Transcriptional Regulation by Mot1 and Spt16

Jason Donald True
Center Point, IN

B.S., University of Southern Indiana, 2009
M.S., Indiana University, 2012
M.S., University of Virginia, 2014

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Biochemistry and Molecular Genetics

University of Virginia
August, 2016

_____

_____

_____

_____

Abstract

The *Saccharomyces cerevisiae* ATPase Mot1 regulates transcription by impacting the distribution and activity of TATA-binding protein (TBP). *In vitro*, Mot1 forms a complex with TBP and DNA, and uses ATP hydrolysis to dissociate TBP from DNA. To gain insight into the Mot1 mechanism, we employed a DNA tethered cleavage approach to map regions of Mot1 in proximity to DNA under different conditions. We present evidence for two conformations of the Mot1 ATPase, the detection of which can be modulated by ATP analogs as well as DNA sequence. We also show using purified complexes that Mot1 dissociation of TBP-DNA is inefficient, suggesting how other transcription factors that bind to TBP may compete with Mot1. In addition, Mot1 and the Spt16 component of the FACT histone chaperone complex, have been shown to physically interact. Here we demonstrate that Mot1 and Spt16 regulate a largely overlapping set of genes in *S. cerevisiae* and physically and genetically interact. Mot1 controlled TBP levels at co-regulated promoters, whereas Spt16 did not. Both Mot1 and Spt16 contribute to TFIIB localization, indicating a convergence on preinitiation complex formation. Globally, Spt16 was required for Mot1 promoter localization, and Mot1 also affected Spt16 localization. Interestingly, we found that Mot1 has a role in establishing or maintaining the occupancy and positioning of nucleosomes at the 5' ends of genes. Spt16 has a broad role in regulating chromatin organization, including those nucleosomes affected by Mot1. These results suggest that Spt16 is required for Mot1 localization to the promoter by establishing a permissive chromatin environment for Mot1. Overall these results suggest that the overlap in Mot1 and Spt16 function arises from a combination of their unique and shared functions in transcription complex assembly.

Dedicated to my best friend Katie Schiermeyer.

Publications

The following papers correspond to the material presented in the indicated chapters.

Chapter II:

Viswanathan, R., **True, J.D.**, and Auble, D.T. (2016). Molecular mechanism of Mot1, a TATA-binding protein (TBP)-DNA dissociating enzyme. Journal of Biological Chemistry doi: 10.1074/jbc.M116.730366.

Chapters III and IV:

True, J.D., Muldoon, J.M., Carver, M.N., Poorey, K., Shetty, S.J., Bekiranov, S., and Auble, D.T. (2016). The Mot1 ATPase and Spt16 histone chaperone co-regulate transcription through preinitiation complex assembly and nucleosome organization. Journal of Biological Chemistry doi: 10.1071/jbc.M116.735134.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADI	ATP-dependent inhibitor of TBP
Bio	Biotinylation
bp	Base pair
BSA	Bovine serum albumin
ChIP	Chromatin immunoprecipitation
ChIP-chip	Chromatin immunoprecipitation and microarray
ChIP-Seq	Chromatin immunoprecipitation and deep sequencing
CTD	C-terminal domain
DI	Downstream initiation
DT	Downstream termination
DTT	Dithiothreitol
Ecu	*Encephalitozoon cuniculi*
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
FACT	Facilitates chromatin transcription
FDR	False discovery rate
FeBABE	Fe(III) (s)-1-(p-bromoacetamidobenzyl) ethylenediamine tetraacetic acid
FOA	Fluroorotic acid
FRET	Förster resonance energy transfer
GTF	General transcription factor
HA	Hemagglutinin

HEAT	Huntington-elongation factor 3-protein phosphatase 2A-target of rapamycin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IGB	Integrated genome browser
IGV	Integrated genome viewer
IP	Immunoprecipitate
KCl	Potassium chloride
KOAc	Potassium acetate
KOH	Potassium hydroxide
MNase	Micrococcal nuclease
MOPS	3-(N-morpholino)propanesulfonic acid
Mot1	Modifier of transcription 1
mRNA	Messenger RNA
NC2	Negative cofactor 2
NDR	Nucleosome depleted region
NFR	Nucleosome free region
NTD	N-terminal domain
OD	Optical density
ORF	Open reading frame
ORGANIC	Occupied regions of genomes from affinity-purified naturally isolated chromatin
PAF	Polymerase associated factors
PCR	Polymerase chain reaction
PIC	Preinitiation complex
Pol II	RNA Polymerase II
PT	Premature termination

qPCR	Quantitative polymerase chain reaction
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
SAGA	Spt-Ada-Gcn5-Acetyltransferase
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
snRNA	Small nuclear RNA
Spt16	Suppressor of transposons 16
Sso	*Sulfolobus solfataricus*
TAF	TBP-associated factor
TAS	Tiling analysis software
TBP	TATA-Binding Protein
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
ts	Temperature sensitive
TSS	Transcription start site
UI	Upstream initiation
WCE	Whole cell extract
WT	Wildtype
YPD	Yeast peptone dextrose

Acknowledgements

Graduate school has been a difficult, enjoyable, and illuminating experience. With that said, I need to thank the following people for everything that they have done during my time as a PhD student:

- Dr. David Auble, for taking me into his lab and taking the time to mentor and push me. He always expected the best out of me which really challenged me. He performed the modeling in Chapter II (Figure 2.6) and the data analysis for the ChIP-Seq and MNase ChIP-Seq experiments in Chapter IV (Figures 4.2, 4.3, and 4.9).
- Dr. Jeff Smith, for being a member of my committee, an expert in yeast genetics, and a wonderful teacher. He was always willing to offer advice to help me out.
- Dr. Patrick Grant, for being a member of my committee and an expert in nucleosomes and nucleosome mapping.
- Dr. Lucy Pemberton, for being a member of my committee, a histone chaperone expert, and a great person to talk to.
- Ramya Viswanathan, for helping me out while we were in lab together and even after she left. We could have some great scientific discussions while still laughing. She performed the 5' Bio bead-bound experiments in Figure 2.1 and generated Figure 2.4*A*.

- Kasia Rowny, for sharing a lab bay with me. Though our time together was brief she brought a unique perspective to science that I am glad she shared.
- Liz Hoffman, for putting up with me since she joined lab; all day, every day.
- Joseph Muldoon, for being a wonderful co-first author. Though we argued a lot, he has truly taught me more than I ever expected. He performed the RNA purification and the data analysis for the tiling arrays (Figures 3.5, 3.6, 3.7, and 3.15).
- Savera Shetty, for performing the ChIP-Seq experiments in Chapter IV (Figures 4.2 and 4.3).
- Melissa Carver, for initiating the experiments for Chapters III and IV and performing the original Mot1-Spt16 co-immunoprecipitation experiment (Figure 3.2*A*)
- Kunal Poorey and Dr. Stefan Bekiranov, for helping with the data analysis in Chapters III and IV.
- Shana, Alicia, Amanda, and Kunal, for being great labmates.
- Ryan, Sheena, Chuck, Emily, Brittany, Kaity, and all of my other science friends. We could randomly have great scientific discussions in the halls and then go out and enjoy our time together too.
- Finally, my family, for being there for me even though I was states away from them. I doubt I would have been able to do all of this without their love.

Chapter I

Introduction

An organism's genetic information is contained within DNA which must be compacted and contained within the nucleus of each cell. DNA has to be converted into RNA through the process of transcription (Nikolov and Burley 1997, Hampsey 1998, Hahn 2004). RNA can be further converted into protein through translation (Jackson et al. 2010). The pathway of DNA being transcribed into RNA and translated into proteins is known as the central dogma of molecular biology (Crick 1970). This dogma has been complicated by the discovery that not all RNA is converted into protein and can instead be functional as non-coding RNA (Kung et al. 2013, Cech and Steitz 2014). Regardless of this, the initial conversion of DNA to RNA is an integral point and undergoes many layers of regulation. Some of the complexities of transcription regulation will be the focus of this dissertation.

Transcription is normally thought of in three phases: initiation, elongation, and termination. The transcription machinery must accurately identify a gene's start site to initiate transcription, elongate through the gene body, and finally terminate at the proper site. A model organism widely used in the transcription field is budding yeast, *Saccharomyces cerevisiae*, a single-celled eukaryote with a conserved transcriptional apparatus. In *S. cerevisiae* there are approximately 6,000 genes that all must be accurately transcribed to enable the yeast cell to properly grow and adapt to its environment (Mackiewicz et al. 2002). RNA polymerases I, II, and III are collectively responsible for transcription of all of these genes, but each polymerase is relegated to a specific group of genes. RNA polymerase I is responsible for transcription of ribosomal RNAs (rRNAs), RNA polymerase II (Pol II) transcribes all messenger RNAs (mRNAs)

and some small nuclear RNAs (snRNAs), while RNA polymerase III transcribes all transfer RNAs (tRNAs) (Vannini and Cramer 2012). The mRNAs transcribed by Pol II will be further translated into proteins, and thus transcription by Pol II is a fundamental area of research.

Assembly of the Preinitiation Complex

Before Pol II can initiate transcription, the assembly of a preinitiation complex (PIC) is required at the promoter of a gene (Hahn 2004). But first, what defines a gene promoter? A gene promoter is the area of DNA where transcription initiation starts (Figure 1.1; Butler and Kadonaga 2002). Promoters in yeast are typically divided into two groups based on the presence or absence of a specific DNA sequence called the TATA box. TATA boxes, consisting of the sequence "TATAWAWR" (W = A/T, R = A/G), are only found at 10-20% of genes in yeast (Basehoar et al. 2004), though a recent study has identified TATA-like elements with up to two mismatches from the TATA consensus sequence (Rhee and Pugh 2012). The TATA box is usually located within 40-100 bases upstream of the TSS, and there can be multiple start sites within this region (Struhl 1987, Guarente 1987, Butler and Kadonaga 2002). Higher eukaryotes have more diversity at the promoter. Promoters of higher eukaryotes can contain other elements such as the initiator, the TFIIB recognition element, or the downstream promoter element (Butler and Kadonaga 2002).

Promoters are not only defined by the underlying DNA sequences but also by the specific patterns of chemical modifications on histones. In yeast, promoter-proximal nucleosomes are typically modified with H3K4me3 or acetylation on either histone H3 or

H4 (Li et al. 2007). Acetylation of the histone tails neutralizes the positive charge of the highly basic histones, which is thought to loosen the contacts between the histones and the DNA backbone, thereby allowing for transcription factors to bind the exposed DNA (Eberharter and Becker 2002). Methylation does not affect the charge of the lysines on histones, and is thought to act by recruiting transcription factors with specific binding domains instead of directly repelling DNA (Li et al. 2007). Along with histone modifications, the first nucleosome downstream of the transcription start site (TSS), denoted as the +1 nucleosome, has been found to be enriched for the histone variant H2A.Z (Li et al. 2007). This variant has been shown to have different interactions with DNA when compared to the canonical H2A. This includes loosening the DNA around the octamer, allowing for easier eviction of this nucleosome and enabling transcription to occur.

Now to initiate transcription, the PIC has to identify the promoter and assemble properly. The PIC consists of the general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH and Pol II (Hahn 2004). The GTFs were identified in fractionations from HeLa nuclear extracts to identify the basal elements required for transcription initiation (Matsui et al. 1980, Dignam et al. 1983, Sawadogo and Roeder 1985, Flores et al. 1989, Flores et al. 1992). They were further characterized by *in vitro* studies to understand their assembly order and functions. The canonical pathway of PIC assembly starts with TFIID or more minimally TATA-binding protein (TBP) binding and bending the promoter DNA (Buratowski et al. 1989). TBP is essential for transcription by all three polymerases even at genes that do not contain a TATA box or TATA-like element (Cormack and Struhl 1992, Joazeiro et al. 1996, Comai et al. 1992). Higher eukaryotes even have multiple forms of TBP or TBP-related factors (Gasch et al. 1990,

Crowley et al. 1993, Rabenstein et al. 1999). After TBP binds the promoter DNA, TFIIA and TFIIB are recruited and stabilize TBP by binding either side of the TBP-DNA complex (Figure 1.2). At this point Pol II is recruited along with TFIIE, TFIIF, and TFIIH and is able to initiate transcription. The GTFs excluding TFIIH are shown in Figure 1.2 from a recent cryo-EM structure of the PIC in complex with Pol II (Plaschka et al. 2016).

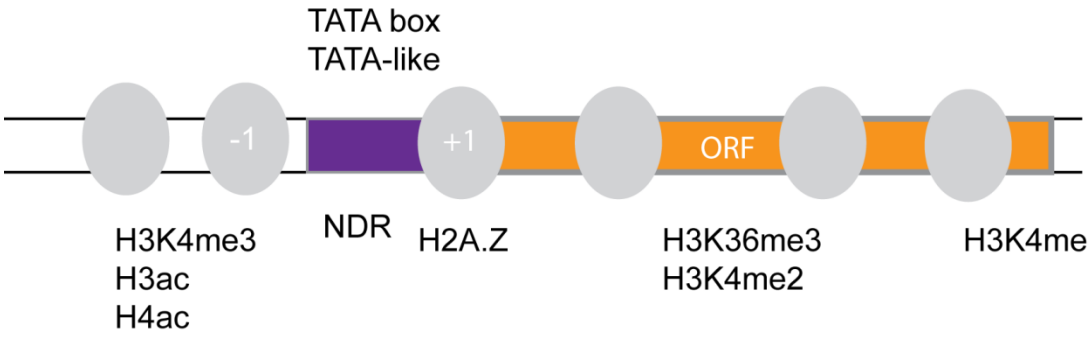

Figure 1.1. Histone modifications and promoter elements associated with an active gene. A cartoon illustration of a "typical" gene is depicted with the promoter (purple box) and open reading frame (ORF; orange box). Nucleosomes are depicted by the gray ovals with the +1 nucleosome specifically labeled. Yeast promoters either contain a TATA box or not, although more recent studies have shown a TATA-like element at previously termed TATA-less promoters. The promoter is typically depleted for nucleosomes and thus termed the nucleosome depleted region (NDR) with the -1 and +1 nucleosomes as boundaries. The promoter-proximal nucleosomes of an active gene are highly acetylated on H3 and H4 and contain H3K4me3. Histones within the ORF have H3K36me3 and H3K4me2, whereas the 3' nucleosomes have H3K4me.

TBP was first crystallized from *Arabidopsis thaliana* as a dimer, secondly from *Saccharomyces cerevisiae* as a monomer, and thirdly from *S. cerevisiae* in complex with DNA (Nikolov et al. 1992, Chasman et al. 1993, Kim et al. 1993). These crystal structures allowed unique insight into how the first component of the PIC actually bound to the promoter DNA. TBP is saddle-shaped and symmetrical and in the absence of DNA can form a dimer. In the presence of DNA, TBP directly binds to the minor groove of the TATA box. TBP has four phenylalanine residues that intercalate between the DNA bases, thus causing two kinks in the DNA and bending the promoter approximately 90° (Horikoshi et al. 1992, Kim et al. 1993). This bending of the DNA is required for the remaining GTFs to assemble along with TBP to form the complete PIC. It is important to note that *in vitro* studies looking at TBP binding to the DNA all use TATA-containing sequences. Analyses of TBP binding to sequences with one mismatch from the TATA sequence have been performed and determined that TBP binding to these sequences was similar to the consensus sequence (Hahn et al. 1989b, Blair et al. 2012). *In vitro* studies have so far been unable to determine the binding affinity of TBP to TATA sequences with more than one mismatch to the consensus, which encompasses the majority of promoters in yeast (Kamenova et al. 2014, Rhee and Pugh 2012, Venters and Pugh 2013).

Both TFIIA and TFIIB bind to TBP-DNA and help to stabilize this interaction by recruitment of the other GTFs and preventing the removal of TBP by other factors(Nikolov et al. 1995, Geiger et al. 1996, Tan et al. 1996), which will be discussed more later. Both domains of TFIIB have been shown to bind to Pol II to help recruit it to the PIC (Pardee et al. 1998, Chen and Hahn 2003). Furthermore, TFIIB is required for proper selection of the start site (Pinto et al. 1992, Berroteran et al. 1994, Li et al. 1994,

Pinto et al. 1994, Bangur et al. 1997). TFIIA specifically enhances interaction of the TAFs with DNA by preventing the interaction of Taf1 with the concave portion of TBP (Weideman et al. 1997, Kokubo et al. 1998, Liu et al. 1998). The interaction between Taf1 and the concave side of TBP would prevent TBP from binding DNA.

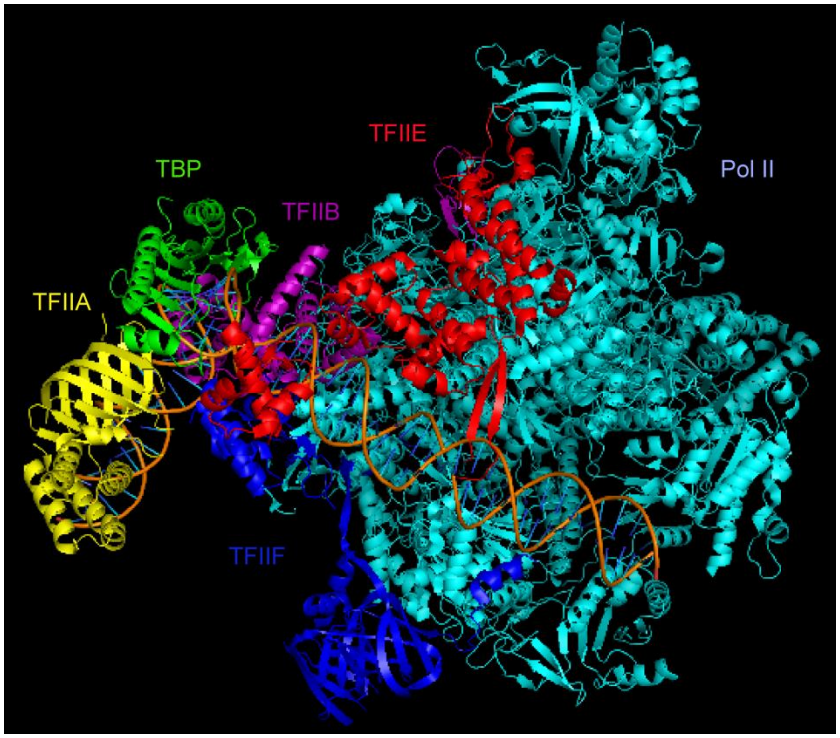

Figure 1.2. Structure of the preinitiation complex. The structures of the general transcription factors (GTFs) that form the preinitiation complex (PIC) are shown (PDB = 5FZ5). TBP (green), TFIIA (yellow), TFIIB (purple), TFIIE (red), TFIIF (blue), and TFIIH (not shown) bind to the promoter DNA (orange and blue) and recruit RNA Polymerase II (Pol II, cyan).

TFIIF and Pol II are subsequently recruited to the TBP-TFIIB-TFIIA-DNA complex. TFIIF interacts with TFIIB, Pol II, and DNA. The interaction with DNA helps to position the DNA in the Pol II active site (Sun and Hampsey 1995, Chung et al. 2003, Forget et al. 2004, Ghazy et al. 2004). TFIIE and TFIIH bind after Pol II and TFIIF have already bound. TFIIE stimulates the activities of TFIIH (Ohkuma et al. 1995, Lee et al. 2000, Okuda et al. 2008). TFIIH has both helicase and kinase activities which are necessary for transcription. The helicase activity of TFIIH helps in promoter melting, the separation of the double stranded DNA to form the transcription bubble (Coin and Egly 1998, Douziech et al. 2000, Takagi et al. 2003). The kinase activity (sometimes referred to as TFIIK) phosphorylates serine 5 of the Pol II CTD, which is required for transcription elongation (Lu et al. 1992).

After transcription has properly initiated, not all of the GTFs dissociate from the promoter. Those remaining behind form the initiation scaffold, which helps for subsequent rounds of reinitiation. Those GTFs that form the scaffold are TFIIA, TFIID, TFIIE, and TFIIH (Yudkovsky et al. 2000). A larger complex known as Mediator is also part of the scaffold. Mediator has several interactions with the GTFs, Pol II, and transcription activators (Kim et al. 1994, Myers and Kornberg 2000, Baek et al. 2006).

Regulation of TBP Localization

As an essential factor required by all three RNA Polymerases, TBP has been widely characterized. TBP can be regulated by multiple factors including TBP associated factors (TAFs), SAGA (Spt-Ada-Gcn5-Acetyltransferase), Mot1, and NC2 (Figure 1.3; Pugh 2000, Auble 2009). Originally, TBP was identified as a large protein

called TFIID, but was further fractionated into TBP and TAFs (Buratowksi et al. 1989, Dynlacht et al. 1991, Pugh and Tijan 1991, Takada et al. 1992). TBP can be recruited and stabilized by these TAFs at sites that lack a TATA box, whereas SAGA recruits TBP to TATA box containing sites (Van Werven et al. 2009). Modifier of transcription 1 (Mot1) and Negative Cofactor 2 (NC2) were both initially characterized as negative regulators of TBP, though further studies have shown that they can have positive effects on TBP recruitment (Davis et al. 1992, Auble et al. 1993, Auble et al. 1994, Meisterernst and Roeder 1991, Muldrow et al. 1999, Collart 1996, Lemaire et al. 2000, Auble 2009, Viswanathan and Auble 2011).

As already mentioned, the original TFIID protein identified in human extracts consisted of TBP and TAFs (Buratowksi et al. 1989, Dynlacht et al. 1991, Pugh and Tijan 1991, Takada et al. 1992). The various TAFs were identified by fractionating extracts from *Drosophila* and human cells (Dynlacht et al. 1991, Pugh and Tijan 1991, Takada et al. 1992). TFIID has been shown to be the predominant form of TBP at TATA-less promoters which are typically the housekeeping genes (Huisinga and Pugh 2004). The TAFs help to stabilize TBP binding to promoters that lack TATA boxes, including through interactions with the Initiator element and the downstream promoter element in *Drosophila* (Verrijizer et al. 1994, Burke and Kadonaga 1997). The various TAFs have been shown to interact with several activators (Goodrich et al. 1993, Hoey et al. 1993, Chiang and Roeder 1995). The Taf1 subunit can also have a repressive effect on transcription by binding to the concave portion of TBP preventing it from binding DNA (Liu et al. 1998). Taf1 has also been shown to compete with TFIIA for binding to TBP (Lieberman and Berk et al. 1994, Kokubo et al. 1998, Ozer et al. 1998, Bagby et al. 2000).

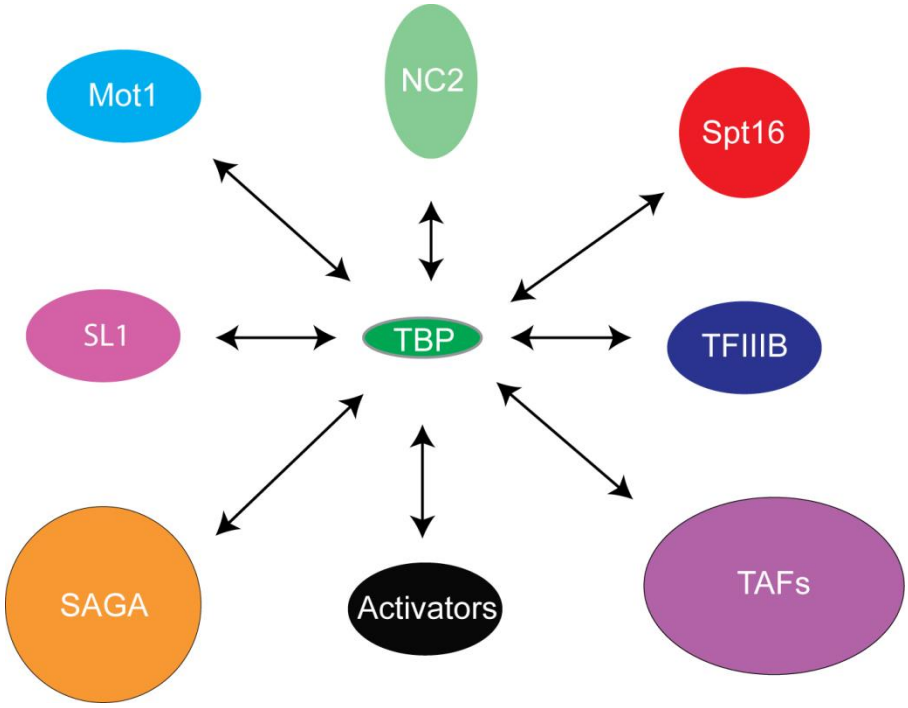

Figure 1.3. Regulators of TBP. TATA-binding protein (TBP) is an integral part of transcription by all three polymerases and thus has many proteins that regulate its localization. SL1 (Pol I), TBP associated factors (TAFs) (Pol II), and TFIIIB (Pol III) are the complexes that TBP forms for transcription by the various polymerases. SAGA localizes TBP to TATA-containing promoters, while TAFs localize TBP to TATA-less promoters. Mot1 and NC2 work cooperatively to remove TBP from TATA-containing promoters and shuttle TBP to TATA-less promoters. Various activators help to localize TBP to promoters including Gal4. Finally, Spt16 indirectly regulates TBP localization by either depositing nucleosomes in the promoter or removing nucleosomes so that TBP can bind.

There is another multi-subunit complex known as SAGA that is important for TBP recruitment to TATA-containing genes (Madison and Winston 1997, Grant et al. 1997,

Bhaumik and Green 2002). These genes are typically stress response genes. The Spt3 subunit of SAGA directly binds TBP and recruits it to promoters (Sterner et al. 1999, Mohibullah and Hahn 2008). The SAGA complex has two enzymatic functions: acetylation and deubiquitination (Roth et al. 2001, Henry et al. 2003). Interestingly, SAGA and TFIID share a subset of TAFs (Grant et al. 1998). The structure of the SAGA complex was solved using electron microscopy (Wu et al. 2004). However, coupling of crosslinking and mass spectrometry depicted a slightly different structural organization of SAGA (Han et al. 2014). These structures were important in understanding how the various subunits fit together to form the whole SAGA complex and perform its various functions (Wu et al. 2004, Han et al. 2014). While initially it was suggested that SAGA only affected a small proportion of genes (i.e. those containing TATA-boxes), there is recent evidence suggesting that SAGA has a wider role in transcriptional regulation (Lee et al. 2000, Bonnet et al. 2014). The HAT activity of SAGA is responsible for histone acetylation at promoters, which is a mark of transcription activation (Robert et al. 2004, Roh et al. 2004, Pokholok et al. 2005, Durant and Pugh 2006, Rosaleny et al. 2007). Furthermore, SAGA deubiquitinates H2B within the coding region of expressed genes (Kohler et al. 2008, Bonnet et al. 2014). Finally the Spt3 subunit was shown to genetically interact with Mot1 through their roles in TBP regulation (Madison and Winston 1997).

In the early fractionations of TAFs, the Negative Cofactor 2 (NC2) complex was first identified (Meisterernst and Roeder 1991). This complex consists of two proteins: NC2α (Bur6) and NC2β (Ydr1/Ncb2). Originally NC2 was found to repress transcription (hence the name) by preventing the binding of TFIIA and TFIIB to TBP (Inostroza et al. 1992, Yeung et al. 1994, Mermelstein et al. 1996, Goppelt et al. 1996). The NC2

complex requires both subunits to bind DNA (Goppelt and Meisterernst 1996), though there is some evidence that the subunits can have different functions (Creton et al. 2002, Van Werven et al. 2008). The *Drosophila* homolog was shown to repress TATA-containing promoters and activate promoters with a downstream promoter element (Willy et al. 2000, Hsu et al. 2008). Footprinting analysis showed that addition of NC2 lessened the protection of DNA by TBP (Schluesche et al. 2007). Furthermore, Förster resonance energy transfer (FRET) analysis showed that NC2 moved TBP on DNA, which suggested that NC2 could act as a TBP shuttle (Schluesche et al. 2007). In conjunction with NC2, Modifier of transcription 1 (Mot1) also regulates TBP localization to the promoter (Auble and Hahn 1993, Auble et al. 1994). The Mot1 ATPase can bind to NC2-TBP-DNA and more will be discussed on Mot1 in the next section (Butryn et al. 2014).

Finally, it was also mentioned that TBP is required by all three RNA Polymerases. While TBP can function as a monomer or in the TFIID complex for Pol II transcription, TBP forms the SL1 complex for RNA Polymerase I transcription and the TFIIIB complex for RNA Polymerase III transcription (Pugh 2000). For Pol II transcription, the GTFs TFIIA and TFIIB strengthen the TBP-DNA interaction which is necessary for PIC formation.

Although the majority of promoters in yeast are TATA-less and therefore TFIID-regulated, TBP has different dynamics than the TAFs as assayed by the Taf1 subunit (Sprouse et al. 2008a). The mobility of TBP was increased by mutation of Taf1, which is consistent with the TFIID complex stabilizing TBP-DNA interactions. TBP's dynamic nature was decreased by mutation of Mot1, which indicates that Mot1 is responsible for the highly mobile property of TBP (Sprouse et al. 2008a). This was specific to Mot1 and

was not shown by mutation of NC2, which was previously shown to shuttle TBP on DNA *in vitro*. TBP dynamics were further explored *in vivo* using crosslinking kinetic analysis (CLK), which showed that mutation of Mot1 decreased the dynamics of TBP (Poorey et al. 2013). This observation was counter-intuitive, since Mot1 removes TBP from DNA. The increase in TBP dynamics was explained by the observation that promoter-bound TBP did not form an active PIC as observed by the decrease in TFIIB (Poorey et al. 2013). Mot1 therefore dissociates TBP bound to weak, spurious, or incorrect locations rather than stably bound TBP incorporated into an active PIC.

The Role of Mot1 in Transcriptional Regulation

Mot1 is a member of the Swi2/Snf2 family of ATPases (Davis et al. 1992, Auble et al. 1994). Initially it was identified genetically as a repressor of transcription in several different yeast genetic screens (Davis et al. 1992, Piatti et al. 1992, Karnitz et al. 1992, Jiang and Stillman 1996, Munn and Reizman 1994, Munn 2000, Voronkova et al. 2006). Biochemically it was originally identified as an ATP-dependent inhibitor of TBP (ADI) and was shown to use ATP hydrolysis to remove TBP from DNA, which was thought to be the mechanism in which Mot1 repressed transcription *in vivo* (Auble and Hahn 1993, Auble et al. 1994). Separately Mot1 was characterized as a TBP-associated factor (Taf170) which interacted with TBP and was independent of the other TAFs that formed TFIID (Poon et al. 1994). It was also shown to be necessary for repression of *LEU2* through its interaction with TBP and the co-repressor Leu3 (Wade and Jaehning 1996). Furthermore, Mot1 was identified in both *Drosophila melanogaster* (helicase 89B) and humans (BTAF1) and has been biochemically characterized in both (Goldman-Levi et al.

1996, Timmers and Sharp 1991, Timmers et al. 1992, Van der Knaap et al. 1997, Chicca et al. 1998). Interestingly, while the ATPase activity of Mot1 was shown to be stimulated by TBP, the ATPase activity of human BTAF1 was only stimulated by TBP-DNA (Auble et al. 1997, Chicca et al. 1998). Adamkewicz et al. 2001). Mot1 was purified as a single polypeptide and was able to function as such (Adamkewicz et al. 2000). This is in contrast to other members of the Swi2/Snf2 ATPase family which are known to function as multi-subunit complex (Eisen et al. 1995, Ryan et al. 2011, Viswanathan and Auble 2011). *MOT1* is an essential gene and the ATPase activity is also required for viability (Davis et al. 1992, Auble et al. 1994).

The NTD was determined to be the portion of Mot1 that interacted with TBP (Figure 1.4; Auble et al. 1997, Darst et al. 2003). This portion of Mot1 contains tetratricopeptide repeats, which were later determined to be a subclass called Huntington-elongation factor 3-protein phosphatase 2A-target of rapamycin (HEAT) repeats (Davis et al. 1992, Wade and Jaehning 1996, Andrade et al. 2000, Darst et al. 2003). While the ATPase activity of Mot1 is required for viability, it is not required for TBP binding (Dasgupta et al. 2002, Adamkewicz et al. 2001). A C-terminal Mot1 deletion mutant exhibited a dominant-negative phenotype most likely due to sequestering TBP through the still intact NTD-TBP interaction (Auble et al. 1994, Auble et al. 1997). To further characterize the Mot1-TBP interaction, mutagenesis of TBP was performed in the temperature sensitive strain *mot1-42* to identify bypass alleles (Sprouse et al. 2008c). As previously mentioned, TBP requires two phenylalanine residues to intercalate into DNA and bend it (Kim et al. 1993). Mutation of one of these residues (F207L) was able to bypass the need for Mot1, suggesting that loosening of the TBP-DNA interaction was responsible for this (Sprouse et al. 2008c). The other mutated

residue (Y185C) is potentially a residue that interacts with TBP and the Negative Cofactor 2 (NC2) complex (Sprouse et al. 2008c).

Crystal structures of Mot1-TBP and Mot1-TBP-NC2-DNA have shed more light on how Mot1 interacts with TBP and DNA (Wollman et al. 2011, Butryn et al. 2015). As already mentioned the N-terminus of Mot1 consists of HEAT repeats that bind to the convex surface of TBP, which was confirmed by both structures. A helical portion of the N-terminus known as the latch can bind to the concave portion of TBP preventing it from binding DNA (Wollman et al. 2011). This interaction with the concave surface of TBP had previously been suggested by a study using BTAF1, which determined that part of the NTD was able to bind to the concave portion of TBP (Pereira et al. 2001).

A recent unpublished crystal structure of the entire Mot1 protein from the Hopfner lab showed the C-terminal Swi2/Snf2 ATPase domain (Butryn and Hopfner unpublished), which was consistent with the previous structure of the Mot1 NTD and the EM structure of the ATPase domain (Wollmann et al. 2011). Previously, the ATPase domain had been shown to come into contact with the upstream DNA and required a 17 base pair handle to remove TBP (Darst et al. 2001, Wollmann et al. 2011). Crystal structures of the related SsoRad54 ATPase domain suggest that the ATPase can adopt two conformations: open, in which there is no ATP-binding pocket; and closed, in which there is an ATP-binding pocket formed (Dürr et al. 2005). These two conformations would require a 180° rotation of the ATPase domain in order to form the ATP-binding pocket. This conformational change had been previously suggested by an experiment that showed that when Mot1-TBP complexes were incubated with an ATP analog ADP-AlF_4, the complex was able to bind to DNA (Darst et al. 2003). This was also suggested

by the observation of two different Mot1-TBP complexes with different kinetics (Sprouse et al. 2008b). More on this will be discussed in Chapter II.

Because *MOT1* is an essential gene, *in vivo* studies utilized temperature-sensitive (ts) mutants to understand the effects of Mot1 (Davis et al. 1992). Locus-specific and genome-wide analyses using various ts alleles of Mot1 showed that Mot1 was necessary for both repression and activation of genes (Timmers et al. 1992, Collart 1996, Madison and Winston 1997, Dasgupta et al. 2002, Geisberg and Struhl 2004, Poorey et al. 2010). Interestingly, the majority of genes that Mot1 represses have a TATA box, while the majority of genes activated by Mot1 lack a canonical TATA box (Collart 1996, Zanton and Pugh 2004, Poorey et al. 2010, Venters et al. 2011). While the biochemical activity of Mot1 suggested a straight-forward mechanism for Mot1 to repress transcription by removing TBP and inhibiting transcription (Auble and Hahn 1992, Auble et al. 1993), the mechanism for Mot1-mediated gene activation was unclear. The results from the *in vivo* analyses indicated that *HIS4* was a Mot1-activated gene, and this was confirmed *in vitro* (Collart 1996, Muldrow et al. 1999). These results showed that when Mot1 was approximately present at a stoichiometric ratio to TBP that it could activate *HIS4* transcription, whereas if Mot1 was present in excess it led to repression (Muldrow et al. 1999).

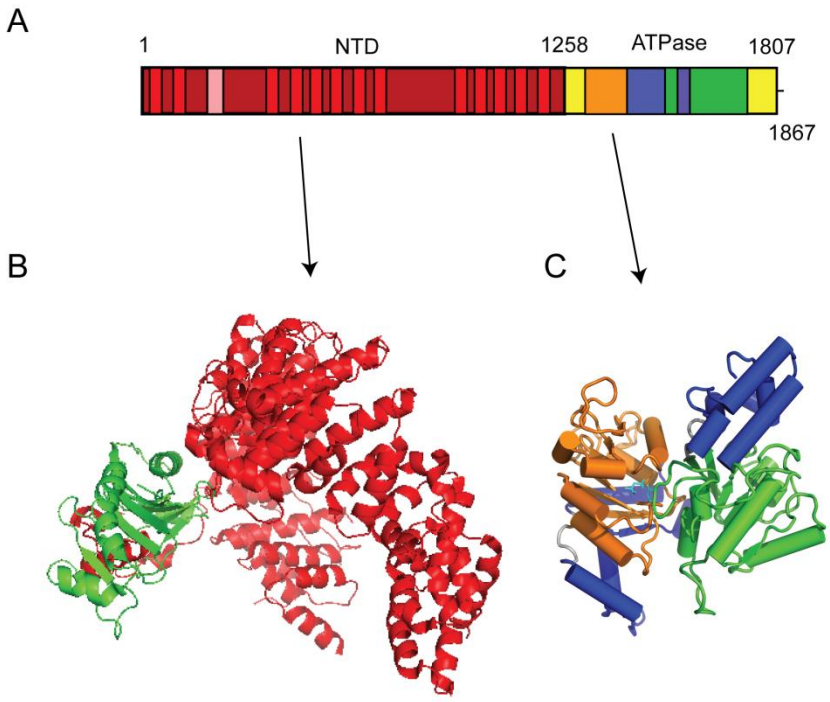

Figure 1.4. Structure of Mot1. *A*, Schematic of the two major domains of Mot1: the N-terminal domain (NTD) comprises roughly two-thirds of the protein (dark red box, residues 1-1258), while the C-terminal Swi2/Snf2 ATPase (yellow box, residues 1258-1807) comprises the remainder. The NTD contains 13 HEAT repeats (bright red boxes; Wollmann et al. 2011) and the latch (pink box). The ATPase domain can be divided into the subdomains 1A (orange box), 2A (green box), 1B, and 2B (blue boxes) *B*, The Mot1 NTD has been crystallized in complex with TBP (picture above; PDB = 3OC3) and also TBP-NC2-DNA (see Appendix A) from *Encephalitozoon cuniculi*. The EcNTD consists of 16 HEAT repeats (red) which bind to the convex surface of TBP (green). *C*, The modeled structure of the Mot1 Swi2/Snf2 ATPase domain. More on the modeling and the ATPase domain will be discussed in Chapter II.

Several models have since been proposed to suggest how Mot1 activates transcription, including remodeling chromatin, forming a unique PIC, removing TBP from improper locations, and removing TBP from high affinity sites to enable a dynamic pool of TBP that is able to bind to lower affinity sites (Topalidou et al. 2004, Geisberg and Struhl 2004, Sprouse et al. 2008a, Zentner and Henikoff 2013). Studies at the *GAL1* gene have shown that Mot1 is required for remodeling of the promoter nucleosomes in order for transcription initiation (Topalidou et al. 2004). However, footprinting analysis at *URA1* showed that mutation of Mot1 did not affect the chromatin structure at this Mot1-activated gene (Sprouse et al. 2008b). *URA1* has also been shown to have a TATA box on the antisense strand of DNA oriented away from the sense direction of transcription that TBP preferentially binds. Therefore, it has been proposed that Mot1 removes incorrectly oriented TBP to enable it to properly bind and initiate transcription (Sprouse et al. 2008b).

It was shown genome-wide that mutation of Mot1 led to a redistribution of TBP from low affinity sites to the high affinity TATA boxes (Zentner and Henikoff 2013), which had been previously suggested by earlier experiments (, Collart 1992, Muldrow et al. 1999, Zanton and Pugh 2004). Analysis of TBP dynamics showed that in the absence of Mot1, TBP is more dynamic. This increase in TBP dynamics in mot1 cells suggested that Mot1 removed TBP from weak-binding sites and therefore stabilized TBP bound to high-affinity sites(Poorey et al. 2013). Locus-specific and genome-wide localization of Mot1 showed a peak of Mot1 at the promoter of genes, regardless of whether Mot1 activates or represses the gene, and an absence of Mot1 in the coding region of genes (Andrau et al. 2002, Dasgupta et al. 2002, Geisberg et al. 2002, Zanton and Pugh 2004). To date only one study has identified Mot1 in coding regions of genes in the absence of TBP

(Geisberg et al. 2002). Due to the lack of corroborating results for Mot1 binding in the ORF, it is possible that this finding has no biological significance.

Separately, both Mot1 and NC2 had been shown to be necessary for repression or activation of genes (Meisterernst and Roeder 1991, Auble and Hahn 1993, Collart 1996, Madison and Winston 1997). Locus-specific experiments showed that Mot1 and NC2 both regulated *HIS3* and *HIS4*, and later genome-wide experiments showed they had a significant overlap in gene regulation (Collart 1996, Madison and Winston 1997, Lemaire et al. 2000, Andrau et al. 2002, Dasgupta et al. 2002, Hsu et al. 2008). ChIP studies also showed that both factors had similar genomic localizations (White et al. 1994, Geisberg et al. 2001, Dasgupta et al. 2002, Geisberg et al. 2002). Characterization of NC2 showed that both subunits were required for binding of NC2 to TBP-DNA (Inostroza et al. 1992, Yeung et al. 1994, Mermelstein et al. 1996, Goppelt et al. 1995, Van Werven et al. 2008). NC2 was able to move TBP along as detected by FRET, which suggests that NC2 could shuttle TBP to promoters of activated genes and move TBP away from repressed genes (Schluesche et al. 2007). NC2 is also perfectly shaped to prevent TFIIA and TFIIB from binding to TBP and forming the PIC, which explains its repressive function (Meisterernst and Roeder 1991, Kamada et al. 2001).

The crystal structure of the Mot1 NTD-NC2-TBP-DNA complex confirmed that this quaternary complex could form (Butryn et al. 2015). Previously it had been suggested that Mot1 and NC2 were antagonistic in binding to the same region of DNA (Geisberg et al. 2001). Though the genome-wide localization profiles and an intact Mot1-NC2-TBP complex purified from whole cell extracts suggested otherwise (Darst et al. 2003). Since the structure of NC2-TBP had previously been solved and Mot1 was known to bind to the convex surface of TBP, it made sense that Mot1 and NC2 could

bind to TBP at the same time (Kamada et al. 2001, Darst et al. 2003, Wollmann et al. 2011). The Mot1 NTD-NC2-TBP-DNA complex demonstrates that TFIIA, TFIIB, SAGA, and TFIID would not be able to bind to TBP, since Mot1 and NC2 bind to the same regions of TBP that those factors bind. (Butryn et al. 2015).

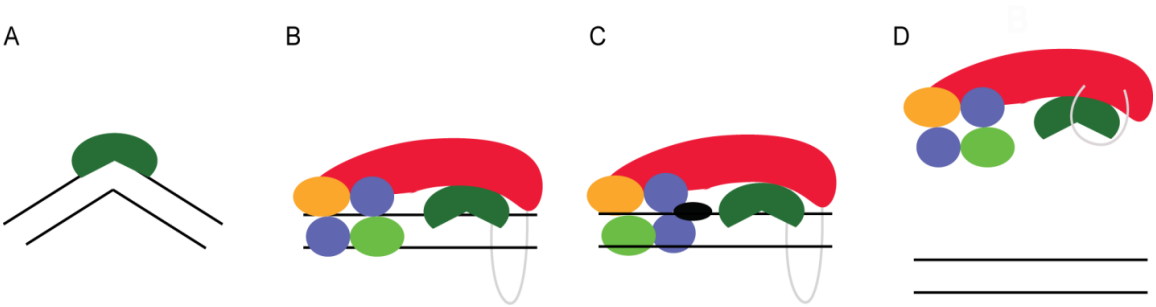

Figure 1.5. Model of Mot1 removal of TBP. *A*, TBP (dark green) binds and bends DNA (black lines). *B*, The Mot1 N-terminal HEAT repeats (red) bind to the convex surface of TBP with the latch (gray) not bound to TBP. The Mot1 ATPase (orange, blue, light green ovals) is positioned in the open conformation on the upstream DNA. *C*, Upon addition of nucleotide (black oval), the Mot1 ATPase transitions to the closed conformation towards TBP. *D*, After several rounds of ATP hydrolysis, TBP is removed from DNA and the Mot1 latch binds to the concave surface of TBP.

The FRET technique allowed for a unique insight into understanding how Mot1 removed TBP from DNA. As previously shown by the crystal structure of TBP-DNA, TBP bends the promoter DNA approximately 90° (Kim et al. 1993). When Mot1 binds to TBP-DNA the FRET study showed that Mot1 straightened the bent DNA somewhat,

potentially priming the TBP for dissociation (Moyle-Heyrman et al. 2012). Importantly, this unbending of the DNA is in the absence of ATP. A model was proposed that after the unbending of DNA that Mot1 then went through multiple rounds of ATP hydrolysis to remove TBP from DNA (Sprouse et al. 2008b, Moyle-Heyrman et al. 2012). The current model of Mot1-mediated removal of TBP from DNA is illustrated in a simple cartoon in Figure 1.5.

Transcription Elongation and the Nucleosomal Obstacle

After transcription has initiated, Pol II has to "escape" the promoter to proceed into productive elongation. The largest subunit of Pol II, Rpb1, has a long flexible tail called the C-Terminal Domain (CTD) that consists of repeats of YSPTSPS (Buratowksi 2009). This CTD plays an integral role in transcription elongation by accumulating modifications, which help to signal that Pol II is now elongating through the gene and recruits proteins that are essential to this process. Serine 5 is first phosphorylated by the TFIIK subunit of TFIIH (Lu et al. 1992). As Pol II progresses through the gene body, Ctk1 phosphorylates serine 2 (Ahn et al. 2004, Smith-Kinnaman et al. 2014). Next, serine 5 is dephosphorylated by Rtr1 (Mosley et al. 2009). These are the two most widely studied modifications of the CTD, however, it is also known that tyrosine 4 and serine 7 can also be phosphorylated, and the prolines can be isomerized (Kouzarides 2007). The roles these modifications play in elongation are not as well defined as serines 2 and 5.

Before serine 5 is dephosphorylated, there is a time when both serine 5 and 2 are phosphorylated. This is essential for the recruitment of the histone

methyltransferase Set2, which requires phosphorylation of both sites for binding (Vojnic et al. 2006). Set2 has been shown to methylate H3K36, which is a mark associated with active transcription (Strahl et al. 2002, Kizer et al. 2005, Kouzarides 2007). It has also been demonstrated that Set2 and H3K36me3 help to suppress cryptic initiation within a gene (Lickwar et al. 2009, Venkatesh et al. 2012). H3K36me3 recruits the Rpd3 histone deacetylase complex, which would then help to reassemble histones on DNA, thereby preventing PICs from inappropriately forming within a gene (Keogh et al. 2005).

The nucleosomes located within the ORF are essential for preventing inappropriate transcription from occurring. However, these nucleosomes also have to be removed in order for Pol II to passage through the ORF to produce the transcript. This is a complex issue that the cell has evolved multiple pathways of regulating. The many layers of nucleosomal regulation include the canonical histones, histone variants, histone modifications, histone chaperones, chromatin remodelers, transcription factors, and finally Pol II. The canonical histones, the H2A.Z variant, and histone modifications have already been introduced previously.

Histone chaperones are a class of proteins that have a high affinity for the histone subunits and have no enzymatic activity (De Koning et al. 2007). These chaperones are specific to the H2A-H2B dimer, the H3-H4 tetramer, or histone variants such as H2A.Z or CENP-A. On the other hand, there are ATP-dependent chromatin remodelers that are members of the Swi2/Snf2 family of ATPases similar to Mot1 (Ryan and Owen-Hughes 2011). Unlike Mot1, the chromatin remodelers utilize ATP hydrolysis to reposition nucleosomes or insert a histone variant into an already established nucleosome.

Two commonly studied histone chaperones in yeast are Nap1 and Asf1. This is because these chaperones have clearly defined histone cargoes; Nap1 binds H2A-H2B and Asf1 binds H3-H4 (Mosammaparast et al. 2001, Mosammaparast et al. 2002, English et al. 2006). Interestingly, while H3-H4 forms a tetramer (Luger et al. 1997), the interaction with Asf1 prevents the tetramer from forming and instead binds H3-H4 as a dimer (English et al. 2006). H2A-H2B normally forms a dimer; however a dimer of Nap1 is able to form an H2A-H2B tetramer, potentially explaining how Nap1 prevents reincorporation of H2A-H2B into the nucleosome (D'Arcy et al. 2013). This could be a feature specific to Nap1 or other H2A-H2B histone chaperones could also function like this.

Another well-characterized histone chaperone is Spt6, which was identified in a yeast genetic screen (Winston et al. 1984, Duina 2011). Spt6 has been shown to assemble nucleosomes *in vitro* and preferentially bind H3-H4 (Bortvin and Winston 1996, Hartzog et al. 1998). *In vivo* experiments showed that at genes with high transcription rates, Spt6 was required for replacement of nucleosomes after Pol II passage, thereby preventing intragenic transcription (Kaplan et al. 2003, Cheung et al. 2008, Jensen et al. 2008). Spt6 binds to the CTD of Pol II specifically phosphorylated at serine 2, which starts accumulating on the CTD during productive elongation (Sun et al. 2010). While Spt6 is required for transcription elongation, it has also been shown to be necessary for nucleosome reassembly at the promoters of some genes after initiation has occurred (Adkins and Tyler 2006, Ivanovska et al. 2011). Spt6 requires the high mobility group (HMG) protein Nhp6 to bind to nucleosomes (Bortvin and Winston 1996). Furthermore, Spt6 also binds to another protein, Spn1, which prevents the association of Spt6 with nucleosomes (McDonald et al. 2010, Diebold et al. 2010). A protein very

similar in function to Spt6 is the histone chaperone Spt16, which will be discussed more in the next section (Duina 2011).

Another important elongation regulator is the PAF complex (Polymerase Associated Factors) (Wade et al. 1996). This complex consists of 5 subunits in yeast: Paf1, Rtf1, Leo1, Cdc73, and Ctr9 (Shi et al. 1996, Shi et al. 1997, Mueller and Jaehning 2002, Squazoo et al. 2002). It has been shown by ChIP to localize to the entire coding region of active genes (Krogan et al. 2002, Pokholok et al. 2002). Unlike Set2 which requires the Pol II CTD to be phosphorylated at serines 5 and 2, the PAF complex associates with all forms of Pol II regardless of the phosphorylation status through interactions with the Cdc73 and Rtf1 subunits (Rozenblatt-Rosen et al. 2005). However, it seems that phosphorylation of another elongation factor Spt5 is required for recruitment of the PAF complex to Pol II (Liu et al. 2009, Zhou et al. 2009). The PAF complex is required for transcription elongation due mostly to two functions: recruitment of Rad6/Bre1 and the full phosphorylation of serine 2 of the Pol II CTD (Jaehning 2010). Rad6/Bre1 ubiquitinates H2BK123 which recruits the Set1 methyltransferase that subsequently methylates H3K4 (Krogan et al. 2003, Ng et al. 2003). The mechanism in which the PAF complex contributes to serine phosphorylation is not fully understood (Mueller et al. 2004); however, this phosphorylation is required for the recruitment of Set2 and subsequent tri-methylation of H3K36 (Krogan et al. 2003). The PAF complex has also been shown to physically and genetically interact with Spt16, indicating the need for cooperation of elongation factors in transcriptional regulation (Krogan et al. 2002, Squazzo et al. 2002, Rondón et al. 2004).

Spt16 Regulates Transcription, Replication, and DNA Repair

Spt16 is a histone chaperone that is part of the FACT (facilitates chromatin transcription) complex (Formosa 2012). Spt16 (or Cdc68) was first identified in a genetic screen in yeast as a suppressor of transposable elements (Malone et al. 1991, Rowley et al. 1991). Genes with transposons integrated into the coding region would normally not be able to be properly transcribed, however mutations of *SPT16* and other *SPT* genes enable these genes to be properly transcribed. Other *SPT* genes include the histone chaperone *SPT6*, the histone genes, and even *SPT15* which encodes TBP (Clark-Adams and Winston 1987, Eisenmann et al. 1989, Hahn et al. 1989a, Malone et al. 1991).

Spt16 forms a heterodimer with Pob3 and together they loosely associate with Nhp6 to form the complete yeast FACT complex (Brewster et al. 1998, Formosa et al. 2001, Formosa 2012). Multiple copies of Nhp6 are required for Spt16-Pob3 to bind to nucleosomes (Ruone et al. 2003). Early *in vitro* studies showed that FACT was required for productive transcription elongation through a nucleosomal template (Orphanides et al. 1998). Many labs have gone on to dissect the roles of both Spt16 and Pob3 in transcription elongation. The domain organization of Spt16 is shown in Figure 1.6. Spt16 consists of an N-terminal domain (NTD; VanDemark et al. 2008), an aminopeptidase-like domain (Stuwe et al. 2008), a dimerization domain, a middle domain, and the C-terminal domain (CTD; O'Donnell et al. 2004). So far only the NTD, peptidase domain, and middle domain have been crystallized (Figure 1.6*B*, *C;* VanDemark et al. 2008, Kemble et al. 2013).

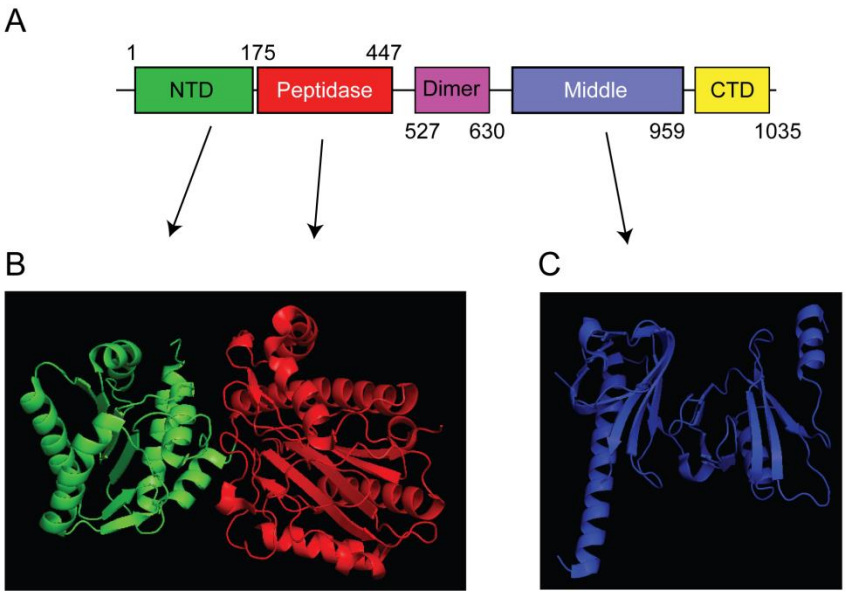

Figure 1.6. Structure of the histone chaperone Spt16. *A*, Schematic of the domains of Spt16. Only the N-terminal domain (NTD, green), peptidase domain (red), and middle domain (blue) have been crystallized so far. The dimerization domain (Dimer, purple) and C-terminal domain (CTD, yellow) are the other two domains in Spt16. *B*, The crystal structure of the Spt16 NTD (green) and peptidase domains (red; PDB = 3BIP). *C*, The crystal structure of the Spt16 middle domain (blue; PDB = 4IOY).

Originally it was thought that Spt16 was a histone chaperone specific to the H2A-H2B dimer (Belotserkovskaya et al. 2003, Ransom et al. 2010, Formosa 2012). However, further biochemical analyses have shown that Spt16 can bind both H2A-H2B dimers and H3-H4 tetramers (Stuwe et al. 2008, VanDemark et al. 2008). The domains which bind to these subunits differ (Winkler et al. 2011, Tsunaka et al. 2016). The NTD and middle domains both bind to H3-H4 tetramers, whereas the acidic CTD binds to

H2A-H2B dimers. The dimerization domain is required for dimerization to Pob3 to form the FACT complex (Formosa 2012).

The genome-wide localization profile of Spt16 showed that Spt16 mostly associates with actively transcribed genes (Mason and Struhl 2003, Saunders et al. 2003). Spt16 peaks immediately downstream of the TSS and continues through the coding region eventually dropping off before the transcription termination site (Mayer et al. 2010). This is unique in comparison to other elongation factors, which typically peak further downstream of Spt16 and exit at the transcription termination site. This has caused some speculation that Spt16 may have a role earlier in elongation and potentially even in late initiation by regulating the +1 nucleosome (Mayer et al. 2010, Belotserkovskaya et al. 2003, Stuwe et al. 2008).

Mutation of Spt16 caused a phenomenon known as cryptic or intragenic initiation (Kaplan et al. 2003, Cheung et al. 2008). This occurs when a promoter element within the ORF is exposed and a PIC is able to assemble within the gene. There are several genes that are widely studied for cryptic initiation including *FLO8* (Duina et al. 2007). This exposure of a cryptic promoter is enabled by the absence of functional Spt16 to replace histones evicted in the wake of Pol II passage (Formosa et al. 2002, Belotserkovskaya et al. 2003, Kaplan et al. 2003, Cheung et al. 2008, Jamai et al. 2009, Xin et al. 2009). Since a PIC forms at this cryptic site, Spt16 can indirectly function in TBP localization. Spt16 has also been implicated in TBP localization to the promoter at a subset of genes (Mason and Struhl 2003, Shimojima et al. 2003, Biswas et al. 2005). In *spt16* cells, TBP localization was decreased at the promoters of genes, potentially through the relocalization of TBP to other exposed sites (Biswas et al. 2005). There is

no evidence that Spt16 directly interacts with TBP and since Spt16 is predominately localized to the coding region, this is most likely an indirect effect of Spt16 on TBP.

The histone variant H2A.Z is enriched at the promoter of active genes (Li et al. 2007). This histone variant has the specific histone chaperone Chz1 (Luk et al. 2007). On top of that, the Swr1 chromatin remodeler is also specific for H2A.Z deposition (Mizuguchi et al. 2003, Korber and Horz 2004). Two independent studies identified Spt16 as another histone chaperone for H2A.Z (Mahapatra et al. 2011, Liu et al. 2014). The first study demonstrated that Spt16 was necessary for the deposition of H2A.Z at tRNA genes (Mahapatra et al. 2011). The second study showed that Spt16 actually is prevented from binding H2A.Z while Chz1 is present, but in strains lacking Chz1 Spt16 was able to bind to H2A.Z nucleosomes (Liu et al. 2014). While Spt16 is linked more to elongation, it is apparent that Spt16 has some roles in transcription initiation as well.

Along with FACT regulating transcription, both Spt16 and Pob3 have been implicated in DNA replication. Originally, this was shown through physical interactions with both proteins to the replicative polymerase DNA Polymerase α in whole cell extracts (Wittmeyer and Formosa 1997). After DNA has been newly synthesized, the immediate problem is that histones have to be replaced along the DNA and new histones also have to be deposited (Annunziato 2005). Several histone chaperones have been shown to be involved in this replication-dependent histone assembly, including the chromatin assembly factor 1 (CAF-1) complex, Asf1, and Rtt106 (Stillman 1996, Tyler et al. 1999, Huang et al. 2005). It was shown that the interaction between Spt16 and H3-H4 was required for deposition of new histones onto DNA during replication, indicating that this is how FACT regulates replication (Foltman et al. 2013, Yang et al. 2016). This interaction

of FACT with the replicative enzymes is mediated by ubiquitination of Spt16 by Rtt101 (Han et al. 2010).

Finally, the FACT complex has been shown to be involved in DNA repair. When DNA is damaged, histones have to be removed to allow for the DNA repair machinery to get to the damaged site of DNA. Histone chaperones are important in replacing histones after the DNA has been properly repaired (Soria et al. 2012). These histone factors include the CAF-1 complex, Asf1, and Spt16 (Polo et al. 2006, Dinant et al. 2013). Spt16 is necessary for the replacement of H2A-H2B dimers specifically, whereas Pob3 is not required for the reassembly of nucleosomes after DNA repair (Dinant et al. 2013). However, both subunits are recruited to the damaged site. In higher eukaryotes, the role of Spt16 in DNA repair has been further characterized. FACT interacts with the histone variant H2A.X, which has been shown to be a marker of DNA damage (Heo et al. 2008, Redon et al. 2002). FACT is involved in the deposition and removal of H2A.X. This interaction with H2A.X is reduced when Spt16 is poly(ADP)-ribosylated by PARP1(Heo et al. 2008). It was demonstrated that the PARP1-dependent poly(ADP)-ribosolyation of Spt16 reduces Spt16's affinity with all chromatin, not just H2A.X (Huang et al. 2006).

Summary

Mot1 and Spt16 are conserved and essential regulators of transcription. Initially, Mot1 was thought to repress transcription by removing TBP from DNA and preventing PIC formation. However, it has been shown that Mot1 is also required for activation of transcription and several models have been suggested. Experiments have shown that Mot1 may have a role in elongation through the prevention of premature termination,

although the predominant localization of Mot1 is to promoters and not coding regions. The Spt16 histone chaperone is part of the FACT complex and has been widely characterized as an elongation factor. Spt16 associates with RNA Polymerase II and reassembles nucleosomes after RNA Polymerase II passages through the coding region. Spt16 is thus important for preventing cryptic initiation. Spt16 has also been implicated in regulation of initiation through its interaction with promoter-proximal nucleosomes, H2A.Z, and effects on TBP localization to the promoter.

Scope of this study

While there have been substantive advances in understanding the mechanism in which Mot1 functions, we still do not understand how the ATPase domain specifically interacts with the upstream DNA. To date there is no published crystal structure of the Mot1 ATPase domain in complex with the upstream DNA, which is important since the upstream DNA is required for Mot1-mediated dissociation of TBP. We previously reported mapping of the ATPase domain to this region of DNA (Wollmann et al. 2011); however, we more clearly defined the subdomains in relation to the DNA. The crystal structure of the related SsoRad54 ATPase domain led to the suggestion that Swi2/Snf2 ATPases could form two different conformations which could be important for their mechanism of action (Dürr et al. 2005). Therefore, we hypothesized that Mot1 could also form these conformations, which would be important for hydrolysis of ATP and subsequent removal of TBP from DNA. All of this will be discussed in Chapter II.

Mot1 and Spt16 have both been widely characterized independently; however, there have been no studies exploring their relationship. It was previously demonstrated

that these two proteins physically interact (Arnett et al. 2008). Mot1 and Spt16 had been separately shown to regulate transcription initiation and elongation (Auble and Hahn 1993, Auble et al. 1994, Poorey et al. 2010, Biswas et al. 2005, Belotserkovskaya et al. 2003), and in order to understand how Mot1 influenced elongation we determined their effects on gene expression. Therefore, we used single mutants and created a double mutant to determine their effects on gene expression using tiling arrays. We also determined their roles in RNA synthesis precision, since Mot1 prevents premature termination (Poorey et al. 2010) and Spt16 prevents cryptic initiation (Cheung et al. 2008). Since both factors have been implicated in TBP localization and PIC formation (Auble and Hahn 1993, Auble et al. 1994, Biswas et al. 2005), we hypothesized that they would overlap in transcription regulation. We determined their effects on components of the PIC and their preferences for promoter elements. All of this will be discussed in Chapter III.

Mot1 and Spt16 have different genomic localization profiles; Mot1 localized predominately to the promoter, whereas Spt16 localized to the coding region of active genes. We hypothesized that these two factors could affect the genome-wide localization of each other based on their physical interaction and overlap in transcription regulation. Spt16 is a bona fide histone chaperone, and Mot1 belongs to a family of proteins that are all chromatin remodelers. One study demonstrated that Mot1 was required for chromatin remodeling at a specific gene (Topalidou et al. 2004), while another study showed it was not required for chromatin organization at another gene (Dasgupta et al. 2005). Therefore we determined if Mot1 had a global role in nucleosome organization and how it compared to Spt16 to determine if they overlapped in nucleosomal regulation. These results will be discussed in Chapter IV.

Chapter II

The Activity and Specificity of Mot1 in Purified Complexes

Most of the data from this Chapter was published in Viswanathan et al. 2016.

The essential *Saccharomyces cerevisiae* ATPase Mot1 globally regulates transcription by impacting the genomic distribution and activity of the TATA-binding protein (TBP). *In vitro*, Mot1 forms a ternary complex with TBP and DNA, and can use ATP hydrolysis to dissociate the TBP-DNA complex. Prior work suggested an interaction between the ATPase domain and a functionally important segment of DNA flanking the TATA sequence. However, how ATP hydrolysis facilitates removal of TBP from DNA is not well understood and several models have been proposed. To gain insight into the Mot1 mechanism, we employed a DNA tethered cleavage approach to map regions of Mot1 in proximity to the DNA under different conditions. We present biochemical evidence for two distinct conformations of the Mot1 ATPase, the detection of which can be modulated by ATP analogs as well as DNA sequence flanking the TATA sequence. We also show using purified complexes that Mot1 dissociation of a stable, high affinity TBP-DNA interaction is surprisingly inefficient, suggesting how other transcription factors that bind to TBP may compete with Mot1. Taken together, these results suggest that TBP-DNA affinity as well as other aspects of promoter sequence influence Mot1 function *in vivo*.

Introduction

Biochemical Analyses of Mot1-Mediated Dissociation of TBP

As discussed in Chapter I, Mot1 was biochemically identified as an ATP-dependent inhibitor of TBP (ADI; Auble and Hahn 1993). It was shown to be a single polypeptide that belonged to the family of Swi2/Snf2 ATPases (Auble et al. 1994, Poon et al. 1994, Adamkewicz et al. 2000). It was capable of binding TBP-DNA complexes, and the association with DNA is dependent on TBP (Auble and Hahn 1993, Auble et al. 1994). Mot1, TBP, and DNA form a 1:1:1 complex that upon addition of ATP is dissociated to the single components (Auble and Hahn 1993, Auble et al. 1994, Poon et al. 1994). Electrophoretic mobility shift assays (EMSAs) allow for detection of the different complexes and dissociation when ATP is added.

Mot1 has been characterized in other eukaryotes, including *Drosophila melanogaster* (dMot1 or 89B helicase; Goldman-Levi et al. 1996) and humans (BTAF1; Timmers and Sharp 1991, Timmers et al. 1992). Interestingly, while the ATPase activity of Mot1 is stimulated by TBP, BTAF1 requires TBP-DNA (Chicca et al. 1998). This indicated some biochemical difference between Mot1 and BTAF1, which could impact their function. In *Drosophila*, it was also shown that 89B helicase could not only bind to TBP but also to the TBP-related factor TRF1 (Adamkewicz et al. 2001). Although there are differences in some of the biochemical properties of the Mot1 orthologs, since Mot1 is well-conserved, understanding its function in one system will provide a framework for Mot1 action in all organisms.

The removal of TBP from DNA is an obvious explanation for the repressive mechanism of Mot1, which would prevent the PIC from forming (see Chapter I).

However, studies also showed that Mot1 was able to activate transcription at certain genes including *HIS4* (Collart 1996, Muldrow et al. 1999). It was shown that these *in vivo* results of gene activation could be recapitulated *in vitro* using low levels of Mot1 that were approximately stoichiometric to TBP (Muldrow et al. 1999). In contrast, higher levels of Mot1 led to repression of the *HIS4* gene (Muldrow et al. 1999). *URA1,* another Mot1-activated gene, was extensively analyzed to understand the Mot1 activation mechanism. *URA1* has two TATA boxes in opposite orientations in its promoter (Sprouse et al. 2008b). Hydroxyl radical experiments were employed to demonstrate that TBP preferentially bound to the TATA box that was facing in the opposite direction from the TSS of *URA1* (Sprouse et al. 2008b). Using a mutated version of TBP that was able to bind to either TATA box in addition with the Mot1-42 ts protein, it was demonstrated that activation of *URA1* expression could be restored, indicating that Mot1 functioned to reorient TBP at the *URA1* promoter (Sprouse et al. 2008b).

As mentioned in Chapter I, when TBP binds to DNA it bends it approximately 90° which is necessary for PIC formation (Starr and Hawley 1991, Lee et al. 1991, Kim et al. 1993). Using FRET it was demonstrated that when Mot1 binds to TBP-DNA, it unbends the DNA (Moyle-Heyrman et al. 2012). The extent to which it unbends the DNA is uncertain. This occurred in the absence of ATP, which could prime the TBP for removal. In the time it takes for Mot1 to remove TBP from DNA, approximately 12-13 molecules of ATP are hydrolyzed (Sprouse et al. 2006, Viswanathan and Auble 2011).

Since it was known that Mot1 removed TBP from DNA using ATP hydrolysis, the next step was to understand how Mot1 performed this function. Other Swi2/Snf2 enzymes had been shown to productively translocate along DNA (Pazin and Kadonaga 1997), which was a provocative model for Mot1 (Auble and Steggerda 1999). However,

with a series of “roadblock” experiments, it was demonstrated that Mot1 did not require a long stretch of DNA in order to translocate and remove TBP, although it still could translocate over a short distance (Auble and Steggerda 1999). Mot1 was also able to remove TBP even when the DNA duplex was crosslinked with psoralen, indicating that it did not require strand separation like helicases (Darst et al. 2001). Furthermore, it was shown that Mot1 did not change the conformation of the DNA by utilizing DNA in mini-circles on which Mot1 was still able to remove TBP (Darst et al. 2001).

Mot1 required ~17 base pairs of double stranded DNA upstream of the TATA box in order to remove TBP (Darst et al. 2001). In addition, it was shown that Mot1 binding to DNA was not affected by removal of the nucleosides from upstream DNA (Darst et al. 2001). However, fluorescence polarization experiments by another lab showed that Mot1 could remove TBP without the requirement for the upstream DNA (Gumbs et al. 2003). This apparent contradiction for the requirement of upstream DNA for Mot1 will be elaborated upon in the discussion. A Swi2/Snf2 ATPase Rad54 from *Sulfolobus solfataricus* was crystallized in complex with DNA showing its requirement for the upstream DNA (Dürr et al. 2005). In addition, it showed that the ATPase domain could interact with the phosphate backbone of the DNA indicating that the ATPase domain did not require a specific sequence to bind (Dürr et al. 2005).

The Two Conformations of Swi2/Snf2 ATPases

As a member of the Swi2/Snf2 family of ATPases, Mot1 has a conserved C-terminal ATPase domain (Gorbalenya et al. 1988, Singleton et al. 2007, Flaus et al. 2006, Davis et al. 1992). This domain can be further classified into four subdomains:

1A, 2A, 1B, and 2B (Dürr et al. 2005). Subdomains 1A and 1B are similar to the domains in the RecA helicase and are therefore called RecA-like subdomains (Singleton et al. 2001). The subdomains 2A and 2B are specific to the Swi2/Snf2 family of ATPases (Dürr et al. 2005). Perhaps the first biochemical evidence of a conformational change in Mot1 was demonstrated when a transition state ATP analog, ADP-AlF_4 was utilized (Darst et al. 2003). When Mot1 was preincubated with TBP, it was unable to bind to DNA. However, addition of ADP-AlF_4 to the Mot1-TBP complex enabled it to bind to DNA suggesting that the binding of the analog changed the conformation of Mot1 in some manner (Darst et al. 2003).

The first tangible evidence of two conformations of Swi2/Snf2 ATPases was discovered when the *Sulfolobus solfataricus* Rad54 ATPase domain was crystallized in complex with DNA (Dürr et al. 2005). This structure showed the ATPase in a conformation where the catalytic residues of the ATPase domain were not close together. In this manner the ATP pocket could not form to bind ATP. Thus it was proposed that this was an “open” conformation and that the domain had to rotate 180° to properly form the ATP binding pocket and the “closed” conformation (Dürr et al. 2005). This also worked well with the “inchworm” mechanism, which suggested that the ATPase moved like an inchworm along DNA in small steps (Velankar et al. 1999). Translocation over such a short distance is difficult to detect, so this model does not contradict the earlier observation that Mot1 does not translocate over a long distance.

The first crystal structure of Mot1 from *Encephalitozoon cuniculi* was only of the Mot1 NTD in complex with TBP (Wollmann et al. 2011). However, negative stained electron micrographs enabled a more complete picture of Mot1 in complex with TBP. This allowed for Mot1-TBP to be superimposed upon DNA showing that the ATPase

domain would come into contact with the upstream DNA as shown by earlier biochemical analyses. Furthermore, the FeBABE hydroxyl radical cleavage assay was employed to show that the Mot1 ATPase domain came into contact with the upstream DNA (Wollmann et al. 2011).

Scope of this Study

Mot1 has been biochemically characterized and the structure of the NTD in complex with TBP-DNA and NC2-TBP-DNA has been solved (Auble et al. 1994, Poon et al. 1994, Adamkewicz et al. 2000, Wollmann et al. 2011, Butryn et al. 2015). However, there are still some parts of the mechanism that are not fully understood. The biochemical characterization has been done mostly using ensemble reactions such as EMSAs. Therefore, we determined how Mot1 functioned in purified complexes to better understand the efficiency of the ATPase. Mot1 had previously been shown to require a 17 base pair handle upstream of the TATA box, although there was some conflicting evidence (Darst et al. 2001, Gumbs et al. 2003). This requirement of the upstream DNA was further supported by FeBABE experiments that showed that the Swi2/Snf2 ATPase domain contacted the upstream DNA (Wollmann et al. 2011). An ATP analog was shown to capture a conformational intermediate in the Mot1 hydrolytic cycle (Viswanathan 2013, Viswanathan et al. 2016). Previously, we had proposed that Mot1 binding to the upstream DNA could be affected by the DNA sequence (Sprouse et al. 2008b). We tested if another ATP analog and various DNA sequences would affect the conformation of Mot1 binding. Overall, this chapter focuses on these mechanistic steps in the Mot1-mediated dissociation of TBP from DNA.

Materials and Method

TBP and Mot1 Purification

Full-length recombinant *Saccharomyces cerevisiae* (*Sc)* TBP was obtained by purification from a bacterial overexpression strain as described previously (Darst et al. 2002, Gupta et al. 2007) and quantified as described previously (Sprouse et al. 2006). Full-length recombinant *Sc* Mot1 was isolated by affinity purification using a yeast overexpression strain as described previously (Darst et al. 2002, Gupta et al. 2007).

DNA Probes

The synthetic deoxyoligonucleotides (oligos) used in this study are listed in Table 2.1. The gapped DNA probes were prepared by annealing combinations of the oligos corresponding to the sequences shown in Figure 1*A*. The oligos used for FeBABE assays (obtained from Life Technologies) were biotinylated on the 5' end of the top strand and hybridized to unmodified bottom strand oligos. The duplexes were formed in 25 mM Tris-Cl pH 8, 10 mM EDTA, 0.1 M NaCl by heating to 100°C for 5 minutes and then slow cooling overnight to room temperature. A two-fold excess of the bottom strand was used during annealing to increase the yield of the duplex product. For testing the catalytic activity of Mot1, oligonucleotides were biotinylated at the 5' end (upstream) or 3' end (downstream) of the TATA box. A two-fold excess of the unmodified strand was added and annealing conditions were the same as those for the FeBABE DNAs. A 20 nucleotide stretch of single-stranded DNA was added to the 5' end of the 5' Bio duplex to create the 5' Tail duplex.

Table 2.1. Oligonucleotides used in this chapter. The phosphorothioates of A, C, G and T are represented by F, O, E and Z respectively.

Name	Sequence
FeB	5’-GGG CGC GAA TTC ACG CGC CAC CCC CTT TTA TAG CCC CCC TTC AGG AAC ACC CGG TCG GGG ATC CCG GGC C-3’
6FeT	5’-BIO-GGC CCG GGA TCC CCG AOC EGG TGT TCC TGA AGG GGG GCT ATA AAA GGG GGT GGC GCG TGA ATT CGC GC-3’
HIS4 6FeT	5’-BIO-CGG GGA TCT GTC GAC CZC EAG AAC AGT AGT ATA CTG TGT ATA TAA TAG ATA TGG AAC-3’
HIS4 6FeB	5’-GTT CCA TAT CTA TTA TAT ACA CAG TAT ACT ACT GTT CTC GAG GTC GAC AGA TCC CCG-3’
Moyle 6FeT	5’-BIO-CGA GAC ACA GAC GTA CGG OCE GGC GCC CCG GAT GGG GGC CTA TAA AAG GGC TGG GCG-3’
Moyle 6FeB	5’-GCT CTG TGT CTG CAT GCC GGC CCG CGG GGC CTA CCC CCG GAT ATT TTC CCG ACC CGC-3’
5’BioT	5’-BIO-GGC CCG GGA TCC CCG ACC GGG TGT TCC TGA AGG GGG GCT ATA AAA GGG GGT GGC GCG TGA ATT CGC GC-3’
5’TailT	5’-BIO-CGA GAC ACA GAC AGA CAG ACG GCC CGG GAT CCC CGA CCG GGT GTT CCT GAA GGG GGG CTA TAA AAG GGG GTG GCG CGT GAA TTC GCG C-3’
EMSA T	5’-CAC GGG GTA CGG CCG GGC GCC CCG GAT GGG GGG CTA TAA AAG GGG CGC-3’
EMSA B	5’-GCG CCC CTT TTA TAG CCC CCC ATC CGG GGC GCC CGG CCG TAC CCC GTG-3’

FeBABE-mediated hydroxyl radical cleavage assay

Six millimolar Fe(III) (s)-1-(p-bromoacetamidobenzyl) ethylenediamine tetraacetic acid (FeBABE) (Dojindo, F279-10) was coupled to 0.7 μM of oligonucleotide hybrid (via specific phosphorothioate modifications) in 20 mM MOPS pH 7.9 for 16 hours at 50°C. The FeBABE-coupled oligonucleotide hybrid was then bound to streptavidin beads (Dynabeads M-280 Streptavidin, 112.05D, Invitrogen) using the biotin linker on the 5’ end of one strand. The beads were then washed twice with 20 mM MOPS pH 7.9 and once with the reaction buffer (4% glycerol, 4 mM Tris-Cl pH 8, 60 mM KCl, 5 mM MgCl2,

100 mg/mL BSA). Mot1-TBP-DNA complexes were formed using buffer conditions identical to those used for EMSA described below. TBP was added to 20 nM final concentration and incubated on a rotator at room temperature for 20 minutes. The beads were then pulled down using a magnet and the supernatant (containing unbound TBP) was removed. The beads were then resuspended in reaction buffer and Mot1 was added to a final concentration of 8 nM, followed by incubation at room temperature for 30 minutes. After this incubation, the beads were pulled down again and washed with reaction buffer without DTT. The beads were then resuspended in 7.5 µL reaction buffer without DTT, and 2.5 µL of 50% glycerol was added as described by Chen and Hahn 2003. To the reactions that contained ADP-AlF_4, 1 µL of 2 mM ADP, 0.5 µL of 50 mM NaF and 0.1 µL of 9 mM $AlCl_3$ were pre-mixed and the mixture was added. To the reactions that contained ADP-BeF_3, 1 µL of 2 mM ADP, 0.1 µL of 50 mM NaF, and 0.12 µL of 9 mM $BeCl_2$ were pre-mixed and the mixture was added. The addition of either analog mixture was followed by a 30 second incubation before initiating the hydroxyl radical cleavage reaction. Hydroxyl radical cleavage was initiated by the addition of 1.25 µL of 50 mM ascorbate to reduce the Fe^{3+} to Fe^{2+}, followed by the addition of 1.25 µL of 50 mM H_2O_2/10 mM EDTA. After incubating for 5 minutes at 37°C, the reaction was stopped using 6 µL of 4X Laemmli protein gel sample buffer and 1 µL of 1M DTT. The products of the cleavage reaction were then resolved on 14% denaturing gels and Western blotted using an antibody specific for the C-terminus of Mot1 (Auble et al. 1997, Darst et al. 2001, Sprouse et al. 2006).

Mapping FeBABE cleavage sites

The molecular weights of the bands on the Western blots were calculated by measuring their migration distance from the well of the gel relative to molecular weight

standards. The average molecular weight of a band was estimated from at least 3 independent blots and the cleavage sites were mapped to the primary sequence using these molecular weight estimates.

Electrophoretic Mobility Shift Assays (EMSAs)

Detection of TBP and Mot1 binding to DNA by native gel electrophoresis was conducted as described (Darst et al. 2001) using a fluorescently labeled fragment of the Adenovirus Major Late Promoter (< 1 nM) and the concentrations of TBP and Mot1 as indicated in the figure legends. Where indicated, ATP was added to 50 μM. Band intensities were quantified using a Typhoon PhosphorImager and ImageQuant software (GE Healthcare). The data presented were obtained by averaging the results of at least two independent experiments.

Immobilized template assays

Oligonucleotide duplexes were coupled to streptavidin magnetic beads and washed as described above for the FeBABE experiments. TBP-DNA and Mot1-TBP-DNA complexes were formed under the same reaction conditions as described above. The complexes were washed once with reaction buffer and split into separate reaction tubes. Each reaction was then resuspended in 20 μL reaction buffer (the same conditions as described for EMSAs). For reactions in which additional Mot1 was added in solution, 2 μL of 382 nM Mot1 was added. ATP was added to the indicated reactions to a final concentration of 100 μM for 5 minutes at room temperature. Where indicated, a

31 bp competitor DNA fragment containing a TATA box sequence was added to a final concentration of 7.5 ng/μL. Then the reactions were again pulled down using a magnet and the supernatant was carefully removed. The beads were resuspended in 20 μL of reaction buffer and half the volume was used for the gel. The beads and supernatant were then added to 2X sample buffer, boiled, and loaded onto 12% or 10% SDS gels and Western blotted for TBP and Mot1 using α-TBP (Reddy and Hahn 1991) and α-Py (Auble et al. 1997) antibodies, respectively.

Molecular Modeling

To visualize the biochemical results in the context of the TBP-DNA complex, several molecular models were developed. Models of the Mot1 ATPase in the 'open' and 'closed' forms were obtained by threading the Mot1 ATPase sequence through published structures of Snf2/Swi2 ATPase domains. The open form model was obtained using the SsoRad54 structure (1Z6A) and CPHmodels (Nielsen et al. 2010). The closed form model was obtained using the *Danio rerio* Rad54 structure (1Z3I) and PROTEUS2 (Montgomerie et al. 2008). Images shown in Figure 2.6 were obtained using MacPyMol 2 (Schrodinger, LLC).

Results

As already mentioned, Mot1 removes TBP from DNA using the energy from ATP hydrolysis (Auble et al. 1994, Poon et al. 1994, Adamkewicz et al. 2000). This has been

studied mostly using EMSAs. Mot1 can remove close to 100% of TBP in the typical EMSA setup, which consists of all the components together in solution. We determined how Mot1 functioned in a purified complex rather than in a mixture in order to understand its catalytic efficiency. In order to do this we utilized the immobilized template assay, which we more commonly refer to as the bead-bound assay. Using this approach, we can assemble the DNA, TBP, Mot1 complex in a stepwise manner, and importantly, only have the bound components remaining when we start the ATPase reaction.

Double-stranded DNA containing a TATA box was biotinylated on one strand of the upstream DNA. The biotinylated DNA could then be bound to streptavidin beads to ensure that only duplex DNA was bound and everything else was washed away. Then TBP could be incubated with the bead-bound DNA and bind to it. The unbound TBP is washed away. Next, Mot1 is incubated with the TBP-DNA-bead complexes and unbound Mot1 is washed away. The end result is a purified DNA-TBP-Mot1 complex attached to streptavidin beads. ATP can then be added to this complex to activate Mot1 and dissociate both Mot1 and TBP.

Figure 2.1 shows a schematic of the duplex DNAs used in these experiments. Both duplexes consist of the same promoter region with a TATA box as indicated, except the 5' Tail duplex has an extra 20 nucleotide single-stranded tail on the 5' end of the top strand. The binding of TBP to the beads was DNA-dependent and the binding of Mot1 to the beads was dependent on TBP-DNA (not shown.) As shown in Figure 2.1*B* (lane 1), TBP bound to the beads in the presence of the 5' Bio duplex. Lanes 2-8 show the TBP that was present in the supernatant after dissociation. Unbiotinylated TATA-containing DNA was used as competitor DNA and did not remove TBP from the beads in

the absence or presence of Mot1 (lanes 2 and 4). TBP was not present in the supernatant when Mot1 was bound to the TBP-DNA complexes (lane 3). Surprisingly, when ATP was added to the TBP-Mot1-DNA complexes, there was no TBP detected in the supernatant. This is in contrast to the large amount of TBP that was dissociated from DNA by Mot1 in the EMSAs. Since TBP can rebind to the bead-bound DNA, we determined if this was the cause of the lack of TBP detected. Competitor DNA was added to the TBP-Mot1-DNA complexes along with ATP, which allowed for detection of TBP in the supernatant. This result suggests that the competitor DNA was able to prevent the displaced TBP from rebinding to the bead-bound DNA. Next, we added excess Mot1 in solution to these reactions to see if increasing the local concentration of Mot1 would increase dissociation of TBP. In fact, the maximal observed dissociation of TBP was achieved when we added ATP, competitor DNA, and excess Mot1 in solution to the Mot1-TBP-DNA complexes. Quantification of this band allowed us to determine that under maximal conditions only ~40% of TBP was removed from DNA. This is in contrast to nearly complete dissociation of TBP detected by EMSAs.

We also determined whether or not Mot1 dissociated from TBP-DNA when we added ATP. Mot1 binding to the beads was dependent on TBP-DNA, and Mot1 was not detected in the supernatant (Figure 2.1*C*, lanes 1 and 4). When ATP was added, Mot1 was detected in the supernatant (lane 5) but some still remained bound to the TBP-DNA beads (lane 2). Excess Mot1 in solution somewhat increased the amount of Mot1 that dissociated from the beads (lane 3) and was present in solution (lane 6).

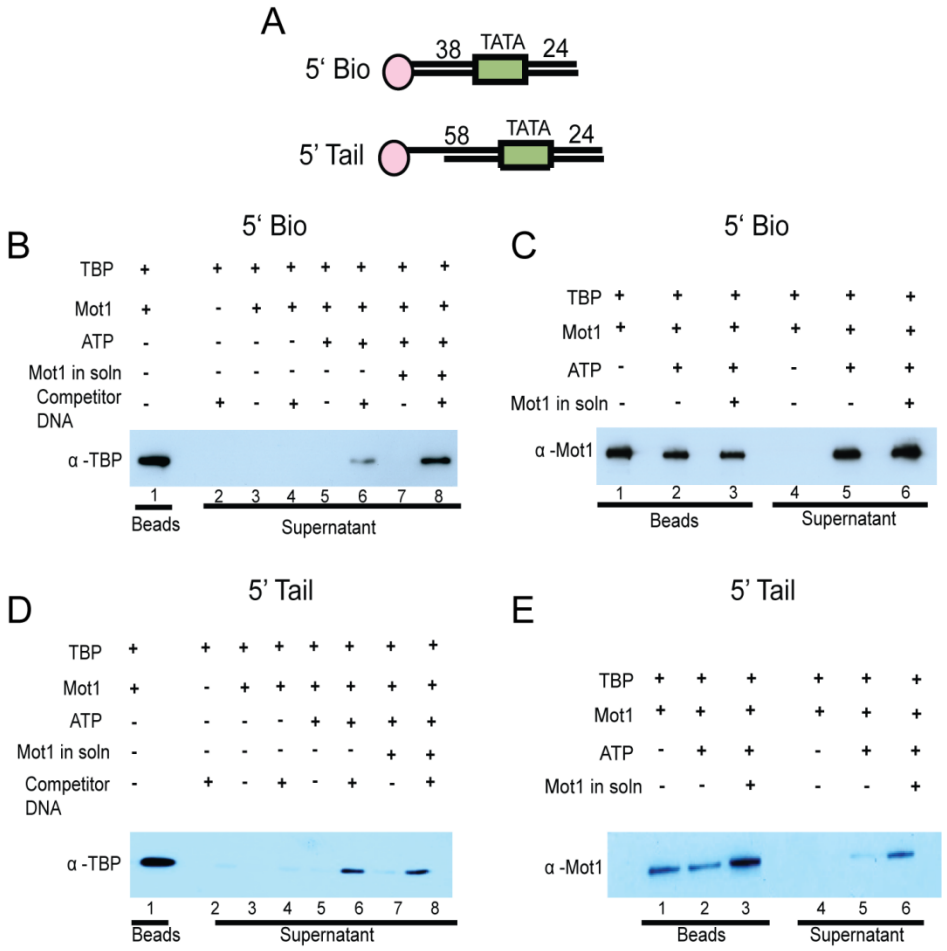

Figure 2.1. Mot1 is inefficient in purified complexes. *A*, Schematic of the oligonucleotide duplexes used for this experiment. The only difference between the 5’ Bio duplex and the 5’ Tail duplex is the presence of the 20 nucleotide single-stranded tail at the 5’ end of the 5’ Tail duplex. *B*, Western blot for TBP showing the presence of TBP on the beads or in the supernatant under the various conditions as indicated by the pluses and minuses above the blot for the 5’ Bio duplex *C*, Western blot for Mot1 similar to *B* for the 5’ Bio duplex. *D*, Western blot for TBP similar to *B* except using the 5’ Tail duplex. *E*, Western blot for Mot1 similar to *C* except using the 5’ Tail duplex.

Since it had been demonstrated by several studies that Mot1 requires upstream DNA to remove TBP (Auble et al. 1994, Auble and Steggerda 1999, Darst et al. 2001, Wollmann et al. 2011), we were concerned that the presence of the streptavidin bead might impact the activity of Mot1. Therefore, we introduced a 20 nucleotide single-stranded tail to the 5' end of the top strand to put some distance between the Mot1-TBP complex and the bead. As already stated, the remaining sequence was exactly the same as the 5' Bio duplex. We then repeated the experiment to determine if TBP or Mot1 dissociation was affected. When the Mot1-TBP-DNA complex was formed and ATP was added, there was no detectable TBP in the supernatant. When competitor DNA was added along with ATP, TBP was detected in the supernatant similar to the results for the 5' Bio duplex. This indicates that the streptavidin bead did not impact the activity of Mot1 and that the inefficiency of Mot1 is intrinsic to the purified complex. Interestingly, when ATP, competitor DNA, and excess Mot1 in solution were added to the 5' Tail duplex, there was no additional increase in TBP dissociation as observed for the 5' Bio duplex. Since the only difference was the presence of the single-stranded tail, it is possible that the excess Mot1 only increases the TBP displacement on shorter DNA duplexes. This will have to be further explored to understand exactly what this means.

Previously we had used the bead-bound assay along with the chemical FeBABE to demonstrate that the ATPase domain of Mot1 came into contact with the upstream DNA (Wollmann et al. 2011). Our lab also used this approach to more specifically map cleavages to the subdomains of the ATPase domain and identify a conformational intermediate trapped by the ATP analog ADP-AlF_4 (Maruta et al. 1993, Viswanathan 2013, Viswanathan et al. 2016). We determined if there was any change in the ATPase

conformation between binding of ATP (as mimicked by the ground state ADP-BeF_3) and hydrolysis (as mimicked by the transition state ADP-AlF_4) (Fisher et al. 1995, Liu et al. 2008). We also determined if this was a DNA sequence specific effect since we had only used one DNA sequence.

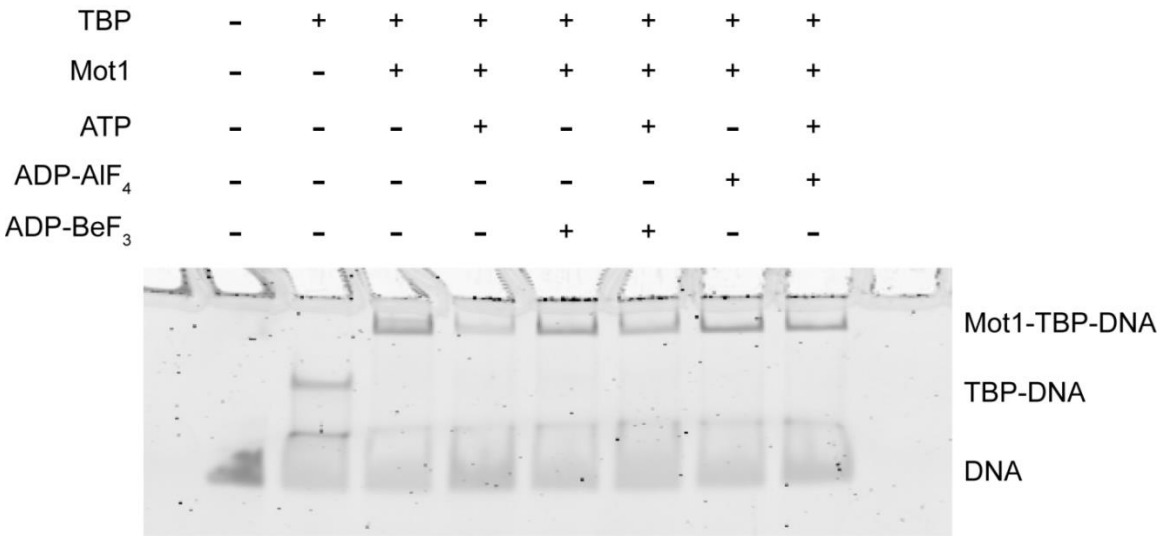

Figure 2.2. ATP analogs prevented ATP from binding to Mot1 and ~50% of TBP-DNA dissociation. Fluorescently labelled DNA was mixed with the indicated components to form complexes. Each complex is marked to the right. ATP analogs were added before ATP to allow for binding to the ATP binding pocket. Reactions were run on a non-denaturing gel to allow for separation of the various complexes. The DNA was visualized using a Typhoon PhosphorImager and ImageQuant software (GE Healthcare).

First, we determined if the ATP analogs bound to the ATPase domain of Mot1 and prevented ATP from binding. To do this we used the EMSA approach which allowed us to quickly identify if there was dissociation of TBP in the presence of the

analogs. As shown in Figure 2.2, when ATP was added to the Mot1-TBP-DNA complex there was ~80% dissociation of the complex. When either ADP-AlF_4 or ADP-BeF_3 was added to the complexes, there was only 18% or 26% dissociation of the Mot1-TBP-DNA complex respectively. When ATP was added after the analogs, there was 37% (ADP-AlF_4) or 45% (ADP-BeF_3) dissociation which was not as great as ATP alone. This indicates that the analogs prevented ATP from binding and ~50% of Mot1 dissociation of TBP.

Now that it was apparent that the ATP analogs would in fact prevent TBP dissociation, we analyzed their effects on the binding of Mot1 to various DNA sequences. The original sequence that we used was based on the Adenovirus Major Late Promoter sequence with phosphorothioates at positions 20 and 22 upstream of the TATA box (6Fe). We decided to use a sequence that we had used in a previous study (Moyle 6Fe; Moyle-Heyrman et al. 2012) and also the *HIS4* promoter sequence (*HIS4* 6Fe) with the phosphorothioates at the same positions, so that we could detect any changes strictly due to the sequence. As shown in Figure 2.3, the cyan arrow points to the unique cleavage product that had been previously observed in the presence of ADP-AlF_4 on the 6Fe probe. The same cleavage product was present in the presence of ADP-BeF_3, although it was present at a lesser amount compared to ADP-AlF_4.As shown in the EMSA (Figure 2.2), there was more dissociation of TBP when ADP-BeF_3 and ATP were added compared to when ADP-AlF_4 and ATP were added together. This could indicate that ADP-BeF_3 did not bind to Mot1 as well as ADP-AlF_4.

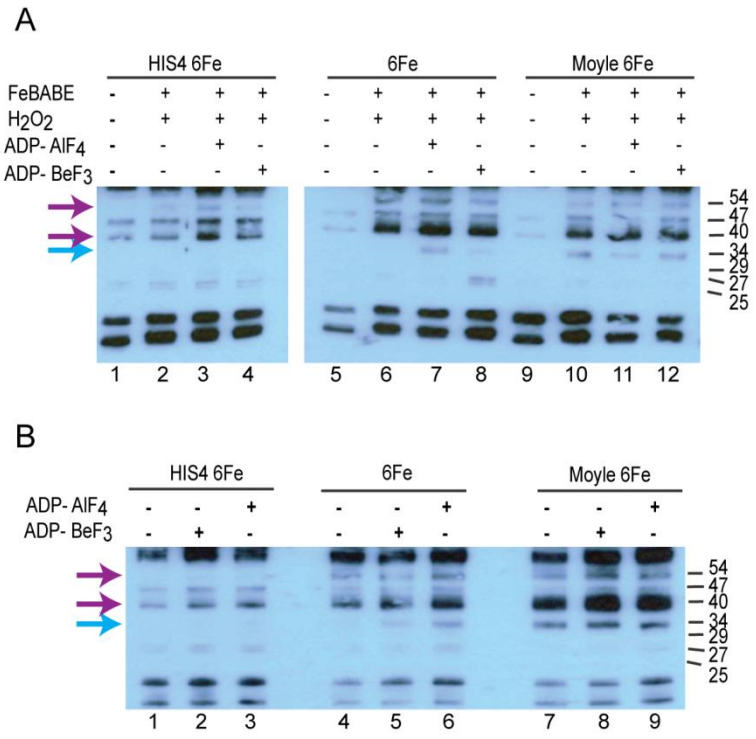

Figure 2.3. DNA sequence affects the binding of the ATPase domain of Mot1 to DNA. Western blots showing the FeBABE-mediated cleavage of Mot1 in reactions containing streptavidin beads, phosphorothioate-conjugated biotinylated DNA, FeBABE, sodium ascorbate, H_20_2, TBP, and Mot1. ADP-AlF_4 and ADP-BeF_3 were added as indicated. The DNA duplexes used are marked on top. The molecular masses of the cleavage products are marked on the right. The cyan arrow marks a unique cleavage product at ~34 kDa, whereas the purple arrows mark cleavage products at ~40 and 54 kDa.

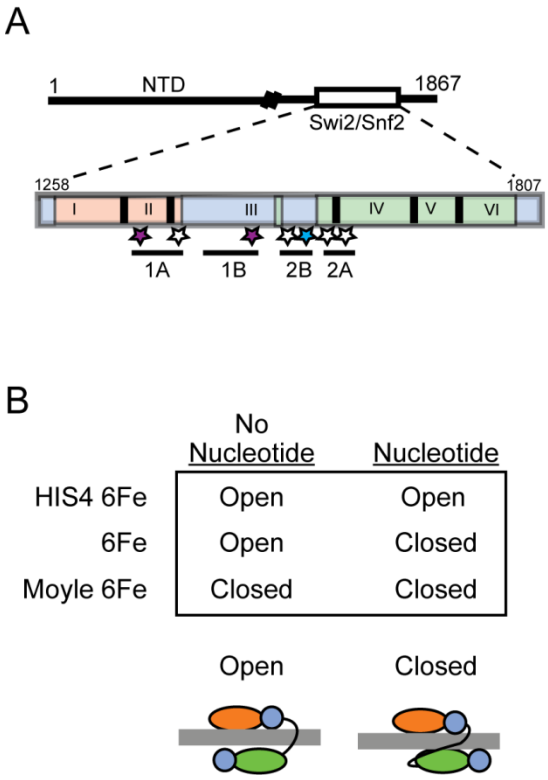

Figure 2.4. DNA sequence affects the conformation of Mot1 that binds upstream DNA. *A*, The primary structure of Mot1 is shown above with the Swi2/Snf2 ATPase blown up below. The seven classical helicase motifs are marked as I-VI (with I encompassing both I and Ia). The subdomains are marked below with the stars indicating the cleavage products from the FeBABE experiments. The cyan star corresponds to the cyan arrow and the purple stars correspond to the purple arrows in Figure 2.3. *B*, Summary of the FeBABE results from the three DNA sequences with a cartoon illustrating the organization of the ATPase domain in the open and closed forms; subdomain 1A (orange), 2A (green), and both 1B and 2B (blue).

Surprisingly, both the *HIS4* 6Fe probe and the Moyle 6Fe probe had different abundances of cleavage products. The cleavage product marked by the cyan arrow was absent for the *HIS4* 6Fe probe in the absence or presence of both ATP analogs. In contrast, this product was present for the Moyle 6Fe probe in the absence or presence of both ATP analogs. This was reproducible as illustrated by the Westerns in *A* and *B*. The purple arrows mark two different cleavage products that differed between the *HIS4* 6Fe probe and the 6Fe and Moyle 6Fe probes (Figure 2.3). This suggests that the DNA sequences can affect how the ATPase domain of Mot1 binds to the upstream DNA.

When these cleavage products were mapped onto the primary sequence of Mot1, they all mapped to the ATPase domain, which was similar to our previous FeBABE results (Figure 2.4*A*; Wollmann et al. 2011). Specifically, the cleavage product marked by the cyan arrow mapped to the 2B subdomain and the cleavage products marked by the purple arrows mapped back to subdomains 1A and 1B. A summary of the results from the three DNA sequences is shown in Figure 2.4*B*. Overall, Mot1 binds to the *HIS4* 6Fe probe in the open conformation regardless of presence of nucleotide, whereas it binds in the closed conformation to the Moyle 6Fe probe. The 6Fe probe favors the open conformation in the absence of nucleotide and the closed conformation in the presence of nucleotide. The differences in the preference for the conformations of Mot1 will be discussed more in the Discussion section of this chapter.

Though Mot1 bound in the open conformation to the HIS4 6Fe probe, previous experiments demonstrated that Mot1 was capable of removing TBP from this DNA sequence. Therefore, we predicted that the closed conformation would be able to form on this probe in the presence of nucleotide. We exposed the Western blots for the *HIS4* 6Fe (Figure 2.5*A*) and 6Fe (Figure 2.5*B*) probes longer to determine if we could detect

the presence of the closed conformation. After exposure of the Western blot for 1 minute, the cleavage product marked by the cyan arrow was not detected. However, exposure for 2.5 minutes yielded detection of this cleavage product in the presence of ADP-AlF_4. Further exposure (5 minutes) enabled detection of this band in the presence of ADP-BeF_3. This cleavage product was also detectable in the absence of nucleotide after exposure for 5 minutes, although the other bands are overexposed, indicating that this is a very small fraction of protein. Results were similar for the 6Fe probe (Figure 2.5*B*) showing that this cleavage product was detectable at higher exposures (2.5 and 5 minutes) even in the absence of nucleotide.

Finally, we used the information discovered from the FeBABE experiments and combined it with structural information to better understand the conformational change. So far the only structures of Mot1 that have been published have only included the N-terminal domain in complex with TBP and in complex with TBP, NC2, and DNA (Wollman et al. 2011, Butryn et al. 2015). Thus, the published literature is currently lacking a structure of the ATPase domain of Mot1 and the full Mot1 protein. However, previously similar Swi2/Snf2 ATPase domains have been crystallized. The Rad54 ATPase from *Sulfolobus solfataricus* has been crystallized in the open conformation (Dürr et al. 2005), and the Rad54 ATPase from *Danio rerio* has been crystallized in the closed conformation (Thoma et al. 2005). Since the Swi2/Snf2 ATPase domains are extremely well conserved, we used these structures to model the Mot1 ATPase domain in both the open and closed conformations (Figure 2.6).

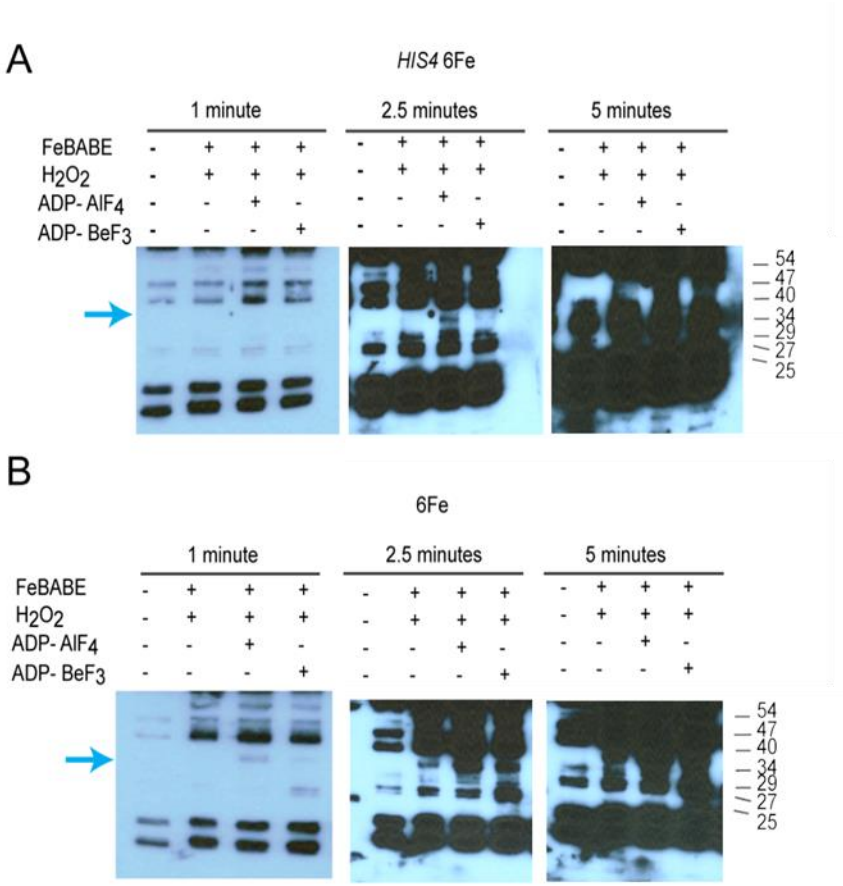

Figure 2.5. Mot1 preferentially binds to *HIS4* 6Fe and 6Fe probes in the open conformation but a small proportion can bind in the closed conformation. *A*, Western blot similar to Figure 2.3. Increasing exposures of Western blots showed cleavage products for the *HIS4* 6Fe probe show the cleavage product marked by the cyan arrow gradually increase indicating that a proportion of Mot1 can bind in the closed conformation on this template in the presence or absence of nucleotide. *B*, Western blot similar to *A* but for the 6Fe probe. Increasing exposures show that the cleavage product marked by the cyan arrow is present in a lower proportion even in the absence of nucleotide.

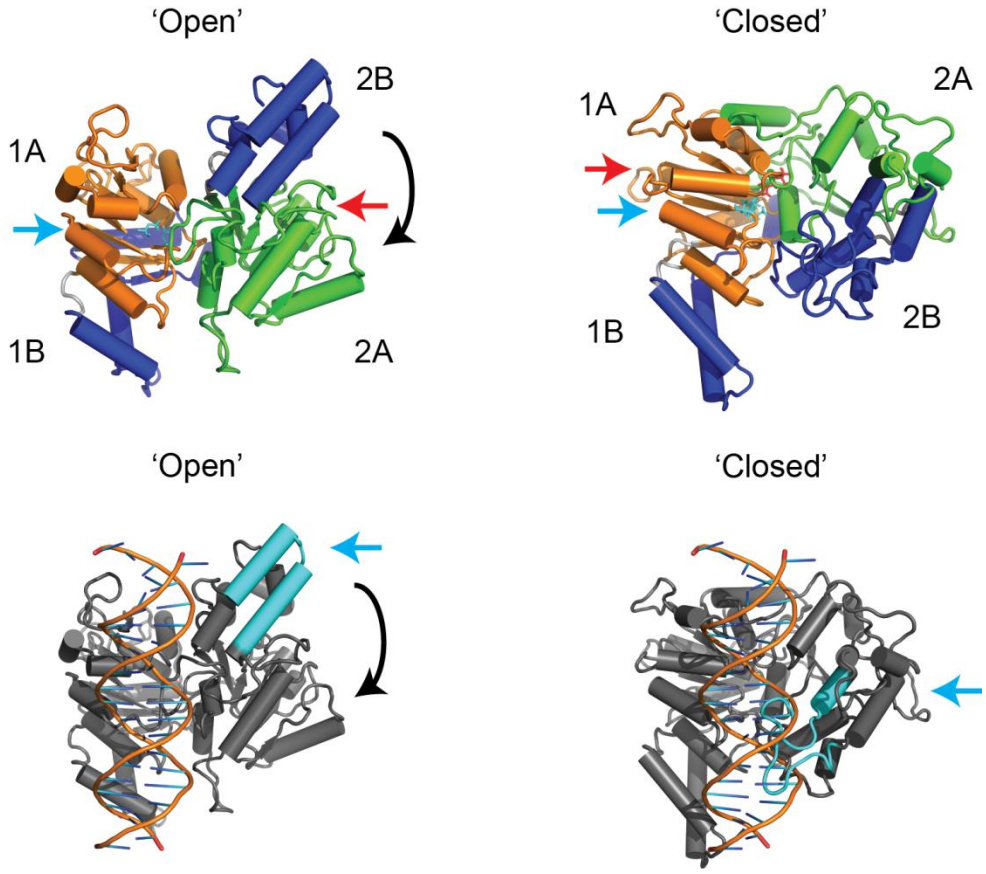

Figure 2.6. Models of the ATPase domain of Mot1 in the open and closed conformations without and with DNA. Molecular models of the Mot1 ATPase domain were generated to better understand the two conformations of Mot1. *A*, The open conformation was modeled on Rad54 from *Sulfolobus solfataricus* in the open conformation (1Z63) using CPHmodels. Subdomains are colored according to Dürr et al. 2005. The catalytic residues of the ATP-binding pocket are marked by the cyan and red spheres and arrows. *B*, The closed conformation was modeled on Rad53 from *Danio rerio* in the closed conformation (1Z3I) using PROTEUS2. The subdomains are colored similar to *A*. *C*, The modeled open structure of the ATPase domain is colored in gray and docked

onto DNA. The cyan region corresponds to the cleavage product marked by the cyan arrow in Figures 2.3 and 2.5. *D*, The closed model docked on DNA similar to *C*. The cyan region now comes into contact with the upstream DNA labeled with phosphorothioates in the 6Fe duplexes. The models were generated by David Auble.

In Figure 2.6*A*,*B* the subdomains are colored differently; subdomain 1A is orange, subdomain 2A is green, and subdomains 2B and 2B are both blue. Three of the residues important for the active site have been highlighted in both models of the ATPase (red and cyan spheres; Dürr et al. 2005). In the open conformation (Figure 2.6*A*), the cyan active site residues are apparent in subdomain 1A, whereas the red spheres are hidden behind subdomain 2A. After subdomain 2B rotates (along the black arrow; Figure 2.6*B*), the red residues are now visible and located next to the cyan residues forming the pocket of the active site.

In Figure 2.6*C*,*D*, the ATPase domain is colored gray and has been docked onto DNA (orange and blue). Since the ATPase domain is known to contact the upstream DNA (Wollmann et al. 2011), this docking allows for easy visualization of how the subdomains rotate along the DNA. We determined where the cleavage site indicated by the cyan arrows in Figures 2.3 and 2.5 mapped to in the Mot1 sequence and the standard deviation from two experiments; this cleaved area is indicated by the cyan color in Figure 2.6. This site is located in subdomain 2B as previously mentioned, and docking it on the DNA allows for a better understanding of why this area was cleaved by the FeBABE reaction. In the open conformation (Figure 2.6*C*), the cyan region is

located further away from the DNA. After subdomain 2B rotates down, the cyan region is now perfectly located along the upstream DNA in the region where the FeBABE was conjugated to the DNA in the 6Fe probes. This provides a clear explanation for why this region is cleaved when the ATPase domain is in the closed conformation instead of the open conformation.

Discussion

Mot1 Activity in a Purified Complex Mot1 was originally identified biochemically as the ATP-Dependent Inhibitor of TBP (ADI; Auble and Hahn 1993, Auble et al. 1994). Since then, several labs have gone on to biochemically ascertain the properties of Mot1 *in vitro*. Among these labs commonly EMSAs were utilized to detect the various complexes of Mot1, TBP, DNA, and other factors. When ATP was added to the Mot1-TBP-DNA complex there was almost complete dissociation of the ternary complex yielding the separate components (Auble et al. 1994, Poon et al. 1994, Adamkewicz et al. 2000). Our setup differs from the typically used EMSA, in that we used purified complexes instead of having a mixture of the protein components during the reaction. Since the stepwise assembly of the Mot1-TBP-DNA complex is well-documented, we assembled it in the proper order to yield the Mot1-TBP-DNA complex bound to the streptavidin beads. Surprisingly, the addition of ATP did not yield detection of any TBP removed from the complex. However, an earlier study suggested that TBP could rebind to DNA, thus preventing our detection of TBP in the supernatant (Sprouse et al. 2006). Nevertheless, addition of a competitor DNA and excess Mot1 in solution only yielded ~40% of the TBP being removed from DNA. As shown in Figure 2.1, I was able to get

~80% of TBP removed from DNA in the typical EMSA format, whereas others are able to produce near complete dissociation of TBP.

So why would Mot1 behave differently using the purified component approach versus using a mixture of the proteins? One possibility is illustrated by the requirement for the additional Mot1 in solution. It has been previously shown that Mot1 requires multiple molecules of ATP to remove TBP (Sprouse et al. 2008b). Upon hydrolysis of ATP, Mot1 falls off the DNA even if it has not removed the TBP. This suggests that increasing the local concentration of Mot1 in solution with certain DNA templates enabled rebinding of Mot1 to the DNA enabling multiple rounds of ATP hydrolysis resulting in TBP being dissociated from the DNA.

Another possibility is that more than one molecule of Mot1 is required to adequately remove TBP from DNA. An unpublished crystal structure of a Mot1 dimer has also led to speculation that Mot1 could function as a dimer (Butryan and Hopfner, unpublished results). As mentioned in Chapter I, Mot1 can function as a single polypeptide, whereas the other members of the Swi2/Snf2 ATPase family require other subunits to perform their functions (Auble et al. 1994, Poon et al. 1994, Eisen et al. 1995, Adamkewicz et al. 2000). Is it possible that more than one molecule of Mot1 is required to remove TBP from DNA rather than a requirement for additional proteins? We could not identify the requirement for multiple molecules of Mot1 using the typical EMSA approach where Mot1 removed almost all of the TBP. The bead-bound approach, however, allowed for us to demonstrate that multiple molecules of Mot1 were required. More work would need to be performed to determine if this idea is correct or not. EMSAs and bead-bound experiments use purified Mot1, TBP, and DNA; however, Mot1 is exposed to other proteins *in vivo*, which means that other proteins

could affect Mot1. Recently, a Mot1-NC2-TBP-DNA structure was published which had been previously demonstrated biochemically by other experiments (Darst et al. 2003, Butryn et al. 2015). These experiments and this structure suggest that the NC2 complex can bind TBP-DNA and allow for Mot1 to bind. Thus it is a possibility that Mot1 does not function solely on TBP-DNA complexes but also on NC2-TBP-DNA complexes. Experiments have been performed to determine if NC2 affects the activity of Mot1 in any way, but so far there have been no definitive results suggesting that NC2 modulates Mot1 activity (Butryn et al. 2015). Even though NC2 does not affect Mot1 activity, Mot1 does interact with other proteins *in vivo*, which leaves other interactions to be explored for their effects on Mot1 (Arnett et al. 2008).

DNA Sequence Affecting Mot1 Conformational Change

Here we showed the presence of two conformations of Mot1 based on our FeBABE approach (Viswanathan et al. 2016). Earlier it was suggested that Mot1 could adopt two different conformations based on results that showed two different types of Mot1-TBP-DNA complexes (Sprouse et al. 2008b). It was proposed that one of these conformations could even act in an ATP-independent manner to dissociate TBP from DNA. The results shown in this chapter further show evidence that Mot1 forms the open and closed conformation similar to other ATPases. The conversion from the open to the closed conformation is required to form the pocket required to bind ATP. Our previous identification of both conformations utilized an ATP analog, ADP-AlF_4, which mimics ATP having been hydrolyzed but not released from the binding pocket (Viswanathan et al. 2016).

The results shown here demonstrate that the ground state ATP analog, ADP-BeF_3, led to similar conformational changes as ADP-AlF_4 (Maruta et al. 1993, Fisher et al. 1995, Liu et al, 2008). This indicates that the binding of nucleotide and hydrolysis do not have significant differences in conformation of the ATPase domain. However, interestingly we observed different proportions of the conformations of the ATPase domain using various DNA templates. The original template, 6Fe, is a TATA-containing sequence based on the Adenovirus Major Late Promoter. The other two templates we used had the same TATA box sequence and phosphorothioates positioned at the same positions upstream of the TATA box. A difference between the templates was the composition of the DNA sequence upstream of the TATA box, which comes into contact with the ATPase domain and could influence the conformation in which this domain binds DNA.

Cleavage of Mot1 on the Moyle 6Fe template in the absence or presence of ATP analogs resulted in a cleavage pattern similar to 6Fe in the presence of ATP analogs. This suggests that Mot1 binds to the Moyle 6Fe template in the closed state regardless of the presence of nucleotide. In contrast, the cleavage pattern of Mot1 on the *HIS4* 6Fe template in the absence or presence of ATP analogs was similar to 6Fe in the absence of ATP analogs. This suggests that Mot1 binds to the *HIS4* 6Fe sequence in the open conformation. More on the *HIS4* 6Fe sequence will be discussed in the next section.

As already mentioned, the three templates differ in the upstream DNA sequences. So how and why would Mot1 binding differ at these various sequences? It is possible that Mot1 has some sequence selectivity that is being illustrated by these different sequences, though previous experiments suggest otherwise (Darst et al. 2001). Moreover, the SsoRad54 ATPase domain interacted with the phosphate backbone of the

DNA, which indicates that the Mot1 ATPase does not have sequence specificity (Durr et al. 2005). It is more likely that Mot1 does not require a specific sequence in the upstream DNA but rather favors different properties of the DNA to determine which conformation binds. This idea is supported by the fact that the GC content of the upstream DNA differs between the three sequences; *HIS4* 6Fe has a GC content of 37%, 6Fe has 68% GC, and Moyle 6Fe has 89% GC. This increase in GC content corresponds to the increase in the closed conformation of Mot1; *HIS4* 6Fe shows no/little of the closed conformation, 6Fe shows the closed conformation in the presence of nucleotide, and Moyle 6Fe shows the closed conformation regardless of nucleotide. The GC content of DNA is well known for decreasing the flexibility of DNA (Travers 2004), so it seems plausible that this is what is affecting the conformation of Mot1. The open conformation could favor more flexible regions of DNA, whereas the closed conformation could favor less flexible regions.

The next question is why these sequences would affect Mot1 binding to DNA. We know that Mot1 preferentially represses TATA-containing promoters and activates TATA-less promoters (Collart 1996, Dasgupta et al. 2002, Poorey et al. 2010, Venters et al. 2011, Zentner and Henikoff 2013). But could the difference in conformational binding also play in to the mechanism of Mot1 *in vivo* and explain how Mot1 distinguishes between gene activation and repression? Previously it was shown that the two types of Mot1-TBP-DNA complexes had different kinetics, with the ATP-independent complex being less stable (Sprouse et al. 2008b). It is possible that the less stable ATP-independent complex is the closed conformation that does not require ATP to cause the transition between open and closed. The closed complex could already be primed for

removal of TBP, thereby potentially decreasing the stability of the ternary complex and increasing dissociation.

Differences in Mot1 Binding and Activity at the HIS4 Promoter

In EMSAs it has been demonstrated that Mot1 can remove almost all of the TBP bound to the *HIS4* promoter (Auble and Hahn 1993). However, the FeBABE experiment utilizing the *HIS4* promoter (Figure 2.3) showed that Mot1 bound in the open conformation even in the presence of nucleotide. As Mot1 was capable of removing TBP from this DNA sequence, Western blots were exposed longer (Figure 2.5) to determine if the cleavage site indicated by the cyan arrow would show up. In fact, using longer exposures, this cleavage site was eventually detected. This indicates that Mot1 preferentially bound to the *HIS4* sequence in a purified complex in the open conformation regardless of nucleotide. The addition of nucleotide did allow for a small proportion of Mot1 to transition to the closed conformation which could be responsible for the removal of TBP.

But why is there a difference between the EMSAs and the FeBABE results? This has already been demonstrated by the near complete removal of TBP by Mot1 as detected by EMSAs; whereas only 40% of TBP is removed using the 6Fe sequence in the bead-bound experiment. There are obviously differences between the two experimental setups that we do not fully understand how they impact Mot1 activity *in vitro*. The *HIS4* EMSA experiments allowed for multiple rounds of Mot1 to bind and hydrolyze ATP, whereas we did not perform a bead-bound using the *HIS4* sequence to see if excess Mot1 in solution would increase the amount of TBP dissociated. These

results could be compared to those from other sequences to determine if there was a difference specific to the *HIS4* sequence. A better quantification of the proportions of Mot1 in the open and closed conformation could be determined too. This would enable a better comparison of the ATPase conformations bound to the various sequences. Longer exposures of the Westerns in Figure 2.5 showed that the closed conformation of Mot1 is detectable on the *HIS4* template in the presence of nucleotide, which was not apparent in Figure 2.3. While differences in Mot1 activity between various DNA templates have not been observed using EMSAs, the bead-bound setup allows for detection of differences in Mot1 binding to the different sequences. Since the FeBABE results demonstrate differences in how Mot1 binds to various sequences, this could enable a better understanding of how Mot1 distinguishes between different promoters to either activate or repress transcription.

Chapter III

Mot1 and Spt16 Co-Regulate Transcription

Data from this Chapter was published in True et al. 2016.

Mot1 is a conserved and essential Swi2/Snf2 ATPase that can remove TATA-binding protein (TBP) from DNA using ATP hydrolysis and in so doing exerts global effects on transcription. Spt16 is also essential and functions globally in transcriptional regulation as a component of the FACT histone chaperone complex. Here we demonstrate that Mot1 and Spt16 regulate a largely overlapping set of genes in *Saccharomyces cerevisiae.* The two factors physically and genetically interact indicating a functional relationship. As expected, Mot1 was found to control TBP levels at co-regulated promoters. In contrast, Spt16 did not affect TBP recruitment. However, bothMot1 and Spt16 contribute to TFIIB localization, indicating a convergence on the formation of the preinitiation complex. These results suggest that the large-scale overlap in Mot1 and Spt16 function arises from a combination of both their unique and shared functions in transcription complex assembly.

Introduction

Mot1 Regulates Initiation and Elongation

Genome-wide analyses of expression changes in Mot1 mutants show that Mot1 is required for repression and activation of transcription as discussed in Chapter I (Collart 1996, Dasgupta et al. 2002, Dasgupta et al., 2005, Poorey et al. 2010, Zentner and Henikoff 2013). Repression of transcription by Mot1 is easily explained, since *in*

vitro it has been demonstrated that Mot1 removes TBP from DNA using ATP hydrolysis (Timmers et al. 1992, Auble and Hahn 1993, Auble et al. 1994, Adamkewicz et al. 2000). Removal of TBP would prevent the PIC forming at the promoter, thereby preventing transcription initiation. Several models have been put forward to explain how Mot1 could activate transcription (Collart 1996, Madison and Winston 1997, Muldrow et al. 1998, Gumbs et al. 2003, Topalidou et al. 2003, Zentner and Henikoff 2013). The current favored model of Mot1-mediated activation is that Mot1 removes TBP from strong TBP-binding sites (TATA boxes) and allows them to be reallocated to weaker binding sites (TATA-less) (Zentner and Henikoff 2013). Other models include Mot1 participating in formation of an alternative PIC or even Mot1 directly or indirectly affecting chromatin organization. Recently it was shown that mutation of Mot1 leads to TBP being more dynamic on DNA as shown by the Crosslinking Kinetics analysis (CLK) (Poorey et al. 2013). This suggests that Mot1 stabilizes TBP at promoters. Genome-wide localization has also shown that Mot1 is predominately localized to promoters of genes it represses and activates, suggesting that Mot1 has a direct role in both types of regulation. This localization coincides with TBP localization at the promoters, and Mot1 has been shown to approach TBP from the upstream DNA, which is consistent with studies showing that the upstream DNA is required for Mot1 function and the results from Chapter II (Darst et al. 2001, Wollmann et al. 2011, Zentner and Henikoff 2013, Viswanathan et al. 2016).

Our lab developed a Matlab-based pipeline to identify four different classes of transcription length defects in mutant cells using data from tiling array analysis (Poorey et al 2010). The four length defects are shown in Figure 3.1 and include upstream initiation (UI), downstream initiation (DI), premature termination (PT), and downstream termination (DT). These defects are based on where the transcript starts and ends

compared to the normal transcript annotations. Mutation of Mot1 was shown to lead to a prominent transcription length defect of premature termination. Because of Mot1's regulation of TBP and localization to the promoter, it is interesting that Mot1 would be necessary for promoting productive Pol II elongation. One study did identify Mot1 localized to the coding regions of certain genes upon activation (Geisberg et al. 2002); however other studies have not yet corroborated this observation (Dasgupta et al. 2002, Zanton and Pugh 2004, Zentner and Henikoff 2013). It is possible that Mot1 indirectly regulates early elongation through an interaction with another protein, or even that it directly regulates elongation even though we do not detect high levels of Mot1 in the coding region.

In order to determine how Mot1 regulates transcription elongation and chromatin organization, we looked for another protein that might overlap with Mot1 and elucidate these functions.

Spt16, an Elongation Factor, Also Regulates Initiation

Spt16, a subunit of the FACT histone chaperone complex, has been implicated in both transcription initiation and elongation (Wittmeyer and Formosa 1997, Brewster et al. 1998, Formosa et al. 2002, Biswas et al. 2005, Jamai et al. 2009, Han et al. 2010, Foltman et al. 2013). Spt16 forms a dimer with Pob3 and together they loosely associate with multiple Nhp6 proteins to bind to a nucleosome (Formosa et al. 2001, Ruone et al. 2003). As mentioned in Chapter I, Spt16 has been widely studied as an elongation factor necessary for Pol II passage and replacement of histones to prevent cryptic initiation (Orphanides et al. 1998, Saunders et al. 2003, Mason et al. 2003,

Kaplan et al. 2003, Cheung et al. 2008, Formosa 2012). However, other studies have suggested that Spt16 also plays an important role in transcription initiation (Mason et al. 2003, Shimojima et al. 2003, Biswas et al. 2005).

Figure 3.1. Classification of transcription length defects. A gene transcript starts at the transcription start site (TSS) and ends at the poly(A) site. Mutation of transcription factors can cause the transcript to start or end at an improper location leading to transcription length defects. The four classes of defects that we have categorized are upstream initiation (UI), downstream initiation (DI) or cryptic initiation, premature termination (PT), and downstream termination (DT).

One study showed that TBP localization to a subset of promoters was perturbed in *spt16-197* cells (Biswas et al. 2005). Spt16 could indirectly regulate the PIC due to its role in preventing the exposure of cryptic initiation sites. Spt16 has also been shown to be important for reorganization of the promoter-proximal nucleosomes (specifically the +1 nucleosome) necessary for proper transcription initiation (Shimojima et al. 2003, Jimeno-González et al. 2006, van Bakel et al. 2013, Liu et al. 2014). Furthermore, Spt16 has a unique localization pattern compared to other elongation factors such as Bur1 and Spt6. Spt16 peaked immediately downstream of the +1 nucleosome, continued through the coding region, and occupancy declined before the termination site where the other elongation factors exited (Mayer et al. 2010). It is therefore possible that the unique genomic localization pattern of Spt16 is due to Spt16's role in initiation and early elongation.

Interestingly enough, a previous study using mass spectrometry identified Spt16 as a Mot1-associated factor (Arnett et al. 2008). The Mot1-Spt16 interaction was confirmed using a co-immunoprecipitation experiment. The proteomics screen also identified Pob3 as interacting with Mot1; however it was below the statistically significant threshold. The final component of the FACT complex, Nhp6, was not identified in this screen. This could indicate that Mot1 and Spt16 have a tighter association than Mot1 does with the other FACT components, and that Pob3 and Nhp6 are not required for the Mot1-Spt16 interaction. However, since Spt16 is not known to function in a complex separate from Pob3, it is also possible that Pob3 plays a role in the Mot1-Spt16 interaction.

Scope of this study

Transcription is a complex process requiring various points of regulation to ensure that the genome is properly transcribed. Mot1 and Spt16 are essential factors in yeast that share a regulatory role in TBP localization to the promoter. Mot1 was initially studied as an initiation factor due to its localization to the promoter; however, the discovery that *mot1-42* cells display premature termination defects suggests that Mot1 also regulates early elongation. Spt16 was considered to be an elongation factor only until it was shown that Spt16 not only regulates TBP localization but also promoter-proximal nucleosomes. It is therefore possible that these two essential factors overlap in transcription regulation, and this interaction could help to explain how Mot1 regulates elongation and how Spt16 regulates initiation. A previous mass spectrometric analysis that identified Spt16 as a Mot1-interacting factor showed that Spt16 and Pob3 both physically associate with Mot1. We hypothesized that Mot1 and Spt16 functioned together in regulating gene expression through their combined actions on initiation and elongation. This chapter focuses on the transcriptional overlap between these two factors and defining their effects on PIC components.

Materials and Methods

Yeast strains and growth conditions

Saccharomyces cerevisiae strains used in this study are listed in Table 1. The wild type (WT) strain AY51 is referred to as *MOT1*-WT, since it was used to generate the *mot1-42* (AY87) strain; these strains were described previously (Sikorski and Hieter 1989, Darst et al. 2003). FY56 is referred to as *SPT16*-WT, since it was used to

generate the *spt16-197* strain (L577); both of these strains were provided by Fred Winston (Malone et al. 1991). The *mot1-42 spt16-197* double mutant was constructed by mating the *spt16-197* strain with a *MOT1* shuffling strain (AY138). Diploids were sporulated, and tetrads were dissected and screened for temperature sensitivity and resistance to kanamycin. Candidates from the screen were then transformed with plasmid pMOT221 (Darst et al. 2003) containing the *mot1-42* allele, and the *MOT1*-WT plasmid was shuffled out by plating on synthetic complete medium without leucine and containing 5-fluoroorotic acid (FOA). *MOT1*-WT, *mot1-42*, *SPT16*-WT, and *spt16-197* cells were all grown at 30°C in yeast extract, peptone, dextrose media (YPD), while *mot1-42 spt16-197* cells were grown at 25°C in YPD. For the RNA isolation and ChIP experiments, all strains were grown at their permissive temperatures to an optical density (OD) ~1.0, heat-shocked with addition of an equal volume of 42°C pre-warmed YPD, and incubated at 35°C for 45 minutes. For spot assays, strains were grown to an OD ~1.0 at their permissive temperatures and then 10-fold serial dilutions were plated on YPD and incubated at 25°C, 30°C, or 35°C for three days.

Table 3.1

Yeast strains used in this chapter

Strain name	Genotype	Reference
YPH499	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1*	Sikorski *et al.* 1989
AY51	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1. mot1Δ::TRP1 pAV20(EE-MOT1, LEU2, CEN ARS)*	Dasgupta *et al.* 2005
AY87	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1. mot1Δ::TRP1 pMot221(mot1-42, LEU2, CEN ARS)*	Dasgupta *et al.* 2005
AY138	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1. mot1Δ::kanMX pMR13(MOT1, URA3, CEN ARS)*	Auble *et al.* 1997
FY56	*MATα, his4-912δ, lys2-128δ, ura3-52*	Malone *et al.* 1991
L577	*MATα, his4-912δ, lys2-128δ, ura3-52, spt16-197*	Malone *et al.* 1991
YMW066	*MATa, ura3-52, trp1-Δ63, his3-Δ200, spt16-197, mot1Δ::kanMX, pMOT221(mot1-42, LEU2,CEN ARS)*	This study
YMW070	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1, MOT1-13XMYC(HIS3), SPT16-HA(TRP1)*	This study

Co-Immunoprecipitation

Yeast strains were grown as above to an OD ~1.0, and cells were then collected and lysed in BA/150 lysis buffer (20 mM HEPES-KOH pH 7.6, 2 mM EDTA, 2 mM EGTA, 0.25% NP-40, 150 mM potassium acetate, 5 mM DTT, and a protease inhibitor cocktail tablet (one tablet per 25 mL buffer) (Roche)). A total of 1 mg protein was immunoprecipitated overnight at 4°C with 15 µL α-myc (9E10, 1 mg/mL), 2.5 µL α-Spt16 (rabbit polyclonal from Tim Formosa), 2.5 µL α-Pob3 (rabbit polyclonal from Tim Formosa), or 2.5 µL α-Nhp6 (rabbit polyclonal from Tim Formosa). Immunoprecipitated samples were incubated with Protein A Sepharose beads (GE Healthcare) and washed with BA/250 lysis buffer (same as BA/150 lysis buffer but with 250 mM potassium acetate). Samples were separated on 4-15% pre-cast Bio-Rad SDS-PAGE gels. After transfer to nitrocellulose membranes, samples were blocked in tris-buffered saline with tween (50 mM Tris HCl, pH 7.6; 150 mM NaCl, 0.05% Tween 20) plus 5% nonfat milk, then incubated with either α-myc (9E10, 1 mg/mL), α-HA (12CA5 Abcam), or α-Spt16 (rabbit polyclonal antibody), followed by incubation with secondary antibody, and developed using the ECL Prime kit (GE Healthcare).

RNA isolation for tiling array analysis and RT-PCR

Strains were grown as above in duplicate for tiling array analysis and in triplicate for RT-PCR. Total RNA was isolated and samples were prepared for tiling array analysis as previously described (Poorey et al 2010). The University of Virginia Microarray Core Facility hybridized the samples to *S. cerevisiae* 1.0 Tiling Arrays (Affymetrix) and generated the raw data. After the total RNA was isolated for RT-qPCR

analysis, the RNA was reverse transcribed with the iScript Select cDNA Synthesis kit (Bio-Rad) according to the manufacturer's instructions, with the exception that twice the amount of starting RNA and reagents were used. Real-time PCR was carried out as previously described (Poorey et al. 2010). Oligonucleotides used are listed in Table 2.

Table 3.2
Primer sequences used in this study

Primer Name	Sequence
RPS23B Promoter F	5' TGC TAA GCA CTA CCG CAT TG 3'
RPS23B Promoter R	5' GAA AGC GTG GAG ACA AGG AG 3'
TRX1 Promoter F	5' CCA AAA CCC TGA AAC TGC AT 3'
TRX1 Promoter R	5' ATT CGC TGG CAG TTT TGA AT 3'
ACT1 Promoter F	5' CAG CTT TTA GAT TTT TCA CGC TTA 3'
ACT1 Promoter R	5' TTT TCG ATC TTG GGA AGA AAA A 3'
INO1 Promoter F	5' GTT GGC GGC AAT GTT AAT TT 3'
INO1 Promoter R	5' CGA CAA CAG AAC AAG CCA AA 3'
GAD1 Promoter F	5' CAC TGA ACT GCA ACG CAC TC 3'
GAD1 Promoter R	5' TTT TAG CAT CGC CAA AAG GT 3'
GND2 Promoter F	5' CGT CAG AAA TTG AAC GTT TCC 3'
GND2 Promoter R	5' GGC ACT CGT GGT TAA AGA GC 3'
RPS23B ORF F	5' GAA CCA CCG AAT GGA GAA GA 3'
RPS23B ORF R	5' GTT GGG CCG AAA ACA ACT AT 3'
TRX1 ORF F	5' GAA GTT GCA AAG GTT GTT GG 3'
TRX1 ORF R	5' TTA GCA GCA ATG GCT TGC TT 3'
LEU9 ORF F	5' AGA AAT TGA ACC CAG AGC GT 3'
LEU9 ORF R	5' TTC AGT AAA TTG GAT AGC GCA 3'
ACT1 ORF F	5' GCA AAA GGA AAT CAC CGC TT 3'
ACT1 ORF R	5' AAG CCA AGA TAG AAC CAC CAA 3'
PCF11 ORF F	5' GGT TGA ATC CTA ATG ACA CCG 3'
PCF11 ORF R	5' TTG CTT GCA GGT TTT TCT GG 3'
INO1 ORF F	5' GGC TGA GCA TGA GGG TAC AT 3'
INO1 ORF R	5' CAA CTT GGT TTG TCC CGA CT 3'
GAD1 ORF F	5' AAC GGA TGG ATC CGA TGA GAA 3'
GAD1 ORF R	5' CAT CTC CGA TCT GAA AAC CA 3'
GND2 ORF F	5' AAT GTG GAG AGG TGG CTG TAT 3'
GND2 ORF R	5' GAA GCG AAG AAC TCG TTG AA 3'

Tiling Array Data Analysis

Integrated Genome Browser compatible files were produced from the raw tiling array data using Tiling Analysis Software (TAS; see Appendix B). Signal profiles were smoothed over a 50 base pair (bp) bandwidth (101 bp sliding window), and differential profiles were smoothed over a 250 bp bandwidth (501 bp sliding window). The Mot1 differential signal was generated by subtracting *mot1-42* signal ($\log_2$ scale) by *MOT1*-WT signal. The Spt16 differential signal was generated by subtracting *spt16-197* signal by *SPT16*-WT signal. Two double mutant differentials were generated by subtracting the *mot1-42 spt16-197* signal by each of the two WT signals. Two types of analysis were conducted with the differential profiles: unbiased and biased. The unbiased analysis considered genome-wide differential expression, whereas the biased analysis considered only the 6,685 annotated genes. For the biased analysis, median differential expression was calculated for each gene, and for the 309 intron-containing genes only the longest exon was considered. Because each mutant produces different magnitudes of changes in transcript levels, different thresholds were applied to each differential profile for classifying genes as under-expressed or over-expressed. These thresholds were determined by comparing the distributions of median differential expression to a normal distribution, and identifying the levels above or below which median gene expression deviated from the normal distribution using quantile-quantile plots. Because both WT strains had relatively high correlations between the datasets (r = 0.74), further analyses for the single and double mutants were conducted using comparisons to *MOT1*-WT.

Chromatin Immunoprecipitation

Yeast strains were grown as described above. After cultures reached an OD ~1.0 and were heat shocked, cells were collected, lysed, and sonicated as previously described (21). Immunoprecipitations were conducted using 1 mg of protein as previously described with either 2.5 µL α-TFIIB (Sua7 rabbit polyclonal) or 5 µL α-TBP (58C9 Abcam). After immunoprecipitation and crosslink reversal, DNA was purified using a Qiagen PCR Purification kit according to the manufacturer's instructions. Purified DNA was then quantified by qPCR using the oligonucleotides (Life Technologies) listed in Table 2 and SYBR Green master mix (Bio-Rad). The mock values were subtracted from the IP values and then normalized to the input.

Analysis of RNA length precision

Categories of RNA length defects were assigned to genes using our previously published method (Poorey et al. 2010). Intron-containing genes were excluded from the analysis. For the *spt16-197* and *mot1-42 spt16-197* differential RNA signals compared to WT, the signal cutoff was raised from 0.30 to 0.36 to account for overall greater magnitudes of differential expression, and to reduce the likelihood of calling false positives. The other parameters were held constant.

All of the genome-wide datasets from this study have been submitted to the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under SuperSeries accession number GSE80235.

Results

In order to validate the Mot1-Spt16 interaction, identified in a previous study (Arnett et al. 2008), a strain was engineered containing *MOT1*-myc and *SPT16*-HA. Mot1-myc or Spt16-HA was immunoprecipitated from the whole cell extracts (WCEs) from these strains. The immunoprecipitated samples were separated on SDS-PAGE gels, transferred to nitrocellulose membranes and probed with antibodies for either myc or HA. Both Mot1-myc and Spt16-HA were detectable in the WCE (Fig. 3.2*A*). The previous study showed that Mot1 and Spt16 co-associate and we were able to recapitulate that result (Fig. 3.2*A*). We also performed a reciprocal IP using a polyclonal Spt16 antibody in WCEs containing Mot1-myc and Spt16-HA. Though this experiment validated the physical association between Mot1 and Spt16 in WCEs, this does not necessarily indicate a direct interaction, since there could be an intermediate factor bridging the two factors.

We next determined if Mot1 interacted with the entire FACT complex or with Spt16. The previous proteomics study showed that Pob3 does indeed interact with Spt16, though this interaction was below their significance threshold (Arnett et al. 2008). Nhp6 was not identified as a Mot1-interacting factor previously. We immunoprecipitated for either Pob3 or Nhp6 which led to pulldown of Mot1-myc, indicating that Mot1 associated with the entire FACT complex (Figure 3.3). Unfortunately, the reciprocal IP for Mot1-myc probed for Pob3 or Nhp6 did not verify this interaction. It is possible that there were some technical problems that could have precluded this experiment from working. It is intriguing that the proteomics study reported that immunoprecipitating for Mot1 yielded a less than significant interaction between Mot1 and Pob3, similar to our

problem showing this interaction. This could indicate that Mot1 and Spt16 are more closely associated than Mot1 is with Pob3 and Nhp6.

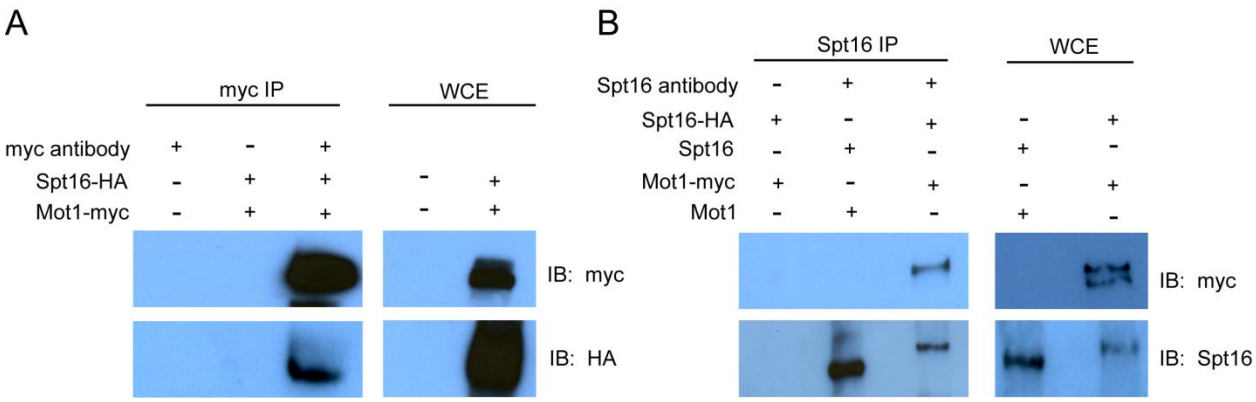

Figure 3.2. Mot1 and Spt16 physically interacted. *A*, Co-immunoprecipitation (Co-IP) of Spt16 and Mot1 in whole cell extracts (WCEs) obtained from the strains with HA-tagged Spt16 and/or myc-tagged Mot1 as indicated. IPs were performed using the myc antibody, and blots were probed with the antibody indicated to the left of each panel. WCE lanes show results with whole cell extracts used as input in the IPs. Western was performed by Melissa Carver. *B*, A similar experiment as *A* using a polyclonal Spt16 antibody in the IP.

Next, we determined if there was a genetic interaction between the two genes in order to identify a functional significance in the physical association between the two proteins. Both *MOT1* and *SPT16* are essential genes and thus warrant the use of the temperature sensitive (ts) alleles *mot1-42* and *spt16-197*. The single mutants were crossed to generate a *mot1-42 spt16-197* double mutant. All three mutants grew on YPD at 25°C comparably (Fig. 3.4 left). However, *spt16-197* cells grew slightly better than the other strains presumably due to a difference in genetic background. Both *mot1-*

42 and *spt16-197* cells were previously shown to be ts at 35°C, consistent with Fig. 3.4 (right). While the single mutants grew well at 30°C, the double mutant was synthetically sick (Fig. 3.4 middle). This synthetic sickness indicated a genetic interaction between *MOT1* and *SPT16* suggesting that the gene products are required for similar functions.

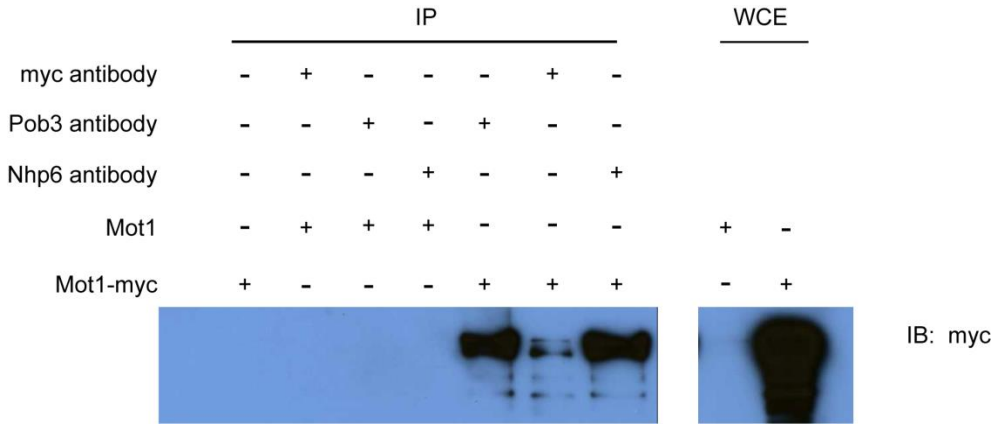

Figure 3.3. Mot1 physically interacted with FACT components Pob3 and Nhp6. Co-immunoprecipitation (Co-IP) of Mot1 and Pob3 or Nhp6 in whole cell extracts (WCEs) obtained from the strains with myc-tagged Mot1 as indicated. IPs were performed using a myc antibody for Mot1, a polyclonal Pob3 antibody, or a polyclonal Nhp6 antibody. Blots were probed with the myc antibody as indicated to the left of each panel. WCE lanes show results with whole cell extracts used as input in the IPs.

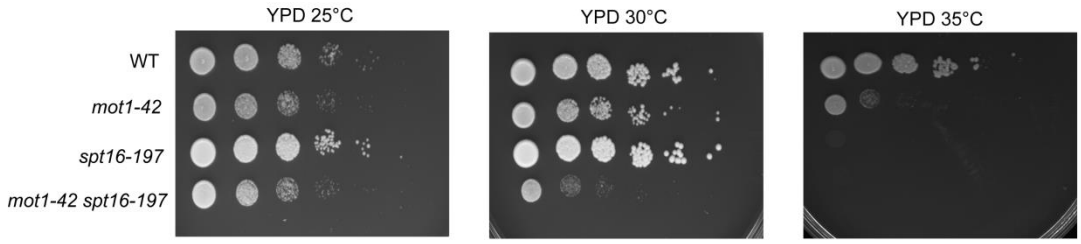

Figure 3.4. Mot1 and Spt16 genetically interacted. Ten-fold serial dilutions of the indicated strains were plated on rich media (YPD) and incubated at the indicated temperatures for three days. The *mot1-42 spt16-197* strain exhibited synthetic sickness at 30°C compared to the other strains.

After confirming the physical and genetic interaction between Mot1 and Spt16, next we determined if there was an overlap in transcription regulation. A previous study used tiling arrays to identify cryptic initiation events in *spt16-197* cells (Cheung et al. 2008). We compared this data to the data generated from our previous tiling array analysis of expression defects in *mot1-42* cells (Poorey et al. 2010). There was a weak correlation between the two datasets, and due to the differences in analysis performed we compared the expression changes between the mutants. The original analysis used sparsely placed probes throughout coding regions for the purpose of identifying cryptic initiation sites. Our *mot1-42* analysis used oligonucleotide probes spaced at 8 base pairs (bp) throughout the entire yeast genome. We thus repeated the *spt16-197* analysis this time using the same tiling arrays from our previous *mot1-42* study which enabled a higher resolution of detection of transcriptional changes. We verified that the results from our *spt16-197* analysis were similar to the results from the previous analysis (Figure 3.5*A*, r = 0.72, $p << 1x10^{-277}$). Next we determined that there was a significant

overlap in expression changes in *mot1-42* and *spt16-197* cells (Figure 3.5*B*, r = 0.34, $p = 1x10^{-277}$). Finally, we determined the effects on expression in the *mot1-42 spt16-197* double mutant. Because the *mot1-42* and *spt16-197* strains have different parental backgrounds, we compared the double mutant expression to both wildtype strains to determine if there were any significant differences between the comparisons. Fig. 3.5*C* shows that the two comparisons have a high correlation (r = 0.74, $p << 1x10^{-277}$). Since the results from the two wildtype strains were similar to each other, we used the *MOT1*-WT data (AY51) for further comparisons.

After obtaining the expression data, we determined which genes were underexpressed (down-regulated) or overexpressed (up-regulated) in the mutants. Comparing the differential expression for each mutant to a standard normal quantile allowed us to identify where the data deviated from the standard normal quantile (Fig 3.6). These points of deviation were chosen as the thresholds for underexpression and overexpression.

Based on these thresholds, we determined that 1,670 genes were down-regulated and 1,653 genes were up-regulated in *mot1-42* cells. A total of 1,650 genes were down-regulated and 1,668 genes were up-regulated in *spt16-197* cells. There was an overlap of 684 genes that were down-regulated and 623 genes that were up-regulated in both *mot1-42* and *spt16-197* cells (Fig. 3.7*A*). The effects on expression in each mutant are shown in Fig. 3.7*B* for all genes. The differential expression is colored based on the key with genes that are down-regulated the most indicated by dark green and up-regulated by dark red. Genes that were similarly affected in either *mot1-42* or *spt16-197* cells tended to show exacerbated defects in the *mot1-42 spt16-197* cells, which agrees with the synthetic lethality.

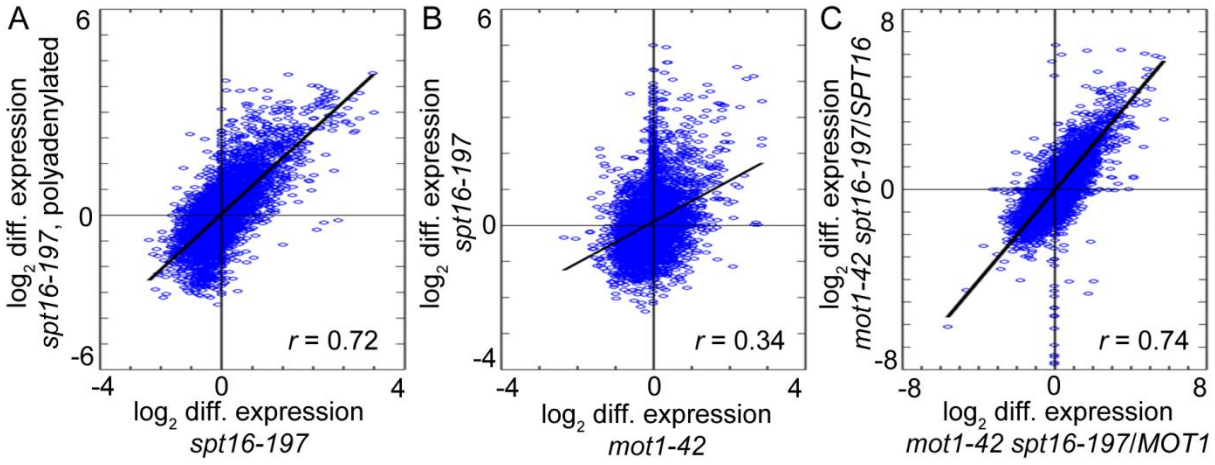

Figure 3.5. Genome-wide expression changes in *mot1-42* and *spt16-197* cells were correlated. *A*, Comparison of differential expression (*spt16-197* cells versus WT) in this study with previously published data (Cheung et al. 2008). The Pearson correlation coefficient r = 0.72 ($p << 1 \times 10^{-277}$) indicates a high correlation. *B*, Comparison of expression changes in *mot1-42* cells (Poorey et al. 2010) and *spt16-197* cells (this study) versus WT. The expression changes are significantly correlated (r = 0.34; $p < 1\text{x}10^{-277}$). *C*, The *mot1-42 spt16-197* double mutant strain was obtained via a cross of the two single mutants, which are derived from different strain backgrounds. The plot shows that differential RNA levels in *mot1-42 spt16-197* cells obtained by comparison to RNA from each of the parental WT strains are highly correlated (r = 0.74; $p << 1\text{x}10^{-277}$). RNA was purified by Joseph Muldoon, the RNA was hybridized to the tiling arrays by the UVA Microarray Core, and figures were generated by Joseph Muldoon.

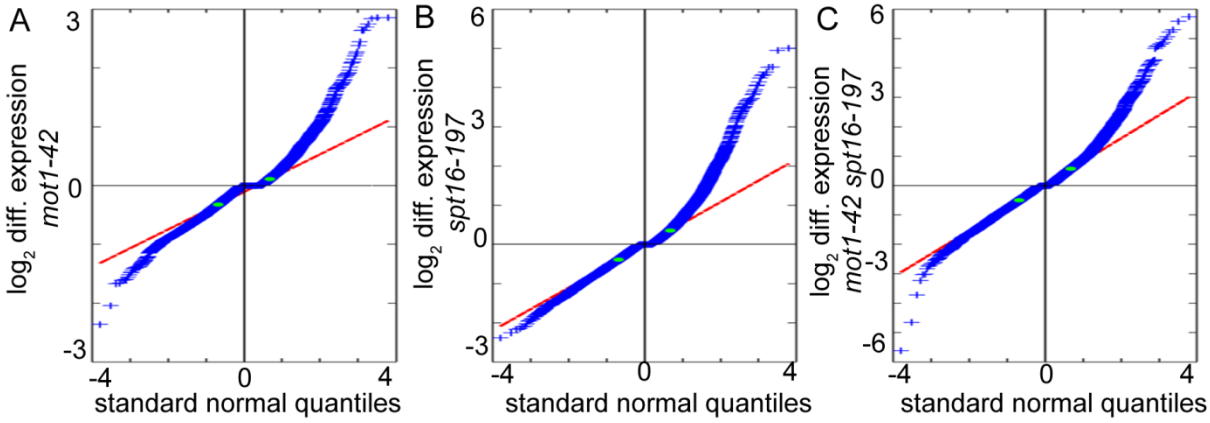

Figure 3.6. Thresholds for overexpression and underexpression were determined using deviation from standard normal quantiles. *A-C*, Differential gene expression thresholds in *mot1-42, spt16-197* and double mutant strains. Distribution of $\log_2$ fold median differential expression in each mutant (blue) is compared to a normal distribution (red line) by plotting standard normal quantiles. Due to differences in magnitudes of expression changes in each mutant, different thresholds were determined for each strain at deviation from the normal distribution (green dots). *A*, Differential expression in *mot1-42* cells, and thresholds for overexpression and underexpression of 0.1100 and -0.3245. *B*, Differential expression in *spt16-197* cells, and thresholds of 0.3540 and -0.3910. *C*, Differential expression in *mot1-42 spt16-197* cells, and thresholds of 0.5705 and -0.5010. Figures were generated by Joseph Muldoon.

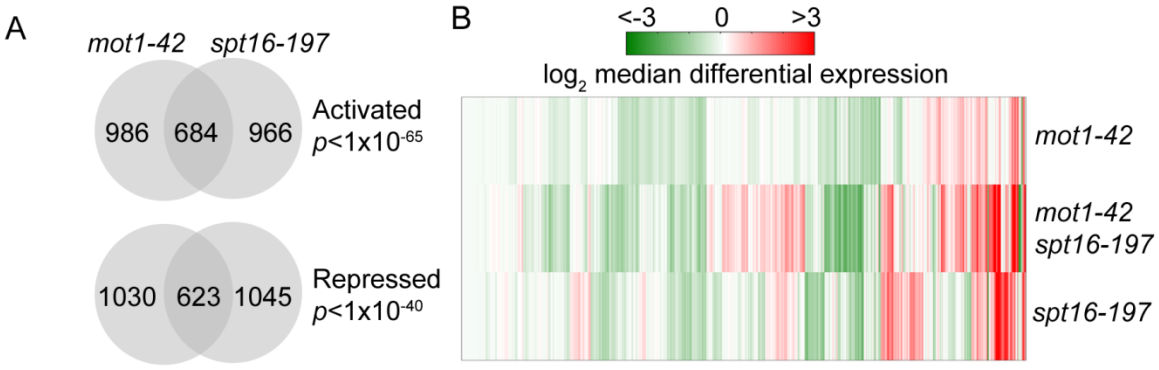

Figure 3.7. Mot1 and Spt16 co-regulate gene expression. *A*, Mot1 and Spt16 co-regulate approximately 1,300 genes. The Venn diagrams show the numbers of genes and their overlaps classified as activated (repressed in the mutants; $p < 10^{-65}$) or repressed (overexpressed in the mutants; $p < 10^{-40}$). *P*-values were determined by a one-tailed Fisher's exact test. *B*, Comparison of expression in the single and double mutants. The heat map shows $\log_2$ median differential expression for 6,685 genes, with vertical lines as genes and colored as in the key. Genes regulated by both factors generally showed similar expression changes in the single mutants that were exacerbated in the double mutant. Analysis was performed by Joseph Muldoon.

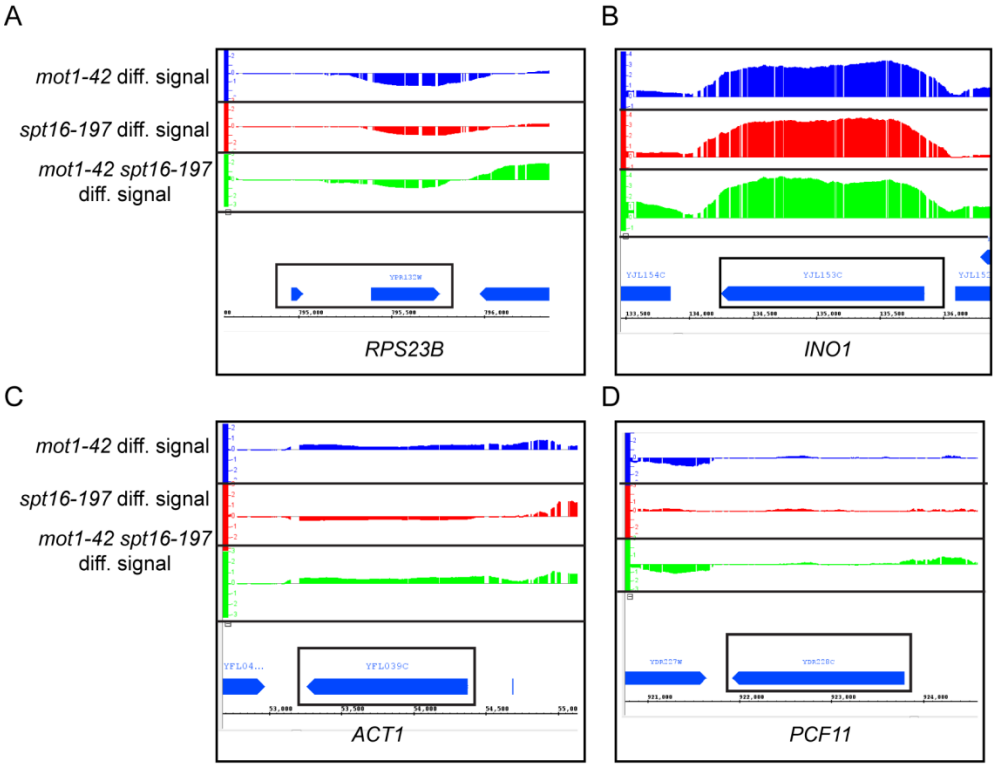

Figure 3.8. Screenshot from the Integrated Genome Browser for expression changes at selected genes in all mutants. *A-D*, Log_2 median differential expression is shown for all three mutants for four selected genes. Each mutant is colored similar to Figure 3.7: *mot1-42* (blue), *spt16-197* (red), and *mot1-42 spt16-197* (green). All mutants are on the same scale but each gene is scaled to show the appropriate expression changes. *A*, Expression changes for the Mot1-Spt16 co-activated gene *RPS23B*. *B*, Expression changes for the co-repressed *INO1* gene. *C*, Expression changes for *ACT1*, which was considered unaffected by the thresholding approach. *D*, Expression changes for *PCF11*, an unaffected control gene.

Next, we selected genes using the Integrated Genome Browser (IGB) to better understand the functional relationship between Mot1 and Spt16. Figure 3.8 shows examples of the different classification of genes: co-activated (*RPS23B*), co-repressed (*INO1*), and unaffected (*ACT1* and *PCF11*). RNA was extracted from heat-shocked cells from all three strains and quantified by RT-qPCR to validate the expression changes visualized in IGB. As shown in Fig. 3.9 all three of the co-activated genes were significantly down-regulated in the mutants and the co-repressed genes were up-regulated as was expected.

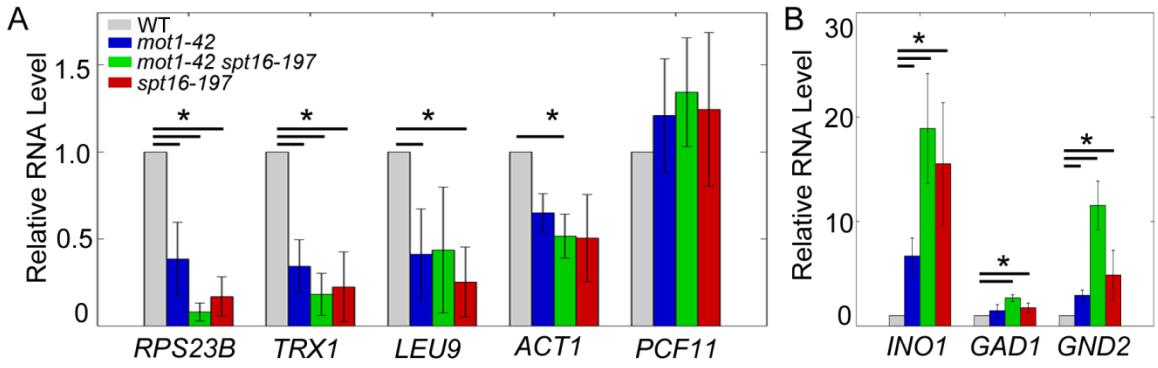

Figure 3.9. Confirmation of RNA expression changes in mutants. *A*, Validation of gene expression effects at selected co-activated genes. The bar graph shows the relative levels of RNA for each gene relative to the level in the WT strain. The RNA levels for the Mot1 and Spt16 co-activated genes *RPS23B, TRX1, LEU9,* and *ACT1* decreased compared to their levels in WT cells. *PCF11* is an unaffected control gene. *B*, RNA levels for the Mot1 and Spt16 co-repressed genes *INO1, GAD1,* and *GND2.* In *A* and *B*, error bars show one standard deviation from at least three biological replicates, and asterisks denote $p < 0.05$ by Student's *t*-test.

ACT1 was initially chosen as a control gene unaffected by mutation of either *MOT1* or *SPT16* according to the tiling array data. However, RT-qPCR showed that *ACT1* was down-regulated in all three mutants, though only reaching statistical significance in *spt16-197* cells. Extensive searching through the tiling array data yielded several candidate control genes: *PDA1*, *MIS1*, *CET1*, and *PCF11*. RT-qPCR demonstrated that only *PCF11* was unaffected in all three mutants and was selected as the control. This difference in tiling array analysis and RT-qPCR is not surprising, since there will be variations in the sensitivity of either approach, which stress the importance of validation.

Mot1 preferentially represses genes that contain TATA boxes and activate genes that lack a TATA box (Collart 1996, Dasgupta et al. 2002, Poorey et al. 2010, Venters et al. 2011, Zentner and Henikoff 2013). Because Mot1 and Spt16 co- regulate of ~1,300 genes, we determined if Spt16 had a similar preference for TATA vs TATA-less promoters. In order to do this, we used a previous classification system from with the strict interpretation of TATA box genes as containing the consensus sequence "TATAWAWR" (W = A/T and R = A/G) and all other genes as TATA-less (Basehoar et al. 2004). The genes were categorized and then sub-divided based on whether they were activated (down-regulated in the mutant) or repressed (up-regulated in the mutant) by each factor. Using a Chi-squared analysis we determined that there were more genes than expected by chance that contained a TATA box and were repressed by either factor (Figure 3.10). Genes that lacked a canonical TATA box and were activated by either factor were also enriched. This indicates that Spt16 has a similar preference as Mot1 for promoters with or without a TATA-box. To further emphasize this point, the same analysis was performed for Set2, which is a histone methyltransferase and a

known elongation factor (Venkatesh et al. 2012). The pattern of enrichment is strikingly different than that for Mot1 or Spt16, showing that Set2 has no preference for promoter elements, which is unsurprising for a bona fide elongation factor. The difference between Spt16 and Set2 suggests that Spt16 has more of a role in transcription initiation than previously appreciated.

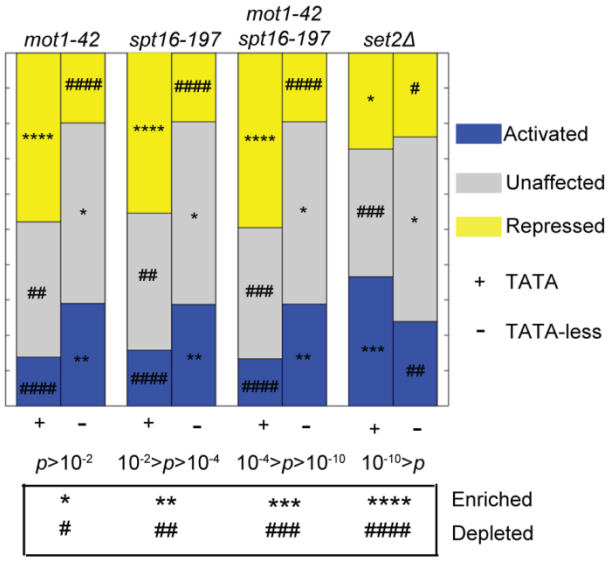

Figure 3.10. Mot1 and Spt16 have similar preferences for promoter elements. Genes are grouped by promoter attribute and proportioned by category of differential expression: co-repressed (yellow), co-activated (blue), or unaffected (gray). The *set2*Δ data (Poorey et al. 2010) is a control that demonstrates a different relationship of TATA-containing and TATA-less genes to a regulatory factor. *P*-values were calculated using a chi-squared test.

We aimed to understand how Mot1 and Spt16 contribute to TBP recruitment at co-regulated promoters. Before examining this though it was necessary to understand some of the conflicting data for TBP recruitment in *mot1* cells. It has been shown previously that TBP levels increased at all promoters in *mot1* cells (Dasgupta et al. 2005), meanwhile another study indicated that TBP levels increased at TATA boxes and decreased at lower affinity sites (Biswas et al. 2006). Normally, ChIP-based methods use formaldehyde to crosslink proteins to DNA in order to lock the proteins in place. Recently, studies have taken a native approach without crosslinking, suggesting that artifacts of formaldehyde crosslinking could hinder the interpretation of results (Zentner and Henikoff 2013). Notably, native IPs showed that TBP levels in *mot1* cells vary depending on the affinity of the binding site for TBP (Dasgupta et al. 2005, Biswas et al. 2006). Formaldehyde-based approaches mostly showed that TBP levels increased at all sites (Dasgupta et al. 2005), potentially due to trapping of weakly bound TBP in *mot1* cells (Poorey et al. 2013). This interpretation would be straight-forward except for the discrepancy still between formaldehyde-based studies. In order to overcome this discrepancy, we examined data from a previous study (Biswas et al. 2006) that yielded contradictory results to our own (Dasgupta et al. 2005). A standard method to normalize ChIP signal is by subtracting the mock signal and dividing by the input DNA. The study that contradicted our data normalized their data to a region on chromosome V that contained no genes, which they termed “Intergenic V” (Biswas et al. 2006) Using our previous TBP ChIP-chip results, we examined the TBP levels at this intergenic V location to determine how TBP was affected here in *mot1* cells (Poorey et al. 2010). Interestingly, we saw a dramatic increase in TBP levels in *mot1* cells compared to WT cells, which we confirmed by ChIP followed by qPCR (4.56-fold). Figure 3.11 shows the ChIP signal obtained from this previous study for three different promoters. In order to

directly compare the values from both studies, we multiplied the ChIP signals from the previous study by the 4.56 increase in TBP at intergenic V to take out this erroneous normalization metric. After correcting the ChIP signals from both studies are comparable showing that there is an increase in TBP at these sites in *mot1* cells. This difference shows the importance of using the correct normalization methods and how it can affect the results.

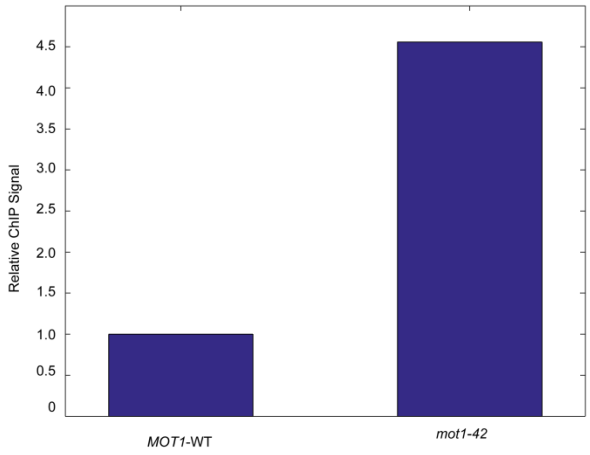

Figure 3.11. TBP levels increased at the "Intergenic V" region in *mot1-42* cells. TBP ChIP was performed in *MOT1*-WT and *mot1-42* cells to determine if there was a change in TBP levels at the Intergenic V region previously used for normalization (Biswas et al. 2006). TBP levels increased 4.56-fold in *mot1-42* cells compared to WT cells. Normalization to this region would therefore perturb the results.

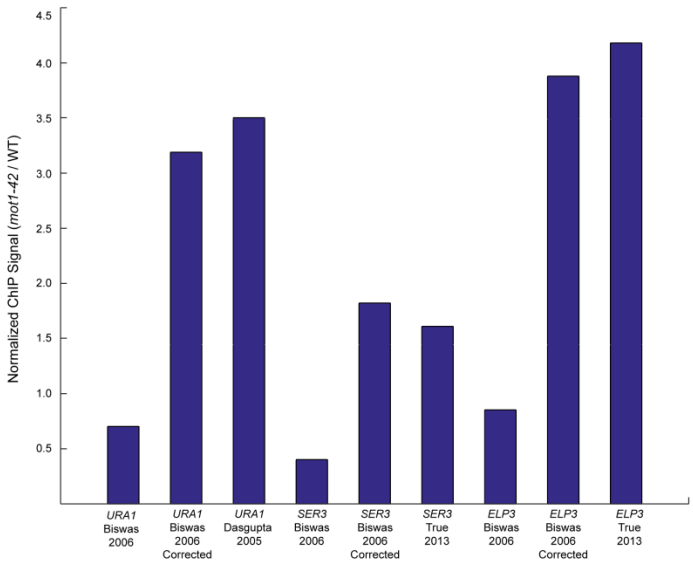

Figure 3.12. Previous contradictory results in TBP localization are consistent with our results after correcting the erroneous normalization. The previous study (Biswas et al. 2006) reported that TBP levels at *URA1*, *SER3*, and *ELP3* all decreased in *mot1-42* cells compared to WT cells after normalization to the Intergenic V region. After removing this incorrect normalization metric, the results for these three genes are consistent with our previously published data (*URA1*; Dasgupta et al. 2005) and new results (*SER3* and *ELP3*).

Next, to determine the effects on TBP localization in *mot1-42*, *spt16-197*, and *mot1-42 spt16-197* cells, we used six of the genes that we validated as Mot1-Spt16 co-regulated genes and an equal number of co-activated (*RPS23B*, *TRX1*, *ACT1*) and co-repressed (*INO1*, *GAD1*, *GND2*). Similar to our previous study (Poorey et al. 2010), TBP levels were increased compared to wildtype at all six genes analyzed (Figure

3.13*A*) in *mot1-42* cells and also in *mot1-42 spt16-197* cells (Poorey et al. 2010). However, TBP levels were unaffected at any of these genes in *spt16-197* cells. This suggests that though Spt16 may regulate TBP recruitment at some genes, Mot1 is responsible for TBP regulation at co-regulated genes.

Next we examined the effects of Mot1 and Spt16 on TFIIB localization to the promoter. Previously, TFIIB has been shown to correlate with expression, whereas TBP does not always correlate with expression (Poorey et al. 2010). Based on this, it was no surprise that TFIIB levels decreased compared to wildtype in all three mutants at co-activated genes and increased compared to wildtype at co-repressed genes (Figure 3.13*B*). Taken together, at co-activated genes in *mot1-42* or *mot1-42 spt16-197* cells there was an increase of TBP and a decrease of TFIIB, indicating that the TBP is non-functional due to improper orientation/positioning.

After analyzing the effects on expression and components of the PIC, we explored the effects on elongation. Using the tiling arrays not only can the magnitude of expression changes be detected but so can transcription length defects as previously explained (Figure 3.1). Previous studies have shown that in *spt16* cells there is an abundance of cryptic initiation (Kaplan et al. 2003, Cheung et al. 2008). When we analyzed the expression data from *spt16-197* cells, we were surprised by the abundance of length defects observed (Figure 3.15). Similar to previous studies we identified DI as the predominant effect in *spt16-197* cells, but there were ~500 genes that had the other three classes of transcription defects. This indicates that Spt16 has a larger role in RNA synthesis precision than previously observed. We also performed this analysis for *mot1-42 spt16-197* cells. The two most predominant classes of defects were PT (736 genes) and DI (774 genes) which were the main defects in the single mutants. This suggests

that Mot1 and Spt16 do not function together in transcription elongation and regulation of transcription length; however both are important in the overall fidelity of RNA synthesis precision.

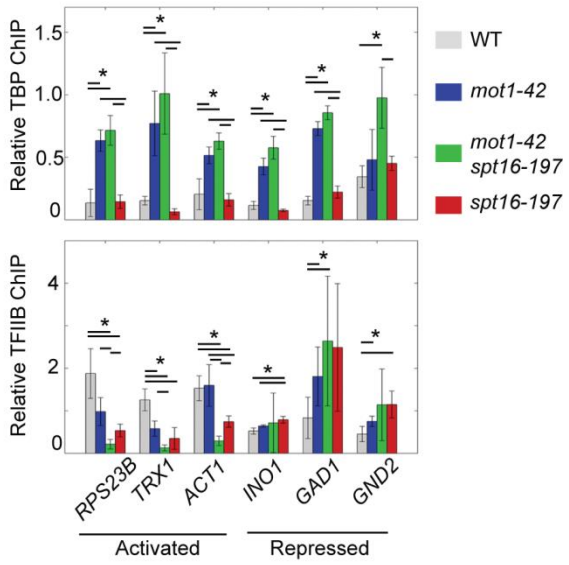

Figure 3.13. Mot1 and Spt16 differentially affected PIC component localization. *A*, To examine effects of Mot1 on TBP levels at co-regulated promoters relative TBP levels in WT and mutant strains were examined at three co-activated and three co-repressed genes. TBP in *mot1-42* cells increased at promoters regardless of promoter attribute, was unaffected in *spt16-197* cells, and in the double mutant cells resembled *mot1-42* cells. *B*, TFIIB recruitment was affected in *mot1-42* and *spt16-197 cells* at co-regulated promoters. Changes in TFIIB levels correlated with changes in RNA in each mutant. In *A* and *B*, error bars are one standard deviation from at least three biological replicates, and asterisks indicate $p < 0.05$.

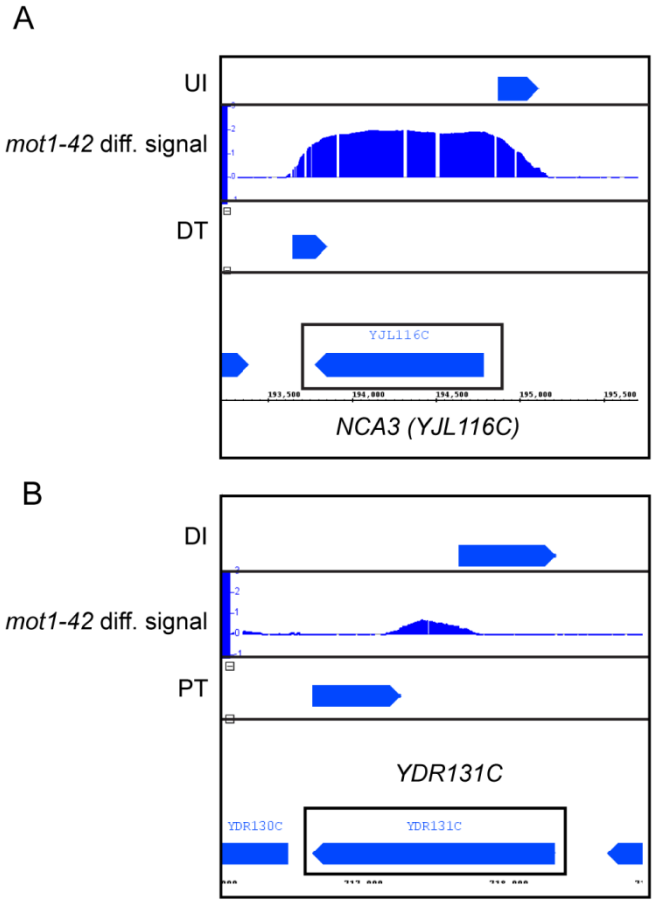

Figure 3.14. Screenshot of transcription length defects in the Integrated Genome Browser. Log_2 median differential expression is shown for the mot1-42 analysis as an example. Length defects were called using the previous analysis pipeline described (Poorey et al. 2010). *A*, Upstream initiation (UI) and downstream termination (DT) were called for this gene *NCA3*. *B*, Downstream initiation (DI) and premature termination (PT) were called for this gene *YDR131C*.

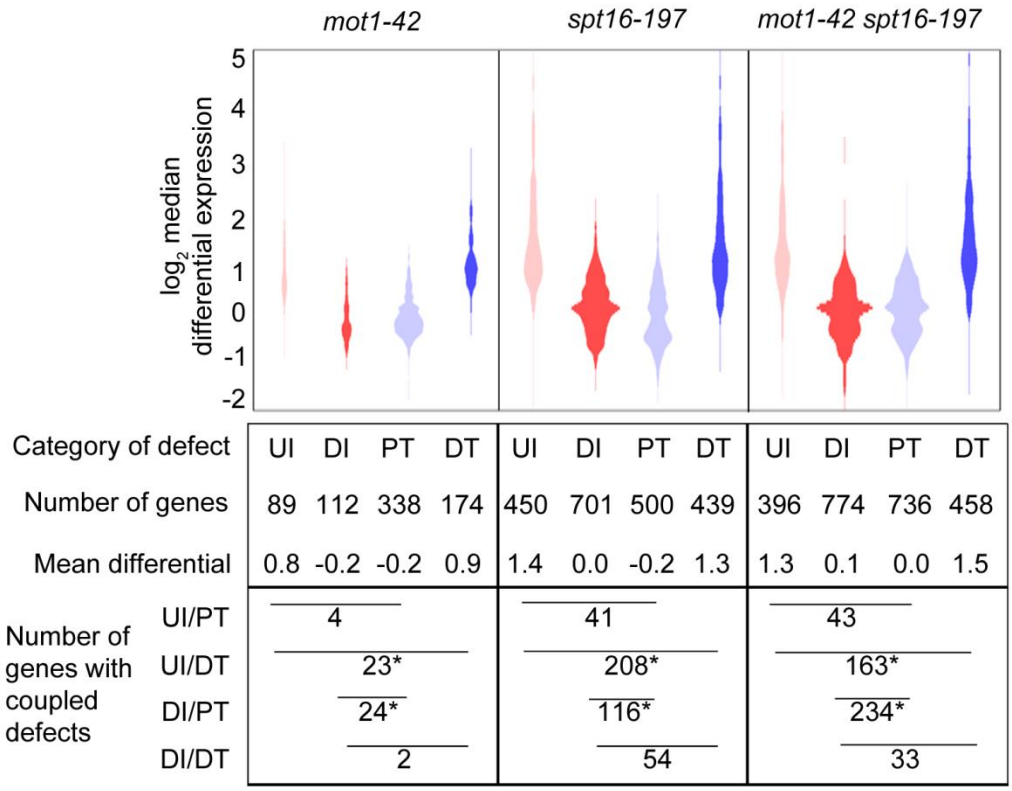

Figure 3.15. Mot1 and Spt16 have distinct and prominent roles in maintaining transcriptional precision. Distributions of differential expression for each length defect category: upstream initiation (UI), downstream or cryptic initiation (DI), premature termination (PT), and downstream termination (DT). Plot areas are proportional to the number of genes. Pairs of categories with similar distributions and mean differentials overlap significantly in affected genes. UI and DT co-occurred significantly in all mutants, as did DI and PT. Asterisks denote $p < 0.001$ from a one-tailed Fisher's exact test. Analysis was performed by Joseph Muldoon.

Discussion

Mot1 and Spt16 Co-Regulate Transcription

Mot1 and Spt16 are essential conserved proteins involved in transcription regulation (Davis et al. 1992, Wittmeyer and Formosa 1997, Poorey et al. 2010., Formosa 2012). Both factors have separately been shown to regulate almost half of the yeast genome (Cheung et al. 2008, Poorey et al. 2010). Our current study is the first to demonstrate that these two factors overlap in regulation of a subset of genes. Approximately 1,300 genes were similarly affected by both factors, and the expression of these genes was further exacerbated in the double mutant. This indicates that at these co-regulated genes both Mot1 and Spt16 are required for proper transcriptional activation or repression. The synthetic lethality of the double mutant is also consistent with this idea. Yeast genetics utilizes growth defects in double mutants or overexpression strains to show that two factors are in the same pathway, antagonize each other, or share redundant roles (Guarente 1993). Synthetic lethality is classically interpreted as the two gene products having a redundant role (Figure 3.16*A*), however, the usage of ts alleles in this study complicates this interpretation, since gene deletions are not able to be used. Both single mutants grow normally at 30°C; however these mutations could still hinder the functions of Mot1 and Spt16 slightly. When both mutants were combined, the synthetic lethality could be a product of hindering both proteins which is enough to impact transcription and cause the decreased viability of the double mutant. This could lead to the possibility of an epistatic interpretation (Figure 3.16*B*). Finally, the two factors could have separate roles initially that converge on the same step (Figure 3.16*C*). As discussed later, the evidence presented in this chapter favors this last interpretation.

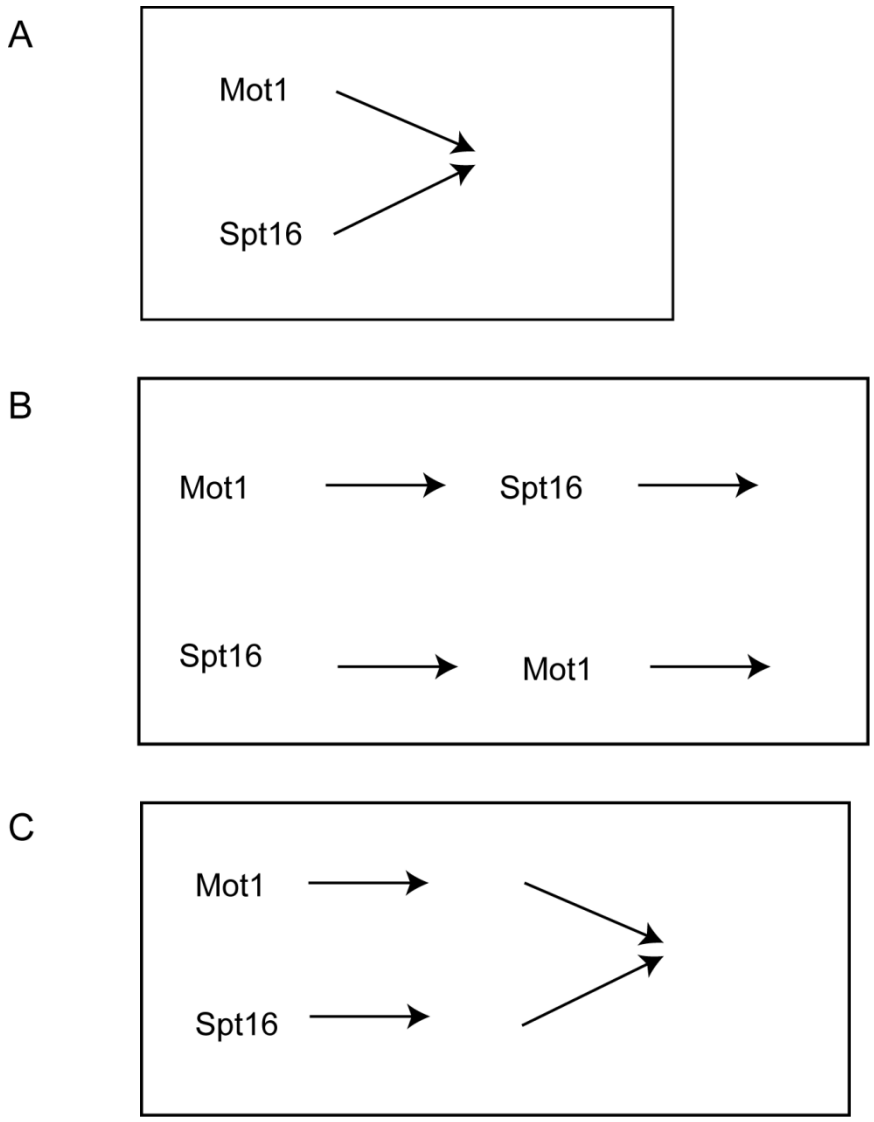

Figure 3.16. Possible pathways interpreted from the synthetic lethality observed in *mot1-42 spt16-197*. *A*, The two proteins could have redundant functions. When one protein is hindered, the other can replace it. In the absence of both, the cells would die. *B*, The two proteins could function epistatically in a pathway. Since temperature-sensitive alleles were used, hindering one protein could still be acceptable for the cell to continue in the

pathway. Hindering both proteins would cause lethality. *C*, The two proteins could have separate roles initially that converge on the same regulatory step.

Both Mot1 and Spt16 are required for activation or repression of specific genes. The overlap in regulation is equally split between co-activated and co-repressed genes. This is interesting because these factors were initially identified as repressive (Mot1) or activating factors (Spt16; Davis et al. 1992, Auble et al. 1993, Orphanides et al. 1999). Factors can have both positive and negative roles on transcription depending on various circumstances including promoter context, the presence of other proteins, or the growth conditions.

The physical interaction between Mot1 and Spt16 is perhaps the most intriguing part of this chapter. Both proteins are thought to have specific chromatin localizations that are separate from each other, however in whole cell extracts they are shown to associate. This association does not necessarily mean that Mot1 and Spt16 are directly binding each other, since there are other factors present in the extracts. To conclusively prove that these two factors directly interact, purified proteins would need to be incubated together and assayed. Regardless of whether this interaction is direct or indirect, it is still interesting to hypothesize about what this association means. Both Pob3 and Nhp6 were shown to associate with Mot1, however the previous proteomics screen showed that both interactions were not as significant as Spt16 with Mot1 (Arnett et al. 2008). It is possible that Mot1 interacts directly with Spt16 which could still be bound to Pob3, explaining the difference in the statistical significance. A previous experiment showed that Mot1 could bind to nucleosomes and the interaction between

Mot1 and Spt16-Pob3 actually prevented Spt16-Pob3 from interacting with Nhp6-nucleosomes (Wells 2012). However, these results have some caveats since the proteins were from different species and may be artificial. It is important to repeat this experiment with proteins from *S. cerevisiae* to understand how these proteins interact with each other and nucleosomes. This same study showed that Mot1 does not directly interact with Nhp6-nucleosomes. This could indicate that the association between Mot1 and Nhp6 is indirect in some manner. If Mot1 interacts with Spt16-Pob3, which prevents them from binding Nhp6-nucleosomes, then another factor must bridge them all together. Purified proteins from the same species need to be used to determine how Mot1, Spt16, Pob3, Nhp6, and nucleosomes interact.

How does TBP play into the interaction between Mot1 and Spt16? This question has not been explored yet. If Mot1 binds TBP, can Mot1 also bind to Spt16 or does the TBP interaction prevent Mot1 from binding Spt16? Since the N-terminal HEAT repeats of Mot1 bind to the convex surface of TBP and the C-terminal Swi2/Snf2 ATPase domain binds upstream DNA, this leaves little room for Mot1 to associate with Spt16 while maintaining these other interactions (Wollmann et al. 2011, Butryn et al. 2015). This again makes an argument for the requirement to determine exactly how Mot1 and Spt16 interact, specifically which domains contribute to this interaction. This would thus allow for a better understanding of how all of these proteins fit together and perform their necessary functions.

Mot1 and Spt16 Have Roles in Transcription Initiation

Mot1 regulates TBP localization genome-wide, however Spt16 has only been implicated in TBP localization at a handful of genes (Gumbs et al. 2003, Biswas et al. 2005, Poorey et al. 2010, Zentner and Henikoff 2013). Mot1 controls TBP levels at promoters and Spt16 plays no role in TBP recruitment at co-regulated genes. It is still possible that Spt16 may be required for TBP localization at genes not regulated by Mot1. It has been suggested that Spt16 plays an indirect role in TBP localization at the promoter due to its role in preventing cryptic initiation within a gene (Saunders et al. 2003). At this point, our results cannot dispute this claim. Exposure of new PIC binding sites could contribute to relocalization of TBP leading to a decrease of promoter-bound TBP.

Further examination of Spt16's role in TBP localization should be performed. This study allows for classification of genes into different subsets based on regulation by either Mot1 or Spt16. The promoters we examined in this study were all Mot1-Spt16 co-regulated genes, so it would be worthwhile to look at promoters for genes that are Spt16-regulated and not affected by Mot1. Furthermore, it would be useful to examine if Spt16 and TBP directly interact by using purified proteins. This would allow us to determine if Spt16 plays a direct or indirect role in TBP recruitment. As already mentioned, an indirect role could be explained by the effect on chromatin organization by mutation of Spt16, which will be discussed more in Chapter IV.

Mot1 and Spt16 affected TFIIB recruitment in a similar manner, correlating with expression changes in the mutants. It is again difficult at this point to determine if Spt16 is involved directly or indirectly in TFIIB localization.

The same argument for an indirect role can be argued here similar to TBP. Mot1's effect on TFIIB is most likely a consequence of its role in TBP localization. Co-activated genes have an increase in TBP levels while there is also a decrease in TFIIB, suggesting that the PIC is not properly formed, preventing proper initiation. In the absence of Mot1, TBP can bind to improper promoter sites or even bind in the wrong orientation which would prevent the PIC from forming.

At this point a simple interpretation of these results is that Mot1 and Spt16 regulate PIC formation albeit at different points at a subset of promoters. Mot1 regulates TBP and downstream TFIIB, while Spt16 affects TFIIB only. Both TBP and TFIIB are required for the PIC to form and transcription to initiate, so this shows how Mot1 and Spt16 can regulate the same complex while maintaining different functions (Figure 3.16*C* and Figure 3.17).

Mot1's Role in Elongation?

This study began in the pursuit of understanding how Mot1 was involved in transcription elongation due to the premature termination defect in *mot1-42* cells. The hope was that looking at another protein with similar functions in both initiation and elongation might lead to this elusive role for Mot1. Transcription length defects were analyzed in *spt16-197* cells and *mot1-42 spt16-197* cells. While the predominant defect in *mot1-42* cells was premature termination, the main defect in *spt16-197* cells was cryptic initiation as previously reported. The double mutant showed prominent defects in both premature termination and cryptic initiation, suggesting that these two proteins have

different roles in elongation. Spt16 helps to replace nucleosomes after Pol II passage thereby preventing intragenic PIC assembly and aberrant transcription (Mason et al. 2003, Saunders et al. 2003, Cheung et al. 2008).

The premature termination defect in *mot1-42* cells is still puzzling. This defect could be due to the cell sensing that transcription has not initiated properly due to improper PIC formation and terminating transcription before it has proceeded too far. This idea would allow Mot1 an indirect role in transcription elongation, since its effect on the promoter could be somehow sensed and relayed to terminate transcription. Mot1 could also be directly involved in early elongation in some manner. Though Mot1 is usually found at the promoter alongside TBP, one report (Geisberg et al. 2002) showed that Mot1 was present in the coding region of certain genes. It is clear that Mot1 functions in transcription initiation and elongation but the exact way in which it does so is still not understood. Future work will be needed to determine specifically how Mot1 effects elongation.

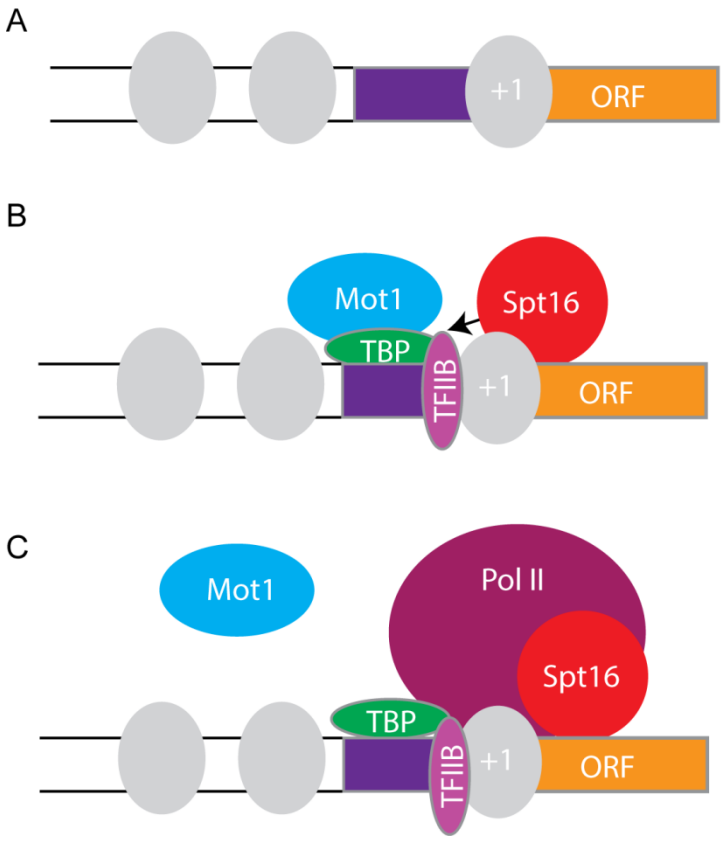

Figure 3.17. Model of Mot1 and Spt16 at the promoter. *A*, The promoter of a gene (purple box) is upstream of the open reading frame (ORF, orange box). Nucleosomes (gray ovals) are located around the promoter region but are typically depleted at the promoter yielding a nucleosome depleted region. *B*, TBP binds to the promoter. Mot1 can either remove TBP, thereby repressing transcription, or help to position the TBP properly to activate transcription. Spt16 does not affect TBP at co-regulated promoters. TFIIB is recruited by the presence of TBP, potentially indirectly due to Mot1. TFIIB is also regulated either positively or negatively by Spt16, most likely due to Spt16's role in the chromatin organization at the promoter. *C*, The rest of the PIC (not shown) and RNA Polymerase II are recruited to the promoter and initiate transcription.

Chapter IV

How Mot1 and Spt16 Affect Localization of Each Other and Nucleosomes

Data from this Chapter was published in True et al. 2016.

The Mot1 ATPase globally regulates TBP localization and localizes to the promoter of genes. The Spt16 histone chaperone contributes to global nucleosome organization and has been shown to be most abundant in the coding region of active genes. Based on the overlap in transcription regulation between these two factors, we determined if they affected the localization of each other. On a global scale, Spt16 was required for Mot1 promoter localization, and Mot1 also affected Spt16 localization to genes. We also determined if Mot1 affected nucleosome organization genome-wide since other members of the Swi2/Snf2 family are chromatin remodelers. Interestingly, we found that Mot1 has an unanticipated role in establishing or maintaining the occupancy and positioning of nucleosomes at the 5' ends of genes. Spt16 has a broad role in regulating chromatin organization in gene bodies, including those nucleosomes affected by Mot1. These results suggest that the Spt16 is required for Mot1 localization to the promoter by establishing a permissive chromatin environment for Mot1 and that both factors contribute to nucleosome organization genome-wide.

Introduction

Mot1 and Spt16 Localization

As previously discussed in Chapter I, Mot1 and Spt16 have distinct genomic localization profiles. However, the results in Chapter III demonstrate that Mot1 and Spt16 physically associate. Previous studies have shown that Spt16 can regulate promoter-proximal nucleosomes and TBP localization to the promoter (Mayer et al. 2010, Stuwe et al. 2008, Biswas et al. 2005). Therefore, we hypothesized that Spt16 and Mot1 could interact at the promoter and potentially affect the localization of each other to co-regulated genes. In order to determine if either factor could regulate the localization of the other factor, we employed ChIP-Seq of Mot1 in *spt16-197* cells and Spt16 in *mot1-42* cells. Because, Mot1 localization to the promoter is dependent on TBP (Auble and Hahn 1994), and TBP localization was unaffected in *spt16-197* cells at co-regulated genes, we hypothesized that there would be no effect on Mot1 localization in *spt16-197* cells. By utilizing ChIP-Seq we were able to determine the effects on a genome-wide scale rather than focusing on specific genes.

Nucleosome Organization and Key Regulators

Studies have been undertaken to determine what the key factors are in determining nucleosome organization, which have identified various factors including DNA sequence, transcription factors, and chromatin remodelers (Kaplan et al. 2009, Zhang et al. 2009). It is well known that poly(dA-dT) tracts disfavor nucleosome

formation (Struhl 1985, Iyer and Struhl 1995, Mai et al. 2000, Anderson and Widom 2001). Two major studies showed that approximately 20% of nucleosome organization *in vivo* is determined by the underlying genomic sequence (Kaplan et al. 2009, Zhang et al. 2009). Transcription factors and chromatin remodelers help to properly position the nucleosomes after they have been initially positioned by the sequence (Kaplan et al. 2009, Zhang et al. 2009).

Two important aspects of nucleosome organization are referred to as nucleosome occupancy and nucleosome positioning (or fuzziness). Nucleosome occupancy refers to the density of histones within a region of DNA in a population of cells (Pugh 2010). Nucleosome positioning on the other hand refers to the precise position of the histone on the DNA (Struhl and Segal 2012). Therefore nucleosomes can be well-positioned with high occupancy (Figure 4.1*A*), well-positioned with low occupancy (Figure 4.1*B*), not positioned with high occupancy (Figure 4.1*C*), or not positioned with low occupancy (Figure 4.1*D*). Nucleosomes can also vary in degrees and have more intermediate occupancy or positioning. An example of a well-positioned nucleosome is the +1 nucleosome immediately downstream of the TSS. The nucleosome depleted region (NDR) is an example of a low occupancy region. Nucleosome occupancy typically anti-correlates with transcription, i.e. active genes have lower nucleosome occupancy and repressed genes have higher nucleosome occupancy. This is not always true since other factors can contribute to transcriptional output.

Genome-wide nucleosome mapping was performed first utilizing microarrays and more recently using deep sequencing (Yuan et al. 2005, Lee at al. 2007, Shivaswamy et al. 2008, van Bakel et al. 2013). Maps have been generated for many different species,

under different conditions, and utilizing different mutant strains to determine how specific proteins affect nucleosome organization. Some of these studies use the enzyme micrococcal nuclease (MNase), which digests unprotected chromatin and leaves only DNA that is bound by proteins. Proper titration of MNase can yield mononucleosomes that allow for mapping to precise genomic locations (Yuan et al. 2005, Tsui et al. 2012, Wal and Pugh 2012). The once-named nucleosome free region (NFR) actually contains highly dynamic or "fragile" nucleosomes, which were detected using varying amounts of MNase (Knight et al. 2014). Now the NFR is more appropriately called the nucleosome depleted region (NDR) (Knight et al. 2014).

Scope of this study

In Chapter III, it was demonstrated that Mot1 and Spt16 co-regulated ~1,300 genes. But, it has been previously shown that both factors have distinct genome-wide localization profiles. Therefore, we determined if these factors influenced the localization of each other genome-wide. We hypothesized that Mot1 had a nucleosomal role at a subset of genes, so we tested this genome-wide using MNase ChIP-Seq. We also compared the effects on nucleosome organization in *mot1-42*, *spt16-197*, and *mot1-42 spt16-197* cells to determine if they affected nucleosomes similarly, which we hypothesized would contribute to their combined efforts in regulating gene expression. Overall, this chapter focuses on the genome-wide localization of Mot1, Spt16, and histone H3 in *mot1* and *spt16* mutants.

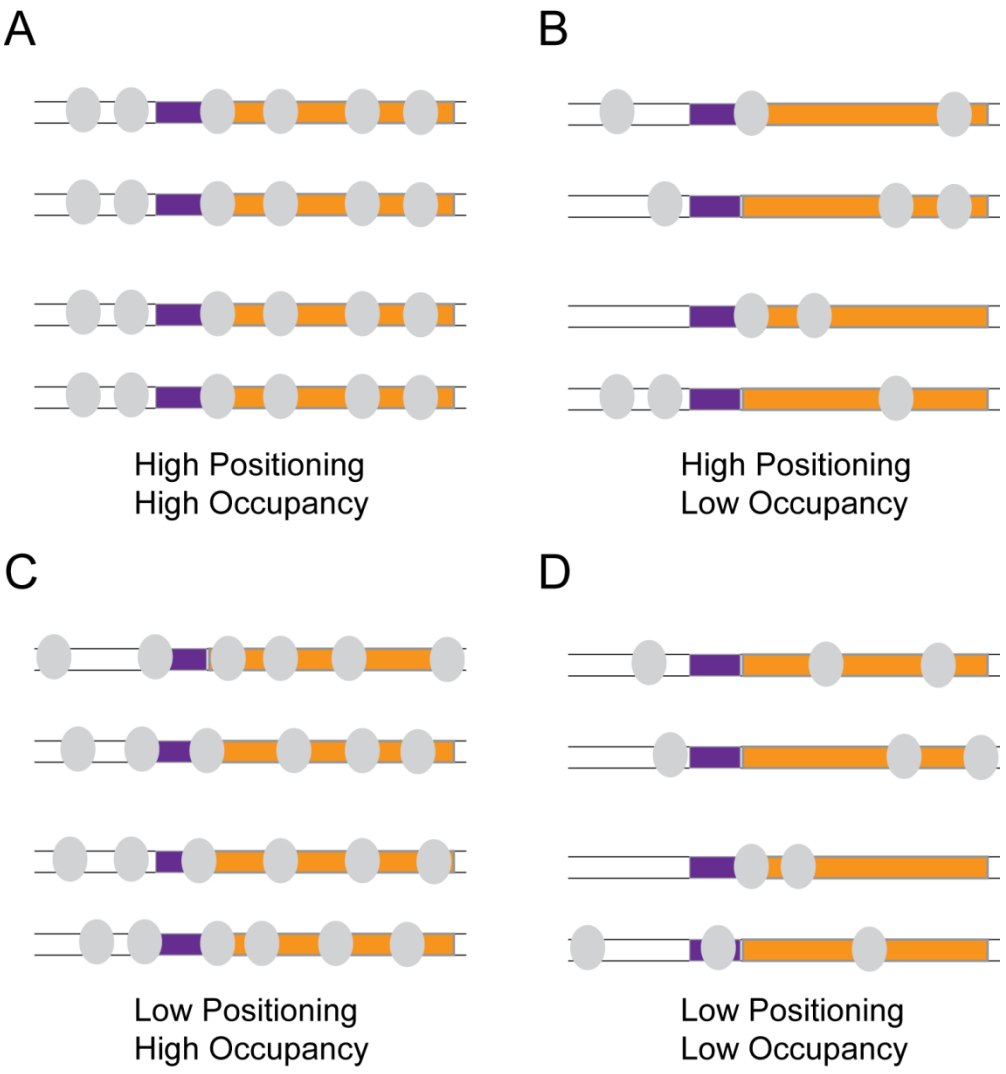

Figure 4.1. Nucleosome occupancy and nucleosome positioning. Nucleosome occupancy refers to the density of histones in a region of DNA, whereas nucleosome positioning refers to the exact location of nucleosomes on the DNA. Examples of nucleosome with high or low positioning and high or low occupancy are shown in *A-D*. *A*, A population of nucleosomes with high positioning and high occupancy. *B*, A population of nucleosomes with high positioning and low occupancy. *C*, A population of nucleosomes with low

positioning and high occupancy. *D*, A population of nucleosomes with low positioning and low occupancy.

Materials and Methods

Yeast strains and growth conditions

Saccharomyces cerevisiae strains used in this study are the same as listed in Table 4.1. *MOT1*-WT, *mot1-42*, *SPT16*-WT, and *spt16-197* cells were all grown at 30°C in yeast extract, peptone, dextrose media (YPD), while *mot1-42 spt16-197* cells were grown at 25°C in YPD. For ChIP, ChIP-Seq, and MNase ChIP-Seq experiments, all strains were grown at their permissive temperatures to an optical density (OD) ~1.0, heat-shocked with addition of an equal volume of 42°C pre-warmed YPD, and incubated at 35°C for 45 minutes.

Chromatin Immunoprecipitation

Yeast strains were grown as described above. After cultures reached an OD ~1.0 and were heat shocked, cells were collected, lysed, and sonicated as described in Chapter III. Immunoprecipitations were conducted using 1 mg of protein as previously described with either 5 µL α-myc (9E10, 1 mg/mL) or 5 µL α-Pol II (8WG16 Covance). After immunoprecipitation and crosslink reversal, DNA was purified using a Qiagen PCR Purification kit according to the manufacturer's instructions. Purified DNA was then quantified by qPCR using the oligonucleotides (Life Technologies) listed in Tables 3.2

and 4.2 and SYBR Green master mix (Bio-Rad). The mock values were subtracted from the IP values and then normalized to the input.

Table 4.1

Yeast strains used in this chapter

Strain name	Genotype	Reference
YPH499	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1*	Sikorski et al. 1989
AY51	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1. mot1Δ::TRP1 pAV20(EE-MOT1, LEU2, CEN ARS)*	Dasgupta et al. 2005
AY87	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1. mot1Δ::TRP1 pMot221(mot1-42, LEU2, CEN ARS)*	Dasgupta et al. 2005
L577	*MATα, his4-912δ, lys2-128δ, ura3-52, spt16-197*	Malone et al. 1991
YJT001	*MATα, his4-912δ, lys2-128δ, ura3-52, MOT1-13xMyc (kanMX)*	This study
YJT002	*MATα, his4-912δ, lys2-128δ, ura3-52, spt16-197, MOT1-13xMyc (kanMX)*	This study
YMW054	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1, mot1Δ::kanMX, SPT16-13xMyc (HIS3), pAV20(EE-MOT1, LEU2, CEN ARS)*	This study
YMW055	*MATa, ura3-52, lys2-801^a, ade2-101^o, trp1-Δ63, his3-Δ200, leu2-Δ1, mot1Δ::kanMX, SPT16-13xMyc (HIS3), pMot221(mot1-42, LEU2, CEN ARS)*	This study
YMW066	*MATa, ura3-52, trp1-Δ63, his3-Δ200, spt16-197, mot1Δ::kanMX, pMOT221(mot1-42, LEU2,CEN ARS)*	This study

Table 4.2

Primer sequences used in this chapter

Primer Name	Sequence
GAL10 ORF F	5' GCA GCC CTG CAA TAC CTA GA 3'
GAL10 ORF R	5' TTC CAC TCA CGA CAC AAA CC 3'

ChIP-Seq

Immunoprecipitated DNA was prepared as above for ChIP. The resulting purified DNA was then processed using the TruSeq ChIP Library Prep Kit from Illumina (IP-202-

1012) according to the manufacturer's instructions. Briefly, DNA fragments were end-repaired and adenylated, adapter 6 was ligated onto the adenylated ends and gel-purified and purified products were enriched and quantified via Qubit. The DNA samples were then analyzed using a Bioanalyzer to ensure that the samples mostly contained a single peak of mononucleosomal DNA. A total of 0.2 picomoles of DNA were sequenced by the UVA Sequencing Core using an Illumina MiSeq instrument, yielding 25.6 +/- 1.6 million 150 bp raw reads per sample. Data were obtained from two independent replicate samples for each strain. ChIP-Seq was performed by Savera Shetty.

Micrococcal Nuclease (MNase) ChIP-Seq

MNase ChIP-Seq was performed as previously described with minor modifications (Wal and Pugh 2012). Strains were grown as described above. After collection, cells were lysed as described above for ChIP. MNase titrations were performed, and samples with the highest proportion of mononucleosomes and without sub-nucleosomal fragments were selected for each sample. Cross-links were reversed with heat, and protein was digested with proteinase K (Thermo Fisher). Samples were then immunoprecipitated with 10 µg α-H3 (Abcam ab1791) overnight at 4°C, and then incubated with protein A Sepharose, washed, and eluted as previously described. The immunoprecipitated material was prepared for sequencing as described above and sequencing data were obtained using the Illumina MiSeq instrument in the UVA Sequencing Core, yielding 27.9 +/- 5.7 million 150 bp raw reads per sample. Data were obtained from two independent replicate samples for each strain.

ChIP-Seq and MNase ChIP-Seq Data Analysis

Reads were aligned to the SacCer3 reference genome using bowtie2-2.2.6 with default mapping settings (Langmead and Salzberg 2012). More than 95% of the raw reads from each dataset were mapped under these conditions. Unmapped reads and reads with mapping quality scores < 30 were removed from the bowtie2-generated bam file using samtools-0.1.19 (Li et al. 2009). This mapping quality threshold resulted in removal of 10% or fewer of the mapped reads. For viewing the data in IGV (Robinson et al. 2011, Thorvaldsdóttir et al. 2013), bam files were converted to BigWig by first sorting and indexing using samtools, then creating .bed and .cov files using BEDTools 2.18.2 (Quinlan and Hall 2010), and finally converting the .cov files to BigWig using the bedGraphToBigWig converter (Kent et al. 2010). Gene average plots were produced using ngs.plot-2.4.7 (Shen et al. 2014). Nucleosome mapping and differential nucleosome analyses were performed using the dpos tool in the DANPOS-2.2.2 package with default settings that include fold-change normalization (multiplying each sample by a scale factor to account for differences in mapped read coverage) rather than quantile normalization as originally described (Chen et al. 2013). Dpos yielded normalized wig files of the processed data. We observed that the total signal nonetheless varied among these normalized datasets by a small (~5%) but not insignificant amount. For this reason, the dpos-generated wig data files were further normalized globally based on the total signal. The gene average plots shown in Figure 4.9 were generated in R from the renormalized wig data by computing the mean signal at base pair resolution for all genes, activated genes, or repressed genes (as indicated) with respect to the TSS. Other statistical analyses and data visualization were performed using R v3.0.2. Data analysis was performed by David Auble.

All of the genome-wide datasets from this study have been submitted to the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under SuperSeries accession number GSE80235.

Results

Because Mot1 and Spt16 regulate overlapping sets of genes, we obtained a clearer picture of their localization patterns and if either protein affected the other's localization. First we performed ChIP-Seq of Mot1-myc in WT and *spt16-197* cells to determine if Spt16 affected Mot1 recruitment. Genes were divided into co-activated or co-repressed to observe any differential effects between gene classes. Genes were computationally aligned based on their transcription start site (TSS, 0 bp) and the average profile was plotted +/- 1200 bp around the TSS (Figure 4.2). The average gene profile showed a peak of Mot1 upstream of the TSS. This is consistent with previous genome-wide and locus-specific experiments (Dasgupta et al. 2002, Geisberg et al. 2002, Zentner and Henikoff 2013). This Mot1 peak differed between co-activated and co-repressed genes. At co-activated genes, Mot1 was more broadly distributed and somewhat overlapping the TSS. At co-repressed genes, the Mot1 peak was more precise and centered at ~400 bp upstream of the TSS. Co-activated genes also had an increase in Mot1 signal within the coding region of the gene. While most studies have shown that Mot1 is absent in the coding region in WT cells, one study detected Mot1 in the coding region of certain genes (Geisberg et al. 2002). It is possible that this signal is due to an artifact from downstream genes or even from highly expressed genes, which

are known to crosslink various proteins non-specifically (Teytelman et al. 2013). However, this could indicate a role for Mot1 in elongation.

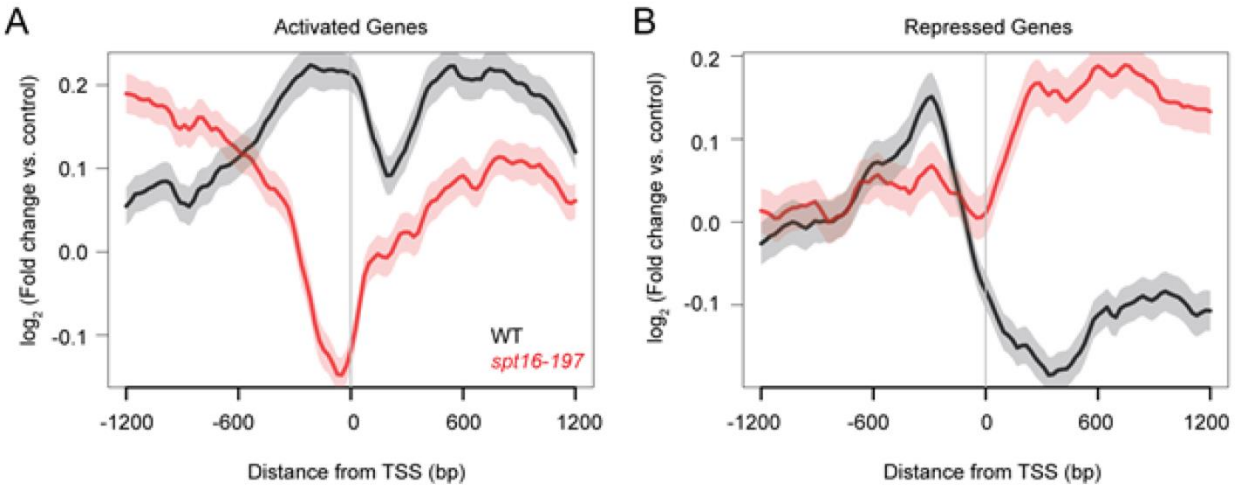

Figure 4.2. Spt16 affected Mot1 localization genome-wide. *A*, Mot1-myc ChIP-Seq in *SPT16*-WT (black line) and *spt16-197* (red line) cells at Mot1-Spt16 co-activated genes. The ChIP signal was plotted as the $\log_2$ of the fold change between the immunoprecipitated sample and the input. Genes were aligned based on their transcription start site (TSS, 0 bp) and signals were averaged +/- 1200 bp around the TSS. The transparent shading around the lines shows the standard error of the mean. *B*, Similar profile for Mot1-myc ChIP-Seq at Mot1-Spt16 co-repressed genes. ChIP-Seq was performed by Savera Shetty and analysis was performed by David Auble.

In *spt16-197* cells, the average gene profile showed that Mot1 localization decreased upstream of the TSS. This could indicate a role for Spt16 in establishing chromatin permissiveness at the promoter, which will be discussed more later. At co-activated genes, Mot1 levels decreased in the coding region, while at co-repressed genes Mot1 levels increased in the coding region, consistent with expression changes (Figure 3.9). Co-repressed genes which are enriched for TATA boxes are known to harbor cryptic initiation sites in the coding region (Cheung et al. 2008). In *spt16-197* cells it is therefore possible that these sites are exposed, recruiting TBP and Mot1 along with it causing an increase in Mot1 signal. Recently it has been shown that Mot1 helps to regulate cryptic initiation by preventing TBP from improperly binding to these exposed intragenic sites (Koster et al. 2014).

Next we determined how Spt16 localization was affected by Mot1. Spt16 localization is presented similarly to Mot1 localization, and the average profile at all genes is consistent with previous studies on Spt16 (Figure 4.3; Kaplan et al. 2003, Cheung et al. 2008, Mayer et al. 2010). Upstream of the TSS is a region that is devoid of Spt16 which is approximately where the nucleosome depleted region (NDR) is found. Since Spt16 is a histone chaperone, it makes sense that there would be a depletion of the factor where there is a lack of nucleosomes. Spt16 peaked immediately downstream of the TSS which could indicate a role for Spt16 in regulating the +1 nucleosome.

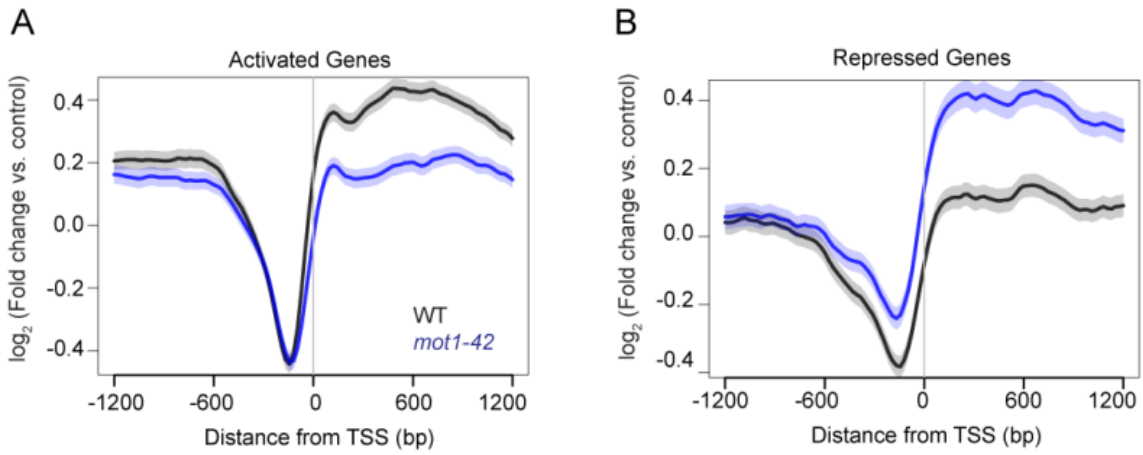

Figure 4.3. Spt16 localized to the coding region and changed consistently with expression changes in *mot1-42* cells. *A*, Spt16-myc ChIP-Seq in *MOT1*-WT and *mot1-42* cells at Mot1-Spt16 co-activated genes plotted similarly to Figure 4.2*A*. *B*, Spt16-myc ChIP-Seq signal at Mot1-Spt16 co-repressed genes similarly to *A*. ChIP-Seq was performed by Savera Shetty and data analysis was performed by David Auble.

Therefore, we examined the Spt16 localization patterns at these subsets of genes. When comparing co-activated genes to co-repressed genes, it is apparent that there was an enrichment of Spt16 at co-activated genes while there was a depletion of Spt16 at co-repressed genes. This could indicate that Spt16 is directly involved in gene activation while a lack of Spt16 leads indirectly to repression. Although solely an absence of Spt16 at a gene does not indicate that it will be repressed, since other genes are not Spt16-regulated. Upon mutation of Mot1, Spt16 localization was perturbed at both subsets of genes. Co-activated genes showed a decrease of Spt16 in the gene body compared to WT levels. This decreased level is comparable to the levels of Spt16

at co-repressed genes in WT cells. Co-repressed genes showed an increase of Spt16 in *mot1-42* cells at the promoter and ORF, which is comparable to levels at co-activated genes in WT cells. This further supports the idea that Spt16 levels correlate with expression levels and that Spt16 repression is due to an absence of the factor at repressed genes. Most likely this also indicates that Mot1 does not directly affect Spt16 localization per se, but instead Spt16 is linked to the level of Pol II recruited to genes, which will be explored shortly.

After examining the genome-wide pattern of Spt16 enrichment, we verified that these general patterns were the same at the genes that we had previously examined for expression, TBP, and TFIIB levels. We looked at Spt16 levels at a region towards the 5' end of the ORF and a region closer to the 3' end of the ORF at four co-activated genes (*RPS23B*, *TRX1*, *LEU9*, and *ACT1*) and three co-repressed genes (*INO1*, *GAD1*, and *GND2*). As shown in Figure 4.4, Spt16 levels were higher at co-activated genes at both regions of the ORF when compared to co-repressed genes, which is consistent with the ChIP-Seq data. Upon mutation of Mot1, Spt16 levels decreased at the 5' ORF at co-activated genes but were not significantly affected at the 3' ORF. Since Spt16 levels declined towards the 3' end of the ORF normally, it is possible that the Spt16 levels are unaffected here in *mot1* cells due to the lower levels present at the 3' end. At *INO1* and *GAD1*, two of the co-repressed genes, Spt16 levels increased in *mot1-42* cells compared to WT cells consistent with the ChIP-Seq data. *GND2* showed no significant change in Spt16 levels, though the expression was increased in *mot1-42* cells. This could be due to insensitivity in the ChIP assay being unable to detect a minor change in Spt16 levels. It is also possible that at *GND2* and potentially other genes that Spt16 levels are not as tightly associated with expression changes. However, all three co-

repressed genes had significant increases in Spt16 levels at the 3' end of the ORF. Mutation of Mot1 could lead to a delayed association of Spt16 in the coding region, which is why Spt16 levels were unaffected at the 5' end of the ORF and increased at the 3' end of the ORF at *GND2*. Kinetics of Spt16 association with genes in *mot1* cells could be explored to test that possibility.

Because Spt16 localization is thought to be linked with the amount of Pol II recruited to a gene (Mason and Struhl 2003, Saunders et al. 2003), we examined if Pol II levels were affected in *mot1-42* cells similar to Spt16 changes. At four co-activated genes, levels of Pol II decreased in *mot1-42* cells compared to WT cells (Figure 4.5*A*), similar to the decrease in expression and Spt16 levels previously observed (Figures 3.9, 4.3, and 4.4). *INO1* and *GAD1* showed a significant increase in Pol II levels as expected by our previous results, though the increase was not very large. *GND2* showed no change in Pol II occupancy in *mot1-42* compared to WT cells consistent with the lack of Spt16 change previously demonstrated. This is intriguing because *GND2* expression increased in *mot1-42* cells due to depression of this gene, however it seems our ChIP assay is not sensitive enough to detect very subtle changes at *GND2*..

To confirm that our ChIP assay could detect large changes in Pol II, we tested the well-documented *GAL10* gene upon induction. *GAL10* is induced by addition of galactose to the media, and thus would be expected to have an increase in Pol II occupancy. We observed an increase in Pol II enrichment at the *GAL10* ORF upon induction (Figure 4.5*B*). This increase in Pol II levels was larger than the increase we observed at the Mot1-Spt16 co-repressed genes. Overall, our ChIP assay could detect a larger increase in Pol II levels at the *GAL10* gene, which indicates that the increase at the co-repressed genes is indeed modest, compared to the increase at *GAL10*.

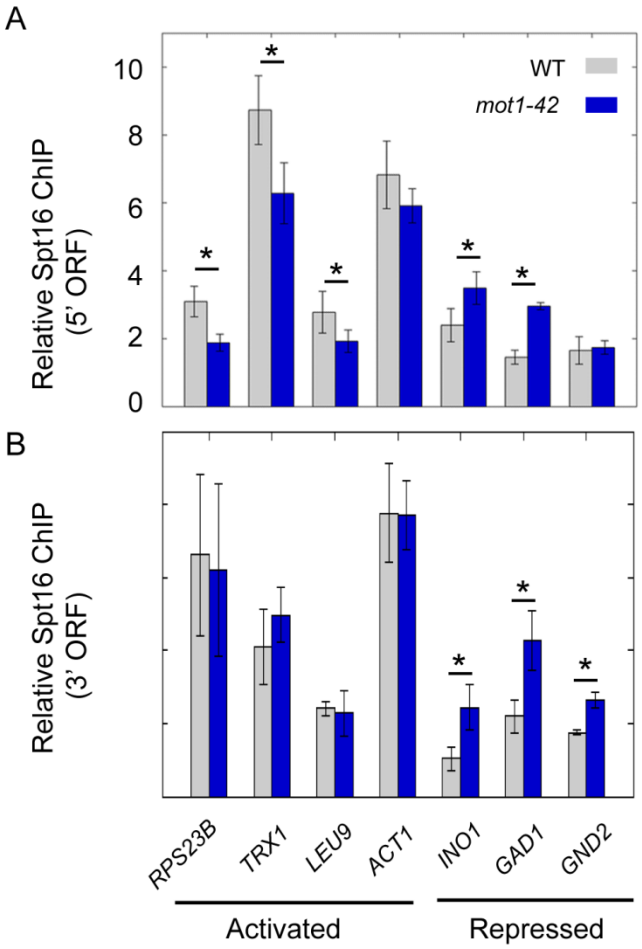

Figure 4.4. Spt16 levels were affected by Mot1 at co-activated and co-repressed genes at the 5' ORF but only in the co-repressed at the 3' ORF. *A*, Spt16-myc ChIP was performed in *SPT16*-myc *MOT1*-WT and *SPT16*-myc *mot1-42* cells at the 5' end of the open reading frame (ORF). Four Mot1-Spt16 co-activated genes and three co-repressed genes were analyzed. *B*, Experiment similar to *A* but for enrichment at the 3' end of the ORF. In *A* and *B*, asterisks indicate $p<0.05$ as determined by a Student's *t*-test using at least three independent replicates.

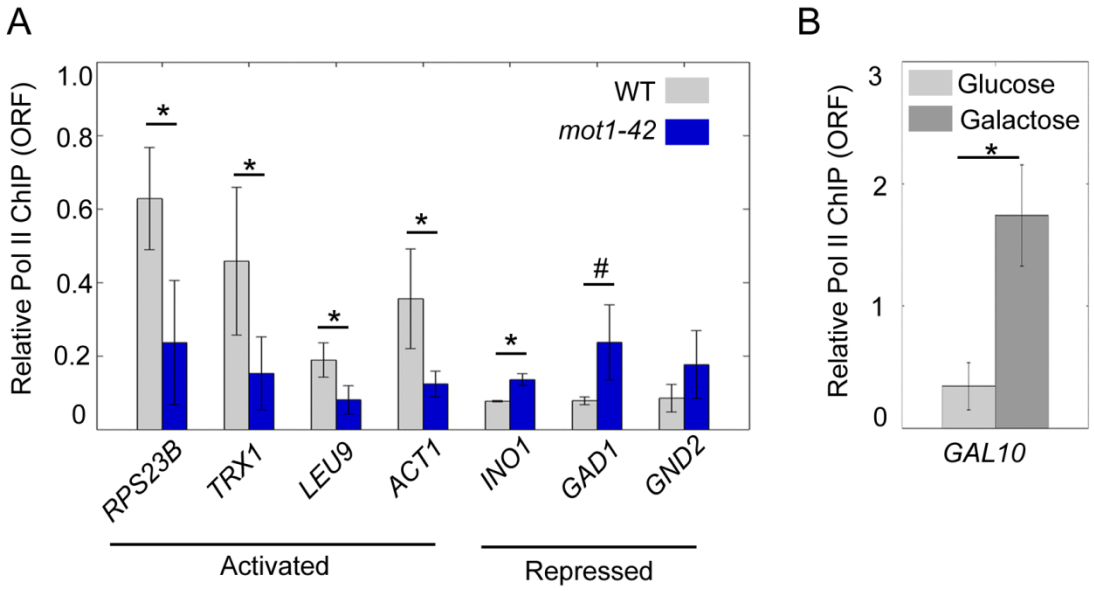

Figure 4.5. RNA Polymerase II levels were consistent with expression changes. *A*, ChIP of RNA Polymerase II (Pol II) in the open reading frame (ORF) of Mot1-Spt16 co-activated and co-repressed genes. *B*, Pol II ChIP at the *GAL10* ORF during glucose repression and galactose induction. In *A* and *B*, the asterisk indicates $p < 0.05$ from a Student's *t*-test. The # indicates $p = 0.055$.

We hypothesized that Spt16 established a permissive chromatin environment at gene promoters that allowed Mot1 to bind, and mutation of Spt16 caused a decrease in Mot1 levels as detected by ChIP-Seq (Figure 4.2). We determined the effects on nucleosomes in the mutant strains with MNase ChIP-Seq (Wal and Pugh 2012). Chromatin was digested by MNase to mononucleosomes, which were subsequently IPed by an antibody to histone H3, enriching for DNA fragments wrapped around histones. It is important to titrate the cells with MNase to achieve ~80% mononucleosomes in order to identify single nucleosomes without over digestion to

subnucleosomal fragments (Tsui et al. 2012, Wal and Pugh 2012). The titration was performed for all strains and replicates as shown in Figure 4.6. The samples with 400 units of MNase were selected as the optimal digestion for all strains and used for producing libraries.

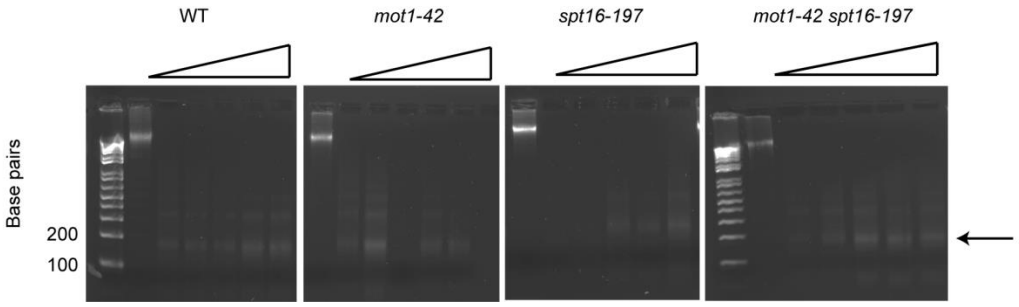

Figure 4.6. Titration of micrococcal nuclease (MNase) to yield optimal mononucleosomal levels in all four strains. Cells from all four strains were treated with increasing amounts of MNase (0, 100, 200, 300, 400, and 500 units) to digest the chromatin. Digested samples were run on a 2% agarose gel to separate the bands. The arrow on the right indicates the mononucleosome band at ~147 bp.

Before proceeding to the library preparation, we verified that the H3 IP had worked. We examined previously published nucleosomal data for *spt16-197* cells, which used MNase-Seq, to determine a gene to use in qPCR analysis (van Bakel et al. 2013). The *GAD1* ORF has a well-positioned nucleosome in WT cells that is depleted in *spt16-197* cells, and this region of the ORF had already been used in the Spt16-myc qPCR analysis (Figure 4.7). Therefore, we used this region of *GAD1* to determine if the IP was

successful in the strains and was depleted in the *spt16-197* cells compared to WT. As shown in Figure. 4.8*A*, there was a 13-fold enrichment of H3 over input and this decreased to only 4-fold enrichment in *spt16-197* cells. There was a 10-fold enrichment in *mot1-42* cells and 3-fold enrichment in *mot1-42 spt16-197* cells. This pattern of enrichment was expected based on the expression levels at *GAD1*, since increased expression correlates with loss of nucleosomes. After the ChIP-Seq library preparation, the enriched samples were validated once again by qPCR at the *GAD1* ORF, and the patterns were the same indicating that the library preparation did not alter the relative levels of DNA (Figure 4.8*B*).

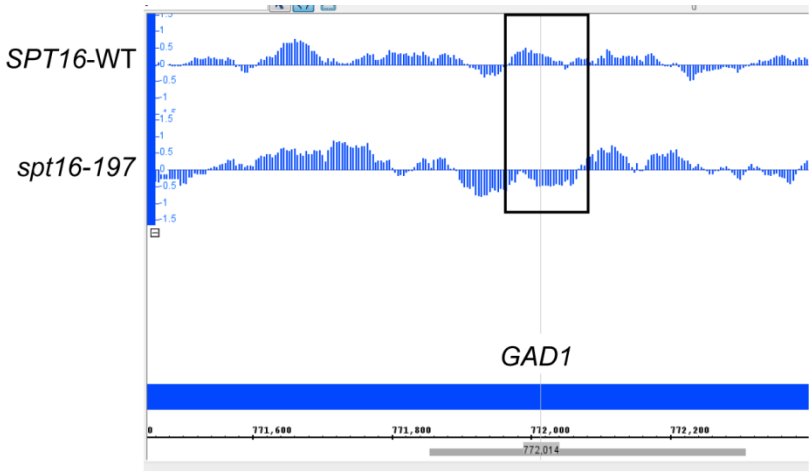

Figure 4.7. The *GAD1* ORF has a nucleosome that is depleted in *spt16-197* cells compared to *SPT16*-WT cells. Publically available nucleosome data (van Bakel et al. 2013) was used from *SPT16*-WT and *spt16-197* cells to determine if the H3 ChIP had worked properly.

Nucleosomal reads were plotted similar to Mot1 and Spt16 ChIP-Seq data with all genes lined up based on their TSS. The nucleosomal pattern for WT cells (Figure 4.9, green line) is similar to previously published data on wildtype nucleosomes (Yuan et al. 2005, Lee at al. 2007, Shivaswamy et al. 2008, van Bakel et al. 2013). The NDR is apparent upstream of the TSS, with well-positioned genic nucleosomes and to some extent upstream nucleosomes. When subdivided into Mot1-Spt16 co-activated and co-repressed genes, these patterns are still generally held. Co-repressed genes had more nucleosomes in the NDR compared to co-activated genes, as expected. Co-activated genes had a very striking peak for the +1 nucleosome compared to co-repressed genes, potentially emphasizing the role of the +1 nucleosome in gene activation.

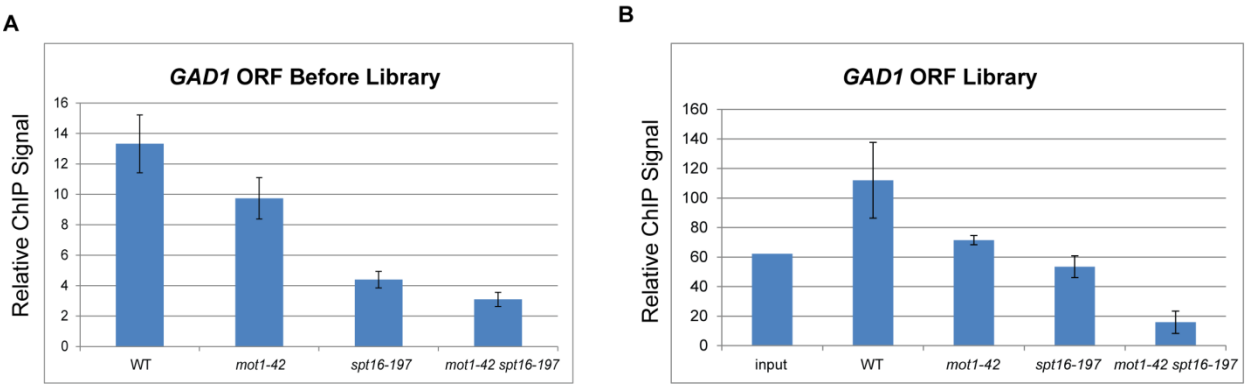

Figure 4.8. H3 ChIP signal in the *GAD1* ORF in all four strains before and after library preparation. *A*, H3 ChIP signal relative to wildtype (WT) input after MNase digestion. *B*, H3 ChIP after ChIP-Seq library preparation from samples in *A*.

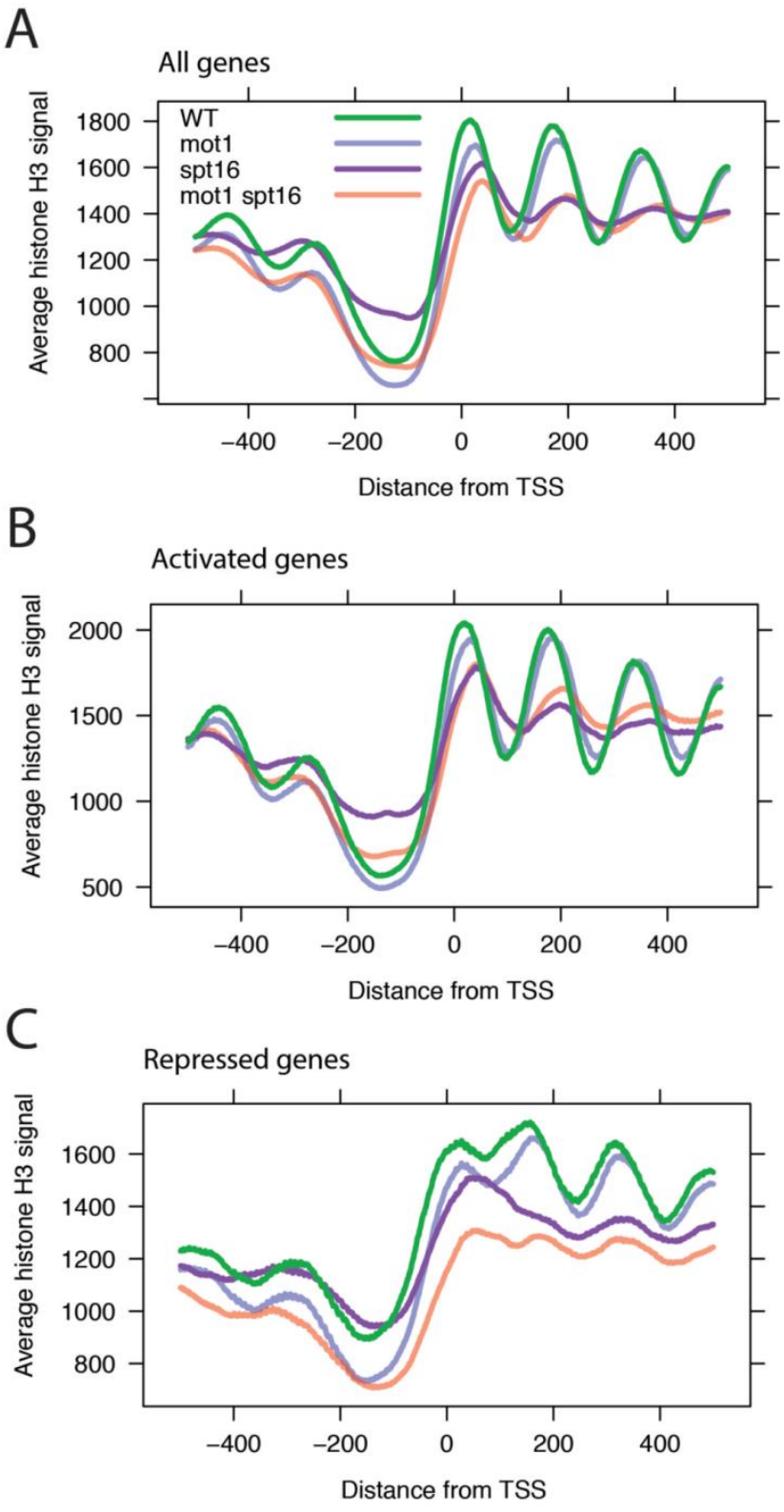
A
All genes
Average histone H3 signal
1800
1600
1400
1200
1000
800
WT
mot1
spt16
mot1 spt16
−400
−200
0
200
400
Distance from TSS
B
Activated genes
Average histone H3 signal
2000
1500
1000
500
−400
−200
0
200
400
Distance from TSS
C
Repressed genes
Average histone H3 signal
1600
1400
1200
1000
800
−400
−200
0
200
400
Distance from TSS

Figure 4.9. Mot1 and Spt16 regulated nucleosome organization genome-wide. *A*, MNase ChIP-Seq signal in all four strains lined up based on the TSS of all genes and extending +/- 600 bp from the TSS. WT nucleosomes (green line) were well-defined, decreased in signal in *mot1-42* cells (blue line), and lost periodicity in *spt16-197* and *mot1-42 spt16-197* cells. *B*, Plot similar to *A* but with only Mot1-Spt16 co-activated genes. *C*, Plot similar to *A* but with only Mot1-Spt16 co-repressed genes. *A-C*, Data analysis was performed by David Auble.

The *spt16-197* nucleosomal pattern has been well established already showing a loss of well-defined nucleosomes (van Bakel et al. 2013). This is consistent with our data showing that there is a loss of nucleosomal patterning genome-wide in *spt16-197* cells. In *mot1-42* cells the -2, -1, NDR, +1, and 2 nucleosomes all decrease, while the remaining nucleosomes are similar to WT levels. Based on the localization pattern of Mot1 with its enrichment in promoters, this is not surprising that Mot1 has an effect only on promoter-proximal nucleosomes. The double mutant cells showed a combined exacerbated effect on nucleosomes; nucleosomal periodicity was lost in the *spt16-197* cells and nucleosome intensity was decreased from both single mutants. This is consistent with the exacerbated effects on expression, TFIIB, and temperature sensitivity in the double mutant as shown in Chapter III.

The average nucleosomal plots shown in Figure 4.9 are useful in understanding general trends in nucleosomes affected by Mot1 and Spt16. However, these plots do not compare individual nucleosomes. To better dissect the roles of both factors in nucleosome organization, we looked at two parameters of nucleosome organization:

occupancy and positioning. Nucleosome occupancy is defined as the amount of histones in a specific location and usually correlates with gene expression. Nucleosome positioning takes into account the exact location of nucleosomes on DNA and how much they deviate from this midpoint, which is known as the fuzziness (Pugh 2010). As mentioned already, both parameters can be affected by DNA sequence and transcription factors.

The program DANPOS allows for identification of nucleosomal peaks in all strains (Chen et al. 2012). These peaks could then be compared for each mutant strain to the WT to determine the change in occupancy or fuzziness. Approximately 67,000 nucleosomal peaks were identified in the mutant strains (Table 4.3). Since each nucleosome is approximately 146 base pairs, this accounts for 9.78 million base pairs of the total yeast genome which is 12.1 million base pairs. Including the linkers means that our average linker is ~34.6 base pairs for a total length of 180.6 base pairs. Our results are consistent with the literature which shows that a nucleosome plus the linker is approximately 200 base pairs (Zhang and Pugh 2011).

Table 4.3
Total number of nucleosomes detected and affected in each strain
Data analysis performed by David Auble

Comparison	Total Peak #	Occupancy Δ (Peak #)	Fuzziness Δ (Peak #)	Peaks with Occupancy Δ that also had Fuzziness Δ	Peaks with Fuzziness Δ that also had Occupancy Δ
mot1 vs WT	66,237	3,763 (5.7%)	8,809 (13.3%)	41.80%	17.90%
spt16 vs WT	67,801	3,756 (5.5%)	11,359 (16.8%)	36.40%	12%
mot1 spt16 vs WT	67,216	4,460 (6.6%)	10,689 (15.9%)	52.20%	21.80%

The total numbers of nucleosomes that had changes in occupancy or fuzziness are listed in Table 4.3. There were approximately 2-3 times more differentially fuzzy nucleosomes than nucleosomes with changes in occupancy in all three mutant strains. This suggests that Mot1 and Spt16 contribute more to nucleosome positioning than nucleosome occupancy. The lower number of nucleosomes with a change in occupancy is consistent with the modest overall effects on expression in *mot1* and *spt16* cells. About half of the nucleosomes that had a change in occupancy also had a change in fuzziness. There was a lower overlap in nucleosomes that had a change in fuzziness that also had a change in occupancy.

Table 4.4

Effects on nucleosomes in mutants compared to wildtype
Data analysis performed by David Auble

Occupancy Change

Dataset Comparison	Increased in mutant	Decreased in mutant	Total	% Decreased in mutant
mot1-42 vs WT	1,315	2,448	3,763	65.1
spt16-197 vs WT	2,291	1,465	3,756	39
mot1-42 spt16-197 vs WT	957	3,503	4,460	78.5

Fuzziness Change

Dataset Comparison	Increased in mutant	Decreased in mutant	Total	% Increased in mutant
mot1-42 vs WT	7,570	1,239	8,809	85.9
spt16-197 vs WT	11,023	336	11,359	97
mot1-42 spt16-197 vs WT	10,484	205	10,689	98.1

The numbers in Table 4.3 only show changes in nucleosome occupancy and fuzziness, but do not specify whether the changes were positive or negative. Table 4.4 shows the nucleosomes that were increased or decreased in occupancy and fuzziness. The majority of nucleosomes in *mot1-42* and *mot1-42 spt16-197* cells decreased in occupancy. The genes that had a decrease in nucleosome occupancy most likely would be genes that are up-regulated in the mutants. However, the genes that are down-regulated in the mutants do not necessarily have to have an increase in nucleosome occupancy, because these genes would have less TBP and TFIIB bound to the promoter and would be down-regulated at that step rather than requiring nucleosomes to shut off transcription. The vast majority of differentially fuzzy nucleosomes had an increase in fuzziness in all three mutant strains (85-98%). This shows that the dominant effect that Mot1 and Spt16 have on nucleosomes is to properly position them.

	Occupancy Δ Overlap (%)	Fuzziness Δ Overlap (%)
mot1 vs. *spt16*	6.6*	35**
mot1 vs. *mot1 spt16*	19.4**	37.9**
spt16 vs. *mot1 spt16*	29.9**	74.6**

Figure 4.10. Changes in nucleosomal occupancy and fuzziness overlapped significantly among the mutants. The overlap in sets of nucleosomes affected in *mot1-42, spt16-197,* and double mutant cells. FDR-corrected *p*-values are < 0.007 (*) and < 2.6×10^{-196} (**). Data analysis performed by David Auble.

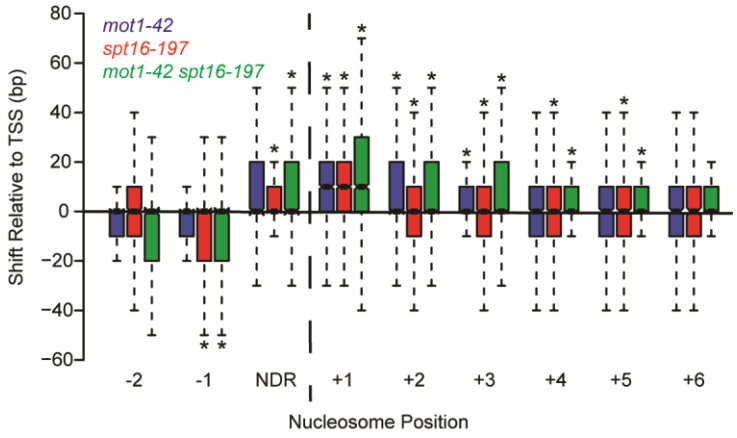

Figure 4.11. Mot1 and Spt16 regulated nucleosome positioning. All nucleosomes with significant changes in fuzziness in each mutant were assigned a nucleosome position based on the midpoint of the nucleosome in WT cells. The position shift was calculated based on how much the midpoint changed in the mutant cells compared to the WT cells in relation to the TSS. Each nucleosome position was grouped together, and the box plots show the distribution of the position shifts. The median is shown by the black line in the middle with the whiskers extending to 1.5X the interquartile range.

Since the focus of this project is to understand how Mot1 and Spt16 co-regulate gene expression, it is necessary to look at the similarly affected nucleosomes in each mutant (Figure 4.10). Only 6.6% of nucleosomes with changes in occupancy were the same between *mot1-42* and *spt16-197* cells which is low but statistically significant. There was a larger overlap in the differentially fuzzy nucleosomes between *mot1-42* and *spt16-197* cells, again emphasizing the importance of both factors in nucleosome

positioning. Both single mutants had significant overlaps with the double mutant, and they had a larger overlap with differentially fuzzy nucleosomes compared to nucleosomes with changes in occupancy.

Since Mot1 and Spt16 are important for nucleosome positioning, we next determined if they specifically regulated nucleosomes with respect to the TSS. The gene average profile in Figure 4.9 suggests that the +1 nucleosome was shifted away from the TSS in all three mutants. This shift in the +1 nucleosome was previously reported for *spt16* cells (van Bakel et al. 2013). The differentially fuzzy nucleosomes were designated -1, +1, +2, etc. based on the distance of the nucleosome midpoint to the TSS. Then the position shifts relative to the TSS were calculated for each mutant compared to WT. A +10 base pair median shift was detected at the +1 nucleosome in *spt16-197* cells, similar to published data (van Bakel et al. 2013). The +1 nucleosome was also shifted 10 base pairs away from the TSS in *mot1-42* and *mot1-42 spt16-197* cells (Figure 4.11). Thus both factors are required for properly positioning the +1 nucleosome and do so in independent manners. The +2 and +3 nucleosomes were significantly shifted in all three mutants, and the -1, +4, and +5 nucleosomes were shifted in *spt16-197* cells. While Spt16 has a more general role in regulating nucleosome position, Mot1 seems to be restricted to promoter-proximal nucleosomes, which makes sense based on its localization to the promoter.

Discussion

Mot1 and Spt16 Have Distinct Localization Profiles but Affect Each Other's Localization

Mot1 and Spt16 have previously been implicated in transcription initiation and elongation (Collart 1996, Muldrow et al. 1999, Poorey et al. 2010, Formosa et al. 2002, Jamai et al. 2009, Biswas et al. 2005). However, their distinct localization profiles have suggested that Mot1 is directly involved in initiation only and Spt16 is only directly involved in elongation. Based on the physical interaction described in Chapter III and the overlap in transcriptional regulation, we determined both factors' genome-wide localization and how they affected each other.

Several studies have utilized ChIP to determine if Mot1 interacted with specific genomic regions and genes that were differentially affected by Mot1. These studies showed that Mot1 localizes to the promoter alongside TBP and is present at the promoters it represses and activates (Andrau et al. 2002, Dasgupta et al. 2002, Geisberg et al. 2002, Zanton and Pugh 2004). Genome-wide ChIP studies have also confirmed Mot1's presence at the promoter of both types of genes (Geisberg and Struhl 2004, Zentner and Henikoff 2013). Our ChIP-Seq corroborates these previous results. Interestingly, though we do see a difference in the peak height and width of Mot1 at the promoter. At repressed genes Mot1 peaked sharply at ~300 base pairs upstream of the TSS. At activated genes Mot1 had a broader peak that is closer to the TSS and even continues somewhat downstream of the TSS. Activated genes also showed slightly more Mot1 in comparison to repressed genes.

As previously mentioned Mot1 preferentially represses TATA-containing promoters and activates TATA-less promoters. It has also been shown that these two classes of promoters have different chromatin landscapes. Furthermore, it has been proposed that the PIC assembles differently at TATA and TATA-less promoters, with the PIC competing with a nucleosome at TATA-containing promoters and assembling in conjunction with the +1 nucleosome at TATA-less promoters (Rhee and Pugh 2012). This could explain why Mot1 has a different pattern at these two classes of genes. TATA-containing genes that are repressed by Mot1 would have a discrete location for TBP to bind and thus Mot1 would bind in a specific location hence the sharp peak upstream of the TSS. TATA-less promoters would have a more ambiguous region for TBP to bind, since there is no defined promoter element for TBP to bind. If the PIC is assembling on top of the +1 nucleosome, this also explains why the peak for Mot1 is closer to the TSS and extends beyond it.

At repressed genes, Mot1 removes TBP by hydrolyzing ATP as discussed in Chapter II. It has been suggested that Mot1's action at activated genes is indirect, however as already mentioned, Mot1 binds at activated genes and to a larger extent than it does at repressed genes. This suggests that Mot1's role at co-activated genes is direct. Mot1 has been shown to help recycle TBP from the high-affinity TATA-binding sites to allow a dynamic pool of TBP to reach the TATA-less promoters thereby allowing for the activation of these genes (Zentner and Henikoff 2013). It has also been shown that Mot1 is required at genes like *URA1* to remove TBP from an improper orientation to help activate the gene (Sprouse et al. 2008b). A combination of Mot1 acting to chaperone TBP from TATA promoters to TATA-less promoters and then properly positioning TBP could explain Mot1's direct role in activation.

One surprising result from the Mot1 ChIP-Seq was the rather large, broad peak of Mot1 in the ORF of activated genes. Previous studies show conflicting results for Mot1 localizing to the ORF (Geisberg et al. 2002, Dasgupta et al. 2002). One study discovered that Mot1 localized to active coding regions and that TBP was not present in the intragenic regions where Mot1 was located. Genome-wide studies however have not shown that Mot1 localizes to the coding regions. The genes that are activated by Mot1 tend to be housekeeping genes or ribosomal genes which have higher transcription rates (Dasgupta et al. 2002, Poorey et al. 2010). It has been shown that almost any transcription factor can be cross-linked to a highly expressed gene without having functional significance (Teytelman et al. 2013). On the other hand, we have previously shown that Mot1 has an unexpected role in transcription elongation, so it is possible that this Mot1 peak is actually real and could indicate Mot1's elusive role in elongation (Poorey et al. 2010). At this point, we cannot distinguish between these alternative explanations for this peak.

First, we would need to divide the activated genes into groups based on expression to determine if there is an effect on Mot1 signal that correlates with the transcription rate. This could help to determine if this is indirect due to cross-linking to highly expressed genes. We could also compare these results to those from the Mot1 ORGANIC-Seq study (Zentner and Henikoff 2013). ORGANIC-Seq takes into account the problems with formaldehyde and instead uses a native approach without cross-linking. When the results from the native approach are divided similarly to our results, we could see if there are any peaks in the ORF in their data that would corroborate our data. We could also divide our data based on length of transcripts to determine if there is enrichment in shorter transcripts. If there are more shorter transcripts in the activated

class it is possible that we are seeing a peak of Mot1 at the beginning of the next gene. If this peak of Mot1 in the ORF has biological significance , then we need to examine if there is TBP located at these genes or cryptic TATA boxes. Finally, it is possible that this peak of Mot1 is a true peak without TBP, as shown in the previous ChIP study that first showed Mot1 in the coding region. This would then require further examination to understand how Mot1 is binding without TBP and if it is associating with other proteins that could lead to a better understanding of Mot1's role in elongation.

Next we determined how the localization of Mot1 was affected by mutation of Spt16. In the absence of Spt16, the Mot1 signal decreased in the coding region of activated genes and increased in the coding region of repressed genes (Figure 4.2). The decrease in activated genes was consistent with the decreased expression. Again, this could be a decrease in non-specific cross-linking due to the decrease in expression, or this could be a real decrease in Mot1 binding. The increase in Mot1 in the coding region of repressed genes could be attributed to the increase in expression. As discussed in Chapter III, genes with TATA boxes in the promoter tend to have more TATA boxes within the coding region. Therefore, the increased signal of Mot1 could be due to exposure of cryptic binding sites. This could be examined further by looking to see if there are cryptic binding sites in these genes and if TBP localizes to these coding regions in the absence of Spt16.

Co-activated and co-repressed genes showed a decrease in Mot1 signal at the promoter in *spt16-1*97 cells (Figure 4.2), although activated genes had a much more pronounced decrease in Mot1. We have previously proposed that Spt16 could affect the chromatin organization at the promoter of genes it regulates. It is therefore possible that Spt16 is required to set up a permissive landscape for Mot1 to be able to bind to the

promoter. This suggests that regardless of whether these factors are involved in activation or repression of these genes, that Spt16 is involved in properly positioning the nucleosomes. This will be expanded upon in the next section of this discussion.

A problem with this explanation is the fact that TBP recruitment to the promoter is unaffected in *spt16-197* cells as shown in Chapter III. Since Spt16decreased Mot1 localization to the promoter of differentially expressed genes, TBP localization should also change. Since the Mot1 localization was assayed using ChIP-Seq and the TBP localization was examined by ChIP-qPCR it is possible there is a difference in sensitivity. Therefore TBP localization should be examined genome-wide in *spt16-197* cells to determine with higher sensitivity how TBP is affected. This could also explain if cryptic binding sites were exposed and if TBP was recruited intragenically in the absence of Spt16.

Our results are very similar to the previously published results which show that Spt16 was enriched in the coding region of genes and depleted at the promoter (Mayer et al. 2010). It is understandable that Spt16 would be depleted at the promoter due to the relative depletion of nucleosomes in this area. The relative enrichment of Spt16 at activated genes compared to repressed genes is also not surprising, since Spt16 has been linked with Pol Ii and highly expressed genes (Mason and Struhl 2003, Saunders et al. 2003). The largest peak of Spt16 was immediately downstream of the TSS. Though Spt16 does not seem to be present at the promoter, the close proximity of this Spt16 peak suggests that Spt16 is important for early elongation and potentially late initiation. Spt16 levels started to increase ~300 bp upstream of the TSS which also suggests that Spt16 could play a role in chromatin organization immediately upstream of the NDR.

In the absence of Mot1, Spt16 levels decreased at activated genes and increased at repressed genes (Figure 4.3). This is consistent with the expression changes at these genes and also with the changes in Pol II levels in the ORF. We verified these changes at a subset of activated and repressed genes which show slight but significant changes in Spt16 levels at the 5' end of the ORF and also at the 3' end of repressed ORFs (Figure 4.4). However, there was no significant change in Spt16 levels at the 3' end of activated genes when analyzed by ChIP-qPCR. This is in contrast to the ChIP-Seq average plot shown in Figure 4.3. Again, this could be due to lack of sensitivity in ChIP-qPCR compared to ChIP-Seq. It is also possible that at some of the activated genes, Spt16 levels are not affected towards the 3' end of a gene and the average plots hide these genes.

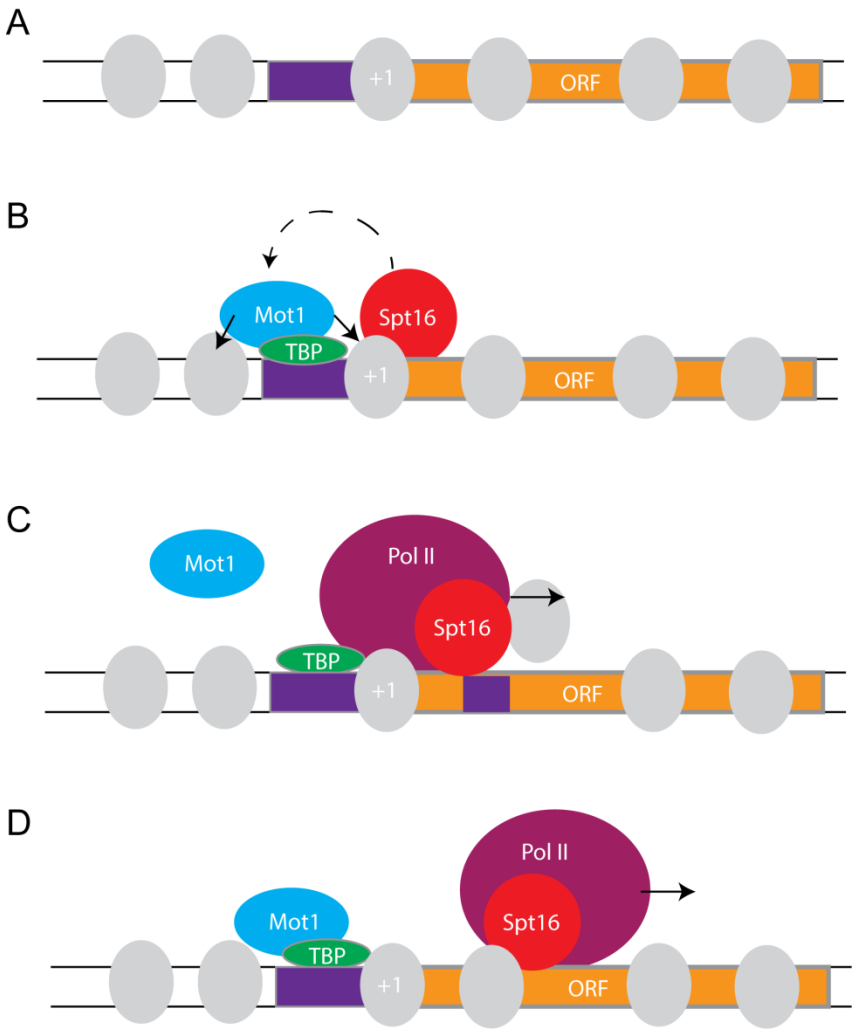

Figure 4.12. Model for Mot1-Spt16 co-regulation of transcription. *A*, A gene with a promoter (purple box), the open reading frame (ORF; orange box), and nucleosomes (gray circles). *B*, Spt16 promotes Mot1 binding to the promoter by reorganizing nucleosomes. Mot1-TBP bind to the promoter. Mot1 regulates the promoter-proximal nucleosomes (black arrows). *C*, After assembly of the PIC, Pol II elongates with Spt16. The purple box in the ORF represents a cryptic site that is exposed when Pol II passages through the gene but is immediately blocked by Spt16 replacing histones after Pol II has elongated. *D*, Pol II continues to elongate and Mot1 can associate with the promoter once again.

Chapter V

Conclusions and Future Directions

Overall, the biochemical and genome-wide studies described here explained more about how Mot1 regulates transcription by itself or through interaction with Spt16. While the general organization of the ATPase domain of Mot1 had been determined using EM and FeBABE (Wollman et al. 2011), we demonstrated that the ATPase domain transitions from an open complex to a closed complex (Figures 2.3-2.6). This conformational change in the ATPase domain is mediated by the presence of nucleotide or the DNA sequence to which it binds (Figure 2.4). These results were the first to show evidence of the two conformations of Mot1, which had been previously alluded to by crystal structures of the SsoRad54 ATPase (Durr et al 2005). In addition to this, we demonstrated that Mot1 and Spt16 similarly regulate a significant number of genes in *S. cerevisiae* (Figures 3.5 and 3.7). Mot1 and Spt16 had different genomic localizations (Figures 4.2 and 4.3); however, Spt16 was required for Mot1 recruitment to the promoter (Figure 4.2), and Mot1 affected Spt16 recruitment to co-regulated genes (Figure 4.3 and 4.4). This was also the first time that Mot1 was demonstrated to affect nucleosome organization at a genome-wide scale (Figures 4.9-4.11, Tables 4.3 and 4.4). While these experiments helped us to better understand transcriptional regulation by Mot1, there are new questions that need to be addressed in the light of these results.

Mot1 in solution increased dissociation of TBP from DNA on the shorter 5' Bio duplex but not on the longer 5' Tail duplex (Figure 2.1). The increase in TBP dissociation in the presence of excess Mot1 led to the conclusion that either increasing the local concentration of Mot1 enabled a rapid binding of multiple Mot1 molecules to

dissociate TBP or that Mot1 dimerized to dissociate TBP. Both of these ideas should be tested. We could vary the concentration of Mot1 in solution to determine if increasing the Mot1 concentration was allowing for multiple Mot1 molecules to bind in succession to dissociate TBP. Lesser concentrations of Mot1 should be less effective, while greater concentrations should increase dissociation. Since there is an unpublished crystal structure of Mot1 forming a dimer, we could also test this experimentally. However, the absence of dimers in gel filtration chromatography suggests that Mot1 does not form a dimer in vivo. A co-immunoprecipitation experiment should be performed with differentially tagged forms of Mot1 to determine if it can form dimers. This would allow for us to conclude if the Mot1 dimer crystal structure was an artifact or had biological significance. I think that Mot1 most likely does not form a dimer and that crystal structure is artefactual.

The sequence dependence of the Mot1 conformational change is the most intriguing part of Chapter II in my opinion (Figures 2.4 and 2.5). Since the three templates used had varying GC content in the upstream DNA, I propose experiments using different combinations of GC content or templates with the same GC content and different sequences. If the conformational change is affected by the structure of the DNA which is impacted by the GC content, then I would expect similar results as we have already shown. Lower GC content would favor the formation of the open complex, whereas higher GC content would favor the formation of the closed complex. Since the Mot1 ATPase interacts with the phosphate backbone of the upstream DNA (Dürr et al. 2005), I do not think the conformational change would be affected by changing the sequences while maintaining the GC content. These experiments were done *in vitro*, but they could be compared to the promoter sequences *in vivo* of Mot1-regulated genes. Do

Mot1-regulated promoters have varying GC content? It is possible that activated and repressed genes have different GC content that could affect the conformation in which Mot1 binds to these promoters. If Mot1 binds in the open conformation and only a low proportion can transition to the closed conformation, does that mean Mot1 is less efficient at removing TBP at these promoters (*HIS4*)? Since *HIS4* is a Mot1-activated gene, this could indicate a difference between activated and repressed genes and how Mot1 distinguishes between them.

Our results demonstrated biochemical evidence of a conformational change in the Mot1 ATPase domain that was dependent on DNA sequence or nucleotide (Figures 2.3-2.6). However, we only show the appearance of a cleavage product that suggests this conformational change. It would be more definitive if we could identify a cleavage product in the open conformation that is lost when the ATPase rotates to form the closed conformation. Based on the closed conformation being cleaved when phosphorothioates are positioned at nucleotides 20 and 22 upstream of the TATA, I would position new phosphorothioates at nucleotides 28 and 30 upstream of the TATA. These positions are based on the model of SsoRad54 in the open and closed conformations (Dürr et al. 2005). This should enable for cleavage of the open conformation, which would be lost when it transitions to the closed complex and provide more support for this model.

Chapters III and IV focus on the functional interaction between Mot1 and Spt16. However, as mentioned in Chapter I, Spt16 is part of the FACT complex. Mutations in Spt16 and Pob3 can lead to different phenotypes (Formosa et al. 2002), but these two factors physically interact and work together in transcription regulation. Therefore, we focused only on the Spt16 component for our studies. We did demonstrate that Mot1

can physically associate with Pob3 and Nhp6 (Figure 3.3). Furthermore, deletion of *MOT1* and both *NHP6* genes is synthetically lethal (Biswas et al. 2006), indicating that these two factors work in parallel pathways. We should determine if *MOT1* and *POB3* also genetically interact. This would further support Mot1 working with the entire FACT complex instead of just functioning with Spt16.

The synthetic lethality observed in Chapter III (Figure 3.4) was not fully explained by our data. However, we observed the synthetic lethality at 30°C and performed the remaining experiments at 35°C. Since *mot1-42* cells and *spt16-197* cells are both ts at 35°C, this would complicate our results. We should initially repeat the expression analysis at 30°C to determine if there was an additive effect on transcription at the co-regulated genes in Figure 3.9. We could then look at TBP and TFIIB levels at those genes at 30°C also. I would expect that the effects on expression and TFIIB would be additive at 30°C which would explain the synthetic lethality.

The physical association between Mot1 and Spt16 demonstrated in Chapter II (Figure 3.2) does not indicate a direct interaction between the two factors. This would need to be determined using purified proteins to observe if this interaction is direct. I would expect that this interaction is direct based on a previous *in vitro* experiment that suggested it was direct (Wells 2012). If this interaction is direct, different portions of the proteins could be incubated together to determine which domains interact. I would also suggest determining how this interaction affects the Mot1-TBP interaction and the Spt16-Pob3-Nhp6-nucleosome interaction. I would hypothesize that the HEAT repeats of Mot1 would interact with Spt16 which prevent Mot1 from interacting with TBP. However, I am unsure which domain of Spt16 would interact with Mot1. Since different domains in

Spt16 are responsible for H2A-H2B and H3-H4 binding, I would expect that Spt16 would still be able to bind some histones but not the entire octamer.

In Chapter IV, we demonstrated that Mot1 and Spt16 contribute to genome-wide nucleosome organization (Figures 4.9-4.11, Tables 4.3 and 4.4). This conclusively demonstrates that Mot1 can affect nucleosome organization on the genomic scale. Previously this was only shown at one gene with contradictory results at another gene (Topalidou et al. 2004, Sprouse et al. 2008b). While Mot1 is not necessary for nucleosome organization at all genes, it does play a significant role in nucleosome regulation. Since the nucleosomes affected by Mot1 were promoter-proximal, I would suggest repeating the MNase ChIP-Seq experiment and using an antibody for Htz1 (H2A.Z). Htz1 is enriched at the +1 nucleosome, and since Mot1 was demonstrated to regulate the positioning and occupancy of the +1 nucleosome, I would expect that it regulates Htz1-containing +1 nucleosomes. The Mot1-Htz1 interaction could be further explored to determine if Mot1 preferentially regulates these nucleosomes compared to canonical nucleosomes. If Mot1 does have a preference for Htz1-containing nucleosomes, then experiments should be done *in vitro* to determine if these nucleosomes are remodeled rather than the canonical octamers that have been studied in these assays.

Finally, I would propose another experiment to better understand Mot1's role in nucleosome regulation. In *mot1* cells, TBP is more dynamic (Poorey et al. 2013), binding and coming off of DNA. This increase in TBP dynamics could affect the nucleosomes in the promoter. If TBP is rapidly binding and coming off the DNA, this could prevent nucleosomes from being able to bind stably. This suggests that the Mot1 effect on nucleosomes is indirect due to TBP dynamics. However, it is also possible that

Mot1 has a direct effect on nucleosomes. I suggest utilizing the Mot1 bypass alleles of TBP to distinguish between these two ideas. TBP F207L enabled cells with a MOT1 deletion to grow (Sprouse 2008c). I would repeat the MNase ChIP-Seq experiment in this strain to determine the effects on nucleosome organization in the absence of Mot1. While this strain has TBP F207L, it also has WT TBP present in the cell. Both the WT and mutant TBPs should not hinder the results, since the cells can grow almost as well as WT (Sprouse 2008c). The point of this experiment is to determine the effects on nucleosomes in the absence of Mot1 while TBP is still present and functioning similar to WT. If Mot1 directly regulated nucleosomes, I would expect the results to be the same as shown in Chapter IV (Figure 4.9) for *mot1-42*. However, if the Mot1 effect on nucleosomes was indirect due to TBP, I would expect the results from this new experiment to more closely resemble WT nucleosomal organization.

Overall, I would say my largest contribution to the field of transcription was showing that Mot1 plays a genome-wide role in nucleosome organization. Before this only two studies with contradictory results had been performed to determine if Mot1 affected nucleosome organization. Furthermore, *in vitro* experiments had not been able to demonstrate that Mot1 was capable of remodeling a nucleosome (David Auble, unpublished results). My other main contribution was showing that DNA sequence could affect the conformation of Mot1 binding to DNA. This had been previously hypothesized by our lab (Sprouse et al. 2008b). The results in Chapter II show that these various sequences impact the conformation, however, whether it is due to differences in flexibility of the upstream DNA or sequence selectivity for the ATPase domain is still not known. I think that these differences in conformational binding could lead to a better

understanding of how Mot1 distinguishes between genes it activates or represses after a thorough look at the GC content of Mot1-regulated promoters.

Appendix A

Gel Detection of NC2 Effects on Mot1 Activity

This appendix serves to document the initial work done using SYPRO Ruby and silver staining to detect the effects of the NC2 complex on Mot1 activity.

As discussed in Chapters I and II, the Mot1 ATPase regulates transcription by its effects on TBP localization and dynamics (Auble et al. 1994, Poorey et al. 2013). The Negative Cofactor 2 (NC2) complex was initially identified through its repressive function on transcription, but was later shown to also be required for transcriptional activation (Davis et al. 1992, Auble et al. 1994). Mot1 and NC2 preferentially repress TATA-containing genes and activate TATA-less genes (Meisterernst and Roeder 1991, Collart 1996, Poorey et al. 2010). Both factors also regulate TBP localization; Mot1 binds TBP and hydrolyzes TBP thereby removing it from DNA, while NC2 prevents TFIIA and TFIIB from binding TBP (Meisterernst and Roeder 1991, Kamada et al. 2001). NC2 can also act as a TBP shuttle to localize TBP to activated genes. Experiments using whole cell extracts enabled isolation of a Mot1-TBP-NC2 complex which suggested that these factors could physically interact. Locus-specific and genome-wide studies showed that both Mot1 and NC2 could localize to promoters at the same time (White et al. 1994, Geisberg et al. 2001, Dasgupta et al. 2002, Geisberg et al. 2002). Furthermore, a recent crystal structure of Mot1-TBP-NC2-DNA was determined (Figure A.1; Butryn et al. 2015). Because of the co-localization of Mot1 and NC2 and their association in a complex, we tested whether NC2 affected the activity of Mot1.

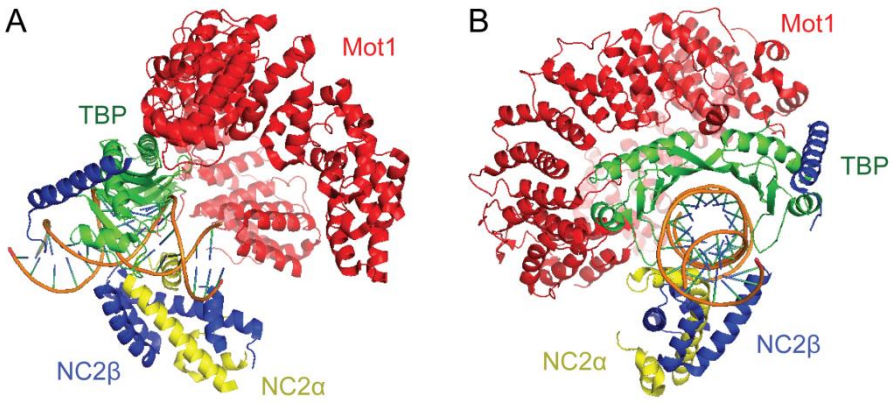

Figure A.1. Crystal structure of Mot1 NTD-NC2-TBP-DNA. *A*, The crystal structure (PDB = 4WZS) of the N-terminal Domain (NTD) of Mot1 (red) in complex with both subunits of NC2 (α, yellow, β, blue), TBP (green), and DNA (orange and blue) is depicted. *B*, A different view of *A* to illustrate different sides that Mot1 and NC2 contact of TBP-DNA.

Materials and Methods

Immobilized Template Assay and Gels

The immobilized template assay was performed in the same manner as described in Chapter II with a few modifications. The biotinylated DNA duplex from Chapter II was used for the initial experiments that are not shown in this dissertation due to the difficulties with this assay as described later. The Mot1, NC2α, NC2β, and TBP proteins were purified from *Encephatlizoon cuniculi* (Ecu) as previously described (Wollmann et al. 2011). The reaction buffer used for the Ecu experiments was 50 mM Tris-HCl pH 8, 100 mM NaCl, 10 mM $MgCl_2$, 4% glycerol, and 1 mM DTT

SYPRO Ruby and Silver Staining

After the gels were done running, they were stained either with SYPRO Ruby fluorescent dye or silver stain. The SYPRO Ruby staining was done as suggested by the manufacturer and is briefly described. The gel was incubated in fixing solution and gently shaken for 15 minutes at room temperature twice. Approximately 60 mL SYPRO Ruby was added to the gel, the dye and gel were microwaved for 30 seconds, shaken for 30 seconds, microwaved again for 30 seconds, and then shaken gently for 5 minutes. This was again microwaved for 30 seconds and shaken for 23 minutes. The gel was washed in water for 30 minutes while shaking and finally washed twice with water. The fluorescent gel was imaged using the Typhoon PhosphorImager (GE Healthcare) using 488 nm excitation and the 610BP30 emission filter. Quantification was done using the ImageQuant software (GE Healthcare).

Results

In Chapter II, I used the immobilized template assay to provide a better understanding of how Mot1 functioned in purified complexes and also to map the region of Mot1 in contact with upstream DNA. This assay utilized Western blots of the components bound to the DNA-bead complex or in the supernatant. For these experiments we used antibodies specific to ScTBP and the C-terminal portion of ScMot1 (Reddy and Hahn 1991, Auble et al. 1997). However, the purified ScNC2 that we have does not contain a tag and no antibody is specific to the protein or the two different

components. Therefore, we adapted the immobilized template assay to work around these problems.

The first change we made to the original assay was to use proteins from *Encephalitozoon cuniculi*, which were used in the crystal structures of both the Mot1 NTD-TBP and Mot1 NTD-NC2-TBP-DNA (Wollmann et al. 2011, Butryn et al. 2015). The EcNC2 complex was purified and had been biochemically shown to function, in contrast to the ScNC2 protein that we had available. Since none of the Ecu proteins were tagged or had antibodies specific to them, we decided to use a fluorescent protein dye, SYPRO Ruby, which would allow for accurate quantitation of nanogram (ng) levels of proteins (Lopez et al. 2000). The EcMot1 and EcNC2 proteins were run along a denaturing gel, stained with SYPRO Ruby, and imaged using the Typhoon PhosphorImager and ImageQuant software (GE Healthcare). Figure A.2 shows an example of a fluorescent gel with various concentrations of both EcMot1 and EcNC2. Both proteins were detected at as low as 5 ng, however this was not as clear as we had originally expected. Higher concentrations of both proteins were easily detected (25+ ng), but this could also be obtained with easier protein dyes such as Coomassie blue.

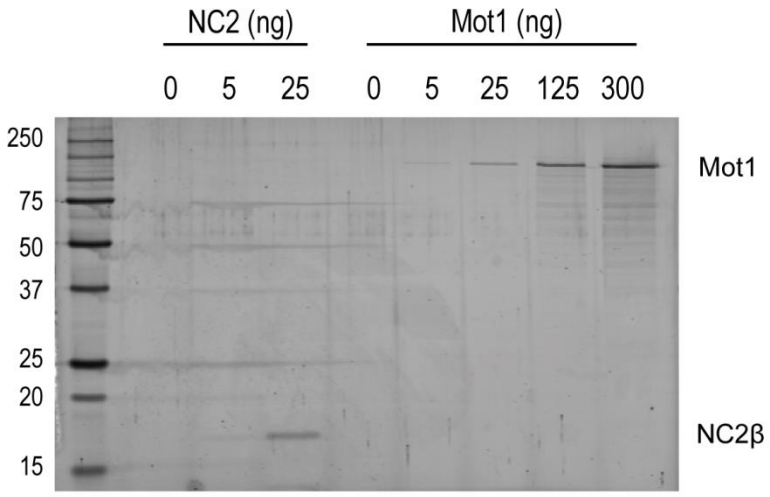

Figure A.2. SYPRO Ruby can detect ~5 ng of NC2 and Mot1. EcNC2 and EcMot1 were separated using a denaturing gel, which was then incubated in SYPRO Ruby. Proteins were barely detected at 5 ng. EcMot1 is approximately 210 kDa and the NC2 complex consists of two subunits which are 10.8 kDa (α) and 17 kDa (β). This gel only shows detection of the larger subunit.

One of the problems with the gel shown in Figure A.2 is that there was only detection of the larger subunit of the NC2 complex. We decided to use a 16% SDS-PAGE gel to enable separation of these two subunits and determine if both could be identified using SYPRO Ruby. As shown in Figure A.3, both subunits of NC2 were well resolved and detected. However, the lowest level detected shown here was 25 ng for NC2B and 50 ng for NC2A. This was extremely troubling and perplexing.

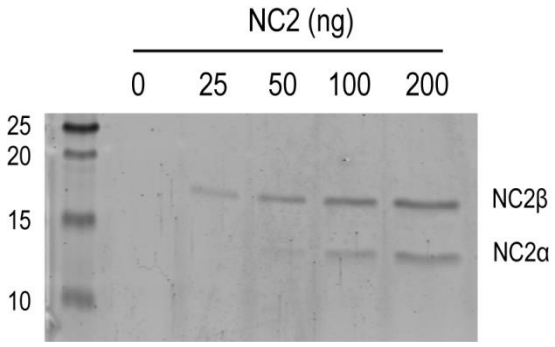

Figure A.3. NC2 subunits can be separated and detected via SYPRO Ruby. A 16% SDS-PAGE gel was utilized to separate the two NC2 subunits. NC2α is 10.8 kDa and NC2β is 17 kDa as shown to the right of the gel. Various concentrations were loaded (as labeled on the top) to determine the range of SYPRO Ruby detection. The lowest level of NC2β detected was 25 ng and NC2α was 50 ng.

Over the course of using the SYPRO Ruby stain, it was evident that there was a lot of variation in the detection of proteins. Because of these technical difficulties, we decided to try another staining method that could potentially be more reliable. Silver staining is commonly used to identify and quantify proteins run on gels. Therefore, I did a comparison of the two staining techniques using the EcMot1 protein. Figure A.4 shows a direct comparison of the same concentrations of EcMot1 stained by either silver stain or SYPRO Ruby. Using silver staining we were able to detect 10 ng of EcMot1. SYPRO Ruby was able to faintly detect 10 ng of EcMot1, but the 50 ng was the lowest level that is more apparent. The SYPRO Ruby variability was again frustrating, whereas silver staining was not able to yield lower levels of detection, which is what we wanted in the first place.

Regardless, we decided to compare the quantification capabilities of the two stains. Using the Typhoon PhosphorImager and ImageQuant (GE Healthcare), we were able to quantify the intensity of all the bands and get a best fit line (Figure A.5). While the silver stain yields a clearer image of the protein bands, the R^2 value for the best fit line was only 0.8252. SYPRO Ruby provided a best fit line with an R^2 value of 0.9999 showing its usefulness for protein quantification.

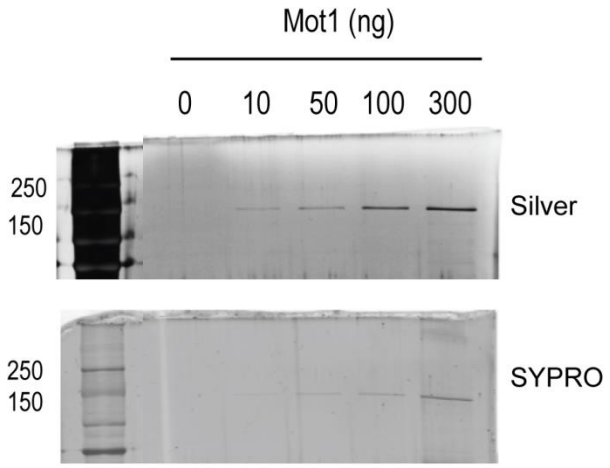

Figure A.4. Silver staining and SYPRO Ruby were able to detect similar amounts of protein. EcMot1 was separated on two different gels in various concentrations as listed on the top. The gels were then stained with either silver stain (top gel) or SYPRO Ruby (bottom gel). The gels were imaged using the Typhoon PhosphorImager (GE Healthcare). As shown both stains were able to detect as low as 10 ng of EcMot1.

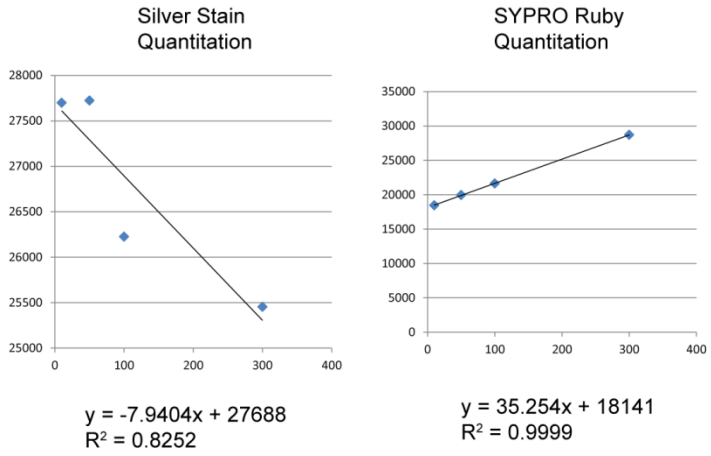

Figure A.5. SYPRO Ruby yielded better quantification than silver staining. The images in Figure A.3 were analyzed using ImageQuant (GE Healthcare) to determine the intensities of the protein bands. Best fit lines were determined for each gel. The silver stained gel yielded a best fit line with an $R^2 = 0.8252$, whereas the SYPRO Ruby stained gel had a better fit with an $R^2 = 0.9999$.

Overall, the SYPRO Ruby and silver staining techniques did not enable us to detect the effects of NC2 activity on Mot1. However, I think a combination of these techniques could be useful in future protein detection assays. While the silver staining procedure requires more components, it is easier to see the lower amounts of protein in the final image. SYPRO Ruby gels are not as clean and clear compared to the silver stained gels; however the quantification from SYPRO Ruby is better than silver stained gels.

Simultaneously while these experiments were underway, another approach was being taken to understand the effects of NC2 on Mot1 activity. David Auble was able to

show that Mot1 was able to dissociate a Mot1-TBP-NC2-DNA complex in the presence of DNA on DNA templates that were blocked on either end with a digoxigenin antibody (Butryn et al. 2015). This was done using the standard EMSA approach and suggests that Mot1 does not require NC2 to move TBP along the DNA to dissociate it (Butryn et al. 2015). Furthermore, this shows that the NC2-TBP-DNA complex is a target of Mot1-mediated dissociation. In conjunction with these experiments, it was demonstrated that NC2 does not affect the ATPase activity of Mot1 (Butryn et al. 2015).

Appendix B

Methods for Computational Analysis of Mot1-Spt16 Co-Regulation

Some of the methods described here were described in brief in True et al. 2016 and Chapter II

This appendix serves to document the various computational methods used for Chapters III and IV in order to assist future computational work in the lab.

Tiling array analysis using Affymetrix TAS

The analysis of Affymetrix tiling arrays is straight-forward thanks to the development of the Tiling Analysis Software (TAS) by Affymetrix. This program is available online for free at http://www.affymetrix.com/estore/partners_programs/programs/developer/TilingArrayTools/index.affx. The data originally obtained from tiling array analysis consists of CEL files for each individual replicate. These files are intensities of the probes that were bound by the transcripts. The CEL files have to be processed in order to visualize the results. Along with the CEL files, Affymetrix also provides BPMAP files that contain the genomic maps for various model organisms and the different builds. For multiple replicates for an experiment, you want to include all of the CEL files for the intensity analysis.

Before running the analysis, we changed a few parameters that were specific to our analyses. Only the signal values are required and can be saved in the BAR format, which can be viewed in the Integrated Genome Browser (IGB). The normalization metric should be changed to 100 with the signal scale set to the $\log_2$ setting. We smoothed the

signal profiles over a 50 base pair (bp) bandwidth (101 bp sliding window), whereas the differential profiles were smoothed over a 250 bp bandwidth (501 bp sliding window). Finally we performed a two-sided test plus or minus 0.3 with a 100 maximum gap and 50 bp minimum run. After these settings are selected, the intensities were analyzed and generated signal files for each strain.

For comparison of two strains, we used this same program to generate differential profiles. This will subtract the signal of one strain from another strain. To do this you have to select two sample comparison analyses and input your treatment replicates and control replicates. Use the settings as above to generate the differential profiles. We generated the Mot1 differential signal in this manner by subtracting the *mot1-42* signal ($\log_2$ scale) from *MOT1*-WT signal. The Spt16 differential signal was generated by subtracting *spt16-197* signal from *SPT16*-WT signal. Two double mutant differentials were generated by subtracting the *mot1-42 spt16-197* signal from each of the two WT signals.

The result is BAR files for each strain and each differential comparison that was performed. These can be loaded into IGB. It is important to load the same genome in IGB that was used for this analysis.

Statistical analysis using R

For portions of this dissertation I used RStudio for statistical analyses. For these analyses I usually input an Excel file with the values that I was comparing for the analysis. In order to do this, you must load the library “xlsx” shown by the following code:

```
library(xlsx)
```

Then you can enter the Excel file with the relevant data into R using this code:

```
nameoffile<-read.xlsx("C://location of file/nameoffile.xlsx",sheetnumber)
```

For Figures 3.9, 4.4, and 4.5, , I utilized a Student's *t*-test. After the relevant data were loaded into R, a Student's *t*-test on the two variables that to be compared (i.e. WT RNA at *RPS23B* to *mot1-42* RNA at *RPS23B*) was used to determine if there was a statistical difference using a cutoff of $p<0.05$. In this example I input the RNA expression values into a variable (either WTRNA or mot1RNA) and then the column names were that of the genes ($RPS23B). The $ calls a column from within the RNA expression lists in this case.

```
t.test(mot1RNA$RPS23B~WTRNA$RPS23B)
```

Another example would be the Wilcoxon test which does not assume that the data is normally distributed or has an equal variance.

```
wilcox.test(mot1RNA$RPS23B~WTRNA$RPS23B)
```

As a side note, there are other tools for determining if there is a significant difference between two samples. One useful calculator is available at http://www.graphpad.com/quickcalcs/ttest1.cfm.

In the comparison of TBP and TFIIB levels in Chapter III (Figure 3.13), I wanted to compare the levels of these components among the mutants and the WT. In order to do this I used a pairwise *t*-test. Input Excel data as described above. This time I

included a column with the strain names (i.e. WT, *mot1-42*, *spt16-197*, double). Then I called the pairwise *t*-test to compare the TBP levels in all 4 strains at the *RPS23B* gene:

```
pairwise.t.test(tbp$RPS23B, tbp$strains)
```

Or the Wilcoxon test can be applies as described above.

```
pairwise.wilcox.test(tbp$RPS23B, tbp$strains)
```

Average localization of Mot1 and Spt16 using Matlab

While the average localization profiles of Mot1 and Spt16 were done using ngs.plot-2.4.7 (Shen et al. 2014), the original results were generated using Matlab.

For analysis in Matlab, I used Spt16 ChIP-chip data (Mayer et al. 2010) and Mot1 ORGANIC-Seq data (Zentner and Henikoff 2013) to map their respective localizations. The original files were GFF (Spt16 ChIP-chip) and BED (Mot1 ORGANIC-Seq), which were reformatted into an Excel sheet before proceeding. Before proceeding, the reformatted data should contain the chromosome number in the first column, the position in the second column, and the signal in the third column. This was then imported into Matlab using the import function. Joseph Muldoon and Kunal Poorey generated a custom Matlab function that can take the localization signals and generate an average profile. This function requires another file that defines the direction of the TSS. Both the function (plot_avg.m) and the file for the TSS direction (TSS_dir) are located on the Auble lab Collab site under Jason's folder. This function will calculate an average localization profile for whatever factor you chose. You can also subdivide this into

different categories based on expression defects, transcription rates, or whatever else you have.

After importing the data into Matlab, you can call the function. We plotted the data +/- 1000 bp from the TSS, which is specified by putting 1000 in for “a”. The window in which the data was smoothed was 20 bp and was defined by “b”.

plot_avg(name of file, TSS_dir, a, b)

The function generated 101 numbers specifying the average signal in each window, which were then saved into a variable (i.e. “spt16_graph”). This data could then be plotted by calling the “plot_loc” function (located on Collab too). Be sure to change in the function which variables are being plotted. The labels, axes, line color, and line width can also be changed. This function generates the final average localization plot.

Nucleosome position shift analysis

This analysis required the DANPOS data (Chen et al. 2013). First the nucleosomes were categorized based on their distance to the TSS. The canonical designations are “-2, -1, +1, +2” etc. depending on whether the nucleosome is upstream or downstream of the start site. We determined that for our data, the +1 nucleosome was -50 to +100 bp from the TSS. Nucleosomes were then binned every 150 bps upstream and downstream. The 150 bp region immediately upstream of the +1 nucleosome was denoted as the nucleosome depleted region (NDR).

After the nucleosomes were grouped, the distance from the nucleosome dyad to the TSS was determined for the WT and mutant cells. The position shift relative to the

TSS was then determined by subtracting the dyad location in the mutant cells from the dyad location in the WT cells. These position shifts can either be positive or negative. Upon determining the position shifts for the nucleosomes, a separate Excel sheet was made showing the shifts for each nucleosome position in the mutant cells. This data was then imported into Matlab. These data were presented in a box and whisker plot, which was generated using:

```
boxplot(nameoffile)
```

Notches were added to the median of the boxes to make them more visible and outliers were removed by ‘Symbol’, k. boxplot(nameoffile, ‘Notch’, on, ‘Symbol’, k)

References

Adamkewicz, J.I., Hansen, K.E., Prud'homme, W.A., Davis, J.L, and Thorner, J. (2001). High affinity interaction of yeast transcriptional regulator, Mot1, with TATA box-binding protein (TBP). The Journal of Biological Chemistry *276*, 11883-11894.

Adamkewicz, J.I., Mueller, C.G.F., Hansen, K.E., Prud'homme, W.A., and Thorner, J. (2000). Purification and enzymatic properties of Mot1 ATPase, a regulator of basal transcription in the yeast Saccharomyces cerevisiae. The Journal of Biological Chemistry *275*, 21158-21168.

Adkins, M.W. and Tyler, J.K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly in the promoter regions. Molecular Cell *3*, 405-416.

Ahn, S.H., Kim, M., Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Molecular Cell *16*, 67-76.

Anderson, J.D. and Widom, J. (2001). Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. Molecular and Cellular Biology *21*, 3830-3839.

Andrade, M.A., Ponting, C.P., Gibson, T.J., and Bork, P. (2000). Homology-based method for identification of protein repeats using statistical significance estimates. Journal of Molecular Biology *298*, 521-537.

Andrau, J.C., Van Oevelen, C.J., Van Teeffelen, H.A., Weil, P.A., Holstege F.C., and Timmers, H.T. (2002). Mot1p is essential for TBP recruitment to selected promoters during in vivo gene activation. The EMBO Journal *1*; 5173-5183.

Annunziato, A.T. (2005). Split decision: what happens to nucleosomes during DNA replication? Journal of Biological Chemistry *280*, 12065-12068.

Arnett, D.R., Jennings, J.L., Tabb, D.L., Link, A.J. and Weil, P.A. (2008). A proteomics analysis of yeast Mot1p protein-protein associations: insights into mechanism. Molecular and Cellular Proteomics *7*, 2090-2106.

Auble, D.T. (2009). The dynamic personality of TATA-binding protein. Trends in Biochemical Sciences *34*, 49-52.

Auble, D.T. and Hahn, S. (1993). An ATP-dependent inhibitor of TBP binding to DNA. Genes and Development *7*, 844-856.

Auble, D.T. and Steggerda, S.M. (1999). Testing for DNA tracking by MOT1, a SNF2/SWI2 protein family member. Molecular and Cellular Biology *19*, 412-423.

Auble, D.T., Hansen, K.E., Mueller, C.G.F., Lane, W.S., Thorner, J., and Hahn, S. (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. Genes and Development *8*, 1920-1934.

Auble, D.T., Wang, D., Post, K.W., and Hahn, S. (1997). Molecular analysis of the SNF2/SWI2 protein family member Mot1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. Molecular and Cellular Biology *17*, 4842-4851.

Baek, H.J., Kang, Y.K., and Roeder, R.G. (2006). Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. Journal of Biological Chemistry *281*, 15172-15181.

Bagby, S., Mal, T.K., Liu, D., Raddatz, E., Nakatani, Y., and Ikura, M. (2000). TFIIA-TAF regulatory interplay: NMR evidence for overlapping binding sites on TBP. FEBS Letters *468,* 149-154.

Bangur, C.S., Pardee, T.S., Ponticelli, A.S. (1997). Mutational analysis of the D1/E1 core helices and the conserved N-terminal region of yeast transcription factor IIB (TFIIB): identification of an N-terminal mutant that stabilizes TATA-binding protein-TFIIB-DNA complexes. Molecular and Cellular Biology *17*, 6784-6793.

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

Belotserkovskaya, R. and Reinberg, D. (2004). Facts about FACT and transcript elongation through chromatin. Current Opinion in Genetics and Development *14*, 139-146.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science *301*, 1090-1093.

Berroteran, R.W., Ware, D.E., and Hampsey, M. (1994). The sua8 suppressors of Saccharomyces cerevisiae encode replacements of conserved residues within the largest subunit of RNA polymerase II and affect transcription start site selection similarly to sua7 (TFIIB) mutations. Molecular and Cellular Biology *14*, 226-237.

Bhaumik, S.R. and Green, M.R. (2002). Differential requirement of SAGA components for recruitment of TATA-box binding protein to promoters in vivo. Molecular and Cellular Biology *22*, 7365-7371.

Biswas, D., Yu, Y., Mitra, D. and Stillman, D.J. (2006). Genetic interactions between Nhp6 and Gcn5 with Mot1 and the Ccr4-Not complex that regulate binding of TATA-binding protein in Saccharomyces cerevisiae. Genetics *172*, 837-849.

Biswas, D., Yu, Y., Prall, M., Formosa, T. and Stillman, D.J. (2005). The yeast FACT complex has a role in transcriptional initiation. Molecular and Cellular Biology *25*, 5812-5822.

Blair, R.H., Goodrich, J.A. and Kugel, J.F. (2012). Single-molecule fluorescence resonance energy transfer shows uniformity in TATA binding protein-induced DNA bending and heterogeneity in bending kinetics. Biochemistry *51*, 7444-7455.

Bonnet, J., Wang, C.Y., Baptista, T., Vincent, S.D., Hsiao, W.C., Stierle, M., Kao, C.F., Tora, L., and Devys, D. (2014). The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. Genes and Development *28*, 1999-2012.

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

Brewster, N.K., Johnston, G.C., Singer, R.A. (1998). Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression. Journal of Biological Chemistry *273*, 21972-21979.

Buratowksi, S. (2009). Progression through the RNA polymerase II CTD cycle. Molecular Cell *36*, 541-546.

Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell *56*, 549-561.

Buratowski, S., Hahn, S., Sharp, P.A. and Guarente, L. (1988). Function of a yeast TATA element-binding protein in a mammalian transcription system. Nature *334*, 37-42.

Burke, T.W. and Kadonaga, J.T. (1997). The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes and Development *11*, 3020-3031.

Butler, J.E. and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes and Development *16*, 2583-2592.

Butryn, A., Schuller, J.M., Stoehr, G., Runge-Wollmann, P., Förster, F., Auble, D.T,. and Hopfner, K-P. (2015). Structural basis for recognition and remodeling of the TBP:DNA:NC2 complex by Mot1. eLife *4*, e07432.

Cech, T.R. and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. Cell *157*, 77-94.

Chasman, D.I., Flaherty, K.M., Sharp, P.A. and Kornberg, R.D. (1993). Crystal structure of yeast TATA-binding protein and model for interaction with DNA. PNAS *90*, 8174-8178.

Chen, H.T. and Hahn, S. (2003). Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. Molecular Cell *12*, 437-447.

Chen, K., Xi, Y., Pan, X., Li, Z,. Kaestner, K., Tyler, J., Dent, S., He, X. and Li, W. (2013). DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Research *23*, 341-351.

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

Chiang, C.M. and Roeder, R.G. (1995). Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science *267*, 531-536.

Chicca, J.J., Auble, D.T., and Pugh, B.F. (1998). Cloning and biochemical characterization of TAF-172, a human homolog of yeast Mot1. Molecular and Cellular Biology *18*, 1701-1710.

Chung, W.H., Craighead, J.L., Chang, W.H., Ezeokonkwo, C., Bareket-Samish, A., Kornberg, R.D., and Asturias, F.J. (2003). RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. Molecular Cell *12*, 1003-1013.

Clapier, C.R. and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annual Review of Biochemistry *78*, 273-304.

Clark-Adams, C.D. and Winston, F. (1987). The SPT6 gene is essential for growth and is required for delta-mediated transcription in Saccharomyces cerevisiae. Molecular and Cellular Biology *7*, 679-686.

Coin, F. and Egly, J.M. (1998). Ten years of TFIIH. Cold Spring Harbor Symposium of Quantitative Biology *63*, 105-110.

Collart, M.A. (1996). The NOT, SPT3, and MOT1 genes functionally interact to regulate transcription at core promoters. Molecular and Cellular Biology *16*, 6668-6676.

Comai, L., Tanese, N., and Tijan, R. (1992). The TATA-binding protein and associated factors are integral components of the RNA Polymerase I transcription factor, SL1. Cell *68*, 965-976.

Cormack, B.P. and Struhl, K. (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. Cell *69*, 685-696.

Creton, S., Svejstrup, J.Q., and Collart, M.A. (2002). The NC2α and β subunits play different roles in vivo. Genes and Development *16*, 3265-3276.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

Crowley, T.E., Hoey, T., Liu, J.-K., Jan, Y.N., and Tijan, R. (1993). A new factor related to TATA-binding protein has highly restricted expression patterns in Drosophila. Nature *361*, 557-561.

D'Arcy, S., Martin, K.W., Panchenko, T., Chen, X., Bergeron, S., Stargell, L.A., Black, B.E., and Luger, K. (2013). Chaperone Nap1 shields histone surfaces used in a nucleosome and can put H2A-H2B in an unconventional tetrameric form. Molecular Cell *51*, 662-677.

Darst, R.P., Dasgupta, A., Zhu, C., Hsu, J-Y., Vroom, A., Muldrow, T., and Auble, D.T. (2003). Mot1 regulates the DNA binding activity of free TATA-binding protein in an ATP-dependent manner. Journal of Biological Chemistry *278*, 13216-13226.

Darst, R.P., Wang, D,. and Auble, D.T. (2001). MOT1-catalyzed TBP-DNA disruption: uncoupling DNA conformational change and role of upstream DNA. The EMBO Journal *20*, 2028-2040.

Dasgupta, A., Darst, R.P., Martin, K.J., Afshari, C.A., and D.T. Auble (2002). Mot1 activates and represses transcription by direct, ATPase-dependent mechanisms. PNAS 99: 2666-2671.

Dasgupta, A., Juedes, S.A., Sprouse, R.O., Auble, D.T. (2005). Mot1-mediated control of transcription complex assembly and activity. The EMBO Journal *24*, 1717-1729.

Davis, J.L, Kunisawa, R., and J. Thorner (1992). A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast Saccharomyces cerevisiae. Mol. Cell. Biol. 12: 1879-1892.

Davison, B.L., Egly, J.M., Mulvihill, E.R. and Chambon, P. (1983). Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature *301*, 680-686.

De Koning, L., Corpet, A., Haber, J.E., and Almouzni, G. (2007). Histone chaperones: an escort network regulating histone network. Nature Structural and Molecular Biology *14*, 997-1007.

Dennehey B. and Tyler, J.K. (2013). Chromatin assembly and disassembly during DNA repair and DNA replication. Chapter 2, Fundamentals of Chromatin.

Diebold, M.L., Koch, M., Loeliger, E,. Cura, V., Winston, F., Cavarelli, J., and Romier, C. (2010). The structure of an Iws1/Sp6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. The EMBO Journal *29*, 3979-3991.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Research *11*, 1475-1489.

Douziech, M., Coin, F., Chipoulet, J.M., Arai, Y., Ohkuma, Y., Egly, J.M., and Coulombe, B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. Molecular and Cellular Biology *20*, 8168-8177.

Duina, A.A. (2011). Histone Chaperones Spt6 and FACT: Similarities and Differences in Modes of Action at Transcribed Genes. Genetics Research International *2011*, 625210.

Duina, A.A., Rufiange, A., Bracey, J., Hall, J., Nourani, A., and Winston, F. (2007). Evidence that the localization of the elongation factor Spt16 across transcribed genes is dependent upon histone H3 integrity in Saccharomyces cerevisiae. Genetics *177*, 101-112.

Durant, M. and Pugh, B.F. (2006). Genome-wide relationships between TAF1 and histone acetyltransferases in Saccharomyces cerevisiae. Molecular and Cellular Biology *26*, 2791-2802.

Dürr, H., Körner, C., Müller, M., Hickmann V., and Hopfner, K.P. (2005). X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. Cell *121*, 363-373.

Dynlacht, B.D., Hoey, T. and Tijan, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell *66*, 563-576.

Eberharter, A. and Becker, P.B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. EMBO Reports *3*, 224-229.

Eisen, J.A., Sweder, K.S., and Hanawalt, P.C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Research *23*, 2715-2723.

Eisenmann, D.M., Dollard, C. and Winston, F. (1989). SPT15, the gene encoding the yeast TATA binding factor TFIID, is required for normal transcription initiation in vivo. Cell *58*, 1183-1191.

English, C.E., Adkins, M.W., Carson, J.J., Churchill, M.E.A., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of Asf1. Cell *127*, 495-508.

Fisher, A.J., Smith, C.A., Thoden, J.B., Smith, R., Sutoh, K., Holden, H.M., Rayment, I. (1995). X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4-. Biochemistry *34*, 8960-8972.

Flaus, A., Martin, D.M., Barton, G.J. and Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic Acids Research *34*, 2887-295.

Flores, O., Lu, H. and Reinberg, D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. Journal of Biological Chemistry *267*, 2786-2793.

Flores, O., Maldonado, E. and Reinberg, D. (1989). Factors involved in specific transcription by mammalian RNA polymerase II. Journal of Biological Chemistry *264*, 8913-8921.

Foltman, M., Evrin, C., De Piccoli, G., Jones, R.C., Edmondson, R.D., Katou, Y., Nakato, R., Shirahige, K. and Labib, K. (2013). Eukaryotic replisome components cooperate to process histones during chromosome replication. Cell Reports *3*, 892-904.

Forget, D. Langelier, M.F., Thérien, C., Trinh, V. and Coulombe, B. (2004). Photo-cross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. Molecular and Cellular Biology *24*, 1122-1131.

Formosa, T. (2012) The role of FACT in making and breaking nucleosomes. Biochimica et Biophysica Acta *1819*, 247-55.

Formosa, T., Eriksson, P., Wittmeyer, J., Ginn, J., Yu, Y. and Stillman, D. J. (2001). Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. The EMBO Journal *20*, 3506–3517.

Formosa, T., Ruone, S., Adams, M. D., Olsen, A. E., Eriksson, P., Yu, Y., Rhoades, A. R., Kaufman, P. D. and Stillman, D. J. (2002) Defects in SPT16 or POB3 (yFACT) in Saccharomyces cerevisiae cause dependence on the Hir/Hpc pathway: polymerase passage may degrade chromatin structure. Genetics *162*, 1557–1571.

Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R.G. and Chua, N.H. (1990). Arabidopsis thaliana contains two genes for TFIID. Nature *346*, 390-394.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., Edelmann, A., Heurtier, M.A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A.M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B. and Superti-Furga, G. (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Geiger, J.H., Hahn, S., Lee, S. and Sigler, P.B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. Science *272*, 830-836.

Geisberg, J.V. and Struhl, K. (2004). Cellular stress alters the transcriptional properties of promoter-bound Mot1-TBP complexes. Molecular Cell *14*, 479-489.

Geisberg, J.V., Holstege, F.C., Young, R.A. and Struhl, K. (2001). Yeast NC2 associates with the RNA polymerase II preinitiation complex and selectively affects transcription in vivo. Molecular and Cellular Biology *21*, 2736-2742.

Geisberg, J.V., Mogtaderi, Z., Kuras, L. and Struhl, K. (2002). Mot1 associates with transcriptionally active promoters and inhibits association of NC2 in Saccharomyces cerevisiae. Molecular and Cellular Biology *22*, 8122-8134.

Ghazy, M.A., Brodie, S.A., Ammerman, M.L., Ziegler, L.M. and Ponticelli, A.S. (2004). Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. Molecular and Cellular Biology *24*, 10975-10985.

Goldman-Levi, R., Miller, C., Bogoch, J. and Zak, N.B. (1996). Expanding the Mot1 subfamily: 89B helicase encodes a new Drosophila melanogaster SNF2-related protein which binds to multiple sites on polytene chromosomes. Nucleic Acids Research *24*, 3121-3128.

Goodrich, J.A., Hoey, T., Thut, C.J., Admon, A. and Tijan, R. (1993). Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell *75*, 519-530.

Goppelt, A. and Meiseterernst, M. (1996). Characterization of the basal inhibitor of class II transcription NC2 from Saccharomyces cerevisiae. Nucleic Acids Research *24*, 4450-4455.

Goppelt, A., Stelzer, G., Lottspeich, F. and Meisterernst, M. (1995). A mechanism for repression of class II gene transcription through specific binding of NC2 to TBP-promoter complexes via heterodimeric histone fold domains. The EMBO Journal *15*, 3105-3116.

Goppelt, A., Stelzer, G., Lottspeich, F., and Meisterernst, M. (1996). A mechanism for repression of class II gene transcription through specific binding of NC2 to TBP-promoter complexes via heterodimeric histone fold domains. The EMBO Journal *15*, 3105-3116.

Gorbalenya, A.E, Koonin, E.V., Donchenko, A.P. and Blinov, V.M. (1988). A conserved NTP-motif in putative helicases. Nature *333*, 22.

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

Grant, P.A., Schieltz, D., Pray-Grant, M.G., Yates, J.R. and Workman, J.L. (1998). The ATM-related cofactor Tra1 is a component of the purified SAGA complex. Molecular Cell *2*, 863-867.

Guarente, L. (1987). Regulatory proteins in yeast. Annual Review of Genetics *21*, 425-452.

Guarente, L. (1993). Synthetic enhancement in gene interaction: a genetic tool come of age. Trends in Genetics *9*, 362-366.

Gumbs, O.H., Campbell, A.M. and Weil, P.A. (2003). High-affinity DNA binding by a Mot1p-TBP complex: implications for TAF-independent transcription. The EMBO Journal *22*, 3131-3141.

Gupta, S., Cheng, H., Mollah, A.K., Jamison, E., Morris, S., Chance, M.R., Khrapunov, S. and Brenowitz, M. (2007). DNA and protein footprinting analysis of the modulation of DNA binding by the N-terminal domain of the Saccharomyces cerevisiae TATA binding protein. Biochemistry *46*, 9886-9898.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nature Structural and Molecular Biology *11*, 394-403.

Hahn, S., Buratowski, S., Sharp, P.A. and Guarente, L. (1989). Isolation of the gene encoding the yeast TATA binding protein TFIID: A gene identical to the SPT15 suppressor of Ty element insertions. Cell *58*, 1173-1181.

Hahn, S., Buratowski, S., Sharp, P.A. and Guarente, L. (1989). Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. PNAS *86*, 5718-5722.

Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiology and Molecular Biology Reviews *62*, 465-503.

Han, Y., Luo, J., Ranish, J. and Hahn, S. (2014). Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex. The EMBO Journal *33*, 2534-2546.

Han, J., Li, Q., McCullough, L., Kettelkamp, C., Formosa, T. and Zhang, Z. (2010) Ubiquitylation of FACT by the cullin-E3 ligase Rtt101 connects FACT to DNA replication. Genes and Development *24*, 1485-1490.

Hartzog, G. A., Wada, T., Handa, H. and Winston, F. (1998). Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. Genes and Development *12*, 357-369.

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

Heo, K., Kim, H., Choi, S.H., Choi, J., Kim, K., Gu, J., Lieber, M.R., Yang, A.S. and An, W. (2008). FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. Molecular Cell *30*, 86-97.

Hoey, T., Weinzierl, R.O.J., Gill, G. Chen, J.-L., Dynlacht, B.D. and Tjian, R. (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell *72*, 247-260.

Horikoshi, M., Bertucciolo, C., Takada, R., Wang, J., Yamamoto, T. and Roeder, R.G. (1992). Transcription factor TFIID induces DNA bending upon binding to the TATA element. PNAS *89*, 1060-1064.

Hsu, J.Y., Juven-Gershon, T., Marr, M.T., 2nd, Wright, K.J., Tjian, R. and Kadonaga, J.T. (2008). TBP, Mot1, and NC2 establish a regulatory circuit that controls DPEdependent versus TATA-dependent transcription. Genes and Development *22*, 2353-2358.

Hsu, J-Y., Juven-Gershon, T., Marr, M.T., Wright, K.J., Tijan, R. and Kadonga, J.T. (2008). TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription. Genes and Development *22*, 2353-2358.

Huang, S., Zhou, H., Katzmann, D., Hochstrasser, M., Atanasova, E. and Zhang, Z. (2005) Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. PNAS *102,* 13410–13415

Huang, J.Y., Chen, W.H., Chang, Y.L., Wang, H.T., Chuang, W.T. and Lee, S.C. (2006). Modulation of nucleosome-binding activity of FACT by poly(ADP-ribosyl)ation. Nucleic Acids Research *34*, 2398-2407.

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

Inostroza, J.A., Mermelstein, F.H., Ha, I., Lane, W.S. and Reinberg, D. (1992). Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. Cell *70*, 477-489.

Ivanovska, I., Jacques, P.É., Rando, O.J., Robert, F. and Winston, F. (2011). Control of chromatin structure by spt6: different consequences in coding and regulatory regions. Molecular and Cellular Biology *31*, 531-541.

Iyer, V. and Struhl, K. (1995). Mechanism of differential utilization of the his3 TR and TC TATA elements. Molecular and Cellular Biology *15*, 7059-7066.

Jackson, R.J., Hellen, C.U. and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nature Reviews Molecular Cell Biology *11*, 113-127.

Jaehning, J.A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? Biochimica et Biophysica Acta *1799*, 379-388.

Jamai, A., Puglisi, A. and Strubin, M. (2009) Histone chaperone spt16 promotes redeposition of the original h3-h4 histones evicted by elongating RNA polymerase. Molecular Cell *35*, 377–383

Jensen, M.M., Christensen, M.S., Bonven, B. and Jensen, T.H. (2008) Requirements for chromatin reassembly during transcriptional downregulation of a heat shock gene in Saccharomyces cerevisiae. FEBS Journal *275*, 2956–2964

Jiang, Y.W. and Stillman, D.J. (1996). Epigenetic effects on yeast transcription caused by mutations in an actin-related protein present in the nucleus. Genes and Development *10*, 604-619.

Jimeno-González, S., Gómez-Herreros, F., Alepuz, P.M. and Chávez, S. (2006). A gene-specific requirement for FACT during transcription is related to the chromatin organization of the transcribed region. Molecular and Cellular Biology *26*, 8710-8721.

Joazeiro, C.A.P., Kassavetis, G.A. and Geiduschek, E.P. (1996). Alternative outcomes in assembly of promoter complexes: the roles of TBP and a flexible linker in placing TFIIIB on tRNA genes. Genes and Development *10*, 725-739.

Kamada, K., Shu, F., Chen, H., Malik, S., Stelzer, G., Roeder, R.G., Meisterernst, M. and Burley, S.K. (2001). Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. Cell *106*, 71-81.

Kamenova, I., Warfield, L. and Hahn, S. (2014). Mutations on the DNA binding surface of TBP discriminate between yeast TATA and TATA-less gene transcription. Molecular and Cellular Biology *34*, 2929-2943.

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

Karnitz, L., Morrison, M. and Young, E.T. (1992). Identification and characterization of three genes that affect expression of ADH2 in Saccharomyces cerevisiae. Genetics *132,* 351-359.

Kemble, D.J., Whitby, F.G., Robinson, H., McCullough, L.L., Formosa, T. and Hill, C.P. (2013). Structure of the Spt16 middle domain reveals functional features of the histone chaperone FACT. Journal of Biological Chemistry *288*, 10188-10194.

Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204-2207.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., Boone, C., Emili, A., Weissman, J.S., Hughes, T.R., Strahl, B.D., Grunstein, M., Greenblatt, J.F., Buratowski, S. and Krogan, N.J. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Kim, Y.J., Björklund, S., Li, Y., Sayre, MH. and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell *77*, 599-608.

Kim, Y., Geiger, J.H., Hahn, S. and Sigler, P.B. (1993). Crystal structure of a yeast TBP/TATA-box complex. Nature *365*, 512-520.

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

Köhler, A., Schneider, M., Cabal, G.G., Nehrbass, U., Hurt, E.. (2008). Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. Nature Cell Biology *10*, 707–715.

Kokubo, T., Swanson, M. J., Nishikawa, J. I., Hinnebusch, A. G. and Nakatani, Y. (1998). The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. Molecular and Cellular Biology *18*, 1003-1012.

Korber, P. and Hörz, W. (2004). In vitro assembly of the characteristic chromatin organization at the yeast PHO5 promoter by a replication-independent extract system. Journal of Biological Chemistry *279*, 35113-35120.

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

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

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S. and Greenblatt, J. (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Molecular and Cellular Biology *23*, 4207-4218.

Kung, J.T.Y., Colognori, D. and Lee, J.T. (2012). Long Noncoding RNAs: Past, Present, and Future. Genetics *193*, 651-669.

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

Lee, D.K., Horikoshi, M. and Roeder, R.D. (1991). Interaction ofTFIID in the minor groove of the TATA element. Cell *67*, 1241-1250.

Lee, D.K., Duan, H.O. and Chang, C. (2000). From androgen receptor to the general transcription factor TFIIH. Identification of cdk activating kinase (CAK) as an androgen receptor NH(2)-terminal associated coactivator. Journal of Biological Chemistry *31*, 9308-9313.

Lemaire, M., Xie, J., Meisterernst, M., Collart, M.A. (2000). The NC2 repressor is dispensable in yeast mutated for the Sin4p component of the holoenzyme and plays roles similar to Mot1p in vivo. Molecular Microbiology *36*, 163-173.

Li, Y., Flanagan, P.M., Tschochner, H. and Kornberg, R.D. (1994). RNA polymerase II initiation factor interactions and transcription start site selection. Science *263*, 805-807.

Bing, L., Carey, M. and Workman, J.L. (2007). The Role of Chromatin during Transcription. Cell *128*, 707-719.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Lickwar, C.R., Rao, B., Shabalin, A.A., Nobel, A.B., Strahl, B.D. and Lieb, J.D. (2009). The Set2/Rpd3S Pathway Suppresses Cryptic Transcription without Regard to Gene Length or Transcription Frequency. PLoS One *4*, e4886.

Lieberman, P.M. and Berk, A.J. (1994). A mechanism for Tafs in transcriptional activation: Activation domain enhancement of TFIID-TFIIA-promoter DNA complex formation. Genes and Development *9*: 995-1006.

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

Liu, F., Putnam, A. and Jankowsky, E. (2008). ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. PNAS *105*, 20209-20214.

Liu, H., Luo, K., Zhou, Z., Mu, Y. and Wan, Y. (2014). Histone chaperone Chz1 facilitates the disfavouring property of Spt16 to H2A.Z-containing genes in Saccharomyces cerevisiae. Biochemical Journal *460*, 387-397.

Liu, H-Y., Badarinarayana, V., Audino, D.C., Rappsilber, J., Mann, M. and Denis, C.L. (1998). The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. The EMBO Journal *17*, 1096-1106.

Lu, H., Zawel, L., Fisher, L., Egly, J.M. and Reinberg D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature *358,* 641–645.

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

Luk, E., Vu, N.D., Patteson, K., Mizuguchi, G., Wu, W.H., Ranjan, A., Backus, J., Sen, S., Lewis, M., Bai, Y. and Wu, C. (2007). Chz1, a nuclear chaperone for histone H2AZ. Molecular Cell *25*, 357-368.

Mackiewicz, P., Kowalczuk, M., Mackiewicz, D., Nowicka, A., Dudkiewicz, M., Laszkiewicz, A., Dudek, M.R. and Cebrat, S. (2002). How many protein-coding genes are there in the Saccharomyces cerevisiae genome? Yeast *19*, 619-629.

Madison, J.M. and Winston, F. (1997). Evidence that Spt3 functionally interacts with Mot1, TFIIA, and TATA-binding protein to confer promoter-specific transcriptional control in Saccharomyces cerevisiae. Molecular and Cellular Biology *17*, 287-295.

Mahapatra, S., Dewari, P.S., Bhardwaj, A. and Bhargava, P. (2011). Yeast H2A.Z, FACT complex and RSC regulate transcription of tRNA gene through differential dynamics of flanking nucleosomes. Nucleic Acids Research *39*, 4023-4034.

Mai, X., Chou, S. and Struhl, K. (2000). Preferential accessibility of the yeast his3 promoter is determined by a general property of the DNA sequence, not by specific elements. Molecular and Cellular Biology *20*,6668–6676.

Malone, E., Clark, C., Chiang, A. and Winston, F. (1991). Mutations in SPT16/CDC68 suppress cis- and trans-acting mutations that affect promoter function in Saccharomyces cerevisiae. Molecular and Cellular Biology *11*, 5710–5717.

Maruta, S., Henry, G.D., Sykes, B.D. and Ikebe, M. (1993). Formation of the stable myosin-ADP-aluminum fluoride and myosin-ADP-beryllium fluoride complexes and their analysis using 19F NMR. Journal of Biological Chemisty *268*, 7093-7100.

Mason, P. B. and Struhl, K. (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.

Matsui, T., Segall, J., Weil, P.A. and Roeder, R.G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. Journal of Biological Chemistry *255*, 11992-11996.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J. and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nature Structural and Molecular Biology *17*, 1272-1278.

McDonald,S.M., Close, D. Xin, H., Formosa, T. and Hill, C.P. (2010). Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. Molecular Cell *40*, 725-735.

Meisterernst, M. and Roeder, R.G. (1991). Family of proteins that interact with TFIID and regulate promoter activity. Cell *67*, 557-567.

Mermelstein, F., Yeung, K., Cao, J., Inostroza, J.A., Erdjument-Bromage, H., Eagelson, K., Landsman, D., Levitt, P., Tempst, P. and Reinberg, D. (1996). Requirement of a corepressor for Dr1-mediated repression of transcription. Genes and Development *10*, 1033-1048.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S. and Wu, C. (2003). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343–348.

Mohibullah, N. and Hahn, S. (2008). Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. Genes and Development *22*, 2994-3006.

Montgomerie, S., Cruz, J.A., Shrivastava, S., Arndt, D., Berjanskii, M. and Wishart, D.S. (2008). PROTEUS2: a web server for comprehensive protein structure prediction and structure-based annotation. Nucleic Acids Research *36*, W202-209.

Mosammaparast, N., Jackson, K.R., Guo, Y., Brame, C.J., Shabanowitz, J., Hunt, D.F. and Pemberton, L.F. (2001). Nuclear import of histone H2A and H2B is mediated by a network of karyopherins. Journal of Cell Biology *153*, 251-262.

Mosammaparast, N., Ewart, C.S. and Pemberton, L.F. (2002). A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. The EMBO Journal *21*, 6527-6538.

Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L. and Washburn, M.P. (2009). Rtr1 is a CTD Phosphatase that Regulates the Transition from Serine 5 to Serine 2 Phosphorylation during the RNA Polymerase II Transcription Cycle. Molecular Cell *34*, 168-178.

Moyle-Heyrman, G., Viswanathan, R., Widom, J. and Auble, D.T. (2012). Two-step mechanism for modifier of transcription 1 (Mot1) enzyme-catalyzed displacement of TATA-binding protein (TBP) from DNA. The Journal of Biological Chemistry *287*, 9002-9012.

Mueller, C.L. and Jaehning, J.A. (2002). Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. Molecular and Cellular Biology.

Mueller, C.L., Porter, S.E., Hoffman, M.G. and Jaehning, J.A. (2004). The Paf1 complex has functions independent of actively transcribing RNA polymerase II. Molecular Cell *14*, 447-456.

Muldrow, T.A., Campbel, A.M., Weil, P.A. and Auble, D.T. (1999). MOT1 can activate basal transcription in vitro by regulating the distribution of TATA binding protein between promoter and nonpromoter sites. Molecular and Cellular Biololgy *19*, 2835-2845.

Munn, A.L. (2000). The yeast endocytic membrane transport system. Microscopy Research and Technique *51*, 547-562.

Munn, A.L. and Riezman, H. (1994). Endocytosis is required for the growth of vacuolar H+-ATPase defective yeast: Identification of six new END genes. The Journal of Cell Biology *127*, 373-386.

Myers, L.C. and Kornberg, R.D. (2000). Mediator of transcriptional regulation. Annual Review of Biochemisty *69*, 729-749.

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

Nielsen, C.B., Cantor, M., Dubchak, I., Gordon, D. and Wang, T. (2010). Visualizing genomes: techniques and challenges. Nature Methods *7*, S5-S15.

Nikolov, D.B. and Burley, S.K. (1997). RNA polymerase II transcription initiation: a structural view. PNAS *94*, 15-22.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G. and Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. Nature *377*, 119-128.

Nikolov, D.B., Hu, S-H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N-H., Roeder, R.G. and Burley, S.K. (1992). Crystal structure of TFIID TATA-box binding protein. Nature *360*, 40-46.

O'Donnell, A.F., Brewster, N.K., Kurniawan, J., Minard, L.V., Johnston, G.C. and Singer, R.A. (2004). Domain organization of the yeast histone chaperone FACT: the conserved N-terminal domain of FACT subunit Spt16 mediates recovery from replication stress. Nucleic Acids Research *32*, 5894-5906.

Ohkuma, Y., Hashimoto, S., Wang, C.K., Horikoshi, M. and Roeder R.G. (1995). Analysisof the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure–function studies of TFIIE-a. Molecular and Cellular Biology *15*, 4856–4866.

Okuda, M., Tanaka, A., Satoh, M., Mizuta, S., Takazawa, M., Ohkuma, Y. and Nishimura, Y. (2008). Structural insight into the TFIIE–TFIIH interaction: TFIIE and p53 share the binding region on TFIIH. The EMBO Journal *27*, 1161-1171.

Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S. and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. Cell *92*, 105-116.

Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M. and Reinberg D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature *400*, 284-288.

Ozer, J., Mitsouras, K., Zerby, D., Carey, M. and Lieberman, P.M. (1998). Transcription factor IIA derepresses TATA-binding protein (TBP)-associated factor inhibition of TBP-DNA binding. Journal of Biological Chemistry *273*, 14293-14300.

Papamichos-Chronakis, M., Watanabe, S., Rando, O.J. and Peterson, C.L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin remodeling enzyme is essential for genome integrity. Cell *144*, 200-213.

Pardee, T.S., Bangur, C.S., Ponticelli, A.S. (1998). The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. Journal of Biological Chemistry *273*, 17859–17864.

Pazin, M.J. and Kadonaga, J.T. (1997). What's up and down with histone deacetylation and transcription? Cell *89*, 325-328.

Pereira, L.A., van der Knaap, J.A., van den Boom, V., van den Heuvel, F.A. and Timmers, H.T. (2001). TAF(II)170 interacts with the concave surface of TATA-binding protein to inhibit its DNA binding activity. Molecular and Cellular Biology *21*, 7523–7534.

Piatti, S., Tazzi, R., Pizzagalli, A., Plevani, P. and Lucchini, G. (1992). Control of DNA synthesis genes in budding yeast: Involvement of the transcriptional modulator MOT1 in the expression of the DNA polymerase α gene. Chromosoma *102*, S107-S113.

Pinto, I., Ware, D.E. and Hampsey, M. (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.

Pinto, I., Wu, W.H., Na, J.G. and Hampsey, M. (1994). Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. Journal of Biological Chemistry *269*, 30569-30573.

Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J. and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. Nature *533*, 353-358.

Pokholok, D.K., Hannett, N.M. and Young, R.A. (2002). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. Molecular Cell *9*, 799-809.

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

Poon, D., Campbell, A.M., Bai, Y. and Weil, P.A. (1994). Yeast Taf170 is encoded by MOT1 and exists in a TATA box-binding protein (TBP)-TBP-associated factor complex distinct from transcription factor IID. Journal of Biological Chemistry *269*, 23135-23140.

Poorey, K., Sprouse, R.O., Wells, M.N., Viswanathan, R., Bekiranov, S. and Auble, D.T. (2010). RNA synthesis precision is regulated by preinitiation complex turnover. Genome Research *20*, 1679-1688.

Poorey, K., Viswanathan, R., Carver, M.N., Karpova, T.S., Cirimotich, S.M., McNally, J.G., Bekiranov, S. and Auble, D.T. (2013). Measuring chromatin interaction dynamics on the second time scale at single-copy genes. Science *342*, 369-372.

Pugh, B.F. (2000). Control of gene expression through regulation of the TATA-binding protein. Gene *255*, 1-14.

Pugh, B.F. (2010). A preoccupied position on nucleosomes. Nature Structural and Molecular Biology *17*, 923.

Pugh, B.F. and Tijan, R. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. Genes and Development *5*, 1935-1945.

Quinlan, A.R. and Hall I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Rabenstein, M.D., Zhou, S., Lis, J.T. and Tjian, R. (1999). TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. PNAS *96*, 4791-4796.

Ransom, M., Dennehey, B.K. and Tyler, J.K. (2010). Chaperoning histones during DNA replication and repair. Cell *140*, 183–195.

Reddy, P. and Hahn, S. (1991). Dominant negative mutations in yeast TFIID define a bipartite DNA-binding region. Cell *65*, 349-357.

Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K. and Bonner, W. (2002). Histone H2A variants H2AX and H2AZ. Current Opinion in Genetics and Development.

Reinberg, D., Horikoshi, M. and Roeder, R.G. (1987). Factors involved in specific transcription in mammalian RNA polymerase II: Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. Journal of Biological Chemistry *262*, 3322-3330.

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

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

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P. (2011). Integrative genomics viewer. Nature Biotechnology *24*, 24-26.

Roh, T.Y., Ngau, W.C., Cui, K., Landsman, D. and Zhao, K. (2004). High-resolution genome-wide mapping of histone modifications. Nature Biotechnology *22*, 1013–1016.

Rondón, A.G., Gallardo, M., García-Rubio, M. and Aguilera, A. (2004). Molecular evidence indicating that the yeast PAF complex is required for transcription elongation. EMBO Reports *5*, 47-53.

Rosaleny, L.E., Ruiz-García, A.B., García-Martínez, J., Pérez-Ortín,1J.E. and Tordera, V. (2007). The Sas3p and Gcn5p histone acetyltransferases are recruited to similar genes. Genome Biology *8*, R119.

Roth, S.Y., Denu, J.M. and Allis, C.D. (2001). Histone acetyltransferases. Annual Review of Biochemistry *70*, 81-120.

Rowley, A., Singer, R.A. and Johnston, G.C. (1991). CDC68, a yeast gene that affects regulation of cell proliferation and transcription, encodes a protein with a highly acidic carboxyl terminus. Molecular and Cellular Biology *11*, 5718-5726.

Rozenblatt-Rosen, O., Hughes, C.M., Nannepaga, S.J., Shanmugam, K.S., Copeland, T.D., Guszczynski, T., Resau, J.H., Meyerson, M. (2005). The parafibromin tumor suppressor protein is part of a human Paf1 complex. Molecular and Cellular Biology *25*, 612-620.

Ruone, S., Rhoades, A.R. and Formosa, T. (2003). Multiple Nhp6 molecules are required to recruit Spt16-Pob3 to form yFACT complexes and to reorganize nucleosomes. Journal of Biological Chemistry *278*, 45288-45295.

Ryan, D.P. and Owen-Hughes, T. (2011). Snf2-family proteins: chromatin remodelers for any occasion. Current Opinion in Chemical Biology *15*, 649-656.

Saunders, A., Werner, J., Andrulis, E.D., Nakayama, T., Hirose, S., Reinberg, D. and List, J.T. (2003). Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. Science *301*, 1094-1096.

Sawadogo, M. and Roeder, R.G. (1985). Factors involved in specific transcription by human RNA polymerase II: Analysis by a rapid and quantitative in vitro assay. PNAS *82*, 4394-4398.

Schluesche, P., Stelzer, G., Piaia, E., Lamb, D.C. and Meisterernst, M. (2007). NC2 mobilizes TBP on core promoter TATA boxes. Nature Structural and Molecular Biology *14*, 1196-1201.

Shi, X., Finkelstein, A., Wolf, A.J., Wade, P.A., Burton, Z.F. and Jaehning, J.A. (1996). Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription. Molecular and Cellular Biology *16*, 669–676.

Shi, X., Chang, M., Wolf, A.J., Chang, C.H., Frazer-Abel, A.A., Wade, P.A., Burton, Z.F. and Jaehning, J.A. (1997). Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme. Molecular and Cellular Biology *17*, 1160–1169.

Shimojima, T., Okada, M., Nakayama, T., Ueda, H., Okawa, K., Iwamatsu, A., Handa, H. and Hirose, S. (2003). Drosophila FACT contributes to Hox gene expression through

physical and functional interactions with GAGA factor. Genes and Development *17*, 1605-1616.

Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M. and Iyer, V.R. (2008). Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. PLoS Biology *6*, e65.

Sikorski, R.S. and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Singleton, M.R., Dillingham, M.S. and Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. Annual Review in Biochemistry *76*, 23–50.

Smith-Kinnaman, W.R., Berna, M.J., Hunter, G.O., True, J.D., Hsu, P., Cabello, G.I., Fox, M.J., Varani, G. and Mosley, A.L. (2014). The interactome of the atypical phosphatase Rtr1 in Saccharomyces cerevisiae. Molecular Biosystems *10*, 1730-1741.

Sprouse, R.O., Brenowitz, M. and Auble, D.T. (2006). Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. The EMBO Journal *25*, 1492-1504.

Sprouse, R.O., Karpova, T.S., Mueller, F., Dasgupta, A., McNally, J.G. and Auble, D.T. (2008a). Regulation of TATA-binding protein dynamic in living yeast cells. PNAS *105*, 13304-13308.

Sprouse, R.O., Shcherbakova, I., Cheng, H., Jamison, E., Brenowitz, M. and Auble, D.T. (2008b). Function and structural organization of Mot1 bound to a natural target promoter. Journal of Biological Chemistry *283*, 24935-24948.

Sprouse, R.O., Wells, M.N. and Auble, D.T. (2008c). TATA-binding protein variants that bypass the requirement for Mot1 in vivo. Journal of Biological Chemistry *284*, 4525-4535.

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

Starr, D.B. and Hawley, D.K. (1991). TFIID binds in the minor groove of the TATA box. Cell *67*, 1231-1240.

Sterner, D.E., Grant, P.A., Roberts, S.M., Duggan, L.J., Belotserkovskaya, R., Pacella, L.A., Winston, F., Workman, J.L. and Berger, S.L.. (1999). Functional organization of the yeast SAGA complex: Distinct components involved in structural integrity,

nucleosome acetylation, and TATA-binding protein interaction. Molecuar and Cellular Biology *19,*86–98.

Stillman, B. (1996). Cell cycle control of DNA replication. Science *274*, 1659-1664.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun,Z.-W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F. and Allis, C.D. (2002). Set2 Is a Nucleosomal Histone H3-Selective Methyltransferase That Mediates Transcriptional Repression. Molecular and Cellular Biology *22*, 1298-1306.

Struhl, K. (1985). Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. PNAS 82, 8419-8423.

Struhl, K. (1987). Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell *49*, 295-297.

Struhl, K. and Segal, E. (2012). Determinants of nucleosome positioning. Nature Structural and Molecular Biology *20*, 267-273.

Stuwe, T., Hothorn, M., Lejeune, E., Rybin, V., Bortfeld, M., Scheffzek, K. and Ladurner, A. G. (2008). The FACT Spt16 “peptidase” domain is a histone H3-H4 binding module. PNAS *105*, 8884–8889.

Sun, Z.W. and Hampsey, M. (1995). Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in Saccharomyces cerevisiae. PNAS *92,* 3127-3131.

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

Takada, R., Nakatani, Y., Hoffmann, A., Kokubo, T., Hasegawa, S., Roeder, R.G. and Horikoshi, M. (1992). Identification of human TFIID components and direct interaction between a 250-kDa polypeptide and the TATA box-binding protein (TFIIDᴛ). PNAS *89*, 11809-11813.

Takagi, Y., Komori, H., Chang, W.H., Hudmon, A., Erdjument-Bromage, H., Tempst, P. and Kornberg, R.D. (2003). Revised subunit structure of yeast transcription factor IIH (TFIIH) and reconciliation with human TFIIH. Journal of Biological Chemistry *278*, 43897-43900.

Tan, S., Hunziker, Y., Sargent, D.F. and Richmond, T.J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. Nature *381*, 127-134.

Teytelman, L., Thurtle, D.M., Rine, J. and van Oudenaarden, A. (2013). Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. PNAS *110*, 18602-18607.

Thoma, N.H., Czyzewski, B.K., Alexeev, A.A., Mazin, A.V., Kowalczykowski, S.C. and Pavletich, N.P. (2005). Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. Nature Structural and Molecular Biology *12*, 350-356.

Thorvaldsdóttir, H., Robinson, J.T. and Mesirov, J. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics *14*, 178-192.

Timmers, H.T.M. and Sharp, P.A. (1991). The mammalian TFIID protein is present in two functionally distinct complexes. Genes and Development *5*, 1946-1956.

Timmers, H.T.M., Meyers, R.M. and Sharp, P.A. (1992). Composition of transcription factor B-TFIID. PNAS *89*, 8140-8144.

Topalidou, I., Papamichos-Chronakis, M., Thireos, G. and Tzamarias, D. (2004). Spt3 and Mot1 cooperate in nucleosome remodeling independently of TBP recruitment. The EMBO Journal *23*, 1943-1948.

Travers, A.A. (2004). The structural basis of DNA flexibility. The Royal Society *362*, 1423-1438.

True, J.D., Muldoon, J.J., Carver, M.N., Poorey, K., Shetty, S.J., Bekiranov, S. and Auble, D.T. (2016). The Mot1 ATPase and Spt16 Histone Chaperone Co-Regulate Transcription through Preinitiation Complex Assembly and Nucleosome Organization. Journal of Biological Chemistry DOI: 10.1074/jbc.M116.735134.

Tsui, K., Durbic, T., Gebbia, M. and Nislow, C. (2012). Genomic approaches for determining nucleosome occupancy in yeast. Methods in Molecular Biology *833*, 389-411.

Tsunaka, Y., Fujiwara, Y., Oyama, T., Hirose, S. and Morikawa, K. (2016). Integrated molecular mechanism directing nucleosome reorganization by human FACT. Genes and Development *30*, 673–686.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T. and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature *402*, 555-560.

Udugama, M., Sabri, A. and Bartholomew, B. (2011). The INO80 ATP-Dependent Chromatin Remodeling Complex Is a Nucleosome Spacing Factor. Molecular and Cellular Biology *31*, 662-673.

van Bakel, H., Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R. and Nislow, C. (2013). A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. PLoS Genetics *9*, e1003479.

Van der Knaap, J.A., Borst, J.W., Van der Vliet, P.C., Gentz, R. and Timmers, H.T.M. (1997). Cloning of the cDNA for the TATA-binding protein-associated factorII170 subunit of transcription factor B-TFIID reveals homology to global transcription regulators in yeast and Drosophila. PNAS *94*, 11827-11832.

Van Werven, F.J., Van Bakel, H., Van Teeffelen, H.A.A.M., Altelaar, A.F.M., Koerkamp, M.G., Heck, A.J.R., Holstege, F.C.P. and Timmers, H.T.M. (2008). Cooperative action of NC2 and Mot1p to regulate TATA-binding protein function across the genome. Genes and Development *22*, 2359-2369.

Van Werven, F.J., Van Teeffelen, H.A.A.M., Holstege, F.C.P. and Timmers, H.T.M. (2009). Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. Nature Structural and Molecular Biology *16*, 1043-1048.

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

Vannini, A. and Cramer, P. (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Molecular Cell *45*, 439-446.

Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S. and Wigley, D.B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. Cell *97*, 75-84.

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

Venters, B.J., Wachi, S., Mavrich, T.N., Andersen, B.E., Jena, P., Sinnamon, A.J., Jain, P., Rolleri, N.S., Jiang, C., Hemeryck-Walsh, C. and Pugh, B.F. (2011). A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. Molecular Cell *41*, 480-492.

Verrijizer, C.P., Yokomori, K., Chen, J.-L. and Tijan, R. (1994). Drosophila TAFII150: Similarity to yeast gene TSM-1 and specific binding to core promoter DNA. Science *264*, 933-941.

Viswanathan, R., True, J.D. and Auble, D.T. (2016). Molecular Mechanism of Mot1, a TATA-Binding Protein (TBP)-DNA Dissociating Enzyme. Journal of Biological Chemistry DOI: 10.1074/jbc.M116.730366.

Viswanathan, R. (2013). Mechanism and Regulation of Transcription Factor Dynamics.

Viswanathan, R. and Auble, D.T. (2011). One small step for Mot1; one giant leap for other Swi2/Snf2 enzymes? Biochimica et Biophysica Acta *1809*, 488-496.

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

Voronkova, V., Kacherovsky, N., Tachibana, C., Yu, D. and Young, E.T. (2006). Snf1-dependent and Snf1-independent pathways of constitutive ADH2 expression in Saccharomyces cerevisiae. Genetics 1*72*, 2123-2138.

Wade, P.A. and Jaehning, J.A. (1996). Transcriptional corepression in vitro: A Mot1p-associated form of TATA-binding protein is required for repression by Leu3p. Molecular and Cellular Biology *16*, 1641-1648.

Wade, P.A., Werel, W., Fentzke, R.C., Thompson, N.E., Leykam, J.F., Burgess, R.R., Jaehning, J.A. and Burton Z.F. (1996). A novel collection of accessory factors associated with yeast RNA polymerase II. Protein Expression and Purification *8*, 85-90.

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

Weideman, C.A., Netter, R.C., Benjamin, L.R., McAllister, J.J., Schmiedekamp, L.A., Coleman R.A. and Pugh, B.F. (1997). Dynamic interplay of TFIIA, TBP and TATA DNA. Journal of Molecular Biology *27*, 61-75.

Wells, M.N. (2012). The Regulation and Analysis of Transcription Preinitiation Complex Dynamics.

White, R.J., Khoo, B.C., Inostroza, J.A., Reinberg, D., Jackson,S.P. (1994). Differential regulation of RNA polymerases I, II, and III by the TBP-binding repressor Dr1. Science *266*, 448-450.

Willy, P.J., Kobayashi, R. and Kadonaga, J.T. (2000). A basal transcription factor that activates or represses transcription. Science *290*, 982-984.

Winkler, D.D., Muthurajan, U.M., Hieb, A.R. and Luger, K. (2011) Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events. Journal of Biological Chemistry *286*, 41883-41892.

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

Wittmeyer, J. and Formosa, T. (1997) The Saccharomyces cerevisiae DNA polymerase alpha catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein. Molecular and Cellular Biology *17*, 4178-4190 .

Wollmann, P., Cui, S., Viswanathan, R., Berninghausen, O., Wells, M.N., Moldt, M., Witte, G., Butryn, A., Wendler, P., Beckmann, R., Auble, D.T. and Hopfner, K-P. (2011). Nature *475*, 403-409.

Wu, P.Y., Ruhlmann, C., Winston, F. and Schultz, P. (2004). Molecular architecture of the S. cerevisiae SAGA complex. Molecular Cell *15*, 199-208.

Xin, H., Takahata, S., Blanksma, M., McCullough, L., Stillman, D.J., Formosa, T. (2009). yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement. Molecular Cell *35*, 365-376.

Yang, J., Zhang, X., Feng, J., Leng, H., Li, S., Xiao, J., Liu, S., Xu, Z., Xu, J., Li, D., Wang, Z., Wang, J. and Li, Q. 2016. The Histone Chaperone FACT Contributes to DNA Replication-Coupled Nucleosome Assembly. Cell Reports *14*, 1128-1141.

Yeung, K.C., Inostroza, J.A., Mermelstein, F.H., Kannabiran, C. and Reinberg, D. (1994) . Structure-function analysis of the TBP-binding protein Dr1 reveals a mechanism for repression of class II gene transcription. Genes and Development *8*, 2097-2109.

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

Yudkovsky, N., Ranish, J.A. and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. Nature *408*, 225-229.

Zanton, S.J. and Pugh, B.F. (2004). Changes in genomewide occupancy of core transcriptional regulators during heat stress. PNAS *101*, 16843-16848.

Zentner, G.E. and Henikoff, S. (2013). Mot1 redistributes TBP from TATA-containing to TATA-less promoters. Molecular and Cellular Biology *33*, 4996-5004.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S. and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. Nature Structural and Molecular Biology *16*, 847-852.

Zhang, Z. and Pugh, B.F. (2011). High-resolution genome-wide mapping of the primary structure of chromatin. Cell *144*, 175-186.

Zhou, K., Kuo, W.H., Fillingham, J. and Greenblatt, J.F. Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. PNAS *106,* 6956-6961.