Twists and Turns of Transcription:
Dynamic interplay between RNA Polymerase II and the nucleosome

Sheila S. Teves

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Steven Henikoff (Chair)

Christine Queitsch

Patrick Paddison

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

ABSTRACT

Twists and Turns of Transcription: Dynamic interplay between RNA Polymerase II and the nucleosome

Sheila S. Teves

Chair of Supervisory Committee:
Steven Henikoff
Affiliate Professor, Department of Genome Sciences

Transcription regulation underlies basic processes essential to life, including differentiation and development, cell-to-cell communication, and response to environmental stimuli. How the cell achieves a precise gene expression system to maintain cellular identity while allowing for plasticity remains an important biological question. As the interface between DNA and DNA-binding factors, chromatin exerts substantial influence on transcriptional regulation through its fundamental unit, the nucleosome. The following thesis addresses the questions of how nucleosomes influence transcriptional regulation, how RNA Polymerase II (Pol II) affects nucleosome stability and dynamics, and how Pol II overcomes the nucleosomal barrier. Using the heat shock response as a model for transcriptional regulation, we found that nucleosomes of activated genes increased in turnover, while those of repressed genes exhibited decreased turnover, suggesting that the act of transcription causes nucleosome turnover. This causality challenges the role of histone modifications in regulating gene expression, as modifications must be re-established after each turnover event. Furthermore, we discovered that the transcription-driven nucleosome turnover is partly mediated by the torsional stress on the DNA generated during Pol II translocation. Nucleosomes were destabilized as Pol II generates positive torsion ahead, and stabilized by the negative torsion behind, providing a mechanism for efficient Pol II progression while maintaining chromatin structure and organization.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

CHAPTER 1

INTRODUCTION

The heat shock response: a case study of chromatin dynamics in gene regulation

Modified from a review published in Biochemistry and Cell Biology

Transcriptional regulation is central to development, environmental response, and disease progression, and occurs at each major stage of the transcription process. In the initial stage, gene specific transcription factors recruit the RNA Polymerase II (Pol II) and general transcription factors to the promoter of the gene, forming the pre-initiation complex (PIC), in response to signal cascades that link transcription with intra- and extra-cellular cues (Baumann et al., 2010). Promoter DNA is then melted, and Pol II becomes dependent on factors that prevent backtracking and arrest (Fish and Kane, 2002). Promoter clearance occurs when Pol II transitions into productive initiation, but it can also pause 30 – 50 base pair (bp) downstream of the transcription start site (TSS) (Cernilogar et al., 2011). For a large percentage of the genes, this pause in elongation serves as the rate-limiting step in gene expression and provides an added layer of regulation (Levine, 2011). Once pausing is relieved, Pol II then enters productive elongation where it encounters an ordered chromatin template. After the whole gene is transcribed, Pol II dissociates from the template, and can then be recycled to begin the process anew (Shandilya and Roberts, 2012). At each stage, a multitude of factors associate with and regulate Pol II, primarily through the C-terminal domain (CTD) of its largest subunit. The CTD consists of tandem hepta-peptide repeats of Y-S-P-T-S-P-S. Each individual residue can be

modified, but phosphorylation of the serine residues is most critical to Pol II function. The hypo-phosphorylated Pol II is recruited to promoters to form the PIC, but the phosphorylation of the fifth serine residue (Ser5) transitions Pol II into productive initiation (Kim et al., 2010). Pausing after initiation is relieved when the second serine (Ser2) becomes phosphorylated (Kim et al., 2010). Aside from the factors that regulate Pol II itself, the CTD also acts as a docking region for many chromatin-related factors that modulate the Pol II template. Many of the mechanisms regulating Pol II, and its interaction with chromatin, have been discovered using model systems for transcriptional regulation such as the heat shock response in *Drosophila*.

The heat shock response has long been a gold standard for studying gene regulation. An evolutionarily conserved defense mechanism, the heat shock response involves a rapid and global transcriptional response to protect the cell against many types of stressors, including heat, cold, oxidative stress, heavy metal and alcohol poisoning (Akerfelt et al., 2010). Under normal conditions, the master heat shock transcription factor HSF exists as an inactive monomer. At the onset of stress, HSF trimerizes and binds to promoters of heat shock protein (*Hsp*) genes. HSF binding signals the release into productive elongation of Pol II that is paused downstream of the TSS, resulting in synchronous activation of stress-inducible genes from 10 – 1000 fold induction within minutes (Lindquist, 1986; Lis, 1998) (Figure 1-1). Concurrent with the activation of *Hsp* genes is the down-regulation of global transcription and a temporary halt in normal translation to prevent the accumulation of misfolded products (Lindquist, 1981; McKenzie et al., 1975; Tissieres et al., 1974). The simultaneous presence of gene induction and repression occurring in a fast system makes the heat shock response ideal for probing the dynamic processes in chromatin that accompany transcriptional regulation.

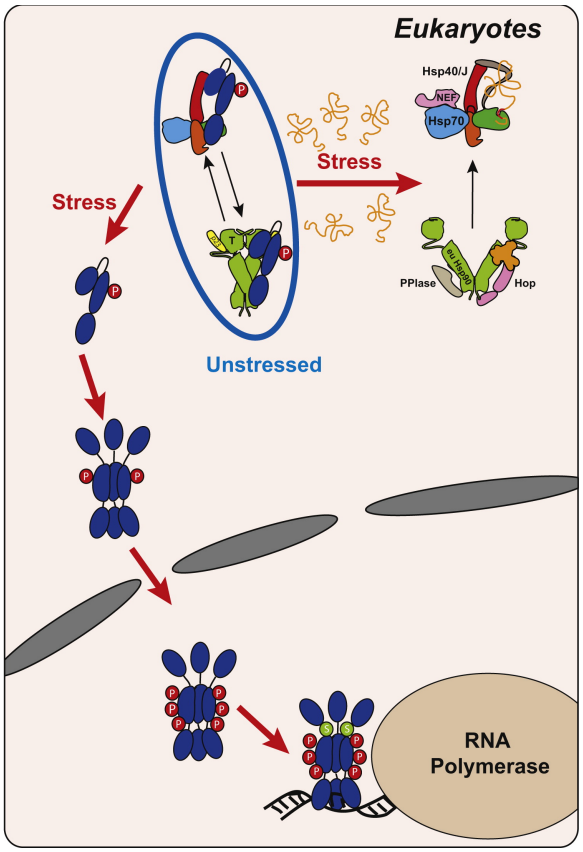

Figure 1-1. Regulation of the Heat Shock Response. In eukaryotes, the heat shock specific transcription factor Hsf1 is maintained in an inactive monomeric form in complexes with the chaperones Hsp70 and Hsp90. The titration of chaperones by massive protein unfolding upon proteotoxic stress will result in the release of Hsf1 from chaperone inactivation. Monomeric Hsf1 then trimerizes, is transported into the nucleus and activates heat shock gene transcription. Transcription of the rest of the genome is halted. Adapted and reprinted with permission from (de Nadal et al., 2011) under license number 3167780584543.

The primary unit of chromatin is the nucleosome, 147 bp of DNA wrapped around an octameric histone complex, which is organized across the genome in a conserved fashion (Luger et al., 1997; Mavrich et al., 2008) (Figure 1-2). Since the majority of eukaryotic DNA exists as nucleosomes, factors must necessarily counteract this packaging to allow Pol II access to the DNA at each stage of the transcription process. In fact, eukaryotes have evolved many redundant mechanisms to allow DNA accessibility during transcription that converge as major hubs of regulation. These include histone post-translational modifications (PTMs), incorporation of histone variants, remodeling by ATP-dependent remodelers, and nucleosome eviction and replacement (Henikoff, 2008). For example, promoters have evolved to restrict nucleosome occupancy through DNA sequences that are anti-nucleosomal (Iyer and Struhl, 2012), and through preferential incorporation of variant histones that form inherently less stable

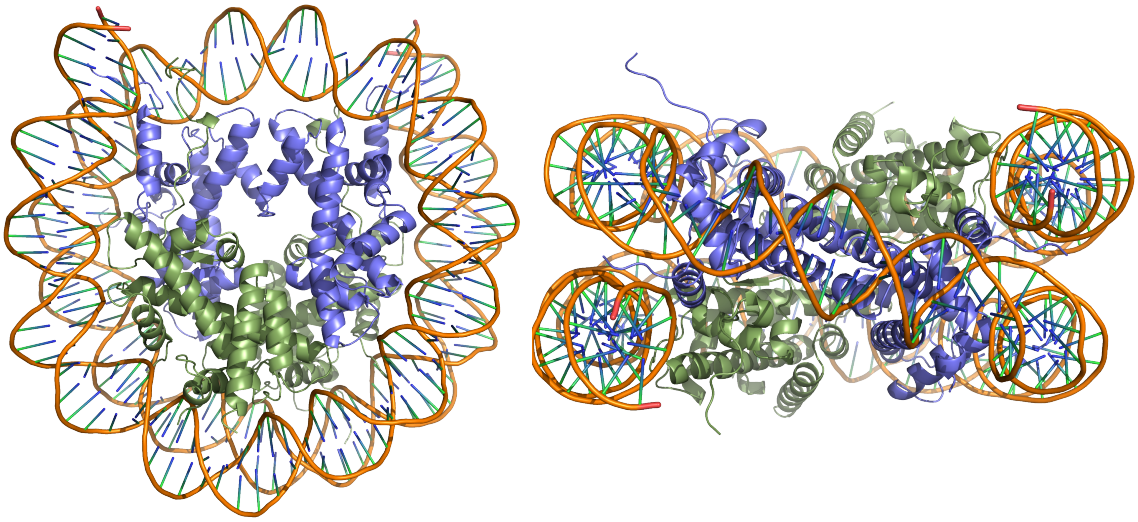

Figure 1-2. Nucleosome structure. The nucleosome is composed of two dimers of H3-H4 (purple) and two dimers of H2A-H2B (green), which combine in an octameric form to wrap 147 bp of DNA (orange). PDB # 1aoi.

nucleosomes (Jin and Felsenfeld, 2007). These properties then allow specific transcription factors to access the promoter region in order to recruit Pol II and associated general transcription factors to form the PIC. During active elongation, many factors that modify chromatin associate with the traveling Pol II, including histone acetyltransferases (HATs) and histone methyltransferases (HMTs) that modify N-terminal tails of genic histones during transcription (Jenuwein and Allis, 2001; Rando, 2012). These PTMs are predicted to alter the conformation of the nucleosome, such as an 'opening' of the nucleosome upon acetylation (Czarnota et al., 1997). Histone chaperones and remodelers also associate with Pol II during elongation to promote histone variant deposition that results in partial to full unwrapping of the nucleosome (Henikoff, 2008). Lastly, Pol II itself generates torsional stress on the DNA that can affect nucleosome dynamics, hinting at a potential mechanism for Pol II transit through the nucleosomal template (Henikoff, 2008).

Here, we examine how the heat shock response can be used as a model system for studying chromatin dynamics during gene regulation. We review the dynamic interplay between

Pol II and the nucleosome during each stage of the transcription process, and highlight the regulatory mechanisms that govern gene activation and repression. We also examine the decades-old question of how Pol II transcribes through a nucleosome. We discuss how the physical properties of Pol II, such as elongation rate, density on a given gene, and torsional stress generation can affect the fate of transcribed nucleosomes. Finally, we provide a description of the following doctoral research study that considers how nucleosomes can regulate Pol II transit, and consequently, gene expression.

Gene activation and the heat shock response

Promoter chromatin dynamics

One of the earliest events in gene activation is the promoter binding of gene specific transcription factors, such as HSF, which serve as transcriptional effectors of intra- and extra-cellular signaling cascades. These factors must access specific sequences in the context of nucleosomal DNA. HSF rapidly binds to promoters of *Hsp* genes upon heat shock, but under normal conditions, the chromatin at *Hsp* gene promoters already exists in an accessible state. Early mapping of chromatin structure using DNAse I, which under limiting conditions preferentially digests highly accessible DNA, revealed that the *Hsp* promoters are hypersensitive to digestion (Costlow and Lis, 1984; Wu, 1980). Furthermore, the hypersensitive regions coincide with the binding sites of sequence-specific factors such as HSF and GAGA factor (Costlow and Lis, 1984; Tsukiyama et al., 1994). DNase I hypersensitivity is independent of HSF binding sites Heat Shock Elements (HSEs), as its deletion or mutation does not significantly alter chromatin architecture on the *Hsp26* gene (Lu et al., 1993). In contrast, changes in the underlying sequence of the GAGA element, TSS, and the pause site for Pol II reciprocally

influence HSF binding at its cognate sites in the *Drosophila Hsp70* gene (Lu et al., 1993; Shopland et al., 1995). The relationship between promoter chromatin dynamics and transcription factor binding has been interrogated genome-wide using HSF ChIP-seq (Guertin and Lis, 2010; Guertin et al., 2012). The presence of DNase I hypersensitive sites, along with hyperacetylation of histones, predict which binding sites will be bound by HSF or not, suggesting that promoter nucleosome dynamics participate in target selection and activation (Guertin and Lis, 2010; Guertin et al., 2012). A comparison of other transcription factor binding and genome-wide DNase I hypersensitive sites suggests that the use of promoter chromatin context to differentiate between binding sites is general (Rhee and Pugh, 2011).

Complementary to DNase I, micrococcal nuclease (MNase) digestion patterns can provide another perspective of promoter chromatin dynamics. MNase is an endonuclease that introduces nicks on exposed double stranded DNA (Desai and Shankar, 2003) and further nibbles ends until it encounters a block in the form of bound protein, such as the nucleosome (Henikoff et al., 2011; Kent et al., 2011). MNase mapping of nucleosomes coupled with genome-wide mapping technologies has revealed a canonical nucleosome organization that is conserved across eukaryotes (Hughes et al., 2012; Mavrich et al., 2008) (Figure 1-3). A prominent aspect of this organization is a nucleosome-depleted region (NDR) near the TSS of most genes, followed by well-positioned nucleosomal arrays within the gene bodies. The NDR results from a highly dynamic nucleosome structure at the TSS that is enriched for the H3.3 and H2A.Z histone variants (Jin and Felsenfeld, 2007). Nucleosome depletion at the TSS is important for reliable gene expression (Bai et al., 2010). The enzymatic action of MNase, however, suggests that any protein will confer protection when bound to the DNA. When coupled with paired-end

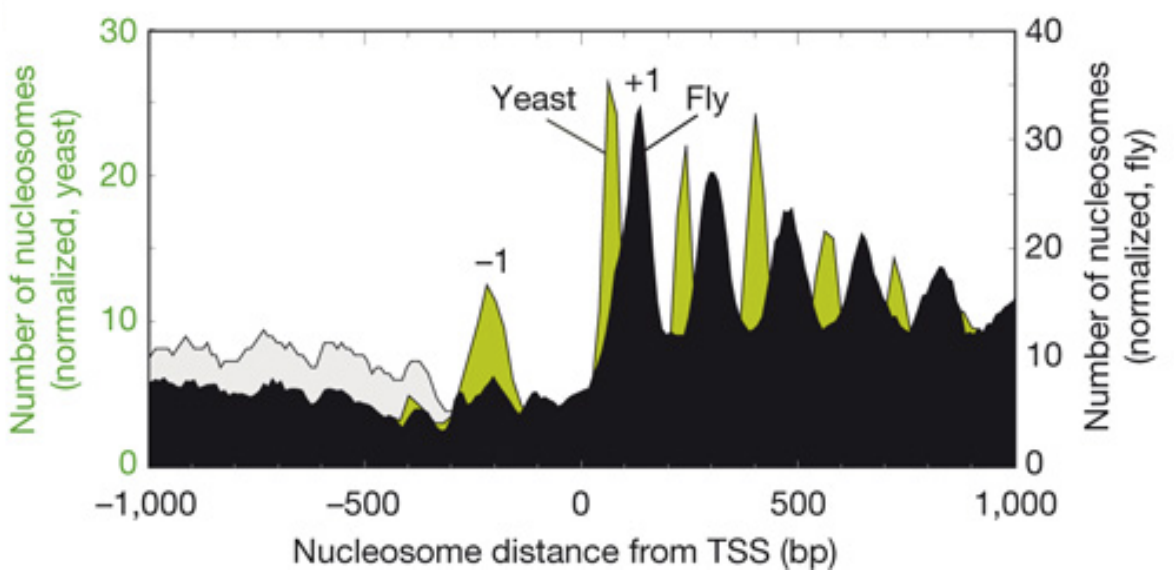

Figure 1-3. Canonical nucleosome organization relative to the TSS. MNase-digested fly chromatin was sequenced. Nearby genes were either included (grey) or eliminated (black) from the analysis, and normalized accordingly. The equivalent Saccharomyces profile is shown in green. Adapted and reprinted with permission from (Mavrich et al., 2008) under license number 3167761032579.

sequencing that can reveal the size of the MNase-digested fragment, MNase can also be used to map any DNA-binding protein, allowing for the visualization at high resolution of both the nucleosomal (~150 bp) and subnucleosomal (< 90 bp) components of chromatin in yeast, where the latter consists primarily of transcription factors and chromatin remodelers (Henikoff et al., 2011). The heat shock response in *Drosophila* cells may prove useful in revealing the dynamics of both nucleosomal and subnucleosomal chromatin components in response to transcriptional perturbation and highlight the dynamic nature of promoter chromatin during gene expression.

Paused RNA Polymerase and chromatin

The *Hsp* genes were among the first genes identified where the rate limiting step for expression occurs at Pol II elongation (Gilmour and Lis, 1986). For decades, the paradigm for gene expression posited that regulation occurs to modulate the recruitment/initiation of Pol II.

That is, once Pol II is recruited to the gene, it fires uniformly to produce the transcript. The *Hsp* genes, however, were known to contain Pol II on the gene body even under non-induced conditions. Furthermore, the Pol II is located about 30 bp downstream of the transcription start site, and was shown to be transcriptionally competent, containing ~30 bp of nascent RNA chain (Rougvie and Lis, 1988). This 'paused' Pol II is stable, persisting for long periods of time, and associating with factors that promote pausing, such as the DRB-sensitivity inducing factor (DSIF), the negative elongation factor (NELF) (Missra and Gilmour, 2010; Wu et al., 2003), and components of the RNA-interference (RNAi) pathway (Cernilogar et al., 2011). Upon activated HSF binding, the kinase positive elongation factor P-TEFb is recruited to the paused Pol II and phosphorylates DSIF, NELF, and Pol II itself at Ser2 of the CTD to transform the paused complex into an actively elongating one (Peterlin and Price, 2006). NELF dissociates from Pol II, and the phosphorylated DSIF subsequently acts to promote elongation as it travels down the gene with Pol II (Yamada et al., 2006) (Figure 1-4). In recent years, promoter-proximal Pol II pausing has emerged as a global mechanism of gene regulation. Genome-wide studies in *Drosophila*, mouse, and human cells have estimated that roughly 30% of the genome is regulated at the level of elongation (Levine, 2011). This list is highly enriched for developmental genes. In one estimate, at least half of the *Drosophila* developmental control genes contain paused Pol II (Levine, 2011). One proposal suggests that regulation at the elongation step has evolved to allow for rapid and synchronous activation of a set of genes to allow for precisely timed development within a population of cells (Levine, 2011).

Several studies have shown that the chromatin architecture of paused genes is markedly different from non-paused genes (Gilchrist et al., 2010; Gilchrist et al., 2008; Weber et al., 2010). The canonical nucleosome organization is evident in non-paused genes, containing an array of

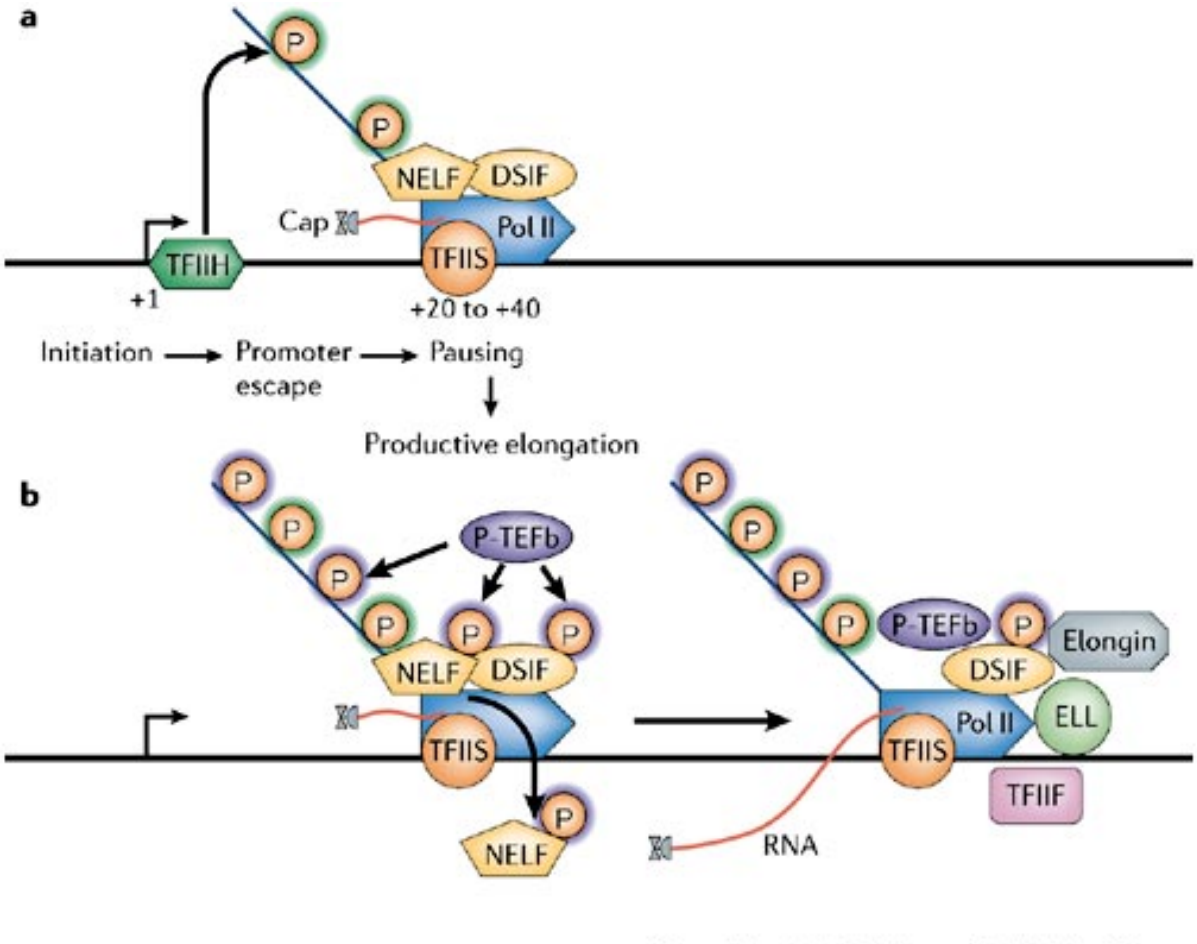

Figure 1-4. Regulation of Pol II pausing. a. TFIIH-mediated phosphorylation of Ser5 of the carboxy-terminal domain (CTD) Pol II occurs on pre-initiation complex formation or before promoter-proximal pausing. DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) probably facilitate Pol II pausing in the promoter-proximal region, and TFIIS also associates with the paused polymerase. TFIIS stimulates the intrinsic RNA-cleavage activity of Pol II to create a new RNA 3'-OH in the Pol II active site after backtracking of the polymerase. Capping enzyme associates with the Ser5-phosphorylated CTD and with Spt5, and the nascent RNA becomes capped during this first stage of elongation. **b**. Positive transcription-elongation factor-b (P-TEFb)-mediated phosphorylation of DSIF, NELF and Ser2 of the Pol II CTD stimulates productive elongation, and the capping enzyme might contribute to this process by counteracting the negative effects of DSIF and NELF. TFIIS facilitates efficient release of Pol II from the pause site by aiding the escape of backtracked transcription complexes. NELF dissociates from the transcription complex and DSIF, TFIIS and P-TEFb track with Pol II along the gene. TFIIF, eleven-nineteen lysine-rich in leukemia (ELL), and elongin, which stimulate Pol II elongation activity, might also associate with the elongation complex. Adapted and reprinted with permission from (Saunders et al., 2006) under license number 3167761215298.

well-positioned nucleosomes within gene bodies. Interestingly, the nucleosomes within paused genes are less well positioned and are lower in occupancy (Gilchrist et al., 2010; Gilchrist et al., 2008; Weber et al., 2010). Also, one study showed that when Pol II pausing is inhibited, paused genes gain a nucleosome at the TSS, suggesting that Pol II elongation and nucleosomes cooperate to maintain regulation of these genes (Gilchrist et al., 2010). However, the regulation occurring at the TSS of paused genes does not explain the disrupted nucleosome organization within their gene bodies. As stated above, one proposed role for regulating Pol II at the elongation step is to provide a mechanism for a fast yet uniform rate of expression. By preventing bursts of multiple Pol IIs at a given time, the organism can then generate synchronous expression of key genes across multiple cells (Levine, 2011). In vitro, such a uniform rate of Pol II progression allows for the survival of the nucleosome, while bursts of multiple Pol IIs result in nucleosome eviction (Bintu et al., 2011; Jin et al., 2010; Kulaeva et al., 2010). Understanding the mechanisms that govern nucleosome organization and dynamics of paused genes may yield insights into the regulation of Pol II pausing.

Nucleosome dynamics and Pol II elongation

Once Pol II enters productive elongation, it faces an array of ordered nucleosomes. Interestingly, gene bodies and exons have higher nucleosome occupancy than introns and intergenic regions (Chen et al., 2010). Yet Pol II moves along chromatin in vivo at a rate comparable to its rate of movement along naked DNA templates, whereas a single nucleosome in vitro presents a formidable barrier to transcriptional elongation (Knezetic and Luse, 1986; Luse and Studitsky, 2011). The question of how Pol II moves through a nucleosome, and the resulting fate of the nucleosome after Pol II has passed through, has been the subject of debate for decades

(Petesch and Lis, 2012). To begin to answer this question, we must first understand nucleosome dynamics.

Most new nucleosomes are formed immediately behind the replication fork and consist of canonical histones. Histones can also be replaced outside of replication and result in the incorporation of special histone variants. Two universal histone variants are deposited during transcription: H2A.Z and H3.3. To replace the H2A/H2B dimer, the nucleosome has to partially unwrap (Schwartz and Ahmad, 2005). However, replacement of the H3 in the central $(H3/H4)_2$ tetramer with the H3.3 variant results in unwrapping of the nucleosome beyond the dyad axis and eviction (Schwartz and Ahmad, 2005). Therefore, H3.3 incorporation marks nucleosome turnover events. Profiling of H3.3 in the *Drosophila* genome has revealed that the highest level of H3.3 deposition occurs within bodies of highly transcribed genes (Mito et al., 2005). The correlation between H3.3 variant deposition and transcription suggests that elongation results in disruption of transcribed nucleosomes. Because there are only 4 amino acids that differ between H3 and H3.3, these experiments were performed using exogenous transgenes that were induced for several days, reflecting steady state levels of deposition. To further define nucleosome dynamics, a new method was developed adapting metabolic labeling of newly synthesized proteins for chromatin profiling (Deal et al., 2010) (Figure 1-5). The methionine analog Azidohomoalanine (Aha) is incorporated into newly synthesized proteins in the absence of methionine. The azide moiety of Aha can then react with an alkyne-adapted biotin linker through a copper catalyzed cycloaddition reaction, providing a biotin tag on all newly synthesized proteins. Isolation of nuclei followed by extraction of MNase-digested chromatin provides the input material for streptavidin pulldown to enrich for nucleosomes containing a newly synthesized histone, thus marking a recent turnover event. This technique is called Covalent

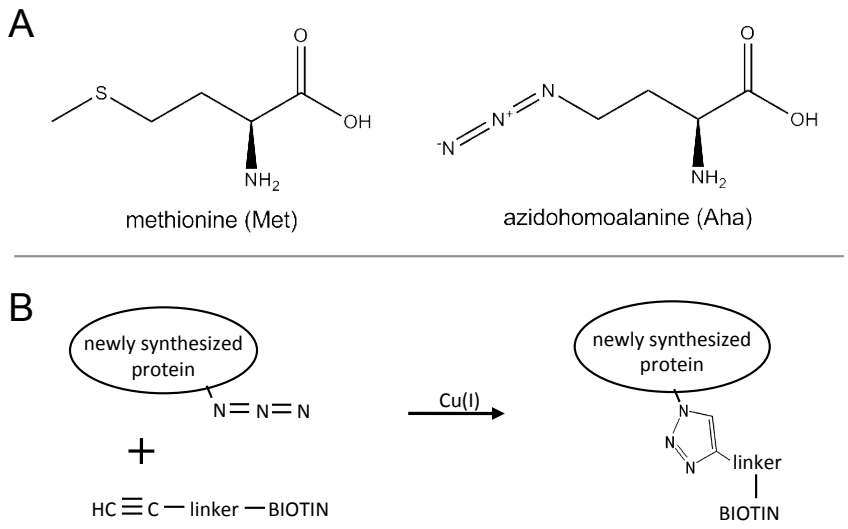

Figure 1-5. CATCH-IT. A) The structure of methionine (left) and the analog azidohomoalanine (right) shows the similarities between the two amino acids. B) In absence of met, Aha is incorporated into newly synthesized proteins. In the presence of Cu(I), the azide group in the Aha reacts with the alkyne group linked to biotin in a cycloaddition reaction to covalently attach biotin to newly synthesized proteins.

Attachment of Tags to Capture Histones and Identify Turnover (CATCH-IT). CATCH-IT signals correspond well with H3.3 incorporation, confirming that H3.3 marks sites of nucleosome turnover within bodies of transcribed genes. Brief Aha pulses allowed for nucleosome turnover times to be estimated. As expected from the H3.3 profiling experiments, the highest expressed genes experience the highest rate of nucleosome turnover, ~20 turnover events per nucleosome during each cell cycle (Deal et al., 2010). These experiments imply that the rate of Pol II elongation is directly related to the rate of nucleosome turnover in *Drosophila* cells, which brings up an important question. Are nucleosomes actively removed ahead of the transcribing Pol II to allow progression? The heat shock response may provide insight into this process as well.

As early as forty years ago, scientists discovered that certain regions of the *Drosophila* polytene chromosomes undergo massive decondensation in response to heat, producing the heat shock puffs (Simon et al., 1985). These puffs are accompanied by changes in DNase I sensitivity and MNase digestion patterns, such that the gene body becomes more sensitive to nuclease

digestion after induction, and the ordered array of nucleosomes becomes disordered (Levy and Noll, 1981; Wu et al., 1979). Furthermore, multiple ATP-dependent chromatin remodelers coordinate in regulating *Hsp* genes in yeast (Erkina et al., 2010; Shivaswamy and Iyer, 2008) as individual nucleosomes become remodeled throughout the genome in response to heat (Shivaswamy et al., 2008). At the *Hsp70* gene, Petesch and Lis discovered that nucleosomes are lost within seconds of heat shock, even before the first transcribing Pol II has passed through (Petesch and Lis, 2008). Using CATCH-IT, we can now study the direct relationship between nucleosome dynamics and transcription.

The fate of transcribed nucleosomes

The question of how Pol II transcribes through nucleosomes, and their resulting fate, has remained unresolved for the last three decades. Nucleosomes have been classically described as barriers to DNA-based processes. Early *in vitro* studies have shown that a nucleosome over the promoter sequence effectively blocks recruitment of Pol II (Lorch et al., 1987; Shaw et al., 1978; Studitsky et al., 1995), and that a well-positioned nucleosome in the path of an engaged Pol II blocks further elongation (Hodges et al., 2009). Using biochemical approaches, many of the factors required in overcoming this barrier have been identified. These factors include an array of histone chaperones, nucleosome remodelers, and general transcription factors (Petesch and Lis, 2012). Furthermore, structural studies and single-molecule assays have provided high resolution models for Pol II traversal through a nucleosome. Single molecule fluorescence resonance energy transfer (FRET) studies have shown substantial fluctuations between DNA and histones near the entry/exit site (Li et al., 2005) that may allow for Pol II access and lead to the loss of H2A/H2B dimers after a single round of Pol II traversal (Bohm et al., 2011; Gansen et al., 2009;

Kireeva et al., 2002; Kuryan et al., 2012). Using optical traps and molecular tweezers, others have shown that the nucleosome can survive Pol II passage through formation of a transient loop between the DNA ahead of Pol II and the histones behind (Bintu et al., 2011; Hodges et al., 2009; Kulaeva et al., 2009). Based on these and other biochemical studies, several in vitro models have been proposed for Pol II traversal through a nucleosome that ultimately center on the resulting fate of the transcribed nucleosome: complete survival of the octamer, partial survival where either one or both H2A/H2B dimers are lost while the H3/H4 tetramer is retained (Kireeva et al., 2002; Kuryan et al., 2012), and lastly, full dissociation of the octamer (Bintu et al., 2011; Jin et al., 2010). How these various models correspond to *in vivo* studies is still unclear. The highly organized structure of nucleosomes surrounding the transcription start site (Mavrich et al., 2008) is suggestive of a mechanism that allows for retaining nucleosomes in a well positioned manner. However, high transcription rates correlate with decreased nucleosome occupancy within the coding regions (Rando, 2012), suggesting that Pol II elongation can lead to eviction of nucleosomes. Also, as discussed above, nucleosomes within the *Hsp70* gene body are lost within seconds of heat shock, prior to the first transcribing Pol II (Petesch and Lis, 2008). How this eviction occurs remains unresolved. Understanding the basic mechanisms for Pol II traversal through nucleosomes may offer insights into gene regulation. Some studies hint that DNA itself may play a more active role during Pol II elongation through chromatin.

Twin supercoiled domain model and chromatin

The double helix nature of the DNA lends itself to torsional stress. Indeed, as Pol II denatures the promoter and elongates, it generates a wave of positive supercoils ahead and negative supercoils behind (Baranello et al., 2009; Liu and Wang, 1987), called the twin-

supercoiled domain model (Figure 1-6). One estimate using mathematical modeling based on experimentally determined physical properties of Pol II and chromatin during transcription suggests that a wave of positive supercoils generated by transcription of 5 bp propagates through chromatin at a rate of 2 orders of magnitude faster than Pol II elongation (Becavin et al., 2010). In yeast, accumulation of unresolved positive supercoils inhibit transcription globally, affecting over 80% of yeast genes (Gartenberg and Wang, 1992; Joshi et al., 2010). Topoisomerases relieve the resulting torsional tension caused by processive enzymes such as Pol II, thereby allowing for continuous and successive rounds of transcription. Interestingly, inhibition of Topoisomerase II and release of torsional tension from DNA nicking lead to decreased DNase I hypersensitivity in the promoters of active beta-globin genes (Villeponteau et al., 1984; Villeponteau and Martinson, 1987), suggesting an intimate connection between DNA

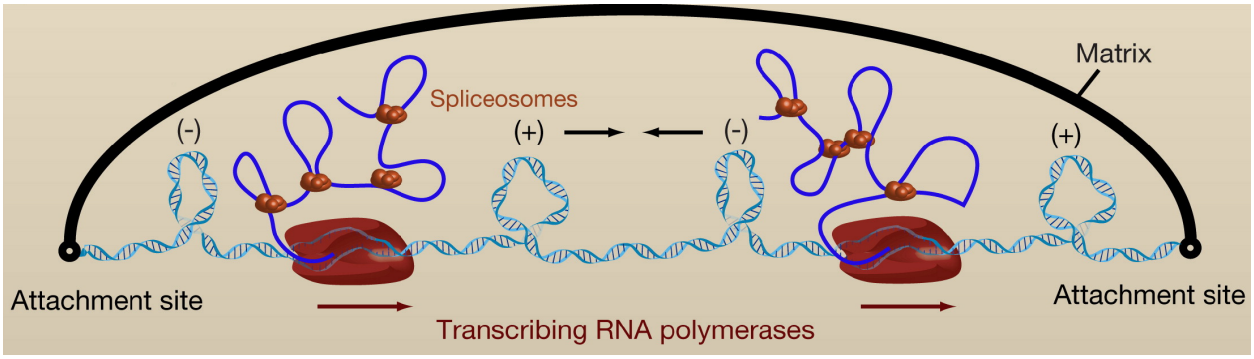

Figure 1-6. DNA Topology and Its Relevance in Transcription. When Pol II is prevented from rotating along the helical axis of the DNA during transcription, positive and negative supercoils accumulate ahead and behind the enzyme, respectively. Multiple factors impede the rotation of RNAP by increasing its hydrodynamic drag. These factors include the nascent RNA strand (blue solid line) and its processing factors, such as spliceosomes. When tandem genes are transcribed, RNAP complexes progress in the same direction on duplex DNA. The DNA domain between them contains both negative and positive supercoils that could diffuse toward each other and subsequently annihilate. The rate at which this process occurs depends on the timescales at which DNA can spin around its own helical axis and the axis defined by plectonemes. Adapted and reprinted with permission from (Koster et al., 2010) under license number 3167770816001.

superhelicity and chromatin structure. Indeed, positive supercoiling of DNA templates has been shown to restrict nucleosome assembly (Gupta et al., 2009), while negative supercoiling promotes assembly (Hizume et al., 2004). When the gene was shown to lose nucleosomes ahead of the initial transcribing Pol II upon heat shock (Petesch and Lis, 2008), Zlatanova and Victor proposed that the Pol II-generated wave of positive supercoils ahead of Pol II destabilizes downstream nucleosomes (Gupta et al., 2009). Paradoxically, active elongation on the *Hsp70* gene does not seem necessary as DRB inhibition of Poll II still resulted in the loss of nucleosomes in the *Hsp70* gene body at heat shocked cells (Petesch and Lis, 2008). However, DRB specifically inhibits the kinase activity of P-TEFb, which acts at the transition of initiation and elongation about 30 bp downstream of the TSS by catalyzing the phosphorylation of Ser-2 on the CTD (Marshall et al., 1996; Marshall and Price, 1995). If the estimate that transcription of 5 bp is sufficient to generate a wave of positive supercoils (Becavin et al., 2010) is accurate, it may actually predict that nucleosomes would still be lost at the *Hsp70* gene even under DRB-inhibition of elongation. Perhaps the pausing of Pol II after transcription of ~30 bp maintains these genes under torsional stress to destabilize genic nucleosomes. In this way, once the signal for elongation is received, Pol II transits through the gene most efficiently. Probing the relationship between Pol II mechanics, DNA structure, and nucleosome organization will further our understanding not only of paused Pol II regulation, but also of the very basic mechanisms of the transcription process itself.

Gene repression and chromatin

Activation of the *Hsp* genes during heat shock occurs concurrent with a rapid and global reduction in transcription, although how this process occurs is largely unknown. In *Drosophila*

salivary glands, Pol II is lost from transcriptionally active, developmentally regulated puff sites upon heat shock (Jamrich et al., 1977), and the loss of Pol II is dependent on an active RNAi machinery (Cernilogar et al., 2011). Interestingly, AGO2-associated small anti-sense RNAs are increased genome-wide upon heat shock, suggesting a role for siRNAs in the *Drosophila* heat shock response (Cernilogar et al., 2011). Similarly, in mammalian systems, global repression is mediated by the heat shock induced expression of noncoding RNAs that function to disrupt contacts between Pol II and promoter DNA (Yakovchuk et al., 2009), implying that loss of Pol II has co-evolved with the heat shock response for efficient global repression. While Pol II is lost from promoters, in yeast, most of the factors in the pre-initiation complex remain bound to promoters after heat shock (Zanton and Pugh, 2006). Such maintenance of the underlying transcriptional machinery may provide the mechanism for efficient recovery after the stress is removed. However, how nucleosome dynamics is affected during gene repression is not fully understood.

The global transcriptional repression that occurs under heat shock is perhaps unique in mechanism because it is necessarily fast, global, and reversible. Upon removal of heat, cells more gradually return to their normal physiological state, including their normal transcription levels. It is possible that the mechanism for reversible gene repression used under heat shock may be separate from a more permanent form of silencing and is more adapted for certain genes that require plastic regulation. For complete gene silencing, cells have evolved redundant mechanisms, many of which take advantage of the inhibitory properties of heterochromatin (Wutz, 2011) such as alterations of histone modifications and recruitment of heterochromatin related proteins like Hp1 (Zeng et al., 2010), leading to a visible condensation of the region, as in the dramatic case of X-inactivation (Haaf et al., 1993; Lee and Bartolomei, 2013).

It is likely that dynamic systems involved in gene expression, such as Pol II and chromatin, serve as feedback mechanisms for regulation. During activation, Pol II-mediated disruption of chromatin, both at promoters and gene bodies, allows for increased DNA accessibility for future transcription events, while the decreased chromatin dynamics inhibit factor access to DNA and thus promoting repression. In this way, minor changes in chromatin dynamics can have important consequences in regulating gene expression levels.

Transcription twists and turns

The following doctoral research study aims to understand the relationship between RNA Polymerase II and chromatin during transcription. In particular, this work attempts to answer the following broad questions: 1) How does differential gene regulation affect chromatin? Conversely, what is the role of chromatin during gene regulation? 2) How does Pol II transcribe through a nucleosomal template *in vivo*? What factors participate in this process? Understanding the basic mechanisms of gene regulation and nucleosome dynamics can lead to greater insights into the epigenetic processes that drive differentiation and development, that allow for environmental adaptation, and that become disregulated during diseases such as cancer.

Chapter 2 describes the various methods used in this doctoral research, with several methods expanded into greater detail in Appendices I and II. In Chapter 3, we illustrate how the heat shock response can be used to study transcription-related chromatin dynamics both for gene activation and repression. We provide evidence for Pol II-dependent nucleosome dynamics during elongation, and for a global chromatin-mediated transcriptional response to stress. We follow up on the mechanisms of transcription-dependent nucleosome dynamics in Chapter 4 by examining the effect of Pol II-generated torsional stress on nucleosome turnover. We present

support for a dynamic role of DNA supercoils generated during transcription in destabilizing nucleosomes ahead of the transcribing Pol II, allowing for efficient elongation. The implications of this doctoral work are put into the perspective of the broader transcription and epigenetics field in Chapter 5, and we describe future avenues for research that follows from this work in Chapter 6.

This work on chromatin and transcription has revealed a dynamic interplay between RNA Polymerase II and the nucleosome that is partly mediated by the DNA itself.

CHAPTER 2

MATERIALS AND METHODS

Cell culture, treatments, and low-salt extraction

Drosophila S2 cells grown to log phase in HYQ-SFX Insect medium (Invitrogen) supplemented with 18 mM L-Glu and harvested as previously described (Henikoff et al., 2009). For heat shocked samples, cells were incubated at 37°C for 15 minutes and immediately harvested. DRB treatment was performed as described previously using 125 μM final concentration in the medium for 10 minutes immediately before the heat shock experiment began (Petesch and Lis, 2008). The following cell treatments were used: control (normal conditions), heat shock for 15 minutes, DRB treatment for 10 minutes, and DRB for 10 minutes followed by 15 minutes of heat shock. Camptothecin (Sigma Aldrich C9911) and ICRF-193 (Sigma Aldrich I4659) were resuspended to 10 mM in DMSO and frozen in aliquots. Final concentration of 10 μM of either drug was added to cell medium for 15 minutes, and cells were harvested immediately. Trimethyl-psoralen (TMP) (Sigma Aldrich T6137) was resuspended in 0.5 mg/mL in ethanol. Nuclear isolation, MNase digestion and low-salt extraction in 80 mM buffer (70 mM NaCl, 10 mM Tris 7.4, 2 mM $MgCl_2$, 2 mM EGTA, 0.1% Triton X-100, 0.5 mM PMSF pH7.5) were performed as described previously (Henikoff et al., 2009) and elaborated in Appendix I.

RNA Polymerase II native ChIP

A 10% aliquot of the 80 mM salt fraction was saved and the remainder was immunoprecipitated overnight with Pol II antibody 8WG16 (Abcam, ab817) at 4°C with constant

agitation. Samples were next mixed with 50 μL of 80 mM buffer-prewashed Protein G Dynabeads (Invitrogen, 100.04D) for 2 hours at 4°C. The bound Dynabeads were washed twice with 80 mM buffer without Triton X-100 and resuspended in 200 μL of the same buffer. For DNA extraction, 1/50th volumes of 5 M NaCl and 0.5 M EDTA were added to both the resuspended beads and saved input material, and total RNA was digested with 0.5 μg of RNase (Roche, 10928100) for 10 minutes at 37°C. SDS was added to a final concentration of 0.5% and protein was digested with 20 μg of Proteinase K (Invitrogen 25530-049) at 70°C for 10 minutes. Samples were extracted twice with Phenol/Chloroform followed by ethanol precipitation of DNA using 20 μg of Glycogen (Roche, 14267332) as carrier. The precipitated DNA was resuspended in 20 μL of 0.1X TE pH8 and quantified using Picogreen quantification kit for modified Solexa paired-end library preparation.

CATCH-IT

CATCH-IT was performed essentially as described (Deal et al., 2010) with the following modifications. Immediately following Azidohomoalanine (Aha) addition, samples were either incubated for 15 minutes at room temperature (control) or heat shocked at 37°C. For experiments with DRB treatment, DRB was added to the cells at 125 μM final concentration during the last 10 minutes of methionine depletion, with the same final concentration maintained during the 15 minute incubation with Aha. For cells treated with either Topoisomerase I or II, drugs were added at final concentration of 10 μM immediately after Azidohomoalanine (Aha) addition and incubated for 15 minutes. Nuclei extraction, biotin coupling, MNase digestion, and streptavidin immunoprecipitation were performed as described previously. The input and streptavidin ChIP

DNA were prepared for paired end Solexa sequencing using the modified library preparation. This procedure is elaborated in Appendix II.

Tri-Methyl Psoralen photobinding

Our protocol was adapted from a previously published assay (Roca). S2 cells were diluted to 1 x 10^7 cells/mL in growth medium, and 2 mL aliquots in small plates were used. TMP was added to cells with and without drug treatments at a final concentration of 2 μg/mL for 10 minutes in the dark. Plates of cells were exposed to 3 kJ m^{-2} of 365 nm light (Fotodyne UV Transilluminator 3-3000 with 15 W bulbs). Cells were then collected and washed with cold 1x PBS and resuspended thoroughly in 0.5 mL of 1x PBS with 0.5% SDS. Proteins were digested using 4 μg of Proteinase K (Invitrogen) for 1 hour at 55°C. DNA was extracted by phenol/chloroform and precipitated with ethanol. After resuspension in 200 μL of H_2O, RNA was digested using 1μg of RNase (Invitrogen) at 37°C for 30 minutes. DNA was extracted and precipitated as before and resuspended to 0.5 μg/μL. To achieve fragment sizes of 100 – 500 bp, 50 μg of DNA was sonicated in a Bioruptor 3x 15 minutes each at high setting with 30 seconds on/off in cold water. Cross-linked fragments were enriched by repeated rounds of denaturation and Exonuclease I (Exo I) digestion. Using 3 μg of sonicated DNA diluted to 250 μL, samples were boiled in water bath for 10 minutes and incubated in ice-water for 2 minutes. To each sample, 30 μL of 10X Exo I buffer and 10 μL of Exo I were added, and digestion was allowed to proceed for 1 hour at 37°C. Samples were boiled and cooled as before, and 10 μL of Exo I was added for a second round of 1 hour digestion. DNA was extracted and precipitated, and concentration was assayed by PicoGreen quantification (Invitrogen).

TMP-seq

We devised a method for producing Solexa libraries to map the precise location of the irreversible TMP interstrand cross-link. Exo I digested DNA samples were subjected to enzymatic reactions for end polishing and 'A'-tailing as previously described (Henikoff et al., 2011). After ligation of PE barcode adapters (Henikoff et al., 2011), the 5' strand was digested using 25 U of λ exonuclease (NEB) for 30 minutes at 37°C. The DNA was purified using Ampure beads, and eluted in 35 μL of H_2O. The resulting 3' strand was used as template for 10 rounds of primer extension in 1x HiFi Phusion buffer, 0.8 mM dNTP, 2 U Phusion DNA Polymerase (NEB), and 40 nM of P7 extension primer (CAAGCAGAAGACGGCATACGA*G - *denotes phosphorothioate linkage) in the following cycling conditions: 95°C for 3 minutes, 10 rounds of linear amplification (95°C for 1 minute, 57°C for 1 minute, 72°C for 3 minutes), 95°C for 1 minute, hold at 8°C. The resulting single-stranded products were purified with Ampure beads, eluted at 35 μL of H_2O, and concentrated down to 17 μL using a vacuum centrifuge. The purified products were then appended with ribo-G in 1x Terminal deoxynucleotidyl Transferase (TdT) buffer, 20 U TdT (NEB), and 4 mM of rGTP in 37°C for 15 minutes. The products were purified using Ampure beads, eluted in 35 μL of H_2O, and concentrated to 18 μL volume as before. The single-stranded ribo-tailed products were ligated to a double stranded adapter with CCC-overhang (oligo 1: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCC, oligo 2: [Phosphate]GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG) as described (Henikoff et al., 2011). Prior to standard PCR for PE library (Henikoff et al., 2011), we added an extension step to create double stranded templates (60°C for 3 minutes, 98°C for 30 seconds, 18 cycles of amplification (98°C for 10 seconds, 65°C for 30 seconds, and 60°C for 30 seconds), 60°C for 5 minutes and

hold at 8°C). The final PCR products were purified with Ampure beads, eluted with 40 μL of H_2O, and quantified by PicoGreen assay.

Biochemical Fractionation for nascent RNA isolation

Biochemical fractionation of chromatin was performed as previously described (Wysocka et al., 2001) with the following changes as described in Weber et al (in preparation). Briefly, ~1 × 10^8 cells were harvested and washed with PBS. Cells were resuspended in 800μL Buffer A (10 mM HEPES [pH 7.9], 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 1 mM dithiothreitol, and protease inhibitor cocktail [Roche], 200 U Superase-in [Ambion]) supplemented with Triton X-100 (0.1% final) before use. After 8 minutes of incubation on ice, nuclei were pelleted (3 min, 3000 × g, 4°C), and lysed in 800μL of Buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM dithiothreitol, and protease inhibitor cocktail [Roche], 200 U Superase-in [Ambion]) supplemented with Triton X-100 (0.1% final) before use. The samples were incubated on ice for 15 min and the pellet was collected by centrifugation as before. Buffer B washing was repeated twice. Pellets were collected and resuspended in Buffer B+ (20 mM EDTA, 0.2 mM EGTA, 2 mM spermine, 5 mM spermidine, 1 mM DTT, protease inhibitor cocktail [Roche], 200 U Superase-in [Ambion]), supplemented with Triton X-100 (0.1% final) before use. Pellets were incubated on ice for 10 minutes and collected by centrifugation as before. Buffer B+ wash was repeated 3 times and the final pellet was resuspended in 720 μL of Buffer B+ with 80 μL 10% SDS. Chromatin associated RNA was isolated by Trizol and quantified spectroscopically. Ribosomal RNA was depleted using Ribo-Zero Magnetic kit (Epicentre).

Nascent RNA-seq

Nascent RNA (200 ng) after rRNA depletion was fragmented to 50 – 300 nucleotides using the RNA Fragmentation Reagents (Ambion). After propanol precipitation, RNA was resuspended in 5 μL and used as input for reverse transcription using the random primers of the SuperScript III First Strand Synthesis System (Life Technologies). Following reverse transcription, sample volumes were increased to 100μL and cleared using Illustra Microspin S-200 HR columns (GE Healthcare). To create the second strand, 50 μL of master mix (0.4× First Strand buffer, 1 mM DTT, 0.8 mM dNTP [with U replacing T], 3× 2nd Strand Buffer, 40 U *E.coli* DNA Pol I [Invitrogen], 10 U *E.coli* DNA Ligase [Invitrogen], 2 U *E.coli* RNAse H [Invitrogen]) was added for a final volume of 150 μL per sample. Reaction proceeded for 2 hours at 16°C, and the resulting double stranded DNA was purified on Qiagen column, eluted in 40 μL, and subjected to modified Paired-End sequencing library protocol (Henikoff et al., 2011). Prior to final PCR amplification, the second strand was digested using 1U UNG AMPerase (Invitrogen) in 40 μL TE 0.1× pH 7.5 for 15 minutes at 37°C. Libraries were amplified 12 to 18 cycles.

Sequencing and data analysis

Solexa sequencing libraries were constructed as described (Henikoff et al., 2011). Cluster generation and 25 rounds of paired-end sequencing was performed by the FHCRC Genomics Shared Resource using the Illumina Hi-Seq 2000. Base calling, data processing and analysis were performed as described (Henikoff et al., 2011). Gene tracks from the sequencing data were visualized using Signalmap (Nimblegen, Inc.). Ends analysis, heat maps and k-means clustering were performed as described (Henikoff et al., 2009). Midpoint-versus-length analysis was

performed as described (Henikoff et al., 2011). Gene Ontology analysis was performed on each k-means cluster using the GeneCodis 2.0 program (Carmona-Saez et al., 2007; Nogales-Cadenas et al., 2009) with a hypergeometric test and false rate discovery rate calculation to correct the p values for multiple testing. For data used in Chapter 3, GEO accession number is GSE30755. For Chapter 4 data, GEO accession number is GSE47795.

For TMP-seq, read 1 was trimmed of the CCC overhangs. The first sequenced nucleotide was mapped onto the genome and fitted with a kernel density function estimation with bandwidth of 20 bp (Gehring M 2009). Ends analyses for TMP-seq data were performed as follows: For each 10-bp interval in a 4 kb region centered at either the TSS or TES, the average signal in that interval was normalized to the average signal in the whole 4 kb region. A running average of a 50 bp-window was used to smoothen the data. The supercoiling difference (SD) was calculated as follows: For each gene, we assigned a supercoiling value (SV) by averaging the TMP-seq values for 1 kb upstream of the TSS and subtracting the average TMP-seq value for 1 kb downstream. $SD_{TopoI} = SV_{TopoI} - SV_{Ctl}$ and $SD_{TopoII} = SV_{TopoII} - SV_{Ctl}$.

CHAPTER 3

HEAT SHOCK REDUCES STALLED RNA POLYMERASE II AND NUCLEOSOME TURNOVER GENOME-WIDE

Modified from an article published in Genes and Development

Summary

Heat shock rapidly induces expression of a subset of genes while globally repressing transcription, making it an attractive system to study alterations in the chromatin landscape that accompany changes in gene regulation. We have characterized these changes in *Drosophila* cells by profiling classical low-salt soluble chromatin, RNA Polymerase II (Pol II), and nucleosome turnover dynamics at single base-pair resolution. With heat shock, low-salt soluble chromatin and stalled Pol II levels were found to decrease within gene bodies, but no overall changes were detected at transcriptional start sites. Strikingly, nucleosome turnover decreased genome-wide within gene bodies upon heat shock in a pattern similar to that observed with inhibition of Pol II elongation, especially at genes involved in the heat shock response. Relatively high levels of nucleosome turnover were also observed throughout the bodies of genes with paused Pol II. These observations suggest that down-regulation of transcription during heat shock involves reduced nucleosome mobility and that this process has evolved to promote heat-shock gene regulation. Our ability to precisely map both nucleosomal and subnucleosomal particles directly from low-salt soluble chromatin extracts to assay changes in the chromatin landscape provides a simple general strategy for epigenome characterization.

Introduction

The heat shock response is a universally conserved reaction to environmental stress that involves rapid transcriptional changes. Inactive monomers of the master heat shock transcription factor HSF trimerize upon heat shock and translocate to the nucleus where they bind to the promoters of heat shock protein (*Hsp*) genes (Akerfelt et al., 2010). HSF binding triggers the release of paused polymerases already engaged at promoters of *Hsp* genes, resulting in fast and synchronous activation of HSF targets (Lee et al., 2008; O'Brien and Lis, 1991). Gene induction varies according to conditions, but this rapid transcriptional response generates 10- to 1000-fold induction of *Hsp* genes within minutes of temperature elevation (Lindquist, 1986). Simultaneously, heat shock results in down-regulation of normal transcription (Jamrich et al., 1977), presumably to prevent accumulation of misfolded translation products (Lindquist, 1986). These rapid genome-wide transcriptional responses make the heat shock system ideal for investigating chromatin alterations that accompany gene regulatory changes.

The basic repeating unit of chromatin is the nucleosome, which consists of an octameric histone protein core that wraps 147 bp of DNA (Luger et al., 1997). Packaging of DNA into nucleosomes can occlude DNA sequences and prevent transcription factor binding, in which case disruption or mobilization of nucleosomes is necessary for gene activation and other processes that require access to regulatory sequences. For example, nucleosomes in the bodies of induced *Hsp70* genes are rapidly lost within seconds of heat shock, and this loss depends on HSF, Poly(ADP-)Ribose Polymerase 1 (PARP1), and GAGA factor (GAF) (Petesch and Lis, 2008). Furthermore, loss of nucleosomes is required for full activation of *Hsp* genes (Petesch and Lis, 2008). Heat shock studies in yeast have revealed a functional interplay between multiple nucleosome remodeling complexes in regulating *Hsp* genes (Erkina et al., 2010; Shivaswamy

and Iyer, 2008), further implicating nucleosome dynamics in gene induction. Despite the wealth of information on *Hsp* gene induction, much less is known about the mechanisms for genome-wide down-regulation of transcription during heat shock. Fluorescence analyses of GFP-tagged Pol II show that Pol II becomes released from DNA upon heat shock, suggesting that direct regulation of Pol II kinetics plays a role in this process (Hieda et al., 2005). In support of this possibility, SINE RNAs that are upregulated during heat shock (Allen et al., 2004; Mariner et al., 2008) have been shown to disrupt contacts between promoter DNA and Pol II (Yakovchuk et al., 2009). However, some evidence exists for a role of nucleosomes in this process as well. In one mammalian study, histone deacetylases HDAC1 and HDAC2 were shown to mediate global histone deacetylation during heat shock (Fritah et al., 2009), indicating that changes in chromatin 'states' correlate with global down-regulation of transcription. Furthermore, repositioning of individual nucleosomes was seen throughout the budding yeast genome in response to heat shock (Shivaswamy et al., 2008). Although these studies revealed effects of gene regulatory changes on chromatin, the mechanistic processes responsible for these changes remain unknown.

To gain insights into the mechanistic role of nucleosomes in gene regulation, we used the heat shock response as a model system to effect global chromatin changes that occur with transcriptional alterations. Specifically, we asked the following questions: 1) How does the chromatin landscape change with alterations in gene expression? 2) What is the interplay between nucleosome dynamics and changes in transcription? And 3) can we gain insight into the mechanisms that govern global gene repression during heat shock and the possible roles that chromatin plays in this process? To address these questions, we have introduced a simple strategy for epigenome characterization based on traditional micrococcal nuclease mapping, chromatin salt fractionation and metabolic labeling of histones to probe the entire low-salt

soluble chromatin landscape and its dynamics at single base-pair resolution. We show that the low-salt soluble chromatin fraction is enriched for distinct subnucleosome-sized chromatin particles relative to total MNase-digested chromatin. During heat shock, the landscapes of both nucleosomal and subnucleosomal chromatin components dramatically change at induced genes. In contrast, the genome-wide distribution of subnucleosomal particles at transcription start sites (TSSs) is maintained during heat shock despite global down-regulation of transcription. Furthermore, genome-wide reduction in stalled Pol II occurs concomitantly with a decrease in nucleosome solubility and turnover within bodies of active genes. Using a Pol II elongation inhibitor, we show that the changes in nucleosome turnover that occur during heat shock resemble the turnover changes caused by direct inhibition of transcription elongation. We conclude that heat shock causes reduced Pol II elongation and decreased nucleosome turnover genome-wide.

Results

Low-salt soluble chromatin is enriched for distinct types of subnucleosomal particles

Using *Drosophila* S2 cells, we digested intact nuclei with micrococcal nuclease (MNase) to obtain mostly mononucleosomes and extracted the low-salt (80 mM) soluble fraction, sometimes referred to as classical 'active' chromatin (Sanders, 1978). We then captured all MNase-protected fragments from both the total MNase digested nuclei and the 80 mM fraction by using paired-end Solexa sequencing without prior size selection, except for removal of unligated adapters with Ampure magnetic beads (Henikoff et al., 2011). The size distribution of paired-end reads from both total nuclei and the 80 mM fraction reveals that fragments as small as ~20 bp could be sequenced and mapped to the genome (Figure 3-1 A-B). Although there is a

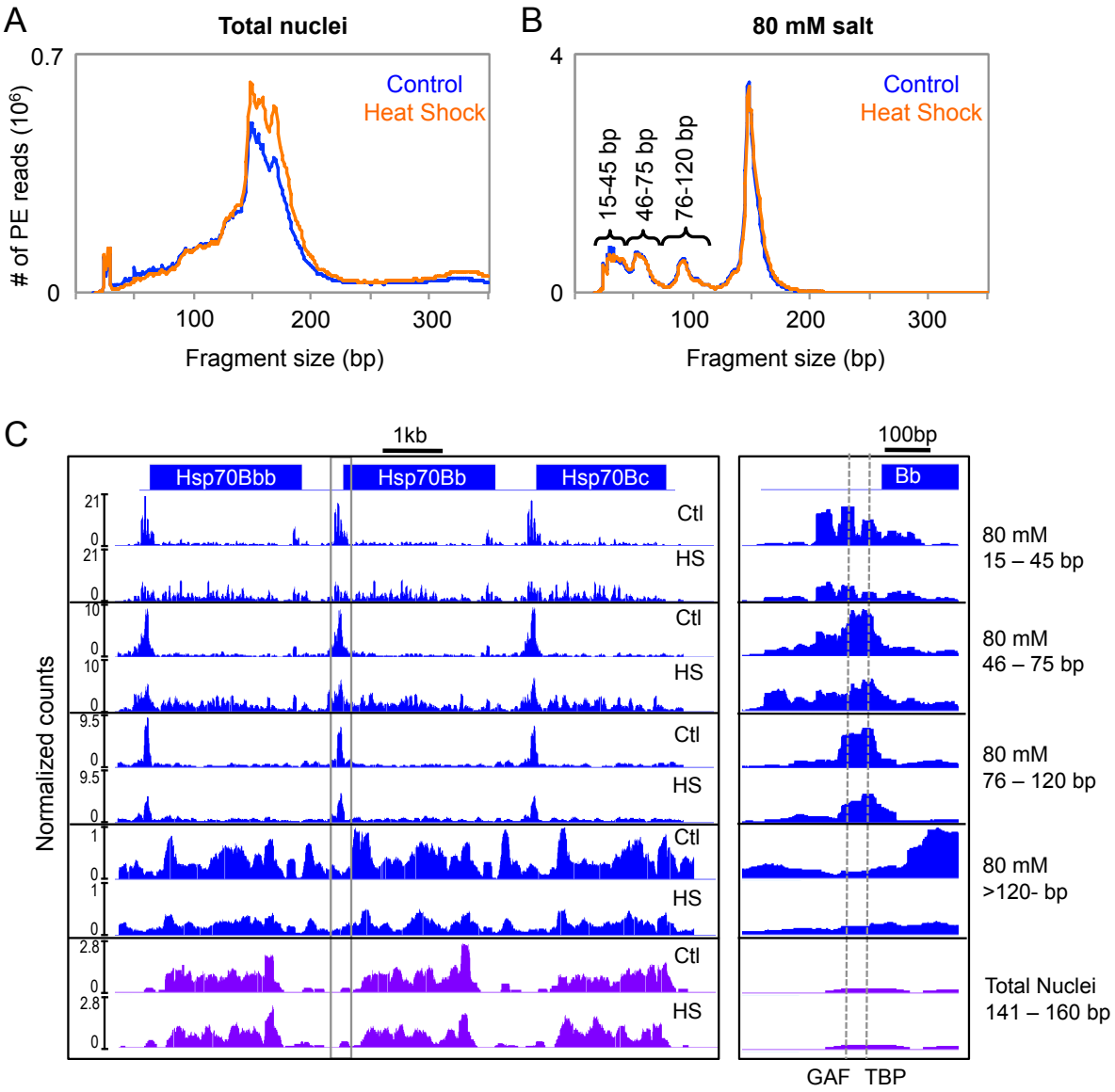

Figure 3-1: Low-salt soluble chromatin is enriched for distinct subnucleosomal particles. Paired-end read length distribution of MNase-digested total nuclei (A) and the 80mM salt extracted fraction (B) before and after heat shock (blue and orange, respectively). (C) Paired-end reads from the different size classes of the 80 mM salt fraction (blue) and the nucleosome-sized fragments of the total nuclei (purple) were mapped onto the *Drosophila melangaster* genome. The mapped reads were converted to normalized counts for each base pair at the Hsp70 gene cluster, and landscapes were displayed with NimbleGen SignalMap. A close-up view of the promoter region of Hsp70Bb that is boxed in (C) is shown on the right panel with the dotted lines representing the binding sites for GAGA Factor (GAF) and TATA-binding protein (TBP). Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

major peak at 147 bp in the total nuclei fraction representing octameric nucleosomes, peaks of larger fragment sizes are also prominent, likely due to histone H1-containing chromatosomes (Simpson, 1978). Interestingly, the size distribution of fragments from the 80 mM salt fraction not only shows a sharp peak for 147–bp nucleosomes, which suggests exclusion of chromatosomes from low-salt soluble chromatin, but also shows distinct peaks for subnucleosomal components not seen in the distribution of fragments from total nuclei. The peaks suggest distinct types of protection of DNA from MNase by particles that are enriched in 80 mM salt-soluble chromatin.

To investigate the distinct size classes in the 80 mM fraction, we parsed the mapped paired-end reads into four length categories based on the peaks in the read size distribution: 15 – 45, 46 – 75, 76 – 120, and > 120 bp. For the total nuclei fraction, we used mapped paired-end reads that are 141 – 200 bp in length to define nucleosomes. We then converted the mapped reads to density counts at single base pair resolution and focused on the three tandem copies of the *Hsp70* gene at the 87C heat shock locus (Figure 3-1 C). In both the total nuclei and the 80 mM fraction, mapped nucleosomes showed a rather complex distribution. With the exception of a few well-positioned nucleosomes, nucleosomes in gene bodies displayed heterogeneous positioning and occupancy profiles. Under normal growth conditions the subnucleosomal particle DNA fragments (15 – 45, 46 – 75, 76 – 120 bp) displayed distributions that are very different from those of nucleosomes. At this locus, the particles were found primarily at the 5' TSSs in regions that are generally depleted of nucleosomes. A closer view of the promoter region revealed several peaks in the 15 - 45 bp class, with two of these peaks corresponding to sites of binding by GAF and TATA-binding protein (TBP) (Figure 3-1C, right panel), indicating that we can map distinct small DNA-binding proteins. Subnucleosomal particles also localized to regions

with no known gene annotations, suggesting that these particles map to potential regulatory regions in the genome (Figure 3-2). Furthermore, the square shape of the peaks of these smaller fragments indicates complete protection and tight positioning of the DNA-binding proteins, in contrast to the rounded peaks of the more mobile nucleosomes. This provides evidence for DNA protection from MNase cleavage by small DNA binding proteins under native conditions.

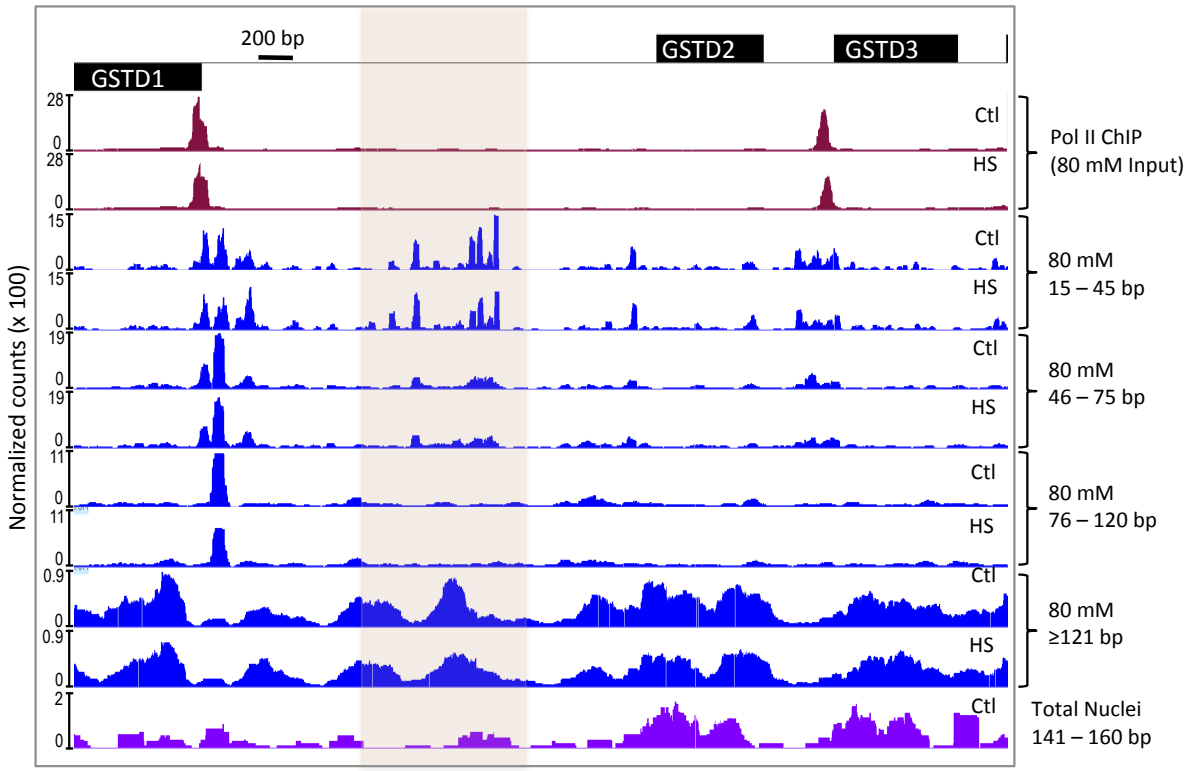

Figure 3-2: Distinct landscapes of stalled Pol II, the 80 mM chromatin salt fraction, and Total Nuclei. Paired-end reads from the Pol II ChIP, 80 mM salt fraction, and total nuclei were parsed into different size classes and mapped onto the genome. The mapped reads were converted to normalized counts for each base pair at the GSTD gene cluster, and landscapes were generated using Nimblegen Signalmap. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

Changes in the Hsp70 chromatin landscape during heat shock

Previous studies have shown that nucleosomes within the *Hsp70* genes are lost within seconds of heat shock (Petesch and Lis, 2008) but we wondered how other components of the chromatin landscape change during induction. Therefore, we harvested cells after 15 minutes at 37 °C and followed the chromatin extraction and Solexa library procedures described above. The size distributions of MNase-protected fragments were highly similar between the control and heat shocked samples for both the total nuclei and the 80 mM salt fraction (Figure 3-1 A-B). We converted the mapped reads for each size class into density counts and compared heat shock with control landscapes over the 87C *Hsp70* locus (Figure 3-1 C). Whereas total nuclei showed only slight changes in nucleosome occupancy with heat shock, the 80 mM salt fraction showed a marked decrease. Interestingly, the prominent peaks of shorter fragments at the TSS were diminished with heat shock induction, whereas occupancy of shorter fragments within gene bodies dramatically increased, which suggests that these shorter fragments derive from the transcriptional machinery itself as it progresses through the *Hsp70* gene.

Distinct genome-wide distributions of low-salt soluble chromatin particles

To examine the genome-wide distribution of each size class, we ordered all genes by expression level and averaged the normalized counts for each expression quintile in the 2–kb regions surrounding both the TSS and the transcription termination site (TTS). We also performed the same analysis on the nucleosome-sized fragments from total nuclei to examine global nucleosome occupancy. The genome-wide distribution of nucleosomes from total nuclei agrees with previous studies that show nucleosome depletion at the TSS, positioning of the +1 nucleosome, and a gradual decrease in positioning further downstream for expressed genes

(Figure 3-3) (Mavrich et al., 2008; Weber et al., 2010). In the 80 mM salt-soluble chromatin fraction, both 76-120 bp and >120 bp fragment size classes exhibit similar strongly phased profiles for the three quintiles of expressed genes, in contrast to the flat profiles that characterize the two inactive gene quintiles (Henikoff et al., 2009). Interestingly, the average profile of 76-120 bp fragments resembles the periodicity of phased nucleosomes (>120 bp) but with narrower peaks, consistent with a shorter fragment size, as if derived from partially unwrapped nucleosomes(Li et al., 2005; Weber et al., 2010). Furthermore, the genome-wide profile of the 76-120 bp class shows a higher peak at the -1 nucleosome position relative to the >120-bp class, which implies that nucleosomes around active promoters might have a higher tendency to be partially unwrapped than their neighbors.

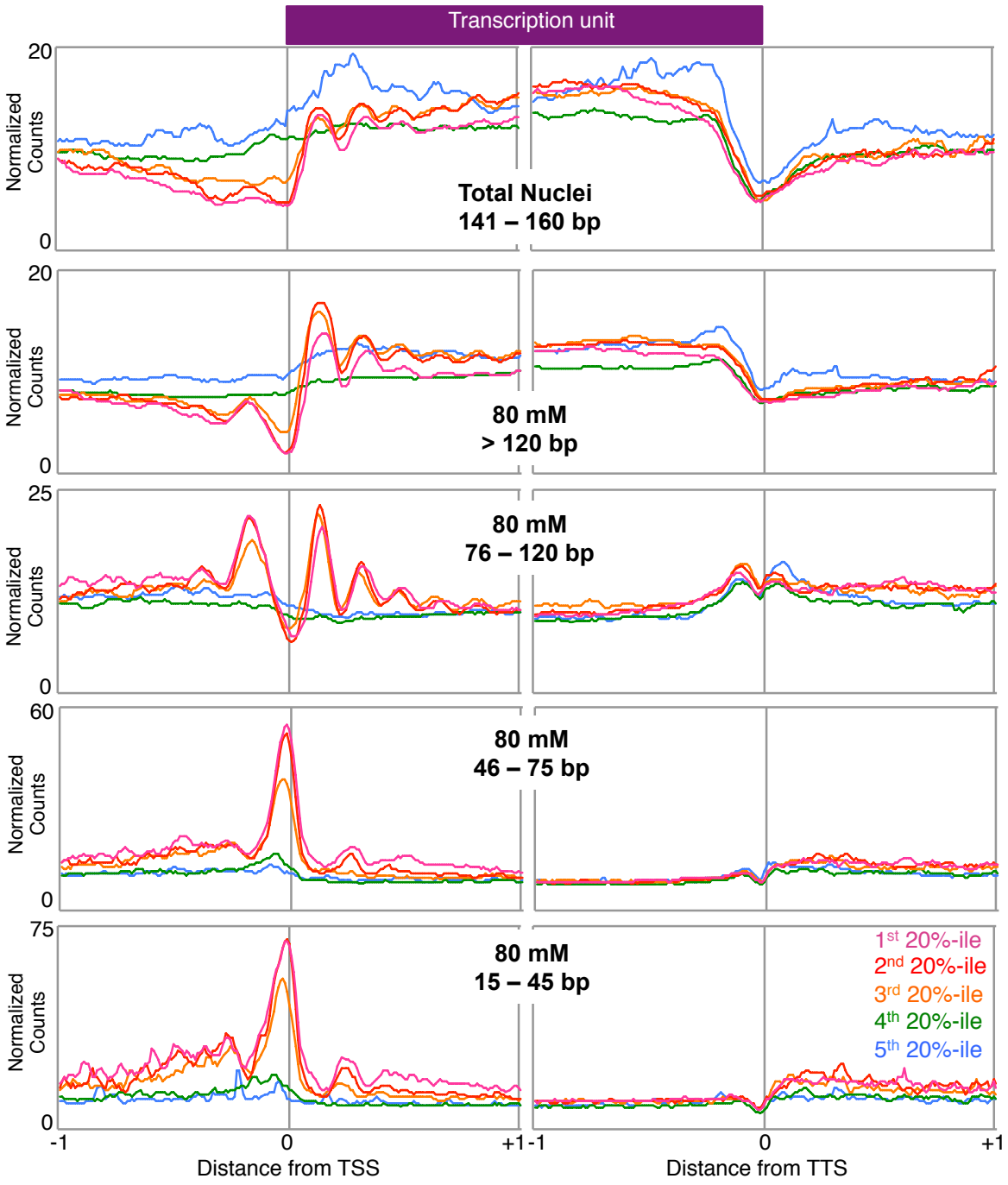

Figure 3-3: Distinct genome-wide distributions of active chromatin particles. Each mapped size class of the 80 mM salt fraction was scaled to normalized counts per 10 bp window. Within each size class, all genes were grouped into quintiles by expression level and the average normalized counts per 10 bp window was determined for each quintile in the 2-kb region immediately surrounding the TSS and TTS. In the Total Nuclei fraction, only the 141-160 bp fragments were used in order to represent the nucleosomal profile. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

We also observed that small fragments (15 – 45 and 46 – 75 bp) are enriched at TSSs genome-wide, similar to *Hsp70* genes. These subnucleosomal particles are found almost exclusively at expressed genes, which suggests that the proteins that protect TSSs from MNase in low-salt soluble chromatin are involved in transcription initiation. Using previously determined genomic sites for various DNA binding proteins, we found that these size classes are enriched for GAF, EZ2 PSC2, and Zeste sites relative to flanking regions (Figure 3-4),

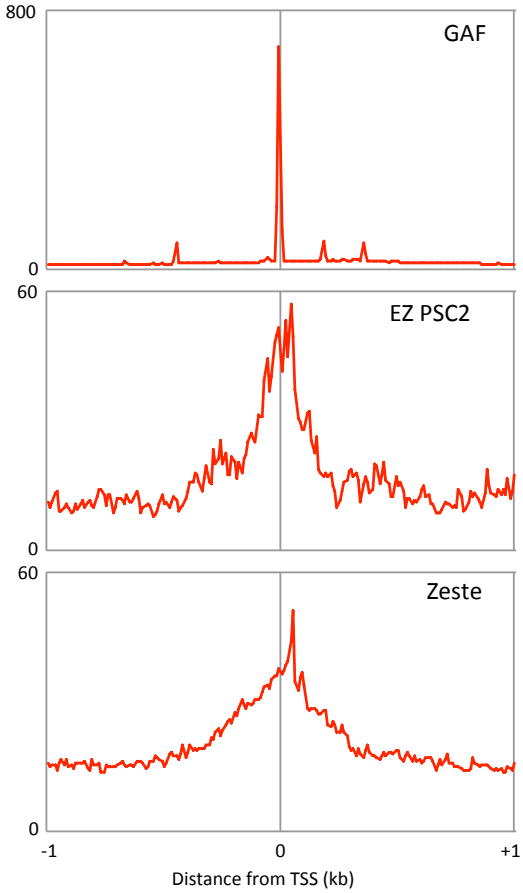

Figure 3-4: Subnucleosomal particles are enriched for binding sites of regulatory proteins. The mapped short fragments (15 – 45 bp) from the 80 mM salt fraction are enriched for GAGA factor (GAF), Polycomb group proteins (ZE PSC2), and Trithorax proteins (Zeste) binding sites. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

suggesting that these small fragments result from MNase protection of general transcription factors such as GAF and protein complexes that generally occupy EZ and Zeste sites.

We next asked how heat shock affects the genome-wide distribution of the subnucleosomal particles. To do this, we compared the average gene profiles of the 15-45 bp size class in heat shocked to control cells (Figure 3-5 A). The plots are superimposable, which indicates that subnucleosomal particles are undisturbed by heat shock. Dividing genes into quintiles based on normal expression level also revealed no differences in the subnucleosomal particle distribution before and after heat shock (Figure 3-6 A). This suggests that heat-shock mediated down-regulation of expression does not occur at the level of recruitment and initiation, in contrast to heat shock induction at *Hsp70* genes, which is accompanied by major changes in subnucleosomal particle occupancy at the TSS.

We also examined changes in the nucleosomal component of low-salt soluble chromatin. We aligned TSSs of normally expressed genes for the nucleosome-sized classes of the 80 mM and total nuclei fractions of heat shocked and control samples (Figure 3-5 B-C). We observed a genome-wide reduction in nucleosome solubility in active genes after heat shock, whereas silent genes showed little to no change (Figure 3-6 B). These observations imply that heat shock results in a decrease in nucleosome salt-solubility within active gene bodies.

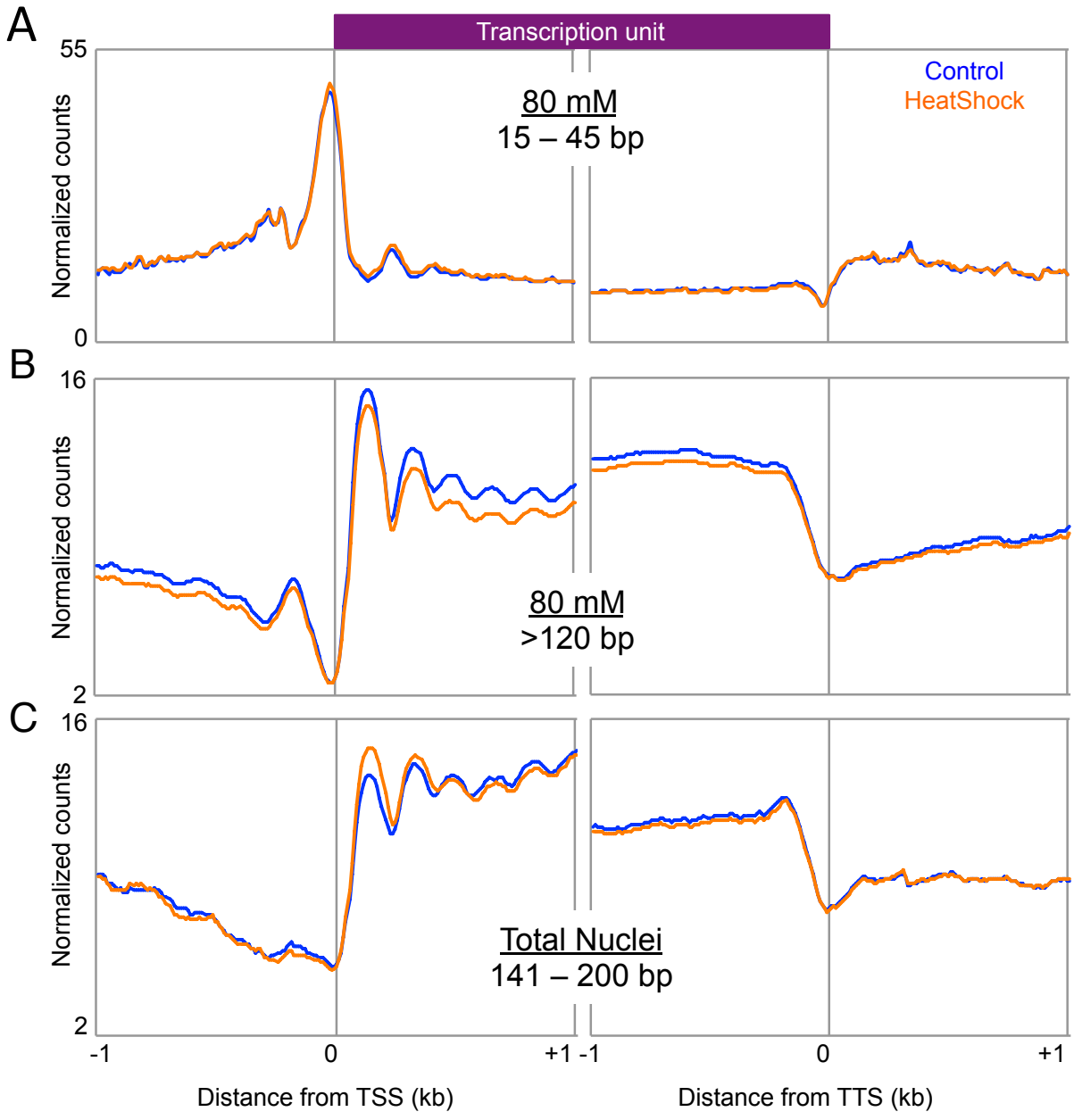

Figure 3-5: Genome-wide effects of heat shock on distinct chromatin components. Ends analysis of: all genes for the 15 – 45 bp size class of the 80 mM fraction (A), all expressed genes for the >120 bp size class of the 80 mM fraction (B), and all expressed genes for the 141 – 200 bp class of the Total Nuclei fraction (C). Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

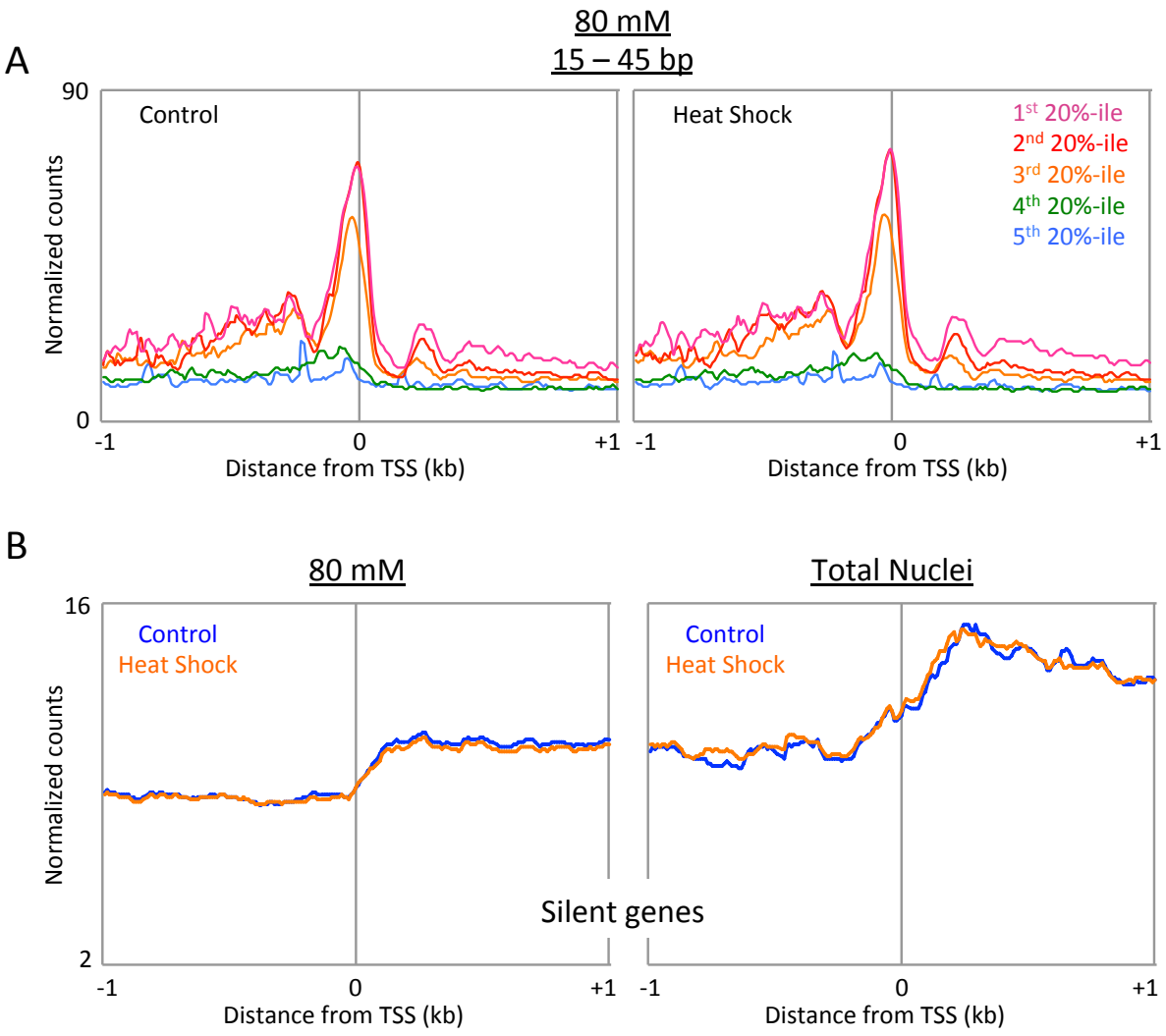

Figure 3-6: Effects of heat shock on distinct chromatin components. (A) Quintile analysis of all genes for the 15 – 45 bp fragments of the 80 mM fraction before (left) and after (right) heat shock. (B) Ends analysis was performed on all silent genes for the nucleosomal fragments of the 80 mM salt fraction and total nuclei over a 2 kb region surrounding the TSS before (blue) and after (orange) heat shock. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

Stalled RNA polymerase II decreases globally during heat shock

Transcriptional changes during heat shock can be detected cytologically in *Drosophila* polytene salivary gland chromosomes within minutes (Ashburner and Bonner, 1979; Belyaeva and Zhimulev, 1976; Bonner and Pardue, 1976), however, standard expression profiling is relatively insensitive to rapid transcriptional changes because newly synthesized mRNA is diluted by steady-state mRNA. A 15 minute heat shock represents only $1/80^{th}$ of the 20 hour S2 cell cycle, which likely accounts for our inability to detect significant expression changes after heat shock by microarray-based profiling, except for heat shock genes (data not shown). A more direct genome-wide approach to observing transcriptional levels is to map Pol II bound to DNA. We had previously shown that low-salt soluble chromatin includes ~50-bp particles that are enriched 30 bp downstream of the TSS in genes that had been found to be regulated at the level of Pol II elongation (paused genes), in contrast to genes matched for expression but lacking paused Pol II (non-paused genes) (Weber et al., 2010). To confirm that these particles represent paused Pol II itself, we performed native chromatin immunoprecipitation (ChIP) with an antibody against the Pol II C-terminal domain (CTD), using 80 mM salt-extracted chromatin as input (low-salt soluble Pol II). DNA was isolated and subjected to paired-end sequencing. When cells were grown under normal conditions, low-salt soluble Pol II was enriched immediately downstream of TSSs of *Hsp70* genes, but nowhere else in the gene bodies (Figure 3-7 A), consistent with previous studies identifying *Hsp* genes as containing paused Pol II (O'Brien and Lis, 1991). We then averaged the normalized counts for all paused and non-paused genes (Muse et al., 2007) in a region of 2 kb surrounding the TSS and found that Pol II was indeed strongly enriched at promoters of paused genes (Figure 3-7 B). We also used 5,6-dichlorobenzimidazole (DRB) to induce Pol II stalling followed by ChIP-seq and found a 40% increase in low-salt

soluble Pol II at the 87C *Hsp70* locus (Figure 3-7 A), further confirming that the soluble chromatin fraction is enriched for paused Pol II at *Hsp70* promoters.

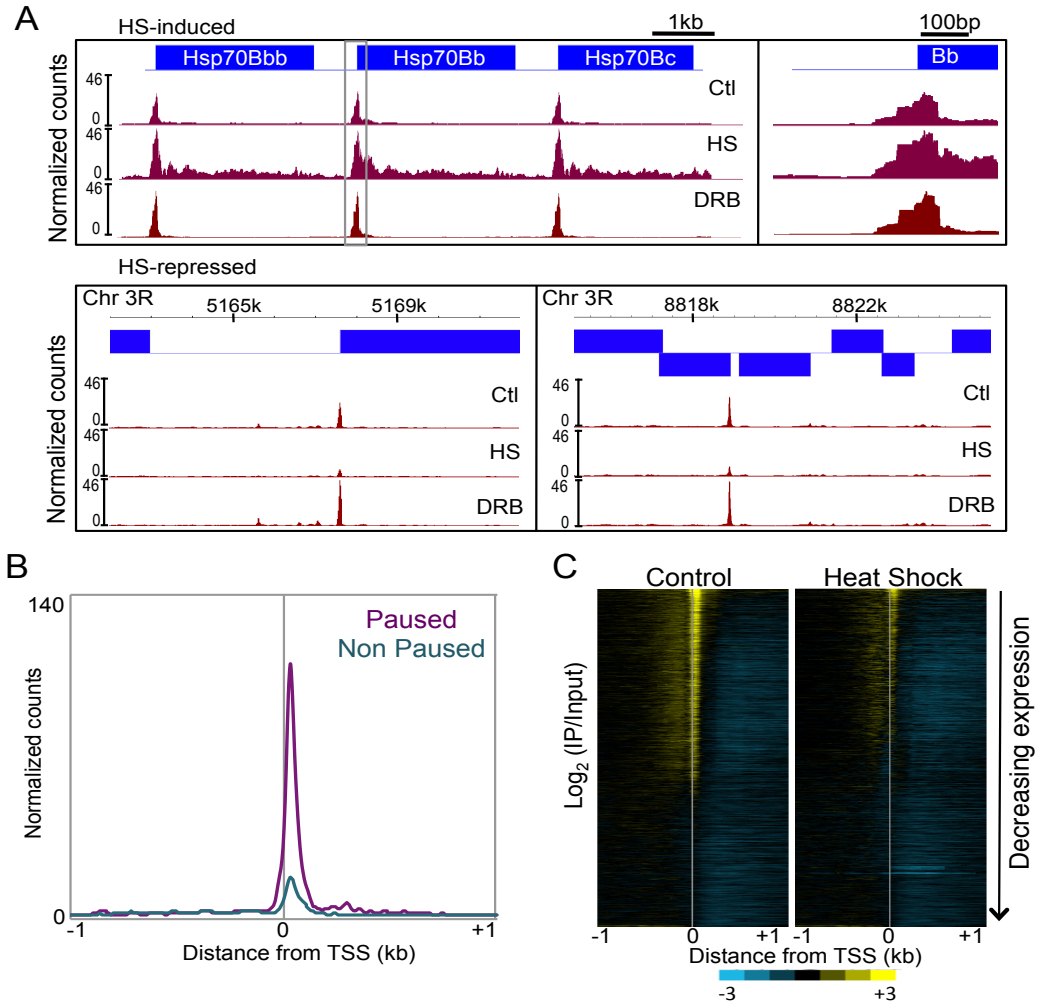

Figure 3-7: Effects of heat shock on low-salt soluble Pol II. (A) The mapped reads of the Pol II ChIP for control, heat shocked, and DRB-treated samples were converted to normalized counts for each base pair at the Hsp70 gene cluster, and landscapes were displayed with NimbleGen SignalMap. A close-up view of the boxed promoter region of Hsp70Bb is shown on the right panel. Landscapes for heat shock repressed genes are also shown for two regions in Chromosome 3R. (B) Average normalized counts for all paused (purple) and non-paused (cyan) genes in a 2-kb region surrounding the TSS. (C) Heat maps generated by Java TreeView (contrast = 4) of the $\log_2$ of Pol II ChIP to 80 mM input ratio in the 2-kb region surrounding the TSS for control (left) and heat shocked (right) samples. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

To visualize the genome-wide distribution of our low-salt soluble Pol II ChIP, we arranged the genes by decreasing level of expression and displayed them as a heat map of the 2-kb region surrounding the TSS (Figure 3-7 C). This showed a striking alignment of low salt soluble Pol II just downstream of the TSS that strongly correlated with expression level, suggesting that we have captured the generally 'stalled' Pol II that includes paused, backtracked, and arrested Pol II species (Levine, 2011). The large fraction of genes characterized by stalled Pol II in the soluble chromatin fraction is consistent with previous studies of paused genes (Core et al., 2008; Gilchrist et al., 2010; Gilchrist et al., 2008; Min et al., 2011; Muse et al., 2007), and supports the proposal that this fraction is enriched in paused Pol II (Weber et al., 2010).

To explore how the stalled Pol II changes genome-wide during heat shock, we performed Pol II ChIP and paired-end sequencing as described above using as input the 80 mM salt fraction extracted from cells heat shocked for 15 minutes at 37 °C. We found that during heat shock, stalled Pol II is present as irregular closely spaced peaks throughout the bodies of *Hsp70* genes (Figure 3-7 A). Most of these stalled Pol II ChIP peaks lined up with the 15-45 bp fragment peaks seen in the heat shocked 80 mM fraction used as input, confirming that a major component of this subnucleosomal chromatin is Pol II itself. This result also implies that Pol II stalls at multiple sites within bodies of the highly induced *Hsp70* gene, reminiscent of a recent report in which mapping of nascent transcripts in yeast showed that stalling throughout gene bodies is a common feature of highly expressed yeast genes (Churchman and Weissman, 2011).

Heat shock causes loss of Pol II from constitutively active puffs in *Drosophila* salivary glands (Jamrich et al., 1977). When we focused on non-HSF target genes containing prominent stalled Pol II peaks under normal conditions, we found that these peaks were greatly decreased upon heat shock but were maintained during DRB Pol II inhibition (Figure 3-7 A). Therefore, we

wondered whether the stalled Pol II that we mapped genome-wide is released from constitutively active genes in S2 cells. Indeed, a heat map display revealed genome-wide decreases in stalled Pol II during heat shock (Figure 3-7 C). To confirm that this decrease resulted from Pol II release rather than reduced solubility of Pol II-bound DNA, we performed quantitative Western blot analysis using a Pol II antibody for the input, unbound, ChIP, and pellet fractions of samples before and after heat shock (Figure 3-8 A). We detected a 50% reduction in insoluble Pol II

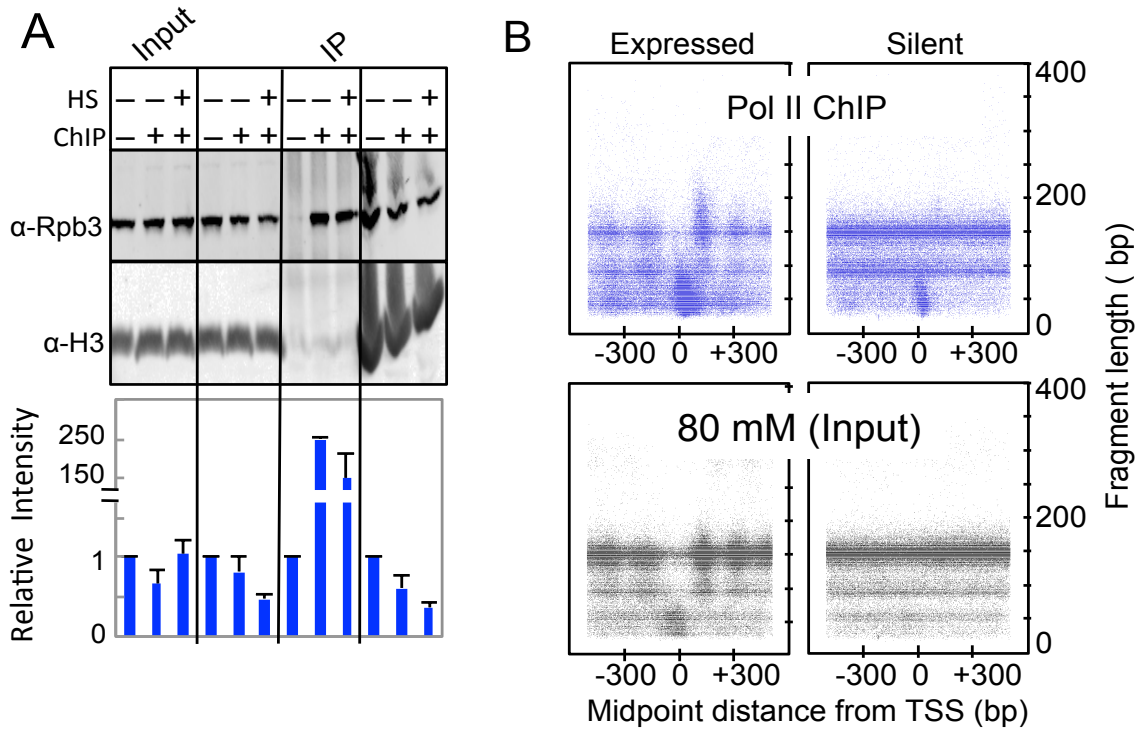

Figure 3-8: Properties of the stalled Pol II (A) Quantitative Western blot from a representative Pol II ChIP experiment. Pol II was probed using an antibody raised against the Rpb3 subunit. Histone H3 was used as a loading control. For both Pol II and H3, bands were quantified using ImageJ software. The ratio of Rpb3 to H3 for each lane was determined, and the values for each fraction (Input, Unbound, IP, and Pellet) were divided by the value for the ChIP no antibody control. (B) For all paired-end reads, the distance between the midpoint of the fragment and the TSS were determined and compared with the length of the fragment. All genes were separated into two groups, expressed and silent, and a random sample of 100,000 data points from each were used in the heat maps. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

during heat shock, which indicates that stalled Pol II became released from the DNA. However, we also detected a decrease in the unbound fraction without significant increase in global Pol II levels in the input fraction. Because the starting material in these experiments is nuclei, it is possible that Pol II that is unaccounted for has shuttled into the cytoplasm or degraded upon heat shock. Nevertheless, this result suggests that reduced Pol II affinity is a mechanism for heat shock mediated down-regulation of expression.

A large proportion of Pol II ChIP mapped reads were nucleosomal in size (Figure 3-9).

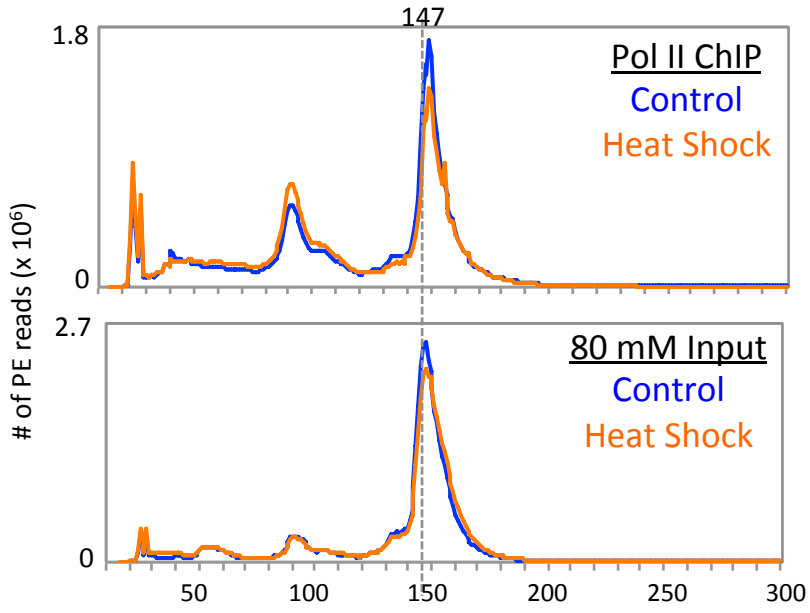

Figure 3-9: Pol II ChIP from 80 mM extracted active chromatin. The paired-end read length distribution of the Pol II ChIP and the corresponding 80 mM Input for samples before (blue) and after (orange) heat shock. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

We sought to determine whether these nucleosomes represent specific associations with stalled Pol II or non-specific background in the ChIP. For both the control Pol II ChIP and the corresponding input 80 mM fraction, we measured the distance between the midpoint of each fragment and the TSS, and plotted this distance versus the length of each fragment in a 1-kb region surrounding the TSS for expressed and silent genes separately (Figure 3-8 B). In the 80 mM salt fraction (the input for Pol II ChIP), the profiles of all size classes became evident in the

expressed genes, including the phased ~150 bp nucleosomes, the phased ~90 bp fragments from internal cleavage of nucleosomes, and the <60 bp fragments predominating at the TSS (Figure 3-8 B, bottom left). Inactive genes of the 80 mM fraction showed no phasing of nucleosomes and a reduced level of subnucleosomal particles at TSSs (Figure 3-8 B, bottom right). In comparison, the Pol II ChIP midpoint-versus-length plot for expressed genes showed an enrichment of small fragments especially downstream of the TSS (Figure 3-8 B, top left). Interestingly, we also observed larger (~200 – 250 bp) fragments centered near the +1 nucleosome, which would represent fragments spanning both stalled Pol II and the +1 nucleosome in the Pol II ChIP, but not detectable in the input material (Figure 3-8 B, left panels). In inactive genes, the ChIP midpoint-versus-length plot showed an enrichment of stalled Pol II at the TSS of some genes, whereas the remaining fragments displayed similar profiles as with the 80 mM input (Figure 3-8 B top right). Taken together, these observations suggest a physical interaction between stalled Pol II and the +1 nucleosome.

Inhibition of Pol II elongation decreases nucleosome turnover within gene bodies

Transcription levels are correlated with levels of nucleosome turnover within gene bodies in *Drosophila* (Deal et al., 2010), but a causal relationship has not been established. If transcriptional elongation is responsible for nucleosome turnover in gene bodies, then reduced elongation during heat shock will lead to a decrease in the level of turnover. To test this prediction, we applied the CATCH-IT (Covalent Attachment of Tags to Capture Histones and Identify Turnover) metabolic labeling procedure (Deal et al., 2010), which measures the kinetics of nucleosome turnover, in control and heat-shocked cells. Cells were fed with Azidohomoalanine (Aha)-containing media either at room temperature (control) or at 37 °C for

15 minutes. Aha-containing proteins within isolated nuclei were then conjugated with an alkyne-adapted biotin linker through a copper-catalyzed cycloaddition reaction. Following MNase digestion, chromatin was extracted in 350 mM NaCl, and Aha-biotin containing nucleosomes were affinity purified using Streptavidin-coated magnetic beads with stringent washing to remove H2A/H2B dimers. DNA from immunoprecipitated tetramers and from the 350 mM input fraction was isolated and prepared for paired-end sequencing as described above. At the *Hsp70* gene cluster, we observed that nucleosomes underwent a dramatic increase in turnover after 15 minutes of heat shock (Figure 3-10 A), suggesting that a large decrease in nucleosome

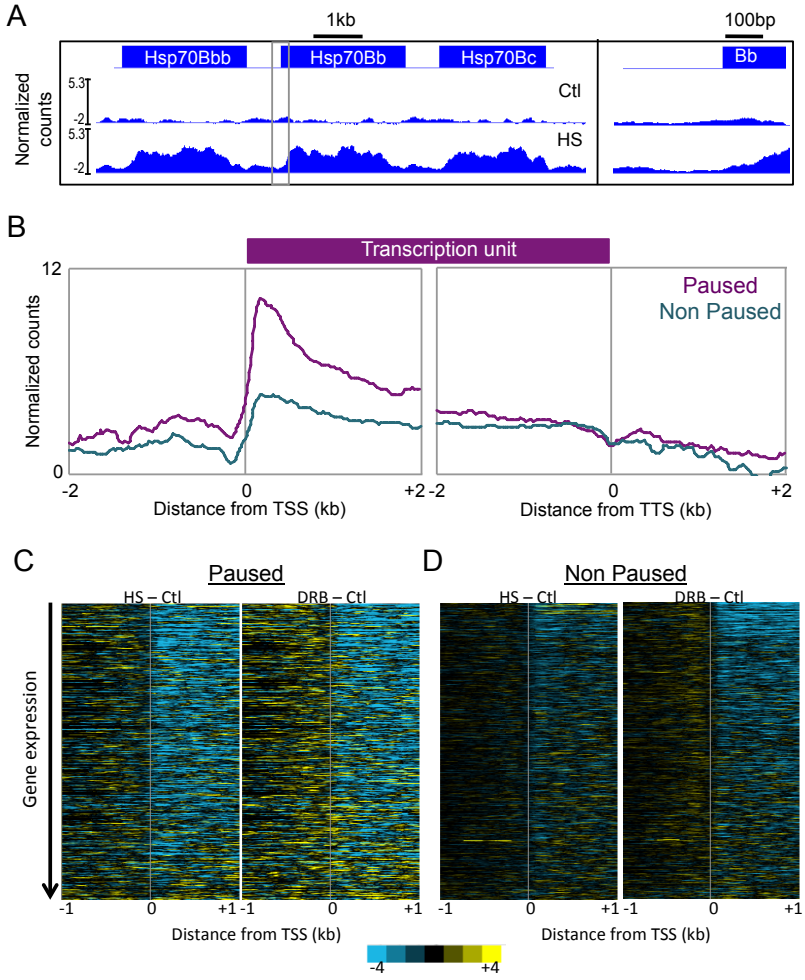

Figure 3-10: Nucleosomes within bodies of paused genes have high turnover relative to non-paused genes. CATCH-IT was performed before and after heat shock, and before and after DRB treatment. Input normalized control samples were subtracted from heat shocked samples (HS – Ctl) and from DRB treated samples (DRB – Ctl) to obtain changes in nucleosome turnover caused by heat shock and DRB treatment, respectively. A) The Hsp70 gene locus was visualized as described as in Figure 1C. (B) Average normalized counts for all paused (purple) and non-paused (cyan) genes in a 2-kb region surrounding the TSS and TTS of the input-normalized CATCH-IT experiment under normal growth conditions. (C) Heat maps generated by Java TreeView (contrast = 4) show changes in nucleosome turnover resulting from heat shock and DRB treatments in paused and non-paused genes ordered by decreasing expression under normal conditions. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

occupancy within seconds of heat shock shown in a previous study (Petesch and Lis, 2008) involves increased turnover. To determine the genome-wide changes in nucleosome turnover caused by heat shock, we subtracted the control from heat shock values and displayed the changes in a heat map with genes ordered by decreasing expression (Figure 3-11 A, left) or by changes in stalled Pol II due to heat shock (Figure 3-12 A). These analyses revealed a decrease in nucleosome turnover for most expressed genes while some previously silent *Hsp* genes showed

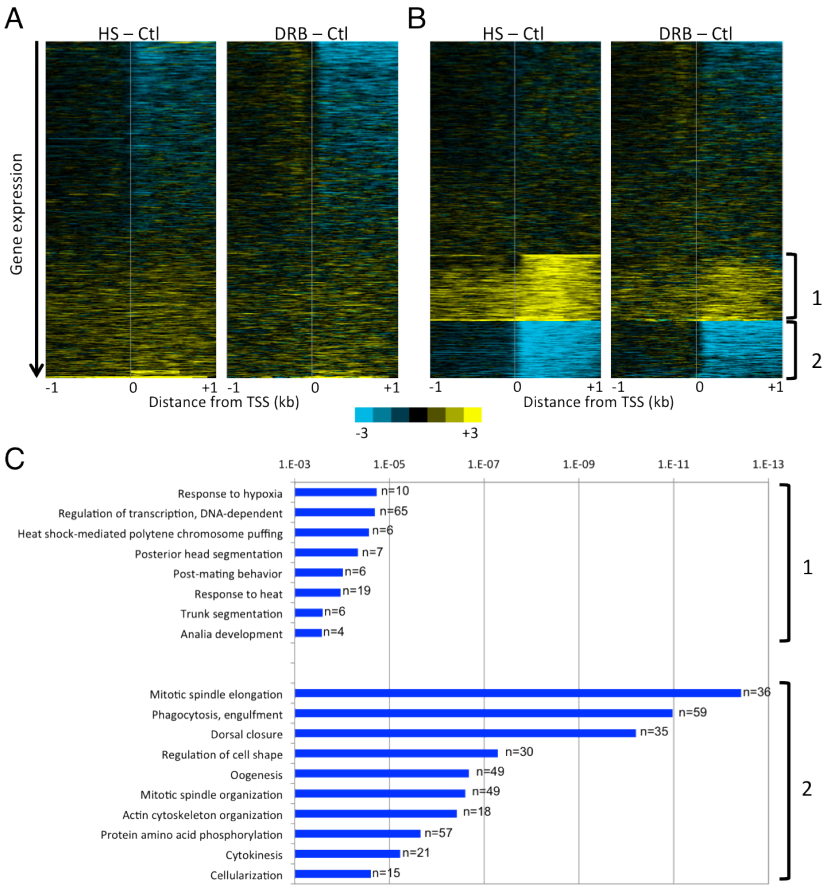

Figure 3-11: Genome-wide changes in nucleosome turnover during heat shock. (A) Heat maps generated by Java TreeView (contrast = 3) show changes in nucleosome turnover resulting from heat shock and DRB treatments, where all genes are ordered by decreasing expression under normal conditions. (B) Data from (A) for HS-Ctl was subjected to unsupervised k-means clustering with k = 3 (Cluster v. 3) and the DRB-Ctl heat map was arranged according to the HS-Ctl clustering. (C) GO term analysis of group 1 and 2 from the k-means clustering of the CATCH-IT data. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

increases in turnover, as well as a correlation between changes in stalled Pol II and nucleosome turnover. We then performed CATCH-IT on DRB-treated cells to ask whether turnover changes result from changes in transcriptional elongation. Indeed, inhibition of elongation with DRB led to a conspicuous reduction in nucleosome turnover (Figure 3-11 A, right), with a remarkable resemblance between the heat maps for heat shocked and DRB-treated cells. Furthermore, when the changes in turnover were displayed as a heat map with genes ordered by changes in stalled Pol II due to DRB treatment, we observed that the greatest turnover changes occur within gene bodies even though the major changes in stalled Pol II occur immediately downstream of the TSS (Figure 3-12 B). This suggests that much of the turnover within gene bodies is caused by the presumably insoluble elongating Pol II. The similar effects of heat shock and DRB inhibition of Pol II on turnover strongly suggests that down-regulation of expression during heat shock occurs after the recruitment of transcriptional machinery.

To further explore the relationship between transcriptional elongation and nucleosome turnover we analyzed the CATCH-IT profiles of paused and non-paused genes under normal conditions (Figure 3-10 B). Notably, paused genes showed consistently higher nucleosome turnover within gene bodies compared to non-paused genes. We then examined the effects of heat shock and DRB on nucleosome turnover for these two gene sets. Both treatments led to substantial decreases in nucleosome turnover within gene bodies of both paused and non-paused genes of matched expression (Figure 3-10 C), which argues that Pol II elongation drives nucleosome turnover regardless of how Pol II is regulated.

Finally, we asked about the molecular functions of the subset of genes that showed higher turnover during either heat shock treatment. To identify genes undergoing nucleosome turnover changes during heat shock treatment, we applied unsupervised k-means clustering with k = 3 to

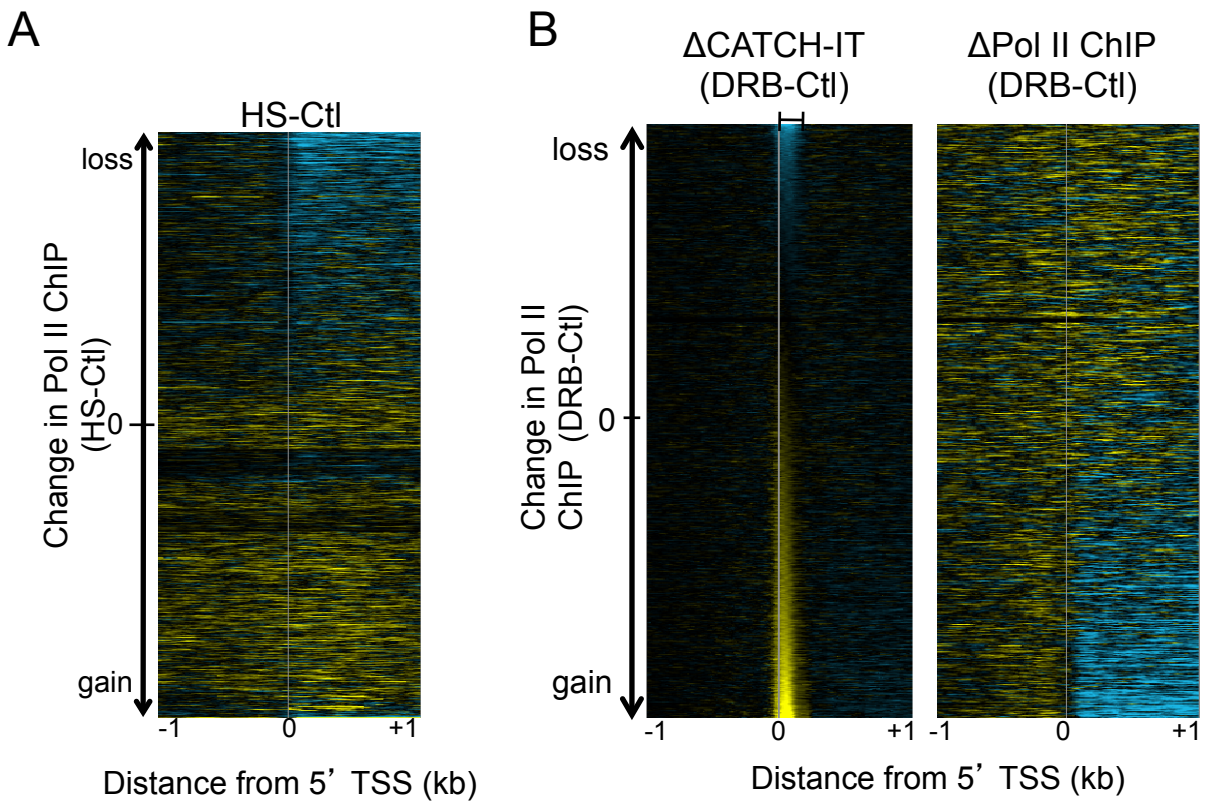

Figure 3-12: Transcription elongation affects nucleosome turnover within gene bodies. (A) Changes in stalled Pol II correlate with changes in nucleosome turnover due to heat shock**.** To show the relationship between changes in stalled Pol II and nucleosome turnover during heat shock, we have arranged the CATCH-IT HS-Ctl heat map by the change in stalled Pol II, from biggest loss (top) to biggest gain (bottom). This confirms that genes that lose the most stalled Pol II decrease the most in nucleosome turnover. (B) DRB inhibition of elongation results in reduced nucleosome turnover within gene bodies. We arranged all genes from biggest loss to biggest gain of stalled Pol II in the region from the TSS to 100 bp downstream (marked in brackets) and displayed the change in turnover as a heatmap, and displayed the CATCH-IT DRB-Ctl heat map accordingly. We see that although the major changes in Pol II ChIP are just downstream of the TSS, the greatest turnover changes are within gene bodies. This suggests that much of the turnover within gene bodies is caused by the presumably insoluble elongating Pol II. See Figure 3-11A for details. Reprinted with permission from Teves and Henikoff (2011) *Genes Dev.*

the heat shock – control data and applied this clustering to the DRB – control data (Figure 3-11 B). Heat map analysis showed a surprisingly similar pattern in nucleosome turnover changes in both treatments, suggesting that transcriptional inhibition has consistent effects on nucleosome dynamics regardless of the cause. Genes with increased turnover downstream of the TSS during both heat shock and DRB treatments were enriched for gene ontology (GO) terms for transcription regulation, and heat shock and hypoxia responses (Figure 3-11 C, group 1). The

largest class of genes that showed a significant increase in nucleosome turnover are involved in DNA-dependent regulation of transcription ($p=6x10^{-4}$, n=65), which suggests that nucleosome turnover at these genes is an inherent feature of the heat shock response. In contrast, the genes that showed reduced nucleosome turnover are predominantly those with developmental and other functions that are expected to be perturbed during heat shock (Figure 3-11 C, group 2). Discovering concordant nucleosome turnover changes at genes in the appropriate functional categories for the heat shock response itself suggests that regulation of turnover is an evolved response to environmental perturbation.

Discussion

We have shown that heat shock induction involves genome-wide changes in chromatin solubility, stalled Pol II occupancy and nucleosome turnover in bodies of expressed genes. Using new methods for epigenome profiling and analysis, we have mapped both nucleosomes and subnucleosomal particles from the soluble chromatin fraction at single base-pair resolution from control and heat shocked cells. In addition, we have determined the stalled Pol II profiles derived from this fraction, and have shown that loss of Pol II at active genes during heat shock correlates with reduced nucleosome turnover, thus providing evidence for a causal relationship between Pol II transit and nucleosome turnover in gene bodies. A similar pattern of reduced nucleosome turnover was also observed following drug inhibition of Pol II elongation, which implies that Pol II sometimes evicts nucleosomes that it encounters during elongation. Loss of stalled Pol II, and subsequent reduction in genic nucleosome turnover, might reflect a general mechanism for repression of diverse genes during heat shock, while maintenance of transcription factor binding at the promoters would facilitate efficient recovery (Zanton and Pugh, 2006).

Several studies that have profiled nucleosomes genome-wide used MNase digestion followed by size selection for nucleosome-protected fragments (Kent et al., 2011; Li and Arnosti, 2010; Mavrich et al., 2008; Xi et al., 2011). However, by using salt fractionation of native chromatin to selectively extract the soluble fraction, by omitting size selection, and by sequencing the paired ends of fragments of all sizes, we have been able to investigate structural properties and dynamics for both nucleosomal and subnucleosomal components of the epigenome. For example, the footprint of subnucleosomal particles can be inferred from the square peaks of short fragments at the promoters of *Hsp70* genes, implying that the bound particles confer complete protection of DNA by being precisely positioned over their binding sites. In contrast, the spontaneous unwrapping and rewrapping of nucleosomes allow for MNase encroachment into nucleosomal DNA (Li et al., 2005) thereby giving nucleosomal peaks rounded shapes that signify average protection. Furthermore, the precise mapping of subnucleosomal particles has allowed us to use the low-salt fraction as starting material for native Pol II ChIP experiments, which indicates that the stalled Pol II can be specifically extracted from the soluble chromatin fraction. By sequencing all MNase-protected fragments from the ChIP and analyzing their midpoint-to-length relationships, we have uncovered direct evidence in support of a proposed interaction between the stalled Pol II and the highly dynamic TSS nucleosome (Gilchrist et al., 2010; Weber et al., 2010). Thus, salt fractionation combined with a modified library preparation protocol and a simple computational analysis provide a powerful tool for investigating the structure, profile and dynamics of the transcriptionally active epigenome.

Using this tool, we have also found that active genes that lose stalled Pol II during heat shock show no overall changes in nucleosomal occupancy, comparable to a yeast study showing

that heat shock does not affect global nucleosome occupancy (Shivaswamy et al., 2008). However, we also observed a genome-wide decrease in the low-salt soluble nucleosomal fraction. How would a decrease in transcription affect nucleosome solubility? Previously, we showed that low salt solubility largely corresponds to nucleosomes containing the conserved histone variant H2A.Z (Weber et al., 2010), which in Arabidopsis is responsive to changes in ambient temperature (Kumar and Wigge, 2010). Therefore, it is possible that transcriptional repression results in decreased H2A.Z occupancy within bodies of expressed genes, and as a result, nucleosome solubility decreases.

As we found to be the case for nucleosome occupancy, heat shock appears to have no effect on the average occupancies of subnucleosomal particles, in contrast to its major effect at the TSSs of heat shock genes. Such maintenance of factor binding at the promoter region of transcribed genes suggests that down-regulation during heat shock occurs 3' of the TSS. Evidence in support of downstream regulation during heat shock comes from the observation that SINE RNAs induced by heat shock disrupt elongating Pol II contacts with promoter DNA but do not disrupt transcription factor recruitment or binding *in vitro* (Yakovchuk et al., 2009). In addition, most of the factors involved in the Pre-Initiation Complex remain bound to promoters after heat shock in budding yeast (Zanton and Pugh, 2006), further supporting the notion that repression occurs after recruitment.

The fact that low salt extraction specifically enriches for stalled Pol II suggests inherent differences between stalled and actively elongating polymerases. Actively elongating Pol II is highly insoluble under standard extraction methods (Cabrerizo et al., 1999; Courvalin et al., 1976; Kimura et al., 1999) and nucleosomes of some transcribed genes remain insoluble even after high salt treatments that solubilize most of the nucleosomes (Henikoff et al., 2009), which

implies an association between these nucleosomes and the large, hydrophobic complexes involved in transcription. This further suggests that stalled Pol II disengages from these large complexes or from transcription factories (Eskiw et al., 2008; Papantonis et al., 2010) making it more soluble during extraction. The large fraction of genes with Pol II signals that we observed suggests that low-salt extraction in general captures stalled Pol II. Such widespread Pol II stalling implies a discontinuous elongation process as has been shown *in vitro* (Herbert et al., 2006), and *in vivo* in *S. cerevisiae* (Churchman and Weissman, 2011). Furthermore, our study showed that heat shock causes partial release of stalled Pol II from DNA just downstream of TSSs, which is an attractive mechanism for transcriptional repression. Pol II release after heat shock has been previously observed in *Drosophila* salivary gland chromosomes (Jamrich et al., 1977). A similar Pol II release process has been proposed in mammals where heat shock results in increased release of bound Pol II from DNA (Hieda et al., 2005).

We also showed that heat shock results in increased nucleosome turnover at heat shock induced genes and decreased turnover elsewhere over gene bodies. These observations imply a causal relationship between transcriptional repression and nucleosome turnover that was previously inferred from correlations between expression levels and H3.3 (Mito et al., 2005) or turnover rates (Deal et al., 2010) over gene bodies. Furthermore, the similarity between heat shock and DRB-induced nucleosome turnover patterns suggests that this causal relationship is an inherent general feature of eukaryotic transcription, although it is unknown whether these effects on transcriptional elongation involve the insoluble RNA polymerase fraction, which cannot be probed by ChIP. *In vitro* studies have revealed that a nucleosome can survive transcription by a single transcribing Pol II but becomes evicted when bursts of multiple Pol IIs pass through (Jin et al., 2010; Kulaeva et al., 2009; Kulaeva et al., 2010). This *in vitro* observation, taken together

with our direct measurements of changes in nucleosome turnover *in vivo*, suggests that both the level of transcription and the rate of Pol II bursting are major determinants of nucleosome eviction in a large portion of the soluble chromatin.

It seems counterintuitive that paused genes would show higher turnover within bodies than non-paused genes (Figure 3-7), insofar as pausing has been proposed to be a mechanism for preventing Pol II bursting (Levine, 2011). However, this and other proposals have only considered events around the TSS, whereas our study is unique in observing features of paused genes throughout gene bodies. One proposal for the immediate disruption of downstream nucleosomes at the 87A *Hsp70* locus during heat shock (Petesch and Lis, 2008) is that the sudden acceleration of Pol II transit elicits a wave of positive supercoiling ahead that causes nucleosomes to unwrap before the topoisomerase swivel can relieve the strain (Zlatanova and Victor, 2009). The fact that ordinary genes show high nucleosome occupancy when they are active (Cui et al., 2010) suggests that transcriptional elongation has evolved such that eviction of nucleosomes upon sudden acceleration of RNA polymerases is minimized. By this rationale, nucleosome eviction would also be minimized by favoring a more uniform rate of polymerase transit, thus reducing damaging acceleration events. Therefore, we speculate that the starts and stops that must occur when RNA Pol II pauses during its transit result in enhanced nucleosome eviction, followed by replacement, which is what we profiled in our CATCH-IT experiments. Paused genes would simply undergo more frequent damaging acceleration events than non-paused genes and so show a higher rate of nucleosome turnover. In this way, events that occur generally within gene bodies and events around the TSSs of paused genes can be seen as resulting from common properties of the Pol II transcriptional apparatus.

CHAPTER 4

RNA POLYMERASE II-GENERATED TORSIONAL STRESS DESTABILIZES NUCLEOSOMES

Summary

As RNA Polymerase II (Pol II) transcribes a gene, it encounters an array of well-ordered nucleosomes. *In vitro*, a single nucleosome presents a barrier to Pol II elongation, but how it traverses through this array *in vivo*, remains an unresolved question. One model proposes that torsional stress generated during Pol II translocation destabilizes nucleosomes ahead of Pol II. Here, we provide a method for high resolution mapping of Pol II-induced underwound DNA using next-generation sequencing, and show that torsional stress is correlated with gene expression in *Drosophila* cells. Inhibition of topoisomerases, enzymes that relieve torsional strain, leads to accumulation of torsional stress, and occurs concomitantly with an increase in nucleosome turnover and salt solubility within gene bodies. Pol II also accumulates at transcription start sites, but Topoisomerase I inhibition results in immediate increase in nascent RNA transcripts whereas Topoisomerase II inhibition shows little change. The elongation-independent effects of torsional stress on nucleosome dynamics within gene bodies suggest that torsional stress contributes to the destabilization of nucleosomes.

Introduction

The wrapping of DNA around octameric histone proteins to form nucleosomes accomplishes the feat of packaging DNA within the nucleus with the consequence of restricting DNA access. This is problematic for DNA-templated processes including replication and transcription (Cairns, 2009). For transcription, cells have evolved to maintain promoter regions relatively depleted of nucleosomes, allowing DNA access to transcription factors and the RNA Polymerase II (Pol II) machinery (Mavrich et al., 2008; Zhang and Pugh, 2011). However, as Pol II leaves the promoter region, it encounters a well-ordered nucleosomal array (Mavrich et al., 2008; Zhang and Pugh, 2011). *In vitro*, a single nucleosome effectively blocks Pol II elongation (Izban and Luse, 1992; Lorch et al., 1987; Shaw et al., 1978), but within the cell, Pol II transcribes chromatin templates with high efficiency. How does Pol II transcribe through the nucleosome? Based on *in vitro* work, several models have been proposed, ranging from complete survival of the nucleosome to complete dissolution (Lavelle, 2007). *In vivo*, transcription has been shown to cause nucleosome turnover within gene bodies (Teves and Henikoff, 2011), perhaps aided by nucleosome remodelers. In the case of *Hsp70* genes in *Drosophila*, gene activation results in decreased nucleosome occupancy prior to passage of the first transcribing Pol II (Petesch and Lis, 2008), hinting at a mechanism for destabilizing nucleosomes ahead of Pol II.

Aside from impacting nucleosome structure and dynamics, Pol II also affects DNA topology. During transcription, the double helix is melted to form the transcription bubble, and resistance to Pol II rotation relative to DNA creates torsional stress that manifests as positive and negative supercoils downstream and upstream of Pol II, respectively, referred to as the twin-supercoiled domain model (Giaever and Wang, 1988; Liu and Wang, 1987). To counteract the

accumulation of torsional stress, topoisomerases resolve supercoils by either introducing a single-stranded nick followed by rotation about the other strand (type I topoisomerases) or creating transient double-stranded breaks and passing another unbroken DNA helix through (type II topoisomerases) (Wang, 2002). Even in the presence of topoisomerases, however, torsional stress has been detected *in vivo*, primarily through the use of the molecule psoralen and its derivatives (Kramer and Sinden, 1997). Psoralen is a DNA intercalating molecule that prefers underwound DNA resulting from negative supercoiling, and in the presence of UV light, induces interstrand crosslinks. Psoralen binding has been used to visualize global unconstrained supercoiling in *Drosophila* polytene chromosomes (Matsumoto and Hirose, 2004), and to map genome-wide supercoiling in yeast (Bermúdez et al., 2010) and human cells (Kouzine et al., 2013; Naughton et al., 2013). Whereas one study showed that supercoiling remodels large scale chromatin domains (Naughton et al., 2013), the low resolution remains a limiting factor in understanding the relationship between transcription, supercoiling and nucleosome dynamics.

Torsional strain induces structural changes on the double helix, which has consequences for DNA binding proteins. For example, in human cells, underwinding of DNA upstream of the c-Myc gene allows for binding of regulatory factors that control c-Myc expression (Kouzine et al., 2004; Kouzine et al., 2008). Nucleosomes, the most abundant DNA-protein structure, are differentially affected by positive and negative supercoiling. *In vitro*, negative supercoiling promotes nucleosome assembly, whereas positive supercoiling inhibits it (Gupta et al., 2009). Furthermore, preferential exchange of nucleosomes from positively to negatively supercoiled DNA has been shown *in vitro* (Clark and Felsenfeld, 1991), suggesting that Pol II-generated supercoiling might participate in nucleosome destabilization during transcription. It remains unknown how Pol II-generated torsional strain *in vivo* affects nucleosome dynamics.

Understanding this interplay has potential clinical implications, as widely used cancer chemotherapeutic anthracycline drugs that intercalate into DNA and induce positive torsion have recently been shown to increase nucleosome turnover and eviction around active promoters (Pang et al., 2013; Yang et al., 2013).

To test the effect of transcription-generated torsional stress on nucleosome dynamics in *Drosophila* cells, we have developed a new paired-end sequencing protocol to measure torsional states genome-wide at single-base resolution. We find that transcribed genes experience higher torsional stress than silent genes, consistent with the twin-supercoiled domain model. After inhibiting topoisomerases individually, we show that torsional stress accumulates surrounding gene bodies, which is accompanied by changes in Pol II levels and transcription genome-wide. Regardless of change in expression levels, nucleosomes within gene bodies increase in turnover, providing *in vivo* evidence for torsion-mediated nucleosome dynamics. The relationship between DNA torsion and nucleosome dynamics suggests an intricate balance between efficient Pol II progression and maintenance of the nucleosomal template.

Results

A high-resolution genome-wide assay for detection of DNA supercoiling states

Several methods have recently been developed for large-scale detection of supercoils in yeast (Bermúdez et al., 2010) and in human cell lines (Kouzine et al., 2013; Naughton et al., 2013) but the relationship between supercoiling and transcription remains vague. For instance, no relationship was seen in yeast (Bermúdez et al., 2010) while studies in human cells found a correlation (Kouzine et al., 2013; Naughton et al., 2013). To address this discrepancy, we achieved single base resolution mapping of DNA supercoils by adapting the micro-array-based

method (Bermúdez et al., 2010) for next generation sequencing. We exposed *Drosophila* S2 cells to Trimethyl-psoralen (TMP), which preferentially intercalates into negatively supercoiled DNA and covalently cross-links both strands upon exposure to 365 nM UV light. Following DNA extraction and shearing to ~250 bp average, we enriched for cross-linked DNA fragments by multiple rounds of denaturation followed by Exonuclease I (Exo I) digestion, which preferentially digests single stranded DNA (Figure 4-1A) (Bermúdez et al., 2010). We then

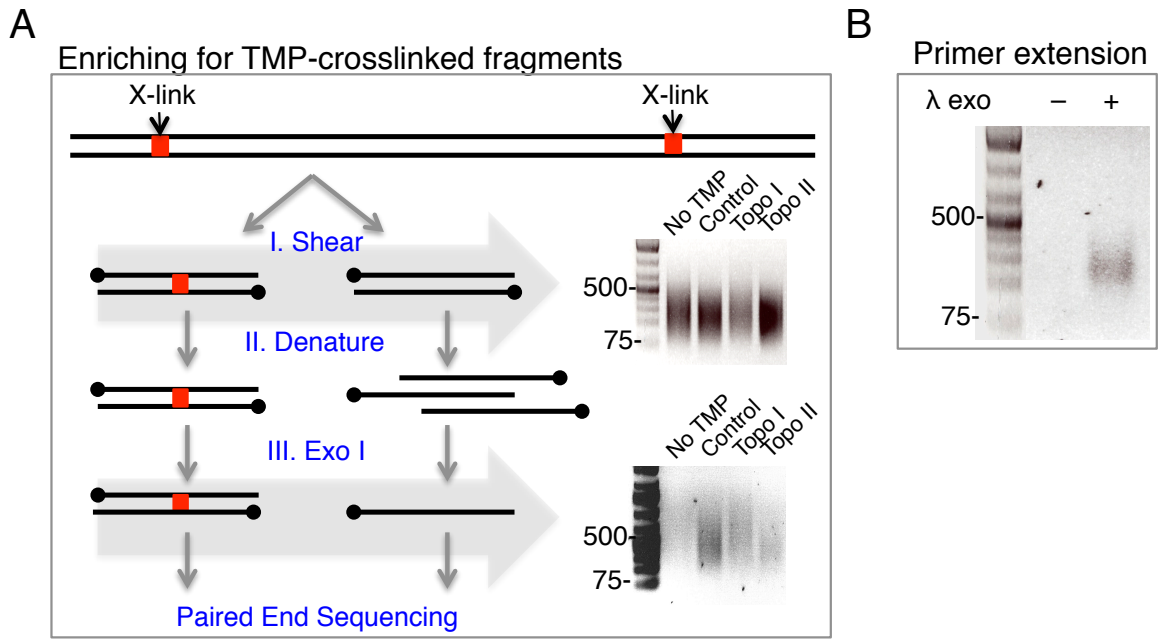

Figure 4-1: Strategy to enrich for TMP-crosslinked fragments. (A) TMP-crosslinked DNA is sheared to an average size of 250 bp (top agarose gel). Samples were denatured in boiling water bath for 10 minutes, snap-cooled in ice water, and digested with Exo I. This process was repeated until the sample without TMP crosslinks have been fully digested (bottom agarose gel). (B) After PE-adapters are ligated (see Figure 1), samples were split into 2 aliquots. The first was subjected directly to primer extension reaction, while the second was 3' strand resected using λ Exo prior to the primer extension reaction. Without λ Exo digestion, no single stranded products were detected.

polished the ends and ligated Illumina barcoded adapters as previously described (Henikoff et al., 2011). Subsequently, we digested the 5' strand of the PE-adapted material with λ exonuclease until the cross-linked nucleotide inhibits further digestion (Figure 4-2 A). Using a primer complementary to the PE adapter, we performed 10 rounds of primer extension that end at the

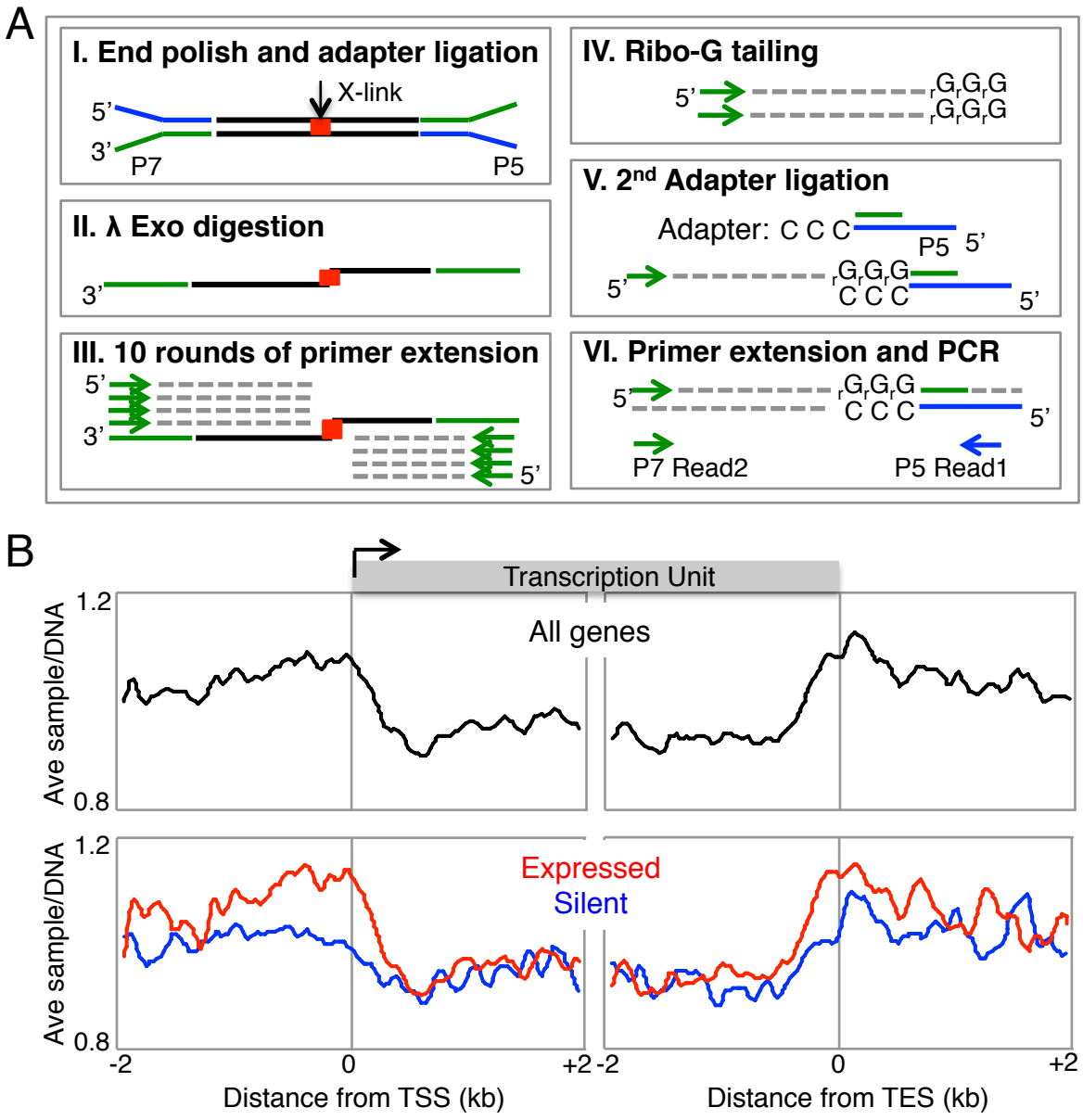

Figure 4-2: High resolution detection of supercoiling states
(A) Strategy for paired-end sequencing of TMP cross-linked DNA fragments. See text for details. (B) The average TMP-seq signal for every 10 bp in a 4 kb region surrounding the TSS and TES of all genes is plotted (top), and for expressed and silent genes separately (bottom).

cross-linked site. When the λ exonuclease digestion was omitted, no single-stranded extension products were observed (Figure 4-1 B). We then extended the ssDNA products with ribo-Gs using Terminal Transferase, and ligated a double stranded adapter that has 5' CCC overhang. After a single round of primer extension, we amplified the resulting libraries using standard Illumina primers. Sequencing from the CCC overhang end allows for nucleotide resolution of the interstrand cross-link (Figure 4-2 A). We refer to this method as TMP-seq. As a control for sequence bias, we added TMP to purified genomic DNA, crosslinked by UV light exposure, and processed in parallel to TMP-treated S2 cells. We then mapped the nucleotide position of the cross-links from samples and DNA controls onto the genome and fit a kernel density distribution around each site (Gehring et al., 2009). A representative region in chromosome 3R is shown (Figure 4-3 A, top) and a zoomed in region at the *Hsp70C* locus shows TMP binding upstream of the transcription start site (TSS) (Figure 4-3 A, bottom), suggesting that promoters of genes experience negative supercoiling. To normalize for sequence biases, we calculated the ratio of each sample to the genomic DNA control. We then averaged the normalized TMP-seq signals around the TSS and transcription end sites (TES) for all genes. This shows that TMP signals are high upstream of genes and low within gene bodies, consistent with the twin supercoiled domain model, which predicts that Pol II generates negative and positive supercoils upstream and downstream of the TSS, respectively (Figure 4-2 B, top). Interestingly, TMP signals rose again at the TES. Actively transcribed genes showed higher TMP levels at the promoter region than silent genes, as predicted by the twin supercoiled domain model (Figure 4-2 B, bottom).

Toposiomerases relieve the torsional stress generated during transcription. We individually inhibited Topo I and Topo II in S2 cells using Camptothecin and ICRF-193, respectively, and measured supercoiling levels using TMP-seq. At the *C* locus, TMP-seq signals

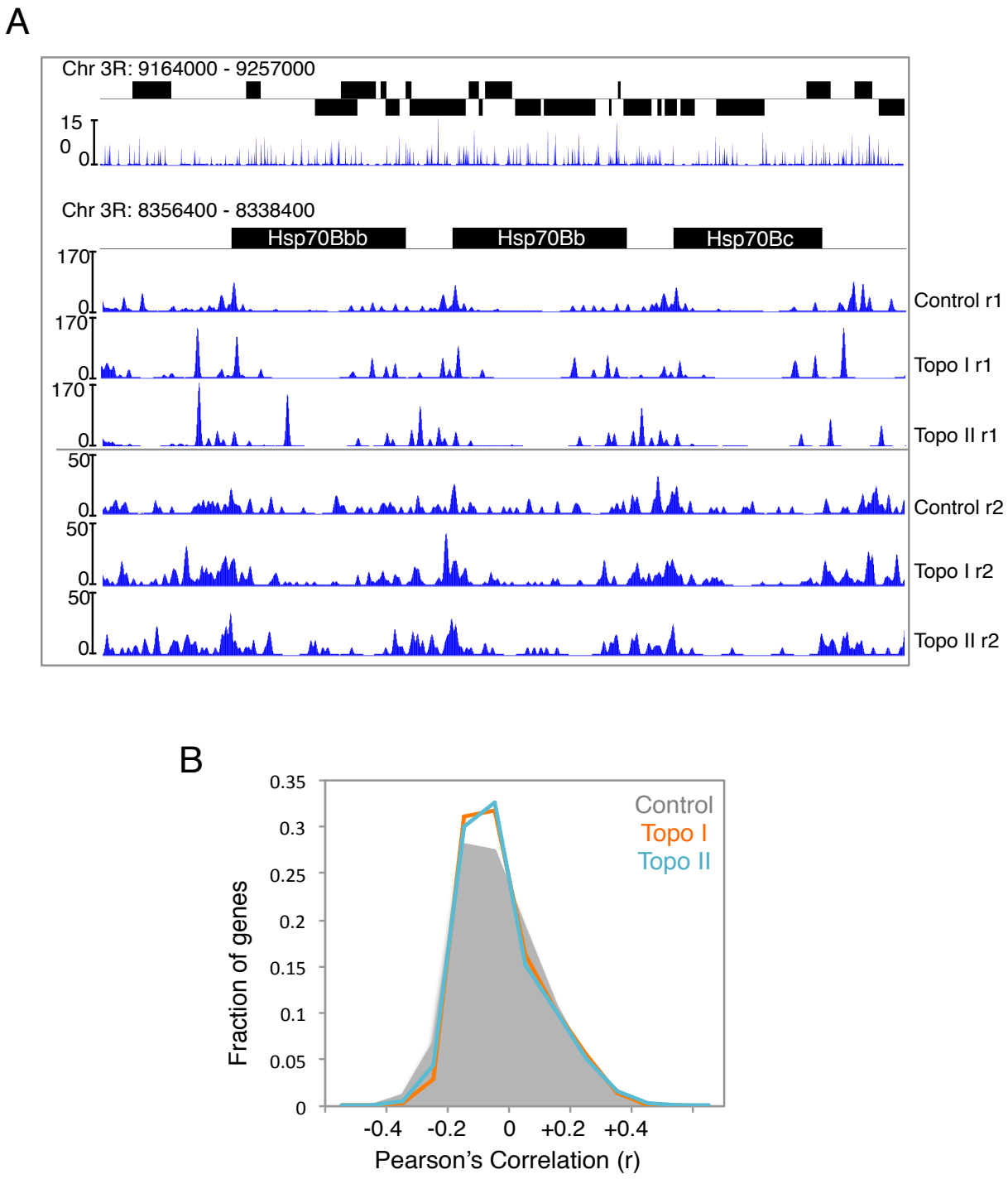

Figure 4-3: TMP-seq largely detects torsion near the TSS. (A) The first sequenced nucleotide after the CCC overhang is mapped onto the genome and fitted with a Kernel density estimator function (bandwidth = 20) (Gehring et al., 2009). A 93 kb representative region in the chromosome 3R is shown at the top. The *Hsp70C* gene locus is shown with replicates for control, Topo I-, and Topo II-inhibited samples. (B) For each gene and for each sample, the Pearson's correlation (r) was determined for the 1 kb region surrounding the TSS between TMP-seq and nucleosome occupancy data, and the distribution of the correlation was plotted.

upstream of the TSS change in distribution and intensity after Topo I or Topo II inhibition (Figure 4-3 A). We then averaged the TMP-seq signals genome-wide and found increased levels at the TSS relative to a no-treatment control, and decreased within gene bodies, suggestive of supercoils, both negative and positive, accumulating in the absence of topoisomerase activity (Figure 4-4 A). This accumulation effect within gene bodies is much larger in transcribed genes (Figure 4-4 B) than in silent genes (Figure 4-4 C), consistent with the known role of both enzymes in transcription. Notably, Topo I inhibition resulted in a larger effect on accumulated torsional stress compared to Topo II inhibition, particularly for expressed genes. This is consistent with previous studies suggesting that Topo I is the major relaxer of transcription-generated torsional stress (Durand-Dubief et al., 2010; Sperling et al., 2011).

Previous studies have shown that, under saturating TMP conditions, nucleosomes inhibit TMP crosslinking (Bermúdez et al., 2010; Potter et al., 1980). Although we used limiting amounts of TMP, it is possible that topoisomerase inhibition results in nucleosome occupancy changes, which in turn are responsible for the observed changes in the TMP-seq profile. We determined nucleosome occupancy for control, Topo I-, and II-inhibited samples, and calculated the Pearson's correlation for each gene between nucleosome occupancy and TMP-seq data in the 1 kb region surrounding the TSS. We then plotted the histogram of Pearson's correlations (Figure 4-3 B), and found that there is an overall slight anti-correlation (36% of genes have $r \leq -0.2$) between nucleosome occupancy and TMP intercalation in control, as expected from previous studies. However, after Topo I or II inhibition, the correlation distributions did not shift, suggesting that the change in TMP-seq profile after topoisomerase inhibition is not a consequence of changes in nucleosome occupancy, but rather is largely due to changes in torsional state.

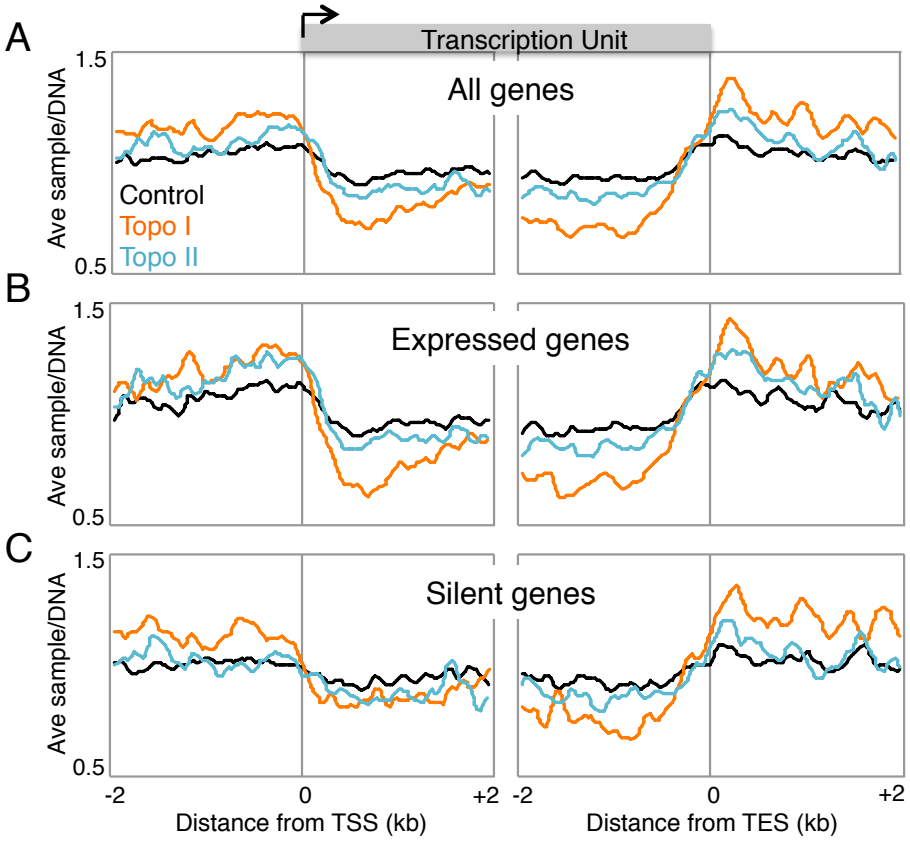

Figure 4-4: Topoisomerase inhibition increases torsional stress genome-wide
(A) Control cells and cells treated with Topo I or II inhibitor were subjected to TMP-seq and the average normalized signals surrounding the TSS and TES were determined. The genes were separated by transcription status, and average TMP-seq profiles for control, Topo I-, or Topo II-inhibited were determined for expressed (B) and silent (C) genes.

Low-salt soluble nucleosomes are sensitive to torsional stress

A previous study has shown that Pol II-generated supercoils affect large-scale chromatin organization, forming underwound and overwound domains (Naughton et al., 2013). We asked if supercoiling can affect chromatin organization at a finer scale, focusing on genic regions, using low salt solubility as measure. Chromatin of genic regions is enriched for DNase I hypersensitivity (Crawford et al., 2004; Keene et al., 1981) and MNase accessibility (Mavrich et al., 2008), reflecting the overall decondensed state of these regions. This state is characterized by reduced inter-nucleosomal interactions such that, after MNase-digestion, these nucleosomes become readily soluble in low salt, whereas condensed chromatin requires high salt for solubilization (Henikoff et al., 2009; Sanders, 1978). We previously showed that low-salt soluble

chromatin, which comprises ~10% of total chromatin, is enriched for both mono-nucleosomes and DNA binding factors at the TSS (Teves and Henikoff, 2011; Weber et al., 2010). We therefore asked whether torsional stress affects low salt solubility of nucleosomes and DNA binding factors. To test this, we inhibited topoisomerases individually, extracted low salt soluble chromatin, and performed paired-end sequencing (Henikoff et al., 2009; Teves and Henikoff, 2011). We then parsed out computationally fragments with lengths greater than 120 bp to represent nucleosomes and fragments 25 – 45 bp in length to represent DNA-bound factors. After mapping these fragment length classes separately onto the genome, we averaged the normalized counts 1 kb surrounding the TSS of all genes for control and Topo I- or Topo II-inhibited cells. However, low salt solubility of nucleosomes dramatically increased within gene bodies while promoter regions experienced decreased solubility (Figure 4-5 B). This observation suggests that, while nucleosome occupancy is subtly affected, torsional stress significantly alters the inter-nucleosomal interactions that determine salt solubility. We then displayed the change in salt solubility for nucleosomal fragments due to Topo I or Topo II inhibition relative to control as a heat map with genes arranged by decreasing mRNA abundance in control. This analysis showed that transcribed genes increase in low salt solubility (Figure 4-5 C), suggesting that torsional stress primarily affects chromatin organization of transcribed genes. As a control, we also performed the same analysis for total nucleosomes to determine changes in occupancy. For total nucleosomes, topoisomerase inhibition led to only a slight increase in overall occupancy near 3' ends of genes (Figure 4-5 A).

We have previously shown that the number of sequenced short fragments in the low salt soluble fraction corresponds to occupancy of DNA binding factors at the TSS (Henikoff et al., 2011; Teves and Henikoff, 2011). To analyze the effect of torsional stress on the occupancy of

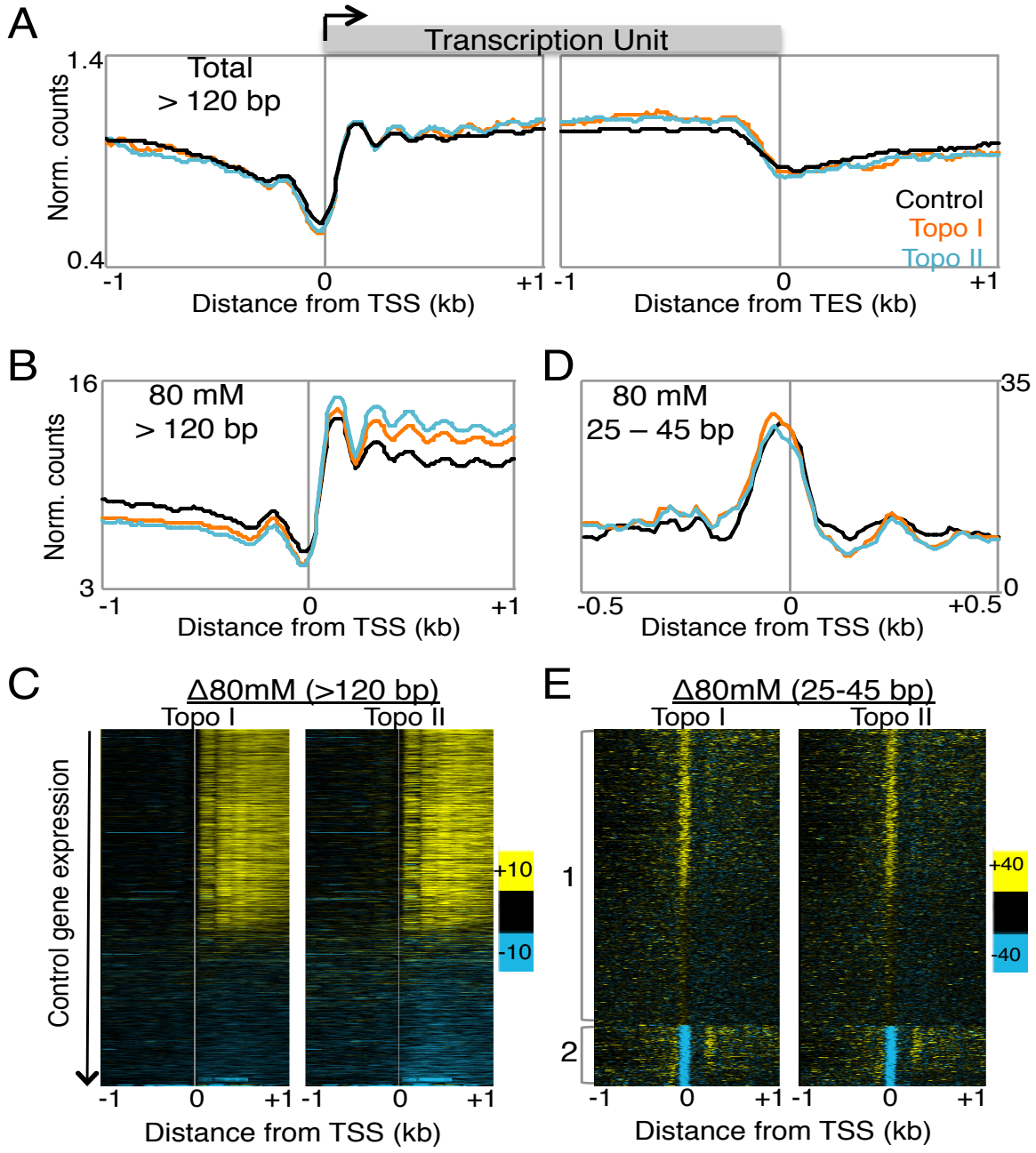

Figure 4-5: Topoisomerase inhibition affects low-salt chromatin fractionation
(A) Total MNase digested chromatin was sequenced as before (Teves and Henikoff, 2011) from treated and untreated cells, and the average signal in all genes for >120 bp reads were plotted. (B) Nuclei from topoisomerase inhibitor treated and untreated cells were isolated and subjected to MNase digestion under native conditions. The low-salt (80 mM) soluble fraction was isolated and sequenced using the modified library preparation protocol (Henikoff et al., 2011). Reads >120 bp were parsed out and mapped onto the genome and the average signals surrounding the TSS was determined. (C) The difference in low-salt soluble signals between control and Topo I- (left), and II-inhibited (right) samples were determined and displayed as heatmaps with genes ordered in decreasing expression in control samples for reads of length >120 bp (D) Same as (B) but for reads of length 25 – 75 bp. (E) Same as (C) but for reads of length 25 – 75 bp.

DNA binding factors, we averaged the signals of the short fragments (25-45 bp) surrounding the TSS of all genes. Topoisomerase inhibition did not alter the overall occupancy at the TSS (Figure 4-5 D), but when the change in occupancy was displayed as heat maps, we detected heterogeneous changes at the TSS of transcribed genes (Figure 4-6 A). Using k-means clustering with k=2 for either Topo I- or Topo II-inhibited samples, 2 distinct groups were detected. The first group (Group 1) comprises of 83% of all genes with varying levels of increased occupancy at the TSS (Figure 4-5 E). The second group (Group 2) comprises of 17% of all genes whose occupancy at the TSS decreased upon either topoisomerase inhibition. Strikingly, there is 78% overlap between genes in Group 2 of Topo I inhibited versus Topo II inhibited samples (Figure 4-6 B), among which are genes encoding for ribosomal components (Figure 4-6 C) known to be highly down-regulated upon topoisomerase inhibition (Brill et al., 1987; Muller et al., 1985), suggesting that the occupancy of factors at the TSS of these genes is sensitive to torsional stress.

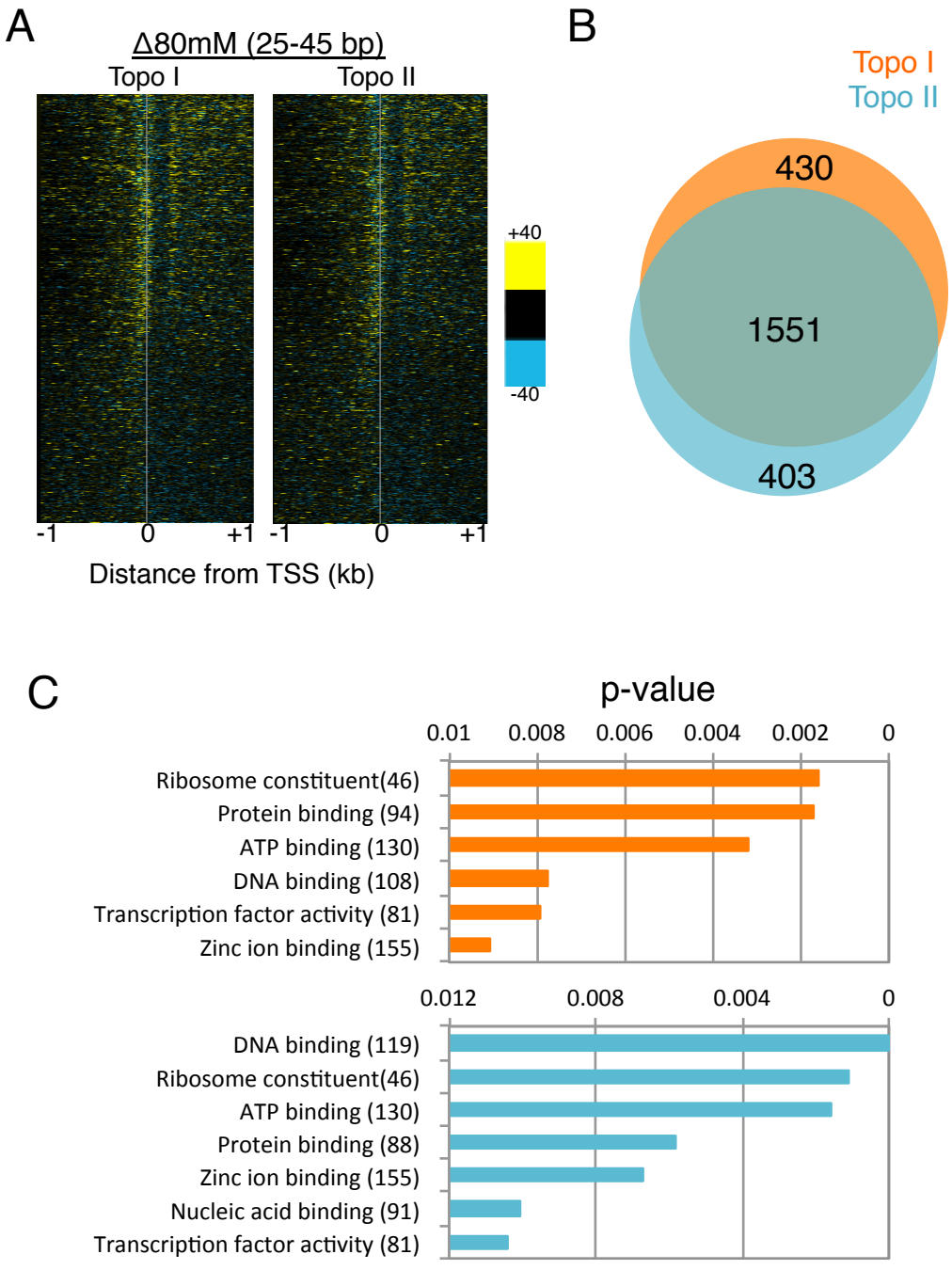

Figure 4-6: Torsional stress affects DNA binding factors at the TSS. (A) For short read fragments (25 – 45 bp), the change in signal after either Topo I (left) or II (right) inhibition was calculated relative to control, and plotted as heat maps with genes ordered by decreasing gene expression in control samples. (B) The overlap between genes in Group 2 (see Figure 3) from Topo I and from Topo II samples were determined and displayed as a Venn diagram using BioVenn (Hulsen et al., 2008). (C) Gene Ontology term analysis was performed on Group 2 genes from Topo I (top) and Topo II (bottom) samples as described (Carmona-Saez et al., 2007) with the number in each GO term listed in parentheses.

Topoisomerase inhibition increases nucleosome turnover within gene bodies

The altered chromatin solubility following topoisomerase inhibition raises the question of how torsional stress affects nucleosome dynamics. Previous studies have shown that positive supercoils inhibit nucleosome assembly *in vitro* (Gupta et al., 2009) and alter chromatin accessibility to nucleases *in vivo* (Lee and Garrard, 1991), leading the authors to propose that positive supercoils disrupt the nucleosomal template within gene bodies. To test this *in vivo*, we inhibited topoisomerases individually and measured nucleosome turnover by CATCH-IT (Deal et al., 2010). After methionine depletion, S2 cells were labeled with the analog Azidohomoalanine in the presence of either Camptothecin or ICRF-193. Both labeling and topoisomerase inhibition were performed for 15 minutes at room temperature, and CATCH-IT was performed as previously described (Deal et al., 2010; Teves and Henikoff, 2011), using paired-end sequencing. Inhibition of Topo I or Topo II led to a decrease in +1 nucleosome turnover but to an increase within gene bodies (Figure 4-7 A), suggesting that nucleosome turnover depends in part on the torsional strain on DNA. Topo I inhibition increased the overall nucleosome turnover within gene bodies but the effect for transcribed genes was greater than for non-transcribed genes (Figure 4-7 B, left). In contrast, Topo II inhibition led to increased nucleosome turnover within gene bodies regardless of expression levels (Figure 4-7 B, right). This implies that Pol II-generated torsional stress increases nucleosome dynamics independent of changes in expression levels.

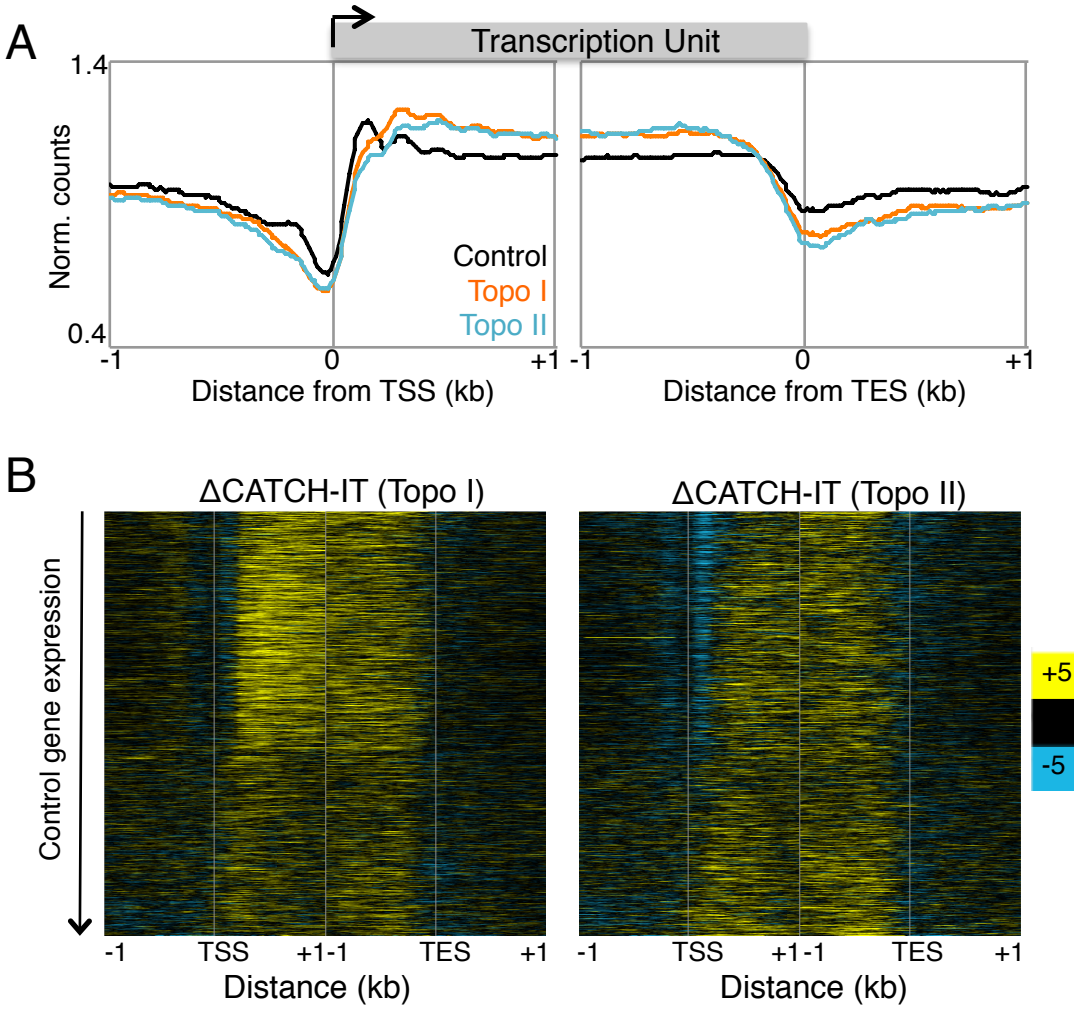

Figure 4-7: Altered nucleosome turnover under torsional stress
(A) Control and treated samples were subjected to CATCH-IT followed by paired end sequencing. Paired-end read fragments of length greater than 120 bp were mapped onto the genome and signals surrounding the TSS and TES were averaged for all genes. (B) The changes in CATCH-IT values were determined by subtracting the control from Topo I- (left) and Topo II-inhibited (right) signals, and were displayed as heatmaps with genes ordered by decreasing expression in control samples.

Topoisomerase inhibition increases Pol II stalling

To study the effects of topoisomerase inhibition on Pol II genome-wide, we measured Pol II levels both before and after topoisomerase inhibition. We profiled the low salt soluble Pol II, which we have shown previously to be highly enriched for the stalled species (Teves and Henikoff, 2011; Weber et al., 2010). Notably, inhibition of either Topo I or Topo II led to overall increase in stalled Pol II just downstream of TSSs (Figure 4-8 A, top), suggesting that topoisomerase inhibition affects the dynamics of Pol II recruitment to, and release from, the promoter. When Pol II levels are displayed as heat maps, increases in stalled Pol II are seen primarily at active genes (Figure 4-8 B). Furthermore, Topo I inhibition resulted in increased Pol II levels within gene bodies compared to control, in contrast to Topo II inhibition (Figure 4-8 A, bottom), suggesting that Topo I inhibition may lead to increased Pol II progression within gene bodies. This aberrant elongation after Topo I inhibition was previously described for the human *Hif-1α* gene and the *DHFR* gene in CHO cells (Baranello et al., 2009; Ljungman, 1996).

We then determined if the increased stalled Pol II at the TSS corresponds to changes in transcriptional activity by measuring the nascent RNA levels before and after topoisomerase inhibition using a biochemical fractionation assay that enriches for chromatin bound nascent RNA (Wysocka et al., 2001) followed by high-throughput sequencing. In parallel, we performed nascent RNA analysis for S2 cells heat shocked at 37°C for 15 minutes to confirm the validity of the assay, as heat shock is well known to rapidly decrease expression levels genome-wide (Jamrich et al., 1977). When we averaged the normalized counts for all genes surrounding the TSS and TES for control, Topo I, and Topo II inhibited samples, we found that Topo I inhibition resulted in overall increased nascent RNA production near the 5' end (Figure 4-8 C), as expected from the increase in Pol II within gene body (Figure 4-8 A, bottom), whereas Topo II inhibition

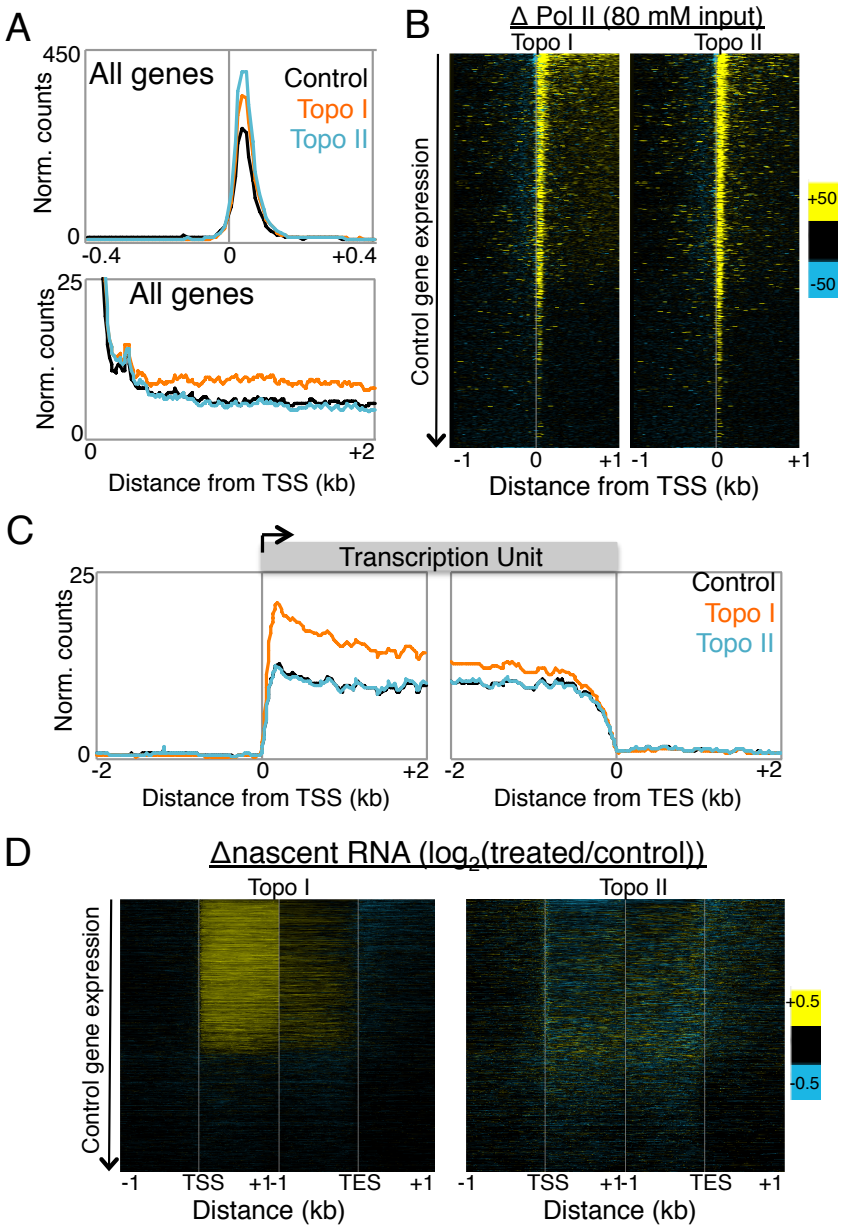

Figure 4-8: Transcriptional effects of topoisomerase inhibition
(A) Pol II ChIP-seq was performed on control and treated cells using the modified library preparation protocol (Henikoff et al., 2011). Reads of length 25 – 75 bp were parsed out and mapped onto the genome, and the average signal 800 bp surrounding the TSS of all genes were determined (top). The same data are shown for the region 2 kb downstream of the TSS (bottom). (B) Changes in Pol II signals were determined by subtracting control signals from Topo I- (left), and Topo II-inhibited (right) Pol II ChIP, and values are presented in heatmap format with genes arranged by decreasing gene expression in control samples. (C) Nascent RNA-seq was performed on control, Topo I-, and Topo II-inhibited cells, and average profiles were determined for 4 kb surrounding the TSS and TES of all genes. (D) The log-ratio of nascent RNA for Topo I- (left) or II-inhibited (right) samples over control was determined and shown in heatmap format with genes ordered by decreasing expression in control samples.

showed no overall change relative to control (Figure 4-8 C). To view gene-by-gene trends, we displayed these changes as heat maps surrounding the TSS and TES. We found that Topo I inhibition affects primarily the transcribed genes, whereas Topo II inhibition results in heterogeneous changes in nascent RNA levels (Figure 4-8 D). Interestingly, the gain in nascent RNA for Topo I inhibition is most prominent within 1 kb of the TSS, whereas the TES shows

little change in expression, consistent with the previously described effect of Topo I inhibition on the *DHFR* gene in CHO cells (Ljungman, 1996). Topo II inhibition primarily down-regulates the most highly expressed genes (Figure 4-8 D), suggesting that for most genes, Topo II is not necessary for expression, but that the highest expressing genes require both Topo I and II. As an assay control, we measured the change in nascent RNA levels after heat shock and found decreased levels genome-wide (Figure 4-9 A, top), consistent with the well known heat shock paradigm of globally reduced expression (Jamrich et al., 1977), as well as increased expression of the highly induced *Hsp* genes (Figure 4-9 A, bottom). As further validation, we repeated the biochemical fractionation assay to enrich for nascent RNA and performed qPCR on select genes that showed increased, decreased, and no change in nascent RNA levels upon Topo I or II inhibition. We found reproducible effects using qPCR as a separate readout (Figure 4-9 C).

From the combined Pol II ChIP and nascent RNA data, we infer that Topo I inhibition leads to aberrant Pol II release into elongation, perhaps only stalling again when supercoils accumulate to prevent further elongation near gene ends. In support of this possibility, long genes show decreased nascent RNA levels at the 3' end compared to short genes after Topo I inhibition (Figure 4-9 B). This suggests that the aberrant release of Pol II after Topo I inhibition eventually stops when torsional stress accumulates.

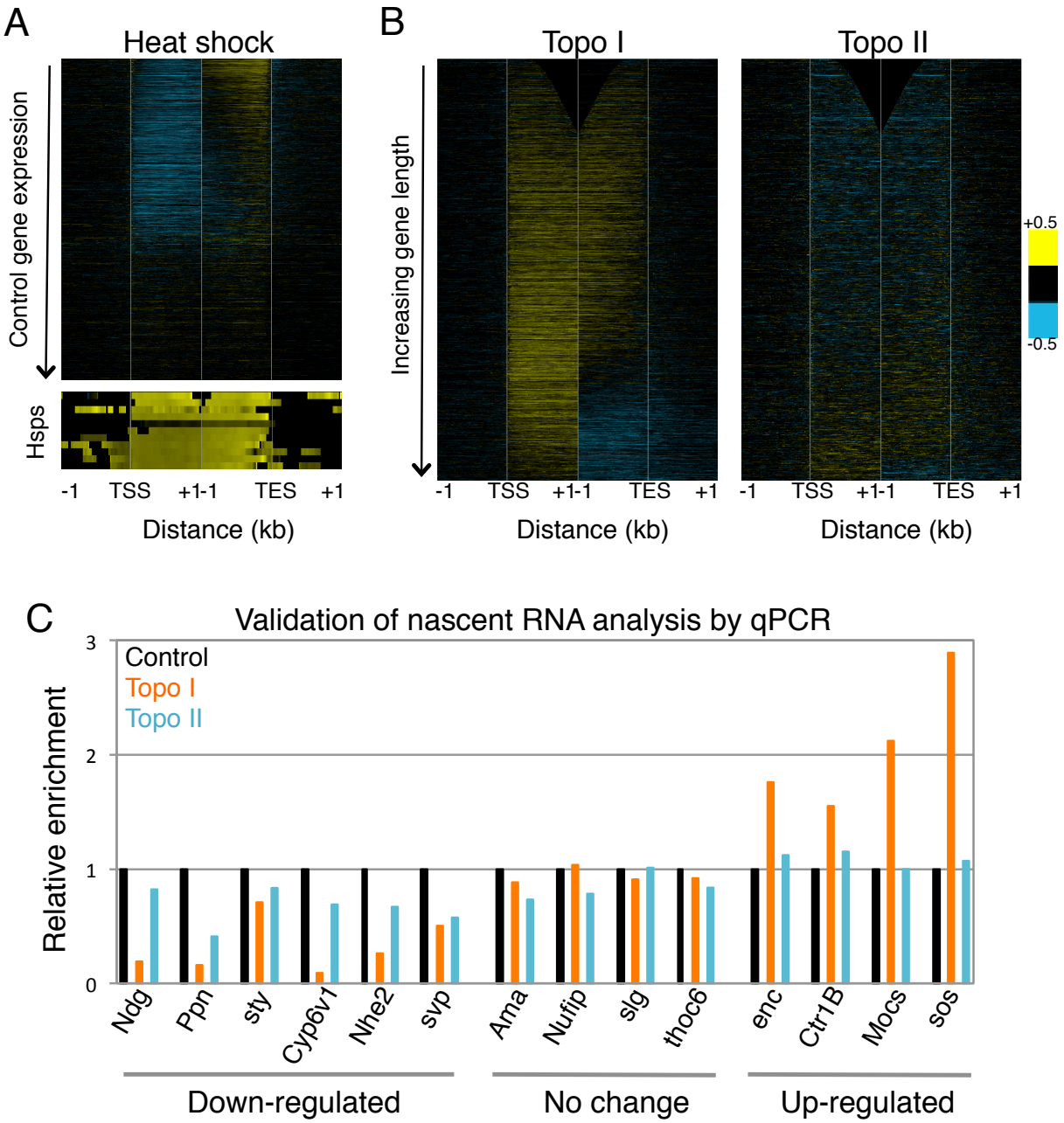

Figure 4-9: Validation of nascent RNA-seq. (A) Heat shocked samples were processed for nascent RNA and sequencing in parallel with untreated control and topoisomerase inhibited samples. The log ratio of heat shocked over control was calculated and displayed as heat map for regions 2 kb surrounding the TSS and TES with all genes arranged in decreasing expression of control samples (top). As expected, the nascent RNA levels decreased for most transcribed genes following 15 minutes of heat shock. Also as expected, the highly induced Hsp genes showed increased expression (bottom). (B) Heat maps are displayed for Topo I- (left) and Topo II-inhibited (right) samples as in Figure 5D but with genes arranged by increasing gene length. The 3' end of long genes experience a decrease in nascent RNA after Topo I inhibition. (C). Based on the sequencing results, 3 groups of genes were selected: down-regulated, no change, and up-regulated after Topo I inhibition. Primers were designed for each of the genes and were used to perform qPCR analysis on replicate nascent RNA preparation. qPCR results show similar trends with sequencing data.

Relationship between supercoiling, nucleosome turnover, and transcription

That topoisomerase inhibition resulted in increased nucleosome turnover regardless of change in nascent RNA levels implies that Pol II passage and torsional stress have separable effects on nucleosome dynamics. To test this possibility, we measured the relative contributions of Pol II elongation (nascent RNA production) and supercoiling to nucleosome turnover. We quantified the correlation between the nascent RNA levels and nucleosome turnover by measuring the correlation for every 10 bp in the first 500 bp downstream of the TSS between the two data sets for control, Topo I- and Topo II-inhibited samples and plotted the cumulative distribution function (CDF) (Figure 4-10). The CDF shows that the correlation between nascent chain production and nucleosome turnover after Topo I inhibition is almost superimposable over

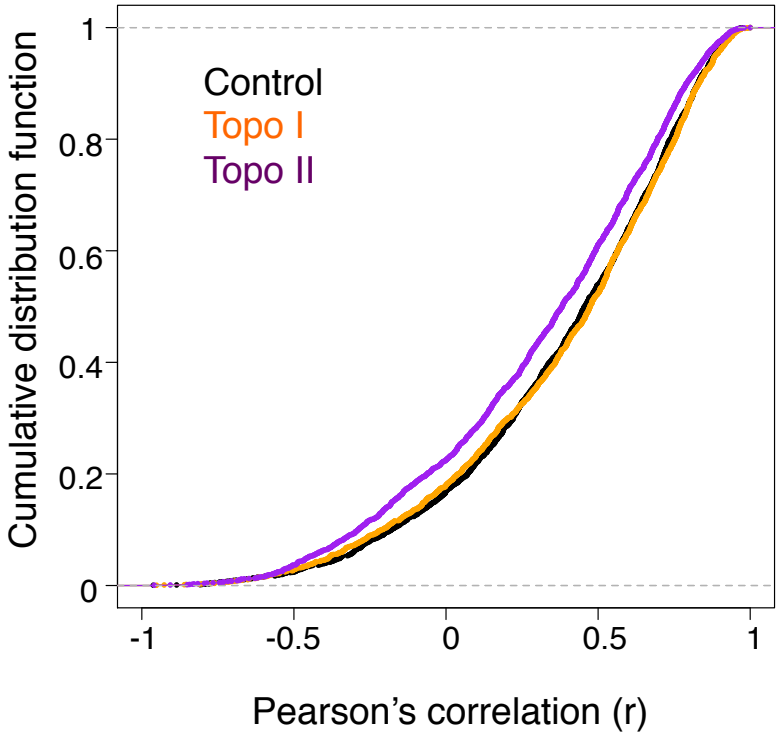

Figure 4-10: Correlation between nucleosome turnover and nascent RNA data. For each gene and for each sample, the Pearson's correlation (r) was determined for the 1 kb region surrounding the TSS between nucleosome turnover and nascent RNA-seq data, and the cumulative distribution function of the correlation was plotted.

that of the control. However, Topo II inhibition led to a decreased correlation between transcription and turnover. This implies that the nucleosome turnover within gene bodies is a cumulative effect of both Pol II-generated supercoils and the elongation passage.

To gain further insight into the relationship between torsional stress, nucleosome turnover, and transcription, we calculated a supercoiling value for each gene by averaging the TMP-seq values for 1 kb upstream of the TSS and subtracting the average TMP-seq value for 1 kb downstream of the TSS. We then calculated the supercoiling difference (SD) by subtracting supercoiling values of control samples from the Topo I- or Topo II-inhibited samples and rank-ordered the genes by the absolute value of SD for either Topo I or Topo II inhibition. We then grouped the genes by quintiles such that the 1st quintile contains genes that have the most SD. For each quintile, we averaged the relative change in stalled Pol II and nascent RNA before and after topoisomerase inhibition. During Topo I inhibition, we found that, while all quintiles showed increased stalled Pol II near the TSS, there was a negative correlation between the relative change of stalled Pol II and SD, with the exception of the 5th (lowest) quintile which experienced the least change in supercoiling (Figure 4-11 A, left). Given the increased nascent RNA after Topo I inhibition (Figure 4-8 C), it is likely that genes with higher SD have more Pol II released into elongation, and thus less stalled Pol II at the TSS. Indeed, as SD increases, there is a higher gain in nascent RNA (Figure 4-11 B, left). In contrast, Topo II inhibition results in high gain of stalled Pol II regardless of SD, again with the exception of the 5th quintile (Figure 4-11 A, right). This is accompanied by a subtle decrease in nascent RNA for the top 4 quintiles (Figure 4-11 B, right), suggesting that Topo II inhibition leads to less Pol II release into productive elongation.

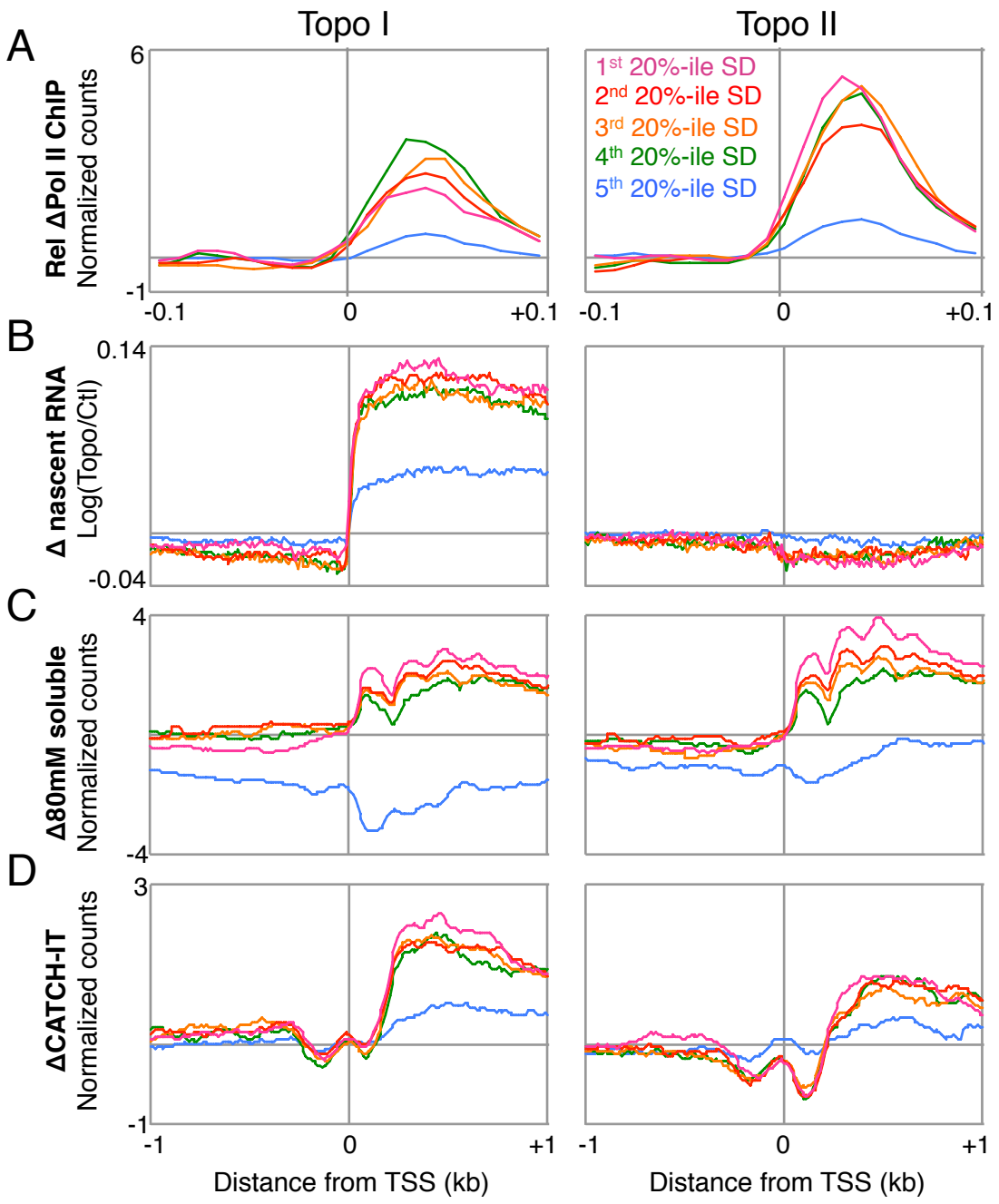

Figure 4-11: Torsional stress affects transcription and nucleosome turnover
Genes were ranked by increasing change supercoiling difference (SD) and grouped into quintiles. For each quintile, the average change in Pol II ChIP (A), nascent RNA (B), low-salt soluble nucleosomes (C), and CATCH-IT (D) surrounding the TSS were plotted.

Lastly, we analyzed the relationship between supercoiling and nucleosome properties. When we examined the average change in low-salt solubility of nucleosomes as a function of SD, we found that increasing SD corresponded with increased low-salt solubility of nucleosomes (Figure 4-11 C). We then examined the relationship between supercoiling and nucleosome turnover. We found that, as in the case of nucleosome low-salt solubility, while all quintiles experienced increased nucleosome turnover relative to the whole genome, the first 4 quintiles increased in turnover relative to the 5th quintile for both Topo I and II inhibition (Figure 4-11 D). These results suggest that once a threshold is reached in torsional stress, nucleosomes become similarly destabilized. This threshold effect of supercoiling on nucleosome salt-solubility and turnover is in contrast to the linear correlation between gene expression and turnover, further implying that supercoiling and Pol II passage have distinct contributions to genic nucleosome properties.

Discussion

In this study, we have examined the effect of Pol II-generated torsional stress *in vivo*. We have developed a genome-wide sequencing based assay to determine DNA supercoiling states genome-wide at single-base resolution in *Drosophila* cells. Using this assay, we have shown that inhibition of topoisomerases leads to accumulation of torsional stress, which is accompanied by changes chromatin properties. Increased torsional stress results in increased low-salt soluble nucleosomes and increased nucleosome turnover. Furthermore, stalled Pol II accumulates immediately downstream of the TSS after inhibition of either Topo I or Topo II, but nascent RNA production is affected differently. Topo I inhibition results in increased nascent RNA levels near the 5' end of genes. In contrast, Topo II inhibition only affects nascent RNA levels of

highly expressed genes, while most genes remain unaltered. We have shown that, despite the differences in transcriptional effects of topoisomerase inhibition, the accumulated torsional stress results in increased nucleosome turnover within gene bodies genome-wide, providing direct evidence for an *in vivo* influence of DNA supercoiling on nucleosome dynamics. Our data support a model whereby the transient wave of positive supercoils downstream of Pol II destabilizes genic nucleosomes to allow progression, and the transient negative supercoils stabilize nucleosome formation behind Pol II to maintain chromatin structure (Lee and Garrard, 1991). In this way, a delicate balance between nucleosomal destabilization, maintenance, and Pol II progression is achieved (Figure 4-12).

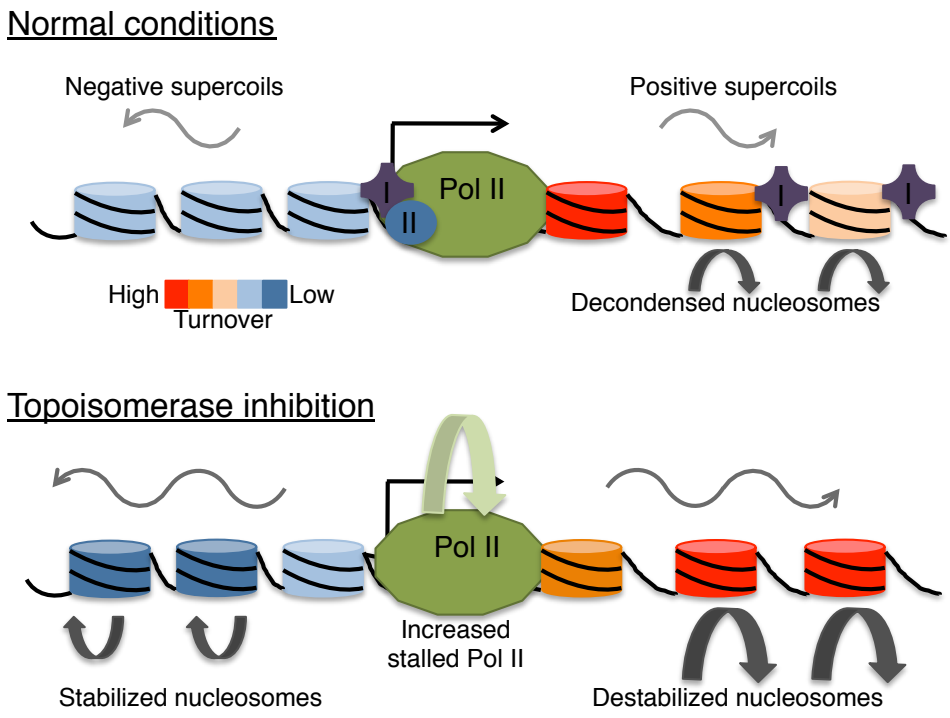

Figure 4-12: Model for transcription-generated torsional stress and nucleosome turnover
Under normal conditions, RNA Polymerase II (Pol II) generates positive supercoils ahead, and negative supercoils behind, as it elongates a transcript along a gene. Topoisomerases I and II prevent the accumulation of torsional stress. Transcription results in nucleosome turnover and decondensed organization (dark gray arrows) of transcribed nucleosomes. Upon topoisomerase inhibition, torsional stress accumulates, resulting in increased Pol II stalling and nucleosome destabilization, both by turnover and low-salt solubility, within gene bodies. This supports a balance between destabilization of nucleosomes for Pol II passage and maintenance of chromatin structure for chromosomal integrity.

Recent studies have used TMP to profile supercoiling states genome-wide in yeast (Bermúdez et al., 2010), and in a portion of the human genome (Kouzine et al., 2013; Naughton et al., 2013). Although supercoiling in yeast was determined genome-wide, the resolution was limited to > 2 kb, which is larger than the average gene length (Bermúdez et al., 2010). Perhaps as a consequence, no association between TMP binding and transcription was detected (Bermúdez et al., 2010). In contrast, studies performed in human cells showed that transcription is a major determinant of supercoiling states (Kouzine et al., 2013; Naughton et al., 2013). Using a new strategy for sequencing TMP-crosslinked fragments, we have shown that, consistent with the twin-supercoiled domain model, promoters and regions upstream of the TSS have high levels of TMP crosslinking, indicative of high negative supercoiling, while gene bodies have low levels of TMP crosslinking, as would be expected when positive supercoils inhibit TMP intercalation. Furthermore, we showed that the level of supercoiling is correlated with expression levels in S2 cells, consistent with previously published data for whole flies (Kobayashi et al., 1998; Matsumoto and Hirose, 2004). However, the average TMP-seq profile in *Drosophila* cells showed a different pattern than in human cells (Kouzine et al., 2013), where TMP-crosslinking peaked at the TSS and diminished beyond 1 kb upstream, leading the authors to conclude that supercoiling is a short-range force. In contrast, we observed high levels of negative supercoiling beyond 2 kb upstream of the TSS. The difference might be due to the prevalent bi-directional transcription seen in mammalian genomes (Core et al., 2008; Seila et al., 2008) that is absent in *Drosophila*. Bi-directional transcription would concentrate negative supercoils behind the Pol IIs at the TSS, while the uni-directional transcription of *Drosophila* promoters would allow for dissipation further upstream. Interestingly, the 3' end of *Drosophila* genes experienced high levels of negative supercoiling, whereas gene bodies upstream of the TES showed low levels of

TMP-crosslinking, suggestive of positive supercoiling (Figure 4-2 C). This 3' end effect was also exacerbated by topoisomerase inhibition (Figure 4-4). In yeast, topoisomerases are required for both transcription termination and the formation of the nucleosome-depleted region at the transcription end sites (Durand-Dubief et al., 2011; Durand-Dubief et al., 2010). Similarly, *Drosophila Hsp70* genes have Topo II both at the 5' and 3' end (Rowe et al., 1986), suggesting that at least with the *Hsp70* gene, events at the 3' end generate sufficient torsional stress to require topoisomerase activity.

It has been generally accepted that inhibition of topoisomerases results in an immediate halt in transcription due to accumulation of torsional stress. Early studies have shown that transcription of rDNA requires topoisomerase activity (Brill et al., 1987; Muller et al., 1985), and *in vitro* transcription assays using viral RNA Polymerases are greatly inhibited when topoisomerase I activity is absent (Bendixen et al., 1990). We have shown that Topo I inhibition results in increased nascent RNA near the 5' TSS for most transcribed genes (Figure 4-5), which seems contradictory to the dogma that topoisomerase inhibition leads to transcription inhibition. However, various other studies inhibiting only Topo I have yielded mixed effects on transcription (Collins et al., 2001; Kretzschmar et al., 1993; Merino et al., 1993). Topo I inhibition has been shown to increase Pol II and nascent RNA near the TSS but not at the 3' end of the *DHFR* gene in cultured CHO cells (Ljungman, 1996), and to increase the release of paused Pol II and anti-sense transcription of *Hif-1* gene in human cells (Baranello et al., 2009). It is possible that the differences in the effect of Topo I inhibition on transcription is dose-dependent. Under high concentrations of Camptothecin, nuclear RNA from HeLa cells is greatly diminished, whereas low concentrations lead to an increased level (Oravec et al., 1972). Therefore, the increase in nascent RNA that we detected may be a response to the low dose of Camptothecin

that we used, and that higher dose would eventually result in decreased expression. Consistent with this interpretation, the 3' end of genes showed much less change in nascent RNA levels after Topo I inhibition. In contrast, Topo II inhibition resulted in decreased nascent RNA levels of only the most highly expressed genes, confirming that Topo I is the main relaxer of transcription-generated supercoils and Topo II acts together with Topo I in highly expressed genes (Durand-Dubief et al., 2010; Sperling et al., 2011)

It is noteworthy that the torsion-induced increase in nucleosome turnover is observed only beyond the first nucleosome downstream of the TSS (+1 nucleosome). In fact, relative to the genome-wide average, the +1 nucleosome showed decreased turnover upon topoisomerase inhibition. The +1 nucleosome has the highest density of stalled Pol II (Muse et al., 2007), suggesting a specialized interaction between Pol II and the +1 nucleosome. Indeed, MNase-protected fragments spanning the +1 nucleosome and the footprint of Pol II are enriched immediately downstream of the TSS (Teves and Henikoff, 2011). It is possible, therefore, that under topoisomerase inhibition, the interaction between Pol II and the +1 nucleosome is further stabilized. This interpretation is consistent with the increased stalled Pol II that we saw at the TSS and the decreased turnover at the +1 nucleosome following topoisomerase inhibition. But how are the effects of torsional stress generated at the TSS propagated beyond the +1 nucleosome? Mathematical modeling of transcription-generated torsion suggests that transcription of 5 bp is sufficient to propagate a wave of positive supercoils at a rate of 2 orders of magnitude faster than Pol II elongation (Becavin et al., 2010). Single molecule experiments on supercoil dynamics have shown that this propagation has two modes: the slower mechanism of diffusion that occurs in short range distances, and the much faster mode of "hopping" where supercoils are propagated in long distances under millisecond time frames (van Loenhout et al.,

2012). Therefore, it is possible that the effect of increased turnover downstream of the +1 nucleosome is a result of torsional stress being propagated in long-range distances through the "hopping" mechanism. In this way, topological changes to DNA during transcription can act as the medium that connects events at the TSS to ones far downstream within gene bodies.

CHAPTER 5

CONCLUSIONS AND PERSPECTIVES

Using new tools to directly study nucleosome turnover, the dissociation of old and subsequent reassembly of new histones, we established the dynamic properties of chromatin in *Drosophila* cells under heat shock. Nucleosomes of activated genes dramatically increased in turnover, while those of repressed genes exhibited decreased turnover, suggesting that the act of transcription causes nucleosome turnover. To test the causality in a different way, we inhibited transcription in these cells using a drug and found a highly similar decrease in turnover genome-wide as during heat shock. Aside from its implications in transcription process, this causality challenges the role of histone modifications in regulating gene expression, as after each turnover event, modifications must be re-established. Our study was also unique in analyzing heat shock-mediated genome-wide repression, providing a mechanistic model for this rapid and reversible process. This included the loss of Pol II from chromatin, which decreased nucleosome turnover within gene bodies, while transcription factors remained promoter-bound, presumably to maintain promoter regions in an active conformation for efficient recovery.

Our discovery that transcription is sufficient to cause nucleosome turnover raised the old question of how Pol II transcribes through a nucleosome. Nucleosome remodelers and elongation factors aid in Pol II progression, but the effect of Pol II-generated torsional stress is largely overlooked. As Pol II translocates, it generates positive (over-twisting) and negative (under-

twisting) supercoils ahead and behind, respectively. *In vitro,* positive supercoiling is inhibitory to nucleosome assembly while negative supercoiling favors nucleosome formation, suggesting a mechanism for the coupling of nucleosome eviction ahead of Pol II with deposition behind to aid progression while maintaining chromatin structure. To test the role of supercoiling in nucleosome dynamics and Pol II progression *in vivo*, we inhibited topoisomerases, enzymes that relax supercoils, and measured the change in nucleosome turnover and Pol II localization in *Drosophila* cells. We found that nucleosome turnover within gene bodies increased significantly, despite the accumulation of Pol II at the TSS, suggesting that DNA supercoils mediate the transcription-dependent nucleosome turnover. By developing a new method to identify supercoiling states genome-wide, we found a net increase in DNA supercoiling, both positive and negative, during topoisomerase inhibition, consistent with the twin-supercoil domain model. Many have speculated on the importance of DNA supercoiling in transcription but this work presents *in vivo* evidence for its active role in coupling Pol II progression and nucleosome dynamics.

From work presented in this thesis dissertation, we conclude that a dynamic interplay between RNA Polymerase II and the nucleosome exists that is partly mediated by the DNA itself. To frame these findings in an evolutionary context, we discuss below examples of gene regulation, DNA topology, and nucleosome dynamics from bacteria to humans.

Dynamic gene regulation

We have used the heat shock response as a system to induce changes in gene expression in *Drosophila* S2 cells, but this evolutionarily conserved stress response is found in all life forms tested thus far, from bacteria, to Archaeal species, to humans. In *E. coli*, the response is mediated

by the transcription factor σ^{32}, which is activated by stresses that induce protein misfolding (Guisbert et al., 2008; Straus et al., 1987; Taylor et al., 1984). Like its eukaryotic counter part HSF, σ^{32} binds to promoters of *Hsp* genes. However, it does so by replacing the main sigma factor σ^{70} in the RNA Polymerase holoenzyme, thus altering Pol function, both in upregulating the *Hsp* genes, and downregulating the rest of the genome (Blaszczak et al., 1995). As we have seen in Chapter 3, eukaryotes have maintained a versatile Pol II regulation to allow for rapid induction and repression. Furthermore, in higher eukaryotes, the disruption of Pol II binding at promoters of non- *Hsp* genes during heat shock has been co-opted by non-coding RNA instead (Allen et al., 2004; Yakovchuk et al., 2009), though the presence of such non-coding RNA in *Drosophila* has yet to be identified.

Another process that depends on dynamic gene regulation is the circadian rhythm, which is the biological cycle of roughly 24 hour period. In mammals, the circadian rhythm is vital for physiology and behavior, governing important processes such as sleep/wake cycles, feeding, body temperature, hormone secretion, and metabolism (Feng and Lazar, 2012) (Figure 5-1). At

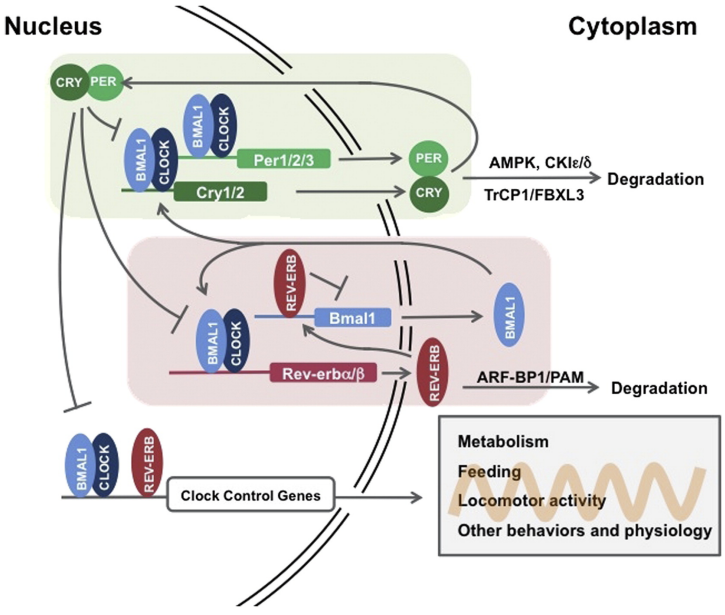

Figure 5-1. The Basic Clock Machinery Consists of Negative Transcriptional-Translational Feedback Loops. In the first loop, BMAL1/CLOCK drives Per/Cry transcription, while PER/CRY binds and inhibits transcriptional activity of BMAL1/CLOCK. In the second loop, BMAL1/CLOCK drives REV-ERB expression, which in turn represses Bmal1 transcription. Both loops are essential for maintaining circadian rhythm. Posttranslational modification, as shown for PER/CRY and REV-ERB, is also important in regulating clock activity. The core clock machinery can drive rhythmic behavioral and physiological activities, such as metabolism. Adapted and reprinted with permission from (Feng and Lazar, 2012) under license number 3172020822412.

the cellular level, the circadian rhythm is primarily controlled by two negative feedback loops. First, BMAL1/CLOCK activates expression of PER/CRY, which in turn binds and inhibits the transcriptional activity of BMAL1/CLOCK. Second, BMAL1/CLOCK activates REV-ERB expression, which is a repressor for BMAL1/CLOCK expression (Feng and Lazar, 2012). Over the last decade, research on the circadian transcription network has revealed an extensive role for chromatin remodeling in regulating changes in gene expression (Aguilar-Arnal and Sassone-Corsi, 2013) (Figure 5-2). The BMAL1/CLOCK complex recruits histone modifying complexes to promoters of target genes. In fact, CLOCK itself is a histone acetyltransferase (HAT)(Doi et al., 2006; Etchegaray et al., 2003; Koike et al., 2012). Furthermore, polycomb proteins associated with CRY are required for gene repression (DiTacchio et al., 2011; Etchegaray et al., 2006; Jones et al., 2010). This results in cyclical changes in histone modifications at target genes, which contributes to overall nucleosome dynamics and chromatin organization. Furthermore, the important role of chromatin in circadian transcription network suggests that chromatin participates in rapid modulation of expression, quite similar to the heat shock response.

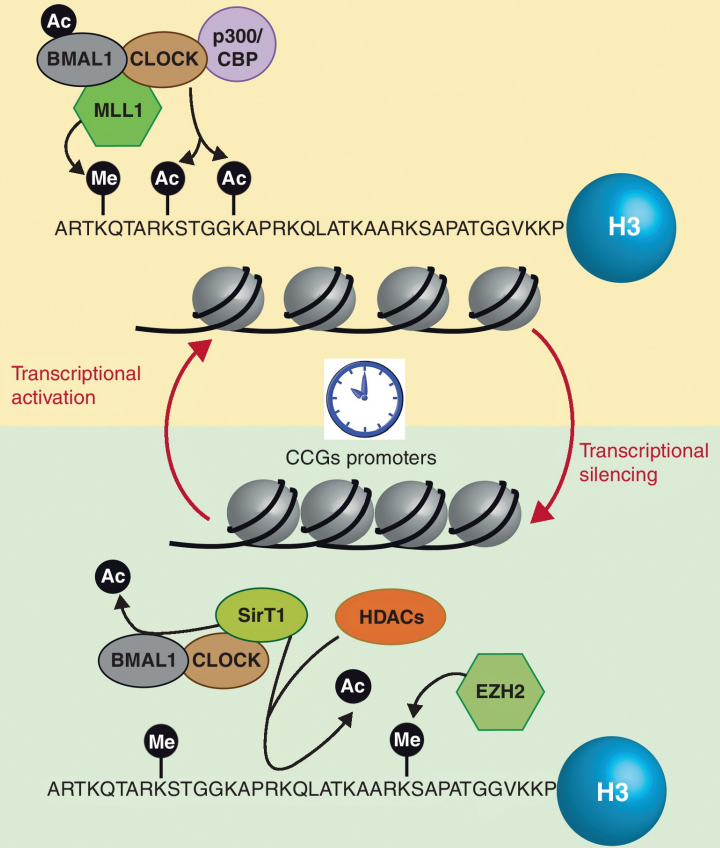

Figure 5-2. Chromatin remodeling in the circadian clock. Chromatin modifying enzymes act in synchrony for fine-tuning clock-controlled gene expression. Transcriptional activators coordinate rhythmic histone hyperacetylation and H3K4 trimethylation at circadian gene promoters, thereby inducing transcription. Conversely, repressors remove acetylation marks and promote a closed state of the chromatin fiber at the clock controlled gene promoters to inhibit transcription. Thus, activator and repressor enzymes act in a very precise synchrony leading to the circadian transcription of about 10–15% of all transcripts. Adapted and reprinted with permission from (Aguilar-Arnal and Sassone-Corsi, 2013) under license number 3172020979898.

DNA topology and nucleosomes throughout evolution

DNA supercoiling, and its influence on gene expression, has been extensively studied in bacterial species (Jeong et al., 2004; Peter et al., 2004). Indeed, a shift in DNA superhelicity, from negatively supercoiled to relaxed, accompanies gene expression changes observed during heat shock in *E. coli* (Lopez-Garcia and Forterre, 2000), suggesting that the link between DNA topology and gene expression predates nucleosome structure. Interestingly, all thermophilic archaeal species have positively supercoiled DNA, and share an archaea-specific topoisomerase, reverse gyrase, that specifically introduces positive supercoils in the genome (Rodriguez and Stock, 2002). Positive supercoiling increases the denaturation temperature of DNA and is believed to be thermoprotective (Kampmann and Stock, 2004). Indeed, when the thermophile

Solfolobus islandicus is heat shocked, the genome becomes even more positively supercoiled (Lopez-Garcia and Forterre, 2000). The direction of change in DNA helicity is similar to that seen in *E. coli* (Lopez-Garcia and Forterre, 2000), suggesting a conserved function of supercoiling in DNA stability. Related to this, others have argued that the shift in temperature alters DNA topology directly (Lopez-Garcia and Forterre, 2000), but given our findings in this thesis research, we suggest that transcription of *Hsp* genes conserved in prokaryotes and archaea could generate sufficient torsional stress and contribute to alter the topological state of the genome.

Bacteria do not form nucleosomes. However, the discovery of a nucleosome-like structure in archaea (Ammar et al., 2012; Pereira et al., 1997) may provide an evolutionary perspective in the relationship between DNA topology and nucleosome dynamics. In archaea, homologs of histones H3 and H4 form tetramers protect about 60 bp of DNA (Pereira et al., 1997), and are organized in the genome similarly to eukaryotes such that the nucleosome-depleted transcription start sites are flanked by well-positioned nucleosomes (Ammar et al., 2012) (Figure 5-3). These archaeal tetramers wrap DNA around once, and have been shown to inhibit basal transcription *in vitro* (Wilkinson et al., 2010). However, these tetramers do not wrap DNA as tightly as eukaryotic nucleosomes do, requiring no chromatin remodelers for transcription (Wilkinson et al., 2010). This observation suggests that nucleosomes have evolved to be more stable, resulting in increased regulatory functions. Interestingly, the archaeal nucleosomes have been shown to constrain both negative and positive supercoils *in vitro* (Musgrave et al., 2000; Musgrave et al., 1991; Reeve et al., 1997; Ronimus and Musgrave, 1996a, b), but it is unclear how the archael nucleosomes affect DNA topology *in vivo*. Although little is yet known about these archaeal nucleosomes, their similarities in structure and organization to

eukaryotic nucleosomes lead to speculations that changes in DNA topology during transcription may also affect archaeal nucleosome stability.

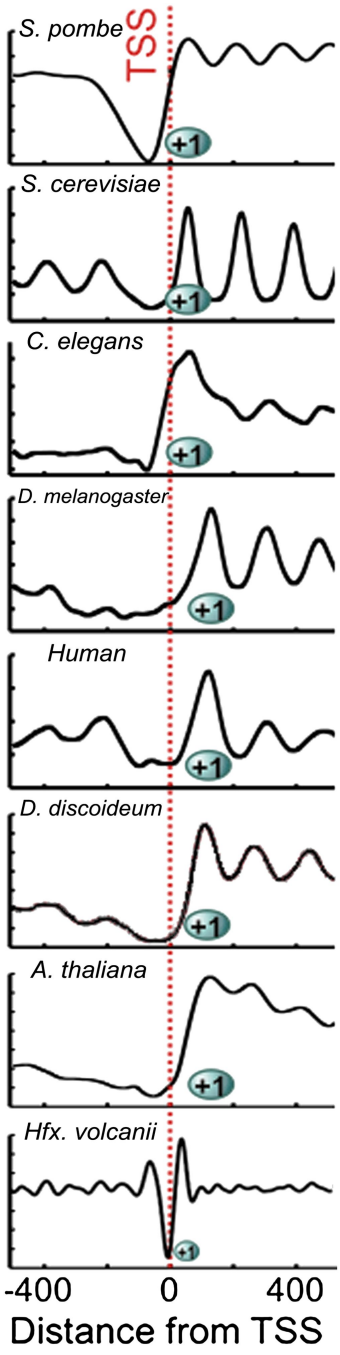

Figure 5-3.
Chromatin architecture is conserved at the 5′ end of transcripts across eukaryotes and archaea.
Due to the smaller size of archaeal nucleosome DNA, the occupancy has a shorter periodicity. Figure adapted with permission from Chang et al. (2012). DOI: http://dx.doi.org/10.7554/eLife.00078.006

CHAPTER 6

FUTURE DIRECTIONS

The research work presented in this dissertation has addressed some of the fundamental questions in chromatin and transcription biology, but also raises several avenues for future research.

Heat shock biology

We have shown that Pol II dissociates from the promoter regions of most genes during heat shock (Chapter 3) and that this is associated with an immediate decrease in transcription (Chapter 4). As we discussed, Pol II dissociation appears to be a well conserved aspect of the heat shock response, from bacteria to mammals. In *E. coli*, the replacement of the main sigma factor σ^{70} with the heat shock specific σ^{32} alters the conformation of RNA Pol to bind to *Hsp* promoters to the exclusion of the rest of the genome (Blaszczak et al., 1995). In mammals, heat shock-induced non-coding RNAs disrupt the interaction between Pol II and promoter DNA in the open complex, resulting in dissociation of Pol II from the DNA (Allen et al., 2004; Yakovchuk et al., 2009). What causes the Pol II dissociation that we observe in *Drosophila* cells upon heat shock? What roles do non-coding RNAs play in this process? A well-known heat shock induced non-coding RNA in *Drosophila* is the heat shock RNA omega gene (hsrω) (Lakhotia, 2011). This expressed gene becomes highly activated upon heat shock, and the resulting non-coding RNA is believed to act as a storage factor for RNA processing proteins,

limiting protein production during heat stress (Lakhotia, 2011, 2012). Could this transcript interact with Pol II itself? What other non-coding RNAs are induced upon heat shock? Non-coding RNAs, and their components, are emerging as vital regulators of the transcription process, and the heat shock response may prove an attractive system in further understanding their roles in gene regulation.

Environmental effects on chromatin and transcription

The heat shock response has provided us with a simple system to study nucleosome and Pol II dynamics during gene regulation. However, temperature shift in the opposite direction also causes transcriptional perturbation, with a different set of challenges. First discovered in bacteria, the cold shock response involves the induction of cold shock protein (Csp) genes and a general suppression of transcription and translation processes (Al-Fageeh and Smales, 2006). Although this process is still relatively unknown, semblances of the bacterial cold shock response have been observed in yeast and mammalian cells (Al-Fageeh and Smales, 2006; Fujita, 1999). Using *Drosophila* S2 cells, we can extensively study the transcriptional and chromatin aspects of the cold shock response. What genes are specifically up-regulated upon cold shock? How is nucleosome dynamics affected by cold shock? Is the genome-wide repression in cold shock similar to that in heat shock? How is Pol II affected? What constitutes recovery from cold shock, and how might cells build tolerance? Using the tools presented in this thesis work, these and other questions on cold shock biology become tractable.

We can further extend our studies to other environmental perturbations. For example, light affects many biological processes because it is used to calibrate the circadian rhythm. As discussed in Chapter 5, the circadian rhythm is controlled at the cellular level by the periodic

transcription of specific genes in a roughly 24-hour cycle. The nucleosomes of circadian-related genes are also periodically modified (Aguilar-Arnal and Sassone-Corsi, 2013). How do these modifications affect the stability of nucleosomes? Would we observe a cyclical wave of nucleosome dynamics corresponding to the circadian period, such that when these genes are repressed, nucleosomes are retained whereas during activated periods, these nucleosomes turnover rapidly? Studies have shown that histone modifying enzymes play an important role in ensuring proper transcription for the circadian rhythm. Do these modifications influence nucleosome stability and dynamics? We will return to this question in detail below, but to study chromatin biology in circadian rhythm, we can apply our CATCH-IT method to the circadian rhythm mouse cell line model developed in the Sassone-Corsi lab (Nakahata et al., 2009). Furthermore, we can study Pol II dynamics in this process as well. Do these genes contain paused Polymerases? When these genes are repressed, does Pol II remain stably bound to the promoter regions? The circadian clock may prove to be an attractive model in further understanding chromatin biology in gene regulation.

Dissecting the role of histone modifications in nucleosome turnover

For the last two decades, histone modifications have been at the forefront of chromatin biology and epigenetics. Modifications associated with transcriptional activation, such as histone acetylation and methylation of lysine 4 on H3 (H3K4), and ones associated with transcriptional repression, such as histone methylation on lysine 9 and lysine 27 on H3 (H3K9 and H3K27), are catalyzed by histone modifying enzymes that are recruited to genes (Smolle and Workman, 2013). However, in the context of transcription-mediated nucleosome turnover, these modifications have to be continually added to newly incorporated histones. Furthermore,

repressive histone marks such as general methylation have been observed at regions that experience higher nucleosome turnover than the genome average, which implies that these modifications also have to be replaced onto new histones for every turnover event. These findings raise the question of precisely what the roles of histone modifications are in the context of nucleosome dynamics and gene regulation. To answer this question, we can knock down individual histone modifying enzymes and measure changes in nucleosome turnover by CATCH-IT, and changes in expression using nascent RNA-seq. By relating histone modifications with nucleosome dynamics, we can begin to dissect the functional effects of histone modifications *in vivo.*

DNA topology and Pol II dynamics

In Chapter 4, we showed that topoisomerase inhibition resulted in increased levels of low salt soluble, stalled Pol II at the TSS. The low salt soluble Pol II, however, constitutes only a fraction of total Pol II in the cell. Although we examined the levels of nascent RNA as measure of Pol II elongation, we lacked the resolution to identify locations of all Pol II as a function of torsional stress. To address this shortcoming, we can map Pol II at single nucleotide resolution during topoisomerase inhibition using a new technique called Nascent Elongating Transcript sequencing (NET-seq) {Churchman, 2011 #9}. This method follows a similar process for enriching for nascent transcripts as described in nascent RNA-seq, but allows for precise identification of the last nucleotide incorporated by Pol II and thereby mapping Pol II at single nucleotide resolution throughout the whole genome. Using this method, we can answer the following questions. Does torsion increase Pol II pausing in front of the nucleosome? Or does the destabilization of nucleosomes during torsional stress leads to a more randomized pausing of

Poll II? We can further examine the transient increase in nascent RNA following Topo I inhibition by looking at Pol II localization as a function in increasing Topo I inhibition. The NET-seq method may provide the resolution needed to study the precise dynamics of Pol II upon torsional stress.

Based on the Topo I and Topo II inhibition on Pol II and nascent RNA levels, we inferred that Topo I is the main relaxer of Pol II-generated torsional stress and that Topo II is mainly required for highly expressed genes (Chapter 4). However, the more direct approach of topoisomerase mapping throughout the genome is needed to solidify this inference. By correlating the localization of each topoisomerase with Pol II *in vivo*, we can begin to dissect the individual roles of topoisomerases in gene regulation.

Concluding remarks

As research continues to shed light on the process of transcriptional regulation, we begin to appreciate that transcription is a balancing act. Pol II, and associated factors, must balance between transcriptional fidelity to maintain identity, and plasticity to respond to a changing environment. Furthermore, chromatin must balance between packaging the DNA within the nucleus and allowing for DNA accessibility for proper transcription. The dynamic interaction between Pol II and chromatin is at the center of this balancing act. Factors that regulate this process, including transcription factors, nucleosome remodelers, histone modifying enzymes, non-coding RNA, and even DNA torsion, primarily act on maintaining this balance. As we frame these interactions in a more dynamic paradigm, we can begin to understand how transcriptional regulation impacts development, how imbalances in this process underlie disease progression, and ultimately, how shifting the balance affects the survival of the cell and even the organism.

APPENDIX I

SALT FRACTIONATION OF NUCLEOSOMES FOR GENOME-WIDE PROFILING

Modified from a chapter published in Methods in Molecular Biology

Summary

Salt fractionation of nucleosomes, a classical method for defining "active" chromatin based on nucleosome solubility, has recently been adapted for genome-scale profiling. This method has several advantages for profiling chromatin dynamics, including general applicability to cell lines and tissues, quantitative recovery of chromatin, base-pair resolution of nucleosomes, and overall simplicity both in concept and in execution. This chapter provides detailed protocols for nuclear isolation, chromatin fragmentation by micrococcal nuclease digestion, successive solubilization of chromatin fractions by addition of increasing concentrations of salt, and genome-wide analyses through microarray hybridization and next generation sequencing.

1. Introduction

Dynamic chromatin organization maintains DNA compaction while allowing for accessibility during active processes such as transcription (Henikoff, 2008). These active processes are regulated through the action of nucleosome remodeling, histone modifications and variants, and chromatin-associated proteins. A variety of methods have been developed to study chromatin dynamics. Traditional methods utilize DNA cleavage systems coupled with chromatin probing, such as DNase I hypersensitivity (Crawford et al., 2004; Hesselberth et al., 2009; Weintraub and Groudine, 1976) and micrococcal nuclease (MNase) (Lee et al., 2007; Tsankov et al., 2010) mapping assays, which respectively measure chromatin accessibility and nucleosome occupancy using nuclease digestion. Other methods rely on chromatin solubility or partitioning differences that occur after formaldehyde cross-linking (FAIRE, Sono-Seq) (Auerbach et al., 2009; Giresi et al., 2007). In addition, there are methods that measure chromatin dynamics directly by measuring nucleosome turnover, either using protein-encoded tags (Dion et al., 2007) or metabolic labeling of histones (Deal et al., 2010). Protein-related information can be obtained using either chromatin immunoprecipitation (ChIP), which relies on affinity capture of the protein of interest (Bonner et al., 1968; Gilchrist et al., 2009; Sikes et al., 2009) or DamID which relies on DNA methylation by tethered Dam methyltransferase (van Steensel et al., 2001). All of these methods for chromatin characterization have been adapted for genome-wide profiling, taking advantage of the extraordinary improvements in microarray and sequencing technologies that have occurred over the past several years.

Another traditional method for assaying chromatin is salt fractionation (Sanders, 1978). Chromatin digested with an enzyme such as MNase can be separated into soluble and insoluble fractions in the presence of physiological Mg^{2+} and low Na^{+} concentrations (Henikoff et al.,

2009). Subsequently, increasing Na^+ concentrations allows for separation of the insoluble fraction into high-salt soluble and insoluble fractions (Sanders, 1978). Low salt concentrations solubilize about 5-10% of chromatin and is composed primarily of mononucleosomes, whereas high salt solubilizes the majority of the nucleosomes (Figure I-1). Genome-wide profiling of

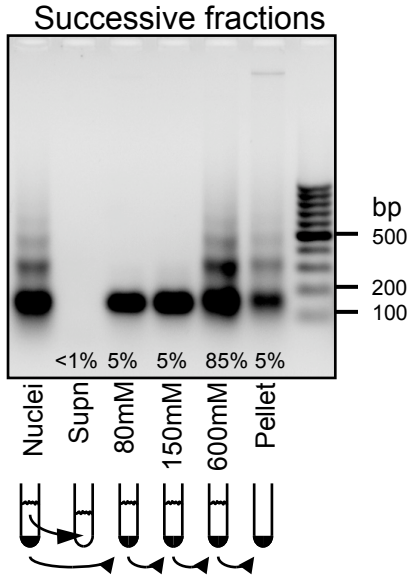

Figure I-1. Size distribution of salt fractions by agarose gel electrophoresis. Drosophila S2 cells were subjected to salt fractionation as described (Henikoff et al., 2009). DNA was extracted from each fraction and electrophoresed on a 1.5% agarose gel. Lane 1 (Nuclei) corresponds to MNase treated total chromatin. Lane 2 (Supn) corresponds to the supernatant after the MNase treated nuclei are pelleted. Lane 3, 4, 5, and 6 are the 80 mM, 150 mM, 600 mM, and insoluble pellet fractions, respectively. The percentage shown indicates the amount of chromatin solubilized in each fraction. The diagram below depicts the process of successive solubilization of chromatin with increasing amounts of salt.

low-salt soluble, high-salt soluble, and high-salt insoluble fractions versus total MNase treated nuclei reveals that salt fractionation can differentially extract chromatin based on distinct physical properties (Henikoff et al., 2009). Highly accessible chromatin is enriched in the low-salt soluble fraction while high salt solubilizes the majority of condensed chromatin, revealing insights into chromatin structure. The insoluble fraction is enriched in transcriptionally active chromatin, rendered insoluble presumably due to its association with large multi-protein complexes. Salt fractionation can therefore be used to map differences in physical properties and organization of chromatin. Furthermore, affinity capture of histones from each fraction can reveal differences in composition and modification of nucleosomes in their respective fraction.

Although originally developed for studying nucleosomes, salt fractionation has recently been used to map paused RNA polymerase (Weber et al., 2010).

Salt fractionation has several advantages over other methods for characterizing chromatin dynamics. No antibodies, transgenes or special treatments are needed, so that salt fractionation can be applied to essentially any eukaryotic cell type, whether from cell lines or tissues. By assaying the low-salt-soluble (active) fraction, the high-salt fraction, and the insoluble ("nuclear matrix") pellet, essentially 100% of native chromatin is characterized. An important advantage of salt fractionation over methods such as X-ChIP, FAIRE, Sono-Seq and DamID is that mononucleosome resolution is achieved, which allows for mapping of active chromatin at single base-pair level using massively parallel sequencing (Figure I-2). Furthermore, the simplicity of the salt fractionation process makes it an attractive method for characterizing epigenomes.

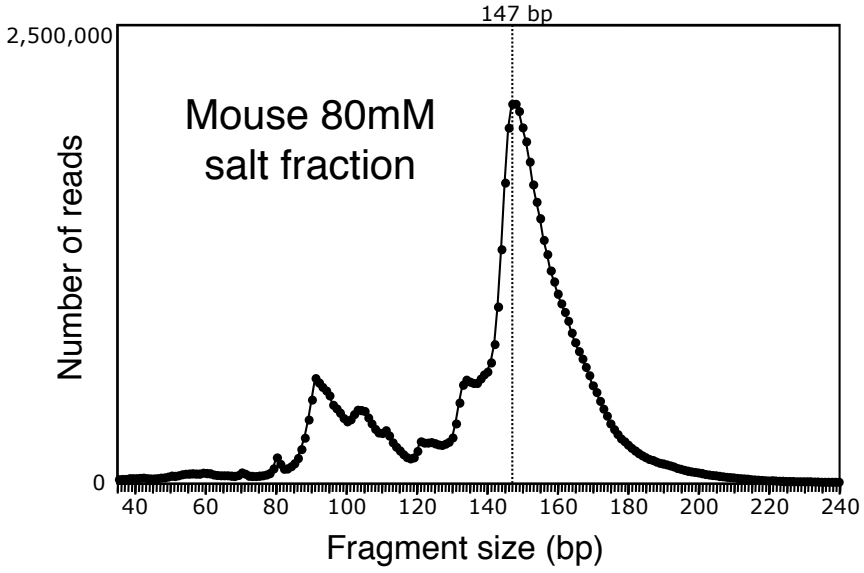

Figure I-2. Size distribution of a low-salt fraction by paired-end Solexa sequencing. Immortalized mouse pre/pro B cells PD31A were subjected to salt fractionation and a paired-end Solexa library was generated following section 3.3.2. Paired-end sequencing was performed in a single lane of an Illumina Hi-Seq 2000 instrument following the Illumina protocol. The length distribution of all 84 million mapped paired-end reads is shown at basepair resolution. Note that the dominant peak of nucleosomal DNA is centered at 147 bp, with smaller sized fragments indicative of MNase nicking within the nucleosome.

We divide the salt fractionation procedure into three stages: 1) preparation of nuclei and MNase digestion, 2) chromatin extraction and DNA isolation, and 3) preparation for genome-scale assays. The use of EGTA, instead of EDTA, retains free Mg^{2+} ions, which are critical for nuclear and chromatin integrity (Sanders, 1978), allowing for ease of nuclear isolation with mild non-ionic detergents. Protein analysis of each fraction can be performed using SDS polyacrylamide gel electrophoresis and immunoblotting. Affinity capture and subsequent DNA isolation can be used to identify changes in nucleosome composition within each fraction. Finally, DNA isolated from each fraction can be used to generate genome-wide profiles of chromatin structure and physical properties using microarray hybridization or next generation sequencing.

2. Materials

2.1. Preparation of nuclei and MNase digestion

1. Mid-log phase cultured cells
2. 14 mL polypropelene tubes
3. Phosphate-Buffered Saline (PBS):
4. Protease-inhibitor tablets, EDTA-free (Roche)
5. *TM2 buffer: 10 mM Tris pH7.4, 2 mM $MgCl_2$, 0.5 mM PMSF
6. 10% NP-40
7. 0.2 M $CaCl_2$
8. MNase 200 U (Sigma Aldrich) resuspended to 0.2 U/μL
9. 0.2 M EGTA

2.2. Chromatin extraction and DNA isolation

10. *80 mM Triton Buffer: 70 mM NaCl, 10 mM Tris 7.4, 2 mM $MgCl_2$, 2 mM EGTA, 0.1% Triton X-100, 0.5 mM PMSF
11. *150 mM Triton Buffer: 140 mM NaCl, 10 mM Tris 7.4, 2 mM $MgCl_2$, 2 mM EGTA, 0.1% Triton X-100, 0.5 mM PMSF
12. *600 mM Triton Buffer: 585 mM NaCl, 10 mM Tris 7.4, 2 mM $MgCl_2$, 2 mM EGTA, 0.1% Triton X-100, 0.5 mM PMSF
13. *TNE buffer: 10 mM Tris pH 7.4, 200 mM NaCl, 1 mM EDTA
14. 5 M NaCl
15. 0.5 M EDTA
16. RNase, DNase-free 500 μg/1 mL solution (Roche)
17. Proteinase K 20 mg/mL RNA grade (Invitrogen)
18. Phenol-Chloroform-Isoamyl alcohol (25:24:1 v/v) (Invitrogen)
19. 200 proof Ethanol
20. TE pH 8
21. Phase-lock gel tubes (Heavy 1.5 mL – 200 tubes) (5 Prime)
22. Glycogen 20 mg/mL (Roche)

2.3. Genome-scale profiling

23. 5' Cy labeled NimbleGen Validated Random 7-mer (Tri-Link)
24. 40 mM dNTPs (10 mM each) (New England Biolabs)
25. Nuclease-Free H_2O (Promega)
26. Klenow fragment 3' → 5' exo- (50U/μL) (New England Biolabs)
27. Isopropanol
28. Paired-end Sample Preparation Kit (Illumina)

29. Qiagen gel purification kit

30. Qiagen MinElute PCR purification kit

2.4 General Equipment

31. Refrigerated table-top centrifuge

32. Non-refrigerated centrifuge for 15 mL conicals (ex. IEC Centra Cl2)

33. End over end eppendorf tube rotator (ex. Labquake Shaker – Thermo Scientific)

34. PCR thermocycler

35. Vortexer

36. Waterbath at 37^{o}C

37. Nanodrop

    * Buffers are supplemented with 1X Protease inhibitor (Roche)

3. Methods

Briefly, cells are harvested and lysed with mild NP-40 to release nuclei. Washed nuclei are subjected to limited MNase digestion to fractionate the chromatin. Successive incubation with buffers containing increasing salt concentrations differentially solubilizes the chromatin. DNA can then be purified from each fraction for genome-wide analysis using microarray hybridization or massively parallel sequencing such as the Illumina platform.

3.1 Preparation of nuclei and MNase digestion

1. Grow *Drosophila* Schneider 2 (S2) cells in preferred media to exponential growth phase and 90% confluency (~2 × 10^6 cells per cm^2; *see* **Notes 1**)

2. Scrape cells off one-75 cm^2 flask and collect them in 14 mL polypropelene round bottom tubes. Pellet the cells from the media in a room-temperature table-top centrifuge (IEC Centra Cl2) for 3 minutes at 2 krpm.
3. Discard the media and wash cells in cold 1X PBS and pellet cells as above.
4. Resuspend cells in 1 mL TM2 Buffer and cool on ice for 3 minutes.
5. To lyse the cells while keeping the nuclei intact, slowly add 60 μL 10% NP-40 while gently vortexing the tube. Incubate the cells on ice for 3 minutes with 5 second gentle vortexing every minute (*see* **Notes 2**).
6. Separate the nuclei from cellular debris by gentle centrifugation (100 rcf) for 10 minutes in a refrigerated table-top centrifuge. The nuclear pellet appears as a white, loose pellet that is easily disrupted. Carefully remove the supernatant to prevent disrupting the nuclear pellet. Wash the nuclei with 1 mL of TM2 Buffer by gently pipetting the buffer into the tube. The pipetting action easily disrupts most of the nuclei pellet and the rest can be fully resuspended by gentle flicking of the tube (*see* **Notes 3**). Pellet the nuclei for 10 minutes at 100 rcf and remove the supernatant as before.
7. Resuspend nuclei in 400 μL TM2 and warm to 37°C in water bath for 5 min.
8. To fractionate the chromatin, add 2 μL of 0.2 M $CaCl_2$ to final concentration of 1 mM and 2.5 μL of MNase (final concentration of 1.25 U/mL). Return to 37°C in water bath for 10 minutes with intermittent mixing to prevent aggregation of nuclei at the bottom of the tube (*see* **Notes 4**).
9. Addition of 4 μL of 0.2 M EGTA to final concentration of 2 mM stops the MNase reaction. Remove 40 μL (10% of reaction) and label as 'Nuclei' for total MNase treated chromatin, and an additional 40 μL for protein analysis. To remove the MNase, pellet the

nuclei for 10 minutes at 100 rcf and carefully remove the supernatant and save as 'Supn' fraction. Remove 30 μL for protein analysis of the Supn fraction. Wash the nuclei carefully in 1 mL of TM2 Buffer, pellet and remove supernatant as above. Proceed with the nuclei pellet to Salt Fractionation (section 3.2).

3.2. Chromatin extraction and DNA isolation

1. From step 9 of section 3.1, resuspend the nuclear pellet in 700 μL of 80 mM Triton Buffer (*see* **Notes 5**) and incubate in constant agitation by placing the tube in a rotator (ex. Labquake Shaker) at 4^{o}C for 2 hours. This releases nucleosomes soluble in 80 mM salt concentration into the supernatant.
2. To extract the low salt soluble nucleosomes, pellet the nuclei at 0.1 rcf for 10 minutes at 4^{o}C and save the supernatant labeled as '80 mM' fraction. The loose nuclear pellet will often be slightly disrupted during this process, causing some of the nuclei to be aspirated with the supernatant. To clear the 80 mM fraction, respin the supernatant for 2 minutes at maximum centrifugation speed and transfer the cleared 80 mM fraction in a new tube. Remove 30 μL for protein analysis.
3. Optional: Resuspend the nuclei from the 80 mM salt extraction with 700 μL 150 mM Triton buffer and incubate with constant agitation at 4^{o}C for 2 hours. Extract and clear the supernatant as in step 3.2.2 to release 150 mM soluble nucleosomes and remove 30 μL for protein analysis. The 150 mM fraction consists of primarily mononucleosomes (Figure I-1) and is enriched at the 5' ends of active genes (Henikoff et al., 2009). This fraction is very similar to the 80 mM fraction but with lower resolution.
4. The low salt buffers solubilize ~5-10% of total chromatin. To solublize the majority of nucleosomes, resuspend the nuclei in 600 mM Triton Buffer and incubate at 4^{o}C in

constant agitation in Nutator for 2 hours to overnight. Extract and clear the supernatant labeled as '600 mM' fraction as in step 2 of section 3.2. Remove 30 μL for protein analysis.

5. The remaining pellet fraction corresponds to ~5 – 10% of chromatin. Resuspend the pellet in 700 μL of TNE buffer and label as 'Pellet' fraction. Hold all fractions on ice prior to DNA extraction.
6. Aliquots of each fraction can be electrophorese on an SDS gel to visualize histones and probe for the presence of specific proteins.
7. ChIP assays can be performed on each salt fraction using standard native ChIP protocols. Save an aliquot of the salt fraction for 'input' DNA and use the remainder for affinity purification.
8. For each fraction, add $1/50^{th}$ volume of 0.5 M EDTA (14 μL). For the Nuclei, Supn, 80 mM and the optional 150 mM fractions, add $1/50^{th}$ volume (14 μL) of 5 M NaCl. The 600 mM and Pellet fractions contain sufficient amounts of NaCl for DNA precipitation purposes.
9. Prepare 1:10 dilution of RNase enzyme in H_2O and add 5 μL of diluted RNase to each fraction. Allow for RNA digestion to proceed for 10 minutes in a 37^oC water bath.
10. To remove proteins, add $1/16^{th}$ volume of 10% SDS (43.75 μL) for a final concentration of 0.625%. Then add 2.5 μL of Proteinase K and incubate at 75^oC for 10 minutes.
11. Extract DNA by adding 1 volume of phenol-chloroform-isoamyl alcohol. Transfer the samples in phase-lock gel tubes and vortex for 2 minutes. Centrifuge the samples at maximum speed in a refrigerated centrifuge for 10 minutes. Transfer the supernatant into

a new phase-lock tube and repeat the extraction one more time. Transfer the aqueous solution into eppendorf tubes.

12. To precipitate the DNA, add 2 μL of Glycogen and 2.5 volumes of ice-cold 100% ethanol. Incubate on ice for 20 minutes and centrifuge at maximum speed for 15 minutes in a refrigerated centrifuge. Remove the supernatant and wash the pellet with 1 mL of ice cold 80% ethanol. Centrifuge the samples for 5 minutes at maximum speed, remove the supernatant, and allow the pellet to dry.
13. Once fully dry, resuspend the DNA with 0.1X TE pH 8. Determine the concentration of DNA using Nanodrop. Electrophorese an aliquot in a 1.5% agarose gel with ethidium bromide.

3.3. Genome-scale profiling

3.3.1 Microarray analysis

1. The following protocol is a modified version of Nimblegen labeling methods specifically adapted to *Drosophila* S2 cells. Bring 0.2 – 1 μg of DNA from each fraction to 20 μL with H_2O and add 20 μL of Cy5 dye in a 0.6 μL thin-walled PCR tubes. Use the same amount of DNA for Nuclei and add 20 μL of Cy3 dye (*see* **Notes 6 and 7**).
2. Incubate each sample at 95^{o}C for 10 minutes on a thermocycler. Immediately place the samples on ice-water for 2 minutes (*see* **Notes 8**).
3. After instant chill, add 5 μL of 50 mM dNTPs, 4 μL of VWR H_2O, and 1 μL of high concentration Klenow fragment. Allow the labeling reaction to proceed for 4.5 hours at 37^{o}C in a thermocycler (*see* **Notes 9**).
4. To stop the labeling reaction, add 5 μL of 0.5 M EDTA. Precipitate the labeled DNA by adding 5.75 μL of 5 M NaCl and 55 μL of isopropanol. Incubate at room temperature for

10 minutes and pellet the DNA at maximum speed for 10 minutes in a room-temperature centrifuge. Wash the colored pellet with ice-cold 80% ethanol and recentrifuge for 2 minutes maximum speed at room temperature. Discard the supernatant carefully and speed-vac dry the samples for at most 5 minutes to prevent over-drying.

5. Resuspend the pellet in 20 μL of VWR H_2O. To quantify the labeled DNA, dilute 0.5 μL of sample in 4.5 μL H_2O and use 1.5 μL for nanodrop measurement.
6. Combine 34 μg of Cy5-labeled salt fraction with 34 μg of Cy3-labeled Nuclei and concentrate the volume to 12.5 μL using a speed-vac. Proceed with Nimblegen hybridization protocol with high density *Drosophila* microarrays.

3.3.2 Paired-end Solexa library preparation

1. Library preparation for sequencing of salt fractions follows closely the Illumina protocol provided with its Paired-End Sample Preparation Kit. Use 500 ng of DNA and follow the Illumina protocol for end repair, 3' adenylation, and adapter ligation.
2. After adapter ligation, samples must be size selected and purified from free adapters. Electrophorese the samples on a 2% agarose gel and excise DNA from 100 – 600 bp range (*see* **Notes 10**). This will isolate double-stranded DNA derived from mono- to tri-nucleosomes. Extract DNA from the agarose gel following Qiagen gel extraction kit using MinElute columns and elute the DNA with 30 uL of EB buffer.
3. To amplify the library, follow Illumina's protocol on PCR amplification and clean up. For salt fractionation, it is not necessary to perform a secondary size selection process after amplification. Quality control analysis varies depending on the sequencing facility's specification, but may include PicoGreen quantification, Bionalyzer analysis, and qPCR

quantification. The resulting paired-end library can be sequenced using one lane of Illumina Genome Analyzer for both single-end and paired-end sequencing.

4. Notes

1. This protocol is specifically designed for S2 cells. However, it can be adapted to any cultured cells from any species, provided that the cells are undergoing exponential growth. Changes in growth phase lead to changes in transcriptional program, which may also lead to changes in chromatin structure. As such, salt fractionation methods are most reproducible and reliable for cells in the same log phase growth. For adherent cells, follow established trypsin conditions for cell harvest and proceed to step 3 of section 3.1.
2. The conditions for cell lysis must be empirically determined for each cell type and for each species. This can be done by altering the concentration of final NP-40 from 0.08 – 0.8% in TM2 solution. Using the lysis protocol described in step 5 of section 3.1, check for complete lysis of the cellular membrane while maintaining nuclei integrity by removing an aliquot and examining the nuclei under a microscope in comparison to intact cells. Alternatively, one can use Trypan-blue exclusion. Remove an aliquot of the lysis, add an equal volume of Trypan-blue solution, and visualize the nuclei under a microscope.
3. The nuclear membrane is sensitive to mechanical disruption, which can lead to lysis and subsequent release of chromatin into the solution. This results in a nuclear pellet that is difficult to resuspend in solution. Formation of clumps in resuspended nuclei is a telltale sign of nuclear lysis. Discard samples and repeat nuclei preparation using less NP-40 in the lysis buffer or gentler handling of nuclei.

4. MNase conditions must be determined empirically for consistent digestion. The optimal conditions yield mostly mononucleosomes with decreasing amounts of di- and tri-nucleosomes (Figure I-1). To determine MNase conditions, prepare nuclei from 150 × 10^6 cells as described in section 3.1 and divide the nuclei into 5-10 aliquots, depending on the number of MNase conditions to be tested. Add increasing amounts of MNase starting with 0.5 U per reaction and incubate each sample in 37^{o}C water bath for 10 minutes. Isolate DNA as described in section 3.3 and electrophorese an aliquot in 1.5% agarose gel. Increasing amounts of MNase should yield increasing intensity of the mononucleosome band at 150 bp. Determine which amount of MNase yields the distinct ladder of mono- to tri-nucleosomes and repeat the conditions to ensure replicability. Furthermore, intermittent mixing of the MNase reaction is important as nuclei can pellet in the span of the 10 minute digestion, which can lead to unequal and incomplete digestion.
5. The concentration of nuclei in the salt buffers is about 2 × 10^5 nuclei per μL of buffer. This ratio is critical for the maintenance of nuclear integrity through the interaction of Mg^{2+} ions with the chromatin and nuclear complex. When adapting this protocol for using less starting number of cells, different cell types, or different species, this ratio must be maintained for proper nuclear integrity.
6. For control purposes, dye swaps may be necessary to determine dye-labeling biases. In these cases, the nuclei fraction can be labeled with Cy5 and salt fractions with Cy3. Alternatively, the Nuclei fraction can be labeled with Cy5 and Cy3 so that hybridization should result in zero enrichment and depletion in the microarray profile. Labeling biases

(Henikoff, 2008) can be identified by lack of inverse correlation with the dye swap, or non-zero profiles in the Nuclei hybridization.

7. The labeling reactions described are half of the total volume of Nimblegen dye labeling protocol. This is sufficient for *Drosophila* samples because of their smaller genome size, but for mammalian systems such as mouse, a full reaction is optimal with 1 μg of starting material. In this case, follow Nimblegen dye labeling protocol closely.
8. It is important that samples are chilled in ice-water for faster and more uniform cooling. Otherwise, efficiency of labeling reaction is decreased.
9. The Nimblegen protocol calls for 2 hour labeling reactions. However, a single labeling reaction does not usually produce enough material needed for hybridization. For the smaller *Drosophila* genome, a longer labeling reaction produces sufficient material for hybridization without introducing labeling biases. However, for larger and more complex mammalian genomes, labeling biases become more pertinent. Therefore, set up 2 or more full volume reactions (Nimblegen protocol) and limit the length to 2 hours.
10. The standard Illumina protocol for size selection of libraries calls for extraction of a relatively small range of sizes for sequencing. However, one advantage of paired-end sequencing is the ability to measure sizes of the sequenced population. This allows mapping of nucleosomes at base-pair resolution. It should be noted that nicks caused by MNase, which is a single-strand-specific endonuclease, will result in primarily ~147-bp fragments from gel-purified mono-, di- and tri-nucleosomes (Weber et al., 2010), with smaller species indicative of internal nicking (Figure I-2).

APPENDIX II

MEASURING GENOME-WIDE NUCLEOSOME TURNOVER USING CATCH-IT

Modified from an chapter published in Methods in Molecular Biology

Summary

The dynamic interplay between DNA binding proteins and nucleosomes underlies essential nuclear processes such as transcription, replication, and DNA repair. Manifestations of this interplay include the assembly, eviction, and replacement of nucleosomes. Hence, measurements of nucleosome turnover kinetics can lead to insights into the regulation of dynamic chromatin processes. In this chapter, we describe a genome-wide method for measuring nucleosome turnover that uses metabolic labeling followed by capture of newly synthesized histones, which we have termed Covalent Attachment of Tagged Histones to Capture and Identify Turnover (CATCH-IT). Although CATCH-IT can be used with any genome-wide mapping procedure, high-resolution profiling is attainable using paired-end sequencing of native chromatin. Our protocol also includes an efficient Solexa DNA sequencing library preparation protocol that can be used for single-base pair resolution mapping of both nucleosome and subnucleosomal particles. We describe the use of these protocols in the context of a *Drosophila* cell line, but also provide the necessary changes for adaptation to other model systems.

1. Introduction

Nucleosomes have evolved to tightly package DNA in chromosomes, which must be mobilized to allow DNA-binding proteins to gain access to their binding sites and perform DNA-templated processes. For DNA to be accessible, nucleosomes must be displaced, partially unwrapped or evicted, and measuring these dynamic events can provide mechanistic insights into the regulation of chromatin-dependent processes. We have recently developed a method that combines kinetic measurement of nucleosome dissociation and replacement (turnover) with genome-wide read-out technologies that we termed Covalent Attachment of Tagged Histones to Capture and Identify Turnover (CATCH-IT) (Deal et al., 2010). In CATCH-IT, a methionine analog containing an azide moiety is incorporated into newly synthesized proteins, which can then be biotinylated through copper-catalyzed cycloaddition reaction with a biotin-alkyne substrate. Nucleosomes containing the newly synthesized biotinylated histones can be isolated using streptavidin-coated beads, and the extracted DNA used to measure the degree of nucleosome turnover genome-wide.

The increasing affordability of short-read massively parallel sequencing potentially enables the study of epigenomic events at single base-pair resolution. Paired-end sequencing using Illumina's Solexa platform (Bentley, 2006) is becoming especially valuable for epigenomic mapping, as it allows precise determination of both fragment lengths and positions. However, current Solexa library preparation protocols were designed for genomics applications, where fragmentation by random shearing and size selection were intended to provide a uniform population of DNA templates for bulk sequencing and so include a gel-based size-selection step to exclude both large and small fragments. However, for many epigenomic mapping applications, random fragmentation and size selection is inappropriate, and the requirement for large amounts

of starting material can be limiting. To address these issues, we have developed a modified Solexa library protocol that removes gel-based size selection and streamlines DNA cleanup to allow for sequencing of as little as ~10 nanograms of starting material with fewer manipulations (Henikoff et al., 2011). Combining this modified protocol with chromatin-probing experiments such as micrococcal nuclease mapping (Henikoff et al., 2011; Kent et al., 2011), salt fractionation (Henikoff et al., 2009), native chromatin immunoprecipitation (Teves and Henikoff, 2011; Weber et al., 2010), and as we discuss in this chapter, CATCH-IT (Deal et al., 2010), allows for single-base pair resolution analyses of chromatin-based processes.

In this chapter, we present detailed protocols and tips to perform CATCH-IT at high resolution using a *Drosophila* cell line and provide illustrative data of expected results. These protocols should be adaptable to cell lines of other organisms, and we highlight the steps where necessary changes should be made when performing CACTH-IT for other systems.

2. CATCH-IT

Below, we provide a detailed protocol for nucleosome turnover analysis through metabolic labeling. The method relies on the depletion of methionine in growth medium followed by incorporation of the methionine analog azidohomoalanine (Aha) into newly synthesized proteins in place of methionine. Following nuclei isolation from Aha-labeled cells, newly synthesized nuclear proteins containing an Aha-azide moiety can be coupled to an alkyne-linked biotin tag through a copper-catalyzed cycloaddition reaction. Chromatin can then be fragmented down to mononucleosomes using Micrococcal Nuclease (MNase), which digests away unprotected DNA fragments. A standard salt extraction step provides the input material for affinity purification of newly synthesized chromatin using streptavidin beads. Because the

turnover rate of H2A/H2B dimers is faster than that of the central $(H3/H4)_2$ tetramer (Rufiange et al., 2007; Thiriet and Hayes, 2005), the immobilized chromatin is washed with a urea-containing solution that strips off the H2A/HB from nucleosomes and also removes virtually all other bound proteins. DNA extracted from the bead-bound material can then be isolated for genome-wide analysis.

2.1 Solutions and Materials

1. *Drosophila* S2 cells maintained in log phase growth
2. Shields and Sang M3 Insect growth medium without methionine
3. Azidohomoalanine (Anaspec cat # 63669)
4. Methionine (Sigma-Aldrich # M9625)
5. Round bottom 30 mL Corex tubes
6. Table-top centrifuge
7. 1x Phosphate-buffered Saline (PBS)
8. *TM2 buffer (10 mM Tris pH 7.5, 2 mM $MgCl_2$)
9. 10% NP-40 (Sigma-Aldrich # 74385)
10. Refrigerated table-top centrifuge
11. *HB125 buffer (0.125 M Sucrose, 15 mM Tris pH 7.5, 15 mM NaCl, 40 mM KCl, 0.5 mM Spermidine, 0.15 mM Spermine)
12. Copper(II) Sulfate pentahydrate (Sigma-Aldrich # C7631)
13. L-Ascorbic Acid (Sigma-Aldrich # A5960)
14. Biotin Alkyne (Invitrogen # B10185)
15. End over end microcentrifuge tube rotator (ex. Labquake Shaker – Thermo Scientific)

16. 0.5 M EDTA
17. 1 M $CaCl_2$
18. 37°C waterbath
19. Micrococcal Nuclease (MNase) 200 U powder, resuspended to 0.2U/μL (Sigma-Aldrich # N3755)
20. *CSB350 buffer (1x PBS, 213 mM NaCl (350 mM total), 2 mM EDTA, 0.1% Triton X-100)
21. Dynabeads M-280 Streptavidin (Invitrogen # 112.05D)
22. Magnetic rack for microcentrifuge tubes
23. *Urea/NaCl Wash buffer (4M Urea, 0.3 M NaCl, 20 mM Tris pH8, 1 mM EDTA)
24. 10% SDS
25. Phenol-Chloroform-Isoamyl alcohol (25:24:1 v/v)
26. Glycogen 20 mg/mL
27. 200 proof Ethanol
28. TE pH 8 (10 mM Tris pH 8, 1 mM EDTA)
29. Complete-mini, Protease Inhibitor Cocktail EDTA-free (Roche # 1 830 170)

*Buffers are supplemented with Complete-mini, protease inhibitor cocktail prior to use.

2.2 Methionine-free growth medium

In many systems, growth medium without methionine is commercially available. However, for insect cell culture systems, this is not the case. We therefore prepare methionine-free Shields and Sang M3 (SSM3-Met) insect medium for use in CATCH-IT experiments. Table 1 lists the components of M3 medium without methionine for 1 L of medium.

1. Add 1.05 g of Bis-Tris to 650 mL of H_2O and adjust pH to 6.8.
2. Combine all solids (salts, vitamins, amino acids, sugar) and grind into fine powder using a mortar and pestle.
3. Slowly add the ground powder to the Bis-Tris-H_2O while mixing using a magnetic stir bar.
4. After all solids have dissolved, re-adjust the pH to 6.8 and bring the final volume to 1 L. Filter sterilize, and store at 4°C.

2.3 Aha-labeling and biotin coupling

1. Grow S2 cells in two 75 cm^2 flasks to late log phase in rich medium. Remove the medium and replace with 7 mL of SSM3-Met medium. Place the flasks back in the incubator for 30 minutes to starve cells of methionine.
2. Remove the medium and replace with 7 mL of SSM3-Met medium supplemented with 4 mM Azidohomoalanine (Aha) for one flask and 4 mM Met for the other flask. Place the flasks back in the incubator for the desired amount of time (20 minutes to several hours).
3. Harvest the cells from each treatment in 30 mL round bottom Corex tubes. Spin at 1200 x g for 3 minutes. Decant the medium and wash cells with 10 mL of 1X PBS, spin again and decant PBS.

*The following steps (4-6) are used to isolate nuclei from *Drosophila* S2 cells. To adapt this method to other systems, the final concentration of NP-40 must be optimized. Alternatively, other standard protocols for nuclei preparation can be substituted.

4. Resuspend the cells in 1 mL of ice-cold TM2 buffer and transfer to 1.5 mL microcentrifuge tubes. Place on ice for 3 minutes.

5. Add 60 uL of 10% NP-40 and vortex for 5 sec at the low-medium setting. Place cells back on ice for 3 minutes, vortex one more time as before and spin out nuclei at 100 x g for 10 minutes at 4°C.
6. Wash nuclei with 1 mL cold TM2 buffer. Pellet nuclei at 100 x g for 10 minutes at 4°C, decant supernatant and resuspend gently in 200 μL of cold HB 125 buffer. If performing a time course, leave each successive sample on ice at this point until all are collected, then proceeed to step 7 with all samples.
7. Prepare reagents as follows for the cycloaddition coupling reaction:
    a. Weigh out 25 mg of $CuSO_4$ and 88 mg of ascorbic acid into 1.5 mL tubes and dissolve each in 1 mL of H_2O to give solutions of 100 mM $CuSO_4$ and 500 mM ascorbic acid. Combine 100 μL of $CuSO_4$ with 100 μL of ascorbic acid. The solution turns yellow as Cu^{2+} is reduced to Cu^{+} by ascorbic acid.
    b. Prepare 20 μL of 2 mM Biotin-PEO-Alkyne by 1:10 dilution of 20 mM stock solution.
8. Place nuclei suspensions at room temperature and then add the following to each 200-μL sample, mixing well after each addition:

    5 μL of 2 mM Biotin-PEO-Alkyne (50 μM final concentration)

    4 μL of $CuSO_4$/ascorbic acid mixture (final concentration is1 mM $CuSO_4$; 5 mM ascorbic acid)
9. Place nuclei suspensions on a microcentrifuge tube rotator at 4°C for 30 minutes. Save 2 μL (1% of total) for western analysis at the end of the procedure.
10. Pellet nuclei at 100 x g for 5 minutes at 4°C.

11. Remove supernatant thoroughly and resuspend nuclei gently in 200 μL of cold HB 125 buffer. Repeat steps 8 through 10 using freshly prepared cycloaddition reagents.

We tested the labeling and coupling process using total cellular extracts from Met- or Aha-labeled cells. We performed Western blot analysis on aliquots saved from each step of the process and probed for biotin using α-Streptavidin antibody (Figure II-1). We found that Aha- but not methionine-labeled cells have incorporated biotin into general cellular proteins. This test also showed that a second biotin coupling is required for saturation of all Aha-labeled proteins.

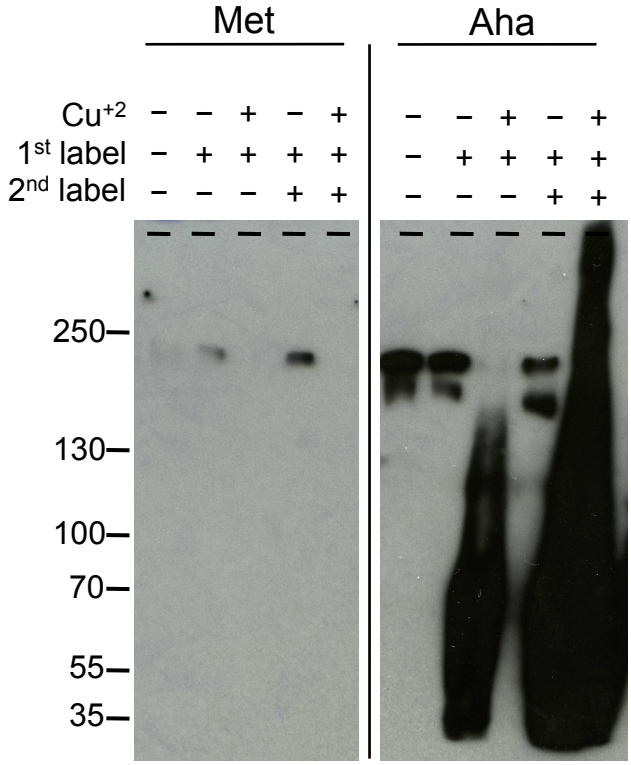

Figure II-1. Incorporation of Aha into cellular proteins. Cells depleted of methionine were incubated in SSM3-Met medium supplemented with 4 mM Met or Aha for 20 minutes. Cells were then lysed in 1% SDS in 1x PBS, and the resulting total protein extract was subjected to biotin coupling as described either with or without Cu^{+}. Aliquots before treatment, after the 1st labeling, and after the 2nd labeling, were subjected to Western blot analysis using α-Streptavidin to visualize biotinylated proteins.

2.4 Chromatin fragmentation and extraction

1. Resuspend nuclei in 250 μL HB 125 with 1 mM EDTA and add $CaCl_2$ to 2 mM final concentration. Place tubes in a 37°C water bath for 5 minutes.

2. Add 2 uL of the MNase solution to each tube in the water bath and mix gently by inversion and flicking several times. Continue digestion for 10 minutes with mixing each minute. This level of digestion should give mostly mononucleosomes (Figure II-2), but this should be optimized for each system by varying the concentration of MNase and/or the length of digestion time.
3. Add EDTA to 2 mM to stop the reaction and place the tubes on ice. Spin at 100 x g for 10 minutes at 4°C.
4. Resuspend nuclei in 300 μL of CSB 350 and mix on a microcentrifuge tube rotator at 4°C for at least one hour to overnight.
5. Centrifuge nuclei at 100 x g for 10 minutes at 4°Cand save the supernatant (soluble chromatin). Clarify the soluble chromatin at by centrifugation at full speed for 5 minutes and move to a fresh tube.
6. Remove a 10 μL aliquot of the soluble chromatin (Input) for DNA purification and 12 μL for Western analysis, coomassie gel and protein concentration measurement. Resuspend the pellet in loading buffer and save for Western analysis.

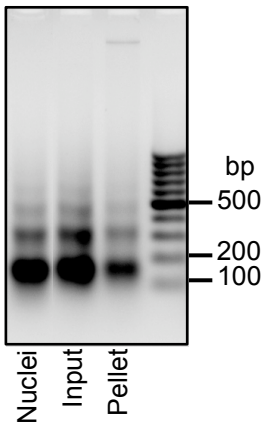

Figure II-2. Nucleosome laddering of MNase-digested chromatin. After biotin coupling, intact nuclei were digested with 0.4 U MNase for 10 minutes and subjected to salt extraction. DNA from total nuclei, 350 mM salt-extracted chromatin, and pellet fractions were electrophoresed on a 1.5% agarose gel and visualized with ethidium bromide.

2.5 Streptavidin affinity capture

1. To the soluble chromatin add 25 μL (beads + buffer, same as original ratio as in stock slurry) of CSB-washed M280 Streptavidin-coated Dynabeads. Rotate for 1.5 hours at 4°C.
2. Place tubes on a Dynal magnet rack for several minutes, then save the supernatant as "unbound" for Western analysis. Resuspend the beads in 700 μL of Urea/NaCl wash buffer and rotate at 4°C for 5 minutes to strip the H2A/H2B dimers from the chromatin.
3. Place tubes on the magnet rack, decant and resuspend in 700 μL of CSB 350. Transfer the beads and buffer to a fresh tube and rotate at 4°C for 5 minutes. Collect and decant the beads and proceed with DNA isolation or resuspend the beads in 20 μL CSB 350 and freeze at -20°C.

Western analysis of total nuclei, input, unbound, and pellet fractions during the Streptavidin affinity capture process (Figure II-3) shows lack of Streptavidin signal in the unbound fraction compared to others, indicating an efficient capture of biotin-coupled, Aha-labeled proteins.

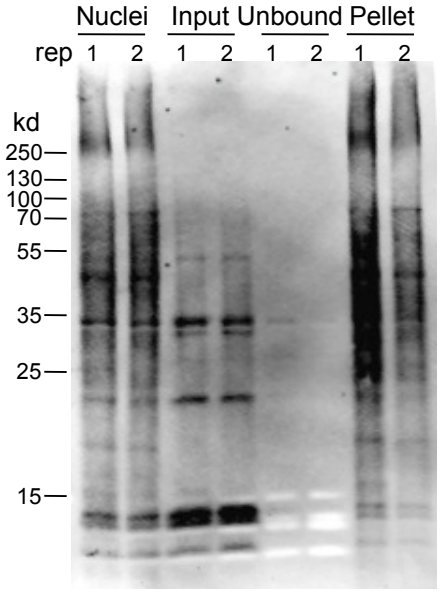

Figure II-3. Efficient capture of biotinylated proteins. CATCH-IT was performed on Drosophila S2 cells in 2 replicates. Aliquots of total nuclei, input, unbound, and pellet fractions were subjected to a Western analysis as in Figure II-1.

2.6 DNA isolation

1. Bring the volumes of input and beads to 200 μL with CSB 350 and add SDS to 0.5%. Add 1 μL of RNAse A and incubate at 37°C for 10 minutes, then add 1 μL of Proteinase K and incubate at 70°C for 10 minutes. Mix beads frequently during each enzyme digestion.
2. Extract the DNA twice with Phenol/Chloroform and precipitate by adding 2.5 volumes of cold ethanol. Wash with 75% ethanol and resuspend in 20 μL of 0.1x TE pH8.
3. Quantify DNA either by using the PicoGreen fluorescence assay (Invitrogen) or by measuring the OD_{260} with a spectrophotometer such as Nanodrop.

3. Modified Solexa Library Preparation

The following protocol for library uses the same enzymological steps as in the established Illumina Paired-End Sample Preparation protocol (Quail et al., 2008), with modifications to the clean-up and purification steps that allow for efficient recovery of fragments as small as 25-bp (Henikoff et al., 2011). In this process, the ends of the starting DNA material are made blunt and 5'-phosphorylated. An 'A' nucleotide is added at the 3' ends to prevent self-ligation and permit specific base-pairing of adapters with a 5' phosphate and a 3' T-overhang The adapted samples are then amplified using primers with 3' complementarity to the adapters, using a 60^{o}C extension step to minimize preferential loss of AT-rich fragments (Lopez-Barragan et al., 2010). In the Illumina protocol, the samples are size-selected following the ligation of adapters both to remove unligated adapters and to isolate a specific subset of the samples for sequencing. This step is designed for samples that are randomly sheared or where single-end

sequencing leads to fragment size ambiguity. However, if chromatin is enzymatically cleaved and paired-end sequenced, the fragment size distribution can reveal valuable features of the chromatin landscape. Therefore, we eliminated the gel-based size-selection step. Instead, we use Agencourt AMPure XP beads both to purify the ligated products from unligated primers and for post-PCR clean-up. Using the dilution factor specified by the manufacturer, the Ampure beads result in a strict size cut-off at 90-100 bp. Because the adapters add ~65 bp to the starting material, insert sizes as small as 25 bp will be present in the library. Another modification is that all QIAgen clean-up steps have been replaced with phenol/chloroform extractions to stop the reactions followed by spin column cleanup to purify the DNA. This modification, combined with the use of low-retention (siliconized) microcentrifuge tubes, minimizes the loss of DNA and allows for lower amounts of starting material to be used.

3.1 Solutions and Materials

1. Low retention 1.5 mL microcentrifuge tubes
2. 10X Annealing buffer (0.5 M NaCl, 0.1M Tris pH 8, 0.01M EDTA)
3. 10X T4 DNA Ligase buffer with 10 mM ATP (New England Biolabs, NEB, # B0202S)
4. 40 mM dNTP (10 mM each)
5. T4 DNA Polymerase 5U/µL (Invitrogen # 100004994)
6. DNA Polymerase I, Large (Klenow) Fragment 50U/µL (NEB # M0210M)
7. T4 Polynucleotide Kinase (PNK) 10U/µL (NEB # M0201L)
8. Phenol/Chloroform/Isoamyl (25:24:1)
9. Illustra MicroSpin S-300 HR Columns (GE Healthcare # 27-5130-01)
10. SpeedVac

11. 10X NEB Buffer 2(NEB # B7002S)
12. 100 mM dATP (Invitrogen # 10216018)
13. Klenow Fragment (3'→5' exo-) 50U/μL (NEB # M0212M)
14. Illumina Paired End (PE) Adapter Mix or independently synthesized adapter mix
15. T4 DNA Ligase (Rapid) and 2X buffer (Enzymatics # L603-HC-L)
16. Agencourt AMPure XP magnetic beads (Agencourt # A63881)
17. Magnetic rack for microcentrifuge tubes
18. 70% Ethanol
19. 0.1X TE (1 mM Tris pH 8, 0.1 mM EDTA)
20. Phusion High Fidelity DNA Polymerase with 5X HF buffer (Finnzymes # F-530L)
21. Illumina Paired End (PE) Primers or independently synthesized primers
22. 10 mM dNTP (2.5 mM each)
23. Thermocycler

3.2 Paired-End Adapter and Primers

We have used both commercially available Paired-End Adapter mix and PCR primers from Illumina and custom-made adapter and primer oligos with comparable results. The following synthesized adapters and primers are compatible with the Illumina Paired-End platform. Therefore, the protocol outlined below does not make a distinction between these options. The oligonucleotide sequences are as follows:

PE Adapter1: [Phosphate]GATCGGAAGAGCGGTTCAGCAGGAATGCCGA*G

PE Adapter2: ACACTCTTTCCCTACACGACGCTCTTCCGATC*T

PE Forward Primer:

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T

PE Reverse Primer:

CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T

These oligonucleotides are PAGE-purified. A phosphorothioate linkage before the last nucleotide as denoted in the sequence (*) is intended to prevent 3' exonuclease activity during the ligation step (Quail et al., 2008).

Adapter oligonucleotides are dissolved to 100 μM in water, mixed in equimolar amounts and annealed by addition of 10X annealing buffer to a final concentration of 50 mM NaCl, 10 mM Tris pH8, 1 mM EDTA. The mixed adapters are annealed by heating to 98^{o}C for 10 minutes in a thermocycler and slowly cooling (-1^{o}C/minutes) to 25^{o}C.

3.3 End repair

1. Measure the starting DNA concentration spectrophotometrically or by fluorescence. A total of 10-500 ng of enzymatically digested DNA can be used to prepare sequencing libraries. Bring the desired starting amount to 20 uL by addition of water or using a SpeedVac without heat to concentrate in a low retention microcentrifuge tube.
2. Prepare the "End Repair" (ER) master mix by combining the following:
    a. 18.5 μL water
    b. 5 μL 10X T4 DNA Ligase buffer
    c. 2 μL 40 mM dNTP

d. 1.5 µL T4 DNA polymerase

e. 0.5 µL Klenow fragment (diluted to 5 U/µL from 50 U/µL stock)

f. 2.5 µL T4 PNK

3. If preparing multiple samples, combine a slight excess of each component of the ER master mix to compensate for pipetting losses during transfers. For example, to make libraries from 8 different samples, combine 8.2 times each of the component in the master mix.
4. Add 30 µL of ER master mix to 20 uL of DNA (10-500 ng).
5. Incubate the sample in a 20°C water bath for 30 minutes. A water bath can be as simple as an ice bucket filled with water that is measured to be 20°C using a thermometer.

3.4 Phenol extraction and column purification

1. Extract the DNA with 50 µL of Phenol/Chloroform/isoamyl. Vortex to thoroughly mix, and separate the aqueous layer by centrifugation at maximum speed for 1 minute using a table top centrifuge. Carefully remove the organic layer from the bottom of the tube.
2. To prepare the Microspin S-300 HR column, snap off the bottom tip, slightly unscrew the cap, and place in tube holder. Spin buffer off for 1 minute at 800 x g and replace the column into a new low-retention microcentrifuge tube.
3. Decant the aqueous layer from step 1 and drain the remaining organic layer by touching the pipette tip to the tube wall. Transfer the aqueous layer into the prepared Microspin S-300 HR column. Extract the DNA by centrifugation for 2 minutes at 800 x g. The resulting eluate will be about 50 – 60 µL in volume.

3.5 A-tailing

1. Reduce the eluate volume to 35 µL in a SpeedVac without heat for about 15 minutes.
2. Prepare the "A-tailing" (A-t) master mix by combining the following:
    a. 5 µL 10X NEB Buffer 2
    b. 10 µL 1 mM dATP
    c. 0.3 µL Klenow exo- (50U/µL)
3. Again, if preparing multiple samples, include a slight excess of each component as in the ER master mix procedure.
4. Add 15.3 µL of the A-t master mix to 35 µL of end-repaired DNA sample
5. Incubate at 37°C for 30 minutes.
6. Extract the DNA by phenol extraction and column purification (section 3.3).

3.6 Adapter ligation and ampure bead purification

1. Reduce the eluate volume to 18 µL in a SpeedVac without heat for about 30 – 45 minutes.
2. Add 1 µL of 1 mM PE Adapter mix. This amount is best used for small amounts of starting material (10 – 50 ng of DNA). When using a large amount of starting material, adjust the adapter amount to give an estimated 10:1 adapter:starting DNA molar ratio.
3. Add 25 µL of 2X Rapid DNA ligase buffer (Enzymatics) and 5 µL of Rapid T4 DNA ligase. If preparing multiple samples, combine the buffer and ligase for a master mix and add 30 µL of this to the DNA-adapter mix.
4. Incubate the sample at 20°C water bath for 15 minutes to ligate adapters.

5. To extract the DNA and remove excess adapters, add 90 μL of AMPure XP magnetic bead slurry to the sample, mix by pipetting 10 times, and hold at room temperature for 5 minutes.
6. Place the sample in a magnetic tube holder until the beads are cleared from the slurry and accumulate on the side (~2 minutes), and aspirate off the solution.
7. While still on the magnet, wash the beads by adding 1 ml 70% ethanol. Aspirate the ethanol after 30 seconds. Repeat this wash one more time, carefully removing excess ethanol.
8. Allow the beads to dry for no more than 5 minutes. Remove the tube from the magnet and add 40 μL of 0.1x TE pH 8. Mix thoroughly and replace the tube on the magnetic tube holder. Transfer the eluate into a new low-retention microcentrifuge tube. This now contains the DNA material with ligated PE adapters.

3.7 PCR amplification and final purification

1. A 20 μL PCR reaction volume is generally sufficient to produce enough product for Paired-End sequencing using the Illumina platform. Prepare a "PCR" master mix by combining the following:
    a. 4 μL 5X Phusion buffer HF
    b. 1.6 μL 10 mM dNTP (2.5 mM each)
    c. 0.4 μL Forward Primer
    d. 0.4 μL Reverse Primer
    e. 0.2 μL Phusion HF Polymerase (Finnzymes)
    f. 8.4 μL H_2O

2. Add 15 μL of the PCR master mix to 5 μL of adapter-ligated material and proceed with PCR using the following cycling parameters:
    a. 98°C for 30 seconds
    b. 12 – 18 cycles of:
        i. 98°C for 10 seconds
        ii. 65°C for 30 seconds
        iii. 60°C for 30 seconds
    c. 60°C for 5 minutes
    d. hold at 8°C
3. Clean up the PCR products by adding 36 μL of AMPure XP beads and following the purification method described above. Elute the sample with 40 μL of 0.1x TE pH8. Check the library on a 2% agarose gel with ethidium bromide. The adapter and PCR primers add about 120 bp onto the starting DNA. Therefore, the distribution of the library should be shifted 120 bp higher than the starting DNA (Figure II-4).
4. Accurately measure the concentration of the library either by fluorescence-based assays such as the PicoGreen (Invitrogen) or by qPCR as outlined in Illumina protocols and dilute for application to the flow cell (2 μM for the Illumina Hi-Seq 2000).

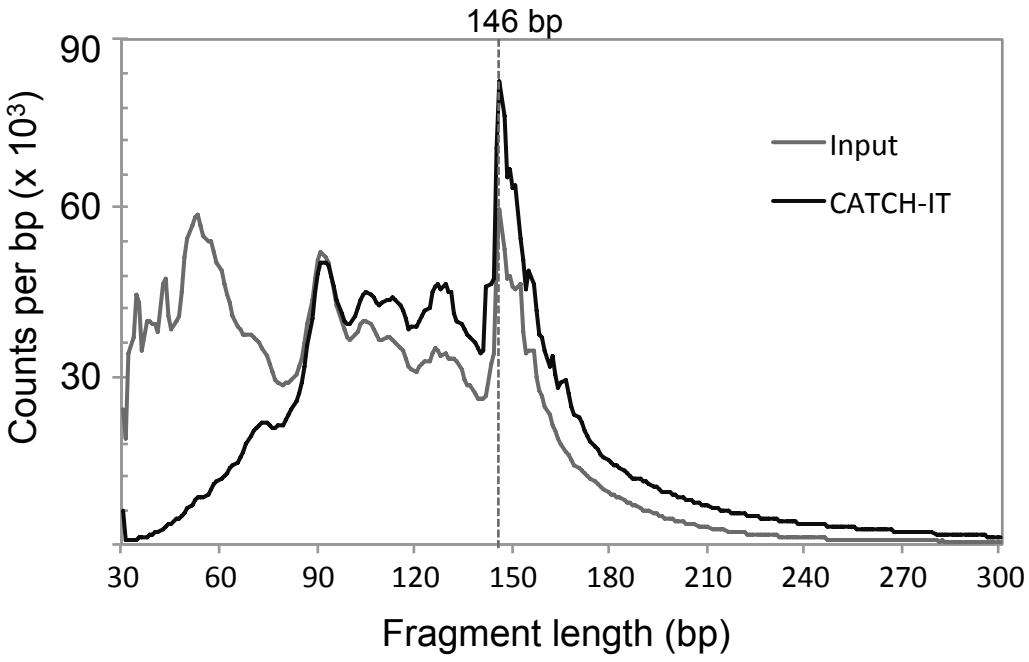

Figure II-4. Length distribution of sequenced fragments. CATCH-IT was performed on Drosophila S2 cells, and a paired-end Solexa library was generated from the input and CATCH-IT material as described here. Paired-end sequencing for each sample was performed in a single lane of an Illumina Hi-Seq 2000 Instrument by the FHCRC Genomics Shared Resource (http://sharedresources.fhcrc.org/core-facilities/genomics), and fragments were mapped to the Drosophila genome using Novoalign alignment program (http://www.novocraft.com/main/index.php). The length distribution of all mapped fragments are shown here at base-pair resolution. The dominant nucleosomal peak is centered at 146 bp.

REFERENCES

Aguilar-Arnal, L., and Sassone-Corsi, P. (2013). The circadian epigenome: how metabolism talks to chromatin remodeling. Curr. Opin. Cell Biol. *25*, 170-176.

Akerfelt, M., Morimoto, R.I., and Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. Nat. Rev. Mol. Cell Biol. *11*, 545-555.

Al-Fageeh, M.B., and Smales, C.M. (2006). Control and regulation of the cellular responses to cold shock: the responses in yeast and mammalian systems. Biochem. J. *397*, 247-259.

Allen, T.A., Von Kaenel, S., Goodrich, J.A., and Kugel, J.F. (2004). The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. Nat. Struct. Mol. Biol. *11*, 816-821.

Ammar, R., Torti, D., Tsui, K., Gebbia, M., Durbic, T., Bader, G.D., Giaever, G., and Nislow, C. (2012). Chromatin is an ancient innovation conserved between Archaea and Eukarya. eLife *1*, e00078.

Ashburner, M., and Bonner, J.J. (1979). The induction of gene activity in drosophilia by heat shock. Cell *17*, 241-254.

Auerbach, R.K., Euskirchen, G., Rozowsky, J., Lamarre-Vincent, N., Moqtaderi, Z., Lefrancois, P., Struhl, K., Gerstein, M., and Snyder, M. (2009). Mapping accessible chromatin regions using Sono-Seq. Proc. Natl. Acad. Sci. *106*, 14926-14931.

Bai, L., Charvin, G., Siggia, E.D., and Cross, F.R. (2010). Nucleosome-Depleted Regions in Cell-Cycle-Regulated Promoters Ensure Reliable Gene Expression in Every Cell Cycle. Dev. Cell *18*, 544-555.

Baranello, L., Bertozzi, D., Fogli, M.V., Pommier, Y., and Capranico, G. (2009). DNA topoisomerase I inhibition by camptothecin induces escape of RNA polymerase II from promoter-proximal pause site, antisense transcription and histone acetylation at the human HIF-1 gene locus. Nucleic Acids Res. *38*, 159-171.

Baumann, M., Pontiller, J., and Ernst, W. (2010). Structure and Basal Transcription Complex of RNA Polymerase II Core Promoters in the Mammalian Genome: An Overview. Mol. Biotechnol. *45*, 241-247.

Becavin, C., Barbi, M., Victor, J.-M., and Lesne, A. (2010). Transcription within Condensed Chromatin: Steric Hindrance Facilitates Elongation. Biophys. J. *98*, 824-833.

Belyaeva, E.S., and Zhimulev, I.F. (1976). RNA synthesis in the Drosophila melanogaster puffs. Cell Differ. *4*, 415-427.

Bendixen, C., Thomsen, B., Alsner, J., and Westergaard, O. (1990). Camptothecin-stabilized topoisomerase I-DNA adducts cause premature termination of transcription. Biochemistry *29*, 5613-5619.

Bentley, D.R. (2006). Whole-genome re-sequencing. Curr. Opin. Genet. Dev. *16*, 545-552.

Bermúdez, I., García-Martínez, J., Pérez-Ortín, J.E., and Roca, J. (2010). A method for genome-wide analysis of DNA helical tension by means of psoralen-DNA photobinding. Nucleic Acids Res. *38*, e182.

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

Blaszczak, A., Zylicz, M., Georgopoulos, C., and Liberek, K. (1995). Both ambient temperature and the DnaK chaperone machine modulate the heat shock response in Escherichia coli by

regulating the switch between sigma 70 and sigma 32 factors assembled with RNA polymerase. EMBO J. *14*, 5085-5093.

Bohm, V., Hieb, A.R., Andrews, A.J., Gansen, A., Rocker, A., Toth, K., Luger, K., and Langowski, J. (2011). Nucleosome accessibility governed by the dimer/tetramer interface. Nucleic Acids Res. *39*, 3093-3102.

Bonner, J., Dahmus, M.E., Fambrough, D., Huang, R.C., Marushige, K., and Tuan, D.Y. (1968). The Biology of Isolated Chromatin: Chromosomes, biologically active in the test tube, provide a powerful tool for the study of gene action. Science *159*, 47-56.

Bonner, J.J., and Pardue, M.L. (1976). The effect of heat shock on RNA synthesis in Drosophila tissues. Cell *8*, 43-50.

Brill, S.J., DiNardo, S., Voelkel-Meiman, K., and Sternglanz, R. (1987). DNA topoisomerase activity is required as a swivel for DNA replication and for ribosomal RNA transcription. NCI monographs, 11-15.

Cabrerizo, M., Bartolome, J., Otero, M., Ruiz-Moreno, M., and Carreno, V. (1999). Sequence variation of hepatitis B virus precore-core open reading frame isolated from serum and liver of children with chronic hepatitis B before and after interferon treatment. J. Med. Virol. *58*, 208-214.

Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. Nature *461*, 193-198.

Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J.M., and Pascual-Montano, A. (2007). GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. Genome Biol. *8*, R3.

Cernilogar, F.M., Onorati, M.C., Kothe, G.O., Burroughs, A.M., Parsi, K.M., Breiling, A., Lo Sardo, F., Saxena, A., Miyoshi, K., Siomi, H.*, et al.* (2011). Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature *480*, 391-395.

Chen, W., Luo, L., and Zhang, L. (2010). The organization of nucleosomes around splice sites. Nucleic Acids Res. *38*, 2788-2798.

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

Clark, D.J., and Felsenfeld, G. (1991). Formation of nucleosomes on positively supercoiled DNA. EMBO J. *10*, 387-395.

Collins, I., Weber, A., and Levens, D. (2001). Transcriptional consequences of topoisomerase inhibition. Mol. Cell. Biol. *21*, 8437-8451.

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.

Costlow, N., and Lis, J.T. (1984). High-resolution mapping of DNase I-hypersensitive sites of Drosophila heat shock genes in Drosophila melanogaster and Saccharomyces cerevisiae. Mol. Cell. Biol. *4*, 1853-1863.

Courvalin, J.C., Bouton, M.M., Baulieu, E.E., Nuret, P., and Chambon, P. (1976). Effect of estradiol on rat uterus DNA-dependent RNA polymerases. Studies on solubilized enzymes. J. Biol. Chem. *251*, 4843-4849.

Crawford, G.E., Holt, I.E., Mullikin, J.C., Tai, D., Blakesley, R., Bouffard, G., Young, A., Masiello, C., Green, E.D., Wolfsberg, T.G., and Collins, F.S. (2004). Identifying gene regulatory

elements by genome-wide recovery of DNase hypersensitive sites. Proc. Natl. Acad. Sci. *101*, 992-997.

Cui, P., Zhang, L., Lin, Q., Ding, F., Xin, C., Fang, X., Hu, S., and Yu, J. (2010). A novel mechanism of epigenetic regulation: nucleosome-space occupancy. Biochem. Biophys. Res. Commun. *391*, 884-889.

Czarnota, G.J., Bazett-Jones, D.P., Mendez, E., Allfrey, V.G., and Ottensmeyer, F.P. (1997). High Resolution Microanalysis and Three-Dimensional Nucleosome Structure Associated with Transcribing Chromatin. Micron *28*, 419-431.

de Nadal, E., Ammerer, G., and Posas, F. (2011). Controlling gene expression in response to stress. Nature Reviews Genetics *12*, 833-845.

Deal, R.B., Henikoff, J.G., and Henikoff, S. (2010). Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. Science *328*, 1161-1164.

Desai, N.A., and Shankar, V. (2003). Single-strand-specific nucleases. FEMS Microbiol. Rev. *26*, 457-491.

Dion, M.F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., and Rando, O.J. (2007). Dynamics of replication-independent histone turnover in budding yeast. Science *315*, 1405-1408.

DiTacchio, L., Le, H.D., Vollmers, C., Hatori, M., Witcher, M., Secombe, J., and Panda, S. (2011). Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. Science *333*, 1881-1885.

Doi, M., Hirayama, J., and Sassone-Corsi, P. (2006). Circadian regulator CLOCK is a histone acetyltransferase. Cell *125*, 497-508.

Durand-Dubief, M., Svensson, J.P., Persson, J., and Ekwall, K. (2011). Topoisomerases, chromatin and transcription termination. Transcription *2*, 66-70.

Durand-Dubief, M.e.l., Persson, J., Norman, U., Hartsuiker, E., and Ekwall, K. (2010). Topoisomerase I regulates open chromatin and controls gene expression in vivo. EMBO J. *29*, 2126-2134.

Erkina, T.Y., Zou, Y., Freeling, S., Vorobyev, V.I., and Erkine, A.M. (2010). Functional interplay between chromatin remodeling complexes RSC, SWI/SNF and ISWI in regulation of yeast heat shock genes. Nucleic Acids Res. *38*, 1441-1449.

Eskiw, C.H., Rapp, A., Carter, D.R., and Cook, P.R. (2008). RNA polymerase II activity is located on the surface of protein-rich transcription factories. J. Cell Sci. *121*, 1999-2007.

Etchegaray, J.P., Lee, C., Wade, P.A., and Reppert, S.M. (2003). Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. Nature *421*, 177-182.

Etchegaray, J.P., Yang, X., DeBruyne, J.P., Peters, A.H., Weaver, D.R., Jenuwein, T., and Reppert, S.M. (2006). The polycomb group protein EZH2 is required for mammalian circadian clock function. J. Biol. Chem. *281*, 21209-21215.

Feng, D., and Lazar, M.A. (2012). Clocks, metabolism, and the epigenome. Mol. Cell *47*, 158-167.

Fish, R.N., and Kane, C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. Biochim. Biophys. Acta *1577*, 287-307.

Fritah, S., Col, E., Boyault, C., Govin, J., Sadoul, K., Chiocca, S., Christians, E., Khochbin, S., Jolly, C., and Vourc'h, C. (2009). Heat-shock factor 1 controls genome-wide acetylation in heat-shocked cells. Mol. Biol. Cell *20*, 4976-4984.

Fujita, J. (1999). Cold shock response in mammalian cells. J Mol Microbiol Biotechnol *1*, 243-255.

Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Toth, K., Langowski, J., and Seidel, C.A. (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET. Proc. Natl. Acad. Sci. *106*, 15308-15313.

Gartenberg, M.R., and Wang, J.C. (1992). Positive supercoiling of DNA greatly diminishes mRNA synthesis in yeast. Proc. Natl. Acad. Sci. *89*, 11461-11465.

Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science *324*, 1447-1451.

Giaever, G.N., and Wang, J.C. (1988). Supercoiling of intracellular DNA can occur in eukaryotic cells. Cell *55*, 849-856.

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.

Gilchrist, D.A., Fargo, D.C., and Adelman, K. (2009). Using ChIP-chip and ChIP-seq to study the regulation of gene expression: Genome-wide localization studies reveal widespread regulation of transcription elongation. METHODS, 1-11.

Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K., 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.

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.

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., and Lieb, J.D. (2007). FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res. *17*, 877-885.

Guertin, M.J., and Lis, J.T. (2010). Chromatin landscape dictates HSF binding to target DNA elements. PLoS Genet. *6*.

Guertin, M.J., Martins, A.L., Siepel, A., and Lis, J.T. (2012). Accurate Prediction of Inducible Transcription Factor Binding Intensities In Vivo. PLoS Genet. *8*, e1002610.

Guisbert, E., Yura, T., Rhodius, V.A., and Gross, C.A. (2008). Convergence of molecular, modeling, and systems approaches for an understanding of the Escherichia coli heat shock response. Microbiol. Mol. Biol. Rev. *72*, 545-554.

Gupta, P., Zlatanova, J., and Tomschik, M. (2009). Nucleosome Assembly Depends on the Torsion in the DNA Molecule: A Magnetic Tweezers Study. Biophys. J. *97*, 3150-3157.

Haaf, T., Werner, P., and Schmid, M. (1993). 5-Azadeoxycytidine distinguishes between active and inactive X chromosome condensation. Cytogenet Cell Genet *63*, 160-168.

Henikoff, J.G., Belsky, J.A., Krassovsky, K., MacAlpine, D.M., and Henikoff, S. (2011). Epigenome characterization at single base-pair resolution. Proc. Natl. Acad. Sci. *108*, 18318-18323.

Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. Nature Reviews Genetics *9*, 15-26.

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., and Ahmad, K. (2009). Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res. *19*, 460-469.

Herbert, K.M., La Porta, A., Wong, B.J., Mooney, R.A., Neuman, K.C., Landick, R., and Block, S.M. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell *125*, 1083-1094.

Hesselberth, J.R., Chen, X., Zhang, Z., Sabo, P.J., Sandstrom, R., Reynolds, A.P., Thurman, R.E., Neph, S., Kuehn, M.S., Noble, W.S., *et al.* (2009). Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. Nat. Methods *6*, 283-289.

Hieda, M., Winstanley, H., Maini, P., Iborra, F.J., and Cook, P.R. (2005). Different populations of RNA polymerase II in living mammalian cells. Chromosome Res. *13*, 135-144.

Hizume, K., Yoshimura, S.H., and Takeyasu, K. (2004). Atomic force microscopy demonstrates a critical role of DNA superhelicity in nucleosome dynamics. Cell Biochem. Biophys. *40*, 249-261.

Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. Science *325*, 626-628.

Hughes, A.L., Jin, Y., Rando, O.J., and Struhl, K. (2012). A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. Mol. Cell *48*, 5-15.

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics *9*, 488.

Iyer, V., and Struhl, K. (2012). Poly(dA:dT),a ubiquitouspromoterelementthat stimulatestranscriptionviaitsintrinsicDNA structure. 1-10.

Izban, M.G., and Luse, D.S. (1992). Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates. J. Biol. Chem. *267*, 13647-13655.

Jamrich, M., Greenleaf, A.L., and Bautz, E.K. (1977). Localization of RNA polymerase in polytene chromosomes of Drosophila melanogaster. Proc. Natl. Acad. Sci. *74*, 2079-2083.

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

Jeong, K.S., Ahn, J., and Khodursky, A.B. (2004). Spatial patterns of transcriptional activity in the chromosome of Escherichia coli. Genome Biol. *5*, R86.

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev. *21*, 1519-1529.

Jin, J., Bai, L., Johnson, D.S., Fulbright, R.M., Kireeva, M.L., Kashlev, M., and Wang, M.D. (2010). Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. Nat. Struct. Mol. Biol. *17*, 745-752.

Jones, M.A., Covington, M.F., DiTacchio, L., Vollmers, C., Panda, S., and Harmer, S.L. (2010). Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. Proc. Natl. Acad. Sci. *107*, 21623-21628.

Joshi, R.S., Pina, B., and Roca, J. (2010). Positional dependence of transcriptional inhibition by DNA torsional stress in yeast chromosomes. EMBO J. *29*, 740-748.

Kampmann, M., and Stock, D. (2004). Reverse gyrase has heat-protective DNA chaperone activity independent of supercoiling. Nucleic Acids Res. *32*, 3537-3545.

Keene, M.A., Corces, V., Lowenhaupt, K., and Elgin, S.C. (1981). DNase I hypersensitive sites in Drosophila chromatin occur at the 5' ends of regions of transcription. Proc. Natl. Acad. Sci. *78*, 143-146.

Kent, N.A., Adams, S., Moorhouse, A., and Paszkiewicz, K. (2011). Chromatin particle spectrum analysis: a method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. Nucleic Acids Res. *39*, e26.

Kim, H., Erickson, B., Luo, W., Seward, D., Graber, J.H., Pollock, D.D., Megee, P.C., and Bentley, D.L. (2010). Gene-specific RNA polymerase II phosphorylation and the CTD code. Nat. Struct. Mol. Biol. *17*, 1279-1286.

Kimura, H., Tao, Y., Roeder, R.G., and Cook, P.R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol. Cell. Biol. *19*, 5383-5392.

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

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell *45*, 95-104.

Kobayashi, M., Aita, N., Hayashi, S., Okada, K., Ohta, T., and Hirose, S. (1998). DNA supercoiling factor localizes to puffs on polytene chromosomes in Drosophila melanogaster. Mol. Cell. Biol. *18*, 6737-6744.

Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K., and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. Science *338*, 349-354.

Koster, D.A., Crut, A., Shuman, S., Bjornsti, M.A., and Dekker, N.H. (2010). Cellular strategies for regulating DNA supercoiling: a single-molecule perspective. Cell *142*, 519-530.

Kouzine, F., Gupta, A., Baranello, L., Wojtowicz, D., Ben-Aissa, K., Liu, J., Przytycka, T.M., and Levens, D. (2013). Transcription-dependent dynamic supercoiling is a short-range genomic force. Nat. Struct. Mol. Biol. *20*, 396-403.

Kouzine, F., Liu, J., Sanford, S., Chung, H.-J., and Levens, D. (2004). The dynamic response of upstream DNA to transcription-generated torsional stress. Nat. Struct. Mol. Biol. *11*, 1092-1100.

Kouzine, F., Sanford, S., Elisha-Feil, Z., and Levens, D. (2008). The functional response of upstream DNA to dynamic supercoiling in vivo. Nat. Struct. Mol. Biol. *15*, 146-154.

Kramer, P.R., and Sinden, R.R. (1997). Measurement of unrestrained negative supercoiling and topological domain size in living human cells. Biochemistry *36*, 3151-3158.

Kretzschmar, M., Meisterernst, M., and Roeder, R.G. (1993). Identification of human DNA topoisomerase I as a cofactor for activator-dependent transcription by RNA polymerase II. Proc. Natl. Acad. Sci. *90*, 11508-11512.

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

Kulaeva, O.I., Hsieh, F.K., and Studitsky, V.M. (2010). RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. Proc. Natl. Acad. Sci. *107*, 11325-11330.

Kumar, S.V., and Wigge, P.A. (2010). H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell *140*, 136-147.

Kuryan, B.G., Kim, J., Tran, N.N.H., Lombardo, S.R., Venkatesh, S., Workman, J.L., and Carey, M. (2012). Histone density is maintained during transcription mediated by the chromatin remodeler RSC and histone chaperone NAP1 in vitro. Proc. Natl. Acad. Sci. *109*, 1931-1936.

Lakhotia, S.C. (2011). Forty years of the 93D puff of Drosophila melanogaster. J. Biosci. *36*, 399-423.

Lakhotia, S.C. (2012). Long non-coding RNAs coordinate cellular responses to stress. Wiley interdisciplinary reviews. RNA *3*, 779-796.

Lavelle, C. (2007). Transcription elongation through a chromatin template. Biochimie *89*, 516-527.

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, J.T., and Bartolomei, M.S. (2013). X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell *152*, 1308-1323.

Lee, M.S., and Garrard, W.T. (1991). Positive DNA supercoiling generates a chromatin conformation characteristic of highly active genes. Proc. Natl. Acad. Sci. *88*, 9675-9679.

Lee, W., Tillo, D., Bray, N., Morse, R.H., Davis, R.W., Hughes, T.R., and Nislow, C. (2007). A high-resolution atlas of nucleosome occupancy in yeast. Nat. Genet. *39*, 1235-1244.

Levine, M. (2011). Paused RNA polymerase II as a developmental checkpoint. Cell *145*, 502-511.

Levy, A., and Noll, M. (1981). Chromatin fine structure of active and repressed genes. Nature *289*, 198-203.

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

Li, L.M., and Arnosti, D.N. (2010). Fine mapping of chromatin structure in Drosophila melanogaster embryos using micrococcal nuclease. Fly *4*.

Lindquist, S. (1981). Regulation of protein synthesis during heat shock. Nature *293*, 311-314.

Lindquist, S. (1986). The heat-shock response. Annu. Rev. Biochem. *55*, 1151-1191.

Lis, J.T. (1998). Promoter-associated Pausing in Promoter Architecture and Postinitiation Transcriptional Regulation. Cold Spring Harb Symp Quant biol *63*, 347-365.

Liu, L.F., and Wang, J.C. (1987). Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. *84*, 7024.

Ljungman, M. (1996). Effect of differential gene expression on the chromatin structure of the DHFR gene domain in vivo. Biochim. Biophys. Acta *1307*, 171-177.

Lopez-Barragan, M.J., Quinones, M., Cui, K., Lemieux, J., Zhao, K., and Su, X.Z. (2010). Effect of PCR extension temperature on high-throughput sequencing. Mol. Biochem. Parasitol. *176*, 64-67.

Lopez-Garcia, P., and Forterre, P. (2000). DNA topology and the thermal stress response, a tale from mesophiles and hyperthermophiles. Bioessays *22*, 738-746.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203-210.

Lu, Q., Wallrath, L.L., Granok, H., and Elgin, S.C. (1993). (CT)n (GA)n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the Drosophila hsp26 gene. Mol. Cell. Biol. *13*, 2802-2814.

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

Luse, D.S., and Studitsky, V.M. (2011). The mechanism of nucleosome traversal by RNA polymerase II: roles for template uncoiling and transcript elongation factors. RNA Biology *8*, 581-585.

Mariner, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F., and Goodrich, J.A. (2008). Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. Mol. Cell *29*, 499-509.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA Polymerase II Elongation Potential by a Novel Carboxyl-terminal Domain Kinase. J. Biol. Chem. *271*, 27176-27183.

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.

Matsumoto, K., and Hirose, S. (2004). Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of Drosophila. J. Cell Sci. *117*, 3797-3805.

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.* (2008). Nucleosome organization in the Drosophila genome. Nature *453*, 358-362.

McKenzie, S.L., Henikoff, S., and Meselson, M. (1975). Localization of RNA from heat-induced polysomes at puff sites in Drosophila melanogaster. Proc. Natl. Acad. Sci. *72*, 1117-1121.

Merino, A., Madden, K.R., Lane, W.S., Champoux, J.J., and Reinberg, D. (1993). DNA topoisomerase I is involved in both repression and activation of transcription. Nature *365*, 227-232.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev. *25*, 742-754.

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. *107*, 11301-11306.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat. Genet. *37*, 1090-1097.

Muller, M.T., Pfund, W.P., Mehta, V.B., and Trask, D.K. (1985). Eukaryotic type I topoisomerase is enriched in the nucleolus and catalytically active on ribosomal DNA. EMBO J. *4*, 1237-1243.

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.

Musgrave, D., Forterre, P., and Slesarev, A. (2000). Negative constrained DNA supercoiling in archaeal nucleosomes. Mol. Microbiol. *35*, 341-349.

Musgrave, D.R., Sandman, K.M., and Reeve, J.N. (1991). DNA binding by the archaeal histone HMf results in positive supercoiling. Proc. Natl. Acad. Sci. *88*, 10397-10401.

Nakahata, Y., Sahar, S., Astarita, G., Kaluzova, M., and Sassone-Corsi, P. (2009). Circadian control of the NAD+ salvage pathway by CLOCK-SIRT1. Science *324*, 654-657.

Naughton, C., Avlonitis, N., Corless, S., Prendergast, J.G., Mati, I.K., Eijk, P.P., Cockroft, S.L., Bradley, M., Ylstra, B., and Gilbert, N. (2013). Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat. Struct. Mol. Biol. *20*, 387-395.

Nogales-Cadenas, R., Carmona-Saez, P., Vazquez, M., Vicente, C., Yang, X., Tirado, F., Carazo, J.M., and Pascual-Montano, A. (2009). GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. Nucleic Acids Res. *37*, W317-322.

O'Brien, T., and Lis, J.T. (1991). RNA polymerase II pauses at the 5' end of the transcriptionally induced Drosophila hsp70 gene. Mol. Cell. Biol. *11*, 5285-5290.

Oravec, M., Kumar, A., and Wu, R.S. (1972). Inhibition of labeling of messenger and nucleoplasmic RNA of HeLa cells by camptothecin. Biochim. Biophys. Acta *272*, 607-611.

Pang, B., Qiao, X., Janssen, L., Velds, A., Groothuis, T., Kerkhoven, R., Nieuwland, M., Ovaa, H., Rottenberg, S., van Tellingen, O*., et al.* (2013). Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. Nat. Commun. *4*, 1908.

Papantonis, A., Larkin, J.D., Wada, Y., Ohta, Y., Ihara, S., Kodama, T., and Cook, P.R. (2010). Active RNA polymerases: mobile or immobile molecular machines? PLoS Biol. *8*, e1000419.

Pereira, S.L., Grayling, R.A., Lurz, R., and Reeve, J.N. (1997). Archaeal nucleosomes. Proc. Natl. Acad. Sci. *94*, 12633-12637.

Peter, B.J., Arsuaga, J., Breier, A.M., Khodursky, A.B., Brown, P.O., and Cozzarelli, N.R. (2004). Genomic transcriptional response to loss of chromosomal supercoiling in Escherichia coli. Genome Biol. *5*, R87.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol. Cell *23*, 297-305.

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). Overcoming the nucleosome barrier during transcript elongation. Trends Genet.

Potter, D.A., Fostel, J.M., Berninger, M., Pardue, M.L., and Cech, T.R. (1980). DNA-protein interactions in the Drosophila melanogaster mitochondrial genome as deduced from trimethylpsoralen crosslinking patterns. Proc. Natl. Acad. Sci. *77*, 4118-4122.

Quail, M.A., Kozarewa, I., Smith, F., Scally, A., Stephens, P.J., Durbin, R., Swerdlow, H., and Turner, D.J. (2008). A large genome center's improvements to the Illumina sequencing system. Nat. Methods *5*, 1005-1010.

Rando, O.J. (2012). Combinatorial complexity in chromatin structure and function: revisiting the histone code. Curr. Opin. Genet. Dev. *22*, 148-155.

Reeve, J.N., Sandman, K., and Daniels, C.J. (1997). Archaeal histones, nucleosomes, and transcription initiation. Cell *89*, 999-1002.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408-1419.

Rodriguez, A.C., and Stock, D. (2002). Crystal structure of reverse gyrase: insights into the positive supercoiling of DNA. EMBO J. *21*, 418-426.

Ronimus, R.S., and Musgrave, D.R. (1996a). A gene, han1A, encoding an archaeal histone-like protein from the Thermococcus species AN1: homology with eukaryal histone consensus sequences and the implications for delineation of the histone fold. Biochim. Biophys. Acta *1307*, 1-7.

Ronimus, R.S., and Musgrave, D.R. (1996b). Purification and characterization of a histone-like protein from the Archaeal isolate AN1, a member of the Thermococcales. Mol. Microbiol. *20*, 77-86.

Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp7- gene of D. melanogaster is transcriptionally egnaged. Cell *54*, 795-804.

Rowe, T.C., Wang, J.C., and Liu, L.F. (1986). In vivo localization of DNA topoisomerase II cleavage sites on Drosophila heat shock chromatin. Mol. Cell. Biol. *6*, 985-992.

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

Sanders, M.M. (1978). Fractionation of nucleosomes by salt elution from micrococcal nuclease-digested nuclei. J. Cell Biol. *79*, 97-109.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nat. Rev. Mol. Cell Biol. *7*, 557-567.

Schwartz, B.E., and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev *19*, 804-814.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849-1851.

Shandilya, J., and Roberts, S.G.E. (2012). The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. BBA - Gene Regulatory Mechanisms *1819*, 391-400.

Shaw, P.A., Sahasrabuddhe, C.G., Hodo, H.G., and Saunders, G.F. (1978). Transcription of nucleosomes from human chromatin. Nucleic Acids Res. *5*, 2999-3012.

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

Shivaswamy, S., and Iyer, V.R. (2008). Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response. Mol. Cell. Biol. *28*, 2221-2234.

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.

Sikes, M.L., Bradshaw, J.M., Ivory, W.T., Lunsford, J.L., McMillan, R.E., and Morrison, C.R. (2009). A streamlined method for rapid and sensitive chromatin immunoprecipitation. J. Immunol. Methods *344*, 58-63.

Simon, J., Sutton, C., Lobell, R., and Glaser, R. (1985). Determinants of heat shock-induced chromosome puffing. Cell *40*, 805-817.

Simpson, R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry *17*, 5524-5531.

Smolle, M., and Workman, J.L. (2013). Transcription-associated histone modifications and cryptic transcription. Biochim. Biophys. Acta *1829*, 84-97.

Sperling, A.S., Jeong, K.S., Kitada, T., and Grunstein, M. (2011). Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II in budding yeast. Proc. Natl. Acad. Sci. *108*, 12693-12698.

Straus, D.B., Walter, W.A., and Gross, C.A. (1987). The heat shock response of E. coli is regulated by changes in the concentration of sigma 32. Nature *329*, 348-351.

Studitsky, V.M., Clark, D.J., and Felsenfeld, G. (1995). Overcoming a nucleosomal barrier to transcription. Cell *83*, 19-27.

Taylor, W.E., Straus, D.B., Grossman, A.D., Burton, Z.F., Gross, C.A., and Burgess, R.R. (1984). Transcription from a heat-inducible promoter causes heat shock regulation of the sigma subunit of E. coli RNA polymerase. Cell *38*, 371-381.

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

Thiriet, C., and Hayes, J.J. (2005). Replication-independent core histone dynamics at transcriptionally active loci in vivo. Genes Dev. *19*, 677-682.

Tissieres, A., Mitchell, H.K., and Tracy, U.M. (1974). Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J. Mol. Biol. *84*, 389-398.

Tsankov, A.M., Thompson, D.A., Socha, A., Regev, A., and Rando, O.J. (2010). The role of nucleosome positioning in the evolution of gene regulation. PLoS Biol. *8*, e1000414.

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.

van Loenhout, M.T.J., de Grunt, M.V., and Dekker, C. (2012). Dynamics of DNA supercoils. Science *338*, 94-97.

van Steensel, B., Delrow, J., and Henikoff, S. (2001). Chromatin profiling using targeted DNA adenine methyltransferase. Nat. Genet. *27*, 304–308.

Villeponteau, B., Lundell, M., and Martinson, H. (1984). Torsional stress promotes the DNAase I sensitivity of active genes. Cell.

Villeponteau, B., and Martinson, H. (1987). Gamma rays and bleomycin nick DNA and reverse the DNase I sensitivity of beta-globin gene chromatin in vivo. Mol. Cell. Biol.

Wang, J.C. (2002). Cellular roles of DNA topoisomerases: a molecular perspective. Nat. Rev. Mol. Cell Biol. *3*, 430-440.

Weber, C.M., Henikoff, J.G., and Henikoff, S. (2010). H2A.Z nucleosomes enriched over active genes are homotypic. Nat. Struct. Mol. Biol. *17*, 1500-1507.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848-856.

Wilkinson, S.P., Ouhammouch, M., and Geiduschek, E.P. (2010). Transcriptional activation in the context of repression mediated by archaeal histones. Proc. Natl. Acad. Sci. *107*, 6777-6781.

Wu, C. (1980). The 5' ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. Nature *286*, 854-860.

Wu, C., Wong, Y., and Elgin, S. (1979). The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. Cell.

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.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nature Reviews Genetics *12*, 542-553.

Wysocka, J., Reilly, P.T., and Herr, W. (2001). Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. Mol. Cell. Biol. *21*, 3820-3829.

Xi, Y., Yao, J., Chen, R., Li, W., and He, X. (2011). Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Res. *21*, 718-724.

Yakovchuk, P., Goodrich, J.A., and Kugel, J.F. (2009). B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. Proc. Natl. Acad. Sci. *106*, 5569-5574.

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 *23*, 227-237.

Yang, F., Kemp, C.J., and Henikoff, S. (2013). Doxorubicin Enhances Nucleosome Turnover around Promoters. Curr. Biol. *23*, 782-787.

Zanton, S.J., and Pugh, B.F. (2006). Full and partial genome-wide assembly and disassembly of the yeast transcription machinery in response to heat shock. Genes Dev. *20*, 2250-2265.

Zeng, W., Ball, A.R., Jr., and Yokomori, K. (2010). HP1: heterochromatin binding proteins working the genome. Epigenetics *5*, 287-292.

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

Zlatanova, J., and Victor, J.-M. (2009). How are nucleosomes disrupted during transcription elongation? HFSP Journal *3*, 373-378.