EXPANDING THE FUNCTION OF HISTONE H3 LYSINE 36 METHYLATION IN *SACCHAROMYCES CEREVISIAE*

Stephen Lee McDaniel

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology.

Chapel Hill
2016

Approved by:

Brian Strahl

Scott Bultman

Jean Cook

Bob Duronio

Bill Marzluff

ABSTRACT

Stephen Lee McDaniel: Expanding the function of histone H3 lysine 36 methylation in *Saccharomyces cerevisiae*
(Under the direction of Brian Strahl)

Eukaryotic DNA is wrapped around an octamer of histone proteins, two each of H2A, H2B, H3, and H4, to form chromatin. The cell must negotiate the chromatin landscape to facilitate all DNA templated processes, including replication, repair, and transcription. The histone proteins themselves can be heavily modified by small chemical moieties like methyl, phospho, or ubiquitin groups, called post-translational modifications (PTMs). PTMs can both alter the electrostatic properties of chromatin, promoting an opening or closing of specific chromatin regions and/or as specific docking sites for effector proteins. The spatial-temporal localization of histone PTMs is highly regulated and when disrupted can lead to a variety of diseases. Methylation of histone H3 at lysine 36 (H3K36me) is a very well conserved and highly regulated histone modification laid down by the histone methyltransferase Set2 in the budding yeast *S. cerevisiae*. H3K36me occurs co-transcriptionally, thus marking actively transcribed genes. Though, H3K36me is associated with active transcription, it actually functions as a repressive mark, recruiting the Isw1b chromatin remodeling complex and the Rpd3S histone deacetylase complex (HDAC) to chromatin following the elongating RNA polymerase II (RNAPII) complex, preventing the binding of RNAPII to cryptic promoters within gene bodies. Here, three new aspects of H3K36me are elucidated. First, a new H3K36me binding protein is characterized, Pdp3. Pdp3 binds to H3K36me and is a

member of the newly described NuA3b histone acetyltrasferase (HAT) complex. The binding of Pdp3 to H3K36me is necessary for the function of NuA3b for in the absence of Pdp3 or H3K36me, NuA3 target genes are down regulated and several other transcriptional defects are observed. Second, the role of the second plant homeodomain (PHD) finger in Rco1 of the Rpd3S complex is elucidated. Like the first PHD finger, it is necessary for Rpd3S localization to chromatin in addition to preventing aberrant transcription from occurring within gene bodies. Finally, a novel role for Set2 and H3K36me in the nutrient stress response is uncovered. Surprisingly, it is found that Set2 genetically interacts with several pathways critical for nutrient response signaling such as the Tor1, Tor2, and Slt2 mitogen-activated protein (MAP kinase) pathways. Without Set2 present, the kinetics and overall levels of signaling in these pathways is altered. Together, the work in this dissertation expands our understanding of the role H3K36me plays in transcription and cellular signaling and hopefully will guide future work in higher eukaryotes to better understand and treat human diseases.

ACKNOWLEDGEMENTS

I would like to begin by thanking my advisor, Dr. Brian Strahl who gave me the freedom to pursue my scientific interests, while at the same time providing me the support to grow as a scientist. I would also like to thank the members of my committee, whose insightful comments always made my science better. I also could not have accomplished this without the help of the members of the Strahl lab, both past and present. Finally, I could not have done this work without the unending support of my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

5-FOA – 5-Fluoroorotic Acid

6-AU – 6-azauracil

ac – acetyl

AI – autoinhibitory domain in Rco1

ATP – adenosine triphosphate

CC – coiled coiled domain

ccRCC – clear cell renal cell carcinoma

CTD – C-terminal domain of RNA polymerase II

CUT – cryptic unstable transcript

DNA – deoxyribonucleic acid

DNMT – DNA methyltransferase

FACT – facilitates chromatin transcription complex

HAT – histone acetyltransferase

HDAC – histone deacetylase complex

H3K4 – histone H3 lysine 4

H3K14 – histone H3 lysine 14

H3K27 – histone H3 lysine 27

H3K36 – histone H3 lysine 36

H3K44 – histone H3 lysine 44

H3K56 – histone H3 lysine 56

H3K79 – histone H3 lysine 79

HMT – histone methyltransferase

HR – homologous recombination

MAPK – mitogen-activated protein kinase

me - methylation

me1 – mono-methylation

me2 – di-methylation

me3 – tri-methylation

MS/MS – tandem mass spectrometry

NHEJ – non-homologous end joining

OD – optical density

ORF – open reading frame

PHD – plant homeodomain

PKC – protein kinase C

PTM – post-translational modification

PWWP – proline-tryptophan-tryptophan-proline domain

RNA – ribonucleic acid

RNAPII – RNA polymerase 2 complex

RT-qPCR – real-time quantitative polymerase chain reaction

SC – synthetic complete media

SD – synthetic depleted media

SID – Sin3 interaction domain in Rco1

SRI – Set2 Rpb1 interaction domain in Set2

SUT – stable unannotated transcript

TAP – tandem affinity purification tag

TCA – trichloroacedic acid

TOR – target of rapamycin

TORC1 – target of rapamycin 1 complex

WCE – whole cell extract

XUT – Xrn1-sensitive unstable transcripts

YPD – yeast extract, peptone, dextrose media

CHAPTER 1 –INTRODUCTION

Animals face two fundamental problems during development: proper packaging of their DNA and proper gene expression. First, they must compact all of their DNA into a very small nucleus (humans have about two meters of DNA which must fit into a ten micron nucleus), while at the same time ensuring that the information contained within the genome can be accessed when needed. Second, they need to ensure that genes are expressed when needed and are repressed when they are not necessary. Throughout development, gene expression programs change as cells move through the cell cycle, respond to stress, or differentiate into different cell types. Gene expression is tightly controlled throughout development to prevent the aberrant expression of oncogenes.

To accomplish these feats, eukaryotic cells package their genomes as chromatin, a complex of DNA and proteins. 147 base pairs of DNA is wrapped around eight histone proteins: two each of H2A, H2B, H3, and H4 (Luger et al., 1997). This DNA histone complex forms the fundamental unit of chromatin—the nucleosome. The positively charged histone proteins have a high affinity for the negatively charged DNA residues, and thus bind DNA readily (Dutnall and Ramakrishnan, 1997), and also interact with other histone proteins. Together with a plethora of specialized proteins that assemble, disassemble, and move nucleosomes along the DNA, the entire genome is packaged into chromatin.

All chromatin is not created equal, however. Broadly, chromatin can be classified into two distinct categories: euchromatin and heterochromatin. Euchromatin is gene rich and is characterized by high levels of gene transcription. It is localized to the middle of the chromosome arms as well as the nuclear interior (Pueschel et al., 2016). Heterochromatin, on the other hand, is gene poor and sees low levels of transcription. It is found at the telomeres and centromeres of chromosomes as well as the exterior of the nuclear space.

While a great deal is known about how chromatin influences DNA templated processes such as transcription, repair, and replication, many questions regarding its function remain. One of the key questions in regards to transcription is how are large numbers of genes turned on and off throughout development? Further, how are precise transcriptional programs quickly and efficiently carried out in the cell? Finally, how are gene expression patterns replicated from parent to daughter cell (and even further across organismal generations)? While certainly not the sole determinant of these functions, post-translational modifications (PTMs) to the histone proteins themselves play a foundational role in the regulation of all DNA templated processes. The work contained in this dissertation focuses specifically on the role of PTMs in transcription.

The Histone Code Hypothesis

Histone PTMs were discovered in the mid-1960s (Allfrey et al., 1964). Then the modifications of histones were limited to lysine and arginine methylation and lysine acetylation. Now, with the tremendous advances in mass spectrometry technology, the diversity of modified residues and the modifications themselves continues to expand

(Zhao and Garcia, 2015). For many years, it was thought that histone modifications played some role in cellular processes, but it was not until 2000 that these ideas were formulated into The Histone Code Hypothesis (Strahl and Allis, 2000). This hypothesis suggests that the differential modification of histone proteins acts as a sort of code that can be "read" by specialized proteins in the nucleus to carry out specific downstream biological functions such as activation or repression of transcription, DNA repair, or DNA replication. These reader or effector proteins have conserved binding domains that are able to recognize specific modified residues and either further modify histones, remodel them, or recruit other proteins or complexes to chromatin (Gardner et al., 2011).

While there many different kinds of histone PTMs, this work will focus on lysine methylation and to a lesser extent on lysine acetylation. Unmodified lysine residues are naturally positively charged. When lysine residues are acetylated, however, that positive charge is neutralized. The same is not true for lysine methylation. Lysine residues can be mono-, di-, or tri-methylated and retain their positive charge regardless of the number of methyl groups added. Generally, acetylation is associated with active transcription due to the decreased electrostatic interactions between the neutralized charge of histones with DNA and decreased inter-nucleosome interactions, while lysine methylation is often associated with repressed areas of transcription, by maintaining the electrostatic interactions between the positively charged methylated lysine residues and the negatively charged DNA bases.

In the budding yeast *Saccharomyces cerevisiae*, there are three known lysine residues that are methylated: histone H3 lysine 4 methylation (H3K4me), histone H3 lysine 36 methylation (H3K36me), and histone H3 lysine 79 methylation (H3K79me). All

three residues and their methylation are conserved from yeast to humans, making budding yeast an attractive model system to study each of these marks. While all three marks are known to function in transcriptional regulation in both yeast and human cells, the precise molecular underpinnings of their role in this and other biological processes remains poorly understood.

One of the key findings in recent years is that each histone PTM has distinct loci of the genome in which it inhabits (Ho et al., 2014; Weiner et al., 2015). In *Saccharomyces cerevisiae*, H3K4me is found primarily at promoters (at enhancers in metazoans), while H3K36me is localized more to the 3' regions of gene bodies. H3K79me is generally uniform across actively transcribed regions of the genome. Because these marks are differentially localized, they also recruit their effector proteins to distinct loci in the genome as well (Yun et al., 2011). Proteins that bind to H3K4me will localize more at the promoters of genes and be more important for transcriptional initiation, while proteins that bind to H3K36me will bind at the 3' ends of genes and be associated with transcriptional elongation. Thus, in addition to providing a diverse set of platforms for effector proteins to bind, histone PTMs also regulate where in the genome these effector proteins can act.

Set2 and H3K36 Methylation

In *Saccharomyces cerevisiae*, there is a single histone methyltransferase (HMT) that acts on H3K36, Set2 (Strahl et al., 2002). Set2 is responsible for all H3K36 methyltransferase activity in yeast, but its homologs in higher eukaryotes are not (Figure 1.1). For example, in *Drosophila melanogaster,* dMes-4 is responsible for H3K36

mono- and di-methyltation (H3K36me1 and H3K36me2), whereas dSet2 is primarily responsible for H3K36 tri-methylation (H3K36me3) (Bell et al., 2007). Because Set2 is responsible for all H3K36me in yeast, and strains where *SET2* has been deleted are viable, yeast is an ideal model system for studying this histone PTM.

The Domain Structure of Set2

The SET Domain

Set2 has several well conserved domains, the first being the SET domain itself, which consists of the AWS-SET-PS motif. This is the catalytic domain of Set2 and carries out the H3K36me transfer. In budding yeast, this domain is at the extreme N-terminus of the protein, which differs in higher eukaryotes. SETD2, the human homolog of Set2, has an extended N-terminus with no conserved domains and no known function. This extended N-terminus is common in non-yeast eukaryotes and could possibly restrict the enzyme's ability to methylate H3K36.

The H4 Interaction Domain

Preceeding the budding yeast Set2 domain there is a small histone H4 interaction domain (Du et al., 2008). This domain has been shown to physically bind to H4 and the corresponding residues in histone H4 (H4K44) are necessary for proper H3K36me *in vivo*. It is currently unknown if the interaction of Set2 with H4 is conserved in higher eukaryotes, though due to the increasing size of Set2 seen throughout evolution, it is likely that the Set2 homologs make multiple histone interactions.

The WW and CC Domains

The SET domain is followed by two protein-protein interaction domains, a WW and CC domain. The WW domain, named so for containing two conserved tryptophan residues, binds to proline rich proteins (Sudol et al., 1995). In the nucleus, proteins with this domain are known to regulate transcription by binding to the C-terminal domain (CTD) of RNA polymerase II (RNAPII) (Sudol et al., 2001). While this domain is conserved in the homologs of Set2, it is currently unknown if it is functional and, if so, what its binding partners are. This is also true of the coiled-coiled (CC) domain which follows the WW domain. This is also a well conserved protein-protein interaction motif, which in certain proteins promotes homo-dimerization (Poon and Mekhail, 2011). Intriguingly, although Set2 originally was purified from yeast and showed a relative mass that is equivalent to a dimer of Set2, co-expression studies employing alternately tagged versions of Set2 that lack the ability to bind to RNAPII (see below) showed no evidence of dimerization (unpublished results).

The SRI Domain

Finally, Set2 contains a Set2-Rpb1 interacting (SRI) domain. This domain binds to the hyper-phosphorylated CTD of RNAPII, specifically the Ser2/Ser5 phosphorylated form (Kizer et al., 2005; Schaft et al., 2003) and is conserved in humans (Hacker et al., 2016; Li et al., 2005). Specifically, it appears that the SRI domain recognizes these phosphorylated residues across two heptapeptide repeats of the CTD (Vojnic et al., 2006).

Importantly, the loss of the SRI domain does not abolish chromatin localization of the Set2 truncation (Youdell et al., 2008). Interestingly, while they do not localize as efficiently to chromatin, ectopic expression of Set2 truncation constructs lacking the SRI domain are able to catalyze H3K36me1 and limited H3K36me2, but are unable to catalyze H3K36me3 (Kizer et al., 2005). This suggests that the Set2/RNAPII interaction is critical in stabilizing Set2 on chromatin long enough to carry out the H3K36me1 to me2 and especially, the me2 to me3 reactions.

The Regulation of Set2 and H3K36me

Serine 2 and Serine 5 Phosphorylation on the CTD of RNAPII

CTD phosphorylation plays a critical role in the regulation of H3K36me and Set2 protein levels. If Set2 is unable to bind to the phosphorylated CTD of RNAPII, H3K36me3 is lost (Kizer et al., 2005). Further, if the CTD is mutated such that the number of Ser5 or Ser2 residues are critically reduced, H3K36me3 is lost (Fuchs et al., 2012; Xiao et al., 2003). Interestingly, Set2 protein levels are also reduced in these mutants, severely so when Ser5 is mutated to alanine (Fuchs et al., 2012). There is also evidence that demonstrates that mutations of *BUR1* a RNAPII CTD Ser2 kinase (Qiu et al., 2009), also disrupt H3K36me3 (Chu et al., 2006). This is also true of the other Ser2 kinase, Ctk1 (Dronamraju and Strahl, 2014). Ctk1 mutants see extreme loss of Set2 protein stability as well as a complete abolishment of H3K36me3 (Dronamraju and Strahl, 2014; Fuchs et al., 2012).

SPT6

Yet another regulator of Set2 and H3K36me is Spt6. Spt6 is a histone chaperone (Bortvin and Winston, 1996) that also binds to the phosphorylated CTD of RNAPII to help facilitate transcription elongation (Diebold et al., 2010; Sun et al., 2010). Mutations in multiple regions of Spt6 result in a loss of H3K36me2 and/or H3K36me3 levels (Dronamraju and Strahl, 2014; Youdell et al., 2008). Like Ctk1, mutations in Spt6 also reduce Set2 protein stability (Dronamraju and Strahl, 2014; Youdell et al., 2008). Interestingly, Ctk1 and Spt6 also mutually regulate each other's protein stability (Dronamraju and Strahl, 2014), forming a feed-forward loop that maintains CTD Ser2 phosphorylation, and consequently, Set2 levels.

The Paf1 Complex

The Paf1 complex, which associates with RNAPII, coordinates multiple histone PTMs, and facilitates transcription, functions upstream of Spt6 and Ctk1. Specific deletions of members of the Paf1 complex (*PAF1* and *CTR9*) lead to reduced levels of Spt6 and Ctk1 (Dronamraju and Strahl, 2014). Spt6 is likely stabilized via direct interaction with Paf1 (Dronamraju and Strahl, 2014), which allows for the co-stabilization of Ctk1. Because a loss of Paf1 complex function leads to a loss in Spt6 and Ctk1 levels, this also affects Set2, which functions further downstream. Deletions of *PAF1* and *CTR9* also result in lower Set2 levels and a loss of H3K36me3 (Dronamraju and Strahl, 2014; Fuchs et al., 2012). Together, these data suggest a model where the Paf1 complex helps to recruit Spt6 and Ctk1 to chromatin. Once recruited to chromatin, Ctk1

can then phosphorylate Ser2 of the CTD of RNAPII. Set2 is then recruited to chromatin via the hyperphosphorylated CTD and is capable of tri-methylating H3K36 (Figure 1.2).

Proteins that Interact with H3K36me

ASF1

Asf1 is a histone chaperone for histones H3 and H4 (Adkins and Tyler, 2004; Schwabish and Struhl, 2006). Asf1 is necessary to deposit H3K56ac histones during transcription and replication (Rufiange et al., 2007). This is important as nucleosomes must be disassembled ahead of the elongating polymerase before it can successfully transcribe a gene. After RNAPII has travelled through a locus, nucleosomes must be replaced. Asf1 is necessary for both the eviction and redeposition of the H3/H4 tetramer *in vivo* (Schwabish and Struhl, 2006). Interestingly, cells deleted for *ASF1* also show a depletion of H3K36me in some strains (Lin et al., 2010), indicating a connection between histone eviction/deposition during transcription and Set2. Further, recent work has shown that *set2Δ* cells show increased H3K56ac, a PTM found on newly synthesized histones, and H4ac levels at the 3' ends of genes in G1 arrested cells (Venkatesh et al., 2012), indicating increased levels of histone exchange. Cells lacking Asf1, on the other hand, demonstrated lower levels of H3K56ac and H4ac, indicating Asf1 and Set2 have opposing function (Venkatesh et al., 2012). Critically, Asf1 shows a decreased affinity for peptides that are H3K36me3 (Venkatesh et al., 2012), suggesting a model where Set2 mediated H3K36me may actively inhibit histone exchange after transcription by preventing Asf1 from binding to and removing H3K36me histones.

IOC4

Ioc4 is a member of the Isw1b ATP-dependent chromatin remodeling complex (Vary et al., 2003). Isw1 is the catalytic subunit of the complex and forms two distinct complexes *in vivo*, one with Ioc4 (Isw1b) and one without (Isw1a). Ioc4 has a PWWP domain, named for to the conserved proline and tryptophan residues, which in humans, have shown preferential binding to H3K36me3 (Vermeulen et al., 2010). The PWWP domain of Ioc4 has also been shown to recognize and bind H3K36me3 (Maltby et al., 2012; Smolle et al., 2012). This facilitates the recruitment of Isw1b to the 3' regions of gene bodies where Isw1 can then slide nucleosomes along the DNA (Smolle et al., 2012). In the absence of Set2 or H3K36me, Isw1b is no longer able to properly associate with chromatin (Maltby et al., 2012; Smolle et al., 2012), creating regions of improperly placed nucleosomes. Finally, though the mechanism is not yet known, Isw1b also functions, in concert with Set2, to prevent histone exchange and limit H4ac (Smolle et al., 2012), again stressing the importance of Set2 and H3K36me in limiting both of these actions once RNAPII has successfully transcribed a gene.

EAF3 and the Rpd3S Complex

Rpd3S, a histone deacetylase complex (HDAC), contains two proteins with reader domains that interact with chromatin: Eaf3 and Rco1. Eaf3 contains a chromo domain that recognizes H3K36me2 (Carrozza et al., 2005; Joshi and Struhl, 2005; Sun et al., 2008; Xu et al., 2008), while Rco1 contains a dual plant homeodomain (PHD) finger motif that is necessary for chromatin engagement (Li et al., 2007a). As discussed in Chapter 3, both PHD fingers are essential for chromatin binding and prefer to bind the

unmodified N-terminus of H3. Interestingly, it has recently been shown that Rco1 exists as a dimer within the Rpd3S complex, adding an additional two chromatin contacts to the complex (Ruan et al., 2016a). Like the Isw1b complex, Eaf3's recognition of H3K36me helps to localize the Rpd3S complex to the 3' ends of gene bodies (Carrozza et al., 2005; Keogh et al., 2005; Li et al., 2007a) where it can deacetylate acetylated H4 tails. In the absence of Set2, Rco1, or Eaf3, or if Eaf3 is unable to bind H3K36me2, H4ac increases at the 3' end of gene bodies (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005; Li et al., 2007a) leading to inappropriate transcription initiation at cryptic promoters within the gene bodies. This is known as cryptic transcription and will be described in detail in the next section.

While Eaf3 and Rco1 play an important role in localizing Rpd3S across the genome, Rpd3S can be recruited to gene bodies independently of Eaf3, Rco1, or H3K36me. Like Set2, Rpd3S can be recruited to chromatin via the phosphorylated CTD of RNAPII, specifically the Ser2/Ser5 dually phosphorylated form (Drouin et al., 2010; Govind et al., 2010). Although, without Eaf3 binding to H3K36me2, or the PHD fingers of Rco1, the Rpd3S complex is catalytically inactive. Recent work has suggested that H3K36me stimulates a conformational change in Rpd3S which increases its affinity for chromatin, perhaps stimulating its enzymatic activity (Ruan et al., 2015).

Interestingly, Rpd3S prefers an H3K36me di-nucleosome substrate (Huh et al., 2012). Rpd3S deacetylase activity is further promoted if these nucleosomes are properly spaced from one another (Lee et al., 2013). The correct spacing is likely ensured by the nucleosome sliding activity of Isw1b, which could explain the increased H4ac levels in *isw1Δ* and *ioc4Δ* cells (Smolle et al., 2012). Without proper spacing, and

even with proper H3K36me, the affinity of Rpd3S for chromatin is decreased to a degree where the naturally low binding affinity of Eaf3 for H3K36me2 is not able to sufficiently engage nucleosomes and stimulate Rpd3's deacetylase activity. Further, Rpd3S also depends on the PHD finger in Rco1 to maintain its localization on chromatin (Li et al., 2007a). If any of these conditions are not met, the net result is increased H4ac levels in gene bodies.

Together, H3K36me functions to maintain nucleosome stability by repelling Asf1 and ensure low levels of H4ac by recruiting the Isw1b chromatin remodeling complex and the Rpd3S HDAC. These activities ensure that RNAPII is not able to engage cryptic promoters within gene bodies (Figure 1.3).

Cryptic Transcription

One of the main functions of Set2 and H3K36me is to prevent cryptic transcription from occurring across the genome. Cryptic transcripts are traditionally defined as transcripts that originate from inside the gene body as opposed to the canonical 5' promoter region (Kaplan et al., 2003). Cryptic transcripts can arise from both the sense and anti-sense direction (Neil et al., 2009; Xu et al., 2009) and seem to occur preferentially at longer genes (Li et al., 2007b). Cryptic transcripts can also be further classified depending on how stable the transcripts are and how they are degraded. Generally, cryptic transcripts are quickly degraded by the exosome (Wyers et al., 2005) and are referred to as cryptic unstable transcripts (CUTs) (Neil et al., 2009; Xu et al., 2009). These transcripts are very unstable and can only be seen when the exosome has been compromised (Wyers et al., 2005), rendering them difficult to study

under normal conditions. Further, their rapid degradation suggests that if left to accumulate within the cell, there may be negative consequences. There is, however, a separate class of transcripts that is much more stable within the cell, the stable unannotated transcripts (SUTs) (Xu et al., 2009). Finally, there is third separate class of transcripts that are reliant on the Xrn1 5'-3' exonuclease for degradation. These are the Xrn1-sensitive unstable transcripts (XUTs) (van Dijk et al., 2011). Genome-wide studies in yeast have shown that cryptic transcripts can arise from within the 3' region of gene bodies, but result from bi-directional transcription events at promoters even more frequently (Neil et al., 2009; Xu et al., 2009).

There are two main mechanisms by which cryptic transcripts are thought to arise: increased histone acetylation and decreased numbers or mis-localization of histones across the gene body. As discussed above, Set2 and H3K36me play an important role in both of these aspects. The main way H3K36me represses cryptic promoters is to limit histone acetylation in gene bodies. It does this by recruiting Rpd3S and Isw1b to the 3' ends of open reading frames (Carrozza et al., 2005; Keogh et al., 2005; Smolle et al., 2012). Rpd3S has five characterized chromatin binding domains: two PHD finger motifs in each copy of Rco1 and the chromo domain of Eaf3. Both PHD fingers and the chromo domain are essential for Rpd3S function (Li et al., 2007a). It is also the case that the uncharacterized N- and C-termini of Rco1 are also critical for Rpd3S function (Ruan et al., 2016b). These regions likely contain the dimerization domains and offer other surfaces for protein-protein contacts necessary for Rpd3S complex integrity, and thus, its proper function.

Isw1b also likely plays a pivotal role in ensuring the gene body remains hypo-acetylated after the elongating RNAPII. A deletion of *ISW1* or *IOC4* results in increased H4ac and increased cryptic transcription (Smolle et al., 2012). This is likely due to the fact that nucleosomes are no longer positioned correctly for Rpd3S to engage a di-nucleosome pair (Lee et al., 2013). Biochemical experiments have demonstrated that Rpd3S has a preferred linker length of ~50 base pairs of DNA (Lee et al., 2013). The addition of Isw1 leads to increased deacetylase activity of Rpd3S (Lee et al., 2013). Isw1b, along with other chromatin remodelers, are likely necessary to accurately prime the chromatin template for deacetylation by Rpd3S. In total, H3K36me serves a double duty both to activate Rpd3S activity through the binding of Eaf3's chromo domain, as well as ensuring that nucleosome spacing is ideal for Rdp3S binding through Isw1b recruitment via the PWWP domain of Ioc4. These functions maintain a repressive chromatin environment behind RNAPII, and reinforce transcription in the sense direction (Churchman and Weissman, 2011).

The other mechanism by which cryptic transcription is repressed is by ensuring that nucleosomes are properly restored after the elongating RNAPII complex has moved through the gene body. In addition to its role in regulating Set2 and H3K36me, Spt6 also plays an important role in this process as well. Spt6 helps to disassemble nucleosomes ahead of RNAPII and reassemble them behind (Kaplan et al., 2003). If Spt6 function is compromised, in addition to a loss of H3K36me, there are also fewer intact nucleosomes on chromatin, allowing for RNAPII access to cryptic promoters (Cheung et al., 2008; DeGennaro et al., 2013). This is also true of other histone chaperones, like Asf1. Here, H3K36me is refractory to Asf1 binding to chromatin, thus

ensuring that once nucleosomes are assembled into chromatin behind RNAPII, they remain so (Venkatesh et al., 2012).

Increased histone acetylation and defects in nucleosome reassembly are likely not the only two mechanisms by which cryptic transcripts arise, however. Surveys of deletions of many chromatin modifying, remodeling, and maintenance factors have shown the ability to promote cryptic transcription (Cheung et al., 2008; Silva et al., 2012). What is also interesting is that the pools of cryptic transcripts produced from these different mutants varies (Smolle et al., 2012), offering further support of different underlying mechanisms regulating cryptic transcript production. What is also curious is that widespread production of cryptic transcripts does not necessarily have a deleterious effect on cell growth. *set2Δ* cells are viable, as are many other deletions that lead to the production of cryptic transcripts. Further, under normal laboratory conditions, a deletion of *SET2* results in relatively few changes to the transcriptome (Lenstra et al., 2011), indicating that production of these transcripts does not drive large-scale transcriptional change. However, there is limited evidence that severe nutrient stress is capable of producing cryptic transcripts in wild-type cells (Cheung et al., 2008). This suggests that cryptic transcripts could act as a defense mechanism or have relevant functions during times of cellular stress.

While cryptic transcripts have historically not had generally characterized functions associated with them, there are currently two thoughts as to their possible effects in the cell. The first is that they could function as transcriptional inhibitors. Genes containing anti-sense transcripts that overlapped with their promoters have recently been shown to have statistically significant drops in full-length transcript

amounts (Huber et al., 2016), though this did not dramatically alter protein levels. With the recent advances in sequencing technology, it has become easier to sequence the transcriptome at great depth, increasing the discoverability and ability to map cryptic transcripts, with the end goal being the elucidation of their function genome wide.

The other function of cryptic transcripts could be the production of cryptic proteins. Using a Spt6 mutation that produces a very strong cryptic transcription phenotype, the Winston lab has observed the translation of a select few cryptic transcripts (Cheung et al., 2008). While there has been no function ascribed to these proteins yet, they raise the intriguing possibility that the increasing complexity that we are just beginning to uncover in the transcriptome could be translated to the proteome. One particularly compelling hypothesis is that these cryptic proteins could behave as dominant negative variants of their full-length counterparts. This is primarily relevant to in-frame, sense, cryptic transcripts as they would theoretically contain fully functional C-terminal domains of a protein, while lacking domains at the N-terminus. Quantitative mass spectrometry data sets examining the proteomes of wild-type and cryptic transcript producing cells will be needed to understand how wide-spread the production of cryptic proteins are and what role they may be playing in the cell.

Other Functions of Set2 and H3K36me and Their Evolutionary Conservation

Aging

In addition to the repression of cryptic transcription, Set2 and H3K36me play several other critical roles in the cell. The first is to regulate aging, likely through the maintenance of transcriptional fidelity across the genome (Sen et al., 2015). The

Berger group showed that a loss of Set2 or a mutation of H3K36 decreased the lifespan of yeast cells and critically, removal of Rph1, an H3K36me3 demethylase, extended the lifespan of these cells significantly. Interestingly, they found that as yeast cells aged, they lost H3K36me genome-wide and a significant increase in cryptic transcription was observed. The loss in H3K36me as yeast cells age could be due to a global loss in histones. It has been observed that older yeast cells can lose as much as 50% of their histones as they age, further increasing the likelihood of transcriptional de-repression genome-wide (Hu et al., 2014). Critically, these functions are likely conserved in higher eukaryotes as the Set2 homolog in *C. elegans*, MET-1, has also been shown to extend the lifespan of worms (Hamilton et al., 2005; Pu et al., 2015; Sen et al., 2015). In summary, H3K36me functions in yeast and worms to regulate the transcriptome by preventing cryptic transcription and thus decreasing the transcriptional noise in the genome. This ensures that genes are properly expressed throughout the organism's lifespan, promoting longevity.

RNA Splicing

The role of H3K36me in regulating splicing was first observed in human cells (Luco et al., 2010). It was found that H3K36me3 regulated the exon choice in the *FGFR2* gene. In epithelial cells exon IIIb is included in the mRNA transcript, whereas in mesenchymal cells, exon IIIc is included. Overexpression or siRNA mediated knockdown of *SETD2* demonstrated a shift in exon inclusion in both cell lines. The mechanism of exon choice was shown to be the recruitment of the splicing factor PTB via an interaction with MRG15, the human *EAF3* homolog. MRG15 was recruited to this

locus in an H3K36me3 dependent manner and when there, recruited PTB to carry out a specific splicing pattern.

The role of SETD2 in splicing has been confirmed genome-wide using next generation sequencing techniques (Guo et al., 2014; Simon et al., 2014). Here, a plethora of RNA splicing defects, including intron retention and aberrantly spliced genes were observed. Interestingly, nucleosome positioning was altered at many sites of RNA processing defects, suggesting that H3K36me's chromatin remodeling function and repression of histone exchange may also be conserved in human cells.

Finally, Set2 has also been shown to regulate splicing in yeast. Mutants lacking H3K36me demonstrated reduced splicing efficiency across the genome (Sorenson et al., 2016). Interestingly, the correct splicing of these genes was dependent on the association of Set2 with the CTD of RNAPII as well as H3K36me2. The importance of the SRI domain suggests that Set2 may act co-transcriptionally to recruit splicing factors while the necessity of H3K36me2 suggests that Rpd3S may also play an important role in this process.

DNA Damage Response

As DNA is packaged into chromatin, histones play a key role in the DNA damage response pathways. Like the above discussions, Set2 and H3K36me are necessary for the DNA damage response, though the mechanism differs between organisms.

In both fission yeast and budding yeast, cells lacking Set2 and H3K36me display strong sensitivity to DNA damaging agents (Jha and Strahl, 2014; Pai et al., 2014; Winsor et al., 2013). In both systems, Set2 regulates chromatin compaction and limits

resection, the creation of single stranded DNA at sights of DNA damage, allowing for the binding of DNA damage response proteins like RPA. Limiting resection promotes non-homologous end joining (NHEJ) (Jha and Strahl, 2014; Pai et al., 2014). In budding yeast it was shown that there were also defects in the DNA damaging response signaling pathways (Jha and Strahl, 2014), suggesting that the chromatin state surrounding the damage site is critical for recruiting DNA damage response proteins and their proper signaling activities. In fission yeast, H3K36me was shown to be cell cycle regulated, further delineating the choice between homologous recombination (HR) and NHEJ (Pai et al., 2014).

In human cells, HR is defective in SETD2 mutant cells (Carvalho et al., 2014; Kanu et al., 2015; Pfister et al., 2014). This seems to be the result of a lack of RPA and RAD51 binding to the DNA damage sites (Kanu et al., 2015; Pfister et al., 2014). Further, upon loss of H3K36me3, DNA damage persists in cells, likely due to an inability of downstream signaling proteins critical for efficient repair, like p53, to be activated (Carvalho et al., 2014).

Together, this work demonstrates that Set2 and H3K36me play an important role in maintaining the integrity and stability of the genome.

DNA Methylation

Finally, H3K36me has been shown to be necessary for establishing DNA methylation in gene bodies. DNA methylation is a critical regulator of gene expression, where hyper-methylation of genes results in repression. Both of the *de novo* DNA methyltranferases, DNMT3a and DMNT3b contain PWWP domains. Both of the PWWP

domains in DMNT3a and DNMT3b have been shown to bind to H3K36me3 (Dhayalan et al., 2010; Rondelet et al., 2016) and are necessary for their DNA methyltransferase activity (Baubec et al., 2015; Dhayalan et al., 2010). In the absence of SETD2, DNA methylation is not able to be targeted to transcribed sequences in the genome (Baubec et al., 2015; Morselli et al., 2015). In addition to the other problems with the transcriptome, a loss of genic DNA methylation is likely to cause a further destabilization in the cellular pool of RNA.

The Role of SETD2 and H3K36me in Cancer

SETD2 has been found to be mutated in up to 15% of patients with clear cell renal cell carcinoma (ccRCC) (Dalgliesh et al., 2010; Gerlinger et al., 2012; Varela et al., 2011). SETD2 is located on chromosome 3p and this region is commonly deleted in ccRCC tumors (Varela et al., 2011). ccRCC tumors are very heterogeneous, but mutations in SETD2 arise frequently and independently in a single tumor (Gerlinger et al., 2012). SETD2 mutations have also been observed in acute leukemias (Mar et al., 2014; Zhu et al., 2014).

While the exact molecular underpinnings of the role SETD2 and H3K36me play in preventing cancer, regulation of the transcriptome and DNA damage repair are likely candidates. As discussed above, SETD2 plays a critical role in maintaining transcriptional fidelity. In its absence, there are mRNA processing defects at as many as 25% of expressed genes across the genome (Simon et al., 2014). Further, it has recently been shown in ccRCC that there are many transcription termination defects in SETD2 mutant cancer cell lines (Grosso et al., 2015). These termination defects lead to

the creation of chimeric transcripts, some involving oncogenes, providing yet another potential mechanism that could lead to cancer development. Together, the defects in mRNA processing, termination defects, and impaired DNA damage signaling likely provide an ideal environment for tumorigenesis.

Finally, it has recently been shown that mutations of the histone H3 lysine 36 residue can also lead to cancer. When this residue is mutated to a methionine at a single copy of the gene, H3K36me is lost genome-wide, leading to tumor formation (Fang et al., 2016; Lu et al., 2016). As H3K36me is lost from the genome, this leads to a redistribution of H3K27me, a PTM laid down by the polycomb repressive complex. Though rather than increasing gene repression globally, this redistribution instead dilutes the polycomb repressive complex 1 across the genome and results in increased gene expression (Lu et al., 2016). This disrupts the normal differentiation progression and locks the cells in an undifferentiated state. Thus, mutating either SETD2 or H3K36 is sufficient to promote cancer progression.

Concluding Thoughts and the Contributions of this Work

H3K36me is one of the few histone PTMs that is conserved from yeast to humans, suggesting that it is extremely important for proper cellular function. Much of the research to date has demonstrated that H3K36me is necessary for transcriptional fidelity across the genome by repressing cryptic transcription. Through the co-transcriptional recruitment of the Isw1b and Rpd3S complexes, H3K36me ensures that nucleosomes are remodeled correctly and that they are deacetylated in the wake of the elongating RNAPII complex. Further, H3K36me ensures that the histone chaperone,

Asf1, cannot remove histones H3 and H4 from chromatin once they have been replaced. Together, these functions act together to ensure that chromatin is in a repressive state behind RNAPII and is refractory to RNAPII binding within the open reading frame.

In addition, this control of nucleosome retention plays a key role in the DNA damage response. In the absence of Set2 and H3K36me, there is increased amounts of DNA resection, which leads to a failure of the DNA damage response system to function properly. While the importance of Set2 is well established in this process, it is currently unknown what role the H3K36me effector proteins play in this process.

Finally, H3K36me is necessary to correctly process mRNAs. Incorrectly spliced and terminated transcripts increase in the absence of H3K36me and likely lead to tumorigenesis. Though, again, it is currently unknown mechanistically how H3K36me properly mediates these processes.

The work in this dissertation seeks to broaden our understanding of the role of H3K36me in the cell. This mission has been undertaken in three separate ways: first, I characterized a novel H3K36me effector protein, Pdp3. This protein is a member of the newly characterized NuA3b histone acetyltransferase (HAT) complex and its recognition of H3K36me is necessary for its function. Second, I further characterized how Rpd3S interacts with chromatin. I found that the previously uncharacterized second PHD finger domain was essential for Rpd3S function and acts cooperatively with the N-terminal PHD finger domain to engage the unmodified H3 N-terminus. Finally, I demonstrate that both Set2 and H3K36me are necessary for the cellular response to nutrient stress. This is likely due to a disruption of the transcriptome that leads the cells poorly able to

effectively respond to cellular stress. Critically, in the absence of Set2, Tor1 and a mitogen-activated protein (MAP) kinase pathway are disrupted and cells are unable to respond properly to nutrient stress. Interestingly, this relationship is likely conserved as human cancer cells treated with mammalian target of rapamycin (mTOR) inhibitors in combination with SETD2 knockdown see significant growth defects and cell death (Zhu et al., 2014). Taken together, this work provides new insight into the function of Set2 and H3K36me in the cell.

Figure Legends

Figure 1.1: Domain Architecture and Function of Set2

Set2 is a 733 amino acid long protein that contains a N-terminal SET domain as well as a C-terminal WW, CC, and SRI domain. There is also a small H4 interaction domain at the N-terminus of the protein. Set2 methylates histone H3 at lysine 36. In *S. cerevisiae,* Set2 carries out mono-, di-, and tri-methylation of this residue.

Figure 1.2: The Regulation of Set2 and H3K36 Methylation

Many factors come together to regulate the stability of Set2 and the production of H3K36me. The Paf1 complex directly interacts with Spt6, stabilizing its protein levels. Spt6 and the CTD Ser2 kinase also interact and mutually regulate each other's protein levels. Spt6 likely helps to recruit Ctk1 to the CTD. Phospho-Ser2 recruits Set2 to the CTD of RNAPII and stabilizes its protein levels while promoting H3K36me3.

Figure 1.3: The Role of H3K36me Effector Proteins

Rpd3S binds H3K36me2 via the chromo domain of Eaf3. This allows for the deacetylation of histones. Isw1b binds H3K36me3 via the PWWP domain of Ioc4 promoting proper nucleosome spacing across the gene body. Finally, Asf1, a histone chaperone, is repelled by H3K36me3, ensuring that nucleosomes are stable once incorporated into chromatin. In wild-type cells, these functions act together to regulate nucleosome position and repress cryptic promoters. In the absence of Set2, nucleosomes are not positioned correctly and normally obscured promoters become active.

Figure 1.1

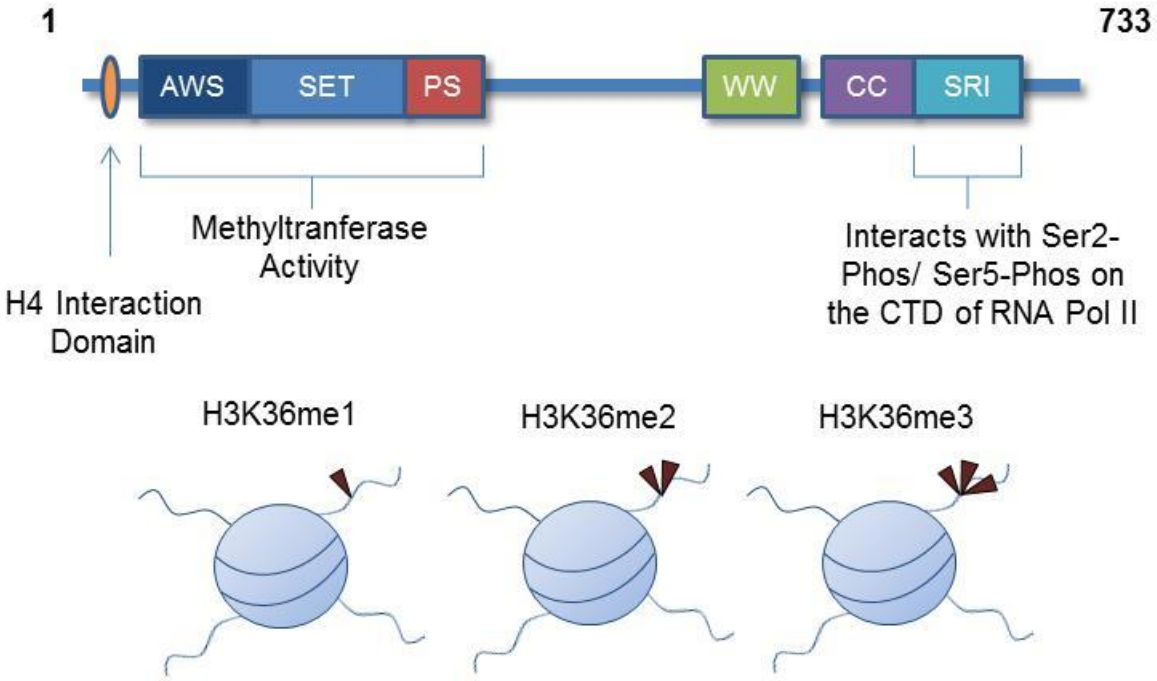

1
733
AWS
SET
PS
WW
CC
SRI
Methyltranferase Activity
H4 Interaction Domain
Interacts with Ser2-Phos/ Ser5-Phos on the CTD of RNA Pol II
H3K36me1
H3K36me2
H3K36me3

Figure 1.2

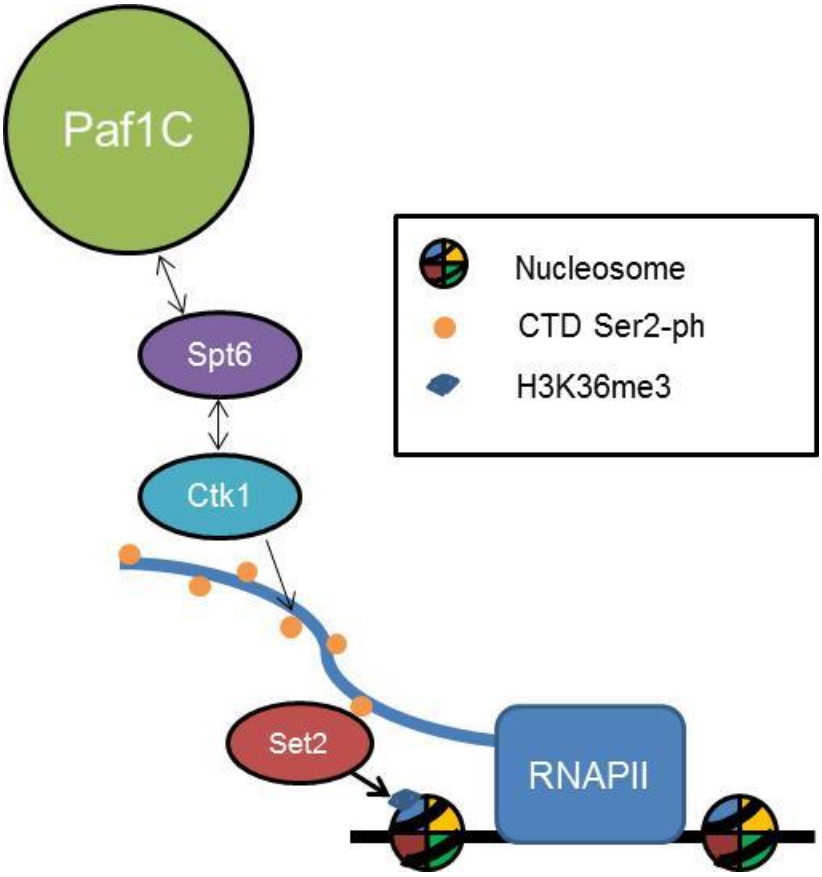
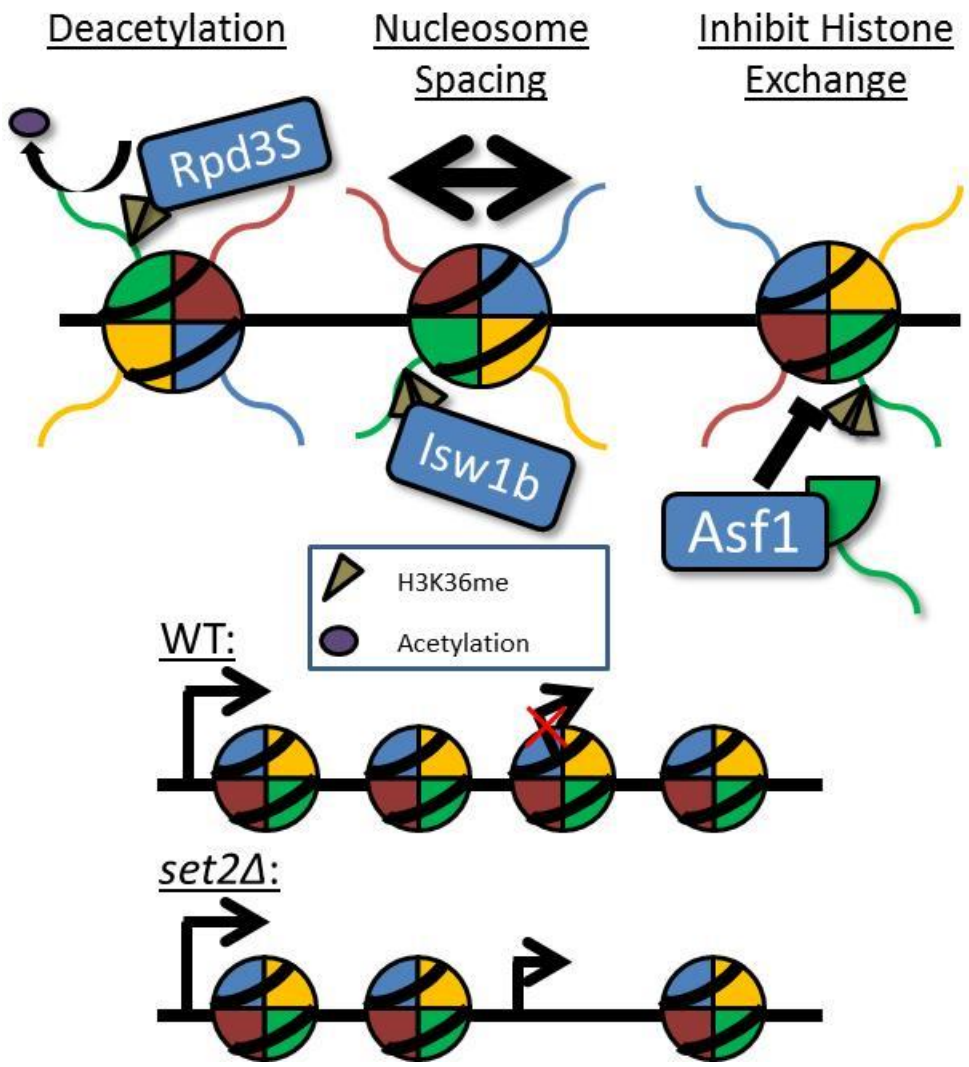

Paf1C
Spt6
Ctk1
Set2
RNAPII
Nucleosome
CTD Ser2-ph
H3K36me3

Figure 1.3


Deacetylation
Nucleosome Spacing
Inhibit Histone Exchange
Rpd3S
Isw1b
Asf1
H3K36me
Acetylation
WT:
set2Δ:

CHAPTER 2 – AN H3K36me3 BINDING PWWP PROTEIN TARGETS THE NuA3 ACETYLTRANSFERASE COMPLEX TO COORDINATE TRANSCRIPTIONAL ELONGATION AT CODING REGIONS[1]

Introduction

Eukaryotic DNA is wrapped around octamers of evolutionarily conserved core histone proteins H3, H4, H2A, and H2B, forming nucleosomes, the fundamental unit of chromatin. Chromatin acts as a barrier to the transcriptional machinery, and therefore precise coordination of nucleosome organization is required for the passage of RNA polymerase II (RNAPII) (Kulaeva et al., 2009; Kulaeva et al., 2010; Smolle et al., 2013). Nucleosome organization is regulated by chromatin remodelers, histone chaperones, and other complexes that enzymatically add, remove, or bind ("write", "erase", or "read", respectively) post-translational modification (PTM) states of histones (Gardner et al., 2011; Petesch and Lis, 2012; Strahl and Allis, 2000; Taverna et al., 2007). Together, these factors ensure that genes remain accessible to transcription factors, activators/co-activators, and RNAPII (Petesch and Lis, 2012). Critically, these factors also function to

[1] Portions of this chapter were adapted from Gilbert, T.M., McDaniel, S.L., Byrum, S.D., Cades, J.A., Dancy, B.C., Wade, H., Tackett, A.J., Strahl, B.D., and Taverna, S.D. (2014). A PWWP domain-containing protein targets the NuA3 acetyltransferase complex via histone H3 lysine 36 trimethylation to coordinate transcriptional elongation at coding regions. Molecular & cellular proteomics : MCP *13*, 2883-2895.

restore chromatin structure following the passage of RNAPII during transcription elongation (Lee and Shilatifard, 2007; Smolle et al., 2013).

Histone PTMs are established at gene loci in a context specific manner, typically defined by the position along a gene (i.e., promoters versus open reading frames (ORFs)), as well as the transcriptional status of that gene (Barrera and Ren, 2006; Fuchs et al., 2009; Gardner et al., 2011; Heyse et al., 2009; Liang et al., 2004; Pokholok et al., 2005; Rando, 2007; Strahl and Allis, 2000). The dual capacity of chromatin complexes to read and write histone PTMs confines certain PTM combinations to discrete regions within genomic loci (Bian et al., 2011; Choy et al., 2001; Klein et al., 2013; Li et al., 2007a; Taverna et al., 2006; Vermeulen et al., 2010). For example, histone H3 is often combinatorially modified by tri-methylation on lysine 4 (H3K4me3) and acetylation, particularly on lysine 14 (H3K14ac), at the 5'-ends of actively transcribed genes (Liang et al., 2004; Pokholok et al., 2005). NuA3, a conserved *S. cerevisiae* histone acetyltransferase (HAT) complex (Eberharter et al., 1998), specifically binds H3K4me3 that is generated by the histone methyltransferase (HMT) Set1, through the plant homeodomain (PHD) finger in the Yng1 subunit (an ortholog of human ING5) (Briggs et al., 2001; Doyon et al., 2006; Howe et al., 2002; John et al., 2000; Martin et al., 2006a; Martin et al., 2006b; Taverna et al., 2006). NuA3 then acetylates H3K14 on the same H3 molecule through the HAT domain of Sas3 (an ortholog of human MYST3) (Doyon et al., 2006; Howe et al., 2002; John et al., 2000; Taverna et al., 2006). Subsequently, additional factors bind either NuA3-catalyzed acetylation or NuA3 subunits themselves, including the chromatin remodeling complex RSC (which contains an H3K14ac-binding bromodomain), and the histone chaperone

complex FACT (which binds to Sas3), to promote transcription initiation at a subset of genes (Ishimi and Kikuchi, 1991; John et al., 2000; Park et al., 2005; Smart et al., 2009; Taverna et al., 2006; VanDemark et al., 2008).

Other histone PTMs positioned within gene bodies further facilitate transcription and maintain transcript fidelity (Fuchs et al., 2009; Lee and Shilatifard, 2007; Smolle et al., 2013; Wagner and Carpenter, 2012). In *S. cerevisiae*, the Set2 HMT localizes to the bodies of actively transcribed genes (Pokholok et al., 2005; Strahl et al., 2002) by physically binding to the hyperphosphorylated C-terminal domain (CTD) of RNAPII (Kizer et al., 2005; Krogan et al., 2003; Li et al., 2003; Li et al., 2002; Schaft et al., 2003; Xiao et al., 2003). CTD binding by Set2 is required for the establishment of H3K36me3 (Kizer et al., 2005). H3K36 methylation correlates with transcription elongation in yeast (Kizer et al., 2005; Krogan et al., 2003; Morris et al., 2005; Pokholok et al., 2005; Schaft et al., 2003) and maintains chromatin integrity by recruiting complexes that collectively restore chromatin structure (Lee and Shilatifard, 2007; Smolle et al., 2013; Wagner and Carpenter, 2012). For example, the histone deacetylase (HDAC) complex, Rpd3S, engages H3K36me2 via the chromodomain of Eaf3 and the PHD finger of Rco1 (Carrozza et al., 2005; Huh et al., 2012; Joshi and Struhl, 2005; Keogh et al., 2005; Li et al., 2007a; Li et al., 2009; Sun et al., 2008; Xu et al., 2008). Rpd3S generates a hypoacetylated environment behind the elongating RNAPII, which compacts chromatin and represses cryptic transcription (Carrozza et al., 2005; Keogh et al., 2005; Li et al., 2007b; Li et al., 2009; Lickwar et al., 2009). H3K36me3 also maintains chromatin integrity by blocking *trans*-histone exchange through at least two mechanisms: the steric reduction of histone chaperone affinity for histone targets (Venkatesh et al., 2012) and

the recruitment of chromatin remodelers that preserve H3K36me3/hypoacetylated histones (Maltby et al., 2012; Smolle et al., 2012; Smolle et al., 2013). Specifically, H3K36me3 precludes Asf1 from depositing newly synthesized histones (Venkatesh et al., 2012) and recruits the ATP-dependent remodeler, Iws1b, via the Pro-Trp-Trp-Pro (PWWP) domain of the Ioc4 subunit, to position nucleosomes in a manner that stimulates Rpd3S activity (Lee et al., 2013; Maltby et al., 2012; Smolle et al., 2012). Recently, human PWWP domain proteins have been shown to bind H3K36me3 (Vermeulen et al., 2010; Wu et al., 2011) and perform a variety of functions, including the regulation of transcription (Laue et al., 2008; Vezzoli et al., 2010), DNA methylation guidance (Dhayalan et al., 2010), and alternative splicing (Pradeepa et al., 2012).

Using recent advances in mass spectrometry, we previously found that Ylr455w, an uncharacterized PWWP domain-containing protein (Krogan et al., 2006; Maurer-Stroh et al., 2003; Tackett et al., 2005b), co-purified with stable members of the NuA3 HAT complex (Taverna et al., 2006). We propose that Ylr455w be called Pdp3: PWWP domain protein in NuA3. Moreover, Pdp3 (Q09842) is also the name of a PWWP domain-containing protein in *S. pombe* that interacts with homologus subunits of the NuA3 complex (Wang et al., 2012).

Here, we further characterize NuA3 and find two functionally distinct forms: NuA3a and NuA3b. NuA3a binds H3K4me3, through the PHD finger of Yng1, and acetylates H3K14 at the 5'-ends of actively transcribed genes to promote transcription initiation (Howe et al., 2002; Martin et al., 2006a; Martin et al., 2006b; Taverna et al., 2006). In contrast, NuA3b contains the unique member, Pdp3, which recruits the complex to chromatin through its PWWP domain binding to H3K36me3. Deletion of the

PDP3 gene decreases NuA3-directed transcription and results in growth defects when combined with transcription elongation mutants, suggesting NuA3b functions in the transcription elongation process. While H3K36me2/3 can act as a repressive mark that protects chromatin integrity during transcription elongation (Lee and Shilatifard, 2007; Smolle et al., 2013; Wagner and Carpenter, 2012), the work described in this study suggests that H3K36me3 can also act to positively regulate transcription elongation through the recruitment of the NuA3b complex via Pdp3.

Materials and Methods

Mass Spectrometric Protein Identification

Pdp3-TAP protein complex purification was performed with *S. cerevisiae* grown to log phase in YPD essentially as described to maintain complex integrity (Tackett et al., 2005a; Tackett et al., 2005b). Proteins co-purifying with Pdp3 were subjected to tandem MS analysis of peptides with a Thermo LTQ-XL mass spectrometer coupled to an Eksigent nanoLC 2D system as described (Smart et al., 2009). Spectral counts and proteins were identified with Mascot.

Database Searching

Tandem mass spectra were extracted by Thermo ExtractMSn version 1.0.0.8. Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.3.01). Mascot was set up to search the SwissProt_57.15 database (selected for *Saccharomyces cerevisiae*, 57.15, 6973 entries), assuming the digestion enzyme was non-specific.

Mascot was searched with a fragment ion mass tolerance of 0.60 Da and a parent ion tolerance of 2.0 Da. Iodoacetamide derivative of cysteine was specified in Mascot as a fixed modification. S-carbamoylmethylcysteine cyclization (N-terminus) of the N-terminus, oxidation of methionine, pyro-carbamidomethyl of the N-terminus and acetyl of asparagine, proline and the N-terminus were specified in Mascot as variable modifications.

Criteria for Protein Identification

Scaffold (version Scaffold_4.0.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 20.0% probability by the Peptide Prophet algorithm (Keller et al., 2002). Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 1 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al., 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

Protein Expression

PDP3 constructs were made with an N-terminal HIS6-pfuMBP$_{(60-434)}$-FLAG tag (pET28a derivative vector obtained from the G. Bowman Laboratory, Johns Hopkins University) and/or an N-terminal Thioredoxin-HIS6-S•tag tag (pET32a vector, Millipore). Proteins were exogenously expressed in BL21 chemically competent *E. coli* (Invitrogen)

after overnight induction with 1 mM IPTG at 18-20 ºC in LB medium. Point mutants were made with the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and expressed as described.

Protein Purification

For pull-down assays, BL21 cells (pET28a constructs) were resuspended in purification buffer (50 mM Tris pH 7.5, 500 mM NaCl, 40 mM imidazole, 10% glycerol, 2 mM β-ME, 1 mM PMSF, 2 mM benzamidine, pH 8.0) and lysed by sonication (Branson). Clarified lysate was nutated with Ni-NTA agarose resin (Invitrogen) for at least 1 hour at 4 °C. Resin was washed with purification buffer and protein was eluted with purification buffer containing 300 mM imidazole. Protein was flash frozen in liquid nitrogen and stored at −80 °C. For fluorescence polarization assays, BL21 cells (pET32a constructs) were resuspended in purification buffer (50 mM Tris pH 7.7, 500 mM NaCl, 10% glycerol, 5 mM DTT, 1 mM PMSF, 2 mM benzamidine) and lysed with a microfluidizer (Watts Fluidair). Clarified lysate was run through a BioScale Mini Profinity IMAC cartridge (BioRad) using an AKTA Purifier system (GE Healthcare). The cartridge was washed with purification buffer containing 12 mM imidazole and protein was eluted with purification buffer containing 125 mM imidazole. Protein was exchanged into gel filtration buffer (50 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol, 5 mM DTT) and separated by a Superdex 200 26/60 column (GE Healthcare) using an AKTA Purifier system (GE Healthcare). Monomeric protein was flash frozen in liquid nitrogen and stored at −80 °C.

In Vivo Pull-down Assays

TAP-tagged *S. cerevisiae* strains were grown to mid-log phase in YPD, cryogenically lysed with a mixer mill (Retsch MM301), and stored at −80 °C. Cells (1 g per pull-down condition) were homogenized (ProScientific) in 650 mM extraction buffer (650 mM NaCl, 20 mM HEPES pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 1 mM PMSF, 0.2% Triton X-100, 1% BSA, 40 mM imidazole) at a ratio of 1 mL buffer per 1 g yeast and nutated for 1 hour at 4 °C. Clarified extracts were diluted to 300 mM NaCl with 'no-salt' extraction buffer, mixed with Ni-NTA agarose resin (Invitrogen) (100 µl per sample pre-coated with Pdp3 protein), and nutated for 30 minutes at 4 °C. Resin was washed 5 times in 300 mM wash buffer (300 mM KCl, 20 mM HEPES pH 7.9, 0.2% Triton X-100, 1% BSA, 40 mM imidazole) and 1 time in buffer containing 10 mM NaCl and 4 mM HEPES pH 7.9. Resin was incubated in 2X SDS-PAGE loading buffer containing 300 mM imidazole for 10 minutes to elute Pdp3-bound proteins. Samples were boiled for 5 minutes, resolved on 8% SDS-polyacrylamide gels, transferred to PVDF membrane, and probed with antibodies recognizing the PrA (DAKO P0450, 1/1500) and FLAG (SIGMA F3165, 1/1000) tags. Immunoblots were visualized using HRP-conjugated secondary antibodies and ECL solution (GE Healthcare). Inputs represent ~0.02-0 .05% of total yeast lysate.

Peptide Synthesis for Peptide Pull-down Assays

Peptides were synthesized as previously described by the C. D. Allis Laboratory (The Rockefeller University) or the UNC Peptide Synthesis and Arraying Core Facility (Rothbart et al., 2012b).

Peptide Pull-down Assays

Steptavidin-coupled Dynabeads (Invitrogen) (25 µl per sample) were incubated with biotinylated histone peptides (1 µg per sample) in binding buffer (20 mM HEPES pH 7.9, 150 mM NaCl, 0.5 mM PMSF, 20% glycerol, 0.2% Triton X-100, 1% BSA) for 1 hour at r.t. Unbound peptide was washed in binding buffer and beads were incubated with purified Pdp3 proteins (20 µg per sample) for 1 hour at r.t. Beads were washed 3 times for 5 minutes each with binding buffer, and 1 time with buffer containing 4 mM HEPES pH 7.9, 10 mM NaCl, 0.5 mM PMSF, 20% glycerol, and 0.2% Triton X-100. Peptide-bound proteins were eluted in boiling 2X SDS-PAGE loading buffer. Samples were resolved on 15% SDS-polyacrylamide gels, transferred to PVDF membrane, and probed with antibodies recognizing the FLAG (SIGMA F3165, 1/1000) and streptavidin (Molecular Probes S-911, 1/10,000) tags. Immunoblots were visualized using HRP-conjugated secondary antibodies and ECL solution (GE Healthcare). Inputs represent .5µg of Pdp3 proteins.

Peptide Synthesis for Fluorescence Polarization Assays

Fluorescent peptides were synthesized using standard Fmoc-solid phase peptide chemistry on a Prelude Peptide Synthesizer (Protein Technologies). Peptides were made on a 0.05 mMol scale with 4 eqs. of amino acids using Rink Amide AM resin (Novabiochem) to generate peptide amides. 5-Carboxyfluorescein (5-FAM) (Chempep) was coupled to the peptides using Lys(ivDde) (Chempep). The ivDde protecting group was orthogonally removed using standard deprotection procedures. The resulting

peptides were cleaved using TFA/thioanisole/water/triisopropylsilane/phenol (87.5:2.5:2.5:2.5:5 v/v) and purified with a Varian Dynamax Microsorb C18 preparative column (Agilent). Purified peptide was lyophilized and its mass was confirmed with an Applied Biosystems Voyager DE-STR MALDI-TOF mass spectrometer (Life Technologies). Of note, to obtain the H3K79me3 5-FAM linked peptide, it was necessary to install Fmoc-(FmocHmb)Phe-OH (Novabiochem) at F84.

Fluorescence Polarization Assays

Binding assays were performed as previously described by (Rothbart et al., 2012a) with the following modifications. Full-length purified wildtype Pdp3, F18A, and W21A proteins were exchanged into FP buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM DTT) and concentrated to ~260-430 µM using Amicon Centrifugal Filter Units MWCO 30,000 (Millipore). Binding assays were performed in a 60 µl volume with 96 well half area black flat bottom non-binding surface plates (Corning). Protein was serially diluted with FP Buffer in 2-fold increments and incubated with 120 nM of 5-FAM labeled histone peptides. Following a 30-minute equilibration period, fluorescence was detected at r.t. with an Infinite M1000 plate reader (Tecan) using a 470 nm excitation filter and 527± 20 nm emission filter. Binding curves were analyzed by the total binding equation $Y= B_{max}* X / (K_d + X) + NS * X + Background$, where $B_{max}= 1$ and non-specific (NS) and background variables are constrained to be equal between peptides, using Prism 5.0 (GraphPad Inc.). Error bars represent the SD of a representative experiment (n= 2) performed in triplicate.

Yeast Strains and Cell Spotting Assays

S. cerevisiae strains were created using heterologous gene replacement (Janke et al., 2004). Strains were grown on YPD or synthetic complete (SC) media as indicated. *BUR1* delete shuffle strains were grown on SC-Ura plates to maintain the wild-type *BUR1* plasmid prior to plating on media containing the drug 5-Fluoroorotic acid (5-FOA) (Boeke et al., 1987). For cell spotting assays, either 0.5 or 2 ODs of cells were 5-fold serially diluted, spotted onto the appropriate plates, and incubated at 30 °C for 2-3 days as indicated.

Immunoblot Analyses for Chromatin Association Assay

5 ODs of cells were isolated and lysed by bead beating in SUTEB buffer (1% SDS, 8 M urea, 10 mM Tris pH 6.8, 10 mM EDTA, 0.01% bromphenol blue) for 3 minutes. The lysates were then boiled for 10 minutes and the supernatant was clarified and isolated. 5 µl of extracts were resolved on 15% SDS-polyacrylamide gels and then transferred to PVDF membrane for 90 minutes at 45 mA. Blots were dried in methanol, washed in TBST (Tris buffered saline with 0.05% Tween 20), and then incubated overnight at 4 °C with the indicated antibodies: Protein A (Sigma Aldrich), G6PDH (Sigma Aldrich), H3 (in house), H4 (Abcam: ab10158), H3K14ac (Millipore: 07-353), H3K36me3 (Abcam: ab9050), and H3K4me3 (Active Motif: 39159). Immunoblots were visualized using HRP-conjugated secondary antibodies and ECL Prime solution (GE Healthcare).

Chromatin Association Assay

Strains were grown overnight in YPD to confluence. Each strain was diluted to 0.1 ODs in 50 mLs of YPD and grown to an OD of ~0.8-1. 40-50 ODs of cells were then isolated, washed with water and SB buffer (1 M Sorbitol, 20 mM Tris.Cl pH 7.4), and frozen at -80 °C until ready for isolation. Cells were resuspended in 1 mL of PSB buffer (20 mM Tris.Cl pH 7.4, 2 mM EDTA, 100 mM NaCl, 10 mM β-ME) and then 1 mL of SB buffer. Cells were then spheroplasted with Zymolyase (Seikagaku Biobusiness) for 30 minutes at r.t. Spheroplasts were spun down at 2000 x *g* and washed with LB (0.4 M Sorbitol, 150 mM potassium acetate, 2 mM magnesium acetate, 20 mM PIPES pH 6.8) twice. TritonX-100 was added to LB (final concentration 1%). Cells were lysed for 15 minutes on ice. Chromatin was isolated by spinning down lysates at 5000 x *g* for 15 minutes. The supernatant was collected and saved as the “soluble” fraction. The chromatin was washed with LB once more and then resuspended into an equal volume to that of the “soluble” fraction. Volume equivalents were resolved on 15% SDS-polyacrylamide gels and subjected to immunoblot analysis.

Relative Transcript Levels

Total RNA was prepared from TAP-tagged YNG1, PDP3, *yng1*Δ, and *pdp3*Δ *S. cerevisiae* strains via Trizol (Life Technologies) and digested with Turbo DNase (Life Technologies). cDNA was synthesized with the Superscript III First Strand Synthesis System (Invitrogen). Differences in the transcript levels of control genes and NuA3 target genes were measured by qPCR using Power SYBR Green PCR Master Mix (Life

Technologies) and a Real Time PCR system (Applied Biosystems v2.1). Primers were designed by Primer3 to target the 5'-end of our genes of interest. Relative transcript levels were calculated using the relative standard curve method. Error bars represent the SEM of a representative experiment (n=4) done in triplicate. Statistical significance was determined by an unpaired two-tailed t-test.

Results

Pdp3 interacts with members of the NuA3 HAT complex

Eaf6, Nto1, Sas3, Taf14, and Yng1 were previously identified as stable members of the *S. cerevisiae* NuA3 HAT complex (Eberharter et al., 1998; Howe et al., 2002; John et al., 2000; Taverna et al., 2006). Pdp3 (Ylr455w), an uncharacterized PWWP domain-containing protein (Krogan et al., 2006; Maurer-Stroh et al., 2003; Tackett et al., 2005b), also showed modest association with the complex (Taverna et al., 2006), as determined through *i*sotopic *d*ifferentiation of *i*nteractions as *r*andom or *t*argeted (i-DIRT) technology (Tackett et al., 2005a). Interestingly, BRPF1, a component of human MOZ/MORF HAT complexes and a homolog of yeast Nto1, contains a PWWP domain that is absent in Nto1 (Doyon et al., 2006; Laue et al., 2008; Vezzoli et al., 2010). We reasoned Pdp3 might function similarly to the PWWP domain of BRPF1 in the NuA3 complex. Genomically TAP-tagged Pdp3 was isolated from *S. cerevisiae* using a method that preserves complex integrity (Tackett et al., 2005a; Tackett et al., 2005b). Proteins co-purifying with Pdp3-TAP were resolved by SDS-PAGE and subjected to MS/MS analysis. Using a 95% protein confidence threshold, 96 proteins were identified by MS/MS (Figure 2.1A). A functional classification of the proteins identified revealed 12

proteins with functions related to transcription (Figure 2.1A). The remaining 84 proteins were “non-specific” co-purifying proteins (e.g., ribosomal, metabolic, nucleolar, heat shock) typically observed in these types of large-scale affinity enrichments (Byrum et al., 2012; Byrum et al., 2013; Smart et al., 2009; Tackett et al., 2005a). Specific proteins co-purifying with Pdp3-TAP include all stable members of NuA3 (Eaf6, Nto1, Sas3, Taf14, Yng1) (Eberharter et al., 1998; Howe et al., 2002; John et al., 2000; Taverna et al., 2006), core histones (H3, H4, H2A, H2B), and components of RNAPII (Rpb2, Rpb4). These co-purifications suggest that Pdp3 is a component of the NuA3 complex and that Pdp3 is involved transcriptional regulation.

To confirm that Pdp3 is a member of the NuA3 complex, we tested the ability of full-length recombinant Pdp3 to pull-down NuA3 subunits from cellular extracts (Figure 2.1B). HIS6-FLAG-tagged Pdp3 was purified from *E. coli* and incubated with lysates from *S. cerevisiae* endogenously expressing TAP-tagged NuA3 subunits. Immunoprecipitated samples were resolved by SDS-PAGE and visualized with an antibody against the TAP tag. HIS6-FLAG-Pdp3 interacts with Yng1, Nto1, and Sas3 (Figure 2.1B), supporting the MS/MS results (Figure 2.1A). Neither Taf14 nor Eaf6 were detected in recombinant Pdp3 pull-downs, which may indicate that these proteins do not interact with Pdp3 (Figure 2.1B, data not shown); however, we cannot exclude the possibility that the TAP tag disrupts native interactions. Also, since Taf14 is a member of multiple chromatin-associated complexes (Kabani et al., 2005), the quantity of soluble Taf14 available to be pulled down in this assay may be insufficient to detect.

Pdp3 specifically engages H3K36me3 through a conserved PWWP domain

H3K36me3 localizes to actively transcribed gene bodies and is associated with transcription elongation in yeast (Kizer et al., 2005; Krogan et al., 2003; Morris et al., 2005; Pokholok et al., 2005; Schaft et al., 2003). BRPF1, a human homolog of yeast Nto1 (Doyon et al., 2006), engages H3K36me3 through a PWWP domain (Laue et al., 2008; Vezzoli et al., 2010) (Figure 2.2) that is absent in Nto1. Interestingly, structural modeling predicts that Pdp3 contains a PWWP domain with aromatic residues critical for methyl-lysine binding, as well as beta strands and alpha helices conserved in other PWWP-domain containing proteins that bind methyl-lysine (Figures 2.2, 2.3A, 2.3B). Because of the structural conservation with the BRPF1 PWWP domain and the homology between the human and yeast HAT complexes, we reasoned that Pdp3 might bind H3K36me3. Therefore, we tested the ability of tri-methylated histone peptides to pull down full-length recombinant Pdp3 (Figure 2.3C). Biotinylated peptides were immobilized on streptavidin resin and incubated with HIS6-FLAG-Pdp3. Immunoprecipitated samples were resolved by SDS-PAGE and visualized with antibodies against FLAG and streptavidin. Pdp3 preferentially engages H3K36me3 compared to all other tri-methylated lysine residues tested (Figure 2.3C).

A conserved aromatic cage within the Pdp3 PWWP domain is required for binding H3K36me3

Like other Royal Family reader modules, PWWP domains employ an aromatic cage to interact with specific tri-methylated histones (Figure 2.2) (Taverna et al., 2007;

Vermeulen et al., 2010; Vezzoli et al., 2010; Yap and Zhou, 2010). For example, BRPF1 requires aromatic residues Y1096, Y1099, and F1147 to coordinate the tri-methylammonium group of H3K36me3 (Figure 2.2) (Vezzoli et al., 2010). These aromatic residues are conserved in the PWWP domain of Pdp3 at positions F18, W21, and F48 (Figures 2.2, 2.3A, 2.3B). To determine whether Pdp3 uses an aromatic cage to bind H3K36me3, we tested the ability of H3K36me3 peptide to pull down full-length recombinant Pdp3 mutated at residues F18, W21, or F48 (Figure 2.3D). All three mutations independently abolish the interaction between Pdp3 and H3K36me3 peptide (Figure 2.3D), suggesting Pdp3 requires a conserved aromatic cage to bind chromatin.

The Pdp3 PWWP domain is necessary and sufficient for binding H3K36me3

We next wanted to determine whether the predicted PWWP domain of Pdp3 was sufficient for H3K36me3 binding. Uniprot defined the Pdp3 PWWP domain as the amino acid residues spanning 7-68 (Figure 2.3B) (Dimmer et al., 2012). However, we found that truncated Pdp3 ($PWWP_{1-74}$) is unable to bind H3K36me3 peptide (Figure 2.3D), suggesting residues beyond the predicted PWWP domain are necessary for Pdp3 function. From the crystal structure of BRPF1 (Vezzoli et al., 2010), we observed that two C-terminal alpha-helices support the aromatic cage and may be critical for engagement of H3K36me3 (Figure 2.2). These alpha-helices are conserved in the structural model of Pdp3 (Figure 2.2) (Kelley and Sternberg, 2009), yet extend beyond the predicted PWWP domain (Figures 2.3A, 2.3B). We created Pdp3 constructs to include increasing segments of the (modeled) C-terminal alpha-helices and tested the ability of H3K36me3 peptide to pull-down $PWWP_{1-110}$, $PWWP_{2-150}$, and $PWWP_{1-219}$. As

shown in Figure 2D, $PWWP_{2-150}$ restores binding to H3K36me3 and thus represents the functional PWWP domain of Pdp3 (Figure 2.3B). These data suggest Pdp3, and likely other PWWP domain proteins, require extended alpha-helical regions for aromatic cage stability and function.

Biophysical characterization of the interaction between Pdp3 and H3K36me3

To biophysically quantitate the specificity of the interaction between Pdp3 and H3K36me3, we performed fluorescence polarization assays using full-length recombinant S·tag-Pdp3 and 5-FAM labeled histone peptides (Figure 2.3E). As expected from our pull-down assays, Pdp3 favors binding to the H3K36me3 peptide over other known targets of PWWP domain proteins, such as H3K79me3 and H4K20me3, with a K_d of 69.51 $\pm$ 3.7 μM (Figure 2.3E) (Qiu et al., 2012; Wu et al., 2011). Notably, this value is one of the lowest reported dissociation constants for PWWP domain proteins that bind histones (Dhayalan et al., 2010; Qiu et al., 2012; Vezzoli et al., 2010; Wen et al., 2014; Wu et al., 2011), which suggests Pdp3 is present in a subset of H3K36me3 enriched chromatin. Pdp3 mutants F18A and W21A bind H3K36me3 peptide ~10 fold weaker than wildtype Pdp3 (Figure 2.3E), further indicating Pdp3 requires the aromatic cage within its PWWP domain to engage H3K36me3. Pdp3 also binds weakly to H3K36me2 peptide with a K_d of ~414 μM (Figure 2.3E). Importantly, the ~6 fold increase in specificity of Pdp3 for H3K36me3 over H3K36me2 suggests Pdp3 has a distinct function from the Rpd3S HDAC complex, which preferentially engages H3K36me2 (Carrozza et al., 2005; Huh et al., 2012; Li et al., 2007a; Li et al., 2009; Youdell et al., 2008) and restores chromatin compaction behind

elongating RNAPII (Carrozza et al., 2005; Keogh et al., 2005; Li et al., 2007b; Li et al., 2009; Lickwar et al., 2009).

Pdp3 requires H3K36me3 to bind chromatin

To further understand the biological role of Pdp3, we next determined whether Pdp3 binds H3K36me3 *in vivo*. We deleted the H3K36me3 methyltransferase, *SET2,* in the *PDP3-TAP* background and used a chromatin association assay to separate soluble proteins from those that bind to chromatin. For added controls, we also deleted the H3K4me3 methyltransferase, *SET1*, and the NuA3 H3K4me3 binding protein, *YNG1*. Strikingly, in the absence of Set2, and thus H3K36me3, we observe an almost complete ablation of Pdp3 binding to chromatin (Figure 2.4A), indicating H3K36me3 is critical for Pdp3 localization *in vivo*. In contrast, Pdp3 remains bound to chromatin in the absence of Set1 (H3K4me3) and Yng1 (Figure 2.4A). The observation that Pdp3 engages chromatin independently of both H3K4me3 and Yng1 suggests that Pdp3 targets NuA3 and other associated proteins to active coding regions, where the majority of H3K36me3 is confined within yeast chromatin (Pokholok et al., 2005; Rando, 2007; Venkatesh et al., 2012). Interestingly, in the absence of Yng1 we note a slight decrease in the amount of chromatin-bound Pdp3 (Figure 2.4A), as well as lower levels of cellular Pdp3 (Figure 2.5A). Together, these data suggest Pdp3 is moderately unstable in *yng1*Δ cells, perhaps due to an inability to be incorporated into NuA3.

We next deleted *SET1, SET2,* and *PDP3* in the *YNG1-TAP* background. In contrast to Pdp3, Yng1 remains bound to chromatin in the absence of either Set1 or Set2 (Figure 2.4B). Furthermore, Yng1 levels are not significantly altered in *pdp3*Δ cells

(Figure 2.5B). These results are consistent with past studies (Chruscicki et al., 2010; Martin et al., 2006a; Martin et al., 2006b) and further suggest that while Yng1 can engage chromatin through both methyl-dependent and independent means, the interaction between Pdp3 and chromatin requires H3K36me3.

Pdp3 is required for NuA3-regulated transcription

Mutation of the Yng1 PHD finger results in genome-wide mislocalization of NuA3 and decreased transcription of NuA3-regulated loci (Taverna et al., 2006). To further assess the function of Pdp3 *in vivo*, we performed RT-qPCR with wildtype, *pdp3*Δ, and *yng1*Δ cells, and calculated the relative transcript levels of NuA3-regulated loci (Figure 2.6A). As expected, in *yng1*Δ cells relative transcript levels of NuA3 target genes show a slight but significant decrease, indicating a positive role for Yng1 in NuA3 activity at these genes (Figure 2.6A) (Doolin et al., 2001; Tadauchi et al., 2001; Taverna et al., 2006). Interestingly, in *pdp3*Δ cells relative transcript levels of NuA3 target genes also decrease (Figure 2.6A). These data indicate that NuA3 binding to both H3K4me3, via Yng1, and H3K36me3, via Pdp3, is required for proper NuA3-directed transcription.

PDP3* and other NuA3 members show genetic interactions with *SET2

Since Pdp3 binds H3K36me3 and is required for proper transcription *in vivo*, we wanted to determine what role, if any, Pdp3 and NuA3 have in the Set2/H3K36 methylation pathway. One key function of H3K36 methylation is to act as a negative regulator of transcription elongation, which is achieved in part through the recruitment of the HDAC Rpd3S (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005; Li

et al., 2007b; Li et al., 2009; Lickwar et al., 2009). Conversely, the cyclin-dependent kinase, Bur1, acts as a positive regulator of transcription through the phosphorylation of several components of elongating RNAPII, including the CTD of RNAPII and the Spt5 C-terminal repeat domain (Keogh et al., 2003; Liu et al., 2009). Under normal conditions deletion of *BUR1* is lethal; however, deletion of *SET2* or *RCO1*, a unique member of Rpd3S, bypasses this lethality (Keogh et al., 2005; Kizer et al., 2005). To explore whether Pdp3 and/or Yng1 contribute to transcription elongation, we examined deletions of *SET2, RCO1, PDP3,* and *YNG1* in the *bur1*Δ background. While the *bur1*Δ*set2*Δ and *bur1*Δ*rco1*Δ strains bypass the lethality of *BUR1* deletion as expected, neither *bur1*Δ*pdp3*Δ nor *bur1*Δ*yng1*Δ cells are viable (Figure 2.6B). This suggests that either the bypass phenotype is too weak to observe, or Pdp3 and Yng1 function as positive elongation factors, and thus would not display a positive growth phenotype in this assay. To distinguish between these two possibilities, we created triple mutant strains with *PDP3* or *YNG1* deleted in the *bur1*Δ*rco1*Δ background. Since this background shows a bypass phenotype, we could now observe positive or negative growth changes resulting from the absence of Pdp3 or Yng1. Significantly, the triple mutant strains both show a decrease in growth as compared to the *bur1*Δ*rco1*Δ background (Figure 2.6B). This suggests that Pdp3 and Yng1 positively regulate transcription elongation. These data are consistent with the idea that NuA3 functions at promoters and gene bodies to facilitate the passage of RNAPII across ORFs.

To further connect NuA3 activity to the transcription elongation pathway, we asked whether the absence of NuA3 complex members resulted in sensitivity or resistance to the transcription elongation inhibitor 6-azauracil (6-AU). Strains containing

a deletion of *SET2,* or downstream effectors, show a 6-AU resistance phenotype (Keogh et al., 2005; Kizer et al., 2005). As expected, strains lacking Set2, Rco1, and Eaf3, a member of the Rpd3S HDAC and NuA4 HAT complexes, display resistance to 6-AU as compared to wild-type (Figure 2.6C). Of all tested NuA3 members, only *pdp3*Δ cells show resistance to 6-AU, a result consistent with the idea that Pdp3 functions in the Set2/H3K36 methylation pathway (Figure 2.6C).

PDP3 **is not synthetically lethal with** ***GCN5***

It has previously been determined that NuA3 complex members Sas3 and Yng1 display synthetic lethality with the HAT Gcn5, (Howe et al., 2001) which indicates Gcn5 and Sas3 likely collaborate to promote gene activation and/or transcription elongation, as Gcn5 is found in both the promoters and transcribed regions of genes (Govind et al., 2007; Sterner et al., 2002). Given the physical interaction and functional overlap between NuA3 and Pdp3, we wanted to determine if Pdp3 is also synthetically lethal with Gcn5. We created *GCN5* shuffle strains that allowed us to delete *YNG1* or *PDP3* in a *gcn5*Δ background. Upon shuffling out the wild-type *GCN5* plasmid, only the *gcn5*Δ*yng1*Δ cells display synthetic lethality (Figure 2.7). Surprisingly, the *gcn5*Δ*pdp3*Δ cells show no deleterious phenotype (Figure 2.7), indicating *PDP3* does not genetically interact with *GCN5*. These results suggest that NuA3 participates in alternate transcriptional pathways, one of which is likely to involve Pdp3 and be distinct from the role of Gcn5 in gene bodies.

Discussion

In this study we characterize a unique form of the NuA3 HAT complex that contains the PWWP domain protein Pdp3. Using mass spectrometric, biochemical, and genetic approaches, our collective findings suggest that NuA3 exists in two functionally distinct forms: NuA3a and NuA3b (Figure 2.8). This nomenclature can be used to distinguish different variations of related protein complexes. For example, while the ATP-dependent chromatin remodeling complexes Isw1a and Isw1b both have the same catalytic protein, Isw1, Isw1b contains a unique PWWP domain protein, Ioc4, required for targeting remodeling activity to H3K36me3 enriched nucleosomes (Maltby et al., 2012; Smolle et al., 2012). For NuA3 complexes, we hypothesize that promoter-associated NuA3a contains the proteins Eaf6, Nto1, Sas3, Taf14, and Yng1, and specifically associates with H3K4me3 using the PHD finger of Yng1 (Figure 2.8A) (Howe et al., 2002; John et al., 2000; Martin et al., 2006a; Martin et al., 2006b; Taverna et al., 2006). NuA3a subsequently acetylates H3K14 through the HAT domain of Sas3, initiating transcription at a subset of genes (Figure 2.8A) (Howe et al., 2002; John et al., 2000; Taverna et al., 2006). In contrast, NuA3b fosters transcription elongation and contains the H3K36me3-binding PWWP domain protein Pdp3, in addition to all stable components of NuA3a (Figure 2.8B). Accordingly, our previous mass spectrometry (i-DIRT) data supports the model that NuA3b is compositionally distinct from NuA3a (Taverna et al., 2006). However, we cannot exclude the alternate possibility that Pdp3 is a member of both NuA3a and NuA3b. To this end, human H3K36me3-binding BRPF proteins resemble a fusion of yeast Nto1 and Pdp3, as if these separate NuA3 proteins are physically linked within the human MOZ/MORF complexes. This brings up the

intriguing possibility that the PWWP domain of Pdp3 (and potentially BRPFs) has a regulated capacity to bind H3K36me3, becoming active or inactive in a context dependent manner.

Our data clearly place Pdp3 and NuA3b within the Set2-dependent transcriptional elongation pathway. Pdp3 binds to H3K36me3, a Set2-catalyzed histone PTM found almost exclusively within transcriptionally active gene bodies (Kizer et al., 2005; Krogan et al., 2003; Morris et al., 2005; Pokholok et al., 2005; Schaft et al., 2003; Strahl et al., 2002); and deletion of the *PDP3* gene results in growth defects when combined with transcription elongation mutants. The interaction of Pdp3 with Rpb4 further supports a role for NuA3b in transcriptional elongation given the links between Rpb4 and actively transcribing RNAPII in *S. pombe* (Kimura et al., 2002). While our data indicate that NuA3b positively regulates transcription elongation, the exact function of NuA3b remains unclear. Since NuA3-directed transcription is decreased in *pdp3*Δ cells, NuA3b could participate in acetylation-dependent nucleosome eviction within the ORF, similar to the proposed function of Gcn5 at coding regions (Ginsburg et al., 2009; Govind et al., 2007; Sanso et al., 2011). While both Gcn5 and Sas3 favor acetylation of H3K14 (Balasubramanian et al., 2002; Grant et al., 1999; Howe et al., 2002; John et al., 2000; Kuo and Andrews, 2013; Taverna et al., 2006; Zhang et al., 1998), unlike NuA3a, NuA3b is not synthetically lethal with Gcn5, suggesting Sas3 within the context of NuA3b may have a distinct target(s) at ORFs (Figure 2.8B). This result and the finding that HBO1 (a related human MYST family HAT) can switch between H4 and H3 acetylation depending on its association with JADE1/2/3 or BRPF1, support the idea that NuA3b may target non-H3K14 substrates (Lalonde et al., 2013). However, future

work is needed to determine the precise mechanism of NuA3b function in the transcription elongation process.

A recurring question in chromatin biology is how the addition or removal of methylation groups on specific lysine residues can alter functional properties of the associated chromatin. In yeast, Set2 catalyzes all three forms of H3K36 methylation (Strahl et al., 2002). To date, H3K36me1 is suggested to function in DNA replication (Pryde et al., 2009), and while H3K36me2 and me3 are both linked to transcriptional elongation, H3K36me2 is essential for Rpd3S recruitment (Carrozza et al., 2005; Huh et al., 2012; Joshi and Struhl, 2005; Keogh et al., 2005; Li et al., 2007a; Li et al., 2009; Sun et al., 2008; Xu et al., 2008) and H3K36me3 is implicated in nucleosomal positioning and repression of *trans*-histone exchange (Maltby et al., 2012; Smolle et al., 2012; Venkatesh et al., 2012). Interestingly, unlike H3K36me2, the establishment of H3K36me3 specifically requires Set2 association with the CTD of RNAPII, and H3K36me3 is positively correlated with the rate of transcription (Pokholok et al., 2005). In this report, we determine that H3K36me3 associates with Pdp3, thereby providing a novel effector protein for this mark in yeast. The finding that the HDAC Rpd3S is linked to H3K36me2 while the HAT complex NuA3b is linked to CTD-dependent H3K36me3, suggests differential methylation of H3K36 may act as a “chromatin switch” to regulate overall levels of transcription. For example, on lowly transcribed genes, the H3K36me2 state may predominate and recruit Rpd3S to maintain a more repressed chromatin environment suitable for low-level transcription. Conversely, on highly transcribed genes, H3K36me3 may predominate and recruit NuA3b to facilitate nucleosome disruption for RNAPII elongation. Alternatively, both H3K36me2 and me3 may be

coordinating HDAC and HAT activities together, within the same coding region, to properly modulate the progression of RNAPII during transcriptional elongation. Although speculative, these intriguing models await testing.

Finally, NuA3 joins a growing list of HAT complexes, which can localize to both promoters and gene bodies by modulating complex components (Ginsburg et al., 2009; Govind et al., 2007; Kremer and Gross, 2009; Kristjuhan and Svejstrup, 2004; Wyce et al., 2007). Whether NuA3b engages a dual H3K4me/H3K36me signature through the combinatorial action of the PHD finger of Yng1 and the PWWP domain of Pdp3, to fine tune transcription initiation and elongation pathways, remains an important question (Taverna et al., 2006) as mutations in chromatin effector proteins are linked to a wide array of epigenetic diseases (Baker et al., 2008; Butler et al., 2012). Relevant to this work, mutations of NuA3 human homologs (Doyon et al., 2006; Vezzoli et al., 2010) are associated with oral squamous cell carcinoma (Cengiz et al., 2010) and acute myeloid leukemia (Avvakumov and Cote, 2007). Accordingly, HATs and many other chromatin-bound protein complexes are gaining favor as pharmaceutical targets (Gunduz et al., 2008; Helin and Dhanak, 2013; Natoli, 2009; Taverna and Cole, 2010). Therefore, future studies of Pdp3 and NuA3 functions in yeast may elucidate how human chromatin effector proteins that are dysregulated in disease can be modulated by drugs targeting the epigenome.

Figure Legends

Figure 2.1: Pdp3/YLR455W is associated with the NuA3 complex, chromatin, and RNAPII

A. Proteins co-purifying with Pdp3-TAP bait that are involved in transcriptional regulation. Bulk MS/MS data are reported for proteins affinity purified by Pdp3 after searching with Mascot. B. Whole cell extracts (WCEs) from the indicated TAP-tagged yeast strains were immunoprecipitated with HIS6-FLAG-Pdp3 treated (+) and untreated (-) resin. WCEs (inputs) and immunoprecipitated samples (IPs) were resolved by SDS-PAGE. The presence of NuA3 complex members was monitored by western blotting.

Figure 2.2: The PWWP domain of Pdp3 is structurally conserved

Crystal structure of human BRPF1 (slate) bound to H3K36me3 peptide (purple) (top) (PDB ID: 2X4w) and the predicted structure of Pdp3 (brown) bound to H3K36me3 peptide (purple) in the same orientation (bottom). Note that Pdp3 is predicted to have similar placement of the three hydrophobic residues (yellow) in BRPF1 that are required for binding to H3K36me3. The Pdp3 structure was modeled on 2X4W using Phyre2, minimized with MOE, and rendered with PyMOL.

Figure 2.3: NuA3 specifically interacts with H3K36me3 through the PWWP domain of Pdp3

A. Clustal W alignment of PWWP domain-containing proteins. Beta sheets (arrows/underlined sequence) and alpha helices (cylinders/grey shading) are annotated. Aromatic cage residues are highlighted in yellow. B. Schematic representation of the

Pdp3 protein. The predicted (red outline) and functional (tan hexagon) PWWP domains are annotated. Aromatic cage residues F18, W21 and F48 are highlighted in yellow. C. Peptide pull-down assays were performed with full-length HIS6-FLAG-Pdp3 and biotinylated histone peptides. Purified protein (input) and immunoprecipitated samples (IPs) were resolved by SDS-PAGE. Direct binding was monitored by western blotting. D. Peptide pull-down assays were performed with full-length HIS6-FLAG-Pdp3, mutants F18A, W21A, and F48A, truncations $PWWP_{(1-74)}$, $PWWP_{(1-110)}$, $PWWP_{(2-150)}$, $PWWP_{(1-219)}$, and $C\text{-}term_{(74-304)}$, and biotinylated histone peptides. Purified proteins (inputs) and immunoprecipitated samples (me0 and 36me3) were resolved by SDS-PAGE. Direct binding was monitored by western blotting. E. Fluorescence polarization assays were used to measure binding affinities of full-length S·tag-Pdp3 and mutants F18A and W21A to the indicated 5-FAM labeled histone peptides. Error bars are the standard deviation from triplicate analyses of a representative experiment (n= 2).

Figure 2.4: Pdp3, but not Yng1, requires H3K36me3 for chromatin association *in vivo* A. TAP-tagged Pdp3 yeast strains and B. TAP-tagged Yng1 yeast strains were biochemically fractionated into chromatin-associated proteins and soluble proteins. Fractions were probed for the presence of Pdp3 and Yng1, respectively. G6PDH and H4 serve as both loading and fractionation controls.

Figure 2.5: Input levels for Pdp3 and Yng1 used in chromatin association assays

A. Whole cell extracts (WCEs) were made from 5 ODs of the indicated TAP-tagged Pdp3 or B. Yng1 yeast strains. WCEs were resolved by SDS-PAGE, transferred to PVDF membrane, and probed with the indicated antibodies.

Figure 2.6: NuA3 requires both Yng1 and Pdp3 to promote transcription

A. RT-qPCR analysis of *pdp3Δ* and *yng1Δ* cells to determine the relative transcript levels of NuA3-target genes and non-target genes (SSA1 and SSB1) as compared to WT. Transcript levels were normalized to Actin expression. Error bars represent the SEM of a representative experiment (n=4) done in triplicate. Asterisks indicate statistical significance as determined by an unpaired two-tailed t-test. * = $p < .05$, ** = $p < .01$, *** = $p < .001$. B. 2 ODs of the indicated yeast strains were 5-fold serially diluted on SC-Ura (left) or SC-Ura + 5-FOA (right) plates and grown at 30 ^{o}C for two or three days, respectively. C. 0.5 ODs of the indicated yeast strains were 5-fold serially diluted onto SC-Ura (left) or SC-Ura + 150 µg/mL 6-AU (right) plates and grown at 30 ^{o}C for two days.

Figure 2.7: The NuA3a complex, but not the NuA3b complex, is synthetically lethal with Gcn5

0.5 ODs of the indicated yeast strains were 5-fold serially diluted onto SC-Ura (top) or SC-Ura + 5-FOA (bottom) plates and grown at 30 ^{o}C for two days.

Figure 2.8: The NuA3 HAT complex has two functionally distinct forms that participate in transcription

A. Model of the NuA3a HAT complex. Yng1 binds to H3K4me3 through its PHD finger, thereby recruiting NuA3a to the promoter regions of actively transcribed genes. Sas3 then acetylates H3K14, leading to transcription initiation at a subset of genes (Howe et al., 2002; John et al., 2000; Martin et al., 2006a; Martin et al., 2006b; Taverna et al., 2006). B. Model of the NuA3b HAT complex. NuA3b contains a unique member, Pdp3. Pdp3 binds to H3K36me3 through its PWWP domain, thereby recruiting NuA3b to the coding regions of actively transcribed genes. Sas3 may then acetylate histones and provide a more open chromatin environment to facilitate the passage of RNAPII during transcription elongation.

Figure 2.1:

A

Protein	Spectral Count	Complex
H2B2	231	Histone
H4	215	Histone
H2A2	89	Histone
H3	42	Histone
Pdp3	145	NuA3
Yng1	84	NuA3
Nto1	60	NuA3
Sas3	59	NuA3
Taf14	43	NuA3
Eaf6	18	NuA3
Rpb4	19	RNAPII
Rpb2	13	RNAPII

B

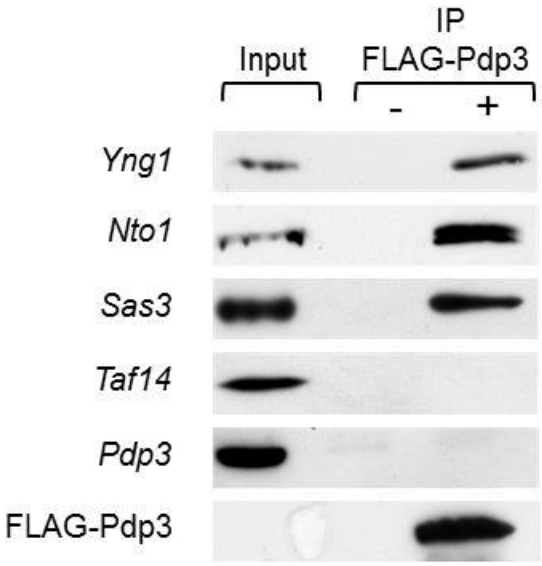


Figure 2.2:

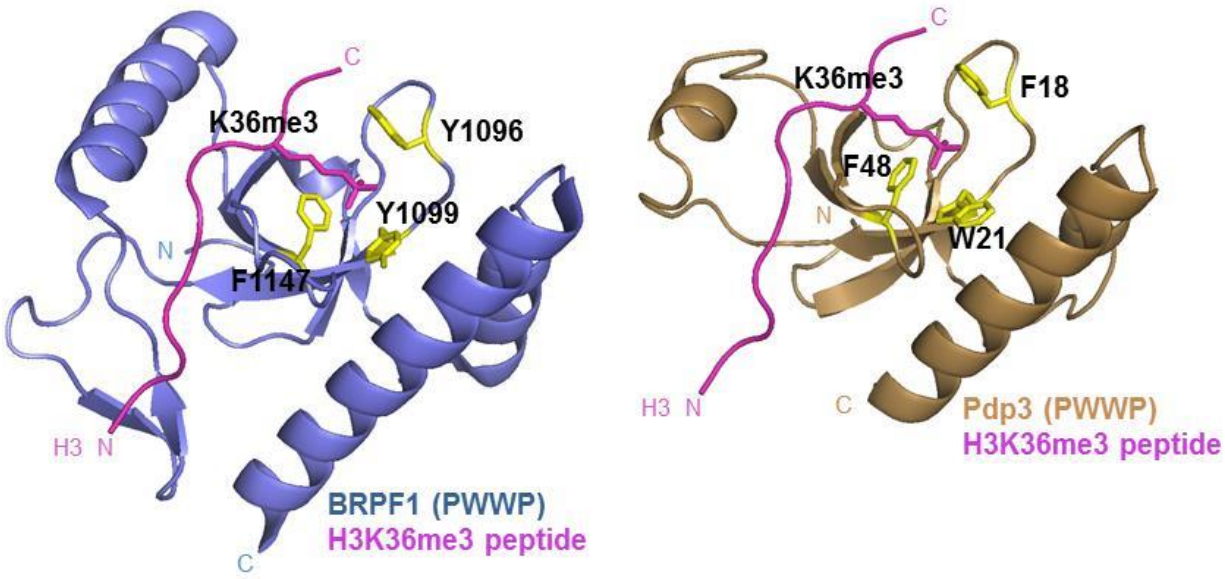

C
K36me3
Y1096
Y1099
N
F1147
H3 N
C
BRPF1 (PWWP)
H3K36me3 peptide
C
K36me3
F18
F48
N
W21
H3 N
C
Pdp3 (PWWP)
H3K36me3 peptide

Figure 2.3:

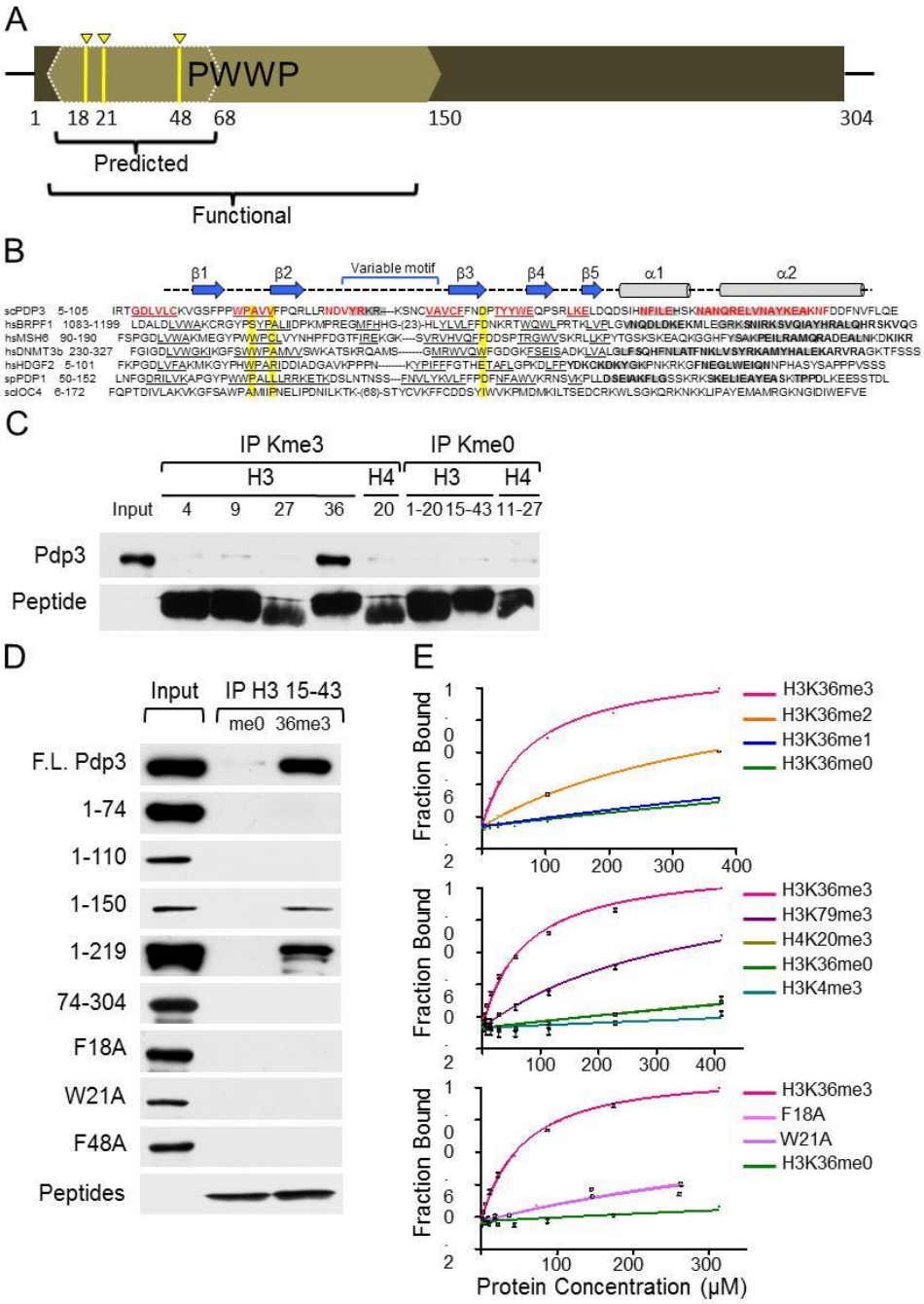

A
PWWP
1
18 21
48
68
150
304
Predicted
Functional
B
β1
β2
Variable motif
β3
β4
β5
α1
α2
scPDP3 5-105 IRTGDLVLCKVGSFPPWPAVVFPQRLLRNDVYRKR---KSNCVAVCFNDPTYYWEQPSRLKELDQDSIHNFILEHSKNANQRELVNAYKEAKNFDDFNVFLQE
hsBRPF1 1083-1199 LDALDLVWAKCRGYPSYPALIIDPKMPREGMFHHG-(23)-HLYLVLFFDNKRTWQWLPRTKLVPLGVNQDLDKEKMLEGRKSNIRKSVQIAYHRALQHRSKVQG
hsMSH6 90-190 FSPGDLVWAKMEGYPWWPCLVYNHPFDGTFIREKGK----SVRVHVQFFDDSPTRGWVSKRLLKPYTGSKSKEAQKGGHFYSAKPEILRAMQRADEALNKDKIKR
hsDNMT3b 230-327 FGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMS-------GMRWVQWFGDGKFSEISADKLVALGLFSQHFNLATFNKLVSYRKAMYHALEKARVRAGKTFSSS
hsHDGF2 5-101 FKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPN--------KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSS
spPDP1 50-152 LNFGDRILVKAPGYPWWPALLLRRKETKDSLNTNSS---FNVLYKVLFFPDFNFAWVKRNSVKPLLDSEIAKFLGSSKRKSKELIEAYEEASKTPPDLKEESSTDL
scIOC4 6-172 FQPTDIVLAKVKGFSAWPAMIIPNELIPDNILKTK-(68)-STYCVKFFCDDSYIWVKPMDMKILTSEDCRKWLSGKQRKNKKLIPAYEMAMRGKNGIDIWEFVE
C
IP Kme3
IP Kme0
H3
H4
H3
H4
Input
4
9
27
36
20
1-20
15-43
11-27
Pdp3
Peptide
D
Input
IP H3 15-43
me0
36me3
F.L. Pdp3
1-74
1-110
1-150
1-219
74-304
F18A
W21A
F48A
Peptides
E
Fraction Bound
H3K36me3
H3K36me2
H3K36me1
H3K36me0
H3K79me3
H4K20me3
H3K4me3
F18A
W21A
Protein Concentration (μM)

Figure 2.4:

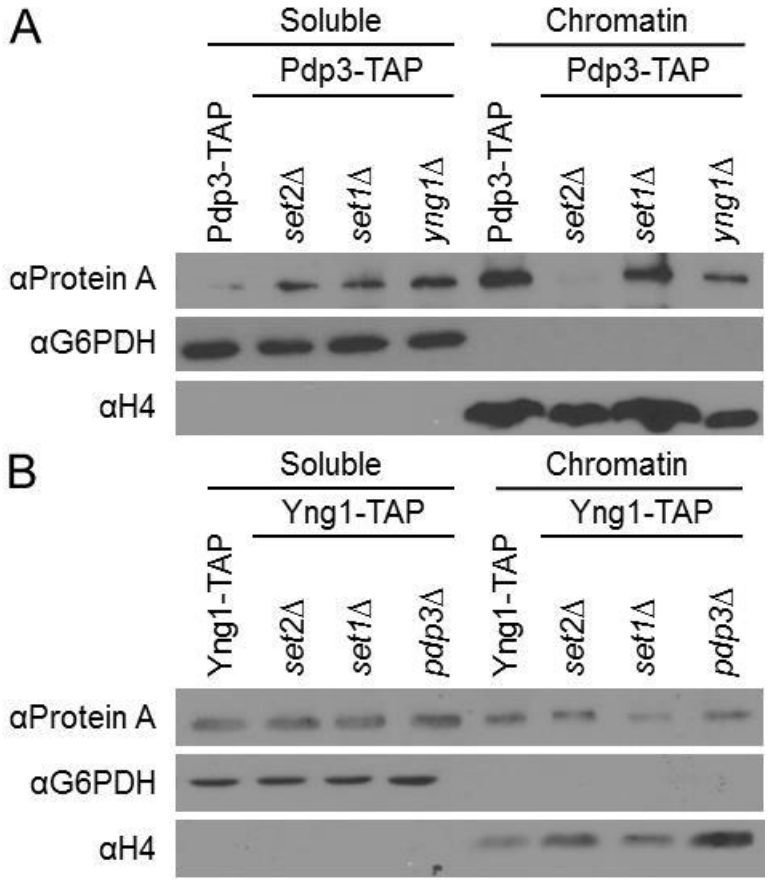

A
Soluble
Chromatin
Pdp3-TAP
Pdp3-TAP
Pdp3-TAP
set2Δ
set1Δ
yng1Δ
Pdp3-TAP
set2Δ
set1Δ
yng1Δ
αProtein A
αG6PDH
αH4
B
Soluble
Chromatin
Yng1-TAP
Yng1-TAP
Yng1-TAP
set2Δ
set1Δ
pdp3Δ
Yng1-TAP
set2Δ
set1Δ
pdp3Δ
αProtein A
αG6PDH
αH4

Figure 2.5:

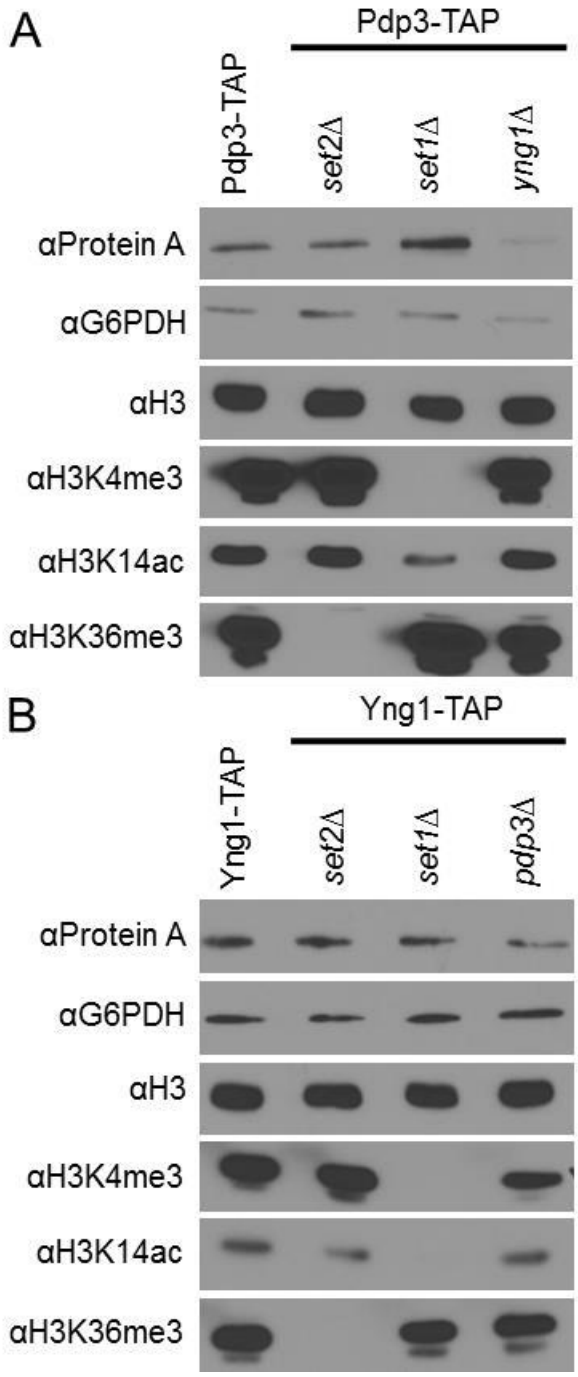

A
Pdp3-TAP
Pdp3-TAP
set2Δ
set1Δ
yng1Δ
αProtein A
αG6PDH
αH3
αH3K4me3
αH3K14ac
αH3K36me3
B
Yng1-TAP
Yng1-TAP
set2Δ
set1Δ
pdp3Δ
αProtein A
αG6PDH
αH3
αH3K4me3
αH3K14ac
αH3K36me3

Figure 2.6:

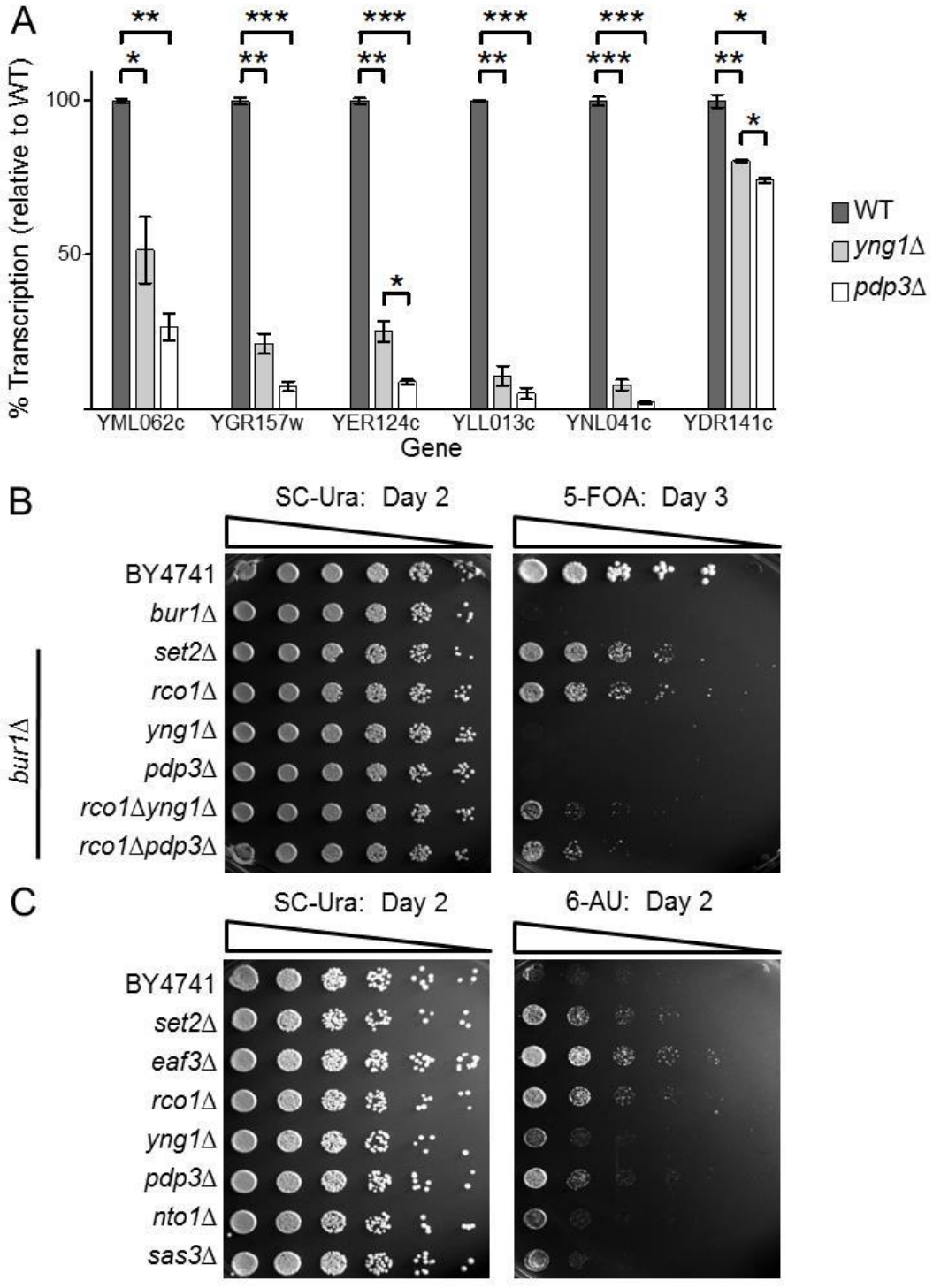

A
% Transcription (relative to WT)
100
50
YML062c
YGR157w
YER124c
YLL013c
YNL041c
YDR141c
Gene
WT
yng1Δ
pdp3Δ
B
SC-Ura: Day 2
5-FOA: Day 3
BY4741
bur1Δ
set2Δ
rco1Δ
yng1Δ
pdp3Δ
rco1Δyng1Δ
rco1Δpdp3Δ
bur1Δ
C
SC-Ura: Day 2
6-AU: Day 2
BY4741
set2Δ
eaf3Δ
rco1Δ
yng1Δ
pdp3Δ
nto1Δ
sas3Δ

Figure 2.7:

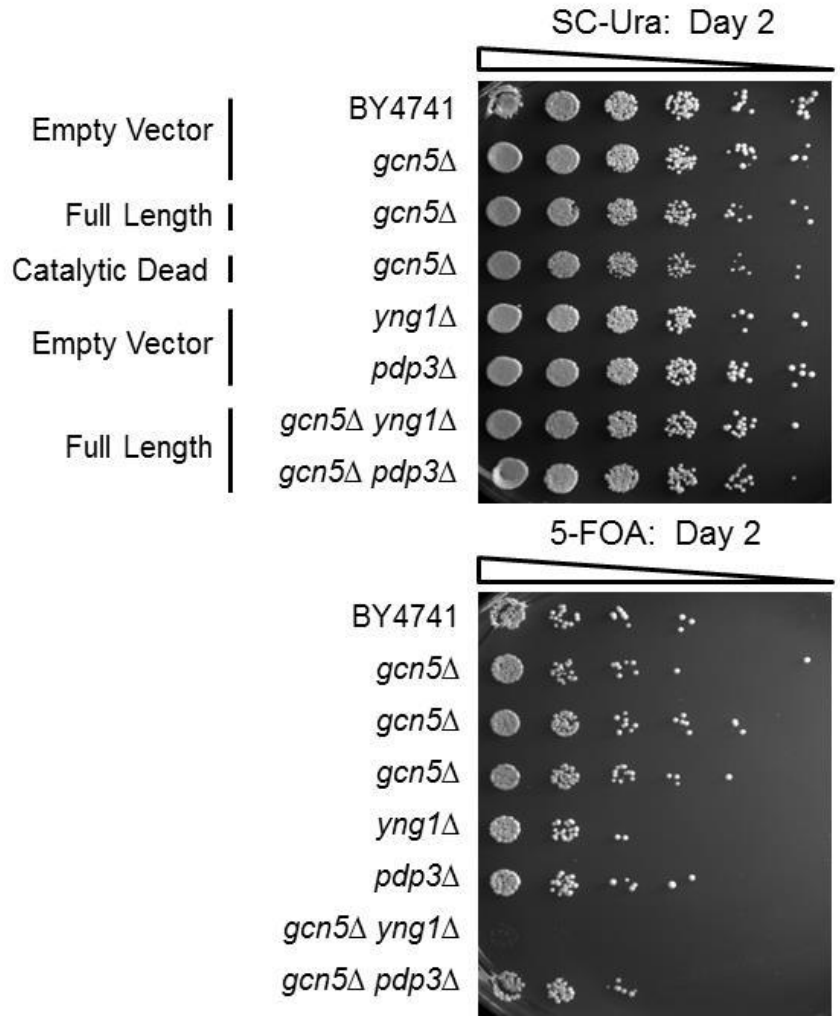
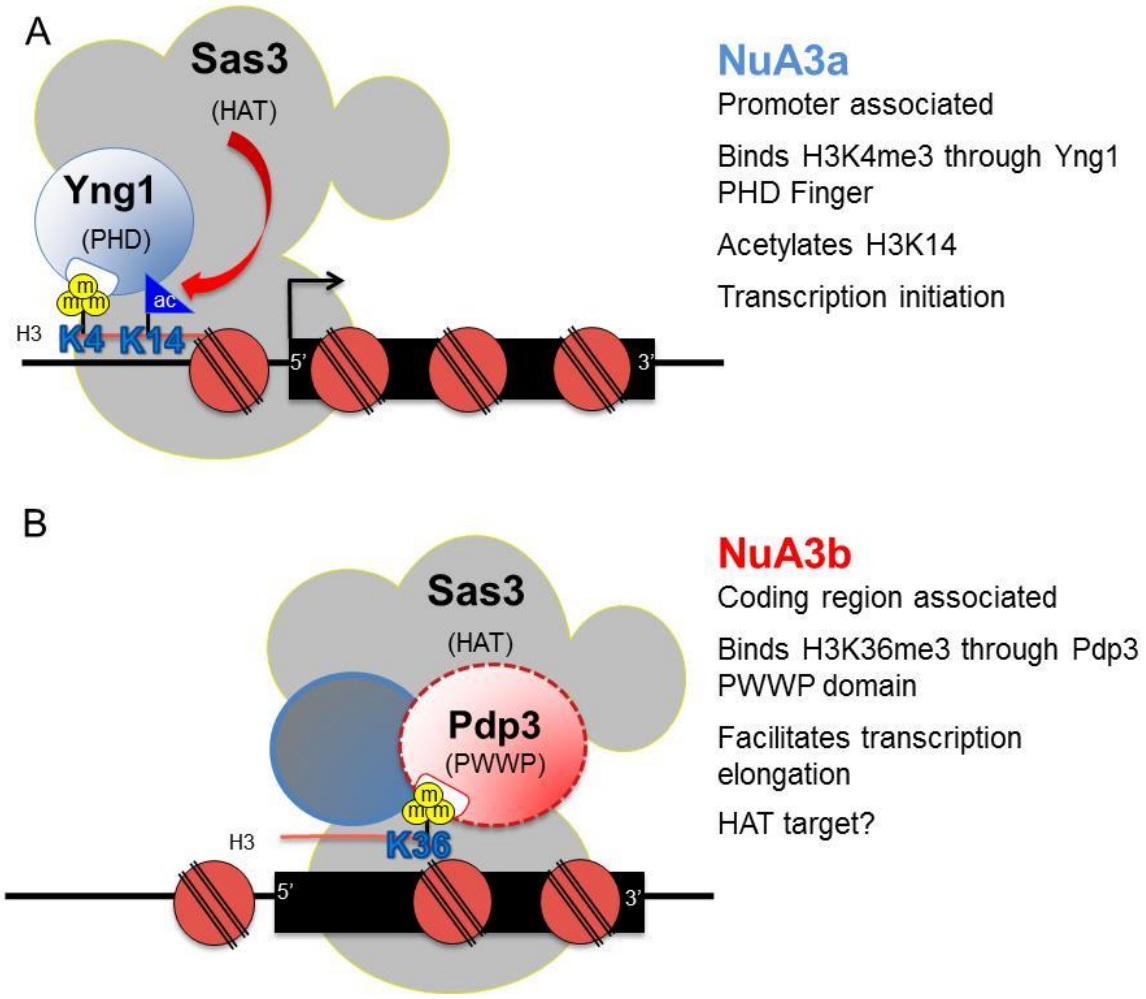

SC-Ura: Day 2
Empty Vector
BY4741
gcn5Δ
Full Length
gcn5Δ
Catalytic Dead
gcn5Δ
Empty Vector
yng1Δ
pdp3Δ
Full Length
gcn5Δ yng1Δ
gcn5Δ pdp3Δ
5-FOA: Day 2
BY4741
gcn5Δ
gcn5Δ
gcn5Δ
yng1Δ
pdp3Δ
gcn5Δ yng1Δ
gcn5Δ pdp3Δ

Figure 2.8:


A
Sas3
(HAT)
Yng1
(PHD)
m
m m
ac
H3
K4 K14
5'
3'
NuA3a
Promoter associated
Binds H3K4me3 through Yng1 PHD Finger
Acetylates H3K14
Transcription initiation
B
Sas3
(HAT)
Pdp3
(PWWP)
m
m m
H3
K36
5'
3'
NuA3b
Coding region associated
Binds H3K36me3 through Pdp3 PWWP domain
Facilitates transcription elongation
HAT target?

CHAPTER 3 – COMBINATORIAL HISTONE READOUT BY THE DUAL PLANT HOMEODOMAIN (PHD) FINGERS OF RCO1 MEDIATES RPD3S CHROMATIN RECRUITMENT AND THE MAINTENANCE OF TRANSCRIPTIONAL FIDELITY[2]

Introduction

Post-translational modifications (PTMs) on histone proteins play a critical role in many DNA-templated processes, particularly the control of gene transcription. Complexes that modify and remodel chromatin to regulate proper transcription contain proteins with conserved recognition domains that bind either modified or unmodified residues within histone proteins (Lalonde et al., 2014; Musselman et al., 2012; Yun et al., 2011). Because these PTMs are dynamically regulated and are targeted to specific locations across the open reading frame of genes, effector proteins/complexes that read these PTMs can be recruited in a spatio-temporal manner to control chromatin structure and RNA polymerase II (RNAPII) elongation during transcription (Bannister and Kouzarides, 2011). For example, histones are hyperacetylated in front of elongating

[2] Portions of this chapter were adapted from McDaniel, S.L., Fligor, J.E., Ruan, C., Cui, H., Bridgers, J.B., DiFiore, J.V., Guo, A.H., Li, B., and Strahl, B.D. (2016). Combinatorial Histone Readout by the Dual Plant Homeodomain (PHD) Fingers of Rco1 Mediates Rpd3S Chromatin Recruitment and the Maintenance of Transcriptional Fidelity. The Journal of biological chemistry *291*, 14796-14802.

RNAPII allowing for nucleosome disassociation and are hypoacetylated behind RNAPII to maintain chromatin structure and prevent inappropriate cryptic transcription (Wagner and Carpenter, 2012). The deacetylation of nucleosomes in transcription is carried out by histone deacetylase complexes (HDAC), which typically have one or more reader domains that are able to engage chromatin. This multi-domain structure allows for recognition of increasingly complex and specific chromatin environments.

Rpd3S, an HDAC that functions in a co-transcriptional manner, has five conserved chromatin-binding domains: a chromodomain in Eaf3, which recognizes Set2-mediated histone H3 lysine 36 methylation (H3K36me) (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005), and four plant homeo-domains (PHDs), two per copy of Rco1, which has recently been shown to form a homo-dimer in Rpd3S (Figure 1A) (Ruan et al., 2016a). The chromodomain of Eaf3 and the N-terminal PHD finger of Rco1 (PHD1) have previously been characterized and are necessary for Rpd3S function and nucleosome engagement (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005; Li et al., 2007a). PHD1 is thought to engage H3 on one nucleosome while the chromodomain of Eaf3 recognizes H3K36me on a neighboring nucleosome, allosterically activating the deacetylase activity of Rpd3 (Lee et al., 2013; Ruan et al., 2015). This activity is necessary to enforce chromatin integrity and transcriptional fidelity across the transcribed regions of genes, thereby preventing the formation of pervasive cryptic unstable transcripts (CUTs) and stable untranslated transcripts (SUTs) (Carrozza et al., 2005; Churchman and Weissman, 2011; Li et al., 2007a; Lickwar et al., 2009). It has been recently shown that the Set2/Rpd3S pathway is particularly important for repressing antisense transcription from divergent promoters

(Churchman and Weissman, 2011). While significant advances have been made in understanding the role of Rpd3S in cells, the precise mechanism by which Rpd3S is targeted to chromatin and mediates its function is still poorly understood.

In this report, we interrogated the role of the C-terminal PHD finger in Rco1 (PHD2), which had not been previously investigated. We show that PHD2 is a functional domain and recognizes the unmodified N-terminus of H3, as does PHD1. Further, mutational analysis shows that nucleosome binding *in vitro* and chromatin association of Rpd3S *in vivo* depend on the function of both PHD fingers. Consistent with this finding, we demonstrate that mutation of either PHD1 or PHD2 leads to chromatin and transcriptional fidelity defects. Together, our data unveils a critical role for two adjacent PHD fingers in coordinating Rpd3S recruitment and function.

Materials and Methods

Mutagenesis

Conserved residues were identified and specifically mutated using the Site Directed Mutagenesis kit (Stratagene) and confirmed via Sanger sequencing.

Immunoblot

A single colony was inoculated overnight to saturation and then diluted to an OD_{600}=0.2 and grown to mid-log phase. Five ODs of cells were isolated and lysed via bead beating in SUTEB (1% SDS, 8 M urea, 10 mM Tris pH 6.8, 10 mM EDTA, 0.01% bromphenol blue) for 3 minutes. Lysates were boiled for 10 minutes and then isolated. Cell debris were removed via centrifugation and the supernatant was isolated. Cleared

lysates were loaded onto 8% SDS-PAGE gels and then transferred to PVDF membrane. Membranes were probed overnight (4^{o}C) with anti-HA (UNC Antibody Core) or anti-G6PDH (Sigma). Immunoblots were visualized using HRP-conjugated secondary antibodies and ECL Prime solution (GE Healthcare).

Alignment and Molecular Modeling

Yeast PHD fingers were isolated from the SMART database and aligned using the Espript 3.0 tool (Gouet et al., 1999). PHD1 and PHD2 sequences were then modeled using the HHpred tool (Soding et al., 2005) and visualized in Pymol.

Purification of GST-tagged PHD fingers

The PHD fingers of Rco1 were purified from SOLUBL21 competent E. coli cells. Cells were grown in the presence of 1mM zinc. Bacteria pellets were lysed in 50mM Tris-HCl, pH 7.4, 130 mM NaCl, 1mM DTT, 1uM $ZnCl_2$, 1mM PMSF, Universal Nuclease for Cell Lysis (Pierce, 1:20,000), 1 Roche Protease Inhibitor Cocktail tablet/50mL, 1 mg/mL lysozyme (Sigma), 0.1% Triton. Cleared lysate was incubated with Pierce glutathione resin for two hours at 4°C. Protein was eluted from the resin in 50mM Tris-HCl, pH 8.0, 130 mM NaCl, 10mM glutathione, 1mM DTT, and 1 uM $ZnCl_2$. Protein was then dialyzed overnight into a storage buffer composed of 50mM Tris-HCl, pH 7.4, 130 mM NaCl, 1mM DTT, and 1uM $ZnCl_2$.

Overnight Peptide Pull-down

1 µg of biotinylated histone peptide (Table S3) was incubated with 1 µg of purified GST-fusion protein in 1mL of Peptide Binding Buffer (PBB, 50mM Tris-HCl, pH 7.5, 130 mM NaCl, 0.1% (v/v) NP-40, 1mM phenylmethylsulphonyl fluoride (PMSF), 1mM DTT, 1uM $ZnCl_2$, and 1 Roche Protease Inhibitor Cocktail tablet/100mL of PBB) overnight at 4°C. After incubation for 1 hour at 4°C with streptavidin beads (Pierce), the beads were washed three times with 1 mL of PBB. Peptide was then eluted from the bead into SDS-loading buffer by boiling for 5 minutes at 98°C. Samples were then subjected to western blot analysis and the membrane was probed with GST antibody (Sigma, 1:4000) for 1hr at room temperature. Peptide loading was assessed by probing the membrane with Streptavidin-HRP (Cell Signaling, 1:5000) for 30 minutes at room temperature.

Rpd3 Complex Isolation

Recombinant Rpd3S complexes were purified from a Sf21 insect cell-based baculovirus expression system as described previously (Ruan et al., 2016a; Ruan et al., 2015). Briefly, freshly passed Sf21 cells were co-infected with individual virus that encodes each subunit of Rpd3S for 48 hrs. Cells were collected and lysed in BV lysis buffer (50 mM HEPES pH7.9; 300 mM NaCl; 2 mM $MgCl_2$; 0.2% Triton X-100; 10% glycerol; 0.5 mM EDTA and freshly added protease inhibitors) on ice for 30 min. Cell lysates were clarified by ultra-centrifugation and incubated with anti-FLAG M2 resin (Sigma) at 4°C for 2 hr. After extensive washing, each complexes were eluted using 500

µg/ml 3xFlag peptides in BV elution buffer (50 mM HEPES pH7.9; 100 mM NaCl; 2 mM $MgCl_2$; 0.02% NP40 and 10% glycerol) and concentrated using Amicon concentrators.

EMSA

Mono-nucleosomes were reconstituted using a 222bp 601-positioning sequence containing DNA template and purified as described previously (Huh et al., 2012; Yun et al., 2012). EMSA reactions were carried out in a 15 µl system containing 10 mM HEPES pH7.8, 50 mM KCl, 4 mM $MgCl_2$, 5 mM DTT, 0.25 mg/ml BSA, 5% glycerol and 0.1 mM PMSF. The samples were incubated at 30°C for 45 min and run on a 3.5% acrylamide (37.5:1, acrylamide:bis-acrylamide) gel at 4°C.

Chromatin Association Assay

Fifty ODs of mid-log phase cells were isolated and fractionated as previously described (Gilbert et al., 2014). Fractions were immunoblotted and probed with anti-HA (UNC Antibody Core), anti-G6PDH (Sigma), or anti-H4 (Millipore).

Rco1 Co-Immunoprecipitation

All strains were grown overnight in SC-Leu. Cultures were diluted to an OD_{600} of 0.2 and grown to log phase in 100 mL of SC-Leu. Cells were pelleted and washed with 50 mL of dH_2O. The pellets were resuspended in 500 µl of Lysis Buffer (Keogh et al., 2006) and split equally into two tubes. Glass beads were added to bring the total volume to 750 µl and samples were vortexed for 12 min, rested for 10 min on ice for a total of two times at 4C. Lysates were collected into fresh tubes via centrifugation and

the lysates were cleared at max speed for 15 min at 4C. Protein concentration was quantified via Bradford assay. An aliquot was taken for input and 1.5 mg/mL of protein was incubated overnight in 1 mL of Lysis Buffer at 4C with 1:1000 dilution of Protein A antibody (Sigma). Antibody was conjugated to IgG Sepharose beads (GE Healthcare) for 2 hours at 4C before being washed with Lysis Buffer and protein eluted with 100 µl of 5x SDS buffer. Samples were boiled at 95C for 5 min before loaded onto an 8% SDS-PAGE gel.

Spotting Assays

All spotting assays were performed with 5-fold serial dilutions of saturated overnight cultures of the indicated strains. Growth was assayed after 2-5 days. All yeast strains used in this study are described in Table S1 and all plasmids are described in Table S2.

Results

Rco1 contains two PHD fingers that bind to the N-terminus of H3

Rco1 is a unique member of the Rpd3S complex, which is defined by a N-terminal PHD finger followed by an auto-inhibitory domain (AI), a Sin3 interaction domain (SID), which associates with the MRG domain of Eaf3 (Ruan et al., 2015), and a second C-terminal PHD finger (Figure 3.1A). Although the first PHD finger is required for nucleosome binding and Rpd3S function (Li et al., 2007a), its second PHD finger (PHD2) remained uncharacterized. To explore whether PHD2 would encode a functional domain, we first performed a sequence alignment of the known PHD fingers

from *S. cerevisiae* (Figure 3.1B) and created structure prediction models in HHpred of PHD1 and PHD2 using a PHD domain from CHD4 (PDB ID: 1MM2) as a scaffold (Figure 3.1C, 3.1D, and 3.1E) (Kwan et al., 2003). Our sequence alignments showed that both PHD fingers of Rco1 contain the necessary conserved cysteine and histidine residues needed for the coordination of two zinc ions – a feature that defines functional PHD fingers. In addition, structural modeling predicted PHD2 as being a folded domain with high structural similarities to PHD1. Interestingly, PHD2 contains a small sequence insertion between the first and second grouping of cysteines and has substituted a conserved cysteine residue for an additional histidine at its C-terminus, thus suggesting PHD2 slightly differs from other PHD fingers.

We next interrogated the ability of both PHD1 and PHD2 to directly associate with histones. Each domain was expressed and purified as a GST fusion and assayed in solution peptide pull-downs for their ability to bind differentially modified biotinylated histone peptides from distinct regions of the H3 N-terminus. As shown in Figure 3.1F, we consistently found that both PHD1 and PHD2 preferentially bound to the N-terminal region (residues 1-20) of the H3 tail. This result is consistent with previous analyses of PHD1 (Kumar et al., 2012; Shi et al., 2007). Interestingly, we also found that tri-methlyation of K4 (H3K4me3) decreases the ability of both PHD1 and PHD2 to bind N-terminal $H3_{1-20}$ peptides, suggesting that these domains bind to the extreme N-terminus of H3 and are affected by N-terminal PTMs. Furthermore, this finding may account, at least in part, for how Rpd3S is restricted from binding to promoter nucleosomes normally marked with H3K4me3.

PHD2 is required for Rpd3S association on chromatin *in vitro*

Our previous studies demonstrated that PHD1 was essential in mediating Rpd3S association on nucleosomes (Li et al., 2007a). Given this, we wondered what the contribution would be, if any, for Rco1's second PHD finger in nucleosome binding or Rco1's homo-dimerization. To ascertain this, we recombinantly expressed the five Rpd3S members and assembled *in vitro* complexes competent for nucleosome binding (Figure 3.2A). In addition to a complete deletion of PHD2 (*rco1*Δ *phd2*), we made a point mutant in PHD2 predicted to disrupt zinc binding and PHD function (*rco1-C417A*). As shown in Figure 3.2A, both the deletion of PHD2 and the C417A point mutation had no effect on Rpd3S complex assembly. Furthermore, mutation of PHD2 showed no defects in Rco1 homo-dimerization by co-immunoprecipitation analysis (Figure 3.3). Surprisingly, even though the integrity of Rpd3S complexes with PHD2 mutants was fully intact, the ability of these complexes to bind nucleosomes was completely abolished (Figure 3.2B) – a result that is identical to the loss of PHD1 (Li et al., 2007a). These results imply that both PHD fingers of Rco1 function in a coordinated fashion to bind nucleosomes.

PHD1 and PHD2 are required for chromatin association *in vivo*

To determine the significance of PHD1 and PHD2 in Rpd3S function in cells, we generated a panel of mutations at conserved residues found in both PHD1 and PHD2 that we predicted to be critical for their function (C275, D276 and H283 in PHD1 and C417, M438, C440, and D441 in PHD2). As shown in Figure 3.2C, nearly wild-type levels of protein were obtained for all of the mutants made in PHD2, but two mutants in

PHD1 proved to be unstable (C275A and H283A), and therefore were not used further in our analyses. We next assessed the ability of these mutants to affect the association of Rpd3S on chromatin *in vivo* using chromatin-association assays. Yeast cells expressing wild-type or mutated versions of Rco1 were fractionated into soluble or chromatin-associated fractions (Figure 3.2D). As expected, wild-type Rco1 was predominantly found in the chromatin fraction. In stark contrast, however, the D276A PHD1 mutant and C417A and C440A PHD2 mutants were unable to maintain association with chromatin. Hence, PHD1 and PHD2 are both required for chromatin association of Rpd3S *in vitro* and *in vivo*.

Loss of PHD1 or PHD2 function in Rco1 leads to chromatin structure and transcriptional fidelity defects

One of the key functions of Rpd3S is to restore chromatin to a hypo-acetylated state after the passage of RNAPII during gene transcription (Wagner and Carpenter, 2012). This locally compacts chromatin structure, thereby preventing bi-directional transcription and RNAPII complexes from binding to cryptic promoter elements along the gene and aberrantly initiating transcription. To monitor cryptic transcription, we employed a yeast strain wherein the *HIS3* gene is fused to a naturally occurring cryptic promoter in the *FLO8* gene (Silva et al., 2012). Importantly, *HIS3* is out of frame with the normal 5' promoter and only produces a functional transcript if the 3' cryptic promoter is used (see schematic in Figure 3.4A). Under growth conditions using media lacking histidine, cells will not grow if chromatin structure is normal. As expected, strains deleted for *RCO1* resulted in a growth phenotype on plates lacking histidine (Figure

3.4B), signifying a significant disruption to the chromatin structure at this locus, resulting in cryptic transcription occurring at the internal promoter. This phenotype is rescued by the addition of wild-type *RCO1*. In complete contrast, however, mutations of PHD1 and PHD2 that disrupt Rpd3S association to chromatin also result in a cryptic transcription phenotype (Figure 3.4B). Comparison of another reporter gene, *STE11*, whose natural cryptic promoter is also fused to the *HIS3* gene showed identical results, thus verifying that the cryptic transcription defect we were observing is not a gene specific effect (Figure 3.5). Together, these results show that without PHD1 or PHD2 function, Rpd3S is unable to engage chromatin and properly regulate chromatin structure during gene transcription.

In addition to cryptic transcription, another assay that has been used for the analysis of chromatin and transcription defects is the *bur1*Δ bypass assay. *BUR1* is an essential kinase that acts positively on transcription by phosphorylating several members of the elongating RNAPII including the C-terminal domain of Rpb1 and the C-terminal repeat domain of Spt5 (Chu et al., 2006; Keogh et al., 2003; Liu et al., 2009). However, in the absence of Set2 and other factors in the *SET2* genetic pathway (e.g., Rpd3S), cells lacking *BUR1* are viable. As expected, loss of *RCO1* resulted in a bypass of lethality that was rescued upon restoring wild-type *RCO1*. Consistent with the role of both PHD1 and PHD2 in Rpd3S function, we observed that mutation of either domain renders cells resistant to the loss of *BUR1* (Figure 3.4C). Together with the cryptic initiation assay, these data show that combinatorial engagement of histone H3 by the PHD1 and PHD2 is critical for Rpd3S chromatin recruitment and function during gene transcription.

Discussion

Based on the work herein and other recent publications, we propose that Rco1 is a critical scaffolding protein that engages the N-termini of histone H3 to stabilize Rpd3S on chromatin, thereby optimally positioning Eaf3 for H3K36me2/me3 binding. With two copies of Rco1 per Rpd3S complex, it is likely that one Rpd3S complex is engaged on two adjacent nucleosomes whereby all four H3 tails are co-occupied to maintain Rpd3S stability on chromatin (see model in Figure 3.4D). Future studies will be needed to resolve whether the PHD domains from the same molecule of Rco1 bind both H3 tails in a single nucleosome, or if they bind a separate H3 tail from two neighboring nucleosomes. Regardless, having Rco1 co-occupy two nucleosomes would further highlight the need and role of Isw1b, which utilizes its ATP-dependent chromatin-remodeling activities to position adjacent nucleosomes in close proximity for Rpd3S binding (Smolle et al., 2012; Venkatesh et al., 2012). Together, these events function to allow Rpd3 to deacetylate histones and maintain chromatin integrity during the transcription process.

Our studies showed that both PHD fingers of Rco1 have a similar preference for binding to the extreme N-terminus of H3, and further, that this binding is highly sensitive to H3K4 tri-methylation. This result may help to provide an explanation for how Rpd3L and Rpd3S binding to discrete regions along the gene are controlled. Rpd3L, which localizes to promoters, does not contain Rco1 but rather two other PHD-containing proteins (Cti6 and Pho23) specific for H3K4me3 (Shi et al., 2007; Wang et al., 2011). This would help to maintain Rpd3L in promoter regions where H3K4me3 is restricted.

In contrast, Rpd3S, which lacks these other PHD-containing proteins for Rco1, is repelled by H3K4me3 thereby restricting this complex to gene bodies. Consequently, this would localize the Rpd3S complex in regions with high levels of H3K36me, which is then recognized by the chromodomain of Eaf3 (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). Thus, the different PHD fingers found in Rpd3L and Rpd3S likely govern their discrete localizations at genes (see model in Figure 3.4D).

Finally, we note that while a previous survey of yeast PHD domains was performed using solution peptide pull-downs (Shi et al., 2007), our studies differ in regards to the ability of PHD1 and PHD2 to bind H3K36me3 – a result that is also true for the characterized PHD1 domain of the human Rco1 counterpart, Pf1 (Kumar et al., 2012). We note that PHD1 and PHD2 expression and maintaining their stability *in vitro* was found to be extremely difficult, and further, that binding and washing conditions greatly impacted weak interactions and non-specific binding. These challenges with studying Rco1 may explain how different observations were observed.

Our results showed that both PHD fingers in Rco1 are required for chromatin targeting and Rpd3S function. This finding strongly argues that Rpd3S is targeted to chromatin via combinatorial readout of H3. We propose that in isolation, PHD1 or PHD2 is not robust enough to maintain stable Rpd3S association to nucleosomes, but rather, is reinforced when both domains operate in unison, and further, when this occurs as a homo-dimer capable of binding four H3 tails. These four histone binding events likely enforce a strict reading of the chromatin environment that requires the appropriate spacing of the di-nucleosome substrate imposed by the Isw1b chromatin remodeling complex.

Finally, we show that loss of H3 binding by either PHD finger of Rco1 results in a disruption of chromatin structure that leads to cryptic transcription and transcriptional defects. Given the significant role these PHD fingers play, it will be interesting to explore whether other chromatin-associated proteins/complexes with multiple PHD fingers behave similarly. In regards to Rpd3L, we predict that this complex's association to promoters will require each of the two PHD domains found in the complex that show binding to H3K4me3 (speculated in Figure 3.4D). Taken together, our results with Rco1 highlight the significance of combinatorial readout in chromatin function and provide further support for the 'histone code' hypothesis (Strahl and Allis, 2000).

Figure Legends

Figure 3.1: The PHD fingers of Rco1 bind to the extreme N-terminus of H3

(A) A schematic representation of the Rco1 protein. PHD fingers are highlighted in green with the AI and SID domains in blue and gray, respectively. *(B)* An alignment of yeast PHD fingers, highlighting the conserved cysteine and histidine residues. *(C)* A molecular model of PHD1. Zinc atoms are in gray. *(D)* A molecular model of PHD2. Zinc atoms are in gray. *(E)* A merge of the models of PHD1 and PHD2 show a high degree of similarity. *(F)* In-solution peptide pull-down assays with PHD1 and PHD2 were carried out with the indicated histone H3 peptides. Both domains bind the unmodified H3 N-terminus and show sensitivity to H3K4me3.

Figure 3.2: Both PHD fingers of Rco1 are necessary for association with chromatin

(A) Coomassie Blue stained gels showing that the Rpd3S complex can be purified intact with Rco1 that is WT, C417A, or lacking PHD2. *(B)* An EMSA assay showing that the PHD2 finger of Rco1 is necessary for Rpd3S association with nucleosomes. *(C)* A western blot of the indicated strains to examine the stability of the mutated *RCO1* constructs. All mutations except C275A and H283A were able to sustain near wild-type levels of Rco1 protein. *(D)* Western blot analysis of soluble and chromatin fractions from the indicated strains. Mutation of conserved residues in either PHD finger renders Rco1 unable to associate with chromatin *in vivo*.

Figure 3.3: Rco1 dimerization is independent of its PHD fingers

(A) Co-immunoprecipitation of endogenous Rco1-TAP with either wild-type or PHD mutant Rco1-HA constructs. Mutation of either PHD finger does not affect Rco1 dimerization.

Figure 3.4: Chromatin structure and transcriptional fidelity requires both PHD fingers in Rco1

(A) A schematic of the *FLO8-HIS3* fusion gene reporter to detect changes in chromatin structure and cryptic transcription. Cells will only grow in the absence of histidine and if the cryptic promoter in front of *HIS3* is utilized. *(B)* Cryptic initiation spotting assay using the *FLO8-HIS3* fusion gene reporter. The indicated strains are spotted in a 5-fold serial dilution from a starting OD_{600} of 0.5. Plates are imaged after 2-5 days. *(C)* A *BUR1* bypass assay reveals transcriptional elongation defects in Rco1 PHD mutations. The indicated strains are spotted in a 5-fold serial dilution from a starting OD_{600} of 2.0. Plates are imaged after 2-3 days. *(D)* A model of how Rpd3S and Rpd3L engage chromatin. Each PHD finger of Rco1 engages the N-terminus of H3 while the chromodomain of Eaf3 recognizes methylated H3K36, which allosterically activates the HDAC activity of Rpd3. The selectivity of each PHD finger for unmodified H3K4 (H3K4me0) may be an important contributing factor in Rpd3S localization to gene bodies, which are not marked with H3K4me3. In contrast, Rpd3L contains two PHD fingers (one each in Cti6 and Pho23) that do bind to H3K4me3, thus enabling this complex to maintain its localization to promoter regions and not gene bodies. These

findings help to further our understanding of combinatorial readout in recruitment of chromatin-associated proteins.

Figure 3.5: Chromatin structure is aberrant in PHD2 mutants

Cryptic initiation spotting assay using the *STE11-HIS3* fusion gene reporter. The indicated strains are spotted in a 5-fold serial dilution from saturated culture. Plates are imaged after 2-5 days.

Figure 3.1:

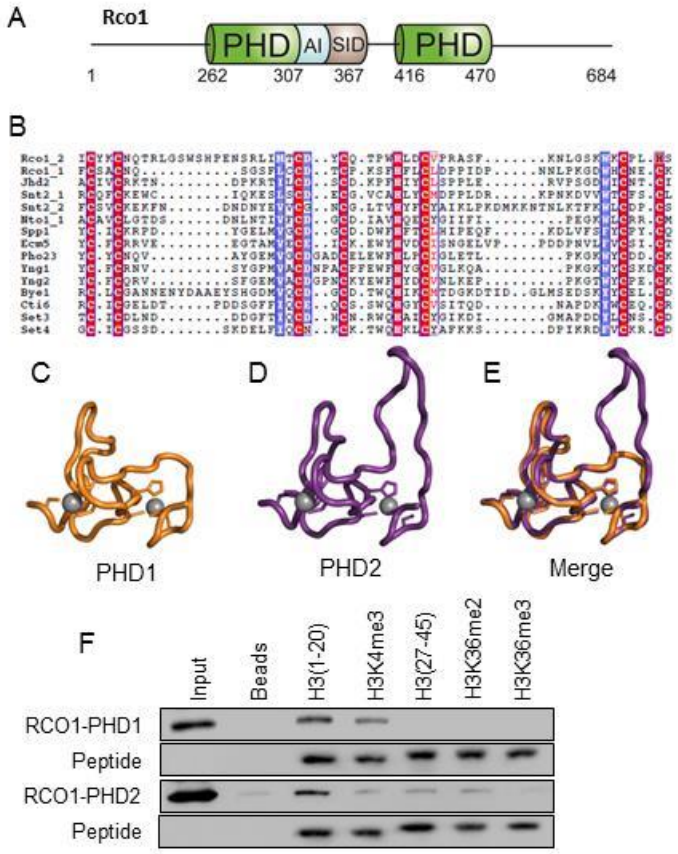

A
Rco1
PHD
AI
SID
PHD
1
262
307
367
416
470
684
B
C
D
E
PHD1
PHD2
Merge
F
Input
Beads
H3(1-20)
H3K4me3
H3(27-45)
H3K36me2
H3K36me3
RCO1-PHD1
Peptide
RCO1-PHD2
Peptide

Figure 3.2:

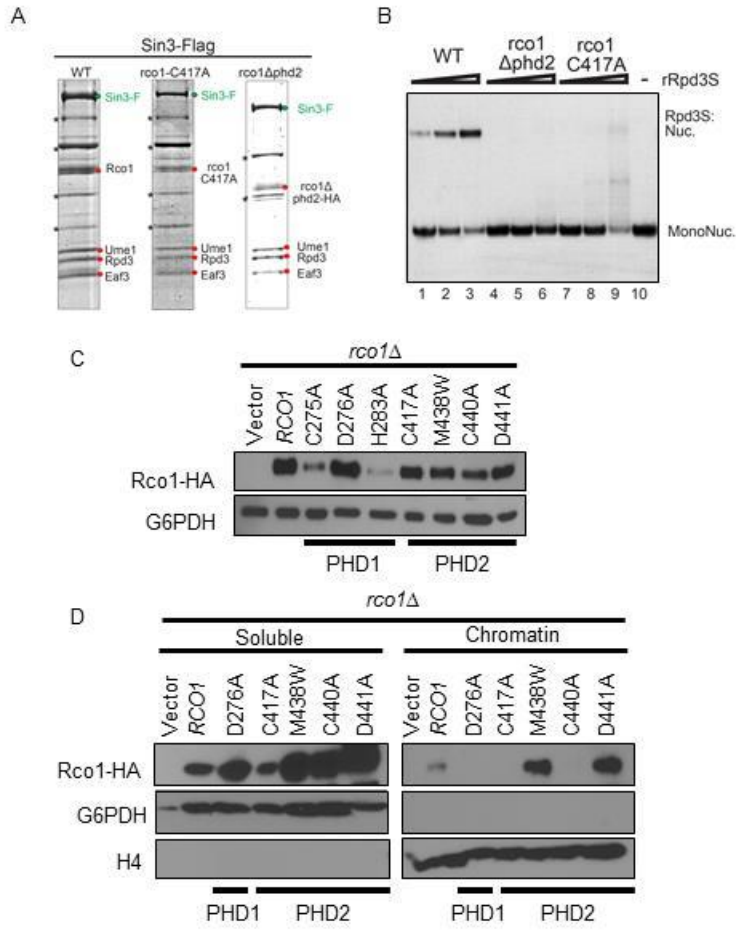

A
Sin3-Flag
WT
rco1-C417A
rco1Δphd2
Sin3-F
Rco1
rco1 C417A
rco1Δ phd2-HA
Ume1
Rpd3
Eaf3
B
WT
rco1 Δphd2
rco1 C417A
rRpd3S
Rpd3S: Nuc.
MonoNuc.
1 2 3 4 5 6 7 8 9 10
C
rco1Δ
Vector
RCO1
C275A
D276A
H283A
C417A
M438W
C440A
D441A
Rco1-HA
G6PDH
PHD1
PHD2
D
rco1Δ
Soluble
Chromatin
Vector
RCO1
D276A
C417A
M438W
C440A
D441A
Rco1-HA
G6PDH
H4
PHD1
PHD2

Figure 3.3:

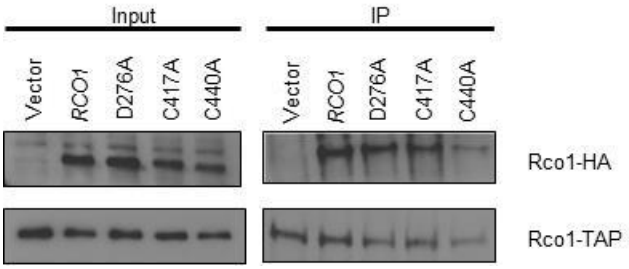

Input
IP
Vector
RCO1
D276A
C417A
C440A
Vector
RCO1
D276A
C417A
C440A
Rco1-HA
Rco1-TAP

Figure 3.4:

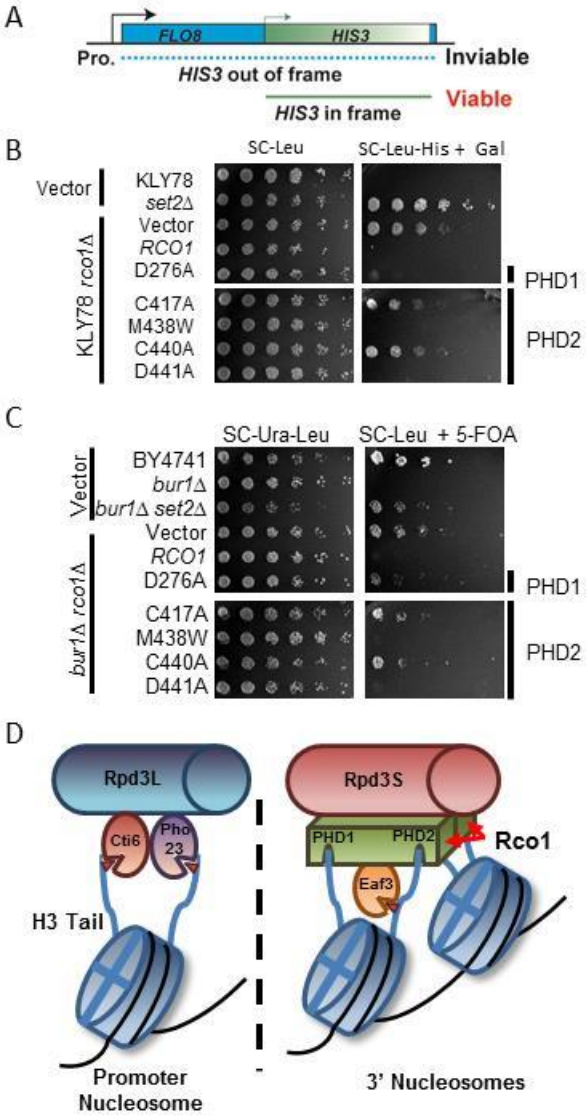
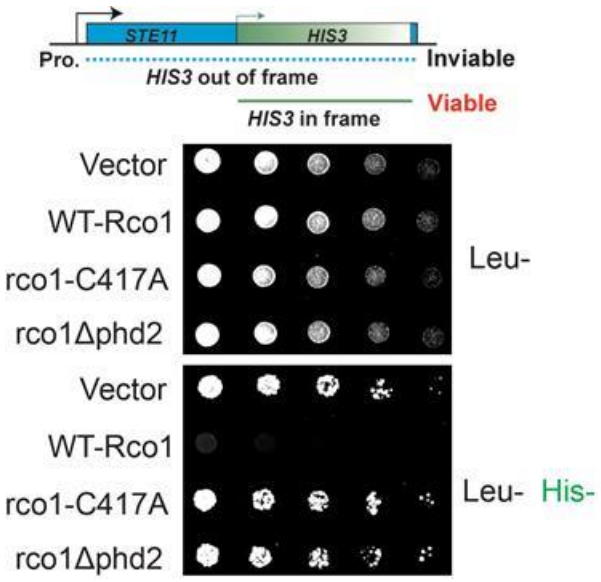

A
FLO8
HIS3
Pro.
Inviable
HIS3 out of frame
HIS3 in frame
Viable
B
SC-Leu
SC-Leu-His + Gal
Vector
KLY78
set2Δ
KLY78 rco1Δ
Vector
RCO1
D276A
C417A
M438W
C440A
D441A
PHD1
PHD2
C
SC-Ura-Leu
SC-Leu + 5-FOA
Vector
BY4741
bur1Δ
bur1Δ set2Δ
bur1Δ rco1Δ
Vector
RCO1
D276A
C417A
M438W
C440A
D441A
PHD1
PHD2
D
Rpd3L
Cti6
Pho 23
H3 Tail
Promoter Nucleosome
Rpd3S
PHD1
PHD2
Rco1
Eaf3
3' Nucleosomes

Figure 3.5:


STE11
HIS3
Pro.
Inviable
HIS3 out of frame
Viable
HIS3 in frame
Vector
WT-Rco1
rco1-C417A
rco1Δphd2
Leu-
Vector
WT-Rco1
rco1-C417A
rco1Δphd2
Leu- His-

CHAPTER 4 – HISTONE H3 LYSINE 36 METHYLATION IS NECESSARY FOR PROPER TRANSCRIPTIONAL REGULATION DURING THE NUTRIENT STRESS RESPONSE IN *SACCHAROMYCES* CEREVISIAE

Introduction:

Eukaryotic DNA complexes with histone proteins to form the basic unit of chromatin, the nucleosome. A vast array of post-translational modifications (PTMs) occur on histones and provide distinct binding sites for a wide variety of effector proteins to interact with the genome and direct essential DNA-based functions such as gene expression and DNA repair (Strahl and Allis, 2000). Set2 is a highly conserved histone methyltransferase that methylates histone H3 at lysine 36 (H3K36) (Strahl et al., 2002; Wagner and Carpenter, 2012). In contrast to higher eukaryotes, Set2 is the sole H3K36 methyltransferase and is responsible for modifying H3K36 with up to three methyl groups, creating mono-, di-, and tri-methylated H3K36 species in the budding yeast *Saccharomyces cerevisiae*. Significantly, increasing evidence in yeast and human cells suggest that each H3K36 methyl state is capable of recruiting distinct effector proteins (Carrozza et al., 2005; Gilbert et al., 2014; Keogh et al., 2005; Smolle et al., 2012; Vermeulen et al., 2010), thereby increasing the signaling capacity of this important histone residue.

A major function of H3K36me in yeast is to repress bi-directional transcription at promoters and the aberrant localization of RNA polymerase II (RNAPII) in gene bodies

– each leading to the production of “cryptic” transcripts (Carrozza et al., 2005; Keogh et al., 2005). In large part, H3K36me prevents the production of cryptic transcripts by recruiting the ISW1b chromatin remodeling complex, via binding of H3K36me3 by the PWWP domain of Ioc4 (Maltby et al., 2012; Smolle et al., 2012), and the Rpd3S histone deacetylase complex, via binding of the chromodomain of Eaf3 to H3K36me2/3 (Carrozza et al., 2005; Joshi and Struhl, 2005; Sun et al., 2008; Xu et al., 2008). Together, these complexes function to deacetylate histones in the wake of the elongating RNAPII complex (Rpd3S) and restore the nucleosomes to their proper positions (ISW1b), thereby preventing RNAPII from aberrantly binding to cryptic promoters inside the gene body and preventing bi-directional transcription – thus maintaining the proper directionality of transcription (Churchman and Weissman, 2011). It is notable that the production of anti-sense transcripts is known to have deleterious effects on normal transcript levels, particularly if they overlap with the native promoter for that gene (Huber et al., 2016). Despite the potential for such deleterious effects, it is surprising that a loss of Set2 and the production of cryptic transcripts that arise from the deletion of *SET2* do not result in global changes in the transcriptome (Lenstra et al., 2011) or loss of cell viability (Strahl et al., 2002); although the loss of transcriptional fidelity in *set2*Δ cells has been shown to result in decreased lifespan in yeast and worms (Sen et al., 2015).

Given the lack of gross transcription changes, but the significance of preventing cryptic transcription in *set2*Δ cells, we asked if Set2 and H3K36me would be important for the transcription and/or transcriptional fidelity of other aspects of cellular biology that have not yet been investigated. In particular, we investigated if the loss of

transcriptional fidelity observed in *set2*Δ cells would have a negative impact on the nutrient stress response, a genetic pathway centered on the highly conserved Tor1 complex (TORC1). During stress conditions, cells must initiate well-timed transcriptional programs to ensure that they are able to quickly and accurately respond to stress. Because cells lacking H3K36me see a global reduction of transcriptional fidelity, we hypothesized that *set2*Δ cells would be unable to properly respond to stress and would show a significant deregulation of their transcriptional response to that stress.

In this report, we interrogated the importance of Set2 and H3K36me in nutrient stress response by carrying out a genetic, molecular, biochemical analysis of WT and *set2*Δ cells. We find that *set2*Δ cells are sensitive to caffeine, a drug that can inhibit Tor1, Tor2, and MAP kinases (Reinke et al., 2006) and rapamycin, a specific inhibitor of the TORC1 (Heitman et al., 1991). Further, we demonstrate that *SET2* deleted cells show synthetic genetic interactions with the *TOR1, TOR2,* flocculation, and protein kinase C (PKC) MAP kinase genetic pathways. Critically, TORC1 and PKC signaling are also disrupted in *set2*Δ cells. Together, this work suggests that under ideal growth conditions, high levels of transcriptional noise is permitted and has little impact upon cellular fitness. However, when cells encounter stress and must execute a precise and rapid transcriptional program to appropriately deal with this stress, a loss of transcriptional fidelity is highly detrimental.

Materials and Methods:

Yeast Strains

Yeast strains were created using standard methods (Janke et al., 2004) employing PCR amplified cassettes with ~50 base pairs of homology to the gene of interest (Janke et al., 2004).

Spotting Assays

Strains were grown in YPD (1% yeast extract, 2% peptone, and 2% glucose) and diluted to an OD_{600} of 0.5 prior to spotting the indicated strains over a 5-fold serial dilution on the indicated plates at 30C for 2-3 days.

H3K36 Methylation Analyses

Protein was isolated from $5*10^7$ cells as previously described (Gilbert et al., 2014). Extracts were loaded onto 15% SDS-PAGE gels and transferred to PVDF. Membranes were incubated overnight with H3 C-term (EpiCypher), H3K36me1 (Abcam 9048), H3K36me2 (Active Motif 39255), H3K36me3 (Abcam 9050), Set2 (in house), or G6PDH (Sigma A9521) antibodies. Membranes were then washed in TBS-Tween (50 mM Tris, 150 mM NaCl, and 0.5% Tween 20), incubated in secondary antibody (Jackson Labs) and then probed with ECL solution (GE Healthcare).

Genome-wide Suppressor Screen

Three isolates of a high-copy number 2µ library of Sau3AI digested genomic fragments (Carlson and Botstein, 1982) were transformed into $1*10^8$ *set2*Δ cells. Transformations were re-suspended in 1 ml of SC-Ura media and plated onto 20 SC-Ura + 20 mM caffeine plates. Plasmids were recovered from the indicated yeast strains using Qiagen mini-prep columns and sequenced by Sanger sequencing using one of two primers flanking the region of insertion: pBR-1: CACTATCGACTACGCGATCA or pBR-2: CGATGCGTCCGGCGTAGA.

Phospho-Protein Analysis

WT or *set2*Δ cells were grown overnight in YPD and diluted to an OD_{600} of 0.2 and grown to an OD_{600} of ~1.0. Upon reaching log phase, cells were isolated and washed with water twice and resuspended in SD media. Ten ODs of cells were isolated at each time point and protein was isolated via TCA extraction as previously described (Fillingham et al., 2008). Extracts were then loaded onto 10% SDS-PAGE gels and transferred to PVDF membrane. Membranes were incubated overnight with the following antibodies: Rps6 (Abcam ab40820), ph-S6K (Cell Signaling 2211S), ph-EIF2α (Cell Signaling 9721S), ph-MAPK (Cell Signaling 9101S), Set2 (in house), and G6PDH (Sigma A9521). Membranes were then washed in TBS-Tween (50 mM Tris, 150 mM NaCl, and 0.5% Tween 20) and then incubated in secondary antibody (Jackson Labs) and then probed with SuperSignal West Fempto Maximum Sensitivity Substrate (Thermofisher).

Results:

Cells Lacking Set2 or H3K36 Methylation are Sensitive to Caffeine and Rapamycin

Recent work from our lab and others show that Set2-mediated H3K36 methylation is also linked to DNA repair (Jha and Strahl, 2014; Pai et al., 2014; Winsor et al., 2013) and mRNA splicing in yeast (Jha and Strahl, 2014; Sorenson et al., 2016), suggesting Set2 and H3K36 methylation may have additional functions in chromatin not yet identified. As Set2 and H3K36 methylation function to suppress cryptic transcription, we reasoned that one of these functions might be to ensure the proper transcriptional fidelity of transcription programs that are both rapid and highly timed, such as the nutrient stress response. To explore this hypothesis, we exposed *set2*Δ cells to caffeine (which inhibits Tor1, Tor2, and MAP kinases, among other things) and rapamycin (a potent and specific inhibitor of the Tor1 complex (TORC1)). As shown in Figure 4.1A, we found *set2*Δ cells from three distinct genetic backgrounds tested show a slow growth phenotype in the presence of caffeine and rapamycin, indicating that the stress response pathway is disrupted in the absence of Set2.

While non-histone substrates for Set2 have not been reported, we could not eliminate the possibility that the growth defects we observed in the *set2*Δ cells on caffeine and rapamycin were due to non-histone targets of Set2. To explore this possibility, we employed histone mutant strains in which H3K36 was mutated to alanine or arginine (H3K36A and H3K36R), thereby creating non-methylatable forms of H3. As shown in Figure 4.1B, the H3K36A and H3K36R mutant strains showed a slow growth phenotype similar to that observed for *set2*Δ when plated on caffeine and rapamycin. Taken together, these data show that Set2-mediated H3K36 methylation is necessary

for yeast cells to respond properly to nutrient stress, rather than methylation of a non-histone substrate.

H3K36me3, but not H3K36me2, is Dispensable for the Nutrient Stress Response

To further understand the role of Set2/H3K36me in the nutrient stress response, we next asked which methyl states of H3K36 were necessary for cells to respond properly to nutrient stress. To do this, we employed a variety of *SET2* mutants that either affect the association of Set2 with RNAPII (*set2-1-618*), the ability of Set2 to interact with the nucleosome (*set2-*Δ*31-39*) or impact its catalytic function (*set2-R195C* and *set2-H199L*) – all of which result in limiting degrees of H3K36 methylation *in vivo* (Figure 4.2). Before performing this analysis, we confirmed that the phenotypes in *set2*Δ cells are specific to a loss of *SET2*, as a WT *SET2* expression construct was able to rescue the growth defect observed in *set2*Δ cells on caffeine (Figure 4.1C) and rapamycin (data not shown). Interestingly, mutants that lack H3K36me3, but not H3K36me1 or me2, were able to rescue the phenotype as well (*set2-*Δ*31-39* and *set2-R195C*). In contrast, the catalytically-dead *SET2* mutant (*set2-H199L*) and a construct with a truncation of the RNAPII interaction domain (*set2-1-618*) (Kizer et al., 2005; Schaft et al., 2003; Vojnic et al., 2006) failed to rescue the phenotype of *set2*Δ cells. These results highlight a critical role for H3K36me1 and H3K36me2 in mediating proper nutrient response, and show that H3K36me3 is dispensable for this phenotype. They also show that the interaction of Set2 with elongating RNAPII is critical for viability during the nutrient stress response.

Multiple H3K36me Effector Complexes are Necessary to Properly Respond to Nutrient Stress

Given H3K36 methylation is critical for responding to caffeine and rapamycin stress, we next asked which of the known effector proteins that are recruited to this mark are responsible for mediating the nutrient stress response. To this end, we first examined the caffeine sensitivity of *EAF3* or *RCO1* – two key members of the Rpd3S histone deacetylase complex – singly or in combination with *SET2*. Unexpectedly, neither *eaf3*Δ nor *rco1*Δ cells showed sensitivity to caffeine when deleted alone (Figure 4.1D). Further, *eaf3*Δ and *rco1*Δ double deletions with *set2*Δ also phenocopied the deletion of *SET2*. These data indicate that the Rpd3S complex is not solely mediating the cellular response to nutrient stress.

Because Rpd3S loss did not phenocopy the caffeine sensitivity observed in *set2*Δ cells, we expanded our search to other effector proteins for H3K36me: Ioc4 of the Isw1b chromatin-remodeling complex (Smolle et al., 2012), and Pdp3, a member of the NuA3b histone acetyltransferase complex which also recognizes H3K36me3 (Gilbert et al., 2014). We also examined the chromatin remodeler *CHD1*, which does not directly associate with H3K36me and functions in a separate genetic pathway (Biswas et al., 2007; Park et al., 2014) as a control. Similar to the deletions of *EAF3* and *RCO1*, single deletions of *PDP3, IOC4,* and *CHD1* did not show sensitivity to caffeine, indicating that the slow growth phenotype observed in *set2*Δ cells is not mediated by any single H3K36me effector complex (Figure 4.1E). In contrast, the *rco1*Δ *ioc4*Δ double mutant strain showed a subtle sensitivity to caffeine, but did not fully recapitulate the phenotype observed for *set2*Δ cells. Unexpectedly, only the *rco1*Δ *ioc4*Δ *pdp3*Δ triple mutant strain

was able to fully phenocopy a deletion of *SET2* (Figure 4.1E). These results are intriguing, as they argue that it is the function of all currently known H3K36me-binding complexes (Rpd3S, Isw1b and NuA3b) that mediate the effects of Set2/H3K36me during stress. NuA3b is necessary for expression of Sas3 target genes, however, our data may suggest a more broad function for this complex in Set2 biology. Interestingly, the *rco1*Δ *ioc4*Δ *chd1*Δ triple mutant also displays strong synthetic sickness on caffeine plates, but did not recapitulate the phenotype observed for *set2*Δ cells in the presence of caffeine. Taken together, our data suggest it is the combination of histone-modifying and chromatin-remodeling factors that signal downstream of H3K36me to promote proper nutrient stress response.

Nutrient and PKC Signaling Pathways Genetically Interact with *SET2*

Having established that *SET2* and H3K36me play a role during the caffeine and rapamycin stress response, we next elucidated which particular genetic pathways *SET2* interacts with during this response. To accomplish this, we transformed *set2*Δ cells with 2µ plasmid overexpression library of genomic fragments (Carlson and Botstein, 1982) and plated the cells on plates with a concentration of caffeine (20 µM) that *set2*Δ cells are normally inviable. This suppression screen yielded genes that fell into one of three distinct pathways: the nutrient signaling pathway (Pho85 and Bmh1), the flocculation MAPK pathway (Sfl1) and the PKC signaling pathway (Lre1, Kkk1 and Slt2) (Figure 4.3A). To explore the results of the suppression screen further, we created single deletions representing all of the recovered genes, along with double deletions of these genes with *SET2*, and analyzed their genetic interactions on caffeine-containing plates

as described above. Genes recovered from all three major pathways in the screen displayed negative genetic interactions with *SET2*, thus further confirming these pathways genetically interact with *SET2* (Figure 4.3B).

SET2 Genetically Interacts with Both *TOR1* and *TOR2*

Although components of the *TOR1* and *TOR2* pathways showed genetic interactions with *SET2* in the suppressor screen above, neither kinase was directly isolated in the screen. To specifically examine if *TOR1* and *TOR2* would show genetic interactions with *SET2*, we combined a deletion of *SET2* with either a deletion of *TOR1* or a temperature sensitive mutant of *TOR2* (*tor2-1),* as *TOR2* is essential in budding yeast, and plated these strains on caffeine and rapamycin. Consistent with our synthetic genetic interactions in Figure 2B, a double deletion of *tor1*Δ and *set2*Δ resulted in a synthetic growth defect on both caffeine and rapamycin plates (Figure 4.3C). This was also the case with the *tor2-1 set2*Δ double mutant strain (Figure 4.3D). Notably, lower concentrations of caffeine and rapamycin were used in these experiments, as the *TOR1* and *TOR2* mutant alleles are extremely sensitive to rapamycin and caffeine respectively; thus, in this context, *set2*Δ cells do not show a significant growth defect on either drug.

Given the differential sensitivities of the *SET2* deletion compared to the *TOR1* and *TOR2* mutants, we next verified that the sensitivity that we were observing with rapamycin in the *set2*Δ cells was occurring through TORC1 and not through an off-target effect of the drug. To examine this, we combined a *SET2* deletion with a deletion of *FPR1*, the key mediator of rapamycin inhibition through its ability to bind this drug and

inhibit TORC1 kinase activity (Heitman et al., 1991). When plated on rapamycin, *set2*Δ *fpr1*Δ cells do not show sensitivity to rapamycin, indicating that the decreased growth rate seen in *set2*Δ cells is due to inhibition of the TORC1 (Figure 4.4). These results further validate our screen and suggest that *SET2* is playing an important role in the nutrient stress response pathway.

Tor1 and MAPK Activity is Aberrant in *set2*Δ Cells

Given *SET2* genetically interacts with both *TOR1* and *TOR2*, we hypothesized that *set2*Δ cells would show significant disruptions in TORC1 and nutrient response signaling in the absence of *SET2* and H3K36me. We shifted WT and *set2*Δ cells from nutrient rich YPD media to media lacking amino acids (SD media) and began collecting samples every 30-60 minutes over a 2 hour time period. Extracts from these samples were then used to examine the protein and phosphorylation levels of multiple signaling proteins in the nutrient signaling pathway: Rps6, the *Saccharomyces cerevisiae* S6K homolog (Gonzalez et al., 2015), EIF2α, and the Slt2 MAPK (involved in PKC signaling). As shown in Figure 4.5, the phosphorylation levels of EIF2α, along with the protein levels of our loading control, G6PDH, were not altered during the stress response in the absence of *SET2*. In striking contrast, however, we found that *set2*Δ cells have much lower levels of Rps6 compared to their isogenic WT counterparts, and further, have drastically increased levels of phosphorylation at the start of the nutrient response. The increased levels of phosphorylation of Rps6 found in *set2*Δ cells is indicative of aberrant TORC1 kinase signaling. In addition, we observed a change and disruption in the pattern and timing of phosphorylation of the MAPK Slt2. Specifically, levels of ph-Slt2 in

*set2*Δ cells begin at WT levels and increase like their WT counterpart, but persist at high levels to 60 minutes after stress whereas in WT cells, phosphorylation levels of Slt2 are more transient reach background levels at 60 minutes. Together, these results show that Set2 is necessary at a biochemical level for proper nutrient response signaling, which is in agreement with our genetic analyses that Set2 and H3K36me are needed for proper nutrient response signaling.

Set2 and the Exosome Independently Regulate the Nutrient Stress Response

If cryptic transcripts and a resulting loss of transcriptional fidelity are truly detrimental to the cellular stress response, then inhibiting the destruction of these aberrant transcripts should exacerbate the growth defects we observe in *set2*Δ cells. To test this hypothesis, we combined deletions of *SET2* with a component of the exosome, *RRP6.* Consistent with this hypothesis, the *set2*Δ *rrp6*Δ double mutant displays a synthetic growth defect on caffeine plates (Figure 4.6). This suggests that the increased levels of cryptic transcripts, which already interfere with proper signaling, are protected in the exosome mutants, and thus are able to further increase the disruption to the nutrient stress response.

Discussion and Future Directions:

Together, these results show that *SET2* is genetically interacting with pathways involving both Tor and MAP kinases. Each kinase is critical for controlling various aspects of the cellular response to stress. While *SET2* has been previously shown to genetically interact with the DNA damage response pathways (Jha and Strahl, 2014;

Pai et al., 2014; Winsor et al., 2013), this is the first demonstration that *SET2* genetically interacts with the Tor complexes in *Saccharomyces cerevisiae*. Further, this is the first evidence that not only shows that *SET2* genetically interacts with these pathways, but also demonstrates that H3K36me is biochemically necessary for proper signaling. Yet, it remains to be seen what the full impact of a loss of *SET2* and H3K36me have on the transcriptional response during this particular stress.

One of the reasons that *set2*Δ cells grows poorly on inhibitors of these pathways could be due to the inability of these cells to signal properly during the stress response. Critically, we observe that *set2*Δ cells have different biochemical ground states when compared to WT cells. While not overtly deleterious to cell growth, this differential ground state could be partially responsible for the delayed phosphorylation kinetics observed in the TORC1 and PKC signaling pathways. In addition, based on the genetic interaction with the exosome and *SET2* during this response, it is equally likely that *set2*Δ cells are also not able to respond properly at the transcriptional level either. Like the differential ground state observed biochemically, there have been subtle differences observed in the transcriptomes of *set2*Δ cells (Lenstra et al., 2011). It is quite possible that these initial differences in the transcriptome could ultimately change the starting pool of signaling proteins, further inhibiting the cell's ability to quickly and accurately respond to nutrient stress.

It is also possible, and likely, that cryptic transcription events could lead to alterations in the transcriptome during nutrient stress. Cryptic transcripts that overlap the promoters of genes are very likely to have negative impacts on the transcript levels of those genes (Huber et al., 2016). Future studies will be needed to identify the cryptic

transcripts produced in *set2*Δ cells to test if any of them are directly responsible for changes in nutrient response gene induction during stress.

While we have no direct molecular evidence that the transcriptome is altered during the nutrient stress response in *set2*Δ cells, the isolation of the flocculation pathway in our screen is indicative of transcriptional deregulation. The flocculation pathway is a MAPK pathway in *Saccharomyces cerevisiae* and is normally transcriptionally repressed in laboratory strains of yeast. It is possible that this pathway was isolated in our screen because the flocculation of cells acts as a defensive mechanism, but we do note that the *sfl1*Δ *set2*Δ double mutant strain displays a strong synthetic flocculation phenotype, whereas neither single mutant alone does (data not shown). Because this pathway is normally repressed, this indicates a high level of transcriptional deregulation in the double mutant and supports the hypothesis that *set2*Δ cells are not able to regulate their transcriptional programs accurately.

*set2*Δ cells alone do not show robust levels of transcriptional degregulation genome-wide (Lenstra et al., 2011). This is also true of many other chromatin remodeling and modifying factors (Lenstra et al., 2011). However, most studies looking at the transcriptomes of these mutants, do so in the context of non-stress conditions. Because relatively few of these mutants are essential or show any growth defects when deleted, it is clear the minor alterations observed to their aberrant transcriptomes are not overtly deleterious. However, this may not be the case during stress conditions. Here, we clearly demonstrate that while *set2*Δ cells are viable under normal growth conditions, they are unable to properly respond to nutrient stress. Whether this response is unique to the nutrient stress response, or can be applied to other types of

stress remains to be seen and will followed up in the future. It is clear, however, that we may not fully appreciate the function of these chromatin factors until we examine them in a broader context outside of ideal laboratory conditions.

Finally, it is likely that the function of Set2 and H3K36me is conserved in higher eukaryotes. H3K36me is one of the few histone modifications that is conserved from yeast to humans. Further, Set2 itself is also highly conserved. Its human homolog, *SETD2* also methylates H3K36, though only the H3K36me2/3 forms (Sun et al., 2005; Yoh et al., 2008). *SETD2* is also one of the most mutated genes in a wide spectrum of cancers (Dalgliesh et al., 2010; Gerlinger et al., 2012; Mar et al., 2014; Varela et al., 2011; Zhu et al., 2014), making it a very attractive target for treatment. Interestingly, TOR inhibitors combined with *SETD2* knockdown
has shown promise in killing leukemia cells (Zhu et al., 2014). It is possible, due to the high level of conservation of both *SET2* and the TOR pathway, that the interactions seen here in *Saccharomyces cerevisiae* are conserved in humans as well.

Figure Legends

Figure 4.1: Cells lacking H3K36me are sensitive to caffeine and rapamycin

A. – D. Five-fold serial dilutions of the indicated strains were plated on SC plates or plates containing caffeine (7-10 mM or rapamycin (8-12.5 nM).

Figure 4.2: Set2 mutants differentially methylate histones

Western blots of the indicated strains were probed with different H3K36me antibodies. H3 and G6PDH serve as loading controls.

Figure 4.3: Nutrient response and PKC signaling pathways genetically interact with *SET2*.

A. A genome-wide high copy suppressor screen carried out revealed three pathways that are able to suppress the lethality of *set2*Δ cells on a high concentration of caffeine (20 mM). B. Five-fold serial dilutions of the indicated strains were plated on control or 5-10 mM caffeine plates. C. and D. Five-fold serial dilutions of the indicated strains were plated on control, caffeine (5 mM), and rapamycin (8 nM) plates.

Figure 4.4: Bypass of the Tor1 pathway renders *set2*Δ cells resistant to rapamycin

The indicated strains were 5-fold serial diluted and plated on control or rapamycin plates (25 nM) and grown for 2 days.

Figure 4.5: Set2 is required to properly activate pathways necessary to respond to nutrient stress

Log phase cells were transferred from nutrient rich YPD media to SD media that lacks amino acids. $1*10^8$ cells were isolated at the indicated times and submitted to immunoblot. G6PDH serves as a loading control.

Figure 4.6: Cryptic transcripts within the *SET2* pathway may be responsible for the growth defects on caffeine and rapamycin

The indicated strains were 5-fold serial diluted and plated on control or caffeine (7 mM) plates and grown for 2-3 days.

Figure 4.1:

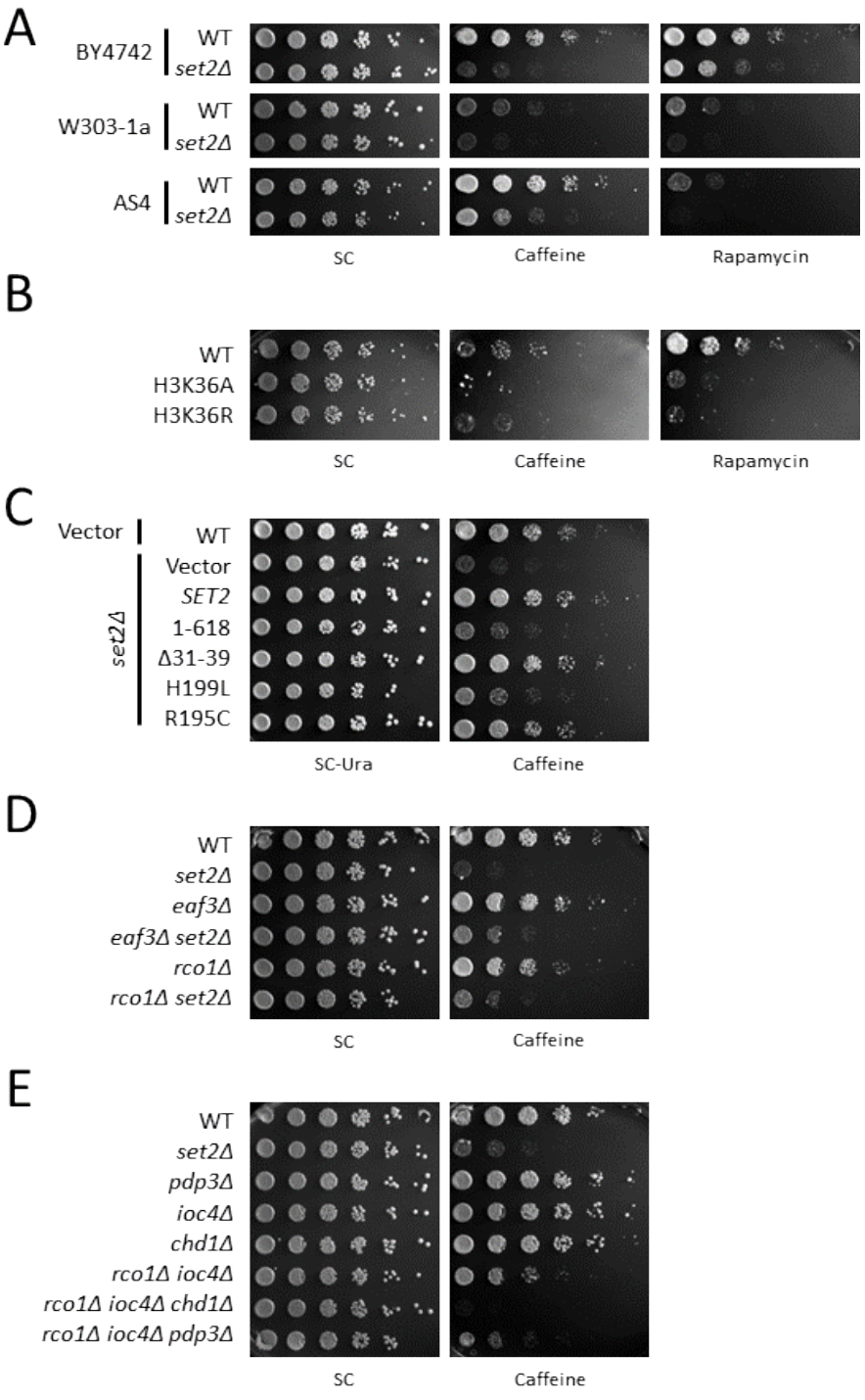

A
BY4742
WT
set2Δ
W303-1a
WT
set2Δ
AS4
WT
set2Δ
SC
Caffeine
Rapamycin
B
WT
H3K36A
H3K36R
SC
Caffeine
Rapamycin
C
Vector
WT
set2Δ
Vector
SET2
1-618
Δ31-39
H199L
R195C
SC-Ura
Caffeine
D
WT
set2Δ
eaf3Δ
eaf3Δ set2Δ
rco1Δ
rco1Δ set2Δ
SC
Caffeine
E
WT
set2Δ
pdp3Δ
ioc4Δ
chd1Δ
rco1Δ ioc4Δ
rco1Δ ioc4Δ chd1Δ
rco1Δ ioc4Δ pdp3Δ
SC
Caffeine

Figure 4.2:

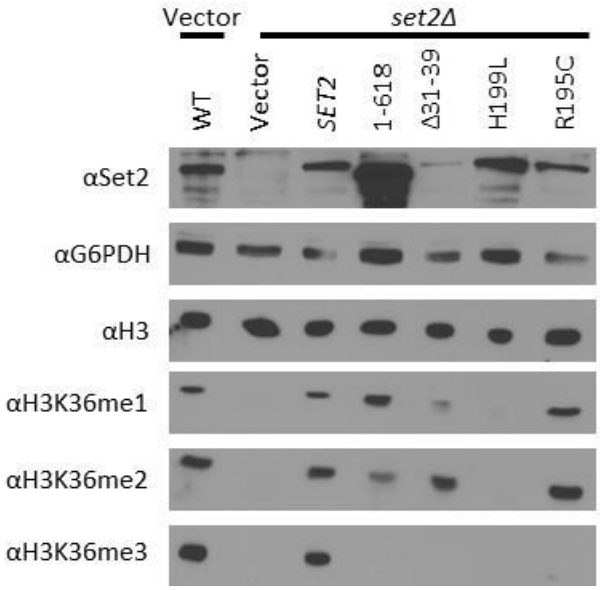

Vector
set2Δ
WT
Vector
SET2
1-618
Δ31-39
H199L
R195C
αSet2
αG6PDH
αH3
αH3K36me1
αH3K36me2
αH3K36me3

Figure 4.3:

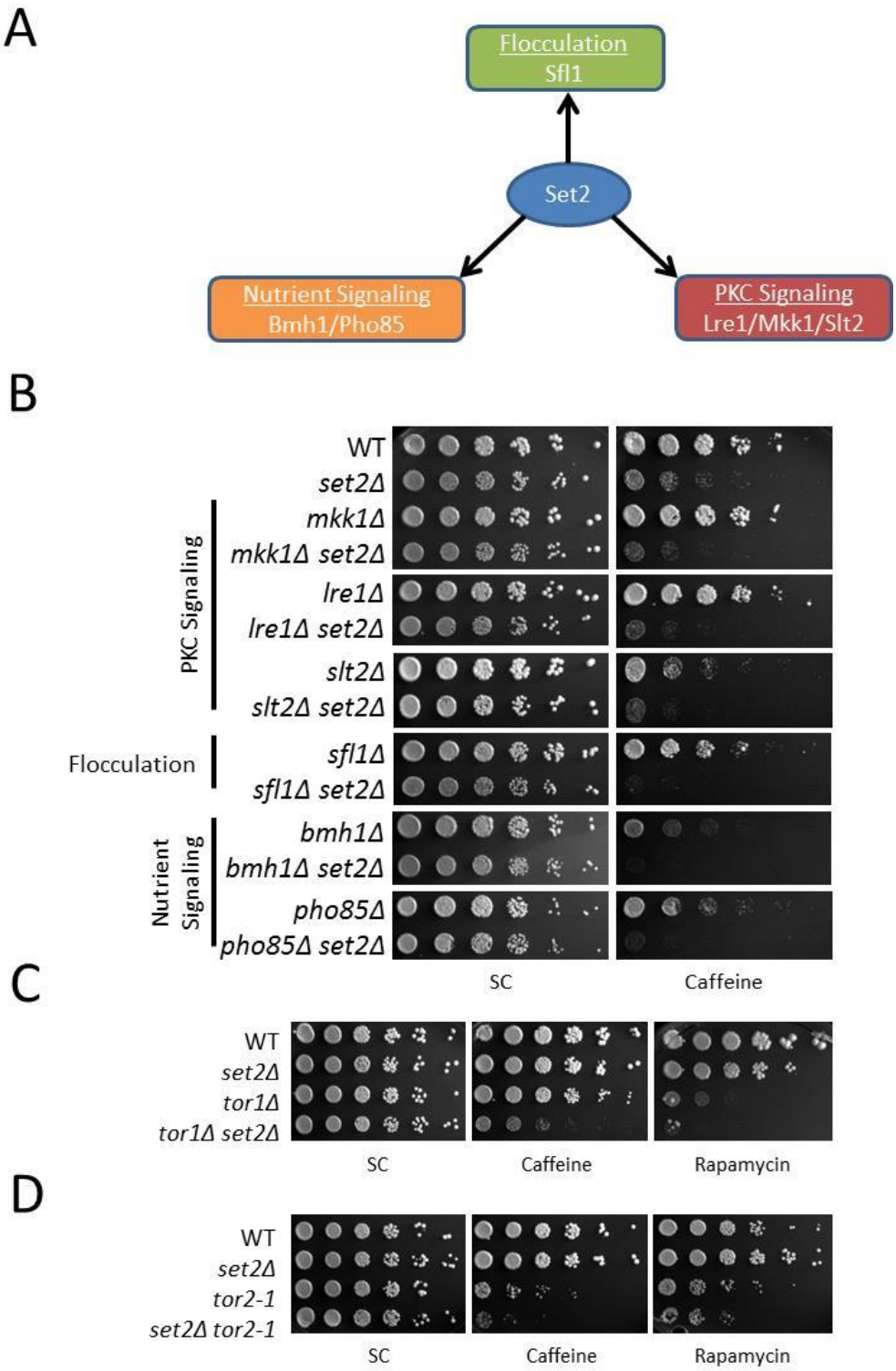

A
Flocculation
Sfl1
Set2
Nutrient Signaling
Bmh1/Pho85
PKC Signaling
Lre1/Mkk1/Slt2
B
WT
set2Δ
mkk1Δ
mkk1Δ set2Δ
lre1Δ
lre1Δ set2Δ
slt2Δ
slt2Δ set2Δ
sfl1Δ
sfl1Δ set2Δ
bmh1Δ
bmh1Δ set2Δ
pho85Δ
pho85Δ set2Δ
PKC Signaling
Flocculation
Nutrient Signaling
SC
Caffeine
C
WT
set2Δ
tor1Δ
tor1Δ set2Δ
SC
Caffeine
Rapamycin
D
WT
set2Δ
tor2-1
set2Δ tor2-1
SC
Caffeine
Rapamycin

Figure 4.4:

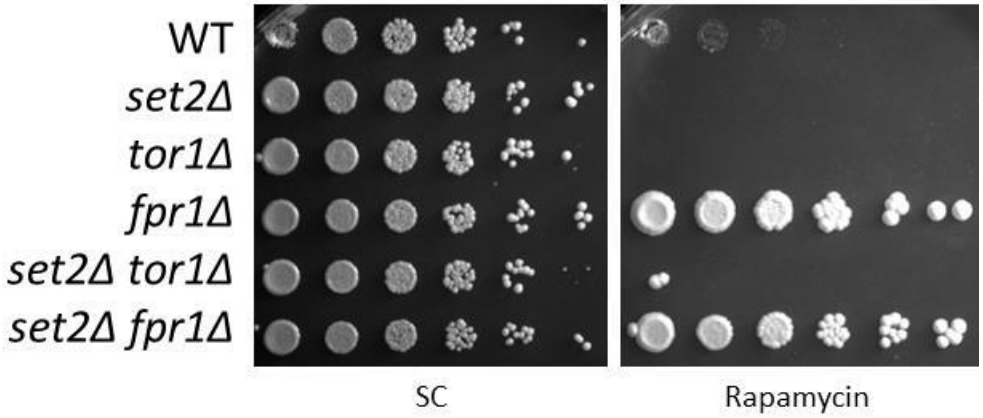


Figure 4.5:

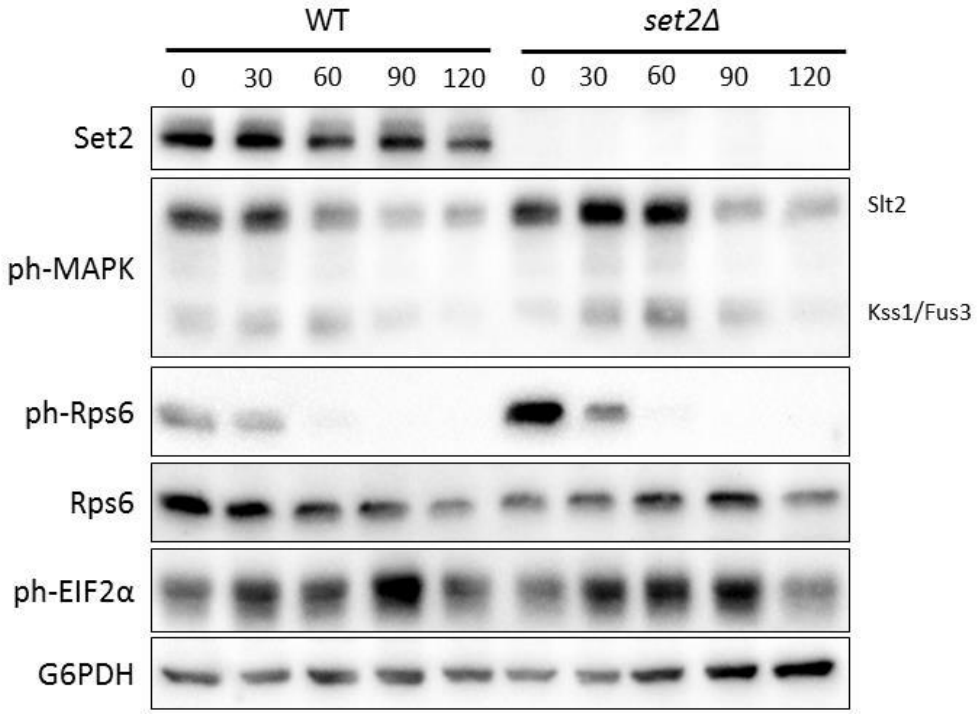
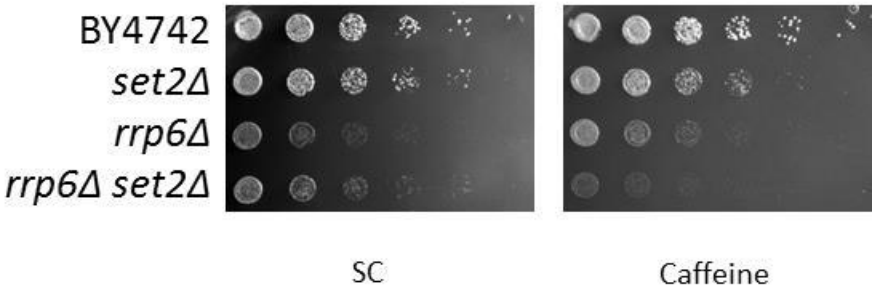

WT
set2Δ
0 30 60 90 120 0 30 60 90 120
Set2
ph-MAPK
Slt2
Kss1/Fus3
ph-Rps6
Rps6
ph-EIF2α
G6PDH

Figure 4.6:



CHAPTER 5 - CONCLUSIONS AND FUTURE DIRECTIONS

Prior to this work, H3K36me was thought primarily to function during transcription as a repressor. This was mainly accomplished via the binding of Rpd3S to H3K36me and deacetylating histones in the wake of the elongating RNAPII complex (Carrozza et al., 2005; Joshi and Struhl, 2005; Sun et al., 2008; Xu et al., 2008). In the absence of H3K36me, histones were rendered hyperacetylated and chromatin structure would become more open, leading to cryptic transcription events (Carrozza et al., 2005; Keogh et al., 2005). Together, this work has expanded the role of H3K36me, demonstrating that while it still functions largely during transcription, its role there is much more multifaceted and important than previously thought.

H3K36me Recruits a Novel Histone Acetyltransferase Complex

H3K36me has previously been linked to histone acetylation via Eaf3's presence in the NuA4 HAT complex (Ginsburg et al., 2014; Sathianathan et al., 2016). Here, we add to Set2's role in histone acetylation by characterizing a novel H3K36me effector protein, Pdp3, which is a member of the NuA3b HAT complex. Like many other PWWP domain containing proteins, Pdp3 prefers to bind H3K36me3 (Vermeulen et al., 2010) and in the absence of Set2 and H3K36me, cannot associate with chromatin. Finally, we demonstrate molecularly and genetically that the interaction between Pdp3 and H3K36me is necessary for proper gene expression.

Yet, there are still several unanswered questions regarding the function of Pdp3 and NuA3b. The first being what is the catalytic target of this complex? NuA3a, a second, promoter bound complex, binds chromatin via the PHD finger of Yng1 and acetylates histone H3 at lysine 14. It is possible that NuA3b, like NuA3a, also acetylates H3K14, but because NuA3b is engaging chromatin in a fundamentally different way; i.e. mainly at H3K36 rather than at H3K4, it is distinctly possible that this complex has an entirely different enzymatic target. This hypothesis is supported by the fact that NuA3b does not display synthetic lethality with the other H3K14 acetyltransferase, Gcn5, while NuA3a does. This change of target would not be unique as human HAT complexes have been shown to change their catalytic targets from H3 to H4 depending on which chromatin binding subunits they contain (Lalonde et al., 2013). It is quite possible that NuA3a/b function similarly, by switching out Yng1 for Pdp3.

Once the catalytic target of NuA3b has been elucidated, it will be easier to identify the broader role of NuA3b. Here, we show that it is necessary for the proper expression of several genes, but do not yet understand precisely how this is regulated and if it truly plays a different role than NuA3a.

The Second PHD Finger of Rco1 is Necessary for Rpd3S Function

In 2007, the Workman group demonstrated that the chromo domain of Eaf3 and the first PHD finger of Rco1 were necessary for Rpd3S function (Li et al., 2007a). Rco1 contains a second, less well conserved PHD finger that was not characterized. Here, we show that both PHD fingers are essential for chromatin association of Rpd3S. Further, they are also essential for the repression of cryptic transcription. Many

chromatin associated complexes contain multiple proteins with putative chromatin binding domains and this work demonstrates that they are all likely functionally relevant. Further, it is also important to note that these chromatin reader proteins can preferentially bind to unmodified residues as well, as do both of the PHD fingers in Rco1. This adds a further layer of complexity to the "histone code" and is particularly relevant in transcription as certain histone modifications only exist in certain regions of the genome. Interestingly, by switching out members that read H3K4me for Rco1 that recognizes unmodified H3K4 (Shi et al., 2007; Wang et al., 2011), the histone deacetylase Rpd3 can deacetylate histones across the whole open reading frame.

This conservation of function is also realized with H3K36me. Because both HAT complexes and HDACs can read H3K36me, the levels of histone acetylation can be finely tuned throughout the transcription process. This supposes a model that H3K36me acts as a beacon of transcriptional memory within the cell, ensuring not only that RNAPII binds to the proper 5' promoters of genes, by facilitating the deacetylation of histones in the middle of gene bodies after transcription, but H3K36me also play a positive role in transcription by opening up chromatin through the recruitment of HAT complexes like NuA4 and NuA3b to chromatin ahead of RNAPII. This model supports a role for H3K36me only after a gene has been transcribed. Pioneering rounds of transcription likely operate under a different mechanism and whether or not H3K36me plays a role during this process is currently unknown. Though because H3K36me1 and H3K36me2 can be laid down in gene bodies independent of RNAPII association (Kizer et al., 2005), it is possible that the lower methylation states could play a role during the

pioneering rounds of transcription while H3K36me3 helps to regulate transcription only after a gene has already been transcribed.

Set2 and H3K36me are Essential for Proper Nutrient Stress Response

It has previously been shown by our lab and others that Set2 plays a role in the DNA damage response pathway (Jha and Strahl, 2014; Pai et al., 2014; Winsor et al., 2013) where it limits the amount of resection at the sites of DNA damage and allows for NHEJ to occur. The work presented here demonstrates that Set2 and H3K36me also are necessary for proper nutrient stress response signaling. We show that H3K36me3 is not necessary for proper nutrient signaling, though H3K36me1/me2 and the association of Set2 with RNAPII is. Further, we provide evidence that Set2 genetically interacts with Tor1 and Tor2, as well as several MAP kinase pathways in the cell. Further, in the absence of Set2, both Tor1 and MAP kinase signaling are disrupted, both in terms of their kinetics as well as their overall levels. Finally, we demonstrate that when the exosome is compromised, in addition to the loss of Set2, cells fare even more poorly under stress conditions, implicating a functional role for cryptic transcription in this process.

The Role of Cryptic Transcription During the Nutrient Stress Response

Of course this leads to two important questions, what role does Set2 play in regulating the transcriptome during (or prior to) the nutrient stress response and is the regulation of cryptic transcription necessary for the proper stress response? In an

attempt to answer these questions, we are in the process of analyzing RNA-seq data from wild-type (WT) and *set2*Δ cells after nutrient stress. Libraries were made from total strand-specific RNA to analyze both sense and anti-sense transcripts across the genome. We are currently in the process of analyzing our data set, but the initial results appear to be promising and will be discussed below.

While previous data showed only ~80 genes were differentially expressed (Lenstra et al., 2011), our data preliminary analysis shows that ~400 genes are differentially expressed between WT and *set2*Δ cells. The difference in the observed number is likely due to the fact that our data relies on RNA-seq data where the genome is sequenced to a very high depth, while the earlier data comes from a micro-array experiment, which is much less sensitive.

Upon stress however, our initial analysis indicates that nearly half of the genes in the genome do not respond properly to nutrient stress. This result suggests that Set2 and H3K36me are critical for proper initiation of the transcriptional programs required to respond to nutrient stress. While it seems that *set2*Δ cells eventually do catch up to their WT counterparts, the initial misfiring of the transcriptome could explain the altered kinetics and gross differences observed in the Tor1 and MAP kinase signaling after nutrient stress in *set2*Δ cells.

Additionally, this highlights a significant gap in the chromatin field. Most experiments looking at the function of chromatin factors do so in the context of ideal growth conditions. Under these conditions, most of these factors are not essential for viability and their loss does not show any overt growth defects. This is true of all three histone methyltransferases in yeast. Set1, Set2, and Dot1 are all non-essential genes.

While a great deal can be learned about their basic functions using standard growth conditions, it is becoming clear that using stress conditions is critical for understanding the true breadth of their role in the cell. Set2, for instance, has until recently only thought to play a role in repressing cryptic transcription. Now it has been shown to be critical for the proper function of two distinct signaling pathways (Jha and Strahl, 2014; Pai et al., 2014; Winsor et al., 2013), raising the possibility of Set2 and H3K36me being important for the function of other conserved pathways in the cell.

Of course, the role of Set2 in repressing cryptic transcription could play an important function in these signaling pathways. Before stress, *set2*Δ cells have sense and anti-sense transcripts that arise at some genes. Interestingly, as the amount of time the cells are undergoing stress increases, the level of cryptic transcripts increases as well. It seems that the stress conditions themselves exacerbate the predisposition of *set2*Δ cells to create aberrant transcripts which could lead to decreased full-length transcripts at these loci. This phenotype is not entirely unexpected as nutrient stress conditions have been previously shown to produce cryptic transcripts in WT cells (Cheung et al., 2008), however, these transcripts were extremely transient and did not increase over time like they do in *set2*Δ cells. While we see robust induction of cryptic transcription during nutrient stress, it is currently unknown if this is a general phenotype during stress or specific to nutrient stress. To test this hypothesis, similar RNA-seq experiments will be needed in a spectrum of stress conditions to track the alterations to the transcriptome as well as the production of cryptic transcripts.

Further, we observe an interesting correlation in that sense and anti-sense transcription seem to co-occur at many genes, originating at the same location. The

Set2 pathway has already been shown to be important in enforcing directionality at promoters (Churchman and Weissman, 2011), and these data suggest that cryptic promoters are behaving in the same fashion. Thus, bi-directional transcription seems to be extremely common in the absence of Set2 and H3K36me.

Together, these data suggest that the increased transcriptional noise due to cryptic transcription could disrupt the finely tuned transcriptional programs necessary for dealing with the nutrient stress response. But, does the increase in cryptic transcription result in changes to the levels of full-length mRNA transcripts? Anti-sense transcription can interfere with sense transcription if it overlaps with the promoters (Huber et al., 2016). It is possible that the dramatic increase in cryptic transcription observed in *set2*Δ cells across the time course could decrease transcript levels, and thus the protein levels, of critical stress response factors. To address this, we are currently determining if a correlation between levels of anti-sense transcription and the fold-change of the full-length transcript is present in our data.

Further, it has also been shown that some cryptic transcripts are translated *in vivo* (Cheung et al., 2008). It is possible that as the cryptic transcript levels increase in *set2*Δ cells, they could begin to be translated, creating an entirely new pool of proteins. The cryptic proteins previously observed do not seem to exist in *set2*Δ cells, but that does not preclude the possibility that they would not be present under stress. The translation of cryptic transcripts could be particularly detrimental during nutrient stress because ribosome production is severely limited (Powers and Walter, 1999). On top of the potential for transcriptional interference due to anti-sense transcription, increased competition of full-length and cryptic transcripts for a limited pool of ribosomes could

further decrease the protein levels of critical nutrient stress response factors. Quantitative mass-spectrometry experiments will be needed to test this hypothesis.

Finally, it is also distinctly possible that the cryptic proteins in and of themselves could be functional, particularly as dominant negative proteins. Because cryptic proteins would begin from the middle of their open reading frames, they would certainly lack many N-terminal domains, but retain their C-terminal domains. As long as the protein was nominally stable, these proteins could act as sponges for protein-protein interactions while lacking their full functionality. In the case of Set2, it is easy to envision a C-terminal fragment of the protein which contains the WW, CC, and SRI domains, but lacks the catalytic SET domain. This cryptic protein could compete with full-length Set2 for a position on the CTD of RNAPII, potentially limiting the levels of H3K36me3 in the cell. Again, quantitative mass-spectrometry experiments will be needed to identify the cryptic proteome and further molecular experiments will be needed to test the function of these cryptic proteins.

Concluding Remarks

Great strides have been recently made in understanding the broader function of H3K36me in the cell. It is clear that it plays a pivotal role in maintaining transcriptional fidelity by ensuring that RNAPII only transcribes from the natural 5' promoters of genes. Together, the work contained here suggests the hypothesis that Set2 and H3K36me function is to maintain the proper "ground state" of the transcriptome, priming the cell to adapt to various kinds of stress. We see that in the absence of Set2, cryptic transcripts arise, the DNA damage response is attenuated (Jha and Strahl, 2014; Pai et al., 2014;

Winsor et al., 2013), and cells cannot respond properly to nutrient stress. Because cells must rapidly and accurately respond to the various environmental stresses they encounter, they likely must limit the transcriptional noise to allow them to quickly establish a programmed response.

Further, we see these roles conserved in higher eukaryotes. Mutations in SETD2 or at H3K36 in humans are a major driver of cancer (Dalgliesh et al., 2010; Gerlinger et al., 2012; Lu et al., 2016; Varela et al., 2011) and likely cause a loss of regulation over the transcriptome, facilitating tumor development. RNA is mis-spliced (Simon et al., 2014) and is not terminated properly (Grosso et al., 2015) in ccRCC cancers and may play a similar role in other cancer types. Thus, it is critical that we learn more about the role of H3K36me in cells in the hopes of developing treatments to these devastating diseases.

REFERENCES

Adkins, M.W., and Tyler, J.K. (2004). The histone chaperone Asf1p mediates global chromatin disassembly in vivo. The Journal of biological chemistry *279*, 52069-52074.

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. Proceedings of the National Academy of Sciences of the United States of America *51*, 786-794.

Avvakumov, N., and Cote, J. (2007). The MYST family of histone acetyltransferases and their intimate links to cancer. Oncogene *26*, 5395-5407.

Baker, L.A., Allis, C.D., and Wang, G.G. (2008). PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutation research *647*, 3-12.

Balasubramanian, R., Pray-Grant, M.G., Selleck, W., Grant, P.A., and Tan, S. (2002). Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. The Journal of biological chemistry *277*, 7989-7995.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell research *21*, 381-395.

Barrera, L.O., and Ren, B. (2006). The transcriptional regulatory code of eukaryotic cells--insights from genome-wide analysis of chromatin organization and transcription factor binding. Current opinion in cell biology *18*, 291-298.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schubeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243-247.

Bell, O., Wirbelauer, C., Hild, M., Scharf, A.N., Schwaiger, M., MacAlpine, D.M., Zilbermann, F., van Leeuwen, F., Bell, S.P., Imhof, A.*, et al.* (2007). Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. The EMBO journal *26*, 4974-4984.

Bian, C., Xu, C., Ruan, J., Lee, K.K., Burke, T.L., Tempel, W., Barsyte, D., Li, J., Wu, M., Zhou, B.O.*, et al.* (2011). Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. The EMBO journal *30*, 2829-2842.

Biswas, D., Dutta-Biswas, R., and Stillman, D.J. (2007). Chd1 and yFACT act in opposition in regulating transcription. Molecular and cellular biology *27*, 6279-6287.

Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods in enzymology *154*, 164-175.

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473-1476.

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 & development *15*, 3286-3295.

Butler, J.S., Koutelou, E., Schibler, A.C., and Dent, S.Y. (2012). Histone-modifying enzymes: regulators of developmental decisions and drivers of human disease. Epigenomics *4*, 163-177.

Byrum, S.D., Raman, A., Taverna, S.D., and Tackett, A.J. (2012). ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. Cell reports *2*, 198-205.

Byrum, S.D., Taverna, S.D., and Tackett, A.J. (2013). Purification of a specific native genomic locus for proteomic analysis. Nucleic acids research *41*, e195.

Carlson, M., and Botstein, D. (1982). Two differentially regulated mRNAs with different 5' ends encode secreted with intracellular forms of yeast invertase. Cell *28*, 145-154.

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.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Carvalho, S., Vitor, A.C., Sridhara, S.C., Martins, F.B., Raposo, A.C., Desterro, J.M., Ferreira, J., and de Almeida, S.F. (2014). SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. eLife *3*, e02482.

Cengiz, B., Gunduz, E., Gunduz, M., Beder, L.B., Tamamura, R., Bagci, C., Yamanaka, N., Shimizu, K., and Nagatsuka, H. (2010). Tumor-specific mutation and downregulation

of ING5 detected in oral squamous cell carcinoma. International journal of cancer Journal international du cancer *127*, 2088-2094.

Cheung, V., Chua, G., Batada, N.N., Landry, C.R., Michnick, S.W., Hughes, T.R., and Winston, F. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS biology *6*, e277.

Choy, J.S., Tobe, B.T., Huh, J.H., and Kron, S.J. (2001). Yng2p-dependent NuA4 histone H4 acetylation activity is required for mitotic and meiotic progression. The Journal of biological chemistry *276*, 43653-43662.

Chruscicki, A., Macdonald, V.E., Young, B.P., Loewen, C.J., and Howe, L.J. (2010). Critical determinants for chromatin binding by Saccharomyces cerevisiae Yng1 exist outside of the plant homeodomain finger. Genetics *185*, 469-477.

Chu, Y., Sutton, A., Sternglanz, R., and Prelich, G. (2006). The BUR1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by SET2. Molecular and cellular biology *26*, 3029-3038.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368-373.

Dalgliesh, G.L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A., Davies, H., Edkins, S., Hardy, C., Latimer, C*., et al.* (2010). Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature *463*, 360-363.

DeGennaro, C.M., Alver, B.H., Marguerat, S., Stepanova, E., Davis, C.P., Bahler, J., Park, P.J., and Winston, F. (2013). Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. Molecular and cellular biology *33*, 4779-4792.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. The Journal of biological chemistry *285*, 26114-26120.

Diebold, M.L., Loeliger, E., Koch, M., Winston, F., Cavarelli, J., and Romier, C. (2010). Noncanonical tandem SH2 enables interaction of elongation factor Spt6 with RNA polymerase II. The Journal of biological chemistry *285*, 38389-38398.

Dimmer, E.C., Huntley, R.P., Alam-Faruque, Y., Sawford, T., O'Donovan, C., Martin, M.J., Bely, B., Browne, P., Mun Chan, W., Eberhardt, R.*, et al.* (2012). The UniProt-GO Annotation database in 2011. Nucleic acids research *40*, D565-570.

Doolin, M.T., Johnson, A.L., Johnston, L.H., and Butler, G. (2001). Overlapping and distinct roles of the duplicated yeast transcription factors Ace2p and Swi5p. Molecular microbiology *40*, 422-432.

Doyon, Y., Cayrou, C., Ullah, M., Landry, A.J., Cote, V., Selleck, W., Lane, W.S., Tan, S., Yang, X.J., and Cote, J. (2006). ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. Molecular cell *21*, 51-64.

Dronamraju, R., and Strahl, B.D. (2014). A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. Nucleic acids research *42*, 870-881.

Drouin, S., Laramee, L., Jacques, P.E., Forest, A., Bergeron, M., and Robert, F. (2010). DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. PLoS genetics *6*, e1001173.

Du, H.N., Fingerman, I.M., and Briggs, S.D. (2008). Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. Genes & development *22*, 2786-2798.

Dutnall, R.N., and Ramakrishnan, V. (1997). Twists and turns of the nucleosome: tails without ends. Structure *5*, 1255-1259.

Eberharter, A., John, S., Grant, P.A., Utley, R.T., and Workman, J.L. (1998). Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. Methods *15*, 315-321.

Fang, D., Gan, H., Lee, J.H., Han, J., Wang, Z., Riester, S.M., Jin, L., Chen, J., Zhou, H., Wang, J.*, et al.* (2016). The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. Science *352*, 1344-1348.

Fillingham, J., Recht, J., Silva, A.C., Suter, B., Emili, A., Stagljar, I., Krogan, N.J., Allis, C.D., Keogh, M.C., and Greenblatt, J.F. (2008). Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Molecular and cellular biology *28*, 4342-4353.

Fuchs, S.M., Kizer, K.O., Braberg, H., Krogan, N.J., and Strahl, B.D. (2012). RNA polymerase II carboxyl-terminal domain phosphorylation regulates protein stability of the Set2 methyltransferase and histone H3 di- and trimethylation at lysine 36. The Journal of biological chemistry *287*, 3249-3256.

Fuchs, S.M., Laribee, R.N., and Strahl, B.D. (2009). Protein modifications in transcription elongation. Biochimica et biophysica acta *1789*, 26-36.

Gardner, K.E., Allis, C.D., and Strahl, B.D. (2011). Operating on chromatin, a colorful language where context matters. Journal of molecular biology *409*, 36-46.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P.*, et al.* (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine *366*, 883-892.

Gilbert, T.M., McDaniel, S.L., Byrum, S.D., Cades, J.A., Dancy, B.C., Wade, H., Tackett, A.J., Strahl, B.D., and Taverna, S.D. (2014). A PWWP domain-containing protein targets the NuA3 acetyltransferase complex via histone H3 lysine 36 trimethylation to coordinate transcriptional elongation at coding regions. Molecular & cellular proteomics : MCP *13*, 2883-2895.

Ginsburg, D.S., Anlembom, T.E., Wang, J., Patel, S.R., Li, B., and Hinnebusch, A.G. (2014). NuA4 links methylation of histone H3 lysines 4 and 36 to acetylation of histones H4 and H3. The Journal of biological chemistry *289*, 32656-32670.

Ginsburg, D.S., Govind, C.K., and Hinnebusch, A.G. (2009). NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5. Molecular and cellular biology *29*, 6473-6487.

Gonzalez, A., Shimobayashi, M., Eisenberg, T., Merle, D.A., Pendl, T., Hall, M.N., and Moustafa, T. (2015). TORC1 promotes phosphorylation of ribosomal protein S6 via the AGC kinase Ypk3 in Saccharomyces cerevisiae. PloS one *10*, e0120250.

Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics *15*, 305-308.

Govind, C.K., Qiu, H., Ginsburg, D.S., Ruan, C., Hofmeyer, K., Hu, C., Swaminathan, V., Workman, J.L., Li, B., and Hinnebusch, A.G. (2010). Phosphorylated Pol II CTD

recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. Molecular cell *39*, 234-246.

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. Molecular cell *25*, 31-42.

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. The Journal of biological chemistry *274*, 5895-5900.

Grosso, A.R., Leite, A.P., Carvalho, S., Matos, M.R., Martins, F.B., Vitor, A.C., Desterro, J.M., Carmo-Fonseca, M., and de Almeida, S.F. (2015). Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. eLife *4*.

Gunduz, M., Gunduz, E., Rivera, R.S., and Nagatsuka, H. (2008). The inhibitor of growth (ING) gene family: potential role in cancer therapy. Current cancer drug targets *8*, 275-284.

Guo, R., Zheng, L., Park, J.W., Lv, R., Chen, H., Jiao, F., Xu, W., Mu, S., Wen, H., Qiu, J.*, et al.* (2014). BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. Molecular cell *56*, 298-310.

Hacker, K.E., Fahey, C.C., Shinsky, S.A., Chiang, Y.J., DiFiore, J.V., Jha, D.K., Vo, A.H., Shavit, J.A., Davis, I.J., Strahl, B.D.*, et al.* (2016). Structure/Function Analysis of Recurrent Mutations in SETD2 Reveals a Critical and Conserved Role for a SET Domain Residue in Maintaining Protein Stability and H3K36 Trimethylation. The Journal of biological chemistry.

Hamilton, B., Dong, Y., Shindo, M., Liu, W., Odell, I., Ruvkun, G., and Lee, S.S. (2005). A systematic RNAi screen for longevity genes in C. elegans. Genes & development *19*, 1544-1555.

Heitman, J., Movva, N.R., and Hall, M.N. (1991). Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. Science *253*, 905-909.

Helin, K., and Dhanak, D. (2013). Chromatin proteins and modifications as drug targets. Nature *502*, 480-488.

Heyse, K.S., Weber, S.E., and Lipps, H.J. (2009). Histone modifications are specifically relocated during gene activation and nuclear differentiation. BMC genomics *10*, 554.

Ho, J.W., Jung, Y.L., Liu, T., Alver, B.H., Lee, S., Ikegami, K., Sohn, K.A., Minoda, A., Tolstorukov, M.Y., Appert, A.*, et al.* (2014). Comparative analysis of metazoan chromatin organization. Nature *512*, 449-452.

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 & development *15*, 3144-3154.

Howe, L., Kusch, T., Muster, N., Chaterji, R., Yates, J.R., 3rd, and Workman, J.L. (2002). Yng1p modulates the activity of Sas3p as a component of the yeast NuA3 Hhistone acetyltransferase complex. Molecular and cellular biology *22*, 5047-5053.

Hu, Z., Chen, K., Xia, Z., Chavez, M., Pal, S., Seol, J.H., Chen, C.C., Li, W., and Tyler, J.K. (2014). Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. Genes & development *28*, 396-408.

Huber, F., Bunina, D., Gupta, I., Khmelinskii, A., Meurer, M., Theer, P., Steinmetz, L.M., and Knop, M. (2016). Protein Abundance Control by Non-coding Antisense Transcription. Cell reports *15*, 2625-2636.

Huh, J.W., Wu, J., Lee, C.H., Yun, M., Gilada, D., Brautigam, C.A., and Li, B. (2012). Multivalent di-nucleosome recognition enables the Rpd3S histone deacetylase complex to tolerate decreased H3K36 methylation levels. The EMBO journal *31*, 3564-3574.

Ishimi, Y., and Kikuchi, A. (1991). Identification and molecular cloning of yeast homolog of nucleosome assembly protein I which facilitates nucleosome assembly in vitro. The Journal of biological chemistry *266*, 7025-7029.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E.*, et al.* (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947-962.

Jha, D.K., and Strahl, B.D. (2014). An RNA polymerase II-coupled function for histone H3K36 methylation in checkpoint activation and DSB repair. Nature communications *5*, 3965.

John, S., Howe, L., Tafrov, S.T., Grant, P.A., Sternglanz, R., and Workman, J.L. (2000). The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex. Genes & development *14*, 1196-1208.

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

Kabani, M., Michot, K., Boschiero, C., and Werner, M. (2005). Anc1 interacts with the catalytic subunits of the general transcription factors TFIID and TFIIF, the chromatin remodeling complexes RSC and INO80, and the histone acetyltransferase complex NuA3. Biochemical and biophysical research communications *332*, 398-403.

Kanu, N., Gronroos, E., Martinez, P., Burrell, R.A., Yi Goh, X., Bartkova, J., Maya-Mendoza, A., Mistrik, M., Rowan, A.J., Patel, H.*, et al.* (2015). SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. Oncogene *34*, 5699-5708.

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

Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Analytical chemistry *74*, 5383-5392.

Kelley, L.A., and Sternberg, M.J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nature protocols *4*, 363-371.

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.

Keogh, M.C., Mennella, T.A., Sawa, C., Berthelet, S., Krogan, N.J., Wolek, A., Podolny, V., Carpenter, L.R., Greenblatt, J.F., Baetz, K.*, et al.* (2006). The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes & development *20*, 660-665.

Keogh, M.C., Podolny, V., and Buratowski, S. (2003). Bur1 kinase is required for efficient transcription elongation by RNA polymerase II. Molecular and cellular biology *23*, 7005-7018.

Kimura, M., Suzuki, H., and Ishihama, A. (2002). Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in Schizosaccharomyces pombe involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. Molecular and cellular biology *22*, 1577-1588.

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. Molecular and cellular biology *25*, 3305-3316.

Klein, B.J., Lalonde, M.E., Cote, J., Yang, X.J., and Kutateladze, T.G. (2013). Crosstalk between epigenetic readers regulates the MOZ/MORF HAT complexes. Epigenetics : official journal of the DNA Methylation Society *9*.

Kremer, S.B., and Gross, D.S. (2009). SAGA and Rpd3 chromatin modification complexes dynamically regulate heat shock gene structure and expression. The Journal of biological chemistry *284*, 32914-32931.

Kristjuhan, A., and Svejstrup, J.Q. (2004). Evidence for distinct mechanisms facilitating transcript elongation through chromatin in vivo. The EMBO journal *23*, 4243-4252.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P.*, et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

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. Molecular and cellular biology *23*, 4207-4218.

Kulaeva, O.I., Gaykalova, D.A., Pestov, N.A., Golovastov, V.V., Vassylyev, D.G., Artsimovitch, I., and Studitsky, V.M. (2009). Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II. Nature structural & molecular biology *16*, 1272-1278.

Kulaeva, O.I., Hsieh, F.K., and Studitsky, V.M. (2010). RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. Proceedings of the National Academy of Sciences of the United States of America *107*, 11325-11330.

Kumar, G.S., Chang, W., Xie, T., Patel, A., Zhang, Y., Wang, G.G., David, G., and Radhakrishnan, I. (2012). Sequence requirements for combinatorial recognition of histone H3 by the MRG15 and Pf1 subunits of the Rpd3S/Sin3S corepressor complex. Journal of molecular biology *422*, 519-531.

Kuo, Y.M., and Andrews, A.J. (2013). Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3. PloS one *8*, e54896.

Kwan, A.H., Gell, D.A., Verger, A., Crossley, M., Matthews, J.M., and Mackay, J.P. (2003). Engineering a protein scaffold from a PHD finger. Structure *11*, 803-813.

Lalonde, M.E., Avvakumov, N., Glass, K.C., Joncas, F.H., Saksouk, N., Holliday, M., Paquet, E., Yan, K., Tong, Q., Klein, B.J.*, et al.* (2013). Exchange of associated factors directs a switch in HBO1 acetyltransferase histone tail specificity. Genes & development *27*, 2009-2024.

Lalonde, M.E., Cheng, X., and Cote, J. (2014). Histone target selection within chromatin: an exemplary case of teamwork. Genes & development *28*, 1029-1041.

Laue, K., Daujat, S., Crump, J.G., Plaster, N., Roehl, H.H., Tubingen Screen, C., Kimmel, C.B., Schneider, R., and Hammerschmidt, M. (2008). The multidomain protein Brpf1 binds histones and is required for Hox gene expression and segmental identity. Development *135*, 1935-1946.

Lee, C.H., Wu, J., and Li, B. (2013). Chromatin remodelers fine-tune H3K36me-directed deacetylation of neighbor nucleosomes by Rpd3S. Molecular cell *52*, 255-263.

Lee, J.S., and Shilatifard, A. (2007). A site to remember: H3K36 methylation a mark for histone deacetylation. Mutation research *618*, 130-134.

Lenstra, T.L., Benschop, J.J., Kim, T., Schulze, J.M., Brabers, N.A., Margaritis, T., van de Pasch, L.A., van Heesch, S.A., Brok, M.O., Groot Koerkamp, M.J.*, et al.* (2011). The specificity and topology of chromatin interaction pathways in yeast. Molecular cell *42*, 536-549.

Li, B., Gogol, M., Carey, M., Lee, D., Seidel, C., and Workman, J.L. (2007a). Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. Science *316*, 1050-1054.

Li, B., Gogol, M., Carey, M., Pattenden, S.G., Seidel, C., and Workman, J.L. (2007b). Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. Genes & development *21*, 1422-1430.

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. The Journal of biological chemistry *278*, 8897-8903.

Li, B., Jackson, J., Simon, M.D., Fleharty, B., Gogol, M., Seidel, C., Workman, J.L., and Shilatifard, A. (2009). Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. The Journal of biological chemistry *284*, 7970-7976.

Li, J., Moazed, D., and Gygi, S.P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. The Journal of biological chemistry *277*, 49383-49388.

Li, M., Phatnani, H.P., Guan, Z., Sage, H., Greenleaf, A.L., and Zhou, P. (2005). Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. Proceedings of the National Academy of Sciences of the United States of America *102*, 17636-17641.

Liang, G., Lin, J.C., Wei, V., Yoo, C., Cheng, J.C., Nguyen, C.T., Weisenberger, D.J., Egger, G., Takai, D., Gonzales, F.A.*, et al.* (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proceedings of the National Academy of Sciences of the United States of America *101*, 7357-7362.

Lickwar, C.R., Rao, B., Shabalin, A.A., Nobel, A.B., Strahl, B.D., and Lieb, J.D. (2009). The Set2/Rpd3S pathway suppresses cryptic transcription without regard to gene length or transcription frequency. PloS one *4*, e4886.

Lin, L.J., Minard, L.V., Johnston, G.C., Singer, R.A., and Schultz, M.C. (2010). Asf1 can promote trimethylation of H3 K36 by Set2. Molecular and cellular biology *30*, 1116-1129.

Liu, Y., Warfield, L., Zhang, C., Luo, J., Allen, J., Lang, W.H., Ranish, J., Shokat, K.M., and Hahn, S. (2009). Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. Molecular and cellular biology *29*, 4852-4863.

Lu, C., Jain, S.U., Hoelper, D., Bechet, D., Molden, R.C., Ran, L., Murphy, D., Venneti, S., Hameed, M., Pawel, B.R.*, et al.* (2016). Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. Science *352*, 844-849.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science *327*, 996-1000.

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.

Maltby, V.E., Martin, B.J., Schulze, J.M., Johnson, I., Hentrich, T., Sharma, A., Kobor, M.S., and Howe, L. (2012). Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. Molecular and cellular biology *32*, 3479-3485.

Mar, B.G., Bullinger, L.B., McLean, K.M., Grauman, P.V., Harris, M.H., Stevenson, K., Neuberg, D.S., Sinha, A.U., Sallan, S.E., Silverman, L.B.*, et al.* (2014). Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. Nature communications *5*, 3469.

Martin, D.G., Baetz, K., Shi, X., Walter, K.L., MacDonald, V.E., Wlodarski, M.J., Gozani, O., Hieter, P., and Howe, L. (2006a). The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. Molecular and cellular biology *26*, 7871-7879.

Martin, D.G., Grimes, D.E., Baetz, K., and Howe, L. (2006b). Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. Molecular and cellular biology *26*, 3018-3028.

Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C.P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends in biochemical sciences *28*, 69-74.

Morris, S.A., Shibata, Y., Noma, K., Tsukamoto, Y., Warren, E., Temple, B., Grewal, S.I., and Strahl, B.D. (2005). Histone H3 K36 methylation is associated with transcription elongation in Schizosaccharomyces pombe. Eukaryotic cell *4*, 1446-1454.

Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S.*, et al.* (2015). In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. eLife *4*, e06205.

Musselman, C.A., Lalonde, M.E., Cote, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. Nature structural & molecular biology *19*, 1218-1227.

Natoli, G. (2009). Control of NF-kappaB-dependent transcriptional responses by chromatin organization. Cold Spring Harbor perspectives in biology *1*, a000224.

Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *457*, 1038-1042.

Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. Analytical chemistry *75*, 4646-4658.

Pai, C.C., Deegan, R.S., Subramanian, L., Gal, C., Sarkar, S., Blaikley, E.J., Walker, C., Hulme, L., Bernhard, E., Codlin, S.*, et al.* (2014). A histone H3K36 chromatin switch coordinates DNA double-strand break repair pathway choice. Nature communications *5*, 4091.

Park, D., Shivram, H., and Iyer, V.R. (2014). Chd1 co-localizes with early transcription elongation factors independently of H3K36 methylation and releases stalled RNA polymerase II at introns. Epigenetics & chromatin *7*, 32.

Park, Y.J., Chodaparambil, J.V., Bao, Y., McBryant, S.J., and Luger, K. (2005). Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. The Journal of biological chemistry *280*, 1817-1825.

Petesch, S.J., and Lis, J.T. (2012). Overcoming the nucleosome barrier during transcript elongation. Trends in genetics : TIG *28*, 285-294.

Pfister, S.X., Ahrabi, S., Zalmas, L.P., Sarkar, S., Aymard, F., Bachrati, C.Z., Helleday, T., Legube, G., La Thangue, N.B., Porter, A.C.*, et al.* (2014). SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. Cell reports *7*, 2006-2018.

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.

Poon, B.P., and Mekhail, K. (2011). Cohesin and related coiled-coil domain-containing complexes physically and functionally connect the dots across the genome. Cell cycle *10*, 2669-2682.

Powers, T., and Walter, P. (1999). Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. Molecular biology of the cell *10*, 987-1000.

Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., and Bickmore, W.A. (2012). Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS genetics *8*, e1002717.

Pryde, F., Jain, D., Kerr, A., Curley, R., Mariotti, F.R., and Vogelauer, M. (2009). H3 k36 methylation helps determine the timing of cdc45 association with replication origins. PloS one *4*, e5882.

Pu, M., Ni, Z., Wang, M., Wang, X., Wood, J.G., Helfand, S.L., Yu, H., and Lee, S.S. (2015). Trimethylation of Lys36 on H3 restricts gene expression change during aging and impacts life span. Genes & development *29*, 718-731.

Pueschel, R., Coraggio, F., and Meister, P. (2016). From single genes to entire genomes: the search for a function of nuclear organization. Development *143*, 910-923.

Qiu, H., Hu, C., and Hinnebusch, A.G. (2009). Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. Molecular cell *33*, 752-762.

Qiu, Y., Zhang, W., Zhao, C., Wang, Y., Wang, W., Zhang, J., Zhang, Z., Li, G., Shi, Y., Tu, X.*, et al.* (2012). Solution structure of the Pdp1 PWWP domain reveals its unique binding sites for methylated H4K20 and DNA. The Biochemical journal *442*, 527-538.

Rando, O.J. (2007). Global patterns of histone modifications. Current opinion in genetics & development *17*, 94-99.

Reinke, A., Chen, J.C., Aronova, S., and Powers, T. (2006). Caffeine targets TOR complex I and provides evidence for a regulatory link between the FRB and kinase domains of Tor1p. The Journal of biological chemistry *281*, 31616-31626.

Rondelet, G., Dal Maso, T., Willems, L., and Wouters, J. (2016). Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. Journal of structural biology *194*, 357-367.

Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-Lovejoy, D., Martinez, J.Y., Bedford, M.T., Fuchs, S.M.*, et al.* (2012a). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. Nature structural & molecular biology *19*, 1155-1160.

Rothbart, S.B., Krajewski, K., Strahl, B.D., and Fuchs, S.M. (2012b). Peptide microarrays to interrogate the "histone code". Methods in enzymology *512*, 107-135.

Ruan, C., Cui, H., Lee, C.H., Li, S., and Li, B. (2016a). Homodimeric PHD-domain containing Rco1 constitutes a critical interaction hub within the Rpd3S histone deacetylase complex. The Journal of biological chemistry.

Ruan, C., Cui, H., Lee, C.H., Li, S., and Li, B. (2016b). Homodimeric PHD Domain-containing Rco1 Subunit Constitutes a Critical Interaction Hub within the Rpd3S Histone Deacetylase Complex. The Journal of biological chemistry *291*, 5428-5438.

Ruan, C., Lee, C.H., Cui, H., Li, S., and Li, B. (2015). Nucleosome contact triggers conformational changes of Rpd3S driving high-affinity H3K36me nucleosome engagement. Cell reports *10*, 204-215.

Rufiange, A., Jacques, P.E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. Molecular cell *27*, 393-405.

Sanso, M., Vargas-Perez, I., Quintales, L., Antequera, F., Ayte, J., and Hidalgo, E. (2011). Gcn5 facilitates Pol II progression, rather than recruitment to nucleosome-depleted stress promoters, in Schizosaccharomyces pombe. Nucleic acids research *39*, 6369-6379.

Sathianathan, A., Ravichandran, P., Lippi, J.M., Cohen, L., Messina, A., Shaju, S., Swede, M.J., and Ginsburg, D.S. (2016). The Eaf3/5/7 Subcomplex Stimulates NuA4 Interaction With Methylated H3K36 And RNA Polymerase II. The Journal of biological chemistry.

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 research *31*, 2475-2482.

Schwabish, M.A., and Struhl, K. (2006). Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. Molecular cell *22*, 415-422.

Sen, P., Dang, W., Donahue, G., Dai, J., Dorsey, J., Cao, X., Liu, W., Cao, K., Perry, R., Lee, J.Y.*, et al.* (2015). H3K36 methylation promotes longevity by enhancing transcriptional fidelity. Genes & development *29*, 1362-1376.

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. The Journal of biological chemistry *282*, 2450-2455.

Silva, A.C., Xu, X., Kim, H.S., Fillingham, J., Kislinger, T., Mennella, T.A., and Keogh, M.C. (2012). The replication-independent histone H3-H4 chaperones HIR, ASF1, and RTT106 co-operate to maintain promoter fidelity. The Journal of biological chemistry *287*, 1709-1718.

Simon, J.M., Hacker, K.E., Singh, D., Brannon, A.R., Parker, J.S., Weiser, M., Ho, T.H., Kuan, P.F., Jonasch, E., Furey, T.S.*, et al.* (2014). Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome research *24*, 241-250.

Smart, S.K., Mackintosh, S.G., Edmondson, R.D., Taverna, S.D., and Tackett, A.J. (2009). Mapping the local protein interactome of the NuA3 histone acetyltransferase. Protein science : a publication of the Protein Society *18*, 1987-1997.

Smolle, M., Venkatesh, S., Gogol, M.M., Li, H., Zhang, Y., Florens, L., Washburn, M.P., and Workman, J.L. (2012). Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. Nature structural & molecular biology *19*, 884-892.

Smolle, M., Workman, J.L., and Venkatesh, S. (2013). reSETting chromatin during transcription elongation. Epigenetics : official journal of the DNA Methylation Society *8*, 10-15.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic acids research *33*, W244-248.

Sorenson, M.R., Jha, D.K., Ucles, S.A., Flood, D.M., Strahl, B.D., Stevens, S.W., and Kress, T.L. (2016). Histone H3K36 methylation regulates pre-mRNA splicing in Saccharomyces cerevisiae. RNA biology *13*, 412-426.

Sterner, D.E., Belotserkovskaya, R., and Berger, S.L. (2002). SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. Proceedings of the National Academy of Sciences of the United States of America *99*, 11622-11627.

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.*, et al.* (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Molecular and cellular biology *22*, 1298-1306.

Sudol, M., Chen, H.I., Bougeret, C., Einbond, A., and Bork, P. (1995). Characterization of a novel protein-binding module--the WW domain. FEBS letters *369*, 67-71.

Sudol, M., Sliwa, K., and Russo, T. (2001). Functions of WW domains in the nucleus. FEBS letters *490*, 190-195.

Sun, B., Hong, J., Zhang, P., Dong, X., Shen, X., Lin, D., and Ding, J. (2008). Molecular basis of the interaction of Saccharomyces cerevisiae Eaf3 chromo domain with methylated H3K36. The Journal of biological chemistry *283*, 36504-36512.

Sun, M., Lariviere, L., Dengl, S., Mayer, A., and Cramer, P. (2010). A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD). The Journal of biological chemistry *285*, 41597-41603.

Sun, X.J., Wei, J., Wu, X.Y., Hu, M., Wang, L., Wang, H.H., Zhang, Q.H., Chen, S.J., Huang, Q.H., and Chen, Z. (2005). Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. The Journal of biological chemistry *280*, 35261-35271.

Tackett, A.J., DeGrasse, J.A., Sekedat, M.D., Oeffinger, M., Rout, M.P., and Chait, B.T. (2005a). I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. Journal of proteome research *4*, 1752-1756.

Tackett, A.J., Dilworth, D.J., Davey, M.J., O'Donnell, M., Aitchison, J.D., Rout, M.P., and Chait, B.T. (2005b). Proteomic and genomic characterization of chromatin complexes at a boundary. The Journal of cell biology *169*, 35-47.

Tadauchi, T., Matsumoto, K., Herskowitz, I., and Irie, K. (2001). Post-transcriptional regulation through the HO 3'-UTR by Mpt5, a yeast homolog of Pumilio and FBF. The EMBO journal *20*, 552-561.

Taverna, S.D., and Cole, P.A. (2010). Drug discovery: Reader's block. Nature *468*, 1050-1051.

Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T.*, et al.* (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Molecular cell *24*, 785-796.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature structural & molecular biology *14*, 1025-1040.

van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Ne, P., Loeillet, S.*, et al.* (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114-117.

VanDemark, A.P., Xin, H., McCullough, L., Rawlins, R., Bentley, S., Heroux, A., Stillman, D.J., Hill, C.P., and Formosa, T. (2008). Structural and functional analysis of the Spt16p N-terminal domain reveals overlapping roles of yFACT subunits. The Journal of biological chemistry *283*, 5058-5068.

Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P., Davies, H., Jones, D., Lin, M.L., Teague, J.*, et al.* (2011). Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature *469*, 539-542.

Vary, J.C., Jr., Gangaraju, V.K., Qin, J., Landel, C.C., Kooperberg, C., Bartholomew, B., and Tsukiyama, T. (2003). Yeast Isw1p forms two separable complexes in vivo. Molecular and cellular biology *23*, 80-91.

Venkatesh, S., Smolle, M., Li, H., Gogol, M.M., Saint, M., Kumar, S., Natarajan, K., and Workman, J.L. (2012). Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. Nature *489*, 452-455.

Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G.*, et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell *142*, 967-980.

Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T., Huntly, B.J., Gottgens, B., and Bycroft, M. (2010). Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. Nature structural & molecular biology *17*, 617-619.

Vojnic, E., Simon, B., Strahl, B.D., Sattler, M., and Cramer, P. (2006). Structure and carboxyl-terminal domain (CTD) binding of the Set2 SRI domain that couples histone H3 Lys36 methylation to transcription. The Journal of biological chemistry *281*, 13-15.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nature reviews Molecular cell biology *13*, 115-126.

Wang, S.S., Zhou, B.O., and Zhou, J.Q. (2011). Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. Molecular and cellular biology *31*, 3171-3181.

Wang, Y., Kallgren, S.P., Reddy, B.D., Kuntz, K., Lopez-Maury, L., Thompson, J., Watt, S., Ma, C., Hou, H., Shi, Y.*, et al.* (2012). Histone H3 lysine 14 acetylation is required for activation of a DNA damage checkpoint in fission yeast. The Journal of biological chemistry *287*, 4386-4393.

Weiner, A., Hsieh, T.H., Appleboim, A., Chen, H.V., Rahat, A., Amit, I., Rando, O.J., and Friedman, N. (2015). High-resolution chromatin dynamics during a yeast stress response. Molecular cell *58*, 371-386.

Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D., Tanaka, K., Ren, Y., Xia, Z., Wu, J*., et al.* (2014). ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature.

Winsor, T.S., Bartkowiak, B., Bennett, C.B., and Greenleaf, A.L. (2013). A DNA damage response system associated with the phosphoCTD of elongating RNA polymerase II. PloS one *8*, e60909.

Wu, H., Zeng, H., Lam, R., Tempel, W., Amaya, M.F., Xu, C., Dombrovski, L., Qiu, W., Wang, Y., and Min, J. (2011). Structural and histone binding ability characterizations of human PWWP domains. PloS one *6*, e18919.

Wyce, A., Xiao, T., Whelan, K.A., Kosman, C., Walter, W., Eick, D., Hughes, T.R., Krogan, N.J., Strahl, B.D., and Berger, S.L. (2007). H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. Molecular cell *27*, 275-288.

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.

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 & development *17*, 654-663.

Xu, C., Cui, G., Botuyan, M.V., and Mer, G. (2008). Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. Structure *16*, 1740-1750.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Munster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033-1037.

Yap, K.L., and Zhou, M.M. (2010). Keeping it in the family: diverse histone recognition by conserved structural folds. Critical reviews in biochemistry and molecular biology *45*, 488-505.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes & development *22*, 3422-3434.

Youdell, M.L., Kizer, K.O., Kisseleva-Romanova, E., Fuchs, S.M., Duro, E., Strahl, B.D., and Mellor, J. (2008). Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. Molecular and cellular biology *28*, 4915-4926.

Yun, M., Ruan, C., Huh, J.W., and Li, B. (2012). Reconstitution of modified chromatin templates for in vitro functional assays. Methods in molecular biology *833*, 237-253.

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

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. The EMBO journal *17*, 3155-3167.

Zhao, Y., and Garcia, B.A. (2015). Comprehensive Catalog of Currently Documented Histone Modifications. Cold Spring Harbor perspectives in biology *7*, a025064.

Zhu, X., He, F., Zeng, H., Ling, S., Chen, A., Wang, Y., Yan, X., Wei, W., Pang, Y., Cheng, H.*, et al.* (2014). Identification of functional cooperative mutations of SETD2 in human acute leukemia. Nature genetics *46*, 287-293.