INVESTIGATING THE ROLE OF THE DNA ENTRY-EXIT SITE OF THE NUCLEOSOME IN TRANSCRIPTION TERMINATION

by

Ashley Elizabeth Hildreth

B.S., North Georgia College & State University, 2012

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts & Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Ashley Elizabeth Hildreth

It was defended on

June 25, 2019

and approved by

Andrea Berman, PhD, Assistant Professor, Dept. of Biological Sciences

Kara Bernstein, PhD, Associate Professor, Dept. of Microbiology & Molecular Genetics

Deborah Chapman, PhD, Associate Professor, Dept. of Biological Sciences

Anthony Schwacha, PhD, Associate Professor, Dept. of Biological Sciences

Committee Chair: Karen Arndt, PhD, Professor, Dept. of Biological Sciences

INVESTIGATING THE ROLE OF THE DNA ENTRY-EXIT SITE OF THE NUCLEOSOME IN TRANSCRIPTION TERMINATION

Ashley Elizabeth Hildreth, PhD

University of Pittsburgh, 2019

Normal cellular function relies on the precise regulation of gene expression. In eukaryotes, DNA is packaged as chromatin, which acts as a barrier between the transcription machinery and genomic material. Chromatin consists of repeating nucleosomes, which contain approximately 147 base pairs of DNA surrounding an octamer of histone proteins H2A, H2B, H3, and H4. Transcription is controlled by factors that remove or modify nucleosomes, allowing RNA polymerase II to contact otherwise occluded DNA. The mechanisms by which nucleosome organization is controlled are well understood in regard to transcription initiation and elongation. Despite a few studies showing that transcription-coupled histone modifications and select chromatin remodelers are important for proper termination, little else is known about the role of chromatin at this step. Therefore, the goal of my dissertation work was to address how nucleosomes contribute to control of transcription termination. I conducted a genetic screen in *Saccharomyces cerevisiae* for histone mutations that cause termination defects. Interestingly, I found that many histone residues required for termination reside in or near the DNA entry-exit site of the nucleosome. The DNA entry-exit site is critical for nucleosome occupancy, a well-studied transcription-coupled histone modification, and prevention of several other transcription-related phenotypes that I have tested. Genome-wide analysis in mutants reveals altered nucleosome occupancy and transcriptional output. To test the hypothesis that stable nucleosomes within termination regions are required to act as physical roadblocks to the polymerase, I integrated a “superbinder” DNA sequence to

position a stable nucleosome at a candidate locus. Nucleosome occupancy increased substantially at the targeted location and suppressed termination read-through of the locus to the level of a wildtype control. Together, these data implicate the DNA entry-exit site as an important player in maintenance of chromatin organization that supports proper transcription, including termination. Further investigation is underway to determine whether the function of the DNA entry-exit site in transcription termination depends on interaction with other chromatin-related factors. These studies will open up exciting new avenues for pursuit by future researchers in the Arndt lab and will help the field to better understand how the critical final step of transcription, termination, is regulated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

It takes a village to raise a graduate student! I'd first like to thank my mentor, Dr. Karen Arndt, for her immutable enthusiasm for the work of her lab. From that very first meeting during my interview at Pitt, I knew Karen was the perfect mentor for me. I sat in her office, nervous about my interviews for the day, and she surprised me by standing up at her dry erase board and excitedly filling it with words and images I didn't yet quite understand, but knew from that moment that I wanted to. Since then, Karen has been a tremendous force in my scientific and personal growth by always encouraging confidence, curiosity, and determination. I am grateful to have studied in her lab during my graduate career, and to have gotten to know her over the past six years.

I would also like to thank Dr. Margaret Shirra, a truly formidable yeast geneticist, for being capable of solving quite literally any problem thrown her way. Peggy manages her own research projects, she keeps the lab organized and stocked, helps all of us troubleshoot when we're at the end of our ropes, and does so as one of the most positive people I've had the pleasure of interacting with from day-to-day. Known for her mantras, she's inspired me to learn something from every "failed" experiment, to just take the vacation sometimes, and to come up with solutions rather than complaints. On a more personal level, Peggy is a champion for local faire, especially food and craft beer. As a homebrewer, Peggy has been greatly influential in helping out my other "research" projects by always having the scoop on the newest beer releases.

I've watched many excellent lab mates come and go on to their next adventures during my time at Pitt. Christine Cucinotta, Beth Raupach, and Branden Van Oss were senior graduate students in the lab when I arrived, and they were all tremendously helpful in getting me acquainted with our lab, department, and graduate student life. Not only were they great friends while we

overlapped in lab, but it has been motivational to me as a younger scientist seeing their successes in graduate school and after. I've spent the most time in the lab with Mitch Ellison, a graduate student a year behind me in the program. Mitch has also been a great inspiration to me, having accomplished so much before even starting his PhD program. He spent time in the Army, finished his undergraduate work and a Master's degree in bioinformatics at lightning speed, and started his family all before beginning his work at Pitt. Mitch has been an excellent teacher in hard work, perseverance, patience, and balance, and I consider myself lucky to have worked with him. He and Brendan McShane, our lab technician, have also been wonderful friends during my time at Pitt, and are two people I'll miss tremendously when I leave. To Mitch, Brendan, and Alex Francette and Sarah Tripplehorn – our newest graduate students – I wish the best of luck in future academic and personal endeavors. I wish the same to our unparalleled undergraduates – Ellie Kerr, Alex Lederer, Logan Russell, Anna Weimer, Rachel Schusteff, Chelsea Guan, Rachel Kocik, Matt Blacksmith, Julia Seraly, Lauren Lotka, Patricia Donehue and Nicole Horan, many of whom have already gone to medical and graduate schools.

For invaluable advice throughout my graduate studies and for discussions that shaped the final draft of this dissertation, I'd like to thank my committee members – Dr. Andrea Berman, Dr. Kara Bernstein, Dr. Debbie Chapman, and Dr. Tony Schwacha. I am so grateful to have received their mentorship throughout my graduate work, especially because their broad interests and areas of expertise always reminded me to step away from my very focused work to see the bigger picture.

Finally, I would like to extend my deepest gratitude and admiration to those not in the Department for support through this process. My parents, Jack and Anita Hildreth, and my sister, Allison Hildreth, have always encouraged and supported me in whatever I planned to do with my future that day, week, or month. My parents trusted me to explore all of the opportunities that came

my way and make my mistakes, but they were always there to pick me back up when I made the really big ones. They taught us to never settle and to always strive for more because, as my mom often reminded us, we “could do anything we set our minds to.”

I would be remiss if I didn’t include my truly very good boy, Cooper, in here somewhere. Coop is a Lab/Border Collie mix that I adopted soon after passing my comprehensive exam. I was at a low in my science (a post-comps, third year slump) when I met Cooper, and he picked me right back up.

I’m grateful to my friends, old and new, who have always put up with my whacky work hours and have listened to me complain, “joke” about quitting, and then rediscover my unrequited love for science – Melissa Plakke and Douglas Hall, Sebastian Echeverri and Stella Chung, Megan Elstner and Kyle Frybarger, Kasi Ives, Laramie Bowcock, and so other many kind, wonderful people that are far too many to list here.

Last, but certainly not least, I’d like to thank from the bottom of my heart my chief partner in crime, my best friend, my drinking buddy, and my soulmate Sam Agosto. He kept me laughing when I didn’t want to, made me dinner when I was too busy to, and reminded me that not meeting my goals on time sometimes meant that I’d set them too high, not that I’d failed. He’s also told me how cool he thinks I am for being a scientist, which has been a major confidence booster. The life we’ve built together so far is something I’m so thankful for and so proud of. We’re engaged and expecting our first child, Parker. I can’t wait to face whatever comes next with them.

This dissertation is dedicated to my great grandmother, Elizabeth Hildreth, and my grandfather, Jack Hildreth, Sr.

1.0 INTRODUCTION

As the first step of gene expression, transcription is critical for the growth and survival of all types of cells. Transcription regulation from initiation to termination is required to produce accurate transcripts for further translation into the proteins that build and maintain the cell. While the initiation and elongation steps of transcription have been extensively studied, and the role of the chromatin template in the regulation of these steps has been intensely characterized, questions remain open as to how termination is regulated not only by various *trans*-factors, but also by the chromatin architecture of the locus being transcribed.

In recent years, genome-wide studies have allowed the field to glimpse a far greater eukaryotic transcriptome than was previously appreciated. A recent review suggests that approximately 85% of the yeast genome, and up to 98% of the human genome, is transcribed (TISSEUR *et al.* 2011). Functions of the many cryptic and antisense transcripts included in this tally have only begun to be characterized, but many of them are important in the regulation of expression of adjacent loci (JACQUIER 2009; HAINER AND MARTENS 2011b). In yeast, an essential termination pathway dependent upon Nrd1, Nab3, and Sen1, terminates many of these pervasive transcripts, suggesting an understanding of this pathway is needed now more than ever before. While this termination pathway is not perfectly conserved, similarities exist between yeast and humans in both protein factors and their functions, suggesting this knowledge can be extended to higher eukaryotes. I anticipate that new knowledge uncovered in this study will have important

clinical implications, especially in understanding cancers where transcriptional interference and aberrant transcription are of particular concern (BURGESS 2014).

1.1 **CHROMATIN STRUCTURE AND FUNCTION**

Eukaryotes must package extraordinary amounts of DNA (~2 meters stretched end-to-end in human cells) into a ~5-20 micron nucleus, requiring extensive condensation of DNA. To facilitate this, the eukaryotic genome is bundled into chromatin. Chromatin serves a dual purpose in eukaryotes: while its benefits in genome packaging are clear, it has become apparent that chromatin also plays important regulatory roles in nuclear processes such as transcription, replication, and recombination.

1.1.1 **Structure of the Nucleosome**

As early as the 1960s, structural and biochemical work had begun to define the means by which eukaryotic cells compact their genomes for nuclear storage. X-ray diffraction and electron microscopy studies of native, condensed DNA consistently report spherical units of approximately 100 Å in diameter or 100 Å thick fibers of DNA, respectively (PARDON *et al.* 1967; RICHARDS AND PARDON 1970). Repetition and periodicity of this molecular unit along the DNA fiber was confirmed by digestion of native chromatin by a calcium-magnesium-responsive endonuclease purified from mammalian nuclei, where sufficiently light digestion resulted in many chromatin fragments that were multiples of a base unit size (HEWISH AND BURGOYNE 1973). Not long after, it was discovered that about 200 bp of DNA and four well-conserved histone proteins – H2A, H2B,

H3, and H4 – were required to generate the X-ray diffraction pattern typical of native chromatin (KORNBERG 1974; KORNBERG AND THOMAS 1974).

The tertiary structure of the small, largely basic histone proteins consists of the histone-fold core and N- and C-terminal extensions. The histone-fold, shared by all four core histone proteins, consists of three alpha-helices and two intervening loops – α1, L1, α2, L2, α3 (Figure 1A) (ARENTS *et al.* 1991; ARENTS AND MOUDRIANAKIS 1995). It is through complementary histone-folds that H2A-H2B and H3-H4 pairs are able to dimerize (Figure 1B). The four-helix bundle that assembles nucleosomes from histone pairs consists of the α2 and α3 helices of two histone-fold heterodimers. To form the H3-H4 heterotetramer, an H3-H3 four-helix bundle assembles from two mirrored H3-H4 heterodimers (Figure 1C). Two H2A-H2B dimers assemble onto the H3-H4 heterotetramer through four-helix bundles between H2B and H4 (Figure 1D), completing the histone core (Figure 1E) (ARENTS *et al.* 1991; ARENTS AND MOUDRIANAKIS 1995; LUGER *et al.* 1997).

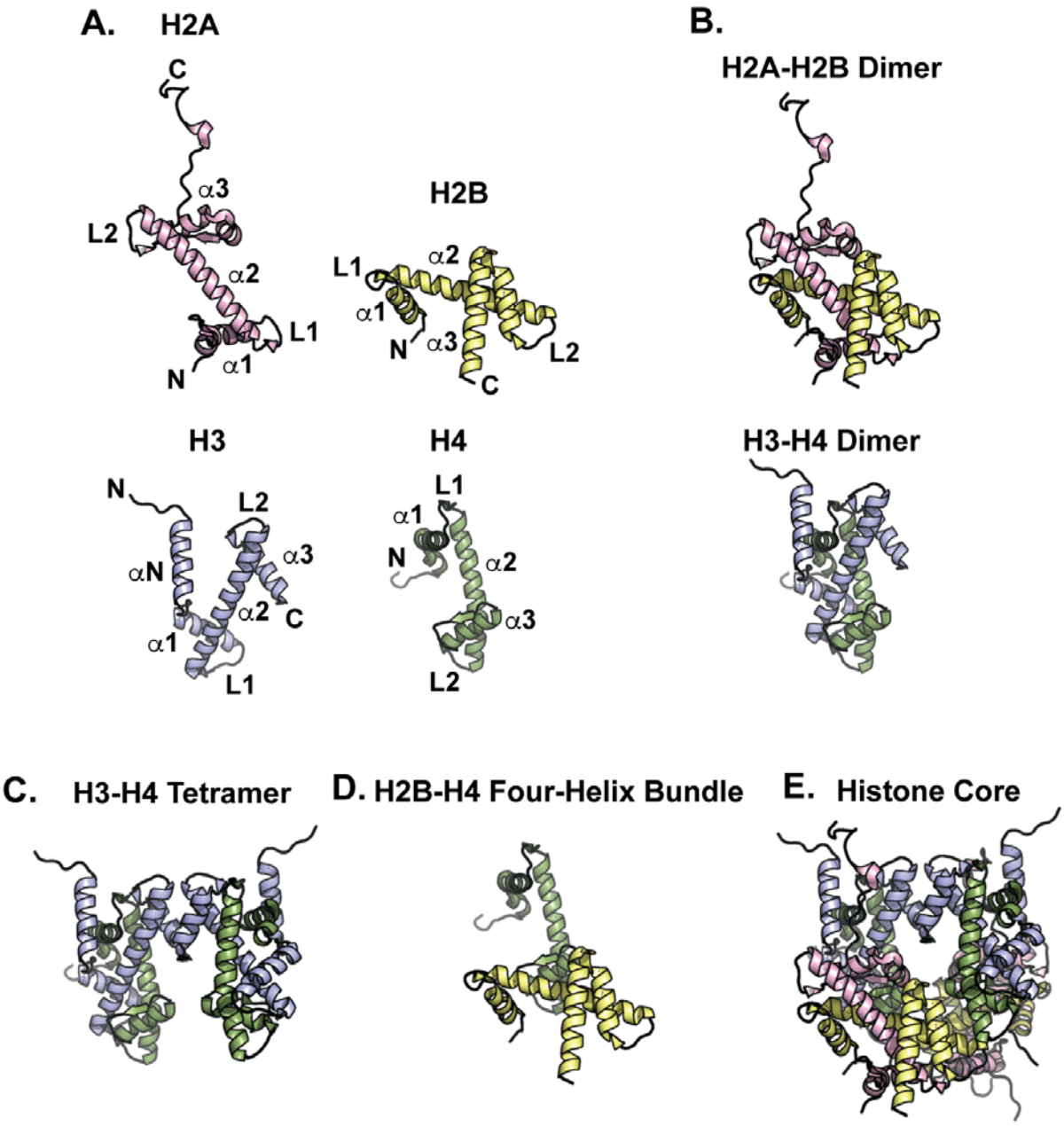

Figure 1. Histone and sub-nucleosome particle structures.

(A) Structures of the individual histone proteins. **(B)** Structures of the H2A-H2B dimer (top) and H3-H4 dimer (bottom). **(C)** Structure of the H3-H4 heterotetramer. **(D)** Structure of the H2B-H4 four-bundle helix. This ultimately assembles H2A-H2B dimers onto the H3-H4 heterotetramer. **(E)** The complete histone core. All structures modified in PyMol from PDB ID 1ID3 (LUGER *et al.* 1997).

Two decades after the original structural investigations of the nucleosome, and just a few years after detailed description of histone protein structure, a high resolution crystal structure of the nucleosome confirmed earlier discoveries and provided exquisite detail of inter- and intramolecular interactions required to maintain this important nucleoprotein (LUGER *et al.* 1997). Consistent with stoichiometry described in earlier studies, the crystal structure revealed that two each of the four core histones are required to form a nucleosome, as described above. DNA interacts with this histone core in two main ways: approximately 120 bp of the 147, on average, interact with histone-fold domains, and the remaining 13 bp on either face of the nucleosome interact with the N-terminal tail and αN helix of H3 (ZHOU *et al.* 2019). These interactions are mediated largely through charge interactions, where the negatively charged phosphodiester DNA backbone is coordinated by a largely positive histone surface (Figure 2) and hydrogen bonding (LUGER *et al.* 1997).

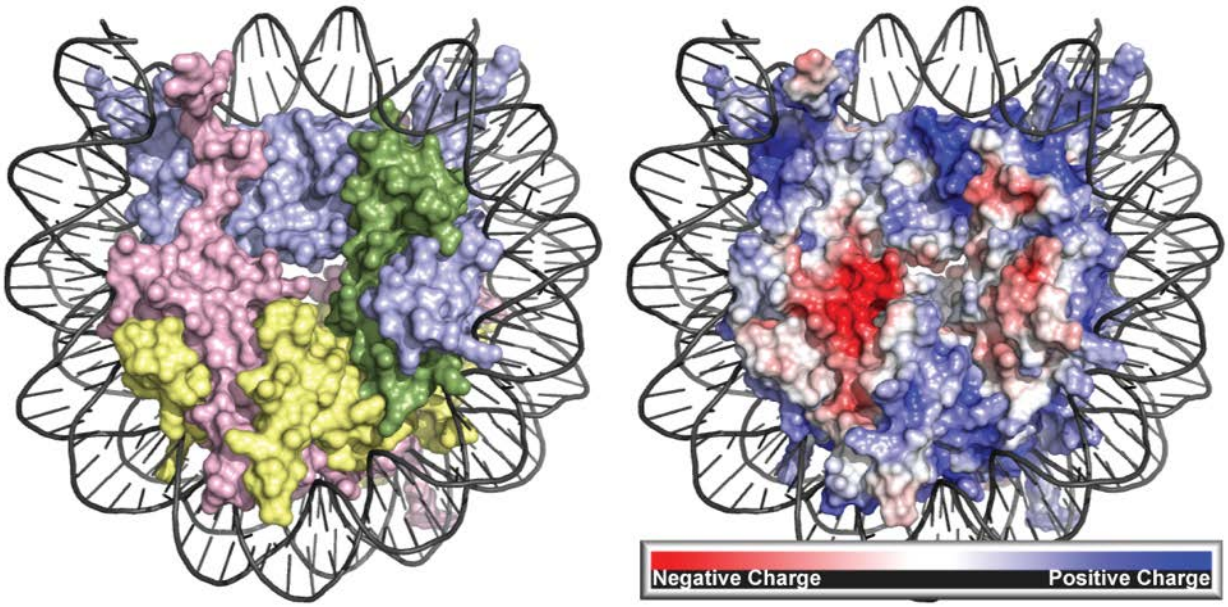

Figure 2. Structure of the complete nucleosome.

(A) Structure of the nucleosome colored by histone protein as in Figure 1. H3 is lilac, H4 is green, H2A is pink, and H2B is yellow. **(B)** Structure of the nucleosome colored by electrostatics. Red indicates acidic residues, blue coloration indicates basic residues. All structures modified in PyMol from PDB ID 1ID3 (LUGER *et al.* 1997).

1.1.2 The DNA Entry-Exit Site and Intrinsic Dynamism of the Nucleosome

1.1.2.1 DNA Entry-Exit Histone-DNA Interactions

The histone core acts as a protein spool for approximately 1.65 turns of double-helix DNA. At the two mirrored surfaces of the nucleosome, DNA makes its first and final direct contacts with the N-terminal tail and αN helix extension of H3 (LUGER *et al.* 1997; LUGER AND RICHMOND 1998). For this reason, this surface, termed the DNA entry-exit site (Figure 3A), is of key importance for the regulation of nucleosome accessibility. The H3 tail inserts into a channel formed by the juxtaposed minor grooves of nucleosome terminal (superhelix location (SHL) 7) and central (SHL

-1) strands, and this amino acid stretch of His39 through Val46 (with the exception of Gly44) directly contacts both (Figure 3B). In addition to these extensive protein-DNA interactions, Arg49 inserts into the minor groove at SHL 6.5, thus stabilizing a rather large section of entry-exit DNA.

Additional nucleosomal stability stems from the binding of non-core histone H1 (also referred to as H5 in early literature) (FRADO *et al.* 1983) at both DNA entry-exit sites of the nucleosome via interaction with the histones and linker DNA. This interaction prevents the binding of other factors to the histone residues or DNA at nucleosome ends by binding directly to those regions and sealing them (ALLAN *et al.* 1980). It's additionally thought that H1 binding contributes to the assembly of higher-order chromatin structures, especially in higher eukaryotes (THOMA AND KOLLER 1977; THOMA *et al.* 1979). Just like the core histones, histone H1 variants can be post-translationally modified and it is perhaps through these modifications that association with nucleosomes can be altered for the formation of heterochromatin, the regulation of gene expression, and the control of development (HERGETH AND SCHNEIDER 2015).

The protein-protein and protein-DNA interactions within the nucleosome yield a stable complex. However, the nucleosome is by no means static, and in fact undergoes a myriad of important chemical and structural changes, depending on local context. The nucleosome itself – especially interactions between its histone proteins and DNA – is quite dynamic. Additionally, chromatin modification factors can act on the nucleosome in various ways, including decorating the histones with post-translational modifications and even functioning as dramatically as to entirely change the positions of nucleosomes on DNA according to the cell's regulatory needs. The nucleosome in its native state is thought to be largely occluded from binding by these modifiers. Therefore, for DNA-templated processes to occur, interactions between DNA and histone proteins must be dynamic.

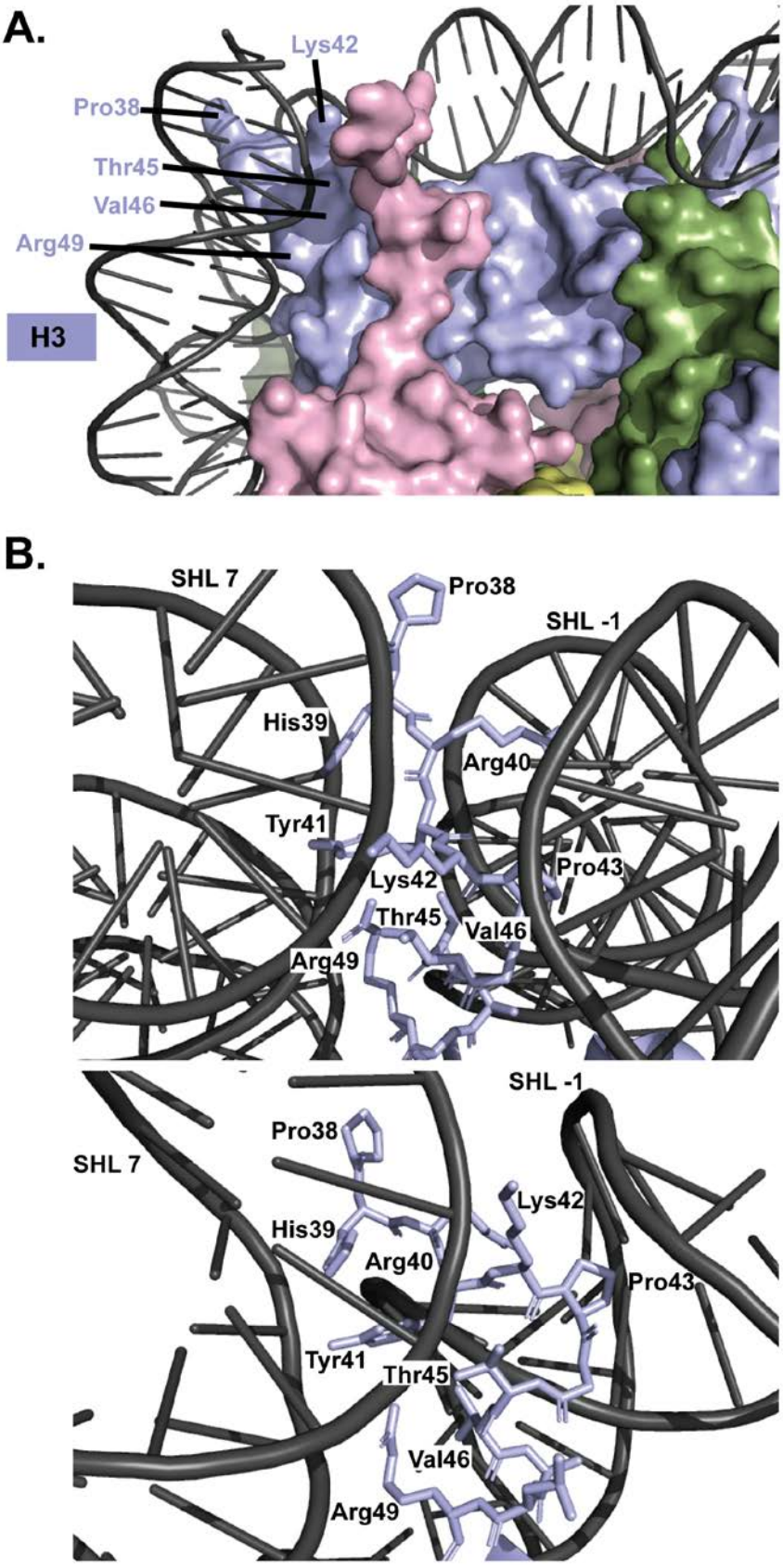
A.
Lys42
Pro38
Thr45
Val46
Arg49
H3
B.
SHL 7
Pro38
SHL -1
His39
Arg40
Tyr41
Pro43
Lys42
Thr45
Val46
Arg49
SHL -1
SHL 7
Pro38
Lys42
His39
Arg40
Pro43
Tyr41
Thr45
Val46
Arg49

Figure 3. DNA-histone interactions in the entry-exit site stabilize the nucleosome.

(A) Zoomed-in nucleosome structure to highlight the DNA entry-exit site. Histone coloration is as in Figure 1. **(B)** Top-down (top) and side-long (bottom) views of insertion of the N-terminal extension of H3 between coils of DNA on the nucleosome. All structures modified in PyMol from PDB ID 1ID3 (LUGER *et al.* 1997).

1.1.2.2 **DNA Breathing**

Early experiments measuring the kinetics of restriction digestion surprisingly reported non-zero equilibrium constants along the entire length of nucleosomal DNA (POLACH AND WIDOM 1995). While naked DNA was digested ~10^2-10^5-fold faster than nucleosomal DNA, restriction sites very near to the end of the particle were digested approximately ~10^2-10^5-fold faster than internal sites near the nucleosome dyad. Equilibrium constants measured for these external restriction sites suggest transient but rapid restriction site exposure via DNA unpeeling from the histone proteins. Another model for nucleosomal DNA exposure that attempts to explain non-zero accessibility of interior sites is that nucleosomal DNA transiently bulges out from interior contacts, leaving DNA ends in contact with the histones (POLACH AND WIDOM 1996).

The former model of DNA spontaneously unwrapping and rewrapping around the histone core is supported by fluorescence resonance energy transfer (FRET) experiments where nucleosomes were assembled with a donor dye molecule positioned at the 5' end of a nucleosome positioning sequence and the acceptor dye at DNA entry-exit site histone residues in either H3 or in nearby H2A (LI AND WIDOM 2004). By measuring loss of FRET efficiency between these fluorescent dye pairs under physiological conditions, it was surmised that DNA ends do indeed move away from, and back to, the histone core. This innate ability of the nucleosome to rapidly change its conformation was also found to be sufficient for invasion and binding of nucleosomal DNA by transcription repressor LexA, and the presence of the DNA-binding protein in fact drove

equilibrium toward the more open state (LI AND WIDOM 2004). Chromatin factors that must invade the nucleosome to modify its structure further likely capitalize on this intrinsic equilibrium to carry out their functions.

A later study used a combination of a very sensitive variation on FRET (stopped-flow FRET) and fluorescence correlation spectroscopy (FCS) on small populations of nucleosomes to measure the kinetics of DNA wrapping and unwrapping from the histone core (LI *et al.* 2005). Using these methods, it was determined that nucleosomes spend more time in their wrapped state (~250 ms) than in the unwrapped state (~10-50 ms). Transitions between these states likely provide biologically relevant time-scales for binding of *trans*-acting factors involved in chromatin modification, as well as an important regulator to processive enzymes like RNA polymerase II (Pol II) during transcription. The unwrapping rate constant, on average, is about 4 s^{-1}, while rewrapping is an exceedingly fast 20-90 s^{-1}. Li *et al.* argued that nucleosome rewrapping within 10-50 ms of unwrapping poses a significant kinetic barrier to the polymerase, which only transcribes ~23 nucleotides per second (SHERMOEN AND O'FARRELL 1991). They proposed that this barrier is likely overcome by elongation factors that poise and stabilize nucleosomes in their open states until Pol II can enter the nucleosome and prevent rewrapping via active transcription.

As a major interaction partner with entry-exit site DNA, it is no surprise that the H3 N-terminal tail plays a major role in stabilizing these DNA ends. FRET and spectroscopy experiments reveal that clipping the H3 tail greatly increases DNA accessibility within the nucleosome (NURSE *et al.* 2013). Interestingly, this study also showed that, while the H4 tail isn't directly involved in restricting DNA breathing, clipping both H3 and H4 tails has an additive effect on DNA accessibility *in vitro*. The authors argued that this is possibly due to post-translational modification crosstalk between the H3 and H4 tails (WANG *et al.* 2009; ZIPPO *et al.* 2009) that may play some

role in stabilizing the H3 tail conformation for DNA interactions. Cryo-EM studies have examined changes to the nucleosome throughout the DNA breathing cycle and have demonstrated shifts in conformation not only of the DNA molecule, but of the histone octamer (BILOKAPIC *et al.* 2018). The first conformational transition consists of DNA bulging away from the αN helix of H3, but remaining in contact with the H3 tail, explaining how loss of the H3 tail might render the nucleosome less stable and more amenable to breathing. Interestingly, this bulged DNA state resembles the structure of the nucleosome with chromatin remodeler Snf2 bound, perhaps suggesting that this partially unwrapped state is required for nucleosome remodeler binding (LIU *et al.* 2017).

Even on symmetrical stretches of DNA, nucleosomes exhibit asymmetrical DNA breathing (NGO *et al.* 2015; CHEN *et al.* 2017). Studies showing this have proposed that this asymmetrical breathing is likely to maintain the integrity of the nucleosome even during processes like transcription. Once DNA reaches its completely breathed state, where approximately 15 bp have severed contact with the histones, a dramatic conformational change occurs within the histone octamer (BILOKAPIC *et al.* 2018). The H3 αN helix on the unwrapped side appears shorter in cryo-EM imaging than in the bulged or unwrapped state and appears to tilt toward the dissociated DNA. This αN helix shift induces changes in positioning of the α1 and α2 helices on the unwrapped side, which tilts the H3 α2 helix on the wrapped side. This conformational change pulls the dyad DNA closer to the histone core, thus condensing and stabilizing DNA on the wrapped side. H2A-H2B conformation also changes to accommodate DNA wrapping. In the unwrapped state, the last histone-DNA contacts are between H2A L2 and H2B L1 with SHL 5.5. These loops tilt toward the DNA to stabilize these interactions, possibly to prevent further unwrapping (BILOKAPIC *et al.* 2018).

1.1.2.3 **Regulatory Post-Translational Modifications of the DNA Entry-Exit Site**

In yeast, work from the lab of Jef Boeke showed that H3K42 makes critical contact with entry-exit site DNA that can be regulated through methylation of this residue (HYLAND *et al.* 2011). Substitution of this residue for an alanine has pleiotropic effects, ranging from deregulated gene expression to sensitivity to drugs that classically identify cell cycle defects (Benomyl, a microtubule destabilizer) and DNA repair defects (hydroxyurea and methyl-methanesulfonate, DNA damaging agents), in line with the DNA entry-exit site's general role in nucleosome stability. Transcription defects in the H3 K42A mutant are myriad and pervasive, including changes in length of about 30% of transcripts, most of which appear 5'- and 3'-extended, and upregulation of about 60% of the transcriptome. Consistent with this residue's position within the DNA entry-exit site of the nucleosome, these effects are in line with destabilization of the "gatekeeping" function of histone-DNA interactions within this region that prevents DNA access, rendering nucleosomes more permissive to transcription. Importantly, this study also attributes further function to the DNA entry-exit site, including a propensity for post-translational modification (methylation of H3K42), and interaction with transcription elongation factors such as Set2, the histone methyltransferase for nearby H3K36, and broadly functioning Pol II elongation factors like the Paf1 complex and Spt6 (HYLAND *et al.* 2011).

In higher eukaryotes, a conserved arginine residue takes the place of the lysine at position 42 of histone H3. Despite the amino acid change, methylation of this residue is conserved from yeast to man, and appears to play the same role in gatekeeping histone-DNA interactions at the DNA entry-exit site to alter nucleosome accessibility (CASADIO *et al.* 2013). *In vivo* work in mouse and human cells, and *in vitro* studies on recombinant human nucleosomes, show that H3R42me^2, redundantly catalyzed by CARM1 and PRMT6, is a direct regulator of transcriptional activity

presumably through disruption of electrostatic interactions between entry-exit DNA and histone H3. *In silico* work supports this model by showing, via molecular dynamics simulations, a profound shift in equilibrium toward a DNA-exposed state when H3K42 is dimethylated, and a greater opening angle between the DNA and histone core compared to an unmodified H3K42 nucleosome (LI AND KONO 2018).

Evidence for direct binding of the histone methyltransferase that modifies H3K42 in yeast has not been detected at the DNA entry-exit site. However, there is genetic and structural evidence for the interaction of some chromatin modifiers with this region. For instance, the histone methyltransferase Set2, which catalyzes methylation of H3K36, has reduced occupancy in yeast cells where the DNA entry-exit site is mutated (ENDO *et al.* 2012). Substitutions of residues in H3 and H2A near the entry-exit site, particularly those that are surface-exposed, directly contact DNA, or make histone-histone contacts, greatly reduce or completely abolish H3K36me$^{2/3}$ (DU *et al.* 2008; DU AND BRIGGS 2010; HAINER AND MARTENS 2011a; ENDO *et al.* 2012). These residues are hypothesized to be critical not only in regulating DNA accessibility at the entry-exit site to allow for Pol II invasion, but also in supporting interactions between the H3-H4 tetramer with H2A-H2B dimers that may be important for further nucleosome remodeling via dimer eviction during active transcription elongation (BELOTSERKOVSKAYA *et al.* 2003). Interestingly, hydrophilic, surface-exposed residues required for H3K36 methylation are modifiable (H3T45, H3R49, and H3R52), suggesting the tempting hypothesis that transient modification of these residues contributes to the ever-growing histone code that carefully coordinates chromatin-templated processes (LAWRENCE *et al.* 2016).

While the modifications of H3R49 and H3R52 have not yet been described, phosphorylation of H3T45 and nearby Y41 have been attributed to facilitate similar nucleosome-

opening functions as other modifications described here, each for specific nuclear processes. Phosphorylation of H3T45 occurs during DNA replication via the S-phase kinase Cdc7-Dbf4 (BAKER *et al.* 2010) and during apoptosis (BREHOVE *et al.* 2015). Nearby, Y41 phosphorylation increases DNA entry-exit site accessibility during transcription, especially via disassembly of promoter chromatin. Alone, H3Y41p increases nucleosome accessibility by about 3-fold. However, in concert with acetylation of H3K56, promoter chromatin disassembly is approximately an order of magnitude more efficient (BREHOVE *et al.* 2015).

H3K56ac is involved in a multitude of nuclear processes, including replication-coupled nucleosome assembly, DNA repair, and transcription (RECHT *et al.* 2006). Mutation of H3K56 has a synthetic growth phenotype in combination with H3 T45A, suggesting parallel roles in DNA replication (BAKER *et al.* 2010). Acetylation of H3K56 requires that H3-H4 interact with the histone chaperone Asf1 prior to binding the histone acetyltransferase, Rtt109, which directly interacts with H3 residues in the DNA entry-exit site and the C-terminal tail of H4 (RECHT *et al.* 2006; FILLINGHAM *et al.* 2008; RADOVANI *et al.* 2013; ZHANG *et al.* 2018). This mark is placed during S-phase and persists through repair of replication-related DNA lesions (CHEN *et al.* 2008; FILLINGHAM *et al.* 2008) where it is thought to be responsible for increasing nucleosome disassembly by the mismatch repair machinery (NORTH *et al.* 2012; BREHOVE *et al.* 2015). Additionally, H3K56ac is found in high levels during promoter chromatin disassembly (WILLIAMS *et al.* 2008). Via studies on *PHO5,* it was determined that H3K56ac is likely involved in the rate-limiting step of promoter chromatin remodeling – removal of the H3-H4 tetramer – and that this process depends on this post-translational modification. Unlike many post-translational histone modifications that occur in the context of chromatin, K56-acetylated H3 is likely incorporated into promoter chromatin to signal for its disassembly and render this process more feasible, since

H3K56ac makes the nucleosome more accessible. Promoter chromatin not marked by H3K56ac, or removal of this acetyl group, is associated with gene repression (XU *et al.* 2007; WILLIAMS *et al.* 2008). Studies in yeast are consistent with the role of H3K56ac in transcriptional programs in human embryonic stem cells (XU *et al.* 2005; XIE *et al.* 2009).

Interestingly, methylation of H3K56 has also been described by mass spectrometry studies (GARCIA *et al.* 2007). *In vivo* and *in vitro* work suggests that this mark is conserved, identified so far in mouse, human, and *C. elegans*, and is found throughout the cell cycle with the exception of S-phase where H3K56 is largely acetylated (JACK *et al.* 2013). This mark, much like H3K9me^3 and H4K20me^3 is enriched in heterochromatin, but remains distinct from these signals by residing specifically in pericentromeric chromatin. Much like these heterochromatic marks, H3K56me^3 is placed by the histone methyltransferase Suv29h and removed by JMJD2 family deacetylases. The tradeoff of H3K56 markings during the cell cycle, where methylation is predominant in all stages but S phase, is thought to occur by newly synthesized H3, marked prior to chromatin incorporation by an acetyl group, turning over "old" methyl-marked H3. It is unclear whether single-cell eukaryotes share this H3 turnover pathway.

1.1.2.4 **Other Chromatin Modification Factors Interact with the DNA Entry-Exit Site**

Post-translational modifications within the DNA entry-exit site not only contribute to loosening of histone-DNA contacts to support nucleosome opening for DNA-templated processes, but they can also act as signals for particular chromatin modifying factors to properly position and time binding. The histone chaperone Rtt106 binds H3-H4 dimers where H3K56 is unmodified, but the affinity of this interaction is greatly increased when the acetylation mark is present (SU *et al.* 2012). Though Rtt106 does not appear to have a canonical acetyl-lysine reader domain, two separate structural domains do recognize H3-H4. The homodimerization domain (DD) of Rtt106 forms a

dimer that binds near the dyad at the H3-H3 four-helix bundle, and the pleckstrin-homology domains (PH) directly bind H3K56ac to participate in nucleosome remodeling for DNA replication, repair, and gene expression (SU *et al.* 2012; ZUNDER *et al.* 2012). A patch of basic residues on Rtt106, counterintuitively, is important for binding H3-H4. Through structural and biochemical work, it's hypothesized that this patch perhaps binds acidic residues on the nucleosome, or with the acidic C-terminal tail of Rtt106 prior to nucleosome binding. The second hypothesis is perhaps the best supported, since the two PH domains (located toward the C-terminus of the protein) are absolutely required for histone binding (ZUNDER *et al.* 2012). Binding to Rtt106 via its PH domains resembles binding of components of the FACT complex, Spt16M (middle domain) (HONDELE *et al.* 2013) and Pob3M (VANDEMARK *et al.* 2006), to histones.

The FACT complex binds histones through contacts between Spt16 and the H3 and H4 N-terminal tails and tail-less H3 and H4 globular domains (BELOTSERKOVSKAYA *et al.* 2003; STUWE *et al.* 2008), the H2A-H2B dimers that this complex is best known to chaperone, and the H3 αN helix (residues 46-65) (HONDELE *et al.* 2013). Pob3 also binds histone proteins in a manner structurally similar to Rtt106 and Spt16 using a PH domain (VANDEMARK *et al.* 2006). Synthetic slow-growth phenotypes occur in Pob3 mutations combined with deletions of the H3 and H4 N-terminal tails, suggesting a role for the flexible extensions of these histones in the interaction between Pob3 and the nucleosome. Interestingly, the FACT interaction with nucleosomes also depends on H3K56ac, though it does not bind only acetylated or deacetylated forms of H3. Instead, the correct balance of available H3K56 and H3K56ac nucleosomes is critical for FACT to perform its chaperoning functions for gene repression and prevention of antisense transcription (MCCULLOUGH *et al.* 2019).

The DNA entry-exit site and immediately adjacent linker DNA is also critical for binding ATP-dependent ISWI-containing chromatin remodelers (KAGALWALA *et al.* 2004; SCHWANBECK *et al.* 2004; STROHNER *et al.* 2005). Hydroxy-radical footprinting studies have shown that ISWI-containing chromatin remodelers like NURF in *D. melanogaster* protect the linker-DNA entry interface, suggesting its binding at that region (SCHWANBECK *et al.* 2004). Common to these chromatin remodelers is the need to stretch across the nucleosome for full binding and remodeling activity. For instance, ISW2 has been shown to bind the H4 tail, near the nucleosome dyad, and the DNA entry-exit site including the H3 tail and linker DNA (KAGALWALA *et al.* 2004). Much like other ISWI-like chromatin remodelers, affinity for the nucleosome by the remodeling factor increases with longer linker DNA. It has been shown biochemically that ISW2 binds asymmetrically to the side of the longer linker and preferentially slides nucleosomes in the direction of its binding to a more central position on a strand of DNA. This is unlike well-studied ISWI-containing chromatin remodeler ACF, which binds the nucleosome more symmetrically (STROHNER *et al.* 2005). Binding of this chromatin remodeler mirrors NURF, ISW2, and others like it, where it capitalizes on spontaneous nucleosome breathing to invade and stabilize the "open" conformation. Using hydroxy-radical footprinting and a combination of fluorescence spectroscopy techniques, studies of ACF defined the ISWI-like "loop-recapture" model (BROWER-TOLAND *et al.* 2002) for nucleosome binding (Figure 4). After DNA breathing loosens histone-DNA contacts at the entry-exit site, DNA is free to interact with neighboring histone surfaces that do not necessarily correspond to the original interaction site. This forms a loop of DNA off of the histone core, which ACF is free to bind to propagate this loop so that the translational position of the nucleosome changes equal to the size, in bp, of the original loop (STROHNER *et al.* 2005).

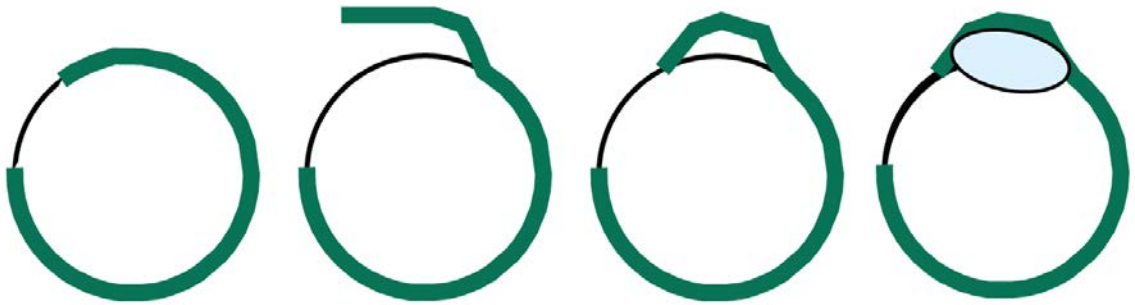

Figure 4. "Loop recapture" model for chromatin remodeling by ISWI-like factors.

DNA unwraps or "breathes" from the DNA entry-exit site, leaving DNA free to interact with neighboring histone surfaces other than the original interaction site. A loop of DNA forms off of the histone core, which chromatin factors may bind. Such a factor (here, a chromatin remodeler, blue oval) may propagate this loop so that the translational position of the nucleosome changes equal to the size, in bp, of the original loop.

1.1.3 **General Chromatin Modifications**

1.1.3.1 **Histone Modifications**

Nucleosomes are post-translationally modified by histone modifying complexes to signal different stages of cellular processes to factors that recognize these modification marks (KOUZARIDES 2007; RANDO AND WINSTON 2012). Histone modifications also alter histone-histone and histone-DNA contacts and make the DNA template more or less accessible. During transcription, the SAGA and NuA4 histone acetyltransferase (HAT) complexes acetylate H3 and H4, respectively, to allow for a more open chromatin state that can be further modified and accessed by the transcription machinery (ALLARD *et al.* 1999; GOVIND *et al.* 2007; WYCE *et al.* 2007; GINSBURG *et al.* 2009). HATs also have roles in other pathways like replication-coupled nucleosome assembly and DNA repair, such as Rtt109 (DRISCOLL *et al.* 2007; HAN *et al.* 2007; JESSULAT *et al.* 2008; IDE *et al.* 2013; VOICHEK *et al.* 2016). NuA3, whose catalytic subunit is Sas3 ("something about silencing"),

is found associated with methylated H3 K4 and H3 K36, as well as the histone chaperone FACT, suggesting its involvement in transcription and replication (JOHN *et al.* 2000; MARTIN *et al.* 2006; MARTIN *et al.* 2017).

In this open chromatin state, transcription elongation factors, histone modifiers, and other chromatin reorganizers interact with Pol II to arrange the chromatin template for effective transcription (HAHN AND YOUNG 2011; RANDO AND WINSTON 2012; TOMSON AND ARNDT 2013). One of these factors is the multifunctional Paf1 complex. The Paf1 complex regulates transcription-coupled ubiquitylation of H2B (H2Bub) by the ubiquitin conjugase, Rad6, and ligase, Bre1 (VAN OSS *et al.* 2017). A trend in histone modification is crosstalk, where some modifications are required for placement of other important marks. For instance, the H2B ubiquitylation mark is needed for downstream methylation of H3 on lysines 4 and 79 (H3K4me and H3K79me) by Set1 and Dot1, respectively (SUN AND ALLIS 2002; KROGAN *et al.* 2003a; NG *et al.* 2003b). H3K4 trimethylation ($H3K4me^3$) marks actively transcribed genes and is enriched at the 5' end of genes (BERNSTEIN *et al.* 2002; SANTOS-ROSA *et al.* 2002; LIU *et al.* 2005; POKHOLOK *et al.* 2005). H3K79me inhibits the binding of Sir proteins, transcriptional silencers, to active loci (VAN LEEUWEN *et al.* 2002; KROGAN *et al.* 2003a; NG *et al.* 2003a). Therefore, this mark doesn't correlate directly to transcription, but appears ubiquitously throughout the genome. Outside of its function in H2Bub, the Paf1 complex also facilitates Set2-mediated methylation of H3K36 (KROGAN *et al.* 2003a; NG *et al.* 2003a; CHU *et al.* 2007). This mark is enriched at the 3' end of genes and recruits the Rpd3S histone deacetylase complex (HDAC) to prevent initiation from cryptic promoters within ORFs (STRAHL *et al.* 2002; CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005; VENKATESH *et al.* 2012).

Table 1. Relevant histone modifications in yeast and higher eukaryotes.

Residue	Modification/Modifier	Abbreviation	Group	Function
H3 K4	Methylation/ COMPASS	H3K4me	Eukaryotes	• Marks active transcription – H3K4me^3 is enriched at 5' ends (BERNSTEIN *et al.* 2002; SANTOS-ROSA *et al.* 2002)
H3 K36	Methylation/ Set2	H3K36me	Eukaryotes	• Marks active transcription – H3K36me^3 is enriched at 3' ends (KROGAN *et al.* 2003b) • Prevents cryptic transcription (CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005)
H3 K42	Methylation/ Unknown	H3K42me	Yeast	• Transcription • DNA damage repair • Cell cycle control (HYLAND *et al.* 2011) • Stabilization of open nucleosome (LI AND KONO 2018)
H3 R42	Methylation/ CARM1, PRMT6	H3R42me	Humans	• Transcription • Stabilization of open nucleosome (CASADIO *et al.* 2013)
H3 T45	Phosphorylation/ Cdc7-Dbf4	H3T45p	Eukaryotes	• Replication (BAKER *et al.* 2010) • Apoptosis (BREHOVE *et al.* 2015)
H3 K56	Acetylation/ Rtt109	H3K56ac	Eukaryotes	• Replication-coupled nucleosome assembly (RECHT *et al.* 2006) • Transcription (XU *et al.* 2005; WILLIAMS *et al.* 2008; XIE *et al.* 2009)
H3 K79	Methylation/ Dot1	H3K79me	Eukaryotes	• Transcriptional silencing (VAN LEEUWEN *et al.* 2002)
H2B K120/K123	Ubiquitylation/ Rad6-Bre1	H2BK120ub/ H2BK123ub/ H2Bub	Humans/ Yeast	• Transcription (XIAO *et al.* 2005) • Promoter-proximal pausing

				(FUCHS *et al.* 2014; WU *et al.* 2014) • Methylation of H3 K4 and K79 (SUN AND ALLIS 2002; KROGAN *et al.* 2003a; NG *et al.* 2003b)

1.1.3.2 **Chromatin Remodelers**

During transcription, ATP-dependent chromatin remodelers can reposition, slide, or remove nucleosomes to facilitate the passage of Pol II. Often, these chromatin remodelers are recruited or perform their functions as a consequence of transcription-coupled histone modifications, discussed above. For instance, the SWI/SNF (SWItch/Sucrose Non-Fermentable) chromatin remodeler is involved in replacing nucleosomes at the promoters of some stress-inducible genes to promote transcription activator binding (BURNS AND PETERSON 1997; RAWAL *et al.* 2018). Histone acetylation by the Gcn5 subunit of the SAGA complex is required for the recruitment of the SWI/SNF complex in yeast (ROBERTS AND WINSTON 1997; QIU *et al.* 2016), and in fact appears to be required for SWI/SNF activity (SYNTICHAKI *et al.* 2000; HASSAN *et al.* 2001; CHANDY *et al.* 2006). Similarly, essential chromatin remodeler RSC (Remodels the Structure of Chromatin), an enzyme in the SWI/SNF family, requires interaction with acetylated lysines through six bromodomains whose mutations are synthetically lethal in combination with a Gcn5 mutant (KASTEN *et al.* 2004; CHATTERJEE *et al.* 2011). In addition to gene expression, RSC has crucial roles in DNA damage repair, replication, and even chromosome segregation (HSU *et al.* 2003; NIIMI *et al.* 2012; CZAJA *et al.* 2014).

Members of the ISWI (Imitation SWItch) family laterally reposition nucleosomes both near nucleosome-depleted regions (NDRs) to activate and inactivate sense and antisense transcription,

as well as along coding regions to promote transcription elongation. Isw2 moves nucleosomes to unfavorable intergenic sequences to prevent both canonical and pervasive transcription (WHITEHOUSE *et al.* 2007). Isw1, on the other hand, activates transcription by moving NDR-adjacent nucleosomes away from transcription factor binding sites (MORILLON *et al.* 2003). Isw1 is also thought to promote proper transcription elongation and termination by maintaining a repressive chromatin state over cryptic initiation sites (ALEN *et al.* 2002; MORILLON *et al.* 2003; VARY *et al.* 2003; TIROSH *et al.* 2010). Members of this family have been shown to contain PHD fingers, which interact with methylated lysines like H3K36 and, in some human ISWI-family members, H3K4me (WYSOCKA *et al.* 2006).

CHD (Chromodomain-Helicase-DNA binding) family remodelers are thought to share some redundant functions with members of the ISWI family (ALEN *et al.* 2002). These remodelers also share some similarities in their chromatin recruitment – one of the distinguishing features of a CHD remodeler is its N-terminal, tandem chromodomains that allow for binding to methylated histones (MARFELLA AND IMBALZANO 2007). Chd1, involved in all stages of transcription, is also recruited to the 5' end of genes by Mediator and the Paf1 complex for elongation-coupled nucleosome turnover (SIMIC *et al.* 2003; LIN *et al.* 2011) and at the 3' end of genes, possibly by H3K36me through the chromodomains, to repress cryptic initiation alongside Isw1 (LI *et al.* 2007; MALTBY *et al.* 2012; SMOLLE *et al.* 2012).

1.1.3.3 **Histone Chaperones**

Canonical histones and histone variants are added and removed from nucleosomes by histone chaperones at various stages of the cell's activities (JANSEN AND VERSTREPEN 2011). The histone chaperone Spt16-Pob3/FACT (Facilitates Chromatin Transcription) disassembles and reassembles nucleosomes for passage of Pol II (MARTIN *et al.* 2018). FACT's primary chaperoning activity is

through destabilization of interactions between H2A-H2B dimers and the H3-H4 tetramer, allowing it to evict H2A-H2B dimers from nucleosomes to render them more open to processes like transcription. FACT does this while maintaining the integrity of histone components for reassembly of nucleosomes in the wake of Pol II (CHEN *et al.* 2018b). Interestingly, FACT can specifically recognize Pol II-disturbed nucleosomes to perform its role in reassembly (MARTIN *et al.* 2018). In addition to its roles in transcription, FACT also regulates DNA replication and repair (WITTMEYER AND FORMOSA 1997; SCHLESINGER AND FORMOSA 2000; KELLER AND LU 2002; HEO *et al.* 2008). FACT is also capable of chaperoning H3-H4 tetramers, however, and has been shown to deposit both dimers and tetramers onto chromatin (BELOTSERKOVSKAYA *et al.* 2003).

Asf1 (Anti-Silencing Factor) chaperones newly synthesized H3-H4 to assemble these histones into the nucleosome in a replication-independent manner with the help of the HIR complex (GREEN *et al.* 2005). Asf1 also contributes to nucleosome assembly during DNA repair and replication (EMILI *et al.* 2001; HU *et al.* 2001; ZHANG *et al.* 2018), functions in transcriptional silencing (ADKINS *et al.* 2004; SHARP *et al.* 2005), and reassembles chromatin post-transcriptionally (SCHWABISH AND STRUHL 2006). Rtt106, discussed above, is a chaperone for histones H3 and H4. This factor binds preferentially to H3 acetylated at lysine 56 and has been shown to act as a repressor for some genes, including the histone genes, by assembling nucleosomes over their promoters (HUANG *et al.* 2005; LI *et al.* 2008).

Spt6 directly interacts with histones (BORTVIN AND WINSTON 1996; VOS *et al.* 2018; ZHU *et al.* 2018) and is required for normal chromatin structure during transcription (KAPLAN *et al.* 2003). In addition to the histones, Spt6 interacts directly with RNA polymerase II and Spt4/5, suggesting its chaperoning function is closely coupled to transcription elongation (ENDOH *et al.* 2004; SDANO *et al.* 2017; DRONAMRAJU *et al.* 2018a). Spt6 and FACT have both been extensively

characterized to promote transcription, but both histone chaperones are also critical for prevention of cryptic transcription by restricting mislocalization of histone H2A.Z, a partially destabilizing histone variant of H2A commonly found at regions of the genome that need to be easily accessed, like promoters (JERONIMO *et al.* 2015; DORIS *et al.* 2018).

1.2 **EUKARYOTIC TRANSCRIPTION**

The first critical step of gene expression is transcription, where genomic DNA must be accurately copied as RNA, a molecule better suited to exit the nucleus and perform such biological functions as providing a template for protein synthesis. Transcription can be broken into three main phases: initiation, elongation, and termination (SHANDILYA AND ROBERTS 2012). While eukaryotic genomes generally encode three DNA-dependent RNA polymerases, Pol II transcribes the vast majority of genes in *S. cerevisiae*, including protein-coding genes and many noncoding loci (KAPLAN 2013). Activity of this massive molecular machine, made up of 12 protein subunits and weighing approximately 1.6 MDa, is tightly regulated by the coordinated actions of several protein transcription factors binding to the chromatin template or the polymerase itself. The largest subunit of Pol II, Rpb1, has a C-terminal domain (CTD) that is differentially phosphorylated throughout the transcription cycle which serves as a dynamic platform for the recruitment of these stage-specific factors to modify the template, co-transcriptionally splice the nascent RNA, and signal for termination at the proper location, among other critical functions (HSIN AND MANLEY 2012; SHANDILYA AND ROBERTS 2012; KAPLAN 2013).

1.2.1 **Initiation**

At the single-gene level, transcription initiation is best understood at TATA-containing promoters, where a predictable assembly of general transcription factors (GTFs) assemble at the promoter to recruit Pol II and promote productive transcription (GRUNBERG AND HAHN 2013). Together, these GTFs form the preinitiation complex (PIC), which makes several important interactions not only with promoter DNA, but with the polymerase to form the scaffold necessary to initiate transcriptional activity. First, the transcription factor TBP (TATA-binding protein) binds the promoter at the TATAWAWR (W=A/T, R=A/G) consensus sequence (BASEHOAR *et al.* 2004). TBP can bind the promoter alone or as part of a larger complex, transcription factor IID (TFIID), which consists of TBP and the TBP-associated proteins (TAFs) (SAINSBURY *et al.* 2015). TFIID contributes to the expression of about 90% of Pol II-transcribed yeast genes (HUISINGA AND PUGH 2004; WARFIELD *et al.* 2017), and is able to do so even at those without canonical TATA sequences due to the ability of the TAFs to bind other core promoter DNA elements (PATEL *et al.* 2018). TFIIA binds the promoter and stabilizes the TBP-DNA complex (SAINSBURY *et al.* 2015).

TFIIB is one of the next factors to join the expanding PIC. In yeast, the essential *SUA7* gene encodes TFIIB. Some mutations to this gene not only dramatically slow cell growth, but also deregulate transcription start site (TSS) selection at many genes (PINTO *et al.* 1992; GOEL *et al.* 2012). Recognition of the TBP-DNA complex by the first of two imperfect, 76 amino acid repeats in the C-terminal domain of TFIIB recruits it to chromatin (HA *et al.* 1993; WANG AND ROBERTS 2010). Importantly, TFIIB then acts to recruit Pol II to the initiation complex through both C-terminal repeats, specifically through arginines 286, 290, and 295. Together, these interactions bridge the gap between promoter DNA and the unphosphorylated, inactive form of Pol II, thus appropriately placing it for promoter escape and productive transcription.

TFIIF interacts with the N-terminal domain of TFIIB (HA *et al.* 1993; LUSE 2012), and this interaction is thought to stabilize TFIIB within the elongation complex (CABART *et al.* 2011). While TFIIF is found associated with only approximately 50% of Pol II in yeast, and it is not absolutely essential for promoter escape and initiation of productive transcription, absence of this transcription factor results in loss of TFIIB association at approximately +12 bp from the transcription start site (TSS). TFIIF helps prevent nonspecific DNA binding by Pol II (CONAWAY *et al.* 1991; GHAZY *et al.* 2004; KHAPERSKYY *et al.* 2008) and, together with TFIIB, helps stabilize the transcription bubble upon its opening (PAN AND GREENBLATT 1994; LUSE 2012).

TFIIE and TFIIH are essential for opening DNA for transcription. Structural and biochemical data suggest that TFIIE alone can stabilize the open bubble at some promoter templates *in vitro* by binding to single-stranded DNA (HOLSTEGE *et al.* 1995). *In vivo*, the major function of TFIIE is thought to be its ability to recruit TFIIH to the PIC and stimulate its activity as the ATP-dependent DNA helicase that opens double-stranded DNA at the promoter (HOLSTEGE *et al.* 1996; CABART AND LUSE 2012). TFIIH functions as a non-canonical DNA helicase and instead acts almost as a "molecular wrench" that anchors on DNA downstream of the promoter and rotates, physically unwinding or melting the double helix to form the transcription bubble (KIM *et al.* 2000; GREBER *et al.* 2017; GREBER *et al.* 2019). Together with TFIIF and TFIIB, TFIIE stabilizes this open DNA conformation to allow for transcription initiation (Figure 5).

Finally, other critical factors that assemble at the PIC and regulate not only various aspects of initiation, but also other downstream transcription events, are the coactivators. Coactivator complexes present at almost all Pol II-transcribed genes include the Mediator and SAGA complexes. Mediator is a large, multi-subunit complex (21 subunits in *S. cerevisiae,* 26 in humans), that acts to help recruit the GTFs and "bridge" communication between GTFs and Pol II (HARPER

AND TAATJES 2018). Many of the GTFs do not contact Pol II directly, so this massive, modular complex helps assemble the complete PIC by recruiting Pol II through direct interaction with unphosphorylated CTD, and then stabilizes critical PIC interactions (ALLEN AND TAATJES 2015). Mediator contributes to retention of Pol II at the PIC until the Cdc7/Kin28 subunit of TFIIH phosphorylates serine 5 of the consensus CTD repeat YSPTSPS, which disrupts the Mediator-Pol II interaction and encourages promoter escape (JERONIMO AND ROBERT 2014; WONG *et al.* 2014).

SAGA is a highly modular coactivator complex that shares many subunits with other complexes, including TFIID (HELMLINGER AND TORA 2017). This large coactivator performs a variety of functions in regulating chromatin and other transcription factors through its multiple modules, including histone acetyltransferase (HAT) and deubiquitinase (DUB) modules for histone modification, a transcription factor binding module, and a splicing module all centered around the core structural module. Contrary to early belief that genes were either TFIID-dominated or SAGA-dominated in their expression (HUISINGA AND PUGH 2004), more recent evidence suggests that these two coactivators are general factors that regulate chromatin state and transcription at almost all yeast genes (BAPTISTA *et al.* 2017; WARFIELD *et al.* 2017).

Genome-wide experiments have recently shown that the PIC components described here can be mapped to promoters containing or lacking the canonical TATA consensus sequence (RHEE AND PUGH 2012). Except for subtle differences in TAF composition and GTF occupancy at some genes, the PIC assembly program is generally conserved at Pol II targets.

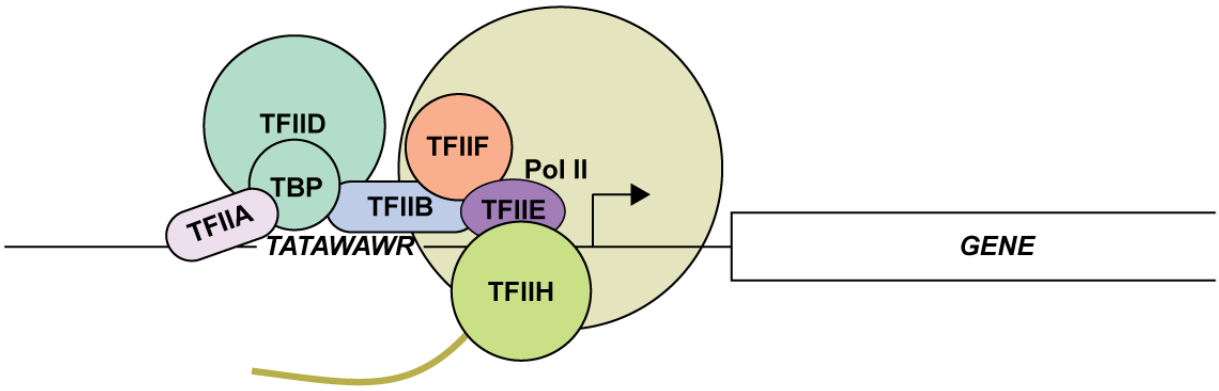

Figure 5. Transcription initiation requires the functions of many factors.

1.2.2 **Elongation**

The highly conserved C-terminal domain (CTD) of Rpb1, the largest subunit of RNA polymerase II, acts as a molecular conductor for timely assembly of the transcription elongation factors that have critical roles in elongation and RNA processing. Depending on the phosphorylation state of residues in the YSPTSPS consensus sequence, factors involved in chromatin modification, co-transcriptional splicing, transcription termination, and RNA processing are recruited at the appropriate time in the elongation process to perform their functions (BURATOWSKI 2009). The start of productive transcription is coordinated by the phosphorylation of Ser5 by a subunit of TFIIH (JERONIMO AND ROBERT 2014; WONG *et al.* 2014), which helps release the polymerase from the preinitiation complex for translocation along the open reading frame. This phosphorylation mark recruits various transcription and early RNA processing factors, including the mRNA capping enzyme (Figure 6A) (FABREGA *et al.* 2003; SUH *et al.* 2010).

Chemical modification is, in fact, a common theme in active transcription. The chromatin template is also highly modified, as discussed above, presumably as a signal to the cell of the transcriptional state of a particular locus. Phosphorylation of the CTD at Ser5, with the help of the

Paf1 complex, is responsible for recruiting Set1, the histone methyltransferase for H3K4, which places methyl marks that demarcate the 5' and 3' ends of actively transcribed genes (Figure 6B) (KROGAN *et al.* 2003a; NG *et al.* 2003b). H3K4me^3 is strongly enriched at the 5' ends of genes, while dimethylated H3K4 appears further downstream. This distinction in histone modification from gene end to end signals factors like histone deacetylases, that bind to or "read" specific histone modifications. Set3, for example, binds H3K4me^2 to restore a more closed, transcriptionally inactive chromatin state by removing acetylation marks from histones H3 and H4 (KIM AND BURATOWSKI 2009).

Upon release of Pol II from the PIC through phosphorylation of Ser5, active transcription begins. Pol II is best described as a Brownian ratchet where rates of translocation and catalysis are closely coupled (DANGKULWANICH *et al.* 2013). This mechanism depends on two main conformational states of the polymerase on DNA – pre- and post-translocated states – that physically couple nucleotide addition with forward progression (GUAJARDO AND SOUSA 1997). Upon NTP incorporation, the polymerase translocates and sequesters the new NTP into its active site where bond formation and pyrophosphate (PP_i) release occurs. PP_i release triggers reset of the polymerase to its pre-translocated state, where it can proceed through another cycle of nucleotide incorporation or enter an off-pathway paused state and backtrack. This retrograde polymerase motion often occurs in response to particular sequences or structural impediments such as DNA-bound proteins and DNA damage (NUDLER 2012). Polymerase pausing can also be regulatory, as in promoter-proximal pausing in higher eukaryotes (CHEN *et al.* 2018a), where Pol II stalls near the 5' end of genes and is poised for signals to activate productive transcription in certain cellular contexts. The most well-known example of such a regulatory pause was described for *Drosophila melanogaster* heat-shock genes, where Pol II associates even with uninduced heat-

shock gene promoters, presumably readying cells for rapid response in the event of stress (GILMOUR AND LIS 1986; CHEN *et al.* 2018a).

Entering these states depends on the evolutionarily conserved trigger loop (TL), which extends beneath the NTP in the nucleotide addition site of Pol II, and, upon incorporation of each correct nucleotide, makes extensive contacts with the NTP (WANG *et al.* 2006; HUANG *et al.* 2010; WANG *et al.* 2013). Through additional intramolecular contacts with other portions of Pol II, a single histidine is placed in the appropriate location to "trigger" phosphodiester bond formation within the Pol II active site. During pausing, altered Pol II intramolecular contacts stabilize the structure of the TL so that it can no longer perform its critical role in nucleotide addition (TOULOKHONOV *et al.* 2007; ZHANG *et al.* 2010; SABA *et al.* 2019).

Transition from early to late transcription elongation, and thus the recruitment of later-acting factors like termination factors, depends on the transition of Pol II CTD phosphorylation from Ser5P to Ser2P (HARLEN AND CHURCHMAN 2017). Some Ser5P persists throughout transcription, but many Ser5 residues in the CTD are dephosphorylated by evolutionarily conserved phosphatases Rtr1 (MOSLEY *et al.* 2009) and Ssu72 (KRISHNAMURTHY *et al.* 2004). Ser2 is phosphorylated by Cdk9, a subunit of positive transcription elongation factor P-TEFb in higher eukaryotes (interestingly, the same factor that relieves promoter-proximal pausing) (NI *et al.* 2004), and redundantly by Ctk1 (CHO *et al.* 2001) and Bur1 (LIU *et al.* 2009; QIU *et al.* 2009) in yeast. The yeast Bur1 kinase was originally described as the kinase for positive transcription factor Spt5 (a subunit of Spt4/5 or DSIF, in higher eukaryotes) (ZHOU *et al.* 2009), whose phosphorylation assists in recruiting the Paf1 complex for co-transcriptional histone modification (LIU *et al.* 2009). The elegant yet simple mechanism by which these Ser2 kinases are recruited to the elongation complex involves their binding to Ser5 phosphorylated CTD (QIU *et al.* 2009). This

requires that Ser2 phosphorylation is placed after Ser5 phosphorylation, ensuring that earlier elongation processes like mRNA capping and proper 5' end histone modification have occurred.

Ser2 phosphorylation balances productive and nonproductive transcription by recruiting factors that remove nucleosome barriers to Pol II (the histone chaperone Spt6) and reduce nucleosome accessibility (histone methyltransferase Set2), respectively. Spt6 is an H3-H4 chaperone that supports transcription elongation by restructuring, and transiently making more accessible, genic nucleosomes for the passage of Pol II (ENDOH *et al.* 2004; SIMS *et al.* 2004; VOS *et al.* 2018). This histone chaperone binds doubly phosphorylated (Ser2 and Ser5) CTD through C-terminal SH2 domains that recognize phosphoserine (YOH *et al.* 2007; SDANO *et al.* 2017), and is recruited through the help of the Paf1 complex (ADELMAN *et al.* 2006). Doubly phosphorylated CTD at Ser2 and Ser5 recruits Set2, the methyltransferase for H3K36 (Figure 6C) (XIAO *et al.* 2003; FUCHS *et al.* 2012). Much like H3K4me, a gradient of H3K36me where trimethylation increases toward the 3' end of a gene signals active transcription to the cell and brings specific transcription factors to the elongation complex (KROGAN *et al.* 2003b; FUCHS *et al.* 2012). Somewhat redundantly to H3K4me^2, H3K36me^3 recruits Rpd3S, another histone deacetylase that closes chromatin against pervasive, or cryptic, transcription (CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005).

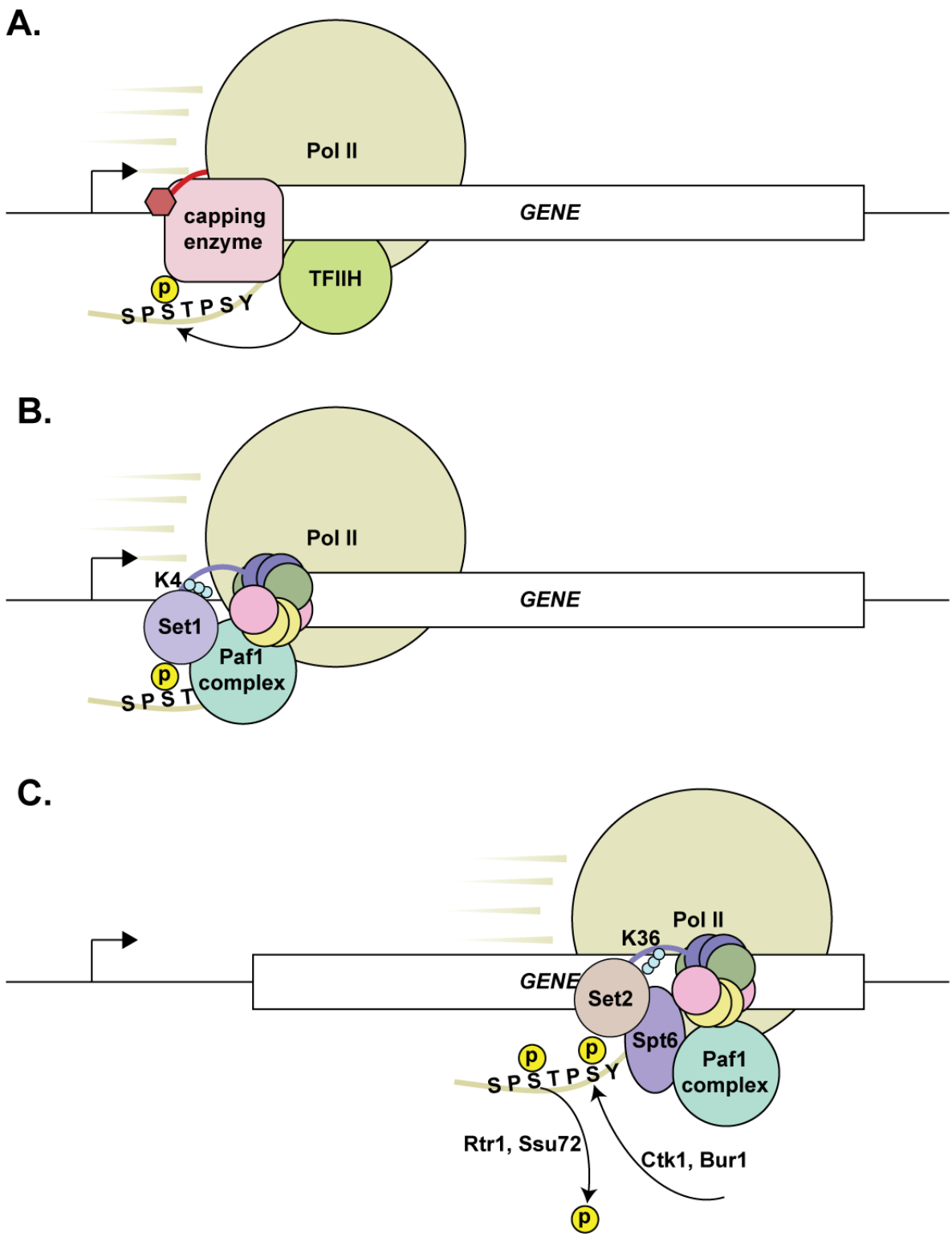
A.
Pol II
capping enzyme
GENE
TFIIH
p
S P S T P S Y
B.
Pol II
K4
Set1
Paf1 complex
GENE
p
S P S T
C.
Pol II
K36
GENE
Set2
Spt6
Paf1 complex
p
p
S P S T P S Y
Rtr1, Ssu72
Ctk1, Bur1
p

Figure 6. Phosphorylation state of the CTD of Pol II recruits factors that regulate transcription elongation.

(A) TFIIH phosphorylates the Pol II CTD at serine 5 to release the polymerase from the preinitiation complex and begin productive elongation. 5' capping machinery binds this phosphorylated CTD and caps (red hexagon) the nascent RNA (red line) co-transcriptionally as it emerges from the active site of Pol II. **(B)** The Set1 histone methyltransferase also recognizes Ser5 phosphorylated CTD and is recruited to chromatin, with the help of the Paf1 complex, to methylate H3 on lysine 4. The cartoon diagram of the nucleosome is colored as in Figure 1. Methyl marks are indicated as small blue circles. Set1 trimethylates H3K4 close to the 5' end and dimethylates H3K4 closer to the 3' end. **(C)** After approximately 500 bp, Rtr1 and/or Ssu72 dephosphorylates some, but not all, Ser5 residues in the CTD. Ctk1 and/or Bur1 places the Ser2P mark. This transition in CTD phosphorylation state recruits Spt6, which in turn helps recruit Set2 (XIAO *et al.* 2003; FUCHS *et al.* 2012). Set2 dimethylates H3K36 in the gene body, and trimethylates H3K36 closer to the 3' end (KROGAN *et al.* 2003b; FUCHS *et al.* 2012).

1.2.3 **Termination**

There are two major, essential pathways of termination by Pol II in yeast: cleavage and polyadenylation-dependent termination of protein-coding genes, and NNS-dependent termination of noncoding transcripts. The complexes required to carry out these pathways respond to different signals from the elongation complex to terminate these different kinds of transcripts, including recognition of differential Pol II CTD phosphorylation states, and binding to unique, specific signals in their target nascent RNAs.

1.2.3.1 **Polyadenylation-Dependent Termination**

Protein-coding genes are transcribed by RNA polymerase II to yield messenger RNAs (mRNAs) that are transported from the nucleus to the cytoplasm to be translated into protein. This class of transcript makes up a large proportion of Pol II-transcribed RNAs, and the pathway responsible for termination of these RNAs is quite well understood. In yeast, termination of mRNAs involves three main complexes: cleavage and polyadenylation factor (CPF), cleavage factor IA (CFIA), and cleavage factor IB (CFIB) (PORRUA AND LIBRI 2015). Together, these complexes form the larger CPF-CF complex that assembles through interactions with signals in the nascent RNA and by binding directly to the body and CTD of Pol II to cleave nascent RNA from the polymerase, recruit factors for 3' end-processing and stabilization, and disengage Pol II from the DNA template.

Transcription of the consensus AAUAAA hexamer (the polyadenylation signal, or PAS), usually found 10-35 bp upstream of the transcript's cleavage and polyadenylation site (PROUDFOOT AND BROWNLEE 1976; FITZGERALD AND SHENK 1981; HU *et al.* 2005) significantly slows and then pauses Pol II activity (OROZCO *et al.* 2002), allowing for more effective binding of termination and 3' end processing factors (GLOVER-CUTTER *et al.* 2008). Several members of the CPF-CF complex are recruited to this PAS in the nascent RNA (GILMARTIN AND NEVINS 1989; TAKAGAKI *et al.* 1989), as well as to a U- or UG-rich sequence downstream of the poly(A) site (SIMONSEN AND LEVINSON 1983; GIL AND PROUDFOOT 1984; HU *et al.* 2005). In addition to recognition of these RNA signals, factors like Pcf11 and Clp1 (subunits of the CFIA complex) bind directly to Ser2 phosphorylated Pol II CTD and to the Pol II core, respectively (PEARSON AND MOORE 2014). Directed by these redundant recruitment mechanisms, CPF subunit Ysh1, an endoribonuclease, cleaves the nascent RNA at the poly(A) site (GARAS *et al.* 2008). This cleavage releases the nascent RNA from the elongation complex with a free 3' hydroxyl, to which the poly(A) polymerase Pap1

adds approximately 70 nucleotides (~200 nt in mammals) (Figure 7A) (KUHN *et al.* 2009; EZEOKONKWO *et al.* 2012). Poly(A) binding proteins – Pab1 and/or Nab2 in yeast – bind the polyadenylate tail and serve to protect the nascent RNA from 3' degradation and assist in nuclear export (HECTOR *et al.* 2002; DUNN *et al.* 2005).

Pol II continues transcribing beyond the poly(A) site until it is disengaged from the DNA by additional factors. During transcription elongation, the active site of Pol II harbors an approximately 8 bp hybrid of single-stranded DNA bound to the RNA transcript being synthesized (KIREEVA *et al.* 2000). This RNA:DNA hybrid is thought to be a major contributor to the stability of the Pol II elongation complex throughout transcription, so its dissociation, predictably, is a critical step in transcription termination (KOMISSAROVA *et al.* 2002). One factor thought to perform this function is the 5'-3' exoribonuclease Rat1/Rai1 (XRN2/DOM3Z in higher eukaryotes), which is recruited to paused polymerase downstream of a poly(A) site through interactions with Rtt103 and 3' end processing factors that bind Ser2P CTD (KIM *et al.* 2004). In the torpedo model of transcription termination, cleavage of the nascent RNA at the poly(A) site is thought to provide a site for the Rat1/Rai1 exonuclease to load onto nascent RNA, where it degrades the RNA in the 5' to 3' direction until it encounters the elongating polymerase (KIM *et al.* 2004; WEST *et al.* 2004; LUO *et al.* 2006). Breaking the RNA:DNA hybrid in this way is thought to dislodge Pol II from the DNA, thus terminating transcription (Figure 7B).

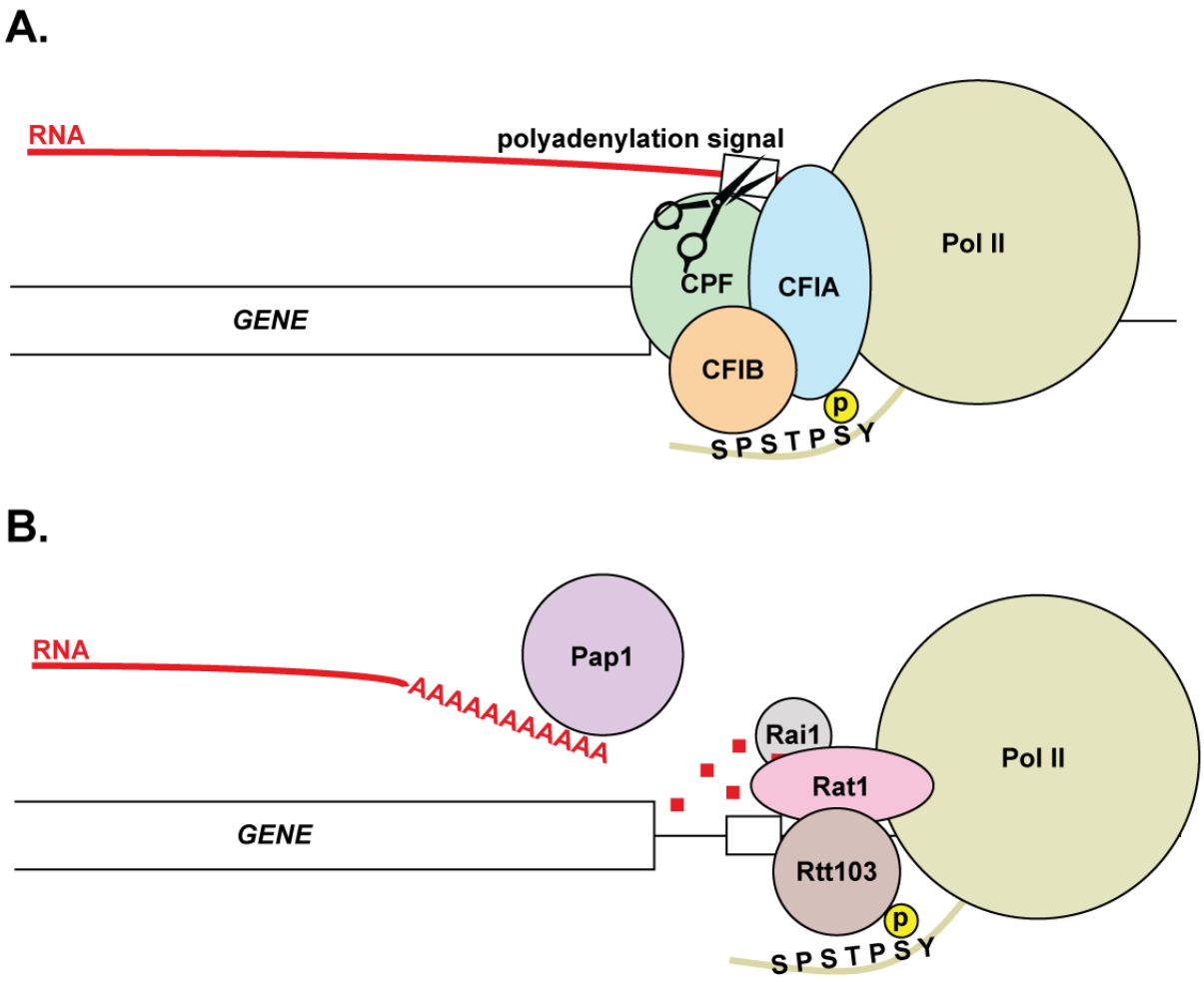

Figure 7. Cleavage and polyadenylation-dependent transcription termination requires a large, multisubunit complex.

(A) The CPF-CF complex recognizes Ser2P CTD and the polyadenylation signal. CPF subunit Ysh1 cleaves the nascent RNA at the PAS and releases the RNA for polyadenylation by Pap1, the poly(A) polymerase in yeast. Poly(A) binding proteins (not shown) bind the polyadenylate tail to protect the 3' end of the nascent RNA. **(B)** Rtt103, a Ser2P-binding protein, helps recruit the Rat1/Rai1 5'-3' exonuribonuclease to the nascent RNA still associated with Pol II. Rat1 degrades the RNA until it catches up with Pol II and dismantles the RNA:DNA hybrid within the active site, releasing Pol II.

1.2.3.2 **NNS-Dependent Termination**

In addition to the polyadenylation-dependent pathway described above, yeast have a second essential termination pathway for its Pol II-transcribed small, noncoding RNAs (discussed in more detail below), short mRNAs, and regulatory transcription attenuation of some mRNAs (STEINMETZ *et al.* 2006; SCHULZ *et al.* 2013). This pathway depends primarily on three main proteins: Nrd1 and Nab3, RNA-binding proteins that bind specific target sequences within the nascent RNA, and Sen1, an RNA-DNA helicase critical for disentangling the hybrid RNA:DNA duplex, thus terminating transcription. The decision to employ one of the two termination pathways largely relies on the length of the transcript: shorter transcripts are associated with Ser5-phosphorylated Pol II, which recruits Nrd1 through its CTD-interacting domain (CID) (Figure 8A) (VASILJEVA *et al.* 2008). Longer transcripts are associated primarily with Ser2-phosphorylated Pol II, which is repressive to the NNS pathway, but activating to polyadenylation-dependent termination (see above) (GUDIPATI *et al.* 2008).

Nrd1 (Nuclear pre-mRNA Downregulation 1) and Nab3 (Nuclear polyAdenylated-RNA Binding 3) are heterogeneous nuclear ribonucleoprotein (hnRNP)-like proteins, a class of proteins that typically has pre-mRNA processing functions in mammals (STEINMETZ AND BROW 1996; CONRAD *et al.* 2000). Nab3 was first identified in a screen for hnRNP-like proteins in yeast (WILSON *et al.* 1994), as it shares sequence homology with mammalian hnRNP-C, and contains several RNA-binding motifs thought to be critical to its role in termination (LOYA *et al.* 2012; LOYA *et al.* 2013). Nrd1 and Sen1 were first linked to termination as *trans*-factors acting on a novel termination sequence element (STEINMETZ AND BROW 1996). Nrd1, also an hnRNP-like protein, has RNA-binding motifs, and an N-terminal CTD-interacting domain (CID) like those found in mammalian SCAF4 and SCAF8 (STEINMETZ AND BROW 1996; VASILJEVA *et al.* 2008).

Through their RRMs, Nab3 recognizes a UCUUG (or as long as UUCUUGUW) consensus sequence, as determined by biochemical and genome-wide studies, and Nrd1 binds to GUA[A/G] (CARROLL *et al.* 2004; CARROLL *et al.* 2007; CREAMER *et al.* 2011; PORRUA *et al.* 2012; WEBB *et al.* 2014). Sen1 (splicing endonuclease 1), the yeast ortholog of mammalian senataxin, is a Superfamily I-like nucleic acid helicase with a wide array of likely roles in the cell, including its original assigned function, tRNA splicing, plus DNA repair, transcription, and RNA processing (WINEY AND CULBERTSON 1988; URSIC *et al.* 1997; URSIC *et al.* 2004; STEINMETZ *et al.* 2006). Together, Nrd1 and Nab3 bind the nascent RNA and heterodimerize, providing the optimal ribonucleoprotein for Sen1 recruitment. Sen1 additionally directly interacts with Pol II (CHINCHILLA *et al.* 2012).

Upon recruitment, Sen1 tracks along the elongating RNA in kinetic competition with Pol II and breaks the stable ternary Pol II-DNA-RNA complex to terminate transcription (HAZELBAKER *et al.* 2013; PORRUA AND LIBRI 2013), a method that mirrors prokaryotic Rho-dependent termination (Figure 8B). Therefore, moderation of Pol II speed during transcription elongation is important for proper termination by the NNS pathway. Because the chromatin template regulates Pol II elongation rate, it is not surprising that factors that affect chromatin structure are also involved in the NNS pathway. Paf1 complex-mediated H2Bub, H3K4me^3, and H3K36me^3 have all been shown to play important roles in regulating proper NNS-dependent termination, particularly preventing terminator readthrough (SHELDON *et al.* 2005; TERZI *et al.* 2011; TOMSON *et al.* 2011; TOMSON *et al.* 2013).

Much like polyadenylation-dependent transcription, the NNS pathway couples termination of target transcripts to their post-transcriptional processing. Upon termination, the CID of Nrd1 releases Pol II and binds the TRAMP (Trf4/Air2/Mtr4 Polyadenylation) complex in a mutually

exclusive manner, bringing it in close proximity of terminating RNAs (TUDEK *et al.* 2014). Nrd1 and the TRAMP complex recruit the nuclear exosome whose active subunit, Rrp6, trims or completely degrades the polyadenylated, nascent RNA, depending on the transcript's downstream function (LACAVA ET AL. 2005; VANACOVA ET AL. 2005; WYERS ET AL. 2005; JIA ET AL. 2011; JIA ET AL. 2012). For example, snoRNAs and other functional RNAs are matured and trimmed by the 3'-to-5' exonuclease complex, while CUTs are completely degraded (ALLMANG *et al.* 1999; VAN HOOF *et al.* 2000; VANACOVA *et al.* 2005; WYERS *et al.* 2005; TUDEK *et al.* 2014).

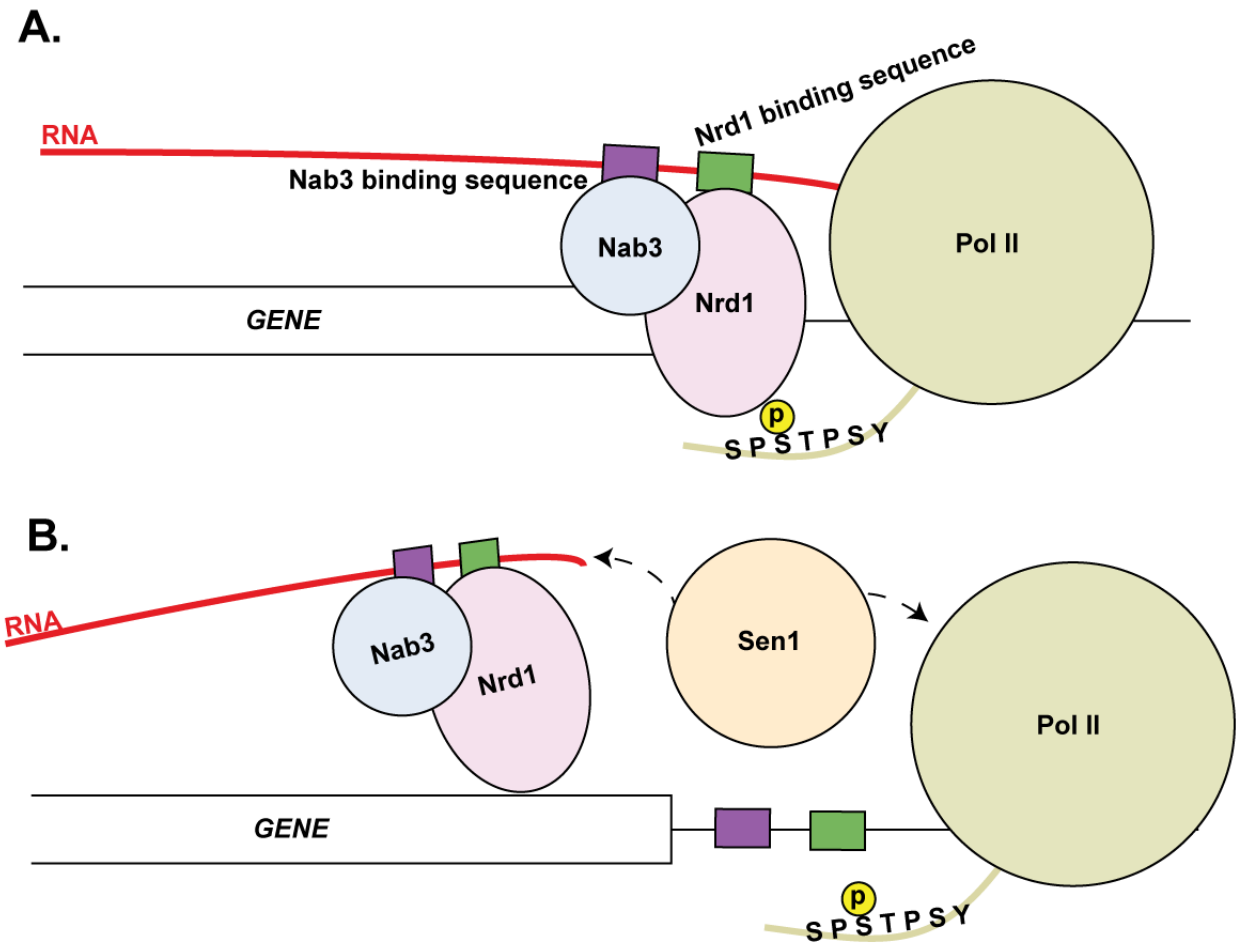

Figure 8. The Nrd1-Nab3-Sen1 pathway terminates many short, noncoding RNAs.

(A) Nrd1 is recruited to nascent RNA via Ser5P CTD and Nrd1 binding sequences in the nascent RNA. Nab3 also recognizes certain signals in the RNA. Together, these two components heterodimerize and recruit Sen1 helicase. **(B)** Once recruited to the nascent RNA, Sen1 translocates along the nascent RNA until it reaches the active site of Pol II. It unwinds the stable RNA:DNA hybrid, releasing the nascent RNA for further processing, and dislodging Pol II.

1.3 **NONCODING TRANSCRIPT TYPES IN YEAST**

In budding yeast, as with all eukaryotes, transcription is pervasive and not restricted to protein-coding genes (JENSEN *et al.* 2013). While classes of RNA types may share some common features, there are unique aspects of the noncoding RNAs (ncRNAs) that set them apart for specialized transcription and RNA processing. Within the ncRNAs, defining any single transcript as a member of one specific class is quite difficult, as many of the classes share overlapping features.

1.3.1 **snoRNAs and snRNAs**

Small nucleolar and small nuclear RNAs (snoRNAs and snRNAs, respectively) are noncoding, functional RNAs largely transcribed by RNA polymerase II in *Saccharomyces cerevisiae.* snRNAs are small, nucleoplasmic RNAs that primarily function as subunits of the spliceosome, the complex that removes intronic sequence from the pre-mRNAs of protein-coding genes (VALADKHAN 2005). snRNAs making up the yeast spliceosome consist of U1, U2, U4, U5, and U6, all of which form small nuclear ribonucleoprotein complexes (snRNPs) critical to the spliceosome's overall structure and function (YAN *et al.* 2015; YAN *et al.* 2016). The snRNA components of the spliceosome act to base pair with intron-exon junctions to guide removal of introns for splicing pre-mRNAs into

mature mRNA transcripts (NILSEN 2003). While transcription of yeast snRNAs is coupled to processing and maturation by NNS-dependent termination, followed by polyadenylation by the TRAMP complex and processing by the nuclear exosome, as discussed above, transcription of these ncRNAs in metazoans differs only slightly from transcription of protein-coding genes. snRNAs depend on the Integrator complex, rather than NNS, for coupling their transcription to 3' processing (BAILLAT *et al.* 2005). Subunits of Integrator share some similarity to components of cleavage and polyadenylation specificity factor (CPSF, CPF in yeast) and interact directly with the CTD of Pol II to carry out very similar cleavage and 3' end processing roles as the NNS complex in yeast.

snoRNAs are broken into two main subclasses: C/D box and H/ACA box snoRNAs. While eukaryotic genomes encode very few snRNAs, there are approximately 76 snoRNAs encoded by the yeast genome and more than 200 in the human genome (LESTRADE AND WEBER 2006; PIEKNA-PRZYBYLSKA *et al.* 2007). Interestingly, it is not only eukaryotic genomes that encode snoRNAs; in fact, they appear to have evolved approximately 2 billion years ago and such ancient origins allow for detection of these small, functional ncRNAs in all three extant domains of life (TERNS AND TERNS 2002). The earliest recognized role for snoRNAs was their ability to post-transcriptionally modify other RNAs, including the rRNAs, whose modifications are required for ribosomal function (DECATUR AND FOURNIER 2003). In addition to the rRNAs, snoRNAs can target snRNAs in eukaryotes (DARZACQ *et al.* 2002), tRNAs in archaea (DENNIS *et al.* 2001), and even some mRNAs in mammals (CAVAILLE *et al.* 2000). The role for snoRNAs in these interactions is as modification-guide RNAs, where the major function of the RNA component is to base pair, and bridge, the modification target and the catalytic portion of the modification RNP (OMER *et al.* 2002; BORTOLIN *et al.* 2003; BAKER *et al.* 2005; CHARPENTIER *et al.* 2005). These

more recent studies of this diverse class of RNAs shows that snoRNAs have evolved functions far beyond what their "nucleolar" name suggests.

1.3.2 **CUTs**

Cryptic unstable transcripts (CUTs) were first uncovered in cells where Rrp6, the catalytic subunit of the nuclear exosome, or the subunits of the TRAMP complex, responsible for polyadenylating ncRNAs, were deleted (WYERS *et al.* 2005; DAVIS AND ARES 2006; HOUALLA *et al.* 2006). Multiple studies have defined CUTs as approximately 200-600 bp in length, often initiating from the promoters of protein-coding genes, and heterogeneous at their 3' ends (NEIL *et al.* 2009; XU *et al.* 2009). Approximately 78% of these genome-wide studies show CUTs arising as divergent transcripts from the nucleosome-free regions (NDRs) of protein-coding genes, possibly as a result of the activity of some general transcription factors involved in PIC formation not having specific polarities. This hypothesis suggests that many promoters are by nature bidirectional and that productive and nonproductive transcripts are differentially retained post-transcriptionally.

Some CUTs, however, are not simply the result of bidirectional transcription initiation and sometimes occur in tandem with protein-coding genes. A well-studied example of this is the *SRG1-SER3* locus, where expression of the *SRG1* noncoding RNA has functional consequence. When grown in the presence of ample serine, yeast highly express *SRG1,* whose TSS is upstream of the TSS for *SER3.* This results in a transcript that overlaps the *SER3* promoter and changes its chromatin architecture, repressing transcription of *SER3* when its gene product is not needed (MARTENS *et al.* 2004; MARTENS *et al.* 2005; HAINER *et al.* 2011; PRUNESKI *et al.* 2011).

1.3.3 **SUTs, XUTs, and NUTs**

Stable unannotated transcripts (SUTs) arose from genome-wide studies on pervasive transcription (XU *et al.* 2009). These transcripts remain unannotated and their functions have not been elucidated, but they appear more stable than CUTs because they can be detected even in the absence of polyadenylation or exosome mutants. Many SUTs are antisense transcripts that may or may not serve regulatory functions. On average, these transcripts are also longer than CUTs at approximately 760 bp. The difference in length may contribute to the major difference between CUTs and SUTs – while CUTs are primarily terminated by the NNS pathway, TRAMP polyadenylated and shuttled to the nuclear exosome for degradation (though other mechanisms for their degradation have been uncovered), SUTs are more often terminated by the cleavage and polyadenylation pathway and exported to the cytoplasm for Xrn1-dependent degradation (THOMPSON AND PARKER 2007; LEE *et al.* 2008).

Xrn1-sensitive unstable transcripts (XUTs) are only detectable upon inactivation of cytoplasmic 5'-3' exoribonuclease, Xrn1 (VAN DIJK *et al.* 2011). These Pol II-transcribed ncRNAs are largely antisense to open reading frames (~66%), and more readily accumulate when yeast cells are grown in stressful media conditions. This suggests that altering the transcriptome via genome-wide antisense transcription could provide an adaptive mechanism for growth in undesirable conditions.

Nrd1-unterminated transcripts (NUTs) are RNAs detected in the cell when the essential termination factor Nrd1 is rapidly depleted from the nucleus (SCHULZ *et al.* 2013). This class of RNAs overlaps others described here, including approximately 80% of snoRNAs, 68% of CUTs, 58% of SUTs, and 37% of XUTs, except that when Nrd1 is depleted, these transcripts are on average 3.8-fold longer than previously annotated transcript isoforms. The approximately 1500

NUTs described by Schulz *et al.* includes divergent and antisense transcripts, providing further evidence for the bidirectional nature of many promoters.

1.3.4 **Cryptic Transcripts, Antisense Transcripts, and SRATs**

Genome-wide spurious transcription can result from loss of nucleosome replacement in the wake of Pol II. For instance, inactivation of Spt6 or Spt16, two transcription-coupled histone chaperones that remove and replace H3-H4 or H2A-H2B for Pol II passage, respectively, results in intragenic transcription initiation (KAPLAN *et al.* 2003). This upregulation of cryptic initiation results from uncovering cryptic promoters through some deregulation of chromatin architecture, allowing for PIC formation in noncanonical locations and transcription initiation. Prevention of cryptic initiation requires suppression of histone turnover and resulting histone acetylation along open reading frames, which cascades from proper placement of H3K36me^3, discussed above (VENKATESH *et al.* 2012).

Antisense transcription initiates from the 3' NDRs of genes, and these transcripts often span the length of ORFs, in the opposite direction, extending to their promoters. Antisense transcripts appear to arise from canonical, 5' promoter-like PICs in the 3' NDR (MURRAY *et al.* 2012). Interestingly, unlike at canonical genes, antisense transcripts originating from these 3' promoter elements are often not associated with divergent gene products, suggesting that antisense transcription is not simply divergent transcription from a downstream sense promoter. Often, sense and antisense transcription at a single locus is mutually exclusive, as in the case of regulatory antisense transcription over the *IME4* locus, which controls the decision of diploid yeast cells to propagate vegetatively or undergo meiosis (HONGAY *et al.* 2006). Larger genome-wide studies have identified more examples of genes whose transcription is tempered by antisense transcription,

even in normal growth conditions where lowly transcribed genes appear to be repressed by previously undetected antisense transcripts (NEVERS *et al.* 2018).

Set2-repressed antisense transcripts (SRATs) are a specific type of antisense transcript whose expression is regulated by Set2-mediated H3K36me (VENKATESH *et al.* 2016). In general, H3K36me is a mark that represses spurious transcription from within genes in a manner that does not affect ongoing transcription of that gene. When methylation of H3K36 is lost, antisense transcripts arise from regions canonically marked by high levels of methylation. While some of these transcripts overlap with known CUTs, SUTs, and XUTs, some are novel and require further analysis to determine potential functions beyond their transcription. Regardless of the potential functions of these recently discovered ncRNAs, the presence of SRATs in the yeast genome suggests that ncRNAs are controlled not only at the level of processing and degradation, but during transcription by chromatin modifications common to coding gene transcription.

1.4 **STUDYING TRANSCRIPTION TERMINATION IN YEAST**

The goal of my project is to identify how the nucleosome, a major regulator of DNA-templated processes, contributes to transcription termination. In light of the importance of the nucleosome in regulating the processes discussed above, whether by acting as a signaling platform via its myriad histone modifications, binding important factors for transcription, replication, DNA repair, etc., or concealing DNA from inappropriate binding by factors, the nucleosome is undoubtedly a central player in nuclear control.

Studying such a process in *Saccharomyces cerevisiae*, the budding yeast, removes complexities of higher eukaryotes such as increased histone gene copy number. The yeast genome

encodes two copies each of the four core histone proteins – H2A, H2B, H3, and H4. H3 and H4 are transcribed together as two divergently-transcribed gene pairs, *HHT1-HHF1* and *HHT2-HHF2,* and H2A and H2B are likewise organized as *HTA1-HTB1* and *HTA2-HTB2* pairs (RANDO AND WINSTON 2012). Compared to higher eukaryotes, whose genomes can encode >50 to even about 100 copies of the histone proteins (humans and *Drosophila*, respectively) (LIFTON *et al.* 1978; MARZLUFF *et al.* 2002), the simple organization of histones in yeast renders manipulation of these genes as quite simple in comparison. Throughout the work described here, strains used have had both chromosomal H3-H4 or H2A-H2B cassettes deleted. H3 and H4, for example, are covered by a *URA3*-marked plasmid expressing wild type *HHT1-HHF1* or *HHT2-HHF2*. Using the plasmid shuffle protocol, a plasmid bearing histone pairs with a target mutation is transformed into such a strain and the *URA3*-marked plasmid is selected against, leaving the mutated copy as the single source of the histone pairs (SIKORSKI AND BOEKE 1991). This method allows for rapid introduction of a wide array of histone mutations to a given genetic background, for instance in order to perform high-throughput screens (Chapter 2).

Importantly, the histone proteins are well conserved from yeast to humans (Figure 9), as are the fundamentals of chromatin modification and transcription (RANDO AND WINSTON 2012), rendering basic information learned here applicable to future studies in higher eukaryotes.

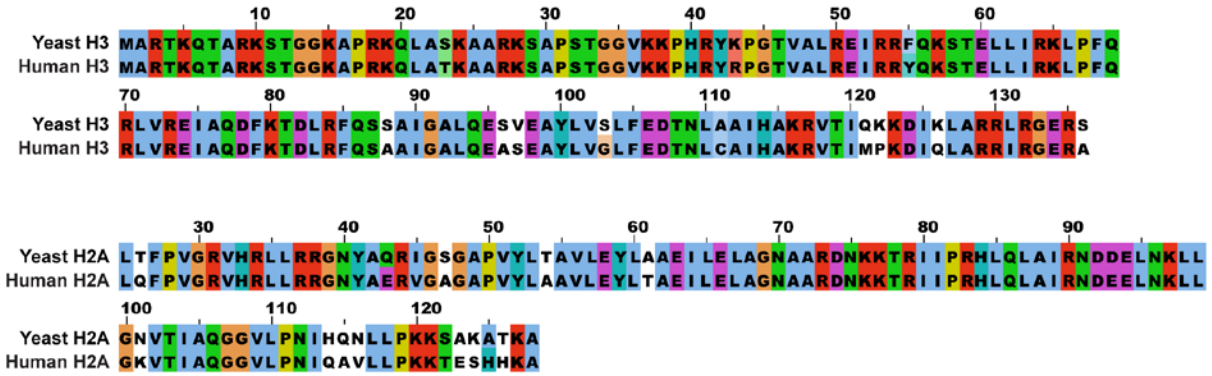

Figure 9. Yeast and human histone proteins are highly conserved.

H3 (top) and H2A (bottom) are numbered with respect to the yeast protein sequence. Yeast and human H3 sequences shown here include amino acids 1-138. Yeast H2A sequence includes amino acids 25-128 and human H2A sequence includes amino acids 24-127.

Finally, care and propagation of yeast is simple, cheap, and requires very little space and media. Large volumes of homogeneous culture can easily be grown, which allows for very rapid, reproducible data collection for high-throughput experiments such as next-generation sequencing. Combined with the conservation of the target process, transcription, and its major components, yeast is the ideal model for defining the mechanism for the nucleosome in transcription termination.

2.0 MUTATIONS TO THE DNA ENTRY-EXIT SITE OF THE NUCLEOSOME IMPAIR TRANSCRIPTION TERMINATION IN SACCHAROMYCES CEREVISIAE

Contents of this chapter will be published as part of a more comprehensive manuscript with portions of Chapter 3. Strain construction for Figures 17 and 18, and the dilution assay for Figure 17 were performed by undergraduate student Julia M. Seraly.

2.1 **INTRODUCTION**

Transcription is the first step of gene expression. This process is highly conserved across taxa, and its precise regulation is required for the development and survival of all cell types. In eukaryotes, the genome is packaged into chromatin, which acts as a regulatory barrier to the transcription machinery. Chromatin is made up of repeated nucleosomes, which consist of approximately 147 bp of DNA surrounding an octamer of core histone proteins H2A, H2B, H3, and H4 (LUGER *et al.* 1997). To faithfully express protein-coding genes and some small, noncoding RNAs (ncRNAs), RNA polymerase II (Pol II) is ushered along this complex template by a host of transcription factors (TEVES *et al.* 2014). Among these factors are histone modifiers that post-translationally decorate the histones with small chemical moieties like methyl or acetyl groups or larger protein groups like ubiquitin (LAWRENCE *et al.* 2016). These marks are individually correlated with specific chromatin-dependent processes such as active transcription, gene silencing, DNA damage repair, and cell cycle progression, and are thought to recruit downstream chromatin modifiers to the proper genomic locations and alter protein-DNA interactions to regulate nucleosome

accessibility (SWYGERT AND PETERSON 2014; LAWRENCE *et al.* 2016). Additionally, histone chaperones and chromatin remodelers can reposition, add, or remove histones or whole nucleosomes, respectively, on the chromatin template to modulate nucleosome barriers to Pol II progression (BARTHOLOMEW 2014). The mechanisms by which these factors modify chromatin for transcription initiation and elongation are well studied, but it is unclear how these factors, and the resulting chromatin structure, affect the final termination step.

Transcription termination is a critical step in gene expression. An unterminated polymerase can continue elongating into neighboring genes, a phenomenon called transcription interference, which results not only in a non-canonical transcript of the initiated gene, but also in improper temporal or spatial expression of downstream genes. Widespread aberrant transcription termination can result from conditions of cell stress (VILBORG *et al.* 2015) or viral infection (NEMEROFF *et al.* 1998; RUTKOWSKI *et al.* 2015), and is prevalent in human cancers (MAHER *et al.* 2009; KANNAN *et al.* 2011; VARLEY *et al.* 2014; GROSSO *et al.* 2015).

In the budding yeast *Saccharomyces cerevisiae*, there are two major Pol II termination pathways: 1) Cleavage and polyadenylation-dependent termination of protein-coding genes, and 2) Nrd1-Nab3-Sen1 (NNS) dependent termination of small ncRNAs and aborted transcripts from protein-coding genes (PORRUA AND LIBRI 2015). Studies in yeast have indicated a role for chromatin as critical players in the transcription termination of a few candidate genes. The Paf1 complex and the histone modifications it mediates have been shown to play important roles in preventing terminator readthrough of some NNS-dependent snoRNAs (SHELDON *et al.* 2005; TERZI *et al.* 2011; TOMSON *et al.* 2011; TOMSON *et al.* 2013). ATP-dependent chromatin remodelers Isw1 and Chd1, which remove and replace nucleosomes to regulate the accessibility of specific DNA sequences, have been linked to transcription termination of some mRNAs (ALEN

et al. 2002; MORILLON *et al.* 2003). More recent work has described nucleosomes and general chromatin regulatory factors like Reb1, Rap1, and Abf1 as widespread physical roadblocks to the progression of Pol II at NNS-dependent ncRNAs (ROY *et al.* 2016).

To further address how the chromatin template contributes to the regulation of this complex process, I screened alanine substitutions of each residue of the core histones H3 and H4, and selected residues in H2A (NAKANISHI *et al.* 2008) for transcription termination defects using a well-characterized termination reporter (CARROLL *et al.* 2004). This screen showed that specific amino acids in H2A, H3, and H4 are required for transcription termination. Most of these residues localize to the DNA entry-exit site of the nucleosome, which is important for transcription-dependent nucleosome occupancy (HAINER AND MARTENS 2011a; HYLAND *et al.* 2011) and trimethylation of H3 at lysine 36 ($H3K36me^3$) (DU AND BRIGGS 2010; ENDO *et al.* 2012). I hypothesized, given the importance of this protein surface in general chromatin maintenance and transcription, that these histone mutants might exhibit a myriad of chromatin- and transcription-related phenotypes. Thus, I assessed these mutant strains for several transcription-related phenotypes and discovered that they do indeed exhibit additional transcriptional defects in addition to their observed termination defect.

2.2 **MATERIALS AND METHODS**

2.2.1 **Yeast Strains and Media**

All *Saccharomyces cerevisiae* strains used in this study are isogenic to FY2, a *GAL2*$^{+}$ derivative of S288C (WINSTON *et al.* 1995). Unless otherwise noted, all experiments were performed with yeast strains lacking endogenous copies of the genes expressing histones H3 and H4 (Table 2) transformed with CEN/ARS plasmids expressing *TRP1*-marked, alanine-substituted histones H3 and H4 (*hht2-HHF2* or *HHT2-hhf2,* respectively) (NAKANISHI *et al.* 2008) or a *TRP1*-marked plasmid expressing the wild type *HHT2-HHF2* cassette. Lithium acetate yeast transformations were performed as previously described (BECKER AND LUNDBLAD 2001). For the genetic screen and following experiments, *URA3*-marked, wild type histone plasmids were shuffled with *TRP1*-marked mutant histone plasmids by sequentially passaging transformants three times on SC-Trp medium containing 2% dextrose and 0.1% 5-FOA.

KY3219 and KY3220 were made by crossing KY2477 (*leu2*Δ*::ADH1-HIS3-CYC1::LEU2*) or KY2474 (*leu2*Δ*::ADH1-SNR47*$_{term}$*-HIS3-CYC1::LEU2*) with KY812 (*hht1-hhf1*Δ*::LEU2, hht2-hhf2*Δ*:KanMX*), respectively. Generation of Spt^{-}, Bur^{-}, cryptic initation, and Sin^{-} (KY3354) reporter strains was by crossing KY811 (*MAT* α) or KY812 (*MAT* **a**) histone shuffle strains to yeast strains harboring the appropriate reporters. The Spt^{-} phenotype was assessed via growth on nutrient-depleted media (SD-His for the *his-912∂* allele and SD-Lys for the *lys2-128∂* allele) (SIMCHEN *et al.* 1984; WINSTON *et al.* 1987; FASSLER AND WINSTON 1988). The Bur^{-} phenotype was detected as growth on media containing sucrose, despite a deletion of the *SUC2* upstream activating sequence (*suc2*Δ*uas*) (PRELICH AND WINSTON 1993). Cryptic initiation was measured via the reporter *GAL1pr-FLO8-HIS3* as growth on media lacking histidine and

containing galactose (CHEUNG *et al.* 2008). Finally, the Sin$^-$ phenotype was detected as growth on galactose-containing media, despite deletion of the chromatin remodeler (*snf2*Δ) critical for modifying chromatin architecture for galactose-induced gene expression (KRUGER *et al.* 1995). Antimycin A, where indicated, prevents yeast from relying on respiration to bypass fermentation of non-dextrose sugars in YP+Galactose or YP+Sucrose.

Yeast strains were otherwise grown in SC-Trp medium containing 2% dextrose or YPD.

2.2.2 **Serial Dilution Assays**

Yeast cultures were grown to saturation overnight at 30°C and washed with sterile water, then resuspended to a concentration of 1 x 10^8 cells/mL. These were serially diluted four more times in water by a factor of five or ten, as indicated in each figure's legend. Three microliters of each dilution were spotted on growth control media (lacking selection) and selective media to assess specific phenotypes. Plates were incubated at 30°C for at least three days and imaged daily.

2.2.3 **Northern Blot Analysis**

Total RNA was isolated from mid-log phase yeast cultures (OD_{600} = 0.8 – 1.0) by hot acid phenol extraction as described previously (COLLART AND OLIVIERO 2001). 20 µg of each sample was analyzed by Northern blot analysis as described previously (SWANSON *et al.* 1991). Double-stranded DNA probes were synthesized by random-prime labeling of PCR fragments with [α-^{32}P]dATP and [α-^{32}P]dTTP (PerkinElmer, Waltham, MA). Probe locations are as follows: *SNR47-YDR042C* (-325 to -33 of *YDR042C*), *SNR48-ERG25* (-746 to -191 of *ERG25*), *SNR13-TRS31* (-

231 to +449 of *TRS31*), *STE11* (+1868 to +2110), and *SCR1* (-242 to +283 of *SCR1*) (primer sequences listed in Table 3).

2.2.4 **Western Blot Analysis**

Total protein was isolated from mid-log phase yeast cultures (OD_{600} = 0.8 – 1.0) by bead beating in 20% trichloroacetic acid. Protein samples were resolved on SDS-PAGE gels and transferred to nitrocellulose or PVDF membrane. The membranes were incubated with primary antibody, including: total H2B (39237 1:5000, Active Motif, Carlsbad, CA), total H3 (ab1791 1:15,000, Abcam, Cambridge, UK), H3K4me^2 (07–030 1:2000, Millipore, Burlington, MA), H3K4me^3 (39159 1:2000, Active Motif, Carlsbad, CA), H3K36me^3 (ab9050 1:1000, Abcam, Cambridge, UK), or G6PDH (A9521 1:30,000, Sigma-Aldrich, St. Louis, MO) as the loading control. Membranes were then incubated with 1:5000 α-rabbit or α-mouse secondary antibody (GE Healthcare, Little Chalfont, UK). Proteins were visualized using chemiluminescence substrate (Pico, PerkinElmer, Waltham, MA) and the ChemiDoc XRS imaging platform (BioRad, Hercules, CA).

For H3K56ac blots (1:5000, P. Kaufman), alkaline lysis protein extraction was performed as previously described (KUSHNIROV 2000). 15% SDS-PAGE gels were run and transferred to nitrocellulose membrane. After incubation with primary antibody, membranes were incubated with α-rabbit secondary antibody and developed as described above.

Signal density for histone post-translational modifications was quantified relative to total H3 signal using ImageJ software, with wild type signal set to one. Quantification of each mark was averaged across three separate experiments, with error bars representing the standard error of the mean (SEM) of the three experiments.

2.2.5 **Chromatin Immunoprecipitation and Quantitative PCR (ChIP-qPCR)**

Chromatin was isolated from 250 mL of yeast cells grown to early log phase (OD_{600} = 0.5 – 0.8) as described previously (SHIRRA *et al.* 2005). ChIPs were performed by incubating sonicated chromatin overnight at 4°C with one of the following primary antibodies: H2A (2 µL 39235, ActiveMotif, Carlsbad, CA) or H3 (5 µL ab1791, Abcam, Cambridge, UK). Antibody-chromatin complexes were purified with Protein A or Protein G conjugated to sepharose beads (30 µL, GE Healthcare, Little Chalfont, UK). DNA was column purified (Qiagen Quick, Qiagen, Hilden, Germany) and analyzed by qPCR with Maxima 2X SYBR Master Mix (ThermoFisher, Waltham, MA). Primer sets analyzed for H3 occupancy include: the coding region of *SNR47* (-35 to +94 of SNR47), downstream of the *SNR47* gene (+225 to +323 of *SNR47),* the 5' coding region of *PMA1* (+214 to +319 relative to the start codon), the 3' coding region of *PMA1* (+2107 to +2194 of *PMA1*), and a transcriptionally inactive region near the telomere of chromosome VI (chromosomal coordinates 269495 to 269598). Levels of protein occupancy at each primer set were calculated using the appropriate primer efficiency raised to the difference between input C_t and immunoprecipitated C_t values. Error bars indicate the standard error of the mean (SEM) of two technical replicates each of three biological replicates.

Table 2. Strains used in Chapter 2.

Strain	MAT	Genotype
KY319	**a**	*spt6-140 his4-912∂ lys2-128∂ ura3-52 trp1Δ63*
KY325	**a**	*spt10Δ::TRP1 his4-912∂ lys2-128∂ leu2Δ1 ura3-52 trp1Δ63 suc2Δuas(-1900/-390)*
KY811	α	*(hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM1 = *URA3/CEN/ARS/HHT2-HHF2*]
KY812	**a**	*(hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY912	**a**	*set2Δ::HIS3 his3Δ200 lys2-128∂ leu2Δ1 ura3-52*
KY934	α	*dot1Δ::HIS3 his3Δ200 leu2Δ1 trp1Δ63*
KY937	α	*set1Δ::HIS3 his3Δ200 lys2-128∂ ura3-52 trp1Δ63*
KY943	**a**	(*hta1-htb1*)Δ::*LEU2* (*hta2-htb2)Δ::TRP1 his3Δ200 lys2-128δ leu2Δ1 ura3-52* [pSAB6 = *URA3/CEN/ARS/HTA1-HTB1*]
KY1437	**a**	*rtt109Δ::KanMX his3Δ200 leu2Δ0 ura3-52*
KY2474	α	*leu2Δ1::ADH1p-SNR47(70)-HIS3-CYC1::LEU2 his3Δ200 lys2-128∂ ura3-52 trp1Δ63*
KY2477	α	*leu2Δ1::ADH1p-HIS3-CYC1::LEU2 his3Δ200 ura3-52*
KY3219	α	*leu2Δ1::ADH1p-HIS3-CYC1::LEU2 (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3220	α	*leu2Δ1::ADH1p-SNR47(70)-HIS3-CYC1::LEU2 (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3354	α	*his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63 (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX snf2Δ::HIS3* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3502	**a**	*his4-912∂ lys2-128∂ leu2Δ1 ura3-52 trp1Δ63 (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3503	α	*his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63 suc2Δuas(-1900/-390) (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3506	α	*his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63 KanMX::GAL1p-FLO8-HIS3 (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]

Table 3. Oligonucleotides used in Chapter 2.

Target		Sequence
SNR47-YDR042C	F	5’-GGCGGTAACGTAAATCAGAGTAGC-3’
Northern probe	R	5’-GAGACCTAGTCGTTTGTTAGCTG-3’
SNR48-ERG25	F	5’-CCTTGGCGCAGAAGACTTTCTCTTC-3’
Northern probe	R	5’-GCATACACAGGCGTACGCATACAAG-3’
SNR13-TRS31	F	5’-CGTAGCGCTGCATATATAATGCG-3’
Northern probe	R	5’-GATGCAGAAGTCGCTGTGCTGGAG-3’
STE11	F	5’-GGATGTCACCAGAGGTGGTCAAAC-3’
Northern probe	R	5’-GATTATCAATACAGGGCCTAGTGCCC-3’
SCR1	F	5'-CAACTTAGCCAGGACATCCA-3'
Northern probe	R	5'-AGAGAGACGGATTCCTCACG-3'
SNR47 Gene	F	5’-CCTTATTATACATTCTCTTGGCGAGTGATC-3
ChIP	R	5’-GTGTTAAAAAGCTATTGTCAAAGTTTGTTTCC-3’
SNR47 Intergenic	F	5’-CGCGTCGGGATAACAAAGCGTAC-3’
ChIP	R	5’-GTACTGCAGATACACAGATTCAAATTTTTCCC-3’
PMA1 5’	F	5’-GCTAGACCAGTTCCAGAAGAATATTTACA-3’
ChIP	R	5’-CAGCCATTTGATTCAAACCGTA-3’
PMA1 3’	F	5’-GAAATCTTCTTGGGTCTATGGATTG-3’
ChIP	R	5’-CAACATCAGCGAAAATAGCGAT-3’
TELVI	F	5’-TGCAAGCGTAACAAAGCCATA-3’
ChIP	R	5’-TCCGAACGCTATTCCAGAAAG-3’

2.3 **RESULTS**

2.3.1 **Identification of Histone Residues Required for Transcription Termination in *Saccharomyces cerevisiae***

The Arndt lab has shown previously that the Paf1 complex, a factor whose cellular roles span a variety of transcription-coupled chromatin transactions, is required for proper transcription termination by the NNS pathway (SHELDON *et al.* 2005; TOMSON *et al.* 2011; TOMSON *et al.* 2013). Additionally, several transcription-dependent histone modifications are involved in regulating this termination pathway (TERZI *et al.* 2011; TOMSON *et al.* 2011; TOMSON *et al.* 2013). Together, these findings suggest that the maintenance of a chromatin state amenable to transcription termination is a highly regulated process that requires the interplay of several factors. Therefore, I asked how chromatin itself regulates the termination of transcription. To begin to address this, I made use of a comprehensive alanine scanning plasmid library (NAKANISHI *et al.* 2008) and a reporter containing the 70 bp termination element of *SNR47* (CARROLL *et al.* 2004) to screen for mutations in histones H3 and H4 that cause defective transcription termination. In short, I generated yeast strains expressing either the termination reporter or a negative control from the *LEU2* chromosomal locus (Figure 10A). Using the plasmid shuffle technique (SIKORSKI AND BOEKE 1991), I transformed these strains with the *TRP1*-marked H3- and H4-expressing library plasmids, and counter-selected on 5-fluoroorotic acid (5-FOA). Passaged transformants were plated on media lacking histidine to assess termination defects. To reduce background readthrough signal, I added 0.5 mM 3-amino-1,2,4-triazole (3-AT), a competitive inhibitor of the *HIS3* gene product, to the media.

Using this method, I identified ten histone mutants, nine in H3 and one in H4, that strongly express *HIS3* from the reporter (Figure 10B). Interestingly, many of the residues altered in these mutants map to the DNA entry-exit site of the nucleosome (Figure 10C). This surface of the nucleosome makes several important DNA contacts (LUGER *et al.* 1997), and has previously been implicated in controlling transcription-coupled nucleosome occupancy (HAINER AND MARTENS 2011a) as well as methylation of H3 lysine 36 (H3K36me) (DU AND BRIGGS 2010; ENDO *et al.* 2012), a mark important to transcription elongation and prevention of spurious cryptic transcription (CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005; VENKATESH AND WORKMAN 2013). These data implicate co-transcriptional histone modification and nucleosome occupancy as major contributors to regulation of transcription termination.

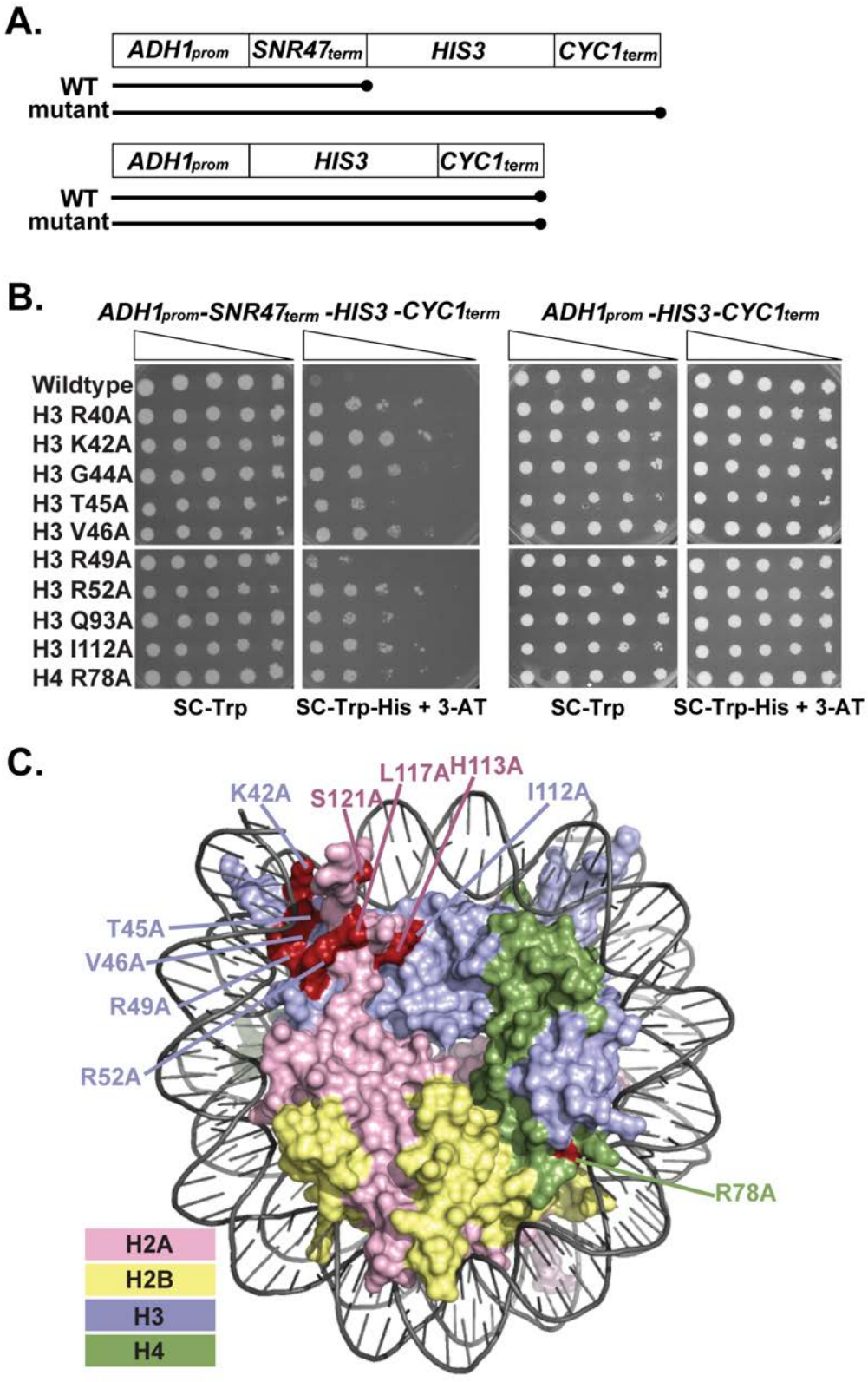

Figure 10. A genetic screen for histone residues required for transcription termination.

(A) Reporter strains contain either a transcription termination reporter (top) or a control transcription cassette lacking a terminator (bottom), integrated at the *LEU2* locus. **(B)** Termination defects in histone mutant strains were assessed by ten-fold serial dilution and plating on selective media. Library plasmids were transformed into both the strain expressing the termination reporter (KY3220) and the strain expressing the reporter control (KY3219). Cells were grown on SC-Trp as a growth control, and on SC-Trp-His + 0.5 mM 3-AT, a competitive inhibitor to the *HIS3* gene product, to select for cells strongly expressing the *HIS3* gene. Plates were incubated at 30°C for 5 days. **(C)** X-ray crystal structure of the nucleosome denoting histone residues required for transcription termination in red. Histones H2A, H2B, H3, and H4 are colored in pink, yellow, lilac, and green, respectively. Structure from PDB ID 1ID3 (LUGER *et al.* 1997).

2.3.2 **Confirmation of Terminator Readthrough at Endogenous Genes**

To determine whether the histone mutants identified in our screen cause transcription readthrough defects at endogenous genes, I performed Northern blot analysis of several snoRNA loci (Figure 11). I radiolabeled probes specific to intergenic regions downstream of *SNR47, SNR48,* and *SNR13,* to detect readthrough transcripts from these loci. The mutants identified in this screen had varying degrees of terminator readthrough as observed by Northern blotting. Interestingly, the mutants tested show varying levels of readthrough at different snoRNA loci, suggesting some degree of locus-specific dependence on each of the DNA entry-exit site residues. It is important to note that, at endogenous genes, H3 R40A has wild type termination at all three loci tested and therefore appears to be a false positive from the original screen.

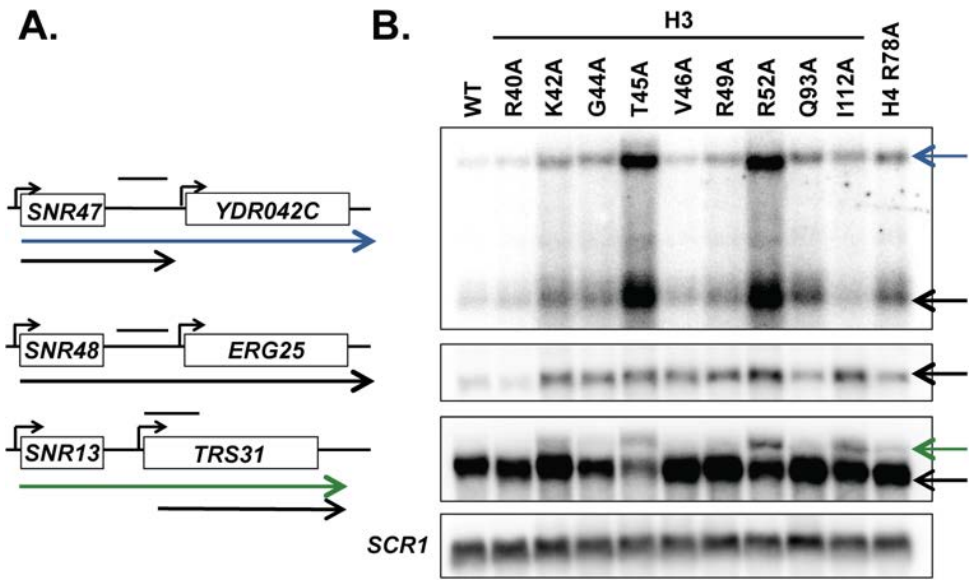

Figure 11. Residues at the DNA entry-exit site in histone H3 are required for transcription termination.

(A) Diagrams of three snoRNA loci analyzed for termination readthrough. The intergenic *SNR47* probe detects two readthrough transcripts, as indicated by the blue and black arrows. The intergenic *SNR48* probe detects a single readthrough transcript. For *SNR13*, the probe anneals to the downstream gene, *TRS31*, and detects a readthrough transcript of *SNR13* (green arrow), as well as the full-length gene product of *TRS31* (black arrow). The black bar over each locus diagram denotes each probe position. **(B)** Northern blot analysis of transcription termination defects in histone H3 and H4 mutants from Figure 10 at endogenous loci. Library plasmids were transformed into a strain lacking the termination reporter, KY812. Blots align with locus diagrams in A. Arrows correspond to arrows below locus diagrams in A, with longer arrows corresponding to slower mobility bands. *SCR1* serves as the loading control.

I was unable to use the *SNR47*$_{term}$*-HIS3* termination reporter to assess the requirement for residues in histones H2A and H2B for proper termination due to genetic marker conflicts. These plasmid libraries are marked by *HIS3*. However, the DNA entry-exit site is composed of portions of histones H3 and H2A. Therefore, I specifically tested H2A residues in and near the DNA entry-

exit site for proper termination by Northern blot analysis (Figure 12). Again, I observed varying termination phenotypes in these H2A mutants, but consistently saw the strongest termination defects in H2A mutants that have substitutions in residues near the sites of the strongest H3 mutants (Figure 10C). Together, these data suggest that the DNA entry-exit site is an important surface for regulating transcription termination.

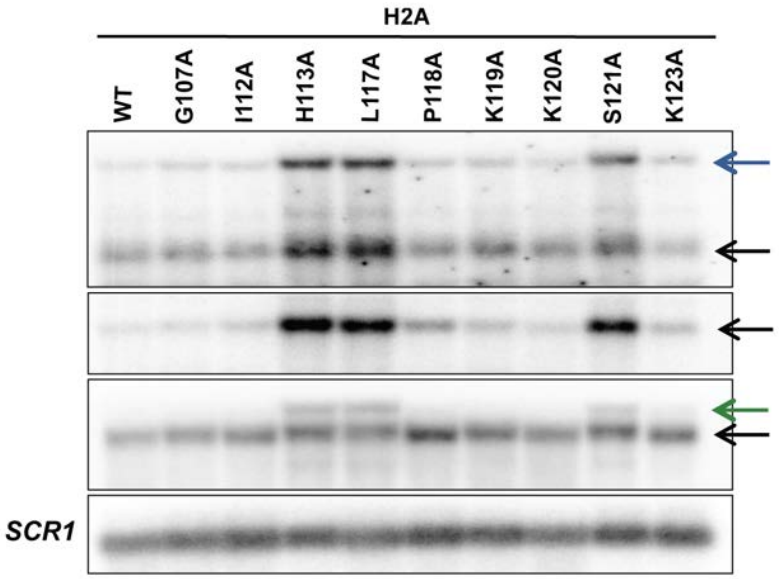

Figure 12. H2A residues in the DNA entry-exit site are critical for transcription termination.

Northern blot analysis of transcription termination defects in histone H2A mutants. Library plasmids were transformed into KY943 and mutant strains were made by plasmid shuffling. Blots align with locus diagrams in Figure 11A. Arrows in this figure correspond to arrows below the locus diagrams in Figure 11A. *SCR1* serves as the loading control.

2.3.3 **Mutations to the DNA Entry-Exit Site Impair Placement of Transcription-Coupled Trimethylation of H3 at K36**

Proper termination by the NNS pathway has been linked to several well-studied, transcription-coupled histone modifications by our lab and others (TERZI *et al.* 2011; TOMSON *et al.* 2011; TOMSON *et al.* 2013), including trimethylation of histone H3 at lysine 4 ($H3K4me^3$) and $H3K36me^3$. These modifications demarcate the 5' and 3' ends of genes, respectively, and are widely recognized as signals of active transcription that help maintain transcription fidelity through recruitment of important factors (RANDO AND WINSTON 2012). Considering the known roles for these methylation marks in transcription, and the proximity of the DNA entry-exit site to the flexible H3 N-terminal tail, I asked whether the requirement for an intact DNA entry-exit site in transcription termination is due to a role in proper placement of these marks.

To further investigate this, I performed Western blot analysis for $H3K4me^2$, $H3K4me^3$, and $H3K36me^3$ in all termination mutants identified in the original genetic screen. Placement of $H3K4me^2$ and $H3K4me^3$ are directly dependent on the Set1 histone methyltransferase, and require monoubiquitylation of H2B by Rad6/Bre1 (SANTOS-ROSA *et al.* 2002; SUN AND ALLIS 2002). Importantly, total H3 levels in termination mutants are not changed. H3K4me levels are, expectedly, completely lost in *bre1*Δ and *set1*Δ control strains but remain unaffected in DNA entry-exit site mutants with the exception of modest but statistically significant decreases of $H3K4me^2$ in H3 V46A and I112A and $H3K4me^3$ in H3 I112A (Figure 13). Termination defects due to alanine substitution of these residues may be partially related to their reduced ability to regulate proper co-transcriptional histone modification. However, for the majority of the H3 and H4 mutants, including those most strongly defective in termination (H3 T45A and H3 R52A), H3K4me is unaffected.

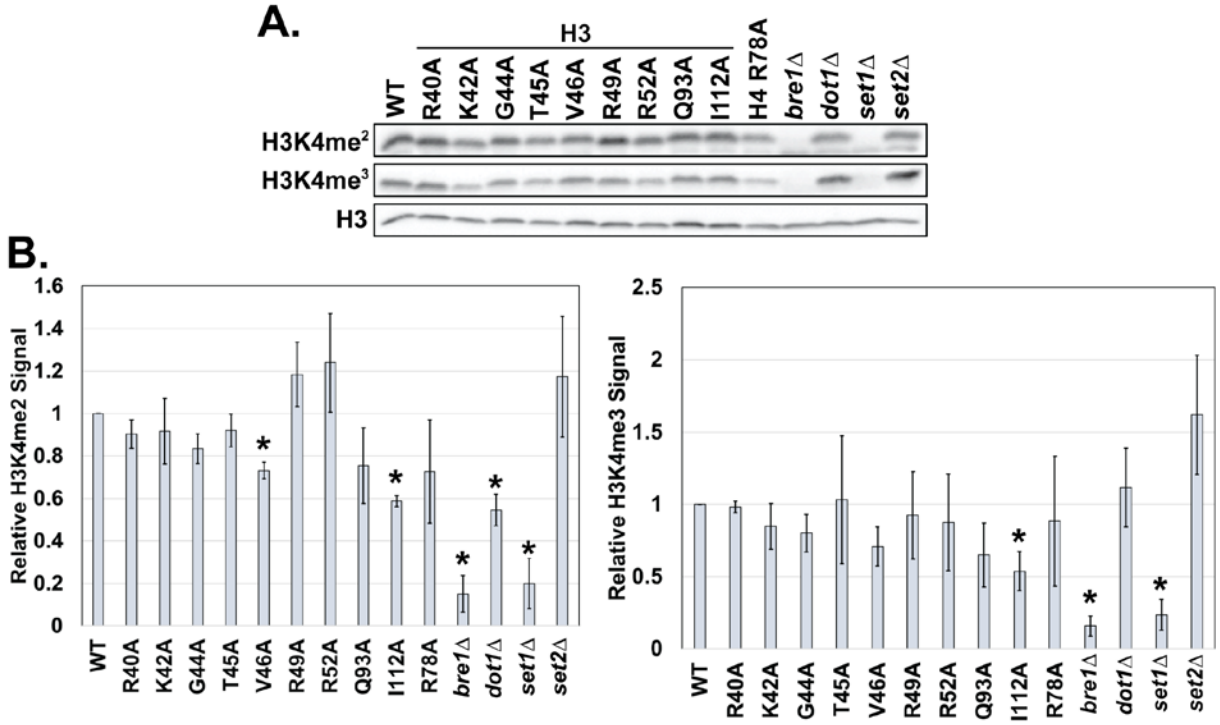

Figure 13. DNA entry-exit site mutants have approximately wild type H3K4me.

(A) Library plasmids were transformed into KY812. *bre1*Δ (KY970)*, dot1*Δ (KY934)*, set1*Δ (KY937), and *set2*Δ (KY912) are histone modifier controls, whose wild type protein products place H2Bub, H3K79me, H3K4me, and H3K36me marks, respectively. Western blotting was performed with antibodies against H3K4me^2, H3K4me^3, total H3. **(B)** Quantification of three biological replicates indicates that levels of H3K4me^2 and H3K4me^3 are relatively unaffected in most DNA entry-exit site mutants. Error bars represent SEM of histone modification signal normalized to total H3. Asterisks represent $p < 0.05$ by a Student's t-test.

Residues in the DNA entry-exit site have been implicated in Set2-dependent H3K36me^3 placement (DU AND BRIGGS 2010; ENDO *et al.* 2012). Further, placement of H3K36me^3 was shown by our lab to be required for proper NNS-dependent termination (TOMSON *et al.* 2013). I tested

whether termination mutants discussed here have defects in placement of this important co-transcriptional modification. Consistent with published results (DU AND BRIGGS 2010; ENDO *et al.* 2012), H3K36me^3 is completely lost in H3 R49A and R52A mutants (Figure 14). Other DNA entry-exit site mutants exhibit reproducible decreases in H3K36me^3 signal. In H2A mutants, levels of methylation at H3K36 are unchanged (data not shown).

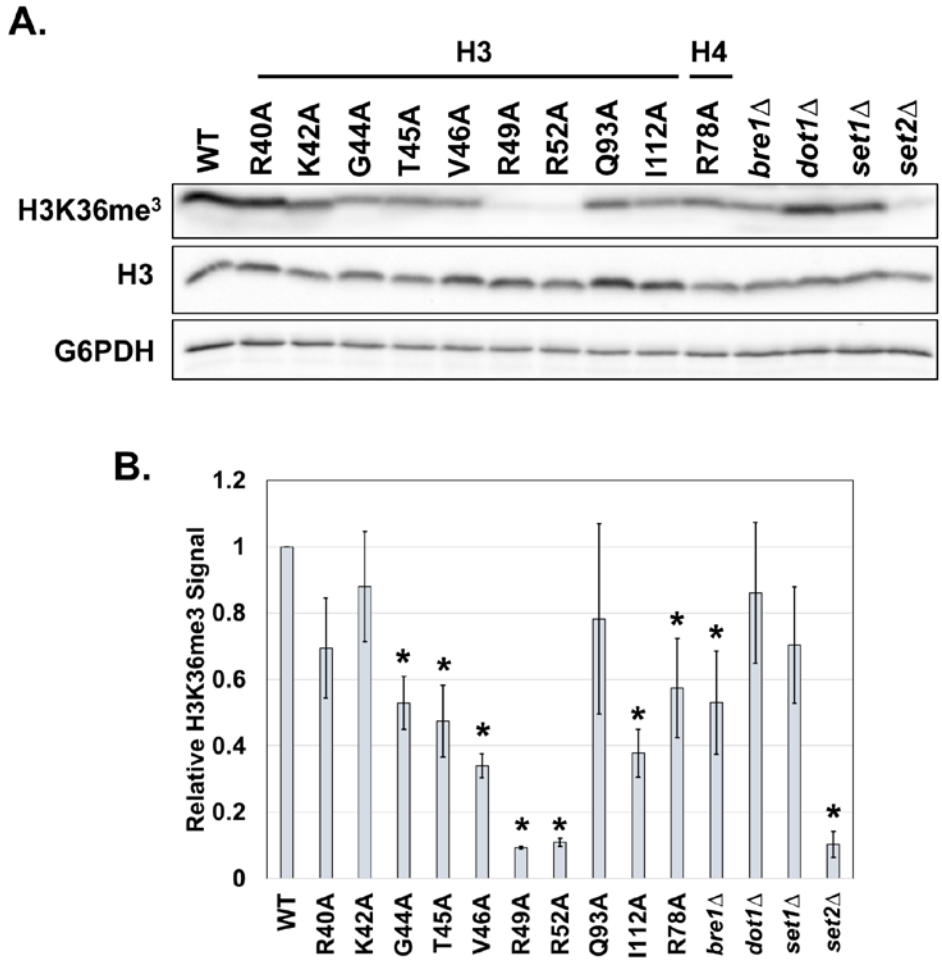

Figure 14. Many termination mutants have defects in H3K36me^3.

(A) Library plasmids were transformed into KY812 for plasmid shuffling. *bre1*Δ (KY970), *dot1*Δ (KY934), *set1*Δ (KY937), and *set2*Δ (KY912) are histone modifier controls, whose wild type protein products place H2Bub, H3K79me, H3K4me, and H3K36me marks, respectively. Western blotting was performed with

antibodies against H3K36me^3, G6PDH, and total H3. Importantly, H3 levels are not greatly increased or decreased in DNA entry-exit site mutants. **(B)** Quantification of three biological replicates indicates that levels of H3K36me^3 levels are decreased in many DNA entry-exit site mutants. Error bars represent SEM of three biological replicates. Histone modification signal is normalized to total H3. Asterisks represent $p < 0.05$ by a Student's t-test.

In addition to the requirement for methylation of histone H3, previous work in our lab has described a requirement for the monoubiquitylation of H2B at lysine 123 (H2BK123ub), another post-translational histone modification that is closely coupled to active transcription, in NNS-dependent transcription termination (TOMSON *et al.* 2011; TOMSON *et al.* 2013). Losses in placement of H2Bub in DNA entry-exit site mutants are consistent with the role for ubiquitylation in termination though, again, strength of the termination defect and the H2Bub defect for a given mutant does not agree (Figure 15).

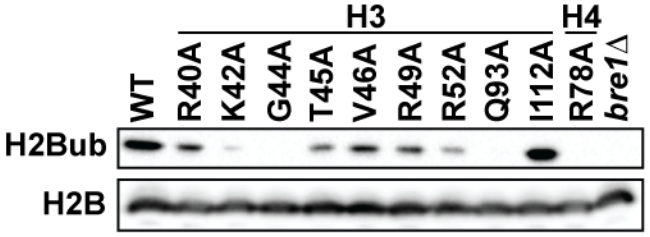

Figure 15. H2Bub appears to be misregulated in DNA entry-exit site mutants.

Western blotting was performed with antibodies against H2BK120ub (equivalent mark in humans to H2BK123ub) and total histone H2B. While similar relative trends in H2Bub have been observed in biological duplicate among the mutants, the wild type control was inconsistent among the two replicates. Further testing will be required to quantitatively assess any H2Bub defects.

Finally, I tested whether termination mutants exhibit changes in acetylation of H3 at lysine 56 (H3K56ac), a mark that is not coupled to transcription, but to deposition of newly synthesized nucleosomes after DNA repair or replication. Recent work implicates the DNA entry-exit site in interacting with Rtt109, the histone acetyltransferase that places this mark (ZHANG *et al.* 2018). However, single-residue alanine substitutions do not alter levels of H3K56ac, suggesting that the termination phenotype caused by these mutants is not due to improper recovery of nucleosomes on chromatin after synthesis or repair (Figure 16). Together, these data implicate the DNA entry-exit site of the nucleosome in several transcription-coupled processes, including placement of various histone modifications.

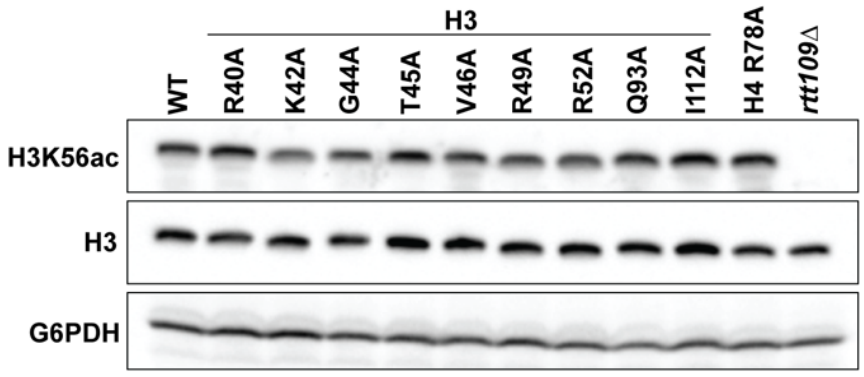

Figure 16. DNA entry-exit site termination mutants do not exhibit changes in H3K56ac.

Library plasmids were transformed into KY812 for plasmid shuffling. *rtt109*Δ (KY1437) is a strain where the histone acetyltransferase for H3K56 is deleted. Western blotting was performed with antibodies against H3K56ac, G6PDH, and total H3. The H3K56ac antibody was a kind gift from Paul Kaufman (UMass).

2.3.4 **Mutations to the DNA Entry-Exit site of the Nucleosome Cause a Variety of Transcription-Related Phenotypes**

Growth-based phenotype testing for defects in transcription using reporter alleles has been thoroughly established by our lab and others. Therefore, with the help of undergraduate student Julia Seraly, I sought to determine whether mutations to the DNA entry-exit site, which is known to contribute to nucleosome stability, cause any other well-studied transcription phenotypes. First, Julia assessed DNA entry-exit site mutants for the Suppressor of Ty (Spt^-) phenotype (SIMCHEN *et al.* 1984; WINSTON *et al.* 1987; FASSLER AND WINSTON 1988) using two commonly published alleles, *lys2-128∂* and *his4-912∂*. This phenotype results from mutations in genes encoding *trans*-factors required for proper transcription, allowing transcription machinery to bypass retrotransposon (Ty or ∂ element) insertion into a gene's upstream regulatory sequence, and thus producing gene products capable of supporting cellular life. Other studies have reported differences in suppression of these two alleles, and indeed we observe an overall more pronounced suppression of the Ty ∂ element upstream of the *HIS4* gene (Figure 17). Suppression of *lys2-128∂* is more varied in the mutants. The mutants primarily exhibit moderate phenotypes with the exception of H3 Q93A, whose severity is comparable to the *spt6-140* control (mutation to a transcription elongation factor and histone chaperone known to show the Spt- phenotype) (WINSTON *et al.* 1984).

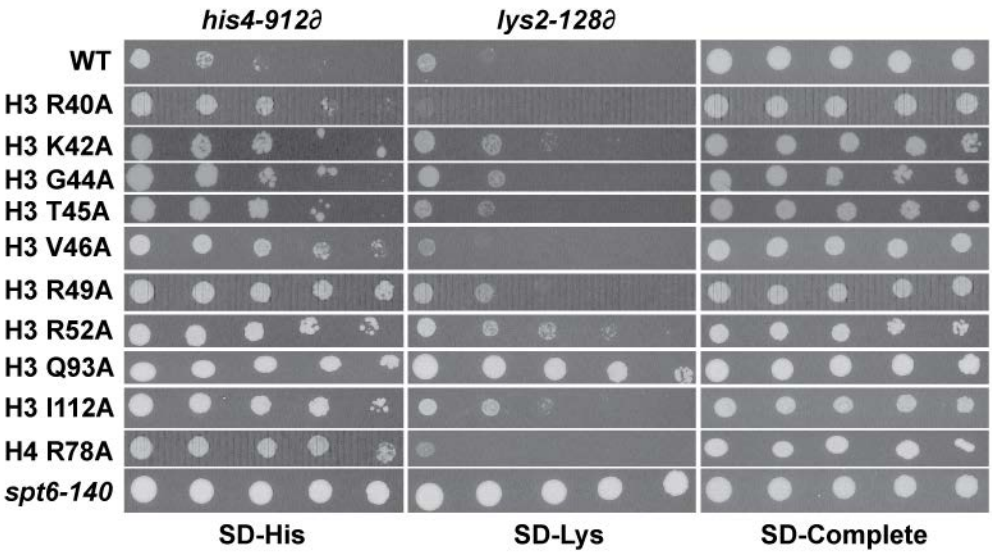

Figure 17. Alanine-substituted DNA entry-exit site residues cause Spt- phenotypes.

The Spt$^-$ phenotype was assessed at two Ty *∂* element insertion mutations by five-fold serial dilution and plating on selective media. Library plasmids were transformed into a strain containing both alleles (KY3502) for plasmid shuffling. The control strain is *spt6-140* (KY319). Cells were grown on SD-Complete as a growth control and on SD-His and SD-Lys to assess suppression of the Ty *∂* element insertions. Plates were incubated at 30°C for 5 days.

Next, we sought to determine whether our termination mutants exhibit the ability to bypass the requirement for upstream activating sequences (UAS) for high levels of expression. Early screens for *BUR* genes identified *HHT1 (BUR5)*, one of the two yeast genes encoding histone H3, as a gene that, when mutated, supports high levels of transcription at the *suc2*Δ*uas(-1900/-390)* allele (PRELICH AND WINSTON 1993). Therefore, we speculated that mutations to the DNA entry-exit site that conceivably reduce nucleosome occupancy on chromatin may exhibit the Bur$^-$ phenotype. When tested, we observed that most of our mutants grow to some degree on YP + Sucrose, indicating the ability to bypass deletion of the *SUC2* UAS (Figure 18). Some mutants, in fact, are comparable to *spt10*Δ, which is the complete deletion of a gene required for replication-

coupled activation of gene promoters (XU *et al.* 2005) and was previously shown to exhibit a strong Bur$^-$ phenotype (PRELICH AND WINSTON 1993).

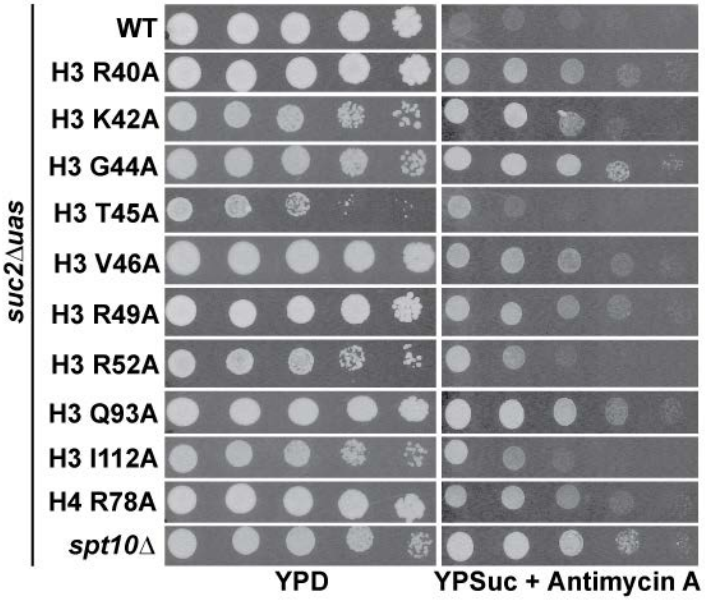

Figure 18. DNA entry-exit site mutants exhibit the Bur$^-$ phenotype.

The Bur$^-$ phenotype was assessed in a *suc2*Δ*uas(-1900/-390)* strain (KY3503) by five-fold serial dilution and plating on selective media. Control strain is *spt10*Δ (KY325). Cells were grown on YPD as a growth control and on YP-Sucrose + 1 μg/mL antimycin A to assess for bypass of the *SUC2* UAS for sufficient levels of expression for growth on media containing sucrose. Plates were incubated at 30°C for 4 days.

Another reporter that reflects a disruption of chromatin structure is the SWI/SNF Independent (Sin$^-$) phenotype. For normal induction of galactose metabolism genes, chromatin remodeling by the SWI/SNF complex at the promoters of those genes is required for gene activation (BURNS AND PETERSON 1997). Mutants that greatly perturb chromatin structure are able to bypass the requirement for this remodeling activity, including mutations to the two copies of

the genes encoding histone H3 (KRUGER *et al.* 1995). DNA entry-exit site substitutions cause very little growth on galactose-containing medium in the context of the *snf2*Δ mutation, suggesting that their effects on transcription are distinct from previously published Sin$^-$ mutants (H3 V117A, Figure 19) (HAINER AND MARTENS 2011a).

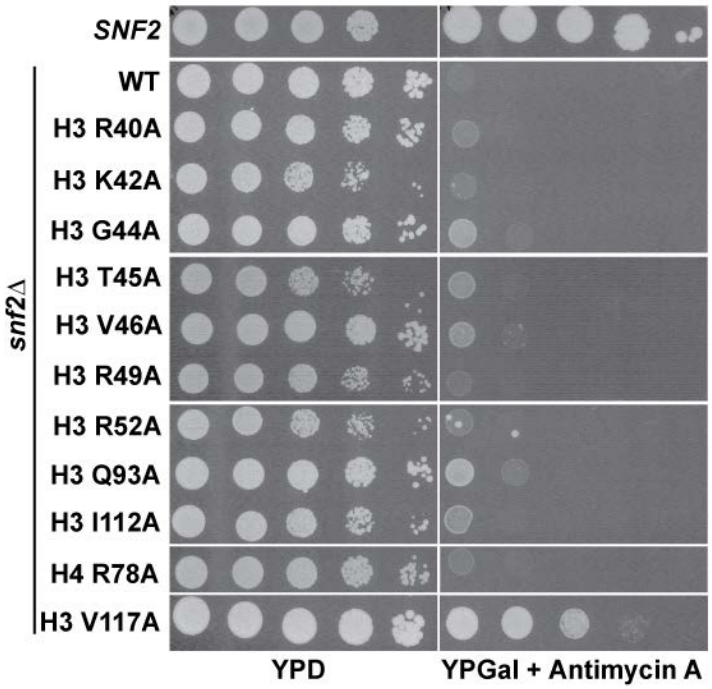

Figure 19. DNA entry-exit site mutants are distinct from Sin$^-$ mutants.

The Sin$^-$ phenotype was assessed in a *snf2*Δ strain (KY3354) by ten-fold serial dilution and plating on selective media. Cells were grown on YPD as a growth control and on YP-Galactose + 1 μg/mL antimycin A to assess bypass of chromatin remodeling for *GAL*-induction. Plates were incubated at 30°C for 3 days.

Cryptic transcription initiation is often detected in genetic backgrounds where improper chromatin architecture permits initiation from a promoter that is normally occluded by nucleosomes in the wild type background (KAPLAN *et al.* 2003). Interestingly, prevention of cryptic initiation requires suppression of histone turnover and resulting histone acetylation along open reading frames, which cascades from proper placement of H3K36me^3 (VENKATESH *et al.*

2012). Led by the findings described above (reduced H3K36me^3, Figure 14), I tested for this phenotype in our mutants via a published cryptic initiation reporter (Figure 20) (CHEUNG *et al.* 2008). In general results of the dilution assay agree with Northern blot analysis of another cryptic initiation-sensitive gene, *STE11* (Figure 21). While most DNA entry-exit site mutants show cryptic initiation via the two assays to some degree, intensities of the phenotypes differ slightly. In general, H3 R49A appears to cause the most pronounced cryptic initiation defect. Interestingly, R52A has a somewhat less pronounced cryptic initiation phenotype, despite its strong H3K36me^3 defect. For the growth-based assay, this may be due to the overall reduced growth rate of the H3 R52A mutant.

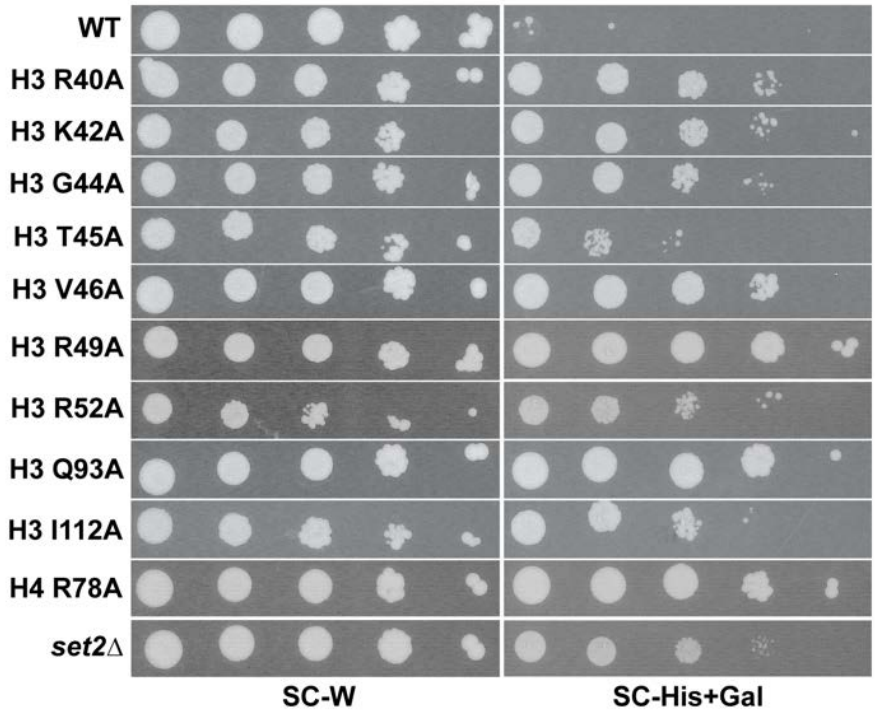

Figure 20. All termination mutants have cryptic initiation phenotypes.

The cryptic initiation phenotype was assessed in a reporter strain (KY3506) by ten-fold serial dilution and plating on selective media. Cells were grown on SC-W as a growth control and on SC-His+Gal to assess for expression of the *GAL1p-FLO8-HIS3* reporter. Plates were incubated at 30°C for 4 days.

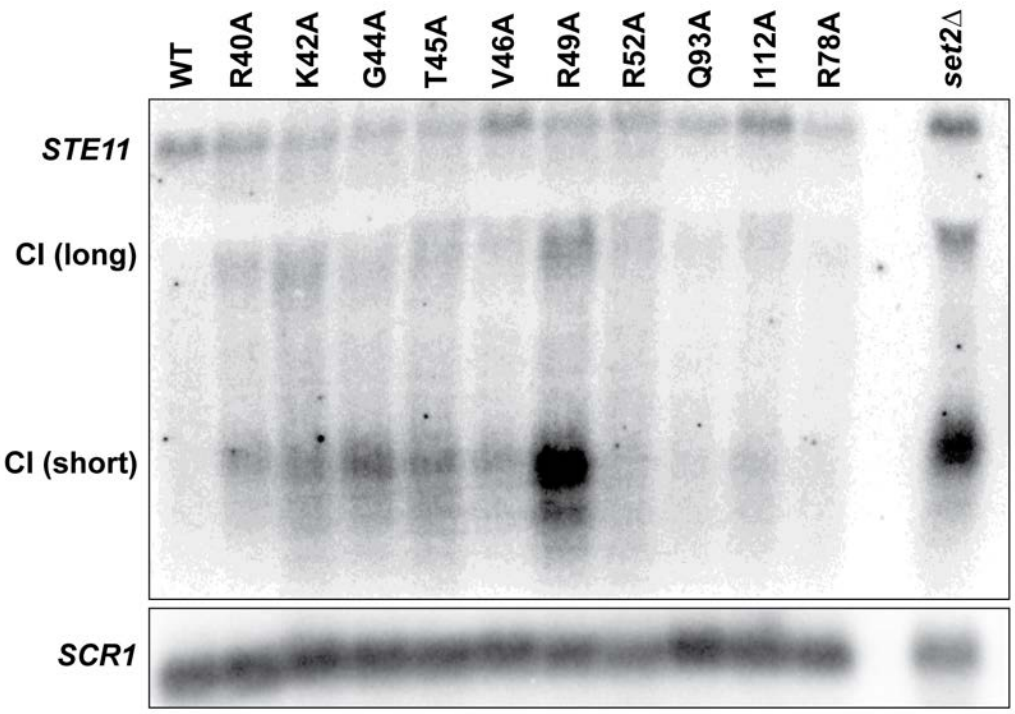

Figure 21. DNA entry-exit site mutants have cryptic initiation defects at *STE11*. Northern blots of DNA entry-exit site mutant plasmids shuffled into KY812. The probe for *STE11* is downstream of cryptic initiation sites internal to the gene; two cryptic initiation (CI) products can be detected. This experiment has been performed in biological triplicate.

2.3.5 **Perturbations to the DNA Entry-Exit Site Reduce Histone Occupancy on Chromatin**

Due to the role of the DNA entry-exit site in nucleosome occupancy, I asked if I could observe this loss at target candidate genes in the termination mutants. Therefore, I performed histone ChIP experiments at various types of loci throughout the yeast genome. First, I tested whether mutations to histone H3 would reduce its chromatin occupancy at the snoRNAs, which provides a possible explanation for defects in transcription termination observed at those loci. I have excluded H3 R40A from this study since it appears wild type for transcription termination at all loci tested by

Northern analysis. Indeed, DNA entry-exit site mutants appear to have reduced H3 occupancy at the *SNR47* gene (Figure 22). Of note, histone occupancy is not perfectly correlated to severity of the termination defect. Additionally, I examined H3 occupancy at a strongly-expressed protein-coding gene, *PMA1*, as well as a transcriptionally silent region of the yeast genome, *TELVI* (Figure 23). These locations also had reduced H3 occupancy in some DNA entry-exit site mutants, suggesting widespread loss in H3 occupancy.

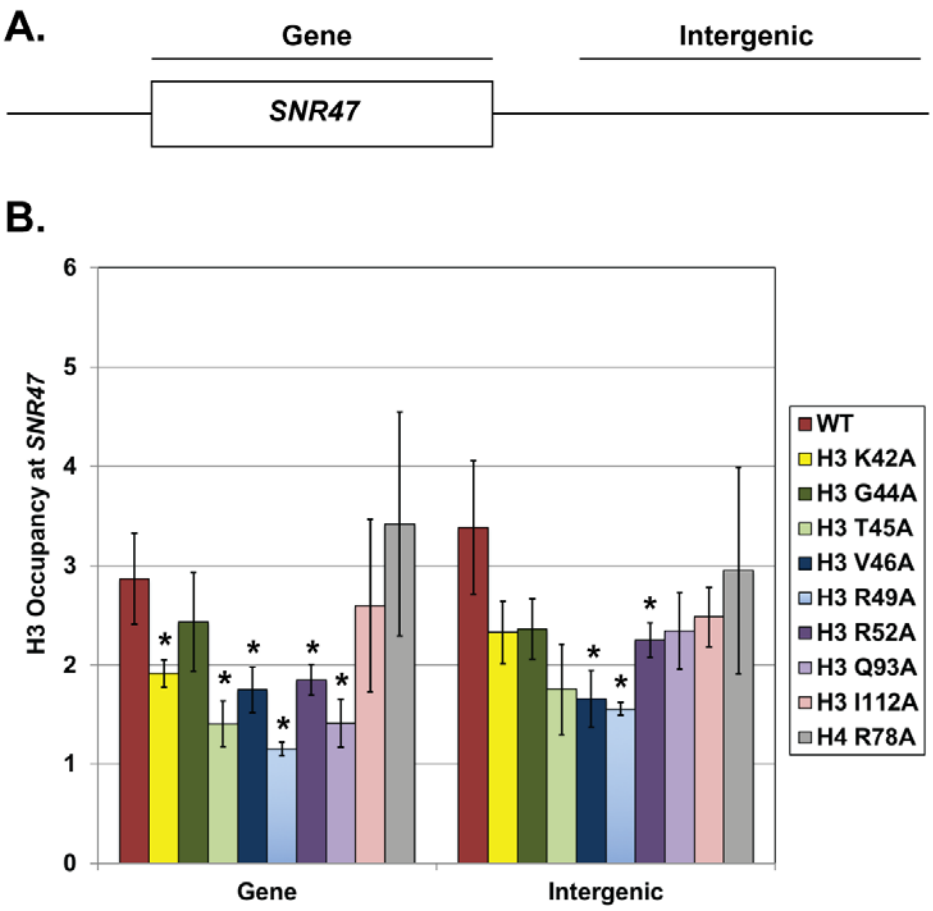

Figure 22. Many DNA entry-exit site mutants have lower H3 occupancy at *SNR47*. **(A)** qPCR product locations over *SNR47*. **(B)** Chromatin immunoprecipitation was performed with an antibody against histone H3. qPCR was performed over *SNR47* or at a downstream location indicated

in A, respectively. The mean of three biological replicates is reported as C_t value of the IP relative to C_t value of the input for each sample. The error bars indicate the SEM. Asterisks represent $p < 0.05$ by a Student's t-test.

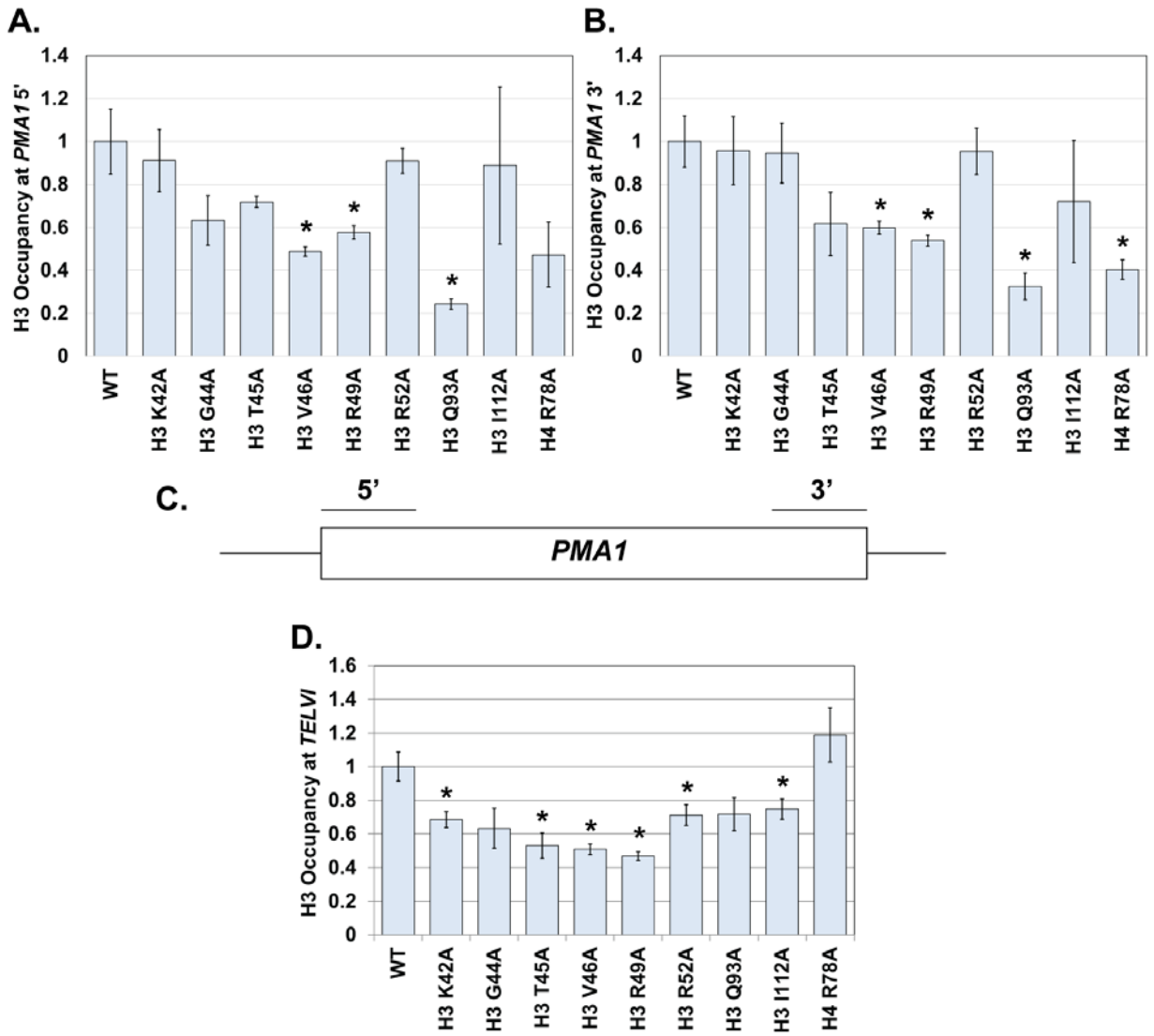

Figure 23. Mutations to the DNA entry-exit site reduce H3 occupancy at protein-coding and silent loci.

(A, B) Chromatin immunoprecipitation was performed with an antibody against histone H3. qPCR was performed over the strongly transcribed *PMA1* gene at a 5' or 3' location indicated in C,

respectively. **(C)** Diagram of *PMA1* locus and qPCR primer set positions. **(D)** ChIP for H3 at a transcriptionally inactive region near *TELVI*. The mean of three biological replicates is reported for each sample. The error bars indicate the SEM. Asterisks represent $p < 0.05$ by a Student's t-test.

2.4 **DISCUSSION**

In this study, I used a comprehensive alanine-scanning histone mutant plasmid library (NAKANISHI *et al.* 2008) and a well-established termination reporter (CARROLL *et al.* 2004) to identify residues in the DNA entry-exit site of the nucleosome required for proper transcription termination (Figures 10-12). A role for this surface of the nucleosome in termination is in line with its history as a regulator of nucleosome accessibility. As early as 1983, Frado *et al.* described the DNA entry-exit site as the binding location for histone H1, a non-core linker histone involved in stabilizing higher-order chromatin structures in higher eukaryotes (FRADO *et al.* 1983). Binding of proteins like H1 at the interface of the DNA entry-exit site and linker DNA between neighboring nucleosomes is hypothesized to oppose nucleosome breathing, where DNA near the entry-exit sites peels away from histones to allow transient DNA accessibility without dissociating the nucleosome entirely (POLACH AND WIDOM 1995).

Subsequent studies have argued breathing as an important feature for activities of chromatin modifiers, like remodelers, that may capture nucleosomes in their open state as a primary means of invasion (LI AND WIDOM 2004; LI *et al.* 2005), and as an important negative regulator of RNA polymerase progression, where empirically-determined speed of DNA rewrapping *in vitro* is hypothesized to pose a kinetic barrier to Pol II (LI *et al.* 2005). *In vivo,* studies on transcriptional regulation of the well-characterized *SRG1-SER3* locus show that mutations to the DNA entry-exit site derepress *SER3* expression in inappropriate nutritional contexts through the loss of critical nucleosome barriers (HAINER AND MARTENS 2011a), supporting the requirement for an only transiently accessible nucleosome in regulating transcription. Consistent with the hypothesis that alanine-substituted residues in the DNA entry-

exit site result in partially destabilized nucleosomes, the termination mutants identified here display phenotypes that are characteristic of disrupted chromatin structure (Figures 17-21).

Previous work in other labs also describes a role for the DNA entry-exit site as a binding surface for factors involved in histone post-translational modifications. One of these, Rtt109, is a histone acetyltransferase required for genome stability and replication-coupled nucleosome synthesis (ZHANG *et al.* 2018). Importantly, despite structural evidence that Rtt109 binds directly to the DNA entry-exit site near its acetylation target H3K56, termination mutants do not display altered H3K56ac levels (Figure 16), suggesting *de novo* nucleosome assembly is likely not impaired in these mutants and is therefore likely not an indirect cause for reduced termination efficiency. Additionally, the DNA entry-exit site has been described as a binding site for the histone methyltransferase Set2, which co-transcriptionally methylates histone H3 at K36 (DU AND BRIGGS 2010; ENDO *et al.* 2012). The methylation activity of Set2 is required for transcriptional fidelity, especially in preventing cryptic initiation from intragenic promoters (CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005). In agreement with previous evidence that perturbations to the DNA entry-exit site causes reduced H3K36me^3 (DU AND BRIGGS 2010; ENDO *et al.* 2012) and that this post-translational histone modification is required for transcription termination (TOMSON *et al.* 2013), I show that several termination mutants in this study indeed have reduced ability to place this mark. Interestingly, the two strongest termination mutants, H3 T45A and R52A, differ in their abilities to support H3K36me^3, suggesting that placement of this modification alone does not fully explain the role of the DNA entry-exit site in transcription termination.

Though no direct interaction between Set1, the histone methyltransferase for H3K4me, and the DNA entry-exit site has been reported, two DNA entry-exit site mutants are impaired in

placement of this transcription-coupled modification. This post-translational histone modification is well conserved, and is both activating and repressive at different loci in different species (STRAHL *et al.* 1999; BRIGGS *et al.* 2001; BERNSTEIN *et al.* 2002; SANTOS-ROSA *et al.* 2002). More recent work has described a role for H3K4me in NNS-dependent transcription termination of several snoRNAs, likely through its role in modulating acetylation levels through recruitment of paired histone acetyltransferase and deacetylase complexes that regulate the opening and closing of chromatin structure, respectively (TERZI *et al.* 2011). Here, I have linked Set1-dependent H3K4 methylation to residues in the DNA entry-exit site by demonstrating that H3 V46A and H3 I112A have reduced H3K4me^{2} (both mutants) and H3K4me^{3} (H3 I112A). Further work is required to better understand this relationship.

Together, these data suggest a requirement for the DNA entry-exit site of the nucleosome in transcription termination. In addition to this defect, the DNA entry-exit site mutants exhibit a myriad of transcription-related phenotypes, including improper placement of H3K4me$^{2/3}$, H3K36me^{3}, and H2Bub, varying Spt^{-} phenotypes at two reporter alleles, the Bur^{-} phenotype, and cryptic initiation (Table 4). As expected, mutations to the DNA entry-exit site also cause loss of histone occupancy on chromatin to various degrees (Figures 22-23), but the loss of occupancy does not correlate with strength of the termination defects of these mutants. It is possible that loss of occupancy, perhaps in addition to loss of other important functions, lead to stronger termination defects in H3 T45A and R52A, for example. These data align to suggest that the DNA entry-exit site has wide roles in chromatin maintenance, and that these pathways are required for productive transcription termination.

Table 4. Phenotypes of DNA entry-exit site mutants.

Strain	Term.	H3K4me	H3K36me	H2Bub	Bur	CI	Sin	Spt*
H3 R40A	WT	WT	WT	--	+++	++	WT	+/-
H3 K42A	++	WT	WT	--	++	++	WT	+/+
H3 G44A	++	WT	-	---	+++	++	-	+/+
H3 T45A	+++	WT	-	--	+	+	-	+/+
H3 V46A	+	-	--	--	+++	++	-	++/WT
H3 R49A	++	WT	---	--	+++	+++	WT	+++/+
H3 R52A	+++	WT	---	--	+	++	WT	+++/++
H3 Q93A	++	WT	WT	---	+++	+++	-	+++/+++
H3 I112A	++	--	--	WT	++	++	WT	+++/+
H4 R78A	++	WT	-	---	+++	+++	WT	+++/WT

WT: wild type-like growth; -/--/---: decreased signal relative to wild type; +/++/+++ increased signal relative to wild type; **his4-912∂/lys2-128∂*

3.0 PROPER TRANSCRIPTION DEPENDS ON NUCLEOSOME-DEPENDENT MAINTENANCE OF CHROMATIN ARCHITECTURE

Contents of this chapter will be published as part of a more comprehensive manuscript with portions of Chapter 2. Construction of the H3 K36A, R52A plasmid used in Figure 31 was performed by Julia Seraly.

3.1 **INTRODUCTION**

Eukaryotic transcription occurs not in the context of naked DNA, but of chromatin, which consists of repeated nucleosomes. These "beads on a string" are made up of approximately 147 bp of DNA wrapped around a histone core. The histone core contains two copies each of H2A, H2B, H3, and H4, where H3-H4 pairs form a heterotetramer to which two H2A-H2B heterodimers bind (LUGER *et al.* 1997). RNA polymerase II, which transcribes most of the *Saccharomyces cerevisiae* genome, requires auxiliary factors to overcome the barrier imposed by nucleosomes. These factors perform a myriad of functions that modify the chromatin landscape, including addition and removal of histone post-translational modifications that alter protein-protein or protein-DNA interactions, replacement of canonical histones with specialized histone variants, or physical repositioning, removal, or replacement of whole nucleosomes or portions of nucleosomes (TEVES *et al.* 2014). Chromatin modifying factors such as these exploit the nucleosome's natural equilibrium between open and closed states, termed "DNA breathing" (LI AND WIDOM 2004). This breathing results from the asymmetrical unwrapping of the first approximately 15 bp of nucleosomal DNA from the

histone core (NGO *et al.* 2015; CHEN *et al.* 2017), giving *trans*-acting factors their initial opportunity to invade the nucleosome and perform further chromatin remodeling or modification (LI AND WIDOM 2004; LIU *et al.* 2017).

The surface of the nucleosome that interacts with this transiently exposed stretch of DNA is the DNA entry-exit site, composed primarily of the αN helix of histone H3, and stabilized by the nearby H2A docking domain and H4 tail (LUGER *et al.* 1997). I have recently discovered that the DNA entry-exit site of the nucleosome is required for termination of RNA polymerase II (Pol II)-transcribed small, noncoding RNAs by the Nrd1-Nab3-Sen1 pathway (Chapter 2). Through a genetic screen, I determined that 13 residues in H3, H2A, and H4, when substituted for alanine, cause terminator readthrough of a reporter (CARROLL *et al.* 2004). I confirmed this defect at several native snoRNA loci, demonstrating varying levels of terminator readthrough at different loci, though H3 T45A and R52A consistently caused the strongest defects.

Given the requirement of the DNA entry-exit site in chromatin accessibility for processes like transcription, I tested the 13 termination mutants for chromatin- and transcription-related phenotypes. In addition to their termination defects, the DNA entry-exit site mutants also exhibit defects in histone modifications associated with active transcription (H3K4me$^{2/3}$, H3K36me^{3}, and H2Bub), as well as reduced histone occupancy at several genomic loci. Some mutants also exhibit Spt^{-} (SIMCHEN *et al.* 1984; WINSTON *et al.* 1987; FASSLER AND WINSTON 1988), Bur^{-} (PRELICH AND WINSTON 1993), and cryptic initiation (KAPLAN *et al.* 2003) phenotypes, growth-based reporter phenotypes that suggest changes in chromatin architecture often associated with deregulation of transcription, consistent with these residues' roles in providing a stable barrier for DNA-templated activities through restriction of DNA breathing.

The DNA entry-exit site has also been implicated in proper transcription-dependent nucleosome occupancy (HAINER AND MARTENS 2011a; HYLAND *et al.* 2011). Based on published data and my previous work, I hypothesized that the observed termination defect in DNA entry-exit site mutants is possibly linked to multiple mechanisms, including impaired H3K36me^3, altered transcription rate, and improper nucleosome occupancy across genes and near terminators. In this study, I show that mutating the DNA entry-exit site of the nucleosome (H3 T45A and R52A) does indeed lead to elevated transcript levels downstream of many yeast genes, consistent with a termination defect. The link between Set2-catalyzed H3K36me remains incompletely understood but making a double mutant histone with alanine substitutions at K36 and R52 results in synthetic lethality, suggesting a relationship between the pathways involving these residues. Genome-wide nucleosome mapping in a DNA entry-exit site mutant (H3 R52A) reveals decreased occupancy compared to wild type cells. This decreased nucleosome occupancy does not correlate with altered transcription elongation rate in DNA entry-exit site mutants as measured by a well-studied, ChIP-based assay (MASON AND STRUHL 2003), but further analysis will be required to better assess elongation rate. Importantly, I find that increasing nucleosome occupancy near the terminator of a select snoRNA gene by integrating a nucleosome superbinder sequence suppresses the termination defect observed upon mutation of the DNA entry-exit site. Overall, these data suggest the DNA entry-exit site plays a major role in the proper termination of transcription.

3.2 **MATERIALS AND METHODS**

3.2.1 **Plasmid Construction**

I constructed a plasmid allowing integration of a histone superbinder (SB) sequence at any location within the yeast genome, selectable with a *URA3* marker, using the pMPY-3xHA backbone (SCHNEIDER *et al.* 1995). I amplified the SB sequence from a TOPO TA vector (HAINER *et al.* 2015) flanked by *NotI* sites or *EcoRI* and *XhoI* sites. These fragments and the pMPY-3xHA vector were digested by the appropriate restriction enzymes and ligated together in a stepwise fashion to produce a plasmid containing SB-*URA3*-SB (KB1479), which can be integrated into the yeast genome, selected for, and then forced to lose the *URA3* marker through loop-out of *URA3* and recombination of the SB sequence on media containing 5-FOA.

A talented undergraduate, Julia Seraly, used site-directed mutagenesis (Agilent 200523) to substitute H3 K36 for an alanine residue in the H3 R52A SHIMA library plasmid (NAKANISHI *et al.* 2008). The resulting plasmid was sequenced to verify the presence of both K36A and R52A substitutions, and to ensure no off-target mutations were made. The resulting double mutant plasmid was used in serial dilution assays to assess genetic interaction of H3 K36 and H3 R52.

3.2.2 **Yeast Strains and Media**

All *Saccharomyces cerevisiae* strains used in this study are isogenic to FY2, a *GAL2*$^{+}$ derivative of S288C (WINSTON *et al.* 1995). Unless otherwise noted, all experiments were performed with KY812 (Table 5) transformed with CEN/ARS plasmids expressing *TRP1*-marked, alanine-substituted histones H3 and H4 (*hht2-HHF2* or *HHT2-hhf2*, respectively) (NAKANISHI *et al.* 2008)

or a *TRP1*-marked plasmid expressing the wild type *HHT2-HHF2* cassette. Lithium acetate yeast transformations were performed as previously described (BECKER AND LUNDBLAD 2001). Wild type histone plasmids were shuffled with *TRP1*-marked mutant histone plasmids by sequentially passaging transformants three times on SC-Trp medium containing 2% dextrose and 0.1% 5-FOA.

Elongation rate assays required a strain with HA-TBP and FLAG-Rpb3, endogenous deletions of the genes encoding histones H3 and H4, and a yeast "long gene" (*YLR454W)* inducibly expressed from the *GAL1* promoter. This strain was constructed via multiple genetic crosses: first KY811 (*hht1-hhf1*Δ*::LEU2, hht2-hhf2*Δ*:KanMX*) by PY408 (*HA-SPT15*), then the resulting strain (*hht1-hhf1*Δ*::LEU2, hht2-hhf2*Δ*:KanMX, HA-SPT15*) by KY2641 (*RPB3::3xFLAG::KanMX, KanMX::GAL1p-YLR454W).* The resulting strain (KY3232) was transformed with the appropriate *TRP1*-marked histone plasmids and shuffled as discussed above to generate wild type and mutant histone strains.

The SB-containing strain was generated by amplifying the SB-*URA3*-SB cassette in two pieces from KB1479. Outside primers had 40 bp of homology to the region downstream of *SNR48,* and universal inside primers annealed to the *URA3* gene. Fragments overlapped such that upon cotransformation into yeast, homologous recombination joined the two fragments and integrated the full-length cassette. Candidate transformants were verified by genotyping through colony PCR and Sanger sequencing. The resulting strain (KY3221) was transformed with the appropriate *TRP1*-marked library plasmids and shuffled as discussed above to generate wild type and mutant histone strains.

To generate a strain for verifying expression levels of H3 from the plasmid encoding the double substitution of K36A and R52A, I crossed a strain expressing HA-tagged *HHT1* (KY2143) with a histone shuffle strain (KY812) to obtain a haploid whose only functional chromosomal copy

is the tagged copy (KY3511, *HHT1-HA::KanMX (hht2-hhf2)*Δ*::KanMX).* Using this strain, I was able to detect expression of H3 from either a wild type, K36A, R52A, or K36AR52A plasmid in a strain with functional, wild type histone H3 to support viability.

Yeast strains were grown in SC-Trp medium containing 2% dextrose for all experiments unless otherwise noted.

3.2.3 **Northern Blot Analysis**

Total RNA was isolated from mid-log phase yeast cultures (OD_{600} = 0.8 – 1.0) by hot acid phenol extraction as described previously (COLLART AND OLIVIERO 2001). 20 µg of each sample was analyzed by Northern blot analysis as described previously (SWANSON *et al.* 1991). Double-stranded DNA probes were synthesized by random-prime labeling of PCR fragments with [α-^{32}P]dATP and [α-^{32}P]dTTP (PerkinElmer, Waltham, MA). PCR fragments are as follows: *SNR47-YDR042C* (-325 to -33 of *YDR042C*), *SNR48-ERG25* (-746 to -191 of *ERG25*), *SNR48* gene (-70 to +92 of *SNR48*), *SNR13-TRS31* (-231 to +449 of *TRS31*), *YLR454W* (+1982 to +2275 of *YLR454W*), *STE11* (+1868 to +2110), and *SCR1* (-242 to +283 of *SCR1*) (primer sequences listed in Table 6).

3.2.4 **Western Blot Analysis**

Total protein was isolated from mid-log phase yeast cultures (OD_{600} = 0.8 – 1.0) by bead beating in 20% trichloroacetic acid. Protein samples were resolved on SDS-PAGE gels and transferred to nitrocellulose or PVDF membrane. The membranes were incubated with primary antibody against HA (12CA5 1:3000, Roche, Basel, Switzerland), total H3 (ab1791 1:15,000, Abcam, Cambridge,

UK), H3K36me^3 (ab9050 1:1000, Abcam, Cambridge, UK), or G6PDH (A9521 1:30,000, Sigma-Aldrich, St. Louis, MO) as the loading control. Membranes were then incubated with 1:5,000 α-mouse or α-rabbit secondary antibody (GE Healthcare, Little Chalfont, UK). Proteins were visualized using chemiluminescence substrate (Pico, PerkinElmer, Waltham, MA) and the ChemiDoc XRS imaging platform (BioRad, Hercules, CA).

3.2.5 **Chromatin Immunoprecipitation and Quantitative PCR (ChIP-qPCR)**

Chromatin was isolated from 250 mL of yeast cells grown to early log phase (OD_{600} = 0.5 – 0.8) as described previously (SHIRRA *et al.* 2005). ChIPs were performed by incubating sonicated chromatin overnight at 4°C with one of the following primary antibodies: HA (25 µL sc-7392 AC, Santa Cruz, Dallas, TX), FLAG (30 µL A2220, Sigma-Aldrich, St. Louis, MO), or H2A (2 µL 39235, ActiveMotif, Carlsbad, CA). Antibody-chromatin complexes were purified with Protein A conjugated to sepharose beads (30 µL, GE Healthcare, Little Chalfont, UK). DNA was column purified (QiaQuick, Qiagen, Hilden, Germany) and analyzed by qPCR with Maxima 2X SYBR Master Mix (ThermoFisher, Waltham, MA). The primer set analyzed for HA-TBP occupancy encompasses *GAL1p-YLR454W (*−194 to +35). For the transcription elongation experiment, primer sets analyzed for Flag-Rpb3 occupancy include: *GAL1p-YLR454W (*−194 to +35), 2 kb (+1982 to +2275), 4 kb (+4069 to +4268), and 8 kb (+7701 to +7850). For superbinder integration, H2A occupancy was measured downstream of *SNR48* (+113 to +375 of *SNR48*). Results are presented as the average of two technical replicates each of three biological replicates, where error bars represent the SEM of these experiments.

3.2.6 **Total RNA Sequencing (RNA-seq)**

Total RNA was isolated from mid-log phase yeast cultures (OD_{600} = 0.8 – 1.0) by hot acid phenol extraction as described previously (COLLART AND OLIVIERO 2001). RNA was DNase treated (TURBO DNase, ThermoFisher, Waltham, MA) and sent to the Health Sciences Sequencing Core at UPMC Children's Hospital for RiboZero treatment, library preparation, and sequencing. Libraries were quantified using Qubit and TapeStation and pooled for paired-end sequencing on an Illumina NextSeq 500. Sequencing reads were aligned with HISAT2 (KIM *et al.* 2015) and low quality reads were filtered using the SAMtools suite (LI *et al.* 2009). Differential expression was determined using DESeq2 (LOVE *et al.* 2014). Heatmaps were generated using deepTools2 (RAMIREZ *et al.* 2014).

3.2.7 **Micrococcal Nuclease Treatment and Sequencing (MNase-seq)**

Mononucleosomes were prepared essentially as previously described (WAL AND PUGH 2012). Briefly, cells were grown in SC medium to OD_{600} = 0.8 and crosslinked with formaldehyde at a final concentration of 1%. 100 mL of cells were pelleted, resuspended in FA buffer (50 mM HEPES/KOH, pH 8.0, 150 mM NaCl, 2.0 mM EDTA, 1.0% Triton X-100, and 0.1% sodium deoxycholate), and lysed by bead beating. Cell extracts containing chromatin were pelleted and resuspended in NP-S buffer (0.5 mM Spermidine, 0.075% IGEPAL, 50 mM NaCl, 10 mM Tris–Cl, pH 7.5, 5 mM $MgCl_2$, 1 mM $CaCl_2$), and then subjected to digestion by MNase (ThermoFisher, Waltham, MA). MNase-treated *Kluyveromyces lactis* chromatin was spiked into *S. cerevisiae* samples at a ratio of 1:10. Mixed mononucleosomal DNA was purified using agarose gel electrophoresis and Freeze N' Squeeze Columns (BioRad, Hercules, CA). Sequencing libraries

were prepared using the NEBNext Ultra II kit and NEBNext Ultra sequencing indexes according to manufacturer's instructions (NEB, Ipswitch, MA). Libraries were quantified using Qubit and TapeStation and pooled for paired-end sequencing on an Illumina NextSeq 500 (UPMC Children's Hospital, Health Sciences Sequencing Core, Pittsburgh, PA). Sequencing reads were aligned with HISAT2 (KIM *et al.* 2015) and low quality reads were filtered using the SAMtools suite (LI *et al.* 2009). Nucleosome positioning was analyzed genome-wide and heatmaps were generated using deepTools2 (RAMIREZ *et al.* 2014). Metagene profiling of nucleosome positioning was performed using DANPOS2 (CHEN *et al.* 2013).

3.2.8 **4tU Labeling of Nascent RNA and Sequencing (4tU-seq)**

4-thiouracil (4tU)-labeled RNA was prepared essentially as previously described (BONNET AND PALANCADE 2014; DUFFY *et al.* 2015). 4tU was added to log phase cultures (OD600 = 0.8 – 1.0) to a final concentration of 0.26 mg/mL in DMSO. Cultures were incubated for 5 min at room temperature with constant agitation. 10 OD units of cells were harvested by centrifugation for 3 min at 3000 rpm at 4°C to pellet cells, and then supernatant was completely removed prior to snap-freezing in liquid nitrogen. To prepare labeled RNA, *S. cerevisiae* cells were mixed 10:1 with labeled *S. pombe* cells in Ribopure lysis buffer (Ambion AM1926) as a spike-in control. RNA was isolated from the cell mixture according to kit instructions. After isolation, 4tU-labeled RNA was biotinylated with 1 µg/mL of MTSEA Biotin-XX (Biotium 90066) for 30 min at room temperature in biotinylation buffer (20 mM HEPES, 1 mM EDTA). Meanwhile, streptavidin beads (Invitrogen 65001) were washed in a high salt wash buffer (100 mM Tris, 10 mM EDTA, 1 M NaCl, 0.05% Tween-20) and blocked (high salt wash buffer, 40 ng/uL glycogen) for an hour prior to use. Phenol chloroform extraction was used to quench biotinylation reactions by removing unincorporated

biotin. Isopropanol precipitated RNA was resuspended in 100 µL nuclease-free water and incubated with blocked streptavidin beads for 15 min. Supernatant was removed and set aside as unlabeled RNA. Beads were washed twice with elution buffer (100 mM DTT, 20 mM HEPES, 1 mM EDTA, 100 mM NaCl, and 0.05% Tween-20) and samples were pooled for concentration via MinElute columns and library preparation using the NEBNext Ultra II kit (NEB E7770 or E7760, directional). Libraries were quantified using Qubit and pooled for paired-end sequencing on an Illumina NextSeq 500 (UPMC Children's Hospital, Health Sciences Sequencing Core, Pittsburgh, PA). Sequencing reads were aligned with HISAT2 (KIM *et al.* 2015) and low quality reads were filtered using the SAMtools suite (LI *et al.* 2009). Spike-in normalization and bigWig generation were performed by deepTools2 (RAMIREZ *et al.* 2014).

3.2.9 **ChIP and Genome-Wide Sequencing (ChIP-seq)**

ChIPs were performed as described above. Upon purification of ChIP DNA, *S. cerevisiae* was mixed 10:1 with immunoprecipitated *K. lactis* DNA. Libraries were built according to NEBNext Ultra II (NEB E7645) manufacturer's instructions. Libraries were quantified using Qubit and pooled for paired-end sequencing on an Illumina NextSeq 500 (UPMC Children's Hospital, Health Sciences Sequencing Core, Pittsburgh, PA). Sequencing reads were aligned with HISAT2 (KIM *et al.* 2015) and low quality reads were filtered using the SAMtools suite (LI *et al.* 2009). Spike-in normalization and bigWig generation were performed by deepTools2 (RAMIREZ *et al.* 2014).

Table 5. Yeast strains used in Chapter 3.

Strain	MAT	Genotype
KY811	α	*(hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM1 = *URA3/CEN/ARS/HHT2-HHF2*]
KY812	**a**	*(hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY2143	α	*leu2Δ1 trp1Δ63 HHT1-HA::KanMX*
KY3221	**a**	*SNR48-SB (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3232	α	*RPB3-3xFLAG::KanMX KanMX::GAL1p-YLR454W HA-SPT15 (hht1-hhf1)Δ::LEU2 (hht2-hhf2)Δ::KanMX his3Δ200 lys2-128∂ leu2Δ1 ura3-52 trp1Δ63* [pDM9 = *URA3/CEN/ARS/HHT1-HHF1*]
KY3511	α	*leu2Δ1 ura3-52 trp1Δ63 HHT1-HA::KanMX (hht2-hhf2)Δ::KanMX*

Table 6. Oligonucleotides used in Chapter 3.

Target	Sequence	
SNR47-YDR042C	F	5’-GGCGGTAACGTAAATCAGAGTAGC-3’
Northern probe	R	5’-GAGACCTAGTCGTTTGTTAGCTG-3’
SNR48-ERG25	F	5’-CCTTGGCGCAGAAGACTTTCTCTTC-3’
Northern probe	R	5’-GCATACACAGGCGTACGCATACAAG-3’
SNR48 Gene	F	5’-GCCTTTTTCTTGAATTGTCAATCCGCCC-3’
Northern probe	R	5’-GTATTCATTGAATGGAGAGTACTTAAAC-3’
SNR13-TRS31	F	5’-CGTAGCGCTGCATATATAATGCG-3’
Northern probe	R	5’-GATGCAGAAGTCGCTGTGCTGGAG-3’
SCR1	F	5'-CAACTTAGCCAGGACATCCA-3'
Northern probe	R	5'-AGAGAGACGGATTCCTCACG-3'
STE11	F	5’-GGATGTCACCAGAGGTGGTCAAAC-3’
Northern probe	R	5’-GATTATCAATACAGGGCCTAGTGCCC-3’
YLR454W	F	5’-GAGACATATCATCCACCCTAG-3’
Northern probe	R	5’-GATTTACCTCTGTCGCATCAC-3’
GAL1p	F	5’-GGGGTAATTAATCAGCGAAGCGATG-3’
ChIP Primers	R	5’-CACTTGTACAGTAGAACATTAATCGG-3’
YLR454W 2 kb	F	5’-GAGACATATCATCCACCCTAG-3’
ChIP Primers	R	5’-GATTTACCTCTGTCGCATCAC-3’
YLR454W 4 kb	F	5’-AGATATTACTCGTTGTTCGTGCC-3’
ChIP Primers	R	5’-TCCCAAAACCCTAGTTTAACAG-3’
YLR454W 8 kb	F	5’-GAGGGTCACAGATCTATTACTTGCCC-3’
ChIP Primers	R	5’-GTTGTGAGTTGCTTCAGTGGTGAAGT-3’
SNR48-SB	F	5’-TTACTATGATTAAACAGACCGAGGGAGAA-3’
ChIP Primers	R	5’-AACGGGTTTAAACGGCTGCC-3’

3.3 **RESULTS**

3.3.1 **Genome-Wide Analysis Demonstrates Readthrough Transcription as a Global Phenomenon in DNA Entry-Exit Site Mutants**

To address whether terminator readthrough observed at specific snoRNA loci (Chapter 2) is a general feature of DNA entry-exit site mutants, I selected strains harboring alanine substitutions that cause the most severe defect by Northern blot analysis (H3 T45A and R52A) and subjected them to RNA sequencing (RNA-seq). My spike-in normalized data reveal that many snoRNA loci are indeed improperly terminated in DNA entry-exit site mutants as measured by 3' extension index (3'EI) (Figure 24). 3'EI is the fold change in RNA signal 150 bp downstream of the annotated transcription end site (TES) in the mutant compared to wild type (TOMSON *et al.* 2013; NEMEC *et al.* 2017). By this metric, 30/73 (41.1%) of Pol II-transcribed snoRNAs are read through at least 1.5-fold higher than wild type in the H3 T45A mutant (Figure 24A) and 43/73 (58.9%) are read through at least 1.5-fold higher than wild type in H3 R52A mutant (Figure 24B).

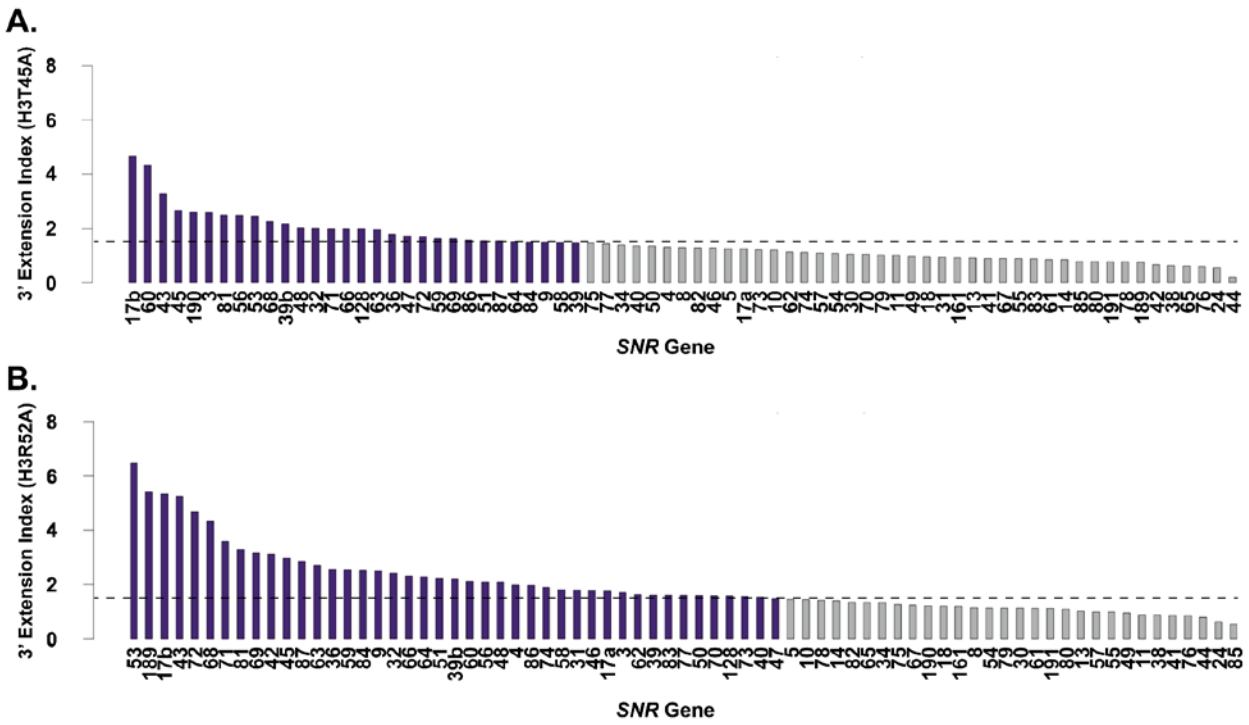

Figure 24. 3' extension index demonstrates global snoRNA readthrough in select entry-exit site mutants.

(A) 3’ extension index measured in the H3 T45A mutant. The mutant/wild type ratio of RNA-seq reads was determined in a window 150 bp downstream of each snoRNA annotated 3’ end. Ratios greater than 1.5 (dotted line) are highlighted in purple. **(B)** 3’ extension index in the H3 R52A mutant.

I additionally asked whether mRNAs are generally improperly terminated when the DNA entry-exit site of the nucleosome is perturbed. To assess this, I centered my RNA-seq data over annotated gene 3’ ends and measured fold-change differences in read counts downstream of these sites in the mutants compared to wild type (Figure 25). In both mutants, ORF transcription is modestly up- or down-regulated at some genes compared to the wild type. Strikingly, however, there is a modest spike of read density immediately downstream of the 3’ end in both mutants by meta-analysis, suggesting that terminator readthrough is a widespread phenomenon in DNA entry-exit site mutants. When transcription levels are assessed in smaller 50 bp windows downstream of

either protein-coding or snoRNA genes, I observe that a large proportion of these readthrough transcripts do extend up to 150 bp from the TES (Figure 26). Together, these results suggest that the DNA entry-exit site is broadly required for termination of transcription at the proper location.

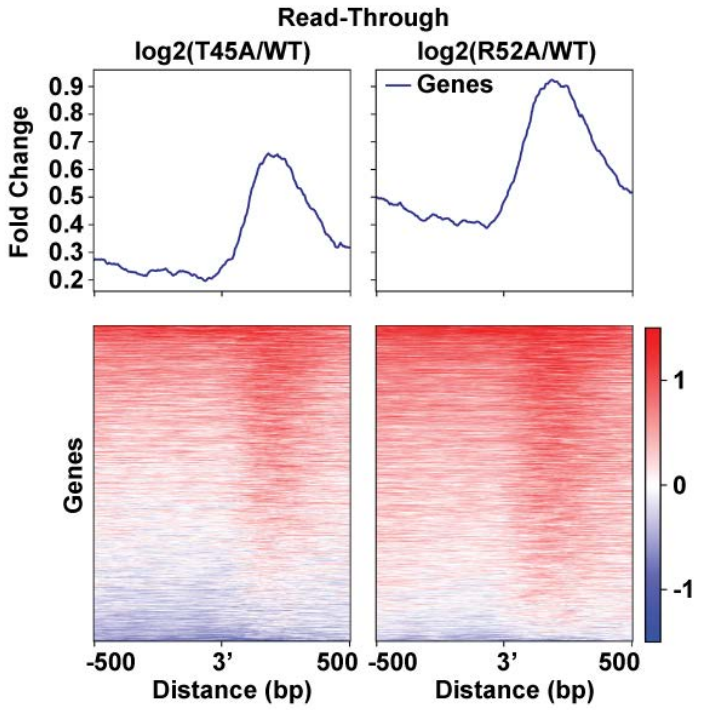

Figure 25. Readthrough of protein-coding genes is a widespread phenomenon in DNA entry-exit site mutants.

Metagene plot (top) and heatmap (bottom) of RNA-seq reads genome-wide. Each gene is represented as a single row in the heatmap. The plot and heatmap are centered around the 3' end of each gene and show 500 bp on either side. Red coloration indicates higher RNA-seq signal in the mutant compared to wild type, and blue coloration indicates downregulation in the mutant. Metaplots indicate greater read density, on average, downstream of genes, suggesting widespread terminator readthrough.

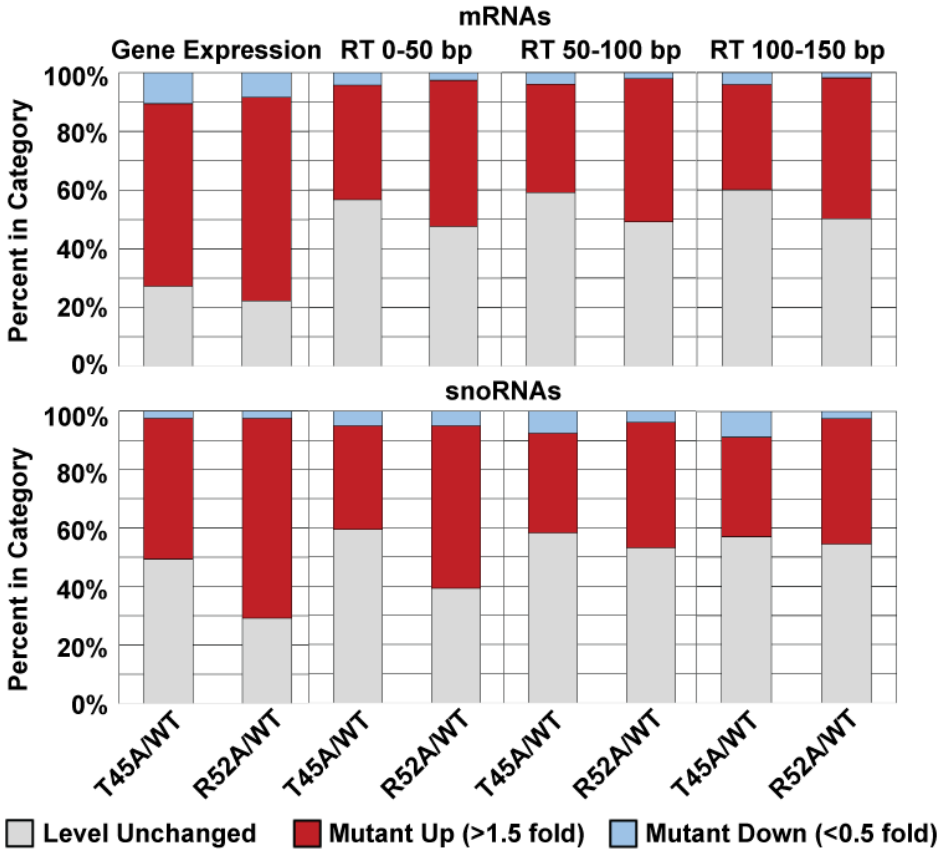

Figure 26. Readthrough of protein-coding genes and snoRNAs extends to 150 bp downstream of the gene.

Percent of genes exhibiting either increased or decreased expression over gene bodies, as well as 0-50 bp, 50-100 bp, and 100-150 bp downstream of genes in H3 T45A and H3 R52A mutants.

To gain a better understanding of the nature of these readthrough transcripts, I generated *de novo* annotations from the wild type and mutant RNA-seq datasets (VENKATESH *et al.* 2016; ELLISON *et al.* 2019). *De novo* annotations were generated from pile-up of raw RNA-seq reads where single transcripts were generated from continuous stretches of reads with a read depth of at least 40 and gaps of no greater than 5 bp. Such an analysis provides assumed transcript isoforms

based only on the RNA sequencing data, not influenced by existing transcript annotations. These annotations reveal 3' extended transcripts at snoRNAs (Figure 27) and protein-coding genes alike.

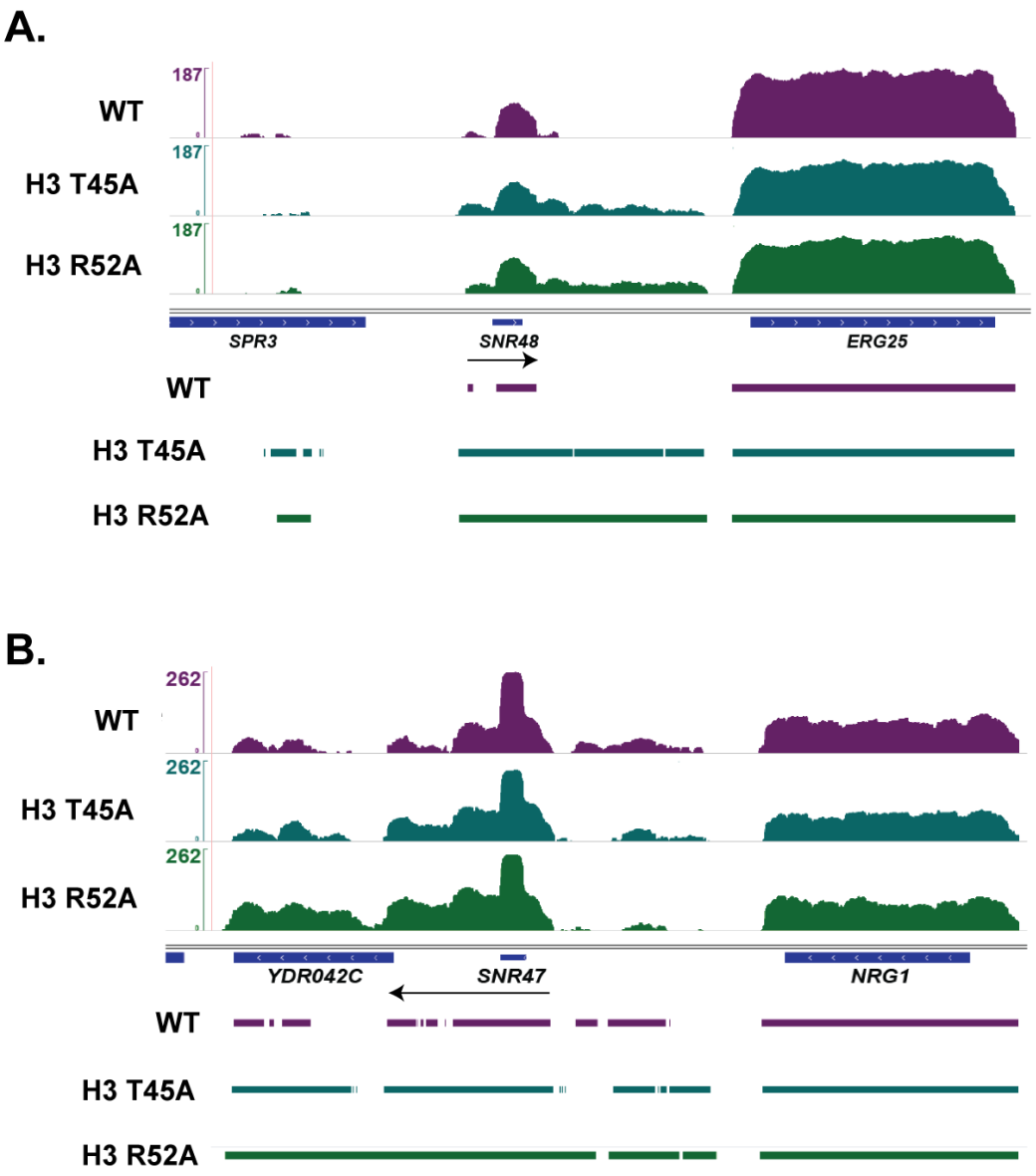

Figure 27. *De novo* transcript annotation shows novel, 3'-extended snoRNA isoforms in termination mutants.

(A) *De novo* annotations across the *SNR48* locus. This locus extends from left to right. **(B)** *De novo* annotations across the *SNR47* locus. The browser tracks represent raw RNA-seq reads and the lines of matching color beneath correspond to the annotations.

3.3.2 **Elongation Rate in DNA Entry-Exit Site Mutants Does Not Correlate with Observed Transcriptional Defects**

Previous work has demonstrated kinetic competition between Sen1 and Pol II for NNS-dependent termination (HAZELBAKER *et al.* 2013). Factors increasing the rate of transcription elongation can cause Pol II to outpace Sen1, resulting in 3' extended transcripts. I hypothesized that the elongation rate in DNA entry-exit site mutants may be faster due to the decrease in nucleosome occupancy. To this end, I have measured elongation rate using a well-characterized ChIP assay (MASON AND STRUHL 2003). Making use of its expression from the *GAL1* promoter, I effectively shut off expression of a yeast long gene, *YLR454W,* by addition of glucose to the growth medium, and then measured Pol II occupancy by FLAG-Rpb3 ChIP-qPCR at four locations over an eight-minute time course (Figure 28A)

Repression of the *GAL1* promoter by Mig1 (Multicopy inhibitor of galactose gene expression 1) is inefficient in some chromatin-related mutants (MALIK *et al.* 2017). Therefore, I monitored proper shutoff of the long gene by additionally performing ChIP assays for HA-tagged TBP at the *GAL1* promoter to ensure proper preinitiation complex (PIC) disassembly upon introduction of glucose, avoiding the risk of the elongation rate results being confounded by reinitiation from the *GAL1* promoter. Reduction in TBP occupancy from zero to eight minutes in glucose indicates that appropriate shutoff of transcription from the *GAL1* promoter occurs at this locus (Figure 28B).

Pol II occupancy across *YLR454W* suggests elongation rates in the mutants are not significantly faster than wild type (Figure 28C-E). In fact, elongation rates in the H3 T45A and H3 R52A mutants appear somewhat slower at this single locus. Importantly, reduction in Pol II occupancy across the time course correlates with a reduction in abundance of the *YLR454W* transcript (Figure 29A). The initial amount of transcript, however, agrees with the lower TBP occupancy in H3 T45A and H3 R52A mutants at zero minutes. To be sure the decrease in TBP occupancy at the *GAL1* promoter is not due to reduced expression of the TBP gene, I performed Western blot analysis and verified that TBP is appropriately expressed in all three genetic backgrounds (Figure 29B). Together, these results suggest that altered rates of transcription elongation do not explain the termination defects observed in DNA entry-exit site mutants. However, further work is required to more thoroughly assess elongation rates in these strains at additional genes.

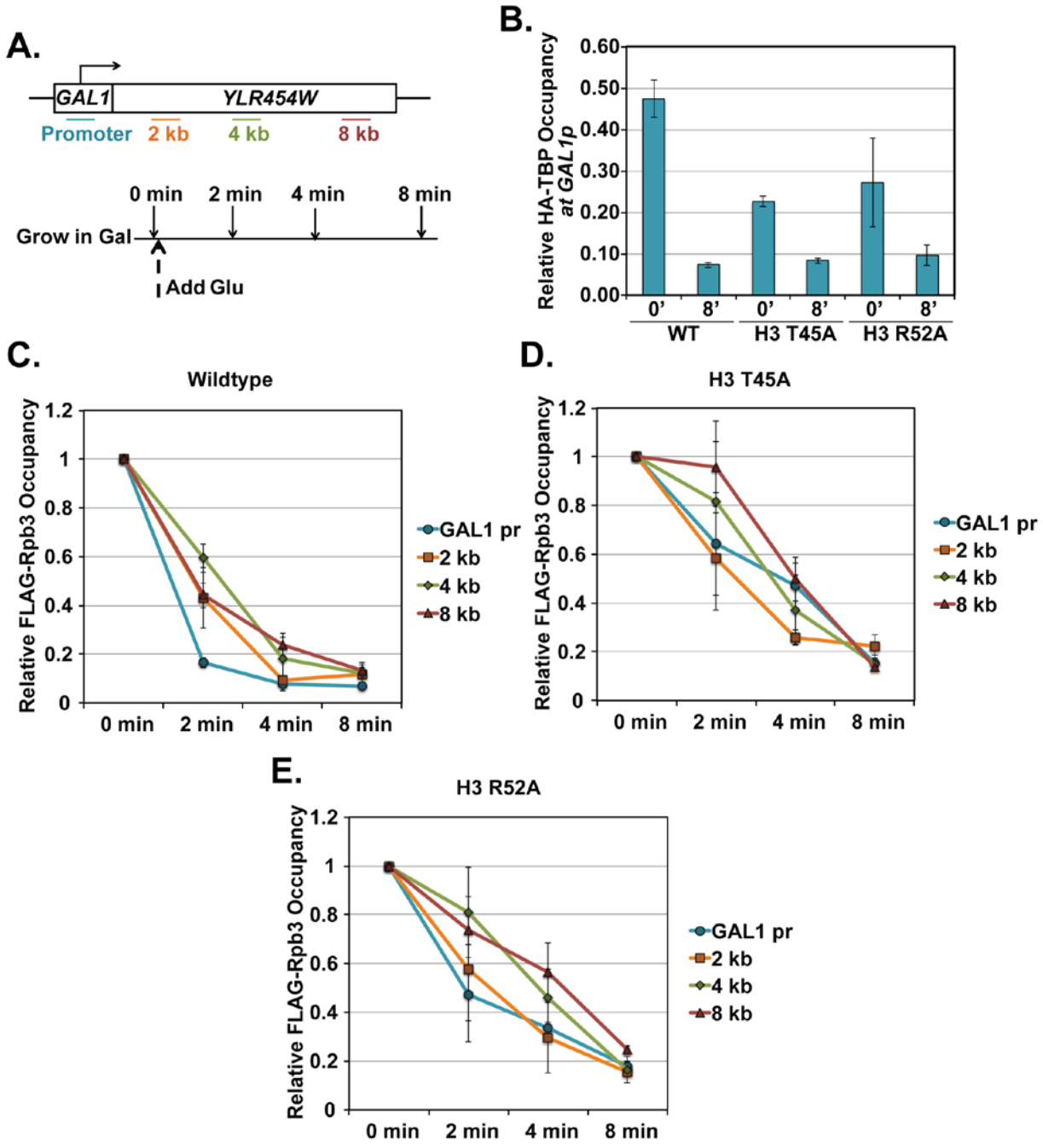

Figure 28. Analysis of transcription elongation rates using a reporter assay demonstrates no significant increase in elongation rates in the H3 T45A and R52A mutants.

(A) Diagram of the *GAL1p-YLR454W* locus. Bars under the locus diagram indicate locations of qPCR primer sets and correspond to the colors in the data graphs in panels C-E. **(B)** Relative TBP occupancy at the *GAL1* promoter in each strain tested by the long gene ChIP experiment. TBP occupancy in the wild

type strain is much higher at 0 minutes, indicating stronger galactose-induced expression of the reporter than in the DNA entry-exit site mutants. Despite the initial difference, TBP occupancy decreases by 8 minutes in all strains. **(C-E)** ChIP-qPCR of FLAG-Rpb3 occupancy at the indicated times after addition of glucose to repress the reporter gene in **(C)** wild type, **(D)** H3 T45A, and **(E)** H3 R52A strains. Each colored line in the graph represents the data for the location that corresponds to the primer location diagram in panel A. Error bars represent the SEM of three independent biological replicates.

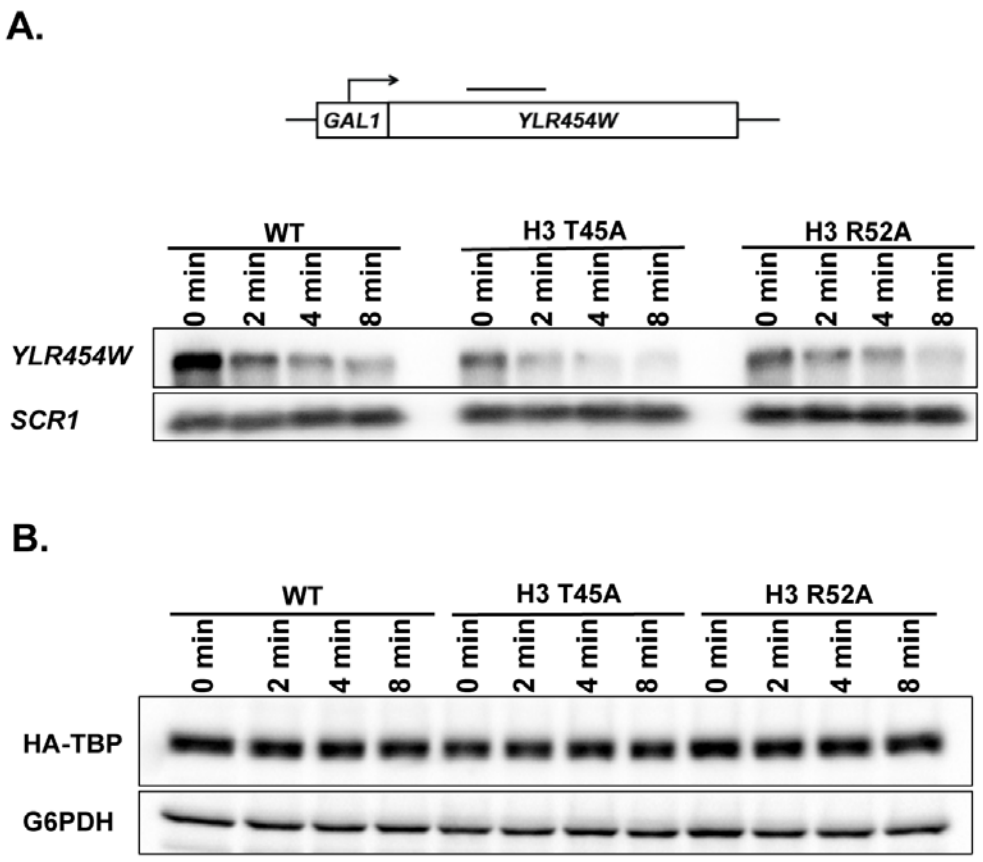

Figure 29. Slower elongation rate in DNA entry-exit site mutants is not due to loss of TBP or failure to shut off the reporter locus.

(A) Northern blot of the *YLR454W* gene at each time point following glucose addition in the elongation rate assay. Location of the Northern probe is as indicated in the locus diagram (top). **(B)** Western blot analysis of HA-TBP levels in DNA entry-exit site mutants compared to wild type. There is no difference in bulk

TBP levels in the histone mutants, suggesting that lower TBP occupancy at the *GAL1* promoter in Figure 28B is likely due to reduced recruitment or retention of TBP, not its expression.

3.3.3 **Defects in Placement of H3K36me^3 by DNA Entry-Exit Site Mutants Do Not Fully Explain the Termination Defect**

I have shown that many DNA entry-exit site mutants have defects in placing transcription-coupled H3K36me^3 (Chapter 2) in agreement with previous literature (DU AND BRIGGS 2010; ENDO *et al.* 2012). This phenotype does not correlate with termination defect strength, however, and the strongest termination mutants do not have equivalent methylation defects (Figure 14, H3 T45A and H3 R52A). Therefore, I sought to identify genetic relationships between the Set2-mediated H3K36me^3 pathway and the DNA entry-exit site. First, I asked, because the DNA entry-exit site is thought to bind Set2 for its methylation activity (DU AND BRIGGS 2010; ENDO *et al.* 2012), if overexpressing *SET2* in DNA entry-exit site mutants might restore the methylation mark to some extent and partially suppress the observed termination defect.

I measured Set2 levels via an HA-tag in wild type cells compared to the DNA entry-exit site mutants, and these Western blots showed no difference in ability to express this histone methyltransferase from an exogenous plasmid (Figure 30A). While there is currently no method for me to examine endogenous Set2 expression, it is likely that Set2 is present in the mutants to restore H3K36me^3, if possible, based on ability to recover the mark in a *set2*Δ strain expressing the plasmid (compare lanes 9 and 10, Figure 30A). Inability of overexpressed Set2 to restore trimethylation at H3K36 may suggest that the DNA entry-exit site provides an important surface for interaction of the histone methyltransferase, or for its placement of the methylation mark.

In addition to its inability to restore proper histone methylation, overexpression of Set2 did not suppress the termination defects of H3 T45A, R49A, or R52A mutants (Figure 30B). This could be due to lack of Set2 binding to a mutated DNA entry-exit site, or independence of the termination defect observed in these mutants from the H3K36me^3 pathway. It is worth noting that, compared to the DNA entry-exit site mutants, the previously reported termination defect in *set2*Δ is mild (TOMSON *et al.* 2013).

Due to the importance of Set2 in maintaining transcription fidelity through suppression of cryptic transcription (CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005; VENKATESH AND WORKMAN 2013), I also tested DNA entry-exit site mutants in combination with overexpressed Set2 for suppression of this defect. Set2 overexpression in the H3 T45A and R49A mutants is unable to suppress cryptic initiation defects at *STE11* (Figure 30C), in line with its inability to promote increased H3K36me^3 in these strains. It appears, however, that cryptic initiation may be somewhat suppressed in H3 R52A whose defect is already, interestingly, more moderate than H3 R49A despite equivalent loss of H3K36me^3.

Considering this relationship, we investigated whether there is a genetic interaction between H3 residues K36 and R52. With the help of undergraduate Julia Seraly, we constructed a histone shuffle plasmid expressing a double histone mutant where both K36 and R52 are substituted for alanines. With the hypothesis that these two residues perhaps work in related pathways concerning transcription regulation, we tested for synthetic relationships. Intriguingly, histone H3 with alanine substitutions at both residues is lethal to the cell (Figure 31A). To assess whether this synthetic lethality is due to lack of expression from the plasmid, we transformed the H3 K36A, H3 R52A double mutant plasmid into yeast cells with tagged endogenous H3 (*HHT1-HA*) and a deletion of the second H3-H4 locus (*hht2-hhf2*Δ). In this context, cells retained the

ability to express the double mutant H3 allele (Figure 31B). Therefore, the synthetic lethality between K36A and R52A could be due to the additive effects of transcriptional deregulation, for example via cryptic initiation due to the K36A substitution, and rampant terminator readthrough due to the R52A mutant.

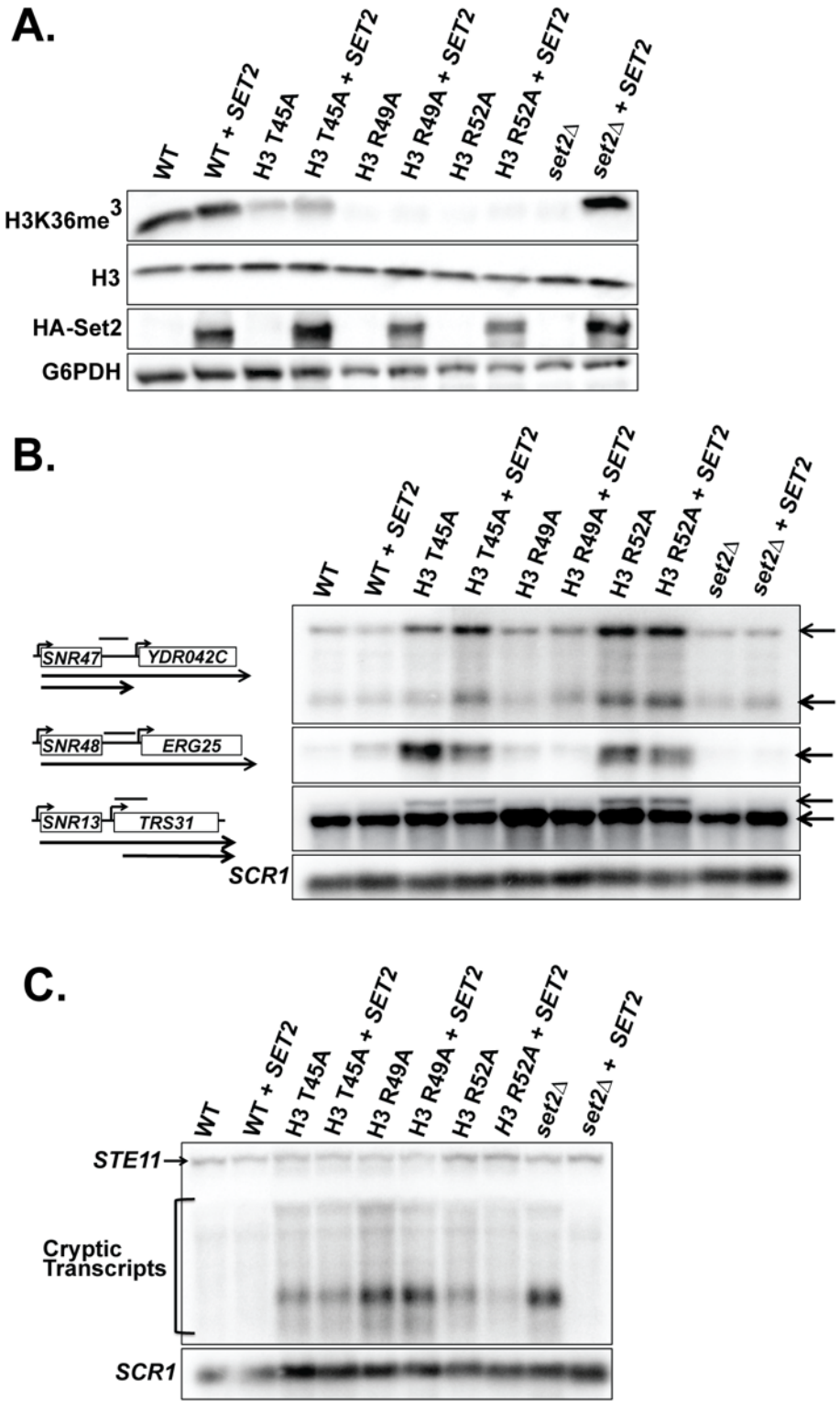

Figure 30. Overexpression of *SET2* in DNA entry-exit site mutants does not restore H3K36me^3 or suppress the termination or cryptic initiation defects of the DNA entry-exit site mutants.

(A) Western blot analysis of overexpression of *SET2* from a high copy, 2μ plasmid over endogenous wild type *SET2*. Blots were probed with antibodies against H3K36me^3, total H3, HA (to detect HA-Set2 from the plasmid), and G6PDH as the loading control. **(B)** Northern blot analysis of termination defects at *SNR47, SNR48,* and *SNR13* in strains from panel A. **(C)** Northern blot analysis of cryptic initiation at *STE11* in strains from panel A.

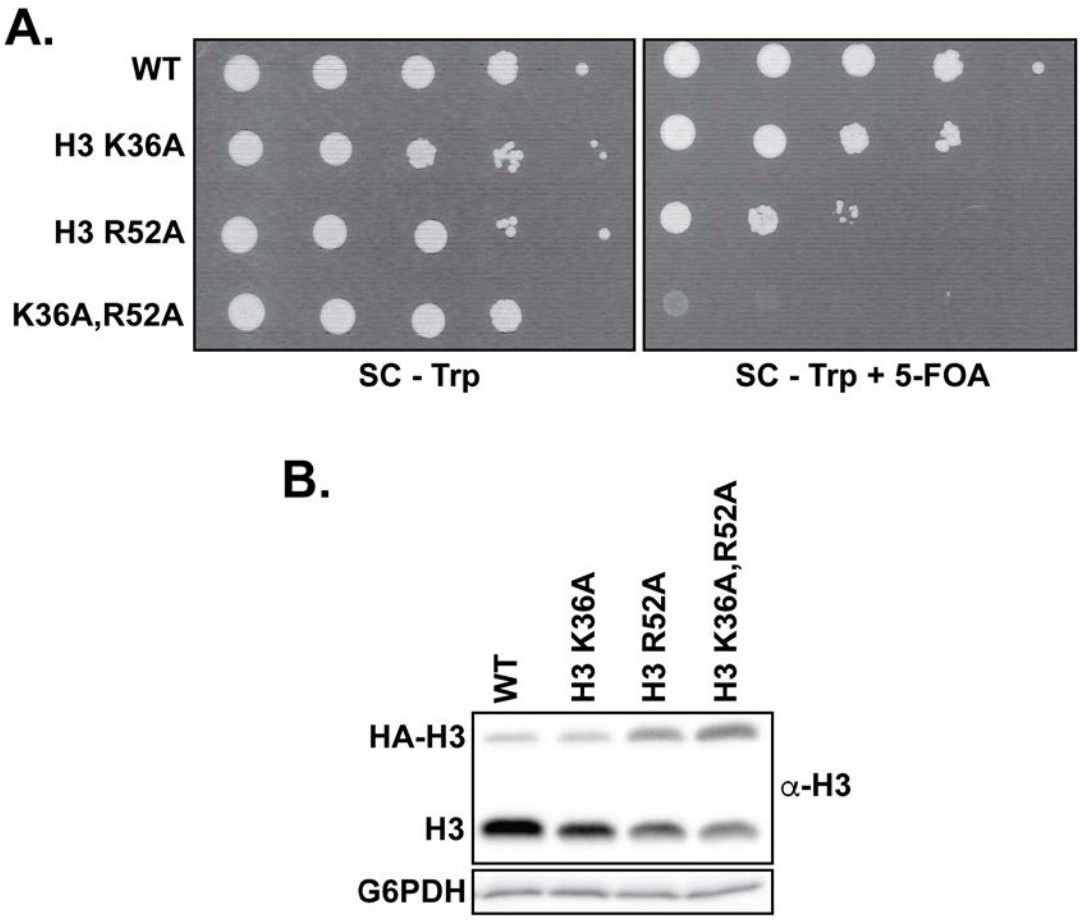

Figure 31. The H3 K36A and R52A substitutions are synthetically lethal.

(A) Strains for yeast dilution assays were constructed by transforming KY812, which contains a wild type, *URA3*-marked H3 expression plasmid, with *TRP1*-marked plasmids encoding wild type H3, single H3 K36A and R52A substitutions, and a double mutant with alanine substitutions at both H3 K36 and R52. Yeast were subjected to a ten-fold serial dilution analysis and plated on SC-Trp with and without 5-FOA.

5-FOA selects against the *URA3*-marked wild type H3 plasmid, leaving the *TRP1* plasmid expressing the indicated allele as the only copy of H3. Yeast expressing the H3 K36A, R52A double mutant as the only source of H3 are not viable. **(B)** Western blot analysis confirming expression of the plasmid-borne histone mutants analyzed in panel A over integrated HA-tagged, wild type H3 from the *HHT1* locus.

3.3.4 **Mutations That Disrupt the DNA Entry-Exit Site Cause Global Changes in Nucleosome Occupancy**

Previous work on the role of the DNA entry-exit site in nucleosome breathing highlights the importance of this surface for regulating nucleosome accessibility during processes like transcription. Work from the Martens lab has associated mutations in the DNA entry-exit site with loss of transcription-coupled nucleosome occupancy, further supporting the role for this nucleosomal surface in restricting access to DNA when necessary (HAINER AND MARTENS 2011a). Led by these previous studies and the preliminary finding that histone levels on chromatin are reduced in DNA entry-exit site mutants (Figures 22, 23), I performed Micrococcal nuclease (MNase) treatment of chromatin coupled to deep sequencing to assess relative positions and occupancy of nucleosomes in the H3 R52A mutant compared to wild type.

First, I treated *S. cerevisiae* chromatin with varying amounts of MNase to assess bulk nucleosome stability, as well as to identify digestion conditions suitable for the sequencing experiment. Based on this preliminary assessment, H3 R52A nucleosomes appear more sensitive to MNase than wild type nucleosomes, as evidenced by reduced levels of di-, tri-, and higher order nucleosomes at comparable treatments (Figure 32).

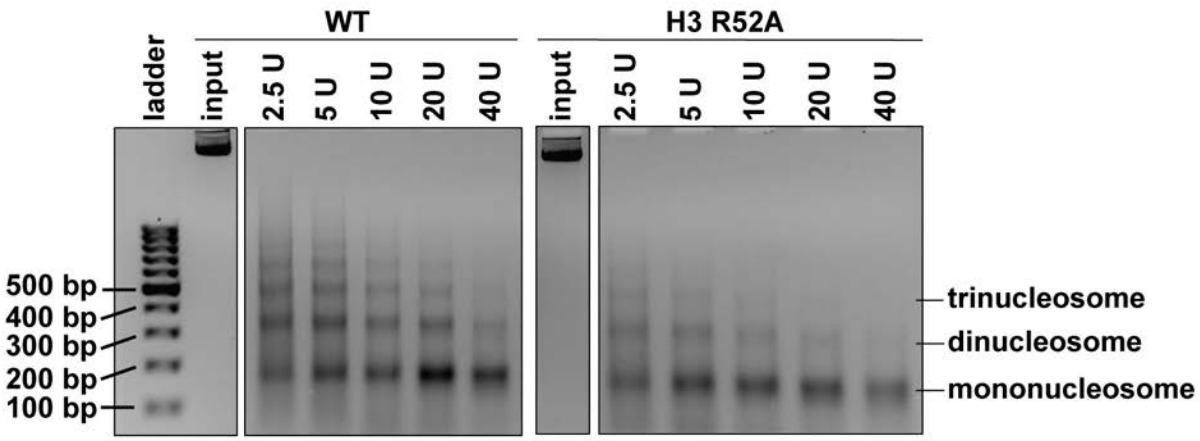

Figure 32. H3 R52A nucleosomes are more sensitive to MNase treatment than wild type nucleosomes. Treatment of total chromatin isolated from cells with wild type nucleosomes or nucleosomes that have incorporated H3 R52A. Chromatin was treated with the indicated amount of MNase and incubated at 37°C for 20 minutes with occasional inversion to mix. Samples were run on a 2% agarose gel for visualization of digestion by ethidium bromide staining.

I proceeded with sequencing of gel-purified mononucleosomes from samples treated with 2.5 U, 20 U, and 40 U MNase that has been spiked with MNase-treated *Kluyveromyces lactis* chromatin. Analysis was performed on the 2.5 U MNase treatments to minimize over-digestion of the mutant and to gain insight on the least chemically perturbed sample available. Nucleosome positions were aligned to the yeast genome, centered on published +1 nucleosome position data (ZHANG *et al.* 2011), and visualized by heatmaps. Differences between wild type and H3 R52A, when viewed this way, appear subtle (Figure 33A). Wild type, as expected, exhibits a strong +1 nucleosome at all genes. Nucleosomes in the H3 R52A mutant have reduced occupancy across all positions and appear to be less well positioned, which is especially visible in the difference heatmap (Figure 33B). This view shows a log2-fold ratio of H3 R52A over wild type signal. Changes in nucleosome position in this view are evident by reduced tightness of +2, +3, etc. bands,

blue signal over nucleosome positions, indicating lower occupancy in the mutant, and increased red signal between nucleosomes indicating that normally internucleosomal regions are more occupied in the mutant. Metagene profiling (Figure 34) depicts this as well, where nucleosomes in the mutants can be seen generally shifting away from the 5’ nucleosome-depleted region (NDR) just upstream of the +1 nucleosome and appear to widen slightly toward the 3’ end of the gene relative to wild type.

Together, these observations suggest that nucleosomes in DNA entry-exit site mutants are indeed altered in their genome-wide occupancy and positioning, possibly contributing to the myriad of phenotypes I have uncovered.

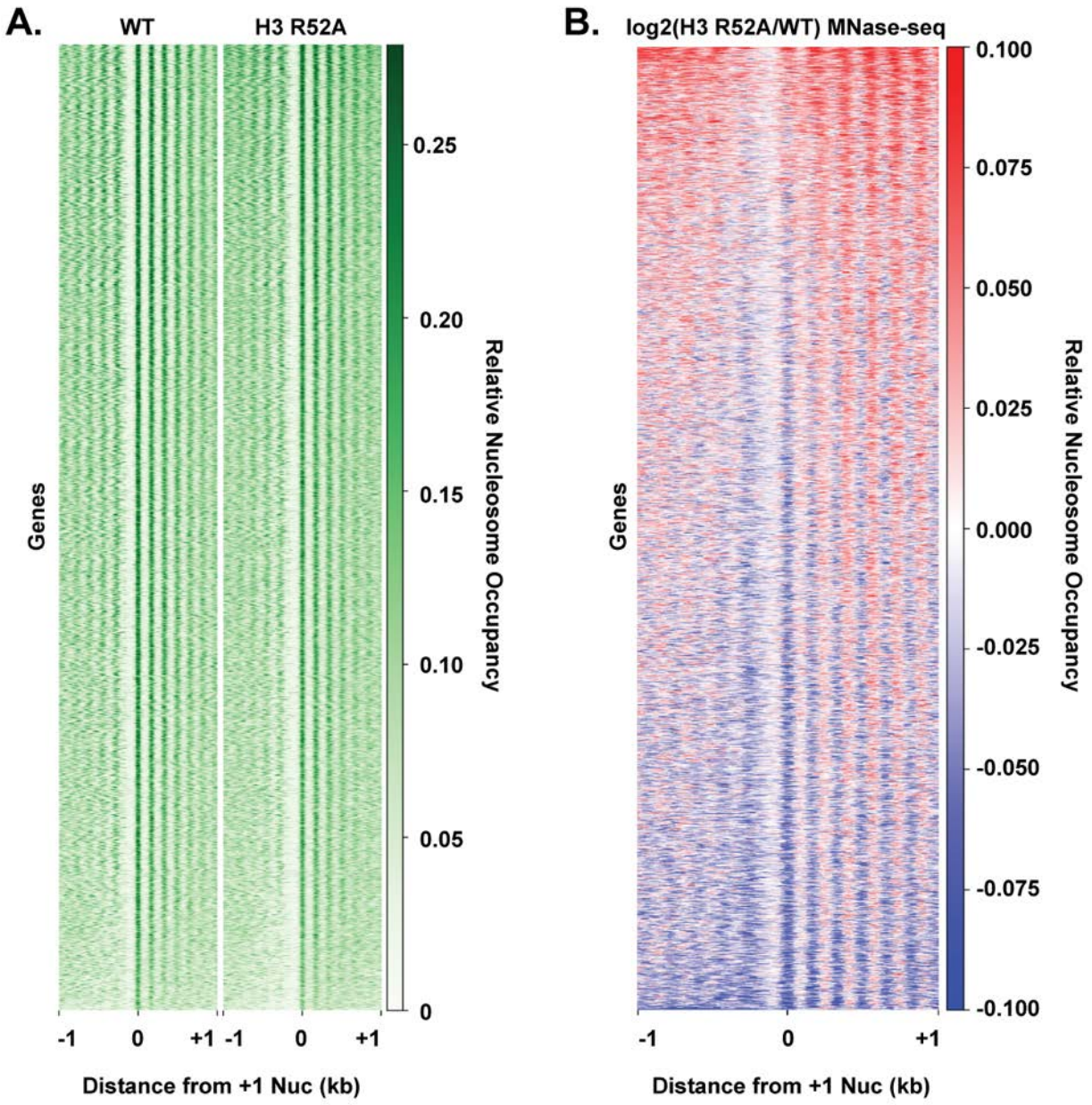

Figure 33. Comparison of nucleosome positioning and occupancy in a wild type strain and the H3 R52A mutant.

(A) Heatmap generated in deepTools2. Each row is a single gene in the yeast genome. Higher green signal means greater nucleosome occupancy, while white signal suggests depletion of nucleosomes. **(B)** Difference heatmap showing the log2-fold ratio of mutant to wild type signal. Blue coloration indicates regions where the mutant has reduced nucleosome occupancy and red suggests that signal is higher in the

mutant. Alternating red and blue signal here shows that nucleosomes are improperly positioned in the H3 R52A strain.

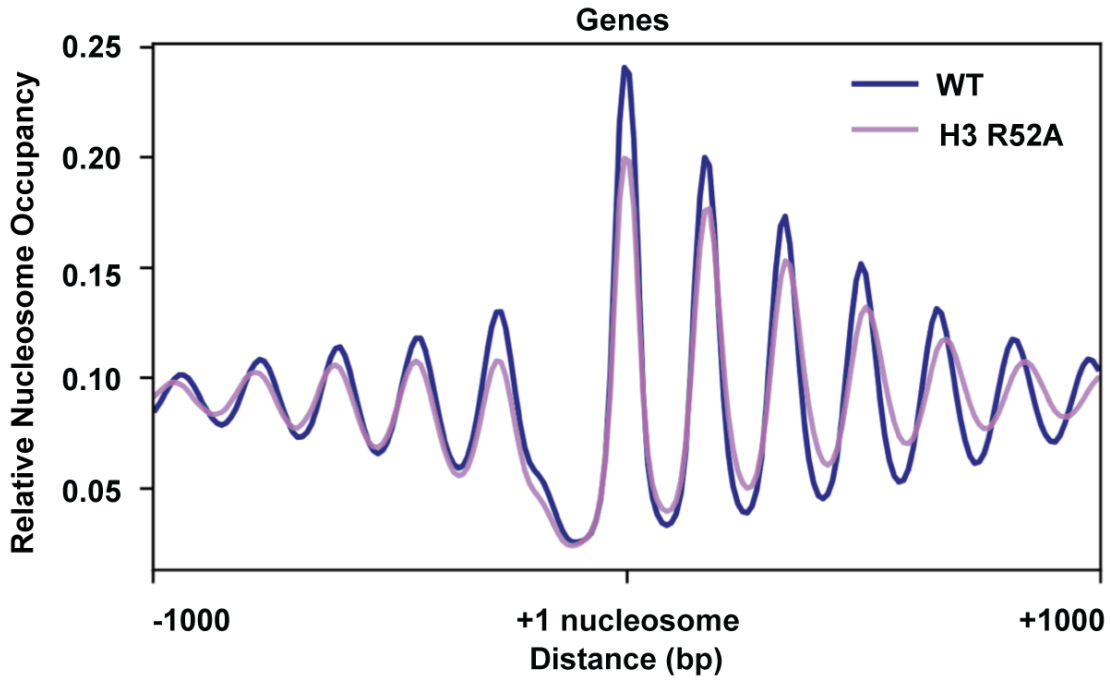

Figure 34. Metagene profiling of H3 R52A nucleosomes compared to wild type reveals reduced occupancy and altered positioning.

Metagene profiling was performed in deepTools2 using the MNase option. Nucleosomes were centered on mapped +1 nucleosome positions (ZHANG *et al.* 2011).

3.3.5 **Comparison of Genome-Wide Mapping Data Reveals Relationships Between Nucleosome Occupancy and Transcription**

To gain a better understanding of nascent transcription versus steady-state RNA levels, and to compare those measurements with nucleosome occupancy obtained from our MNase-seq studies, I performed 4-thiouracil labeling coupled to deep sequencing, as well as Pol II ChIP-seq. 4tU sequencing makes use of a uracil analog to label nascent RNA, which then allows it to be

biotinylated and purified under stringent conditions for sequencing, and Pol II ChIP-seq allows for sequencing specifically of DNA bound by Pol II. Comparison of RNA-seq, 4tU-seq, and MNase-seq reveals that RNA signal is higher where nucleosome occupancy is low, supporting the idea that mutation of the DNA entry-exit site perhaps increases nucleosome accessibility for transcription and termination readthrough (Figure 35).

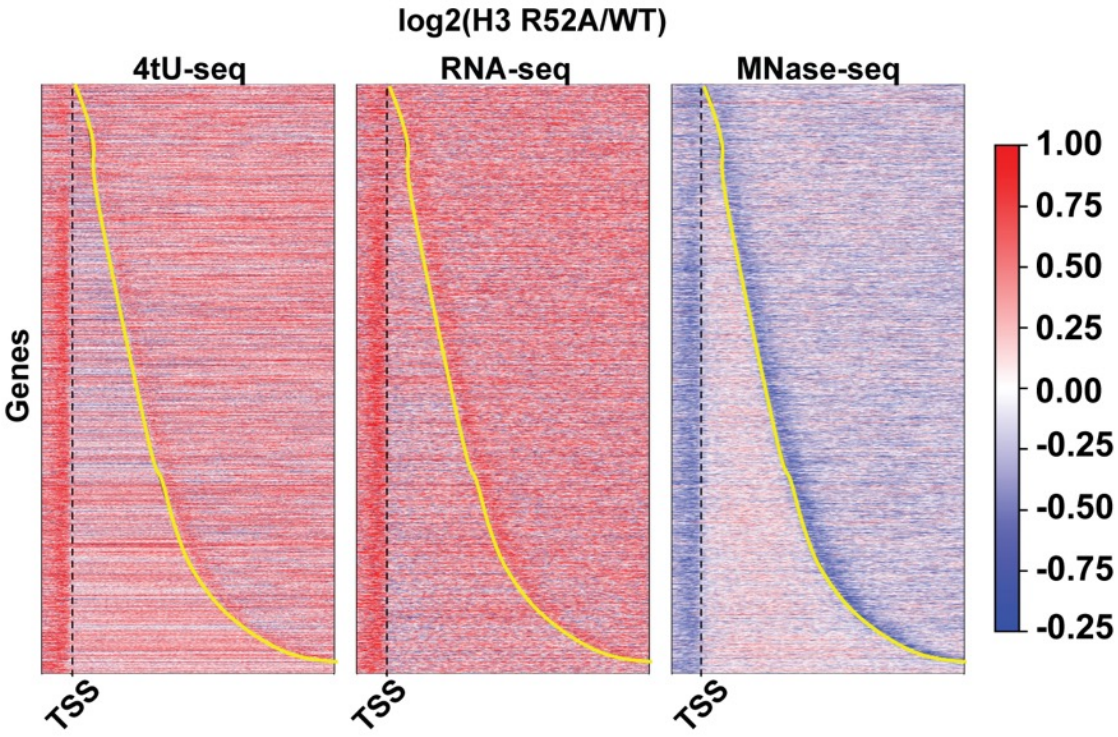

Figure 35. Comparison of RNA levels with nucleosome occupancy suggests RNA is higher where nucleosome occupancy is reduced.

Heatmaps of genes (each row) centered on the TSS and ordered by gene length. The yellow line indicates the 3' end of each gene. In general, regions upstream of the TSS and downstream of the 3' end are enriched for RNA and depleted of nucleosomes, suggesting a relationship between nucleosome occupancy and transcription efficiency.

Analysis of Pol II ChIP-seq data reveals several interesting patterns. When aligned to annotated +1 nucleosome locations, the nucleosomes appear to pose a barrier to the elongating polymerase, especially in the wild type strain. Pol II density increases and decreases periodically with the nucleosomes, suggesting perhaps a slight pause before proceeding through them. This patterning is diminished in H3 R52A, suggesting that this nucleosome barrier may be reduced, as predicted (Figure 36). Perhaps as an extension of this idea, Pol II density is decreased downstream of genes in the H3 R52A mutant compared to wild type (Figure 37). If Figure 36 does indeed suggest lower nucleosome barriers to transcription in the mutant, it is possible that elongation proceeds at a rate that prevents adequate pausing near 3' ends to efficiently terminate Pol II, explaining the reduced Pol II occupancy in this region, despite RNA levels often appearing increased downstream of genes in this mutant (Figure 35). Overall reduction of Pol II occupancy in the H3 R52A mutant is supported by previous work by the Horikoshi lab (ENDO *et al.* 2012).

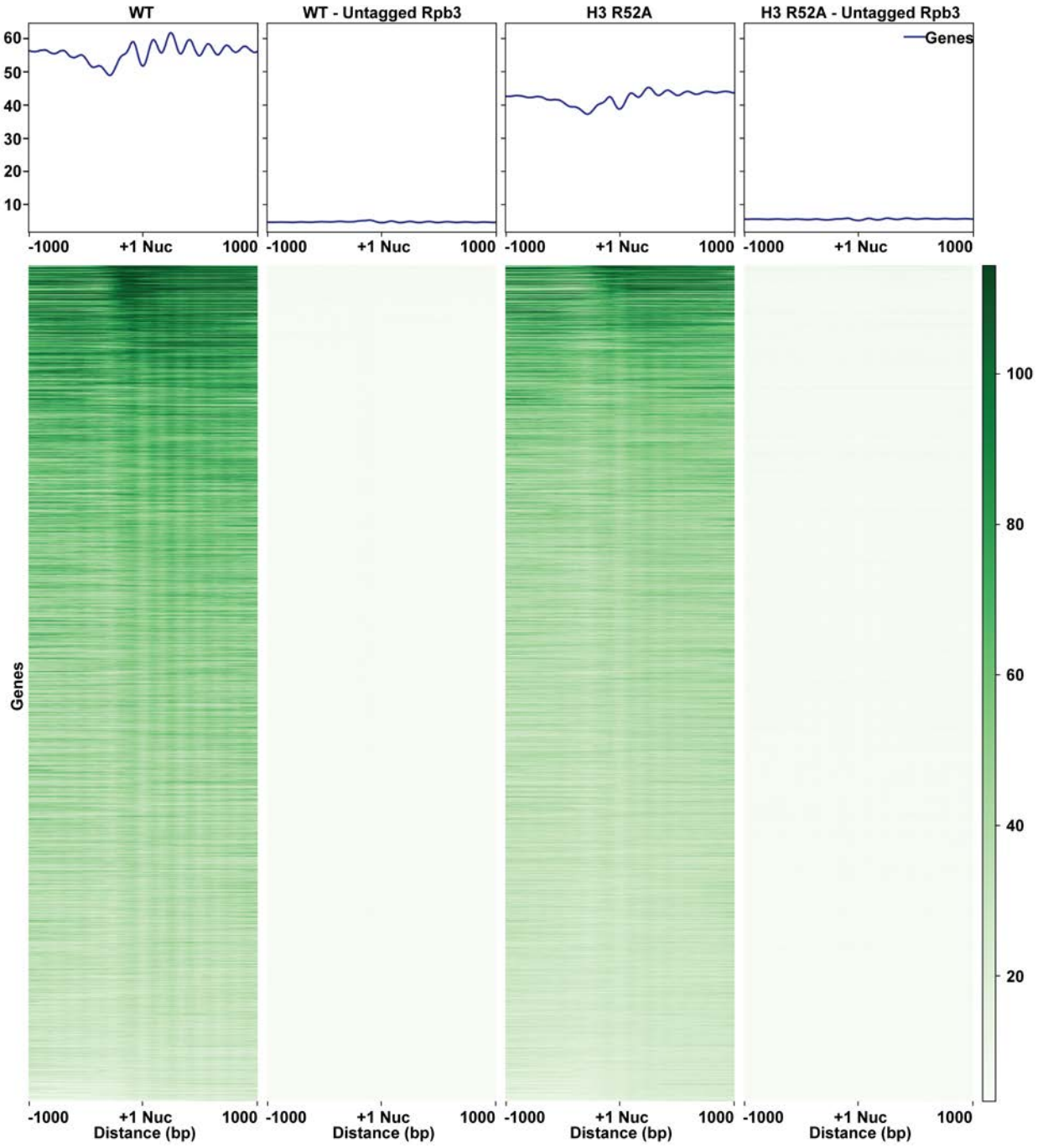

Figure 36. Pol II patterning across genes changes in the H3 R52A mutant.

Heatmaps of genes (each row) centered on published +1 nucleosome maps (ZHANG *et al.* 2011). RNA polymerase II appears periodic across a gene possibly due to the requirement to transcribe through nucleosome barriers. This periodicity is diminished in the H3 R52A mutant, suggesting the barrier is reduced. This periodic Pol II density has been observed in our lab previously (M. Ellison, unpublished).

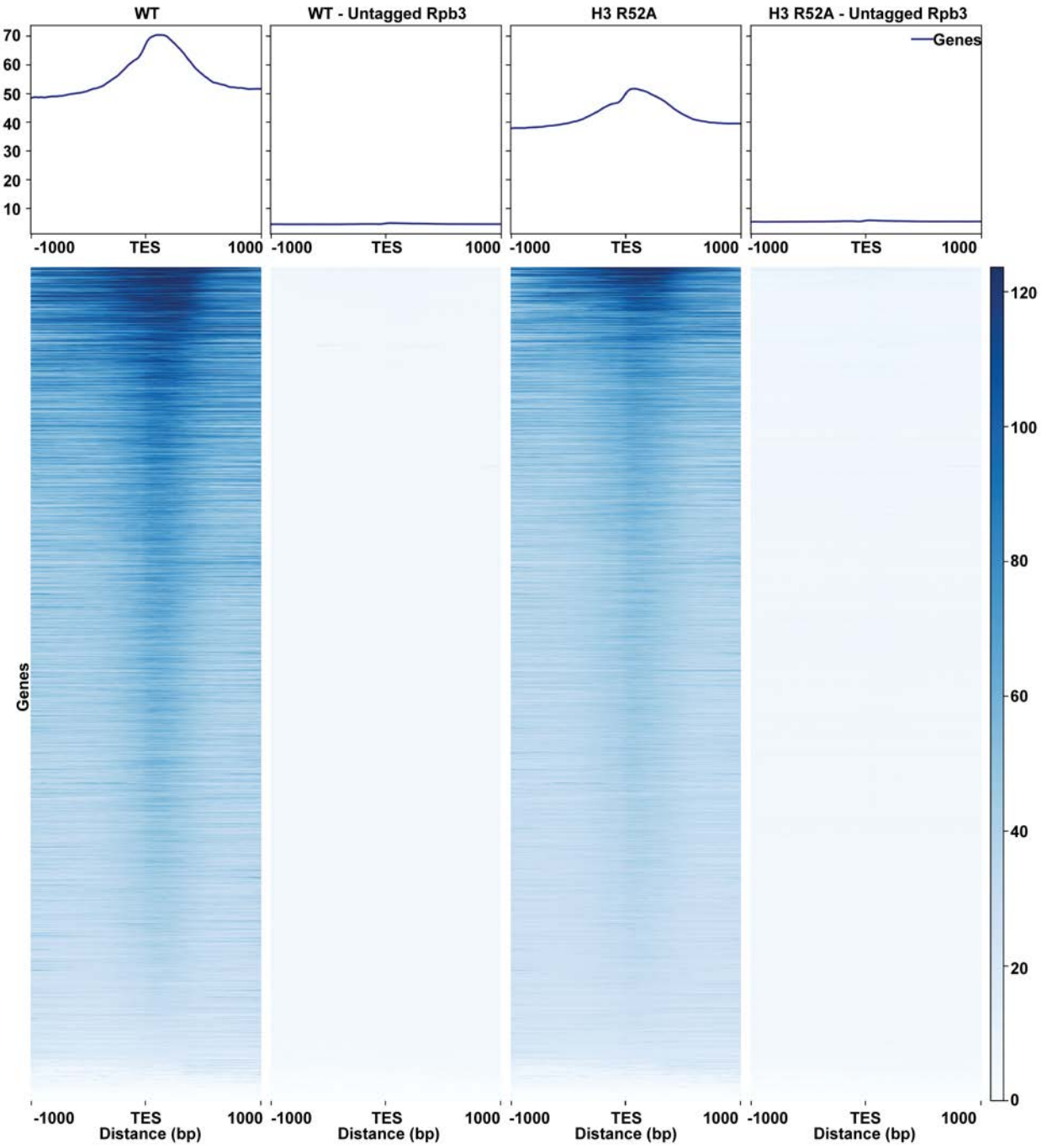

Figure 37. Pol II occupancy decreases downstream of TESes.

Heatmaps of genes (each row) centered on annotated TES sites. RNA polymerase II occupancy is low across genes in general in the H3 R52A mutant and, unlike in wild type, Pol II density does not increase dramatically downstream of transcribed genes. It is possible that this indicates fewer polymerases pausing near the 3' ends of genes, thus likely not properly terminating transcription.

3.3.6 **Stably Positioning a Nucleosome in the *SNR48* Terminator Suppresses the Termination Defect of the H3 R52A Mutant**

Informed by my MNase-seq data, I hypothesized that nucleosome occupancy around terminators is critical for proper termination by Pol II. To test this, I inserted a 133 bp sequence, termed the superbinder sequence, that very stably binds histone proteins (WANG *et al.* 2011) downstream of the *SNR48* termination site as determined by our *de novo* transcriptome assembly data (Figure 38). To monitor nucleosome occupancy at this location, I performed H2A ChIP coupled to qPCR using primers conserved in strains containing and lacking the superbinder sequence (Figure 39A). These data suggest that nucleosome occupancy is increased approximately 3-fold in wild type cells and 2.5-fold in mutant cells (Figure 39B). Consistent with my hypothesis, this increase in nucleosome occupancy downstream of *SNR48* suppresses the termination defect in H3 R52A to the level of a wild type cell lacking the superbinder (Figure 39C). Interestingly, increased nucleosome occupancy at this terminator is also sufficient to suppress basal level detectable transcript observed in the wild type strain.

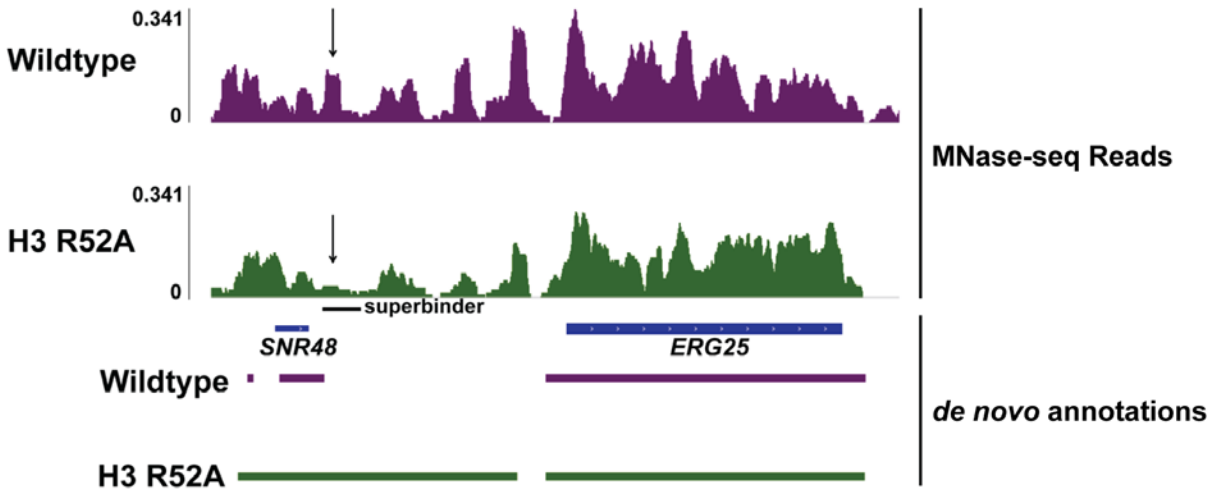

Figure 38. Using MNase-seq coupled to RNA-seq-based *de novo* transcript assembly to determine the best site for integration of the superbinder sequence.

MNase-seq data from a 2.5 U digestion visualized in IGV (top) compared to *de novo* annotations generated from RNA-seq data from the same strains (bottom). Arrows indicate lost nucleosome occupancy in the H3 R52A strain compared to the wild type, which is coincident with the 3' end of the transcript in the wild type strain. The arrow corresponds to the location of integration of the superbinder sequence (black bar under browser shot). For scale, *SNR48* is 113 bp.

I ensured specificity of the readthrough transcript in the context of the superbinder by additionally probing Northern blots for the *SNR48* gene, which can detect the mature *SNR48* transcript and the readthrough transcript (Figure 40). Together, these data suggest that proper transcription termination is regulated in part by nucleosome occupancy and that the DNA entry-exit site of the nucleosome is critical for maintaining the proper chromatin landscape.

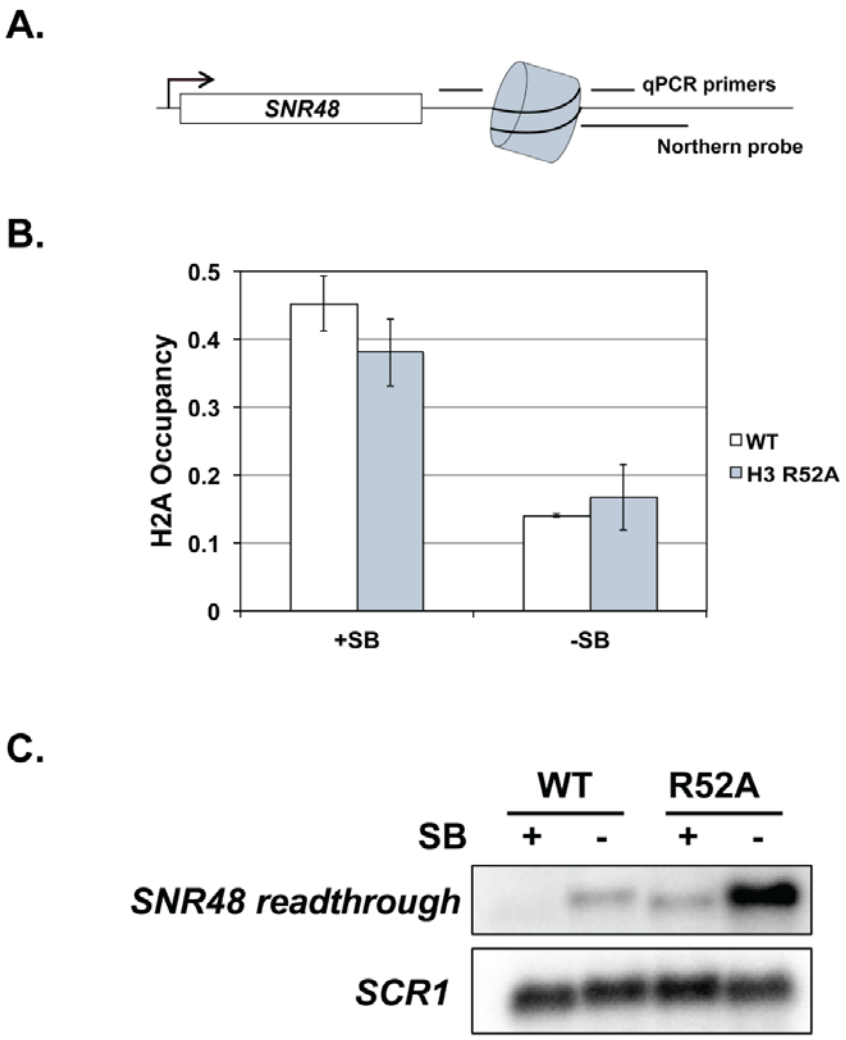

Figure 39. Integration of the superbinder sequence downstream of *SNR48* suppresses the termination defect of H3 R52A.

(A) Cartoon diagram of "superbinder nucleosome" location (blue wheel) and locations of qPCR primers (for ChIP) and the Northern probe (for Northern blot analysis of *SNR48* readthrough). **(B)** ChIP with an antibody against histone H2A. Histone occupancy increases ~2.5-fold in +SB strains. **(C)** Northern blot analysis of *SNR48* readthrough in wild type and H3 R52A strains with or without the superbinder sequence. Integration of the superbinder reduces readthrough of the *SNR48* locus in mutant and wild type strains.

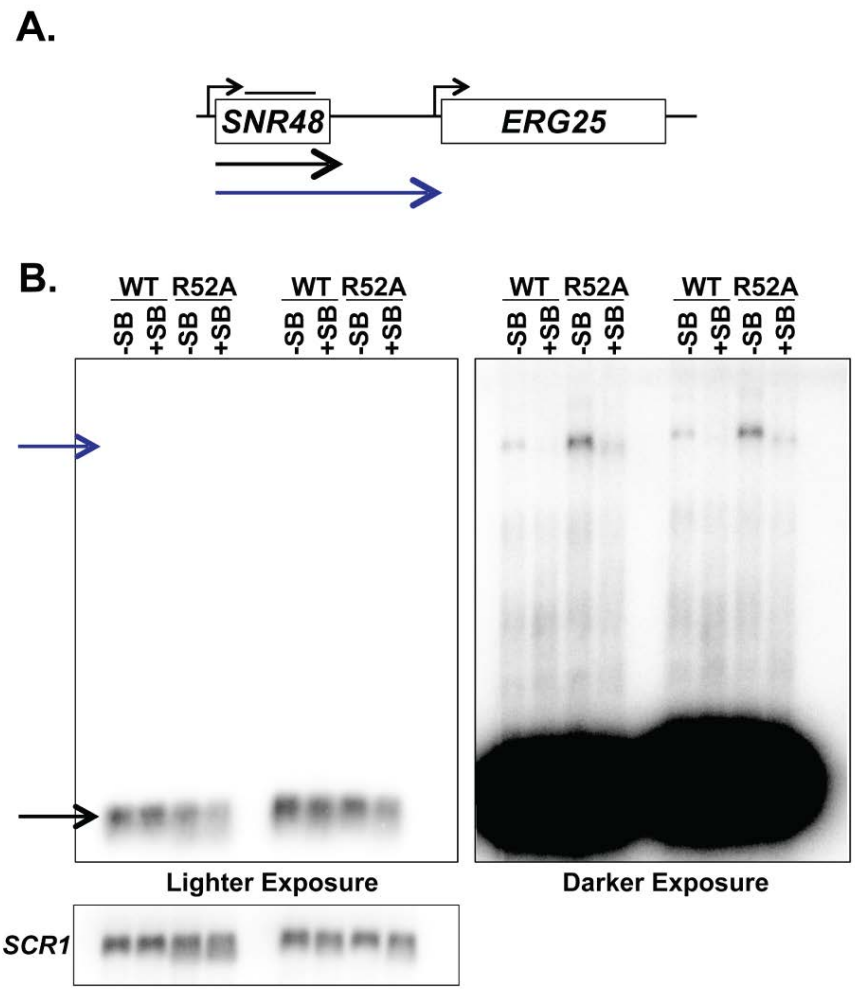

Figure 40. Readthrough of *SNR48* in strains with and without the superbinder sequence are detectable with a Northern probe for *SNR48.*

(A) Location of Northern blot probe over the *SNR48* locus (black bar over *SNR48*) and the observable transcript types. **(B)** Northern blot analysis of transcripts associated with the *SNR48* gene. Lighter exposure reveals only the highly expressed *SNR48* gene (black arrow), while a darker exposure reveals the associated readthrough transcripts (blue arrow). The same pattern of reduction in the strains containing the superbinder, as in Figure 39C, are observed here.

3.4 **DISCUSSION**

In this study, I have shown that maintenance of chromatin architecture is critical for regulating transcription and its termination genome-wide. Based on phenotypic characterization of DNA entry-exit site residues that, when mutated, cause termination defects (Chapter 2), I considered three main hypotheses regarding the role of the entry-exit site in transcription termination and tested these in the two most termination-defective mutants: H3 T45A and H3 R52A. First, work from the Buratowski lab suggests that transcription elongation rate is closely coupled to fidelity of transcription termination (HAZELBAKER *et al.* 2013). Hypothesizing that impairing the DNA entry-exit site reduces nucleosome stability, thus reducing the barrier that nucleosomes in such a strain might pose to Pol II during transcription, I tested transcription rate in H3 T45A and H3 R52A by a well-characterized, ChIP-based assay (MASON AND STRUHL 2003). By this assay, elongation rate in the two mutants tested appears very similar to – if not somewhat slower than – wild type yeast cells. Interestingly, this is in contrast to Pol II ChIP (not shown) and Pol II ChIP-seq (Figures 36, 37) that may suggest altered Pol II kinetics in the H3 R52A strain.

Second, based on previous work from our lab and others (DU AND BRIGGS 2010; ENDO *et al.* 2012; TOMSON *et al.* 2013), I hypothesized that Set2-mediated H3K36me^3 may directly involve the DNA entry-exit site. Interestingly, overexpression of *SET2* in mutant strains does not increase K36me^3 in mutants tested, nor does this suppress their termination defects or cryptic initiation defects to a significant degree. This inability to increase trimethylation of H3 K36 perhaps indicates that mutation of the DNA entry-exit site blocks interaction of Set2 with the nucleosome, or that these mutations have affected another requirement for H3K36me^3. Interestingly, a similar result was published from the Strahl and Mellor labs where *SET2* overexpression was unable to restore trimethylation of H3 K36 in the absence of histone chaperone Spt6 (YOUDELL *et al.* 2008).

It is possible that the relationship between the DNA entry-exit site and placement of this mark depends on Spt6 in a way that has not yet been tested.

Finally, related to the role of the DNA entry-exit site in restricting DNA breathing, I asked whether alanine substitution in the entry-exit site surface could effectively reduce nucleosome stability, or alter nucleosome organization, to the extent that transcription is affected genome wide. I explored this possibility by performing MNase-seq in wild type and H3 R52A strains. Indeed, I observed that global nucleosome occupancy is reduced and nucleosome positioning changes when the DNA entry-exit site is mutated. Interestingly, the position change consists of nucleosome peak centers shifting away from 5' NDRs and more 3' nucleosomes becoming fuzzier, or less well positioned, in the H3 R52A mutant compared to wild type. This shares some similarity in regards to the loss in occupancy and direction of nucleosome shifting with MNase-seq profiles of the *spt6-1004* mutant (DORIS *et al.* 2018), which also exhibits loss of H3K36me^3, increased cryptic initiation, and reduced nucleosome reassembly (CHEUNG *et al.* 2008).

Of note, the reduction in nucleosome occupancy seems apparent not only in the MNase-seq data, but in the Pol II ChIP-seq data as well (Figure 36). Compared to wild type, in which Pol II occupancy across a gene is associated with high, nucleosome-like periodicity, that periodicity is diminished in the H3 R52A mutant. This suggests that Pol II does not accumulate around nucleosomes as in wild type, perhaps demonstrating the relationship between the weakened nucleosome in a DNA entry-exit site mutant and relief of a stringent barrier against Pol II progression, allowing its relatively unperturbed progression. While this aligns well with the work published in Hazelbaker *et al.*, it is not supported by the elongation rate experiment performed here (Figure 28).

To directly test the hypothesis that nucleosomes are specifically required for acting as a physical barrier to RNA polymerase II to promote efficient termination, I integrated a superbinder sequence downstream of *SNR48.* This sequence has been shown previously to increase histone occupancy at the location of choice (WANG *et al.* 2011; HAINER *et al.* 2015), driving formation of a highly stable nucleosome. Integration of this sequence downstream of *SNR48* increases histone occupancy by about 2.5-fold in wild type and H3 R52A strains and effectively suppresses terminator readthrough at this locus in both strains (Figure 39). This supports the hypothesis that a stable nucleosome is a sufficient barrier to Pol II progression that perhaps drives the Pol II pause necessary for termination factors to perform their respective activities.

Together, these data suggest that the relationship between the DNA entry-exit site and regulation of transcription termination involves its intrinsic role in gating nucleosome accessibility and, perhaps, its additional roles in mediating cotranscriptional histone modifications required for altering chromatin structure.

4.0 IDENTIFYING GENETIC INTERACTORS WITH THE DNA ENTRY-EXIT SITE OF THE NUCLEOSOME

Contents of this chapter were performed with the help of Arndt lab undergraduate Lauren Lotka. Further analysis, experimentation, and investigation of candidates from this screen are required and are the goal of Lauren's Summer 2019 Colella Fellowship under the mentorship of Dr. Margaret Shirra.

4.1 **INTRODUCTION**

The DNA entry-exit site of the nucleosome regulates nucleosome accessibility during processes such as transcription, replication, recombination, and DNA repair (Introduction). This surface is composed of the αN terminal helix extension of histone H3, and is stabilized by the docking domain region of histone H2A (LUGER *et al.* 1997). Through nucleosome breathing, where the 13-15 base pairs of linker-adjacent, nucleosomal DNA transiently wraps and unwraps from the histone core surface, DNA becomes accessible to factors like histone chaperones, chromatin remodelers, and histone modifiers, among other proteins, that must invade the nucleosome core to perform their respective activities (LI AND WIDOM 2004; LI *et al.* 2005). An intact DNA entry-exit site effectively restricts access to such factors, including RNA polymerase II, until cellular conditions are appropriate for transcription of particular loci. When this surface of the nucleosome is altered by mutations, however, the nucleosome becomes more accessible and a multitude of cellular

processes are affected (DU AND BRIGGS 2010; HAINER AND MARTENS 2011a; HYLAND *et al.* 2011; ENDO *et al.* 2012).

My work so far suggests that the DNA entry-exit site is required for transcription termination (Chapters 2 and 3). Alanine substitution of residues in H3 near entry-exit DNA causes termination defects genome-wide, and substitutions in H2A cause readthrough defects at least at select snoRNAs (Figures 11, 12, 24, and 25). These mutations also cause a myriad of other transcription-related defects, including deregulation of histone modification marks, phenotypes common to mutants with altered chromatin states, and reduced nucleosome occupancy genome-wide. Previous work from our lab and others has drawn connections between the DNA entry-exit site and known regulators of transcription termination. This includes the requirement for the entry-exit site in Set2-mediated H3K36me^3 (DU AND BRIGGS 2010; ENDO *et al.* 2012), which our lab has shown as a critical pathway for Nrd1-Nab3-Sen1 (NNS)-dependent termination (TOMSON *et al.* 2013). Additionally, many entry-exit residues are required for maintaining transcription-coupled nucleosome occupancy (HAINER AND MARTENS 2011a). Transcription elongation rate, the regulation of which depends on factors such as appropriate chromatin architecture, has been shown to dictate the location of the termination window for NNS-terminated transcripts via kinetic competition between the Sen1 helicase and RNA polymerase II (HAZELBAKER *et al.* 2013).

I have shown that H3K36me^3 and nucleosome occupancy defects in DNA entry-exit site mutants may contribute to their termination defects. However, the mechanisms linking these pathways is not yet clear. Therefore, we sought to identify factors that interact with the DNA entry-exit site of the nucleosome to better understand its roles in histone modification, nucleosome occupancy, and transcription termination. To screen for positive and negative genetic interactions with a mutated DNA entry-exit site (H3 R52A), we employed Saturated Transposon Analysis in

Yeast (SATAY) (MICHEL *et al.* 2017). Adapted from prokaryotic screens for essential genes (CHRISTEN *et al.* 2011), this method makes use of a plasmid-borne *miniDS* transposon and the complementary *Ac* transposase from maize under galactose-inducible expression for driving transposition in *Saccharomyces cerevisiae* (WEIL AND KUNZE 2000; LAZAROW et al. 2012). From large cultures of cells, this method generates millions of transposed clones, thus yielding saturated libraries of transposon insertions to assess cellular responses to given genetic or environmental contexts. Where transposon insertion is not tolerated – suggesting the particular environmental conditions are toxic to cells in the absence of a particular gene product, or if a negative genetic interaction exists with a background mutation of interest – yeast cells are inviable, removing them from the pool of transposed clones. Coupled to genome-wide sequencing, locations of transposition events are uncovered, and instances of tolerated and non-tolerated events are differentiated based on relative sequencing coverage (non-tolerated transposition events will appear as gaps in sequencing data). In this way, a comprehensive, unbiased map of genetic interactions in a desired context is generated for further investigation.

4.2 **MATERIALS AND METHODS**

4.2.1 **Yeast Strains and Media**

All *Saccharomyces cerevisiae* strains used in this study are isogenic to FY2, a *GAL2*$^{+}$ derivative of S288C (WINSTON *et al.* 1995). Performing SATAY analysis requires a complete deletion of *ADE2*, therefore, I crossed KY809 (*MAT* α *(hht1-hhf1)*Δ*::LEU2 (hht2-hhf2)*Δ*::HIS3*) and KY812 (*MAT* **a** *(hht1-hhf1)*Δ*::LEU2 (hht2-hhf2)*Δ*::KanMX*) to generate a histone shuffle diploid. I used one-step gene replacement to swap the *ADE2* gene with the *NatMX* cassette, and then dissected the resulting heterogeneous diploid for haploids with histone shuffle capabilities, as well as the new *ade2*Δ*::NatMX* allele. All experiments were performed with this resulting strain (KY2881) transformed with a CEN/ARS plasmid expressing *hht2R52A-HHF2* or a wild type *HHT2-HHF2* from a *TRP1*-marked plasmid (NAKANISHI *et al.* 2008). Lithium acetate yeast transformations were performed as previously described (BECKER AND LUNDBLAD 2001). *URA3*-marked wild type histone plasmids covering the base strain were shuffled with *TRP1*-marked mutant histone plasmids by sequentially passaging transformants three times on SC-Trp medium containing 2% dextrose and 0.1% 5-FOA.

4.2.2 **Screen for Synthetic Lethal Interactions with Saturated Transposon Analysis in Yeast (SATAY)**

4.2.2.1 **Transposon Library Generation**

The SATAY protocol was adapted from (MICHEL *et al.* 2017). KY2881 was transformed with pBK549 (Figure 41) and selected on SD-Ura. Single transformants were struck on SD-Ura and

SD-Ade to obtain Ura+ Ade- single colonies, ensuring no spurious recombination events had occurred (repairing the *ADE2* gene within the plasmid and allowing for premature growth on media lacking adenine). A patch of cells derived from a single colony was used to inoculate 150-200 mL of SD-Ura + 0.2% glucose + 2% raffinose, which was incubated at 30°C until saturation (OD_{600} = ~3.5). A small aliquot of the culture was diluted 40,000X and plated on control plates to count instances of overall plasmid loss (comparing SD-Complete and SD-Ura plates) and uninduced transposition (SD-Ade + 2% glucose). Cells were pelleted, washed two times with sterile water, and resuspended to an OD_{600} = 39. 200 µL of this concentrated culture was spread on approximately 100 SD-Ade + 2% galactose plates to induce expression of the transposase encoded on pBK549. Once dried, plates were restored to their original sleeves in stacks of approximately 12-15, loosely sealed to prevent drying, and incubated at 30°C for 2-3 weeks.

To determine length of incubation time, plates were visualized by microscopy and counted once every several days. During this time, plates were inspected for contamination and discarded if necessary. To count, a 1 cm^2 box was drawn on 3-5 representative plates per strain background. Dense colonies of cells in this region were counted at 160X magnification for each plate, and the number of cells counted on each plate was averaged. This average count was multiplied by 55 cm^2, to estimate the number of cells per plate, and recorded for each count. Cells were harvested when the number of colonies per plate changed by less than 5% between counts.

Transposon libraries were harvested by pipetting 2 mL of sterile water to each plate and using a sterilized glass rake to scrape and resuspend colonies. All resuspended colonies were collected in a 100 mL flask and diluted 1:100 for OD_{600} measurement. The cell suspension was used to inoculate 2 L of prewarmed SD-Ade + 2% glucose to OD_{600} = 0.2. The culture was allowed to grow to saturation, and then harvested by pelleting. Cell pellets were washed once with sterile

water, supernatant was decanted, and cells were resuspended in remaining liquid. The concentrated cells were aliquoted at 500 µL per microcentrifuge tube, spun briefly to remove excess supernatant, and flash frozen for storage.

4.2.2.2 **DNA Library Generation**

Cell pellets were thawed on ice and resuspended in 500 µL Cell Breaking Buffer (2% Triton-X, 1% SDS, 100 mM NaCl, 100 mM Tris, and 1 mM EDTA). Resuspended pellets were aliquoted into microcentrifuge tubes in 250 µL aliquots. 200 µL phenol:chloroform:isoamyl alcohol (25:24:1) and 300 µL glass beads were added to each sample. Samples were vortexed for 10 minutes (30 seconds vortex and 30 seconds rest on ice). After lysis, 200 µL TE (10 mM Tris and 1 mM EDTA) was added to each sample. Samples were centrifuged for 5 minutes at 14,000 rpm at 4°C. The aqueous phase (~400 µL) was transferred to a fresh tube and 1 mL cold 100% ethanol was added. Samples were inverted several times to mix and incubated at -80°C for one hour. After precipitation, samples were pelleted for 15 minutes at 14,000 rpm at 4°C. Supernatant was decanted and pellets were washed with 1 mL cold 70% ethanol. Pellets were spun once more at 14,000 rpm for 5 minutes at room temperature. DNA pellets were dried completely, resuspended in 200 µL RNase A (250 µg/mL in TE), and incubated at 37°C for 30 minutes. Upon completion of digestion, 500 µL of 100% ethanol and 20 µL sodium acetate (3 M, pH 5.2) was added to each sample. Again, samples were incubated at -80°C for one hour. After precipitation, samples were pelleted for 15 minutes at 14,000 rpm at 4°C. Supernatant was decanted and pellets were washed with 1 mL cold 70% ethanol. Pellets were spun once more at 14,000 rpm for 5 minutes at room temperature. Pellets were dried completely, resuspended in 100 µL nuclease free water, and incubated at 65°C for 10 minutes to ensure complete resuspension. For each sample, 5 µL of a 10X

dilution was run on a 0.6% agarose gel against titration of a DNA sample of known concentration to estimate DNA concentration for each sample.

2 µg genomic DNA from each sample was digested in parallel with 50 units of *DpnII* and 50 units of *NlaIII* in 50 µL for 16 hours at 37°C. Restriction enzymes were heat-inactivated at 65°C for 20 minutes. For each sample, 3 µL of the resulting digests were run on a 2% agarose gel where a smear whose center is approximately 500 bp is observed. Digested fragments were self-circularized with 25 Weiss units of T4 DNA Ligase (NEB, Ipswitch, MA) for 6 hours at 22°C in a volume of 400 µL. Samples were precipitated with 40 µL sodium acetate (3 M, pH 5.2), 1 mL cold 100% ethanol, and 1 µL glycogen (20 mg/mL) overnight at -80°C. DNA was pelleted at 4°C by spinning 14,000 rpm for 20 minutes. Pellets were washed with 1 mL 70% ethanol and centrifuged again at 4°C for 10 minutes. Pellets were allowed to dry completely, and then were resuspended in 50 µL nuclease free water. Samples were divided into 10 µL aliquots in PCR tubes and amplified using primers containing Illumina P5 and P7 adaptor sequences. Each individual sample (for instance, WT *DpnII*, WT *NlaIII*, mutant *DpnII*, mutant *NlaIII*) was amplified with a separate P7-containing oligonucleotide with a unique barcoding index sequence to allow for multiplexing several samples on a single flow cell (Table 8). Resulting PCR reactions were pooled for each sample, purified, and run on a gel to verify library builds. Both digestion schemes yield smears from approximately 250 bp – 1.2 kb, with the highest density signal around 500 bp. For *DpnII* digests, a band corresponding to cut pBK549 appears at 465 bp. In *NlaII*-digested samples, this band appears at 867 bp. Additional bands may be visible within the smear.

250 – 1200 bp smears were gel purified to exclude adaptor-dimers from sequencing. Samples were pooled in equal concentrations and submitted to UPMC Children's Hospital

Sequencing Core for sequencing on the Illumina NextSeq 500. Due to the nature of the library build, samples were submitted with custom sequencing and indexing primers (Table 8).

4.2.2.3 **Data Analysis**

Demultiplexed FASTQ files were obtained from the UPMC Children's Hospital Sequencing Core and were first trimmed of non-genomic, transposon-associated reads using CutAdapt (MARTIN 2010). The -m 20 option was used to output only reads greater than 20 bp to avoid non-unique mapping to the *S. cerevisiae* genome. The trimmed output sequences were aligned to the sacCer3 (UCSC) yeast genome with HISAT2 (KIM *et al.* 2015) and low quality reads were filtered using the SAMtools suite (LI *et al.* 2009). The deepTools2 suite (RAMIREZ *et al.* 2016) was used to normalize each library by counts per million (CPM) and generate bigWig files for viewing in Integrative Genomics Viewer (IGV 2.5.x) (THORVALDSDOTTIR *et al.* 2013). Genome browser tracks show combined *DpnII* and *NlaIII* data sets for a given biological sample. Sequencing reads were counted over genes using featureCounts (LIAO *et al.* 2014). To estimate fold-enrichment of transposon insertion in the mutant compared to wild type, we averaged the three H3 R52A replicates and divided read counts from wild type to read counts across the three replicates of the mutant. To prevent division errors, 1 read was added to each count prior to calculating the enrichment score.

Table 7. Yeast strains used in Chapter 4.

Strain	MAT	Genotype
KY2881	**a**	*(hht1-hhf1)*Δ*::LEU2 (hht2-hhf2)*Δ*::KanMX his3*Δ*200 lys2-128∂ leu2*Δ*1 ura3-52 trp1*Δ*63 ade2*Δ*::NatMX* [pDM1 = *URA3/CEN/ARS/HHT2-HHF2*]

Table 8. Oligonucleotides used in Chapter 4.

Target	Sequence
P5_*miniDS*	5’-AATGATACGGCGACCACCGAGATCTACtccgtcccgcaagttaaata-3’
*miniDS*_P7_Index1	5’-CAAGCAGAAGACGGCATACGAGAT**CGAGTAAT**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index2	5’-CAAGCAGAAGACGGCATACGAGAT**TCTCCGGA**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index3	5’-CAAGCAGAAGACGGCATACGAGAT**AATGAGCG**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index4	5’-CAAGCAGAAGACGGCATACGAGAT**GGAATCTC**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index5	5’-CAAGCAGAAGACGGCATACGAGAT**TTCTGAAT**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index6	5’-CAAGCAGAAGACGGCATACGAGAT**ACGAATTC**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index7	5’-CAAGCAGAAGACGGCATACGAGAT**AGCTTCAG**acgaaaacgaacgggataaa-3’
*miniDS*_P7_Index8	5’-CAAGCAGAAGACGGCATACGAGAT**GCGCATTA**acgaaaacgaacgggataaa-3’
688_*miniDs*SEQ1210	5’-tttaccgaccgttaccgaccgttttcatcccta-3’
Custom_IndexSEQ	5’-GGTTTTCGATTACCGTATTTATCCCGTTCGTTTTCGT-3’

Uppercase text: P5/P7 sequence; lowercase text: *miniDS* sequence; bold text: barcoding index

4.3 **RESULTS**

4.3.1 **Pilot of SATAY Library Generation in Strains Expressing Wild Type and Mutant DNA Entry-Exit Site H3 Alleles on Plasmids**

The original SATAY screen is technically very simple. The strain of interest is transformed with a plasmid encoding a transposase/transposon (*Ac/Ds)* system that, upon plating on media containing galactose, inserts transposons into the yeast genome (WEIL AND KUNZE 2000; LAZAROW et al. 2012). Selection on SD-Ade is driven by excision of the *miniDS* sequence from the plasmid-borne *ADE2* gene. Resulting double strand break repair of *ADE2* restores a functional copy of this gene to a cell with a deleted chromosomal copy. However, this repair system limits the number of recovered transposon clones to those cells successful in repairing *ADE2.* After initial publication of the SATAY protocol (MICHEL *et al.* 2017), the Kornmann lab engineered a modified, higher efficiency *Ac/Ds* transposition plasmid (Figure 41). pBK549 maintains all of the features of the original plasmid, pBK257, except that the halves of *ADE2* on either side of the *miniDs* transposon include sequence to direct homologous recombination of the *ADE2* gene upon excision of the transposon. This relatively error-free repair mechanism drives higher efficiency of recovery of a wild type *ADE2* locus and, therefore, far more selectable transposon clones.

Using the new SATAY system, we generated about 9.36×10^5 clones in a strain with deleted for the chromosomal genes for H3 and H4 (*hht1-hhf1*Δ*, hht2-hhf2*Δ) and expressing H3 R52A from a plasmid. We also generated transposon libraries with approximately 1.31×10^6 clones from a matched strain expressing wild type H3 from a plasmid to account for background effects of plasmid-based histone expression. Notably, the modified SATAY protocol, using strains derived from BY4741, suggests that approximately 2 million clones can be isolated on 20-40 SD-Ade

plates. Numbers of clones reported here are the result of approximately 100 plates, suggesting that FY2 derivatives, especially those harboring plasmids, may have reduced transposition efficiency in comparison to BY4741. We also found from this preliminary screen that additional plasmids in SATAY query strains, even low-copy plasmids, act as transposon sinks that reduce chromosomal transposon reads (Figure 42). This is especially the case in the presence of the wild type H3 expression plasmids that may be able to tolerate transposition better than plasmids expressing the H3 R52A mutant allele (Figure 42B).

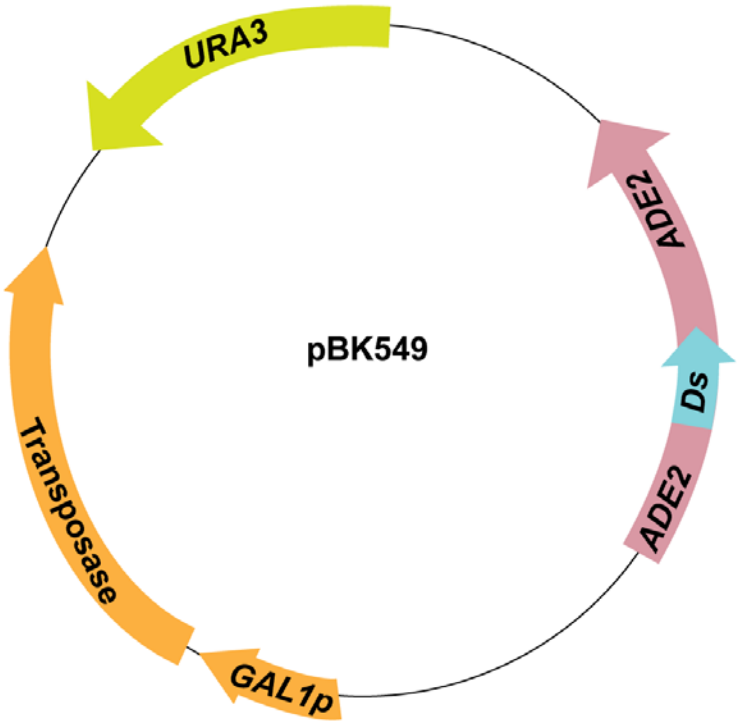

Figure 41. Simple diagram of SATAY plasmid (pBK549).

URA3 acts as the selectable marker for transformation of pBK549. The *AC* transposase is driven by the *GAL1* promoter, so high levels of transposition is induced by growth on galactose. The *ADE2* gene is interrupted by the *miniDS* maize transposon. Excision of the transposon results in repair of *ADE2* and ability to grow on SD-Ade, the selection media for transposon clones.

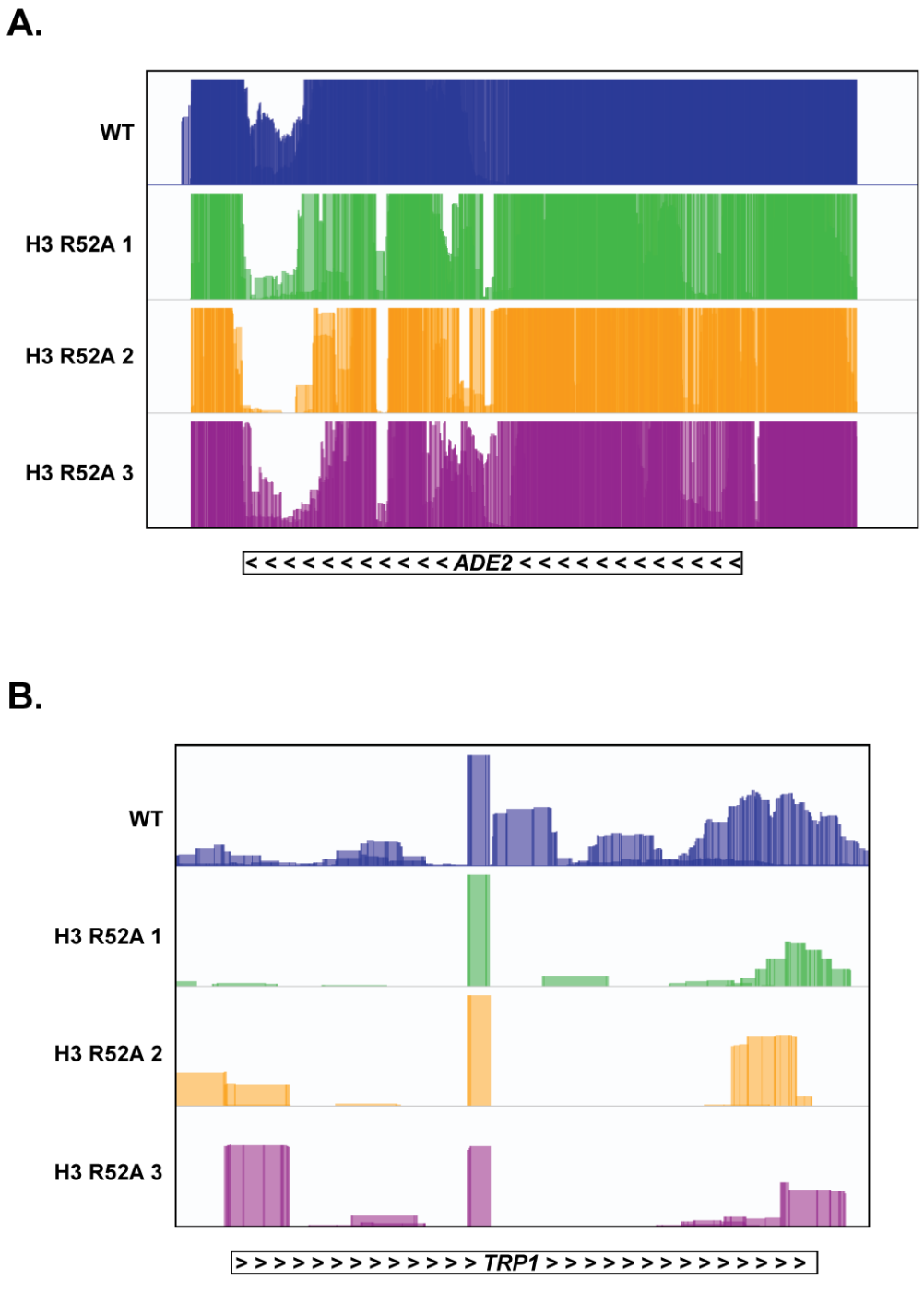

Figure 42. pBK549 and additional plasmids are targets for transposon insertion.

(A) Browser window of the *ADE2* gene, which is encoded on the pBK549 SATAY plasmid. In all conditions, thousands of reads align to this highly abundant plasmid-borne sequence in SATAY sequencing

libraries. This reduces effective signal elsewhere in the genome. Data from the Kornmann lab also shows a higher density of reads over SATAY-associated genes, such as *ADE2*, than surrounding genes (MICHEL *et al.* 2017). **(B)** To a lesser extent, additional plasmids in SATAY query strains reduce signal. Especially for the wild type H3 plasmid, shown here as blue tracks, the *TRP1* marker shows high sequencing coverage.

4.3.2 **Identification of Potential Positive and Negative Genetic Interactors with the DNA Entry-Exit Site**

Analysis of our SATAY pilot screen with the H3 R52A DNA entry-exit site mutant reveals several interesting paths to follow genetically. Subjecting the top 10% of potential synthetic lethal interactors (enrichment of over 10-fold) to Gene Ontology (GO) function analysis with a p-value cutoff of $p \leq 0.01$ results in a single category, "Unannotated," making up 101 of 572 (17.7%) of these genes (p-value = 3.89e-05). Interestingly, several of these dubious ORFs overlap genes of known function in processes related to the DNA entry-exit site (Table 9, Figure 43A), perhaps suggesting the synthetic relationship is due to affecting the expression of the known gene. Additionally, perhaps mistakenly, *CTK3*, a subunit of the CTD Kinase-1 (CTDK-1) complex, is included in this list of "unannotated" ORFs.

Table 9. Unannotated ORFs overlap genes relevant to DNA entry-exit site function.

Dubious ORF	Overlapped Gene	ORF Function	Gene Function	Reads WT	Reads R52A	Ratio
YNL235C	*SIN4*	Unknown	Mediator Complex	49.78	1.15	43
CTK3/YML112W	N/A	CTDK-1	N/A	40.41	1	40
YDL071C	*BDF2*	Unknown	Transcription (TFIID)	117.34	3.96	30
YGL042C	*DST1*	Unknown	TFIIS	25.80	1	26
VPS65/YLR322W	*SFH1*	Unknown	RSC Complex	22.66	1	23
YDR290W	RTT103	Unknown	Pol II Termination	23.38	1.34	17

From the top 10% of potential synthetic lethal interactors, I identified several genes of known function related to transcription and RNA processing (Table 10, Figure 43B). Of interest, at least two members each of the COMPASS, NuA4, and SAGA (Figure 43C) complexes are represented in this list, suggesting that these complexes potentially play some role in regulating transcriptional activity related to the function of the DNA entry-exit site. The COMPASS complex, of which the catalytically active subunit is Set1, is responsible for cotranscriptionally methylating histone H3 at lysine 4 (BRIGGS *et al.* 2001). NuA4, whose Esa1 subunit is the only essential histone acetyltransferase in the yeast genome, is the primary HAT for histone H4 (ALLARD *et al.* 1999; GINSBURG *et al.* 2009). SAGA has many transcription-related roles, including interaction with TBP, for which both Spt3 and Spt8 are required, and histone acetylation and deubiquitylation (HELMLINGER AND TORA 2017).

Table 10. General transcription-related genes in the top 10% of transposon enrichment in WT over mutant (synthetic lethality).

Gene	Complex, Function	Reads WT	Reads R52A	Ratio
HTA1	Histone H2A	154.99	1.78	87
BRE2	COMPASS, H3K4me	140.06	1.62	87
EAF7	NuA4 HAT	97.22	1.61	61
CTK3	CTDK-1	40.41	1	40
THO1	nuclear RNA binding	47.92	2	24
SDC1	COMPASS, H3K4me	18.39	1.03	18
SPT8	SAGA Complex	700.64	41.08	17
YNG2	NuA4 HAT	183.29	13	14
PAF1	Paf1 Complex	55.75	4.01	14
EAF1	NuA4 HAT	58.15	4.18	14
SPT3	SAGA Complex	238.29	18.27	13

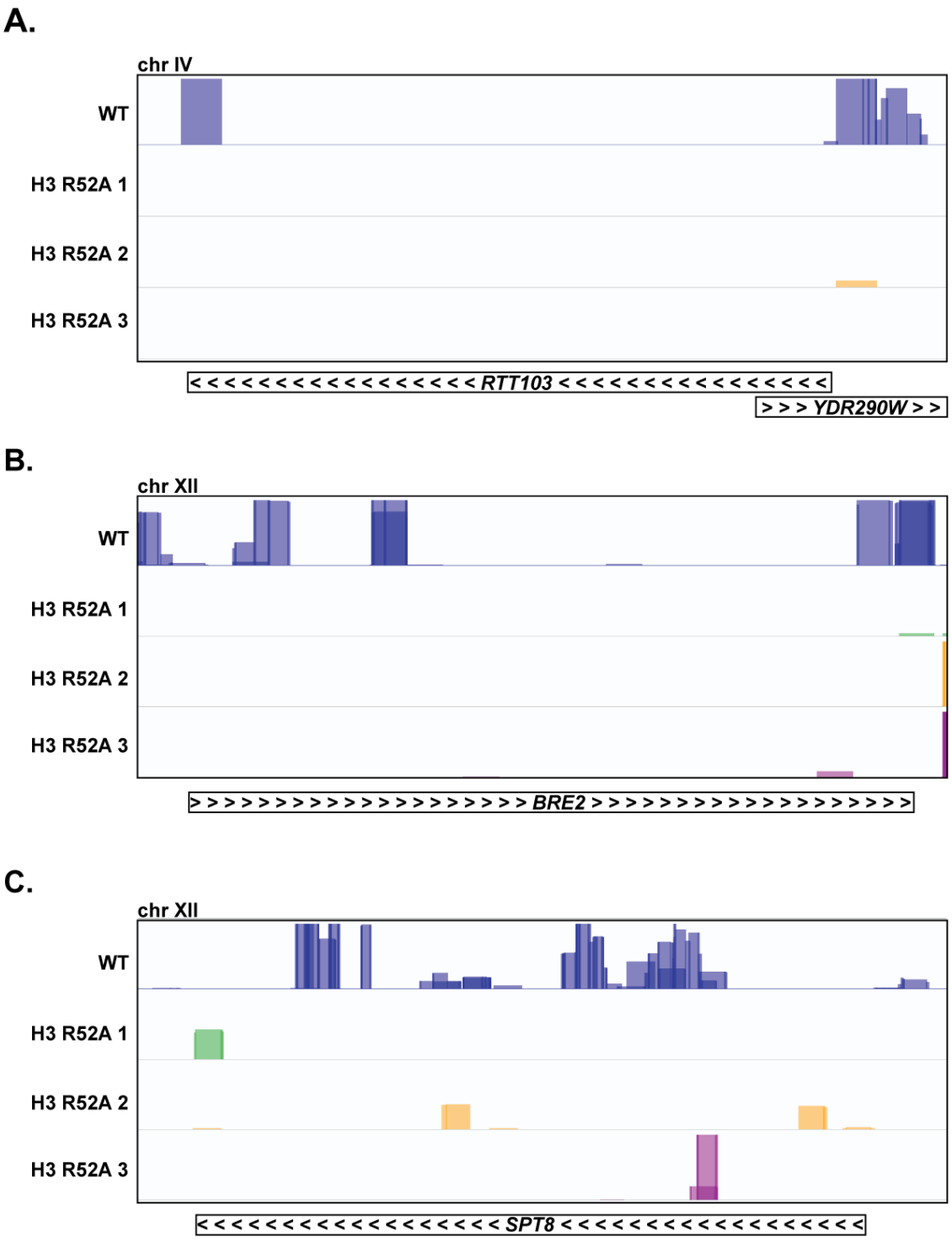

Figure 43. Browser views of unannotated ORFs and genes involved in transcription that are synthetically lethal with H3 R52A.

Gene browser views of **(A)** the *YDR290W* dubious ORF that overlaps the promoter and 5' end of *RTT103*, **(B)** COMPASS subunit *BRE2*, and **(C)** SAGA subunit *SPT8*. For all panels, transposon reads are enriched in wild type over replicates of the mutant.

I additionally sought transposon insertions that suppress, or confer better growth, on H3 R52A mutants. Subjecting the 10% of genes with the lowest enrichment scores (wild type relative to mutant reads) to GO analysis returned 52 ribosome-associated genes (9.1%, p-value = 2.05e-09) and 64 genes (11.2%, p-value = 3.94e-07) in the "structural molecule activity" category, which includes ribosomal proteins, mitochondrial proteins, and nuclear pore complex proteins. I have shown that H3 R52A causes widespread transcription deregulation, so it is possible that mutations to the ribosome helps the cell to compensate for altered transcript levels.

Among the top 10% of potential suppressors are also several genes related to known chromatin functions, including genes encoding members of the Paf1 complex, Rad6 (the E2 for ubiquitylation of H2BK123), and members of the SAGA complex, among others (Table 11, Figure 44). Interestingly, SAGA subunits arise in both the synthetic lethal and suppressor groups. Synthetic lethal genes, mentioned above, are *SPT3* and *SPT8*, which encode SAGA subunits thought to interact with TBP (SERMWITTAYAWONG AND TAN 2006; LAPRADE *et al.* 2007). Suppressor interactions arise from transposon insertions in *SPT20*, which encodes a structural core subunit of SAGA and related to histone acetyltransferase function (MARCUS *et al.* 1996; ROBERTS AND WINSTON 1996), and *ADA2*, which encodes a subunit of the HAT module (CANDAU *et al.* 1997; STERNER *et al.* 2002; SUN *et al.* 2018).

Table 11. Genes that potentially suppress the H3 R52A mutant, as identified by higher transposition levels.

Gene	Complex, Function	Reads WT	Reads R52A	Ratio
SWD3	COMPASS, H3K4me	1.38	334.45	0.004
BUR2	Bur1-Bur2 Kinase	1.41	72.11	0.020
CDC73	Paf1 Complex	1.71	56.47	0.030
HDA2	Hda1 HDAC	1	23.74	0.042
RAD6	H2Bub	1	17.85	0.056
SPT20	SAGA Complex	1.91	32.50	0.059
ADA2	SAGA Complex	1	11.34	0.089
CTR9	Paf1 Complex	13.75	143.77	0.096

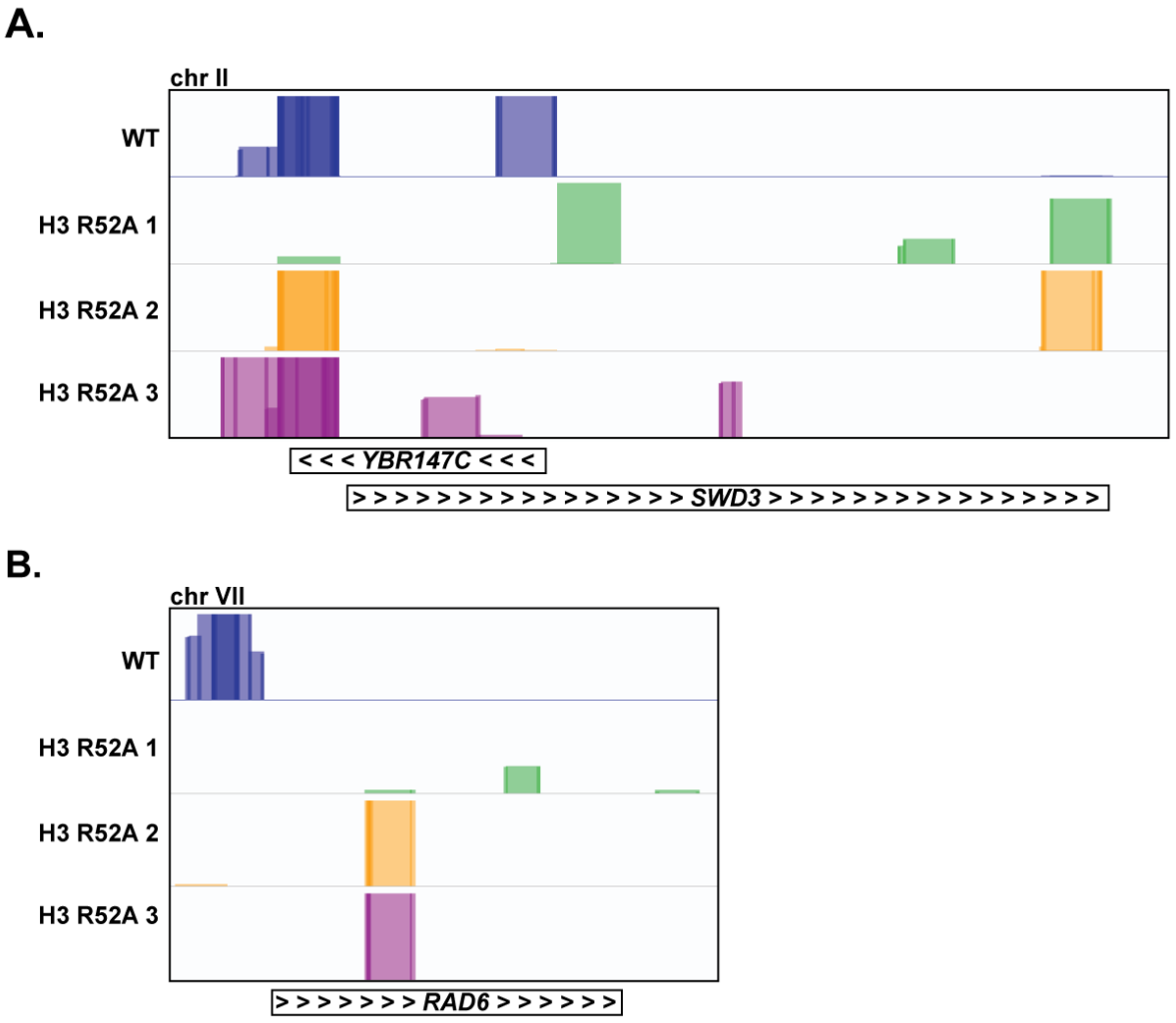

Figure 44. Genes for which mutations potentially suppress H3 R52A and lead to greater transposition efficiency in the mutant than the wild type.

Gene browser views of **(A)** *SWD3*, a subunit of COMPASS, and **(B)** H2Bub E2, *RAD6*. For all panels, transposon reads are enriched in mutant replicates over the wild type.

4.4 **DISCUSSION**

4.4.1 **Technical Troubleshooting Will Generate More Saturated, Comprehensive Libraries in Future Studies**

This study piloted the use of a technique to generate saturated transposon insertions of the yeast genome in either an environmental context or in the presence of a mutation of choice (MICHEL *et al.* 2017). Our initial attempt at this screen was successful in preliminarily identifying potential genetic interactors with the DNA entry-exit site, but verification of these relationships requires further, directed study. Some of the shortcomings of the current data lie in the relatively small transposon library we generated in comparison to recommendations from the published protocol, being approximately 2-fold less saturated at $9x10^5$ (H3 R52A) and $1.3x10^6$ (wild type) clones compared to $2x10^6$ clones. Due to the poor transposition efficiency related to strain background differences, the additional plasmid in our query strains, and perhaps the health of the H3 R52A mutant, generation of larger libraries would require scaling up and increasing original plating number.

It may be possible to optimize this protocol for future use through construction of an integration vector for the SATAY system into the query strain of choice, thus eliminating dependence on copy number of pBK549 for coverage of other portions of the yeast genome. Additionally, where possible, use of query strains whose genetic manipulation of interest (a

mutation, deletion, or overexpression) is integrated into the yeast genome as opposed to plasmid-borne eliminates the need for any plasmids in the strain of choice, driving all transposition events toward genomic regions. On top of this, it may benefit future screening attempts to switch strain backgrounds to BY4741, into which Dr. Margaret Shirra plans to attempt integrating the H3 R52A mutant. H3 R52A remains a unique case where integration of such a mutation into the yeast genome, even at both copies of the H3-H4 histone loci, may be lethal to the cells. Further work will be required to determine if a viable strain can be generated.

4.4.2 **Preliminary Evidence for Positive and Negative Genetic Interactors with the DNA Entry-Exit Site**

So far, our results demonstrate that some dubious ORFs (unannotated genes) that overlap known, transcription related genes, may be synthetically lethal with H3 R52A. Perhaps most interesting from this list is *YDR290W,* which overlaps termination-related gene *RTT103.* Rtt103, which binds Ser2 phosphorylated Pol II CTD, recruits Rat1/Rai1 for Pol II termination (KIM *et al.* 2004). Where termination readthrough is already observed, a mutation such as *rtt103*Δ that impairs Rat1/Rai1 recruitment, and likely causes delayed transcription termination, seems a logical mechanism to exacerbate termination defects to the point of lethality. Relatedly, loss of *CTK3,* a member of the CTDK-1 complex that phosphorylates Ser2 on the CTD of Pol II, is synthetically lethal with H3 R52A. It is possible that this interaction is related to the interaction between the DNA entry-exit site and Rtt103, where reduction of Ser2P CTD may reduce recruitment of later termination factors that "fail-safe" faulty NNS-dependent termination (RONDON *et al.* 2009; COLIN *et al.* 2014; ROY *et al.* 2016).

Components of histone modification complexes COMPASS, NuA4, and SAGA may also be synthetically lethal with H3 R52A as indicated by our preliminary SATAY results. Interestingly, transposition within *SWD3*, another member of COMPASS, may enhance growth of the mutant. Set1 and its methylation of H3K4 have previously been described as requirements for termination by the NNS pathway (TERZI *et al.* 2011). This work showed synthetic growth defects and exacerbated termination defects in strains harboring a Nrd1 mutation alongside a *SET1* deletion or with the unmodifiable H3K4R allele. The Paf1 complex also has subunits in the synthetic lethal and suppressor categories. Earlier studies by the Arndt lab have linked the role of the Paf1 complex in H2BK123ub to NNS-dependent termination, which is a prerequisite for placement of H3K4me (SHELDON *et al.* 2005; TOMSON *et al.* 2011; TOMSON *et al.* 2013). While it's not entirely clear why members of these cotranscriptional histone modification complexes are divided between two starkly different genetic categories with respect to the H3 R52A mutant, their presence further supports a dependence on modulation of histone modification for early termination by the NNS complex. Future work will be needed to confirm these genetic interactions.

Our preliminary data suggest histone acetylation complexes likewise genetically interact positively and negatively with the DNA entry-exit site. NuA4 catalyzes acetylation of H4 and H2A.Z (KEOGH *et al.* 2006) to regulate cellular events such as transcription (DOYON AND COTE 2004), chromosome segregation (KROGAN *et al.* 2004), and DNA repair (DOYON AND COTE 2004; SQUATRITO *et al.* 2006). The Tip60 complex in humans, homologous to the yeast NuA4 complex, also has non-histone targets (SAPOUNTZI *et al.* 2006), suggesting that the same may be true of yeast though such a role has not yet been described. More recent studies have shown that *EAF1,* which falls into the category of synthetically lethal with H3 R52A, is required for NuA4 complex integrity (MITCHELL *et al.* 2008). Its mutation here presumably reduces NuA4 function, which suggests that

loss of regulation of nucleosome accessibility through critical acetylation marks is detrimental to cells with a defective nucleosome DNA entry-exit site, likely rendering nucleosomes constitutively more accessible.

The requirement for a balance of nucleosome accessibility is supported by histone acetyltransferase complex mutations potentially suppressing H3 R52A growth phenotypes, as with components of the SAGA complex and Hda2, a subunit of the Hda1/2/3 histone deacetylase complex. Past work on the interplay between sense and antisense transcription, and regulators of that process, showed that the Hda1/2/3 complex deacetylates nucleosomes at the NDR and 5' end of genes to silence sense transcription and allow for antisense transcription to occur (CAMBLONG *et al.* 2007; CASTELNUOVO *et al.* 2014). In a genetic background where we have observed widespread antisense transcription (Chapter 3), the genetic interaction between H3 R52A and mutation of *HDA2* proposes a mechanism where relief of this sense transcription repression through loss of the Hda1/2/3 complex balances the upregulation of antisense transcription in a nucleosome deficient DNA entry-exit site mutant.

Together, initial findings of our pilot SATAY screen suggest that multiple factors known to be involved in other stages of transcription may influence termination in a way that requires the DNA entry-exit site of the nucleosome. Future work on these known and unknown factors alike will provide not only more in-depth mechanistic insight into the role of the entry-exit site for transcription termination, but, excitingly, may also result in a preliminary understanding of the functions of unannotated yeast genes.

5.0 CONCLUSIONS AND FUTURE DIRECTIONS

Setting out, the goal of my project was to identify how the nucleosome, a major regulator of DNA-templated processes, contributes to transcription termination. Several previous studies hinted that the nucleosome is an important player in this process, as in other stages of transcription, but a mechanism for its role had not yet been defined (ALEN *et al.* 2002; TERZI *et al.* 2011; TOMSON *et al.* 2011; TOMSON *et al.* 2013; ROY *et al.* 2016). In light of this, I sought to approach investigating roles for the nucleosome in this important stage of transcription by querying the requirement for particular nucleosomal surfaces that might bind specific chromatin or transcription factors involved in termination. My work in the Arndt lab began as a comprehensive screen for histone H3 and H4 mutations that cause termination defects. From that initial screen, we have learned much more about one particular surface of the nucleosome, the DNA entry-exit site, and how it functions to regulate nucleosome accessibility for processes like transcription termination.

5.1 THE DNA ENTRY-EXIT SITE OF THE NUCLEOSOME IS REQUIRED FOR TRANSCRIPTION TERMINATION GENOME-WIDE

To investigate the requirement for specific histone residues in transcription termination, I performed a genetic screen using an alanine-scanning histone mutant plasmid library (NAKANISHI *et al.* 2008) and a published termination reporter (CARROLL *et al.* 2004) (Figure 11). Truly the most profound, encouraging result from this screen was that alanine-substituted residues scoring as termination-defective mutants concentrated primarily to a single surface: the DNA entry-exit

site. The DNA entry-exit site, while not explicitly linked to transcription termination in previous literature, is richly published for its role as a regulator of nucleosome accessibility. The entry-exit site possesses the innate ability to mediate DNA breathing, where DNA near the nucleosome ends peels away from histones to allow transient accessibility without dissociating the nucleosome completely (POLACH AND WIDOM 1995). Complementary to its role in DNA breathing, this region of the nucleosome is known to facilitate gating of this transient DNA accessibility, for instance via binding non-core histone H1, especially important in higher eukaryotes for stabilizing higher-order chromatin structure (FRADO *et al.* 1983).

Later studies have suggested DNA breathing as an important vantage point for chromatin modifiers such as remodelers, histone chaperones, and histone modifiers, that may capture nucleosomes in their breathed state as a preliminary means of nucleosomal access (LI AND WIDOM 2004; LI *et al.* 2005). Recent work has described this surface of the nucleosome as a potential binding site for H3K36 histone methyltransferase Set2 (DU AND BRIGGS 2010; ENDO *et al.* 2012). Rtt109, a histone acetyltransferase required for genome stability and replication-coupled nucleosome synthesis, has also been demonstrated by X-ray crystallography to work with histone chaperone Asf1 to unwind the H3 αN helix to bind the nucleosome surface and acetylate H3 K56 (ZHANG *et al.* 2018). In addition to Asf1, the histone chaperone FACT is suspected to use the DNA entry-exit site's accessibility to invade the nucleosome to reach its target H2A-H2B dimers (HONDELE *et al.* 2013).

It is not unreasonable to think that these findings extend to transcription termination, where increased accessibility of the nucleosome through mutation of the DNA entry-exit site may eliminate an important barrier to the elongating polymerase. Genome-wide analysis by steady-state RNA-seq and nascent 4tU-seq suggest that the termination readthrough phenotype of select

DNA entry-exit site mutants, H3 T45A and R52A, is global (Figures 24-27). In light of this, I hypothesized that the general functions of the DNA entry-exit site, to mediate nucleosome accessibility and to interact with specific chromatin factors, might contribute to its role in transcription termination and set out to address more mechanistically its involvement in this process.

5.2 SET2-MEDIATED H3K36ME^3 REQUIRES RESIDUES WITHIN THE DNA ENTRY-EXIT SITE

Substitution of H3 T45, R49, or R52 for alanine results in reduced Set2 occupancy on chromatin by ChIP-qPCR (ENDO *et al.* 2012). Because of this and the loss of H3K36me^3 in some DNA entry-exit site mutants, this surface of the nucleosome has been described as a critical interaction partner for Set2, the histone methyltransferase responsible for catalyzing this mark (DU AND BRIGGS 2010; ENDO *et al.* 2012). Set2-mediated H3 K36 trimethylation is required for transcriptional fidelity, especially in preventing cryptic initiation from intragenic promoters (CARROZZA *et al.* 2005; JOSHI AND STRUHL 2005; KEOGH *et al.* 2005). In agreement with previous evidence that perturbations to the DNA entry-exit site cause reduced H3K36me^3 (DU AND BRIGGS 2010; ENDO *et al.* 2012) and that this post-translational histone modification is required for transcription termination (TOMSON *et al.* 2013), I showed that several termination-defective mutants indeed have reduced ability to place this mark. Interestingly, the two strongest termination mutants, H3 T45A and R52A, differ significantly in their abilities to support H3K36me^3, suggesting that this modification alone does not fully explain the roles of all residues in the DNA entry-exit site in transcription termination.

In H3 T45A, R49A, and R52A-expressing strains, where Set2 occupancy is reduced on chromatin, overexpression of *SET2* in mutant strains does not increase K36me^3 (Figure 30). Overexpression of *SET2* in a wild type strain likewise fails to increase trimethylation at H3 K36. Further, this does not suppress the termination defects or cryptic initiation defects observed in these mutants to a significant degree. These results may suggest that mutation of the DNA entry-exit site prohibits recruitment of Set2 to the nucleosome, that these residues are normally required for catalysis of trimethylation on H3 K36, or that these mutations have affected another requirement for H3K36me^3. Spt6, a histone chaperone involved in maintaining chromatin organization during active transcription, is required for Set2 activity (YOUDELL *et al.* 2008; LICKWAR *et al.* 2009; DRONAMRAJU *et al.* 2018b). In previous work, *SET2* overexpression was unable to restore trimethylation of H3 K36 in an *spt6-1004* mutant that seems to lose its interaction with the nucleosome (YOUDELL *et al.* 2008).

The strongest transcription termination mutant to also have an H3K36me^3 defect is H3 R52A. Considering this relationship, I investigated via genetic interaction whether H3 K36 and R52 function in similar pathways with the help of Julia Seraly. Intriguingly, histone H3 with alanine substitutions at both residues is lethal to the cell (Figure 31A). This synthetic lethality could conceivably be the result of a nucleosome that is too permissive to support viability, perhaps allowing for toxic levels of spurious transcription, suggesting that K36 and R52 function through related, but distinct, pathways.

5.3 **MUTATIONS TO THE DNA ENTRY-EXIT SITE CAUSE DECREASED NUCLEOSOME OCCUPANCY AND ALTERED POSITIONING**

In vivo studies on transcriptional regulation of the well-characterized *SRG1-SER3* locus show that histone amino acid substitutions at the DNA entry-exit site loosen nucleosome barriers against inappropriate ncRNA expression (HAINER AND MARTENS 2011a). The identification of histone mutations that disrupt the the DNA entry-exit site in a screen for transcription termination defects is consistent with such an idea: if stable nucleosomes are required near the 3' ends of genes to physically halt the polymerase for efficient termination, then loosening this chromatin barrier should allow RNA polymerase II to bypass canonical termination sites, leading to readthrough. Consistent with the hypothesis that alanine-substituted residues in the DNA entry-exit site results in partially destabilized nucleosomes, the termination mutants identified here display phenotypes that are characteristic of disrupted chromatin structure (Figures 17-21).

I further explored this possibility by comparing MNase-seq patterns of wild type and H3 R52A strains. In the entry-exit site mutant, nucleosome occupancy is reduced and nucleosome positioning changes globally. This positioning change consists of nucleosomes shifting away from 5' NDRs and more 3' nucleosomes becoming fuzzier, or less well positioned, in the H3 R52A mutant compared to wild type (Figures 33, 34). Nucleosome profiles of the *spt6-1004* mutant (DORIS *et al.* 2018), which also exhibits loss of H3K36me^3, increased cryptic initiation, and reduced nucleosome reassembly (CHEUNG *et al.* 2008), share very similar characteristics, including a loss in nucleosome occupancy, a shift away from 5' NDRs, and increased fuzziness toward the 3' end of genes. This could suggest that Spt6 mediates the involvement of the DNA entry-exit site in its global functions, but such a hypothesis requires further investigation.

The altered chromatin landscape of the H3 R52A mutant is also apparent in my Pol II ChIP-seq data (Figure 36). In a wild type strain, Pol II occupancy across a gene is associated with high, nucleosome-like periodicity. This suggests that wild type nucleosomes act as the appropriate barrier to the progression of Pol II, and that the polymerase is likely seen briefly pausing between nucleosomes prior to their partial disassembly. This periodicity is diminished in the H3 R52A mutant (reduced peak height in the metagene plot and reduced sharpness of nucleosome-like banding in the heatmap), suggesting that Pol II does not accumulate around nucleosomes as in wild type. This seems to demonstrate the relationship between the weakened nucleosome in a DNA entry-exit site mutant and relief of a stringent barrier against Pol II progression, allowing its relatively unperturbed progression.

Relatedly, the NNS-dependent termination pathway relies on the kinetic relationship between the RNA-DNA helicase, Sen1, and elongating RNA polymerase II for termination of transcription in the appropriate termination window (HAZELBAKER *et al.* 2013). These findings closely couple transcription elongation rate to termination efficiency. If impairing the DNA entry-exit site reduces nucleosome stability, it is a reasonable assumption that the polymerase may elongate faster in these mutant strains, thus contributing to terminator readthrough. I tested transcription rate in the H3 T45A and H3 R52A mutants by a well-characterized, ChIP-based assay (MASON AND STRUHL 2003), but the results of this assay suggest that the elongation rates in H3 T45A and R52A mutants are in fact somewhat slower than wild type yeast cells. Interestingly, this is in contrast to Pol II ChIP (not shown) and Pol II ChIP-seq (Figures 36, 37) that may suggest increased Pol II kinetics in the H3 R52A strain. Further work will be necessary to uncover the inconsistencies in these findings.

5.4 **INCREASING NUCLEOSOME OCCUPANCY SUPPRESSES TERMINATION DEFECTS AT A SNORNA GENE**

To directly test whether nucleosomes are specifically required for acting as roadblocks to RNA polymerase II to promote efficient termination, I integrated a superbinder sequence downstream of *SNR48.* This sequence has previously been successful in increasing histone occupancy in a site-directed manner (WANG *et al.* 2011; HAINER *et al.* 2015), driving formation of a highly stable nucleosome. Integration of this sequence at *SNR48* successfully increases histone occupancy by about 2.5-fold in wild type and H3 R52A strains and effectively suppresses terminator readthrough at this locus in both strains (Figure 39). This, along with my Pol II ChIP-seq data, supports the hypothesis that a stable nucleosome is a sufficient barrier to Pol II progression that perhaps drives the Pol II pause necessary for termination factors to perform their respective activities.

In the future, it would be beneficial to test this idea at other loci. I have designed an integration scheme to insert the superbinder sequence downstream of the *SNR47* gene as well, where loss of a specific nucleosome in the H3 R52A mutant, compared to wild type, is not apparent. If the same pattern of increased histone occupancy coupled to reduced termination readthrough is observed, this may suggest that increasing nucleosome stability, regardless of genomic location, can contribute to halting transcription. If a nucleosome in this position is unable to suppress the termination defect at *SNR47,* it is possible that the DNA entry-exit site has multiple roles in regulating transcription termination and that these roles may be applied in a locus-specific manner rather than globally.

5.5 **CHROMATIN FACTORS MAY GENETICALLY INTERACT WITH THE DNA ENTRY-EXIT SITE TO SUPPORT ITS FUNCTION IN TRANSCRIPTION**

Preliminary results from our SATAY screen suggests that known transcription- and chromatin-related genes may participate in pathways related to DNA entry-exit site function. While this screen bears repeating with a larger transposon library and deeper sequencing, our current data hints at genetic relationships between the DNA entry-exit site and known termination factors, histone modifiers, transcription elongation factors, and Pol II CTD kinases. These findings suggest that the DNA entry-exit site, while sufficient to mediate DNA breathing, does not act in isolation. Other factors involved in all three stages of transcription may influence termination in a way that requires the DNA entry-exit site of the nucleosome. Replication of the SATAY results and further analysis will allow for more in-depth mechanistic insight into the role of the entry-exit site for transcription termination.

5.6 **CONCLUDING REMARKS**

Studies discussed here and in the cited literature reveal the DNA entry-exit site of the nucleosome as a major player in when, where, and how factors gain access to the nucleosome. Intact, it is an efficient gatekeeper for processes that require DNA contact, including DNA replication, repair, recombination, and transcription. How the entry-exit site coordinates all of these delicate processes requires further intensive study, but it is tempting to speculate, due to the charge composition of this site, that post-translational modification of entry-exit site residues will continue to emerge as a theme for modulating open and closed states. Further, whether *trans*-acting factors specifically

interact with modified or unmodified entry-exit site residues to mediate these processes requires more investigation. It will be exciting to see how genetic, biochemical, and structural work builds upon our current knowledge, and advances our understanding of this critical nucleosomal surface in years to come.

APPENDIX A

Contents of this Appendix are adapted with permission under the terms of the Creative Commons CC BY license from (CUCINOTTA *et al.* 2019). The sections of the article regarding MNase sequencing – for which I generated samples, analyzed sequencing data, and wrote the methods and results – are included here. I have supplied an independent, abbreviated introduction section.

A.1 BRIEF INTRODUCTION

Nucleosomes serve two major roles in eukaryotic cells – they package DNA into the nucleus and also act as regulators of DNA-templated processes like replication, repair, recombination, and transcription. Interaction of a primarily basic histone protein surface with the negatively charged DNA phosphate backbone is a major stabilizer of the nucleosome's structure. However, the histone core assembles in such a manner that forms a surface exposed patch of acidic residues primarily in H2A and H2B. In recent years, this acidic patch has been described as a critical regulatory surface for many factors that interact with the nucleosome, including the LANA peptide of Kaposi's sarcoma virus (BARBERA *et al.* 2006), the deubiquitylation module of SAGA (MORGAN *et al.* 2016), and the ubiquitylation modules of RNF168 and PRC1 (LEUNG *et al.* 2014; MATTIROLI *et al.* 2014; MCGINTY *et al.* 2014), among others. The acidic patch has also been shown to facilitate higher-order chromatin condensation through interaction with the H4 tails of adjacent nucleosomes (KALASHNIKOVA *et al.* 2013). All of these interactions contribute to modification of the nucleosome for specific nuclear processes required for cellular function.

Recently, work from our lab and the Kohler lab has described important roles for the nucleosome acidic patch in ubiquitylation and deubiquitylation of H2B at lysine 123 (H2Bub) in yeast (CUCINOTTA *et al.* 2015; GALLEGO *et al.* 2016). Monoubiquitylation of H2B at this residue is catalyzed by the E2-E3 complex Rad6-Bre1 (ROBZYK *et al.* 2000; SUN AND ALLIS 2002; WOOD *et al.* 2003a) and requires a functional Paf1 complex (Paf1C) for its placement (NG *et al.* 2003a; WOOD *et al.* 2003b). H2Bub is a cotranscriptional histone modification (FUCHS *et al.* 2014) thought to enhance processivity of RNA polymerase II through its roles in stimulating the activities of other elongation factors, like histone chaperone FACT (PAVRI *et al.* 2006), in supporting downstream transcription-coupled modifications such as H3K4me and H3K79me (DOVER *et al.* 2002; SUN AND ALLIS 2002), and in regulation of appropriate nucleosome organization at promoters and over gene bodies (BATTA *et al.* 2011). Led by findings discussed in the manuscript not addressed in detail here, including the discovery of direct interactions between the acidic patch and the Paf1 complex and FACT, both known factors in chromatin reorganization during transcription, we sought to better understand the genome-wide chromatin landscape of an acidic patch mutant (H2A E57A) via MNase-sequencing.

A.2 MATERIALS AND METHODS

Mononucleosomes were prepared essentially as described in (WAL AND PUGH 2012). Briefly, biological duplicates of wild type and H2A E57A mutant cells were grown in appropriate SC medium to an OD_{600} of 0.8 and crosslinked with formaldehyde at a final concentration of 1%. 100 mL of cells were pelleted, resuspended in FA buffer (50 mM HEPES-KOH, pH 8.0, 150 mM NaCl, 2.0 mM EDTA, 1.0% Triton X-100, 0.1% sodium deoxycholate), and lysed by bead beating. Following centrifugation, the chromatin-containing pellet was resuspended in NP-S buffer (0.5 mM spermidine, 0.075% IGEPAL, 50 mM NaCl, 10 mM Tris-Cl, pH 7.5, 5 mM $MgCl_2$, 1 mM $CaCl_2$), and then subjected to digestion by MNase (Thermo Scientific 88216) (Figure 45). Mononucleosomal DNA was purified from treatments with 2.5 (low), 20 (mid), and 40 (high) U MNase digestions using agarose gel electrophoresis and Freeze N' Squeeze Columns (BioRad 7326166). A fixed amount of MNase-digested, gel purified *Kluyveromyces lactis* DNA was added to each sample for spike-in normalization to assess occupancy changes using a method described in (HU *et al.* 2014). Sequencing libraries were prepared using the NEBNext Ultra II kit (NEB E7645) and NEBNext Ultra sequencing indexes (NEB E7335) according to manufacturer's instructions. Libraries were quantified using Qubit and TapeStation and pooled for paired-end sequencing on an Illumina NextSeq 500 (UPMC Children's Hospital, Health Sciences Sequencing Core). Sequencing reads were aligned with HISAT2 (KIM *et al.* 2015; PERTEA *et al.* 2016) to the sacCer3 reference genome (LANGMEAD AND SALZBERG 2012) and low quality reads were filtered using the SAMtools suite (LI *et al.* 2009). Reads from biological duplicate samples were merged for downstream analysis. Reproducibility of the MNase-seq data for the two biological replicates, both wild type and H2A E57A, is demonstrated through heatmaps, biplots, and a Pearson correlation plot in Figure 45B-D. Nucleosome positioning was analyzed genome-wide relative to

published +1 nucleosome dyad positions (ZHANG *et al.* 2011) using the MNase option in the deepTools2 suite (RAMIREZ *et al.* 2014). Counts represent tag centers (3 bp) of uniquely mapped fragments 135 – 160 bp in length, counted in 1 bp windows and averaged over 50 bp. Heatmaps and metagene plots were generated using deepTools2 (RAMIREZ *et al.* 2014). Genome browser images were generated in Integrative Genomics Viewer (IGV 2.4.13) (ROBINSON *et al.* 2011; THORVALDSDOTTIR *et al.* 2013). Data presented here are lowly digested samples (2.5 U) to preserve and most accurately represent steady-state chromatin architecture.

In the course of this analysis, a duplication of chromosome VIII was detected in the H2A E57A MNase-seq datasets. Using whole genome sequencing, we independently verified the disomy of chrVIII in this strain and screened additional H2A E57A strains, generated by plasmid shuffle, for aneuploidies. In this way, we identified a strain with normal chromosome composition and subjected it, along with the original H2A E57A strain, to H3 ChIP-seq analysis. A comparision of the H3 ChIP-seq datasets between the disomic and non-disomic H2A E57A strains confirmed that chromatin architecture phenotypes were the same. We are, therefore, confident in our MNase-seq datasets. We removed chromosome VIII reads from MNase-seq analysis of both wild type and H2A E57A datasets to reduce any artifacts related to overrepresentation of chromosome VIII nucleosomes in meta-analyses.

A.3 RESULTS AND DISCUSSION

We have shown that the nucleosome acidic patch directly interacts with subunits of two transcription elongation factors, Paf1C and FACT, that regulate nucleosome occupancy and modification. Led by these findings, we set out to address whether mutations to the acidic patch alter nucleosome positions and occupancy genome-wide using MNase-seq analysis of a strain in which the only copy of H2A is the viable, H2BK123ub-defective E57A allele (CUCINOTTA *et al.* 2015). The spike-in normalized MNase-seq data revealed a general reduction in nucleosome occupancy in the H2A E57A mutant at all genes (Figure 46A-C) and a shift of nucleosomes toward the 5' nucleosome-depleted region (NDR) in the H2A E57A mutant (Figure 46B). Over gene bodies, this 5'-shift in nucleosome position in the mutant relative to wild type increases from a modest 3 base pair shift, on average, at the +1 nucleosome to a more dramatic 12 base pair shift at the +4 nucleosome when all genes are considered.

Due to the dynamic nature of chromatin during transcription, we asked whether this nucleosome shift in the H2A E57A mutant was influenced by gene expression level. Therefore, we partitioned our MNase-seq data according to experimentally defined expression quintiles (HOLSTEGE *et al.* 1998) and plotted read density averages in each of these subsets as metagene plots. In agreement with previous studies (BINTU *et al.* 2011; NOCETTI AND WHITEHOUSE 2016), nucleosome positions at the most highly expressed genes in the wild type strain were shifted in the 5' direction when compared to nucleosome positions at the most lowly expressed genes (Figure 46D, compare peak positions of the solid and dotted pink lines). The extent of nucleosome shifting in the H2A E57A mutant was similar to that of wild type when gene expression level was considered (Figure 46D, compare peak positions of the solid and dotted purple lines). In addition to this 5' shift in position, we also observed a general loss of positioning at more 3' nucleosomes

in the mutant as evidenced by greater peak widths. For instance, at the +3 nucleosome (Figure 46E, compare pink lines to purple), the width at half peak height is 63 base pairs on average in the wild type compared to 69 base pairs in the H2A E57A mutant. Increased peak width suggests "fuzziness" in positioning, or a pronounced inconsistency in nucleosome phasing within the population of mutant yeast compared to wild type. While pervasive, the changes in nucleosome positioning, as measured by 5' shift or peak width, observed in the H2A E57A were not strongly correlated with gene expression level (Figure 46E).

H2B ubiquitylation levels generally correspond with transcription levels and nucleosome occupancy, and the acidic patch promotes H2Bub (XIAO *et al.* 2005; TANNY *et al.* 2007; BATTA *et al.* 2011; CUCINOTTA *et al.* 2015; GALLEGO *et al.* 2016). Therefore, we asked whether changes in chromatin architecture relate to the reduced levels of ubiquitylation in the H2A E57A mutant. We subset our MNase-seq data based on our published H2Bub ChIP-exo data (VAN OSS *et al.* 2016) and found that nucleosome positions at highly ubiquitylated genes mirror the 5' shift of highly expressed genes, consistent with the relationship between H2Bub and gene expression (Figure 46F) (MINSKY *et al.* 2008; KIM *et al.* 2009; FUCHS *et al.* 2014). We also observed, in agreement with previous data (BATTA *et al.* 2011), that wild type nucleosomes exhibit higher occupancy over highly ubiquitylated genes. This difference in occupancy is diminished in the H2A E57A strain, likely as a result of the inability of this acidic patch mutant to promote placement of the H2Bub mark.

Together, these results demonstrate that the acidic patch influences nucleosome occupancy and positioning on active genes through mechanisms that likely depend on its multiple interactions with transcription elongation factors, histone modifiers, and other chromatin-associated factors.

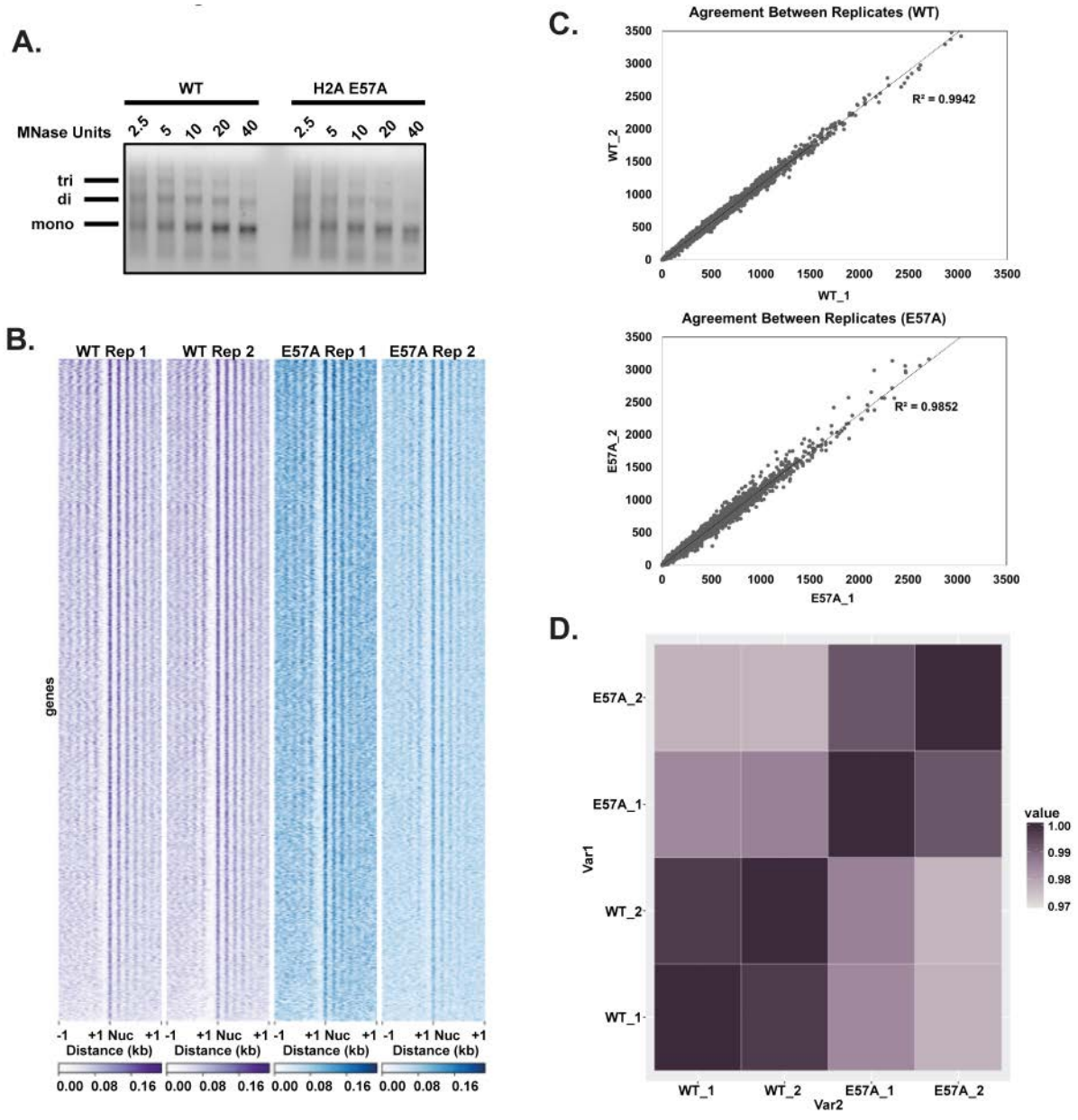

Figure 45. MNase titration of cell lysates and analysis of MNase-seq data reproducibility.

(A) Increasing units of MNase were added to cell lysates. Bottom band shows the mononucleosome species that was purified and subjected to library preparation and paired-end sequencing. **(B)** Heatmaps of nucleosome positions determined in biological duplicate by MNase-seq analysis of strains expressing plasmid-encoded wild type H2A or H2A E57A as the only source of H2A. Sequences are aligned to the +1 nucleosome of annotated yeast genes. **(C)** Biplots of MNase sequencing signal averaged over each

annotated yeast gene comparing biological replicates of wild type (top) or H2A E57A (bottom). R^2 values are indicated to the right of the line of best fit. **(D)** Pearson correlation heat maps of MNase-seq data sets from all four samples. Darker color indicates higher correlation.

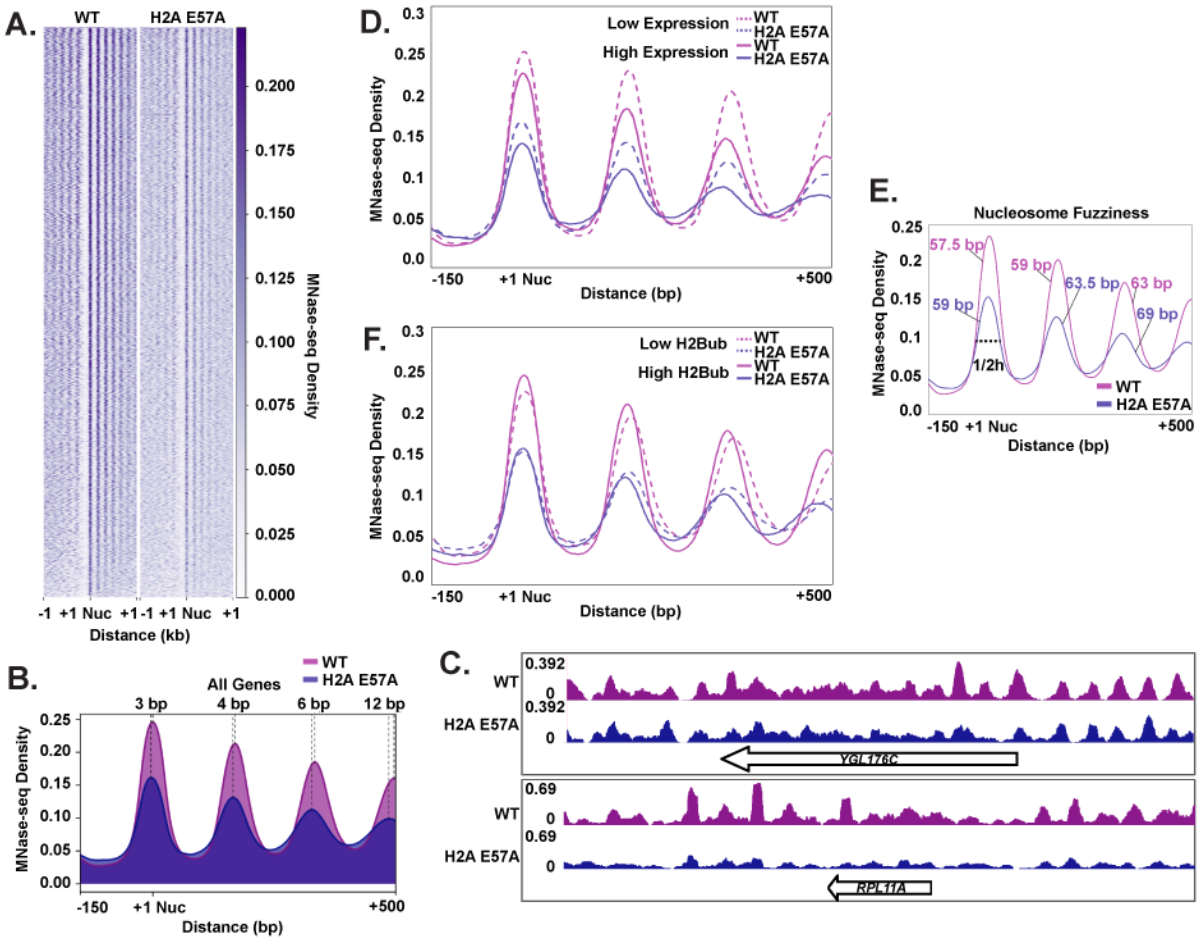

Figure 46. An intact nucleosome acidic patch is required for proper nucleosome occupancy and positioning.

(A) Heatmaps of nucleosome positions as determined by MNase-seq analysis of strains expressing plasmid-encoded wild type H2A or H2A E57A as the only source of H2A. Sequences are aligned to the +1 nucleosome of ORFs (ZHANG *et al.* 2011). Heatmaps represent averages of two biological replicates. **(B)** Metaplot of nucleosome positions at all genes in the indicated strains. **(C)** Spike-in normalized MNase-seq reads at two individual genes. **(D)** Comparison of nucleosome positions in WT and H2A E57A strains at

lowly expressed and highly expressed genes (bottom 20% and top 20%) (HOLSTEGE *et al.* 1998). **(E)** Comparison of peak widths at ½ peak height in WT and the H2A E57A mutant. Greater peak width suggests poorer nucleosome positioning on average across genes. **(F)** Comparison of nucleosome positions at genes with high and low levels of H2Bub (bottom and top quintiles) (VAN OSS *et al.* 2016).

REFERENCES

Adelman, K., W. Wei, M. B. Ardehali, J. Werner, B. Zhu *et al.*, 2006 Drosophila Paf1 modulates chromatin structure at actively transcribed genes. Mol Cell Biol 26**:** 250-260.

Adkins, M. W., S. R. Howar and J. K. Tyler, 2004 Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. Mol Cell 14**:** 657-666.

Alen, C., N. A. Kent, H. S. Jones, J. O'Sullivan, A. Aranda *et al.*, 2002 A role for chromatin remodeling in transcriptional termination by RNA polymerase II. Mol Cell 10**:** 1441-1452.

Allan, J., P. G. Hartman, C. Crane-Robinson and F. X. Aviles, 1980 The structure of histone H1 and its location in chromatin. Nature 288**:** 675-679.

Allard, S., R. T. Utley, J. Savard, A. Clarke, P. Grant *et al.*, 1999 NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. EMBO J 18**:** 5108-5119.

Allen, B. L., and D. J. Taatjes, 2015 The Mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol 16**:** 155-166.

Allmang, C., J. Kufel, G. Chanfreau, P. Mitchell, E. Petfalski *et al.*, 1999 Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J 18**:** 5399-5410.

Arents, G., R. W. Burlingame, B. C. Wang, W. E. Love and E. N. Moudrianakis, 1991 The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A 88**:** 10148-10152.

Arents, G., and E. N. Moudrianakis, 1995 The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc Natl Acad Sci U S A 92**:** 11170-11174.

Baillat, D., M. A. Hakimi, A. M. Naar, A. Shilatifard, N. Cooch *et al.*, 2005 Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell 123**:** 265-276.

Baker, D. L., O. A. Youssef, M. I. Chastkofsky, D. A. Dy, R. M. Terns *et al.*, 2005 RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. Genes Dev 19**:** 1238-1248.

Baker, S. P., J. Phillips, S. Anderson, Q. Qiu, J. Shabanowitz *et al.*, 2010 Histone H3 Thr 45 phosphorylation is a replication-associated post-translational modification in S. cerevisiae. Nat Cell Biol 12**:** 294-298.

Baptista, T., S. Grunberg, N. Minoungou, M. J. E. Koster, H. T. M. Timmers *et al.*, 2017 SAGA Is a General Cofactor for RNA Polymerase II Transcription. Mol Cell 68**:** 130-143 e135.

Barbera, A. J., J. V. Chodaparambil, B. Kelley-Clarke, V. Joukov, J. C. Walter *et al.*, 2006 The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. Science 311**:** 856-861.

Bartholomew, B., 2014 Regulating the chromatin landscape: structural and mechanistic perspectives. Annu Rev Biochem 83**:** 671-696.

Basehoar, A. D., S. J. Zanton and B. F. Pugh, 2004 Identification and distinct regulation of yeast TATA box-containing genes. Cell 116**:** 699-709.

Batta, K., Z. Zhang, K. Yen, D. B. Goffman and B. F. Pugh, 2011 Genome-wide function of H2B ubiquitylation in promoter and genic regions. Genes Dev 25**:** 2254-2265.

Becker, D. M., and V. Lundblad, 2001 Introduction of DNA into yeast cells. Curr Protoc Mol Biol Chapter 13**:** Unit13 17.

Belotserkovskaya, R., S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky *et al.*, 2003 FACT facilitates transcription-dependent nucleosome alteration. Science 301**:** 1090-1093.

Bernstein, B. E., E. L. Humphrey, R. L. Erlich, R. Schneider, P. Bouman *et al.*, 2002 Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci U S A 99**:** 8695-8700.

Bilokapic, S., M. Strauss and M. Halic, 2018 Histone octamer rearranges to adapt to DNA unwrapping. Nat Struct Mol Biol 25**:** 101-108.

Bintu, L., M. Kopaczynska, C. Hodges, L. Lubkowska, M. Kashlev *et al.*, 2011 The elongation rate of RNA polymerase determines the fate of transcribed nucleosomes. Nat Struct Mol Biol 18**:** 1394-1399.

Bonnet, A., and B. Palancade, 2014 Regulation of mRNA trafficking by nuclear pore complexes. Genes (Basel) 5**:** 767-791.

Bortolin, M. L., J. P. Bachellerie and B. Clouet-d'Orval, 2003 In vitro RNP assembly and methylation guide activity of an unusual box C/D RNA, cis-acting archaeal pre-tRNA(Trp). Nucleic Acids Res 31**:** 6524-6535.

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

Brehove, M., T. Wang, J. North, Y. Luo, S. J. Dreher *et al.*, 2015 Histone core phosphorylation regulates DNA accessibility. J Biol Chem 290**:** 22612-22621.

Briggs, S. D., M. Bryk, B. D. Strahl, W. L. Cheung, J. K. Davie *et al.*, 2001 Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev 15**:** 3286-3295.

Brower-Toland, B. D., C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson *et al.*, 2002 Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. Proc Natl Acad Sci U S A 99**:** 1960-1965.

Buratowski, S., 2009 Progression through the RNA polymerase II CTD cycle. Mol Cell 36**:** 541-546.

Burgess, D. J., 2014 Transcription: Interference from near and far. Nat Rev Genet 15**:** 144.

Burns, L. G., and C. L. Peterson, 1997 The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo. Mol Cell Biol 17**:** 4811-4819.

Cabart, P., and D. S. Luse, 2012 Inactivated RNA polymerase II open complexes can be reactivated with TFIIE. J Biol Chem 287**:** 961-967.

Cabart, P., A. Ujvari, M. Pal and D. S. Luse, 2011 Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. Proc Natl Acad Sci U S A 108**:** 15786-15791.

Camblong, J., N. Iglesias, C. Fickentscher, G. Dieppois and F. Stutz, 2007 Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell 131**:** 706-717.

Candau, R., J. X. Zhou, C. D. Allis and S. L. Berger, 1997 Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function in vivo. EMBO J 16**:** 555-565.

Carroll, K. L., R. Ghirlando, J. M. Ames and J. L. Corden, 2007 Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. RNA 13**:** 361-373.

Carroll, K. L., D. A. Pradhan, J. A. Granek, N. D. Clarke and J. L. Corden, 2004 Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol Cell Biol 24**:** 6241-6252.

Carrozza, M. J., B. Li, L. Florens, T. Suganuma, S. K. Swanson *et al.*, 2005 Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 123**:** 581-592.

Casadio, F., X. Lu, S. B. Pollock, G. LeRoy, B. A. Garcia *et al.*, 2013 H3R42me2a is a histone modification with positive transcriptional effects. Proc Natl Acad Sci U S A 110**:** 14894-14899.

Castelnuovo, M., J. B. Zaugg, E. Guffanti, A. Maffioletti, J. Camblong *et al.*, 2014 Role of histone modifications and early termination in pervasive transcription and antisense-mediated gene silencing in yeast. Nucleic Acids Res 42**:** 4348-4362.

Cavaille, J., K. Buiting, M. Kiefmann, M. Lalande, C. I. Brannan *et al.*, 2000 Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc Natl Acad Sci U S A 97**:** 14311-14316.

Chandy, M., J. L. Gutierrez, P. Prochasson and J. L. Workman, 2006 SWI/SNF displaces SAGA-acetylated nucleosomes. Eukaryot Cell 5**:** 1738-1747.

Charpentier, B., S. Muller and C. Branlant, 2005 Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. Nucleic Acids Res 33**:** 3133-3144.

Chatterjee, N., D. Sinha, M. Lemma-Dechassa, S. Tan, M. A. Shogren-Knaak *et al.*, 2011 Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms. Nucleic Acids Res 39**:** 8378-8391.

Chen, C. C., J. J. Carson, J. Feser, B. Tamburini, S. Zabaronick *et al.*, 2008 Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. Cell 134**:** 231-243.

Chen, F. X., E. R. Smith and A. Shilatifard, 2018a Born to run: control of transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol 19**:** 464-478.

Chen, K., Y. Xi, X. Pan, Z. Li, K. Kaestner *et al.*, 2013 DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res 23**:** 341-351.

Chen, P., L. Dong, M. Hu, Y. Z. Wang, X. Xiao *et al.*, 2018b Functions of FACT in Breaking the Nucleosome and Maintaining Its Integrity at the Single-Nucleosome Level. Mol Cell 71**:** 284-293 e284.

Chen, Y., J. M. Tokuda, T. Topping, S. P. Meisburger, S. A. Pabit *et al.*, 2017 Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. Proc Natl Acad Sci U S A 114**:** 334-339.

Cheung, V., G. Chua, N. N. Batada, C. R. Landry, S. W. Michnick *et al.*, 2008 Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS Biol 6**:** e277.

Chinchilla, K., J. B. Rodriguez-Molina, D. Ursic, J. S. Finkel, A. Z. Ansari *et al.*, 2012 Interactions of Sen1, Nrd1, and Nab3 with multiple phosphorylated forms of the Rpb1 C-terminal domain in Saccharomyces cerevisiae. Eukaryot Cell 11**:** 417-429.

Cho, E. J., M. S. Kobor, M. Kim, J. Greenblatt and S. Buratowski, 2001 Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev 15**:** 3319-3329.

Christen, B., E. Abeliuk, J. M. Collier, V. S. Kalogeraki, B. Passarelli *et al.*, 2011 The essential genome of a bacterium. Mol Syst Biol 7**:** 528.

Chu, Y., R. Simic, M. H. Warner, K. M. Arndt and G. Prelich, 2007 Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. EMBO J 26**:** 4646-4656.

Colin, J., T. Candelli, O. Porrua, J. Boulay, C. Zhu *et al.*, 2014 Roadblock termination by reb1p restricts cryptic and readthrough transcription. Mol Cell 56**:** 667-680.

Collart, M. A., and S. Oliviero, 2001 Preparation of yeast RNA. Curr Protoc Mol Biol Chapter 13**:** Unit13 12.

Conaway, R. C., K. P. Garrett, J. P. Hanley and J. W. Conaway, 1991 Mechanism of promoter selection by RNA polymerase II: mammalian transcription factors alpha and beta gamma promote entry of polymerase into the preinitiation complex. Proc Natl Acad Sci U S A 88**:** 6205-6209.

Conrad, N. K., S. M. Wilson, E. J. Steinmetz, M. Patturajan, D. A. Brow *et al.*, 2000 A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. Genetics 154**:** 557-571.

Creamer, T. J., M. M. Darby, N. Jamonnak, P. Schaughency, H. Hao *et al.*, 2011 Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. PLoS Genet 7**:** e1002329.

Cucinotta, C. E., A. E. Hildreth, B. M. McShane, M. K. Shirra and K. M. Arndt, 2019 The nucleosome acidic patch directly interacts with subunits of the Paf1 and FACT complexes and controls chromatin architecture in vivo. Nucleic Acids Res.

Cucinotta, C. E., A. N. Young, K. M. Klucevsek and K. M. Arndt, 2015 The Nucleosome Acidic Patch Regulates the H2B K123 Monoubiquitylation Cascade and Transcription Elongation in Saccharomyces cerevisiae. PLoS Genet 11**:** e1005420.

Czaja, W., P. Mao and M. J. Smerdon, 2014 Chromatin remodelling complex RSC promotes base excision repair in chromatin of Saccharomyces cerevisiae. DNA Repair (Amst) 16**:** 35-43.

Dangkulwanich, M., T. Ishibashi, S. Liu, M. L. Kireeva, L. Lubkowska *et al.*, 2013 Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. Elife 2**:** e00971.

Darzacq, X., B. E. Jady, C. Verheggen, A. M. Kiss, E. Bertrand *et al.*, 2002 Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. EMBO J 21**:** 2746-2756.

Davis, C. A., and M. Ares, Jr., 2006 Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 103**:** 3262-3267.

Decatur, W. A., and M. J. Fournier, 2003 RNA-guided nucleotide modification of ribosomal and other RNAs. J Biol Chem 278**:** 695-698.

Dennis, P. P., A. Omer and T. Lowe, 2001 A guided tour: small RNA function in Archaea. Mol Microbiol 40**:** 509-519.

Doris, S. M., J. Chuang, O. Viktorovskaya, M. Murawska, D. Spatt *et al.*, 2018 Spt6 Is Required for the Fidelity of Promoter Selection. Mol Cell 72**:** 687-699 e686.

Dover, J., J. Schneider, M. A. Tawiah-Boateng, A. Wood, K. Dean *et al.*, 2002 Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem 277**:** 28368-28371.

Doyon, Y., and J. Cote, 2004 The highly conserved and multifunctional NuA4 HAT complex. Curr Opin Genet Dev 14**:** 147-154.

Driscoll, R., A. Hudson and S. P. Jackson, 2007 Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science 315**:** 649-652.

Dronamraju, R., A. J. Hepperla, Y. Shibata, A. T. Adams, T. Magnuson *et al.*, 2018a Spt6 Association with RNA Polymerase II Directs mRNA Turnover During Transcription. Mol Cell 70**:** 1054-1066 e1054.

Dronamraju, R., J. L. Kerschner, S. A. Peck, A. J. Hepperla, A. T. Adams *et al.*, 2018b Casein Kinase II Phosphorylation of Spt6 Enforces Transcriptional Fidelity by Maintaining Spn1-Spt6 Interaction. Cell Rep 25**:** 3476-3489 e3475.

Du, H. N., and S. D. Briggs, 2010 A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. J Biol Chem 285**:** 11704-11713.

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

Duffy, E. E., M. Rutenberg-Schoenberg, C. D. Stark, R. R. Kitchen, M. B. Gerstein *et al.*, 2015 Tracking Distinct RNA Populations Using Efficient and Reversible Covalent Chemistry. Mol Cell 59**:** 858-866.

Dunn, E. F., C. M. Hammell, C. A. Hodge and C. N. Cole, 2005 Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. Genes Dev 19**:** 90-103.

Ellison, M. A., A. R. Lederer, M. H. Warner, T. N. Mavrich, E. A. Raupach *et al.*, 2019 The Paf1 Complex Broadly Impacts the Transcriptome of Saccharomyces cerevisiae. Genetics.

Emili, A., D. M. Schieltz, J. R. Yates, 3rd and L. H. Hartwell, 2001 Dynamic interaction of DNA damage checkpoint protein Rad53 with chromatin assembly factor Asf1. Mol Cell 7**:** 13-20.

Endo, H., Y. Nakabayashi, S. Kawashima, T. Enomoto, M. Seki *et al.*, 2012 Nucleosome surface containing nucleosomal DNA entry/exit site regulates H3-K36me3 via association with RNA polymerase II and Set2. Genes Cells 17**:** 65-81.

Endoh, M., W. Zhu, J. Hasegawa, H. Watanabe, D. K. Kim *et al.*, 2004 Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. Mol Cell Biol 24**:** 3324-3336.

Ezeokonkwo, C., M. A. Ghazy, A. Zhelkovsky, P. C. Yeh and C. Moore, 2012 Novel interactions at the essential N-terminus of poly(A) polymerase that could regulate poly(A) addition in Saccharomyces cerevisiae. FEBS Lett 586**:** 1173-1178.

Fabrega, C., V. Shen, S. Shuman and C. D. Lima, 2003 Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. Mol Cell 11**:** 1549-1561.

Fassler, J. S., and F. Winston, 1988 Isolation and analysis of a novel class of suppressor of Ty insertion mutations in Saccharomyces cerevisiae. Genetics 118**:** 203-212.

Fillingham, J., J. Recht, A. C. Silva, B. Suter, A. Emili *et al.*, 2008 Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Mol Cell Biol 28**:** 4342-4353.

Fitzgerald, M., and T. Shenk, 1981 The sequence 5'-AAUAAA-3'forms parts of the recognition site for polyadenylation of late SV40 mRNAs. Cell 24**:** 251-260.

Frado, L. L., C. V. Mura, B. D. Stollar and C. L. Woodcock, 1983 Mapping of histone H5 sites on nucleosomes using immunoelectron microscopy. J Biol Chem 258**:** 11984-11990.

Fuchs, G., D. Hollander, Y. Voichek, G. Ast and M. Oren, 2014 Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. Genome Res 24**:** 1572-1583.

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

Gallego, L. D., M. Ghodgaonkar Steger, A. A. Polyansky, T. Schubert, B. Zagrovic *et al.*, 2016 Structural mechanism for the recognition and ubiquitination of a single nucleosome residue by Rad6-Bre1. Proc Natl Acad Sci U S A 113**:** 10553-10558.

Garas, M., B. Dichtl and W. Keller, 2008 The role of the putative 3' end processing endonuclease Ysh1p in mRNA and snoRNA synthesis. RNA 14**:** 2671-2684.

Garcia, B. A., S. B. Hake, R. L. Diaz, M. Kauer, S. A. Morris *et al.*, 2007 Organismal differences in post-translational modifications in histones H3 and H4. J Biol Chem 282**:** 7641-7655.

Ghazy, M. A., S. A. Brodie, M. L. Ammerman, L. M. Ziegler and A. S. Ponticelli, 2004 Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. Mol Cell Biol 24**:** 10975-10985.

Gil, A., and N. J. Proudfoot, 1984 A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3'-end formation. Nature 312**:** 473-474.

Gilmartin, G. M., and J. R. Nevins, 1989 An ordered pathway of assembly of components required for polyadenylation site recognition and processing. Genes Dev 3**:** 2180-2190.

Gilmour, D. S., and J. T. Lis, 1986 RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. Mol Cell Biol 6**:** 3984-3989.

Ginsburg, D. S., C. K. Govind and A. G. Hinnebusch, 2009 NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5. Mol Cell Biol 29**:** 6473-6487.

Glover-Cutter, K., S. Kim, J. Espinosa and D. L. Bentley, 2008 RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat Struct Mol Biol 15**:** 71-78.

Goel, S., S. Krishnamurthy and M. Hampsey, 2012 Mechanism of start site selection by RNA polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIIH. J Biol Chem 287**:** 557-567.

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

Greber, B. J., T. H. D. Nguyen, J. Fang, P. V. Afonine, P. D. Adams *et al.*, 2017 The cryo-electron microscopy structure of human transcription factor IIH. Nature 549**:** 414-417.

Greber, B. J., D. B. Toso, J. Fang and E. Nogales, 2019 The complete structure of the human TFIIH core complex. Elife 8.

Green, E. M., A. J. Antczak, A. O. Bailey, A. A. Franco, K. J. Wu *et al.*, 2005 Replication-independent histone deposition by the HIR complex and Asf1. Curr Biol 15**:** 2044-2049.

Grosso, A. R., A. P. Leite, S. Carvalho, M. R. Matos, F. B. Martins *et al.*, 2015 Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. Elife 4.

Grunberg, S., and S. Hahn, 2013 Structural insights into transcription initiation by RNA polymerase II. Trends Biochem Sci 38**:** 603-611.

Guajardo, R., and R. Sousa, 1997 A model for the mechanism of polymerase translocation. J Mol Biol 265**:** 8-19.

Gudipati, R. K., T. Villa, J. Boulay and D. Libri, 2008 Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. Nat Struct Mol Biol 15**:** 786-794.

Ha, I., S. Roberts, E. Maldonado, X. Sun, L. U. Kim *et al.*, 1993 Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev 7**:** 1021-1032.

Hahn, S., and E. T. Young, 2011 Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics 189**:** 705-736.

Hainer, S. J., W. Gu, B. R. Carone, B. D. Landry, O. J. Rando *et al.*, 2015 Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. Genes Dev 29**:** 362-378.

Hainer, S. J., and J. A. Martens, 2011a Identification of histone mutants that are defective for transcription-coupled nucleosome occupancy. Mol Cell Biol 31**:** 3557-3568.

Hainer, S. J., and J. A. Martens, 2011b Transcription of ncDNA: Many roads lead to local gene regulation. Transcription 2**:** 120-123.

Hainer, S. J., J. A. Pruneski, R. D. Mitchell, R. M. Monteverde and J. A. Martens, 2011 Intergenic transcription causes repression by directing nucleosome assembly. Genes Dev 25**:** 29-40.

Han, J., H. Zhou, B. Horazdovsky, K. Zhang, R. M. Xu *et al.*, 2007 Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science 315**:** 653-655.

Harlen, K. M., and L. S. Churchman, 2017 The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat Rev Mol Cell Biol 18**:** 263-273.

Harper, T. M., and D. J. Taatjes, 2018 The complex structure and function of Mediator. J Biol Chem 293**:** 13778-13785.

Hassan, A. H., K. E. Neely and J. L. Workman, 2001 Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. Cell 104**:** 817-827.

Hazelbaker, D. Z., S. Marquardt, W. Wlotzka and S. Buratowski, 2013 Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. Mol Cell 49**:** 55-66.

Hector, R. E., K. R. Nykamp, S. Dheur, J. T. Anderson, P. J. Non *et al.*, 2002 Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. EMBO J 21**:** 1800-1810.

Helmlinger, D., and L. Tora, 2017 Sharing the SAGA. Trends Biochem Sci 42**:** 850-861.

Heo, K., H. Kim, S. H. Choi, J. Choi, K. Kim *et al.*, 2008 FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. Mol Cell 30**:** 86-97.

Hergeth, S. P., and R. Schneider, 2015 The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO Rep 16**:** 1439-1453.

Hewish, D. R., and L. A. Burgoyne, 1973 Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem Biophys Res Commun 52**:** 504-510.

Holstege, F. C., E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner *et al.*, 1998 Dissecting the regulatory circuitry of a eukaryotic genome. Cell 95**:** 717-728.

Holstege, F. C., D. Tantin, M. Carey, P. C. van der Vliet and H. T. Timmers, 1995 The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. EMBO J 14**:** 810-819.

Holstege, F. C., P. C. van der Vliet and H. T. Timmers, 1996 Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. EMBO J 15**:** 1666-1677.

Hondele, M., T. Stuwe, M. Hassler, F. Halbach, A. Bowman *et al.*, 2013 Structural basis of histone H2A-H2B recognition by the essential chaperone FACT. Nature 499**:** 111-114.

Hongay, C. F., P. L. Grisafi, T. Galitski and G. R. Fink, 2006 Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell 127**:** 735-745.

Houalla, R., F. Devaux, A. Fatica, J. Kufel, D. Barrass *et al.*, 2006 Microarray detection of novel nuclear RNA substrates for the exosome. Yeast 23**:** 439-454.

Hsin, J. P., and J. L. Manley, 2012 The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev 26**:** 2119-2137.

Hsu, J. M., J. Huang, P. B. Meluh and B. C. Laurent, 2003 The yeast RSC chromatin-remodeling complex is required for kinetochore function in chromosome segregation. Mol Cell Biol 23**:** 3202-3215.

Hu, F., A. A. Alcasabas and S. J. Elledge, 2001 Asf1 links Rad53 to control of chromatin assembly. Genes Dev 15**:** 1061-1066.

Hu, J., C. S. Lutz, J. Wilusz and B. Tian, 2005 Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. RNA 11**:** 1485-1493.

Hu, Z., K. Chen, Z. Xia, M. Chavez, S. Pal *et al.*, 2014 Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. Genes Dev 28**:** 396-408.

Huang, S., H. Zhou, D. Katzmann, M. Hochstrasser, E. Atanasova *et al.*, 2005 Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. Proc Natl Acad Sci U S A 102**:** 13410-13415.

Huang, X., D. Wang, D. R. Weiss, D. A. Bushnell, R. D. Kornberg *et al.*, 2010 RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. Proc Natl Acad Sci U S A 107**:** 15745-15750.

Huisinga, K. L., and B. F. Pugh, 2004 A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol Cell 13**:** 573-585.

Hyland, E. M., H. Molina, K. Poorey, C. Jie, Z. Xie *et al.*, 2011 An evolutionarily 'young' lysine residue in histone H3 attenuates transcriptional output in Saccharomyces cerevisiae. Genes Dev 25**:** 1306-1319.

Ide, S., K. Saka and T. Kobayashi, 2013 Rtt109 prevents hyper-amplification of ribosomal RNA genes through histone modification in budding yeast. PLoS Genet 9**:** e1003410.

Jack, A. P., S. Bussemer, M. Hahn, S. Punzeler, M. Snyder *et al.*, 2013 H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization. PLoS One 8**:** e51765.

Jacquier, A., 2009 The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat Rev Genet 10**:** 833-844.

Jansen, A., and K. J. Verstrepen, 2011 Nucleosome positioning in Saccharomyces cerevisiae. Microbiol Mol Biol Rev 75**:** 301-320.

Jensen, T. H., A. Jacquier and D. Libri, 2013 Dealing with pervasive transcription. Mol Cell 52**:** 473-484.

Jeronimo, C., and F. Robert, 2014 Kin28 regulates the transient association of Mediator with core promoters. Nat Struct Mol Biol 21**:** 449-455.

Jeronimo, C., S. Watanabe, C. D. Kaplan, C. L. Peterson and F. Robert, 2015 The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. Mol Cell 58**:** 1113-1123.

Jessulat, M., M. Alamgir, H. Salsali, J. Greenblatt, J. Xu *et al.*, 2008 Interacting proteins Rtt109 and Vps75 affect the efficiency of non-homologous end-joining in Saccharomyces cerevisiae. Arch Biochem Biophys 469**:** 157-164.

Jia, H., X. Wang, J. T. Anderson and E. Jankowsky, 2012 RNA unwinding by the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex. Proc Natl Acad Sci U S A 109**:** 7292-7297.

Jia, H., X. Wang, F. Liu, U. P. Guenther, S. Srinivasan *et al.*, 2011 The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. Cell 145**:** 890-901.

John, S., L. Howe, S. T. Tafrov, P. A. Grant, R. Sternglanz *et al.*, 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 Dev 14**:** 1196-1208.

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

Kagalwala, M. N., B. J. Glaus, W. Dang, M. Zofall and B. Bartholomew, 2004 Topography of the ISW2-nucleosome complex: insights into nucleosome spacing and chromatin remodeling. EMBO J 23**:** 2092-2104.

Kalashnikova, A. A., M. E. Porter-Goff, U. M. Muthurajan, K. Luger and J. C. Hansen, 2013 The role of the nucleosome acidic patch in modulating higher order chromatin structure. J R Soc Interface 10**:** 20121022.

Kannan, K., L. Wang, J. Wang, M. M. Ittmann, W. Li *et al.*, 2011 Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. Proc Natl Acad Sci U S A 108**:** 9172-9177.

Kaplan, C. D., 2013 Basic mechanisms of RNA polymerase II activity and alteration of gene expression in Saccharomyces cerevisiae. Biochim Biophys Acta 1829**:** 39-54.

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

Kasten, M., H. Szerlong, H. Erdjument-Bromage, P. Tempst, M. Werner *et al.*, 2004 Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. EMBO J 23**:** 1348-1359.

Keller, D. M., and H. Lu, 2002 p53 serine 392 phosphorylation increases after UV through induction of the assembly of the CK2.hSPT16.SSRP1 complex. J Biol Chem 277**:** 50206-50213.

Keogh, M. C., S. K. Kurdistani, S. A. Morris, S. H. Ahn, V. Podolny *et al.*, 2005 Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell 123**:** 593-605.

Keogh, M. C., T. A. Mennella, C. Sawa, S. Berthelet, N. J. Krogan *et al.*, 2006 The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes Dev 20**:** 660-665.

Khaperskyy, D. A., M. L. Ammerman, R. C. Majovski and A. S. Ponticelli, 2008 Functions of Saccharomyces cerevisiae TFIIF during transcription start site utilization. Mol Cell Biol 28**:** 3757-3766.

Kim, D., B. Langmead and S. L. Salzberg, 2015 HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12**:** 357-360.

Kim, J., M. Guermah, R. K. McGinty, J. S. Lee, Z. Tang *et al.*, 2009 RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell 137**:** 459-471.

Kim, M., N. J. Krogan, L. Vasiljeva, O. J. Rando, E. Nedea *et al.*, 2004 The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature 432**:** 517-522.

Kim, T., and S. Buratowski, 2009 Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. Cell 137**:** 259-272.

Kim, T. K., R. H. Ebright and D. Reinberg, 2000 Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science 288**:** 1418-1422.

Kireeva, M. L., N. Komissarova, D. S. Waugh and M. Kashlev, 2000 The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J Biol Chem 275**:** 6530-6536.

Komissarova, N., J. Becker, S. Solter, M. Kireeva and M. Kashlev, 2002 Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Mol Cell 10**:** 1151-1162.

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

Kornberg, R. D., and J. O. Thomas, 1974 Chromatin structure; oligomers of the histones. Science 184**:** 865-868.

Kouzarides, T., 2007 Chromatin modifications and their function. Cell 128**:** 693-705.

Krishnamurthy, S., X. He, M. Reyes-Reyes, C. Moore and M. Hampsey, 2004 Ssu72 Is an RNA polymerase II CTD phosphatase. Mol Cell 14**:** 387-394.

Krogan, N. J., K. Baetz, M. C. Keogh, N. Datta, C. Sawa *et al.*, 2004 Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. Proc Natl Acad Sci U S A 101**:** 13513-13518.

Krogan, N. J., J. Dover, A. Wood, J. Schneider, J. Heidt *et al.*, 2003a The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell 11**:** 721-729.

Krogan, N. J., M. Kim, A. Tong, A. Golshani, G. Cagney *et al.*, 2003b Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol 23**:** 4207-4218.

Kruger, W., C. L. Peterson, A. Sil, C. Coburn, G. Arents *et al.*, 1995 Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. Genes Dev 9**:** 2770-2779.

Kuhn, U., M. Gundel, A. Knoth, Y. Kerwitz, S. Rudel *et al.*, 2009 Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. J Biol Chem 284**:** 22803-22814.

Kushnirov, V. V., 2000 Rapid and reliable protein extraction from yeast. Yeast 16**:** 857-860.

LaCava, J., J. Houseley, C. Saveanu, E. Petfalski, E. Thompson *et al.*, 2005 RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell 121**:** 713-724.

Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat Methods 9**:** 357-359.

Laprade, L., D. Rose and F. Winston, 2007 Characterization of new Spt3 and TATA-binding protein mutants of Saccharomyces cerevisiae: Spt3 TBP allele-specific interactions and bypass of Spt8. Genetics 177**:** 2007-2017.

Lawrence, M., S. Daujat and R. Schneider, 2016 Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet 32**:** 42-56.

Lazarow, K., M. L. Du, R. Weimer and R. Kunze, 2012 A hyperactive transposase of the maize transposable element activator (Ac). Genetics 191**:** 747-756.

Lee, A., K. D. Hansen, J. Bullard, S. Dudoit and G. Sherlock, 2008 Novel low abundance and transient RNAs in yeast revealed by tiling microarrays and ultra high-throughput sequencing are not conserved across closely related yeast species. PLoS Genet 4**:** e1000299.

Lestrade, L., and M. J. Weber, 2006 snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic Acids Res 34**:** D158-162.

Leung, J. W., P. Agarwal, M. D. Canny, F. Gong, A. D. Robison *et al.*, 2014 Nucleosome acidic patch promotes RNF168- and RING1B/BMI1-dependent H2AX and H2A ubiquitination and DNA damage signaling. PLoS Genet 10**:** e1004178.

Li, B., M. Gogol, M. Carey, D. Lee, C. Seidel *et al.*, 2007 Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. Science 316**:** 1050-1054.

Li, G., M. Levitus, C. Bustamante and J. Widom, 2005 Rapid spontaneous accessibility of nucleosomal DNA. Nat Struct Mol Biol 12**:** 46-53.

Li, G., and J. Widom, 2004 Nucleosomes facilitate their own invasion. Nat Struct Mol Biol 11**:** 763-769.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25**:** 2078-2079.

Li, Q., H. Zhou, H. Wurtele, B. Davies, B. Horazdovsky *et al.*, 2008 Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. Cell 134**:** 244-255.

Li, Z., and H. Kono, 2018 Investigating the Influence of Arginine Dimethylation on Nucleosome Dynamics Using All-Atom Simulations and Kinetic Analysis. J Phys Chem B 122**:** 9625-9634.

Liao, Y., G. K. Smyth and W. Shi, 2014 featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30**:** 923-930.

Lickwar, C. R., B. Rao, A. A. Shabalin, A. B. Nobel, B. D. Strahl *et al.*, 2009 The Set2/Rpd3S pathway suppresses cryptic transcription without regard to gene length or transcription frequency. PLoS One 4**:** e4886.

Lifton, R. P., M. L. Goldberg, R. W. Karp and D. S. Hogness, 1978 The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. Cold Spring Harb Symp Quant Biol 42 Pt 2**:** 1047-1051.

Lin, J. J., L. W. Lehmann, G. Bonora, R. Sridharan, A. A. Vashisht *et al.*, 2011 Mediator coordinates PIC assembly with recruitment of CHD1. Genes Dev 25**:** 2198-2209.

Liu, C. L., T. Kaplan, M. Kim, S. Buratowski, S. L. Schreiber *et al.*, 2005 Single-nucleosome mapping of histone modifications in S. cerevisiae. PLoS Biol 3**:** e328.

Liu, X., M. Li, X. Xia, X. Li and Z. Chen, 2017 Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. Nature 544**:** 440-445.

Liu, Y., L. Warfield, C. Zhang, J. Luo, J. Allen *et al.*, 2009 Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. Mol Cell Biol 29**:** 4852-4863.

Love, M. I., W. Huber and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15**:** 550.

Loya, T. J., T. W. O'Rourke and D. Reines, 2012 A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3. Nucleic Acids Res 40**:** 7476-7491.

Loya, T. J., T. W. O'Rourke and D. Reines, 2013 Yeast Nab3 protein contains a self-assembly domain found in human heterogeneous nuclear ribonucleoprotein-C (hnRNP-C) that is necessary for transcription termination. J Biol Chem 288**:** 2111-2117.

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

Luger, K., and T. J. Richmond, 1998 DNA binding within the nucleosome core. Curr Opin Struct Biol 8**:** 33-40.

Luo, W., A. W. Johnson and D. L. Bentley, 2006 The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. Genes Dev 20**:** 954-965.

Luse, D. S., 2012 Rethinking the role of TFIIF in transcript initiation by RNA polymerase II. Transcription 3**:** 156-159.

Maher, C. A., C. Kumar-Sinha, X. Cao, S. Kalyana-Sundaram, B. Han *et al.*, 2009 Transcriptome sequencing to detect gene fusions in cancer. Nature 458**:** 97-101.

Malik, I., C. Qiu, T. Snavely and C. D. Kaplan, 2017 Wide-ranging and unexpected consequences of altered Pol II catalytic activity in vivo. Nucleic Acids Res 45**:** 4431-4451.

Maltby, V. E., B. J. Martin, J. M. Schulze, I. Johnson, T. Hentrich *et al.*, 2012 Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. Mol Cell Biol 32**:** 3479-3485.

Marcus, G. A., J. Horiuchi, N. Silverman and L. Guarente, 1996 ADA5/SPT20 links the ADA and SPT genes, which are involved in yeast transcription. Mol Cell Biol 16**:** 3197-3205.

Marfella, C. G., and A. N. Imbalzano, 2007 The Chd family of chromatin remodelers. Mutat Res 618**:** 30-40.

Martens, J. A., L. Laprade and F. Winston, 2004 Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature 429**:** 571-574.

Martens, J. A., P. Y. Wu and F. Winston, 2005 Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. Genes Dev 19**:** 2695-2704.

Martin, B. J., K. L. McBurney, V. E. Maltby, K. N. Jensen, J. Brind'Amour *et al.*, 2017 Histone H3K4 and H3K36 Methylation Independently Recruit the NuA3 Histone Acetyltransferase in Saccharomyces cerevisiae. Genetics 205**:** 1113-1123.

Martin, B. J. E., A. T. Chruscicki and L. J. Howe, 2018 Transcription Promotes the Interaction of the FAcilitates Chromatin Transactions (FACT) Complex with Nucleosomes in Saccharomyces cerevisiae. Genetics 210**:** 869-881.

Martin, D. G., D. E. Grimes, K. Baetz and L. Howe, 2006 Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. Mol Cell Biol 26**:** 3018-3028.

Martin, M., 2010 Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnet.journal 17.

Marzluff, W. F., P. Gongidi, K. R. Woods, J. Jin and L. J. Maltais, 2002 The human and mouse replication-dependent histone genes. Genomics 80**:** 487-498.

Mason, P. B., and K. Struhl, 2003 The FACT Complex Travels with Elongating RNA Polymerase II and Is Important for the Fidelity of Transcriptional Initiation In Vivo. Molecular and Cellular Biology 23**:** 8323-8333.

Mattiroli, F., M. Uckelmann, D. D. Sahtoe, W. J. van Dijk and T. K. Sixma, 2014 The nucleosome acidic patch plays a critical role in RNF168-dependent ubiquitination of histone H2A. Nat Commun 5**:** 3291.

McCullough, L. L., T. H. Pham, T. J. Parnell, Z. Connell, M. B. Chandrasekharan *et al.*, 2019 Establishment and Maintenance of Chromatin Architecture Are Promoted Independently of Transcription by the Histone Chaperone FACT and H3-K56 Acetylation in Saccharomyces cerevisiae. Genetics 211**:** 877-892.

McGinty, R. K., R. C. Henrici and S. Tan, 2014 Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. Nature 514**:** 591-596.

Michel, A. H., R. Hatakeyama, P. Kimmig, M. Arter, M. Peter *et al.*, 2017 Functional mapping of yeast genomes by saturated transposition. Elife 6.

Minsky, N., E. Shema, Y. Field, M. Schuster, E. Segal *et al.*, 2008 Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. Nat Cell Biol 10**:** 483-488.

Mitchell, L., J. P. Lambert, M. Gerdes, A. S. Al-Madhoun, I. S. Skerjanc *et al.*, 2008 Functional dissection of the NuA4 histone acetyltransferase reveals its role as a genetic hub and that Eaf1 is essential for complex integrity. Mol Cell Biol 28**:** 2244-2256.

Morgan, M. T., M. Haj-Yahya, A. E. Ringel, P. Bandi, A. Brik *et al.*, 2016 Structural basis for histone H2B deubiquitination by the SAGA DUB module. Science 351**:** 725-728.

Morillon, A., N. Karabetsou, J. O'Sullivan, N. Kent, N. Proudfoot *et al.*, 2003 Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. Cell 115**:** 425-435.

Mosley, A. L., S. G. Pattenden, M. Carey, S. Venkatesh, J. M. Gilmore *et al.*, 2009 Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. Mol Cell 34**:** 168-178.

Murray, S. C., A. Serra Barros, D. A. Brown, P. Dudek, J. Ayling *et al.*, 2012 A pre-initiation complex at the 3'-end of genes drives antisense transcription independent of divergent sense transcription. Nucleic Acids Res 40**:** 2432-2444.

Nakanishi, S., B. W. Sanderson, K. M. Delventhal, W. D. Bradford, K. Staehling-Hampton *et al.*, 2008 A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. Nat Struct Mol Biol 15**:** 881-888.

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

Nemec, C. M., F. Yang, J. M. Gilmore, C. Hintermair, Y. H. Ho *et al.*, 2017 Different phosphoisoforms of RNA polymerase II engage the Rtt103 termination factor in a structurally analogous manner. Proc Natl Acad Sci U S A 114**:** E3944-E3953.

Nemeroff, M. E., S. M. Barabino, Y. Li, W. Keller and R. M. Krug, 1998 Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol Cell 1**:** 991-1000.

Nevers, A., A. Doyen, C. Malabat, B. Neron, T. Kergrohen *et al.*, 2018 Antisense transcriptional interference mediates condition-specific gene repression in budding yeast. Nucleic Acids Res 46**:** 6009-6025.

Ng, H. H., S. Dole and K. Struhl, 2003a The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. J Biol Chem 278**:** 33625-33628.

Ng, H. H., F. Robert, R. A. Young and K. Struhl, 2003b Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell 11**:** 709-719.

Ngo, T. T., Q. Zhang, R. Zhou, J. G. Yodh and T. Ha, 2015 Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. Cell 160**:** 1135-1144.

Ni, Z., B. E. Schwartz, J. Werner, J. R. Suarez and J. T. Lis, 2004 Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. Mol Cell 13**:** 55-65.

Niimi, A., A. L. Chambers, J. A. Downs and A. R. Lehmann, 2012 A role for chromatin remodellers in replication of damaged DNA. Nucleic Acids Res 40**:** 7393-7403.

Nilsen, T. W., 2003 The spliceosome: the most complex macromolecular machine in the cell? Bioessays 25**:** 1147-1149.

Nocetti, N., and I. Whitehouse, 2016 Nucleosome repositioning underlies dynamic gene expression. Genes Dev 30**:** 660-672.

North, J. A., J. C. Shimko, S. Javaid, A. M. Mooney, M. A. Shoffner *et al.*, 2012 Regulation of the nucleosome unwrapping rate controls DNA accessibility. Nucleic Acids Res 40**:** 10215-10227.

Nudler, E., 2012 RNA polymerase backtracking in gene regulation and genome instability. Cell 149**:** 1438-1445.

Nurse, N. P., I. Jimenez-Useche, I. T. Smith and C. Yuan, 2013 Clipping of flexible tails of histones H3 and H4 affects the structure and dynamics of the nucleosome. Biophys J 104**:** 1081-1088.

Omer, A. D., S. Ziesche, H. Ebhardt and P. P. Dennis, 2002 In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. Proc Natl Acad Sci U S A 99**:** 5289-5294.

Orozco, I. J., S. J. Kim and H. G. Martinson, 2002 The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause in vivo and then to release stochastically from the template. J Biol Chem 277**:** 42899-42911.

Pan, G., and J. Greenblatt, 1994 Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. J Biol Chem 269**:** 30101-30104.

Pardon, J. F., M. H. Wilkins and B. M. Richards, 1967 Super-helical model for nucleohistone. Nature 215**:** 508-509.

Patel, A. B., R. K. Louder, B. J. Greber, S. Grunberg, J. Luo *et al.*, 2018 Structure of human TFIID and mechanism of TBP loading onto promoter DNA. Science 362.

Pavri, R., B. Zhu, G. Li, P. Trojer, S. Mandal *et al.*, 2006 Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell 125**:** 703-717.

Pearson, E., and C. Moore, 2014 The evolutionarily conserved Pol II flap loop contributes to proper transcription termination on short yeast genes. Cell Rep 9**:** 821-828.

Pertea, M., D. Kim, G. M. Pertea, J. T. Leek and S. L. Salzberg, 2016 Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 11**:** 1650-1667.

Piekna-Przybylska, D., W. A. Decatur and M. J. Fournier, 2007 New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. RNA 13**:** 305-312.

Pinto, I., D. E. Ware and M. Hampsey, 1992 The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. Cell 68**:** 977-988.

Pokholok, D. K., C. T. Harbison, S. Levine, M. Cole, N. M. Hannett *et al.*, 2005 Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122**:** 517-527.

Polach, K. J., and J. Widom, 1995 Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. J Mol Biol 254**:** 130-149.

Polach, K. J., and J. Widom, 1996 A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites. J Mol Biol 258**:** 800-812.

Porrua, O., F. Hobor, J. Boulay, K. Kubicek, Y. D'Aubenton-Carafa *et al.*, 2012 In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. EMBO J 31**:** 3935-3948.

Porrua, O., and D. Libri, 2013 A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. Nat Struct Mol Biol 20**:** 884-891.

Porrua, O., and D. Libri, 2015 Transcription termination and the control of the transcriptome: why, where and how to stop. Nat Rev Mol Cell Biol 16**:** 190-202.

Prelich, G., and F. Winston, 1993 Mutations that suppress the deletion of an upstream activating sequence in yeast: involvement of a protein kinase and histone H3 in repressing transcription in vivo. Genetics 135**:** 665-676.

Proudfoot, N. J., and G. G. Brownlee, 1976 3' non-coding region sequences in eukaryotic messenger RNA. Nature 263**:** 211-214.

Pruneski, J. A., S. J. Hainer, K. O. Petrov and J. A. Martens, 2011 The Paf1 complex represses SER3 transcription in Saccharomyces cerevisiae by facilitating intergenic transcription-dependent nucleosome occupancy of the SER3 promoter. Eukaryot Cell 10**:** 1283-1294.

Qiu, H., R. V. Chereji, C. Hu, H. A. Cole, Y. Rawal *et al.*, 2016 Genome-wide cooperation by HAT Gcn5, remodeler SWI/SNF, and chaperone Ydj1 in promoter nucleosome eviction and transcriptional activation. Genome Res 26**:** 211-225.

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

Radovani, E., M. Cadorin, T. Shams, S. El-Rass, A. R. Karsou *et al.*, 2013 The carboxyl terminus of Rtt109 functions in chaperone control of histone acetylation. Eukaryot Cell 12**:** 654-664.

Ramirez, F., F. Dundar, S. Diehl, B. A. Gruning and T. Manke, 2014 deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42**:** W187-191.

Ramirez, F., D. P. Ryan, B. Gruning, V. Bhardwaj, F. Kilpert *et al.*, 2016 deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44**:** W160-165.

Rando, O. J., and F. Winston, 2012 Chromatin and transcription in yeast. Genetics 190**:** 351-387.

Rawal, Y., R. V. Chereji, H. Qiu, S. Ananthakrishnan, C. K. Govind *et al.*, 2018 SWI/SNF and RSC cooperate to reposition and evict promoter nucleosomes at highly expressed genes in yeast. Genes Dev 32**:** 695-710.

Recht, J., T. Tsubota, J. C. Tanny, R. L. Diaz, J. M. Berger *et al.*, 2006 Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. Proc Natl Acad Sci U S A 103**:** 6988-6993.

Rhee, H. S., and B. F. Pugh, 2012 Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature 483**:** 295-301.

Richards, B. M., and J. F. Pardon, 1970 The molecular structure of nucleohistone (DNH). Exp Cell Res 62**:** 184-196.

Roberts, S. M., and F. Winston, 1996 SPT20/ADA5 encodes a novel protein functionally related to the TATA-binding protein and important for transcription in Saccharomyces cerevisiae. Mol Cell Biol 16**:** 3206-3213.

Roberts, S. M., and F. Winston, 1997 Essential functional interactions of SAGA, a Saccharomyces cerevisiae complex of Spt, Ada, and Gcn5 proteins, with the Snf/Swi and Srb/mediator complexes. Genetics 147**:** 451-465.

Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander *et al.*, 2011 Integrative genomics viewer. Nat Biotechnol 29**:** 24-26.

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

Rondon, A. G., H. E. Mischo, J. Kawauchi and N. J. Proudfoot, 2009 Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol Cell 36**:** 88-98.

Roy, K., J. Gabunilas, A. Gillespie, D. Ngo and G. F. Chanfreau, 2016 Common genomic elements promote transcriptional and DNA replication roadblocks. Genome Res 26**:** 1363-1375.

Rutkowski, A. J., F. Erhard, A. L'Hernault, T. Bonfert, M. Schilhabel *et al.*, 2015 Widespread disruption of host transcription termination in HSV-1 infection. Nat Commun 6**:** 7126.

Saba, J., X. Y. Chua, T. V. Mishanina, D. Nayak, T. A. Windgassen *et al.*, 2019 The elemental mechanism of transcriptional pausing. Elife 8.

Sainsbury, S., C. Bernecky and P. Cramer, 2015 Structural basis of transcription initiation by RNA polymerase II. Nat Rev Mol Cell Biol 16**:** 129-143.

Santos-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein *et al.*, 2002 Active genes are tri-methylated at K4 of histone H3. Nature 419**:** 407-411.

Sapountzi, V., I. R. Logan and C. N. Robson, 2006 Cellular functions of TIP60. Int J Biochem Cell Biol 38**:** 1496-1509.

Schlesinger, M. B., and T. Formosa, 2000 POB3 is required for both transcription and replication in the yeast Saccharomyces cerevisiae. Genetics 155**:** 1593-1606.

Schneider, B. L., W. Seufert, B. Steiner, Q. H. Yang and A. B. Futcher, 1995 Use of polymerase chain reaction epitope tagging for protein tagging in Saccharomyces cerevisiae. Yeast 11**:** 1265-1274.

Schulz, D., B. Schwalb, A. Kiesel, C. Baejen, P. Torkler *et al.*, 2013 Transcriptome surveillance by selective termination of noncoding RNA synthesis. Cell 155**:** 1075-1087.

Schwabish, M. A., and K. Struhl, 2006 Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. Mol Cell 22**:** 415-422.

Schwanbeck, R., H. Xiao and C. Wu, 2004 Spatial contacts and nucleosome step movements induced by the NURF chromatin remodeling complex. J Biol Chem 279**:** 39933-39941.

Sdano, M. A., J. M. Fulcher, S. Palani, M. B. Chandrasekharan, T. J. Parnell *et al.*, 2017 A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. Elife 6.

Sermwittayawong, D., and S. Tan, 2006 SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. EMBO J 25**:** 3791-3800.

Shandilya, J., and S. G. Roberts, 2012 The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim Biophys Acta 1819**:** 391-400.

Sharp, J. A., G. Rizki and P. D. Kaufman, 2005 Regulation of histone deposition proteins Asf1/Hir1 by multiple DNA damage checkpoint kinases in Saccharomyces cerevisiae. Genetics 171**:** 885-899.

Sheldon, K. E., D. M. Mauger and K. M. Arndt, 2005 A Requirement for the Saccharomyces cerevisiae Paf1 complex in snoRNA 3' end formation. Mol Cell 20**:** 225-236.

Shermoen, A. W., and P. H. O'Farrell, 1991 Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. Cell 67**:** 303-310.

Shirra, M. K., S. E. Rogers, D. E. Alexander and K. M. Arndt, 2005 The Snf1 protein kinase and Sit4 protein phosphatase have opposing functions in regulating TATA-binding protein association with the Saccharomyces cerevisiae INO1 promoter. Genetics 169**:** 1957-1972.

Sikorski, R. S., and J. D. Boeke, 1991 In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. Methods Enzymol 194**:** 302-318.

Simchen, G., F. Winston, C. A. Styles and G. R. Fink, 1984 Ty-mediated gene expression of the LYS2 and HIS4 genes of Saccharomyces cerevisiae is controlled by the same SPT genes. Proc Natl Acad Sci U S A 81**:** 2431-2434.

Simic, R., D. L. Lindstrom, H. G. Tran, K. L. Roinick, P. J. Costa *et al.*, 2003 Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. EMBO J 22**:** 1846-1856.

Simonsen, C. C., and A. D. Levinson, 1983 Analysis of processing and polyadenylation signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. Mol Cell Biol 3**:** 2250-2258.

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

Smolle, M., S. Venkatesh, M. M. Gogol, H. Li, Y. Zhang *et al.*, 2012 Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. Nat Struct Mol Biol 19**:** 884-892.

Squatrito, M., C. Gorrini and B. Amati, 2006 Tip60 in DNA damage response and growth control: many tricks in one HAT. Trends Cell Biol 16**:** 433-442.

Steinmetz, E. J., and D. A. Brow, 1996 Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. Mol Cell Biol 16**:** 6993-7003.

Steinmetz, E. J., C. L. Warren, J. N. Kuehner, B. Panbehi, A. Z. Ansari *et al.*, 2006 Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell 24**:** 735-746.

Sterner, D. E., X. Wang, M. H. Bloom, G. M. Simon and S. L. Berger, 2002 The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex. J Biol Chem 277**:** 8178-8186.

Strahl, B. D., P. A. Grant, S. D. Briggs, Z. W. Sun, J. R. Bone *et al.*, 2002 Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol Cell Biol 22**:** 1298-1306.

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

Strohner, R., M. Wachsmuth, K. Dachauer, J. Mazurkiewicz, J. Hochstatter *et al.*, 2005 A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling. Nat Struct Mol Biol 12**:** 683-690.

Stuwe, T., M. Hothorn, E. Lejeune, V. Rybin, M. Bortfeld *et al.*, 2008 The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module. Proc Natl Acad Sci U S A 105**:** 8884-8889.

Su, D., Q. Hu, Q. Li, J. R. Thompson, G. Cui *et al.*, 2012 Structural basis for recognition of H3K56-acetylated histone H3-H4 by the chaperone Rtt106. Nature 483**:** 104-107.

Suh, M. H., P. A. Meyer, M. Gu, P. Ye, M. Zhang *et al.*, 2010 A dual interface determines the recognition of RNA polymerase II by RNA capping enzyme. J Biol Chem 285**:** 34027-34038.

Sun, J., M. Paduch, S. A. Kim, R. M. Kramer, A. F. Barrios *et al.*, 2018 Structural basis for activation of SAGA histone acetyltransferase Gcn5 by partner subunit Ada2. Proc Natl Acad Sci U S A 115**:** 10010-10015.

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

Swanson, M. S., E. A. Malone and F. Winston, 1991 SPT5, an essential gene important for normal transcription in Saccharomyces cerevisiae, encodes an acidic nuclear protein with a carboxy-terminal repeat. Mol Cell Biol 11**:** 4286.

Swygert, S. G., and C. L. Peterson, 2014 Chromatin dynamics: interplay between remodeling enzymes and histone modifications. Biochim Biophys Acta 1839**:** 728-736.

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

Takagaki, Y., L. C. Ryner and J. L. Manley, 1989 Four factors are required for 3'-end cleavage of pre-mRNAs. Genes Dev 3**:** 1711-1724.

Tanny, J. C., H. Erdjument-Bromage, P. Tempst and C. D. Allis, 2007 Ubiquitylation of histone H2B controls RNA polymerase II transcription elongation independently of histone H3 methylation. Genes Dev 21**:** 835-847.

Terns, M. P., and R. M. Terns, 2002 Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. Gene Expr 10**:** 17-39.

Terzi, N., L. S. Churchman, L. Vasiljeva, J. Weissman and S. Buratowski, 2011 H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. Mol Cell Biol 31**:** 3569-3583.

Teves, S. S., C. M. Weber and S. Henikoff, 2014 Transcribing through the nucleosome. Trends Biochem Sci 39**:** 577-586.

Thoma, F., and T. Koller, 1977 Influence of histone H1 on chromatin structure. Cell 12**:** 101-107.

Thoma, F., T. Koller and A. Klug, 1979 Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol 83**:** 403-427.

Thompson, D. M., and R. Parker, 2007 Cytoplasmic decay of intergenic transcripts in Saccharomyces cerevisiae. Mol Cell Biol 27**:** 92-101.

Thorvaldsdottir, H., J. T. Robinson and J. P. Mesirov, 2013 Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14**:** 178-192.

Tirosh, I., N. Sigal and N. Barkai, 2010 Widespread remodeling of mid-coding sequence nucleosomes by Isw1. Genome Biol 11**:** R49.

Tisseur, M., M. Kwapisz and A. Morillon, 2011 Pervasive transcription - Lessons from yeast. Biochimie 93**:** 1889-1896.

Tomson, B. N., and K. M. Arndt, 2013 The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. Biochim Biophys Acta 1829**:** 116-126.

Tomson, B. N., E. M. Crisucci, L. E. Heisler, M. Gebbia, C. Nislow *et al.*, 2013 Effects of the Paf1 complex and histone modifications on snoRNA 3'-end formation reveal broad and locus-specific regulation. Mol Cell Biol 33**:** 170-182.

Tomson, B. N., C. P. Davis, M. H. Warner and K. M. Arndt, 2011 Identification of a role for histone H2B ubiquitylation in noncoding RNA 3'-end formation through mutational analysis of Rtf1 in Saccharomyces cerevisiae. Genetics 188**:** 273-289.

Toulokhonov, I., J. Zhang, M. Palangat and R. Landick, 2007 A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. Mol Cell 27**:** 406-419.

Tudek, A., O. Porrua, T. Kabzinski, M. Lidschreiber, K. Kubicek *et al.*, 2014 Molecular basis for coordinating transcription termination with noncoding RNA degradation. Mol Cell 55**:** 467-481.

Ursic, D., K. Chinchilla, J. S. Finkel and M. R. Culbertson, 2004 Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. Nucleic Acids Res 32**:** 2441-2452.

Ursic, D., K. L. Himmel, K. A. Gurley, F. Webb and M. R. Culbertson, 1997 The yeast SEN1 gene is required for the processing of diverse RNA classes. Nucleic Acids Res 25**:** 4778-4785.

Valadkhan, S., 2005 snRNAs as the catalysts of pre-mRNA splicing. Curr Opin Chem Biol 9**:** 603-608.

van Dijk, E. L., C. L. Chen, Y. d'Aubenton-Carafa, S. Gourvennec, M. Kwapisz *et al.*, 2011 XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature 475**:** 114-117.

van Hoof, A., P. Lennertz and R. Parker, 2000 Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol Cell Biol 20**:** 441-452.

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

Van Oss, S. B., C. E. Cucinotta and K. M. Arndt, 2017 Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. Trends Biochem Sci 42**:** 788-798.

Van Oss, S. B., M. K. Shirra, A. R. Bataille, A. D. Wier, K. Yen *et al.*, 2016 The Histone Modification Domain of Paf1 Complex Subunit Rtf1 Directly Stimulates H2B Ubiquitylation through an Interaction with Rad6. Mol Cell 64**:** 815-825.

Vanacova, S., J. Wolf, G. Martin, D. Blank, S. Dettwiler *et al.*, 2005 A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol 3**:** e189.

VanDemark, A. P., M. Blanksma, E. Ferris, A. Heroux, C. P. Hill *et al.*, 2006 The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. Mol Cell 22**:** 363-374.

Varley, K. E., J. Gertz, B. S. Roberts, N. S. Davis, K. M. Bowling *et al.*, 2014 Recurrent read-through fusion transcripts in breast cancer. Breast Cancer Res Treat 146**:** 287-297.

Vary, J. C., Jr., V. K. Gangaraju, J. Qin, C. C. Landel, C. Kooperberg *et al.*, 2003 Yeast Isw1p forms two separable complexes in vivo. Mol Cell Biol 23**:** 80-91.

Vasiljeva, L., M. Kim, H. Mutschler, S. Buratowski and A. Meinhart, 2008 The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol 15**:** 795-804.

Venkatesh, S., H. Li, M. M. Gogol and J. L. Workman, 2016 Selective suppression of antisense transcription by Set2-mediated H3K36 methylation. Nat Commun 7**:** 13610.

Venkatesh, S., M. Smolle, H. Li, M. M. Gogol, M. Saint *et al.*, 2012 Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. Nature 489**:** 452-455.

Venkatesh, S., and J. L. Workman, 2013 Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development. Wiley Interdiscip Rev Dev Biol 2**:** 685-700.

Vilborg, A., M. C. Passarelli, T. A. Yario, K. T. Tycowski and J. A. Steitz, 2015 Widespread Inducible Transcription Downstream of Human Genes. Mol Cell 59**:** 449-461.

Voichek, Y., R. Bar-Ziv and N. Barkai, 2016 A role for Rtt109 in buffering gene-dosage imbalance during DNA replication. Nucleus 7**:** 375-381.

Vos, S. M., L. Farnung, M. Boehning, C. Wigge, A. Linden *et al.*, 2018 Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature 560**:** 607-612.

Wal, M., and B. F. Pugh, 2012 Genome-wide mapping of nucleosome positions in yeast using high-resolution MNase ChIP-Seq. Methods Enzymol 513**:** 233-250.

Wang, B., A. V. Predeus, Z. F. Burton and M. Feig, 2013 Energetic and structural details of the trigger-loop closing transition in RNA polymerase II. Biophys J 105**:** 767-775.

Wang, D., D. A. Bushnell, K. D. Westover, C. D. Kaplan and R. D. Kornberg, 2006 Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell 127**:** 941-954.

Wang, X., G. O. Bryant, M. Floer, D. Spagna and M. Ptashne, 2011 An effect of DNA sequence on nucleosome occupancy and removal. Nat Struct Mol Biol 18**:** 507-509.

Wang, Y., and S. G. Roberts, 2010 New insights into the role of TFIIB in transcription initiation. Transcription 1**:** 126-129.

Wang, Z., C. Zang, K. Cui, D. E. Schones, A. Barski *et al.*, 2009 Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell 138**:** 1019-1031.

Warfield, L., S. Ramachandran, T. Baptista, D. Devys, L. Tora *et al.*, 2017 Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIID. Mol Cell 68**:** 118-129 e115.

Webb, S., R. D. Hector, G. Kudla and S. Granneman, 2014 PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol 15**:** R8.

Weil, C. F., and R. Kunze, 2000 Transposition of maize Ac/Ds transposable elements in the yeast Saccharomyces cerevisiae. Nat Genet 26**:** 187-190.

West, S., N. Gromak and N. J. Proudfoot, 2004 Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature 432**:** 522-525.

Whitehouse, I., O. J. Rando, J. Delrow and T. Tsukiyama, 2007 Chromatin remodelling at promoters suppresses antisense transcription. Nature 450**:** 1031-1035.

Williams, S. K., D. Truong and J. K. Tyler, 2008 Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. Proc Natl Acad Sci U S A 105**:** 9000-9005.

Wilson, S. M., K. V. Datar, M. R. Paddy, J. R. Swedlow and M. S. Swanson, 1994 Characterization of nuclear polyadenylated RNA-binding proteins in Saccharomyces cerevisiae. J Cell Biol 127**:** 1173-1184.

Winey, M., and M. R. Culbertson, 1988 Mutations affecting the tRNA-splicing endonuclease activity of Saccharomyces cerevisiae. Genetics 118**:** 609-617.

Winston, F., D. T. Chaleff, B. Valent and G. R. Fink, 1984 Mutations affecting Ty-mediated expression of the HIS4 gene of Saccharomyces cerevisiae. Genetics 107**:** 179-197.

Winston, F., C. Dollard, E. A. Malone, J. Clare, J. G. Kapakos *et al.*, 1987 Three genes are required for trans-activation of Ty transcription in yeast. Genetics 115**:** 649-656.

Winston, F., C. Dollard and S. L. Ricupero-Hovasse, 1995 Construction of a set of convenient Saccharomyces cerevisiae strains that are isogenic to S288C. Yeast 11**:** 53-55.

Wittmeyer, J., and T. Formosa, 1997 The Saccharomyces cerevisiae DNA polymerase alpha catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein. Mol Cell Biol 17**:** 4178-4190.

Wong, K. H., Y. Jin and K. Struhl, 2014 TFIIH phosphorylation of the Pol II CTD stimulates mediator dissociation from the preinitiation complex and promoter escape. Mol Cell 54**:** 601-612.

Wood, A., N. J. Krogan, J. Dover, J. Schneider, J. Heidt *et al.*, 2003a Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell 11**:** 267-274.

Wood, A., J. Schneider, J. Dover, M. Johnston and A. Shilatifard, 2003b The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem 278**:** 34739-34742.

Wu, L., L. Li, B. Zhou, Z. Qin and Y. Dou, 2014 H2B ubiquitylation promotes RNA Pol II processivity via PAF1 and pTEFb. Mol Cell 54**:** 920-931.

Wyce, A., T. Xiao, K. A. Whelan, C. Kosman, W. Walter *et al.*, 2007 H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. Mol Cell 27**:** 275-288.

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

Wysocka, J., T. Swigut, H. Xiao, T. A. Milne, S. Y. Kwon *et al.*, 2006 A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 442**:** 86-90.

Xiao, T., H. Hall, K. O. Kizer, Y. Shibata, M. C. Hall *et al.*, 2003 Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev 17**:** 654-663.

Xiao, T., C. F. Kao, N. J. Krogan, Z. W. Sun, J. F. Greenblatt *et al.*, 2005 Histone H2B ubiquitylation is associated with elongating RNA polymerase II. Mol Cell Biol 25**:** 637-651.

Xie, W., C. Song, N. L. Young, A. S. Sperling, F. Xu *et al.*, 2009 Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. Mol Cell 33**:** 417-427.

Xu, F., K. Zhang and M. Grunstein, 2005 Acetylation in histone H3 globular domain regulates gene expression in yeast. Cell 121**:** 375-385.

Xu, F., Q. Zhang, K. Zhang, W. Xie and M. Grunstein, 2007 Sir2 deacetylates histone H3 lysine 56 to regulate telomeric heterochromatin structure in yeast. Mol Cell 27**:** 890-900.

Xu, Z., W. Wei, J. Gagneur, F. Perocchi, S. Clauder-Munster *et al.*, 2009 Bidirectional promoters generate pervasive transcription in yeast. Nature 457**:** 1033-1037.

Yan, C., J. Hang, R. Wan, M. Huang, C. C. Wong *et al.*, 2015 Structure of a yeast spliceosome at 3.6-angstrom resolution. Science 349**:** 1182-1191.

Yan, C., R. Wan, R. Bai, G. Huang and Y. Shi, 2016 Structure of a yeast activated spliceosome at 3.5 A resolution. Science 353**:** 904-911.

Yoh, S. M., H. Cho, L. Pickle, R. M. Evans and K. A. Jones, 2007 The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. Genes Dev 21**:** 160-174.

Youdell, M. L., K. O. Kizer, E. Kisseleva-Romanova, S. M. Fuchs, E. Duro *et al.*, 2008 Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. Mol Cell Biol 28**:** 4915-4926.

Zhang, J., M. Palangat and R. Landick, 2010 Role of the RNA polymerase trigger loop in catalysis and pausing. Nat Struct Mol Biol 17**:** 99-104.

Zhang, L., A. Serra-Cardona, H. Zhou, M. Wang, N. Yang *et al.*, 2018 Multisite Substrate Recognition in Asf1-Dependent Acetylation of Histone H3 K56 by Rtt109. Cell 174**:** 818-830 e811.

Zhang, Z., C. J. Wippo, M. Wal, E. Ward, P. Korber *et al.*, 2011 A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. Science 332**:** 977-980.

Zhou, K., G. Gaullier and K. Luger, 2019 Nucleosome structure and dynamics are coming of age. Nat Struct Mol Biol 26**:** 3-13.

Zhou, K., W. H. Kuo, J. Fillingham and J. F. Greenblatt, 2009 Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. Proc Natl Acad Sci U S A 106**:** 6956-6961.

Zhu, F., L. Farnung, E. Kaasinen, B. Sahu, Y. Yin *et al.*, 2018 The interaction landscape between transcription factors and the nucleosome. Nature 562**:** 76-81.

Zippo, A., R. Serafini, M. Rocchigiani, S. Pennacchini, A. Krepelova *et al.*, 2009 Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. Cell 138**:** 1122-1136.

Zunder, R. M., A. J. Antczak, J. M. Berger and J. Rine, 2012 Two surfaces on the histone chaperone Rtt106 mediate histone binding, replication, and silencing. Proc Natl Acad Sci U S A 109**:** E144-153.