The Pennsylvania State University

The Graduate School

The Eberly College of Science

ROLE OF CHROMATIN IN PROMOTER PROXIMAL PAUSING IN DROSOPHILA

A Dissertation in

Biochemistry, Microbiology and Molecular Biology

By

Bhavana Achary

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2013

The dissertation of Bhavana Achary was reviewed and approved* by the following:

David S Gilmour
Professor of Biochemistry and Molecular Biology
Dissertation Advisor
Chair of Committee

Song Tan
Professor of Biochemistry and Molecular Biology

Joseph C Reese
Professor of Biochemistry and Molecular Biology

Yanming Wang
Associate Professor of Biochemistry and Molecular Biology

Pamela Hankey
Professor of Immunology

Scott Selleck
Professor of Biochemistry and Molecular Biology
Head of the Department of Biochemistry and Molecular Biology

*Signatures are on file in the Graduate School

ABSTRACT

RNA polymerase II pauses in the promoter proximal region of thousands of genes. Amongst the various factors that contribute to pausing, chromatin architecture in the promoter proximal region is thought to be one of them. *In vitro* experiments have shown that nucleosomes can impede elongation of RNA Pol II. Genome wide maps of Pol II and nucleosomes in *Drosophila* indicate a dynamic interplay between the two. Currently, there are two prevalent models regarding the relationship between paused Pol II and nucleosome positioning. The first model indicates that a positioned nucleosome could contribute to pausing of RNA Pol II. The second model suggests that it is the paused Pol II that contributes to determining nucleosome organization in the promoter proximal region of a gene. In this study, I asked the question, what is the effect on nucleosome position when the position of paused Pol II is changed? To answer this, I determined the position of paused Pol II in embryos from wild type flies and flies with a mutant form of RNA Pol II. Permanganate footprinting experiments showed that the mutant RNA Pol II is paused closer to the transcription start site than in wild type. I mapped the nucleosomes in wild type and mutant to see if the shift in the paused Pol II affected the nucleosome position. The results indicate that the position of +1 nucleosome does not change when the position of paused RNA Pol II is shifted.

Pausing is widely recognized as an important step in regulation of gene expression and the mechanism of pausing and the factors involved are under investigation. The *Drosophila hsp70* gene has long been regarded as a model gene to study promoter proximal pausing. Transcription of the *hsp70* gene involves distinct steps, which include establishment of a paused Pol II, activation and release of the paused Pol II into productive elongation. Various factors are implicated in regulating these distinct steps. I have investigated the effects of depleting some of these factors on transcription of *hsp70*. Using a beta-galactosidase reporter assay as an initial

screen, I chose several factors that when depleted, showed reduced *hsp70* transcription. Some of the factors are known to be important in transcription of *hsp70*, such as the activator HSF and the Pol II CTD kinase, PTEF-b. One surprising result was the effect of depletion of HDAC3. HDAC3 depletion resulted in a decreased level of *hsp70* expression, both in the beta-gal assay and mRNA measurements. Upon further investigation, I observed that HDAC3 depletion affected RNA Pol II recruitment and impaired pausing on the *hsp70* gene.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
mRNA	messenger RNA
RNA Pol II, Pol II	RNA Polymerase II
NELF	Negative Elongation Factor
DSIF	DRB sensitivity inducing factor
GAF	GAGA Factor
HSF	Heat shock factor
P-TEFb	Positive transcription elongation factor b
Cdk9	Cyclin dependent kinase 9
CycT	Cyclin-T
Rpd3	Reduced potassium dependency 3
HDAC3	Histone deacetylase 3
SMRTER	SMRT-related ecdysone receptor interacting factor
ELL	Eleven nineteen lysine-rich leukemia
NURF	Nucleosome remodeling factor
HS	Heat shock
NHS	Non heat shock
$KMnO_4$	Potassium permanganate
TSS	Transcription start site

bp	base pair(s)
M/mM/μM/nM	molar/millimolar/micromolar/nanomolar
Mg/μg/ng	milligram(s)/microgram(s)/nanogram(s)
ml/μl	milliliter(s)/microliter(s)
min	minute(s)

ACKNOWLEDGEMENTS

The saying that “it takes a village to raise a child” can be applied to my years in graduate school. This work would not have been possible without the support and encouragement of numerous individuals and I want to take this opportunity to express my gratitude to them.

Firstly, I would like to thank my advisor Dr. David Gilmour for his advice and help throughout graduate school, even during the times I did not work under him. His guidance, patience and faith in my abilities have helped me tackle challenging problems. Needless to say, his mentorship has been integral to my own development as a researcher.

I would also like to thank my committee members Dr. Joseph Reese, Dr. Song Tan, Dr. Yanming Wang and Dr. Pamela Hankey for their guidance and input at various stages of my research work.

I have received help from all members of Centre for Gene Regulation and thank them for it. In particular I would like to thank Christine and Kuangyu who have allowed me to pick their brains numerous times and Vinesh for his help with my experiments.

I would like to thank all the Gilmour lab members for making it a supportive and interesting place to work. I have enjoyed my discussions both scientific and otherwise during long hours at lab. I would like to thank Saikat for his friendship over the years, Jian for his help particularly with the bioinformatic analyses, Bede for being a good bench partner, Laura for being ready to share a cup of coffee, Yijun for the time she spent as a listening board, Mike for his always interesting opinions and Doug for being the cheerful guy he is. I especially cherish my friendship with Anamika that was forged in lab and continues beyond the shared time spent at the bench.

I want to extend my heartfelt thanks to my friends who have always found time for me. To Anu, for never failing to pick up my call and always lifting my spirit. To Shak, for the steering hand and considerate soul. To Saras, for showing me that I could be a good roommate. To Saras and Raj for the good memories of hiking trips, conversations and meals shared. To Rekha and Murali for making me a part of their family, I will always be indebted to you for your love and affection.

I am immensely grateful to Bala for being by side throughout, for staying up with me while I wrote my dissertation, for his patience during the frustrating and anxious times and his "gentle pushes" when I slacked. It would have been very difficult to see this through without your love and encouragement. My in-laws have been very supportive and helped me continue my work and I am thankful for their support and prayers.

I have been blessed to have teachers who have guided me throughout my career. I am especially thankful to Asha Rao and Dr. Pasupathy for recognizing my potential and encouraging me to pursue my studies.

Last but not the least, my parents have been an immeasurable source of strength and support over these far too many years away from home. My mom's comforting words and my dad's sincere belief and confidence in me has kept me going through difficult times. They have been with me throughout this journey and I would like to dedicate this thesis to them.

Chapter 1

Introduction

Chromatin is the physiological template for transcription by RNA Polymerase II in higher eukaryotes. The basic structural unit of chromatin is the nucleosome. The packaging of DNA into chromatin can greatly limit the access of transcription factors to regulatory elements and affect the ability of RNA Pol II to transcribe across the gene. It thus serves as an important mode of regulation of gene expression, especially at the stages of initiation and elongation of transcription.

There are three well-defined stages in transcription. Transcription is initiated by recruitment of RNA Pol II to the promoter region, which is followed by elongation across the body of the gene. Termination of transcription results in release of the RNA transcript and dissociation of the transcription machinery from the DNA template. The traditional view of transcription being regulated primarily at the stage of recruitment of RNA Pol II and hence initiation has been challenged in recent years. It has emerged that regulation of elongation can impact gene expression at several steps. Recent studies have highlighted the existence of an additional stage of transcription in metazoans, which occurs after initiation of transcription and prior to elongation across the body of the gene. This stage has been characterized as promoter proximal pausing of Pol II. The positioning of nucleosomes and the covalent modifications on the histones within the nucleosomes can affect all these distinct stages of transcription.

1-1 Role of nucleosome positioning in transcription.

Nucleosomes tend to occupy specific locations across the gene. This is especially true for the first nucleosome downstream of the TSS, commonly referred to as the +1 nucleosome. Many

factors exert their influence on gene regulation by affecting the +1 nucleosome. For example, depletion of several chromatin-associated factors such as FACT, Chd1 and Paf1 affected the levels of the +1 nucleosome on the induced *Drosophila hsp70* gene (Petesch and Lis, 2008).

1-1.1 Role in regulation of initiation.

Nucleosomes serve as an impediment to transcription by regulating access of DNA binding factors to DNA elements. There are several examples where the positioned nucleosome is removed either by disassembly of the histones or remodeling by chromatin remodelers, prior to initiation of gene expression. In yeast, nucleosome disruption by activity of chromatin remodelers on the promoters of *PHO5* and *GAL1* genes precedes activation of the gene in response to stimulus (Almer et al., 1986; Fedor et al., 1989). Genome wide nucleosome maps in yeast show that the +1 nucleosome overlaps with the TSS, which signifies the need for remodeling of the nucleosome for initiation to occur (Albert et al., 2007). Similarly, the -1 nucleosome is evicted in the presence of Pol II and thus, the width of the NFR shows a positive correlation to transcription levels (Venters and Pugh, 2009). A similar correlation between the width of the NFR and transcription levels is also observed on the human genome (Ozsolak et al., 2007; Schones et al., 2008).

1-1.2 Role in regulation of elongation

Positioned nucleosomes can affect transcription elongation by blocking access to factors that regulate elongation and nucleosomes downstream of the TSS also act as physical barriers to the elongating RNA Pol II. *In vitro* experiments have shown that the presence of nucleosomes inhibits transcription and processivity of purified RNA Pol II (Izban and Luse, 1991a). *In vivo*,

elongation is facilitated by ATP-dependent chromatin remodelers and histone chaperones. For example, the chromatin remodeler, Swi/Snf is required for elongation on heat shock genes in yeast (Schwabish and Struhl, 2007). Swi/Snf is also recruited to the HIV LTR by the activator protein, Tat (Tréand et al., 2006). Tat functions as an activator by increasing the efficiency of elongation of RNA Pol II (Jones, 1997). The human *c-myc* gene shows the presence of Pol II that has already initiated transcription, and is localized at the promoter proximal region under both active and repressed states. However, under repressed conditions, the upstream region of the promoter is occupied by nucleosomes that could presumably block access to regulatory factors necessary for activation (Albert et al., 1997).

1-1.3 Role in pausing

In vitro experiments have shown that nucleosomes increase sequence specific pausing of RNA Pol II. Transcription on a human *hsp70* promoter assembled into a nucleosomal template showed increased pausing of Pol II (Brown et al., 1996). On the *Drosophila hsp70* gene, mapping of the MNase protected region before and after heat shock showed that nucleosome disassembly precedes transcription. This would indicate that the chromatin structure prior to heat shock possibly helps stabilize the paused Pol II (Petesch and Lis, 2008). However the mechanism by which the chromatin structure contributes to stabilizing the paused Pol II is unclear as the first nucleosome is located about 200 base pairs downstream from the paused Pol II.

Genome wide nucleosome maps in *Drosophila* and human cells detect a positioned nucleosome downstream of the paused Pol II (Mavrich et al., 2008a; Schones et al., 2008). This indicates that nucleosomes may play an influential role in pausing on a genome wide scale. However, this model has been challenged by results that suggest that contrary to the above model, it is instead the RNA Pol II that influences nucleosome positioning (Gilchrist et al., 2010).

1-2 Role of histone modifications in transcription

Covalent modifications of the N-terminal tails of histones contributes to the roles of chromatin in regulation of gene expression (Li et al., 2007a). The various modifications of the histone tails that can occur are acetylation, methylation, ubiquitination, phosphorylation and ADP-ribosylation. Histone modifications can act to alter the histone-DNA contacts and/or serve to recruit transcription factors that bind to the modified histone tails.

Overall, higher levels of acetylation are associated with actively transcribed genes. Acetylation of histones creates an open chromatin structure by "loosening" up the nucleosome structure or by recruiting factors that disassemble the nucleosomes. For example, activation of the HIV pro-virus promoter, is accompanied by acetylation of the nucleosome present downstream of the TSS, leading to disruption of the nucleosome and subsequent transcription elongation (Emiliani et al., 1998; Van Lint et al., 1996). Genome wide distributions of histone modifications in human CD4+ T-cells show increased levels of histone acetylation at the enhancer and promoter regions where they presumably facilitate access of DNA binding factors (Wang et al., 2008).

Acetylation of specific residues can also affect packaging of chromatin. H4K16 acetylation prevents folding of chromatin into a higher order structure (Shogren-Knaak et al., 2006). H4K16 acetylation is also involved in dosage compensation in flies. The X-chromosome in flies is hyper-acetylated at H4K16 by the MSL complex (Male specific lethal)(Gelbart and Kuroda, 2009). This hyper-acetylation results in a 2-fold increase in transcription of X-linked genes in male flies (Straub et al., 2005).

Histone modifications such as acetylation, methylation and phosphorylation of specific residues can be present in different combinations across a gene. For example, in *Drosophila* active promoter and promoter proximal regions are marked by tri and di methylation of H3K4 (H3K4me2/me3), H3K9 acetylation and H3K27 acetylation while H3K36me3 marks the body of

the gene (Kharchenko et al., 2011). These combinations of histone modifications possibly function to recruit factors that regulate transcription. The state of chromatin defined by enrichment for H3K27ac, H3K4me1 and H3K18ac is associated with paused genes and genes involved in development (Gilchrist and Adelman, 2012). The different histone modifications possibly define regions that are recognized by different sets of proteins and show different modes of regulation. Indeed, binding patterns of chromatin associated factors show correlations with the histone modifications mentioned above and have been used to "color-code" chromatin (Filion et al., 2010). These different chromatin states defined in these various studies reflect the importance of histone modifications in gene regulation.

1-2.1 Role of histone modifications in initiation

Tri-methylation of H3K4 by the *Drosophila* Set1 complex, correlates with gene expression levels (Ardehali et al., 2011). This modification is thought to regulate transcription by acting as a binding platform for a number of transcription factors. For example, the TAF3 subunit of TFIID binds to H3K4me3 via its PHD domain (Vermeulen et al., 2007). It is also required for recruitment of factors such as the chromatin remodeler NURF and Chd1 and mRNA processing factors involved in splicing towards the 5' ends of genes (Sims et al., 2007; Wysocka et al., 2006). The *Drosophila* Set1 complex is rapidly recruited to *hsp70* upon heat shock induction and dSet1 is thought to be involved in productive elongation on this gene. Depletion of dSet1 reduced the levels of Pol II released into the body of the gene. The accumulation of *hsp70* mRNA after 10 minutes of heat shock was significantly less in dSet1 depleted cells than in control cells (Ardehali et al., 2011).

1-2.2 Role of histone modifications in elongation

The bodies of actively transcribed genes show high levels of tri-methylated H3K36. In yeast, H3K36 tri-methylation is implicated in repression of cryptic transcription. Methylation of H3K36 is carried out by Set2 complex in yeast. The H3K36 tri-methylated residue results in recruitment of the histone deacetylase complex, Rpd3S, which deacetylates histones in the wake of elongating Pol II. The deacetylation of histones by Rpd3s is thought to "reset" the chromatin by facilitating the reassembly of nucleosomes after Pol II has transcribed across the gene body and thus prevents spurious transcription from cryptic promoters (Carrozza et al., 2005).

H3K36me3 is found preferentially over the exonic regions of transcribed genes in flies and worms (Kharchenko et al., 2011; Kolasinska-Zwierz et al., 2009). H3K36me3 is thought to function in alternate splicing mechanism by recruitment of the splicing factor, PTB, which results in inclusion of the PTB-dependent exons (Luco et al., 2010).

1-2.3 Role of histone modifications in pausing

Histone modifications can affect elongation via regulation of the paused state. On the human *FOSL1* gene, serum induction results in phosphorylation of H3Ser10 residue in the enhancer region by the kinase PIM1. The adaptor protein 14-3-3 recognizes the phosphorylated serine residue and recruits the H4K16 acetyl transferase MOF. The combinatory signals of phosphorylated H3Ser10 and acetylated H4K16 provides a platform for binding of the bromodomain containing protein, Brd4. Binding of Brd4 results in recruitment of PTEF-b and subsequent release of paused Pol II at the FOSLI gene (Zippo et al., 2009). In a later section, a similar histone cross-talk mechanism is described that functions on the *Drosophila hsp70* gene.

On the *FOSLI* gene, modification on H3Ser10 promoted the modification on H4K16 and influenced transcription in a combinatorial fashion. However, modification of a specific residue can also function to prevent modification on other lysine residues and affect transcription. For example, the presence of H3K20me3 prevents the binding of MOF and hence acetylation of H4K16 on the *TMS1* gene (Kapoor-Vazirani et al., 2011). The *TMS1* gene in certain human breast cancer cells shows the presence of paused Pol II. The gene is maintained in a repressed state in these cells by the presence of H3K20me3, which peaks upstream of the TSS. Depletion of the methyl-transferase SUV420H2, responsible for H3K20 methylation lead to increased levels of H4K16 acetylation. This leads to subsequent release of paused Pol II and up-regulation of the *TMS1* gene (Kapoor-Vazirani et al., 2011).

1-3 What is pausing?

As pausing of RNA Pol II is emerging as a prevalent mode of regulation of transcription, much focus has been directed towards elucidating the mechanism of pausing and the factors involved.

To study pausing, it is first important to define the state of paused Pol II. Pausing is a self-reversible state in which RNA Pol II ceases transcription for a period of time and is capable of resuming its intrinsic enzymatic reaction of addition of ribonucleotides on its own (Coppola et al., 1983). This state is distinct from the arrested state of Pol II, where the enzyme cannot resume the enzymatic reaction due to misalignment of the 3' end of the transcript at the active site and needs additional factors to resume elongation. There are additional factors that affect pausing of Pol II as discussed in the subsequent sections, but they do not affect the intrinsic enzymatic activity of Pol II. Pausing of Pol II was first reported in *in vitro* studies using nuclear extracts from HeLa cells (Coppola et al., 1983). Transcription complexes assembled on the adenovirus

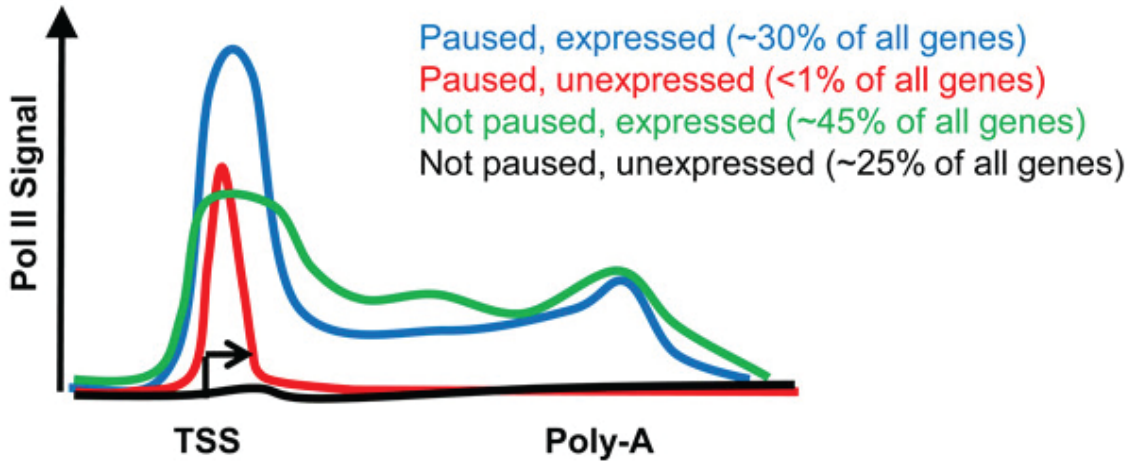
promoter under low NTP concentrations produced short 20 nt long RNA. These short transcripts could be chased to full-length transcripts with addition of NTPs. These transcripts were the result of pausing of Pol II close to the promoter and this state of Pol II was designated as promoter proximally paused Pol II.

Pausing of Pol II was discovered to occur *in vivo* on the *Drosophila hsp70* gene. UV crosslinking studies detected high levels of Pol II at the 5' end of the heat shock gene, *hsp70* (Gilmour and Lis, 1986). Nuclear run-on assays showed this Pol II molecule to be transcriptionally engaged (Rougvie and Lis, 1988). Permanganate foot-printing experiments, which detected the transcription bubble associated with transcriptionally engaged Pol II, detected Pol II paused at 20-30 nucleotides downstream of the TSS (Giardina et al., 1992).

The advent of DNA sequencing technologies led to the discovery that Pol II accumulates at the 5'end of genes in *Drosophila* and mammalian cells. The initial ChIP-chip studies showed that the Pol II distribution across genes followed three different patterns (Guenther et al., 2007; Muse et al., 2007). First, a peak of Pol II at the 5' end followed by an even distribution of Pol II across the gene. Second, a peak of Pol II at the 5' end with absence of Pol II on the gene and third, absence of 5' peak with an even distribution of or absence of Pol II across the gene (Figure 1-1). The ratio of Pol II levels at the 5' end and body of the gene is described as the pausing index and led to classification of genes as paused or non-paused based on the value of their pausing index. This method however, is ambiguous since the percentage of genes paused, based on their pausing index can vary from 30-90% (Adelman and Lis, 2012).

Pausing is influenced by a number of factors including DNA sequences, nascent transcript and protein factors. *In vitro* experiments have shown that Pol II has difficulty in transcribing certain DNA sequences and pauses in these regions (Kireeva et al., 2005). The secondary structure of the RNA is also thought to influence pausing by preventing backtracking

of Pol II on the template (Zamft et al., 2012). Various protein factors have been identified that regulate establishment of and release from the paused state.



Adapted from Adelman and Lis (2012), Nature Reviews Genetics

Figure **1-1**: Distribution of Pol II across the Drosophila genome.

Genome wide ChIP-seq experiments show that the distribution of Pol II on genes can be divided into 4 categories. Accumulation of Pol II at the 5' region of the gene indicates paused Pol II (blue and red traces). Paused but inactive genes show a 5' peak of Pol II and an absence of Pol II in the body of the gene (red trace). Paused and active genes show peak of Pol II at 5' end of the gene in addition to significant levels of Pol II in body of the gene (blue trace). Genes that are actively transcribed and not paused show a relatively even distribution of Pol II across the gene (green trace). Pol II is not detected on silent genes (black trace). Figure adapted from Adelman and Lis (2012).

1-3.1 Factors involved in setting up paused Pol II

Many of the factors and processes involved in pausing have been studied in the context of the *Drosophila hsp70* gene. Hence, I describe the factors that are involved in transcription of *hsp70* gene, which serves as the model gene to study regulation of pausing (Figure 1-2).

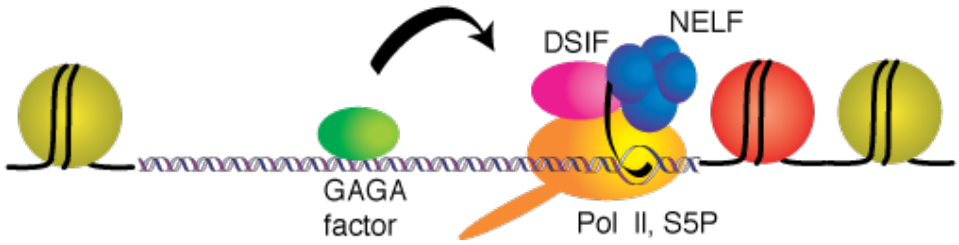


Figure **1-2**: Factors involved in setting up paused Pol II on the *hsp70* gene in *Drosophila*
The key factors involved in setting up the paused Pol II on *hsp70* include GAGA factor, NELF and DSIF. GAGA factor recruits the chromatin remodeler NURF and opens up the chromatin, allowing transcription initiation to occur. GAGA factor also recruits the pausing factor NELF. NELF and DSIF pause Pol II in the promoter proximal region of *hsp70* by binding to the elongation complex and the nascent transcript.

1-3.1.1 GAGA factor

Under non-heat shock conditions, the general transcription factor TFIID and DNA binding protein, GAGA factor are bound to the promoter of the *hsp70* gene (Shopland et al., 1995). GAGA factor recruits the chromatin remodeler NURF, that functions to open up the chromatin structure (Tsukiyama et al., 1995). The interaction between GAGA factor and TFIID could serve to recruit Pol II and initiate transcription. Early studies had shown that presence of GAGA elements in the promoter region of *hsp70* promoter fused to a silent gene *yp1*, resulted in pausing of Pol II on this gene (Lee et al., 1992). Mutations of the GAGA element lead to reduced pausing. However, it was not established if this effect was due to reduced initiation of transcription due to absence of GAGA factor binding or GAGA factor was directly involved in pausing of RNA Pol II (Lee et al., 1992). A number of recent reports have shown that GAGA factor is associated with highly paused genes (Gilchrist et al., 2008; Lee et al., 2008; Li and Gilmour, 2013). Recent experiments have shown that GAGA factor associates with the pausing factor NELF and depletion of GAGA factor decreases pausing (Li et al., 2013).

1-3.1.2 NELF

After initiation of transcription, Pol II transcribes about 20-40 nucleotides and pauses. Pausing of Pol II is brought about by the activities of two factors, NELF and DSIF.

NELF has four subunits, NELF A, B, D and E. NELF binds to DSIF/Pol II complex and was initially thought to promote pausing by binding to the nascent RNA that extrudes from the Pol II complex, via the RRM (RNA recognition motif) present on the NELF-E subunit (Yamaguchi et al., 2002). However, that model has been challenged by results that show DSIF, not NELF binds the 20-22 nucleotide long nascent transcript. DSIF could first bind to the

elongation complex and then recruit NELF to pause Pol II in the promoter proximal region (Missra and Gilmour, 2010). Depletion of NELF reduces pausing on *hsp70* gene and also on a genome wide scale (Wu et al., 2003; Gilchrist et al., 2010). Recent results show that NELF is also involved in re-establishment of the paused state on the *hsp70* gene when the cells recover from heat shock, by affecting the dissociation of the heat shock factor HSF (Ghosh et al., 2011).

1-3.1.3 DSIF

DSIF is hetero-dimeric protein that is conserved from yeast to mammals. It has two subunits Spt4 and Spt5 (Yamaguchi et al., 1999). *In vitro* studies show that DSIF along with NELF function as a negative regulator of transcription (Cheng and Price, 2007). Phosphorylation of DSIF by P-TEFb converts DSIF to a positive elongation factor (Fujinaga et al., 2004a; Wada et al., 1998).

DSIF and NELF are involved in pausing Pol II on the *hsp70* gene *in vivo* (Wu et al., 2003). Upon *hsp70* activation, DSIF is detected across the gene at rates mirroring Pol II suggesting that it tracks along with Pol II (Ni et al., 2004). DSIF interacts with the Capping enzyme (CE) and it is thought that DSIF/NELF-mediated pausing serves as a checkpoint for recruitment of CE and proper capping of the nascent transcript (Mandal et al., 2004; Sims et al., 2004).

1-3.2 Factors affecting release from pausing

1-3.2.1 P-TEFb

P-TEFb is a kinase that has two subunits, Cdk9, the kinase subunit and cyclin T1. It phosphorylates NELF and DSIF (Price, 2000), and it is recruited to *hsp70* gene by the heat shock factor, HSF (Lis et al., 2000). P-TEFb also phosphorylates the Pol II CTD (C-terminal Domain) at serine 2, which is a hallmark of the elongating form of Pol II. The level of elongating Pol II was reduced within minutes of treatment with Flavopiridol, a P-TEFb specific inhibitor and resulted in decreased levels of mRNA produced after heat shock (Ni et al., 2004, 2008). This indicates that activity of P-TEFb is required for release of paused Pol II into productive elongation. P-TEFb activity is also required for proper processing of mRNA as upon inhibition of its activity, there is an increase in the level of non-adenylated transcripts (Ni et al., 2004).

1-3.2.2 HSF

The heat shock factor, HSF, is the master regulator of transcription of the heat shock genes. Upon heat shock, HSF undergoes trimerization and binds to the heat shock elements within seconds and strongly activates the heat shock genes (Boehm et al., 2003; Sarge et al., 1993). Binding of HSF is dependent on the chromatin landscape defined by the presence of GAGA factor and paused Pol II (Shopland et al., 1995). Binding of HSF also correlates with the presence of active chromatin marks H3K4me3 and H3K9ac, which are present prior to HSF binding. However, it is not known if binding of HSF is dependent on the presence of these chromatin marks (Guertin and Lis, 2010).

1-3.3 Chromatin modifiers involved in pausing

1-3.3.1 PARP

Recent results have shown the Poly ADP-ribose Polymerase (PARP) participates in the loss of nucleosomes that accompanies transcription of the *hsp70* gene upon heat shock (Petesch and Lis, 2008). The activity of PARP could result in destabilization of the nucleosomes, generating a chromatin structure more conducive to transcription. Modification of H2A affects PARP activity. H2AK5 is acetylated by the Tip60 acetyl transferase and this acts as a signal for PARP activation. Loss of Tip60 activity inhibits the nucleosome loss on *hsp70* gene and reduces *hsp70* mRNA levels (Petesch and Lis, 2012). This is an example of how histone modification linked to nucleosome organization is important in regulating *hsp70* transcription.

1-3.3.2 Set1

The *Drosophila* Set1 carries out H3K4 tri-methylation near the TSS of transcribed genes. Depletion of Set1 resulted in decreased levels of *hsp70* mRNA (Ardehali et al., 2011). Depletion of Set1 also resulted in increased retention of Pol II at the promoter region and hence is implicated in the release of paused Pol II into the body of the gene. Decrease in levels of H3K4me3 probably affects the recruitment of factors involved in making the chromatin more accessible to RNA Pol II. A potential candidate is the chromatin remodeler, Chd1 that has a chromo-domain by which it can bind to methylated histones. Depletion of Chd1 leads to increased levels of nucleosomes across the *hsp70* gene prior to heat shock and increased retention of the +1 nucleosome after a 2 minute heat shock (Petesch and Lis, 2012). Chd1 localized to heat shock puffs supports the idea that it is involved in heat shock gene transcription (Stokes et al., 1996).

1-3.3.3 Jil-1

Phosphorylation of H3Ser10 in flies is implicated in activation of heat shock genes, by affecting recruitment of elongation factors. Upon heat shock, the H3Ser10 kinase Jil-1 is recruited to the *hsp70* gene (Karam et al., 2010). Phosphorylation of H3Ser10 results in recruitment of the adaptor protein 14-3-3 and the histone acetyl transferase, Elp3 to heat shock puffs. Elp3 acetylates H3K9 and was identified in yeast as part of the Elongator complex and functions in elongation. Jil-1 was thought to function at *hsp70* gene by recruiting P-TEFb in a manner similar to that observed on the human *FOSL1* gene (Ivaldi et al., 2007; Zippo et al., 2009).

However, it was later shown that there is no reduction in Pol II binding on polytene chromosome in Jil-1 mutants and no significant decrease in *hsp70* mRNA levels (Cai et al., 2008). Recent genome wide mapping of Jil-1 binding showed that while Jil-1 is present at active genes, the level of Jil-1 is not proportional with the levels of elongating Pol II bound on the active genes or the transcript levels (Regnard et al., 2011).

1-4 Scope and significance of thesis

The examples cited in the previous sections are representative of the growing number of studies highlighting the importance of chromatin structure in regulating transcription and more specifically in regulating promoter proximal pausing. The objective of my research project was to study the role of chromatin structure on pausing with a specific focus on the influence of the +1 nucleosome in pausing and the role of chromatin associated proteins on release from the paused state.

Several *in vitro* and *in vivo* studies lend credence to the hypothesis that the +1 nucleosome plays an important role in setting up the paused Pol II. *In vivo* maps of nucleosome positions correlate with the position of paused Pol II and *in vitro* studies show that presence of nucleosomes can result in pausing of Pol II (Izban and Luse, 1991b; Kireeva et al., 2005; Mavrich et al., 2008a). However, there are other studies that refute the hypothesis that positioned nucleosomes contribute to pausing. Nucleosome profiles on genes that are silent in tissue culture cells but are expressed in embryos show that genes without paused Pol II have higher levels of nucleosomes over the promoter (Gilchrist et al., 2010). Analysis of the published nucleosome maps from *Drosophila* embryos by a different group showed that the nucleosome profiles for paused genes and active, non-paused genes are similar (Rach et al., 2011). In my studies, I attempted to test this hypothesis by changing the location of the paused Pol II and asked if it affected the positioning of the +1 nucleosome. To do so, I carried out genome wide permanganate footprinting in wild type and a mutant fly line in which the Pol II pauses closer to the TSS. I then mapped the nucleosomes in these two fly lines to study if the position of paused Pol II correlated with the position of the +1 nucleosome. My results indicate that the +1 nucleosome is not shifted when the position of paused Pol II is changed.

Pausing is regulated at the stage of setting up the paused Pol II and the release from paused state. Release from paused state is thought to be a rate-limiting step in transcription of paused genes. I carried out a directed RNAi screen to identify factors that could be involved in this step. I employed a beta-galactosidase reporter gene fused to an *hsp70* promoter to assay for defects in activation of *hsp70* gene upon depletion of various factors. Surprisingly, I found that depletion of histone deacetylase, HDAC3 inhibited activation of *hsp70*. ChIP experiments showed that the recruitment of Pol II is lower upon heat shock and permanganate footprinting assay indicated that pausing is impaired.

The results presented here add to our understanding of how chromatin and its associated factors affect pausing and lend itself to directing further research into mechanisms involved in regulation of the paused state.

Chapter 2

Materials and Methods

2-1 Experiments carried out in salivary glands.

2-1.1 Beta-galactosidase staining assay

Third instar larvae from control and RNAi flies mated with yw;Z243,1824 flies were heat shocked at 37^0C for 30 minutes. The larvae were then allowed to recover at room temperature for 30 minutes to allow for synthesis of beta-galactosidase. Salivary glands were dissected in dissection buffer (130 mM NaCl, 5 mM KCl, 1.5 mM $CaCl_2$) on a silicone platform. After several pairs of salivary glands were collected, the dissection buffer was pipetted away. 100 μl of 1% gluteraldehyde solution diluted in dissection buffer was added and glands were incubated in this solution for 5 minutes. The gluteraldehyde solution was removed and glands were washed with dissection buffer by adding 100 μl of buffer to the platform and pipetting it away. The glands were transferred to 100 μl of X-gal stain solution in a 1.7 ml tube and placed on a rotator for 2 hours. Stain solution was made fresh prior to use by addition of 100 μl of X-gal (20 mg/ml X-gal dissolved in dimethylformamide) to 1 ml of stain solution (10 mM $NaPO_4$ (pH 7.2), 150 mM NaCl, 1 mM $MgCl_2$, 5.7 mM K4[FeII $(CN)_6$], 6 mM K_3[FeIII $(CN)_6$], 0.3% Triton x-100) pre-warmed at 37^0C for 5 minutes. After staining, the glands were carefully transferred to the silicone dish and washed with dissection buffer as described. The glands were placed in drop of 100% glycerol and photographed under a stereo microscope.

2-1.2 Reverse transcription PCR analysis

Salivary glands were dissected from third instar larvae in dissection buffer and placed in 10 μl of S2 insect media in 0.25 ml thin walled tube. Once 10 pairs of glands were collected, they were either processed directly for non heat shocked samples or heat shocked at 37^0C for 10 minutes in a heated PCR block.

Total RNA from 10 pairs of salivary glands per sample was extracted using 200 μl Trizol reagent following the manufacturer's protocol (Invitrogen). RNA was dissolved in 20 μl DEPC treated distilled, deionized water (ddH_2O) and was quantified using the Nanodrop. Equal amounts of RNA from each sample were run on a 2% agarose gel to ensure the integrity of samples and the concentrations of RNA in each sample.

200 ng of RNA was used to generate cDNA using 0.5 μM oligodT and 150 ng of random hexamers (Invitrogen). Samples were heated to 65^0C for 5 minutes and then chilled on ice. Reverse transcription was carried out at 37^0C for one hour with 100 units of MMLV-RT (Promega) in the presence of 1X transcription buffer (50 mM Tris-Cl, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$), 10 μM DTT, 0.2 units of RNasin (Promega). The reaction was stopped by incubation at 75^0C for 5 minutes.

cDNA was analyzed by quantitative real time PCR analysis with gene specific primers. Standard curves were generated by serial dilution of RNA and cDNA was generated from the diluted series of RNA. Amounts of cDNA in each sample were quantified relative to the standard curve and normalized against *rp49* cDNA in each sample to correct for differences in total RNA in each sample.

2-1.3 Chromatin Immunoprecipitation in salivary glands

ChIP was done as described by Ghosh et al., 2011. 10 pairs of salivary glands were isolated in dissection buffer (130 mM NaCl, 5 mM KCl, 1.5 mM $CaCl_2$) and then transferred to 10 µl of Drosophila Schneider media in a thin-walled 0.25 ml PCR tube. I found that using thin-walled tubes facilitated rapid heat shock and consistency of the heat shock treatments to different samples, in comparison to heat shocking whole larvae followed by dissection of salivary glands. Following heat shock, 90 µl of ice-cold dissection buffer was added. The glands were cross-linked by addition of 2.7 µl of 37% formaldehyde (EMD chemicals) followed by incubation on ice for 5 minutes, agitated intermittently to keep the glands in suspension and transferred to room temperature for 7 minutes. These sequential incubations were carried out to allow for better diffusion of the formaldehyde into the glands. The cross-linking reaction was quenched by addition of 5.4 µl of 2.5 M glycine to a final concentration of 125 mM and samples were placed on ice for 2 minutes. The glands were transferred to a 1.5 ml tube by puncturing the 0.25 ml tube and spinning the contents into the larger tube. The glands were pelleted at low speed for 2 minutes at 4^0C and supernatant was discarded. The glands were homogenized in 100 µl sonication buffer (20 mM Tris, pH 8.0, 0.5 % SDS, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM PMSF and protease inhibitor cocktail). The homogenization was carried out in three sequential steps: incubation at room temperature for 10 minutes, vortexed for 10 minutes and physically homogenized with a micro-pestle. These steps were important to achieve complete lysis of the crosslinked material. The homogenized chromatin was sonicated at 4^0C in a Diagenode Bioruptor at full frequency for 15 minutes at a 30 second on/30 second off cycle to give an average fragment size of 400 base pairs. The sonicated chromatin was centrifuged at 13000 rpm for 7 minutes at 4^0C to remove cell debris. The supernatant was transferred to a fresh tube, flash frozen and stored at -80^0C.

20 μl of the sonicated chromatin was used for each immunoprecipitation with a single antibody. The immunoprecipitations were carried out in 0.65 ml tubes. The chromatin was diluted with 380 μl IP buffer (0.5% Triton X-100, 2mM EDTA, 20mM Tris 8.0, 150 mM NaCl, 10% glycerol); then pre-cleared for 2 hours at 4^0 C with 15 μl of 50% slurry of Protein A Sepharose beads equilibrated in IP buffer + 1mg/ml acetylated BSA. Beads were centrifuged at low speed (1000 rpm) and supernatant transferred to a fresh tube. Immunoprecipitation was carried out with 4 μl of anti-Rpb3 (rabbit Ab), 2 μl of HDAC3 antibody (guinea-pig Ab, gift from Dr. Mannervik), 4 μl of IgG (pre-screening serum). Samples were incubated with antibody on a rotator overnight at 4^0C.

15 μl of 50% slurry of pre-equilibrated Protein A sepharose beads was added to the chromatin and incubated on a rotating platform for 2 hours at 4^0C. Beads were spun down and subjected to a series of washes for 5 minutes each and 1minute centrifugation at 1000 rpm between washes. The beads were washed once with low salt wash buffer (0.1% SDS, 1% TritonX-100, 2 mM Tris-Cl (pH-7.8) and 150 mM NaCl), thrice with high salt wash buffer (0.1% SDS, 1 % TritonX-100, 2 mM EDTA, 20 mM Tris-Cl pH-7.8 and 500 mM NACl) and once with lithium chloride wash buffer (0.1% SDS, 1% TritonX-100, 1 μM EDTA, 0.01 mM Tris-Cl pH-7.8, 0.25 M LiCl, 1% NP-40 and 1% Sodium deoxycholate). A second lithium chloride wash was carried out overnight at 4^0C. The beads were transferred to a fresh tube and washed twice with TE. The cross-linked protein-DNA complexes were eluted from the beads in 2 sequential steps of 15 minutes each with 100 μl each of Elution Buffer (1% SDS and 0.1M $NaHCO_3$). 8 μl of 5M NaCl was added to the eluates and reversal of cross-links was carried out for 3.5 hours at 65^0C. At the end of 3.5 hours, 15 μl of 1M Tris-Cl, pH 7.8 was added to the samples and the samples were treated with 2 μl of Proteinase K (10 mg/ml) for 30 minutes at 65^0C. To monitor the percent of DNA immunoprecipitated, reversal of crosslinks and proteinase K treatment was also done on

20 μl of chromatin that had not been immunoprecipitated. Extractions with phenol/chloroform/isoamyl alcohol, 25:24:1, and then chloroform followed by ethanol precipitation at room temperature was carried out. The DNA was dissolved in 50 μl of 1/2x TE. Serial dilutions of the input DNA control, generated from 20 μl of chromatin that was not immunoprecipitated was done to generate a standard curve against which measurements of immunoprecipitated DNA was carried out. 4 μl of DNA per sample was analyzed by qRT-PCR with appropriate primers.

2-1.4 Permanganate footprinting in salivary glands

Permanganate footprinting in salivary glands on individual genes was performed as described previously (Ghosh et al., 2011). 10 pairs of dissected salivary glands were used for each sample. Dissected glands were placed in S2 media prior to heat shock. Heat shocking of glands was carried out in 10 μl of S2 media in a thin-walled 0.25 ml tube at 37^0C. Heat shocked glands were immediately placed on ice and treated with 100 μl of 20 mM $KMnO_4$ for 2 minutes. The reaction was stopped by addition of an equal volume of 2X $KMnO_4$ stop solution (20 mM Tris-Cl (pH 8.0), 20 mM NaCl, 40 mM EDTA (pH 8.0), 1% SDS, 0.4 M beta-mercaptoethanol). The glands were allowed to lyse at room temperature. The lysates were treated with 50 μg of Proteinase K for 2 hours at 37^0C. The DNA was extracted with equal volume of phenol, phenol:chloroform:isoamyl alcohol (in a 25:24:1 ratio) and then chloroform. DNA was precipitated with $1/10^{th}$ volume of 3M sodium acetate and 2.5 volumes of ethanol and dissolved in 20 μl T.E (10 mM Tris-Cl (pH 7.5), 1 mM EDTA (pH 8.0)).

Permanganate treated DNA in a volume of 90 μl was treated with 10 μl of piperidine for 30 minutes at 90^0C. At the end of the reaction, 300 μl of water was added and sequential organic

extractions with isobutanol were carried out to remove piperidine. DNA was precipitated as described earlier and quantified.

Naked DNA was generated in a similar manner described by treating purified genomic DNA with 20 mM $KMnO_4$ for 1 minute. To generate G/A markers, purified genomic DNA in a volume of 20 μl was treated with 50 μl of 99% formic acid. The solution was incubated at 15^0C for 5 minutes. The reaction was stopped by addition of cold depurination stop solution (0.3M sodium acetate and 50 μg/ml tRNA). DNA was precipitated and subjected to piperidine cleavage as described above. Ligation mediated PCR (LM-PCR) was carried out with 150 ng of purified DNA as described earlier (Gilmour and Fan, 2009).

2-1.5 Immunofluorescence analysis of polytene chromosomes

Polytene chromosome squashes were prepared as previously described (Champlin et al., 1991). Two pairs of salivary glands were used for preparation of one slide. For heat shocked samples, whole larvae were heat shocked instead of dissected glands.

Larvae were dissected in dissection buffer and then transferred to a drop (50 μl) of Solution A (15 mM Tris-Cl (pH 7.4), 60 mM KCl, 15 mM spermine, 1.5 mM spermidine and 1% Triton X-100) for 30 seconds, followed by transfer to Solution B (15 mM Tris-Cl (pH 7.4), 60 mM KCl, 15 mM spermine, 1.5 mM spermidine, 1% Triton X-100 and 3.7% formaldehyde) for 30 seconds and final transfer to Solution G (50% glacial acetic acid). The glands were immediately transferred to a drop (9 μl) of Solution G on a siliconized coverslip and incubated for 3 minutes. A glass slide was carefully placed over the coverslip and turned over. The coverslip was gently tapped allowing for some movement of the coverslip to break open the nuclei. Chromosomes from RNAi depleted glands were at times more fragile and the use of the eraser end of a pencil to gently move the coverslip around instead of tapping on it greatly helped

prevent the chromosomes from breaking during the spreading process. Once the spreading of the chromosomes was observed, the slides were frozen in liquid nitrogen. The coverslip was removed and the slide was placed in a coplin jar with 95% ethanol, overnight at 4^0C.

Staining of slides was carried out at room temperature. The slides were rehydrated in TBST (150 mM NaCl, 10 mM Tris-Cl (pH 7.5), 0.03% Triton X-100) for 10 minutes. The slides were then washed once in TBST for 5 minutes. The slide was carefully placed on a coverslip with 20 µl of blocking solution (10% Fetal Bovine Serum in TBST) and incubated at room temperature for one hour. The slides were then washed with TBST thrice for 5 minutes each. Primary antibodies were diluted in 10% FBS in TBST and 20 µl was used for staining of each slide. HDAC3 antibodies were used at 1:50 dilution and the antibodies were a gift from Dr Mattias Mannervik. Rpb3 antibodies were used at a 1:100 dilution. Staining with primary antibodies was carried out for 2 hours at room temperature. The slides were then washed thrice with TBST for 10 minutes each. Appropriate fluorescently tagged secondary antibodies were used at a 1:250 dilution and 20 µl was used per slide. The slides were incubated at room temperature for 90 minutes and then washed thrice with TBST. The slides were then washed with TBS containing 2 ng/ml Hoechst for 20 minutes and then washed with TBS for 20 minutes. The slides were then mounted onto a coverslip with 20 µl of mounting solution (100 mM Tris-Cl (pH 8.5), 80% glycerol, 2% n-propylgallate). Chromosome spreads were visualized by fluorescence microscopy.

2-2 Experiments carried out in embryos

2-2.1 Crosslinking of embryos for nucleosome mapping (heptane method)

Collections of 12-18 AED (after egg deposition) embryos were dechorionated with 50% bleach solution for 90 seconds. Next, they were washed with embryo wash buffer (0.03% Triton X-100, 140 mM NaCl) and then extensively with water. One gram of embryos was transferred to Falcon tube with 10 ml ChIP-FIX solution (50 mM HEPES (pH 7.6), 100 mM NaCl, 0.1mM EDTA, 40 mM EGTA, 2% formaldehyde) and 30 ml heptane. The solution was vigorously shaken for 15 minutes on a vortexer. Embryos were spun down and the aqueous phase was discarded. 10 ml of PBST (140mM NaCl, 2.5 mM KCl, 10 mM Na_2HPO_4 (dibasic), 1.8 mM KH_2PO_4 (monobasic), 0.01% Triton x-100) and 0.125 mM Glycine was added to stop crosslinking. The solution was vigorously shaken for an additional 5 minutes and then the embryos were washed thrice with PBST for 5 minutes in each wash. One gram of crosslinked embryos were homogenized in 5 ml of homogenization buffer (0.3 M Sucrose, 10 mM HEPES (pH 7.6), 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA) to which DTT, sodium bisulfite, PMSF were added fresh to a final concentration of 1 mM, 1 mM and 0.2 mM respectively. Embryos were homogenized in a glass dounce with 10 strokes of loose pestle and 15 strokes of tight pestle and the homogenate was centrifuged at 2000 rpm for 10 minutes at 4^0C in HB4 rotor. The crude nuclear pellet was further processed using the method described in preparation of crude nuclei (Section 2.3).

The crosslinking efficiency using this procedure was low and I was unable to immunoprecipitate H3. This procedure was used successfully on 0-12 hour embryos (Mavrich et al., 2008a), but the older age of the embryos used in my experiments were less permeable to formaldehyde using this technique. Nevertheless, the nucleosome distributions obtained from

embryos that were processed by this method are in good agreement with nucleosome distributions obtained from tissue culture cells that were crosslinked and immunoprecipitated for H3 (Figure 3-13, Chapter 3). Other studies have also been reported that do not employ a crosslinking step in preparation of MNase treated nuclei (Teves and Henikoff, 2011).

2-2.2 Crosslinking of embryos by homogenization

Collections of 12-18 AED embryos were dechorionated with 50% bleach solution for 90 seconds. Next, they were washed with embryo wash buffer (0.03% Triton X-100, 140 mM NaCl) and then extensively with water. Embryos were weighed in a pre-weighed 15 ml Falcon tube. The weight of the collection of embryos varied from 1 to 3 grams. Collections of embryos weighing more than 3 grams were split into more than one tube for further processing. The embryos were then transferred to a 7 ml dounce by adding 5 ml of Buffer A1z (15 mM HEPES (pH 7.5), 15 mM NaCl, 60 mM KCl, 4 mM $MgCl_2$, 0.5% Triton X-100, 0.5 mM DTT, Protease inhibitor cocktail-2.5 μl). Embryos were crosslinked with 2% formaldehyde for 15 minutes while being homogenized on ice. 270 μl of 37% formaldehyde was added and the embryos were then homogenized with 5-10 strokes of the loose and tight pestle respectively (till the pestle could move freely). The reaction was quenched by addition of 0.125 M glycine after 15 minutes. The solution was mixed well and incubated on ice for a minimum of 3 minutes. Solution was filtered through Mira cloth into a 15 ml polypropylene oak ridge style tubes and centrifuged at 4500 rpm (3000g) for 5 minutes, in a HB-4 rotor. Supernatant was discarded and pellet was washed thrice by resuspending nuclei in 5 ml of Buffer A1z and centrifuging at 6500 rpm in a HB-4 rotor.

2-2.3 Preparation of crude nuclei from cross-linked embryos

The pellet of homogenized, cross-linked embryos was transferred to 7 ml dounce with 5 ml of Nuclear Buffer A+ 0.3M Sucrose (60 mM KCl, 15 mM NaCl, 0.1 mM EGTA, 0.5 mM DTT, 0.1 mM PMSF, 15 mM Tris-Cl (pH 7.4), 5 mM $MgCl_2$, 0.3M Sucrose, 0.5% NP-40). 40 μl of 25% Triton was added to the 5 ml of Nuclear Buffer A in the dounce to a final concentration of 0.2% and the pellet was homogenized using a glass dounce. Solution was centrifuged at 6500 rpm, for 10 minutes, at 4^0C in HB4 rotor. The crude nuclei pellet was re-suspended again in 5 ml of Nuclear Buffer A + 0.3M Sucrose. This was layered on top of 10 ml of Nuclear Buffer A +1.7M Sucrose and centrifuged in the cold at 10,000 rpm for 15 minutes. Supernatant was carefully removed and discarded. The nuclear pellet was resuspended in NPS+B-Me buffer such that 0.1 gm of embryos was resuspended in 120 μl of NPS+B-Me buffer. (NPS Buffer: 0.5 mM Spermidine, 0.075% IGEPAL, 50 mM NaCl, 10 mM Tris-Cl (7.5), 5 mM $MgCl_2$, 1 mM $CaCl_2$.). Beta-mercaptoethanol was added fresh to a final concentration of 1 mM.

2-2.4 MNase treatment

After standardizing conditions such that the major band obtained corresponded to mononucleosomal sized DNA, the conditions used for MNase treatment for all the subsequent MNase digestion experiments were as follows. 0.1 gm of embryos in 120 μl of NPS+B-Me buffer were treated with 10 μl of MNase at a concentration of 5 units/μl. The reaction was allowed to proceed for 10 minutes at 25^0C (in the thermocycler) with intermittent mixing to resuspend the nuclei that settle to the bottom of the tube. The digestion was stopped by adding 10 μl of MNase Stop solution (2.8 μl of 0.5 M EDTA+ 2.8 μl of 10% SDS + 4.4 μl of NPS+B-Me buffer).

The pellet was mildly sonicated (5 cycles of 30 seconds on, 30 seconds off at medium strength in the Bioruptor) to release DNA from the insoluble chromatin. Earlier protocols (as used in Mavrich et al., 2008) spun down the nuclei after sonication and discarded the supernatant. This step was not included as it was observed that a small fraction of mononucleosomal DNA is released into the supernatant after MNase treatment. This could arise from the region corresponding to the readily accessible chromatin present at the promoter. Indeed, it has been reported that chromatin easily released by MNase treatment under low salt conditions are associated with active genes (Teves and Henikoff, 2011).

The sonicated chromatin (140 μl) was spun down for 7 minutes at 13,000 rpm at 4^0C. The supernatant was transferred to a fresh tube. The pellet was saved to check if the DNA was efficiently released by sonication. Analysis of the DNA recovered from the pellet and the supernatant on an agarose gel indicated that greater than 90% of DNA was in the supernatant and very little remained in the pellet after sonication.

MNase treated chromatin was incubated at 65^0C for a minimum of 4 hours and not more than 16 hours to reverse the protein-DNA crosslinks. Samples were treated with RNAse (2mg/ml) for 30 minutes at 37^0C. The proteins were degraded by treatment with Proteinase K (0.25 mg/ml) at 37^0C for 2 hours. Successive extractions with phenol, phenol:chloroform:isoamyl alcohol (25:24:1 ratio) and finally chloroform were carried out. . One μl of 10 mg/ml of glycogen was added to help visualize the pellet. The DNA was precipitated with $1/10^{th}$ volume sodium acetate, 1 μl of 10 mg/ml glycogen and 2.5 volumes of ethanol.

DNA was run on a 2% gel at 120 volts for one hour. The band corresponding to mononucleosomal DNA (around 150 bp) was gel purified using gel purification kit according to manufacturer's instructions (Qiagen). The gel purified DNA was taken through the steps involved in generating a library for high-throughput sequencing.

2-2.4.1 Library preparation of mononucleosomal DNA for sequencing on the Illumina Hiseq2000 machine

100 ng of gel purified mononucleosomal DNA was used for library preparation. First, the ends of DNA fragments were phosphorylated with T4 polynucleotide kinase (PNK) enzyme in the presence of T4 DNA Ligase Buffer (1x T4 DNA Ligase Buffer- 50 mM Tris-Cl (pH 7.5), 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, 25 μg/ml BSA) in a 20 μl reaction volume, at 37^0C for 30 minutes. The enzyme was heat inactivated at 65^0C for 15 minutes. Next, an "A" base was added to the 3' end of DNA by incubation with dATP and Klenow exo- in the presence of NEB Buffer 2 (50 mM NaCl, 10 mM Tris-Cl (pH 7.9), 10 mM Mg Cl_2, 1 mM DTT) at 37^0C for 30 minutes. Enzyme was heat inactivated at 75^0C for 10 minutes. The total reaction volume for the A-tailing reaction was 30μl (20μl from the previous ligase reaction).

Exa1 and Exa2 adaptors required for the subsequent sequencing of the DNA were ligated at the same time. Adaptors were used in 15-fold excess over the amount of mononucleosomal DNA. Ligation was carried out at 25^0C for 1 hour with T4 Ligase obtained from Dr. Song Tan and in the presence of T4 DNA ligase buffer. The total reaction volume at this step was 50 μl (30 μl template DNA+ 20 μl reaction volume for the ligation of the adaptors). Oligos that include a 6 base pair unique sequence used to tag each sample ("barcode") were ligated onto the Exa2 adaptor in a subsequent step. The products of the ligation reaction were purified to remove excess adaptors (AMPure beads from Beckman Coulter were used to purify the ligation products). The DNA was then run on a 6% polyacrylamide gel to further gel purify the DNA ligated to the adaptors (approximate size is 250 basepairs). Gel-purified DNA was suspended in 100 μl of ddH_2O and $1/3^{rd}$ was used for one PCR reaction with P1.2 and P7 primers. The volume of the PCR reaction was 100 μl and at the end of the PCR reaction, 15 μl was run on a 2% gel. The number of cycles (18 cycles) was kept to the minimum that was required to visualize a band of

250 bp in a 2% agarose gel. In a final step, the PCR amplified DNA was further gel purified from a 2% agarose gel using a gel purification kit (Qiagen) and submitted for sequencing.

2-2.5 Permanganate ChIP-seq in embryos

2-2.5.1 Permanganate treatment of crosslinked embryos

Crosslinking was carried out on homogenized 12-18 AED embryos as described in section (2-2.3). The crude pellet of nuclei following the crosslinking reaction and subsequent washes was resuspended in dissection buffer (130 mM NaCl, 5 mM KCl, 1.5 mM $CaCl_2$). The volume of dissection buffer used was based on the initial weight of the embryos such that 100 mg of embryos were resuspended in 150 µl of buffer. The suspension was treated with equal volume of 20 mM potassium permanganate ($KMnO_4$) for 1 minute on ice. The reaction was stopped by addition of beta-mercaptoethanol (0.2M beta-mercaptoethanol, 100 mM EDTA) to quench the permanganate. The nuclei were pelleted by a centrifugation at 1000 rpm for 7 minutes at 4^0C and the pellet was washed twice with dissection buffer. The nuclei were then resuspended in sonication buffer (0.5% SDS, 20 mM Tris-Cl (pH 8.0), 2 mM EDTA, 0.5 mM EGTA, 0.5 mM PMSF, protease inhibitors at 1:1000 dilution) such that 150 mg of embryos were in a volume of 250 µl. The nuclei were incubated on ice for 10 minutes to allow for lysis and then sonicated in the Bioruptor (30 cycles of 30 seconds on, 60 seconds off at high power settings). The sonicated chromatin was centrifuged for 7 minutes at 13000 rpm at 4^0C, the supernatant was flash frozen in liquid nitrogen and stored at -80^0C and the pellet of cell debris was discarded.

Permanganate treatment of uncrosslinked embryos was carried out to analyze individual genes (Gilmour and Fan, 2009; Wang et al., 2007). Approximately 10-15 mg of embryos were dechorionated and resuspended in 50 µl of dissection buffer. Embryos were briefly homogenized

with a motorized pestle and treated with equal volume of 20 mM $KMnO_4$ for 1 minute on ice. The reaction was stopped by addition of equal volume of 2X Stop solution as mentioned in section 2-1.4. The samples were then purified using procedures similar to that described for salivary glands (Section 2-1.4).

2-2.5.2 ChIP for Pol II with permanganate treated chromatin.

The equivalent of 85 mg of embryos from one sample was used for immunoprecipitation of Pol II and the subsequent library preparation of DNA fragments to be sequenced. Based on estimates of the weight of an embryo and the number of cells in an 12-18 hour old embryo, this is equivalent to 50 million cells for one sample. While this is in excess, it is prudent to start with a higher number of cells due to the loss of DNA during the multitude of steps in this protocol.

The steps involved in immunoprecipitation of RNA Pol II are the same as described for ChIP in salivary glands till the step of final washes with TE. Briefly the protocol is as follows. The chromatin is pre-cleared and incubated with antibody overnight. The antibody-protein complex is pulled down with Protein A sepharose beads. The beads are extensively washed and transferred to a fresh tube. The subsequent enzymatic steps were carried out in low retention eppendorf tubes (0.65 ml tubes).

2-2.5.3 Library preparation with permanganate treated and immunoprecipitated chromatin

The initial steps in the preparation of library up until treatment with piperidine were carried out on beads. After each enzymatic step, the beads were washed in the same manner as in a ChIP experiment (1 wash with low salt buffer, 3 washes with high salt buffer and 1 wash with lithium chloride buffer, as described in ChIP on salivary glands protocol)

The ends of the DNA bound to the beads were kinased and polished with T4 PNK and T4 DNA polymerase enzymes. The 3' ends of the DNA were "A-tailed" and next ligation with the Exa2 adaptor was carried out. Ligation of the "indexed" primer to the Exa2 adaptor was carried out in a subsequent step. These reactions were carried out on a rotortoque at the bench, as it is important to keep the beads in suspension. At the end of each reaction, the beads were washed by addition of 400 μl of wash buffer and placed on a rotator for 5 minutes and beads were pelleted by centrifugation at 1000 rpm for 2 minutes. The ligated adaptor was then extended in the "fill-in" reaction by phi29 enzyme.

The DNA was then eluted from the beads with elution buffer (1% SDS, 0.1 M $NaHCO_3$) as described for ChIP with salivary glands. Reverse crosslinking of the protein-DNA complexes was carried out and the DNA was precipitated.

The eluted DNA was treated with piperidine that cleaves at the oxidized thymines generated by the permanganate reaction and this results in the 5' end of the DNA fragment being available for ligation of the next adaptor in the subsequent steps in preparation of the library. Piperidine treated DNA was subjected to extraction with isobutanol and then ether and finally precipitated with ethanol.

The DNA precipitated from the piperidine cleavage reaction was then subjected to a primer extension reaction. DNA was denatured and primer extension with P7 primer, complementary to the Exa2 adaptor was carried out with phi29 polymerase to generate double stranded DNA. The 5' end of the DNA fragment not ligated to Exa2 corresponded to the site of the oxidized "T". "A" bases were added to the 3' ends of the DNA. This helped in the ligation of the Exa1 adaptor in the next step.

Next, Exa1 adaptor was ligated. Ligation of Exa1 to the end of DNA fragment that corresponds to the oxidized T was driven by the absence of a phosphate group on Exa2 adaptor that was ligated earlier and the presence of "T" overhang at the 3' end of Exa1 adaptor.

PCR amplification of the purified DNA with P7 and P1.2 (sequencing primer) primers was carried out. P1.2 primer contains the sequence that allows the DNA to be bound to the flow-cell during the sequencing reaction and hence is referred to as the sequencing primer. Pfu enzyme was used for PCR amplification as it gave me a better yield as compared to other recommended enzymes such as Phusion enzyme. The PCR products (between 200 to 300 basepairs) were gel purified from a 2% agarose gel using a gel purification kit and according to manufacturer's instructions (Qiagen). The samples were then submitted for sequencing.

2-3 Analysis of sequencing data

Samples were sequenced on Illumina Hiseq2000 machine and reads were mapped to Drosophila reference genome 5.2, using the BWA program (Li and Durbin, 2010), allowing for up to 6 mismatches. For the permanganate ChIP-seq data, each read was assigned to a position that is one nucleotide upstream from the 5' end where the piperidine cleavage occurred. The numbers of uniquely mapped reads for each sample were as follows:

1.Permanganate ChIP-seq (WT-A= 2370596, WT-B= 4661695, C4-A = 8667423,C4-B= 4661695)

2. MNase treated embryos (crosslinked by heptane method) (WT-A = 40,719,161, WT-B= 42,512,217, C4-A= 42,881,086, C4-B= 40,116,422)

3. MNase treated embryos (crosslinked by homogenization) (WT-A= 25,005,382, WT-B= 25,976,134, C4-A= 23,362,583, C4-B= 24,773,885)

The letters A and B represent biological replicates. The percentage of uniquely mapped reads was approximately 70% for each sample. For the subsequent analyses that determine location of paused Pol II and the nucleosome distributions in WT and C4 embryos, I combined the reads from the biological replicates.

40% of the permanganate ChIP-seq data mapped to a T. This is similar to the 55% of reads that mapped to a T in the permanganate ChIP-seq experiments Jian Li and Ho Sung Rhee obtained from Drosophila tissue culture cells and was sequenced with a SOLiD sequencing machine. The genomic data corresponding to each read was obtained using tools on Galaxy (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010). This gave me the identity of the nucleotide associated with each read and I filtered the reads to retain only the reads that corresponded to a thymine. This then represented the reactive "T". The reactive "T's" were mapped to a reference point on the gene and I used the transcription start site (TSS) as the reference point. This was done using Bedtools (Quinlan and Hall, 2010). Binning of data was carried out using scripts written by Jian Li and composite plots were drawn using Excel. Clustering of data was carried out with Cluster 3.0 program and heat maps were generated with Java TreeView package (Saldanha, 2004).

To evaluate the reproducibility of the data, I analyzed each biological replicate from control (WT) and mutant (C4) embryos (described in Chapter 3) separately using the bioinformatics methods described above. I show the composite plots of the permanganate reactivity on 2 groups of genes that differ in the behavior of C4 and WT RNA Pol II for each pair of replicates in Figure 2-1. The positions of permanganate reactivity overlap well in the biological replicates from control (WT) and mutant (C4) embryos. I computed the Pearson correlation coefficient for the biological replicates and there is a positive correlation between the replicates in both WT and C4. The set of genes used for this analysis have been filtered to exclude the genes that had levels of permanganate reactivity falling below a value of 3-fold over the reactivity detected in intergenic regions. I used this same set of gene in the subsequent analyses described in Chapter 3. I summed up the reads in a region 200 nucleotides upstream and downstream of the TSS for each gene in each sample and compared the sum in each biological replicate. The

Pearson correlation coefficient of determination (R^2) for WT-A and WT-B is 0.910 (n= 3,444) and the R^2 value for C4-A and C4-B is 0.833 (n=3,444).

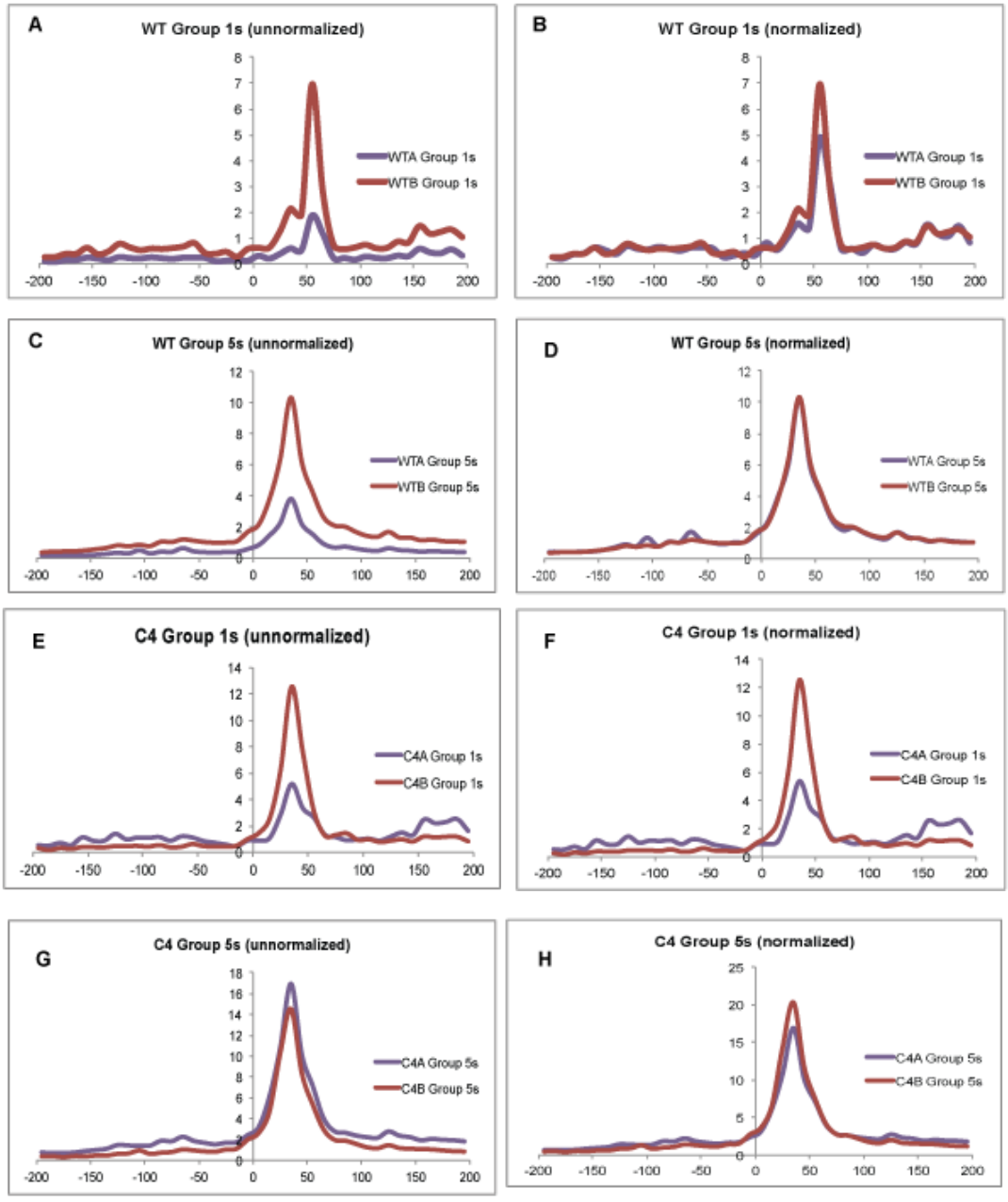


Figure **2-1**: Composite plot of permanganate reactivity in biological replicates of WT and C4 embryos.

(A-D) Composite plot of permanganate reactivity on different groups of genes (as defined in Figure 3-6, Chapter 3) in biological replicates of permanganate ChIP-seq data from WT embryos. WT-A has 2.5 fold lower number of sequencing reads than WT-B. Panel A and C show the composite plot derived from the raw reads divided by the number of genes in each group. Panel B and D show the composite plot derived from the raw reads divided by the number of genes in each group after WT-A has been normalized to correct for differences in total number of reads. **(E-H)** Composite plot of permanganate reactivity on different groups of genes (as defined in Figure 3-6, Chapter 3) in biological replicates of permanganate ChIP-seq data from C4 embryos. C4-A has 1.2-fold lower number of reads than C4-B. Panel E and G show the composite plot derived from the raw reads divided by the number of genes in each group. Panel F and H show the composite plot derived from the raw reads divided by the number of genes in each group after C4-A has been normalized to correct for differences in total number of reads.

Chapter 3

Relationship between nucleosome positioning and paused Pol II

3-1 Introduction

Eukaryotic DNA is organized into a condensed structure called chromatin. Chromatin consists of DNA, wrapped around histone octamers composed of two units of H2A, H2B, H3, and H4. This organization results in wrapping of 147 base pairs of DNA in about 1.7 helical turns about the histone octamer surface, resulting in a structural unit called the nucleosome. These nucleosomes are arranged in an array across the gene.

3-1.1 Nucleosome positioning

Nucleosome positioning refers to the preferred position of a nucleosome with respect to the underlying DNA sequence. Nucleosome positioning has two features: translational and rotational positioning. Translational positioning is the result of the underlying DNA sequence that determines the position of the histone octamer along the DNA. Nucleosomes that occupy a fixed position with respect to the underlying sequence of DNA are translationally positioned. Rotational positioning is the result of the intrinsic curvature of DNA that defines which side of the DNA faces the histone octamer. When the orientation of the DNA helix with respect to the surface of the histone octamer is fixed, the nucleosome is rotationally positioned.

3-1.2 Role of nucleosome positioning in gene regulation

The organization of the DNA into chromatin influences the process of transcription of the DNA template. *In vitro* experiments have shown that nucleosomes act as a physical roadblock to the transcribing Pol II (Izban & D S Luse, 1991). Experiments using templates, which incorporate DNA sequences with high affinity for histones such as the 601 positioning sequence, showed that the nucleosomal block could be relieved only in the presence of high NTP concentrations and/or high salt concentrations (Bondarenko et al., 2006). The presence of a nucleosome on a DNA template that has sequences which Pol II has an inherent difficulty in transcribing through, increases the propensity of Pol II to become arrested (Kireeva et al., 2005).

In vivo, RNA Pol II is able to overcome the nucleosome barrier and transcription is accompanied by loss of nucleosomes on highly transcribed genes (Petesch and Lis, 2008). To overcome the nucleosomal barrier, RNA Pol II is assisted by several factors. For example, FACT a histone chaperone protein that displaces H2A/H2B dimers, is involved in transcription of the *hsp70* gene (Orphanides et al., 1998). Spt6, another histone chaperone protein is also required for efficient elongation of the *hsp70* gene in *Drosophila* (Ardehali et al, 2009). Pol II can backtrack several nucleotides along the DNA template when it runs into a nucleosome. Backtracking can result in misalignment of the 3' end of RNA with the active site and induce an arrested state of Pol II where it is unable to spontaneously resume elongation. TFIIS is a protein that stimulates the cleavage activity of RNA Pol II, thus allowing it overcome the arrested state (Kireeva et al., 2005).

3-1.2.1 Nucleosomes regulate access to DNA sequence elements

In addition to posing a physical barrier to Pol II, nucleosomes can also influence transcription by regulating access to DNA elements recognized by DNA binding factors. One of several examples is the phosphate regulated *PHO5* promoter in yeast (Almer et al., 1986). The gene is repressed when the concentration of phosphate is high. Under these conditions, the *PHO5* promoter has an array of well-positioned nucleosomes. The nucleosomes prevent the transcription machinery from binding to the TATA box. Upon activation, under limiting phosphate conditions, the nucleosomes upstream of the TSS are evicted, exposing additional binding sites for the activator (PHO4) and also allows the transcription machinery to access the promoter elements such as the TATA box (Almer et al., 1986).

The heat shock genes in *Drosophila* provide an example of an alternate scenario where the changes in chromatin structure occur prior to receipt of an activating signal. Under non-heat shock conditions, binding of the DNA binding protein GAGA factor to the *hsp70* promoter results in recruitment of the chromatin remodeler NURF that functions to open up the chromatin allowing initiation to occur (Badenhorst et al., 2002; Tsukiyama and Wu, 1995).

The *hsp26* gene differs from the other heat shock genes in *Drosophila* due to the presence of a positioned nucleosome in the promoter region. Two DNAse hypersensitive sites flank the positioned nucleosome. The HSE elements to which the Heat Shock Factor binds upon activation are present in the DNAse sensitive regions. The nucleosome present in this region is thought to bring the two HSE's close to each other and promote cooperative binding of HSF to the promoter (Lu et al., 1995).

3-1.3 Chromatin landscape of eukaryotic genes

Advances in DNA sequencing have allowed us to progress from drawing inferences from studies of a handful of genes to deriving global maps of locations and interactions between factors involved in transcription. Genome wide sequencing of nucleosomal DNA has revealed general features of the nucleosome organization on a typical gene. The promoter region is generally depleted of nucleosomes. There is a NFR (Nucleosome Free Region) just upstream of the TSS and at the 3' end of the gene. There is a regular array of nucleosomes in the body of the gene, with a well-positioned +1 nucleosome. The +1 nucleosome has the tightest positioning and its histone tails are often acetylated and /or methylated. The nucleosomes further away from the TSS are more de-localized and hence "fuzzier" (Jiang and Pugh, 2009; Radman-Livaja and Rando, 2010).

Features of the DNA sequence influence the ability of DNA to wrap around the histone octamer and thus contribute to positioning of the nucleosome. Stretches of DNA comprised of the dinucleotide sequence AA/TT, spaced at 10 base pair intervals were found to be enriched in well-positioned nucleosomes in yeast, especially at the +1 nucleosome location (Mavrich et al., 2008b). However stretches of DNA with poly dA or poly dT sequences are refractory to nucleosome positioning. In yeast, it has been suggested that the combination of nucleosome positioning sequences at the +1 nucleosome location and the poly dA:dT stretches at the promoter serve to provide a barrier against which the downstream nucleosome array is arranged (Mavrich et al., 2008b).

Comparisons of *in vivo* nucleosome maps with positions of *in vitro* reconstituted nucleosome show a high correlation (Kaplan et al., 2009). A primary factor for the high correspondence is thought to be nucleosome exclusion at the ends of genes as there is very good overlap of the nucleosome free regions in the *in vivo* and *in vitro* maps (Radman-Livaja and

Rando, 2010). Biochemical experiments with reconstitution of chromatin and whole cell extracts showed that while there is a strong correlation between consensus location of poly dA:dT tracts and +1 nucleosome position, recapitulation of *in vivo* nucleosome positions beyond the +1 position requires the activity of ATP-dependent chromatin remodelers. Thus, DNA sequence is not the sole determinant of nucleosome position and instead plays a greater role in defining the nucleosome free regions than in defining the locations of nucleosomes (Zhang et al., 2011).

The differences in the position of the +1 nucleosome relative to the TSS in different species could reflect differences in modes of regulation. In yeast, the position of the +1 nucleosome overlaps with the TSS. For initiation to occur, the +1 nucleosome would have to be moved, either by displacement or remodeling. Transcription regulation in yeast is thought to occur primarily at the level of initiation. In flies and humans, however, the +1 nucleosome is situated downstream of the TSS, and does not block access of the transcription machinery (Jiang and Pugh, 2009). Indeed, in metazoans, a large proportion of the genes have Pol II that has initiated transcription and is in a paused state. Density profiles of RNA Pol II across the gene in *Drosophila* and human cells, show that Pol II is present at higher levels at the 5' ends of genes (Guenther et al., 2007; Muse et al., 2007). This enrichment of Pol II near the TSS is seen on active genes, irrespective of transcription levels. Since ChIP-chip studies cannot distinguish between Pol II that is bound at the initiation complex and actively transcribing Pol II, genome wide experiments mapping the RNA products associated with Pol II were carried out (Core et al., 2008; Nechaev et al., 2010). These studies along with the Pol II ChIP-seq experiments support the conclusion that promoter-proximal peak of Pol II on most genes represents paused Pol II (Nechaev et al., 2010).

Comparison of the distribution of nucleosomes with that of Pol II in *Drosophila* embryos, revealed an interesting correspondence between paused Pol II and the +1 nucleosome. Genes with paused Pol II had a 10 bp downstream shift in position of nucleosome compared to genes without

paused Pol II (Mavrich et al., 2008a). This raises the possibility that on paused genes, Pol II collides with the +1 nucleosome, causing it to shift further downstream. A ChIP on chip for Pol II associated nucleosomes showed a selective enrichment at the +1 position. This enrichment was not observed on genes lacking paused Pol II. This observation indicates that on genes that are paused, Pol II is in contact with the +1 nucleosome. The authors conclude that the +1 nucleosome aids in establishment of the paused state (Mavrich et al., 2008a).

An alternative model stating that the paused Pol II helped define the nucleosome positions was raised by results showing that loss of paused Pol II resulted in increased occupancy of nucleosomes over the promoter (Gilchrist et al., 2010). Genes with a high pausing index (ratio of Pol II at the promoter and in the body of the gene) are overall depleted of nucleosomes, while genes with a low pausing index have higher levels of nucleosomes. Pausing of Pol II is orchestrated by the activity of NELF, so depletion of NELF results in decreased pausing. Treatment of cells with NELF RNAi resulted in nucleosomes moving over and occluding the promoters (Gilchrist et al., 2008). An additional example of loss of Pol II correlating with increased nucleosomes occupancy was observed in the comparison of paused genes in embryos and tissue culture cells. Genes that were paused in embryos, but did not have Pol II bound in tissue culture cells, showed increased nucleosome occupancy in tissue culture cells (Gilchrist et al., 2010). The model proposed by the authors is that in *Drosophila*, the paused Pol II serves to keep the promoters in an open chromatin state. The promoters of *Drosophila* genes are GC rich regions that promote nucleosome assembly. This is in contrast to the AT rich regions in yeast promoters that disfavor nucleosome assemble and thus help keep the promoters open. The authors (Gilchrist et al., 2010) propose a model that in *Drosophila*, instead of the DNA sequence contributing to the nucleosome free region at the promoters, it the presence of Pol II that maintains the open chromatin structure at the promoters. In the absence of paused Pol II the nucleosomes would move towards the nucleosome favored GC rich regions over the promoter.

3-1.4 Test of correlation between pausing and nucleosome positioning

The genome wide results that correlate pausing of Pol II and nucleosome positioning imply a dynamic interplay between the two. To test if there is interplay between the Pol II and the nucleosome, I employed a strategy wherein I varied the location of the Pol II and assessed whether this changed the location of the nucleosomes. To do so, I took advantage of our finding that a mutant form of Pol II (C4) shows an upstream shift in permanganate footprints on paused genes. Pol II isolated from the C4 fly line was previously shown to transcribe at half the rate of wild-type (Chen et al., 1996; Coulter and Greenleaf, 1985). This fly line has a point mutation in the largest subunit of Pol II (Greenleaf et al., 1980). If the location of the +1 nucleosome was dictated by where the Pol II paused, then I predict that the +1 nucleosome would be shifted towards the transcription start site in the C4 mutant. If there were no shift in the nucleosome array, then this would indicate that pausing of Pol II was not dependent on a contact between Pol II and the +1 nucleosome.

I carried out genome wide permanganate footprinting in 12-18 hour old embryos from WT and C4 flies. I used 12-18 hour old embryos, as I did not see a difference in permanganate footprinting between WT and C4 in younger (0-6, 6-12 hour old) embryos. Permanganate footprinting is a more definitive assay for detecting paused Pol II than standard techniques for detection of Pol II such as ChIP. It relies on the oxidation of single stranded thymine in the transcription bubble and hence reports more directly on the location of the active, transcriptionally engaged Pol II (Gilmour and Fan, 2009). I also mapped genome-wide the locations of nucleosomes in wild-type and C4 mutant embryos. My results indicate that a change in the position of paused Pol II does not cause a change in the position of the +1 nucleosome.

3-2 Results

3-2.1 Permanganate footprinting on individual genes detects a change in position of paused Pol II in C4 embryos

Earlier results from our lab had shown that the rate of elongation affected where pausing occurred on the gene. *In vitro* experiments with lowered levels of NTP slowed RNA Pol II, causing it to pause closer to the TSS (Li et al., 2013). A similar shift was observed in permanganate footprinting experiments carried out on salivary glands from the C4 mutant fly line (Li et al., 2013). The C4 fly line expresses a form of Pol II that exhibits a 2-fold reduction in elongation rate *in vitro* (Chen et al., 1996). It was initially characterized as an alpha- amanitin insensitive mutant (Greenleaf et al., 1980).

I planned to perform permanganate footprinting genome-wide but first I had to determine if I could detect a change in the location of the C4 Pol II in embryos and establish that I could do Pol II ChIP's in *Drosophila* embryos. Both of these methods have been widely used on tissue culture cells and salivary glands, but cells lack the C4 mutation and salivary glands provided limited amounts of material. I carried out permanganate footprinting in 12-18 hour embryos from control (yw) and the mutant flies (C4). A change in the position of paused Pol II is observed in C4 embryos on a number of genes shown in Figure 3-1.

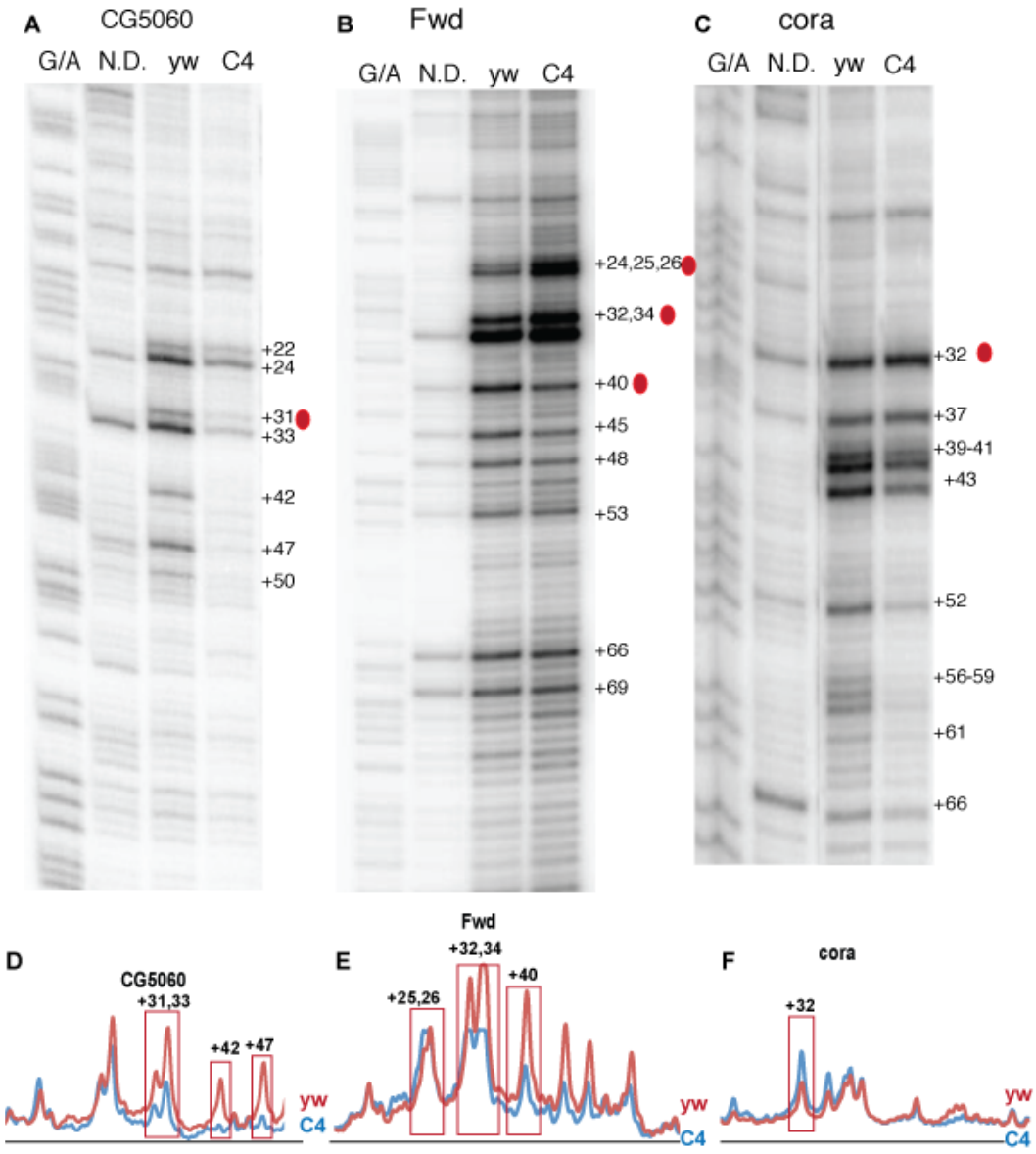


Figure **3-1**: Permanganate footprinting on individual genes show a change in position of paused Pol II in C4 embryos.

(A-C) Permanganate footprinting on individual genes were carried out in 12-18 hour old embryos from yw and C4 flies. Lane 1 is the G/A marker, Lane 2 is naked DNA (N.D.) treated with permanganate. Lanes 3 and 4 are permanganate treated embryos from yw and C4 embryos respectively. **(D-F)** Densitometric traces of the permanganate footprints in Panel A-C respectively. The traces were drawn using ImageQuant software after normalization to correct for differences in background. Detailed image is shown in Appendix A.

Permanganate footprinting maps the hyper reactive thymines in the single stranded DNA associated with the open transcription bubble. The assay thus, detects the transcriptionally engaged, paused Pol II at near base pair resolution. To adapt the technique to generate a genome wide map of paused Pol II, our lab developed the permanganate ChIP-seq procedure that combines permanganate footprinting with a ChIP for Pol II (Li et al., 2013). This technique was developed in tissue culture cells. Adapting the procedure for embryos posed some challenges. The embryos were not efficiently cross-linked with the standard procedures of using heptane to render the embryos permeable to formaldehyde. The older embryos I used in this procedure have a developed cuticle, so I homogenized the embryos in the presence of formaldehyde. Using this approach, I was able to detect Pol II at the promoter of *hsp70* (Figure 3-2D).

Permanganate ChIP-seq technique can detect paused Pol II at nearly the same resolution as a traditional foot-printing experiment. The technique combines permanganate treatment with immunoprecipitation of Pol II to get selective enrichment of transcriptionally engaged Pol II. The embryos are treated with permanganate and sonicated. One of the sequencing adaptors is ligated to the sonicated ends. The permanganate reactive thymines are cleaved with piperidine and the second sequencing adaptor is ligated to the piperidine cleaved ends of the DNA fragments. These ends are then mapped using high throughput sequencing. The majority of the sequencing reads (40%) mapped to thymines and only the reads mapped to thymines were used in subsequent analyses.

Figure 3-2A shows an example of paused Pol II detected by a traditional permanganate footprinting experiment and the sequencing reads associated with the same region in the genome wide maps are shown in Figure 3-2B. The loss of signal downstream from the TSS and increase in intensity of bands closer to TSS (+30,31) is observed in the sequencing reads obtained from the permanganate ChIP-seq experiment in Figure 3-2B. Figure 3-3 is an example where the permanganate ChIP-seq results do not correspond completely with the results from the traditional

footprinting technique. A number of bands (+28-29, +42-44) that are present in the gel image (Figure 3-3A) are not detected by the ChIP-seq method (Figure 3-3B). This discrepancy could be due to different cross-linking efficiencies of different sequences. Early *in vitro* experiments suggest that crosslinking of GC rich regions would be favored compared to AT rich regions, as reaction of formaldehyde with the exocyclic amino group of cytosine, adenine and guanine is favored over the reaction of formaldehyde with the endocyclic imino group of thymine (McGhee and von Hippel, 1975a, 1975b). However these studies with carried out with pure nucleic acids and there could be differences in reaction rates of formaldehyde with nucleotides in the presence of proteins.

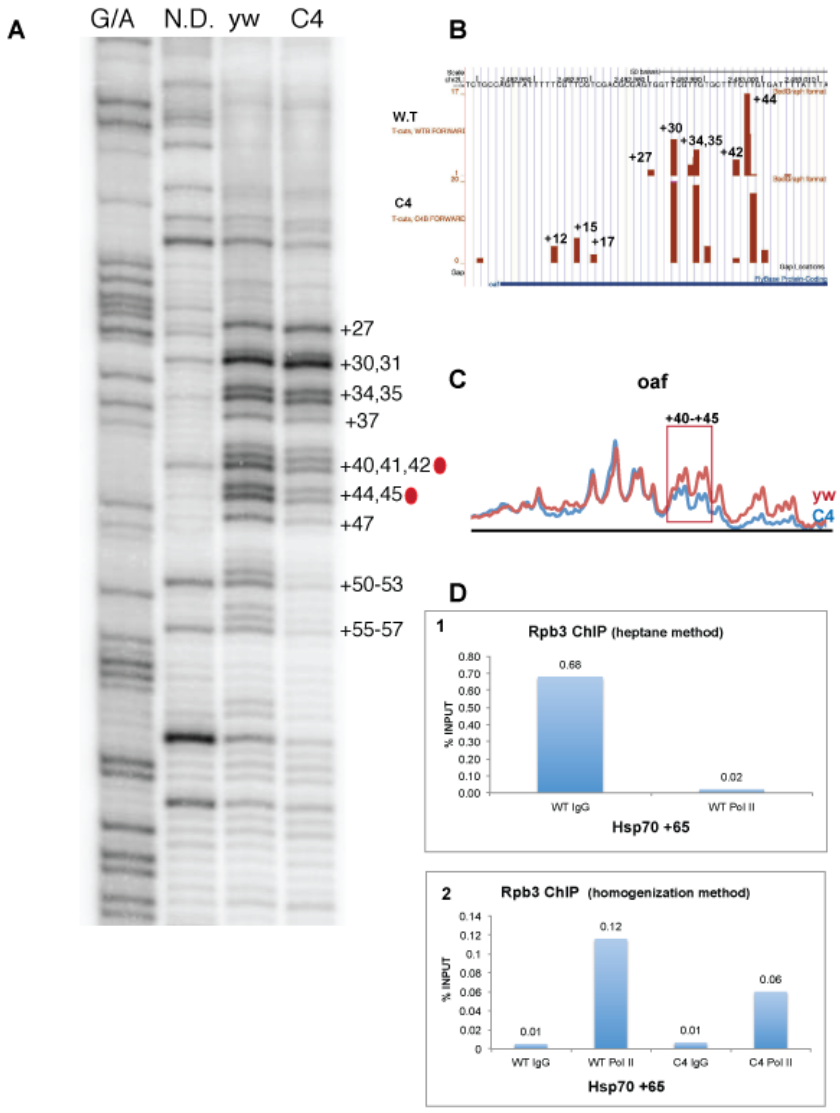


Figure **3-2**: Comparison of genome wide map of hypersensitive thymines with permanganate footprinting on individual genes.

(A) Permanganate footprinting on *oaf* gene were carried out in 12-18 hour old embryos from yw and C4 flies shown in Panel A. Lane 1 is the G/A marker, Lane 2 is naked DNA (N.D.) treated with permanganate. Lane 3 and 4 are permanganate treated embryos. Panel A footprints on the *oaf* gene as detected by LM-PCR. **(B)** UCSC genome browser screenshot showing footprint on oaf gene from the genome wide permanganate ChIP-seq data. **(C)** Densitometric traces of permanganate footprints in yw and C4 embryos shown in (A) after normalization to correct for differences in background. Detailed image is shown in Appendix A. **(D)** Traditional Pol II ChIP experiment with permanganate treated and crosslinked embryos generated using heptane while crosslinking or homogenization during crosslinking of embryos.

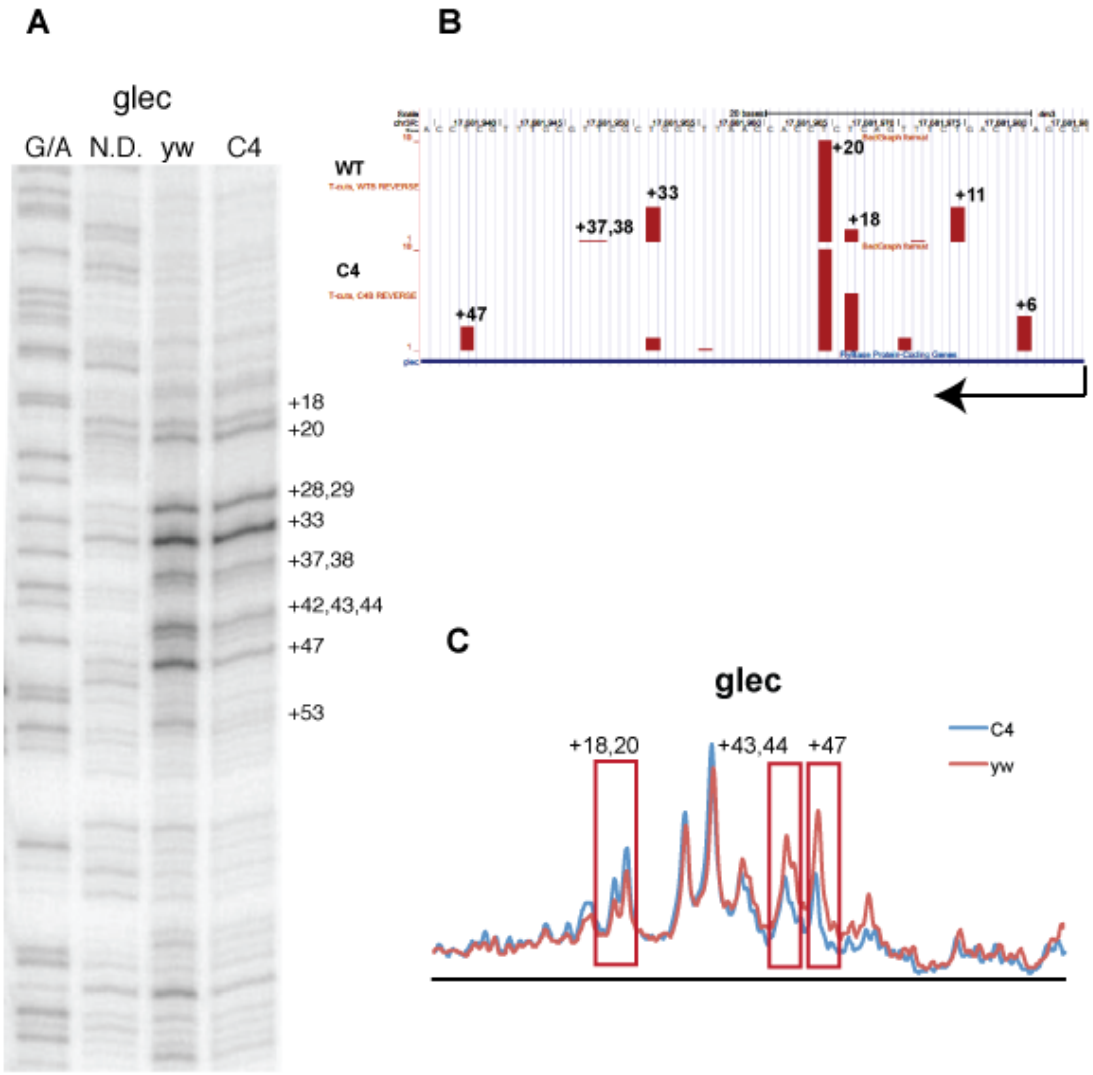


Figure **3-3**: Comparison of genome wide map of hypersensitive thymines with permanganate footprinting on individual genes.

(A) Permanganate footprinting on *glec* gene were carried out in 12-18 hour old embryos from yw and C4 flies. Lane 1 is the G/A marker, Lane 2 is naked DNA (N.D.) treated with permanganate. Lane 3 and 4 are permanganate treated embryos. **(B)** UCSC genome browser screenshot showing footprint on *glec* gene from the genome wide permanganate ChIP-seq data. **(C)** Densitometric traces of permanganate footprints in yw and C4 embryos shown in (A) after normalization to correct for differences in background. Detailed image is shown in Appendix A

3-2.2 Genome wide maps of permanganate sensitive thymines detect changes in positions of paused Pol II in WT and C4 embryos

Mapping the reactive T's on the genes relative to the TSS recapitulates the position of paused Pol II as observed by other experiments, including ChIP for Pol II (Muse et al., 2007). Reads map to the promoter proximal region of genes, about 20-50 nucleotides downstream of TSS (Figure 3-4). The heat map displays the reads mapped relative to the TSS. Each row corresponds to a gene and the genes are ordered by the sum of their reads in the region displayed (-50 to +100). I detect paused Pol II in similar regions in WT and C4 (Figure 3-4).

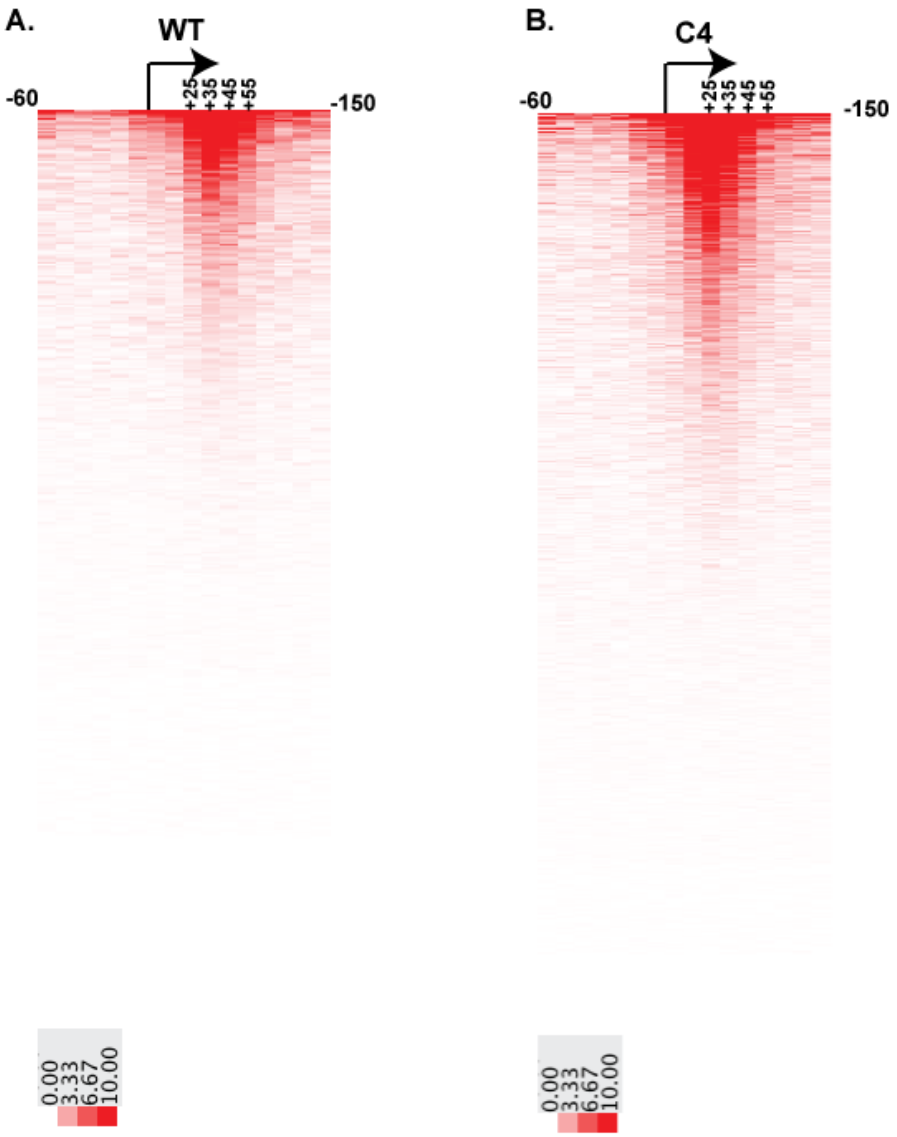


Figure **3-4**: Genome wide permanganate foot-printing in WT and C4 embryos.

Reads corresponding to reactive T's were mapped in 10 base pair bins and aligned to the TSS. Genes without neighboring genes in a 300 bp region were selected (13950 genes out of 16411protein coding genes). **(A)** Heat map displaying all of the reads in -50 to +100 region in WT embryos and ranked according to sum of reads in -50 to +100 region. Each row represents a gene. The reads were normalized against the sum of reads in the selected region for all of the genes in the heat map. **(B)** Heat map displaying all of the reads in -50 to +100 region in C4 embryos and displayed according to order of genes in panel A. The reads in C4 embryos were normalized against the total number of reads in WT to correct for differences in total number of sequencing reads in each data set.

In order to identify genes that show a difference in the position of paused Pol II, sequencing reads corresponding to reactive T's were mapped in 10 base pair bins. Reads in each bin were normalized against total reads in the region corresponding to 200 base pairs upstream and downstream of the TSS of each gene. Normalization was carried out so that the resulting maps reported on differences in positions of paused Pol II on a particular gene in wild-type and C4 samples and not relative intensities. Normalized reads in WT were subtracted from corresponding regions of the same genes in C4. These differences were subjected to K-means clustering and divided into groups, using the Cluster 3.0 program. The difference in positions of reactive T's is displayed as a heat map where the intensity of the color represents the value of the difference (Figure 3-5A). Red indicates a positive value and green, a negative value. The genes are grouped together on the basis of the shift in position of paused Pol II in the C4 embryos. Group 1 shows the genes with the most definitive change in the positions of reactive T's in C4, where the red stripe shows an increase in reactivity at that position in C4 and the green stripe represents decreased reactivity. This group represents the genes where the position of Pol II is shifted closer to the TSS in C4 embryos.

The panel on the right (Fig 3-5B) shows individual composite plots for T reactivity in WT and C4. The composite plots were generated by summing the reads in each 10 base pair bin for all the genes within a particular group. The sum of reads was divided by the number of genes in each group to correct for the different number of genes in each group. In Group 1, Pol II in C4 is shifted 20 bp upstream compared to WT. Group 5 is comprised of genes, which do not show a change in paused Pol II in C4 embryos.

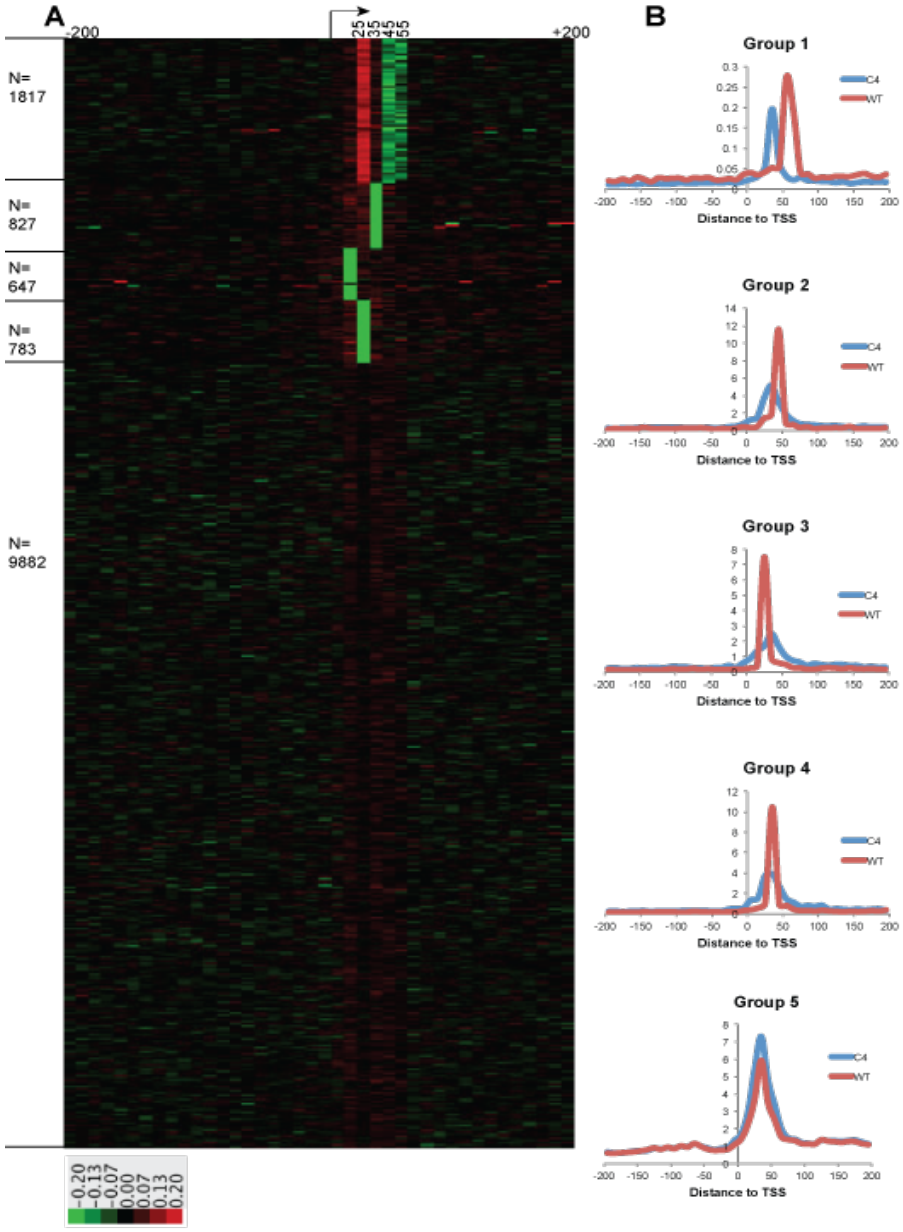


Figure **3-5**: Genome wide map of permanganate footprints in C4 and control (WT) shows groups of genes with a distinct shift in position of paused Pol II.

The reads in each data set were binned in 10 base pair bins. The reads in each bin in a region of -200 to +200 were normalized against the sum of reads in that region for each gene. Differences in normalized reads from C4 and WT were subjected to K-means clustering using the Cluster 3.0 program. N indicates number of genes in each group. Each row represents a gene aligned to its TSS. **(A)** The difference in permanganate reactivity between C4 and WT is represented in the heat map. The intensity of the color is a measure of the difference in the signal

from C4 and WT. **(B)** Composite plots derived from raw tags, binned in 10 bp bins, showing the position of the paused Pol II in C4 and WT, within each. The x-axis shows the distance from TSS and the y-axis is number of reads/number of genes in each cluster. The composite plot showing the nucleosome dyad positions on Groups 1-5 is shown in Appendix B.

The heat map reveals information about the location of Pol II but not the level of Pol II. In contrast, the composite plot shows the position and level of paused Pol II on the different groups of genes that show a shift in C4 embryos. The height of the peak represents the average number of reads in that region, within the group. The average reads in Group 1 were significantly lower than in the other groups, suggesting that these genes on average had at least 10-fold less Pol II paused in the promoter proximal region than the other groups. This raised a serious concern. Embryos consist of many cell types, and it was possible that many of the genes in Group 1 might be harboring a paused Pol II in a small number of cells. If this was the case, the MNase cutting patterns could be detecting the nucleosome organization in cells on genes that were largely devoid of paused Pol II. Consequently, the permanganate patterns would be unrelated to the MNase cutting patterns. To overcome this possibility, I filtered out the genes that had levels of permanganate reactivity falling below a value of 3-fold over the reactivity detected in intergenic regions. Applying this filter identified a set of genes that had similar levels of Pol II. Clustering the differences in position of Pol II in C4 and WT, as described earlier, detected groups of genes with a distinct shift in permanganate footprints (Figure 3-6). These groups of genes showing a change in the position of Pol II provided the basis for further analyses of corresponding changes in nucleosome positions.

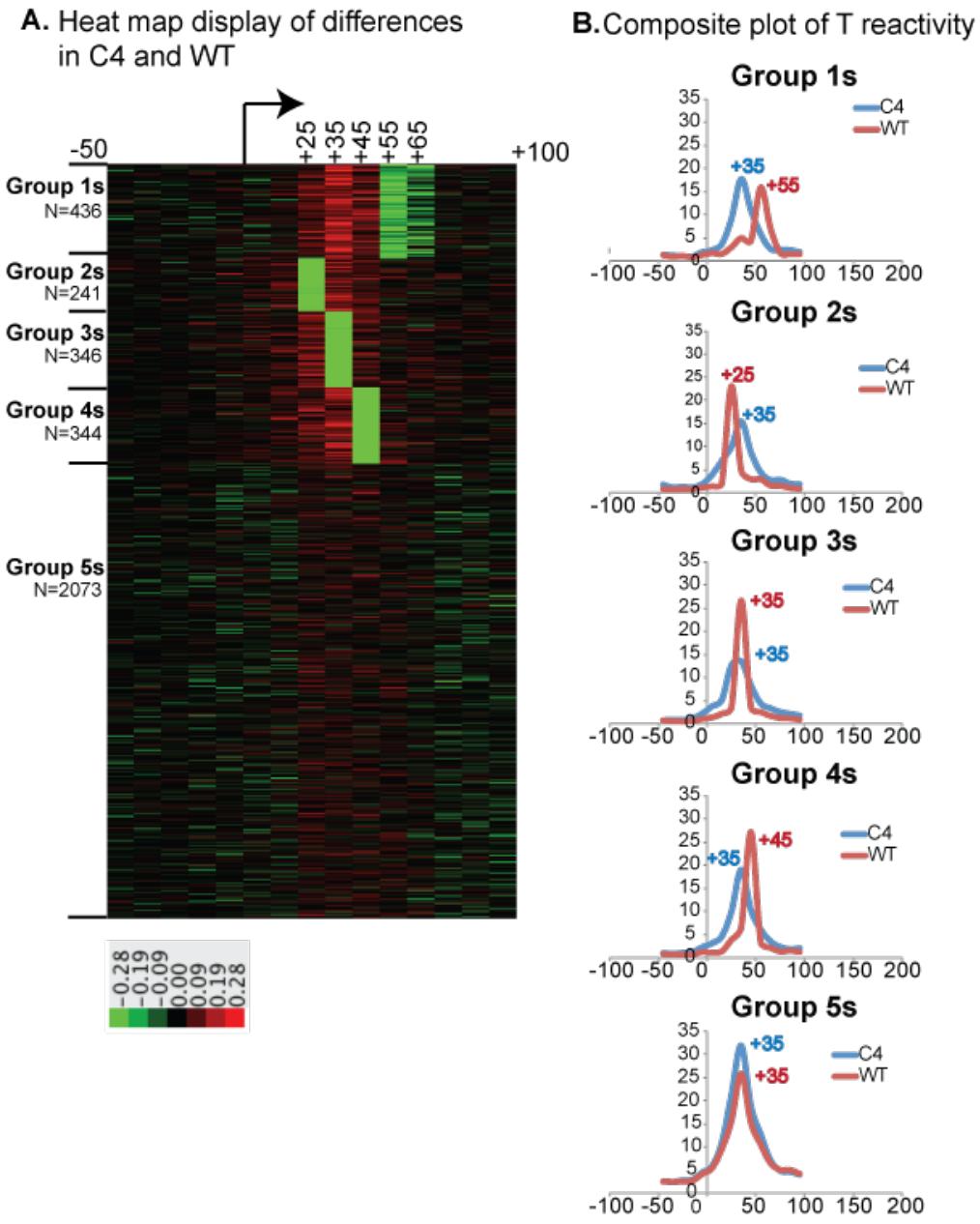


Figure **3-6**: Genome wide map of permanganate footprints on genes with significant levels of Pol II (3 fold over intergenic) in C4 and WT.

Genes with significant levels of Pol II above background (3-fold over intergenic) were selected for this analysis (3440 genes). The difference in normalized reads in the region from -50 to +100 base pairs of the TSS in C4 (mutant) and WT (control) was clustered using Cluster 3.0 program. **(A)** Heat map depicting the difference in signal from C4 and WT. N indicates the

number of genes in each group and the letter “s” indicates the groups of genes with significant levels of Pol II (to distinguish them from the groups shown in Figure 3-5). Red color indicates gain of signal in C4 over WT and green color indicates loss of signal in C4 compared to WT. **(B)** Composite plots showing the position of Pol II in C4 and WT. The x-axis indicates distance from TSS and the y-axis indicates the sum of the number of reads/number of genes in each group.

3-2.3 Shift in paused Pol II does not change position of nucleosomes

Recent results show that loss of paused Pol II causes nucleosomes to move over the promoter towards their intrinsically preferred location, thus occluding the promoter from the transcriptional machinery (Gilchrist et al., 2010). This model suggests that a change in position of paused Pol II would result in a change in nucleosome position. To test this model, I mapped nucleosomes across the genome in WT and C4 embryos. I treated nuclei from 12-18 hour old embryos with MNase and digested the chromatin to mononucleosomal sized DNA. I then ligated sequencing adaptors to generate a library of DNA fragments, which was sequenced on the Illumina sequencing machine. Figure 3-7 is a composite plot of nucleosomal arrays on the groups of genes that showed a shift in paused Pol II in Figure 3-6. The plot uses the center of the nucleosome to show the positioned nucleosomes. The center is calculated by shifting the nucleosome tags corresponding to upstream and downstream edges of nucleosome (relative to the TSS) by 73 base pairs in the appropriate direction. This shift is based on the assumption that the size of the mononucleosomal particle is 147 base pairs and MNase digests the upstream and downstream edges of the nucleosome symmetrically. There is a positioned array of nucleosomes on all 5 groups of genes (Figure 3-7). All 5 groups of genes also have paused Pol II (Figure 3-6B). This is in agreement with results that show Pol II bound genes have an ordered array of nucleosomes while genes not bound by Pol II show lower levels of positioned nucleosomes (Gilchrist and Adelman, 2012).

An upstream shift in the position of paused Pol II did not cause a shift in nucleosome positions towards the TSS, even on the group of genes that show a 20 bp shift in location of paused Pol II (Group 1s from Figure 3-6). If we consider a simple model where Pol II serves as a bookend to the nucleosome array thus defining the array's location on paused genes, then the 20 bp shift in location of paused Pol II should result in a shift of nucleosomes in the direction of the

shift in Pol II. However, I did not observe a shift in nucleosome position on genes with a shifted location of paused Pol II. My results indicate that location of Pol II is not the sole determinant of nucleosome positions.

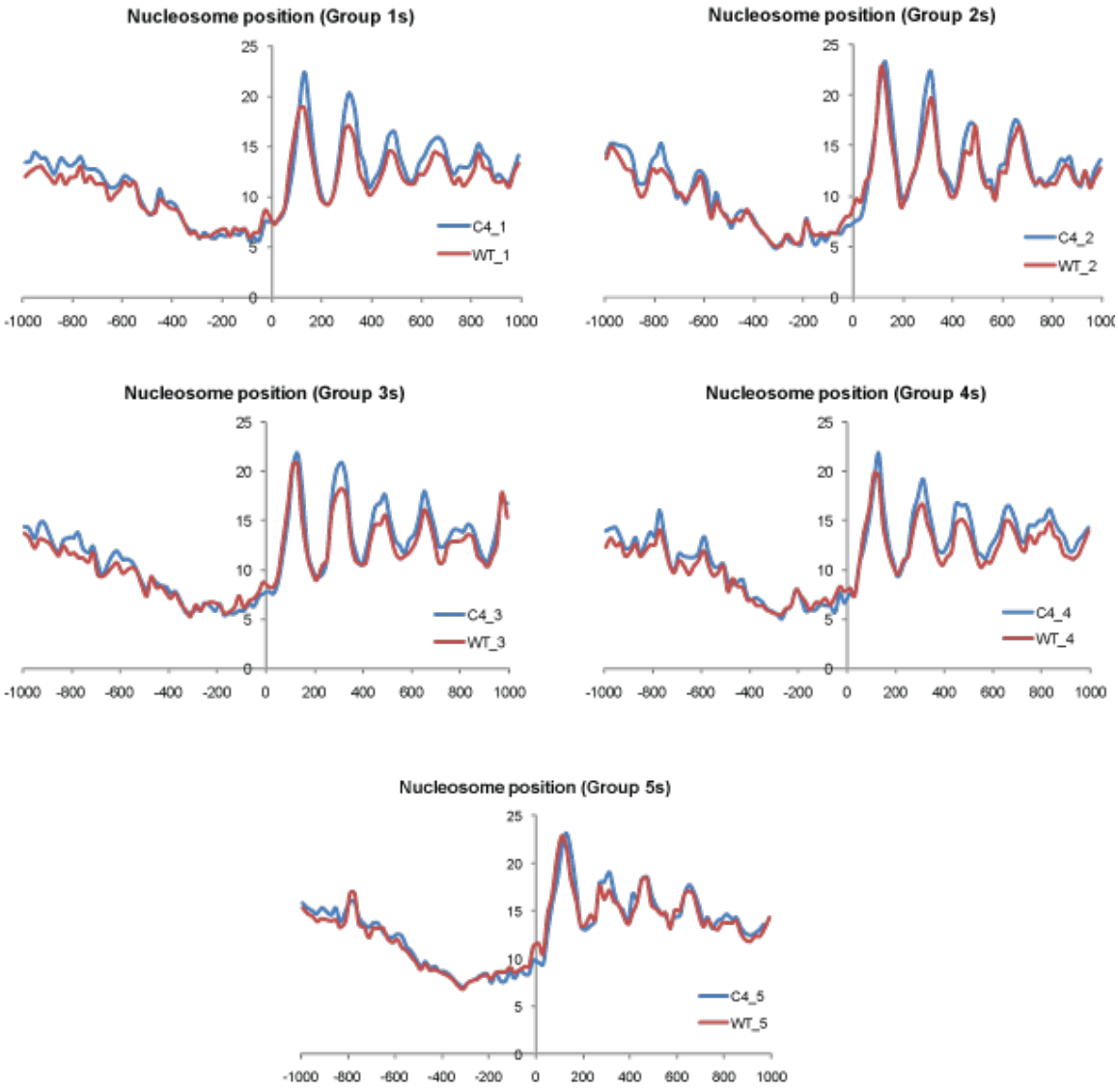


Figure **3-7**: Genes that show a shift in position of paused Pol II do not show a shift in position of nucleosomes (dyad positions on C4 and WT embryos).

Composite plot of nucleosome dyad positions on individual groups as defined in Figure 3-6. The reads are mapped in 20 basepair bins and aligned to the TSS. The x-axis is the distance from TSS and the y-axis is number of reads/number of genes in each group.

3-2.4 Shift in paused Pol II changes the sensitivity of individual borders of nucleosomes to MNase

Several *in vitro* and *in vivo* experiments show that the nucleosome acts as a physical barrier to Pol II elongation. Genome wide studies comparing the position of nucleosome with that of paused Pol II on a number of genes indicated that in *Drosophila,* the 5' edge of the +1 nucleosome is located approximately 10 base pairs downstream from the front edge of paused Pol II (Mavrich et al., 2008a). ChIP-chip of mononucleosomal-sized DNA associated with Pol II suggested that paused Pol II is in contact with the upstream border of the +1 nucleosome. The conventional method to show the nucleosome organization depicts the dyad position. This method however presumes that the digestion on individual borders is symmetrical and would be inaccurate if there were differences in the sensitivity of individual borders to MNase. If contact between Pol II and +1 nucleosome is changed, this could cause differences in the MNase sensitivity of the upstream border of the +1 nucleosome. To address this possibility, I analyzed separately the upstream and downstream borders of the +1 nucleosome on the groups of genes that differ in the location of paused Pol II as defined in Figure 3-6. The results are displayed in Figure 3-8.

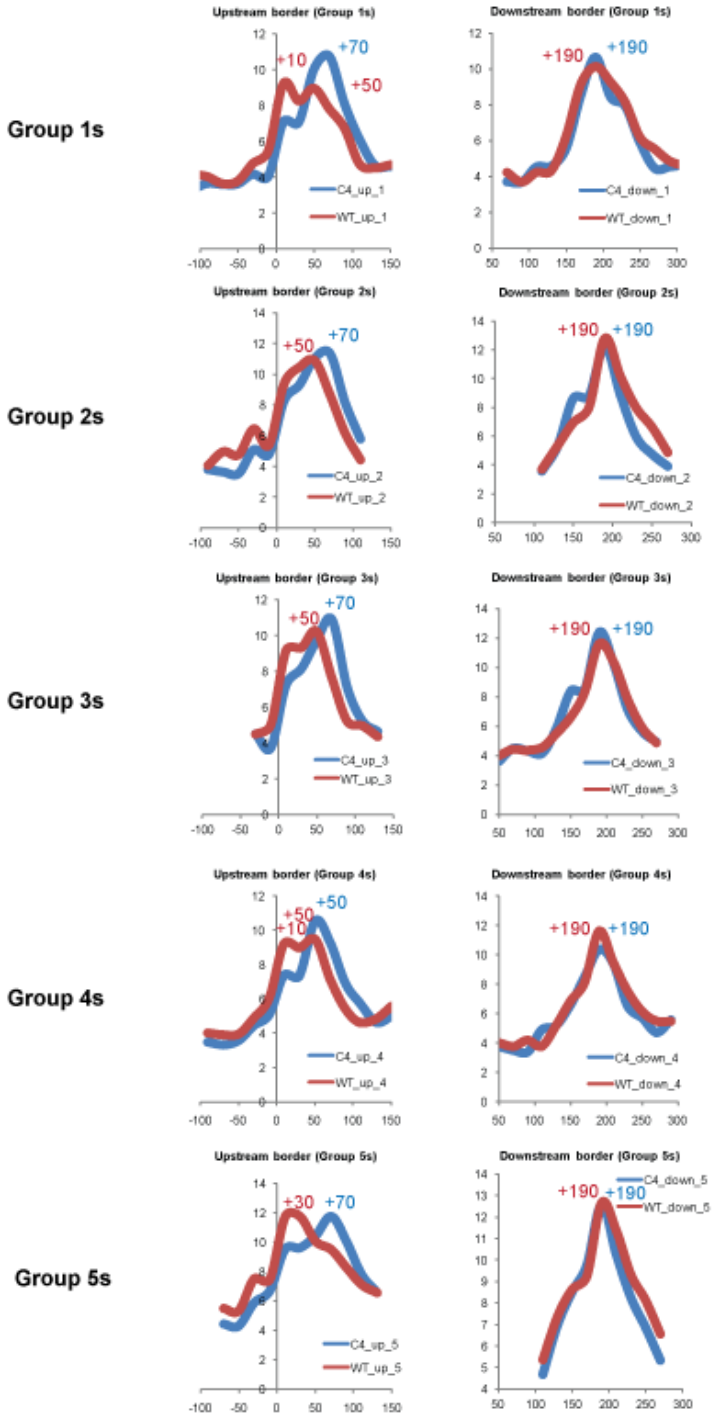


Figure **3-8**: - Nucleosome borders on genes with groups, which show a shift in paused Pol II in C4.

Upstream and downstream borders of +1 nucleosome on genes within each group as defined by the difference in location of paused Pol II in C4 and WT (as shown in Figure 3-6)

Data is derived from nucleosome maps displayed in Figure 3-7. The binned reads corresponding to the downstream border in C4 were lower than in WT. The reason is unclear, as I do not observe any difference in the upstream border or on the reads that have been shifted to generate the presumed dyad. Hence, I multiplied them by 1.8 to compensate for the difference. This does not change the conclusion regarding the position of the border as it changes the peak height not the location. All subsequent analyses with the downstream borders have this correction.

The upstream border of +1 nucleosome in Group 1s is shifted 20 base pairs farther downstream from the TSS in the C4 embryos than in WT embryos. In contrast, the downstream border is the same for the C4 and WT embryos. This argues against the possibility that the difference in upstream border is due to over-digestion by MNase in C4 embryos. To address the possibility that the differences in the upstream border of +1 nucleosome is due to experimental variation and is unrelated to the presence of paused Pol II, I asked if the nucleosome borders in WT and C4 differ on genes with and without paused Pol II (Figure 3-9). I ranked all genes that lacked a neighboring gene within 300 base pairs on the sum of T reactivity in the +15 to +65 region and divided them into 4 groups. Group 1 consists of genes with levels of T reactivity equal to or below the 25th percentile and Group 4 consists of genes with levels of T reactivity equal to or above the 75th percentile. I display the composite plot of the upstream and downstream borders of the nucleosomes in the two groups in Figure 3-9. The shift in the upstream border is only observed on genes with the high levels of paused Pol II and is absent on the genes with no to low levels of paused Pol II. The shift in the upstream border is hence not an artifact due to differences in MNase digestion, as that would have resulted in differences observed on +1 nucleosomes on all genes and not preferentially on genes with paused Pol II.

At present, it is difficult to explain the differences in the MNase digestion of the upstream border because it is observed for all 5 groups irrespective of how the Pol II behaves in these groups. Indeed, group 5s argues that the difference in cutting is not dependent on Pol II shifting its location since the C4 and WT Pol II remain in the same location for this group. For this reason I based my conclusion about the relationship between the location of the Pol II and the first nucleosome on the location of the downstream border in the two types of embryos. The coincidental locations of the downstream borders of +1 nucleosome in C4 and WT indicate that the +1 nucleosome is not shifted in C4.

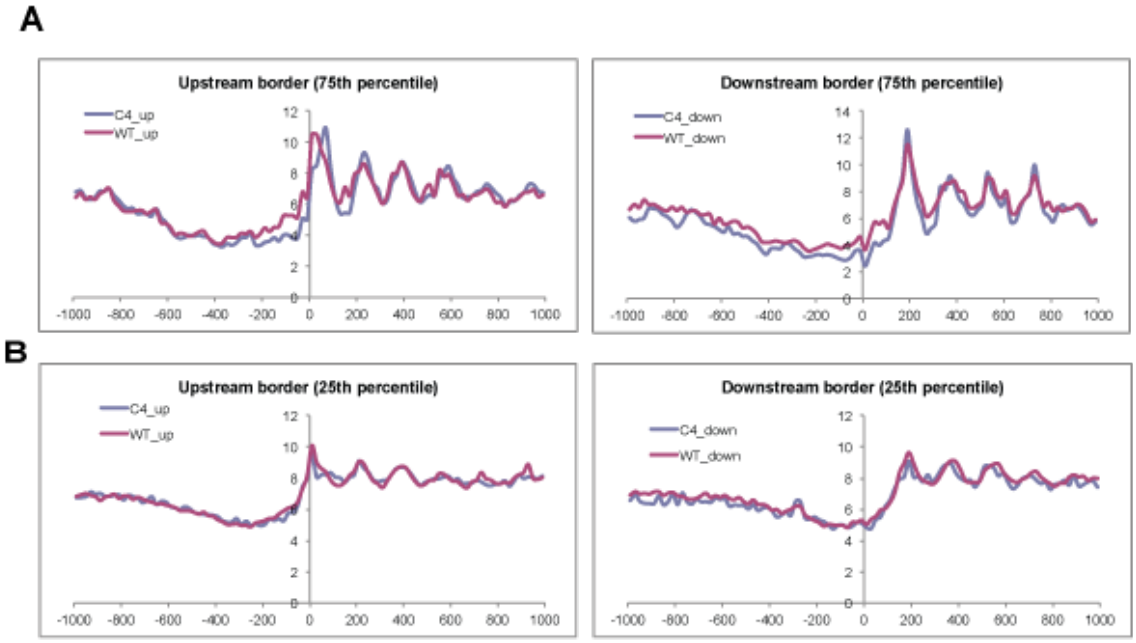


Figure **3-9**: Shift in nucleosome border is observed on genes with paused Pol II and not on genes without paused Pol II.

Composite plot of upstream and downstream borders on nucleosomes in WT and C4 embryos grouped into quartiles based on the sum of reads in +15 to +65 region, where Pol II is paused. **(A)** Group of genes with high levels of paused Pol II (75% percentile). **(B)** Group of genes with low levels of paused Pol II (25% percentile).

3-2.5 Positioned nucleosome downstream of paused Pol II in WT embryos

Another way to investigate the relationship between pausing and the chromatin structure is to group the genes in wild-type embryos according to the location of the paused Pol II and determine if these locations have any relationship to the location of the +1 nucleosome. Using genes with permanganate reactivity 3-fold above background, I clustered the genes in WT based on the permanganate reactivity in a region 50 nucleotides upstream to 100 nucleotides downstream of the TSS. Genes were grouped on the basis of the position of paused Pol II in the promoter proximal region in WT embryos.

Considering the leading edge of Pol II to be 16 bp ahead of the transcription bubble, the edge of Pol II in the 4 groups (A-D) defined in Figure 3-10, would be +41, +51, +61 and +71 respectively (Gnatt et al., 2001; Mavrich et al., 2008). The downstream border of the +1 nucleosome is at 190 nucleotides downstream from the TSS on Groups A, B and C. This is in agreement with the location of downstream border of +1 nucleosome mapped by paired-end sequencing in tissue culture cells (Li and Gilmour, 2013). Considering the size of the DNA fragment to be 147 basepairs, the upstream edge of the +1 nucleosome on genes in Groups A, B, C would be at +43. If we consider the front edge of the Pol II on genes in Group A to be at +41, this would put the front edge of Pol II in Group A in close proximity with upstream border of +1 nucleosome. In Groups B and C, the position of front edge of paused Pol II that is calculated to be at +51 and +61 respectively is beyond the upstream border of +1 nucleosome on these genes.

While the position of the +1 nucleosome appears to be unchanged in Groups A, B and C, the position of +1 nucleosome in Group D is shifted 20 bp further downstream. Pol II is also paused further downstream in Group D, at +55. Thus, it would appear that there is a correlation between change in position of paused Pol II and change in position of +1 nucleosome. While the position of +1 nucleosome appears to be unrelated to position of paused Pol II on the majority of

genes, it is possible that on genes where Pol II is paused further away from the TSS, the +1 nucleosome is involved in pausing Pol II by acting as a barrier. The correlation between the position of the +1 nucleosome and the position of paused Pol II is in agreement with early nucleosome maps in *Drosophila* that showed a positioned nucleosome downstream of paused Pol II (Mavrich et al., 2008a). Recent experiments that detect transcriptionally engaged Pol II by mapping the 3' ends of nascent transcripts on a genome wide scale have identified groups of promoters that are classified as promoter proximal or promoter distal based on position of paused Pol II (Kwak et al., 2013). The 3' ends of nascent transcripts on distally paused genes map to a position within the +1 nucleosome suggesting that on these genes the +1 nucleosome could play a direct role in pausing Pol II. However, on the majority of genes that have Pol II paused closer to the TSS and classified as proximal promoters, pausing appears to be independent of nucleosome positioning (Kwak et al., 2013). These results are in agreement with my observation that the position of +1 nucleosome is unchanged in Groups A, B and C even though the position of paused Pol II is changed. However in Group D where Pol II is paused further away from the TSS, the position of +1 nucleosome is also changed.

The +1 nucleosome could aid in stabilizing the paused Pol II. However, since position of paused Pol II can be changed without a concomitant change in +1 nucleosome, the nucleosome is not the sole determining factor of where the Pol II pauses.

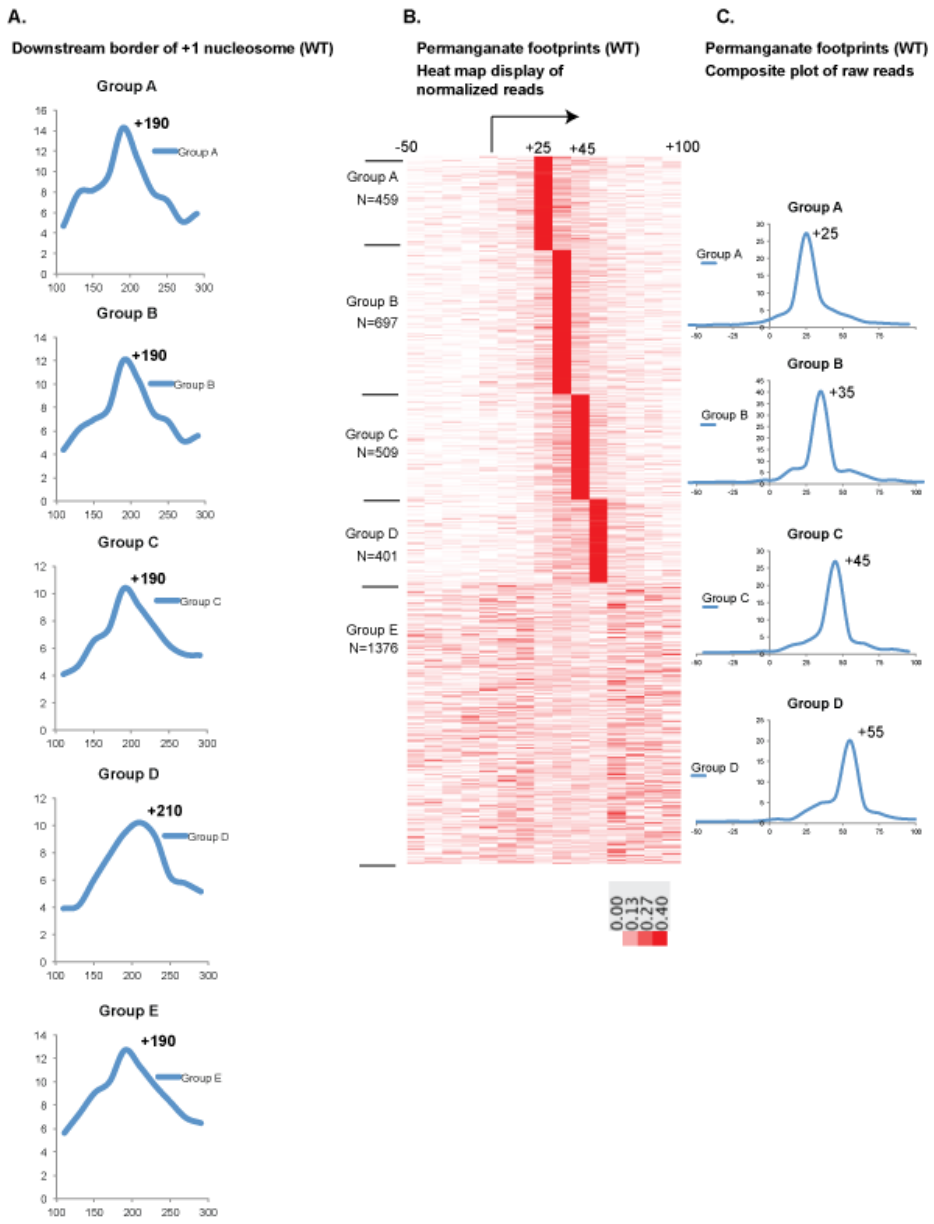


Figure **3-10**: Positioned nucleosome downstream of paused Pol II in WT embryos.

Genes in WT embryos were subjected to K-means clustering and grouped based on the location of paused Pol II. **(A)** Composite plot of downstream border on +1 nucleosome in groups ordered by location of paused Pol II. **(B)** Heat map of normalized reads aligned to TSS. Reads have been normalized against the sum of reads in -50 to +100 region for each gene. **(C)** Composite plot of the T reactivities in different groups defined in Panel B. The reads in the composite plot have been normalized against the number of genes in each group.

3-2.6 Promoters associated with two DNA binding factors, GAF and M1BP, show differences in paused Pol II and nucleosome distributions in WT embryos

GAF is associated with genes with high levels of paused Pol II (Gilchrist et al., 2010; Lee et al., 2008). Permanganate ChIP-seq experiments carried out in our laboratory on tissue culture cells have also demonstrated that genes with high levels of paused Pol II were associated with GAF (Li and Gilmour, 2013). These experiments also showed that presence of a second DNA binding factor M1BP, also correlated with presence of paused Pol II (Li and Gilmour, 2013). M1BP is thought to function as a transcriptional activator and functions mainly on active housekeeping genes. Genes associated with GAGA factor have low levels of nucleosomes that are not well positioned, while M1BP bound genes have high levels of well-positioned nucleosomes in tissue culture cells (Gilchrist et al., 2010; Li and Gilmour, 2013).

In order to examine the pattern of nucleosome positioning on genes bound by GAGA factor or M1BP in embryos, I divided the genes that have 3 fold enrichment of Pol II over intergenic regions into three groups; GAGA factor bound genes; M1BP bound gene and genes that are not bound by either GAGA factor or M1BP. The list of genes bound by GAGA factor or M1BP was obtained from data shared by a former graduate student (Li and Gilmour, 2013; Li et al., 2013). Figure 3-11 is a display of the upstream and downstream borders on nucleosomes on these groups. The nucleosome patterns on GAGA factor bound or M1BP bound genes in WT embryos shown in Figure 3-11 are similar to nucleosome patterns observed in S2 cells (Li and Gilmour, 2013; Li et al., 2013). Genes with M1BP have a well-positioned array of nucleosomes as depicted by the clear oscillating pattern observed on this group of genes. On the other hand, genes that lack M1BP and GAF have positioned nucleosomes but they are more delocalized compared to M1BP bound genes, since the peak heights are lower and the depth of the valleys in the plots are shallower and the peaks are less defined. GAF genes have a low nucleosome occupancy that disordered relative to the TSS.

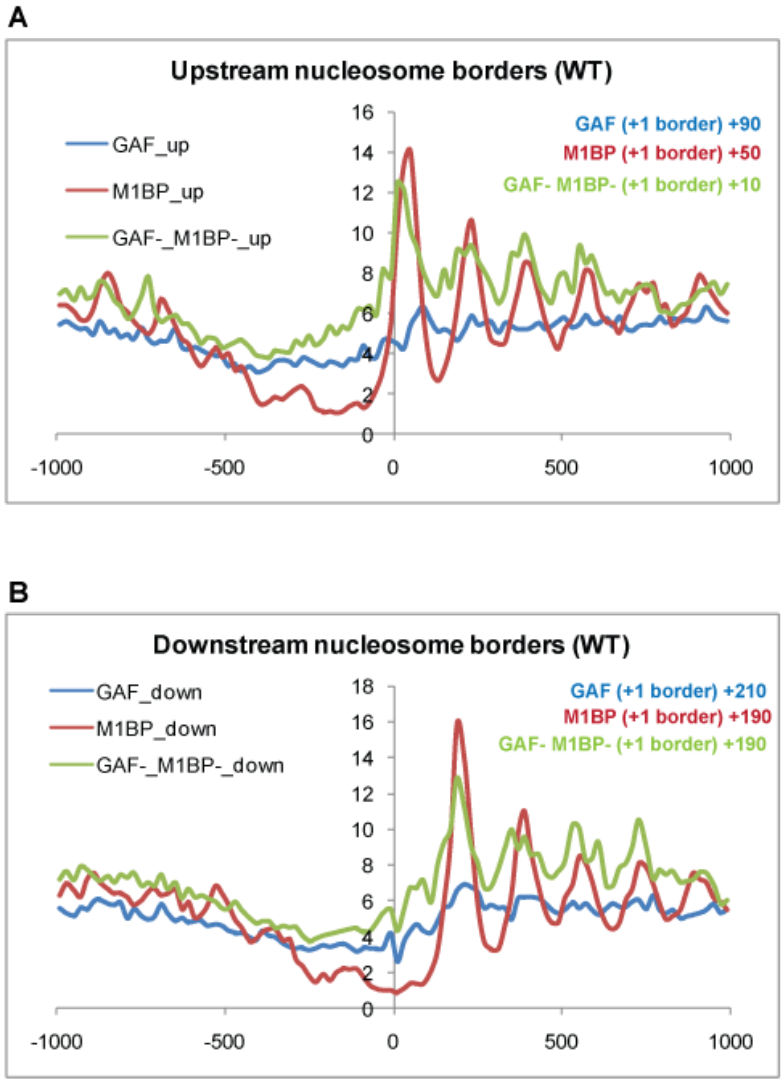


Figure **3-11**: Nucleosome distributions on different groups of promoters in WT.

Composite plot of individual borders of nucleosomes on genes associated with GAGA factor (number of genes = 716) or M1BP (number of genes = 615) or genes that do not associate with either (number of genes = 2023) in WT embryos. Plots were derived from raw tags mapped in 20 base pair bins. They have been normalized for the number of genes in each group. Panel A shows upstream border of nucleosomes and Panel B shows the downstream border. The numbers indicate the position of the peak of +1 nucleosome borders.

To examine if there is a correlation between positions of paused Pol II and nucleosomes in the GAGA factor bound genes or M1BP bound genes or genes that bound by neither of these factors, I determined the distribution of GAGA factor or M1BP bound promoters and promoters that are not bound by either within the group of genes that showed maximum shift in permanganate footprints in C4 embryos compared to WT (Group 1s, Figure 3-6). As a control I also determined distributions of the differently bound promoters (bound by GAGA factor or M1BP or by neither) within the group of genes that showed no shift in permanganate footprint in C4 embryos compared to WT. Next, I compared the downstream border of +1 nucleosome in WT and C4 embryos in the subdivided groups described above (Figure 3-12).

While it is difficult to precisely define the position of the peak corresponding to the border of +1 nucleosomes in GAGA factor bound genes as the nucleosomes are not well-positioned in this group as seen in Figure 3-11, the borders of the +1 nucleosome on GAGA factor bound genes are observed at similar locations in WT and C4 embryos. The non-uniform peaks corresponding to the downstream borders of +1 nucleosome in WT and C4 embryos in GAGA factor bound genes could be a reflection of the delocalized nature of nucleosomes on this group. The location of downstream border of +1 nucleosome on genes bound by M1BP is coincidental in WT and C4 embryos. The group of genes not bound by GAGA factor or M1BP show a difference in the peaks corresponding to the downstream borders in C4 embryos. The downstream border of +1 nucleosome is shifted further downstream in C4 embryos. The reasons for a downstream shift in position of +1 nucleosome on the group of genes not bound by GAGA factor or M1BP and which show a shift in location of Pol II is unclear.

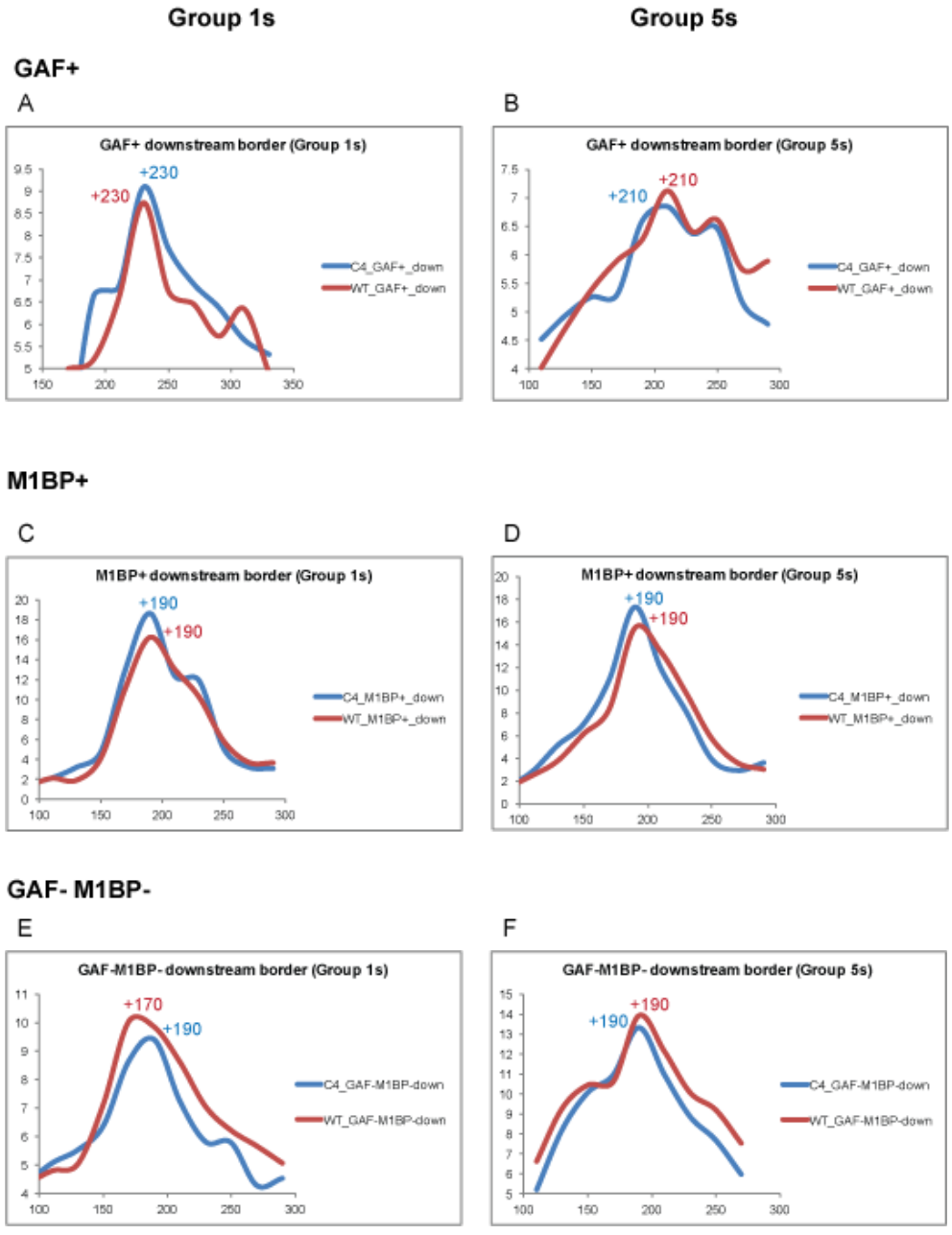


Figure **3-12**: Comparison of downstream border of +1 nucleosome distribution on different promoters in WT and C4.

Composite plot of downstream border of +1 nucleosomes on the group of genes that showed a shift in location of paused Pol II and those that did not (Group 1s and 5s from Figure 3-6). The two groups are further divided into groups of genes associated with GAGA factor or

M1BP or genes that do not associate with either in WT and C4 embryos. Plots were derived from raw tags mapped in 20 base pair bins. They have been normalized for the number of genes in each group. (A-B) Border of nucleosomes on GAGA factor bound genes. Number of genes in Panel A=81 and Panel B=437. (C-D) Borders on M1BP bound genes. Number of genes in Panel C = 91 and Panel D= 339. (E-F) Borders on genes not bound by GAGA factor or M1BP. Number of genes in Panel E = 251 and Panel F = 2239. Panels A, C, E correspond to genes that belong to Group 1s. Panels B, D, F correspond to genes that belong to Group 5s.

3-3 Discussion

Pausing of RNA Pol II in the promoter proximal region is now considered to be a prevalent feature on metazoan genes. Establishment and release from the paused state offer important regulatory steps in transcription on these genes. The *in vitro* experiments demonstrating the presence of a nucleosome in the path of Pol II results in Pol II being paused on the DNA template combined with the *in vivo* results that show transcription by Pol II is accompanied by disassembly of nucleosomes implicate nucleosome organization as an important feature that could affect pausing of Pol II in cells.

3-3.1 Pausing of Pol II in C4 embryos

The difference plot of T-reactive regions in C4 and WT embryos show different groups of genes pause at different locations and a substantial number of genes show a shift in permanganate footprints in C4 compared to WT embryos (Figure 3-6). The genes that show a shift in the location of paused Pol II in C4 embryos do not show any correlation to promoter elements, such as TATA box, DPE, Inr (data shown in Appendix C). There are several factors that could influence the location of pausing of Pol II which include sequence of the template, rate of incorporation of NTPs, rate of translocation and protein factors that function to pause Pol II. *In vitro* experiments have shown that there are intrinsic pause sites on naked DNA suggesting that sequence of the template could influence location of pausing (Kireeva et al., 2005). The stability of the RNA-DNA hybrid in the transcription bubble is also thought to affect pausing (Nechaev et al., 2010). When the RNA-DNA hybrid in the upstream position is more thermodynamically favorable than the hybrid at the 3' end of nascent RNA, reverse translocation occurs and RNA Polymerase backtracks. Backtracking of Pol II can result in a paused or arrested state.

Some models on pausing suggest that the probability of a backtracked pause depends on competition between tendency of an unstable elongation complex to backtrack due to weak RNA-DNA hybrid and the rate of incorporation of NTPs (Bai et al., 2004). The C4 form of RNA Pol II shows higher Km for UTP in elongation assays and this could account partly for the slower rate of elongation in these flies (Coulter and Greenleaf, 1985). The slower rate of elongation in C4 embryos could result in increased tendency of Pol II to backtrack and hence pause closer to the TSS.

However, recent analyses have shown that there is no correlation between regions of thermodynamically stable RNA-DNA hybrids and location of paused Pol II (Li et al., 2013). Additionally depletion of TFIIS, a factor that helps Pol II overcome the backtracked/arrested state does not change the location of permanganate footprints corresponding to paused Pol II (Saikat Ghosh, unpublished results). These results suggest that backtracking of Pol II into a thermodynamically favorable location based on the stability of the RNA-DNA hybrid is not the sole determinant of the location of paused Pol II.

The nucleotide addition cycle is a fundamental step in transcription elongation. The elongation complex oscillates between the pre-translocated state where the newly added nucleotide still occupies the nucleotide addition site (i+1) and the post-translocated state where the nucleotide addition site is vacant and available for the next NTP. Restriction of the oscillation between these two states can cause Pol II to pause. If the rate of translocation is slower than the nucleotide addition cycle, Pol II can pause. The pausing observed in this scenario is independent of the thermodynamic stability of the RNA-DNA hybrid (Bochkareva et al., 2011). The C4 mutation is thought to slow down transcription by affecting the translocation of Pol II (Chen et al., 1996). A combinatory effect of rate of NTP incorporation and the change in the equilibrium of the pre and post-translocated states of Pol II could dictate the location of paused Pol II.

Pausing *in vivo* involves the pausing factors, NELF and DSIF. Depletion of NELF results in loss of paused Pol II (Wu et al., 2005). Results from our lab and others have shown that the DNA binding factor, GAGA factor is also associated with paused genes (Gilchrist et al., 2010; Li and Gilmour, 2013). *In vitro* experiments conducted in our lab studying GAGA factor showed that it interacts with NELF and could serve as a scaffold to allow NELF and DSIF to bind to the elongating Pol II and bring about pausing (Li et al., 2013). It is proposed that the location of paused Pol II is influenced by the competition between the elongation rate of transcription and the rate at which NELF can capture the elongating Pol II (Li et al., 2013). Slower rates of elongation would allow NELF more time to bind to Pol II and result in pausing the complex closer to the TSS. This is observed in the C4 mutant where Pol II pauses closer to the TSS on a number of genes (Figure 3-1, 3-2, 3-4 to 3-6).

The C4 embryos show a broadened footprint of Pol II in comparison to WT (Composite plots in Figure 3-5B and 3-6B). C4 Polymerase is less efficient in transcribing through intrinsic blocks on DNA template (Chen et al., 1996). This could account for the broadened footprint in C4, as it is slower in transcribing through the regions that act as blocks to transcription. This would increase the "dwell time" of C4 Pol II on these regions and increase its tendency to get stuck at these locations while WT Pol II is able to efficiently transcribe through these regions.

3-3.2 Nucleosome organization in C4 and WT embryos

Genome wide maps of Pol II and nucleosomes have implied a causal relationship between the two and suggest two distinct models regarding the relationship between the two. The presence of a highly positioned +1 nucleosome downstream of Pol II and a correspondence between the location of the +1 nucleosome with location of the paused Pol II suggests that the +1 nucleosome acts as a barrier that helps pause the Pol II molecule in the promoter proximal region

(Mavrich et al., 2008). This model is supported by the *in vitro* experiments that show that nucleosomes act as a barrier to the elongating Pol II (Izban & D S Luse, 1991). The second model suggests that paused Pol II influences the chromatin landscape and the presence of paused Pol II maintains the promoter region in an open chromatin conformation by preventing nucleosomes from sliding over to their preferred location on the promoter (Gilchrist et al., 2010). This study attempts to investigate the relationship between paused Pol II and +1 nucleosome by posing the question, what is the effect on nucleosome position when the position of paused Pol II is changed?

My results show that the position of the +1 nucleosome is not altered when the location of paused Pol II is shifted. In the C4 mutant, Pol II is paused closer to the transcription start site while there is no change in nucleosome position, thus no barrier against which the Pol II pauses (Figure 3-7). The shift in paused Pol II did not result in the +1 nucleosome sliding towards the TSS. Paused Pol II thus, does not appear to influence the positioning of the +1 nucleosome. Additional support for the conclusion that position of Pol II and nucleosomes are independent of each other on the majority of genes is gained from the comparisons of +1 nucleosome positions in WT embryos on genes that differed in the location of paused Pol II. This showed that nucleosome position was unchanged even when the position of Pol II was changed (Groups A-C in Figure 3-10) with the exception of Group D. Group D is comprised of genes where Pol II was paused distal to the promoter. In Group D the position of +1 nucleosome is also shifted downstream compared to genes where Pol II is paused proximal to the promoter. Based on the position of paused Pol II and +1 nucleosome, the forward edge of Pol II would appear to be in contact with the +1 nucleosome. The genes in Group D could represent the smaller group of genes where positioned nucleosome has a greater influence on pausing than promoter bound factors such as NELF or GAGA (Kwak et al., 2013; Li et al., 2013).

The results that show that the change in location of paused Pol II is not accompanied by a change in nucleosome positions argues against a role of paused Pol II in determining the location of the +1 nucleosome, and argues that pausing is not merely a consequence of nucleosome architecture. This is consistent with early *in vitro* experiments showing that Pol II can be paused even in the absence of a downstream nucleosome (Benjamin and Gilmour, 1998), and challenge another *in vitro* study showing that it was necessary to reconstitute chromatin on DNA *in vitro* to observe pausing (Brown et al., 1996). Pausing occurs mainly due to the activity of factors such as NELF, DSIF and GAGA factor (Lee et al., 2008). While the nucleosome could play a role in stabilizing the paused Pol II on certain genes where paused Pol II is located at the border of the +1 nucleosome (Figure 3-10), it is not essential for pausing to occur.

3-3.2.1 Analysis of +1 nucleosome borders

The presence of paused Pol II ahead of the +1 nucleosome affects the sensitivity of the upstream and downstream borders of the nucleosome to MNase, differently (Figure 3-9). Hence, I analyzed the individual borders of +1 nucleosome in C4 and WT embryos. This analysis showed that the upstream border of +1 nucleosome in C4 embryos is shifted further away from the TSS as compared to upstream border on WT embryos, while the downstream border is unchanged (Figure 3-8).

Early nucleosome maps showed a difference in position of +1 nucleosome on paused and non –paused genes (Mavrich et al., 2008). The location of paused Pol II in the promoter proximal region differs on different genes. Grouping the genes that showed Pol II paused at similar locations in WT embryos, further refined the early analysis on paused and non-paused genes. As seen with the comparisons with C4, the downstream border of the +1 nucleosome is unchanged on groups that show a difference in position of paused Pol II. This would seem contradictory to

the published data that suggests that position of +1 nucleosome is different on paused genes (Mavrich et al., 2008). The early analyses were based on defining the position of +1 nucleosome based on the dyad position and not the individual borders. My observation that the digestion of the two borders of the +1 nucleosome is different and is affected by the presence of Pol II (Figure 3-9), would argue that defining location of +1 nucleosome solely on basis of the dyad calculated with an underlying assumption regarding the size of the DNA is flawed.

Based on the analysis of the downstream borders of the +1 nucleosome, it is centered approximately 120 bp from TSS in WT embryos. The downstream border of +1 nucleosome tends to be located at +190, which is in agreement with the position of downstream border in S2 cells (Li and Gilmour, 2013). The conventional method of calculating the center places it at 130 bp from the TSS. These positions are different from what was reported for paused genes in *Drosophila* embryos (+145 for paused genes, +135 for non-paused genes). The variations in these numbers could reflect differences in methods of analyses. Analysis of the data from Maverich et al., by a different group, who filtered the genes to include those that were expressed in the 0-12 hour period of embryogenesis, placed the center of the +1 nucleosome at +125 (Rach et al., 2011). Given these differences resulting from determining the position of nucleosomes based on presumed center of the nucleosome, this method is inadequate to define the location of the nucleosome, especially the +1 nucleosome. It would of interest to see if analyzing individual borders of the +1 nucleosome in these published data sets helps to resolve the discrepancies and assess how well the downstream borders coincide.

While the +1 nucleosome does not seem to be instrumental in setting up the paused Pol II, it could contribute to pausing on the distally paused Pol II (Figure 3-10, Group D). Pausing of Pol II in the region closer to the TSS is influenced by the rate at which NELF captures the elongating Pol II (Li et al., 2013). The distally paused genes could represent the group of genes

where Pol II has escaped capture by NELF and the +1 nucleosome could act as a barrier against which Pol II pauses on this group of genes.

3-3.3 Nucleosome positioning

Nucleosome maps in different species reveal that while the general features of nucleosome organization are similar, some of the notable differences in *Drosophila* are the absence of a -1 nucleosome and presence of GC rich sequences in the NFR region, which would promote nucleosome assembly.

Studies on nucleosome positioning have shown that one of the primary determinants of nucleosome positioning is the sequence of DNA. Certain sequences promote nucleosome assembly and certain sequences disfavor nucleosome assembly (Segal et al., 2006). However, comparisons between *in vitro* and *in vivo* nucleosome maps in yeast showed higher correlation for the NFR region than for positioned nucleosomes present over the body of genes. Recapitulation of nucleosomal arrays *in vitro* required the presence of ATP dependent chromatin remodelers (Zhang et al., 2011). This would indicate that trans-factors have a significant role in dictating nucleosome positions.

One of the trans-acting factors that could contribute to nucleosome positioning in *Drosophila* is the chromatin remodeler ISWI. ISWI localizes to regions near the TSS (Sala et al., 2011). Loss of ISWI causes delocalization of the positioned nucleosomes downstream of TSS. Thus, in *Drosophila*, the chromatin remodeler, ISWI could play a role in positioning nucleosomes on a number of genes.

Another factor that could affect nucleosome architecture is RNA Pol II (Schones et al., 2008). In *Drosophila,* genes that were silent in S2 cells but showed the presence of paused Pol II in embryos showed increased occupancy of nucleosomes at the promoter region (Gilchrist et al.,

2010). Thus it has been hypothesized that Pol II influences the nucleosome architecture by maintaining promoters in a nucleosome free state. This hypothesis is supported by the observation that on genes that lose paused Pol II, there is increased occupancy of nucleosomes over the promoter (Gilchrist et al., 2008). The proposed model would predict that a shift in location of paused Pol II towards the TSS would cause the nucleosomes to move towards the preferred nucleosome positioning sequences near the promoter. My results show that a change in the location of paused Pol II does not result in sliding of the nucleosomes towards the TSS suggesting that Pol II does not influence positioning of the +1 nucleosome in *Drosophila*.

3-3.4 Potential limitations

The initial analyses carried out on all genes, irrespective of levels of Pol II, showed that I could detect defined nucleosomal arrays on genes with very low levels of Pol II, which possibly did not have a paused Pol II in all cells (Figure 3-5). MNase digestion on genes that possibly do not have a paused Pol II raised a concern that MNase digestion patterns were unrelated to the permanganate footprints detected on this set of genes. Filtering out genes that had levels of Pol II below 3 fold over intergenic helped ensure that I was analyzing the nucleosome organization on genes that had Pol II paused in the majority of cells. The filtered genes had similar levels of Pol II (Figure 3-6).

A potential limitation of these experiments was that the nucleosome mapping was done in embryos that were not efficiently crosslinked. Crosslinking was carried out as per standard procedures of using heptane to make the embryos permeable to formaldehyde. This method did not result in efficient crosslinking of the older aged embryos used in this experiment. ChIP for H3 from these crosslinked embryos had very low yields, which alerted me to the possibility of inefficient cross-linking in these older embryos. Although the nucleosome maps used in this study

are generated from the inefficiently cross-linked embryos, the composite plot depicting the nucleosome organization is comparable to that observed with crosslinked samples from S2 cells (Figure 3-13).

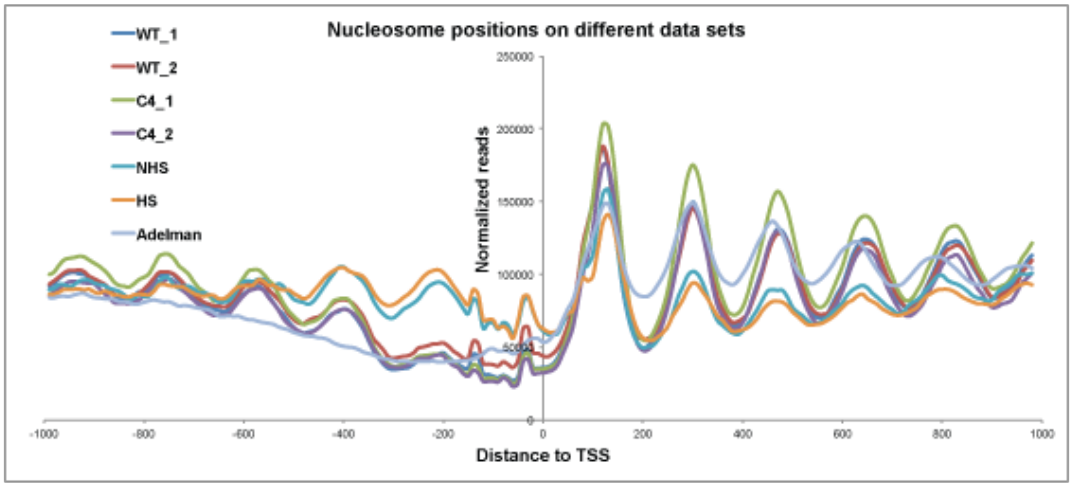


Figure **3-13**: Comparison of nucleosome positions with nucleosome maps derived from cross-linked cells and with samples generated from H3 immunoprecipitated DNA.

Composite plot of nucleosome positions (Peak corresponds to dyad position) from uncrosslinked (WT_1,WT_2,C4_1,C4_2), crosslinked and not immunoprecipitated with H3 antibody (Adelman), crosslinked and precipitated with H3 antibody (NHS, HS). The reads from all genes with no neighbors in 300 bp (13950) were mapped in 20 base pair bins and aligned to the TSS and normalized to correct for differences in total number of reads in each sample. The x-axis is the distance from TSS and the y-axis is normalized reads (against total number of reads in each sample). WT_1,WT_2, C4_1, C4_2, represent biological replicates of WT and C4 respectively. The trace annotated as Adelman corresponds to data obtained from Adelman et al., 2010 (Gilchrist et. al, 2010). The traces annotated as NHS, HS, represent nonheat shocked and heat shocked data and were obtained from tissue culture cell by Bede Portz.

I attempted to generate nucleosome maps from embryos that were crosslinked using an alternative method in which embryos were homogenized with a dounce, while cross-linking. This improved the cross-linking efficiency. Cross-linked chromatin was digested with MNase to mononucleosomal sized DNA. This was further size-selected for mononucleosomal sized DNA prior to and after ligation of sequencing adaptors, in a manner similar to published protocols for nucleosome mapping in S2 cells. However, instead of observing an oscillating pattern of MNase cuts diagnostic of nucleosomes, most of the cuts occurred in the promoter region. The borders mapped to a region spanning the TSS, which is contrary to published data that shows the TSS in *Drosophila* is maintained in a nucleosome free region (Gilchrist et al., 2010). The presence of these borders correlated with the presence of paused Pol II, which led me to conclude that the region spanned by the borders is not a nucleosome and instead the initiation complex at the TSS. The heat map display of the borders in the crosslinked data show faint lines corresponding to the nucleosomes in the body of the gene but the display is dominated by the signal at the TSS (Figure 3-14). The harsher conditions needed for crosslinking the older embryos probably resulted in greater crosslinking of initiation complexes than the nucleosomes.

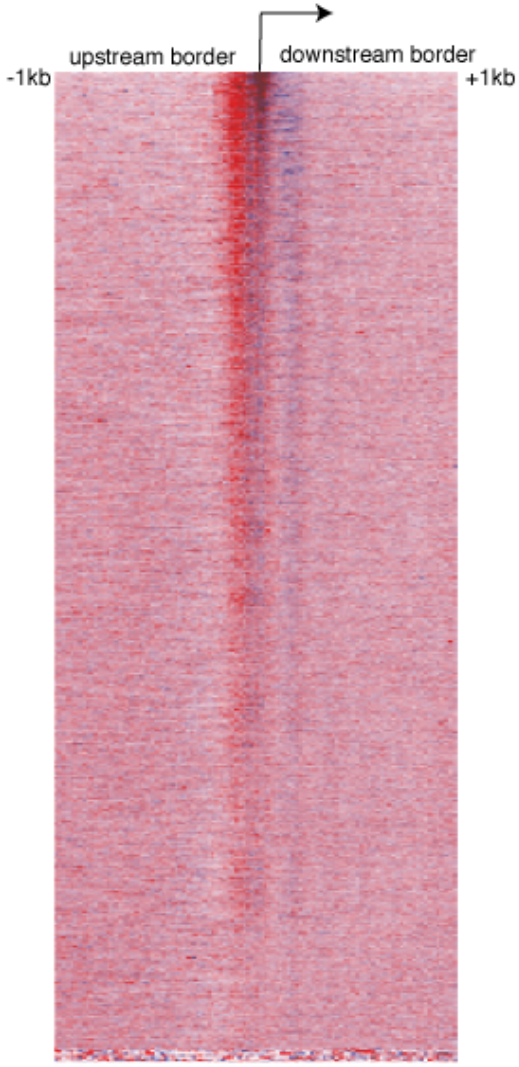


Figure **3-14**: Upstream and downstream borders on mononucleosomal sized fragments from crosslinked embryos.

Borders of 147 base pair fragments isolated after MNase treatment of crosslinked embryos. The chromatin was not immunoprecipitated with any histone antibody. All genes with no neighbors (13950) were ranked on their sum of T-reactivity in +15 to +65 region and the upstream and downstream borders of the mononucleosomal sized fragment on the ranked list of genes is shown in the display. The upstream border is shown in red and the downstream border is shown in blue.

There is precedence for generating nucleosome maps from size-selected mononucleosomal DNA that has not been immunoprecipitated for H3 (Gilchrist et al., 2010; Teves and Henikoff, 2011). Indeed, Teves and Henikoff observed MNase resistant fragments generated from the promoter regions although these tended to be smaller than mononucleosome sized fragments (Teves and Henikoff, 2011). It is possible that my crosslinking conditions stabilized complexes at promoters and resulted in MNase protected fragments that were similar in size to a nucleosome. The nucleosome maps used in this study were from embryos that were not efficiently cross-linked. There is a possibility that the nucleosome positions in uncross-linked samples are not as sharply defined in comparison to cross-linked data, leading to greater fuzziness in the data. However, the composite plots for nucleosome positions generated from cross-linked samples that were immunoprecipitated for H3, cross-linked samples that were not immunoprecipitated for H3 and my uncross-linked samples overlap very well (Figure 3-13). This suggests that the data from the uncross-linked embryos are not fuzzier than data from crosslinked nucleosomes. The biological replicates for the WT and C4 samples also show high overlap.

These nucleosome maps were derived from single end sequencing where only one end of the DNA fragment is sequenced. Paired end sequencing, which sequences both ends of a single DNA fragment, would offer a more definitive assessment of the nucleosome borders. The presence of two peaks on the composite plots of upstream borders and the generally uniform peaks on plots of the downstream border hint at the possibility of two populations of DNA fragments that contribute to defining the +1 nucleosome. This is especially apparent in the border of +1 nucleosome in Group 5s in Figures 3-8 and 3-12. Different populations of fragments could arise due to asymmetric digestion of the upstream border of +1 nucleosome. The presence of proteins, such as Pol II in front of the nucleosome could affect the sensitivity to MNase. *In vitro* experiments have shown that Pol II can transcribe into the nucleosome (Bondarenko et al., 2006) . Recent analyses in our lab have detected the presence of smaller sized DNA fragment localizing

to the +1 position, which reflects the presence of an “altered” nucleosomal particle at that position (Li and Gilmour, 2013). The multiple peaks on the composite plot of the upstream border at the +1 position could reflect the two populations of differently digested nucleosomes. Paired end sequencing would enable us to determine accurately the actual size of the sequenced fragment rather than relying on extrapolation from single end reads that are used to define borders in a population rather than on an individual DNA fragment. This would help to determine if there are differently sized nucleosomes as would be predicted by results that suggest Pol II collides with the +1 nucleosome.

Chapter 4

HDAC3 affects transcription of *hsp70* gene.

4-1 Introduction

Transcription involves distinct steps. On the *hsp70* gene in *Drosophila*, these distinct steps include establishment of a paused Pol II, activation and release of the paused Pol II, productive elongation by Pol II and termination of transcription. The un-induced *hsp70* gene was one of the earliest genes shown to have a transcriptionally engaged Pol II, paused at the +20 to +30 region (Rougvie and Lis, 1988). It is now known that pausing occurs on numerous genes in *Drosophila* (Muse et al., 2007; Lee et al., 2008). Various factors are implicated in regulating the establishment of and release from the paused state.

In the inactive state, the Pol II has initiated transcription on the *hsp70* gene and is stably paused in the promoter proximal region (Rougvie & Lis, 1988). When induced, these paused polymerases rapidly progress into the elongation phase, resulting in a several hundred fold induction of the heat shock genes (Lis et al., 1981).

4-1.1 Setting up the paused Pol II on *hsp70*

Early UV-crosslinking experiments on the *hsp70* gene showed that Pol II occupied the promoter proximal region and was in a elongation competent form (Gilmour et al., 1986; Rougvie et al.,1988). To set up a paused Pol II, the promoter needs to be in an open chromatin structure to allow formation of the pre-initiation complex by recruitment and binding of the general transcription factors and Pol II (Orphanides and Reinberg, 2000). On the *hsp70* gene, two factors

are thought to be involved in this step. GAGA factor is a transcription factor that binds to GA/CT elements present on the *hsp70* promoter. Biochemical experiments on reconstituted chromatin showed that binding of GAGA factor resulted in a DNAse hypersensitive region on the promoter and was discovered to be due to the activity of an ATP-dependent chromatin remodeler, NURF (Tsukiyama and Wu, 1995; Tsukiyama et al., 1994). GAGA factor binding resulted in recruitment of NURF, a member of the ISWI family of remodelers, to the promoter, which mobilizes the nucleosomes and opens up the chromatin. GAGA factor is essential for initiation and pausing of RNA Pol II as mutations in GAGA elements result in loss of paused Pol II (Lee et al., 1992; Li et al., 2013).

Biochemical analysis of transcription elongation identified two factors, NELF and DSIF, that function to pause RNA Pol II (Renner et al., 2001; Cheng et al., 2007; Yamaguchi et al., 1999). Depletion of NELF reduces the level of paused pol II on the *Drosophila hsp70* gene (Wu et al., 2003). More recent genome wide studies have shown that depletion of NELF causes the level of Pol II at the promoters of several thousand genes to decrease, thus implicating NELF as the factor that pauses RNA Pol II (Gilchrist et al., 2010).

4-1.2 Activation of *hsp70*

The master activator heat shock factor (HSF), is essential for heat shock induced activation of *hsp70* (Boehm et al., 2003; Lis and Wu, 1993). Upon heat shock, HSF undergoes homotrimerization, phosphorylation and binds to the heat shock elements in the promoter of *hsp70* (Fernandes et al., 1995; Rabindran et al., 1993). The activity of NURF is thought to be required for activation of the heat shock gene as it functions to expose the heat shock factor binding sites so that they are accessible to heat shock factor during heat shock (Badenhorst et al.,

2002; Tsukiyama et al., 1994). HSF interacts with the TRAP80 subunit of mediator complex, which serves to recruit additional Pol II to the *hsp70* promoter (Park et al., 2001).

In addition to the activity of HSF, release of paused Pol II requires the function of P-TEFb (Ni et al., 2008). P-TEFb is a cyclin dependent kinase that was characterized in biochemical experiments as the kinase responsible for rendering the reaction sensitive to the transcription inhibitor, DRB (Marshall et al., 1995). P-TEFb is recruited to the *hsp70* gene upon heat shock, where it alleviates the NELF-mediated pausing of Pol II (Lis et al., 2000). P-TEFb phosphorylates NELF and DSIF, resulting in release of NELF and converting DSIF into a positive elongation factor that then tracks with Pol II (Fujinaga et al., 2004b; Yamada et al., 2006).

P-TEFb also phosphorylates Ser2 of the Pol II CTD, which is a hallmark of elongating Pol II. The C-terminal domain (CTD) of the largest subunit of Pol II contains repeats of a heptapeptide sequence, YSPTSPS. The serines in the 2nd and 5th positions are the primary targets of kinases during the transcription cycle. Different stages of transcription are marked by distinct phosphorylation states of the CTD of Pol II. In metazoans, Pol II in the pre-initiation complex is hypo-phosphorylated, and in the paused state Pol II is phosphorylated at serine 5 by TFIIH.

The CTD of Pol II also serves as a platform for binding and activity of a number of factors that are required for processing of precursor mRNA into mature mRNA. Phosphorylation of serine 5 results in recruitment of the mRNA capping enzyme (Cho et al., 1997). The serine 2 phosphorylated form of Pol II interacts with 3' end processing factors such as Pcf11 (Lunde et al., 2010). Inhibition of P-TEFb impairs the 3' end processing of the *hsp70* mRNA (Ni et al., 2004).

The transcription elongation factor, Spt6 is also recruited to the *hsp70* gene upon heat shock. Recruitment of Spt6 is dependent on HSF (Andrulis et al., 2000). Spt6 is thought to aid the pioneering RNA Pol II to transcribe across a nucleosomal template, as depletion of Spt6 resulted in lower levels of Pol II following short times of heat shock but did not show reduced levels of

Pol II after steady state transcription is established (Ardehali et al., 2009). FRAP experiments helped establish that Spt6 depletion causes a 2 fold decrease in the transcription rate on the *hsp70* gene (Ardehali et al., 2009). These results that show that a histone chaperone is important in promoting transcription of *hsp70* indicates that the chromatin organization on this gene is refractory to elongation.

4-1.3 Chromatin organization on the *hsp70* gene

The chromatin architecture of the un-induced *hsp70* gene shows 4 distinct regions. The first is a nucleosome free region at the promoter, second is the first nucleosome downstream of the TSS, which is centered at +330, third, a low level of poorly positioned nucleosomes in the body of the gene and fourth, a nucleosome free region at the 3' end of the gene (Petesch and Lis, 2008). Within 30 seconds of heat shock Lis and colleagues observed a decrease in the level of MNase protection across the gene (Petesch and Lis, 2008). This decrease is thought to be transcription independent since based on earlier results, it takes approximately 2 minutes for RNA Pol II to transcribe to the end of *hsp70* gene (Boehm et al., 2003). In addition, the loss of nucleosomes was observed when transcription was inhibited by treatment with DRB. By 2 minutes after heat shock, there is a complete loss of nucleosomes across the gene. This loss of nucleosomes is dependent on transcription.

4-4.4 Factors affecting chromatin organization after heat shock

Three factors were identified as important for the loss of nucleosomes on the *hsp70* gene after heat shock. Depletion of HSF, GAGA factor and PARP inhibited the heat shock-dependent loss of nucleosomes during heat shock. PARP is an enzyme that catalyzes formation of long

chains of poly-ADP ribose units. PARP functions in formation of heat shock induced 'puffs' by displacing histones, which bind to the negatively charged PAR polymers (Tulin and Spradling, 2003). The chains of poly-ADP ribose units are thought to form a "transcription compartment" which allows for retention of factors like Pol II, Spt6 and to some extent P-TEFb, during prolonged heat shock (Zobeck et al., 2010). Evidence for the formation of a "transcription compartment" comes from fluorescence recovery experiments carried out in *Drosophila* cells expressing a fluorescently tagged Pol II subunit (Yao et al., 2007). The authors observed that even though there was a steady increase in the amount of mRNA being produced during heat shock, recovery of the fluorescent signal from GFP tagged Pol II molecules from photo-bleached cells after a prolonged heat shock (20 minutes) slowed down compared to recovery after bleaching in the initial stages of activation (2 minutes). The authors also showed that the level of Pol II binding to *hsp70* gene as detected by ChIP is maximal after 5 minutes of heat shock and decreased upon longer periods of heat shock. However, there was no decrease in intensity of fluorescently tagged Pol II molecules after an initial increase in intensity in accordance with the increase observed in the Pol II ChIP experiments, suggesting that Pol II molecules not actively engaged in transcription were being retained at the heat shock loci in a "transcription compartment" (Yao et al., 2007). Formation of the compartment is proposed to facilitate recycling of factors during prolonged heat shocks (Zobeck et al., 2010).

PARP activity also results in loss of nucleosomes on the *hsp70* gene (Petesch and Lis, 2008). Upon heat shock, there is a loss of nucleosomes on the *hsp70* gene that occurs in two phases. The first phase occurs before Pol II has transcribed to the end of the gene and the second phase is dependent on transcription. Depletion of PARP or inhibition of its activity affects both phases of nucleosome loss (Petesch and Lis, 2008).

Amongst the other factors that affected the chromatin structure of the *hsp70* gene, the histone deacetylase HDAC3 was identified as a factor that maintained the chromatin structure of

the *hsp70* gene prior to heat shock. Depletion of HDAC3 resulted in a decrease but not a complete loss of nucleosomes, similar to the transcription independent loss observed after a 1-minute heat shock (Petesch and Lis, 2008). HDAC3 is required to maintain PARP in an inactive state under non heat shock conditions and its depletion causes relocation of PARP from the promoter to the body of the gene with a concomitant increase in PAR under non-heat shock conditions (Petesch and Lis, 2012). HDAC3 depletion in tissue culture cells showed a 33% reduction in *hsp70* mRNA after a 20 minute heat shock (Ardehali et al., 2009).

HDAC3 maintains *hsp70* gene in a deacetylated state prior to heat shock. Decrease in HDAC3 levels or its activity resulted in an increase in H2AK5 acetylation, which in turn acts as an activating signal for PARP (Petesch and Lis, 2012).

4-4.4.1 HDAC3

HDAC3 belongs to the Class 1 Histone deacetylases. HDAC3 was initially identified in human cells, by sequence alignments of expressed sequence tags from the NCBI database, with the yeast RPD3 sequence (Emiliani et al., 1998). Of the Class I HDACs, HDAC3 is the only HDAC that contains both a nuclear localization signal (NLS) and a nuclear export signal (NES), and it is found both in the nucleus and cytoplasm (Gregoretti et al., 2004; Johnson et al., 1998). HDAC3 is associated the hormone receptor co-repressor complexes, N-CoR and SMRT and forms a very stable complex. N-CoR and SMRT function to recruit HDAC3 to its target genes and also to activate HDAC3 (Guenther et al., 2001). SMRTER is the *Drosophila* homolog of the nuclear receptor co-repressor complexes, SMRT and N-CoR (Tsai et al., 1999).

In *Drosophila*, HDAC3 functions during development through interactions with Ebi, a co-repressor of Snail that is required for mesoderm formation. HDAC3 depletion impaired Snail-mediated repression in cells (Qi et al., 2008). Loss of HDAC3 results in increased apoptosis and

is essential for development as loss of HDAC3 is lethal during larval/pupal stages (Zhu et al., 2008).

Most studies implicate histone deacetylases in repression of expression (Kouzarides, 2007; Yang and Seto, 2003). Acetylation of histones is correlated with gene expression as acetylation is thought to loosen up the chromatin. The acetylated residues also act as sites for binding of factors essential for transcription. Conversely, deacetylation is thought to prevent binding of such factors and thus repress transcription (Deckert and Struhl, 2002). However, a number of recent studies suggest that histone deacetylases can also positively regulate transcription (Dovey et al., 2010; Reynolds et al., 2013; Smith, 2008).

The genome wide localization of HDAC3 suggests a role of HDAC3 in elongation. HDAC3 is associated with actively transcribed genes (Wang et al., 2009). Genome wide mapping of HDAC3 in *Drosophila* shows that HDAC3 binds in the bodies of actively transcribed genes that are marked by the H3K36me3 histone modification (Nègre et al., 2011).

My study attempted to identify additional factors involved in regulation of *hsp70* expression. I developed a beta-gal reporter assay to assess the effects of depletion of various factors implicated in *hsp70* transcription. My screen identified HDAC3 and its associated factor SMRTER as positive regulators of expression of *hsp70* gene. Depletion of HDAC3 resulted in decreased expression and reduced levels of Pol II at the *hsp70* gene. HDAC3 is also involved in stably pausing Pol II in the promoter proximal region of *hsp70* gene.

4-2 Results

4-2.1 RNAi-screening for factors that affect induction of an *hsp70* reporter transgene

I set out to identify transcription factors involved in activation of *hsp70* gene in *Drosophila* by using RNAi to deplete specific factors and an *hsp70*-beta-galactosidase reporter gene to rapidly assess the effects of the depletion on heat shock induction of the reporter gene. I choose to develop this assay in salivary glands of *Drosophila* 3rd instar larvae for two reasons. One, it provided a quick visual evaluation of the effect of depletion of factors on *hsp70* expression. Two, preliminary experiments indicated that the effect of RNAi in salivary glands were more robust than what I observed in tissue culture cells (Appendix D). Depletion of some of the factors identified in this screen did not affect *hsp70* expression in tissue culture cells to the extent that I observed in salivary glands. Similar results have been reported for RNAi-mediated depletion of other factors (Murawska et al., 2011). Moreover, we could easily monitor the effects of the RNAi on the morphology of the tissue and on the global appearance of the polytene chromosomes.

RNAi fly-lines targeting various factors were mated with the Z243, 1824 line to deplete specific factors only in the salivary glands. The Z243, 1824 fly-line contains two transgenes (Figure 4-1A). The Z243 transgene is a reporter gene consisting of the *hsp70* promoter fused to a beta-gal gene. The 1824 transgene expresses GAL4 only in the salivary glands. Mating the Z243, 1824 fly to a fly targeting a specific protein by a GAL4-induced RNAi transgene results in depletion of the protein in salivary glands of the progeny. The progeny were heat shocked for 30 minutes to activate the *hsp70* transgene and then allowed to recover for 30 minutes to allow synthesis of beta-galactosidase. The larvae were dissected and the salivary glands were fixed in gluteraldehyde followed by incubation in X-gal solution for 2 hours. Production of blue color

indicates expression of the *hsp70*-beta gal transgene and hence normal activation of the *hsp70* promoter. Figure 4-1B shows the staining of salivary glands from a mating of Z243,1824 (driver) to yw, a fly line that does not bear any RNAi transgene. It serves as a positive control as it is expected to produce high levels of beta-galactosidase. The intense X-gal staining indicates high levels of beta-galactosidase being produced and hence activation of the *hsp70* transgene.

To evaluate this assay, I determined if depletion of factors that are known to play a positive role in *hsp70* induction showed reduced blue color (Figure 4-1C- E, I-J). HSF is the activator protein that binds to the heat shock elements in the promoter of *hsp70* and activates transcription (Boehm et al., 2003; Fernandes et al., 1995). P-TEFb is a kinase whose activity is important in release of the paused Pol II and productive elongation (Ni et al., 2008) . As expected, depletion of HSF or the subunits of P-TEFb, Cdk9 and Cyclin T1, resulted in significantly diminished levels of X-gal staining in salivary glands from heat-shocked larvae (Figure 4-1C-E). In addition, depletion of Cdk9 or Cyclin T significantly diminished the size of the glands.

Additional factors that have been shown to be involved in *hsp70* transcription were assayed. NURF is a chromatin remodeler that functions to open up the chromatin structure at the *hsp70* promoter region to allow for initiation and formation of the paused complex (Badenhorst et al., 2002; Tsukiyama and Wu, 1995). ELL is an elongation factor that promotes elongation by suppressing transient pausing of Pol II in the body of the gene (Gerber et al., 2001). It co-localizes at the *hsp70* loci with serine 2 phosphorylated form of Pol II and depletion of ELL is shown to reduce *hsp70* mRNA levels (Smith et al., 2008). Depletion of NURF and ELL showed modest reduction in the transcription of the transgene (Figure 4-1F-G).

Rpd3 is an HDAC present in the Sin3 complex. A subunit of Sin3, SAP18, interacts with GAGA factor, which as already discussed, is involved in *hsp70* transcription. Genome wide ChIP-chip data showed Rpd3 is present at the 5' ends of genes (Nègre et al., 2011). Hence, I speculated that Rpd3 could play a role in *hsp70* transcription, possibly by maintaining it in a

paused state. Immunofluorescence staining of polytene chromosomes from Rpd3 depleted glands showed that it was depleted from the chromosomes indicating that the RNAi against Rpd3 was efficient (Appendix E). Depletion of Rpd3 did not affect activation of *hsp70* as the X-gal staining in Rpd3 depleted and heat shocked glands appears similar to the control (Figure 4-1H). HDAC3, also a histone deacetylase was shown to be important in maintaining the chromatin structure on *hsp70* gene (Petesch and Lis, 2008). Depletion of HDAC3 was expected to show no change or a modest effect on *hsp70* activation. Surprisingly depletion of HDAC3, a histone deacetylase, resulted in significant reduction in X-gal staining (Figure 4-1-I). Depletion of SMRTER, a HDAC3 associated co-factor also showed reduced X-gal staining (Figure 4-1-J).

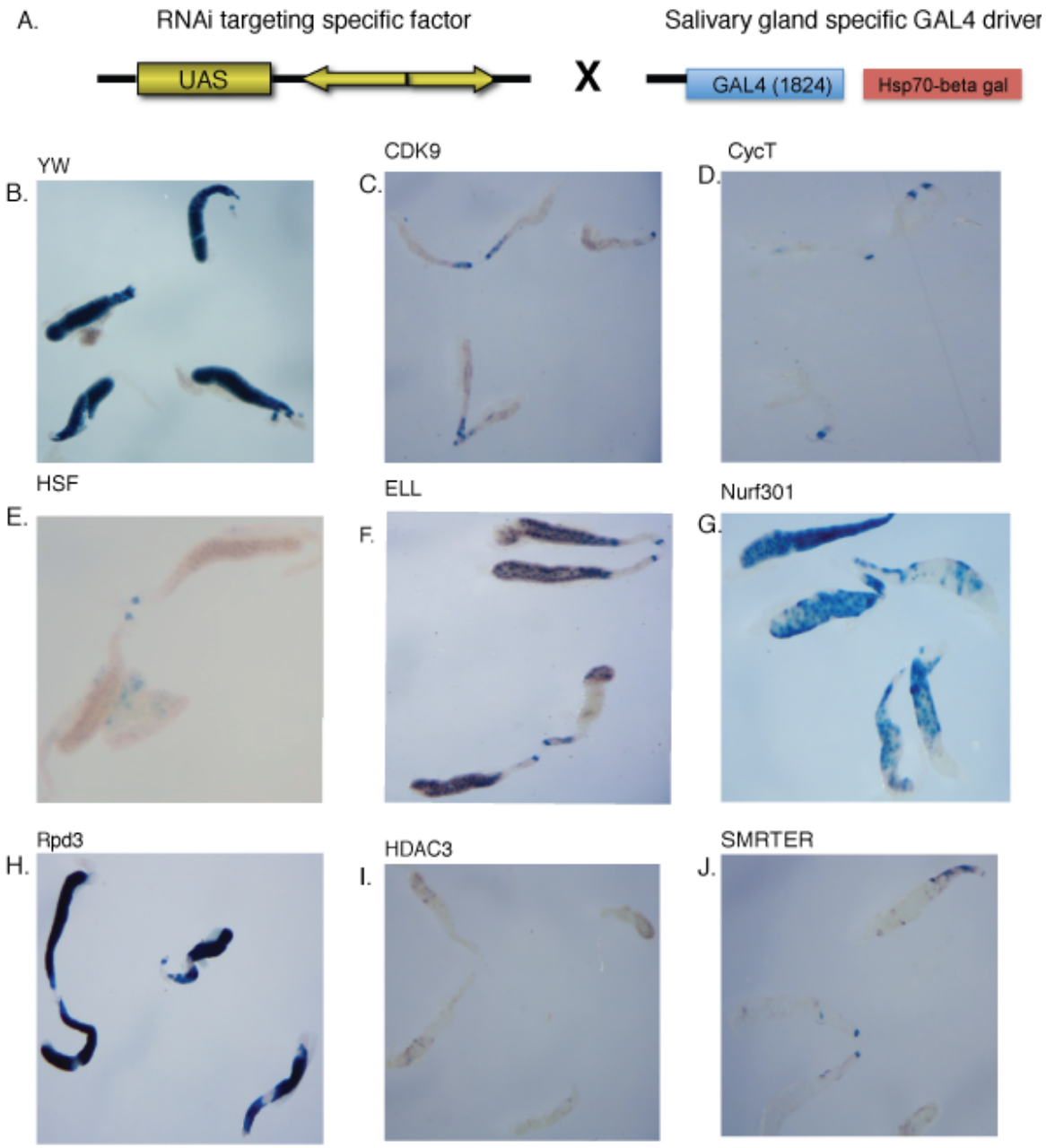


Figure **4-1**:Inhibition of heat shock induction of an *hsp70*-beta-gal reporter gene by specific RNAi's salivary glands.

(A) Schematic of the mating between an RNAi fly-line with a Gal4-regulated RNAi transgene and a fly containing an *hsp70-beta-gal* reporter gene (Z243) and a transgene expressing Gal4p in salivary glands. Third instar larvae from the mating were heat shocked for 30 minutes and then recovered at 22^0C for 30 minutes. The dissected salivary glands were incubated in X-gal solution for 2 hours. **(B)** Salivary glands from a control mating of yw with Z243,1824. yw is a

control fly line that lacks the RNAi transgene. Depletion of Cdk9, Cyclin T1 (subunits of P-TEFb), HSF (heat shock activator) inhibits activation of the *hsp70* transgene (Figure 4-1 C-E), as seen by the very low levels of X-gal staining in the salivary glands. Depletion of ELL or Nurf301 has a more modest effect. Depletion of Rpd3 (HDAC1) does not inhibit heat shock induction of the transgene. Depletion of HDAC3 or SMRTER shows reduced expression of beta-gal (Figure 4-1 I-J).

4-2.2 HDAC3 depletion results in reduced transcription of the endogenous *hsp70* gene

The decrease in beta-galactosidase in HDAC3 depleted glands could reflect defective transcription and/or translation of the transgene. To determine if HDAC3 affected transcription, I measured the levels of mRNA produced after heat shock. Dissected glands were heat shocked for 10 minutes and 30 minutes and RT-PCR was performed on RNA isolated from the salivary glands. The mRNA levels are expressed as a ratio of *hsp70* to *rp49* mRNA, which serves as an internal standard, to correct for variation in recovery of RNA. I chose *rp49* as an internal standard as its mRNA levels are not affected by HDAC3 or SMRTER depletion. Also, *rp49* mRNA has been used in earlier studies as an internal standard for measurement of *hsp70* mRNA levels (Ardehali et al., 2009; Ghosh et al., 2011). Depletion of HDAC3 and SMRTER showed a decrease in levels of the endogenous *hsp70* mRNA after heat shock and the effect is more pronounced at 10 minutes of heat shock compared to 30 minutes (Figure 4-2). A second RNAi line targeting HDAC3 (HDAC3_2 in Figure 4-2) showed similar reductions in *hsp70* mRNA. The 2 fold decrease in mRNA level after 30 minutes of heat shock is in agreement with a 33% reduction in mRNA after depletion of HDAC3 in S2 cells reported earlier (Ardehali et al., 2009).

As a further validation of my results, I also measured the level of induction of the reporter transgene, using primers that targeted the beta-gal gene (Figure 4-3). I see a similar decrease in the activation of the *hsp70*-beta-gal transgene as observed in the endogenous *hsp70* gene.

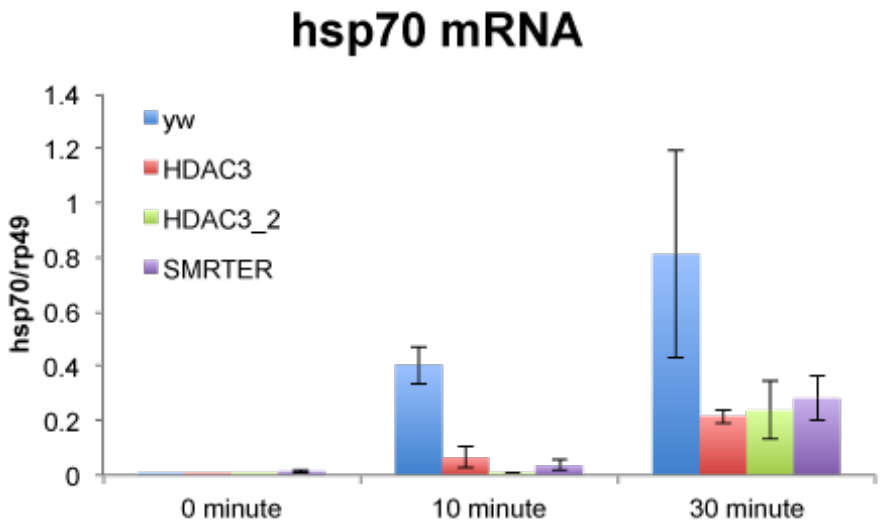


Figure **4-2**: Depletion of HDAC3 inhibits heat shock induction of the endogenous *hsp70* gene.

RT-PCR analysis of *hsp70* mRNA levels in salivary glands from RNAi lines targeting HDAC3, SMRTER or control (yw) larvae. Results from the second RNAi fly line targeting HDAC3 are shown (HDAC3_2). HDAC3 and SMRTER RNAi fly lines were obtained from Vienna Drosophila RNAi Centre and the stock number for HDAC3 is VDRC-KK 107073, for HDAC3_2 is VDRC 20814 and for SMRTER is VDRC-KK 106701. Salivary glands were heat shocked at 37^0C for 10 minutes or 30 minutes. Total RNA was isolated, followed by cDNA synthesis. *hsp70* was detected by qPCR and expressed as a ratio to *rp49*, which serves as an internal standard. Results are from three independent experiments and error bars represent SEM.

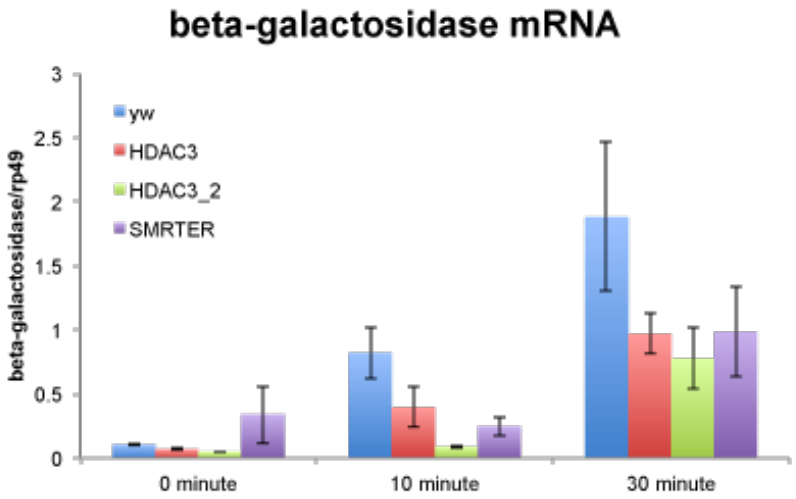


Figure **4-3**: Depletion of HDAC3 inhibits heat shock dependent synthesis of *hsp70*-beta gal transgene mRNA.

RT-PCR analysis of beta-galactosidase transgene mRNA levels in salivary glands from RNAi lines targeting HDAC3, SMRTER or control (yw) larvae. Results from the second RNAi fly line targeting HDAC3 are shown (HDAC3_2). Salivary glands were heat shocked at 37^0C for 10 minutes or 30 minutes. Total RNA was isolated. cDNA library was generated. Beta-galactosidase mRNA was detected by qPCR and expressed as a ratio to rp49, which serves as an internal standard. Results are from three independent experiments and error bars represent SEM.

4-2.3 HDAC3 is present at the heat shock puffs

To determine if HDAC3 associates with the heat shock genes, I stained polytene chromosomes from heat-shocked larvae with antibodies against HDAC3 and Rpb3 (subunit of Pol II). I observed HDAC3 and Pol II staining of the heat shock puffs in salivary glands from heat shocked control larvae (Figures 4-4B, 4-4C). To determine if the staining by the HDAC3 antibody is specific for HDAC3, I monitored HDAC3 staining on polytene chromosomes from heat shocked larvae expressing HDAC3 RNAi (Figure 4-5B). Staining for HDAC3 is absent at the heat shock puffs in the HDAC3 depleted salivary glands, thus confirming the depletion of HDAC3 by RNAi. Depletion of HDAC3 did not affect binding of Pol II at the heat shock puffs on the polytene chromosomes as monitored by immunofluorescence (Figure 4-5C). This could reflect a limitation of the immunofluorescence experiments as it offers a qualitative evaluation of binding of proteins to genes and quantifying differences in staining is difficult. In *Drosophila*, activity of the enzyme PolyADP-Ribose Polymerase (PARP) is essential for "heat shock puff" formation on polytene chromosomes (Tulin and Spradling, 2003). It has been proposed that the poly-ADP ribose (PAR) polymers that are formed by enzymatic activity of PARP results in retention of Pol II and other transcription factors at heat shock gene loci (Zobeck et al., 2010). It is possible that the Pol II detected at the heat shock puffs in the HDAC3 depleted glands are molecules of Pol II trapped in the PAR polymers and not engaged in active transcription.

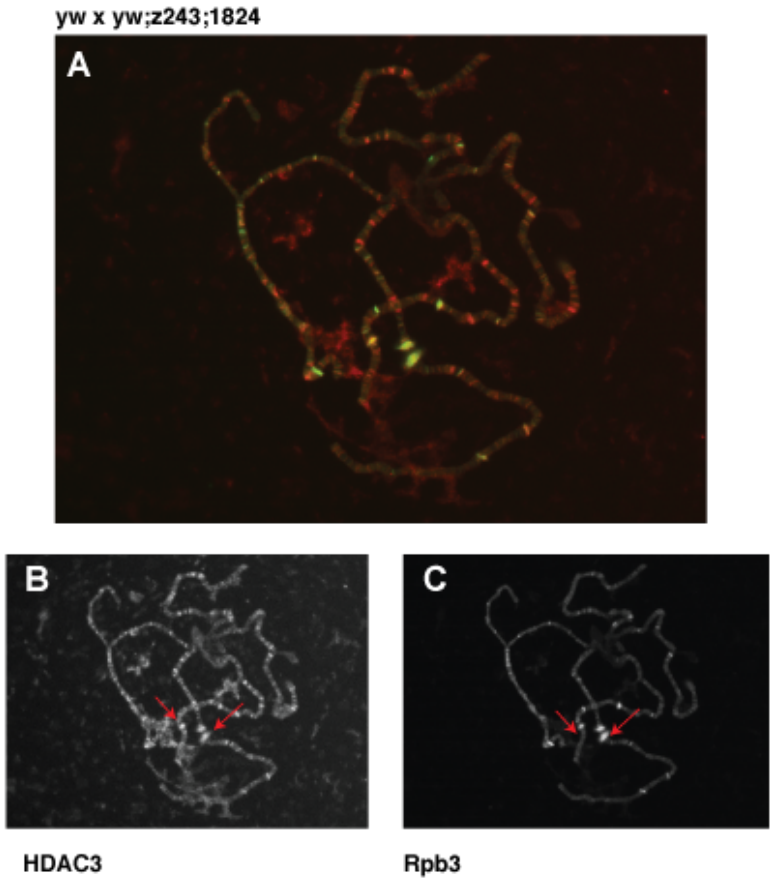


Figure **4-4**: Immunostaining of polytene chromosomes detects HDAC3 at heat shock puffs.

Polytene chromosomes squashes were prepared from salivary glands of control (yw x yw;Z243,1824) third instar larvae that were heat shocked for 10 minutes. **(A)** overlaid image of staining for HDAC3 and Pol II in yw x yw;Z243,1824 larva. Lower panel shows individual staining for HDAC3 **(B)** and Pol II **(C)**. HDAC3 was stained with an anti-guinea pig primary antibody (gift from M.Mannervik) at 1:50 dilution. Pol II was stained with anti-rabbit primary antibody against Rpb3, at a 1:100 dilution. The primary antibodies were visualized with a Alexa-568 conjugated anti-guinea pig and Alexa-488 conjugated anti-rabbit antibody.

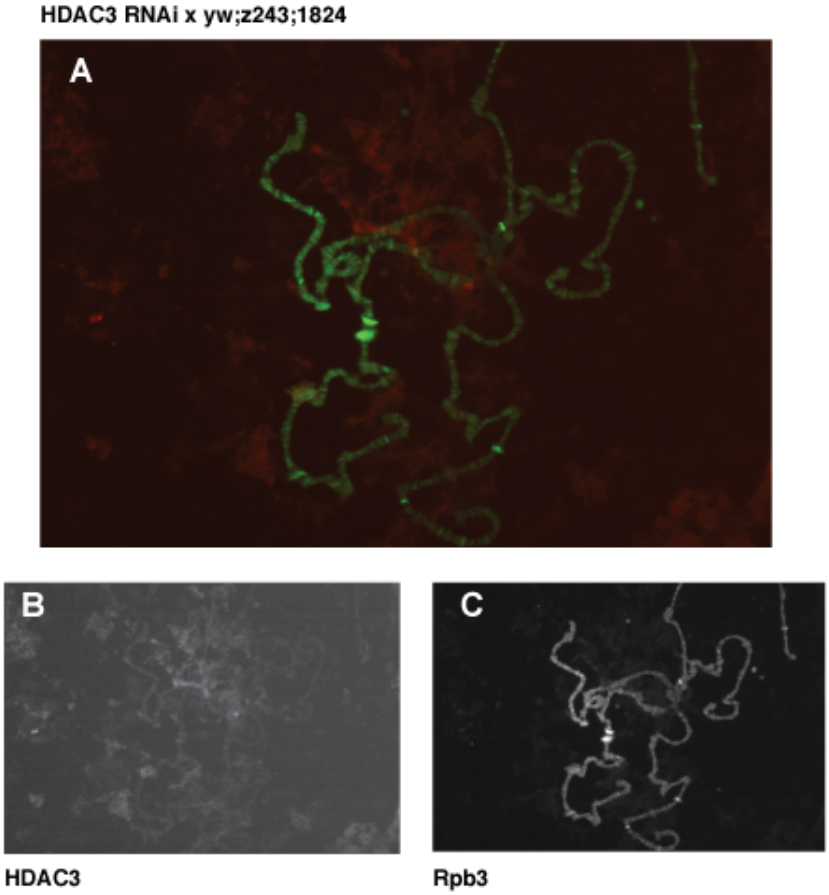


Figure **4-5**: Loss of HDAC3 staining on polytene chromosomes in HDAC3 depleted glands.

Polytene chromosomes squashes were prepared from salivary glands of HDAC3 depleted (HDAC3 x yw;Z243,1824) third instar larvae that were heat shocked for 10 minutes. **(A)** Overlaid image of staining for HDAC3 and Pol II in yw x yw;Z243,1824 larva. Lower panel shows individual staining for HDAC3 **(B)** and Pol II **(C)**. HDAC3 was stained with an anti-guinea pig primary antibody (gift from M.Mannervik) at 1:50 dilution. Pol II was stained with anti-rabbit primary antibody against Rpb3, at a 1:100 dilution. The primary antibodies were visualized with a Alexa-568 conjugated anti-guinea pig and Alexa-488 conjugated anti-rabbit antibody.

4-2.4 Depletion of HDAC3 reduces the level of Pol II on the *hsp70* gene during heat shock

Although depletion of HDAC3 greatly diminished the levels of *hsp70* mRNA, it did not affect the staining of Pol II at the heat shock puffs. A limitation of immunofluorescence experiments is that it offers only a qualitative evaluation of binding of proteins to genes. To assess the levels of bound Pol II at a quantitative level, I carried out ChIP in salivary glands for Rpb3 (subunit of Pol II) under non-heat shock conditions and after 10 minutes of heat shock. The levels of Pol II detected were well above background (IgG control in Figure 4-6). I examined three regions of the gene corresponding to the 5' promoter proximal region (+65), the body of the gene (+861) and the 3' end of the gene (+1976).

In the control, under uninduced conditions, I detect the paused Pol II at the promoter proximal region (+65 region in Figure 4-6A). There are very low levels of Pol II in the body and 3' end of the gene (+861 and +1976 regions in Figure 4-6A). Upon heat shock, there is a 2-fold increase in the level of Pol II associated with the *hsp70* promoter region of control glands but no increase for the HDAC3 and SMRTER depleted glands (Compare Figures 4-6A and 4-6B).

While Pol II levels in the paused region and body of the gene in heat shocked glands are lower for HDAC3 and SMRTER depleted glands than control glands, they show no significant difference at the 3' end of the gene (+1976 region, Figure 4-6B). Published results show that Pol II starts to pile up at the 3' end of the *hsp70* gene, just after the polyadenylation site. This is thought to be due to slowing down of polymerase in this region to allow for 3' end processing and termination of transcription (Ni et al., 2008). The similar levels of Pol II at the +1976 region could be the result of the build-up of Pol II towards the 3' end of the gene. I did not analyze the region downstream from the polyadenylation site because there are 3 copies of the *hsp70* gene arranged in a tandem array and the region downstream from the polyadenylation site is in close proximity to another *hsp70* promoter.

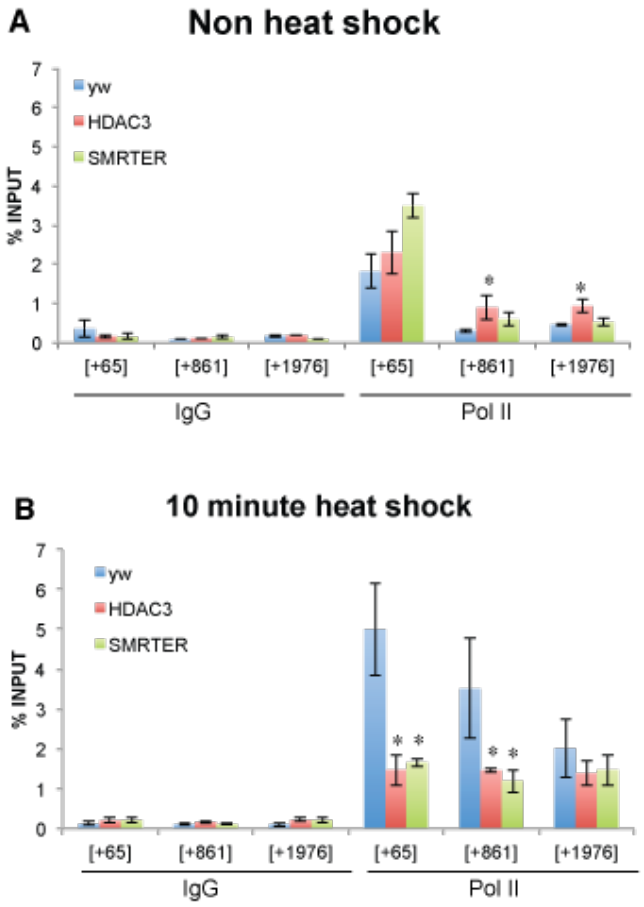


Figure **4-6**: ChIP for Pol II in control, HDAC3 and SMRTER depleted salivary glands.

ChIP for Pol II from salivary glands under **(A)** non-heat shock conditions or **(B)** heat shocked for 10 minutes. Numbers on the x-axis represent the center of the amplified region on the *hsp70* gene. Y-axis represents percentage of input that was recovered in the ChIP. Error bars represent SEM of three independent experiments (* indicate a p value <0.03, t-test).

4-2.5 Depletion of HDAC3 or SMRTER does not affect rate of transcription

The ChIP results show that levels of Pol II are reduced at the *hsp70* gene in the HDAC3 and SMRTER depleted glands. The reduced levels of Pol II could be the result of a reduced rate of transcription. While changes in levels of Pol II can be detected by ChIP, it does not assay for changes in rates of transcription. Permanganate footprinting after various times of heat shock allows us to track the movement of Pol II and thus determine if rate of transcription is impaired in HDAC3 or SMRTER depleted glands.

Permanganate footprinting detects transcriptionally engaged Pol II. The thymines in the single stranded DNA in the transcription bubble are more reactive to oxidation by permanganate than the thymines in double stranded DNA. In *Drosophila*, hyper-reactive thymines at +22 and +34 at the *hsp70* promoter reveal the presence of paused Pol II (Giardina et al., 1992). Upon heat shock, Pol II is released from the pause and transcribes into the body of the gene, as is seen by the increased permanganate hyper reactivity at sites beyond the pause site, for example at +76, +101-102 (Figure 4-7A). The single base pair resolution of this assay makes it ideal to track Pol II across the gene, *in vivo*.

Within a minute after heat shock, I observe increased levels of permanganate reactivity in the control glands, at +7-8, which reflects newly initiating Pol II (Figure 4-7A, 4-7B, 4-9A). I also observe increase in reactivity beyond the pause site corresponding to Pol II that has transcribed further downstream. Permanganate footprints associated with the newly initiating Pol II and Pol II that has transcribed beyond the pause site are also observed in the heat shocked HDAC3 or SMRTER depleted glands at similar time points relative to the control. Next, I monitored the appearance of Pol II at the 3' end of the gene (Figure 4-8A, 4-8B, 4-9B). The permanganate reactivity at +2485-2486, +2506, +2509 was higher in control glands that were

heat shocked for 4, 6 and 10 minutes of heat shock compared to glands that were heat shocked for shorter times such as 0, 1 and 2 minutes (Figure 4-8A).

Pol II is detected at the 5' end at 2 minutes after heat shock but does not appear at the 3' end of the gene until 4 minutes and hence I can estimate that it takes Pol II approximately 2 minutes to transcribe the *hsp70* gene. This is in accordance with published results that show the rate of transcription of *hsp70* gene to be 1.25 kb/min (Boehm et al., 2003; Ghosh et al., 2011; Zobeck et al., 2010). There is no detectable difference in the permanganate reactivity at the 3' end in control and HDAC3 depleted glands at 4 minutes suggesting that the rate of transcription is not impaired in the absence of HDAC3 (compare Figure 4-8A and Figure 4-8B). Similar results are observed in SMRTER depleted glands as is evident in the appearance of bands at the 3' end at 4 minutes of heat shock (Figure 4-9B).

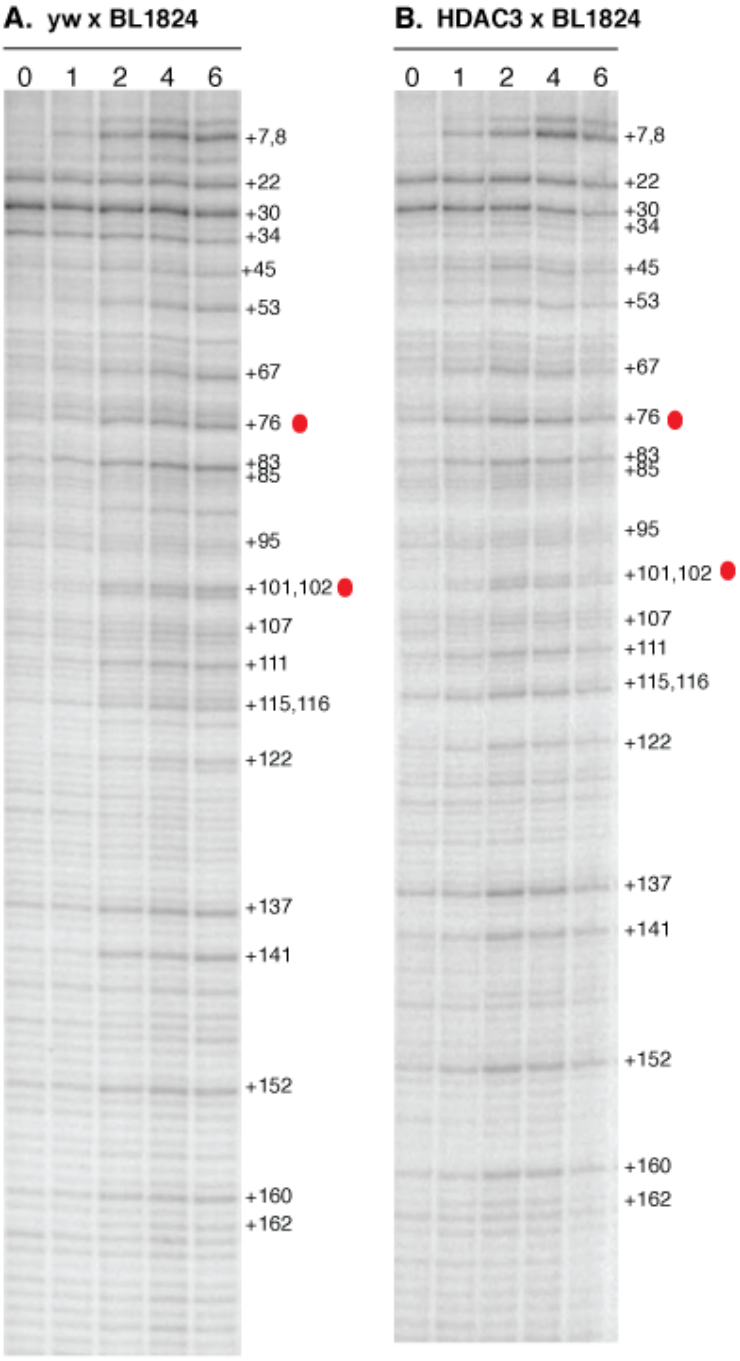
4-2.6 Pausing of Pol II is impaired in HDAC3 or SMRTER depleted glands

In addition to the permanganate footprints under uninduced conditions, pausing of Pol II in the promoter proximal region of *hsp70* is also detected by the ChIP assay, as shown by the high level of Pol II at the +65 region and absence of Pol II in the body (+861 region) of the gene in the control glands (Figure 4-6A). If the Pol II is not stably paused, it could escape into the body of the gene and this would result in higher levels of Pol II in the body of the gene. This is evident in the body of the gene in HDAC3 depleted glands as there is modest increase in level of Pol II in the body of the gene in comparison to the control glands at 0 minutes of heat shock.

The results of the ChIP assay in HDAC3 depleted glands are in contrast to the permanganate footprinting results at 0 minutes of heat shock, as the permanganate reactivity at 0 minutes of heat shock in HDAC3 depleted glands is similar to the control glands. Inherent differences in the two

assays could explain the discrepancy. Permanganate footprinting detects Pol II at near base-pair resolution, while the ChIP assay measures the density of Pol II averaged over a broader region of few hundred nucleotides.

While the permanganate footprinting experiments do not show any differences in reactivity in HDAC3 depleted and control glands, I detect increased reactivity downstream of the paused site in SMRTER depleted glands prior to heat shock. For example, permanganate footprints at +76, +101-102 are observed at 2 minutes of heat shock in control glands, whereas similar footprints are observed in SMRTER depleted glands at 0 and 1 minute of heat shock (Figure 4-9). This would indicate that depletion of SMRTER affects pausing on *hsp70* prior to heat shock induction.


A. yw x BL1824
0 1 2 4 6
+7,8
+22
+30
+34
+45
+53
+67
+76
+83
+85
+95
+101,102
+107
+111
+115,116
+122
+137
+141
+152
+160
+162
B. HDAC3 x BL1824
0 1 2 4 6
+7,8
+22
+30
+34
+45
+53
+67
+76
+83
+85
+95
+101,102
+107
+111
+115,116
+122
+137
+141
+152
+160
+162

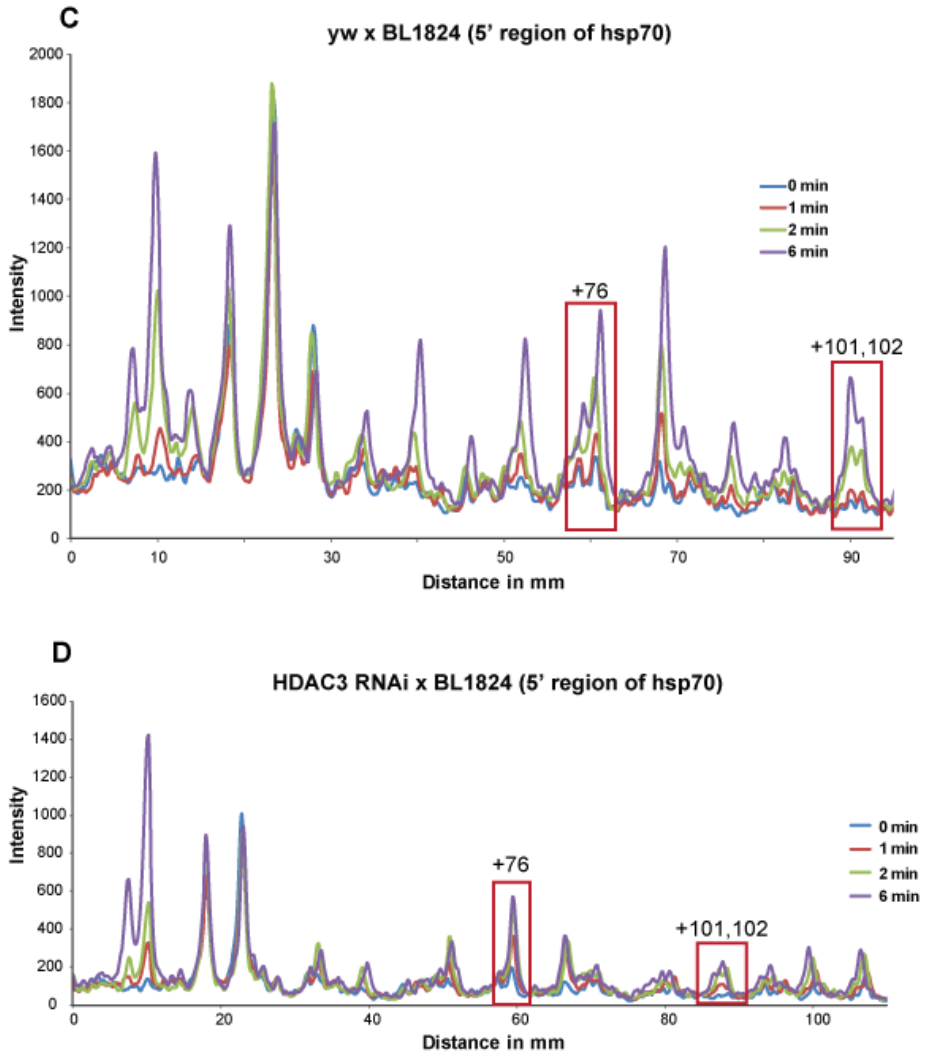
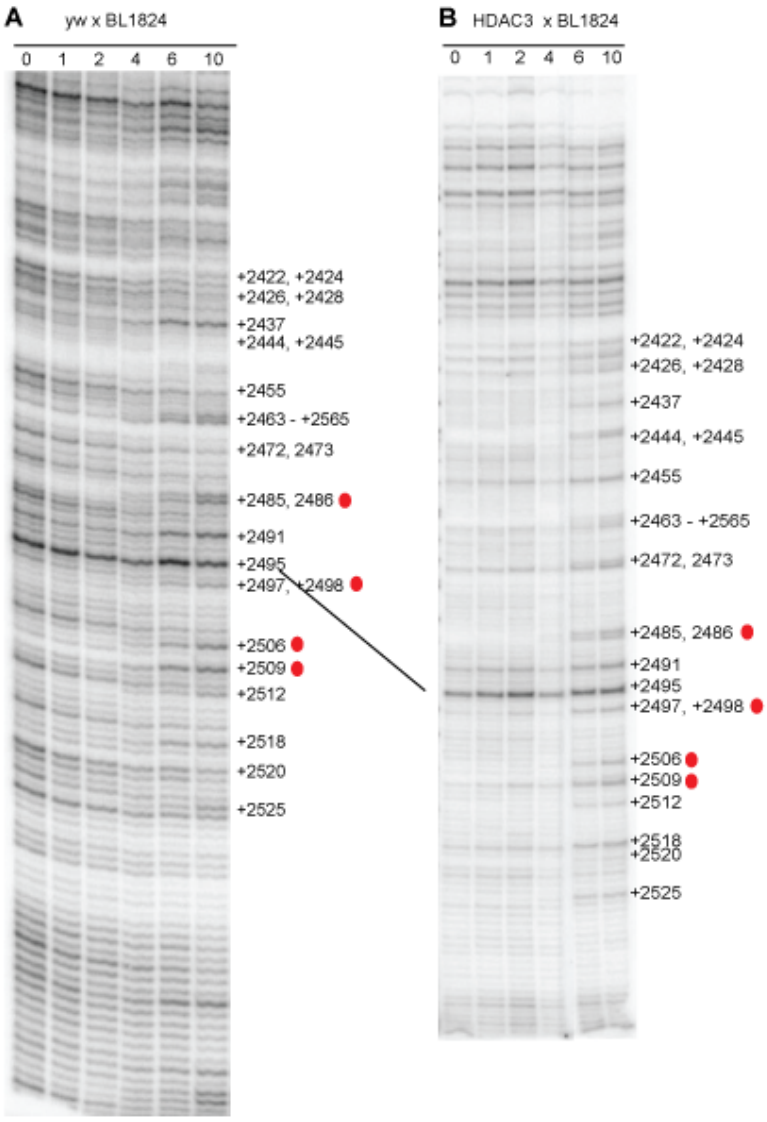


Figure **4-7**: Permanganate foot printing of the promoter proximal region of *hsp70* in control and HDAC3 depleted salivary glands.

Permanganate footprinting in salivary glands isolated from control (yw x BL1824) **(A)** or HDAC3 (HDAC3 x BL1824) depleted glands **(B)**. Dissected glands were heat shocked for indicated number of minutes. LM-PCR was carried out with primers corresponding to the promoter proximal region of *hsp70* gene. The bottom panels show the intensities of the bands in the 0 minute, 1 minute, 2 minute and 6 minute heat shocked lanes after correcting for differences in background. The densitometric traces were drawn using Imagequant software (GE Healthcare). The gel shown is representative of 2 independent experiments.


A
yw x BL1824
0 1 2 4 6 10
+2422, +2424
+2426, +2428
+2437
+2444, +2445
+2455
+2463 - +2565
+2472, 2473
+2485, 2486
+2491
+2495
+2497, +2498
+2506
+2509
+2512
+2518
+2520
+2525
B
HDAC3 x BL1824
0 1 2 4 6 10
+2422, +2424
+2426, +2428
+2437
+2444, +2445
+2455
+2463 - +2565
+2472, 2473
+2485, 2486
+2491
+2495
+2497, +2498
+2506
+2509
+2512
+2518
+2520
+2525

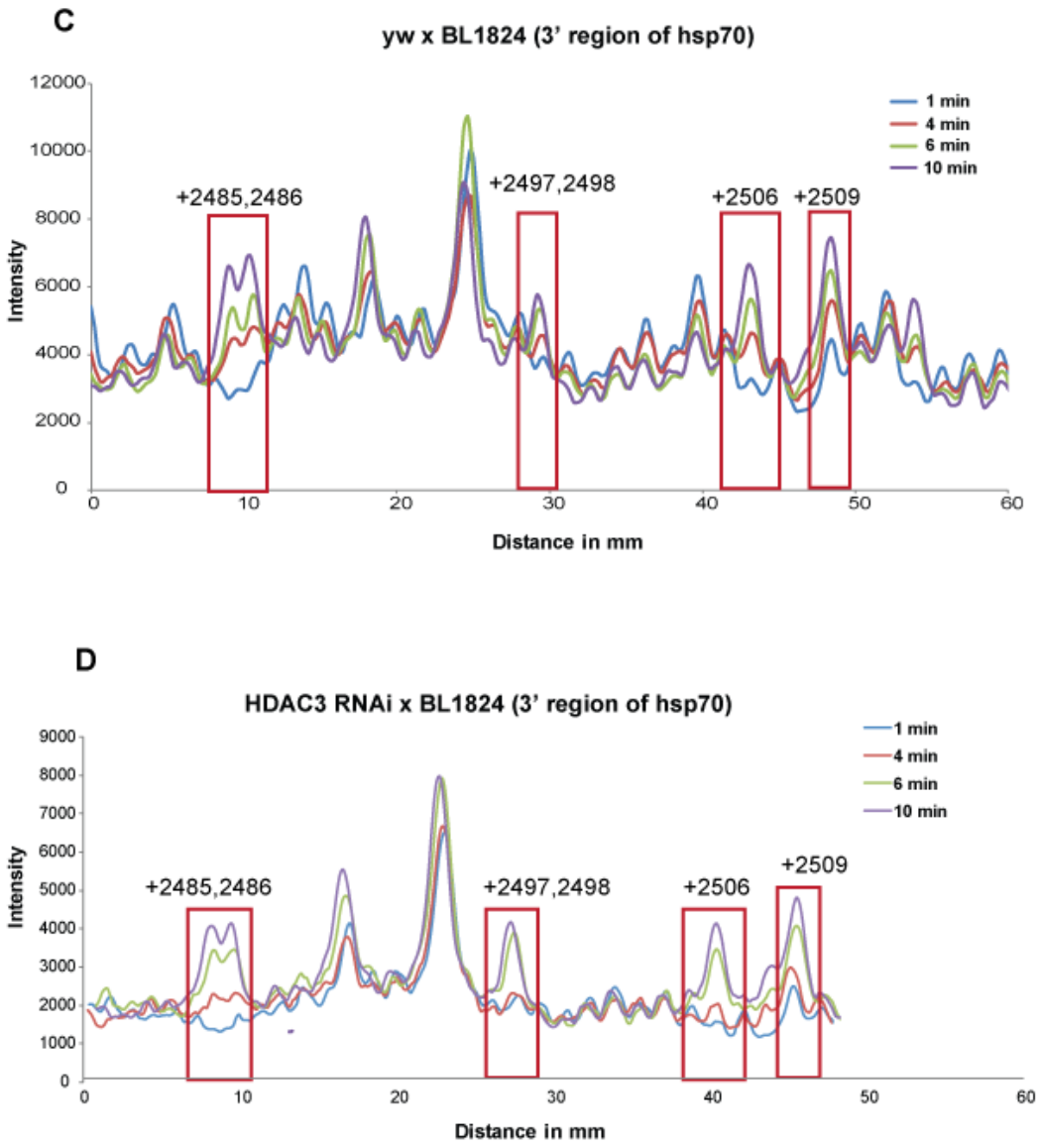
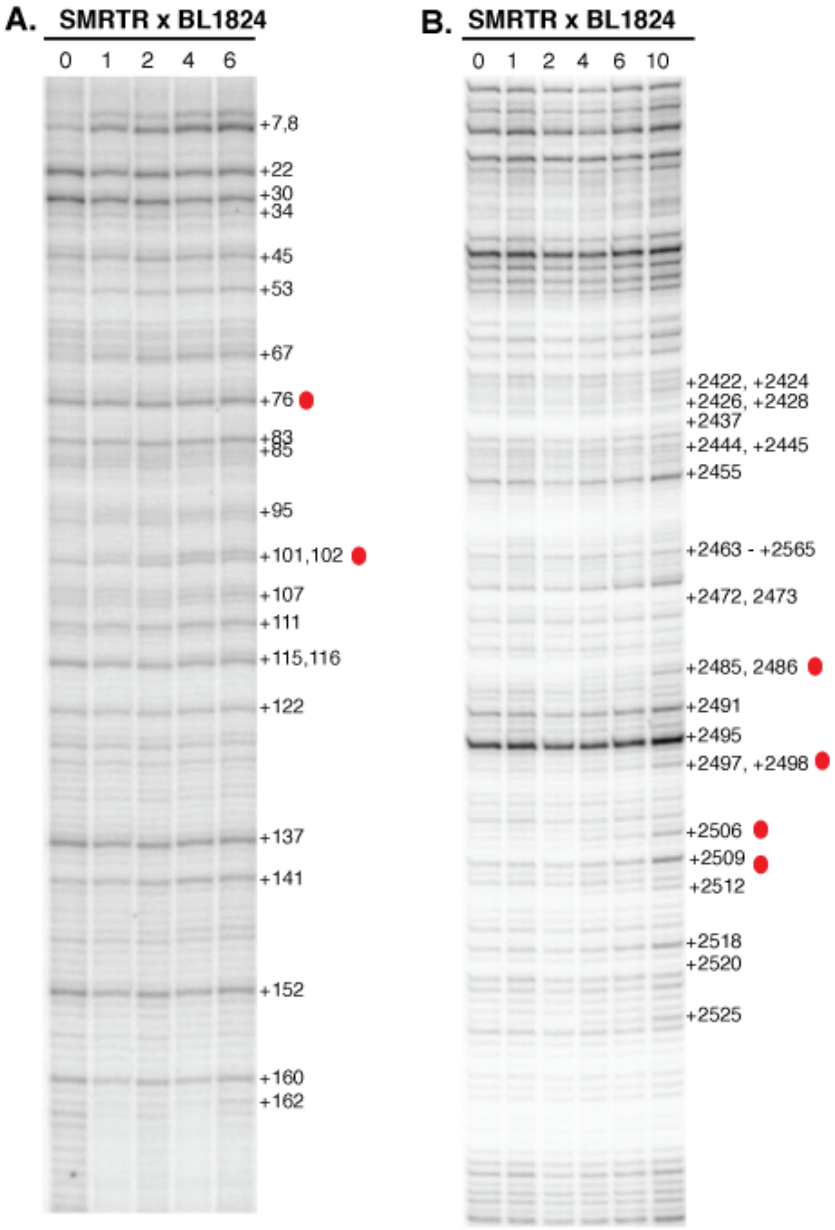


Figure **4-8**: Permanganate footprinting in 3' end of *hsp70* gene in control and HDAC3 depleted glands.

Permanganate footprinting in salivary glands isolated from control (yw x BL1824) **(A)** or HDAC3 (HDAC3 x BL1824) depleted glands **(B)**. Dissected glands were heat shocked for indicated number of minutes; with 0 minutes being non heat shocked condition. LM-PCR was carried out with primers corresponding to the 3' end of *hsp70* gene. The bottom panels show the intensities of the bands in the 1-minute, 4 minute, 6 minute and 10 minute heat shocked lanes after correction for differences in the background. The densitometric traces were drawn using Imagequant software (GE Healthcare) after correction for differences in the background. The gel shown is representative of 2 independent experiments.


A.
SMRTR x BL1824
0 1 2 4 6
+7,8
+22
+30
+34
+45
+53
+67
+76
+83
+85
+95
+101,102
+107
+111
+115,116
+122
+137
+141
+152
+160
+162
B.
SMRTR x BL1824
0 1 2 4 6 10
+2422, +2424
+2426, +2428
+2437
+2444, +2445
+2455
+2463 - +2565
+2472, 2473
+2485, 2486
+2491
+2495
+2497, +2498
+2506
+2509
+2512
+2518
+2520
+2525

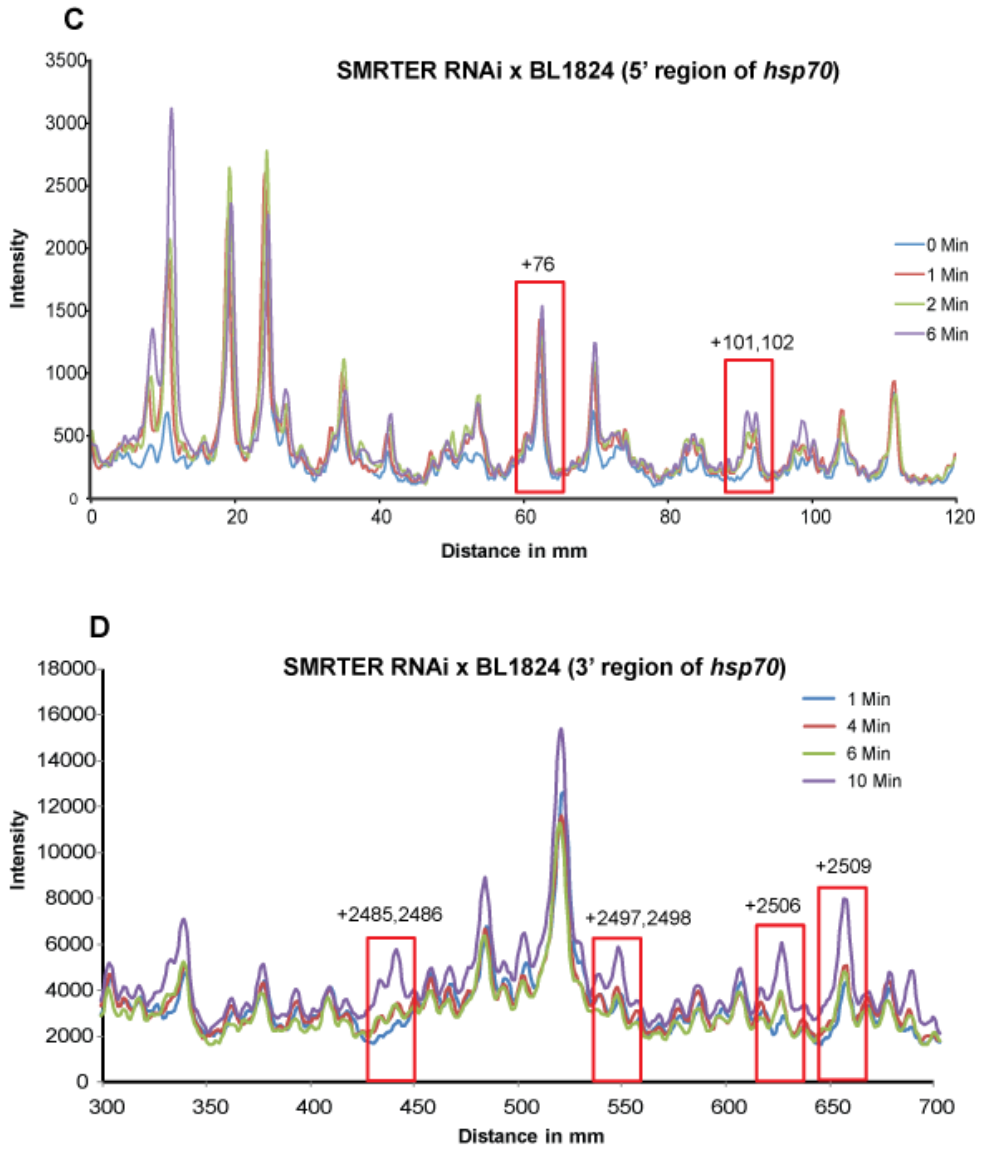


Figure **4-9**: Permanganate footprinting at promoter proximal and 3' end of *hsp70* gene in SMRTER depleted glands.

(A-B) Permanganate footprinting in salivary glands isolated from SMRTER depleted glands (SMRTER x BL1824). Dissected glands were heat shocked for indicated number of minutes; with 0 minutes being non heat shocked condition. LM-PCR was carried out with primers corresponding to the 5'end and 3' end of *hsp70* gene. The gel shown is representative of 2 independent experiments. (C-D) Densitometric traces of intensities of the bands after correction for differences in the background. (C) Intensities of the bands in 5' region of *hsp70* and (D) Intensities of the bands in the 3' region of *hsp70* gene. The densitometric traces were drawn using Imagequant software (GE Healthcare).

4-2.7 HDAC3 is present on bodies of active genes

Recent genome wide data shows HDAC3 to be present in the bodies of genes. This is in contrast to the localization of the other HDACs, notably HDAC1, which is present at the promoter region (Nègre et al., 2011). I obtained the ModENCODE data for genome wide ChIP for HDAC3 and displayed an alignment of this data with Pol II ChIP seq data (Gilchrist et al., 2010) as a heat map (Figure 4-10). HDAC3 (in blue) is present on the genes with Pol II (in red) and is absent from the genes without Pol II. Thus, HDAC3 associates with active genes. Interestingly, HDAC3 appears to be concentrated in the body of the gene while Pol II is concentrated in the promoter region.

There is also a correlation between presence of HDAC3 and H3K36me3 modification in the bodies of genes (Nègre et al., 2011). H3K36 methylation is a histone modification that is associated with elongating Pol II on actively transcribed genes (Kharchenko et al., 2011).

Taken together, the presence of HDAC3 on the bodies of active genes and its correlation with H3K36me3, suggest a role for HDAC3 in elongation.

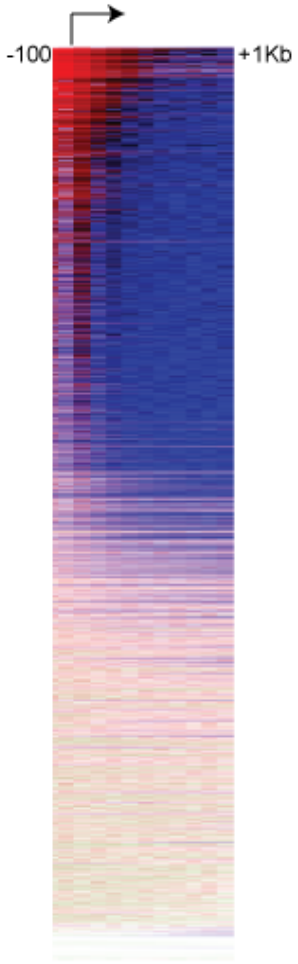


Figure **4-10**: Genome wide localization of HDAC3.

Genome wide ChIP-chip data for HDAC3 was obtained from the modEncode data base, aligned to the TSS, and binned in 100 base pair bins. Genome wide data for Pol II was obtained from Adelman 2010 and ranked based on the sum of intensities in the region of -150 to +150. Data for HDAC3 was aligned against the Pol II data. HDAC3 is shown in blue and Pol II in red.

4-3 Discussion.

4-3.1 Beta-gal screening assay for factors involved in *hsp70* transcription

I used a beta-galactosidase reporter assay in a directed RNAi screen to survey for factors that affect *hsp70* activation. With the availability of RNAi lines targeting thousands of different proteins, it will be possible to screen for RNAi transgenes that impair heat shock induction of an *hsp70* reporter gene. The beta-gal gene fused to the *hsp70* promoter served as an ideal reporter gene, as it allowed for a rapid and easily visualized method to detect defects in expression in the various RNAi lines. However, the assay required some troubleshooting before it was reliable. Initially, I used a reporter gene (D7) that had been used in our laboratory previously that included the region spanning from -194 to +84 of the *hsp70* gene (Wu et al., 2001). The staining results using this reporter were inconsistent as the control yw glands did not always stain well. I tried carrying out the experiment in fat bodies as X-gal staining of fat bodies as had been reported in our lab (Wang et al., 2005), but staining was poor. Finally, I switched to using a reporter gene (Z243), which spans a larger region of the *hsp70* gene (-194 to +250) (Simon and Lis, 1987). This reporter gene gave consistent results and showed good staining of control glands.

As positive controls, I depleted factors that are known to be essential for *hsp70* transcription, such as the activator HSF and subunits of the Pol II kinase, P-TEFb. Depletion of these factors showed decreased expression of the *hsp70*-betagal transgene. Amongst the other factors that have been reported to affect activation of *hsp70*, depletion of elongation factor ELL showed some reduction in expression of the transgene. ELL is an elongation factor, which stimulates elongation by reducing transcriptional pausing in the body of the gene (Shilatifard et al., 1996). In *Drosophila*, ELL localizes to the heat shock puffs during heat shock and depletion of ELL reduced *hsp70* mRNA levels (Gerber et al., 2001) Depletion of chromatin remodeler

Nurf301, which is important in initiation of transcription and binding of HSF to polytene chromosomes after heat shock resulted in a modest decrease in transcription (Badenhorst et al., 2002; Tsukiyama et al., 1994). These results indicate that our screen is able to detect known regulators of the *hsp70* gene.

Treatment of flies with histone deacetylase inhibitor, TSA showed increased expression of heat shock genes in the absence of heat shock (Chen T et al., 2002). *Drosophila* has two nuclear HDACs, Rpd3 and HDAC3, that function as part of co-repressor complexes and regulate gene expression (Gregoretti et al., 2004). I saw no effect of depleting Rpd3 on *hsp70*-beta gal expression. Depletion of Rpd3 led to loss of Rpd3 from polytene chromosomes indicating that the RNAi was efficient (Appendix E).

In contrast, depletion of HDAC3 greatly inhibited expression of the *hsp70*-beta gal reporter gene. HDAC3 was recently shown to affect chromatin organization and its depletion resulted in decreased nucleosome occupancy on the uninduced *hsp70* gene (Petesch and Lis, 2008). Based on these results, I anticipated that HDAC3 would function prior to heat shock to maintain *hsp70* gene in an inactive state. Hence, the finding that depleting HDAC3 or SMRTER reduced expression of the *hsp70*-beta gal transgene was unexpected. The decrease in expression in HDAC3 or SMRTER depleted glands was comparable to depletion of the activators HSF, P-TEFb.

4-3.2 HDAC3 is involved in activation of *hsp70*

The results from the beta-gal screen indicated that depletion of HDAC3 or SMRTER severely impaired expression of the *hsp70*-beta gal transgene. This result was further corroborated by RT-PCR analysis of the endogenous *hsp70* and the beta-gal transgene that showed a significant reduction in levels of mRNA produced in HDAC3 and SMRTER depleted

glands (Figure 4-2, 4-3). Additionally immunostaining of polytene chromosomes from control glands with HDAC3 antibody showed that I could detect HDAC3 at the heat shock puffs (Figure 4-4). These results suggest a role for HDAC3 in expression of the *hsp70* during heat shock.

To further investigate how HDAC3 and SMRTER are involved in regulation of transcription of *hsp70*, I carried out ChIP for Rpb3 a subunit of Pol II at 0 minutes and 10 minutes of heat shock. The ChIP results showed that there is an increase in levels of Pol II in control glands after heat shock. This is in agreement with published results (Ni et al., 2008) and represents an increase in the density of Pol II due to recruitment of additional Pol II upon heat shock. However the levels of Pol II in the promoter and body of the *hsp70* gene in HDAC3 or SMRTER depleted glands upon heat shock is lower than in the control. This suggests that depletion of HDAC3 or SMRTER results in reduced recruitment of Pol II to *hsp70* gene upon heat shock.

Recruitment of Pol II to the heat shock gene promoters is dependent on the binding of the heat shock activator, HSF. Recent research in human cells has shown that the binding of HSF to DNA is influenced by its acetylation status. Acetylation of HSF results in reduced binding to DNA and causes it to dissociate. In human cells, acetylation of HSF is carried out by the HAT p300/CBP and deacetylation is carried out by SIRT1 (Westerheide et al., 2009). Depletion of SIRT1 resulted in reduced binding of HSF and decreased induction of the heat shock genes (Westerheide et al., 2009). It is thought that CBP functions in a similar manner in *Drosophila* as depletion of CBP resulted in HSF being bound to the heat shock gene promoter for a longer period of time (Ghosh et al., 2011). The corresponding deacetylase that functions on HSF in *Drosophila* has not yet been identified. I posit that HDAC3 and SMRTER could function in recruitment of Pol II by counteracting acetylation of HSF by CBP. In the absence of HDAC3 or SMRTER, the equilibrium of HDAC/HAT activity is shifted towards increased acetylation of

HSF by CBP. This could cause a decrease in HSF binding and hence lowered level of Pol II recruitment.

4-3.3 Depletion of HDAC3 or SMRTER leads to inefficient pausing of *hsp70* gene

To evaluate if depletion of HDAC3 or SMRTER also affected the rate of transcription of *hsp70*, I carried out permanganate footprinting on salivary glands at various times after heat shock. Upon heat shock, I could detect Pol II transcribing past the pause site and to the 3' end of the gene at similar time points in the heat shocked control and HDAC3 or SMRTER depleted glands (Figures 4-7, 4-8 and 4-9). This indicates that depletion of HDAC3 or SMRTER does not affect rate of transcription of *hsp70*.

While the permanganate footprinting results at the various time points were similar in HDAC3 depleted and control glands, the ChIP for Pol II showed that under uninduced conditions, pausing of Pol II could be defective in the HDAC3 depleted glands. The higher levels of Pol II in the body of *hsp70* in the HDAC3 depleted glands under non-heat shock conditions could arise from Pol II that is not stably paused at the 5' end of the gene. The discrepancy between the results of the permanganate footprinting experiments and the ChIP assay could be due to inherent differences in the techniques as mentioned earlier. The permanganate footprinting experiment is a sensitive assay that detects the presence of Pol II at single base-pair resolution. The ChIP experiments measure the levels of Pol II over a broad region spanning few hundred nucleotides. It is also possible that the RNAi targeting the different proteins does not deplete the protein equally in all cells. This is evident in the uneven decrease in staining with X-gal in the HDAC3 or SMRTER depleted glands in Figure 4-1. The stochastic nature of the RNAi could also contribute to the differences in the results of the permanganate footprinting and ChIP assays.

The permanganate footprinting experiment in the SMRTER depleted glands at 0 minutes of heat shock shows an increase in reactivity beyond the pause site suggesting that pausing of Pol II is affected in absence of SMRTER. It is possible that this effect in SMRTER depleted glands is not related to HDAC3, as the permanganate footprints in HDAC3 depleted glands are similar to the control. SMRT, the human homolog of SMRTER also interacts with Sin3 repressor complex, which includes histone deacetylase Rpd3 (Nagy et al., 1997). It is possible that in absence of HDAC3, SMRTER functions with a different HDAC to stably pause Pol II on *hsp70*. It has been shown that HDAC3 in mammals requires the co-repressor complex, N-CoR or SMRT to be active *in vivo* (Guenther et al., 2001; You et al., 2013) and hence, in absence of SMRTER, HDAC3 probably does not function at the *hsp70* gene. *Drosophila* has only one nuclear hormone receptor co-repressor complex and the functional homolog of N-CoR and SMRT in *Drosophila* is SMRTER (Tsai et al., 1999).

Additionally, HDAC3 and SMRTER were observed to co-immunoprecipitate with NELF (unpublished result from Anamika Missra, Jian Li, Greg Kothe, and Doug Baumann). Hence HDAC3 might be regulating the activity of NELF, and loss of NELF-activity with the depletion of HDAC3 could result in inefficient pausing.

Additional support for this hypothesis is provided by the even distribution of Pol II that I detect across the *hsp70* gene during heat shock. Pol II pauses on the *hsp70* gene even after induction and release from paused state is a rate limiting step in transcription of this gene (Ni et al., 2008). Pausing of Pol II after activation is evident in the heat shocked yw glands as seen in the persistence of the +65 peak even after heat shock (Figure 4-6B). If recruitment of Pol II was the only step impaired in activation of *hsp70* in HDAC3 or SMRTER depleted glands and pausing occurred normally, the ratio of Pol II in the paused region to the body of the gene would be the same as in control. However, this ratio is decreased in HDAC3 or SMRTER depleted glands. The distribution of Pol II is level across the gene in the absence of HDAC3 or SMRTER.

This would indicate that in the absence of HDAC3 or SMRTER, pausing after activation is reduced and is not a rate-limiting step in release of Pol II into the body of the gene. Thus, HDAC3 and SMRTER appear to be involved in pausing of RNA Pol II on the *hsp70* gene.

4-3.4 HDAC3 and SMRTER could play a dual role in transcription of the *hsp70* gene

My results lead me to hypothesize that HDAC3 has two functions at the *hsp70* gene (Figure 4-11). Firstly, HDAC3 or SMRTER could function to maintain a stably paused Pol II on *hsp70* prior to heat shock. Secondly, HDAC3 or SMRTER could function in recruitment of Pol II after heat shock by affecting the acetylation status of the heat shock activator HSF,

My data suggests that pausing is disturbed in HDAC3 or SMRTER depleted glands. Pausing is considered to be a checkpoint in the transition into productive elongation. It has been suggested that one of the functions of promoter proximal pausing is to allow for co-transcriptional capping of mRNA. The capping enzyme interacts with DSIF, which functions along with NELF to pause Pol II (Pei and Shuman, 2002). Impaired capping of the mRNA would result in destabilization of the mRNA. This could also explain the significant decrease in beta-galactosidase activity observed in HDAC3 or SMRTER depleted glands (Figure 4-1). Upon induction of heat shock, deficient pausing results in decreased dwell time of RNA Pol II at the promoter region. Release from pause site is not rate limiting in HDAC3 or SMRTER depleted glands. Thus, the regulatory steps at the checkpoint are bypassed. When the glands are heat shocked, the transcribed mRNAs are not properly processed, leading to reduced level of transcripts.

In the HDAC3 or SMRTER depleted glands, in addition to decreased pausing, I also see lower level of Pol II recruited to the gene after heat shock. When the organism is subjected to heat stress, the master regulator protein, HSF trimerizes and binds to the heat shock elements in

the promoter of heat shock genes. This binding of HSF, in humans, is regulated by the dual activities of the HDAC, SIRTI and the HAT, CBP. Acetylation of HSF inhibits its DNA binding. In *Drosophila*, depletion of CBP resulted in increased binding of HSF (Ghosh et al., 2011). I propose that HDAC3 and SMRTER complex could be involved in the deacetylation of HSF. This could result in decreased recruitment of Pol II due to decreased levels of HSF bound to the promoter.

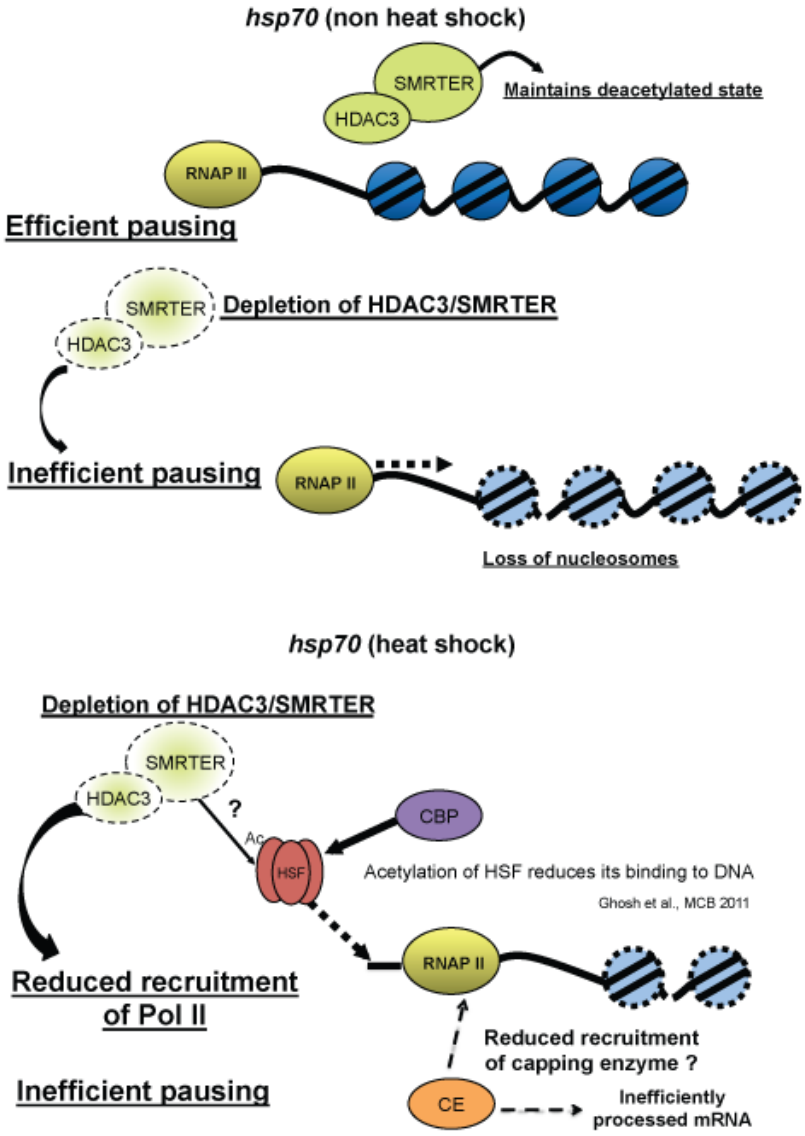


Figure **4-11**: Model for activity of HDAC3 in transcription of *hsp70*

Under non-heat conditions, HDAC3 helps maintain the gene in a repressed state by keeping the H2AK5 residue in nucleosomes in a deacetylated state. HDAC3 and SMRTER help maintain the Pol II in an efficiently paused state, allowing for efficient capping of the 5' transcripts associated with the paused Pol II complex. When the gene is activated, recruitment of Pol II is defective in HDAC3 or SMRTER depleted glands. This could be due to decreased binding of HSF, due to increased acetylation by CBP.

Overall, my results point to a positive role of HDAC3 in transcription of *hsp70* gene.

This proposed positive role of HDAC3 in regulation of *hsp70* expression is in contrast to established roles of histone deacetylases in repression of transcription. However there are exceptions to the rule. For example, histone deacetylase activity is required for recruitment of Pol II in a select group of Interferon β (IFN-β) stimulated genes in human fibroblast cells (Sakamoto et al., 2004). The effect of histone deacetylase inhibitors on reducing recruitment of Pol II to some IFN-β target genes is mediated by IRF-9 (Interferon Regulatory Factor 9), though the exact mechanism is not known (Sakamoto et al., 2004). Similarly histone deacetylase inhibitors affect induction of genes that are targets of the signal transducer and activator STAT5, which plays an important role in cytokine signaling. Histone deacetylase inhibitors prevent induction of STAT5 target genes by preventing recruitment of basal transcription machinery (Rascle et al., 2003). While these early studies used HDAC inhibitors that targeted multiple HDACs, a recent study showed that treatment of human HEK293 cells with HDAC3 specific inhibitor blocked induction of NF-κB p65 responsive genes. HDAC3 functions as a co-activator for induction of these NF-κB p65 target genes by removal of inhibitory acetyl groups from the p65 subunit (Ziesché et al., 2012). Inhibition of HDACs in *Drosophila* S2 cells resulted in decreased expression of select groups of genes thus suggesting a positive role of HDACs in gene expression on those genes, although it was not shown if this was a direct effect (Foglietti et al., 2006). Genome wide ChIP-chip experiments in flies at various stages of development have shown that all 5 HDACs are enriched at active genes and HDAC3 binding sites were enriched in the bodies of transcribed genes that were marked by the histone modification H3K36me3, which is associated with active genes (Nègre et al., 2011). In human cells HDAC3 localizes to the promoter regions of active genes (Wang et al., 2009). These results from genome wide localization studies allude to a role of HDACs in active transcription.

Chapter 5

Discussion

Chromatin is the physiological template for RNA Pol II and plays an influential role in regulation of transcription. My work has focused on studying the influence of chromatin on establishment and release of paused Pol II.

5-1 Factors affecting pausing of Pol II

5-1.1 Nucleosome positioning

Several *in vitro* studies have shown that the nucleosome acts as a strong impediment to the transcribing Pol II ((Izban and Luse, 1991a; Kireeva et al., 2005). Genome wide experiments mapping the nucleosomes in *Drosophila* have highlighted the role of chromatin structure in pausing (Gilchrist et al., 2010; Mavrich et al., 2008a). One model proposed is that the positioned nucleosome acts as a barrier to Pol II, which contributes to pausing (Mavrich et al., 2008a). The second model proposes that the paused Pol II acts as a barrier to nucleosomes. This is suggested by the results that on certain genes in the absence of paused Pol II, nucleosomes move over to the nucleosome positioning sequences in the promoter (Gilchrist et al., 2010). A prediction from the second model would be that a change in the position of paused Pol II further upstream would be accompanied by a change in the nucleosome positions as well. My experiments mapping the position of paused Pol II and the +1 nucleosome in the C4 mutant fly line show that while the position of pausing is shifted upstream in C4 compared to the control, there is no accompanying change in the position of +1 nucleosome.

Additionally to test the model that the position of paused Pol II is influenced by +1 nucleosome, I grouped genes on the basis of position of paused Pol II and analyzed the position of +1 nucleosome in these groups. I observed that on majority of genes, there is no correlation between a change in position of paused Pol II and position of +1 nucleosome, with the exception of the distally paused genes (Figure 3-10). On this group of genes, (Group D in Figure 3-10) where Pol II is paused at approximately 55 nucleotides downstream of the TSS, the position of +1 nucleosome is also shifted downstream compared to genes where Pol II is paused closer to the TSS. This is in accordance with recent reports that suggest that on distally paused genes, the nucleosome plays a greater role in pausing while on the proximally paused genes, factors other than the nucleosome play a major role in pausing Pol II (Kwak et al., 2013; Li and Gilmour, 2013). The results presented in Figure 3-10 can be further improved by using a peak-calling program to better define the position of the transcription bubble and the nucleosome borders and calculate the distance between them. A more rigorous test of the role of +1 nucleosome in pausing of Pol II in the distally paused genes would be to vary the location of +1 nucleosome by engineering a transgene with a nucleosome positioning sequence and determining the location of paused Pol II on these genes.

5-1.2 Additional factors affecting pausing

Amongst the other factors shown to affect pausing are factors such as NELF, DSIF and GAGA factor (Gilchrist et al., 2008; Lee et al., 2008). A number of recent studies have reported that majority of genes with high pausing indices show the presence of GAGA elements at their promoters (Gilchrist et al., 2010; Kwak et al., 2013; Li et al., 2013; Rach et al., 2011). Binding of GAGA factor to the promoters also influences the position of paused Pol II (Kwak et al., 2013; Li et al., 2013). GAGA factor binds NELF allowing NELF to be poised to bind to the elongation

complex (Li et al., 2013). NELF in association with DSIF pauses Pol II by binding to the nascent transcript as it emerges from the elongation complex (Missra and Gilmour, 2010). It is proposed that the location of pausing is controlled by the rate at which NELF and DSIF bind to the elongation complex and the rate of Pol II elongation (Li et al., 2013). In addition to GAGA factor, DNA binding proteins that bind to their cognate elements in the promoter are also shown to influence position of pausing (Kwak et al., 2013). It is also observed that there is a positive correlation between promoter DNA elements located closest to their consensus positions and pausing closer to the TSS (Kwak et al., 2013). I compared the distribution of promoter elements in the different groups defined by position of paused Pol II in WT and C4 embryos in Figure 3-6. I observed no enrichment of any specific promoter element in the groups of genes that showed a shift in location of paused Pol II in C4 embryos compared to WT. I also did not observe any enrichment or depletion of promoter elements in the group of genes that did not show a shift in C4 embryos (Appendix C). The results of these analyses indicate that these sequence elements do not influence the behavior of Pol II in C4 embryos.

The sequence of the transcribed DNA was also thought to influence pausing (Nechaev et al., 2010). The authors proposed that Pol II pauses in a region where the stability of RNA-DNA hybrid was high (Nechaev et al., 2010). However recent analyses of the free energy profiles of the RNA-DNA hybrid in paused genes show no correlation between regions of low free energy change associated with stable RNA-DNA hybrids and the location of paused Pol II (Li et al., 2013). Additionally, Pol II pauses at different locations on the same genes in the two different fly lines analyzed (C4 and WT) indicating that DNA sequence is not the sole determinant of location of pausing (Figure 3-6) (Li et al., 2013).

Although DNA sequence alone does not determine location of pausing, *in vitro* experiments have shown that Pol II pauses at intrinsic pause sites on naked DNA (Kireeva et al., 2005). The C4 polymerase is less efficient in transcribing beyond these intrinsic pause sites (Chen

et al., 1996). The inability of the mutant form of Pol II in C4 embryos to overcome the intrinsic pause sites could also contribute to the change in pausing locations observed in C4 embryos compared to WT. It could also explain the spread of permanganate reactivity observed in C4 embryos (Figure 3-6B).

5-2 Role of chromatin modifiers on pausing

My results show that HDAC3 affects transcription of the *hsp70* gene. However, the target of HDAC3 activity at this gene is not known.

5-2.1 Effect on pausing

HDAC3 depleted glands showed a defect in pausing (Figure 4-6, Chapter 4). Co-immunoprecipitation experiments show that HDAC3 associates with NELF in *Drosophila* nuclear extracts (Anamika Missra, Greg Kothe, Doug Baumann, Jian Li unpublished results). Acetylation of NELF could result in decreasing its activity. If that is true, then depletion of NELF should show decreased pausing and hence increased Pol II downstream of the paused site. Indeed permanganate footprints in NELF depleted cells show a modest shift of Pol II downstream of the paused site (Li et al., 2013). Biochemical experiments with unmodified NELF and acetylated NELF can help further test this hypothesis. Identification of the acetylation sites on NELF and mutating them would help determine their effect on pausing. Additionally, ChIP experiments with NELF antibodies in HDAC3 depleted glands can be carried out to test if HDAC3 activity affects association of NELF with the *hsp70* promoter. Immunofluorescence of polytene chromosomes in control glands will help assess if HDAC3 and NELF co-localize on heat shock gene loci. Further,

if association of NELF is affected by HDAC3, it can also be monitored by immunostaining polytene chromosomes with NELF antibodies in HDAC3 depleted glands.

5-2.2 Effect on release from paused state

Specific histone modifications and combinations thereof could serve as a signal for regulation of the paused state. Several studies have shown acetylation of specific residues on histones to be linked with release from paused state and consequent activation of the gene. It is thought to do so by activation or recruitment of co-activators. For example, a number of studies in human cells have linked acetylation of H4K16 to release of paused Pol II (Zippo et al., 2009; Kapoor-Vazirani et al., 2011). While similar studies have not been reported in *Drosophila*, to the best of my knowledge, it is reasonable to hypothesize that a similar mechanism could play a role on paused genes in *Drosophila*. A recent report has shown increased staining with H4K16ac antibodies in imaginal discs of larvae expressing HDAC3 RNAi (Lv et al., 2012). HDAC3 depletion could result in increased levels of H4K16ac that results in premature release of and hence a destabilized paused Pol II as detected by the higher levels of Pol II in the body of the uninduced *hsp70* gene by ChIP, in the HDAC3 depleted glands (Figure 4-6, Chapter 4). This can be tested by monitoring levels of H4K16 acetylation at the *hsp70* promoter in HDAC3 depleted glands, prior to heat shock.

5-2.3 Effect on activation

A number of studies have shown that histone deacetylases play a positive role in transcription (Smith, 2008). For example, HDAC activity is important in regulation of interferon responsive genes. Inhibition of histone deacetylases resulted in reduced recruitment of RNA Pol

II to the target genes (Nusinzon and Horvath, 2005). Thus, histone deacetylases have been shown to play a positive role in transcription of these genes though their exact targets remain unknown.

My results show that HDAC3 depletion resulted in lowered levels of Pol II at the promoter proximal region of induced *hsp70* gene (Figures 4-6, Chapter 4). A model raised by my results is that HDAC3 functions to deacetylate a factor involved in recruitment of Pol II (Figure 4-11, Chapter 4). A possible candidate is the heat shock factor, HSF. HSF is the master regulator of heat shock gene activation and binding of HSF results in the recruitment of several factors involved in transcription, including Pol II (Zobeck et al., 2010) Acetylation of HSF by the histone acetyl transferase CBP, decreased its binding to the *hsp70* gene (Westerheide et al., 2009). I posit that a function of HDAC3 could be to counteract CBP acetylation, thus increasing the occupancy of HSF at the promoters. To test this model, we could ChIP for HSF in the HDAC3/ SMRTER depleted glands. Based on my hypothesis, I would expect decreased binding of HSF to the *hsp70* promoter upon prolonged heat shock in HDAC3 or SMRTER depleted glands.

Another potential target of HDAC3 is the CTD kinase P-TEFb. Acetylation of lysine residues K44 and K48 in the catalytic domain of Cdk9 reduced its kinase activity and inhibited transcription of a reporter gene controlled by a minimal HIV-1 promoter (Sabò et al., 2008). The acetylated form of Cdk9 is associated with the transcriptionally inactive HIV provirus and the level of acetylated Cdk9 decrease and levels of unacetylated Cdk9 increase upon activation. HDAC3 was shown to associate with P-TEFb (Fu et al., 2007). However, the association of HDAC3 and consequent deacetylation of K44 was shown to inhibit Cdk9 kinase and transcriptional activity (Fu et al., 2007). A potential caveat in this study is that the authors did not show that acetylation of K44 residue increased the association of Cdk9 to the HIV promoter. It has been suggested that the difference in behavior of acetylated Cdk9 in the two published reports could be attributed to the different lysine residues being acetylated and it is possible that

acetylation of K44 alone functions differently than acetylation of both K44 and K48 (Sabò et al., 2008).

The results showing association of HDAC3 with P-TEFb and inhibition of its activity by acetylation of K44, K48 residues in the catalytic core suggest that a potential target of HDAC3 could be P-TEFb. While I observe lower levels of Pol II on the induced *hsp70* gene, the antibody used does not distinguish between the different phosphorylated forms of Pol II. If P-TEFb activity is affected in HDAC3 depleted glands, then the levels of serine 2 phosphorylated Pol II should reflect this. This can be investigated by carrying out a ChIP experiment with serine 2 phosphorylated Pol II specific antibodies in HDAC3 depleted glands after normalization for differences in levels of total Pol II. In addition, ChIP for Cdk9 or CycT subunits of P-TEFb in HDAC3 depleted glands can also address the question if HDAC3 activity affects association of P-TEFb with the *hsp70* gene.

HDAC3 is present in the coding region of active genes (Kharchenko et al., 2011). In *Drosophila*, it could be involved in maintaining the nucleosomes in a deacetylated state on active genes. Hyperacetylation in absence of histone deacetylases could allow for quicker passage of RNA Pol II across the gene, which could affect the activity of other factors associated with elongating Pol II such as those involved in processing of mRNA. However, I do not observe a difference in elongation rates of RNA Pol II on the *hsp70* gene upon induction as measured by permanganate footprinting at various times after heat shock. This could be a reflection of the different behaviors of the pioneering RNA Pol II molecule and the subsequent RNA Pol II molecules. The first wave of RNA Pol II that transcribes through could be more competent in elongation than the subsequent molecules. This could be due to improper modifications of the RNA Pol II that follows after the first one or decreased associations of factors that aid in elongation.

My results suggest that chromatin plays a role in regulation of pausing in an unexpected manner. While the positioned +1 nucleosome does not play a significant role in pausing of Pol II, the factors associated with chromatin like HDAC3 could play a role in regulation of pausing, at the stage of release from the paused state.

Appendix A

Quantification of Permanganate footprinting

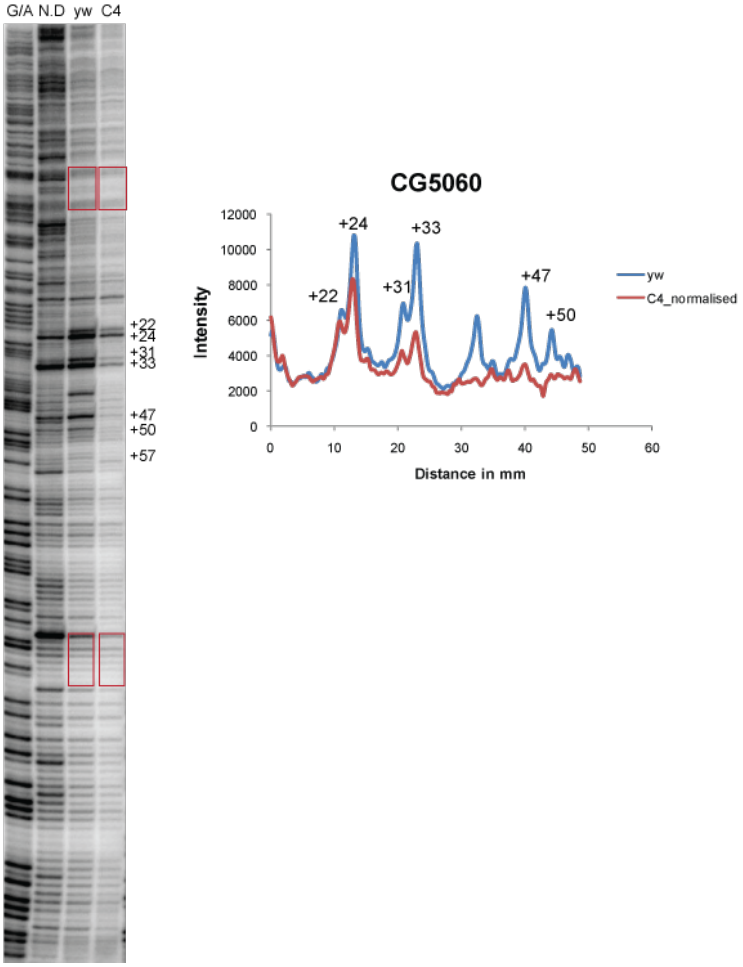


Figure **A-1**: Permanganate footprinting and quantification of permanganate reactivity on *CG5060.* The panel on the left shows permanganate footprinting carried out on *CG5060*. The red boxes indicate the regions I used to quantify differences in background using ImageQuant software. I analyzed two different regions of the gel as indicated and obtained similar differences in the background intensity of lanes corresponding to yw and C4. I chose regions of the gene that did not show the presence of footprints or were similar in yw, C4 and naked DNA lanes. The panel on the right shows a trace of the intensities after normalization for differences in background.

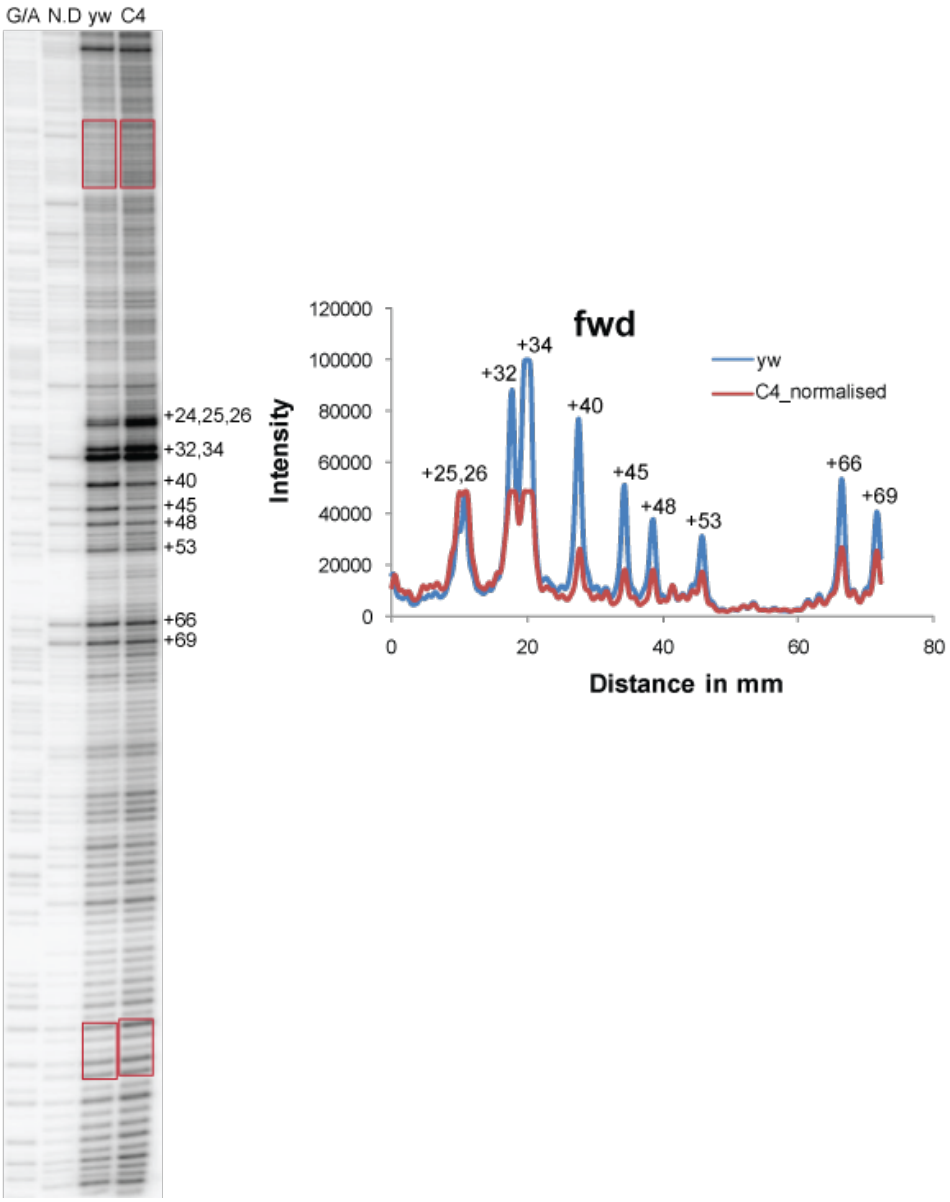


Figure **A-2**: Permanganate footprinting and quantification of permanganate reactivity on *fwd*. The panel on the left shows permanganate footprinting carried out on *fwd*. The red boxes indicate the regions I used to quantify differences in background using ImageQuant software. I analyzed two different regions of the gel as indicated and obtained similar differences in the background intensity of lanes corresponding to yw and C4. The panel on the right shows a trace of the intensities after normalization for differences in background.

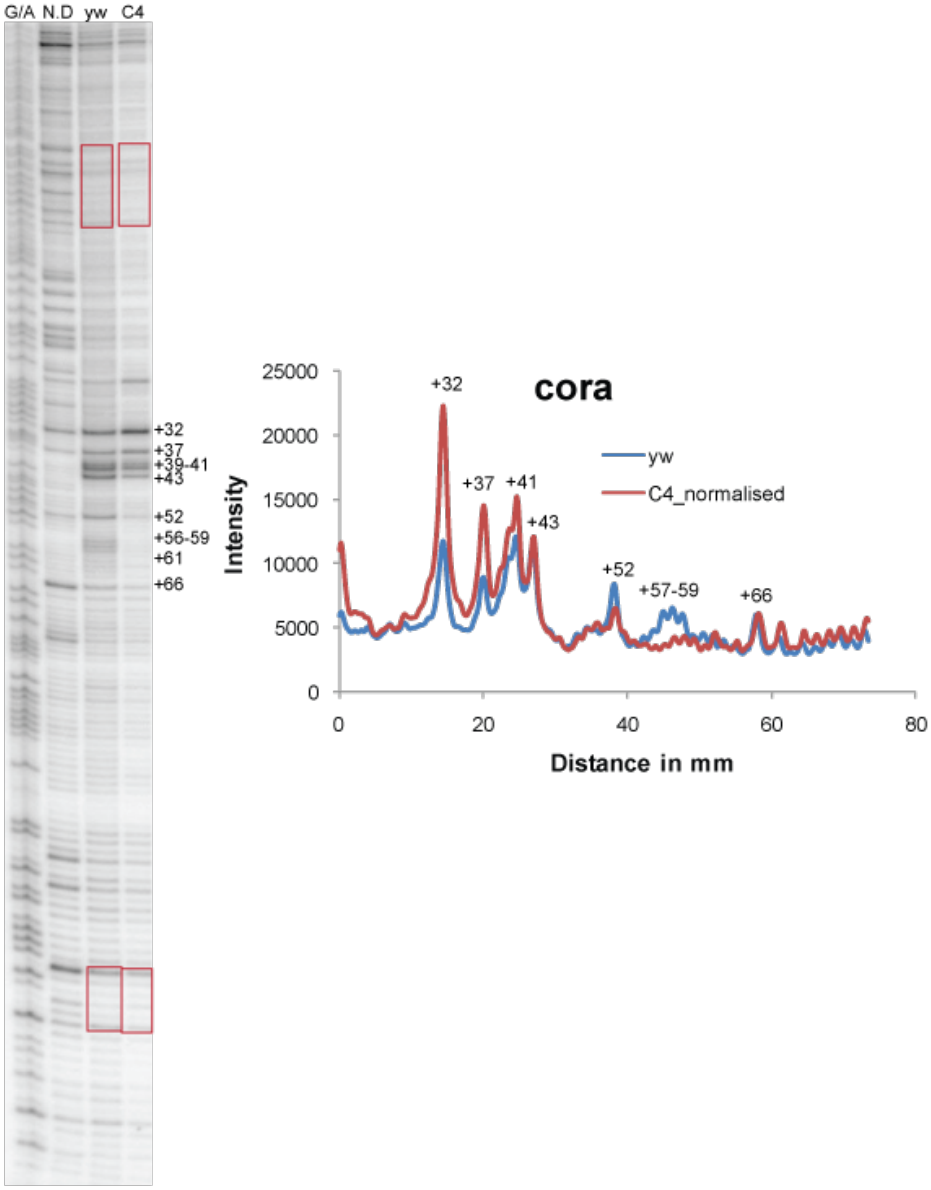


Figure **A-3**: Permanganate footprinting and quantification of permanganate reactivity on *cora.* The panel on the left shows permanganate footprinting carried out on *cora* . The red boxes indicate the regions I used to quantify differences in background using ImageQuant software. I analyzed two different regions of the gel as indicated and obtained similar differences in the background intensity of lanes corresponding to yw and C4. The panel on the right shows a trace of the intensities after normalization for differences in background.

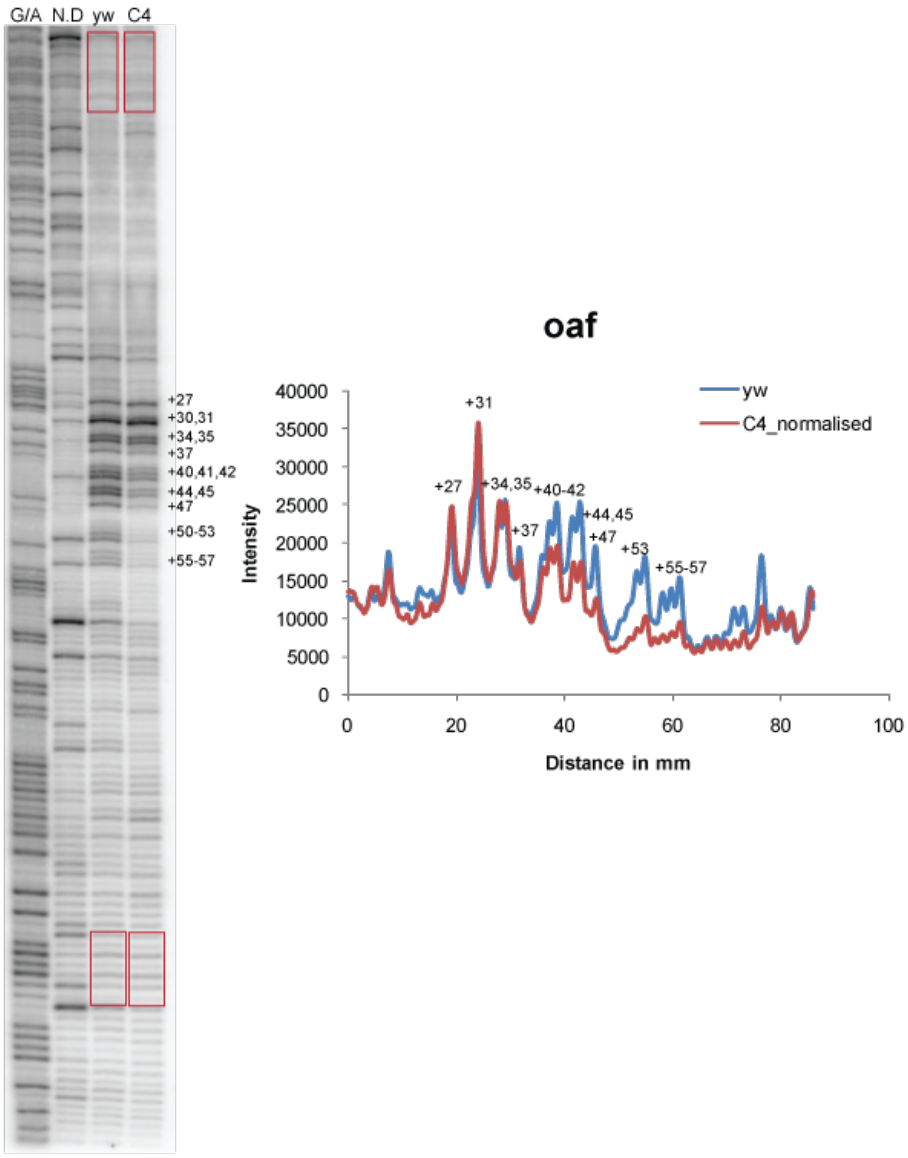


Figure **A-4**: Permanganate footprinting and quantification of permanganate reactivity on *oaf*. The panel on the left shows permanganate footprinting carried out on *oaf* . The red boxes indicate the regions I used to quantify differences in background using ImageQuant software. I analyzed two different regions of the gel as indicated and obtained similar differences in the background intensity of lanes corresponding to yw and C4. The panel on the right shows a trace of the intensities after normalization for differences in background.

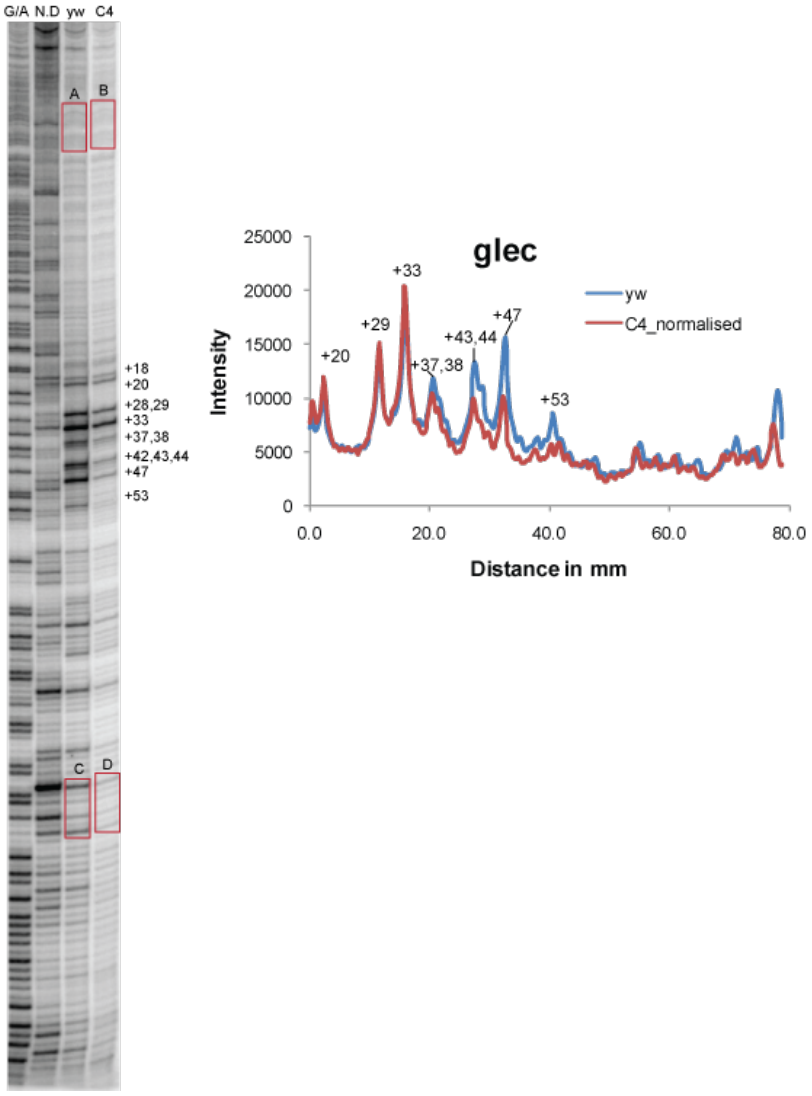


Figure **A-5**: Permanganate footprinting and quantification of permanganate reactivity on *glec*
The panel on the left shows permanganate footprinting carried out on *glec*. The red boxes indicate the regions I used to quantify differences in background using ImageQuant software. I analyzed two different regions of the gel as indicated and obtained similar differences in the background intensity of lanes corresponding to yw and C4. The panel on the right shows a trace of the intensities after normalization for differences in background.

Appendix B

Nucleosome position on unfiltered group of genes in WT and C4 embryos

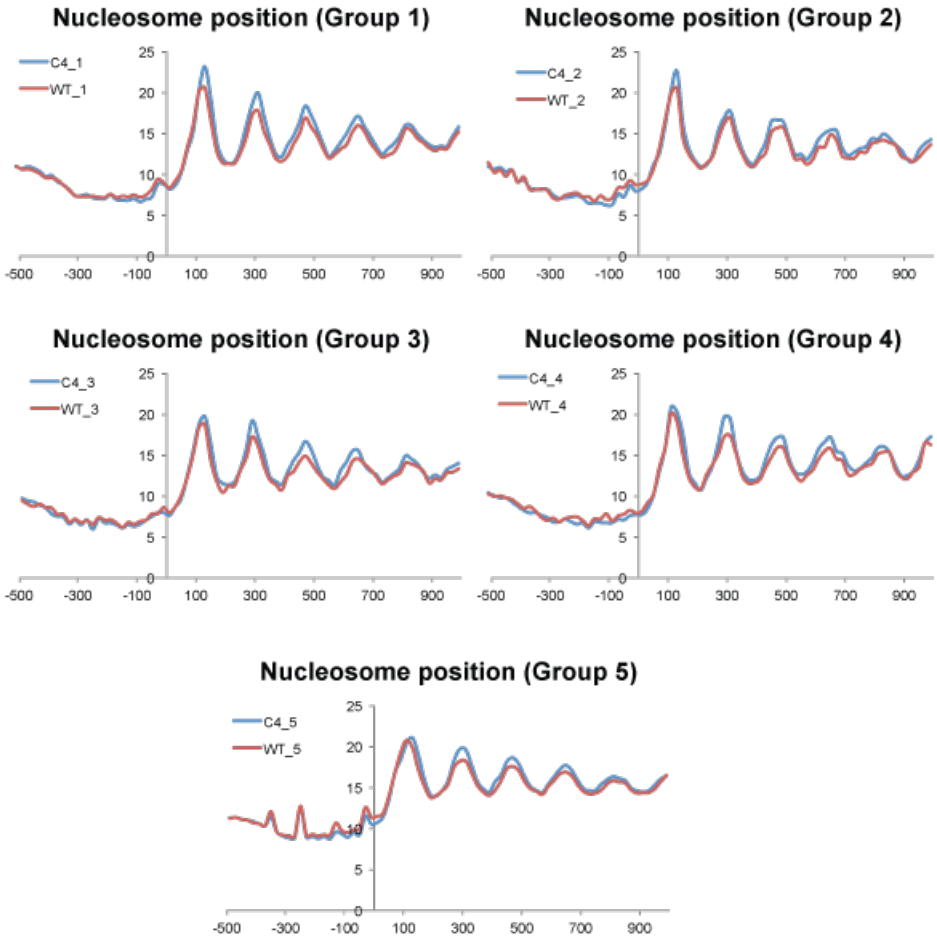


Figure **B-1**: Nucleosome positions on unfiltered genes in WT and C4 embryos.

Composite plot of nucleosome dyad positions on individual groups as defined in Figure 3-5. The reads are mapped in 20 basepair bins and aligned to the TSS. The x-axis is the distance from TSS and the y-axis is number of reads/number of genes in each group.

Appendix C

Distribution of promoter elements on paused genes in WT and C4 embryos

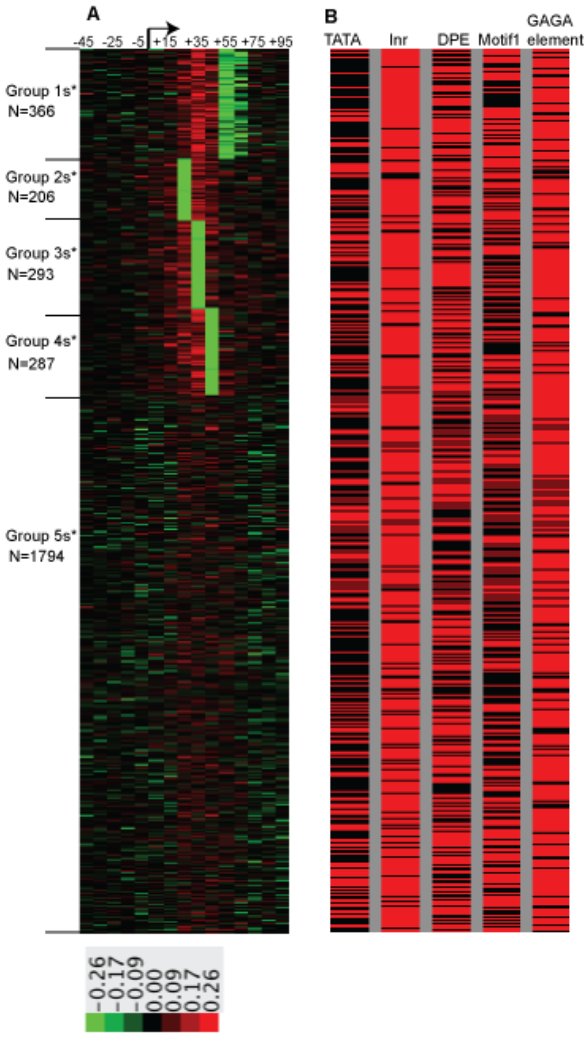


Figure **C-1**: Distribution of promoter elements on groups of genes defined by difference in location of paused Pol II in WT and C4 embryos.

Panel A is a treeview display of the difference in permanganate reactivity on genes with significant levels of Pol II above background, in WT and C4 embryos, similar to Figure 3-6. Panel B is a display of the distribution of different promoter elements associated with the genes in the different groups shown in Panel A.

Appendix D

HDAC3 depletion in S2 cells

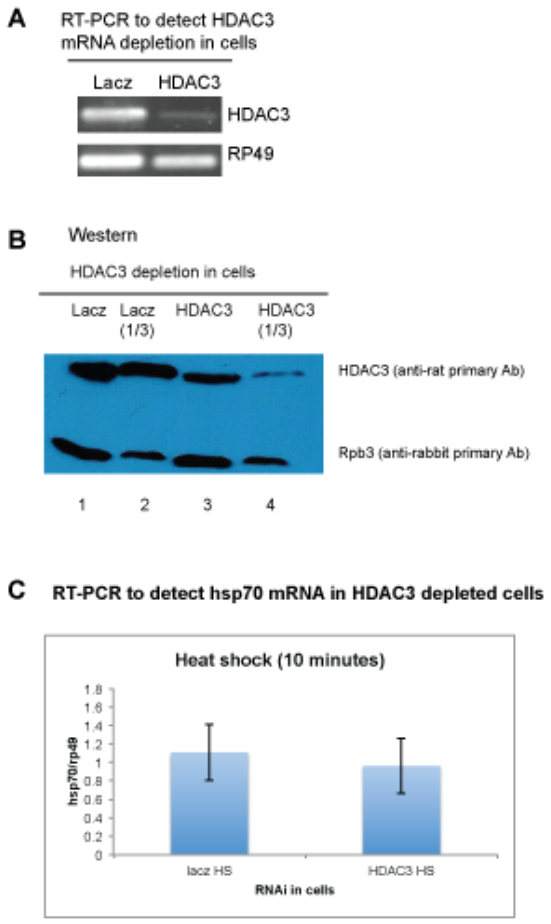


Figure **D-1**: Depletion of HDAC3 in S2 cells.

A. S2 cells were treated with 25 μg of dsRNA against HDAC3 for 5 days. Cells were treated with dsRNA against LacZ as a control. RT-PCR with HDAC3 specific primers was carried out to monitor depletion of the protein. RP49 served as an internal standard. **B.** Western blot with HDAC3 specific antibodies was carried out on protein lysates obtained from control and HDAC3

depleted S2 cells. Rpb3 was used as a loading control. Lanes 2 and 4 had 0.3 % of the amount of protein loaded in the lanes 1 and 3. **C.** RT-PCR analysis of *hsp70* mRNA levels in control and HDAC3 depleted cells. Cells were heat shocked at 37^0C for 10 minutes. Total RNA was isolated followed by cDNA synthesis. *hsp70* was detected by qPCR and expressed as a ratio to *rp49*, which serves as an internal standard. Results are from two independent experiments and error bars represent range of values.

Appendix E

Immunostaining of polytene chromosomes in Rpd3 depleted glands

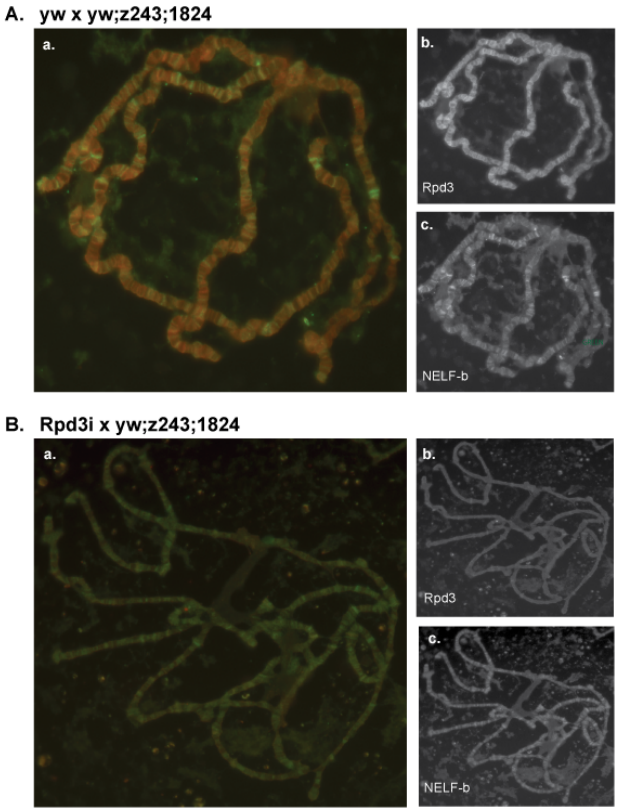


Figure **E-1**: Loss of Rpd3 staining on polytene chromosomes in Rpd3 depleted glands. **(A)** Polytene chromosomes from salivary glands of control (yw x yw; Z243,1824) third instar larvae **(a)** Overlaid image of staining for Rpd3 and NELF-b, subunit of NELF. Panels to the right show individual staining for Rpd3 **(b)** and NELF-b **(c)**. **(B)** Polytene chromosomes from salivary glands of Rpd3-depleted (Rpd3 x yw;Z243,1824) third instar larvae. **(a)** Overlaid image of staining for Rpd3 and NELF-b in Rpd3 x yw;Z243,1824 larvae. Panels to the right show individual staining for Rpd3 **(b)** and NELF-b **(c)**. Rpd3 was stained with an anti-rabbit primary antibody at 1:100 dilution. NELF-b was stained with anti-guinea pig primary antibody at a 1:10 dilution. The primary antibodies were visualized with an Alexa-488 conjugated anti-guinea pig and Alexa- 568 conjugated anti-rabbit antibody.

Table **1-1**: List of LM-PCR primers

Gene name	Primer name	Sequence	Tm	Annealing temperature
glec	glec+197R_LM1	5'-GTACGCATTTTAGGCATTG	54.4	51.0
glec	glec+188R_LM2	5'-TTAGGCATTGCAATTGCAG	58.4	55.0
glec	glec+178R_LM3	5'-CAATTGCAGTTTAAATCGGTTTTCT	61.7	58.0
fwd	fwd +198R_LM1	5'-ACATCGTCGATACACAAGG	54.3	51.0
fwd	fwd +186R_LM2	5'-ACAAGGGTGCACGAATGA	58.6	55.0
fwd	fwd +173R_LM3	5'-AATGAAGCGGCGCGTACT	61.9	58.0
cora	cora+188R_LM1	5'-ATTTTATTTATAAACACACTGCTG	52.7	51.0
cora	cora+179R_LM2	5'-ATAAACACACTGCTGCCGAAT	59.7	55.0
cora	cora+173R_LM3	5'-ACACTGCTGCCGAATTGC	61.0	58.0
CG5060	CG5060+195R_LM1	5'-TTGTATTTGCAAACGAATTTA	53.8	51.0
CG5060	CG5060+185R_LM2	5'-AAACGAATTTAATTTGCGACAC	58.3	55.0
CG5060	CG5060+173R_LM3	5'-TTTGCGACACCTAAAAATCGAA	61.8	58.0

Table **1-2**: List of PCR primers.

Gene name	Primer name	Sequence	Tm
hsp70	hsp70Bc+1916F	5’-ACTGGACGAGGCCGACAAG	60
hsp70	hsp70Bc+2036R	5’-GCGAGTGAGCTCCTCCATCT	59
hsp70	87C hsp70 +16/+115 F	5’-GCAAAGTGAACACGTCGCTAAG	56
hsp70	87C hsp70 +16/+115 R	5’-ATTGATTCACTTTAACTTGCACTTTACTG	54
rp49	RealT Rp49 F	5’-GCGTCGCCGCTTCAAG	58
rp49	RealT Rp49 R	5’-CAGCTCGCGCACGTTGT	59
rp49	rp49-A	5'-TACAGGCCCAAGATCGTGAA	56
rp49	rp49-B	5'-ACGTTGTGCACCAGGAACTT	57
hsp70	87C hsp70-825/897 F	5'-GGTGAGCGCAATGTGCTTATC	56
hsp70	87C hsp70 825/897 R	5'-AGCGCACCTCGAACAGAGAT	58

Table **1-3**: List of sequencing primers.

Primer	Sequence	Tm
Exa2_NoP_sense	5’ /Phos/AGACGTGTGCTCTTCCGATC*T	
Exa2_NoP_antisense	5’-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC	
ExA1_NoP_sense	5'GGACACTCTTTCCCTACACGACGCTCTTCCGATC*T	
ExA1_NoP_antisense	5’GATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT	
P7 Primer	5’-GAGCATACGGCAGAAGACGAAC	58
P1.2 Primer	5’AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGC	59
Index I-06	5’CAAGCAGAAGACGGCATACGAGAT**CTTGTA**GTGACTGGAGTTC	
Index I-07	5’CAAGCAGAAGACGGCATACGAGAT**GATCAG**GTGACTGGAGTTC	
Index I-08	5’CAAGCAGAAGACGGCATACGAGAT**GCCAAT**GTGACTGGAGTTC	58
Index I-09	5’CAAGCAGAAGACGGCATACGAGAT**GGCTAC**GTGACTGGAGTTC	59
Index I-10	5’CAAGCAGAAGACGGCATACGAGAT**TAGCTT**GTGACTGGAGTTC	56
Index I-11	5’CAAGCAGAAGACGGCATACGAGAT**TGACCA**GTGACTGGAGTTC	57
Index I-12	5’CAAGCAGAAGACGGCATACGAGAT**TTAGGC**GTGACTGGAGTTC	55
Index I-13	5’CAAGCAGAAGACGGCATACGAGAT**AAACAT**GTGACTGGAGTTC	56

Table **1-4**: List of fly-lines.

RNAi fly line	Stock number
HSF	VDRC37699
Cdk9	VDRC 30449
CycT	VDRC 37562
Nurf301	VDRC 4665
HDAC3	VDRC-KK107073
HDAC3_2	VDRC 20184
SMRTER	VDRC-KK 106701
Rpd3	VDRC 30600

References

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II : emerging roles in metazoans. Nat. Rev. Genet. *13*, 720–731.

Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C., and Pugh, B.F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. Nature *446*, 572–576.

Albert, T., Mautner, J., Funk, J.O., Hörtnagel, K., Pullner, A., and Eick, D. (1997). Nucleosomal structures of c-myc promoters with transcriptionally engaged RNA polymerase II. Mol. Cell. Biol. *17*, 4363–4371.

Almer, A., Rudolph, H., Hinnen, A., and Hörz, W. (1986). Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. EMBO J. *5*, 2689–2696.

Andrulis, E.D., Guzmán, E., Döring, P., Werner, J., and Lis, J.T. (2000). High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. Genes Dev. *14*, 2635–2649.

Ardehali, M.B., Yao, J., Adelman, K., Fuda, N.J., Petesch, S.J., Webb, W.W., and Lis, J.T. (2009). Spt6 enhances the elongation rate of RNA polymerase II in vivo. EMBO J. *28*, 1067–1077.

Ardehali, M.B., Mei, A., Zobeck, K.L., Caron, M., Lis, J.T., and Kusch, T. (2011). Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. EMBO J. *30*, 2817–2828.

Badenhorst, P., Voas, M., Rebay, I., and Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. Genes Dev. *16*, 3186–3198.

Bai, L., Shundrovsky, A., and Wang, M.D. (2004). Sequence-dependent kinetic model for transcription elongation by RNA polymerase. J. Mol. Biol. *344*, 335–349.

Benjamin, L.R., and Gilmour, D.S. (1998). Nucleosomes are not necessary for promoter-proximal pausing in vitro on the Drosophila hsp70 promoter. Nucleic Acids Res. *26*, 1051–1055.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. Curr. Protoc. Mol. Biol. Ed. by Frederick M Ausubel Al *Chapter 19*, Unit 19.10.1–21.

Bochkareva, A., Yuzenkova, Y., Tadigotla, V.R., and Zenkin, N. (2011). Factor-independent transcription pausing caused by recognition of the RNA – DNA hybrid sequence. EMBO J. *31*, 630–

639.

Boehm, A.K., Saunders, A., Werner, J., and Lis, J.T. (2003). Transcription Factor and Polymerase Recruitment, Modification, and Movement on dhsp70 In Vivo in the Minutes following Heat Shock. Mol. Cell. Biol. *23*, 7628–7637.

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

Brown, S.A., Imbalzano, A.N., and Kingston, R.E. (1996). Activator-dependent regulation of transcriptional pausing on nucleosomal templates. Genes Dev. *10*, 1479–1490.

Cai, W., Bao, X., Deng, H., Jin, Y., Girton, J., Johansen, J., and Johansen, K.M. (2008). RNA polymerase II-mediated transcription at active loci does not require histone H3S10 phosphorylation in Drosophila. Dev. Cambridge Engl. *135*, 2917–2925.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M.P., et al. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592.

Champlin, D.T., Frasch, M., Saumweber, H., and Lis, J.T. (1991). Characterization of a Drosophila protein associated with boundaries of transcriptionally active chromatin. Genes Dev. *5*, 1611–1621.

Chen, Y., Chafin, D., Price, D.H., and Greenleaf, A.L. (1996). Drosophila RNA polymerase II mutants that affect transcription elongation. J. Biol. Chem. *271*, 5993–5999.

Cheng, B., and Price, D.H. (2007). Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. J. Biol. Chem. *282*, 21901–21912.

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

Coppola, J.A., Field, A.S., and Luse, D.S. (1983). Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. Proc. Natl. Acad. Sci. U. S. A. *80*, 1251–1255.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science. *322*, 1845–1848.

Coulter, D.E., and Greenleaf, A.L. (1985). A mutation in the largest subunit of RNA polymerase II alters RNA chain elongation in vitro. J. Biol. Chem. *260*, 13190–13198.

Deckert, J., and Struhl, K. (2002). Targeted recruitment of Rpd3 histone deacetylase represses transcription by inhibiting recruitment of Swi/Snf, SAGA, and TATA binding protein. Mol. Cell. Biol. *22*, 6458–6470.

Dovey, O.M., Foster, C.T., and Cowley, S.M. (2010). Emphasizing the positive: A role for histone deacetylases in transcriptional activation. Cell Cycle *9*, 2700–2701.

Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998). Characterization of a human RPD3 ortholog, HDAC3. Proc. Natl. Acad. Sci. U. S. A. *95*, 2795–2800.

Fernandes, M., Xiao, H., and Lis, J.T. (1995). Binding of heat shock factor to and transcriptional activation of heat shock genes in Drosophila. Nucleic Acids Res. *23*, 4799–4804.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212–224.

Foglietti, C., Filocamo, G., Cundari, E., De Rinaldis, E., Lahm, A., Cortese, R., and Steinkühler, C. (2006). Dissecting the biological functions of Drosophila histone deacetylases by RNA interference and transcriptional profiling. J. Biol. Chem. *281*, 17968–17976.

Fu, J., Yoon, H.-G., Qin, J., and Wong, J. (2007). Regulation of P-TEFb Elongation Complex Activity by CDK9 Acetylation. Mol. Cell. Biol. *27*, 4641–4651.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004a). Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Mol. Cell. Biol. *24*, 787–795.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004b). Dynamics of Human Immunodeficiency Virus Transcription : P-TEFb Phosphorylates RD and Dissociates Negative Effectors from the Transactivation Response Element. Mol. Cell. Biol. *24*, 787–795.

Gelbart, M.E., and Kuroda, M.I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. Dev. Cambridge Engl. *136*, 1399–1410.

Gerber, M., Ma, J., Dean, K., Eissenberg, J.C., and Shilatifard, A. (2001). Drosophila ELL is associated with actively elongating RNA polymerase II on transcriptionally active sites in vivo. EMBO J. *20*, 6104–6114.

Ghosh, S.K.B., Missra, A., and Gilmour, D.S. (2011). Negative elongation factor accelerates the rate at which heat shock genes are shut off by facilitating dissociation of heat shock factor. Mol. Cell. Biol. *31*, 4232–4243.

Giardina, C., Pérez-Riba, M., and Lis, J.T. (1992). Promoter melting and TFIID complexes on Drosophila genes in vivo. Genes Dev. *6*, 2190–2200.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., et al. (2005). Galaxy: A platform for interactive large-scale genome analysis. Genome Res. *15*, 1451–1455.

Gilchrist, D. A, and Adelman, K. (2012). Coupling polymerase pausing and chromatin landscapes for precise regulation of transcription. Biochim. Biophys. Acta. *1819 (7)*, 700-706.

Gilchrist, D. a, Nechaev, S., Lee, C., Ghosh, S.K.B., Collins, J.B., Li, L., Gilmour, D.S., and Adelman, K. (2008). NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev. *22*, 1921–1933.

Gilchrist, D. a, Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540–551.

Gilmour, D.S., and Fan, R. (2009). Detecting transcriptionally engaged RNA polymerase in eukaryotic cells with permanganate genomic footprinting. Methods San Diego Calif *48*, 368–374.

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

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science. *292*, 1876–1882.

Goecks, J., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. *11*, R86.

Greenleaf, A.L., Weeks, J.R., Voelker, R.A., Ohnishi, S., and Dickson, B. (1980). Genetic and biochemical characterization of mutants at an RNA polymerase II locus in D. melanogaster. Cell *21*, 785–792.

Gregoretti, I. V, Lee, Y.-M., and Goodson, H. V (2004). Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. J. Mol. Biol. *338*, 17–31.

Guenther, M.G., Barak, O., and Lazar, M.A. (2001). The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. Mol. Cell. Biol. *21*, 6091–6101.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Guertin, M.J., and Lis, J.T. (2010). Chromatin Landscape Dictates HSF Binding to Target DNA Elements. PLoS Genet. *6*, 15.

Ivaldi, M.S., Karam, C.S., and Corces, V.G. (2007). Phosphorylation of histone H3 at Ser10 facilitates RNA polymerase II release from promoter-proximal pausing in Drosophila. Genes Dev. *21*, 2818–2831.

Izban, M.G., and Luse, D.S. (1991a). Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. Genes Dev. *5*, 683–696.

Izban, M.G., and Luse, D.S. (1991b). Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. Genes Dev. *5*,

683–696.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. Nat. Rev. Genet. *10*, 161–172.

Johnson, C.A., Barlow, A.L., and Turner, B.M. (1998). Molecular cloning of Drosophila melanogaster cDNAs that encode a novel histone deacetylase dHDAC3. Gene *221*, 127–134.

Jones, K. A. (1997). Taking a new TAK on Tat transactivation. Genes Dev. *11*, 2593–2599.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362–366.

Kapoor-Vazirani, P., Kagey, J.D., and Vertino, P.M. (2011). SUV420H2-mediated H4K20 trimethylation enforces RNA polymerase II promoter-proximal pausing by blocking hMOF-dependent H4K16 acetylation. Mol. Cell. Biol. *31*, 1594–1609.

Karam, C.S., Kellner, W. A, Takenaka, N., Clemmons, A.W., and Corces, V.G. (2010). 14-3-3 mediates histone cross-talk during transcription elongation in Drosophila. PLoS Genet. *6*, e1000975.

Kharchenko, P. V, Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480–485.

Kireeva, M.L., Hancock, B., Cremona, G.H., Walter, W., Studitsky, V.M., and Kashlev, M. (2005). Nature of the nucleosomal barrier to RNA polymerase II. Mol. Cell *18*, 97–108.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nat. Genet. *41*, 376–381.

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

Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950–953.

Lee, C., Li, X., Hechmer, A., Eisen, M., Biggin, M.D., Venters, B.J., Jiang, C., Li, J., Pugh, B.F., and Gilmour, D.S. (2008). NELF and GAGA Factor are linked to promoter-proximal pausing at many genes in Drosophila. Mol. Cell. Biol. *28*, 3290–3300.

Lee, H., Kraus, K.W., Wolfner, M.F., and Lis, J.T. (1992). DNA sequence requirements for generating paused polymerase at the start of hsp70. Genes Dev. *6*, 284–295.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics *26*, 589–595.

Li, J., and Gilmour, D.S. (2013). Distinct mechanisms of transcriptional pausing orchestrated by

GAGA factor and M1BP, a novel transcription factor. EMBO J. *32*, 1829–1841.

Li, B., Carey, M., and Workman, J.L. (2007a). The role of chromatin during transcription. Cell *128*, 707–719.

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

Li, J., Liu, Y., Rhee, H.S., Ghosh, S.K.B., Bai, L., Pugh, B.F., and Gilmour, D.S. (2013). Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. Mol. Cell *50*, 711–722.

Van Lint, C., Emiliani, S., Ott, M., and Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. EMBO J. *15*, 1112–1120.

Lis, J., and Wu, C. (1993). Protein Traffic on the Heat Shock Promoter : Parking , Stalling , and Trucking Along. Cell *74*, 1–4.

Lis, J., Mason, P., Peng, J., and Price, D. (2000). P-TEFb kinase recruitment and function at heat shock loci. Genes Dev. *14*, 792–803.

Lis, J.T., Neckameyer, W., Dubensky, R., and Costlow, N. (1981). Cloning and characterization of nine heat-shock-induced mRNAs of Drosophila melanogaster. Gene *15*, 67–80.

Lu, Q., Wallrath, L.L., and Elgin, S.C. (1995). The role of a positioned nucleosome at the Drosophila melanogaster hsp26 promoter. EMBO J. *14*, 4738–4746.

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

Lunde, B.M., Reichow, S.L., Kim, M., Suh, H., Leeper, T.C., Yang, F., Mutschler, H., Buratowski, S., Meinhart, A., and Varani, G. (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. Nat. Struct. Mol. Biol. *17*, 1195–1201.

Lv WW, Wei HM, Wang DL, Ni JQ, S.F. (2012). Depletion of histone deacetylase 3 antagonizes PI3K-mediated overgrowth through the acetylation of histone H4 at lysine 16. J Cell Sci. *125*, 5369-5378.

Mandal, S.S., Chu, C., Wada, T., Handa, H., Shatkin, A.J., and Reinberg, D. (2004). Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *101*, 7572–7577.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J. Biol. Chem. *270*, 12335–12338.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C., et al. (2008a). Nucleosome organization in the Drosophila genome.

Nature *453*, 358–362.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008b). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res. *18*, 1073–1083.

McGhee, J.D., and von Hippel, P.H. (1975a). Formaldehyde as a probe of DNA structure. I. Reaction with exocyclic amino groups of DNA bases. Biochemistry *14*, 1281–1296.

McGhee, J.D., and von Hippel, P.H. (1975b). Formaldehyde as a probe of DNA structure. II. Reaction with endocyclic imino groups of DNA bases. Biochemistry *14*, 1297–1303.

Missra, A., and Gilmour, D.S. (2010). Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the Drosophila RNA polymerase II transcription elongation complex. Proc. Natl. Acad. Sci. U. S. A. *107*, 11301–11306.

Murawska, M., Hassler, M., Renkawitz-Pohl, R., Ladurner, A., and Brehm, A. (2011). Stress-Induced PARP Activation Mediates Recruitment of Drosophila Mi-2 to Promote Heat Shock Gene Expression. PLoS Genet. *7*, 15.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511.

Nagy, L., Kao, H.Y., Chakravarti, D., Lin, R.J., Hassig, C.A., Ayer, D.E., Schreiber, S.L., and Evans, R.M. (1997). Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell *89*, 373–380.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science *327*, 335–338.

Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the Drosophila genome. Nature *471*, 527–531.

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

Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.-R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. Mol. Cell. Biol. *28*, 1161–1170.

Nusinzon, I., and Horvath, C.M. (2005). Unexpected roles for deacetylation in interferon- and cytokine-induced transcription. J. Interferon Cytokine Res. *25*, 745–748.

Orphanides, G., and Reinberg, D. (2000). RNA polymerase II elongation through chromatin. Nature *407*, 471–475.

Orphanides, G., Leroy, G., Chang, C., Luse, D.S., and Reinberg, D. (1998). FACT , a Factor that Facilitates Transcript Elongation through Nucleosomes. Cell *92*, 105–116.

Ozsolak, F., Song, J.S., Liu, X.S., and Fisher, D.E. (2007). High-throughput mapping of the chromatin structure of human promoters. Nat. Biotechnol. *25*, 244–248.

Park, J.M., Werner, J., Kim, J.M., Lis, J.T., and Kim, Y.J. (2001). Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. Mol. Cell *8*, 9–19.

Pei, Y., and Shuman, S. (2002). Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. J. Biol. Chem. *277*, 19639–19648.

Petesch, S.J., and Lis, J.T. (2008). Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. Cell *134*, 74–84.

Petesch, S.J., and Lis, J.T. (2012). Activator-induced spread of poly(ADP-ribose) polymerase promotes nucleosome loss at Hsp70. Mol. Cell *45*, 64–74.

Price, D.H., 2000. Minireview P-TEFb , a Cyclin-Dependent Kinase Controlling Elongation by RNA Polymerase II. Mol. Cell. Biol., *20*, 2629–2634.

Qi, D., Bergman, M., Aihara, H., Nibu, Y., and Mannervik, M. (2008). Drosophila Ebi mediates Snail-dependent transcriptional repression through HDAC3-induced histone deacetylation. EMBO J. *27*, 898–909.

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

Rabindran, S.K., Haroun, R.I., Clos, J., Wisniewski, J., and Wu, C. (1993). Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science *259*, 230–234.

Rach, E. A Winter, D.R., Benjamin, A.M., Corcoran, D.L., Ni, T., Zhu, J., and Ohler, U. (2011). Transcription initiation patterns indicate divergent strategies for gene regulation at the chromatin level. PLoS Genet. *7*, e1001274.

Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome positioning: how is it established, and why does it matter? Dev. Biol. *339*, 258–266.

Rascle, A., Johnston, J.A., and Amati, B. (2003). Deacetylase Activity Is Required for Recruitment of the Basal Transcription Machinery and Transactivation by STAT5 Deacetylase Activity Is Required for Recruitment of the Basal Transcription Machinery and Transactivation by STAT5. Mol. Cell. Biol. *23*, 4162–4173.

Regnard, C., Straub, T., Mitterweger, A., Dahlsveen, I.K., Fabian, V., and Becker, P.B. (2011). Global analysis of the relationship between JIL-1 kinase and transcription. PLoS Genet. *7*, e1001327.

Renner, D.B., Yamaguchi, Y., Wada, T., Handa, H., and Price, D.H. (2001). A highly purified RNA polymerase II elongation control system. J. Biol. Chem. *276*, 42601–42609.

Reynolds, N., O'Shaughnessy, A., and Hendrich, B. (2013). Transcriptional repressors: multifaceted regulators of gene expression. Development *140*, 505–512.

Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell *54*, 795–804.

Sabò, A., Lusic, M., Cereseto, A., and Giacca, M. (2008). Acetylation of conserved lysines in the catalytic core of cyclin-dependent kinase 9 inhibits kinase activity and regulates transcription. Mol. Cell. Biol. *28*, 2201–2212.

Sakamoto, S., Potla, R., and Larner, A.C. (2004). Histone deacetylase activity is required to recruit RNA polymerase II to the promoters of selected interferon-stimulated early response genes. J. Biol. Chem. *279*, 40362–40367.

Sala, A., Toto, M., Pinello, L., Gabriele, A., Di Benedetto, V., Ingrassia, A.M.R., Lo Bosco, G., Di Gesù, V., Giancarlo, R., and Corona, D.F. V (2011). Genome-wide characterization of chromatin binding and nucleosome spacing activity of the nucleosome remodelling ATPase ISWI. EMBO J. *30*, 1766–1777.

Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246–3248.

Sarge, K.D., Murphy, S.P., and Morimoto, R.I. (1993). Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization , acquisition of DNA-Binding activity , and nuclear localization and can occur in the absence of stress. Mol. Cell. Biol. *13*, 1392–1407.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887–898.

Schwabish, M.A., and Struhl, K. (2007). The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo. Mol. Cell. Biol. *27*, 6987–6995.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature *442*, 772–778.

Shilatifard, A., Lane, W.S., Jackson, K.W., Conaway, R.C., and Conaway, J.W. (1996). An RNA polymerase II elongation factor encoded by the human ELL gene. Science *271*, 1873–1876.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844–847.

Shopland, L.S., Hirayoshi, K., Fernandes, M., and Lis, J.T. (1995). HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and

RNA polymerase II binding sites. Genes Dev. *9*, 2756–2769.

Simon, J.A., and Lis, J.T. (1987). A germline transformation analysis reveals flexibility in the organization of heat shock consensus elements. Nucleic Acids Res. *15*.

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

Sims, R.J., Millhouse, S., Chen, C.-F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. Mol. Cell *28*, 665–676.

Smith, C.L. (2008). A shifting paradigm: histone deacetylases and transcriptional activation. BioEssays News Rev. Mol. Cell. Dev. Biol. *30*, 15–24.

Smith, E.R., Winter, B., Eissenberg, J.C., and Shilatifard, A. (2008). Regulation of the transcriptional activity of poised RNA polymerase II by the elongation factor ELL. Proc. Natl. Acad. Sci. U. S. A. *105*, 8575–8579.

Stokes, D.G., Tartof, K.D., and Perry, R.P. (1996). CHD1 is concentrated in interbands and puffed regions of Drosophila polytene chromosomes. Proc. Natl. Acad. Sci. U. S. A. *93*, 7137–7142.

Straub, T., Gilfillan, G.D., Maier, V.K., and Becker, P.B. (2005). The Drosophila MSL complex activates the transcription of target genes. Genes Dev. *19*, 2284–2288.

Teves, S.S., and Henikoff, S. (2011). Heat shock reduces stalled RNA polymerase II and nucleosome turnover genome-wide. Genes Dev. *25*, 2387–2397.

Tréand, C., du Chéné, I., Brès, V., Kiernan, R., Benarous, R., Benkirane, M., and Emiliani, S. (2006). Requirement for SWI/SNF chromatin-remodeling complex in Tat-mediated activation of the HIV-1 promoter. EMBO J. *25*, 1690–1699.

Tsai, C., Kao, H., Yao, T., Mckeown, M., Evans, R.M., Hughes, H., and Jolla, L. (1999). SMRTER , a Drosophila Nuclear Receptor Coregulator , Reveals that EcR-Mediated Repression Is Critical for Development. Mol. Cell *4*, 175–186.

Tsukiyama, T., and Wu, C. (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. Cell *83*, 1011–1020.

Tsukiyama, T., Becker, P.B., and Wu, C. (1994). ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. Nature *367*, 525–532.

Tulin, A., and Spradling, A. (2003). Chromatin loosening by poly(ADP)-ribose polymerase (PARP) at Drosophila puff loci. Science *299*, 560–562.

Venters, B.J., and Pugh, B.F. (2009). A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. Genome Res. *19*, 360–371.

Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W.M.P., Van Schaik, F.M.A., Varier, R.A., Baltissen, M.P.A., Stunnenberg, H.G., Mann, M., and Timmers, H.T.M. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell *131*, 58–69.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. EMBO J. *17*, 7395–7403.

Wang, X., Lee, C., Gilmour, D.S., and Gergen, J.P. (2007). Transcription elongation controls cell fate specification in the Drosophila embryo. Genes Dev. *21*, 1031–1036.

Wang, Y.V., Tang, H., David, S., and Gilmour, D.S. (2005). Identification in vivo of different rate-limiting steps associated with transcriptional activators in the presence and absence of a GAGA element identification in vivo of different rate-limiting steps associated with transcriptional activators in the presence and absence of GAGA element. Mol. Cell. Biol. *25*, 3543–3552.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell *138*, 1019–1031.

Westerheide, S.D., Anckar, J., Stevens, S.M., Sistonen, L., and Morimoto, R.I. (2009). Stress-inducible regulation of heat shock factor 1 by the deacetylase SIRT1. Science *323*, 1063–1066.

Wu, C.-H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes Dev. *17*, 1402–1414.

Wu, C.-H., Lee, C., Fan, R., Smith, M.J., Yamaguchi, Y., Handa, H., and Gilmour, D.S. (2005). Molecular characterization of Drosophila NELF. Nucleic Acids Res. *33*, 1269–1279.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86–90.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Mol. Cell *21*, 227–237.

Yamaguchi, Y., Wada, T., Watanabe, D., Takagi, T., Hasegawa, J., and Handa, H. (1999). Structure and function of the human transcription elongation factor DSIF. J. Biol. Chem. *274*, 8085–8092.

Yamaguchi, Y., Inukai, N., Narita, T., Wada, T., and Handa, H. (2002). Evidence that Negative Elongation Factor Represses Transcription Elongation through Binding to a DRB Sensitivity-Inducing Factor / RNA Polymerase II Complex and RNA Evidence that Negative Elongation Factor Represses Transcription Elongation through Bindin. Mol. Cell. Biol. *22*, 2918–2927.

Yang, X., and Seto, E. (2003). Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. Curr. Opin. Genet. Dev. *13*, 143–153.

Yao, J., Ardehali, M.B., Fecko, C.J., Webb, W.W., and Lis, J.T. (2007). Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. Mol. Cell *28*, 978–990.

You, S.-H., Lim, H.-W., Sun, Z., Broache, M., Won, K.-J., and Lazar, M. A (2013). Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo. Nat. Struct. Mol. Biol. *20*, 182–187.

Zamft, B., Bintu, L., Ishibashi, T., and Bustamante, C. (2012). Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. Proc. Natl. Acad. Sci. U. S. A. 1–6.

Zhang, Z., Wippo, C.J., Wal, M., Ward, E., Korber, P., and Pugh, B.F. (2011). A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. Science *332*, 977–980.

Zhu, C.C., Bornemann, D.J., Zhitomirsky, D., Miller, E.L., O'Connor, M.B., and Simon, J. a (2008). Drosophila histone deacetylase-3 controls imaginal disc size through suppression of apoptosis. PLoS Genet. *4*, e1000009.

Ziesché, E., Kettner-Buhrow, D., Weber, A., Wittwer, T., Jurida, L., Soelch, J., Müller, H., Newel, D., Kronich, P., Schneider, H., et al. (2012). The coactivator role of histone deacetylase 3 in IL-1-signaling involves deacetylation of p65 NF-κB. Nucleic Acids Res. 1–20.

Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. (2009). Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. Cell *138*, 1122–1136.

Zobeck, K.L., Buckley, M.S., Zipfel, W.R., and Lis, J.T. (2010). Recruitment timing and dynamics of transcription factors at the Hsp70 loci in living cells. Mol. Cell *40*, 965–975.

VITA

Bhavana Achary

EDUCATION:

- Ph.D., Biochemistry and Molecular Biology, The Pennsylvania State University
- M.Sc., Biochemistry, Mumbai University, December 2001
- B.Sc., Chemistry/Biochemistry, Mumbai University, December 1999

PUBLICATIONS AND PRESENTATIONS:

- RNAi screen in Drosophila larvae identifies histone deacetylase 3 as a positive regulator of the hsp70 heat shock gene expression during heat shock. Manuscript submitted to BBA-Gene Regulatory Mechanisms
- Relationship between nucleosome organization and RNA Polymerase II pausing. Manuscript under preparation
- Poster presentation at Summer Symposium on Chromatin and Epigenetic Regulation of Transcription, Pennsylvania State University, 2011