influence the cellular output and functional states of the genome. New experimental and theoretical approaches, including visualization technologies, analysis of chromatin dynamics, and system-wide computational modelling, are necessary to approach this challenge.

In closing, it is important to point out that chromatin fundamentally affects human health and disease directly (Bhaumik et al., 2007). It is becoming clear that chromatin changes are not only crucial for normal development, but also for maintaining human health. Traditionally, to understand diseases and find therapies, research has mainly focused on genetic changes such as amplifications, deletions and point mutations. However, impairing chromatin modifications, such as DNA and histone methylation, also lead to occurrence and progression of human disease (Feinberg, 2007; Petronis, 2010). Advances in chromatin research will eventually lead to therapeutics and treatments for diseases caused by misregulation of chromatin and thus contribute to human health.

Bibliography

- Allis, C. D., Jenuwein, T., Reinberg, D. and Caparros, M. L. (2007). Epigenetics. 1st edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. → pages 176
- Bell, O., Wirbelauer, C., Hild, M., Scharf, A. N. D., Schwaiger, M., MacAlpine, D. M., Zilbermann, F., van Leeuwen, F., Bell, S. P., Imhof, A., Garza, D., Peters, A. H. F. M. and Schübeler, D. (2007). Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in *Drosophila*. *The EMBO journal* 26, 4974–84. → pages 180
- Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J., Gingeras, T. R., Schreiber, S. L. and Lander, E. S. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169–81. → pages 180
- Bhaumik, S. R., Smith, E. and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14, 1008–16. → pages 184
- Bird, A., Yu, D., Pray-Grant, M., Qiu, Q., Harmon, K., Megee, P., Grant, P., Smith, M. and Christman, M. (2002). Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* 419, 411–5. → pages 177
- Bittner, C., Zeisig, D., Zeisig, B. and Slany, R. (2004). Direct physical and functional interaction of the NuA4 complex components Yaf9p and Swc4p. *Eukaryot Cell* 3, 976–83. → pages 177
- Chatterjee, C. and Muir, T. W. (2010). Chemical approaches for studying histone modifications. *The Journal of biological chemistry* 285, 11045–50. → pages 183
- Chen, C.-C., Carson, J. J., Feser, J., Tamburini, B., Zabaronick, S., Linger, J. and Tyler, J. K. (2008). Acetylated lysine 56 on histone H3 drives chromatin assem-

bly after repair and signals for the completion of repair. *Cell* 134, 231–43. → pages 177

Deal, R. B., Henikoff, J. G. and Henikoff, S. (2010). Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* (New York, NY) 328, 1161–4. → pages 183

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002). Capturing chromosome conformation. *Science* (New York, NY) 295, 1306–11. → pages 184

Downs, J. A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., Kron, S. J., Jackson, S. P. and Côté, J. (2004). Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol Cell* 16, 979–90. → pages 177

Draker, R. and Cheung, P. (2009). Transcriptional and epigenetic functions of histone variant H2A.Z. *Biochem Cell Biol* 87, 19–25. → pages 182

Edmunds, J. W., Mahadevan, L. C. and Clayton, A. L. (2008). Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *The EMBO journal* 27, 406–20. → pages 180

Ehrenhofer-Murray, A. E. (2004). Chromatin dynamics at DNA replication, transcription and repair. *Eur J Biochem* 271, 2335–49. → pages 176

Feinberg, A. P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature* 447, 433–40. → pages 184

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 Dev* 21, 2018–29. → pages 179

Frederiks, F., Tzouros, M., Oudgenoeg, G., van Welsem, T., Fornerod, M., Krijgsveld, J. and van Leeuwen, F. (2008). Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. *Nat Struct Mol Biol* 15, 550–7. → pages 178, 179

- Jeltsch, A. and Rathert, P. (2008). Putting the pieces together: histone H2B ubiquitylation directly stimulates histone H3K79 methylation. *Chembiochem* 9, 2193–5. → pages 179
- Kobor, M., Venkatasubrahmanyam, S., Meneghini, M., Gin, J., Jennings, J., Link, A., Madhani, H. and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* 2, E131. → pages 182
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. → pages 178
- Latham, J. A. and Dent, S. Y. R. (2007). Cross-regulation of histone modifications. *Nat Struct Mol Biol* 14, 1017–1024. → pages 182
- Li, B., Jackson, J., Simon, M. D., Fleharty, B., Gogol, M., Seidel, C., Workman, J. L. and Shilatifard, A. (2009). Histone H3 lysine 36 dimethylation (H3K36me₂) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *The Journal of biological chemistry* 284, 7970–6. → pages 180
- Li, H., Ilin, S., Wang, W., Duncan, E., Wysocka, J., Allis, C. and Patel, D. (2006). Molecular basis for site-specific read-out of histone H3K4me₃ by the BPTF PHD finger of NURF. *Nature* 442, 91–5. → pages 181
- Lu, P. Y. T., Lévesque, N. and Kobor, M. S. (2009). NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and components. *Biochem Cell Biol* 87, 799–815. → pages 177
- 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, 812–6. → pages 179
- Meneghini, M. D., Wu, M. and Madhani, H. D. (2003). Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* 112, 725–36. → pages 182

- Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787–800. → pages 184
- Nakanishi, S., Lee, J. S., Gardner, K. E., Gardner, J. M., Imai Takahashi, Y., Chandrasekharan, M. B., Sun, Z.-W., Osley, M. A., Strahl, B. D., Jaspersen, S. L. and Shilatifard, A. (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *J Cell Biol* 186, 371–7. → pages 179
- Ng, H. H., Robert, F., Young, R. A. and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* 11, 709–19. → pages 180
- O’Neill, L. P. and Turner, B. M. (2003). Immunoprecipitation of native chromatin: NChIP. *Methods* 31, 76–82. → pages 182
- O’Neill, L. P., VerMilyea, M. D. and Turner, B. M. (2006). Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nat Genet* 38, 835–41. → pages 183
- Papamichos-Chronakis, M., Krebs, J. E. and Peterson, C. L. (2006). Interplay between Ino80 and Swr1 chromatin remodeling enzymes regulates cell cycle checkpoint adaptation in response to DNA damage. *Genes Dev* 20, 2437–49. → pages 177
- Peña, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., Zhao, R. and Kutateladze, T. G. (2006). Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442, 100–3. → pages 181
- Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 465, 721–7. → pages 184
- Pokholok, D., Harbison, C., Levine, S., Cole, M., Hannett, N., Lee, T., Bell, G., Walker, K., Rolfe, P., Herbolzheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. and Young, R. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 122, 517–27. → pages 180

- Raisner, R., Hartley, P., Meneghini, M., Bao, M., Liu, C., Schreiber, S., Rando, O. and Madhani, H. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* 123, 233–48. → pages 177
- Rando, O. J. (2007). Global patterns of histone modifications. *Curr Opin Genet Dev* 17, 94–9. → pages 181
- Rao, B., Shibata, Y., Strahl, B. D. and Lieb, J. D. (2005). Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. *Mol Cell Biol* 25, 9447–59. → pages 180
- Recht, J., Tsubota, T., Tanny, J., Diaz, R., Berger, J., Zhang, X., Garcia, B., Shabanowitz, J., Burlingame, A., Hunt, D., Kaufman, P. and Allis, C. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc Natl Acad Sci U S A* 103, 6988–93. → pages 177
- Ruthenburg, A., Li, H., Patel, D. and Allis, C. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* 8, 983–94. → pages 178, 182
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–11. → pages 180
- Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C. and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* 6, 73–7. → pages 180
- Schulze, J. M., Wang, A. Y. and Kobor, M. S. (2010). Reading chromatin: Insights from yeast into YEATS domain structure and function. *Epigenetics : official journal of the DNA Methylation Society* 5. → pages 176
- Schwartz, S. and Ast, G. (2010). Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing. *The EMBO journal* 29, 1629–36. → pages 181

- Shahbazian, M., Zhang, K. and Grunstein, M. (2005). Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. *Mol Cell* *19*, 271–7. → pages 178
- Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* *38*, 1348–54. → pages 184
- Spektor, T. M. and Rice, J. C. (2009). Identification and characterization of post-translational modification-specific binding proteins in vivo by mammalian tethered catalysis. *Proc Natl Acad Sci USA* *106*, 14808–13. → pages 177
- Suganuma, T. and Workman, J. L. (2008). Crosstalk among Histone Modifications. *Cell* *135*, 604–7. → pages 182
- Szerlong, H., Hinata, K., Viswanathan, R., Erdjument-Bromage, H., Tempst, P. and Cairns, B. R. (2008). The HSA domain binds nuclear actin-related proteins to regulate chromatin-remodeling ATPases. *Nat Struct Mol Biol* *15*, 469–76. → pages 177
- Taverna, S. D., Ilin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L. P., Chait, B. T., Patel, D. J., Aitchison, J. D., Tackett, A. J. and Allis, C. D. (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol Cell* *24*, 785–96. → pages 181
- Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., Chait, B. T., Hunt, D. F. and Allis, C. D. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proc Natl Acad Sci USA* *104*, 2086–91. → pages 183
- van Attikum, H., Fritsch, O. and Gasser, S. M. (2007). Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *The EMBO journal* *26*, 4113–25. → pages 177

- Vassetzky, Y., Gavrilov, A., Eivazova, E., Priozhkova, I., Lipinski, M. and Razin, S. (2009). Chromosome conformation capture (from 3C to 5C) and its CHIP-based modification. *Methods Mol Biol* 567, 171–88. → pages 184
- Wan, Y., Saleem, R. A., Ratushny, A. V., Roda, O., Smith, J. J., Lin, C.-H., Chiang, J.-H. and Aitchison, J. D. (2009). Role of the histone variant H2A.Z/Htz1p in TBP recruitment, chromatin dynamics, and regulated expression of oleate-responsive genes. *Mol Cell Biol* 29, 2346–58. → pages 182
- Wu, W.-H., Wu, C.-H., Ladurner, A., Mizuguchi, G., Wei, D., Xiao, H., Luk, E., Ranjan, A. and Wu, C. (2009). N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex. *The Journal of biological chemistry* 284, 6200–7. → pages 177
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H. and Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121, 859–72. → pages 181
- Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J.-P., Steinmetz, A. and Shen, W.-H. (2008). Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in *Arabidopsis thaliana*. *Mol Cell Biol* 28, 1348–60. → pages 180
- Youdell, M. L., Kizer, K. O., Kisseleva-Romanova, E., Fuchs, S. M., Duro, E., Strahl, B. D. and Mellor, J. (2008). Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol Cell Biol* 28, 4915–26. → pages 180
- Young, N. L., Plazas-Mayorca, M. D. and Garcia, B. A. (2010). Systems-wide proteomic characterization of combinatorial post-translational modification patterns. *Expert Rev Proteomics* 7, 79–92. → pages 183
- Zanton, S. J. and Pugh, B. F. (2006). Full and partial genome-wide assembly and disassembly of the yeast transcription machinery in response to heat shock. *Genes Dev* 20, 2250–65. → pages 182, 183

Zhang, H., Roberts, D. N. and Cairns, B. R. (2005). Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* 123, 219–31. → pages 177

Zhao, Z., Tavoosidana, G., Sjölander, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K. S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S. and Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* 38, 1341–7. → pages 184