CONSERVATION AND FUNCTION OF THE HISTONE METHYLTRANSFERASE SET2

Stephanie A. Morris

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2007

Approved by:
Advisor: Brian D. Strahl
Reader: Kerry S. Bloom
Reader: Jeanette G. Cook
Reader: Henrik G. Dohlman
Reader: William F. Marzluff

ABSTRACT

STEPHANIE A. MORRIS: Conservation and Function of the Histone Methyltransferase Set2
(Under the direction of Brian D. Strahl)

Histone methylation is an important post-translational modification involved in the regulation of eukaryotic gene expression. While many methylation sites on histone proteins have been identified to play roles in both gene activation and repression, the enzymes mediating these modifications and their exact functions are just beginning to be discovered. In the budding yeast *Saccharomyces cerevisiae,* methylation of histone H3 at lysine 36 (H3K36) by the histone methyltransferase Set2 has been linked to the process of transcription elongation. Previous findings indicate that through an interaction with the elongating RNA polymerase, Set2 targets H3K36 for methylation in the coding region of genes. However, the exact functions for this enzyme and its modification were largely unknown. In these studies, I demonstrate that Set2 methylation of H3K36 is highly conserved and associated with elongating RNA polymerase II in organisms distinct from budding yeast. These results reveal that Set2 and H3K36 methylation have a conserved role in the transcription elongation process. Furthermore, I have contributed to the finding that Set2 regulates global histone acetylation patterns by recruiting a small Rpd3 deacetylase (Rpd3S) complex to the coding region of genes. This is among one of the first studies to identify a functional mechanism for Set2-mediated H3K36 methylation in transcription elongation. Finally, I have identified a novel and conserved modification on H3K36. Independent of being methylated, my studies

reveal H3K36 is acetylated by the transcriptional co-activator Gcn5 at promoter regions. Collectively, these results suggest that distinct modifications on H3K36 play diverse roles in the transcription process.

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Brian Strahl. Not only has he been a great mentor, but his constant enthusiasm and positive energy have made my experience as a graduate student a very good one. I would also like to thank all of the current and past members of the Strahl laboratory. Having spent so much time together, they have become my second family. A special thanks to Nick and Kelby whose opposing personalities brought balance to the satellite lab.

With the fast-pace of research in the chromatin and transcription fields, collaborations have become an essential component of the research process. Therefore, I must give a big thank you to my collaborators. I have been quite fortunate to have worked with some of the top laboratories and to have made such great connections with fellow scientists that I know will last for years to come. Thank you to my committee members as well. They have been very supportive throughout this process.

Last, but not least, I would like to thank my friends and family for their support and understanding. I especially would like to thank my family who has always stood by the decisions I have made and were confident that I would be successful at whatever I decided to do in life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

A	alanine
ac	acetylation
acetyl-CoA	acetyl coenzyme A
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
bp	base pair
C	carboxyl
CBP	calmodulin binding protein
ChIP	chromatin immunoprecipitations
chip	DNA microarray analyses
cpm	counts per minute
CTD	C-terminal domain
DNA	deoxyribonucleic acid
dpm	disintegrations per minute
EMM	Edinburgh's minimal media
ECL	enhanced chemiluminescence
EDTA	ethylendiamine tetraacetic acid
FACT	facilitates chromatin transcription
H	histone
HA	hemagglutinin epitope tag
HAT	histone acetyltransferase

HDAC	histone deacetylase
HMT	histone methyltransferase
HRP	horseradish peroxidase
IPTG	isopropyl-D-thiogalactoside
ING	inhibitor of growth
IP	immunoprecipitation
K	lysine
kb	kilobase pair
kDa	kilodalton
L	liter
LTQ-FT	linear quadrupole ion trap-Fourier transform
ml	milliliter
mM	millimolar
M	molar
me	methylation
mRNA	messenger RNA
MS	mass spectrometry
MYND	Myeloid translocation protein 8, Nervy, and DEAF-1
N	amino
nl	nanoliter
nm	nanometer
NP-40	nonidet P40
NuA3	nucleosome acetyltransferase of H3

NuA4	nucleosome acetyltransferase of H4
NURF	nucleosome remodeling factor
ORF	open reading frame (coding region)
P	phosphorylation
PCR	polymerase chain reaction
PHD	plant homeodomain
PMSF	phenylmethylsufonyl fluoride
PVDF	polyvinylidene difluoride
R	arginine
RNA	ribonucleic acid
RNAPII	RNA polymerase II
RSC	remodel the structure of chromatin
S	serine
SAGA	Spt-Ada-Gcn5-acetyltransferase
SAM	S-adenosyl-methionine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SC-HIS	synthetic complete lacking histidine
SC-URA	synthetic complete lacking uracil
SD-URA	synthetic defined lacking uracil
Ser	serine
SET	Su(var)r3-9, En(zeste), and Trithorax
SWI/SNF	mutants defective in mating type switching (swi) and sucrose fermentation (snf)
T	threonine

TAP	tandem affinity purification
UTR	untranslated region
WCE	whole-cell extract
WT	wild-type
Y	tyrosine
YEA	yeast extract supplemented with adenine
YPD	yeast extract-peptone-dextrose
α	against
Δ	delete
µg	microgram
µl	microliter
^{3}H	tritium
µCi	microcuries

CHAPTER 1

INTRODUCTION

Organization and Regulation of Chromatin

In eukaryotes, the regulation of chromatin structure is essential for all DNA-templated processes such as DNA replication, repair and transcription. It is the combination of DNA and proteins in the nucleus of the cell that forms chromatin (Kornberg, 1974). Within this structure, access to the underlying DNA is extremely restrictive which serves as a means of not only compacting and protecting DNA, but as a method of regulating protein-DNA interactions (Kornberg and Lorch, 1999; Zheng and Hayes, 2003). The basic subunit of chromatin, the nucleosome, plays a major role in this compaction. Each nucleosome consists of ~147 bp of DNA wrapped twice around an octamer of histone proteins which contains two copies each of the four core histones H2A, H2B, H3 and H4 (Luger et al., 1997). The core histones consist predominantly of a structured globular domain from which extend highly charged, unstructured basic amino (N)-terminal domain tails. These N-terminal tails are vital for the condensation of chromatin and mediate interactions between and within nucleosomes (Fig. 1.1A). Besides playing a structural role, the involvement of the N-terminal histone tails in the folding of nucleosomes into-higher order structures, such as the 30-nm chromatin fiber and higher-order tertiary structures, play essential roles in the regulation of gene activity by creating permissive (euchromatin) and repressive (heterochromatin) environments for gene expression (Kornberg and Lorch, 1999; Peterson and Laniel, 2004). In the absence of proper

nucleosome assembly, spontaneous DNA damage and loss of cell viability occurs, further supporting the importance of histone-DNA interactions and the formation of chromatin (Gunjan et al., 2005).

With chromatin being inhibitory to DNA-templated processes, the cell has devised several mechanisms to overcome this inhibition. Among these mechanisms is the remodeling of chromatin, which uses the energy of ATP-hydrolysis to shift the position of nucleosomes along genes. Chromatin-remodeling complexes such as SWI/SNF disrupt DNA-histone contacts while ISWI complexes such as NURF and CHRAC use an alternative method of chromatin restructuring involving the sliding of intact nucleosomes (Cairns et al., 1994; Cote et al., 1994; Hamiche et al., 1999; Langst et al., 1999). In either case, chromatin structure is sufficiently altered to allow the binding of factors required for the activation of transcription (Imbalzano et al., 1994).

A second method of regulation, the incorporation of histone variants, has come to light as an important mechanism of chromatin regulation in the past few years. Given their role in the packaging of DNA, histones are among the most highly conserved eukaryotic proteins. However, despite their high conservation, several organisms produce specialized histone variants that can be incorporated into nucleosomes, potentially altering nucleosomal stability. Two such variants exist for the histone H2A, H2A.X and H2A.Z, which are highly conserved from yeast to humans and are involved in DNA double-strand break repair and transcription, respectively (Redon et al., 2002). Variants also exist for the histone H3 protein including the centromeric-specific histone CENP-A/Cse4 that is crucial for kinetochore assembly. H3.3 in mammalian cells is deposited independent of replication at active genes during transcription

and is the only H3 variant in yeast (Ahmad and Henikoff, 2002; Smith, 2002). Interestingly, a variant for the most highly conserved histone, H4, has not been reported.

The third mechanism, the covalent modification of histone proteins, occurs on all four core histone proteins and has been shown to control many aspects of chromatin structure and function. A majority of these marks occur on the N-terminal tails of histone proteins. Given the importance of the N-terminal tails in the folding of nucleosomes (Kan et al., 2007), histone modifications are likely to play a major role in DNA regulation through control of chromatin organization and folding. In this chapter, I will review what was currently known about these modifications at the beginning of my studies with an emphasis on the roles of these marks in the process of transcription.

Post-translational Histone Modifications

With the discovery of DNA and chromatin in the late 19th century (Miescher, 1871; Miescher, 1874), it has only been in the past 40 years that modifications on histone proteins have been identified (Allfrey et al., 1964; Murray, 1964). A number of different highly conserved post-translational modifications have been found to occur on histone proteins and include acetylation of lysine (K) residues, methylation of lysines and arginines (R), phosphorylation of serines (S) and threonines (T), as well as ubiquitylation and sumoylation of lysines. Additionally, lysine residues can be mono-, di-, or trimethylated while arginines can be mono- or dimethylated (symmetric or asymmetric). The majority of these post-translational modifications occur primarily on the basic, highly charged N-terminal tails with some recently found in the globular domains (Fig. 1.1B) (Cosgrove et al., 2004; Zhang et al., 2003). Several of these modifications play roles in cellular processes such as DNA repair,

replication, transcription, and mitosis. They are thought to function in one of two ways either directly by altering chromatin structure by disrupting histone-DNA contacts leading to the "unraveling" of chromatin or indirectly by serving as binding platforms for the recruitment of protein complexes. This second mechanism of action has led to the "histone code hypothesis" in which it has been proposed that the combination of histone modifications act as a code to recruit regulatory proteins that elicit distinct biological effects (Jenuwein and Allis, 2001; Strahl and Allis, 2000).

Of all the different types of post-translational modifications that occur on histones, we have learned the most about lysine acetylation and methylation with the majority of studies focused on acetylation and the dynamics of its regulation. One way acetylation functions is to neutralize the positive charge of lysines, leading to the disruption or loosening of histone-histone interactions and DNA-histone interactions that allows regulatory proteins to gain access to DNA (Sterner and Berger, 2000). Acetylation has been linked to several cellular processes which include DNA repair, chromatin assembly, cell-cycle progression and transcription (Clarke et al., 1999; van Attikum and Gasser, 2005; Wang et al., 1997; Ye et al., 2005). This modification also functions to recruit regulatory proteins that modify chromatin structure to control gene transcription (Kurdistani et al., 2004). Generally, histone acetyltransferases target several lysines for modification and in some cases, are functionally redundant with one another (Howe et al., 2001; Wittschieben et al., 2000). However, acetylation is highly dynamic and can be reversed within minutes by histone deacetylases (Sun et al., 2003; Waterborg, 2001). Similar to acetyltransferases, histone deacetylases target several lysines for the removal of acetyl groups and have a small degree of functional overlap (Robyr et al., 2002).

Unlike, acetylation, methylation of lysines is considered a more stable modification that does not significantly affect the charge of histone proteins, but may primarily function to recruit regulatory proteins. Evidence gathered clearly shows that histone methylation plays roles in the activation and repression of transcription as well as in DNA repair and other cellular processes (Iizuka and Smith, 2003; Lee et al., 2005; Sanders et al., 2004). Additionally, the ability to mono-, di-, or trimethylate lysine residues expands on the possible functions for each methyl state in these different cellular processes. Yet, the biological roles of these modifications, especially methylation, have not been entirely determined.

Gene Transcription

Interestingly, the majority of what we know about the function of histone modifications is in the context of gene transcription. The organization of DNA into chromatin plays a vital role in the regulation of this process. In eukaryotes, transcription of all protein-coding genes is carried out by the 12-subunit RNA polymerase II (RNAPII) complex and is characterized by three main phases: initiation, elongation, and termination. In addition to the requirement of transcription factors for the recruitment of RNAPII's enzymatic activity, regulatory proteins are needed to decompact the inhibitory chromatin structure. Thus, RNAPII serves as a platform for several messenger RNA (mRNA) processing factors and chromatin-modifying enzymes that interact with the polymerase during transcription. A major component of RNAPII responsible for these interactions is the long unstructured C-terminal domain (CTD) on its largest subunit, Rpb1 (Buratowski, 2003; Hahn, 2004).

Of the three RNA polymerases, RNAPII is the only one to contain a CTD, which consists of the repeated heptapeptide sequence YSPTSPS (52 repeats in humans and 26 in yeast). *In*

vivo, the CTD is highly phosphorylated and thought to interact with many proteins. However, all of these factors do not bind to the CTD at the same time. Specifically, phosphorylation of the 5th serine (Ser5) in the heptapeptide repeat by the kinase TFIIH (Kin28) occurs at the 5' end of genes during transcription initiation and recruits such proteins as the capping enzyme complex (Cho et al., 1997; Komarnitsky et al., 2000). Conversely, phosphorylation of serine 2 (Ser2) by the kinase pTEFb/Ctk1 marks transcription elongation and occurs in the coding region and 3' end of genes. Phosphorylation at this residue has been shown to interact with histone methyltransferases, RNA processing, and termination factors (Ahn et al., 2004; Hampsey and Reinberg, 2003; Komarnitsky et al., 2000). Consequently, these different phosphorylation events distinguish between the different phases of transcription and regulate protein recruitment.

Histone Modifications in Gene Transcription

Although the modifications of histones have been linked to the regulation of gene expression, it has only been in the past decade that functions for these marks in transcription have begun to be revealed. The identification of the first histone acetyltransferase (HAT), Gcn5, which was originally identified as a transcription co-activator, gave the first hint of a possible role for histone acetylation in gene activation (Brownell et al., 1996). Interestingly, the majority of these histone-modifying enzymes exist in multiprotein complexes that regulate not only their histone and residue specificity, but also their targeting to genes. In a complex with other proteins (i.e. SAGA complex), Gcn5 has been shown to target its histone acetyltransferase activity to the promoters of genes through an interaction with gene-specific activators such as Gcn4 (Brown et al., 2001; Grant et al., 1999). Several studies suggest that

targeted recruitment of HAT complexes by activators to promoter regions leads to hyperacetylation and transcriptional activation while recruitment of histone deacetylase (HDAC) complexes to promoters by repressors leads to hypoacetylation and transcriptional repression (Fig. 1.2) (Brown et al., 2001; Cosma, 2002; Kurdistani et al., 2002; Robert et al., 2004). This has clearly been demonstrated not only by SAGA, but also the histone deacetylase Rpd3 which, in complex with the co-repressor Sin3, is targeted to the promoters of genes (Kadosh and Struhl, 1997; Kasten et al., 1997). In addition to being targeted to promoters, HATs and HDACs may modify histones in an untargeted, global manner regulating the general levels of acetylation genome-wide (Vogelauer et al., 2000). Collectively, these studies highlight the association of histone acetylation and deacetylation with the regulation of gene activity.

While histone acetylation/deacetylation is correlated with the state of gene activity, this correlation may not be due to direct effects on chromatin structure. In many cases, histone modifications are involved in the regulation of chromatin structure by acting as binding platforms for the recruitment of protein complexes as has been proposed by the "histone code hypothesis". Specifically, bromodomain-containing proteins can recognize and bind acetyl marks (Kasten et al., 2004; Martinez-Campa et al., 2004; Matangkasombut et al., 2000). Several such proteins have been found in chromatin-remodeling complexes such as SWI/SNF and RSC, thus indicating a strong link between the recruitment of chromatin-remodeling activities during transcriptional activation and histone acetylation (Agalioti et al., 2002; Hassan et al., 2002; Kasten et al., 2004). Additionally, bromodomains have been found in acetyltransferases themselves, such as Gcn5 and can function to stabilize the binding of other chromatin-remodeling activities (Syntichaki et al., 2000).

Similar to bromodomain-containing proteins, chromodomain-containing proteins have been found to bind to histone methyl marks. The best example of this interaction is that of the chromodomain of the heterochromatin protein HP1 which specifically binds to methylated H3K9 to regulate transcriptional silencing and the formation of heterochromatin (Bannister et al., 2001; Lachner et al., 2001). Although modifications such as H3K9 methylation appear to function as repressive marks, not all methylation events are involved in transcriptional repression. Methylation at K4 on histone H3, specifically trimethylation, localizes to sites of transcriptional activation (Santos-Rosa et al., 2002; Schneider et al., 2004). In mammalian cells, H3K4 methylation prevents the association of the repressive NuRD complex, presumably preventing the methylation of H3K9 (Zegerman et al., 2002). However, these positive and negative functional mechanisms of methylation are specific to multicellular eukaryotes. In budding yeast, the functions of these modifications are less defined. The chromodomain of the chromatin-remodeler Chd1, as a component of the SAGA complex, was found to specifically interact with methylated H3K4 (Pray-Grant et al., 2005). However, this finding has proven to be controversial as subsequent research suggests it is the human Chd1 and not the yeast protein that may interact with H3K4 methylation (Flanagan et al., 2005; Sims et al., 2005).

Indeed, the other three chromodomain-containing proteins in yeast, New1, Esa1, and Eaf3 have the potential to function through the binding of methyl marks. However, at the initiation of the studies carried out in this dissertation, the functional mechanisms of these proteins had not been identified. Furthermore, several other potential chromatin-regulating domains were beginning to emerge as potential binders of methylated residues, such as the tudor and WD40-repeat domains, suggesting that there may be a wide range of domains and

associated complexes that may "read" histone modifications (Huyen et al., 2004; Sanders et al., 2004; Wysocka et al., 2005).

Histone Methylation and SET Domain-containing Proteins

With the discovery of the enzymes responsible for histone methylation, we have begun to learn about the functions of these marks in transcription. In eukaryotes, there are at least five lysine residues on the histone H3 and H4 tails that can be methylated (Lachner et al., 2003). Of these methyl marks, the methylation of H3K4, H3K36, and H3K79 have been linked to the activation of transcription while methylation of H3K9, H3K27, and H4K20 have been linked to gene repression (Lachner et al., 2003; Sims et al., 2003). The first histone lysine methyltransferases identified were the mammalian Suv39h1 proteins, which methylate histone H3K9 creating a platform for the recruitment of heterochromatin protein (HP1) (Bannister et al., 2001; Lachner et al., 2001; Rea et al., 2000). In the Suv39h1 proteins, it was discovered that methyltransferase activity is mediated through the highly conserved SET domain, a motif of approximately 130 amino acids initially identified in the three *Drosophila* genes, the suppressor of position effect variegation Su(var)3-9, the *Polycomb*-group protein En(zeste), and the homeotic gene regulator Trithorax (Jenuwein et al., 1998). For members of the Suv39h family, the SET domain alone is not sufficient for catalytic activity, but requires adjacent cysteine-rich domains (pre-SET and post-SET) suggesting that these regions may be required for the activity of other SET domain-containing methyltransferases (Rea et al., 2000). Yet, several proteins that are missing one or both of these domains have been found to contain histone methyltransferase activity leading to the identification of a series of SET domain proteins and methyltransferase families (Kouzarides, 2002).

In addition to SUV39H1, there are more than 70 gene sequences containing SET domains in mammals (Rea et al., 2000). In contrast, the budding yeast *Saccharomyces cerevisiae* contains seven SET domain-containing genes (Schultz et al., 1998). Interestingly, only methylation at H3K4, H3K36, and H3K79 are found in yeast (Perrod and Gasser, 2003). With the exception of the non-SET domain-containing Dot1 protein which methylates H3K79, both H3K4 and H3K36 are methylated by the SET proteins Set1 and Set2, respectively (Briggs et al., 2001; Feng et al., 2002; Strahl et al., 2002; van Leeuwen et al., 2002). Set1, as part of the COMPASS complex, methylates H3K4 at the 5' ends of genes (Briggs et al., 2001; Krogan et al., 2002; Miller et al., 2001; Nagy et al., 2002; Roguev et al., 2001). Consistent with the location of this histone mark, Set1 interacts with the Ser5 phosphorylated form of RNAPII (Krogan et al., 2003a; Ng et al., 2003). Methylation of H3K4, as previously described, prevents the recruitment of a repressive complex in mammals and recruits chromatin modifying activities, thereby linking its function to the activation of transcription. In yeast, methylation of this site has been implicated in both transcription activation and gene silencing dependent on the genomic location of this modification and methyl state (Bernstein et al., 2002; Briggs et al., 2001; Bryk et al., 2002; Kouzarides, 2002; Santos-Rosa et al., 2002; Zegerman et al., 2002). How H3K4 methylation is involved in activation and silencing is currently not understood.

Set2 and Histone H3 Lysine 36 Methylation

As a founding member of the Set2 family of proteins, Set2 is the first methyltransferase identified to specifically methylate nucleosomal H3K36 in yeast (Strahl et al., 2002). Members of this family are characterized by an AWS domain preceding the SET domain

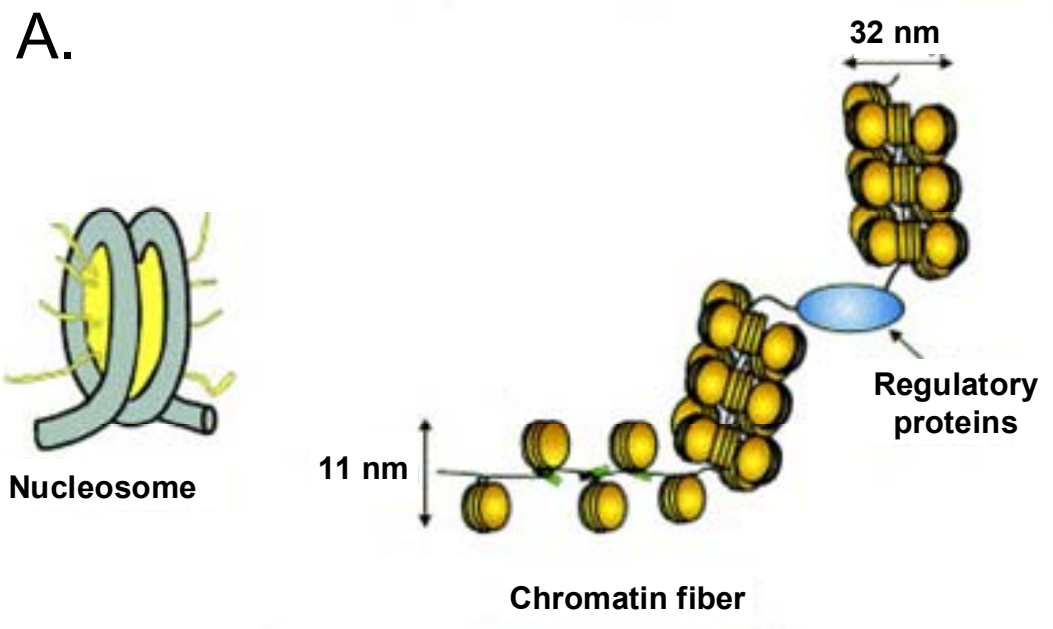
which is followed by a post-SET domain. Similar to the pre-SET domain in the Suv39h family, the AWS domain is a cysteine-rich sequence which may play a role in enzymatic activity. Based on sequence similarity, other members of this family include the mammalian proteins NSD1, NSD2, NSD3, and HIF1 (HYPB) as well as the *Drosophila* protein Ash1. Ash1 has been demonstrated to be a multicatalytic methyltransferase that activates transcription by methylating H3K4, H3K9, and H4K20 (Beisel et al., 2002). The NSD proteins are highly related to one another and have all been implicated in the development of human cancers, suggesting a role for these proteins in cell growth and differentiation (Angrand et al., 2001; Jaju et al., 2001; Stec et al., 1998). Besides Set2, NSD1 can methylate H3K36 as well as H4K20, but this activity has only been detected *in vitro* (Rayasam et al., 2003). The methyltransferase activities of NSD2, NSD3, and HIF1 are not known, but based on structural similarity they may also modify H3K36.

Recently, studies have demonstrated that a domain in the C-terminus of budding yeast Set2 associates with the Ser2 phosphorylated form of RNAPII during transcription elongation suggesting that Set2 may play a role in the transcription elongation process (Kizer et al., 2005; Krogan et al., 2003b; Li et al., 2003; Li et al., 2002; Schaft et al., 2003; Xiao et al., 2003). Mutations of the RNAPII CTD at the Ser2 position or deletion of the enzyme, Ctk1, responsible for phosphorylating the CTD during elongation abolishes this interaction as well as Set2-mediated H3K36 methylation (Li et al., 2003; Xiao et al., 2003). These results indicate that Set2 activity is targeted to the body of genes through an RNAPII interaction (Fig. 1.3). Consistent with this idea, evidence shows that Set2 and H3K36 methylation associate with the coding and 3' ends of genes (Krogan et al., 2003b; Schaft et al., 2003; Xiao et al., 2003). Additionally, Set2 has been found to genetically interact with a number of

other elongation factors, thus further supporting a role for its activities in transcription elongation (Krogan et al., 2003b). However, how Set2 functions in the transcription elongation process was not known. One clue comes from findings that Set2 and H3K36 methylation can repress transcription. In one case, Set2 was artificially tethered to a test promoter and this led to strong transcription repression (Strahl et al., 2002). In another case, deletion of Set2 specifically relieved the basal repression of the *GAL4* gene in yeast (Landry et al., 2003). In both cases, the repression by Set2 was at least partially dependent on having an active SET domain.

Although data suggests that Set2 is involved in transcription, the exact roles of Set2 and H3K36 methylation in gene expression remain unclear. In the following chapters, I present data describing novel functions for Set2 and H3K36 methylation in transcription. Chapter 2 describes my work on the conservation of Set2 and H3K36 methylation. In this study, I show that Set2-mediated methylation of H3K36 during transcription elongation occurs in a number of organisms outside of budding yeast. This work is the first to show a conserved interaction between Set2 and the elongating polymerase. Chapter 3, as part of a much larger report, describes my contributions to the first identified role for Set2-mediated H3K36 methylation. I will present evidence to show that Set2 regulates global histone acetylation patterns in the coding region of genes by recruiting a small Rpd3 deacetylase (Rpd3S) complex. Finally, in chapter 4, I demonstrate that H3K36 is a site of acetylation independent of its methylation and suggest that both modifications play distinct roles in the process of transcription.

Figure 1.1


A.
32 nm
Regulatory proteins
11 nm
Nucleosome
Chromatin fiber

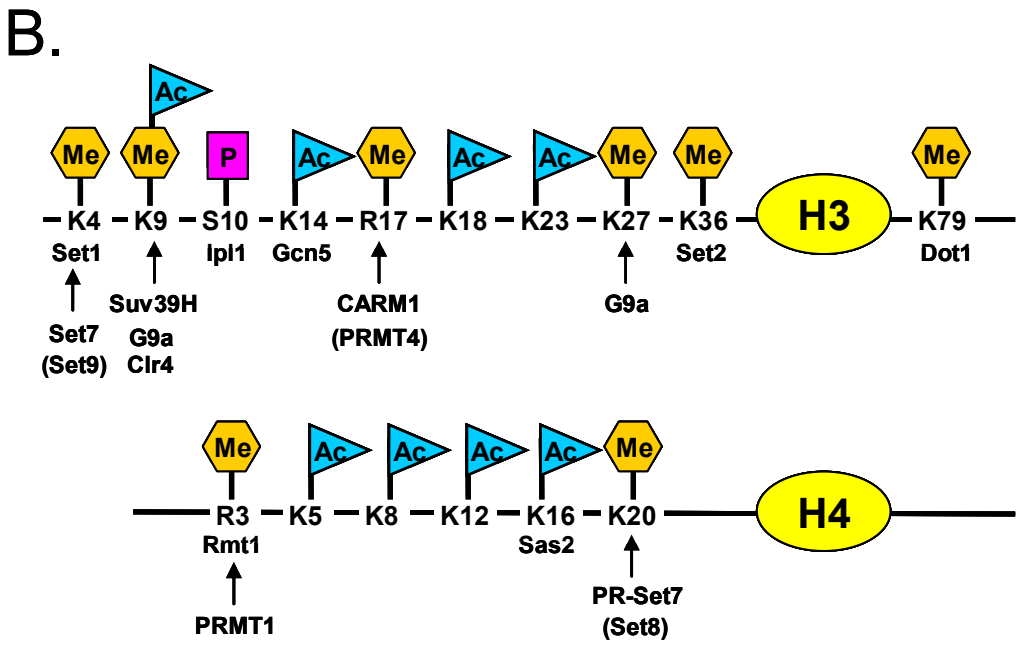

B.
Ac
Me
Me
P
Ac
Me
Ac
Ac
Me
Me
Me
K4
K9
S10
K14
R17
K18
K23
K27
K36
H3
K79
Set1
Ipl1
Gcn5
Dot1
Set2
Set7
(Set9)
Suv39H
G9a
Clr4
CARM1
(PRMT4)
G9a
Me
Ac
Ac
Ac
Ac
Me
R3
K5
K8
K12
K16
K20
H4
Rmt1
Sas2
PRMT1
PR-Set7
(Set8)

Figure 1.1. Chromatin and the variety of post-translational histone modifications. *(A)* Arrangement of nucleosomes into chromatin and folding into higher-order structures. Pictured (left) is the nucleosome which consists of ~147 bp of DNA wrapped around a histone octamer core from which extend the highly-charged N-terminal domain tails. To the right is an illustration of the compaction of the 11-nm chromatin fiber into a 32-nm fiber interspersed with regulatory proteins (reprinted, with modification, from Morales et al., 2001). *(B)* Overview of selected modifications on the N-terminal tails of histones H3 and H4. Post-translational modifications include methylation (Me), acetylation (Ac), and phosphorylation (P). Proteins in *Saccharomyces cerevisiae* that modify the indicated modifications are denoted immediately below each residue, while proteins from other organisms are below arrows directed towards the appropriate targets. Alternate protein names are indicated in parentheses.

Figure 1.2

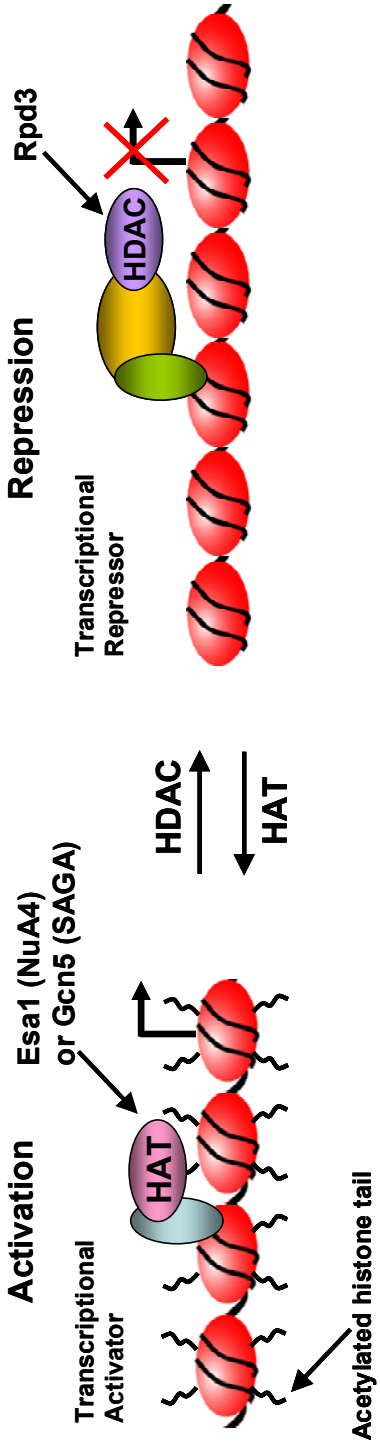

Repression
Rpd3
HDAC
Transcriptional
Repressor
HDAC
HAT
Activation
Esa1 (NuA4)
or Gcn5 (SAGA)
HAT
Transcriptional
Activator
Acetylated histone tail

Figure 1.2. Targeting of histone acetyltransferases and deacetylases during transcription. The targeting of histone acetyltransferases (HAT) such as Esa1 and Gcn5 to the promoters of genes during transcription by co-activators. Acetylation of histones by HATs results in the disruption of nucleosomes and gene activation. The recruitment of histone deacetylases (HDAC) such as Rpd3 by co-repressors leads to transcription repression by removing acetylation marks from histones in the promoters of genes.

Figure 1.3

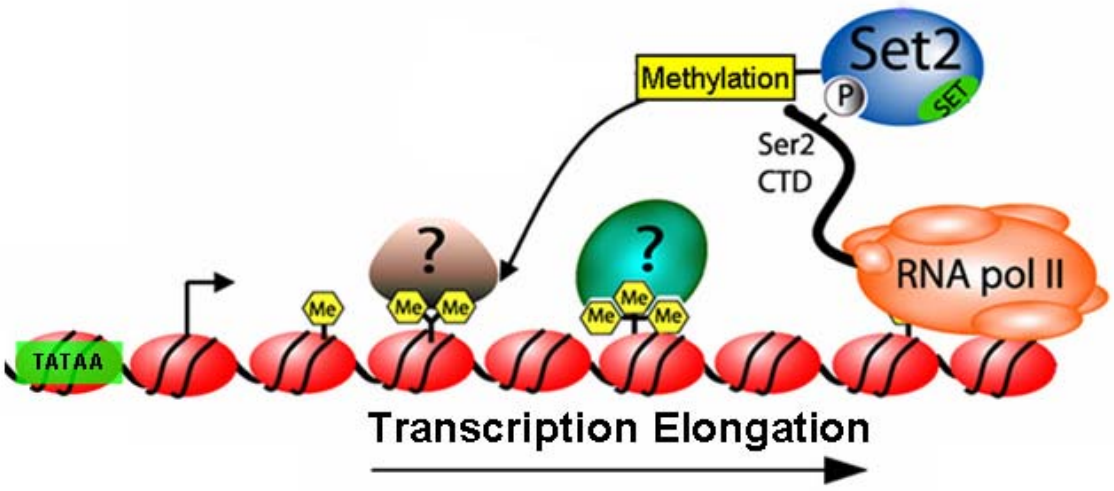

Methylation
Set2
P
SET
Ser2
CTD
RNA pol II
?
?
Me
Me
Me
Me
Me
Me
TATAA
Transcription Elongation

Figure 1.3. Model of Set2 regulation. Interaction between the Serine 2 phosphorylated C-terminal domain (CTD) of RNA polymerase II (RNA Pol II) and Set2 targets the methyltransferase activity of Set2 to the coding region of genes during transcription elongation. Set2-mediated mono-, di-, or trimethylation of H3K36 may then subsequently serve as recognition marks for the recruitment of protein modules.

Acknowledgements

The image in Figure 1.1A has been reprinted, with modification, from Biochimie, vol. 83, "Chromatin structure and dynamics: Functional implications", pp 1029-1039, Copyright (2001), with permission from Elsevier. The authors of this work included Violette Morales, Claire Giamarchi, Catherine Chaileux, Françoise Moro, Véronique Marsud, Sophie Le Ricousse, and Hélenè Richard-Foy.

References

Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. Cell *111*, 381-392.

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell *9*, 1191-1200.

Ahn, S. H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell *13*, 67-76.

Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). Acetylation And Methylation Of Histones And Their Possible Role In The Regulation Of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.

Angrand, P. O., Apiou, F., Stewart, A. F., Dutrillaux, B., Losson, R., and Chambon, P. (2001). NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. Genomics *74*, 79-88.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. (2002). Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature *419*, 857-862.

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 U S A *99*, 8695-8700.

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*, 3286-3295.

Brown, C. E., Howe, L., Sousa, K., Alley, S. C., Carrozza, M. J., Tan, S., and Workman, J. L. (2001). Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. Science *292*, 2333-2337.

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*, 843-851.

Bryk, M., Briggs, S. D., Strahl, B. D., Curcio, M. J., Allis, C. D., and Winston, F. (2002). Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in S. cerevisiae by a Sir2-independent mechanism. Curr Biol *12*, 165-170.

Buratowski, S. (2003). The CTD code. Nat Struct Biol *10*, 679-680.

Cairns, B. R., Kim, Y. J., Sayre, M. H., Laurent, B. C., and Kornberg, R. D. (1994). A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. Proc Natl Acad Sci U S A *91*, 1950-1954.

Cho, E. J., Takagi, T., Moore, C. R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev *11*, 3319-3326.

Clarke, A. S., Lowell, J. E., Jacobson, S. J., and Pillus, L. (1999). Esa1p is an essential histone acetyltransferase required for cell cycle progression. Mol Cell Biol *19*, 2515-2526.

Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat Struct Mol Biol *11*, 1037-1043.

Cosma, M. P. (2002). Ordered recruitment: gene-specific mechanism of transcription activation. Mol Cell *10*, 227-236.

Cote, J., Quinn, J., Workman, J. L., and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science *265*, 53-60.

Feng, Q., Wang, H., Ng, H. 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*, 1052-1058.

Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature *438*, 1181-1185.

Grant, P. A., Eberharter, A., John, S., Cook, R. G., Turner, B. M., and Workman, J. L. (1999). Expanded lysine acetylation specificity of Gcn5 in native complexes. J Biol Chem *274*, 5895-5900.

Gunjan, A., Paik, J., and Verreault, A. (2005). Regulation of histone synthesis and nucleosome assembly. Biochimie *87*, 625-635.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol *11*, 394-403.

Hamiche, A., Sandaltzopoulos, R., Gdula, D. A., and Wu, C. (1999). ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. Cell *97*, 833-842.

Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell *113*, 429-432.

Hassan, A. H., Prochasson, P., Neely, K. E., Galasinski, S. C., Chandy, M., Carrozza, M. J., and Workman, J. L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell *111*, 369-379.

Howe, L., Auston, D., Grant, P., John, S., Cook, R. G., Workman, J. L., and Pillus, L. (2001). Histone H3 specific acetyltransferases are essential for cell cycle progression. Genes Dev *15*, 3144-3154.

Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., 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*, 406-411.

Iizuka, M., and Smith, M. M. (2003). Functional consequences of histone modifications. Curr Opin Genet Dev *13*, 154-160.

Imbalzano, A. N., Kwon, H., Green, M. R., and Kingston, R. E. (1994). Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature *370*, 481-485.

Jaju, R. J., Fidler, C., Haas, O. A., Strickson, A. J., Watkins, F., Clark, K., Cross, N. C., Cheng, J. F., Aplan, P. D., Kearney, L.*, et al.* (2001). A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. Blood *98*, 1264-1267.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol Life Sci *54*, 80-93.

Kadosh, D., and Struhl, K. (1997). Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. Cell *89*, 365-371.

Kan, P. Y., Lu, X., Hansen, J. C., and Hayes, J. J. (2007). The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. Mol Cell Biol *27*, 2084-2091.

Kasten, M., Szerlong, H., Erdjument-Bromage, H., Tempst, P., Werner, M., and Cairns, B. R. (2004). Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. Embo J *23*, 1348-1359.

Kasten, M. M., Dorland, S., and Stillman, D. J. (1997). A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators. Mol Cell Biol *17*, 4852-4858.

Kizer, K. O., Phatnani, H. P., Shibata, Y., Hall, H., Greenleaf, A. L., and Strahl, B. D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol *25*, 3305-3316.

Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev *14*, 2452-2460.

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R. D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285-294.

Kouzarides, T. (2002). Histone methylation in transcriptional control. Curr Opin Genet Dev *12*, 198-209.

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. J Biol Chem *277*, 10753-10755.

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

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C.*, et al.* (2003b). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Kurdistani, S. K., Robyr, D., Tavazoie, S., and Grunstein, M. (2002). Genome-wide binding map of the histone deacetylase Rpd3 in yeast. Nat Genet *31*, 248-254.

Kurdistani, S. K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. Cell *117*, 721-733.

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*, 116-120.

Lachner, M., O'Sullivan, R. J., and Jenuwein, T. (2003). An epigenetic road map for histone lysine methylation. J Cell Sci *116*, 2117-2124.

Landry, J., Sutton, A., Hesman, T., Min, J., Xu, R. M., Johnston, M., and Sternglanz, R. (2003). Set2-catalyzed methylation of histone H3 represses basal expression of GAL4 in Saccharomyces cerevisiae. Mol Cell Biol *23*, 5972-5978.

Langst, G., Bonte, E. J., Corona, D. F., and Becker, P. B. (1999). Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. Cell *97*, 843-852.

Lee, D. Y., Teyssier, C., Strahl, B. D., and Stallcup, M. R. (2005). Role of protein methylation in regulation of transcription. Endocr Rev *26*, 147-170.

Li, B., Howe, L., Anderson, S., Yates, J. R., 3rd, and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *278*, 8897-8903.

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. J Biol Chem *277*, 49383-49388.

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

Martinez-Campa, C., Politis, P., Moreau, J. L., Kent, N., Goodall, J., Mellor, J., and Goding, C. R. (2004). Precise nucleosome positioning and the TATA box dictate requirements for the histone H4 tail and the bromodomain factor Bdf1. Mol Cell *15*, 69-81.

Matangkasombut, O., Buratowski, R. M., Swilling, N. W., and Buratowski, S. (2000). Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. Genes Dev *14*, 951-962.

Miescher, F. (1871). Ueber die chemische Zusammen setzung der Eiterzellen. Med-Chem Unters *4*, 441-460.

Miescher, F. (1874). Die Spermatozoen einiger Wirbelthiere. Ein Beitrag zur Histochemie. Verh Natforsch Ges Basel *6*, 138-208.

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 U S A *98*, 12902-12907.

Morales, V., Giamarchi, C., Chailleux, C., Moro, F., Marsaud, V., Le Ricousse, S., and Richard-Foy, H. (2001). Chromatin structure and dynamics: functional implications. Biochimie *83*, 1029-1039.

Murray, K. (1964). The Occurrence Of Epsilon-N-Methyl Lysine In Histones. Biochemistry *127*, 10-15.

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 U S A *99*, 90-94.

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*, 709-719.

Perrod, S., and Gasser, S. M. (2003). Long-range silencing and position effects at telomeres and centromeres: parallels and differences. Cell Mol Life Sci *60*, 2303-2318.

Peterson, C. L., and Laniel, M. A. (2004). Histones and histone modifications. Curr Biol *14*, R546-551.

Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., and Grant, P. A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434-438.

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. Embo J *22*, 3153-3163.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K., and Bonner, W. (2002). Histone H2A variants H2AX and H2AZ. Curr Opin Genet Dev *12*, 162-169.

Robert, F., Pokholok, D. K., Hannett, N. M., Rinaldi, N. J., Chandy, M., Rolfe, A., Workman, J. L., Gifford, D. K., and Young, R. A. (2004). Global position and recruitment of HATs and HDACs in the yeast genome. Mol Cell *16*, 199-209.

Robyr, D., Suka, Y., Xenarios, I., Kurdistani, S. K., Wang, A., Suka, N., and Grunstein, M. (2002). Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. Cell *109*, 437-446.

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. Embo J *20*, 7137-7148.

Sanders, S. L., Portoso, M., Mata, J., Bahler, J., Allshire, R. C., and Kouzarides, T. (2004). Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell *119*, 603-614.

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

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*, 2475-2482.

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*, 73-77.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A *95*, 5857-5864.

Sims, R. J., 3rd, Chen, C. F., Santos-Rosa, H., Kouzarides, T., Patel, S. S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem *280*, 41789-41792.

Sims, R. J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends Genet *19*, 629-639.

Smith, M. M. (2002). Centromeres and variant histones: what, where, when and why? Curr Opin Cell Biol *14*, 279-285.

Stec, I., Wright, T. J., van Ommen, G. J., de Boer, P. A., van Haeringen, A., Moorman, A. F., Altherr, M. R., and den Dunnen, J. T. (1998). WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphy gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. Hum Mol Genet *7*, 1071-1082.

Sterner, D. E., and Berger, S. L. (2000). Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev *64*, 435-459.

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

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*, 1298-1306.

Sun, J. M., Spencer, V. A., Chen, H. Y., Li, L., and Davie, J. R. (2003). Measurement of histone acetyltransferase and histone deacetylase activities and kinetics of histone acetylation. Methods *31*, 12-23.

Syntichaki, P., Topalidou, I., and Thireos, G. (2000). The Gcn5 bromodomain co-ordinates nucleosome remodelling. Nature *404*, 414-417.

van Attikum, H., and Gasser, S. M. (2005). The histone code at DNA breaks: a guide to repair? Nat Rev Mol Cell Biol *6*, 757-765.

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

Vogelauer, M., Wu, J., Suka, N., and Grunstein, M. (2000). Global histone acetylation and deacetylation in yeast. Nature *408*, 495-498.

Wang, L., Mizzen, C., Ying, C., Candau, R., Barlev, N., Brownell, J., Allis, C. D., and Berger, S. L. (1997). Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. Mol Cell Biol *17*, 519-527.

Waterborg, J. H. (2001). Dynamics of histone acetylation in Saccharomyces cerevisiae. Biochemistry *40*, 2599-2605.

Wittschieben, B. O., Fellows, J., Du, W., Stillman, D. J., and Svejstrup, J. Q. (2000). Overlapping roles for the histone acetyltransferase activities of SAGA and elongator in vivo. Embo J *19*, 3060-3068.

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*, 859-872.

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*, 654-663.

Ye, J., Ai, X., Eugeni, E. E., Zhang, L., Carpenter, L. R., Jelinek, M. A., Freitas, M. A., and Parthun, M. R. (2005). Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. Mol Cell *18*, 123-130.

Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. J Biol Chem *277*, 11621-11624.

Zhang, L., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. Chromosoma *112*, 77-86.

Zheng, C., and Hayes, J. J. (2003). Structures and interactions of the core histone tail domains. Biopolymers *68*, 539-546.

CHAPTER 2

CONSERVATION OF SET2-MEDIATED HISTONE H3 LYSINE 36 METHYLATION

Abstract

Set2 methylation of histone H3 at lysine 36 (H3K36) has recently been shown to be associated with RNA polymerase II (RNAPII) elongation in *Saccharomyces cerevisiae*. However, whether this modification is conserved and associated with transcription elongation in other organisms is not known. Here we report the identification and characterization of the Set2 ortholog responsible for H3K36 methylation in the fission yeast *Schizosaccharomyces pombe*. We find that similar to the budding yeast enzyme, *S. pombe* Set2 is also a robust nucleosomal-selective H3 methyltransferase that is specific for K36. Deletion of the *S. pombe set2*$^{+}$ gene results in a complete abolishment of H3K36 methylation as well as a slow-growth phenotype on plates containing synthetic medium. These results indicate that Set2 is the sole enzyme responsible for this modification in fission yeast and is important for cell growth under stressed conditions. Using the chromatin immunoprecipitation (ChIP) assay, we demonstrate that H3K36 methylation in *S. pombe* is associated with the transcribed regions of RNAPII-regulated genes and is absent from regions that are not transcribed by RNAPII. Consistent with a role for Set2 in transcription elongation, we find that *S. pombe* Set2 associates with the hyperphosphorylated form of

RNAPII and can fully rescue H3K36 methylation and RNAPII interaction in budding yeast cells deleted for Set2. These results, along with our finding that H3K36 methylation is highly conserved among eukaryotes, imply a conserved role for this modification in the transcription elongation process.

Introduction

Covalent histone modifications represent a major mechanism by which cells regulate the structure and function of chromatin. A number of different post-translational modifications are known to occur on histones, including acetylation, methylation, phosphorylation, ubiquitylation and more recently, sumoylation (Berger, 2002; Holde, 1988; Peterson and Laniel, 2004; Shiio and Eisenman, 2003). While the majority of these modifications are restricted to the flexible N- and C-terminal 'tail' domains of these proteins, a significant number of these modifications have been identified in their highly structured globular domains (Cosgrove et al., 2004; Zhang et al., 2003). The function of these modifications are not well understood, but it is becoming increasingly clear that they coordinate their effects in the form of a histone code to regulate the complex and diverse activities associated with DNA in chromatin (Jenuwein and Allis, 2001; Strahl and Allis, 2000; Turner, 2000).

A large body of work now shows that histone methylation plays a key role in the regulation of chromatin structure and function. In particular, studies show that the methylation of lysine and/or arginine residues regulates diverse cellular functions such as transcriptional repression and activation, heterochromatin formation, X-inactivation and polycomb-mediated gene silencing (Cao and Zhang, 2004; Grewal and Rice, 2004; Iizuka and Smith, 2003; Kouzarides, 2002; Lee et al., 2004; Zhang and Reinberg, 2001). More recently, studies have revealed an unexpected role for histone methylation in the process of transcription elongation by RNA polymerase II (RNAPII). In the budding yeast *Saccharomyces cerevisiae*, the histone methyltransferases Set1 and Set2, which catalyze H3 lysine 4 (H3K4) and lysine 36 (H3K36) methylation, respectively, have been found to be associated with the elongation competent form of RNAPII (Gerber and Shilatifard, 2003;

Hampsey and Reinberg, 2003). While Set1 association is dependent on the Kin28 kinase, which phosphorylates the serine 5 (Ser5) position of the C-terminal domain (CTD) of RNAPII (Gerber and Shilatifard, 2003; Hampsey and Reinberg, 2003; Krogan et al., 2003a; Ng et al., 2003), Set2 association and methylation is dependent on Ctk1, which phosphorylates the serine 2 position of the CTD (Krogan et al., 2003b; Li et al., 2003; Li et al., 2002; Schaft et al., 2003; Xiao et al., 2003). Although the precise function of these enzyme associations with RNAPII is still unclear, it is believed that H3K4 and H3K36 methylation function in the elongation process at different stages of the transcription elongation cycle (Shilatifard, 2004; Sims et al., 2004).

To date, the association of Set2 with elongating RNAPII has only been demonstrated in *S. cerevisiae*. Whether this enzyme has a conserved role and associates with RNAPII in other organisms is not known. In the fission yeast *Schizosaccharomyces pombe*, Set1-mediated H3K4 methylation is preferentially enriched at the euchromatic loci, in particular at the regions containing open reading frames (Noma et al., 2001; Noma and Grewal, 2002). Moreover, a potential ortholog of the *S. cerevisiae* Set2 has been previously identified in *S. pombe* (Noma and Grewal, 2002). In this report, we characterize the fission yeast *S. pombe* Set2 (SpSet2) and find that this enzyme is a robust H3K36 methyltransferase that mediates nucleosomal-selective methylation. Similar to what is found in budding yeast, H3K36 methylation in *S. pombe* is restricted to the coding region of active genes, and we show that the SpSet2 enzyme interacts with RNAPII and restores H3K36 methylation in *S. cerevisiae* when the endogenous *SET2* gene is deleted. These studies, and the fact that H3K36 methylation is conserved across eukaryotes, suggest a highly conserved role for H3K36 methylation in transcription.

Material and Methods

Yeast Strains

The *S. pombe* yeast strains used in these studies are SP1173 (wild-type; *h- leu1-32 his2 ura4 ade6-216*) and SPK549 (*set2Δ*; *h+ leu1-32 ura4 set2::kanMX6 cen1::ura4 ade6-210*). For the growth assays, *S. pombe* strains SPK131 (wild type; *h- leu1-32 his2 ura4 Rint2::ura4 ade6-216*) and SPK612 (*set2Δ*; *h- leul-32 his2 ura4 Rint2::ura4 set2:kanMX6 ade6-210*) were used. The *set2Δ* strain was constructed by a PCR-based method using a *kanMX6* module to replace the *S. pombe set2*$^{+}$ gene as described (Bahler et al., 1998). Deletion was confirmed by PCR and Southern blot analysis. SP1173 was used to genomically 3XFLAG epitope tag *S. pombe set2*$^{+}$ as described (Bahler et al., 1998) producing the SpSet2-3FLAG strain SPK653 (*h+ leu1-32 ura4 set2-3XFLAG-kan cen1::ura4 ade6-210*). *S. cerevisiae* wild-type and *set2Δ* strains in the BY4742 background were obtained from Research Genetics.

Histone Preparation

Histones were prepared as previously described (Strahl et al., 2001). Briefly, nuclei were isolated by detergent lysis and low-speed centrifugation from 293T cells grown at 37 ºC in Dulbeco's modified Eagle's medium (DMEM) containing 5% fetal bovine serum (FBS). Histones were extracted from nuclei by either DNase I or acid extraction. Wild-type *Tetrahymena thermophila* was grown in enriched 1% proteose peptone, and macronuclear histones were isolated from vegetatively growing cells (Strahl et al., 2001). Nuclei were isolated from wild-type *S. cerevisiae* cells grown in yeast extract-peptone-dextrose (YPD) medium and histones were acid-extracted from isolated nuclei as described (Strahl et al., 2001).

Sequence Analysis (Set2 Homology Searching)

Database searches and protein sequence identifications were performed with BLAST (Altschul et al., 1990) and PSI-BLAST (Altschul et al., 1997). Sequences were aligned and a phylogeny tree was calculated using the neighbor-joining method in the program Clustal X (Jeanmougin et al., 1998). The resulting dendrogram was displayed using the tree drawing program NJplot (Perriere and Gouy, 1996).

Cloning of S. pombe SET2 and Generation of Expression Constructs

Using primers specific to the ORF of *set2*$^+$ in *S. pombe* (*S. pombe* GeneDB ID: SPAC29B12.02c), full-length *set2*$^+$ with a C-terminal FLAG epitope tag inserted just before the stop codon was PCR amplified from genomic DNA. The resulting product was cloned into either the pCAL-n (Stratagene) bacterial or the PN823 yeast expression plasmids. The SpSet2 coding region was sequenced for accuracy. The resulting SpSet2-PN823 (SpSet2-FLAG) plasmid, which is driven by the *ADH1* promoter, was transformed into the BY4742 *S. cerevisiae set2Δ* strain. As a control, the PN823 plasmid without the ORF of *set2*$^+$ (empty vector) and ScSet2-FLAG were transformed into wild-type and *set2Δ* BY4742 strains. Wild-type full-length ScSet2-FLAG bacterial and yeast expression constructs have been previously described (Bryk et al., 2002). Transformants were selected on synthetic complete (SC)-Ura plates.

Expression of Recombinant SpSet2

Plasmids expressing either SpSet2-FLAG or empty vector were transformed into BL21 (DE3) cells. Five ml cell cultures were grown to an optical density (OD_{600}) of 0.8-1.0 in Luria Broth (LB) media supplemented with ampicillin (100 µg/ml), followed by addition of 1 mM isopropyl-D-thiogalactoside (IPTG) for 3 h at 30 ºC. Harvested cells were resuspended

in 600 µl lysis buffer (50 mM Tri-HCl, pH 8.0, 0.1% TritonX-1000, 350 mM NaCl, 10% glycerol, 1 mg/ml lysozyme, 1 mM phenylmethylsulfonyl fluoride [PMSF], and 2 µg/ml leupeptin, aprotinin, and pepstatin). Lysates were prepared by sonication as previously described (Bryk et al., 2002).

In Vitro Histone Methyltransferase Assays

Histone methylation assays were performed as previously described with minor modifications (Bryk et al., 2002). Briefly, 1 µl of bacterial lysate was incubated with either 1.25 µg recombinant H3 (rH3), 5 µg chicken core histones, 5 µg chicken oligonucleosomes, or 5 µg H3 synthetic peptide along with 1 µCi S-Adenosyl-L-[methyl-^{3}H]methionine (^{3}H-SAM, 69.8 Ci/mmol, Amersham Biosciences) in methyltransferase buffer (final concentration 50 mM Tris, pH 9.0, 10% glycerol, 1 mM PMSF, and 2 µg/ml leupeptin, pepstatin, and aprotinin) for 30 min at 30 ºC in a total volume of 10 µL. 2 µl of the reaction was spotted on p81 Whatman paper while the remainder was analyzed by SDS-PAGE followed by Coomassie staining and fluorography. Identical reactions were performed in parallel using non-radiolabeled SAM (40 µM, Sigma) and analyzed by SDS- PAGE followed by western blotting with the α-H3K36me2 antibody (Upstate, catalog # 07-274).

Nuclear and Whole-cell Lysate Extractions

For nuclei extractions, wild-type and *set2Δ* S. *pombe* strains were grown in 1 L yeast extract supplemented with adenine (YEA) to a final OD_{600} between 2.0 and 2.5 prior to harvesting. Transformed *S. cerevisiae* strains were grown to a final OD_{600} between 2.0 and 2.5 in 200 ml SC-Ura prior to harvesting. Nuclei were extracted by Dounce homogenization from these cell pellets as previously described (Edmondson et al., 1996). Yeast whole-cell extracts (WCE) were prepared from 20 ml cultures grown to a final OD_{600} between 2.5 and

3.0 as described (Briggs et al., 2001) and only differed in the breaking buffer used for cell disruption (50 mM Tris, pH 8.0, 300 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 0.1% NP40, 0.5 mM EDTA, 10% glycerol, 2 mM PMSF, phosphatase inhibitor cocktail I (5 µl, Sigma), and 2 µg/ml pepstatin, aprotinin, and leupeptin).

Electrophoresis and Immunoblot Analyses

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and immunoblot analyses were performed using procedures and reagents from GE Healthcare. The anti-H3K36me (α-H3K36me, catalog # ab9048), H3K36me3 (α-K3K36me3, catalog # ab9050) and C-terminus of H3 (α-H3, catalog # ab1791) rabbit polyclonal antibodies were obtained from Abcam and used at dilutions of 1:1,000, 1:10,000, and 1:20,000, respectively. All other histone modification-detecting antibodies (rabbit) were obtained from Upstate Biotechnology Inc. and used at the following dilutions: 1:3,500-1:5,000 for H3K36me2 (α-H3K36me2, catalog # 07-274), 1:20,000 for H3K4me2 (α-H3K4me2, catalog # 07-030), 1:10,000 for H3K9ac (α-H3K9ac, catalog # 07-352), and 1:40,000 for H3K4me3 (α-H3K4me3, catalog # 07-413). Mouse monoclonal anti-FLAG antibody (M2; Sigma, catalog # F1804) was used at 1 µg/ml. Anti-polymerase CTD antibodies 8WG16 (Unmod CTD, catalog # MMS-126R) and H14 (Ser5 phosphorylation, catalog # MMS-134R) were from Covance Inc. and used at dilutions of 1:500 and 1:10,000-1:30,000, respectively. The IgM H14 antibody was detected using horseradish peroxidase (HRP)-conjugated donkey anti-mouse IgM at 1:5,000 (Jackson ImmunoResearch Laboratories). Typically, 20- 50 µg of WCE or 20-100 µg nuclei were resolved on SDS-PAGE gels (8% for RNAPII and FLAG blots or 13-15% for histone modification blots), followed by transfer to polyvinylidene difluoride (PVDF) membranes, and immunoblot analyses.

FLAG Immunoprecipitations

For FLAG immunoprecipitations, 2.0 mg of each WCE was incubated with 12.5 µl of pre-equilibrated anti-FLAG affinity beads (Anti-FLAG M2 agarose, Sigma) for 2 hours at 4°C. After three washes in extraction buffer, the bead-bound proteins were analyzed by immunoblot analysis using the antibodies and dilutions indicated above.

Growth Assays

The effects of media on cell growth were tested by growing wild-type and *set2Δ* strains (SPK131 and SPK612) in rich yeast extract supplemented with adenine (YEA) medium to a final OD_{600} of 0.8 and plating 1:10 serial dilutions of cells on either YEA or Edinburgh's minimal medium supplemented with amino acids (EMM). Plates were incubated at 30 ºC for 3-5 days.

Chromatin Immunoprecipitations Analyses

Chromatin immunoprecipitation assays were performed as previously described (Xiao et al., 2003). Briefly, WCEs were prepared from formaldehyde-fixed wild-type and *set2Δ S. pombe* strains grown in 100 ml YEA media to a final OD_{600} between 1.0 and 1.5. Extracts were sonicated to shear chromatin followed by immunoprecipitation (IP) using Protein A Sepharose (GE Healthcare) with anti-H3K36me2 (α-H3K36me2) at 3 µl/IP. Following washes and DNA elution, cross-links were reversed and DNA was extracted for amplification using standard PCR methods. Specific regions in the promoter and coding regions of the following genes were amplified: *ADE6*, *PMA1*, and *ACT1*. As a control, we used a primer pair to the K-region found in the mating-type loci. Primer sequences are available upon request. The results represent the ratio of immunoprecipitated (IP) DNA to

input DNA (Input) normalized to the IP/Input ratio from the mating-type loci associated region (K-region).

Results

Histone H3K36 Methylation is Highly Conserved

While H3K36 methylation has been demonstrated in budding yeast, its presence and relative abundance in other organisms has not been well established. To determine the conservation and relative abundance of H3K36 methylation in several diverse organisms, we isolated histones from budding yeast, *Tetrahymena*, chicken erythrocyte nuclei and human 293T cells and probed them for H3K36 methylation using an anti-dimethyllysine 36 antiserum (α-H3K36me2). For comparison and as a control, we used an antibody specific to dimethylation at histone H3 lysine 4 (α-H3K4me2). As shown in Fig. 2.1, we found that H3K36 dimethylation was present in all of the organisms analyzed, although their relative abundances varied between species. It is interesting that in *Tetrahymena*, H3K36 dimethylation appears to be less abundant compared to the levels of this modification found in yeast, chicken and humans (Fig. 2.1). However, a possible reason for this observation may be an inability of the H3K36 dimethyl antibody to efficiently recognize *Tetrahymena* H3. In yeast, chicken, and humans, H3K36 is immediately preceded by the amino acid valine while the predominant form of H3 in *Tetrahymena* (H3.1) contains an isoleucine that precedes K36 (GGVK$_{36}$KPH vs. GGIK$_{36}$KPH). Thus, this amino acid substitution may decrease this antibody's ability to effectively recognize H3K36 methylation in the context of its surrounding residues in *Tetrahymena*. Nonetheless, mass spectrometry analysis confirms that H3K36 is indeed mono-, di-, and trimethylated in this organism (C.D. Allis and D. Hunt, personal communication), although the relative amounts of these methyl forms in *Tetrahymena* H3 are not known. In addition to these results, previous studies have shown the existence of H3K36 methylation in humans, chicken and sea urchin by protein sequencing

(Holde, 1988). Furthermore, we and others have determined that H3K36 is also methylated in *Drosophila*, *Neurospora*, and *C. elegans* (B.D.S., unpublished results and Han et al., 2003). Thus, H3K36 methylation is found in a broad range of distinct eukaryotes.

With the finding that H3K36 methylation is highly conserved, we next asked whether Set2 homologs could be identified in these different organisms. Using the AWS (Associated with SET), SET, and post-SET domains (amino acids 63-260) of the budding yeast Set2 (ScSet2) protein as bait in a PSI-BLAST search, we found a significant number of proteins bearing similar sequence structures to Set2, and assembled them in a hierarchical family tree (Fig. 2.2). As documented in the figure, the ScSet2 protein was most similar to the *S. pombe* and *Neurospora crassa* (NCU00269.1) Set2 proteins (33% and 43% sequence identity, respectively) (also see Noma and Grewal, 2002). While not as highly conserved, the domain structure of ScSet2 is found in a number of other proteins found in a variety of diverse organisms. Strikingly, in addition to the AWS, SET, post-SET, and WW domain, more complex eukaryotes have a large number of additional domains and sequences such as PHD fingers and HMG domains, indicating that these putative Set2 homologs may carry out additional chromatin-related functions. It is notable that the mouse homolog of human NSD1 (Nsd1) has been shown to mediate H3K36 methylation *in vitro* (Rayasam et al., 2003), suggesting that the other proteins listed in Fig. 2.2 may be bonafide H3K36-methylating homologs. However, it was not known whether any of these proteins methylate H3K36 and associate with RNAPII.

S. Pombe Set2 is a Robust Methyltransferase Specific for H3K36

To determine if the link between H3K36 methylation and transcription elongation might be conserved, we characterized the Set2 protein thought to be responsible for H3K36

methylation in *S. pombe*. We chose to focus on *S. pombe* because many proteins in this organism have been found to be more similar to their mammalian counterparts than to their complements in *S. cerevisiae* (Sipiczki, 2000). In addition, the role of H3K36 methylation in this organism has not been investigated.

We first asked whether this protein is an active histone methyltransferase (HMT) and whether it catalyzes H3K36 methylation. To determine this, we cloned the *S. pombe* protein into a pCAL-n expression construct, expressed it in *E. coli,* and tested the recombinant protein in HMT assays using S-Adenosyl-L-[*methyl*-^{3}H] methionine (^{3}H-SAM) as a cofactor. As shown in Fig. 2.3A, SpSet2 showed a robust HMT activity towards nucleosomal substrates, and to a lesser extent, free core histones in filter binding assays. In contrast, this enzyme showed little activity towards free histone H3 (Fig. 2.3A). To determine the histone specificity of this methyltransferase, a portion of the HMT assays involving nucleosomal substrates were electrophoresed on a 15% SDS-PAGE gel and examined by fluorography. The results revealed that histone H3 was the only histone methylated (Fig. 2.3B). We next performed "cold" HMT assays with SpSet2 using unlabeled cofactor, followed by Western blot analysis with an antibody specific for H3K36 dimethylation to determine if SpSet2 was specific for H3K36. Results showed a significant immunoreactivity towards H3K36 di-methylation in the presence of SpSet2 (Fig. 2.3C). In contrast, no immunoreactivity was witnessed after these HMT assays with antibodies directed against either H3 lysine 79 dimethylation or H3K4 dimethylation (data not shown). To further verify the site specificity of SpSet2, we examined H3 synthetic peptides that were either unmodified or trimethylated at K36 in filter binding assays. Although the overall level of activity towards H3 peptides was low (compare ^{3}H incorporation levels between panels A and D), it was still sufficient to

determine whether this activity could be blocked with an H3K36 methylated peptide. As shown in Fig. 2.3D, SpSet2 was able to methylate an H3 peptide of residues 27-46, but not that of an H3 N-terminal peptide (residues 1-21). Importantly, a matched 27-46 peptide that was trimethylated at H3K36 was not a substrate (Fig. 2.3D). These data demonstrate that SpSet2 is a robust nucleosomal-selective HMT specific for H3K36 methylation.

Additionally, we characterized the methyltransferase activity of the *N. crassa* Set2 protein, NcSet2, *in vitro*. Like SpSet2, this protein has a similar sequence and domain organization to budding yeast Set2 suggesting that it may be responsible for methylating H3K36. However, *N. crassa* is a more developmentally complex multicellular eukaryote leading to the possibility that NcSet2 may have additional functions than that of the proteins identified in *S. pombe* and budding yeast. In collaboration with the Selker laboratory (U. of Oregon), we bacterially expressed the N-terminus of NcSet2 (amino acids 1-372) and tested the recombinant protein in HMT assays. Similar to SpSet2, NcSet2 is a nucleosomal histone methyltransferase that targets H3K36 for methylation (Appendix). In a companion report, the Selker laboratory goes on to further show that NcSet2-mediated methylation of H3K36 is required for proper development in *N. crassa* (Adhvaryu et al., 2005).

Set2 and H3K36 Methylation are Associated with Transcription Elongation in S. Pombe

We next asked whether SpSet2 is responsible for *in vivo* H3K36 methylation in *S. pombe* and whether it associates with elongating RNAPII. To address the first point, we deleted the *set2*$^{+}$ gene from *S. pombe* and used these cells, along with the wild-type (WT) control, to generate purified nuclei for subsequent western blot analyses. As shown in Fig. 2.4A, deletion of *set2*$^{+}$ resulted in a complete abolishment of H3K36 methylation (mono-, di-, and trimethylation), but not H3K4 methylation or H3K9 acetylation, in bulk histones, indicating

that SpSet2 is the sole enzyme in fission yeast responsible for this modification. We also examined the *set2*$^{+}$ deletion (*set2Δ)* strain for growth defects and found that while *set2Δ* cells grew normally on rich YEA medium, it showed a strong growth defect in synthetic medium (EMM), which is nutrient depleted compared to YEA (Fig. 2.4B). These data reveal an important role for Set2 in cell growth under deprived nutrient or stressed conditions. Similar slow-growth phenotypes on minimal medium have been described for the deletion of other factors involved in transcription and translation (Akiyoshi et al., 2001; Smith et al., 1999).

To determine if the SpSet2 enzyme would be associated with RNAPII, we tagged SpSet2 at its C-terminus with a triple FLAG epitope (SpSet2-3FLAG) and then used this epitope to perform co-immunoprecipitation experiments to monitor the association of unmodified or hyperphosphorylated RNAPII. The different forms of RNAPII were monitored using antibodies 8WG16 and H14, which recognize unmodified or Ser5 phosphorylated CTD respectively. As shown in Fig. 2.4C, immunoprecipitation of SpSet2-3FLAG resulted in strong immunoreactivity of the Ser5 phosphorylated CTD form of RNAPII. No unmodified RNAPII could be detected in these immunoprecipitates, although unmodified RNAPII could be readily detected in the "input" extracts. These data demonstrate that SpSet2 is associated with the elongating form of RNAPII in *S. pombe*. This result is also consistent with the finding that a region in the C-terminus of the SpSet2 protein contains similarity (17/42% identity/similarity) to a region in ScSet2 that was found to mediate association of Set2 with the phosphorylated polymerase (Kizer et al., 2005).

Next, we asked whether H3K36 methylation is associated with the transcribed region of active genes in *S. pombe*. To address this, we used an H3K36 dimethylation-specific antiserum ChIP assays to examine the abundance and distribution of this modification over

genes. Consistent with observations in budding yeast (Krogan et al., 2003b; Schaft et al., 2003; Xiao et al., 2003), we found that H3K36 methylation was highly enriched over the transcribed region of several active genes tested (Fig. 2.5). In contrast, non-transcribed regions of telomeric and mating type loci were found to be devoid of this methyl mark (data not shown and Fig. 2.5). These data strongly suggest that H3K36 methylation, mediated by SpSet2, is associated with the elongation process in *S. pombe*.

Given the strong similarities found between the budding and fission yeast Set2 proteins, we finally asked if the fission yeast H3K36-methylating enzyme could complement for the loss of Set2 in budding yeast cells. To examine this, we cloned the *S. pombe set2*$^{+}$ gene, containing a C-terminal FLAG tag, into a budding yeast expression construct (under the control of the *ADH1* promoter) and expressed this protein in *set2Δ* cells. As a control, a similar expression construct containing the budding yeast *SET2* gene was included. As shown in Fig. 2.6, full-length SpSet2 could be readily detected by western blot analysis using an anti-FLAG antibody. The *S. pombe* protein runs with a slower migration compared to the budding yeast Set2, as this protein is slightly larger than its budding yeast counterpart. We then purified nuclei from these strains and examined the levels of H3K36 methylation on bulk histones. Significantly, we found that the *S. pombe* Set2 protein could restore H3K36 methylation in *set2Δ* cells. This result suggests that SpSet2 forms a stable interaction with RNAPII in budding yeast. To examine this idea further, we performed similar co-IP studies as described above and found that SpSet2 efficiently associates with the elongating form of budding yeast RNAPII, similar to its budding yeast counterpart. This result shows that these enzymes are interchangeable, thereby supporting a notion that Set2 and H3K36 methylation is functionally conserved.

Discussion

Similar to the conservation found between histone protein sequences among eukaryotes, their covalent modifications are also highly conserved (Holde, 1988). Yet, a looming question has been whether these modifications perform the same functions in all of these different organisms, or do they have distinct functions that have arisen through evolutionary change? To date, several sites of histone methylation have been associated with active transcription. These include the methylation of H3 at lysines 4, 36, and 79 (Gerber and Shilatifard, 2003; Hampsey and Reinberg, 2003; Krogan et al., 2003b; Litt et al., 2001; Noma et al., 2001; Shilatifard, 2004; Strahl et al., 1999). In *S. cerevisiae*, the enzymes responsible for H3K4 and H3K36 methylation have been found associated with the elongating form of RNAPII (Hampsey and Reinberg, 2003; Shilatifard, 2004). This intriguing observation implies a novel role for histone methylation in the elongation phase of transcription. However, whether these enzymes are conserved and have similar functions in organisms outside of budding yeast has not been fully investigated.

In this report, we characterize the fission yeast enzyme responsible for H3K36 methylation and provide evidence that this modification in *S. pombe* is coupled with the transcription elongation process. Given that budding yeast is evolutionarily distinct from fission yeast, this result suggests that Set2 has a conserved function in transcriptional regulation. While our study has focused on H3K36 methylation, several studies have characterized Set1 homologs responsible for H3K4 methylation outside of budding yeast (Briggs et al., 2001; Byrd and Shearn, 2003; Milne et al., 2002; Nishioka et al., 2002; Wysocka et al., 2003). Although the link between H3K4 methylation and transcription elongation in these organisms is not well defined, there are a number of similarities found

between Set1/H3K4 methylation among eukaryotes that suggests a conserved role for this modification similar to what is found for Set2/H3K36 methylation. First, Set1-mediated H3K4 methylation is associated primarily with euchromatic regions in all organisms, with a distribution pattern on genes in metazoans and in *S. pombe* that is very similar to what is observed in budding yeast (Schneider et al., 2004). Secondly, comparative studies of human, budding and fission yeast Set1 have shown that these proteins have nearly identical complex compositions, and in the case of *S. pombe*, Set1's methylation is dependent on the ubiquitin-conjugating enzyme Rad6 (Roguev et al., 2003). Additionally, a Set1 homolog in humans, MLL2, has been shown to associate with the Ser5 phosphorylated form of RNAPII, similar to the budding yeast counterpart (Hughes et al., 2004). Collectively, these data suggest that the functional conservation of H3K4 and H3K36 methylation in RNAPII-mediated transcription is highly conserved.

It is significant to mention that while *S. pombe* has only one Set2-like homolog, our BLAST searches revealed that most multicellular organisms have a number of Set2-like enzymes (see Fig. 2.2 and data not shown). Although the function of these other enzymes are not known, it is intriguing to speculate whether all of these putative H3K36-methylating enzymes could be associated with the elongating polymerase in their respective organisms, or if some of these Set2-like proteins are involved in other biological processes outside of RNAPII-coupled transcription. While it will take in-depth characterization of each putative Set2 homolog to determine their role(s) in chromatin regulation, one clue to suggest that H3K36 methylation in more complex eukaryotes may have distinct functions outside of a role with RNAPII is the fact that only one Set2 homolog from any given species appears to contain a prototypical SRI domain, which is the domain required for Set2 to mediate its

association with the phosphorylated CTD (Kizer et al., 2005). Given that not all Set2 homologs contain this domain, we speculate that H3K36 methylation will have a broad range of activities in chromatin in addition to a conserved role with the transcribing polymerase.

In summary, we demonstrate that SpSet2 is a true ortholog of the budding yeast Set2 enzyme, and that this enzyme and H3K36 methylation is linked to the transcription elongation process in *S. pombe*. While our characterization studies are limited to *S. pombe*, an accompanying paper shows similar findings for the *Neurospora* Set2 homolog (Adhvaryu et al., 2005). Furthermore, a link between H3K36 methylation and CTD phosphorylation has been suggested in *C. elegans* (Han et al., 2003), and recent evidence shows a correlation of H3K36 methylation with active genes in metazoans (Bannister et al., 2005). Taken together, these results indicate a highly conserved role for H3K36 methylation in transcriptional regulation.

Figure 2.1

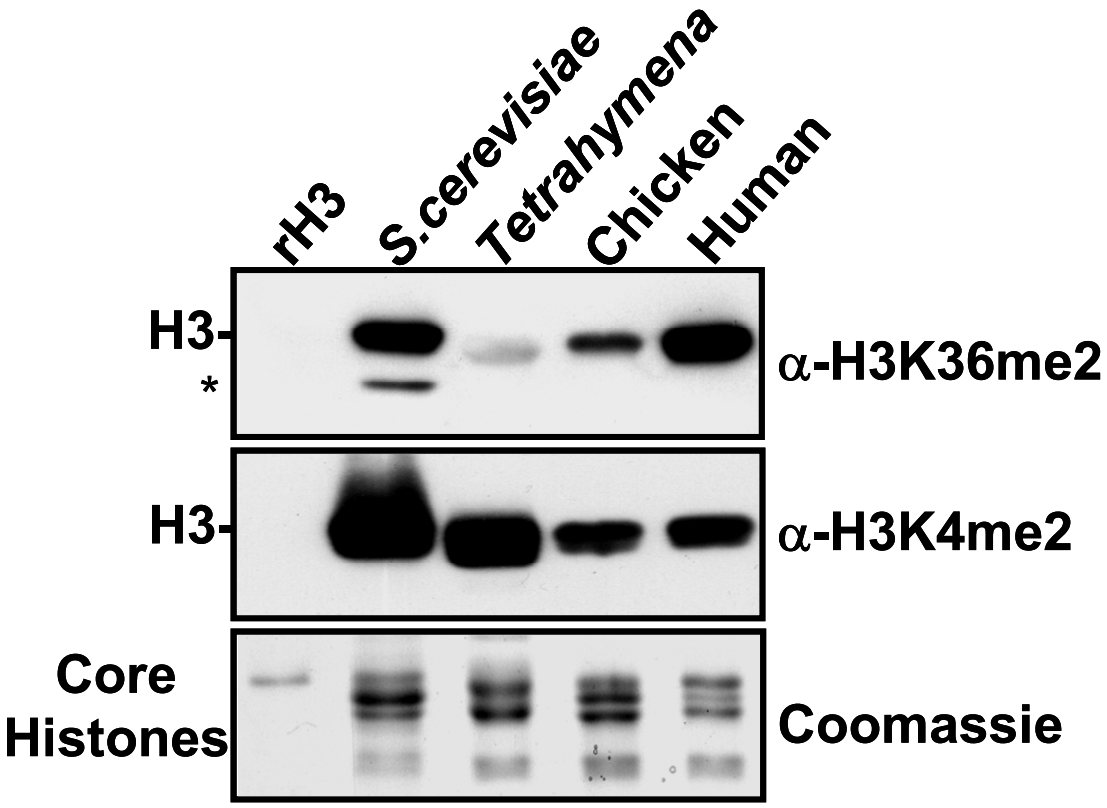

rH3
S.cerevisiae
Tetrahymena
Chicken
Human
H3
*
α-H3K36me2
H3
α-H3K4me2
Core Histones
Coomassie

Figure 2.1. Conservation and abundance of histone H3K36 methylation. 1 μg of *Xenopus* recombinant histone H3 and 5 μg of total core histones from the species indicated were resolved on 15% SDS-PAGE gels, transferred to PVDF membrane, and probed with α-H3K36me2 and α-H3K4me2. Identical samples were examined in parallel by Coomassie staining to show histone loading. Asterisk indicates H3 breakdown products that are typically observed.

Figure 2.2

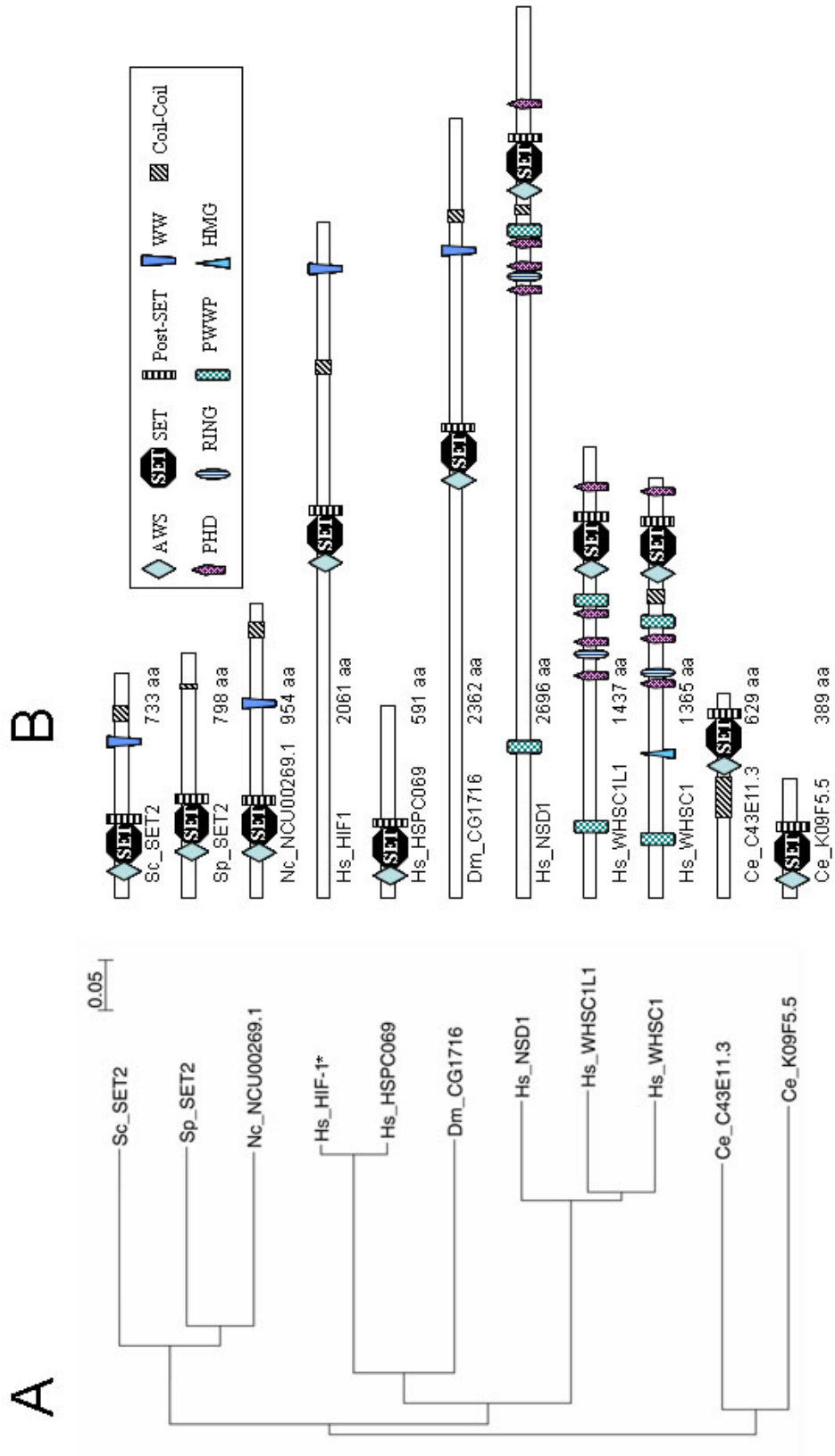

A
0.05
Sc_SET2
Sp_SET2
Nc_NCU00269.1
Hs_HIF-1*
Hs_HSPC069
Dm_CG1716
Hs_NSD1
Hs_WHSC1L1
Hs_WHSC1
Ce_C43E11.3
Ce_K09F5.5
B
AWS
SET
Post-SET
WW
Coil-Coil
PHD
RING
PWWP
HMG
Sc_SET2 733 aa
Sp_SET2 798 aa
Nc_NCU00269.1 954 aa
Hs_HIF1 2061 aa
Hs_HSPC069 591 aa
Dm_CG1716 2362 aa
Hs_NSD1 2696 aa
Hs_WHSC1L1 1437 aa
Hs_WHSC1 1365 aa
Ce_C43E11.3 629 aa
Ce_K09F5.5 389 aa

Figure 2.2. Conservation of Set2 proteins among eukaryotes. *(A)* Phylogenetic tree of Set2 and its putative homologs. Set2 proteins were identified in different species by using PSI-BLAST and then clustered into groups based on amino acid sequences using the phylogenetic analysis program Clustal X (ExPASy/SIB accession nos: Sc_SET2, P46995; Sp_SET2, O14026; Nc-NCU00269.1, Q7RZU4, Hs_HIF-1, Q9BYW2; Hs_HSPC069, Q9NZW9; Dm_CG1716, Q9VYD1; Hs_NSD1, Q96L73; Hs_WHSC1L1, Q9BZ95; Hs_WHSC1, O96028; Ce_C43E11.3, Q8I7H3; Ce_K09F5.5, Q21404). The scale bar equals a distance of 0.05 amino acids. Set2 protein abbreviations: *S. cerevisiae*, Sc; *S. pombe*, Sp; *N. crassa*, Nc; *Homo Sapiens*, Hs; *Drosophila*, Dm; and *C. elegans*, Ce. Asterisk indicates that this protein is also known as HYPB. *(B)* Schematic domain representation of Set2 proteins identified from the alignment in *(A)*. Protein names and lengths in amino acids are noted beneath each protein.

Figure 2.3

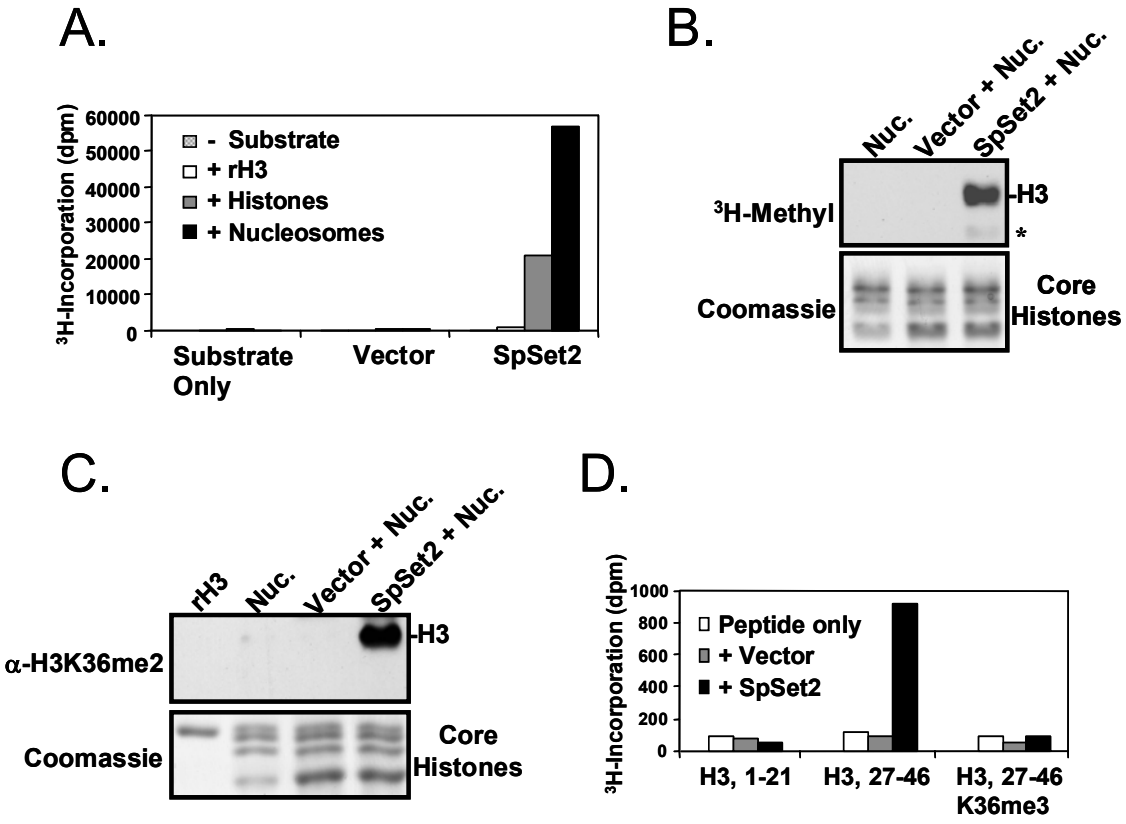

A.
3H-Incorporation (dpm)
60000
50000
40000
30000
20000
10000
0
- Substrate
+ rH3
+ Histones
+ Nucleosomes
Substrate Only
Vector
SpSet2
B.
Nuc.
Vector + Nuc.
SpSet2 + Nuc.
3H-Methyl
-H3
*
Coomassie
Core Histones
C.
rH3
Nuc.
Vector + Nuc.
SpSet2 + Nuc.
α-H3K36me2
-H3
Coomassie
Core Histones
D.
3H-Incorporation (dpm)
1000
800
600
400
200
0
Peptide only
+ Vector
+ SpSet2
H3, 1-21
H3, 27-46
H3, 27-46 K36me3

Figure 2.3. Set2 from *Schizosaccharomyces pombe* is a robust nucleosomal-selective H3K36-specific methyltransferase. *(A)* Bacterial lysates containing recombinantly expressed SpSet2 (or vector only control) were incubated with recombinant H3 (rH3), chicken core histones, or oligonucleosomes and ^{3}H-labeled S-adenosyl-methionine (^{3}H-SAM). ^{3}H incorporation was analyzed by the filter-binding assay and monitored by scintillation counting. *(B)* Reaction products from HMT assays using oligonucleosomes (Nuc.) and SpSet2 were resolved on an SDS-PAGE 15% gel and examined by Coomassie staining (lower panel) and fluorography (upper panel). Asterisk indicates the H3 breakdown product that is typically observed. *(C)* Oligonucleosomes or rH3 were incubated with SpSet2 and cold SAM in an HMT assay followed by immunoblotting with the α-H3K36me2 antibody (upper panel). Parallel reactions were performed and examined by Coomassie staining to monitor loading (lower panel). *(D)* Filter-binding assays were performed as in *(A)* using bacterial lysates with or without SpSet2 and H3 peptides either unmodified (H3, 1-21 or H3, 27-46) or trimethylated at H3K36 (H3, 27-46 K36me3) in the presence of ^{3}H-SAM.

Figure 2.4

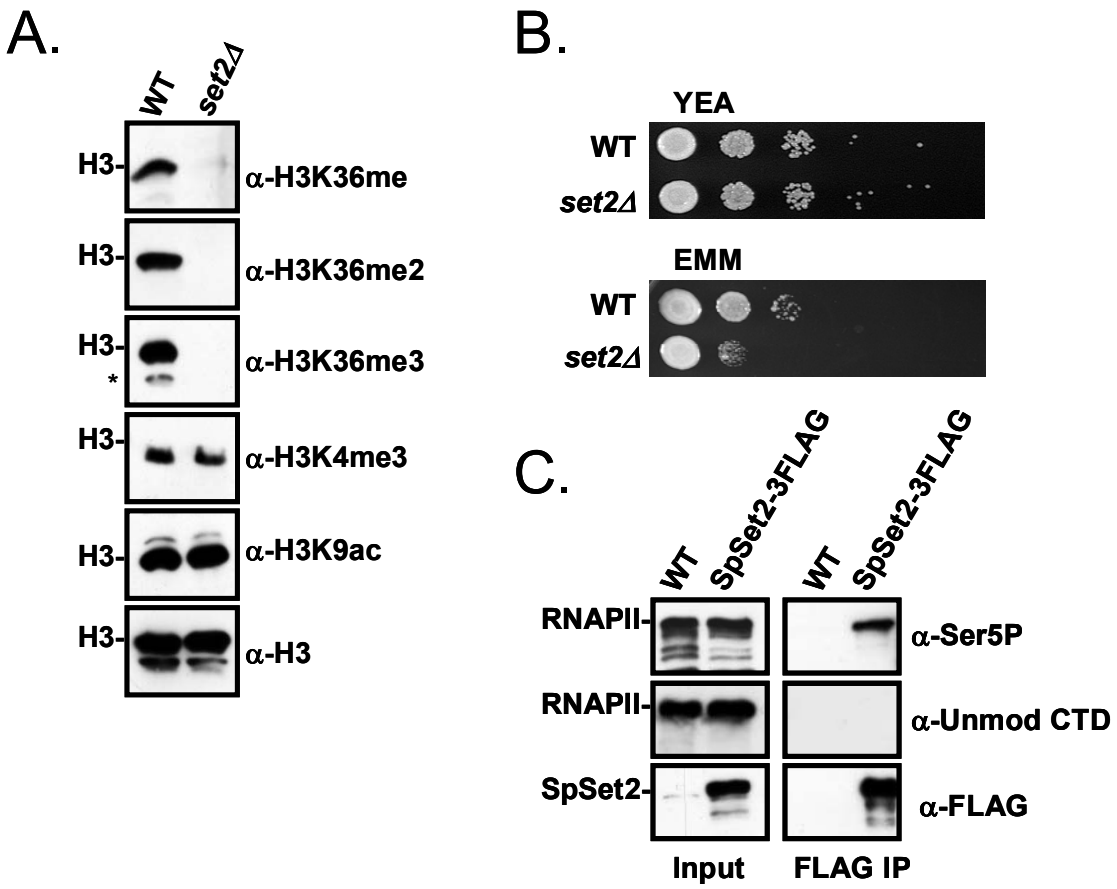

A.
WT
set2Δ
H3
α-H3K36me
H3
α-H3K36me2
H3
*
α-H3K36me3
H3
α-H3K4me3
H3
α-H3K9ac
H3
α-H3
B.
YEA
WT
set2Δ
EMM
WT
set2Δ
C.
WT
SpSet2-3FLAG
WT
SpSet2-3FLAG
RNAPII
α-Ser5P
RNAPII
α-Unmod CTD
SpSet2
α-FLAG
Input
FLAG IP

Figure 2.4. Set2 is responsible for mediating global H3K36 methylation in *S. pombe*. *(A) S. pombe* nuclear extracts prepared from wild-type and *set2Δ* strains were probed with antibodies against H3K36 mono-, di-, and trimethylation. An antibody specific for the C-terminus of H3 was used as a loading control. Antibodies specific for H3K9 acetylation and H3K4 trimethylation were used as additional controls. Asterisk indicates H3 breakdown products that were observed. *(B)* A slow-growth phenotype develops in the absence of *S. pombe set2*$^{+}$ under nutrient deprived conditions. Wild-type or *set2Δ* cells were spotted at serial dilutions of 1:10 and grown at 30ºC on rich medium (YEA) for 3 days or minimal medium (EMM) for 5 days before being photographed. *(C)* SpSet2 interacts with the hyperphosphorylated form of RNAPII. Whole-cell extracts (WCEs) prepared from wild-type or genomically tagged SpSet2 (SpSet2-3FLAG) were immunoprecipitated with anti-FLAG antibody followed by immunoblotting with antibodies directed against unmodified CTD (8WG16, α-unmod CTD), Ser5 phosphorylated CTD (H14, α- Ser5P), or FLAG (α-FLAG). The locations of RNAPII and SpSet2-3Flag are indicated. The input WCEs were also examined by immunoblot analysis to monitor the presence of RNAPII and SpSet2-FLAG.

Figure 2.5

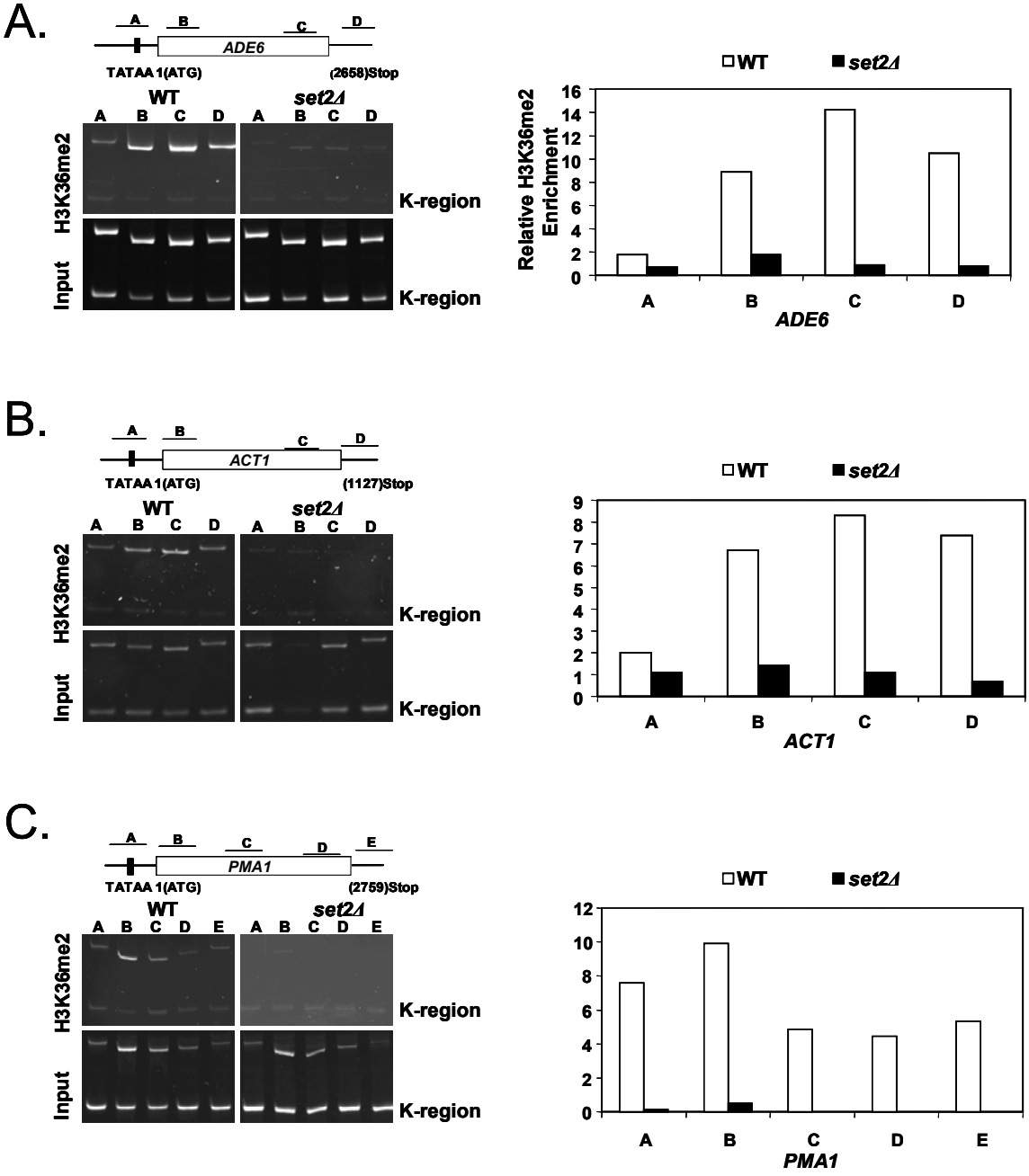

A.
A
B
C
D
ADE6
TATAA 1(ATG)
(2658)Stop
WT
set2Δ
A B C D A B C D
H3K36me2
K-region
Input
K-region
WT
set2Δ
Relative H3K36me2 Enrichment
16
14
12
10
8
6
4
2
0
A
B
C
D
ADE6
B.
A
B
C
D
ACT1
TATAA 1(ATG)
(1127)Stop
WT
set2Δ
A B C D A B C D
H3K36me2
K-region
Input
K-region
WT
set2Δ
9
8
7
6
5
4
3
2
1
0
A
B
C
D
ACT1
C.
A
B
C
D
E
PMA1
TATAA 1(ATG)
(2759)Stop
WT
set2Δ
A B C D E A B C D E
H3K36me2
K-region
Input
K-region
WT
set2Δ
12
10
8
6
4
2
0
A
B
C
D
E
PMA1

Figure 2.5. Set2-mediated H3K36 methylation is preferentially associated with the transcribed region of active *S. pombe* genes. Left Panels: *(A-C)* Chromatin immunoprecipitation assays were used to monitor the location of H3K36 dimethylation on actively transcribed genes (*ADE6*, *ACT1*, and *PMA1*) in wild-type and *set2Δ* strains using an H3K36me2-specific antibody. DNA from enriched precipitates (IP) were isolated and used in PCR reactions with promoter and coding region-specific primer pairs for the indicated genes. A DNA fragment from the silent mating-type loci (K-region) of *S. pombe* known to lack modifications associated with active genes (H3K4 methylation and H3K14 acetylation; see 34) was used as a control to normalize and calculate the relative enrichment of gene sequences in immunoprecipitated samples. Right Panels: Quantification of the ChIP results shown in *(A-C)*. Relative enrichment values shown on the y axes were calculated by dividing the ratio of band intensities for IP DNA/K-region with the ratio of intensities for the input DNA/K-region. Gels and graphs are representative experiments from three independent repeats.

Figure 2.6

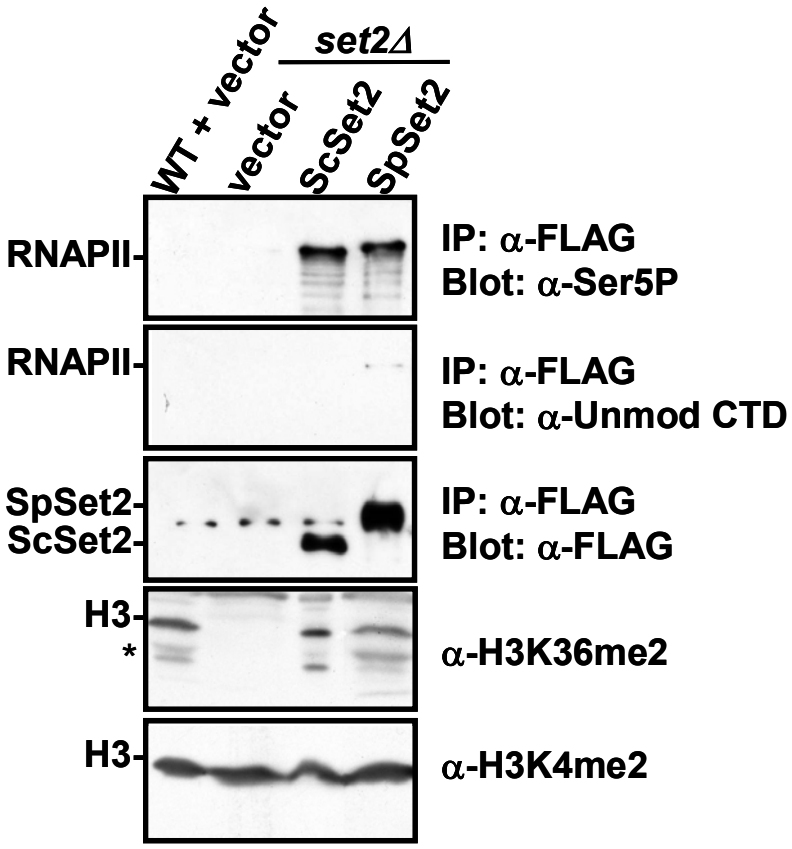

WT + vector
set2Δ
vector
ScSet2
SpSet2
RNAPII-
IP: α-FLAG
Blot: α-Ser5P
RNAPII-
IP: α-FLAG
Blot: α-Unmod CTD
SpSet2-
ScSet2-
IP: α-FLAG
Blot: α-FLAG
H3-
*
α-H3K36me2
H3-
α-H3K4me2

Figure 2.6. SpSet2 rescues H3K36 methylation in *S. cerevisiae.* Whole-cell extracts (WCEs) were prepared from *S. cerevisiae* wild-type and *set2Δ* strains that were transformed with an empty vector or the indicated Set2-FLAG constructs. These extracts were immunoprecipitated with α-FLAG antibody followed by immunoblot analysis using antibodies directed against unmodified CTD (8WG16, α-unmod CTD), Ser5 phosphorylated CTD (H14, α-Ser5P), or FLAG (α-FLAG). The locations of RNAPII and Set2-Flag proteins are indicated. All input extracts showed equivalent levels of RNAPII (data not shown). Nuclear extracts prepared from these same strains were immunoblotted with an antibody directed against H3K36 dimethylation or an antibody specific for H3K4 dimethylation (used as a loading control). Asterisk indicates H3 breakdown products that were observed.

Acknowledgments

This work has been published under the title "Histone H3 K36 methylation is associated with transcription elongation in *Schizosaccharomyces pombe*" in Eukaryotic Cell, vol. 4: pp 1446-1454, 2005. Permission to include this manuscript in my dissertation was granted by the American Society for Microbiology. The authors, in addition to myself (lead author), were Yoichiro Shibata, Ken-ichi Noma, Yuko Tsukamoto, Erin Warren, Brenda Temple, Shiv Grewal, and Brian Strahl. The data presented in Figure 2.1 was generated by Brian Strahl, data in panel A of Figure 2.2 was generated by Brenda Temple, data in Figure 2.5 was generated by Yoichiro Shibata, and the H3K36me2/H3K4me2 blots in Figure 2.6 were generated by Erin Warren. Ken-ichi Noma created the *set2Δ* and SpSet2-3FLAG strains in the *S. pombe* background.

References

Adhvaryu, K. K., Morris, S. A., Strahl, B. D., and Selker, E. U. (2005). Methylation of histone H3 lysine 36 is required for normal development in Neurospora crassa. Eukaryot Cell *4*, 1455-1464.

Akiyoshi, Y., Clayton, J., Phan, L., Yamamoto, M., Hinnebusch, A. G., Watanabe, Y., and Asano, K. (2001). Fission yeast homolog of murine Int-6 protein, encoded by mouse mammary tumor virus integration site, is associated with the conserved core subunits of eukaryotic translation initiation factor 3. J Biol Chem *276*, 10056-10062.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Bahler, J., Wu, J. Q., Longtine, M. S., Shah, N. G., McKenzie, A., 3rd, Steever, A. B., Wach, A., Philippsen, P., and Pringle, J. R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast *14*, 943-951.

Bannister, A. J., Schneider, R., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2005). Spatial distribution of di-and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem.

Berger, S. L. (2002). Histone modifications in transcriptional regulation. Curr Opin Genet Dev *12*, 142-148.

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*, 3286-3295.

Bryk, M., Briggs, S. D., Strahl, B. D., Curcio, M. J., Allis, C. D., and Winston, F. (2002). Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in S. cerevisiae by a Sir2-independent mechanism. Curr Biol *12*, 165-170.

Byrd, K. N., and Shearn, A. (2003). ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc Natl Acad Sci U S A *100*, 11535-11540.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev *14*, 155-164.

Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat Struct Mol Biol *11*, 1037-1043.

Edmondson, D. G., Smith, M. M., and Roth, S. Y. (1996). Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. Genes Dev *10*, 1247-1259.

Gerber, M., and Shilatifard, A. (2003). Transcriptional elongation by RNA polymerase II and histone methylation. J Biol Chem *278*, 26303-26306.

Grewal, S. I., and Rice, J. C. (2004). Regulation of heterochromatin by histone methylation and small RNAs. Curr Opin Cell Biol *16*, 230-238.

Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell *113*, 429-432.

Han, Z., Saam, J. R., Adams, H. P., Mango, S. E., and Schumacher, J. M. (2003). The C. elegans Tousled-like kinase (TLK-1) has an essential role in transcription. Curr Biol *13*, 1921-1929.

Holde, K. E. v. (1988). Chromatin (New York: Springer-Verlag).

Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacharjee, A., Biondi, C. A., *et al.* (2004). Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell *13*, 587-597.

Iizuka, M., and Smith, M. M. (2003). Functional consequences of histone modifications. Curr Opin Genet Dev *13*, 154-160.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., and Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. Trends Biochem Sci *23*, 403-405.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-1080.

Kizer, K. O., Phatnani, H. P., Shibata, Y., Hall, H., Greenleaf, A. L., and Strahl, B. D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol *25*, 3305-3316.

Kouzarides, T. (2002). Histone methylation in transcriptional control. Curr Opin Genet Dev *12*, 198-209.

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

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C.*, et al.* (2003b). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Lee, D. Y., Teyssier, C., Strahl, B. D., and Stallcup, M. R. (2004). Role of Protein Methylation in Regulation of Transcription. Endocr Rev.

Li, B., Howe, L., Anderson, S., Yates, J. R., 3rd, and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *278*, 8897-8903.

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. J Biol Chem *277*, 49383-49388.

Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science *293*, 2453-2455.

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., and Hess, J. L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell *10*, 1107-1117.

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*, 709-719.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C. D., Tempst, P., and Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes Dev *16*, 479-489.

Noma, K., Allis, C. D., and Grewal, S. I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science *293*, 1150-1155.

Noma, K., and Grewal, S. I. (2002). Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. Proc Natl Acad Sci U S A *99 Suppl 4*, 16438-16445.

Perriere, G., and Gouy, M. (1996). WWW-query: an on-line retrieval system for biological sequence banks. Biochimie *78*, 364-369.

Peterson, C. L., and Laniel, M. A. (2004). Histones and histone modifications. Curr Biol *14*, R546-551.

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. Embo J *22*, 3153-3163.

Roguev, A., Schaft, D., Shevchenko, A., Aasland, R., Shevchenko, A., and Stewart, A. F. (2003). High conservation of the Set1/Rad6 axis of histone 3 lysine 4 methylation in budding and fission yeasts. J Biol Chem *278*, 8487-8493.

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*, 2475-2482.

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*, 73-77.

Shiio, Y., and Eisenman, R. N. (2003). Histone sumoylation is associated with transcriptional repression. Proc Natl Acad Sci U S A *100*, 13225-13230.

Shilatifard, A. (2004). Transcriptional elongation control by RNA polymerase II: a new frontier. Biochim Biophys Acta *1677*, 79-86.

Sims, R. J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev *18*, 2437-2468.

Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? Genome Biol *1*, REVIEWS1011.

Smith, K. N., Iwanejko, L., Loeillet, S., Fabre, F., and Nicolas, A. (1999). Disruption and functional analysis of seven ORFs on chromosome IV: YDL057w, YDL012c, YDL010w, YDL009c, YDL008w (APC11), YDL005c (MED2) and YDL003w (MCD1). Yeast *15*, 1255-1267.

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

Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R., and Allis, C. D. (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. Curr Biol *11*, 996-1000.

Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proc Natl Acad Sci U S A *96*, 14967-14972.

Turner, B. M. (2000). Histone acetylation and an epigenetic code. Bioessays *22*, 836-845.

Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N., and Herr, W. (2003). Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev *17*, 896-911.

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*, 654-663.

Zhang, L., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. Chromosoma *112*, 77-86.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev *15*, 2343-2360.

CHAPTER 3

ROLE OF SET2-MEDIATED HISTONE H3 LYSINE 36 METHYLATION IN TRANSCRIPTION ELONGATION

Abstract

Post-translational histone modifications, such as lysine acetylation and methylation, play major roles in the regulation of chromatin during eukaryotic gene transcription. In the budding yeast *Saccharomyces cerevisiae*, the histone methyltransferase Set2 is the sole enzyme responsible for methylating lysine 36 of histone H3 (H3K36). Through an interaction with the elongating RNA polymerase II, the activity of Set2 is localized to the coding region of genes suggesting that Set2 and H3K36 methylation are involved in transcription elongation. However, the role of Set2 in transcription remains unclear. Here, we report a function for Set2 and H3K36 methylation in the regulation of histone acetylation in the coding region of genes. Using a combination of genetic interaction studies, Set2 was found to be genetically similar to the histone deacetylase Rpd3 and the chromodomain protein Eaf3, members of the small histone deacetylase complex Rpd3S. These data suggest that these proteins function in the same biological pathway. Consistent with this idea, we find that deletion of *SET2*, *RPD3*, and *EAF3* result in similar growth defects in the presence of the nucleotide-depleting drug 6-azauracil. Using chromatin immunoprecipitation (ChIP), we further demonstrate that similar to a deletion of *EAF3*, a mutation of *SET2* or H3K36 results in increased acetylation in the coding region of genes and functions upstream of the Rpd3S complex. In agreement with these findings, TAP-purified Rpd3S complex fails to

interact with nucleosomes in the absence of *SET2* indicating that Set2 and H3K36 methylation are required for the recruitment of this complex to chromatin. Collectively, these findings identify a function for Set2-mediated H3K36 methylation in the maintenance of deacetylated histones in the coding region of genes.

Introduction

The modification of histone proteins is an important mechanism of regulating chromatin accessibility during transcription. There are a number of different post-translational modifications known to occur on histones which include phosphorylation, ubiquitylation, methylation and acetylation (Berger, 2002; Holde, 1988). In particular, the dynamic regulation of lysine acetylation has been extensively linked to transcriptional activity (Grant, 2001). Histone acetyltransferases (HATs) can be recruited to the promoters of genes through an interaction with DNA-bound activators in which the acetylation of histones can alter chromatin structure or recruit additional factors leading to transcription activation, while the removal of acetylation by targeted histone deacetylases (HDACs) has been linked to repression (Brown et al., 2001; Cosma, 2002; Kurdistani et al., 2002; Robert et al., 2004). Additionally, HATs and HDACs can act in an untargeted, global manner to regulate acetylation levels genome-wide (Vogelauer et al., 2000). Together, these mechanisms of recruitment create a general pattern of higher acetylation in the promoter regions of genes in comparison to coding regions (Liu et al., 2005; Pokholok et al., 2005).

Unlike acetylation, histone methylation is not as well studied in the process of transcription. Linked to transcription activation and repression, methylation occurs in the promoter and coding regions of genes, with recent studies revealing a role for histone methylation in transcription elongation (Cao and Zhang, 2004; Grewal and Rice, 2004; Iizuka and Smith, 2003). In the budding yeast *Saccharomyces cerevisiae*, the histone methyltransferases Set1 and Set2, which catalyze H3 lysine 4 (H3K4) and lysine 36 (H3K36), respectively, have been found to associate with the elongating form of RNA polymerase II (RNAPII) (Gerber and Shilatifard, 2003; Hampsey and Reinberg, 2003). Set1

association and activity is dependent on the serine 5 (Ser5) phosphorylated form of the RNAPII C-terminal domain (CTD) at the 5' end of genes (Krogan et al., 2003a; Ng et al., 2003), while Set2 associates with the Ser2 phosphorylated CTD and methylates histones in the coding region of genes (Krogan et al., 2003b; Li et al., 2003; Li et al., 2002; Schaft et al., 2003; Xiao et al., 2003). Given the location of these modifications, it appears that they both function at different stages of transcription elongation. However, the exact functions of these enzymes and modifications remain unclear.

To gain a better understanding of how Set2 functions in transcription, our collaborators used a combination of gene expression profiles and synthetic genetic array analyses to identify protein-coding genes that may function in the same biological process as Set2 (Keogh et al., 2005; Krogan et al., 2003b; Tong et al., 2001). A high degree of similarity was found between *SET2*, *RPD3*, and *EAF3*. Rpd3, a yeast histone deacetylase, targets all 4 histones for deacetylation and functions as a part of a large multiprotein complex (Kurdistani et al., 2002; Suka et al., 2001; Zhang et al., 1998). As part of this complex, Rpd3 acts as a transcriptional repressor when targeted to promoters by the DNA-binding repressor Ume6 (Rundlett et al., 1998). Rpd3 also exists in a smaller, uncharacterized deacetylase complex, Rpd3S, that contains the unique subunits Rco1 and Eaf3 (Carrozza et al., 2005; Keogh et al., 2005). Eaf3, a component of both the NuA4 acetyltransferase and Rpd3 deacetylase complexes plays a role in regulating the histone acetylation patterns in the promoter and coding regions of genes (Eisen et al., 2001; Reid et al., 2004). In yeast, Eaf3 is not essential for growth and appears to affect the transcription of a subset of genes (Eisen et al., 2001). Interestingly, Eaf3 contains a structural motif that can bind methylated lysines in histones, the chromodomain (Bottomley, 2004; Brehm et al., 2004).

In this report, we demonstrate that Set2 functions in the same pathway as Rpd3 and Eaf3, members of the Rpd3S deacetylase complex. Like the chromodomain protein Eaf3, Set2-mediated H3K36 methylation regulates histone acetylation patterns in the coding region of genes. In the absence of *SET2*, Rpd3S does not interact with chromatin suggesting that H3K36 methylation by Set2 is required for the recruitment of Rpd3S through the chromodomain of Eaf3. This recruitment, in turn, is required to maintain histones in a deacetylated state in the wake of transcribing RNAPII during transcription elongation. Collectively, these results identify a functional mechanism for Set2 in the regulation of gene expression.

Materials and Methods

Yeast Strains and Plasmids

All gene deletion (*set2Δ*, *eaf3Δ*, *rpd3Δ*, *rco1Δ*, and *dst1Δ*) and wild-type (WT) strains are in the BY4741 background and were obtained from Open Biosystems. The wild-type histone H3 and H3K36-to-alanine point mutation (K36A) strains are in the WZY42 background and have been described previously (Kizer et al., 2005). The H2B-FLAG strain is in the BY4742 background (Xiao et al., 2005). For gene disruption of *SET2* in the H2B-FLAG strain, *SET2* was deleted by high efficiency transformation using a PCR product amplified from genomic DNA of *SET2* which had already been replaced by the *KanMX* gene (Research Genetics). The pRS313-Eaf3-3HA-SSN6 (Eaf3-3HA) and pRS313-3HA-SSN6 plasmids were kind gifts of Michael Keogh (Albert Einstein College of Medicine). The TAP-purified Rpd3S complex (Rco1-CBP) was kindly provided by Nevan Krogan (UCSF).

6-azauracil Growth Assays

Yeast strains (WT, *set2Δ*, *eaf3Δ*, *rpd3Δ*, and *dst1Δ*) were transformed with the *URA3*$^+$ CEN plasmid pRS316 and grown in synthetic defined medium lacking uracil (SD-URA) to a final OD_{600} between 1.0 and 2.0. 10-fold serial dilutions were plated on SD-URA medium with or without 100 µg/ml 6-azauracil (6AU (Aldrich)) and incubated at 30ºC for 2-3 days.

Cloning of RCO1 and Generation of Expression Construct

Using primers specific to the ORF of *RCO1* and 305 bp upstream of this gene, full-length *RCO1,* including regulatory regions, was PCR amplified from genomic DNA. The resulting product was cloned into the pRS313-3HA-SSN6 yeast expression plasmid which adds three copies of the HA epitope to the C-terminus of *RCO1*. The *RCO1* regulatory and coding regions were sequenced for accuracy. The resulting pRS313-Rco1-3HA-SSN6 (Rco1-3HA)

plasmid, which is driven by the native promoter of *RCO1*, was transformed into the BY4741 *S. cerevisiae rco1Δ* strain. As a control, the pRS313-3HA-SSN6 plasmid without the ORF of *RCO1* (empty vector) was transformed into wild-type and *rco1Δ* BY4741 strains. Transformants were selected on synthetic complete (SC)-His plates.

Whole-cell Lysate Extractions

S. cerevisiae strains transformed with HA expressing plasmids were grown to a final OD_{600} between 1 and 2 in 50-100 ml SC-His prior to harvesting while deletion strains (*set2Δ*, *eaf3Δ*, *rpd3Δ*, and corresponding WT) were grown in 10 ml yeast extract-peptone-dextrose (YPD) to a final OD_{600} of 1. Yeast whole-cell extracts (WCE) were prepared as described (Briggs et al., 2001) and only differed in the breaking buffer used for cell disruption (50 mM Tris, pH 8.0, 300 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 0.1% NP40, 0.5 mM EDTA, 10% glycerol, 2 mM PMSF, phosphatase inhibitor cocktail I (5 µl, Sigma), and 2 µg/ml pepstatin, aprotinin, and leupeptin).

Electrophoresis and Immunoblot Analyses

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblot analyses were performed using procedures and reagents from GE Healthcare. The anti-H3K36me3 (α-K3K36me3, catalog # ab9050), Eaf3 (α-Eaf3, catalog # ab4467), and Rpd3 (α-Rpd3, catalog # ab18085) rabbit polyclonal antibodies were obtained from Abcam and used at a dilution of 1:2,000. The following antibodies (rabbit) were obtained from Upstate Biotechnology, Inc. and used at the indicated dilutions: 1:500 for calmodulin binding protein epitope tag (α-CBP, catalog # 07482), 1:15,000 for the C-terminus of H3 (α-H3, catalog # 07-690), 1:3,500 for H3K14ac (α-H3K14ac, catalog # 07-353), and 1:2,000 for HA (α-HA, catalog # 05-902). The Set2 antiserum (α-Set2, 1:10,000 dilution) was

developed by immunizing a rabbit with a recombinant, bacterially expressed fragment of the N-terminus of Set2 (amino acids 1-261). Anti-polymerase CTD antibody H14 (Ser5 phosphorylation, catalog # MMS-134R) was from Covance Inc. and used at a dilution of 1:30,000. The IgM H14 antibody was detected using horseradish peroxidase (HRP)-conjugated donkey anti-mouse IgM at 1:5,000 (Jackson ImmunoResearch Laboratories). The enhanced chemiluminescence (ECL) streptavidin-HRP conjugate was used to detect biotinylated peptides and was used at a dilution of 1:80,000 (GE Healthcare, code RPN1231). Typically, 20- 50 µg of WCE were resolved on SDS-PAGE gels (8%-10% for RNAPII, Set2, Rco1, Rpd3, and Eaf3 or 13-15% for histone modification blots), followed by transfer to polyvinylidene difluoride (PVDF) membranes, and immunoblot analyses.

Chromatin Immunoprecipitation Analyses

Chromatin immunoprecipitation assays were performed as previously described (Xiao et al., 2003). Briefly, WCEs were prepared from formaldehyde-fixed WT, *set2Δ*, *eaf3Δ,* and K36A strains grown in 100 ml YPD media to a final OD_{600} between 1.0 and 1.5. WCEs were prepared, as well, from transformed WT, *rco1Δ,* and *rco1Δ* +Rco1-3HA strains grown under similar conditions in SC-His media. Extracts were sonicated to shear chromatin followed by immunoprecipitation (IP) using Protein A Sepharose (GE Healthcare) and the following antibodies as indicated: α-H3K14ac (3 µl/IP, Upstate Biotechnology, catalog # 07-353), α-H4ac (2 µl/IP, Upstate Biotechnology, catalog # 06-598), α-H3K9/14ac (3 µl/IP, Upstate Biotechnology, catalog # 06-599), and α-H3K9ac (3 µl/IP, Upstate Biotechnology, catalog # 07-352). Following washes and DNA elution, cross-links were reversed and DNA was extracted for amplification using standard PCR methods. Specific regions in the promoter and coding regions of the *SCC2* gene were amplified. As an internal control, we used a

primer pair specific to a subtelomeric region approximately 500 bp from the end of chromosome VI-R (TEL) which is not transcribed and hypomodified at sites associated with transcriptional activation (Rundlett et al., 1998; Suka et al., 2001). Primer sequences are available upon request. The results represent the ratio of immunoprecipitated (IP) DNA to input DNA (input) normalized to the IP/input ratio from the TEL subtelomeric region.

In Vitro Binding Studies

For each peptide pull-down reaction, 8 µl of pre-equilibrated streptavidin sepharose (GE Healthcare, high performance) was incubated for 30 min at room temperature with 3 nM of the following biotinylated peptides,: unmodified histone H3 (amino acids 27-46), K36me histone H3 (amino acids 27-46), K36me2 histone H3 (amino acids 27-46), and K36me3 histone H3 (amino acids 27-46). Peptides were either obtained from Upstate Biotechnology or synthesized and verified by mass spectrometry at the University of North Carolina Microprotein Sequencing and Peptide Synthesis Facility. After three washes in IP buffer (50 mM Tris, pH 8.0, 400 mM NaCl, 5 mM EDTA, 2 mM PMSF, phosphatase inhibitor cocktail I (Sigma), and 2 µg/ml pepstatin, aprotinin, and leupeptin), bead-bound peptides were resuspended in 225 µl IP buffer to which 75 µl of TAP-purified Rpd3S complex was added. Reactions were incubated for 2.5 hr at 4 °C and then washed three times in IP buffer. Bead-bound proteins were analyzed by SDS-PAGE (10% gels) followed by immunoblotting with the indicated antibodies. Similar peptide pull-downs were carried out using Eaf3-3HA WCEs with slight modifications. After three washes in 150 mM NaCl IP buffer, bead-bound peptides were resuspended in a total volume of 500 µl with 2 mg WCE. Reactions were incubated overnight at 4 °C and then washed three times in IP buffer followed by SDS-PAGE analysis.

For nucleosome pull-downs, yeast WCEs were prepared from 100 ml cultures of wild-type H2B-FLAG and *set2Δ* H2B-FLAG strains grown to a final OD_{600} between 0.8 and 1.0 in YPD as described above. For chromatin fragments, WCEs were pulse sonicated 6 times for 6 pulses each (30% output and 90% duty cycle) and clarified by centrifugation at 13,000 rpm for 15 min. For each pull-down reaction, 2 mg of WCE was mixed with 12.5 µL of pre-equilibrated anti-Flag beads (M2 agarose, Sigma) for 2 hours at 4 °C. After three washes in IP buffer (50 mM Tris, pH 8.0, 400 mM NaCl, 5 mM EDTA, 2 mM PMSF, phosphatase inhibitor cocktail I (Sigma), and 2 µg/ml pepstatin, aprotinin, and leupeptin), bead-bound chromatin was resuspended in 300 µl IP buffer to which 200 µL of TAP-purified Rco1 complex was added. Reactions were incubated overnight at 4 °C and then washed three times in IP buffer. Bead-bound proteins were analyzed by SDS-PAGE (12% gels) followed by immunoblotting with the indicated antibodies.

HA Immunoprecipitations

For HA immunoprecipitations, 2.0 mg of each WCE was incubated with 12.5 µl of pre-equilibrated monoclonal anti-HA agarose (Sigma) for 2 hours at 4°C. After three washes in IP buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 10% glycerol, 0.1% NP-40, 2 mM PMSF, phosphatase inhibitor cocktail I (Sigma), and 2 µg/ml pepstatin, aprotinin, and leupeptin), the bead-bound proteins were analyzed by immunoblot using the antibodies and dilutions indicated above.

Results

Members of the Rpd3S Complex and Set2 have Similar Growth Phenotypes

Given that previous studies have implicated Set2 in having a role in transcription elongation, and that Rpd3 and Eaf3 are genetically similar to Set2, we wanted to determine if Rpd3 and Eaf3 of the Rpd3S complex would also be involved in this process. To test for the involvement of these proteins in transcription elongation, we assayed deletion strains for growth defects in the presence of the nucleotide-depleting drug 6-azauracil (6AU). Growth phenotypes observed in the presence of this drug are frequently used as indicators of defects in transcription elongation (Exinger and Lacroute, 1992). Wild-type and matched *rpd3Δ*, *eaf3Δ*, and *set2Δ* strains transformed with the *URA3* expressing yeast plasmid pRS316 (a requirement for the 6AU assay) were grown on medium with or without 6AU (Smith et al., 2000). The colony growth of each strain was monitored over several days and compared to those of the control plate lacking 6AU. As previously reported, deletion of *SET2* confers resistance to the drug 6AU (Fig. 3.1A) (Kizer et al., 2005; Xiao et al., 2005). Similarly, *eaf3Δ* and *rpd3Δ* were both resistant to this drug in comparison to wild-type cells, while a control lacking the yeast homolog of TFIIS, *dst1Δ*, displayed sensitivity to 6AU as previously observed (Archambault et al., 1992). Interestingly, the deletion of *RPD3* was not as resistant to 6AU as either the *eaf3Δ* or *set2Δ*. This may be due in part to the involvement of Rpd3 in the large Rpd3 complex which has been linked to the regulation of transcription activation and not elongation. Nonetheless, these results indicate that Rpd3 and Eaf3 are involved in the transcription elongation process. Furthermore, the resistance phenotype observed in the absence of Rpd3 and Eaf3 suggest that these proteins may play a similar role in this process to that of Set2.

Set2 Regulates Histone Acetylation Patterns on Genes

With the finding that *SET2*, *RPD3*, and *EAF3* deletions exhibit similar growth phenotypes and genetically interact, we next wanted to investigate how these gene products may function in the same biological pathway. Previous studies of Eaf3 have shown that this protein is responsible for regulating global histone acetylation patterns on histones H3 and H4 (Reid et al., 2004). Generally, acetylation of H3 and H4 are higher in the promoters of genes in comparison to the coding regions. In the absence of Eaf3, this pattern is reversed such that H3 and H4 acetylation are dramatically increased in the coding region of genes and decreased at promoters without affecting global levels of acetylation (Reid et al., 2004). As a member of the Rpd3S complex, the loss of Eaf3 may interfere with the targeting of this deacetylase complex to the coding region of genes. This may be due, in part, to the chromodomain of Eaf3 which has been described as a motif that can specifically interact with methylated histones (Bannister et al., 2001; Lachner et al., 2001). Furthermore, Set2 methylates H3K36 in the coding region of genes, suggesting that Set2, through recruitment of the Rpd3S complex, may regulate acetylation levels in this region of genes as well.

To initially address this possibility, whole-cell extracts prepared from wild-type and matched *set2Δ*, *eaf3Δ*, and *rpd3Δ* strains were analyzed by immunoblot to examine the effects of these deletions on global levels of H3 acetylation and methylation of H3K36. As shown in Fig. 3.1B, similar to *eaf3Δ*, *set2Δ* does not affect global levels of H3K14 acetylation, while *rpd3Δ* leads to a subtle increase in this modification. This is not surprising as Rpd3 is a major global deacetylase that targets H3K14 among other sites on H3 and H4 (Vogelauer et al., 2000). As expected, H3K36me3 was lost in the *set2Δ*. Interestingly, methylation of H3K36 occurs independently of Eaf3 or Rpd3 (Fig. 3.1B) suggesting that, if

they do function in the same pathway, the Rpd3S complex functions downstream of H3K36 methylation.

To look at the effects of *set2Δ* on the histone acetylation patterns of individual genes, we used chromatin immunoprecipitation (ChIP) to examine the distribution and levels of H3K14 acetylation on *SCC2*, a constitutively active gene modified by Set2 (Kizer et al., 2005). In the case of wild-type cells, H3K14ac was enriched in just the promoter and 5' region of *SCC2* in comparison to a subtelomeric control region (Figure 3.2). Upon deletion of *SET2*, H3K14 acetylation dramatically increased in the coding region. Significantly, this increase mirrored that of *eaf3Δ* supporting the hypothesis that Set2-mediated H3K36 methylation recruits the Rpd3S complex to genes. We further analyzed *SCC2* for acetylation on H4 and H3K9 using antibodies specific to tetra-acetylated H4 (H4K5, K8, K12, and K16) and H3 diacetylated at H3K9 and 14. As reported for the deletion of *EAF3*, the deletion of *SET2* lead to a major increase in H4 acetylation and a less pronounced, but significant, increase in H3 acetylation (Reid et al., 2004) (Fig. 3.3A-B). Others have shown that acetylation at H3K9 is the predominant epitope recognized by the H3K9/14 antibody (Zhang et al., 1998). When we used an antibody specific for H3K9 (Fig. 3.3C), the levels of H3K9 acetylation seemed to be unaffected by the deletion of *SET2* or *EAF3* suggesting that Rpd3 regulates acetylation of specific H3 sites in the coding region of genes.

Since the loss of Set2 resulted in increased acetylation in the coding region of *SCC2*, we wanted to determine if this increase was due specifically to methylation at H3K36. To look at the possibility that H3K36 methylation regulates acetylation in the coding region of genes, we performed H4 acetylation and H3K9/14 acetylation ChIPs in extracts from an H3K36 point mutant (K36A) strain. As shown in Fig. 3.3D, mutation of H3K36 resulted in

increased acetylation of H3 and H4 comparable to the *set2Δ* strain. Similar effects were observed on *STE11*, *FLO8*, and *PMA1* indicating that this regulation is not specific to one gene (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). Collectively, these results indicate that Set2-mediated methylation of H3K36 is important for the regulation of acetylation levels in the coding region of genes.

Set2-mediated H3K36 Methylation Recruits the Rpd3S Deacetylase Complex to Nucleosomes

To determine if the increased acetylation in the absence of *SET2* and *EAF3* is due to the recruitment of the Rpd3S complex by methylated H3K36, we initially analyzed the interaction of the Rpd3S complex with H3K36me peptides. TAP-purified Rpd3S complex was incubated with different H3 peptides in which H3K36 was either unmodified or methylated. We further used peptides that were either mono-, di-, or trimethylated at H3K36 to identify which form specifically interacted with the Rpd3S complex. As shown in Fig. 3.4A, the Rpd3S complex interacted with all of the H3 peptides without specificity even under the most stringent salt conditions. This result may have been due to a missing accessory factor required for the specificity of this interaction lost during purification of the Rpd3S complex. To investigate this possibility, whole cell extracts expressing an HA-tagged form of Eaf3 were incubated with bead-bound-H3K36 peptides followed by immunoblot analysis. Similar to the binding studies using the purified Rpd3S complex, we could not detect a specific interaction (Fig. 3.4B). This led us to believe that the binding of Eaf3 to H3K36me requires additional histone H3 sequence and/or perhaps other histones in the context of a nucleosome. To test this hypothesis, we isolated nucleosomes from H2B-FLAG wild-type and *set2Δ* strains and incubated them with the purified Rpd3S complex. As shown in Figure 3.4C, the Rpd3S complex interacted strongly with nucleosomes from wild-

type and not *set2Δ* cells indicating that the recruitment of the Rpd3S complex to chromatin is dependent on Set2 activity.

Although the Rpd3S complex requires H3K36me to interact with chromatin, we also considered that it may not be directly binding this modification, but may be recruited to the coding region of genes through an interaction with the elongating polymerase or a direct interaction with Set2. To explore this possibility, we immunoprecipitated an HA-tagged form of Rco1, a unique member of the Rpd3S complex, and analyzed associated proteins by immunoblot. As shown in Fig. 3.5A, we did not detect an interaction between Rco1 and Set2 or the elongating polymerase. However, we did co-purify members of the Rpd3S complex, Eaf3 and Rpd3 (Fig. 3.5C). These results support the idea that it is the chromodomain of Eaf3 that links this complex to H3K36-methylated chromatin. Significantly, other studies have gone on to demonstrate that the chromodomain of Eaf3 can bind specifically to di- and trimethylated H3K36 (Carrozza et al., 2005; Shi et al., 2006).

Discussion

Histone-modifying enzymes such as acetyltransferases and methyltransferases have long been linked to the regulation of transcriptional activation and repression. However, little is know about how these enzymes function in this process, and even less is know about the role of histone modifications in the process of transcription elongation. In this report, we identify a function for the yeast histone methyltransferse Set2, which methylates H3K36 in the coding region of genes. Previous studies of Set2 and H3K36 methylation have implied a role for this modification in transcription elongation (Krogan et al., 2003b; Li et al., 2003; Li et al., 2002; Schaft et al., 2003; Xiao et al., 2003). Here, we identify a specific role for Set2-mediated H3K36 methylation in the recruitment of the Rpd3S deacetylase complex to the coding region of genes. The activities of the Rpd3 complex, in turn, are required to maintain histones in the coding region in a deacetylated state.

While our studies have focused on the identification of a function for Set2, several studies have examined the role of Rpd3 in transcription. Through targeted recruitment to promoters by the DNA-binding protein Ume6, Rpd3 has been shown to exert gene-specific repression of transcription (Kadosh and Struhl, 1998; Rundlett et al., 1998). Furthermore, this enzyme has been linked to a more global, untargeted role in deacetylating histones genome-wide not only in the promoters of genes, but in coding regions as well (Kurdistani et al., 2002; Vogelauer et al., 2000). These previous results suggest that Rpd3 can be recruited by additional, unknown factors or may bind directly to histones. Our results present such an alternative mechanism of recruitment in which Set2-mediated H3K36 methylation serves as a binding platform for the Rpd3S complex. Also, these findings are consistent with previous reports of Set2 repressing transcription when artificially tethered to a promoter (Strahl et al.,

2002). Given the observations here, this repression might be explained, at least in part, by the recruitment of the Rpd3S complex by the activity of Set2.

Interestingly, Rpd3 is not the only histone deacetylase linked to the deacetylation of coding regions. Hos2, a class I histone deacetylase like Rpd3, has been found to deacetylate histones in the coding region of genes and has been implicated in the activation of transcription (Rundlett et al., 1996; Wang et al., 2002). As a predominant member of the Set3 complex, Hos2 is one of two histone deacetylases in this complex (Hst1 is the other) and is required for the complex's integrity along with Set3 (Pijnappel et al., 2001). Consistent with the importance of Hos2 in the Set3 complex, Set3 is also important in the activation of transcription suggesting that Hos2 may function to deacetylate histones in the coding region of genes as part of the Set3 complex (Wang et al., 2002). Significantly, Set2 has been shown to genetically interact with all members of the Set3 complex indicating that Set2 functions in some way similar to the Set3 complex. (Krogan et al., 2003b). Taken together, these data support a role for multiple HDACs in the regulation of histone acetylation in the coding region of genes. It remains to be determined exactly how each HDAC contributes to this regulation, but definitely supports an important role for histone acetylation during transcription elongation.

With these findings, the question then becomes why is it important to maintain low acetylation levels in the coding region of genes? We know from previous studies that histones act as a barrier to effective transcription elongation (Bondarenko et al., 2006; Kireeva et al., 2002). In order for RNAPII to transcribe through the coding region of genes, chromatin undergoes significant rearrangements by chromatin-modifying enzymes involving the disassembly of histones in front of the polymerase and reassembly of histones behind the

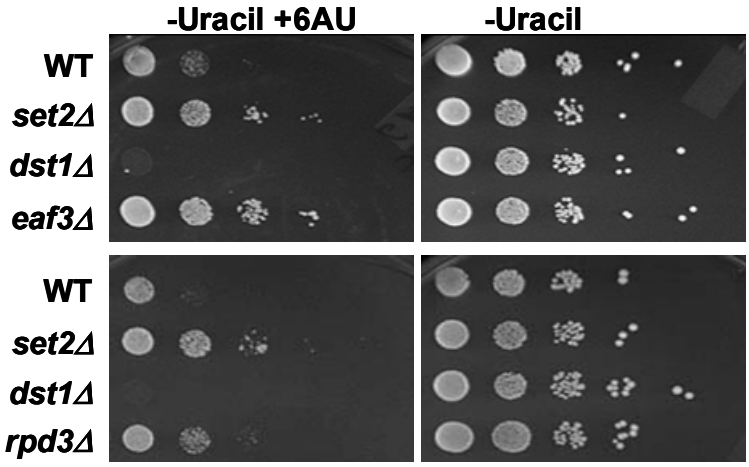
passing polymerase (Lee et al., 2004; Schwabish and Struhl, 2004). In the absence of appropriate reassembly by such factors as the histone chaperones, FACT (facilitates chromatin transcription) and Spt6, DNA sequences recognized as cryptic promoters are exposed leading to inappropriate transcription initiation in the coding region of genes (Belotserkovskaya and Reinberg, 2004; Kaplan et al., 2003; Svejstrup, 2003). Similarly, acetylation plays a role in the disruption of chromatin and is required in the coding region of genes for effective transcription (Kristjuhan et al., 2002; Walia et al., 1998). In a recent report, the absence of Set2 and members of the Rpd3S complex have also been shown to result in cryptic initiation highlighting the importance of the removal of acetylation in the wake of the transcribing polymerase (Carrozza et al., 2005). If acetylation levels remain elevated, chromatin remains in a state permissive for transcription initiation. These findings are consistent with a previous report of RNAPII accumulation in the coding region of genes in the absence of Set2 (Kizer et al., 2005). Thus, Set2-mediated methylation of H3K36 is required to maintain the coding region of genes in a state repressive to transcription initiation by recruiting the Rpd3S deacetylase complex.

Although a role for Set2 in the recruitment of the Rpd3S complex has been identified, several questions remain unanswered. There is still a question of how Eaf3 functions in the NuA4 histone acetyltransferase complex. Does Eaf3 target NuA4 to promoters by binding a specific methylation mark such as H3K4 methylation which is found at the 5' end of genes? If so, how does Eaf3, in its respective complexes, target specific gene regions? Answers to these questions may reside in associated protein members unique to each complex such as Rco1 in the Rpd3S complex. The absence of Rco1 results in the same increased acetylation and cryptic initiation in the coding region of genes as the absence of Eaf3 suggesting that it

plays some role in stabilizing the complex or targeting (Carrozza et al., 2005; Keogh et al., 2005). There is also the question of how other histone deacetylases contribute to the regulation of histone acetylation in the coding region of genes. Do members of the Set3 deacetylase complex play a role in the suppression of cryptic initiation? Furthermore, since Set2 and H3K36 methylation are highly conserved, how well conserved is this mechanism of cryptic initiation suppression in higher eukaryotes? With all of these unanswered questions, there is much to learn from future investigations about the regulation of histone acetylation and transcription elongation.

Figure 3.1

A.


-Uracil +6AU
-Uracil
WT
set2Δ
dst1Δ
eaf3Δ
WT
set2Δ
dst1Δ
rpd3Δ

B.

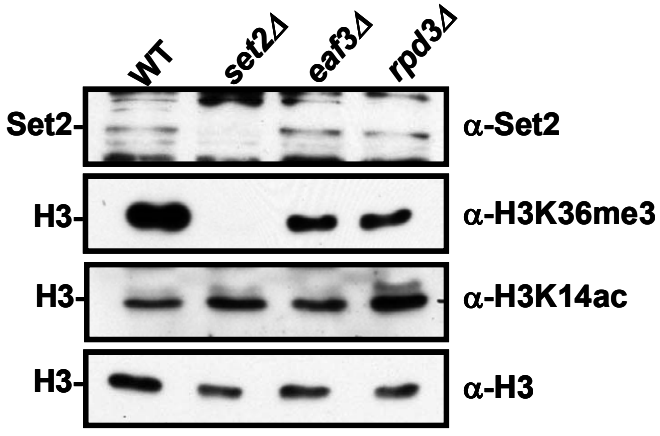
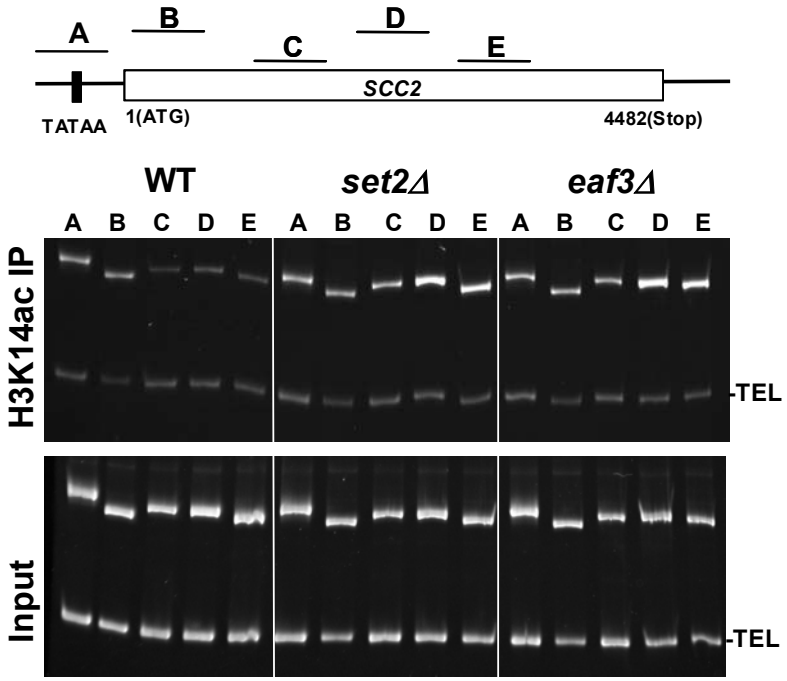

WT
set2Δ
eaf3Δ
rpd3Δ
Set2
α-Set2
H3
α-H3K36me3
H3
α-H3K14ac
H3
α-H3

Figure 3.1. Deletion of *RPD3, EAF3,* and *SET2* result in similar resistance to the transcription elongation drug 6-azauracil. *(A)* Wild-type, *set2Δ*, *eaf3Δ*, *rpd3Δ*, or *dst1Δ* strains were transformed with the *Ura3*$^{+}$ plasmid pRS316 and spotted at a serial dilution of 1:10 on synthetic defined-uracil medium with or without 6-azauracil (6AU, 100 μg/ml). Strains were grown at 30 ºC for 2-3 days before being photographed. *(B)* The Rpd3S complex is not responsible for the stability or activity of Set2. Whole-cell extracts (WCEs) prepared from wild-type (WT), *set2Δ*, *eaf3Δ*, and *rpd3Δ* strains were probed with antibodies against Set2, H3K36me3, and H3K14ac. An antibody specific for the C-terminus of H3 (α-H3) was used as a loading control. The location of Set2 and histone H3 are indicated.

Figure 3.2

A.


A
B
C
D
E
SCC2
TATAA
1(ATG)
4482(Stop)
WT
set2Δ
eaf3Δ
A B C D E A B C D E A B C D E
H3K14ac IP
TEL
Input
TEL

B.

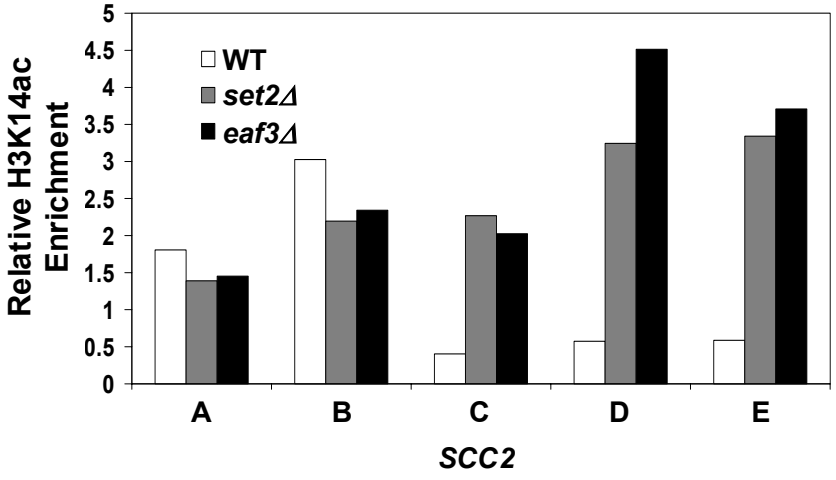

Relative H3K14ac Enrichment
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
WT
set2Δ
eaf3Δ
A
B
C
D
E
SCC2

Figure 3.2. Similar to Eaf3, Set2 regulates histone acetylation patterns on genes. *(A)* Chromatin immunoprecipitations (ChIP) assays were used to monitor the location of H3K14 acetylation on the actively transcribed gene *SCC2* in wild-type, *set2Δ*, or *eaf3Δ* strains using an H3K14ac-specific antibody. DNA from enriched precipitates (IP) were isolated and used in PCR reactions with promoter and coding region-specific primer pairs for the *SCC2* gene (middle). A DNA fragment from an untranscribed subtelomeric region on chromosome VI-R (TEL) was used as a loading control to normalize and calculate the relative enrichment of gene sequences in immunoprecipitated samples. A schematic of the *SCC2* gene and primer pairs are indicated (top). *(B)* Relative enrichment values shown on the y axes were calculated by dividing the ratio of band intensities for IP DNA/TEL with the ratio of intensities for the input DNA/TEL. The gel and graph are representative of two independent repeats.

Figure 3.3

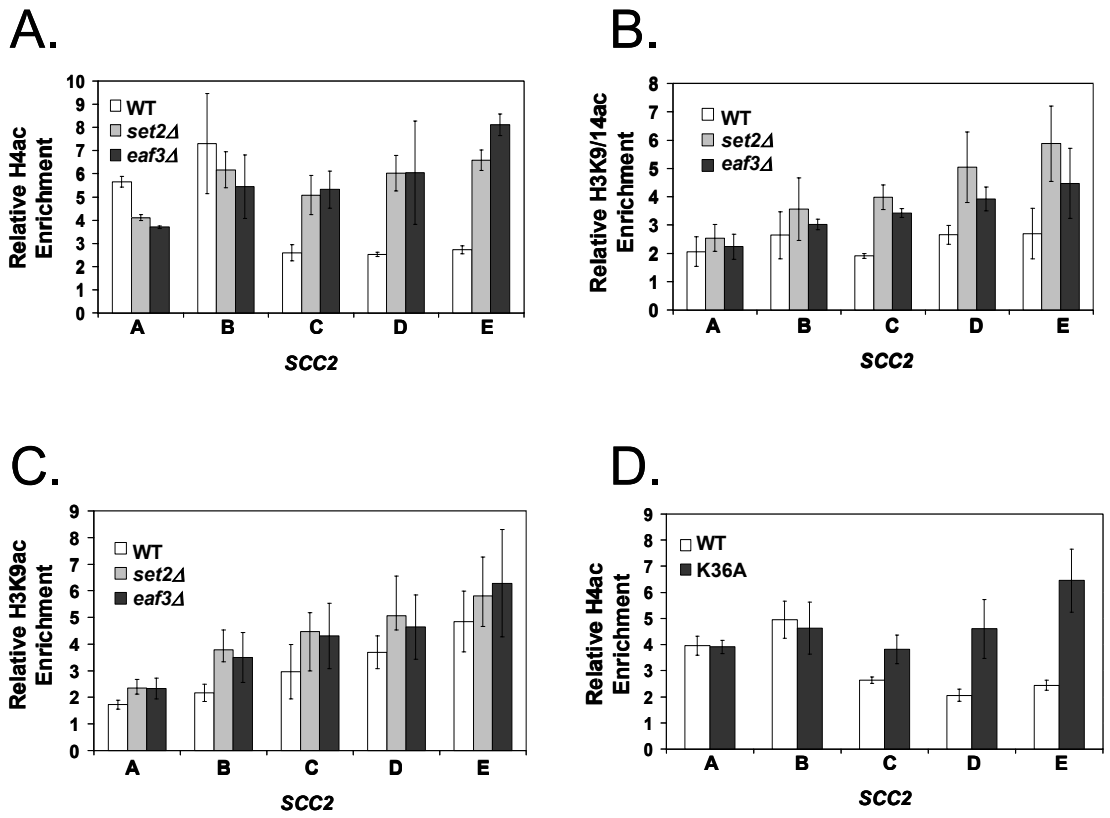

A.
Relative H4ac Enrichment
WT
set2Δ
eaf3Δ
A
B
C
D
E
SCC2
B.
Relative H3K9/14ac Enrichment
WT
set2Δ
eaf3Δ
A
B
C
D
E
SCC2
C.
Relative H3K9ac Enrichment
WT
set2Δ
eaf3Δ
A
B
C
D
E
SCC2
D.
Relative H4ac Enrichment
WT
K36A
A
B
C
D
E
SCC2

Figure 3.3. Set2 regulates H3 and H4 acetylation in the transcribed region of genes via H3K36 methylation. *(A-C)* Wild-type, *set2Δ*, or *eaf3Δ* strains were analyzed by ChIP for levels of the indicated histone modifications on the *SCC2* gene as in Fig. 3.2. The data shown represents the average of two independent experiments. *(D)* Wild-type or H3K36 mutant (K36A) strains were analyzed by ChIP for H4ac. The date shown represents the average of three independent experiments. The standard error of the mean is indicated.

Figure 3.4

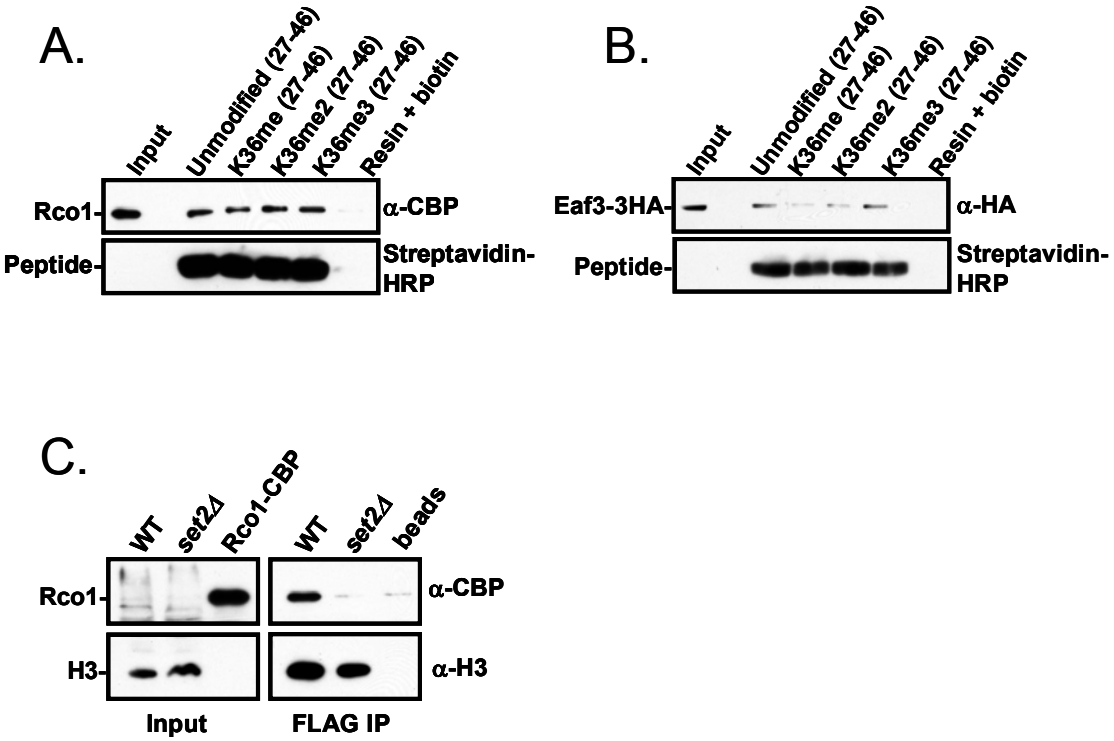

A.
Input
Unmodified (27-46)
K36me (27-46)
K36me2 (27-46)
K36me3 (27-46)
Resin + biotin
Rco1-
α-CBP
Peptide-
Streptavidin-HRP
B.
Input
Unmodified (27-46)
K36me (27-46)
K36me2 (27-46)
K36me3 (27-46)
Resin + biotin
Eaf3-3HA-
α-HA
Peptide-
Streptavidin-HRP
C.
WT
set2Δ
Rco1-CBP
WT
set2Δ
beads
Rco1-
α-CBP
H3-
α-H3
Input
FLAG IP

Figure 3.4. Set2-mediated H3K36 methylation is required for the Rpd3S deacetylase complex interaction with chromatin. *(A)* Biotinylated peptides were bound to streptavidin sepharose followed by incubation with TAP-purified Rpd3S complex (Rco1-CBP). Bead-bound proteins were analyzed by immunoblot for associated Rpd3S complex (α-CBP). Biotinylated peptide loading was monitored by cutting the bottom portion of the protein gel in *(A)* and transferring to PVDF followed by immunobloting using HRP-conjugated streptavidin. *(B)* Experiments were carried out as in *(A)* except bound peptides were incubated with extracts from cells containing HA-tagged Eaf3. *(C)* Nucleosomes purified from FLAG-tagged H2B wild-type and *set2Δ* strains were incubated with TAP-purified Rpd3S complex (Rco1-CBP). Bead-bound proteins were analyzed by immunoblot for bound nucleosomes (α-H3) and associated Rpd3S complex (α-CBP). The location of Rco1 and H3 are indicated. The input WCEs and Rpd3S complex were also examined by immunoblot to monitor the presence of nucleosomes and Rco1.

Figure 3.5

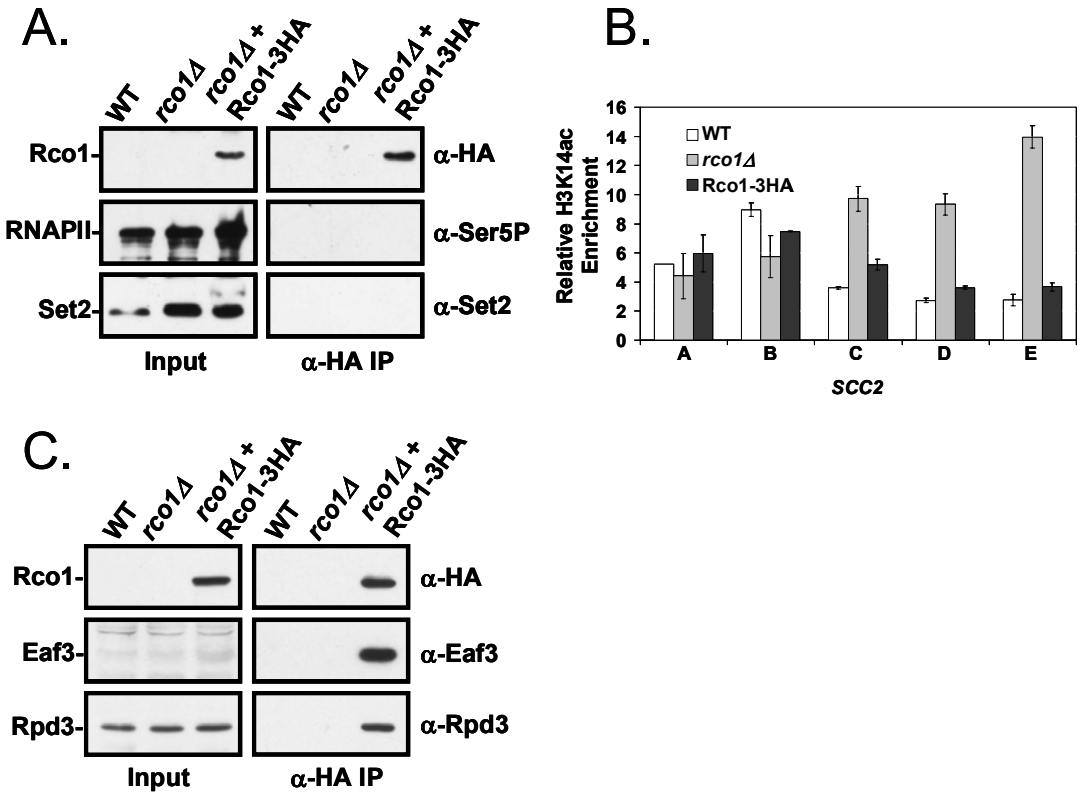

A.
WT
rco1Δ
rco1Δ + Rco1-3HA
WT
rco1Δ
rco1Δ + Rco1-3HA
Rco1-
RNAPII-
Set2-
α-HA
α-Ser5P
α-Set2
Input
α-HA IP
B.
Relative H3K14ac Enrichment
WT
rco1Δ
Rco1-3HA
16
14
12
10
8
6
4
2
0
A
B
C
D
E
SCC2
C.
WT
rco1Δ
rco1Δ + Rco1-3HA
WT
rco1Δ
rco1Δ + Rco1-3HA
Rco1-
Eaf3-
Rpd3-
α-HA
α-Eaf3
α-Rpd3
Input
α-HA IP

Figure 3.5. The Rpd3S complex does not interact with Set2 or elongating RNAPII. *(A)* Whole-cell extracts prepared from WT and *rco1Δ* strains that were transformed with an empty vector or the Rco1-3HA construct were immunoprecipitated with α-HA agarose followed by immunoblot analysis using antibodies directed against Rco1 (α-HA), Ser5 phosphorylated CTD (H14, α- Ser5P), and Set2 (α-Set2). The locations of Rco1, RNAPII, and Set2 are indicated. The input WCEs were also examined by immunoblot to monitor the presence of Rco1, RNAPII and Set2. *(B)* Tagged Rco1 functionally restores acetylation patterns on genes. Wild-type and *rco1Δ* strains that were transformed with an empty vector or the Rco1-3HA construct were analyzed by ChIP for H3K14ac levels on the *SCC2* gene. The data shown represents the average of two independent experiments. The standard error of the mean is indicated. *(C)* Rco1 associates with members of the Rpd3S complex. α-HA immunoprecipitations were performed as in *(A)* followed by immunoblot analysis using antibodies against Rco1 (α-HA), Eaf3 (α-Eaf3), and Rpd3 (α-Rpd3). The locations of Rco1, Eaf3, and Rpd3 are indicated. The input WCEs were also examined by immunoblot to monitor the presence of Rco1, Eaf3, and Rpd3.

Acknowledgements

The data in panel A of Figure 3.1 was generated by Kelby Kizer. The data presented in panel D of Figure 3.5 was published in a collaborative paper and has been reprinted from Cell, vol. 123, "Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex", pp 593-605, Copyright (2005), with permission from Elsevier. The authors of this work included Michael-Christopher Keogh, Siavash Kurdistani, Seong Hoon Ahn, Vladimir Podolny, Sean Collins, Maya Schuldiner, Kayu Chin, Thanuja Punna, Natalie Thompson, Charles Boone, Andrew Emili, Jonathan Weissman, Timothy Hughes, Brian Strahl, Michael Grunstein, Jack Greenblatt, Stephen Buratowski, Nevan Krogan, and myself. The pRS313-Eaf3-3HA-SSN6 (Eaf3-3HA) and pRS313-3HA-SSN6 plasmids were kind gifts of Michael Keogh (Albert Einstein College of Medicine). The TAP-purified Rpd3S complex (Rco1-CBP) was kindly provided by Nevan Krogan (UCSF).

References

Archambault, J., Lacroute, F., Ruet, A., and Friesen, J. D. (1992). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. Mol Cell Biol *12*, 4142-4152.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Belotserkovskaya, R., and Reinberg, D. (2004). Facts about FACT and transcript elongation through chromatin. Curr Opin Genet Dev *14*, 139-146.

Berger, S. L. (2002). Histone modifications in transcriptional regulation. Curr Opin Genet Dev *12*, 142-148.

Bondarenko, V. A., Steele, L. M., Ujvari, A., Gaykalova, D. A., Kulaeva, O. I., Polikanov, Y. S., Luse, D. S., and Studitsky, V. M. (2006). Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. Mol Cell *24*, 469-479.

Bottomley, M. J. (2004). Structures of protein domains that create or recognize histone modifications. EMBO Rep *5*, 464-469.

Brehm, A., Tufteland, K. R., Aasland, R., and Becker, P. B. (2004). The many colours of chromodomains. Bioessays *26*, 133-140.

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*, 3286-3295.

Brown, C. E., Howe, L., Sousa, K., Alley, S. C., Carrozza, M. J., Tan, S., and Workman, J. L. (2001). Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. Science *292*, 2333-2337.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev *14*, 155-164.

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*, 581-592.

Cosma, M. P. (2002). Ordered recruitment: gene-specific mechanism of transcription activation. Mol Cell *10*, 227-236.

Eisen, A., Utley, R. T., Nourani, A., Allard, S., Schmidt, P., Lane, W. S., Lucchesi, J. C., and Cote, J. (2001). The yeast NuA4 and Drosophila MSL complexes contain homologous subunits important for transcription regulation. J Biol Chem *276*, 3484-3491.

Exinger, F., and Lacroute, F. (1992). 6-Azauracil inhibition of GTP biosynthesis in Saccharomyces cerevisiae. Curr Genet *22*, 9-11.

Gerber, M., and Shilatifard, A. (2003). Transcriptional elongation by RNA polymerase II and histone methylation. J Biol Chem *278*, 26303-26306.

Grant, P. A. (2001). A tale of histone modifications. Genome Biol *2*, REVIEWS0003.

Grewal, S. I., and Rice, J. C. (2004). Regulation of heterochromatin by histone methylation and small RNAs. Curr Opin Cell Biol *16*, 230-238.

Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell *113*, 429-432.

Holde, K. E. v. (1988). Chromatin, In Springer Series in Molecular Biology, A. Rich, ed. (New York: Springer-Verlag), pp. 111-148.

Iizuka, M., and Smith, M. M. (2003). Functional consequences of histone modifications. Curr Opin Genet Dev *13*, 154-160.

Joshi, A. A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell *20*, 971-978.

Kadosh, D., and Struhl, K. (1998). Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. Mol Cell Biol *18*, 5121-5127.

Kaplan, C. D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096-1099.

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.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Kireeva, M. L., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M., and Studitsky, V. M. (2002). Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. Mol Cell *9*, 541-552.

Kizer, K. O., Phatnani, H. P., Shibata, Y., Hall, H., Greenleaf, A. L., and Strahl, B. D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol *25*, 3305-3316.

Kristjuhan, A., Walker, J., Suka, N., Grunstein, M., Roberts, D., Cairns, B. R., and Svejstrup, J. Q. (2002). Transcriptional inhibition of genes with severe histone h3 hypoacetylation in the coding region. Mol Cell *10*, 925-933.

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

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C.*, et al.* (2003b). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Kurdistani, S. K., Robyr, D., Tavazoie, S., and Grunstein, M. (2002). Genome-wide binding map of the histone deacetylase Rpd3 in yeast. Nat Genet *31*, 248-254.

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*, 116-120.

Lee, C. K., Shibata, Y., Rao, B., Strahl, B. D., and Lieb, J. D. (2004). Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat Genet *36*, 900-905.

Li, B., Howe, L., Anderson, S., Yates, J. R., 3rd, and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *278*, 8897-8903.

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. J Biol Chem *277*, 49383-49388.

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.

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*, 709-719.

Pijnappel, W. W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Seraphin, B., Aasland, R., and Stewart, A. F. (2001). The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. Genes Dev *15*, 2991-3004.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E.*, et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Reid, J. L., Moqtaderi, Z., and Struhl, K. (2004). Eaf3 regulates the global pattern of histone acetylation in Saccharomyces cerevisiae. Mol Cell Biol *24*, 757-764.

Robert, F., Pokholok, D. K., Hannett, N. M., Rinaldi, N. J., Chandy, M., Rolfe, A., Workman, J. L., Gifford, D. K., and Young, R. A. (2004). Global position and recruitment of HATs and HDACs in the yeast genome. Mol Cell *16*, 199-209.

Rundlett, S. E., Carmen, A. A., Kobayashi, R., Bavykin, S., Turner, B. M., and Grunstein, M. (1996). HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. Proc Natl Acad Sci U S A *93*, 14503-14508.

Rundlett, S. E., Carmen, A. A., Suka, N., Turner, B. M., and Grunstein, M. (1998). Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3. Nature *392*, 831-835.

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*, 2475-2482.

Schwabish, M. A., and Struhl, K. (2004). Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. Mol Cell Biol *24*, 10111-10117.

Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M. R.*, et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C., and Boeke, J. D. (2000). A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein family. Proc Natl Acad Sci U S A *97*, 6658-6663.

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*, 1298-1306.

Suka, N., Suka, Y., Carmen, A. A., Wu, J., and Grunstein, M. (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. Mol Cell *8*, 473-479.

Svejstrup, J. Q. (2003). Transcription. Histones face the FACT. Science *301*, 1053-1055.

Tong, A. H., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C. W., Bussey, H.*, et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Vogelauer, M., Wu, J., Suka, N., and Grunstein, M. (2000). Global histone acetylation and deacetylation in yeast. Nature *408*, 495-498.

Walia, H., Chen, H. Y., Sun, J. M., Holth, L. T., and Davie, J. R. (1998). Histone acetylation is required to maintain the unfolded nucleosome structure associated with transcribing DNA. J Biol Chem *273*, 14516-14522.

Wang, A., Kurdistani, S. K., and Grunstein, M. (2002). Requirement of Hos2 histone deacetylase for gene activity in yeast. Science *298*, 1412-1414.

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*, 654-663.

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

Zhang, W., Bone, J. R., Edmondson, D. G., Turner, B. M., and Roth, S. Y. (1998). Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. Embo J *17*, 3155-3167.

CHAPTER 4

MODIFICATIONS ON HISTONE H3 LYSINE 36: IDENTIFICATION OF HISTONE H3 LYSINE 36 ACETYLATION AS A HIGHLY CONSERVED HISTONE MODIFICATION

Abstract

Histone lysine (K) acetylation is a major mechanism by which cells regulate the structure and function of chromatin, and new sites of acetylation continue to be discovered. Here we identify and characterize histone H3K36 acetylation (H3K36ac). By mass spectrometric analyses of H3 purified from *Tetrahymena thermophila* and *Saccharomyces cerevisiae* (yeast), we find that H3K36 can be acetylated or methylated. Using an antibody specific to H3K36ac, we show that this modification is conserved in mammals. In yeast, genome-wide ChIP-chip experiments show that H3K36ac is localized predominantly to the promoters of RNA polymerase II-transcribed genes, a pattern inversely related to that of H3K36 methylation. The pattern of H3K36ac localization is similar to that of other sites of H3 acetylation, including H3K9ac and H3K14ac. Using histone acetyltransferase complexes purified from yeast, we show that the Gcn5-containing SAGA complex that regulates transcription specifically acetylates H3K36 *in vitro*. Deletion of *GCN5* completely abolishes H3K36ac *in vivo*. These data expand our knowledge of the genomic targets of Gcn5, show H3K36ac is highly conserved, and raise the intriguing possibility that the transition between H3K36ac and H3K36me acts as an "acetyl/methyl switch" governing chromatin function along transcription units.

Introduction

In eukaryotes, the regulation of chromatin structure modulates all DNA-templated processes such as DNA replication and transcription. One major mechanism that regulates the structure and function of chromatin is the covalent modification of histones. A number of different post-translational modifications are known to occur on histones, including acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation (Berger, 2002; Holde, 1988; Nathan et al., 2006; Peterson and Laniel, 2004; Shiio and Eisenman, 2003). While the majority of these modifications are restricted to the flexible N- and C-terminal 'tail' domains of these proteins, a significant number of these marks have been identified in their highly structured globular domains (Cosgrove et al., 2004; Zhang et al., 2003). The function of many of these modifications is not well understood, but several of them have been linked to transcriptional activation and repression, DNA repair, and cell-cycle regulation (Davie and Spencer, 1999; Lachner and Jenuwein, 2002; van Attikum and Gasser, 2005; Wei et al., 1999).

Acetylation and methylation of histone lysine residues, in particular, have been shown to play key roles in the regulation of chromatin structure and function, with the majority of these modifications occurring on histone H3 (Felsenfeld and Groudine, 2003). Acetylation is highly dynamic and has been linked to cellular processes such as transcriptional activation, DNA repair, as well as chromatin assembly (Masumoto et al., 2005; Wang et al., 1997; Ye et al., 2005). Methylation, in contrast, has been considered a stable modification that regulates transcriptional repression and activation, transcriptional elongation, heterochromatin formation, X-inactivation and polycomb-mediated gene silencing (Cao and Zhang, 2004; Grewal and Rice, 2004; Iizuka and Smith, 2003; Kouzarides, 2002; Lee et al., 2005;

Shilatifard, 2004; Sims et al., 2004; Zhang and Reinberg, 2001). However, recent studies have revealed that histone lysine methylation can also be enzymatically reversed (Klose et al., 2006; Shi et al., 2004; Tsukada et al., 2006; Whetstine et al., 2006).

An additional layer of functional complexity in the acetylation or methylation of lysine residues arises from the finding that lysine residues can be targeted for both acetylation and methylation, but not simultaneously. Specifically, it has been shown in mammals and fission yeast that methylation at lysine 9 on H3 (H3K9) serves as a binding site for the recruitment of the chromodomain protein HP1, initiating the formation of heterochromatin (Bannister et al., 2001; Nakayama et al., 2001). However, acetylation on H3K9 prevents this site from being methylated, thus requiring deacetylation prior to methylation for proper heterochromatin formation (Nicolas et al., 2003). This interplay between acetylation, deacetylation and methylation at the same site demonstrates a dynamic relationship between gene activation and repression that has the potential to occur at other histone lysine residues. However, the possibility of "dual" modifications (here "dual" denoting the choice of being either one modification or another) occurring on lysine residues outside of H3K9 has not been widely explored (Zhang et al., 2002).

Previous studies have characterized lysine 36 of histone H3 (H3K36) as a site of methylation mediated by the methyltransferase Set2 in the budding yeast *Saccharomyces cerevisiae* (Strahl et al., 2002). In its methylated form, H3K36 functions in the process of transcriptional elongation and occurs predominantly in the coding region of genes (Shilatifard, 2004; Sims et al., 2004; Xiao et al., 2003). Here, we show that in addition to being a site of methylation, H3K36 can also be a target for acetylation. We find that acetylation at H3K36 is conserved in mammals and, in yeast, is localized predominantly to

the promoters of RNA polymerase II-transcribed genes. We also find that the Gcn5-containing SAGA complex specifically acetylates H3K36 *in vitro* and is required for H3K36ac *in vivo*. Collectively, these results identify H3K36 acetylation as a conserved modification that likely functions in transcription. Because H3K36 is also a site of methylation, these data raise the intriguing possibility that the transition between H3K36ac and H3K36me represents a novel "chromatin switch" involved in the regulation of gene transcription.

Materials and Methods

Yeast Strains

The wild-type histone H3 and H3K36-to-alanine point mutation (K36A) strains are in the WZY42 background and have been described previously (Kizer et al., 2005). The TAP-tagged *GCN5* and *SAS3* strains are in the BY4741 background and were obtained from Open Biosystems. The wild-type (DY150), *gcn5Δ* (DY5925) (Yu et al., 2003), and *sas3Δ* (DY8179) strains are in the W303 background and were kindly provided by David Stillman (University of Utah).

Histone Preparation

Yeast, *Tetrahymena* and mammalian histones were acid-extracted from purified nuclei as previously described (Hake et al., 2006; Strahl and Allis, 2000). For mass spectrometric analyses, *Tetrahymena* and yeast histone H3 were purified by reverse phase (RP)-HPLC as previously described (Medzihradszky et al., 2004; Recht et al., 2006). Typically, acid-extracted core histones were separated on a C8 column (220 X 4.6 mm, Aquapore RP-300; PerkinElmer) using a linear ascending solvent B gradient of 35–60% over 75 min at 1.0 ml/min on a Beckman Coulter System Gold 126 Pump Module and 166 Detector (solvent A was 5% acetonitrile/0.1% TFA in water, and solvent B was 90% acetonitrile/0.1% TFA in water). Peak fractions corresponding to H3 were collected, dried and resuspended in water. A portion of each fraction was used to confirm the presence of purified histones by gel electrophoresis followed by Coomassie blue staining. The remainder of the fraction was used for mass spectrometric (MS) analyses.

Mass Spectrometric Analyses

Mass spectrometric analyses of *Tetrahymena* and yeast H3 were performed as previously described with slight modifications (Hake et al., 2006). Briefly, *Tetrahymena* H3 was digested with trypsin prior to propionylation while yeast H3 lysines were blocked by propionylation followed by digestion with either trypsin or chymotrypsin. Derivatization of H3 with propionylation reagent converts internal lysine residues (monomethylated and endogenously unmodified residues) and the amino terminus to propionyl amides causing the blockage of trypsin cleavage on the C-terminal side of lysines allowing cleavage to occur only C-terminal to arginine (Clayton et al., 2000). Digestion of *Tetrahymena* H3 prior to propionylation allowed for the generation of peptides distinct from that of yeast H3 improving the chances of detecting post-translational modifications. Following digestion and propionylation reactions, samples were gradient eluted (Agilent 1100 Series) directly into a Finnigan linear quadrupole ion trap-Fourier transform (LTQ-FT) mass spectrometer (Thermo Electron) at a flow rate of 100 nl/min operated in the MS/MS data-dependent mode. All MS/MS data were manually validated by inspection of b- and y- type ions.

Electrophoresis and Immunoblot Analyses

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and immunoblot analyses were performed using procedures and reagents from GE Healthcare. The anti-H3K36me3 (α-H3K36me3) rabbit polyclonal antibody was obtained from Abcam (catalog # ab9050) and used at a dilution of 1:2,000-1:5,000. All other rabbit histone modification-specific antibodies were obtained from Upstate Biotechnology and used at the following dilutions: 1:10,000 for H3K18ac (α-H3K18ac, catalog # 07-354), 1:5,000 for H3K14ac (α-H3K14ac, catalog #07-353), and 1:15,000- 1:30,000 for the C-terminus of H3

(α-H3, catalog # 07-690). The H3K36ac antiserum (α-H3K36ac, 1:1,000-1:10,000 dilution) was developed by immunizing a rabbit with a synthetic KLH-conjugated peptide specific for acetylation at H3K36 (C-APATGGVKacKPH). Typically, histones or synthetic histone H3 peptides were resolved on SDS-PAGE gels (15% for histones and 10% for peptides), followed by transfer onto polyvinylidene difluoride (PVDF) membranes (Millipore). In some cases, membranes were stained with Ponceau S (Sigma-Aldrich) to ensure proper protein transfer. After incubation with primary antibody and addition of a horseradish peroxidase (HRP)-conjugated secondary antibody (GE Healthcare), membranes were incubated with enhanced chemiluminescence (ECL)-Plus substrate (GE Healthcare), and proteins were detected by exposure to X-ray films.

For peptide competition experiments, the H3K36ac antibody was incubated overnight at 4 ºC with 5 μg/ml of the following peptides: unmodified histone H3 (amino acids 1-21), K9/14ac histone H3 (amino acids 1-21), K14ac histone H3 (amino acids 1-21), unmodified histone H3 (amino acids 27-46), and K36ac histone H3 (amino acids 27-46). Peptides were either obtained from Upstate Biotechnology or synthesized and verified by mass spectrometry at the University of North Carolina Microprotein Sequencing and Peptide Synthesis Facility.

Chromatin Immunoprecipitation (ChIP) and DNA Microarray (ChIP-chip) Analyses

ChIP assays were performed as previously described (Xiao et al., 2003). Antibodies and amounts used in the immunoprecipitations (IPs) are as follows: α-H3K36ac (2 μl/IP, Upstate Biotechnology, catalog # 07-540), α-H3K36me2 (3 μl/IP, Upstate Biotechnology, catalog # 07-274), and α-H3K9/14ac (3 μl/IP, Upstate Biotechnology, catalog # 06-599). Following DNA recovery, the ChIP-enriched DNAs were amplified as described (Bohlander et al.,

1992). Briefly, two initial rounds of DNA synthesis with T7 DNA polymerase using primer 1 (5'-GTTTCCCAGTCACGATCNNNNNNNNN-3') was followed by 25 cycles of PCR with primer 2 (5'-GTTTCCCAGTCACGATC-3'). Cy3-dUTP or Cy5-dUTP were then incorporated directly with an additional 25 cycles of PCR using primer 2. Direct microarray hybridizations of H3K36ac ChIP vs. H3K36me2 ChIP or H3K9/K14ac ChIP vs. H3K36me2 ChIP were performed using standard procedures (Iyer et al., 2001). This method allowed for the direct comparison between the histone modification patterns, which showed that H3K36ac enrichment was preferentially in the promoter region of genes while H3K36me2 enrichment was preferentially in the transcribed regions. Additional control experiments in which the reference DNA were from H3 ChIPs (for standard nucleosome occupancy normalization), H3K36ac ChIPs from a H3K36 point mutation strain (K36A), or genomic DNA also demonstrated the H3K36ac enrichment to be promoter specific (data not shown). Following hybridizations, the arrays were scanned with a GenePix 4000B scanner and data was extracted with Genepix 5.0 software. Data were normalized such that the median $\log_2$ ratio value for all quality elements on each array equaled zero, and the median of pixel ratio values was retrieved for each spot. Only spots of high quality by visual inspection, with at least 50 pixels of quality data (regression $R^2 > 0.5$) were used for analysis. All data was log-transformed before further analysis. For ChIP-chip data analyses, the $\log_2$ ratio of each spot was transformed to a z-score using the formula $z_x = (X-\mu)/\sigma$, where X is a retrieved spot value, μ is the mean of all retrieved spots from one array, and σ is the standard deviation of all retrieved spots from that same array. Z-scores from three biological replicates were averaged. Raw data can be obtained from the University of North Carolina Microarray

Database at https://genome.unc.edu. Data are also available through GEO (accession number GSE5544).

Purification of Native Yeast Histone Acetyltransferase (HAT) Complexes

For purification of SAGA and NuA3, 4 L of yeast cells containing Gcn5-TAP or Sas3-TAP were grown to an OD_{600} between 1.0 and 2.0. Cells were disrupted by glass bead-beating in a breaking buffer consisting of 50 mM Tris-HCl (pH 8.0), 350 mM NaCl, 10% glycerol, 0.1% triton-X-100 and protease inhibitors (PMSF, aprotinin, leupeptin, and pepstatin) followed by clarification of the extract by ultracentrifugation at 25,000 rpm for 1 h at 4 ºC as described (Grant et al., 1999). TAP-tagged proteins were purified in a one-step procedure from the extracts by directly binding the calmodulin binding protein (CBP) component of the TAP tag to calmodulin resin (Stratagene; 200 µl beads) in the presence of $CaCl_2$ (1 mM final concentration) at 4 ºC for 4 hours. All wash and elution steps were performed in 0.8 X 4 cm Polyprep chromatography columns (BioRad). Protein-bound resin was washed two times with breaking buffer containing $CaCl_2$ at a final concentration of 1 mM. Following washes, bound proteins were eluted in twelve 250 µl fractions with elution buffer (50 mM Tris-HCl (pH 8.0), 350 mM NaCl, 10% glycerol, 2 mM EGTA, and protease inhibitors). Generally, the peak of complex elution was found in fractions 2 and 3. Purified complexes were analyzed by Coomassie blue staining, immunoblot analysis with an anti-Protein A antibody (Sigma-Aldrich) and *in vitro* HAT assays (see below). Additionally, an untagged wild-type strain was included in the purification procedure as a control to confirm that there were no contaminating activities due to nonspecific protein interactions with the calmodulin resin.

In Vitro Histone Acetyltransferase Assays

Acetyltransferase assays were performed as described previously with some modifications (Grant et al., 1999; Mizzen et al., 1999). Briefly, 3 µl of TAP-purified SAGA or NuA3 complex (fraction 2) was incubated with either 2 µg chicken core histones, 2 µg recombinant mononucleosomes, or 2 µg synthetic H3 peptides along with 30 µM unlabeled acetyl coenzyme A (acetyl-CoA, Sigma-Aldrich) in acetyltransferase buffer (50 mM Tris (pH 7.4), 10% glycerol, and protease inhibitors). Reactions were allowed to proceed for 30 min at 30ºC in a total volume of 12 µl and analyzed by SDS-PAGE followed by Coomassie blue staining or immunoblotting. For radiolabeled HAT assays, 3 µl of TAP-purified SAGA or NuA3 complex was incubated with 2 µg synthetic H3 peptides along with 0.125 µCi [^{3}H] acetyl coenzyme A (2-10 Ci/mmol, GE Healthcare) in acetyltransferase buffer. Reactions were allowed to proceed as described above followed by spotting of reaction products onto p81 Whatman paper (Fisher Scientific) and monitoring ^{3}H incorporation by scintillation counting (filter-binding assay). Recombinant mononucleosomes were a gift from Song Tan (Pennslyvania State University) and consisted of *Xenopus* core histones that were bacterially expressed, purified, and reconstituted with the NucB region of the MMTV promoter.

Results

Histone H3 is Acetylated at Lysine 36 in Tetrahymena and Yeast

Methylation at H3K36 is a highly conserved modification found in eukaryotes ranging from *Tetrahymena* to humans, but whether other types of modifications occur at this residue was unknown. To determine if novel modifications occur at H3K36, we analyzed H3 from *Tetrahymena* by tandem mass spectrometry. This organism, a ciliated protozoan, has proven useful in the discovery of other "ON" histone covalent modifications, notably H3K4me (Strahl et al., 1999). *Tetrahymena* H3 was purified by RP-HPLC from acid-extracted bulk histones and digested with both trypsin and chymotrypsin followed by chemical derivatization using propionic anhydride reagent (Syka et al., 2004). Digests were analyzed by on-line nanoflowLC-MS/MS on a linear quadrupole ion trap-Fourier transform (LTQ-FT) mass spectrometer. Using this platform, we determined that H3K36 is acetylated.

Shown in Fig. 4.1A is the MS/MS spectrum of a peptide produced from the propionylated tryptic digest of *Tetrahymena* histone H3. The $[M+2H]^{2+}$ of the parent ion is shown at m/z 542.3062. The high mass accuracy of the LTQ-FT mass spectrometer can easily distinguish between trimethylation and acetylation on peptides ($\Delta m = 0.0364$ Da). The accurate mass of the parent ion recorded was found to be consistent with acetylation on this peptide (+0.18 ppm) and not trimethylation. Importantly, we were able to also detect H3K36 mono-, di-, and trimethylation on other H3 peptides (data not shown), confirming that this residue can be methylated or acetylated.

We additionally surveyed histone H3 purified from budding yeast to determine if H3K36 acetylation might be present in this distinct unicellular organism. Using the approach mentioned above for *Tetrahymena* H3, we identified that yeast H3K36 was also acetylated

(Fig. 4.1B). H3K36 acetylation was observed on a number of different peptides. Fig. 4.1B displays the tandem mass spectrum of the $[M+2H]^{2+}$ ion at m/z 823.4552 from a propionylated digest of one such identified H3 peptide. This peptide was identified to be the 27-40 amino acid H3 fragment and it was found to contain acetylation at both H3K27 and H3K36. These results identify a novel acetylation event at H3K36 that is conserved between *Tetrahymena* and yeast.

To further investigate the occurrence of this modification, we raised an antibody against a synthetic peptide acetylated at H3K36 and analyzed *Tetrahymena* and yeast histones by immunoblot analysis. As shown in Fig. 4.2A, the α-H3K36ac antibody efficiently recognized the RP-HPLC purified *Tetrahymena* H3, originally analyzed by mass spectrometry. We determined the antibody to be specific for H3K36ac, as only a synthetic H3 peptide acetylated at H3K36 could selectively compete away the signal detected by the α-H3K36ac antibody (Fig. 4.2A). Other unmodified or acetylated H3 peptides did not compete for this antibody's detection of H3.

Next, we determined if this antibody could specifically recognize acetylation at H3K36 in yeast. Yeast histones isolated by acid-extraction from a wild-type or H3K36 point mutant (K36A) strain were analyzed by immunoblot analysis using the H3K36ac-specific antibody. Similar to *Tetrahymena*, the α-H3K36ac antibody detected H3 in a wild-type yeast strain but much less so in a strain where H3K36 was mutated to alanine (Fig. 4.2B). However, we note that high concentrations of histones loaded in these assays results in the ability of the α-H3K36 antibody to weakly recognize the backbone of H3 (Fig. 4.2B, K36A lane). Nonetheless, these results confirm our mass spectrometry findings that H3K36ac exists in both *Tetrahymena* and yeast.

H3K36 Acetylation is Preferentially Enriched in the Promoters of RNA Polymerase II-transcribed Genes Genome-wide

Next, we used a ChIP-chip approach to determine the genomic distribution of H3K36ac and how it compared to the distribution of methylation found at H3K36 (Pokholok et al., 2005; Rao et al., 2005). The α-H3K36ac-specific antibody was used in ChIP reactions from yeast whole cell extracts. Enriched genomic DNA fragments were treated with RNase, amplified, and labeled fluorescently. DNA enriched from H3K36me2 ChIPs were prepared in a similar manner, and both samples were hybridized on the same array. Three independent sets of ChIPs were performed. Using this method, we directly compared the patterns of H3K36ac and H3K36me2 using arrays that tiled continuously over the entire genome at a resolution of ~1 kb. We observed a preferential enrichment of H3K36ac at 5′ regulatory regions (bidirectional and unidirectional promoters) relative to coding regions (ORFs) and 3′ UTRs (non-coding region downstream of two convergently transcribed genes) (Fig. 4.3A). Significantly, the H3K36ac pattern was found to be inversely related to that of H3K36me2, which occurs preferentially in the coding region and 3′ UTR of genes (Krogan et al., 2003; Rao et al., 2005; Xiao et al., 2003).

We then compared the location of the H3K36ac modification to that of other well-characterized acetylation sites on H3; namely H3K9 and H3K14 acetylation (H3K9/14ac). Using the same extracts, ChIPs were performed with an antibody that recognizes diacetylated H3K9/14, and the enriched DNA was amplified, labeled, and hybridized to DNA microarrays as described above. These experiments revealed that H3K9/K14ac was localized to promoters in a pattern strikingly similar to the pattern we observed for H3K36ac genome-wide (data not shown and see Fig. 4.3B). These data were also consistent with the H3K9/14 acetylation results obtained by others (Liu et al., 2005; Pokholok et al., 2005). We further

examined the distribution of H3K36ac, H3K9/K14ac, and H3K36me2 using a high-resolution microarray containing probes that covered all of chromosome III at a resolution of 200 bp (and at 100 bp resolution for 1/3 of the chromosome). For both H3K36ac and H3K9/K14ac, the level of acetylation enrichment drops sharply as a function of distance from translational initiation site while H3K36me2 increases (Fig. 4.3B). These data are fully consistent with our analysis using lower-resolution arrays and confirm that these acetyl marks occur preferentially upstream of coding regions while H3K36me2 occurs primarily in the coding region of genes.

We next asked whether H3K36ac associates with genomic regions other than those characteristic of RNA polymerase II (RNAPII) promoters. We found that genomic regions which were transcriptionally silent under our growth conditions, or regions not transcribed by RNAPII including centromeres, telomeres, and silent mating type loci, contained low levels of H3K36ac and H3K9/K14ac (Fig. 4.3C). These results indicate that H3K36ac is associated strictly with active RNAPII regulatory sequences, and suggest that like H3K36me, H3K36ac may function in RNAPII-mediated gene transcription.

The Gcn5-containing SAGA Histone Acetyltransferase Complex Mediates H3K36 Acetylation

We next sought to identify the responsible histone acetyltransferase(s) (HATs) that deposits this mark. Given H3K36 acetylation is enriched in the promoter regions of RNAPII-transcribed genes, and has a pattern of acetylation similar to that of H3K9/14 acetylation, we initially focused on the Gcn5-containing SAGA histone acetyltransferase (HAT) complex that mediates acetylation to the N-terminus of H3 (Grant et al., 1999). Furthermore, we noticed that the amino acid sequence surrounding H3K36 is very homologous to a preferred Gcn5 consensus site of acetylation found at H3K14 (Kuo et al.,

1996) (compare STGG**K**^{14}AP vs. STGGV**K**^{36}KP; underlined residues indicate homology and bold "K" indicates the residue targeted for acetylation).

To test if SAGA would mediate H3K36ac, we TAP-purified SAGA from yeast whole extracts using a tagged form of Gcn5 and then incubated the complex with either unmodified or modified H3 synthetic peptides along with unlabeled acetyl-CoA (acetyl donor) in a HAT assay. Following the HAT reactions, the products were electrophoresed on 10% SDS-PAGE gels and analyzed by immunoblot with the α-H3K36ac antibody. After incubation with purified SAGA, a previously unmodified H3 27-46 amino acid peptide was recognized by the H3K36ac-specific antibody (Fig. 4.4A). Importantly, when these same assays were performed using a radiolabeled form of acetyl-CoA, SAGA was readily able to acetylate an N-terminal H3 peptide (residues 1-21 containing H3K9 and H3K14) as well as the unmodified 27-46 peptide, but showed no activity toward a matched 27-46 peptide acetylated at H3K36 (Fig. 4.4B).

To learn more about the physiological relevance of SAGA-mediated H3K36 acetylation, we next asked if SAGA could acetylate H3K36 in the context of other histone proteins and DNA. Again HAT assays were performed using TAP-purified SAGA in which the complex was incubated with either free chicken core histones or recombinant mononucleosomes. As shown in Fig. 4.4C, SAGA was capable of acetylating both substrates at H3K36.

We also tested the NuA3 HAT complex for an activity that specifically acetylates H3K36. Like SAGA, NuA3 has been identified as a HAT complex that specifically targets H3 for acetylation (Grant et al., 1997). Previous work shows that the catalytic subunit of NuA3, Sas3, has overlapping activities with Gcn5 and can target H3K14 for acetylation *in vivo* (Howe et al., 2001). Unlike SAGA, however, we found that TAP-purified NuA3 was

unable to acetylate H3K36 in the context of H3 synthetic peptides. However, NuA3 could acetylate H3K36 in the context of core histones and nucleosomes (data not shown).

Given the ability of SAGA and NuA3 to catalyze H3K36ac, we next asked whether either one of these HATs was required for H3K36ac *in vivo*. We therefore purified bulk core histones from *GCN5* and *SAS3* deletion strains and analyzed them for H3K36ac by immunoblot analysis. As shown in Fig. 4.4D, acetylation at H3K36 was abolished in the *GCN5* deletion, but not in the *SAS3* deletion. These results reveal that Gcn5 is the major HAT responsible for H3K36ac *in vivo*.

H3K36 Acetylation is Conserved in Mammalian Cells

Although our studies showed that several unicellular organisms contained H3K36ac, we wanted to determine how conserved this modification would be among several diverse multicellular organisms. We therefore isolated histones by acid-extraction from *Tetrahymena*, yeast, mouse, and human cells and probed them for the presence of H3K36ac. As shown in Fig. 4.5A, we found H3K36ac in all of the organisms analyzed, although the relative abundance varied between species. H3K36ac appeared to be most abundant in *Tetrahymena* and yeast, followed by human and mouse. The mouse embryonic fibroblasts used for this study apparently harbor a very low level of H3K36ac, which is consistent with our mass spectrometry experiments that did not detect H3K36ac in histones isolated from these cells (Garcia et al., 2007). In agreement with our human results, a recent mass spectrometry proteomics survey reported the existence of H3K36ac in human HeLa cells (Kim et al., 2006). Although differences exist in the relative abundance of H3K36ac found in these organisms, and perhaps within different cell types, these results reveal that H3K36ac occurs in organisms as diverse as yeast and humans. Our results also underscore differences

in the employment of histone marks in different organisms and in the importance of applying independent assays to assess them.

Discussion

In this report, we identify and characterize a novel site of acetylation on histone H3 at lysine 36. This site was previously determined to be mono-, di-, and trimethylated in a broad range of eukaryotic organisms, and we find that acetylation at this residue is also highly conserved. Furthermore, we have determined that H3K36ac is mediated by the Gcn5-containing SAGA complex in yeast, and is preferentially enriched in the promoter regions of RNAPII-transcribed genes genome-wide. Although the exact function of this modification remains to be elucidated, our data suggests that it is involved in the transcription process.

We and others have found several clues as to how Gcn5 can target H3K36 to acetylate this site. First, previous studies have shown that Gcn5 in isolation can only target H3K14 (Grant et al., 1999; Kuo et al., 1996). However, in its native SAGA complex, this enzyme has an expanded substrate range on H3 that includes H3K9, K18 and K23 (Grant et al., 1999). H3K36ac was not detected in this previous study as only H3 synthetic peptides containing amino acid residues from 5 to 28 were investigated. Our data, therefore, reveal an expanded site utilization pattern by SAGA. Second, the amino acid sequence immediately surrounding H3K36 is very similar to that of H3K14, which is a preferred site of Gcn5 acetylation (compare STGG**K**^{14}AP vs. STGGV**K**^{36}KP; underlined sequences show identity while bold lysines are the acetyl accepting residues). Structural studies of Gcn5 in complex with an H3 peptide containing H3K14 have identified several critical residues immediately surrounding H3K14 (glycine 13 and proline 16) that are important for substrate recognition and high affinity binding (Clements et al., 2003; Poux and Marmorstein, 2003; Rojas et al., 1999). Importantly, these key residues (G-**K**^{14}-X-P) are conserved in the sequence surrounding H3K36. Thus, the high similarity between residues surrounding H3K14 and

H3K36 may explain how Gcn5 is capable of acetylating H3K36 *in vitro* and *in vivo*. Given H3K9/14 acetylation patterns overlap with H3K36ac, it is likely that SAGA mediates a broad H3 acetylation pattern when recruited to gene promoters.

Interestingly, while we found that H3K36ac was present in *Tetrahymena,* yeast, and mammalian cells, the general abundance of this modification varied greatly among the organisms we analyzed (Fig. 4.5A). While many possible reasons could account for these differences, one plausible explanation may be due to the fact that budding yeast contains only a single H3 isoform (H3.3), whereas *Tetrahymena* and mammalian cells contain multiple H3 variants (H3.1, H3.2 and H3.3; see Fig. 4.5B). Since H3.1 and H3.2 in flies and mammals are generally associated with silenced chromatin (which make up a large proportion of their genomes), these histone variants are not likely to be H3K36 acetylated and/or associated with active transcription. Thus, much of the bulk-isolated histones from these more complex organisms contain H3 isoforms that harbor "OFF" marks. This idea, along with our H3K36ac observations, is consistent with earlier studies that show that these same multicellular organisms have much lower levels of H3K4 methylation (a modification associated with transcriptionally competent chromatin) compared to those observed in yeast (Briggs et al., 2001). Given the sequence surrounding H3K36 is highly conserved between these organisms (Fig. 4.5B), the differences observed between species is not likely due to the inability of our antibody to efficiently recognize the site of H3K36.

That H3K36ac is mediated by Gcn5 and is found in the promoters of RNAPII-regulated genes suggests that acetylation at H3K36 may play a role in gene transcription. Previous studies of SAGA and Gcn5 indicate that this enzyme complex is activator recruited and targets acetylation at specific promoters during transcriptional activation (Gregory et al.,

1998; Kuo et al., 1998; Sterner et al., 1999; Wang et al., 1997). While the exact function of H3K36ac is unknown, we speculate that H3K36ac acts in concert with other Gcn5-mediated sites of acetylation to properly regulate transcriptional induction. Such cooperativity would be in agreement with prior studies that have shown the importance of Gcn5-mediated acetylation of multiple sites on H3 for normal cell growth and transcriptional activation (Zhang et al., 1998). Additionally, the acetylation of H3K36 by SAGA, as part of this complex's expanded targeting, is consistent with studies on H4 that show that a cumulative effect of acetylation is associated with the promoters of active genes (Dion et al., 2005).

In contrast to H3K36ac, methylation at H3K36 is found in the coding region of genes and is involved in transcriptional elongation. Recent reports have identified a function for this modification in maintaining an environment within coding regions that is repressive to the activation of intragenic transcription by recruiting the deacetylase complex Rpd3S (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). An important question, therefore, is whether acetylation and methylation activities compete for the same target sites, such as H3K36. Well documented is the finding that H3K9 is subject to either methylation or acetylation (see below), and a report co-published with this one reveals that a significant number of lysines targeted for acetylation are also targeted for methylation, and vise versa (Garcia et al., 2007). These results raise the intriguing possibility that functional interplay between two or more posttranslational modifications at a single lysine residue is a general phenomenon that drives distinct biological effects within chromatin. To date, the best example of functional interplay between methylation and acetylation is with H3K9, at which acetylation is removed prior to methylation by Su(var)3-9 in the promoters of genes (Schotta et al., 2002). This activity is required for the recruitment of HP1/Swi6 that leads to the

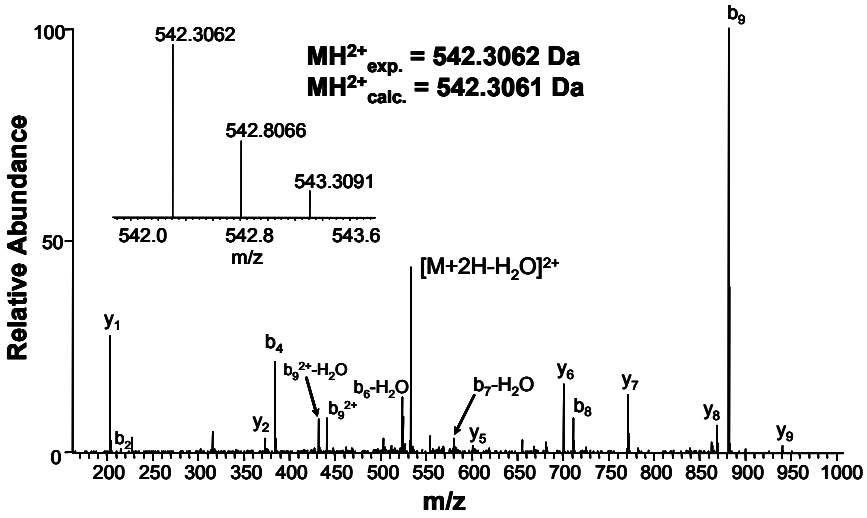
formation of heterochromatin in both fission yeast and multicellular eukaryotes. Although our analyses do not reveal if functional interplay occurs at this site, future investigations will aim to determine whether SAGA and Set2 regulate transcription initiation events through competition for H3K36.

Figure 4.1

A.

Tetrahymena H3

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	
	144	215	312	383	484	541	598	711	881	1083
Pr-S	**A**	**P**	**A**	**T**	**G**	**G**	**I**	**K^{ac}**	**K-Pr**	
1083	940	869	772	701	600	543	486	373	203	
	y_9	y_8	y_7	y_6	y_5	y_4	y_3	y_2	y_1	



B.

Yeast H3

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	b_{13}	
	227	314	385	482	569	670	727	784	883	1053	1237	1334	1471	1645
Pr-K^{ac}	**S**	**A**	**P**	**S**	**T**	**G**	**G**	**V**	**K^{ac}**	**K-Pr**	**P**	**H**	**R**	
1645	1419	1332	1261	1164	1077	976	919	862	763	593	409	312	175	
	y_{13}	y_{12}	y_{11}	y_{10}	y_9	y_8	y_7	y_6	y_5	y_4	y_3	y_2	y_1	

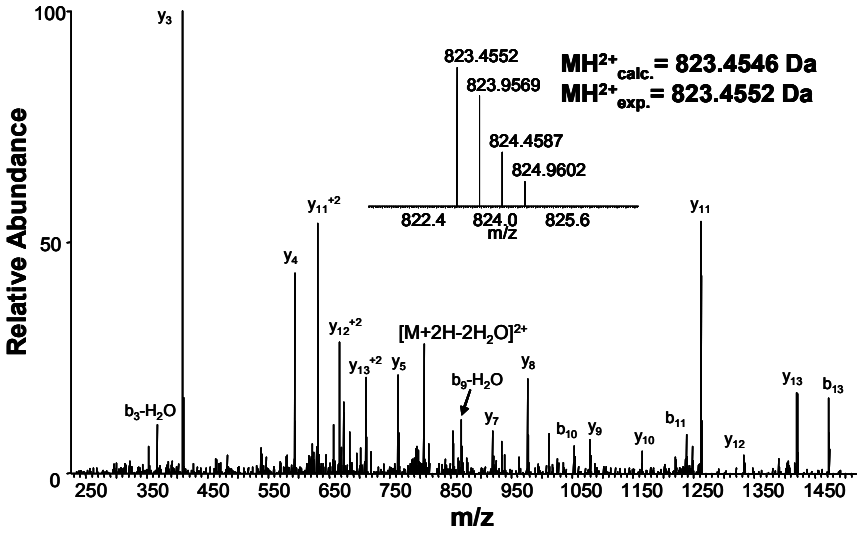


Figure 4.1. Identification of H3K36 acetylation in *Tetrahymena* and yeast by mass spectrometry. *(A)* Displayed is the MS/MS fragmentation spectrum of the $[M+2H]^{2+}$ parent ion from an H3 peptide (histone H3, amino acids 29-37) derived from a propionylated digest of *Tetrahymena* H3. Inset shows full MS of parent ion (whole peptide), from which the fragmentation spectrum was taken, at 542.3062 m/z. Scale for the y-axis of inset represents relative abundance of the parent ion and is identical to the fragmentation spectrum y-axis. Accurate mass (inset) indicates acetylation and not trimethylation on this peptide (+0.18 ppm error), while fragment ions (full spectrum) show H3K36 as the acetylation site. Experimentally observed ($MH^{2+}_{exp.}$) and calculated masses ($MH^{2+}_{calc.}$) for this acetylated peptide are indicated. Above the spectrum is the peptide sequence in which predicted b-type ions, which contain the amino terminus of the peptide, are immediately above the sequence. Predicted y-type ions, which contain the carboxyl terminus, are immediately below the sequence. Ions observed in the spectrum are underlined and represent masses associated with the fragmented peptides from the MS/MS analyses. Note the addition of propionyl groups (Pr) which add 56 Da on amino terminus residues and unmodified lysine residues due to chemical derivatization with propionic anyhydride reagent (see experimental procedures for explanation). *(B)* Same as in *(A)* except peptides were derived from digested RP-HPLC purified yeast H3. Displayed is the MS/MS fragmentation spectrum of the $[M+2H]^{2+}$ parent ion at m/z 823.4552. This peptide was identified as the 27-40 fragment from yeast histone H3. Inset shows full MS of parent ion at m/z 823.4552. Accurate mass indicates the addition of two acetylation modifications and not trimethylation on this peptide (+0.73 ppm error), while fragment ions show H3K27 and H3K36 as the acetylation sites. Experimentally observed ($MH^{2+}_{exp.}$) and calculated masses ($MH^{2+}_{calc.}$) for this acetylated peptide are

indicated. As in *(A)*, b- and y-type ions observed in the spectrum are underlined and the peptide contains the addition of propionyl groups (Pr) on unmodified lysine and amino terminus residues.

Figure 4.2

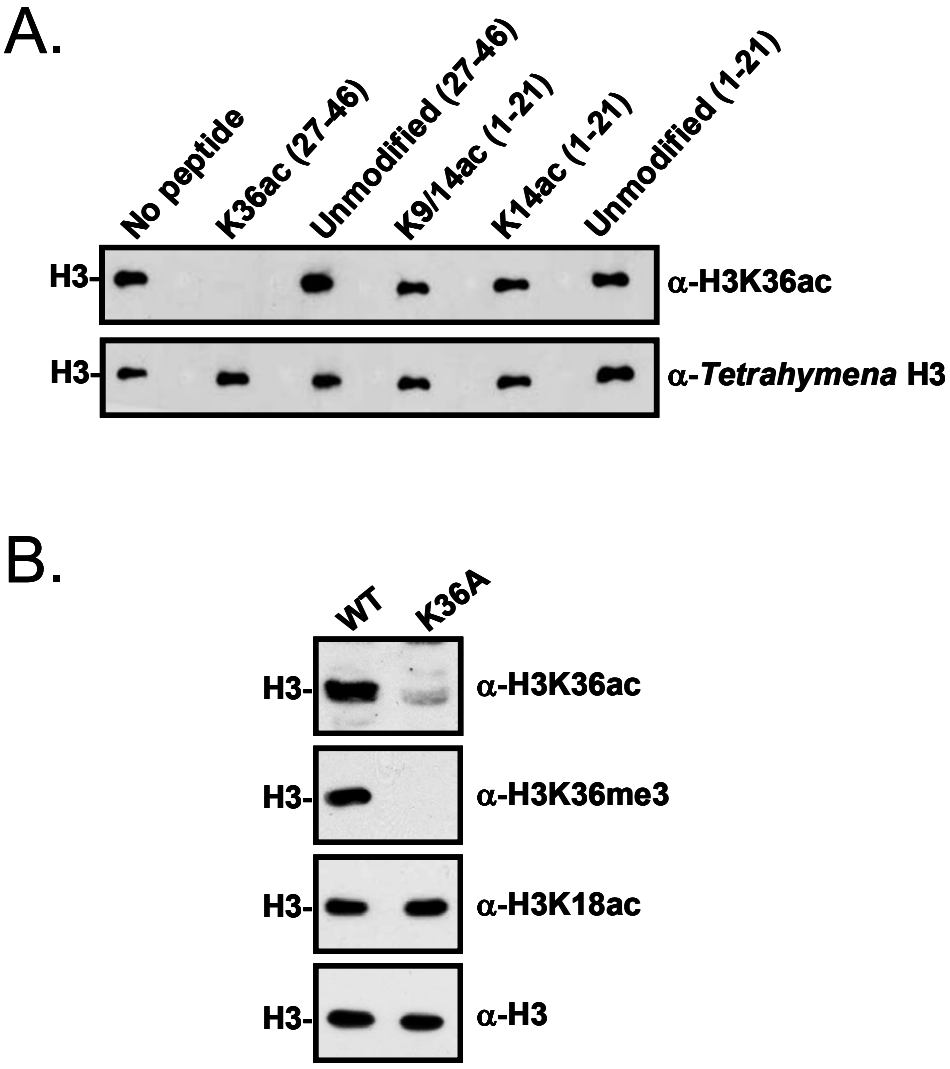

A.
No peptide
K36ac (27-46)
Unmodified (27-46)
K9/14ac (1-21)
K14ac (1-21)
Unmodified (1-21)
H3-
α-H3K36ac
H3-
α-Tetrahymena H3
B.
WT
K36A
H3-
α-H3K36ac
H3-
α-H3K36me3
H3-
α-H3K18ac
H3-
α-H3

Figure 4.2. Detection of H3K36 acetylation in *Tetrahymena* and yeast using a specific antiserum. *(A)* An antibody specific to H3K36 acetylation recognizes *Tetrahymena* H3. RP-HPLC *Tetrahymena* H3 (same as used in Fig. 4.1A) was loaded onto adjacent lanes and resolved on a 15% SDS-PAGE gel. Following transfer to a polyvinylidene difluoride (PVDF) membrane, each lane was separated and probed with an α-H3K36 acetyl antiserum (α-H3K36ac) that was preincubated with different unmodified or modified H3 synthetic peptides as indicated. The same blots were stripped and reprobed with an antibody specific for *Tetrahymena* H3 (α-*Tetrahymena* H3) as a loading control. *(B)* The α-H3K36ac antibody specifically recognizes H3K36 acetylation in yeast. Acid-extracted histones prepared from a wild-type or H3K36 point mutant yeast strain (K36A) were resolved on a 15% SDS-PAGE gel, transferred to a PVDF membrane, and probed for H3K36ac. An antibody specific for the C-terminus of H3 (α-H3) was used as a loading control. Antibodies specific for H3K18 acetylation (α-H3K18ac) and H3K36 trimethylation (α-H3K36me3) were used as additional controls.

Figure 4.3

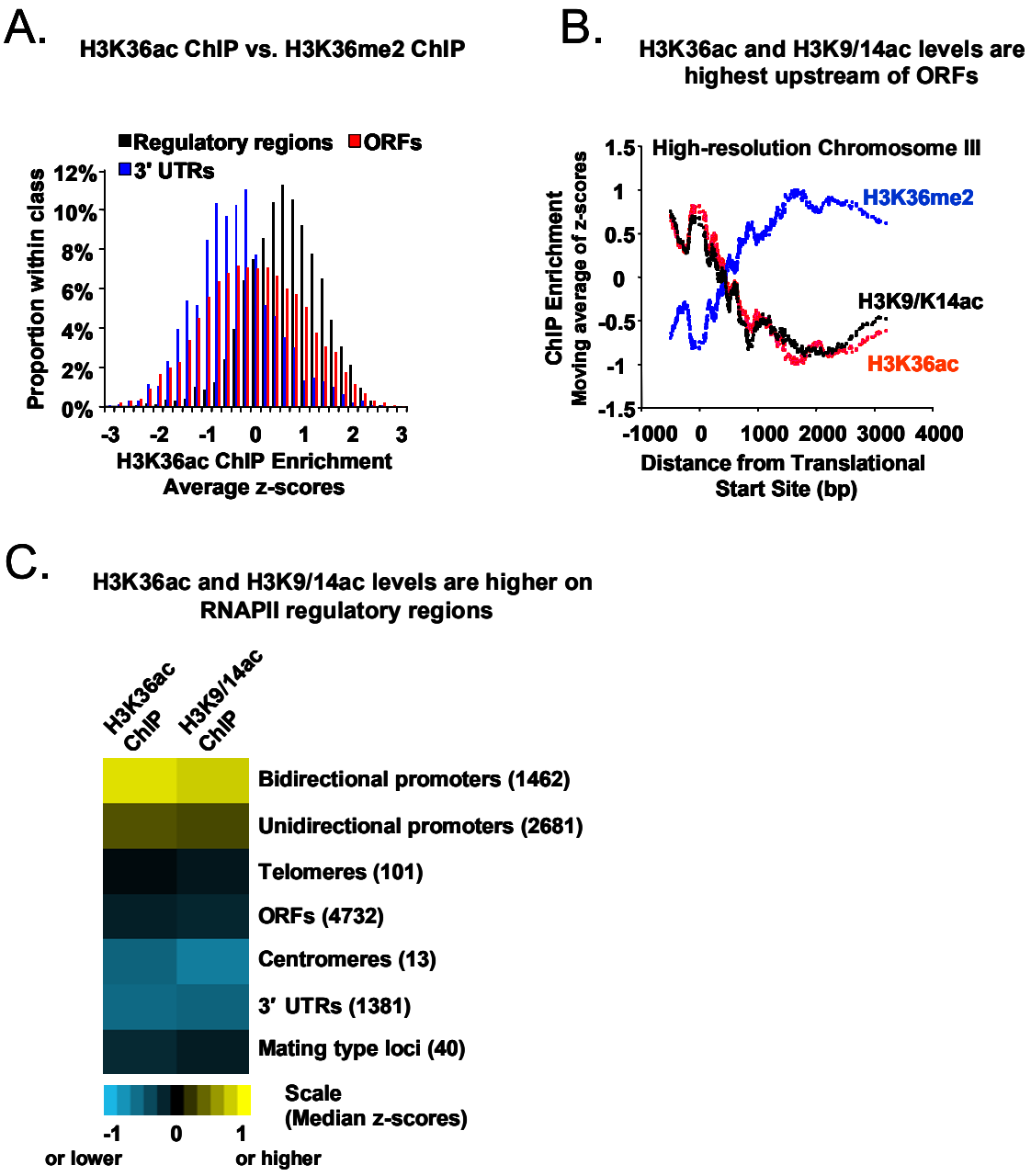

A.
H3K36ac ChIP vs. H3K36me2 ChIP
Regulatory regions
ORFs
3' UTRs
Proportion within class
12%
10%
8%
6%
4%
2%
0%
-3
-2
-1
0
1
2
3
H3K36ac ChIP Enrichment
Average z-scores
B.
H3K36ac and H3K9/14ac levels are highest upstream of ORFs
High-resolution Chromosome III
ChIP Enrichment
Moving average of z-scores
1.5
1
0.5
0
-0.5
-1
-1.5
H3K36me2
H3K9/K14ac
H3K36ac
-1000
0
1000
2000
3000
4000
Distance from Translational Start Site (bp)
C.
H3K36ac and H3K9/14ac levels are higher on RNAPII regulatory regions
H3K36ac ChIP
H3K9/14ac ChIP
Bidirectional promoters (1462)
Unidirectional promoters (2681)
Telomeres (101)
ORFs (4732)
Centromeres (13)
3' UTRs (1381)
Mating type loci (40)
Scale
(Median z-scores)
-1
or lower
0
1
or higher

Figure 4.3. H3K36 acetylation is localized predominantly to the promoters of RNA polymerase II-transcribed genes genome-wide. *(A)* The distribution of average z-scores (units are standard deviation from the mean) for 5′ regulatory regions (black), ORFs (red), and 3' UTRs (blue) derived from ChIP-chip experiments in which H3K36ac ChIPs were compared directly to H3K36me2 ChIPs. Thus, the H3K36ac and H3K36me2 ratios shown here are inversely related. Similar promoter enrichment results for H3K36ac were obtained when the H3K36ac ChIPs were compared to a genomic DNA reference, or references composed of histone H3 ChIPs (experimental procedures). *(B)* A moving-average plot (window size=40, step size=1) of average z-scores from three independent experiments comparing H3K36ac ChIPs (red) , H3K9/14ac ChIPs (black) and H3K36me2 ChIPs (blue) on a high resolution DNA microarray covering all of chromosome III. ChIP enrichment is plotted as a function of the distance from the translational start site among genes greater than 1 kb in length. *(C)* H3K36ac distribution genome-wide. Colors (scale at bottom) represent the median of all z-scores recorded from all arrayed elements in the indicated functional class (labeled on right, number of elements indicated in parentheses). Data were derived from three independent replicates.

Figure 4.4

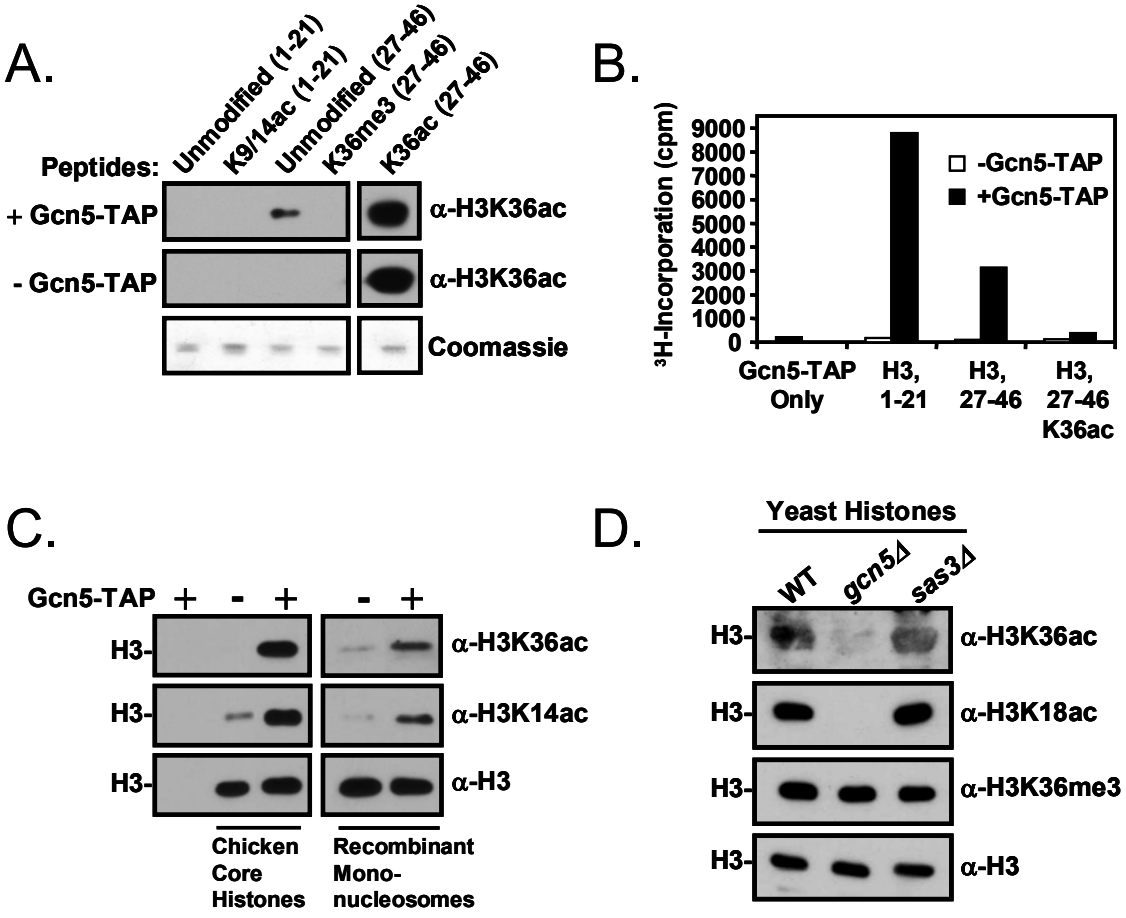
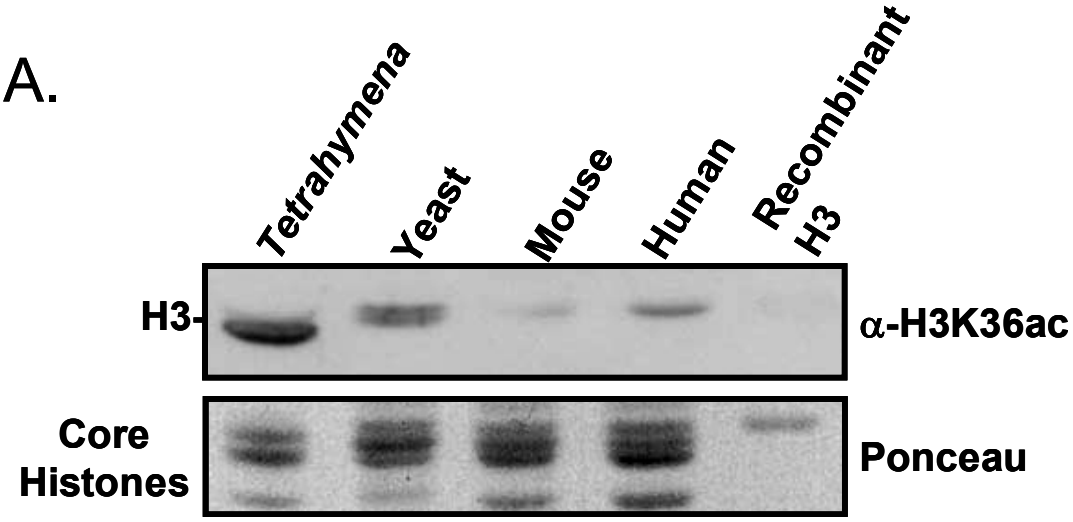

A.
Peptides:
Unmodified (1-21)
K9/14ac (1-21)
Unmodified (27-46)
K36me3 (27-46)
K36ac (27-46)
+ Gcn5-TAP
- Gcn5-TAP
α-H3K36ac
α-H3K36ac
Coomassie
B.
³H-Incorporation (cpm)
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-Gcn5-TAP
+Gcn5-TAP
Gcn5-TAP Only
H3, 1-21
H3, 27-46
H3, 27-46 K36ac
C.
Gcn5-TAP + - + - +
H3-
α-H3K36ac
H3-
α-H3K14ac
H3-
α-H3
Chicken Core Histones
Recombinant Mono-nucleosomes
D.
Yeast Histones
WT
gcn5Δ
sas3Δ
H3-
α-H3K36ac
H3-
α-H3K18ac
H3-
α-H3K36me3
H3-
α-H3

Figure 4.4. The Gcn5-containing SAGA complex acetylates H3K36. *(A)* Shown are the results of a HAT assay in which TAP-purified SAGA complex was incubated with either unmodified or modified H3 synthetic peptides along with unlabeled acetyl coenzyme A (acetyl-CoA). Reaction products were resolved on a 10% SDS-PAGE gel, transferred to a PVDF membrane, and analyzed by immunoblot for H3K36ac. An H3 synthetic peptide acetylated at H3K36 was used as a control for antibody detection. Parallel reactions were performed and examined by Coomassie staining to monitor loading (lower panel). *(B)* Displayed is a graph representing the results of a HAT assay in which TAP-purified SAGA complex was incubated with either unmodified or modified H3 synthetic peptides along with [^{3}H] labeled acetyl-CoA. ^{3}H incorporation was analyzed by filter-binding assay and monitored by scintillation counting. *(C)* Shown are the results of a HAT assay in which TAP-purified SAGA complex was incubated with either chicken core histones or recombinant mononucleosomes along with unlabeled acetyl-CoA as in *(A)*. Reaction products were resolved on a 15% SDS-PAGE gel, transferred to PVDF membrane, and analyzed by immunoblot for H3K36ac. Parallel reactions were performed and analyzed by immunoblot for H3K14ac as a control. The same blot was stripped and reprobed with an antibody specific for the C-terminus of H3 (α-H3) to monitor loading. Note slight antibody detection of histone H3 backbone in the absence of TAP-purified SAGA in the recombinant (unmodified) mononucleosomes reactions. *(D)* Gcn5 is responsible for mediating H3K36ac in yeast. Acid-extracted histones prepared from wild-type, *gcn5Δ* or *sas3Δ* strains were resolved on a 15% SDS-PAGE gel, followed by transfer to a PVDF membrane, and analyzed by immunoblot for the presence of H3K36ac. An antibody specific for the C-terminus of H3

(α-H3) was used as a loading control. Antibodies specific for H3K18ac and H3K36me3 were used as additional controls.

Figure 4.5

A.



B.

	30 * 45
Tetrahymena H3.2	...PATGGIKKPHRFRPGT...
Tetrahymena H3.3	...PVSGGVKKPHKFRPGT...
Yeast H3.3	...PSTGGVKKPHRYKPGT...
Mouse/Human H3.1	...PATGGVKKPHRYRPGT...
Mouse/Human H3.2	...PATGGVKKPHRYRPGT...
Mouse/Human H3.3	...PSTGGVKKPHRYRPGT...

Figure 4.5. Histone H3K36 acetylation is conserved in mammals. *(A)* Acid-extracted histones from *Tetrahymena*, yeast (*S. cerevisiae*), mouse (mouse embryonic fibroblasts) and human (HEK293) cells, along with recombinant H3 from *Xenopus*, were resolved on a 15% SDS-PAGE gel, transferred to a PVDF membrane and probed for H3K36ac (upper panel). Prior to immunoblot analysis, the membrane was Ponceau S stained to confirm equal loading of protein (lower panel). *(B)* Alignment of histone H3 protein sequences (amino acids 30-45) from different eukaryotic species. Divergent residues are highlighted in gray boxes. Asterisk indicates the position of lysine 36 in the H3 sequence.

Acknowledgements

This work has been published under the title "Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification" in the Journal of Biological Chemistry, vol. 282: pp 7632-7640, 2007. Permission to include this manuscript in my dissertation was granted by the American Society for Biochemistry and Molecular Biology. The authors, in addition to myself (lead author), were Bhargavi Rao, Benjamin Garcia, Sandra Hake, Robert Diaz, Jeffrey Shabanowitz, Donald Hunt, C. David Allis, Jason Lieb, and Brian Strahl. The data presented in Figure 4.1 was generated by Benjamin Garcia, data in Figure 4.3 was generated by Bhargavi Rao, and data in panel A of Figure 4.5 was generated by Sandra Hake. Recombinant mononucleosomes and HAT deletion yeast strains were kind gifts of Song Tan (Pennslyvania State University) and David Stillman (University of Utah), respectively. The H3K36ac antibody was produced with the help of Upstate Biotechnology.

References

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Berger, S. L. (2002). Histone modifications in transcriptional regulation. Curr Opin Genet Dev *12*, 142-148.

Bohlander, S. K., Espinosa, R., 3rd, Le Beau, M. M., Rowley, J. D., and Diaz, M. O. (1992). A method for the rapid sequence-independent amplification of microdissected chromosomal material. Genomics *13*, 1322-1324.

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*, 3286-3295.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev *14*, 155-164.

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*, 581-592.

Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C. (2000). Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. Embo J *19*, 3714-3726.

Clements, A., Poux, A. N., Lo, W. S., Pillus, L., Berger, S. L., and Marmorstein, R. (2003). Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. Mol Cell *12*, 461-473.

Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat Struct Mol Biol *11*, 1037-1043.

Davie, J. R., and Spencer, V. A. (1999). Control of histone modifications. J Cell Biochem *Suppl 32-33*, 141-148.

Dion, M. F., Altschuler, S. J., Wu, L. F., and Rando, O. J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. Proc Natl Acad Sci U S A *102*, 5501-5506.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Garcia, B. A., Hake, S. B., Diaz, R. L., Kauer, M., Morris, S. A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B. D., Allis, C. D., and Hunt, D. F. (2007). Organismal differences in post-translational modifications in histones H3 and H4. J Biol Chem *282*, 7641-7655.

Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F.*, et al.* (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes Dev *11*, 1640-1650.

Grant, P. A., Eberharter, A., John, S., Cook, R. G., Turner, B. M., and Workman, J. L. (1999). Expanded lysine acetylation specificity of Gcn5 in native complexes. J Biol Chem *274*, 5895-5900.

Gregory, P. D., Schmid, A., Zavari, M., Lui, L., Berger, S. L., and Horz, W. (1998). Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the PHO5 promoter in yeast. Mol Cell *1*, 495-505.

Grewal, S. I., and Rice, J. C. (2004). Regulation of heterochromatin by histone methylation and small RNAs. Curr Opin Cell Biol *16*, 230-238.

Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D. P., Allis, C. D., and Hunt, D. F. (2006). Expression patterns and post-translational modifications associated with mammalian histone H3 variants. J Biol Chem *281*, 559-568.

Holde, K. E. v. (1988). Chromatin (New York: Springer-Verlag).

Howe, L., Auston, D., Grant, P., John, S., Cook, R. G., Workman, J. L., and Pillus, L. (2001). Histone H3 specific acetyltransferases are essential for cell cycle progression. Genes Dev *15*, 3144-3154.

Iizuka, M., and Smith, M. M. (2003). Functional consequences of histone modifications. Curr Opin Genet Dev *13*, 154-160.

Iyer, V. R., Horak, C. E., Scafe, C. S., Botstein, D., Snyder, M., and Brown, P. O. (2001). Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. Nature *409*, 533-538.

Joshi, A. A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell *20*, 971-978.

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.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L.*, et al.* (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell *23*, 607-618.

Kizer, K. O., Phatnani, H. P., Shibata, Y., Hall, H., Greenleaf, A. L., and Strahl, B. D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol *25*, 3305-3316.

Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature.

Kouzarides, T. (2002). Histone methylation in transcriptional control. Curr Opin Genet Dev *12*, 198-209.

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C.*, et al.* (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature *383*, 269-272.

Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E., and Allis, C. D. (1998). Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. Genes Dev *12*, 627-639.

Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. Curr Opin Cell Biol *14*, 286-298.

Lee, D. Y., Teyssier, C., Strahl, B. D., and Stallcup, M. R. (2005). Role of protein methylation in regulation of transcription. Endocr Rev *26*, 147-170.

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.

Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature *436*, 294-298.

Medzihradszky, K. F., Zhang, X., Chalkley, R. J., Guan, S., McFarland, M. A., Chalmers, M. J., Marshall, A. G., Diaz, R. L., Allis, C. D., and Burlingame, A. L. (2004). Characterization of Tetrahymena histone H2B variants and posttranslational populations by electron capture dissociation (ECD) Fourier transform ion cyclotron mass spectrometry (FT-ICR MS). Mol Cell Proteomics *3*, 872-886.

Mizzen, C. A., Brownell, J. E., Cook, R. G., and Allis, C. D. (1999). Histone acetyltransferases: preparation of substrates and assay procedures. Methods Enzymol *304*, 675-696.

Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

Nathan, D., Ingvarsdottir, K., Sterner, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., Whelan, K. A., Krsmanovic, M., Lane, W. S., Meluh, P. B.*, et al.* (2006). Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes Dev *20*, 966-976.

Nicolas, E., Roumillac, C., and Trouche, D. (2003). Balance between acetylation and methylation of histone H3 lysine 9 on the E2F-responsive dihydrofolate reductase promoter. Mol Cell Biol *23*, 1614-1622.

Peterson, C. L., and Laniel, M. A. (2004). Histones and histone modifications. Curr Biol *14*, R546-551.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E.*, et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Poux, A. N., and Marmorstein, R. (2003). Molecular basis for Gcn5/PCAF histone acetyltransferase selectivity for histone and nonhistone substrates. Biochemistry *42*, 14366-14374.

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*, 9447-9459.

Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., Garcia, B. A., Shabanowitz, J., Burlingame, A. L., Hunt, D. F.*, et al.* (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*, 6988-6993.

Rojas, J. R., Trievel, R. C., Zhou, J., Mo, Y., Li, X., Berger, S. L., Allis, C. D., and Marmorstein, R. (1999). Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. Nature *401*, 93-98.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. Embo J *21*, 1121-1131.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shiio, Y., and Eisenman, R. N. (2003). Histone sumoylation is associated with transcriptional repression. Proc Natl Acad Sci U S A *100*, 13225-13230.

Shilatifard, A. (2004). Transcriptional elongation control by RNA polymerase II: a new frontier. Biochim Biophys Acta *1677*, 79-86.

Sims, R. J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes Dev *18*, 2437-2468.

Sterner, D. E., Grant, P. A., Roberts, S. M., Duggan, L. J., Belotserkovskaya, R., Pacella, L. A., Winston, F., Workman, J. L., and Berger, S. L. (1999). Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. Mol Cell Biol *19*, 86-98.

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

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*, 1298-1306.

Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. Proc Natl Acad Sci U S A *96*, 14967-14972.

Syka, J. E., Marto, J. A., Bai, D. L., Horning, S., Senko, M. W., Schwartz, J. C., Ueberheide, B., Garcia, B., Busby, S., Muratore, T.*, et al.* (2004). Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications. J Proteome Res *3*, 621-626.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

van Attikum, H., and Gasser, S. M. (2005). The histone code at DNA breaks: a guide to repair? Nat Rev Mol Cell Biol *6*, 757-765.

Wang, L., Mizzen, C., Ying, C., Candau, R., Barlev, N., Brownell, J., Allis, C. D., and Berger, S. L. (1997). Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. Mol Cell Biol *17*, 519-527.

Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A., and Allis, C. D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. Cell *97*, 99-109.

Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467-481.

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*, 654-663.

Ye, J., Ai, X., Eugeni, E. E., Zhang, L., Carpenter, L. R., Jelinek, M. A., Freitas, M. A., and Parthun, M. R. (2005). Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. Mol Cell *18*, 123-130.

Yu, Y., Eriksson, P., Bhoite, L. T., and Stillman, D. J. (2003). Regulation of TATA-binding protein binding by the SAGA complex and the Nhp6 high-mobility group protein. Mol Cell Biol *23*, 1910-1921.

Zhang, K., Tang, H., Huang, L., Blankenship, J. W., Jones, P. R., Xiang, F., Yau, P. M., and Burlingame, A. L. (2002). Identification of acetylation and methylation sites of histone H3 from chicken erythrocytes by high-accuracy matrix-assisted laser desorption ionization-time-of-flight, matrix-assisted laser desorption ionization-postsource decay, and nanoelectrospray ionization tandem mass spectrometry. Anal Biochem *306*, 259-269.

Zhang, L., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. Chromosoma *112*, 77-86.

Zhang, W., Bone, J. R., Edmondson, D. G., Turner, B. M., and Roth, S. Y. (1998). Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. Embo J *17*, 3155-3167.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev *15*, 2343-2360.

CHAPTER 5

DISCUSSION

Histone methylation is a highly conserved modification occurring in organisms ranging from yeast to humans. Although it has been implicated in the regulation of chromatin in processes such as DNA repair and transcription, the exact functions of this modification are just beginning to be discovered. In the preceding chapters, I have focused my research on the histone methyltransferase Set2 and the methylation of its target site, H3K36, in the process of transcription elongation. Through my studies, I have demonstrated that not only is Set2-mediated methylation of H3K36 conserved, but also contributed to the identification of a role for Set2 and H3K36 in the regulation of histone acetylation patterns in the coding region of genes. Furthermore, I have identified a novel and conserved modification on H3K36. Independent of methylation, H3K36 is acetylated at the promoters of genes by the transcriptional co-activator Gcn5. In this chapter, I will address some of the unanswered questions that remain following the completion of my work. Additionally, I will discuss the findings published subsequent to my own studies in an attempt to gain a better understanding of how Set2 and histone methylation are involved in the regulation of gene expression.

Conserved Role for Set2 and H3K36 Methylation in Transcription Elongation

While I have demonstrated in chapter 2 that Set2 methylation of H3K36 is conserved in organisms such as the fission yeast *S. pombe* and the multicellular eukaryote *N. crassa,* there is still the question of whether Set2 methylates H3K36 in higher eukaryotes and, if so, whether this methylation is mediated through an interaction with the transcribing polymerase. Initial studies identified NSD1, a mammalian member of the Set2 family, as an H3K36 methyltransferase *in vitro,* but whether it functions as such *in vivo* has yet to be determined (Rayasam et al., 2003). In my studies, we identified several potential Set2 homologs by sequence analyses. Significantly, a report found di- and trimethylation of H3K36 in the coding region of active metazoan genes supporting the hypothesis that one or more of the Set2 homologs identified may be functioning as H3K36-specific methyltransferases in higher eukaryotes (Bannister et al., 2005). In line with these findings, a report following my study identified another mammalian Set2 homolog, HYPB, as an active H3K36 methyltransferase that interacts with elongating RNAPII (Sun et al., 2005). Interestingly, the interaction between HYPB and RNAPII is mediated through a domain at its C-terminus similar to the RNAPII interaction domain identified in Set2 (SRI) (Li et al., 2005b; Sun et al., 2005). Of the mammalian Set2 homologs, HYPB is the only protein to contain the SRI domain implying that it may be the lone H3K36 methyltransferase in mammals coupled to RNAPII transcription. It is an attractive possibility that in higher eukaryotes, the other Set2 homologs may have unique biological functions. Yet, it remains to be determined what the roles of the other mammalian Set2 homologs are in the regulation of chromatin and gene expression.

In addition to identifying a conserved role for Set2 in H3K36 methylation, I have also identified in budding yeast a role for Set2 in recruiting the Rpd3S complex via the

chromodomain protein Eaf3 to the coding region of genes. This complex maintains nucleosomes in a hypoacetylated state and suppresses cryptic initiation of transcription. Thus, the obvious question becomes whether or not this role for Set2 and H3K36 methylation is conserved in higher eukaryotes. To date, there are very few studies that have investigated the conservation of this function. However, like Set2, Eaf3 is conserved from yeast to humans and studies of its homologs provide clues about the function of Eaf3 in other organisms. For example, in *S. pombe,* the Eaf3 homolog Alp13 is required for chromosome segregation and genomic integrity (Nakayama et al., 2003), while the closest mammalian homolog, MRG15, regulates embryonic development and cell proliferation (Tominaga et al., 2005). Importantly, studies of both Eaf3 homologs have identified them as members of histone deacetylase complexes (Gavin et al., 2002; Nakayama et al., 2003). Studies of MRG15, in particular, have shown that this protein is a stable component of both the hNuA4 acetyltransferase and Sin3/HDAC deacetylase complexes, which are both homologous to the yeast NuA4 and Rpd3 complexes, respectively (Doyon et al., 2004). In complex with Sin3, MRG15 has been shown to repress transcription (Yochum and Ayer, 2002). Interestingly, the chromodomain of MRG15 was recently shown to interact specifically with H3K36 di- and trimethylation (Zhang et al., 2006). Taken together, these data strongly support a conserved role for H3K36 methylation in the recruitment of a histone deacetylase complex in higher eukaryotes. Future investigations will provide more clues as to whether or not either Alp13 or MRG15 are involved in transcription elongation. However, a recent report, unexpectedly, offers an alternative view of H3K36 methylation and deacetylation in mammals, in which an H3K36 methyltransferase interacts directly with a repressive Sin3/HDAC deacetylase complex (Brown et al., 2006). Unlike members of the Set2 family,

this methyltransferase, Smyd2, contains a SET domain split into two parts by a MYND domain, a zinc-finger motif that mediates protein-protein interactions (Brown et al., 2006). These results open up the intriguing possibility that, in higher eukaryotes, proteins, potentially outside of Set2, may function in a similar manner with a repressive deacetylase complex independent of an interaction with elongating RNAPII.

Histone Acetylation and the Process of Transcription Elongation: Consequences of Misregulation

Although histone acetylation is well studied, the majority of research has focused on the link between this modification and transcriptional activity at the promoters of genes. As we gain a better understanding of transcription elongation, a few studies have highlighted a role for histone acetylation during this process. For example, studies have demonstrated the association of the Elongator HAT complex with transcribing RNAPII (Wittschieben et al., 1999) and a requirement for acetylation in the coding region of genes for transcriptional activity (Kristjuhan et al., 2002; Wang et al., 2002). With the finding that Set2 couples histone deacetylation and transcription elongation, the role of histone acetylation in transcription has been shown to be more complex in which a certain level must be maintained to allow passage of the polymerase without inappropriate initiation of transcription.

Exactly how histone acetylation is involved in the alteration of chromatin in the coding region of genes in not well known, but it is presumed to involve the disruption of histone-DNA and histone-histone contacts increasing the accessibility of DNA to polymerase. In fact, a recent study has shown that histone acetylation in the coding regions may serve as a binding platform for the recruitment of chromatin remodeling complexes which, in turn, alter

the positions of nucleosomes allowing the passage of RNAPII (Carey et al., 2006). There are also questions of which HATs are responsible for the acetylation in the coding region of genes. Genome-wide studies have shown that HATs such as Gcn5 and Esa1 are found predominantly at the promoters of genes suggesting that perhaps the Elongator complex is the only active acetyltransferase in the coding regions (Robert et al., 2004). However, a recent study has shown that Gcn5-mediated acetylation, as a member of the SAGA complex, in the coding region of the *GAL1* gene is involved in the eviction of nucleosomes (Govind et al., 2007). As to why Gcn5 has not been detected before in the coding region of genes is not known. It may have to do with the activity of the gene analyzed (*GAL1* is a highly transcribed gene) in combination with how Gcn5 is targeted to coding regions which may make detection difficult. Nonetheless, these findings add support to the role of histone acetylation in the process of transcription elongation. They also bring several questions to the forefront concerning the HATs responsible for acetylation in the coding region of genes and the potential mechanisms utilized to regulate the rearrangement of chromatin.

As much as we are interested in which HATs function in the coding region of genes, there is also a great deal of interest in the downstream consequences of unchecked acetylation in the absence of Set2, specifically the initiation of cryptic transcription. Unlike the FACT complex, the loss of Set2 or Eaf3 does not affect cell viability or growth suggesting that it may play a minor role in the suppression of cryptic transcription (Belotserkovskaya and Reinberg, 2004; Eisen et al., 2001; Strahl et al., 2002). A recent report would suggest that this lack of phenotype has something to do with the types and number of genes affected by the loss of Set2. Li *et al.* (2007b) have shown that although Set2 and the Rpd3 complex regulate coding region acetylation genome-wide, longer genes

with a low transcription frequency are the most dependent on the Set-Rpd3S pathway to prevent cryptic initiation (Li et al., 2007b). These investigations further propose that highly transcribed genes would have a high density of elongating RNAPII which would sterically hinder the binding of transcription factors required for the initiation of cryptic transcription. Thus, the effect of increased histone acetylation in the absence of Set2, cryptic initiation, is not genome-wide which may explain why Set2 does not have any obvious effects on cell viability.

Alternatively, there may be more than one mechanism of controlling cryptic transcription. The main concern being that the cryptic transcripts produced could be translated into aberrant proteins. As has been shown previously, many of these transcripts are polyadenylated, a requirement for protein translation (Carrozza et al., 2005). Yet, it seems that in addition to the repression of transcription from cryptic promoters, yeast cells have evolved a backup, quality control mechanism involving post-transcriptional degradation of these transcripts (Wyers et al., 2005). Interestingly, it seems that even under wild-type conditions cryptic transcription is widely occurring in the yeast genome (Wyers et al., 2005). Indeed, this occurrence seems to be conserved as many cryptic transcripts from mammalian chromosomes were found using microarray tiling experiments (Johnson et al., 2005). It remains to be seen if any of these cryptic transcripts escape degradation and serve some functional purpose. Nevertheless, mutations of proteins such as Set2 leading to abnormal levels of cryptic transcription will most likely have a more pronounced effect in higher eukaryotes which have longer genes. It will be interesting to see what future investigations of Set2 and cryptic transcription in higher eukaryotes will reveal.

Histone Methylation Binding Motifs

While a portion of my studies have focused on the recruitment of the chromodomain protein Eaf3, several novel domains within chromatin-regulating complexes have emerged as binders of methylated histones (Kim et al., 2006; Shi et al., 2007). For example, the tudor domain of the 53BP1 protein binds to methylation of H3K79 at double-strand DNA breaks (Huyen et al., 2004) while, the WD40-repeat domain of the WDR5 protein in hSet1 complexes associates with H3K4 methylation and is essential for H3K4 trimethylation in vertebrae development (Wysocka et al., 2005). Interestingly, in the past year, the PHD domain has surfaced as a major motif that recognizes and binds methylated histone residues. Initially identified in two plant homeodomain proteins, the PHD domain has been found in many transcriptional regulators (Aasland et al., 1995; Bienz, 2006). Specifically, members of the *S. cerevisiae* and human ING (for inhibitor of growth) family of tumour suppressor proteins which contain PHD domains have been shown to bind to H3K4 di- and trimethylation (Shi et al., 2006). In the case of ING2, a member of the human mSin3A/HDAC1 histone deacetylase complex, the PHD domain of this protein stabilizes the deacetylase complex at the promoters of genes linking H3K4me3 to gene repression (Shi et al., 2006). Additionally, PHD proteins have been implicated in the recruitment of chromatin remodeling activity through an association with methylated H3K4 (Wysocka et al., 2006). Thus, histone methylation serves as a modification capable of recruiting diverse chromatin activities including histone-modifying complexes to genes.

As it turns out, Rco1, the only unique member of Rpd3S complex, is a PHD domain protein. At the conclusion of my studies related to the Rpd3S complex, the only thing known about this protein was that it and Eaf3 require each other for stable association with the

Rpd3S complex suggesting that it may play some role in stabilizing the complex (Carrozza et al., 2005). With the finding that PHD domain proteins can interact with methylated histones, a recent report has shown that Rco1 plays a much bigger role within the Rpd3S complex. Specifically, Li *et al.* (2007a) have shown that it is the combination of the Eaf3 chromodomain and the PHD domain of Rco1 that determines the specificity and affinity of the Rpd3S complex for nucleosomes in the coding region of genes (Li et al., 2007a). Significantly, this report indicates that the NuA4 acetyltransferase complex, of which Eaf3 is also a member, does not recognize H3K36me3 presumably because it lacks Rco1. Additional proteome-wide analyses have identified other PHD domains capable of interacting with H3K36 methylation such as the PHD domain in the Nto1 protein, a member of the NuA3 acetyltransferase complex (Shi et al., 2007). These findings demonstrate that methylation of H3K36 may recruit multiple effector complexes.

Regulation of the Distribution of H3K36 Methylation and Acetylation on Genes

Until recently, histone methylation was considered a stable, irreversible modification, but, with the discovery of histone demethylases, has been revealed to be dynamically regulated (Klose et al., 2006; Shi et al., 2004; Tsukada et al., 2006; Whetstine et al., 2006). It has only been this past year that histone demethylases specific for H3K36 methylation in budding yeast have been identified (Fang et al., 2007; Klose et al., 2007; Tu et al., 2007). Specifically, the histone demethylase Rph1 targets H3K36 di- and trimethylation while the histone demethylase Jhd1 targets H3K36 mono- and dimethylation indicating that all three methyl states can be reversed on H3K36 (Fang et al., 2007; Klose et al., 2007; Tu et al., 2007). With the finding that H3K36 can be acetylated or methylated in the promoter and

coding region of genes, respectively, the identification of histone demethylases may be vital to the distribution of these modifications on genes.

As is, studies have previously shown a link between Set2 and the inhibition of transcription initiation at the *GAL1* gene (Biswas et al., 2006). Moreover, transient trimethylation of H3K36 at the promoter of the *MET16* gene coincides with the recruitment of the NuA4 histone acetyltransferase complex and, with the onset of transcription, the distribution of H3K36 di- and trimethylation changes with methylation accumulating in the coding region of this gene (Morillon et al., 2005). These results suggest methylation of H3K36 at promoters may be gene-specific as H3K36 methylation has been found genome-wide to occur primarily in the coding regions (Pokholok et al., 2005; Rao et al., 2005). Yet, it is possible that there may be some transient H3K36 methylation at promoters that is regulated by histone demethylases. It remains to be determined what effect histone demethylation has on the acetylation of H3K36. It would be interesting to see what happens to the distribution of this modification in the absence of functional histone demethylases.

Additionally, there may be other mechanisms involved in the prevention of H3K36 methylation in the promoters of genes such as the requirement for an interaction between Set2 and elongating RNAPII for H3K36 methyltransferase activity (Krogan et al., 2003; Li et al., 2003; Schaft et al., 2003; Xiao et al., 2003). Another mechanism may be the incorporation of the H2A variant Htz1 into the promoter nucleosomes of inactive genes which inhibits Set2-mediated methylation (Li et al., 2005a). In turn, methylation of H3K36 may maintain a boundary between the promoter and coding region of genes by preventing the spreading of H3K36 acetylation, similar to the RNAPII CTD kinase Ctk1 which prevents the spreading of H3K4 methylation into the coding region of genes from the promoter (Xiao et

al., 2007). Questions still remain about the role of H3K36 acetylation in transcription. Data presented in chapter 4 support a role for this modification in the activation of transcription. It remains to be seen if SAGA and Set2 compete for H3K36 and how disruption of either complex affects the modification of H3K36 and the process of transcription.

Concluding Remarks

Transcription has proven to be a well-orchestrated process with several key and minor players each playing a role in the ordered production of functional RNA. With each new study, the modification of histone proteins, especially histone methylation, has been shown to be an important mechanism of chromatin regulation in the process of transcription. In my study of the histone methyltransferase Set2, I have found that the methylation of H3K36 is not simply a modification associated with transcription elongation. The modification of H3K36, not unlike the process of transcription itself, is carefully regulated leading to distinct biological effects involved in the regulation of transcription (Fig. 5.1). Although questions still remain regarding the acetylation of H3K36 and its role in transcription, my work has contributed to the identification of a role for Set2-mediated H3K36 methylation in transcription elongation. Through recruitment of the Rpd3S deacetylase complex, Set2 and H3K36 methylation serve as one mechanism of suppressing cryptic initiation in the coding region of genes. Interestingly, there are multiple mechanisms in place to maintain appropriate transcription elongation highlighting the interplay between histone modifiers, histone chaperones, and chromatin remodelers. However, future investigations will be required to determine the exact role of H3K36 acetylation, and several other histone modifications, in the process of transcription. With the identification of new binding motifs,

the possible roles for histone modifications in the regulation of chromatin continue to expand.

Figure 5.1

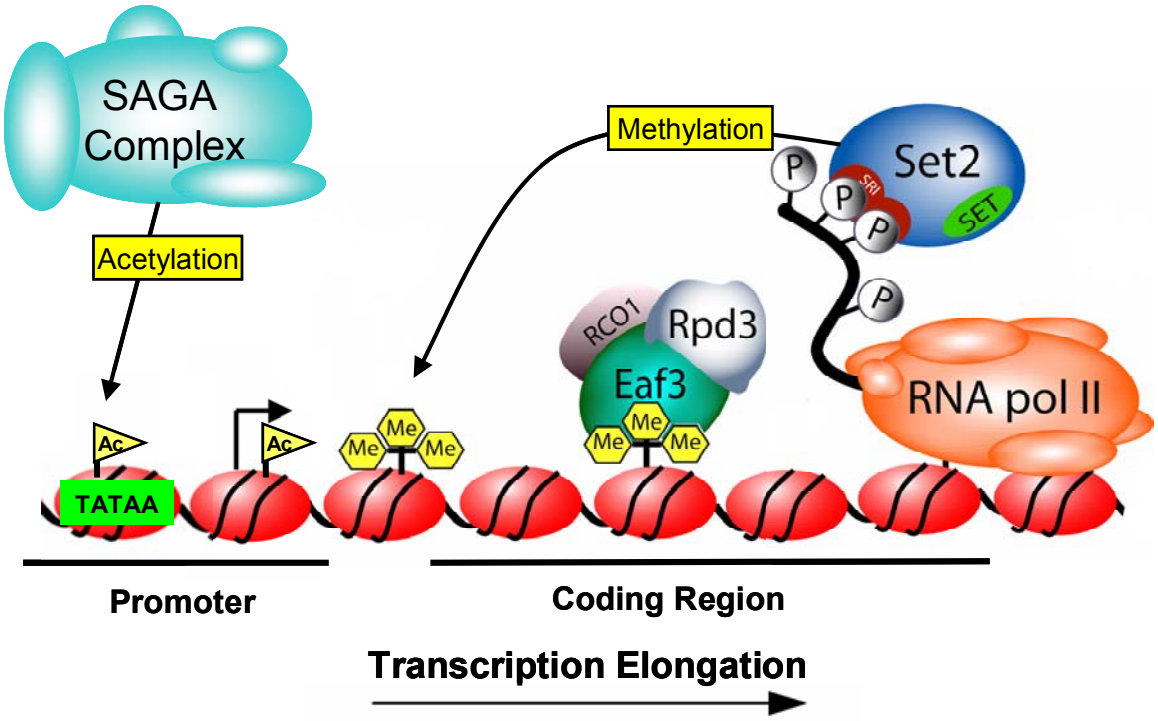
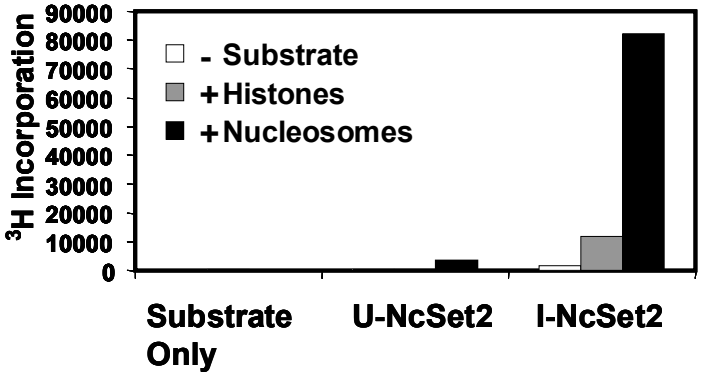

SAGA
Complex
Acetylation
Methylation
Set2
SRI
SET
P
P
P
P
RCO1
Rpd3
Eaf3
Me
Me
Me
Me
Me
Ac
Ac
TATAA
RNA pol II
Promoter
Coding Region
Transcription Elongation

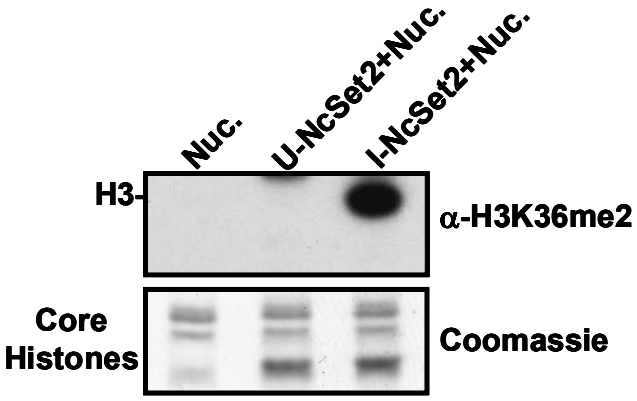
Figure 5.1. Model of H3K36 modifications and transcription elongation. At the promoters of genes, the Gcn5-containing SAGA complex targets H3K36 for acetylation. Interaction between Serine 2 phosphorylated C-terminal domain (CTD) of RNA polymerase II (RNA Pol II) and Set2 targets the methyltransferase activity of Set2 to the coding region of genes during transcription elongation. Set2-mediated H3K36 di- or trimethylation (trimethylation is pictured here) then serve as recognition marks for the recruitment of the Rpd3S deacetylase complex which, in turn, removes acetyl marks behind the transcribing polymerase maintaining the coding region of genes in a hypoacetylated state.

Appendix: Set2 from *Neurospora crassa* is a robust nucleosomal-selective H3K36-specific methyltransferase

A.



B.



Appendix. Set2 from *Neurospora crassa* is a robust nucleosomal-selective H3K36-specific methyltransferase. *(A)* Lysates containing recombinant NcSet2 (amino acids 1-372) from uninduced (U-NcSet2) or IPTG-induced (I-NcSet2) bacterial expression strains were incubated with chicken core histones or oligonucleosomes and ^{3}H-labeled S-adenosyl-methionine (^{3}H-SAM). ^{3}H incorporation was analyzed by the filter-binding assay and monitored by scintillation counting. *(B)* Oligonucleosomes (Nuc.) were incubated with NcSet2 and cold SAM in an HMT assay followed by immunoblotting with the α-H3K36me2 antibody (upper panel). Parallel reactions were performed and examined by Coomassie staining to monitor loading (lower panel).

The above figure has been reprinted, with modification, from a collaborative paper from Eukaryotic Cell, vol. 4, "Methylation of histone H3 lysine 36 is required for normal development in *Neurospora crassa*", pp 1455-1464 with permission from the American Society for Microbiology. The authors of this work included Keyur Adhvaryu, Brian Strahl, Eric Selker and myself.

References

Aasland, R., Gibson, T. J., and Stewart, A. F. (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends Biochem Sci *20*, 56-59.

Bannister, A. J., Schneider, R., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2005). Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem *280*, 17732-17736.

Belotserkovskaya, R., and Reinberg, D. (2004). Facts about FACT and transcript elongation through chromatin. Curr Opin Genet Dev *14*, 139-146.

Bienz, M. (2006). The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci *31*, 35-40.

Biswas, D., Dutta-Biswas, R., Mitra, D., Shibata, Y., Strahl, B. D., Formosa, T., and Stillman, D. J. (2006). Opposing roles for Set2 and yFACT in regulating TBP binding at promoters. Embo J *25*, 4479-4489.

Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D., and Tucker, P. W. (2006). Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. Mol Cancer *5*, 26.

Carey, M., Li, B., and Workman, J. L. (2006). RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation. Mol Cell *24*, 481-487.

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*, 581-592.

Doyon, Y., Selleck, W., Lane, W. S., Tan, S., and Cote, J. (2004). Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. Mol Cell Biol *24*, 1884-1896.

Eisen, A., Utley, R. T., Nourani, A., Allard, S., Schmidt, P., Lane, W. S., Lucchesi, J. C., and Cote, J. (2001). The yeast NuA4 and Drosophila MSL complexes contain homologous subunits important for transcription regulation. J Biol Chem *276*, 3484-3491.

Fang, J., Hogan, G. J., Liang, G., Lieb, J. D., and Zhang, Y. (2007). The S. cerevisiae histone demethylase Jhd1 fine-tunes the distribution of H3K36me2. Mol Cell Biol.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M.*, et al.* (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141-147.

Govind, C. K., Zhang, F., Qiu, H., Hofmeyer, K., and Hinnebusch, A. G. (2007). Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. Mol Cell *25*, 31-42.

Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., 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*, 406-411.

Johnson, J. M., Edwards, S., Shoemaker, D., and Schadt, E. E. (2005). Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. Trends Genet *21*, 93-102.

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.

Klose, R. J., Gardner, K. E., Liang, G., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2007). Demethylation of histone H3K36 and H3K9 by Rph1: a vestige of an H3K9 methylation system in Saccharomyces cerevisiae? Mol Cell Biol.

Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature.

Kristjuhan, A., Walker, J., Suka, N., Grunstein, M., Roberts, D., Cairns, B. R., and Svejstrup, J. Q. (2002). Transcriptional inhibition of genes with severe histone h3 hypoacetylation in the coding region. Mol Cell *10*, 925-933.

Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C.*, et al.* (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol *23*, 4207-4218.

Li, B., Gogol, M., Carey, M., Lee, D., Seidel, C., and Workman, J. (2007a). Combined action of PHD and Chromo domains direct the Rpd3S HDAC to transcribed chromatin. Science *In press*.

Li, B., Gogol, M., Carey, M., Pattenden, S., Seidel, C., and Workman, J. (2007b). Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. Genes Dev *In press*.

Li, B., Howe, L., Anderson, S., Yates, J. R., 3rd, and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *278*, 8897-8903.

Li, B., Pattenden, S. G., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J., and Workman, J. L. (2005a). Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. Proc Natl Acad Sci U S A *102*, 18385-18390.

Li, M., Phatnani, H. P., Guan, Z., Sage, H., Greenleaf, A. L., and Zhou, P. (2005b). Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. Proc Natl Acad Sci U S A *102*, 17636-17641.

Morillon, A., Karabetsou, N., Nair, A., and Mellor, J. (2005). Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. Mol Cell *18*, 723-734.

Nakayama, J., Xiao, G., Noma, K., Malikzay, A., Bjerling, P., Ekwall, K., Kobayashi, R., and Grewal, S. I. (2003). Alp13, an MRG family protein, is a component of fission yeast Clr6 histone deacetylase required for genomic integrity. Embo J *22*, 2776-2787.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E.*, et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

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*, 9447-9459.

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. Embo J *22*, 3153-3163.

Robert, F., Pokholok, D. K., Hannett, N. M., Rinaldi, N. J., Chandy, M., Rolfe, A., Workman, J. L., Gifford, D. K., and Young, R. A. (2004). Global position and recruitment of HATs and HDACs in the yeast genome. Mol Cell *16*, 199-209.

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*, 2475-2482.

Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M. R.*, et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

Shi, X., Kachirskaia, I., Walter, K. L., Kuo, J. H., Lake, A., Davrazou, F., Chan, S. M., Martin, D. G., Fingerman, I. M., Briggs, S. D.*, et al.* (2007). Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. J Biol Chem *282*, 2450-2455.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

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*, 1298-1306.

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. J Biol Chem *280*, 35261-35271.

Tominaga, K., Kirtane, B., Jackson, J. G., Ikeno, Y., Ikeda, T., Hawks, C., Smith, J. R., Matzuk, M. M., and Pereira-Smith, O. M. (2005). MRG15 regulates embryonic development and cell proliferation. Mol Cell Biol *25*, 2924-2937.

Tompa, R., and Madhani, H. D. (2007). Histone H3 Lysine 36 Methylation Antagonizes Silencing in Saccharomyces cerevisiae Independently of the Rpd3S Histone Deacetylase Complex. Genetics *175*, 585-593.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Tu, S., Bulloch, E. M., Yang, L., Ren, C., Huang, W. C., Hsu, P. H., Chen, C. H., Liao, C. L., Yu, H. M., Lo, W. S.*, et al.* (2007). Identification of histone demethylases in Sacchromyces cerevisiae. J Biol Chem.

Wang, A., Kurdistani, S. K., and Grunstein, M. (2002). Requirement of Hos2 histone deacetylase for gene activity in yeast. Science *298*, 1412-1414.

Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467-481.

Wittschieben, B. O., Otero, G., de Bizemont, T., Fellows, J., Erdjument-Bromage, H., Ohba, R., Li, Y., Allis, C. D., Tempst, P., and Svejstrup, J. Q. (1999). A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. Mol Cell *4*, 123-128.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J., Regnault, B., Devaux, F., Namane, A., Seraphin, B.*, et al.* (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725-737.

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*, 859-872.

Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P.*, et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature.

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*, 654-663.

Xiao, T., Shibata, Y., Rao, B., Laribee, R. N., O'Rourke, R., Buck, M. J., Greenblatt, J. F., Krogan, N. J., Lieb, J. D., and Strahl, B. D. (2007). The RNA polymerase II kinase Ctk1 regulates positioning of a 5' histone methylation boundary along genes. Mol Cell Biol *27*, 721-731.

Yochum, G. S., and Ayer, D. E. (2002). Role for the mortality factors MORF4, MRGX, and MRG15 in transcriptional repression via associations with Pf1, mSin3A, and Transducin-Like Enhancer of Split. Mol Cell Biol *22*, 7868-7876.

Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q., and Ding, J. (2006). Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. Nucleic Acids Res *34*, 6621-6628.